٢٥٦
٧٦٤

تجليد صالح الدقر
تلفون ٢٢٩٧٧


311.2
H21tA
5 Jun 68
21 Nov 68
18 DEC 1968
JAFET
- 9 DEC 992

311.2
H21t
C.1

طرق الإحصاء

تأليف

الدكتور محمد مظلوم حمدى

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

بكالوريوس الشرف فى التجارة من جامعة فؤاد الأول
بكالوريوس الشرف فى الاقتصاد من جامعة لندن
دكتوراه فى الإحصاء من جامعة لندن
أستاذ الاقتصاد السياسى بكلية التجارة فى جامعة فاروق الأول

الطبعة الثانية

١٣٧١هـ — ١٩٥١م

حقوق الطبع محفوظة للمؤلف

ملتزم الطبع والنشر
دار المعارف بمصر

مقدمة الطبعة الأولى

بسم الله الرحمن الرحيم الذى بفضله نتحرر من أصفاد الجهل ونسلك سبل الهدى والرشاد وبقوته نستعين على تذليل ما يواجهنا من شدائد وصعاب .

أقدم إلى طلبتى الأعزاء مؤلفى الأول فى علم الإحصاء راجياً أن يكون الله قد سدد خطاى وأفلح مسعاى ووفقنى إلى خدمتهم وتهيئة أسباب البحث والاطلاع لهم .

وبعد فلم أقصد أن يكون هذا الكتاب بحثاً متعمقاً فى نظرية الإحصاء . بل إنى تعمدت أن أجعله عرضاً مبسطاً لأهم طرق الإحصاء بحيث يشمل المناهج التى وضعتها لطلبة كليتى التجارة والزراعة بجامعة فاروق الأول فى هذه المادة وبحيث يفيد منه كل من تعنيه دراسة الإحصاء . وعلى الرغم من أنى قد توسعت فى بعض موضوعات الكتاب إلى الحد الذى يجاوز العرض المبسط إلا أنى وضعت نصب عينى حاجة الطالب المبتدئ فى غالبية الموضوعات التى عالجتها .

وليس هذا مجال ترديد أهمية مادة الإحصاء وحاجتنا إلى المزيد من المعلومات عنها . ولكن إذا « أحصينا » عدد الكتب الأجنبية عن طرق ونظرية الإحصاء لفزعنا حقاً من افتقار مكتبتنا العربية إلى مؤلفات فى هذه المادة . ولولا المجهود العظيم الذى بذله أستاذى الفاضل الدكتور عبد المنعم ناصر الشافعى فى إخراج الكتاب الأوحد فى علم الإحصاء لظل هذا الباب غير مطروق إلى يومنا هذا . وسيجد القارئ أنى قد استعنت بكثير من أفكار هذا الأستاذ الكبير الذى كان أول من لقننى هذا المادة خصوصاً فيما يتعلق بصياغة

المصطلحات الفنية باللغة العربية كما استعنت بأفكار عدد كبير من المؤلفين الأجانب خصوصاً أولئك الذين تتلمذت عليهم فى انجلترا ، وجنيت من علمهم الغزير محصولا وفيرا كان له الفضل فى إخراج هذا الكتاب .

وأنتهز هذه الفرصة لأتوجه بشكرى الخالص إلى زميلى وتلميذى من قبل حضرة سعد إبراهيم برغوث أفندى المعيد بكلية التجارة وعضو بعثتها لدراسة الإحصاء بأمريكا وتلميذى البار على يوسف أفندى للمساعدة القيمة التى قدماها لى فى عمل بعض الرسوم البيانية الواردة فى هذا الكتاب . كما أشكر تلميذى الفاضل سعد أفندى زغلول الذى قام بعمل الشكل الخاص بالعرض التصويرى وحضرة محمد على حسن أفندى الذى توج الكتاب بخطه الجميل .

والله أسأل أن يوفقنا جميعاً إلى خدمة وطننا المحبوب .

يناير سنة ١٩٤٨ محمد مظلوم حمدى

مقدمة الطبعة الثانية

منذ أن نفدت الطبعة الأولى من كتاب طرق الإحصاء ، أتيحت لى فرصة مراجعته مراجعة شاملة للوقوف على أوجه النقص فيه . وقد وجدت أنه من الضرورى ، إتماماً للفائدة ، إضافة بعض الموضوعات والتوسع فى البعض الآخر . فأعدت كتابة الباب الثانى الخاص بجمع البيانات الإحصائية وأضفت إليه ملحقاً عن أهم مصادر الإحصاءات المصرية . وأضفت إلى الباب الرابع جزءاً جديداً عن منحنى لورنز نظراً لأهميته فى دراسة إحصاءات الدخل والملكية . كذلك ألحقت بالباب الثامن بعض الإضافات الضرورية . وفى الباب الثالث عشر أضفت جزءاً جديداً عن الارتباط بين السلسلات الزمنية . وقد رأيت تحقيقاً لرغبة أبداها كثير من الطلبة أن أضيف إلى الكتاب نماذجاً من الأسئلة التى وضعتها للامتحانات المختلفة . وقد حذفت الجزء الخاص بإيجاد عزوم التوزيعات التكرارية بالتكامل من الباب السابع نظراً لأن أهميته لا تكاد تتعدى الناحية النظرية البحت .

وقد انتهزت فرصة مراجعة الكتاب لتصحيح بعض الأخطاء التى وردت فى الطبعة الأولى . وإنى لأتقدم بشكرى الخالص لتلميذى الوفى وزميلى الآن حضرة رمزى زكى سيد أحمد أفندى المعيد بقسم الاقتصاد بكلية التجارة الذى عاونى معاونة صادقة فى مراجعة الكتاب وأبدى ملاحظات هامة .

أسأل الله أن يهيئ لنا السبيل السوى .

نوفمبر سنة ١٩٥١ محمد مظلوم حمدى

الباب الأول

الطريقة الإحصائية

الطريقة الإحصائية هى إحدى طرق وصف ومقارنة المجموعات والظواهر المتغيرة وإثبات الحقائق العلمية المتصلة بها شأنها فى ذلك شأن أية طريقة من طرق الاستنتاج المنطقى كالطريقة التجريبية مثلا . غير أنها تختلف عن طرق الإثبات الأخرى فى أنها تعتمد اعتماداً كلياً على التعبير الرقمى عن الظواهر . فالظواهر الغير القابلة للقياس العددى أو التى لا يمكن التعبير عنها بالأرقام بأى صورة كانت لا تكون موضوعاً للدراسات الإحصائية . على أنه يجب ألا يفهم من قصر استعمال الطريقة الإحصائية على الناحية العددية للظواهر أن تطبيقها محدود المجال ، إذ الواقع أن هذه الطريقة أصبحت الآن فى مقدمة طرق البحث العلمى فى أى فرع من فروع الدراسة . وقليل من التأمل كفيل بأن يقنعنا بأن معظم الظواهر التى نشاهدها فى حياتنا العملية أو التى نقابلها فى دراساتنا المختلفة يمكن التعبير عنها بالأرقام إما عن طريق القياس المباشر كالطول والسن والوزن والضغط ودرجة الحرارة والثمن والطلب والدخل والتكاليف والربح أو بدلالة أى وحدات أو رتب أو مستويات نحددها كالذكاء والجدارة والقوة الشرائية للنقود .

وفكرة الإحصاء بالمعنى العادى المألوف عبارة عن جمع بيانات عن المجموعات والظواهر المختلفة والتعبير عنها فى صورة رقمية . وهى بهذا المعنى ليست بالفكرة الحديثة بل إنها كانت معروفة منذ أقدم عصور التاريخ . غير أن جمع هذه البيانات قلما يكون فى حد ذاته غاية نسعى إليها بل فى إنه

فى معظم الأحيان وسيلة أو خطوة أولية نرمى بها إلى تحقيق هدف معين . وكثيراً ما يكون هذا الهدف هو وصف المجموعات والظواهر المتغيرة إما لمجرد التعرف عليها والوقوف على طريقة تغيرها أو لغرض مقارنتها بظواهر أخرى ومحاولة استنباط علاقات عامة تربط بعضها بالبعض الآخر . من الواضح إذن أن عمل الإحصائى أو أى شخص يستخدم الطريقة الإحصائية كإحدى طرق الإثبات العلمى لا ينتهى بجمع البيانات الإحصائية عن الظواهر التى يبحثها بل الأصح أن نقول أن عمله يبتدىء بجمع هذه البيانات ولا ينتهى إلا بعد أن تتم تنقيتها مما قد يعلق بها من شوائب ثم ترتيبها وتحليلها وتفسيرها . فالبيانات الإحصائية للباحث العلمى هى بمثابة المواد الأولية للصانع والطرق الإحصائية له بمثابة الأدوات والآلات التى يستخدمها لتحويل المواد الأولية إلى سلع صالحة لإشباع الحاجات . وكما أن المواد الأولية تختلف فى جودتها فتختلف قيمة السلع المنتجة منها تبعاً لذلك فإن البيانات الإحصائية تختلف أيضاً فى دقة تمثيلها للظواهر الموضوعة تحت البحث وتختلف تبعاً لذلك قيمة النتائج التى نحصل عليها منها . والمواد الأولية مهما بلغت جودتها فلا تصلح فى شكلها الأصلى لإشباع الحاجات التى نشعر بها بل لا بد من خلق منفعتها أو زيادتها بتحويلها إلى الشكل المناسب الذى يكسبها صفة الصلاحية لإشباع الحاجات وبالمثل فإن البيانات الإحصائية أو الأرقام الأولية لا تصلح لإبراز أية حقيقة علمية عن الظواهر إلا بعد أن يتم تحليلها بالطرق الإحصائية المختلفة لتعيين الحدود المنطقية التى تجيز لنا استنتاج أى علاقات أو قوانين تخضع لها هذه الظواهر . فمن الخطأ أن نكتفى بتسجيل بعض البيانات عن ظاهرة معينة ثم نندفع إلى عمل استنتاجات خاصة بها متخطين بذلك مرحلة تحليل البيانات ووضعها فى الصورة التى تجيز عمل هذه الاستنتاجات .

وقد يكون التعبير الرقمى عن الظواهر من أقوى وسائل الإقناع والإثبات فإن منطق الأرقام منطق موضوعى مستقل عن الاعتقاد أو الميل الشخصى

للباحث . غير أنه قد يكون أيضاً من أشد وسائل التضليل إذا أسىء فهمه أو أسىء استعمالة .

