مبادئ
الإحصاء

al-Shāfiʿī, ʿAbd al-Munʿim

الجزء الأول

Mabādiʾ al-Iḥṣāʾ

تأليف

عبد المنعم ناصر الشافعي

B. Sc. (Hons.) ; B. Com. ; Ph. D. ; F.S.S.

بكالوريوس الشرف فى العلوم الرياضية . بكالوريوس فى التجارة .
دكتوراه فى الفلسفة . زميل فى الجمعية الملكية للاحصاء بلندن
وعضو الجمعية الأمريكية للاحصاء ؛ أستاذ الإحصاء التطبيقى
بكلية التجارة بجامعة فؤاد الأول .

الطبعة الثانية

ملتزمة النشر والطبع
مكتبة النهضة المصرية
٩ شارع عدلى باشا - القاهرة

القاهرة
مطبعة مصر شركة مساهمة مصرية
٤٠ شارع نوبار باشا (سابقا شارع الدواوين)
١٩٤٨

بسم الله الرحمن الرحيم

المقصود من هذا الكتاب أن يكون مرجعاً سهلا لطالب الإحصاء ، يحتوى على المبادىء الأولية والقواعد البسيطة التى ينبنى عليها هذا العلم . ولم أقصد به التعمق فى شرح الأسس النظرية والرياضية التى يتناولها البحث الإحصائى ، تاركا ذلك إلى فرصة أخرى . وقد حاولت ما استطعت أن أتحاشى التعرض لهذه المسائل ؛ وقد تجاهلتها عمداً فى بعض النقط ، للتخفيف عن القارىء المبتدىء .

وقد اجتهدت أيضاً أن أقتصد فى استعمال الرموز الجبرية والطرق الرياضية ، فلم ألجأ إليها إلا حيثما وجدتها أوضح بياناً وأقصر تعبيراً من غيرها ؛ وهذه البراهين الرياضية جعلتها فى حواشى الصفحات ، بعيدة عن صلب الكلام ، حتى لا تقطع على القارىء تفكيره إذا اختار أن يهملها . ولذلك فإنى أعتذر للقراء غير الميّالين إلى الرياضة عن إدخال هذا القدر اللازم منها ، وإلى غيرهم إذا لم يجدوا كفايتهم منها . على أن ما يحتاج إليه القارىء لفهم هذه البراهين لا يتعدى القواعد الجبرية البسيطة .

وسيرى القارىء أنى استعنت بأفكار كثيرة اقتبستها من قراءتى للكتب الانجليزية والأمريكية ؛ خصوصاً فى موضوع الأرقام القياسية الذى شرحه الأستاذ ارفنج فيشر بإسهاب فى كتابه المشهور عن الأرقام القياسية ، وفى موضوع الارتباط عن الأستاذ ف . ميلز ، والمستر ج . يول ، وغيرهم من العلماء الإحصائيين مثل كارل بيرسون وآرثر بولى ، فى هذه الموضوعات وغيرها .

وإذا كان هذا الكتاب أول اجتهاد فى هذا العلم باللغة العربية — وأغلب ظنى أن هذا صحيح — فإنى أرجو أن أكون قد وفقت بهذا المجهود المتواضع إلى قضاء حاجة شعرت بها حيناً ، وكثيراً ما ألح على زملائى وتلامذتى للسعى فى قضائها ، فكانت تحول واجباتى الأخرى دون ذلك .

وإنى مدين بالشكر الوافر إلى حضرة محمد سمير إبراهيم افندى ، صديقى وتلميذى من قبل ، حيث قام بعمل الرسوم والأشكال الواردة فى هذا الكتاب ، وبذل فى ذلك عناية كبيرة حتى أخرجها على غاية من الإتقان والجمال . وكذلك أشكر حضرة عبد الله إبراهيم درويش افندى حيث قام بمراجعة حلول الأمثلة وحساباتها ، وقراءة الأصول وإبداء بعض الانتقادات المفيدة .

ولا يفوتنى أن أذكر مع وافر الشكر والثناء ما لقيته من رجال مطبعة مصر فى أثناء طبع هذا الكتاب من العناية والاهتمام ، حيث لم يدخروا وسعاً فى إخراجه على أحسن وجه ، رغم ما به من صعوبات فنية كبيرة .

المؤلف

عبد المنعم ناصر الشافعى

كلية التجارة

أول يناير سنة ١٩٣٩

مقدمة الطبعة الثانية

فى هذه الطبعة قد حاولت الاحتفاظ بالخطة الذى سبق أن رسمتها لهذا الكتاب ، وهى أن يحتوى على المبادىء الأولية لعلم الإحصاء . وقد انتهزت هذه الفرصة لتصحيح بعض الأخطاء التى وقعت فى الطبعة الأولى ، وكذلك إضافة بابين جديدين : أحدهما عن العينات الصغيرة تضمن بعض التطبيقات البسيطة للنتائج التى وصل إليها استُيُودِنْتْ ؛ والآخر عن تحليل التباين وهى طريقة إحصائية مفيدة جداً وضعها الأستاذ ر . ا . فيشر وطبقها هو وغيره فى انجلترا وأمريكا وغيرها فى العشرين سنة الأخيرة فى كثير من البحوث التجريبية ، خصوصاً التجارب الزراعية . وقد أعيد ترتيب أبواب الكتاب بعض الشىء .

وإنى أشكر جميع إخوانى وزملائى إذ تكرموا بإبداء بعض الملاحظات والانتقادات المفيدة .

جامعة فؤاد الأول — كلية التجارة
١٥ أبريل سنة ١٩٤٨

المؤلف

الفهرس

مبادئ الاحصاء

الباب الأول

مقدمة

تعريف علم الاحصاء

١ — الإحصاء علم يبحث فى طريقة جمع الحقائق الخاصة بالظواهر العلمية والاجتماعية ، وكيفية تسجيلها فى صورة قياسية رقمية ، وتلخيصها بطريقة يسهل بها معرفة اتجاهات هذه الظواهر وعلاقات بعضها ببعض ؛ ويبحث أيضاً فى دراسة هذه العلاقات والاتجاهات ، واستخدامها فى تفهم حقيقة الظواهر ومعرفة القوانين التى تسير تبعاً لها .

منشأ فكرة الاحصاء

٢ — والإحصاء ، بمعنى الحصر والعد ، فكرة قديمة يرجع منشأها إلى عهد بعيد فى تاريخ المدنية الإنسانية . ويظهر أن أول من قام بتطبيق هذه الفكرة واستخدامها فى تدبير سياسة الدول هم قدماء المصريين ، حيث قام بناة الأهرام بعمل تعداد لسكان مصر وثروتها ، واستخدموا النتائج فى تنظيم مشروع البناء . وكذلك عمل رمسيس الثانى تعداداً آخر للسكان تمهيداً لعملية إقطاع الأراضى وتوزيعها على السكان بطريقة عادلة . وفى العصور الوسطى نجد أن الملوك ورؤساء القبائل ، ومن بينهم الخليفة المأمون ، قاموا بمثل هذه العملية بين حين وآخر ليتعرفوا عدد ما لديهم من الرجال ومقدرتهم على الدفاع عن أوطانهم أو مهاجمة الغير .

الاحصاء كعلم يبحث فى جمع الحقائق الخاصة لشئون الدولة

٣ — وعندما تدرج الإنسان فى مدنيته، وتعددت مرافق الحياة، أصبحت مشاكلها أكثر تعقيداً، واستخدمت فكرة الإحصاء بالتدريج فى نواح كثيرة؛ وكانت فى كل حالة مساعداً كبيراً فى الاهتداء إلى حقيقة الأمور واتجاهات جميع الظواهر الاقتصادية والاجتماعية والعلمية البحتة. وكان استخدام الإحصاء فى المبدأ مقصوراً على الأعمال الخاصة بشئون الدولة. كما يدل على ذلك الأصل اللغوى فى اسم هذا العلم وهو بالإنجليزية (Statistics) و بناؤه (Stat-ist-ics)، وهو مشتق من كلمة (State) أى «الدولة»، ومعناه «مجموعة الحقائق الخاصة بشئون الدولة».

استخدام الاحصاء فى الأبحاث العلمية والشئون الخاصة

٤ — ولم يلبث أن انتشر استخدام هذا العلم فى نواح مختلفة، وتبينت فائدته كطريقة سليمة من طرق البحث العلمى الدقيق. ولم يقتصر تطبيقه على النواحى التى تهتم بها الحكومات فى تدبير سياستها وتصريف شئونها العامة، بل تعداها إلى جميع الظواهر الاقتصادية والاجتماعية والعلمية البحتة، وكذلك شؤون الأفراد والهيئات الخاصة التى لا تمت للحكومة بصلة ما.

الاحصاء الآن علم مستقل له نظرياته وقواعده

٥ — وكان مما ساعد على سعة تطبيق هذا العلم ونشر تعاليمه أن توفر على دراسته عدد كبير من العلماء النابغين، فبحثوا نظرياته و بنوها على أسس علمية صحيحة، وهذبوا طرقه العملية على ضوء هذه النظريات، والخبرة العملية التى اكتسبوها من أبحاثهم. وكان من بين هؤلاء بعض فحول الرياضيين، مثل عائلة برنوللى، وفردريك جاوس الألمانى، ولابلاس الفرنسى، وكيتليه البلجيكى، وجولتون الانجليزى؛ وفى وقتنا الحاضر لاينسى فضل كارل پيرسون، وآرثر بولى، وأدنى يول فى انجلترا، و إرڤنج فيشر فى أمريكا [١]، وغيرهم فى البلاد الأخرى. وهؤلاء الحديثون قد تفرغوا لدراسة علم الاحصاء، واستنباط نظرياته، وكشف غوامضه، وتطبيق هذه النظريات فى العلوم الاقتصادية والاجتماعية والعلوم

(١) Bernoulli; F. Gauss; Laplace; Quetlet; F. Galton; Karl Pearson; A. Bowley; U. Yule; Irving Fisher.

الطبيعية والحيوية . و بفضل هـذه المجهودات قد تكونت لدينا الآن ثروة عظيمة من النظريات العلميـة والطرق العملية ، يتكون منها علم مستقل جليـل الشأن ، يشتغل به عدد كبير من كبار العلماء ، وتنمو هـذه الثروة العلمية كل يوم بأبحاثهم المتواصلة ، وأبحاث تلاميذهم .

كان الاحصاء فى مبدأ نشأته لا يستخدم الطريقة الرقمية

٦ — كان علم الإحصاء فى بداية نشأته الحديثة يعنى فقط بجمع البيانات التى تهم الحكومة ؛ وكان القائمون بهذا العمل يهتمون بحفظ هذه البيانات وتسجيلها فى دفاتر الحكومة بحيث يمكن الرجوع إليها واستخدامها للاهتداء بها فى تصريف أمور الدولة ورسم سياستها . ولم يراع فى مبدأ الأمر أن يكون تسجيل هذه الحقائق بطريقة رقمية كما هو المشاهد فى الإحصاء الذى عهدنا به الآن ، بل كان هذا التسجيل يقتصر على وصف تلك الحقائق بالكلمات العادية بدون الالتجاء إلى استخدام الأرقام لتحديد هـذه الأوصاف تحديداً دقيقاً . ور بما كانت أول خطوة فى هـذا الاتجاه هى التى اتخذها آنكرسون Ankerson [1] المؤرخ الدانيمركى ، حيث رسم جدولا يبين حالات بعض الممالك الأوربيـة فى كتاب نشره فى سنة ١٧٤١ عن خمس عشرة دولة . ولم يستعمل آنكرسون الأرقام فى هـذا الجدول بالذات ، بل كان يضع أوصافاً لفظية أمام كل مملكة فى الجدول ؛ ولكن على كل حال كانت الخطوة التالية لهذه سهلة ، وهى استخدام بعض الأرقام للدلالة على صفات معينة .

استخدام الطريقة الرقمية وظهور علم الحساب السياسى فى القرن ١٧ عند الانجليز

٧ — بعد ذلك ظهرت بالتدريج أفضلية استخدام الطريقة الرقميـة للدلالة على الظواهر التى كان يطاب مراقبتها وتسجيل أحوالها ، لما فيها من تمام الوضوح ودقة التعبير — و بذلك عم استخدام الأرقام . وكانت الظواهر المختلفة تقاس كمًّا و يعبر عن مقاييسها بأعداد حسابية . ومن ثم أخذ الإحصاء شكلاً جديداً عبر عنه الانجليز فى القرن السابع عشر بالحساب السياسى أى (Political Arithmetic)

(١) أنظر (Westergaard, "*Contribution to the History of Statistics*" 1932, p. 12)

وكان هذا يتناول عدد المواليد وعدد الوفيات ، وعدد السكان ، ومقدار ثروتهم ودخلهم ، ومقدار الضرائب المتحصلة ، ومقدار الناتج من المحاصيل المختلفة ، وهكذا . وكان المشتغلون بهذا الفن فى ذلك الوقت يأملون الوصول بواسطته إلى معرفة مقدرة الممالك المختلفة على الإنتاج أو الحرب مثلاً ، إذا هم قارنوها بمملكة معينة عُرِف عدد سكانها وكمية محصولها ودرجة خصبها .

استخدام النظريات الرياضية فى استنباط القوانين الاحصائية وتدعيم نتائجها

٨ — ولما أخذ علم الاحصاء هذا الشكل الجديد ، كان من السهل الاستعانة بالنظريات الرياضية فى تفهم مسائله وشرح ما غمض منها ، وتحليلها للوصول إلى حقيقتها . وقد ساهم فى ذلك العلماء دانيال برنوللى وفردريك جاوس ولابلاس بقسط كبير (من سنة ١٧٠٠ إلى ١٨٢٠) خصوصاً فى تطبيق نظرية الاحتمالات على المسائل الاحصائية . واستنباط القوانين الاحصائية المبنية على هذه النظرية . والحقيقة أن استخدام هذه النظريات الرياضية والنتائج المبنية عليها أكسب الاحصاء صبغة علمية ، وجعل منه علماً مستقلاً محترماً ، وجذب إليه اهتمام العلماء ؛ فأنشأوا جمعيات علمية للاحصاء فى البلاد المختلفة ، وعقدوا مؤتمرات دولية لمناقشة مسائل هذا العلم ، وأصدروا المجلات العلمية مملوءة بالأبحاث الجديدة فيه . وفى أوائل القرن الماضى أنشأت الحكومات أقساماً للاحصاء بوزارات التجارة والصحة ؛ وكذلك خصصت الجامعات أقساماً بها لدراسة هذا العلم وعمل الأبحاث فيه .

علم الاحصاء

٩ — وكان طبيعياً أن يبحث هؤلاء العلماء وتلاميذهم عن تطبيقات النظريات ، فطرقوا ميادين متعددة ، بعيدة عن المجال الأصلى الذى نشأ به العلم ، ألا وهو شئون الدولة . وحصلوا فى كل حالة على نتائج علمية خطيرة يؤيدها الواقع الملموس . وهذا مما ساعد على سعة انتشاره ، وبدد ما كان

عند بعض النـاس من الشكوك فى نتائجه ونظرياته ؛ فأقبلوا على طلبه واستيعاب قواعده ، واستخدموه فى بحث المسائل العلمية المختلفة .

استخدام الاحصاء فى العلوم الطبيعية البحتة

١٠ — ومن هذه النواحى الجديدة التى استخدم فيها علم الاحصاء نذكر علوم الفلك والوراثة والبيولوجيا وعلم النفس . فى علم الفلك يستخدم علم الإحصاء فى تحليل مشاهدات أرصاد الكواكب والنجوم ، وكذلك المشاهدات الخاصة بأحوال الجو وتقلباته فى علم المتريولوجيا ، واستخدام النتائج الإحصائية فى تعليل الظواهر الجوية واستنباط قوانينها ، وتطبيقها فى التنبؤ بأحوال الجو .

وفى علم الأحيـاء (البيولوجيا) تستخدم الطرق الإحصائية فى دراسة الأجناس والفصائل المختلفة من الحيوان والنبات ، ومعرفة خواص كل جنس ، التى تميزه عن غيره ، ومقدار اختلاف مفردات الجنس الواحد فى أية خاصة معينة . فمثلا نرى أن الذكور فى الجنس البشرى أطول قامة من الإناث فى المتوسط ، مع أن الذكور فيما بينهم يختلفون فى الطول إلى درجة ما ، وكذلك الإناث . ونرى أيضاً أن بعض أصناف القطن أطول تيلة من الأصناف الأخرى ، وأن هذا لا يمنع من أن لوزات الصنف الواحد قد تحتوى على شعيرات تختلف فى طولها إلى درجة معينة — وكل هذه خواص تميز الأصناف بعضها من بعض .

وفى علم الوراثة كان علم الإحصاء من أهم العوامل فى تقدمه و بنائه على أسس علمية متينة ، حيث يدرس العلماء العلاقات بين خواص الأب والابن فى الحيوان والنبات بالطريقة الإحصائية ، ويمكن بذلك تمييز الصفات والخواص المتوارثة من المكتسبة ، وتحديد الظروف التى تحيط بعوامل الوراثة تحديداً دقيقاً لكل صفة أو خاصة . وباستخدام الطريقة الإحصائية يمكن دراسة أثر العوامل الوراثية بعضها فى بعض ، ومعرفة أيها أشد أثراً من الآخر فى التوريث . فمثلا

نرى أن ضعف القوى العقليـة فى الإنسان صفة تنتقل بالوراثة بدرجة أشد من بعض أنواع الصمم ، كما يتبين لنا من إحصاء عدد الحالات التى تنتقل فيها كل من العاهتين من الأب إلى الإبن بطريق الوراثة .

وقد استخدم علماء النفس الطرق الإحصائية فى قياس ذكاء الأشخاص والتميز بين المجموعات المختلفة من الأجناس البشرية ، وفى دراسة العلاقة بين ذكاء الشخص ومهارته فى بعض النواحى دون الأخرى ؛ وأمكنهم بواسطتها أيضاً دراسة تأثير البيئات المختلفة فى الأشخاص وسرعة نضوجهم العقـلى . وكان من البارزين فى هذا الميدان الأستاذ سپيرمان (Spearman) الانجليزى .

الاحصـــاء فى العـــلوم الاقتصـــادية يرجـــع إلى زمن بعيـــد

١١ — وفى ميدان العلوم الاقتصادية يعتبر علم الإحصاء بمثابة أحد العناصر الأساسية التى يكوّن منها العلماء نظرياتهم ، والمحكّ الذى يختبرون به منه النظريات ليتبينوا صلاحيتها لتفسير الظواهر الاقتصـادية والاجتماعية المشـاهدة فى الواقع . فبواسطة عمل إحصـاءات التجارة الخارجية مثلاً يمكن دراسـة تأثير الضرائب الجمركية على الإنتـاج الداخلى ، وعلى مستوى الأسـعار . و بعمل إحصـاءات عن كمية النقد المتداول و كمية الائتمان ، يمكن درس حالة الأسعار وما يتبعها من رواج فى التجارة ونشاط فى الأعمال .

الاحصـاء فى دوائر الأعمال الخاصـــة

١٢ — أما فى دوائر الأعمال المـالية والصناعية والتجارية ، وخصوصاً فى الأعمال ذات النطاق الـكبير ، فنجد البيانات الإحصائية هى المرشد الأول الذى يهتدى به المشتغلون بهذه الأعمال فى رسم خططهم المالية أو الصناعية أو التجارية ؛ وهم يهتمون بها جداً ولا يبخلون فى الإنفاق عليها ، لأنهم يعتبرون الإحصاءات بمثابة « ترمومتر » يقيس لهم التغييرات التى تحدث و يسجلها بدقة . فنجد الماليين مثلاً يعنون بالإحصاءات التى تدل على حركة النقد وكمية الائتمان و كمية المصدر من

الأوراق المالية الجديدة والقروض ونحو ذلك ، حتى يكونوا على بينة بحالة السوق المالية ، فلا يؤخذوا على غرة .

وكذلك المنتج يراقب الإحصاءات عن كمية المنتج والمخزون من السلع التى يشتغل بها ، وكذلك عن المواد الخام التى ينتظر أن يحتاج إليها ؛ ومن ناحية أخرى يراقب إحصاءات البيع والتصريف ، ويتفقدها ليرى مواطن الضعف فيتلافاها ، ويوازن بين كمية الإنتاج والتوزيع .

وفى الأعمال التجارية يهتم أصحابها بإحصاءات الأسعار وحركتها هبوطاً وصعوداً ، وكذلك كمية المعروض من السلع والمطلوب منها ، وقوة الجماهير على الشراء ، والسعى فى استغلال هذه القوة الشرائية وتوجيهها بقدر الإمكان نحو شراء سلعهم .

استخدام الإحصاء فى المسائل الاجتماعية

١٣ — وفى العلوم الاجتماعية والسياسية تستخدم الاحصاءات كأداة لقياس درجة رفاهية الشعب ورقى مستوى معيشته وثقافته . فنجد المشتغلين بهذه المسائل الاجتماعية يجمعون البيانات الإحصائية لمعرفة مستوى الأجور وتقدير الثروة الأهلية ؛ وبواسطة هذه البيانات ومقارنتها بالبيانات المعروفة عن أسعار الحاجيات تقدر القوة الشرائية للسكان ، وكمية مايستهلكونه من الأشياء ؛ ويعرف مستوى معيشتهم . وكذلك تجمع البيانات الإحصائية للدلالة على الحالة الصحية للسكان والتعليم والبطالة وغير ذلك من المسائل الاجتماعية الخطيرة .

الإحصاء يساعد على تحديد النقط الأساسية التى تستلزم البحث والعلاج

١٤ — وفى جميع هذه العلوم كثيراً ما يعرض للباحث بعض المسائل المعقدة حيث يلتبس عليه الأمر ، فلا يعرف أى العوامل أو الظواهر أقرب صلة بالموضوع الذى يبحث فيه وأشدها أثراً فيه ؛ فيلجأ إلى تحليل البيانات الإحصائية والنتائج المستنبطة منها ، ويهتدى بذلك إلى تعيين النقط الأساسية ، فيوليها اهتمامه ويعنى

ببحثها وملاحظتها دون غيرها ، وبذلك يحصر بحثه فى دائرة ضيقة ، فيكون مثمراً ومؤدياً إلى أحسن النتائج .

المـــراجع

Bowley, A. L., *Elements of Statistics*, Chapter 1.

Connor, L. R., *Strtistics in Theory and Practice*, Chapter 1.

King, W. I., *Elements of Statistical Method*, Chapter 1.

Mills, F. C., *Statistical Methods*, Chapter 1.

Secrist, H., *Introduction to Statistical Methods*, Chapter 1.

Westergaard, H., *Contributions to the History af Statistics.*

الباب الثاني

الطريقة الإحصائية

١٥ — تكلمنا فى الباب السابق عن تعريف علم الإحصاء، ومنشأ فكرته وتطورها، والمكان الذى يشغله هذا العلم بين العلوم فى الوقت الحاضر، وعن استعماله فى العلوم الاقتصادية والطبيعية؛ وسنتكلم فى هذا الباب وما يليه عن الطريقة التى يستخدمها هذا العلم فى المسائل التى يعالجها.

عملية الاحصاء تبدأ بملاحظة الظواهر التى نبحثها وجمع الحقائق عن أحوالها

١٦ — تبدأ عمليـة الإحصاء بمشاهدة الظواهر التى نبحثها فى الظروف المختلفة — مكانية كانت هـذه الظروف أو زمانية — وتسجيل هذه المشاهدات بطريقة يسهل بهـا الرجوع إليها وتحليلها واستنباط القوانين التى تسير تبعاً لهـا الظواهر التى نبحثها. فإذا أردنا مثلا أن نعرف ما يحمله نبات القطن من اللوز، فإننا نأخذ بعض النبـاتات فى حقل ما ونلاحظ ما يحمله كل واحد من اللوزات. ثم ننتقل إلى حقل آخر ونأخذ غيرها ونلاحظ ما تحمله هـذه من اللوز. ولكى يكون عندنا فكرة أصح نلاحظ نباتات من أصناف مختلفة من القطن. وأحسن من ذلك أيضاً أننا نكرر هذه الملاحظات فى محاصيل أعوام مختلفة وأراضٍ مختلفة، حتى يكون حكمنا أقرب ما يكون إلى الصحة. وهكذا نشاهد الظاهرة التى نبحثها فى ظروف مختلفة، وندون ملاحظاتنا عنها.

عملية العد في الإحصاء هي العملية الابتدائية : الإحصاء ليس بمجرد علم عد

١٧ — والطريقة المتبعة في الإحصاء هي — كما يدل عليها الاسم العربي — طريقة العدّ . ففي كل أبحاثنا الإحصائية نعبر عن مشاهداتنا للظواهر تعبيراً رقمياً ، ونسجل هذه الملاحظات في صورة رقمية . ففي المثال الذي قدمناه في البند السابق نعد لوزات القطن التي يحملها كل نبات ؛ كما أننا في إحصاء السكان نعد الأشخاص المقيمين في كل جهة من جهات المملكة ؛ وفي إحصاءات العمل نعد العمال المشتغلين في كل صناعة وعدد العاطلين أيضاً ، وهكذا . وعملية العدّ هذه هي خطوة أساسية في الأعمال الإحصائية ، مما جعل بعض الناس يعرف علم الإحصاء بأنه علم العد (Science of Counting) ؛ ولكن هذا تعريف قاصر لايفي بالمعنى ، وذلك لأن عملية العد ماهي إلا العملية الابتدائية . ومع أنها خطوة أساسية في العمل فإنها ليست كل شيء ، كما سيتضح لنا فيما يلي .

البيانات الأولية

١٨ — وهذه الحقائق التي نجمعها عن الظواهر التي نريد بحثها ونسجلها في صيغة رقمية ، نسميها البيانات الأولية (Primary Data) التي نستخدمها في البحث للوصول إلى الحقيقة . وهذه البيانات تكون بأيدينا بمثابة للمواد الأولية بيد الصانع ؛ فهو يعالجها ويهذبها حسب ماتقتضيه قواعد حرفته ؛ ويخرج منها سلعة جديدة قد تختلف كل الاختلاف في شكلها وتركيبها عن المادة المشتقة منها . وكذلك الإحصائي : فهو يعالج هذه البيانات الأولية ، بالتحليل أو التركيب ، ويستخرج منها بيانات ثانوية يستعملها في أبحاثه . ففي مسألة القطن التي ذكرناها مثلا نأتي بالبيانات الأولية عن عدد اللوز وعدد الشجيرات التي وقعت تحت ملاحظتنا ، ومساحة الأراضي المزروعة . ومن هذه المعلومات الأولية يمكن استخراج متوسط عدد اللوزات التي يحملها كل نبات ، ويمكن أيضاً معرفة متوسط عدد الشجيرات المزروعة في وحدة المساحة في كل حقل ، ومتوسط غلة الفدان

من المحصول . و بواسطة هـذه البيانات الثانوية يمكن دراسة العلاقة بين كثافة الزرع وعدد اللوز على الشجيرات ، ومحصول الفدان من القطن الشعر . وقد تستعمل هذه العلاقة بعـد دراستها ومعرفة حقيقتها فى دراسة أخرى أكثر تعقيداً ، وهكذا .

الطـرق التى نتبعها فى جمع البيـانات الأوليـة وفى اسـتخراج المعـلومات الثانوية تختلف عن بعضها

١٩ — ولا يخفى أن الخطوات والطرق الفنيـة التى تستخدم فى الحصول على المعلومات الثانوية تختلف عن الطرق المتبعـة فى جمع المعلومات الأولية ، إذ هما مسألتان مختلفتـان شكلاً وموضوعاً ؛ والصعوبات التى تعترضنا فى جمع البيـانات الأولية تخالف تلك التى تصادفنا فى تلخيص هذه البيـانات أو تركيبها لاستخراج البيانات الثانوية . فعنـد جمع الحقائق الأولية مثـلا يجب أن نتفق على الوحدة التى نستعملها فى العد ، ونعين الأشياء التى يتناولها العد ، والمصادر التى نعتمد عليها فى إمدادنا بهذه المعلومات ، والنظام الذى نتبعه فى جمعها ، وهكذا . أما فى تركيب هذه البيانات أو تحليلها فالصعوبة هنا من الوجهة النظرية والحسابية .

فإذا أردنا مثلا أن ندرس مستوى الأسعار العام فى بلد ما ، نبدأ بجمع البيانات الأولية عن الأسعار ؛ ولهـذا يجب أولا أن نحدد أى السلع التى نجمع أسعارها ، ثم الوحدات التى نسـتخدمها فى بيان الأسعار ، ثم الهيئـات أو الأشخاص الذين نلجأ إليهم فى معرفة هذه الأسعار ، وكيفية إرسالها منهم إلينا لتصلنا فى مواعيد معينة ، وهكذا .

أما فى تلخيص هذه البيانات الأولية بعد ورودها ، فالمسألة حينئذ تنحصر فى كيفية تركيب هذه البيانات لتنتج لنـا الملخص المطلوب عن مستوى الأسعار ، بحيث يكون ممثلا عادلا لحالة السوق ، يأخذ فى الاعتبار السلع المختلفة ، كل منها حسب أهميته . و بعد الاتفاق على هذا ، واختيار الصيغة الملائمة ، تصادفنا بعض

الصعوبات فى الأعمال الحسابية التى ربما تجعل العمل مرهقاً يكلفنا فوق طاقتنا. وكل هذه صعوبات تختلف فى طبيعتها وطرق معالجتها عن الصعوبات الأولى. وسنقتصر فى هذا الباب على دراسة المبادىء التى نتبعها فى جمع البيانات الأولية، ونؤجل النظر فى كيفية معالجة هذه البيانات واستخدامها فى دراستنا للظواهر التى نبحثها إلى الأبواب التالية.

يجب أن نحدد المسائل التي نريد بحثها والظواهر التي تتصل بها ثم نجمع البيانات على هذا الأساس

٢٠ — من البدهى أننا عند دراسة ظاهرة معينة نحدد، بقدر الإمكان، مجال البحث حتى يكون منتجاً، ونقتصر فى دراستنا على الظواهر التى لها علاقة بالمسألة التى نبحثها بالذات؛ وهذه نشرع فى ملاحظتها ونترك غيرها جانباً. ومما يساعدنا فى ذلك تحديد المسألة التى تواجهنا بالذات، ومعرفة الغرض الذى نرمى إليه بعمل هذ االبحث. فإذا أردنا مثلاً عمل تقدير لمحصول القطن المصرى هذا العام، فالعناصر الأساسية فى هذا الموضوع هى، بلا شك، مجموع المساحات المزروعة قطناً فى جميع أنحاء القطر، وتقسيم هذه المساحات حسب الأصناف المختلفة من القطن والجهات، وحالة الزراعة فى هذا العام بالذات — من حيث التبكير والتأخير مثلا، ومن حيث الجو وتأثيره، ومن حيث الآفات وشدة وقعها والنجاح فى مكافحتها والتغلب عليها. وبدهى أن هذه هى العناصر التى لها علاقة بالمحصول وتؤثر فى مقداره؛ وكل ماعداها فهو إما لاعلاقة له بالمحصول أو ذو علاقة بعيدة أو غير مباشرة، ولا يمكن أن يكون له أى تأثير محسوس.

يصعب أحياناً تحديد العناصر المتصلة بالموضوع فنعمل دراسة تمهيدية لمعرفتها

٢١ — وربما لا تكون هذه الخطوة الابتدائية سهلة كما تبدو فى المثال الذى أخذناه، خصوصاً فى المسائل المعقدة حيث لا نعرف بسهولة أى العناصر ذا صلة قريبة أو بعيدة بالظاهرة التى ندرسها، ونحتاج إلى تفكير طويل قبل الاهتداء إلى تعيين النواحى التى نبحث فيها ونجمع عنها البيانات. فلوأردنا مثلادراسة موضوع

البطالة بين العمال بقصد الاهتـداء إلى حل يخفف وطأتها أو يمحوها ، فقـد يظن البعض أن الظاهرة الأكثر اتصالا بها هو مستوى الأجور ؛ ويقول آخر مستوى الأسعار ، أو كمية النقد ، أو كمية الإنتاج ، أو حركة التصدير ، وهكذا . وفى مثل هذه الأحوال يتعين علينا دراسة هـذه الظواهر مجتمعة أو منفردة ، وأثر كل واحدة منها ، حتى نتبين أيها تتصل بموضوعنا ، فنخصها بالعناية والدرس ونهمل ما عداها .

عملية العـدد تتطلب تعيين وحـدة

٢٢ — قلنا إن البحث الإحصائى يبدأ بعملية العد . وبدهى أن هذه العملية معناها تحليل الشىء المعـدود إلى وحدات ، ومعرفة عدد الوحدات التى يتركب منها ، والدلالة على هذا العدد بواسطة الأرقام الحسابية المعروفة ؛ كما لو أراد شخص أن يعد ما لديه من النقود فهو يعبر عن هذه الكمية بدلالة القرش أو الجنيه كوحدة .

الوحـدات الحسـابية مفروض أنها متطابقة

فلإجراء عملية العد ، إذن ، يجب أن نتفق أولا على وحدة ، ثم نبحث عن عدد ما يحتويه الشىء المعدود من هـذه الوحدات . والأصل فى الوحدات التى تضاف إلى بعضها لتكون شيئاً معيناً ، أن تكون كلها متماثلة متشابهة متساوية من جميع الوجوه ، أو كما يقول الرياضيون « متطابقة » ؛ فنحن مثـلا لا نضيف وحدات من البرتقـال إلى وحدات من المنجـة ، وإن كنا نتجاوز أحياناً ونضيف عدد صناديق البرتقـال على عدد صناديق اليوسفى المصدرة إلى الخارج على اعتبـار أنهما « صادرات موالح » . ومع أن كثيراً منا يعتقد أن « للذكر مثل حظ الأنثيين » ، فهم لا يغضبون كثيراً ولا قليلاً إذا قيل لهم إن تعداد القطر المصرى فى سنة ١٩٣٧ كان ١٥٫٩٠٤٫٥٢٥ نسـمة ، حتى ولو كان من السـكان ٧٩٤٧١٩٣ ذكوراً و٧٩٥٧٣٣٢ إناثاً . والحقيقـة أننا نكلف الأشـياء ضد طباعها إذا تطلبنا أن تكون الوحدات متساوية تماماً قبل أن نعدها سوياً . فمن ذا الذى رأى شيئين

متساويين من جميع الوجوه ؟ فيجب ، إذن ، أن نتجاوز عن الفروق البسيطة التى نشاهدها بين المفردات المتشابهة ونعتبرها وحدات متساوية للسهولة ، ونضيفها إلى بعضها أو نطرحها ، كما نجمع ونطرح الأعداد فى عملياتنا الحسابية العادية .

في علم الإحصاء ننظر فقط إلى المجموعة ولا نهتم بالمفردات ، فالفرد قد يشذ عن المجموعة

٢٣ — وهذا التجاوز ألزم مايكون فى عملية العدّ فى علم الإحصاء ، لأن البحث الإحصائى يتناول دائماً المجموعات الكبيرة المكونة من مفردات كثيرة العدد ، ويعتمد فى أحكامه ونتائجه على أنها تستنبط من دراسة عدد كبير من الحالات ومجموعات تشتمل على مفردات كثيرة . ولذلك لا يهتم الإحصائى بالمفردات ويهملها فى ذاتها ؛ ولا ينظر إليها إلا من حيث أنها تكون مجموعة معينة ، لها خواص ومميزات معينة ، ربما لا تكون ظاهرة بوضوح فى بعض المفردات ، أو تكون بعض المفردات شاذة نوعاً عن المجموعة . والسبب فى وجهة النظر هذه — وهى أساسية فى البحث الإحصائى على العموم — أن المفردات فى كل مجموعة تكون كل واحدة منها ، بحكم انفرادها ، واقعة تحت تأثير ظروف خاصة ؛ وهذه الظروف تؤثر فى هذه المفردة تأثيرات مختلفة كماً وكيفاً . فلو كنا نبحث مقدار ماينتجه شجر البرتقال من الثمر مثلاً ، لا نقتصر على معرفة ما تنتجه شجيرة معينة ، لأنها قد تكون غرست عفواً فى تربة غنية بالغذاء ، أو بالعكس تكون انتابتها آفة دون غيرها ، لسبب طارىء ، وهكذا . والحقيقة أن علم الإحصاء لايبحث فى دراسة الأفراد لذاتها ؛ وإنما يقصد منه دراسة المجموعات ومعرفة خواصها ومميزاتها — وذلك باستقراء مفرداتها وإبراز صفاتها المشتركة التى تميزها ، كمجموعة ، عن سائر المفردات والمجموعات الأخرى .

يجب تقسيم المجموعة إلى أجزاء متجانسة قبل عدها

٢٤ — والباحث الإحصائى الموفق يمكنه أن يميز بسهولة — حسب خبرته وقوة ملاحظته — بين الفوارق الظاهرية التى يشاهدها بين المفردات التى تقع تحت

ملاحظته فيتغاضى عنها ، والفوارق الحقيقية بين هذه المفردات فيأخذها فى الاعتبار ؛ ويقسم المفردات إلى مجموعات على أساس هــذه الفوارق . فمن يبحث فى أجور مجموعة كبيرة من العمال مثلا ، ربما يجد بعض هؤلاء العمال يشتغلون فى حرف فنيـة والبعض الآخر يشتغلون فى أعمال عادية ؛ ولعلمه بأن هـذا الفارق له دخل كبير فى تحـديد الأجور فهو يقسم هؤلاء العمال إلى مجموعتين ، ويبحث أجور كل منهما على حدة ؛ وذلك على اعتبار أن أفراد المجموعة الأولى يخـالفون أفراد المجموعة الثانية ، ولا يصح إضافة هؤلاء إلى هؤلاء كوحدات متساوية . وكذلك من يبحث فى طول شعيرات القطن ، وأحضر لهذا البحث عدداً كبيراً من لوزات القطن المختلفة ، يمكنه أن يهمل الاختـلافات التى بين هذه اللوزات من حيث مكان الزرع ، وثقافة زراعها ، وكميـة السماد المستعملة ، وذلك بدون أن ينشأ عن هذا الإهمال خطأ كبير فى النتيجة التى يصل إليها ؛ ولـكنه لا يمكنه أن يتجاهل الفوارق بين هـذه اللوزات من حيث الصنف — قطن سكلاريدس أو أشمونى أو غير ذلك . ويجب أولا أن يقسم هذه المجموعة الـكبيرة إلى مجموعات أصغر حسب الصنف ، ثم يبحث فى مفردات كل مجموعة على حدة حيث تـكون كل من هذه المجموعات الصغيرة متجانسة وخاليـة من المفردات الغريبة . وإذا لم يفعل ذلك فلا بد أن يحصل على نتائج غير دقيقة ومضللة .

تحـديد معنى الوحدة يكون بتعيين الصفات المهمـة التى يجب توافرها فى كل مفردة

٢٥ — والخطوة الأساسـية فى تحديد معنى الوحدة التى نستعملها فى عملية العدّ هى تعيين الصفات الرئيســـية التى إذا توافرت فى مفردة عددناها وحدة من الوحدات ، والصـفات الأخرى التى لا نهتم لها ، وجدت أو لم توجد ، فى أى واحدة من هذه الوحدات . وواضح أن جميع الوحدات التى سنحصل عليها طبقاً لهذا لن تـكون كلها متساوية من جميع الوجوه ، وإنما تشترك فقط فى الصفات

التى عددناها رئيسية ؛ وربما اختلفت بعض الوحدات عن البعض الآخر فى الصفات الأخرى التى اعتبرناها غير مهمة .

وبدهى أن تعيين الصفات الرئيسية وغيرها يتوقف بدوره على ظروف المسألة ، والغرض الذى نرمى إليه من عملية العدّ ، والصعوبات العملية التى ربما نلاقيها فى التنفيذ . فإذا أردنا مثلاً إحصاء الماشية فى قطر زراعى ، عددنا الأبقار والجاموس عموماً . أما إذا كان الغرض من هذا الإحصاء تقدير الناتج من اللبن فى العام ، فطبعاً نعد الإناث منها فقط ، وربما اقتصرنا على الحلوب من هذه وتركنا غيرها . وكذلك إذا أردنا إحصاء الآلات الميكانيكية المستعملة فى الصناعة ، وكان الغرض من هذا معرفة مقدرة المؤسسات الصناعية على الإنتاج ، عددنا كل الآلات الموجودة وقوتها بالحصان البخارى مثلاً ؛ فى حين لو كان الغرض من هذا الإحصاء هو تقدير الإنتاج الحاصل فعلاً فإننا لا نعد إلا الآلات المدارة فعلا ونترك المعطلة منها .

الوحدات ولو اشتركت فى الصفات الرئيسية يصح أن تتفاوت فى هذه الصفات

٢٦ — وكثيراً ما نجد عملياً أن هذه الوحدات التى اعتبرناها متساوية ، تقريباً ، لاشتراكها فى الصفات الرئيسية التى نعينها لهذا الغرض ، لا تساوى بعضها بعضاً فى هذه الصفات ، بل تتفاوت فيما بينها تفاوتاً قد يكون كبيراً . فعند تقدير المنتج من اللبن فى العام مثلاً قلنا إننا نعد الإناث من الأبقار والجاموس ؛ والأفضل أن نقتصر على الحلوب منها ليكون التقدير أدنى إلى الصواب . ولكننا نعلم أن كمية اللبن ونوعه يختلفان بين البقر والجاموس . ونعلم أيضاً أن الأبقار الحلوب تتفاوت فيما بينها فى كمية اللبن التى تدرّه كل واحدة فى اليوم : فمنها ما تحلب قدحاً ومنها ما تحلب أربعة ، وكذلك فى الجاموس . وعلاوة على ذلك فكمية اللبن فى كليهما تختلف باختلاف الموسم والفصول ، وبحسب الجو ونوع الغذاء ، وغير ذلك من الاختلافات .

الوحدة الاحصائية هى رمز ينوب عن عدة مفردات تشترك فى صفة معينة

٢٧ — أمام هذه الحقائق العملية نرى أنه من العبث أن نحاول تعريف وحدة العدّ التى نستعملها بأنها تساوى فى صفاتها واحدة معينة من المفردات المطلوب عدّها ؛ والأفضل أن نعتبر الوحدة الإحصائية رمزاً يدل على أى واحدة من المفردات التى تشترك فى صفة أو عدة صفات معينة ، حتى ولو كان بين هذه المفردات تفاوت فى هذه الصفة أو الصفات . وليس من الضرورى أن يكون لهذا الرمز وجود ذاتى فى الواقع .

وعلى هذا الاعتبار يمكننا أن نقول إن عدد سكان القطر المصرى فى سنة ١٩٣٧ كان ١٥٫٩٠٤٫٥٢٥ نسمة ؛ ونفهم أن الوحدة هنا رمز يدل فى نفس الوقت على الفلاح فى قريته ، والتاجر ساكن المدينة ، والمرأة فى منزلها ، واليافع فى مدرسته ، والصانع فى عمله ، وهكذا . والكل يشتركون فى صفة واحدة وهى أنهم كانوا جميعاً « على قيد الحياة فى منتصف الليلة الواقعة بين يومى ٢٦ و ٢٧ مارس سنة ١٩٣٧ فى الأراضى المصرية » .

وكذلك نقول إن مقدار الوارد إلى مصر من أجهزة الراديو (للاستقبال) فى سنتى ١٩٣٦ و ١٩٣٧ كان على الترتيب ١٥٢٦٧ و ٢٠٠٨٠ جهازاً . ونعلم أن هذه الأجهزة تختلف بعضها عن بعض فى النوع والقيمة والحجم ؛ ولا يمنعنا هذا من أن نستنتج أن استعمال هذه الأجهزة فى مصر زاد زيادة محسوسة بين السنتين المذكورتين .

لدراسة الصفات نقيسها و نعبر عنها بأرقام

٢٨ — وعندما ندرس صفة معينة نلاحظ عدداً من المفردات التى تظهر فيها هذه الصفة؛ ونجتهد ، بقدر الإمكان ، أن نقيسها لنحصل على تعبير رقمى لها نستخدمه فى مقارنة المفردات المختلفة من حيث هذه الصفة . ولإجراء عملية القياس لا بد أن نختار وحدة قياس مناسبة ، سهلة ، عملية ، دقيقة بقدر الإمكان . فإذا أردنا بحث

طول القامة عند مجموعة من الأشخاص مثلاً ، نقيس طول كل واحد منهم بالسنتيمتر أو بالقدم والبوصة إذا فضلنا ؛ وبذلك نحصل على قياسات لهذه الصفة فى صورة رقمية يمكن بواسطتها التمييز بين الأفراد من حيث أطوالهم . وكذلك إذا أردنا دراسة الوزن أو العمر فإننا نقيس وزن كل فرد بالرطل أو بالكيلوجرام ، والعمر بالسنة أو بالشهر ، وهكذا .

وواضح أن دراسة الأشياء بواسطة قياسها ، والتعبير عنها بصورة رقمية ، هى أحسن طريقة ممكنة ، وهى الطريقة الوحيدة للبحث العلمى الدقيق .

بعض الصفات يمكن قياسها والبعض يتعذر أو يستحيل قياسها مباشرة

٢٩ — عملية القياس ليست ممكنة فى جميع الأحوال : فهناك بعض الصفات أو الأشياء يمكن قياسها قياساً مباشراً بدون أدنى صعوبة ، كما نرى فى صفات الطول والوزن والعمر وأثمان الأشياء ، وأجر العامل (النقدى) وطول ساعات العمل وهكذا . وفى مثل هذه الصفات لا نجد صعوبة فى اختيار وحدة القياس التى نستعملها . وبعض الأشياء يتعذر قياسها ويصعب تحديد وحدة لقياسها ؛ وبعضها لا يمكن قياسها بالمرة . فإذا كان لدينا ثلاثة رجال مثلاً أمكننا أن نعرف تواريخ ميلادهم ، ونحسب أعمارهم بالسنين وكسورها ؛ ولكن يصعب علينا قياس صفة الصحة أو المرض بينهم قياساً مباشراً ؛ ويصح أن نعبر عن هذا بطريق غير مباشر بأن نقيس الوزن أو ضغط الدم أو النبض . ولكننا هنا لا نقيس صفة الصحة التى نقصدها بالذات وإنما نقيس صفة أو صفات أخرى غيرها نعلم أن بينها وبين الصفة المقصودة ارتباطاً وثيقاً .

أما إذا كنا نبحث فى صفة مثل الديانة أو الجنسية أو الحرفة التى يزاولها كل منهم ، فلا يمكننا أبداً قياس هذه الصفات ولا تعيين وحدات لقياسها ؛ وكل ما يمكن عمله أن نقول إن الأول ديانته : مسلم ، وجنسيته : مصرى ، وحرفته : نجار مثلاً ؛ وهكذا للثانى والثالث ، بحسب أنواع الديانة والجنسية والحرفة .

دقة المقاييس تتوقف على الأجهزة المستعملة والغرض المقصود من البحث

٣٠ — وقبل الشروع فى عملية قياس الصفة التى نبحثها يجب أن نقرر إلى أى درجة من الدقة نسير فى هذه العملية . وهذا يتوقف طبعاً على دقة الأجهزة التى نستعملها ودرجة حساسيتها ؛ ويتوقف أيضاً على الشخص الذى يقوم بعملية القياس وما يبذله من الوقت والمجهود فى سبيل الحصول على مقاييس دقيقة . ومن ناحية أخرى يتوقف أيضاً على مقدار الشىء الذى نقيسه ، والظروف الأخرى المحيطة به . ومهما كان فمن المعلوم أن الدقة التامة مستحيلة على البشر . ولكن هذا لا يمنعنا طبعاً من أن نتوخاها ونبذل فى سبيلها كل ما يمكننا من وقت ومجهود حسب حاجتنا إليها . فنحن نكتفى مثلا أن نعرف وزن جسمنا لأقرب كيلو جرام أو لأقرب نصف كيلو جرام على الأكثر . ولا نهتم كثيراً ولا قليلا إذا كان الوزن الحقيقى أكبر أو أقل مما يسجله الميزان بعشرة أو عشرين أو مائة جرام . ولا نهتم بأن نزن أنفسنا على موازين حساسة تعطينا الوزن لأقرب جرام أو نصف جرام ، لأنها تكلفنا أكثر من الخمسة مليمات التى نضعها فى الميزان العادى ، ولأن دقتها الزائدة عما نحتاج إليه لا تبرر الزيادة فى الثمن . وكذلك نقيس أطوالنا لأقرب بوصة أو سنتيمتر ؛ فى حين أننا نقيس الضغط الجوى لأقرب مليمتر على تدريج البارومتر .

وفى بعض الأحيان نحصل على مقاييس أكثر دقة مما نحتاج إليه فعلا فى أعمالنا . وهذه المقاييس نقربها بالطريقة المعتادة ، وهى أن نهمل الأجزاء التى تقل عن نصف الوحدة المستعملة ، ونجبر الكسور التى تزيد عن النصف إلى الواحد الصحيح .

قواعد جمع البيانات الاحصائية

٣١ — ذكرنا أن البحث الإحصائى يبتدىء بجمع الحقائق والبيانات الأولية عن الظواهر التى نريد دراستها ، وهذا يكون بطريق العدّ والقياس لوضع هذه الحقائق فى صورة عددية . ولتنظيم هذه العملية نتبع بعض القواعد العامة .

ولتوضيح هذه القواعد نأخذ مثالاً عملياً تظهر فيه الخطوات المهمة .

لنفرض أننا نجمع بيانات عن أجور العمال المشتغلين فى الصناعة فى القطر المصرى مثلاً ، وأن الغرض من هذا البحث هو معرفة مستوى المعيشة بين طبقة العمال .

تعـريف الوحدات التى يتناولها البحث

نبدأ بتعريف من هو العامل الذى سيتناوله البحث ، وهو كل شخص يشتغل لحساب غيره فى إحدى العمليات الصناعية مقابل أجر نقدى لا يزيد عن خمسين قرشاً فى اليوم مثلاً . وعند جمع البيانات نقتصر على من ينطبق عليه هذا التعريف ؛ فنترك المشتغلين فى التجارة أو الزراعة مثلاً ، وكذلك من يشتغلون لحسابهم الخاص ، ومن يشتغلون بالماهية الشهرية أو السنوية ، والذين يزيد أجرهم عن الحد المعين .

وكذلك نحدد معنى الأجر ووحدة القياس التى نستعملها ، لنتفق مثلاً على أن المعنى المقصود بالأجر فى هذا البحث هو جملة ما يتحصل عليه العامل فى اليوم (أو الأسبوع) مقابل قيامه بالعمل ، وأن هذا الكسب يقدر نقدياً بالقروش عن اليوم الواحد .

توحيد نظام وميعاد جمع البيانات عن الصناعات والجهات المختلفة

٣٢ — ويجب أن تكون البيانات التى نجمعها عن العمال فى الصناعات والجهات المختلفة على نظام واحد حتى يمكن مقارنتها ببعضها وإضافة الأرقام بعضها إلى بعض لكى نحصل على المجموع الكلى والمتوسط العمومى . وذلك بأن نقسم العمال حسب الصناعات التى يشتغلون فيها وحسب الجهات . ويجب أن تكون البيانات كلها مجموعة فى وقت واحد حتى يمكن مقارنتها من حيث الزمن ، إذ لا تكون المقارنة صحيحة إذا وضعنا متوسط الأجور فى القاهرة فى سنة ما بجانب متوسط الأسكندرية المحسوب من بيانات جمعت فى سنة أخرى .

طبع كشوف وتوزيعها لجمع البيانات بواسطتها

٣٣ — ولضمان هذا التوحيد عملياً نجمع البيانات بواسطة كشوف أو استمارات نطبعها على نسق واحد ، ونوزعها على الجهات المختلفة لاستيفاء البيانات المطلوبة. وهذه الطريقة تضمن : أولاً أن تكون الأسئلة المطلوب الإجابة عليها موحدة

فى جميع الحالات ، وفى الوقت نفسه تسهل عمليـة جمع البيانات . وعند تصميم هـذه الكشوف ، واختيار الأسئلة التى يطلب الإجابة عليها ، يجب أن نأخذ فى الاعتبار درجة ثقافة الأشخاص الذين سنعتمد عليهم فى إجابة الأسئلة المطلوبة وملء الكشوف . فنصوغ الأسئلة فى صورة سهلة الفهم محددة المعنى ، بحيث لا يحتمل السؤال أكثر من جواب واحد يعطى البيان المطلوب معرفته .

٣٤ — ولتوضيح ذلك نورد هنا صورة من الاستمارة المستخدمة فى إحصاء تجريبى للأجور وساعات العمل فى بعض الصـناعات بالقاهرة عن شهر مايو سنة ١٩٣٨ :

اسم المصنع .... العنـوان ————————

نوع الصناعة .... ساعات العمل فى المصنع ....

يوم العطلة الأسبوعى .... مواعيـد دفع الأجور ....

رقم مسلسل	العـامل				الأجـر				استقطاع غرامات ومواد تالفـة	الأجـر الاضافى		جملة أيام الشغل فى شهر مايو	جملة الكسب فى شهر مايو ٩٣٨	ملاحظات
	جنسيته	نوعه	سنه	وظيفته	باليوم	بالأسبوع	بالشهر	بالقطعة		الساعات	الأجر			

المـــراجع

Bowley, A. L., *Elementary Manual of Statistics*, Chapter II.

Bowley, A. L., *Elements of Statistics*, Chapters II., III.

Connor, L. R., *Statistics in Theory and Practice*, Chapters II, IV.

Secrist, H., *Introduction to Statistical Methods*, Chapters II, IV.

الباب الثالث

طرق عرض البيانات

توضيح البيانات الإحصائية وطرق عرضها

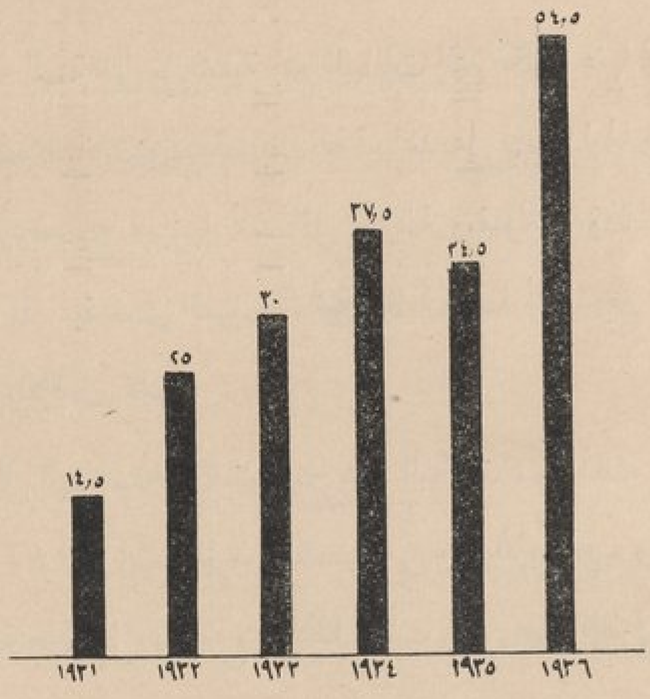
٣٥ — المهمة التى تلى عملية جمع البيانات التى تكلمنا عنها فى الباب السابق هى ترتيب هذه البيانات وتنسيقها بطريقة تساعد على فهم مدلولها والاستفادة منها، ثم عرضها وتوضيح معانيها بوسائل طريفة ومقبولة. وذلك لأن البيانات الإحصائية إذا سيقت فى الصورة الرقمية الجافة ربما لا تشجع الشخص العادى على قراءتها والإقبال عليها.

والوسائل التى نستخدمها لتوضيح هذه البيانات، وكذلك طريقة عرضها، تتوقف على نوع البيانات والغرض المقصود من هذا الإيضاح، والحقائق التى نريد إبرازها بصفة خاصة؛ وسنشرح فى هذا الباب بعض هذه الطرق.

تمثيل الأرقام هندسياً بواسطة خطوط ذات أطوال متناسبة

٣٦ — يكون لدينا أحياناً سلسلة من الأرقام تدل مثلاً على المستهلك من القطن الخام فى مصنع معين فى عدة سنين متتالية، أو على قيمة الصادرات أو الواردات لبلد معين فى سلسلة من السنين المتتالية أيضاً. مثل هذه السلسلات الزمنية يمكن توضيحها هندسياً بواسطة أعمدة عريضة، أو مستطيلات رأسية، تتناسب ارتفاعاتها مع الأرقام التى تمثلها هذه الأعمدة أو المستطيلات للسنين المختلفة. وهذه توضع بجانب بعضها بطريقة مناسبة يسهل بها عمل مقارنات بين السنين المختلفة بمجرد النظر و بسرعة.

وفى الشكل المرافق (رقم ١) نرى توضيحاً لكميات المنتج محلياً فى القطر المصرى من المنسوجات القطنية فى السنين ١٩٣١ — ١٩٣٦ مقدراً بملايين الأمتار المربعة[1]. وبالقاء نظرة سريعة على هذا الشكل يمكن للقارىء أن يأخذ فكرة واضحة عن حركة الإنتاج المحلى فى هذه الصناعة فى المدة المذكورة . ويلاحظ أن بساطة الشكل ووضوحه مما يساعد على سهولة المقارنة بين السنين وسرعة رسوخ هذه الحقائق فى الذهن بدون عناء .



(شكل ١)

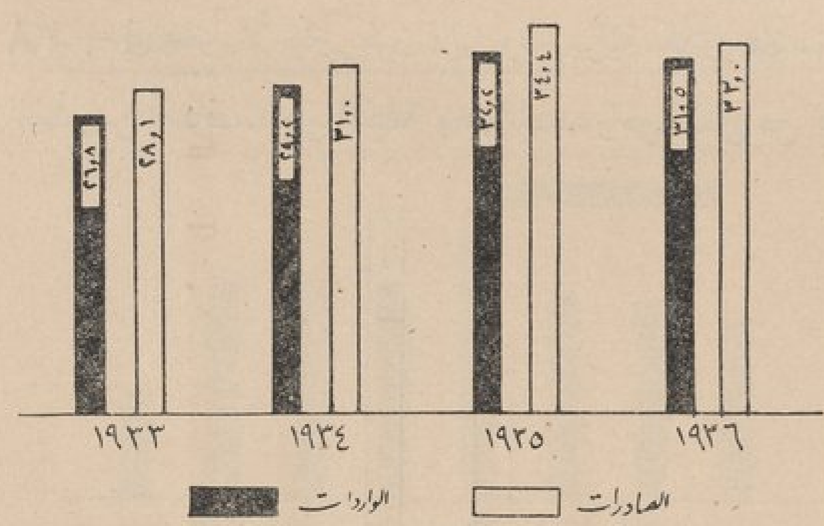
إنتاج المنسوجات فى مصر فى السنين ١٩٣١ — ١٩٣٦ بملايين الأمتار المربعة

أعمدة مزدوجة تمثل ظاهرتين فى شكل واحد

٣٧ — وفى بعض الأحيان يكون لدينا بيانات مزدوجة لعدة سنين ، مثل قيم الصادرات والواردات المصرية فى سنين متتالية ، أو مقادير محصول القطن والمساحات المزروعة قطناً فى سلسلة من السنين . وفى هذه الحالة نرسم أمام كل سنة

(١) الأرقام مأخوذة عن تقرير المستر سيلوس الملحق التجارى البريطانى فى مصر . انظر مجلة غرفة الاسكندرية التجارية ، عدد يونيه ١٩٣٨ ص ٢٥ . الرقم ٢٠ فى أعلى العمولة لسنة ١٩٢٣ خطأ وصحته ٣٠ . وكذلك رقم السنة صحته ١٩٣٣ .

عمودين متجاورين يمثلان قيمتى الظاهرتين فى هذه السنة ، بحيث يكون طول كل منهما متناسباً مع الرقم الذى يمثله . وهنا يحسن أن نميز بين هذه الأعمدة ، بألوان مختلفة أو بالتظليل مثلاً ، وذلك منعاً للالتباس . ونرى (فى شكل ٢) بياناً يوضح الصادرات والواردات المصرية فى السنين ١٩٣٣ — ٣٦ بهذه الأعمدة المزدوجة [الأرقام من نفس المرجع] .



(شكل ٢)

الصادرات والواردات المصرية بملايين الجنيهات

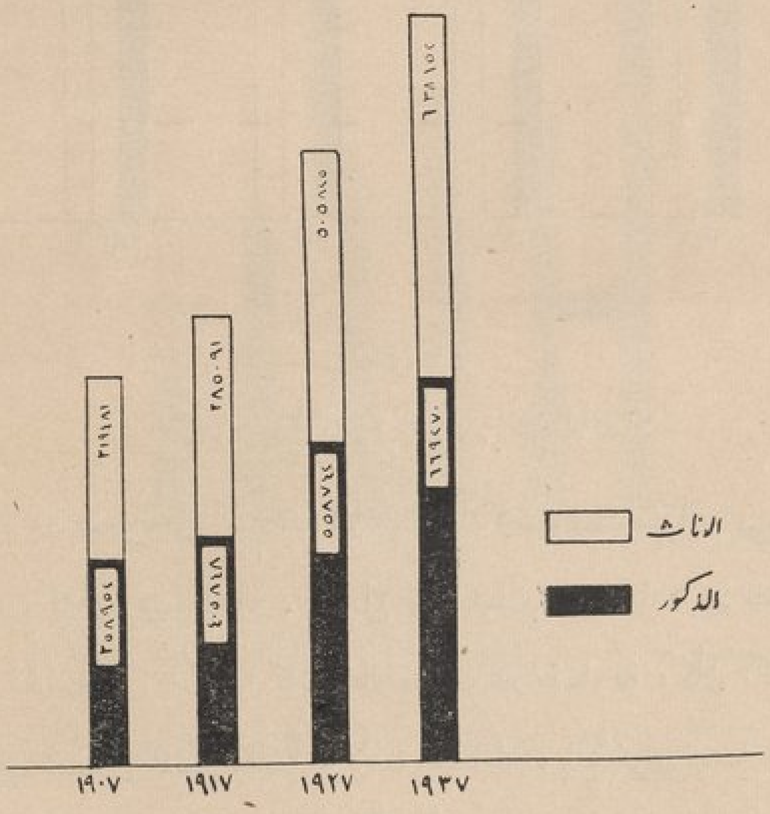
ولا نجد صعوبة فى رسم هذه الأعمدة إذا كانت وحدات القياس للظاهرتين متساوية ، كما فى هذه الحالة . فكل من الصادرات والواردات مقدرة قيمتها بملايين الجنيهات . أما إذا كانت الوحدات غير متساوية ، كأن تكون إحدى الظاهرتين محصول القطن مقدراً بآلاف القناطير ، والأخرى مساحة الأراضى المزروعة قطناً مقدرة بالأفدنة ، فيلزم أن نأخذ هذا فى الاعتبار عند تحديد أطوال الأعمدة التى نرسمها لكل من هاتين الظاهرتين .

ويصح أن نستخدم هذه الطريقة أيضاً إذا كان لدينا ثلاث سلسلات من القيم لثلاث ظواهر ، مثل مساحة الأراضى المزروعة أرزاً بالأفدنة ومقدار المحصول

بالطن، ومقدار المصدر من الأرز كل سنة. وفى هذه الحالة نرسم أمام كل سنة ثلاثة أعمدة تمثل هذه المقادير الثلاثة. ولكن يخشى أن تؤدى كثرة الأعمدة بهذا الشكل إلى زيادة التعقيد وضياع الفائدة المرجوة، وهى الإيضاح مع البساطة. ولهذا السبب أيضاً لا يستحسن استخدام هذه الطريقة إذا كان لدينا أرقام لعدد كبير من السنين، ويحسن أن نلجأ إلى طريقة أخرى.

أعمدة مقسمة إلى أجزاء

٣٨ — فى بعض الأحيان يكون لدينا أرقام جزئية تكوّن جملة عامة؛ مثلاً أرقام المصدر من أصناف القطن المختلفة وجملة المصدر من القطن عن كل سنة؛



(شكل ٣)

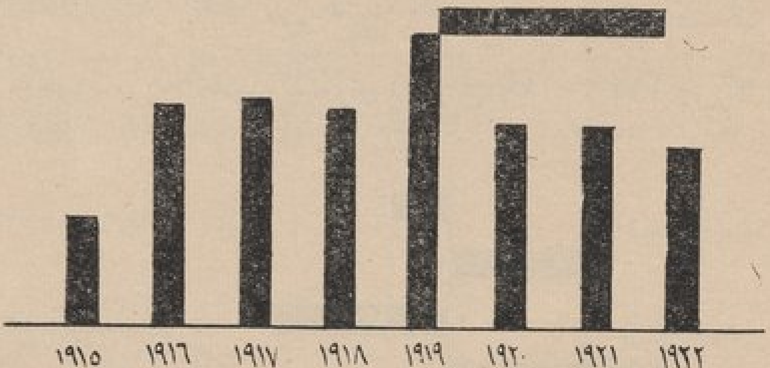
عدد سكان القاهرة من ذكور وإناث

أو عدد السكان الذكور وعدد السكان الإناث وجملة السكان فى بلد معين، وهكذا. يمكننا توضيح هذه البيانات فى شكل واحد بواسطة أعمدة « تجميعية »

يتكون كل عمود من أجزاء — مميزة عن بعضها بألوان مختلفة — كل منها يمثل رقماً من الأرقام الجزئية ؛ ومجموع هذه الأجزاء — وهو طول العمود الكلى — يمثل رقم الجملة . ونرى (فى شكل ٣) تطبيق ذلك لتوضيح كيفية نمو عدد سكان القاهرة حسب تعدادات السنين : ١٩٠٧ و ١٩١٧ و ١٩٢٧ و ١٩٣٧ .

كسر الأعمدة الطويلة

٣٩ — وفى حالة ما يكون بعض الأعمدة أطول بكثير من الأعمدة الأخرى يحسن أن نكسر الجزء الزائد من العمود ونكمله أفقياً لمسافة مساوية ، حتى يمكن أن يسعه فراغ الورقة . ونرى هذه الطريقة موضحة فى شكل ٤ الذى يبين



(شكل ٤)

الرقم السنوى لسعر القطن فى المدة ١٩١٥ — ١٩٢٢ بالنسبة إلى سعر ١٩١٣ كأساس

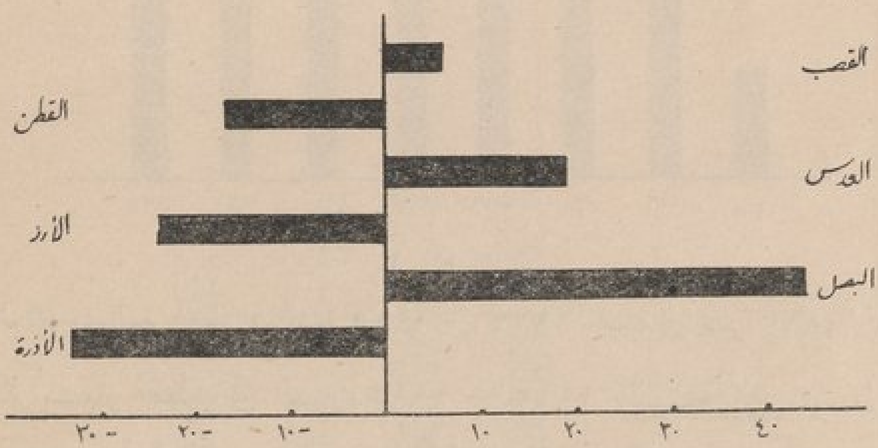
حركة أسعار القطن فى السنين ١٩١٥ — ١٩٢٢ بالنسبة إلى متوسط سعره فى سنة ١٩١٣ كأساس = ١٠٠ . [الأرقام مأخوذة من « الإحصاء السنوى العام » لسنة ١٩٣٥ — ١٩٣٦ صفحة ٥٠٤] .

وإذا كانت المسافة الأفقية الموجود فى الورقة لا تكفى فيصح أن نكسر العمود مرة ثانية إلى أسفل ونستكمل الطول اللازم . وإذا كانت الأعمدة الطويلة كثيرة يصح أن نصعد بها إلى ارتفاع معين ؛ والأجزاء الباقية نثنيها على شكل أقواس من دوائر . وعلى كل حال يجب أن نترك هذه التفاصيل للتصرف الشخصى بحسب ظروف كل حالة .

خطوط أفقية

٤٠ — فى بعض المسائل نستعمل خطوطاً أفقية لتمثيل البيانات الإحصائية. ولتوضيح هذه الطريقة نستخدمها لبيان مقادير الزيادة أو النقص (فى المائة) فى أسعار بعض المحاصيل الزراعية المصرية فى سنة ١٩٣٥ بالنسبة إلى أسعارها فى سنة ١٩١٣ كأساس (يساوى ١٠٠).

نرسم محوراً رأسياً نبدأ منه القياس. ولكل محصول نرسم خطاً أفقياً يتناسب طوله مع مقدار الزيادة أو النقص (فى المائة) فى السعر. ونرسم كل الخطوط التى تمثل الزيادة على يمين المحور الرأسى والخطوط التى تمثل النقص على يساره. ويحسن أن نرسم فى أسفل الشكل محوراً أفقياً نبين عليه مقياس الرسم كما هو مبين فى شكل ٥ [الأرقام مأخوذة من الإحصاء السنوى العام ١٩٣٥—١٩٣٦ ص ٥٠٤].



(شكل ٥)

مقادير الزيادة والنقص فى أسعار بعض المحاصيل فى سنة ١٩٣٥ عنها فى سنة ١٩١٣

المساحات : قطاعات الدائرة

٤١ — يمكن أن نستخدم المساحات بدل الخطوط أو الأعمدة لتمثيل البيانات. هنا تكون المساحات متناسبة مع الأرقام التى تمثلها، كما فى الخطوط أو الأعمدة.

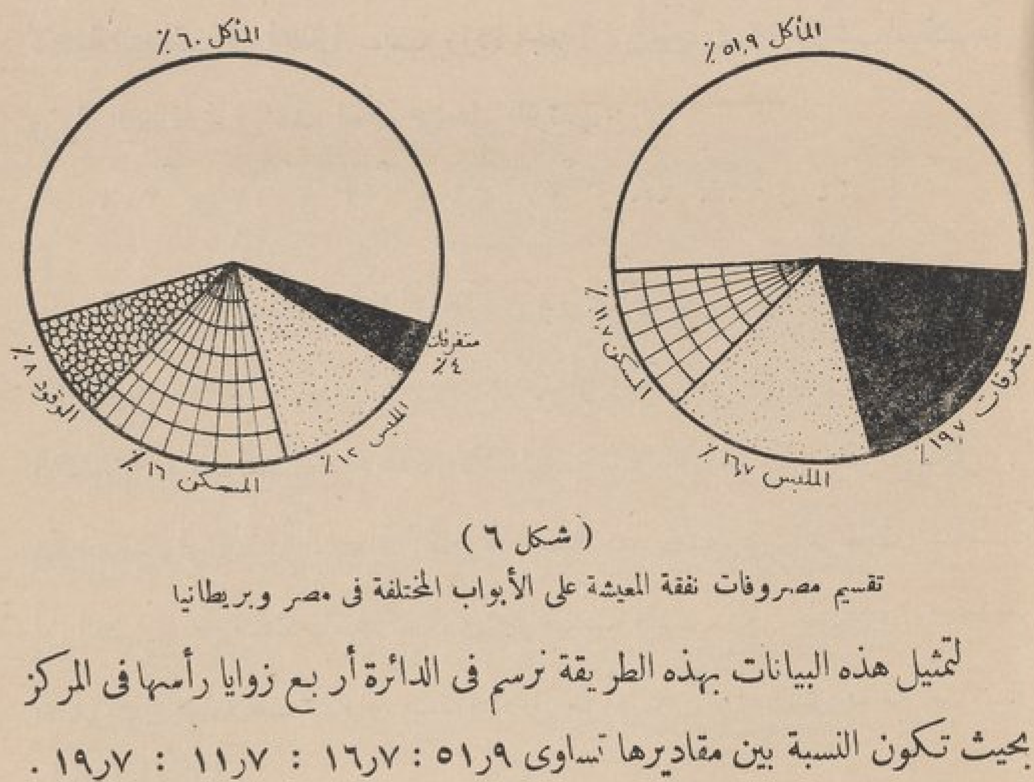
وربما كانت الدائرة أبسط الأشكال وأحسنها بياناً فى حالة استعمال المساحات

كطريقة للتوضيح ، خصوصاً إذا كانت الأرقام المراد توضيحها عبارة عن نسب مئوية من كمية واحدة.

فى هــذه الحالة تمثل الجملة العموميــة بالمساحة الـكلية للدائرة ، ونقسمها إلى قطاعات تتلاقى فى المركز بحيث تكون مساحاتها متناسبة مع المقادير الجزئيــة التى تكون الجملة العمومية . وهــذه القطاعات نميزها عن بعضها بألوان مختلفــة لزيادة الإيضاح .

خذ مثلا مصروفات الأسرة العادية وتوزيعها بين الأشياء المختلفة الضرورية للمعيشة . تبين من بحث عملته مصلحة الإحصاء المصرية فى سنة ١٩٢٠ أن الأسرة العادية توزع مصروفاتها على الأبواب الرئيسية بنسـبة ٥١ر٩ ٪ للأكل، و١٦ر٧ ٪ للملابس، و ١١ر٧ ٪ للمسكن، والباقى أى ١٩ر٧ ٪ للمتفرقات[1].

القطر المصرى بريطانيا



(شكل ٦)

تقسيم مصروفات نفقة المعيشة على الأبواب المختلفة فى مصر وبريطانيا

لتمثيل هذه البيانات بهذه الطريقة نرسم فى الدائرة أربع زوايا رأسها فى المركز بحيث تكون النسبة بين مقاديرها تساوى ٥١ر٩ : ١٦ر٧ : ١١ر٧ : ١٩ر٧ .

(١) هذه النسب عدات بعد الحرب الأخير .

وبما أن مساحة قطاع الدائرة تتناسب مع زاوية رأسه ، ينتج أن النسبة بين مساحات القطاعات المرسومة بهذا الشكل تساوى نفس النسبة أى ٥١٫٩ : ١٦٫٧ : ١١٫٧ : ١٩٫٧ .

وبما أن مجموع الزوايا التى يمكن رسمها فى مركز الدائرة يساوى أربع قوائم أى ٣٦٠° ، فنقسم العدد ٣٦٠ بالنسب المطلوبة تنتج الزوايا الآتية ومجموعها يساوى ٣٦٠ درجة :

٢٤ ٥٠ َ ١٨٦° ، ١٢ ً ٧ َ ٦٠° ، ١٢ ً ٧ َ ٤٢° ، ١٢ ً ٥٥ َ ٧٠°

ونرى هذه القطاعات فى الشكل المرافق (رقم ٦) ، وهو يوضح هذا التقسيم فى مصر . وبجانبه أيضاً تقسيم المصروفات المعيشية فى بريطانيا حيث النسب[1] هى ٦٠ ٪ للأكل و ١٢ ٪ للملابس و ١٦ ٪ لإيجار المسكن و ٨ ٪ للوقود والإنارة (وهذا باب جديد لا يوجد له نظير فى المصروفات المصرية ، وذلك نظراً لأهمية التدفئة فى انجلترا بسبب برودة الجو) ؛ وأخيراً ٤ ٪ للأشياء المتنوعة .

وزوايا القطاعات فى هذه الحالة هى على الترتيب :

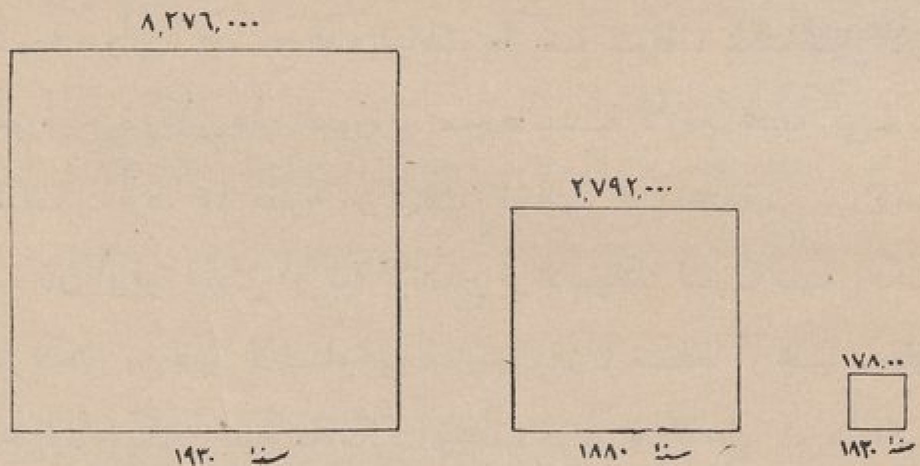
٢١٦° و ١٢ ٤٣° و ٣٦ َ ٥٧° و ٤٨ َ ٢٨° و ٢٤ َ ١٤° .

استخدام المربعات : مساحاتها تتناسب مع مربعات الأضلاع

٤٢ — ويصح استخدام أشكال غير الدائرة فى طريقة التمثيل بالمساحات ، مثل المربع ؛ وهو أسهلها وأحسنها ولو أن طريقة الدائرة تمتاز عنه فى أغلب الأحيان . وعند استخدام هذه الأشكال يحسن رسم شكل منفرد لكل رقم وتوضع بجوار بعضها لسهولة المقارنة . ويلاحظ عند رسم المربعات المتناسبة أن النسب بين مساحاتها تساوى مربعات النسب بين أضلاعها ؛ فالمربع الذى طول ضلعه سنتيمتران مساحته أربعة أمثال المربع الذى ضلعه سنتيمتر واحد ، وهكذا .

(١) هذه النسب عدلت ابتداء من يوليه سنة ١٩٤٧

لنأخذ مثلا محصول القطن المصرى فى السـنين ١٨٣٠ و ١٨٨٠ و ١٩٣٠ وهو ، على الترتيب ، ١٧٨٠٠٠ و ٢٧٩٢٠٠٠ و ٨٢٧٦٠٠٠ من القناطير[1] . ونرى فى شكل ٧ ثلاثة مربعات مساحاتها تتناسب مع هـذه الأعداد ، أى أن الأضلاع تتناسب مع جذورها . وهذه الجذور بنسبة ١ : ٣٫٩٦ : ٦٫٧٨ تقريباً .



(شكل ٧)

محصول القطن المصرى بالقنطار فى سنى ١٨٣٠و١٨٨٠و١٩٣٠

استخدام الأشكال المجسمة لتمثيل البيانات

٤٣ — يمكننا أيضاً تمثيـل البيانات الإحصائية بصورة مجسمة بواسـطة استخدام الأشكال الهندسيه المجسمة المعروفة ، مثل الكرة أو المكعب ؛ وهنا تكون النسبة بين أحجام الأجسام التى تمثل أعداداً معينة تساوى النسبة بين هذه الأعداد . ويلاحظ فى هذه الحالة أن حجم الجسم ، مثل المكعب أو الكرة ، يتناسب مع مكعب الضلع ، فى حالة المكعب ؛ أونصف القطر ، فى حالة الكرة . فلو أردنا تمثيـل الأرقام المذكورة فى البند السابق بواسطة ثلاثة مكعبات ، كانت أضـلاعها متناسبة مع الجـذور التكعيبية للأعداد ١٧٨٠٠٠ و ٢٧٩٢٠٠٠ و ٨٢٧٦٠٠٠ أى بنسـبة ١ : ٢٫٥٠ : ٣٫٥٨ تقريباً . وكذلك لو أردنا تمثيلها

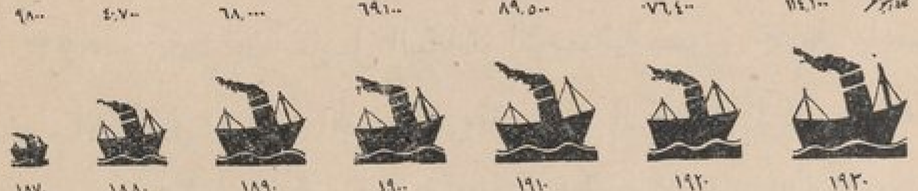
---

(١) الأرقام مأخوذة عن تقرير الملحق التجارى البريطانى السابق ذكره (حاشية ص ٢٦)

بكرات كانت النسبة بين أنصاف أقطارها تساوى هذه النسبة نفسها .

الصــور والأشكال

٤٤ — يمكن أن نستبدل بهذه الأشكال الهندسية المسطحة أو المجسمة رسوماً أو صوراً معينة تكون لهـا دلالة خاصة ذات صلة متينة بالموضوع الذى نتكلم فيه . فلو حصلنا مثلاً على أرقام لعدد البواخر التى مرت بقنـاة السويس فى عدة سنين متتاليـة ، وأردنا توضيـح هذه البيانات بطريقـة مشوقة ، يمكننا تمثيل الأرقام بصورة باخرة وتكبير هذه الصورة أو تصغيرها بنسـبة الأرقام المختلفة التى لدينا . وكذلك إذا أردنا عمل مقارنة بين الممالك (أو التواريخ) المختلفة من حيث تعداد السـكان أو قوة الجيش أو كمية الإنتـاج أو الاستهلاك لسلعـة معينة وهكذا ، نختار لـكل من هذه الأشياء التى نقـارنها رمزاً (مسطحاً أو مجسماً ، حقيقياً أو خياليـاً) يكون بسيطاً ما أمكن ، واضحاً ، سريع الدلالة على الشىء أو السلعة المقصودة بالمقـارنة ؛ ثم نكبر أو نصـغر هذا الرسم لـكل مملـكة (أو تاريخ) بقدر ما يناسب الرقم الخاص بهذه المملـكة (أو التاريخ) .



(شكل ٨)
عدد البواخر التى مرت فى قناة السويس من سنة ١٨٧٠ إلى سنة ١٩٣٠ (١٩١٣ = ١٠٠ كأساس)

وتحديد المقاسـات لـكى تتناسب مع الأرقام التى تمثلها بهذه الرسوم والصور يكون بنفس القواعد السـابق ذكرها فى البنـدين السـابقين على حسب كون الرمز أو الصـورة يدل على شكل مسطح أو مجسم . فإذا كان الأول أخذت النسب كما فى حالة المسـاحات المذكورة فى بند ٤٢ . و إذا كان الثـانى اتبعت قاعدة تناسب الحجوم (بند ٤٣) .

وفى شكل ٨ نرى تطبيقاً لهذه الطريقة لتوضيح الأرقام (القياسية بالنسبة إلى سنة ١٩١٣ كأساس يساوى ١٠٠) الخاصة بعدد البواخر التى مرت فى قناة السويس من سنة ١٨٧٠ إلى سنة ١٩٣٠ . وفى هذا الرسم اعتبرنا الأشكال مسطحة ، ولذلك فأبعاد هذه القطع متناسبة مع الأرقام على حسب قاعدة المساحات .

ويلاحظ أن هذه الطريقة مشوقة ، فهى تعطى للبيانات الرقمية الجافة صورة واقعية واضحة ، تفهم بسهولة وترسخ فى الذهن بسرعة وبدون عناء عقلى كبير . وفيها مجال واسع للتصرف فى اختيار الوسائل الإيضاحية الناجحة التى تساعد فى إبراز الحقائق بطريقة جذابة تترك أثراً فى نفس القارىء لا يمحى بسهولة ولو طال الزمن .

٤٥ — الطرق التى ذكرناها هنا ماهى إلا أمثلة لما يمكن عمله فى هذه الناحية ؛ والمجال متسع للتصرف الشخصى فى كل حالة بحسب الظروف المحيطة بها ، مثل البيانات المطلوب عرضها والأوساط التى تعرض فيها . ويلاحظ أن الطرق المتقدم ذكرها نتوخى فيها السهولة والوضوح خصوصاً من ناحية المقارنة (الزمانية أو المكانية) . ولا نهتم كثيراً بتحفيظ الأرقام نفسها لأن هذا يكلف الشخص العادى عناء ربما ينفره من الموضوع .

تبويب البيانات وعمل الجداول

التقسيم إلى مجموعات متجانسة

٤٦ — رأينا عند الكلام فى طرق جمع البيانات الإحصائية ، فى الباب السابق ، أن هذه البيانات تأتينا من مصادر مختلفة ، علاوة على أنها تتناول عدة نواح وعناصر مرتبطة بعضها ببعض إلى درجة ما . ويتعذر على أى شخص — أو يستحيل عليه — أن يلم بهذه البيانات ويستوعبها ليهتدى إلى الحقيقة بدون أن يجمع شتاتها . وهذا يكون بتبويبها وتقسيمها إلى فروع أو مجموعات متجانسة .

ويراعى فى هذا التبويب أن تشمل المجموعة الواحدة كل المفردات المتحدة فى صفة معينة من الصفات المهمة فى الموضوع ، أو عدة صفات مرتبطة ببعضها . ولا بأس من تقسيم هذه المجموعات الرئيسية إلى فروع ، وهذه إلى أقسام إذا اقتضى الحال ، وهكذا .

ففى إحصاء للأجور مثل الذى أشرنا إليه فى بند ٣٤ يصح أن نقسم البيانات التى نحصل عليها عن الأجور إلى أقسام بحسب الصناعة التى يشتغل فيها العامل ، أو بحسب الجهة التى فيها مكان العمل . و إذا كانت المجموعات التى نحصل عليها كبيرة العدد وتسمح بتقسيم آخر ، يمكن أن نقسم العمال فى كل صناعة إلى فنيين وغير فنيين مثلا ، وهكذا .

وعلى كل حال فطريقة التبويب وتعيين الصفات أو الخواص التى تتخذ أساساً لهذا التبويب لابد تتوقف على الغرض المقصود من عمل الإحصاء ، والتفاصيل التى نحصل عليها عند جمع البيانات اللازمة .

عمل الجداول العادية

٤٧ — بعد تقرير النظام الذى نتبعه فى التبويب . وتعيين الصفات التى تميز المفردات التابعة لكل مجموعة ، نرصد البيانات التى حصلنا عليها فى جدول مناسب يوضح هذه المجموعات والصفات المميزة لها .

والجدول العادى عبارة عن ترتيب خاص يبين تقسيم البيانات من ناحيتين معينتين . فيمكننا رسم جدول يبين تقسيم المصانع الموجودة بمدينة الإسكندرية مثلاً : أولاً بحسب الجهة أو القسم من المدينة الكائنة فيه هذه المصانع ؛ وثانياً من ناحية كون هذه المصانع تستخدم عمالاً أو لا تستخدم أحداً . والجدول الآتى يبين هذا التقسيم حسب تعداد سنة ١٩٢٧ [١] .

---

(١) انظر التعداد الصناعى والتجارى لسنة ١٩٢٧ (ص ١١٢)

(جدول ١) عدد المصانع بمدينة الإسكندرية فى سنة ١٩٢٧

القسم	مصانع بها مستخدمون	مصانع ليس بها مستخدمون	جملة
العطارين	١٣٦٦	٤٧٨	١٨٤٤
الجمرك	٨٠٧	٤٥٣	١٢٦٠
كرموز	٧٩١	٤٩٨	١٢٨٩
اللبان	٨٠٠	٣٥٧	١١٥٧
المنشية	٨٢٧	٤٠٠	١٢٢٧
مينا البصل	٣٦٩	٢١٦	٥٨٥
محرم بك	٤٦٥	٢١٢	٦٧٧
الرمل	٣٤٥	١٥٦	٥٠١
جملة	٥٧٧٠	٢٧٧٠	٨٥٤٠

ويصح تقسيم أحد الأقسام إلى فروع جزئية ، فمثلا المصانع التى بها مستخدمون يمكن تقسيمها إلى فئات بحسب عدد المستخدمين : واحدة تشمل المصانع التى تستخدم من ١ إلى ٤ مثلا ؛ وأخرى تشمل المصانع التى تستخدم من ٥ إلى ٩ ؛ وثالثة للمصانع التى تستخدم ١٠ فأكثر، وهكذا . وهنا يمكن تقسيم العمود الثانى فى هذا الجدول إلى أربعة أعمدة جزئية : واحد لكل من هذه الفئات ، والرابع للجملة . وبالطبع هذا التقسيم لا يمكن إذا لم تكن لدينا البيانات مفصلة .

والقاعدة العامة فى تصميم الجداول هى أن ننظر إلى تقسيم البيانات

ن ناحيتين فقط ؛ ونجعل لكل قسم من أقسام الناحية الأولى عموداً خاصاً تسجل فيه الأرقام الخاصة به ؛ ونجعل لكل قسم من أقسام الناحية الثانية سطراً أفقياً تدون فيه البيانات الخاصة به . وكل بيان لابد أن يكون له صفتان : الأولى تعين القسم الذى ينتمى إليه من أقسام الناحية الأولى ؛ والصفة الثانية تعين القسم الخاص من أقسام الناحية الثانية . وعلى ذلك فكل بيان يرصد فى الجدول عند ملتقى العمود والسطر اللذين تعينهما الصفتان . ولا يمكن أن يرصد البيان الواحد فى أكثر من مكان واحد إلا إذا كان تقسيم الصفات غير محدد ؛ وهذا عيب كبير فى تصميم الجدول ، وخطأ يؤدى إلى الخلط والالتباس . فى الجدول السابق مثلا نجد ٧٩١ مصنعاً أمام قسم كرموز ، وتحت المصانع التى بها مستخدمون . ومعنى ذلك أن هناك ٧٩١ مصنعاً كلها كائنة فى دائرة هذا القسم وكلها معين فيها مستخدمون . فمن الخطأ وضع هذا البيان فى أى مكان آخر فى الجدول . ويجب أن يشمل السطر الأول مثلاً كل المصانع التى فى دائرة قسم العطارين دون سواها ، ويوضع فى العمود (٢) كل المصانع التى بها مستخدمون فقط .

فرز البيانات عملية أساسية لعمل الجداول

٤٨ — هذه الأرقام التى نراها فى جدول (١) لايمكن الحصول عليها مباشرة من البيانات التى ترد لنا من أصحاب المصانع أو المصادر الأخرى عند عمل الإحصاءات الخاصة أو التعدادات العامة . وقد قلنا إننا عند جمع البيانات الإحصائية من هذا النوع ، نطلب من الأشخاص أو الهيئات ملء كشوف مطبوعة نرسلها لهم . وهذه الكشوف عندما نحصل عليها لابد من « فرزها » بحسب الأنواع التى يشملها التقسيم المتفق عليه ؛ ثم نعد مفردات كل قسم على حدة فنحصل على الأرقام التى نراها فى مثل هذا الجدول . فى حالة المصانع مثلاً ، نفرز الكشوف أولاً على حسب الجهات . وفى كل جهة نفصل المصانع التى بها مستخدمون من غيرها ، ونعد كل قسم على حدة ؛ ثم نضع الأرقام الناتجة فى الجدول .

وهـذه العملية يمكن إجراؤها بسهولة إذا كان عدد الكشوف صغيراً ، وكانت البيانات بسيطة وغير معقدة . ولكنها ، فيما عدا ذلك ، صعبة جداً ومرهقة للغاية ، ولا يمكن إجراؤها إلا باستخدام الوسائل الآلية الآتى شرحها . وسنشرح الآن طريقة يدوية سهلة يمكن استخدامها فى الإحصاءات الصغيرة ، التى لا تتعدى بضع مئات .

نرسم الجدول الذى نقسم على نظامه البيانات ، ونجعل الخانات ، فى الأعمدة والسطور ، واسعة نوعاً . ثم نتناول الكشوف التى لدينا واحداً بعد واحد ؛ ونعين لكل كشف الخانة التى يدخل تحتها بحسب البيانات المذكورة فيه ، ونضع فى هذه الخانة إشارة نصطلح عليها (نقطة بسيطة بالقلم أو شرطة صغيرة مثل — أو /) . وعدد هـذه الإشارات فى كل خانة يدل على عدد المفردات التى تخصها بحسب التقسيم المتبع فى الجدول . فلو اتبعنا هذه الطريقة لعمل جدول (١) مثلا (فى الواقع لا نستعمل هـذه الطريقة فى مثل هذه الحالة لأن العدد كبير جداً) فسنجد فى الخانة ملتقى السطر الأول والعمود الثالث ٤٧٨ إشارة من هذا النوع . وهذا يدل على وجود مصانع بهذا العدد فى دائرة العطارين ليس بها مستخدمون . ويمكن تسمية هذه العملية « تفريغ » البيانات فى الجدول .

ولتسهيل عملية العدّ يحسن ، عند وضع الإشارات فى الخانات المناسبة ، أن نجعل كل خمس منها بجوار بعضها كوحدة للعدّ منفصلة عن غيرها من الوحدات بفراغ بسيط (مثل ///// ///// /////) . ويكون عدد الإشارات فى الخانة يساوى عدد هذه الوحدات مضروباً فى ٥ مضافاً إليه الباقى . والأحسن أن نشطب على كل أربعة خطوط أو شرط بالخط الخامس هكذا فتكون المجموعة على شكل « حزمة » تدل على خمس مفردات ، وهـذا أوضح وأكثر اقتصاداً للفراغ الموجود فى الجدول .

الوسائل الآليــة للتبويب

طريقة البطاقات المثقوبة

٤٩ — قلنا إن هذه الطريقة اليدوية لايمكن استخدامها فى الإحصاءات الكبيرة ، ولا بد من الاستعانة بالوسائل الآليــة لتسهيلها . وأحسن الآلات المستخدمة لهذا الغرض هى الآلات المبنيــة على نظام البطاقات المثقوبة . ويوجد شركتان تصنعان هذه الآلات فى الوقت الحاضر ، وهما « هُولَرِيثْ » و « پَوَرْزْ — ساماس »[1] ؛ وهما تحتكران هذه الآلات فى جميع بلاد العالم .

والفكرة الأساسية فى هذه الآلات هى رصد البيانات المطلوب فرزها أو تبويبها أولاً على بطاقات خاصة مقسمة إلى عدد من الأعمدة ، كل منها به عشرة سطور مرقومة من ٠ إلى ٩ ، (كما فى شكل ٩) .

(شكل ٩)

بطاقة من ٢٦ عموداً مستعملة لإحصاء الأجور

تثقيب البطاقات بحسب البيانات

وهذا الرصد يكون بعمل ثقوب فى هذه البطاقة فى مواضع معينة بحسب البيان المطلوب رصده . ليكن هذا البيان هو أجر عامل معين فى مصنع

(١) آلات (Powers-Samas, Hollerith) . الأشكال ٩ — ١٣ معارة من شركة بورز — مع الشكر . والطريقة تسمى بالإنجليزية : (Punched-Card System)

فى القاهرة و يساوى ١٧٥ مليما فى اليوم مثلا . نعين أربعة أعمدة فى البطاقة لرصد الأجور (من ١٤ إلى ١٧ فى شكل ٩) ؛ ونثقب فى العمود الأول من اليمين (خانة الآحاد) الرقم ٥ ، وفى العمود الثانى (خانة العشرات) الرقم ٧ ، وفى الثالث (خانة المئات) الرقم ١ . وهكذا كل البيانات الخاصة بهذا العامل، مثل الاستقطاعات وصافى الكسب وعدد الأيام ، نرصدها على نفس البطاقة بعد تعيين أعمدة خاصة لكل نوع من البيانات . وفى حالة البيانات غير الرقمية ، مثل مكان المصنع : (القاهرة أو الإسكندرية) الخ ، أو نوع الصناعة : (مثلا النسيج أو الأحذية الخ) ، يجب أن نستخدم أرقاماً معينة تدل على هذه الأماكن أو الصناعات ، فالقاهرة مثلا نعطيها رقم ١ والإسكندرية رقم ٢ وهكذا . وكذلك فى الصناعات وغيرها من البيانات اللفظية . ونجد فى شكل ١٠ بطاقة مثقوبة ومعدّة لعملية الفرز والتبويب.

(شكل ١٠)

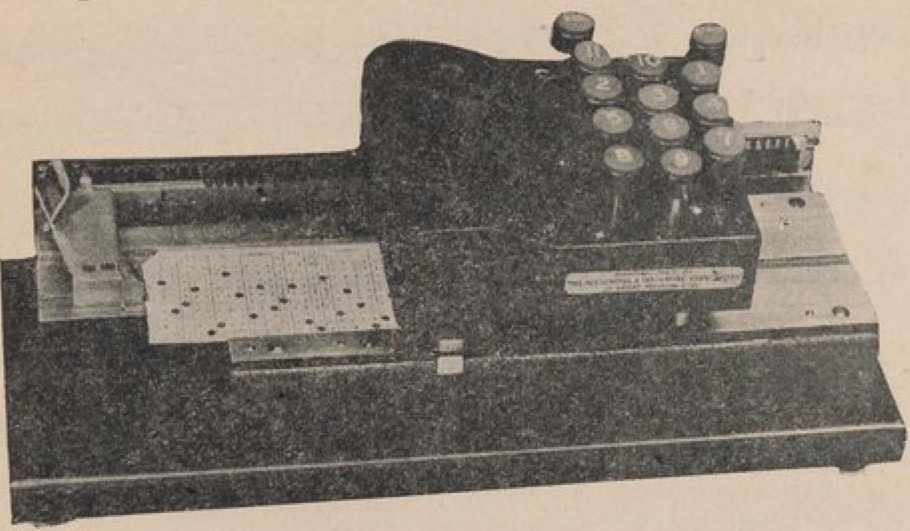
بطاقة مثقوبة معدة للفرز أو التبويب وقد سجلت بها بيانات خاصة بعملية بيع فى محل تجارى

ويوجد آلات خاصة لعمل هذه الثقوب ، وآلات أخرى لمراجعتها والتأكد من صحة مطابقتها للبيانات المطلوب رصدها . ويمكن عمل هذه الثقوب ومراجعتها بسرعة كبيرة ، تبلغ نحو ٣٠٠ بطاقة فى الساعة . ونرى آلة التثقيب فى شكل ١١ .

الفرز الآلى

٥٠ — بعد تثقيب البطاقات ومراجعتها توضع فى آلة الفرز لفصل الأنواع المختلفة . ويمكن شرح الفكرة فى عمل هذه الآلة باختصار كما يأتى :

توضع البطاقات فى مستودع خاص ؛ ومنه تخرج واحدة بعد الأخرى ، فتمر فوق سطح معدنى ، ويمسها من أعلى مشط معدنى أيضاً ، عدد أسنانه يساوى عدد الأعمدة فى البطاقة . وكل سن متصلة كهربائياً بالسطح المعدنى أسفل البطاقة بواسطة دائرة كهربائية تظل مفتوحة (لا يمر فيها تيار) ما دامت البطاقة (وهى من ورق عازل لا يوصل الكهربائية) حائلة بين السن والسطح المعدنى .

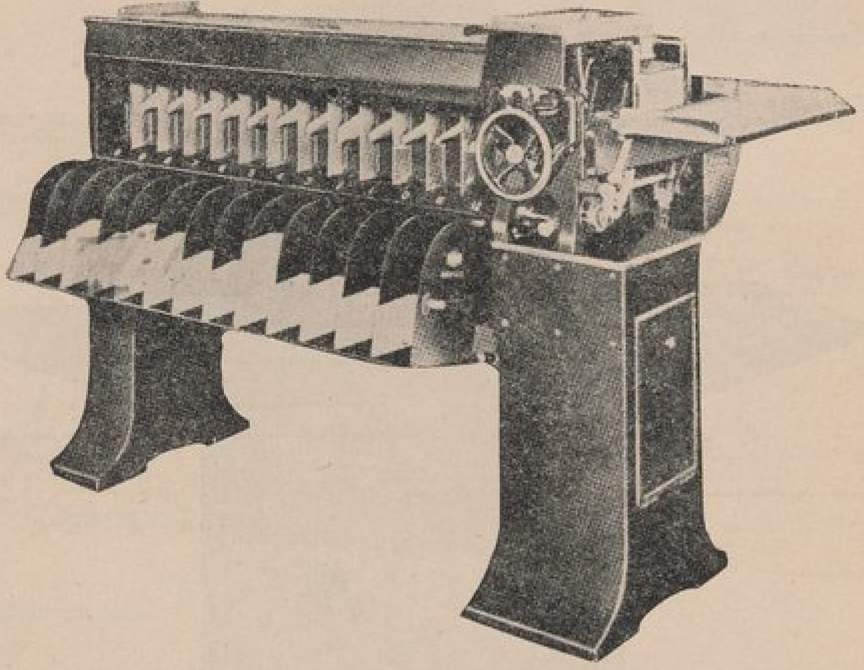


(شكل ١١)
آلة لتثقيب البطاقات

وإذا كانت البطاقة مثقوبة ، عند الرقم ٦ مثلاً ، فعند مرور الثقب تحت السن يحصل تماسٌ بين هذا الأخير والسطح المعدنى يقفل الدائرة الكهربائية عن طريق هذا الثقب عينه . وعند مرور التيار فى هذه الدائرة الكهربائية ينتقل تأثيره إلى صف من الصناديق (عددها عشرة ، مرقومة من ٠ إلى ٩) ، فيفتح غطاء الصندوق رقم ٦ فتقع فيه هذه البطاقة عند مرورها ؛ وكذلك تقع فيه كل بطاقة مثقوبة عند رقم ٦ فى هذا العمود . وهكذا نجمع فى كل من الصناديق العشرة البطاقات الخاصة به .

والمفروض هنا أننا نفرز البطاقات على أساس ثقوب كلها موجودة فى عمود واحد كل مرة . لنفرض أننا نريد فصل البطاقات حسب الجهات (القاهرة

والإسكندرية أو . .) ؛ ولتكن البيانات الخاصة بالجهات مرصودة بثقوب فى العمود العاشر من البطاقات . فى هذه الحالة نفرز على العمود رقم ١٠ ونترك كل الأعمدة الباقية ، أى أن السنّ رقم ١٠ فقط من المشط المعدنى هى التى تغلق الدائرة الكهربائية وتفتحها ؛ والأسنان الباقية تعلق عن العمل مؤقتاً .



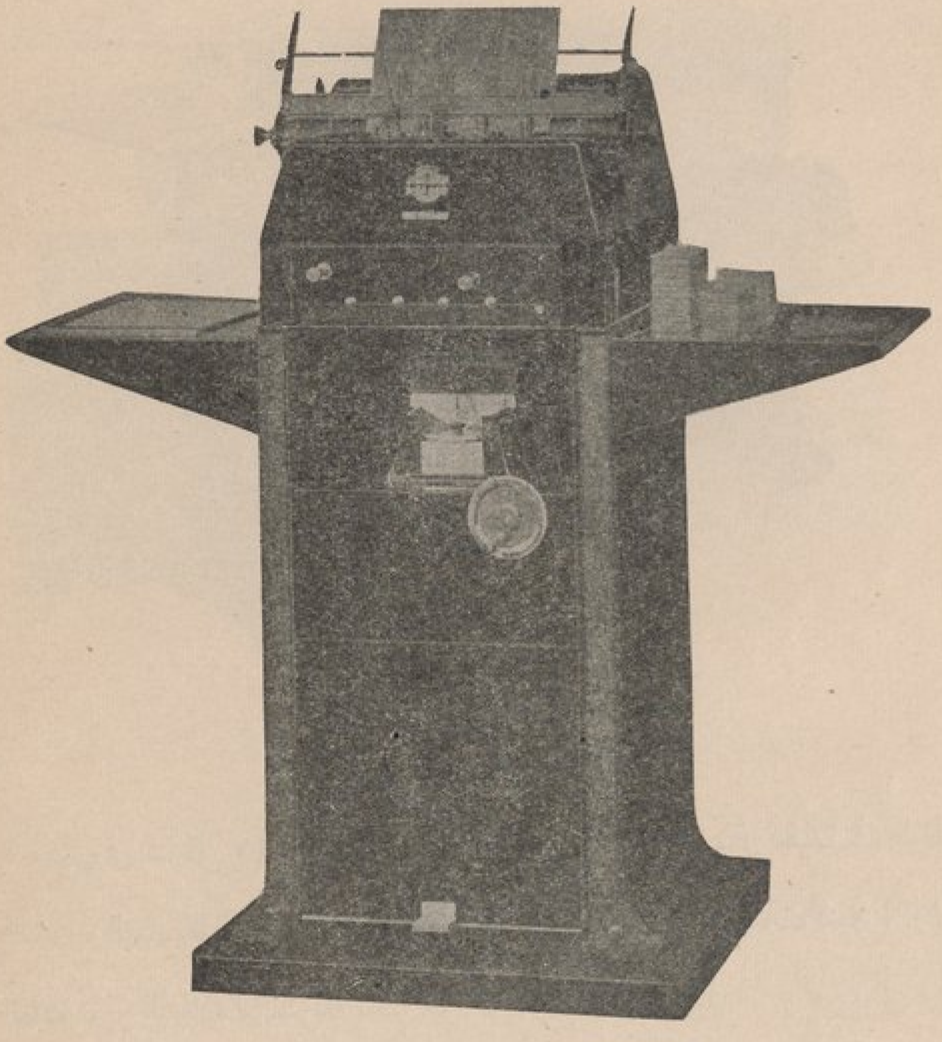
(شكل ١٢)
آلة كهربائية لفرز البطاقات

ونرى فى شكل ١٢ صورة لهذه الآلة ، وهى بالطبيعة أكثر تعقيداً فى الواقع من الشرح البسيط الذى قدمناه ؛ وكل يوم يضاف إليها تحسينات فنية . وهى فى العادة تفرز البطاقات بسرعة تبلغ نحو ٤٠٠ بطاقة فى الدقيقة . ويوجد فى هذه الآلات عدادات واحد على كل صندوق لمعرفة عدد البطاقات التى تنزل فيه ، وعدّاد لمعرفة جملة البطاقات المفروزة . وفى بعضها تتصل هذه العدادات بأداة تدوّن هذه الأرقام فى ورق ، قصداً فى المجهود الذى يبذل فى الكتابة ، ولتفادى الخطأ الذى قد يحصل عند قراءة الأرقام على العدادات وكتابتها . وهذه النقطة الأخيرة مهمة جداً من الناحية العملية .

آلات لعمل الجداول والتبويب

٥١ — العملية التى تلى فرز البطاقات، وربما كانت أشق منها وأكثر تعرضاً للخطأ، هى عملية تبويب البيانات، أى تفريغها من البطاقات المفروزة فى جداول مناسبة، وجمع الأرقام الخاصة بكل مجموعة أو فئة.

يوجد الآن آلات للقيام بهذه الأعمال كلها بغاية الدقة والسرعة؛ ولولاها



(شكل ١٣)

ما أمكن عمل الإحصاءات والتعدادات الكبيرة الواسعة النطاق. وهى آلات عمل الجداول (١).

والفكرة فى هذه الآلات أن تمر البطاقات المفروزة فى الآلة، واحدة بعد الأخرى، من مستودع خاص. وعند دخول البطاقة تمر بين السطح المعدنى السابق

(١) اسمها بالإنجليزية Tabulating Machines

ذكره من أسفل ومشط الأسنة من أعلى . وتتصل كل واحدة من هذه الأسنة بذراع يحمل حروف الكتابة . والمشط « يشعر » ، بواسطة الأسنة ، أين توجد الثقوب على البطاقة ، وينقل هذا « الشعور » إلى الأذرع التى تحمل الحروف : فتكتب ٦ مثلا إذا كان هناك ثقب عند ٦ فى البطاقة ، وهكذا . وتمر البطاقات تباعاً تحت المشط وتكتب الأرقام بدورها على الورق ، وتجمع على سابقتها فى الوقت نفسه ، حتى إذا انتهت مجموعة البطاقات كتب مجموع الأرقام من نفسه بدون احتياج إلى شخص يقوم بعملية الجمع . و بعد ذلك تأتى مجموعة أخرى ، وتنتهى وغيرها وغيرها ؛ وفى النهاية يكتب المجموع الكلى من نفسه أيضاً .

وفى شكل ١٣ صورة لواحدة من هذه الآلات . وهى ، طبعاً ، فى غاية التعقيد والإتقان ، وأقل ما توصف به أنها قطعة ثمينة من ثمرات التقدم العلمى فى هذا العصر . وسرعتها فى الكتابة تبلغ حوالى ١٠٠ سطر فى الدقيقة . وتوجد آلات أسرع من ذلك وآلات أكثر تعقيداً تقوم بعمليات كثيرة ومتنوعة فى وقت واحد و بسرعة هائلة .

ولا شك أن هذه الآلات قد أدت خدمات كبيرة فى سبيل تقدم علم الإحصاء العملى ، وساعدت كثيراً على الانتفاع بنتائج الإحصاءات والتعدادات الواسعة النطاق ، إذ جعلت من الممكن إخراج هذه النتائج فى وقت قصير جداً و بمجهود أقل وأيسر كثيراً مما كانت تتطلبه بدونها .

المراجع

Bowley, A.L., : *Elementary Manual of Statistics*, Chapter II.

Bowley, A.L., : *Elements of Statistics*, Chapters II., III.

Connor, L.R., : *Statistics in Theory and Practice*, Chapters II., IV.

Secrist, H., : *Introduction to Statistical Methods*, Chapters II., IV.

الباب الرابع

الرسوم البيانية

تعريف الخط البيانى

٥٢ — المنحنى البيانى هو خط يرسم بطريقة معينة لتوضيح العلاقة بين ظاهرتين أو كميتين متغيرتين . وبواسطته يرى الإنسان بسهولة كيف تتغير إحدى الظاهرتين مع الأخرى أو تبعاً لها .

وهذه الخطوط البيانية مستعملة كثيراً فى جميع العلوم ، خصوصاً التى تبحث فى الظواهر عن طريق مشاهدتها وتقديرها رقمياً . ولذلك سنشرح باختصار طريقة رسمها وخواصها .

رسم الخطوط البيانية

٥٣ — لنأخذ مثلا كمية الإنتاج المحلى فى مصر من المنسوجات القطنية فى السنين ١٩٣١—١٩٣٧ ، مقدراً بملايين الأمتار المربعة، وها هى الأرقام [1] :

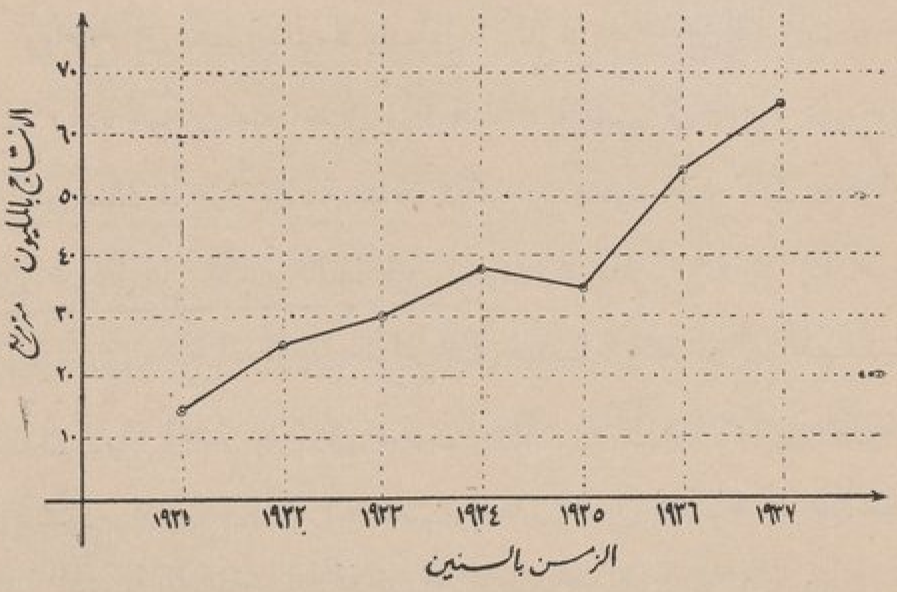
السنين	١٩٣١	١٩٣٢	١٩٣٣	١٩٣٤	١٩٣٥	١٩٣٦	١٩٣٧
الإنتاج	١٤ر٥	٢٥ر٠	٣٠ر٠	٣٧ر٥	٣٤ر٥	٥٤ر٥	٦٥ر٠٠

لرسم خط بيانى لكمية الإنتاج فى هذه السنين نأخذ ورقة مقسمة ونرسم عليها محورين متعامدين . نسمى ماتقى المحورين « نقطة الأصل » . ونأخذ السنين على المحور الأفقى مبتدئين من اليسار إلى اليمين . لذلك نحدد عليه سبع نقط

(١) مأخوذة عن تقرير الملحق البريطانى المذكور سابقاً فى هامش ص ٢٦ (رقم ١٩٣٧ تقريبى) .

على مسافات متساوية من بعضها ، وهى تمثل السنين من ١٩٣١ إلى ١٩٣٧ على الترتيب من اليسار إلى اليمين .

نقيس على المحور الرأسى مسافات متساوية تمثل الوحدات لقياس كمية الإنتاج. وليكن طول هذه المسافات المتساوية مناسباً بحيث يمكن تعيين نقطة فى فراغ الورقة تمثل كل كمية من كميات الإنتاج التى عندنا .



(شكل ١٤)

الإنتاج المحلى من المنسوجات القطنية فى مصر فى السنين ١٩٣١—١٩٣٧

أمام كل سنة تقيم عموداً على المحور الأفقى . ونقيس على المحور الرأسى (ابتداء من نقطة الأصل إلى أعلى) مسافة تساوى كمية الإنتاج فى هذه السنة ، ونرسم من نهايتها خطاً يوازى المحور الأفقى فيقابل العمود السابق ذكره فى نقطة وحيدة . هذه النقطة تدل ، فى نفس الوقت ، على السنة وعلى كمية الإنتاج فى هذه السنة .

الخط الذى يصل بين هذه النقط هو الخط البيانى للإنتاج فى هذه المدة . وهو يوضح كيفية تغير الإنتاج مع الزمن فى إثناء هذه المدة . ويكفى إلقاء نظرة

سريعة على الرسم (شكل ١٤) لكى تطبع هذه الصورة واضحة فى الذهن ، ويفهم معناها بسرعة و بدون عناء . وفى كثير من الأحيان يساعدنا الرسم البيانى ، بوضوحه وسهولته ، فى ملاحظة الخواص المهمة للظواهر التى نبحثها والعلاقات التى بينها . وهذا ربما لايتيسر لنا بالتأمل فى جداول مزدحمة بالأرقام ، حتى ولو أطلنا النظر إليها .

الخط البيانى المستوى يمثل العلاقة بين ظاهرتين فقط

وواضح أن الخط البيانى لا يتناول أكثر من ظاهرتين فى وقت واحد . لأننا نرسم محورين متعامدين ، كما قلنا ، ونقيس كل ظاهرة على محور . فإذا كان هناك ظاهرة ثالثة متغيرة لابد من وجود محور ثالث خاص بها يكون عمودياً على الاثنين السابقين . وهذا لا يمكن رسمه إلا إذا خرجنا عن مستوى سطح الورقة إلى الفراغ الذى فوقها . والنتيجة أن الخط البيانى يكون فى الفراغ المجسم بدلا من مستوى الورقة . ولو أن هذا ممكن تصوره عقلا وعمل نماذج مجسمة له ، إلا أنه صعب ومعقد . فنحن نقتصر هنا على الخطوط البيانية المستوية التى تبين العلاقة بين ظاهرتين فقط .

يراعى أن يكون الشكل مقبولا

٥٤ — والغرض الرئيسى من الرسم البيانى هو توضيح العلاقة بين الظاهرتين اللتين نبحثهما . فيجب الاهتمام بالناحية الفنية للأشكال التى نرسمها ، بحيث يكون منظرها العام مقبولا شائقاً ، خالياً من التعقيد بقدر الإمكان . وهذا يتوقف طبعاً على خبرة الشخص وذوقه .

مقياس الرسم على كل محور يناسب أرقام الظاهرة التى تقاس عليه

ويحسن أن يكون الخط البيانى واقعاً بالقرب من المحورين ما أمكن ، حتى يسهل مقارنة مواقع النقط عليه بالتدريج على كل منهما ، ولئلا يضيع فراغ الورقة بدون فائدة . ولهذا يجب أن نختار مقياسى الرسم على المحورين مناسبين للبيانات التى عندنا للظاهرتين . وليس من الضرورى أن يكون المقياسان

على المحورين متساويين ، فقد رأينا فى المثال السابق (شكل ١٤) أن وحدة الطول على المحور الأفقى تمثل سنة واحدة ، فى حين أن وحدة الطول على المحور الرأسى أخذناها تمثل ١٠ ملايين من الأمتار المربعة من القماش .

ولا يتحتم أن نبدأ القياس على أى المحورين من الصفر عند نقطة الأصل التى هى ملتقى المحورين ، بل يصح أن نبدأ بأصغر قيمة عندنا . وقد ابتدأنا على محور السنين فى الشكل السابق بالسنة ١٩٣١ بجوار نقطة الأصل . ولو حتمنا الابتداء بالسنة ١ بدلها لاحتجنا إلى مسافة طولها ١٩٣٠ سنتيمتراً حتى نصل إلى موقع السنة ١٩٣١ على المحور . ومع ذلك لا فائدة منها لعدم وجود بيانات عن الإنتاج فى هذه السنين ، ولا يمكن رسم أى شىء فى هذا الفراغ .

مقياس الرسم يؤثر فى مظهر تغيرات الظاهرة

٥٥ — الشكل الذى يأخذه الخط البيانى صعوداً وهبوطاً يتغير تبعاً لمقياس الرسم الذى نأخذه على كل من المحورين . والمقصود بمقياس الرسم هنا هو طول المسافة التى نأخذها على المحور الرأسى ، مثلا ، لتمثل الوحدة المستعملة فى قياس الظاهرة المأخوذة على هذا المحور . ففى الشكل السابق ، مثلا ، أخذنا مقياس الرسم على المحور الرأسى مسافة طولها أقل من مليمتر واحد لكل مليون متر مربع من النسيج ، وعلى المحور الأفقى سنتيمتر تقريباً لكل سنة .

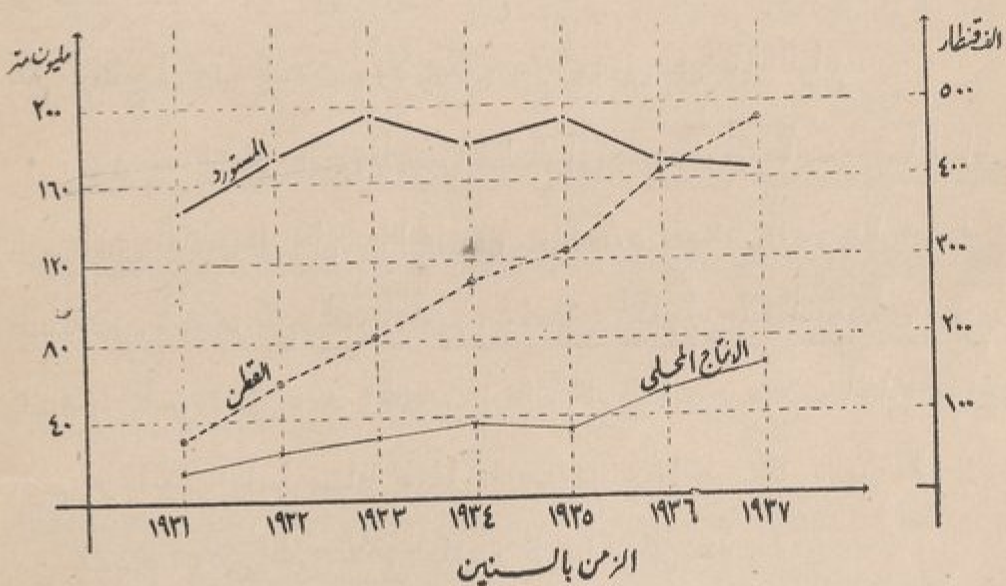
المقياس الكبير يبالغ فى مظهر التغيرات والمقياس الصغير يضعف من أهميتها

إذا كان مقياس الرسم على المحور الرأسى كبيراً بالنسبة للمقياس على المحور الأفقى ، فإن أى زيادة ، ولو بسيطة ، فى كمية الإنتاج تسبب ارتفاعاً كبيراً — نسبياً — فى الخط البيانى ؛ وانخفاض صغير فى الإنتاج يسبب هبوطاً لمسافة كبيرة فى المنحنى . إذن لو أخذنا مقياس الرسم كبيراً على المحور الرأسى تظهر التذبذبات فى المنحنى عنيفة . أى أن المقياس الكبير يبالغ فى شدة التغيرات التى تطرأ على الظاهرة . وبالعكس : المقياس الصغير يضعف من حدة هذه التغيرات فى نظر القارىء ، و يعمل على تمهيد المنحنى و إظهاره خالياً من التذبذبات العنيفة .

رسم أكثر من خط بيانى واحد فى شكل

٥٦ — نحتاج أحياناً لدراسة ظاهرتين أو أكثر من حيث تغيرهما بالنسبة للزمن أو لظاهرة أخرى مشتركة (أصلية) . مثلاً كمية الإنتاج المحلى من المنسوجات القطنية فى مدة معينة ، وكمية المستورد منها من الخارج فى نفس المدة ، وكمية القطن الخام المستهلك مثلا .

يمكن رسم خطين بيانيين أو أكثر فى نفس الشكل ، كل منها يوضح العلاقة بين ظاهرة من هذه ، والظاهرة المشتركة — وهى الزمن بالسنين فى هذا المثال .



(شكل ١٥)

الإنتاج المحلى والمستورد من المنسوجات والقطن المستهلك محلياً

فى هذه الحالة نأخذ الظاهرة المشتركة (الأصلية) على المحور الأفقى . ثم نرسم لكل واحدة من الظواهر الأخرى (التابعة) خطاً بيانياً يوضح علاقتها مع الظاهرة الأصلية ، بنفس الطريقة السابق شرحها : فنأخذ الظاهرة التابعة على المحور الرأسى ، ونقيسها عليه بمقياس رسم خاص بها . ولعدم الالتباس نميز هذه الخطوط بألوان أو نظم مختلفة .

إذا كانت وحدات الظواهر التابعـــة متفقة فيمكن عمل تدريج واحد على المحور الرأسى يستعمل للجميع . أما إذا كانت الوحدات مختلفة — كأن تكون إحدى الظواهر مقدرة بالقناطير مثلا والأخرى بالأمتار المربعة — فلا بد من عمل تدريج خاص لكل واحدة على المحور الرأسى فى الشكل . ويحسن حينئذ رسم خطين رأسيين متجاورين أو أكثر ، يبين على كل منها تدريج خاص بظاهرة واحدة من الظواهر التابعة . ويمكن وضع أحد الخطوط على يمين الشكل ، زيادة فى الإيضاح ومنعاً للالتباس .

ونرى (فى شكل ١٥) ثلاثة خطوط بيانية تصور فى نفس الرسم كمية الإنتاج المحلى من المنسوجات القطنية . وكمية المستورد من الخارج ، وكميـة القطن المستهلك محليـاً للغزل ، فى المدة ١٩٣١ — ١٩٣٧ ؛ والأرقام هى كالآتى [١] :

جدول (٢) المنتج محلياً والمستورد من المنسوجات والقطن المستهلك محلياً .

السـنة	الإنتاج المحلى مليون متر مربع	المستورد مليون متر مربع	القطن الخام ألف قنطار
١٩٣١	١٤ر٥	١٤٧ر٠	٧٨ر٥٠٠
١٩٣٢	٢٥ر٠	١٧٤ر٠	١٤٩ر٧٠٠
١٩٣٣	٣٠ر٠	١٩٧ر٠	٢٠٤ر٨٠٠
١٩٣٤	٣٧ر٥	١٨٢ر٠	٢٧٤ر١٠٠
١٩٣٥	٣٤ر٥	١٩٣ر٥	٣٠٦ر٦٠٠
١٩٣٦	٥٤ر٥	١٦٩ر٥	٤١١ر٨٠٠
١٩٣٧	٦٥ر٠	١٦٤ر٠	٤٧٨ر٤٠٠

(١) عن تقرير الملحق التجارى البريطانى (أرقام ١٩٣٧ تقريبية)

ويلاحظ فى الشكل أن مقياس الرسم للإنتاج المحلى والمستورد متساويان ، ومقياس الرسم لكمية القطن يخالفهما ؛ وقد أظهرناه منفرداً على يمين الشكل . ويلاحظ أيضاً أنه بينما نرى حركة الإنتاج المحلى والقطن المستهلك فى ازدياد مستمر نجد أن كمية المستورد أخذت فى الزيادة أولا ثم عادت فهبطت . وهناك فرق آخر وهو أن الزيادة فى القطن الخام المستهلك أسرع بكثير من الزيادة فى المنتج محلياً من النسيج ؛ والسبب فى ذلك أن بعض القطن المستهلك يغزل فقط ولا ينسج بل يصدر إلى الخارج فى صورة غزل .

استعمال ورق المربعات للرسم البيانى العادى

٥٧ — فى مثل المسائل المتقدمة نستعمل ورق المربعات العادى لرسم الخطوط البيانية المطلوبة . وهذا الورق ، كما نعلم ، مقسم بخطوط متوازية على أبعاد متساوية فى كل من الاتجاهين الأفقى والرأسى ، بحيث ينتج من تقابل هذه الخطوط مربعات متساوية الأضلاع قائمة الزوايا . وهذا الورق نستعمله فى المسائل العادية حيث نريد بيان العلاقة بين القيم المتناظرة للمتغيرين تحت البحث . ففى المثال المذكور فى بند ٥٣ درسنا العلاقة بين مقادير الإنتاج ، وهو المتغير « التابع » مَقِيساً بعدد الأمتار المربعة ، والزمن ، وهو المتغير « المتبوع » أو « المستقل » مَقِيساً بعدد السنين . وكل نقطة على الخط البيانى لها « إحداثيان » يمثلان قيمتين متناظرتين للمتغيرين : الإحداثى الأفقى يمثل السنة ، أى قيمة المتغير المستقل ، والإحداثى الرأسى يمثل كمية الإنتاج فى تلك السنة أى قيمة المتغير التابع المناظرة لها .

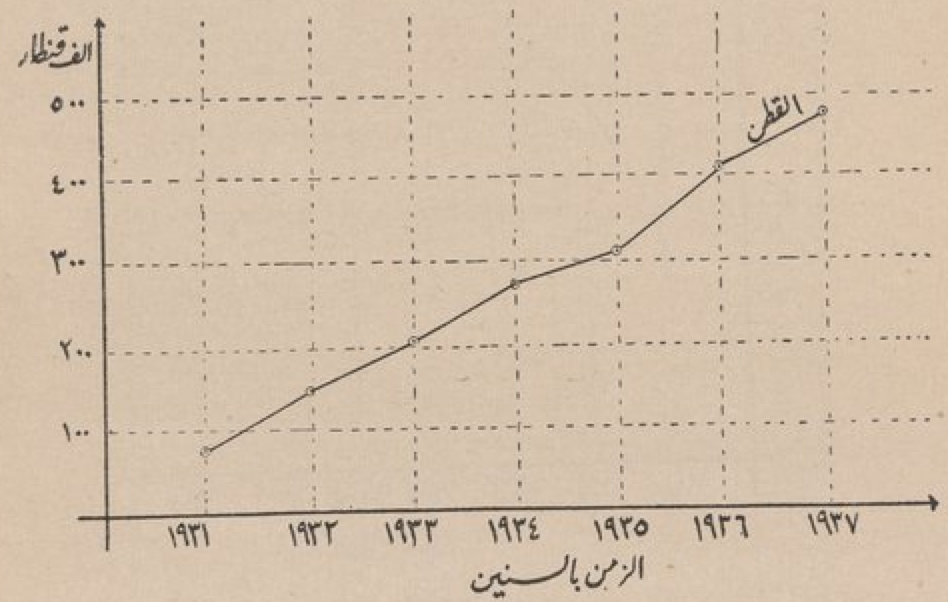
خط بيانى لقيم متغير مع لوغاريتمات قيم متغير آخر

٥٨ — فى بعض المسائل نريد دراسة العلاقة بين قيم أحد المتغيرين ولوغاريتمات القيم المناظرة للمتغير الثانى . ولرسم خط بيانى يمثل هذه العلاقة يمكننا استعمال الورق العادى . نأخذ قيم المتغير الأول على المحور الأفقى . ونستخرج لوغاريتمات قيم المتغير الثانى من جداول اللوغاريتمات بالطريقة العادية ؛ ونأخذ هذه

اللوغاريتمات على المحور الرأسى ، ونرصد النقط فى الشكل كالمعتاد ، ونصل بينها بخط يكون هو الخط البيانى المطلوب .

الورق اللوغاريثمى

ولكن هذه الطريقة عقيمة ومطولة . فعملية استخراج اللوغاريتمات من الجداول عملية متعبة ومعرضة للخطأ فى قراءة الأرقام أو نقلها . ويمكن تفادى هذا كله باستعمال ورق « لوغاريثمى » فيه المحور الرأسى مقسم إلى مسافات تساوى لوغاريتمات الأعداد الطبيعية ١ ، ٢ ، ٣ ، ... بواسطة خطوط أفقية بعرض الصفحة . وباستعمال هذا الورق نستغنى عن عملية استخراج اللوغاريتمات من الجداول ، ونرصد النقط فى الشكل من القيم المعطاة مباشرة ، ونحصل بسهولة على الخط المطلوب .



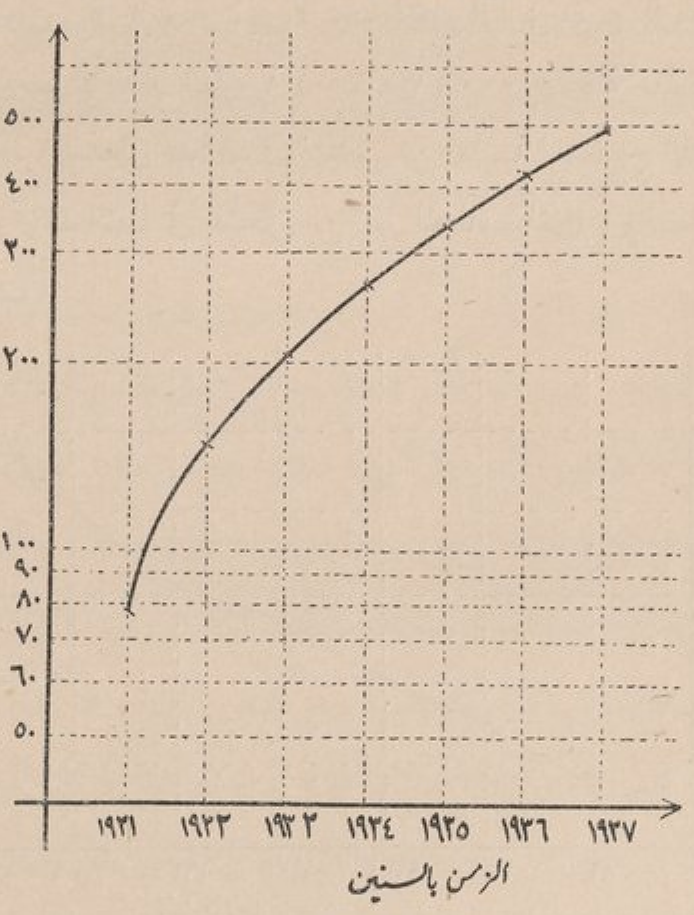
(شكل ١٦)

المستهلك من القطن فى المصانع المحلية فى مصر . تقسيم ورق عادى [١]

٥٩ — يجب أن نلاحظ هنا أن التقسيمات اللوغاريثمية على المحور الرأسى المقابلة للأعداد ١٠٠ و ٢٠٠ و ٣٠٠ مثلا لا تكون على أبعاد متساوية كما فى الورق العادى (شكل ١٦) لأننا نعلم أن لوغاريتم ٢٠٠ لا يساوى ضعف لوغاريتم ١٠٠

(١) فى هذا الشكل بالذات ظهرت التقاسيم مستطيلة بدلا من مربعة ، وذلك لاختلاف مقياس الرسم على المحورين .

فالمسافة التى تمثل لو ٢٠٠ على المحور الرأسى لا تساوى ضعف المسافة التى تمثل لو ١٠٠ على نفس المحور. والمسافة التى تمثل لو ٣٠٠ لا تساوى المسافة التى تمثل لو ٢٠٠ مرة ونصفاً لنفس السبب، وهكذا. ولكن لو ١٠ ولو ١٠٠ ولو ١٠٠٠ ولو ١٠٠٠٠ تكون على مسافات متساوية على المحور لأنها تساوى ١ و ٢ و ٣ و ٤ على الترتيب.



(شكل ١٧)

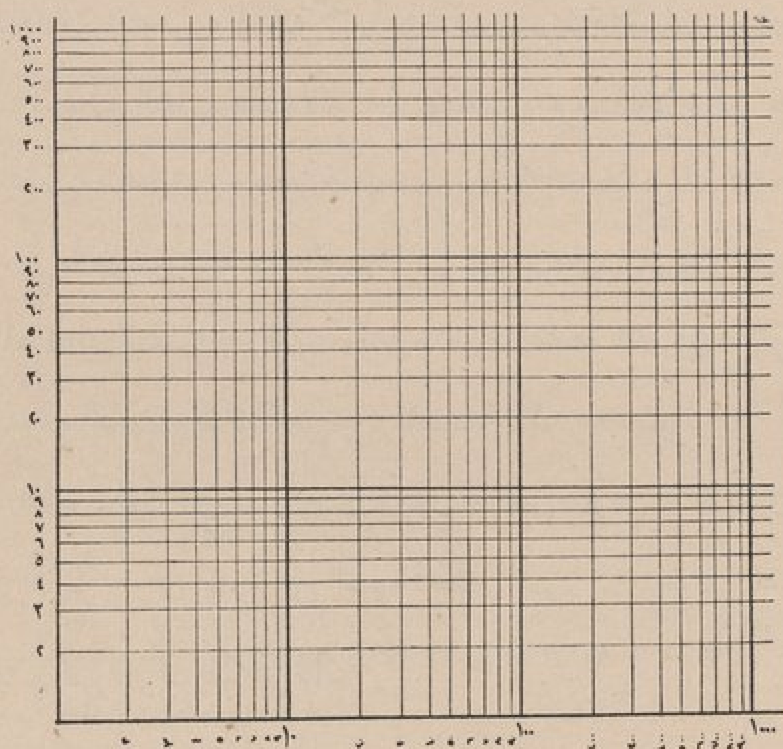
المستهلك من القطن بآلاف القناطير فى المصانع المحلية فى مصر (تقسيم نصف لوغاريثمى)

ينتج من هذا أن شكل الخط البيانى لظاهرتين، المرسوم على ورق مربعات عادى، يخالف شكل الخط البيانى المرسوم لنفس الظاهرتين على ورق لوغاريثمى. ويتضح ذلك من مقارنة الشكلين ١٦ و ١٧، حيث رسمنا فى الأول خطاً بيانياً على ورق مربعات عادى لكمية القطن الخام المستهلك فى مصانع الغزل والنسج المصرية فى السنين ١٩٣١ — ١٩٣٧، وفى الثانى (شكل ١٧) نجد نفس الأرقام

مرصودة على ورق « نصف لوغاريثمى » (أى فيه التدريج اللوغاريثمى على محور واحد فقط وهو الرأسى) . فبينما نجد الخط البيانى فى شكل ١٦ قريباً من الاستقامة ، ومُقعَّراً قليلا إلى أعلى، نجد الخط البيانى لنفس الأرقام فى شكل ١٧ منحنياً ومحدباً إلى أعلى . ويصح أن يكون الخط البيانى مستقيما على الورق العادى ومنحنياً على الورق اللوغاريثمى ، أو العكس .

الورق اللوغاريثمى المـــزدوج

٦٠ — ويصح أيضاً أن نستعمل تقسيماً لوغاريثميًّا على المحورين معاً ؛ وذلك إذا أردنا دراسة العلاقة بين لوغاريثمات قيم المتغير الأول ولوغاريثمات



(شكل ١٨)
تقسيم لوغاريثمي مزدوج

قيم الثانى . ولرسم الخط البيانى المطلوب فى هذه الحالة نرصد النقط على الشكل مباشرة من القيم الأصلية الموجودة لدينا بواسطة التدريج غاريثمى على كل من المحورين .

ونرى فى شكل ١٨ نموذجاً من التقسيم اللوغاريتمى المزدوج على المحورين فى وقت واحد. ويلاحظ أن الورقة هنا ليست مقسمة إلى مربعات متساوية ومتجاورة كما فى ورق المربعات العادى، بل إلى مربعات متداخل بعضها فى بعض لها قطر مشترك؛ ويتركب كل واحد منها من مربعين مشتركين معه فى القطر، يكملهما مستطيلان جانبيان.

الخطوط البيانية تكون فى العادة كثيرة التعاريج

٦١ — هذه الخطوط البيانية التى نرسمها — بناء على بيانات مأخوذة من التجارب العملية — تكون فى العادة كثيرة التذبذب وغير «ممهدة»؛ وذلك لأن مشاهدات التجارب العملية كثيراً ما تكون معرضة لتقلبات المصادفة. وهذه التقلبات لا يمكن التغلب عليها لأنها، بطبيعتها، عرضية ولا يمكن التنبؤ بها. فمثلا قد يخطىء الباحث — بدون قصد طبعاً — فى قراءة مقياس الحرارة أو الضغط، أو قد يخطىء الجهاز نفسه فى حالة ما لسبب غير معروف، أو يطرأ ظرف طارىء على الظاهرة التى نبحثها — فيزيد مقدارها زيادة خارقة للقانون، أو ينقص، وهكذا. هذه التقلبات الناشئة عن مجرد المصادفة تكون «عشوائية»[1] أى بدون ضابط. فتارة تكون بالزيادة وأخرى تكون بالنقص، وغير مقصودة فى أى اتجاه دون الآخر. وهى تسمى على العموم **أخطاء التجربة**؛ وأحياناً نسميها **تقلبات المصادفة** و **تغيرات عرضية أو فجائية**[2].

---

(١) عشواء مؤنث أعشى، بمعنى عمياء مؤنث أعمى. وهى مستعملة هنا بمعنى الكلمة الانجليزية (Random). وقد اقترح هذه الترجمة الدكتور عبد العزيز القوصى؛ وهو يستعملها فى أبحاثه الإحصائية الخاصة بعلم النفس. والمعنى واضح من قول الشاعر:

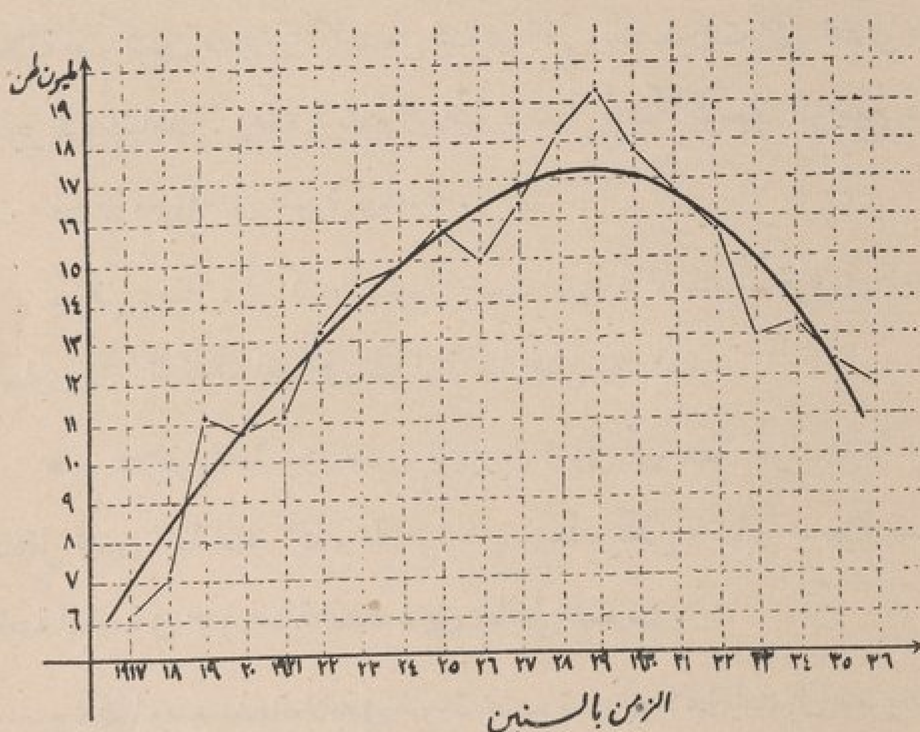
رأيت المنايا خبط عشواء: من تصب تمته؛ ومن تخطىء يعمر فيهرم

(٢) وبالانجليزية Experimental Errors ، Random Fluctuations ، Casual or Accidental Variations. على الترتيب.

تمهيد الخطوط البيانية باليد

٦٢ — وللتخلص من هـذه التذبذبات « نمهد » الخط البيانى الذى نحصل عليه من واقع البيانات الأصلية المأخوذة من التجربة والمشاهـدات ؛ وذلك بأن نغضّ النظر عن هـذه التذبذبات العرضية ، ونرسم خطاً ممهداً يتمشى مع الخط الأصلى ويهمل التفاصيل والتذبذبات الصغيرة .

وهذا يمكن عمله بالنظر بدون صعوبة . فبعد رصد النقط فى الشكل — من واقع الأرقام المعطاة لنا — نرسم باليـد خطاً يتوسط بين هذه النقط دون التقيد بالمرور بها جميعاً ، إذا كان هذا المرور يسبب تعريجات كثيرة فى الخط تفسد تمهيده؛



(شكل ١٩)

صافى حمولة السفن الإنجليزية المارة بقناة السويس

ونرى فى شكل ١٩ الخط البيانى لصافى حمولة السفن الإنجليزية المارة بقناة السويس فى المدة ١٩١٧ — ١٩٣٦ ؛ وفى نفس الشكل الخط الممهد بهذه الطريقة .

ويلاحظ فى هذه الطريقة أننا تركنا الحرية للشخص فى رسم الخط الممهـد

حسب ما يتراءى له ، وهذا طبعاً يتوقف على خبرته و إلمامه بظروف الظاهرة التى نبحثها ؛ ويتوقف أيضاً على دقته فى الرسم ومرانه ؛ و بناء على هذا ينتظر أن نحصل على خطوط ممهدة مختلفة إذا كلفنا أشخاصاً مختلفين بعملية التمهيد . وهذا هو العيب الذى يؤخذ على هذه الطريقة . وسنعود إلى شرح طرق أدق من هذه فى المستقبل .

معادلة الخط البيانى

وضع العلاقة فى صورة جبرية

٦٣ — تكلمنا فى هذا الباب على رسم الخطوط البيانية كوسيلة لتوضيح العلاقة بين كميتين متغيرتين كما تتمثل لنا فى القيم والتقديرات التى نحصل عليها بالتجربة والمشاهدة . وهذه الرسوم البيانية تستعمل أيضاً كوسيلة لتوضيح علاقة بين متغيرين معروفة فى صورة رياضية .

لنفرض ، مثلا ، أن كميتين تتغيران بحيث إن إحداهما تتبع الأخرى فى تغيرها ، وتساوى دائماً ثلاثة أمثالها فى المقدار زائداً خمس وحدات .

هذه علاقة واضحة بين هذين المتغيرين ؛ وفى أى لحظة نعرف قيمة المتغير المستقل تتعين قيمة المتغير التابع المناظرة لها . وهذه يمكن حسابها بسهولة بضرب القيمة المعلومة فى ٣ و إضافة العدد ٥ إلى حاصل الضرب .

و يمكننا وضع هذه العلاقة فى صورة مختصرة جداً وواضحة باستعمال بعض الرموز الرياضية البسيطة هكذا :

لنرمز للكمية المتغيرة الأولى (المستقلة) بالرمز الجبرى س والكمية الأخرى (التابعة) بالرمز ص . ومعنى هذا الرمز هو ، كما نعلم فى قواعد الجبر العادية ، أن الحرف س يدل على كمية متغيرة تأخذ قيما متعددة ومختلفة فى ظروف مختلفة ؛ ومجموعة هذه القيم كلها ، مهما كان عددها ، يرمز إليها بالحرف س و كذلك الحرف ص .

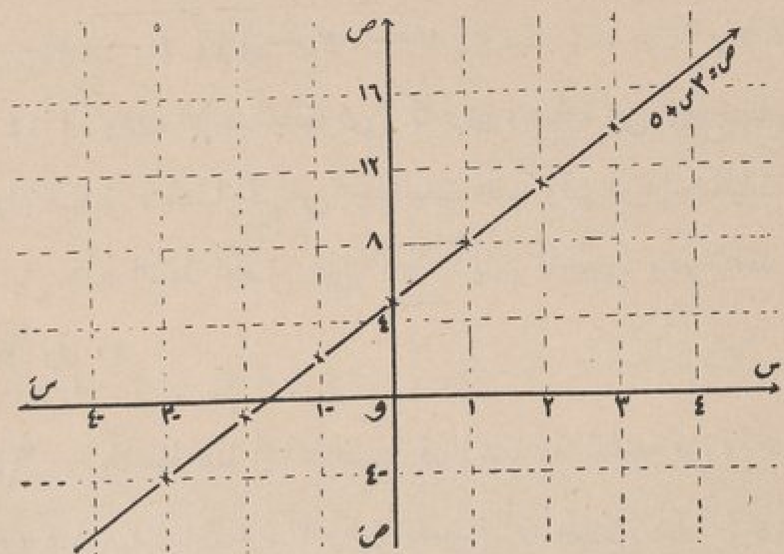
فإذا وضعنا

ص = ٣س + ٥ . . . . (١)

فهذا معناه أن س كمية متغيرة و ص كمية متغيرة أيضاً وتابعة لها ، بحيث إن قيمة ص تساوى دائماً ثلاثة أمثال قيمة س مضافاً إلى ذلك العدد ٥ . وهذه بلا شك ، صورة مختصرة واضحة للعلاقة التى نحن بصددها .

رسم الخط البيانى للعلاقة الرمزية

٦٤ — لكى نصور هذه العلاقة فى شكل خط بيانى نفرض عدة قيم مختلفة يأخذها المتغير المستقل س ؛ ولكل منها نحسب القيمة المناظرة التى يأخذها المتغير



(شكل ٢٠)

الخط البيانى للعلاقة ص = ٣ س + ٥

التابع ص بحكم هذه العلاقة المفروضة بينهما . ثم نضع هذه القيم التناظرة فى جدول . ومن هذا الجدول نرسم الخط البيانى بالطريقة المعتادة السابق شرحها فى أول هذا الباب . وفى العادة نأخذ قيم المتغير المستقل س على المحور الأفقى ، وقيم المتغير التابع ص على المحور الرأسى . وتقاس القيم الموجبة فى الاتجاه من اليسار إلى اليمين على المحور الأفقى ، ومن أسفل إلى أعلى على المحور الرأسى . و بناء على ذلك تقاس القيم السالبة فى الاتجاه المضاد على كل من المحورين .

نفرض أن س تأخذ القيم — ٣ ، — ٢ ، — ١ ، ٠ ، ١ ، ٢ ، ٣ على الترتيب ، ولكل واحدة نحسب قيمة ص المناظرة من العلاقة المفروضة وهى ص = ٣ س + ٥ ، فنحصل على الجدول الآتى :

س : — ٣ ، — ٢ ، — ١ ، ٠ ، ١ ، ٢ ، ٣

ص : — ٤ ، — ١ ، ٢ ، ٥ ، ٨ ، ١١ ، ١٤

نأخذ محورين متعامدين سَ و س ، صَ و ص متقاطعين فى نقطة الأصل و . ثم نرصد فى الشكل النقط التى إحداثياتها هى :

(— ٣ ، — ٤) و (— ٢ ، — ١) و . . . . . . و (٢ ، ١١) و (٣ ، ١٤) . وهذه الإحداثيات هى ، كما نرى ، عبارة عن أزواج القيم المتناظرة فى الجدول السابق ، باعتبار قيم س إحداثيات أفقية وقيم ص إحداثيات رأسية . وبتوصيل هذه النقط نحصل على الخط البيانى المطلوب وهو الخط الموضح فى الشكل رقم ٢٠

نحصل على نفس الخط إذا فرضنا قيما أخرى للمتغير س

٦٥ — يلاحظ هنا أننا فرضنا قيما اختيارية للمتغير س ، وهى — ٣ ، — ٢ ، — ١ ، ٠ ، ١ ، ٢ ، ٣ وكان يصح أن نفرض قيما أخرى غير هذه إما واقعة بينها أو بعيدة عنها فى أحد الطرفين ؛ وكان يصح أيضاً أن نأخذ أكثر منها عدداً أو أقل . ومهما كانت القيم التى نختارها والقيم المناظرة لها للمتغير التابع ص ، فالخط البيانى الذى نحصل عليه هو نفس الخط وينطبق عليه أو يقع على امتداده من أحد الطرفين أو الآخر . وهذا يمكن إثباته بالتجربة بسهولة . وهذا ، فى الحقيقة ، هو الواجب لأن هذا الخط البيانى يمثل علاقة ثابتة لا تتغير بين قيم س وقيم ص المناظرة لها . فيجب أن يكون الخط الذى نحصل عليه واحداً مهما كانت القيم التى نعطيها للمتغير الأصلى س .

ويلاحظ أيضاً أن أى نقطة على الخط ، خلاف النقط التى تمثل القيم المفروضة ،

تتوفر فيها هذه العلاقة ص = ٣ س + ٥ . فإذا أخذنا نقطة مثل حـ فسنجد (إذا كان الرسم دقيقاً) أن إحداثيها الرأسي يساوى ثلاثة أمثال إحداثيها الأفقى زائداً العدد ٥ ؛ أى أن العلاقة المفروضة مستوفاة . وهكذا إذا أخذنا أى نقطة أخرى على الخط أو امتداده من إحدى الناحيتين أو الأخرى .

الخط البيانى هو مسار نقطة تتحرك بشرط معين

و بعبارة أخرى نقول إن الخط المرسوم فى شكل ٢٠ هو الخط البيانى للعلاقة ص = ٣ + ٥ ، أو « هو الخط الذى معادلته هى ص = ٣ س + ٥ » ، أو « هو مسار النقطة التى تتحرك فى المستوى بحيث إن إحداثيها الرأسى ، فى أى موضع لها ، يساوى ثلاثة أمثال إحداثيها الأفقى زائداً العدد ٥ » ، أو « هو المحل الهندسى لنقطة تتحرك بهذا الشرط » .

معادلة المستقيم

كل معادلة من الدرجة الأولى يمثلها خط مستقيم

٦٦ — نلاحظ أيضاً أن الخط المرسوم فى (شكل ٢٠) خط مستقيم ، وأن معادلة هذا الخط ، وهى ص = ٣ س + ٥ ، من الدرجة الأولى بالنسبة إلى س و ص ، أى أنها لا تحتوى على قوى أعلى من الأولى لأى واحد من المتغيرين فلا نرى فيها حدوداً تحتوى على س^{٢} أو س^{٣} . . . ولا ص^{٢} أو ص^{٣} . . . وهذا التوافق ، بين صفة استقامة الخط البيانى وكون معادلته من الدرجة الأولى ، ليس مصادفة فى هذه المسأله فقط . ويمكن إثبات أن كل معادلة من الدرجة الأولى يمثلها خط مستقيم ، وأن كل خط مستقيم تمثله معادلة من الدرجة الأولى . ولكن المقام هنا لا يسمح بإيراد هذا البرهان ، ويجب أن يرجع القارىء إلى أحد الكتب فى الهندسة التحليلية[١] .

---

(١) انظر كتاب « الهندسة التحليلية » تأليف نصيف سعيد وصادق بشارة (١٩٣٥) الجزء الأول بند ١٣ و ١٤

الصورة العامة لمعادلة المستقيم

٦٧ — وكل معادلة من الدرجة الأولى بالنسبة إلى س و ص يمكن وضعها فى صورة بسيطة مكونة من ثلاثة حدود : حد يحتوى على س (مضروبة فى كمية ثابتة أو صفر) ؛ وحد يحتوى على ص (مضروبة أيضاً فى كمية ثابتة أو صفر) ؛ وحد مطلق ، خال من س و ص (وهو عبارة عن كمية ثابتة ؛ ويصح أن يكون صفراً) . والصورة العامة لمعادلة المستقيم هى المسماة «المعادلة الصريحة» ، للمستقيم وهى:

$$ص = م س + حـ$$

حيث س هى المتغير المستقل ، و ص المتغير التابع ، و م كمية ثابتة تسمى « ميل المستقيم على محور السينات » ؛ والحد المطلق حـ كمية ثابتة أيضاً . ويلاحظ هنا أن ص وحدها فى طرف من طرفى المعادلة الصريحة ، وباقى الحدود فى الطرف الآخر . ففى المعادلة السابقة مثلا : ص = ٣ س + ٥ نرى أن

$$م = ٣ \quad و \quad حـ = ٥ .$$

أى أن ميل هذا المستقيم على محور السينات يساوى ٣ ؛ وهو يساوى ظل الزاوية التى يصنعها هذا المستقيم مع محور السينات [1] ؛ وأن الكمية حـ = ٥ تساوى الجزء الذى يقطعه هذا المستقيم من محور الصادات ، كما هو واضح فى الشكل ، حيث نجد المستقيم يقطع محور ص عند النقطة (٠ ، ٥) .

وإذا كان لدينا معادلة مثل ١٥ س + ٣ ص — ١٢ = ٠ ، يمكن وضعها فى الصورة العامة بقسمة الطرفين على معامل ص وهو ٣ ، ثم ننقل السينات والحد المطلق إلى الطرف الأيسر فينتج أن :

$$ص = — ٥ س + ٤ .$$

---

(١) هذا يظهر فى الشكل ٢٠ عندما نأخذ فى الاعتبار اختلاف مقياس الرسم على المحورين فى هذا الشكل بالذات ، حيث السنتيمتر على محور س يساوى وحدة من وحدات س ، فى حين أنه على محور ص يساوى ٤ وحدات من وحدات ص .

فالمعادلة ، إذن ، تمثل مستقيماً ميله — ٥ ، ويقطع من محور الصادات جزءاً يساوى ٤ وحدات من وحدات ص .

معادلة القطع المكافىء

٦٨ — نأخذ الآن معادلة من الدرجة الثانية بالنسبة إلى س ، ونرسم الخط البيانى لها . ولنأخذ معادلة صريحة للسهولة ، فيها ص وحدها فى الطرف الأيمن ، والسينات والحد المطلق فى الطرف الأيسر . منحنى القطع المكافىء

المطلوب رسم الخط البيانى الذى معادلته :

$$ص = س^{٢} + ٢$$

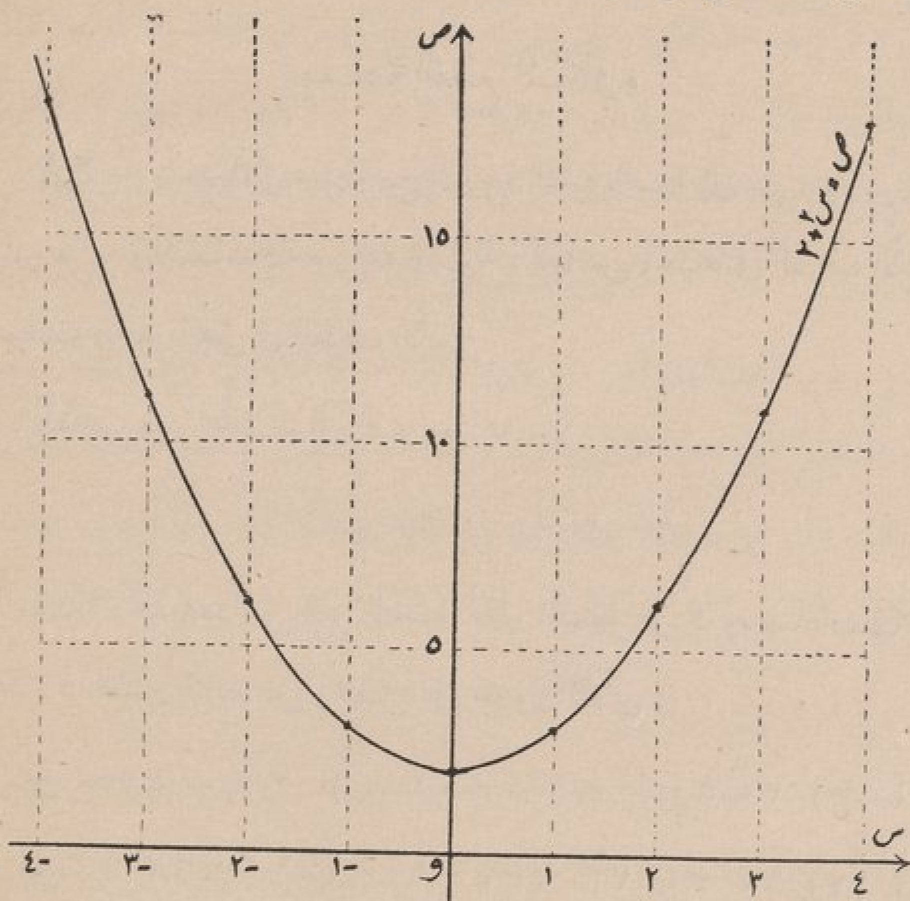
لهذا نأخذ عدداً من القيم المناسبة للمتغير المستقل س ؛ ومنها نحسب القيم المناظرة لها للمتغير التابع ص ونحصل على جدول كالآتى :

س :	— ٤ ،	— ٣ ،	— ٢ ،	— ١ ،	٠	١ ،	٢ ،	٣ ،	٤ ،
ص :	١٨ ،	١١ ،	٦	٣ ،	٢	٣ ،	٦ ،	١١ ،	١٨ ،

نرسم المحورين ونختار مقياسين مناسبين للرسم عليهما ، ثم نرصد النقط من واقع القيم المتناظرة الموجودة فى الجدول السابق ، ونصل بينها فنحصل على الخط البيانى المطلوب .

ويلاحظ أن هذا المنحنى متماثل بالنسبة إلى محور الصادات كما نرى فى شكل ٢١ . وهذا هو المنتظر كما هو واضح من المعادلة المعطاة ومن القيم الموجودة فى الجدول والمحسوبة بواسطة هذه المعادلة . فالقيمتان — ٣ و + ٣ للمتغير س مثلا ، تعطيان نفس القيمة للمتغير ص وهى ١١ ؛ لأن س تظهر فى الطرف الأيسر للمعادلة فى شكل $س^{٢}$ فقط ؛ وهذا التربيع لا يجعل للإشارة أثراً فى القيمة النهائية .

هذا المنحنى يسمى **القطع المكافئ** . ويكون متماثلا بالنسبة لمحور الصادات إذا لم تحتو المعادلة على س مفردة غير مربعة .

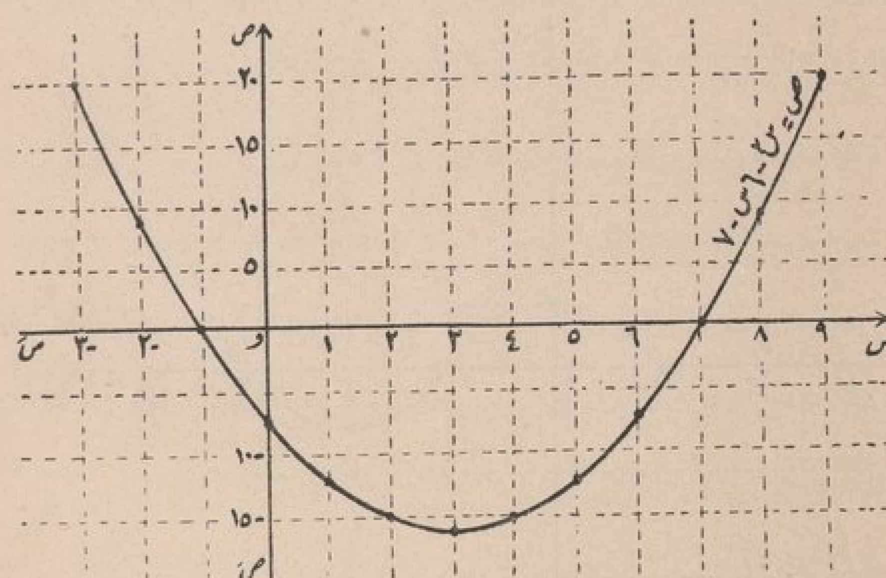


(شكل ٢١)
قطع مكافىء متماثل

أما إذا احتوت المعادلة على س بدون تربيع ، فلا يكون القطع متماثلا بالنسبة إلى محور الصادات ، ولكن محور تماثله حينئذ يكون مستقيما رأسياً غير محور الصادات . ونرى فى شكل ٢٢ قطعاً مكافئاً معادلته $ص = س^{٢} - ٦ س - ٧$ ، وهو متماثل حول مستقيم رأسي يوازى محور الصادات ويبعد عنه بمسافة تساوى ثلاث وحدات من وحدات س . وهذا المستقيم يقابل القطع فى نقطة تحت محور السينات ، إحداثياها + ٣ و — ١٦ ، كما هو واضح من الشكل . وهذه النقطة تسمى **رأس القطع المكافئ** ؛ والمستقيم يسمى **محور القطع المكافئ** .

ويصح أن يكون محور « تماثل » القطع المكافىء موازياً لمحور السينات أو منطبقاً عليه . وتكون معادلته حينئذ على الصورة المتقدمة مع وضع س بدل ص ووضع ص مكان س .

وقد رأينا القطع المكافىء فى هذين الشكلين مقعراً إلى أعلى ورأسه فى أسفل . وذلك لأن معامل س^{٢} فى المعادلة موجب فى كلتا الحالتين . أما إذا كان معامل س^{٢} فى معادلة القطع المكافىء سالباً فإنه يظهر فى الشكل مقلوباً ، ويكون محدباً إلى أعلى وتكون رأسه أعلى نقطة فيه . وهذه خواص



(شكل ٢٢)

قطع مكافىء محور تماثله يوازى محور ص

ثابتة لهذا المنحنى ، ويحسن تذكرها فهى تساعدنا على تعرف المنحنيات والحكم من شكلها على صورة المعادلة التى تمثلها . فنعلم مثلا أن كل معادلة من الدرجة الثانية يمثلها قطع مكافىء ، وأن كل قطع مكافىء له رأس واحدة فقط . وهذه تكون على شكل نهاية عظمى أو صغرى للمنحنى على حسب ما تكون إشارة س^{٢} سالبة أو موجبة على الترتيب .

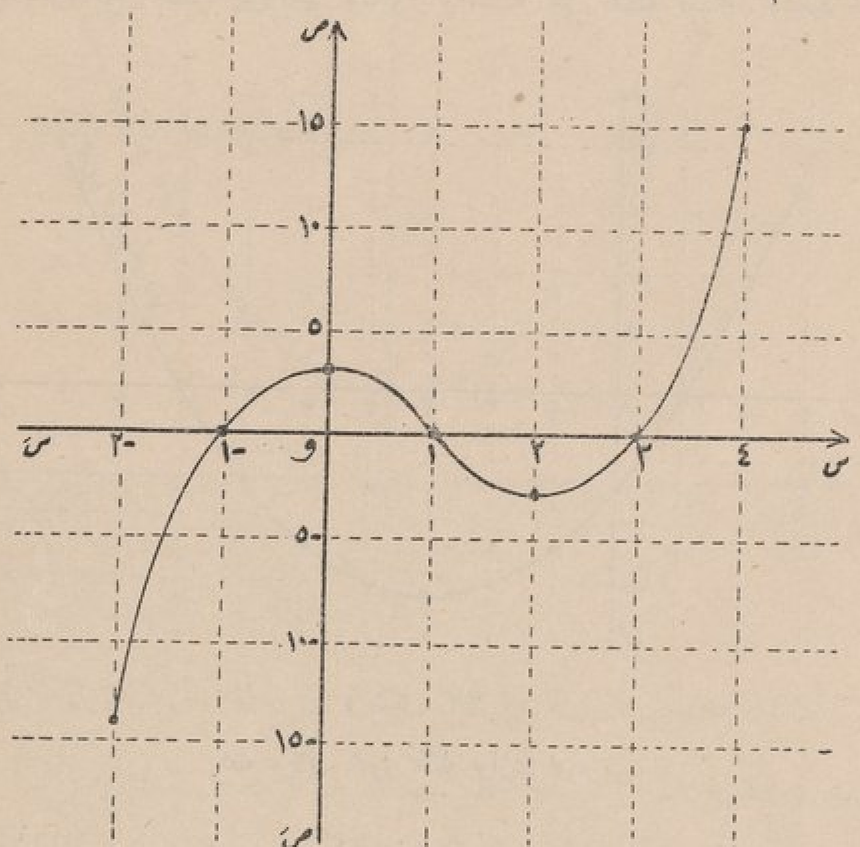
منحنى الدرجة الثالثة

٦٩ — نأخذ الآن معادلة من الدرجة الثالثة بالنسبة إلى س ، ونرسم المنحنى البيانى لها لكى ندرس بعض الخواص العامة التى تميزه .