ومع أنه قلما نجد الآن أى بحث علمى خال من البيانات الإحصائية سواء أكان هذا البحث اجتماعيا أو تاريخياً أو جغرافياً أو اقتصادياً أو زراعياً أو طبياً فإن سوء استعمال هذه البيانات شائع بدرجة تدعو إلى الأسف حقاً . ويرجع ذلك إما إلى عدم فهم طبيعة الطرق الإحصائية أو إلى معرفة القليل المضلل عنها .

فقد ذكرنا أن طرق الإحصاء هى للباحث العلمى بمثابة الأدوات والآلات للصانع . وهل ننتظر من الصانع أن يؤدى عملا منتجاً إلا إذا فهم طبيعة الأدوات والآلات التى يستخدمها ؟ ومن أخطر وسائل التضليل هو محاولة التوصل إلى نتائج علمية من واقع البيانات الأولية مباشرة . وقد بينا أن ذلك لا يختلف كثيراً عن محاولة تقديم المواد الأولية للمستهلك فى شكلها الأصلى مباشرة . ومع ذلك فكثيراً ما نجد مثل الإحصاءات الآتية فى الأبحاث العلميـــة : —

جدول رقم ١

إيرادات السكك الحديدية وعدد السيارات فى القطر المصرى

السنة	إيرادات السكك الحديدية بالجنيهات	عدد السيارات
١٩٢٨	٣,٠٤١,١٢١	٢٤ ٥٠٥
١٩٢٩	٢,٩٦٢,٥٦٩	٢٧ ٧١٥
١٩٣٠	١,٩٦٩,٥٥٠	٣٠ ٠٧٩
١٩٣١	١,٣١٥,٠٣١	٢٩ ٩٩٦
١٩٣٢	١,٤٦٩,٣٤٣	٣٧ ٦٦٥

المصدر : الإحصاء السنوى العام للقطر المصرى .

يقدم الباحث هذه البيانات أو هذه الأرقام الأولية على علاتها دون تحليلها بأية طريقة إحصائية ويستنتج منها مباشرة أن أزمة السكك الحديدية التى تتمثل فى هبوط إيراداتها ترجع إلى اشتداد منافسة وسائل النقل بالطرق التى تتمثل فى زيادة عدد السيارات فى نفس الوقت الذى يحدث فيه الهبوط . ربما نعتبر مسرفين فى النقد إذا نعتنا هذه الطريقة بأنها طريقة خاطئة فى الاستنتاج . ولكن يمكننا على الأقل أن نصفها بعدم الكمال للأسباب الآتية :-

أولا - أن عدد السنين التى سجلت أرقامها لا يكفى لتبرير الوصول إلى مثل تلك النتيجة العامة . فإذا اختبرنا الأرقام الخاصة بعدد كبير من السنين فقد لا نجد مثل هذا الارتباط[1] بين الظاهرتين .

ثانياً - أنه حتى على فرض وجود ارتباط بين الظاهرتين فقد يكون هذا الارتباط راجعاً إلى عوامل المصادفة وهو على أى حال لا يكفى لإثبات علاقة السببية بين الظاهرتين .

ثالثاً - قد تكون هناك أسباب أخرى خلاف ازدياد عدد السيارات تؤثر فى إيرادات السكك الحديدية كالكساد الشديد الذى أصاب كافة المرافق الاقتصادية من سنة ١٩٢٩ إلى سنة ١٩٣٢ .

رابعاً - إذا جاز أن نعزو هبوط إيرادات السكك الحديدية إلى زيادة عدد السيارات فلا بد أن تكون المقارنة على أساس عدد السيارات التى يحتمل أن تحول جزءاً من حركة البضائع والركاب عن السكك الحديدية وليس العدد الإجمالى للسيارات بما فى ذلك السيارات الخصوصية التى نستبعد أن يكون عددها مسئولا عن نقص حركة نقل البضائع والركاب بالسكك الحديدية على هذا النحو . والواقع أن الزيادة فى عدد السيارات الموضحة بالجدول السابق تكاد

---

(١) نقصد مبدئياً بالارتباط مصاحبة القيم الكبيرة للظاهرة الأولى للقيم الكبيرة للظاهرة الأخرى أو مصاحبة القيم الكبيرة للظاهرة الأولى للقيم الصغيرة للظاهرة الأخرى . لشرح المعنى التفصيلى للارتباط راجع الباب الخاص بهذا الموضوع .

ترجع كلية إلى زيادة عدد السيارات الخصوصية الأمر الذى يهدم حجة الباحث فى استنتاج علاقة السببية بين الظاهرتين من أساسها .(١)

هناك مئات أخرى من الأمثلة التى تدل على سوء استعمال البيانات الإحصائية يستطيع الإنسان أن يستخرجها من الأبحاث العلمية المنشورة فعلا . غير أنه يكفى أن نشير هنا فى إيجاز إلى أهم الأخطاء الشائعة فى هذا المجال :ـ

أولا ـ التحيز

من المبادئ المعروفة فى علم النفس أن الإنسان يميل دون قصد منه إلى الانتهاء إلى النتائج التى تقبلها نفسه . فكأنه يصل إلى النتيجة أولا وبعد ذلك يحاول إيجاد الأرقام التى تثبت صحة هذه النتيجة . ولا غبار على هذا المسلك إذا كان لدى الباحث أدلة أخرى قوية مستقلة عن الطريقة الإحصائية تؤيد نتيجته . أى إذا كانت هذه الطريقة الأخيرة مستخدمة لمجرد تأييد ما يكاد يكون ثابتاً فعلا بطرق أخرى .

- أما إذا كانت هذه هى الطريقة الأساسية التى يعتمد عليها فى الإثبات فلا شك أنه من الخطأ أن يصل إلى النتيجة قبل أن توصله الأرقام إليها . ويقول الاستاذ يول (Yule) أنه قلما يخلو إنسان من الميل الطبيعى نحو التحيز فى صوره المختلفة فى اختيار البيانات الإحصائية وأن هذه الصفة الإنسانية متأصلة إلى درجة تجعل من المستحيل التغلب عليها حتى بالمران أو الحذر الشديد(٢) . وتتضح أهمية التحيز عندما نحاول وصف مجموع معين (كالقمح المصرى مثلا) عن طريق اختيار عينة منه تحتوى على عدد محدود من

---

(١) هذه الأرقام والاستنتاجات التى بنيت عليها كانت النقطة الأساسية فى بحث عن أزمة السكك الحديدية نشرته مجلة L'Egypte contemporaine عدد ١٣٩ يناير سنة ١٩٣٣

(٢) راجع كتاب « مقدمة لنظرية الإحصاء » لمؤلفيه يول وكندال ص ٣٣٧ .

المفردات كما هو الحال فى كل الأبحاث الإحصائية تقريباً . فعلى الرغم من أنه يبدو من السهل اختيار عينة اعتباطية خالية من التحيز إلا أن التجارب دلت على أن مثل هذا الاختيار المحايد غير ممكن عمليا بدون الالتجاء إلى بعض الطرق التى سنبحثها فى موضوع العينات .

ثانياً ـ استعمال بيانات غير كافية

طرق الإحصاء تبحث فى الظواهر المتغيرة فقط . والظواهر المتغيرة مفروض نظرياً أنها تتخذ عدداً غير محدود من القيم ، وعلى ذلك فمحاولة وصف الظاهرة عن طريق تسجيل عدد صغير من هذه القيم ربما يؤدى إلى نتائج خاطئة . فلو حاولنا تقدير طول نبات القمح فمن الخطأ أن نكتفى باختيار خمسة أو عشرة نباتات اعتباطاً ونسجل أطوالها ، إذ من المحتمل أن تكون الخمسة أو العشرة نباتات جميعها شاذة بعيدة عن الطول النموذجى أو الطول المتوسط . أما لو اخترنا عينة عدد مفرداتها خمسمائة أو ألف فإن احتمال الخطأ يقل كثيراً . وسنرى فيما بعد كيف أن دقة المقاييس الإحصائية تزداد بزيادة عدد مفردات العينة .

ثالثاً ـ عمل مقارنة بين بيانات لا يصح المقارنة بينها .

من أوضح الأمثلة على ذلك مقارنة الأجور النقدية فى دولتين مختلفتين ، أو فى الدولة الواحدة فى أوقات مختلفة ، واعتبار الأجر النقدى المرتفع دليلا على ارتفاع مستوى المعيشة دون أن نأخذ فى الحسبان اختلاف مستوى الأسعار فى الحالتين . أو مقارنة قطعتين من الأراضى الزراعية واعتبار المحصول الأعلى دليلا على خصوبة أعظم دون مراعاة نفقات الزراعة التى بذلت فى الحالتين . ويمكن بوجه عام أن نقول أن مقارنة التغير المطلق فى قيم الظواهر تكون أدعى

إلى الخطأ والتضليل من مقارنة التغير النسبى . فإذا وجدنا مثلا أن عدد سكان مدينة القاهرة زاد بمقدار ١٠٠,٠٠٠ نسمة فى حين أن عدد سكان مدينة بنها زاد بمقدار ٢٠,٠٠٠ فلا يجوز أن نستنتج من ذلك أن سكان المدينة الأولى يتزايدون بسرعة أكبر من الثانية فالعبرة فى مثل هذه الحالات بالزيادة النسبية وليس بالزيادة المطلقة .

ومن العوامل المهمة الواجب مراعاتها عند عمل المقارنة بين الظواهر هو تحديد وحدات المقارنة تحديداً دقيقاً لا يترك مجالا للشك أو الإبهام . فإذا كنا نقيس الإنتاج اليومى لمصنعين تمهيداً لمقارنة الكفاية الانتاجية لكل منهما فيجب أن نتنبه إلى احتمال اختلاف عدد ساعات العمل اليومى فى كل منهما الأمر الذى يجعل أساس المقارنة فى حاجة إلى التعديل . كما أنه لو كنا نقارن أرقام الإنتاج فى مصنع معين من شهر لآخر فقد تختلف عدد أيام العمل فى كل شهر ليس فقط لاختلاف عدد أيام الشهر من ٢٨ إلى ٣١ يوماً بل لوقوع أيام العطلات العامة فى أشهر دون الأخرى . كذلك إذا كان موضوع البحث يتعلق بمقارنة نوع معين من المنتجات الزراعية أو الصناعية بنوع آخر يجب أن نتأكد من أن وحدات المقارنة تصف نفس الشىء فى الحالتين . فلا يكفى أن نقول إن مصر استوردت فى سنة ١٩٢٩ ما قيمته ١ مليون جنيه من الجلود فى حين أنها استوردت فى سنة ١٩٣٩ ما قيمته نصف مليون جنيه فقط دون أن نحدد بدقة المقصود من كلمة « الجلود » — هل هى الجلود الخام أو هل تشمل الجلود المدبوغة أو البضائع المصنوعة من الجلد ، وهل الرقم الخاص بسنة ١٩٣٩ يصف نفس أنواع السلع المشار إليها فى إحصاءات سنة ١٩٢٩ ؟ .

تطور الطريقة الإحصائية وظهور الإحصاء كعلم مستقل

ذكرنا أن التعبير الرقمى عن الظواهر كان معروفاً منذ أقدم العصور وكذلك

فإن الطريقة الإحصائية فى الاثبات لا تعتبر فناً حديثاً . فقد كانت مستعملة منذ القرن السابع عشر ولكن فى شكلها البدائى الذى لا يتعدى جمع البيانات الأولية والاعتماد عليها بحالتها الأصلية أو بعد تعديل طفيف فى الإثبات . وقد بينا الخطأ الذى قد يقع فيه الباحث إذا اكتفى بهذه البيانات دون تحليلها . ولتجنب سوء استعمال الطريقة الاحصائية وضعت عدة قواعد وقوانين إحصائية لإرشاد الباحث العلمى إلى كيفية استخدام الارقام التى يسجلها عن الظواهر فى إثبات الحقائق وتقدير قيمة النتائج التى يتوصل إليها بالطريقة الإحصائية . أو بعبارة أخرى لتعيين درجة الثقة التى يمكن أن يضعها فى هذه النتائج . وقد كثرت هذه القواعد والقوانين وأصبحت فى حد ذاتها حقائق ثابتة رياضياً مما جعل الإحصاء علماً قائماً بذاته وليس مجرد طريقة من طرق الإثبات . على أن الإحصاء ليس له مجال مستقل للتطبيق بل إن مواده الأولية — الأرقام — لا بد وأن تستمد من الظواهر التى تبحث فيها العلوم الأخرى وبنوع خاص مجموعة العلوم البيولوجية والاجتماعية والاقتصادية وأى علم آخر يبحث فى الظواهر المتغيرة التى يمكن قياسها والتعبير عنها تعبيراً رقمياً .

على أنه مما لا شك فيه أن مجال تطبيق قواعد ونظريات علم الإحصاء فى الأبحاث البيولوجية بوجه عام والتجارب الزراعية بنوع خاص أعظم من مجال تطبيقها فى أى فرع من فروع الدراسة . وقد تمت القواعد والقوانين الإحصائية التى تجد مجال تطبيقها الرئيسى فى العلوم البيولوجية حتى لقد أصبحت بدورها تكون علماً مستقلا هو علم الـ Biometry أو الإحصاء البيولوجى . وطبيعة علم الإحصاء تشابه إلى حد كبير جداً علم الرياضة البحتة لا سيما من جهة علاقته بالعلوم الأخرى كما أنها تعتمد عليه اعتماداً يكاد يكون كليا . فالرياضة علم مستقل قائم بذاته ومع ذلك فإن مجال تطبيقه هو فى العلوم الأخرى كالطبيعة والميكانيكا والاقتصاد وعلى وجه العموم جميع الدراسات التى تبحث فى الظواهر التى يمكن قياسها والتعبير عنها تعبير رمزياً . وسواء

اعتبرنا كلا من الرياضة أو الإحصاء علماً مستقلا أم طريقة علمية فالفرق شكلى بحت لا يؤثر فى أهمية هذه الدراسات التى أصبحت الأساس الذى تستند إليه معظم الأبحاث العلميـــة الراقية فى شتى فروع النشاط الفكرى .

ولا بد أن نشيد هنا بفضل علماء الرياضة الذين وضعوا الأسس والقوانين التى تكون فى مجموعها نظرية الإحصاء ونخص بالذكر باسكال وأسرة برنوللى وفردريك جاوس . إذ يرجع لباسكال وبعض أعضاء أسرة برنوللى الفذة الفضل فى وضع عدة قوانين عظيمة الفائدة فى دراسة الاحتمالات التى تقوم عليها معظم الطرق الإحصائية الراقية والتى لولا تطبيقها فيها لما اكتسبت الإحصاء هذه الصبغة العلمية ولظلت النتائج التى نحصل عليها باستخدام هذه الطرق تقريبية لا يصح الاعتماد عليها . فمعظم الأبحاث الإحصائية ترمى كما ذكرنا إلى تحديد صفات المجموعات الكبيرة التى تحتوى على عدد لا نهاية له من المفردات ، عن طريق دراسة عينات محدودة مختارة من بين هذه المفردات . وما دمنا نجرى البحث على العينة وليس على المجموع الأصلى فلا شك أننا نتخبط فى الظلام وأن نتائجنا تكون طبعاً مثاراً للشك لأنها لا تخلو من عنصر « التخمين » . غير أن تطبيق نظرية الاحتمالات فى دراسة العينات أضاء لنا السبيل فأصبحنا نعرف الحدود التى لا يمكن أن يتجاوزها الخطأ المحتمل تسربه إلى نتائجنا . وبذلك زادت ثقتنا فى هذه النتائج زيادة عظيمة . ولا نظن أننا نغالى إذا قلنا أن نظرية الاحتمالات هى جوهر علم الإحصاء . أما العالم الفذ فردريك جاوس فيكفى أن نشير هنا إلى أنه مكتشف قانون المنحنى الطبيعى أو منحنى الخطأ المتماثل الذى لا تكاد تخلو أى طريقة إحصائية راقية من الاعتماد عليه . وإذا كان فضل هؤلاء العلماء فى بناء دعائم علم الإحصاء لا ينكر فإن الشخص الذى شيد البناء الإحصائى الشامخ على تلك الدعائم الرياضية المتينة هو كارل بيرسون فهو بحق الإحصائى الأول كما أنه أول من وضع أسس ذلك الفرع من الدراسات الإحصائية البيولوجية

وهو الـ Biometrics الذى أصبحت معظم بحوث الوراثة والمحاصيل والتطور ترتكز عليه . ومن العلماء المعاصرين الذين ساهموا فى تعلية البناء الذى شيده بيرسون الأساتذة أدنى يول وموريس كندال وأرثر بولى وفيشر وغيرهم ممن لا يتسع المجال للتنويه بفضلهم على هذا العلم .

الباب الثانى

جمع وتبويب البيانات الإحصائية

أشرنا فى الباب السابق إلى أن عمل الإحصائى يبتدىء بجمع البيانات الإحصائية عن الظواهر وأنه يتناول بعد ذلك عرض هذه البيانات ثم تحليلها وتفسيرها . وسنفرد هذا الباب لطرق جمع وتبويب البيانات الإحصائية ثم نختتمه بعرض موجز لأهم مصادر الإحصاءات المصرية .

لا يمكن وضع قواعد عامة يتعين اتباعها عند عمل أى بحث إحصائى . فلكل بحث ظروفه الخاصة التى تحدد مقدار الجهد والمال الممكن بذلهما فى سبيل التوصل إلى النتائج المطلوبة . ومن أهم المسائل التى تعرض للباحث عند تحديد نطاق العمل هو الاختيار بين جمع البيانات الإحصائية بنفسه أو الاعتماد على البيانات المنشورة فعلا والتى جمعت فى الماضى لأغراض أخرى . ولا شك أن الطريقة الأولى أفضل بكثير ، بل إن الثانية لا تعدو كونها بديلا يلجأ إليه الباحث لعجزه عن الحصول على الموارد التى تسمح له باتباع الطريقة الأصلية . فالبحث الإحصائى يفقد كثيراً من قيمته إذا لم يكن الباحث ملماً بكافة الظروف التى جمعت فيها البيانات التى يتخذها أساساً لبحثه حتى يستطيع تقدير درجة دقتها والوقوف على أوجه النقص التى قلما تخلو منها البيانات الإحصائية الأولية . وتتوقف طريقة جمع البيانات إلى حد كبير على نوع البحث الذى نتابعه . ففى بعض الابحاث الإحصائية كتلك المتصلة بالتجارب الزراعية والبيولوجية يجب أن يشرف الباحث بنفسه على جمع البيانات التى تستمد من مشاهدة الظواهر وقياسها وتسجيل

الأرقام التى تعبر عن بعض الحقائق المتصلة بها . وفى بعض الأبحاث الأخرى كتلك المتعلقة بالإحصاءات الحيوية كإحصاءات المواليد والوفيات يجد الباحث نفسه مضطرا إلى الاعتماد على البيانات التى تجمعها السلطات الحكومية المختصة والتى يتطلب جمعها وقتاً ومالا يتعذر على غير الحكومات انفاقهما .