المطلوب رسم المنحنى الذى معادلته هى :

$$ص = س^{٣} - ٣س^{٢} - س + ٣$$

نفرض عدة قيم مناسبة للمتغير س ونحسب قيم ص المناظرة لها ؛ ومن هذه القيم المتناظرة نرسم المنحنى بالطريقة العادية . وهو ، كما نرى فى شكل ٢٣ ، خط ينحنى



(شكل ٢٣)

منحنى معادلته من الدرجة الثالثة

على نفسه مرتين : مرة محدب إلى أعلى بشكل نهاية عظمى ، ومرة محدب إلى أسفل بشكل نهاية صغرى . وفى الجزء الأيسر يبتدىء المنحنى من أسفل و يصعد حتى يصل إلى نهاية عظمى ؛ ثم ينحنى إلى أسفل حتى يصل إلى نهاية صغرى ، و بعدها

يصعد ثانياً و يستمر فى الصعود بدون حد . أى أن المنحنى يغير اتجاه سيره مرتين فقط ، بخلاف القطع المكافىء من الدرجة الثانية ، إذ يرجع فى مسيره مرة واحدة فقط .

و ينتج من ذلك أن منحنى الدرجة الثالثة له نقطتا « رجوع » وليس أكثر من ذلك : واحدة عندها نهاية عظمى ، والأخرى عندها نهاية صغرى . ولا يمكن أن تكون النهايتان من نوع واحد ومتتاليتين .

فى هذا المنحنى نجد النهاية العظمى أولا وتليها النهاية الصغرى إذا اتجهنا مع المنحنى فى الاتجاه سَ و س ، أى الاتجاه الذى تزيد فيه س؛ والعكس يحصل إذا كانت إشارة $س^٣$ سالبة فى معادلة المنحنى (أى مخالفة لإشارة ص) .

عدد نقط الرجوع فى أى منحنى

٧٠ — و إذا رسمنا منحنياً من الدرجة الرابعة فلن نجد له أكثر من ثلاث نقط رجوع مهما كان . وعلى العموم فعدد نقط الرجوع لأى منحن درجته ن مثلا ، يساوى (ن — ١) على الأكثر ، و يصح أن يكون أقل من هذا ولكن لا يزيد . وهذه النظرية يمكن إثباتها ولكن لا يتسع المجال لإثباتها هنا؛ و يحسن بالقارىء أن يرجع إلى بعض كتب الجبر أو التفاضل والتكامل .

منحنى الخطأ المعتدل

٧١ — المطلوب رسم المنحنى الذى معادلته هى :

$$ص = هـ^{—س^٢}$$

حيث هـ هى الكمية ٢٫٧١٨٢٨١٨٢٨٤ (وهى أساس اللوغاريثمات الطبيعية المستعملة فى الرياضة)[١].

---

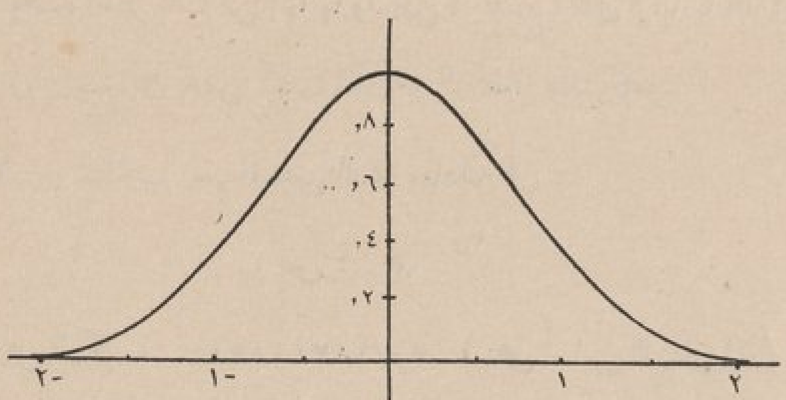
(١) هذه الكمية هـ مشهورة جداً فى العلوم الرياضية :

$$هـ = ١ + ١ + \frac{١}{٢!} + \frac{١}{٣!} + \frac{١}{٤!} + \ldots \text{ إلى } \infty .$$

ويرمز لها فى الإنجليزية بالحرف e ، وتسمى المتسلسلة الأسية (Exponential Series) . أنظر كتاب الجبر الابتدائى ، جزء ٢ ، تأليف هول ونايت .

نفرض عدة قيم مناسبة للمتغير س ، ونحسب لكل منها القيمة المناظرة للمتغير ص . ويلاحظ أن س هنا مربعة وموجودة فى الأس ترفع إليها الكمية هـ . ويوجد جداول[1] للقوى المختلفة لهـــذه الكمية نستخرج منها القيم المطلوبة فنحصل على الجدول الآتى ، مع العلم بأن قيمة ص لا تتغير بتغير إشارة س :

س	ص	س	ص
	٠ر١	٨ر	٥٢٧٣ر
١ر	٩٩٠١ر	٩ر	٤٤٤٩ر
٢ر	٩٦٠٨ر	٠ر١	٣٦٧٩ر
٣ر	٩١٣٩ر	٢ر١	٢٣٦٩ر
٤ر	٨٥٢٢ر	٥ر١	١٠٥٤ر
٥ر	٧٧٨٨ر	٠ر٢	٠١٨٣ر
٦ر	٦٩٧٧ر	٥ر٢	٠٠١٩ر
٧ر	٦١٢٦ر	٠ر٣	٠٠٠١ر



(شكل ٢٤)

ويلاحظ أن هـــذا المنحنى متماثل بالنســبة إلى محور الصادات الذى يقسمه إلى نصفين متطابقين . ويلاحظ أيضاً أن طرفيه يتقاربان شيئاً فشيئاً من محور س

(١) أنظر مثـــلا C.G. Knott, *Four-Figure Mathematical Tables*, p. 40.

ولا يقطعانه بل يمسانه فى نقطتين بعيدتين (عند ∞ من الطرفين) .

وهذا المنحنى مشهور جداً فى علم الإحصاء ؛ وهو أساسى جداً فى دراسة النظريات التى ينبنى عليها هذا العلم . وهذا المنحنى معروف بجملة أسماء نذكرمنها[١] « منحنى جاوس » ، نسبة إلى العالم الألمانى كارل ف . جاوس الذى استنبطه رياضياً ودرس خواصه . ويسمى أيضاً « منحنى الخطأ المتماثل » ، و « منحنى التكرار المعتدل » ، وهكذا . وسنعود إلى شرح بعض خواصه فى مناسبات أخرى .

ولو رسمنا المنحنى الذى معادلته مثلا .

$$ص = هـ^{-(س-٣)^٢}$$

نجد أنه منحن متماثل أيضاً ولكن محور تماثله لا ينطبق على المحور الرأسى ، بل يوازيه ويبعد عنه بمقدار ثلاث وحدات من وحدات س .

المنحنى التكرارى غير المتماثل

٧٢ — المطلوب رسم المنحنى الذى معادلته هى

$$ص = س^٢ . هـ^{-س}$$

إبتداء من س = . إلى س = ∞ ، مع العلم بأن هـ لها نفس المعنى المذكور فى البند السابق .

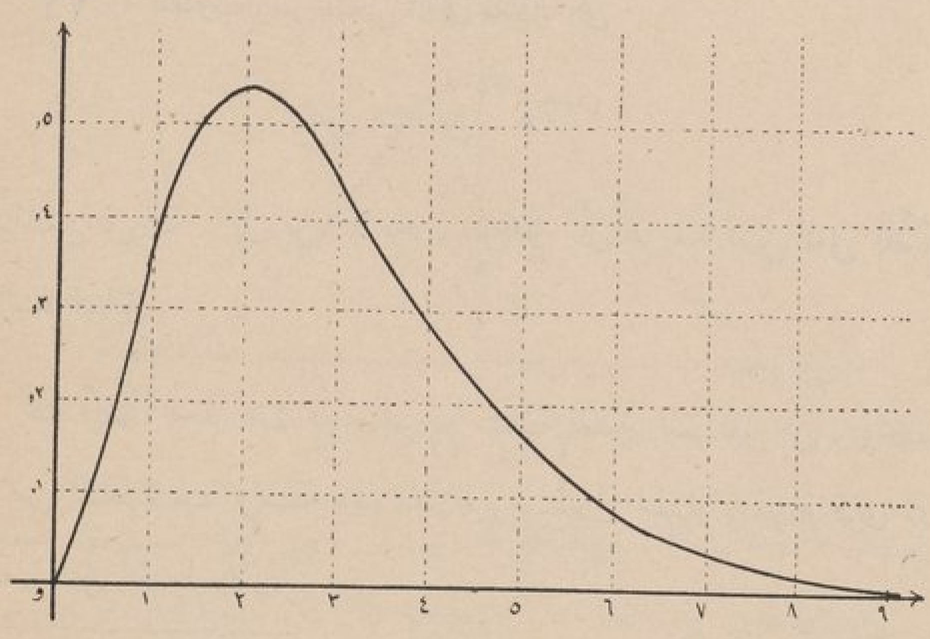
هنا أيضاً نأخذ عدة قيم مناسبة متتالية للمتغير المستقل س ، ونحسب قيم ص المناظرة لها بواسطة هذه المعادلة واستعمال الجداول لمعرفة قوى هـ ، فنحصل على الجدول الآتى :

---

(١) بالإنجليزية Gaussian Curve ، Symmetrical Curve of Error ، Normal Frequency Curve.

س	ص	س	ص
٠		١ر٥	ر٥٠٢٠
ر١	ر٠٠٩٠	٢ر٠	ر٥٤١٢
ر٢	ر٠٣٢٧	٢ر٥	ر٥١٣١
ر٣	ر٠٦٦٧	٣ر٠	ر٤٤٨٢
ر٤	ر١٠٧٢	٣ر٥	ر٣٧٠٠
ر٥	ر١٤١٦	٤ر٠	ر٢٩٢٨
ر٦	ر١٩٧٦	٤ر٥	ر٢٢٤٨
ر٧	ر٢٤٣٣	٥ر٠	ر١٦٧٥
ر٨	ر٢٨٧٦	٦ر٠	ر٠٩٠٠
ر٩	ر٣٢٩٣	٨ر٠	ر٠١٩٢
١ر٠	ر٣٦٧٩	١٠ر٠	ر٠٠٤٥

ومن واقع هذا الجدول نرسم المنحنى كالمعتاد فنحصل على (الشكل ٢٥)



(شكل ٢٥)

منحنى تكرارى غير متماثل

ويلاحظ من الشكل أن هذا المنحنى غير متماثل ، وأنه فى صعوده أسرع منه فى هبوطه . وهو يمس محور السينات من الطرف الأيمن .

وهذا المنحنى أيضاً من المنحنيات المعروفة فى علم الإحصاء ؛ وله أهمية كبيرة هو الآخر من الناحية النظرية .

توفيق المنحنيات

إيجاد معادلة رياضية لمنحن مرسوم من واقع بيانات تجريبية

٧٣ — تكلمنا فى هذا الباب عن كيفية رسم المنحنيات البيانية بمعرفة القيم المتناظرة للمتغيرين اللذين نريد دراسة العلاقة بينهما . وهذه القيم المتناظرة حصلنا عليها إما من التجربة والمشاهدة ، وإما عن طريق علاقة رمزية معروفة بين المتغيرين . وفى الجزء الأخير كنا نرسم الخطوط البيانية بعد معرفة العلاقة بين المتغيرين ؛ والآن تبقى مسألة لم نعالجها بعد ، ألا وهى :

« إذا عرفنا القيم المتناظرة للمتغيرين بطريق التجربة والمشاهدة ورسمنا المنحنى البيانى من واقعها ، فهل يمكننا أن نتوصل إلى معرفة العلاقة الرياضية بين المتغيرين ؟ وكيف يكون ذلك ؟ » . أو بعبارة أخرى :

« إذا عرفنا من نتائج التجارب والمشاهدات عدداً من أزواج القيم المتناظرة لمتغيرين ، فما هى أحسن صورة رياضية نصوِّر بها العلاقة بين هذين المتغيرين حتى تكون موافقة أحسن ما يمكن للبيانات التى حصلنا عليها من التجربة ؟ »

والبحث فى هذه المسألة يتلخص فى البحث عن أحسن معادلة رياضية تربط هذين المتغيرين بحيث توافق نتائج التجربة أحسن ما يمكن — أو تتعارض معها أقل ما يمكن . وأقترح التعبير عن ذلك بالكلمة **توفيق المنحنيات**[١] .

---

(١) الكلمة الإنجليزية هى (Curve Fitting)

يوجد طريقتان لتوفيق المنحنيات وهما[١] : طريقة المربعات الصغرى وطريقة العزوم . والطريقة الأولى هى الأسهل من الناحيـة العملية والنظرية ؛ وسنكتفى هنا بشرح هـذه الطريقة . ويجب على القارىء أن يرجع إلى بعض الكتب الرياضية لمعرفة الطريقة الثانية[٢] .

٧٤ — لنفرض أن القيم التى حصلنـا عليها من التجربة للمتغيرين س و ص هى :

$$س_١ ، س_٢ ، س_٣ ، \cdots\cdots ، س_ن$$

$$و \quad ص_١ ، ص_٢ ، ص_٣ ، \cdots\cdots ، ص_ن$$

نرسم محورين متعامدين ونرصد فى الشكل النقط التى إحداثياتها :

$$(س_١ ، ص_١) ، (س_٢ ، ص_٢) ، \cdots\cdots\cdots ، (س_ن ، ص_ن)$$

بالنظر إلى مواقع هـذه النقط فى الشكل يمكننـا أن نرى إذا ما كان يوافقها خط مستقيم أو خط منحنٍ من الدرجة الثانيـة أو الثالثة أو الرابعـة الخ ، وذلك بحسب ما نعرفه من خواص وأشكال هذه المنحنيات ذات الدرجة الثانية أو الثالثة أو الرابعة أو . . التى لاحظناها فى هذا الباب (بند ٦٨ و بنـد ٦٩) . فإذا اتفقنا على أن هذه النقط البيانية يوافقها خط مستقيم ، نشرع فى البحث عن معادلته بحيث يكون ، كما قلنا ، موافقاً أحسن ما يمكن لهذه النقط .

لنفرض أن معادلة[٣] هذا المستقيم على العموم هى :

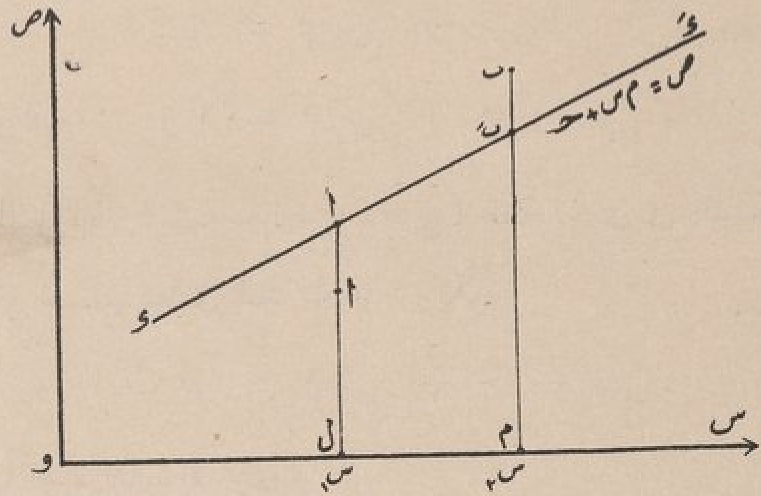
$$ص = م س + ح \quad \cdots \qquad (١) ،$$

(١) The Method of Least Squares, and The Method of Moments

(٢) انظر كتاب H. Rietz "*Handbook of Mathematical Statistics*" (1924) p. 68.

(٣) هـذه المعادلة التى فرضنا ، بناء على رأينا وخبرتنا ، أنها توافق البيـانات التى حصلنا عليها بالتجربة والمشاهدة ، نسميها « المعادلة الاعتبارية » Empirical Equation

حيث م كمية ثابتة مستقلة عن س و ص وكذلك ح ؛ وهما ، كما نعلم ، ميل المستقيم على المحور الأفقى ، وطول الجزء الذى يقطعه من المحور الرأسى . وطبعاً يتعين المستقيم تعييناً تاماً متى علمنا قيمتى م و ح . ونحن نبحث الآن عن هاتين القيمتين اللتين تجعلان المستقيم يوافق البيانات المشاهدة بالتجربة أحسن ما يمكن .



(شكل ٢٦)

لنفرض أن النقطة ا ($س_1$ ، $ص_1$) مثلا هى إحدى هذه النقط فى الشكل ، وأن المستقيم ء ءَ هو المستقيم الذى معادلته ص = م س + ح ؛ ولنفرض أن هذه النقطة وقعت عفواً ، لسبب طرأ فى أثناء التجارب التى أجريت ، بعيدة نوعاً عن هذا المستقيم . نرسم الخط الرأسى ل ا اَ ماراً بالنقطة ا يقابل المستقيم فى النقطة اَ . فيكون البعد ا اَ هو مقدار « انحراف » هذه النقطة ا عن الوضع الذى كان يجب أن تكون فيه لو أنها تمشت تماماً مع هذا القانون الاعتبارى الذى تمثله المعادلة (١) . ويصح أن نسميه أيضاً « الخطأ التجريبى » .

نرى من الشكل أن الانحراف

ا اَ = ل اَ — ل ا

= ل اَ — $ص_1$

حيث ل اَ هو الإحداثى الرأسى للنقطة اَ الواقعة على الخط . والإحداثى الأفقى لهذه النقطة اَ يساوى نفس الإحداثى الأفقى للنقطة ا أى يساوى $س_١$. و بما أن النقطة اَ واقعة على المستقيم الذى معادلته ص = م س + ح ، فلا بد أن إحداثياها الرأسى والأفقى يحققان هذه المعادلة ؛ و بذلك يكون :

$$اَ = م\,س_١ + ح$$

$$\therefore \text{الانحراف} \quad ا\,اَ = م\,س_١ + ح - ص_١ \text{، ...... (٢) .}$$

و بالمثل إذا كانت ب ($س_٢$ ، $ص_٢$) نقطة أخرى فى الشكل ، فإن انحرافها عن المستقيم يساوى البعد ب بَ ؛ ويكون

$$ب\,بَ = م\,س_٢ + ح - ص_٢ .$$

وهكذا مع جميع النقط الأخرى .

وطبيعى أن الخط ص = م س + ح يكون أوفق ما يمكن لتمثيل هذه النقط كلما كانت هذه الأخطاء أو الانحرافات صغيرة فى المقدار ، بصرف النظر عن كونها موجبة أو سالبة . و بعبارة أخرى نقول إن هذا المستقيم يكون أفضل ما يمكن إذا كان مجموع مربعات هذه الأخطاء أصغر ما يمكن[١] ؛ أى أن المقدار

$$(م\,س_١ + ح - ص_١)^٢ + (م\,س_٢ + ح - ص_٢)^٢ + \cdots$$
$$+ (م\,س_ن + ح - ص_ن)^٢$$
$$= \text{مج}\,(م\,س + ح - ص)^٢$$

يكون أصغر ما يمكن .

---

(١) أول من وضع هذه النظرية العالم الفرنسى لجندر Legendre سنة ١٨٠٦ ؛ ومن بعده العالمان لابلاس Laplace وجاوس Gauss وغيرهما . ولإثباتها انظر كتاب :
Whittaker and Robinson, *"Calculus of Observations"*, 1929 p. 209.

ولكى يكون المقـدار $\text{مجـ}\,(\text{م}\,\text{س}+\text{ح}-\text{ص})^{2}$ أصـغر ما يمكن ، يجب أن يكون مجموع الانحرافات نفسها يساوى صـــفراً ؛ وفى الوقت نفسه يجب أن يكون مجموع حواصل ضرب هذه الانحرافات ، كل فى قيمة الإحداثى الأفقى للنقطة ، يساوى صفراً أيضاً [١] ؛ أى أن :

$$(\text{م}\,\text{س}_{1}+\text{ح}-\text{ص}_{1})+(\text{م}\,\text{س}_{2}+\text{ح}-\text{ص}_{2})+\cdots+(\text{م}\,\text{س}_{\text{ن}}+\text{ح}-\text{ص}_{\text{ن}})=0$$

و

$$\text{س}_{1}(\text{م}\,\text{س}_{1}+\text{ح}-\text{ص}_{1})+\text{س}_{2}(\text{م}\,\text{س}_{2}+\text{ح}-\text{ص}_{2})+\cdots+\text{س}_{\text{ن}}(\text{م}\,\text{س}_{\text{ن}}+\text{ح}-\text{ص}_{\text{ن}})=0$$

أو بالاختصار :

$$\text{م}\,.\,\text{مجـ}\,\text{س}+\text{ن}\,\text{ح}=\text{مجـ}\,\text{ص}\quad\cdots\quad(3)\,،$$

$$\text{م}\,.\,\text{مجـ}\,\text{س}^{2}+\text{ح}\,\text{مجـ}\,\text{س}=\text{مجـ}\,\text{س}\,\text{ص}\quad\cdots\quad(4)\,.$$

وبما أن جميع قيم س وقيم ص معلومة ، وهى القيم المشاهدة فى التجربة ، وكذلك ن ، يكون المجهول فى هاتين المعادلتين الآتيتين (٣) و (٤) هما قيمتا م و ح فقط . فنحل المعادلتين بالطرق الجبرية المعروفة ونستخرج قيمتى م و ح ونعوضهما فى المعادلة (١) نحصل على معادلة المستقيم المطلوب .

٧٥ — لنأخذ مثلا القيم الآتية لكل من س و ص ، باعتبار س تمثل عمر الطفل بالسنة و ص تمثل وزنه بالكيلو جرام . والمطلوب توفيق خط مستقيم ليمثل العلاقة بين السن والوزن عند الأطفال [٢] .

مثال توفيق مستقيم

---

(١) يمكننا إثبات ذلك بسهولة باستخدام نظرية النهايات الكبرى والصغرى فى حساب التفاضل والتكامل . نبحث عن النهاية الصغرى للكمية $\text{مجـ}\,(\text{م}\,\text{س}+\text{ح}-\text{ص})^{2}$ باعتبارها دالة المتغيرين م و ح . نفاضلها بالنسبة إلى ح وحدها ونضع النتيجة تساوى صفراً ، فنحصل على المعادلة (٣) . ثم نفاضل الكمية بالنسبة إلى م وحدها ، ونضع النتيجة تساوى صفراً فنحصل على المعادلة (٤)

(٢) هذه الأرقام مأخوذة من بعض الإحصاءات الطبية عن الأطفال (أولاد) فى ألمانيا — ١٩٣٧

س :	٥	٦	٧	٨	٩	١٠	١١	١٢	١٣	١٤	١٥
ص :	٢٢ر٥	٢٤ر٥	٢٥ر٧٥	٢٧ر٢٥	٢٩ر٥	٣١ر٢٥	٣٣ر٢٥	٣٥ر٧٥	٣٨ر٧٥	٤١ر٧٥	٤٦ر٧٥

نفرض أن معادلة المستقيم المطلوب هى ص = م س + حـ ، ونوجد قيمتى م و حـ بحل المعادلتين الآتيتين .

مح ص = م . مح س + ن حـ

مح س ص = م . مح س^{2} + حـ . مح س

ولتسهيل العمل فى إيجاد قيم مح س ، مح ص ، مح س^{2} ، مح س ص ، نرتب العمل فى جدول كالآتى :

	س	ص	س^{2}	س . ص
	٥	٢٢ر٥٠	٢٥	١١٢ر٥٠
	٦	٢٤ر٥٠	٣٦	١٤٧ر٠٠
	٧	٢٥ر٧٥	٤٩	١٨٠ر٢٥
	٨	٢٧ر٢٥	٦٤	٢١٨ر٠٠
	٩	٢٩ر٥٠	٨١	٢٦٥ر٥٠
	١٠	٣١ر٢٥	١٠٠	٣١٢ر٥٠
	١١	٣٣ر٢٥	١٢١	٣٦٥ر٧٥
	١٢	٣٥ر٧٥	١٤٤	٤٢٩ر٠٠
	١٣	٣٨ر٧٥	١٦٩	٥٠٣ر٧٥
	١٤	٤١ر٧٥	١٩٦	٥٨٤ر٥٠
	١٥	٤٦ر٧٥	٢٢٥	٧٠١ر٢٥
المجموع	١١٠	٣٥٧ر٠٠	١٢١٠	٣٨٢٠ر٠٠

وبالتعويض فى المعادلتين الآتيتين ، مع العلم بأن ن = ١١ ،

∴ ٣٥٧ = ١١٠ م + ١١ حـ .... (١) ؛

و ٣٨٢٠ = ١٢١٠ م + ١١٠ حـ .... (٢) .

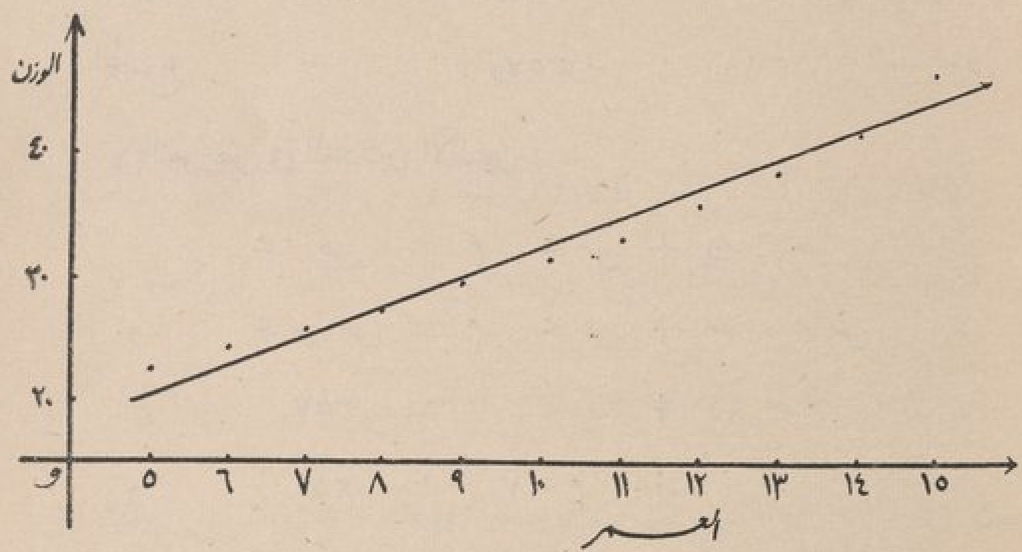
∴ م = $٢\frac{٣}{١١}$ = ٢٫٢٧٣ ؛

و حـ = $\frac{١٠٧}{١١}$ = ٩٫٧٢٧ .

∴ معادلة المستقيم المطلوب هى :

ص = ٢٫٢٧٣ س + ٩٫٢٢٧ ... ... (٣) ،

باعتبار أن ص هى وزن الطفل بالكيلو جرام و س عمره بالسنين ، ونراه مرسوماً فى (شكل ٢٧) .



(شكل ٢٧)

٧٦ — يمكن تسهيل العمل الحسابى فى هذا المثال لو أننا أخذنا قيم س مقيسة ابتداء من القيمة الوسطى فى السلسلة المعطاة (أى العمر ١٠) بدلا من قياسها ابتداء من الصفر . وحينئذ نعبر عن الأعمار هكذا : طريقة لاختصار العمل الحسابى

العمر	س	ص	س^{٢}	س . ص
٥	— ٥	٢٢ر٥٠	٢٥	— ١١٢ر٥٠
٦	— ٤	٢٤ر٥٠	١٦	— ٩٨ر٠٠
٧	— ٣	٢٥ر٧٥	٩	— ٧٧ر٢٥
٨	— ٢	٢٧ر٢٥	٤	— ٥٤.٥٠
٩	— ١	٢٩ر٥٠	١	— ٢٩ر٥٠
١٠	٠	٣١ر٢٥	٠	٠
١١	١	٣٣ر٢٥	١	٣٣ر٢٥
١٢	٢	٣٥ر٧٥	٤	٧١ر٥٠
١٣	٣	٣٨ر٧٥	٩	١١٦ر٢٥
١٤	٤	٤١ر٧٥	١٦	١٦٧ر٠٠
١٥	٥	٤٦ر٧٥	٢٥	٢٣٣ر٧٥
المجموع	٠	٣٥٧ر٠٠	١١٠	٥٠ر٠٠

وبالتعويض فى المعادلتين الآتيتين :

مح ص = م . مح س + ن . ح ،

و مح س ص = م . مح س^{٢} + ح . مح س ،

∴ ٣٥٧ = ٠ + ١١ ح ،

و ٢٥٠ = ١١٠ م + ٠ .

∴ م = ٢ر٢٧٣ ،

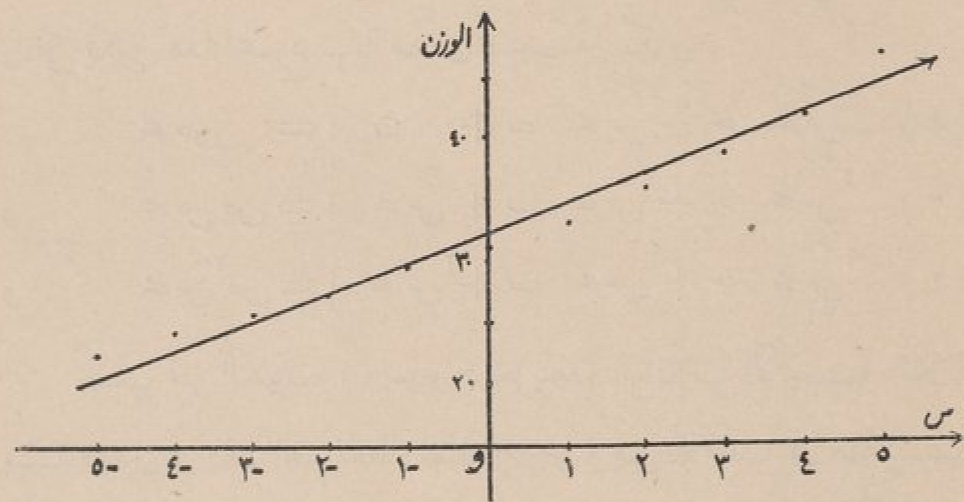
و ح = ٣٢ر٤٥٥ .

∴ معادلة المستقيم فى هذه الحالة هى :

ص . = ٢ر٢٧٣ س + ٣٢ر٤٥٥ ،

حيث ص هى وزن الطفل بالكيلو جرام ، و س تدل على الفرق ، مقدراً

بالسنين ، بين العمر الحقيقي والعمر ١٠ سنوات . ونرى المستقيم مرسوماً في (شكل ٢٨) ؛ ويلاحظ أن ميل المستقيم واحد في هذين الشكلين . والفرق بين الشكلين هو أن المحور الرأسي نقل من مكانه فقط ، فتغير طول الجزء الذي يقطعه المستقيم من هذا المحور [١] .



(شكل ٢٨)

توفيق منحن من الدرجة الثانية

٧٧ — إذا رأينا أن القيم المشاهدة للمتغيرين س ، ص يوافقها خط منحن من الدرجة الثانية ، نوفق لها منحنياً معادلته من الدرجة الثانية على الصورة

$$\text{ص} = \text{ا} + \text{ب}\,\text{س} + \text{ح}\,\text{س}^{٢} \quad \ldots \quad (١)،$$

حيث ا ، ب ، ح ثلاث كميات ثابتة مجهولة ، نبحث عن قيمها التي تجعل هذا المنحني يوافق القيم المشاهدة أحسن ما يمكن . وهذه القيم هي :

$$(\text{س}_{١}\text{، ص}_{١})\text{، }(\text{س}_{٢}\text{، ص}_{٢})\text{، }\ldots\text{، }(\text{س}_{ن}\text{، ص}_{ن}) .$$

---

(١) إذا كان عدد قيم س المعطاة في المسألة زوجياً ، وكانت القيم على مسافات متساوية من بعضها ، نأخذ نقطة الأصل في منتصف الفترة بين قيمتي س المتوسطتين في السلسلة .

بمثل البرهان الذى اتبعناه فى بند ٧٤ بالنسبة للخط المستقيم ، نثبت هنا أن مجموع مربعات الأخطاء أو الانحرافات عن هذا المنحنى ، هو :

$$(\text{ا} + \text{ب}\,\text{س}_1 + \text{حـ}\,\text{س}_1^2 - \text{ص}_1)^2 + (\text{ا} + \text{ب}\,\text{س}_2 + \text{حـ}\,\text{س}_2^2 - \text{ص}_2)^2 + \dots = \text{مجـ}(\text{ا} + \text{ب}\,\text{س} + \text{حـ}\,\text{س}^2 - \text{ص})^2 \text{ ؛}$$

ولكى يكون هذا المجموع نهاية صغرى ، يجب أن يكون :

$$\text{مجـ}\,\text{ص} = \text{ا}\,.\,\text{ن} + \text{ب}\,.\,\text{مجـ}\,\text{س} + \text{حـ}\,.\,\text{مجـ}\,\text{س}^2 \quad ..(٢)\text{،}$$

$$\text{و}\quad \text{مجـ}\,\text{س}\,\text{ص} = \text{ا}\,.\,\text{مجـ}\,\text{س} + \text{ب}\,.\,\text{مجـ}\,\text{س}^2 + \text{حـ}\,.\,\text{مجـ}\,\text{س}^3 \quad ..(٣)\text{،}$$

$$\text{و}\quad \text{مجـ}\,\text{س}^2\text{ص} = \text{ا}\,.\,\text{مجـ}\,\text{س}^2 + \text{ب}\,.\,\text{مجـ}\,\text{س}^3 + \text{حـ}\,.\,\text{مجـ}\,\text{س}^4 \quad ..(٤)\,.$$

نحسب قيم الكميات ا ، ب ، حـ بحل هذه المعادلات الآتيـــة الثلاثة ، مستخدمين فى ذلك القيم المعطاة للمتغيرين س و ص كما فعلنا فى حالة المستقيم . ولكن العمل الحسابى هنا يكون أطول بالطبع ، ولذلك يستحسن دائماً أن نتبع طريقة مختصرة مثل التى شرحناها فى بند ٧٦ ، وذلك بأن نختار مبدأ قياس السينات من الوسط (إذا كانت قيمها المعطاة متدرجة بفترات منتظمة) .

مثال لتوفيق منحن من الدرجة الثانية

٧٨ — لنفرض أن القيم المشاهدة للمتغيرين س و ص هى :

س :	٣	٥	٧	٩	١١	١٣	١٥	١٧	١٩
ص :	١٢	١٦	١٩	٢١	٢٣	٢٠	١٨	١٥	١٠

ولتكن المعادلة المطلوبة هى :

$$\text{ص} = \text{ا} + \text{ب}\,\text{س} + \text{حـ}\,\text{س}^2 \qquad (١)$$

لتسهيل العمل نأخذ مبدأ قياس القيم السينية عند ١١ حتى يكون هناك قيم موجبة وأخرى سالبة نحوها . ونرتب العمل كما فى الجدول الآتى :

س	سَ	ص	سَ^{٢}	سَ^{٣}	سَ^{٤}	سَ ص	سَ^{٢} ص
٣	—٨	١٢	٦٤	—٥١٢	٤٠٩٦	—٩٦	٧٦٨
٥	—٦	١٦	٣٦	—٢١٦	١٢٩٦	—٩٦	٥٧٦
٧	—٤	١٩	١٦	—٦٤	٢٥٦	—٧٦	٣٠٤
٩	—٢	٢١	٤	—٨	١٦	—٤٢	٨٤
١١	٠	٢٣	٠	٠	٠	٠	٠
١٣	٢	٢٠	٤	٨	١٦	٤٠	٨٠
١٥	٤	١٨	١٦	٦٤	٢٥٦	٧٢	٢٨٨
١٧	٦	١٥	٣٦	٢١٦	١٢٩٦	٩	٥٤٠
١٩	٨	١٠	٦٤	٥١٢	٤٠٩٦	٨٠	٦٤٠
المجموع	٠	١٥٤	٢٤٠	٠	١١٣٢٨	—٢٨	٣٢٨٠

وبالتعويض فى المعادلات

مجـ ص = ا . ن + ب . مجـ سَ + حـ مجـ سَ^{٢} ،

و مجـ سَ ص = ا . مجـ سَ + ب مجـ سَ^{٢} + حـ . مجـ سَ^{٣} ،

و مجـ سَ^{٢} ص = ا . مجـ سَ^{٢} + ب مجـ سَ^{٣} + حـ . مجـ سَ^{٤} ،

∴ ١٥٤ = ٩ ا + ٠ + ٢٤٠ حـ (٢) ،

و —٢٨ = ٠ + ٢٤٠ ب + ٠ (٣) ،

و ٣٢٨٠ = ٢٤٠ ا + ٠ + ١١٣٢٨ حـ (٤) .

∴ ب = —٠٫١١٦٧ .

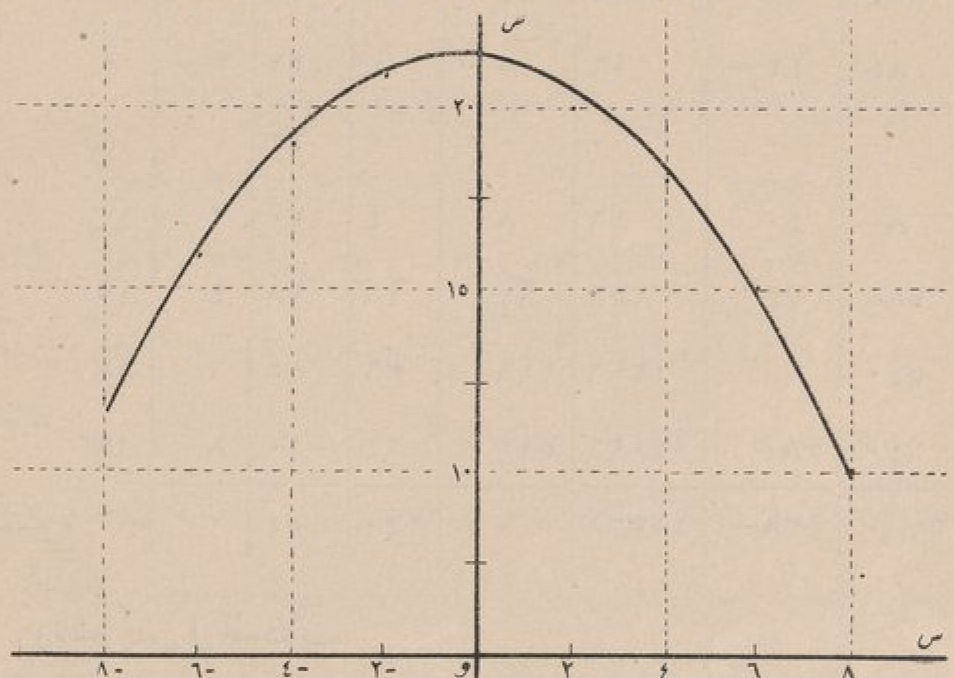
ومن المعادلتين (٢) و (٤)

∴ ا = ٢١٫٥٨٣ ، حـ = —٠٫١٦٧٧ .

وتكون معادلة المنحنى المطلوب هى :

ص = ٢١ر٥٨٣ — ١١٦٧ر س + — ١٦٧٧ر س^{٢}

وفى (شكل ٢٩) نجد هذا المنحنى مرسوماً ونرى النقط الأصلية المعطاة إحداثياتها واقعة على مقربة من المنحنى ، و بعضها فوق المنحنى و بعضها تحته .



(شكل ٢٩)

توفيق معادلة من الدرجة الثالثة أو أعلى

٧٩ — وبالمثل إذا أردنا توفيق منحن من الدرجة الثالثة ، نفرض المعادلة

ص = ا + ب س + حـ س^{٢} + ء س^{٣}

ونوجد قيم ا ، ب ، حـ ، ء بحل الأربع المعادلات الآتية :

مجـ ص = ا . ن + ب مجـ س + حـ . مجـ س^{٢} + ء . مجـ س^{٣} ،

و مجـ س ص = ا . مجـ س + ب مجـ س^{٢} + حـ . مجـ س^{٣} + ء . مجـ س^{٤} ،

و مجـ س^{٢} ص = ا . مجـ س^{٢} + ب مجـ س^{٣} + حـ . مجـ س^{٤} + ء . مجـ س^{٥} ،

و مجـ س^{٣} ص = ا . مجـ س^{٣} + ب مجـ س^{٤} + حـ . مجـ س^{٥} + ء . مجـ س^{٦} .

ولكن العمل بهذه الطريقة فى توفيق المنحنيات ذات الدرجات العليا

يستلزم حسابات مرهقة ، مما يضعف من فائدتها فى بعض الحالات . ولكن الطرق الأخرى ليست أسهل من هذه بكثير .

المراجع

أبو زهرة و باخوم : **التفاضل والتكامل** (١٩٣٥) .

سليم أمين حداد : **مسائل فى علم الاحصاء** الرسم البيانى .

محمد على حجاب وآخرون : **كتاب دروس الرياضة** (١٩٣٨) ، الأبواب ١٠ — ١٢ .

نصيف سعيد وصادق بشاره : **الهندسة التحليلية** (١٩٣٥) الباب الثانى .

هول ونايت : **الجبر الابتدائى** — الجزء الثانى

Bowley, A.L., : *Elements of Statistics*, Chapter VII.
Karsten, : *Charts and Graphs.*
Mills, F.C. : *Statistical Methods*, Chapter II.
Rietz, H. : *Handbook of Mathematical Statistics*, Chapter IV.
Whittaker and Robinson : *Calculus of Observations*, Chapter IX.

الباب الخامس

التوزيع التكرارى

كل ظاهرة متغيرة تأخذ عدة قيم إما مختلفة أو متقاربة

٨٠ — إذا نحن شاهدنا ظاهرة متغيرة فى ظروفها الزمانية أو المكانية المختلفة ، وأخذنا فى كل حالة وقعت تحت ملاحظتنا بياناً رقمياً بمقدار هذه الظاهرة المتغيرة فى تلك الحالة ، نحصل بالطبيعة على عدة قيم لهذا المتغير الذى نبحثه . هذه القيم يصح أن تكون كلها مختلفة بعضها عن بعض ، ويجوز أن يكون بعضها متساوياً أو قريباً من التساوى .

فلو تتبعنا مثلاً أسعار القطن فى أثناء موسم معين فى بورصة مينا البصل ، وجمعنا أسعار الإقفال فى أيام الموسم لحصلنا على عدة مقادير مختلفة للسعر : بعضها أكبر أو أصغر من الأسعار الباقية ، و بعضها متوسط بين الطرفين .

وإذا كنا ندرس درجة ذكاء تلاميذ فرقة معينة ، واختبرنا لذلك ذكاء كل تلميذ فى هذه الفرقة نحصل أيضاً على مقاييس مختلفة لذكاء هؤلاء التلاميذ . وكذلك إذا قسنا طول القامة لكل فرد من مجموعة رجال فى سن واحدة ، نحصل على مقادير مختلفة للأطوال .

ترتيب القيم تصاعدياً أو تنازلياً

٨١ — ولكى ندرس مثل هذه الظاهرة المتغيرة يصح أن نرتب مجموعة القيم التى نحصل عليها بالمشاهدة أو التجربة ترتيباً تصاعدياً أو تنازلياً تمهيداً لهذه الدراسة . ولكن هذا الترتيب لن يساعدنا كثيراً ، خصوصاً إذا كان عدد القيم كبيراً ، حيث يتعذر على الإنسان أن يحصر ذهنه فى مجموعة كبيرة من الأرقام

المختلفة ، ولا يمكنه أن يستوعبها ليستنبط منها خواص المجموعة واتجاهاتها .

تقسيم مجموعة القيم إلى فئات

٨٢ — ولتسهيل هذه العملية يحسن أن نقسم المجموعة الأصلية إلى مجموعات جزئية تشمل كل واحدة عدداً من القيم المتقاربة من بعضها ، التى يمكننا اعتبارها « متساوية » تقريباً بعد إهمال الفروق البسيطة التى بينها .

هذه المجموعات الجزئية نسمى كلا منها فئة . وكل فئة ، إذن ، تشمل جميع المفردات أو القيم التى تقع بين حدين مناسبين نعينهما حسب ما يتراءى لنا ، وهذا يتوقف طبعاً على ظروف المسألة واعتبارات أخرى سيأتى شرحها . ولكننا على العموم سنعتبر كل القيم داخل الفئة الواحدة كأنها جميعاً متساوية ، ونصرف النظر عن الفروق البسيطة التى تكون بين هذه المفردات فى الواقع .

لنفرض مثلا أننا نبحث فى أطوال مجموعة من الأشخاص ؛ وبعد ترتيب هذه الأطوال ترتيباً تصاعدياً وجدنا أن أقصر شخص فيها طوله ١٤١ سم ، وأن أطول شخص طوله ١٩٧ سم .

ولنفرض أننا قسمنا هذه المجموعة إلى فئات : الأولى تشمل كل من طولهم ١٤٠ سم وأقل من ١٥٠ سم ؛ وتشمل الثانية كل من طولهم ١٥٠ سم وأقل من ١٦٠ سم ؛ وهكذا إلى الفئة الأخيرة (السادسة) وتشمل كل الأطوال الواقعة بين ١٩٠ سم وأقل من ٢٠٠ سم .

توزيع المفردات على الفئات

٨٣ — بعد الاتفاق على عدد الفئات التى تنقسم إليها المجموعة الأصلية وتعيين الحدين الأدنى والأعلى لكل فئة ، نوزع المفردات على هذه الفئات ونضع كل مفردة فى الفئة المناسبة لها . ثم نعد المفردات الموجودة فى كل فئة ، ونضع هذا العدد أمام كل فئة فى جدول كالآتى :

جدول ٣ — توزيع أطوال مجموعة من الرجال

فئات الأطوال بالسنتيمتر	مراكز الفئات بالسنتيمتر	عدد الأشخاص فى كل فئة
١٤٠ وأقل من ١٥٠	١٤٥	٨
١٥٠ « ١٦٠	١٥٥	١٩
١٦٠ « ١٧٠	١٦٥	٣١
١٧٠ « ١٨٠	١٧٥	٣٧
١٨٠ « ١٩٠	١٨٥	٢٥
١٩٠ « ٢٠٠	١٩٥	١١
	المجموع	١٣١

نفرض أن كل مفردة فى الفئة تساوى مركزها

٨٤ — بعد توزيع المفردات على الفئات بهذا الشكل تضيع معالم القيم الأصلية التى وضعت المفردات فى الفئات بمقتضاها ؛ وحينئذ لا نعرف شيئاً عن أى مفردة سوى أنها واحدة فى فئة معينة محدودة بحدين معينين هما ١٦٠ سم و ١٧٠ سم مثلا . وهذه المفردة تساوى جميع المفردات الأخرى التى معها فى نفس الفئة — وعددها جميعاً ٣١ فى هذا المثال . ولا يمكننا من واقع هذا الجدول أن نحدد طول أى شخص موجود فى أى فئة مثل الفئة ١٦٠ — ١٧٠ . ولا نعلم إذن ما إذا كان طوله قريباً من ١٦٠ سم أو هو أقرب إلى ١٧٠ سم . ولذلك نفرض عملياً أن جميع الأشخاص فى هذه الفئة طولهم = ١٦٥ سم ، أى يساوى مركز الفئة وهو منتصف المدى بين الحدين الأعلى والأدنى . ومع أن هذا الفرض ربما يحتاج إلى تبرير ، إذ أن الحكم فى هذه المسألة لا بد يتوقف على كيفية توزيع المفردات عموماً ، ولكنه بالرغم من ذلك فرض مناسب وعملى ولا يبعد كثيراً عن الصواب فى أغلب الأحيان . أو قل إن الخطأ الناشىء عن هذا الفرض لا يؤثر كثيراً فى النتيجة التى نحصل عليها ، خصوصاً عندما يكون مدى الفئة قصيراً .

الجـدول التكرارى . التـوزيع التكرارى .

٨٥ — وعلى ذلك نجد فى هذه الفئة ٣١ شخصاً طولهم جميعاً ١٦٥ سم . أى أن الطول ١٦٥ سم « تكرر » ٣١ مرة فى هذه المجموعة . والجدول حينئذ يعطينا الأطوال المختلفة — مراكز الفئات — وعدد مرات « تكرار » كل واحد منها فى هذه المجموعة . أو ، بالاختصار ، يعطى الأطوال وتكراراتها . ولذلك نسمى هـذا الجدول **الجدول التكرارى**[١] وهو يبين ما نسميه **التوزيع التكرارى**[٢] لأطوال هذه المجموعة . وهذه الفئات التى تنقسم إليها المجموعة نسميها **فئات تكرارية**[٣] .

ويلاحظ أن هذا الجدول التكرارى يعطينا صورة واضحة ومختصرة لكيفية توزيع قيم الظاهرة التى نبحثها . فمثلا نرى أن المجموعة تتركز فى الفئتين ١٦٠ وأقل من ١٧٠ و ١٧٠ وأقل من ١٨٠ ، وعلى الأخص فى الأخيرة . ونرى أيضاً أن القيم المتطرفة قليلة التكرار نسبياً ، وعلى ذلك ينتظر أن يكون الطول النموذجى لهذه المجموعة واقعاً فى إحدى الفئتين المذكورتين .

كيفية تعيين عدد الفئات ومواقعها

٨٦ — هناك عدة اعتبارات يجب أن نراعيها عند تعيين الفئات التى نقسم إليها المجموعة لكى نحصل على جدول تكرارى مناسب . ومن هذه الاعتبارات ما يختص بعدد الفئات وطول الفترة ومنها ما يختص بتحديد مواقع الفئات أى تعيين حدودها الدنيا والعليا .

عدد الفئات يناسب عدد المفردات وطول المدى

وعند تحديد عدد الفئات يجب أن ننظر إلى طول المدى بين أصغر وأكبر قيمة فى المجموعة . وهذا المدى يُقسَّم إلى عدد مناسب من الأقسام طولها أو مداها معقول .

ويجب أيضاً أن ننظر إلى عدد المفردات كلها التى فى المجموعة ، ونراعى

(١) Frequency Table . (٢) Frequency Distribution . (٣) Frequency Groups .

أن هـذا العدد يكفى للتوزيع على الفئات بحيث تنـال كل واحدة عدداً معقولا من المفردات ، وإلا أخذنا عدداً أصغر من الفئات .

طـول مدى الفئـة يسهل العمل ولـكنه ليس دقيقاً

ويلاحظ أنه كلما كثر عدد الفئات كان العمل الحسابى اللازم أصعب وأطول . وبالعكس إذا صغرنا عدد الفئات كانت العمليات الحسابية أقصر وأسهل ، ولـكن على حساب الدقة فى النتيجة . وذلك لأنه كلما صغر عدد الفئات كان مدى الفئـة طويلا . وينشأ عن هذا عدم تحقق الغرض بأن مركز الفئـة يمثل جميع المفردات التى فيها (انظر بند ٨٤) ، إذ أن بعض القيم فى الفئة الطويلة المدى تكون بعيدة عن مركزها ، ونكون مخطئين كثيراً إذا اعتبرناها منطبقـة عليه . فيجب إذن عند تحديد عدد الفئات التى ينقسم إليها المدى أن نأخذ فى الاعتبـار درجة الدقة التى نبتغيها ، إذ أن عدد الفئات يعين طول الفترة فى كل فئة ، وبالتالى يعين مقدار كسور الفترة التى نهملها حينما نفرض أن كل القيم متجمعة عند مركز الفئة .

المغـالاة فى تضييق الفئات تسبب عـدم انتظـام التكرارات

٨٧ — ولـكن يجب ألا نغـالى فى تضييق فترات الفئات طلباً للدقة . لأن هذا ينشأ عنـه زيادة عدد الفئـات أكثر من اللزوم ، فلا يكفى عدد المفردات للتوزيع على هذه الفئات الـكثيرة الناتجة . ونحصل حينئذ على توزيع غير منتظم للتـكرارات ونرى هذا واضحاً من الجدول الآتى ، وهو يبين الدرجات التى حصل عليها عدد من التلاميذ (١٥٣٤) فى امتحان معين[١] .

(١) درجات مادة الحسـاب فى امتحـان شهادة الدراسة الثـانوية (قسم ثان) فى يونيه سنة ١٩٣٤ . يلاحظ كثرة العدد عند الدرجة ٣ ، وهى درجة النجاح فى هذا العلم . ويظهر بوضوح أن كثيراً من هؤلاء كانوا راسبين ثم « جبروا » .

جدول ٤ — جدول تكرارى لتوزيع درجات تلاميذ فى امتحان معين :

الدرجة	عدد التلاميذ الحاصلين عليها	الدرجة	عدد التلاميذ الحاصلين عليها	الدرجة	عدد التلاميذ الحاصلين عليها
صفر	٧	٥ر٥	٧٨	١١ر٠	٢٦
ر٥	٩	٦ر٠	٧٣	١١ر٥	٢٣
١ر٠	١٨	٦ر٥	٨٧	١٢ر٠	٣٤
١ر٥	٣٣	٧ر٠	٧٦	١٢ر٥	١٩
٢ر٠	٦١	٧ر٥	٨٩	١٣ر٠	١٢
٢ر٥	٢٣	٨ر٠	٨٦	١٣ر٥	١٢
٣ر٠	١١٦	٨ر٥	٧٧	١٤ر٠	٧
٣ر٥	٨٥	٩ر٠	٦٧	١٤ر٥	٨
٤ر٠	٨١	٩ر٥	٤٩	١٥ر٠	٨
٤ر٥	١٠٥	١٠ر٠	٤٧		
٥ر٠	٨٠	١٠ر٥	٣٨	المجموع	١٥٣٤

ويلاحظ أن التكرارات فى هذا الجدول متذبذبة بدون انتظام . وذلك لكثرة عدد الفئات (٣١ فئة) وقلة عدد المفردات الموزعة عليها نسبياً . ولو جعلنا مدى الفئة درجة كاملة بدل نصف درجة ، يقل عدد الفئات إلى النصف ، ونحصل على توزيع أكثر انتظاماً للتكرارات ، وهو كما يأتى :

جدول ٥ — التوزيع التكرارى للدرجات فى جدول ٤
فى فئات مدى فترتها درجة كاملة

الدرجات	عدد التلاميذ الحاصلين عليها	الدرجات	عدد التلاميذ الحاصلين عليها
صفر وأقل من ١	١٦	٨ وأقل من ٩	١٦٣
١ « ٢	٥١	٩ « ١٠	١١٦
٢ « ٣	٨٤	١٠ « ١١	٨٥
٣ « ٤	٢٠١	١١ « ١٢	٤٩
٤ « ٥	١٨٦	١٢ « ١٣	٥٣
٥ « ٦	١٥٨	١٣ « ١٤	٢٤
٦ « ٧	١٦٠	١٤ « ١٥	١٥
٧ « ٨	١٦٥	١٥ ٠	٨
		الجملة	١٥٣٤

ويلاحظ أن التكرارات فى هذا الجدول تتغير بانتظام أكثر منها فى الجدول السابق ، ولو أن هنا بعض التذبذب أيضاً . ولكنه ليس شديداً (١) . وهذه التذبذبات تتلاشى إذا جعلنا مدى الفئة درجتين بدل درجة واحدة . ونبين ذلك فى الجدول الآتى :

(١) أنظر شكل ٣٢ صفحة ١٠٣

جدول ٦ — التوزيع التكرارى السابق

فى فئات مدى فتراتها درجتان

الدرجات	عدد التلاميذ الحاصلين عليها	الدرجات	عدد التلاميذ الحاصلين عليها
صفر وأقل من ٢	٦٧	٨ وأقل من ١٠	٢٧٩
٢ « ٤	٢٨٥	١٠ « ١٢	١٣٤
٤ « ٦	٣٤٤	١٢ « ١٤	٧٧
٦ « ٨	٣٢٥	من ١٤ إلى ١٥	٢٣
		الجملة	١٥٣٤

ولا نجد فى هذه التكرارات تذبذباً بل نجدها تزداد بالتدريج حتى تصل إلى نهاية كبرى ثم تقل بعد ذلك ، بالتدريج أيضاً .

تعيين حدود الفئات

٨٨ — أما بخصوص تحديد مبادىء الفئات ونهاياتها ، فهذا يتوقف على نوع البيانات التى لدينا ودرجة تفصيلها وطريقة جمعها .

ويستحسن على العموم أن نجعل طول الفترة أو المدى متساوياً فى كل الفئات لأن هذا يسهل العمل الحسابى والرسم . ولكن فى بعض المسائل تكون البيانات مفصلة فى جزء ومجملة فى جزء آخر من المجموعة . وفى مثل هذه الحالات لا يمكن عمل فئات متساوية ، كما نجد فى حالة البيانات الخاصة بالملكية العقارية . وفيها نجد الملكيات الصغيرة مفصلة أكثر من الملكيات الكبيرة ، كما نرى فى الجدول الآتى وهو يبين التوزيع التكرارى للملكيات فى سنة ١٩٣٦[١] بالقطر المصرى

(١) انظر الإحصاء السنوى العام ١٩٣٥ — ١٩٣٦ ص ٣٣٠

جدول ٧ — توزيع الملكية العقارية فى مصر فى سنة ١٩٣٦

فئات الملكية بالفدان	عدد الملاك (التكرار)
أكثر من ٠ إلى ١	١٦٧٧٥٣٦
« ١ « ٥	٥٦٤٧٠٠
« ٥ « ١٠	٨٤٦١٧
« ١٠ « ٢٠	٣٩٦٤٣
« ٢٠ « ٣٠	١٢٤٢٥
« ٣٠ « ٥٠	٩٣٧٤
« ٥٠ فداناً	١٢٤٢٠
الجملة	٢٤٠٠٧١٥

الفئات غير المتساوية فى طول المدى

ويلاحظ أن هذه الفئات غير متساوية المدى . وذلك لأن الملكيات الصغرى كثيرة العدد ، ويمكن تقسيمها إلى عدة فئات دون أن يتضاءل نصيب كل فئة من عدد المفردات أو التكرارات . كما أنها بطبيعتها وطبيعة أصحابها تستوجب الدرس بالتفصيل . بخلاف الملكيات الكبيرة فهى قليلة العدد لا تتحمل التجزئة إلى فئات كثيرة قصيرة المدى . هذا فضلاً عن أنه لو اتبعنا نفس التقسيم فى كل الجدول لاحتجنا إلى عمل جدول من خمسين فئة أو أكثر ، وهى عملية متعبة جداً وبدون مبرر ؛ ويجوز أن يكون نصيب بعض هذه الفئات المفصلة ضئيلا جداً أو منعدماً .

ويلاحظ أيضاً فى هذا الجدول أن الفئة الأخيرة ليس لها حد أعلى . وهذا هو ما نعبر[١] عنه بقولنا إن الجدول **مفتوح من أعلى**، وفى بعض الأحيان يكون الجدول مفتوح الطرفين . وعند ذلك تكون الفئة الأولى فى الجدول معروفاً

(١) بالانجليزية Open-end Table

حدها الأعلى فقط ، كأن نقول عن الفئة الأولى فى الجدول المذكور فى بند ٨٣ إنها « أقل من ١٥٠ » ولا نذكر ١٤٠ .

فئات خاصة للتفصيل فى بعض المناطق

٨٩ — فى بعض المسائل يطلب بيانات تفصيلية عن بعض أجزاء من المدى الذى تتغير فيه الظاهرة . مثلاً فى دراسة الحالة الصحية لبلد ما نعنى بدراسة أعمار الأشخاص عند الوفاة ، وعلى الخصوص أعمار الأطفال دون السنة من العمر . وفى هذه الحالات نعمل فئة خاصة ، فى الجدول التكرارى لأعمار المتوفين للأطفال ، ونجعل مبدأها العمر صفر ونهايتها سنة واحدة . ونتبعها بفئة أخرى تنتهى عند سنتين . و بعد هذه نعمل فئات أطول مدى حيث التفصيل غير مطلوب .

نختار مبادىء الفئات بحيث تخفف أثر الخطأ فى جمع البيانات

٩٠ — يراعى أيضاً عند تحديد مبادىء الفئات الظروف التى جمعت فيها البيانات الأصلية ، ومقدار دقة هذه البيانات وما فيها من أخطاء أو تقريب ، حتى يمكن اختيار نظام للفئات يلافى هذه الأخطاء أو يخفف من أثرها فى النتائج .

فمثلا عند جمع بيانات عن الأعمار فى تعدادات السكان ، نجد كثيراً من الناس يذكرون أعمارهم لأقرب عشر سنين ، ولا يهتمون — أو لا يرغبون — بذكر الأعمار بدقة ، فتجد أحدهم يذكر أن عمره ٣٠ سنة ، فى حين أن عمره الحقيقى ٢٨ أو ٣٣ سنة مثلاً .

فللتخفيف من أثر هذا الخطأ فى البيانات ، يمكننا جعل مدى فئات الأعمار حوالى هذه الأرقام الشائعة : ٢٠ ، ٣٠ ، ٤٠ ، ... فنأخذ مبادىء الفئات ١٥ — ، ٢٥ — ، ٣٥ وهكذا . وبذلك تشمل هذه الفئات جميع الأشخاص الذين ينتمون إليها حقيقة . بخلاف الحالة إذا جعلنا مبادىء الفئات ٢٠ — ، ٣٠ — ، ... ؛ فربما يدخل فى الفئة ٣٠ — أناس كثيرون عمرهم الحقيقى ٢٧ أو ٢٨ أو ٢٩ إلخ ؛ وينتج عن هذا خطأ كبير فى العمل .

ونجـد مثالا آخر لهـذه الأخطاء فى كيفية توزيع الدرجات فى الجداول المذكورة فى بند ٨٧ . حيث لاحظنا فى الجدول ٥ تضخما كبيراً فى تكرار الدرجة ٣. وهذا بدون شك على حساب الفئتين المجاورتين لهذه — وعلى الأخص فئة الدرجة ٢٫٥ . والسبب فى ذلك واضح كما ذكرنا فى الحاشـية فى صفحة ٨٨ . وهو أن كثيراً من التلاميذ الحاصلين فى الأصل على درجتين ونصف فقط يعطون نصف درجة للنجاح ، فيصبحوا عندنا فى فئة الدرجة ٣ . ويجوز أن يدخل هـذه الفئة مفردات كانت فى الأصل فى فئة الدرجة ٢ ونقلت إلى ٣ لنفس السبب . وبذلك يتضخم تكرار الفئة ٣ كما نرى فى الجدول المذكور . (صفحة ٨٩) .

ويمكن تلافى أثر هـذا الخطأ العملى بأن نختار حدود الفئـات بحيث تكون الدرجة ٣ واقعـة داخل فئة وليست فى طرفها . وأحسن نظام هو كالمقترح فى جدول ٧ حيث الفئة الثانية تمتد من ٢ إلى ٣٫٥ . وهذا يضمن أن كل الحالات التى كانت فى الأصل ٢ أو ٢٫٥ أو ٣ أو ٣٫٥ توضع فى المكان الصحيح ، حتى ولو كانت وضعت خطأ لسبب من الأسباب .

٩١ — وطريقة التعبير عن مبادىء الفئات ونهاياتها سهلة ؛ ويحسن توحيد الطريقة لمنع الالتباس . وسنستعمل فى هذا الكتاب الصور الآتية لتدل على المعانى المبينة : التعبير عن مبادىء ونهايات الفئات

(١) — ٥ وأقل من ١٠ ، تكتب ٥ — للاختصار

١٠ « « ١٥ « ١٠ — «

١٥ « « ٢٠ « ١٥ — «

ومعنى هـذا أن الفئـة مداها ٥ وحدات ، والفئة الأولى تشمل كل القيم التى تساوى ٥ أو تزيد عنها بحيث تقل عن ١٠ . وعلى ذلك فالقيمة ١٠ تدخل فى الفئة

التى تليها ، وهكذا . ويكون مركز هذه الفئة المتوسط بين حديها ، أى ٧ر٥ للأولى و ١٢ر٥ للثانية و ١٧ر٥ للثالثة ، وهكذا .

(٢) — أكثر من ٥ إلى ١٠ ، تكتب — ١٠
» ١٠ » ١٥ » — ١٥
» ١٥ » ٢٠ » — ٢٠

ومعنى هـذا أن الفئة مداها ٥ وحدات أيضاً ، وأن حدها الأعلى ١٠ للفئة الأولى و ١٥ للثانية و ٢٠ للثالثة ، وهكذا . ومراكزها نصف فترة قبل هذه الحدود أى ٧ر٥ و ١٢ر٥ و ١٧ر٥ على الترتيب . وتشمل الفئة الأولى جميع القيم التى أكثر من ٥ حتى ١٠ . أما القيم التى أكبر من ١٠ فتدخل فى الفئة التالية لها ، وهكذا .

(٣) — أقل من ٥ تكتب < ٥
أكثر من ٥٠ » > ٥٠

والفئة الأولى < ٥ تدل على فئـة مفتوحة من أسفل ، ومحدودة من أعلى ؛ ولكن مداها غير معروف . وهى تشـمل جميع القيم التى أقل من ٥ ، ولا تشمل ٥ نفسها . وهذه عادة تكون فى مبدأ الجدول التكرارى ، وعندئذ يسمى جدولاً مفتوحاً من أسفل .

والفئـة « أكثر من ٥٠ أى > ٥٠ » فئـة مفتوحة من أعلى ، ومحدودة من أسفل ، ومداها غير معروف أيضاً . وهى تشمل جميع القيم التى أكبر من ٥٠ ، ولا تشمل ٥٠ نفسها . وهذه عادة تكون فى نهاية الجدول التكرارى وعندئذ يسمى الجدول مفتوحاً من أعلى .

ويصح أن يكون الجدول مفتوح الطرفين إذا ابتـدأ وانتهى بمثلى هاتين الفئتين .

(٤) — الصورة : ١٠
٢٠
٣٠
٤٠

تدل على مراكز فئات ، كل فئة مداها عشر وحدات . وكل فئة تبدأ خمس وحدات (نصف المدى) قبل المركز وتنتهى بعد المركز بخمس وحدات أيضاً . فالفئة الأولى هى ٥ — والثانية ١٥ — والثالثة ٢٥ — وهكذا .

(٥) — الوضع : ٤ — ٨
٨ — ١٢
١٢ — ١٦

بدون ذكر أى شىء يعين مبادىء الفئات أو نهاياتها ، وضع مبهم ويجب تحديده أو استبداله بإحدى الصور المحددة السابقة . إذ لايمكن هنا تقرير الفئة التى توضع فيها القيمة ١٢ مثلا : هل توضع فى الفئة الثانية أو الثالثة من هذه الفئات .

(٦) — الوضع : ٥ — ٩
١٠—١٤
١٥—١٩

معناه أن الفئة مداها من ٥ إلى ٩ وكسورها ، والثانية مداها من ١٠ إلى ١٤ وكسورها . فهو إذن صورة أخرى للتعبير المذكور فى (١) وهو يستعمل كثيراً فى الجداول التكرارية للأعمار ونحوها .

التغير المتصل

٩٢ — إذا كانت الظاهرة التى نشاهدها ونسجل قيمها المختلفة ، تمهيداً لدراستها بواسطة الجدول التكرارى ، تتغير تغيراً « متصلا » — مثال ذلك أطوال الأشخاص أو أعمارهم — فإن قيمها تنتشر فى المدى بين القيمتين الصغرى والكبرى فى المجموعة بالتدريج بدون أن تتجمع فى بعض النقط دون الأخرى ؛ لأن التغير المتصل معناه أن الكمية المتغيرة لا « تقفز » من قيمة إلى أخرى أكبر منها ، ولكنها « تنمو » بالتدريج بحيث تمر بكل القيم المتوسطة . فمثلا إذا كانت درجة الحرارة فى الصباح ٢٢° وفى الظهر ٢٨٫٥° ، فمعنى ذلك أنها كانت فى وقت من الأوقات بين الصبح والظهر ٢٥° مثلا ، وفى وقت آخر كانت ٢٧٫٤° مثلا ، وهكذا . بدليل أن عمود الزئبق فى الترمومتر الذى يقيس درجة الحرارة لا يمكن أن يقفز فجأة من ٢٢° إلى ٢٨٫٥° ، بل هو يتمدد شيئاً فشيئاً ماراً بكل نقطة على تدريج الترمومتر من ٢٢° إلى ٢٨٫٥° .

السلسلة المتصلة

ومجموعة القيم التى نحصل عليها لظاهرة مثل هذه نسميها **سلسلة متصلة** (١) . ومثل هذه المجموعة توضع فى جدول تكرارى « متصل » أى أن الحد الأعلى لكل فئة يلاصق الحد الأدنى للفئة التى تليها مباشرة ؛ وكل قيمة بالقرب من هذين الحدين المتلاصقين تكون تابعة لفئة واحدة منهما فقط .

ومثال ذلك الجدول التكرارى للأعمار . إذ أن عمر الشخص يزيد بالتدريج بحيث إننا نجد أشخاصاً أعمارهم تساوى كل الأعمار التى بين ٢٥ و ٣٠ سنة مثلا . وكذلك التوزيع التكرارى للأطوال أو الأوزان وهكذا . وفى مثل هذه التوزيعات لا نجد صعوبة عملية أو نظرية فى فرض قيم كل فئة تساوى مركز الفئة .

السلسلة المنفصلة

٩٣ — وبعكس ذلك نجد بعض الظواهر تتغير **تغيراً منفصلا** ، أى أن مقدار الظاهرة « يقفز » من قيمة إلى أخرى فجأة بدون أن يتدرج فى القيم

(١) بالإنجليزية Continuous Series

الواقعة بينهما . مثال ذلك : عدد ما عند الرجل من أطفال ، وعدد ما فى المصنع من عمال ، وعدد ما تحمله شجرة القطن من اللوز ، وهكذا . فبينما نجد أناساً كثيرين عند الواحد منهم ٣ أطفال ، وكثيرين غيرهم عند الواحد منهم ٤ أطفال ، لا نجد أحداً عنده ٣ر٧ من الأطفال مثلا . ونجد كثيراً من شجرات القطن تحمل كل واحدة ١٨ لوزة ، وكثيراً غيرها تحمل كل واحدة ١٩ لوزة مثلا ، ولا نجد شجرة واحدة تحمل ١٨و٣ أو ١٨و٦ من اللوزات .

ومثل هذه المجموعة نسميها سلسلة منفصلة[1] . والجدول التكرارى لها يكون منفصلا . أى أن الفئات المتتالية تكون منفصلة وغير متلاصقة . فنجد مثلا أن التوزيع التكرارى لعدد العمال فى مصانع القاهرة (حسب التعداد الصناعى لسنة ١٩٢٧) كما يأتى :

جدول ٨ — توزيع تكرارى لعدد العمال فى مصانع القاهرة سنة ١٩٢٧

فئات عدد العمال فى كل مصنع	التكرار أى عدد المصانع التى بها هذا العدد
٠	٥٥٩١
من ١ إلى ٤	٨٨٥٢
« ٥ « ٩	١٤٤٨
١٠ عمال فأكثر	٩٩٤
الجملة	١٦٨٨٥

ويلاحظ أن الفئة الأولى تشمل كل المصانع التى ليس بها مستخدمون ؛ وتشمل الفئة الثانية جميع المصانع التى بها عامل أو اثنان أو ثلاثة أو أربعة .

---

(١) بالانجليزية Discrete or Discontinuous Series

كما يلاحظ أيضاً أن الحد الأعلى لكل فئة منفصل عن الحد الأدنى للتى تليها ، وذلك لعدم وجود مصانع فيها عدد العمال واقع بين هذه الحدود : ٤ و٥ أو ٩ و ١٠ مثلا .

ويجب أن نفرق بين هـذه الحالة والحالة المذكورة فى (٦) من بنـد ٩٠ . والفرق هو أن المتغير هنا منفصل والفئة تمتـد من ٥ إلى ٩ فقط على حين أن المتغير فى الوضع (٦) المذكور متصل والفئة تمتد من ٥ إلى ٩ وكسورها ؛ مثلا : ٩ سنين و٣ أشهر أو ١٠ أو ١١ شهراً ، أو أى شىء أقل من ١٠ سنوات مهما كان .

توضيح الجداول التكرارية بالرسم

٩٤ — يمكن توضيح الجدول التكرارى بتمثيله على شكل هندسى . ونشرح الآن بعض الطرق المستعملة لهذا الغرض .

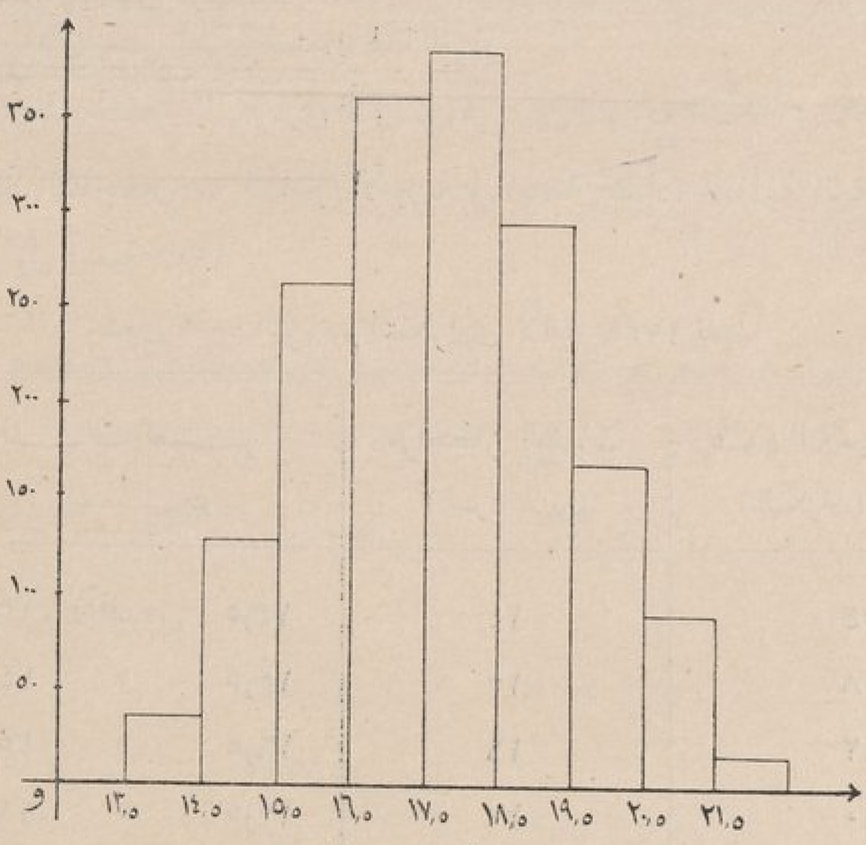
لنأخذ الجدول التكرارى الآتى ، وهو يبين التوزيع التكرارى لأعمار التلاميذ الناجحين من المدارس الأميرية فى امتحان شهادة الدراسة الثانوية (قسم أول) سنة ١٩٣٣

جدول ٩ — التوزيع التكرارى لأعمار ١٧٣٩ تلميذاً

فئات العمر بالسنين	مراكز الفئات (العمر التقريبى)	عدد التلاميذ (التكرار)
١٣٫٥ وأقل من ١٤٫٥	١٤	٣٤
١٤٫٥ « ١٥٫٥	١٥	١٢٨
١٥٫٥ « ١٦٫٥	١٦	٢٦٢
١٦٫٥ « ١٧٫٥	١٧	٣٦٠
١٧٫٥ « ١٨٫٥	١٨	٣٨٦
١٨٫٥ « ١٩٫٥	١٩	٢٩٤
١٩٫٥ « ٢٠٫٥	٢٠	١٦٧
٢٠٫٥ « ٢١٫٥	٢١	٩٢
٢١٫٥ « ٢٢٫٥	٢٢	١٦
		المجموع ١٧٣٩

مدرج التكرار (الهستوجرام)

٩٥ — نرسم محورين متعامدين ، ونأخذ المحور الأفقى لقياس الأعمار بالسنين ، والمحور الرأسى لقياس التكرارات . ثم نقسم المحور الأفقى إلى ٩ أقسام متساوية تمثل الفئات التسع للأعمار ، ونكتب مبادىء الفئات أمام هذه التقاسيم . ثم نأخذ على المحور الرأسى مقياس رسم يناسب أرقام التكرارات التى عندنا فى الجدول .



(شكل ٣٠)

مدرج تكرارى أو هيستوجرام لأعمار ١٧٣٩ تلميذاً

نرسم أمام كل فئة مستطيلا رأسياً تتناسب مساحته مع رقم التكرار الخاص بالفئة ، بحيث تمتد قاعدة المستطيل على المحور الأفقى من أول الفئة إلى آخرها ، كما هو مبين بالحدين الأدنى والأعلى المذكورين فى الجدول .

فإذا كانت الفئات متساوية المدى ، كما هو الحال فى هـذا المثال ، فلا بد أن تكون قواعد المستطيلات متساوية ؛ وحينئذ تكون النسب بين ارتفاعها تساوى النسب بين التكرارات المذكورة فى الجدول .

وإذا لم تكن الفئات متساوية المدى فتكون مساحات هـذه المستطيلات (القاعدة × الارتفاع) هى التى تتناسب مع أرقام التكرار . ويكون طول قاعدة المستطيل يساوى طول فترة الفئة التى يمثلها ، وارتفاعه يساوى خارج قسمة تكرار هذه الفئة على طول فترتها .

وفى كلتا الحالتين نحصل على شكل مدرج يشبه تدريج السلم ، ويسمى[1] **هيستوجرام أو مدرج التكرار** . وهو يمثل التوزيع التكرارى الموجود بالجدول فى شكل هندسى .

الفئات المفتوحة لا يرسم لها مستطيلات

٩٦ — وإذا كان الجدول « مفتوحاً » من أحد الطرفين أو من كليهما ، فلا يمكن رسم مستطيل يمثل تكرار الفئة المفتوحة . لأن الفئة المفتوحة لا يعرف مداها بالضبط حيث لا نعرف إلا حدًّا واحداً من حدّيها . ولذلك لا يمكن معرفة طول القاعدة التى ينشأ عليها المستطيل المطلوب ؛ وبالتالى لا نعرف ارتفاعه أيضاً .

وعلى ذلك نهمل الفئات المفتوحة — سواء فى أول الجدول أو فى آخره — ولا نرسم لها مستطيلات فى الهيستوجرام . وإلا فيتعين علينا أن نفرض لها حدوداً تقريبية — على ضوء ما نعرفه عن الجدول والتكرارات من خبرتنا الخاصة ، إذا كان لدينا أى معلومات مفيدة — وبواسطتها نعرف طول القاعدة ونشىء عليها المستطيل . ولكن الأسلم فى العادة أن نترك هذه الفئات المفتوحة ولا نتعرض لها .

مساحات المستطيلات تمثل التكرارات

ويلاحظ أن مجموع مساحات المستطيلات فى الهيستوجرام العادى يمثل مجموع التكرارات كلها . لأن كل مستطيل فى الشكل ، مساحته تمثل تكرار فئة من الفئات ؛ ومجموع هذه المساحات إذن يمثل مجموع التكرارات — ما عدا الفئات

(١) بالانجليزية Histogram

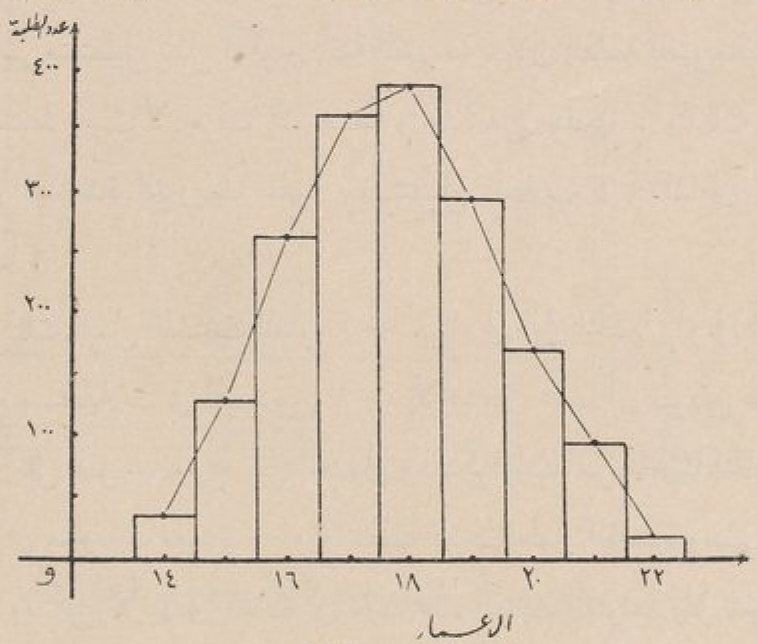
المفتوحة إذا وجدت ولم يرسم لها مستطيلات ؛ وحينئذ يجب استبعاد تكرارات هذه الفئات من حسابنا .

المضلع التكرارى

٩٧ — توجد طريقة أخرى لتوضيح التوزيع التكرارى نستخدمها فى حالة تساوى جميع الفئات فى المدى ، وهى رسم **المضلع التكرارى** [١] .

لذلك نأخذ مراكز الفئات كإحداثيات أفقيـة ، والتـكرارات المناظرة لها كإحداثيات رأسية . ثم نرصد فى الشكل نقطاً بهذه الإحداثيات ، ونصل هذه النقط بخط منكسر فنحصل على ما نسميه المضلع التكرارى .

وفى هذه الحالة أيضاً نهمل الفئات المفتوحة إذا وجدت ، وذلك لعدم معرفة مراكزها ، وعدم معرفة الإحداثيات الأفقية للنقط التى تمثل تكرارات هذه الفئات فى المضلع — إلا إذا لجأنا إلى فرض حدود لهذه الفئات لتعيين مراكزها .



(شكل ٣١)

المدرج التكرارى ومعه المضلع التكرارى لأعمار ١٧٣٩ تلميذاً

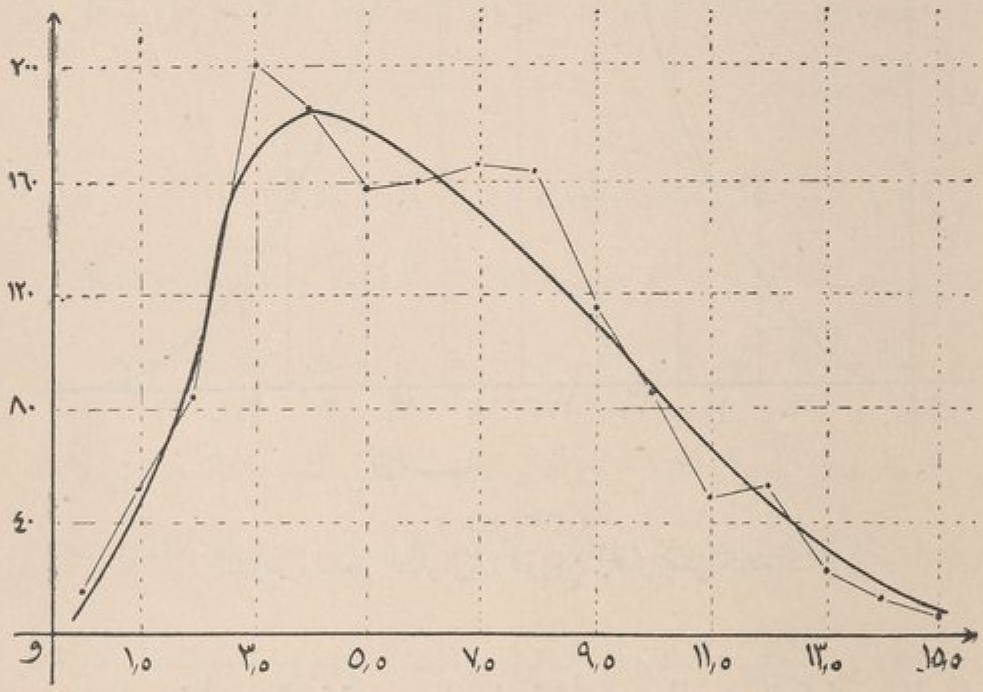
ويلاحظ أن المضلع التكرارى يمكن رسمه من واقع المدرج التكرارى . فهذه النقط التى نعينها فى المضلع التكرارى هى فى الحقيقة منتصفات القواعد العليا

(١) بالانجليزية Frequency Polygon

للمستطيلات فى المدرج التكرارى . فإذا وصلنا بين هذه المنتصفات بخط منكسر حصلنا على نفس المضلع التكرارى السابق .

المضلع والمدرج متساويان فى المساحة

٩٨ — ويجب أن نلاحظ هنا أن المساحة المحدودة بالمضلع التكرارى تساوى المساحة المحدودة بمستطيلات المدرج التكرارى . فإذا تأملنا فى شكل ٣١ نجد أن كل ضلع فى المضلع يقطع جزءاً مثلثاً من المستطيل ويضيف مثلثاً آخر . والجزء المقطوع يساوى الجزء المضاف تماماً . وعلى ذلك تكون مساحة المضلع التكرارى ممثلة بالضبط مجموع التكرارات كما قلنا عن مساحة المدرج التكرارى .



(شكل ٣٢)

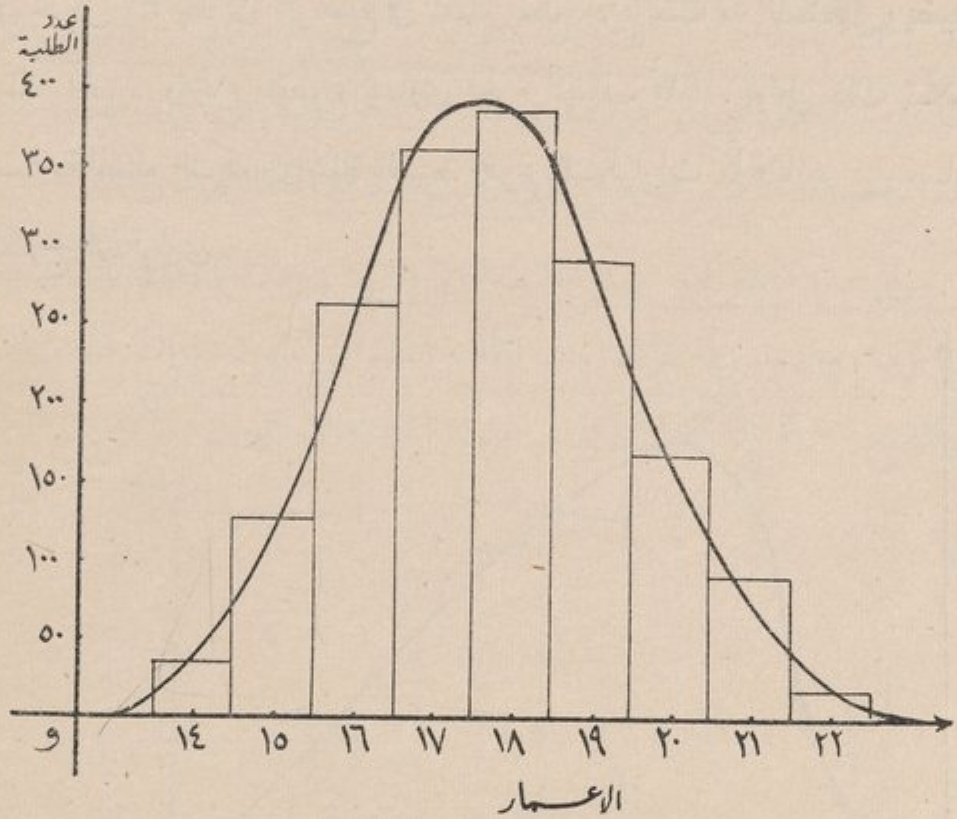
المضلع التكرارى والمنحنى التكرارى لتوزيع درجات بعض التلاميذ

المنحنى التكرارى

٩٩ — إذا نحن رسمنا المضلع التكرارى ثم مهدنا انكساراته نحصل على خط منحن خال من الذبذبات الفجائية ، ونسمى[1] هذا الخط **المنحنى التكرارى**

(١) بالانجليزية Frequency Curve

وهو يستعمل كطريقة أخرى لتمثيل التوزيعات التكرارية فى شكل هندسى واضح . وهى فى الحقيقة أحسن الطرق المستعملة لهذا الغرض بشرط أن يكون الرسم دقيقاً . ونرى فى شكل ٣٢ المضلع التكرارى والخط الممهد أو المنحنى التكرارى لتوزيع الدرجات المذكور فى جدول ٥ صفحة ٩٠ ، وقد أخذنا مراكز الفئات هنا عند ½ و ١½ و ٢½ وهكذا .



(شكل ٣٣)

المدرج والمنحنى التكراريين لتوزيع أعمار بعض التلاميذ

مساحة المنحنى لاتساوى مساحة المضلع أوالمدرج تماماً

ويلاحظ أن هناك فرقاً بين المساحة المحدودة بالمنحنى التكرارى والمساحة المحدودة بالمضلع أو المدرج التكرارى ، كما هو واضح من شكل ٣٣ حيث رسمنا المدرج والمنحنى التكرارى لنفس التوزيع الموجود فى جدول ٩ ، ونرى فى الشكل أن المنحنى يقطع أجزاء من أعالى مستطيلات المدرج ، ويضيف أجزاء غيرها . وكذلك نرى فى شكل ٣٢ أن المنحنى يقطع بعض أجزاء المضلع و يضيف غيرها ؛

ولكن يصح أن تكون مساحة الأجزاء المقطوعة أكبر أو أصغر من مساحة الأجزاء المضافة . وينشأ عن ذلك طبعاً اختلاف بين مساحة المنحنى ومساحة المدرج أو المضلع .

المدرج يؤول في النهاية إلى المنحنى التكرارى

١٠٠ — ومن الواضح أنه كلما كان عرض المستطيلات ضيقاً كانت مساحة الأجزاء المقطوعة منها أو المضافة إليها صغيرة . ومن باب أَوْلَى يكون الفرق بين مساحات الأجزاء المقطوعة والمضافة صغيراً أيضاً . وعلى ذلك إذا كانت فئات التكرار قصيرة المدى ، كان الفرق صغيراً بين مساحة المنحنى ومساحة المدرج . وفى النهاية تتساوى المساحتان ، وينطبق المنحنى على المدرج الذى يؤول حينئذ إلى خط مُسنَّن كالمنشار الدقيق ، بعد أن كان خطاً منكسراً على شكل سلم عريض الدرجات .

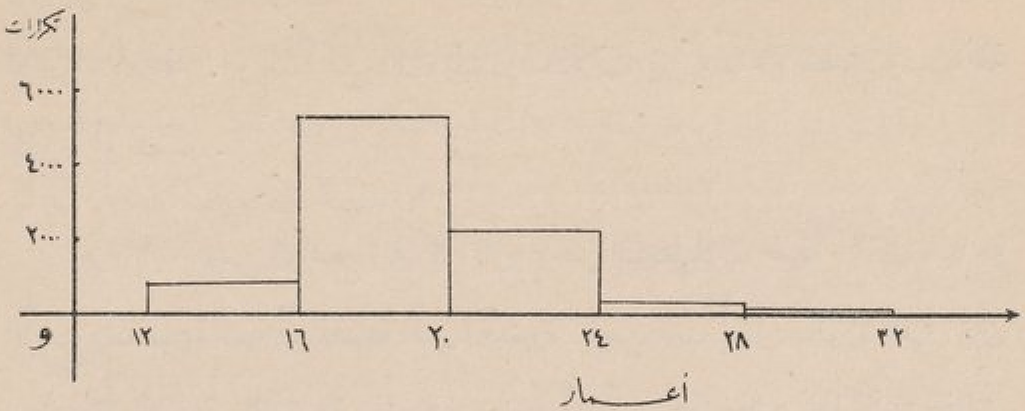
التدرج من الهستوجرام إلى المنحنى التكرارى

١٠١ — ويمكن توضيح ذلك عملياً بأن نأخذ توزيعاً تكرارياً غزيراً ذا فئات واسعة المدى ، ثم نضيق الفئات شيئاً فشيئاً ونرسم المدرج التكرارى فى كل حالة ، فنرى كيف يؤول المدرج العريض إلى منحن ممهد فى النهاية .

لنأخذ مثلا الجدول التكرارى الآتى ، وهو يعطى التوزيع التكرارى لأعمار مجموعة من الأشخاص عددهم ٨٤٤١ ، مقسمين إلى فئات مدى الواحدة ٤ سنين .

ونرى فى شكل ٣٤ المدرج التكرارى لهذا التوزيع . ونلاحظ الارتفاع الفجائى لبعض المستطيلات ثم هبوطها فجأة أيضاً .

ونلاحظ بهذه المناسبة أن شكل المدرج وعدم تناسقه يدل على أن تقسيم الفئات غير مناسب .



(شكل ٣٤)

مدرج تكرارى ، لتوزيع ذى فئات واسعة المدى

جدول ١٠ — توزيع أعمار ٨٤٤١ شخصاً

في فئات مداها ٤ سنين

فئات العمر بالسنين	عدد الأشخاص (التكرار)	فئات العمر بالسنين	عدد الأشخاص (التكرار)
١٢—	٧٩١	٢٨—	٢٧
١٦—	٥١٧٢	٣٢—	٦
٢٠—	٢٢٢٣	٣٦—	٤
٢٤—	٢١٦	٤٠ وأقل من ٤٤	٢
		الجملة	٨٤٤١

14
12 - 16
18
16 - 20

نأخذ الآن نفس الأشخاص ونقسمهم إلى فئات ذات فترات أضيق .

وبالرجوع إلى البيانات الأصلية التى جمعناها عن أعمار هؤلاء الأشخاص نجد أن التوزيع التكرارى لأعمارهم ، مقسمين فى فئات ذات فترات طولها سنتان ، هو كالمبين بالجدول الآتى :

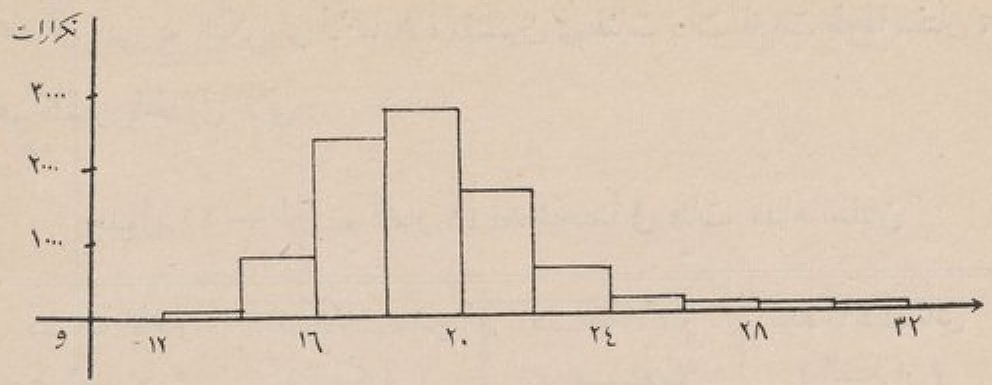
جدول ١١ — توزيع أعمار ٨٤٤١ شخصاً فى فئات مداها سنتان

فئات الأعمار (بالسنين)	عدد الأشخاص (التكرار)	فئات الأعمار (بالسنين)	عدد الأشخاص (التكرار)
١٢ —	٢٧	٢٨ —	٢٠
١٤ —	٧٦٤	٣٠ —	٧
١٦ —	٢٤٠٤	٣٢ —	٤
١٨ —	٢٧٦٨	٣٤ —	٢
٢٠ —	١٦٤٥	٣٦ —	٢
٢٢ —	٥٧٨	٣٨ —	٢
٢٤ —	١٦٦	٤٠ —	١
٢٦ —	٥٠	٤٢ وأقل من ٤٤	١
		الجملة	٨٤٤١

وفى شكل ٣٥ نرى المدرج التكرارى الذى يمثل هذا التوزيع .

ويلاحظ أن الهيستوجرام فى هذه الحالة أكثر تدريجاً وتناسقاً منه فى شكل ٣٤ ، وذلك لأن التكرارات الغزيرة توزعت على فئتين ؛ وقد كانت أولا موجودة فى فئة واحدة .

وأخيراً نأخذ نفس المجموعة مقسمة إلى فئات أضيق ، مداها سنة واحدة

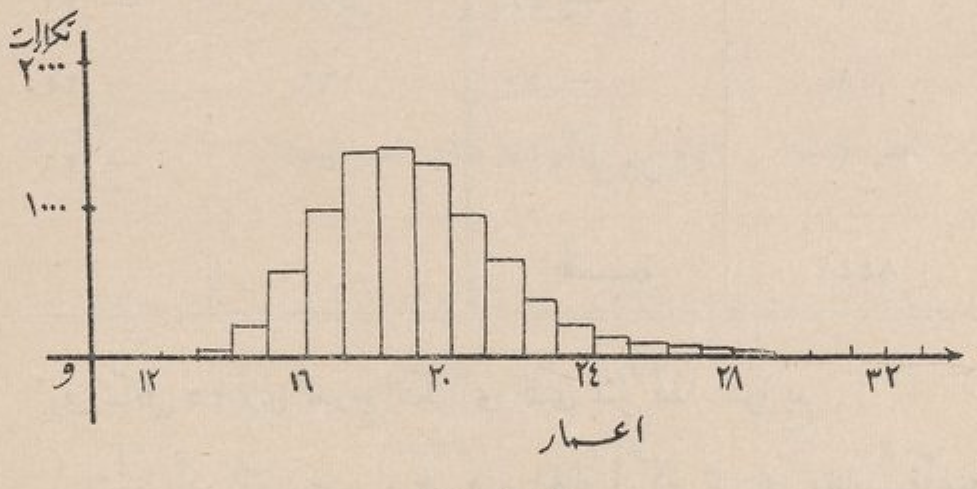


(شكل ٣٥)

مدرج تكرارى للتوزيع ذى فئات مداها سنتان

بدل سنتين . والجدول التكرارى الآتى (رقم ١٢) يبين توزيع أعمار هذه المجموعة (١) فى فئات طول الفترة فيها سنة واحدة .

وفى شكل ٣٦ نرى الهيستوجرام الذى يمثل هذا التوزيع .



(شكل ٣٦)

هيستوجرام ذو فئات طول فترتها سنة واحدة

---

(١) فى سنة ١٩٣٤ كانت أعمار المتقدمين لامتحان شهادة الدراسة الثانوية (قسم أول) موزعة هكذا مع تغيير مبادىء الفئات هنا إلى مراكز فئات .

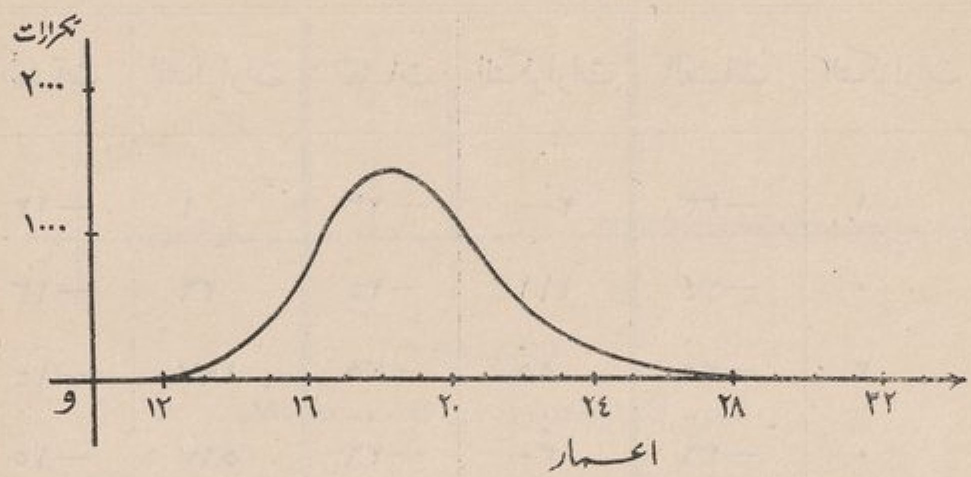
جدول ١٢ — التوزيع التكرارى لأعمار ٨٤٤١ شخصاً
فى فئات طول فترتها سنة واحدة .

الفئات	التكرارات	الفئات	التكرارات	الفئات	التكرارات
١٢—	١	٢٣—	٢٠٠	٣٣—	١
١٣—	٢٦	٢٤—	١١١	٣٤—	٠
١٤—	١٩٧	٢٥—	٥٥	٣٥—	٢
١٥—	٥٦٧	٢٦—	٣٠	٣٦—	٠
١٦—	٩٩٢	٢٧—	٢٠	٣٧—	٢
١٧—	١٤١٢	٢٨—	١٣	٣٨—	١
١٨—	١٤٣٨	٢٩—	٧	٣٩—	١
١٩—	١٣٣٠	٣٠—	٢	٤٠—	١
٢٠—	٩٨٧	٣١—	٥	٤١—	٠
٢١—	٦٥٨	٣٢—	٣	٤٢—	١
٢٢—	٣٧٨			الجملة	٨٤٤١

ونرى أن الهيستوجرام فى هذه الحالة أكثر تمهيداً منه فى الحالتين المتقدمتين ، بسبب تضييق الفترات الناتج عن تقسيم الفئات . وذلك لأن تقسيم فئة إلى اثنتين ينشىء فئة جديدة ذات تكرار متوسط بين الفئتين الأصليتين المتجاورتين ؛ وهذا يخفف من حدة الانتقال من تكرار صغير إلى تكرار كبير فجأة .

ولو جعلنا الفترة نصف سنة لحصلنا على هيستوجرام أكثر تمهيداً من هذا الأخير . وفى النهاية نحصل على منحن ممهد كالذى نراه فى شكل ٣٧ .

ويجب ملاحظة عدم المبالغة فى تقسيم الفئات تقسيما دقيقاً ؛ إذ يخشى حينئذ أن عدد المفردات الموجودة لدينا يصبح قليلا ولا يكفى للتوزيع على الفئات الكثيرة



(شكل ٣٧)

منحن تكرارى لتوزيع أعمار ٨٤٤١ شخصاً

الناتجة ، فنحصل على توزيع غير منتظم كما رأينا فى المثال المذكور فى بند ٨٧ . وإذا أمكن زيادة عدد المفردات كلما زاد عدد الفئات بسبب تجزئتها ، فلا بد أن نحصل فى النهاية على المنحنى الممهد .

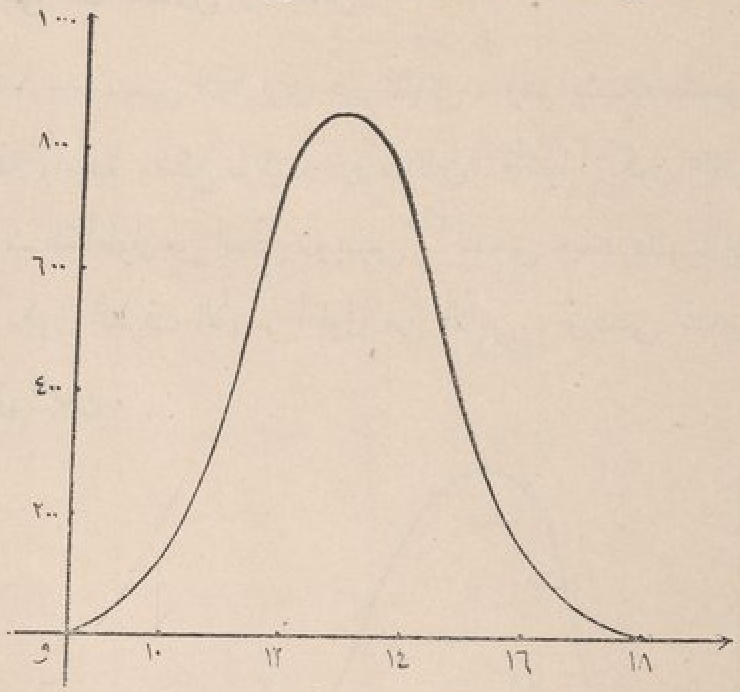
الأشكال المختلفة للمنحنيات التكرارية

١٠٢ — تختلف المنحنيات التكرارية بعضها عن بعض فى الشكل . وهى تتبع فى ذلك كيفية تغير الظواهر التى تمثلها والمقادير التى تأخذها هذه الظواهر عند تغيرها . وفى الواقع يعتبر شكل المنحنى التكرارى لأى مجموعة من المفردات من خواص هذه المجموعة التى تميزها عن غيرها .

ويمكن تقسيم المنحنيات التكرارية من حيث شكلها العام إلى الأنواع الرئيسية الآتى بيانها .

١٠٣ — المنحنى التكرارى المعتدل ، وهو متماثل ، كما نرى فى شكل ٣٨ ويشبه الناقوس العادى . وله محور تماثل رأسى يمر بنقطة النهاية العظمى له ، ويقسمه

إلى جزءين متطابقين . وقد سبقت الإشارة إليه فى بند ٧١ (شكل ٢٤) ، وسنعود إلى شرح بعض خواصه فى مناسبة أخرى . ونجد فى شكل ٣٨ مثالاً عملياً لهذا المنحنى ، حيث نرى المنحنى الممهد للتوزيع التكرارى لأعمار التلاميذ المذكور فى بند ٩٤ ، ويتضح من الشكل أن هذا التوزيع معتدل ومتماثل .



(شكل ٣٨) — منحن تكرارى معتدل

وليست كل المنحنيات التكرارية من هذا النوع متشابهة . فقد يكون المنحنى ضيقاً ومرتفعاً كما فى شكل ٣٨ مثلاً أو قصيراً ومفرطحاً كما فى شكل ٢٤ صفحة ٦٨ ، وهذه الاختلافات تتوقف طبعاً على التوزيعات التكرارية وكيفية تغير الظواهر التى تمثلها هذه التوزيعات ؛ وكذلك على مقياس الرسم المأخوذ على المحورين .

وفى الغالب نجد أن الكميات والظواهر التى تتغير تبعاً لأسباب طبيعية أو حيوية أو وراثية ، تتوزع مقاديرها توزيعاً معتدلا ، إذا كانت المجموعات متجانسة وليست خليطاً .

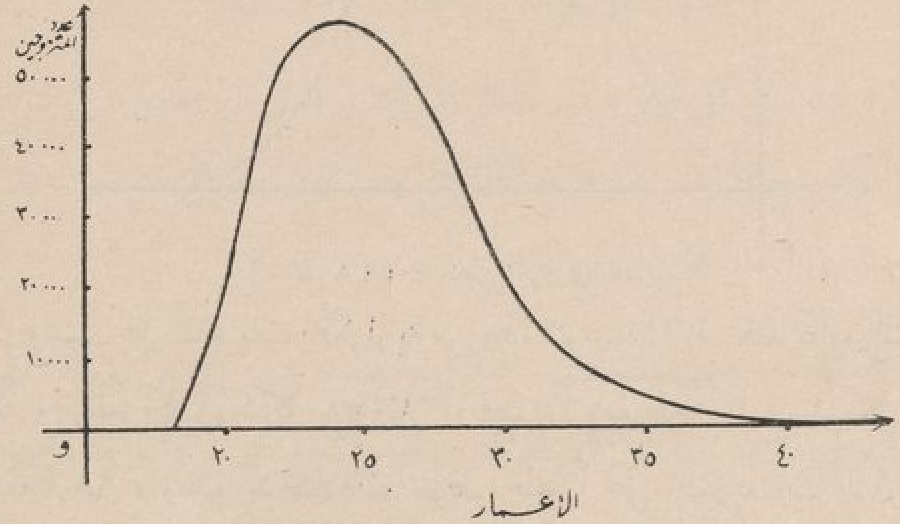
منحنى الخطأ المتماثل

ومن الأمثلة العملية لهذا المنحنى ، توزيع مقادير أخطاء التجارب العملية ، إذا كانت هذه الأخطاء بمحض المصادفة ولم تكن متحيزة فى ناحية معينة .

ففى هذه الحالة نجد أن تكرار خطأ قدره + ٠٣ر٠ مثلا فى قياس كمية معينة (وزن جسم معلوم مثلا) يساوى تكرار الخطأ – ٠٣ر٠ ، وهكذا . فعندما نرسم منحنياً تكرارياً لمقادير الخطأ نجده معتدلا . وهذا هو السبب فى تسمية[1] هذا المنحنى **منحنى الخطأ المعتدل أو المتماثل**.

المنحنى الملتوى إلى اليسار

١٠٤ — المنحنى التكرارى غير المتماثل ، وهو يشابه المنحنى السابق فى أن له قمة واحدة ولكن طرفيه غير متماثلين . فأحياناً يكون الطرف الأيمن ممتداً إلى مسافة أطول من اليسار ، ويسمى[2] المنحنى حينئذ **ملتوياً إلى اليسار** ؛ وأحياناً يكون الطرف الأيسر أطول من الأيمن ، ويسمى المنحنى حينئذ **ملتوياً إلى اليمين** .



(شكل ٣٩)

المنحنى التكرارى لأعمار مجموعة من الرجال عند زواجهم

فالمنحنى الملتوى إلى اليسار يكون صعوده إلى القمة سريعاً وهبوطه منها بطيئاً ؛ والعكس فى المنحنى الملتوى إلى اليمين .

وفى شكل ٣٩ نرى المنحنى التكرارى لأعمار مجموعة من الرجال (العزاب)

---

(١) بالانجليزية Normal or Symmetrical Curve of Error.

(٢) بالإنجليزية Skeiv to the Left

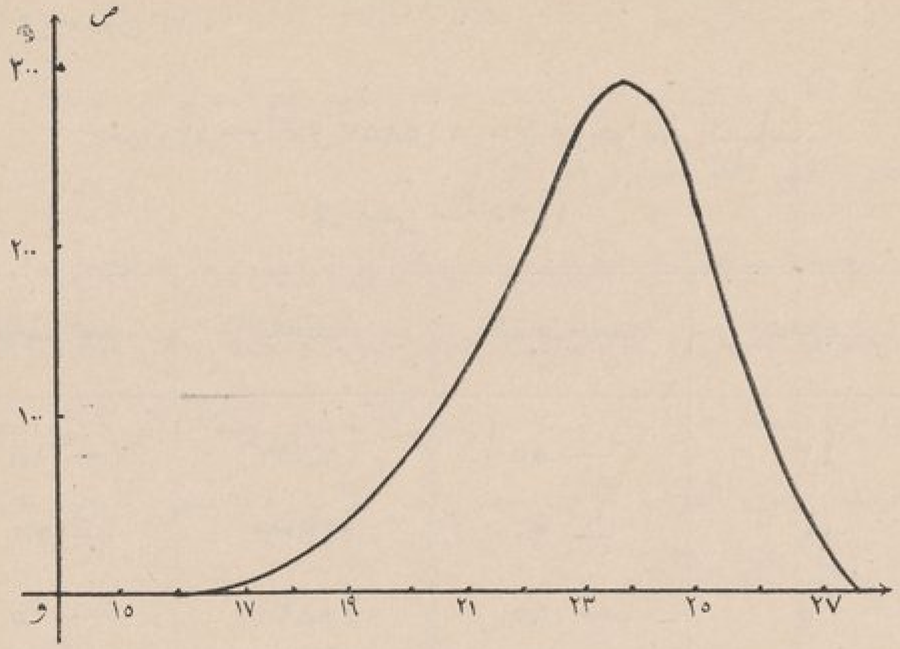
الذين تزوجوا من آنسات فى القطر المصرى فى سـنة ١٩٣٥ ، وتوزيع الأعمار مبين فى جدول ١٣ .

جدول ١٣ — أعمار ١١٥٨٥٧ رجلا تزوجوا من آنسات فى مصر سنة ١٩٣٥

فئات العمر	التكرار	فئات العمر	التكرار
١٨ —	٧٨٥٧	٥٥ —	٤٢
٢٠ —	٥٤٣٣٣	٦٠ —	١٩
٢٥ —	٤٠٨٢١	٦٥ —	٥
٣٠ —	٩٥٧٨	٧٠ —	٥
٣٥ —	٢٤٩٢	٧٥ —	٢
٤٠ —	٤٥٤	٨٠ فأكثر	٣
٤٥ —	١٦١		
٥٠ —	٨٥	الجملة	١١٥٨٥٧

ونلاحظ أن المنحنى يصعد بسرعة من الصفر (عند ١٧ سنة) إلى ٧٨٥٧ مباشرة ثم إلى ٥٤٣٣٣ — وهى القمة — ثم يهبط بالتدريج و ببطء إلى الصفر بعد مسافة على المحور الأفقى مقدارها ٦٠ سـنة . والسبب فى الصعود السريع فى هذه الحالة أن السن القانونية للزواج عند الرجال هى ١٨ سنة ؛ وعلى ذلك لايتزوج أحد قبل هذه السن . وعند بلوغ هذه السن يبادر الكثيرون بالزواج .

وفى شـكل ٤٠ نرى منحنياً تكرارياً ملتوياً إلى اليمين وعكس المنحنى فى شكل ٣٩



(شكل ٤٠)

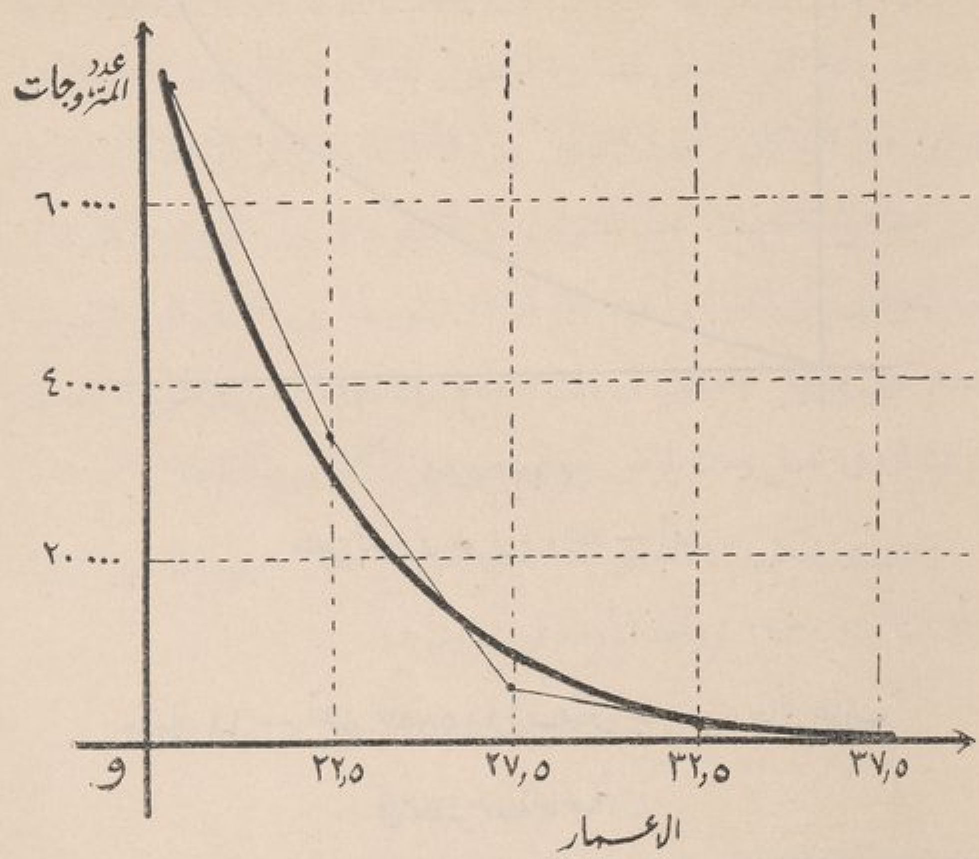
منحن تكرارى ملتو الى اليمين

المنحنى ذو الفرع الواحد

١٠٥ — المنحنى التكرارى ذو الفرع الواحد ؛ وهو إما نازل من اليسار إلى اليمين حيث تكون التكرارات كبيرة عند القيم الصغيرة ، و إما صاعد من اليسار إلى اليمين — وهنا تكون القيم الصغيرة نادرة وتكراراتها صغيرة بعكس القيم الكبيرة فهى شائعة .

ونرى فى شكل ٤١ منحنياً ذا فرع واحد نازلاً إلى اليمين ؛ وهو يبين التوزيع التكرارى لأعمار الآنسات اللاتى تزوجن من رجال (عزاب) فى القطر المصرى سنة ١٩٣٥ . ونرى الأرقام فى الجدول رقم ١٤ .

والسبب فى هذا الوضع أن السن القانونية لزواج الإناث هو ١٦ سنة ، فلا تتزوج منهن واحدة قبل هذه السن . و يظهر من الأرقام والمنحنى الذى يمثلها

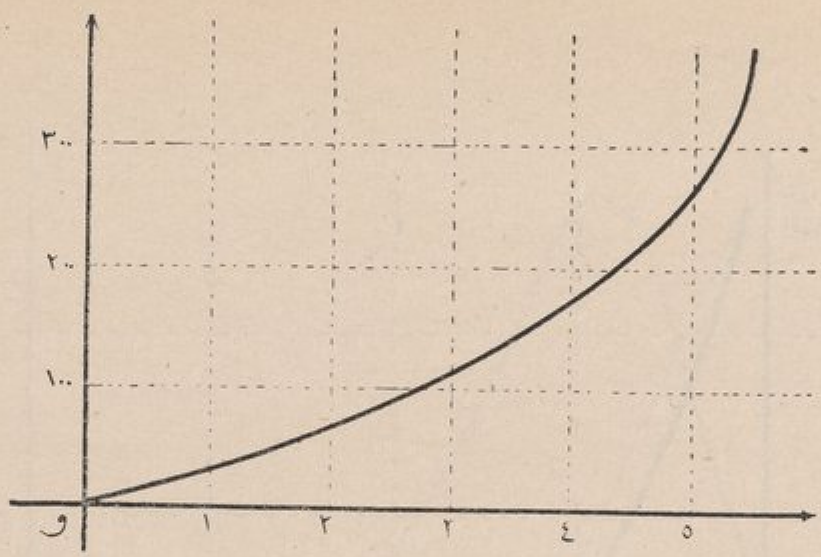


(شكل ٤١)

منحن تكراری لتوزيع أعمار المتزوجات من الآنسات فی مصر سنة ١٩٣٥

أن الآنسات يتزوج أغلبهن قبل سن العشرين ؛ ولا يبقى منهن بدون زواج بعد سن الخامسة والعشرين إلا نسبة ضئيلة (حوالى ٥ ٪) . بخلاف الرجال فإن نسبة صغيرة فقط منهم يتزوجون قبل سن العشرين ، ويبقى حوالى ٥٠ ٪ منهم بعد سن الخامسة والعشرين بدون زواج .

ونرى فی شكل ٤٢ منحنياً تكرارياً ذا فرع واحد صاعداً من اليسار إلى اليمين .



(شكل ٤٢)

منحنى تكرارى ذو فرع واحد — أيمن

جدول ١٤ — أعمار ١١٥٨٥٧ آنسة تزوجن من رجال عزاب

فى مصر سنة ١٩٣٥

فئات العمر	التكرار	فئات العمر	التكرار
١٦ —	٧٤٣٦١	٤٥ —	٤٨
٢٠ —	٣٤٣٢٣	٥٠ —	٢٧
٢٥ —	٦٠٢٠	٥٥ —	١٥
٣٠ —	٧٧٩	٦٠ —	٥
٣٥ —	١٩٧	٦٥ —	٢
٤٠ —	٧٥	٧٠ —	٥
		الجملة	١١٥٨٥٧

منحنى تكرارى ذو شعبتين

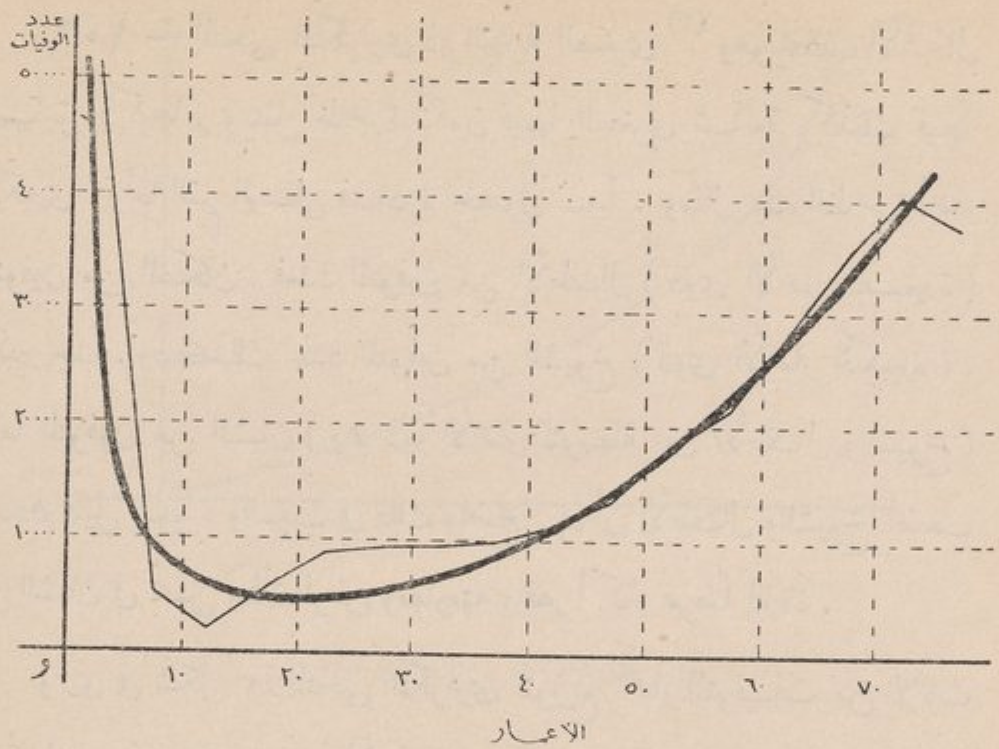
١٠٦ — المنحنى التكرارى ذو النهاية الصغرى [١] وهو يخالف الأشكال السابق ذكرها . ويمثل ظاهرة تكون قيمها الصغرى شائعة وكذلك قيمها الكبرى . وأما القيم الوسطى فنادرة الحصول نسبياً . ومثال هذه الظاهرة أعمار المتوفين من السكان . فعدد المتوفين من الأطفال (ذوى الأعمار الصغيرة) كبير جداً . وكذلك عدد المتوفين من الشيوخ (ذوى الأعمار الكبيرة) . وأما المتوفون من الشبان (وهم ذوو الأعمار المتوسطة بين الأطفال والشيوخ) فعددهم قليل نسبياً . والسبب فى ذلك واضح : إذ أن الأطفال والشيوخ أضعف من الشبان فى تحمل وطأة المرض ومقاومته ، فهم أكثر تعرضاً للوفاة .

ونرى فى شكل ٤٣ المنحنى التكرارى لتوزيع أعمار المتوفيات من الإناث فى ألمانيا سنة ١٩٣٠ ؛ ونجد الأرقام فى جدول ١٥ .

جدول ١٥ — توزيع أعمار المتوفيات (بين عمر ٠ و ٧٥)
فى ألمانيا سنة ١٩٣٠

فئات السن	التكرار	فئات السن	التكرار	فئات السن	التكرار
أقل من ١ سنة	٤١٠٩١	٢٥—	٩٥٦٩	٥٥—	٢١٦٢٧
١ —	١٠٤٨٦	٣٠—	٩٦٥٨	٦٠—	٢٨٢٥٤
٥ —	٥٢٢٤	٣٥—	٩٩٧٦	٦٥—	٣٥٠٩٢
١٠—	٢٣٣٠	٤٠—	١١١٣٦	٧٠—	٤٠٢٧٢
١٥—	٥٧٠٩	٤٥—	١٣١٤٢	٧٥—	٣٧٨٣٦
٢٠—	٨٨٠٣	٥٠—	١٨٠٥٩	الجملة	٣٠٨٢٦٤

(١) بالإنجليزية U-Shaped Curve



(شكل ٤٣)

المنحنى التكرارى لأعمار المتوفيات (بين عمر ٠ و ٧٥) فى ألمانيا سنة ١٩٣٠

المنحنى التكرارى المتجمع

١٠٧ — المنحنى التكرارى المتجمع[١] وهو يبين تراكم التكرارات فى الفئات المتتالية . ولبيان طريقة رسمه نأخذ المثال المذكور فى بند ٩٤ (صفحة ٩٩) . و بواسطة الأرقام المعطاة فى جدول ٩ نكون جدول التكرار المتجمع المبين فى صفحة ١١٩.

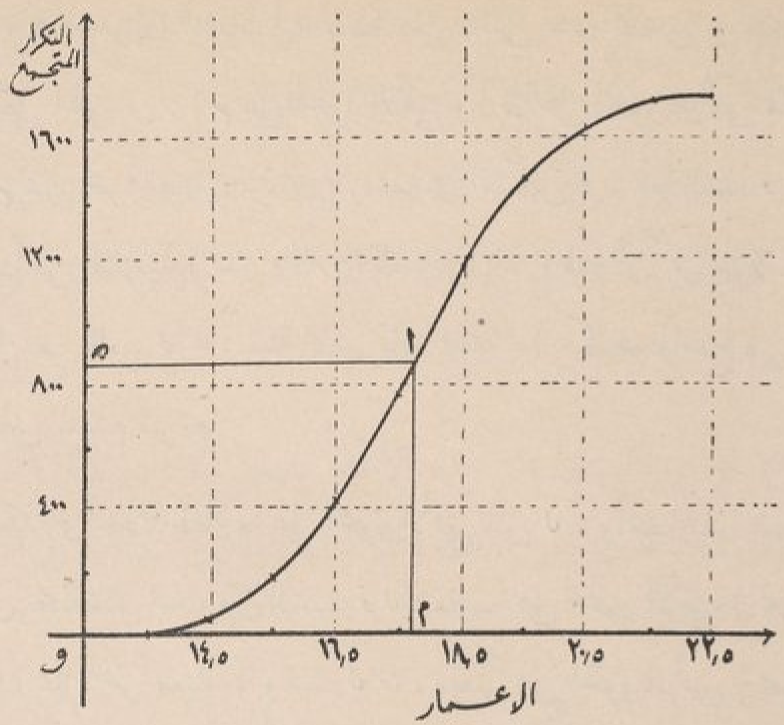
وطريقة حساب التكرارات المتجمعة واضحة ، حيث نجمع تكرار كل فئة على مجموع تكرارات الفئات التى قبلها ؛ والمجموع يمثل عدد المفردات التى أقل من الحد الأعلى لتلك الفئة .

رسم المنحنى الصاعد

ونرسم المنحنى من واقع أرقام التكرارات المتجمعة والحدود العليا للفئات . فنأخذ هذه الحدود العليا كإحداثيات أفقية ، والتكرارات المتجمعة المناظرة لها كإحداثيات رأسية . والخط البيانى الذى يصل بين النقط المرصودة بهذه الإحداثيات هو المنحنى التكرارى المتجمع (الصاعد) المطلوب .

---

(١) بالأنجليزية Cumulative Frequency Curve = Ogive



(شكل ٤٤)

المنحنى التكرارى المتجمع الصاعد لأعمار ١٧٣٩ تلميذاً

جدول ١٦ — التكرارات المتجمعة لأعمار ١٧٣٩ تلميذاً

فئات العمر بالسنين	عدد التلاميذ (التكرار)	الحدود العليا للفئات	التكرار المتجمع (الصاعد)
		أقل من ١٣٫٥	٠
١٣٫٥ وأقل من ١٤٫٥	٣٤	« ١٤٫٥	٣٤
١٤٫٥ —	١٢٨	« ١٥٫٥	١٦٢
١٥٫٥ —	٢٦٢	« ١٦٫٥	٤٢٤
١٦٫٥ —	٣٦٠	« ١٧٫٥	٧٨٤
١٧٫٥ —	٣٨٦	« ١٨٫٥	١١٧٠
١٨٫٥ —	٢٩٤	« ١٩٫٥	١٤٦٤
١٩٫٥ —	١٦٧	« ٢٠٫٥	١٦٣١
٢٠٫٥ —	٩٢	« ٢١٫٥	١٧٢٣
٢١٫٥ واقل من ٢٢٫٥	١٦	« ٢٢٫٥	١٧٣٩
الجملة	١٧٣٩		

نقط المنحنى ومعنى إحداثياتها

١٠٨ — وإذا أخذنا أية نقطة مثل ا على هذا المنحنى وأسقطنا منها عمودين على المحورين : ا م على المحور الأفقى و ا ن على المحور الرأسى مثلا ، فإن البعد و م يمثل عمراً معيناً (= ١٧٫٧ سنة فى الشكل) . أما البعد و ن فهو يمثل عدداً من التلاميذ (= ٨٧٠ فى الشكل) . والمهم أن كل هؤلاء التلاميذ عمرهم أقل من العمر ١٧٫٧ سنة الذى يمثله البعد و ا ، كما هو واضح من طريقة رسم المنحنى المتقدم شرحها .

وهكذا لو أخذنا نقطة ما على المحور الأفقى مثل ل ، وأقمنا منها عموداً يقابل المنحنى فى نقطة مثل ب ، فإن البعد و ل ، مقيساً على المحور الأفقى (محور قياس الأعمار) ، يمثل عمراً معيناً ؛ والبعد ل ب ، مقيساً على المحور الرأسى (محور قياس التكرارات المتجمعة) ، يمثل عدداً من الأشخاص أعمارهم جميعاً دون العمر و ل .

وأخيراً ، لو أخذنا أية نقطة على المحور الرأسى مثل ك ، وأقمنا منها عموداً على هذا المحور الرأسى ليقابل المنحنى فى نقطة حـ مثلا ، فإن البعد و ك ، مقيساً على المحور الرأسى ، يمثل عدداً من التلاميذ عمرهم جميعاً أقل من العمر الذى يمثله البعد ك حـ على المحور الأفقى .

خواص المنحنى الصاعد لها فائدة عملية

١٠٩ — وهذه الخواص التى للمنحنى التكرارى المتجمع ذات فائدة كبيرة ، ونستخدمها كثيراً كما يتبين لنا فى المستقبل . ويجب التنبه إلى نقطة مهمة عند رسم هذا المنحنى . وهى أن الإحداثيات الأفقية هى الحدود العليا للفئات التكرارية التى عندنا . و إغفال هذه الخاصة ينشأ عنه أخطاء فى العمل .

وهذا المنحنى نسميه المنحنى الصاعد لأن تراكم التكرارات دائماً فى ازدياد . وعلى ذلك فالمنحنى دائماً فى الصعود ، بالغاً ذروته عند آخر نقطة فيه . وهذه النقطة أعلى نقطة على المنحنى لأن إحداثيها الرأسى يساوى مجموع تكرارات

الفئات جميعها ، فهو بالضرورة أكبر إحداثى رأسى . و إحداثيها الأفقى هو الحد الأعلى للفئة الكبرى ، فهو كذلك أكبر إحداثى أفقى .

و يلاحظ ازدياد سرعة صعود المنحنى فى وسطه ، وذلك لأن الفئات الوسطى فى العادة تكراراتها أكبر من تكرارات الفئات المتطرفة . وعلى ذلك تكون سرعة زيادة التكرارات المتجمعة أكبر فى الوسط .

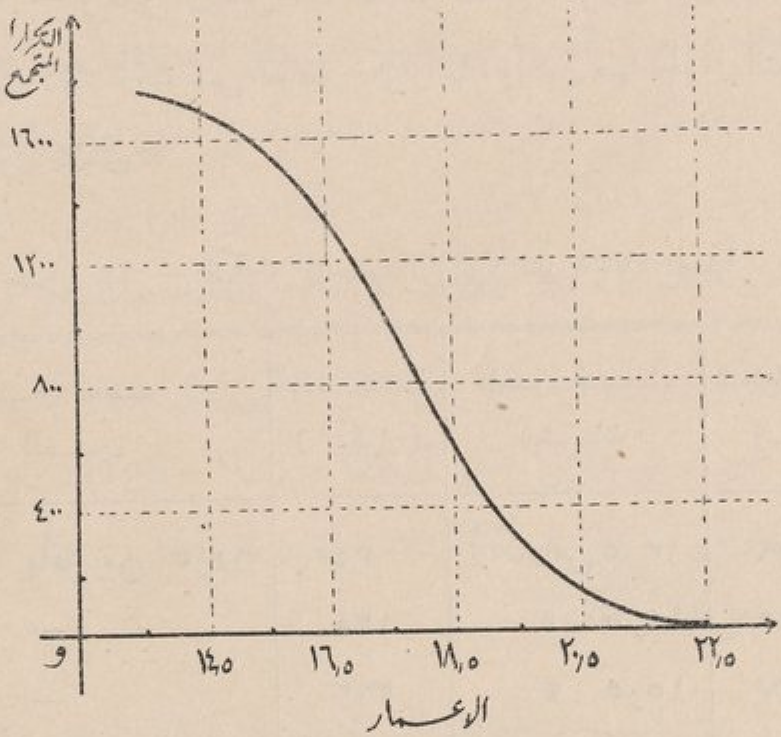
المنحنى التكرارى المتجمع النازل

١١٠ — ويمكن أن نمثل هذه الأعمار بمنحن تكرارى نازل (متجمع) أيضاً . ولبيان ذلك نأخذ نفس الفئات وتكراراتها ، ونكون منها الجدول التكرارى المتجمع النازل كما يأتى :

جدول ١٧ — التكرار المتجمع النازل لأعمار ١٧٣٩ تلميذاً

فئات العمر بالسنين	عدد التلاميذ (التكرار)	الحدود الدنيا للفئات	التكرار المتجمع النازل
١٣ر٥ وأقل من ١٤ر٥	٣٤	أكبر من ١٣ر٥	١٧٣٩
١٤ر٥ —	١٢٨	« ١٤ر٥	١٧٠٥
١٥ر٥ —	٢٦٢	« ١٥ر٥	١٥٧٧
١٦ر٥ —	٣٦٠	« ١٦ر٥	١٣١٥
١٧ر٥ —	٣٨٦	« ١٧ر٥	٩٥٥
١٨ر٥ —	٢٩٤	« ١٨ر٥	٥٦٩
١٩ر٥ —	١٦٧	« ١٩ر٥	٢٧٥
٢٠ر٥ —	٩٢	« ٢٠ر٥	١٠٨
٢١ر٥ — ٢٢ر٥	١٦	« ٢١ر٥	١٦
الجملة	١٧٣٩	« ٢٢ر٥	٠

وطريقة تكوين التكرار المتجمع النازل هى أننا نبدأ بالمجموع الكلى للتكرارات أمام مبدأ الفئة الأولى ؛ ثم نطرح تكرار الفئة الأولى فيكون الباقى عدد المفردات التى أكبر من الحد الأدنى للفئة الثانية ، وهكذا مع باقى الفئات . والأفضل أن نبدأ بوضع صفر أمام الحد الأعلى للفئة الأخيرة ، ثم نضيف تكرار كل فئة إلى مجموع تكرارات الفئات التى أسفلها حتى نصل إلى المجموع الكلى أمام الفئة الأولى .



(شكل ٤٥)

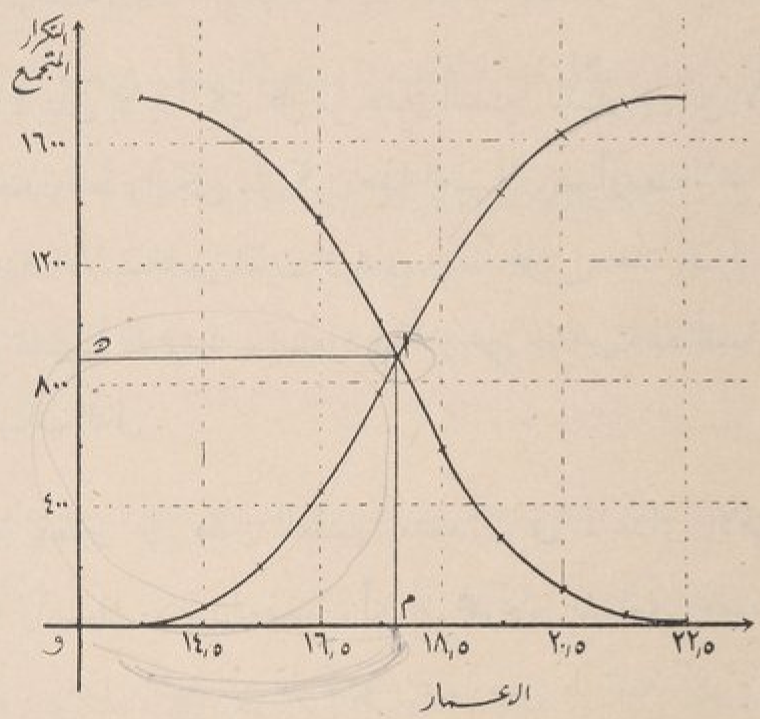
المنحنى التكرارى المتجمع النازل لأعمار ١٧٣٩ تلميذاً

رسم المنحنى وخواصه

ونرسم المنحنى أيضاً من واقع هذه الأرقام . فنأخذ الحدود الدنيا للفئات الأصلية كإحداثيات أفقية ، والتكرارات المتجمعة المتناقصة كإحداثيات رأسية . ثم نصل بين النقط المرصودة من واقع هذه الإحداثيات ، فنحصل على المنحنى المطلوب كما فى شكل ٤٥

ولهذا المنحنى خواص مماثلة لخواص المنحنى الصاعد التى ذكرناها فى بند ١٠٨ . فكل نقطة عليه يمثل إحداثيها الرأسى عدداً من الأشخاص عمرهم جميعاً أكبر من العمر الذى يمثله إحداثيها الأفقى مقيساً على المحور الأفقى (وهو محور قياس الأعمار) ، وهذا ناتج طبعاً من تعريف المنحنى وطريقة رسمه .

ويلاحظ أن المنحنى دائماً فى الهبوط من اليسار إلى اليمين ؛ ويكون هبوطه فى الوسط أسرع من هبوطه فى الطرفين . وذلك لنفس السبب الذى ذكرناه فى حالة المنحنى الصاعد .



(شكل ٤٦)

المنحنى التكرارى المتجمع الصاعد والمنحنى النازل لأعمار ١٧٣٩ تلميذاً

نقطة تقابل المنحنيين الصاعد والنازل

١١٠ — وإذا رسمنا المنحنيين فى شكل واحد وبنفس مقياس الرسم على المحورين فلا بد أن يتقابلا فى نقطة . وهذه النقطة لها خاصة مفيدة عملياً (وهى أن إحداثيها الرأسى يساوى نصف مجموع التكرارات كلها أى $\frac{1}{2}$ × ١٧٣٩

فى هذا المثـال) . والسبب فى ذلك واضح . فبما أن هذه النقطة ا على المنحنى الصاعد (وليكن إحداثيها الأفقى و م كما فى شكل ٤٦) يكون إحداثيها الرأسى م ا يساوى عدد الأشخاص الذين عمرهم أقل من العمر و م ، كما قلنا فى بند ١٠٧ . وبما أنها فى الوقت نفسه واقعة على المنحنى النازل ، فإحداثيها الرأسى نفسه يمثل عدداً من الأشخاص أعمارهم جميعاً أكبر من العمر و م أيضاً ، كما ذكرنا فى بند ١٠٩ ، وعلى ذلك فعدد الأشخاص الذين عمرهم أكبر من العمر و م ، يساوى عدد الأشخاص الذين عمرهم أقل من هذا العمر نفسه .

وهذا لايتأتى إلا إذا كان كل من هذين العددين المتساويين من الأشخاص يساوى نصف العدد الكلى . ويكون هذا العمر إذن هو أوسط الأعمار كلها . ويكون عدد الأشخاص الذين يزيدون سناً عن هذه السن يساوى عدد الذين تنقص أعمارهم عنه . وسنعود للكلام على خواص هذه القيمة الوسطى للعمر فى الباب التالى .

وكل نقطتين على هذين المنحنيين متحدتين فى الإحداثى الأفقى يكون إحداثياهما لرأسيان مكملين لبعضهما ، أى أن مجموعهما يساوى المجموع الكلى للتكرارات . وهذا واضح من طريقة رسمهما ومن طريقة تكوين التكرار المتجمع الصاعد والنازل .

١١٢ — يجب التنبه إلى الفرق بين هذا المنحنى التجميعى والمنحنى التكرارى العادى . فالإحداثيات الرأسية فى الأول تدل على مجموع تكرارات . وهذا المجموع فى المنحنى التكرارى العادى تمثله مساحة ذلك المنحنى ، كما ذكرنا فى بند ٩٥ و بند ٩٩ عند الكلام على الهستوجرام والمنحنى التكرارى المشتق منه . العلاقة بين المنحنى المتجمع والمنحنى التكرارى العادى

فإذا أردنا معرفة عدد المفردات التى أقل من مقدار معين ، نأخذ على المحور الأفقى طولا يساوى هذا المقدار ، ونقيم من نهايته عموداً يقابل المنحنى التكرارى المتجمع فى نقطة ؛ طول هذا العمود يساوى عدد المفردات المطلوب . أما فى المنحنى التكرارى العادى فنرسم هذا العمود بنفس الطريقة ليقابل ذلك المنحنى فى نقطة . وتكون المساحة المحصورة بين المنحنى والمحور الأفقى و الواقعة على يسار العمود المرسوم تمثل عدد المفردات التى أقل من هذا المقدار المعين .

وهذه العلاقة هى ما يعبر عنه فى الرياضة بأن المنحنى المتجمع هو « تكامل » المنحنى العادى ؛ أو أن المنحنى العادى هو « تفاضل » المنحنى المتجمع . ونكتفى بهذه الإشارة إذ أن نطاق هذا الكتاب لا يسمح بالإسهاب فى هذا الموضوع .

المنحنى التكرارى المتعدد القمم

١١٣ — نجد فى بعض الأحيان مجموعات تكرارية يكون المنحنى التكرارى لها ذا قمتين أو أكثر . وهذا بخلاف المجموعات العادية ذات التكرار المنتظم التى فيها يزداد تكرارات الفئات بالتدريج حتى تصل إلى نهاية كبرى — تمثلها قمة المنحنى — ثم تتناقص بعد ذلك بالتدريج ، ولا تزيد ثانياً .

ونرى مثل هذا المنحنى المتعدد القمم فى شكل ٤٧ ؛ وهو يمثل التوزيع التكرارى المبين فى جدول رقم ١٨

خلط المجموعات المختلفة بسبب تعدد القمم أحياناً

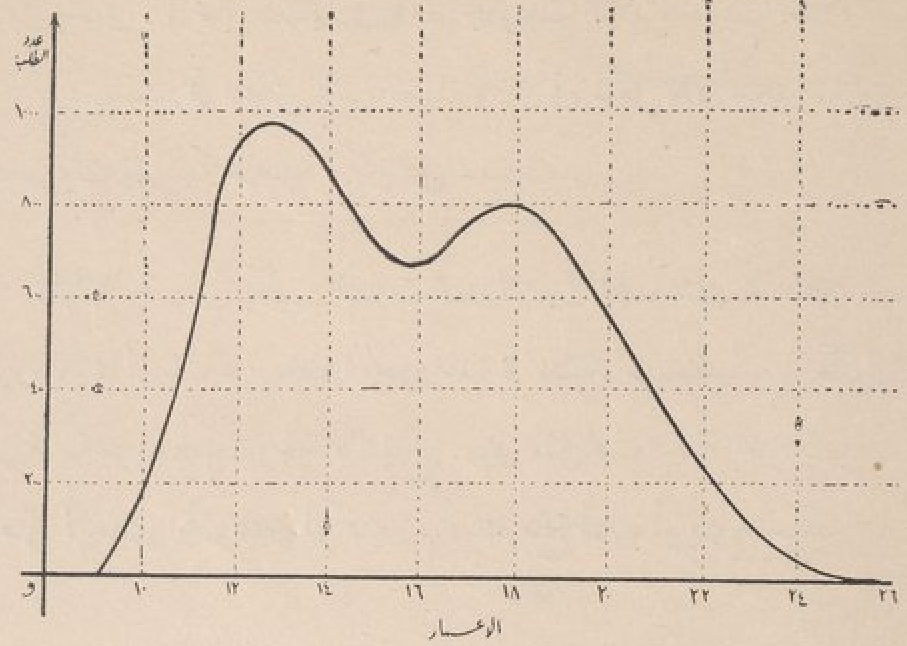
١١٤ — ومثل هذه المجموعات تكون مفرداتها غير متجانسة ، كأن تكون خليطاً بين مجموعتين مختلفتين أو أكثر . وهذا الاختلاف بين المجموعات المركبة ينشأ عنه ازدياد فى تكرارات الفئات الأولى حتى تصل إلى نهاية كبرى ثم تنقص . ولكن لا تلبث أن تأخذ التكرارات فى الزيادة مرة ثانية حتى تصل إلى نهاية كبرى ثانية — وربما ثالثة أو أكثر . وهذا التذبذب يظهر أثره فى المنحنى التكرارى فتتعدد القمم .

جدول ١٨ — توزيع أعمار تلاميذ المدارس الأميرية
المتقدمين للامتحانات العامة فى سنة ١٩٣٥

العمر	عدد التلاميذ	العمر	عدد التلاميذ
٩	١	١٩	٧٦٠
١٠	١٥٠	٢٠	٥٨٠
١١	٥٧٠	٢١	٣٧٧
١٢	٩٥٠	٢٢	٢٢٣
١٣	٩٨٧	٢٣	١١٢
١٤	٨٢٤	٢٤	٣٠
١٥	٦٦٨	٢٥	١١
١٦	٦٥٦	٢٦	٦
١٧	٧٧٠	٢٧	٢
١٨	٨٠٠	المجموع	٨٤٩٥

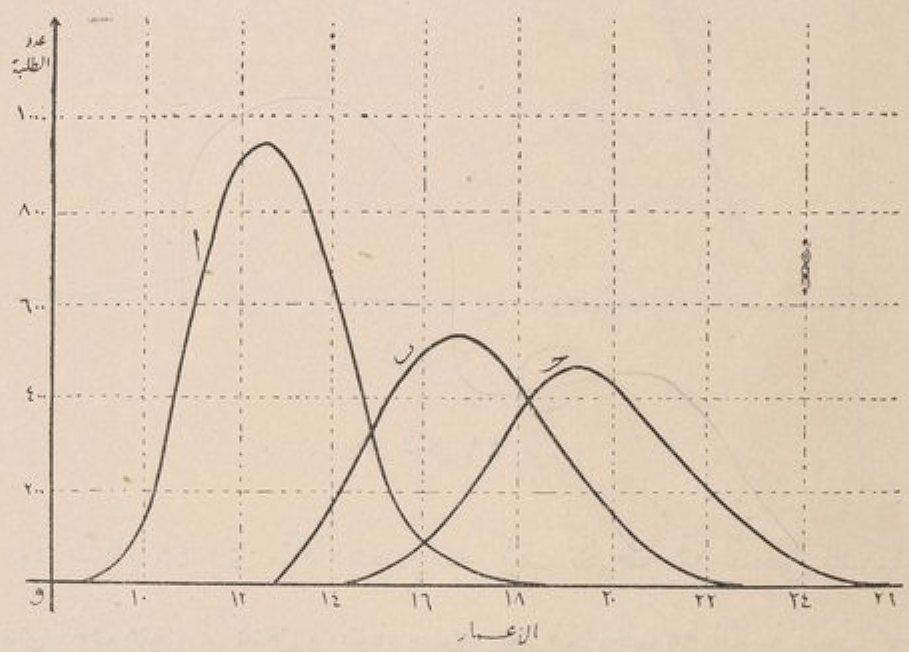
تعدد القمم دليل على عدم التجانس

وفى جميع الأحوال يؤخذ تعدد القمم فى المنحنى التكرارى الممهد دليلا على عدم تجانس المجموعة التى يمثلها هذا المنحنى ؛ وبالتالى على شدة التباين والتفاوت بين مفرداتها . ومن هذا يمكن الاهتداء إلى زيادة البحث فى مفردات مثل هذه المجموعات ، لعلّنا نكتشف خاصة معينة تميز بعض المفردات عن البعض الآخر ؛ وبذلك تنقسم المجموعة الأصلية إلى مجموعتين (أو أكثر) يكون كل منها على حدة أكثر تجانساً من المجموعة الكبرى . فإذا رجعنا مثلا إلى مصدر الأرقام الموجودة فى جدول ١٨ ، نجد أن هؤلاء التلاميذ خليط من ثلاثة أنواع متباينة .



(شكل ٤٧)

منحن تكرارى ذو قمتين



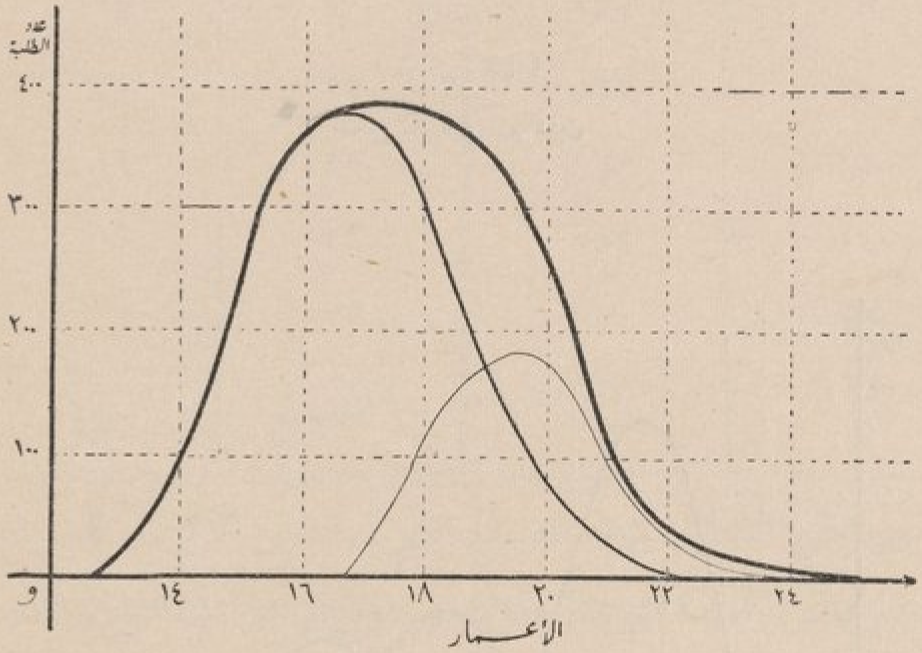
(شكل ٤٨)

المنحنيات التكرارية لثلاث مجموعات مختلفة

وهم المتقدمون إلى شهادة الدراسة الابتدائية ، وشهادتى الدراسة الثانوية بقسميها الأول والثانى . وأعمار التلاميذ فى هذه المجموعات الثلاث تختلف اختلافاً كبيراً . وهذا يسبب تعدد قمم المنحنى عند خلطها . ونرى فى شكل ٤٨ المنحنيات التكرارية للمجموعات المركّبة الأصلية ، كل على حدة .

خلط المجموعات يسبب أحياناً اختلال التماثل أو تفرطح القمة

١١٥ — حينما يكون التفاوت معتدلا بين المجموعتين المركبتين لمجموعة كبيرة ، ربما لا ينشأ عن خلطها وجود عدة قمم متميزة من بعضها ، بل تكون هذه القمم قريبة من بعضها لدرجة الاندماج . وفى مثل هذه الحالة يختل تماثل المنحنى الأصلى لإحدى المجموعتين المركبتين ، أو تتفرطح قمته ، أو ينتفخ جذعه .



(شكل ٤٩)

وفى شكل ٤٩ نرى مثالا يتضح منه كيف تتفرطح قمة المنحنى التكرارى لإحدى المجموعتين المركبتين وينتفخ جذعه بعد ضم المجموعة الثانية إليها . ونجد فى جدول ١٩ تكرارات المجموعتين الأصليتين والمجموعة الكلية . وهما عبارة

جدول ١٩ — توزيع أعمار مجموعتين من التلاميذ
والمجموعة المركبة منهما معاً

الأعمار بالسنين	تكرارات المجموعة الأولى	تكرارات المجموعة الثانية	تكرارات المجموعة المركبة منهما
١٣	١١		١١
١٤	٩٣		٩٣
١٥	٢٤٨		٢٤٨
١٦	٣٦٩		٣٦٩
١٧	٣٩٠	٧	٣٩٧
١٨	٢٧٤	١١٥	٣٨٩
١٩	١٨٧	١٧٠	٣٥٧
٢٠	٨٧	١٩١	٢٧٨
٢١	٢٩	١٠٦	١٣٥
٢٢	٣	٣١	٣٤
٢٣	٢	٧	٩
٢٤	٠	٣	٣
٢٥	٠	١	١
الجملة	١٦٩٣	٦٣١	٢٣٢٤

عن التوزيع التكرارى لأعمار الناجحين من تلاميذ المدارس الأميرية وتلاميذ التعليم الأولى فى امتحان شهادة الدراسة الثانوية قسم أول سنة ١٩٣٤

ونرى فى الشكل ثلاثة منحنيات : ا (الداخلى الأيسر) و ب (الداخلى الأيمن) و حـ (الخارجى) وهى تمثل المجموعة الأولى والثانية والثالثة على الترتيب .

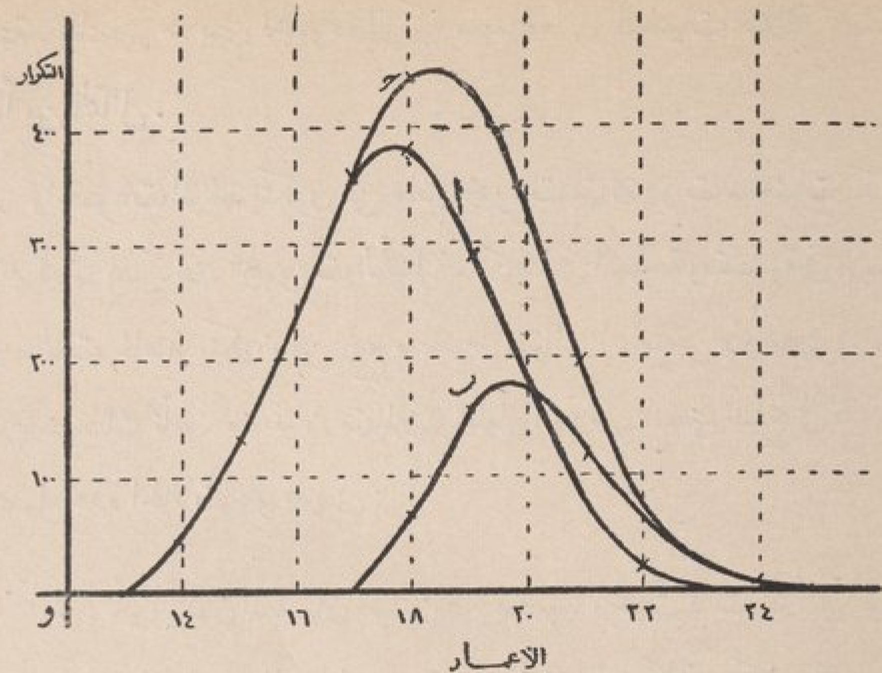
شكل المنحنى التكرارى يدل على درجة التفاوت بين المفردات

١١٦ — يتضح لنا من هذه الأمثلة العملية التى قدمناها بشأن خلط المجموعات أن عدم تماثل المنحنى التكرارى قد ينشأ عن اختلاط المفردات وعدم تجانسها ؛ وكذلك تفرطح قمته وانتفاخ جـذعه . وهذه الظواهر تدل أيضاً على وجود التباين والتفاوت بين مفردات المجموعات التى تمثلها هذه المنحنيات الملتوية أو المفرطحة أو السميكة . وهكذا يمكن الاستدلال من شكل المنحنى التكرارى لمجموعة ما ، على بعض خواص مفرداتها كدرجة تجانسها وتقاربها من بعضها ، أو تفاوتها وتشتتها من بعضها . وسنعود إلى الكلام فى هذه النقطة بتوسع أكثر فى مناسبة أخرى . ولذلك يحسن التنبه إلى فائدة المنحنيات التكرارية من هذه الوجهة .

خلط المجموعات غير المتماثلة قد ينتج مجموعة متماثلة

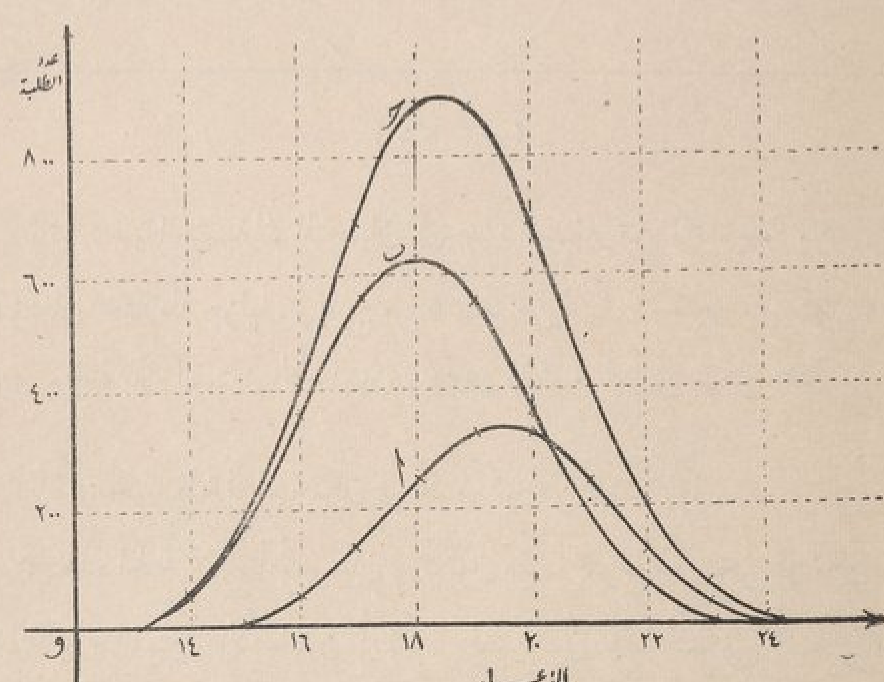
١١٧ — وكما أن خلط المجموعات يؤدى أحياناً إلى اختلال تماثل منحنى المجموعة المركبة منها ، يصح أن هذا الخلط بين مجموعتين يسوى ما فى إحداها أو كليهما من التواء ، فتنتج منهما مجموعة متماثلة أو أقرب إلى التماثل من أيهما على حدة . ونرى مثالا لذلك فى شكل ٥٠ حيث نرى ثلاثة منحنيات منها ا ذو التواء بسيط و ب ذو التواء أكبر نوعاً . والمنحنى حـ المركب منهما أقرب إلى التماثل من كليهما . والمنحنى ا هنا يمثل توزيع أعمار تلاميذ المدارس الأميرية الناجحين فى امتحان شهادة الدراسة الثانوية قسم أول سنة ١٩٣٣ ، والمنحنى ب يبين توزيع أعمار مجموعة أخرى من الطلبة .

١١٨ — وقد تنتج لدينا مجموعة متماثلة من خلط مجموعتين متماثلتين ، كما نرى فى شكل ٥١ ؛ حيث المنحنى ا يبين توزيع أعمار المتقدمين إلى شهادة



(شكل ٥٠)

الدراسة الثانوية قسم ثان (علمى) من المدارس الأميرية فى سنة ١٩٣٣ ، والمنحنى ب يبين توزيع أعمار المتقدمين من هذه المدارس فى امتحان قسم أول فى نفس

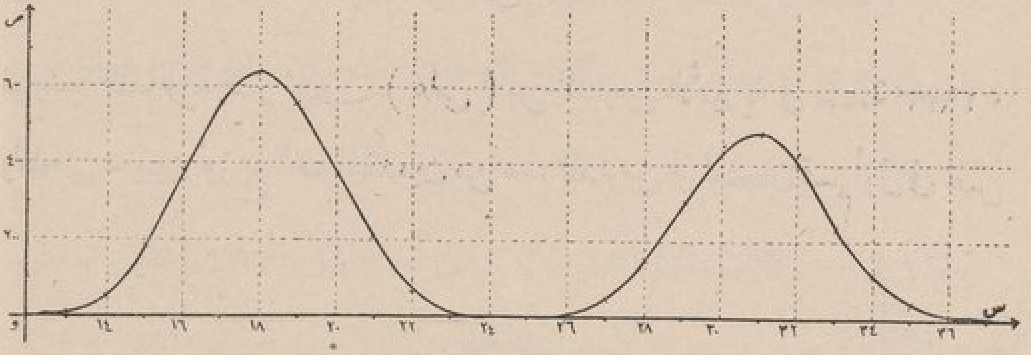


(شكل ٥١)

السنة . والمنحنى ح يبين المجموعة المركبة منهما معاً . والمنحنيات الثلاثة كلها قريبة جداً من التماثل .

والمجموعات المتماثلة يتكون من خلطها مجموعات متماثلة ذوات منحنيات تكرارية متماثلة أيضاً إذا كانت المجموعات الأصلية متكافئة فى المساحة ومتشابهة فى الشكل . وفى حالات خاصة يتكون من جمع منحنيين متماثلين ، منحن ذو قمتين أو ملتو . ولتوضيح ذلك نأخذ منحنيين متماثلين ونخلطهما ونرسم المنحنى الناتج فى كل حالة . وتفصيل هذه الحالات هو كما يأتى :

أولا : حينما يكون المنحنيان بعيدين عن بعضهما ، بحيث لا يشتركان فى أى جزء من قاعدتيهما . هنا يبقى المنحنيان مستقلين ولا يختلطان كما فى شكل ٥٢

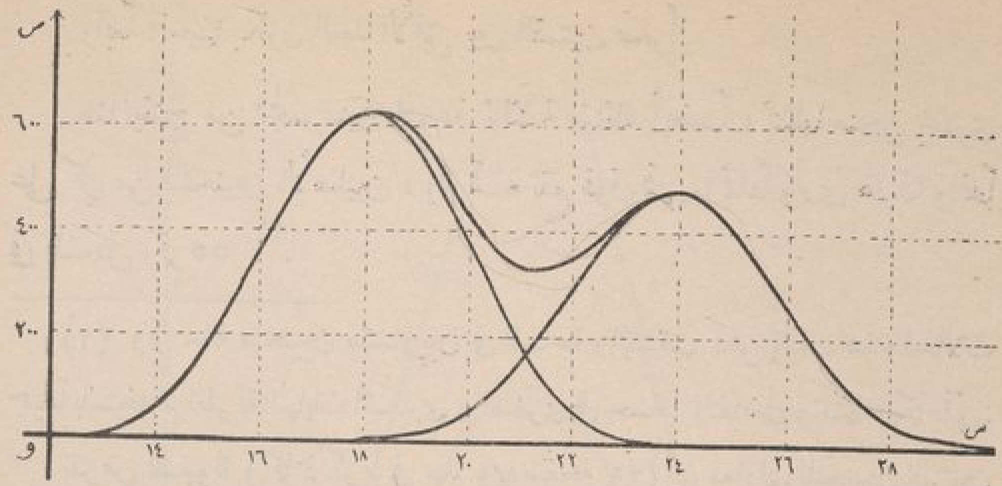


(شكل ٥٢)

ثانياً : حينما يشترك المنحنيان فى جزء صغير من قاعدتيهما . هنا تختلط المجموعتان اختلاطاً جزئياً ، و ينتج من هذا منحن ثالث ذو قمتين ، و يكون له نهاية صغرى واقعة بين القمتين الأصليتين . ونرى ذلك موضحاً فى شكل ٥٣

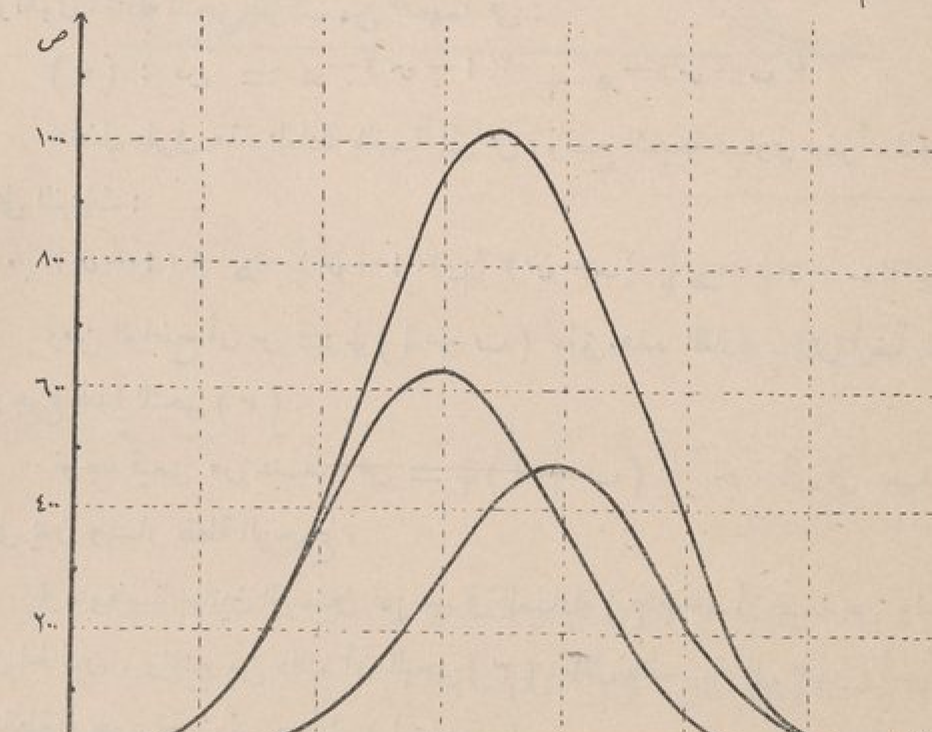
ثالثاً : حينما يشترك المنحنيان فى جزء كبير من قاعدتيهما .

فى هذه الحالة تمتزج المجموعتان و يتكون منهما مجموعة لها منحن ذو قمة واحدة ، واقعة بين القمتين الأصليتين أيضاً . وقد يكون هذا المنحنى الناتج متماثلا كما فى شكل ٥٤ الذى يبين التوزيعين التكراريين لأعمار تلاميذ المدارس الأميرية



(شكل ٥٣)

المتقدمين إلى امتحان شهادة الدراسة الثانوية قسم أول وقسم ثان فى سنة ١٩٣٣ ومن الرسم نرى أن الإحداثى الأفقى لقمة المنحنى الأول عند ١٨ سنة ، والثانى



(شكل ٥٤)

عند ٢٠ سنة . والمنحنى الناتج من المجموعتين فى هذه الحالة متماثل أيضاً (تقريباً) والإحداثى الأفقى لقمته عند ١٩ تقريباً . وفى كثير من الأحوال ينتج منحن غير متماثل .

رابعاً : حينما يكون البعد الأفقى بين القمتين صفراً

هنا ينتج من المجموعتين مجموعة ثالثة متماثلة أيضاً ، يمثلها منحن يحتوى على كل من المنحنيين الأصليين ، وله قمة تقع فوق قمتيهما تماماً ونرى هذا واضحاً فى الشكل رقم ٥٥ [١].

---

(١) وفى حالة منحنيين متساويين فى المساحة والتشتت يمكن إثبات هذه الحالات جميعاً باستخدام نظرية النهايات الكبرى والصغرى فى حساب التفاضل والتكامل كما يأتى:

نفرض للسهولة (كما ذكرنا فى بند ٧١ صفحة ٦٩) أن معادلتى المنحنيين المتماثلين هما كما يأتى:

$$(١) : \text{ص}_{١} = \text{هـ}^{-(\text{س}-\text{ا})^{٢}}$$

$$\text{و}(٢) : \text{ص}_{٢} = \text{هـ}^{-(\text{س}-\text{ب})^{٢}} \quad , \quad \text{ب} > \text{ا}$$

وتكون معادلة المنحنى المركب من جمعهما هى:

$$(٣) : \text{ص} = \text{هـ}^{-(\text{س}-\text{ا})^{٢}} + \text{هـ}^{-(\text{س}-\text{ب})^{٢}}$$

نفاضل طرفى هذه المعادلة بالنسبة إلى س ، ونضع النتيجة تساوى صفراً ، لنحصل على النهايات:

$$\therefore (\text{س}-\text{ا}) \,.\, \text{هـ}^{-(\text{س}-\text{ا})^{٢}} + (\text{س}-\text{ب}) \,.\, \text{هـ}^{-(\text{س}-\text{ب})^{٢}} = ٠$$

ومن الواضح أن $\text{س} = \frac{١}{٢}(\text{ا}+\text{ب})$ تحقق هذه المعادلة ، وهى أيضاً نقطة رجوع لهذا المنحنى (٣).

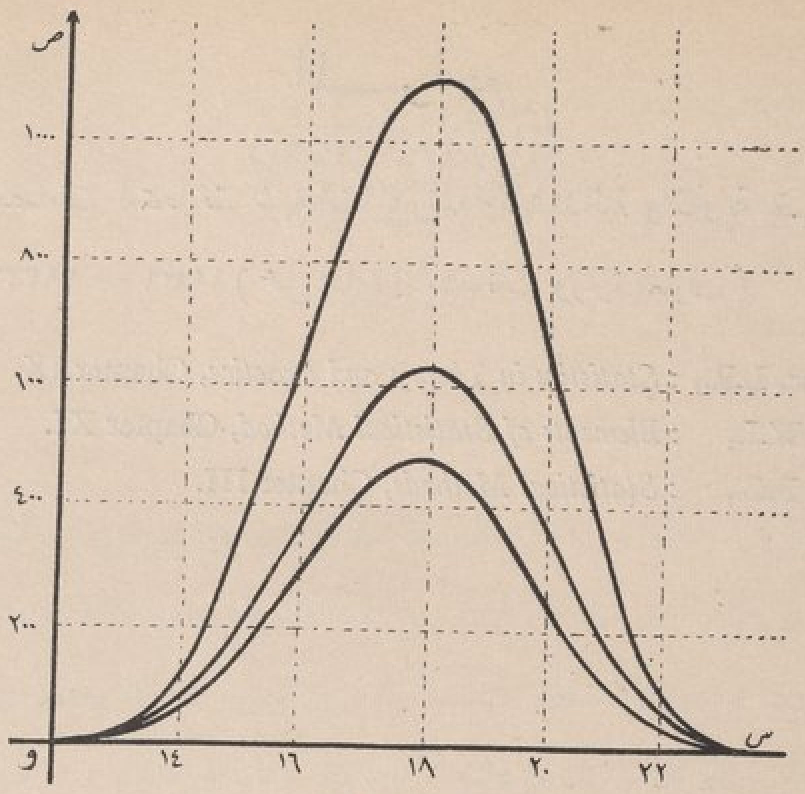
نوجد قيمتى ص عندما $\text{س} = \frac{١}{٢}(\text{ا}+\text{ب}) \pm \text{حـ}$ ، أى على بعد حـ على يمين ويسار نقطة الرجوع.

لو عوضنا بهاتين القيمتين عن س فى المعادلة (٣) نجد أن قيمة ص واحدة فى الحالتين. وينتج من ذلك أن المنحنى (٣) متماثل بالنسبة إلى محور رأسى يمر بنقطة الرجوع حيث $\text{س} = \frac{١}{٢}(\text{ا}+\text{ب})$.

ولو فاضلنا ص مرة أخرى بالنسبة إلى س ، نحصل على المشتقة الثانية وهى:

$$\text{ص}'' = -٢\,\text{هـ}^{-(\text{س}-\text{ا})^{٢}} - ٢\,\text{هـ}^{-(\text{س}-\text{ب})^{٢}} + ٤(\text{س}-\text{ا})^{٢}$$
$$\times\, \text{هـ}^{-(\text{س}-\text{ا})^{٢}} + ٤(\text{س}-\text{ب})^{٢} \,.\, \text{هـ}^{-(\text{س}-\text{ب})^{٢}}.$$

فلو عوضنا عن س بقيمتها عند نقطة الرجوع أى $\text{س} = \frac{١}{٢}(\text{ا}+\text{ب})$ ، (*)



(شكل ٥٥)

(*) ∴ ص″ = [− ٤ + ٢ (ب − ا)^{٢}] × هـ مرفوعة إلى القوة − $\left(\frac{\text{ب} - \text{ا}}{٢} \right)^{٢}$

أى أن ص″ موجبة إذا كان (ب − ا)^{٢} > ٢ ، وسالبة إذا كان (ب − ا)^{٢} < ٢ .

∴ نقطة الرجوع تكون نهاية صغرى إذا كان ب − ا > $\sqrt{٢}$ ويكون للمنحنى (٣) حينئذ قمتان ، كما فى ثانياً (وأولا) .

وإذا كان ب − ا < $\sqrt{٢}$ ، تكون نقطة الرجوع نهاية كبرى ، ويكون للمنحنى نهاية كبرى واحدة ، أى قمة واحدة . ويكون المنحنى الناتج فى هذه الحالة متماثلا وذا قمة واحدة . ولكن يجب ملاحظة أن معادلة هذا المنحنى تختلف فى تركيبها عن معادلة أى واحد من المنحنيين الأصليين .

المـــراجع

إحصاءات امتحانات شهادات الدراسة الابتـدائية والثانوية بقسميها فى السنين ١٩٣٣ — ١٩٣٦ (عمل إدارة الامتحانات بوزارة المعارف)

Connor, L.R., : *Statistics in Theory and Practice*, Chapter IX.

King, W.I., : *Elements of Statistical Method*, Chapter XI.

Mills, F.C., : *Statistical Methods*, Chapter III.

الباب السادس

المتوسطات

١١٩ — أتينا فى الأبواب السابقة بالطرق المتبعة فى جميع البيانات الإحصائية الأولية التى نبنى عليها بحثنا ، وكيفية تبويبها وتنظيمها لمساعدة الفكر على استيعابها ثم توضيحها بواسطة الأشكال الهندسية والرسوم البيانية . وقد توصلنا فى الباب السابق إلى وضع مختصر للمجموعات الإحصائية منظمة فى فئات ، يوضحها منحن تكرارى ، يساعد على إبراز بعض خواص هذه المجموعات .

يجب تعيين نموذج يمثل كل مجموعة

١٢٠ — ولكن هذه الصورة الأخيرة لا زالت ينقصها خطوة ضرورية تكملها ، ألا وهى البحث عن نموذج يمثل المجموعة الإحصائية ومفرداتها ، أو معيار تقاس بالنسبة إليه مفردات هذه المجموعة ، ، وتقارن بواسطته المجموعة كلها بالنسبة إلى المجموعات الإحصائية الأخرى .

فالمدرس لفرقة معينة من التلاميذ مثلا يراهم أمامه مختلفين فى الذكاء وفى المقدرة على التحصيل ، وربما أمكنه تقسيمهم إلى طبقات أو (فئات) من حيث ذكائهم وتحصيلهم . ولكنه زيادة على ذلك يحتاج إلى تعيين نموذج أو معيار لذكاء هذه الفرقة حتى يمكنه مقارنتها بغيرها من الفرق ، ولكى يمكنه أيضاً المفاضلة بين تلاميذ هذه الفرقة عينها لمعرفة مرتبة كل تلميذ بالنسبة إلى كل واحد من الآخرين ، وكذلك بالنسبة إلى النموذج العمومى أو المعيار الممثل للفرقة جميعها .

وكذلك صاحب المصنع وهو فى صدد تحديد الأجور لعماله على أساس إنتاجى مثلا يتفقد العمال فى أحد أقسام المصنع فيجدهم مختلفين فى الكفاية الإنتاجية : كل على حسب خبرته ومرانه وذكائه وصحته وغير ذلك من الظروف . فهو وإن أمكنه تقسيم عمال هذا القسم من المصنع إلى مراتب من حيث الكفاية وسرعة الإنتاج ، لا يزال محتاجاً إلى معرفة الكمية النموذجية أو المعيارية للانتاج فى هذا القسم ليقيس بها العمال فيه بالنسبة إلى بعضهم ، وليقارن بين هذا القسم فى مجموعه والأقسام الأخرى .

والطبيب أيضاً الذى يبحث مقدار ضغط الدم مثلا . فهو يختبر لذلك جماعة من الأشخاص ويقيس ضغط الدم عند كل واحد منهم ، فإذا به يختلف من شخص لآخر تبعاً لظروفه المختلفة مثل السن والحالة الصحية والتغذية والحالة العصبية وهكذا . وهو يحتاج فى هذا البحث إلى نموذج يمثل الجماعة لمقارنتهم بغيرهم أو بعضهم ببعض .

لايمكن مقارنة المجموعات بدون معرفة نماذج لها

١٢١ — وهكذا يمكننا تعداد هذه الأمثلة فى جميع نواحى الحياة العلمية أو العملية ، ونصل فى كلها إلى نفس النتيجة : ألا وهى أنه لايمكن مقارنة المجموعات الإحصائية بعضها ببعض مقارنة صحيحة بدون الالتجاء إلى مقارنتها بواسطة نماذج أو معايير تمثلها تمثيلا صحيحاً .

القيمة النموذجية لمجموعة تقع متوسطة بين طرفيها

وبما أن كل مجموعة من المقادير — كميات إنتاج عدد من العمال مثلا — تتكون من سلسلة متدرجة من الأرقام بعضها صغير والبعض أكبر ، فمن الواضح أن النموذج الذى يمثل هذه الكميات لا بد أن يكون واقعاً فى الوسط بين أصغر وأكبر قيمة . وإلا فقد صفة تمثيل المجموعة كلها . فهو إذن قيمة متوسطة بين طرفى المجموعة تتلخص فيه صفات مفرداتها وتتمثل فيها .

علم الإحصاء ليس علم

١٢٢ — والبحث فى هذه المتوسطات الإحصائية(١) وخواصها والأغراض التى تؤديها من أهم أبحاث هذا العلم ، ويشغل جزءاً كبيراً من اهتمام الباحث

---

(١) تسمى بالانجليزية Statistical Averages

المتوسطات فقط

الإحصائى فى أى مسألة يعالجها . وقد حدا ببعض الكتاب أن يجعل علم الإحصاء كله عبارة عن بحث فى المتوسطات دون سواها ، فعرفوه خطأ بأنه «علم المتوسطات». ولكن هذه التسمية قاصرة . إذ أن هذا البحث ما هو إلا ناحية من النواحى التى يتناولها العلم — حتى ولو كانت ناحية مهمة جداً .

تعريف المتوسط الإحصائى لمجموعة

١٢٣ — المتوسط الإحصائى لمجموعة من القيم التى تأخذها كمية متغيرة ، هو إذن عبارة عن قيمة تمثل هذه السلسلة من القيم أحسن تمثيل ، بحيث يمكن اتخاذها دليلا مميزاً لهذه المجموعة عن غيرها ، فنعرف بواسطتها الاتجاه الذى تأخذه هذه القيم فى مجموعتها . والغرض من استعماله فى أبحاثنا هو الاستغناء به عن استقراء مفردات المجموعة كلها ؛ لأن المفردات تتعرض بعضها إلى ظروف خاصة فتعطينا فكرة خاطئة عن المجموعة واتجاهها ؛ فضلا عن أن هذه الطريقة صعبة ومستحيلة عملياً فى الإحصاءات الكبيرة .

على أي أساس نختار المتوسط ليمثل المجموعة

١٢٤ — درجة اعتمادنا على المتوسط الإحصائى لمجموعة تتوقف على طريقة اختياره ، وتوفر صفات تمثيل المجموعة فيه ، وهى التى من أجلها اخترناه . وعلى ذلك فلا بد من التفكير فى الأساس الذى نبنى عليه اختيارنا للمتوسط ، حتى تتوفر فيه الشروط المطلوبة ، وهى صحة تمثيل المجموعة . ولو تأملنا فى ذلك من الناحية المنطقية أو الرياضية ، أو من الناحية العملية ، نجد أمامنا عدة أسس تبدو كلها معقولة ، وكل منها يصلح لأن يبنى عليه اختيار المتوسط الذى نبحث عنه . وهذه نذكرها فيما يلى .

الوسط الحسابى يصلح لأن يمثل المجموعة

١٢٥ — يمكن أن ننظر إلى صفة التمثيل أو النموذجية بمعنى أنها ضمان تعادل الزيادة فى بعض المفردات مع النقص فى المفردات الأخرى ، وتلافى أثر الصدف فى بعض الحالات دون الأخرى. ففى الرياضة مثلا نستخدم الوسط الحسابى(١)

(١) بالإنجليزية Arithmetic Mean

بين عدة كميات للتخلص من هذه التغيرات العرضية ، والحصول بذلك على قيمة متوسطة تمثل المجموعة الأصلية . وهذا على فرض أن الوسط الحسابى للقيم المختلفة التى يأخذها متغير هو القيمة الحقيقية لهذا المتغير . وهذا فرض معقول فى حد ذاته — وفعلا يمكن تبريره رياضياً فى بعض الحالات . وعلى هذا الأساس إذن يمكن اختيار الوسط الحسابى كأنه المتوسط المطلوب .

الوسط الهندسى يصلح أيضاً كمتوسط

١٢٦ — وغير الوسط الحسابى نستخدم أيضاً الوسط الهندسى[(١)] فى بعض الأحيان . والوسط الهندسى بين أى كميتين أو أكثر يقع أيضاً بين القيمة الكبرى والقيمة الصغرى . وعلى ذلك فيمكن اختياره أساساً لإيجاد المتوسط الإحصائى المطلوب .

الوسط التوافقى

وكذلك نرى الوسط التوافقى (سيأتى تعريفه) بين أى كميتين واقعاً بينهما ، فيمكن إذن أن يعتبر ممثلا لهما فى بعض الأحيان .

القيمة الأكثر شيوعاً تصلح لأن تمثل المجموعة

١٢٧ — إذا نظرنا إلى صفة النموذجية بمعنى أن نموذج المجموعة هو الشىء الأكثر شيوعاً فيها أو الذى يتكرر أكثر من غيره — وهذه وجهة نظر معقولة أيضاً — نصل من هذا الاعتبار إلى أساس آخر . فالقيمة التى نعتبرها « متوسطاً » على هذا الأساس هى القيمة الأكثر شيوعاً فى المجموعة ؛ أو التى تكرارها أكبر من تكرار أى قيمة أخرى فى المجموعة . والمتوسط — بهذا المعنى يسمى[(٢)] **المنوال** .

القيمة الوسطى أى الوسيط تصلح أيضاً

١٢٨ — يصح أيضاً أن نعتبر المتوسط الذى يمثل المجموعة هو القيمة الوسطى فيها ، بحيث إذا رتبت مفرداتها ، تصاعدياً أو تنازلياً ، كانت هى فى الوسط تماماً ،

---

(١) بالإنجليزية Geometric Mean

(٢) اسمه بالإنجليزية (Mode) . وأظن أن صاحب هذه الترجمة هو الأستاذ سليم أمين حداد بك أو الأستاذ على مصطفى مشرفة باشا . وأولهما أرجح .

وعدد المفردات التى قبلها يساوى عدد المفردات التى بعدها . وبذلك يكون عدد المفردات التى أكبر من القيمة الوسطى يساوى تماما عدد المفردات التى أصغر منها . ولا شك أن هذا اعتبار وجيه وله قيمته . والمتوسط بهذا المعنى يسمى بالإنجليزية (The Median) ، وأقترح تسميته « **الوسيط** » [1] أى أن الوسيط هو القيمة التى تقسم المجموعة إلى شطرين متكافئين من حيث العدد : الشطر الأول منها جميع مفرداته أصغر من الوسيط والشطر الثانى جميع مفرداته أكبر من الوسيط .

كل واحد من هذه المتوسطات له ميزة خاصة وهى تختلف فيما بينها إلا نادراً جداً

١٢٩ — نرى حينئذ أن هناك عدة اعتبارات يمكن اتخاذها أساساً لاختيار نوع المتوسط الذى نرمى إليه ، وكلها تؤدى إلى صفات لازمة يجب توافرها فى نوع المتوسط الذى نختاره . وحبذا لو اجتمعت هذه الصفات فى نوع واحد . ولكن هذا لايمكن — مع الأسف . وسنرى فيما بعد أن بعضاً منها فقط يمكن توافره فى متوسط واحد — وهذا فى حالة خاصة فقط : وهى حينما يكون التوزيع التكرارى للقيم متماثلا ، حيث يتساوى الوسط الحسابى والمنوال والوسيط ، وتجتمع صفاتها فى متوسط واحد .

طرق حساب المتوسطات

إيجاد الوسط الحسابى

١٣٠ — لإيجاد الوسط الحسابى لمجموعة من القيم ، نوجد حاصل جمع هذه القيم كلها ثم نقسمه على عددها . فالوسط الحسابى للأعمار :

٤٤ و ٤٢ و ٤١ و ٤٥ و ٤٨ و ٤٧ سنة ،

يساوى $\frac{٤٤ + ٤٢ + ٤١ + ٤٥ + ٤٨ + ٤٧}{٦} = \frac{٢٦٧}{٦}$

$=$ ٤٤٫٥ سنة .

---

(١) الفعل وسط الشىء بمعنى كان فى وسطه . واسم الفاعل منه واسط بمعنى كائن فى وسطه ؛ والصفة المشبهة « وسيط » تدل على حالة الثبوت .

ومن الواضح أنه يمكن اختصار هذا العمل نوعاً بأن نقول إنه

$$\text{يساوى } ٤٠ + \frac{٤ + ٢ + ١ + ٥ + ٨ + ٧}{٦} = ٤٠ + \frac{٢٧}{٦}$$

$$= ٤٤{,}٥ \text{ سنة .}$$

أى أننا نأخذ العدد ٤٠ مشتركا فى كل القيم ، ثم نبحث عن الوسط الحسابى للفروق بين الأعداد الأصلية وهذا الرقم المشترك ، أى

٤ و ٢ و ١ و ٥ و ٨ و ٧ ؛

ثم نضيف هذا الوسط الحسابى للفروق إلى العدد المشترك ٤٠ ، فينتج المتوسط المطلوب .

وواضح أيضاً أنه كان يمكن اختيار عدداً مشتركا غير ٤٠ هذا ، إذا وجدنا عدداً آخر أكثر مناسبة أو أسهل فى العمل . وعلى كل حال فيمكن إثبات هذه القضية بصفة عامة كما يأتى .

استخدام وسط فرضى

١٣١ — لنفرض أن مجموعة القيم عددها $ن$ ، وأن هذه القيم هى :

$$س_١ ، س_٢ ، س_٣ ، \cdots ، س_ن .$$

ولنرمز إلى الوسط الحسابى المطلوب إيجاده بالرمز $\overline{س}$. وعلى حسب تعريف الوسط الحسابى يكون :

$$\overline{س} = \frac{س_١ + س_٢ + س_٣ + \cdots س_ن}{ن} .$$

لتكن ا كمية اختيارية ، أياً كانت .

$$\therefore \overline{س} = ا + \frac{س_١ + س_٢ + \cdots + س_ن}{ن} - ا$$

$$= ا + \frac{س_١ + س_٢ + \cdots + س_ن - ن ا}{ن}$$

$$= ا + \frac{(س_١ - ا) + (س_٢ - ا) + (س_٣ - ا) + \cdots + (س_ن - ا)}{ن}$$

وذلك بتوزيع ن ا فى البسط على السينات (وعددها ن) كل واحدة نطرح منها ا . وعلى ذلك فالوسط الحسابى لأى عدد من الكميات ، يساوى أى كمية اختيارية مثل ا ، مضافاً إليها الوسط الحسابى للفروق بين الكميات الأصلية وهذه الكمية ا ، التى نسميها « **الوسط الفرضى** » .

ويمكن وضع هذه النتيجة فى صورة مختصرة جداً إذا وضعنا

$$مج\ س = س_١ + س_٢ + س_٣ + \cdots + س_ن ؛$$

$$ومج (س - ا) = (س_١ - ا) + (س_٢ - ا) + \cdots + (س_ن - ا) ؛$$

$$\therefore \bar{س} = \frac{مج\ س}{ن} = ا + \frac{مج (س - ا)}{ن}$$

وهذه العلاقة مفيدة جداً من الناحية العملية ، وتستخدم لتسهيل العمليات الحسابية .

١٣٢ — ننتقل الآن إلى شرح طريقة إيجاد الوسط الحسابى لمجموعة كبيرة من القيم منتظمة فى شكل توزيع تكرارى . لنأخذ مثلا التوزيع التكرارى الموجود فى جدول ٩ (صفحة ٩٩) ، ونوجد الوسط الحسابى للأعمار . إيجاد الوسط الحسابى للتوزيع التكرارى

نلاحظ مبدئياً أننا فى التوزيع التكرارى نعتبر كل المفردات التى فى فئة تكرارية واحدة جميعها متساوية ؛ وكل منها تساوى مركز الفئة (أنظر بند ٨٣) . وبناء على ذلك نعتبر مراكز الفئات هى الأعمار التى نريد إيجاد وسطها الحسابى ،

مع العلم بأن لكل منها تكراراً مخالفاً لتكرار الآخر . وعدد القيم يساوى مجموع التكرارات كلها ، وهو ١٧٣٩ .

جدول ٢٠ — إيجاد الوسط الحسابى لأعمار ١٧٣٩ تلميذاً .

فئة العمر بالسنين	مركز الفئة (العمر = س)	عدد التلاميذ التكرار = ك	حاصل الضرب س × ك
١٣٫٥ وأقل من ١٤٫٥	١٤	٣٤	٤٧٦
١٤٫٥ « ١٥٫٥	١٥	١٢٨	١٩٢٠
١٥٫٥ « ١٦٫٥	١٦	٢٦٢	٤١٩٢
١٦٫٥ « ١٧٫٥	١٧	٣٦٠	٦١٢٠
١٧٫٥ « ١٨٫٥	١٨	٣٨٦	٦٩٤٨
١٨٫٥ « ١٩٫٥	١٩	٢٩٤	٥٥٨٦
١٩٫٥ « ٢٠٫٥	٢٠	١٦٧	٣٣٤٠
٢٠٫٥ « ٢١٫٥	٢١	٩٢	١٩٣٢
٢١٫٥ « ٢٢٫٥	٢٢	١٦	٣٥٢
		١٧٣٩	٣٠٨٦٦

حسب تعريف الوسط الحسابى ، يجب أن نوجد حاصل جمع القيم أو الأعمار كلها وعددها ١٧٣٩ : منها العمر ١٤ مكرراً ٣٤ مرة ، والعمر ١٥ مكرراً ١٢٨ مرة ، والعمر ١٦ مكرراً ٢٦٢ مرة وهكذا .

وعلى ذلك يجب ضرب هذه الأعمار فى تكراراتهـا المناظرة لها حتى نحصل على حاصل الجمع الصحيح . وهذا نقسمه على مجموع التكرارات ، أى ١٧٣٩ ، وهو عدد القيم فى هذه الحالة .

∴ الوسط الحسابى $= \frac{٣٠٨٦٦}{١٧٣٩} =$ ١٧٫٧٤٩ سنة .

ويجب ملاحظة أننا لا نقسم على عدد القيم التى تكررت (وهى ١٤ و ١٥ و ١٦ و ٠٠ و ٢٢ ؛ أى ٩ قيم) ، بل نقسم على ١٧٣٩ وهو مجموع التكرارات .

إيجاد الوسط الحسابى باستخدام وسط فرضى

١٣٣ — الطريقة المتقدمة فى البند السابق متعبة نوعاً . ويمكن اختصارها وتسهيل العمليات الحسابية فيها باستخدام العلاقة التى أثبتناها فى بند ١٣١ ، وهى اختيار وسط فرضى ، وإيجاد الوسط الحسابى للفروق بين هذا الوسط الفرضى وبين الأعمار ، ثم إضافة هذا الوسط الحسابى للفروق إلى الوسط الفرضى فينتج الوسط الحسابى للأعمار نفسها ، وهو الذى نرمى إليه .

ويمكن شرح خطوات العمل فى جدول كالآتى :

جدول ٢١ — إيجاد الوسط الحسابى بالطريقة المختصرة باستعمال وسط فرضى

الفئات (بالسنين)	مركز الفئة (العمر = س)	عدد التلاميذ (التكرار = ك)	الانحرافات عن ١٨ (ح)	ضرب ح × ك
١٣٫٥ —	١٤	٣٤	—٤	—١٣٦
١٤٫٥ —	١٥	١٢٨	—٣	—٣٨٤
١٥٫٥ —	١٦	٢٦٢	—٢	—٥٢٤
١٦٫٥ —	١٧	٣٦٠	—١	—٣٦٠
١٧٫٥ —	١٨	٣٨٦	٠	٠
١٨٫٥ —	١٩	٢٩٤	١	٢٩٤
١٩٫٥ —	٢٠	١٦٧	٢	٣٣٤
٢٠٫٥ —	٢١	٩٢	٣	٢٧٦
٢١٫٥ —	٢٢	١٦	٤	٦٤
		١٧٣٩		—٤٣٦

نختار وسطاً فرضياً مناسبا ، وليكن العمر ١٨ ، أى مركز الفئة (١٧٫٥ و أقل من ١٨٫٥) . ويلاحظ أن تكرار هذه الفئة أكبر تكرار . وسنرى أن هذا الاختيار يساعدنا كثيراً فى تخفيف عمليات الضرب والعمليات الحسابية الأخرى . ولذلك يحسن دائما أن نختار الوسط الفرضى عند مركز الفئة ذات التكرار الأكبر ، أو فئة قريبة منها .

نحسب « انحرافات » الأعمار عن هذا الوسط الفرضى . وذلك بأن نطرح هذا الوسط من جميع الأعمار وهى مراكز الفئات فى الجدول ؛ ونحتفظ بالإشارة الجبرية أمام كل انحراف . نرمز إلى الانحراف بالحرف ح .

بما أن المطلوب الآن هو حساب الوسط الحسابى للانحراف ح ؛ وبما أن هذه الانحرافات (وهى — ٤ و — ٣ و — ... إلى ٣ و ٤) قد تكرر كل منها عدداً معيناً من المرات يخالف تكرارات الانحرافات الأخرى ،

∴ يجب أن نضرب كل انحراف فى عدد مرات تكراره ، وهو طبعاً يساوى تكرار القيمة الأصلية المناظرة له فى الجدول . ثم نجمع هذه الحواصل جمعاً جبرياً ، ونقسم الناتج على مجموع التكرارات ، وهو ١٧٣٩ ، ونضيف الناتج إضافة جبرية إلى الوسط الفرضى ، فنحصل على الوسط الحسابى المطلوب .

∴ الوسط الحسابى $= ١٨ + \frac{-٤٣٦}{١٧٣٩}$

$= ١٨ - ٠٫٢٥١$

$=$ ٧٤٩ و ١٧ سنة ؛

وهو نفس الناتج الذى حصلنا عليه من قبل .

والمشاهد أن اختيار الوسط الفرضى أمام الفئات غزيرة التكرار جعل الانحرافات الصغيرة تضرب فى هذه التكرارات الكبيرة . وهذا مما يساعد طبعاً

على تخفيف العمل الحسابى . ولا شك أن استخدام هـذه الطريقة أسهل بكثير من الطريقة المتقدمة فى البند السابق .

طريقـة تغيير الوحـدات

١٣٤ — فى بعض المسائل تكون فترات الفئات كبيرة المـدى ومتساوية فى الوقت نفسه . وفى هذه الأحوال يمكننا إجراء تسهيل آخر فى العمل الحسابى زيادة عن المتقدم . لنأخذ مثلا التوزيع التكرارى الآتى :

جدول ٢٢ — إيجاد الوسط الحسابى لأجور ٧٤٣٢ عاملا
باستعمال الطريقة المختصرة

فئات الأجـر (بالقرش)	مركز الفئة (الأجر=س)	عـدد العمال (التكرار=ك)	الانحراف عن ١٩٫٥ ح	الانحراف الجـديد ى	ضرب ك × ى
١٢ وأقل من ١٥	١٣٫٥	٢٢٤	—٦	—٢	—٤٤٨
١٥ « ١٨	١٦٫٥	١٩٧١	—٣	—١	—١٩٧١
١٨ « ٢١	١٩٫٥	٣٧٥٥	٠	٠	٠
٢١ « ٢٤	٢٢٫٥	١٢٣٦	٣	١	١٢٣٦
٢٤ « ٢٧	٢٥٫٥	١٩٦	٦	٢	٣٩٢
٢٧ « ٣٠	٢٨٫٥	٤٠	٩	٣	١٢٠
٣٠ « ٣٣	٣١٫٥	١٠	١٢	٤	٤٠
		٧٤٣٢			—٦٣١

لنأخذ فى هذا المثال الوسط الفرضى يساوى ١٩٫٥ ، وهو مركز الفئة الأكبر تكراراً . ثم نحسب انحرافات الأجور الأخرى عن هـذا الأجر ، فنجدها فى العمود ح عبارة عن مكررات العدد ٣ ؛ وهو يساوى طول فترة الفئة . فبدل أن نضرب هذه الانحرافات فى التكرارات ، يمكننا اختزالها نوعاً بقسمها جميعاً على العدد المشترك ٣ . فنحصل على أرقام أصغر وأسهل فى الضرب . وهذه نسميها

الانحرافات الجديدة ونرمز لها بالحرف ں . ويلاحظ أن الوحدة المقيسة بها هذه الانحرافات ليست قرشاً واحداً بل ثلاثة قروش . بدليل أن الانحراف فى السطر الأول يساوى — ٦ فى عمود ح ، بينما هو نفسه يساوى — ٢ فى عمود ف . وهكذا فى كل الانحرافات . أى أن ح = ٣ ف .

وتكرارات الانحرافات ف هى طبعاً نفس تكرارات الانحرافات القديمة ح . وعلى ذلك نضرب الانحرافات الجديدة ف فى التكرارات ك ، ونوجد المجموع الجبرى لها . ثم نقسم هذا المجموع على عدد العمال ، وهو مجموع التكرارات أى ٧٤٣٢ ، ينتج الوسط الحسابى للانحرافات ف ، وهو :

$$- \frac{٦٣١}{٧١٣٢} = - ٠٫٨٤٩ \quad \text{وحدة من وحدات ں}$$

∴ الوسط الحسابى للأجور بالقروش يساوى

$$١٩٫٥ + (- ٠٫٨٤٩ \times ٣) = ١٩٫٥ - ٢٫٥٤٧$$

$$= ١٩٫٢٥ \quad \text{قرشاً (تقريباً)}$$

ويجب ملاحظة هذه الخطوة الأخيرة ، وهى إرجاع الوسط الحسابى للانحرافات الجديدة ف إلى الوحدات الأصلية ، وذلك بضربه فى الرقم المشترك الذى قسمنا عليه عند الانتقال من الانحرافات القديمة ح إلى الانحرافات الجديدة ف .

ولكن هذا التسهيل الأخير المشروح أعلاه ، لا يفيدنا كثيراً فى حالة ما يكون الجدول التكرارى ذا فئات غير منتظمة ، كما نرى فى جدول ٨ (صفحة ٩٨) مثلاً ، حيث طول الفترة غير متساوٍ فى الفئات المختلفة ؛ حتى ولا نجد بين أطوال الفترات هناك عاملا مشتركا يمكن القسمة عليه كما فعلنا فى هذا المثال الأخير .

الوسط الحسابى لا يمكن إيجاده من الجداول المفتوحة

١٣٥ — فى الجدول التكرارى المفتوح (أنظر بند ٩١ «٣» صفحة ٩٥) لا يمكن إيجاد الوسط الحسابى . لأن الجدول المفتوح من أحد الطرفين أو كليهما يحتوى على فئة أو اثنتين لا يعرف مركزهما . وعلى ذلك لا يمكن إجراء خطوة ضرب مركز الفئة فى تكرارها لإيجاد الوسط الحسابى المطلوب . فى الجدول الآتى مثلا يمكننا ضرب تكرار الفئة الأولى فى صفر وتكرار الثانية فى مركزها وهو ٢,٥ ،

جدول ٢٣ — جدول تكرارى ، مفتوح من أعلى

لتوزيع عدد العمال فى مصانع القاهرة سنة ١٩٢٧

فئات عدد العمال فى المصنع	(عدد المصنع التكرار)
٠	٥٥٩١
من ١ إلى ٤	٨٨٥٢
من ٥ إلى ٩	١٤٤٨
١٠ فأكثر	٩٩٤
الجملة	١٦٨٨٥

والثالثة فى ٧ . ولكن لا نعلم فى أى مقدار نضرب تكرار الفئة الأخيرة لأننا لا نعلم حدها الأعلى . وهكذا لا يمكننا إيجاد الوسط الحسابى بدقة فى مثل هذه الحالة .

ولكن هذا لا يمنعنا من أن نفرض حداً أعلى لهذه الفئة الأخيرة يكون مناسبا حسب ما يتراءى لنا من خبرتنا وإلمامنا بظروف هذه المسألة — إذا كان لدينا هذا الإلمام . وبذلك يمكن تحديد مركز هذه الفئة على وجه التقريب ، وبالتالى إيجاد الوسط الحسابى على وجه التقريب أيضاً .

ويصح فى بعض الأحيان أن نهمل الفئة المفتوحة بالمرة إذا كان تكرارها ضئيلا بالنسبة لمجموع التكرارات ، ولا يؤثر هذا الحذف تأثيراً كبيراً فى النتيجة . وبذلك يكون لدينا الفئات الباقية جميعها محددة ، فنحسب الوسط الحسابى على أساسها . والنتيجة هنا تقريبية أيضاً ، وتتوقف درجة الدقة فيها على مقدار الخطأ الذى يترتب على إهمال الفئات المفتوحة .

وربما يكون الأسلم فى مثل هذه الأحوال أن نترك الوسط الحسابى ونحصل على المتوسط أو النموذج المطلوب عن طريق المتوسطات الأخرى التى لا تحتاج إلى معرفة حدود الفئات المتطرفة وهذه هى الوسيط والمنوال كما سيأتى بعد .

إيجاد المنوال

١٣٦ — نشرح الآن بعض الطرق المختلفة التى نستخدمها لإيجاد المنوال . ولنأخذ نفس التوزيع التكرارى لأعمار ١٧٣٩ تلميذاً المبين فى جدول ٢٠ (صحيفة ١٤٤) ؛ ونوجد العمر المنوالى لهذه المجموعة بهذه الطرق .

توزيع أعمار ١٧٣٩ تلميذاً

الفئــات	مراكز الفئات	التكرارات
١٣٫٥ —	١٤	٣٤
١٤٫٥ —	١٥	١٢٨
١٥٫٥ —	١٦	٢٦٢
١٦٫٥ —	١٧	٣٦٠
١٧٫٥ —	١٨	٣٨٦
١٨٫٥ —	١٩	٢٩٤
١٩٫٥ —	٢٠	١٦٧
٢٠٫٥ —	٢١	٩٢
٢١٫٥ —	٢٢	١٦

المنوال يساوى مركز الفئة المنوالية تقريباً

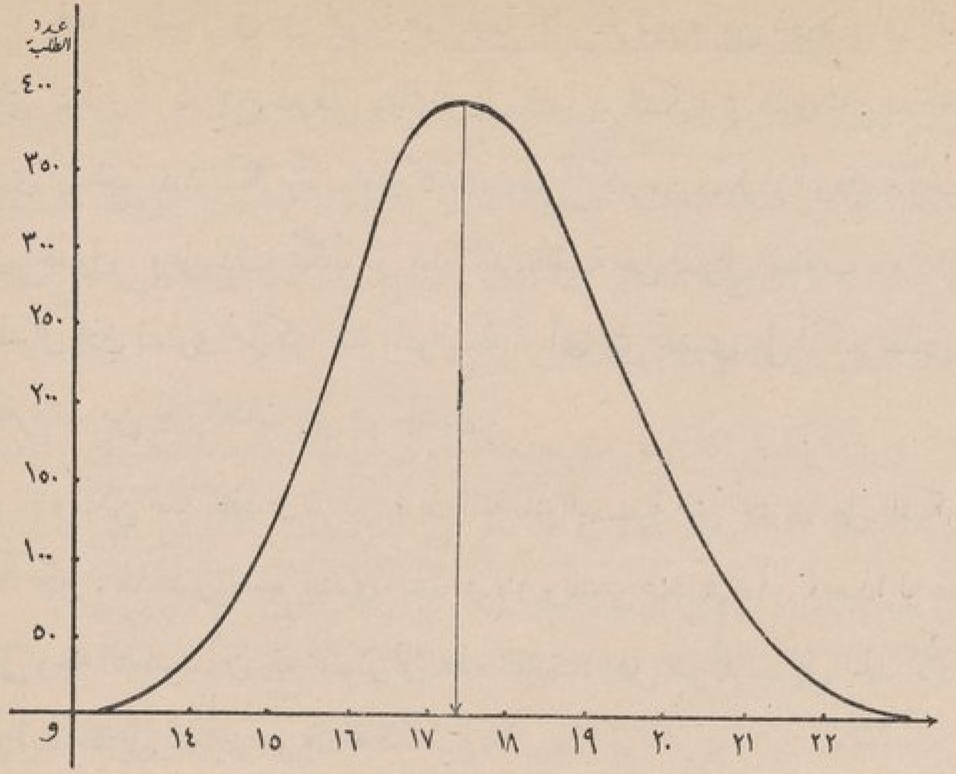
قلنا إن المنوال فى أى مجموعة هو القيمة الأكثر شيوعاً من غيرها ، أى القيمة التى تتكرر أكثر من غيرها . وبالنظر إلى التوزيع التكرارى الذى نحن بصدده ، نرى أن العمر ١٨ تكراره يساوى ٣٨٦، وهو أكبر من تكرار أى قيمة أخرى فى الجدول . وعلى ذلك يمكننا أن نعتبر هذه القيمة هى المنوال المطلوب . ويكون المنوال إذن يساوى مركز الفئة المنوالية . أى التى تحتوى على أكبر عدد من المفردات من بين الفئات التى فى الجدول .

ولكن هذا التقدير تقريبى ، إذ الفئة المنوالية التى تحتوى على التكرار الأكبر ، لها مدى واسع يبتدىء عند ١٧,٥ وينتهى عند ١٨,٥ . وحبذا لو علمنا على وجه التحديد أين يقع المنوال فى هذه الفترة : هل هو أقرب إلى الحد الأدنى أو الحد الأعلى ؟ وأين هو بالدقة ؟

ولتحديد موقع المنوال بالدقة نلجأ إلى طرق أخرى نشرحها فيما يلى : —

إيجاد المنوال بالرسم

١٣٧ — لو رسمنا المنحنى التكرارى الممهد لهذا التوزيع لاحظنا أنه يصعد بالتدريج حتى يبلغ ذروته ثم يهبط بعد ذلك . وبما أن الإحداثى الرأسى لأى نقطة على منحن تكرارى يدل على تكرار القيمة التى يمثلها الإحداثى الأفقى لنفس النقطة ، يتضح لنا أن ارتفاع قمة المنحنى عن المحور الأفقى ، وهو الإحداثى الرأسى لها ، يمثل تكرار القيمة التى يمثلها الإحداثى الأفقى للقمة . فإذا أسقطنا عموداً من قمة المنحنى على المحور الأفقى نعرف الإحداثى الأفقى للقمة ، وهو يساوى القيمة التى تكرارها أكبر من تكرار أى قيمة أخرى . وهى المنوال المطلوب . ونرى فى (شكل ٥٦) المنحنى التكرارى للتوزيع . ومنه نجد أن الإحداثى الأفقى للقمة يساوى البعد و ا وهو يساوى ١٧و٤ على مقياس المحور الأفقى . وهذا البعد يساوى المنوال المطلوب .



(شكل ٥٦)
تعيين المنوال بالرسم من المنحنى التكرارى

والنتيجة فى هذه الحالة لا بد تتوقف على الرسم وكيفية تمهيده ، لأن موقع القمة فى الشكل يتعين بعد التمهيد . ولذلك لا نضمن دقة هذه الطريقة إلا بعد التأكد من دقة الرسم .

وحبذا لو علمنا المعادلة الرياضية لهذا المنحنى ؛ إذ يمكن بواسطتها حساب قيمة المنوال بأحسن ما يمكن من الدقة . ولكن البحث فى إيجاد هذه المعادلة بحث شاق ومرهق ؛ علاوة على أنه يقتضى إلماماً ببعض القواعد والنظريات الرياضية ، الخاصة بمسألة توفيق المنحنيات التى شرحناها بالاختصار فى آخر الباب الرابع .

إيجاد المنوال بالحساب

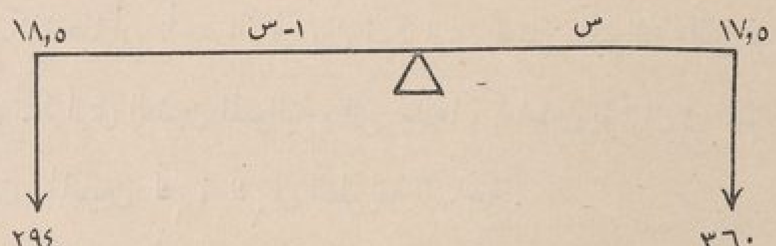
١٣٨ — توجد طريقة حسابية ، تقريبية نوعاً ، يمكن استخدامها لإيجاد المنوال . وهى تنبنى على استخدام تكرارى الفئتين المجاورتين للفئة المنوالية ،

من الجانبين العلوى والسفلى لها ، كوسيلة لتعرف موقع المنوال داخل الفترة . فى هذا المثال نجد تكرارات الفئة المنوالية والفئتين المحيطتين بها هى :

الفئـــة	التكرار
١٦,٥ وأقل من ١٧,٥	٣٦٠
١٧,٥ » ١٨,٥	٣٨٦
١٨,٥ » ١٩,٥	٢٩٤

فالمنوال إذن يقع بين ١٧,٥ و ١٨,٥ . ليكن ١٧,٥ + س . ويتعين المنوال تماماً إذا عرفنا قيمة هذا المجهول س . ولمعرفة قيمـة س ، نقول إن التكرارين ٣٦٠ و ٢٩٤ أشـــبه بقوتين تؤثران فى المنوال وتتجاذبانه بحيث إن الأقوى منهما تجعـله أدنى إلى حد الفئة المنواليـة القريب منها ، مع وجوده دائماً داخل فترة الفئة المنواليّة ولا يتعداها . ويكون موقع المنوال فى هذه الفـترة إذن بحيث يتعادل تأثير القوتين مع بعضهما .

وعلى ذلك يكون مدى الفترة (وطوله يساوى سنة واحدة) يشبه ساقاً طولها يساوى ١ (أى طول الفترة) ، علق فى طرفيه ثقلان : أحدهما يساوى ٣٦٠ ، والآخر ٢٩٤ . والمطلوب تعيين محور الارتكاز لهذه الرافعة حتى تتعادل هاتان القوتان .



وعلى ذلك لحساب قيمة س نقول :

٣٦٠ × س = ٢٩٤ × (١ — س) (١)

٣٦٠ س = ٢٩٤ - ٢٩٤ س

س (٣٦٠+٢٩٤) = ٢٩٤

$$\therefore \quad س = ١ \times \frac{٢٩٤}{٣٦٠ + ٢٩٤} \qquad (٢)$$

$$= ٠٫٤٥٠ \quad \text{تقريباً}$$

$$\therefore \quad \text{يكون المنوال} = ١٧٫٥ + ٠٫٤٥$$

$$= ١٧٫٩٥ \quad \text{سنة .}$$

ولو كان مدى الفترة غير الواحد الصحيح ، كنا نضع مقداره (١٢ شهراً مثلا) بدل ١ فى الطرف الأيسر للمعادلتين (١) و (٢) .

وعلى العموم ، إذا كان تكرار الفئة قبل المنوالية يساوى $ك_١$ ، وتكرار الفئة بعد المنوالية يساوى $ك_٢$ ، وكان طول الفترة فى الفئة المنوالية يساوى ل ، والحد الأدنى لها يساوى ا ، فإن المنوال المطلوب

$$= ا + ل \times \frac{ك_٢}{ك_١ + ك_٢} .$$

وهذه الطريقة تقريبية كما قلنا ، ولكنها سهلة من الناحية العملية ، وبسيطة من الناحية النظرية .

١٣٩ — وقد اقترح بيرسون (Karl Pearson) طريقة أخرى لحساب المنوال ، تعتمد أيضاً على تكرارات الفئات الثلاثة ، المنوالية والمحيطتين بها . طريقة الفروق كارل بيرسون

والفكرة هنا أن نأخذ الفرق بين تكرارى الفئتين المنوالية والتى قبلها ، أمام الفرق بين تكرارى الفئتين المنوالية والتى بعدها ، كعاملين يؤثران فى المنوال ، بدل التكرارين الجانبيين $ك_١$ ، $ك_٢$ فى الطريقة السابقة .

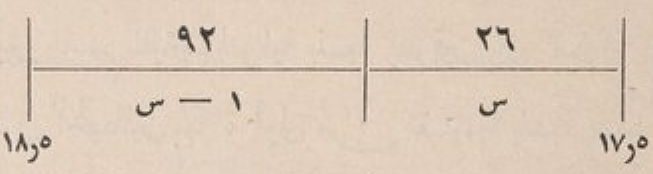
وهذا على اعتبار أن التغير فى التكرار من فئة إلى الفئة التى تليها ، يكون دليلا أدق على قرب أو بعد المنوال ، من نفس تكرارى الفئتين المجاورتين للمنوالية.

وبتطبيق هذه القاعدة على المثال الذى نحن بصدده نجد :

الفئــة	التكرار	الفرق
١٦,٥ وأقل من ١٧,٥	٣٦٠	
		٢٦
١٧,٥ « ١٨,٥	٣٨٦	
		٩٢
١٨,٥ « ١٩,٥	٢٩٤	

ويظهر فى العمود الأخير من هذا الجدول الفروق بين تكرار الفئة المنوالية وتكرارى الفئتين المحيطتين بها من الجانبين .

وهنا يكون المنـوال واقعاً بين حدى الفئة المنواليـــة أيضاً ، وليكن ١٧,٥ + س ، وتعيين قيمة س يكون بتقسيم المسافة بين ١٧,٥ و ١٨,٥ بنسبة ٢٦ : ٩٢



$$\therefore \quad \text{س} \times ٩٢ = ٢٦\,(١ - \text{س})\,؛$$

$$\text{أى أن} \quad \text{س} = ١ \times \frac{٢٦}{٢٦ + ٩٢}\,،$$

$$= ,٢٢$$

$$\therefore \quad \text{المنوال} = ١٧,٥ + ٠,٢٢$$

$$= ١٧,٧٢ \ \text{سنة} .$$

وعلى العموم : إذا كان تكرار الفئة المنوالية يساوى ك ، وتكرار التى قبلها $\text{ك}_١$ ، وتكرار التى بعدها $\text{ك}_٢$ ؛ وكان طول الفترة المنوالية يساوى ل ، والحد الأدنى لها يساوى ا ، فعلى أساس هذه الطريقة :

$$\text{المنوال} = \text{ا} + \text{ل} \times \frac{\text{ك} - \text{ك}_١}{(\text{ك} - \text{ك}_١) + (\text{ك} - \text{ك}_٢)}$$

$$= \text{ا} + \text{ل} \times \frac{\text{ك} - \text{ك}_١}{٢\,\text{ك} - \text{ك}_١ - \text{ك}_٢}$$

طريقة بيرسون أدق من السابقة

١٤٠ — ولو تأملنا فى هذا ، وجدنا أن الفكرة الأساسية فى طريقة بيرسون أقرب للصواب من فكرة الطريقة السابقة لها . وذلك لأنه إذا كان الفرق بين تكرارى فئتين متتاليتين صغيرا ، وكان المنوال واقعا فى أحدهما ، فأنه يكون أقرب إلى الحد المشترك بين هاتين الفئتين مما لو كان الفرق بين تكراريهما كبيراً .

و بعبارة أخرى : إذا كان الفرق صغيراً بين تكرار فئة منواليّة وتكرار فئة مجاورة لها ، فهذا معناه أن تكرار هـذه الفئة المجاورة أقرب من النهاية العظمى للتكرارات (وهى تساوى تكرار المنوال) مما لو كان هذا الفرق عظيماً . وعلى ذلك فالمنوال يكون أقرب إلى هذه الفئة المجاورة مما لو كان الفرق بين تكرارها وتكرار الفئة المنوالية كبيراً .

وهـذا يبرر تقسيم الفترة المنوالية بنسبة الفرقين بين تكرار الفئة المنواليـة وتكرارى الفئتين المحيطتين بها ، أولى من تقسيمها بنسبة تعتمد على تكرارى هاتين الفئتين .

وليس غريباً أن نحصل على نتيجتين مختلفتين لقيمة المنوال بهاتين الطريقتين ، وهذا هو المنتظر إذ أنهما على أساسين مختلفين . والواقع كما قلنا ، أن كليهما تقريبيان ، وواحدة أحسن من الأخرى . فليس معنى الاختلاف أن إحداهما خطأ والأخرى صحيحة ، بل الواحدة أقرب إلى الصواب من الأخرى .

إيجاد المنوال بتعديل الفترات

١٤١ — إذا نحن غيرنا فترات الفئـات فى أى توزيع تكرارى عادى ، وعدلنا مواقع حدودها ، فإن هذا يحدث تغييرا فى التكرارات ؛ وفى الغالب يحدث تغيـيراً فى موقع الفئة المنواليـة . ويمكن استخدام هـذه الظاهرة فى البحث عن المنوال وذلك ، وذلك بأن نغير مواقع الفئات وفتراتها ، ونحصل على عدة جداول

تكرارية مختلفة لنفس المجموعة من القيم . وكل جدول يظهر فيه فئة ذات تكرار أكبر يقع فيها المنوال . وبمقارنة الفئات المنوالية وفتراتها التى نحصل عليها من الجداول التكرارية المختلفة يمكن الاهتداء إلى موقع المنوال ، ويكون إذ ذاك فى المنطقة المشتركة بين فترات الفئات المنوالية للجداول التكرارية المختلفة .

لنأخذ مثلاً توزيع الأعمار بين الناجحين فى شهادة الدراسة الثانوية قسم ثان سنة ١٩٣٦ ، وهو كما يأتى :

جدول ٢٤ — توزيع أعمار ٢٢٠٢ طالباً

فى فئات مختلفة المدى

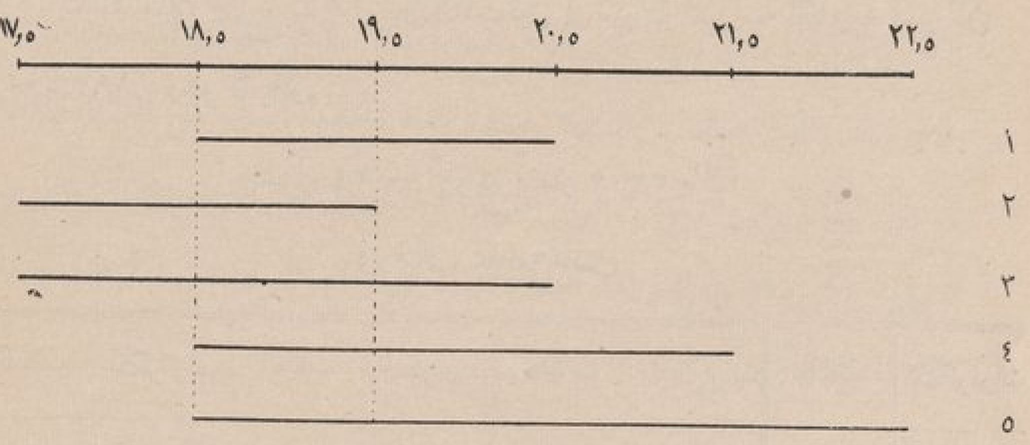
الفئات	التكرارات	الفئات	التكرارات	الفئات	التكرارات	الفئات	التكرارات
١٤٫٥ —	١١٤	١٤٫٥ —	١٥	١٤٫٥ —	٤٠٦	١٤٫٥ —	١٥
١٦٫٥ —	٦٩٠	١٥٫٥ —	٣٩١	١٧٫٥ —	١١٧٣	١٥٫٥ —	٧٨٩
١٨٫٥ —	٧٧٥	١٧٫٥ —	٨٤٠	٢٠٫٥ —	٥٠٢	١٨٫٥ —	١٠٣٥
٢٠٫٥ —	٤٢٤	١٩٫٥ —	٥٩٣	٢٣٫٥ —	٩٨	٢١٫٥ —	٢٩٨
٢٢٫٥ —	١٣٤	٢١٫٥ —	٢٤٢	٢٦٫٥ —	١٤	٢٤٫٥ —	٤٨
٢٤٫٥ —	٤٢	٢٣٫٥ —	٨٠	٢٩٫٥ —	٩	٢٧٫٥ —	١٧
٢٦٫٥ —	١١	٢٥٫٥ —	٢٤			٣٠٫٥ —	٠
٢٨٫٥ —	٣	٢٧٫٥ —	٨				
٣٠٫٥ —	٩	٢٩٫٥ —	٩				
	٢٢٠٢		٢٢٠٢		٢٢٠٢		٢٢٠٢

الفئات	التكرارات
١٤٫٥ —	٨٠٤
١٨٫٥ —	١١٩٩
٢٢٫٥ —	١٧٦
٢٦٫٥ —	١٤
٣٠٫٥ —	٩
	٢٢٠٢

ومن هذه الجداول التكرارية نرى أن المنوال واقع فى الوقت نفسه فى خمس مناطق مختلفة ، وهى على الترتيب :

١٨٫٥ — ٢٠٫٥ ؛ ١٧٫٥ — ١٩٫٥ ؛ ١٧٫٥ — ٢٠٫٥ ؛ ١٨٫٥ — ٢١٫٥ ؛ ١٨٫٥ — ٢٢٫٥ .

وإذا وضعنا هذه المناطق فوق بعضها فى نظام كالآتى :



نرى بوضوح أن المنطقة المشتركة بينها جميعاً هى ١٨٫٥ — ١٩٫٥ ، فلا بد أن يكون المنوال داخل هذه المنطقة . ويلاحظ أن مدى هذه المنطقة أقل من مدى أى منطقة من المناطق المشتركة . وعلى ذلك قد حصلنا على موقع للمنوال أكثر تحديداً من المواقع التى عرفناها بواسطة أى جدول تكرارى على حدته . وبالطبع يمكننا تحديد الموقع أكثر من ذلك إذا كانت البيانات التى لدينا عن أعمار هؤلاء الطلبة مفصلة تفصيلاً أكثر يسمح بعمل فئات أضيق .

تعدد المناويل

١٤٢ — إذا كان المنحنى التكرارى متعدد القمم [١] ، فمعنى ذلك أن له أكثر من منوال واحد . ويمكننا إيجاد كل منها باستخدام تكرارى الفئتين المجاورتين لكل منوال بالطرق التى شرحناها فى بند ١٣٦ أو ١٣٨ ، ولكن على كل حال فالمنوال لمثل هذا المنحنى لا يكون له فائدة كبيرة من حيث تمثيل المجموعة ، حيث قد ذكرنا سابقاً أن مثل هذه المجموعة تكون غير متجانسة ، وفى الغالب تكون

(١) يسمى بالإنجليزية Multimodal أو Bimodal إذا كان ذا قمتين

خليطاً بين مجموعتين مختلفتين أو أكثر . وعلى ذلك فمعنى المنوال الواحد لهذه المجموعة يكون بعيداً عن المقصود من دراسة المتوسطات . والأفضل أن نوجد المناويل المختلفة فى مثل هذا المنحنى ، وهى فى الواقع تساوى (بالتقريب) المناويل الأصلية للمجموعات الجزئية المركبة لهذه المجموعة الكبيرة .

اختلاف قيم المنوال المحسوبة بالطرق المختلفة

١٤٣ — شرحنا الآن عدة طرق لإيجاد المنوال ، وقد رأينا أن النتائج التى نحصل عليها تخالف بعضها أحياناً . ويجب ألا نأخذ هذا دليلا على التناقض أو غلط بعض الطرق وصحة الأخرى . فكل هذه الطرق عبارة عن طرق تقريب للوصول إلى القيمة الحقيقية للمنوال . والتقريب معناه أن هناك بعض الخطأ ولكنه يهمل لعدم أهميته . والخطأ إما أن يكون زيادة أو نقصا . وهذا هو السر فى اختلاف النتائج فى حساب المنوال بهذه الطرق المختلفة .

ويلاحظ أننا لم نشاهد مثل هذه الاختلافات فى إيجاد الوسط الحسابى بالطرق المختلفة التى شرحناها . وذلك لأن الوسط الحسابى ذو معنى محدود فهو مجموع القيم التى لدينا مقسوماً على عددها . والقيم معلومة عندنا وكذلك عددها معلوم . أما فى حالة المنوال — وهى القيمة التى تكرارها أكبر من تكرار أى قيمة أخرى — فنحن نجهل القيمة ونجهل هذا التكرار الأكبر . ومن ثم كان عدم التحديد والالتجاء إلى الطرق المختلفة للتوصل إلى الغرض . ولو علمنا هذا التكرار لأمكننا تحديد القيمة مباشرة ، بدون الالتجاء إلى الطرق المساعدة والوقوع فى أخطائها .

وهذا الغموض وهذه الصعوبة فى حساب المنوال مما يعيبه ويجعله أقل فائدة واستعمالا من الوسط الحسابى .

إيجاد الوسيط لمجموعة من القيم

١٤٤ — طريقة إيجاد الوسيط لمجموعة من القيم تنبنى على ترتيبها تصاعدياً (أو تنازلياً) والقيمة الوسطى ، التى يسبقها ويليها عددان متساويان من القيم ، هى الوسيط .

إذا كان عدد القيم صغيراً فمن السهل ترتيبها تصاعدياً ومعرفة الوسيط ؛ وهو القيمة الوسطى التى ترتيبها يساوى ½ (ﻫ + ١) ، حيث ﻫ هى عدد القيم جميعها . فالعمر الوسيط للمجموعة الآتية مثلا :

١٧   ١٨٫٥   ٢٠   ٢١   ٢١٫٥   ٢٢   ٢٣٫٥

هو ٢١ سنة ؛ وهى القيمة الرابعة فى هذه السلسلة التصاعدية ، أى التى ترتيبها ½ (٧ + ١) ، حيث ٧ تمثل عدد القيم الموجودة .

و إذا كان عدد القيم زوجياً ، فلا نجد قيمة وسطى يحيط بها عددان متساويان من القيم ، بل نجد قيمتين فى الوسط ، كما فى مجموعة الأجور الآتية مثلا :

٧٫٥   ٩   ١٠٫٥   ١١   ١٢   ١٣٫٥   ١٤   ١٥   قرشاً ،

حيث القيمتان الرابعة والخامسة فى السلسلة ، أى ١١ و ١٢ قرشاً ، يسبقهما ثلاث قيم أصغر منهما ، و يليهما ثلاث أخر أكبر منهما . ففى مثل هذه الحالة نعتبر الوسط الحسابى بينهما هو الوسيط ، كما لو كانت هناك قيمة وهمية واقعة بين القيمتين الرابعة والخامسة ، وترتيبها ٤ ½ ، أى ½ (٨ + ١) ، بوضع ٨ بدل ﻫ فى العبارة ½ (ﻫ + ١) المذكورة أعلاه . وهذه العبارة نستخدمها عملياً لمعرفة ترتيب الوسيط عندما تكون مفردات القيم معلومة وعددها ﻫ صغيراً ومرتبة تصاعدياً ، أو تنازلياً فيما بينها . أما من الناحية النظرية فإن ترتيب الوسيط لمفردات عددها ﻫ هو ½ ﻫ وليس ½ (ﻫ + ١) . وعلى كل حال فالفرق طفيف لا يؤثر كثيراً خصوصاً إذا كانت ﻫ عدداً كبيراً . و بناء على ذلك نستخدم دائماً ½ ﻫ فى إيجاد ترتيب الوسيط .

إيجاد الوسيط لتوزيع تكرارى كبير

١٤٥ — إذا كان عدد القيم فى مجموعة ما كبيراً ، ونظمناها فى فئات على شكل جدول تكرارى ، يمكن ترتيبها تصاعدياً فى جدول تكرارى متجمع يضع الفئات وراء بعضها ، و يبين تراتيب القيم فى السلسلة . و بواسطة هذا التكرار

المتجمع يمكن معرفة قيمة الوسيط بعد معرفة ترتيبه . ولتوضيح ذلك نأخذ التوزيع التكرارى السابق ، ونوجد التكرار المتجمع كما شرحنا فى بند ١٠٧ صفحة ١١٨

جدول ٢٥ — التكرارات المتجمعة لأعمار ١٧٣٩ تلميذاً :

فئات العمر بالسنين	عـدد التلاميذ (التكرار)	الحدود العليا للفئات	التكرار المتجمع (الصاعد)
١٣،٥ أقل من ١٤،٥	٣٤	أقل من ١٤،٥	٣٤
١٤،٥ « ١٥،٥	١٢٨	« ١٥،٥	١٦٢
١٥،٥ « ١٦،٥	٢٦٢	« ١٦،٥	٤٢٤
١٦،٥ « ١٧،٥	٣٦٠	« ١٧،٥	٧٨٤
١٧،٥ « ١٨،٥	٣٨٦	« ١٨،٥	١١٧٠
١٨،٥ « ١٩،٥	٢٩٤	« ١٩،٥	١٤٦٤
١٩،٥ « ٢٠،٥	١٦٧	« ٢٠،٥	١٦٣١
٢٠،٥ « ٢١،٥	٩٢	« ٢١،٥	١٧٢٣
٢١،٥ « ٢٢،٥	١٦	« ٢٢،٥	١٧٣٩
الجملة	١٧٣٩		

ومن هذا الجدول نرى أن ال ٣٤ قيمة الأولى جميعها أقل من ١٤،٥ سنة ؛ وال ١٦٢ قيمة الأولى كلها أقل من ١٥،٥ . وهكذا .

وطبقاً لتعريف الوسيط ، يكون ترتيبه فى هـذه السلسلة يساوى $\frac{1}{2}$ (١٧٣٩) = ٨٦٩ $\frac{1}{2}$ ، أى بوضع ن = ١٧٣٩ فى العبارة $\frac{1}{2}$ ن .

الآن نبحث عن الوسيط وهو القيمة التى ترتيبها ٨٦٩ $\frac{1}{2}$ فى هـذه السلسلة .

وبالتأمل فى الجدول التكرارى المتجمع نجد أن الـ ٧٨٤ قيمة الأولى كلها أقل من ١٧,٥ سنة ، فى حين أن الـ ١١٧٠ قيمة الأولى جميعها أقل من ١٨,٥ سنة .

∴ العمر الوسيط ، وترتيبه ٨٦٩,٥ ، لا بد واقع بين ١٧,٥ و ١٨,٥ .

ليكن ١٧,٥ + س . ويبقى إذن معرفة قيمة المجهول س .

نقول إن فرقاً فى الترتيب من ٧٨٤ إلى ١١٧٠ يقابله فرق فى القيمة من ١٧,٥ إلى ١٨,٥ ،

∴ فرق فى الترتيب من ٧٨٤ إلى ٨٦٩,٥ يقابله فرق فى القيمة قدره س .

وهذه عملية تناسب[1] بسيطة نضعها على الصورة الآتية :

$$\frac{١١٧٠ - ٧٨٤}{٨٦٩,٥ - ٧٨٤} = \frac{١٨,٥ - ١٧,٥}{س} \qquad (١)$$

$$∴ \quad س = ١ \times \frac{٨٦٩,٥ - ٧٨٤}{١١٧٠ - ٧٨٤} \qquad (٢)$$

$$= ٠,٢٢٢ \text{ تقريباً}$$

$$∴ \quad \text{الوسيط} = ١٧,٥ + ٠,٢٢٢$$

$$= ١٧,٧٢٢ \text{ سنة .}$$

ولو كانت فترة الفئة غير ١ ، أى عدداً آخر مثل ١٢ شهراً ، كنا نضع هذا العدد بدل ١ فى المعادلة (٢) .

بالصدفة كان عدد القيم فى هذا المثال فردياً (= ١٧٣٩) ، ولذلك كان ترتيب الوسيط عدداً كسرياً (= ٨٦٩,٥) . ولكن إذا كان العدد الأصلى زوجياً فإن ترتيب الوسيط يكون عدداً صحيحاً . مثلاً ٨٧٠ إذا كان عدد المفردات

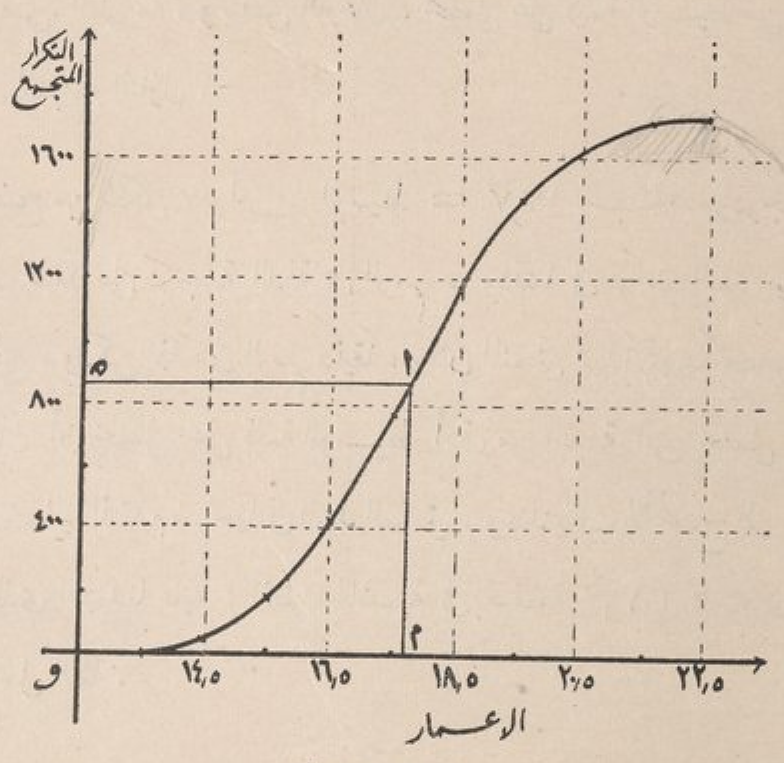
---

(١) هذا التناسب على فرض أن المفردات الموجودة فى الفئة ١٧,٥ — ١٨,٥ التى يقع فيها الوسيط ، متدرجة بانتظام داخل هذه الفترة . وهذا فرض عملى مناسب ، وفيه شىء من التقريب طبعاً . ولكن الخطأ فى هذا صغير ، خصوصاً إذا كانت فترة الفئة صغيرة .

١٧٤٠ بدل ١٧٣٩ ، ولكن هذا لا يؤثر فى الطريقة ، حيث نضع هذا الترتيب ٨٧٠ بدل الترتيب ٨٦٩,٥ فى المعادلتين (١) و (٢) .

إيجاد الوسيط بالرسم

١٤٦ — يمكن الاستعانة بالمنحنى التكرارى المتجمع فى معرفة قيمة الوسيط . بما أن الوسيط هو القيمة التى يسبقها نصف عدد القيم ويليها النصف الآخر ؛ وبما أن كل نقطة على المنحنى التكرارى المتجمع الصاعد يكون إحداثيها الرأسى يساوى عدد القيم السابقة فى الترتيب (والأقل فى المقدار) للقيمة التى يمثلها الإحداثى الأفقى لهذه النقطة (كما قلنا فى بند ١٠٧) .



(شكل ٥٧)

تعيين الوسيط بالرسم

∴ نرسم المنحنى الصاعد لهذا التوزيع ، شكل ٥٧ ، ونعين نقطة عليه ، ا مثلا ، إحداثيها الرأسى يساوى نصف عدد القيم أى $\frac{1}{2}$ × ١٧٣٩ .

والإحداثى الأفقى لهذه النقطة ا ، وهو و م فى الشكل ، يساوى الوسيط المطلوب . لأنه عبارة عن قيمة يوجد قبلها قيم أصغر منها ، عددها نصف عدد القيم كلها .

لذلك نأخذ البعد و ى على المحور الرأسى يساوى ½ × ١٧٣٩ ؛ ونرسم من ى موازياً لمحور السينات فيقطع المنحنى فى نقطة ا . نسقط منها العمود ا م على المحور الأفقى ، فيكون البعد و م ، مقيساً بوحدات المحور الأفقى ، يساوى الوسيط المطلوب .

و بنفس الطريقة ، و بنفس البرهان ، نحصل على قيمة الوسيط من المنحنى التكرارى المتجمع النازل .

ويتضح من شكل ٥٧ أن الوسيط = ١٧,٧ سنة . ودرجة الدقة هنا تتوقف طبعاً على درجة الدقة فى الرسم ، ودقتنا فى قراءة البعد و م على المحور الأفقى . ولكن إذا كان الرسم دقيقاً ، وكان المنحنى ممهداً تمهيداً مضبوطاً ، ففى الإمكان أن نحصل على قيمة للوسيط أدق من القيمة التى نحصل عليها بالطريقة الحسابية المتقدم شرحها فى البند السابق . وعند ذلك لا نحتاج إلى الفرض التقريبى الذى اعتمدنا عليه (أنظر الحاشية فى صفحة ١٦٢) فى تحرى موقع الوسيط داخل الفئة .

الربيعان الأدنى والأعلى

١٤٧ — نحتاج أحياناً لمعرفة القيمة التى تقع عند ربع السلسلة الصاعدة أو النازلة لمجموعة القيم التى عندنا . أى أن ترتيبها فى السلسلة الصاعدة يساوى ربع عدد القيم أو ثلاثة أرباعه على الترتيب . ومعرفة هاتين القيمتين تعطينا فكرة مفيدة عن كيفية توزيع القيم فى السلسلة .

وأقترح تسميتها [1] **الربيع الأدنى والربيع الأعلى**

فلإيجاد الربيع الأدنى للأعمار فى هذا المثال نحسب ترتيبه وهو يساوى $\frac{1}{4}$ (١٧٣٩) ، أى بوضع ١٧٣٩ بدل ﻫ فى العبارة $\frac{1}{4}$(ﻫ) .

∴ ترتيب الربيع الأدنى = ٤٣٤,٧٥

∴ الربيع الأدنى نفسه يقع بين ١٦,٥ و ١٧,٥ . لأننا نرى من جدول التكرار المتجمع أن الـ ٤٢٤ قيمة الأولى كلها أصغر من ١٦,٥ ، وأن الـ ٧٨٤ قيمة الأولى جميعها أصغر من ١٧,٥ . فإذا فرضنا أن :

الربيع الأدنى = ١٦,٥ + س

يكون $\frac{\text{س}}{١٧,٥ - ١٦,٥} = \frac{٤٣٤,٧٥ - ٤٢٤}{٧٨٤ - ٤٢٤}$

∴ س = ١ × $\frac{١٠,٧٥}{٣٦٠}$

= ٠٣ و

∴ الربيع الأدنى = ١٦,٥٣ سنة .

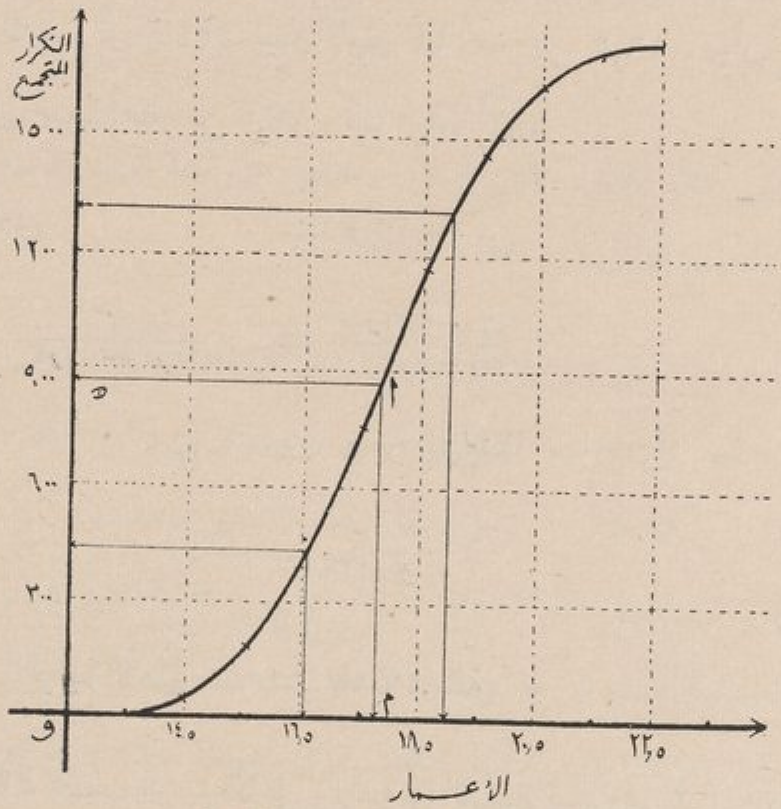
وكذلك نحسب الربيع الأعلى وترتيبه فى السلسلة يساوى $\frac{3}{4}$ (١٧٣٩) ، أى ١٣٠٤,٢٥ ؛ وهو يقع فى الفئة ١٨,٥ — ١٩,٥ ؛ ونحسبه بنفس الطريقة . فينتج أنه يساوى ١٨,٩٦ سنة .

---

(١) الفعل ربع الشىء بمعنى كان فى ربعه أو قسمه إلى أرباع . واسم الفاعل رابع ؛ والصفة المشبهة ربيع ، بمعنى الذى يقسم الشىء إلى أرباع . وأظن أن هذه أنسب وأسهل من « الوسط الربعى » الذى اقترحها البعض . وهذان الربيعان يسميان بالإنجليزية Lower Quartile and Upper Quartile على الترتيب .

إيجاد الربيعين بالرسم

١٤٨ — يمكن الحصول على قيمتى الربيع الأدنى والأعلى من رسم المنحنى التكرارى المتجمع بنفس الطريقة التى استعملناها فى إيجاد الوسيط من نفس الرسم. فنأخذ على المحور الرأسى بعداً يساوى $\frac{1}{4}$ (١٧٣٩) ونرسم خطاً أفقياً يقابل المنحنى الصاعد فى نقطة . والإحداثى الأفقى لهذه النقطة يساوى الربيع الأدنى . وإذا أخذنا بعداً على المحور الرأسى يساوى $\frac{3}{4}$ (١٧٣٩) حصلنا على الربيع الأعلى بنفس الطريقة . ونرى ذلك موضحاً فى الشكل الآتى :



(شكل ٥٨ إيجاد الربيعين بالرسم)

ويلاحظ أن إيجاد الربيعين غير ممكن بهذه السهولة من المنحنى التكرارى العادى ، لأن المنحنى التكرارى المتجمع يساعدنا فى هذه المسألة بالذات بمناسبة أن الإحداثيات الرأسية تمثل التكرارات المتجمعة ؛ فى حين أنها فى المنحنى التكرارى العادى تمثل الإحداثيات الرأسية تكرارات القيم منفردة ، كما سبق أن أشرنا فى بند ١١٢ (صفحة ١٢٥) .

العلاقة بين الربيعين والوسيط

١٤٩ — عدد القيم المحصورة بين الربيعين الأدنى والأعلى يساوى نصف عدد القيم كلها . وهذا واضح ، لأن ربع القيم أقل من الربيع الأدنى ، وربعها أكبر من الربيع الأعلى . وسنرى أن هذه الخاصة لها أهميتها .

الفرق بين الربيع الأعلى والوسيط لا يساوى الفرق بين الأخير والربيع الأدنى ، إلا إذا كان المنحنى التكرارى متماثلا ، بالمعنى الذى ذكرناه سابقا (بند ١٠٣) . وهذه خاصة من خواص المنحنى المتماثل ؛ وسنستعملها فيما بعد لاختبار درجة تماثل المنحنيات أو عدم تماثلها .

والفرق بين هذين المقدارين يكون صغيراً فى المنحنى القريب من التماثل ؛ كما نرى فى المثال الذى نحن بصدده حيث نجد :

الربيع الأعلى — الوسيط = ١٨,٩٦ — ١٧,٧٢٢

= ١,٢٣٨

الوسيط — الربيع الأدنى = ١٧,٧٢٢ — ١٦,٥٣

= ١,١٩٢

وهما متساويان تقريباً .

العشير والمئين

١٤٨ — يمكننا أيضاً حساب القيمة التى تقع عند عشر السلسلة من أسفل أو من أعلى ، وهذه نسميها[١] **العشير** الأدنى أو الأعلى على الترتيب . وذلك بطريقة تشابه طريقة حساب الوسيط والربيعين .

وكذلك يمكن تقسيم السلسلة إلى أجزاء مئوية[٢] **مئينات** وإيجاد قيمة أى جزء منها بنفس الطريقة . ولا نغالى فى التقسيم إلى هذه الأجزاء الصغيرة إلا إذا كان عدد المفردات فى المجموعة غزيراً ويكفى لهذا التقسيم .

ويمكن إيجاد هذه القيم العشرية والمئوية بالرسم أيضاً كما تقدم فى حالة الوسيط والربيعين .

---

(١) بالانجليزية Decile (٢) بالانجليزية Centiles

تعريف الوسط الهندسى

١٥١ — الوسط الهندسى لكميتين هو الجذر التربيعى لحاصل ضربهما . وقياساً على ذلك فالوسط الهندسى لمجموعة من الكميات عددها ن يساوى الجذر النونى لحاصل ضرب هذه الكميات جميعها . ويمكن وضع هذا بصورة مختصرة كما يأتى : نفرض أن القيم هى

$$\text{س}_{١} \text{ ، } \text{س}_{٢} \text{ ، } \text{س}_{٣} \text{ ، } \cdots \text{ ، } \text{س}_{\text{ن}} \text{ .}$$

ونفرض أن الوسط الهندسى لها يساوى هـ مثلا .

$$\therefore \text{هـ} = \sqrt[\text{ن}]{\text{س}_{١} \times \text{س}_{٢} \times \text{س}_{٣} \times \cdots \times \text{س}_{\text{ن}}}$$

$$= \left[\text{س}_{١} \times \text{س}_{٢} \times \cdot\cdot \times \text{س}_{\text{ن}}\right]^{\frac{١}{\text{ن}}}$$

كيفية إيجاده عملياً

ولتسهيل حسابه عملياً نستخدم نظرية اللوغاريتمات فنحسب لوغاريتم الكمية التى فى الطرف الأيسر لهذه المعادلة ؛ ثم نبحث عن العدد المقابل لهذا اللوغاريتم ينتج العدد هـ المطلوب .

$$\therefore \text{لو هـ} = \text{لو}\left[\text{س}_{١} \times \text{س}_{٢} \times \cdot\cdot \times \text{س}_{\text{ن}}\right]^{\frac{١}{\text{ن}}}$$

$$= \frac{١}{\text{ن}} \text{لو}\left[\text{س}_{١} \times \text{س}_{٢} \times \cdot\cdot \times \text{س}_{\text{ن}}\right]$$

$$= \frac{١}{\text{ن}}\left[\text{لو س}_{١} + \text{لو س}_{٢} + \text{لو س}_{٣} + \cdot\cdot + \text{لو س}_{\text{ن}}\right] \text{ ،}$$

لأن لوغاريتم حاصل الضرب ، الذى فى الطرف الأيسر ، يساوى مجموع لوغاريتمات الكميات $\text{س}_{١}$ ، $\text{س}_{٢}$ ، ... ، $\text{س}_{\text{ن}}$.

وينتج إذن أن لوغاريتم الوسط الهندسى لأى عدد من الكميات يساوى الوسط الحسابى للوغاريتمات هذه الكميات .

ونرى من هذا أن حساب الوسط الهندسى لعدة قيم يكون شاقاً خصوصاً إذا كان عددها كبيراً . وهذه الصعوبة العملية من أكبر عيوب الوسط الهندسى ، وتجعل مجال استخدامه محدوداً جداً . ولكنه مع ذلك لا يخلو من بعض مزايا

تفضله عن أغلب المتوسطات السابقة فى بعض المسائل ، كما سنرى فى المستقبل .
وأهم استعمال له فى إيجاد متوسط التغير فى مستوى الأسعار .

تعريف الوسط التوافقى

١٥٢ — يمكن تعريف الوسط التوافقى بين عدة كميات كما يأتى :

لتكن الكميات الأصلية عددها ن وهى :

$$\text{س}_1 \text{ ، } \text{س}_2 \text{ ، } \text{س}_3 \text{ ، } \ldots \text{ ، } \text{س}_{\text{ن}}$$

ولنرمز للوسط التوافقى بينها بالرمز ف ، فطبقاً لتعريف الوسط التوافقى تكون العلاقة بين الوسط التوافقى ف والكميات س هى :

مقلوب الوسط التوافقى يساوى الوسط الحسابى لمقلوبات الكميات :

أى أن $$\frac{1}{\text{ف}} = \frac{1}{\text{ن}}\left(\frac{1}{\text{س}_1} + \frac{1}{\text{س}_2} + \cdots + \frac{1}{\text{س}_{\text{ن}}}\right).$$

أو $$\frac{\text{ن}}{\text{ف}} = \text{مجـ}\left(\frac{1}{\text{س}}\right).$$

» $$\text{ف} = \frac{\text{ن}}{\text{مجـ}\left(\frac{1}{\text{س}}\right)}.$$

ولكن حساب هذا الوسط عقيم ومتعب كما نرى من هذه المعادلات . فهو يستلزم حساب مقلوبات الكميات وجمعها ثم حساب متوسطها . وربما كان حساب الوسط الهندسى أسهل فى بعض الأحيان . وهذه الصعوبة العملية من أكبر عيوب الوسط التوافقى ، فضلا عن أن فكرته معقدة وصعبة الفهم . وهو يستعمل فى حالات معدودة فقط . ومنها مثلا حساب متوسط سعر السلع حينما نذكر قيمة وحدة النقود بالنسبة إلى السلع . كأن نقول اشتريت ٥ برتقالات بقرش ثم اشتريت ٧ بقرش آخر ، وتسعة بقرش ثالث . فإذا قلنا إن المتوسط هو ٧ برتقالات بقرش كان هذا هو الوسط التوافقى لسعر البرتقال وليس الوسط الحسابى .

لأن أسعار البرتقال التى اشترينا بها هى فى الحقيقة $\frac{١}{٥}$ قرش و $\frac{١}{٧}$ قرش و $\frac{١}{٩}$ قرش للبرتقالة الواحدة فى المرات الثلاثة على الترتيب . والوسط التوافقى لهذه الأسعار هو $\frac{١}{٧}$ قرش للواحدة أى ٧ برتقالات بقرش . أما الوسط الحسابى للأسعار فهو $\frac{١}{٣}$ ($\frac{١}{٥}$ + $\frac{١}{٧}$ + $\frac{١}{٩}$) قرش للبرتقالة الواحدة .

اختلاف مفردات المجموعة فى الأهمية

١٥٣ — نلاحظ فى جميع المتوسطات المتقدمة أننا كنا نعامل جميع المفردات فى المجموعة معاملة المساواة ، فلا نرجح واحدة على الأخرى ولا نهمل واحدة أو نعزز أخرى . ففى الوسط الحسابى مثلا نجمع جميع المفردات بدون استثناء . وكذلك فى الوسط التوافقى والوسط الهندسى ، والمنوال والوسيط أيضاً .

وهذا الإجراء نتبعه طبعاً فى كل الأحوال التى لا يكون فيها تفاضل بين المفردات فى الأهمية . ولكن إذا أعطينا مجموعة وعلمنا أن بعض مفرداتها فعلاً أهم من بعضها الآخر ، أو أن قيمةً أرجح من أخرى لسبب من الأسباب ، وأردنا إيجاد المتوسط لهذه المجموعة ، لا يسعنا إلا أن نأخذ فى الاعتبار هذا الاختلاف فى الأهمية بين المفردات . فهل نعتبر مثلاً سعر القمح فى ساحل روض الفرج بالقاهرة — وليكن ١٤٤ قرشاً للأردب مثلا — على قدم المساواة مع سعر القمح فى سوق ريفية صغيرة — وليكن ١٤٠ قرشاً — ونقول إن متوسط سعر القمح = $\frac{١٤٤ + ١٤٠}{٢}$ أى ١٤٢ قرشاً فى ذلك اليوم ؟ مع علمنا بأن ساحل روض الفرج يحصل فيه تعامل بآلاف الأرادب كل يوم ويجتمع فيه كبار التجار ، وأن السوق الأخرى لا تعقد إلا يوماً كل أسبوع ولا يباع فيها إلا بعض عشرات الأرادب ؟ .

الوسط المرجح

١٥٤ — فى مثل هذه الأحوال يجب أن نحسب **الوسط المرجح** (١)

---

(١) اسمه بالانجليزية The Weighted Mean ؛ بعض الناس يسمونه الوسط المعدل ؛ ولكن هذه الترجمة غير مناسبة ، وتوقعنا فى التباس فى معنى التعديل . والبعض يسمونه الوسط المثقل ، وهى ترجمة حرفية معيبة .

الذى يأخذ فى الاعتبار الأهمية النسبية للمفردات ، حيث ترجح كل مفردة بما يتناسب وأهميتها بالنسبة إلى غيرها .

ويجب أن نفكر فى كيفية قياس الأهمية النسبية للمفردات المختلفة فى المجموعة، كل حالة وما يناسبها . ومقدار الأهمية النسبية لكل مفردة نسميه[1] **وزناً** ونستعمله لترجيح هذه المفردة بالنسبة إلى غيرها عند حساب الوسط المرجح .

تعريف الوسط المرجح

الوسط الحسابى المرجح يساوى مجموع حواصل ضرب القيم فى أوزانها ، مقسوماً على مجموع الأوزان . فإذا كانت القيم هى :

$$س_١ ، س_٢ ، س_٣ ، \dots ، س_ن$$

وأوزانها $و_١ ، و_٢ ، و_٣ ، \dots ، و_ن$ على الترتيب ،

فإن الوسط المرجح لهذه القيم حسب التعريف

$$= \frac{س_١ \cdot و_١ + س_٢ \cdot و_٢ + \dots + س_ن \cdot و_ن}{و_١ + و_٢ + \dots + و_ن}$$

أو $$= \frac{مج(س و)}{مج و} .$$

لنفرض مثلا أن أسعار المواد الغذائية ارتفعت بنسبة ١٠ ٪ ، وأسعار الملابس بنسبة ١٥ ٪ وإيجار المساكن بقدر ٢٠ ٪ — هنا لا يصح أن نقول إن متوسط الزيادة فى نفقات المعيشة هو $\frac{١٠+١٥+٢٠}{٣} = ١٥$ ٪ ، مع علمنا بأن ما تنفقه كل أسرة على الطعام يعادل تقريباً ضعف ما تنفقه على المسكن ، وأربعة أمثال ما تنفقه على الملبس . بل الواجب أن نرجح الزيادة فى الطعام والزيادة فى المسكن بقدر أهميتهما بالنسبة إلى الزيادة فى الملابس . فنقول إن الطعام أربعة أمثال الملابس ، والمسكن ضعفها فى الأهمية . وعلى ذلك نضرب زيادة الملابس فى ١ وزيادة المسكن فى ٢ وزيادة الطعام فى ٤ ، ونقسم على مجموع هذه الأوزان وهو ٧.

(١) بالانجليزية Weight.

$$\therefore \text{الوسط المرجح للزيادة} = \frac{١٠ \times ٤ + ١٥ \times ١ + ٢٠ \times ٢}{٧}$$

$$= ١٣٫٥٧ \text{ ٪}$$

وهذه نتيجة مخالفة للمتوسط الأول بدون أوزان . وطبيعى أننا نحصل على نتيجة مخالفة لهذه إذا استعملنا أوزاناً غير ٤ و ١ و ٢ المستعملة هنا .

اختيار الأوزان المناسبة

١٥٥ — رأينا أنه من الواجب استعمال أوزان مختلفة للتعبير عن الأهمية النسبية للمفردات فى المجموعة ، إذا ثبت لنا اختلاف هذه المفردات فى الأهمية . ويجب إذن أن نختار هذه الأوزان بحيث تعبر بدقة عن درجة أهميـة كل قيمة أو مفردة فى المجموعة . ولا بد طبعاً أن نأخذ فى الاعتبار الناحية الرئيسية التى تتجلى فيها هذه الأهمية ، وعلاقتها بالموضوع الذى سنستعمل فيه الوسط المرجح الذى نبحث عنه . وإذا كانت للمفردات أهمية خاصـة فى ناحية لا علاقة لها بهذا الموضوع ، ولا تؤثر فيه ، فلا شأن لنا بها . ويجب أن نتركها ونبحث عن أهمية القيم فى الناحية التى لها مساس بموضوع بحثنا الذى سنستخدم فيه الوسط المرجح المطلوب .

لنفرض مثلاً أن صانعاً يستخدم فى صناعته خامات من معادن النحاس والزنك والقصدير ويشترى كل يوم بمبلغ ٦٠ قرشاً من النحاس و ٣٠ قرشاً من الزنك و ١٠ قروش من القصدير . ولنفرض أيضاً أن أسعار هـذه المعادن ارتفعت يوماً ما بنسبة ٥٠٪ للنحاس و ٣٠٪ للزنك و ١٠٠٪ للقصدير .

فلو أردنا فى مثل هذه الحالة معرفة متوسط الزيادة فى أسعار هذه المعادن لاستخدامه فى تقدير الزيادة فى تكاليف الخامات لهذا الصانع وفى ثمن بيع منتجاته ، نحسب المتوسط المرجح للزيادات ٥٠٪ و٣٠٪ و١٠٠٪ التى طرأت على أسعار هذه المعادن ؛ وتكون الأوزان المناسـبة فى هذه الحالة هى مقـادير ما كان ينفقه هذا الصـانع على هذه المعـادن كل يوم ، أى ٦٠ قرشاً للنحاس و ٣٠ للزنك و ١٠ للقصدير ، لأن هذه المقادير هى أحسن مقياس للأهمية النسبية بين هذه المعادن .

وعلى ذلك يكون المتوسط المرجح للزيادة فى أسعار المعادن

$$= \frac{٥٠ \times ٦٠ + ٣٠ \times ٣٠ + ١٠٠ \times ١٠}{٦٠ + ٣٠ + ١٠}$$

$$= ٤٩ \,\%.$$

ولكن إذا كان هناك صانع آخر يستخدم نفس هذه المعادن ولكن بنسب مختلفة عن الصانع الأول ، كأن يكون مقدار ما ينفقه على شرائها هو ٢٠ قرشاً للنحاس و٣٠ للزنك و٥٠ للقصدير مثلا ، فإن المتوسط المرجح فى هذه الحالة يكون

$$= \frac{٥٠ \times ٢٠ + ٣٠ \times ٣٠ + ١٠٠ \times ٥٠}{٣٠ + ٢٠ + ٥٠}$$

$$= ٦٩ \,\%.$$

أى أن الزيادة فى أسعار المعادن تؤثر على الصانعين تأثيرين مختلفين .

ومن هذا المثال يتضح لنا أن الأهمية الكبرى التى للنحاس مثلا عند الصانع الأول ، لا شأن لنا بها عند تقدير متوسط الزيادة فى الأسعار لاستخدامه فى تقدير منتجات الصانع الثانى .

الوسط المرجح يتغير بتغير الأوزان

١٥٦ — الوسط المرجح لأى عدد من القيم المعلومة ، يتغير طبعاً حسب الأوزان المستعلة فى استخراجه . ولكن الوسط المرجح ليس حساساً لهذه التغيرات فى الأوزان ، أى أن تغييراً كبيراً فى الوزن لا يحدث تأثيراً كبيراً فى النتيجة النهائية ، كما لو تغيرت إحدى القيم التى نحسب لها المتوسط . و يتضح هذا من المثال الآتى :

لنأخذ القيم	١٥ و ١٢ و ٢٠	
وأوزانها	٤ و ١ و ٢	على الترتيب .

∴ الوسط المرجح بهذه الأوزان

$$= \frac{١٥ \times ٤ + ١٢ \times ١ + ٢٠ \times ٢}{٤ + ١ + ٢} = ١٦ .$$

نفرض أننا غيرنا وزن القيمة الأولى من ٤ إلى ٣ ، وتركنا القيم كما هى .

$$\therefore \text{الوسط المرجح} = \frac{١٥ \times ٣ + ١٢ \times ١ + ٢٠ \times ٢}{٣ + ١ + ٢}$$

$$= ١٦{,}١٧$$

فنرى أن تغيير قدره ٢٥ ٪ فى الوزن غيَّر الوسط من ١٦ إلى ١٦,١٧ ، أى بنسبة ١ ٪ تقريباً .

نفرض الآن أننا تركنا الأوزان كما هى وغيرنا إحـدى القيم ، الأولى مثلا ، من ١٥ إلى ١٢ .

$$\therefore \text{الوسط المرجح} = \frac{١٢ \times ٤ + ١٢ \times ١ + ٢٠ \times ٢}{٤ + ١ + ٢}$$

$$= ١٤{,}٢٨$$

ونرى هنا أن تغييراً قدره ٢٠ ٪ فى إحدى القيم غير الوسط المعدل من ١٦ إلى ١٤,٢٧ ، أى بقدر ١٠,٨ ٪ تقريباً . وهذا أكبر بكثير من تأثير التغيير فى الوزن والسبب فى ذلك واضح . فنحن إذا أنقصنا الوزن من ٤ إلى ٣ ينقص البسط تبعاً لذلك وينقص المقام أيضاً . وهـذا النقص فى قيمة المقام يعوض شيئاً من النقص فى قيمة البسط ، ولا يتغـير الكسر تغيراً كبيراً ؛ وكذلك الأمر إذا زاد الوزن . أما إذا أنقصنا القيمة ١٥ (أو أى قيمة أخرى) ؛ وتركنا الأوزان كما هى ؛ فإن قيمة البسـط تنقص وتبقى قيمة المقـام كما هى . وبذلك تتغير قيمة الكسر تغيراً كبيراً . وكذلك فى حالة الزيادة .

الدقة فى تقدير الأوزان غير مهمة

والذى نستفيده من هذا أنه لا يلزم عند اختيار الأوزان أن نتحرى منتهى الدقة فى معرفة قيمتها ، ولكن يكفى أن نأخذها مقربة بقدر الإمكان ؛ ولا يؤثر ذلك كثيراً فى النتائج التى نحصل عليها للسبب الذى ذكرناه . وهذا تسهيل كبير يذكر فى الوسط المرجح .

الأوزان عديمة الأثر

١٥٧ — نصادف أحياناً (وهذا نادر) أن نختار أوزاناً ليس لها أثر ، بمعنى أن الوسط المرجح بها وبدونها واحد لا يتغير .

لنأخذ مثلا القيم ١٢ و ١٠ و ٢٠ مثلا ؛ ولتكن أوزانها على الترتيب هى ٤ ، ١ ، ٢ .

$$\text{الوسط الحسابى} = \frac{١٢ + ١٠ + ٢٠}{٣} \quad \text{بدون أوزان ،}$$

$$= ١٤$$

$$\text{والوسط المرجح} = \frac{١٢ \times ٤ + ١٠ \times ١ + ٢٠ \times ٢}{٤ + ١ + ٢} = \frac{٩٨}{٧}$$

$$= ١٤ \ \text{أيضاً .}$$

والواقع أن هذا ليس مجرد مصادفة . إنما هى خاصة رياضية تربط مجموعة القيم ١٢ و ١٠ و ٢٠ بمجموعة الأوزان ٤ و ١ و ٢ . وكل مجموعة أخرى من ثلاث قيم يمكن أن نجد لها مجموعة من ثلاثة أوزان « عديمة الأثر » كهذه . وهذه القضية يمكن إثباتها جبرياً بدون صعوبة [1]

---

(١) لإثبات ذلك نفرض أن القيم المعلومة هى ا ، ب ، حـ وأن الأوزان عديمة الأثر هى س ، ص ، ع .

فالمطلوب إيجاد س ، ص ، ع بحيث إن الوسط الحسابى يساوى الوسط المرجح للقيم ا ، ب ، حـ .

أى أن $\frac{١}{٣}$ (ا + ب + حـ) = $\frac{\text{ا س + ب ص + حـ ع}}{\text{س + ص + ع}}$ حسب تعريف الوسط المرجح .

∴ ا (ص + ع − ٢س) + ب (ع + س − ٢ص) + حـ (س + ص − ٢ع) = ٠

أى ا ل + ب م + حـ ن = ٠ مثلا .

ولكن ل + م + ن = ٠

∴ م (ب − ا) + ن (حـ − ا) = ٠ وبالمثل م (ب − حـ) + ل (ا − حـ) = ٠

∴ ل : م : ن = (ب − حـ) : (حـ − ا) : (ا − ب) .

فإذا علمنا قيم ا ، ب ، حـ ، ولتكن ١٢ و ١٠ و ٢٠ ، كما فى المثال المذكور ، نعوض بدلها فى الطرف الأيسر من هذه المعادلة الأخيرة فنحصل على النسبة ل : م : ن = − ١٠ : ٨ : ٢ . ونرى بسهولة أن س ، ص ، ع تحقق هذه المعادلة إذا كانت : س : ص : ع = ٤ : ١ : ٢ وهو المطلوب .

الوسط التوافقى المرجح

١٥٨ — لإيجاد الوسط التوافقى المرجح نضرب مقلوب كل قيمة فى وزن هذه القيمة ، ثم نقسم مجموع حواصل الضرب هذه على مجموع الأوزان ؛ والناتج يساوى مقلوب الوسط التوافقى المرجح . فلو أخذنا القيم

$$س_{١} \quad س_{٢} \quad س_{٣} \quad . \; . \; . \quad س_{ن} \quad ،$$

وأوزانها $$و_{١} \quad و_{٢} \quad و_{٣} \quad . \; . \; . \quad و_{ن} \quad ،$$

ورمزنا للوسط التوافقى المرجح بالحرف ق

يكون $$\frac{١}{ق} = \frac{\frac{و_{١}}{س_{١}} + \frac{و_{٢}}{س_{٢}} + \ldots + \frac{و_{ن}}{س_{ن}}}{و_{١} + و_{٢} + \ldots + و_{ن}}$$

الوسط الهندسى المرجح

١٥٩ — ولإيجاد الوسط الهندسى المرجح[1] نرفع كل قيمة إلى قوة تساوى وزنها ، ونضرب هذه القوى للقيم المختلفة ، ثم نأخذ الجذر الذى رتبته تساوى مجموع الأوزان لهذا الحاصل .

نفرض أن الوسط الهندسى المرجح يساوى هـ ، وأن القيم وأوزانها كما تقدم ، فيكون

$$هـ = \left[(س_{١})^{و_{١}} \times (س_{٢})^{و_{٢}} \times \ldots \times (س_{ن})^{و_{ن}} \right]^{\frac{١}{مجـ و}}$$

$$\therefore \quad لو\ هـ = \frac{و_{١} \cdot لو س_{١} + و_{٢} \cdot لو س_{٢} + \ldots + و_{ن} \cdot لو س_{ن}}{و_{١} + و_{٢} + \ldots + و_{ن}}$$

أى أن لوغاريتم الوسط الهندسى المرجح يساوى الوسط المرجح للوغاريتمات القيم الأصلية . وترجيح اللوغاريتمات يكون بضرب لوغاريتم كل قيمة فى وزن هذه القيمة ، ثم جمع هذه الحواصل وقسمة الناتج على مجموع الأوزان . ولكن هذا الوسط قليل الأهمية نظراً لكثرة تعقيده ، وصعوبة حسابه عملياً .

---

(١) بالانجليزية Weighted Geometric Meian

المنوال والوسيط المرجحان

١٦٠ — هذا ويمكن تعميم فكرة الترجيح لتشمل المنوال والوسيط فنحصل على « المنوال المرجح » (١) « والوسيط المرجح » (٢) . فبعد اختيار الأوزان والاتفاق عليها يكون المنوال هو القيمة الأكبر وزناً من جميع القيم الأخرى . وهذا يتمشى مع الفكرة الأصلية للمنوال بمعنى أنه القيمة الأكثر شيوعاً من غيرها .

وأما الوسيط المرجح فنجده بسهولة بعد أن نرتب القيم تصاعدياً ونضع أمام كل واحدة وزنها المتفق عليه . فالقيم فى هذا الوضع تشابه القيم فى الجدول التكرارى وأمامها تكراراتها . ونعتبر الأوزان هنا بمثابة التكرارات هناك ، ونجرى العمل بالطريقة المعروفة .

ولكن استعمال هذين الوسطين المرجحين نادر ، وهما لذلك قليلا الأهمية .

المقارنة بين المتوسطات المختلفة

١٦١ — يحسن قبل أن نترك هذا الباب أن نبحث فى العلاقة بين هذه المتوسطات المختلفة والمقارنة بينها من حيث تأديتها للغرض المقصود منها : ألا وهو تمثيل المجموعات الإحصائية المختلفة .

الوسط الحسابى يتأثر بالقيم الشاذة

١٦٢ — رأينا أن الوسط الحسابى يساوى مجموع القيم مقسوماً على عددها ، وأن القيم تؤخذ جميعها على علاتها بصرف النظر عما يكون بينها من تفاوت فى الأهمية . ولذلك فإن الوسط الحسابى يتأثر كثيراً بالقيم المتطرفة ، خصوصاً فى المجموعات الصغيرة العدد . وقد يعطى فكرة خطأ عن المجموعة إذا كانت إحدى قيمها بالصدفة كبيرة جداً (أو صغيرة جداً) بالنسبة لباقى القيم . ففى المجموعة الآتية للأجور مثلا :

---

(١) بالانجليزية Weighted Mode (٢) بالانجليزية Weighted Median

٧,٥ ، ٨ ، ٨,٥ ، ٩ ، ١٠,٥ ، ١١ ، ٢١,٥ قرشاً ،

نلاحظ أن أغلبها حوالى ٩ قروش ، ولكن أحد هذه الأجور كبير جداً ، أى ٢١,٥ قرشاً ؛ وربما كان وجوده فى هذه المجموعة عن طريق المصادفة فقط ؛ والوسط الحسابى لهذه الأجور هو ١١ قرشاً تقريباً . وواضح أنه لا يمثل المجموعة تماماً ، إذ يظهر أن القيم تتجمع حول ٩ ، والفرق بين ٩ و ١١ ناتج من شذوذ الأجر ٢١,٥ وبعده عن باقى الأجور الأخرى . ولو كانت المجموعة أكبر عدداً مما هى الآن ، ربما شملت أجراً منخفضاً جداً (أقل من ٧,٥ المذكور هنا) يعادل تأثير هذا الأجر الشاذ ، فيكون الوسط الحسابى أقرب إلى الحقيقة من ١١

عيوب الوسط الحسابى

١٦٣ — والوسط الحسابى ، كما قلنا سابقاً ، لا يساعدنا فى التوزيعات التكرارية المفتوحة . وفى مثل هذه التوزيعات يجب أن نلجأ إلى الوسيط أو المنوال .

ويجب ألا نعتمد على الوسط الحسابى بمفرده فى تمثيل المجموعات غير المتجانسة ومتعددة المناويل . لأن الاقتصار على ذكر الوسط الحسابى فى مثل هذه الأحوال — أو أى متوسط آخر — فيه تضليل ؛ ويجب أن نذكر هذه الحقيقة إلى جانب الوسط الحسابى إذا عرفنا قيمته . ولكن الأحسن هو حساب هذه المناويل وهذا أوضح وأدق ، وأحسن دلالة على صفات المجموعة ، من الاقتصار على الوسط الحسابى الذى لا يمثلها تمثيلا صحيحاً .

والوسط الحسابى مضلل أيضاً فى حالة التوزيع التكرارى ذى النهاية الصغرى (بند ١٠٦ شكل ٤٣ صحيفة ١١٨) . لأنه فى هذه المجموعات يكون فى الوسط بين أكبر قيمة وأصغر قيمة . ويكون تكراره أصغر تكرار أو بالقرب منه . وعلى ذلك فلا يمكن أن نأخذه كنموذج للمجموعة .

ميزات الوسط الحسابى

١٦٤ — ومن أكبر ميزات الوسط الحسابى سهولته وبساطة الفكرة المبنى عليها، ووضوحها. فمن السهل على كل إنسان فهم المعنى المقصود بالوسط الحسابى. بخلاف المتوسطات الأخرى، فهى تجمع إلى التعقيد فى المعنى والفكرة صعوبة عملية فى حسابها ومعرفة مقاديرها بالدقة. ولذلك فالوسط الحسابى أكثر المتوسطات استعمالا، رغم ما فيه من عيوب فى بعض الأحيان.

ومن الناحية النظرية أيضاً، نجد فكرة الوسط الحسابى — أى مجموع القيم على عددها — تساعدنا كثيراً فى الأبحاث الرياضية الخاصة بهذه الموضوعات: فمن السهل استنباط خواص مهمة ومفيدة للوسط الحسابى بطرق رياضية يصعب تطبيقها فى حالة المتوسطات الأخرى. ولهذا نجد أن الوسط الحسابى قد تناولته أبحاث نظرية من نواح متعددة، بخلاف المتوسطات الأخرى.

بعض ميزات المنوال وعيوبه

١٦٥ — نلاحظ أننا عند حساب المنوال فى أى توزيع تكرارى لا نهتم بالقيم أو الفئات المتطرفة فى الجدول. بل نعنى فقط بالثلاث فئات ذات التكرارات الكبرى. وهذه تكون فى وسط الجدول فى العادة — إذا كان التوزيع ذا منوال واحد فقط. وهذه الخاصة التى يختص بها المنوال — والوسيط أيضا — تساعدنا فى دراسة التوزيعات التكرارية المفتوحة من أحد الطرفين أو كليهما، كما ذكرنا سابقاً. وهذه ميزة نحتاج إليها أحياناً.

والمنوال لا ينفعنا فى دراسة المنحنى التكرارى ذى النهاية الصغرى، ولا فى المنحنى ذى الفرع الواحد — صاعداً كان أو هابطاً — والسبب فى ذلك واضح إذ أن هذه المنحنيات ليس لها منوال بالمعنى الذى نفهمه كما فى المنحنى التكرارى العادى. ومع أننا نجد فى كل من هذه التوزيعات التكرارية قيمة تكرارها أكبر من تكرار أى قيمة أخرى، إلا أنها تكون قيمة طرفية، لا يسبقها

قيم أو لا يليها قيم (كما نرى فى الأشكال ٤١ و ٤٢ و ٤٣) ولكننا لا نعتمد كثيراً على القيم الطرفية فى تمثيل المجموعات .

إذا ضربنا المنوال فى عدد القيم كلها لا نحصل على مجموع هذه القيم كما فى الوسط الحسابى . فمثلا إذا عرفنا الوسط الحسابى لأجور العمال أمكننا معرفة جملة مايكسبه هؤلاء العمال من أجور بضرب الوسط فى عددهم . ولكن لا يمكننا ذلك مع المنوال إلا إذا كان التوزيع التكرارى متماثلا ، حيث يكون المنوال يساوى الوسط الحسابى — والوسيط أيضاً .

هذا وقد ذكرنا سابقاً أن الصعوبة العملية فى حساب قيمة المنوال والأخطاء التى تتعرض لها فى ذلك ، تجعلنا لا نلجأ إليه كثيراً فى أعمالنا ؛ ولو أن الفكرة النظرية للمنوال فكرة صحيحة ووجيهة فى ذاتها .

أهم ميزات الوسيط

١٦٦ — من ميزات الوسيط أننا فى حساب قيمته لا نهتم بمقادير القيم مثل ما نهتم بترتيب هذه القيم بين بعضها : الأصغر فالأكبر . وهذه الخاصة يمكن الانتفاع بها لتصحيح خطأ الوسط الحسابى عند تأثره بالقيم المتطرفة كما ذكرنا سابقاً (بند ١٦٢) . ففى مجموعة الأجور التى ذكرناها ، وهى :

٧,٥ ، ٨ ، ٨,٥ ، ٩ ، ١٠,٥ ، ١١ ، ٢١,٥ قرشاً

لانهتم بمقادير هذه الأجور عند حساب الوسيط ، بل نأخذ الأجر الأوسط منها وهو ٩ قروش هنا . وبذلك لا يتأثر المتوسط بالقيمة الشاذة ٢١,٥ التى وجدت فى المجموعة صدفة ، وهكذا نأخذ فكرة صحيحة عن الأجور أحسن مما لو استعملنا الوسط الحسابى .

من خواص الوسيط أيضاً أن مجموع الفروق (بدون إشارة) بين المفردات والوسيط يكون أقل من مجموع الفروق بين نفس المفردات وأى قيمة أخرى من قيم المجموعة ، غير الوسيط . فلو أخذنا المجموعة المذكورة أعلاه مثلا ، ووسيطها

٩ قروش ، نجد أن مجموع الفروق بدون إشارة ، بين هذه الأجور والعدد ٩ ، يساوى ١٩ . ولو أخذنا الأجر ١٠,٥ مثلا بدل ٩ ، نجد هذا المجموع يساوى ٢٠,٥ ؛ أو ٢٢ إذا أخذنا ٨ — ويمكن إثبات هذه الخاصة عموماً بدون صعوبة[١]

العلاقة بين الوسط والمنوال والوسيط

١٦٧ — فى التوزيع التكرارى المعتدل نجد أن الوسط الحسابى والوسيط والمنوال كلها متساوية . وهذه الخاصة تنتج من معنى التماثل . فهذا المنحنى ينقسم إلى نصفين متطابقين بواسطة الخط الرأسى الذى يمر بالقمة ، أى محور التماثل . وهذا يثبت أن المنوال يساوى الوسيط وهما عبارة عن مسقط قمة المنحنى على المحور الأفقى . ونظراً لتماثل توزيع التكرارات نرى أن القيم متساوية البعد عن المنوال (من أعلى ومن أسفل) متساوية التكرار أيضاً . وهذا يثبت أن المنوال هو الوسط الحسابى أيضاً .

وإذا كان الفرق بين هذه المتوسطات صغيراً فى منحنى تكرارى معين فهذا يدل على أن هذا المنحنى قريب من التماثل . ومثال ذلك المنحنى التكرارى فى شكل ٣٣ (صفحة ١٠٤ . التوزيع فى جدول ٩ صفحة ٩٩) — وجدنا أن الوسط الحسابى لهذا التوزيع التكرارى لأعمار ١٧٣٩ تلميذاً هو ١٧,٧٤٩ سنة (صفحة ١٤٥) ، والوسيط ١٧,٧٢٣ (صفحة ١٦٢) ، والمنوال ١٧,٧٢ (صفحة ١٥٥) .

وعلى العموم يكون الوسيط واقعاً بين الوسط الحسابى والمنوال . وعندما يكون المنحنى التكرارى قريباً من التماثل جداً ، يكون الفرق بين الوسط والوسيط ثلث الفرق بين الوسط والمنوال تقريباً :

الوسط — الوسيط = $\frac{1}{3}$ (الوسط — المنوال) ، تقريباً .

ويجب ملاحظة أن هذه العلاقة تقريبية وليس من اللازم تحققها بالضبط فى أى حالة معينة .

---

(١) أنظر كتاب Whittaker & Robinson *"Calculus of Observations,"* (1929), p. 197.

المـــراجع

Bowley, A.L., : *Elementary Manual of Statistics*, Chapter III.

Bowley, A.L., : *Elements of Statistics*, Chapter V.

Connor, R.L.,: *Statistics in Theory and Practice*, Chapter X.

Jones, D.C., : *First Course in Statistics*, Chapters IV, V.

King, W.I., : *Elements of Statistical Method*, Chapter XII.

Mills, F.C., : *Statistical Methods*, Chapter IV.

Secrist, H., : *Statistical Methods*. Chapter IX.

الباب السابع

التشتت

١٦٨ — « التشتت » [١] فى أى مجموعة من القيم ، يقصد به التباعد بين مفرداتها أو التفاوت أو الاختلاف بينها . وهذا التشتت يكون صغيراً بالطبع إذا كان التفاوت بين مفردات القيم قليلاً ، أى إذا كانت القيم قريبة من بعضها . ويكون التشتت كبيراً إذا كان التفاوت بينها كبيراً ، أى إذا كانت القيم بعيدة عن بعضها . معنى التشتت

وعلى ذلك يمكننا أن نتخذ مقدار التشتت ، قليلاً كان أو كبيراً ، كدليل على تجمع القيم وقربها من بعضها أو على تفرقها وتباعدها عن بعضها . وهكذا يكون لدينا مقياس لمقدار تجانس المجموعات الإحصائية ، أو عدم تجانسها .

لا شك أن صفة التجانس من عدمه . صفة تهمنا معرفتها فى كل مجموعة ندرسها . ومعرفة المتوسطات لا تغنينا عنها . فحبذا لو أمكننا قياس التشتت بطريقة تعبر — بدقة و بسهولة — عن هذه الصفة المهمة التى نقصدها .

١٦٩ — يوجد عدة طرق إحصائية لقياس التشتت تختلف فيما بينها فى الدقة والسهولة فى العمل ، وفى الأساس النظرى الذى تبنى عليه . مقاييس التشتت

أسهل هذه الطرق عملياً هى طريقة قياس المدى [٢] بين أصغر وأكبر قيمة

---

(١) بالإنجليزية Dispersion

(٢) بالإنجليزية Range

فى المجموعة ، واعتبار طول هذا المدى كمقياس للتشتت فى هذه المجموعة . فالقيم الآتية مثلا :

٢٨ ، ٣١ ، ٣٢ ، ٣٢,٥ ، ٣٣ ، ٣٥ ، ٣٥,٥ ، ٣٦ ، ٣٧,٥ ، ٤٠

تنتشر فى مدى قدره من ٢٨ إلى ٤٠ ، أى ١٢ ، ونعتبرها أكثر تشتتاً من المجموعة الآتية :

٣٢ ، ٣٢,٣ ، ٣٢,٧ ، ٣٣,٥ ، ٣٦ ، ٣٦,٢ ، ٣٦,٨ ، ٣٧,٤ ، ٣٨ .

لأن هذه المجموعة الأخيرة تنتشر فى فترة أصغر ، وقدرها ٦ فقط ؛ وعلى ذلك فالقيم هنا أقرب إلى بعضها منها فى المجموعة الأولى .

هذه الطريقة غير دقيقة ومضللة

ولكن هذه الطريقة ، وإن كانت سهلة عملياً ، ليست دقيقة ؛ وأحياناً تكون مضللة . حيث قد يسبب وجود قيمة متطرفة فى المجموعة زيادة كبيرة فى طول المدى ، يستدل منها — خطأ — على وجود تشتت كبير بين المفردات ، مع أن جميعها فى الواقع متجمعة بالقرب من بعضها ما عدا هذه القيمة الشاذة ، كما يظهر فى المجموعة الآتية للأعمار مثلا :

١٨ ، ١٨,٥ ، ١٩ ، ٢٠ ، ٢٠,٥ ، ٢١ ، ٣١ سنة .

فوجود العمر الشاذ ٣١ فى المجموعة يجعل المدى ١٣ سنة ، وتبدو الأعمار كأنها مشتتة كثيراً مع أن أغلبيتها متجمعة حوالى ٢٠ سنة .

المدى بين الربيعين

١٧٠ — يمكن أن نتفادى هذا العيب بأن نحدد المدى بحيث لا يتأثر بالقيم الشاذة والمتطرفة . لذلك نأخذ المدى بين الربيعين الأعلى والأدنى كمقياس للتشتت . وبهذه الطريقة نتخلص من تأثير القيم المتطرفة التى تكون أحياناً متأثرة بظروف خاصة لا تسرى على باقى القيم فى المجموعة .

ومن ناحية أخرى نعلم أن نصف عدد المفردات فى المجموعة ينحصر

بين هذين الربيعين . وهـذا هو الجزء المهم من المجموعة الذى يكون فيه ازدحام المفردات أكثر ما يمكن . وهو يحتوى على الأشياء المهمة فى المجموعة مثل الوسط الحسابى والمنوال والوسيط وباقى المتوسطات . وعلى ذلك فاتخاذ هذه المنطقة لقياس التشتت أفضل من الطريقة المتقدمة .

وفى العادة نأخذ مقياس التشتت يساوى نصف المدى بين الربيعـين [(١)] ، بدلا من المدى كله .

وحساب هذا المقياس سهل طبعاً ، حيث نحسب الربيعين الأعلى والأدنى باحدى الطرق المعروفة ، ونأخـذ نصف الفرق بينهما . ففى التوزيع التكرارى لأعمار ١٧٣٩ طالباً المذكور فى جـدول ٢٥ (صفحة ١٦١) مثلا ، وجـدنا (صفحة ١٦٥) أن الربيع الأدنى للأعمار يساوى ١٦,٥٣ سـنة والربيع الأعلى لها يساوى ١٨,٩٦ سنة . ومقياس التشتت بهذه الطريقة يساوى إذن

$$\frac{1}{2}\left(١٨,٩٦ - ١٦,٥٣ \right) = ١,٢١٥ \text{ سنة .}$$

أما بالطريقة السابقة ، فنجد من الجدول التكرارى (جدول ٢٥) أن أصغر قيمة فى المجموعة كلها هى ١٣,٥ سنة ، وأكبر قيمة فى الجدول هى ٢٢,٥ سنة . وعلى ذلك فالمدى بين أصغر وأكبر قيمة يساوى ٩ سـنين ، وهو مقياس آخر للتشتت فى هذه المجموعة .

اختلاف مقاييس التشتت

١٧١ — يلاحظ هنا اختلاف مقياسى التشتت لنفس المجموعة : أحـدهما يساوى ٩ سـنين والآخر ١,٢١٥ سنة فقط . وهـذا هو الواجب أن يكون رغم ما يمكن أن يقال بوجوب تطابق النتيجتين باعتبار أنهما مقياس لشىء واحد وهو تشتت مجموعة واحـدة . ولكن السبب فى هذا الاختلاف هو أننا نقيس هـذا

(١) اسمه بالإنجليزية Semi-Inter Quartile Range.

الشيء الواحد على أسس مختلفة . كما لو أردنا مثلا قياس درجة النمو فى مجموعة من التلاميذ . فقد نأخذ طول قامة الشخص كدليل على درجة نموه ؛ وهذا يقاس بالسنتيمترات . وقد نأخذ وزنه كدليل على درجة النمو أيضاً ؛ ونقيس الوزن بالكيلو جرام أو الرطل مثلا . وطبعاً لا ننتظر أن يكون الرقم الدال على وزن الشخص هو نفس الرقم الدال على طوله ، باعتبار أنهما قياسان لشيء واحد وهو درجة النمو .

وهذا ينطبق أيضاً على الطرق الذى سنذكرها لقياس التشتت . فهى طرق مبنية على أسس مختلفة ، ولا ينتظر أن تؤتى نفس النتائج .

والمهم فى هذه المسألة أننا إذا أردنا مقارنة التشتت فى مجموعتين ، يجب أن نقيسه فيهما بطريقة واحدة ثم نقارن النتيجتين . ولا نقارن مقياس التشتت فى إحداهما محسوباً بطريقة معينة ، بمقياس التشتت فى المجموعة الثانية محسوباً بطريقة أخرى .

قياس التشتت حول قيمة معينة

١٧٢ — فى الطريقتين المتقدمتين نقيس تشتت القيم فيما بينها . ولكننا فى دراسة المتوسطات ، التى نستخدمها كنماذج تمثل المجموعات ، نحتاج إلى معرفة تشتت المجموعة حول هذه المتوسطات ، وخصوصاً الوسط الحسابى .

ولقياس التشتت حول أى قيمة معينة نتخذ هذه القيمة كمركز ، ونبحث فى الفروق بينها وبين مفردات القيم كل على حدة . والتشتت حول هذه القيمة يكون كبيراً أو صغيراً حسب ما تكون هذه الفروق كبيرة أو صغيرة فى مجموعها . وهناك فكرتان للبحث فى هذه الفروق واستخدامها لقياس التشتت . والآن نشرح هاتين الفكرتين والطريقة العملية لتطبيق كل منهما .

الفرق بين أى قيمة فى مجموعة ومتوسط هذه المجموعة نسميه عادة **انحراف** هذه النقطة عن المتوسط.(١)

(١) بالانجليزية Deviation from the Mean

طريقة الانحراف المتوسط

١٧٣ — الطريقة الأولى هى المسماة [١] بطريقة **الانحراف المتوسط**، وفيها نوجد انحراف كل قيمة فى المجموعة عن الوسط الحسابى لها ؛ ونجرد هذه الانحرافات من الإشارات الجبرية ونجمعها ؛ ثم نقسم حاصل الجمع على عدد هذه الانحرافات . وبالطبع هذا العدد يساوى تماماً عدد القيم الأصلية فى المجموعة . وبذلك نحصل على الانحراف المتوسط المطلوب وهو مقياس لتشتت المجموعة حول وسطها الحسابى .

لنأخذ مثلا مجموعة الأجور الآتية (بالقروش) :

٩ ، ١٠٫٥ ، ١١ ، ١٢٫٥ ، ١٣ ، ١٣٫٥ ، ١٤٫٥

الوسط الحسابى لهذه الأجور ١٢ قرشاً وانحرافاتها عن هذا الوسط ، بدون إشارات ، هى على الترتيب :

٣ ، ١٫٥ ، ١ ، ٠٫٥ ، ١ ، ١٫٥ ، ٢٫٥ .

والوسط الحسابى لهذه الانحرافات (مجموعها مقسوماً على عددها) ، يساوى ١٫٥٧ وهو الانحراف المتوسط المطلوب ، وهو التشتت حول الوسط الحسابى .

والسر فى أننا أهملنا إشارات هذه الانحرافات ، هو أننا ننظر إلى الانحراف باعتباره مجرد فرق بين القيمة والمتوسط ، بصرف النظر عن كون هذا الفرق بالنقص أو بالزيادة ، لأن التشتت الذى نريد قياسه لا يميز بين النقص والزيادة عن المتوسط ، بل يهتم فقط بمقدار البعد عنه .

الانحراف المتوسط فى التوزيع التكرارى

١٧٤ — فى المثال السابق أخذنا مجموعة من القيم عددها صغير . والآن نشرح كيفية إيجاد الانحراف المتوسط لمجموعة كبيرة مقسمة إلى فئات فى جدول تكرارى .

(١) بالانجليزية Mean Deviation

لنأخذ نفس التوزيع التكرارى لأعمار ١٧٣٩ تلميذاً ، ونحسب الانحراف المتوسط للأعمار (عن الوسط الحسابى) .

جدول ٢٦ — إيجاد الانحراف المتوسط للأعمار

الأعمار	التكرارات ك	الانحرافات عن ١٨	انحرافات ع بدون إشارة	ضرب ع . ك
١٤	٣٤	—٤	٤	١٣٦
١٥	١٢٨	—٣	٣	٣٨٤
١٦	٢٦٢	—٢	٢	٥٢٤
١٧	٣٦٠	—١	١	٣٦٠
١٨	٣٨٦	٠	٠	٠
١٩	٢٩٤	١	١	٢٩٤
٢٠	١٦٧	٢	٢	٣٣٤
٢١	٩٢	٣	٣	٢٧٦
٢٢	١٦	٤	٤	٦٤
	١٧٣٩			٢٣٧٢

الوسط الحسابى لهذه الأعمار يساوى ١٧و٧٤٩ سنة (كما رأينا فى بند ١٣٢ صفحة ١٤٥) فلإيجاد الانحراف المتوسط يجب أن نطرح هذا الوسط الحسابى من كل من القيم ١٤ و ١٥ و ١٦ و ٠٠ و ٢١ و ٢٢ ، لكى نحصل على انحرافاتها عنه (وهى ، بدون إشارة = ٣٫٧٤٩ ، ٢٫٧٤٩ ، ١٫٧٤٩ ، ٠٠٠ ، ٣٫٢٥١ ، ٤٫٢٥١ على الترتيب) . ثم بعد ذلك نضرب كلا من هذه الانحرافات فى التكرار المناظر له ، ونجمع ونقسم الناتج على ١٧٣٩ ينتج الانحراف المتوسط .

ولكن هذه الطريقة تكون متعبة جداً لأن هذه الانحرافات كثيرة الأرقام

وعمليات ضربها فى تكراراتها تكون مرهقة . ولذلك نستخدم وسطاً فرضياً كما فعلنا فى إيجاد الوسط الحسابى . نحسب الانحرافات عن هذا الوسط الفرضى ، ونجردها من إشاراتها ؛ ثم نضربها فى التكرارات المناظرة ، ونجمع الحواصل كما هو مبين فى الجدول السابق .

العدد الناتج ، وهو ٢٣٧٢ ، يساوى مجموع انحرافات القيم المعطاة عن الوسط الفرضى ١٨ ، ولكن المطلوب هو مجموع الانحرافات عن الوسط الحسابى ١٧٫٧٤٩ ، وعلى ذلك يجب تصحيح الرقم ٢٣٧٢ الذى حصلنا عليه . وهذا التصحيح يكون كما يأتى :

الأعمار التى أصغر من الوسط الحسابى ، وهى ١٤ و ١٥ و ١٦ و ١٧ (ومجموع تكراراتها = ٣٤ + ١٢٨ + ٢٦٢ + ٣٦٠ = ٧٨٤) طرحناها فى الجدول من العدد ١٨ ، وكان يجب طرحها من ١٧٫٧٤٩ ، وعلى ذلك فكل من الانحرافات ح التى حصلنا عليها فى الجدول أمام هذه الأعمار أكبر من اللزوم بمقدار ٠٫٢٥١ (أى ١٨٫٠ — ١٧٫٧٤٩) . ومجموعها إذن أكبر من اللزوم بمقدار ٧٨٤ × ٠٫٢٥١

أما الأعمار التى أكبر من الوسط الحسابى ، وهى ١٨ و ١٩ و ٢٠ و ٢١ و ٢٢ (ومجموع تكراراتها = ٣٨٦ + ٢٩٤ + ١٦٧ + ٩٢ + ١٦ = ٩٥٥) فقد طرحنا منها العدد ١٨ فى الجدول . وكان يجب أن نطرح منها ١٧٫٧٤٩ . فالانحرافات ح التى حصلنا عليها أمام هذه الأعمار ، كل منها أقل من اللزوم بمقدار ٠٫٢٥١ (أى الفرق بين الوسطين الفرضى والحسابى) ومجموعها إذن أقل من اللزوم بمقدار ٩٥٥ × ٠٫٢٥١

ويكون إذن المجموع الذى حصلنا عليه أكبر من اللزوم بمقدار ٧٨٤ × ٢٥١ و ٠ ، وأقل من اللزوم بمقدار ٩٥٥ × ٢٥١ و ٠

∴ ٢٣٧٢ — ٧٨٤ × ٢٥١ر + ٩٥٥ × ٢٥١ر٠

هو مجموع الانحرافات عن الوسط الحسابى المطلوب.

وعلى العموم يمكننا استنباط القاعدة الآتية من هذا المثال المتقدم وهى :

لنفرض أن :

$\bar{\text{س}}$ = الوسط الحسابى لأى توزيع تكرارى =١٧ر٧٤٩ هنا

و = الوسط الفرضى لهذا التوزيع = ١٨ر٠ «

ع = الوسط الحسابى—الوسط الفرضى ، أى $\bar{\text{س}}$ — و =—٢٥١ر٠ «

ك$_1$ = عدد القيم التى هى أصغر من $\bar{\text{س}}$ = ٧٨٤ «

ك$_2$ = « « « « أكبر من $\bar{\text{س}}$ = ٩٥٥ «

ح = مجموع الانحرافات عن الوسط الحسابى (بدون إشارة) = ؟ «

م = « « « « الفرضى « = ٢٣٧٢ «

∴ ح = م + ع (ك$_1$ — ك$_2$) . . . . (١)

وبالتعويض عن م و ع ، ك$_1$ وك$_2$ بقيمها فى هذا المثال ،

∴ ح = ٢٣٧٢ — ٢٥١ ر (٧٨٤ — ٩٥٥)

= ٢٣٧٢ + ٩٢١ ر ٤٢

= ٩٢١ ر ٢٤١٤

∴ الانحراف المتوسط = $\frac{\text{٢٤١٤,٩٢١}}{\text{١٧٣٩}}$

= ١,٣٩ سنة ؛

وهو مقياس التشتت المطلوب :

يجب أن نختار الوسط الفرضى فى نفس الفئة مع الحسابى

١٧٥ — عند تطبيق هذه العلاقة (١) المذكورة أعلاه ، يجب أن يكون الوسط الحسابى والوسط الفرضى كلاهما داخل حدود فئة واحدة ، وإلا كان هناك ثلاث مجموعات من القيم : مجموعة أصغر من الوسط الحسابى والفرضى معاً ،

ومجموعة أخرى أكبر من الوسط الحسابى والفرضى معاً ، ومجموعة ثالثة واقعة بين الوسطين أكبر من أحدهما وأقل من الآخر . وفى هذه المجموعات يكون التصحيح مخالفاً للنظام الذى ذكرناه فى هذا المثال واستنبطنا منه القاعدة المشار إليها .

الانحراف المتوسط عن الأوساط الأخرى

١٧٦ — يمكن بنفس هذه الطريقة طبعاً أن نحسب الانحراف المتوسط عن الوسيط أو المنوال أو أى متوسط آخر ، مع مراعاة وضع قيمة هذا المتوسط بدل قيمة الوسط الحسابى فى العلاقة (١) بند ١٧٤ ، ومراعاة الشرط المذكور فى بند ١٧٥ .

وإذا حسبنا الانحراف عن الوسيط فسنجده أقل من الانحراف المتوسط عن الوسط الحسابى أو أى متوسط آخر . وقد ذكرنا السبب فى ذلك فى آخر الباب السابق (بند ١٦٧) ، وهو أن مجموع أبعاد القيم فى مجموعة عن وسيطها (وهو مجموع الانحرافات بدون إشارة) أقل من مجموع أبعادها عن أى متوسط آخر . وبالتالى يكون الانحراف المتوسط عن الوسيط له نفس الصفة .

تعريف الانحراف المعيارى

١٧٧ — أهم مقاييس التشتت وأكثرها استعمالا هو المسمى (١) **الانحراف المعيارى** . وهو مقياس للتشتت حول الوسط الحسابى . وهو عبارة عن الجذر التربيعى لمتوسط مربعات انحرافات القيم عن الوسط الحسابى .

لنفرض على العموم أن عدد القيم فى مجموعة إحصائية هو ن ، وأن هذه القيم هى :

$$س_١ ، س_٢ ، س_٣ ، . . . ، س_ن$$

---

(١) يسمى بالانجليزية Standard Deviation . ويسمى أحياناً جذر متوسط مربع الخطأ Root-Mean Square Error

وليكن $\bar{س}$ هو الوسط الحسابى لهذه المجموعة .

∴ انحرافات هذه القيم عن الوسط الحسابى تكون :

$س_١ - \bar{س}$ ، $س_٢ - \bar{س}$ ، $س_٣ - \bar{س}$ ، ... ، $س_ن - \bar{س}$.

نربع هذه الانحرافات ونقسم مجموع هذه المربعات على عددها ، وهو ن طبعاً ، فنحصل على متوسط مربعات الانحرافات عن الوسط الحسابى ، وهو

$$= \frac{١}{ن} [(س_١ - \bar{س})^٢ + (س_٢ - \bar{س})^٢ + \ldots + (س_ن - \bar{س})^٢]$$

$$= \frac{\sum (س - \bar{س})^٢}{ن} \quad \ldots\ldots\ldots\ldots (١)$$

∴ الانحراف المعيارى لقيم س ، ونرمز له بالحرف ع[1] ، هو

$$ع = \sqrt{\frac{\sum (س - \bar{س})^٢}{ن}} \quad \ldots\ldots\ldots (٢)$$

الكمية $ع^٢$ نسميها[2] **التباين** .

١٧٨ — السر فى تربيع الانحرافات هنا هو أننا نريد التخلص من الإشارات السالبة فى بعض الانحرافات ، وهى التى أهملناها عند حساب الانحراف المتوسط ، للسبب الذى ذكرناه حينئذ (بند ١٧٣)

نربع الانحرافات لنتخلص من الإشارة ثم نأخذ جذر المتوسط لنرجع للوحدات الأصلية

والسبب فى أخذ الجذر التربيعى فى المعادلة (٢) لإيجاد الانحراف المعيارى ، هو أننا نريد الرجوع إلى الوحدات الأصلية بعد تربيع الانحرافات .

---

(١) فى الكتب الأفرنجية يرمز للانحراف المعيارى بالحرف الأغريقى «σ» واسمه سيجما (Sigma) .

(٢) بالإنجليزية (Variance) ؛ ويرمز له أحياناً بالحرف الأغريقى u (ميو) .

وذلك ليكون التشتت مقيساً بنفس الوحـــدات المقيسة بها القيم المعطاة $س_١$ ، $س_٢$ ، . . . . ، $س_ن$. وقد لاحظنا فى المثال الذى أخذناه فى بند ١٦١ أن مقياس تشتت الأعمار معبر عنه بالسنين ، وهى نفس الوحدات المقيسة بها الأعمار .

الانحراف المعيارى لتوزيع تكرارى

١٧٩ — المعادلة (٢) فى بند ١٦٦ تعبر عن الانحراف المعيارى لمجموعة من المفردات $س_١$ ، $س_٢$ ، . . . . ، $س_ن$ بصرف النظر عن كونها تحتوى على بعض القيم المتساوية أى المتكررة ، أو أن جميع القيم مختلفة . وطبعاً إذا كان هناك عدد من القيم المتساوية ، أى قيمة متكررة عدداً معيناً من المرات ، فإننا نضرب مربع انحراف هذه القيمة فى عدد مرات التكرار ؛ ومجموع حواصل الضرب يساوى مجموع مربعات انحرافات هذه القيم المتساوية .

وعلى ذلك فلايجاد الانحراف المعيارى فى جدول تكرارى ، نضرب تكرار كل فئة فى مربع انحراف مركزها عن الوسط الحسابى ؛ ومجموع هذه الحواصل يساوى مجموع مربعات انحرافات القيم الأصلية عن الوسط . و بقسمة هذا على عدد المفردات كلها ، أى مجموع التكرارات ، واستخراج الجذر التربيعى للناتج ، نحصل على الانحراف المعيارى المطلوب .

ولتوضيح ذلك عملياً نحسب الانحراف المعيارى للتوزيع التكرارى الذى أخذناه سابقاً ، وهو توزيع أعمار ١٧٣٩ تلميذاً .

نرى الخطوات موضحة فى جدول ٢٧ ، حيث نجد فى العمود (١) الحدود الدنيا لفئات الأعمار ؛ وفى العمود (٢) نجد مراكز هـذه الفئات ؛ وفى العمود (٣) نجد التكرارات ك . وبما أن الوسط الحسابى لأعمار هؤلاء التلاميذ هو ١٧٫٧٤٩ سنة (كما وجدنا فى بند ١٣٣ صفحة ١٤٦) ، فلايجاد انحرافات القيم عن هذا الوسط نطرح ١٧٫٧٤٩ من القيم ١٤ ، ١٥ ، ١٦ ، ... على التوالى ، فنحصل على الانحرافات المكتوبة فى عمود (٤) .

ولكننا نرى أن هذه الانحرافات عن الوسط الحسابى مكونة من أرقام كثيرة وتربيعها سيعطينا أرقاماً أكثر تعقيداً . هذا فضلا عن أن عمليات ضرب هذه الأعداد الكبيرة فى نفسها ، ثم ضرب النواتج فى التكرارات ، تكون شاقة وتستغرق وقتاً طويلا . ولذلك نختصر العمل ونترك هذه الانحرافات عن الوسط

جدول ٢٧ — حساب الانحراف المعيارى لأعمار

١٧٣٩ تلميذاً من الجدول التكرارى

الفئات	مراكز الفئات س	التكرارات ك	انحرافات س عن الوسط الحسابى	انحرافات عن وسط فرضى : ع	(٣) × (٥) ع × ك	(٦) × (٥) ع^{2} . ك
(١)	(٢)	(٣)	(٤)	(٥)	(٦)	(٧)
١٣ر٥—	١٤	٣٤	— ٣ر٧٤٩	— ٤	—١٣٦	٥٤٤
١٤ر٥—	١٥	١٢٨	— ٢ر٧٤٩	— ٣	—٣٨٤	١١٥٢
١٥ر٥—	١٦	٢٦٢	— ١ر٧٤٩	— ٢	—٥٢٤	١٠٤٨
١٦ر٥—	١٧	٣٦٠	— ر٧٤٩	— ١	—٣٦٠	٣٦٠
١٧ر٥—	١٨	٣٨٦	ر٢٥١	٠	٠	٠
١٨ر٥—	١٩	٢٩٤	١ر٢٥١	١	٢٩٤	٢٩٤
١٩ر٥—	٢٠	١٦٧	٢ر٢٥١	٢	٣٣٤	٦٦٨
٢٠ر٥—	٢١	٩٢	٣ر٢٥١	٣	٢٧٨	٨٢٨
٢١ر٥—	٢٢	١٦	٤ر٢٥١	٤	٦٤	٢٥٦
		١٧٣٩			—٤٣٦	٥١٥٠

الحسابى ونختار وسطاً فرضياً مناسباً . ونحسب الانحرافات عنه ، كما فعلنا عند حساب

الانحراف المتوسط ، ونجد فى العمود (٦) حواصل ضرب هذه الانحرافات — فى التكرارات ؛ وفى عمود (٧) نجد حواصل ضرب مربعاتها فى التكرارات ؛ ومن هذا نحصل على مجموع مربعات الانحرافات عن هذا الوسط الفرضى، وهو ٥١٥٠

ولمعرفة مجموع مربعات الانحرافات عن هـذا الوسط الحسابى ، نستخدم العلاقة [1] الآتية :

---

(١) لإثبات هذه العلاقة نفرض أن مفردات القيم هى : $\text{س}_{١}$ ، $\text{س}_{٢}$ ، ...، $\text{س}_{\text{ن}}$ ؛ وأن الوسط الحسابى لها يساوى $\bar{\text{س}}$ ، وأن الوسط الفرضى يساوى و مثلا .

$\therefore\ \bar{\text{س}} = \frac{١}{\text{ن}}(\text{مجـ س})$ ؛ أى $\text{مجـ س} - \text{ن}\,\bar{\text{س}} = ٠$.

نضع $\text{ح} = \bar{\text{س}} - \text{و}$ = الفرق بين الوسطين الحسابى والفرضى .

$\therefore$ انحراف القيمة $\text{س}_{١}$ عن الوسط الفرضى هو $\text{س}_{١} - \text{و}$.

لكن $\text{س}_{١} - \text{و} = \text{س}_{١} - \bar{\text{س}} + \bar{\text{س}} - \text{و} = \text{س}_{١} - \bar{\text{س}} + \text{ح}$

$\therefore\ (\text{س}_{١} - \text{و})^{٢} = (\text{س}_{١} - \bar{\text{س}})^{٢} + \text{ح}^{٢} + ٢\,\text{ح}\,(\text{س}_{١} - \bar{\text{س}})$ ؛

وبالمثل $(\text{س}_{٢} - \text{و})^{٢} = (\text{س}_{٢} - \bar{\text{س}})^{٢} + \text{ح}^{٢} + ٢\,\text{ح}\,(\text{س}_{٢} - \bar{\text{س}})$ ؛

» . . . . . .

» $(\text{س}_{\text{ن}} - \text{و})^{٢} = (\text{س}_{\text{ن}} - \bar{\text{س}})^{٢} + \text{ح}^{٢} + ٢\,\text{ح}\,(\text{س}_{\text{ن}} - \bar{\text{س}})$

وبجمع هذه المتساويات

$\therefore\ \text{مجـ}(\text{س} - \text{و})^{٢} = \text{مجـ}(\text{س} - \bar{\text{س}})^{٢} + \text{ن}\,\text{ح}^{٢} + ٢\,\text{ح}\,(\text{مجـ س} - \text{ن}\,\bar{\text{س}})$.

ولكن $\text{مجـ س} - \text{ن}\,\bar{\text{س}} = ٠$. فإذا كتبنا $\text{ن}\,\text{ل}^{٢}$ بدلا من $\text{مجـ}(\text{س} - \text{و})^{٢}$ للاختصار

و $\text{مجـ}(\text{س} - \bar{\text{س}})^{٢} = \text{ن}\,\text{ع}^{٢}$

$\therefore\ \text{ن}\,\text{ل}^{٢} = \text{ن}\,\text{ع}^{٢} + \text{ن}\,\text{ح}^{٢}$

أو $\text{ل}^{٢} = \text{ع}^{٢} + \text{ح}^{٢}$

وينتج من هذا أيضاً أن $\text{ن}\,\text{ل}^{٢} > \text{ن}\,\text{ع}^{٢}$ دائماً ، إلا إذا كان $\text{ح} = ٠$.

أى أن مجموع مربعات الانحرافات عن الوسـط الحسابى دائماً أصغر من مجموع مربعات الانحرافات عن أى وسط فرضى ، إلا إذا كان الوسطان متساويين . وهذه خاصة للوسط الحسابى تقابلها خاصة مثلها للوسيط ذكرناها فى بند ١٧٤ .

$$\text{ﻣ ل}^{٢} = \text{ﻣ ع}^{٢} + \text{ﻣ ح}^{٢}$$

حيث ﻣ ل^{٢} = مجموع مربعات انحرافات القيم عن الوسط الفرضى = ٥١٥٠ هنا
و ﻣ ع^{٢} = « « « « الحسابى = ؟ «
و ح = الفرق بين الوسطين (الحسابى — الفرضى) = — ٢٥١ر« «
و ﻣ = عدد المفردات جميعها ، أى مجموع التكرارات = ١٧٣٩ «

∴ ٥١٥٠ = ١٧٣٩ ع^{٢} + ١٧٣٩ (— ٢٥١ ر)^{٢}

∴ $\frac{٥١٥٠}{١٧٣٩}$ = ع^{٢} + ٠٦٣٠٠١ر

∴ ع^{٢} = ٢و٨٩٨٥

∴ ع = ١و٧٠٢ سنة ؛

وهو الانحراف المعيارى للأعمار فى هذه المجموعة .

نلاحظ أن النسبة بين الانحراف المتوسط (بند ١٧٢) والانحراف المعيارى لهذه المجموعة ، تساوى $\frac{٤}{٥}$ تقريباً . وهذه خاصة معروفة من خواص التوزيعات المتماثلة أو القريبة من التماثل .

طريقة تغيير الوحدات

١٨٠ — إذا كانت فترات الفئات فى التوزيع التكرارى غير الواحد الصحيح فيمكننا أن نستخدم وحدات جديدة للانحرافات ، كما فعلنا فى بند ١٣٢ (صحيفة ١٤٧) . ولبيان ذلك نأخذ نفس التوزيع التكرارى للأجور المذكور فى جدول ٢٢ هناك .

فى هذا الجدول نجد فئات الأجور فى عمود (١) ومراكزها فى عمود (٢) وتكراراتها فى العمود (٣) . ونجد الانحرافات عن الوسط الفرضى ١٩ر٥ ، فى العمود (٤) . ونلاحظ هنا أن هذه الانحرافات تحتوى على العدد ٣ كعامل

مشترك (وهو يساوى طول فترة الفئات) ؛ فنقسم هذه الانحرافات على هذا العامل فنحصل على الانحرافات الجديدة فى العمود (٥) ، وهى مقيسة بوحدات كل منها تساوى ثلاثة قروش .

جدول ٢٧ — حساب الانحراف المعيارى لأجور
٧٤٣٢ عاملا ، بالطريقة المختصرة

فئات الأجر بالقرش	مراكز الفئات	التكرار ك	الانحرافات عن وسط فرضى: ع	الانحرافات الجديدة ٠ ف	(٣) × (٥) ك × ف	(٥) × (٦) ك ف ٢
(١)	(٢)	(٣)	(٤)	(٥)	(٦)	(٧)
١٢—	١٣٫٥	٢٢٤	—٦	—٢	—٤٤٨	٨٩٦
١٥—	١٦٫٥	١٩٧١	—٣	—١	—١٩٧١	١٩٧١
١٨—	١٩٫٥	٣٧٥٥	٠	٠	٠	٠
٢١—	٢٢٫٥	١٢٣٦	٣	١	١٢٣٦	١٢٣٦
٢٤—	٢٥٫٥	١٩٦	٦	٢	٣٩٢	٧٨٤
٢٧—	٢٨٫٥	٤٠	٩	٣	١٢٠	٣٦٠
٣٠—	٣١٫٥	١٠	١٢	٤	٤٠	١٦٠
		٧٤٣٢			—٦٣١	٥٤٠٧

ونجد فى العمود (٦) حواصل ضرب هـذه الانحرافات فى التكرارات . ومجموع هذا العمود نستخدمه لإيجاد الفرق بين الوسط الفرضى والوسط الحسابى ، لمعرفة هذا الأخير .

ونستعمل أرقام هذا العمود أيضاً كخطوة أولى للحصول على أرقام العمود (٧) وهى عبارة عن مربعات الانحرافات مضروبة فى التكرارات . ومجموع هذا العمود هو مجموع مربعات الانحرافات عن الوسط الفرضى ، بوحدات جديدة

ولإيجاد الفرق بين الوسطين بالوحدات الجديدة ، نقسم مجموع العمود (٦) على مجموع التكرارات فى العمود (٣) .

∴ الفرق بين الوسطين $= \text{ح}$

$= - \frac{\text{٦٣١}}{\text{٧٤٣٢}}$

$= - \text{٠٨٤٩ر}$ بالوحدات الجديدة .

وباستخدام العلاقة $\text{ن ل}^{2} = \text{ن ع}^{2} + \text{ن ح}^{2}$ ،

∴ $\text{٥٤٠٧} = \text{٧٤٣٢ ع}^{2} + \text{٧٤٣٢} (- \text{٠٨٤٩ر})^{2}$ بالوحدات الجديدة

∴ $\frac{\text{٥٤٠٧}}{\text{٧٤٣٢}} = \text{ع}^{2} + \text{٠٠٧٢ر}$

∴ $\text{ع}^{2} = \text{٧٢٧٦ر} - \text{٠٠٧٢ر}$

$= \text{٧٢٠٤ر}$

∴ $\text{ع} = \text{٨٥ر}$ تقريباً ، بالوحدات الجديدة .

∴ $= \text{٣} \times \text{٨٥ر}$ « الأصلية

$= \text{٥٥و٢}$ قرشاً ؛

وهو الانحراف المعيارى للأجور .

مجموع مربعات القيم

١٨١ — ويمكننا إيجاد الانحراف المعيارى لمجموعة من القيم إذا عرفنا مجموع مربعات هذه القيم ، والوسط الحسابى لهذه القيم . ولبيان ذلك نفرض أن القيم هى على العموم :

س_{1} ، س_{2} ، س_{3} ، . . . . . . ، $\text{س}_{\text{ن}}$ ؛

وأن وسطها الحسابى $\bar{\text{س}}$ ، حيث $\text{ن} \bar{\text{س}} = \text{مجـ س}$ ؛ وأن الانحراف المعيارى $\text{ع}_{\text{س}}$. انحراف القيمة س_{1} عن الوسط الحسابى يساوى $\text{س}_{1} - \bar{\text{س}}$.

$$\therefore\ (\text{س}_1 - \bar{\text{س}})^2 = \text{س}_1^2 + \bar{\text{س}}^2 - ٢\,\text{س}_1\,\bar{\text{س}}$$

$$\therefore\ (\text{س}_2 - \bar{\text{س}})^2 = \text{س}_2^2 + \bar{\text{س}}^2 - ٢\,\text{س}_2\,\bar{\text{س}}$$

. . . .

$$\therefore\ (\text{س}_{\text{ن}} - \bar{\text{س}})^2 = \text{س}_{\text{ن}}^2 + \bar{\text{س}}^2 - ٢\,\text{س}_{\text{ن}}\,\bar{\text{س}}$$ ؛ وبالجمع

$$\therefore\ \text{مج}(\text{س} - \quad = \text{مج}\,\text{س}^2 + \text{ن}\,\bar{\text{س}}^2 - ٢\,\bar{\text{س}}\,\text{مج}\,\text{س} .$$

$$\therefore\ \text{ن}\,\text{ع}^2 = \text{مج}\,\text{س}^2 + \text{ن}\,\bar{\text{س}}^2 - ٢\,\bar{\text{س}}\,(\text{ن}\,\bar{\text{س}}) .$$

$$\therefore\ = \text{مج}\,\text{س}^2 - \text{ن}\,\bar{\text{س}}^2 \text{ ، } . . . . \quad (١)$$

$$\therefore\ \text{مج}\,\text{س}^2 = \text{ن}\,\text{ع}^2 + \text{ن}\,\bar{\text{س}}^2 \quad (٢) .$$

$$\text{أى}\quad \tfrac{١}{\text{ن}}\,\text{مج}\,\text{س}^2 = \text{ع}^2 + \bar{\text{س}}^2 \ . . . \quad (٣) .$$

وهذه المعادلة (٣) مهمة ، وسنحتاج إلى استخدامها فى كثير من الأحيان

ويلاحظ أنه من الممكن استنباطها من المعادلة المذكورة فى بند ١٧٩ ، بوضع الوسط الفرضى و $=$. ، وعليه يكون الفرق بين الوسطين ح $= \bar{\text{س}} -. = \bar{\text{س}}$.

ويتضح من هذا أننا لو علمنا مجموع مربعات القيم ومتوسطها الحسابى ، أمكننا إيجاد انحرافها المعيارى بسهولة .

الانحراف المعيارى لمجموعة كبيرة مركبة من مجموعات صغيرة

١٨٢ — إذا علمنا الانحرافات المعيارية لعدد من المجموعات ، أمكننا استنباط الانحراف المعيارى للمجموعة الكلية المكونة من هذه المجموعات . فمثلا لو تكونت مدرسة من أربع فرق ، وعلمنا الانحراف المعيارى لأعمار كل فرقة على حدة ، وكذلك الوسط الحسابى لأعمارها ، وعدد التلاميذ فيها يمكننا بسهولة استنباط الانحراف المعيارى للأعمار فى المدرسة كلها من هذه المعلومات . وذلك باستخدام العلاقة الآتية ، وهى تربط الانحراف المعيارى العمومى (للمجموعة الكلية) بالانحرافات المعيارية للمجموعات الجزئية ومتوسطاتها الحسابية . وهذه العلاقه هى :

$$\text{ں}\,\text{ع}_{\text{س}}^{٢} = \text{مجـ}\,\text{ن}\,\text{ع}^{٢} + \text{ں}\,\text{ع}_{\text{م}}^{٢} \quad \ldots \qquad (١)$$

حيث ں = العدد الإجمالى للمفردات فى المجموعة الكلية ،

و $\text{ع}_{\text{س}}$ = الانحراف المعيارى للمجموعة الكلية ،

و $\text{ن}_{١}$ ، $\text{ن}_{٢}$ ، $\text{ن}_{٣}$ ، . . . ، $\text{ن}_{\text{ر}}$ عدد المفردات فى المجموعات الجزئية الأولى والثانية ..،

و $\text{ع}_{١}$ ، $\text{ع}_{٢}$ ، $\text{ع}_{٣}$ ، . . . ، $\text{ع}_{\text{ر}}$ الانحرافات المعيارية لهذه المجموعات الجزئية ،

و $\text{ع}_{\text{م}}$ الانحراف المعيارى للمتوسطات الحسابية للمجموعات،

$$\text{و ں} = \text{ن}_{١} + \text{ن}_{٢} + \text{ن}_{٣} \ldots + \text{ن}_{\text{ر}} = \text{مجـ}\,\text{ن}$$

ولإثبات صحة هذه العلاقة نفرض أن مفردات المجموعات هى كما يأتى :

الأولى : $\text{ا}_{١}$ ، $\text{ا}_{٢}$ ، ... ، $\text{ا}_{\text{ن}_{١}}$ ، عددها $\text{ن}_{١}$ ، ووسطها الحسابى $\text{م}_{١}$ ، وانحرافها المعيارى $\text{ع}_{١}$

الثانية : $\text{ب}_{١}$ ، $\text{ب}_{٢}$ ، ... ، $\text{ب}_{\text{ن}_{٢}}$ » $\text{ن}_{٢}$ » » $\text{م}_{٢}$ » » $\text{ع}_{٢}$

الثالثة : $\text{حـ}_{١}$ ، $\text{حـ}_{٢}$ ، ... ، $\text{حـ}_{\text{ن}_{٣}}$ » $\text{ن}_{٣}$ » » $\text{م}_{٣}$ » » $\text{ع}_{٣}$

. . . .

الأخيرة : $\text{ى}_{١}$ ، $\text{ى}_{٢}$ ، ... ، $\text{ى}_{\text{ن}_{\text{ر}}}$ » $\text{ن}_{\text{ر}}$ » » $\text{م}_{\text{ر}}$ » » $\text{ع}_{\text{ر}}$

$\therefore$ $\text{مجـ}\,\text{ا} = \text{ن}_{١}\,\text{م}_{١}$ ، و $\text{مجـ}(\text{ا} - \text{م}_{١})^{٢} = \text{ن}_{١}\,\text{ع}_{١}^{٢}$ ؛ وبالمثل للمجموعات الأخرى .

الوسط الحسابى للمجموعة الكلية

$$= \frac{\text{مجـ}\,\text{ا} + \text{مجـ}\,\text{ب} + \ldots + \text{مجـ}\,\text{ى}}{\text{ن}_{١} + \text{ن}_{٢} + \ldots + \text{ن}_{\text{ر}}} = \overline{\text{س}} \text{ ، مثلا ،}$$

$$= \frac{\text{ن}_{١}\,\text{م}_{١} + \text{ن}_{٢}\,\text{م}_{٢} + \ldots + \text{ن}_{\text{ر}}\,\text{م}_{\text{ر}}}{\text{ن}_{١} + \text{ن}_{٢} + \ldots + \text{ن}_{\text{ر}}}$$ وهو الوسط الحسابى للميات .

أى أن الوسط الحسابى للمجموعة الكلية هو فى الوقت نفسه الوسط الحسابى لمتوسطات المجموعات الجزئية ، وهو المنتظر بطبيعة الحال.

∴ الانحراف المعيارى لهذه المتوسطات بين نفسها هو $ع_م$ حيث

$$ه\, ع_م^{٢} = ن_{١}\,(م_{١} - \bar{س})^{٢} + ن_{٢}\,(م_{٢} - \bar{س})^{٢} + \dots$$
$$+\, ن_{ه}\,(م_{ه} - \bar{س})^{٢}$$
$$= مجـ\, ن\,(م - \bar{س})^{٢} \quad \dots \qquad (٢)$$

الانحراف المعيارى المجموعة الأولى هو $ع_{١}$ ، حيث

$$ن_{١} ع_{١}^{٢} = مجـ\,(ا - م_{١})^{٢}.$$

وإذا أخذنا $\bar{س}$ كوسط فرضى لهذه المجموعة ، ووضعنا $ح_{١} = م_{١} - \bar{س}$ ،

ورمزنا لمجموع مربعات الانحرافات عن الوسط الفرضى بالرمز $ن_{١} ل_{١}^{٢}$

$$\therefore\ مجـ\,(ا - \bar{س})^{٢} = ن_{١} ل_{١}^{٢} = ن_{١} ع_{١}^{٢} + ن_{١} ح_{١}^{٢} \dots \ (\text{بند } ١٧٧)$$
$$= ن_{١} ع_{١}^{٢} + ن_{١}\,(م_{١} - \bar{س})^{٢}$$

$$\text{وكذلك } مجـ\,(ب - \bar{س})^{٢} = ن_{٢} ع_{٢}^{٢} + ن_{٢}\,(م_{٢} - \bar{س})^{٢}$$
$$\text{»}\quad مجـ\,(حـ - \bar{س})^{٢} = ن_{٣} ع_{٣}^{٢} + ن_{٣}\,(م_{٣} - \bar{س})^{٢}$$

$$\dots \qquad \dots \qquad \dots$$

$$\text{»}\quad مجـ\,(ى - \bar{س})^{٢} = ن_{ه} ع_{ه}^{٢} + ن_{ه}\,(م_{ه} - \bar{س})^{٢}.$$

يلاحظ أن مجموع الأطراف اليمنى لهذه المتساويات يساوى مجموع مربعات انحرافات جميع المفردات ا ، ب ، . . ، ى عن الوسط العام $\bar{س}$ ، ويساوى $ه\, ع_{س}^{٢}$.

$$\therefore\ ه\, ع_{س}^{٢} = مجـ\, ن\, ع^{٢} + مجـ\, ن\,(م - \bar{س})^{٢}؛ \ \text{وباستخدام العلاقة } (٢)$$

$$\therefore\ = مجـ\, ن\, ع^{٢} + ه\, ع_م^{٢}،$$

وهى النتيجة المطلوبة .

$$\therefore\ مجـ\, ن\, ع^{٢} = ه\, ع_{س}^{٢} - ه\, ع_م^{٢}، \quad \dots \qquad (٣).$$

وهذا معناه أن مجموع مربعات انحرافات المفردات ، كل واحدة عن الوسط

الحسابي لمجموعتها ، يساوى مجموع مربعات انحرافاتها عن الوسط الحسابي العمومى ناقصاً مجموع مربعات انحرافات الأوساط الحسابية للمجموعات عن هذا الوسط العمومى نفسه . وهذه نتيجة مهمة سوف سنحتاج إليها فيما بعد .

ولو قسمنا طرفى هذه المتساوية على $ك = مجـ\ ن$ ينتج أن

$$ع^2_{س} = \frac{مجـ\ ن\ ع^2}{مجـ\ ن} + ع^2_{م} \qquad (٤) .$$

وباستخدام تعريف التباين المذكور فى ص ١٩٢ يكون $ع^2_{س}$ هو التباين العام لجميع المفردات ، و $ع^2_{م}$ هو تباين المتوسطات ، ويكون $\frac{مجـ\ ن\ ع^2}{مجـ\ ن}$ هو المتوسط المرجح لتباينات المجموعات الجزئية . وبذلك يمكن التعبير عن هذه العلاقة الأخيرة (٤) بالعبارة الآتية :

التباين العام = متوسط التباينات + تباين المتوسطات ؛

وهذه العلاقة لها تطبيقات عملية مهمة نراها فيما بعد .

عزوم التوزيع التكرارى

١٨٣ — لنفرض مجموعة من القيم فى توزيع تكرارى ، وهى :

$$س_1 ،\ س_2 ،\ س_3 ،\ \ldots$$

وتكراراتها $ك_1 ،\ ك_2 ،\ ك_3 ،\ \ldots$ على الترتيب ، حيث $مجـ\ ك = ك$.

وليكن الوسط الحسابي لهذه القيم $\bar{س}$ والانحراف المعيارى $ع_{س}$.

الكمية $\frac{1}{ن}\ مجـ\ س\ .\ ك$ = العزم الأول للتوزيع التكرارى حول نقطة الأصل أو الصفر ،

$$= م'_1 \quad .\ \ .\ \ .\ \ .\ \ (١) .$$

والكمية $\frac{1}{ن}\ مجـ\ س^2\ .\ ك$ = العزم الثانى للتوزيع حول نقطة الأصل ،

$$= م'_2 \quad .\ \ .\ \ .\ \ .\ \ (٢) .$$

وعلى العموم نقول :

$\frac{1}{\text{ن}}\ \text{مجـ}\ \text{س}^{\text{ر}} . \text{ك}$ = العزم الرائى للتوزيع حول نقطة الأصل ،

$= \bar{\text{م}}_{\text{ر}}$ ، . . . . (٤) .

ويلاحظ أن $\bar{\text{م}}_{1} = \bar{\text{س}}$ = الوسط الحسابى ،

وأن $\bar{\text{م}}_{2} = \text{ع}^{2} + \bar{\text{س}}^{2}$ [أنظر معادلة (٣) بند ١٨١]

وإذا أخذنا الوسط الحسابى كمحور للعزوم ، وضربنا التكرارات فى انحرافات القيم عن الوسط الحسابى — أى بعدها عنه — نحصل على العزوم المختلفة حول الوسط الحسابى . وهذه نرمز لها بالحرف مـ .

$\therefore \frac{1}{\text{ن}}\ \text{مجـ}\ (\text{س} - \bar{\text{س}}) . \text{ك}$ = العزم الأول حول الوسط الحسابى .

$= \text{م}_{1}$ ، . . . . (٥) .

وكذلك $\frac{1}{\text{ن}}\ \text{مجـ}\ (\text{س} - \bar{\text{س}})^{2} . \text{ك}$

= العزم الثانى حول الوسط الحسابى .

$= \text{م}_{2}$ ، . . . . (٦) .

وعموماً : $\frac{1}{\text{ن}}\ \text{مجـ}\ (\text{س} - \bar{\text{س}})^{\text{ر}} . \text{ك}$

= العزم الرائى حول الوسط الحسابى .

$= \text{م}_{\text{ر}}$ ، . . . (٧) .

ويلاحظ أن م_{1} = صفراً ، و $\text{م}_{2} = \text{ع}^{2}$.