وثمة مشكلة أخرى تعرض للباحث عندما يضع الخطة العامة للبحث وهى الاختيار بين بحث « المجموع » كله بحثا شاملا وبين الاكتفاء بانتخاب عينة تمثل هذا المجموع بقدر الإمكان . وقد كانت طريقة بحث المجموع كله هى السائدة حتى عهد قريب خصوصاً فى الأبحاث الإحصائية الرسمية . غير أن تقدم الدراسات الإحصائية الخاصة بالعينات وما ترتب عليه من ازدياد درجة الثقة فى النتائج التى يمكن الحصول عليها من دراسة العينات الممثلة للمجموع جعل هذه الطريقة تزداد فى الأهمية إلى حد أنها كادت تطغى على الطريقة الأصلية . ومن مزايا طريقة البحث بالعينة أنها توفر فى الوقت والمال اللذين يتعين بذلهما فى البحث . كما أنها تجعل من السهل استخدام نفر قليل من الأخصائيين المدربين على جمع البيانات الاحصائية . أما طريقة بحث المجموع كله ، فعلاوة على ما تتطلبه من وقت ومال فإنها تحتم استخدام عدد ضخم من جامعى البيانات الذين قلما تتوفر فيهم الصفات اللازمة والخبرة الكافية والذين يتعذر تدريبهم وإعدادهم لأداء مهمتهم على الوجه الأكمل . وبالإضافة إلى ذلك فإن طريقة البحث بالعينة تفوق طريقة بحث المجموع فى أنها تهيىء الفرصة لاختيار عدد محدود من الأشخاص الذين تستمد منهم البيانات فتأتى إجاباتهم على الأسئلة الموجهة إليهم دقيقة واضحة خالية من التحريف والتضليل فى حين أن الطريقة الأخرى تلجىء الباحث إلى توجيه الأسئلة إلى كل من يشملهم المجموع على اختلاف درجة ثقافتهم ودرجة إيمانهم بأهمية البحث ودرجة ثقتهم فى الغرض الذى يهدف إليه . وعلى الرغم مما لطريقة البحث بالعينة من مزايا عديدة فإن طريقة بحث المجموع

كله ما زالت تستخدم فى حالات قليلة يكون ميدان البحث فيها ضيقاً بطبيعته والمجموع ذاته محدوداً بحيث لا يكون ثمة مبرر للالتجاء إلى طريقة العينة . على أن أهم حالة تستخدم فيها طريقة بحث المجموع كله هى التعدادات الرسمية مثل تعداد السكان وإحصاء الإنتاج الصناعى والزراعى ، إذ فى مثل هذه الحالات تحتاج الحكومة إلى بيانات تفصيلية وافية تشمل جميع السكان والموارد الإنتاجية ، ومن الواضح أن مثل هذه البيانات التفصيلية لا يمكن الحصول عليها بطريقة العينة . غير أنه نظراً لما يبذل فى هذه التعدادات من وقت وجهد ومال فإن الحكومة تكتفى بإجرائها مرة واحدة كل عشرة سنوات . وقد استخدمت الحكومة المصرية طريقة العينات فى البحث الذى أجرته لتقدير عدد مستحقى المعاشات والإعانات طبقاً لقانون الضمان الاجتماعى . وعلى الرغم من أن عدد مفردات العينات المختارة لم يكن يتجاوز نصف فى المائة من عدد مفردات المجموع ورغم أن الطريقة التى اتبعت فى اختيار العينات لم تكن الطريقة المثلى فإن البحث قد أتى بالنتائج المرجوة . وتلجأ الحكومات أحياناً إلى طريقة مختلطة تجمع بين بحث المجموع وبحث العينة . ففى تعداد سكان الولايات المتحدة توجه الأسئلة السهلة الواضحة إلى جميع السكان بينما يختار منهم نسبة صغيرة (١٪ مثلا) يطلب منها الإجابة على الأسئلة التى تحتاج إلى عناية وتفكير .

ويمكن بوجه عام أن نوجز طرق جمع البيانات الإحصائية فيما يلى :ـ

أولا ـ طريقة العد

تتلخص هذه الطريقة فى استخدام عدد من الموظفين (العدادين) الذين يمرون بالمنازل منزلا منزلا أو بالمصانع أو بالمحال التجارية للحصول على بعض المعلومات من السكان أو أصحاب المصانع أو المحال . وأحياناً يطلب هؤلاء العدادون من المحال التى يزورونها الإجابة على الأسئلة التى يوجهونها لهم شفهياً

ثم يسجلون هم البيانات بأنفسهم وأحيانا يوزعون قوائم تشتمل على عدة أسئلة (Questionnaires) ويطلبون من السكان أو أصحاب المحال الإجابة كتابة على الأسئلة الموجهة إليهم . وهذه هى الطريقة المتبعة فى تعداد سكان القطر المصرى .

ثانياً — طريقة المراسلة .

هذه الطريقة لا تختلف عن السابقة إلا فى أن قوائم الأسئلة التى أشرنا إليها ترسل بطريق البريد . وتستعمل هذه الطريقة عادة فى إحصاءات الرأى العام أو الاستفتاءات العامة (Public Opinion Surveys) التى تقوم بها بعض المؤسسات الخاصة بذلك أو الجرائد الكبرى .

وأهم ما يجب مراعاته عند استخدام إحدى الطريقتين السابقتين هو أن تكون الأسئلة الموجهة إلى الأفراد سهلة وواضحة ومختصرة بحيث يتسنى للشخص العادى الإجابة عليها دون عناء . فلا بد أن نتذكر أنه ليس هناك حافز قوى يدفع الشخص إلى تكبد مشقة كبيرة فى الإجابة على الأسئلة التى توجه إليه . ولضمان الحصول على إجابات صحيحة يجب تفسير الأسئلة الغامضة أو تلك التى يحتمل أن تكون الإجابة عليها غير مستوفاة . فمثلا إذا أردنا سؤال شخص عن حالته المدنية فالأفضل أن نضع السؤال فى الصيغة الآتية :—

هل أنت : متزوج
أو أعزب
أو أرمل
أو مطلق

ونطلب منه التأشير أمام الحالة التى تنطبق عليه بوضع علامة (×) أمامها. وذلك لأننا إن لم نضع السؤال بهذا التفصيل فمن المحتمل جدا أن يكتفى

الشخص الموجه إليه السؤال بأن يذكر ما إذا كان متزوجاً أو غير متزوج تاركا الإشارة إلى الحالتين الأخيرتين مع أهميتهما فى الوقوف على الحالة الاجتماعية للسكان . ويحسن ، كلما أمكن ذلك ، أن توضع الأسئلة فى صيغة تجعل من الممكن الإجابة عليها بنعم أو لا .

ثالثاً — طريقة التسجيل :

هذه هى الطريقة التى تتبعها الحكومات عادة فى الحصول على البيانات التى تعتبرها ضرورية للوقوف على سير الحالة الصحية أو الاجتماعية أو الاقتصادية للبلاد . وتتلخص فى إلزام جميع أفراد الشعب بتبليغ السلطات المسئولة عن وقوع بعض الأحداث الاجتماعية الهامة كالوفاة والميلاد والزواج والطلاق . وأحياناً تكلف الحكومات فئة خاصة فقط من الأفراد بمدها ببعض البيانات . ونذكر على سبيل التمثيل تكليف أصحاب المصانع أو المحال التجارية بمد الجهات المختصة ببيانات عن طبيعة صناعاتهم أو تجارتهم وعدد العمال المشتغلين بها وما إلى ذلك من البيانات التى تحتاج إليها الحكومة . وقد يكون التسجيل عاماً ومستمراً كما فى حالة تسجيل المواليد والوفيات . وقد ينشأ فى ظروف خاصة وينتهى بزوالها مثل تكليف الممولين بمد الحكومة ببيانات عما فى حيازتهم من الأرصدة الأجنبية أو الأصول المقومة بالعملة الأجنبية كما حدث فى مصر عندما أرادت الحكومة حصر الأرصدة الاسترلينية المستحقة لأفراد الشعب المصرى ، أو كما حدث إبان الحرب الماضية فى بريطانيا عندما أرادت الحكومة البريطانية حصر ممتلكات الشعب البريطانى فى أمريكا للاستيلاء عليها .

وعلى الرغم من أن لطريقة التسجيل عيب واضح ينشأ من تهاون بعض الأفراد فى أداء واجبهم فى التبليغ عن الأحداث الاجتماعية حال وقوعها ، أو تهرب بعضهم من مد السلطات المختصة ببيانات عن أنفسهم وعن أموالهم

عندما يطلب منهم ذلك ، إلا أن هذه الطريقة تكون مصحوبة عادة بفرض عقوبات رادعة تحد من ميل البعض إلى التهرب من أداء الواجب المطلوب .

مراجعة وتبويب البيانات الإحصائية

مراجعة البيانات .

بعد أن ننتهى من جمع البيانات الأولية لا بد من مراجعة القوائم التى ترد إلينا قبل أن نجرى عليها أى عملية أخرى . فعملية المراجعة قد تكشف عن أخطاء أو نقص فى الإجابة كما أنها قد تبين فى الحال أن كثيراً من البيانات التى لدينا لا يمكن الاعتماد على دقتها أو قبولها على علاتها . ومن المعتاد أن أن تكون البيانات التى قام بجمعها عدادون متمرنون أكثر مطابقة للواقع من البيانات التى ترد بطريقة المراسلة .

فالعداد اللبق يمكنه أن يستخلص من الشخص الذى يوجه إليه الأسئلة معلومات أكثر وأدق من تلك التى يرسلها الشخص بمحض إرادته بطريق البريد . ومع ذلك فمهما بلغت درجة الدقة المتبعة فى جمع البيانات فإنها لا تغنى عن مراجعة القوائم الواردة بدقة وعناية . وتختلف طريقة المراجعة باختلاف نوع البيانات الإحصائية والاهداف التى جمعت من أجلها . على أن هناك بعض القواعد ذات التطبيق العام وسنشير إليها فيما يلى بإيجاز .

لا بد أن يثق القائم بالعمل من أن القوائم الواردة كاملة ، ويجب أن تعزل تلك التى لاينطبق عليها هذا الوصف . كذلك يجب أن تفحص الإجابات لمعرفة ما إذا كانت معقولة ومتمشية مع بعضها ، أو بعبارة أخرى للتأكد من خلوها من التناقض . ولا بد أن تعزل جميع الإجابات المتناقضة أو تلك التى يعتبرها القائم بعملية المراجعة بعيدة الاحتمال جداً توطئة لمحاولة تصحيحها بقدر الإمكان . فإذا لم يكن من الممكن تصحيحها فيجب أن ترد إلى مصدرها

الأصلى أو تهمل كلية ويستعاض عنها بغيرها .

ولعملية المراجعة فائدة كبرى فى الحالات التى تحتاج فيها البيانات الإحصائية إلى بعض التعديل الذى يحقق الانسجام بين الإجابات وتجانس الوحدات المستعملة . فمثلا إذا كانت القوائم تشتمل على سؤال خاص بالسن ، فقد يعطى بعض الأشخاص سنهم إلى أقرب سنة والبعض الآخر إلى أقرب شهر وهكذا . وفى مثل هذه الحالة يجب تحقيق التجانس بين الوحدات المستعملة فى جميع الإجابات قبل استخدام هذه البيانات . وقد نسأل بعض المصانع عن الأجور التى تدفع لعمالها فيعطينا البعض أجراً يومياً والبعض الآخر أجراً أسبوعياً أو شهرياً . ومن الواضح أنه لا بد فى مثل هذه الحالة أن ترد جميع الأجور إلى أساس واحد ، يومياً كان أو أسبوعياً أو شهرياً .