المقارنة بين تشتت المجموعات المختلفة

١٨٤ — مقاييس التشتت التى بحثناها فى هذا الباب يعبر كل منها عن التشتت بين مفردات مجموعة معينة . ولكنها لا تصلح على علاتها لمقارنة التشتت فى مجموعتين مختلفتين أو أكثر . لأنها معبر عنها بوحدات مطلقة هى نفس الوحدات المقيسة بها القيم الأصلية فى المجموعات ، كما رأينا مثلا فى بند ١٨٠ حيث وجدنا الانحراف المعيارى للأجور يساوى ٢,٥٥ قرشاً ، وفى بند ١٧٤ وجدنا الانحراف المتوسط للأعمار يساوى ١,٣٩ سنة . ولا يمكننا طبعاً أن نقارن بين ٢,٥٥ قرشاً

و ١,٣٩ سنة ، ونقول أن تشتت الأجور أكبر أو أقل من تشتت الأعمار .

ومن جهة أخرى ، نعلم أن الانحراف المعيارى لأى مجموعة يقيس تشتتها حول وسطها الحسابى . فإذا كان لدينا مجموعتان لهما وسطان حسابيان مختلفان : مثلا فرقتان من التلاميذ فيهما الوسطان الحسابيان للأعمار ١٥ سنة فى الأولى و١٨ سنة فى الثانية ، لا يمكننا مقارنة الانحرافين المعياريين لهما مباشرة ، بدون عمل تصحيح يعادل اختلاف وسطيهما الحسابيين .

١٨٥ — معامل الاختلاف [1] لأى مجموعة هو خارج قسمة الانحراف المعيارى على الوسط الحسابى ، مضروباً فى العدد ١٠٠ . فهو يعبر إذن عن الانحراف المعيارى (أو التشتت) فى صورة نسبة مئوية من الوسط الحسابى . وعلى ذلك يمكن استخدامه فى مقارنة التشتت للمجموعات المختلفة ، حتى ولو لم تتساو أوساطها الحسابية ، أو اختلف نوع الوحدات المستعملة فى قياس مفرداتها . معامل الاختلاف

ففى التوزيع التكرارى لأعمار ١٧٣٩ تلميذاً نرى أن معامل الاختلاف بين هذه الأعمار .

$$= \frac{١,٧٠٢}{١٧,٧٤٩} \times ١٠٠$$

$$= ٩,٦ \, .$$

وفى التوزيع التكرارى للأجور المذكورة فى بند ١٨٠ ، نجد معامل الاختلاف فى هذه المجموعة .

$$= \frac{٢,٥٥}{١٩,٢٥} \times ١٠٠$$ (انظر الوسط الحسابى فى بند ١٣٢)

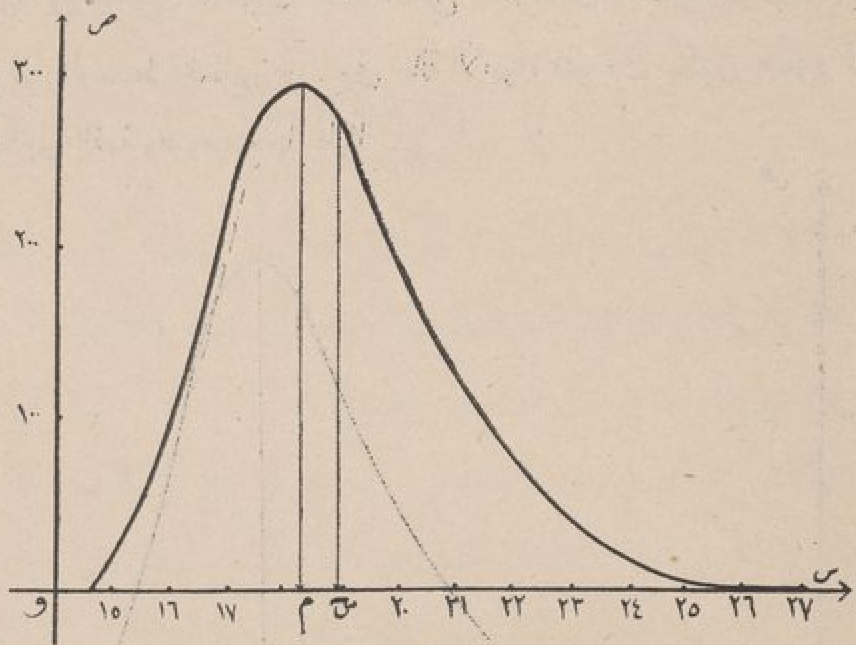
$$= ١٣,٢ \, .$$

∴ درجة الاختلاف بين الأجور (فى مجموعة ٧٤٣٢ عاملا) أكبر من درجة اختلاف الأعمار (فى مجموعة ١٧٣٩ تلميذاً) . ويلاحظ أن العددين ٩,٦ و ١٣,٢ لا تمييز لهما ، وهما فى الواقع نسبتان مئويتان .

(١) يسمى بالانجليزية Coefficient of Variation

قياس الالتواء

١٨٦ — ذكرنا فى الباب الخامس أن المنحنيات التكرارية بعضها متماثل وبعضها غير متماثل أو ملتو . والآن نبحث فى كيفية قياس درجة الالتواء فى المنحنيات التكرارية العادية ، أى التى يكون لها قمة واحدة ، كما فى الأشكال ٣٢ و ٣٣ و ٣٧ (صفحات ١٠٣ و ١٠٤ و ١١٠) . أما المنحنيات ذات الفرع الواحد (شكل ٤١ و ٤٢) أو ذات الفرعين (شكل ٤٣) فلا محل للبحث فيها من هذه الناحية .



(شكل ٥٩)
منحن تكرارى ملتو إلى اليسار التواء موجباً

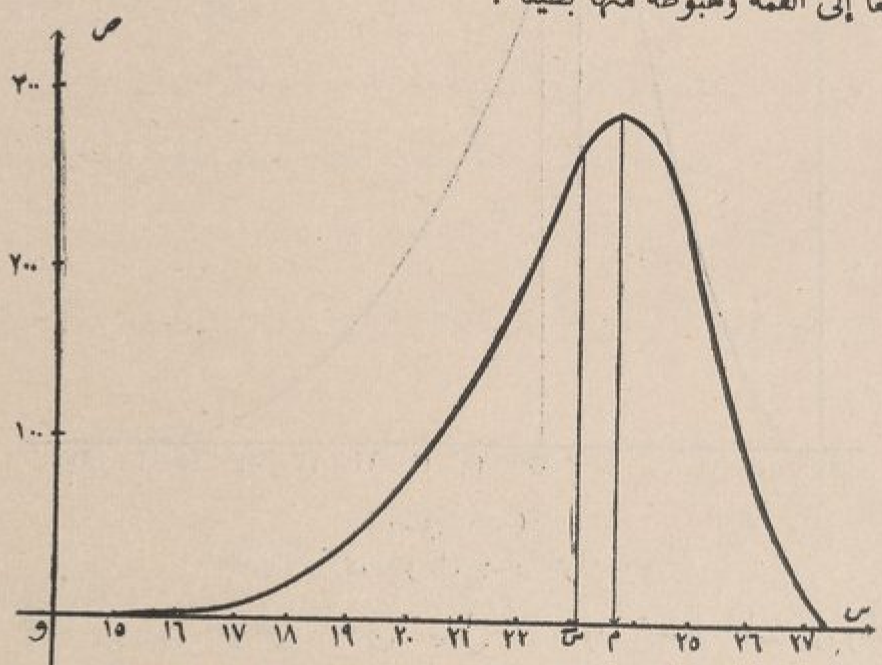
قلنا إن من خواص المنحنى المتماثل أن يكون الوسط الحسابى والمنوال متساويين . وهذه الخاصة لا توجد فى المنحنى غير المتماثل . يمكننا إذن اتخاذ هذه الخاصة أساساً لقياس الالتواء فالفرق بين الوسط الحسابى والمنوال يساوى صفراً إذا كان المنحنى متماثلا ؛ ويكون قريباً من الصفر إذا كان المنحنى قريباً من التماثل ؛ ويكون الالتواء شديداً إذا كان هذا الفرق كبيراً . وهذا الفرق نقيسه بالنسبة إلى الانحراف المعيارى للتوزيع التكرارى الذى نبحثه ، وذلك لأننا نعتبر الانحراف المعيارى هو الانحراف « النموذجى » أو « العادى » لمفردات المجموعة

عن الوسط الحسابى فنعتبره وحدة لقياس الانحرافات ؛ وننسب إليه انحراف المنوال عن الوسط الحسابى .

و بناء على ذلك يكون مقياس الالتواء[١] هو ى$_١$.

$$ى_١ = \frac{\text{الوسط الحسابى — المنوال}}{\text{الانحراف المعيارى}} = \frac{\bar{س} - م}{ع}$$

١٨٧ — إذا كان الوسط الحسابى أكبر من المنوال كان الالتواء موجباً كما نرى فى شكل ٥٩ حيث نجد المنحنى منحرفاً إلى اليسار ، ونرى المنوال م أصغر من الوسط الحسابى $\bar{س}$. وفى حالة الالتواء الموجب يكون صعود المنحنى سريعاً إلى القمة وهبوطه منها بطيئاً . الالتواء الموجب



(شكل ٦٠) منحن تكرار ملتوى إلى اليمين التواء سالباً

١٨٨ — إذا كان الوسط الحسابى للتوزيع التكرارى أصغر من المنوال ، كان الالتواء سالباً . والمنحنى التكرارى فى هذه الحالة يكون منحرفاً إلى اليمين حيث تكون قمته بعيدة عن المحور الرأسى ، و بعدها ينزل المنحنى بسرعة . ونرى فى شكل ٦٠ منحنياً ذا التواء سالب . ونرى فى الشكل أن المنوال أبعد من الوسط الالتواء السالب

(١) معناها بالانجليزية (Skewness) .

الحسابى عن المحور الرأسى ، أى أن المنوال م أكبر من الوسط الحسابى $\overline{س}$.

١٨٩ — يمكن أن نستعيض عن المقدار $\overline{س}$ — م المذكور فى مقياس الالتواء ، بالمقدار ٣ (الوسط الحسابى — الوسيط) إذا كان المنحنى قريباً من التماثل . وقد ذكرنا سابقاً (بند ١٦٧ صفحة ١٨١) أن :

الوسط الحسابى — الوسيط = $\frac{١}{٣}$ (الوسط — المنوال) تقريباً .

و بناء على ذلك يكون لدينا مقياس تقريبى للالتواء نستعمله فى حالة المنحنيات القريبة من التماثل ؛ وهو :

$$ى_١ = \frac{٣ \text{ (الوسط — الوسيط)}}{\text{الانحراف المعيارى}}$$

عيوب هذين المقياسين

١٩٠ — والعيب الكبير فى هذين المقياسين ى ، وى$_١$ أنهما يقتضيان معرفة الانحراف المعيارى للتوزيع التكرارى ؛ ونعلم أن حساب الانحراف المعيارى يحتاج إلى مجهود حسابى كبير نوعاً ، خصوصاً فى المجموعات كثيرة العدد . وإذا كان التوزيع التكرارى مفتوحاً من أحد الطرفين أو من كليهما ، لا يمكن معرفة الوسط الحسابى ولا الانحراف المعيارى ، ولا يمكن استخدامهما — وهذا عيب آخر . وكلاهما تقريبى : فالأول ى يحتاج إلى معرفة المنوال ؛ وهذا صعب التحديد كما نعلم ، ولا نعرفه إلا بالتقريب . والمقياس الثانى ى$_١$ مقرب أيضاً كما ذكرنا ، ومجال استعماله مقصور على بعض المنحنيات فقط .

مقياس بولى للالتواء

١٩١ — هناك مقياس آخر للالتواء . والفكرة الأساسية فى هذا المقياس هى أن الربيعين فى التوزيع التكرارى المتماثل متساويا البعد عن الوسط .

لنفرض أن الوسيط ، فى أى توزيع تكرارى ، يساوى ط ، وأن الربيع الأدنى = ع$_١$ ، والربيع الأعلى = ع$_٢$.

نضع $ف_٢ = ع_٢ - ط$ و $ف_١ = ط - ع_١$

∴ فى المنحنى المتماثل يكون $ف_٢ = ف_١$.

وعلى ذلك فالكمية $ف_٢ - ف_١$ تساوى صفراً فى المنحنيات المتماثلة ، وقريبة من صفر فى المنحنيات القريبة من التماثل . وتكون كبيرة إذا كان الالتواء شديداً .

وهنا ننسب الكمية $ف_٢ - ف_١$ إلى الكمية $ف_٢ + ف_١$ ، وهى تساوى الفرق بين الربيعين ($ف_٢ + ف_١ = ع_٢ - ع_١$). وقد سبق أن استخدمنا الفرق بين الربيعين كمقياس للتشتت (بند ١٧٠ صفحة ١٨٤) ، وعلى ذلك يكون مقياس[١] الالتواء بهذه الطريقة هو :

$$ى_٢ = \frac{ف_٢ - ف_١}{ف_٢ + ف_١}$$

وميزة هذا المقياس سهولة حساب الوسيط والربيعين بدقة أحسن من المنوال ، وأسهل من الانحراف المعيارى .

ويلاحظ أن الالتواء يكون سالباً إذا كان $ف_٢$ أصغر من $ف_١$ ، أى إذا كان الربيع الأعلى أقرب إلى الوسيط من الربيع الأدنى . وهذا يوافق تماماً المنحنى المنحرف بقمته إلى اليمين شكل ٦٠ . وعلى ذلك سيتفق المقياسان $ى_١$ ، $ى_٢$ فى إشارة الالتواء السالب .

وكذلك يتفق المقياسان فى إشارة الالتواء الموجب حينما يكون $ف_٢$ أكبر من $ف_١$ ، أى حينما يكون الربيع الأعلى أبعد عن الوسيط من الربيع الأدنى ، (شكل ٥٩) .

---

(١) هذا المقياس اقترحه A. L. Bowley أما المقياس الأول فينسب إلى Karl Pearson .
انظر كتاب F. C. Mills, *Statistical Methods* صفحة ١٦٦ طبعة ١٩٢٤

١٩٢ — نأخذ التوزيع التكرارى الآتى ، ونحسب مقاييس الالتواء بهاتين الطريقتين . وهو توزيع أعمار الناجحين فى شهادة الدراسة الثانوية (قسم ثان) قسم العلوم سنة ١٩٣٦ .

جدول ٢٨ — إيجاد الوسط الحسابى والمنوال والوسيط والربيعين والانحراف المعيارى لأعمار ١٣٤٤ تلميذاً .

الأعمار س	التكرار ك	انحرافات ع	ع . ك	ع^{2} . ك	حدود عليا	تكرار متجمع
١٥	١١	— ٤	— ٤٤	١٧٦	أقل من ١٥٫٥	١١
١٦	٨٠	— ٣	— ٢٤٠	٧٢٠	« ١٦٫٥	٩١
١٧	٢٣٦	— ٢	— ٤٧٢	٩٤٤	« ١٧٫٥	٣٢٧
١٨	٢٩٢	— ١	— ٢٩٢	٢٩٢	« ١٨٫٥	٦١٩
١٩	٢٧٤	٠	٠	٠	« ١٩٫٥	٨٩٣
٢٠	١٨٥	١	١٨٥	١٨٥	« ٢٠٫٥	١٠٧٨
٢١	١٤٠	٢	٢٨٠	٥٦٠	« ٢١٫٥	١٢١٨
٢٢	٦٤	٣	١٩٢	٥٧٦	« ٢٢٫٥	١٢٨٢
٢٣	٣٢	٤	١٢٨	٥١٢	« ٢٣٫٥	١٣١٤
٢٤	٢٣	٥	١١٥	٥٧٥	« ٢٤٫٥	١٣٣٧
٢٥	٣	٦	١٨	١٠٨	« ٢٥٫٥	١٣٤٠
٢٦	٣	٧	٢١	١٤٧	« ٢٦٫٥	١٣٤٣
٢٧	١	٨	٨	٦٤	« ٢٧٫٥	١٣٤٤
	١٣٤٤		— ١٠١	٤٨٥٩		

الوسط الحسابى = ١٨٫٩٢٥ ، والمنوال = ١٨٫٢٥٧ ، والانحراف المعيارى = ١٫٩٠٠

الوسيط = ١٨٫٦٩٥ ، والربيع الأدنى = ١٧٫٥٣٢ ، والربيع الأعلى = ٢٠٫١٢٦

$$\therefore \quad ى_{1} = \frac{١٨٫٩٢٥ - ١٨٫٢٥٧}{١٫٩٠٠} = ٠٫٣٥٢$$

$$ى'_{1} = \frac{٣\,(١٨٫٩٢٥ - ١٨٫٦٩٥)}{١٫٩٠٠} = ٠٫٣٦٣$$

$$ى_٢ = \frac{١٫٤٣١ - ١٫١٦٣}{٢٫٥٩٤} = ٠٫١٠٣$$

ويلاحظ أن المقياس الثانى يساوى الأول تقريباً . وأن الثلاثة متفقة فى الإشارة الموجبة ، دلالة على أن المنحنى ذو التواء موجب .

ولكن لا ينتظر أن يكون المقياس الثالث $ى_٣$ يساوى المقياس الأول أو الثانى ؛ لأن الفكرة الأساسية فيهما تختلف . وعلى ذلك لا نتوقع أن نحصل من المقياسين على نفس النتيجة . ويجب أن نتنبه إلى هذه النقطة عند مقارنة الالتواء فى مجموعات مختلفة ، فنقيس الالتواء فى كل المجموعات بنفس الطريقة ، حتى تكون المقارنة على أساس مشترك ، وإلا كانت المقارنة خطأ .

قياس الالتواء باستخدام العزم الثالث

١٩٣ — يوجد مقياس ثالث للالتواء وهو

$$ى_٣ = \frac{\sqrt[٣]{م_٣}}{ع}$$

حيث ع هى الانحراف المعيارى، و$م_٣$ هي العزم الثالث حول الوسط الحسابى .

أى أن

$$ى_٣ = \frac{\sqrt[٣]{مجـ ك (س - \bar{س})^٣}}{ع}$$

والفكرة الأساسية فى هذا المقياس هى أن العزم الثالث للتوزيع المتماثل حول وسطه الحسابى يساوى صفراً . وكلما كان هذا العزم قريباً من الصفر ، كان المنحنى قريباً من التماثل . وبالعكس إذا كان العزم الثالث كمية كبيرة (موجبة أو سالبة) كان التواء المنحنى شديداً .

ولكن هذا المقياس لا يستعمل كثيراً ، لصعوبة حسابه عملياً .

المــــراجع

Bowley, A.L., : *Elements of Statistics*, Chàpter VI.

Connor, L.R.,: *Statistics in Theory and Practice*, Chàpters XI, XII.

Jones, C., : *First Course in Statistics*, Chapters VI, VII.

Mills, F.C., : *Statistical Methods*, Chapter V.

الباب الثامن

الارتبـــــاط

تعريف الارتباط بين متغيرين

١٩٤ — الارتبـاط [١] بين ظاهرتين أو كميتين متغـيرتين معناه وجود علاقة بينهما ، بحيث إذا تغيرت إحداهما فى اتجاه معين فإن الثانية تميل إلى التغير فى اتجاه معين أيضـاً . ويصح أن يكون تغـير الظاهرتين فى اتجاه واحد ، أو فى اتجاهين متضادين . وفى الحالة الأولى نسمى الارتباط « طردياً » ، حيث إذا زادت واحدة تميل الثانيـــة إلى الزيادة أيضاً ؛ وإذا نقصت الأولى تميل الثانية إلى النقص أيضاً . وفى الحالة الثانية نسمى الارتباط « عكسياً » ، بمعنى أنه إذا تغيرت الـكمية الأولى بالزيادة ، تميل الثانية فى تغيرها إلى النقص ؛ والعكس بالعكس .

ليس من المحتم تلازم أو تساوى التغير فى الظاهرتين

١٩٥ — ولا يتحتم لوجود الارتباط أن كل زيادة تحصل فى أحد المتغيرين لا بد أن يصحبها زيادة فى المتغير الآخر (أو نقص فى حالة الارتباط العكسى) ؛ أو أن يكون التغير فيهما بنسبة واحدة . على أن هذا إذا تحقق يكون دلالة على شدة الارتباط والعلاقة بين المتغيرين ؛ ولـكن قد يتأتى أن يزيد المتغير الأول مثلاً ، ونظراً لظروف طارئة ، فى حالة معينة ، ينقص المتغير الثانى ، على خلاف ما تقضى به العلاقة الطردية المفروضة بينهما . ولـكن المهم فى الموضوع أنه فى أغلب الحالات نجد الزيادة فى المتغير الأول مصحوبة بزيادة فى الثانى ، فى حالة الارتباط الطردى (أو بنقص فى حالة الارتباط العكسى) ؛ ونجد النقص فى أحدهما مقروناً

(١) معناه بالإنجليزية Correlation وقد سماه بعض الناس « علاقة مشتركة » ؛ ولـكننا نرى كلمة ارتباط أحسن وأبسط .

بنقص فى الآخر (أو زيادة فى الحالة العكسية) . ولا تكون النسبة بين التغيرين ثابتة فى كل الأحوال التى تقع تحت ملاحظتنا ، ولكنها تتراوح حول مقدار معين .

وهذا هو السبب فى قولنا إن وجود الارتباط معناه أن أحد المتغيرين « يميل » إلى مصاحبة الثانى فى تغيره على وجه العموم .

ويلاحظ أن هذا التعريف أيضاً تتمثل فيه وجهة النظر الإحصائية . إذ أننا لا نرجع إلى الحالات الفردية عند استنباط القوانين أو القواعد التى تسير عليها الظواهر ، ولكننا نهتم فقط بالاتجاه العام الذى تأخذه المجموعات الكبيرة . وذلك لنتفادى أخطاء المصادفات وتقلبات الظروف الشاذة .

معنى وجود الارتباط

١٩٦ — يتضح لنا مما تقدم أن الارتباط بين متغيرين معناه أن التغير فى أحدهما يكون فى العادة مصحوباً بتغير فى الآخر ، وأن هناك علاقة معينة بين اتجاهى التغير فيهما — طردية أو عكسية .

ولكن وجود ارتباط بين ظاهرتين متغيرتين ليس دليلا على أن إحداهما نتيجة للأخرى، أو أن التغير فى واحدة تابع للتغير فى الأخرى ولا ينشأ إلا بسببه . بل هو يشير فقط إلى احتمال وجود هذه العلاقة . لأن هذه العلاقة ما هى إلا نوع خاص من أنواع العلاقات التى يدل الارتباط على وجودها . وهى كما يأتى :

الحالة الأولى — أن يكون أحد المتغيرين نتيجة مباشرة للثانى .

علاقة سببية مباشرة

ومثال ذلك الارتباط بين سعر أى سلعة فى السوق وكمية المطلوب منها فى السوق . إذ أن ارتفاع الأسعار ينتج عنه مباشرة هبوط فى كمية المستهلك من هذه السلعة ؛ وانخفاضها يسبب ازدياداً فى الطلب .

علاقة سببية غير مباشرة

الحالة الثانية — يكون أحد المتغيرين سبباً غير مباشر للثانى ، يؤثر فيه بوساطة عامل ثالث أو أكثر . فارتفاع الرسوم الجمركية على المنسوجات القطنية مثلا ،

يسبب ارتفاعاً فى أسعارها الداخلية . وهـذا الارتفاع يمكن أصحاب المصانع المحلية من رفع أجور عمالهم . فالارتباط بين مقدار الرسوم الجمركية فى هذه الحالة ومستوى الأجور ناتج من علاقة سببية غير مباشرة بينهما .

عامل واحد يـؤثر فى المتغيرين

الحالة الثالثة — أن يكون كل من المتغيرين المرتبطين نتيجة لعـامل ثالث ، مشترك بينهما ، يؤثر فيهما فى وقت واحد ، فيكون كل تغـير فى أحدهما مصحوباً بتغير فى الآخر . مثال ذلك الارتباط بين أسعار سلعتين تستهلكهما طبقـة معينة من السكان . فإن أسـعارهما تكون مرهونة بالحالة الاقتصادية لهؤلاء السكان : فيرتفع سعر كل منهما إذا زادت القوة الشرائية لهم ، ويهبط السعران معاً إذا نقصت قوتهم الشرائية بسبب انتشار البطالة بينهم أو لأى سبب آخر . وكذلك الارتباط بين أسعار سلعتين تردان من بلد بعيدة جداً ، بحيث تكون نفقات إنتاجهما ضئيلة بالنسبة إلى نفقات النقل — فنجد أسـعار هاتين السلعتين ترتفع وتنخفض معاً ، تبعاً لتغيرات نفقات النقل . وكذلك إذا كانتـا تنتجان من مادة خام واحدة ، رئيسية فى كل منهما ، بحيث تكوّن الجزء الأكبر من نفقات الإنتـاج فيهما — فأسعارهما ترتفع أو تنخفض معاً تبعاً لأسعار هذه المـادة الرئيسية .

بعض العوامل مشتركة بين المتغيرين

الحالة الرابعة — أن يكون ضمن العوامل التى تؤثر فى أحد المتغيرين والعوامل التى تؤثر فى الآخر ، عامل مشترك أو أكثر . مثلا لو اختبرنا عدداً من التلاميـذ فى مادتين مثل الجغرافيا والتاريخ ، نجد ارتباطاً شديداً بين درجات هاتين المادتين . والسبب فى ذلك أن مقدرة الشخص ونبوغه أو ضعفه فى مادة التـاريخ يتوقف على استعداده العام أو ذكائه ، وعلى مقدرة أخرى نوعية خاصة بهذا العلم وطرائقه . ونبوغ أى شخص أو ضعفه فى علم الجغرافيـا يتوقف أيضاً على استعداده العـام أو ذكائه ، وبجانب ذلك على مقدرة خاصة بعلم الجغرافيا تخالف تلك المقدرة اللازمة

للنبوغ فى علم التاريخ . وعلى ذلك نجد هنا عاملاً مشتركاً — وهو الذكاء الشخصى أو الاستعداد العام — يؤثر فى الظاهرتين ، علاوة على عوامل أخرى خاصة بكل ظاهرة على حدة وليست مشتركة .

ومثال ذلك أيضاً سلعتان فى السوق تدخل فى إنتاجهما مادة خام أو أكثر بصفة رئيسية ، علاوة على مواد أخرى خاصة بكل سلعة ، ولا تدخل فى الأخرى . فالارتباط الذى نجده بين أسعار هاتين السلعتين ناتج من وجود عوامل مشتركة بينهما ضمن العوامل التى تؤثر فى كل واحدة .

شدة الارتباط لا تحدد نوع العلاقة

ومهما كان الارتباط بين المتغيرين شديداً ، فهو لا يكفى بمفرده لمعرفة نوع العلاقة بينهما . ولا بد لتحديد نوعها من الاستعانة بمعلوماتنا الخاصة ، وإلمامنا بظروف هذين المتغيرين . وعلى كل حال فنوع هذه العلاقة محصور فى الأنواع الأربعة التى ذكرناها . وفى كل نوع منها يصح أن يكون الارتباط شديداً أو ضعيفاً .

لقياس الارتباط نميز بين أقسامه وهى : الارتباط والاقتران والتوافق

١٩٧ — أول خطوة فى دراسة الارتباط هى أن نبحث فى كيفية قياسه ، والتعبير عنه فى صورة رقمية تساعدنا فى عمل المقارنات بين الحالات المختلفة التى يظهر فيها الارتباط . ولو تأملنا فى الحالات المختلفة التى يمكن أن تعرض لنا عند دراسة الارتباط ، وجدناها تشمل عدة أنواع . وهذه يمكن تقسيمها إلى ثلاثة أنواع [١] كما يأتى :

---

(١) اقتران = Association . توافق = Contingency . ارتباط بسيط = Simple Correlation

أولا — العلاقة بين ظواهر يمكن أن تقاس ويعبر عنها فى صورة رقمية . وهذه العلاقة نسميها « ارتباطاً » . ومثال ذلك العلاقة بين طول الشخص ووزنه ، و بين سعر السلعة وكمية المطلوب منها ، و بين كمية المحصول وكمية السماد المستعمل فى حقل معين ، وهكذا .

ثانياً — العلاقة بين ظواهر لا يمكن قياسها رقمياً . وهذه العلاقة نسميها « الاقتران » . ومثال ذلك العلاقة بين جنسية الشخص وديانته ، و بين لون الزهرة ورائحتها ، و بين نوع الشخص (ذكر أو أنثى) ونوع العمل الذى يقوم به (صناعى أو تجارى) ، وهكذا .

ثالثاً — العلاقة بين ظواهر بعضها يقاس رقمياً و بعضها لا يقاس . وهذه نسميها « توافقاً » . ومثال ذلك العلاقة بين نوع القطن (سكلاريدس وأشمونى الخ) وطول تيلته بالسنتيمتر ، و بين نوع الحرفة التى يزاولها العامل وأجره بالقرش ، وهكذا .

الارتباط البسيط والتعدد والجزئى

١٩٨ — لنأخذ أولا العلاقة بين الظواهر المقيسة ، أى الارتباط . ونجد هنا أن البحث يتفرع إلى فرعين[1] :

١ — الارتباط بين ظاهرتين أو متغيرين فقط . و يسمى « الارتباط البسيط » .

٢ — الارتباط بين ظاهرة وظاهرتين مجتمعتين أو منفردتين ، أو أكثر من ظاهرتين . ويسمى « الارتباط المتعدد » أو « الارتباط الجزئى » ، على الترتيب . وسنبدأ ببحث الارتباط البسيط وطريقة قياسه .

---

(١) ارتباط متعدد = Multiple Correlation . ارتباط جزئى Partial Correlation .

لقياس الارتباط نحصل على قيم متناظرة للمتغيرين

١٩٩ — عند دراسة الارتباط البسيط بين كميتين متغيرتين ، نبدأ بمشاهدة هاتين الكميتين فى عدد من الحالات ، وفى كل حالة من هذه نقيس كلا منهما ، وندون القيم المتناظرة لهما . وإذا كان عدد الحالات التى وقعت تحت ملاحظتنا ن مثلا ، حصلنا على قيم عددها ن للمتغير الأول ومثلها للمتغير الثانى ، تناظرها واحدة لواحدة . أى أننا نحصل على ن من أزواج القيم المتناظرة لهذين المتغيرين .

وللسهولة فى الكلام والتفكير ، نفرض أن المتغير الأول س ، والمتغير الثانى المرتبط به ص . ونفرض أن القيم التى حصلنا عليها للمتغير الأول فى هذه التجربة هى :

$$\text{س}_1 \text{ ، } \text{س}_2 \text{ ، } \text{س}_3 \text{ ، } . \quad . \quad . \text{ ، } \text{س}_{\text{ن}} \text{ ؛}$$

وأن القيم المناظرة لها للمتغير الثانى هى على الترتيب :

$$\text{ص}_1 \text{ ، } \text{ص}_2 \text{ ، } \text{ص}_3 \text{ ، } . \quad . \quad . \text{ ، } \text{ص}_{\text{ن}} \text{ .}$$

ولنفرض أيضاً أن الوسطين الحسابيين لهاتين المجموعتين من القيم هما على الترتيب $\bar{\text{س}}$ و $\bar{\text{ص}}$ ؛ وأن الانحرافين المعياريين لهما هما ع ، $\text{ع}'$.
وحسب تعريف الوسط الحسابى والانحراف المعيارى :

$$\text{يكون ن } \bar{\text{س}} = \text{مجـ س} \qquad \text{و} \quad \text{ن } \bar{\text{ص}} = \text{مجـ ص ؛}$$

$$\text{و} \quad \text{ن ع}^2 = \text{مجـ}(\text{س} - \bar{\text{س}})^2 \quad \text{و} \quad \text{ن ع}'^2 = \text{مجـ}(\text{ص} - \bar{\text{ص}})^2$$

مقياس الارتباط مشتق من قيم س وقيم ص على السواء

٢٠٠ — هذه القيم السينية والصادية هى كل ما لدينا من المعلومات فى هذا الموضوع ، وسنعتمد عليها كلياً فى قياس الارتباط بين هذين المتغيرين س و ص ؛ فلا بد إذن أن يكون مقياس الارتباط الذى نستخدمه مشتقاً

من هذه القيم . ولا بد أيضاً أن تدخل قيم س وقيم ص على قدم المساواة في الصيغة العامة لهذا المقياس . وبما أننا نستخدم جميع ما لدينا من المعلومات عن هذين المتغيرين لقياس الارتباط بينهما ، فسنجد أن جميع هذه القيم السينية والصادية تدخل فى حسابنا بدون تفضيل بين أى قيمة وأخرى .

مقياس الارتباط الذى نستعمله فى هذه الحالة نسميه « معامل الارتباط »(١) .

والآن نتكلم فى كيفية اشتقاقه من هذه القيم الموجودة بالتجربة والقياس ، طبقاً للتعريف الذى أوردناه (أنظر بند ١٩٤) ، والمعنى الذى نقصده من فكرة الارتباط .

الارتباط يقاس بواسطة التغيرات الحاصلة فى الظاهرتين

٢٠١ — قلنا إن الارتباط بين متغيرين معناه أن التغير فى أحدهما يكون — على العموم — مصحوباً بتغير فى الآخر ، بمعنى أن الزيادة فى أحدهما تكون — على العموم — مصحوبة مثلاً(٢) بزيادة فى الثانى ، والنقص فى الأول يكون مصحوباً بنقص فى الثانى أيضاً .

وإذا تكلمنا عن التغير فى كمية مثل س (أو ص) ، فمقدار هذا التغير يساوى الفرق بين القيمة التى تأخذها س ومقدار معين يعتبر أساساً . وأحسن أساس نختاره لقياس التغير فى قيم س هو بلا شك الوسط الحسابى لهذه القيم السينية ، أى $\bar{س}$. وكذلك فى حالة ص نختار $\bar{ص}$.

وعلى ذلك تكون التغيرات فى قيم س وقيم ص هى :

$$س_١ - \bar{س} ، س_٢ - \bar{س} ، س_٣ - \bar{س} ، \dots \dots ، س_ن - \bar{س} ؛$$

$$و \quad ص_١ - \bar{ص} ، ص_٢ - \bar{ص} ، ص_٣ - \bar{ص} ، \dots \dots ، ص_ن - \bar{ص} .$$

(١) فى الإنجليزية Coefficient of Correlation ويرمز له بالحرف ر .

(٢) أو تكون مصحوبة بنقص فى حالة الارتباط العكسى .

وهذه هى ، بعبارة أخرى ، انحرافات القيم عن وسطها الحسابى ؛ وهذه الانحرافات إذن هى التى سنعتمد عليها فى قياس الارتباط .

تقسيم الانحرافات على الانحراف المعيارى

٢٠٢ — ولكن هذه الانحرافات لا يمكن مقارنتها ببعضها كما هى على هذه الصورة . لأنه إذا فرضنا أن س تدل على عمر الشخص بالسنين مثلا و ص تدل على وزنه بالكيلوجرام ، فإن الانحرافات السينية تكون مقيسة بالسنين ، فى حين أن الانحرافات الصادية تكون مقيسة بالكيلوجرامات . وعلاوة على ذلك فإن التشتت مختلف فى مجموعتى قيم س و ص ، مما يفسد المقارنة بين هذه الانحرافات على علاتها .

لذلك نقسم الانحرافات السينية على الانحراف المعيارى لقيم س ، والانحرافات الصادية على الانحراف المعيارى لقيم ص ، حيث إن الانحراف المعيارى هو ، كما سبق أن ذكرنا فى مناسبة أخرى (صفحة ٢٠٥) ، الوحدة التى تقاس بها الانحرافات و بواسطتها تمكن مقارنتها . وهكذا نحصل على انحرافات سينية وانحرافات صادية خالية من كل تمييز ومعبر عنها بوحدات يمكن مقارنتها .

وهذه الانحرافات هى :

$$\frac{\text{س}_1 - \bar{\text{س}}}{\text{ع}} \text{ ، } \frac{\text{س}_2 - \bar{\text{س}}}{\text{ع}} \text{ ، } \frac{\text{س}_3 - \bar{\text{س}}}{\text{ع}} \text{ ، } \ldots\ldots \text{ ، } \frac{\text{س}_\text{ن} - \bar{\text{س}}}{\text{ع}} \text{ ؛}$$

$$\frac{\text{ص}_1 - \bar{\text{ص}}}{\text{ع}'} \text{ ، } \frac{\text{ص}_2 - \bar{\text{ص}}}{\text{ع}'} \text{ ، } \frac{\text{ص}_3 - \bar{\text{ص}}}{\text{ع}'} \text{ ، } \ldots\ldots \text{ ، } \frac{\text{ص}_\text{ن} - \bar{\text{ص}}}{\text{ع}'} \text{ .}$$

متوسط حاصل ضرب الانحراف السينى والانحراف الصادى المناظر له

٢٠٣ — لنفرض أن هناك ارتباطاً شديداً بين س و ص . إذن لو تغيرت س وزادت زيادة كبيرة (أو نقصت كثيراً) وانحرفت بذلك عن وسطها الحسابى انحرافاً كبيراً ، كان ذلك مصحوباً بتغير كبير أيضاً فى ص وانحراف كبير عن وسطها الحسابى أيضاً ؛ ونتج عن ذلك أن يكون حاصل ضرب هذين الانحرافين

كبيراً (موجباً إذا كان الارتباط طردياً ، أو سالباً إذا كان عكسياً) . وإذا تغيرت س بمقدار صغير فقط ، كان التغير فى ص صغيراً أيضاً بحكم الارتباط الشديد بينهما ، وكان حاصل ضرب الانحرافين صغيراً ؛ ولكن هذا الحاصل الصغير يكون فقط فى الحالات التى تكون فيها قيم س و ص قريبة من وسطيهما الحسابيين . وعلى العموم يكون متوسط حواصل ضرب الانحرافات كبيراً فى حالة الارتباط الشديد .

أما إذا كان الارتباط ضعيفاً ، وكانت كل من س و ص تتغير مستقلة عن الأخرى ، فيصح أن تكون س كبيرة جداً ، وانحرافها عن وسطها كبيراً جداً ، دون أن يظهر لذلك أى أثر فى ص ، لضعف الصلة بينهما . وعلى ذلك يكون حاصل ضرب انحرافى س و ص صغيراً فى المتوسط .

وعلى ذلك نأخذ متوسط حاصل ضرب انحرافى س و ص عن وسطيهما الحسابيين ، كمقياس للارتباط بينهما ، كما نقول فى علم الطبيعة إن قوة الجاذبية بين جسمين تتناسب مع حاصل ضرب كتلتيهما ، و بين قطبى مغناطيسين تساوى حاصل ضرب شدتيهما . ويكون معامل الارتباط [1] إذن يساوى

$$\text{ر} = \frac{1}{\text{ن}}\left[\frac{\text{س}_1 - \bar{\text{س}}}{\text{ع}} \times \frac{\text{ص}_1 - \bar{\text{ص}}}{\bar{\text{ع}}} + \frac{\text{س}_2 - \bar{\text{س}}}{\text{ع}} \times \frac{\text{ص}_2 - \bar{\text{ص}}}{\bar{\text{ع}}} + \dots + \frac{\text{س}_\text{ن} - \bar{\text{س}}}{\text{ع}} \times \frac{\text{ص}_\text{ن} - \bar{\text{ص}}}{\bar{\text{ع}}}\right]$$

$$= \frac{\text{مج}(\text{س} - \bar{\text{س}})\,.\,(\text{ص} - \bar{\text{ص}})}{\text{ن ع}\,.\,\bar{\text{ع}}} \quad ، \qquad (١) .$$

وهذا هو أهم المقاييس المستعملة فى دراسة الارتباط . وهذه الصيغة هى الصيغة الأساسية لمعامل الارتباط ؛ وسنرى جملة صيغ أخرى مشتقة منها بعمليات جبرية بسيطة .

---

(١) هذا المعامل وضعه Karl Pearson وسماه « معامل حاصل ضرب العزوم للارتباط »
Product Moment Coefficient of Correlation

حدود قيم المعامل ر وإشارته الجبرية

٢٠٤ — ويلاحظ أن جميع قيم س وقيم ص تدخل فى هذا المعامل على قدم المساواة كما قلنا من قبل (بند ٢٠٠) .

ومن المهم جداً أن نلاحظ أن هذا المعامل لا يزيد أبداً عن الواحد الصحيح ؛ وهو دائماً محصور[١] فى الفترة ما بين — ١ و + ١ . ويكون موجباً إذا كان الارتباط طردياً ، وسالباً إذا كان عكسياً . وإذا كان ر = . كان الارتباط منعدماً ، أى أن الظاهرتين مستقلتان عن بعضهما تماماً .

لأنه فى حالة الارتباط الطردى تقترن قيم س الكبيرة ، على العموم ،

---

(١) يمكن إثبات ذلك بأن نبرهن على أن $ر^2$ دائماً أقل من ١ ، وذلك كما يأتى :

نكتب $ل_1 = س_1 - \bar{س}$ ، و $ل_2 = س_2 - \bar{س}$ ، .. للاختصار

و $م_1 = ص_1 - \bar{ص}$ ، و $م_2 = ص_2 - \bar{ص}$ ، ... «

∴ $ن ع^2 = ل_1^2 + ل_2^2 + .. + ل_ن^2$ ؛ و $ن ع'^2 = م_1^2 + م_2^2 + .. + م_ن^2$:

$ر^2 = [ل_1 م_1 + ل_2 م_2 + .. + ل_ن م_ن]^2 \div (ل_1^2 + ل_2^2 + .. + ل_ن^2)(م_1^2 + م_2^2 + .. + م_ن^2)$

البسط $= ل_1^2 م_1^2 + ل_2^2 م_2^2 + ل_3^2 م_3^2 + .. + 2 ل_1 م_1 (ل_2 م_2 + ل_3 م_3 + ...) + 2 ل_2 م_2 (ل_3 م_3 + ..) + ...$

المقام $= ل_1^2 م_1^2 + ل_2^2 م_2^2 + ل_3^2 م_3^2 + ... + ل_1^2 (م_2^2 + م_3^2 + ..) + ل_2^2 (م_1^2 + م_3^2 + ..) + ...$

∴ المقام — البسط $= ل_1^2 م_2^2 + ل_2^2 م_1^2 + ل_1^2 م_3^2 + ل_3^2 م_1^2 + ...$
$- 2 ل_1 م_1 ل_2 م_2 - 2 ل_1 م_1 ل_3 م_3 - ...$
$= (ل_1 م_2 - ل_2 م_1)^2 + (ل_1 م_3 - ل_3 م_1)^2 + (..)^2 + ...$
= مجموع مربعات كاملة = كمية موجبة أو صفراً

∴ البسط $\le$ من المقام . وعلى ذلك $ر^2 \le 1$ ؛

ويكون ر محصوراً بين —١ و +١ ، ويصح أن $ر = \pm 1$.

بقيم ص الكبيرة أيضاً ، وكذلك تقترن قيم س الصغيرة بقيم ص الصغيرة . أى أن انحرافات س و ص عن وسطيهما الحسابيين تكون متحدة الإشارة . فتكون حواصل ضرب هذه الانحرافات موجبة ومجموع هذه الحواصل يكون موجباً ، وحينئذ تكون ر موجبة .

أما فى حالة الارتباط العكسى فتقترن قيم س الكبيرة بقيم ص الصغيرة على العموم ، والعكس بالعكس . وعلى ذلك تكون انحرافات س عن وسطها الحسابى مخالفة فى الإشارة لانحرافات ص . ويكون حاصل ضرب هذه الانحرافات سالباً ، وتكون ر سالبة بناء على ذلك .

وإذا كان بعض قيم س الكبيرة مقترنة بقيم كبيرة للمتغير ص و بعضها مقترناً بقيم صغيرة ، فإن بعض الحواصل يكون موجباً و بعضها يكون سالباً ، مما يجعل مجموع هذه الحواصل صغيراً أو صفراً ، فتكون ر صغيرة أو صفراً ، دلالة على ضعف الارتباط أو انعدامه .

صيغة أخرى لمعامل الارتباط

٢٠٥ — يمكننا ، إذا أردنا ، أن نستعيض (١) عن العبارة

---

(١) من الواضح أن $\text{مج}(\text{س}-\overline{\text{س}})(\text{ص}-\overline{\text{ص}}) = \text{مج}\,\text{س}\,\text{ص} - \text{ن}\,\overline{\text{س}}\,\overline{\text{ص}}$

لأن $(\text{س}_1-\overline{\text{س}})(\text{ص}_1-\overline{\text{ص}}) = \text{س}_1\text{ص}_1 + \overline{\text{س}}\,\overline{\text{ص}} - \overline{\text{س}}\,\text{ص}_1 - \overline{\text{ص}}\,\text{س}_1$ ،

و $(\text{س}_2-\overline{\text{س}})(\text{ص}_2-\overline{\text{ص}}) = \text{س}_2\text{ص}_2 + \overline{\text{س}}\,\overline{\text{ص}} - \overline{\text{س}}\,\text{ص}_2 - \overline{\text{ص}}\,\text{س}_2$ ،

. . . .

و $(\text{س}_\text{ن}-\overline{\text{س}})(\text{ص}_\text{ن}-\overline{\text{ص}}) = \text{س}_\text{ن}\text{ص}_\text{ن} + \overline{\text{س}}\,\overline{\text{ص}} - \overline{\text{س}}\,\text{ص}_\text{ن} - \overline{\text{ص}}\,\text{س}_\text{ن}$ ؛

وبالجمع

$$\therefore\ \text{مج}(\text{س}-\overline{\text{س}})(\text{ص}-\overline{\text{ص}}) = \text{مج}\,\text{س}\,\text{ص} + \text{ن}\,\overline{\text{س}}\,\overline{\text{ص}} - \overline{\text{س}}(\text{مج}\,\text{ص}) - \overline{\text{ص}}(\text{مج}\,\text{س})$$

$$= \text{مج}\,\text{س}\,\text{ص} + \text{ن}\,\overline{\text{س}}\,\overline{\text{ص}} - \overline{\text{س}}\cdot(\text{ن}\,\overline{\text{ص}}) - \overline{\text{ص}}\cdot(\text{ن}\,\overline{\text{س}})$$

$$= \text{مج}\,\text{س}\,\text{ص} - \text{ن}\,\overline{\text{س}}\,\overline{\text{ص}} \ .$$

مجـ $(س - \bar{س})\,(ص - \bar{ص})$ بالعبارة مجـ $س\,ص - ن\,\bar{س}\,\bar{ص}$.

وبذلك يصبح معامل الارتباط :

$$ر = \frac{\text{مجـ}\ س\,ص - ن\,\bar{س}\,\bar{ص}}{ن\ ع \cdot ع'} \quad \text{،} \quad \ldots \ \ldots \qquad (٢) \ .$$

وهـذه الصورة تكون أحياناً أسهل فى الحساب من الصورة المتقدمة ، لأن الوسـط الحسابى $\bar{س}$ أو $\bar{ص}$ عادة يكون عدداً كسرياً ذا أرقام كثيره ؛ وينتج من ذلك أيضاً أن الانحرافات $س - \bar{س}$ و $ص - \bar{ص}$ تكون أعداداً مكونة من أرقام كثيرة ، فتجعل العمليات الحسابية معقدة .

صيغة ثالثة للمعامل ر

٢٠٦ — ويمكننا أيضاً، إذا أردنا ، أن نحسب الانحرافات السينية عن وسط فرضى مناسب بدل الوسط الحسابى ، وكذلك الانحرافات الصادية نحسبها عن وسط فرضى آخر بدل وسطها الحسابى . وهذه الطريقة نلجأ إليها فى بعض الأحيان إذا وجدنا أن عمليات الحساب فى إيجاد مجـ $(س - \bar{س})$ $\times (ص - \bar{ص})$ ، أو مجـ $س\,ص - ن\,\bar{س}\,\bar{ص}$ ، صعبة أو طويلة . وحينئذ[1] يكون معامل الارتباط :

---

(١) هنا أيضاً

مجـ $(س - \bar{س})\,(ص - \bar{ص}) = \text{مجـ}\,(س - و)\,(ص - و') - ن\,ح\,ح'$

لأن $(س_١ - \bar{س})\,(ص_١ - \bar{ص}) = (س_١ - و + و - \bar{س})$
$\times\,(ص_١ - و' + و' - \bar{ص})$
$= (س_١ - و - ح)\,(ص_١ - و' - ح')$
$= (س_١ - و)\,(ص_١ - و') - ح\,(ص_١ - و')$
$- ح'\,(س_١ - و) + ح\,ح'$

وكذلك $(س_٢ - \bar{س})\,(ص_٢ - \bar{ص}) = (س_٢ - و)\,(ص_٢ - و') - ح\,(ص_٢ - و')$
$- ح'\,(س_٢ - و) + ح\,ح'$

و .. ..

و $(س_ن - \bar{س})\,(ص_ن - \bar{ص}) = (س_ن - و)\,(ص_ن - و') - ح\,(ص_ن - و')$
$- ح'\,(س_ن - و) + ح\,ح' \qquad (*)$

ر = [مجـ (س − و) (ص − وَ) − ن ح حَ] / ن ، ...... (٣) ،

بفرض أن و = الوسط الفرضى السينى ، و ح = س̄ − و ،
وأن وَ = « « الصادى ، و حَ = ص̄ − وَ .

معادلة الفروق للمعامل ر

٢٠٧ — ولو عوضنا فى المعادلة (٢) عن العبارة مجـ س ص بالعبارة[1] :

½ [مجـ س^{٢} + مجـ ص^{٢} − مجـ (س − ص)^{٢}]

نحصل على صيغة ثالثة لمعامل الارتباط ، نسميها[2] **معادلة الفروق** لحساب معامل الارتباط ، تمييزاً لها عن معادلات حاصل الضرب السابقة ، حيث إنها تحتوى على الفروق المربعة مجـ (س − ص)^{٢} بدل حواصل الضرب مجـ (س ص) أو مجـ (س − س̄) (ص − ص̄) ، أو مجـ (س − و) (ص − وَ).

وهذه المعادلة هى :

(*) ∴ مجـ (س − س̄) (ص − ص̄) = مجـ (س − و) (ص − وَ) − ح (مجـ ص − ن وَ)
− حَ (مجـ س − ن و) + ن ح حَ
= مجـ (س − و) (ص − وَ) − ح (ن ص̄ − ن وَ)
− حَ (ن س̄ − ن و) + ن ح حَ
= مجـ (س − و) (ص − وَ) − ح . ن حَ
− حَ . ن ح + ن ح حَ
= مجـ (س − و) (ص − وَ) − ن ح حَ

(١) واضح أن هاتين العبارتين متساويتان . لأن

(س − ص)^{٢} = س^{٢} + ص^{٢} − ٢ س ص ،

∴ مجـ (س − ص)^{٢} = مجـ س^{٢} + مجـ ص^{٢} − ٢ مجـ س ص ،

أى أن ٢ مجـ س ص = مجـ س^{٢} + مجـ ص^{٢} − مجـ (س − ص)^{٢}.

(٢) بالإنجليزية Difference Equation for the Correlation Coefficient

$$\text{ر} = \frac{\text{مجـ س}^{٢} + \text{مجـ ص}^{٢} - \text{مجـ}(\text{س} - \text{ص})^{٢} - ٢\,\text{ن}\,\bar{\text{س}}\,\bar{\text{ص}}}{٢\,\text{ن}\,\text{ع}\,\text{عَ}} \quad \ldots (٤) .$$

وبما أن $\text{مجـ س}^{٢} = \text{ن ع}^{٢} + \text{ن}\,\bar{\text{س}}^{٢}$ ،

و $\text{مجـ ص}^{٢} = \text{ن عَ}^{٢} + \text{ن}\,\bar{\text{ص}}^{٢}$ ،

فلو عوضنا بهذه القيم عن $\text{مجـ س}^{٢}$ و $\text{مجـ ص}^{٢}$ فى المعادلة (٣) نحصل على صيغة خامسة لمعامل الارتباط وهى :

$$\text{ر} = \frac{١}{٢\,\text{ع}\,\text{عَ}}\left[\text{ع}^{٢} + \text{عَ}^{٢} + (\bar{\text{س}} - \bar{\text{ص}})^{٢} - \frac{١}{\text{ن}}\,\text{مجـ}(\text{س} - \text{ص})^{٢}\right] \quad (٥)$$

٢٠٨ — وفى بعض الأحيان يفضل استخدام إحدى هاتين المعادلتين فى حساب معامل الارتباط ، نظراً لسهولة حساب الفروق (س — ص) وتربيعها ، عن ضرب القيم نفسها أو ضرب انحرافاتها . فنرى مثلا فى الحالة الخاصة التى تكون فيها قيم ص هى نفس قيم س ولكن بترتيب مخالف ، أن هذه المعادلة الأخيرة تؤول إلى صورة بسيطة جداً . لأن الوسطين الحسابيين $\bar{\text{س}}$ و $\bar{\text{ص}}$ يكونان حينئذ متساويين ، وكذلك الانحرافان المعياريان . وتؤول المعادلة حينئذ إلى الصورة الآتية:

حينما تكون قيم س هى قيم ص بترتيب مختلف

$$\text{رَ} = ١ - \frac{\text{مجـ}(\text{س} - \text{ص})^{٢}}{٢\,\text{ن}\,.\,\text{ع}^{٢}} \quad (٦)$$

حيث وضعنا رَ بدل ر فى هذه الحالة الخاصة ، ووضعنا ع لتقوم مقام الانحرافين المعياريين المتساويين .

٢٠٩ — وإذا كانت قيم س تدل على تراتيب مجموعة من الأشخاص (عددهم ن) فى مسابقة معينة ، فى امتحان مادة ما مثلا ، وكانت فى الوقت نفسه قيم ص تدل على تراتيبهم فى مسابقة أخرى ، كانت قيم س ، وكذلك قيم ص ، عبارة عن الأعداد الطبيعية ١ ، ٢ ، ٣ . . . إلى ن ، مرتبة ترتيباً خاصاً .

الارتباط بين التراتيب . معامل سبيرمان

أى أن الوسط الحسابى لقيم س ، أو ص ، يساوى

$$\bar{\text{س}} = \frac{\text{١}}{\text{ن}} (\text{١} + \text{٢} + \text{٣} + \ldots + \text{ن}) = \frac{\text{١}}{\text{٢}} (\text{ن} + \text{١})$$

والانحراف المعيارى لها ع حيث[١].

$$\text{ن ع}^{\text{٢}} = \text{مج س}^{\text{٢}} - \text{ن}\ \bar{\text{س}}^{\text{٢}}$$

$$= \text{١}^{\text{٢}} + \text{٢}^{\text{٢}} + \text{٣}^{\text{٢}} + \ldots + \text{ن}^{\text{٢}} - \frac{\text{ن}}{\text{٤}} (\text{ن} + \text{١})^{\text{٢}}$$

$$= \frac{\text{ن (ن + ١) (٢ن + ١)}}{\text{٦}} - \frac{\text{ن}}{\text{٤}} (\text{ن} + \text{١})^{\text{٢}}$$

$$= \frac{\text{١}}{\text{١٢}} \,.\, \text{ن (ن + ١) (ن − ١)}$$

أى أن $\text{ع}^{\text{٢}} = \frac{\text{١}}{\text{١٢}} \,.\, (\text{ن}^{\text{٢}} - \text{١})$

وبالتعويض بهذه القيمة عن ع² فى المعادلة (٦) نحصل على معاملى الارتباط فى هذه الحالة الخاصة جداً ، وهو فى الحقيقة معامل الارتباط بين تراتيب الأشخاص فى المسابقتين . فإذا رمزنا له بالحرف الخاص ر تمييزاً له عن المعامل العام ر ، يكون

$$\text{ر} = \text{١} - \frac{\text{٦ مج (س − ص)}^{\text{٢}}}{\text{ن (ن}^{\text{٢}} - \text{١)}} \quad \ldots \ldots \quad (\text{١٦}) .$$

وقد استنبط اسبيرمان[٢] هـذا المعامل واستخدمه فى أبحاثه الخـاصة فى علم النفس ، وسيأتى ذكره فى مناسبة أخرى .

---

(١) لمعرفة مجموع مربعات الأعداد ١² + ٢² + ٣² + ... انظر مثـلا كتاب الجبر العالى تأليف هول ونايت باب جمع حدود المتسلسلات .

(٢) يسمى هـذا المعامل بالانجليزية Spearman's Rank-Coeffficient of Correlation ويرمز له بالحرف الأغريقى ρ (رّو) تمييزاً له عن المعـامل العادى الذى يرمز له بالحرف r أنظر مقالة W. Stephenson فى صفحة ٣٣٥ فى عدد يناير سنة ١٩٣٤ من مجلة : *The British Journal of Psychology*, General, Section, vol, XXIV, Part, 3.

حساب ر بواسطة الانحراف المعيارى للفرق والمجموع

٢١٠ — هناك معادلتان أخريان لحساب معامل الارتباط ر . وهما تعتمدان على معرفة الانحرافين المعياريين لقيم س وقيم ص ، والانحراف المعيارى للفرق بين القيم المتناظرة من س و ص ، أو الانحراف المعيارى للمجموع .

وهاتان المعادلتان هما (١) :

---

(١) لإثبات هاتين العلاقتين نفرض أن :

قيم س هى $س_{١} ، س_{٢} ، \dots ، س_{ن}$ ، وسطها الحسابى $\bar{س}$ ، وانحرافها المعيارى ع

« ص « $ص_{١} ، ص_{٢} ، \dots ، ص_{ن}$ « « $\bar{ص}$ ، « « $ع'$

الفروق « $ف_{١} = س_{١} - ص_{١} ، ف_{٢} = س_{٢} - ص_{٢} ، \dots$

وسطها الحسابى $\bar{ف} = \bar{س} - \bar{ص}$

والمجاميع « $ح_{١} = س_{١} + ص_{١} ، ح_{٢} = س_{٢} + ص_{٢} ، \dots$

وسطها الحسابى $\bar{ح} = \bar{س} + \bar{ص}$

$$\therefore \quad ن\, ع_{ف}^{٢} = \text{مج}(ف - \bar{ف})^{٢} ، \text{ و } ن\, ع_{ح}^{٢} = \text{مج}(ح - \bar{ح})^{٢}$$

$$\therefore \quad ف_{١} - \bar{ف} = (س_{١} - ص_{١}) - (\bar{س} - \bar{ص})$$

$$\text{و } ح_{١} - \bar{ح} = (س_{١} + ص_{١}) - (\bar{س} + \bar{ص})$$

$$ف_{١} - \bar{ف} = (س_{١} - \bar{س}) - (ص_{١} - \bar{ص})$$

$$\text{و } ح_{١} - \bar{ح} = (س_{١} - \bar{س}) + (ص_{١} - \bar{ص})$$

$$\therefore \quad (ف_{١} - \bar{ف})^{٢} = (س_{١} - \bar{س})^{٢} + (ص_{١} - \bar{ص})^{٢}$$
$$- ٢\, (س_{١} - \bar{س})\, (ص_{١} - \bar{ص}) ،$$

$$\text{و } (ف_{٢} - \bar{ف})^{٢} = (س_{٢} - \bar{س})^{٢} + (ص_{٢} - \bar{ص})^{٢}$$
$$- ٢\, (س_{٢} - \bar{س})\, (ص_{٢} - \bar{ص}) ،$$

$$\dots \quad \dots \quad \dots \quad \dots \quad \dots ،$$

$$(ف_{ن} - \bar{ف})^{٢} = (س_{ن} - \bar{س})^{٢} + (ص_{ن} - \bar{ص})^{٢}$$
$$- ٢\, (س_{ن} - \bar{س})\, (ص_{ن} - \bar{ص})$$ ؛ وبالجمع :

$$\therefore \quad ع_{ف}^{٢} = ع^{٢} + ع'^{٢} - ٢ \,.\, ر \,.\, ع \,.\, ع' ،$$

وبالمثل بالنسبة إلى المجاميع ح :

$$\therefore \quad ع_{ح}^{٢} = ع^{٢} + ع'^{٢} + ٢ \,.\, ر \,.\, ع \,.\, ع' .$$

$$ر = \frac{ع^{٢} + ع'^{٢} - ع_{ف}^{٢}}{٢\ ع\ .\ ع'} \quad ...\ ...\ ...\ ...\ ...\ (٧)\ ؛$$

أو

$$ر = \frac{ع_{ح}^{٢} - ع^{٢} - ع'^{٢}}{٢\ ع\ .\ ع'} \quad ...\ ...\ ...\ ...\ ...\ (٨)\ ؛$$

حيث ع و ع' هما الانحرافان المعياريان السينى والصادى على الترتيب ؛ و $ع_{ف}$ هو الانحراف المعيارى للفرق (س — ص) ؛ و $ع_{ح}$ هو الانحراف المعيارى للمجموع (س + ص) .

وفى بعض الأحيان نجد حساب ر بواسطة إحدى هاتين المعادلتين — وخصوصاً الأولى منهما — أسهل منه بواسطة المعادلات السابقة .

حساب المعامل ر عمليـاً

٢١١ — لنأخذ مثالاً عملياً ونحسب معامل الارتباط بهذه الطرق المختلفة . نأخذ مثلا أرقام (١) نسبة البطالة فى انجلترا وقيمة صادراتها بملايين الجنيهات فى العشر السنوات ١٩٢١ — ١٩٣٠ ، ونحسب معامل الارتباط بين هاتين الظاهرتين الاقتصاديتين . ونقتصر هنا على عشر قيم فقط لكل من المتغيرين ، حتى نرى خطوات العمل واضحة . ولكن فى المسائل العادية يجب ألا يقل عدد الحالات التى نبحثها عن ثلاثين حالة . وإلا كان تأثير الحالات الشاذة كبيراً جداً ، يحدث خطأ كبيراً فى النتيجة . ونجد الأرقام فى جدول ٢٩ .

نرى من الجدول أن $\overline{س}$ = ١٢٫٤١ ، و $\overline{ص}$ = ١٫٧١٤ ؛

وأن مجـ (س — $\overline{س}$)^{٢} = ١٠ ع^{٢} = ٥٨٫٤٢٩ ؛

و مجـ (ص — $\overline{ص}$)^{٢} = ١٠ ع'^{٢} = ٣٨٣٤٦٫٩ ؛

---

(١) الأرقام مأخوذة عن Recueil International de Statistique, 1919-1930 طبعة المعهد الدولى للاحصاء فى لاهاى سنة ١٩٣٤ (صحيفتى ١١٢ — ١٩٣)

و مج (س — س̄) (ص — ص̄) = — ٨٧٣٫٦١٠ ؛

∴ ع = ٢٫٤١٧٢ ، و عَ = ٦١٫٩٢٤٩ ؛

∴ ر = — ٨٧٣٫٦١٠ / (١٠ × ١٤٩٫٦٨٤٩) ، حسب المعادلة (١) بند ٢٠٣

= — ٠٫٥٨٤ ؛

أى أنه ارتباط عكسى . وهذا هو المنتظر فى مثل هذه الحالة . إذ أن زيادة قيمة الصادرات تكون على العموم مصحوبة بزيادة فى كميتها ، أى بزيادة فى الإنتاج الصناعى ونقص فى عدد العمال العاطلين .

جدول ٢٩ — حساب معامل الارتباط بين نسبة البطالة وقيمة الصادرات فى انجلترا ، باستخدام الصيغة الأساسية

السنة	نسبة البطالة س	قيمة الصادرات ص	س — س̄	ص — ص̄	(س — س̄)٢	(ص — ص̄)٢	(س — س̄) × (ص — ص̄)
١٩٢١	١٧٫٠٪	٧٠٣	٤٫٥٩	— ١١٫١	٢١٫٠٦٨١	١٢٣٫٢١	— ٥٠٫٩٤٩
٢٢	١٤٫٣	٧٢٠	١٫٨٩	٥٫٩	٣٫٥٧٢١	٣٤٫٨١	١١٫١٥١
٢٣	١١٫٧	٧٦٧	— ٫٧١	٥٢٫٩	٫٥٠٤١	٢٧٩٨٫٤١	— ٣٧٫٥٥٩
٢٤	١٠٫٣	٨٠١	— ٢٫١١	٨٦٫٩	٤٫٤٥٢١	٧٥٥١٫٦١	— ١٨٣٫٣٥٩
٢٥	١١٫٣	٧٧٣	— ١٫١١	٥٨٫٩	١٫٢٣٢١	٣٤٦٩٫٢١	— ٦٥٫٣٧٩
٢٦	١٢٫٥	٦٥٣	٫٠٩	— ٦١٫١	٫٠٠٨١	٣٧٣٣٫٢١	— ٥٫٤٩٩
٢٧	٩٫٧	٧٠٩	— ٢٫٧١	— ٥٫١	٧٫٣٤٤١	٢٦٫٠١	١٣٫٨٢١
٢٨	١٠٫٨	٧٢٤	— ١٫٦١	٩٫٩	٢٫٥٩٢١	٩٨٫٠١	— ١٥٫٩٣٩
٢٩	١٠٫٤	٧٢٠	— ٢٫٠١	٥٫٩	٤٫٠٤٠١	٣٤٫٨١	— ١١٫٨٥٩
٣٠	١٦٫١	٥٧١	٣٫٦٩	— ١٤٣٫١	١٣٫٦١٦١	٢٠٤٧٧٫٦١	— ٥٢٨٫٠٣٩
المجموع	١٢٤٫١	٧١٤١			٥٨٫٤٢٩٠	٣٨٣٤٦٫٩٠	— ٨٧٣٫٦١٠
المتوسط	١٢٫٤١	٧١٤٫١			٥٫٨٤٢٩	٣٨٣٤٫٦٩	

طريقة أخرى لحساب ر

٢١٢ — يمكننا أن نتفادى الكسور الموجودة فى الوسطين الحسابيين والانحرافات عنهما ، وعمليات ضربها الموجودة فى الأعمدة الثلاثة الأخيرة من جدول ٢٩ ، بأن نسير فى العمل بالخطوات المبينة فى الجدول الآتى :

جدول ٣٠ — حساب معامل الارتباط بين نسبة البطالة وقيمة الصادرات فى انجلترا

باستخدام الصيغة (٢) صحيفة ٢٢٣

السنة	نسبة البطالة س	قيمة الصادرات ص	$\text{س}^{٢}$	$\text{ص}^{٢}$	س × ص
١٩٢١	١٧٫٠	٧٠٣	٢٨٩٫٠٠	٤٩٤٢٠٩	١١٩٥١٫٠
٢٢	١٤٫٣	٧٢٠	٢٠٤٫٤٩	٥١٨٤٠٠	١٠٢٩٦٫٠
٢٣	١١٫٧	٧٦٧	١٣٦٫٨٩	٥٨٨٢٨٩	٨٩٧٣٫٩
٢٤	١٠٫٣	٨٠١	١٠٦٫٠٩	٦٤١٦٠١	٨٢٥٠٫٣
٢٥	١١٫٣	٧٧٣	١٢٧٫٦٩	٥٩٧٥٢٩	٨٧٣٤٫٩
٢٦	١٢٫٥	٦٥٣	١٥٦٫٢٥	٤٢٦٤٠٩	٨١٦٢٫٥
٢٧	٩٫٧	٧٠٩	٩٤٫٠٩	٥٠٢٦٨١	٦٨٧٧٫٣
٢٨	١٠٫٨	٧٢٤	١١٦٫٦٤	٥٢٤١٧٦	٧٨١٩٫٢
٢٩	١٠٫٤	٧٢٠	١٠٨٫١٦	٥١٨٤٠٠	٧٤٨٨٫٠
٣٠	١٦٫١	٥٧١	٢٥٩٫٢١	٣٢٦٠٤١	٩١٩٣٫١
مجموع . .	١٢٤٫١	٧١٤١	١٥٩٨٫٥١	٥١٣٧٧٣٥	٨٧٧٤٦٫٢
متوسط .	١٢٫٤١	٧١٤٫١			

$$\therefore\ \text{مج س}^{٢} = \text{١٥٩٨٫٥١} = \text{١٠ ع}^{٢} + \bar{\text{س}}^{٢} \text{ ،}$$

$$\therefore\ \text{١٠ ع}^{٢} = \text{١٥٩٨٫٥١} - \text{١٠}\ (\text{١٢٫٤١})^{٢} \text{ ،}$$

$$= \text{٠٥٨٫٤٢٩}$$

$$\text{و مج ص}^{٢} = \text{٥١٣٧٧٣٥} = \text{١٠ ع}^{٢} + \text{١٠}\ \bar{\text{ص}}^{٢} \text{ ،}$$

$$\therefore\ \text{١٠ ع}^{٢} = \text{٥١٣٧٧٣٥} - \text{١٠}\ (\text{٧١٤٫١})^{٢} \text{ ،}$$

$$= \text{٠٣٨٣٤٦٫٩}$$

$$\text{مجـ س ص} = ٨٧٧٤٦{,}٢ \quad \text{و} \quad \text{ن}\ \bar{\text{س}}\ \bar{\text{ص}} = ١٠ \times ٨٨٦١{,}٩٨١$$

$$\therefore \text{مجـ س ص} - \text{ن}\ \bar{\text{س}}\ \bar{\text{ص}} = - ٨٧٣{,}٦١$$

$$= \text{مجـ}\ (\text{س} - \bar{\text{س}})\ (\text{ص} - \bar{\text{ص}})\ .$$

$$\therefore \quad \text{ر} = - ٫٥٨٤$$

ويلاحظ أن العمليات الحسابية هنا كبيرة أيضاً ، ولكننا نحصل على نفس القيم للانحرافين المعياريين ومعامل الارتباط .

حساب ر باختيار وسطين فرضيين

٢١٣ — نشرح الآن طريقة حساب ر باختيار وسط فرضى مناسب لكل من س و ص . ويستحسن اختيار إحدى القيم المعطاة كوسط فرضى ، بحيث تكون قريبة من الوسط الحسابى . ولو اخترنا وسطى س و ص الفرضيين فى سطرين مختلفين فى الجدول ، حصلنا على انحرافين كل منهما يساوى صفراً ، ينتجان حاصلى ضرب كل منهما يساوى صفراً أيضاً . وهذا مما يسهل عمليات الضرب . لنأخذ الوسط الفرضى لقيم س يساوى ١٢٫٥ (وهى نسبة البطالة فى سنة ١٩٢٦) والوسط الفرضى لقيم ص يساوى ٧٠٩ (وهى قيمة الصادرات فى سنة ١٩٢٧) . وهذان الوسطان قريبان من الوسطين الحسابيين ، وهما ١٢٫٤١ و ٧١٤٫١ على الترتيب .

ونرى خطوات العمل موضحة فى الجدول رقم ٣١ :

$$\therefore \text{مجـ}\ (\text{س} - ١٢{,}٥)^{٢} = ٥٨{,}٥١ = ١٠\ \text{ع}^{٢} + ١٠\ (\bar{\text{س}} - ١٢{,}٥)^{٢}.$$

$$\therefore ١٠\ \text{ع}^{٢} = ٥٨{,}٥١ - ١٠\ (- ٫٠٩)^{٢}\ ،$$

$$= ٥٨{,}٤٢٩\ .$$

و مجـ (ص — ٧٠٩)^٢ = ٣٨٦٠٧ = ١٠ عَ^٢ + ١٠ (صَ — ٧٠٩)^٢

∴ ١٠ عَ^٢ = ٣٨٦٠٧ — ١٠ (٥,١)^٢

= ٣٨٣٤٦,٩ .

و مجـ (س — ١٢,٥) (ص — ٧٠٩) — ١٠ (١٢,٤١ — ١٢,٥) (٧١٤,١ — ٧٠٩)

= — ٨٧٨,٢ — ١٠ × — ٠,٠٩ × ٥,١ ،

= — ٨٧٣,٦١ .

∴ ر = [مجـ (س — و) (ص — وَ) — ن ح حَ] ÷ [ن ع . عَ]

= — ٨٧٣,٦١ ÷ ١٤٩٦,٨٤٩

= — ٠,٥٨٤ ؛

وهى نفس النتائج التى حصلنا عليها من قبل [١] .

هذه الطريقة أخصر من الأولى دائماً ومن الثانية أحياناً

واضح من هذا المثال أن الطريقة الأخيرة أسهل من الطريقتين المتقدمتين، وأنه يمكن للإنسان أن يختار أوساطاً فرضية مناسبة تختصر العمل الحسابى إلى حد كبير . والواقع أننا نستعمل هذه الطريقة فى أغلب الحالات إن لم يكن كلها ؛ وإنما أوردنا الطرق الثلاث لمعرفة خطوات العمل فى كل طريقة ؛ وليتبين أيضاً أن النتائج التى نصل إليها واحدة مهما كانت الطريقة المتبعة . وعلى كل حال فالاختيار بين هذه الطرق الثلاث يتوقف على ظروف المسألة التى نحن بصددها وسهولة الأرقام المركبة منها أو صعوبتها ؛ فنستخدم الطريقة الثانية حينما تكون قيم س ، ص قليلة الأرقام وخالية من الكسور ، حيث تسهل عمليات الضرب والتربيع ؛ أما الطريقة الأولى فلا تستخدم إلا نادراً .

---

(١) يمكن للقارىء أن يحسب ر بواسطة إحدى المعادلتين المذكورتين فى بند ٢١٠ فنأخذ بدل العمود الأخير فى جدول ٣١ قيم الفروق س — ص مثلا ، ثم نربعها ونحسب انحرافها المعيارى ، ونعوضه فى المعادلة (٧) بند ٢١٠

جدول ٣١ — حساب معامل الارتباط بين نسبة البطالة وقيمة الصادرات فى انجلترا
بطريقة استخدام وسطين فرضيين حسب المعادلة (٣) صحيفة ٢٢٤

السنة	نسبة البطالة س	قيمة الصادرات ص	س — ١٢٫٥	ص — ٧٠٩	(س — ١٢٫٥)٢	(ص — ٧٠٩)٢	(س — ١٢٫٥) × (ص — ٧٠٩)
١٩٢١	١٧٫٠	٧٠٣	٤٫٥	−٦	٢٠٫٢٥	٣٦	−٢٧٫٠
٢٢	١٤٫٣	٧٠٠	١٫٨	١١	٣٫٢٤	١٢١	١٩٫٨
٢٣	١١٫٧	٧٦٧	−٫٨	٥٨	٫٦٤	٣٣٦٤	−٤٦٫٤
٢٤	١٠٫٣	٨٠١	−٢٫٢	٩٢	٤٫٨٤	٨٤٦٤	−٢٠٢٫٤
٢٥	١١٫٣	٧٧٣	−١٫٢	٦٤	١٫٤٤	٤٠٩٦	−٧٦٫٨
٢٦	١٢٫٥	٦٥٣	٠	−٥٦	٠	٣١٣٦	٠
٢٧	٩٫٧	٧٠٩	−٢٫٨	٠	٧٫٨٤	٠	٠
٢٨	١٠٫٨	٧٢٤	−١٫٧	١٥	٢٫٨٩	٢٢٥	−٢٥٫٥
٢٩	١٠٫٤	٧٢٠	−٢٫١	١١	٤٫٤١	١٢١	−٢٣٫١
٣٠	١٦٫١	٥٧١	٣٫٦	−١٣٨	١٢٫٩٦	١٩٠٤٤	−٤٩٦٫٨
المجموع	١٢٤٫١	٧١٤١			٥٨٫٥١	٣٨٦٠٧	−٨٧٨٫٢
المتوسط	١٢٫٤١	٧١٤٫١					

يجب ألا يقل عدد الحالات التى نحسب منها عن ٣٠

٢١٤ — ونكرر هنا أن عدد الحالات التى أخذناها فى هذا المثال التوضيحى صغير جداً ، ويجعل النتيجة تحت رحمة أخطاء المصادفات إلى حد كبير. لأن زيادة كبيرة تحدث عن طريق المصادفة فى قيمة الصادرات فى أى سنة لسبب ما (لارتفاع الأسعار فجأة مثلا ، أو فى نسبة البطالة) ، قد يترتب عليها خطأ كبير فى النتيجة . ولذلك يجب أن يكون عدد الحالات التى نأخذها كبيراً — ٣٠ على الأقل — حتى يكون هناك فرصة لتعادل تأثير الحالات الشاذة مع بعضها .

ولكن إذا زاد عدد الحالات كثيراً ، فلا شك أن العمل الحسابى يكون مرهقاً للغاية لو احتفظنا بكل قيمة على انفرادها مع نظيرتها كما فعلنا فى المثال السابق . فلا بد إذن أن نفكر فى طريقة مختصرة نستخدمها عندما يكون لدينا

عدد كبير من الحالات — ٢٠٠ أو ٣٠٠ مثلا — بحيث لا تكلفنا مجهوداً كبيراً زيادة عن اللزوم ، وفى الوقت نفسه تعطينا نتائج دقيقة .

جدول الارتباط

٢١٥ — لنفرض أننا نبحث فى العلاقة بين عمر الرجل وعدد ما عنده من الأطفال ؛ وأننا بحثنا حالة ٢٠٠ رجل متزوج ، وعرفنا عمر كل رجل وعدد أطفاله ، فحصلنا على البيان الآتى :

رقم مسلسل	الاسم	العمر	عدد الأطفال
١	أحمد	٢٨	٣
٢	إبراهيم	٢٢	٠
٣	حسين	٤١	٤
..	..	..	..
..	..	..	..
٢٠٠	يوسف	٣٨	٦

نقسم هؤلاء الرجال إلى فئات مناسبة من حيث أعمارهم ولتكن هذه الفئات هى : ٢٠ وأقل من ٢٥ ؛ ٢٥ وأقل من ٣٠ و ... و ٤٠ وأقل من ٤٥ سنة .

وكذلك نقسمهم إلى فئات من حيث عدد الأطفال . وليكن عدد الأطفال فى هذه الفئات هو : ٠ و ١ و ٢ و ... و ٦ . ثم نوزع هؤلاء الرجال فى « **جدول تكرار مزدوج** » أو « **جدول ارتباط** [١] » كالآتى :

(١) اسمه بالانجليزية Correlation Table or Double Frequency Table

جدول ٣٢ — توزيع تكرارى مزدوج لأعمار وعدد أطفال ٢٠٠ رجلا(١)

العمر س / عدد الأطفال ص	٢٠ —	٢٥ —	٣٠ —	٣٥ —	٤٠ —	المجموع
٠	٦	٨	٩	١		٢٤
١	١	٢٥	١١	٧	٣	٤٧
٢	٤	١٣	١٥	٢٠	٦	٥٨
٣		٦	١٨	١٠	٧	٤١
٤		١	٦	٨	٥	٢٠
٥			١	٢	٥	٨
٦				٢		٢
المجموع	١١	٥٣	٦٠	٥٠	٢٦	٢٠٠

فنرى مثلا أن أحمد يدخل فى الخانة ملتقى العمود الثالث بالسطر الخامس لأن عمره ٢٨ سنة (أى فى الفئة ٢٥ — ٣٠) وعدد أطفاله ٣ ، وبالمثل يوضع إبراهيم فى الخانة ملتقى العمود الثانى بالسطر الثانى لأن عمره ٢٢ سنة وليس عنده أطفال بالمرة . وهكذا إلى آخر الكشف حيث يوضع يوسف فى الخانة الأخيرة من العمود الخامس . وبعد توزيع كل الرجال على الخانات بهذه الطريقة نعد الحالات التى وقعت فى كل خانة ، فنحصل على التكرارات الموجودة فى الجدول أعلاه . وهو جدول تكرارى لأن العدد ١٣ مثلا ، الذى نراه

(١) البيانات مأخوذة من بحث عمله المؤلف عن الحالة المعيشية للعمال فى القاهرة سنة ١٩٣٥

فى ملتقى العمود الثالث بالسطر الرابع ، هو عدد الرجال الذين عمرهم بين ٢٠ و ٢٥ سنة وعند كل واحد منهم ٢ من الأطفال .

أما عمود المجاميع الأخير فهو عبارة عن التوزيع التكرارى لعدد الأطفال عند هؤلاء الرجال ، حيث منهم ٢٤ ليس عندهم أطفال بالمرة ، و ٤٧ عند كل منهم طفل واحد ، و ٥٨ عند كل منهم اثنان ، وهكذا .

وكذلك السطر الأخير ، فهو توزيع تكرارى لأعمار هؤلاء الرجال : فنرى منهم ١١ رجلا أعمارهم ٢٠ وأقل من ٢٥ سنة ، و ٥٣ رجلا أعمارهم ٢٥ وأقل من ٣٠ سنة . وهكذا .

٢١٦ — لحساب معامل الارتباط هنا نوجد أولا الوسطين الحسابيين والانحرافين المعياريين للأعمار وعدد الأطفال ، وذلك باستخدام التوزيعين التكراريين فى السطر الأخير والعمود الأخير . حساب معامل الارتباط من الجدول المزدوج

جدول ٣٣ — إيجاد الوسط الحسابى والانحراف المعيارى
لأعمار الرجال فى جدول ٣٢

فئات العمر	مراكز الفئات س	التكرار ك	س — ٣٢٫٥ / ح	ك. ح	ك. ح٢
٢٠—	٢٢٫٥	١١	—١٠	—١١٠	١١٠٠
٢٥—	٢٧٫٥	٥٣	—٥	—٢٦٥	١٣٢٥
٣٠—	٣٢٫٥	٦٠	٠	٠	٠
٣٥—	٣٧٫٥	٥٠	٥	٢٥٠	١٢٥٠
٤٠—	٤٢٫٥	٢٦	١٠	٢٦٠	٢٦٠٠
		٢٠٠		١٣٥	٦٢٧٥

∴ الوسط الحسابى للأعمار هو :

$\bar{س} = ٣٢,٥ + \frac{١٣٥}{٢٠٠}$

$= ٣٣,١٧٥$ سنة .

و $٦٢٧٥ = ٢٠٠ ع^٢ + ٢٠٠ (٦٧٥,)^٢$

$\therefore ع^٢ = ٣١,٣٧٥ - ٤٥٥٦,$

$= ٣٠,٩١٩٤$

$\therefore ع = ٥,٥٦٠٥$ سنة .

جدول ٣٤ — إيجاد الوسط الحسابى والانحراف المعيارى
لعدد الأطفال عند الرجال فى جدول ٣٢

عدد الأطفال ص	التكرار كَ	ص — ٢ عَ	كَ . عَ	كَ . عَ^٢
٠	٢٤	—٢	—٤٨	٩٦
١	٤٧	—١	—٤٧	٤٧
٢	٥٨	٠	٠	٠
٣	٤١	١	٤١	٤١
٤	٢٠	٢	٤٠	٨٠
٥	٨	٣	٢٤	٧٢
٦	٢	٤	٨	٣٢
	٢٠٠		١٨	٣٦٨

∴ الوسط الحسابى لعدد الأطفال هو :

$$\overline{\text{ص}} = ٢ + \frac{١٨}{٢٠٠}$$

$$= ٢٫٠٩ \text{ طفلا}$$

$$\text{و} \qquad ٣٦٨ = ٢٠٠\,\text{عَ}^{٢} + ٢٠٠\,(٫٠٩)^{٢}$$

$$\therefore \quad \text{عَ}^{٢} = ١٫٨٤ - ٫٠٠٨١$$

$$= ١٫٨٣١٩$$

$$\therefore \quad \text{عَ} = ١٫٣٥٣٥ \text{ طفلا .}$$

ولحساب معامل الارتباط فى مثل هذه المسائل ، نأخذ وسطين فرضيين مناسبين للسينات والصادات ، ونحسب حواصل ضرب الانحرافات عن هذين الوسطين ، ونستخدم المعادلة .

$$\text{ر} = \frac{\text{مج}\,\overset{\text{ح}}{(\text{س} - \text{و})}\,\overset{\text{حَ}}{(\text{ص} - \text{وَ})} - \text{ن ح حَ}}{\text{ن . ع . عَ}}$$

$$\frac{\text{مج ح حَ} - \text{ن ح حَ}}{\text{ن ع عَ}}$$

ويستحسن عملياً أن نختار مركز إحدى الفئات السينية كوسط فرضى للسينات ، وكذلك فى الصادات .

لنأخذ الوسط الفرضى للأعمار و = ٣٢٫٥ سنة ، والوسط الفرضى لعدد الأطفال وَ = ٢ .

$$\therefore \text{ح} = \overline{\text{س}} - \text{و} = ٣٣٫١٧٥ - ٣٢٫٥ \quad = ٠٫٦٧٥$$

$$\text{و} \; \text{حَ} = \overline{\text{ص}} - \text{وَ} = ٢٫٠٩ - ٢٫٠٠ \quad = ٫٠٩$$

$$\therefore \text{ن ح حَ} = ٢٠٠ \times ٫٦٧٥ \times ٫٠٩ = ١٢٫١٥٠$$

نحسب الانحرافات السينية عن ٣٢٫٥ ، والانحرافات الصادية عن ٢ . فنطرح ٣٢٫٥ من مراكز الفئات السينية المبينة فى رؤوس الأعمدة فى جدول الارتباط ، ونطرح ٢ من مراكز الفئات الصادية المبينة فى العمود الأيمن من نفس الجدول . وطبعاً يكون الانحراف السينى لأى خانة فى الجدول هو نفس الانحراف السينى للعمود الذى تقع فيه هـذه الخانة ؛ ويكون الانحراف الصادى لها نفس الانحراف الصادى للسطر الذى تقع فيه أيضاً .

لكل خانة نحسب حاصل الضرب الآتى :

تكرار الخانة × انحرافها السينى × انحرافها الصادى

ومجموع هـذه الحواصل لجميع الخانات يساوى الكمية التى نبحث عنها ، وهى مجـ (س — و) (ص — وَ) .

إذا اخترنا الوسطين الفرضيين عند مركزى فئتين ، فسنجد عملياً أن الحواصل فى عمود الوسط الفرضى السينى ، وفى سطر الوسط الفرضى الصادى ، كلها تساوى صفراً ، لأن الانحراف السينى فى هـذا العمود صفر ، والانحراف الصادى فى هذا السطر صفر أيضاً . فيمكننا إذن أن نهمل هذا العمود وهذا السطر بالمرة ، ونشطبهما من الجدول .

ولتسهيل عمليات الضرب وضمان الدقة وعدم السهو فى العمـل ، نكتب الانحراف السينى لكل خانة فى ركن معين منها ، والانحراف الصادى لها فى ركن آخر معين أيضاً ، وحاصل ضرب الاثنين فى التكرار فى ركن ثالث من الخانة نفسها . ويجب تمييز الكتابة فى هذه الأركان المختلفة فنستعمل ألواناً مختلفة من الحبر أو الرصاص مثلا ، منعاً للالتباس . ونرى هذه الخطوات موضحة فى الجدول الآتى :

ويلاحظ أننا هنا كتبنا الانحرافات السينية فى الركن الجنوبى الغربى من كل خانة ، والانحراف الصادى فى الركن الجنوبى الشرقى . وكتبنا الحاصل النهائى فى الركن الشمالى الشرقى .

جدول ٣٥ — حساب معامل الارتباط من جدول تكرارى مزدوج باختيار وسطين فرضيين

ص \ س	٢٢,٥	٢٧,٥	٣٢,٥	٣٧,٥	٤٢,٥	المجموع	الحواصل
٠	١٢٠ ٦ ٢- (١٠-)	٨٠ ٨ ٢- (٥-)	٩	١٠- ١ ٢- (٥)		٢٤	٢٠٠ ١٠ -
١	١٠ ١ ١- (١٠-)	١٢٥ ٢٥ ١- (٥-)	١١	٣٥- ٧ ١- (٥)	٣٠- ٣ ١- (١٠)	٤٧	١٢٥ ٦٥ -
٢	٤	١٣	١٥	٢٠	٦	٥٨	
٣		٣٠- ٦ ١ (٥-)	١٨	٥٠ ١٠ ١ (٥)	٧٠ ٧ ١ (١٠)	٤١	١٢٠ ٣٠ -
٤		١٠- ١ ٢ (٥-)	٦	٨٠ ٨ ٢ (٥)	١٠٠ ٥ ٢ (١٠)	٢٠	١٨٠ ١٠ -
٥			١	٣٠ ٢ ٣ (٥)	١٥٠ ٥ ٣ (١٠)	٨	١٨٠
٦				٤٠ ٢ ٤ (٥)		٢	٤٠
المجموع	١١	٥٣	٦٠	٥٠	٢٦	٢٠٠	٨٥٥ ١١٥ -
الحواصل	١٣٠	٢٠٥ ٤٠ -		٢٠٠ ٤٥ -	٣٢٠ ٣٠ -	٨٥٥ ١١٥ -	٧٤٠

ويلاحظ أن مجموع الحواصل واحد ، سواء حسبناه أفقياً بالسطور أو رأسياً بالأعمدة . ويحسن مراجعة المجموعين على بعضهما منعاً للخطأ . والمجموع النهائى يساوى ، كما قلنا من قبل ،

$$\text{مج}(س - و)\,(ص - وَ) = ٧٤٠$$

$$\therefore\ r = \frac{٧٤٠ - ١٢{,}١٥}{٢٠٠ \times ٥{,}٥٦٠٥ \times ١{,}٣٥٣٥}$$

$$= \frac{٧٢٧{,}٨٥}{١٥٠٥{,}٢٢٧}$$

$$= ٠{,}٤٨٤$$

وهذا المعامل يدل على وجود علاقة طردية بين سن الرجل وعدد الأطفال الذين يعولهم ؛ وهذا هو المنتظر بطبيعة الحال . ومقدار المعامل ر فى هذه الحالة يدل على أن الارتباط شديد نوعاً .

طريقة المجموعات القطرية

٢١٧ — نشرح الآن طريقة أسهل لحساب ر من الجدول التكرارى المزدوج ، وهى مبنية على الفكرة التى شرحناها فى بند ٢١٠ . وهى كما يأتى :

أقطار الفروق المتساوية

فى جدول ٣٢ صفحة ٢٣٥ ، نأخذ الأقطار النازلة من اليمين إلى اليسار(١) ، وهى التى تبدأ من الخانات ذات القيم الصغرى لكل من س و ص وتنتهى بالقيم الكبرى لهما . ونجمع تكرارات الخانات الواقعة على كل قطر من هذه الأقطار المتوازية . وطبعاً ستكون الخانات المتتالية على أى قطر ، مشتركة فى ركن واحد دائماً . ونأخذ حواصل جمع التكرارات لكل قطر على حدة ، ونكتبها بالترتيب حسب هذه الأقطار .

وفيما يلى بيان بتكرارات الفئات الواقعة على الأقطار بالترتيب من أعلى إلى أسفل . وسنكتب تكرارات كل قطر حسب مواقعها على نفس القطر فى جدول ٣٢ :

---

(١) هذا طبعاً على فرض أن ترتيب فئات س فى الجدول تصاعدى من اليمين إلى اليسار ، وفئات ص متصاعدة من فوق إلى تحت . وهذا هو النظام المعتاد . إلا أنه أحياناً قد يعكس ترتيب فئات س أو ص أو هما معاً . وعلى كل حال فالمقصود هو الأقطار التى تبدأ من ناحية القيم الصغيرة للمتغيرين معاً ، وتنتهى فى ناحية القيم الكبيرة لهما .

القطر الأول	١ + ٣	= ٤
« الثانى	٩ + ٧ + ٦	= ٢٢
« الثالث	٨ + ١١ + ٢٠ + ٧	= ٤٦
« الرابع	٦ + ٢٥ + ١٥ + ١٠ + ٥	= ٦١
« الخامس	١ + ١٣ + ١٨ + ٨ + ٥	= ٤٥
« السادس	٤ + ٦ + ٦ + ٢	= ١٨
« السابع	١ + ١ + ٢	= ٤
		٢٠٠

نعتبر فئات س كأنها مراتب متتالية لقيم س، وفئات ص كأنها مراتب متتالية لقيمها، فيكون لدينا خمس مراتب سينية وسبع صادية. وبالتأمل فى الخانات التى على القطر الأول، نجد فى كل منها أن الفرق بين مرتبتى س و ص يساوى + ٣، وفى القطر الثانى الفرق يساوى + ٢، ...، وهكذا إلى القطر السابع حيث الفرق بين مرتبتى س و ص فى كل خانة يساوى — ٣. وهكذا نسمى هذه الأقطار **أقطار الفروق المتساوية**.

٢١٨ — يمكننا إذن اعتبار حواصل جمع التكرارات على هذه الأقطار كأنها تكرارات لهذه الفروق المختلفة بين مراتب س و ص، ثم نطبق الفكرة التى أوردناها فى بند ٢١٠، بخصوص الانحراف المعيارى للفرق بين قيم س و ص. ونسير فى العمل كما فى الجدول الآتى:

$$\therefore\ ٢٠٠\ ع_{ى}^{٢} = ٣٢٣ - \frac{(٩)^{٢}}{٢٠٠}$$

$$= ٣٢٢{,}٥٩٥$$

$$= ح \quad \text{مثلا.}$$

جدول ٣٦ — حساب الانحراف المعيارى للفرق بين مرتبتى س ، ص .

الفروق ن	التكرارات ك	ك . ن	ك . ن^{٢}
٣	٤	١٢	٣٦
٢	٢٢	٤٤	٨٨
١	٤٦	٤٦	٤٦
٠	٦١	٠	٠
—١	٤٥	—٤٥	٤٥
—٢	١٨	—٣٦	٧٢
—٣	٤	—١٢	٣٦
	٢٠٠	٩	٣٢٣

يلاحظ أنه كان من الممكن الاستغناء عن كتابة العمود الأول من هذا الجدول ، وكتابة صفر أمام التكرار الأوسط ، أو أى تكرار قريب منه ؛ وكتابة — ١ و — ٢ و— ٣ إلى أعلى و + ١ و + ٢ و + ٣ إلى أسفل ، أو العكس، بدون تأثير فى النتيجة النهائية .

٢١٩ — ونعمل مثل ذلك فى فئات س باعتبارها مراتب ، تكراراتها هى الأرقام الموجودة فى السطر الأخير من جدول ٣٢ . وكذلك فى فئات ص باعتبارها مراتب وتكراراتها موجودة فى العمود الأخير فى نفس الجدول . فنرى من جدول ٣٧

$$\text{أن :}\quad ٢٠٠\ ع^{٢} = ٢٥١ - \frac{(٢٧)^{٢}}{٢٠٠}$$

$$= ٢٤٧٫٣٥٥$$

$$= \text{ا}\quad \text{مثلا .}$$

جدول ٣٧ — حساب الانحراف المعيارى لمراتب س

مراتب س	تكرارات ك	ح	ح . ك	ح . ك
١	١١	٢—	٢٢—	٤٤
٢	٥٣	١—	٥٣—	٥٣
٣	٦٠	٠	٠	٠
٤	٥٠	١	٥٠	٥٠
٥	٢٦	٢	٥٢	١٠٤
	٢٠٠		٢٧	٢٥١

جدول ٣٨ — حساب الانحراف المعيارى لمراتب ص

مراتب ص	تكرارات كَ	حَ	حَ . كَ	حَ٢ . كَ
١	٢٤	٢—	٤٨—	٩٦
٢	٤٧	١—	٤٧—	٤٧
٣	٥٨	٠	٠	٠
٤	٤١	١	٤١	٤١
٥	٢٠	٢	٤٠	٨٠
٦	٨	٣	٢٤	٧٢
٧	٢	٤	٨	٣٢
	٢٠٠		١٨	٣٦٨

وبالمثل نرى من جدول ٣٨ أن :

$$\text{٢٠٠ ع}'^{\text{٢}} = \text{٣٦٨٠} - \frac{(\text{١٨})^{\text{٢}}}{\text{٢٠٠}}$$

$$= \text{٣٦٦,٣٨}$$

$$= \text{ب} .$$

$$\text{لكن} \quad \text{ر} = \frac{\text{ع}^{\text{٢}} + \text{ع}'^{\text{٢}} - \text{ع}_{\text{ف}}^{\text{٢}}}{\text{٢ ع . ع}'} \text{ ، بند ٢١٠ (٧)}$$

$$\therefore \quad \text{ر} = \frac{\text{ا} + \text{ب} - \text{ح}}{\text{٢} \sqrt{\text{ا ب}}}$$

$$= \frac{\text{٢٤٧,٣٥٥} + \text{٣٦٦,٣٨} - \text{٣٢٢,٥٩٥}}{\text{٢} \sqrt{\text{٩٠٦٢٥,٩٢٤٩}}}$$

$$= \frac{\text{٢٩١,١٤٠}}{\text{٦٠٢,٠٨٢٨}}$$

$$= \text{٠,٤٨٤}$$

وهى نفس النتيجة التى حصلنا عليها من قبل . ولا شك أن هذه الطريقة أسهل بكثير من الطريقة السابقة فى هذا المثال ، ولكنها فى بعض الأحيان ربما تكون أطول . ويمكننا اختصار العمل إلى حد كبير لو استعملنا ورقة مقسمة إلى مربعات وسرنا فى العمل كما فى جدول ٣٩[1] .

أقطار المجاميع المتساوية

٢٢٠ — يمكننا أيضاً أن نحسب ر من المعادلة (٨) بند ٢١٠ ، أى باستخدام فكرة الانحراف المعيارى لمجموع مرتبتى س و ص ، بدل الفرق بينهما كما فعلنا فى الطريقة السابقة . فإذا أخذنا الأقطار المتعامدة مع الأقطار السابقة ، وجدنا أن مجموع مرتبتى س و ص فى كل الخانات التى على قطر واحد ثابت . وهذه الأقطار إذن هى **أقطار المجاميع المتساوية** ؛ وهذه بيانها كما يأتى :

---

(١) انظر كتاب *How to Calculate Correlation* تأليف G. Thompson 1924

جدول (٣٩) لحساب معامل الارتباط

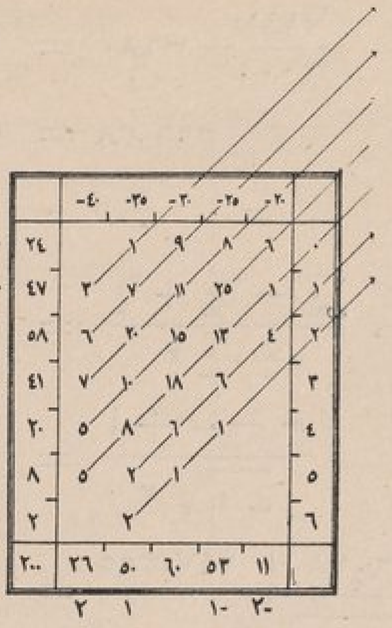
	حك	ح	ك
٣٦	-١٢	-٣	٤
٨٨	-٤٤	-٢	٢٢
٤٦	-٤٦	-١	٤٦
.	.	.	٦١
٤٥	٤٥	١	٤٥
٧٢	٣٦	٢	١٨
٣٦	١٢	٣	٤
٣٢٣	-٩		٢٠٠

مربع = ٨١

حـ = ٣٢٣ - $\frac{٨١}{٢٠٠}$

= ٣٢٣ - ٤٠٥,

= ٣٢٢,٥٩٥



-٣	-٤٨	٩٦
-١	-٤٧	٤٧
.	.	.
١	٤١	٤١
٢	٤٠	٨٠
٣	٢٤	٧٢
٤	٨	٣٢
	١٨	٣٦٨

مربع = ٣٢٤

٧٢٩ = مربع ٢٧ = -٢٢ -٥٢ . ٥٠ ٥٢

٢٥١ = ٤٤ ٥٣ . ٥٠ ١٠٤

ا = ٢٥١ - $\frac{٧٢٩}{٢٠٠}$ = ٢٥١ - ٣,٦٤٥

= ٢٤٧,٣٥٥

ب = ٣٦٨ - $\frac{٣٢٤}{٢٠٠}$

= ٣٦٨ - ١,٦٢

= ٣٦٦,٣٨٠

ا = ٢٤٧,٣٥٥

ا+ب = ٦١٣,٧٣٥

-حـ = ٣٢٢,٥٩٥

ا+ب-حـ = ٢٩١,١٤٠

القطر	مجموع المرتبتين	تكرارات الخانات		جملة التكرار
الأول	٢	٦	=	٦
الثانى	٣	١ + ٨	=	٩
الثالث	٤	٤ + ٢٥ + ٩	=	٣٨
الرابع	٥	١٣ + ١١ + ١	=	٢٥
الخامس	٦	٦ + ١٥ + ٧	=	٢٨
السادس	٧	١ + ١٨ + ٢٠ + ٣	=	٤٢
السابع	٨	٦ + ١٠ + ٦	=	٢٢
الثامن	٩	١ + ٨ + ٧	=	١٦
التاسع	١٠	٢ + ٥	=	٧
العاشر	١١	٢ + ٥	=	٧
		مجموع التكرارات	=	٢٠٠

الانحراف المعيارى لمجموع المرتبتين

٢٢١ — نحسب الانحراف المعيارى لمجموع من هذا التوزيع التكرارى كما حسبنا الانحراف المعيارى للفرق فى بند ٢١٨ و بيان ذلك كما يأتى :

جدول ٤٠ — حساب الانحراف المعيارى لمجموع مرتبتى س ، ص

مجموع المرتبتين	التكرار ك	الانحراف ح	ك . ح	ك . ح^{٢}
٢	٦	—٤	—٢٤	٩٦
٣	٩	—٣	—٢٧	٨١
٤	٣٨	—٢	—٧٦	١٥٢
٥	٢٥	—١	—٢٥	٢٥
٦	٢٨	٠	٠	٠
٧	٤٢	١	٤٢	٤٢
٨	٢٢	٢	٤٤	٨٨
٩	١٦	٣	٤٨	١٤٤
١٠	٧	٤	٢٨	١١٢
١١	٧	٥	٣٥	١٧٥
	٢٠٠		+٤٥	٩١٥

$$\therefore \text{٢٠٠ ع}_{م}^{٢} = \text{٩١٥} - \frac{(\text{٤٥})^{٢}}{\text{٢٠٠}}$$

$$= \text{٩٠٤٫٨٧٥}$$

$$= \text{م مثلا}$$

وبالمثل نحسب الكميتين ا و ب كما فعلنا فى بند ٢١٩ ، ونعوض فى المعادلة :

$$\text{ر} = \frac{\text{ع}^2_{\text{م}} - \text{ع}^2 - \text{ع}'^2}{2\ \text{ع}\ .\ \text{ع}'} \quad \text{... بند ٢١٠ (٨) ،}$$

$$= \frac{\text{م} - \text{ا} - \text{ب}}{2\sqrt{\text{ا ب}}}$$

$$= \frac{\text{٩٠٤٫٨٧٥} - \text{٦١٣٫٧٣٥}}{2\sqrt{\text{٩٠٦٢٥٫٩٢٤٩}}}$$

$$= \frac{\text{٢٩١٫١٤٠}}{\text{٦٠٢٫٠٨٢٨}}$$

$$= \text{٠٫٤٨٤}$$

وهى نفس النتيجة التى حصلنا عليها من قبل . ولكن هذه الطريقة فى الغالب تكون أطول من الطريقة المتقدمة ، إذ من الواضح أن الجدول التكرارى لمجاميع مرتبتى س و ص يكون أطول من الجدول التكرارى للفرق بينهما ؛ فقد رأينا هنا أن مجموع المرتبتين يتغير من ٢ إلى ١١ ، فى حين أن الفرق بينهما يتغير من — ٣ إلى + ٣ .

٢٢٢ — فى الطريقتين الأخيرتين لحساب معامل الارتباط ، تكلمنا فقط عن مرتبتى س و ص والفرق بينهما أو مجموعهما ؛ ولم نتكلم أبداً عن القيم الفعلية التى تأخذها س أو ص . وما كان ينبغى لنا أن نهتم بهذه القيم هنا ؛ لأن مجموع قيمتى س و ص ، أو الفرق بينهما ، عدد ليس له معنى ، إذ لا معنى لمجموع عمر رجل (مقدراً بالسنين) وعدد أطفاله ؛ ولا معنى للفرق بينهما أيضاً . وقد تخلصنا من هذه الصعوبة بأن قسمنا القيم إلى مراتب ، يمكن طرحها أو جمعها ، وحصلنا بسهولة على نفس النتيجة التى حصلنا عليها باستخدام القيم نفسها . فهل من الممكن تعميم هذه الفكرة واستخدامها فى كل المسائل ؟

إن أمكن هذا فلا شك أننا نختصر كثيراً من العمل الحسابى الذى يستلزمه حساب معامل الارتباط من القيم نفسها .

يجب أن نلاحظ هنا أن مراتب س و ص منتظمة حيث كانت الفئات السينية وكذلك الصادية متساوية فى الفترة . بمعنى أن الفرق بين المرتبة الأولى والثانية مثلاً يساوى الفرق بين الثالثة أو الرابعة ، أو أى مرتبتين متتاليتين . وهو يساوى ٥ سنين فى حالة س و ١ طفل فى حالة ص . ويجب أن نميز بين هذه الحالة والحالات الأخرى التى تكون فيها الفروق بين المراتب غير متساوية ، التى نسميها حينئذ « تراتيب » بدل مراتب .

ومن الواضح طبعاً أنه من الممكن حساب معامل الارتباط باستخدام مراتب القيم ، إذا كانت هذه القيم موزعة فى فئات ذات فترات متساوية ، سواء بالطريقة الأساسية التى شرحناها أولا ، أو بطريقة المجموعات القطرية [١] .

معامل الارتباط بين تراتيب س و ص

٢٢٣ — أما فى حالة عدم تساوى الفترات بين القيم المتتالية . فنستخدم معامل الارتباط بين التراتيب الذى وضعه اسبيرمان (صحيفة ٢٢٦) حيث نستعمل تراتيب قيم س وتراتيب قيم ص بدل القيم نفسها .

وهذه المعادلة هى كما رأينا فى بند ٢٠٩ :

$$ر = ١ - \frac{٦ \text{ مج } ف^{٢}}{ن(ن^{٢} - ١)}$$

حيث ر هى معامل الارتباط (وهو هنا تقريبى بالنسبة إلى المعامل ر) ؛

(١) يمكن للطالب أن يثبت ذلك لنفسه بأن يأخذ جدول ٣٥ ويكتب ١ ، ٢ ، ٣ ، ... لقيم س بدلا من ٢٢٫٥ ، ٢٧٫٥ ، ٣٢٫٥ ، ... ولن يؤثر ذلك فى النتيجة النهائية لمعامل الارتباط لأن هذا يكون بمثابة اختيار وحدات جديدة للظاهرة س كل منها تساوى ٥ سنين وقياس العمر من ١٧٫٥ بدلا من صفر .

و ن هى عدد الحالات ؛

و ى هى الفرق بين ترتيب س و ص ، كل فى مجموعتها .

كيفية حساب ر عمليـاً

٢٢٤ — لنحسب هـذا المعامل ر بين تراتيب س وص فى المثـال الوارد فى بنـد ٢١٠ (صفحة ٢٢٨) ، حيث س تدل على نسـبة البطالة بين العمال فى انجلترا ، و ص تدل على قيمة الصادرات ، فى السنين ١٩٢١ — ١٩٣٠

نرتب أولا قيم س وقيم ص تصاعدياً (أو تنازلياً) لنعرف ترتيب كل قيمـة بين أخواتها . ومن الضرورى أن يكون نظام الترتيب واحداً فى مجموعتى س و ص . إما تصاعدى فى الاثنتين معاً أو تنازلى ، حتى نمـيز بين الارتباط الطردى والارتباط العكسى .

القيم السينية مرتبة تصاعدياً ، وتراتيبها هى كما يأتى :

٩٫٧ ، ١٠٫٣ ، ١٠٫٤ ، ١٠٫٨ ، ١١٫٣ ، ١١٫٧ ، ١٢٫٥ ، ١٤٫٣ ، ١٦٫١ ، ١٧٫٠٠ ؛

١ ، ٢ ، ٣ ، ٤ ، ٥ ، ٦ ، ٧ ، ٨ ، ٩ ، ١٠ ؛

وكذلك القيم الصادية وتراتيبها ، تصاعدياً أيضاً ، هى :

٥٧١ ، ٦٥٣ ، ٧٠٣ ، ٧٠٩ ، ٧٢٠ ، ٧٢٠ ، ٧٢٤ ، ٧٦٧ ، ٧٧٣ ، ٨٠١ ؛

١ ، ٢ ، ٣ ، ٤ ، ٥½ ، ٥½ ، ٧ ، ٨ ، ٩ ، ١٠ .

يلاحظ هنا كيف عالجنـا القيمة المتـكررة ٧٢٠ ، فأعطينا كل مفردة منهما نفس الترتيب وهو ٥½ حيث إن هذه القيمة المتـكررة تشـغل بالفعل المـكانين الخامس والسادس فى السلسلة ، فتعطى كلا منهما رتبة تساوى المتوسط بين ٥ ، ٦ .

نضع هـذه التراتيب بدل القيم نفسها الواردة فى جدول ٢٩ (صفحة ٢٢٩) ونحسب ر . وها هى تلك الخطوات موضحة فى الجدول الآتى :

جدول ٤١ — حساب معامل الارتباط ر

بين تراتيب نسبة البطالة وقيمة الصادرات فى انجلترا

السنة	نسبة البطالة س	قيمة الصادرات ص	تراتيب س	تراتيب ص	الفرق بين الترتيبين ف	ف٢
١٩٢١	١٧ر٠٠	٧٠٣	١٠	٣	٧	٤٩
٢٢	١٤ر٣	٧٢٠	٨	$5\frac{1}{2}$	$2\frac{1}{2}$	٦ر٢٥
٢٣	١١ر٧	٧٦٧	٦	٨	— ٢	٤
٢٤	١٠ر٣	٨٠١	٢	١٠	— ٨	٦٤
٢٥	١١ر٣	٧٧٣	٥	٩	— ٤	١٦
٢٦	١٢ر٥	٦٥٣	٧	٢	٥	٢٥
٢٧	٩ر٧	٧٠٩	١	٤	— ٣	٩
٢٨	١٠ر٨	٧٢٤	٤	٧	— ٣	٩
٢٩	١٠ر٤	٧٢٠	٣	$5\frac{1}{2}$	$-2\frac{1}{2}$	٦ر٢٥
٣٠	١٦ر١	٥٧١	٩	١	٨	٦٤
						٢٥٢ر٥

وبالتعويض فى المعادلة مجـ ف٢ = ٢٥٢ر٥ ، ن = ١٠

$$\therefore \text{ر} = ١ - \frac{٦ \times ٢٥٢ر٥}{١٠ \, (١٠٠ - ١)}$$

$$= ١ - \frac{١٥١٥}{٩٩٠}$$

= — ٥٣٠ر،

بينما ر = — ٥٨٤ر.

ولا شك أن فى هذه الطريقة تسهيلاً كبيراً واختصاراً للعمل الحسابى ، حيث نستعيض عن القيم الأصلية ذات الأرقام الكثيرة ، بتراتيبها ذات الأرقام المختصرة ، علاوة على أننا غير محتاجين إلى حساب الوسطين الحسابيين أو الانحرافين المياريين . وقد رأينا أن تأثير هذا الاختصار على النتيجة النهائية ليس كبيراً جداً ؛ وربما لايزيد عن ١٠ أو ١٥ ٪ ، كما هو واضح من مقارنة قيمتى ر ، ر .

معامل الائتلاف

٢٢٥ — يوجد طريقة أخرى أكثر اختصاراً من السابقة لقياس العلاقة أو الارتباط بين ظاهرتين مقيستين ، وذلك باستخدام معادلة وضعها المستر ج . أ . يول . وهذه المعادلة تعطى مقياس العلاقة الذى نسميه (١) **معامل الائتلاف** ونرمز له بالحرف ل وهو :

$$\text{ل} = \frac{\sqrt{\text{ا ء}} - \sqrt{\text{ب حـ}}}{\sqrt{\text{ا ء}} + \sqrt{\text{ب حـ}}}$$

حيث ا = عدد الحالات التى فيها كل من س و ص فوق المتوسط

« ء = « « « « « س و ص تحت «

« ب = « « « « « قيمة س تحت و ص فوق «

« حـ = « « « « « س فوق و ص تحت «

(١) هذا المعامل يسمى G. U. Yule's "Coefficient of Colligation" ويرمز له بالحرف الأغريقى ω «أوميجا» — أنظر مقالته فى :

(Journal of the Royal Statistical Society, (Vol. 75 (1911-1912), pp. 579-652)

وانظر أيضا انتقادات K. Pearson and Heron

فى مجلة Biometrika, Vol. 9 (1913) pp. 159-315 على هذا المعامل ، وعلى معامل الاقتران الذى سيأتى ذكره فى بند ٢٢٨ .

ولحساب هذا المعامل نحسب المتوسط لقيم س وقيم ص ثم نوزع قيم س وص في جدول مزدوج من أربع خانات فقط مثل الجـدول المبين ، ثم نحسب عدد الحالات التي تقع في كل خانة . ففي المسألة التي بأيدينا مثلا ، نعلم أن الوسط الحسابي لقيم س (أي نسبة البطالة) هو ١٢٫٤١ ، والوسط الحسابي لقيم ص هو ١٫٧١٤ .

قيم س

قيم ص	فوق المتوسط	تحت المتوسط
فوق المتوسط	ا	ب
تحت المتوسط	حـ	ى

∴ عدد الحالات التي فيها س فوق المتوسط يساوي ٤ ، كما يتضح من جدول ٤١ .

من هـــذه الحالات واحدة فيها ص فوق متوسطها ، وثلاث فيها ص تحت متوسطها .

وعدد الحالات التي فيها س أقل من المتوسط يساوي ٦ طبعاً .

ومن هذه الحالات نجـد خمساً فيها ص فوق المتوسـط ، وواحدة فيها ص تحت المتوسط .

وعلى ذلك يكون توزيع الحالات العشر كما يأتي :

قيم س

قيم ص	+	—
+	١	٥
—	٣	١

حيث تدل العلامة + على فوق المتوسط ، والعلامة — على تحت المتوسط

$$\therefore \quad ل = \frac{\sqrt{١ \times ١} - \sqrt{٥ \times ٣}}{\sqrt{١ \times ١} + \sqrt{٥ \times ٣}}$$

$$= \frac{١ - ٣ر٨٧٣٠}{١ + ٣ر٨٧٣٠}$$

$$= - ر٥٩٠ \text{ ؛}$$

بينما ر = — ر٥٣٠ ،

و ρ = — ر٥٨٤ .

ونلاحظ أن هذه المقاييس الثلاثة كلها متقاربة ، وكلها متحدة فى الإشارة السالبة ، دلالة على أن العلاقة عكسية .

معامل الائتلاف من جدول الارتباط

٢٢٦ — ويمكن تطبيق هذه الطريقة الأخيرة بسهولة فى حالة جدول الارتباط ، حيث يكون عدد الحالات كبيراً . وذلك بأن نرسم خطاً رأسياً فى الجدول يقسم الفئات السينية إلى جزءين : أقل من الوسط الحسابى وأكبر منه . وكذلك نرسم خطاً أفقياً يقسم الفئات الصادية إلى قسمين أيضاً : تحت المتوسط وفوقه .

وهذان المحوران يقسمان الجدول إلى أربعة أقسام ؛ ومجموع تكرارات الخانات الموجودة فى كل قسم تعطينا الأعداد ا ، ب ، حـ ، ء المطلوبة لحساب معامل الائتلاف .

ففى جدول ٣٢ مثلا ، نعلم أن $\bar{س}$ = ١٧٥ر٣٣ سنة ، و $\bar{ص}$ = ٠٩ر٢ طفلا . وعلى ذلك فالخط الرأسى يقع بين الفئة السينية الثالثة (وهى التى مركزها ٥ر٣٢ أقل من $\bar{س}$) والرابعة . والخط الأفقى يقع بين الفئة الصادية الثالثة (فيها قيم ص = ٢ أقل من $\bar{ص}$) والرابعة .

مجاميع تكرارات الخانات فى الأقسام الأربعة هى كما يأتى :

		س +	س −
ص	+	٣٩	٣٢
	−	٣٧	٩٢

$$\therefore \quad ل = \frac{\sqrt{٣٩ \times ٩٢} - \sqrt{٣٢ \times ٣٧}}{\sqrt{٣٩ \times ٩٢} + \sqrt{٣٢ \times ٣٧}}$$

$$= \frac{٥٩٫٨٩٩٩ - ٣٤٫٤٠٩٣}{٥٩٫٨٩٩٩ + ٣٤٫٤٠٩٣}$$

= ٢٧٠ر ،

ولكن ر = ٤٨٤ر ، (أنظر صحيفة ٢٤١) .

٢٢٧ — ويلاحظ أن الفرق بين المعاملين ل و ر فى هذه الحالة كبير نوعاً . ولو أن فئات الأعمار كانت أضيق مدى مما هى فى جدول ٣٢ ، لأمكن تقسيم الأعمار فوق المتوسط وتحته ، تقسيما أدق مما فعلنا هنا ، وحصلنا بذلك على قيمة للمعامل ل أدق من ٢٧٠ر . وأقرب إلى ر منها . فنرى هنا أن اتساع مدى الفترة فى فئات الأعمار فى هذا المثال جعلنا نعتبر ٦٠ رجلا (الموجودين فى الفئة السينية ٣٠ وأقل من ٣٥) أعمارهم جميعاً تحت المتوسط لأن مركز الفئة ٥ر٣٢ والمتوسط ١٧٥ر٣٣ ، مع أننا نعلم أن عدداً من هؤلاء لا بد أعمارهم أكبر من ١٧٥ر٣٣ حيث أن الفئة تمتد إلى ٣٥ ، ولو كان المدى أقل من ٥ سنوات لأمكننا تقسيم هؤلاء الستين بدقة أكثر .

ويلاحظ أن المعاملين متحدان فى الإشارة الموجبة . حيث يدل كل منهما على أن العلاقة طردية بين السن وعدد الأطفال ، كما هو منتظر .

الاقتران بين الصفات غير المقيسة

٢٢٨ — تكلمنا فى البنود السابقة عن طرق قياس العلاقة بين الظواهر المتغيرة التى يمكن قياسها رقمياً . ويبقى إذن أن نبحث فى طرق قياس العلاقة بين

الصفات التى لا يمكن قياسها . وقد سبق أن ذكرنا أننا نسمى هذه العلاقة « الاقتران » بين الصفات . ومقياس هذه العلاقة نسميه « معامل الاقتران » .

لنفرض أننا نبحث فى العلاقة بين جنسية التاجر (فى القاهرة مثلا) ونوع العمل الذى يقوم به . ولنفرض أن البيانات التى لدينا تقسم التجار من حيث الجنسية إلى صنفين فقط : مصريون وأجانب ؛ وتقسم الأعمال التجارية نوعين أيضاً : أعمال مالية صرفة ، وأعمال تجارية للبيع والشراء . ففى القاهرة مثلا نجد عدد المشتغلين بالأعمال التجارية فى سنة ١٩٢٧ يساوى ٩٣٧٠٠ . من هؤلاء ٨٢٢٠٠ مصريون و ١١٥٠٠ أجانب . ومن المصريين ٦٧٠٠ يشتغلون بالأعمال المالية و ٧٥٥٠٠ بالمتاجرة . ومن الأجانب ٢٠٠٠ يشتغلون بالأعمال المالية و ٩٥٠٠ بأعمال المتاجرة . فهل هناك علاقة بين جنسية الشخص ونوع العمل الذى يقوم به ، وما مقياس هذه العلاقة ؟

هنا نستخدم « معامل الاقتران » الذى وضعه ج . أ . يول[١] لقياس العلاقة بين الصفات التى لا تقاس . ولذلك نرتب الأرقام المعطاة فى « جدول الاقتران » كما يأتى :

جدول الاقتران بين الجنسية ونوع العمل

		الجنسية	
		مصريون	أجانب
العمل	مالى	٦٧٠٠	٢٠٠٠
	متاجرة	٧٥٥٠٠	٩٥٠٠

(١) أنظر كتاب : G. U. Yule "*Introduction to the Theory of Statistics*" (1937)
وانظر أيضا مقالته فى مجلة *Phil. Trans. Roy Soc.*, Series A, Vol. 194 (1900), p. 257.

$$\text{معامل الاقتران} = \frac{٦٧٠٠ \times ٩٥٠٠ - ٢٠٠٠ \times ٧٥٥٠٠}{٦٧٠٠ \times ٩٥٠٠ + ٢٠٠٠ \times ٧٥٥٠٠}$$
$$= ٠٫٤٠٧$$

ولا ضرورة هنا للإشارة الجبرية ؛ لأن المعامل إذا دل على اقتران بين الجنسية المصرية وأعمال المتاجرة ، فهو يدل فى الوقت نفسه على « تنافر » بين الجنسية المصرية والأعمال المالية . والمفهوم من هذا المعامل أن هناك علاقة أو ارتباطاً بين صفة الجنسية المصرية وصفة العمل التجارى ، ومثلها بين صفة الجنسية الأجنبية وصفة العمل المالى .

وعلى العموم إذا كان توزيع الأعداد فى جدول الاقتران هو :

ا	ب
حـ	ء

يكون معامل الاقتران

$$ن = \frac{ا ء - ب حـ}{ا ء + ب حـ} ، \ldots \qquad (١) ؛$$

مع صرف النظر عن الإشارة الجبرية . وبدهى أن هذا المعامل دائماً أقل من ١ ، وإذا كانت ن = صفراً أو قريبة منه كان ذلك دليلا على عدم وجود الاقتران أو ضعفه .

٢٢٩ — إذا كانت إحدى الظاهرتين اللتين نبحث العلاقة بينهما ، معامل التوافق
أو كلتيهما ، تنقسم إلى أكثر من نوعين ، فإن معامل الاقتران لا يساعدنا فى هذه الحالة . وعندئذ نستخدم « معامل التوافق » الذى وضعه كارل بيرسون[١] لقياس العلاقة بين الصفات غير المقيسة ، أو بين صفات بعضها يقاس بالأرقام وبعضها لا يقاس .

---

(١) اسمه بالإنجليزية Karl Pearson's "Coefficient of Contingency" ؛ أنظر كتاب Whittaker and Robinson: *Calculus of Observations*, (1929), p. 338.
وهناك يرمز لمعامل التوافق بالحرف الأغريقى Φ ، الذى نستبدل به الحرف العربى قَ .

ولتعريف هذا المعامل نأخذ المثال الآتى :

فى سنة ١٩٣٤ تقدم إلى امتحان شهادة الدراسة الثانوية (قسم أول) طلبة من ١٤٤ مدرسة . وكانت هذه المدارس مقسمة (من حيث الإدارة الفنية) إلى ثلاث أنواع : س (مدارس أميرية) ، و ص (مدارس خاضعة لتفتيش وزارة المعارف) ، وع (مدارس غير خاضعة لهذا التفتيش) ؛ ثم قسمت هذه المدارس حسب نسبة النجاح إلى خمس رتب : ا ، ب ، حـ ، ى ، هـ ، فيها نسب النجاح كما يأتى :

ا : ٨٠ وأقل من ١٠٠ ٪
ب : ٦٠ « ٨٠ ٪
حـ : ٤٠ « ٦٠ ٪
ى : ٢٠ وأقل من ٤٠ ٪
هـ : ٠ « ٢٠ ٪

وكان توزيع المدارس ال ١٤٤ على هـذه المجموعات المزدوجة كما هو مبين فى « جدول التوافق » الآتى :

جدول ٤٢ — توزيع ١٤٤ مدرسة حسب النوع والرتبة

		رتب المدارس					المجموع
		ا	ب	حـ	ى	هـ	
نوع المدارس	س	١	١٨	٩	١	٠	٢٩
	ص	٢	٠	٢	٢٧	١٦	٤٧
	ع	٢	٢	٨	٢٢	٣٤	٦٨
المجموع		٥	٢٠	١٩	٥٠	٥٠	١٤٤

حيث العدد الموجود فى كل خانة من هذا الجدول يدل على عدد المدارس (أو التكرار) التى تجتمع فيها الصفتان المبينتان .

٢٣٠ — ولحساب معامل التوافق ، نربع تكرار كل خانة ونقسمه على حاصل ضرب المجموعين الرأسى والأفقى للعمود والصف الملتقيين فى هذه الخانة ، كما هو مبين فى الجدول الآتى : حساب معامل التوافق

جدول ٤٣ — حساب معامل التوافق بين نوع المدرسة ورتبتها

		رتب المدارس					المجموع
		ا	ب	ح	ى	ه	
نوع المدارس	س	١ / (٥ × ٢٩)	٣٢٤ / (٢٠ × ٢٩)	٨١ / (١٩ × ٢٩)	١ / (٥٠ × ٢٩)	٠	٫٧١٣١
	ص	٤ / (٥ × ٤٧)	٠	٤ / (١٩ × ٤٧)	٧٢٩ / (٥٠ × ٤٧)	٢٥٦ / (٥٠ × ٤٧)	٫٤٤٠٦
	ع	٤ / (٥ × ٦٨)	٤ / (٢٠ × ٦٨)	٦٤ / (١٩ × ٦٨)	٤٨٤ / (٥٠ × ٦٨)	١١٥٦ / (٥٠ × ٦٨)	٫٥٤٦٥
المجموع		٠٫٣٥٦	٫٥٦١٥	٫٢٠١٠	٫٤٥٣٢	٫٤٤٨٩	١٫٧٠٠٢

فلو رمزنا للمجموع الكلى ١٫٧٠٠٢ بالحرف ح ، يكون معامل التوافق الذى يسميه بيرسون **جذر متوسط مربع التوافق**[١] .

$$\text{ق}' = \sqrt{\text{ح} - ١} \quad . \; . \; . \; . \; . \; . \quad (١)$$

(١) بالانجليزية Root-Mean Square Contingency.

ولكن أغلب الإحصائيين[١] يستخدمون معادلة أخرى لمعامل التوافق وهى :

$$\text{ق} = \sqrt{\frac{\text{م} - ١}{\text{م}}} \quad \ldots \ldots \ldots \quad (٢) ،$$

كل من هذين المعاملين ق و قَ موجب[٢] ؛ ولكنهما غير متساويين ، لأن :

$$\text{ق}^{٢} = \frac{\text{نَ}^{٢}}{\text{نَ}^{٢} + ١} \quad \ldots \ldots \ldots \quad (٣) ،$$

و

$$\text{قَ}^{٢} = \frac{\text{ن}^{٢}}{١ - \text{ن}^{٢}} \quad \ldots \ldots \ldots \quad (٤) ،$$

وكل منهما يمكن استخدامه كمقياس للعلاقة بين الصفات المبينة فى الجدول .

و بناء على ذلك يكون معامل التوافق بين نوع المدارس ورتبتها فى المسألة التى نحن بصددها هو :

$$\text{قَ} = \sqrt{٫٧٠٠٢} ،$$

$$= ٠٫٨٤ ،$$

فى حين أن

$$\text{ق} = \sqrt{\frac{٫٧٠٠٢}{١٫٧٠٠٢}} .$$

$$= ٫٦٤ ،$$

وهذا معناه أن هناك علاقة بين نوع المدرسة ونسبة النجاح بين طلبتها ، وإلا كان كل من المعاملين صفراً ، أو ما يقرب من الصفر .

---

(١) واضح أن المعامل ق يكون دائماً أقل من الواحد . وهو أكثر استعمالا من المعامل قَ .

(٢) أنظر كتاب G. Thompson *"How To Measure Correlation,"* (1924), p. 20
حيث يستعمل الحرف C بدل ق . أو كتاب
Sargant Florence *"Statistical Methods"*, (1929). p. 505
ويلاحظ أن فلورنس يذكر المعادلة «٢» فى شكل مخالف . ولكن يمكن تحويرها بسهولة حتى تأخذ الشكل المذكور أعلاه .

المـــراجع

Bowley, A.L., : *Elements of Statistics*, Chapter VI Part II.

Connor, L.R., : *Statistics in Theory and Practice*, Chapter XV.

Florence, S.P.,: *Statistical Methods*, Chapter IX.

Jones, C., : *First Course in Statistics*, Chapter X.

King, W. : *Statistical Method*, Chapter XVII.

Mills, F.C., : *Statistical Methods*, Chapter X.

Secrist, H., : *Statistical Methods*, Chapter XIII.

Thompson, G.: *How to Measure Correlation*,

Whittaker and Robinson : *Calculus of Observations*, Chapter XII.

الباب التاسع

الارتبـــــاط

خطوط الانحدار المستقيمة والمنحنية

قائدة دراسة الارتباط هى فى إمكان تقدير إحدى الظاهرتين إذا عرفت الأخرى

٢٣١ — تكلمنا فى الباب السابق عن الارتباط ومعناه وكيفية قياسه فى الحالات المختلفة . ومقاييس الارتباط التى نحصل عليها بهذه الطرق تعرفنا درجة العلاقة بين الكميتين أو الظاهرتين المتغيرتين . ولكن هذه المعرفة لا تفيدنا كثيراً فى دراسة هاتين الظاهرتين إذا اقتصرت فائدتها على إثبات وجود هذه العلاقة أو عدم وجودها ، وقياسها إن وجدت . لأن المفهوم عادة من وجود علاقة أو ارتباط بين متغيرين أننا إذا علمنا قيمة أحدهما فى حالة ما ، أمكننا — بناء على وجود هذه العلاقة أو الارتباط — أن نقدر ، ولو بالتقريب ، قيمة المتغير الثانى ؛ وأن تقديرنا هذا يكون أقرب إلى الحقيقة كلما كان الارتباط شديداً . والآن نبحث فى هذه الناحية من موضوع الارتباط .

ازدياد عدد الأطفال مع العمر

٢٣٢ — رأينا فى المثال العملى الذى أخذناه فى بند ٢١٥ (جدول ٣٢ صحيفة ٢٣٥) أن هناك علاقة طردية بين عمر الرجل وعدد أطفاله . أى أن عدد الأطفال يزيد ، على وجه العموم ، بزيادة العمر . ومعنى ذلك أن رجلا عمره ٣٠ سنة مثلا يكون عدد أطفاله فى العادة أكبر من عدد أطفال رجل عمره ٢٥ سنة فقط ؛ وأن رجلا عمره ٣٥ سنة يكون عدد أطفاله ، فى العادة ، أكبر من عدد أطفال رجل عمره ٣٠ سنة فقط ، وهكذا . على أن هذا لا يمنع طبعاً

من أن رجلا معيناً عمره ٢٥ سـنة مثلا عدد أطفاله ، لسبب ما ، أكبر من عدد أطفال رجل آخر عمره ٣٠ سنة . ولـكن هذا إن تحقق فلن يكون إلا نادراً .

إذن يكون متوسط عدد الأطفال عنـد مجموعة الرجال الذين سنهم بين ٢٥ و٣٠ سنة مثلا ، أكبر من متوسط عدد الأطفال عند مجموعة الرجال الذين تنحصر أعمارهم بين ٢٠ و ٢٥ سـنة فقط . وهكذا فى فئات الأعمار المتتـالية ، التى نراها فى الجدول رقم ٣٢ . و يمكننا[١] بسهولة حساب متوسط عدد الأطفال لـكل مجموعة من هذه المجموعات الخمسة وهى :

مجموعة الرجال الذين أعمارهم تساوى ٢٢٫٥ سنة ، متوسط عدد أطفالهم = ٠٫٨٢

» » » » ٢٧٫٥ » » » » = ١٫٣٨

» » » » ٣٢٫٥ » » » » = ٢٫٠٧

» » » » ٣٧٫٥ » » » » = ٢٫٦٢

» » » » ٤٢٫٥ » » » » = ٣٫١٢

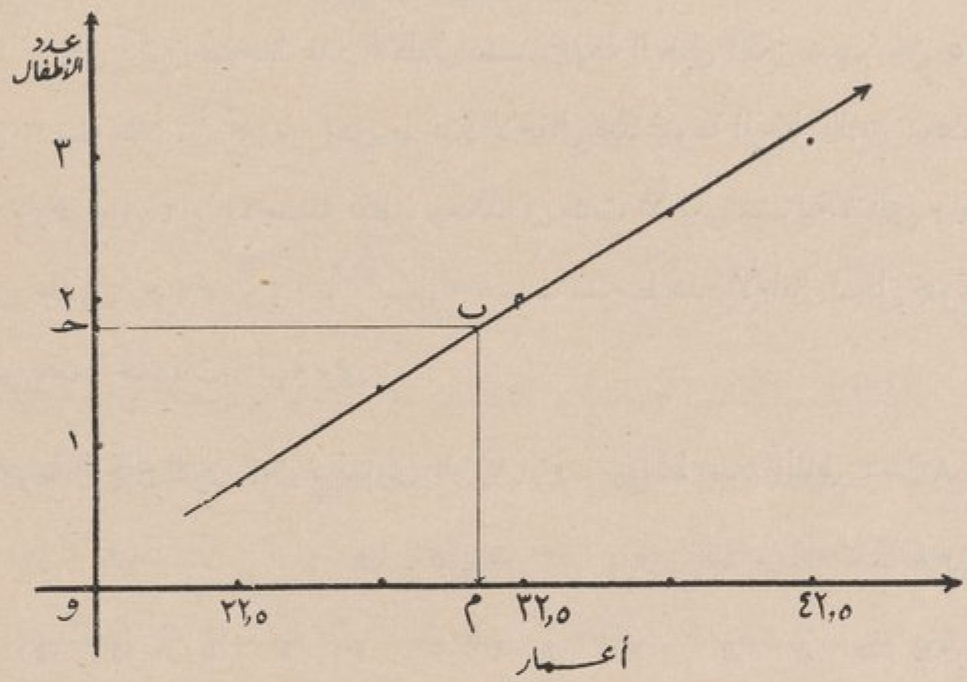
ونرى هنـا بوضوح كيف يزداد متوسط عدد الأطفال بالتـدريج مع زيادة عمر الرجل .

٢٣٣ — لنرسم محورين متعامدين ، ونقيس على المحور الأفقى الأعمار مثلا ، ونقيس عدد الأطفـال على المحور الرأسى . ثم نرصد النقط التى إحداثياتها الأفقية رسم خط بيانى

(١) كل عمود فى جدول ٣٢ عبارة عن توزيع تـكرارى لعـدد الأطفال المبينة فئـاته فى العمود الأيمن من الجـدول . ففى كل خانة فى أى عمود نضرب تكرارها فى عدد الأطفال المبين أمام هـــذه الخانة ، ونجمع هذه الحواصل ونقسم على المجموع الـكلى للعمود ، فنحصل على متوسط عدد الأطفال لهذا العمود : مثلا العمود الثالث نجد المتوسط :

= (٨ × ٠ + ٢٥ × ١ + ١٣ × ٢ + ٦ × ٣ + ١ × ٤) ÷ ٥٣ = ١٫٣٨ طفلا .

هى ٢٢,٥ ، ٢٧,٥ ، ٣٢,٥ ، ٣٧,٥ ، ٤٢,٥ ؛ وإحداثيـاتها الرأسـية هى على الترتيب ٨٢, ، ١,٣٨ ، ٢,٠٧ ، ٢,٦٢ ، ٣,١٢ .



(شكل ٦١) متوسط عدد الأطفال والعمر

ثم نصل بين هذه النقط بخط ممهد على قدر الإمكان ؛ ويظهر من شكل ٦١ أن الخط البيانى فى هذه الحالة مستقيم تقريباً .

هذا الخط إذن يبين العلاقة بين عمر الرجل ومتوسط عدد ما عنـده من الأطفـال . وهو وإن كان مرسـوماً من واقع خمس نقط معينة فقط ، ألا وهى النقط المقـابلة للأعمار ٢٢,٥ ، ٢٧,٥ ، ٠٠ ، ٤٢,٥ ، فهو يمكننا من معرفة ماذا يكون متوسط عدد الأطفـال عنـدما يكون عمر الرجل ٣١ أو ٣٩ سنة مثلا ، أو أى عمر آخر . وذلك بأن نحدد على المحور الأفقى بعداً ، مثل و م يمثل ٣١ سنة (أو ٣٩ أو أى عمر نريده) ثم نقيم من م عموداً على المحور الأفقى فيقابل الخط البيانى فى نقطة مثل ب . والإحداثى الرأسى لهذه النقطة ، وهو م ب = و حـ فى الشكل ، هو متوسط عدد الأطفال عندما يكون العمر يساوى ٣١ سنة (أو ٣٩ أو إلخ) .

خط انحدار عدد الأطفال على العمر

٢٣٤ — هذا الخط البيانى نسميه(١). **خط انحدار عدد الأطفال على العمر**. وهو كما رأينا يصور العلاقة بين العمر ومتوسط عدد الأطفال، وبواسطته يمكن تقدير عدد الأطفال عند أى رجل إذا علم عمر هذا الرجل. وهذا التقدير يعطى متوسط ما عند أمثال هذا الرجل من الأطفال؛ وهو تقدير يقرب من الحقيقة كلما كان الارتباط شديداً بين العمر وعدد الأطفال.

فائدة خط الانحدار فى دراسة الارتباط

٢٣٥ — خط الانحدار إذن يحقق لنا الفائدة التى نرجوها من دراسة العلاقة بين كميتين متغيرتين؛ فهو يصور لنا هذه العلاقة فى شكل هندسى منظور، نرى فيه كيفِ تميل إحدى الكميتين إلى متابعة الأخرى فى تغيرها — إما طردياً وإما عكسياً. وحبذا لو أمكننا أيضاً تصوير هذه العلاقة أو الارتباط فى صورة جبرية أو تحليلية، إذ أن هذا يكون بلا شك أوفى وأتم، من الوجهتين النظرية والعملية.

وهذا الأمر ميسور لنا، حيث قد علمنا (فى الباب الرابع) أن الأشكال البيانية يمكن الدلالة عليها بمعادلات جبرية. فلنبحث إذن فى استنباط معادلة خط الانحدار، مستقيماً كان أو منحنياً؛ وإذا حصلنا عليها فقد حصلنا على الصورة الجبرية أو التحليلية المطلوبة للعلاقة بين المتغيرين تحت البحث.

رسم شكل الانتشار

٢٣٦ — المقصود من خط الانحدار هو كما قلنا تصوير العلاقة بين المتغيرين فى صورة جبرية تحليلية. لنفرض أن قيم المتغيرين هى كالمعتاد:

---

(١) يسمى بالانجليزية Regression Line of the Number of Children on the age of the father وأول من وضع هذه التسمية هو فرانسيس جولتون Francis Galton وترجمتها بالعربية «ارتداد». وسيظهر فيما يلى السبب الذى جعله يستخدم هذا المعنى، والسبب الذى جعلنا نفضل كلمة «انحدار». على أن المعنى الذى تؤديه كلمة ارتداد أو Regression يقتصر على ناحية واحدة فقط للفكرة التى يصورها خط الانحدار.

$\text{س}_1 \text{ ، } \text{س}_2 \text{ ، } \text{س}_3 \text{ ، } \ldots\ldots \text{ ، } \text{س}_{\text{ن}}$

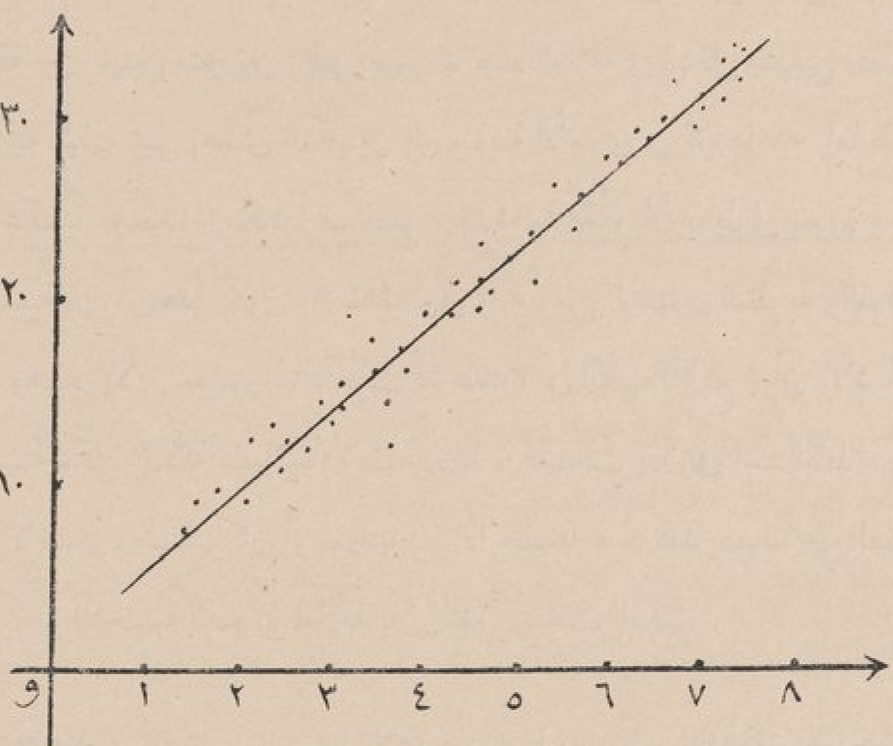
و $\text{ص}_1 \text{ ، } \text{ص}_2 \text{ ، } \text{ص}_3 \text{ ، } \ldots\ldots \text{ ، } \text{ص}_{\text{ن}}$ ؛

وأن الوسطين الحسابيين هما $\bar{\text{س}}$ و $\bar{\text{ص}}$ ، والانحرافين المعياريين هما ع و عَ .

∴ ن $\bar{\text{س}}$ = مج س ، ن $\bar{\text{ص}}$ = مج ص ؛

و ن $\text{ع}_{\text{س}}^2$ = مج $(\text{س} - \bar{\text{س}})^2$ ، ن $\text{ع}_{\text{ص}}^2$ = مج $(\text{ص} - \bar{\text{ص}})^2$.

نرسم محورين متعامدين ونرصد فى الشكل النقط التى إحداثياتها الأفقية تساوى



(شكل ٦٢)

شكل انتشار نقط على خط مستقيم

قيم س ، وإحداثياتها الرأسية تساوى قيم ص المناظرة لها ، أى النقط التى إحداثياتها هى : $(\text{س}_1 \text{ ، } \text{ص}_1)$ ، $(\text{س}_2 \text{ ، } \text{ص}_2)$ ، $(\text{س}_3 \text{ ، } \text{ص}_3)$ ، . . . . . ، $(\text{س}_{\text{ن}} \text{ ، } \text{ص}_{\text{ن}})$.

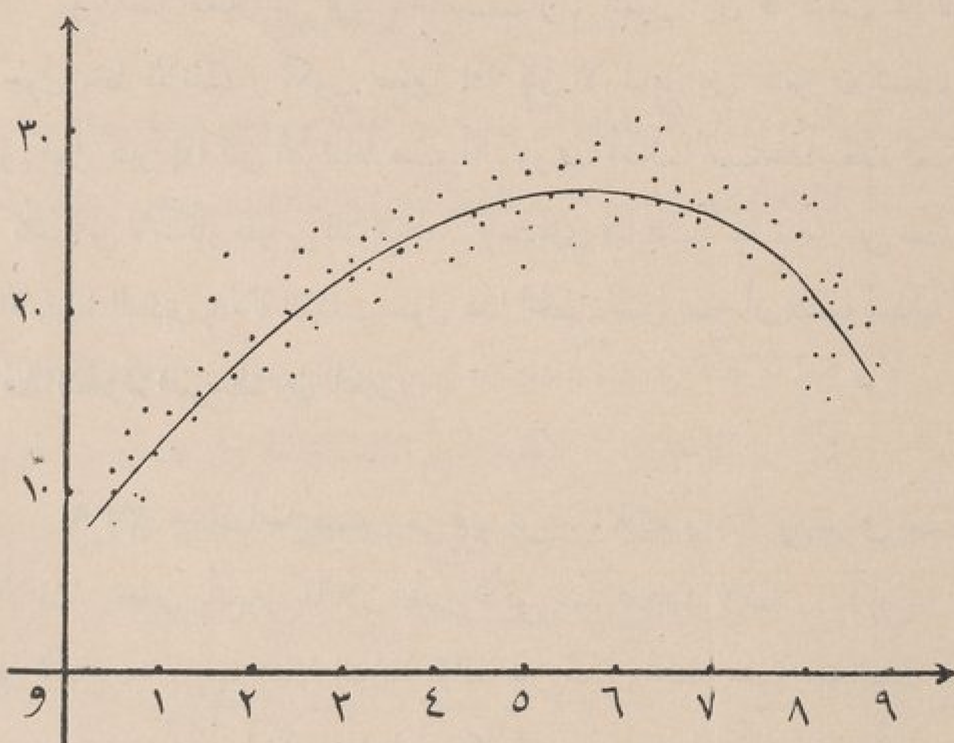
يتكون من هذه النقط البيانيـــة **شكل الانتشار** [1] كالذى نراه فى (شكل ٦٢) حيث نجد النقط منتشرة فى الشكل ومبعثرة نوعاً . و بطبيعة الحال

(١) يسمى بالإنجليزية : Scatter Diagram

إذا كانت هناك علاقة تربط المتغيرين س و ص ، فسيكون أثرها أن تنتشر هذه النقط بشكل منتظم يسير على قاعدة معينة . أما إذا كانت النقط فى هذا الشكل مبعثرة حيثما اتفق و بدون أى نظام ملحوظ ، فهذا دليل على أن العلاقة بين المتغيرين معدومة أو ضعيفة جداً . والخط الذى تنتشر النقط حوله بانتظام نسميه **خط الانتشار**

خط الانتشار إما مستقيم أو غير مستقيم

٢٣٧ — فى بعض الأحيان نجد أن النقط فى شكل الانتشار تنتظم فى خط مستقيم ، أو ما يقرب من خط مستقيم كما نرى فى شكل ٦٢ ؛ وأحياناً تنتشر هذه النقط على خط غير مستقيم ، ولكنه خط منتظم . وهو إما منحن ذو نهاية واحدة



(شكل ٦٣)

شكل انتشار نقط على خط منحن من الدرجة الثانية

(عظمى أو صغرى) ، كما نرى فى (شكل ٦٣) ؛ وتكون إذ ذاك معادلته من الدرجة الثانية كما رأينا فى بند ٧٨ من الباب الرابع . و إما أن يكون منحنياً ذا نهايتين أو أكثر ، فتكون معادلته من الدرجة الثالثة أو من درجة أعلى .

تشـتت النقط حـول خط الانتشار

٢٣٨ — ولا يلزم أن تكون جميع النقط البيانية فى شكل الانتشار واقعة على خط الانتشار أو ملاصقة له ، إذ يصح أن تشذ نقطة معينة ، لسبب ما ، وتقع بعيدة عن هذا الخط . على أنه إذا كان عدد كبير من النقط واقعاً على هذا الخط أو ملاصقاً له ، وكانت النقط الباقية على مقربة منه ، كان ذلك دليلا واضحاً على شدة الارتباط بين المتغيرين . وبالعكس إذا كان كثيراً ما تشذ النقط وتبعد عن هذا الخط ، ولم يكن هناك خط آخر أقرب إلى النقط من هذا ، دل ذلك على ضعف الارتباط بين المتغيرين أو عدم وجوده .

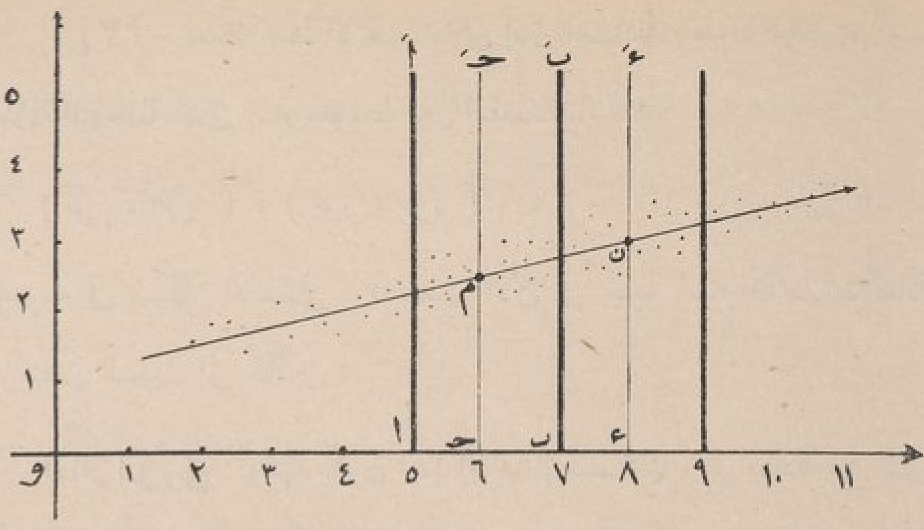
ويمكننا التعبير عن هذا باختصـــار ، فنقول : إن « تشتت » النقط حول خط الانتشار يكون صغيراً إذا كان الارتباط بين المتغيرين شديداً ، ويكون كبيراً إذا كان الارتباط ضعيفاً . وربما أمكننا استخدام هذه الفكرة كأساس لابتكار مقياس للارتباط . وسنرى فيما بعد كيف نعبر عن معامل الارتباط العادى بدلالة التشتت حول هذا الخط ، الذى يصح أن نسميه مبدئياً [١] **خط العلاقة المتوسطة** بين المتغيرين .

خط الانتشار هو نفس خط الانحدار

٢٣٩ — لنأخذ قيمتين من قيم س ، مثلا ٥ و ٧ ؛ ونرسم فى شكل الانتشار خطين رأسيين يقابلان المحور الأفقى عند ٥ وعند ٧ مثلا ا اَ و ب بَ فى شكل ٦٤ ، جميع النقط البيانية الواقعة بين هذين الخطين لها إحداثيات أفقية بين ٥ و ٧ ، أى أنها تكون فئة من فئات س ، حدها الأدنى ٥ والأعلى ٧ ، ومركزها ٦ طبعاً . ويمكن اعتبار أن الإحداثيات الأفقية لجميع هذه النقط تساوى ٦ تقريباً ، كالمعتاد فى التوزيعات التكرارية . نرسم خطاً رأسياً يمر بمركز هذه الفئة مثل حـ حـَ ، يقابل خط الانتشار أو خط العلاقة المتوسطة فى نقطة م

---

(١) بالانجليزية Line of Average Relationship.



(شكل ٦٤)

خط الانتشار هو نفس خط الانحدار

مثلا . فمن الواضح أن الإحداثى الرأسى لهذه النقطة م ، وهو = حـ م ، يساوى متوسط الإحداثيات الرأسية لجميع النقط البيانية الواقعة بين الخطين ا اَ و ب بَ .

لأن خط الانتشار يمر دائماً متوسطاً بين النقط التى حوله — وهذه هى وظيفته طبعاً — ولأن الإحداثيات الأفقية لجميع النقط الواقعة بين هذين الخطين ، أصبحت تساوى ٦ ، و بذلك تركزت جميع هذه النقط على الخط الرأسى حـ حَ .

وكذلك إذا أخذنا الفترة ٧ — ٩ لقيم س ، ورسمنا الخط ء ءَ ليقابل خط الانتشار عند ن ، يكون ء ن مساوياً لمتوسط الإحداثيات الرأسية للنقط الواقعة فى الفترة ٧ — ٩ .

وهكذا فكل نقطة من النقط التى يتألف منها خط الانتشار ، يمثل إحداثيها الأفقى قيمة سينية ، و يمثل إحداثيها الرأسى متوسطاً صادياً مناظراً لهذه القيمة السينية ؛ وعلى ذلك يكون هذا الخط هو خط انحدار ص على س ، حسب خواص خط الانحدار المذكورة فى بندى ٢٣٤ و ٢٣٥ .

معادلة خط الانحدار

٢٤٠ — معادلة خط الانحدار هى إذن معادلة خط الانتشار ، أو خط العلاقة المتوسطة الذى يسير متوسطاً بين النقط البيانية :

$$(\text{س}_1 ، \text{ص}_1) ، (\text{س}_2 ، \text{ص}_2) ، \ldots ، (\text{س}_{\text{ن}} ، \text{ص}_{\text{ن}}) .$$

المرسومة فى شكل الانتشار . فلنبحث الآن فى كيفية استنباط هذه المعادلة . هنا ينقسم البحث إلى حالتين :

الحالة الأولى : وفيها يكون خط الانحدار مستقيما ، ذا معادلة من الدرجة الأولى بالنسبة إلى س و ص ، ويسمى الارتباط حينئذ « الارتباط ذا الانحدار المستقيم » أو « الارتباط المستقيم » .

الحالة الثانية : وفيها يكون خط الانحدار غير مستقيم ، وتكون معادلته من الدرجة الثانية ، أو من درجة أعلى . ويسمى الارتباط حينئذ « الارتباط غير المستقيم »(١) .

معادلة خط الانحدار المستقيم

٢٤١ — فى الحالة الأولى ، حالة الارتباط المستقيم(٢) ، تكون معادلة خط انحدار ص على س هى معادلة أفضل خط مستقيم يوافق النقط البيانية المعطاة وهى :

---

(١) الارتباط المستقيم يسمى Linear Correlation أو Normal Correlation وغير المستقيم Non-Linear Correlation

(٢) لإيجاد معادلة هذا المستقيم نفرض أنها (من الدرجة الأولى) هى :

$$\text{ص} = \text{م}\,\text{س} + \text{حـ} \quad \ldots \quad (١)$$

ونوجد قيمتى م وحـ من المعادلتين الآتيتين :

$$\text{مجـ ص} = \text{م} \,.\, \text{مجـ س} + \text{ن حـ} \quad (٢)$$

$$\text{و} \quad \text{مجـ س ص} = \text{م} \,.\, \text{مجـ س}^{٢} + \text{حـ} \,.\, \text{مجـ س} \quad (٣)$$

ولكن $\text{مجـ س} = \text{ن}\,\overline{\text{س}}$ و $\text{مجـ ص} = \text{ن}\,\overline{\text{ص}}$ و $\text{مجـ س}^{٢} = \text{ن ع}^{٢} + \text{ن}\,\overline{\text{س}}^{٢}$ ، بند ١٧٩ (*)

$(\text{س}_1 ، \text{ص}_1) ، (\text{س}_2 ، \text{ص}_2) ، \ldots ، (\text{س}_\text{ن} ، \text{ص}_\text{ن})$.

ولإيجاد هذه المعادلة ، نوفق خطاً مستقيماً لهذه النقط بنفس الطريقة التى شرحناها فى بند ٧٥ (صحيفة ٧٥) . ونجد بسهولة أن معادلة خط انحدار ص على س هى :

$$\text{ص} - \bar{\text{ص}} = \text{ر} \frac{\text{ع}'}{\text{ع}} (\text{س} - \bar{\text{س}}) \quad \ldots \quad (١) ،$$

حيث $\bar{\text{س}}$ = الوسط الحسابى لقيم س المعطاة فى المسألة ،

و $\bar{\text{ص}}$ = « « « ص « « ،

و ر = معامل الارتباط بين المتغيرين س و ص ،

و ع = الانحراف المعيارى السينى ،

و $\text{ع}'$ = « « الصادى .

والكمية $\text{ر} \frac{\text{ع}'}{\text{ع}}$ تسمى[1] **معامل الانحدار للمتغير ص على المتغير س**

وهذا الخط كما قلنا فى بند ٢٣٩ يعبر عن العلاقة بين قيم المتغير س ومتوسطات القيم الصادية التى تقترن بكل منها . أى أن الإحداثى الأفقى لأى نقطة

(*) فينتج من (٢) أن $\bar{\text{ص}} = \text{م} \, \bar{\text{س}} + \text{ح}$ أى أن $\text{ح} = \bar{\text{ص}} - \text{م} \, \bar{\text{س}}$

وينتج من (٣) أن $\text{مج س ص} = \text{م ن} (\text{ع}^2 + \bar{\text{س}}^2) + \text{ح} \, . \, \text{ن} \, \bar{\text{س}}$

$$= \text{ن م ع}^2 + \text{ن م} \, . \, \bar{\text{س}}^2 + \text{ن} \, \bar{\text{س}} \, \bar{\text{ص}} - \text{ن م} \, \bar{\text{س}}^2 \quad \ldots (٤)$$

ولكن $\text{ن} \, . \, \text{ر} \, . \, \text{ع} \, \text{ع}' = \text{مج س ص} - \text{ن} \, \bar{\text{س}} \, \bar{\text{ص}}$ معادلة (٢) بند ٢٠٥

$\therefore \quad = \text{ن م ع}^2$

$\therefore \quad \text{م} = \text{ر} \frac{\text{ع}'}{\text{ع}}$ وبالتالى $\text{ح} = \bar{\text{ص}} - \text{ر} \frac{\text{ع}'}{\text{ع}} \bar{\text{س}}$.

وبالتعويض عن م و ح فى المعادلة المفروضة $\text{ص} = \text{م س} + \text{ح}$ تنتج المعادلة المطلوبة .

(١) بالإنجليزية Coefficient of Regression of Y upon X وتسمى أحياناً نسبة الانحدار أى Regression Ratio of Y on X

على هذا المستقيم ، يمثل قيمة معينة للمتغير س ، والإحداثي الرأسي يمثل متوسط قيم المتغير ص التي تقترن بهذه القيمة السينية الخاصة .

ففي مسألة الارتباط بين أعمار الرجال وأعداد أطفالهم (جدول ٣٢) لو أخذنا س تمثل العمر و ص تمثل عدد الأطفال ، نجد أن معادلة خط انحدار عدد الأطفال على العمر هي :

$$ص - ٢,٠٩ = ٤٨٤, \times \frac{١,٣٥٣٥}{٥,٥٦٠٥} (س - ٣٣,١٧٥) ،$$

أي $$ص = ٠,١١٧٨ س - ١,٨١٨٠ \quad .... \quad (٢) ،$$

حيث س = عمر الرجل

و ص = متوسط عدد الأطفال عند الرجال الذين عمرهم يساوي س .

فلو أخذنا رجلا عمره ٣١ سنة مثلا ، نجد أن عدد أطفاله في المتوسط يكون

$$ص = ٠,١١٧٨ \times ٣١ - ١,٨١٨٠ ،$$

$$= ١,٨٣٣٨ \text{ طفلا .}$$

ويمكن الوصول إلى نفس النتيجة بالرسم ، حيث نرسم المستقيم الذي معادلته ص = ١١٧٨, س — ١,٨١٨ ونأخذ على المحور الأفقي البعد و ا = ٣١ سنة ونرسم ا ب رأسياً ليقابل المستقيم في نقطة ب يكون إحداثيها الرأسي ا ب = و ح = ١,٨٣٣٨ طفلا (شكل ٦١ صحيفة ٢٦٤).

٢٤٢ — أما معادلة خط انحدار س على ص فنحصل عليها بنفس الطريقة بوضع س بدل ص ووضع ص بدل س في كل خطوة من البرهان السابق . وهذه المعادلة هي : خط انحدار س على ص

$$س - \overline{س} = ر \frac{ع}{ع} (ص - \overline{ص}) \quad ... \quad ... \quad ... \quad (١) .$$

هنا ص تمثل أى قيمة معينة للمتغير ص ، فى حين أن س تمثل متوسط القيم السينية التى تقترن بهذه القيمة الصادية ؛ والكمية $ر \frac{ع}{ع}$ هى معامل انحدار س على ص ، أو نسبة انحدار س على ص .

وفى مسألة الأعمار وعدد الأطفال ، نجد معادلة خط انحدار الأعمار على عدد الأطفال هى :

$$\text{س} - ٣٣٫١٧٥ = ٫٤٨٤ \times \frac{٥٫٥٦٠٥}{١٫٣٥٣٥} (\text{ص} - ٢٫٠٩)$$

$$\text{أى} \quad \text{س} = ١٫٩٨٨٤ \, \text{ص} + ٢٩٫٠١٩٢ \quad (١) \, .$$

فلو أردنا مثلا معرفة ماذا يكون متوسط عمر الرجل الذى عنده ٣ من الأطفال ، نعوض فى هذه المعادلة ص = ٣ ،

$$\therefore \quad \text{س} = ١٫٩٨٨٤ \times ٣ + ٢٩٫٠١٩٢$$

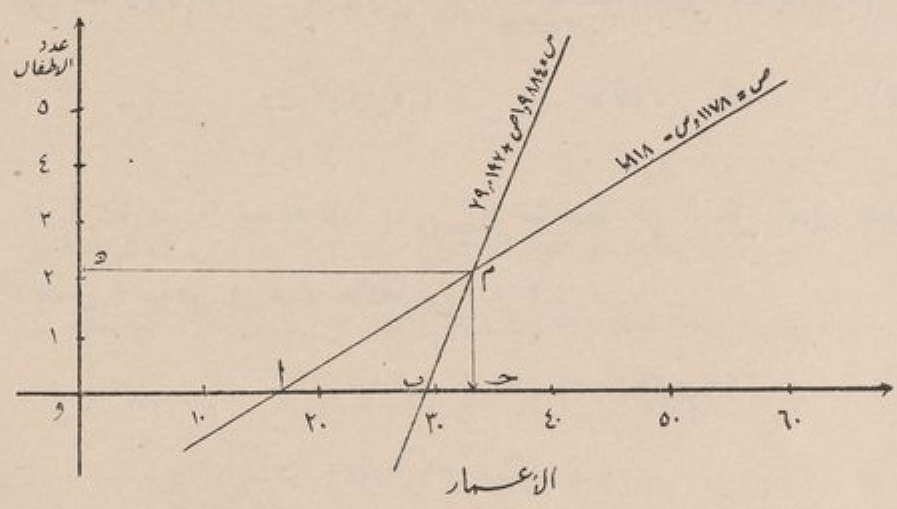
$$= ٣٤٫٩٨٤٤ \, \text{سنة} \, .$$

وهذا معناه أن الرجال الذين عندهم ٣ أطفال يكون متوسط أعمارهم ٣٥ سنة تقريباً .

خواص خطوط الانحدار

٢٤٣ — إذا كان الارتباط بين المتغيرين طردياً ، نجد أن خطى الانحدار يميلان على المحور الأفقى ميلا موجباً ؛ أى أن كلا منهما يصنع زاوية حادة مع محور س ، ظلها يساوى كمية موجبة . أما إذا كان الارتباط عكسياً (ر سالبة) ، فكل منهما يصنع زاوية منفرجة مع محور س ، وظلها يساوى كمية سالبة . وهذا واضح من صورة معادلتى خط الانحدار حيث تظهر الكمية ر مضروبة فى س (أو فى ص) ، حيث تكون إشارتها الجبرية هى إشارة الميل أو ظل الزاوية التى يصنعها المستقيم مع المحور الأفقى (أو الرأسى) .

و إذا رسمنا خطى الانحدار فى شكل واحد من واقع معادلتيهما ، نجد أنهما يتقابلان فى النقطة التى إحداثيها الأفقى يساوى الوسط الحسابى السينى ، و إحداثيها الرأسى يساوى الوسط الحسابى الصادى . والسبب فى ذلك واضح ، إذ نرى أن النقطة ($\overline{\text{س}}$ ، $\overline{\text{ص}}$) تحقق معادلتى خطى الانحدار فى وقت واحد . ففى مسألة الأعمار وعدد الأطفال نجد الخطين يتقابلان فى النقطة م (٣٣٫١٧٥ ، ٢٫٠٩) كما نرى فى شكل ٦٥ .



(شكل ٦٥)
خطا انحدار الأعمار وعدد الأطفال

وهذان الإحداثيان هما كما نعلم الوسطان الحسابيان للأعمار وأعداد الأطفال .

حاصل ضرب معامل الانحدار يساوى ر^2

٢٤٤ — بالنظر إلى معادلتى خطى الانحدار وهما :

$\text{ص} = \text{ر}\,\frac{\text{ع}'}{\text{ع}}\,\text{س} + (\overline{\text{ص}} - \text{ر}\,\frac{\text{ع}'}{\text{ع}}\,\overline{\text{س}})$ ، (وهو خط انحدار ص على س)

$\text{س} = \text{ر}\,\frac{\text{ع}}{\text{ع}'}\,\text{ص} + (\overline{\text{س}} - \text{ر}\,\frac{\text{ع}}{\text{ع}'}\,\overline{\text{ص}})$ ، (« « « س على ص)

نجد أن حاصل ضرب معامل س فى الأولى ، ومعامل ص فى الثانية (أى معاملى الانحدار) يساوى ر^2 ، وهو مربع معامل الارتباط .

∴ فى مسألة الأعمار والأطفال ، حيث المعادلتان هما :

ص = ١١٧٨, س — ١٫٨١٨٠ ،

و   س = ١,٩٨٨٤ ص + ٢٧و٠١٩٢ ،

يكون $ر^{٢}$ = ١١٧٨و٠ × ١و٩٨٨٤ = ٢٣٤٢و٠ تقريباً

ومن ثم ر = ٤٨٤و٠

وفى شكل ٦٥ نرى أن حاصل ضرب هذين المعاملين هو

$$ر^{٢} = \frac{\text{ح م}}{\text{ا ح}} \times \frac{\text{ب ح}}{\text{ح م}} ،$$

$$= \frac{\text{ب ح}}{\text{ا ح}} ،$$

ومن هذه النسبة بين البعدين ب ح و ا ح نحسب قيمة ر .