تبويب البيانات

بعد أن تتم مراجعة القوائم للتأكد من صلاحيتها للغرض المطلوب تفرز هذه القوائم كى تقسم إلى مجموعات متجانسة . وبديهى أن الغرض الذى نسعى إليه من عمل البحث الإحصائى — أو النتيجة التى نرغب فى استخلاصها من البيانات المجموعة — هو الذى يحدد لنا صفة التجانس هذه . فإذا كان هدفنا من عمل البحث الإحصائى هو الوقوف على كيفية توزيع العمال فى بلد ما على فروع النشاط الإنتاجى المختلفة ، فإن صفة التجانس هنا هى الاشتراك فى مهنة أو حرفة أو صناعة معينة . وعليه نقسم البيانات الموجودة لدينا إلى مجموعات تشتمل الواحدة منها على جميع العمال المشتركين فى حرفة أو صناعة معينة . فنضع العمال الزراعيين فى مجموعة ، والصناعيين فى مجموعة أخرى وهكذا . ويمكن أن نعيد التقسيم بعد ذلك بطريقة أكثر تفصيلا حسب الغرض المطلوب .

وقبل أن نتعرض لشرح الطرق العملية للتبويب نذكر بعض الأسس الشائعة التى تقوم عليها هذه العملية :—

(١) التبويب الزمنى

واستخدام هذا الأساس فى التبويب يعنى فصل جميع البيانات الخاصة بوحدة زمنية معينة — ولتكن شهر أو سنة مثلا — على حدة .

(٢) التبويب الجغرافى .

واستخدام هذا الأساس يتطلب تقسيم البيانات إلى مجموعات كل منها خاص بوحدة جغرافية معينة — أى بمكان معين .

(٣) التبويب الكمى .

ويعنى تقسيم البيانات الإحصائية إلى فئات تضم كل فئة منها قيمة معينة أو مدى محدوداً من قيم الظاهرة الموضوعة تحت البحث . ولنفرض أن لدينا بيانات عن أعمار السكان . يمكن تبويب هذه البيانات على أساس وضع جميع السكان الذين فى سن العشرين فى فئة معينة والذين فى سن الحادية والعشرين فى فئة أخرى وهكذا . ويمكن طبعاً أن يتسع مدى الفئات فتضم الفئة الأولى جميع السكان الأقل من سن الخامسة وتضم الثانية جميع السكان الواقعة أعمارهم بين الخامسة والعاشرة وهكذا . ويتوقف مدى الفئة أو اتساعها على الغرض الذى نسعى إليه من البحث ودرجة الدقة المطلوبة . وسنعود إلى موضوع تقسيم البيانات الأولية إلى فئات بتفصيل أكثر فى الباب الخاص بالتوزيعات التكرارية .

(٤) التبويب الوصفى .

ويقصد به تقسيم البيانات إلى مجموعات تشترك مفردات كل منها فى صفة خاصة تعتبر ذات أهمية فى موضوع البحث . فمثلا فى بحث عن البطالة يصح تقسيم العمال إلى عمال نظاميين يعملون كل الوقت ، وعمال يعملون بعض الوقت فقط ، وعمال عاطلين .

وإذا ما انتهينا من تقسيم البيانات على النحو السابق نعمل جدولا يحتوى على عدد من الخانات يماثل التقسيمات التى اتفقنا عليها ثم تفرغ فيه البيانات .

ولتوضيح طريقة تفريغ البيانات فى الجدول ، نفرض أن لدينا قوائم خاصة بعدد من العمال الذين ينتمون إلى صناعات مختلفة ونريد أن نقسمهم حسب حالة العمل (أى طبقاً لما إذا كانوا يعملون كل الوقت أو بعض الوقت أو لا يعملون إطلاقاً) .

نبدأ بفرز القوائم حسب نوع الصناعة ، ويحسن أن يرمز لكل صناعة برقم معين لتسهيل العمل . بعد ذلك نتناول القوائم واحدة تلو الأخرى ونرصد البيانات الموجودة بها فى الخانة الخاصة بذلك فى الجدول ، وذلك بوضع إشارة يتفق عليها فى الخانة المذكورة ولتكن / . ونكرر هذه العملية حتى ننتهى من تفريغ جميع القوائم فى الجدول . ولإرشاد العين فى عملية العد نضع كل أربع إشارات بجوار بعضها هكذا //// . أما الإشارة الخامسة فتشطب الأربعة السابقة لها على هذا النحو 𝍸 . وهكذا يكون لدينا مجموعات من الإشارات كل منها يحتوى على خمسة (𝍸 𝍸 𝍸 𝍸) متباعدة عن بعضها نوعاً ويمكن عدها بسهولة تامة . وفى النهاية يتخذ الجدول الذى أفرغت فيه البيانات الصورة الآتيـــة : —

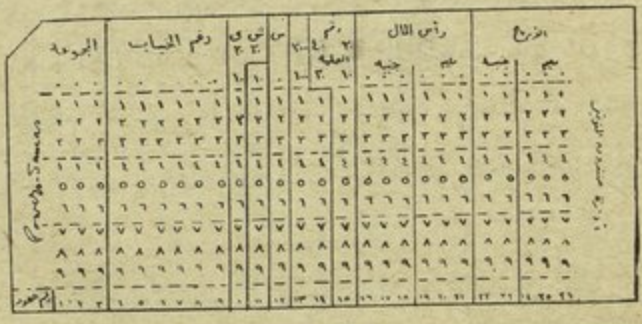
نوع الصناعة	العمال النظاميون	العمال الذين يعملون بعض الوقت	العمال العاطلون
(١)	١٧	١٢	٣
(٢)	٢٦	٧	٦
(٣)	١٥	١	٢
(٤)	٢٥	١٠	٤

واضح من الجدول السابق أن عدد العمال النظاميين فى الصناعة رقم (١) هو ١٧ ، وعدد العمال الذين يعملون بعض الوقت فى الصناعة ذاتها هو ١٢ ، وعدد العاطلين ٣ ، أما عدد العمال فى الصناعة رقم ٢ حسب التقسيم السابق فهو ٢٦ ، ٧ ، ٦ على التوالى . أى أنه يمكن إعادة تكوين الجدول بالصورة الآتيـــة :—

التبويب الآلى :

من الواضح أن طريقة التبويب السابقة تكون مرهقة جدا فى العمل لو أن عدد القوائم المطلوب تبويب البيانات المسجلة عليها كان كبيراً . هذا فضلا عن أن احتمال الخطأ فى عملية التبويب يكون كبيراً جداً . فليس من المعقول أن يقوم شخص واحد (أو عدد من الأشخاص) بفرز ٥٠,٠٠٠ أو ١٠٠,٠٠٠ قائمة وتفريغ محتوياتها فى الجدول بالطريقة السابقة دون أن يتسرب الخطأ إلى عمله . فإذا تذكرنا أن عدد القوائم الخاصة بالتعداد العام للسكان تعد بالملايين ،

تبين لنا بكل وضوح أن الطريقة السابقة لا تصلح مطلقاً فى مثل هذه الإحصاءات الكبيرة . ولذلك لا بد فى مثل تلك الحالات من الالتجاء إلى آلات التبويب التى بلغت الآن درجة من الدقة والإتقان تدعو إلى الإعجاب حقاً . وأشهر طرق التبويب الآلى طريقة البطاقات المثقوبة (Punch card System) والفكرة الأساسية فى طريقة التبويب الآلى هى رصد البيانات المطلوب تبويبها على بطاقات خاصة مقسمة إلى عدد من الخانات ، يختلف اتساع كل منها حسب الحاجة . وتحتوى كل خانة على عدد من الأعمدة المكونة من الأرقام المسلسة من ٠ إلى ٩ ، ويختلف عدد الأعمدة حسب اتساع الخانة . فبعض الخانات تحتوى على عمود واحد فقط من الأرقام ، وبعضها على اثنين أو أكثر . وسنرى سبب اختلاف اتساع الخانات وبالتالى عدد الأعمدة الموجودة داخل كل منها عند شرح طريقة تفريغ البيانات على البطاقة . ولكى نعطى القارىء فكرة عن هذه البطاقات نورد فيما يلى شكل إحداها قبل أن تسجل عليها البيانات ثم ننتقل إلى شرح طريقة تسجيل البيانات عليها .



(شكل رقم ١)

إحدى البطاقات التى تستعملها إدارة صندوق التوفير بمصلحة البريد فى تبويب البيانات

طريقة استعمال البطاقات .

المبدأ الأساسى الذى تنبنى عليه طريقة البطاقات المثقوبة هو الاستعاضة

عن البيانات اللفظية التى تصف أى ظاهرة برموز رقمية . وهذه الطريقة تشابه إلى حد ما طريقة الشفرة المستعملة فى المراسلات البرقية . غير أنها تختلف عنها فى أنه ليس هناك طريقة موحدة متفق عليها للاستعاضة عن الكلمات بالرموز كما هو الحال فى طريقة الشفرة ، بل إن هذه الرموز يختارها الباحث بنفسه فى كل مرة لتناسب موضوع البحث الذى يجريه . فمثلا إذا كنا نعمل بحثاً إحصائياً يتطلب تقسيم السكان من حيث الجنس ، يصح أن نرمز للذكر بالرقم ١ وللأنثى بالرقم ٢ . كذلك إذا كنا نعمل بحثاً يتطلب بيان صلة كل شخص برب الأسرة التى يعيش فيها ، يصح أن نرمز إلى رب الأسرة نفسه بالرقم ١ والإبن بالرقم ٢ والإبنة بالرقم ٣ والأخ بالرقم ٤ والأخت بالرقم ٥ وهكذا . وعلى ذلك إذا فحصنا بطاقة خاصة بهذا البحث ووجدنا مسجلا عليها فى الخانة الخاصة بالصلة برب الأسرة الرقم ٢ ، نعلم أن الشخص الذى تخصه هذه البطاقة هو ابن رب الأسرة التى يعيش فيها دون حاجة إلى تسجيل أى بيانات لفظية . ومن الواضح أن هناك بعض البيانات التى تعطى فى الأصل فى صورة رقمية مثل السن أو الأجر أو عدد حجرات المسكن ، وفى مثل هذه الحالات لا نحتاج طبعاً إلى الاستعاضة عنها برموز أخرى بل تسجل الأرقام كما هى على البطاقة . ويظهر مما تقدم أن الباحث مطلق الحرية فى اختيار الرموز التى يراها مناسبة فى حدود الحيز الموجود على البطاقـــة .

بعد أن نختار الرموز المناسبة يحسن أن نسجلها فى دليل خاص كى يسهل علينا عكس العملية السابقة أى رد الرموز إلى البيانات الأصلية التى تنوب عنها كما يحدث فى حالة « فك الشفرة » .

تبدأ بعد ذلك عملية تسجيل البيانات (التى أصبحت الآن رموزاً رقمية) على البطاقات . ويتم هذا التسجيل لا بكتابة الرقم الخاص على البطاقة (فالبطاقة كما ذكرنا مطبوع عليها فعلا عدة أعمدة من الأرقام المسلسلة من ٠ إلى ٩)

وإنما بثقب البطاقة فى موضع الرقم الذى نرغب فى تسجيله وذلك باستعمال آلة خاصة بالثقب (Key Punch) موضحة فى الشكل رقم ٢ . ويكون هذا الثقب بمثابة تسجيل البيان على البطاقة .

ولكى نوضح طريقة تفريغ البيانات من القوائم عملياً نفرض أن البيانات الموضحة فيما يلى كانت مسجلة فى إحدى القوائم التى تستعملها إدارة صندوق التوفير بمصلحة البريد .

حساب رقم : ٦٥ ٤٣٥ مكتب : شبرا مصر

مجموعة : ٨٠٣

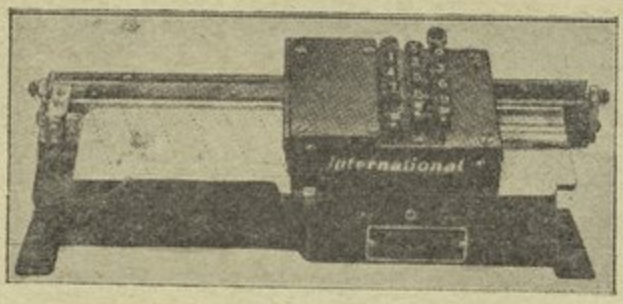
حساب خاص بحضرة : محمد بدوى نيازى أفندى

المهنة والعنوان : خبير ضرائب ٩٧ شارع شبرا

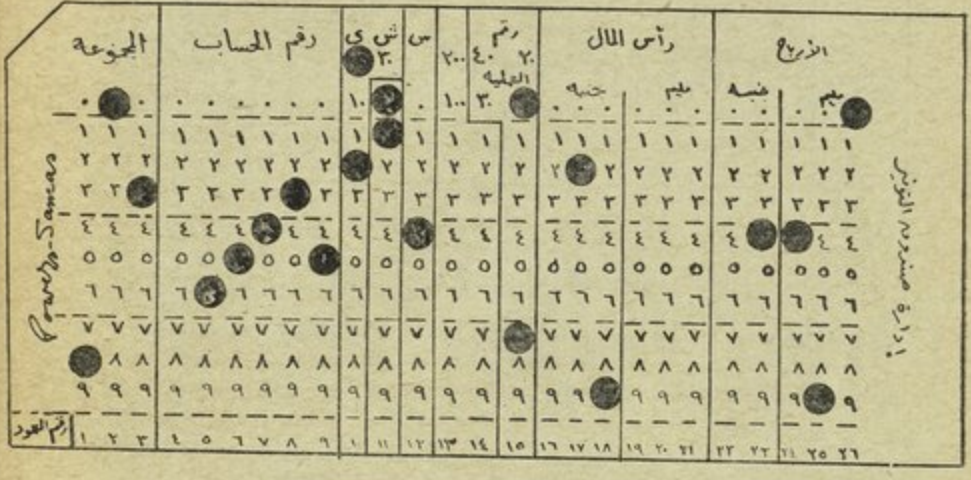
الرصيد				العملية				التاريخ	رقم العملية
الأرباح		رأس المال		الأرباح		رأس المال			
مليم	جنيه	مليم	جنيه	مليم	جنيه	مليم	جنيه		
٤٩٠	٤	٠٠٠	٣٢٩					١٩٤٤/١١/٢١	
				٤٩٠	٤	٠٠٠	٢٩	١٩٤٤/١١/٢٢	١٧
٠٠	٠٠		٣٠٠						

يظهر من البيانات السابقة أن رصيد المودع كان فى ١٩٤٤/١١/٢١ ٣٢٩ جنيهاً . وأنه فى ١٩٤٤/١١/٢٢ سحب ٢٩ جنيهاً من رأس المال علاوة على جميع الأرباح المستحقة له فى هذا التاريخ وقدرها ٤ جنيه و ٤٩٠ مليم . ويمكن تسجيل سحب هذا المبلغ على إحدى البطاقات الخاصة بالسحب[١] كالآتى (انظر شكل ٣) .

(١) يلاحظ أن إدارة صندوق التوفير تستعمل بطاقات خاصة بالإيداع مماثلة لتلك الخاصة بالسحب وإنما من لون مخالف . وتسجيل هذه العمليات على البطاقات يساعدها على حصر جميع المبالغ التى أودعت أو سحبت فى أى يوم من أيام السنة بسهولة .

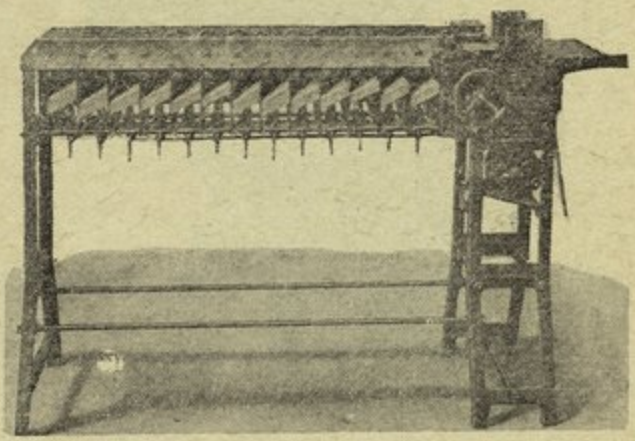


(شكل رقم ٢)
آلة ثقب البطاقات



(شكل رقم ٣)
البطاقة بعد تسجيل البيانات عليها باستعمال آلة الثقب

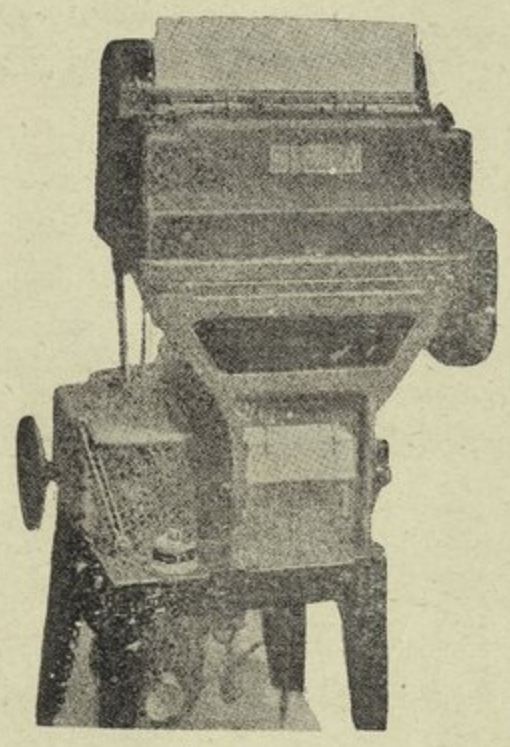
بعد أن ننتهى من تسجيل البيانات بالطريقة السابقة تفرز هذه البطاقات عن طريق تمريرها فى آلات كهربائية خاصة بالفرز (Mechanical Sorters) مثل تلك الموضحة فى الشكل رقم ٤ .



(شكل رقم ٤)
آلة الفرز الكهربائية

ويمكن لهذه الآلات فرز البطاقات عن طريق موضع الثقوب بمعدل ١٨٠٠٠ بطاقة فى الساعة . وكما ذكرنا فإن كل ثقب يدل على بيانات معينة .

تمر البطاقات بعد ذلك فى آلات التبويب (Tabulating Machines) مثل تلك الموضحة فى الشكل رقم ٥ .



(شكل رقم ٥)
آلة التبويب

ويضيق هذا المجال عن شرح تفاصيل عمل هذه الآلات ولكننا سنصف عملها وصفاً موجزاً يعطى القارىء فكرة عامة عنها .

لنفرض أن لدينا عدداً من البطاقات الخاصة ببعض الذكور وبعض الإناث وأردنا فرز كل مجموعة على حدة . نعرف أن بطاقات الذكور تكون مثقوبة فى العمود الخاص بالجنس عند رقم معين نختاره وليكن ١ مثلا . بينما تكون بطاقات الإناث مثقوبة عند رقم آخر وليكن ٢ مثلا . ينحصر عمل آلة الفرز فى هذه الحالة فى تقسيم البطاقات إلى مجموعتين إحداهما خاصة بالذكور والأخرى خاصة بالإناث ووضع كل مجموعة فى خانة خاصة من الخانات

الموجودة بها بطريقة آلية . وتتم هذه العملية عن طريق وضع البطاقات فى مستودع معد لها بآلة الفرز تمر منه الواحدة تلو الأخرى فوق سطح معدنى هو أحد نهايتى دائرة كهربائية . وبالآلة ذراع خاص له أسنان دقيقة مثل أسنان المشط هو عبارة عن النهاية الأخرى للدائرة الكهربائية . وفى أثناء تمرير البطاقة على السطح المعدنى تمسها أسنان هذا الذراع من أعلى ، وتظل الدائرة الكهربائية مفتوحة ما دام ورق البطاقة يعزل النهايتين عن بعضهما . فإذا ما مر الذراع بموضع الثقب الموجود بالبطاقة تقفل الدائرة الكهربائية ويحدث هذا تأثيراً خاصاً من شأنه أن تسقط البطاقة فى إحدى خانات آلة الفرز . أو على وجه التحديد تسقط البطاقة فى الخانة التى تحمل نفس الرقم المثقوب(١) وكذلك تسقط فى نفس هذه الخانة جميع البطاقات المثقوبة فى نفس الموضع ، أى جميع البطاقات الخاصة بالذكور مثلا . أما البطاقات المثقوبة فى موضع آخر — بطاقات الإناث — فتسقط فى خانة أخرى . وبذلك يتم فصل كل مجموعة من البطاقات على حدة . ومن الممكن طبعاً أن نأخذ نفس هذه البطاقات لنعيد فرزها على أساس آخر ، أساس السن مثلا .