تشتت النقط حول المستقيم

٢٤٥ — لنبحث[١] الآن فى تشتت النقط المعطاة

($س_{١}$ ، $ص_{١}$) ، ($س_{٢}$ ، $ص_{٢}$) ، . . . ($س_{ن}$ ، $ص_{ن}$) ،

حول خط الانتشار أو خط الانحدار المستقيم الذى معادلته :

ص = م س + ح ،

تلك الفكرة التى نوهنا عنها فى آخر بند ٢٣٨ ؛ ولنبحث فيما إذا كان هناك أى علاقة بين مقدار هذا التشتت ومعامل الارتباط بين المتغيرين س و ص .

قلنا فى الباب السابع ، إن التشتت حول الوسط الحسابى يقاس بالجذر التربيعى لمتوسط مربعات الانحرافات عن هذا الوسط . وكذلك نقول هنا إن تشتت هذه النقط حول الخط ص = م س + ح يقاس بالجذر التربيعى لمتوسط مربعات انحرافات النقط عن الخط .

عند توفيق المستقيم ص = م س + ح ، قلنا (بند ٧٥ صفحة ٧٥) إن انحراف النقطة ا ($س_{١}$ ، $ص_{١}$) مثلا عن هذا المستقيم ، وليكن $ح_{١}$ مثلا، هو

---

(١) يحسن للقارىء المبتدىء أن يترك هذا البند وما يليه فى هذا الباب .

$ع_١ = م س_١ + ح - ص_١$ . . . (١) ؛

وكذلك النقطـــة ب ($س_٢$ ، $ص_٢$) يكون انحرافها عن المستقيم ، وليكن $ع_٢$ ، هو

$ع_٢ = م س_٢ + ح - ص_٢$ (٢) ؛

وهكذا مع جميع النقط .

∴ مجموع هذه الانحرافات

$مج ع = مج (م س + ح - ص)$ ،

$= صفراً$ (٣) .

لأن $مج ص = م مج س + ن ح$ من معـادلات توفيق المسـتقيم

$ص = م س + ح$.

لو ضربنا طرفى المعادلة (١) فى $ع_١$ والمعادلة (٢) فى $ع_٢$ وهكذا

∴ $ع_١^٢ = م . ع_١ س_١ + ح . ع_١ - ع_١ ص_١$ ،

و $ع_٢^٢ = م . ع_٢ س_٢ + ح . ع_٢ - ع_٢ ص_٢$ ،

. . . .

و $ع_ن^٢ = م ع_ن س_ن + ح . ع_ن - ع_ن ص_ن$ ؛ وبالجمع

∴ $مج ع^٢ = م . مج ع س + ح . مج ع - مج ع ص$ ، ... (٤) .

ولكن $مج ع س = مج س (م س + ح - ص)$

$= مج (م س^٢ + ح س - س ص)$

$= م . مج س^٢ + ح . مج س - مج س ص$

$= صفراً$.

لأن $مج س ص = م مج س^٢ + ح مج س$ ، من معادلات التوفيق ،

وكذلك $ح . مج ع = صفراً$.

$$\therefore \quad \text{مح}\,\text{ح}^{٢} = -\,\text{مح}\,\text{ح}\,\text{ص} \text{ ،}$$

و بضرب طرفى معادلة (١) فى $\text{ص}_{١}$ وهكذا ينتج أن

$$\text{مح}\,\text{ح}^{٢} = -\,\text{مح}\,(\text{م}\,.\,\text{س}\,\text{ص} + \text{حـ}\,\text{ص} - \text{ص}^{٢})$$

$$= \text{مح}\,\text{ص}^{٢} - \text{م}\,\text{مح}\,\text{س}\,\text{ص} - \text{حـ}\,.\,\text{مح}\,\text{ص} \ldots \qquad (٥)$$

$$\text{ولكن } \text{مح}\,\text{ص}^{٢} = \text{ن}\,\text{عَ}^{٢} + \text{ن}\,\bar{\text{ص}}^{٢} \text{ ؛}$$

$$\text{و } \text{حـ}\,.\,\text{مح}\,\text{ص} = \text{حـ}\,.\,\text{ن}\,\bar{\text{ص}} = (\bar{\text{ص}} - \text{م}\,\bar{\text{س}})\,.\,\text{ن}\,\bar{\text{ص}} \text{ ،}$$

$$= \text{ن}\,\bar{\text{ص}}^{٢} - \text{م}\,.\,\text{ن}\,\bar{\text{س}}\,\bar{\text{ص}} \text{ ؛}$$

$$\text{و } \quad \text{م} = \text{ر}\,\frac{\text{عَ}}{\text{ع}}$$

فلو كتبنا $\text{ن}\,\text{سـهـ}^{٢}$ بدل $\text{مح}\,\text{ح}^{٢}$ فى معادلة (٥) وعوضنا هذه القيم فيها

$$\therefore \quad \text{ن}\,\text{سـهـ}^{٢} = \text{ن}\,\text{عَ}^{٢} + \text{ن}\,\bar{\text{ص}}^{٢} - \text{م}\,(\text{مح}\,\text{س}\,\text{ص} - \text{ن}\,\bar{\text{س}}\,\bar{\text{ص}}) - \text{ن}\,\bar{\text{ص}}^{٢}$$

$$= \text{ن}\,\text{عَ}^{٢} - \text{ر}\,.\,\text{ر}\,.\,\frac{\text{عَ}}{\text{ع}}\,(\text{ن}\,.\,\text{ر}\,\text{ع}\,.\,\text{عَ}) \text{ ،}$$ [انظر معادلة (١) (بند ٢٠٥)

$$\therefore \quad \text{سـهـ}^{٢} = \text{عَ}^{٢}\,(١ - \text{ر}^{٢}) \quad \ldots \ldots \qquad (٦)$$

$$\therefore \quad \text{ر} = \sqrt{١ - \frac{\text{سـهـ}^{٢}}{\text{عَ}^{٢}}} \quad \ldots \ldots \ldots \qquad (٧)$$

وهكذا قد أثبتنا أن هناك علاقة جبرية بين معامل الارتباط ر والتشتت سه ، للنقط المعطاة حول خط الانحدار أو خط الانتشار ، على فرض أن هذا الخط مستقيم . ومن الواضح أن ر تكون كبيرة إذا كانت سه صغيرة « أى التشتت صغير » ، والعكس بالعكس . وهذا هو ما أشرنا إليه فى بند ٢٣٨ .

الخطأ المعيارى لقيم س

٢٤٦ — هذا المقدار سه يشابه تماماً الانحراف المعيارى لمجموعة من القيم حول متوسطها الحسابى ، إذ أن الانحراف المعيارى يساوى الجذر التربيعى

لمجموع مربعات الانحرافات عن الوسط مقسوماً على عدد القيم ن . وكذلك سہ َ تساوى الجذر التربيعى لمجموع مربعات انحرافات النقط عن الخط (الذى سميناه خط العلاقة المتوسطة) مقسوماً على عدد النقط وهو ن . وللتمييز نسميه [1] **الخطأ المعيارى** للقيم الصادية ، ونرمز له بالرمز سہص أو للسهولة بالحرف سہ َ ، وذلك لأن هذه الانحرافات هى فى الحقيقة انحرافات القيم الفعلية التى يأخذها المتغير ص فى التجربة ، عن القيم النظرية التى كان يجب أن يأخذها هذا المتغير ، لو أنه اتبع القانون الاعتبارى الذى تمثله المعادلة المفروضة :

$$ص = م\ س + حـ$$

حساب ر من الخطأ المعيارى لقيم ص

٢٤٧ — هكذا نرى أنه من الممكن حساب معامل الارتباط بين متغيرين مثل س و ص ، بأن نوفق خط العلاقة المتوسطة بينهما ، ثم نحسب الفرق بين قيمة ص المشاهدة فى التجربة وقيمتها النظرية ، المحسوبة على أساس هذه المعادلة التى وفقناها . ثم نحسب الخطأ المعيارى سہ َ^{٢} بتربيع هذه الفروق وقسمة مجموع هذه المربعات على عددها . ونحسب أيضاً الانحراف المعيارى عَ للقيم الصادية نفسها المشاهدة من التجربة . ثم نحسب ر من المعادلة

$$سہ َ^{٢} = عَ^{٢} (١ - ر^{٢}) .$$

وهذه الطريقة أفضل وأعم ، وأحياناً تكون أسهل من الطرق العادية ؛ والحقيقة أننا لا نحتاج عملياً لحساب القيم النظرية من المعادلة التى نوفقها ثم نطرحها من القيم المشاهدة ؛ ولكننا نحصل على سہ َ^{٢} مباشرة من المعادلة (٥) بند ٢٤٥ وهى

$$ن\ سہ َ^{٢} = مجـ\ ح^{٢}$$

$$= مجـ\ ص^{٢} - م\ مجـ\ س\ ص - حـ\ مجـ\ ص .$$

---

(١) فى الإنجليزية Standard Error Sy,

ويلاحظ أن الكميات مجـ ص و مجـ س ص معلومة لنا حيث استخدمناها فى توفيق المعادلة الأصلية ص = م س + حـ . ومتى عرفنا الانحراف المعيارى للقيم الصادية ، وهو $\bar{\text{ع}}$ ، أمكننا بسرعة حساب مجـ $\text{ص}^{٢}$ من المعادلة

$$\text{مجـ ص}^{٢} = \text{ن}\ \bar{\text{ع}}^{٢} + \text{ن}\ \overline{\text{ص}}^{٢} \ .$$

خط الانحدار غير مستقيم

٢٤٨ — ننتقل الآن إلى البحث فى الارتباط غير المستقيم الذى فيه يكون خط الانحدار غير مستقيم ، بل منحنياً من الدرجة الثانية أو أعلى .

لنفرض أولا أن معادلة خط الانحدار من الدرجة الثانية بالنسبة إلى س ، وأنها على الصورة

$$\text{ص} = \text{ا} + \text{ب س} + \text{حـ س}^{٢} \ ، \ . \ . \ ، \qquad (١)$$

حيث ا ، ب ، حـ ثلاث كميات ثابتة ، نختارها بحيث يكون هذا المنحنى (١) أوفق المنحنيات لتمثيل القيم المشاهدة المعطاة وهى :

$$(\text{س}_{١} ، \text{ص}_{١}) ، (\text{س}_{٢} ، \text{ص}_{٢}) ، \ . \ . \ . \ . \ ، \ (\text{س}_{\text{ن}} ، \text{ص}_{\text{ن}}) \ .$$

و بطبيعة الحال تكون قيم ا ، ب ، حـ هذه تحقق المعادلات الثلاث :

$$\text{مجـ ص} = \text{ا} \ . \ \text{ن} + \text{ب} \ . \ \text{مجـ س} + \text{حـ} \ . \ \text{مجـ س}^{٢} \ . \ . \ (٢) \ ،$$

$$\text{مجـ س ص} = \text{ا} \ . \ \text{مجـ س} + \text{ب} \ . \ \text{مجـ س}^{٢} + \text{حـ} \ . \ \text{مجـ س}^{٣} \ . \ . \ (٣) \ ،$$

$$\text{مجـ س}^{٢}\text{ص} = \text{ا} \ . \ \text{مجـ س}^{٢} + \text{ب} \ . \ \text{مجـ س}^{٣} + \text{حـ} \ . \ \text{مجـ س}^{٤} \ . \ . \ (٤) \ ،$$

وهذه المعادلات الثلاث كما نعلم ، كافية لتعيين قيم المجاهيل الثلاثة ا ، ب ، حـ .

وهكذا نحصل على معادلة خط الانحدار للمتغير ص على المتغير س ؛ ويبقى أن نعرف كيف نستخدم هذه المعادلة لقياس الارتباط بين هذين المتغيرين كما فعلنا فى حالة الارتباط المستقيم .

٢٤٩ — لنبحث فى تشتت النقط ($س_١$ ، $ص_١$) ، . . . ، ($س_ن$ ، $ص_ن$) حول هذا الخط الاعتبارى : $ص = ا + ب\,س + حـ\,س^٢$. تشتت حول خط منحن

لنفرض أن $ح_١$ هو انحراف النقطة ($س_١$ ، $ص_١$) عن هذا المنحنى ،

$$\therefore \quad ح_١ = ا + ب\,س_١ + حـ\,س_١^٢ - ص_١ \quad . \quad . \quad . \quad (١) ،$$

ولو ضربنا هذه المتساوية فى $س_١$ و $س_١^٢$ و $ص_١$ و $ح_١$ على التوالى ،

$$\therefore \quad ح_١\,س_١ = ا\,س_١ + ب\,س_١^٢ + حـ\,س_١^٣ - س_١\,ص_١ \; . \; (٢) ،$$

$$و \quad ح_١\,س_١^٢ = ا\,س_١^٢ + ب\,س_١^٣ + حـ\,س_١^٤ - س_١^٢\,ص_١ \; \ldots \; (٣) ،$$

$$و \quad ح_١\,ص_١ = ا.\,ص_١ + ب.\,س_١\,ص_١ + حـ\,س_١^٢\,ص_١ - ص_١^٢ \; (٤) ،$$

$$و \quad ح_١^٢ = ا.\,ح_١ + ب\,ح_١\,س_١ + ح_١\,س_١^٢ - ح_١\,ص_١ \; . \; (٥) ،$$

وهكذا بالنسبة إلى جميع النقط الأخرى ($س_٢$ ، $ص_٢$) ، . . . ، ($س_ن$ ، $ص_ن$) .

فيكون مجموع مربعات انحرافات النقط جميعها عن هذا الخط يساوى :

$$مج\,ح^٢ = ن\,سـ^٢$$

$$= ا\,مج\,ح \quad + حـ\,مج\,س^٢\,ح \; - \; مج\,ح\,ص \; ، \; . \; . \; . \; (٦) .$$

ولكن $مج\,ح = ا\,ن \quad + ب\,مج\,س + حـ\,مج\,س^٢ - مج\,ص$ معادلة (١) أعلاه ،

$= صفراً$ « (٢) بند ٢٤٨.

و $مج\,ح\,س = ا.\,مج\,س + ب\,مج\,س^٤ + حـ.\,مج\,س^٣ - مج\,س\,ص$ « (٢) أعلاه ،

$= صفراً$ « (٣) بند ٢٤٨.

و $مج\,ح\,س^٢ = ا.\,مج\,س^٢ + ب.\,مج\,س^٣ + حـ\,مج\,س - مج\,س^٢\,ص$ « (٣) أعلاه ،

$= صفراً$ « (٤) بند ٢٤٨.

∴ ن سَ^{2} = مجـ ح^{2}

= — مجـ ح ص ؛ ومن معادلة (٤) أعلاه :

= مجـ ص^{2} — ا مجـ ص — ب مجـ س ص — حـ . مجـ س^{2} ص ، (٧)،

ويلاحظ أن الكميات مجـ س^{2} ص و مجـ س ص و مجـ ص هى نفس الكميات التى استخدمناها فى توفيق المنحنى وحساب قيم الكميات ا و ب و حـ . ولسنا محتاجين لحسابها من جديد . أما الكمية مجـ ص^{2} فهى معروفة بدلالة الوسط الحسابى والانحراف المعيارى لقيم المتغير ص ، حيث مجـ ص^{2} = ن عَ^{2} + ن صَ^{2} .

و بناء على ذلك يمكننا بسهولة معرفة مقدار تشتت النقط المعطاة حول خط الانحدار ، أو خط العلاقة المتوسطة ، الذى معادلته التحليلية هى :

ص = ا + ب س + حـ س^{2} .

والكمية سَ نطلق عليها أيضاً اسم **الخطأ المعيارى** لقيم ص كما قلنا فى بند ٢٤٦ .

دليل الارتباط

٢٥٠ — فى حالة الارتباط المستقيم وجدنا أن العلاقة الجبرية بين الخطأ المعيارى سَ ومعامل الارتباط ر هى :

ر^{2} = ١ — $\frac{\text{سَ}^{2}}{\text{عَ}^{2}}$ ، (١) ،

حيث عَ هى الانحراف المعيارى لقيم ص .

وقياساً على ذلك ، إذا كان ط يدل على معامل الارتباط فى هذه الحالة ، حالة الارتباط غير المستقيم ، تكون العلاقة الجبرية بين ط و سَ و عَ هى :

ط^{2} = ١ — $\frac{\text{سَ}^{2}}{\text{عَ}^{2}}$ (٢) ،

ويكون ط مقياس الارتباط غير المستقيم . وللتمييز بينه و بين ر ، معامل لارتباط المستقيم ، نسميه[١] **دليل الارتباط** .

وبوضع مجـ ص^{٢} = ن ع^{٢} + ن $\overline{\text{ص}}^{٢}$ فى المعادلة (٧) بند ٢٤٩ ،

واستبدال سـ^{٢} فى المعادلة (٢) أعلاه بقيمتها

∴ ط^{٢} = $\frac{١}{\text{ن ع}^{٢}}$ [ا مجـ ص + ب مجـ س ص + جـ . مجـ س^{٢}ص + ... − ن $\overline{\text{ص}}^{٢}$]

خط انحدار من الدرجة الثالثة

٢٥١ — وكذلك لو كان خط الانحدار ذا معادلة من الدرجة الثالثة مثل

ص = ا + ب س + جـ س^{٢} + ء س^{٣} . . . (١) .

يمكننا بنفس البرهان السابق إثبات أن الخطأ المعيارى فى هذه الحالة أيضاً هو سـ حيث

ن سـ^{٢} = مجـ ص^{٢} — ا مجـ ص — ب . مجـ س^{٢}ص — جـ مجـ س^{٢}ص — ء . مجـ س^{٣} ص؛

ويكون دليل الارتباط هو ط أيضاً ، حيث

ط^{٢} = ١ — $\frac{\text{سـ}^{٢}}{\text{ع}^{٢}}$

خط الانحدار من أى درجة

٢٥٢ — وهكذا لو كان خط الانحدار من أى درجة فوق الثالثة . وعلى العموم نفرض أن معادلة خط الانحدار من أى درجة ، ونفرض أن هذه المعادلة هى

ص = ا + ب س + جـ س^{٢} + ء س^{٣} + هـ س^{٤} + . . . (٢) ،

حيث ا ، ب ، جـ ، ء ، هـ ، . . . كميات ثابتة مستقلة عن كل من س و ص ، ويصح أن يكون بعضها سالباً أو موجباً أو صفراً .

---

(١) بالإنجليزية Index of Correlation ويرمز له بالحرف الأغريقى ρ (روه) .

وباتباع نفس الخطوات كما فى بند ٢٤٩ ، يمكننا بسهولة إثبات أن الخطأ المعيارى (وهو مقياس التشتت) هو سـَ حيث:

$$\therefore\ \text{سَ}^{2} = \text{مح ص}^{2} - \text{ا . مح ص} - \text{ب . مح س ص} - \text{حـ . مح س}^{2}\text{ص}$$

$$- \text{ء مح س}^{3}\text{ص} - \text{هـ . مح س}^{4}\text{ص} - \ldots \quad (٣) .$$

وفى حالة ما يكون خط الانحدار مستقيما ، تكون المعادلة (٢) من الدرجة الأولى فقط وتكون حـ = ء = هـ = . . . = صفراً . و إذا كان خط الانحدار من الدرجة الثانية كانت ء = هـ = . . . = صفراً . وهكذا . وفى كل هذه الأحوال يكون دليل الارتباط ، وهو مقياس العلاقة بين المتغيرين س و ص ، هو ط حيث :

$$\text{ط}^{2} = ١ - \frac{\text{سَ}^{2}}{\text{عَ}^{2}}$$

ط = ر فى حالة الانحدار المستقيم فقط

٢٥٣ — ويلاحظ أن دليل الارتباط ط يساوى معامل الارتباط ر فى حالة واحدة فقط ، ألا وهى حالة الارتباط المستقيم حيث يكون خط الانحدار مستقيما ، معادلته من الدرجة الأولى ، وحيث تؤول المعادلة (٣) فى البند السابق إلى الصورة .

$$\therefore\ \text{سَ}^{2} = \text{مح ص}^{2} - \text{ا . مح ص} - \text{ب مح س ص ،}$$

حيث ا هنا بدل حـ ، و ب بدل م المستعملتين فى بند ٢٤٥ ، ومنه ينتج

أن $\text{سَ}^{2} = \text{ع}^{2}(١ - \text{ر}^{2})$ ، كما فى معادلة (٦) بند ٢٤٥

$$\text{ط}^{2} = ١ - (١ - \text{ر}^{2})$$

$$= \text{ر}^{2} .$$

ولكن هذه العلاقة بين ط و ر ليست صحيحة فى حالة الانحدار غير المستقيم

على اختلاف أنواعه ، حيث تكون معادلة الخطأ المعيارى معقدة ، كما رأينا فى (٣) بند ٢٥٢ ؛ ولا يمكن اختصارها كما فعلنا هنا فى حالة الانحدار المستقيم .

ط أعم وأدق من ر فى قياس الارتباط

٢٥٤ — وكذلك العلاقة بين سـَ و ر ، وهى

$$\text{سـَ}^{2} = \text{ع}^{2} \left(1 - \text{ر}^{2} \right)$$

ليست صحيحة إلا فى حالة الانحدار المستقيم فقط . وعلى ذلك لا يمكن الاعتماد على معامل الارتباط ر كمقياس دقيق للارتباط فى حالات الانحدار غير المستقيم بل يجب فى هذه الأحوال استخدام ط . لأن العلاقة بين سـَ و ط هى دائماً

$$\text{سـَ}^{2} = \text{عَ}^{2} \left(1 - \text{ط}^{2} \right)$$

مهما كانت صورة معادلة خط الانحدار . ولذلك يعتبر دليل الارتباط أعم وأدق من معامل الارتباط فى قياس العلاقة بين متغيرين .

ر هى معامل الاستقامة لخط الانحدار

٢٥٥ — وهذا هو السبب فى أننا نرى حديثاً أن بعض الإحصائيين ينصح بعدم استعمال ر بتاتاً لقياس الارتباط ، ويقترح استخدامه فقط كمقياس لدرجة استقامة الانحدار ، وتسميته **معامل الاستقامة**[1] . وذلك لأنه إذا كان ر كبيراً كان مقدار سـَ صغيراً ، وبالتالى يكون تشتت النقط حول الخط المستقيم صغيراً ، مما يدل على أن هذا المستقيم يوافق جيداً النقط المعطاة . أى أن خط الانحدار قريب من المستقيم . وإذا كان ر صغيراً ، كان التشتت سـَ حول هذا المستقيم كبيراً ، مما يدل على أن هذا المستقيم لا يوافق هذه النقط جيداً . أى أن خط الانحدار بعيد عن الاستقامة .

---

(١) بالانجليزية Coefficient of Linearity ؛ أنظر بحث الأستاذ M. Frèchet فى مجلة Revue de l'Institut International de Statistique, 1933, Part 4, pp. 16—23.

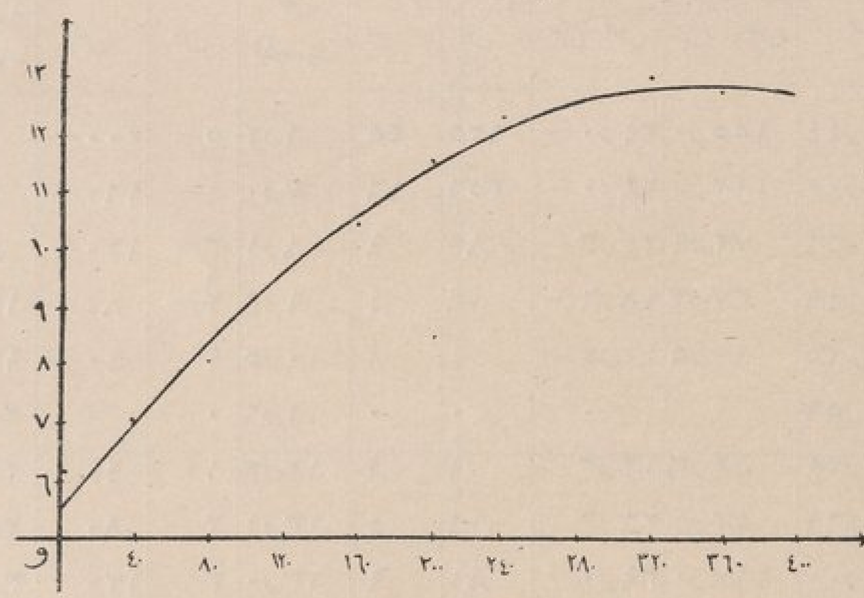
المقارنة بين ط و ر عملياً

٢٥٦ — نأخذ الآن مثالا ونحسب معامل الارتباط ودليل الارتباط ، ونقارن بينهما فى قياس العلاقة بين المتغيرين .

عملت تجربة فى عدد من الحقول بقصد دراسة العلاقة بين كمية السماد المستعمل وكمية المحصول الناتج ، فكانت الأرقام كما يأتى[1] :

كمية السماد للفدان بالكيلوجرام	٠	٤٠	٨٠	١٢٠	١٦٠	٢٠٠	٢٤٠	٢٨٠	٣٢٠	٣٦٠	٤٠٠
محصول الفدان بالأردب	٦٫٢	٧٫٠	٨٫١	٩٫٣	١٠٫٥	١١٫٦	١٢٫٣	١٣٫١	١٣٫	١٢٫٨	١٢٫٥

لنرمز لكمية السماد بالحرف س وللمحصول بالحرف ص . ثم نرسم محورين متعامدين ونرصد النقط التى إحداثياتها الأفقية تساوى قيم س ، وإحداثياتها الرأسية تساوى قيم ص على الترتيب ، كما فى شكل ٦٦ . ويتضح



(شكل ٦٦)

العلاقة بين كمية السماد والمحصول

---

(١) فى هذا المثال النظرى أخذنا ١١ حقلا فقط وذلك لتوضيح خطوات العمل . ولكن عملياً يجب ألا يقل العدد عن حوالى ٣٠ ؛ وذلك لكى نتفادى أخطاء المصادفات .

من مواقع هذه الإحدى عشرة نقطـة أن أحسن خط يوافقها تكون معـادلته من الدرجة الثانية . لنفرض هذه المعادلة هى :

ص = ا + ب س + حـ س^{٢} (١)

حيث ص تساوى كمية المحصول و س تمثل كمية السماد . ونوجد قيم ا ، ب ، حـ كالمعتاد من المعادلات .

مجـ ص = ١١ ا + ب . مجـ س + حـ مجـ س^{٢} (٢) ،

و مجـ س ص = ا . مجـ س^{٢} + ب . مجـ س^{٢} + حـ . مجـ س^{٣} (٣) ،

و مجـ س^{٢}ص = ا . مجـ س^{٢} + ب . مجـ س^{٣} + حـ . مجـ س^{٤} (٤) ،

ولتسهيل حل هذه المعادلات لإيجاد قيم ا ، ب ، حـ ، نبدأ قياس السينات من الوسط حتى يكون بعضها سالباً والبعض موجباً فتختصر ، ولا سيما أن قيم س المعطاة واقعة على مسافات متساوية من بعضها . ونجد خطوات العمل موضحة فى الجدول الآتى :

جدول ٤٤ — توفيق منحن من الدرجة الثانية وحساب دليل الارتباط

كمية السماد	س	سَ	ص المحصول	سَ^{٢}	سَ^{٤}	سَ ص	سَ^{٢} ص	ص^{٢}
٠	—٢٠٠	—٥	٦ر٢	٢٥	٦٢٥	—٣١ر٠	١٥٥ر٠	٣٨ر٤٤
٤٠	—١٦٠	—٤	٧ر٠	١٦	٢٥٦	—٢٨ر٠	١١٢ر٠	٤٩ر٠٠
٨٠	—١٢٠	—٣	٨ر١	٩	٨١	—٢٤ر٣	٧٢ر٩	٦٥ر٦١
١٢٠	—٨٠	—٢	٩ر٣	٤	١٦	—١٨ر٦	٣٧ر٢	٨٦ر٤٩
١٦٠	—٤٠	—١	١٠ر٥	١	١	—١٠ر٥	١٠ر٥	١١٠ر٢٥
٢٠٠	٠	٠	١١ر٦	٠	٠	٠	٠	١٣٤ر٥٦
٢٤٠	٤٠	١	١٢ر٣	١	١	١٢ر٣	١٢ر٣	١٥١ر٢٩
٢٨٠	٨٠	٢	١٣ر١	٤	١٦	٢٦ر٢	٥٢ر٤	١٧١ر٦١
٣٢٠	١٢٠	٣	١٣ر٠	٩	٨١	٣٩ر٠	١١٧ر٠	١٦٩ر٠٠
٣٦٠	١٦٠	٤	١٢ر٨	١٦	٢٥٦	٥١ر٢	٢٠٤ر٨	١٦٣ر٨٤
٤٠٠	٢٠٠	٥	١٢ر٥	٢٥	٦٢٥	٦٢ر٥	٣١٢ر٥	١٥٦ر٢٥
المجموعات		٠	١١٦ر٤	١١٠	١٩٥٨	٧٨ر٨	١٠٨٦ر٦	١٢٩٦ر٣٤

و بالتعويض فى المعادلات (٢) ، (٣) ، (٤) أعلاه

∴ ١١٦٫٤ = ١١ ا + ٠ + ١١٠ حـ (٤) ،

و ٧٨٫٨ = ٠ + ١١٠ ب + ٠ (٥) ،

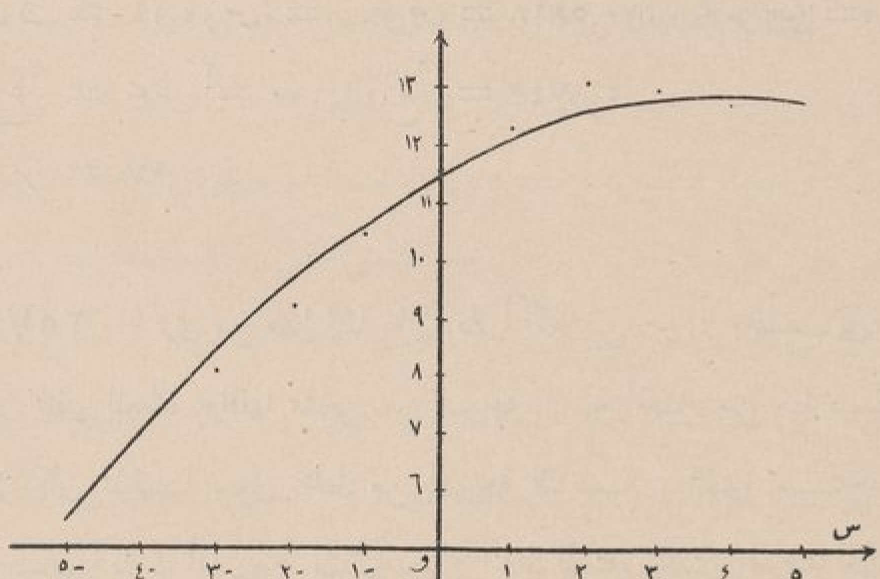
و ١٠٨٦٫٦ = ١١٠ ا + ٠ + ١٩٥٨ حـ (٦) ،

∴ ا = ١١٫٤٨٣٩ ، ب = ٠٫٧١٦٤ ، حـ = − ٠٫٠٩٠٢٠٠

∴ معادلة خط الانحدار هى :

ص = ١١٫٤٨٣٩ + ٠٫٧١٦٤ سَ − ٠٫٠٩٠٢ سَ^{٢} (٧) ،

حيث ص تساوى المحصول بالأرادب ، فى حين أن سَ تدل على كمية السماد محسوبة ابتداء من ٢٠٠ كيلو جرام ، بوحدات كل منها تساوى ٤٠ كيلو جراماً . ونرى هذا المنحنى مرسوماً فى شكل ٦٧ .



(شكل ٦٧)

ولكن ١١ سـ^{٢} = مجـ س^{٢} − ا مجـ ص − ب مجـ ص − حـ مجـ س^{٢} ص

∴ = ١٢٩٦٫٣٤ − ١٣٣٦٫٧٢٦٠ − ٥٦٫٤٥٢٣ + ٩٨٫٠١١٣

= ١٫١٧٣٠

$$\therefore \quad \bar{\text{س}}^{2} = \text{١٠٦٦ر}$$

$$\therefore \quad \text{مج ص}^{2} = \text{١١ } \bar{\text{ع}}^{2} + \text{١١ } \bar{\text{ص}}^{2} \text{ و ١١ } \bar{\text{ص}} = \text{١١٦٫٤}$$

$$\therefore \quad \bar{\text{ع}}^{2} = \frac{\text{١}}{\text{١١}} \times \text{١٢٩٦٫٣٤} - (\text{١٠٫٥٨١٨})^{2} \text{ ،}$$

$$= \text{٥٫٨٧٤٢ .}$$

$$\therefore \quad \text{ط}^{2} = \text{١} - \frac{\text{١٠٦٦ر}}{\text{٥٫٨٧٤٢}}$$

$$= \text{٩٨١٨ر}$$

$$\therefore \quad \text{ط} = \text{٩٩ ر}$$

ولحساب معامل الارتباط ر ، تستخدم المعادلة

$$\text{ر} = \frac{\text{مج س ص} - \text{ن } \bar{\text{س}}\,\bar{\text{ص}}}{\text{ن} \cdot \text{ع} \cdot \bar{\text{ع}}} \text{ ،}$$

حيث ن = ١١ ، و $\bar{\text{س}}$ = ٠ ، و $\bar{\text{ص}}$ = ١٠٫٥٨١٨ ، و مج س ص = ٧٨٫٨

و ١١ع^{2} = مج س^{2} = ١١٠ ، و $\bar{\text{ع}}^{2}$ = ٥٫٨٧٤٢

$$\therefore \quad \text{ر} = \text{٩٧ر}$$

قيمة ط تتوقف على أنوع المنحنى المفروض كخط الانحدار

٢٥٧ — نرى من هذا المثال أن ط أكبر من ر ؛ والسبب فى ذلك هو أن القيم المعطاة يوافقها منحن من الدرجة الثانية أفضل من خط مستقيم. ولذلك كان تشتتها حول الخط من الدرجة الثانيـة ، الذى اسـتخدمناه فى حساب ط ، أقل من تشتتها حول الخط المستقيم الذى افترضنا موافقته لتمثيل هذه القيم ، فى حساب المعامل ر .

ولو رأينا فى بادىء الأمر — خطأً — أن القيم المعطاة يوافقها منحن من الدرجة الثالثـة ، بدلا من الدرجة الثانية ، وحسبنا دليل الارتباط على هذا الأساس ، وليكن طَ مثلا ، لوجدنا أن طَ أصغر من ط والسبب فى ذلك أن

تشتت القيم حول هذا الخط الأخير أكبر من تشتتها حول الخط الأصلى ذى الدرجة الثانية ، الذى يوافقها فى الحقيقة أفضل من الخط ذى الدرجة الثالثة .

وبالمثل ، لو أن هـذه القيم المعطاة يوافقها خط مستقيم أحسن من أى خط آخر ، وأننا رأينا (خطأ ، لسبب ما) أن نوفق لها منحنياً من الدرجة الثانيـة ، فحسبنا ط على هذا الأساس ، لوجدنا أن معامل الارتباط ر أكبر من الدليل ط . والسبب فى ذلك هو نفس السبب المتقـدم : ألا وهو أن تشتت القيم حول الخط الجيد أقل من تشتتها حول أى خط آخر أقل موافقة من الأول . والقـاعدة إذن هى أن مقياس الارتباط (ر أو ط) المحسوب على أساس الخط الجيد ، أكبر من المقياس المحسوب على أساس أى خط آخر أقل منه توفيقاً .

قياس الارتباط فى النهاية مسألة اعتبارية يدخل فيها العـامل الشخصى

٢٥٨ — مقياس الارتباط الذى نحصل عليه فى أى مسألة معينـة يتوقف إذن على نوع المنحنى الذى نعتبره صالحاً لتمثيل العـلاقة المتوسطة بين المتغيرين . ومسألة اختيار المنحنى الأصلح ، وإن كانت تبدو سهلة فى ظاهرها ، حيث يسهل فى بعض الأحيان الاتفـاق على نوع المنحنى الذى يوافق مجموعة من القيم المشاهدة لمتغيرين أحسن من غيره ، كثيراً ما تكون صعبة أو ملتبسة لا يمكن الوصول فيها إلى رأى قاطع . وفى مثل هذه الأحوال لا بد أن يترك أمر اختيار المنحنى الأصلح إلى تقديرنا واعتبارنا . وبهذه الطريقة يدخل العامل الشخصى فى أبحاثنا ويتحكم فى المقياس الذى نحصل عليه لدرجة العلاقة بين المتغيرين تحت البحث .

وهكذا نرى أن مسألة قياس الارتباط بين متغيرين علمت لنا مجموعة من القيم المشاهدة لهما ، تسـتند فى النهاية على أمور اعتبارية — إلى حدٍّ ما على الأقل . وكنا نود لو أن هـذا العامل الشخصى لم يتدخل فى حكمنا على الأشـياء ، ولـكن للأسف لا توجد قاعدة مضبوطة نرجع إليهـا فى معرفة أى نوع من المنحنيات

يصلح لتمثيل مجموعة من القيم المعينة أفضل من غيره . وهذه حقيقة لا نستطيع في الوقت الحاضر أن نكذبها أو ننكرها .

نسبة الارتباط

٢٥٩ — إذا كان لدينا جدول تكرارى مزدوج لمتغيرين مثل س و ص ، ورأينا من توزيع قيمهما فى هذا الجدول أن خط الانحدار بينهما لا يمكن أن يكون مستقيماً ، فلا يمكننا والحالة هذه الاعتماد على معامل الارتباط ر كمقياس دقيق للعلاقة بين المتغيرين . ولا يمكننا أيضاً حساب دليل الارتباط ط ، لأن هذا لا بد له من توفيق منحن لقيم ص بالنسبة إلى س ؛ وهذا غير ميسور ، إذ أن قيم س وقيم ص ليست عندنا منفردة ، بل مقسمة إلى مجموعات فى فئات تكرارية مزدوجة . وقد وضع كارل بيرسون مقياساً سماه [١] **نسبة الارتباط** ، لاستخدامه فى مثل هذه الأحوال . ولنرمز لهذا المقياس بالحرف ى . ونحسبه كما يأتى :

أولا : نحسب المتوسط الصادى لكل عمود (أى فى كل فئة من فئات س) ؛

ثانياً : فى كل خانة من أى عمود ، نحسب مربع انحراف قيمة ص عن الوسط الحسابى الصادى فى هذا العمود ؛

ثالثاً : نجمع مربعات الانحرافات الصادية لكل عمود ، فنحصل على تشتت الصادات فى هذا العمود حول وسطها الحسابى ؛

رابعاً : نجمع هذه التشتتات للأعمدة كلها فنحصل على مجموع مربعات انحرافات القيم الصادية كل عن الوسط الحسابى فى عمودها . لنفرض أن هذا المجموع يساوى ن $\sigma^2_{م\,ص}$ ، حيث ن يساوى عدد الحالات فى الجدول كله .

---

(١) اسمه بالإنجليزية Correlation Ratio ، ويرمز له بالحرف الأغريقى η (إيتا) أنظر كتاب Whittaker and Robinson : *Calculus of Observations*, (1929), p. 336.

خامساً : نحسب الانحراف المعيارى لقيم ص ؛ وليكن هذا يساوى عَ كالمعتاد

سادساً : نحسب ى من المعادلة

$$\text{ى}^{\text{٢}} = \text{١} - \frac{\text{سـ}^{\text{٢}}_{\text{م ص}}}{\text{عَ}^{\text{٢}}} \qquad (\text{١})\,.$$

ويلاحظ أن هذه المعادلة على صورة معادلات ط و ر ، ما عدا التشتت $\text{سـ}_{\text{م ص}}$ هنا بدل التشتت البسيط سـَ هناك .

٢٦٠ — ولكن حساب التشتت $\text{سـ}_{\text{م ص}}$ بالطريقة المتقدمة متعب ويستلزم مجهوداً كبيراً ، ولذلك نحسبه بطريقة مختصرة غير مباشرة ، بواسطة حساب الانحراف المعيارى لنفس المتوسطات الحسابية للأعمدة التى أوجدناها فى الخطوة الأولى أعلاه . لنفرض أن هذا الانحراف المعيار يساوى $\text{ع}_{\text{م}}$ مثلا . حساب التشتت $\text{سـ}_{\text{م ص}}$ عملياً

وبما أن مجموع مربعات انحرافات عدد من القيم كل واحدة عن الوسط الحسابى لمجموعتها ، يساوى مجموع مربعات انحرافاتها عن الوسط الحسابى العمومى ، ناقصاً مجموع مربعات انحرافات الأوساط الحسابية للمجموعات عن هذا الوسط العمومى نفسه[١] ،

$$\therefore \text{ن} \,.\, \text{سـ}^{\text{٢}}_{\text{م ص}} = \text{ن}\,\text{عَ}^{\text{٢}} - \text{ن}\,\text{ع}^{\text{٢}}_{\text{م}} \qquad (\text{٢})\,،$$

$$\therefore \text{ى}^{\text{٢}} = \text{١} - \frac{\text{عَ}^{\text{٢}} - \text{ع}^{\text{٢}}_{\text{م}}}{\text{عَ}^{\text{٢}}} \qquad (\text{٣})\,،$$

$$\therefore \text{ى} = \frac{\text{ع}_{\text{م}}}{\text{عَ}} \qquad (\text{٤})\,،$$

وهذه العلاقة أبسط من المعادلة (١) ، من ناحية الشكل ومن الناحية

---

(١) انظر المعادلة (٣) فى بند ١٨٢ صفحة ٢٠١ ، حيث نضع ن $\text{سـ}^{\text{٢}}_{\text{م ص}}$ بدلا من مح ن $\text{ع}^{\text{٢}}$ هناك .

العملية الحسابية أيضاً ، حيث تؤول المسألة إلى حساب المتوسطات الحسابية للأعمدة ، وحساب انحرافها المعيارى ، والانحراف المعيارى لعموم الصادات ؛ ونسبة الارتباط هى خارج قسمة الأخيرين .

نسبة الارتباط تساوى دليل الارتباط فى النهاية

٢٦١ — وإذا تأملنا فى الفكرة الأساسية التى ينبنى عليها تعريف نسبة الارتباط ، نجد أن هذه النسبة تساوى (تقريباً) دليل الارتباط . فقد قلنا إن د سهم^{٢} س ، المستعملة فى حساب ى ، هى مجموع مربعات انحرافات القيم الصادية ، كل عن الوسط الحسابى لمجموعتها ، فى الجدول التكرارى المزدوج ؛ وكل مجموعة من هذه القيم الصادية تقترن بقيمة معينة للمتغير س ، ألا وهى مركز الفئة السينية المبينة أعلى كل عمود من أعمدة هذا الجدول التكرارى . أنظر مثلا فى جدول ٣٢ (صحيفة ٢٣٥) حيث نجد فى العمود الثالث منه ٥٣ رجلا عمرهم جميعاً يساوى ٢٧,٥ سنة . ولكن عدد ماعندهم من الأطفال يختلف : فمنهم ثمانية ليس عندهم أطفال بالمرة ، ومنهم ٢٥ عندهم طفل واحد ، وهكذا . ومتوسط ماعند هؤلاء من الأطفال يساوى ١,٣٨ طفل .

وقلنا إن د سهَ^{٢} المستعملة فى حساب ط ، تساوى مجموع مربعات انحرافات القيم الصادية ، كل عن القيمة النظرية المحسوبة على أساس معادلة خط الانتشار ، أو خط العلاقة المتوسطة (وهو فى النهاية خط الانحدار) ، الذى فرضنا أنه يمثل العلاقة بين المتغيرين س ، ص . ولو رجعنا إلى شكل ٦٤ (صحيفة ٢٦٩) ، وما قلناه فى بند ٢٣٩ ، وجدنا أن هذه القيمة النظرية هى فى النهاية متوسط القيم الصادية الواقعة فى هذه الفئة . ومن ذلك يتضح أن سهَ و سهم ، وبالتالى ط ، ى ، متساويتان فى النهاية ؛ وأن الفرق بينهما عملياً يكون صغيراً ، خصوصاً لو كانت الفئات السينية فى جدول الارتباط ضيقة المدى ، وكذلك فئات ص .

الفرق بين دليل الارتباط ونسبة الارتباط ى ، لم يخرج إذن عن كونه تقريباً فى حدود الدقة ، بقصد تسهيل العمل الحسابى . فتستعمل ط حينما يكون لدينا قيم منفردة متناظرة للمتغيرين س و ص ، عددها صغير لا يستلزم مجهوداً حسابياً كبيراً . أما إذا كان لدينا عدد كبير جداً من قيم س و ص فقسمناها إلى فئات ووضعناها فى جدول تكرارى مزدوج ، تعذر علينا استخدام ط ونلجأ حينئذ إلى حساب نسبة الارتباط ى من واقع هذا الجدول .

ويلاحظ بهذه المناسبة أن استخدام ى يغنينا عن عملية توفيق المنحنى اللازم لحساب ط ؛ ويعفينا من مواجهة المشكلة المرتبطة بهذه العملية والتى تكلمنا عنها فى بند ٢٥٧ ، ألا وهى اختيار أصلح المنحنيات لتمثيل البيانات المعطاة لنا .

حساب نسبة الارتباط عملياً

٢٦٢ — نأخذ الآن مثالا مختصراً لتوضيح خطوات العمل فى حساب ى . فى جدول ٤٥ نجد التوزيع التكرارى المزدوج لأعمار الرجال المتزوجين (من آنسات) وأعمار زوجاتهم يوم الزواج . والملاحظ عادة أن الآنسات يتزوجن فى مقتبل العمر ، ولا يبقى منهن لسن متأخرة إلا القليل ، ومهما بقين فهن لا يتزوجن بعد سن معينة ، إلا نادراً جداً . أما الرجال فكثيراً ما يتزوجون فى السنين المتأخرة من العمر .

نحسب المتوسط الصادى لكل عمود ، وهو متوسط أعمار زوجات الرجال الذين من سن واحدة . ويلاحظ أن هذه المتوسطات تزيد بسرعة فى بادىء الأمر مع أعمار الرجال ، ولكن معدل هذه الزيادة لا يلبث أن يهبط ويتلاشى تدريجياً ، مما يدل على أن أعمار الزوجات من الآنسات يزيد تبعاً لأعمار أزواجهن إلى حد معين ثم يقف . أى أن خط انحدار ص على س ليس مستقيما . ومن هذا يتضح أن معامل الارتباط ر لا يصلح فى هذه الحالة ، وينبغى إذن حساب نسبة الارتباط ى .

جدول ٤٥ — توزيع تكرارى مزدوج لأعمار الرجال = س ،

وأعمار زوجاتهم = ص ؛ عند الزواج

س / ص	٢٠ -	٣٠ -	٤٠ -	٥٠ -	٦٠ -	٧٠ -	٨٠ -	المجموع
١٥ -	٣٥	٦٣	٢٢	٢				١٢٢
٢٠ -	٢٥	٧٢	٣٣	١٣	٥	١		١٤٩
٢٥ -	١٧	٥١	٣٠	١٦	٨	٢	١	١٢٥
٣٠ -	٢	٣٣	٢٨	١٩	١٥	٧	٣	١٠٧
٣٥ -	١	١٨	٢٠	١٥	٧	٤	٠	٦٥
٤٠ -		٣	١٢	١٠	٥	١	١	٣٢
المجموع	٨٠	٢٤٠	١٤٥	٧٥	٤٠	١٥	٥	٦٠٠
متوسط ص	٢١٫٨١٢٥	٢٥٫٠٠	٢٨٫٤٣١٠	٣١٫٦٣٣٣	٣٢٫٣٧٥٠	٣٣٫١٦٦٦	٣٣٫٥٠٠	

لذلك نحسب الانحراف المعيارى عم لهذه المتوسطات الحسابية الصادية ؛ وكذلك الانحراف المعيارى لقيم ص ؛ وتكون ى تساوى خارج قسمة الأول على الثانى .

لإيجاد الانحراف المعيارى لقيم ص ، نستخدم التوزيع التكرارى لها ، كما هو مبين بالجدول السابق .

جدول ٤٦ — حساب الانحراف المعيارى لأعمار الزوجات

الفئات	مراكزها ص	التكرار ك	الانحرافات ح	ك . ح	ك . ح^{٢}
١٥—	١٧٫٥	١٢٢	١٠—	١٢٢٠—	١٢٢٠٠
٢٠—	٢٢٫٥	١٤٩	٥—	٧٤٥—	٣٧٢٥
٢٥—	٢٧٫٥	١٢٥	٠	٠	٠
٣٠—	٣٢٫٥	١٠٧	٥	٥٣٥	٢٦٧٥
٣٥—	٣٧٫٥	٦٥	١٠	٦٥٠	٦٥٠٠
٤٠—	٤٢٫٥	٣٢	١٥	٤٨٠	٧٢٠٠
		٦٠٠		٣٠٠—	٣٢٣٠٠

∴ الوسط الحسابى العمومى لقيم ص يساوى

$$\bar{ص} = ٢٧٫٥ - \frac{٣٠٠}{٦٠٠} ،$$

$$= ٢٧ \quad \text{سنة} .$$

$$\text{و} \quad ٣٢٣٠٠ = ٦٠٠ \, \acute{ع}^{٢} + ٦٠٠ \, (٠٫٥)^{٢}$$

$$\therefore \quad \acute{ع}^{٢} = ٥٣٫٥٨٣٣$$

$$\therefore \quad \acute{ع} = ٧٫٣٢٠٦ \ \text{سنة} .$$

ولحساب الانحراف المعيارى عم للمتوسطات الصادية حول الوسط الحسابى العمومى للصادرات وهو ٢٧ سنة ، نربع انحراف كل منها عن ٢٧ ، ونضرب هذا المربع فى عدد المفردات التى يمثلها المتوسط . ونجد الخطوات موضحة فى الجدول الآتى :

جدول ٤٧ — حساب الانحراف المعيارى للمتوسطات

المتوسط الصادى	تكرار العمود ك	الانحراف عن ٢٧ ، أى ح	ح^٢	ك . ح^٢
٢١,٨١٢٥	٨٠	—٥,١٨٧٥	٢٦,٩١٠٢	٢١٥٢,٨١٢٥
٢٥,٠٠٠٠	٢٤٠	—٢,٠٠	٤,٠٠٠٠	٩٦٠,٠٠٠٠
٢٨,٤٣١٠	١٤٥	١,٤٣١٠	٢,٠٤٧٨	٢٩٦,٩٢٥٣
٣١,٦٣٣٣	٧٥	٤,٦٣٣٣	٢١,٤٦٧٥	١٦١٠,٠٦٠١
٣٢,٣٧٥٠	٤٠	٥,٣٧٥٠	٢٨,٨٩٠٦	١١٥٥,٦٢٥٠
٣٣,١٦٦٦	١٥	٦,١٦٦٦	٣٨,٠٢٧٠	٥٧٠,٤٠٤٣
٣٣,٥٠٠٠	٥	٦,٥٠٠٠	٤٢,٢٥٠٠	٢١١,٢٥٠٠
	٦٠٠			٦٩٥٧,٠٧٧

∴ ٦٠٠ $ع_م^٢$ = ٦٩٥٧,٠٧٧ ،

و $ع_م$ = ٣,٤٠٥٢ .

∴ ى = $\frac{٣,٤٠٥٢}{٧,٣٢٠٦}$ ،

= ٠,٤٦٥

ولو حسبنا معامل الارتباط بين س و ص نجد أنه يساوى ٠,٤٥٠ ومن ذلك يظهر لنا أن ى أكبر من ر . وهذا هو الواجب أن يكون ، حيث قد رأينا أن خط الانحدار فى هذه المسألة غير مستقيم . أى أننا لو كنا وفقنا خطين للقيم المعطاة فى المسألة ، أحدهما مستقيم معادلته من الدرجة الأولى ، والآخر منحن معادلته من الدرجة الثانية ، كنا نجد تشتت النقط حول المستقيم أكبر من تشتتها

حول الخط الآخر . وعلى ذلك فمقياس الارتبـاط المبنى على فرض أن الانحـدار مستقيم ، وهو معامل الارتباط ر ، يكون أصغر من المقياس الآخر الذى لا يستند على هذا الفرض الخطأ .

نسبة الارتباط أعم من المعامل ر

٢٦٣ — ولو كان الانحـدار فى هذه المسألة مستقيما ، كنا نجد أن نسـبة الارتباط تساوى معامل الارتباط . وعلى ذلك نقول إن نسبة الارتباط ى مقياس أعم من معـامل الارتباط ر ، وتعطينا نتائج أدق فى الأحوال التى يخفق فيهـا معامل الارتباط فى قياس العلاقة بين المتغيرين وتصويرها على حقيقتها . ويمكننا إذن الاعتماد على نسبة الارتباط ى فى جميع الأحوال ؛ بخلاف معامل الارتباط لا نستعمله إلا إذا علمنا أن الانحـــدار بين المتغيرين مستقيم ، أو قريب جداً من الاستقامة .

خط انحـدار س على ص

٢٦٤ — يلاحظ أننا قصرنا الـكلام فيما سبق على خط انحدار المتغير ص على المتغـير س ، وبحثنا فى توفيق معادلات خطوط الانحـدار على صورة تعطى المتغـير ص منفرداً فى أحد طرفى المعادلة ، وقوى س وحدها فى الطرف الآخر ، كما فى بند ٢٥٦ مثلا حيث :

$$ص = ١١{,}٤٨٣٩ + {,}٧١٦٤\,س - ٠و٠٩٠٢\,س^{٢} .$$

أى أننا اعتبرنا س كمتغير مسـتقل واعتبرنا ص كمتغير تابع له . وحسـبنا المقاييس ر أو ط أو ى على هذا الأساس ، فحسبنا التشتت حول خط انحـدار ص على س .

أما إذا كانت معادلة خط الانحدار تعطى س فى صورة متغير تابع إلى ص ، أى خط انحدار س على ص ، فإننا نحسب $سه^{٢}$ بدل $سهَ^{٢}$ ، أى الخطأ المعيارى

لقيم س بدل الخطأ المعيارى لقيم ص ؛ ونستعمل ع بدل عَ ، و $\text{ع}_{\text{م س}}$ بدل $\text{ع}_{\text{م ص}}$. فنحسب مقياس الارتباط من إحدى المعادلات الآتية :

$$\text{ر}^{\text{٢}} \text{ أو } \text{ط}^{\text{٢}} = \text{١} - \frac{\text{سـ}}{\text{ع}^{\text{٢}}}$$

و
$$\text{ى}^{\text{٢}} = \text{١} - \frac{\text{سـ}^{\text{٢}}_{\text{م س}}}{\text{ع}^{\text{٢}}}$$

أى
$$\text{ى} = \frac{\text{ع}_{\text{م س}}}{\text{ع}}$$

حيث $\text{ع}_{\text{م س}}$ هى الانحراف المعيارى للمتوسطات السينية .

كل مقاييس الارتباط تبنى على أســاس التشتت حول خط الانحدار

٢٦٥ — والخلاصة التى نخرج بها من هذا البــاب والباب الســابق ، هى أن العلاقة بين كميتين متغيرتين يمكن قياسها بعدة مقاييس أهمها ثلاثة ، وهى معــامل الارتبــاط ر ، ودليل الارتبــاط ط ، ونسبة الارتبــاط ى . وأن الاختيار بين هذه الثلاثة يتوقف على نوع العلاقة بين المتغيرين ، أو بالأحرى على شكل خط الانحدار . والنقطة المهمة فى هذا الموضوع هى أن الفكرة الأساسية فى هذه المقــاييس الثلاثة أصلها واحد فى كل الأحوال ، وهو مقدار تشتت القيم المعطاة للمتغيرين حول خط العلاقة المتوسطة بينهما ، أى خط الانحدار .

المــــراجع

Bowley, A.L., : *Elements of Statistics,* Part II. Chapters VI., VII.

Mills, F.C., : *Statistical Methods*, Chapter XII.

الباب العاشر

الارتباط المتعدد والارتباط الجزئى

---

٢٦٦ — فى البابين الثامن والتاسع درسنا الارتباط بين ظاهرتين فقط ، واستنبطنا مقاييس مختلفة لقياس العلاقة بين المتغيرين ، على فرض أنهما لا يتأثران بأى عوامل أخرى خارجية عنهما .

ولكن هذا الفرض قليلا ما يتحقق عملياً ، خصوصاً فى المسائل الاقتصادية والاجتماعية ، وفى المسائل العلمية البحتة أيضاً ، حيث نرى الظاهرة التى نبحثها تتأثر بعوامل كثيرة بدرجات مختلفة . ولدراسة العلاقة بين هذه الظاهرة والظواهر الأخرى مجتمعة أو منفردة ، يجب علينا التفكير فى مقاييس للارتباط تأخذ فى الحسبان هذه الظروف الجديدة ، التى لم نعرها اهتماماً من قبل .

وسنشرح فى هذا الباب باختصار مقاييس الارتباط التى يمكن استعمالها فى مثل هذه المسائل . وسنقتصر على دراسة الارتباط بين الظواهر التى يمكن قياسها رقمياً .

الارتباط المتعدد والارتباط الجزئى

٢٦٧ — لنفرض أننا ندرس العلاقة بين ظاهرة مثل كمية محصول الفدان (من القمح أو الأذرة أو . . .) وظواهر أخرى مثل كمية السماد المستعمل وكمية مياه الرى ومتوسط درجة حرارة الجو أثناء النمو . كل واحد من هذه العوامل الأخيرة يؤثر فى كمية المحصول ، وكل منها يتغير تبعاً لمؤثرات خاصة وقد يكون فى تغيره مستقلا عن العاملين الآخرين أو مرتبطاً بهما أو بأحدهما .

وكمية المحصول تتغير تبعاً لهذه العوامل الثلاثة . أى أن كمية السماد (س) وكمية مياه الرى (ص) ودرجة الحرارة (ع) كلها متغيرات إما مستقلة أو مرتبطة ، وكمية المحصول (م) متغير تابع لها جميعاً .

إذا أردنا قياس العلاقة بين كمية المحصول م والظواهر الأخرى س و ص و ع مجتمعةً ، أى أنها تتغير جميعاً فى وقت واحد بدون قيد ولا شرط ، فنستخدم لذلك ما نسميه [1] **معامل الارتباط المتعدد** . أما العلاقة بين كمية المحصول م وأحد العوامل الأخرى — مثل س — منفرداً ، أى بفرض أن العاملين الآخرين ص و ع يبقيان ثابتين أثناء تغير س و م ، فتسمى [2] **الارتباط الجزئى** ؛ وهذا نقيسه باستخدام **معامل الارتباط الجزئى** .

نحصل من التجربة على قيم متناظرة للمتغيرات

٢٦٨ — لدراسة الارتباط المتعدد وقياسه عملياً ، نبدأ كالمعتاد بمشاهدة الظواهر المتغيرة التى نبحثها فى عدة ظروف أو حالات مختلفة . وفى كل حالة ندون قيمة كل من هذه المتغيرات الأربعة م و س و ص و ع ؛ فنحصل بذلك على جدول به أربع سلسلات من القيم المتناظرة لهذه المتغيرات . فإذا أجرينا عدة تجارب فى حقول مختلفة (أو سنين مختلفة) على هذا المحصول ، وكانت قيم كل من س و ص و ع تختلف بعضها عن بعض من حقل إلى آخر ، وكذلك تتغير م تبعاً لذلك ، حصلنا لكل من هذه المتغيرات على قيم بعدد حقول التجارب التى لدينا ، وليكن ن مثلا .

توفق معادلة تربط هذه المتغيرات

من هذه القيم المشاهدة للمتغيرات م و س و ص و ع ، نوفق معادلة جبرية تمثل العلاقة بين م كمتغير تابع وبين س و ص و ع كمتغيرات مستقلة ،

(١) بالانجليزية Coefficient of Multiple Correlation

(٢) بالانجليزية Coefficient of Partial Correlation

أحسن ما يمكن . وذلك بطريقة المربعات الصغرى أو بأى طريقة أخرى .

مثال معادلة من الدرجة الأولى

٢٦٩ — نأخذ على سبيل التمثيل الحالة البسيطة التى فيها تكون المعادلة بين م و س و ص و ع من الدرجة الأولى ، ولنفرض أننا وفقنا هذه المعادلة بطريقة المربعات الصغرى فكانت كما يأتى :

م = ا س + ب ص + حـ ع + ء ، (١)

حيث ا و ب و حـ و ء كميات عددية ثابتة مستقلة عن المتغيرات م و س و ص و ع . وهذه الكميات نختارها بحيث تجعل هذه المعادلة أصلح من أى معادلة أخرى لتمثيل القيم المعطاة . وهذه الكميات نستخرجها ، كما نعلم فى طريقة توفيق المنحنيات ، من المعادلات[١] الآنية الآتية :

مجـ م = ا . مجـ س + ب . مجـ ص + حـ . مجـ ع + ن ء ، (٢)

و مجـ م س = ا . مجـ س^{٢} + ب . مجـ س ص + حـ . مجـ ع س + ء . مجـ س (٣) ،

و مجـ م ص = ا . مجـ س ص + ب . مجـ ص^{٢} + حـ . مجـ ع س + ء . مجـ ص (٤) ،

و مجـ م ع = ا . مجـ س ع + ب . مجـ ص ع + حـ . مجـ ع^{٢} + ء . مجـ ع (٥) ،

من هذه المعادلات نستخرج قيم ا و ب و حـ و ء ؛ وبذلك نحصل على المعادلة (١) التى تربط المتغير م بالمتغيرات س و ص و ع . وهذه المعادلة هى معادلة « خط » العلاقة المتوسطة بين م والثلاثة المتغيرات مجتمعة .

---

(١) يلاحظ أن كيفية تركيب هذه المعادلات هى كما يأتى : نضرب المعادلة (١) فى معامل ء فيها ثم نجمعها لجميع الحالات تنتج المعادلة (٢) ؛ ثم نضرب (١) فى معامل ا أى فى س ، ثم نجمع فنحصل على المعادلة (٣) . ثم نضرب (١) فى معامل ب ، أى ص ونجمع فنحصل على (٤) . وأخيراً نضرب (١) فى معامل حـ ، أى فى معامل ع ، ونجمع فنحصل على (٥) وهذه الطريقة تتمشى مع الطريقة التى شرحناها سابقاً .

حساب الخطأ المعيـــارى لمعـادلة العلاقة المتوسطة

٢٧٠ — نحسب تشتت القيم المعطاة ، أو الخطأ المعيارى لقيم م ، المشاهدة حول هذا « الخط » كما فعلنا فى الارتباط العادى فى الباب السابق . لذلك نحسب انحرافات قيم م عن هذه المعادلة . فإذا فرضـــنا أن ح_1 هو خطأ القيمة م_1 انحرافها عن هذه المعادلة :

$$\therefore \text{ح}_1 = \text{ا}\,\text{س}_1 + \text{ب}\,\text{ص}_1 + \text{حـ}\,\text{ع}_1 + \text{ء} - \text{م}_1 \qquad (٦)$$

ونضرب طرفى هذه المعادلة فى الكميات م_1 و س_1 و ص_1 و ع_1 و ح_1 على التوالى .

$$\therefore \text{ح}_1\,\text{م}_1 = \text{ا} \,.\, \text{م}_1\text{س}_1 + \text{ب} \,.\, \text{م}_1\text{ص}_1 + \text{حـ} \,.\, \text{م}_1\text{ع}_1 + \text{ء} \,.\, \text{م}_1 - \text{م}_1^2 \qquad (٧) ،$$

$$\text{و}\quad \text{ح}_1\text{س}_1 = \text{ا} \,.\, \text{س}_1^2 + \text{ب} \,.\, \text{س}_1\text{ص}_1 + \text{حـ} \,.\, \text{ع}_1\text{س}_1 + \text{ء} \,.\, \text{س}_1 - \text{م}\,\text{س}_1 \qquad (٨) ،$$

$$\text{و}\quad \text{ح}_1\text{ص}_1 = \text{ا} \,.\, \text{س}_1\,\text{ص}_1 + \text{ب} \,.\, \text{ص}_1^2 + \text{حـ} \,.\, \text{ع}_1\text{ص}_1 + \text{ء} \,.\, \text{ص}_1 - \text{م}_1\text{ص}_1 \qquad (٩) ،$$

$$\text{و}\quad \text{ح}_1\text{ع}_1 = \text{ا} \,.\, \text{س}_1\,\text{ع}_1 + \text{ب} \,.\, \text{ص}_1\,\text{ع}_1 + \text{حـ}\,\text{ع}_1^2 + \text{ء} \,.\, \text{ع}_1 - \text{م}_1\text{ع}_1 \qquad (١٠) ،$$

$$\text{و}\quad \text{ح}_1^2 = \text{ا} \,.\, \text{ح}_1\text{س}_1 + \text{ب} \,.\, \text{ح}_1\text{ص}_1 + \text{حـ} \,.\, \text{ح}_1\text{ع}_1 + \text{ء}\,\text{ح}_1 - \text{م}_1\,\text{ح}_1 \qquad (١١) ،$$

ولو جمعنا كلا من هذه المعادلات بالنسبة إلى جميع قيم المتغير م

$$\therefore \text{مجـ}\,\text{ح} = \text{ا} \,.\, \text{مجـ}\,\text{س} + \text{ب} \,.\, \text{مجـ}\,\text{ص} + \text{حـ} \,.\, \text{مجـ}\,\text{ع} + \text{ن}\,\text{ء} - \text{مجـ}\,\text{م} ،$$

أنظر المعادلة (٢) أعلاه ؛

$$\text{و}\quad \text{مجـ}\,\text{ح}\,\text{م} = \text{ا} \,.\, \text{مجـ}\,\text{م}\,\text{س} + \text{ب}\,\text{مجـ}\,\text{م}\,\text{ص} + \text{حـ} \,.\, \text{مجـ}\,\text{م}\,\text{ع} + \text{ء} \,.\, \text{مجـ}\,\text{م} - \text{مجـ}\,\text{م}^2 \qquad (١٢) ،$$

و مج ع س = ا . مج س^{٢} + ب مج س ص + حـ . مج ع س

+ ء . مج س — مج م س

= صفراً   أنظر معادلة (٣) .

وبالمثل مج ع ص = مج ع ع = صفراً .

وبضرب طرفى المعادلة (٦) فى ع$_{١}$ وجمعها لكل القيم

∴ مج ع^{٢} = ا . مج ع س + ب . مج ع ص + حـ . مج ع ع

+ ء . مج ع — مج ع

= — مج م ع ،   (١٣) ،

ولو وضعنا مج ع^{٢} = ن سه$_{م}^{٢}$ ، ووضعنا ع$_{م}$ للانحراف المعيارى لقيم المتغير م ورمزنا للوسط الحسابى لها بالحرف مَ

∴ مج م^{٢} = ن ع$_{م}^{٢}$ + ن مَ^{٢}

∴ ن سه$_{م}^{٢}$ = — مج م ع

= مج م^{٢} — ا . م س — ب . مج م ص — ب . مج م ع

— ء . مج م .

٢٧١ — وقياساً على تعريف الارتباط بدلالة الخطأ المعيارى سه الذى استخدمناه فى الباب السابق ، يكون معامل الارتباط المتعدد بين م والمتغيرات س و ص و ع مجتمعة هو ر$_{م.س ص ع}$ ، حيث

معامل الارتباط المتعدد

ر$^{٢}_{م.س ص ع}$ = ١ — $\frac{\text{سه}_{م}^{٢}}{\text{ع}_{م}^{٢}}$

∴ ر$^{٢}_{م.س ص ع}$ = $\frac{\text{ا . مج م س + ب . مج م ص + حـ . مج م ع + ء . مج م — ن . مَ}^{٢}}{\text{ن . ع}_{م}^{٢}}$

وبذا نحصل على معامل الارتباط المتعـــدد المطلوب ؛ وهو يقيس العلاقة بين

المتغير م والمتغيرات الأخرى كلها مجتمعةً . أى أنه يقيس لنا درجة تأثير المتغيرات س و ص و ع مع بعضها فى كمية المحصول م .

تعريف معامل الارتباط الجزئى

٢٧٢ — بقى أمامنا أن نعالج المسألة الأخرى ، وهى قياس العلاقة بين المتغير م وأحد المتغيرات الأخرى ، على فرض بقاء المتغيرين الآخرين ثابتين . فنريد مثلا معرفة معامل الارتباط بين كمية المحصول وكمية السماد المستعمل ، على فرض أن كمية مياه الرى ودرجة الحرارة لا تتغيران . هذا المعامل يسمى[١] **معامل الارتباط الجزئى أو الصافى .**

طريقة استبعاد العوامل الغريبة

٢٧٣ — يمكننا إيجاد معامل الارتباط الجزئى بأن ننتقى من نتائج التجارب التى نجريها جميع الحالات التى فيها كمية مياه الرى ثابتة ، وكذلك درجة حرارة الجو ، ولا يتغير إلا كمية السماد المستعمل وكمية المحصول . ومن هذه الحالات نحسب معامل الارتباط الجزئى من واقع الأرقام التى تدل على قيم م و س فى هذه الحالات دون غيرها .

وهذا الإجراء سهل فى بعض المسائل ؛ والواقع أن هذا هو المتبع عادة فى التجارب الزراعية . فى هذه المسألة مثلا يمكننا تثبيت عاملى الرى ودرجة الحرارة بغاية السهولة بأن نأخذ حقول التجارب كلها فى منطقة زراعية واحدة ، فنضمن بذلك تساوى درجة الحرارة فى كل منها ؛ ونروى كل الحقول عدداً معيناً من المرات وفى نفس المواعيد وبنفس المقدار ، فنضمن بذلك تساوى كمية مياه الرى فى كل الحقول . ونغير كمية السماد الموضوع فى هذه الحقول ، ونلاحظ المحصول فى كل منها : وبذلك يمكننا استبعاد تأثير جميع العوامل ويبقى عامل السماد وحده

(١) بالانجليزية Coefficient of Partial or Net Correlation

الذى يؤثر فى كمية المحصول . ونضمن حينئذ أن كل تغيير فى كمية المحصول يكون ناشئاً عن تغير مناظر له فى كمية السماد . فنحسب معامل الارتباط بين م و س من واقع الأرقام التى نحصل عليها من هذه الحقول ؛ ويكون هذا المعامل هو معامل الارتباط الجزئى بين كمية المحصول وكمية السماد ، بفرض أن الرى ودرجة الحرارة لم يتغيرا .

استبعاد العوامل الغريبة يمكن فى العلوم الطبيعية ، ومستحيل فى المسائل الاجتماعية

٢٧٤ — وهذه الطريقة — طريقة عزل العامل الذى نريد دراسته واستبعاد العوامل الأخرى من المجال — هى الطريقة العلمية المتبعة فى أبحاث العلوم البحتة مثل الكيمياء والطبيعة وعلوم الحياة وغيرها . وهى سليمة طبعاً من الناحية المنطقية والنظرية ؛ ولا غبار عليها من الناحية العملية فهى سهلة وميسورة فى تلك العلوم ، ولو أن استبعاد العوامل الغريبة ومحو أثرها غالباً ما يكلف الباحث مجهوداً كبيراً ونفقات طائلة . ولكن هذه الطريقة غير ميسورة بل مستحيلة فى دراسة الظواهر الاقتصادية والاجتماعية ، حيث إن هذه الظواهر تكون دائماً متأثرة بعوامل شتى ، متشابكة مع بعضها لا يمكن عزلها أو استبعاد بعضها . ولذلك يتعين على الباحث الاقتصادى أن يفكر فى استنباط طريقة أخرى لا تعتمد على استبعاد أثر العوامل الغريبة بتثبيتها أثناء إجراء التجربة .

حاصل ضرب معاملى الانحدار

٢٧٥ — رأينا فى بند ٢٤٤ (صحيفة ٢٧٤) أن حاصل ضرب معامل انحدار ص على س ، فى معامل انحدار س على ص يساوى $ر^٢$ ، أى مربع معامل الارتباط بينهما . وقياساً على هذا يمكننا إيجاد معامل الارتباط الجزئى الذى نحن بصدده .

نفرض أن لدينا المتغير م والمتغيرات الثلاثة س و ص و ع كما فى المسألة السابقة . ولتكن معادلة « خط » العلاقة المتوسطة بين هذه المتغيرات هى :

$$م = ا_{٢١} س + ب_{٣١} ص + حـ_{٤١} ع + ء_{١} \text{ ، } \ldots \text{ (١) .}$$

حيث $ا_{٢١}$ ، $ب_{٣١}$ ، $حـ_{٤١}$ ، $ء_{١}$ كميات ثابتة معينة .

هذه المعادلة هى فى الواقع معادلة انحـدار بين م والمتغيرات الأخرى ، كما سبق أن قلنا من قبل ؛ فلو ثبتنا ص و ع وجعلنا س فقط تتغير و م تبعاً لها ، كما هو مفروض فى تعريف الارتباط الجزئى ، آلت هذه المعادلة إلى الصورة :

$$م = ا_{٢١} س + \text{كمية ثابتة ، . . .} \qquad \text{(٢) .}$$

وهذه هى معادلة خط العلاقة المتوسطة بين م و س ، أى خط انحدار م على س . وتكون $ا_{٢١}$ إذن هى معامل انحـدار المتغير م على المتغير س ، باعتبار أن س متغير مستقل و م متغير تابع له .

نوفق معادلة أخرى تعبر عن المتغير س بدلالة المتغيرات م ، ص ، ع ، من واقع القيم المعطـاة لنا من التجربة ؛ وذلك بنفس الطريقة التى اتبعنـاها فى توفيق المعادلة (١) بند ٢٦٩ . ولتكن هذه المعادلة الجديدة هى :

$$س = ا_{١٢} م + ب_{٣٢} ص + حـ_{٤٢} ع + ء_{٢} \ldots \qquad \text{(٣) ؛}$$

حيث كل من $ا_{١٢}$ ، $ب_{٣٢}$ ، $حـ_{٤٢}$ ، $ء_{٢}$ كمية ثابتة مستقلة عن المتغيرات س ، م ، ص ، ع .

إذا ثبتنا ص ، ع وجعلنا م ، س فقط تتغيران ، تؤول هذه المعادلة إلى معادلة خط انحدار س على م ، وتكون $ا_{١٢}$ هى أيضاً معامل انحدار س على م .

وهكذا نحصل على معاملى الانحدار بين المتغيرين م و س ، كل على الآخر باعتبار المتغيرين الآخرين ص و ع ثابتين . وهذان المعاملان هما $ا_{٢١}$ و $ا_{١٢}$ ،

ويكون معامل الارتباط الجزئى بين م و س ، بفرض ص و ع ثابتين ، هو[١] $ر_{١٢.٣٤}$ حيث :

$$ر^{٢}_{١٢.٣٤} = ا_{١٢} \, . \, ا_{٢١} \quad ، \quad \dots \quad \dots \qquad (٤)$$

وكذلك لإيجاد معامل الارتباط الجزئى بين م و ص بفرض س و ع ثابتين ، نوفق معادلة تعطى ص بدلالة م و س و ع ولتكن هى :

$$ص = ا_{٣٢} س + ب_{٣١} م + حـ_{٣٤} ع + ء_{٣} \, . \qquad (٥) \, .$$

ويكون إذن معامل الارتباط الجزئى بين م و ص هو $ر_{١٣.٢٤}$ حيث :

$$ر^{٢}_{١٣.٢٤} = ب_{١٣} \, . \, ب_{٣١} \qquad (٦) \, .$$

وأخيراً نوفق معادلة بين ع والمتغيرات الأخرى ولتكن هى :

$$ع = ا_{٤٢} س + ب_{٤٣} ص + حـ_{٤١} م + ء_{٤} \, . \qquad (٧) \, ،$$

ويكون معامل الارتباط الجزئى بين م و ع هو $ر_{١٤.٢٣}$، حيث :

$$ر^{٢}_{١٤.٢٣} = حـ_{١٤} \, . \, حـ_{٤١} \qquad (٨) \, .$$

استنباط المعاملات ر فى حالة خاصة

٢٧٦ — نأخذ حالة خاصة ونوجد معامل الارتباط الجزئى لتوضيح الخطوات :

لنفرض للسهولة أن لدينا متغيرين مستقلين فقط وهما س و ص ، ومتغيراً ثالثاً تابعاً لهما وهو م . وأن لدينا ن من القيم لكل من هذه المتغيرات الثلاثة ، وهى على التوالى :

---

(١) يدل الرمز المرقوم $ر_{١٢.٣٤}$ على معامل الارتباط الجزئى بين المتغيرين الأول والثانى (م و س) بفرض أن الثالث والرابع (ص و ع) ثابتان ، وهكذا .

$$س_١ \quad س_٢ \quad س_٣ \quad \cdots \quad \cdots \quad س_ن ،$$

$$و \quad ص_١ \quad ص_٢ \quad ص_٣ \quad \cdots \quad \cdots \quad ص_ن ،$$

$$و \quad م_١ \quad م_٢ \quad م_٣ \quad \cdots \quad \cdots \quad م_ن ،$$

ولنفرض ، للسهولة أيضاً ، أن هذه القيم مقيسة ابتـداء من الوسط الحسابى لكل منها ؛ أى أن

$$مجـ س = مجـ ص = مجـ م = صفراً ؛$$

$$وكذلك \quad \bar{س} = \bar{ص} = \bar{م} = صفراً ،$$

∴ تـكون الانحرافات المعيارية لهذه المتغيرات هى $ع_١$ و $ع_٢$ و $ع_٣$ حيث

$$ن ع_١^٢ = مجـ م^٢ ، ن ع_٢^٢ = مجـ س^٢ ، ن ع_٣^٢ = مجـ ص^٢ .$$

ويكون أيضاً معامل الارتباط بين م و س مثلا يساوى $ر_{٢١}$ ، حيث

$$ر_{٢١} = \frac{مجـ م س}{ن ع_١ ع_٢} \quad \cdots \quad \cdots \quad معادلة (١) \ بند ٢٠٣ ،$$

نوفق معادلة بين م و س و ص من الدرجة الأولى ، من واقع هذه القيم المعطاة . ولتـكن هذه المعادلة هى :

$$م = ا س + ب ص + حـ \quad \cdots \quad \cdots \quad (١) ،$$

حيث كل من ا و ب و حـ كمية ثابتة ، تستخرج من المعادلات المعروفة فى توفيق المنحنيات وهى :

$$مجـ م = ا . مجـ س + ب مجـ ص + ن حـ \quad (٢) ،$$

$$و مجـ س م = ا . مجـ س^٢ + ب مجـ س ص + حـ . مجـ س \quad (٣) ،$$

$$و مجـ ص م = ا . مجـ س ص + ب مجـ ص^٢ + حـ . مجـ ص \quad (٣) ،$$

وبوضع مح س = مح ص = مح م = . ، ينتج أن حـ = . ؛

وبالتعويض عن مح س^{٢} و مح ص^{٢} و مح م^{٢} بقيمها ، ووضع :

$$\text{مح م س} = \text{ن}\ \text{ر}_{٢١} \cdot \text{ع}_{١}\,\text{ع}_{٢}$$

$$\text{و} \quad \text{مح م ص} = \text{ن}\ \text{ر}_{٣١} \cdot \text{ع}_{١}\,\text{ع}_{٣}$$

$$\text{و} \quad \text{مح س ص} = \text{ن}\ \text{ر}_{٣٢} \cdot \text{ع}_{٢}\,\text{ع}_{٣}$$

$$\therefore \quad \text{ر}_{٢١} \cdot \text{ع}_{١} = \text{ا} \cdot \text{ع}_{٢} + \text{ب} \cdot \text{ر}_{٣٢}\,\text{ع}_{٣} \quad (٥)،$$

$$\text{و} \quad \text{ر}_{٣١} \cdot \text{ع}_{١} = \text{ا} \cdot \text{ر}_{٣٢}\,\text{ع}_{٢} + \text{ب} \cdot \text{ع}_{٣} \quad (٦)، \ .$$

$$\therefore \quad \text{ع}_{١}\,(\text{ر}_{٢١} - \text{ر}_{٣١} \cdot \text{ر}_{٣٢}) = \text{ا}\,(١ - \text{ر}_{٣٢}^{٢})\,\text{ع}_{٢}$$

$$\text{و} \quad \text{ع}_{١}\,(\text{ر}_{٣١} - \text{ر}_{٢١} \cdot \text{ر}_{٣٢}) = \text{ب}\,(١ - \text{ر}_{٣٢}^{٢})\,\text{ع}_{٣}$$

∴ معادلة انحدار م على س و ص ، هى

$$\frac{\text{م}}{\text{ع}_{١}} = \frac{\text{ر}_{٢١} - \text{ر}_{٣١} \cdot \text{ر}_{٣٢}}{١ \quad \text{ر}_{٣٢}^{٢}} \cdot \frac{\text{س}}{\text{ع}_{٢}} + \frac{\text{ر}_{٣١} - \text{ر}_{٣١} \cdot \text{ر}_{٣٢}}{١ - \text{ر}_{٣٢}^{٢}} \cdot \frac{\text{ص}}{\text{ع}_{٣}} ، \ldots (٧) \ .$$

وبمثل ذلك نوفق معادلة انحدار س على م وص ، وهى :

$$\frac{\text{س}}{\text{ع}_{٢}} = \frac{\text{ر}_{٢١} - \text{ر}_{٣١} \cdot \text{ر}_{٣٢}}{١ \quad \text{ر}_{٣١}^{٢}} \cdot \frac{\text{م}}{\text{ع}_{١}} + \frac{\text{ر}_{٣٠} \quad \text{ر}_{٢١} \cdot \text{ر}_{٣٢}}{١ - \text{ر}_{٣١}^{٢}} \cdot \frac{\text{ص}}{\text{ع}_{٣}} \ldots (٨) \ .$$

ومن هاتين المعادلتين نرى أن معاملى انحدار م على س و س على م هما

على التوالى :

$$\frac{\text{ع}_{١}}{\text{ع}_{٢}} \cdot \frac{\text{ر}_{٢١} - \text{ر}_{٣١} \cdot \text{ر}_{٣٢}}{١ - \text{ر}_{٣٢}^{٢}} \quad \text{و} \quad \frac{\text{ع}_{٢}}{\text{ع}_{١}} \cdot \frac{\text{ر}_{٢١} - \text{ر}_{٣١}\ \text{ر}_{٣٢}}{١ - \text{ر}_{٣١}^{٢}} \ .$$

وينتج إذن أن معامل الارتباط الجزئى بين م و س هو الوسط الهندسى

لهذين المعاملين . . . . (٩) ،

$$\therefore \quad \text{ر}_{٢١ \cdot ٣} = \frac{\text{ر}_{٢١} - \text{ر}_{٣١}\ \text{ر}_{٣٢}}{\sqrt{(١ - \text{ر}_{٣١}^{٢})\,(١ - \text{ر}_{٣٢}^{٢})}}$$

حيث $\text{ر}_{٢١}$ ، $\text{ر}_{٣١}$ و $\text{ر}_{٣٢}$ هى معاملات الارتباط العادية بين م و س ، م و ص ، س و ص على التوالى .

وهكذا أمكننا الحصول على معامل الارتباط الجزئى بين متغيرين (م و س) بفرض ثبات الثالث (ص) ، بدلالة معاملات الارتباط العادية بين أزواج بسيطة من المتغيرات .

معامل الارتباط الجزئى من رتب على

٢٧٧ — وإذا كان لدينا عدد من المتغيرات أكبر من ثلاثة ، وأردنا معامل الارتباط الجزئى بين اثنين منها مع تثبيت المتغيرات الباقية ، نستخرج أولا معاملات الارتباط العادية بين أزواج بسيطة من المتغيرات . ومن هـذه نستنبط معامل الارتباط بين متغيرين مع تثبيث واحد غيرهما ، باستخدام العلاقة (٩) من البند السابق . نسمى هذا معاملاً من الرتبة الأولى مثلا $\text{ر}_{٣٠٢١}$ ، حيث يوجد متغير واحد ثابت ، وهو الثالث المعبر عنه بالرقم ٣ بعد النقطة .

وبواسـطة معاملات الرتبة الأولى نستنبط معاملات الرتبة الثانية ، وهى معاملات الارتباط الجزئى مع تثبيت متغيرين . مثلا $\text{ر}_{٤٣٠٢١}$ هو معامل الارتباط الجزئى بين المتغيرين الأول والثانى ، بفرض تثبيت الثالث والرابع .

ومن هـذا نستنبط معاملات الرتبة الثالثة وهكذا . والعلاقة بين معاملات الرتب المتتالية يمكن وضعها فى الصورة الآتية (١) :

$$\text{ر}_{٤٣٠٢١} = \frac{\text{ر}_{٣٠٢١} - \text{ر}_{٣٠٤١} \cdot \text{ر}_{٣٠٤٢}}{\sqrt{(١ - \text{ر}^{٢}_{٣٠٤١})(١ - \text{ر}^{٢}_{٣٠٤٢})}} \text{ ،}$$

$$\text{ر}_{٥٤٣٠٢١} = \frac{\text{ر}_{٤٣٠٢١} - \text{ر}_{٤٣٠٥١} \cdot \text{ر}_{٤٣٠٥٢}}{\sqrt{(١ - \text{ر}^{٢}_{٤٣٠٥١})(١ - \text{ر}^{٢}_{٤٣٠٥٢})}} \text{ ،}$$

(١) أنظر كتاب G.U.Yule, *Introduction to the Theory of Statistics* (1937) p. 262
وكتاب Mills, C.F. *Statistical Methods*, (1924) p. 508

وعلى العموم يكون :

$$\text{ر}_{\text{١٢٠٣٤٥}\cdots\text{ن}} = \frac{\text{ر}_{\text{١٢٠٣٤}\cdots(\text{ن}-\text{١})} - \text{ر}_{\text{١ن٠٣٤٥}\cdots(\text{ن}-\text{١})} \cdot \text{ر}_{\text{٢ن٠٣٤}\cdots(\text{ن}-\text{١})}}{\sqrt{\left(\text{١} - \text{ر}^{\text{٢}}_{\text{١ن٠٣٤}\cdots(\text{ن}-\text{١})}\right)\left(\text{١} - \text{ر}^{\text{٢}}_{\text{٢ن٠٣٤}\cdots(\text{ن}-\text{١})}\right)}}$$

فى حالة علاقة غيرخطية نأخذ متغيراً جديداً

٢٧٨ — يلاحظ أننا فرضنا فى هذا الباب أن العلاقة بين المتغير م والمتغيرات س ، ص ، ع علاقة خطية ، أى أنها من الدرجة الأولى . ومن الممكن تعميم هذه الطريقة لتشمل الحالات الأخرى التى تكون فيها العلاقة بين م و بعض المتغيرات ، أو كلها ، غير خطية . فلنفرض مثلا أن المتغير س يظهر فى شكل $\text{س}^{\text{٣}}$ فى معادلة العلاقة المتوسطة [المعادلة رقم (١) بند ٢٦٩] ؛ أى أنه يظهر فى الدرجة الثالثة بدلا من الدرجة الأولى . فى هذه الحالة يمكننا أن نستبدل المتغير $\text{س}^{\text{٣}}$ بمتغير جديد مثل ل ؛ وتكون قيم ل حينئذ تساوى تكعيب قيم س التى حصلنا عليها من التجربة . ونوفق المعادلة بين المتغير التابع م والمتغيرات المستقلة ل ، ص ، ع بنفس الطريقة المستخدمة فى توفيق المعادلة العادية .

ويجوز طبعاً أن تكون العلاقة بين المتغير الجديد ل والمتغير المستبدل ، بشكل آخر غير ل = $\text{س}^{\text{٣}}$ ، مثلا ل = $\text{س}^{\text{٤٫٥}}$ ، أو لو س ، أو $\text{ا}^{\text{س}}$ ، أو . . . ؛ ولكن البحث فى هذه المسائل يخرجنا عن نطاق هذا الكتاب ، لأنه يستلزم إلماماً ببعض النظريات الرياضية الصعبة والمعقدة التى لا محل لها هنا .

٢٧٩ — وقد فرضنا ضمناً فى هذا البحث أيضاً ، أن المتغيرات التى نبحثها ذات توزيع تكرارى متماثل . وهذا أيضاً فرض لا يتحقق إلا فى حالات خاصة فقط ، وإنما اخترناه هنا للسهولة .

المـــراجع

Bowley, A.L., : *Elements of Statistics*, Part II, Chapter VIII.

Mills, F.C., : *Statistical Methods*, Chapter XV.

Rietz, H., : *Handbook of Mathematical Statistics*, Chapter IX.

Whittaker and Robinson: *Calculus of Observations*, Chapter XII.

Yule, G.U., : *Introduction to the Theory of Statistics*, Chapter XIV

الباب الحادي عشر

الأرقام القياسية

تعريف الرقم القياسي

٢٨٠ — الرقم القياسي هو عبارة عن رقم نسبي ، أو ملخص لعدة أرقام نسبية ، ينشأ لبيان وقياس الحركة أو التغير فى أى ظاهرة معينة بالنسبة إلى أساس معين . ولتركيب الرقم القياسي لأى ظاهرة ، نكون نسبة مئوية بين القيمة المقارنة لهذه الظاهرة والقيمة الأخرى لها ، المعتبرة أساساً للمقارنة . فالرقم القياسي مثلا لسعر القمح هذا العام بالنسبة إلى سعره فى سنة ١٩٢٠ كأساس ، يساوى خارج قسمة السعر الحالى على السعر فى سنة ١٩٢٠ (سنة الأساس) مضروباً فى العدد ١٠٠. وكذلك الرقم القياسي لسعر القطن فى بورصة مينا البصل هذا الأسبوع بالنسبة إلى سعره فى بورصة ليفربول كأساس ، يساوى خارج قسمة السعر فى مينا البصل على السعر فى ليفربول ، مضروباً فى العدد ١٠٠ أيضاً . وهكذا .

استخدام الأرقام القياسية فى الاقتصاد

٢٨١ — والأرقام القياسية كثيرة الاستعمال فى الأبحاث الاقتصادية جميعها ، وفى غيرها من المسائل العلمية أيضاً . وهى أداة نافعة جداً فى تصوير التغيرات التى تطرأ على الظواهر الاقتصادية المختلفة ، خصوصاً تلك الظواهر المركبة من عدة عوامل متغيرة فى وقت واحد . ومثال ذلك المستوى العمومى للأسعار فهو عبارة عن ملخص عن أسعار جميع السلع ؛ وكل سلعة تحيط بها ظروف خاصة بها تعمل على تغيير أسعارها ، وظروف مشتركة تجعلها تتبع الحركة العامة للسوق . فالرقم القياسي لمستوى الأسعار يكوّن لنا فكرة ملخصة ودقيقة وواضحة عن

تغيرات أسعار هذه السلع فى مجموعها . وبدون هـذا الرقم لا يمكننا دراسة الحالة العامة للسوق ، ومستوى الأسعار وتأثيرهما فى الحالة الاقتصادية للبلد .

والحقيقة أن مسائل الأسعار هى أهم النواحى التى نستخدم فيها الأرقام القياسية ؛ إلا أن هناك عدة مسائل اقتصادية واجتماعية نحتاج فى دراستها إلى استخدام الأرقام القياسية . ومثال ذلك إنشاء رقم قياسى للنشاط الصناعى أو التجارى ، ورقم قياسى للأجور ، ورقم قياسى لمستوى المعيشة ونفقتها ، وغير ذلك من المسائل الهامة . ولكننا سنقصر الكلام على تركيب الأرقام القياسية للأسعار ؛ والمفهوم طبعاً أن الطرق المستعملة فى مسائل الأسعار تنطبق بذاتها ، مع تغيير بسيط فى معنى الرموز ، على المسائل الأخرى مثل الأجور أو الإنتاج ونحوهما.

تعيين سنة الأساس

٢٨٢ — يمكننا تركيب أرقام قياسية — لمستوى الأسعار مثلا — على عدة صور ؛ وسنشرح الآن بعض الطرق المختلفة لتركيب الأرقام القياسية . ويجب أن نلاحظ فى مبدأ الأمر أن النتائج التى نحصل عليها بهذه الطرق لن تكون متطابقة ، ولو أنها مشتقة من نفس البيانات .

وعند إنشاء أى رقم قياسى لا بد أن نتفق على الأساس الذى سنتخذه لتركيب هذا الرقم القياسى ، فنأخذ مثلا (١) سنة معينة (أو فترة أخرى أكبر أو أصغر من سنة) ونعتبرها أساساً . وهذه نسميها **سنة الأساس** أو **القاعدة** . وفى العادة تكون سنة (أو فترة) الأساس سابقة للسنة التى نريد مقارنتها ، ولكن أحياناً يكون المطلوب رقم قياسى للأسعار فى سنة ١٩٣٥ مثلا بالنسبة إلى سنة ١٩٣٨ كأساس . ولذلك سنتكلم عن السنة (أو البلد) الأساسية والسنة المقارنة ، وعن الأسعار الأساسية والأسعار المقارنة ، وهكذا .

---

(١) يصح طبعاً أن يكون مكاناً معيناً كما لو أردنا تكوين الرقم القياسى لمستوى الأسعار فى الأسكندرية بالنسبة إلى القاهرة كأساس . هنا الأساس هو القاهرة ، ولا ذكر للزمن أو لسنة الأساس التى تسمى بالإنجليزية Base Year

٢٨٣ — لنفرض أن لدينا عدة سلع يتركب منها الرقم القياسى ، ولتكن الأسعار الأساسية لهذه السلع هى : الرموز المستعملة

$\text{ع}_{٠}$ ، $\text{ع}'_{٠}$ ، $\text{ع}''_{٠}$ ، $\text{ع}'''_{٠}$ ، . . . بعدد السلع الموجودة .

ولتكن الكميات الأساسية لهذه السلع هى على التوالى :

$\text{ك}_{٠}$ ، $\text{ك}'_{٠}$ ، $\text{ك}''_{٠}$ ، $\text{ك}'''_{٠}$ ، . . .

ولنفرض أن الأسعار والكميات المقارنة لهذه السلع هى على التوالى :

$\text{ع}_{١}$ ، $\text{ع}'_{١}$ ، $\text{ع}''_{١}$ ، $\text{ع}'''_{١}$ . . .

و $\text{ك}_{١}$ ، $\text{ك}'_{١}$ ، $\text{ك}''_{١}$ ، $\text{ك}'''_{١}$ . . .

إذا قسمنا السعر المقارن لأى سلعة على سعرها الأساسى ، وضربنا الناتج فى ١٠٠ ، حصلنا على ما نسميه[١] **منسوب السعر** لهذه السلعة . وإذا رمزنا للمنسوب بالحرف س يكون

$$\text{س} = \frac{\text{ع}_{١}}{\text{ع}_{٠}} \times ١٠٠ \quad \text{و} \quad \text{س}' = \frac{\text{ع}'_{١}}{\text{ع}'_{٠}} \times ١٠٠$$ ، وهكذا مع باقى السلع .

ونشرح الآن بعض الأسس التى يمكن أن يبنى عليها تركيب أرقام قياسية للأسعار بالمعنى المقصود فى تعريف الرقم القياسى . ولن نتعرض هنا إلى المفاضلة بين هذه الأسس ، ولا إلى كيفية اختيار أصلحها .

---

(١) اسمه بالإنجليزية Price Relative . بعض الناس يستعمل كلمة « السعر النسبى » . أنظر الإحصاء السنوى العام لسنة ١٩٣٥ — ١٩٣٦ صحيفة ٤٩٠ ، ولكن هذه الكلمة فى رأيى لا تؤدى المعنى هنا . لأن هذا ليس سعراً ، بل نسبة .

الرقم التجميعي البسيط

٢٨٤ — الرقم التجميعي البسيط للأسعار [١] . ولتركيب هـذا الرقم نقسم حاصل جمع الأسعار المقارنة على حاصل جمع الأسعار الأساسية لكل السلع كما هي وبدون مفاضلة بينها أو ترجيح البعض على البعض . وهذا الرقم هو إذن

$$\frac{ع_{١} + عَ_{١} + عً_{١} + \cdots}{ع_{.} + عَ_{.} + عً_{.} + \cdots} = \frac{\text{مجـ}\ ع_{١}}{\text{مجـ}\ ع_{.}} \text{ ، } \cdots \cdots \cdots \cdots \ (١)$$

لنأخذ مثلا سلع القطن والقمح والفول والشعير فى سنتى ١٩٣١ و ١٩٣٥ ، ونحسب الرقم القياسي لأسعار هذه المحاصيل فى سنة ١٩٣٥ بالنسبة إلى سنة ١٩٣١ كأساس :

جدول ٤٨ — أسعار [٢] محاصيل القطن والقمح والفول والشعير

فى سنتى ١٩٣١ و ١٩٣٥

السـنة	القطن بالقنطار	القمح بالأردب	الفول بالأردب	الشعير بالأردب
	م . ح	م . ح	م . ح	م . ح
١٩٣١	٣,٦٢٥	١,٢١٩	١,٣٩٩	٠,٦٤٣
١٩٣٥	٣,٩٦٠	١,٥٣١	١,٣٤٣	٠,٨٠٢

$$\therefore \frac{\text{مجـ}\ ع_{١}}{\text{مجـ}\ ع_{.}} = \frac{٣,٩٩٠ + ١,٥٣١ + ١,٣٤٣ + ٠,٨٠٢}{٣,٦٢٥ + ١,٢١٩ + ١,٣٩٩ + ٠,٦٤٣}$$

$$= \frac{٧,٦٣٦}{٦,٨٨٦} .$$

---

(١) بالانجليزية Simple Aggregative Index

(٢) أنظر الإحصاء السنوى العام ١٩٣٥ — ١٩٣٦ صفحتى ٤٠٩ و ٤١٠ هذه الأرقام خاصة بمحاصيل زراعة الأملاك الأميرية .

$$\therefore \quad \frac{\text{مجـ ع}_{1}}{\text{مجـ ع}_{0}} \times ١٠٠ = ١١٠٫٨٩ \text{ ،}$$

أى أن مستوى أسعار هذه المحاصيل ارتفع ١٠٫٨ ٪ فى سنة ١٩٣٥ عما كان عليه فى سنة ١٩٣١.

الرقم التجميعي سهل الحساب ولـكنه لا يعطى للسلع أهميتها الحقيقية

٢٨٥ — وهذا الرقم التجميعى البسيط هو أسهل الأرقام القياسية عملاً وتركيباً ، إذ نضع الأسعار على علاتها وبدون تحريف أو تعديل مهما كانت . ولـكن هذه البساطة والسهولة هما فى نفس الوقت عيب يؤخذ عليه . لأن جميع السلع هنا تعامل نفس المعاملة ، بدون ترجيح أو تعزيز بعضها بما يتناسب وأهميتها بجانب غيرها .

هذا فضلاً عن أن اختلاف الوحدات المستعملة فى تسعيرات السلع المختلفة ، وما ينشأ عنه من تكبير أو تصغير السعر $ع_{0}$ أو $ع_{1}$ ، يعطى بعض السلع أهمية مفتعلة ليست لها . فمثلاً إذا كانت ع سعر الخبز ، وهو ٣٠ مليما للأقة ، وكانت $\acute{ع}$ سعر الفحم ، وهو ١٨٠ قرشاً للطن ، وجدنا أن ع صغيرة جداً بالنسبة إلى $\acute{ع}$ ؛ وبذلك نعطى لسعر الفحم وتغيراته وزناً وأهمية أكبر من سعر الخبز ، الذى هو فى الحقيقة أولى بهذه الأهمية .

الرقم التجميعى المرجح

٢٨٦ — ولتصحيح هذا العيب فى الرقم التجميعى البسيط ، نرجح أسعار السلع المختلفة بأوزان تتناسب وأهمية هذه السلع . فنستخدم الـكميات المنتجة (أو المستهلـكة) من هذه السلع كأوزان . ويصح أن نستعمل الـكميات الأساسية أو الـكميات المقـارنة كأوزان ، وبذلك نحصل على صـيغتين للرقم التجميعى المرجح ، وهما :

ا — الرقم التجميعى المرجح بالـكميات الأساسية ، وهو :

$$\frac{ع_{1}ك_{0} + \acute{ع}_{1}\acute{ك}_{0} + \ddot{ع}_{1}\ddot{ك}_{0} + \cdots}{ع_{0}ك_{0} + \acute{ع}_{0}\acute{ك}_{0} + \ddot{ع}_{0}\ddot{ك}_{0} + \cdots} = \frac{\text{مجـ ع}_{1}ك_{0}}{\text{مجـ ع}_{0}ك_{0}} \quad \cdots (١٢) ؛$$

٢ — الرقم التجميعى المرجح بالكميات المقارنة ، وهو :

$$\frac{\text{ع}_{١}\,\text{ك}_{١} + \text{ع}'_{١}\,\text{ك}'_{١} + \text{ع}''_{١}\,\text{ك}''_{١} + \cdots}{\text{ع}_{٠}\,\text{ك}_{١} + \text{ع}'_{٠}\,\text{ك}'_{١} + \text{ع}''_{٠}\,\text{ك}''_{١} + \cdots} = \frac{\text{مج ع}_{١}\,\text{ك}_{١}}{\text{مج ع}_{٠}\,\text{ك}_{١}} \quad \cdots \ (\text{٢ ب}) \ .$$

ويلاحظ فى كلتا الحالتين ا و ب أن الأوزان المستعملة فى البسط هى نفسها المستعملة فى المقام .

ولتطبيق هذين الرقمين نأخذ كميات هـذه المحاصيل الأربعة (فى أراضى مصلحة الأملاك الأميرية أيضاً) وهى المبينة فى الجدول الآتى :

جدول ٤٩ — كميات محاصيل القطن والقمح والشعير

فى سنتى ١٩٣١ و ١٩٣٥

السنة	القطن	القمح	الفول	الشعير
	قناطير	أرادب	أرادب	أرادب
١٩٣١	٨٢٥٨	٣٨٠٩	٢٥٨٩	٣١٧٩
١٩٣٥	٩٥٥٩	٤٥٧٦	٣٠٠٦	٤١٩٩

هنا نجـد

$$\frac{\text{مج ع}_{١}\,\text{ك}_{٠}}{\text{مج ع}_{٠}\,\text{ك}_{٠}} = \frac{٣٫٩٦٠ \times ٨٢٥٨ + ١٫٥٣١ \times ٣٨٠٩ + ١٫٣٤٣ \times ٢٥٨٩ + ٫٨٠٢ \times ٣١٧٩}{٣٫٦٢٥ \times ٨٢٥٨ + ١٫٢١٩ \times ٣٨٠٩ + ١٫٣٩٩ \times ٢٥٨٩ + ٫٦٤٣ \times ٣١٧٩}$$

$$= \frac{٤٤٥٥٩٫٨٤٤}{٤٠٢٤٤٫٥٢٩}$$

$$\therefore \quad ١٠٠ \times \frac{\text{مج ع}_{١}\,\text{ك}_{٠}}{\text{مج ع}_{٠}\,\text{ك}_{٠}} = ١١٠٫٧٢ \ .$$

وهذا الرقم القياسى يساوى ، بالتقريب ، الرقم الذى حصلنا عليه بدون أوزان .

وبالمثل نجد أن الرقم القياسى بالأوزان ب هو :

$$\text{١٠٠} \times \frac{\text{مجـ } ع_{1} \cdot ك_{1}}{\text{مجـ } ع_{0} \cdot ك_{1}} = \text{١١٠٫٨٨} \;\text{؛}$$

وهى أيضاً نتيجة مخالفة لـكل من النتيجتين السـابقتين ، ولو أن الفرق صغير . ولـكنه يصح أن يكون أ كبر من ذلك إذا زاد عدد السلع ، أو إذا كان الاختلاف بين الكميات الأساسية والكميات المقارنة أ كبر مما هو فى هذا المثال .

ويلاحظ أن الرقم البسيط يقع فى الوسط بين الرقمين (٢ ا) و (٢ ب) .

الرقم القياسى الأمثل

٢٨٧ — يصح أن نجمع بين نظام الأوزان المستعمل فى الرقم (٢ ا) والنظام المستعمل فى (٢ ب) ، فنحصل على رقم جديد يكون أ كثر اعتدالا وأقل تحيزاً من كل منهما . فمثلا لو أخذنا الوسط الهندسى لهذين الرقمين نحصل على ما يسميه الأستاذ إرفنج فيشر **الرقم القياسى الأمثل** [1] ، وهو

$$\sqrt{\frac{\text{مجـ } ع_{1} \cdot ك_{0}}{\text{مجـ } ع_{0} \cdot ك_{0}} \times \frac{\text{مجـ } ع_{1} \cdot ك_{1}}{\text{مجـ } ع_{0} \cdot ك_{1}}} \quad \dots \quad \dots \qquad (\text{٣})$$

وسنرى فيما بعد أن هذا الرقم يستحق هذه التسمية حقيقة ، حيث تجتمع فيه كل الصـفات المطلوبة فى الرقم القياسى الصحيح ، ويخلو من العيوب التى تشوب الأرقام القياســـية الأخرى . وهذا ما يجعله فى المـكان الأول بين جميع الأرقام القياسية والمثل الأعلى لها .

ولو حسبنا الرقم القياسى فى المثال الذى بأيدينا على أساس هذه المعادلة نجد أنه يساوى $\sqrt{\text{١٢٢٧٦٫٦٣٣٦}} = \text{١١٠٫٨٠٠}$

رقم آخر يجمع بين الرقمين ٢ ا ، ٢ ب

٢٨٨ — ويصح أن نجمع بين نظامى الأوزان فى (٢ ا) و (٢ ب) فى صورة أخرى ، فنأخذ الوزن لـكل سلعة يساوى الوسط الحسابى (أو أى وسط آخر) للـكميتين $ك_{0}$ و $ك_{1}$ ، فنحصل على الصورة الآتية :

---

(١) يعرف هذا الرقم باسم Irving Fisher's "Ideal Index."

$$\frac{\text{مح ع}_1 \,(\text{ك}_. + \text{ك}_1)}{\text{مح ع}_. \,(\text{ك}_. + \text{ك}_1)} \qquad (٣\ ا)\ .$$

وهذه الصيغة (٣ ا) أسهل فى العمل الحسابى من الصيغة (٣) ، ولكن هذه الأخيرة تفضلها من عدة وجوه نشرحها فى مناسبة أخرى . و إذا أخذنا الوسط الهندسى بدلا من الحسابى نحصل على الصورة الآتية وهى أصعب فى الحساب

$$\frac{\text{مح ع}_1 \sqrt{\text{ك}_. \,.\, \text{ك}_1}}{\text{مح ع}_. \sqrt{\text{ك}_. \,.\, \text{ك}_1}} \qquad (٣\ ب)\ .$$

المتوسط البسيط للمناسيب

٢٨٩ — يمكننا تركيب أرقام قياسية أيضاً باستخدام مناسيب الأسعار بدل الأسعار نفسها ، فنحسب الوسط الحسابى ، أو الهندسى أو التوافقى (بسيطاً أو مرجحاً بأوزان مناسبة) لمناسيب الأسعار التى سبق أن عرفناها فى بند ٢٨٣ ، فإذا فرضنا أن هذه المناسيب للسلع المختلفة ، وعددها ن ، هى :

$$\text{س} \ ، \ \text{س}' \ ، \ \text{س}'' \ ، \ \ldots$$

يكون الوسط الحسابى البسيط لهذه المناسيب هو

$$\frac{\text{س} + \text{س}' + \text{س}'' + \cdots}{\text{ن}} = \frac{\text{مح س}}{\text{ن}} \qquad (٤)\ .$$

والوسط الهندسى البسيط للمناسيب

$$= \sqrt[\text{ن}]{\text{س} \times \text{س}' \times \text{س}'' \times \cdots} \qquad (٥)\ .$$

والوسط التوافقى المناسيب هو ف مثلا ، حيث

$$\frac{\text{ن}}{\text{ف}} = \frac{1}{\text{س}} + \frac{1}{\text{س}'} + \frac{1}{\text{س}''} + \ldots \qquad (٦)\ .$$

وفى المثال الذى بأيدينا نرى أن هذه المناسيب هى على التوالى :

منسوب سعر القطن $= \frac{٣{,}٩٦٠}{٣{,}٦٢٥} \times ١٠٠ = ١٠٩{,}٢٤$ ؛

« « القمح $= \frac{١{,}٥٣١}{١{,}٢١٩} \times ١٠٠ = ١٢٥{,}٥٩$ ؛

منسوب سعر الفول $= \frac{١،٣٤٣}{١،٣٩٩} \times ١٠٠ = ٩٦،٠٠$ ؛

« « الشعير $= \frac{٠،٨٠٢}{٠،٦٤٣} \times ١٠٠ = ١٢٤،٧٣$.

∴ الوسط الحسابى البسيط لها $= ١١٣،٨٩$ ؛

و « التوافقى « « $= ١١٣،٠٦$ ؛

و « الهندسى « « $= ١١٣،٢١$.

و بناء على ذلك يكون مستوى أسعار هذه المحاصيل ارتفع فى سنة ١٩٣٥ عنه فى سنة ١٩٣١ بمقدار ١٣،٨٩ أو ١٣،٠٦ أو ١٣،٢١ فى المائة ، حسب الصيغة التى نختارها للرقم القياسى : الوسط الحسابى أو التوافقى أو الهندسى (البسيط). وهذه النتائج مختلفة فيما بينها ، وتخالف أيضاً النتائج التى حصلنا عليها من قبل. وهذا بالرغم من أن كلها ترمى إلى تصوير شىء واحد : ألا وهو مستوى أسعار هذه المحاصيل فى سنة ١٩٣٥ بالنسبة إلى سنة ١٩٣١ ، ونكتفى هنا بالإشارة إلى هذا الاختلاف وسنتكلم عن منشئه فيما بعد.

ترجيح المناسيب حسب أهمية السلع

٢٩٠ — هذه المتوسطات البسيطة لا تفرق بين مناسيب السلع المختلفة بل تعاملها جميعاً نفس المعاملة ، مع العلم بأن بعض السلع يزيد فى الأهمية عن البعض الآخر. وعلى ذلك فالأرقام القياسية التى نحصل عليها بالمتوسطات البسيطة لا تصور الحالة على حقيقتها فتعطى نتائج مضللة أو خاطئة. ولذلك يستحسن تعديل هذه المتوسطات باستخدام أوزان تتناسب مع أهمية السلع ، نرجح بها المناسيب الخاصة بها.

وأحسن شىء نقيس به أهمية السلعة هو قيمتها ، أى حاصل ضرب سعرها فى كميتها. ولكن أى كمية وأى سعر ؟ فقد عرفنا أن لكل سلعة سعراً أساسياً

وسعراً مقارناً ، وكذلك كمية أساسية وكمية مقارنة . فمن الممكن إذن أن نختار أحد التوافيق الآتية وهى :

ا — السعر الأساسى × الكمية الأساسية = $\text{ع}_{\cdot} \times \text{ك}_{\cdot} = \text{م}_{\cdot}$ مثلا ؛

ب — » » × » المقارنة = $\text{ع}_{\cdot} \times \text{ك}_{١} = \text{مـ}_{\cdot}$ » ؛

حـ — » المقارن × » الأساسية = $\text{ع}_{١} \times \text{ك}_{\cdot} = \text{مـ}_{١}$ » ؛

ى — » » × » المقارنة = $\text{ع}_{١} \times \text{ك}_{١} = \text{م}_{١٢}$ » ؛

وعلى ذلك يمكن تركيب رقم قياسى من المناسيب على الصور الآتية :

وسط حسابى مرجح بأوزان ا وهو $\frac{\text{مجـ س م}_{\cdot}}{\text{مجـ م}_{\cdot}}$ . . . (١٤ ا) ،

» » » » ب » $\frac{\text{مجـ س مـ}_{\cdot}}{\text{مجـ مـ}_{\cdot}}$ . . . (٤١ ب) ،

» » » » حـ » $\frac{\text{مجـ س مـ}_{١}}{\text{مجـ مـ}_{١}}$ . . . (٤١ حـ) ،

» » » » ى » $\frac{\text{مجـ س م}_{١٢}}{\text{مجـ م}_{١٢}}$ . . . (٤١ ى) .

يلاحظ أنه بوضع $\text{س} = \frac{\text{ع}_{١}}{\text{ع}_{\cdot}}$ يتضح أن (١٤) هو نفس (١٢) المذكور فى بند ٢٨٦ ، وأن (٤ ب) هو نفس (٢ ب) .

٢٩١ — وبالمثل نحصل على الوسط التوافقى المرجح ، بهذه الأوزان الأربعة ؛ ويكون مقلوب الوسط التوافقى يساوى : الوسط التوافقى المرجـح

$$\frac{١}{\text{مجـ م}_{\cdot}} \times \text{مجـ} \left(\frac{\text{م}_{\cdot}}{\text{س}} \right) \quad . . . \quad (١٥) ،$$

أو

$$\frac{١}{\text{مجـ مـ}_{\cdot}} \times \text{مجـ} \left(\frac{\text{مـ}_{\cdot}}{\text{س}} \right) \quad . . . \quad (٥ ب) ،$$

أو

$$\frac{١}{\text{مجـ مـ}_{١}} \times \text{مجـ} \left(\frac{\text{مـ}_{١}}{\text{س}} \right) \quad . . . \quad (٥ حـ) ،$$

أو

$$\frac{١}{\text{مجـ م}_{١٢}} \times \text{مجـ} \left(\frac{\text{م}_{١٢}}{\text{س}} \right) \quad . . . \quad (٥ ى) .$$

الوسط الهندسى المرجح

٢٩٢ — أما فى الوسط الهندسى المرجح فتظهر الأوزان كأسس ترفع إليها المناسيب ، حيث قد قلنا فى باب المتوسطات إن لوغاريتم الوسط الهندسى لأى مجموعة من القيم ، يساوى الوسط الحسابى للوغاريتمات هذه القيم . وعلى ذلك فالأرقام القياسية المركبة على أساس الوسط الهندسى المرجح (بأوزان ا و ب و حـ و ء) هى :

ا — بأوزان م. هو $\sqrt[\text{مجـ م}_0]{(\text{س})^{\text{م}_0} \times (\text{س}')^{\text{م}'_0} \times (\text{س}'')^{\text{م}''_0} \times \ldots}$ (١٦) ؛

ب — « مـ هو $\sqrt[\text{مجـ مـ}_0]{(\text{س})^{\text{مـ}_0} \times (\text{س}')^{\text{مـ}'_0} \times (\text{س}'')^{\text{مـ}''_0} \times \ldots}$ (١٦ ب) ؛

حـ — « مـ١ هو $\sqrt[\text{مجـ مـ}_1]{(\text{س})^{\text{مـ}_1} \times (\text{س}')^{\text{مـ}'_1} \times (\text{س}'')^{\text{مـ}''_1} \times \ldots}$ (١٦ حـ) ؛

ء — « م١ هو $\sqrt[\text{مجـ م}_1]{(\text{س})^{\text{م}_1} \times (\text{س}')^{\text{م}'_1} \times (\text{س}'')^{\text{م}''_1} \times \ldots}$ (١٦ ء)

يوجد ستة أسس للأرقام القياسية

٢٩٣ — وهكذا يكون لدينا ستة أسس لتركيب أرقام قياسية وهى : (١) تجميع الأسعار نفسها ؛ (٢) الوسط الحسابى للمناسيب ؛ (٣) الوسط التوافقى ؛ (٤) الوسط الهندسى . ونذكر على سبيل الحصر أيضاً (٥) الوسيط و (٦) المنوال لهذه المناسيب ، ولكن استعمالها نادر جداً .

وفى كل من هذه الحالات يصح أن نستخدم أوزاناً (نختارها بعدة طرق لترجيح السلع المختلفة بما يتناسب وأهميتها ، أو لا نستخدم أوزاناً بالمرة .

دخول عنصر الزمن فى المقارنة على أساس قديم

٢٩٤ — عند ما نستخدم الأرقام القياسية لبيان حركة الأسعار — أو أى ظاهرة أخرى — أثناء مدة طويلة ، يدخل فى المسألة عنصر جديد وهو الزمن ؛ حيث إذا طالت المدة بين السنة أو الفترة المعتبرة أساساً والسنة المقارنة ، فإن الزمن

كفيل بأن يحدث تغييراً كبيراً فى الظروف المحيطة بالسلع التى نبحثها والتى يتركب منها الرقم القياسى . فقد يحدث مثلا أن بعض السلع التى كانت شائعة الاستهلاك فى سنة الأساس قل استهلاكها أو انعدم فى الأزمنة الحديثة ؛ أو بالعكس قد توجد سلع لم تكن معروفة من قبل كما نرى فى حالة سلعة مثل الحرير الصناعى ومنسوجاته . أو على الأقل تتغير الأهمية النسبية بين السلع التى تدخل فى تركيب الرقم القياسى . وهذا إذا حصل — ولا بد أن شيئاً من هذا يحصل إذا طالت الفترة — لا يمكن معه الاطمئنان إلى صحة المقارنة، وتنعدم بذلك الفائدة التى من أجلها فكرنا فى إنشاء الرقم القياسى . ويجب إذن عند البحث فى الرقم القياسى للأسعار أو خلافها ، بالنسبة إلى سنة بعيدة ، أن نتأكد من استمرار الظروف على حالها — ولو تقريباً — أثناء فترة المقارنة .

نستعمل طريقة السلسلة إذا تغيرت الظروف

٢٩٥ — وإذا لم نضمن ثبات هذه الظروف الاقتصادية ، فيحسن أن نركب سلسلة من الأرقام القياسية كل سنة بالنسبة إلى سابقتها كأساس ، وهذه السلسلة يمكن استخدامها لمراقبة حركة الأسعار من سنة إلى أخرى تسبقها أو تليها .

ولتوضيح طريقة تركيب هذه السلسلة (١) ، نفرض أن لدينا أسعار السلع فى عدة سنين متتالية مثل ١٩٢٠ و١٩٢١ و١٩٢٢ و.. ؛ ولتكن هذه الأسعار كما يأتى :

	$ع_{.}$ ،	$ع_{١}$ ،	$ع_{٢}$ ،	... ...	للسلعة الأولى ؛
و	$عَ_{.}$ ،	$عَ_{١}$ ،	$عَ_{٢}$ ،	... ...	« الثانية ؛
و	$عً_{.}$ ،	$عً_{١}$ ،	$عً_{٢}$ ،	... ...	« الثالثة ؛

وهكذا لباقى السلع .

نحسب الرقم القياسى للأسعار فى سنة ١٩٢١ بالنسبة إلى سنة ١٩٢٠ ، طبقاً لأى معادلة من المعادلات التى شرحناها فى هذا الباب . ولنرمز إلى هذا الرقم

(١) تسمى بالإنجليزية Chain System

للاختصار بالرمز[1] ع١٠ . ثم نحسب الرقم القياسى لأسعار سنة ١٩٢٢ بالنسبة إلى سنة ١٩٢١ كأساس ؛ وليكن هذا الرقم هو ع٢١ . ونحسب أيضاً رقم سنة ١٩٢٣ بالنسبة إلى سنة ١٩٢٢ وليكن ع٣٢ ؛ وهكذا مع باقى السنين . فنحصل على سلسلة الأرقام القياسية الآتية :

ع١٠ ، ع٢١ ، ع٣٢ ، ع٤٣ ، . . .

وهى تعبر عن مستوى الأسعار فى كل سنة بالنسبة إلى التى قبلها مباشرة كأساس . وبذلك يكون الأساس فى هذه السلسلة **متحركا**[2] ، وليس ثابتاً كما لو اتخذنا سنة ١٩٢٠ أساساً لكل السنين .

ميزة هذه الطريقة على السابقة

٢٩٦ — وعند حساب أى واحد من هذه الأرقام يمكننا أن ندخل فى الاعتبار أى تغيير يطرأ على الظروف المحيطة بهذه السلع . فنكيف تركيب الرقم بما يلائم الحالة فى كل سنة . فمثلا ندخل فى الرقم أى سلعة عظم شأنها فى السوق بعد أن كانت قليلة الأهمية ؛ أو نهمل سلعة أخرى ضعفت حركتها وفقدت أهميتها فى السوق بعد أن كانت تلعب دوراً خطيراً بين السلع من قبل . وبالجملة نعدل الأهمية النسبية بين السلع الداخلة فى الرقم بما يتناسب مع الظروف .

وهذه المرونة غير موجودة فى الرقم القياسى ذى القاعدة الثابتة ، الذى تنسب فيه الأسعار فى السنة المقارنة إلى السنة الأساسية مباشرة ، وبذلك ترجح هى أو مناسيبها بأوزان ثابتة على طول السنين ، فلا تتمشى مع الظروف الاقتصادية التى تحيط بالسلع الداخلة فى تركيب الرقم .

---

(١) هذا الرمز ينطق هكذا : عين صفر واحد ، أى مستوى الأسعار للسنة ١ ، بالنسبة إلى السنة ٠ كأساس ة وكذلك ع٢١ عين واحد اثنين ، أى أسعار السنة ٢ بالنسبة إلى السنة ١ وهكذا .

(٢) يسمى بالإنجليزية Shifting or Moving Base

تثبيت الأساس في طريقة السلسلة

٢٩٧ — وإذا أردنا الرجوع إلى أساس ثابت ، فيمكننا بسهولة أن نعبر عن مستوى الأسعار فى هذه السنين بالنسبة إلى سنة ١٩٢٠ (السنة الأساسية) بطريق غير مباشر كما يأتى :

∵ $ع_{١٠}$ هو مستوى الأسعار فى سنة ١٩٢١ بالنسبة إلى سنة ١٩٢٠

و $ع_{٢١}$ « « « « ١٩٢٢ « « ١٩٢١

∴ حاصل الضرب

$$ع_{١٠} \times ع_{٢١}$$

يعبر عن مستوى الأسعار فى سنة ١٩٢٢ بالنسبة إلى سنة ١٩٢٠ (عن طريق سنة ١٩٢١) . وكذلك حاصل الضرب

$$(ع_{١٠} \times ع_{٢١}) \times ع_{٣٢}$$

يعبر عن مستوى الأسعار فى سنة ١٩٢٣ بالنسبة إلى سنة ١٩٢٠ ، (عن طريق سنتى ١٩٢١ و ١٩٢٢) .

وهكذا نحصل على الرقم القياسى لمستوى الأسعار فى كل سنة بواسطة الرقم القياسى للسنة التى قبلها مباشرة ؛ فتتكون لدينا سلسلة من الأرقام القياسية تعبر عن مستوى الأسعار منسوباً إلى أساس ثابت (وهو سنة ١٩٢٠ هنا) . وبالطبع تكون هذه الأرقام على العموم مخالفة (١) للأرقام القياسية المحسوبة بالطريقة المباشرة ، أى بنسبة كل سنة إلى سنة الأساس مباشرة .

مثال لحساب الأرقام القياسية عملياً

٢٩٨ — نأخذ مثلا أسعار أربعة المحاصيل المصرية : القمح والأذرة والفول والشعير فى السنين ١٩٣٠ — ١٩٣٥ ، ونحسب سلسلة الأرقام القياسية لمستوى أسعار هذه المحاصيل بالطرق المختلفة . وللسهولة نستعمل الوسط الحسابى البسيط للمناسيب .

(١) إلا فى حالة استعمال الوسط التجميعي البسيط للأسعار ، وفى حالة ماتتغير أسعار جميع السلع بنسبة واحدة .

جدول ٥٠ — أسعار بعض الحبوب المصرية [1]

في السنين ١٩٣٠ — ١٩٣٥

	الأسعار بالقرش للأردب					
	١٩٣٠	١٩٣١	١٩٣٢	١٩٣٣	١٩٣٤	١٩٣٥
القمح	١٢٩	١٣٢	١٠٧	١١٢	١٥٠	١٤٣
الأذرة	٩٥	٧٧	٥٩	١٠٤	١٠٦	٧٣
الفول	١٨٣	١٥٩	٨٥	٨٨	١٥٥	١٥٣
الشعير	٧٠	٨٨	٥٦	٤٦	٩٠	٨٥

الوسط الحسابي للمناسيب بأساس ثابت

أولا : توجد الأرقام القياسية للأسعار بالنسبة إلى سنة ١٩٣٠ كأساس ثابت ؛ فنحسب في كل سنة منسوب السعر لكل سلعة بالنسبة إلى السنة الأساسية ثم نوجد الوسط الحسابي للمناسيب في كل سنة ينتج الرقم القياسي للأسعار. وهذه العمليات مبينة في الجدول الآتي :

جدول ٥١ — مناسيب أسعار أربعة محاصيل

في السنين ١٩٣١ — ١٩٣٥ بالنسبة إلى سنة ١٩٣٠ كأساس

	١٩٣٠	١٩٣١	١٩٣٢	١٩٣٣	١٩٣٤	١٩٣٥
القمح	١٠٠	١٠٢,٣٣	٨٢,٩٥	٨٦,٨٢	١١٦,٢٨	١١٠,٨٥
الأذرة	١٠٠	٨١,٠٥	٦٢,١١	١٠٩,٤٧	١١١,٥٨	٧٦,٨٤
الفول	١٠٠	٨٦,٨٩	٤٦,٤٥	٤٨,٠٩	٨٤,٧٠	٨٣,٦١
الشعير	١٠٠	١٢٥,٧١	٨٠,٠٠	٦٥,٧١	١٢٨,٥٧	١٢١,٤٣
متوسط المناسيب	١٠٠	٩٩,٠٠	٦٧,٨٨	٧٧,٥٢	١١٠,٢٨	٩٨,١٨

(١) أنظر الإحصاء السنوى العام ١٩٣٥ — ١٩٣٦ . صفحة ٥٠٤ .

فالأرقام القياسية للأسعار فى السنتين ١٩٣١ — ١٩٣٥ بالنسبة إلى سنة ١٩٣٠ كأساس ثابت ، هى على الترتيب :

٩٩٫٠٠ ، ٦٧٫٨ ، ٧٧٫٥ ، ١١٠٫٣ ، ٩٨٫٢

الرقم القياسى بأساس متحرك

ثانياً : نوجد الأرقام القياسية على نظام السلسلة ذات الأساس المتحرك ؛ فنحسب منسوب سعر السلعة فى كل سنة بالنسبة إلى سابقتها مباشرة . والمنسوب بهذا المعنى الخاص نسميه **النسيب**[1] . فيكون نسيب سعر القمح فى سنة ١٩٣٥ مثلا ، هو خارج قسمة السعر فى سنة ١٩٣٥ على السعر فى سنة ١٩٣٤ ، مضروباً فى ١٠٠ . ونرمز له بالرمز $س_{٥٤}$ مثلا .

والرقم القياسى المطلوب يكون إذن هو الوسط الحسابى لأنسبة السلع فى كل سنة . أى أن سنة ١٩٣٥ يكون رقمها القياسى بطريقة السلسلة هو :

$$ع_{٥٤} = \frac{١}{٤} (مجـ\ س_{٥٤})$$

جدول ٥٢ — أنسبة أسعار أربعة المحاصيل فى السنين
١٩٣١ — ١٩٣٥ ، أساس متحرك

	١٩٣١	١٩٣٢	١٩٣٣	١٩٣٤	١٩٣٥
القمح ...	١٠٢٫٣٣	٨١٫٠٦	١٠٤٫٦٧	١٣٣٫٩٣	٩٥٫٣٣
الأذرة ...	٨١٫٠٥	٧٦٫٦٣	١٧٦٫٢٥	١٠١٫٩٣	٦٨٫٨٧
الفول ...	٨٦٫٨٩	٥٣٫٤٦	١٠٣٫٥٣	١٧٦٫١٣	٩٨٫٧١
الشعير ...	١٢٥٫٧١	٦٣٫٦٣	٨٢٫١٤	١٩٥٫٦٦	٩٤٫٤٥
متوسط الأنسبة ...	٩٩٫٠٠	٦٨٫٧٠	١١٦٫٦٥	١٥١٫٩١	٨٩٫٣٤

وعلى ذلك فالأرقام القياسية للأسعار فى السنين ١٩٣١ — ١٩٣٥ ، كل بالنسبة إلى سابقتها ، هى على الترتيب :

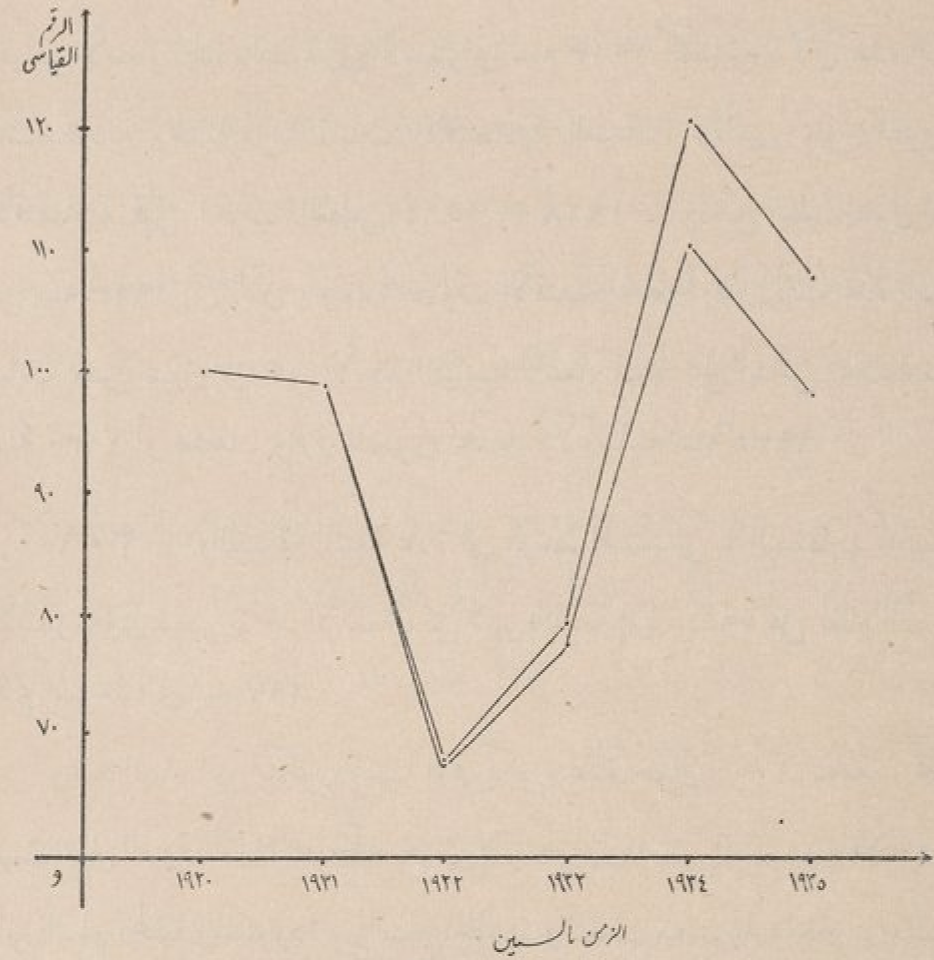
(١) بالإنجليزية Link Relative

٩٩ و ٦٨,٧ و ١١٦,٥ و ١٥١,٩ و ٨٩,٣ ؛

أى $\text{ع}_{\text{١٠}}$ و $\text{ع}_{\text{٢١}}$ و $\text{ع}_{\text{٣٢}}$ و $\text{ع}_{\text{٤٣}}$ و $\text{ع}_{\text{٥٤}}$

وإذا أردنا الرجوع إلى أساس ثابت (١٩٣٠ مثلاً) ، نحسب الحواصل $\text{ع}_{\text{١٠}} \times \text{ع}_{\text{٢١}}$ ، $\text{ع}_{\text{١٠}} \times \text{ع}_{\text{٢١}} \times \text{ع}_{\text{٣٢}}$ ، $\text{ع}_{\text{١٠}} \times \text{ع}_{\text{٢١}} \times \text{ع}_{\text{٣٢}} \times \text{ع}_{\text{٤٣}}$ ، . . . فينتج الرقم القياسى لكل سنة بالنسبة إلى سنة ١٩٣٠ كأساس (غير مباشر) . وهذه الأرقام هى :

١٠٠,٠٠ ، ٩٩,٠٠ ، ٦٨,٠ ، ٧٩,٣ ، ١٢٠,٥ ، ١٠٧,٧



(شكل ٦٨)

الأرقام القياسية لأسعار الحبوب بالنسبة إلى سنة ١٩٣٠

ونلاحظ أن هذه الأرقام مخالفة للتى حصلنا عليها فى جدول ٥١ ، حيث نسبنا كل سنة إلى سنة الأساس مباشرة . ويلاحظ أن الاختلاف يزيد كلما بعدنا عن سنة الأساس ، ونرى ذلك موضحاً فى شكل ٦٨

بعض الأرقام القياسية المهمة

الرقم القياسى لأسعار الجملة

٢٩٩ — الرقم القياسى لأسعار الجملة [١] من أهم الأرقام القياسية المستعملة فى المسائل الاقتصادية على العموم . ونجد الحكومات فى كل البلاد المتمدنة تقوم بعمل ونشر هذا الرقم للدلالة على حركة التجارة ، وعلى الحالة الاقتصادية العامة فى البلد .

سنة ١٩١٣ كانت تؤخذ كأساس

٣٠٠ — وقد كانت أغلب الحكومات إلى عهد قريب تنشىء الرقم القياسى لأسعار الجملة بالنسبة إلى الأسعار فى سنة ١٩١٣ كأساس ؛ لأن هذه السنة تعتبر عادية وخالية من التقلبات الاقتصادية العنيفة ؛ ولأنها تمثل الظروف الاقتصادية قبل الحرب العظمى ١٩١٤ — ١٩١٨ . ولكن نظراً لطول المدة من سنة ١٩١٣ إلى الآن ، وتغير الظروف الاقتصادية تغيراً كبيراً أثناء هذه المدة ، بدأت بعض الدول فى عمل الأرقام القياسية لأسعار الجملة على أساس أحدث مثل سنة ١٩٣٠ أو بعدها . والرقم المصرى الجديد [٢] أساسه سنة ١٩٣٥ .

الطريقة المتبعة فى رقم أسعار الجملة

الوسط الهندسى البسيط

٣٠١ — والطريقة المتبعة عادة هى الوسط الهندسى (البسيط) لمناسيب الأسعار ؛ وفى مصر تؤخذ الأنسبة ، كل شهر بالنسبة إلى سابقه ، على نظام السلسلة الذى شرحناه فى بند ٢٩٧ .

وعدد السلع التى يتكون منها الرقم يكون عادة حوالى ١٠٠ سلعة ، تختار بحيث تمثل السوق تمثيلا صحيحاً ، بحيث لا يتحيز إلى بعض السلع دون الأخرى ، مثل السلع المحلية والمستوردة ، والسلع الجاهزة والنصف مصنوعة والخام ، والسلع الزراعية والصناعية ، وهكذا .

---

(١) بالإنجليزية Whole Sale Prices Index Number

(٢) أنظر الإحصاء السنوى العام ١٩٣٥ — ١٩٣٦ صفحة ٤٩٠ . وقد بدأت مصلحة الإحصاء المصرية تنشر أرقاماً قياسية منسوبة إلى الفترة يونية — يولية — أغسطس سنة ١٩٣٩ كأساس ، وهى الفترة السابقة للحرب العظمى الثانية ١٩٣٩ — ١٩٤٥ .

ونكتفى هنا بهذا الوصف البسيط ؛ وسنعود إلى شرح هذا الرقم بالتفصيل فى المستقبل .

الرقم القياسى لنفقة المعيشة

٣٠٢ — **الرقم القياسى لنفقة المعيشة** [١] ، وهو يقوم مقام الرقم القياسى لأسعار التجزئة فى أغلب الأبحاث ، وهو من أهم الأرقام القياسية المستعملة أيضاً ، حيث إنه يمس النواحى الاجتماعيـة والتجارية والصناعية للدولة فى نفس الوقت . وربمـا يعتبر أهم من الرقم القياسى لأسـعار الجملة الذى يختص بناحية واحدة من نواحى النشاط الاقتصادى للبلد ، ألا وهى تجارة الجملة .

والمقصود من إنشاء هذا الرقم القياسى هو قياس التغيرات التى تطرأ على نفقة المعيشـة بسبب تغيرات أسـعار الحاجيات الضرورية للحياة ، والتى يستهلكها السواد الأعظم من السكان ، لـكى نكون دائماً على بينة من الحالة المعيشية للشعب ودرجة رفاهيته ؛ ولنعلم مبلغ كفاية الأجور فى توفير أسباب المعيشة ؛ ومقدرة الناس على شراء السلع والحاجيات التى يستهلكونها .

والأساس المتفق عليه فى إنشاء هذا الرقم ، فى كل الدول تقريباً ، كان أسعار هذه الحاجيات قبل الحرب العظمى الأولى أى سـنة ١٩١٣ ، لنفس السبب الذى ذكرناه . ولـكن كثيراً من الدول غيرت هذا الأساس إلى ما قبل الحرب العالمية الثانية ١٩٣٩ — ١٩٤٥ .

نستخدم طريقة الوسط المرجح

٣٠٣ — والطريقة المتبعة فى تركيب الرقم القياسى لنفقة المعيشة ، هى الوسط المرجح لمناسيب أسـعار هذه الحاجيات فى الشهر أو السنة المقـارنة ، على أساس أسعارها فى سنة ١٩١٣

والأوزان المستعملة فى ترجيح مناسيب الأسعار تتناسب مع أهمية السلع . وتقاس أهمية السلعة بمقدار ما يخصص للإنفاق عليها من الدخل الـكلى .

(١) اسمه بالإنجليزية Cost of Living Index Number

تقسيم السلع فى أربع مجموعات أو خمس

٣٠٤ — وفى العادة تقسم الحاجيات التى تدخل فى تركيب الرقم القياسى لنفقة المعيشة (من سلع وخدمات وغيرها) إلى مجموعات ؛ ويكوّن لكل منها رقم قياسى خاص يبين التغيرات فى أسعارها على حدة ومن هذه الأرقام الجزئية يكون الرقم القياسى العام الذى يشمل كل المجموعات . وهذه المجموعات خمس فى أغلب البلاد ؛ وهى تمثل البنود الرئيسية للإنفاق وهى : المأكل ، والملبس والمسكن ، والإضاءة ، والتدفئة (للبلاد الباردة) ، والمصروفات النثرية التى لاتدخل تحت أى واحد من هذه البنود .

وتقسيم الدخل بين هذه المجموعات ، وبين المفردات الداخلة فى كل مجموعة مبنى على أساس ميزانية الأسرة النموذجية ، وكيفية توزيع دخلها فى شراء ما كانت تستهلكه من الحاجيات قبل الحرب — أو فى سنة ١٩٢٠ بالنسبة لمصر (١) ، وفى سنة ١٩٣٧ — ١٩٣٨ بالنسبة لانجلترا .

معنى الرقم القياسى لنفقة المعيشة

٣٠٥ — وبناء على ذلك يكون معنى الرقم القياسى لنفقة المعيشة ، وليكن ٣٠٧ مثلاً ، كما كان فى مصر فى شهر مايو سنة ١٩٤٣ بالنسبة إلى سنة ١٩١٣ ، هو أن المعيشة التى كانت تتكلف ١٠٠ قرش فى سنة ١٩١٣ أصبحت تتكلف ٣٠٧ قرشاً فى مايو سنة ١٩٤٣ . وهذا طبعاً لا يقرر أن الناس لا زالوا يعيشون كما عاشوا فى سنة ١٩١٣ أو أنهم يستهلكون الآن ما كانوا يستهلكونه فى تلك السنة . وهذه فى الحقيقة نقطة ضعف فى طريقة تركيب هذا الرقم ، وفى غيره أيضاً ، نظراً لطول المدة التى مضت من سنة ١٩١٣ إلى الآن .

وقد رأينا فى بند ٤١ (صفحة ٣٢) النسب المئوية لتقسيم الدخل بين البنود

---

(١) أنظر مثلا الإحصاء السنوى العام ٩٣١ — ١٩٣٢١ صفحة ٤٣٩ . أو أنظر تقرير نفقات المعيشة الملحق « بالإحصائية الشهرية الزراعية » شهر نوفمبر سنة ١٩٢٠ .

الأربعة المهمة فى مصر وانجلترا (خمسة بنود) ، وهى الأرقام المستعملة فى تركيب الرقم القياسى لنفقة المعيشة فى البلدين . (وقد أصبح عدد البنود ٨ فى انجلترا من يولية ١٩٤٧) .

الرقم القياسى للانتاج

٣٠٦ — يلى هذين الرقمين فى الأهمية **الرقم القياسى للانتاج الصناعى**[١] (أو الزراعى) ؛ ونجد كثيراً من البلاد الصناعية فى الوقت الحاضر تمنى بعمل ونشر أرقام قياسية للانتاج الصناعى بأقسامه المختلفة ، وخصوصاً منذ سنة ١٩٣٠

وهذا الرقم يستخدم للدلالة على حالة النشاط الصناعى فى الدولة . وهو يبين التغير فى كمية المنتج من الصناعات المختلفة ، بصرف النظر عن القيمة النقدية لهذه المنتجات . والمقصود بالكمية ، مجردة عن القيمة ، هو عدد الوحدات المنتجة فى كل صناعة ؛ لأن هذا العدد يدل بدقة على درجة النشاط أو الركود فى الأعمال الصناعية ؛ ويدل أيضاً على حالة العمل والبطالة بين جمهور العمال ؛ ويدل على مقدار الكفاية الإنتاجية للصناعات المختلفة والعمال المشتغلين فيها ؛ وغير ذلك من الظواهر الاقتصادية المهمة .

ومثل هذا يقال عن الإنتاج الزراعى . وفى مصر يوجد رقم قياسى لهذا الإنتاج الزراعى تنشره مصلحة عموم الإحصاء[٢] ، ولكن لا يوجد رقم قياسى للانتاج الصناعى .

طريقة إنشاء الرقم القياسى للانتاج

٣٠٧ — الطريقة المتبعة فى تركيب هذا الرقم القياسى تشابه طريقة الرقم القياسى لنفقة المعيشة ، حيث نتفق على سنة عادية نعتبرها أساساً ؛ ونكوّن لكل صناعة منسوباً لكمية الإنتاج فيها فى السنة المقارنة بالنسبة للسنة الأساسية .

---

(١) اسمه بالإنجليزية Index Number of the Phyisical Volume of Production

(٢) أنظر مثلا الإحصاء السنوى العام ١٩٣٥ — ١٩٣٦ صحيفة ٣٨٤ .

وهذه المناسيب ترجح بأوزان تتناسب وأهمية الصناعات ، ويستخرج الوسط المرجح لهذه المناسيب ، وهو الرقم القياسى المطلوب .

وتقاس أهمية كل صناعة بمقدار صافى إنتاجها [١] أثناء فترة زمنية معينة تعتبر عادية وهادئة (تؤخذ سنة الأساس عادة) ؛ والأوزان المستعملة تتناسب مع هذه المقادير .

بعض الصعوبات العملية

٣٠٨ — الصعوبات التى نلاقيها فى عمل هذا الرقم القياسى هى أولا من ناحية الحصول على بيانات عن الصناعات المختلفة يمكن مقارنتها بسهولة على مرور السنين . لأن المنتجات الصناعية تتغير فى شكلها وماهيتها وفى طريقة صنعها ، على مرور الزمن . وعلى ذلك فمن الصعب مقارنة عدد الوحدات المنتجة هذا العام بعدد الوحدات المنتجة فى نفس المصنع ، أو نفس الصناعة ، من مدة خمس سنين مثلا .

وثانياً نلاقى صعوبة أخرى فى جمع بيانات يعتمد عليها من عدد كبير من الصناعات والمؤسسات ، أو فى مدة قصيرة تسمح بعمل الرقم القياسى للإنتاج ونشره بسرعة ، قبل مرور مدة طويلة على الفترة المشمولة فى تركيب الرقم .

رغم هذه الصعوبات يجب عمل هذا الرقم

٣٠٩ — ولكن هذه الصعوبات ، وإن لم يمكن التغلب عليها تماماً ، يجب ألا تقعدنا عن تركيب هذا الرقم بأى طريقة كانت ، لما له من الفائدة العظيمة فى تصوير حركة النشاط الصناعى ، ومعرفة مواطن الركود ومعالجتها بسرعة ، ومعرفة نواحى النشاط واستغلالها . ونكتفى هنا بالإشارة إلى هذه الصعوبات وسوف نعود إلى شرحها بالتفصيل والطرق المستعملة فى بعض البلاد للتغلب على هذه الصعوبات العملية .

(١) أو بعدد المشتغلين فيها .

٣١٠ — خلاف هذه الأرقام القياسية ، نجد بعض البلاد تكوّن أرقاماً قياسية لأسعار الأوراق المالية وللأجور وغيرها من الظواهر الاقتصادية والاجتماعية المهمة . وسيأتي الكلام على هذه الأرقام وكيفية تركيبها فيما بعد .

المـــراجع

Bowley, A.L.,	: *Elements of Statistics*	Chapter	IX
Conner, L.R.	: *Statistics in Theory and Practice*	»	XV
Mills, F.C.,	: *Statistical Methods*	»	VI
Secrist, H.	: *Statistical Methods*	»	XV

الباب الثاني عشر

اختبار الأرقام القياسية

توجد عدة صيغ مختلفة للرقم القياسي

٣١١ — تكلمنا فى الباب السابق عن الصيغ المختلفة المستعملة فى تركيب الأرقام القياسية ؛ وشرحنا طريقة بنائها من البيانات التى يمكننا الحصول عليها ، مثل أسعار وكميات السلع . وقد أوردنا هذه المعادلات أو الصيغ المختلفة على اعتبار أن كلا منها يعبر بطريقة ما عن المعنى أو الغرض المقصود من الرقم القياسى : ألا وهو رقم نسبى ملخص يقيس مقدار التغير فى ظاهرة أو عدة ظواهر بين لحظة وأخرى أو بين بلد وأخرى .

النتائج تختلف حسب الصيغ المستعملة

٣١٢ — وقد رأينا أن الأرقام القياسية المحسوبة على أساس هذه المعادلات المختلفة تؤدى إلى نتائج مختلفة أيضاً ، أى أنها تصور الأحوال تصويراً مختلفاً على حسب المعادلة أو الصيغة التى نستعملها . وهذا بطبيعة الحال يؤدى بنا إلى التساؤل : أى هذه الصيغ أو المعادلات صحيح ، وأيها خطأ ؟

وكيف نختبر أى واحدة منها لنتبين صحتها أو خطأها ، ومقدار هذا الخطأ ؟

وكيف نختار الأصلح من بينها جميعاً ؟

٣١٣ — والمفاضلة بين صيغ الأرقام القياسية المختلفة لا بد تتناول الناحيتين النظرية والعملية . وقد بحث الأستاذ إرفنج فيشر [١] فى الأسس النظرية لكيفية

(١) أنظر كتابه I. Fisher. *The Making of Index Numbers*

اختبار الأرقام القياسية وتصحيحها ، وإليه يرجع الفضل الأكبر فى هذا الموضوع ؛ وسنشرح هنا باختصار الاختبارات التى يقترحها وكيفية تطبيقها .

الرموز المستعملة فى البحث

٣١٤ — لنفرض ، كما سبق ، أن أسعار السلع وكمياتها فى الفترة (أو البلد) الأساسية هى كما يلى على الترتيب :

$$\text{ع}_{٠} ، \text{ع}'_{٠} ، \text{ع}''_{٠} ، \ldots ،$$

$$\text{و} \quad \text{ك}_{٠} ، \text{ك}'_{٠} ، \text{ك}''_{٠} ، \ldots ؛$$

وأن هذه الأسعار والكميات فى السنة (أو البلد) المقارنة هى :

$$\text{ع}_{١} ، \text{ع}'_{١} ، \text{ع}''_{١} ، \ldots ،$$

$$\text{و} \quad \text{ك}_{١} ، \text{ك}'_{١} ، \text{ك}''_{١} ، \ldots .$$

وإذا كان هناك سنة مقارنة أخرى نضع الرقم ٢ بدل الرقم ١ بجانب هذه الحروف . وعلى العموم ندل على السنة الأساسية بالرقم ٠ ، والسنين الأخرى بالأرقام ١ و ٢ و ٣ و ٠٠ ، إذا وجدت . فإذا استعملنا الحرف س للدلالة على مناسيب الأسعار ، كان $\text{س}_{١٠}$ يدل على منسوب سعر السنة ١ على أساس السنة ٠ ، وكان $\text{س}_{٢٠}$ يدل على منسوب السعر فى السنة ٢ بالنسبة إلى السنة ٠ كأساس . وبالمثل $\text{س}_{٢١}$ يدل على منسوب السعر للسنة ٢ على أساس السنة ١ ، وهكذا . أى أن الرقم الأول يدل على السنة المعتبرة أساساً ، والرقم الثانى يعين السنة المقارنة .

اختبار ترتيب وضع السلع

٣١٥ — من البدهى أن الرقم القياسى الصحيح يجب ألا تتغير قيمته إذا تغير ترتيب أسعار السلع فى المعادلة الجبرية للرقم ؛ فمثلا فى الرقم التجميعى البسيط ، وهو المذكور فى المعادلة (١) بند ٢٨٥ ، :

$$\frac{\text{ع}_{١} + \text{ع}'_{١} + \text{ع}''_{١} + \cdots}{\text{ع}_{٠} + \text{ع}'_{٠} + \text{ع}''_{٠} + \cdots} ،$$

لا تتغير قيمة هذا الكسر إذا غيرنا ترتيب وضع الأسعار ع ، عَ ، عً ، . . في البسط أو في المقام . وعلى ذلك تكون هذه المعادلة ، أو الصيغة مستوفية لهذا الشرط . وفي الواقع نجد أن كل الصيغ التي ذكرناها في الباب السابق تستوفي هذا الشرط .

الانعكاس في الزمن

الانعكاس في الزمن أو في المكان

٣١٦— إذا أخذنا سلعة معينة وحسبنا الرقم القياسي لسعرها في سنة ١٩٣٨ بالنسبة إلى سعرها في سنة ١٩٣٠ مثلاً ، فمن البدهي أن هذا الرقم القياسي يساوي مقلوب الرقم القياسي لسعر هذه السلعة في سنة ١٩٣٠ بالنسبة إلى سعرها في سنة ١٩٣٨ كأساس . وإلا كان هذا الرقم القياسي خاطئاً ، ولا يؤدي المعنى أو الغرض الذي نرمي إليه من استخدام الأرقام القياسية . وكذلك لو كنا نبحث في الأسعار في مكانين مختلفين بدل سنتين .

وعلى ذلك يجب أن ينقلب الرقم القياسي عندما نعكس ترتيب السنين ، أو الأماكن المنسوبة والمنسوب إليها . فلو كان الرقم القياسي للأسعار في القاهرة ، مثلاً ، بالنسبة إلى الأسكندرية يساوي ١٢٥ ٪ ، يجب أن يكون الرقم القياسي للأسكندرية بالنسبة إلى القاهرة ٨٠ ٪ . وهذا واضح لأنه إذا كانت النسبة بين سعر سلعة في القاهرة وسعرها في الأسكندرية هي ١٢٥ ٪ ، أي كنسبة ٥ : ٤ ، كانت النسبة بين سعرها في الأسكندرية وسعرها في القاهرة تساوي ٤ : ٥ ، أي ٨٠ ٪ . وما يقال عن سلعة واحدة هنا يقال عن عدة سلع .

وبالاختصار يجب أن يكون حاصل ضرب الرقمين القياسيين المتبادلين في الزمن (أو المكان) يساوي ١ .

إذن لو أخذنا صيغة من الصيغ التى نعرفها ، وحسبنا بواسطتها الرقم القياسى للأسعار فى سنة ١٩٣٨ بالنسبة إلى سنة ١٩٣٠ مثلا ، ثم عكسنا الترتيب الزمنى وحسبنا ، بواسطتها أيضاً ، الرقم القياسى للأسعار فى سـنة ١٩٣٠ بالنسبة إلى سنة ١٩٣٨ ، يجب أن يكون حاصل ضرب النتيجتين يساوى الواحد . و إلا فتكون هذه الصيغة غير صالحة للاستعمال كرقم قياسى دقيق ، و يجب نبذها .

البديل الزمنى

٣١٧ — البديل الزمنى (١) لأى رقم قياسى هو نفس الرقم محسوباً بطريقة عكسية ، بحيث نعتبر السنة المقارنة هى السنة الأساسية ، والسنة الأساسية هى السنة المقارنة . أى أننا نضع الأسعار $\text{ع}_{١}$ بدل $\text{ع}_{٠}$ والعكس بالعكس ، فى معادلة الرقم القياسى ؛ وكذلك $\text{ك}_{١}$ بدل $\text{ك}_{٠}$ إذا وجدت .

∴ الرقم $\frac{\text{مح ع}_{١}}{\text{مح ع}_{٠}}$ يكون بديله الزمنى $\frac{\text{مح ع}_{٠}}{\text{مح ع}_{١}}$

وهذا البديل يعبر عن أسعار السنة . بالنسبة إلى أسعار السنة ١ .

وكذلك الصيغة $\frac{\text{مح ع}_{١}\,\text{ك}_{٠}}{\text{مح ع}_{٠}\,\text{ك}_{٠}}$ يكون بديلها الزمنى $\frac{\text{مح ع}_{٠}\,\text{ك}_{١}}{\text{مح ع}_{١}\,\text{ك}_{١}}$.

» » $\frac{\text{مح ع}_{١}\,\text{ك}_{١}}{\text{مح ع}_{٠}\,\text{ك}_{١}}$ » » » $\frac{\text{مح ع}_{٠}\,\text{ك}_{٠}}{\text{مح ع}_{١}\,\text{ك}_{٠}}$.

شرط الانعكاس فى الزمن

٣١٨ — ولكى يكون أى رقم مستوفياً شرط **الانعكاس فى الزمن** (٢) ، يجب إذن أن يكون حاصل ضرب هذا الرقم فى بديله الزمنى يساوى الواحد تماماً . فلنختبر إذن بعض المعادلات التى حصلنا عليها فى الباب السابق لنرى إذا ما كانت تفى بهذا الشرط أو لا .

---

(١) يسمى بالإنجليزية Time Reciprocal

(٢) اسمه بالإنجليزية Time Reversal Test

اختبار الأرقام التجميعية

٣١٩ — من الواضح أن الرقم التجميعى البسيط يستوفى هذه الشروط ، لأن :

$$\frac{\text{مح ع}_{٠}}{\text{مح ع}_{١}} \times \frac{\text{مح ع}_{١}}{\text{مح ع}_{٠}} \ \text{(وهو البديل الزمنى)} = ١$$

أما الرقم التجميعى المرجح ، سواء بكميات السنة الأساسية أو بكميات السنة المقارنة ، فلا يستوفى هذا الشرط . لأن

$$\frac{\text{مح ع}_{١}\ \text{ك}_{٠}}{\text{مح ع}_{٠}\ \text{ك}_{٠}} \times \text{بديله الزمنى وهو} \ \frac{\text{مح ع}_{٠}\ \text{ك}_{١}}{\text{مح ع}_{١}\ \text{ك}_{١}} \neq ١ \ ؛$$

$$\text{وكذلك} \ \frac{\text{مح ع}_{١}\ \text{ك}_{١}}{\text{مح ع}_{٠}\ \text{ك}_{١}} \times \text{» \quad » \quad »} \ \frac{\text{مح ع}_{٠}\ \text{ك}_{٠}}{\text{مح ع}_{١}\ \text{ك}_{٠}} \neq ١ \ \text{أيضاً .}$$

الرقم الأمثل ينعكس فى الزمن

٣٢٠ — أما الصيغة المشتقة من هاتين الأخيرتين ، والتى يسميها فيشر الرقم القياسى الأمثل كما ذكرنا فى بند ٢٨٧ ، فهى تستوفى هذا الشرط . لأن هذه الصيغة وهى :

$$\sqrt{\frac{\text{مح ع}_{١}\ \text{ك}_{٠}}{\text{مح ع}_{٠}\ \text{ك}_{٠}} \times \frac{\text{مح ع}_{١}\ \text{ك}_{١}}{\text{مح ع}_{٠}\ \text{ك}_{١}}} \ \text{، بديلها الزمنى هو} \ \sqrt{\frac{\text{مح ع}_{٠}\ \text{ك}_{١}}{\text{مح ع}_{١}\ \text{ك}_{١}} \times \frac{\text{مح ع}_{٠}\ \text{ك}_{٠}}{\text{مح ع}_{١}\ \text{ك}_{٠}}}$$

وواضح أن حاصل ضرب هاتين الكميتين = ١ تماماً .

وكذلك الصيغة التى أوردناها فى بند ٢٨٨ ، وهى :

$$\frac{\text{مح ع}_{١}\ (\text{ك}_{٠} + \text{ك}_{١})}{\text{مح ع}_{٠}\ (\text{ك}_{٠} + \text{ك}_{١})}$$

تنعكس فى الزمن أيضاً .

الوسط الحسابى البسيط أو المرجح ، لا ينعكس فى الزمن

٣٢١ — الوسط الحسابى للمناسيب ، سواء كان بسيطاً أو مرجحاً بأى الأوزان التى ذكرناها فى بند ٢٩٠ ، لا ينعكس فى الزمن . فالوسط البسيط مثلا ، وهو $\frac{١}{\text{ن}}$ (مح س) حيث $\text{س} = \text{ع}_{١} \div \text{ع}_{٠}$ ، لا ينعكس لأن

$$\frac{١}{\text{ن}}\left(\frac{\text{ع}_{١}}{\text{ع}_{٠}} + \frac{\text{ع}'_{١}}{\text{ع}'_{٠}} + \frac{\text{ع}''_{١}}{\text{ع}''_{٠}} + \cdots\right) \times \frac{١}{\text{ن}}\left(\frac{\text{ع}_{٠}}{\text{ع}_{١}} + \frac{\text{ع}'_{٠}}{\text{ع}'_{١}} + \frac{\text{ع}''_{٠}}{\text{ع}''_{١}} + \cdots\right) \neq ١ \ .$$

ويلاحظ هنا أن البديل الزمنى للوسط الحسابى للمناسيب يساوى مقلوب وسطها التوافقى .

وكذلك الوسط المرجح بالأوزان ا المذكورة فى بند ٢٩٠ ، وهو

$$\frac{\text{مج}\ \text{س}\ \text{م}_{.}}{\text{مج}\ \text{م}_{.}} = \frac{\text{مج}\ \text{ع}_{1}\ \text{ك}_{.}}{\text{مج}\ \text{ع}_{.}\ \text{ك}_{.}} \text{، لأن } \text{س} = \frac{\text{ع}_{1}}{\text{ع}_{.}} \text{ ، } \text{م}_{.} = \text{ع}_{.}\ \text{ك}_{.} \text{ ؛}$$

وقد أثبتنا ، فى البند السابق ، أن هذا الأخير لا ينعكس . وبالمثل نجد $\frac{\text{مج}\ \text{س}\ \bar{\text{م}}_{.}}{\text{مج}\ \bar{\text{م}}_{.}}$ لا ينعكس أيضاً . وكذلك الوسطان المرجحان بالأوزان حـ ، ء المذكوران فى بند ٢٩٠ ، لا ينعكسان فى الزمن .

الوسط التوافقى لا ينعكس فى الزمن

٣٢٢ — والوسط التوافقى للمناسيب ، بسيطاً كان أو مرجحاً بأى أوزان ، لا ينعكس أيضاً ؛ بدليل أن مقلوبه يساوى البديل الزمنى للوسط الحسابى ، كما لاحظنا فى البند السابق . أى أن الوسط الحسابى هو البديل الزمنى لمقلوب الوسط التوافقى . وعدم انعكاس أحدهما يقتضى عدم انعكاس الآخر .

الوسط الهندسى البسيط ينعكس فى الزمن

٣٢٣ — أما الوسط الهندسى للمناسيب فلا ينعكس فى الزمن إلا إذا كان بسيطاً غير مرجح بأى نوع من الأوزان . بدليل أن الوسط الهندسى البسيط ، وهو

$$\sqrt[\text{ن}]{\frac{\text{ع}_{1}}{\text{ع}_{.}} \times \frac{\bar{\text{ع}}_{1}}{\bar{\text{ع}}_{.}} \times \frac{\bar{\bar{\text{ع}}}_{1}}{\bar{\bar{\text{ع}}}_{.}} \times \cdots} \text{، بديله الزمنى هو } \sqrt[\text{ن}]{\frac{\text{ع}_{.}}{\text{ع}_{1}} \times \frac{\bar{\text{ع}}_{.}}{\bar{\text{ع}}_{1}} \times \frac{\bar{\bar{\text{ع}}}_{.}}{\bar{\bar{\text{ع}}}_{1}} \times \cdots}$$

ومن الواضح أن حاصل ضرب هاتين الكميتين يساوى ١ تماماً .

وإذا أخذنا أى واحد من الأوساط الهندسية المرجحة المذكورة فى بند ٢٩٢ نجد أنه لا ينعكس . فمثلاً نأخذ :

$$\sqrt[\text{مجـ م}_٠]{(س)^{م_٠} \times (سَ)^{مَ_٠} \times (سً)^{مً_٠} \times \cdots}$$

$$\text{أى} \left[\left(\frac{ع_١}{ع_٠}\right)^{ع_٠ ك_٠} \times \left(\frac{عَ_١}{عَ_٠}\right)^{مَ كَ_٠} \times \left(\frac{عً_١}{عً_٠}\right)^{مً كً_٠} \cdots \right]^{\frac{١}{\text{مجـ ك}_٠ ع_٠}}$$

ولو وضعنا $ع_١$ بدل $ع_٠$ ، $ع_٠$ بدل $ع_١$ فى هـذه المعادلة ، وكذلك بالنسبة إلى $ك_٠$ و $ك_١$ ، نحصل على البديل الزمنى لهذه المعـادلة . ومن الواضح أن حاصل ضرب هذه المعادلة فى هذا البديل الزمنى لا يساوى الواحد أبداً .

نستبعد المعادلات التى لا تنعكس فى الزمن .

٣٢٤ — هـذا الاختبار الزمنى إذن أداة نافعة نختبر بهـا الأرقام القياسية فى أشـكالها المختلفة . ونستبعد منها ما لا يصلح فلا نستعمله ، ونستبقى منها ما كان مستوفياً لشروط هذا الاختيـار . وبذلك نضمن صحة المقارنة التى نجريها بواسـطة هذه الأرقام القياسية . أو على الأقل نضمن سلامتها من الأخطاء المترتبة على عدم استيفاء الرقم القياسى لشروط هذا الاختبار .

ولـكن هذا الاختبار لا يمس إلا ناحيـة واحدة فقط ، وهى نسبة الأسعار فى تاريخين (أو مكانين) مختلفين . وهناك ناحية أخرى لا تقل عن هذه فى الأهمية . فالأرقام القياسية ، كما قلنا سابقاً ، لا تستخدم فقط لبيان حركة أسعار السلع ، بل تسـتخدم أيضاً فى بيان التغيرات فى كميات السلع ، وفى غيرهـا من الظواهر مثل الأجور والإنتاج وهكذا .

منسوب سعر سلعـة فى منسوب كميتها يسـاوى منسـوب القيمـة

٣٢٥ — ولو حسبنا الرقم القياسى (منسوب) لسعر سلعة واحدة فى سنتين مختلفتين واحدة بالنسـبة إلى الأخرى ، وحسبنا أيضاً الرقم القيـاسى (منسوب) لـكمية هذه السلعة فى نفس السنتين ، فمن البدهى أن حاصل ضرب هذين الرقمين يساوى النسبة بين قيمتى هذه السلعة فى نفس السنتين ، واحدة بالنسبة إلى الأخرى ، وإلا كان الرقم القياسى المستعمل خاطئاً فى تصويره للأحوال .

شـــرط الانعكــاس فى المعــامل

٣٢٦ — وقياساً على ذلك إذا استعملنا أى صيغة من الصيغ التى نعرفها للأرقام القياسية فى حساب الرقم القياسى لأسعـار عدة سلع فى سنتين مختلفتين واحـدة بالنسبة إلى الأخرى ، واستعملنا نفس الصيغة فى حساب الرقم القيـاسى لـكميات هذه السلع فى نفس السـنتين ، يجب أن يكون حاصل ضرب هذين الرقمين القياسـيين ، للأسعار والـكميات ، مساوياً للنسبة بين قيم هـذه السلع (القيمة = الـكمية × السعر) فى نفس السنتين تحت البحث . وخلاف ذلك تـكون الصيغة المستعملة فى حساب هذين الرقمين القياسيين صيغة خاطئة ولا تصلح للاستعمال . وهكذا يكون لدينا اختبار آخر نمتحن بواسطته ما عندنا من الأرقام القياسية ، ونستبعد منهـا ما لا يستوفى هذا الشرط .

هذا الشرط يسمى شرط **الانعكـاس فى المعامل أو الانعكـاس المعاملى**[١] ويمكن وضعه فى صورة مختصرة كما يأتى :

نفرض أن الأسعار الأساسية والمقارنة هى ع. و ع١ ؛ وكذلك الكميات الأساسية والمقارنة ك. و ك١ ، فيكون مجموع قيم هذه السلع فى السنة الأساسـية والسنة المقارنة هو على الترتيب مج ع. ك. و مج ع١ ك١ .

∴ يكون منسوب القيم $= \dfrac{\text{مج ع}_1\ \text{ك}_1}{\text{مج ع}_0\ \text{ك}_0}$.

ونفرض أن **البديل المعاملى** لأى معادلة من معادلات الأرقام القياسـية ، هو نفس المعادلة موضوعاً فيهـا ع بدل ك وك بدل ع ، أى نضع الـكمية بدل السعر والعكس بالعـكس . فإذا كانت المعادلة الأصلية تعـبر عن رقم قياسى للأسعار ، يكون البديل المعاملى هو الرقم القياسى للكميات ، مركباً بنفس الطريقة ، والعكس بالعكس .

---

(١) بالانجليزية Factor Reversal Test

وبذلك يكون الشرط الواجب تحققه فى كل رقم قياسى ليكون صالحاً للاستعمال هو :

$$\text{الرقم القياسى} \times \text{البديل المعاملى له} = \frac{\text{مجـ ع}_{١}\,\text{ك}_{١}}{\text{مجـ ع}_{٠}\,\text{ك}_{٠}} \text{ ،}$$

وهذا هو شرط الانعكاس المعاملى .

تطبيق الاختبارين على كل الأرقام القياسية

٣٢٧ — وهكذا يكون لدينا اختباران أساسيان ، نستخدمهما فى انتقاء الأرقام القياسية الصالحة واستبعاد غيرها . والرقم الصالح للاستعمال إذن هو الرقم الذى يستوفى هذين الاختبارين فى وقت واحد . فنختبر كل الأرقام التى عرفناها بالنسبة للانعكاس الزمنى ونستبعد ما لا ينعكس منها ؛ ثم نختبر الباقى بالنسبة إلى الانعكاس المعاملى . وسواء اختبرنا الأرقام بهذا الترتيب أو بالعكس ، فالنتيجة واحدة فى الحالتين . وبذلك نحصل على الأرقام القياسية التى تستوفى الاختبارين معاً ، فتكون صالحة للاستعمال .

٣٢٨ — الآن نختبر بعض الأرقام القياسية التى عرفناها ، بالنسبة إلى الانعكاس المعاملى . وسنكتفى هنا باختبار تلك الأرقام التى ثبت لنا أنها تنعكس فى الزمن . وطريقة اختبار أى رقم هى أن نضربه فى بديله المعاملى ؛ فإذا كان حاصل الضرب مساوياً منسوب القيم ، كان الرقم منعكساً .

اختبار بعض الأرقام

٣٢٩ — الرقم التجميعى البسيط (ينعكس فى الزمن) لا ينعكس فى المعامل لأنه يساوى

$$\frac{\text{مجـ ع}_{١}}{\text{مجـ ع}_{٠}} \text{ ، وبديله المعاملى يساوى } \frac{\text{مجـ ك}_{١}}{\text{مجـ ك}_{٠}} \text{ .}$$

$$\text{ولكن } \quad \frac{\text{مجـ ع}_{١}}{\text{مجـ ع}_{٠}} \times \frac{\text{مجـ ك}_{١}}{\text{مجـ ك}_{٠}} \neq \frac{\text{مجـ ع}_{١}\,\text{ك}_{١}}{\text{مجـ ع}_{٠}\,\text{ك}_{٠}} \text{ .}$$

∴ هذا الرقم لا ينعكس فى المعامل .

الرقم الأمثل ينعكس فى المعامل أيضاً

٣٣٠ — نأخذ الرقم الأمثل الذى وضعه فيشر ، وهو

$\sqrt{\frac{\text{مجـ}\,ع_1\,ك_0}{\text{مجـ}\,ع_0\,ك_0} \times \frac{\text{مجـ}\,ع_1\,ك_1}{\text{مجـ}\,ع_0\,ك_1}}$ ، وبديله المعاملى وهو $\sqrt{\frac{\text{مجـ}\,ك_1\,ع_0}{\text{مجـ}\,ك_0\,ع_0} \times \frac{\text{مجـ}\,ك_1\,ع_1}{\text{مجـ}\,ك_0\,ع_1}}$

وواضح أن حاصل ضرب هاتين الكميتين $= \frac{\text{مجـ}\,ع_1\,ك_1}{\text{مجـ}\,ع_0\,ك_0}$.

أى أن هذا الرقم الأمثل ينعكس فى المعامل وفى الزمن أيضاً . والواقع أنه الرقم الوحيد ، بين الأرقام التى ذكرناها كلها الذى يستوفى شرطى الانعكاس معاً . وهو لهذا السبب أحسن الأرقام القياسية وأصحها من الناحية النظرية — والعملية أيضاً ، إلى حد ما . ومن ثم كانت تسميته[1] **بالرقم الأمثل**

الوسط الهندسى لا ينعكس فى المعامل

٣٣١ — الوسط الهندسى البسيط لا ينعكس فى المعامل ، لأنه يساوى :

$\sqrt[ن]{\frac{ع_1}{ع_0} \times \frac{\bar{ع}_1}{\bar{ع}_0} \times \cdots}$ ؛ وبديله المعاملى يساوى $\sqrt[ن]{\frac{ك_1}{ك_0} \times \frac{\bar{ك}_1}{\bar{ك}_0} \times \cdots}$ ؛

وحاصل ضرب هاتين الكميتين $\neq \frac{\text{مجـ}\,ع_1\,ك_1}{\text{مجـ}\,ع_0\,ك_0}$ منسوب القيم .

ومن السهل إثبات أن باقى الأرقام القياسية التى ذكرناها فى هذا الباب وفى الباب السابق ، بسيطة كانت أو مرجحة بأى نوع من الأوزان لا تنعكس فى المعامل .

ومن هذا يظهر لنا أن الاختبار الثانى — الانعكاس فى المعامل — هو فى الواقع اختبار « أصعب » من الاختبار الأول أو الزمنى ، حيث بعض الأرقام تمر فى الاختبار الأول ولا تمر فى الثانى .

أخطاء الأرقام القياسية

٣٣٢ — لنبحث الآن فيما يترتب على عدم استيفاء شروط هذين الاختبارين .

---

(١) بالانجليزية Ideal Index

الوسط الحسابى فى بديله الزمنى أكبر من ١

لنأخذ مثلا الوسط الحسابى (البسيط) للمناسيب . فقد رأينا أن بديله الزمنى يساوى مقلوب الوسط التوافقى ؛ وأنه لاينعكس فى الزمن . وعلى ذلك فخارج قسمة الوسط الحسابى على الوسط التوافقى لا يساوى ١ ، فهل هو أكبر أو أقل من ١ ؟

وجدنا فى المثال الذى أخذناه فى الباب السابق (صحيفة ٣٢١) أن الوسط الحسابى لمناسيب أسعار أربعة محاصيل هو ١١٣٫٨٩٪ ، وأن الوسط التوافقى لها ١١٣٫٠٦ ، خارج قسمة هذين (الأول على الثانى) يساوى ١٫٠٠٧ ، أى أن حاصل الضرب (الوسط × بديله الزمنى) فى هذا المثال أكبر من ١

والحقيقة أن حاصل ضرب الوسط الحسابى للمناسيب فى بديله الزمنى دائماً يكون أكبر من ١ ، وهذا يمكن إثباته جبرياً بسهولة [١] . وبالعكس نجد أن الوسط التوافقى مضروباً فى بديله الزمنى يعطى ناتجاً أقل من ١ ، ويمكننا أيضاً إثبات مثل هذا بالنسبة للوسط المرجح بأوزان ا ، ب ، حـ ، أو ء (بند ٢٩٠) . وللوسط التوافقى أيضاً [٢] .

---

(١) لإثبات ذلك نضرب الوسط الحسابى فى بديله الزمنى ، مع العلم بأن البديل الزمنى للمنسوب س هو ١ ÷ س

$$\therefore \frac{\text{١}}{\text{ن}}\left(\text{س} + \text{سَ} + \text{سً} + \cdots\right) \times \frac{\text{١}}{\text{ن}}\left(\frac{\text{١}}{\text{س}} + \frac{\text{١}}{\text{سَ}} + \frac{\text{١}}{\text{سً}} + \cdots\right)$$

$$= \frac{\text{١}}{\text{ن}^{\text{٢}}}\left[\text{ن} + \frac{\text{س}}{\text{سَ}} + \frac{\text{سَ}}{\text{س}} + \frac{\text{س}}{\text{سً}} + \frac{\text{سً}}{\text{س}} + \cdots\right]$$

$$\therefore \text{حاصل الضرب} = \frac{\text{١}}{\text{ن}^{\text{٢}}}\left[\text{ن} + \frac{\text{ن}\,(\text{ن} - \text{١})}{\text{٢}}\ \text{حدا مثل}\ \left(\frac{\text{س}}{\text{سَ}} + \frac{\text{سَ}}{\text{س}}\right)\right.$$

وهو عدد توافيق الحروف س ، سَ ، سً ... ، مثنى مثنى . وكل من هذه الحدود أكبر من ٢ . لأن $\text{حـ} + \frac{\text{١}}{\text{حـ}} > \text{٢}$ ، لأن $\text{حـ}^{\text{٢}} + \text{١} > \text{٢حـ}$ لأن $(\text{حـ} - \text{١})^{\text{٢}} > \text{٠}$]

$$\therefore \text{حاصل الضرب} > \frac{\text{١}}{\text{ن}^{\text{٢}}}\left[\text{ن} + \text{ن}^{\text{٢}} - \text{ن}\right]$$ ؛ أى $> \text{١}$ وهو المطلوب .

(٢) البرهان فى هذه الحالة أكثر تعقيداً منه فى الحالة المذكورة . أنظر ص ٢٨٦ من كتاب I. Fisher, *The Making of Index Numbers.*

تحيز الوسطين الحسابى والتوافقى

٣٣٣— وينتج من ذلك أن الوسط الحسابى للمناسيب ، البسيط أو المرجح ، **متحيز إلى أعلى**[١] ، والوسط التوافقى **متحيز إلى أسفل**[٢] ، ومعنى ذلك أن الوسط الحسابى للمناسيب ، فى دلالته على مستوى الأسعار فى سنة معينة بالنسبة إلى أخرى ، يميل إلى إظهار هذا المستوى أعلى مما هو فى الحقيقة . وبالعكس مع الوسط التوافقى ، فهو يميل إلى تصوير مستوى الأسعار أقل مما هو فى الحقيقة .

فنرى إذن أن عدم انعكاس الرقم القياسى فى الزمن ، وهو ناتج من خطأ فى صيغته أو كيفية تركيبه ، يترتب عليه تحيز الرقم ، وعدم دقته فى تصوير الأشياء على حقيقتها . ونخطىء إذا نحن اعتمدنا على مثل هذا الرقم .

وكذلك عدم الانعكاس فى المعامل يترتب عليه تحيز الأرقام القياسية .

تحيز الأوزان

٣٣٤ — وقد ينشأ التحيز أيضاً عن نوع الأوزان المستعملة فى ترجيح الرقم القياسى ، وكيفية إدخالها أو استعمالها فى المعادلة التى تمثله . فلنأخذ مثلا الوسط الحسابى للمناسيب ، وهو كما نعلم بدون أوزان $= \frac{\text{مجـ س}}{\text{ن}}$

$$= \frac{1}{\text{ن}}\left[\frac{\text{ع}_1}{\text{ع}_0} + \frac{\text{ع}'_1}{\text{ع}'_0} + \frac{\text{ع}''_1}{\text{ع}''_0} + \cdots\right]$$

فاذا رجحناه بالأوزان $\text{م}_1 = \text{ع}_1\text{ك}_0$ يصبح

$$\frac{\text{مجـ س م}_1}{\text{مجـ م}_1} = \frac{1}{\text{مجـ م}_1}\left[\left(\frac{\text{ع}_1}{\text{ع}_0}\right)\text{ع}_1\text{ك}_0 + \left(\frac{\text{ع}'_1}{\text{ع}'_0}\right)\text{ع}'_1\text{ك}'_0 + \cdots\right] ،$$

أى أننا نرجح المناسيب الكبيرة ، التى فيها البسط ع_1 كبيراً بالنسبة إلى ع_0 ، بأوزان كبيرة أيضاً ؛ وفى الوقت نفسه نرجح المناسيب الصغيرة ، التى فيها

---

(١) بالإنجليزية Biased Upwards (٢) Biased Downwards

$ع_{١}$ صغيرة بالنسبة إلى $ع_{٠}$ ، بأوزان صغيرة أيضاً . وبذلك تتحيز للمناسيب ضد المناسيب الصغيرة . وينتج إذن أن هذا الرقم متحيز إلى أعلى بسبب الأوزان .

وكذلك الوسط المرجح بالأوزان $م_{١٢} = ع_{١} ك_{١}$ ، أى الأوزان ء بند ٢٩٠ يكون متحيزاً إلى أعلى أيضاً .

ومن هـذا يتبين لنا أن الوسط الحسابى لمناسيب الأسعار المرجح بأوزان $مـ_{١}$ أو $م_{١٢}$ ، يكون بطبيعة تركيبه متحيزاً إلى أعلى حتما سواء فى ذلك أكانت الأسعار فى ارتفاع أو فى هبوط .

المرجح ا أو ب ، أصغر من المرجح حـ أو ء

٣٣٥ — وإذا قارنا بين متوسط المناسيب المرجح ا أو ب (أى بأوزان $م_{٠}$ أو $مـ_{٠}$) ومتوسطها المرجح حـ أو ء (أوزان $مـ_{١}$ أو $م_{١}$) ، نجد أن المتوسط الأول دائماً أصغر من المتوسط الثانى . وذلك سواء كانت حركة الأسعار فى هبوط أو صعود .

لأنه عند استعمال الأوزان $م_{٠}$ أو $مـ_{٠}$ نضرب المناسيب $\frac{ع_{١}}{ع_{٠}}$ فى $ع_{٠} ك_{٠}$ أو فى $ع_{٠} ك_{١}$ ، على الترتيب . وبذلك تختفى الأسعار $ع_{٠}$ وتبقى الحواصل $ع_{١} ك_{٠}$ أو $ع_{١} ك_{١}$ أما عند استعمال الأوزان $مـ_{١}$ أو $م_{١}$ فنضرب نفس المناسيب $\frac{ع_{١}}{ع_{٠}}$ فى $ع_{١} ك_{٠}$ أو $ع_{١} ك_{١}$. على الترتيب . وتكون النتيجة ، كما قلنا فى البند السابق ، أن المناسيب الكبيرة تضرب فى أوزان كبيرة أيضاً ، والمناسيب الصغيرة تضرب فى أوزان صغيرة ، وبذلك يكون المتوسط الناتج أكبر من المتوسط الأول حتما ، سواء ارتفعت الأسعار أو هبطت .

أما العلاقة بين المتوسطين المرجحين بأوزان ا و ب ، أو بين المتوسطين المرجحين حـ و ء ، فلا يمكن تحديدها بطريقة جازمة كما فعلنا هنا ، إذ يصح أن يكون ا أكبر أو أصغر من المتوسط ب . وكذلك المتوسطان حـ و ء

تحيز الوسط الهندسى المرجح

٣٣٦ — لنبحث الآن فى الوسط الهندسى المرجح وما يترتب على عدم انعكاسه فى الزمن .

نرمز إلى الوسط الهندسى البسيط بالحرف هـ ، وللوسط الهندسى المرجح بالأوزان أ و ب و حـ و د بالرموز الآتية [١] على الترتيب :

$$\text{ها} ، \text{هب} ، \text{هج} ، \text{هد}$$

ونرمز للبديلات الزمنية لهذه المتوسطات بالرموز

$$\text{ها}_1 ، \text{هب}_1 ، \text{هج}_1 ، \text{هد}_1$$

فنعلم أن

$$\text{ها} = \sqrt[\text{مج}\,\text{م}_0]{(\text{س})^{\text{م}_0} \times (\text{سَ})^{\text{مَ}_0} \times (\text{سً})^{\text{مً}_0} \times \cdots} \quad (1)\text{؛}$$

وأن

$$\text{ها}_1 = \sqrt[\text{مج}\,\text{م}_1]{(\text{س})^{-\text{م}_1} \times (\text{سَ})^{-\text{مَ}_1} \times (\text{سً})^{-\text{مً}_1} \times \cdots} \quad (2)\text{؛}$$

لأن $\text{م}_0 = \text{ع}_0\,\text{ك}_0$ ، وبديلها الزمنى $\text{م}_1 = \text{ع}_1\,\text{ك}_1$ ،

و $\text{س} = \frac{\text{ع}_1}{\text{ع}_0}$ » $= \frac{\text{ع}_0}{\text{ع}_1} = \text{س}^{-1}$.

فلكى ينعكس ها فى الزمن يجب أن يكون $\text{ها} \cdot \text{ها}_1 = 1$

أو بعبارة أخرى ، يجب أن يكون لوغاريتم هذا الحاصل يساوى صفراً .

وإذا كان $\text{ها} \cdot \text{ها}_1 = 1$ ، فإن $\text{لو ها} \cdot \text{ها}_1 = 0$

» » $\text{ها} \cdot \text{ها}_1 < 1$ » $\text{لو ها} \cdot \text{ها}_1 < 0$

» » $\text{ها} \cdot \text{ها}_1 > 1$ » $\text{لو ها}\ \text{ها}_1 > 0$

---

(١) تنطق هذه الرموز هكذا : هـ ألف ، هـ باء ، هـ جيم ، هـ دال .

ولكن

$$\text{لو}\ \text{ها} = \frac{1}{\text{مجـ}\ \text{م}_{.}} \left(\text{م}_{.}\ \text{لو}\ \text{س} + \text{م}'_{.}\ \text{لو}\ \text{س}' + \text{م}''_{.}\ \text{لو}\ \text{س}'' + \cdots \right) \cdots (٣) ؛$$

$$\text{و} - \text{لو}\ \text{ها}_{1} = \frac{1}{\text{مجـ}_{1}} \left(\text{م}_{1}\ \text{لو}\ \text{س} + \text{م}'_{1}\ \text{لو}\ \text{س}' + \text{م}''_{1}\ \text{لو}\ \text{س}'' + \cdots \right) \cdots (٤) .$$

والطرف الأيسر فى كل من هاتين المتساويتين عبارة عن وسط مرجح للوغاريتمات المناسيب ، بأوزان $\text{م}_{.}$ و م_{1} على الترتيب . ونرى فى (٤) ، كما فى البندين السابقين ، أن المناسيب الكبيرة مرجحة بأوزان كبيرة أيضاً ، والمناسيب الصغيرة مرجحة بأوزان صغيرة أيضاً .

وينتج من ذلك أن الطرف الأيسر من المتساوية (٤) أكبر من الطرف الأيسر فى (٣) . وسواء فى ذلك ارتفعت الأسعار أو هبطت .

$$\therefore\ \text{لو}\ \text{ها} . \text{ها}_{1} = \text{لو}\ \text{ها} + \text{لو}\ \text{ها}_{1}$$

$$= \text{كمية سالبة} .$$

$$\therefore\ \text{ها} . \text{ها}_{1} < 1 .$$

وعلى ذلك يكون الوسط الهندسى المرجح بأوزان ا متحيزاً إلى أسفل دائماً .

٣٣٧ — نأخذ الوسط المرجح هب ، ونثبت بنفس الطريقة أنه متحيزاً إلى أسفل أيضاً . المعادلة الجبرية لهذا الوسط هى كما نعلم : الوسط هب متحيز لأسفل

$$\text{هب} = \sqrt{(\text{س})^{\text{م}_{.}} \times (\text{س}')^{\text{م}'_{.}} \times (\text{س}'')^{\text{م}''_{.}} \times \cdots} \qquad (١)$$

وبديله الزمنى هو

$$\text{هب}_{1} = \sqrt[\text{مجـ}_{1}]{(\text{س})^{-\text{م}_{1}} \times (\text{س}')^{-\text{م}'_{1}} \times (\text{س}'')^{-\text{م}''_{1}} \times \cdots} \qquad (٢)$$

حيث $\text{م}_{.} = \text{ع}_{.}\,\text{ك}_{١}$ ، وبديلها الزمنى $\text{م}_{١} = \text{ع}_{١}\,\text{ك}_{.}$ ؛

و $\text{س} = \frac{\text{ع}_{.}}{\text{ع}_{١}}$ « « $\frac{\text{ع}_{.}}{\text{ع}_{١}} = \text{س}^{-١}$.

$$\therefore \text{لو هب} = \frac{١}{\text{مج}\,\text{م}_{.}}\left(\text{م}_{.}\,\text{لوس} + \bar{\text{م}}_{.}\,\text{لو}\bar{\text{س}} + \bar{\bar{\text{م}}}_{.}\,\text{لو}\bar{\bar{\text{س}}} + \ldots\right) \ldots (٣) \text{؛}$$

$$\text{و} -\text{لوهب}_{١} = \frac{١}{\text{مج}\,\text{م}_{١}}\left(\text{م}_{١}\,\text{لوس} + \bar{\text{م}}_{١}\,\text{لو}\bar{\text{س}} + \bar{\bar{\text{م}}}_{١}\,\text{لو}\bar{\bar{\text{س}}} + \ldots\right) \ldots (٤) .$$

وهنا أيضاً نرى أن المناسيب الكبيرة مرجحة بأوزان كبيرة ، والعكس بالعكس . وينتج من ذلك أن الطرف الأيسر من (٤) أكبر من نظيره فى (٣) .

$$\therefore \quad \text{لو هب}\,.\,\text{هب}_{١} = \text{كمية سالبة}$$

ويكون $\quad \text{هب}\,.\,\text{هب}_{١} < ١$

أى أن الوسط المرجح **هب** متحيز إلى أسفل .

٣٣٨ — وبنفس البرهان نثبت أن الوسطين **هج** و **هد** متحيزان إلى أعلى ، أى أن	الوسطان هج وهد متحيزان إلى أعلى

$$\text{هج} \times \text{هج}_{١} > ١ \quad \text{؛}$$

$$\text{و} \quad \text{هد} \times \text{هد}_{١} > ١ \quad \text{،}$$

حيث $\quad \text{هج} = \sqrt[\text{مج}\,\text{م}_{١}]{(\text{س})^{\text{م}_{١}} \times (\bar{\text{س}})^{\bar{\text{م}}_{١}} \times \ldots}$

وبديله الزمنى $\quad \text{هج}_{١} = \sqrt[\text{مج}\,\text{م}_{.}]{(\text{س})^{-\text{م}_{.}} \times (\bar{\text{س}})^{-\bar{\text{م}}_{.}} \times \ldots}$

وكذلك $\quad \text{هد} = \sqrt[\text{مج}\,\text{م}_{١}]{(\text{س})^{١} \times (\bar{\text{س}})^{\bar{\text{م}}_{١}} \times \ldots}$

وبديله الزمنى $\text{هـ}_{١} = \sqrt[\text{مجـ م.}]{(\text{س})^{-\text{م}_{٢.}} \times (\text{س})^{-\text{م}_{٣.}} \times \ldots}$

٣٣٩ — واضح من هذه البراهين أن تحيز الوسط الهندسى ناتج عن الأوزان المستعملة ، لاسيما وأننا نعرف أن الوسط الهندسى من الأصل ينعكس فى الزمن إذا لم يرجح بأى نوع من الأوزان . وخاصة الانعكاس فى الزمن التى يمتاز بها الوسط الهندسى البسيط ، هى السبب فى تفضيله على غيره من الأرقام القياسية فى كثير من الأبحاث الخاصة بالأسعار . وقد ذكرنا سابقاً أنه الرقم القياسى المطبق عملياً فى كثير من الدول لبيان حركة أسعار **الجملة** .

عدم الانعكاس فى المعامل أكثر تعقيداً

٣٤٠ — تكلمنا هنا عن الأخطاء التى تنشأ عن عدم انعكاس الرقم القياسى فى الزمن . ولم نبحث فى الأخطاء الأخرى الناشئة عن عدم الانعكاس فى المعامل . والبحث فى هذه الأخطاء أكثر تعقيداً من البحث الذى أوردناه فى البنود الأخيرة من هذا الباب .

الاختبار الدورى

٣٤١— عدا الاختبارين اللذين بحثناهما فى هذا الباب ، يوجد اختبار ثالث يسمى (١) **الاختبار الدورى** . والفكرة الأساسية فى هذا الاختبار تتلخص فيما يأتى : إذا كان لدينا أسعار سلعة واحدة فى ثلاث سنين (أو أماكن) معينة ، السنين ١ و ٢ و ٣ مثلا ، وكانت مناسيب أسعارها لهذه السنين هى :

$$\text{س}_{٢١} \quad \text{و} \quad \text{س}_{٣٢} \quad \text{و} \quad \text{س}_{٣١}$$

حيث الرقم الأول يدل على السنة المعتبرة أساساً ، فإننا نجد دائماً أن :

$$\frac{\text{ع}_{٢}}{\text{ع}_{١}} \times \frac{\text{ع}_{٣}}{\text{ع}_{٢}} = \text{س}_{٢١} \times \text{س}_{٣٢} = \text{س}_{٣١} \text{ ؛}$$

أى أن

$$\text{س}_{٢١} \times \text{س}_{٣٢} \times \text{س}_{١٣} = ١$$

(١) بالإنجليزية Circular Test

والاختبار الدورى يرمى إلى أن الرقم القياسى لأسعار عدة سلع يجب أن يستوفى هذا الشرط . أى أنه ، بفرض ع هى الرقم القياسى ، يجب أن يكون

$$ع_{٢١} \times ع_{٣٢} \times ع_{١٣} = ١$$

الاختبار الدورى مقبول شكلا

٣٤٢ — هذا الاختبار يبدو معقولا فى ظاهره ، ولكنه من الناحية النظرية خطأ ولا يمكن أن يستوفى هذا الشرط بأى رقم قياسى صالح . بل الواجب أن الرقم القياسى الصحيح لا يستوفى هذا الاختبار . وذلك لأن أوزان السلع التى تدخل فى تركيب الرقم القياسى فى السنة ٢ بالنسبة إلى السنة ١ مثلا ، تخالف الأوزان المستعملة لنفس السلع بين السنتين ٢ و ٣ ، أو السنتين ٣ و ١ ، هذا فضلا عن أن السلع نفسها قد تختلف ، حيث تدخل سلع جديدة وتسقط سلع قديمة . خصوصاً عند مقارنة الأسعار بين الممالك المختلفة (بدل السنين) ، حيث يجوز أن تكون السلع المهمة فى المقارنة بين البلدين ١ و ٢ مثلا ، والتى تدخل فى تركيب الرقم القياسى بينهما ، غير مهمة ، أو غير موجودة بالمرة ، فى المقارنة بين المملكتين ٢ و ٣ أو بين ٣ و ١ ، وعلى ذلك لا يصح نظرياً أن يتحقق الاختبار الدورى للأرقام القياسية بين هذه البلاد . والواجب أن هذا الاختبار لا يتحقق فى مثل هذه الأحوال .

الاختبار الدورى خطأ نظرياً . وهو صحيح فقط إذا كانت الأوزان واحدة

وأى رقم قياسى يتحقق فيه هذا الاختبار يكون **غير صحيح** ، ويجب رفضه وعدم الاعتماد عليه بالمرة .

أما إذا كانت الأوزان المستعملة واحدة فى الحالتين فالشرط يتحقق ، ولكن هذا نادر جداً لا فائدة منه عملياً .

الأرقام القياسية الجيدة قريبة من تحقيق الاختبار

٣٤٣ — وبالرغم من هذا ، نجد فى الواقع أن أحسن الأرقام القياسية وأصلحها من الناحية النظرية ، أقرب من غيرها تحقيقاً لهذا الاختبار الخاطىء ، ولكنها لا تحققه تماماً بالطبع . فمثلاً نجد [1] أن الرقم الأمثل ، الذى سبقت

(١) أنظر كتاب I. Fisher: *The Making of Index Numbers* 1927, p. 277

الإشارة إليه — وهو أحسن الأرقام القياسية المعروفة وأصحها وأدقها — أدنى [1] إلى تحقيق هذا الاختبار من بعض الأرقام القياسية الأخرى . ولكنه لا يحقق الاختبار تماماً . ويلاحظ أيضاً أن الأرقام القياسية الرديئة بعيدة عن تحقيق هذا الاختبار أى أن حاصل الضرب

$$ع_{٢١} \times ع_{٣٢} \times ع_{١٣}$$

يختلف كثيراً عن ١ وهو القيمة المطلوبة حسب شروط الاختبار .

والخلاصة أن وجود فرق كبير بين هذا الحاصل والواحد الصحيح ، أو عدم وجود فرق بالمرة ، يدل على خطأ الرقم القياسى وعدم صلاحيته للاستعمال .

---

(١) أى أن حاصل الضرب $ع_{٢١} \times ع_{٣٢} \times ع_{١٣}$ أقرب إلى ١ فى حالة الرقم الأمثل منه فى حالة الأرقام الأخرى .

تعديل الأرقام القياسية

الأرقام التى لا تنعكس فى الزمن أو المعامل يمكن تعديلها لتنعكس

٣٤٤ — عرفنا فى هذا الباب الطرق التى نستعملها لاختبار الأرقام القياسية من حيث جودتها وصلاحيتها للتعبير بدقة عن الفكرة التى نرمى إليها باستخدام الرقم القياسى . و بتطبيق اختبارى الانعكاس فى الزمن وفى المعامل على الأرقام القياسية المعروفة لنا ، أمكننا أن نحكم على كل منها ، فرفضنا بعضها واستبقينا البعض الآخر .

ولكن هذين الاختبارين كانا فى الواقع بمثابة امتحان قاس لهذه الأرقام ، فكانت النتيجة أن رفضنا جميعها ، ماعدا واحداً فقط ، ألا وهو الرقم الأمثل الذى يقترحه فيشر . فهل من سبيل إلى تصحيح هذه الأرقام المرفوضة حتى تستوفى هذين الاختبارين ؟

جميع الأرقام القياسية يمكن **تعديلها** حتى تنعكس فى الزمن أو فى المعامل . وسنشرح هنا طرق التعديل ونطبقها على بعض الأرقام القياسية المعروفة .

تعريف المقلوب الزمنى

٣٤٥ — **المقلوب الزمنى** [١] لأى رقم قياسى هو عبارة عن خارج قسمة الواحد الصحيح على بديله الزمنى . فمثلا الرقم التجميعى المرجح بأوزان $ك_{٠}$ نعلم أنه لا ينعكس فى الزمن ؛ ومعادلته هى كما نعلم :

$$\frac{\text{مجـ}\ ع_{١}\ ك_{٠}}{\text{مجـ}\ ع_{٠}\ ك_{٠}}$$

ومعادلة بديله الزمنى هى

$$\frac{\text{مجـ}\ ع_{٠}\ ك_{١}}{\text{مجـ}\ ع_{١}\ ك_{١}}$$

(١) اسمه بالانجليزية Time Antithesis

فينتج أن المقلوب الزمنى معادلته هى :

$$\frac{\text{مج ع}_1\text{ ك}_1}{\text{مج ع}_.\text{ ك}_1} \cdot$$

وكذلك الوسط الحسابى للمناسيب المرجح بأوزان م. مثلا ، وبديله الزمنى ، معادلتاهما على الترتيب هما :

$$\frac{\text{مج س م}_.}{\text{مج م}_.} \quad \text{و} \quad \frac{\text{مج } \left(\frac{\text{م}_1}{\text{س}}\right)}{\text{مج م}_1} \cdot$$

.·. معادلة المقلوب الزمنى لهذا الوسط هى :

$$\frac{\text{مج م}_1}{\text{مج } \left(\frac{\text{م}_1}{\text{س}}\right)} \cdot$$

المقلوب الزمنى يساوى الرقم نفسه إذا كان منعكساً

وإذا كان الرقم الأصلى ينعكس فى الزمن ،كان المقلوب الزمنى مساوياً للرقم نفسه . ويكون شرط الانعكاس فى الزمن هو أن الرقم يساوى مقلوبه الزمنى ، أى ١ ÷ البديل الزمنى .

فالرقم التجميعى البسيط مثلا ، نعلم أنه ينعكس فى الزمن ، ومعادلته هى :

$\frac{\text{مج ع}_1}{\text{مج ع}_.}$ ، وبديله الزمنى معادلته هى $\frac{\text{مج ع}_.}{\text{مج ع}_1}$

.·. مقلوبه الزمنى يساوى $\frac{\text{مج ع}_1}{\text{مج ع}_.}$ = الرقم الأصلى نفسه .

٣٤٦ — الوسط الهندسى بين أى رقم قياسى ومقلوبه الزمنى ، ينعكس فى الزمن حتى ولو كان الرقم الأصلى لا ينعكس . فمثلا :

الرقم $\frac{\text{مج ع}_1\text{ ك}_.}{\text{مج ع}_.\text{ ك}_.}$ ، ومقلوبه الزمنى يساوى $\frac{\text{مج ع}_1\text{ ك}_1}{\text{مج ع}_.\text{ ك}_1}$

والوسط الهندسى لهذين هو :

$$\sqrt{\frac{\text{مجـ}\ ع_{١}\ ك_{٠}}{\text{مجـ}\ ع_{٠}\ ك_{٠}} \times \frac{\text{مجـ}\ ع_{١}\ ك_{١}}{\text{مجـ}\ ع_{٠}\ ك_{١}}}$$

وهو الرقم الأمثل المعروف لنا . ونعلم أنه ينعكس فى الزمن . وهكذا فى أى رقم آخر نختاره ، نجد أن الجذر التربيعى لحاصل ضربه فى مقلوبه الزمنى ينعكس فى الزمن . والحقيقة أن هذه الخاصة ناتجة مباشرة من تعريف المقلوب الزمنى .

الانعكاس فى الزمن خاصة لازمة للمقلوب الزمنى

إذ لو كان ا أى رقم قياسى ، وكان مقلوبه الزمنى ب ، فإن المقلوب الزمنى لهذا الأخير هو ا نفسه . لأنه حسب تعريف المقلوب الزمنى يكون

$$ب = \frac{١}{\text{بديل ا الزمنى}} ،$$

$$\therefore \text{بديل ب الزمنى} = \frac{١}{ا} ،$$

$$\text{أى أن}\qquad ا = \frac{١}{\text{بديل ب الزمنى}} ،$$

$$= \text{مقلوب ب الزمنى} ،$$

وبناء على ذلك يكون المقلوب الزمنى لحاصل الضرب ا ب يساوى مقلوب ا × مقلوب ب ، أى ب . ا ؛ ويكون إذن $\sqrt{ا\ ب}$ رقماً ينعكس فى الزمن حيث إن مقلوبه الزمنى يساويه نفسه .

وهكذا يمكننا تعديل أى رقم قياسى لكى ينعكس فى الزمن ، بأن نوجد الوسط الهندسى بينه و بين مقلوبه الزمنى . وهـذا الوسط الهندسى نفسه يكون هو الرقم القياسى المعدل .

المقلوب المعاملى

٣٤٧ — وبمثل هذه الطريقة تعدل الأرقام القياسية لكى تنعكس فى المعامل . فنوجد **المقلوب المعاملى** [1] للرقم ؛ فيكون الوسط الهندسى بين الرقم ومقلوبه المعاملى مستوفياً لشرط الانعكاس فى المعامل . والمقلوب المعاملى لأى رقم هو خارج قسمة منسوب القيم $\frac{\text{مح ع}_{١}\,\text{ك}_{١}}{\text{مح ع}_{٠}\,\text{ك}_{٠}}$ على البديل المعاملى للرقم نفسه .

لنأخذ مثلا الرقم التجميعى المرجح بأوزان $\text{ك}_{٠}$ ؛ وهو كما نعلم لا ينعكس فى المعامل . هذا الرقم وبديله المعاملى هما على الترتيب :

$$\frac{\text{مح ع}_{١}\,\text{ك}_{٠}}{\text{مح ع}_{٠}\,\text{ك}_{٠}} \quad \text{و} \quad \frac{\text{مح ك}_{١}\,\text{ع}_{٠}}{\text{مح ك}_{٠}\,\text{ع}_{٠}} \,.$$

∴ المقلوب المعاملى ، على حسب التعريف ، يكون

$$= \frac{\text{مح ع}_{١}\,\text{ك}_{٠}}{\text{مح ع}_{٠}\,\text{ك}_{٠}} \div \frac{\text{مح ك}_{١}\,\text{ع}_{٠}}{\text{مح ك}_{٠}\,\text{ع}_{٠}}$$

$$= \frac{\text{مح ع}_{١}\,\text{ك}_{١}}{\text{مح ع}_{٠}\,\text{ك}_{١}}$$

ويكون الوسط الهندسى بين هذا الأخير والرقم الأصلى ، يساوى

$$\sqrt{\frac{\text{مح ع}_{١}\,\text{ك}_{٠}}{\text{مح ع}_{٠}\,\text{ك}_{٠}} \times \frac{\text{مح ع}_{١}\,\text{ك}_{١}}{\text{مح ع}_{٠}\,\text{ك}_{١}}}$$

وهذا هو الرقم الأمثل الذى نعرف أنه ينعكس فى المعامل .

الانعكاس فى المعامل نتيجة تعريف المقلوب

٣٤٨ — وهذه الخاصة أيضاً ، مثل خاصة الانعكاس فى الزمن ، نتيجة مباشرة لتعريف المقلوب المعاملى . فلنفرض مثلا أن ا هو أى رقم قياسى وأن ح هو مقلوبه المعاملى .

$$\therefore \text{ح} = \frac{\text{مح ع}_{١}\,\text{ك}_{١}}{\text{مح ع}_{٠}\,\text{ك}_{٠}} \div \text{(البديل المعاملى للرقم الأصلى ا)} \quad \cdots (١)$$

---

(١) اسمه بالإنجليزية Factor Autithesis

∴ البديل المعاملي للرقم حـ

$$= \frac{\text{مج}\ \text{ك}_{١}\ \text{ع}_{١}}{\text{مج}\ \text{ك}_{٠}\ \text{ع}_{٠}} \div (\text{الرقم ا نفسه}) \qquad (٢)$$

$$\therefore\ ا = \frac{\text{مج}\ \text{ع}_{١}\ \text{ك}_{١}}{\text{مج}\ \text{ع}_{٠}\ \text{ك}_{٠}} \div (\text{البديل المعاملي للرمز حـ}) \qquad (٣)$$

= المقلوب المعاملي للرمز حـ .

أى أن ا و حـ كل منهما المقلوب المعاملي للآخر . وينتج من ذلك أن المقلوب المعاملي لحاصل الضرب ا حـ هو حـ ا . إذن يكون الوسط الهندسى $\sqrt{\text{ا حـ}}$ رقماً قياسياً ينعكس فى المعامل ، لأن :

$\sqrt{\text{ا حـ}}$ × بديله المعاملي

$$= \sqrt{\text{ا حـ}} \times \sqrt{\text{بديل ا المعاملي} \times \text{بديل حـ المعاملي}}$$

$$= \frac{\text{مج}\ \text{ع}_{١}\ \text{ك}_{١}}{\text{مج}\ \text{ع}_{٠}\ \text{ك}_{٠}}$$ ، أنظر المعادلتين (١) و (٢) أعلاه ؛

= منسوب القيم .

تعديل الرقم القياسى من حيث الزمن والمعامل

٣٤٩ — بينا أن الوسط الهندسى بين أى رقم ومقلوبه الزمنى يستوفى شرط الانعكاس فى الزمن ، حتى ولو كان الرقم الأصلى لا يستوفى هذا الشرط . وكذلك بينا أن الوسط الهندسى بين أى رقم قياسى ومقلوبه المعاملي ينعكس فى المعامل ، ولو كان الرقم الأصلى لا ينعكس . وبذلك توصلنا إلى طريقة لتعديل أى رقم لا ينعكس فى الزمن أو فى المعامل . أما إذا كان الرقم لا ينعكس فى الزمن ولا فى المعامل ، فتلزم معالجته مرتين : الأولى بالنسبة إلى الزمن ، والثانية بالنسبة للمعامل . وهذا يكون بإيجاد الوسط الهندسى بين هذا الرقم ومقلوبه الزمنى ، فينتج لنا رقم ينعكس فى الزمن . ثم نعالج هذا بالنسبة للمعامل ، فنوجد

الوسط الهندسى بينه و بين مقلوبه المعاملى . وهذا الأخير ينعكس فى المعامل وفى الزمن أيضاً . أى أن المعالجة الأخيرة بخصوص الانعكاس فى المعامل لا تؤثر فى صفة الانعكاس فى الزمن المكتسبة من المعالجة الأولى .

ولإثبات ذلك نفرض أن ا أى رقم قياسى لاينعكس فى الزمن ولا فى المعامل، ونفرض أن مقلوبه الزمنى هو ب ، فينتج أن $\sqrt{\text{ا ب}}$ ينعكس فى الزمن .

ونفرض أن $\text{ا}'$ هو بديل معاملى ا ؛ وأن $\text{ب}'$ هو بديل معاملى ب . فينتج أن $\text{ب}'$ هو مقلوب زمنى $\text{ا}'$ ؛ لأن

$\text{ب}'$ = بديل معاملى ب ،

= بديل معاملى للمعادلة $\left(\frac{1}{\text{بديل زمنى ا}} \right)$

= بديل زمنى للمعادلة $\left(\frac{1}{\text{بديل معاملى ا}} \right)$

= بديل زمنى $\left(\frac{1}{\text{ا}'} \right) = \frac{1}{\text{بديل زمنى } \text{ا}'}$

= مقلوب زمنى $\text{ا}'$

$\therefore$ $\sqrt{\text{ا}'\text{ب}'}$ ينعكس فى الزمن ؛ وهو بديل معاملى $\sqrt{\text{ا ب}}$.

$\therefore$ مقلوب معاملى $\sqrt{\text{ا ب}}$ هو خارج قسمة منسوب القيم على $\sqrt{\text{ا ب}}$.

$\therefore$ الوسط الهندسى بين $\sqrt{\text{ا ب}}$ ومقلوبه المعاملى

$$= \left[\sqrt{\frac{\text{ا ب}}{\text{ا}'\text{ب}'}} \times \frac{\text{م ع}_1 \text{ ك}_1}{\text{م}_0 \text{ ع}_0 \text{ ك}_0} \right]^{\frac{1}{2}} .$$

هذا ينعكس فى المعامل ، طبقاً للبند ٣٤٧ ، وهو ينعكس أيضاً فى الزمن ، لأن كل جزء من أجزائه : $\sqrt{\text{ا ب}}$ و $\sqrt{\text{ا}'\text{ب}'}$ ومنسوب القيم ، ينعكس فى الزمن على حدته . وهو إذن ينعكس فى الزمن وفى المعامل فى نفس الوقت .

ترتيب التعديل بالنسبة إلى الزمن أو المعامل غير مهم

٣٥٠ — وسواء فى ذلك إذا أجرينا التعديل بالنسبة إلى الزمن أولا ثم بالنسبة إلى المعامل ، أو العكس . والنتيجة التى نصل إليها فى الحالتين واحدة كما يتضح مما يأتى :

لنفرض أيضاً أن ا أى رقم قياسى لا ينعكس فى الزمن ولا فى المعامل ، وأن $\overline{\text{ا}}$ بديله المعاملى ؛ وأن ب ، $\overline{\text{ب}}$ هما المقلوبان الزمنيان لهما على الترتيب .

∴ الوسط الهندسى بين ا ومقلوبه المعاملى

$$\left[\frac{\text{ا}}{\overline{\text{ا}}} \times \frac{\text{ح}\ \text{ع}_{1}\ \text{ك}_{1}}{\text{ح}\ \text{ع}_{0}\ \text{ك}_{0}}\right]^{\frac{1}{2}} \text{ ؛}$$

وهذا ينعكس فى المعامل . والمقلوب الزمنى لهذا الأخير هو

$$\left[\frac{\overline{\text{ب}}}{\text{ب}} \times \frac{\text{ح}\ \text{ع}_{1}\ \text{ك}_{1}}{\text{ح}\ \text{ع}_{0}\ \text{ك}_{0}}\right]^{\frac{1}{2}} .$$

∴ الوسط الهندسى بين هذين هو :

$$\left[\sqrt{\frac{\text{ا}\ \text{ب}}{\overline{\text{ا}}\ \overline{\text{ب}}}} \times \frac{\text{ح}\ \text{ع}_{1}\ \text{ك}_{1}}{\text{ح}\ \text{ع}_{0}\ \text{ك}_{0}}\right]^{\frac{1}{2}} \text{ ،}$$

وهذا الرقم ينعكس فى المعامل وفى الزمن معاً ، وقد حصلنا عليه بتعديل ا بالنسبة للمعامل أولا ، ثم تعديل الناتج بالنسبة إلى الزمن . ولو عالجنا ا بالنسبة إلى الزمن أولا لحصلنا على $\sqrt{\text{ا}\ \text{ب}}$. و بتعديل هذا بالنسبة إلى المعامل نحصل على

$$\left[\sqrt{\frac{\text{ا}\ \text{ب}}{\overline{\text{ا}}\ \overline{\text{ب}}}} \times \frac{\text{ح}\ \text{ع}_{1}\ \text{ك}_{1}}{\text{ح}\ \text{ع}_{0}\ \text{ك}_{0}}\right]^{\frac{1}{2}}$$ ، وهى نفس النتيجة السابقة .

تعديل صيغ الأرقام يسبب تعقيدها وصعوبتها

٣٥١ — يمكننا إذن أن نعدل أى رقم قياسى حتى ينعكس فى الزمن وفى المعامل أيضاً ، وهذا التعديل هو من حيث الصيغة أو التركيب . و بذلك يمكننا أن نجعل جميع صيغ الأرقام القياسية التى عرفناها صالحة للاستعمال .

ولكن هذا التعديل — مع الأسف — يعطينا فى جميع الأحوال صيغاً أكثر تعقيداً فى التركيب ، وأشد إرهاقاً فى الحساب ، من الصيغ الأصلية المقصود تعديلها وهذا التعقيد وما ينشأ عنه من زيادة فى أعباء العمل الحسابى واحتمال الخطأ فيه ، ينال من فائدة هذا التعديل ويقلل من أهميته من الناحية العملية . ولا يسلم من هذا التعقيد إلا الصيغ التجميعية التى ذكرناها فى بند ٢٨٦ ، حيث ينتج من تعديلها الرقم الأمثل ، أو

$$\sqrt{\frac{\text{مح}\,ع_1\,ك_0}{\text{مح}\,ع_0\,ك_0} \times \frac{\text{مح}\,ع_1\,ك_1}{\text{مح}\,ع_0\,ك_1}} \; .$$

أما من حيث السرعة فى العمل الحسابى فهذا الرقم أفضل من أغلبية الأرقام كلها (٧٥ ٪ منها تقريباً)[١] .

تعديل الأرقام بتهجين أوزانها

٣٥٢ — نبحث الآن فى تعديل صيغ الأرقام القياسية بمعالجة الأوزان المستعملة لترجيح الأسعار الداخلة فى تركيب هذه الأرقام .

وجدنا فى الصيغ التجميعية (بند ٢٨٥ و بند ٢٨٦) أننا نرجح الأسعار إما بالكميات $ك_0$ أو بالكميات $ك_1$ ؛ ووجدنا فى كل حالة أن الصيغة الناتجة من استعمال واحد أو الآخر من هذين النظامين لا تنعكس فى الزمن ولا فى المعامل . ويمكننا تحسين هذه النتيجة لو استعملنا « هجناً » من هذين النوعين من الأوزان . فقد رأينا مثلا (بند ٣٢٠) أن الصيغة

$$\frac{\text{مح}\,ع_1\,(ك_0 + ك_1)}{\text{مح}\,ع_0\,(ك_0 + ك_1)}$$

تنعكس فى الزمن ؛ فضلا عن أنها تأخذ فى الاعتبار ظروف السنتين المقارنة والأساسية من حيث تحديد الأوزان . ولو أنها لا تنعكس فى المعامل ولكنها أسهل فى الحساب من الرقم الأمثل .

(١) أنظر I. Fisher *The Making of Index Numbers*, 1927

التهجين الهندسى والتوافقى

٣٥٣ — ويمكننا تهجين الأوزان بطريقة أخرى ، كأن نأخذ الوسط الهندسى أو التوافقى بين $ك_{٠}$ و $ك_{١}$ بدل الوسط الحسابى الذى اخذناه هنا . ولكن العمل الحسابى اللازم هنا يكون أصعب منه لو أخذنا الوسط الحسابى . وفى كل هذه الأحوال نحصل على صيغ تنعكس فى الزمن ، مثل الصيغة المذكورة .

وكذلك فى الصيغ الأخرى غير التجميعية ، مثل الوسط الحسابى أو الهندسى (البسيط أو المرجح) لمناسيب الأسعار ، يمكننا تهجين الأوزان . فنعرف أن الأوزان المستعملة فى هذه الصيغ هى :

$$م_{٠} = ع_{٠}\,ك_{٠} \quad و \quad م_{١} = ع_{١}\,ك_{١} ،$$

$$و \quad مـ_{٠} = ع_{٠}\,ك_{١} \quad و \quad مـ_{١} = ع_{١}\,ك_{٠} .$$

والتهجين بين الأوزان $م_{٠}$ و $م_{١}$ ، أو بين $مـ_{٠}$ و $مـ_{١}$ ، يعطى نتائج تنعكس فى الزمن ، سواء كان التهجين على نظام الوسط الحسابى أو التوافقى أو الهندسى . وأما التهجين بين م و مـ فينتج صيغاً لا تنعكس فى الزمن . وهذا واضح ، حيث أن حاصل جمع (أو ضرب) $ع_{٠}\,ك_{٠}$ و $ع_{١}\,ك_{١}$ — أو $ع_{٠}\,ك_{١}$ و $ع_{١}\,ك_{٠}$ — عبارة عن كمية متماثلة بالنسبة إلى الرقمين ٠ و ١ . يوجد أحدهما حيثما وكيفما يوجد الآخر . فلا نتغير هذه الكمية إذا ما أخذ كل منهما مكان الآخر عند تبديل سنة الأساس بسنة المقارنة . بخلاف حاصل جمع (أو ضرب) $ع_{٠}\,ك_{١}$ و $ع_{٠}\,ك_{٠}$ — أى $مـ_{٠}$ مع $م_{١}$ أو $مـ_{١}$ مع $م_{٠}$ — فهو كمية غير متماثلة بالنسبة للرقمين ٠ و ١ ، فنرى هنا مثلا أن صفراً موجود ثلاث مرات : اثنتان مع ع وثالثة مع ك ، بينما ١ يوجد مرة فقط مع ك ولا يوجد مع ع أبداً . وهذه الكمية تتغير عند تبديل السنة الأساسية بالسنة المقارنة ، فلا ينعكس الرقم فى الزمن .

تعديل الأرقام بتهجين معادلاتها

٣٥٤ — ويمكننا تعديل الأرقام القياسية لكى تنعكس فى الزمن، وفى الوقت نفسه نجمع بين ميزات الأوزان المختلفة، بإجراء التهجين بين الصيغ نفسها المستعملة فيها هذه الأوزان، بدلا من تهجين الأوزان. والتهجين هنا أيضاً يكون إما حسابياً، أو توافقياً، أو هندسياً، حيث نوجد الوسط الحسابى أو التوافقى أو الهندسى بين الصيغتين المراد تهجينهما، للجمع بين ميزات الأوزان فيهما. وأحسن مثال لذلك هو تهجين الرقمين التجميعيين المرجحين بالكميات $ك_{.}$ و $ك_{1}$ وهما

$$\frac{\text{مج}\ ع_{1}\ ك_{.}}{\text{مج}\ ع_{.}\ ك_{.}} \quad \text{و} \quad \frac{\text{مج}\ ع_{1}\ ك_{1}}{\text{مج}\ ع_{.}\ ك_{1}}$$

حيث نتج من تهجينهما هندسياً الرقم الأمثل :

$$\sqrt{\frac{\text{مج}\ ع_{1}\ ك_{.}}{\text{مج}\ ع_{.}\ ك_{.}} \times \frac{\text{مج}\ ع_{1}\ ك_{1}}{\text{مج}\ ع_{.}\ ك_{1}}}$$

تهجين المعادلات هندسياً أفضل

٣٥٥ — التهجين الهندسى هنا أفضل من الحسابى أو التوافقى. لأن عمليات الجمع المستعملة فى هذين الأخيرين تكون عقبة كبيرة فى اختصار الكسور التى نقابلها فى القسمة على البديلات الزمنية أو المعاملية، لاستخراج المقلوبات الزمنية والمعاملية للأرقام. وهذا مما يمنع تحقيق شرط الانعكاس المنشود.

٣٥٦ — يمكننا إذن أن نعدل صيغ الأرقام القياسية لكى تنعكس فى الزمن بتهجين الأوزان أو الصيغ نفسها. وهذه الطريقة تخالف طريقة المقلوب الزمنى، ولكننا نحتاج فيها إلى انتقاء الصيغ أو الأوزان المناسبة لكى تنتج من التهجين صيغ تنعكس فى الزمن، بخلاف الطريقة الأولى فهى صريحة تؤدى إلى النتيجة المطلوبة مباشرة، حيث نوجد البديل الزمنى للرقم المراد تعديله، ونقسم هذا الأخير على البديل الزمنى، ونستخرج الجذر التربيعى، فنحصل على الرقم الجديد المعدل.

المـــراجع

Fisher, I., : *The Making of Index Numbers*,
Chapters IV, V, VII, VIII, XIII.
Mills, F.C., : *Statistical Methods*, Chapter VI.
Rietz, H., : *Handbook of Mathematical Statistics*, Chapter XII.

الباب الثالث عشر

المفاضلة بين الأرقام القياسية

يوجد لدينا عدد كبير من الصيغ المختلفة للأرقام القياسية

٣٥٧ — عرفنا فى الباب الحادى عشر الصيغ المختلفة لتركيب الأرقام القياسية ؛ وفى الباب السابق بحثنا فى صلاحية كل من هذه الصيغ لتأدية الغرض المقصود من إنشاء الرقم القياسى . وعند تطبيق الاختبارات التى أجريناها على هذه الأرقام القياسية حصلنا على مجموعة جديدة من الصيغ : المقلوب الزمنى والمقلوب المعاملى لكل الصيغ التى اختبرناها أو حاولنا تعديلها .

وكل واحدة من هذه الصيغ أمكننا تعديلها وتصحيحها لكى تستوفى شروط الانعكاس فى الزمن وفى المعامل ، وبذلك أصبحت جميعها صالحة للاستعمال — نظرياً . وتبقى الآن مسألة المفاضلة بين هذه الأرقام ، واختيار أحسنها من الوجهة النظرية والعملية أيضاً .

الصيغ الأصلية للأرقام القياسية

٣٥٨ — يمكن تقسيم صيغ الأرقام القياسية التى حصلنا عليها إلى ست مجموعات رئيسية حسب نوع المتوسط المستعمل فى تركيبها . وهذه المجموعات هى :

١ — مجموعة الوسط الحسابى لمناسيب الأسعار :

(١) بسيط ، (٢) مرجح بأوزان م. ؛ و (٣) مرجح بأوزان مـ. ؛

و (٤) مرجح بأوزان مـ١ ؛ و (٥) مرجح بأوزان م١ .

ولكل من هذه الخمسة مقلوب معاملى ؛ فتكون جملتها عشرة .

٢ — مجموعة الوسط التوافقى لمناسيب الأسعار :

بسيط أو مرجح بأوزان $\text{م}_{٠}$ أو $\text{مـ}_{٠}$ أو $\text{مـ}_{١}$ أو $\text{م}_{١}$.
ولكل من هذه الخمسة مقلوب معاملى ، وجملتها عشرة أيضاً

٣ — مجموعة الوسط الهندسى للمناسيب :

بسيط ومرجح بأوزان $\text{م}_{٠}$ أو $\text{مـ}_{٠}$ أو $\text{مـ}_{١}$ أو $\text{م}_{١}$
ولكل منها مقلوب معاملى .

٤ — مجموعة الوسيط للمناسيب :
بسيط أو مرجح(١) بأى نوع من الأوزان المذكورة ومقلوباتها المعاملية .

٥ — مجموعة المنوال المناسيب :

بسيط أو مرجح(١) ، ومقلوباتها المعاملية .

٦ — الأرقام التجميعية للأسعار .

(١) بسيط ؛ و (٢) مرجح بأوزان $\text{ك}_{٠}$ ، و (٣) مرجح بأوزان $\text{ك}_{١}$.
ولكل من هذه مقلوب معاملى أيضاً .

عدد الصيغ الأصلية

٣٥٩ — ومن هذا البيان يظهر أن عدد الصيغ الأصلية يساوى ٢٨ ، يشتق منها ٢٨ أخرى كمقلوبات معاملية لها ، فيكون المجموع ٥٦ ؛ ولكن يجب أن نلاحظ أن بعض الصيغ مشتركة ، فقد لاحظنا مثلا فى بند ٢٩٠ أن الصيغة (٢) من المجموعة ١ ، تطابق الصيغة (٢) من المجموعة ٦ ، ومثال ذلك كثير بين المجموعات المذكورة . ولذلك فهذه الصيغ ليست كلها مختلفة .

الصيغ الناتجة من التعديل

٣٦٠ — وعندما نعدل هذه الصيغ لتنعكس فى الزمن — ما عدا التى تنعكس من نفسها مثل الهندسى البسيط والتجميعى البسيط — نحصل على مجموعة أخرى من الأرقام القياسية عددها حوالى الخمسين . وعندما نعدلها

(١) أنظر معنى الوسيط المرجح والمنوال المرجح فى باب المتوسطات بند ١٥٨

بالنسبة إلى المعامل نحصل على مجموعة ثالثة؛ وعند تعديلها بالنسبة للزمن والمعامل معاً نحصل على مجموعة رابعة من الصيغ . وهكذا عندما نجرى التعديلات الأخرى مثل تهجين الأوزان أو تهجين الصيغ مع بعضها ، هندسياً أو حسابياً أو توافقياً ، يزداد العدد إلى ١٧٠ وهذا ينقص إلى ١٣٤ بعد استبعاد المكررات [١] .

٣٦١ — من بين هذه المجموعة الكبيرة نريد اختيار رقم واحد يجمع بين السهولة فى الحساب والسلامة من التعقيد فى الصورة والخطأ النظرى ، بحيث يكون حساساً يظهر التغيرات التى تطرأ على الأسعار من وقت لآخر ، و يقيسها باعتدال و بدون تحيز .

رفض الوسيط والمنوال ومشتقاتهما

٣٦٢ — وبالتأمل فى المجموعات الست التى ذكرناها فى بند ٣٥٨ وما يشتق منها ، لا نتردد فى أن نرفض مجموعتى الوسيط والمنوال بسبب ضعف حساستهما . لأن وسيط مجموعة من مناسيب الأسعار (أو المنوال) هو فى الواقع منسوب سلعة واحدة فى وسط السلسلة . وقد علمنا أن الوسيط لا يتأثر مهما تغيرت القيم التى تحته ، ما دامت فى تغيرها لا تزال أقل منه ؛ ومهما تغيرت القيم التى فوقه ، ما دامت لا تزال أكبر منه . فمجموعة المناسيب الآتية مثلا :

٨٥ ، ٩٢ ، ٩٩ ، ١٠٦ ، ١١١ ، ١١٩ ، ١٢٠ ، ١٢١ ، ١٢٥

وسيطها ١١١ ، وهذا الوسيط يبقى كما هو لو تغيرت المناسيب الأخرى وأصبحت كما يأتى مثلا :

٩٢ ، ٩٤ ، ١٠٢ ، ١٠٣ ، ١١١ ، ١١٢ ، ١٢٥ ، ١٣١ ، ١٤٥

ومن الواضح أن مستوى الأسعار فى المجموعة الثانية أرفع بكثير منه فى الأولى . وبالرغم من ذلك فإن الوسيط لا « يحس » بهذا الفرق الكبير

---

(١) أنظر كتاب فيشر المشار إليه فى ص ٢٠٢ .

ولا يتأثر به . فيجب إذن أن نرفض الوسيط ولا نعتمد عليه فى قياس تغيرات مستوى الأسعار . وكذلك نرفض المنوال وكل ما يشتق منه ومن الوسيط .

رفض الصيغ غير المرجحة

٣٦٣ — نرفض أيضاً الأرقام ذات الصيغ البسيطة غـير المرجحة ، حتى ولو كانت معدلة لتستوفى الاختبارين . لأنها لا تفاضل بين السلع المختلفة بما يتناسب وأهميتها . وبدهى أننا لو عرفنا الكميات $ك_{٠}$ و $ك_{١}$ للسلع فالأولى أن نستخدمها كأوزان لإنشاء رقم قياسى مرجح يأخذ فى الاعتبار الأهمية النسبية للسلع المختلفة، بدل استخدامها فى تعديل رقم بسيط لينعكس فى الزمن أو فى المعامل ، وهو لا يفرق بين السلع المهمة وغيرها .

تطبيـــق الاختبارين

٣٦٤ — الخطوة التالية فى عمليـــة الاختبار واضحة ، ألا وهى تطبيق اختبـارى الانعكاس فى الزمن وفى المعامل . فنرفض جميع الصيغ التى لا تحقق شروط هذين الاختبارين معاً . وبذلك يبقى عدد صغير جداً من الأرقام القياسـية — ١٣ فقط [1] . وهـذه الثلاثة عشر هى نتيجة تعديل الأرقام الآتية فى الزمن وفى المعامل :

الوسط الحسابى للمناسيب مرجح بأوزان $م_{١}$ أو $م_{١}$ أو $\sqrt{م_{٠} . م_{١}}$ ، وهمن بين الأول والثانى ؛ والوسط الهندسى بأوزان $م_{٠}$ أو $م_{١}$ أو $\sqrt{م_{٠} . م_{١}}$ وهجن بين الأول والثـانى ؛ والرقم التجميعى بأوزان $ك_{٠}$ أو $(ك_{٠} + ك_{١})$ أو الوسـط التوافقى بين $ك_{٠}$ ، $ك_{١}$ ، أو الهندسى $\sqrt{ك_{٠} . ك_{١}}$: أو بأوزان $\frac{م_{٠} + م_{١}}{ع_{٠} + ع_{١}}$ وهذه الأرقام غـير متساوية طبعاً ؛ ولـكنها تعطى نتائج قريبة من بعضها ،

---

(١) أنظر كتاب فيشر صفحة ٢١٩ ، وصفحات ٤٦٥ — ٤٨٥ حيث توجد معادلات هذه الأرقام وغـيرها . ويرى القارىء معادلات هذه الأرقام الثلاثة عشر فى آخر الـكتاب صفحة ٣٧٥ .

بحيث يمكننا اعتبارها متساوية عملياً لو أهملنا الفروق الصغيرة فى المنازل العشرية البعيدة .

الرقم الأمثل أفضل من الجميع

٣٦٥ — وبما أن جميع هذه الأرقام تستوفى اختبارى الانعكاس فى الزمن ، فالمفاضلة بينها تكون على أساس البساطة فى التركيب والسهولة فى العمل والحساب . ولو تأملنا فى هذه الصيغ (صحيفة ٣٧٥) وجدنا أن أبسطها تركيباً وأسهلها جميعاً هى الصيغة التجميعية التى عرفناها بالرقم الأمثل ، وهى

$$\sqrt{\frac{\text{مح ع}_1\text{ ك}_.}{\text{مح ع}_.\text{ ك}_.} \times \frac{\text{مح ع}_1\text{ ك}_1}{\text{مح ع}_.\text{ ك}_1}} \;.$$

ومن حيث السرعة فى الحساب فهى أسرع من الجميع .

بعض الاعتبارات العملية

٣٦٦ — هكذا نختار بين الأرقام القياسية من حيث الصيغة أو طريقة التركيب . ولكن صيغة الرقم القياسى ليست كل شىء فيه ، بل توجد اعتبارات أخرى وشروط يجب توافرها فيه قبل أن نعتمد عليه اعتماداً كلياً فى تصوير حركة مستوى الأسعار (أو الإنتاج أو أى ظاهرة أخرى) . وأهم هذه الاعتبارات الأخرى من الناحية العملية هى :

١ — ما هى السلع التى تدخل فى تركيب الرقم ، وما هى السلع التى يمكن إهمالها .

٢ — عدد السلع التى نأخذها .

٣ — دقة البيانات الخاصة بالأسعار والأوزان وإمكان الحصول عليها .

٤ — السرعة فى الحساب .

نوع السلع التى تدخل فى الرقم

٣٦٧ — وبخصوص نوع السلع التى ندخلها فى تركيب الرقم ، يجب أن نأخذ فى الحسبان الغرض الذى من أجله ننشىء الرقم القياسى ،

ونوع الرقم الذى نريده فإذا كنا مثلا ننشىء رقماً قياسياً لأسعار الجملة ، لاستخدامه للدلالة على حركة السوق العامة والنشاط التجارى على العموم ، فيجب أن السلع التى نأخذها تمثل جميع السلع المتداولة ذات الأهمية فى السوق ، فلا نتحيز للسلع الصناعية مثلا ضد السلع الاستهلاكية ، أو للسلع المستوردة ضد السلع الوطنية أو العكس ، أو المنتجات الزراعية ضد المنتجات الصناعية ، وهكذا .

أما إذا أردنا رقماً قياسياً لأسعار التجزئة ، فلا يدخل فيه إلا السلع التى تباع بالتجزئة أى السلع الاستهلاكية . وهذه يكون أغلبها سلعاً تامة الصنع جاهزة للاستهلاك ، أو سلعاً زراعية استهلاكية مثل القمح والبطاطس واللحوم وغيرها .

تقسيم السلع إلى مجموعات متجانسة

٣٦٨ — ويحسن تقسيم السلع هنا إلى مجموعات ، تمتاز المفردات فى كل منها بصفات خاصة ، ذات أهمية فى الناحية التى يتناولها الرقم القياسى المراد إنشاؤه . ففى الرقم القياسى العام لأسعار الجملة مثلا ، نقسم السلع إلى مجموعات المواد الغذائية والمواد الخام المستعملة فى الصناعة ، والمواد النصف المصنوعة ، والمنتجات الجاهزة ، ومجموعة منتجات التعدين . ويصح أن تقسم أى واحدة من هذه المجموعات إلى مجموعات فرعية زيادة فى التفصيل .

وعلى كل حال فالأساس الذى نبنى عليه التقسيم إلى المجموعات يختلف حسب الغرض الذى يستعمل فيه الرقم القياسى . فيصح أن يكون التقسيم مبنياً على أساس نوع المادة الرئيسية فى السلع ؛ كأن نقول مثلا مجموعة منتجات الحديد والصلب ، ومجموعة لمنتجات المعادن الأخرى ، ومجموعة لمنتجات القطن ، أو الصوف أو غيره . أو أن يكون على أساس طوائف المستهلكين من أفراد أو صناعات ؛ أو أن يكون التقسيم على أساس جغرافى ، كما نفعل أحياناً حيث نقسم السلع إلى مستوردة ووطنية .

في كل مجموعة تأخذ السلع غير المتشابهة في الحركة

٣٦٩ — وعند اختيار السلع من أى مجموعة ، يجب أن تكون السلع المختارة تمثل الحركات أو الاتجاهات المختلفة للأسعار فى هذه المجموعة . فإذا كان هناك عدة سلع ، تسير أسعارها دائماً فى اتجاه واحد هبوطاً أو صعوداً ، فيكفى أن نأخذ منها واحدة أو اثنتين لتمثيلها ، بدل أن نأخذها جميعها وندخلها فى الرقم القياسى . لأن إدخال الكل لا يؤثر فى النتيجة أى تأثير يذكر بجانب المشقة الحسابية الناتجة عن ذلك .

أما السلع التى تتغير أسعارها فى اتجاهات مختلفة أو بنسب مختلفة ، فهذه يجب إدخالها فى الرقم القياسى ، حتى تتمثل فيه هذه الاتجاهات المتباينة ، فيكون أدق فى تصوير الحالة على حقيقتها .

الترجيح غير المباشر يأخذ عدة تسعيرات لنفس السلعة

٣٧٠ — ومن الواضح أن كثرة السلع المأخوذة من أى مجموعة تعمل على تعزيز هذه المجموعة وترجيح التغيرات التى تحصل فى أسعار سلعها . وبالمثل إذا أخذنا جملة تسعيرات لنفس السلعة وأدخلنا هذه التسعيرات فى الرقم القياسى ، كما لو كانت تسعيرات لسلع مختلفة ، فإن هذا يكون بمثابة **ترجيح غير مباشر**(١) لهذه السلعة وتعزيز لأهميتها على غيرها .

عدد السلع فى الرقم القياسى

٣٧١ — وبدهى أن عدد السلع يزيد فى دقة الرقم القياسى ، ويجعله أقرب إلى الصحة فى تمثيل الحركة العامة لمستوى الأسعار . ويجب إذن أن يكون عدد السلع التى ندخلها فى الرقم القياسى كبيراً ، بحيث لا يقل عن حوالى ٣٠ سلعة . ومعلوم طبعاً أن زيادة السلع ينشأ عنها زيادة كبيرة فى العمليات الحسابية

---

(١) هذا هو المتبع فى إنشاء الرقم القياسى الجديد لأسعار الجملة فى مصر . أنظر الإحصاء السنوى العام سنة ١٩٣٥ — ١٩٣٦ صفحة ٤٩٠ ، حيث يسمونه « تثقيلا غير مباشر » وهى فى رأيى ترجمة ركيكة للعبارة الإنجليزية Indirect Weighting

خصوصاً فى الأرقام القياسية ذات الصيغ المعقدة . على أن الزيادة فى دقة الرقم القياسى الناشئة عن زيادة عدد السلع لا تتناسب معها تماماً . حيث إذا زاد عدد السلع إلى أربعة أمثاله ، نقص الخطأ المحتمل للرقم القياسى إلى النصف فقط . أى أن دقته تزيد إلى الضعف فقط . وكذلك إذا زاد عدد السلع إلى تسعة أمثاله ، زادت الدقة إلى ثلاثة أمثال فقط . والسبب فى ذلك أن الخطأ فى الرقم القياسى — حسب نظرية الاحتمالات — يتناسب مع $\frac{1}{\sqrt{ن}}$ ، حيث ن يساوى عدد السلع .

وعلى ذلك لا يجدى كثيراً أن نزيد عدد السلع طلباً للدقة ؛ ويكفى فى المسائل العادية أن نأخذ حوالى ٥٠ سلعة ، على ألا يكون العدد أقل من ٢٠ ، وليس هناك أى فائدة — عملياً — من أخذ عدد يزيد على ٢٠٠ سلعة .

الخطأ فى بيانات الأسعار أعظم خطراً من الخطأ فى تقدير الأوزان

٣٧٢ — إذا حصل خطأ فى جمع البيانات الخاصة بالأسعار فلا بد أن يظهر تأثير فى النتيجة النهائية للرقم القياسى المحسوب من هذه البيانات . أما إذا كان هناك خطأ فى الأوزان المستعملة فإنه لا يؤثر كثيراً فى النتيجة . وذلك لأن الوزن المضروب فى السعر أو فى منسوبه ، يوجد فى البسط وفى المقام أيضاً . فإذا تغير بالزيادة مثلا ، زاد البسط والمقام معاً — زيادة مختلفة على العموم — وكان تأثير ذلك فى قيمة الكسر صغيراً نسبياً ؛ وكذلك إذا نقص الوزن . أما إذا أخطأنا فى السعر أو منسوب السعر ، فإن البسط يتغير وحده دون المقام ، فيكون التغير فى قيمة الكسر أكبر منه فى الحالة الأولى .

وعلى ذلك يجب أن نعتنى كل العناية فى جمع بيانات الأسعار ، وحساب مناسيبها بدقة ؛ ولا بأس من تقريب الأوزان إلى أعداد صحيحة لتسهيل العمليات الحسابية .

الأرقام التجميعية أسرع فى الحساب من غيرها

٣٧٣ — أما من حيث السرعة فى العمليات الحسابية فنجد على العموم أن الأرقام القياسية ذات الصيغ التجميعية أسهل فى حسابها من الأرقام الأخرى، وأسهلها جميعاً وأسرعها فى الحساب هو الرقم التجميعى البسيط ؛ ويليه فى ذلك الرقم الذى معادلته :

$$\frac{\text{مج ع}_1 (\text{ك}_0 + \text{ك}_1)}{\text{مج ع}_0 (\text{ك}_0 + \text{ك}_1)} \text{ ؛}$$

وهو يجمع بين السرعة والانعكاس فى الزمن . وكذلك يعطى للسلع أهميتها حسب أوزان معتدلة ، تأخذ فى الاعتبار ظروف السنة الأساسية وظروف السنة المقارنة ، ولكنه لا ينعكس فى المعامل .

ويلى ذلك فى السرعة الرقم الأمثل ، وهو كما نعلم يستوفى شروط الانعكاس ولكنه أكثر تعقيداً من هذا .

استخدام الكميات كأوزان غير ممكن إذا كانت غير معروفة

٣٧٤ — وثمة اعتبار آخر يجب الالتفات إليه حينما نختار صيغة الرقم القياسى وهو خاص بالبيانات اللازم الحصول عليها لحساب الرقم القياسى ؛ ففى مصر مثلا لا يمكننا الآن الحصول على بيانات دقيقة من التجار عن الكميات المباعة أو المنتجة من أغلب السلع ، إذ أنه من السهل جداً على أى شخص أن يحصل على بيان بسعر السلعة بمجرد سؤال التاجر الذى يبيعها ، ولكن هذا التاجر نفسه لا يعرف كمية ما باعه منها هذا العام مثلا إلا بعد الرجوع إلى دفاتره إذا كانت لديه دفاتر منظمة ؛ ومن باب أولى لا يعرف ما باعه غيره من التجار أو كمية المنتج أو المخزون منها . هذا فضلا عن أن هذا البيان بالذات يعتبره التاجر سراً من أسراره وليس عنده الاستعداد للإدلاء به للغير .

ففى مثل هذه الأحوال لا بد من استبعاد جميع الأرقام القياسية التى تدخل فيها ك .

٣٧٥ — نورد هنا معادلات الأرقام القياسية التى تنعكس فى الزمن وفى المعامل المذكورة فى صفحة ٣٦٩ ، وعددها ١٣ .

فى هذه المعادلات قد وضعنا الرموز الآتية للاختصار :—

$$\text{س} = \frac{\text{ع}_١}{\text{ع}_٠} = \text{منسوب السعر ، } \text{ص} = \frac{\text{ك}_١}{\text{ك}_٠} = \text{منسوب الكمية}$$

$$\text{و } \text{ق} = \frac{\text{مجـ ع}_١\text{ك}_١}{\text{مجـ ع}_٠\text{ك}_٠} = \text{منسوب القيم ،}$$

$$\text{و } \text{م}_٠ = \text{ع}_٠\text{ك}_٠ \text{ ، } \text{مـ}_٠ = \text{ع}_٠\text{ك}_١ \text{ ، } \text{مـ}_١ = \text{ع}_١\text{ك}_٠ \text{ ، } \text{م}_١ = \text{ع}_١\text{ك}_١$$

المجموعة الأولى : الوسط الحسابى مرجحاً بأوزان $\text{مـ}_١$ أو $\text{م}_١$ أو $\sqrt{\text{م}_٠\text{م}_١}$ ، وهجن بين الأول والثانى ، كل منها معـدلا فى الزمن والمعامل .

وهذه هى على الترتيب :

$$(١)-\left[\frac{\text{مجـ س مـ}_١}{\text{مجـ مـ}_١} \times \left(\text{ن} \div \frac{\text{مجـ ص مـ}_٠}{\text{مجـ مـ}_٠}\right) \times \frac{\text{مجـ مـ}_٠}{\text{مجـ}\left(\frac{\text{مـ}_٠}{\text{س}}\right)} \times \left(\text{ن} \div \frac{\text{مجـ مـ}_١}{\text{مجـ}\left(\frac{\text{مـ}_١}{\text{ص}}\right)}\right)\right]^{\frac{١}{٤}} .$$

$$(٢)-\left[\frac{\text{مجـ س م}_١}{\text{مجـ م}_١} \times \left(\text{ن} \div \frac{\text{مجـ ص م}_١}{\text{مجـ م}_١}\right) \times \frac{\text{مجـ م}_٠}{\text{مجـ}\left(\frac{\text{م}_٠}{\text{س}}\right)} \times \left(\text{ن} \div \frac{\text{مجـ م}_٠}{\text{مجـ}\left(\frac{\text{م}_٠}{\text{ص}}\right)}\right)\right]^{\frac{١}{٤}} .$$

$$(٣)-\left[\frac{\text{مجـ س}\sqrt{\text{م}_١\text{م}_٠}}{\text{مجـ}\sqrt{\text{م}_١\text{م}_٠}} \times \left(\text{ن} \div \frac{\text{مجـ ص}\sqrt{\text{م}_١\text{م}_٠}}{\text{مجـ}\sqrt{\text{م}_١\text{م}_٠}}\right) \times \frac{\text{مجـ}\sqrt{\text{م}_١\text{م}_٠}}{\text{مجـ}\left(\frac{\sqrt{\text{م}_١\text{م}_٠}}{\text{س}}\right)} \times \left(\text{ن} \div \frac{\text{مجـ}\sqrt{\text{م}_١\text{م}_٠}}{\text{مجـ}\left(\frac{\sqrt{\text{م}_١\text{م}_٠}}{\text{ص}}\right)}\right)\right]^{\frac{١}{٤}} .$$

(٤) — $\sqrt{(١) \times (٢)}$ السابق ذكرهما .

المجموعة الثانية : الوسط الهندسى مرجحاً بأوزان مـ_0 أو م_1 أو $\sqrt{\text{م}_0 . \text{م}_1}$ ، وهى بين الأول والثانى ، كل منها معدلا فى الزمن والمعامل .

وهى على الترتيب :

$$(٥) - \left[\sqrt[\text{مجـ م}_0]{(\text{س})^{\text{م}_0} \times (\bar{\text{س}})^{\bar{\text{م}}_0} \times ..} \times \left\{ \text{ن} \div \sqrt[\text{مجـ م}_0]{(\text{ص})^{\text{م}_0} \times (\bar{\text{ص}})^{\bar{\text{م}}_0} \times ..} \right\} \right.$$

$$\left. \times \sqrt[\text{مجـ م}_1]{(\text{س})^{\text{م}_1} \times (\bar{\text{س}})^{\bar{\text{م}}_1} .. \times} \left\{ \text{ن} \div \sqrt[\text{مجـ م}_1]{(\text{ص})^{\text{م}_1} \times (\bar{\text{ص}})^{\bar{\text{م}}_1} \times ..} \right\} \right]^{\frac{1}{٤}}$$

$$(٦) - \left[\sqrt[\text{مجـ مـ}_0]{(\text{س})^{\text{مـ}_0} \times (\bar{\text{س}})^{\bar{\text{مـ}}_0} .. \times} \left\{ \text{ن} \div \sqrt[\text{مجـ مـ}_0]{(\text{ص})^{\text{مـ}_0} \times (\bar{\text{ص}})^{\bar{\text{مـ}}_0} \times ..} \right\} \right.$$

$$\left. \times \sqrt[\text{مجـ مـ}_1]{(\text{س})^{\text{مـ}_1} \times (\bar{\text{س}})^{\bar{\text{مـ}}_1} .. \times} \left\{ \text{ن} \div \sqrt[\text{مجـ مـ}_1]{(\text{ص})^{\text{مـ}_1} \times (\bar{\text{ص}})^{\bar{\text{مـ}}_1} \times ..} \right\} \right]^{\frac{1}{٤}}$$

$$(٧) - \left[\sqrt[\text{مجـ}\sqrt{\text{م}_0 . \text{م}_1}]{(\text{س})^{\sqrt{\text{م}_0 . \text{م}_1}} \times ... \times} \left\{ \text{ن} \div \sqrt[\text{مجـ}\sqrt{\text{م}_0 . \text{م}_1}]{\text{ص}^{\sqrt{\text{م}_0 . \text{م}_1}} \times ..} \right\} \right]^{\frac{1}{٤}}$$

(٨) — $\sqrt{(٥) \times (٦)}$ السابق ذكرها .

المجموعة الثالثة : الرقم التجميعى مرجحاً بأوزان ك_0 ، أو بالوسط الحسابى بين ك_0 و ك_1 ، أو بالوسط الهندسى ، أو التوافقى بينهما ، أو بالأوزان $\frac{\text{م}_0 + \text{م}_1}{\text{ع}_0 + \text{ع}_1}$ ، كل منها معدلا فى الزمن والمعامل .

وهى على الترتيب :

$$(٩) - \sqrt{\frac{\text{مجـ}\ ع_١ ك_٠}{\text{مجـ}\ ع_٠ ك_٠} \times \frac{\text{مجـ}\ ع_١ ك_١}{\text{مجـ}\ ع_٠ ك_١}} \text{ ، الرقم الأمثل}$$

$$(١٠) - \left[\frac{\text{مجـ}\ ع_١ (ك_٠ + ك_١)}{\text{مجـ}\ ع_٠ (ك_٠ + ك_١)} \times \left\{ ن \div \frac{\text{مجـ}\ ك_١ (ع_٠ + ع_١)}{\text{مجـ}\ ك_٠ (ع_٠ + ع_١)} \right\}\right]^{\frac{١}{٢}}$$

$$(١١) - \left[\frac{\text{مجـ}\ ع_١ \sqrt{ك_٠ ك_١}}{\text{مجـ}\ ع_٠ \sqrt{ك_٠ ك_١}} \times \left\{ ن \div \frac{\text{مجـ}\ ك_١ \sqrt{ع_٠ ع_١}}{\text{مجـ}\ ك_٠ \sqrt{ع_٠ ع_١}} \right\}\right]^{\frac{١}{٢}}$$

$$(١٢) - \left[\frac{\text{مجـ}\ ع_١ \left(\frac{٢}{\frac{١}{ك_٠}+\frac{١}{ك_١}}\right)}{\text{مجـ}\ ع_٠ \left(\frac{٢}{\frac{١}{ك_٠}+\frac{١}{ك_١}}\right)} \times \left\{ ن \div \frac{\text{مجـ}\ ك_١ \left(\frac{٢}{\frac{١}{ع_٠}+\frac{١}{ع_١}}\right)}{\text{مجـ}\ ك_٠ \left(\frac{٢}{\frac{١}{ع_٠}+\frac{١}{ع_١}}\right)} \right\}\right]^{\frac{١}{٢}}$$

$$(١٣) - \left[\frac{\text{مجـ}\ ع_١ \left(\frac{م_٠ + م_١}{ع_٠ + ع_١}\right)}{\text{مجـ}\ ع_٠ \left(\frac{م_٠ + م_١}{ع_٠ + ع_١}\right)} \times \left\{ ن \div \frac{\text{مجـ}\ ك_١ \left(\frac{م_٠ + م_١}{ك_٠ + ك_١}\right)}{\text{مجـ}\ ك_٠ \left(\frac{م_٠ + م_١}{ك_٠ + ك_١}\right)} \right\}\right]^{\frac{١}{٢}}$$

المـــراجع

CONNOR, L.R., : *Statistics in Theory and Practice,* Chapter XX.

FISHER I., : *The Making af Index Numbers,* Chapters XV-XVII

MILLS, F.C., : *Statistical Methods,* Chapter VI.

الباب الرابع عشر

نظرية العينات

الأعمال الحسابية فى الاحصاء تزداد إرهاقاً فى المجموعات الكبيرة

٣٧٦ — بحثنا فى الأبواب السابقة من هذا الكتاب ، القواعد والطرق الإحصائية التى نستخدمها عند دراسة مجموعة من المفردات ، لنتعرف خواصها التى تميزها عن غيرها من المجموعات ، تلك الخواص الإحصائية مثل المتوسط الذى يمثل مفردات المجموعة كنموذج لها ، أو الانحراف المعيارى الذى يقيس تشتتها حول هذا المتوسط أو تقاربها منه ، أو معامل الارتباط الذى يقيس العلاقة بين ظاهرتين متغيرتين ، وغير ذلك من المقاييس الإحصائية التى عرفناها . وفى كل هذه المسائل التى بحثناها رأينا أن المجهود الفنى والعقلى يتضاعف ويزداد عبؤه أضعافاً مضاعفة كلما زاد عدد المفردات التى تدخل فى البحث ، فضلا عن التكاليف المادية التى ترهق الباحث فوق طاقته ، وتحول بينه و بين إنجاز بحثه و إخراج النتائج التى يريد معرفتها . وكأننا والحالة هذه إذا أردنا بحث مجموعة كبيرة من المفردات لمعرفة خواصها ، لا نستطيع القيام بالعمليات الحسابية المرهقة، ولا نطيق تحمل النفقات الضرورية لهذا البحث ، فنترك المسألة بالمرة عاجزين يائسين من حلها.

نكتفى عملياً ببحث عينة صغيرة من المجتمع

٣٧٧ — وهناك طريقة عملية نلجأ إليها فى هذه المسائل ، ونصيب نجاحاً يبرر الاعتماد عليها ، ألا وهى أننا نأخذ من المجموعة الكبيرة المطلوب بحثها ، التى نسميها[1] **المجتمع الأصلى** ، مجموعة صغيرة نسميها[2] **عينة** . ثم نبحث هذه

---

(١) بالإنجليزية Original or Parent Population (٢) Sample

العينة الصغيرة بدلا من المجتمع الكبير ؛ ونقول إن الخواص الإحصائيـة للمجتمع الأصلى تساوى الخواص الإحصائية للعينة تقريباً . ونعتمد فى ذلك على أن هذه العينة التى أخذناها عينة صريحة مطلقة لا تحيز فيها ولا محاباة ، أى أننا أخذناها من بين مفردات المجتمع خبط عشواء وحيثما اتفق . أى كما نسميها [1] **عينة عشوائية** .

نطبق صفات العينة على المجتمع

٣٧٨ — لنفرض مثـلا أننا نبحث فى أجور مجموعـة كبيرة من العمال بقصد معرفة المتوسط العام للأجور بين هؤلاء العمال ، وأننا أخذنا تسهيـلا للعمل عينة عشوائية من هؤلاء العمال وحسبنا متوسط الأجور فى هذه العينة فوجـدناه عشرة قروش مثـلا . فينتج من ذلك ، حسب ما قلناه فى البنـد السابق ، أن المتوسط العام للأجور هو عشرة قروش أيضاً — تقريبـاً . ولـكن يصح لنا أن نتساءل : كيف لنا أن نعمم صفات عينة جزئية على مجتمع كبير ؟ وأين يكون متوسط هذه العينة من متوسط المجتمع ؟ وأى ثقة نضعها فى هذه النتيجة التى وصلنا إليها بطريق العينة ، وما الذى يبرر هـذه الثقة ؟ ومتى يكون حكمنا من العينـة على المجتمع صحيحاً أو قريباً من الصحة ، ومتى يكون بعيداً عنها ؟

دخول عوامل المصـادفة فى اختيار العينـة

٣٧٩ — ولم لا يكون حكمنا هذا صحيحاً وقد حرصنا عنـــد أخذ العينة أن تـكون عادلة ، فلم تتحيز إلى العمال ذوى الأجور العاليـة أو العكس ، بل أخذنا بعض العمال خبط عشواء وحيثما اتفق وكونا منهم العينة . أليس فى ذلك ترك العنان لعوامل المصادفة تأخـذ مجراها الطبيعى حتى إذا أخذنا فى العينـة رجلا ذا أجر عال جداً بمجرد المصادفة ، فإن المصادفة نفسها كفيلة أيضاً بأن تعطينا فى العينة رجـلا آخر يكون ذا أجر منخفض ، فيتعادل أجراً هـذين الرجلين

(١) بالإنجليزية Random Sample

ولا يتأثر متوسط الأجور فى العينة . والجواب على هذا هو ، بكل بساطة ، أننا لا نضمن ذلك التعادل ، ولا يمكننا الجزم بحصوله فى أى حالة معينة إذا ترك للمصادفة وحدها . وكل ما يمكن أن نقوله هو أنه يحتمل حصول هذا التعادل ، ويجوز أن تجود المصادفة باختيار الرجل الثانى فى العينة حتى يحصل هذا التعادل . وهكذا ندخل فى موضوع جديد — موضوع الاحتمالات — ويتوقف الفصل فى مسألتنا على تحديد معنى هذا العنصر الجديد ، وإيجاد طريقة دقيقة لقياسه .

الاحتمالات وحسابها

معنى الاحتمال فى اللغة الدارجة

٣٨٠ — الاحتمال فى لغتنا الدارجة لفظ يدل على الشك وعدم التحديد . شىء لا يقاس ولا يقدر تقديراً دقيقاً ، هو الإبهام بعينه . وهو بهذا المعنى الدارج لا يصلح لدراستنا التحليلية الرياضية أو الإحصائية التى تستلزم التحديد والدقة . بيد أننا لو تأملنا قليلا يمكننا أن نستخلص من هذا المعنى الدارج معنى آخر يساعدنا فى دراستنا . لننظر مثلاً لماذا نقول إن زهرة النرد (١) إذا رميت حيثما اتفق وتركت وشأنها تلف وتدور وتنقلب ، فهى ترسو أخيراً على واحد من أوجهها الستة ، يحتمل أن يكون « الجهار » ، مثلاً . ما الذى دعانا أن نقول إن « الجهار يحتمل أن يظهر » ؟ الجواب أننا إذ نرمى هذه الزهرة ونشاهدها تلف وتدور حول نفسها نجهل تماماً على أى وجه تستقر . وكل ما نعلمه من أمرها يقيناً أنها لا بد أن تستقر على واحد فقط من الأوجه الستة . ولا يمكننا أن نحدد هذا الوجه لأننا نعلم أن هناك عوامل وظروف تدعو الزهرة

---

(١) زهرة النرد العادى عبارة عن مكعب صغير من مادة متجانسة الكثافة ، نقشت على الأوجه الستة منه نقط عددها ١ و ٢ و ٣ و ٤ و ٥ و ٦ ، ونسمى الأوجه حسب عدد النقط : « يك ، دو ، سيه ، جهار ، بيش ، شيش » — على الترتيب .

أن تستقر على الجهار ، وعوامل وظروف أخرى تدعوها أن تستقر على غيره ؛ وأن كلا من العوامل المساعدة على ظهور « الجهار » أو غيره من الأوجه متساوية فى القوة والأهمية .

تحديد معنى الاحتمال

٣٨١ — ولنتأمل قليلا فى هذا الموقف العويص ، فإذا بنا نستخلص من هذا الإبهام والشك فكرة محدودة لا إبهام فيها . وهى أن النسبة بين احتمال ظهور « الجهار » واحتمال عدم ظهوره ، أى ظهور وجه آخر من الأوجه الخمسة الباقية للزهرة ، تساوى النسبة بين العددين ١ و ٥ ، لأن كلا من هذه الأوجه الخمسة له نفس الفرصة فى الظهور مثل « الجهار » تماماً .

مثال عملى للاحتمالات

٣٨٢ — وهذا مثال عملى آخر . لنأخذ مجموعة من أوراق اللعب [1] وقد خلطت أوراقها خلطاً تاماً . نسحب منها ورقة واحدة حيثما اتفق ، ونبحث فى احتمالات كونها تسعة مثلا ، بصرف النظر عن شكلها أو لونها . هنا نسحب الورقة من المجموعة خبط عشواء ، ولا نعلم صنفها إذ نسحبها ؛ ولكننا نعلم أن هناك فى المجموعة أربع ورقات من صنف التسعات موزعة حيثما اتفق ، يصح أن تسحب أى واحدة منها ، كما نعلم أن هناك ٤٨ ورقة من أصناف أخرى ، ويصح أن تسحب أى واحدة منها أيضاً . ولما كان السحب خبط عشواء ، فبديهى أن لكل ورقة فى المجموعة نفس الفرصة فى أن تكون هى المسحوبة . فمن الواضح إذاً أن النسبة بين احتمال سحب التسعة واحتمال عدم سحبها ، أى سحب غيرها ، هى كالنسبة بين العددين ٤ و ٤٨

النسبة بين عدد طرق الوقوع وعدد طرق الامتناع

٣٨٣ — ومن ناحية أخرى نرى فى هذه التجربة أن هناك أربعة سبل أو طرق يتحقق بها الأمر الذى نبحث فيه ، ألا وهو سحب ورقة ٩ ، كما أن هناك

---

(١) المسماة فى العادة « كتشينة » وفيها ٥٢ ورقة مقسمة إلى ١٣ صنفا ؛ وكل صنف من أربعة أشكال ولونين .

٤٨ طريقة يمتنع بها هذا الأمر عينه . وكل هذه الطرق — وعددها ٥٢ — جائزة بدرجة واحدة .

وبالاختصار نقول إن النسبة بين احتمال وقوع حدث معين واحتمال امتناع هذا الحدث ، تساوى النسبة بين عدد الطرق التى يمكن أن يقع بها ، وعدد الطرق التى يمكن أن يمتنع بها ، على فرض أن طرق الوقوع ، وطرق الامتناع كلها جائزة بدرجة واحدة ، كما رأينا فى المثالين السابقين . فإذا علمنا أن حدثاً معيناً يمكن أن يقع بطرق عددها ا ، ويمكن أن يمتنع بطرق عددها ب ، تكون النسبة بين احتمال وقوعه واحتمال امتناعه تساوى $\frac{ا}{ب}$.

التعريف الجبرى للاحتمال

٣٨٤ — ولما كان لكل حدث حالتان : إما الوقوع أو الامتناع ، فمن المؤكد أن حالة منهما تحصل . فإذا جعلنا الواحد الصحيح مقياساً للتأكيد ، يكون احتمال وقوع الحدث يساوى $\frac{ا}{ا + ب}$ واحتمال امتناعه يساوى $\frac{ب}{ا + ب}$ ، ومجموع الاحتمالين يساوى الواحد الصحيح . أى أن مقياس احتمال الوقوع ، أو الامتناع ، كسر أقل من الواحد الصحيح الذى يدل على التأكيد ؛ وكلما قرب هذا الكسر من الواحد زاد الاحتمال وقرب من أن يكون تأكيداً .

أمثلة على حساب الاحتمالات

٣٨٥ — هكذا نصل إلى معنى محدد للاحتمال يمكن قياسه بسهولة وبدقة فى أى حالة معينة . وذلك بأن نبحث فى عدد الطرق التى يمكن بها حصول الحدث الذى نحن بصدده ، وعدد الطرق التى يمتنع بها ؛ وما دامت طرق الحصول والامتناع كلها جائزة بدرجة واحدة ، نطبق القاعدة المذكورة فى البند السابق ، فنحصل على المقياس المطلوب لاحتمال وقوع الحدث أو امتناعه .

ولنأخذ الآن بعض أمثلة تطبيقية :

مثال (١)— رميت زهرة حيثما اتفق ، والمطلوب إيجاد احتمال رمى خمس نقط أو أكثر .

هذا يمكن أن يتحقق بظهور « البيش » أو « الشيش » (٥ أو ٦)، أى بطريقتين ؛ ويمكن أن يمتنع بظهور « يك » أو « دو » أو « سيه » أو « جهار » — أى بأربع طرق.

$$\therefore \text{احتمال رمى خمس نقط أو أكثر يساوى } \frac{٢}{٢ + ٤} = \frac{١}{٣} .$$

مثال (٢) — رميت زهرتان فى وقت واحد ، فما هو احتمال أن يحصل الرامى من الزهرتين معاً على نقط مجموعها ٧ (أو بالاختصار احتمال رمى ٧).

يمكن الحصول على سبع نقط بأى واحدة من الطرق الآتية :

١	من	الزهرة	الأولى	مع	٦	من	الثانية
٢	»	»	»	»	٥	»	»
٣	»	»	»	»	٤	»	»
٤	»	»	»	»	٣	»	»
٥	»	»	»	»	٢	»	»
٦	»	»	»	»	١	»	»

أى بست طرق فقط.

ولكن عدد الأوضاع الممكنة للزهرتين معاً هو على العموم ٦ × ٦ ، أى ٣٦ ؛ لأن كل وجه من أوجه الزهرة الأولى يمكن أن يقترن بأى وجه من أوجه الزهرة الثانية. والأوضاع الستة المذكورة آنفاً هى ضمن هذه الستة والثلاثين ، والأوضاع الباقية لا تعطى سبع نقط.

$$\therefore \text{احتمال الحصول على سبع نقط } = \frac{٦}{٣٦} = \frac{١}{٦} .$$

مثال (٣) — ذهب رجل لحضور اجتماع ، وطلب من خادمه أن يقابله بالسيارة ساعة الانصراف . وذهب الخادم إلى مكان الاجتماع فى الموعد المحدد

فإذا به بناء كبير له خمسة أبواب على أربعة شوارع مهمة ، منها اثنان على الشارع البحرى . ففى أى شارع يفضل الانتظار ؟

هنا يقف الخادم فى حيرة ، فلم يخبره سيده من أى باب سيخرج ، ويرى أمامه الناس متدفقين من كل الأبواب على السواء . فهو يتبع الحكمة لو أنه فكر فى احتمال خروج سيده فى كل من هذه الشوارع ، وانتظر فى الشارع الأكثر احتمالا .

احتمال خروجه فى الشارع البحرى $\frac{٢}{٢ + ٣} = \frac{٢}{٥}$ ، حيث فيه بابان يمكن أن يخرج سيده من أى واحد منهما على السواء ، وثلاثة أبواب أخرى .

واحتمال الخروج من أى شارع من الشوارع الأخرى = $\frac{١}{٥}$ ، لأن فى كل منها باب واحد فقط يمكن الخروج منه .

فالأفضل للخادم أن ينتظر سيده فى الشارع البحرى ، حيث احتمال مقابلته أكبر من احتمال مقابلته فى أى شارع آخر .

احتمال حصول حدثين مستقلين معاً

٣٨٦ — فى هذه الأمثلة بحثنا فى احتمال حصول حدث واحد أو امتناعه . لكننا فى كثير من المسائل نحتاج إلى معرفة احتمال حصول حدثين معاً (أو أكثر من حدثين) . وسنقتصر هنا على بحث الحالة التى يكون فيها الحدثان مستقلين عن بعضهما تماماً ، بمعنى أن حصول واحد منهما (أو امتناعه) لا يؤثر مطلقاً فى احتمال حصول الثانى . كما لو رمينا زهرة مرة وحصلنا منها على « يك » ، فهذا لا يؤثر مطلقاً فى احتمال ظهور « اليك » ثانياً إذا رميناها مرة ثانية .

لنفرض أن حدثين مستقلين عن بعضهما ، احتمال الأول منهما ح واحتمال الثانى حَ (نقصد احتمال الحصول) . فاحتمال حصولهما معاً يساوى حاصل ضرب الاحتمالين ، أى ح × حَ . لأن الحدث الأول قد يتحقق وقد يمتنع .

وفى الحالات التى يقع فيها الحدث الأول ، يقع الثانى فى بعضها و يمتنع فى الباقى .
فيكون احتمال حصولهما معاً يساوى (١) حاصل الضرب ح . حَ

إذا رمينا زهرتين مثلا فاحتمال أن كلا منهما تأتى « بالشيش » يساوى $\frac{١}{٣٦}$.
لأن احتمال « الشيش » من الأولى وحدها يساوى $\frac{١}{٦}$ ، وكذلك احتمال ظهور « الشيش » من الثانية يساوى $\frac{١}{٦}$ أيضاً . وما تأتى به إحدى الزهرتين لا يؤثر مطلقاً فى احتمال ما تأتى به الأخرى .

احتمال ثلاثة أحداث مستقلة أو أكثر

٣٨٧ — وكذلك إذا كان هناك ثلاثة أحداث مستقلة عن بعضها واحتمالاتها ح و حَ و حً على الترتيب ، فإن احتمال حصولها جميعاً يساوى حاصل الضرب ح . حَ . حً . وهكذا مهما تعددت الأحداث يكون احتمال حصولها جميعاً يساوى حاصل ضرب احتمالاتها فى بعضها ، بفرض أنه لا علاقة بين أى اثنين منها . ومن الواضح طبعاً أن احتمال وقوع أمرين معاً أصغر من احتمال وقوع أحدهما بمفرده ، وذلك لأن كلا من الاحتمالين أصغر من الواحد الصحيح وحاصل ضربهما أصغر من أيهما . و إذا ما كثر عدد هذه الأحداث تضاءل احتمال حصولها معاً .

---

(١) يمكننا إثبات ذلك جبريا بسهولة . نفرض أن الحدث الأول يمكن أن يقع بطرق عددها ا ، ويمتنع بطرق عددها ب ؛ ونفرض أن الثانى يقع بطرق عددها اَ ويمتنع بطرق عددها بَ .

$$\therefore \text{ح} = \frac{\text{ا}}{\text{ا} + \text{ب}} \qquad \text{و} \qquad \text{حَ} = \frac{\text{اَ}}{\text{اَ} + \text{بَ}}$$

ومن الواضح أن عدد الطرق التى يمكن أن يحصل بها الحدثان معاً يساوى ا . اَ ؛ وأن عدد الطرق الممكنة كلها للحدثين مع بعضهما يساوى (ا + ب) . (اَ + بَ) بما فيها الطرق التى يحصل بها الحدثان مجتمعين أو منفردين أو يمتنعان معاً .

$$\therefore \text{احتمال حصولهما معاً} = \frac{\text{ا اَ}}{(\text{ا} + \text{ب}) . (\text{اَ} + \text{بَ})} = \text{ح} . \text{حَ} .$$

مثال (١) — دعا رجل ثلاثة ضيوف إلى وليمة . فما هو احتمال أن الثلاثة يلبون الدعوة جميعاً ، إذا علم أن احتمال مجىء الأول $\frac{٣}{٤}$ واحتمال مجىء الثانى $\frac{١}{٣}$ واحتمال مجىء الثالث $\frac{١}{٢}$ ، وأنه لا علاقة لواحد منهم بآخر .

نظراً لعدم وجود أى علاقة بين هؤلاء الضيوف و بعضهم ، نرى أن مجىء واحد منهم لا يؤثر مطلقاً فى احتمال مجىء واحد آخر . وعلى ذلك فاحتمال مجىء الثلاثة هو حاصل ضرب الاحتمالات الثلاثة ، أى يساوى

$$\frac{٣}{٤} \times \frac{١}{٣} \times \frac{١}{٢} = \frac{١}{٨}$$

وهو احتمال أضعف من احتمال مجىء أى واحد على حدة .

مثال (٢) — ا و ب يلعبان الزهر فيرمى اللاعب زهرة ويكسب إذا رمى ٤ أو أكثر ، ويأخذ من صاحبه قرشاً ، ويتنازل عن اللعب إذا رمى أقل . فإذا بدأ ا باللعب وكان معه ثلاثة قروش فما هو احتمال أن يخسرها جميعها دون أن يكسب مرة واحدة .

هذا يتحقق لو بدأ ا اللعب فخسر ثم لعب ب فكسب ثلاث مرات متتالية حتى نفد ما يملكه ا .

احتمال الكسب فى أى رمية $= \frac{٣}{٦} = \frac{١}{٢}$

∴ احتمال أن ا يخسر عندما يبدأ اللعب $= \frac{١}{٢}$

و احتمال أن ب يكسب ثلاث مرات متتالية $= (\frac{١}{٢})^{٣}$

∴ احتمال أن ا يخسر ثم يكسب ب ثلاث مرات متتالية $= (\frac{١}{٢})^{٤}$

مثال (٣) — ما هو احتمال أن ا يخسر باستمرار وفى آخر فرصة يكسب فيسترد ما فقد دون أن يخسر ثانياً ؟

معنى ذلك أن ا يبدأ اللعب فيخسر ، ثم يرمى ب فيكسب مرتين متتاليتين ويخسر فى الثالثة فيحتفظ ا بقرشه الأخير ، ثم يلعب فيكسب مرتين متتاليتين و يسترد ما خسره أولا

∴ الاحتمال المطلوب $= \frac{١}{٢} \times (\frac{١}{٢})^{٢} \times \frac{١}{٢} \times (\frac{١}{٢})^{٢}$

$$= \frac{١}{٦٤}$$

احتمال الكسب فى عدد ر من المرات فقط والفشل فى الباقى

٣٨٨ — لنفرض أن احتمال الكسب فى لعبة معينة يساوى ح واحتمال الفشل أو الخسارة يساوى ل (ح + ل = ١) . فإذا حاولها لاعب ن من المرات فما هو احتمال أنه يكسب ر مرات ويخسر فى باقى المحاولات ، بصرف النظر عن ترتيب مرات الكسب والخسارة ؟

يمكننا اختيار المحاولات الناجحة ، وعددها ر ، من بين جميع المحاولات وعددها ن بطرق عددها

$$ق^{ن}_{ر} = \frac{\underline{|ن}}{\underline{|ر} \cdot \underline{|ن-ر}}$$

احتمال الكسب فى ر مرات متتالية يساوى

$$ح \times ح \times \dots ر \text{ مرات} = ح^{ر} ،$$

لأن كل محاولة مستقلة عن الأخرى ، والنجاح فى أى محاولة لا يؤثر مطلقاً فى احتمال نجاح أو فشل أى محاولة أخرى .

احتمال الفشل فى ن — ر مرات متتالية يساوى

$$ل \times ل \dots \text{إلى } (ن - ر) \text{ مرات} = ل^{ن-ر}$$

∴ احتمال الكسب فى أى ر من المرات والخسارة فى الباقى يساوى

$$ق^{ن}_{ر} \cdot ح^{ر} \cdot ل^{ن-ر}$$

مثال (١) رمى لاعب سبع زهرات مرة واحـدة ، فما هو احتمال حصوله على « شيش » من أربع زهرات فقط ؟

رمى سـبع زهرات هو بمثابة عمل سبع محاولات . واحتمال النجاح فى كل محاولة يساوى احتمال ظهور «شيش» من الزهرة أى $\frac{١}{٦}$ ؛ واحتمال الفشل يساوى $\frac{٥}{٦}$

$$\therefore \text{الاحتمال المطلوب} = {}^{٧}ق_{٤} \times \left(\frac{١}{٦}\right)^{٤} \times \left(\frac{٥}{٦}\right)^{٣}$$

$$= \frac{٤٣٧٥}{٢٧٩٩٣٦}$$

مثال (٢) دفع رجل ريال عن كل جواد من سـتة جياد من خيل السباق على أن يأخـذ ريالين عن كل جواد يربح ، عـلاوة على الريال الذى دفعه ، فما هو احتمال أنه يسترد ما دفعه بدون أن يكسب أو يخسر ؛ مع العلم بأن احتمال نجاح أى جواد من الستة يساوى $\frac{١}{٣}$ ؟

جملة ما دفعه ٦ ريالات ، ويمكنه أن يستردها كاملة إذا ربح اثنـان فقط من الستة جياد .

احتمال نجاح أى جوادين من الستة يساوى

$${}^{٦}ق_{٢} \times \left(\frac{١}{٣}\right)^{٢} \times \left(\frac{٢}{٣}\right)^{٤} = \frac{١٦٠}{٢٧٩}$$

توزيع ذات الحدين التكرارى

٣٨٩ — بتطبيق النتيجة التى حصلنا عليها فى بند ٣٨٨ ، نرى أن احتمالات الكسب عنـدما $ر = ن$ و $(ن - ١)$ و $(ن - ٢)$ و ٠٠٠ و ٢ و ١ و صفر ، هى على التوالى حدود المفكوك

$$(ح + ل)^{ن}$$

طبقاً لنظرية ذات الحدين . وهى مبينة فى الجدول الآتى :

احتمال نجاح ر من المرات حينما ر تأخذ قيما من ٠ إلى ن

قيمة ر	الاحتمال
ن	$\text{ح}^{\text{ن}}$
ن — ١	$^{\text{ن}}\text{ق}_{\text{ن}-1} \,.\, \text{ح}^{\text{ن}-1} \,.\, \text{ل}$
ن — ٢	$^{\text{ن}}\text{ق}_{\text{ن}-2} \,.\, \text{ح}^{\text{ن}-2} \,.\, \text{ل}^{2}$
⋮	⋮
٢	$^{\text{ن}}\text{ق}_{2} \,.\, \text{ح}^{2} \,.\, \text{ل}^{\text{ن}-2}$
١	$^{\text{ن}}\text{ق}_{1} \,.\, \text{ح} \,.\, \text{ل}^{\text{ن}-1}$
٠	$\text{ل}^{\text{ن}}$

وواضح أن مجموع هذه الاحتمالات كلها يساوى $(\text{ح} + \text{ل})^{\text{ن}} = 1$. وهذا هو المنتظر حيث إن هذا الجدول يشمل كل ما يمكن أن يحصل عند محاولة هذه اللعبة ن مرات .

توزيع ذات الحدين متماثل إذا كان ح = ل

٣٩٠ — وإذا تذكرنا أن معاملات الحدود فى مفكوك ذات الحدين تكون متساوية فى الحدود المتساوية البعد عن طرفى المفكوك ، أى أن .

$$^{\text{ن}}\text{ق}_{1} = {}^{\text{ن}}\text{ق}_{\text{ن}-1} \text{ ، } {}^{\text{ن}}\text{ق}_{2} = {}^{\text{ن}}\text{ق}_{\text{ن}-2} \text{ ، } \ldots$$

نجد أنه فى حالة ما يكون $\text{ح} = \text{ل} = \frac{1}{2}$ ، تكون هذه الاحتمالات المذكورة فى البند السابق موزعة توزيعاً متماثلا ، بمعنى أن الأول يساوى الأخير ، والثانى يساوى

قبل الأخير وهكذا . أى أن التوزيع الذى يمثله المفكوك

$$\left(\frac{1}{2} + \frac{1}{2} \right)^{ن}$$

يكون توزيعاً متماثلا ، بحيث لو رسمنا خطاً بيانياً فيه الإحداثيات الأفقية تمثل قيم ر والإحداثيات الرأسية تمثل الاحتمالات المناظرة لها ، حصلنا على منحن متماثل حول محور رأسى يقسمه إلى نصفين متطابقين .

أما إذا اختلفت قيمتا ح و ل فإن هذا التماثل يختل ، ونحصل على منحن ملتوٍ ، ويزداد هذا الالتواء كلما زاد الفرق بينهما .

على أن هذا المنحنى [1] سواء كان متماثلا أو ملتوياً ، ليس معتدلا بمعنى الاعتدال الذى ذكرناه فيما سبق ، ولكنه يتقرب إلى صفة الاعتدال المعروفة كلما كبرت ن . ويمكن القول — نظرياً على الأقل — إنه مهما كان الاختلاف بين ح و ل كبيراً فإن تكبير قيمة ن كفيل بأن يصلح الالتواء الناشىء فى المنحنى عن الاختلاف بين قيمتى ح و ل ويقرب المنحنى بقدر ما نريد من الاعتدال .

الوسط الحسابى لعدد مرات النجاح

٣٩١ — نرى من ذلك أن قيم ر — وهى عدد مرات النجاح أو الكسب — تختلف من صفر إلى ن وهى عدد المحاولات كلها؛ ونرى أن هذه القيم تختلف فى مقادير احتمالاتها . ومن الواضح أن القيمة الأكثر احتمالا هى الأرجح وهى الأكثر تكراراً إذا جربنا عددا كبيراً من المرات . فلايجاد متوسط قيم ر نرجح كل قيمة بمقدار احتمالها ، فنضرب كل قيمة فى الاحتمال ونقسم على مجموع

---

(١) هذا المنحنى الناتج نسميه منحنى ذات الحدين ، وبالإنجليزية Binomial Curve ؛ وتوزيع ذات الحدين يسمى Binomial Distribution، بخلاف المنحنى المعتدل المسمى Normal Curve. ولإثبات تقرب الأول إلى الثانى بتكبير قيمة ن أنظر كتاب: G. Yule and Kendall: *Introduction to the Theory of Statistics* (1946) p. 177.

الاحتمالات وهو يساوى ١ . فينتج[1] أن متوسط قيم ر يساوى ه ح .

أى أن الوسط الحسابى لعدد مرات الكسب يساوى عدد المحاولات مضروباً فى احتمال الكسب فى أى محاولة ، باعتبار أن نجاح أى محاولة لا يؤثر فى احتمال نجاح أو فشل أى محاولة أخرى ، و بفرض أن احتمال النجاح واحد فى كل المحاولات .

مثال (١) — رمى لاعب ١٢ زهرة دفعة واحدة وكرر هذه التجربة عدداً من المرات . فما هو متوسط عدد ما يحصل عليه من « الشيشات » لكل مرة ؟

احتمال رمى « شيش » من أى زهرة يساوى $\frac{١}{٦}$.

واللاعب إذ يرمى ١٢ زهرة دفعة واحدة كأنه يعمل ١٢ محاولة

∴ متوسط عدد ما يحصل من « الشيشات » فى كل رمية يساوى

$$١٢ \times \frac{١}{٦} = ٢$$

مثال (٢) — إذا اعتبر ظهور « جهار » أو أكثر من الزهرة نجاحاً . فما هو متوسط عدد الزهرات الناجحة فى المثال السابق ؟

احتمال النجاح هنا يساوى $\frac{٣}{٦} = \frac{١}{٢}$

وعدد المحاولات يساوى ١٢ أيضاً .

∴ متوسط عدد مرات النجاح يساوى

$$١٢ \times \frac{١}{٢} = ٦$$

---

(١) لإثبات ذلك نقول إن متوسط قيم ر

$$= \text{ه} \cdot \text{ح}^{\text{ن}} + (\text{ه} - ١) \cdot \text{ق}_{\text{ن}-١} \cdot \text{ح}^{\text{ن}-١} \cdot \text{ل}$$

$$+ (\text{ه} - ٢) \cdot {}^{\text{ن}}\text{ق}_{\text{ن}-٢} \cdot \text{ح}^{\text{ن}-٢} \cdot \text{ل}^{٢} + \cdots$$

$$\cdots + ١ \cdot {}^{\text{ن}}\text{ق}_{١} \cdot \text{ح} \cdot \text{ل}^{\text{ن}} + ٠ \times \text{ل}^{\text{ن}}$$

$$= \text{ه}\,\text{ح} \left[\text{ح}^{\text{ن}-١} + (\text{ه} - ١) \cdot \text{ح}^{\text{ن}-٢}\,\text{ل} + \cdots + \text{ل}^{\text{ن}-١} \right]$$

$$= \text{ه}\,\text{ح}\,(\text{ح} + \text{ل})^{\text{ن}-١}$$

$$= \text{ه}\,\text{ح} .$$

أى أن اللاعب سيحصل **فى المتوسط** على ست زهرات ناجحة فى كل رمية .

معنى الوسط الحسابى . تجربة ولدون

٣٩٢ — والمقصود بقولنا « إن الوسط الحسابى لعدد الزهرات الناجحة يساوى ٦ » هو أننا إذا قمنا فعلاً بتجربة فقد نحصل على أكثر أو أقل من ٦ زهرات ناجحة ، وربما نحصل على ست بالضبط إذا جرّبنا مراراً .

ولو كررنا هذه التجربة عدداً كبيراً من المرات فإن المتوسط يكون ٦ أو قريباً منها .

وقد قام و . ف . وِلْدونْ بتجربة (١) مثل هذه فرمى ١٢ زهرة دفعةً واحدة ٤٠٩٦ مرة ، ولاحظ فى كل رمية عدد الزهرات الناجحة (النجاح هو ظهور « جهار » أو « بيش » أو « شيش ») فحصل فى بعض الرميات على ٦ زهرات وفى بعضها على ٤ و ٥ و ٧ و ٨ وغيرها ، كما هو مبين فى التوزيع التكرارى الآتى :

جــدول ٥٣

عدد الزهرات الناجحة	التكرار المشاهد
٠	٠
١	٧
٢	٦٠
٣	١٩٨
٤	٤٣٠
٥	٧٣١
٦	٩٤٨
٧	٨٤٧
٨	٥٣٦
٩	٢٥٧
١٠	٧١
١١	١١
١٢	٠
	٤٠٩٦

(١) تجربة Weldon انظر W.F.R. F.C. Mills, *Statistical Methods*, (1924), p. 523.
أو G.U. Yule, *Introduction to the Theory of Statistics* (1937) p. 351, p. 424.
أو F.Y. Edgeworth, *Encyclopedia Britannica*, 11th. ed. Vol. 22, p. 39
وراجع بعض الملاحظات على هذه التجربة فيما يلى (بند ١٣٥) .

ومن هذه التجربة العملية نرى أن العدد يختلف من ١ إلى ١١ ، ونلاحظ أن العدد الأكثر تكراراً هو ٦ وهو الوسط الحسابى الذى حصلنا عليه نظرياً . والوسط الحسابى لهذا التوزيع التكرارى يساوى ٦,١٣٩ وهو قريب من الوسط النظرى ٦ ؛ ويكون أقرب من ذلك لو كان عدد الرميات أكبر من العدد هنا ، هذا بفرض أن الزهرات ليس بها خطأ طبعاً .

إيجاد الانحراف المعيارى

٣٩٣ — وكما أوجدنا الوسط الحسابى ن ح يمكننا[1] أن نوجد الانحراف

---

(١) لكى نوجد هذا الانحراف المعيارى لقيم ر — عدد المحاولات الناجحة فى التجربة انظر بند ٣٨٩ — نضرب مربع كل قيمة فى احتمال هذه القيمة (الذى يمثل تكرارها) ؛ ثم نطرح مربع الوسط الحسابى ن ح من مجموع هذه الحواصل ، فينتج مربع الانحراف المعيارى ، طبقاً للعلاقة المعروفة (انظر بند ١٨١ معادلة (٢)) . فبفرض أن احتمالات القيم ن ، ن — ١ ، هى و ن ، و ن — ١ ، ... ، ينتج أن مجموع حاصل ضرب مربعات القيم فى احتمالاتها أى مجـ ر^{٢} . و

= ن^{٢} . ح^{ن} + (ن — ١)^{٢} . ن . ح$^{ن—١}$ ل + (ن — ٢)^{٢} . $\frac{\text{ن (ن — ١)}}{٢}$ ح$^{ن—٢}$ ل^{٢} + ...

... + ١ × ن ح ل$^{ن—١}$ + ٠ × ل^{ن}

= ن ح [(ن — ١) ح { ح$^{ن—٢}$ + (ن — ٢) ح$^{ن—٣}$ ل + ... + ل$^{ن—٢}$ }

+ { ح$^{ن—١}$ + (ن — ١) ح$^{ن—٢}$ ل + ... + ل$^{ن—١}$ }]

= ن ح [(ن — ١) ح . (ح + ل)$^{ن—٢}$ + (ح + ل)$^{ن—١}$]

= ن ح [ن ح — ح + ١] ، وبوضع ١ — ح = ل

= (ن ح)^{٢} + ن ح ل .

$\therefore$ مربع الانحراف المعيارى ع^{٢} = مجـ و . ر^{٢} — (ن ح)^{٢}

$\therefore$ ع = $\sqrt{\text{ن ح ل}}$.

المعيارى ، وهو الذى يقيس لنا درجة التباين والاختلاف بين الأعداد التى يصح أن نحصل عليها . هذا الانحراف المعيارى $ع = \sqrt{ن\,ح\,ل}$.

مثال : المطلوب إيجاد الانحراف المعيارى لعدد الزهرات الناجحة فى مثال (٢) من بند ٣٩١ .

احتمال النجاح (أى ظهور « جهار » أو « بيش » أو « شيش ») فى أى زهرة يساوى $\frac{٣}{٦} = \frac{١}{٢}$ ؛ وعدد المحاولات يساوى عدد الزهر فى كل رمية وهو ١٢ .

وبما أن احتمال الفشل $ل = ١ - ح = \frac{١}{٢}$ ،

∴ الانحراف المعيارى المطلوب هو

$$\sqrt{ن\,ح\,ل} = \sqrt{١٢ \times \frac{١}{٢} \times \frac{١}{٢}}$$

$$= \sqrt{٣}$$

$$= ١{,}٧٣٢ \text{ تقريباً .}$$

تطبيق هذه النتائج فى دراسة العينات

استخدام الوسط الحسابى والانحراف المعيارى

٣٩٤ — هذا الانحراف المعيارى مهم جداً فى بحث العينات . فنعلم أن الانحراف المعيارى لأى توزيع تكرارى يقيس تشتت المفردات حول الوسط الحسابى ، ومقدار بعدها عنه أو قربها منه فى المتوسط . فى التوزيع التكرارى المعتدل نعلم[1] أن ٦٨٫٣ ٪ من المفردات لا تختلف عن الوسط الحسابى بأكثر من الانحراف المعيارى ، وأن ٩٥٫٥ ٪ من المفردات لا تختلف عن هذا الوسط بأكثر من ضعف الانحراف المعيارى ، وأن ٩٩٫٧ ٪ من المفردات

---

(١) انظر المذكرة عن المنحنى التكرارى المعتدل فى آخر الكتاب فى الملحق الرياضى .

لا يزيد الفرق بينها و بين الوسط الحسابى (سالباً كان أو موجباً) عن ثلاثة أمثال الانحراف المعيارى .

فمتى عرفنا الوسط الحسابى لأى مجموعة ، وهو ن ح ، والانحراف المعيارى لها وهو √ن ح ل ، نقول (على فرض أن هـذه المجموعة معتدلة أو قريبـة من الاعتدال) إن ٩٩٫٧٪ من مفرداتها لا تختلف عن الوسط الحسابى ن ح بأ كثر من ٣ √ن ح ل ، أى أنها تقع فى المنطقة ن ح ± ٣ √ن ح ل . أو بعبارة أخرى نقول إنه لا يقع خارج هذه المنطقة إلا ما كان شاذاً أو نادراً جداً (بنسبة ١ : ٣٧٠ تقريباً) . أما الفروق التى أقل من ٣ √ن ح ل . فيصح أن نتغاضى عنهـا ونعزوها إلى عوامل المصادفة . ولو أن بعض الإحصائيين يتردد فى ذلك نوعا ، و يقتصر على الفروق التى أصغر من ٢ √ن ح ل . ولكن الأسلم أن نأخذ الحد الفاصل يساوى ٣ √ن ح ل ، حيث إنه من المؤكد تقريباً أن عوامل المصـادفة وحدها لاتحدث فرقاً أو خطأ أ كبر من هذا المقدار . و بناء على ذلك فالمفردات التى تختلف عن الوسط الحسـابى لمجموعة (معتدلة) بأقل من ٣ √ن ح ل لا نعتبرها شاذة .

متوسط نسبة النجـــاح وانحرافهـــا المعيـــارى

٣٩٥ — و بدل أن نتكلم عن الوسط الحسابى والانحراف المعيارى لعـدد المحاولات الناجحـة فى أى تجر بة ، يمكننا أن نتكلم عن متوســط نسبة النجاح فى التجر بة ، والانحراف المعيارى لهذه النسبة .

و بديهى أن متوســط نسبة النجاح يساوى الوسط الحسـابى لعدد المرات الناجحة وهو ن ح ، مقسوماً على العـدد ن . أى أن الوسط الحسابى لنسبة النجاح يساوى ح .

و كذلك يكون الانحراف المعيارى لنسـبة النجاح هو الانحراف المعيـارى

الأصلى $\sqrt{ن ح ل}$. مقسوماً على العـدد ن نفسه ، فيكون الانحراف المعيارى لنسبة النجاح يساوى

$$\frac{\sqrt{ن ح ل}}{ن} = \sqrt{\frac{ح ل}{ن}}$$

أمثلة على العينات

٣٩٦ — نأخذ الآن بعض الأمثلة الإحصائية لتوضيح طريقة استخدام هذه النتائج فى المسائل العملية .

مثال (١) — فى سنة ١٩٤٢ كان عدد المواليد فى المملكة المصرية ٦٥٨٣٢٤ منهم ٣٤٣٦٣٣ ذكور . وفى نفس السنة كان عـدد المواليد بمديرية الجيزة ٣٢٢٧٧ منهم ١٦٦٨٣ ذكور . فهل يستدل من هذه الأرقام أن هنـاك اختلاف حقيق بين مديرية الجيزة وباقى المملكة المصرية فى نسبة الذكور بين المواليد .

نفرض أن مديرية الجيزة عينة عشوائية من سكان المملكة ليس فيها تحيز ، ويسرى على كل سكانها ما يسرى على سكان جميع المملكة .

احتمال أن أى مولود فى المملكة يأتى مولوداً ذكراً يساوى ح ،

حيث $ح = \frac{٣٤٣٦٣٣}{٦٥٨٣٢٤}$

$= ٠,٥٢٢$ تقريباً .

وعلى اعتبار أن مواليـد الجيزة وعددهم ٣٢٢٧٧ بمثابة محاولات لأمر معين احتمال نجاحه فى كل محاولة هو ح = ٠,٥٢٢

∴ عدد الذكور المنتظر من هؤلاء = ن ح ، حيث ن = ٣٢٢٧٧

و الانحراف المعيارى لهذا العـدد = $\sqrt{ن ح ل}$ ، حيث ل = ٠,٤٧٨

∴ ن ح = ٣٢٢٧٧ × ٠,٥٢٢

= ١٦٨٤٩ تقريباً .

و √ن ح ل = √٣٢٢٧٧ × ٥٢٢ر × ٤٧٨ر

= ٨٩ تقريباً .

ولكن عـدد الذكور فعلا هو ١٦٦٨٣ و يختلف عن العدد المنتظر بمقـدار ١٦٨٤٩ — ١٦٦٨٣ أى ١٦٦ .

و بما أن هذا الفرق ١٦٦ أقل من ٣ × ٨٩ = ٢٦٧ ، أى ثلاثة أمثال الانحراف المعيارى لعـدد الذكور المنتظر ، فإن هـذا الفرق يمكن أن يعزى للصدف .

و بناء على ذلك لا يمكن القول بأن هناك اختلاف حقيقى بين مديرية الجيزة والمملكة المصرية فى نسبة المواليد الذكور .

مثال (٢) — كان عدد مواليد المملكة المصرية فى سنة ١٩٤٢ هو كما ذكرنا فى المثال السابق ٦٥٨٣٢٤ منهم ٣٤٣٦٣٣ ذكور . فهل يستدل من ذلك على أن نسبة المواليـد الذكور لا تساوى النصف بالضبط و إنما تختلف عنه اختلافاً حقيقياً .

قد يقول البعض إن نسـبة الذكور بين المواليد هى فى الحقيقـة ٥٠ ٪ و إنما الاختـلاف الذى قد نلاحظه فى بعض الأحوال ناشىء عن مجرد المصادفة وليس له وجود حقيقى . ونحن الآن نختبر صحة هذا الفرض بأن نبحث إذا ما كان الواقع يبرر هذا الفرض أو يناقضه .

على فرض أن نسبة المواليد الذكور إلى مجموع المواليد = $\frac{1}{2}$ تماماً ، يكون احتمال كون المولود ذكراً = $\frac{1}{2}$ ؛ و بنـاء على ذلك إذا اعتبرنا كل حالة ميلاد كأنهـا محاولة وأن النجاح هو كون المولود ذكراً ، فإن احتمال النجاح = $\frac{1}{2}$ ، وكذلك احتمال الفشل = $\frac{1}{2}$.

∴ عدد المحاولات فى هذه المسألة يساوى عدد المواليد كلهم

∴ $ن = ٦٥٨٣٢٤$ ، $ح = \frac{١}{٢}$ ، $ل = \frac{١}{٢}$.

∴ $ن ح = ٣٢٩١٦٢$ وهو العدد المنتظر من الذكور

و $\sqrt{ن ح ل} = \sqrt{٦٥٨٣٢٤ \times \frac{١}{٤}}$

$= ٤٠٦$ وهو الانحراف المعيارى لهذا العدد المنتظر من الذكور .

ولكن عدد المواليد الذكور فعلا كان ٣٤٣٦٣٣

∴ الفرق $٣٤٣٦٣٣ - ٣٢٩١٦٢ = ١٤٣٧١$ ،

وهو أكبر من الانحراف المعيارى ثلاثين مرة ، فلا يمكن أن يعزى مثل هذا الفرق لمجرد الصدفة . و بناء على ذلك فإن الفرض بأن نسبة المواليد الذكور هى ٥٠ ٪ فرض لا يمكن أن يبرره الواقع (إلا نادراً جداً جداً ..) فهو خطأ و يجب رفضه.

مثال (٣) — فى سنة ١٩٣٤ كانت نسبة النجاح فى مجموع المدارس الأميرية فى امتحان شهادة الدراسة الثانوية قسم أول تساوى ٦٢٫٨ ٪ ، وكانت نسبة النجاح فى مدرسة فؤاد الأول الثانوية ٦٠٫٨ ٪ ، وكان عدد من حضروا الامتحان منها ١٥٨ تلميذاً . فهل فى هذه الأرقام ما يثبت أن مدرسة فؤاد الأول حقيقة تحت المتوسط من بين المدارس الأميرية أو أن هذا الفرق راجع إلى المصادفات ؟ و يصح أن ينعكس الوضع بالنسبة إلى هذه المدرسة فى سنة أخرى .

نعتبر مدرسة فؤاد الأول بمثابة عينة من المدارس الأميرية على العموم حيث النسبة المتوسطة للنجاح تساوى ٦٢٫٨٪ ، أو ٠٫٦٢٨ ومعنى ذلك أن المدارس الأميرية الأخرى عينات ونسبها تتراوح حول هذا المتوسط . احتمال نجاح أى تلميذ

من المدارس الأميرية على العموم كان فى تلك السنة ٠,٦٢٨ — وهو متوسط نسبة النجاح التى ذكرناها فى بند ٣٩٦ ، وبما أن العـدد فى هذه العينة هو ١٥٨ تلميذاً ، يكون الانحراف المعيارى لنسبة النجاح هو

$$ع = \sqrt{\frac{ح\ ل}{ن}} = \sqrt{\frac{١}{١٥٨} \times ٠,٦٢٨ \times ٠,٣٧٢}$$

$$= ٠,٠٣٨٦ \text{ تقريباً}$$

$$\therefore \quad ٣ع = ٠,١١٥٨ \quad «$$

ولـكن الفرق بين نسبة المدرسة والنسبة العامة هو

$$٦٢٨, - ٦٠٨, = ٠,٠٢$$

وهو فرق **أقل** من الانحـراف المعيارى نفسـه ، فضلاً عن ثلاثة أمثاله . وينتج من ذلك أن هذا الفرق يصـح أن يكون بمجرد المصـادفة ، ويصح أن المصادفات تأتى بعكس ذلك فى السنين الأخرى . وفعلاً نجد [1] أن النتيجة العامة للمدارس الأميرية فى سـنة ١٩٣٣ كانت ٥٨ ٪ وكانت نتيجة مدرسة فؤاد الأول ٦٦,٧ ٪ .

الخطأ المعيارى كمقيـاس للفروق المعنوية وغيرها

٣٩٧ — وهكذا يمكننا أن نعـدد هذه الأمثلة لتطبيق هاتين النتيجتين المشهورتين ، ألا وهما : أن الوسط الحسابى لعدد مرات النجاح يساوى ن ح ، والانحراف المعيارى يسـاوى $\sqrt{ن\ ح\ ل}$ ، أو بعبارة أخرى ، متوسط نسبة النجاح يساوى ح ، وانحرافها المعيارى $\sqrt{\frac{ح\ ل}{ن}}$. وفى كل مسألة يعطينا الوسط

---

(١) انظر الإحصاء العام لامتحان شهادة الدراسة الثانوية (قسم أول لسنتى ١٩٣٣ و١٩٣٤) عمل وزارة المعارف العمومية (صفحتى ٢ و ٣) .

الحسابى الرقم المنتظر ، والانحراف المعيارى يعطينا مقياساً لدرجة الاختلاف أو التباين الذى يمكن أن تأتى به المصادفات بين عينة وأخرى . أو بعبارة أخرى يعطينا الحد الفاصل بين الاختلافات والفروق التى يمكن أن نعزوها إلى المصادفات فلا نهتم بها ، والفروق الأخرى التى لا يمكن اعتبارها ناشئة عن المصادفات نظراً لعظمها ، فنستدل منها على وجود عوامل أخرى . وللتمييز نقول **فروق ظاهرية** (١) أى صغيرة ، **وفروق معنوية** (٢) .

وهكذا نرى أن الانحراف المعيارى هو بمثابة مسطرة أو أداة القياس التى نقيس بها الفروق أو الأخطاء لنعلم مقدار أهميتها وما تدل عليه . و يسمى أحياناً **الخطأ المعيارى** (٣) أو **معيار الخطأ** ، وهو المقياس الأساسى ، والمفضل على غيره .

الخطأ المتعادل

٣٩٨— و يستخدم بعض الإحصائيين مقياساً آخر نسميه (٤) **الخطأ المتعادل**. وهو مشتق من الانحراف المعيارى أو الخطأ المعيارى الذى رمزنا له بالحرف ع ، وهذا الخطأ الجائز أو المتعادل يساوى

$$ى = ٠,٦٧٤٥ \text{ ع}$$
$$= \tfrac{٢}{٣} \text{ ع تقريباً}$$

---

(١) بالإنجليزية Apparent or Insignificant Differences

(٢) بالإنجليزية Significant Difference

(٣) بالإنجليزية Standard Error

(٤) بالإنجليزية Probable Error ، وترجمتها الحرفية « الخطأ المحتمل » . وهى ترجمة يؤخذ عليها عيب العبارة الإنجليزية ولا تؤدى المعنى المقصود من تعريف هذا الخطأ كما هو مذكور أعلاه . فاقترحت واستعملت العبارة « الخطأ الجائز » مدة من الزمن . ثم اقترح الأستاذ الدكتور مشرفة باشا العبارة « خطأ متعادل » فى حديث معى حوالى سنة ١٩٣٧ — وهى فى نظرى أفضل .

والأصل فى اختيار هذا المقياس أن من خواص[١] التوزيع التكرارى المعتدل (التى ذكرنا بعضها فى بند ٣٩٤) أن ٥٠ ٪ من مجموع المفردات فى هذا التوزيع لا تختلف عن الوسط الحسابى (بالعجز أو بالزيادة) بأكثر من المقدار ٠٫٦٧٤٥ ع . وهذا معناه أن احتمال كون الفرق بين مفردة مّا والوسط الحسابى أقل من ٠٫٦٧٤٥ ع يساوى $\frac{1}{2}$ ؛ وكذلك احتمال أن هذا الفرق أو الخطأ يزيد عن ٠٫٦٧٤٥ ع ، يساوى $\frac{1}{2}$. وبعبارة أخرى : احتمال أن الخطأ فى أى مفردة (أى الفرق بينها وبين الوسط الحسابى للمجموعة كلها) يقل عن ء يساوى احتمال أنه يزيد عن ء نفسها . ومن ثم كانت التسمية الإنجليزية Probable Error ولكن هذه التسمية ليست موفقة مع الأسف ؛ لأنها لا تؤدى المعنى المقصود منها والمذكور أعلاه ، حيث كل خطأ « محتمل » الحصول . ولكن كثرة الاستعمال والصفة التاريخية أكسبت هذه العبارة معنى اصطلاحياً معروفاً ومستقلا عن — بل متناقضاً مع — المعنى المفهوم من الكلمتين المركبتين لها . ويظهر أن بعض الإحصائيين رأوا حديثاً أن يتخلصوا من هذه الورطة بأن يقلعوا[٢] عن استعمال هذا المقياس ء = $\frac{2}{3}$ ع. فهم يفضلون الآن الاعتماد على المقياس الأصلى وهو الخطأ المعيارى ع نفسه ، ويستخدمون المقدار ٣ ع كما قدمنا .

٣٩٩ — ولا يفوتنا أن نذكر القارىء بالملاحظة التى أشرنا إليها قبلا (فى بند ٣٩٤) بخصوص استخدام هذه المقاييس ، ألا وهى أن هذه النتائج التى نصل إليها مبنية على فرض أن التوزيع التكرارى للكمية ن ح (وهى تساوى العدد المنتظر أو متوسط عدد مرات النجاح فى العينات المختلفة التى يصح

---

(١) أنظر المذكرة عن المنحنى التكرارى المتماثل فى الملحق الرياضى فى آخر الكتاب .

(٢) أنظر كتاب G.U. Yule, *Introduction to the Theory of Statistics*, 1937 p. 353, art. 19.09.

أن نأخذها من المجتمع الذى نبحثه) هو توزيع معتـدل أو على الأقل قريب من الاعتدال ؛ أما إذا كان هناك ما يشير إلى أن هـذا التوزيع التكرارى بعيد عن الاعتدال فيجب أن نحترس عند استخدام الحـدود ع و ٢ ع و ٣ ع والاحتمالات المناظرة لها التى ذكرناها (فى بند ٣٩٥) . ففى حالة التوزيع غير المعتدل يجب تعديل هذه النسب (٦٨,٣٪ ، ٩٥,٥٪ ، ٩٩,٧٦٪) وربما نحتاج فى بعض الأحيان إلى اتخاذ مقياس آخر غير ٣ ع الذى أخذناه فى الأمثلة التى شرحناها . وربما احتجنا إلى طريقة مخالفة بالمرة للتى استخدمناها فى تلك الأمثلة . ومع أن غالبية المسائل التى نعالجها تنطبق عليها هذه القاعدة ولو إلى حد ما ، فيجب التأكد فى كل حالة من أن التوزيع ليس بعيداً جـداً عن الاعتدال . ويكفى للتأكد من هذا الاعتـدال أن يكون احتمال النجـاح ح يساوى ٠,٥ أو قريباً منها ؛ وإن كانت ح بعيـدة نوعاً عن ٠,٥ فيجب أن تـكون ن كبيرة كبراً كافياً ليكون التوزيع قريباً من الاعتدال كما قلنا فى بند ٣٩٠ .

المـــراجع

Bowley, A.L. : *Elementary Manual of Statistics*, Chapter VII.

— : *Elements of Statistics*, Chapter, II, Part II.

Connor, L.R. : *Statistics in Theory and Practice*, Chapter XIV.

Hall and Knight: *Higher Algebra*, Chapter XXXII.

Jones, D.C. : *First Course in Statistics*, Chapter XII.

Mills, F.C. : *Statistical Methods*, Chapters XV, XVI.

Secrist, H. : *Statistical Methods*, Chapter XI.

Yule, G.U. : *Introduction to the Theory of Statistics*, Chapters XVII—XXI.

الباب الخامس عشر

التوزيعات التكرارية للعينات

٤٠٠ — شرحنا فى الباب السابق بعض نتائج نظرية الاحتمالات، وكيفية تطبيق هذه النتائج فى مسائل العينات. وهناك تطبيقات مهمة أخرى لهذه النتائج، وسنشرح فى هذا الباب التطبيقات الخاصة بالتوزيع التكرارى للعينات.

تجربة ولدون

نعود لحظة إلى تجربة ولدونْ حيث رمى ١٢ زهرة دفعة واحدة ٤٠٩٦ مرة، ولاحظ فى كل رمية عدد الزهرات الناجحة، باعتبار النجاح هو ظهور أحد الأوجه الثلاثة: « جهار » و « بيش » و « شيش ». وحصل ولدونْ من هذه التجربة على توزيع تكرارى لعدد الزهرات الناجحة وهو المبين فى الجدول الآتى:

ولو نظرنا إلى عمود التكرارات المشاهدة فى هذا الجدول نجد مثلاً أن ولدونْ حصل على خمس زهرات ناجحة فى ٧٣١ رمية من الرميات كلها. والآن يصح لنا أن نتساءل:

لو أن رجالا كثيرين غير ولدون قاموا بمثل هذه التجربة، فرمى كل واحد منهم ١٢ زهرة ٤٠٩٦ مرة، فهل كانوا يحصلون على خمس زهرات ناجحة ٧٣١ مرة كما فعل ولدون؟ وإذا اختلفت نتائجهم فحول أى وسط تتراوح هذه النتائج، وما هى حدود الاختلاف؟ أو بعبارة أخرى ما هو الوسط الحسابى والانحراف المعيارى لتكرار هذه الفئة من الجدول التكرارى. وهذان كافيان لتحديد مركز اختلاف نتائج هؤلاء الرجال ودرجة هذا الاختلاف، كما قلنا فى الباب السابق.

جدول ٥٤ — تجربة ولدون . رمى ١٢ زهرة ٤٠٩٦ مرة

عدد الزهرات الناجحة	التكرار المشاهد فى التجربة	التكرار النظرى حسب الاحتمالات
٠	٠	١
١	٧	١٢
٢	٦٠	٦٦
٣	١٩٨	٢٢٠
٤	٤٣٠	٤٩٥
٥	٧٣١	٧٩٢
٦	٩٤٨	٩٢٤
٧	٨٤٧	٧٩٢
٨	٥٣٦	٤٩٥
٩	٢٥٧	٢٢٠
١٠	٧١	٦٦
١١	١١	١٢
١٢	٠	١
	٤٠٩٦	٤٠٩٦

ويمكننا معرفة هذا الوسط الحسابى وهذا الانحراف المعيارى بتطبيق القاعدة المعروفة ن ح و $\sqrt{\text{ن ح ل}}$ ، لو عرفنا الاحتمال ح ؛ وهو فى هذه الحالة يساوى احتمال الحصول على خمس زهرات ناجحة من ١٢ نرميها حيثما اتفق ، وهذا يساوى

$$^{١٢}\text{ق}_{٥} \times \left(\frac{١}{٢}\right)^{٥} \left(\frac{١}{٢}\right)^{٧} = \frac{٧٩٢}{٤٠٩٦}$$

وذلك بتطبيق القاعدة التى أثبتناها فى بند ٣٨٨ ، مع العلم بأن احتمال نجاح أى زهرة يساوى $\frac{٣}{٦} = \frac{١}{٢}$.

∴ احتمال الحصول على ٥ زهرات ناجحة فى أى رمية يساوى $^{١٢}\text{ق}_{٥} \left(\frac{١}{٢}\right)^{١٢}$

و بما أن عدد الرميات هنا يساوى ٤٠٩٦ ، ينتج أن الوسط الحسابى لعدد مرات الحصول على ٥ زهرات ناجحة يساوى

$$٤٠٩٦ \times ق_{٥}^{١٢} \times (\tfrac{١}{٢})^{١٢} = ٧٩٢$$

وهذا هو التكرار النظرى ، أو التكرار المتوسط الذى ينتظر أن نحصل عليه حسب نظرية الاحتمالات .

والانحراف المعيارى يساوى

$$\sqrt{ن\ ح\ ل} = \sqrt{٤٠٩٦ \times \frac{٧٩٢}{٤٠٩٦} \times \frac{٣٣٠٤}{٤٠٩٦}}$$

$$= ٢٥{,}٣ \text{ تقريباً}$$

وبذلك نحصل على مركز « تذبذب » تكرار هذه الفئة و « سعة ذبذبته » .

وبالمثل بالنسبة إلى باقى الفئات فى الجدول ، فنحصل على التكرارات النظرية التى نراها فى العمود الثالث من الجدول ٥٤ .

التغير فى تكرار أى فئة على حساب تكرارات الفئات الأخرى

٤٠١ — كل واحدة من هذه التجارب يمكننا اعتبارها كعينة . ولكل عينة توزيع تكرارى . وفى هذه التوزيعات التكرارية نرى أن تكرار الفئة ٥ مثلاً ، يختلف من عينة إلى أخرى و يتراوح حول قيمة معينة (هى ٧٩٢ فى هذه الحالة). ولما كانت كل التوزيعات متساوية فى مجموع التكرارات — وهو ٤٠٩٦ فى كل حالة — فمن الواضح أنه إذا زاد تكرار هذه الفئة ٥ فى عينة ما ، تكون هذه الزيادة على حساب فئة أو فئات أخرى فى نفس العينة . وهذا يشير إلى أنه لابد من وجود علاقة أو ارتباط بين تكرارات الفئات المختلفة فى نفس العينة ، والتغيرات التى تحصل فى هذه التكرارات ؛ ومن المهم دراسة هذه العلاقة والوقوف على حقيقتها .

٤٠٢ — و بوجه عام نفرض أن لدينا مجتمعاً كبيراً ، وأن مفردات هذا المجتمع لها توزيع تكرارى مقسم إلى فئات كالمعتاد . ولتكن تكرارات هذه الفئات هى :

$$ص_{١} ،\ ص_{٢} ،\ ص_{٣} ،\ ...... ،\ ص_{م} ،\ ......$$

حيث مجـ ص = حـ = عدداً كبيراً

نأخذ من هذا المجتمع الكبير عينة صغيرة (بدون تحيز) عدد مفرداتها ن مثلا ، ونفرض أن هذه المفردات موزعة فى فئات تكرارية مثل فئات المجتمع الأصلى تماماً . ولتكن تكرارات هذه الفئات هى :

$$ك_{١} ،\ ك_{٢} ،\ ك_{٣} ،\ ... ،\ ك_{م} ،\ ...$$

حيث مجـ ك = ن = عدداً صغيراً نسبياً

إذا عرفنا التكرارات ص ومجموعها حـ فى المجتمع الأصلى ، يكون احتمال أن أى مفردة من مفرداته تقع فى الفئة الميمية مثلاً يساوى $\frac{ص_{م}}{حـ}$.

أما إذا كنا لا نعرف هذه التكرارات ص ، وهذا هو الغالب فى كل هذه المسائل ، فيمكننا أن نستعيض عنها بالتكرارات ك التى نجدها فى العينة . وهذا جائز على وجه التقريب مادامت العينة عشوائية ولا تحيز فيها . فيكون لدينا قيمة تقريبية للاحتمال وهى $\frac{ك_{م}}{ن}$. ويكون الوسط الحسابى لتكرار الفئة الميمية يساوى $ك_{م}$ تقريباً .

و بناء على القاعدة المعروفة للانحراف المعيارى ($= \sqrt{ن\ ح\ ل}$) يكون الانحراف المعيارى لتكرار هذه الفئة الميمية ، وهو $ع_{م}$ مثلاً ، يساوى

$$ع_{م} = \sqrt{ك_{م} \left(١ - \frac{ك_{م}}{ن} \right)} \qquad (١)$$

وهذا يعطينا مقياساً لمدى الاختلاف فى تكرار أى فئة بين عينة وأخرى .

حساب معامل الارتباط بين تكرارى فئتين فى عينة

٤٠٣ — وإذا أخذنا من هذا المجتمع عدة عينات ففى أغلب الأحوال يتغير تكرار هذه الفئة الميمية من عينة إلى أخرى . وقد قلنا إنه إذا زاد هذا التكرار فى عينة ما عن الوسط الحسابى المقرر ، فإن هذه الزيادة تكون على حساب

فئة أو فئات أخرى . أى أن الزيادة فى تكرار فئة يصحبها نقص فى تكرار فئة أو فئات أخرى . وننتظر بناء على ذلك أن معامل الارتباط بين تكرارى أى فئتين فى عينة — الميمية واللامية مثلا — يكون سالباً . ولو رمزنا لمعامل الارتباط بين تكرارى هاتين الفئتين بالرمز $\text{ر}_{\text{م ل}}$ يمكننا [1] إثبات أن

$$\text{ر}_{\text{م ل}} = - \frac{1}{\text{ن}} \times \frac{\text{ك}_{\text{م}} \cdot \text{ك}_{\text{ل}}}{\text{ع}_{\text{م}} \cdot \text{ع}_{\text{ل}}}$$

حيث $\text{ك}_{\text{م}}$ و $\text{ك}_{\text{ل}}$ هما تكرار للفئتين الميمية واللامية على الترتيب . و $\text{ع}_{\text{م}}$ و $\text{ع}_{\text{ل}}$ هما الانحرافان المعياريان لهذين التكرارين (أنظر المعادلة (١) من البند السابق) و ن عدد مفردات العينة .

---

(١) لإثبات ذلك نفرض أن تكرار الفئة الميمية زاد عن الوسط الحسابى أو التكرار المنتظر $\text{ك}_{\text{م}}$ بمقدار يساوى $\text{ف}_{\text{م}}$ مثلا . هذا الفرق يجب أن يعوض بنقص فى مجموع تكرارات الفئات الأخرى ، حتى يكون المجموع الكلى لتكرارات العينة يساوى ن كما هو . فلو فرضنا أن $\text{ف}_{\text{ل}}$ هو نصيب الفئة اللامية من هذا النقص ، يكون

$$\text{ف}_{\text{ل}} = - \text{ف}_{\text{م}} \times \frac{\text{ك}_{\text{ل}}}{\text{ن} - \text{ك}_{\text{م}}}$$

ويكون $$\text{ف}_{\text{ل}} \cdot \text{ف}_{\text{م}} = - \frac{\text{ك}_{\text{ل}} \cdot \text{ك}_{\text{م}}}{\text{ن}} \cdot \frac{\text{ف}_{\text{م}}^{2}}{\text{ك}_{\text{م}} \left(1 - \frac{\text{ك}_{\text{م}}}{\text{ن}}\right)}$$

$$= - \frac{\text{ك}_{\text{ل}} \cdot \text{ك}_{\text{م}}}{\text{ن}} \cdot \frac{\text{ف}_{\text{م}}^{2}}{\text{ع}_{\text{م}}^{2}}$$

ولو فرضنا أننا أخذنا عينات كثيرة ينتج فى كل واحدة فرق $\text{ف}_{\text{م}}$ بين تكرار الفئة الميمية ووسطه الحسابى ، وفرق آخر $\text{ف}_{\text{ل}}$ بين تكرار الفئة اللامية ووسطه الحسابى . ومن بين هذه العينات نحصل على مج $\text{ف}_{\text{ل}} \cdot \text{ف}_{\text{م}}$. ونعرف أن مج $\text{ف}_{\text{م}}^{2}$ يساوى مربع الانحراف المعيارى $\text{ع}_{\text{م}}$ لتكرار الفئة الميمية مضروباً فى عدد العينات المأخوذة . وبقسمة مج $\text{ف}_{\text{ل}} \cdot \text{ف}_{\text{م}}$ على حاصل ضرب الانحرافين المعياريين $\text{ع}_{\text{م}}$ و $\text{ع}_{\text{ل}}$ ، نحصل على النتيجة المطلوبة .

الوسط الحسابى للعينة

الوسط الحسابى للعينة وانحرافه المعيارى

٤٠٤ — نبحث الآن فى مقدار الوسط الحسابى الذى نحصل عليه من دراسة مفردات العينة ، وقربه أو بعده عن الوسط الحسابى للمجتمع الأصلى المأخوذة منه العينة : وطريقة هذا البحث أن نأخذ عدة عينات ونحسب لكل منها وسطها الحسابى ، ثم نوجد الانحراف المعيارى لهذه الأوساط الحسابية ، فيعطينا مقياساً لدرجة اختلافها . ومن ثم نعرف مقدار بعد أى واحد منها عن الوسط الحسابى العمومى للمجتمع الذى هو فى الوقت نفسه المتوسط الحسابى لهذه المتوسطات العيِّنية .

لنفرض أن العينة النموذجية التى تمثل المجتمع الأصلى تمثيلا صحيحاً تحتوى على مفردات عددها ن موزعة فى جدول تكرارى ينقسم إلى عدة فئات تكرارية مراكزها هى :

$$س_1 \text{ و } س_2 \text{ و } س_3 \text{ و} \ldots$$

ونفرض أن تكرارات هذه الفئات فى هذه العينة هى على الترتيب :

$$ك_1 \text{ و } ك_2 \text{ و } ك_3 \text{ و} \ldots$$

∴ الوسط الحسابى لقيمة س فى هذه العينة يساوى

$$\bar{س} = \frac{س_1 ك_1 + س_2 ك_2 + \ldots}{ك_1 + ك_2 + \ldots} \text{ ، وبوضع } \sum ك = ن$$

$$= \frac{1}{ن} \times \sum س ك \qquad (1)$$

نأخذ عينة ثانية من نفس المجتمع ونقسم مفرداتها فى جدول تكرارى بنفس الفئات . وستكون تكرارات هذه العينة فى الغالب مخالفة لتكرارات العينة الأولى . لنفرض أن هذه التكرارات هى :

$$ك_1 + ف_1 \text{ ، } ك_1 + ف_2 \text{ ، } ك_3 + ف_3 \text{ ، } \ldots$$

حيث مج ف = صفراً ، لأن مجموع التكرارات واحد فى العينتين = ن .

الوسط الحسابى لهذه العينة الثانية يختلف فى الغالب عن الوسط الحسابى للعينة الأولى . لنفرض أنه يساوى $\overline{س} + ى$ ، حيث ى كمية صغيرة .

$$\overline{س} + ى = \frac{س_١(ك_١ + ف_١) + س_٢(ك_٢ + ف_٢) + \cdots}{ن} \qquad (٢)$$

حيث مج (ك + ف) = مج ك = ن = عدد مفردات العينة

و بطرح المعادلتين (١) و (٢) ينتج أن

$$ى = \frac{١}{ن}\,(ف_١ . س_١ + ف_٢ . س_٢ + \cdots) \qquad (٣)$$

= انحراف الوسط الحسابى للعينة الثانية عن الوسط الحسابى $\overline{س}$.

فلايجاد الانحراف المعيارى للأوساط الحسابية للعينات ، نحسب لكل عينة انحراف وسطها الحسابى عن الوسط النموذجى $\overline{س}$. ونربع هذه الانحرافات ، ونقسم مجموع مربعاتها على عدد العينات — وليكن ح مثلا — ينتج مربع الانحراف المعيارى لهذه الأوساط الحسابية ، الذى يقيس لنا تشتتها حول الوسط النموذجى $\overline{س}$ ، وهو الوسط الذى تتذبذب حوله متوسطات العينات ، وهو يساوى الوسط الحسابى للمجتمع الأصلى .

هذا الانحراف المعيارى للأوساط الحسابية يساوى $ع_و$. ويمكننا[١] إثبات أن

$$ع_و = \frac{ع.}{\sqrt{ن}} \qquad (٤)$$

---

(١) لإثبات ذلك نربع طرفى المعادلة (٣) أعلاه

$$\therefore \quad ى^٢ = \frac{١}{ن^٢}\,(ف_١^٢ س_١^٢ + ف_٢^٢ س_٢^٢ + \cdots$$

$$+ ٢ ف_١ ف_٢ . س_١ س_٢ + ٢ ف_١ ف_٣ . س_١ س_٣ + \cdots)$$

والحروف س واحدة فى جميع العينات . أما ى و $ف_١$ فتتغير من عينة إلى أخرى كذلك $ف_٢$ . . .

*

حيث $\text{ع}_{.}$ هى الانحراف المعيارى لقيم س فى العينة النموذجية أو المجتمع الأصلى و ن عدد مفردات العينة . وهذه النتيجة مهمة جداً ولها تطبيقات كثيرة .

$$\therefore\quad \text{٠٫٦٧٤٥}\ \text{ع}_{\text{و}} = \text{٠٫٦٧٤٥}\ \frac{\text{ع}_{.}}{\sqrt{\text{ن}}}$$

وهو الخطأ المتعادل للوسط الحسابى للعينة .

$$* \quad \therefore\ \text{مج}\ \text{د}^{2} = \frac{1}{\text{ن}^{2}}\Big(\text{س}_{1}^{2}\ \text{مج}\ \text{ف}_{1}^{2} + \text{س}_{2}^{2}\ \text{مج}\ \text{ف}_{2}^{2} + \cdots$$
$$+\ \text{٢}\ \text{س}_{1}\text{س}_{2}\ \text{مج}\ \text{ف}_{1}\text{ف}_{2} + \text{٢}\ \text{س}_{1}\text{س}_{3}\ \text{مج}\ \text{ف}_{1}\text{ف}_{3} + \cdots\Big)$$

حيث $\text{مج}\ \text{ف}_{1}^{2}$ = مجموع مربعات فروق تكرار الفئة الأولى فى العينات كلها ،

$$= \text{ح}\ .\ \text{ع}_{1}^{2} = \text{ح}\ .\ \text{ك}_{1}\left(1 - \frac{\text{ك}_{1}}{\text{ن}}\right)$$ ، ح عدد العينات ،

و ك_{1} متوسط تكرار الفئة الأولى .

$$\text{و}\ \text{مج}\ \text{ف}_{1}\text{ف}_{2} = -\ \text{ح}\ \frac{\text{ك}_{1}\text{ك}_{2}}{\text{ن}}$$ (أنظر $\text{ف}_{\text{ل}}\ \text{ف}_{\text{م}}$ فى حاشية بند ٤٠٣)

$$\text{و}\ \text{مج}\ \text{د}^{2} = \text{ح}\ .\ \text{ع}_{\text{و}}^{2}$$

$$\therefore\ \text{ع}_{\text{و}}^{2} = \frac{1}{\text{ن}^{2}}\Big[\text{س}_{1}^{2}\ \text{ك}_{1}\left(1 - \frac{\text{ك}_{1}}{\text{ن}}\right) + \text{س}_{2}^{2}.\ \text{ك}_{2}\left(1 - \frac{\text{ك}_{2}}{\text{ن}}\right)$$
$$+ \cdots - \text{٢}\ \text{س}_{1}\text{س}_{2} \times \frac{1}{\text{ن}}\ \text{ك}_{1}\text{ك}_{2}$$
$$-\ \text{٢}\ \text{س}_{1}\text{س}_{3}\ .\ \frac{1}{\text{ن}}\ .\ \text{ك}_{1}\text{ك}_{3} - \cdots\Big]$$
$$= \frac{1}{\text{ن}^{2}}\Big[\text{س}_{1}^{2}\text{ك}_{1} + \text{س}_{2}^{2}\text{ك}_{2} + \text{س}_{3}^{2}\text{ك}_{3} + \cdots$$
$$-\ \frac{1}{\text{ن}}\left(\text{س}_{1}^{2}\text{ك}_{1}^{2} + \text{س}_{2}^{2}\text{ك}_{2}^{2} + \cdots\right.$$
$$\left. +\ \text{٢}\ \text{س}_{1}\text{س}_{2}\text{ك}_{1}\text{ك}_{2} + \cdots\right).\Big]$$
$$= \frac{1}{\text{ن}^{2}}\Big[\text{ن}\ \text{ع}_{.}^{2} + \text{ن}\ \overline{\text{س}}^{2}$$
$$-\ \frac{1}{\text{ن}}\left(\text{س}_{1}\text{ك}_{1} + \text{س}_{2}\text{ك}_{2} + \cdots\right)^{2}\Big]$$

$$\text{لأن}\ \text{مج}\ \overline{\text{س}^{2}}\ \text{ك} = \text{ن}\ \text{ع}_{.}^{2} + \text{ن}\ \text{س}_{.}^{2}$$ أنظر العلاقة (٢) بند ١٨١

$$\therefore\ \text{ع}_{\text{و}}^{2} = \frac{1}{\text{ن}}\ \text{ع}_{.}^{2}$$

حيث $\text{ع}_{.}$ هو الانحراف المعيارى لقيم $\text{س}_{.}$ فى العينة الأولى المأخوذة كنموذج .

غير أنه فى كثير من الأحيان لا يتيسر لنا معرفة $ع_{و}$ وهى الانحراف المعيارى للعينة النموذجية أو المجتمع الأصلى، فنستعيض عنها بالانحراف المعيارى ع للعينة التى بأيدينا.

فائدة الخطأ المعيارى فى تحديد دقة الوسط الحسابى

٤٠٥ — وهذه المقاييس تقيس لنا مقدار الخطأ فى الوسط الحسابى للعينة ؛ أى مقدار بعده عن الوسط الحسابى للمجتمع الأصلى المأخوذة منه العينة . حيث نقول طبقاً للقاعدة التى شرحناها (فى بند ٣٩٤) ، إن الوسط الحسابى العمومى لا يختلف عن الوسط الحسابى للعينة بأكثر من المقدار $\pm$ ٣ $ع_{و}$.

وهكذا قد حصلنا على دليل دقيق نهتدى بواسطته عند تعميم خواص العينة وتطبيقها على المجتمع الأصلى ، ونعرف مدى الخطأ الذى يصح أن نقع فيه .

خطأ الوسط الحسابى يقل كلما كبرت العينة

٤٠٦ — ويلاحظ أن $ع_{و}$ أى الخطأ المعيارى للوسط الحسابى ، وكذلك الخطأ المتعادل له ، ينقص كلما زادت ن ، أى كلما كبرت العينة . أى أن دقة النتائج التى نحصل عليها من العينة تزداد كلما كبرت العينة . وهذا معقول ، حيث إنه كلما كبرت العينة وزاد عدد مفرداتها ، كانت تمثل المجتمع تمثيلاً أدق وتجمع بين عناصره المختلفة . وكما أن الخطأ فى الوسط الحسابى يتناسب مع الكمية $\frac{1}{\sqrt{ن}}$ ، يصح أن نقول إن دقة الوسط الحسابى تتناسب مع مقلوب هذه الكمية ، أى مع $\sqrt{ن}$.

٤٠٧ — مثال : الجدول الآتى يبين التوزيع التكرارى لأوزان ١٢١٢ رجلاً اسكتلندياً [١] .

فماذا يكون الوسط الحسابى لأوزان الرجال الأسكتلنديين على العموم ؟

الوسط الحسابى للأوزان فى هذه العينة يساوى ١٦٤,٨٦٠ رطلاً ، والانحراف المعيارى يساوى ٢٠,٣٢٥ رطلاً . فينتج أن الانحراف (الخطأ) المعيارى للوسط الحسابى هو :

---

(١) الأرقام مأخوذة من تقرير اللجنة الأنثربولوجية فى الجمعية البريطانية سنة ١٨٨٣ ، انظر كتاب G.U. Yule, *Introduction to the Theory of Statistics*, (1937), p. 111.

جدول ٥٥ — توزيع تكرارى لأوزان ١٢١٢ رجلاً أسكتلندياً

فئات الأوزان بالأرطال	التكرار	فئات الأوازن بالأرطال	التكرار
١٠٠ وأقل من ١١٠	١	١٨٠ وأقل من ١٩٠	١٢٥
١١٠—	٨	١٩٠—	٦٧
١٢٠—	٢٢	٢٠٠—	٢٤
١٣٠—	٦٣	٢١٠—	١٤
١٤٠—	١٧٣	٢٢٠—	٧
١٥٠—	٢٥٥	٢٣٤—	٤
١٦٠—	٢٧٥	٢٤٠—	٢
١٧٠—	١٦٨	٢٥٠—	٤

$$ع_و = \frac{٢٠,٣٢٥}{\sqrt{١٢١٢}} = ٠,٥٨٦ \text{ رطلا .}$$

وذلك باعتبار أن الانحراف المعيارى للأوزان فى هذه العينـــة التى بأيدينا ، وهو ٢٠,٣٢٥ رطلاً ، يساوى تقريباً الانحراف المعيارى لأوزان جميع الرجال الأسكتلنديين .

∴ الخطأ المعيارى لهذا الوسط الحسابى = ٠,٥٨٦ ؛ والخطأ المتعادل له يساوى $\frac{٢}{٣}$ هذا المقـدار أى ٠,٣٧٨ ؛ و بناء على ذلك نقول إن الوسط الحسابى للأوزان عند الرجال الأسكتلنديين على العموم هو :

$$١٦٤,٨٦٠ \pm ٠,٣٧٨$$

بمعنى أن احتمال كونه خارج هذه المنطقة يساوى احتمال كونه داخلها . أو نقول إن الوسط الحسابى العمومى للأوزان لا يختلف عن ١٦٤,٨٦٠ رطلاً بأكثر من ١,٧٥٨ رطلا إلا نادراً جداً (بنسبة ٣ فى الألف) ، لأن :

$$٣\ ع_{و} = ٣ \times ٠{,}٥٨٦$$

$$= ١{,}٧٥٨ \quad \text{رطلاً ،}$$

ونعلم أن ٩٩,٧ ٪ من العينات لها أوساط حسابية لا تختلف عن الوسط الحسابي الحقيقي لأوزان الرجال عموماً بأكثر من ٣ $ع_{و}$ وهذا = ١,٧٥٨ رطلاً .