فإذا كان البيان الذى سيعمل الفرز على أساسه يحتل أكثر من عمود واحد على البطاقة — كما لو أردنا فرز بطاقات صندوق التوفير الموضحة بشكل رقم (٣) على أساس رقم المجموعة — فلا بد من تكرار عمليه الفرز بقدر عدد الأعمدة التى يحتلها البيان وذلك لأن الآلة لا تستطيع الفرز إلا على أساس عمود واحد فقط فى كل مرة .

بالنظر إلى شكل (٣) نجد أن البيان الخاص برقم المجموعة يحتل ثلاثة أعمدة على البطاقة هى الأعمدة رقم ١ ، ٢ ، ٣ على التوالى . فإذا أردنا فصل

---

(١) يلاحظ أن آلة الفرز تحتوى على صف من الخانات المرقومة من ٠ إلى ٩ ولكل خانة غطاء يفتح من تلقاء نفسه عند إقفال الدائرة الكهربائية أى عند ما تمس أسنان الذراع السطح المعدنى من خلال الثقب الموجود على البطاقة .

جميع البطاقات التى تنتمى إلى المجموعة رقم ٨٠٣ على حدة ، نبدأ أولا بإعداد الآلة للفرز على أساس العمود رقم (١) وهو العمود الخاص برقم المئات فى البيان الخاص برقم المجموعة فتقسم لنا الآلة جميع البطاقات التى أدخلت فيها إلى عشرة مجموعات تعزل كل منها على حدة فى إحدى خانات آلة الفرز . نتناول بعد ذلك جميع البطاقات التى سقطت فى الخانة رقم ٨ (أى التى يكون رقم المجموعة فيها يبدأ بالرقم ٨) لنعيد فرزها على أساس العمود رقم (٢) وهو الخاص برقم العشرات فى الخانة الخاصة برقم المجموعة ، فتقسم لنا الآلة هذه البطاقات إلى عشرة مجموعات أخرى فنأخذ منها تلك التى سقطت فى الخانة رقم (٠) فتكون جميعها مسجل عليها الرقم ٨ فى خانات المئات والرقم (٠) فى خانة العشرات . ثم نتناول هذه البطاقات لفرزها مرة أخرى على أساس العمود رقم (٣) وهو الخاص برقم الآحاد فى البيان الخاص برقم المجموعة . وبعد أن تنتهى عملية الفرز نأخذ جميع البطاقات التى سقطت فى الخانه رقم (٣) فتكون جميعها تحمل الرقم ٨٠٣ فى الخانة الخاصة برقم المجموعة .

ويلاحظ أننا لا نحتاج إلى عد هذه البطاقات لمعرفة مجموعها . فإن آلة الفرز تحتوى على « عداد » يسجل عدد البطاقات التى تسقط فى كل خانة بطريقة أوتوماتيكية . كما أن هذه الآلة لا تحتاج إلى ملاحظة دقيقة أثناء تأدية عملها فهى تقف من تلقاء نفسها عند ما يتم فرز جميع البطاقات الموضوعة بها .

أما آلة التبويب فعملها الرئيسى هو إعادة كتابة الأرقام المسجلة على البطاقات بطريقة الثقب على ورق خاص معد بها . أى أنه إذا مررنا بها بطاقة مثقوبة فى موضع الرقم ٦ فى العمود رقم ١٤ فإنها تسجل هذا البيان كتابة على لفائف الورق المعدة بها . وبالإضافة إلى ذلك فهى تقوم بجمع الأرقام التى تسجلها بطريقة أوتوماتيكية . فمثـــلا لو مررنا بها البطاقات الخاصة بإدارة صندوق التوفير المسجل على كل منها بيان خاص بالمبلغ المسحوب

(أو المودع) فإنها تعطينا فى الحال مجموع المبالغ التى سحبت (أو أودعت) فى أى يوم . ويمكن لآلة التبويب معالجة ٣٥٠٠ بطاقة فى الساعة كما أنها تستطيع تسجيل سبعة مجاميع مختلفة فى نفس الوقت . وهذه الآلة مبنية على نفس فكرة الدائرة الكهربائية التى سبق أن أشرنا إليها وإن كانت أكثر تعقيداً من آلة الفرز .

ملحق للباب الثانى

أهم مصادر الإحصاءات المصرية

١ ــ التعدادات (كل عشر سنوات)

(ا) تعداد سكان المملكة المصرية :
الجزء الأول ــ كراسات تعداد كل من المحافظات والمديريات .
الجزء الثانى جداول عامة موضح فيها ــ على سبيل التمثيل وليس الحصر ــ توزيع السكان حسب فئات الأعمار وتوزيع السكان حسب الجنسية والمهنة وما إلى ذلك .

(ب) التعداد الصناعى والتجارى

(ح) التعداد الزراعى العام .

٢ ــ الإحصاءات الصحية والمدنية

(ا) نشرة عن المواليد والوفيات والأمراض المعدية والحالة الجوية باللغتين العربية والإنجليزية وهذه النشرة أسبوعية .

(ب) نشرة مواليد ووفيات المحافظات وعواصم المديريات باللغتين العربية والفرنيسية وهذه النشرة شهرية .

(ح) نشرة عن المواليد والوفيات والأمراض المعدية والزواج والطلاق باللغتين العربية والفرنسية وتصدر هذه النشرة مرة كل ثلاثة شهور .

(د) الإحصاءات الصحية باللغتين العربية والإنجليزية وهذه النشرة سنوية

(هـ) نشرة مواليد ووفيات نواحى المملكة المصرية وتصدر مرة كل ثلاث سنوات.

٣ ــ الزراعة والاقتصـــاد

(ا) نشرة يومية باللغتين العربية والإنجليزية عن أثمان وموجودات الغلال بأسواق روض الفرج وأثر النبى والإسكندرية ورشيد .

(ب) نشرة أسبوعية باللغتين العربية والإنجليزية عن أسعار أهم الحاجيات الأولية بالجملة والتجزئة فى أسواق القاهرة .

(حـ) نشرة القطن ــ أسبوعية باللغتين العربية والإنجليزية .

(د) النشرة الشهرية للإحصاءات الإقتصادية والزراعية ، باللغتين العربية والإنجليزية .

(هـ) نشرة الأرقام القياسية لأسعار الجملة ونفقات المعيشة بالنسبة للحرب العالمية الأخيرة وهذه النشرة شهرية .

(و) نشرة أوراق العملة المصدرة والنقود المعاونة لها المتداولة وتصدر كل شهر باللغتين العربية والفرنسية .

(ز) نشرة الأرقام القياسية العامة لأسعار الجملة والتجزئة ونفقات المعيشة بالنسبة لأسعار ما قبل الحرب العالمية الأولى والحرب العالمية الثانية وهذه النشرة شهرية باللغتين العربية والفرنسية .

٤ ــ التجارة الخارجية

(ا) الملخص الشهرى لتجارة الخارجية باللغتين العربية والإنجليزية .

(ب) نشرة (أرقام أولية) عن التجارة الخارجية مبيناً بها أهم الأصناف والبلاد وهذه النشرة شهرية .

(جـ) نشرة سنوية عن تجارة مصر الخارجية باللغتين العربية والإنجليزية .

(د) ملخص حركة التجارة طبقاً للدليل الدولى ــ سنوى باللغة الإنجليزية .

٥ – الملاحة البحرية

(ا) نشرة حركة الملاحة الخارجية فى موانىء الإسكندرية وبورسعيد والسويس باللغتين العربية والفرنسية – شهرية .

(ب) نشرة حركة الملاحة والتجارة البحرية فى الثغور المصرية وفى قناة السويس وتصدر مرة فى كل ثلاثة شهور باللغتين العربية والإنجليزية .

(حـ) إحصائية سنوية لحركة الملاحة والتجارة البحرية فى الثغور المصرية وفى قناة السويس – سنوية باللغتين العربية والإنجليزية .

٦ – إحصاءات متنوعة

(ا) نشرة السيارات باللغة الإنجليزية – شهرية وسنوية .

(ب) إحصاء الأجور وساعات العمل بالمملكة المصرية باللغتين العربية والإنجليزية ويصدر مرة كل ستة أشهر .

(حـ) الإحصاء السنوي للجيب باللغتين العربية والفرنسية .

(د) إحصاء شركات المساهمة التى يوجد استغلالها الرئيسى فى مصر ويصدر سنوياً .

(هـ) إحصاء سنوى عن الموظفين والمستخدمين بالحكومة والهيئات العامة .

(و) الإحصاء السنوي العام باللغتين العربية والفرنسية .

(ز) الإحصاء العام لمعاهد التعليم بالمملكة المصرية ويصدر كل ثلاث سنوات باللغتين العربية والفرنسية .

(حـ) إحصاء الإنتاج الصناعى ويصدر كل ثلاث سنوات باللغتين العربية والإنجليزية .

الباب الثالث

طرق عرض البيانات الإحصائية

أبسط طريقة لعرض البيانات الإحصائية هى أن تدمج هذه البيانات – أى الأرقام الأولية – فى صيغة كتابية . فنقول مثلا « بلغت قيمة الرسوم التى حصلتها مصلحة الجمارك حتى أول أكتوبر سنة ١٩٥١ مبلغ ٩٩,٧٥٥,١٠٤ جنيهاً فى حين بلغت هذه الرسوم فى نفس المدة من العام الماضى ٩٧,٢٥٥,١٠٠ جنيهاً أى أن الرسوم المحصلة حتى أول أكتوبر من السنة الحالية تزيد بمقدار ٢ ١/٢ مليون جنيه تقريباً عن تلك المحصلة فى نفس المدة من السنة الماضية . وترجع معظم الزيادة إلى الرسم الإضافى على الكحول الذى بلغ المتحصل منه ١ ١/٢ مليون جنيه تقريباً » . وعلى الرغم من أن هذه الطريقة فى عرض البيانات الإحصائية تمتاز على غيرها من الطرق بالسهولة التامة الراجعة إلى إمكان توضيح الأرقام بعبارات تفسيرية كلما دعت الحاجة إلى ذلك إلا أنها لا تعد بين الطرق الفنية فى العرض الإحصائى ، فهى طريقة مطولة وعقيمة وتتطلب وقتاً طويلا فى القراءة الأمر الذى يجعل الملل يتسرب إلى القارىء وقلما يكلف الإنسان نفسه مشقة الاطلاع على إحصاءات معروضة بهذه الكيفية . ومن الواضح أنه يتعذر عرض بيانات خاصة بعدد كبير من السنين بهذه الطريقة ولذلك نكتفى بمجرد الإشارة إليها .