الوسط الحسابي لا يكفي وحده بدون ذكر الخطأ

٤٠٨ — فى كل المسائل التى يطلب فيها إيجاد الوسط الحسابى لمجموعة ما ، نوجد الوسط الحسابى بالطرق المعتادة ، ويجب أن نحسب أيضاً الخطأ المعيارى (أو الخطأ المتعادل إذا فضلنا) لهذا الوسط الحسابى حتى يكون عندنا فكرة عن مقدار دقة هذه النتيجة ، ومقدار الخطأ الذى يصح أن نقع فيه إذا أردنا تطبيق هذه النتيجة على مجموعات أكبر وأعم من المجموعة أو العينة المستخدمة فى إيجاد الوسط الحسابى .

وعلى العموم يحسن أن نشفع الوسط الحسابى الذى نستخرجه لأى مجموعة أو عينة نبحثها ، بذكر الخطأ المتعادل أو الخطأ المعيارى ، وبدون ذكر أحد هذه المقاييس يكون بحثنا ناقصاً . وفى العادة يذكر الوسط الحسابى وبجانبه الخطأ المتعادل وبينهما علامة $\pm$ ، كما نرى فى الوضع المبين فى المسألة السابقة . والأفضل أن نستعمل ٣ $ع_{و}$ بدلا من الخطأ المتعادل .

الاختلاف بين متوسطى عينتين

٤٠٩ — متى يكون الفرق بين متوسطى عينتين راجعاً إلى المصادفة ؟ ومتى يكون هذا الفرق دالاً على وجود اختلاف حقيقى أو معنوى بين المجموعتين ؟ هذه مسألة ذات أهمية كبيرة فى الأبحاث التجريبية ، ويتوقف على الفصل فيها نتائج هامة .

فمثلاً نرى أن فرقة من التلاميذ متوسط درجاتهم فى امتحان معين يساوى ٢٢ درجة ، ونرى فرقة أخرى متوسط درجاتهم فى نفس الامتحان ٢٤ درجة . فهل هذا فرق يستلفت النظر أو هو راجع إلى المصادفة ؟ هل هو فرق حقيقى يدل على انحطاط مستوى تلاميذ الفرقة الأولى بالنسبة إلى الثانية بسبب إهمال المدرس أو التلاميذ أو . . . ، أو هو فرق ظاهرى أتت به المصادفة هكذا فى هذه المرة وربما ينعكس الوضع فى امتحان آخر ؟

ونرى مثلا أن متوسط الأجور فى مصنع يساوى ١٢ قرشاً فى اليوم ، وفى مصنع آخر مشابه له ، وفى نفس المدينة نجد المتوسط يساوى ١٠ قروش فقط . فهل هذا الفرق حقيقى أو ظاهرى ؟

وهناك مجموعة من الأطفال يتبعون نظاماً فى التغذية ومتوسط وزن الواحد ٣٢ كيلو جراماً ، ومجموعة أخرى تتبع نظاماً آخر متوسط الوزن فيهم ٣٥ كيلو جراماً . فهل هذا الفرق راجع إلى المصادفة فلا يلتفت إليه ؛ أو هو فرق حقيقى فيدل على أفضلية النظام الثانى للتغذية على الأول ؟

وهناك جماعة من المرضى مثلا حقنوا بمصل معين وكانت نسبة الوفيات بينهم ١٧ ٪ ، ومجموعة أخرى من المرضى بنفس المرض ولم يحقنوا بهذا المصل ولا بغيره ، وكانت نسبة الوفيات بينهم ٢٣ ٪ مثلا . فهل الفرق بين النتيجتين راجع إلى المصادفة فلا نهتم له ، أو هو فرق حقيقى ومعنوى فيدل على نجاح المصل فى العلاج ؟ .

وهكذا يمكننا تعداد هذه الأمثلة فى نواح مختلفة من البحث العلمى . ولا شك أن شيئاً كثيراً يتوقف على حكمنا على أصل الفرق بين العينتين . ولو أخطأنا فى الحكم فلا بد أن يترتب على هذا الخطأ نتائج خطيرة . ومن ذلك يتبين لنا أنه من المهم أن يكون لدينا مقياس نقيس به مثل هذه الفروق بين العينات حتى نتأكد من كنهها ومعنويتها .

نبحث عن تشتت الفروق

٤١٠ — والوسيلة الوحيدة التى نتبعها لبحث هذه الفروق هى أن نبحث عن انحرافها المعيارى الذى يقيس درجة تشتتها أو اختلافها من تجربة إلى أخرى . وبواسطة هذا المقياس نحكم على أى فرق نشاهده فنعرف إذا كان صغيراً تسمح به المصادفات فى الأحوال العادية أو كبيراً لا تأتى به المصادفات إلا نادراً جداً .

نأخذ أزواجاً من العينات فنحصل على توزيع تكرارى للفروق

٤١١ — لنفرض أننا نأخذ أزواجاً من العينات من مجتمع كبير ونحسب لكل عينة الوسط الحسابى ، ثم نحسب الفرق بين الوسطين الحسابيين لكل زوج من العينات . وليكن هذا الفرق يساوى ف مثلا . فإذا تعددت هذه الأزواج من العينات نحصل على فروق مختلفة فى المقادير ، وهذه الفروق المختلفة تكوّن توزيعاً تكرريًا ، له وسط حسابى وانحراف معيارى ، كأى توزيع تكرارى عادى . وحينئذ نحصل على ما نريده لو أمكننا معرفة الانحراف المعيارى لهذه الفروق وليكن يساوى $ع_ف$ مثلا ؛ لأن هذا الانحراف المعيارى $ع_ف$ يقيس لنا درجة اختلاف هذه الفروق ف ، وهو المعيار الذى نحكم بواسطته على أى فرق معين لنعلم إذا ما كان ظاهريًا فنعزوه إلى المصادفات ، أو حقيقيًا فنهتم له .

الأوساط الحسابية فى أزواج العينات مستقلة عن بعضها . معامل الارتباط بينها يساوى صفراً

٤١٢ — ولما كانت هذه العينات التى نأخذها كلها عشوائية وغير متحيزة بأى شكل ما ، ففى أى زوج من العينات يكون الوسط الحسابى لإحداهما مستقلا عن الوسط الحسابى للأخرى ولا علاقة بينهما مطلقاً . فإذا رمزنا للأوساط الحسابية للعينات الأول من هذه الأزواج بالحروف $س_١$ ، $س_٢$ ، $س_٣$ ، . . . ، وللأوساط الحسابية للعينات الثانية بالحروف $ص_١$ ، $ص_٢$ ، $ص_٣$ ، . . . ، ينتج أن س ، ص كميتان متغيرتان مستقلتان عن بعضهما . أى أن معامل الارتباط بينهما يساوى صفراً .

حساب الانحراف المعيارى للفروق

٤١٣ — نرمز إلى الانحراف المعيـــارى للأوساط الحسابية $س_١$ ، $س_٢$ ، . . . بالرمز $ع_و$ ، وللانحراف المعيارى للأوساط الحسابية $ص_١$ ، $ص_٢$ ، . . . بالرمز $عَ_و$ ، وبالرمز $ع_ف$ إلى الانحراف المعيارى للفروق بين الأزواج المتناظرة من هذه الأوساط . وهذه الفروق [١] هى :

$$ف_١ = س_١ - ص_١ ، ف_٢ = س_٢ - ص_٢ ، . . .$$

الانحراف المعيارى لهذه الفروق وهو $ع_ف$ نستخرجه من المعادلة المعروفة (أنظر الحاشية فى صفحة ٢٢٧ بند ٢١٠) وهى :

$$ع_ف^٢ = ع_و^٢ + عَ_و^٢ - ٢ ر . ع_و . عَ_و ، \quad \text{وبما أن } ر = ٠$$

$$\therefore \quad = ع_و^٢ + عَ_و^٢ ، \quad \text{ومن معادلة (٤) بند ٤٠٤}$$

$$= \frac{ع_١^٢}{ن_١} + \frac{ع_٢^٢}{ن_٢}$$

حيث $ع_١$ و $ع_٢$ هما الانحرافان المعياريان للعينتين ، و $ن_١$ و $ن_٢$ عدد ما فيهما من المفردات .

وهكذا حصلنا على الانحراف المعيارى للفرق بين الوسطين الحسابيين لعينتين [٢] بدلالة الانحراف المعيارى للوسط الحسابى فى العينة الأولى والانحراف المعيارى للوسط الحسابى فى العينة الثانية . وقد عرفنا (بند ٤٠٤) طريقة إيجاد هذين الأخيرين بدلالة الانحرافين المعياريين لنفس العينتين وعدد المفردات فى كل منهما .

---

(١) يلاحظ أن الوسط الحسابى لهذه الفروق يساوى صفراً ، وأن توزيعها التكرارى معتدل ، لأن توزيع كل من س و ص على حدة معتدل .

(٢) إذا أردنا الخطأ المعيارى للفرق بين متوسط إحدى العينتين ومتوسط العينتين عند ضمهما ليكونا مجموعة واحدة معاً وليكن $عَ_ف^٢$ = $ع^٢ \times \frac{ن_٢}{ن_١ (ن_١ + ن_٢)}$

أنظر كتاب G. Yule, p. 388.

إذا زاد الفرق عن ثلاثة أمثال الانحراف المعيارى ، يعتبر معنوياً

٤١٤ — والمقياس التى نعتبره كحد فاصل بين الفروق الظاهرية التى ترجع إلى المصادفات ، والفروق المعنوية التى تنشأ عن عوامل أخرى غير عوامل المصادفة ، هو كالمعتاد ثلاثة أمثال الخطأ المعيارى ، أى ٣ $\text{ع}_{\text{ف}}$. فإذا حسبنا الفرق بين الوسطين الحسابيين لعينتين ، ثم حسبنا الانحراف (أو الخطأ) المعيارى لهذا الفرق ، فوجدنا الفرق أكبر من ثلاثة أمثال انحرافه المعيارى كان ذلك دليلا على أن الفرق خارج عن مجال المصادفة ، وأنه يدل على اختلاف حقيقى بين العينتين .

مثال لتطبيق هذه النتيجة

٤١٥ — الجدول الآتى يبين التوزيع التكرارى لأوزان ٧٣٨ رجلا من مواليد مقاطعة ويلز فى انجلترا . المطلوب إيجاد الوسط الحسابى لأوزانهم ، ومقارنتهم كمجموعة بالرجال الأسكتلنديين المذكورة أوزانهم فى بند ٤٠٧ جدول ٥٥ صفحة ٢١٤ .

الوسط الحسابى للأوزان فى هذه المجموعة يساوى ١٥٧٫٨٠٥ رطلا والانحراف المعيارى للأوزان يساوى ١٨٫٨١٦ رطلا .

∴ الانحراف المعيارى للوسط الحسابى فى هذه المجموعة هو $\text{ع}_{\text{و}}$

$$\text{ع}_{\text{و}} = \frac{١٨٫٨١٦}{\sqrt{٧٣٨}}$$

$$= ٫٦٩٢٦$$

وفى مجموعة الأسكتلنديين وجدنا أن الوسط الحسابى للأوزان يساوى ١٦٤٫٨٦ . والانحراف المعيارى لهذا الوسط الحسابى

$\text{ع}_{\text{و}} = ٫٥٨٦$ رطلا .

الفرق بين الوسطين الحسابيين يساوى

$$١٦٤{,}٨٦٠ - ١٥٧{,}٨٠٥ = ٧{,}٠٥٥$$ رطلا ؛

والانحراف المعيارى لهذا الفرق يساوى $ع_ى$ حيث

$$ع_ى^٢ = ع_و^٢ + ع_{و'}^٢$$

$$= ٨٢٣٦{,}$$

$$\therefore ع_ى = ٩٠٧{,}$$

جدول ٥٦ — أوزان ٧٣٨ رجلا من ويلز

فئات الوزن بالرطل	التكرار	فئات الوزن بالرطل	التكرار
١٠٠ وأقل من ١١٠	٢	٢٠٠ وأقل من ٢١٠	٧
١١٠—	١٠	٢١٠—	٨
١٢٠—	٢٣	٢٢٠—	١
١٣٠—	٦٨	٢٣٠—	٢
١٤٠—	١٥٣	٢٤٠—	٠
١٥٠—	١٧٨	٢٥٠—	١
١٦٠—	١٣٤	٢٦٠—	٠
١٧٠—	١٠٢	٢٧٠—	٠
١٨٠—	٣٤	٢٨٠—	١
١٩٠—	١٤		
		المجموع	٧٣٨

ومن ذلك ينتج أن الفرق بين متوسطى هاتين العينتين وهو ٧,٠٥٥ رطلا ، أكبر بكثير جداً من ثلاثة أمثال انحرافه المعيارى . وعلى ذلك لا يمكن أن يكون هذا الفرق ناشئاً من اختلاف المصادفة ، ولا بد أن يكون راجعاً إلى فروق حقيقية بين الجنسين الأسكتلندى (أثقل وزناً) والغالى (أخف وزناً) .

مقارنة نسبتى النجاح فى مدرستين

٤١٦ — فى سنة ١٩٣٦ تقدم من مدرسة الأورمان الابتدائية ٧٠ تلميذاً فى امتحان الشهادة الابتدائية ؛ وكانت نسبة النجاح ٨٤٫٣ ٪ ، وفى نفس السنة تقدم من مدرسة أسيوط الابتدائية ٧٢ تلميذاً ؛ وكانت نسبة النجاح ٨٣٫٣ ٪ ، فهل هذا الفرق بين نسبتى النجاح فى المدرستين ظاهرى أو معنوى ؟

الفرق بين النسبتين يساوى ٠٫٠١

والانحراف المعيارى لهذا الفرق هو $ع_{ف}$ مثلاً . وليكن الانحراف المعيارى لنسبتى النجاح فى مدرستى الأورمان وأسيوط هما $ع_{١}$ و $ع_{٢}$ على الترتيب . و بتطبيق قاعدة الانحراف المعيارى للنسبة (بند ٣٩٥) وهى أنه يساوى .

$$\sqrt{\frac{ح\ ل}{ن}}$$

حيث ح = احتمال النجاح ، و ل = احتمال الفشل ، و ن = عدد المحاولات

$$\therefore \quad ع_{١}^{٢} = \frac{١}{٧٠} \times \frac{٨٤٣}{١٠٠٠} \times \frac{١٥٧}{١٠٠٠} = ,٠٠١٨٩٠٧١$$

وكذلك

$$ع_{٢}^{٢} = \frac{١}{٧٢} \times \frac{٨٣٣}{١٠٠٠} \times \frac{١٦٧}{١٠٠٠} = ,٠٠١٩٣١٠٨$$

$$\therefore \quad ع_{ف}^{٢} = ع_{١}^{٢} + ع_{٢}^{٢} = ,٠٠٣٨٢١٧٩$$

$$\therefore \quad ع_{ف} = ,٠٦١٨$$

ومن الواضح أن الفرق بين نسبتى النجاح وهو ٠٫٠١ أقل من الانحراف المعيارى للفرق فضلاً عن ثلاثة أمثاله . و ينتج من ذلك أن هذا الفرق راجع إلى المصادفة ولا يدل على شىء .

مقارنة معدل المواليد فى بلدين

٤١٧ — فى سنة ١٩٣٥ كان معدل الوفيات ٢١٫٢ فى الألف من السكان فى محافظة دمياط و ٢٩٫٤ فى محافظة السويس . فإذا علم أن التعداد

التقديرى لسكان هاتين المحافظتين فى تلك السـنة هو ٣٩٠٠٠ و ٤٧٠٠٠ على الترتيب ، فهل من الجائز أن يكون الفرق بين معدلى الوفيات فى المحافظتين راجعاً إلى المصادفة .

الفرق بين المعدلين أو النسبتين يساوى

$$ف = ٠٢٩٤, — ٠٢١٢, = ٠٠٨٢,$$

والانحراف المعيـارى لهذا الفرق هو $ع_ف$. والانحرافان المعيـاريان لهذين المعدلين هما $ع_١$ و $ع_٢$ للسويس ودمياط على الترتيب .

$$ع_١^٢ = \frac{١}{٤٧٠٠٠} \times \frac{٢٩٤}{١٠٠٠٠} \times \frac{٩٧٠٦}{١٠٠٠٠} = ٦٠٧١ \ldots\ldots ٠,$$

$$ع_٢^٢ = \frac{١}{٣٩٠٠٠} \times \frac{١}{١٠٠٠٠} \times \frac{٩٧٨٨}{١٠٠٠٠} = ٥٣٢١ \ldots\ldots ٠,$$

$$ع_ف^٢ = ع_١^٢ + ع_٢^٢ \qquad = ١١٣٩٢ \ldots\ldots ٠,$$

$$ع_ف = ١٠٦ \;٠٠,$$

وواضح من هذه الأرقام أن الفرق ف أكبر بكثير من ثلاثة أمثال الانحراف المعيـارى له . وينتج من ذلك أن هذا الفرق لا يمكن أن يعزى إلى عوامل المصـــادفة ، بل هو يدل على اختلاف معنوى بين المحافظتين من الناحية الصحية .

المقاييس الإحصائية الأخرى

٤١٨ — عرفنا الخطأ المعيارى للوسط الحسابى للعينة ($ع_و = \frac{ع}{\sqrt{ن}}$ بند ٤٠٤) ؛ فأمكننا أن نهتدى بواسطته إلى معرفة الوسط الحسابى للمجتمع . وكذلك بالنسـبة إلى الانحراف المعيارى للعينة ، ومعامل الاختلاف ، ومقياس الالتواء ، ومعامل الارتباط ، وغيرها من المقاييس الإحصائية التى نعرفها — يمكننا

أخطاء معيارية لجميع المقاييس الإحصـائية المعروفة

أن نوجد الخطأ المعيارى لكل منها . وهذا نستخدمه فى معرفة خواص المجتمع الأصلى المأخوذة منه العينة ، وفى المقارنة بين العينات المختلفة ، كما فعلنا فى حالة الوسط الحسابى . ونكتفى هنا بإيراد القوانين الخاصة بهذا الأخطاء المعيارية ، بدون أن نتعرض إلى البحث الرياضى فى كيفية الحصول عليها [١] . لأن هذا يخرجنا عن مجال هذا الكتاب .

جدول ٥٧ — الأخطاء المعيارية لبعض المقاييس الإحصائية المهمة [٢]

بفرض ن = عدد المفردات فى العينة و ع. الانحراف المعيارى للمجتمع الأصلى

المقياس	الخطأ المعيارى	ملاحظات
١ – التكرار فى اى فئة ، ص	$\sqrt{ص\,(١ - \frac{ص}{ن})}$	
٢ – الوسط الحسابى	$\frac{ع.}{\sqrt{ن}}$	
٣ – الوسيط	$\frac{١{,}٢٥٣٣١}{\sqrt{ن}}$ ع.	سواء كان المجتمع الأصلى معتدلا أو لا بفرض أن المجتمع الأصلى **معتدل**
٤ – الربيع	$\frac{١{,}٣٦٢٦٣}{\sqrt{ن}}$ ع.	» » » »
٥ – الانحراف المعيارى : ع	$\frac{ع.}{\sqrt{٢ن}}$	» » » »
٦ – العزم الثانى حول الوسط • أو التباين $ع^٢$	$ع.^٢ \sqrt{\frac{٢}{ن}}$	» » » »

(١) لإثبات هذه القوانين أنظر كتاب
G.U. Yule: *Introduction to the Theory of Statistics*, 1946, pp.380-411

(٢) يجد القارىء الأخطاء المعيارية لبعض المقاييس الأخرى فى كتاب G. U. Yule: المذكور آنفاً .

أو فى كتاب D.C. Jones: *First Course in Statistics*, (1927) p. 163.

أو فى كتاب H. Rietz: *Handbook of Mathematical Statistics*, (1924) p. 77

المقياس	الخطأ المعيارى	ملاحظات
٧ — العزم الثالث حول الوسط : م٣	$\text{ع}^{٣}\sqrt{\frac{٦}{\text{ن}}}$	بفرض أن المجتمع الأصلى معتدل
٨ — العزم الرابع حول الوسط : م٤	$\text{ع}^{٤}\sqrt{\frac{٩٦}{\text{ن}}}$	» » » »
٩ — معامل الاختلاف : ف	$\frac{\text{ف}}{\sqrt{٢\text{ن}}}\left[١+\frac{٢\text{ف}^{٢}}{١٠٠٠٠}\right]^{\frac{١}{٢}}$	» » » »
١٠ — معامل الارتباط : ر	$\frac{١-\text{ر}^{٢}}{\sqrt{\text{ن}}}$	» » » »
١١ — معامل الانحدار : $\text{ر}\frac{\text{ع}}{\bar{\text{ع}}}$	$\frac{\text{ع}}{\bar{\text{ع}}}\cdot\sqrt{\frac{١-\text{ر}^{٢}}{\text{ن}}}$	» » » »
١٢ — س المحسوبة من معادلة خط الانحدار : $\text{س}-\bar{\text{س}}=\text{ر}\frac{\text{ع}}{\bar{\text{ع}}}(\text{ص}-\bar{\text{ص}})$	$\text{ع}\sqrt{١-\text{ر}^{٢}}$	» » » »
١٣ — ص المحسوبة من معادلة خط الانحدار : $\text{ص}-\bar{\text{ص}}=\text{ر}\frac{\bar{\text{ع}}}{\text{ع}}(\text{س}-\bar{\text{س}})$	$\bar{\text{ع}}\sqrt{١-\text{ر}^{٢}}$	» » » »
١٤ — نسبة الارتباط : ى	$\frac{١-\text{ى}^{٢}}{\sqrt{\text{ن}}}$	
١٥ — المسافة بين المنوال والوسط الحسابى	$\text{ع}\cdot\sqrt{\frac{٣}{٢\text{ن}}}$	توزيع غير متماثل
١٦ — مقياس الالتواء ى١	$\sqrt{\frac{٣}{٢\text{ن}}}$	» » »

كيفية اختبار المقاييس الإحصائية بواسطة الخطأ المعيارى

٤١٩ — وطريقة استخدام هذه الأخطاء المعيارية هى نفس الطريقة التى شرحناها فى حالة الوسط الحسابى والانحراف المعيارى . ويجب إذاً فى كل العينات أو المجموعات التى نبحثها إحصائياً ونستخرج منها أى واحد من هذه المقاييس الإحصائية ، أن نحسب لكل واحد منها الخطأ المعيارى له ونذكره بجانبه ،

لكى نعرف درجة دقتـه من جهة ، ومن جهة أخرى نهتدى إلى قيمـة تقريبية لهذا المقياس ، إذا حسب من المجتمع الأصلى الذى أخذنا منه العينة التى بحثت فعلاً .

وعند مقارنة أى مقياس إحصائى فى عينتين ، نحسب الانحراف المعيارى لهذا المقياس فى كل من العينتين بمساعدة الخطأ المعيارى له المذكور فى جدول ٥٧ ؛ ثم نوجد الفرق بين قيمتى هذا المقياس فى العينتين ، ونحسب الانحراف المعيارى لهذا الفرق باستخدام المعادلة المعروفة (أنظر بند ٤١٢) .

$$ع_{ف}^{٢} = ع_{١}^{٢} + ع_{٢}^{٢}$$

حيث $ع_{ف}$ = الانحراف المعيارى للفرق ،

و $ع_{١}$ = « « للمقياس فى العينة الأولى ،

و $ع_{٢}$ = « « « « « « الثانية ،

فإذا كان الفرق ف أكبر من ثلاثة أمثال انحرافه المعيارى $ع_{ف}$ ،كان ذلك دليلاً على وجود اختلاف حقيقى بين العينتين لا يمكن أن يعزى إلى المصادفة .

اختبار الفرق بين التشتت فى مجموعتين

٤٢٠ — وجدنا من التوزيع التكرارى لأوزان الرجال الأسكتلنديين (المذكور فى بند ٤٠٧ صحيفة ٤١١) أن الانحراف المعيارى للأوزان فى هذه المجموعة (وعددها ١٢١٢) يساوى ٢٠٫٣٢٥ رطلاً . فماذا يكون مقدار الانحراف المعيارى للأوزان فى الرجال الأسكتلنديين على العموم .

نعلم أن الخطأ المعيارى للانحراف المعيارى يساوى $\frac{ع_{.}}{\sqrt{٢ ن}}$

ولعدم معرفتنا $ع_{.}$ نضع بدلها ع المحسوبة من العينة نفسها .

∴ الخطأ المعيارى فى هذه الحالة يساوى

$$\frac{٢٠٫٣٢٥}{\sqrt{٢٤٢٤}} = ٠٫٤٣١ \text{ تقريباً ،}$$

∴ الانحراف المعيارى للأوزان على العموم لا يختلف كثيراً عن

$$٢٠{,}٣٢٥ \pm ٣ \times ٠{,}٤٣١$$

أى $$٢٠{,}٣٢٥ \pm ١{,}٢٩٣ \text{ رطلا}$$

بمعنى أن احتمال كونه فى هذه المنطقة أكبر من ٠,٩٩٧ (أنظر بند ٣٩٤).

مثال — وجدنا أن الانحراف المعيارى لمجموعة الرجال (٧٣٨) من ويلز يساوى ١٨,٨١٦ رطلاً . فهل هذا يدل على أن التفاوت فى الوزن عند الأسكتلنديين أكبر منه عند أهل ويلز ؟

الفرق بين الانحرافين المعياريين فى المجموعتين هو

$$ف = ٢٠{,}٣٢٥ - ١٨{,}٨١٦$$

$$= ١{,}٥٠٩$$

نقارن هذا الفرق بانحرافه المعيارى $ع_{ف}$ حيث

$$ع_{ف}^{٢} = \left[\frac{٢٠{,}٣٢٥}{\sqrt{٢ \times ١٢١٢}}\right]^{٢} + \left[\frac{١٨{,}٨١٦}{\sqrt{٢ \times ٧٣٨}}\right]^{٢}$$

$$= ٠{,}١٧٠٥ + ٠{,}٢٣٩٩$$

$$\therefore ع_{ف} = ٠{,}٦٤٠٦$$

$$و\ ٣ ع_{ف} = ١{,}٩٢١٨$$

∴ الفرق ف أقل من ثلاثة أمثال انحرافه المعيارى ، فمن الجائز أن يكون راجعاً إلى المصادفة . ولا يمكن إذاً أن نستدل منه على وجود فرق بين هاتين المجموعتين فى التشتت . ومما يعزز هذا الرأى أن المنطقتين

$$٢٠{,}٣٢٥ \pm ١{,}٢٩٣ \quad ، \quad ١٨{,}٨١٦ \pm ١{,}٤٧٠$$

بينهما جزء مشترك .

مثال على اختبار معامل الارتباط

٤٢١ — وجدنا فى مثال معامل الارتباط بين عمر الرجل وعدد ما عنده من الأطفال (أنظر صفحة ٢٤١) أن معامل الارتباط ر = ٠,٤٨٤ وكان عدد الرجال فى الجدول التكرارى يساوى ٢٠٠ ، فماذا ينتظر أن يكون مقدار ر على العموم ، باعتبار أن هذه المجموعة عينة عشوائية من الرجال المتزوجين ؟

الخطأ المعيارى لمعامل الارتباط ر هو $ع_ر$ حيث

$$ع_ر = \frac{١ - ر^٢}{\sqrt{ن}} = \frac{١ - (٠,٤٨٤)^٢}{\sqrt{٢٠٠}} = ٠,٠٥٤٣$$

$$٣ ع_ر = ٠,١٦٢٩$$

$$\therefore \quad ر = ٠,٤٨٤ \pm ٠,١٦٢٩$$

بمعنى أن احتمال وقوع ر فى هذه المنطقة يزيد عن ٠,٩٩٧

ويلاحظ هنا أن ر تساوى تسعة أمثال انحرافها المعيارى $ع_ر$ تقريباً ؟ وهى تدل على ارتباط شديد نوعاً . وبهذه المناسبة نذكر أن معامل الارتباط إذا نقص عن ثلاثة أمثال انحرافه المعيارى فلا يعتبر دليلا على وجود ارتباط بين المتغيرين ، فلا نهتم به . ولذلك يجب حساب الانحراف المعيارى $ع_ر$ فى كل مسألة لمعامل الارتباط ، حتى نتبين مقدار أهمية المعامل ر الذى نحصل عليه .

ومثل ذلك ينطبق على نسبة الارتباط ى ، حيث انحرافها المعيارى $ع_ى$ يشابه $ع_ر$ فى الصورة كما رأينا فى جدول ٥٧ .

اختبار النتائج التجريبية

اختبار نتائج التجارب ومعرفة تحيزها من عدمة

٤٢٢ — يمكننا تطبيق هذه القواعد فى اختبار بعض النتائج التى نحصل عليها بالتجربة . لنأخذ مثلا تجربة ولدون التى ذكرناها فى بند ٣٩٢ من الباب السابق ولننظر فيما تدل عليه النتيجة التى حصل عليها من رمى الزهر .

رأينا أن الوسط الحسابى لعدد الزهرات الناجحة يساوى ٦,١٣٩ كما هو محسوب من واقع التوزيع التكرارى المشاهد . وبحساب الانحراف المعيارى لهذا التوزيع التكرارى نجده يساوى ١,٧١٢.

ونعلم فى الوقت نفسه أن نظرية الاحتمالات تقضى بأن يكون الوسط الحسابى لعدد الزهرات الناجحة يساوى ٦ ، كما هو واضح من تماثل التوزيع التكرارى النظرى الذى نجده فى العمود الثانى من جدول ٥٤ . فعلى اعتبار أن التوزيع التكرارى المشاهد هو بصفة عينة متوسطها ٦,١٣٩ فهل يمكن أن يكون هذا الفرق ٦,١٣٩ — ٦ = ٠,١٣٩ راجعاً إلى عوامل المصادفة ، وإلا فما معنى وجود هذا الفرق ؟

الخطأ المعيارى للوسط الحسابى فى هذه الحالة يساوى

$$\frac{ع}{\sqrt{ن}} = \frac{١,٧١٢}{\sqrt{٤٠٩٦}}$$

$$= ٠٢٧,$$

ومن الواضح أن الفرق ١٣٩, أكبر من خمسة أمثال انحرافه المعيارى؛ ونحن لا نعزو الفرق إلى المصادفة إلا إذا كان أقل من ثلاثة أمثال الانحراف المعيارى . وعلى ذلك فلابد أن تكون هناك عوامل أخرى غير المصادفة سببت وجود هذا الفرق الكبير . وتفسير ذلك أن الزهرات المستعملة فى التجربة — أو على الأقل كيفية رميها — كانت متحيزة بشكل من الأشكال ، حتى إنها أدت إلى هذه النتيجة الغريبة .

٤٢٣ — هذا الحكم الذى وصلنا إليه بخصوص هذه التجربة وعدم مطابقة نتائجها المشاهدة لما تقضى به النظرية ، كان يمكننا الوصول إليه من طريق آخر ، لو أننا تأملنا فى الفروق بين التكرارات النظرية والتكرارات المشاهدة . انحراف نتائج التجربة عن النظرية المعروضة

فلو حسبنا الفرق بين التكرارين فى كل فئة وربعناه ، كما نفعل عادة فى الانحرافات لنتخلص من الإشارة الجبرية التى لا تهمنا فى هذا البحث ، ثم قسمنا هذا المربع على التكرار النظرى للفئة ، حصلنا بذلك على مقياس للخطأ النسبى فى التكرار. وهذا الخطأ يكون صغيراً على العموم كلما قربت النتيجة المشاهدة من مطابقة النظرية . فلو جمعنا هذه الأخطاء النسبية فى كل الفئات حصلنا على كمية جديدة يمكننا استخدامها فى تقدير درجة مطابقة النتائج التجريبية للفروض النظرية . وبديهى أنه كلما كبر مقدار هذه الكمية ، بعدت التجربة عن تحقيق النظرية ومطابقتها . وبالعكس كلما صغرت هذه الكمية كان ذلك دليلاً على مطابقة الفرض النظرى للواقع المشاهد .

مجموع الأخطاء النسبية x^2 أو كا٢

٤٢٤ — ولو أجرينا هذا بالنسبة إلى هذه التجربة فربعنا الفروق بين التكرارات المشاهدة والنظرية (المذكورة فى جدول ٥٣) ثم قسمنا كلا من هذه المربعات على التكرار النظرى للفئة ، حصلنا [١] بعد الجمع على الكمية ٣٣٫٨١٠٤ ؛ ولنرمز إلى هذه الكمية بالحرف كا٢ بدلاً من الحرف الأغريقى x^2 (كاى تربيع) وهو الرمز المستعمل فى الكتب الأفرنجية .

ولكن معرفة القيمة المطلقة للكمية كا٢ أو x^2 فى أى مسألة معينة لا تكفى لمعرفة ما نريد ، ألا وهو درجة التطابق بين نتائج التجربة ونتائج النظرية . ويلزم لإتمام البحث معرفة التوزيع التكرارى لقيم الكمية كا٢ أو x^2 .

٤٢٥ — وقد اهتدى كارل بيرسون [٢] إلى معرفة قانون التوزيع التكرارى

---

(١) أنظر كتاب G.U. Yule, p 424.

(٢) أنظر K. Pearson: *Philosophical Magazine*, Vol. 50, Series 5,1900, p. 157.

ولأجل الجداول أنظر أيضاً K. Pearson: *Tables for Statisticians and Biometricians*, *part* I, pp. 26 - 28

والجدول الموجود فى آخر الكتاب

لقيم الـكمية $كا^٢$ أو $x^٢$ ، واستخدام هذا التوزيع فى حساب الاحتمال الخاص بكل قيمة لها ، وأنشأ جدولاً يعطى الاحتمال المناظر لكل قيمة ، حسب ظروفها . وبذلك يكون لدينا مقياس لقياس درجة مطابقة النتائج العملية للنظريات المؤسسة عليها هذه التجارب . وهذا يسمى [١] **اختبار حسن المطابقة** .

التوزيع التكرارى لقيم $كا^٢$

٤٢٦ — ولكى نشرح معنى هذا الاختبار نفرض أن جدولاً به فئات عددها ن أو خانات كما فى الجداول التكرارية العادية أو المزدوجة مثل جدول الارتباط والاقتران وغيرها ونفرض أن التكرارات النظرية فى هذه الخانات هى :

$$ص_١ ، ص_٢ ، ص_٣ ، . . . ص_ر . .$$

وأن التكرارات المشاهدة فى هذه الخانات هى على التوالى :

$$ك_١ ، ك_٢ ، ك_٣ ، . . . . ، ك_ر . . .$$

$$\therefore كا^٢ = مجـ \frac{(ك_ر - ص_ر)^٢}{ص_ر} = مجـ \frac{ف_ر^٢}{ص_ر}$$

حيث $ف_ر$ = الفرق بين التكرارين المشاهد والنظرى فى الفئة (أو الخانة) الرائية . وقد أثبت بيرسون أن التوزيع التكرارى لقيم $كا^٢$ تمثله معادلة من الصورة :

$$ا \,.\, كا^م \,.\, هـ^{-\frac{١}{٢} كا^٢}$$

حيث ا كمية ثابتة و م كمية ثابتة أيضاً تتوقف على عدد الخانات ن .

احتمالات الكمية $كا^٢$

٤٢٧ — ومن هذه المعادلة حسب بيرسون لكل قيمة للكمية $كا^٢$ الاحتمال ح طبقاً للتعريف الآتى :

---

(١) بالإنجليزية Test for Goodness of Fit.

ع = احتمال أننا نحصل من عينة عشوائية على قيمة للكمية كا^{٢} تساوى القيمة التى حصلنا عليها فعلا أو أكبر منها .

ففى المسألة التى ذكرناها فى بند ٤٢٤ مثلا ، وجدنا أن كا^{٢} = ٣٣٫٨١٠٤ . وبالكشف فى الجداول نجد أن ع = ٠٫٠٠١٨ أى أنه من النادر جداً أن نحصل على مثل هذا الانحراف (كا^{٢} = ٣٣٫٨١٠٤) أو أكبر منه ، على فرض عدم وجود تحيز . أو بعبارة أخرى أن فرض عدم التحيز لا يبرره الواقع .

وعلى العموم إذا كانت ع كبيرة فإنها تدل على مطابقة جيدة بين المشاهد والنظرية المفروضة ، وبالتالى تدل مبدئياً على وجاهة الفرض وعدم وجود دليل يناقضه . وإذا كانت ع صغيرة فإنها تدل على عدم مطابقة التجربة للنظرية المفروضة وهذا يؤوَّل بأحد أمرين : إما أن الفرض صحيح ولكن التجربة متحيزة وغير عشوائية ، وإما أن التجربة صحيحة وعشوائية كما يجب ولكن الفرض خاطىء لا يبرره الواقع المشاهد .

وعلى كل حال [١] يجب ألا نستدل من كبر قيمة ع على أن النظرية التى افترضناها صحيحة قطعا ، ولكن ذلك يدل على أنه لا يوجد فى البيانات التى بأيدينا ما يناقض صحة هذا الفرض .

---

(١) أنظر كتاب G.U. Yule & M.G. Kendall: *An Introduction to the Theory of Statistics,* (1946) p. 423, 22.19.

المـــراجع

Jones, D.C. : *First Course in Statistics*, Chapters XII—XIV

Pearson, K. : *Tables for Statisticians and Biometricians*

Rietz H. : *Handbook of Mathematical Statistics*, Chapter V

Tippett : *The Methods of Statistics*, Chapter III

Yule, G.U. : *Introduction to the Theory of Statistics*, Chapters 20, 21.

الباب السادس عشر

العينات الصغيرة

٤٢٨ — تكلمنا فى الباب السابق عن العينات بصفة عامة ، وعرفنا أن الخطأ المعيارى لأى مقياس إحصائى نحسبه من العينة يتوقف على حجم العينة ، فيزيد كلما كانت ن صغيرة وينقص كلما كبرت . ورأينا أيضاً أن الخطأ المعيارى للوسط الحسابى والانحراف المعيارى وغيرهما من المقاييس الإحصائية المهمة ، يتوقف على ع. وهو الانحراف المعيارى للمجتمع الأصلى المأخوذة منه العينة . ونظراً لأننا لا نعرف مقداره فى أغلبية المسائل التى نعالجها فنحن مضطرون إلى الاستعاضة عنه بالانحراف المعيارى ع المحسوب من العينة نفسها ، على اعتبار أن ع = ع. تقريباً .

ولكن هذا الاعتبار يكون خاطئاً بدرجة كبيرة إذا كانت ن صغيرة ، حيث قد رأينا أن الخطأ المعيارى للانحراف المعيارى يساوى $\frac{\text{ع.}}{\sqrt{\text{٢ن}}}$.

فلا بد إذن أن نبحث عن تقدير آخر للانحراف المعيارى ع. يكون أدق من ع لنستخدمه فى حالة العينات الصغيرة .

ع. أكبر من ع فى المتوسط

٤٢٩ — لنفرض أن لدينا عدداً كبيراً من العينات حجم كل منها ن ومتوسطاتها م$_{١}$ ، م$_{٢}$ ، م$_{٣}$ ، . . . وانحرافاتها المعيارية ع$_{١}$ ، ع$_{٢}$ ، ع$_{٣}$ ، . . . وليكن المتوسط العام للمجتمع المكون من جميع هذه العينات فى مجموعة

واحدة كبيرة ، هو $\bar{\text{س}}$ ، وليكن الانحراف المعيارى لهذا المجتمع هو $\text{ع}_{.}$. والمطلوب الآن معرفة العلاقة بين هذا الانحراف المعيارى للمجتمع والانحرافات المعيارية للعينات المأخوذة منه .

و بتطبيق البرهان المستخدم فى بند ١٨٢ مع وضع $\text{ع}_{.}$ بدل $\text{ع}_{\text{س}}$ هناك ، ووضع $\text{ن}_{١} = \text{ن}_{٢} = \text{ن}_{٣} = \ldots = \text{ن}$ ، حيث العينات كلها متساوية فى عدد المفردات ، ينتج أن

$$\text{ع}_{.}^{٢} = \frac{\text{مج ن ع}^{٢}}{\text{مج ن}} + \text{ع}_{\text{م}}^{٢} \quad \ldots \ldots \ldots \ldots \quad (١) \text{ ،}$$

حيث $\text{ع}_{\text{م}}$ هى الانحراف المعيارى للمتوسطات الحسابية للعينات ، وقد عرفنا أنه يساوى $\text{ع}_{\text{م}} = \frac{\text{ع}_{.}}{\sqrt{\text{ن}}}$

$$\therefore \quad \text{ع}_{.}^{٢} = \frac{\text{مج ن ع}^{٢}}{\text{مج ن}} + \frac{١}{\text{ن}} \text{ع}_{.}^{٢}$$

$$\therefore \quad \frac{\text{مج ن ع}^{٢}}{\text{مج ن}} = \frac{\text{ن} - ١}{\text{ن}} \text{ع}_{.}^{٢} \quad \ldots \ldots \ldots \ldots \ldots \quad (٢) \text{ .}$$

ولكن الطرف الأيمن من هذه المعادلة الأخيرة عبارة عن متوسط الكمية $\text{ع}^{٢}$ فى جميع العينات . أى أن $\text{ع}^{٢}$ المحسوبة من العينة تساوى فى المتوسط المقدار $\frac{\text{ن} - ١}{\text{ن}} \text{ع}_{.}^{٢}$.

و بناء على ذلك نقول إن

$$\text{ع}_{.}^{٢} = \frac{\text{ن}}{\text{ن} - ١} \text{ع}^{٢} \quad \ldots \ldots \ldots \ldots \ldots \quad (٣) \text{ ،}$$

فى المتوسط ، وكذلك

$$\text{ع}_{.} = \text{ع} \sqrt{\frac{\text{ن}}{\text{ن} - ١}} \quad \ldots \ldots \ldots \ldots \ldots \quad (٤) \text{ ،}$$

فى المتوسـط . وهكذا يثبت لنا أن $\bar{ع}$ فى المتوسط لا تساوى ع فقط ، بل ع × كمية أكبر من الواحد الصحيح وهى $\sqrt{\frac{ن}{ن-١}}$ ؛ أى أن الانحراف المعيارى للعينة وهو ع يكون فى المتوسـط أصغر من الانحراف المعيارى ع. المجتمع الأصلى المأخوذة منه العينة .

تقدير ع. من العينـة

٤٣٠ — ويتضح من ذلك أن الأفضل لنا أن نأخذ الكمية ع $\sqrt{\frac{ن}{ن-١}}$ بدلا من ع كتقدير دقيق للانحراف المعيارى الأصلى ع. .

وعلى كل حال فمن الواضح أن الفرق بينهما يتضاءل كلما كبرت العينة ، حيث تكون ن كبيرة والكمية $\sqrt{\frac{ن}{ن-١}}$ قريبة جداً من الواحد الصحيح .

وإذا كانت مفردات العينة هى

$$س_١ ، س_٢ ، س_٣ ، . . . س_ن ،$$

ومتوسطها الحسابى $\bar{س} = \frac{١}{ن} . مج س$ ، فإن انحرافها المعيارى هو ع ، حيث

$$ع^٢ = \frac{١}{ن} مج (س - \bar{س})^٢$$

حسب تعريف الانحراف المعيارى لأى مجموعة من المفردات .

ويمكننا إذن تقدير قيمة الانحراف المعيـــارى للمجتمع الأصلى ع. من المعادلة :

$$ى^٢ = \frac{ن}{ن-١} ع^٢ ، \text{حيث ى هى تقدير ع.}$$

$$= \frac{١}{ن-١} . مج (س - \bar{س})^٢ ،$$

أى أننا نقسم مجموع مربعات انحرافات مفردات العينـة عن وسطها الحسابى على (ن — ١) بدلا من ن .

هذه الكمية ى^{٢} نسميها[1] التباين التقديرى أو تقـدير التباين المحسوب من مفردات العينة .

درجات الحرية

٤٣١ — نبحث الآن فى السبب الذى جعل تقدير ع. أو مربعها ع^{٢} المحسوب على أساس $\frac{١}{ن - ١}$. مج (س — $\overline{س}$)^{٢} ، أدق وأفضل من تقديرها المحسوب من $\frac{١}{ن}$ مج (س — $\overline{س}$)^{٢} .

نعلم أن الأصل فى حساب ع.^{٢} للمجتمع الأصلى هو أننا نجمع مربعات انحرافات المفردات عن الوسط الحسابى للمجتمع وهو $\overline{س}$. ، بفرض أن هذا معروف لدينا من قبل ، ثم نقسم حاصل الجمع على عدد المفردات لنحصل على متوسـط مربع الانحراف للمفردة الواحـدة ، وذلك على اعتبار أن كل مفردة مستقلة تماماً عن باقى المفردات ، وبالتالى يكون انحرافها عن $\overline{س}$. مستقلا تماماً عن انحرافات المفردات الأخرى .

ولكننا فى حالة عينة عادية مفرداتها

$$س_{١} ، س_{٢} ، س_{٣} ، . . . ، س_{ن-١} ، س_{ن}$$

ووسـطها الحسـابى $\overline{س}$ حيث ن $\overline{س}$ = مج س ، نحسب انحرافات المفردات عن الوسط الحسـابى للعينة نفسها ؛ وفى هذه الحالة لا يكون انحراف

---

(١) بالإنجليزية Estimated Variance

أو Variance Estimate calculated from the sample data

أى مفردة مستقلاً عن انحرافات المفردات الأخرى ، لأن هذه الانحرافات مرتبطة ببعضها بالعلاقة المعروفة مح (س — $\overline{س}$) = ٠ ، أو :

$$(س_١ - \overline{س}) + (س_٢ - \overline{س}) + \cdots + (س_{ن-١} - \overline{س}) + (س_ن - \overline{س}) = ٠$$

وهى ناتجة من تعريف الوسط الحسابى للعينة : مح س = ن $\overline{س}$. وهذه العلاقة اللازم توافرها بين الانحرافات عبارة عن قيد أو شرط يقيد هذه الانحرافات فى تغيرها ، فلا يتغير أحدها إلا على حساب الباقى .

و إذا تأملنا فى هذه العلاقة نجد أنها تسمح لنا بحرية اختيار مقادير ن — ١ من هذه الانحرافات فقط ، و بتحديد مقادير هذه يتحدد مقدار الانحراف الأخير تحديداً كاملاً بموجب هذه العلاقة . هذا العدد (ن — ١) أو د نسميه[(١)] **عدد درجات الحرية** المتروكة لنا فى تحديد مقادير الانحرافات اختيارياً ، وهو فى الحقيقة عدد المصادر المختلفة المستقلة التى ينشأ عنها تباين المفردات و يتكون منها المجموع مح (س — $\overline{س}$)^٢ . فنحن نقسم هذا المقدار الأخير على عدد درجات الحرية التى أحدثته لنحصل على تقدير دقيق للتباين بين مفردات المجتمع المأخوذة منه العينة .

حساب عدد درجات الحرية

٤٣٢ — وعدد درجات الحرية يساوى دائماً عدد الأشياء التى نبحث فى تغيرها ناقصاً عدد القيود التى تربط هذه الأشياء بعضها ببعض . ففى الحالة التى ذكرناها فى البند السابق كنا نبحث فى متوسط مربعات انحرافات المفردات عن وسطها الحسابى ؛ وكان عدد الانحرافات ن و بينها قيد واحد يربطها ، ألا وهو أن مجموعها يساوى صفراً (نشأ عن كونها مقيسة من الوسط الحسابى

---

(١) بالإنجليزية Degrees of Freedom ؛ والقيد يسمى Constraint وسنرمز لعدد درجات الحرية بالرمز د تمييزاً له عن حجم العينة ن .

المستخرج من المفردات نفسها) ، ولذلك وجدنا عدد درجات الحرية يساوى

ن — ١ = ٮ

في جدول تكراري

وفى حالة توزيع عدد معين ن من المفردات على فئات جدول تكرارى عددها ل مثلاً ، يمكننا تعيين مقادير تكرارات ل — ١ فئة من هذه اختياريا ، وبذلك يتحدد تكرار الفئة الباقية ، لأن مجموع التكرارات ثابت وهو ن . وعلى ذلك يكون عدد درجات الحرية لتكرارات الفئات هو ل — ١ ، ويلاحظ أن هناك قيد واحد فقط يربط هذه التكرارات ألا وهو أن مجموعها ثابت يساوى ن .

في جدول الاقتران وجدول التوافق

وكذلك فى حالة توزيع مفردات فى جدول الاقتران أو جدول التوافق حيث تكون المجاميع الأفقية (للسطور) ثابتة وكذلك المجاميع الرأسية للأعمدة ، وتتغير تكرارات الخانات فى حدود هذه الشروط . فإذا كان عدد الأعمدة ل وعدد السطور م مثلاً يكون عدد الخانات كلها ل × م ؛ ولكن عدد درجات الحرية لتكرارات هذه الخانات يساوى (ل — ١) × (م — ١) ، لأن عدد الخانات التى يمكن ملؤها اختياريا فى أى سطر هو ل — ١ ، لكى يكون مجموع تكرارات السطر مساويا للمجموع الموجود فى الهامش ؛ وكذلك عدد الخانات الممكن ملؤها اختياريا فى أى عمود هو م — ١ . وعلى ذلك يكون عدد الخانات الممكن ملؤها اختياريا فى الجدول كله يساوى (ل — ١) . (م — ١) . وبعبارة أخرى نقول إن عدد القيود المختلفة التى تفرضها المجاميع الهامشية على تكرارات الفئات ، يساوى ل + م — ١ . وعلى ذلك يكون عدد درجات الحرية لتكرارات هذه الخانات يساوى

ل م — (ل + م — ١) = (ل — ١) × (م — ١) .

ففى حالة جدول الاقتران المكون من ٤ خانات ٢ × ٢ يكون هناك درجة حرية واحدة فقط . وهذا واضح حيث إنه يكفى تحديد تكرار خانة واحدة فقط و به تتحدد تكرارات الخانات الثلاث الباقيـة فى الجدول ، باسـتخدام المجاميع التى فى الهامشين الجانبى والسفلى للجدول .

توزيـع ت لواضـعه Student

٤٣٣ — نبحث الآن فى طريقة لاستخدام متوسط العينة للاستدلال على الوسط الحسـابى للمجتمع ، ونقدر مدى الخطأ الذى نتعرض له فى هذا الاستدلال . لنفرض أن مفردات العينة هى على العموم :

$$\text{س}_{١} \text{ ، } \text{س}_{٢} \text{ ، . . . ، } \text{س}_{\text{ن}} \text{ ،}$$

وأن متوسطها الحسابى $\bar{\text{س}} = \frac{١}{\text{ن}}$ مج س كالمعتاد . ولتكن $\bar{\text{س}}_{.}$ و $\text{ع}_{.}$ هما الوسط الحسـابى والانحراف المعيـارى للمجتمع الأصلى المـأخوذة منه العينة ، وكلاهما مجهول .

نحسب من العينة الـكمية ى من واقع المعادلة المذكورة فى بند ٤٣٠ وهى :

$$\text{ى}^{٢} = \frac{١}{\text{ن} - ١} \text{ مج } (\text{س} - \bar{\text{س}})^{٢} \text{ ،}$$

وهى كما قلنا من قبل تعتبر أحسن تقـدير لدينا للانحراف الأصلى $\text{ع}_{.}$. و بنـاء على ذلك يكون معيار الخطأ للوسط الحسابى هو كما قلنا

$$\text{ع}_{\text{و}} = \frac{١}{\sqrt{\text{ن}}} \text{ ع}_{.} = \frac{١}{\sqrt{\text{ه} + ١}} \text{ ع}_{.} \qquad (١) \text{ ،}$$

$$= \frac{١}{\sqrt{\text{ن}}} \text{ ى} = \frac{١}{\sqrt{\text{ه} + ١}} \text{ ى تقريباً} \qquad (٢) \text{ .}$$

نأخذ الـكمية الآتية [١]

---

(١) هذه الـكمية هى المعروفة فى الـكتب الأفرنجية بالرمز t . وأول من أدخلها وبحث فى خواصها هو *Student* فى سنة ١٩٠٨ فى مجلة *Biometrika* .

$$ت = \frac{\bar{س} - \bar{س}_{.}}{ى}\sqrt{ن + ١} \qquad (٣) .$$

وهى كما نرى تساوى الفرق بين وسط العينة والمتوسط العمومى للمجتمع الأصلى ، مقسوماً على الانحراف المعيارى (التقديرى) أو الخطأ المعيارى (التقديرى) للوسط الحسابى للعينة . أى أنها تساوى الفرق بين متوسط العينة والمتوسط الأصلى ، مقيساً بوحدات تساوى الخطأ المعيارى التقديرى (المحسوب من العينة) لمتوسط العينة .

هذه الكمية ت تتوزع توزيعاً تكرارياً متماثلا يقرب من الاعتدال كلما كبرت ه ، كما يظهر من المعادلة الآتية التى تبين التوزيع التكرارى لقيم ت فى العينات المختلفة (على فرض أن المجتمع الأصلى المأخوذ من العينات معتدل [1]) :

$$ص = ص_{.}\left[١ + \frac{ت^{٢}}{ن}\right]^{-\frac{ن+١}{٢}} \qquad (٤) .$$

حيث $ص_{.}$ كمية ثابتة مستقلة عن ت ، $ن + ١ = ه$ وهى عدد المفردات فى العينة أو حجمها . ومن الواضح أن هذا المنحنى متماثل حول ت = صفرا ، وأن $ص_{.}$ هى قيمة ص عندما ت = صفرا ، وهى أكبر قيمة تأخذها ص .

ومن الممكن أن نأخذ $ص_{.}$ بحيث تكون المساحة الكلية للمنحنى (٤) من ت = − ∞ إلى ت = + ∞ ، تساوى الواحد الصحيح ؛ وعند ذلك تكون مساحة أى جزء من المنحنى محدود بمستقيمين رأسيين يساوى بالضبط احتمال وجود قيمة ت بين الحدين الذين يحدان هذا الجزء

(١) وعلى كل حال فالمعروف أن التوزيع التكرارى للفرق $س - \bar{س}_{.}$ معتدل وله وسط حسابى يساوى صفراً .

من المنحنى ، شأنه فى ذلك شأن المنحنيات التكرارية العادية التى درسناها من قبل .

جداول احتمالات القيم ت

٤٣٤ — وقد أنشأ استيودنت مكتشف هذه المعادلة (٤) وغيره جداول تعطى قيم ت والاحتمالات المناظرة لكل منها مع قيم ىہ . وأحد هذه الجداول[١] يعطى الاحتمال ح١ المناظر للقيمة ت١ ، و يساوى احتمال حصولنا على مقدار يساوى ت١ أو أكبر منه للكمية

$$ت = \frac{\bar{س} - \bar{س}_.}{ى} \sqrt{ىہ + ١}$$

من أى عينة عشوائية . فإذا كان ح١ صغيراً جداً فهو يدل على أن هذه القيمة ت١ نادرة جداً ، ولا نحصل على مثلها أو أكبر منها فى أى عينة عشوائية إلا نادراً جداً. وهذا يدل على أن الفرق $\bar{س} - \bar{س}_.$ الذى نشأت منه ت١ فرق معنوى ولا يمكن أن نعزوه إلى مجرد الصدف . وهكذا يمكننا الاستدلال من متوسط العينة على متوسط المجتمع ، ونعرف احتمال صحة هذا الاستدلال .

و بعض الجداول[٢] يعطى الاحتمال ل١ = ١ — ح١ وهو احتمال حصولنا على قيمة أصغر من ت١ ، وهى الجداول التى أنشأها استيودنت نفسه فى سنة ١٩٠٨ وأتمها فى سنة ١٩٢٥ .

وقد أنشأ الأستاذ ر . ا . فيشر جدولا مختصراً[٣] يعطى قيم ت إذا علم

(١) أنظر مثلا صفحة ٤٨٨ من كتاب : Peters and Van Voorhis *Statistical Procedures and their Mathematical Bases* طبعة سنة ١٩٤٠

(٢) نشرت هذه الجداول فى مجلة Metron (روما — إيطاليا) سنة ١٩٢٥ (Vol. 5, part 3) ويجدها القارىء مختصرة إلى ثلاثة أرقام عشرية فى كتاب G. U. Yule: and M. G. Kendal: *An Introduction to the Theory of Statistics*, (1946) p. 536.

(٣) أنظر مثلا صفحة ١٦٦ من كتابه *Statistical Methods for Research Workers* أو صفحة ١٧٣ من كتاب Peters and Van Voorhis المذكور أعلاه ، وفى غيرهما .

كل من الاحتمال ح وعدد درجات الحرية ن . والاحتمال ح = ٢ ح وهو يشمل قيم ت ≥ ت١ و ≤ — ت١ .

٤٣٥ — نأخذ بعض الأمثلة لتوضيح طريقة استخدام ت وجداولها المختلفة.

مثال (١) — قيست أطوال عشرة أشخاص فكانت أطوالهم بالسنتيمترات هى :

١٦٩ ، ١٧١ ، ١٧٢ ، ١٧٣ ، ١٧٣ ، ١٧٥ ، ١٧٦ ، ١٧٦ ، ١٧٨ ، ١٨٠ .

فالمطلوب حساب ت لهذه المجموعة وإيجاد احتمال حصولنا على قيمة لها تساوى هذه أو أكبر منها ، على فرض أن المتوسط العام للأطوال $\bar{س}$ = ١٧٥ سنتيمتراً .

نحسب أولا الوسط الحسابى $\bar{س}$ لهذه المفردات العشر ؛ وإذا أخذنا ١٧٣ كوسط فرضى لتبسيط الحساب يكون

$$\bar{س} = ١٧٣ + \frac{١}{١٠} (-٤-٢-١+٠+٠+٢+٣+٣+٥+٧)$$

$$= ١٧٤{,}٣ \text{ سنتيمتراً .}$$

$$مح (س - \bar{س})^٢ = مح (س - ١٧٣)^٢ - ١٠ (١٧٤{,}٣ - ١٧٣)^٢ ،$$

$$= ١١٧ - ١٦{,}٩$$

$$= ١٠٠{,}١$$

ولكن $ى^٢ = \frac{١}{ن - ١} مح (س - \bar{س})^٢$ ، حيث ن = ١٠ ،

$$\therefore \quad ى = \frac{١}{٣} \sqrt{١٠٠{,}١} \text{ ،}$$

$$= ٣{,}٣٣٥ \text{ .}$$

وبالتعويض عن س، ن، ى فى المعادلة

$$ت = \frac{\bar{س} - \bar{س}.}{ى} \sqrt{ن + ١}\text{، حيث } ن = ﻫ - ١ = ٩\text{،}$$

$$\therefore \quad ت = \frac{١٧٤{,}٣ - ١٧٥}{٣{,}٣٣٥} \cdot \sqrt{١٠}$$

$=$ ٠٫٦٦٤ بصرف النظر عن الإشارة الجبرية .

وبالنظر إلى جدول [1] الاحتمالات ح لقيم ت ، نجد أنه عندما يكون عدد درجات الحرية ن = ٩، تكون

ح = ٫٢٨١٧ عندما ت = ٫٦ .

و ح = ٫٢٥٠٨ « ت = ٫٧ .

$\therefore$ عندما تكون ت = ٫٦٦٤ كما فى المسألة

$$\text{يكون} \quad ح = ٫٢٨١٧ - (٫٢٨١٧ - ٫٢٥٠٨) \times \frac{٫٦٦٤ - ٫٦}{٫٧ - ٫٦}$$

$$= ٫٢٦١٩$$

ويكون ٢ ح = ٫٥٢٣٨

أى أننا نحصل على القيمة ت = ٫٦٦٤ أو أكبر منها ٢٦ مرة فى كل ١٠٠ مرة ، وفى ٢٦ مرة أخرى نحصل على قيمة أصغر من — ٫٦٦٤ ؛ أى أننا نحصل فى ٥٢ ٪ من الحالات على قيم للكمية ت أكبر من ٫٦٦٤ عددياً . وهذا معناه أن الفرق بين متوسط العينة والمتوسط العمومى ليس معنوياً بل هو راجع إلى الصدف ، لأن المتفق عليه أن الفرق لا يعتبر معنوياً إلا إذا كانت ٢ ح أقل من ٫٠٥ أو ٫٠١ إذا أخذنا بالأحوط . فإذا قررنا إذن أن المتوسط العمومى للأطوال (فى المجتمع

---

(١) كتاب Peters and Van Voorhis صحيفة ٤٨٨ ، يلاحظ أنه يستعمل حرف n لعدد درجات الحرية بدل ن وحرف n' لعدد مفردات العينة بدل ﻫ عندنا . وبعض الكتب الأخرى تستعمل الحرف الإغريقى ν (نبو) لعدد درجات الحرية ن .

الأصلى المأخوذة منه هذه العينة التى بحثناها هنا) هو ١٧٥ سنتيمتراً ، فليس فى البيانات التى بحثناها ولا النتائج التى توصلنا إليها بواسطتها ما يشير إلى أن هذا الفرض خطأ أو بعيد عن الصواب .

وهكذا أمكننا أن نحدد مكان المتوسط العمومى بشىء من الدقة باستخدام البيانات المعلومة من العينة و بغير حاجة إلى معرفة ع أو أى مقياس آخر للمجتمع الأصلى .

مثال (٢) — زرع ا و ب قمحاً فى حقليهما مدة من السنين و بينما كان ا يزرع بدون سماد كان ب يضع سماداً فى حقله بمعدل جوال واحد (ثمنه جنيهان) لكل فدان . وكانت محصولاتهما من هذه الزراعة كما يلى :

المحصول بالأردب للفدان

	عند ا	عند ب
فى السنة الأولى	٤٫٨	٦٫٢
« الثانية	٤٫٦	٥٫٨
« الثالثة	٥٫٠	٦٫٧
« الرابعة.	٥٫٢	٦٫٩
« الخامسة	٤٫١	٥٫٠
« السادسة	٤٫٥	٥٫٤

فإذا كان متوسط سعر القمح فى هذه السنين هو ٣ جنيهات للأردب ، فهل هناك ربح مضمون بالتسميد ؟

العبرة هنا بالفرق بين المحصولين ، فإذا كان ثمن القمح الزائد فى المحصول

يساوى فى المتوسط قيمة السماد فليس هناك ربح . وقيمة السماد هنا ٢٠٠ قرشاً وهى بالصدفة تساوى ثمن ثلثى أردب من القمح . فإذا كان متوسط الفرق فى المحصول ثلثى أردب فى الفدان فقط فليس هناك ربح . ولا يكون هناك ربح مضمون إلا إذا كان الفرق فى المتوسط بين المحصولين أكبر من $\frac{٢}{٣}$ أردب .

نبحث إذن فى قيمة الفرق بين المحصولين فى المتوسط ونعتبر الفروق فى الست سنوات المذكورة كعينة من مجتمع كبير لهذه الفروق ، ونبحثها لنعرف المتوسط العمومى لهذا المجتمع ، فإذا كان المتوسط أكبر من $\frac{٢}{٣}$ كان الربح مضموناً .

هذه الفروق الستة هى بالأرادب :

١٫٤ ، ١٫٢ ، ١٫٧ ، ١٫٧ ، ٠٫٩ ، ٠٫٩

ومتوسطها الحسابى $\bar{س} = ١٫٣$ أردباً .

$$\therefore \quad \text{مج}\,(س - \bar{س})^{٢} = ٠٫٦٦$$

وعدد درجات الحرية $\quad ه = ن - ١ = ٥$.

$$\therefore \quad ى^{٢} = ٫١٣٢$$

$$\therefore \quad ى = ٫٣٦٣٣$$

لنفرض أنه ليس هناك ربح مضمون ، أى أن متوسط الفرق إذا أجرينا هذه التجربة عدداً كبيراً من السنين يساوى $\frac{٢}{٣}$. أى أننا نفرض مبدئياً أن $\bar{س}_{.} = \frac{٢}{٣}$ ، وننظر فى احتمال صحة هذا الفرض .

هنا $\quad ت = \dfrac{\bar{س} - \bar{س}_{.}}{ى}\sqrt{ه + ١}$

$$= \frac{١٫٣ - ٫٦٦٦٧}{٫٣٦٣٣}\sqrt{٦}$$

$$= ٤٫٢٨$$

هنا نستعمل مثلا جدول استيودنت نفسه الذى يعطى ل وهو احتمال حصولنا على ت أصغر من ٤٫٢٨ بنفس درجات الحرية وهى هنا ٥ ، ونجد أن

عندما ت = ٤٫٢ ، ل = ٫٩٩٦

و « ت = ٤٫٣ ، ل = ٫٩٩٦

∴ عندما ت = ٤٫٢٨

يكون ل = ٫٩٩٦

∴ ٢ (١ — ل) = ٫٠٠٨

أى أننا نحصل على قيمة أكبر من ٤٫٢٨ عدديًا (≤ ٤٫٢٨ أو ≥ — ٤٫٢٨) ٨ مرات فى كل ألف وهذا نادر جداً ، مما يدل على أن الفرض الذى افترضناه ، وهو أن متوسط الفرق = $\frac{٢}{٣}$ (أى أن الربح ليس مضمونًا) فرض لا يؤيده الواقع بل يناقضه . و بناء على ذلك نقول إن هناك ربح مضمون بالتسميد .

مثال (٣) — إذا كان سعر السماد فى المثال السابق ٣ جنيهات للجوال (بدلا من جنيهين) ، فهل يعود التسميد بربح للزارع ؟

هنا ثمن السماد يساوى ثمن أردب من القمح تماماً . فإذا كان الفرق فى المحصول يساوى أردباً لكل فدان فى المتوسط فلن يكون هناك ربح مضمون يعود به التسميد .

لنفرض أن متوسط الفرق فى المحصول هو أردب واحد ونبحث فيما إذا كان هذا الفرض يقرره الواقع المشاهد أو يناقضه ويدحضه كما وجدنا فى المثال السابق .

باستخدام نفس البيانات المذكورة ووضع س. = ١ بدل $\frac{٢}{٣}$

$$\therefore \quad ت = \frac{١٫٣ - ١}{٠٫٣٦٣٣}\sqrt{٦}$$

$$= ٢٫٠٢١$$

وإذا رجعنا مثلا إلى جدول ر. ا. فيشر عند ں = ٥ نجد أن

ح = ٢ ح = ٠٫١ عندما ت = ٢٫٠١٥

و ح = ٢ ح = ٠٫٠٥ « ت = ٢٫٥٧١

أى أن ح التى تناظر ت = ٢٫٠٢١ تقع بين ٠٫١ ، ٠٫٠٥ ، وهى تكاد تساوى ٠٫١ ؛ أى أننا نحصل على القيمة ت = ٢٫٠٢١ أو أكبر منها (عددياً) عشر مرات فى كل مائة تقريباً .

وهذا احتمال صغير نوعاً . ولكنه ليس صغيراً جداً لحد المعنوية إذا اعتبرنا حد المعنوية عند ح = ٠٫٠٥ كما هو المتفق عليه ، أى أن قيمة ت لا تعتبر معنوية إلا إذا كانت ح المناظرة لها تساوى ٠٫٠٥ على الأكثر . وعلى هذا الاعتبار تكون ت = ٢٫٠٢١ التى حصلنا عليها فى هذه المسألة ليست معنوية وليس هناك إذن فى البيانات التى بأيدينا عن المحصول بالتسميد وبدونه ما يناقض الفرض الأساسى الذى بدأنا به وهو أن مقدار الفرق فى المتوسط يساوى أردباً واحداً . ومعنى ذلك أن هناك شك فى أن يعود التسميد على الزارع بربح مضمون .

هذا ما أمكننا استنباطه فى حدود البيانات التى لدينا ، ولعلنا نستطيع إذا أضفنا إليها تجارب أخرى أن نصل إلى رأى نطمئن إليه .

مستوى المعنوية

٤٣٦ — رأينا فى الأمثلة السابقة أن حكمنا فى هـذه المسائل يعتمد أساساً على تحـديد مستوى معـين للاحتمال ح (أو ح أو ل) لا يزيد عنه . وهذا هو ما نسميه [١] **مستوى المعنوية** . وقد ذكرنا فى المثال الثالث أن المتفق عليه هو أن ح = ٠٫٠٥ وهذا يعادل ح = ٠٫٠٢٥ أو ل = ٠٫٩٧٥ حسب الجداول التى نستخدمها . و إذا أخذنا بالأحوط نعتبر مستوى المعنوية عند ح = ٠٫٠١ أى ح = ٠٫٠٠٥ ، ل = ٠٫٩٩٥ ونظراً لأن مستوى المعنوية هو أساس الحكم فى الموضوع فقد نظم فيشر جدوله [٢] على أساس عـدد درجات الحرية ىه (ىه = ١ ، ٢ ، ٣ ، . . . إلى ٣٠) وعـدد قليل من مستويات المعنوية (ح=٠٫١ ، ٠٫٢ ، ٠٫٠٥ ، ٫١ ، ٫٢ ، ٫٠٠ إلى ٠٫٩) . والجدول يعطى قيم ت التى تصل إلى أى واحد من مستويات المعنوية بعدد معين من درجات الحرية . و بهذه الطريقة أمكنه أن يختصر الجـدول اختصاراً كبيراً بدون أن يخل به .

مقارنة الوسط الحسابى فى عينتين

الفـرق بين الوسطين

٤٣٧ — يمكننا اسـتخدام الـكمية ت والجداول الخاصة بها لاختبار معنوية الفرق بين المتوسطين الحسابيين لعينتين عشوائيتين ، لنعرف ما إذا كان من الجائز اعتبارهما من مجتمع واحد أو لا .

---

(١) يسمى بالإنجليزية Significance Level

(٢) أنظر أيضاً جداوله الحديثة التى تحتوى على تفاصيل أ كثر من هذه لـكل من ىه و ح
Fisher and Yates: *Statistical Tables for Biological, Agricultural and Medical Research* (1943).
وأيضاً صفحة ٣٢٠ من كتاب Pearson and Bennet: *Statistical Methods* (1942)

لنفرض أن مفردات العينتين هى :

$\text{س}_1 ، \text{س}_2 ، \ldots ، \text{س}_{\text{ن}_1}$ للعينة الأولى ومتوسطها $\bar{\text{س}}_1$ ،

و $\text{س}'_1 ، \text{س}'_2 ، \ldots ، \text{س}'_{\text{ن}_2}$ للعينة الثانية ومتوسطها $\bar{\text{س}}_2$.

نحسب مجموع مربعات انحرافات المفردات كل واحدة عن الوسط الحسابى لمجموعتها، فنحصل على المجموع الآتى :

$$\text{مجـ}(\text{س} - \bar{\text{س}}_1)^2 + \text{مجـ}(\text{س}' - \bar{\text{س}}_2)^2$$

وعدد درجات الحرية لهذا المجموع يساوى

$$\text{ن}_1 + \text{ن}_2 - 2$$

لأن عدد درجات الحرية فى العينة الأولى $\text{ن}_1 - 1$ وفى الثانية $\text{ن}_2 - 1$ ، ومجموع مربعات انحرافات مفردات العينتين كل واحدة عن الوسط الحسابى لعينتها، يساوى مجموع العددين . وينتج من ذلك أن تقدير التباين للمجتمع المحسوب من المفردات الموجودة فى هاتين العينتين هو ى^2 حيث

$$\text{ى}^2 = \frac{1}{\text{ن}_1 + \text{ن}_2 - 2}\left[\text{مجـ}(\text{س} - \bar{\text{س}}_1)^2 + \text{مجـ}(\text{س}' - \bar{\text{س}}_2)^2\right].$$

وقد سبق أن قلنا (بند ٤١٣) إن مربع الانحراف المعيارى للفرق بين كميتين تتغيران مستقلتين عن بعضهما هو :

$$\text{ع}_{\text{ف}}^2 = \text{ع}_1^2 + \text{ع}_2^2$$

حيث ع_1^2 هو تباين الكمية الأولى و ع_2^2 تباين الكمية الثانية . فعند دراسة الفرق بين متوسطى عينتين عشوائيتين وهما $\bar{\text{س}}_1$ ، $\bar{\text{س}}_2$ تكون ع_1^2 هى مربع الانحراف المعيارى للوسط $\bar{\text{س}}_1$ وتساوى (١) $\frac{1}{\text{ن}_1}\text{ع}^2$ ، وتكون ع_2^2 مربع

---

(١) طبقاً للمعادلة $\text{ع}_{\text{و}}^2 = \frac{1}{\text{ن}}\text{ع}^2$ فى بند ٤٠٤ فى الباب السابق .

الانحراف المعيارى للوسط $\bar{س}_٢$ وتساوى $\frac{١}{ن_٢}ع_{.}^{٢}$ حيث $ع_{.}$ هى الانحراف المعيارى للمجتمع الأصلى المأخوذ منه العينتين ، ونستعيض عنه بالتقدير ى المحسوب من مفردات العينتين وهو أحسن تقدير عندنا .

$$\therefore \quad ع_{ف}^{٢} = ى \left(\frac{١}{ن_١} + \frac{١}{ن_٢} \right)$$

$$= \frac{ن_١ + ن_٢}{ن_١ ن_٢} ى^{٢} .$$

و إذا قسمنا الفرق $ف = \bar{س}_١ - \bar{س}_٢$ على انحرافه المعيارى $ع_{ف}$ ورمزنا لخارج القسمة بالرمز المعروف ت نجد أن

$$ت = \frac{\bar{س}_١ - \bar{س}_٢}{ى} \sqrt{\frac{ن_١ \; ن_٢}{ن_١ + ن_٢}} ،$$

و إذا تذكرنا (١) أن الوسط الحسابى للفرق ف فى جميع أزواج العينات العشوائية من هذا المجتمع الواحد ، يساوى صفراً نرى أن ت فى هذه المعادلة عبارة عن خارج قسمة الفرق بين قيمة مشاهدة لكمية معينة (وهى $ف = \bar{س}_١ - \bar{س}_٢$) والوسط الحسابى العمومى لها (وهو صفر هنا) ، على الانحراف المعيارى لهذه الكمية . أى أن ت فى هذه المعادلة هى نفس ت الواردة فى بند ٤٣٣ مع تغيير فى معنى الرموز فى الطرف الأيسر . و بناء على ذلك فإن التوزيع التكرارى للكمية

$$ت = \frac{\bar{س}_١ - \bar{س}_٢}{ى} \sqrt{\frac{ن_١ \cdot ن_٢}{ن_١ + ن_٢}}$$

هو نفس التوزيع المذكور فى بند ٤٣٣ ، ونستخدم نفس الجداول التى ذكرناها لاختبار معنوية قيمة ت التى نحصل عليها من أى تجربة ، مع ملاحظة أن عدد

---

(١) هذا الفرق يتوزع توزيعاً تكرارياً معتدلا حول الوسط الحسابى $\bar{ف}$ = صفراً .

درجات الحرية لقيمة ت عند دراسة الفرق بين متوسطى عينتين حجماهما $ن_1$ و $ن_2$ ، يساوى $ن_1 + ن_2 - 2$. فإذا وجدنا فى تجربة ما أن ت معنوية نستدل من ذلك على أن الفرق بين المتوسطين معنوى ؛ وعلى ذلك نقول إن العينتين ليستا من مجتمع واحد ، وبينهما اختلافاً حقيقياً لا يمكن تفسيره بإرجاعه إلى مجرد الصدفة .

٤٣٨ — نأخذ مثالين لتوضيح كيفية تطبيق هذه النتائج :

مثال (١) — كانت درجات مجموعتين ا و ب من الطلبة فى امتحان معين كالآتى حيث كانت ا مكونة من ٩ طلبة و ب من ٨ :

درجات المجموعة ا : ٦٧ ، ٧٧ ، ٦٨ ، ٧٥ ، ٧٧ ، ٧٩ ، ٧٧ ، ٧٣ ، ٦٧

» » ب : ٦٦ ، ٦٦ ، ٧٠ ، ٦٦ ، ٧٠ ، ٦٧ ، ٦٥ ، ٧١ .

فهل هناك فرق حقيقى بين مستوى إجابات الطلبة فى المجموعتين .

نلاحظ مبدئياً أن كل هذه الدرجات فوق ٦٥ فمن الممكن أن نطرح ٦٥ من كل واحد من هذه الأعداد ، ونقارن مجموعتى الفروق بدلا من الأعداد نفسها ، وفى ذلك تسهيل كبير للعمل الحسابى . وهذا طبعاً لا يؤثر فى النتيجة لأننا نعتمد فى بحثنا على انحرافات المفردات عن وسطها الحسابى ، وطرح عدد ثابت (أو إضافته) من جميع المفردات لا يؤثر مطلقاً فى هذه الانحرافات .

وبإجراء هذا الطرح تصبح المفردات فى المجموعتين كما يلى :

المجموعة ا : ٢ ، ١٢ ، ٣ ، ١٠ ، ١٢ ، ١٤ ، ١٢ ، ٨ ، ٢ ؛

» ب : ١ ، ١ ، ٥ ، ١ ، ٥ ، ٢ ، ٠ ، ٦ .

وبحساب الوسطين الحسابيين نجد أن :

$$\bar{\text{س}}_{١} = ٨{,}٣٣٣ \qquad ، \qquad \bar{\text{س}}_{٢} = ٢{,}٦٢٥$$

$$\text{مج}\,(\text{س} - ٨{,}٣٣٣)^{٢} = ١٨٤{,}٠٠$$

$$\text{مج}\,(\text{سَ} - ٢{,}٦٢٥)^{٢} = ٢٧{,}٨٧٥$$

$$\overline{٢٢١{,}٨٧٥}$$

$$\therefore\ \text{ى}^{٢} = \frac{١}{\text{ن}_{١} + \text{ن}_{٢} - ٢}\left[\text{مج}\,(\text{س} - \bar{\text{س}}_{١})^{٢} + \text{مج}\,(\text{سَ} - \bar{\text{س}}_{٢})^{٢}\right]$$

$$= \frac{٢٢١{,}٨٧٥}{١٥} = ١٤{,}٧٩١٧$$

$$\therefore\ \text{ى} = ٣{,}٨٤٣٤$$

$$= \frac{\bar{\text{س}}_{١} - \bar{\text{س}}_{٢}}{\text{ى}} \sqrt{\frac{\text{ن}_{١} \cdot \text{ن}_{٢}}{\text{ن}_{١} + \text{ن}_{٢}}}$$

$$= \frac{٨{,}٣٣٣ - ٢{,}٦٢٥}{٣{,}٨٤٣٤} \sqrt{\frac{٩ \times ٨}{١٧}}$$

$$= ٠٣{,}٠٥٦$$

وبالرجوع إلى جداول فيشر بعدد ١٥ لدرجات الحرية نجد أن ت لا تصل إلى هذه القيمة إلا نادراً جداً جداً — أقل من ١ ٪ . [وإذا رجعنا إلى جداول استيودنت نجد أن ت تزيدعن هذه القيمة أقل من ٨ مرات فى كل ألف مرة]. وهذا يدل على أن الفرق بين مجموعتى الطلبة فرق معنوى بدرجة عالية ولا يمكن تفسيره بإرجاعه إلى الصدف . ولايمكن إذن أن نعتبر المجموعتين من الطلبة كأنهما عينتان من نفس المجتمع .

مثال (٢) — قيست أطوال ٦ من البحارة انتخبناهم حيثما اتفق فكانت

أطوالهم بالبوصة ٦٣ ، ٦٥ ، ٦٨ ، ٦٩ ، ٧١ ، ٧٢ . وقيست أطوال عشرة من الجنود فكانت أطوالهم بالبوصة هى :

٦١ ، ٦٢ ، ٦٥ ، ٦٦ ، ٦٩ ، ٦٩ ، ٧٠ ، ٧١ ، ٧٢ ، ٧٣

فهل فى هذه الأرقام ما يؤيد الفرض القائل بأن البحارة أطول قامة من الجنود .

نقارن بين متوسطى الأطوال فى المجموعتين . وبما أن كل الأطوال أكبر من ٦٠ بوصة فيصح أن نطرح ٦٠ من جميع الأعداد ، وبذلك تصبح المفردات كما يلى :

البحارة : ٣ ، ٥ ، ٨ ، ٩ ، ١١ ، ١٢ وسطها الحسابى $\bar{\text{س}}_{١} = ٨$

الجنود : ١ ، ٢ ، ٥ ، ٦ ، ٩ ، ٩ ، ١٠ ، ١١ ، ١٢ ، ١٣ » $\bar{\text{س}}_{٢} = ٧٫٨$

$$\text{مجـ}(\text{س} - ٨)^{٢} = ٢٥ + ٩ + ٠ + ١ + ٩ + ١٦ = ٦٠$$

$$\text{مجـ}(\text{س} - ٧٫٨)^{٢} = \text{مجـ}(\text{س} - ٨)^{٢} - ١٠\,(٨ - ٧٫٨)^{٢}$$

$$= ٤٩ + ٣٦ + ٩ + ٤ + ١ + ١ + ٤ + ٩ + ١٦ + ٢٥ - ٠٫٤$$

$$= ١٥٣٫٦$$

$$\therefore \quad \text{ى}^{٢} = \frac{١}{٦ + ١٠ - ٢}\,(٦٠ + ١٥٣٫٦) = ١٥٫٢٥٧١$$

$$\therefore \quad \text{ى} = ٣٫٩٠٦٠$$

$$\text{ولكن} \quad \text{ت} = \frac{\bar{\text{س}}_{١} - \bar{\text{س}}_{٢}}{\text{ى}} \sqrt{\frac{\text{ن}_{١} \cdot \text{ن}_{٢}}{\text{ن}_{١} + \text{ن}_{٢}}}$$

$$\therefore \quad \text{ت} = \frac{٠٫٢}{٣٫٩٠٦٠} \sqrt{\frac{٦ \times ١٠}{١٦}}$$

$$= ٠٫٠٩٩$$

وبالرجوع إلى جـــدول فيشر لقيم ت بدرجات حرية عددها ١٤ ، نجد أن هذه القيمة ٠٫٠٩٩ أصـغر جداً من أن تكون معنوية ، حيث ح لهذه القيمة أكبر من ٠٫٩ ، ولـكى تكون ت معنوية بهذا العدد من درجات الحرية يجب أن تساوى ٢٫١٤٥ على الأقل .

وينتج من ذلك أنه لا يمكن القول بأن البحارة أطول قامةً من الجنود .

تعريف القيمة المعنوية للكمية ت

٤٣٩ — نرى من جـدول فيشر أنه يوجد لـكل مستوى ح من مستويات المعنوية ولـكل عدد ن من درجات الحرية قيمة معينة للـكمية ت هى $ت_{ح ن}$ مثلاً (وهى تقع فى ملتقى العمود ح والسطر ن فى الجـدول) نسميها [1] القيمة المعنوية للكمية ت ، بمعنى أن ت لا تعتبر معنوية إلا إذا كانت تساوى $ت_{ح ن}$ أو تزيد عنها . ففى المثال الأول من البند السابق مثلاً حيث اتخذنا ح = ٠٫٠١ مستوى للمعنوية وكانت ن = ١٥ نجد أن

$$ت_{ح ن} = ت_{٠٫٠١ ، ١٥} = ٢٫٩٥$$

ولذلك اعتبرنا ت = ٣٫٠٥٦ المحسـوبة من العينـة معنوية لأنهـا أكبر من $ت_{٠٫٠١ ، ١٥}$.

أما فى المثال الثانى فـكان عدد درجات الحرية ١٤ وباتخاذ مستوى المعنوية عند ح = ٠٫٠١ نجد أن القيمة المعنوية $ت_{٠٫٠١ ، ١٤}$ = ٢٫٩٨ وهذه أكبر بكثير من قيمة ت التى حصلنا عليها من العينة وهى ٠٫٠٩٩ وبناءً على ذلك اعتبرنا ت = ٠٫٠٩٩ غير معنوية ، وقررنا بناءً على ذلك أنه لا يوجد فرق معنوى بين أطوال البحارة وأطوال الجنود .

---

(١) بالإنجليزية Significant value of t . والرمز $ت_{ح ن}$ يقرأ « ت نون و حاء » والرمز $ت_{٠٫٠١ ، ١٥}$ يقرأ « ت ١٥ وواحد من مائة » .

معنوية الوسط الحسابى للعينة

الفرق بين متوسط العينة والمجتمع يتوقف على ت

٤٤٠ — رأينا فى تعريف الكمية ت وتطبيقاتها أنها عبارة عن الفرق بين متوسط العينة ومتوسط المجتمع الأصلى ، مقسوماً على الخطأ المعيارى (التقديرى) لوسط العينة ؛ أى أنها تساوى هذا الفرق مقدراً بوحدات تساوى الخطأ المعيارى التقديرى المحسوب من مفردات العينة . وهذا واضح من المعادلة (٣) بند ٤٤٣ ، وهى :

$$\text{ت} = \frac{\bar{\text{س}} - \bar{\text{س}}_{.}}{\text{ى}} \sqrt{\text{ن}}$$

ولو كتبنا للاختصار $\text{ى}'$ بدلا من $\frac{\text{ى}}{\sqrt{\text{ن}}}$ وهو الخطأ المعيارى التقديرى ،

$$\therefore \quad \text{ت} = \frac{\bar{\text{س}} - \bar{\text{س}}_{.}}{\text{ى}'} \,.$$

$$\therefore \quad \bar{\text{س}} - \bar{\text{س}}_{.} = \text{ت}\ \text{ى}' \,.$$

أى أن الفرق بين متوسط العينة ومتوسط المجتمع يساوى ت مرات الخطأ المعيارى المقدر من العينة .

وبما أن ت يصح أن تكون سالبة أو موجبة دون أن يؤثر ذلك على الاحتمال (لأن التوزيع التكرارى لقيم ت متماثل بالنسبة إلى ت = صفراً كما رأينا فى بند ٤٣٣) فيمكننا أن نقول إن

$$\bar{\text{س}}_{.} = \bar{\text{س}} \pm \text{ت}\ \text{ى}' \,.$$

أى أن الوسط الحسابى للمجتمع يقع فى حدود منطقة تحيط بالوسط $\bar{\text{س}}$ المحسوب من العينة ، وتمتد من $\bar{\text{س}} - \text{ت}\ \text{ى}'$ إلى $\bar{\text{س}} + \text{ت}\ \text{ى}'$. وهكذا نحدد موقع

الوسط الحسابى للمجتمع فى منطقة حول متوسط العينة ، يضيق مداها أو يتسع تبعاً لقيمة ت المحسوبة من مفردات العينة .

٤٤١ — وقد رأينا أن قيم ت تتبع احتمالات مختلفة حسب عدد درجات الحرية ، ففى جدول فيشر نرى مثلاً أن

$$ت_{٠١ر٠ ، ٢٤} = ٢ر٨٠$$

أى أن قيمة ت المحسوبة من عينة عدد مفرداتها ٢٥ (وعدد درجات الحرية ٢٤) لا تتعدى ٢ر٨ إلا مرة واحدة فى كل مائة مرة . ففى مثل هذه الحالة يكون

$$\bar{س}_{.} = \bar{س} \pm ٢ر٨٠ \ \bar{ى}$$

بمعنى أن $\bar{س}_{.}$ لا تخرج عن حدود هـذه المنطقة إلا مرة فى كل مائة مرة . فإذا قررنا أن متوسط المجتمع $\bar{س}_{.}$ يقع فى الفترة $\bar{س} \pm ٢ر٨٠ \ \bar{ى}$ يكون حكمنا هذا صواباً فى ٩٩ ٪ من الحالات ويكون خطأ فى ١ ٪ فقط .

مواضـــع وحدود الثقة

٤٤٢ — وهكذا يمكننا تقدير الوسط الحسابى للمجتمع بتحديد فترة نرجح أنه يكون فى داخلها ، وحساب احتمال وجوده داخلها أو خارجها . ويتوقف طول هذه الفترة وحدّاها الأعلى والأدنى واحتمال وجود المتوسط بينهما أو خارجهما على مقدار الكمية ت والخطأ المعيارى التقديرى ى ($= \bar{ى} \sqrt{ن}$) . وكلاهما محسوب من مفردات العينة .

هذه الفـترة التى نرجح وجود المتوسط العـام داخلها والتى تبـدأ عند $\bar{س} - ت \ \bar{ى}$ وتنتهى عند $\bar{س} + ت \ \bar{ى}$ نسميها [١] **الفترة موضع الثقة**

(١) بالإنجليزية Confidence or Fiducial Interval

والمقداران $\bar{س} \pm ت\, ى\bar{}$ نسميهما[1] **حدا الثقة** ، واحتمال وجود المتوسط بين هذين الحدين نسميه[2] **درجة الثقة** .

مثال — فى تجربة زراعية على نوع من أنواع القمح كان مقدار المحصول فى عشرة حقول كما يأتى بالأردب للفدان :

٥ر٤ ، ٤ر٧ ، ٥ر٦ ، ٥ر٣ ، ٦ر٢ ، ٤ر٩ ، ٥ر٨ ، ٥ر٤ ، ٦ر٠ ، ٥ر٧

فما مقدار متوسط محصول الفدان من هذا النوع على العموم .

هذه العشر مفردات متوسطها الحسابى $\bar{س}$ = ٥ر٥ والمطلوب معرفة المتوسط العام $\bar{س}$. للمجتمع المكوّن من جميع الحقول المزروعة بهذا النوع من القمح، على اعتبار هذه الحقول العشرة عينة عشوائية من هذا المجتمع .

$$مج\,(س - \bar{س})^{٢} = ٠{,}٠١ + ٠{,}٦٤ + ٠{,}٠١ + ٠{,}٠٤ + ٠{,}٤٩ + ٠{,}٣٦$$

$$+ ٠{,}٠٩ + ٠{,}٠١ + ٠{,}٢٥ + ٠{,}٠٤ = ١{,}٩٤$$

$$\therefore \quad ى^{٢} = \frac{١{,}٩٤}{٩} = ٠{,}٢١٥٥$$

$$\therefore \quad ى = ٠{,}٤٦٤٢ \text{ أردباً}$$

$$\therefore \quad \bar{ى} = \frac{ى}{\sqrt{١٠}}$$

= ١٤٦٧ر٠ أردباً = الخطأ المعيارى التقديرى .

وبالرجوع إلى جداول ت على أساس مستوى المعنوية ع = ٠١ر٠ وعدد درجات الحرية ٩ نجد أن $ت_{٩,٠١}$ = ٣ر٢٥ .

---

(١) بالإنجليزية Confidence or Fiducial Limits

(٢) « Confidence Level

∴ $ت_{٠٫٠١،٩} \times \bar{ى} = ٠٫٤٧٦٧ = ٠٫٤٨$ أردبا تقريباً

∴ الفترة موضع الثقة هى ٥ر٥ ± ٤٨ر أردباً ودرجة الثقة هنا ٩٩ ٪ ، وحدّا الثقة هما ٥ر٠٢ و ٥ر٩٨ أردباً .

أى أننا نعتبر متوسط المحصول من هـذا النوع على العموم واقعاً بين ٥ر٠٢ و ٥ر٩٨ أردباً للفدان ويكون هذا الحكم صواباً فى ٩٩ ٪ من المرات وخطأ فى ١ ٪ .

و إذا اعتمدنا على مستوى المعنوية ٠٫٠٥ بدلا من ٠١ر فنجد فى جدول ت عند هذا المستوى وذلك العدد ٩ من درجات الحرية أى $ت_{٠٫٠٥،٩} = ٢٫٢٦٢$ ، و بناء عليه يكون

$$ت_{٠٫٠٥،٩} \times \bar{ى} = ٠٫٣٣١٨ \text{ أردباً}$$

وتكون حـدود الثقة ٥ر١٧ و ٥ر٨٣ أردباً وفترة الثقة ٥ر١٧ — ٥ر٨٣ ، ودرجـة الثقة ٩٥ ٪ ، بمعنى أننا نقول إن المتوسط العمومى لمحصول الفـدان من هذا الصنف يقع بين ٥ر١٧ و ٥٫٨٣ أردباً ، ويكون هذا الحكم صحيحاً فى ٩٥ ٪ من الحالات وخطأ فى ٥ ٪ .

و يلاحظ من هذا المثال أنه كلما ضاقت فترة الثقة ضعف احتمال صحة الحكم ، و يزداد هذا الاحتمال كلما اتسعت الفترة ، وهذا معقول طبعاً .

تطبيق ت فى اختبار معنوية معامل الارتباط

٤٤٣ — يمكننا اسـتخدام ت فى اختبار معنوية معامل الارتباط $ر$ الذى نحصل عليه من العينات الصغيرة المأخوذة من مجتمع توزيعه التكرارى معتدل وفيه معامل الارتباط الحقيقى $\bar{ر}$ يساوى صفراً . فى مثل هذه الحالة يكون التوزيع

التكرارى لقيم ر المحسوبة من عينات عشوائيـة مأخوذة من هذا المجتمع ، له المعادلة البسيطة الآتية :

$$\text{ص} = \text{ص}_{\cdot}\,(\text{١} - \text{ر}^{\text{٢}})^{\frac{\text{ن} - \text{٤}}{\text{٢}}}$$

حيث ن عدد مفردات العينة و ص. كمية مستقلة عن ر . وإذا وضعنا

$$\text{ت} = \frac{\text{ر}}{\sqrt{\text{١} - \text{ر}^{\text{٢}}}} \cdot \sqrt{\text{ن} - \text{٢}}$$

يمكن إثبات أن ت هذه لها نفس التوزيع التكرارى للكمية ت الذى اكتشفه استيودنت وتكلمنا عنه فى البنود السابقة فى هذا الباب . وعدد درجات الحرية هنا يساوى ن — ٢ ، حيث فى كل عينة نحسب ر باستخدام الوسطين الحسابيين $\bar{\text{س}}$ و $\bar{\text{ص}}$ من نفس البيانات الموجودة فى العينة ، ولذلك نفقد درجتين من درجات الحرية . وعلى ذلك يمكن اختبار معنوية ت بالطريقة المعروفة وبالتالى نختبر معنوية أى معامل ارتباط ر نحصل عليه من عينة . فإذا كان معنوياً دل ذلك على أن هناك ارتباط حقيقةً فى المجتمع المأخوذة منه العينة .

مثال — فى عينة عشوائية تحتوى على ١٥ زوجاً من قيم المتغيرين س و ص وجد أن معامل الارتباط ر = — ٠ر٥ فهل هذا يدل على وجود ارتباط فعلا فى المجتمع .

نفرض مبدئياً أن هذا المجتمع ليس فيه ارتباط (أى أن $\bar{\text{ر}}$ = ٠) ثم نبحث فيما إذا كانت الصدف وحدها كفيلة بإعطائنا قيمة مثل التى وجدناها وهى ر = — ٠ر٥ لذلك نحسب ت من المعادلة المذكورة أعلاه وهى :

$$\text{ت} = \text{ر} \cdot (\text{١} - \text{ر}^{\text{٢}})^{-\frac{\text{١}}{\text{٢}}} \cdot \sqrt{\text{ن} - \text{٢}} \text{ ،}$$

$$= -\frac{\text{١}}{\text{٢}} \cdot \left(\frac{\text{٣}}{\text{٤}}\right)^{-\frac{\text{١}}{\text{٢}}} \cdot \sqrt{\text{١٣}}$$

$$= -\sqrt{\frac{\text{١٣}}{\text{٣}}} = \sqrt{\text{٤٫٣٣٣}}$$

$$= \text{٢٫٠٨١}$$

بصرف النظر عن الإشارة ، ودرجات الحرية ١٣ . وبالرجوع إلى جدول فيشر لقيم ت ، نجد أن $\text{ت}_{\text{٠٫٠٥ ، ١٣}} = \text{٢٫١٦٠}$ ، $\text{ت}_{\text{٠٫٠١ ، ١٣}} = \text{٢٫٦٥٠}$. وينتج من ذلك أن ت = ٢٫٠٨١ ليست معنوية . وبناءً عليه نرى أن الصدف كفيلة وحدها بأن تأتى بقيمة لمعامل الارتباط ر مثل التى حصلنا عليها من هذه العينة وهى — ٠٫٥ حتى ولو كان المجتمع الأصلى فيه $\bar{\text{ر}} = \text{٠}$ ؛ وهذا يدل على أن ر = — ٠٫٥ التى حصلنا عليها فى العينة ليست معنوية ويصح أن نحصل عليها صدفةً من مجتمع ليس فيه ارتباط .

[يلاحظ أن ت = ٢٫٠٨١ قريبة من المعنوية $\text{ت}_{\text{٠٫٠٥ ، ١٣}} = \text{٢٫١٦٠}$ فيجوز اعتبار أن ر = — ٠٫٥ تقريباً معنوية ، ولكن لا يمكن الاعتماد على هذا .]

حالة المجتمع الذى فيه $\bar{\text{ر}}$ لاتساوى صفراً

٤٤٤ — هذه الطريقة لا يمكن استخدامها إذا علمنا أن المجتمع الأصلى فيه ارتباط ، أى حينما يكون $\bar{\text{ر}} \neq \text{٠}$ ، حيث يكون التوزيع التكرارى للمعامل ر فى العينات العشوائية أكثر تعقيداً من الذى ذكرناه فى البند السابق . ولاختبار معنوية معامل الارتباط ر فى العينات المأخوذة من مجتمع فيه معامل الارتباط

الحقيقى $\bar{ر} \neq ٠$ ، وضع ر . ا . فيشر مقياساً آخر نسميه [١] بالعربية فز وعرّفه بالمعادلة

$$فز = \frac{١}{٢} لو_{هـ} \frac{١ + ر}{١ - ر}$$

وباستخدام اللوغاريثمات العادية للأساس ١٠ يكون [٢]

$$فز = ١٫١٥١٣ لو_{١٠} \frac{١ + ر}{١ - ر}$$

وبحث ر . ا . فيشر فى التوزيع التكرارى لهذه الكمية فز فوجد [٣] أنه توزيع تكرارى معتدل تقريباً وسطه الحسابى يساوى

$$\overline{فز} = \frac{١}{٢} لو_{هـ} \frac{١ + \bar{ر}}{١ - \bar{ر}} = ١٫١٥١٣ لو_{١٠} \frac{١ + \bar{ر}}{١ - \bar{ر}}$$

وانحرافه المعيارى يساوى $\frac{١}{\sqrt{ن - ٣}}$ ، حيث ن تساوى عدد مفردات العينة .

مثال (١) — حسب معامل الارتباط ر فى عينة مفرداتها ٢٥ فوجد أن ر = ٥٨ر٠ فهل هذا يدل على أن المجتمع الأصلى للعينة فيه ارتباط بين المتغيرين .

لنفرض مبدئياً أن المجتمع عديم الارتباط ، أى أن $\bar{ر} = ٠$ ، فهل يمكن أن نحصل على مثل هذه القيمة ر = ٥٨ر٠ من عينة عشوائية حجمها ٢٥ من هذا المجتمع ؟ . على هذا الفرض يكون

---

(١) يسمى بالإنجليزية ، R. A. Fisher's z– Transformation ، ويرمز له بالحرف z . ولكن هذا الرمز مع الأسف يلتبس مع z أخرى لفيشر نفسه . ولذلك نقترح الرمز فز — ينطق فازين — لهذه ، وللاخرى ظ كما سنرى فى الباب التالى .

(٢) معلوم أن $لو_{هـ} س = لو_{١٠} س$ مقسوماً على $لو_{١٠} هـ$ ، وهذا الأخير يساوى ٤٣٤٢٩ر٠

(٣) أنظر كتاب يول كندال صحيفة ٤٥١ مع (طبعة ١٩٤٦)

$$\text{فز} = ١٫١٥١٣ \text{ لو}_{١٠} \frac{١ + ٥٨ر٠}{١ - ٥٨ر٠}$$

$$= ٦٥١١ر٠$$

وبما أن $\bar{\text{ر}}$ = صفراً يكون متوسط فز صفراً .

∴ فز فى العينة تنحرف عن متوسطها فى جميع العينات بمقدار ٦٥١١ر٠ ، وهذا الانحراف نقارنه بالانحراف المعيارى للكمية فز نفسها وهو كما قلنا

$$\frac{١}{\sqrt{٢٢}} = ٢١٣٣ر٠ \text{ لأن } ن = ٢٥ :$$

∴ فز = ٦٥١١ر٠ أكبر من ثلاثة أمثال انحرافها المعيارى ؛ وهى إذن معنوية . وينتج من ذلك أن هذه العينة لا يمكن أن نحصل عليها من مجتمع عديم الارتباط

∴ ر = ٥٨ر٠ من عينة حجمها ٢٥ تدل على وجود ارتباط فى المجتمع .

مثال (٢) — فى المثال السابق ابحث الفرض بأن المجتمع الأصلى فيه ارتباط معاملة $\bar{\text{ر}}$ = ٨ر٠

$$\text{هنا نحسب} \quad \text{فز} = ١٫١٥١٣ \text{ لو} \frac{١ + \text{ر}}{١ - \text{ر}} = ٦٥١١ر٠$$

$$\text{ومتوسطها الحسابى} \quad \bar{\text{فز}} = ١٫١٥١٣ \text{ لو} \frac{١ + \bar{\text{ر}}}{١ - \bar{\text{ر}}}$$

$$= ١٫١٥١٣ \text{ لو} \frac{١ + ٨ر٠}{١ - ٨ر٠}$$

$$= ١٫٠٩٨٥$$

∴ فز − $\overline{\text{فز}}$ = ١٫٠٩٨٥ − ٠٫٦٥١١

= ٠٫٤٧٤٤

وهذا الفرق نقارنه بالانحراف المعيارى ٠٫٢١٣٣ فنجده أقل من ثلاثة أمثاله فهو ليس معنوى . وبناء عليه فالفرض أن $\overline{\text{ر}}$ = ٠٫٨ لايوجد فى نتيجة العينة ما يناقضه .

مثال (٣) — المطلوب تطبيق فز لاختبار معنوية ر = − ٠٫٥ التى حصلنا عليها فى المثال المذكور فى البند السابق من عينة حجمها ١٥ .

هنا فز = ١٫١٥١٣ لو$_{\text{١٠}}$ $\dfrac{\text{١} - \text{٠٫٥}}{\text{١} + \text{٠٫٥}}$ لأن ر سالبة

= − ٠٫٥٤٩٢

ومتوسطها $\overline{\text{فز}}$ = ١٫١٥١٣ لو $\dfrac{\text{١} - \text{٠}}{\text{١} + \text{٠}}$

= صفراً ؛

وانحرافها المعيارى $\dfrac{\text{١}}{\sqrt{\text{١٢}}}$ = ٠٫٢٨٨٧ لأن ن = ١٥ .

∴ فز − $\overline{\text{فز}}$ = ٠٫٥٤٩٢

وهذا أقل من ضعف الانحراف المعيارى فهو إذن ليس معنوياً . ومعنى ذلك أن ر = − ٠٫٥ من عينة حجمها ١٥ يصح أن نحصل عليها من مجتمع عديم الارتباط فيه $\overline{\text{ر}}$ = ٠ وعلى ذلك فالفرض بأن $\overline{\text{ر}}$ = ٠ لا يوجد فى نتيجة العينة ما يناقضه .

فى حالة العينات الـكبيرة

٤٤٥ — وطبعا لو كانت ن فى هـذه العينات كبيرة كنا نستخدم المعادلة التى ذكرناها فى جدول ٥٧ للانحراف المعيارى لمعامل الارتباط ر وهى

$$\frac{١ - ر^{٢}}{\sqrt{ن}} ،$$

ونختبر معنوية ر بمقارنتها بهذا الانحراف المعيارى . ولـكن فى العينات الصغيرة نتعرض لخطأ كبير لو اعتمدنا على هذه المعادلة .

أما بخصوص نسبة الارتباط ى التى نحسبها من جدول الارتباط المزدوج، فيمكننا اختبار معنويتهـا بطرق تحليل التباين التى سنشرحهـا فى الباب التالى (أنظر بند ٤٦٢ والتمرين رقم ١٥٨ فى آخر الجزء الثانى من الكتاب) .

المـــراجع

Fisher, R.A.: *Statistical Medhods for Research Workers*, Chapters.

Fisher and Yates: *Statistical Tables for Biological, Agricultural and Medical Research.*

Pearson, K.: *Tables for Statisticians and Biometricians.*

Pearson and Bennet: *Statistical Methods*, Chapter 17.

Petzers and Van Voorhis: *Statistical Procedures and their Mathematical Bases*, Chapter 6.

Snedecor, G.W.: *Statistical Methods*, Chapters 2, 3, 6,

"Student": *Collected Papers.*

Yule, G.U. and Kendall, M.G.: *An Introduction to the Theory of Statistics*, Chapter 23.

الباب السابع عشر

تحليل التباين

مقارنة التشتت في مجموعتين لا يمكن استخدام ت

٤٤٦ — رأينا فى الباب السابق الطرق التى نستخدمها فى العينات الصغيرة لكى نستدل من متوسط العينة على الحدود التى يقع بينها متوسط المجتمع ، أو لنختبر معنوية الفرق بين متوسطى عينتين لنعرف ما إذا كانتا من مجتمع واحد أو من مجتمعين مختلفين ، كل ذلك بالاعتماد على القيم أو البيانات الموجودة فى العينات وبدون توقف على معرفة المعالم الإحصائية للمجتمع (١) . واستخدمنا لذلك الكمية ت التى وضعها استيودنت وتوزيعها التكرارى الذى أوجده والجداول التى أنشأها هو وغيره لحساب احتمالات القيم المختلفة لتلك الكمية . والآن نبحث فى طريقة أخرى لاختبار الفرق بين التشتت فى مجموعتين أو عينتين ، إذ أن الكمية ت وتوزيعها التكرارى لا ينطبقان على مقاييس التشتت .

٤٤٧ — لنفرض على العموم أن لدينا مجموعتان أو عينتان كما يلى :

الأولى مفرداتها $س_١ ، س_٢ ، . . . ، س_{ن_١}$ وعددها $ن_١$ ومتوسطها $\bar{س}_١$

و الثانية « $سَ_١ ، سَ_٢ ، . . . ، سَ_{ن_٢}$ « $ن_٢$ « $\bar{س}_٢$.

---

(١) اقترح استخدام هذه العبارة ترجمة للعبارة الإنجليزية Statistical Parameters of The Population وهى الخواص الاحصائية التى تميز مجتمعاً عن غيره ومنها الوسط الحسابى والانحراف المعيارى لهذا المجتمع وغيرهما . ومفردها معلم ومعناه الخاصة أو الصفة المميزة . وأشكر حضرة الدكتور عبد العزيز عبد المجيد حيث أوضح فى هـذا المعنى إذ يقال معلم الطريق أو المكان بمعنى العلامة أو الصفة التى تميزه عن غيره .

نحسب التباين التقديرى ى^{٢} فى هاتين الحالتين . فى الحالة الأولى يكون :

$$ى_{١}^{٢} = \frac{١}{ن_{١}} مج(س — \bar{س}_{١})^{٢}$$ ، وله درجات حرية عددها $ن_{١} = ح_{١} — ١$ ؛

وفى الحالة الثانية يكون :

$$ى_{٢}^{٢} = \frac{١}{ن_{٢}} مج(س — \bar{س}_{٢})^{٢}$$ ، وله درجات حرية عددها $ن_{٢} = ح_{٢} — ١$.

وكل من ى$_{١}^{٢}$ و ى$_{٢}^{٢}$ عبارة عن تقدير للتباين فى المجتمع الأصلى المأخوذ منه العينة، فإذا اختلف هذان المقداران اختلافاً حقيقياً كان ذلك دليلا على أن العينتين مأخوذتين من مجتمعين مختلفين حقيقة من حيث التباين :

٤٤٨ — ولو كانت هاتان العينتان كبيرتين لأمكننا بحث الفرق بين انحرافيهما المعياريين ع$_{١}$ و ع$_{٢}$ واختبار معنويته بالطريقة التى أوردناها فى بند ٤١٣ فى الباب الخامس عشر . ولكن صغر ح$_{١}$ و ح$_{٢}$ وعدم معرفة الانحراف المعيارى ع$_{ }$ للمجتمع الأصلى يعرضنا لخطأ كبير فى الاستعاضة عن ع بالانحرافين ع$_{١}$ و ع$_{٢}$ وتكون النتيجة التى نصل إليها غير موثوق بها . توزيع ظ التكرارى

ولذلك فكر ر . ا . فيشر فى دراسة النسبة بين التباينين ى$_{١}^{٢}$ و ى$_{٢}^{٢}$ بدلا من الفرق بينهما ، إذ رأى أن هذه النسبة توضح الفرق بين الكميتين أكثر من باقى طرحهما . ولذلك وضع الرمز[1] ظ بالتعريف الآتى :

---

(١) نستخدم هذا الحرف للدلالة على الرمز z المستعمل فى الكتب الافرنجية تحت عنوان Fisher's z - Distribution . أنظر مثلا صفحة ٤٤٣ من كتاب Yule and Kendall السابق ذكره . نلاحظ هنا أن هـ هى الكمية التى سبقت الإشارة إليها فى بند ٧١ صفحة ٦٦ ، وهى المسماة أساس اللوغاريتمات الطبيعية وتساوى ٢٫٧١٨٢٨١٨ ولوغاريتمها للأساس المعتاد ١٠ يساوى ٠٫٤٣٤٢٩ .

$$\text{ظ} = \frac{١}{٢}\ \text{لوهـ}\ \frac{\text{ى}_{١}^{٢}}{\text{ى}_{٢}^{٢}} = \text{لوهـ ى}_{١} - \text{لوهـ ى}_{٢}$$

وبحث فى التوزيع التكرارى لقيم ظ فوجد أن معادلته هى

$$\text{ص} = \text{ص}_{.} \times \text{هـ}^{\text{ن}_{١}\text{ظ}} \times \left[\text{ن}_{١}\,\text{هـ}^{٢\text{ظ}} + \text{ن}_{٢}\right]^{-\frac{\text{ن}_{١}+\text{ن}_{٢}}{٢}}$$

حيث $\text{ن}_{١}$ تساوى عدد درجات الحرية للتباين الأكبر $\text{ى}_{١}^{٢}$ و $\text{ن}_{٢}$ عدد درجات الحرية للتباين الأصغر $\text{ى}_{٢}^{٢}$، و $\text{ص}_{.}$ كمية ثابتة نختارها بحيث تجعل المساحة الكلية للمنحنى تساوى الواحد الصحيح .. وبواسطة هذه المعادلة تمكن فيشر من حساب احتمالات قيم ظ المختلفة تبعاً لقيم $\text{ن}_{١}$ و $\text{ن}_{٢}$. ونظراً لصعوبة تبويب قيم ظ تبعاً للمتغيرات الثلاثة $\text{ن}_{١}$ و $\text{ن}_{٢}$ والاحتمال ح ، فقد اكتفى بالاحتمالات المهمة ٠٥ر ، ٠١ر ، ٠٠١ر٠ وحسب قيم ظ المعنوية لكل منها تبعاً لدرجات الحرية $\text{ن}_{١}$ و $\text{ن}_{٢}$ ؛ ورتبت هذه القيم فى ثلاثة جداول واحد لكل مستوى من مستويات المعنوية المذكورة [١]. وبواسطة هذه الجداول يمكننا اختبار معنوية أى قيمة للكمية ظ نحصل عليها فى أى تجربة، حيث الاحتمال ح فى الجداول يساوى احتمال حصولنا على مثل هذه القيمة أو أكبر منها فى المعاينة العشوائية .

مثال — فى إحدى تجارب تغذية الحيوان [٢] استحضر الباحث عشر أزواج من الخنازير وقسمها إلى مجموعتين ا ، ب وأعطى لكل حيوان ١٠٠ رطل من العلف ، فكانت الزيادة فى وزن الحيوانات فى المجموعتين كما يأتى بالأرطال :

---

(١) أنظر هذه الجداول فى صفحة ٢٤٩ من كتاب R. A Fisher : *Statistical Methods* أو صفحة ٥٣٨ من كتاب يول وكندال .

(٢) أنظر تجربة Crampton فى صفحة ٤٦ من كتاب *Statistical Methods* تأليف G. W. Snedecor

المجموعة ا : ١٧ ، ٢٢ ، ٢٢ ، ١٥ ، ٢٤ ، ٢٢ ، ٢١ ، ٢١ ، ١٧ ، ٢١

« ب : ٢٠ ، ٢١ ، ٢١ ، ٢٢ ، ٢٤ ، ٢٢ ، ٢٣ ، ٢١ ، ٢٢ ، ٢٣ .

والمطلوب معرفة ما إذا كان هناك اختلاف حقيقي بين هاتين المجموعتين فى التباين .

لتخفيف العمل الحسابى يمكن طرح العدد ٢٠ من كل هذه الأرقام بدون أى تأثير على النتيجة التى نبحث عنها ، فتصبح المفردات كما يأتى :

س : —٣ ، ٢ ، ٢ ، —٥ ، ٤ ، ٢ ، ١ ، ١ ، —٣ ، ١

سَ : ٠ ، ١ ، ١ ، ٢ ، ٤ ، ٢ ، ٣ ، ١ ، ٢ ، ٣

وينتج من ذلك أن $\bar{س}_{١} = ٠{,}٢$ ، $\bar{س}_{٢} = ١{,}٩$.

$\therefore$ $مج\,(س - \bar{س}_{١})^{٢} = مج\,س^{٢} - ١٠\,\bar{س}_{١}^{٢} = ٧٤ - {,}٤ = ٧٣{,}٦$ ؛

و $مج\,(سَ - \bar{س}_{٢})^{٢} = مج\,سَ^{٢} - ١٠\,\bar{س}_{٢}^{٢} = ٤٩ - ٣٦{,}١ = ١٢{,}٩$.

$$\therefore \quad ى_{١}^{٢} = \frac{١}{ن_{١} - ١}\, مج\,(س - \bar{س}_{١})^{٢}$$

$$= ٨{,}١٧٨$$

$$و \quad ى_{٢}^{٢} = \frac{١}{ن_{٢} - ١}\, مج\,(سَ - \bar{س}_{٢})^{٢}$$

$$= ١{,}٤٣٣$$

$$\therefore \quad ظ = \frac{١}{٢}\, لو_{هـ} \frac{٨{,}١٧٨}{١{,}٤٣٣} = \frac{٢{,}٣٠٢٦}{٢}\, لو_{١٠} \frac{٨{,}١٧٨}{١{,}٤٣٣} \text{ ،}$$

لأن اللوغاريتم الطبيعى لأى كمية (أى لوغاريتمها للأساس هـ) يساوى [١]

---

(١) إذا كان $س = لو_{هـ}\, ا$ مثلا ، فإن $هـ^{س} = ا$ حسب تعريف اللوغاريتم.

ومن ذلك ينتج أن $لو_{١٠}\, ا = لو_{١٠}\,(هـ^{س}) = س\,.\,لو_{١٠}\, هـ = ٠{,}٤٣٣٢٩ \times س$

$\therefore$ $لو_{هـ}\, ا = س = \frac{١}{٠{,}٤٣٤٢٩} \times لو_{١٠}\, ا = ٢{,}٣٠٢٦\, لو_{١٠}\, ا$

لوغاريتم $_{هـ}$ نفس الكمية للأساس ١٠ مضروباً فى العدد ٢٫٣٠٢٦ وهو
١ ÷ لو$_{١٠}$ هـ .

$$\therefore \quad ظ \;=\; ١٫١٥١٣ \;\big[\, لو_{١٠}\, ٨٫١٧٨ \;-\; لو_{١٠}\, ١٫٤٣٣ \,\big]$$
$$=\; ٠٫٨٧٠٨$$

وبالرجوع إلى جداول فيشر بدرجات الحرية $ن_{١} = ن_{٢} = ٩$ نجد أن ظ$_{٠٫٠٥}$ = ٠٫٥٧٦٢ تقريباً وأن ظ$_{٠٫٠١}$ = ٠٫٨٤١٠ تقريباً ، وهما قيمتا ظ المعنويتان فى المستويين ح = ٠٫٠٥ ، ٠٫٠١ للمعنوية على الترتيب . ومن ذلك يتضح أن ظ = ٠٫٨٧٠٨ التى حصلنا عليها فى هذه التجربة معنوية بدرجة أعلى من ٠٫٠١ ؛ أى أننا نحصل عليها فى المعاينات العشوائية أقل من مرة واحدة فى كل مائة مرة . وهذا دليل على أن هاتين المجموعتين تختلفان اختلافاً حقيقياً فى التباين .

نسبة التباين

٤٤٩ — ونظراً لصعوبة استخراج اللوغاريتمات وتحويلها من لوغاريتمات طبيعية بالأساس هـ إلى لوغاريتمات حسابية للأساس ١٠ ، فكر ج . و . سنديكور فى تبويب قيم النسبة $ى_{١}^{٢} : ى_{٢}^{٢}$ نفسها تبعاً لدرجات الحرية $ن_{١}$ ، $ن_{٢}$ عند مستويَى المعنوية ح = ٠٫٠٥ ، ٠٫٠١ ؛ وتسمى (١) هذه النسبة **نسبة التباين** ونرمز لها بالحرف ف ؛ وطبعاً تكون

$$ف = \frac{ى_{٢}^{٢}}{ى_{١}^{٢}} = هـ^{٢ظ}$$

لأن ظ = $\frac{١}{٢}$ لو$_{هـ}$ ف كما رأينا فى تعريف الكمية ظ فى البند السابق .

---

(١) بالانجليزية G. W. Snedecor's "Variance Ratio". أنظر مثلا كتابه « الطرق الإحصائية » صفحة ٢١٩ .

وعلى هذا الأساس أنشأ جدولا جديداً يعطى قيم ف المعنوية على المستوى ح = ٠٥ر٠ والمستوى ح = ٠١ر٠ مبوبة تبعاً لعدد مناسب من أعداد درجات الحرية $ن_١$ ، $ن_٢$. وبالرجوع إلى هذا الجدول يمكننا اختبار معنوية أى نسبة نحصل عليها للتباينين $ى_١^٢$ ، $ى_٢^٢$ فى أى تجربة . وهذا الجدول يغنينا عن استخراج اللوغاريتمات فى كل مسألة .

مثال — عمل اختبار لمجموعتين صغيرتين من الطلبة فكانت درجاتهم كما يأتى :

المجموعة الأولى : ٦٧ ، ٤٨ ، ٥٥ ، ٧٢ ، ٧٦ ، ٦١ ؛ ٧٠ ، ٥٨ ، ٦٣ ؛

« الثانية : ٦٣ ، ٦٩ ، ٦٤ ، ٥٩ ، ٧١ ، ٦٥ ، ٥٨ ، ٦٦ ، ٦٠ ، ٧٠ .

فهل يستدل من نتائج هذا الاختبار على أن إحدى المجموعتين أكثر تفاوتاً من الأخرى .

نحسب التباينين التقديريين $ى_١^٢$ ، $ى_٢^٢$ لهاتين المجموعتين من الدرجات ونحسب النسبة بينهما ف ، ونختبر معنويتها باستخدام جداول سنديكور . ولتخفيف الحساب نطرح العدد ٦٣ من درجات المجموعة الأولى و ٦٥ من الثانية ، فتصبح مفردات المجموعتين كما يلى :

س : ٤ ، — ١٥ ، — ٨ ، ٩ ، ١٣ ، — ٢ ، ٧ ، — ٥ ، ٠ ؛

سَ : — ٢ ، ٤ ، — ١ ، — ٦ ، ٦ ، ٠ ، — ٧ ، ١ ، — ٥ ، ٥ .

∴ $\bar{س}_١ = \frac{١}{٣}$ و $\bar{س}_٢ = -$ ٥ر٠ وهما الوسطان الحسابيان للمجموعتين .

∴ $مج(س - \bar{س}_١)^٢ = مج\, س^٢ - ٩\, \bar{س}_١^٢ = ٦٣٣ - ١ = ٦٢٢$ ؛

و $مج(سَ - \bar{س}_٢)^٢ = مج\, سَ^٢ - ١٠\, \bar{س}_٢^٢ = ١٩٣ -$ ٥ر٢ = ٥ر١٩٠ .

$$\therefore \quad ى_{١}^{٢} = \frac{٦٢٢}{٨} = ٧٧{,}٧٥ \text{ ،}$$

$$\text{و} \quad ى_{٢}^{٢} = \frac{١٩٠{,}٥}{٩} = ٢١{,}١٦٧ \text{ ،}$$

$$\text{و} \quad ف = \frac{٧٧{,}٧٥٠}{٢١{,}١٦٧} = ٣{,}٦٧ \text{ .}$$

وبالرجوع إلى جدول سنديكور عند $ن_{١}$ = ٨ و $ن_{٢}$ = ٩ ، نجد أن $ف_{٠٥{,}٠}$ = ٣٫٢٩ ، $ف_{٠١{,}٠}$ = ٥٫٤٧ ؛ وعلى ذلك فإن النسبة ف = ٣٫٦٧ التى حصلنا عليها هنا معنوية لحد معين ، فهى واقعة بين المستويين ٥ ٪ و ١ ٪ وأقرب إلى الأول منها إلى الثانى . ويستدل من ذلك على أن نتائج هذا الاختبار تشير إلى أن التباين أو التفاوت بين الطلبة فى المجموعة الثانية أقل منه فى المجموعة الأولى ، ويستحسن الحصول على معلومات إضافية إذا أردنا أن نصل إلى حكم قاطع نطمئن إلى صحته تمام الاطمئنان .

وطبعاً لو حسبنا ظ بدلا من ف ورجعنا إلى جداول ظ نحصل على نفس النتيجة ، فكما رأينا فى المثال السابق وفى تعريف ظ و ف

$$ظ = \frac{١}{٢} \text{ لو}_{هـ} ف = ١{,}١٥١٣ \text{ لو}_{١٠} (٣{,}٦٧)$$

$$= ٠{,}٦٥٠١$$

وبالرجوع إلى جداول ظ عند $ن_{١}$ = ٨ و $ن_{٢}$ = ٩ نجد أن :

$$ظ_{٠٥{,}٠} = ٠{,}٥٨٦٢ \quad ، \quad ظ_{٠١{,}٠} = ٠{,}٨٤٩٤$$

وعلى ذلك فالقيمة ظ = ٠٫٦٥٠١ التى حصلنا عليها واقعة بين المستويين ٥ ٪ و ١ ٪ ، وهى أقرب إلى الأول منها إلى الثانى . وهى نفس النتيجة التى وصلنا إليها عن طريق ف بمجهود حسابى أقل .

تحليل تباين مجتمع مكون من عدة مجموعات

تطبيق فكرة النسبة ف فى مقارنة المجموعات

٤٥٠ — هذه الأداة ف أو ظ التى رأينا استخدامها فى مقارنة التباين فى مجموعتين لها تطبيق مفيد جداً فى ميدان البحوث التجريبية ، خصوصاً فى ميدان التجارب الزراعية والصناعية . ففى كثير من الأحيان يقوم الباحث بعمل تجربة على عدد من المفردات يعالجها معالجات مختلفة لدراسة الآثار المترتبة على هذه المعالجات ، كأن يعطى أنواعاً مختلفة من العلف لمجموعات من الغنم مثلا ، بقصد معرفة أثر الأنواع المختلفة فى زيادة الوزن . أو يزرع أنواعاً مختلفة من القمح مثلا تحت ظروف واحدة ويقارن بين متوسط المحصول فى كل نوع ليعرف ما إذا كان هناك اختلاف حقيقى بين هذه الأنواع . وفى الصناعة مثلا قد يستورد مصنع معين إحدى مواده الخام من عدة مصادر مختلفة ، ويختبر عدداً من القطع المنتجة من كل نوع من أنواع الخام بقصد معرفة إذا ما كان هناك اختلاف حقيقى ناشىء من اختلاف مصدر الخام . وكذلك فى أبحاث التربية وعلم النفس نأتى بعدة مجموعات من التلاميذ ونقارن درجاتهم فى اختبار معين لنعرف ما إذا كانت الاختلافات المشاهدة بين درجاتهم ناشئة عن اختلافات حقيقية بين المستويات العلمية لهذه المجموعات .

تحليل التباين فى العينة إلى مركبتين

٤٥١ — لنفرض على العموم أن لدينا عينة كبيرة نوعاً مفرداتها $س_١$ ، $س_٢$ ، . . . وعددها يساوى ن ومتوسطها $\bar{س}$ وانحرافها المعيارى ع فيكون بطبيعة الحال

$$ن \bar{س} = مجـ س \quad ، \quad ن ع^٢ = مجـ (س - \bar{س})^٢ .$$

وإذا قسمنا هذه المفردات إلى ح من المجموعات الجزئية وكانت أعداد هذه المفردات

$$ن_١ ، ن_٢ ، ... ، ن_ر ، ... ، ن_هـ$$

وكانت المتوسطات الحسابية لهذه المجموعات والانحرافات المعيارية لها هى على الترتيب

$$م_١ ، م_٢ ، ... ، م_ر ، ... ، م_هـ$$

و $$ع_١ ، ع_٢ ، ... ، ع_ر ، ... ، ع_هـ$$

يكون $$ن_ر ع_ر^٢ = مج (س - م_ر)^٢ ،$$

للمجموعة الرائية ، وهكذا فى باقى المجموعات . أما فى العينة الكلية فإن انحراف المفردة عن الوسط الحسابى العام $\bar{س}$ هو $س - \bar{س}$ وبوضع

$$س - \bar{س} = (س - م_ر) + (م_ر - \bar{س})$$

نحصل على النتيجة الآتية (كما فعلنا فى بند ١٨٢ مع تغيير بسيط فى الرموز) :

$$ن ع^٢ = مج (س - \bar{س})^٢$$

$$= مج ن_ر ع_ر^٢ + مج ن_ر (م_ر - \bar{س})^٢ \quad ... (١)$$

وبالقسمة على $ن = مج ن_ر$ فى الطرفين

$$\therefore \quad ع^٢ = \frac{مج ن_ر ع_ر^٢}{مج ن_ر} + \frac{مج ن_ر (م_ر - \bar{س})^٢}{مج ن_ر}$$

$$= \bar{ع^٢} + ع_م^٢$$

حيث $\bar{ع^٢}$ هى الوسط المرجح للكميات $ع_ر^٢$ ، كل واحدة مرجحة بوزن يساوى عدد مفردات مجموعتها $ن_ر$ ، و $ع_م^٢$ هو مربع الانحراف المعيارى للمتوسطات .

وهكذا نرى أن التباين العام فى العينة الأصلية يتكون من جزأين ، أو يمكن تحليله إلى مركبتين ، هما متوسط التباين بين مفردات المجموعة الواحدة والتباين

بين متوسطات هذه المجموعات . وللاختصار يصح أن نسمى هاتين المركبتين على الترتيب تباين المفردات وتباين المجموعات .

مقدار كل مركبة بالنسبة إلى الأخرى يتوقف على طريقة انتخاب المجموعات

٤٥٢ — ومعنى ذلك أن هاتين المركبتين تكملان بعضهما ، فإذا زادت إحداهما نقصت الأخرى ليبقى مجموعهما ثابتاً وهو تباين العينة الأصلية ، ومهما كانت الطريقة المتبعة فى تقسيم المفردات إلى المجموعات الجزئية فإن مجموع المركبتين يكون ثابتاً لا يتغير .

وعلى ذلك إذا روعى عند تقسيم العينة إلى مجموعاتها الجزئية أن ننتخب المفردات المتشابهة من نوع ونجعلها وحدها فى مجموعة واحدة ، ونضع المفردات المتشابهة من نوع آخر وحدها فى مجموعة ثانية وهكذا ، فتكون نتيجة ذلك أن تكون مفردات كل مجموعة قريبة بعضها إلى بعض ويكون التباين بين مفردات المجموعة الواحدة صغيراً . وفى الوقت نفسه تكون المجموعات متميزة بعضها عن بعض ، فتكون الأوساط الحسابية لها مختلفة بعضها عن بعض اختلافاً محسوساً ، فيكون تباين المتوسطات كبيراً نسبياً .

أما إذا أجرينا تقسيم مفردات العينة إلى المجموعات الجزئيـة عشوائياً وحيثما اتفق ، فإن كل مجموعة تضم بين مفرداتها نصيبها العادل من كل نوع بدون أى تحيز . وبذلك تكون مفردات كل مجموعة خليطاً من كل الأنواع الموجودة فى العينة ، ويكون التباين بين مفردات المجموعة الواحدة كبيراً فى المتوسط . وفى الوقت نفسه تكون المتوسطات الحسابية للمجموعات قريبة إلى بعضها فيكون تباينها صغيراً نسبياً .

الفائدة العملية لهذه النتيجة

٤٥٣ — ومعنى ذلك أنه إذا كانت لدينا عينة مقسمة إلى مجموعات جزئية ، يمكننا تحليل التباين فيها إلى مركبتيه ؛ وإذا وجدنا أن تباين المتوسطات كبير نسبياً

(أكبر مما تأتى به الصدف فى المعاينات العشوائية) حكمنا بأن تقسيم العينة إلى مجموعات لم يكن عشوائياً ؛ أى أن المجموعات مختلفة عن بعضها اختلافاً حقيقياً ولا يمكن اعتبارها عينات عشوائية من نفس المجتمع .

وهكذا يكون لدينـا أداة لاختبار معنوية الاختلافات التى نشاهدها بين المجموعات التى يتكون منها مجتمع واحد ، لنعرف ما إذا كانت هذه المجموعات عشوائية أو مختارة تبعاً لنظام معين .

٤٥٤ — ولإجراء هذه المقارنة بين $\bar{\text{ع}}^{٢}$ و $\text{ع}_{\text{م}}^{٢}$ يصح أن نبحث عن قيمة النسبة $\text{ع}_{\text{م}}^{٢} : \bar{\text{ع}}^{٢}$ ، أو أى نسبة أخرى بين المركبتين ، فى حالة تقسيم العينة إلى مجموعات عشوائيـــة ، وتكون هذه النسبة بمثابة المعيار الذى نقيس عليه فى أى مسـألة معينة . ونظراً لأن عدد درجات الحرية له تأثير فى مقادير هذه التباينات فى المعاينات العشوائية فيجب أن يكون المعيار الذى نقيس عليه يأخذ فى الاعتبار عدد درجات الحرية .

انقسام درجات الحـرية إلى مركبتين

٤٥٥ — بالرجوع إلى المعادلات المذكورة فى بند ٤٥١ وهى

$$\text{ن}\,\text{ع}^{٢} = \text{ن}\,\text{ع}_{\text{م}}^{٢} + \text{مجـ}\,\text{ن}_{\text{ر}}\,\text{ع}_{\text{ر}}^{٢}$$

و $\text{مجـ}(\text{س} - \bar{\text{س}})^{٢} = \text{مجـ}\,\text{ن}_{\text{ر}}(\text{م}_{\text{ر}} - \bar{\text{س}})^{٢} +$ مجموع أمثال $\left[\text{مجـ}(\text{س} - \text{م}_{\text{ر}})^{٢}\right]$

نرى أن كل واحد من هذه الحدود له عدد من درجات الحرية يختلف عن الآخر :

فالمجموع $\text{مجـ}(\text{س} - \bar{\text{س}})^{٢}$ له $\text{ن} - ١$ من درجات الحرية ، حيث إن عدد مفرداته يساوى ن وليس بين مفرداته إلا علاقة واحدة أو قيـد واحد وهو $\text{مجـ}(\text{س} - \bar{\text{س}}) = ٠$.

والمجموع $\text{مجـ}\,\text{ن}_{\text{ر}}(\text{م}_{\text{ر}} - \bar{\text{س}})^{٢}$ وهو مجموع مربعات انحرافات المتوسطات

الجزئية $م_ر$ عن متوسطها العمومى $\overline{س}$ له ح — ١ من درجات الحرية، حيث إن عدد المتوسطات يساوى عدد المجموعات وهو ح ، ونفقد درجة حرية واحدة لأن $\overline{س}$ هى متوسط قيم م .

وأخيراً مجموع أمثال [$مج (س — م_ر)^٢$] عدد مفرداته ن وفيه المتوسطات $م_ر$ وعددها ح محسوبة من نفس المفردات ، وعلى ذلك يكون عدد درجات الحرية ن — ح .

ومن ذلك نرى أن العدد الكلى ن — ١ لدرجات الحرية ينقسم مثل التباين إلى جزأين يكملان بعضهما كما يلى :

$$ن — ١ = ح — ١ + ن — ح$$

والجزء الأول منهما يقترن بالمركبة الأولى للتباين والثانى يقترن بالمركبة الثانية .

تقدير تباين المجتمع بالقسمة على عدد درجات الحرية

٤٥٦ — وعلى ذلك إذا نظرنا للعينة الكلية كعينة عشوائية من مجتمع كبير تباينه الحقيقى $ع_.^٢$ ، فإننا نحصل على ثلاثة تقديرات مختلفة لهذا التباين بقسمة مجموع مربعات الانحرافات على عدد درجات الحرية (بند ٤٣١) :

$$\therefore \quad \frac{١}{ن — ١} مج (س — \overline{س})^٢ = ى_١^٢ \text{ مثلا}$$

هو التقدير الأول للتباين $ع^٢$.

وعلى فرض أننا قسمنا هذه العينة تقسيما عشوائية إلى مجموعات جزئية عددها ح ومتوسطاتها $م_١$ ، $م_٢$ ، . . . $م_ح$ وأحجامها $ن_١$ ، $ن_٢$ ، . . . ، $ن_ح$

فإن التقدير الثانى للتباين الأصلى $ع_.^٢$ يكون :

$$\frac{1}{\text{ح} - 1} \,\text{مج}\, \text{ن}_{\text{ر}} \left(\text{م}_{\text{ر}} - \text{س}\right)^{2} = \text{ى}_{1}^{2} \ \text{مثلا}.$$

وبالمثل يكون التقدير الثالث هو

$$\frac{1}{\text{ن} - \text{ح}} \,\text{مج} \left[\text{مج} \left(\text{س} - \text{م}_{\text{ر}}\right)^{2}\right] = \text{ى}_{2}^{2} \ \text{مثلا}.$$

والأصل فى هذين التقديرين الأخيرين ى_{1}^{2} ، ى_{2}^{2} أن يكونا متساويين أو قريبين إلى التساوى فى حالة التقسيم العشوائى للعينة إلى مجموعاتها الجزئية . أى أن النسبة

$$\frac{\text{ى}_{1}^{2}}{\text{ى}_{2}^{2}} = \text{ف}$$

تساوى الواحد الصحيح فى حالة التقسيم العشوائى ، أو قل إنها لا تختلف عن الواحد إلا بقدر ما تأتى به الصدف فى المعاينات العشوائية . فإذا وجدنا فى مسألة معينة أن ف تختلف اختلافاً معنوياً عن ١ فنستدل من ذلك على أن المجموعات غير عشوائية بل تختلف بعضها عن بعض اختلافاً حقيقياً .

وهكذا تؤول المسألة فى آخر الأمر إلى المقارنة بين تباينين ، لكل منهما عدد من درجات الحرية .

ويمكننا إذن تطبيق فكرة التوزيع التكرارى لنسبة التباين ف أو المقياس ظ السابق ذكرها . ولذلك نستخدم جداول ج . و . سنديكور أو ر . ا . فيشر لاختبار معنوية ف أو ظ حسب درجات الحرية ومستويات المعنوية المعروفة .

حساب ف عملياً

٥٤٧ — لتطبيق هذه النتائج عملياً نأخذ بعض أمثلة تجريبية ونحسب ف أو ظ فى كل حالة ، ونشرح الخطوات التى نتخذها فى كل حالة . ونبدأ بالحالة التى فيها المجموعات الجزئية كلها متساوية فى العدد ، وهذه أبسط فى الحساب طبعاً من الحالة العامة لأن هنا

حالة المجموعات المتساوية العدد

$$\text{ن}_{١} = \text{ن}_{٢} = \cdot\cdot = \text{ن}_{ر} = \cdots = \text{ن مثلا .}$$

فيما يلى بيان بأوزان ٣٥ حملا كلها متساوية فى العمر ومقسمة إلى خمس مجموعات حسب نوع ونظام التغذية التى أعطيت لها، وهى ا ، ب ، حـ ، ى ، هـ . وكل مجموعة مكونة من ٧ حيوانات أرقامها ١ ، ٢ ، ٠٠٠ ، ٧ :

أوزان المجموعة ا بالرطل قائم	:	٣٥ ، ٣٩ ، ٣٦ ، ٤٠ ، ٣٨ ، ٤١ ، ٣٧
« « ب « «	:	٤٠ ، ٣٨ ، ٤١ ، ٤٥ ، ٤٧ ، ٣٩ ، ٤٤
« « حـ « «	:	٤٣ ، ٥١ ، ٤٥ ، ٥٣ ، ٤٩ ، ٤٧ ، ٤٨
« « ى « «	:	٥٢ ، ٥٧ ، ٦٠ ، ٥١ ، ٥٤ ، ٥٩ ، ٥٢
« « هـ « «	:	٦١ ، ٦٣ ، ٥٧ ، ٥٩ ، ٥٦ ، ٦٤ ، ٦٠

والمطلوب معرفة ما إذا كان نوع ونظام التغذية لهما أثر حقيقى على الوزن .

فلكى نحسب التقديرين $\text{ى}_{١}^{٢}$ و $\text{ى}_{٢}^{٢}$ نحسب المجموعين

$\text{مجـ ن}_{ر}\,(\text{م}_{ر} - \bar{\text{س}})^{٢}$ و $\text{مجـ}\left[\,\text{مجـ}\,(\text{س} - \text{م}_{ر})^{٢}\,\right]$ ونقسم كلا منهما على عدد درجات الحرية التى يقترن بها . وهذا يتطلب حساب المتوسط العام $\bar{\text{س}}$ والمتوسطات الجزئية $\text{م}_{ر}$ ثم حساب الانحرافات $\text{س} - \text{م}_{ر}$ و $\text{م}_{ر} - \bar{\text{س}}$ وتربيعها وجمع مربعاتها .

طريقة مختصرة للحساب

٤٥٨ — ويمكننا تخفيف الحساب نوعاً بوضع $\text{ن}_1 = \text{ن}_2 = \cdots = \text{ن}$ واستخدام العلاقة المذكورة فى بند ٤٥٥ وهى

$$\text{مجـ}\,(\text{س} - \bar{\text{س}})^2 = \text{مجـ}\,\text{ن}\,(\text{م}_\text{ر} - \bar{\text{س}})^2 + \text{مجـ}\left[\text{مجـ}\,(\text{س} - \text{م})^2\right]$$

فنحسب المجموع الأخير والأكثر تعقيداً بحساب الأول والثانى وطرحهما .

والأفضل عملياً فى حساب $(\text{مجـ}\,\text{س} - \bar{\text{س}})^2$ أن نحسب مربعات المفردات س نفسها حيث يمكننا الاستعانة على ذلك بجداول مربعات الأعداد ، ونحصل على ما نريد باستخدام العلاقة المعروفة .

$$\text{مجـ}\,(\text{س} - \bar{\text{س}})^2 = \text{مجـ}\,\text{س}^2 - \text{ى}\,\bar{\text{س}}^2$$

$$= \text{مجـ}\,\text{س}^2 - \frac{1}{\text{ى}}\cdot(\text{مجـ}\,\text{س})^2$$

لأن $\text{مجـ}\,\text{س} = \text{ى}\,\bar{\text{س}}$ كما نعلم .

وكذلك نحسب المجموع الآخر $\text{مجـ}\,\text{ن}\,(\text{م}_\text{ر} - \bar{\text{س}})^2$ باستخدام نفس الفكرة وتحويله إلى مربعات أعداد لكى يمكن استخدام جداول المربعات . ويمكننا[1] بسهولة إثبات أن هذا المجموع

$$\text{مجـ}\,\text{ن}\,(\text{م}_\text{ر} - \bar{\text{س}})^2 = \frac{1}{\text{ن}}\,\text{مجـ}\,\text{ل}_\text{ر}^2 - \frac{1}{\text{ى}}\,(\text{مجـ}\,\text{س})^2 ،$$

حيث $\text{ل}_\text{ر}$ يســـاوى مجموع مفردات المجموعة الرائيـــة ، وطبعـــاً يكون

---

(١) مع ملاحظة أن عملية الجمع هنا تتناول المجموعات وعددها ح ، وملاحظة أن $\text{ح}\,\text{ن} = \text{ى}$ ، وأن $\text{ن}\,\text{م}_\text{ر} = \text{ل}_\text{ر}$ ، نرى أن :

$$\text{مجـ}\,\text{ن}\,(\text{م}_\text{ر} - \bar{\text{س}})^2 = \frac{1}{\text{ن}}\,\text{مجـ}\,\text{ن}^2\,(\text{م}_\text{ر} - \bar{\text{س}})^2 = \frac{1}{\text{ن}}\,\text{مجـ}\,(\text{ل}_\text{ر} - \text{ن}\,\bar{\text{س}})^2$$

$$= \frac{1}{\text{ن}}\left[\text{مجـ}\,\text{ل}_\text{ر}^2 - 2\,\text{ن}\,\bar{\text{س}}\cdot\text{مجـ}\,\text{ل}_\text{ر} + \text{ح}\,\text{ن}^2\,\bar{\text{س}}^2\right]$$

$$= \frac{1}{\text{ن}}\left[\text{مجـ}\,\text{ل}_\text{ر}^2 - 2\,\text{ن}\left(\frac{\text{مجـ}\,\text{س}}{\text{ى}}\right)(\text{مجـ}\,\text{س}) + \text{ن}\cdot\text{ح}\,\text{ن}\cdot\left(\frac{\text{مجـ}\,\text{س}}{\text{ى}}\right)^2\right]$$

$$= \frac{1}{\text{ن}}\,\text{مجـ}\,\text{ل}_\text{ر}^2 - \frac{1}{\text{ى}}\,(\text{مجـ}\,\text{س})^2 ،\ \text{لأن}\ \text{ح}\,\text{ن} = \text{ى}\,.$$

$ل_ر = ن\ م_ر$ ، ويكون $مجـ\ ل_ر = مجـ\ س$ ، وهو مجموع المفردات كلها التى تتكون منها المجموعات .

وهكذا تؤول عملية حساب المجموعين المطلوبين إلى حساب ثلاثة عناصر بسيطة هى :

(١) — مجموع مربعات المفردات كلها وهو $مجـ\ س^٢$ ،

(٢) — مربع مجموع المفردات كلها وهو $(مجـ\ س)^٢$ ،

(٣) — مجموع المقادير (حاصل جمع مفردات المجموعة)^٢ لكل المجموعات وهو $مجـ\ ل_ر^٢$.

وباستخدام جداول مربعات الأعداد يمكننا إنجاز هذه العمليات بسهولة . ويحسن ترتيب العمل فى جدول كالآتى :

رقم الحيوان	المجموعات ا	ب	ح	ء	هـ	كل العينة	
١	٣٥	٤٠	٤٣	٥٢	٦١		
٢	٣٩	٣٨	٥١	٥٧	٦٣		
٣	٣٦	٤١	٤٥	٦٠	٥٧		
٤	٤٠	٤٥	٥٣	٥١	٥٩		
٥	٣٨	٤٧	٤٩	٥٤	٥٦		
٦	٤١	٣٩	٤٧	٥٩	٦٤		
٧	٣٧	٤٤	٤٨	٥٢	٦٠		
(١) مجموع المفردات : ل =	٢٦٦	٢٩٤	٣٣٦	٣٨٥	٤٢٠	١٧٠١	= مجـ س
(٢) عدد « : ن =	٧	٧	٧	٧	٧	٣٥	= ن ح = ن
(٣) مربع مجموع « : $ل^٢$ =	٧٠٧٥٦	٨٦٤٣٦	١١٢٨٩٦	١٤٨٢٢٥	١٧٦٤٠٠	٢٨٩٣٤٠١	= $(مجـ\ س)^٢$
(٤) = (٣) ÷ (٢) =	١٠١٠٨	١٢٣٤٨	١٦١٢٨	٢١١٧٥	٢٥٢٠٠	٨٢٦٦٨٫٦	= $\frac{١}{ن}(مجـ\ س)^٢$
(٥) مجموع مربعات المفردات =	١٠١٣٦	١٢٤١٦	١٦١٩٨	٢١٢٥٥	٢٥٢٥٢	٨٥٢٥٧	= $مجـ\ س^٢$

$$\frac{\text{١}}{\text{ن}}\ \text{مجـ}\ \text{ل}_{\text{ر}}^{\text{٢}} = \frac{\text{١}}{\text{٧}} \times \text{٥٩٤٧١٣}$$

$$= \text{٨٤٩٥٩}$$

$$\therefore\ \text{مجـ}\ (\text{س} - \bar{\text{س}})^{\text{٢}} = \text{مجـ}\ \text{س}^{\text{٢}} - \frac{\text{١}}{\text{ن}}\ (\text{مجـ}\ \text{س})^{\text{٢}}$$

$$= \text{٨٥٢٥٧} - \text{٨٢٦٦٨٫٦} = \text{٢٥٨٨٫٤}$$

$$\text{مجـ}\ \text{ن}\ (\text{م}_{\text{ر}} - \bar{\text{س}})^{\text{٢}} = \frac{\text{١}}{\text{ن}}\ \text{مجـ}\ \text{ل}_{\text{ر}}^{\text{٢}} - \frac{\text{١}}{\text{ن}}\ (\text{مجـ}\ \text{س})^{\text{٢}}$$

$$= \text{٨٤٩٥٩} - \text{٨٢٦٦٨٫٨} = \text{٢٢٩٠٫٤}$$

$$\therefore\ \text{مجـ}\left[\text{مجـ}(\text{س} - \text{م}_{\text{ر}})^{\text{٢}}\right] = \text{مجـ}(\text{س} - \bar{\text{س}})^{\text{٢}} - \text{مجـ}\ \text{ن}\ (\text{م}_{\text{ر}} - \bar{\text{س}})^{\text{٢}} = \text{٢٩٨٫٠}$$

وهذه النتائج يمكن ترتيبها فى جدول كالآتى يوضح مصادر التباين ودرجات الحرية المقترنة بكل واحد منها ؛ وتقديرات التباين $\text{ى}_{\text{١}}^{\text{٢}}$ ، $\text{ى}_{\text{٢}}^{\text{٢}}$ ، $\text{ى}_{\text{٣}}^{\text{٢}}$ نحصل عليها بقسمة مجموع المربعات على عدد درجات الحرية .

جدول تحليل التباين إلى مركبتين

مصدر التباين	مجموع المربعات	عدد درجات الحرية	التباين التقديرى
التباين بين المجموعات	$\text{مجـ}\ \text{ن}\ (\text{م}_{\text{ر}} - \bar{\text{س}})^{\text{٢}} = \text{٢٢٩٠٫٤}$	$\text{حـ} - \text{١} = \text{٤}$	$\text{ى}_{\text{١}}^{\text{٢}} = \text{٥٧٢٫٦}$
« داخل «	$\text{مجـ}[\text{مجـ}(\text{س} - \text{م}_{\text{ر}})^{\text{٢}}] = \text{٢٩٨٫٠}$	$\text{ن} - \text{حـ} = \text{٣٠}$	$\text{ى}_{\text{٢}}^{\text{٢}} = \text{٩٫٩}$
« الكلى «	$\text{مجـ}\ (\text{س} - \bar{\text{س}})^{\text{٢}} = \text{٢٥٨٨٫٤}$	$\text{ن} - \text{١} = \text{٣٤}$	$\text{ى}_{\text{٣}}^{\text{٢}} = \text{٧٦٫١}$

$$\therefore\quad \text{ف} = \frac{\text{ى}_{\text{١}}^{\text{٢}}}{\text{ى}_{\text{٢}}^{\text{٢}}} = \text{٥٧٫٨٨}$$

وبالرجوع إلى جدول سنديكور عند درجات الحرية ٤ للتباين الأول و ٣٠ للثانى ، نجد أن

$$\text{ف}_{\text{ر٠٥}} = \text{٢ر٦٩} \quad ، \quad \text{ف}_{\text{ر٠١}} = \text{٤ر٠٢} .$$

ومعنى ذلك أن ف = ٥٧ر٩ التى حصلنا عليها فى هذه المسألة معنوية بدرجة عالية جداً — أعلى من ١ ٪ بكثير — بمعنى أننا لا نحصل على مثلها فى المعاينات العشوائية إلا أقل من مرة واحدة فى كل مائة مرة .

اختصار الحساب بطرح عدد مناسب من جميع المفردات

٤٥٩ — كان من الممكن تخفيف العمل الحسابى بتصغير الأعداد بطرح عدد معين مثل ٥٠ من جميع هذه الأوزان ، حيث تؤول هذه الأوزان إلى مفردات واقعة بين ٠ و ١٥ مما يجعل مربعاتها أعداداً صغيرة . وبعض المفردات بعد الطرح يكون سالباً والآخر موجباً ، مما يجعل المجاميع $\text{ل}_{\text{ر}}$ صغيرة أيضاً وكذلك مربعاتها . وطبعاً هذا العمل لا يؤثر مطلقاً فى النتيجة النهائية لأن التباينات تحسب من الانحرافات عن المتوسط ، وهذه الانحرافات لا تتأثر بطرح أو إضافة عدد معين لكل المفردات . ويمكن للقارىء التأكد من ذلك بحل هذه المسألة على سبيل التمرين بعد طرح العدد ٥٠ ، أو أى عدد آخر مناسب ، من كل المفردات .

المجموعات غير المتساوية فى عدد المفردات

الحساب هنا أطول قليلا

٤٦٠ — هذه الحالة نعالجها بنفس الطريقة التى شرحناها فى المثال السابق ، غير أن حساب التباين بين المجموعات سيكون أطول نوعاً ، لأن اختلاف عدد المفردات فى المجموعات يجعل $\text{ن}_{\text{ر}}$ مختلفة من مجموعة إلى أخرى ، فيكون

$$\text{مجـ}\ \text{ن}_{\text{ر}} \left(\text{م}_{\text{ر}} - \bar{\text{س}} \right)^{٢} = \text{مجـ}\ \frac{١}{\text{ن}_{\text{ر}}} \cdot \text{ل}_{\text{ر}}^{٢} - \frac{١}{\text{ن}} \left(\text{مجـ}\ \text{س} \right)^{٢} .$$

وعلى ذلك يتعين علينا حساب $\frac{١}{ن_ر}$. ل$_ر^٢$ لكل مجموعة على حدة ، بقسمة مربع مجموع المفردات فى هذه المجموعة على عدد مفرداتها ، بدلا من قسمة هذه المربعات على نفس العدد ن فى كل المجموعات . ونوضح الخطوات فى المثال الآتى [١] :

مثال عملى لحساب ف فى هذه الحالة

٤٦١ — فى تجربة زراعية أجراها م . ت . جنكينز و ج . و . سنديكور على ستة أنواع من الذرة كل منها يمثله عدد من السلالات ، كان محصول السلالات من كل نوع [٢] كما يأتى :

النوع الأول : ٧ر٣ ، ٥ر٤ ، ٧ر٤ ، ٧ر٤ ، ٥ر٠ ، ٥ر٩ ، ٦ر٤ ، ٦ر٣ ، ٥ر٠ ، ٦ر١ ، ٧ر٩ ، ٧ر٥ ؛

النوع الثانى : ٧ر٧ ، ٥ر٤ ، ٥ر٢ ، ٤ر٠ ؛

النوع الثالث : ٦ر٩ ، ٦ر٨ ، ٧ر٦ ، ٨ر١ ، ٩ر٤ ، ١٢ر٠ ، ١٥ر٩ ، ٧ر٤ ، ٩ر٠ ، ٥ر٢ ، ٩ر٢ ، ٨ر٦ ؛

النوع الرابع : ٩ر٦ ، ٧ر٨ ، ٩ر٦ ، ٧ر٧ ، ٨ر٢ ، ٧ر٣ ، ١١ر٣ ، ٩ر٥ ، ٨ر٨ ، ٨ر٤ ، ٦ر٨ ؛

النوع الخامس : ٤ر٨ ، ٩ر٢ ، ٨ر٥ ، ٨ر٨ ، ٧ر٩ ، ٥ر٩ ، ٩ر٢ ؛

النوع السادس : ٤ر٣ ، ٨ر٤ ، ٦ر٦ ، ٤ر٩ ، ٥ر٨ ، ٧ر٦ ، ٣ر٧ .

---

(١) أنظر تمرين رقم ١٠ر١٥ فى كتاب
G. W. Snedecor *Statistical Methods*, (1946), P. 235

(٢) أسماء هذه الأنواع بالإنجليزية هى على الترتيب المذكور من اليسار لليمين
Four County, Silver King, Iodent, Lancaster, Osterland, Clark
والمحصول مقدر بوحدات Buhsels per Acre

والمطلوب حساب قيمة ف (= ٥٫١٩ أى ١٧٫٥٥ ÷ ٣٫٣٨) .

لتسهيل العمل الحسابى نضرب كل هذه الأرقام فى العدد ١٠ (كأنما نقيس المحصول بوحدات جديدة تساوى عشر الوحدات الأصلية) ، ونطرح العدد ٧٠ من جميع هذه الأرقام ، ونرتب العمل فى جدول كالآتى :

رقم السلالة	الأنواع						كل العينة
	١	٢	٣	٤	٥	٦	
١	٣	٧	-١	٢٦	-٢٢	-٢٧	
٢	-٢٥	-١٦	-٢	٨	٢٢	١٤	
٣	٤	-١٨	٦	٢٦	١٥	-٤	
٤	٤	-٣٠	١١	٧	١٨	-٢١	
٥	-٢٠		٢٤	١٢	٩	-١٢	
٦	-١١		٥٠	٣	-١١	٦	
٧	-٦		٨٩	٤٣	٢٢	-٣٣	
٨	-٧		٤	٢٥			
٩	-٢٠		٢٠	١٨			
١٠	-٩		-١٨	١٤			
١١	٩		٢٢	-٢			
١٢	-١٣		١٦				
(١) مجموع المفردات ل	-٩١	-٥٧	٢٢١	١٨٠	٥٣	-٧٧	٢٢٩ = مج س
(٢) عدد « ن	١٢	٤	١٢	١١	٧	٧	٥٣ = ن
(٣) مربع المجموع ل^٢	٨٢٨١	٣٢٤٩	٤٨٨٤١	٣٢٤٠٠	٢٨٠٩	٥٩٢٩	٥٢٤٤١ = (مج س)^٢
(٤) = (٣) ÷ (٢)	٦٩٠٫٠٨	٨١٢٫٢٥	٤٠٧٠٫٠٨	٢٩٤٥٫٤٥	٤٠١٫٢٩	٨٤٧	٩٨٩٫٥٠ = ١/ن (مج س)^٢
(٥) مجموع مربعات المفردات	٢٠٠٣	١٥٢٩	١٢٦٣٩	٤٦١٦	٢٢٠٣	٢٦٥١	٢٥٦٤١ = مج س^٢

مج ١/ن ل^٢ = ٩٧٦٦٫١٥

∴ مج (س — س̄)^٢ = مج س^٢ — ١/ن (مج س)^٢

= ٢٥٦٤١ — ٩٨٩٫٥٠ = ٢٤٦٥١٫٥٠

و مجـ $ن_ر$ ($م_ر$ — $\bar{س}$)^{٢} = مجـ $\frac{١}{ن_ر}$ $ل_ر^{٢}$ — $\frac{١}{ن}$ (مجـ س)^{٢}

= ٩٧٦٦,١٥ — ٩٨٩,٥ = ٨٧٧٦,٦٥

∴ مجـ [مجـ (س — $م_ر$)^{٢}] = ٢٤٦٥١,٥٠ — ٨٧٧٦,٦٥ = ١٥٨٧٤,٨٥

و بقسمة كل مجموع للمربعات على عدد درجات الحرية الخاصة به نحصل على مركبات التباين $ى_١^{٢}$ ، $ى_٢^{٢}$ ، $ى_٣^{٢}$ وهى مبينة فى الجدول الآتى :

جدول تحليل التبـاين حسب مصادره

مصدر التبـاين (١)	مجموع المربعـــات (٢)	عدد درجات الحرية (٣)	التبـاين التقديرى (٢) ÷ (٣)
التباين بين المجموعات	مجـ $ن_ر$ ($م_ر$ — $\bar{س}$)^{٢} = ٨٧٧٦,٦٥	حـ — ١ = ٥	$ى_١^{٢}$ = ١٧٥٥,٣٣
« داخل «	مجـ [مجـ (س — $م_ر$)^{٢}] = ١٥٨٧٤,٨٥	ن — حـ = ٤٧	$ى_٢^{٢}$ = ٣٣٧,٧٦
التباين الـكلى	مجـ (س — $\bar{س}$)^{٢} = ٢٤٦٥١,٥٠	ن — ١ = ٥٢	$ى_٣^{٢}$ = ٤٧٤,٠٧

∴ ف = $\frac{ى_١^{٢}}{ى_١^{٢}}$ = $\frac{١٧٥٥,٣٣}{٣٣٧,٧٦}$

= ٥,١٩ .

وبالرجوع إلى جداول سنديكور عنـد درجات الحرية $د_١$ = ٥ و $د_٢$ = ٤٧، وملاحظـة أن الجدول يحتوى فقط على $د_٢$ = ٤٦ و ٤٨ ولا يحتوى على قيم ف عندما $د_٢$ = ٤٧ بالذات ، نرى أن

$ف_{٠,٠٥}$ تقع بين ٢,٤٢ و ٢,٤١

و $ف_{٠,٠١}$ « « ٣,٤٤ و ٣,٤٢

ويتضح من ذلك أن ف = ٥,١٩ التى حصلنا عليها فى هـذه المسألة معنوية بدرجة أعلى من ٠,٠١ وهذا دليل واضح على أن هناك اختلاف حقيقى بين هذه الأنواع الستة من الذرة فى المحصول .

يلاحظ أنه لو كنا أجرينا الحساب على أساس الوحدات الأصلية المعطاة فى المسألة كنا نحصل على ١٧,٥٥٣٣ لتقدير التباين $ى_١^٢$ ، وعلى ٣,٣٧٧٦ للتقدير $ى_٢^٢$ بدلا من ١٧٥٥,٣٣ و ٣٣٧,٧٦ على الترتيب ؛ وذلك لأن الوحدات الجديدة عُشْر الوحدات الأصلية ، فيكون مقـدار التباين مقاساً بها يساوى مقداره مقاساً بالوحدات الأصلية مضروبا فى ١٠٠ . ولـكن هـذا على كل حال لا يؤثر فى قيمة النسبة ف .

نصل إلى نفس النتيجـــة باستخدام ظ

٤٦٢ — ومن الممكن استخدام ظ بدلا من ف ونصل إلى نفس النتيجة طبعاً باستخدام جداول ر . ا. فيشر لقيم ظ ، لأن

$$ظ = \frac{١}{٢} لو_هـ \frac{ى_١^٢}{ى_٢^٢}$$ ، وكما قلنا فى بند ٤٤٨

$$= ١,١٥١٣ \; لو_{١٠} \; ٥,١٩$$

$$= ٠,٨٢٢٥$$

وبالرجوع إلى جداول فيشر لقيم ظ مع $ن_١ = ٥$ و $ن_٢ = ٤٧$ ، نجـــد أن

	$ظ_{٠,٠٥}$	واقعة بين	٠,٤٦٤٨	و ٠,٤٣١١
و	$ظ_{٠,٠١}$	« «	٠,٦٥٤٠	و ٠,٦٠٢٨

وعلى ذلك فالقيمة ظ = ٠,٨٢٢٥ معنوية بدرجة أعلى من ٠,٠١ وهى نفس

النتيجة التى وصلنا إليها عن طريق ف . [و يلاحظ أن جدول فيشر عند مستوى المعنوية ٠٫٠٠١ يعطى لدرجات الحرية ٥و٤٧ قيمة ظ٫٠٠١ واقعة بين ٠٫٨١٧٤ و ٠٫٧٧٩٨ مما يدل على أن ظ = ٠٫٨٢٢٥ معنوية بدرجة عالية جداً و يعزز الحكم بأن هذه الأنواع من الذرة تختلف عن بعضها البعض فى المحصول] .

٤٦٣ — ولا يفوتنا التنبيه إلى أن الأساس المبنى عليه هذا الجدال الذى شرحناه فى تحليل التباين هو أن المجموعات الجزئية التى نقارنها عبارة عن عينات عشوائية من مجتمع معتدل ، ولا بد من الاحتياط من هذه الناحية ؛ على أن توزيع قيم ف يتأثر قليلا بما قد يكون فى هذه المجتمعات من التواء بسيط . وليس فى وسعنا الآن أن نخوض فى تفاصيل هذا الموضوع فى هذا الكتاب ، ولا يمكننا أن نلم فيه بجميع التطبيقات والمباحث التى نستخدم فيها هذه الأداة الإحصائية المفيدة التى سميناها[1] تحليل التباين . و إنما نكتفى هنا بهذه المسائل البسيطة والتطبيقات السهلة التى أوضحناها بالأمثلة المتقدمة ؛ وعلى من يريد التعمق فى هذا الموضوع والإلمام بكل تطبيقاته ودقائقه أن يرجع إلى كتب أخرى أكثر تخصصاً وأدق تفصيلا ، مثل كتابى ر . ا فيشر و ج . و . سنديكور وغيرهما . وفى العادة يفضل استخدام جدول سنديكور لقيم نسبة التباين ف على المستويين ٠٫٠١ ، ٠٫٠٥ من مستويات المعنوية ، وهى مبوبة تبعاً لدرجات الحرية $ن_١$ و $ن_٢$ كما ذكرنا . ومن الواضح أن العمل أسهل فى حساب ف منه فى حساب ظ . و إذا أراد القارىء اختبار معنوية نسبة التباين على مستوى ٠٫٠٠١ فلا بد من حساب ظ والرجوع إلى جدول فيشر لقيم ظ على هذا المستوى العالى .

---

(١) بالانجليزية Analysis of Variance

المــــراجع

Barlow's : Tables of Squares. Square Roots, etc.

Fisher, R. A. : *Statistical Methods for Research Workers*, Chapters VII, VIII.

Pearson and Bennet : *Statistical Methods*, Chapter XIX.

Peter and Van Voorhis : *Statistical Procedures etc.*, Chapter XII.

Snedecor, G. W. : *Statistical Methods*, Chapters X, XI.

Yule and Kendall : *Introduction to the Theory of Statistics*, Chapter XIX.

الباب الثامن عشر

الاستكمال وتمهيد البيانات الإحصائية

إنشاء الجداول الرياضية

٤٦٤ — الجداول الرياضية والإحصائية تعطينا فى العادة سلسلتين من القيم المتناظرة لكميتين متغيرتين : إحداهما نسميها المتغير المستقل ، ويرمز لها عادة بالحرف س ، والأخرى نسميها المتغير التابع ، ويرمز لها بالحرف ص . والغرض من إنشاء هذه الجداول هو توفير العمليات الحسابية اللازم إجراؤها عندما نعرف قيمة المتغير س ونريد معرفة قيمة المتغير التابع ص . فبواسطة جدول اللوغاريتمات مثلاً يمكننا فى أى وقت معرفة لوغاريتم أى عدد بدقة و بسرعة .

الجداول تشتمل على بعض قيم س فقط

٤٦٥ — ولكن الجداول العادية لا تعطى قيم الدالة ص المناظرة إلى جميع قيم س الممكن تصورها ، لأن هذا مستحيل عملياً . و إنما يقتصر فى هذه الجداول على سلسلة مختصرة من قيم س حتى يسهل إخراجها والانتفاع بها . فنجد مثلاً جداول اللوغاريتمات العادية تعطى لوغاريتمات الأعداد الطبيعية من ١ إلى ١٠٠٠٠ مثلاً ، وجداول الفائدة المركبة تعطى جملة الجنيه بسعر معين لمدد تتغير من ١ إلى ١٠٠ سنة مثلاً ، أو تعطى جملة الجنيه فى مدة معينة بأسعار مختلفة مثل ١ ٪ و $\frac{1}{2}$١ ٪ و ٢ ٪ و ... إلى ٩ ٪ مثلاً .

استكمال الجداول لإعطاء قيم متوسطة

٤٦٦ — ومن البديهى أننا فى مسائلنا العملية لا تتقيد ظروفنا بهذه القيم فقد نريد مثلاً جملة الجنيه بسعر ٤ ٪ لمدة ١٢٫٢٥ من السنين فلا نجدها

فى الجدول ؛ و إنما نجد جملة الجنيه بالسعر المطلوب لمدة ١٢ سنة ، وهى ١٫٦٠١٠٣ ، ولمدة ١٣ سنة وهى ١٫٦٦٥٠٧ . كما أنه يصح أن يطلب معرفة جملة الجنيه لمدة عشر سنوات مثلاً بسعر ٤ $\frac{١}{٤}$ ٪ فلا نجدها فى الجدول ، حيث لا نجد الجملة إلا بسعر ٤ ٪ وتساوى ١٫٤٨٠٢٤ ، والجملة بسعر ٤ $\frac{١}{٢}$ ٪ ، وهى ١٫٥٥٢٩٧ .

أمام هذه الصعوبة نريد أن نستكمل الجدول لكى يعطينا القيمة المطلوبة . وسنذكر هنا بعض الطرق البسيطة المستعملة فى الاستكمال ونشرح باختصار الأسس النظرية التى تنبنى عليها هذه الطرق . ويجب على القارىء الذى يريد استقصاء هذا الموضوع أن يرجع إلى بعض الكتب (١) الخاصة به .

وسنبدأ بشرح الطرق المستعملة فى حالة ما تكون قيم س على فترات متساوية من بعضها .

طريقة التقسيم التناسبى للفروق

٤٦٧ — أبسط طريقة للاستكمال هى أن نقول فى مثال جملة الجنيه المطلوبة لمدة ١٢٫٢٥ سنة بسعر ٤ ٪ ، إن هذه الجملة لا بد أن تكون محصورة بين جملة الجنيه لمدة ١٢ سنة ، أى ١٫٦٠١٠٣ ، وجملته لمدة ١٣ سنة وهى ١٫٦٦٥٠٧ . ولتحديد موقعها بالضبط فى هذه الفترة نجرى عملية التناسب البسيطة الآتية :

بفرض أن الجملة المطلوبة هى ص ، يكون

$$\frac{\text{ص} - ١٫٦٠١٠٣}{١٫٦٦٥٠٧ - ١٫٦٠١٠٣} = \frac{١٢٫٢٥ - ١٢}{١٣ - ١٢}$$

فينتج أن $\qquad$ ص = ١٫٦١٧٠٤

وهى فعلا واقعة بين ١٫٦٠١٠٣ و ١٫٦٦٥٠٧ .

---

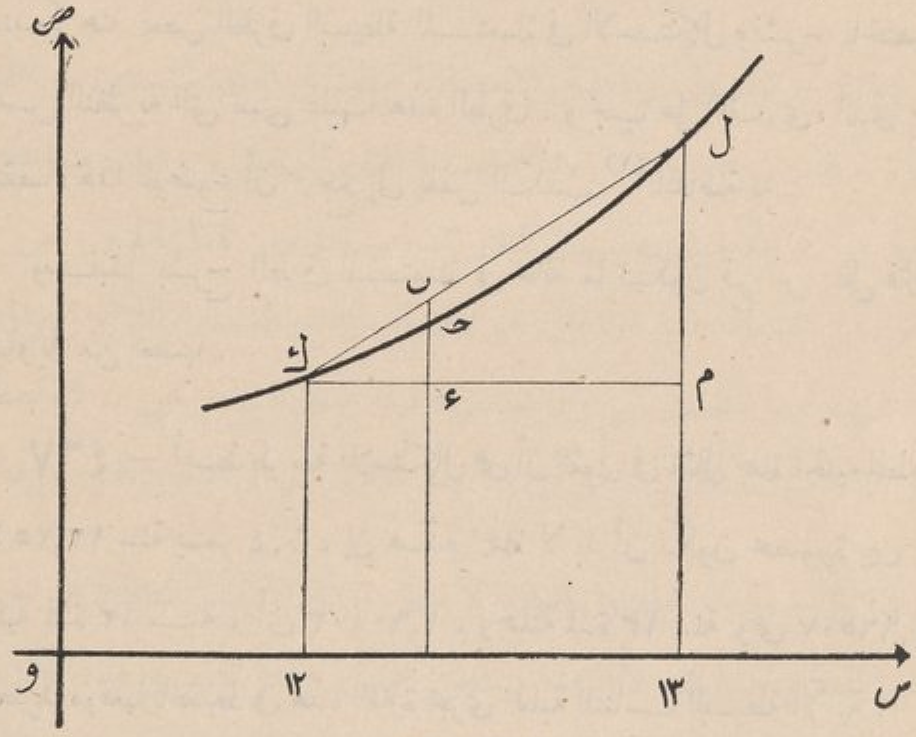
(١) أنظر كتاب Boole : *Finite Differences*

أو كتاب Whittaker and Robinson : *Calculus of Observations*, (1929) pp. 1—70

وهذا الموضوع يسمى بالإنجليزية Interpolation

هذه الطريقة تفرض أن ص دالة مستقيمة

٤٦٨ — ولكن هذه الطريقة تقريبية جداً لأنها تفرض ضمناً أن الكمية ص تتغير تغيراً منتظماً أو بسرعة منتظمة، أثناء هذه الفترة ما بين س = ١٢ سنة و س = ١٣ سنة، أى أنها تفرض أن الخط البيانى بين ص و س فى هذه الفترة مستقيم. حيث إن هذا التقسيم التناسبي الذى ذكرناه لا يكون صحيحاً إلا فى حالة الخط المستقيم كما يتضح من الشكل الآتى :



(شكل ٦٩)

فالقيمة التى نحصل عليها بهذه الطريقة تساوى البعد ا ب فى الشكل[1]، فى حين أن القيمة المطلوبة تساوى البعد ا حـ. على أن هذه الطريقة لا بأس بها إذا كان الخط البيانى قريباً من الاستقامة فى هذه الفترة، أى بين النقطتين ك و ل ؛ أو إذا كانت الفترة نفسها صغيرة بحيث يمكن اعتبار القوس ك ل كأنه خط مستقيم.

(١) النقطة ا لم تظهر فى الرسم وهى مسقط النقطة ب على المحور الأفقى : وهى ملتقى العمود ب حـ ء بالمحور الأفقى بين ١٢ و ١٣.

المدة بالسنين س	جملة الجنيه بسعر ٤ ٪ ص	الفروق △ ص	الفروق الثانية $\triangle^{2}$ ص
٩	١٫٤٢٣٣١		
		٠٫٠٥٦٩٣	
١٠	١٫٤٨٠٢٤		٠٫٠٠٢٢٨
		٠٫٠٥٩٢١	
١١	١٫٥٣٩٤٥		٠٫٠٠٢٣٧
		٠٫٠٦١٥٨	
١٢	١٫٦٠١٠٣		٠٫٠٠٢٤٦
		٠٫٠٦٤٠٤	
١٣	١٫٦٦٥٠٧		٠٫٠٠٢٥٧
		٠٫٠٦٦٦١	
١٤	١٫٧٣١٦٨		٠٫٠٠٢٦٥
		٠٫٠٦٩٢٦	
١٥	١٫٨٠٠٩٤		

جملة الجنيه لا تتغير بسرعة منتظمة

٤٦٩ — ولو بحثنا فى جدول الفوائد المركبة لوجدنا أن الدالة ص ، وهى جملة الجنيه ، لا تتغير « بسرعة منتظمة » كما تفترض هـذه الطريقة ، لأن الفروق بين قيم ص المتتالية فى الجدول ليست متساوية كما هو واضح من الجدول أعلاه ، مما يدل على أن ص لا تزيد بمعدل ثابت . حتى هـذه الفروق نفسها لا تزيد بمعدل ثابت ، كما يتبين عنـد حساب فروق الفروق التى نجـدها فى العمود الأخير من الجدول السابق . على أن هـذه الفروق الأخيرة أكثر انتظاماً فى تغيرها ، بدليل أننا لو حسبنا الفروق الثالثة وجدناها كلها متساوية تقريباً .

إنشاء جـدول الفـروق

٤٧٠ — وتمهيداً لعملية الاستكمال ننشىء جدول الفروق للدالة التى نريد استكمالها ، وذلك بأن نحسب الفروق بين قيمها المتتالية ، فنطرح كل قيمة من التى تليها طرحاً جبرياً ، وهذه الفروق نسميها **الفروق الأولى**[1] للدالة ص ، ونرمز

(١) بالإنجليزية First Differences

لها بالرمز $\triangle$ ص (دِلْتا ص) . ثم نحسب فروق هذه الفروق بطرح كل فرق من الذى يليه مباشرة طرحاً جبرياً أيضاً . وهذه تسمى **الفروق الثانية**[١]، ونرمز لها بالرمز $\triangle^{2}$ ص (دلتا تربيع ص) . ونستمر فى حساب الفروق حتى نحصل على فروق متساوية — بالضبط أو بالتقريب — فتكون الفروق التى بعدها تساوى صفرا — أو قريبة من الصفر . وبذلك نكون ما يسمى **جدول الفروق**[٢] .

لنأخذ مثلا الجدول الآتى :

س	ص	$\triangle$ ص	$\triangle^{2}$ ص	$\triangle^{3}$ ص	$\triangle^{4}$ ص	$\triangle^{5}$ ص
١	١١					
		١٥				
٢	٢٦		٥٠			
		٦٥		٦٠		
٣	٩١		١١٠		٢٤	
		١٧٥		٨٤		٠
٤	٢٦٦		١٩٤		٢٤	
		٣٦٩		١١٨		٠
٥	٦٣٥		٣٠٢		٢٤	
		٦٧١		١٣٢		٠
٦	١٣٠٦		٤٣٤		٢٤	
		١١٠٥		١٥٦		٠
٧	٢٤١١		٥٩٠		٢٤	
		١٦٩٥		١٨٠		
٨	٤١٠٦		٧٧٠			
		٢٤٦٥				
٩	٦٥٧١					

(١) بالإنجليزية Second Differences

(٢) » Difference Table

الفروق $\triangle$ ص ورتبها المختلفة

٤٧١ — وعلى العموم إذا كانت قيم س المعروفة فى الجدول ، مرتبة تصاعدياً هى :

$$س_{.} ، س_{١} ، س_{٢} ، س_{٣} ، \ldots\ldots ، س_{ن}$$

وكانت قيم ص المناظرة لها فى الجدول هى على الترتيب

$$ص_{.} ، ص_{١} ، ص_{٢} ، ص_{٣} ، \ldots\ldots ، ص_{ن}$$

نقول إن الفروق الأولى لقيم ص المتتالية هى :

$$ص_{١} - ص_{.} = \triangle ص_{.}$$
$$ص_{٢} - ص_{١} = \triangle ص_{١}$$
$$ص_{٣} - ص_{٢} = \triangle ص_{٢}$$

وهكذا . وتكون الفروق الثانية ، أو فروق الفروق ، هى :

$$\triangle ص_{١} - \triangle ص_{.} = \triangle^{٢} ص_{.}$$
$$\triangle ص_{٢} - \triangle ص_{١} = \triangle^{٢} ص_{١}$$
$$\triangle ص_{٣} - \triangle ص_{٢} = \triangle^{٢} ص_{٢}$$

وبالمثل تكون الفروق الثالثة هى :

$$\triangle^{٢} ص_{١} - \triangle^{٢} ص_{.} = \triangle^{٣} ص_{.}$$
$$\triangle^{٢} ص_{٢} - \triangle^{٢} ص_{١} = \triangle^{٣} ص_{١}$$
$$\triangle^{٢} ص_{٣} - \triangle^{٢} ص_{٢} = \triangle^{٣} ص_{٢}$$

وهكذا نحسب الفروق من أى رتبة .

التعبير الرمزى

٤٧٢ — ويمكننا التعبير عن هذه الفروق المتتالية $\triangle$ ص بدلالة قيم ص المتتالية بطريقة رمزية مختصرة . وذلك كما يأتى :

$$\triangle ص_{.} = ص_{١} - ص_{.}$$

$$\Delta^{٢}\text{ص}_{٠} = \Delta\,\text{ص}_{١} - \Delta\,\text{ص}_{٠}$$

$$= \text{ص}_{٢} - ٢\,\text{ص}_{١} - \text{ص}_{٠}$$

= مفكوك $(\text{ص} - ١)^{٢}$ بعد وضع $\text{ص}_{٢}$ بدل ص
و $\text{ص}_{٠}$ بدل ١

$$\Delta^{٣}\text{ص}_{٠} = \Delta^{٢}\text{ص}_{١} - \Delta\,\text{ص}_{٠}$$

$$= \Delta\,\text{ص}_{٢} - \Delta\,\text{ص}_{١} - (\Delta\,\text{ص}_{١} - \Delta\,\text{ص}_{٠})$$

$$= \text{ص}_{٣} - ٣\,\text{ص}_{٢} + ٣\,\text{ص}_{١} - \text{ص}_{٠}$$

= مفكوك $(\text{ص} - ١)^{٣}$

وهكذا لأى درجة من الفروق . ونجد على العموم أن

$\Delta^{\text{ن}}\text{ص}_{٠}$ = مفكوك $(\text{ص} - ١)^{\text{ن}}$ بوضع $\text{ص}_{\text{ر}}$ بدل $\text{ص}^{\text{ر}}$ ، و $\text{ص}_{٠}$ بدل ١ .

ويمكننا قلب هـذه العملية فنعبر عن قيم ص المتتالية بدلالة فروق إحدى القيم مثل $\text{ص}_{٠}$ وذلك كما يأتى :

$\text{ص}_{١} = \text{ص}_{٠} + \Delta\,\text{ص}_{٠}$ ، و بوضع مختصر نقول :

$$= (١ + \Delta)\,\text{ص}_{٠}$$

$$\text{ص}_{٢} = \text{ص}_{١} + \Delta\,\text{ص}_{١}$$

$$= \text{ص}_{٠} + \text{ص}_{٠} + \Delta(\text{ص}_{٠} + \Delta\,\text{ص}_{٠})$$

$$= \text{ص}_{٠} + ٢\,\Delta\,\text{ص}_{٠} + \Delta^{٢}\text{ص}_{٠}$$

$$= (١ + \Delta)^{٢}\,\text{ص}_{٠}$$

$$\text{ص}_{٣} = \text{ص}_{٢} + \Delta\,\text{ص}_{٢}$$

$$= \text{ص}_{٠} + ٢\,\Delta\,\text{ص}_{٠} + \Delta^{٢}\,\text{ص}_{٠}$$

$$+ \Delta\,(\text{ص}_{٠} + ٢\,\Delta\,\text{ص}_{٠} + \Delta^{٢}\,\text{ص}_{٠})$$

$$= \text{ص}_{٠} + ٣\,\Delta\,\text{ص}_{٠} + ٣\,\Delta^{٢}\,\text{ص}_{٠} + \Delta^{٣}\,\text{ص}_{٠}$$

$$= (١ + \Delta)^{٣}\,\text{ص}_{٠}$$

وهكذا حتى نصل إلى أى قيمة صادية نريدها ، وعلى العموم

$$\text{ص}_{\text{ر}} = (١ + \Delta)^{\text{ر}}\,\text{ص}_{٠} \quad \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots \quad (١)$$

وليس معنى ذلك أن الطرف الأيسر عبارة عن الكمية $(١ + \Delta)^{\text{ر}}$ مضروبة فى الكمية $\text{ص}_{٠}$. ولكن المقصود هو وضع رمزى مختصر معناه أننا نفك القوس $(١ + \Delta)^{\text{ر}}$ طبقاً لنظرية ذات الحدين المعروفة ، ثم نلصق $\text{ص}_{٠}$ بكل واحد من حدود هذا المفكوك . لأن $(١ + \Delta)^{\text{ر}}$ وحدها ليس لها معنى ولا أى واحد من حدود مفكوك هذا القوس . فالحد الثالث مثلا هو $\frac{\text{ر}(\text{ر} - ١)}{٢} . \Delta^{٢}$ لا معنى له بمفرده . ولكن إذا ألصقنا به $\text{ص}_{٠}$ فأصبح $\frac{\text{ر}(\text{ر} - ١)}{٢} . \Delta^{٢}\,\text{ص}_{٠}$ ، يتضح معناه بعد هذه الإضافة (المضاف والمضاف إليه فى النحو) ، وهو أننا نضرب الكمية $\frac{\text{ر}(\text{ر} - ١)}{٢}$ فى الكمية $\Delta^{٢}\,\text{ص}_{٠}$ التى هى الفرق الثانى للقيمة $\text{ص}_{٠}$. ولكن Δ أو $\Delta^{٢}$ أو . . لا معنى لها بمفردها

وهذه المعادلة تسمى قانون جريجورى ونيوتن للاستكمال [١] .

٤٧٣ — يمكننا استخدام المعادلة $\text{ص}_{\text{ر}} = (١ + \Delta)^{\text{ر}}\,\text{ص}_{٠}$ لمعرفة قيم للمتغير ص غير موجودة فى الجدول الذى عندنا . فلو أخذنا الجدول المذكور فى بند ٤٧١ مثلا ، واعتبرنا $\text{س}_{٠} = ١$ وهى أول قيمة سينية يحتوى عليها الجدول ، تكون $\text{س}_{٨} = ٩$ ، وبالمثل تكون $\text{ص}_{٠} = ١١$ و $\text{ص}_{٨} = ٦٥٧١$.

تطبيق هذه المعادلة فى تكميل الجدول

---

(١) بالإنجليزية Gregory-Newton Formula

والجدول لا يعطينا قيماً صادية ولا سينية أعلى من هاتين . ولكن بواسطة العلاقة $\text{ص}_{\text{ر}} = (\text{١} + \triangle)^{\text{ر}}\ \text{ص}_{\text{٠}}$ وجدول الفروق يمكننا معرفة $\text{ص}_{\text{٩}}$ و $\text{ص}_{\text{١٠}}$ و $\text{ص}_{\text{١١}}$ و . . .

$$\text{ص}_{\text{٩}} = (\text{١} + \triangle)^{\text{٩}}\ \text{ص}_{\text{٠}}$$

$$= \Big[\text{١} + \text{٩}\triangle + \frac{\text{٩}(\text{٩}-\text{١})}{\text{٢}}\triangle^{\text{٢}} + \frac{\text{٩}\times\text{٨}\times\text{٧}}{\text{١}\times\text{٢}\times\text{٣}}\triangle^{\text{٣}} + \frac{\text{٩}\times\text{٨}\times\text{٧}\times\text{٦}}{\text{١}\times\text{٢}\times\text{٣}\times\text{٤}}\triangle^{\text{٤}}$$

$$+ \frac{\text{٩}\times\text{٨}\times\text{٧}\times\text{٦}\times\text{٥}}{\text{١}\times\text{٢}\times\text{٣}\times\text{٤}\times\text{٥}}\triangle^{\text{٥}} + \ldots\Big]\ \text{ص}_{\text{٠}}$$

$$= \text{ص}_{\text{٠}} + \text{٩}\triangle\text{ص}_{\text{٠}} + \text{٣٦}\triangle^{\text{٢}}\text{ص}_{\text{٠}} + \text{٨٤}\triangle^{\text{٣}}\text{ص}_{\text{٠}} + \text{١٢٦}\triangle^{\text{٤}}\text{ص}_{\text{٠}} + .$$

ونقف فى المفكوك عند الحد الذى يحتوى على $\triangle^{\text{٤}}$ ، لأننا نعلم من جدول الفروق أن الفروق الخامسة وما بعدها تساوى صفراً . وبالتعويض عن فروق $\text{ص}_{\text{٠}}$ من الجدول ، وهى فى أوائل الأعمدة المتتالية ،

$$\therefore\ \text{ص}_{\text{٩}} = \text{١١} + \text{٩}\times\text{١٥} + \text{٣٦}\times\text{٥٠} + \text{٨٤}\times\text{٦٠} + \text{١٢٦}\times\text{٢٤}$$

$$= \text{١٠٠١٠}$$

وهكذا أمكننا تكميل الجدول خطوة ، وبالمثل يمكننا حساب $\text{ص}_{\text{١٠}}$ وغيرها .

تطبيق القانون فى حالة قيم كسرية

٤٧٤ — ولا يقتصر تطبيق هذه المعادلة على الحالات التى تكون فيها ر عدداً صحيحاً موجباً ، مثل ٩ كما فى المثال السابق ، بل هى تنطبق أيضاً فى الحالات الأخرى[١]. فنعلم أن نظرية ذات الحدين تنطبق[٢] فى حالات الأسس الكسرية والسالبة ويمكن بواسطتها إيجاد مفكوك مثل $(\text{١} + \text{س})^{\text{ن}}$ حينما يكون الأس ن عدداً كسرياً أو سالباً . والمفكوك حينئذ لا ينتهى وهو :

$$(\text{١}+\text{س})^{\text{ن}} = \text{١} + \text{ن س} + \frac{\text{ن}(\text{ن}-\text{١})}{\text{٢}}\text{س}^{\text{٢}} + \frac{\text{ن}(\text{ن}-\text{١})(\text{ن}-\text{٢})}{\text{١}\times\text{٢}\times\text{٣}}\text{س}^{\text{٣}} + \ldots \text{إلى } \infty$$

---

(١) لإثبات هذا الانطباق راجع الملحق الرياضى فى آخر الكتاب .

(٢) أنظر كتاب Hall and Knight : *Higher Algebra*

ويمكننا إثبات [١] أن قانون نيوتن فى الاستكمال

$$\text{ص}_{\text{ر}} = (١ + \triangle)^{\text{ر}}\ \text{ص}_{٠}$$

$$= \text{ص}_{٠} + \text{ر}\triangle\,\text{ص}_{٠} + \frac{\text{ر}(\text{ر}-١)}{٢}\triangle^{٢}\,\text{ص}_{٠}$$

$$+ \frac{\text{ر}(\text{ر}-١)(\text{ر}-٢)}{١\times٢\times٣}\triangle^{٣}\,\text{ص}_{٠} + \ldots \text{ إلى } \infty$$

ينطبق أيضاً فى حالة ما تكون ر عدداً كسرياً أو سالباً . وفى هذه الحالة يكون معنى $\text{ص}_{\text{ر}}$ ، مثلاً $\text{ص}_{٢٫٣}$ ، هو تلك القيمة الصادية الواقعة بين $\text{ص}_{٢}$ و $\text{ص}_{٣}$ والتى تناظر قيمة س الواقعة بين $\text{س}_{٢}$ و $\text{س}_{٣}$ التى نرمز لها بالرمز $\text{س}_{٢٫٣}$ ، بحيث يكون

$$\text{س}_{٢٫٣} - \text{س}_{٢} = ٠٫٣ \times (\text{س}_{١} - \text{س}_{٢})\ ،$$

أى أن $\text{س}_{٢٫٣}$ تزيد عن $\text{س}_{٢}$ بمقدار ٠٫٣ من الفترة .

وبالمثل إذا كانت ر سالبة ، فمثلاً $\text{ص}_{-٠٫٥}$ هى قيمة س التى تناظر $\text{س}_{-٠٫٥}$ ، وهى القيمة السينية التى تسبق $\text{س}_{٠}$ بنصف فترة .

وبهذا المعنى الجديد يمكننا استخدام قانون نيوتن

$$\text{ص}_{\text{ر}} = (١ + \triangle)^{\text{ر}}\ \text{ص}_{٠}$$

على نطاق أوسع . فنستكمل الجدول فى أى موضع ، ونوجد قيمة الدالة ص المناظرة إلى أى قيمة نريدها ، سواء بعد نهاية الجدول أو قبل بدايته أو بين سطوره .

٤٧٥ — فى الجدول المذكور فى بند ٤٧٠ المطلوب إيجاد قيمة ص حينما س تساوى ٣٫٤ مثال ر عدد كسرى موجب

---

(١) أنظر كتاب Whittaker and Robinson: *Calculus of Observations*, (1929), Art. 8, p. 10. أو راجع الملحق الرياضى فى آخر الكتاب

نعتبر $\text{س}_{.} = ١$ فى أول الجدول و $\text{ص}_{.} = ١١$

نفرض أن القيمة السينية المطلوبة هى $\text{س}_{\text{ر}} = ٣ر٤$

$$\therefore\ \text{س}_{\text{ر}} - \text{س}_{.} = ٣ر٤ - ١$$

$= ٢ر٤$ من الفترات السينية

$$\therefore\ \text{س}_{\text{ر}}\ \text{هى}\ \text{س}_{٢ر٤}$$

وبناءً عليه تكون القيمة الصادية المطلوبة هى $\text{ص}_{\text{ر}} = \text{ص}_{٢ر٤}$

$$\therefore\ \text{ص}_{٢ر٤} = (١ + \triangle)^{٢ر٤}\ \text{ص}_{.}$$

$$= \Big[\, ١ + ٢ر٤\,\triangle + \frac{٢ر٤ \times ١ر٤}{٢}\triangle^{٢} + \frac{٢ر٤ \times ١ر٤ \times ر٤}{٢ \times ٣}\triangle^{٣}$$
$$+ \frac{٢ر٤ \times ١ر٤ \times ر٤ \times -ر٦}{٢ \times ٣ \times ٤}\triangle^{٤} \,.\Big]\ \text{ص}_{.}$$

$$= \text{ص}_{.} + ٢ر٤\,\triangle\,\text{ص}_{.} + ١ر٦٨\,\triangle^{٢}\,\text{ص}_{.} + ر٢٢٤\,\triangle^{٣}\,\text{ص}_{.}$$
$$- ر٠٣٣٦\,\triangle^{٤}\,\text{ص}_{.}$$

$$= ١١ + ٢ر٤ \times ١٥ + ١ر٦٨ \times ٥٠ + ر٢٢٤ \times ٦٠$$
$$- ر٠٣٣٦ \times ٢٤$$

$$= ١٤٣ر٦٣٣٦$$

٤٧٦ — فى هذا المثـال اعتبرنا أول قيمة فى الجدول هى $\text{س}_{.}$. ويمكننا اعتبار أى قيمة أخرى فى الجدول هى $\text{س}_{.}$ ، واعتبار القيمة الصادية المناظرة لها $\text{ص}_{.}$. حل المثـال السابق بأخذ ر سالبة

لنأخذ الآن $\text{س}_{.} = ٤$ مثلاً ، و $\text{ص}_{.} = ٢٦٦$ ، ولتكن القيمة السينية المطلوبة هى ٣ر٤ أيضاً ونرمز لهـا بالرمز $\text{س}_{\text{ر}}$.

$$\therefore\ \text{س}_{\text{ر}} - \text{س}_{.} = ٣ر٤ - ٤$$

$= -\,ر٦$ من الفترة السينية

$\therefore$ $س_ر$ هى $س_{-٦ر}$. وتكون $ص_ر$ هى $ص_{-٦ر}$.

$$\therefore \quad ص_{-٦ر} = (١ + \triangle)^{-٦ر} ص_٠$$

$$= \left[١ - ٦ر \triangle + \frac{-٦ر \times -١ر٦}{٢} \triangle^{٢} + \frac{-٦ر \times -١ر٦ \times -٢ر٦}{١ \times ٢ \times ٣} \triangle^{٣} \right.$$

$$\left. + \frac{-٦ر \times -١ر٦ \times -٢ر٦ \times -٣ر٦}{١ \times ٢ \times ٣ \times ٤} \triangle^{٤} \right] ص_٠$$

$$= ص_٠ - ٦ر \triangle ص_٠ + ٤٨ر \triangle^{٢} ص_٠ - ٤١٦ر \triangle^{٣} ص_٠$$

$$+ ٣٧٤٤ر \triangle^{٤} ص_٠$$

$$= ٢٦٦ - ٦ر \times ٣٦٩ + ٤٨ر \times ٣٠٢ - ٤١٦ر \times ١٣٢$$

$$+ ٣٧٤٤ر \times ٢٤ .$$

ويجب ملاحظة أننا عوضنا هنا $ص_٠ = ٢٦٦$ ، وعوضنا عن $\triangle ص_٠$ ، $\triangle^{٢} ص_٠$ ، . بفروق هــذه القيمة الخاصة ٢٦٦ ، وهــذه الفروق هى على الترتيب ٣٦٩ و ٣٠٢ و ١٣٢ و ٢٤ ؛ وينتج أن

$$ص_{-٦ر} = ١٤٣ر٦٣٣٦ ،$$

وهى نفس النتيجة التى حصلنا عليها من قبل . ويتضح لنا من هذا المثال أن اختيار موقع $س_٠$ فى الجــدول مسألة اختيارية . وإنما نلاحظ أن الفروق التى عوضناها فى هذه المرة كانت أكثر أرقاماً من الفروق التى عوضناها فى الحل السابق عندما أخذنا $س_٠$ و $ص_٠$ فى أول الجـدول . فيحسن أن نختــار $س_٠$ و $ص_٠$ فى كل مسألة بعد إنشاء جدول الفروق حتى نرى أسهل الطرق فنتبعها . وعلى كل حال يجب أن نختــار $س_٠$ و $ص_٠$ بحيث يوجد الفروق المطلوبة . فلا يمكن هنا مثــلا أن نأخذ $س_٠ = ٨$ و $ص_٠ = ٤١٠٦$ لأن الجــدول لا يحتوى على فروق $ص_٠$ هذه المطلوبة فى الحساب .

٤٧٧ — فى الأمثلة السابقة كانت الفروق كلها موجبة ، وكانت الفروق الرابعة متساوية تماماً والخامسة صفراً . ولكن الفروق يصح أن تكون سالبة كلها أو بعضها . وفى بعض الأحيان لا نصل إلى فروق متساوية تماماً كما فى المثال الآتى :

المطلوب حساب لوغاريتم العدد ٣٫١٢ بواسطة الجدول الآتى :

العدد	اللوغاريتم	Δ	Δ^2	Δ^3	Δ^4	Δ^5
٣٫٠	٫٤٧٧١٢١٣					
		١٤٢٤٠٤				
٣٫١	٫٤٩١٣٦١٧		—٤٥٢١			
		١٣٧٨٨٣		٢٧٧		
٣٫٢	٫٥٠٥١٥٠٠		—٤٢٤٤		—٢٢	
		١٣٣٦٣٩		٢٥٥		٣
٣٫٣	٫٥١٨٥١٣٩		—٣٩٨٩		—٢٥	
		١٢٩٦٥٠		٢٣٠		٨
٣٫٤	٫٥٣١٤٧٨٩		—٣٧٥٩		—١٧	
		١٢٥٨٩١		٢١٣		
٣٫٥	٫٥٤٤٠٦٨٠		—٣٥٤٦			
		١٢٢٣٤٥				
٣٫٦	٫٥٥٦٣٠٢٥					

لم نكتب الشرطة العشرية والأصفار فى الفروق توفيراً للكتابة ؛ والمفهوم طبعاً أن الرقم الأيمن فى أعمدة الفروق فى المنزلة العشرية السابعة مثل الرقم الأيمن فى عمود اللوغاريتمات . ونرى الفروق الخامسة صغيرة ولا لزوم للاستمرار .

هنا نعتبر س٠ = ٣٫٠ و ص٠ = ٠٫٤٧٧١٢١٣

نفرض أن $٣٫١٢ = \text{س}_{\text{ر}}$

$$\therefore\ \text{س}_{\text{ر}} - \text{س}_{.} = ٣٫١٢ - ٣٫٠ = ٠٫١٢$$

$$= ١٫٢ \ \text{فترة سينية}.$$

$\therefore$ $\text{س}_{\text{ر}}$ هى $\text{س}_{١٫٢}$ والمطلوب إذاً هو $\text{ص}_{١٫٢}$

$$\text{ص}_{١٫٢} = (١ + \Delta)^{١٫٢}\ \text{ص}_{.}$$

$$= \Big[١ + ١٫٢\,\Delta + \frac{١٫٢ \times ٠٫٢}{٢}\Delta^{٢} + \frac{١٫٢ \times ٠٫٢ \times -٠٫٨}{٢ \times ٣}\Delta^{٣}$$

$$+ \frac{١٫٢ \times ٠٫٢ \times -٠٫٨ \times -١٫٨}{٢ \times ٣ \times ٤}\Delta^{٤}$$

$$+ \frac{١٫٢ \times ٠٫٢ \times -٠٫٨ \times -١٫٨ \times -٢٫٨}{٢ \times ٣ \times ٤ \times ٥}\Delta^{٥} + \dots \Big]\ \text{ص}_{.}$$

$$= \text{ص}_{.} + ١٫٢\ \Delta\,\text{ص}_{.} + ٠٫١٢\ \Delta^{٢}\text{ص}_{.} - ٠٫٠٣٢\ \Delta^{٣}\text{ص}_{.}$$

$$+ ٠٫٠١٤٤\ \Delta^{٤}\text{ص}_{.} - ٠٫٠٠٨٠٦٤\ \Delta^{٥}\text{ص}_{.}$$

$$= ٠٫٤٧٧١٢١٣ + ١٫٢ \times ٠٫٠١٤٢٤٠٤ - ٠٫١٢ \times ٠٫٠٠٠٤٥٢١$$

$$- ٠٫٠٣٢ \times ٠٫٠٠٠٠٢٧٧ + ٠٫٠١٤٤ \times -٠٫٠٠٠٠٠٢٢$$

$$- ٠٫٠٠٨٠٦٤ \times ٠٫٠٠٠٠٠٠٣ .$$

ومن الواضح أن الحدين الأخيرين لا تأثير لهما على النتيجة . وكان من الممكن الاستغناء عنهما . و بإجراء عمليات الضرب مع التقريب إلى الخانة العشرية السابعة ينتج أن

$$\text{ص}_{١٫٢} = \text{لو}\ ٣٫١٢$$

$$= ٠٫٤٩٤١٥٤٦ .$$

الاستكمال فى فترات سينية غير متساوية

الفروق المنقسمة

٤٧٨ — قانون نيوتن وجريجورى الذى شرحناه لا يصلح إلا فى الجداول التى فيها قيم س على فترات متساوية من بعضها . لأن برهان ذلك القانون مبنى على هذا الفرض كأساس [١] . ولكننا نحتاج فى كثير من الأحيان إلى الاستكمال فى فترات غير متساوية . وفى مثل هذه الأحوال إذا حسبنا الفروق بين قيم ص المتتالية ، فإن هذه الفروق إذا أخذت على علاتها لا تكون منتظمة . ولذلك فإننا نكون منها فروقاً أخرى تسمى [٢] **الفروق المنقسمة** وهذه نحسبها بأن نقسم الفرق بين قيمتين متتاليتين من قيم ص على الفرق بين قيمتى س المناظرتين لهما .

ولتوضيح كيفية حساب هذه الفروق ، نفرض أن قيم س المتتالية مرتبة تصاعدياً ، هى :

$$\text{ا}_{.} \text{ ، } \text{ا}_{1} \text{ ، } \text{ا}_{2} \text{ ، } \ldots$$

وأن قيم ص المناظرة لهذه هى على التوالى :

$$\text{د}(\text{ا}_{.}) \text{ ، } \text{د}(\text{ا}_{1}) \text{ ، } \text{د}(\text{ا}_{2}) \text{ ، } \ldots$$

مع فرض أن $\text{ا}_1 - \text{ا}_{.}$ لايساوى $\text{ا}_2 - \text{ا}_1$ وهكذا .

نقسم الفرق $\text{د}(\text{ا}_1) - \text{د}(\text{ا}_{.})$ على الفرق $(\text{ا}_1 - \text{ا}_{.})$ ينتج الفرق المنقسم للقيمة $\text{د}(\text{ا}_{.})$. ولو رمزنا له بالرمز $\text{د}(\text{ا}_{.} \text{ ، } \text{ا}_1)$ يكون :

---

(١) أنظر كتاب Whittaker and Robinson : *Calculus of Observations* Art. 2, p. 2.; Art. 11, p. 20.

(٢) بالإنجليزية *Divided Differences*

$$د\,(ا_{٠}\,،\,ا_{١}) = \frac{د\,(ا_{١}) - د\,(ا_{٠})}{ا_{١} - ا_{٠}}$$ = الفرق المنقسم الأول للقيمة $د\,(ا_{٠})$.

وكذلك تكون الفروق المنقسمة الأول للقيم $د\,(ا_{١})$ و $د\,(ا_{٢})$ و .. هى على الترتيب:

$$د\,(ا_{١}\,،\,ا_{٢}) = \frac{د\,(ا_{٢}) - د\,(ا_{١})}{ا_{٢} - ا_{١}}$$

$$و \quad د\,(ا_{٢}\,،\,ا_{٣}) = \frac{د\,(ا_{٣}) - د\,(ا_{٢})}{ا_{٣} - ا_{٢}}$$

$$و \quad د\,(ا_{٣}\,،\,ا_{٤}) = \frac{د\,(ا_{٤}) - د\,(ا_{٣})}{ا_{٤} - ا_{٣}}$$

وهكذا .

٤٧٩ — ولحساب الفروق المنقسمة الثانية نطرح أولا كل فرق من الذى يليه ، ثم نقسم باقى الطرح على الفرق بين قيمتى س الطرفيتين . الفروق الثانية وأعلى

فإذا رمزنا للفرق المنقسم الثانى للقيمة $د\,(ا_{٠})$ بالرمز $د\,(ا_{٠}\,،\,ا_{١}\,،\,ا_{٢})$ يكون

$$د\,(ا_{٠}\,،\,ا_{١}\,،\,ا_{٢}) = \frac{د\,(ا_{١}\,،\,ا_{٢}) - د\,(ا_{٠}\,،\,ا_{١})}{ا_{٢} - ا_{٠}}$$ ؛ وكذلك يكون

$$د\,(ا_{١}\,،\,ا_{٢}\,،\,ا_{٣}) = \frac{د\,(ا_{٢}\,،\,ا_{٣}) - د\,(ا_{١}\,،\,ا_{٢})}{ا_{٣} - ا_{١}}$$ هو الفرق الثانى للقيمة $د\,(ا_{١})$

$$و \quad د\,(ا_{٢}\,،\,ا_{٣}\,،\,ا_{٤}) = \frac{د\,(ا_{٣}\,،\,ا_{٤}) - د\,(ا_{٢}\,،\,ا_{٣})}{ا_{٤} - ا_{٢}}$$ هو الفرق الثانى للقيمة $د\,(ا_{٢})$

وهكذا .

ومن هذه الفروق الثانية نكون الفروق الثالثة بنفس الطريقة :

$$د\,(ا_٠ ، ا_١ ، ا_٢ ، ا_٣) = \frac{د\,(ا_١ ، ا_٢ ، ا_٣) - د\,(ا_٠ ، ا_١ ، ا_٢)}{ا_٣ - ا_٠}$$

= الفرق المنقسم الثالث للقيمة $د\,(ا_٠)$.

$$و\quad د\,(ا_١ ، ا_٢ ، ا_٣ ، ا_٤) = \frac{د\,(ا_٢ ، ا_٣ ، ا_٤) - د\,(ا_١ ، ا_٢ ، ا_٣)}{ا_٤ - ا_١}$$

= الفرق المنقسم الثالث للقيمة $د\,(ا_١)$.

وهكذا لجميع القيم $د\,(ا_٢)$ ، $د\,(ا_٣)$ ، . . . وبمثل ذلك نحصل على الفروق الأعلى . ونحسب هـــذه الفروق حتى نصل إلى فروق متساوية فى جدول الفروق المنقسمة :

قيم س	قيم ص	فروق أولى	فروق ثانية	فروق ثالثة	فروق رابعة
$ا_٠$	$د\,(ا_٠)$				
		$د\,(ا_٠ ، ا_١)$			
$ا_١$	$د\,(ا_١)$		$د\,(ا_٠ ، ا_١ ، ا_٢)$		
		$د\,(ا_١ ، ا_٢)$		$د\,(ا_٠ ، ا_١ ، ا_٢ ، ا_٣)$	
$ا_٢$	$د\,(ا_٢)$		$د\,(ا_١ ، ا_٢ ، ا_٣)$		$د\,(ا_٠ ، ا_١ ، ا_٢ ، ا_٣ ، ا_٤)$
		$د\,(ا_٢ ، ا_٣)$		$د\,(ا_١ ، ا_٢ ، ا_٣ ، ا_٤)$	
$ا_٣$	$د\,(ا_٣)$		$د\,(ا_٢ ، ا_٣ ، ا_٤)$		$د\,(ا_١ ، ا_٢ ، ا_٣ ، ا_٤ ، ا_٥)$
		$د\,(ا_٣ ، ا_٤)$		$د\,(ا_٢ ، ا_٣ ، ا_٤ ، ا_٥)$	
$ا_٤$	$د\,(ا_٤)$		$د\,(ا_٣ ، ا_٤ ، ا_٥)$		
		$د\,(ا_٤ ، ا_٥)$			
$ا_٥$	$د\,(ا_٥)$				

قانون نيوتن للفترات غير المتساوية

٤٨٠ — الخواص الجبرية لهذه الفروق المنقسمة تشابه خواص الفروق العادية التى عرفناها فى حالة الفترات المتساوية . وقد استنبط نيوتن قانوناً للاستكمال فى الفترات غير المتساوية باستخدام هذه الفروق المنقسمة . ونكتفى هنا بذكر هذا القانون بدون محاولة إثباته (١) .

نفرض أن القيمة السينية المطلوب استكمال الجدول عندها هى س ، وأن قيمة ص المناظرة لها والمطلوب معرفتها هى د (س) هذه القيمة الصادية نحسبها من القانون الآتى :-

$$\text{د}(\text{س}) = \text{د}(\text{ا}_{٠}) + (\text{س} - \text{ا}_{٠}) \,.\, \text{د}(\text{ا}_{٠}\text{،}\text{ا}_{١})$$

$$+ (\text{س} - \text{ا}_{٠})(\text{س} - \text{ا}_{١}) \,.\, \text{د}(\text{ا}_{٠}\text{،}\text{ا}_{١}\text{،}\text{ا}_{٢})$$

$$+ (\text{س} - \text{ا}_{٠})(\text{س} - \text{ا}_{١})(\text{س} - \text{ا}_{٢}) \,.$$

$$\text{د}(\text{ا}_{٠}\text{، ا}_{١}\text{، ا}_{٢}\text{، ا}_{٣})$$

$$+ \ldots\ldots$$

$$+ (\text{س} - \text{ا}_{٠})(\text{س} - \text{ا}_{١})\ldots(\text{س} - \text{ا}_{ر-١}) \,.$$

$$\,.\,\text{د}(\text{ا}_{٠}\text{،ا}_{١}\text{،}\ldots\text{،ا}_{ر}) \,.$$

حيث د (ا. ، ا₁ ، ، ا_ر) هو الفرق المنقسم ذو الرتبة ر ، بفرض أن الفروق الرائية متساوية ، والفروق التى بعدها صفر .

مثال : المطلوب حساب قيمة ص حينما س = ٣ من واقع الجدول الآتى :

س	٤	٥	٧	١٠	١١	١٣
ص	٤٨	١٠٠	٢٩٤	٩٠٠	١٢١٠	٢٠٢٨

من هذه القيم المتناظرة ننشىء جدول الفروق المنقسمة ونستخدم قانون نيوتن :

(١) انظر الملحق الرياضى فى آخر الكتاب .

س	ص				
٤	٤٨				
		٥٢			
٥	١٠٠		١٥		
		٩٧		١	
٧	٢٩٤		٢١		٠
		٢٠٢		١	
١٠	٩٠٠		٢٧		٠
		٣١٠		١	
١١	١٢١٠		٣٣		
		٤٠٩			
١٣	٢٠٢٨				

نعتبر د (ا.) هى ٤٨ ، ونضع س فى القانون تساوى ٣ ، ونعوض عن الفروق المختلفة .

∴ د (٣) = ٤٨ + (٣ — ٤) × ٥٢ + (٣ — ٤)(٣ — ٥) × ١٥

+ (٣ — ٤)(٣ — ٥)(٣ — ٧) × ١

= ١٨

صورة أخرى لقانون نيوتن

٤٨١ — هنـاك صورة أخرى لقانون نيوتن المذكور فى البند السابق . وهذه الصورة تفضل فى بعض الأحيان من الناحية العملية ؛ وهى كما يأتى :

د (س) = د (ا.)

+ (س — ا.) { د(ا. ا$_١$) + (س — ا$_١$) [د(ا. ، ا$_١$ ، ا$_٢$)

+ (س — ا$_٢$) (د(ا. ، ا$_١$ ، ا$_٢$ ، ا$_٣$) + . . .)] } .

ولتوضيح كيفية استخدام هذه الصورة عملياً نأخذ المثال الآتى :

قيم س : ٢ ٣ ٥ ٦ ٨ ٩

قيم ص : ٣١ ٩٦ ٦٤٠ ١٣١١ ٤١١١ ٦٥٧٦

والمطلوب حساب د (٤)

قيم س	قيم ص				
٢	٣١				
		٦٥			
٣	٩٦		٦٩		
		٢٧٢		١٦	
٥	٦٤٠		١٣٣		١
		٦٧١		٢٢	
٦	١٣١١		٢٤٣		١
		١٤٠٠		٢٨	
٨	٤١١١		٣٥٥		
		٢٤٦٥			
٩	٦٥٧٦				

نعتبر د (ا) = ٣١ ، ونطبق القانون فى الصورة الجديدة .

∴ د (٤) = ٣١ + (٤ − ٢) } ٦٥ + (٤ − ٣) ×

[٦٩ + (٤ − ٥) (١٦ + (٤ − ٦) × ١)] {

= ٣١ + (٤ − ٢) } ٦٥ + ١ [٦٩ − ١٤] {

= ٢٧١ .

٤٨١ — هنـاك قانون آخر استنبطه لاجرانج الفرنسى سـنة ١٧٩٥[١] ، يستعمل أيضاً للاستكمال فى الفترات غير المتساوية . وهـــذا القانون يعطى القيمة المطلوبة د (س) للمتغير ص بدلالة القيم المعطاة نفسها وليس بدلالة الفروق . قانون لاجرانج

بفرض أن قيم س الموجودة فى الجدول هى $\text{ا}_{٠}$ ، $\text{ا}_{١}$ ، . . . $\text{ا}_{ر}$ ؛

وأن قيم ص المناظرة لها فى الجدول هى د ($\text{ا}_{٠}$) ، د ($\text{ا}_{١}$) ... د ($\text{ا}_{ر}$) ؛

وأن قيمة س المطلوب استكمالها هى س ، نحسب د (س) من القانون :

$$\text{د}(\text{س}) = \frac{(\text{س}-\text{ا}_{١})\,(\text{س}-\text{ا}_{٢})\,\cdots\,(\text{س}-\text{ا}_{ر})}{(\text{ا}_{٠}-\text{ا}_{١})\,(\text{ا}_{٠}-\text{ا}_{٢})\,\cdots\,(\text{ا}_{٠}-\text{ا}_{ر})}\,.\,\text{د}\,(\text{ا}_{٠})$$

$$+ \frac{(\text{س}-\text{ا}_{٠})\,(\text{س}-\text{ا}_{٢})\,\cdots\,(\text{س}-\text{ا}_{ر})}{(\text{ا}_{١}-\text{ا}_{٠})\,(\text{ا}_{١}-\text{ا}_{٢})\,\cdots\,(\text{ا}_{١}-\text{ا}_{ر})}\,.\,\text{د}\,(\text{ا}_{١})$$

$$+ \cdots$$

$$+ \frac{(\text{س}-\text{ا}_{٠})\,(\text{س}-\text{ا}_{١})\cdots(\text{س}-\text{ا}_{ر-١})}{(\text{ا}_{ر}-\text{ا}_{٠})\,(\text{ا}_{ر}-\text{ا}_{١})\cdots(\text{ا}_{ر}-\text{ا}_{ر-١})}\,.\,\text{د}\,(\text{ا}_{ر})$$

على فرض أن الفروق المنقسمة من الرتبة (ر + ١) تساوى صـفراً . ويجب قبل تطبيق هذا القـانون فى أى مسألة معينـة أن نكون جدول الفروق المنقسمة لنتبين أى الفروق تساوى صفراً .

مثال : أوجد د (٤) فى المثال المذكور فى البند السابق بتطبيق قانون لاجرانج . فى هذا المثال نجد الفروق الرابعة متساوية والخامسة تساوى صفراً .

$$\therefore \text{د}(٤) = \frac{(٤-٣)\,(٤-٥)\,(٤-٦)\,(٤-٨)}{(٢-٣)\,(٢-٥)\,(٢-٦)\,(٢-٨)} \times ٣١ + \frac{(٤-٢)\,(٤-٥)\,(٤-٦)\,(٤-٨)}{(٣-٢)\,(٣-٥)\,(٣-٦)\,(٣-٨)} \times ٩٦$$

$$+ \frac{(٤-٢)\,(٤-٣)\,(٤-٦)\,(٤-٨)}{(٥-٢)\,(٥-٣)\,(٥-٦)\,(٥-٨)} \times ٦٤٠ + \frac{(٤-٢)\,(٤-٣)\,(٤-٥)\,(٤-٨)}{(٦-٢)\,(٦-٣)\,(٦-٥)\,(٦-٨)} \times ١٣١١$$

---

(١) أنظر كتاب Whittaker and Robinson : *Calculus of Observations*, p. 28.

$$+ \frac{(٤-٢)\,(٤-٣)\,(٤-٥)\,(٤-٦)}{(٨-٢)\,(٨-٣)\,(٨-٥)\,(٨-٦)} \times ٤١١١ \,.$$

$$= -\frac{٣١}{٩} + \frac{٢٥٦}{٥} - \frac{٥١٢٠}{٩} + \frac{١٣١١}{٣} + \frac{٤١١١}{٤٥}$$

$$= ٢٧١$$

وهى نفس النتيجة التى حصلنا عليها من قبل .

استخدام الفروق الوسطى

٤٨٣ — يلاحظ فى كل القوانين التى ذكرناها أننا نستعمل الفروق الخاصة بقيمة معينة من قيم ص — سواء فى الفروق العادية أو الفروق المنقسمة — وهذا يستلزم أن نتبع هذه الفروق فى خط مائل . وقد لاحظنا أيضاً أن جدول الفروق يكون على شكل مثلث تقريباً ، وأننا إذا أخذنا $ص_{.}$ قيمة غير الأولى أو الثانية أو الثالثة فربما لا نجد لها الفروق المطلوبة فى الجدول .

وهذه المسألة دعت إلى التفكير فى إمكان الاستكمال باستخدام[١] **الفروق الوسطى** أو المركزية ، أى الفروق التى تقع على خط أفقى واحد فى وسط جدول الفروق . وهذا الإجراء أسهل فى كثير من الأحيان ومريح عملياً ، فضلاً عن أننا نضمن وجود الفروق المطلوبة للحساب فى وسط الجدول ، وربما لا نجدها إذا ابتدأنا من أعلى أو من أسفل .

وهناك بعض القوانين للاستكمال باستخدام الفروق الوسطى ، يكفى أن نذكر منها بعضها . وفى هذه القوانين نرمز للقيم السينية التى فى الجدول بالرموز

$$س_{-٣} \text{ ، } س_{-٢} \text{ ، } س_{-١} \text{ ، } س_{.} \text{ ، } س_{١} \text{ ، } س_{٢} \text{ ، } س_{٣} \; . \; .$$

---

(١) هذه الفروق تسمى بالإنجليزية Central Differences

وللقيم الصادية المناظرة لها بالرموز

$$\text{ص}_{-٣} \text{ ، } \text{ص}_{-٢} \text{ ، } \text{ص}_{-١} \text{ ، } \text{ص}_{٠} \text{ ، } \text{ص}_{١} \text{ ، } \text{ص}_{٢} \text{ ، } \text{ص}_{٣}$$

وننشىء جدول الفروق كالمعتاد .

٤٨٤ — من هذه القوانين قانون يعرف باسم [١] نيوتن وجاوس ، وهو خاص بالفترات السينية المتساوية .

بفرض أن $\text{س}_{\text{ر}}$ هى القيمة المطلوب استكمال الجدول عندها ، وأن $\text{ص}_{\text{ر}}$ هى القيمة المطلوبة ، حيث ر تساوى عدد الفترات السينية (صحيح أو كسرى) بين $\text{س}_{٠}$ ، $\text{س}_{\text{ر}}$ فإن القيمة المطلوبة $\text{ص}_{\text{ر}}$ هى :

$$\text{ص}_{\text{ر}} = \text{ص}_{٠} + \text{ر} \,.\, \triangle \text{ص}_{٠} + \frac{\text{ر}(\text{ر}-١)}{\underline{|٢}} \,.\, \triangle^{٢} \text{ص}_{-١}$$

$$+ \frac{(\text{ر}+١)\,\text{ر}\,(\text{ر}-١)}{\underline{|٣}} \,.\, \triangle^{٣} \text{ص}_{-١}$$

$$+ \frac{(\text{ر}+١)\,\text{ر}\,(\text{ر}-١)(\text{ر}-٢)}{\underline{|٤}} \,.\, \triangle^{٤} \text{ص}_{-٢}$$

$$+ \frac{(\text{ر}+٢)(\text{ر}+١)\,\text{ر}\,(\text{ر}-١)(\text{ر}-٢)}{\underline{|٥}} \,.\, \triangle^{٥} \text{ص}_{-٢}$$

$$+ \ldots$$

والفروق المطلوبة هنا تقع قريبة من الخط الأفقى المار بالقيمة $\text{ص}_{٠}$

٤٨٥ — وقد وجد نيوتن قوانين أخرى للاستكمال بالفروق الوسطى منها القانون الآتى المعروف بقانون نيوتن و بسل [٢] ، وهو قانون نيوتن وبسل

---

(١) بالإنجليزية Newton-Gauss Formula

(٢) اسمه بالإنجليزية Newton-Bessel Formula

$$ص_ر = \tfrac{١}{٢}(ص_٠ + ص_١) + (ر - \tfrac{١}{٢}) \triangle ص_٠$$

$$+ \frac{ر(ر-١)}{\underline{\lfloor ٢}} \cdot \frac{١}{٢}(\triangle^٢ ص_{-١} + \triangle^٣ ص_٠)$$

$$+ \frac{ر(ر-١)(ر-\frac{١}{٢})}{\underline{\lfloor ٣}} \cdot \triangle^٣ ص_{-١}$$

$$+ \frac{(ر+١)\,ر\,(ر-١)(ر-٢)}{\underline{\lfloor ٤}} \cdot \frac{١}{٢}(\triangle^٤ ص_{-٢} + \triangle^٤ ص_{-١})$$

$$+ \cdots$$

ويلاحظ أن هذا القانون متماثل حول $ص_{\frac{١}{٢}}$.

٤٨٦ — ويمكننا تحوير قانون نيوتن وجاوس المذكور فى بنـد ٤٨٤ : قانون نيوتن استرلنج

$$ص_ر =$$

$$ص_٠ + ر\cdot\triangle ص_٠ + \tfrac{١}{٢}\, ر(ر-١)\,\triangle^٢ ص_{-١}$$

$$+ \tfrac{١}{٦}(ر+١)\,ر\,(ر-١)\cdot\triangle^٣ ص_{-١}$$

$$+ \tfrac{١}{٢٤}(ر+١)\,ر\,(ر-١)(ر-٢)\,\triangle^٤ ص_{-٢} + \cdots$$

ونغير ترتيب الحدود فنحصل على الصورة

$$ص_ر = ص_٠ + ر\,[\triangle ص_٠ - \tfrac{١}{٢}\triangle^٢ ص_{-١}] + \frac{١}{\underline{\lfloor ٢}}\cdot ر^٢\triangle^٢ ص_{-١}$$

$$+ \frac{ر(ر^٢-١)}{\underline{\lfloor ٣}}\,[\triangle^٣ ص_{-١} - \tfrac{١}{٢}\triangle^٤ ص_{-٢}]$$

$$+ \frac{١}{\underline{\lfloor ٤}}\, ر^٢(ر^٢-١)\cdot\triangle^٤ ص_{-٢} + \cdots$$

وباستخدام العلاقات المعروفة .

$$\triangle^٢ ص_{-١} = \triangle ص_٠ - \triangle ص_{-١}$$

و $$\triangle^٤ ص_{-٢} = \triangle^٣ ص_{-١} - \triangle^٣ ص_{-٢}$$

وهكذا . . . . ، والتعويض منها عن الفروق الزوجية الآتيــة داخل الأقواس ، نحصل على القانون نفسه فى صورة جديدة :

$$\text{ص}_{\text{ر}} = \text{ص}_{٠} + \frac{١}{٢}\,\text{ر}\left(\triangle\,\text{ص}_{٠} + \triangle\,\text{ص}_{-١}\right) + \frac{١}{٢}\,\text{ر}^{٢}.\triangle^{٢}\,\text{ص}_{-١}$$

$$+ \frac{\text{ر}\,(\text{ر}^{٢} - ١^{٢})}{٢\,.\,٣!}\left(\triangle^{٣}\text{ص}_{-١} + \triangle^{٣}\text{ص}_{-٢}\right) + \frac{\text{ر}^{٢}(\text{ر}^{٢} - ١)}{٤!}\,\triangle^{٤}\,\text{ص}_{-٢}$$

$$+ \frac{\text{ر}\,(\text{ر}^{٢} - ١^{٢})\,(\text{ر}^{٢} - ٢^{٢})}{٢\,.\,٥!}\left(\triangle^{٥}\,\text{ص}_{-٢} + \triangle^{٥}\,\text{ص}_{-٣}\right)$$

$$+ \frac{\text{ر}^{٢}\,(\text{ر}^{٢} - ١^{٢})\,(\text{ر}^{٢} - ٢^{٢})}{٦!}\,\triangle^{٦}\,\text{ص}_{-٣} + \ldots$$

والقانون فى هذه الصــورة يسمى [١] قانونــ نيوتن واستيرلنج ، حيث استنبطه نيوتن حوالى سنة ١٧١١ ، ومن بعده استرلينج فى سنة ١٧٣٠

٤٨٧ — هذا ويوجد كما قلنا عدة قوانين أخرى للاستكمال يتميز بعضها عن بعض حسب نوع المسألة المطلوب معالجتها والجـداول المطلوب استكمالهـا . ولا يتسع المقام هنـا لسرد جميعها فنكتفى بهذه القوانين التى ذكرناها . ونأخذ الآن بعض الأمثلة لتوضيح طرق استخدام هذه القوانين .

مثال : المطلوب إيجاد ص حينما س = ٣٣٫٦ من الجدول الآتى :

س	٣١	٣٢	٣٣	٣٤	٣٥	٣٦
ص	٩٢٣٥٢١	١٠٤٨٥٧٦	١١٨٥٩٢١	١٣٣٦٣٣٦	١٥٠٠٦٢٥	١٦٧٩٦١٦

أولا — نستخدم قانون نيوتن وجاوس ، ونعتـــبر $\text{س}_{٠} = ٣٣$ ، $\text{ص}_{٠} = ١١٨٥٩٢١$ ، ونعتبر $\text{س}_{\text{ر}} = ٣٣٫٦$ فتــكون $\text{س}_{\text{ر}}$ هى $\text{س}_{٠٫٦}$.

---

(١) بالانجليزية Newton-Stirling Formula

س	ص	$\triangle$	$\triangle^2$	$\triangle^3$	$\triangle^4$	$\triangle^5$
٣١	٩٢٣٥٢١					
		١٢٥٠٥٥				
٣٢	١٠٤٨٥٧٦		١٢٢٩٠			
		١٣٧٣٤٥		٧٨٠		
٣٣	١١٨٥٩٢١		١٣٠٧٠		٢٤	
		١٥٠٤١٥		٨٠٤		٠
٣٤	١٣٣٦٣٣٦		١٣٨٧٤		٢٤	
		١٦٤٢٨٩		٨٢٨		
٣٥	١٥٠٠٦٢٥		١٤٧٠٢			
		١٧٨٩٩١				
٣٦	١٦٧٩٦١٦					

$$\therefore \text{ص}_{٠,٦} = ١١٨٥٩٢١ + ٦\,\triangle\,\text{ص}_{.} - \frac{٦, \times ٤,}{٢}\,\triangle^{٢}\text{ص}_{-١}$$

$$- \frac{١,٦ \times ٦, \times ٤,}{٦}\,\triangle^{٣}\text{ص}_{-١}$$

$$+ \frac{١,٦ \times ٦, \times ٤, \times ١,٤}{٢٤}\,\triangle^{٤}\text{ص}_{-٢}\,.$$

$$= ١١٨٥٩٢١ + ٦, \times ١٥٠٤١٥ - ١٢, \times ١٣٠٧٠$$

$$- ٠٦٤, \times ٨٠٤ + ٠٢٢٤, \times ٠٢٤$$

$$= ١٢٧٤٥٥٠,٦٨١٦$$

ويلاحظ هنا أننا استخدمنا الفروق فى السطر المار بالقيمة السينية ٣٣ ، أى س. ، ولم نخرج عن هذا السطر إلى مرتين إلى أسفل .

ثانياً — نستخدم قانون نيوتن و بسل ، ونعتبر أيضاً س. = ٣٣ ، يناظرها

ص. = ١١٨٥٩٢١

ونعتبر $س_ر = ٣٣٫٦$ ، فتكون $س_ر$ هى $س_{٠٫٦}$.

$$\therefore ص_{٠٫٦} = \frac{١}{٢} (ص_. + ص_١) + ٫١ \times \triangle ص_.$$

$$- \frac{٫٦ \times ٫٤}{٤} (\triangle^٢ ص_{-١} + \triangle^٢ ص_.)$$

$$- \frac{٫٦ \times ٫٤ \times ٫١}{٦} . \triangle^٣ ص_{-١} + \frac{١٫٦ \times ٫٦ \times ٫٤ \times ١٫٤}{٢٤ \times ٢}$$

$$(\triangle^٤ ص_{-٢} + \triangle^٤ ص_{-١}) .$$

$$= \frac{١}{٢} \left(\begin{matrix} ١١٨٥٩٢١ \\ +١٣٣٦٣٣٦ \end{matrix} \right) + ٫١ \times ١٥٠٤١٥ - ٫٠٦ \left(\begin{matrix} ١٣٠٧٠ \\ +١٣٨٧٤ \end{matrix} \right)$$

$$- ٫٠٠٤ \times ٨٠٤ + ٫٠١١٢ \left(\begin{matrix} ٢٤ \\ +٢٤ \end{matrix} \right) ؛$$

$$= ١٢٧٤٥٥٠٫٦٨١٦ ؛$$

وهى نفس النتيجــة التى حصلنا عليها باســتخدام قانون نيوتن وجاوس . ويلاحظ أننا استخدمنا القيم والفروق الموجودة فى الجدول على السطرين المارين بالقيمتين السينيتين ٣٣ و ٣٤ ، أى $س_.$ و $س_١$ أى على جانبي $س_{\frac{١}{٢}}$ وهذا كما قلنا فى البند السابق أن هذا القانون متماثل بالنسبة إلى $ص_{\frac{١}{٢}}$.

ثالثاً — نستخدم قانون نيوتن واستيرلنج ؛ ونعتبر $س_. = ٣٣$ يناظرها $ص_. = ١١٨٥٩٢١$. ونعتــبر $س_ر = ٣٣٫٦$ فتكون ، $ر = ٠٫٦$ ونعوض فى القانون :

$$\therefore ص_{٠٫٦} = ص_. + \frac{٫٦}{٢} (\triangle ص_. + \triangle ص_{-١}) + \frac{٫٣٦}{٢} \triangle^٢ ص_{-١}$$

$$- \frac{٫٦ \times ٫٦٤}{٢ \times ٦} (\triangle^٣ ص_{-١} + \triangle^٣ ص_{-١}) - \frac{٫٣٦ \times ٫٦٤}{٢٤} \triangle^٤ ص_{-٢}$$

$$= ١١٨٥٩٢١ + ٠٫٣ \left(\begin{matrix} ١٣٧٣٤٥ \\ +١٥٠٤١٥ \end{matrix} \right) + ٠٫١٨ \times ١٣٠٧٠$$

$$- ٠٫٠٣٢ \left(\begin{matrix} ٧٨٠ \\ +٨٠٤ \end{matrix} \right) - ٠٫٠٠٩٦ \times ٢٤ ؛$$

$$= ١٢٧٤٥٥٠٫٦٨١٦ ،$$

وهى نفس النتيجة التى حصلنا عليها من قبل .

ويلاحظ هنا أيضاً أننا استخدمنا فقط الأرقام الموجودة فى جدول الفروق على السطر المتوسط المار بالقيمة $\text{س}_{٠}$ والسطرين المحيطين به ؛ أى أن هذا القانون متماثل بالنسبة إلى $\text{ص}_{٠}$.

تفصيل المجموعة فى الفئات الواسعة

٤٨٨ — هناك تطبيق لقانون نيوتن واسترلنج وضعه ج . كِنج[1] . إذا علم تقسيم السكان فى بلد معينة فى فئات الأعمار ، وكان مدى الفترة فى كل فئة خمس سنوات مثلاً ، فالمطلوب معرفة عدد السكان فى كل سنة من العمر داخل الفترة .

لنفرض أن عدد السكان فى الفئات المتتالية هو … $\text{ك}_{١}$ ، $\text{ك}_{٠}$ ، $\text{ك}_{-١}$ ، . .
وأن مدى الفئة خمس سنوات كما قلنا . ونفرض أن تعدادات السكان فى السنين المختلفة من العمر هى … $\text{ص}_{-٢}$ ، $\text{ص}_{-١}$. $\text{ص}_{٠}$ ؛ $\text{ص}_{١}$ ، $\text{ص}_{٢}$ ، …

$$\therefore \text{ يكون } \text{ك}_{١} = \text{ص}_{٧} + \text{ص}_{٦} + \text{ص}_{٥} + \text{ص}_{٤} + \text{ص}_{٣}$$

$$\text{و} \quad \text{ك}_{٠} = \text{ص}_{٢} + \text{ص}_{١} + \text{ص}_{٠} + \text{ص}_{-١} + \text{ص}_{-٢}$$

$$\text{و} \quad \text{ك}_{-١} = \text{ص}_{-٣} + \text{ص}_{-٤} + \text{ص}_{-٥} + \text{ص}_{-٦} + \text{ص}_{-٧}$$

والمطلوب هو معرفة $\text{ص}_{٠}$ بدلالة المجاميع ك

رأينا فى قانون نيوتن واسترلنج أن

$$\text{ص}_{ر} = \text{ص}_{٠} + ر \, \frac{\Delta \text{ص}_{٠} + \Delta \text{ص}_{-١}}{٢} + \frac{١}{٢} ر^{٢} \Delta^{٢} \text{ص}_{-١}$$

---

(١) أنظر G. King: *Journal of Institute of Actuaries*, 43 (1909)pp. 109; 50, p.32.
أو كتاب Whittaker and Robinson: *Calculus of Observations*) (1929) pp. 57-60.

$$+ \frac{\text{ر}(\text{ر}^{2}-1)}{6} \cdot \frac{\Delta^{3}\,\text{ص}_{-2} + \Delta^{3}\,\text{ص}_{-1}}{2}$$

$$+ \frac{\text{ر}^{2}(\text{ر}^{2}-1)}{24}\,\Delta^{4}\,\text{ص}_{-2} + \cdots\cdots$$

و بوضـع $\text{ل}_{\text{ر}} = \text{ص}_{\text{ر}} + \text{ص}_{-\text{ر}}$ ، و إهمال فروق الرتبة الرابعة وما يليها ،

$$\therefore \quad \text{ل}_{\text{ر}} = 2\,\text{ص}_{.} + \text{ر}^{2}\,\Delta^{2}\,\text{ص}_{-1}$$

$$\therefore \quad \text{ك}_{.} = \text{ص}_{.} + \text{ل}_{1} + \text{ل}_{2}$$

$$= 5\,\text{ص}_{.} + 5\,\Delta^{2}\,\text{ص}_{-1} \qquad (1)$$

$$\text{و} \quad \text{ك}_{1} + \text{ك}_{-1} = \text{ل}_{3} + \text{ل}_{4} + \text{ل}_{5} + \text{ل}_{6} + \text{ل}_{7}$$

$$= 10\,\text{ص}_{.} + 135\,\Delta^{2}\,\text{ص}_{-1} \qquad (2)$$

و ينتج من المعادلتين (١) و (٢) أن

$$125\,\text{ص}_{.} = 27\,\text{ك}_{.} - (\text{ك}_{-1} + \text{ك}_{1}) \qquad (3)$$

وهذه المعادلة تعطينا القيمة المطلوبة $\text{ص}_{.}$ بدلالة القيم المعلومة $\text{ك}_{.}$ ، ك_{1} ، ك_{-1}

و إذا عوضنا $\Delta^{2}\,\text{ك}_{-1} = \text{ك}_{-1} - 2\,\text{ك}_{.} + \text{ك}_{1}$ ينتج أن :

$$\text{ص}_{.} = 0{,}2\,\text{ك}_{.} - 0{,}008\,\Delta^{2}\,\text{ك}_{-1} \cdots \qquad (4)$$

مثال : فى سنة ١٩٢٧ كان عدد سكان القاهرة فى فئات العمر ١٥ — ١٩ و ٢٠ — ٢٤ و ٢٥ — ٢٩ هو على الترتيب ١٠٧٨٢١ و ١٠١٦٠٥ و ٩٨٩٤٧ نسمة . والمطلوب معرفة عدد السكان الذين عمرهم ٢٢ سنة وأقل من ٢٣ (أى مركز الفئة ٢٠ — ٢٤) .

ننشىء جدول الفروق وتطبق القانون السابق ، معادلة (٤) .

$ك_{-١} = ١٠٧٨٢١$

$- ٦٢١٦$

$ك_{.} = ١٠١٦٠٥$ ٣٥٥٨

$- ٢٦٥٨$

$ك_{١} = ٩٨٩٤٧$

$\therefore$ $ص_{.} = ٢ر \times ١٠١٦٠٥ - ٠٠٨ر \times ٣٥٥٨$

$= ٢٠٢٩٣$

٤٨٩ — وقد استنبط كنج أيضاً معادلة [١] أخرى غير (٤) فى البند السابق ، بفرض استبقاء الفروق الرابعة $\triangle^{٤}$ ك و إهمال السادسة وما بعدها . وهذه المعادلة هى :

$$ص_{.} = ٢ر ك_{.} - ٠٠٨ر \triangle^{٢} ك_{-١} + ٠٠٠٨٩٦ر \triangle^{٤} ك_{-٢}$$

تمهيد البيانات الإحصائية

٤٩٠ — البيانات التى نحصل عليها عملياً عندما نجرى تجربة لدراسة العلاقة بين ظاهرتين متغيرتين ، تكون فى العادة معرضة لأخطاء تجريبية وتذبذبات عرضية لا حكم لها ولا ضابط . فإذا حسبنا مثلاً معدل الوفيات بين جماعة من الموظفين بالنسبة إلى طول مدة الخدمة بالسنين ، نجد أن هذا المعدل يتذبذب بين سنة وأخرى بالرغم من أن له اتجاهاً معيناً . والخلاصة أننا عند التأمل فى سلسلة من القيم مثل هذه لا يمكننا أن نأخذ أى واحدة منها على علاتها ونعتمد عليها تماماً ونثق فى صحتها وخلوها من هذه التأثيرات العرضية .

نتائج التجارب العمليـة تتأثر بعوامل عرضية

---

(١) راجع برهان هذه المعادلة فى كتاب Whittaker and Robinson صحيفة ٥٩

البيانات التجريبية لا يمكن الاعتماد عليها كما هى

٤٩١ — فإذا أردنا استخدام بعض القيم من هذه السلسلة فى عمل حسابات خاصة ، كما لو كانت شركة من شركات التأمين تفكر فى التأمين على حياة هؤلاء الموظفين مثلاً ، يجب أن نهذب هذه القيم المشاهدة بحيث نتخلص من هذه التأثيرات العرضية ونحصل بذلك على القيم « الحقيقية » أو « النظرية » أو « القيم الأكثر احتمالا » فى الحصول ، حتى إذا اعتمدنا عليها فى حساباتنا لا يكون الخطأ فيها كبيراً أو معدوماً إذا أمكن . وهذه هى العملية التى نسميها **تمهيد أو تدريج البيانات** .

قوانين للتمهيد

٤٩٢ — لنفرض أن سلسلة القيم التى حصلنا عليها بالتجربة هى :

$$ص_١ ، ص_٢ ، ص_٣ ، . . . . ، ص_ن$$

ونفرض أن أى قيمة $ص_ر$ من هذه تصبح بعد تمهيدها $= \bar{ص}_ر$. والمطلوب إذن معرفة $\bar{ص}_ر$ بدلالة القيم العادية المشاهدة كلها أو بعضها . وسنذكر هنا بعض القوانين المستعملة فى حساب القيم الممهدة من واقع القيم المشاهدة بدون التعرض لبراهين (١) هذه القوانين .

$$\bar{ص}_٠ = ص_٠ - \frac{٣}{٣٥} \Delta^٤ ص_{-٢} \qquad (١)$$

$$\bar{ص}_٠ = \frac{١}{٢١} [٧ ص_٠ + ٦ (ص_١ + ص_{-١})$$

$$+ ٣ (ص_٢ + ص_{-٢}) - ٢ (ص_٣ + ص_{-٣})] \qquad (٢)$$

ويلاحظ أن القانون الأول يستخدم $ص_٠$ وأربع قيم أخرى محيطة بها لحساب القيمة الممهدة $\bar{ص}_٠$ التى تناظرها . بينما القانون الثانى يستخدم ست قيم غير $ص_٠$ ؛ وهو أدق ولكنه أكثر تعقيداً .

---

(١) لإثبات هذه القوانين وغيرها أنظر كتاب Whittaker and Robinson, p. 291
والعملية تسمى Smoothing or Graduation of Data

وبديهى أنه يمكن اعتبار أى قيمة من التى عندنا هى ص. وحساب القيمة المهدة المناظرة لها . غير أنه بالقرب من طرفى الجدول الذى عندنا لا يمكننا أن نستعمل هذين القانونين لتمهيد القيمة الأولى والثانية أو الأخيرة والتى قبلها .

ويوجد غير هذين قوانين كثيرة ولكنها أكثر تعقيداً . والفكرة الأساسية هنا هى توفيق منحنيات لمجموعات من القيم المتتالية (خمس قيم كما فى القانون الأول ، أو ٧ كما فى الثانى ، أو أكثر كما فى القوانين الأخرى) ويصح أن تكون هذه المنحنيات من الدرجة الأولى ، أو من الدرجة الثانية أو من درجة أعلى . والقانونان السابقان على أساس منحن من الدرجة الثانية .

المراجع

Bowley A.L.: *Elements of Statistics*, Chapter X

Connor L.R.: *Statistics in Theory and Practice*, Chapter XVII

Whittaker and Robinson : *Calculus of Observations*, Chap. I-III, XI.

Yule, G.U. : *Introduction to the Theory of Statistics*, Chapter XXIV.

ملحق رياضى

المنحنيات التكرارية

لنفرض أن كمية متغيرة مثل س تأخذ عدة قيم مختلفة فى حالات مختلفة ؛ ولنفرض مثلاً أننا نبحث فى أعمار مجموعة معينة من الأشخاص ، وأن س تمثل عمر الشخص بالسنين — ففى هذه الحالة تتغير قيمة س من شخص إلى آخر ؛ ولكن اختلاف قيم س لا يمنع من أن يكون بعض الأشخاص متقاربين فى العمر كأن يعتبروا من سن واحدة . أى أن بعض القيم التى تأخذها س تكون متقاربة من بعضها لدرجة اعتبارها متساوية على وجه التقريب . فمثلاً نعتبر جميع الأشخاص الذين تنحصر أعمارهم بين ١٧ سنة و ٦ شهور ، و١٨ سنة و ٦ شهور متساوين فى السن ، وأن عمر الواحد منهم ١٨ سنة (بوجه التقريب) .

وبهذه الطريقة تنقسم مجموعة الأشخاص التى نحن بصددها إلى عدة مجموعات صغيرة ، أو فئات للسن ، تحتوى كل فئة على عدد من الأشخاص معتبرين فى سن واحدة ، مع أنهم فى الحقيقة — لو راعينا الدقة — ذوو أعمار مختلفة نوعاً ، إلا أن هذا الاختلاف صغير يمكن إهماله . وعلى ذلك يمكن اعتبار قيم الكمية المتغيرة س منقسمة إلى فئات أو فترات كل فئة لها مبدأ ولها نهاية : مثلاً من ٦ — ١٧ إلى ٦ — ١٨ و ٦ — ١٨ إلى ٦ — ١٩ إلخ ، وهما الحدان الأدنى والأعلى للفئة أو الفترة . وكل فئة تحتوى على عدد من القيم نعتبرها كلها متساوية تقريباً ؛ وكل منها يساوى القيمة التى فى منتصف الفترة . وعند ذلك نقول إن هذه القيمة الخاصة تكررت عدداً من المرات يساوى عدد القيم الواقعة فى هذه الفترة .

وليس من الضرورى أن يكون عدد القيم واحداً فى كل فترة ، فيختلف عدد القيم أى التكرار ، حسب موقع الفترة ، أى بحسب القيمة التى فى منتصفها .

ويختلف توزيع مفردات المجموعة على الفئات بحسب القانون الذى تتغير تبعاً له الكمية س . وفى الحقيقة يعتبر هذا التوزيع وشكله من مميزات الكمية المتغيرة التى نبحثها . ويستخدم كطريقة للدلالة على القانون الذى تسير عليه الكمية س فى تغيرها .

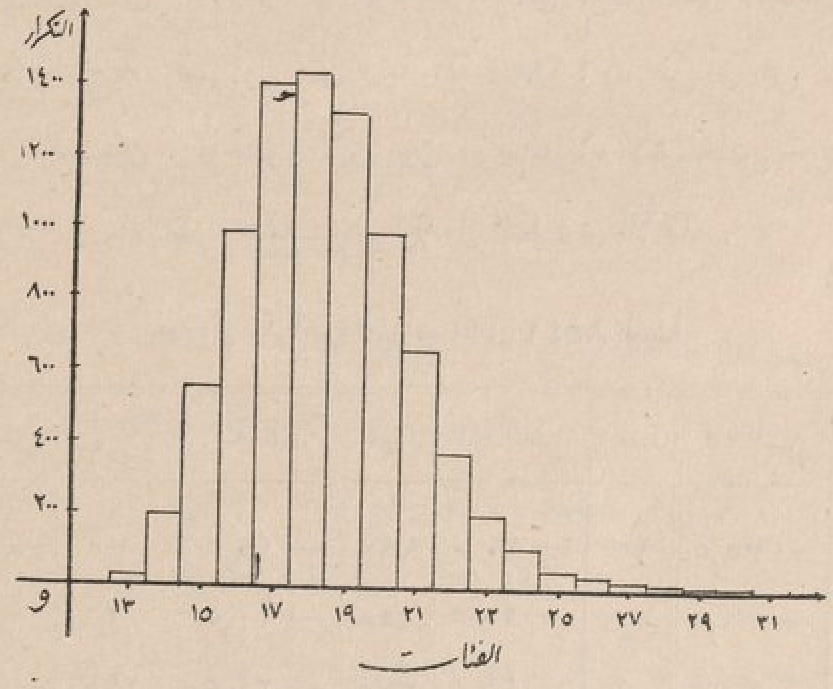
التمثيل الهندسى لتوزيع القيم

الآن نجتهد فى رسم شكل بيانى للتوزيع ونستخدمه لبحث خواصه وخواص الكمية المتغيرة س . لنأخذ مثلا مجموعة الأشخاص الذين تقدموا لامتحان شهادة الدراسة الثانوية (قسم أول) فى يونية سنة ١٩٣٤ ؛ ونفرض أن س تمثل عمر الشخص بالسنين . والجدول الآتى يبين توزيع هؤلاء الأشخاص — وعددهم ٨٤٤١ — بين الأعمار المختلفة ، وهو الجدول التكرارى للأعمار .

جدول تكرارى لتوزيع أعمار ٨٤٤١ تلميذاً

فئة السن	عدد الأشخاص	فئة السن	العدد	فئة السن	العدد	فئة السن	العدد
١١,٥ —	١	١٩,٥ —	٩٨٧	٢٧,٥ —	١٣	٣٥,٥ —	١
١٢,٥ —	٢٦	٢٠,٥ —	٦٥٨	٢٨,٥ —	٧	٣٦,٥ —	١
١٣,٥ —	١٩٧	٢١,٥ —	٣٧٨	٢٩,٥ —	٢	٣٧,٥ —	٠
١٤,٥ —	٥٦٧	٢٢,٥ —	٢٠٠	٣٠,٥ —	٥	٣٨,٥ —	١
١٥,٥ —	٩٩٢	٢٣,٥ —	١١١	٣١,٥ —	٢	٣٩,٥ —	١
١٦,٥ —	١٤١٢	٢٤,٥ —	٥٥	٣٢,٥ —	٠	٤٠,٥ —	٠
١٧,٥ —	١٤٣٨	٢٥,٥ —	٣٠	٣٣,٥ —	٢	٤١,٥ —	١
١٨,٥ —	١٣٣٠	٢٦,٥ —	٢٠	٣٤,٥ —	١		
						المجموع	٨٤٤١

ولبيان هـذا التوزيع هندسياً نرسم محورين متعـامدين نقيس قيم س على الأفقى منهمـا فنقسمه إلى فترات مدى الواحدة يساوى الوحـدة ، ونرسم على كل فترة مستطيلا رأسياً تمثل مساحته عدد الأشخاص الذين فى هذه الفترة ؛ فنرى أن القواعد العليا لهذه المستطيلات المتجاورة تكون خطاً منكسراً ؛ ونلاحظ أن هـذه المستطيلات تعتر فى الارتفاع إلى حد أقصى ثم ينقص ارتفاعها تدريجياً حتى تتلاشى ، وهـذا الشكل يبين التوزيع التكرارى لهذه الأعمـار ، ويسمى الهيستوجرام .



(شكل ٧٠)

ونلاحظ مبدئياً أن :

١ — ارتفاع المستطيل يتوقف على بعده عن المحور الرأسى ، أى أن الارتفاع يتغير مع تغير قيمة س عند منتصف القاعدة السفلى للمستطيل .

٢ — مساحة كل مستطيل تمثل عدد القيم الواقعة بين مبـدأ ونهاية الفترة

المقام عليها المستطيل ؛ أى أن مجموع مساحات المستطيلات كلها يمثل عدد القيم كلها — وهو ٨٤٤١ فى هذا المثال .

٣ — ينتج من (٢) أن

$$\frac{\text{مساحة المستطيل ا ح}}{\text{مجموع مساحات المستطيلات كلها}} = \frac{١٤١٢}{٨٤٤١}$$

= احتمال أن أى شخص حيثما اتفق يكون عمره بين ١٦٫٥ و ١٧٫٥ وكذلك احتمال أن شخصاً ما يكون عمره بين ٢٢٫٥ و ٢٣٫٥ مثلا هو $\frac{٢٠٠}{٨٤٤١}$ وهكذا .

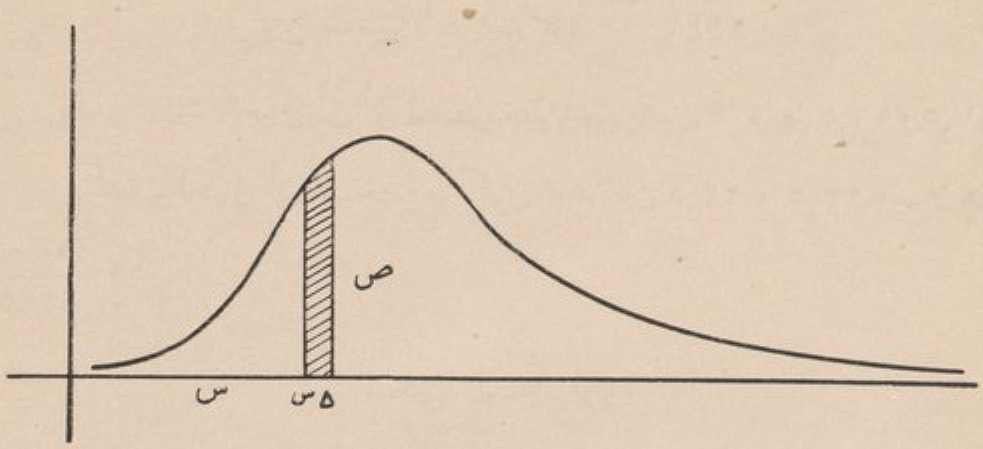
المنحنى التكرارى

لنفرض الآن على وجه العموم أن مبدأ الفترة هو س ونهايتها س + △ س (حيث △ س كمية صغيرة موجبة) . أى أن مداها = △ س وهو عرض المستطيل . ثم نفرض أن △ س تصغر ، فنجد أن هذه المستطيلات المتجاورة تضيق ؛ وهذا الضيق يقلل من حدة أركانها العليا ، حتى إذا ما صغرت △ س وقربت من الصفر تلاشت هذه الأركان ، وحصلنا فى النهاية على منحن ممهد متواصل نسميه **المنحنى التكرارى** .

ومعنى تصغير △ س فى المثال السابق هو بالطبع أننا نقسم كل فئة إلى فئتين صغيرتين مثلا ، كل واحدة منهما مداها ٦ أشهر بدلا من سنة كاملة . ثم نحصى عدد الأشخاص الذين يقعون فى كل نصف سنة على حدة ؛ وبذا يتكون جدول به ٦٢ فئة بدل ٣١ ، يمثله هيستوجرام مكون من ٦٢ مستطيلا — ثم نقسم كل فئة بعد ذلك إلى فئتين فنحصل على ١٢٤ فئة و ١٢٤ مستطيلا ، وهكذا . وبفرض أن عدد الأشخاص فى الأصل كبير جداً ، فسنجد فى كل فئة من الفئات

الصغيرة الجديدة عدداً محسوساً من الأشخاص يمثله مستطيل ضيق جداً في الهيستوجرام .

وإذا استمررنا في هذا التقسيم إلى النهاية يؤول الهيستوجرام إلى منحن ممهد متواصل كالمبين في شكل ٢



شكل (٧١)

ويكون عدد القيم المحصورة بين س ، س + △ س تمثله مساحة الشريط الضيق الذي عرضه △ س وارتفاعه ص ، ويبعد عن محور الصادات بمقدار س . وهنا يكون مساحة المنحنى كله تمثل عدد القيم كلها وليكن ن مثلاً (٨٤٤١ في المثال المذكور) . وإذا فرضنا على العموم أن معادلة المنحنى المبين في الشكل هي :

ص = ك (س)

حيث ص دالة للمتغير س نرمز لها بالرمز ك (س) ، ونسميها [١] **دالة التكرار** ؛ لأن عدد أو تكرار القيم الواقعة بين س ، س + △ س يساوي ص × △ س ، ويكون مجموع التكرارات كلها يساوي مجموع مساحات هذه الأشرطة أي مساحة المنحنى .

(١) بالإنجليزية Frequency Function

$$\therefore\ \text{ن} = \int \text{ص} \,.\, \text{ء س} = \text{مجموع التكرارت كلها أو عدد القيم كلها}$$

$$\frac{\text{مساحة الشريط}}{\text{مساحة المنحنى كله}} = \frac{\text{ص} \,.\, \triangle \text{س}}{\int \text{ص} \,.\, \text{ء س}} = \frac{\text{ص} \,.\, \triangle \text{س}}{\text{ن}}$$

$$= \text{احتمال وقوع قيمة س بين الحدين س ، س} + \triangle\text{س}$$

ولهذا نسمى[1] $\frac{\text{ص}}{\text{ن}}$ أو صـ **دالة الاحتمال**.

وواضح أن تكامل دالة الاحتمال يساوى

$$\int \frac{\text{ص}}{\text{ن}} \,.\, \text{ء س} = \frac{1}{\text{ن}} \int \text{ص} \,.\, \text{ء س} = 1 \,.$$

ويلاحظ أن الفرق بين دالتى التكرار والاحتمال هو فى وجود المقدار ن فى المقام. وقد عرفنا أن ن تساوى مجموع التكرارات أو مساحة المنحنى التكرارى.

وصورة الدالة التكرارية ك (س) تتوقف على ظروف تغير الكمية س. وهى تعتبر من مميزات هذه الكمية، وتعتبر دالة التكرار هى القانون الرياضى الذى تتغير تبعاً له الكمية س. ويوجد عدة صور لهذه الدوال فى علم الإحصاء. وهى:

$$1 - \text{ص} = \text{ص}_{0} \left(1 + \frac{\text{س}}{\text{ا}}\right)^{\text{ا م}} \,.\, \left(1 + \frac{\text{س}}{\text{ب}}\right)^{\text{ب م}}$$

$$2 - \text{ص} = \text{ص}_{0} \left(1 - \frac{\text{س}^{2}}{\text{ا}^{2}}\right)^{\text{ا م}}$$

$$3 - \text{ص} = \text{ص}_{0} \left(1 + \frac{\text{س}}{\text{ا}}\right)^{\text{ا م}} \,.\, \text{هـ}^{-\text{م س}}$$

$$4 - \text{ص} = \text{ص}_{0} \left(1 + \frac{\text{س}^{2}}{\text{ا}^{2}}\right)^{-\text{م}} \,.\, \text{هـ}^{-\text{ن ظا}^{-1} \frac{\text{س}}{\text{ا}}}$$

$$5 - \text{ص} = \text{ص}_{0}\, \text{س}^{-\text{ب}} \,.\, \text{هـ}^{-\frac{\text{حـ}}{\text{س}}}$$

---

(١) بالإنجليزية Probability Function

$$٦ - ص = ص. (س - ١)^{ب} . س^{-ح}$$

$$٧ - ص = ص. هـ^{- \frac{(س - ١)^٢}{ب}}$$

حيث هـ هى أساس اللوغاريتم الطبيعى أى أن $هـ = ١ + ١ + \frac{١}{\lfloor ١} + \frac{١}{\lfloor ٢} + \ldots \infty$. وكل الحروف ا ، ب ، م ، ح ، ص. تدل على كميات ثابتة مستقلة عن س و ص هى تختلف بحسب كل حالة .

وهذه الصورة المختلفة استنبطها كارل بيرسون بطريقة رياضية عامة ، ووضعها بهذا الترتيب ، وهى تشمل كل المنحنيات التكرارية المعروفة . ونرى الصورة الأخيرة تمثل منحنى الخطأ الطبيعى أو المنحنى المعتدل (الذى سبقت الإشارة إليه فى الجزء الأول) . ونظراً لأهميته فى الأبحاث الإحصائيـة نشرح كيفية استنباط هذه المعادلة .

معادلة المنحنى التكرارى المعتدل

إذا رمينا ١٢ زهرة ن مرات ، باعتبار النجاح هو ظهور جهار أو بيش أو شيش ، فإن احتمال الحصول على زهرات ناجحة عددها ر فى أى رميـة هو

$$^{١٢}ق_{ر} \times \left(\frac{١}{٢} \right)^{ر} \left(\frac{١}{٢} \right)^{١٢ - ر}$$

والاحتمالات المختلفة لقيم ر هى على التوالى حدود المفكوك :

$$\left(\frac{١}{٢} + \frac{١}{٢} \right)^{١٢}$$

طبقاً لنظرية ذات الحدين . وإذا رمينا هـذه الزهرات ٤٠٩٦ مرة (كما فعـل وِلْدُونْ فى تجربتـه) فإن التكرارات النظرية لقيم ر هى على التوالى حدود المفكوك .

$$٤٠٩٦ \left(\frac{١}{٢} + \frac{١}{٢} \right)^{١٢}$$

وهذه التكرارات تكون متماثلة حول القيمة $\text{ر} = ٦$ كما نرى فى جدول ١٥ (صفحة ٩٢) والمنحنى التكرارى هو منحن يمثل توزيعاً تكرارياً مثل هذا ؛ والمطلوب استنباط معادلة هذا المنحنى على العموم .

لنفرض أن ص تساوى احتمال نجاح ن + س من الزهرات وعدم نجاح ن − س من الزهرات . أى أننا نرمى ٢ ن من الزهرات كل مرة ، على فرض أن احتمال نجاح أى زهرة = $\frac{١}{٢}$ واحتمال فشلها $\frac{١}{٢}$ أيضاً .

$$\therefore \quad \text{ص} = {}^{\text{٢ن}}\text{ق}_{\text{ن}+\text{س}} \left(\tfrac{١}{٢}\right)^{\text{ن}+\text{س}} \left(\tfrac{١}{٢}\right)^{\text{ن}-\text{س}} \qquad (١)$$

ولو ضربنا هذا الاحتمال فى $٢^{\text{٢ن}}$ حصلنا على التكرار .

$$\text{ص} = \frac{\underline{|\text{٢ن}}}{\underline{|\text{ن}+\text{س}}\ \underline{|\text{ن}-\text{س}}} \times \frac{١}{٢^{\text{٢ن}}} \qquad (٢)$$

$$\text{وبما أن } \underline{|\text{٢ن}} = \frac{٢^{\text{٢ن}}\ \underline{|\text{ن}}\ .\ \underline{|\text{ن}}}{\sqrt{\text{ن}\ .\ \text{ط}}} \text{ تقريباً}$$

(حيث ط = النسبة التقريبية المعروفة ، بفرض ن كبيرة[1])

$$\therefore \text{ص} = \frac{١}{\sqrt{\text{ن ط}}} \times \frac{\left(١-\frac{١}{\text{ن}}\right)\left(١-\frac{٢}{\text{ن}}\right)\ldots\left(١-\frac{\text{س}-١}{\text{ن}}\right)}{\left(١+\frac{١}{\text{ن}}\right)\left(١+\frac{٢}{\text{ن}}\right)\ldots\left(١+\frac{\text{س}}{\text{ن}}\right)} \quad \ldots\ldots \quad (٣)$$

$$\therefore \text{لو}\left(\text{ص}\sqrt{\text{ن ط}}\right) = \text{لو}\left(١-\frac{١}{\text{ن}}\right) - \text{لو}\left(١+\frac{١}{\text{ن}}\right) + \text{لو}\left(١-\frac{٢}{\text{ن}}\right)$$

$$- \text{لو}\left(١+\frac{٢}{\text{ن}}\right) + \ldots$$

---

(١) هذه المعادلة وضعها (Wallis) وهى صحيحة إذا أهملنا الحدود التى تحتوى على $\frac{١}{\text{ن}}$.
انظر كتاب (Bowley, *Elements of Statistics*, p. 434.) .

$$+\,لو\left(1-\frac{س}{ن}\right)-لو\left(1+\frac{س}{ن}\right)-لو\left(1-\frac{س}{ن}\right).$$

$$=-2\left(\frac{1}{ن}+\frac{1}{3ن^{3}}+\ldots\right)-2\left(\frac{2}{ن}+\frac{8}{3ن^{3}}+\ldots\right)+\ldots$$

$$-2\left(\frac{س}{ن}+\frac{س^{3}}{3ن^{3}}+\ldots\right)-لو\left(1-\frac{س}{ن}\right).$$

$$=-2\cdot\frac{1+2+3+\ldots+س}{ن}-$$

$$-2\cdot\frac{1+2^{3}+3^{3}+\ldots+س^{3}}{3ن^{3}}-\ldots$$

$$-\frac{2}{2ر+1}\cdot\frac{1^{2ر+1}+2^{2ر+1}+\ldots+س^{2ر+1}}{ن^{2ر+1}}$$

$$-\,لو\left(1-\frac{س}{ن}\right).$$

وبما أن [1] $\dfrac{1^{ر}+2^{ر}+3^{ر}+\ldots+س^{ر}}{س^{ر+1}}=\dfrac{1}{ر+1}+\dfrac{1}{2س}+\dfrac{ر}{12س^{2}}$

$$+\ldots \qquad (٤)$$

$$\therefore\ لو\left(ص\sqrt{نط}\right)=-\frac{س(س+1)}{ن}-\frac{2}{3}\cdot\frac{س^{2}(س+1)^{2}}{4ن^{3}}-\ldots$$

$$-\frac{2}{2ر+1}\cdot\frac{س^{2ر+2}+\ldots}{(2ر+2)ن^{2ر+1}}$$

$$+\left(\frac{س}{ن}+\frac{س^{3}}{3ن^{2}}+\ldots\right).$$

$\therefore$ $=-\dfrac{س^{2}}{ن}+$ حدود تحتوى على $\dfrac{1}{\sqrt{ن}}$ أو قوى أعلى

---

(١) أنظر Whittaker and Robinson: *Calculus of Observations* (1930), p. 138

ولكن الانحراف المعيارى لعدد مرات النجاح فى هذه المسألة هو ع حيث

$$\text{ع}^{2} = 2\,\text{ن} \times \tfrac{1}{2} \times \tfrac{1}{2} = \tfrac{1}{2}\,\text{ن}$$

∴ $\text{لو}\left(\text{ص}.\text{ع}\sqrt{2\text{ط}}\right) = -\tfrac{1}{2} \cdot \frac{\text{س}^{2}}{\text{ع}^{2}}$ بعد إهمال $\frac{1}{\sqrt{\text{ن}}}$ وقواها العالية .

$$\therefore \quad \text{ص} = \frac{1}{\text{ع}\sqrt{2\text{ط}}} \cdot \text{هـ}^{-\frac{\text{س}^{2}}{2\text{ع}^{2}}} \qquad (٥)$$

= احتمال أن عدد الزهرات الناجحة يساوى ن + س

وبما أن احتمال النجاح يساوى احتمال الفشل ، ينتج أن احتمال نجاح ن + س من الزهرات يساوى تماماً احتمال نجاح ن — س منها .

$$\therefore \quad \text{ص} = \frac{1}{\text{ع}\sqrt{2\text{ط}}} \cdot \text{هـ}^{-\frac{2}{2\text{ع}^{2}}}$$

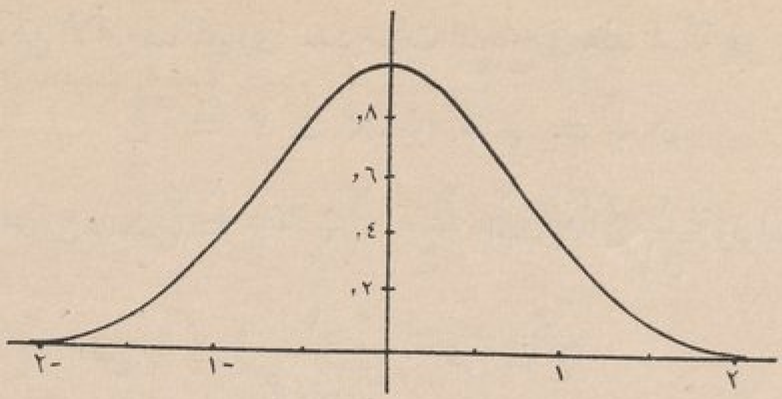
= احتمال أن عدد الزهرات الناجحة يزيد ، أو يقل ، عن المتوسط وهو ن بمقدار يساوى س (على فرض أن ن عدد كبير) .

وإذا جعلنا ع = ١ ، أو بعبارة أخرى ، إذا أخذنا وحدة قياس السينات تساوى ع ، تؤول المعادلة إلى صورة أبسط وهى :

$$\text{ص} = \frac{1}{\sqrt{2\text{ط}}} \cdot \text{هـ}^{-\frac{\text{س}^{2}}{2}} \qquad (٦)$$

والمنحنى البيانى لهذه المعادلة متماثل بالنسبة إلى محور الصادات كما نرى فى شكل ٧٢ ، ويسمى المنحنى الاحتمالى أو المنحنى المعتدل .

وبما أن دالة الاحتمال ودالة التكرار لا تختلفان عن بعضهما إلا فى وجود كمية ثابتة مضروبة أو مقسوم عليها ، نرى أنه من الممكن اعتبار المعادلة (٥) أو (٦) بعد ضربها فى كمية ثابتة مثل ص. ، تمثل المنحنى التكرارى المعتدل .



شكل **(٧٢)**
منحن معتدل متماثل بالنسبة إلى محور ص

و إذا غيرنا نقطة الأصل إلى (ا ، ٠) نحصل على الصورة العامة لمعادلة المنحنى التكرارى المتماثل وهى :

$$\text{ص} = \frac{\text{ص}_{\cdot}}{\sqrt{\text{٢ ط}}} \cdot \text{هـ}^{-\frac{(\text{س}-\text{ا})^{٢}}{\text{٢ ع}^{٢}}} \qquad (٧)$$

حيث ص. = مجموع التكرارات .

و ع = الانحراف المعيارى للتوزيع التكرارى ،

و ا = الوسط الحسابى للتوزيع التكرارى ، ويساوى أيضاً المنوال والوسيط ؛

و ط = النسبة التقريبية المعتادة .

ويمكننا إثبات خواص هذه الثوابت كما يأتى :

(١) إيجاد مساحة المنحنى .

$$\text{مساحة المنحنى} = \int_{-\infty}^{+\infty} \text{ص} \cdot \text{ءس} \cdot$$

$$= \frac{\text{ص}_{\cdot}}{\text{ع} \sqrt{\text{٢ ط}}} \int_{-\infty}^{+\infty} \text{هـ}^{-\frac{(\text{س}-\text{ا})^{٢}}{\text{٢ ع}^{٢}}} \cdot \text{ءس} \cdot$$

وبوضع $\frac{\text{س}-\text{ا}}{\sqrt{\text{٢}} \cdot \text{ع}} = \text{ل}$ مثلا ، $\text{ءس} = \sqrt{\text{٢}} \cdot \text{ع} \cdot \text{ءل}$

$$\therefore \text{المساحة} = \frac{\text{ص}_{\cdot}}{\sqrt{\text{ط}}} \int_{-\infty}^{+\infty} \text{هـ}^{-\text{ل}^{٢}} \,.\, \text{ء ل}$$

$$= \frac{٢\text{ص}_{\cdot}}{\sqrt{\text{ط}}} \int_{0}^{\infty} \text{هـ}^{-\text{ل}^{٢}} \,.\, \text{ء ل}$$

وبوضع $\text{ل}^{٢} = \text{س} \therefore \text{ء س} = ٢\,\text{ل} \,.\, \text{ء ل} = ٢\,\text{س}^{\frac{١}{٢}} \,.\, \text{ء ل}$

$$\therefore \text{المساحة} = \frac{\text{ص}_{\cdot}}{\sqrt{\text{ط}}} \int_{0}^{\infty} \text{س}^{\frac{١}{٢}-١}\, \text{هـ}^{-\text{س}} \,.\, \text{ء س}$$

$$= \frac{\text{ص}_{\cdot}}{\sqrt{\text{ط}}} \times \Gamma\left(\tfrac{١}{٢}\right)$$

$$= \text{ص}_{\cdot} \qquad (٨)$$

حيث $\Gamma(\text{ن})$ هى دالة جاما (Gamma Function) . ومعلوم[٧] أن $\Gamma\left(\tfrac{١}{٢}\right) = \sqrt{\text{ط}}$.

(٢) إيجاد المنوال

المنوال هو قيمة س ذات التكرار الأكبر، أى التى تجعل ص نهاية عظمى :

نفاضل المعادلة (٧) بالنسبة إلى س ونضع النتيجة تساوى صفراً .

$$\therefore \frac{\text{ء ص}}{\text{ء س}} = -\,\text{ح} \,.\, \frac{٢(\text{س}-١)}{٢\,\text{ع}^{٢}} \,.\, \text{هـ}^{-\frac{(\text{س}-١)^{٢}}{٢\,\text{ع}^{٢}}}$$

$= $ صفراً حينما $\text{س} = ١$.

$$\therefore \text{المنوال} = ١ \qquad (٩)$$

---

(١) $\Gamma(\text{ن}+١) = \underline{|\text{ن}}$ فى حالة ما تكون ن عدداً صحيحاً موجباً . وعلى العموم سواء كانت ن هكذا أو لا فإن $\Gamma(\text{ن}+١) = \text{ن}\,\Gamma(\text{ن})$. انظر كتاب
H. Lamb : *Infinitesimal Calculus*, 1929 p. 230, Ex. 15.

(٣) إيجاد الوسط الحسابى

الوسط الحسابى هو مجموع حواصل ضرب القيم كل فى تكرارها ، مقسوماً على مجموع التكرارات كلها . وهذا يساوى عزم المنحنى حول محور الصادات مقسوماً على مساحة المنحنى الكلية ، كما هو واضح من شكل ٧١

$$\therefore \text{الوسط الحسابى} = \frac{\sum (\text{ص} . \triangle \text{س}) . \text{س}}{\sum \text{ص} . \triangle \text{س}}$$

$$= \frac{\int_{-\infty}^{+\infty} \text{ص} . \text{س} . \text{ءس}}{\int_{-\infty}^{+\infty} \text{ص} . \text{ءس}}$$

$$= \frac{1}{\text{ع} \sqrt{2 \text{ط}}} \int_{-\infty}^{+\infty} \text{س} . \text{هـ}^{-\frac{(\text{س} - \text{ا})^2}{2 \text{ع}^2}} . \text{ءس}$$

$$= \frac{1}{\text{ع} \sqrt{2 \text{ط}}} \int_{-\infty}^{+\infty} (\text{س} - \text{ا} + \text{ا}) . \text{هـ}^{-\frac{(\text{س} - \text{ا})^2}{2 \text{ع}^2}} . \text{ءس}$$

$$= \frac{\text{ا}}{\text{ع} \sqrt{2 \text{ط}}} \int_{-\infty}^{+\infty} \text{هـ}^{-\frac{(\text{س} - \text{ا})^2}{2 \text{ع}^2}} . \text{ءس}$$

$$+ \frac{1}{\text{ع} \sqrt{2 \text{ط}}} \int_{-\infty}^{+\infty} (\text{س} - \text{ا}) . \text{هـ}^{-\frac{(\text{س} - \text{ا})^3}{2 \text{ع}^2}} . \text{ءس}$$

وبما أن التكامل الأخير يساوى صفراً (لأنه دالة فردية يمثلها منحن نصفه فوق محور السينات ونصفه الآخر تحت محور السينات ، فيكون مجموع مساحتيهما صفراً) يبقى فقط التكامل الأول .

و بنفس الطريقة التى اتبعناها فى (١) نثبت أن التكامل الأول يساوى ا

$$\therefore \text{الوسط الحسابى} = \text{ا} \quad . \quad . \quad . \quad . \qquad (10)$$

(٤) إيجاد الوسيط

حيث إن المنحنى متماثل حول محور رأسى يمر بقمتـه (وإحداثيها الأفقى يساوى ا = المنوال = الوسط الحسابى) فينتج أن نصف مسـاحة المنحنى واقع بعد س = ا ، والنصف الآخر قبل س = ا . أى أن الوسيط يسـاوى ا .

(٥) إيجاد الانحراف المعيارى

نعلم أن مربع الانحراف المعيارى يسـاوى مجموع حواصل ضرب مربعات الانحرافات عن الوسط الحسـابى فى التكرارات ، مقسوماً على مجموع التكرارات . وبما أن الوسط الحسابى هنا يسـاوى ا ، ومساحة المنحنى تساوى ص. تساوى مجموع التكرارات ،

∴ مربع الانحراف المعيارى

$$= \frac{\text{مج}\,(\text{س} - \text{ا})^{٢} \,.\, \text{ص} \,\triangle\, \text{س}}{\text{مساحة المنحنى}} = \frac{١}{\text{ص.}} \int_{-\infty}^{+\infty} \text{ص}\,(\text{س} - \text{ا})^{٢} \,.\, \text{ءس}$$

$$= \frac{١}{\text{ع}\,\sqrt{٢\,\text{ط}}} \int_{-\infty}^{+\infty} (\text{س} - \text{ا})^{٢} \times \text{هـ}^{-\frac{(\text{س} - \text{ا})^{٢}}{٢\,\text{ع}^{٢}}} \,.\, \text{ءس}$$

وبوضع $\text{م} \,.\, \sqrt{٢} \,.\, \text{ع} = \text{س} - \text{ا}$ و $\text{ءس} = \sqrt{٢}\,\text{ع} \,.\, \text{ءم}$

∴ مربع الانحراف المعيارى .

$$= \frac{٢\,\text{ع}^{٢}}{\sqrt{\text{ط}}} \int_{-\infty}^{+\infty} \text{م}^{٢} \,.\, \text{هـ}^{-\text{م}^{٢}} \,.\, \text{ءم}$$

$$= \frac{٤\,\text{ع}^{٢}}{\sqrt{\text{ط}}} \int_{.}^{\infty} \text{م}^{٢}\, \text{هـ}^{-\text{م}^{٢}} \,.\, \text{ءم}$$

وبوضع $\text{م}^{٢} = \text{ل}$ ى $\text{ءل} = ٢ \,.\, \text{ل}^{\frac{١}{٢}} \,.\, \text{ءم}$.

$$\therefore \quad = \frac{٢\,ع^{٢}}{\sqrt{ط}} \int_{.}^{\infty} ل^{\frac{١}{٢}} \,.\, هـ^{-ل} \,.\, ى\,ل$$

$$= \frac{٢\,ع^{٢}}{\sqrt{ط}} \,\Gamma\left(\tfrac{٣}{٢}\right) .$$

$$= ع^{٢}$$

لأن $\Gamma\left(\tfrac{٣}{٢}\right) = \tfrac{١}{٢} \,.\, \Gamma\left(\tfrac{١}{٢}\right) = \tfrac{١}{٢} \sqrt{ط}$

∴ الانحراف المعيارى = ع .

منحنى الخطأ الطبيعي أو المعتدل

من أهم التطبيقات ، بل من أولى المسائل التى استخدم فيها المنحنى التكرارى المعتدل ، استخدام معادلتـه لتمثيل التوزيع التكرارى للأخطاء التى تقع فى قياس بعض الظواهر الطبيعية أو الفلكية . فمن المعلوم أنه عند قيـاس طول معين مثلاً يخطىء الإنسـان فى قياسه فيقدره بأقل من الحقيقـة أو بأكثر ، أى أن الخطأ يكون أحياناً موجباً وأحياناً سـالباً . ومن الطبيعي أنه إذا لم يكن هناك تحيز إلى ناحية خاصة فإن الأخطاء السالبة تكون فى عددها تساوى الأخطاء الموجبة ، وتكون كلها موزعة توزيعاً تكرارياً متماثلا حول الوسط الحسابى لهـا .

و إذا وضعنا ا = صفراً فى معادلة المنحنى التكرارى المعتدل ، وعبرنا بالرمز س عن مقدار الخطأ الواقع فى القياس تكون معادلة المنحنى التكرارى للأخطاء ، أى معادلة منحنى الخطأ الطبيعى هى :

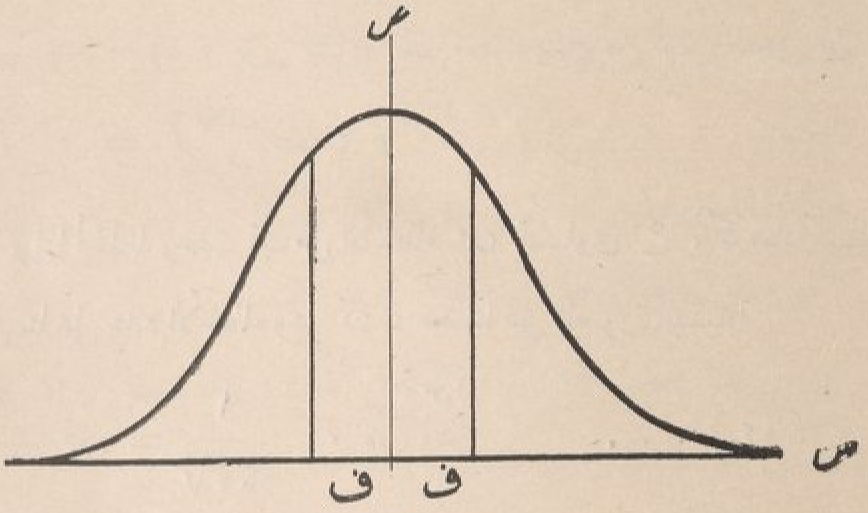
$$ص = \frac{ن}{ع\sqrt{٢\,ط}} \,.\, هـ^{-\frac{س^{٢}}{٢\,ع^{٢}}}$$

حيث ﮟ = عدد الأخطاء كلها

و س = مقدار الخطأ ،

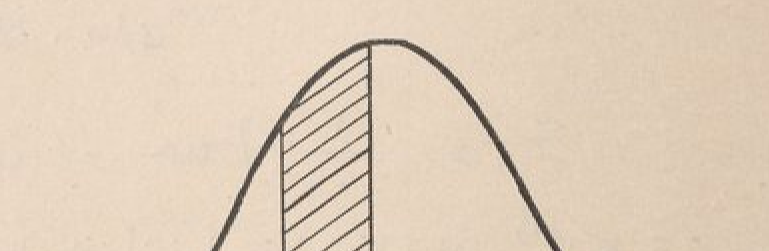
و ع = الانحراف المعياري للأخطاء .

وإذا رسمنا هذا المنحنى نجده متماثلا حول محور الصادات ، أي أن الوسط الحسابي للأخطاء = صفراً ، كما في الشكل رقم (٧٣) .



شكل (٧٣)

وقد عرفنا فيما تقدم أن عدد الأخطاء التي تقع بين س ، س + △ س هو ص. △ س بشرط ألا تكون △ س كبيرة فيكون الشريط عريضاً ، وإذا



شكل (٧٤)

المطلوب معرفة عدد الأخطاء التي تنحصر بين $س_1$ ، $س_2$ ، حيث $س_2 - س_1$ كبير ، يجب علينا أن نحسب المساحة من $س_1$ إلى $س_2$ وهي :

$$= \int_{س_1}^{س_2} ص \cdot ء س$$

ونعلم أن مساحة المنحنى كلها $= \int_{-\infty}^{+\infty} ص \cdot ء س = ن =$ عدد الأخطاء كلها

وإذا أخذنا $س_1$ و $س_2$ على جانبي المحور ومتساويتي البعد عنه ، أي أن $س_1 = -ف$ ، $س_2 = +ف$ مثلا ، ينتج من تماثل المنحنى أن

$$\int_{س_1}^{س_2} = \int_{-ف}^{+ف} = ٢ \int_{0}^{ف} ص \cdot ء س \quad \ldots \quad (١٣)$$

وإذا أخذنا وحدة قياس الأخطاء س تساوي ع ، كما فعلنا سابقاً ، ثم قسمنا على عدد الأخطاء وهو ن ، حصلنا على الصورة البسيطة

$$ص' = \frac{١}{\sqrt{٢ط}} \cdot هـ^{-\frac{س^2}{٢}} \quad \ldots \quad (١٤)$$

حيث س تدل على مقدار الخطأ مقيساً بوحدات كل منها يساوي الانحراف المعياري ، و $ص'$ تدل على احتمال حصول هذا الخطأ س (سالباً كان أو موجباً) .

∴ احتمال حصول خطأ لا يزيد مقداره عن المقدار ف مثلا ، بصرف النظر عن الإشارة ، يساوي

$$\int_{-ف}^{+ف} ص' \cdot ء س = ٢ \int_{0}^{ف} \frac{١}{\sqrt{٢ط}} \cdot هـ^{-\frac{س^2}{٢}} \cdot ء س \quad \ldots \quad (١٥)$$

ومن المهم إذن أن نعرف مقدار التكامل الأخير[1] لكل قيمة نأخذها ف

---

(١) توجد جداول مختصرة في آخر الكتاب (ص ٥٥٣) لقيم هذا التكامل (مقسومة على ٢) . وتوجد جداول أخرى مفصلة لهذا التكامل وغيره في كتاب
Rietz, H. *Handbook of Mathematical Statistics*, pp. 209—219,
وفي كتاب : *Tables for Statisticians and Biometricians*, Part I.

وقد اجتهد كثيرون فى عمل جداول لقيم هذا التكامل حسب قيم ف .

ويوجد أيضاً جداول للتكامل المتمم لهذا وهو

$$\frac{٢}{\sqrt{٢\text{ط}}}\int_{\text{ف}}^{\infty} \text{هـ}^{-\frac{\text{س}^٢}{٢}} . \text{ءس}$$

ومعلوم أن

$$٢\left(\int_{٠}^{\text{ف}} + \int_{\text{ف}}^{\infty}\right) = ٢\int_{٠}^{\infty} \frac{١}{\sqrt{٢\text{ط}}} . \text{هـ}^{-\frac{\text{س}^٢}{٢}} \text{ءس}$$

$$= ١$$

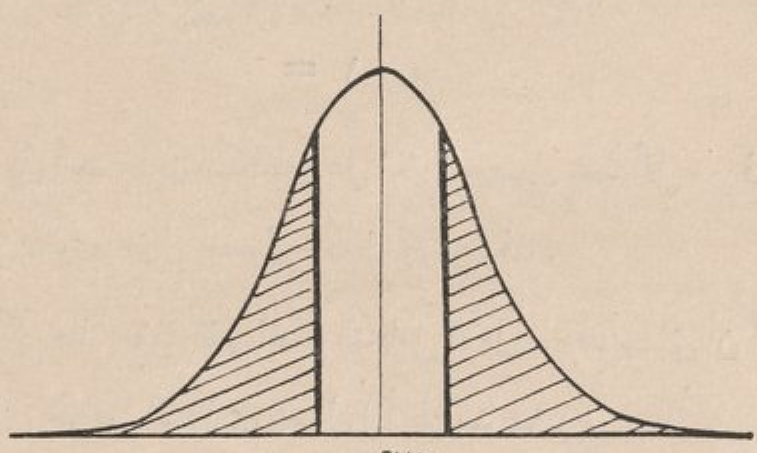
فإذا دل أحد هذين التكاملين على احتمال حصول خطأ أقل من ف عددياً ، فإن الآخر يساوى احتمال حصول خطأ أكبر من ف .

ونذكر هنا بعض القيم المهمة للتكامل الأول ، وقد سبق أن أشرنا إليها ضمن المنحنى خواص التكرارى المعتدل (أنظر بند ٣٩٤)

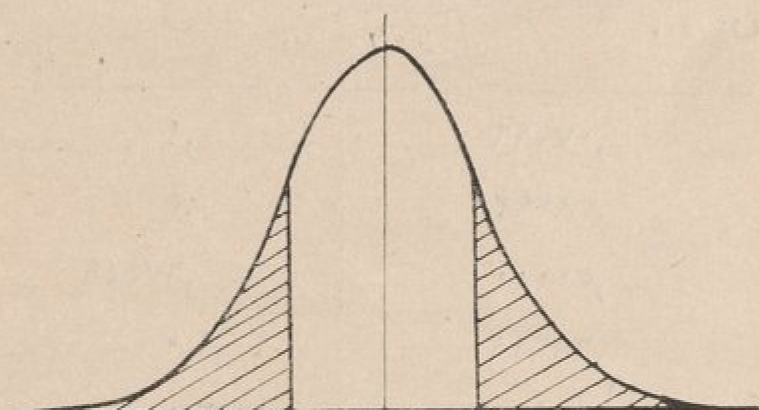
ف	$\frac{٢}{\sqrt{٢\text{ط}}}\int_{٠}^{\text{ف}} \text{هـ}^{-\frac{\text{س}^٢}{٢}} . \text{ءس}$
ر١	ر٠٧٩٦٦
ر٥	ر٣٨٢٩٢
ر٦٧٤٤٩	ر٥٠٠٠٠
١ر٠	ر٦٨٢٦٨
٢ر٠	ر٩٥٤٥٠
٣ر٠	ر٩٩٧٣٠
٤ر٠	ر٩٩٩٩٤

ويلاحظ أن الوحدة المقيس بها الخطأ ف هى الانحراف المعيارى كما قلنا من قبل.

ومعنى ذلك أن ٥٠٪ من الأخطاء لا يزيد مقدارها العددى على ٦٧٤٤٩ر من الانحراف المعيارى ، وأن ٦٨٫٢٦٨٪ منها لا يزيد عن نفس الانحراف المعيارى ، و ٩٥٫٤٥ ٪ منها لا يزيد على ضعفه ، و ٩٩٫٧٣ ٪ منها لا تزيد على ثلاثة أمثاله . وأخيراً ٩٩٫٩٩٤٪ منها لا يزيد على ٤ ع . وهذه هى المناطق التى سبق أن أشرنا إليها . ونكرر هنا أن هذه الحدود يشترط فيها أن التوزيع التكرارى للأخطاء معتدل ، أو قريب من الاعتدال .

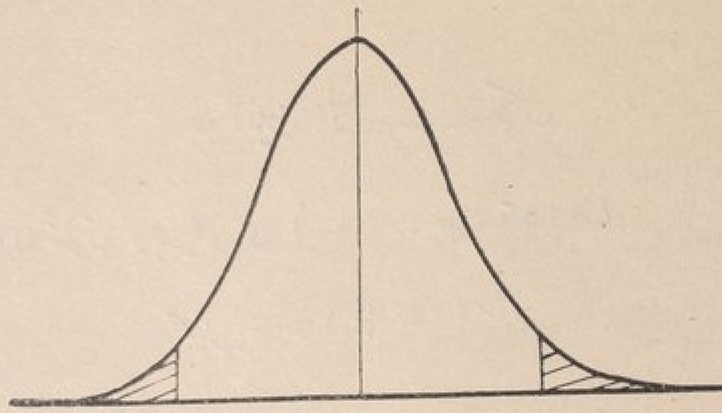


(شكل ٧٥)



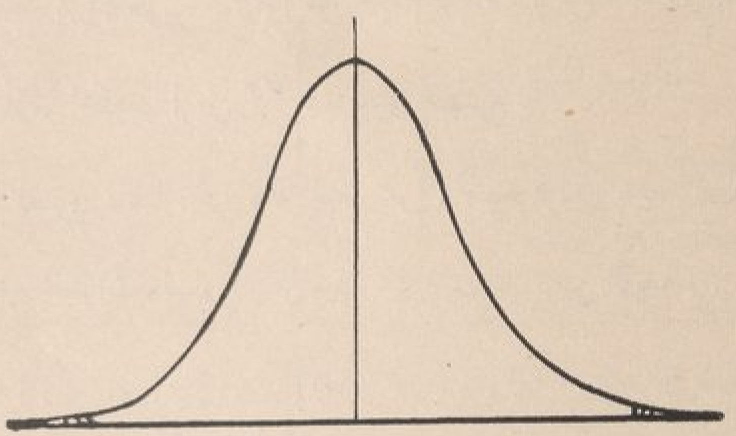
(شكل ٧٦)

ونرى ذلك موضحاً فى الأشكال ٧٥ و ٧٦ و ٧٧ و ٧٨ ؛ حيث تدل المساحة غير المظللة على قيمة التكامل الأول ، وتدل المساحة المظللة على قيمة التكامل المتمم له .



٢

(شكل ٧٧)



٣

(شكل ٧٨)

وفى هـذه الأشكال نرى أن المساحة غير المظللة تساوى على الترتيب ٥٠٪ و ٣ر٦٨٪ و ٥ر٩٥٪ و ٧ر٩٩٪ من المساحة الكلية للمنحنى .

قوانين الاستكمال

قانون نيوتن وجريجورى (بند ٤٧٢ صفحة ٤٩٥)

$$ص_{ر} = (١ + \triangle)^{ر}\, ص \, .$$

يمكننا إثبات هذا القانون كما يأتى :

تمهيد : نأخذ المقدار

$$س(س - ١)(ص - ٢)(س - ٣) \cdots (س - ر + ١)$$

ونرمز له بالرمز المختصر $[س]^{ر}$ ، ويسمى **مضروباً**(١) .

هذا المضروب عبارة عن دالة للمتغير س . نوجد فروق هذه الدالة ، باعتبار أن س تتغير بخطوات متساوية كل منها = ١ ، ابتداء من القيمة ا :

$$[ا]^{ر} = ا(ا - ١)(ا - ٢) \ldots (ا - ر + ١)$$

$$\therefore [ا + ١]^{ر} = (ا + ١)\, ا\, (ا - ١) \ldots (ا - ر + ٢)$$

$$\therefore \triangle\, [ا]^{ر} = [ا + ١]^{ر} - [ا]^{ر}$$

$$= ا(ا-١)(ا-٢) \ldots (ا-ر+٢)\{(ا+١) - (ا-ر+١)\}$$

$$= ا(ا-١)(ا-٢) \ldots (ا-ر+٢) \,.\, ر$$

$$= ر\,[ا]^{ر-١} \qquad (١)$$

$$\therefore \frac{\triangle\,[س]^{ر}}{\lfloor\!\underline{ر}} = \frac{[س]^{ر-١}}{\lfloor\!\underline{ر-١}} \qquad (٢)$$

---

(١) بالإنجليزية Factorial

$$\text{أو}\quad \frac{[\text{س}+1]^{\text{ر}}}{\underline{|\text{ر}}} = \frac{[\text{س}]^{\text{ر}}}{\underline{|\text{ر}}} + \frac{[\text{س}]^{\text{ر}-1}}{\underline{|\text{ر}-1}}$$

ويتضح من معادلة (١) أن الفرق $\triangle^{\text{ن}}[\text{س}]^{\text{ن}}$ يساوى مقداراً خالياً من س ، وأن $\triangle^{\text{ن}+1}[\text{س}]^{\text{ن}} =$ صفراً .

نأخذ الآن أى متحددة[1] من درجة ن بالنسبة إلى س مثل
$\text{ص} = \text{ل}(\text{س}) = \text{ا}_{0}\text{س}^{\text{ن}} + \text{ا}_{1}\text{س}^{\text{ن}-1} + \text{ا}_{2}\text{س}^{\text{ن}-2} + \ldots + \text{ا}_{\text{ن}}$.

هـذه المتحددة يمكن وضعها فى صورة مجموع مضروبات [س] ، فيسهل حساب الفروق المتتالية لها باستخدام العـلاقة (١) ، مع العلم بأن الفروق من رتبة ن تكون ثابتة ومن رتبة أعلى تنعدم . لو كانت $\text{ل}_{\text{ر}}(\text{س})$ تمثل أى متحددة من الدرجة ر بالنسبة إلى س ، وقسمناها على (س − ن + ر) كان الخارج متحددة من درجة ر − ١ ، ولتكن $\text{ل}_{\text{ر}-1}(\text{س})$ مثلا ، والباقى حداً مطلقاً م مثلا وهكذا لو قسمنا $\text{ل}_{\text{ر}-1}(\text{س})$ على (س − ن + ر − ١) حصلنا على متحددة من درجة (ر − ٢) وحد مطلق آخر .

$$\therefore\ \text{ل}_{\text{ن}}(\text{س}) = \text{ب} + [\text{س}]^{1} . \text{ل}_{\text{ن}-1}(\text{س})$$
$$= \text{ب} + \text{حـ}[\text{س}]^{1} + [\text{س}]^{2}\,\text{ل}_{\text{ن}-2}(\text{س})$$
$$= \ldots\ldots$$
$$= \text{ب} + \text{حـ}[\text{س}]^{1} + \text{ء}[\text{س}]^{2} + \text{هـ}[\text{س}]^{3}$$
$$+\ \text{و}[\text{س}]^{4} + \ldots + \text{ك}[\text{س}]^{\text{ن}} ;\quad (4)$$

(١) ترجمة "Polynomial" . ومعناها دالة ذات عدة حدود ؛ ولذلك اقترح استخدام كلمة « متحددة » على وزن « متشـعبة » أى كثيرة الشعب ؛ لأن المتحددة هى دالة كثيرة الحدود .

حيث ب ، حـ ، . . . ك حدود مطلقة خالية من س .

نفرض الآن أن لدينا جـدولا به قيم للمتغير س وقيم للمتغـير التابع ص ، ونفرض أن ص = د (س) = متحددة من درجـة ن ، وأن قيم س على فترات متساوية كل منها تساوى ف مثلا ، أى أن قيم س و ص هى :

س = ١ ، ١ + ف ، ١ + ٢ ف ، ١ + ٣ ف ، ...

ص = د (١) ، د (١ + ف) ، د (١ + ٢ ف) ، د (١ + ٣ ف) ، ...

= ص. ، ص١ ، ص٢ ، ص٣ ، ...

والمطلوب معرفة قيمة ص التى تناظر قيمـــة المتغير س تساوى ١ + ر ف مثلا ، حيث ر كسر أقل من ١ . هـذه القيمة المطلوبة د (١ + ر ف) = صر ، وهى متحددة من درجة ن بالنسـبة إلى ر ، يمكن وضعها بدل ل (ر) فى المعادلة (٤)

∴ △ صر = حـ + ٢ ء [ر] + ٣ هـ [ر]^٢

+ ٤ و [ر]^٣ + ... + ن . ك . [ر]^(ن-١)

و △^٢ صر = ٢ ء + ٢ × ٣ . هـ [ر] + ..

+ ن (ن − ١) . ك . [ر]^(ن-٢)

و △^٣ صر = لـ٣ . هـ + ٤ × ٣ × ٢ [ر] + ..

+ ن (ن − ١) (ن − ٢) . ك . [ر]^(ن-٣)

ونحصل على قيم ب ، حـ ، ء ... ك بوضع ر = ٠ فينتج أن

ب = د (١) = ص.

حـ = △ ص.

$$ء = \frac{1}{2} \triangle^{2} ص_{.}$$

$$هـ = \frac{1}{6} \triangle^{3} ص_{.}$$

. . . .

$$ك = \frac{1}{\underline{|ن}} \triangle^{ن} ص_{.}$$

$$\therefore ص_{م} = د\,(1 + م ف)$$

$$= ص_{.} + \triangle ص_{.}\,[م] + \frac{1}{2} \triangle^{2} ص_{.}\,[م]^{2}$$

$$+ \frac{1}{\underline{|3}} \triangle^{3} ص_{.}\,[م]^{3} + \cdots\cdot$$

$$+ \frac{1}{\underline{|ن}} \triangle^{ن} ص_{.}\,[م]^{ن}\,.$$

$$= ص_{.} + م\,.\,\triangle ص_{.} + \frac{م\,(م-1)}{2}\,.\,\triangle^{2} ص_{.}$$

$$+ \frac{م\,(م-1)\,(م-2)}{\underline{|3}} \triangle^{3} ص_{.} + \cdots$$

$$+ \frac{م(م-1)(م-2)\cdots(م-ن+1)}{\underline{|ن}}\,.\,\triangle^{ن} ص_{.}$$

. . . . . . (٥)

ولا لزوم للاستمرار بعد هذا الحد لأن الفروق من رتبة أعلى من ن تساوى صفراً ، حيث قد فرضنا أن ص متحددة من رتبة ن .

ولكننا فى مسائل الاستكمال لا نعلم شيئاً عن شكل الدالة ص = د (س) ، ولا عن درجتها ؛ وكل ما نعرفه أن الفروق من رتبة معينة متساوية أو منعدمة . فإذا وجدنا فى جدول الفروق لدالة ما أن الفروق الخامسة مثلا تساوى صفراً ، وأردنا أن نستكمل هذا الجدول لمعرفة $ص_{ر} = ص_{٢٫٧}$ مثلا ، قلنا إن الدالة ص = د (س) يمثلها منحن ممهد فى الفترة السينية بين س = ٢ و س = ٣ ، ونفرض للسهولة أن هذا المنحنى تمثله متحددة من الدرجة الرابعة بالنسبة إلى س ، تأخذ القيم المعلومة فى الجدول وهى :

د (ا) ، د (ا + ب) ، د (ا + ٢ ب) ، د (ا + ٣ ب) ،...

أى $ص_{٠}$ ، $ص_{١}$ ، $ص_{٢}$ ، $ص_{٣}$ ، ...

مثلا ، حينما س تأخذ القيم :

ا ، ا + ب ، ا + ٢ ب ، ا + ٣ ب ،...

وبواسطة النتيجة (٥) يمكننا كتابة هذه المتحددة فوراً وهى :

$$ص_{ر} = ص_{٠} + ر \,\triangle\, ص_{٠} + \frac{ر(ر-١)}{\underline{|٢}} \triangle^{٢} ص_{٠}$$

$$+ \frac{ر(ر-١)(ر-٢)}{\underline{|٣}} \triangle^{٣} ص_{٠} + \dots$$

$$+ \frac{ر(ر-١)(ر-٢)\cdots(ر-ن+١)}{\underline{|ن}} \cdot \triangle^{ن} ص_{٠}$$

$$= (١ + \triangle)^{ر} ص_{٠}$$

حيث نَفُكّ القوس $(١ + \triangle)^{ر}$ حسب نظرية ذات الحدين ، ونضع كل الفروق التى من رتبة أعلى من ن تساوي صفراً . وبذلك تثبت صحة القانون

$$ص_{ر} = (+ \triangle)^{ر} ص_{٠}$$

قانون نيوتن للفروق المنقسمة (بند ٤٨٠ صفحة ٥٠٥)

لنفرض أن القيم السينية والصادية الموجودة فى الجدول هى على الترتيب :

$$س = ا_{٠} ، ا_{١} ، ا_{٢} ، ا_{٣} ، ا_{٤} \ldots$$

$$ص = د(ا_{٠}) ، د(ا_{١}) ، د(ا_{٢}) ، د(ا_{٣}) ، د(ا_{٤}) \ldots$$

و أن الفروق المنقسمة من الرتبة الرابعة منعدمة ، أو تساوى صفراً تقريباً .
أى أن الفروق المنقسمة الثالثة متساوية .

$$\therefore د(ا_{٠}، ا_{١}، ا_{٢}، ا_{٣}) = د(ا_{١}، ا_{٢}، ا_{٣}، ا_{٤}) = د(ا_{٢}، ا_{٣}، ا_{٤}، ا_{٥}) = \ldots$$

$$\therefore د(س، ا_{٠}، ا_{١}، ا_{٢}) = د(ا_{٠}، ا_{١}، ا_{٢}، ا_{٣})$$

وبناء على تعريف الفروق المنقسمة من الرتبة ٢

$$\therefore د(س، ا_{٠}، ا_{١}) = د(ا_{٠}، ا_{١}، ا_{٢}) + (س - ا_{٢}) \cdot د(س، ا_{٠}، ا_{١}، ا_{٢}) .$$

وبناء على تعريف الفروق من رتبة ١

$$\therefore د(س، ا_{٠}) = د(ا_{٠}، ا_{١}) + (س - ا_{١}) \cdot د(س، ا_{٠}، ا_{١}، ا_{٢})$$

$$\therefore \quad = د(ا_{٠}، ا_{١}) + (س - ا_{١})(س - ا_{٢}) \cdot د(ا_{٠}، ا_{١}، ا_{٢}، ا_{٣})$$

$$ولكن\ د(س) = د(ا_{٠}) + (س - ا_{٠}) \cdot د(س، ا_{٠})$$

$$\therefore \quad = د(ا_{٠}) + (س - ا_{٠}) \cdot د(ا_{٠}، ا_{١})$$

$$+ (س - ا_{٠})(س - ا_{١}) \cdot د(ا_{٠}، ا_{١}، ا_{٢})$$

$$+ (س - ا_{٠})(س - ا_{١})(س - ا_{٢}) .$$

$$د(ا_{٠}، ا_{١}، ا_{٢}، ا_{٣}) :$$

وهذا هو قانون نيوتن فى حالة الفروق الثالثة ، ويمكن تعميم هـذا البرهان ونحصل على القانون المعروف .

قانون لاجرانج للفروق المنقسمة (بند ٤٨٢ صحيفة ٥٠٨)

نثبت هذا القانون كما يأتى :

نعلم أن

$$\text{د}(\text{ا}_{٠}،\text{ا}_{١}) = \frac{\text{د}(\text{ا}_{١}) - \text{د}(\text{ا}_{٠})}{\text{ا}_{١} - \text{ا}_{٠}}$$

$$= \frac{\text{د}(\text{ا}_{١})}{\text{ا}_{١} - \text{ا}_{٠}} + \frac{\text{د}(\text{ا}_{٠})}{\text{ا}_{٠} - \text{ا}_{١}}$$

وبالمثل

$$\text{د}(\text{ا}_{١}،\text{ا}_{٢}) = \frac{\text{د}(\text{ا}_{٢})}{\text{ا}_{٢} - \text{ا}_{١}} + \frac{\text{د}(\text{ا}_{١})}{\text{ا}_{١} - \text{ا}_{٢}}$$

وهكذا .

وكذلك $\text{د}(\text{ا}_{٠}،\text{ا}_{١}،\text{ا}_{٢})$

$$= \frac{\text{د}(\text{ا}_{١}،\text{ا}_{٢}) - \text{د}(\text{ا}_{٠}،\text{ا}_{١})}{\text{ا}_{٢} - \text{ا}_{٠}}$$

$$= \frac{\text{د}(\text{ا}_{٠})}{(\text{ا}_{٠} - \text{ا}_{١})(\text{ا}_{٠} - \text{ا}_{٢})} + \frac{\text{د}(\text{ا}_{١})}{(\text{ا}_{١} - \text{ا}_{٠})(\text{ا}_{١} - \text{ا}_{٢})}$$

$$+ \frac{\text{د}(\text{ا}_{٢})}{(\text{ا}_{٢} - \text{ا}_{٠})(\text{ا}_{٢} - \text{ا}_{١})}$$

و $\text{د}(\text{ا}_{٠}،\text{ا}_{١}،\text{ا}_{٢}،\text{ا}_{٣})$

$$= \frac{\text{د}(\text{ا}_{٠})}{(\text{ا}_{٠} - \text{ا}_{١})(\text{ا}_{٠} - \text{ا}_{٢})(\text{ا}_{٠} - \text{ا}_{٣})} + \frac{\text{د}(\text{ا}_{١})}{(\text{ا}_{١} - \text{ا}_{٠})(\text{ا}_{١} - \text{ا}_{٢})(\text{ا}_{١} - \text{ا}_{٣})}$$

$$+ \frac{\text{د}(\text{ا}_{٢})}{(\text{ا}_{٢} - \text{ا}_{٠})(\text{ا}_{٢} - \text{ا}_{١})(\text{ا}_{٢} - \text{ا}_{٣})} + \frac{\text{د}(\text{ا}_{٣})}{(\text{ا}_{٣} - \text{ا}_{٠})(\text{ا}_{٣} - \text{ا}_{١})(\text{ا}_{٣} - \text{ا}_{٢})}$$

وهكذا للفروق من أى رتبة .

فإذا كان لدينا جدول ما ، وكانت الفروق المنقسمة من الرتبة ٣ متساوية مثلاً ، كانت الفروق من الرتبة الرابعة صفراً .

فلو أردنا استكمال هذا الجدول عنـد القيمة السينية س مثلا ، وفرضنا أن القيمة المطلوبة هى د (س) كان الفرق الرابع

$$\text{د}(\text{س}، \text{ا}_{٠}، \text{ا}_{١}، \text{ا}_{٢}، \text{ا}_{٣}) = ٠$$

ولـكن هذا الفرق يمكن وضعه فى صورة مثل المذكورة أعلاه فهو يساوى :

$$\frac{\text{د}(\text{س})}{(\text{س}-\text{ا}_{٠})(\text{س}-\text{ا}_{١})(\text{س}-\text{ا}_{٢})(\text{س}-\text{ا}_{٣})} + \frac{\text{د}(\text{ا}_{٠})}{(\text{ا}_{٠}-\text{س})(\text{ا}_{٠}-\text{ا}_{١})(\text{ا}_{٠}-\text{ا}_{٢})(\text{ا}_{٠}-\text{ا}_{٣})}$$

$$+ \frac{\text{د}(\text{ا}_{١})}{(\text{ا}_{١}-\text{س})(\text{ا}_{١}-\text{ا}_{٠})(\text{ا}_{١}-\text{ا}_{٢})(\text{ا}_{١}-\text{ا}_{١})} + \frac{\text{د}(\text{ا}_{٢})}{(\text{ا}_{٢}-\text{س})(\text{ا}_{٢}-\text{ا}_{٠})(\text{ا}_{٢}-\text{ا}_{١})(\text{ا}_{٢}-\text{ا}_{٣})}$$

$$+ \frac{\text{د}(\text{ا}_{٣})}{(\text{ا}_{٣}-\text{س})(\text{ا}_{٣}-\text{ا}_{٠})(\text{ا}_{٣}-\text{ا}_{١})(\text{ا}_{٣}-\text{ا}_{٢})}$$

$$= \text{صفراً}$$

وينتج من ذلك أن القيمة المطلوبة

$$\text{د}(\text{س}) = \frac{(\text{س}-\text{ا}_{١})(\text{س}-\text{ا}_{٢})(\text{س}-\text{ا}_{٣})}{(\text{ا}_{٠}-\text{ا}_{١})(\text{ا}_{٠}-\text{ا}_{٢})(\text{ا}_{٠}-\text{ا}_{٣})} \,.\, \text{د}(\text{ا}_{٠})$$

$$+ \frac{(\text{س}-\text{ا}_{٠})(\text{س}-\text{ا}_{٢})(\text{س}-\text{ا}_{٣})}{(\text{ا}_{١}-\text{ا}_{٠})(\text{ا}_{١}-\text{ا}_{٢})(\text{ا}_{١}-\text{ا}_{٣})} \,.\, \text{د}(\text{ا}_{١})$$

$$+ \frac{(\text{س}-\text{ا}_{٠})(\text{س}-\text{ا}_{١})(\text{س}-\text{ا}_{٣})}{(\text{ا}_{٢}-\text{ا}_{٠})(\text{ا}_{٢}-\text{ا}_{١})(\text{ا}_{٢}-\text{ا}_{٣})} \,.\, \text{د}(\text{ا}_{٢})$$

$$+ \frac{(\text{س}-\text{ا}_{٠})(\text{س}-\text{ا}_{١})(\text{س}-\text{ا}_{٢})}{(\text{ا}_{٣}-\text{ا}_{٠})(\text{ا}_{٣}-\text{ا}_{١})(\text{ا}_{٣}-\text{ا}_{٢})} \,.\, \text{د}(\text{ا}_{٣})$$

وهكذا نثبت القانون فى الحالة العامة حينما تكون الفروق المنعدمة من أى رتبة .

قانون نيوتن وجاوس للفروق الوسطى (بند ٤٨٤ صفحة ٥١٠)

نفرض أن قيم س المعطاة فى الجدول (على فترات متساوية تساوى ف) هى

. . . ، ا — ف ، ا ، ا + ف ، ا + ٢ ف ، . . .

وأن قيم ص المناظرة لها هى على الترتيب

. . . د (ا — ف) ، د (ا) ، د (ا + ف) ، د (ا + ٢ ف) ، . . .

نعوض ا$_{٠}$ = ا ، ا$_{١}$ = ا + ف ، ا$_{٢}$ = ا — ف ، ا$_{٣}$ = ا + ٢ ف ،

ا$_{٤}$ = ا + ٣ ف ، وهكذا ، فى قانون نيوتن للفروق المنقسمة وهو

د (س) = د (ا$_{٠}$) + (س — ا$_{٠}$) د (ا$_{٠}$ ، ا$_{١}$)

+ (س — ا$_{٠}$) (س — ا$_{١}$) . د (س . ا$_{٠}$ ، ا$_{١}$ ، ا$_{٢}$)

+ (س — ا$_{٠}$) (س — ا$_{١}$) (س — ا$_{٢}$) .

× د (ا$_{٠}$ ، ا$_{١}$ ، ا$_{٢}$ ، ا$_{٣}$)

+ . . .

حيث س = ا + ر ف

∴ د (ا + ر ف) = د (س) =

د (ا) + (س — ا) . د (ا ، ا + ف)

+ (س — ا) (س — ا — ف) . د (ا ، ا + ف ، ا — ف)

+ (س — ا) (س — ا — ف) (س — ا + ف) .

د (ا ، ا + ف ، ا — ف ، ا + ٢ ف)

+ (س — ا) (س — ا — ف) (س — ا + ف) (س — ا — ٢ ف)

× د (ا ، ا + ف ، ا — ف ، ا + ٢ ف ، ا — ٢ ف) + . . . (١)

ويمكن التعبير عن الفروق المنقسمة الموجودة فى هذه المعادلة بالفروق العادية كما يأتى :

$$\text{د}(\text{ا}\text{،}\ \text{ا}+\text{ف}) = \frac{١}{\text{ف}}\ \triangle\ \text{د}(\text{ا})$$

$$\text{و}\quad \text{د}(\text{ا}\text{،}\ \text{ا}+\text{ف}\text{،}\ \text{ا}-\text{ف}) = \frac{١}{\underline{|٢}\ .\ \text{ف}^{٢}}\ .\ \triangle^{٢}\ \text{د}(\text{ا}-\text{ف})$$

$$\text{د}(\text{ا}\text{،}\ \text{ا}+\text{ف}\text{،}\ \text{ا}-\text{ف}) = \frac{١}{\underline{|٣}\ .\ \text{ف}^{٣}}\ \triangle^{٣}\ \text{د}(\text{ا}-\text{ف})$$

وهكذا . فينتج أن المعادلة (١) تأخذ الصورة :

$$\text{د}(\text{ا}+\text{ر}\,\text{ف})$$

$$= \text{د}(\text{ا}) + \text{ر}\ .\ \triangle\ \text{د}(\text{ا})$$

$$+ \frac{\text{ر}(\text{ر}-١)}{\underline{|٢}}\ .\ \triangle^{٢}\ \text{د}(\text{ا}-\text{ف})$$

$$+ \frac{(\text{ر}+١)\text{ر}(\text{ر}-١)}{\underline{|٣}}\ .\ \triangle^{٣}\ \text{د}(\text{ا}-٢\,\text{ف})$$

$$+ \frac{(\text{ر}+١)\text{ر}(\text{ر}-١)(\text{ر}-٢)}{\underline{|٤}}\ .\ \triangle^{٤}\ \text{د}(\text{ا}-٢\,\text{ف})$$

$$+ \ldots$$

وبوضع $\text{ص}_{\text{ر}}$ بدل $\text{د}(\text{ا}+\text{ر}\,\text{ف})$ ، و$\text{ص}_{-١}$ بدل $\text{د}(\text{ا}-\text{ف})$ ،

وهكذا ، تنتج الصورة المعروفة .

قانون نيوتن وبسل (بند ٤٨٥ صحيفة ٥١٠)

فى قانون نيوتن وجاوس :

$$\text{ص}_{\text{ر}} = \text{ص}_{٠} + \text{ر} \cdot \triangle \text{ص}_{٠} + \tfrac{١}{٢}\,\text{ر}\,(\text{ر} - ١)\,\triangle^{٢}\text{ص}_{-١}$$

$$+ \tfrac{١}{٦}\,(\text{ر}+١)\,\text{ر}\,(\text{ر}-١)\cdot\triangle^{٣}\text{ص}_{-١}$$

$$+ \tfrac{١}{٢٤}\,(\text{ر}+١)\,\text{ر}\,(\text{ر}-١)(\text{ر}-٢)\cdot\triangle^{٤}\text{ص}_{-٢} + \cdots$$

نعوض عن $\tfrac{١}{٢}\triangle \text{ص}_{٠}$ ، $\tfrac{١}{٢}\triangle^{٢}\text{ص}_{-١}$ ، $\tfrac{١}{٢}\triangle^{٤}\text{ص}_{-٢}$ ، .. بقيمها الآتية :

$$\text{ص}_{٠} = \text{ص}_{١} - \triangle \text{ص}_{٠}$$

$$\triangle^{٢}\text{ص}_{-١} = \triangle^{٢}\text{ص}_{٠} - \triangle^{٣}\text{ص}_{-١}$$

$$\triangle^{٤}\text{ص}_{-٢} = \triangle^{٤}\text{ص}_{-١} - \triangle^{٥}\text{ص}_{-٢}$$

. . . .

فتصبح المعادلة على الصورة المذكورة فى بند ٤٨٥

تم الجزء الأول

جـــدول قيم كا٢

مبوبة تبعاً لاحتمالات (ح) ودرجات حرية (ن)

ح / ن	,٩٩	,٩٨	,٩٥	,٩٠	,٥	,١	,٠٥	,٠٢	,٠١	ح / ن
١	,٠٠٠١٥٧	,٠٠٠٦٢٨	,٠٠٣٩٣	,٠١٥٨	,٤٥٥	٢,٧٠٦	٣,٨٤١	٥,٤١٢	٦,٦٣٥	١
٢	,٠٢٠١	,٠٤٠٤	٠,١٠٣	٠,٢١١	١,٣٨٦	٤,٦٠٥	٥,٩٩١	٧,٨٢٤	٩,٢١٠	٢
٣	٠,١١٥	,١٨٥	,٣٥٢	,٥٨٤	٢,٣٦٦	٦,٢٥١	٧,٨١٥	٩,٨٣٧	١١,٣٤٥	٣
٤	٠,٢٩٧	,٤٢٩	,٧١١	١,٠٦٤	٣,٣٥٧	٧,٧٧٩	٩,٤٨٨	١١,٦٦٨	١٣,٢٧٧	٤
٥	,٥٥٤	,٧٥٢	١,١٤٥	١,٦١٠	٤,٣٥١	٩,٢٣٦	١١,٠٧٠	١٣,٣٨٨	١٥,٠٨٦	٥
٦	,٨٧٢	١,١٣٤	١,٦٣٥	٢,٢٠٤	٥,٣٤٨	١٠,٦٤٥	١٢,٥٩٢	١٥,٠٣٣	١٦,٨١٢	٦
٧	١,٢٣٩	١,٥٦٤	٢,١٦٧	٢,٨٣٣	٦,٣٤٦	١٢,٠١٧	١٤,٠٦٧	١٦,٦٢٢	١٨,٤٧٥	٧
٨	١,٦٤٦	٢,٠٣٢	٢,٧٣٣	٣,٤٩٠	٧,٣٤٤	١٣,٣٦٢	١٥,٥٠٧	١٨,١٦٨	٢٠,٠٩٠	٨
٩	٢,٠٨٨	٢,٥٣٢	٣,٣٢٥	٤,١٦٨	٨,٣٤٣	١٤,٦٨٤	١٦,٩١٩	١٩,٦٧٩	٢١,٦٦٦	٩
١٠	٢,٥٥٨	٣,٠٥٩	٣,٩٤٠	٤,٨٦٥	٩,٣٤٢	١٥,٩٨٧	١٨,٣٠٧	٢١,١٦١	٢٣,٢٠٩	١٠
١١	٣,٠٥٣	٣,٦٠٩	٤,٥٧٥	٥,٥٧٨	١٠,٣٤١	١٧,٢٧٥	١٩,٦٧٥	٢٢,٦١٨	٢٤,٧٢٥	١١
١٢	٣,٥٧١	٤,١٧٨	٥,٢٢٦	٦,٣٠٤	١١,٣٤٠	١٨,٥٤٩	٢١,٠٢٦	٢٤,٠٥٤	٢٦,٢١٧	١٢
١٣	٤,١٠٧	٤,٧٦٥	٥,٨٩٢	٧,٠٤٢	١٢,٣٤٠	١٩,٨١٢	٢٢,٣٦٢	٢٥,٤٧٢	٢٧,٦٨٨	١٣
١٤	٤,٦٦٠	٥,٣٦٨	٦,٥٧١	٧,٧٩٠	١٣,٣٣٩	٢١,٠٦٤	٢٣,٦٨٥	٢٦,٨٧٣	٢٩,١٤١	١٤
١٥	٥,٢٢٩	٥,٩٨٥	٧,٢٦١	٨,٥٤٧	١٤,٣٣٩	٢٢,٣٠٧	٢٤,٩٩٦	٢٨,٢٥٩	٣٠,٥٧٨	١٥
١٦	٥,٨١٢	٦,٦١٤	٧,٩٦٢	٩,٣١٢	١٥,٣٣٨	٢٣,٥٤٢	٢٦,٢٩٦	٢٩,٦٣٣	٣٢,٠٠٠	١٦
١٧	٦,٤٠٨	٧,٢٥٥	٨,٦٧٢	١٠,٠٨٥	١٦,٣٣٨	٢٤,٧٦٩	٢٧,٥٨٧	٣٠,٩٩٥	٣٣,٤٠٩	١٧
١٨	٧,٠١٥	٧,٩٠٦	٩,٣٩٠	١٠,٨٦٥	١٧,٣٣٨	٢٥,٩٨٩	٢٨,٨٦٩	٣٢,٣٤٦	٣٤,٨٠٥	١٨
١٩	٧,٦٣٣	٨,٥٦٧	١٠,١١٧	١١,٦٥١	١٨,٣٣٨	٢٧,٢٠٤	٣٠,١٤٤	٣٣,٦٨٧	٣٦,١٩١	١٩
٢٠	٨,٢٦٠	٩,٢٣٧	١٠,٨٥١	١٢,٤٤٣	١٩,٣٣٧	٢٨,٤١٢	٣١,٤١٠	٣٥,٠٢٠	٣٧,٥٦٦	٢٠
٢١	٨,٨٩٧	٩,٩١٥	١١,٥٩١	١٣,٢٤٠	٢٠,٣٣٧	٢٩,٦١٥	٣٢,٦٧١	٣٦,٣٤٣	٣٨,٩٣٢	٢١
٢٢	٩,٥٤٢	١٠,٦٠٠	١٢,٣٣٨	١٤,٠٤١	٢١,٣٣٧	٣٠,٨١٣	٣٣,٩٢٤	٣٧,٦٥٩	٤٠,٢٨٩	٢٢
٢٣	١٠,١٩٦	١١,٢٩٣	١٣,٠٩١	١٤,٨٤٨	٢٢,٣٣٧	٣٢,٠٠٧	٣٥,١٧٢	٣٨,٩٦٨	٤١,٦٣٨	٢٣
٢٤	١٠,٨٦٥	١١,٩٩٢	١٣,٨٤٨	١٥,٦٥٩	٢٣,٣٣٧	٣٣,١٩٦	٣٦,٤١٥	٤٠,٢٧٠	٤٢,٩٨٠	٢٤
٢٥	١١,٥٢٤	١٢,٦٩٧	١٤,٦١١	١٦,٤٧٣	٢٤,٣٣٧	٣٤,٣٨٢	٣٧,٦٥٢	٤١,٥٦٦	٤٤,٣١٤	٢٥
٢٦	١٢,١٩٨	١٣,٤٠٩	١٥,٣٧٩	١٧,٢٩٢	٢٥,٣٣٦	٣٥,٥٦٣	٣٨,٨٨٥	٤٢,٨٥٦	٤٥,٦٤٢	٢٦
٢٧	١٢,٨٧٩	١٤,١٢٥	١٦,١٥١	١٨,١١٤	٢٦,٣٣٦	٣٦,٧٤١	٤٠,١١٣	٤٤,١٤٠	٤٦,٩٦٣	٢٧
٢٨	١٣,٥٦٥	١٤,٨٤٧	١٦,٩٢٨	١٨,٩٣٩	٢٧,٣٣٦	٣٧,٩١٦	٤١,٣٣٧	٤٥,٤١٩	٤٨,٢٧٨	٢٨
٢٩	١٤,٢٥٦	١٥,٥٧٤	١٧,٧٠٨	١٩,٧٦٨	٢٨,٣٣٦	٣٩,٠٨٧	٤٢,٥٥٧	٤٦,٦٩٣	٤٩,٥٨٨	٢٩
٣٠	١٤,٩٥٣	١٦,٣٠٦	١٨,٤٩٣	٢٠,٥٩٩	٢٩,٣٣٦	٤٠,٢٥٦	٤٣,٧٧٣	٤٧,٩٦٢	٥٠,٨٩٢	٣٠

جدول فيشر لقيم ت

مبوَّبة تبعاً لاحتمالات « ح » ودرجات حرية « ن » معلومة[1]

ح / ن	٫١	٫٠٥	٫٠٢	٫٠١	ح / ن	٫١	٫٠٥	٫٠٢	٫٠١
١	٦٫٣١٤	١٢٫٧٠٦	٣١٫٨٢١	٦٣٫٦٥٧	٢١	١٫٧٢١	٢٫٠٨٠	٢٫٥١٨	٢٫٨٣١
٢	٢٫٩٢٠	٤٫٣٠٣	٦٫٩٦٥	٩٫٩٢٥	٢٢	١٫٧١٧	٢٫٠٧٤	٢٫٥٠٨	٢٫٨١٩
٣	٢٫٣٥٣	٣٫١٨٢	٤٫٥٤١	٥٫٨٤١	٢٣	١٫٧١٤	٢٫٠٦٩	٢٫٥٠٠	٢٫٨٠٧
٤	٢٫١٣٢	٢٫٧٧٦	٣٫٧٤٧	٤٫٦٠٤	٢٤	١٫٧١١	٢٫٠٦٤	٢٫٤٩٢	٢٫٧٩٧
٥	٢٫٠١٥	٢٫٥٧١	٣٫٣٦٥	٤٫٠٣٢	٢٥	١٫٧٠٨	٢٫٠٦٠	٢٫٤٨٥	٢٫٧٨٧
٦	١٫٩٤٣	٢٫٤٤٧	٣٫١٤٣	٣٫٧٠٧	٢٦	١٫٧٠٦	٢٫٠٥٦	٢٫٤٧٩	٢٫٧٧٩
٧	١٫٨٩٥	٢٫٣٦٥	٢٫٩٩٨	٣٫٤٩٩	٢٧	١٫٧٠٣	٢٫٠٥٢	٢٫٤٧٣	٢٫٧٧١
٨	١٫٨٦٠	٢٫٣٠٦	٢٫٨٩٦	٣٫٣٥٥	٢٨	١٫٧٠١	٢٫٠٤٨	٢٫٤٦٧	٣٫٧٦٣
٩	١٫٨٣٣	٢٫٢٦٢	٢٫٨٢١	٣٫٢٥٠	٢٩	١٫٦٩٩	٢٫٠٤٥	٢٫٤٦٢	٢٫٧٥٦
١٠	١٫٨١٢	٢٫٢٢٨	٢٫٧٦٤	٣٫١٦٩	٣٠	١٫٦٩٧	٢٫٠٤٢	٢٫٤٥٧	٢٫٧٥٠
١١	١٫٧٩٦	٢٫٢٠١	٢٫٧١٨	٣٫١٠٦	٣٥		٢٫٠٣٠		٢٫٧٢٤
١٢	١٫٧٨٢	٢٫١٧٩	٢٫٦٨١	٣٫٠٥٥	٤٠		٢٫٠٢١		٢٫٧٠٤
١٣	١٫٧٧١	٢٫١٦٠	٢٫٦٥٠	٣٫٠١٢	٤٥		٢٫٠١٤		٢٫٦٩٠
١٤	١٫٧٦١	٢٫١٤٥	٢٫٦٢٤	٢٫٩٧٧	٥٠		٢٫٠٠٨		٢٫٦٧٨
١٥	١٫٧٥٣	٢٫١٣١	٢٫٦٠٢	٢٫٩٤٧	٦٠		٢٫٠٠٠		٢٫٦٦٠
١٦	١٫٧٤٦	٢٫١٢٠	٢٫٥٨٣	٢٫٩٢١	٧٠		١٫٩٩٤		٢٫٦٤٨
١٧	١٫٧٤٠	٢٫١١٠	٢٫٥٦٧	٢٫٨٩٨	٨٠		١٫٩٩٠		٢٫٦٣٨
١٨	١٫٧٣٤	٢٫١٠١	٢٫٥٥٢	٢٫٨٧٨	٩٠		١٫٩٨٧		٢٫٦٣٢
١٩	١٫٧٢٩	٢٫٠٩٣	٢٫٥٣٩	٢٫٨٦١	١٠٠		١٫٩٨٤		٢٫٦٢٦
٢٠	١٫٧٢٥	٢٫٠٨٦	٢٫٥٢٨	٢٫٨٤٥	∞	١٫٦٤٥	١٫٩٦٠	٢٫٣٢٦	٢٫٥٧٦

(١) مختصر عن جدوله الموجود فى كتابه « الطرق الإحصائية » وجداول أخرى أيضاً .

جدول المنحنى التكرارى المعتدل

تبويب قيم الاحداثى ص $= \frac{١}{\sqrt{٢ط}} . هـ^{-\frac{س^٢}{٢}}$ والاحتمالين

م $= \int_{\cdot}^{س}$ ص . ء س ، ح = ١ — ٢م تبعاً لقيم س ، وهى انحراف أى مفردة عن الوسط الحسابى بوحدات تساوى الانحراف المعيارى . وتكون م تساوى احتمال وجود أى مفردة بين صفر ، + س ؛ و ح تساوى احتمال وجودها خارج الحدود ± س ، أى أكبر من س عددياً .

س	ص	م	ح	س	ص	م	ح
٠ر٠	٠ر٣٩٨٩٤	٠ر٠٠٠٠٠	١ر٠٠٠٠٠	٢ر٥	ر٠١٧٥٣	٠ر٤٩٣٧٩	ر٠١٢٤٢
٠ر١	ر٣٩٦٩٥	ر٠٣٩٨٣	ر٩٢٠٣٤	٢ر٦	ر٠١٣٥٨	ر٤٩٥٣٤	ر٠٠٩٣٢
ر٢	ر٣٩١٠٤	ر٠٧٩٣٦	ر٨٤١٤٨	٢ر٧	ر٠١٠٤٢	ر٤٩٦٥٣	ر٠٠٦٩٣
ر٣	ر٣٨١٣٩	ر١١٧٩١	ر٧٦٤١٨	٢ر٨	ر٠٠٧٩٢	ر٤٩٧٤٤	ر٠٠٥١١
ر٤	ر٣٦٨٢٧	ر١٥٥٤٢	ر٦٨٩١٦	٢ر٩	ر٠٠٥٩٥	ر٤٩٨١٣	ر٠٠٣٧٣
ر٥	ر٣٥٢٠٧	ر١٩١٤٦	ر٦١٧٠٨	٣ر٠	ر٠٠٤٤٣	ر٤٩٨٦٥	ر٠٠٢٧٠
ر٦	ر٣٣٣٢٢	ر٢٢٥٧٥	ر٥٤٨٥١	٣ر١	ر٠٠٣٢٧	ر٤٩٩٠٣	ر٠٠١٩٤
ر٧	ر٣١٢٢٥	ر٢٥٨٠٤	ر٤٨٣٩٣	٣ر٢	ر٠٠٢٣٨	ر٤٩٩٣١	ر٠٠١٣٧
ر٨	ر٢٨٩٦٩	ر٢٨٨١٤	ر٤٢٣٧١	٣ر٣	ر٠٠١٧٢	ر٤٩٩٥٢	ر٠٠٠٩٧
ر٩	ر٢٦٦٠٩	ر٣١٥٩٤	ر٣٦٨١٢	٣ر٤	ر٠٠١٢٣	ر٤٩٩٦٦	ر٠٠٠٦٧
١ر٠	ر٢٤١٩٧	ر٣٤١٣٤	ر٣١٧٣١	٣ر٥	ر٠٠٠٨٧	ر٤٩٩٧٧	ر٠٠٠٤٧
١ر١	ر٢١٧٨٥	ر٣٦٤٣٣	ر٢٧١٣٣	٣ر٦	ر٠٠٠٦١	ر٤٩٩٨٤	ر٠٠٠٣٢
١ر٢	ر١٩٤١٩	ر٣٨٤٩٣	ر٢٣٠١٤	٣ر٧	ر٠٠٠٤٢	ر٤٩٩٨٩	ر٠٠٠٢٢
١ر٣	ر١٧١٣٧	ر٤٠٣٢٠	ر١٩٣٦٠	٣ر٨	ر٠٠٠٢٩	ر٤٩٩٩٣	ر٠٠٠١٤
١ر٤	ر١٤٩٧٣	ر٤١٩٢٤	ر١٦١٥١	٣ر٩	ر٠٠٠٢٠	ر٤٩٩٩٥	ر٠٠٠١٠
١ر٥	ر١٢٩٥٢	ر٤٣٣١٩	ر١٣٣٦١	٤ر٠	ر٠٠٠١٣	ر٤٩٩٩٧	ر٠٠٠٠٦
١ر٦	ر١١٠٩٢	ر٤٤٥٢٠	ر١٠٩٦٠	٤ر١	ر٠٠٠٠٩	ر٤٩٩٩٨	ر٠٠٠٠٤
١ر٧	ر٠٩٤٠٥	ر٤٥٥٤٣	ر٠٨٩١٣	٤ر٢	ر٠٠٠٠٦	ر٤٩٩٩٩	ر٠٠٠٠٣
١ر٨	ر٠٧٨٩٥	ر٤٦٤٠٧	ر٠٧١٨٦	٤ر٣	ر٠٠٠٠٤	ر٤٩٩٩٩	ر٠٠٠٠٢
١ر٩	ر٠٦٥٦٢	ر٤٧١٢٨	ر٠٥٧٤٣	٤ر٤	ر٠٠٠٠٢	ر٤٩٩٩٩	ر٠٠٠٠١
٢ر٠	ر٠٥٣٩٩	ر٤٧٧٢٥	ر٠٤٥٥٠	٤ر٥	ر٠٠٠٠٢		ر٠٠٠٠١
٢ر١	ر٠٤٣٩٨	ر٤٨٢١٤	ر٠٣٥٧٣	٤ر٦	ر٠٠٠٠١		ر٠٠٠٠٠
٢ر٢	ر٠٣٥٤٧	ر٤٨٦١٠	ر٠٢٧٨١	٤ر٧	ر٠٠٠٠١		
٢ر٣	ر٠٢٨٣٣	ر٤٨٩٢٨	ر٠٢١٤٥	٤ر٨	ر٠٠٠٠٠		
٢ر٤	ر٠٢٢٣٩	ر٤٩١٨٠	ر٠١٦٤٠	٤ر٩	ر٠٠٠٠٠		

دليل

بأسماء الأعلام والاصطلاحات الافرنجية

A

B

C

D

E