أما الطرق الفنية فى عرض البيانات الإحصائية فتنقسم إلى قسمين رئيسين :

أولا – العرض الجدولى – أو عرض البيانات فى شكل جداول .

ثانياً – العرض البيانى – وتشتمل هذه الطريقة على عدة فروع أهمها :–

١ ــ الرسوم البيانية .

٢ ــ طريقة الأشكال الهندسية .

٣ ــ العرض التصويرى .

وسنشرح جميع هذه الطرق بعضها فى شىء من التفصيل والبعض الآخر بإيجاز حسب ما تتطلبه كل حالة .

العرض الجدولى

عرض البيانات الإحصائية فى شكل جداول طريقة سهلة شائعة الإستعمال ، وقد رأينا فى الباب السابق كيف أن هناك آلات لتبويب البيانات ودرجها فى شكل جداول الأمر الذى يسهل كثيراً من مهمة الإحصائى . وأهم ما يجب مراعاته فى عمل الجداول هى أن تكون واضحة وسهلة بحيث يمكن تفهم محتوياتها بعد دراسة سريعة. هذه هى القاعدة الأساسية فى عمل الجداول . وهناك بضع نقط يجب الإشارة إليها وكلها ترمى إلى تحقيق القاعدة السابقة من حيث وضوح الجدول وسهولة فهم محتوياته وأهم هذه النقط هى :ــ

أولا ــ عنوان الجدول .

لا بد أن يحمل كل جدول إحصائى عنوانا مميزاً لتعيينه وتعريفه . ويجب أن يكتب العنوان بدقة واختصار كى يعطى القارىء فكرة موجزة سريعة عن محتويات الجدول . ويحسن أن يعطى الجدول رقماً خاصاً لسهولة الرجوع إليه.

ثانياً ــ المذكرات الافتتاحية والتفسيرية .

تكتب المذكرات الافتتاحية عادة تحت العنوان بحروف صغيرة إذا دعت الحاجة إليها . والمقصود من كتابتها إيضاح محتويات الجدول بوجه عام . فمثلا

إذا عملنا جدولا يتعلق بالاحتياطى العام للحكومة المصرية نكتب تحت العنوان بحروف صغيرة مذكرة تبين أن فائض الإيرادات السنوية يضاف كل عام إلى رصيد خاص يكون هذا الاحتياطى . والمذكرة الافتتاحية تعتبر ضرورية فى مثل هذه الحالة لأن نظام الاحتياطى هذا غير مستعمل فى غالبية الدول . على أن المذكرات الافتتاحية نادرة الاستعمال جداً لأن عنوان الجدول وطريقة تقسيمة يكفيان عادة لتوضيح الغرض منه .

أما المذكرات التفسيرية فتكتب فى أسفل الجدول والمقصود منها تفسير رقم أو بيان خاص فى صلب الجدول . وجرت العادة على أن يوضع فوق الرقم أو البيان المراد تفسيره إشارة معينة (* مثلا) تظهر أن هناك تفسيراً فى أسفل الجدول .

ثالثاً — مذكرات المصادر .

تكتب هذه إما تحت العنوان أو فى أسفل الجدول كما هو المتبع عادة ، والغرض منها بيان المصادر التى استقيت منها البيانات التى يحتويها الجدول . وكتابة هذه المذكرات من أهم الأمور الواجب مراعاتها فى عمل الجدول فهى تزيد من ثقة القارئ فى البيانات المعروضة فضلا عن أنها تمكنه من الرجوع إلى هذه المصادر الأصلية لاستيفاء بعض المعلومات التى قد يرغب فى الاطلاع عليها .

رابعاً — تقريب الأرقام .

لكى تكون الأرقام مبسطة سهلة يحسن تقريبها إما إلى أقرب عدد صحيح أو ألف أو مليون خصوصاً إذا كان الجدول سيعرض على الشخص العادى وكثيراً ما تحذف أصفار الألوف والملايين ويبين ذلك فى خانات البيان الموجودة فى رأس الجدول أو فى مذكرة تفسيرية .

خامساً – حجم الجدول وشكله العام .

يحسن بوجه عام أن يراعى فى تكوين الجدول ألا يكون طويلا وضيقاً أو قصيراً ومتسعاً . فإذا كان الجدول يحتوى على عمودين أو ثلاثة فقط فلايجوز أن نجعل طوله يمتد إلى عشرين أو ثلاثين سطراً . كذلك لا يجوز عمل جدول يحتوى على عدة تقسيمات أفقية فى حين أن الأرقام الواردة فى صلبه لا تشغل أكثر من سطر واحد أو سطرين . فلو فرض أن كان لدينا بيانات خاصة بعدد كبير من السنين وأردنا عرضها فى شكل جدولى فيمكننا تقسيم الجدول إلى قسمين (أو أكثر) على أن يفصل كل قسم عن الآخر بخط مزدوج كما هو الحال فى الجدول رقم ٢ . وقد يدعو الحال إلى استخدام طريقة الورقة المطوية إذا كان فراغ الصفحة لا يكفى لعمل جدول يحتوى على جميع البيانات المطلوب عرضها . والنموذج الآتى يوضح بعض النقط السابق ذكرها .

جدول رقم ٢

محصول القطن المصرى ومتوسط إنتاج الفدان

١٨٨٠ – ١٨٩١

السنة	محصول القطن الإجمالى « بآلاف القناطير »	متوسط إنتاج الفدان « بالقنطار »	السنة	محصول القطن الإجمالى « بآلاف القناطير »	متوسط إنتاج الفدان « بالقنطار »
١٨٨٠	٢٧٧٦	٣,٠١	١٨٨٦	٢٩٣١	٢,٨٨
١٨٨١	٢٩١٢	٢,٦٨	١٨٨٧	٢٩٣٧	٢,٩٨
١٨٨٢	٢٢٤٨ *	٢,١٠	١٨٨٨	٢٧٢٣	٢,٤٩
١٨٨٣	٢٦٩٤	٢,٧٧	١٨٨٩	٣١٨٣	٣,٠٦
١٨٨٤	٣٦١٦	٣,١٣	١٨٩٠	٤٠٧٢	٣,٩٧
١٨٨٥	٢٩٢٣	٢,١٩	١٨٩١	٤١٩٥	٣,٩٩

المصدر : الإحصاء السنوى العام للقطر المصرى عن سنة ١٩٠٩ ص ١١٦

* لاحظ هبوط المحصول ومتوسط إنتاج الفدان نظراً لحالة الاضطراب التى اجتازتها البلاد فى هذه السنة.

سادساً – وحدات القياس .

إذا لم تكن وحدات القياس واضحة من تلقاء نفسها كوحدات الأعمار مثلا ، يجب إيضاحها إما بمذكرة افتتاحية أو فى أسفل الخانات الخاصة بالبيان .

وعلى وجه الخصوص يجب توضيح ما إذا كانت الأرقام الواردة فى الجدول تمثل وحدات مطلقة أم نسب مئوية . وفى حالة استخدام هذه النسب الأخيرة يجب أن نبين بوضوح طبيعة النسب المئوية – هل هى نسب زيادة أو نقص ؟ وهل الزيادة (أو النقص) شهرية أو سنوية أو عشر سنوية ؟

سابعاً – تسطير الجدول .

يفضل معظم الإحصائيين التقليل من التسطير فى الجدول بوجه عام . وبنوع خاص يجب عند استخدام التسطير الرأسى ألا يترك أحد جانبى الجدول مفتوحاً (أى بدون تسطير) بينما يكون الجانب الآخر مغلقاً . أما التسطير الأفقى فيجب تجنبه فى صلب الجدول ومحاولة قصر استعماله على فصل أجزاء الجدول عن بعضها ، كفصل خانة المجموع عن بقية الجدول مثلا . وإذا كان الجدول يشتمل على بيانات يراد تأكيد المقارنة بينها كما فى حالة كون هذه البيانات تمثل قيم ظاهرة معينة فى عهدين مختلفين – قبل وبعد الحرب مثلا – فيحسن استخدام الخط المزدوج لفصل البيانات الخاصة بكل عهد على حدة كما هو موضح بالجدول رقم ٣ .

جدول رقم ٣

النتائج المالية لشركة مصر للغزل والنسج

قبل و بعد منحها الإعانة الحكومية*

١٩٢٨ — ١٩٣٨

السنة	رأس المال المدفوع « بالجنيهات »	الاحتياطيات « بالجنيهات »	الأرباح الموزعة « بالجنيهات »
١٩٢٨	٣٠٠,٠٠٠	—	—
١٩٢٩	٣٠٠,٠٠٠	—	—
١٩٣٠	٣٠٠,٠٠٠	٤٣٣٠	—
١٩٣١	٤٥٠,٠٠٠	٢٨٠٧	—
١٩٣٢	٤٥٠,٠٠٠	١٦٧٠٩	٢٢٥٠٠
١٩٣٣	٥٠٠,٠٠٠	٢٨٩١٧	٣٠٠٠٠
١٩٣٤	٨٠٠,٠٠٠	٩٩٠٦٥	٤٨٠٠٠
١٩٣٥	٨٠٠,٠٠٠	١٠٩٣٨٥	٤٨٠٠٠
١٩٣٦	١,٠٠٠,٠٠٠	١٥٨٩٩٥	٤٨٠٠٠
١٩٣٧	١,٠٠٠,٠٠٠	١٦٧٦٤٢	٦٠٠٠٠
١٩٣٨	١,٠٠٠,٠٠٠	١٧٨١٧٤	٧٠٠٠٠

المصدر : إحصاء الشركات المساهمة التى يوجد عملها الرئيسى بمصر .

ثامناً — إرشاد العين وتسهيل القراءة .

يحسن فى عمل الجداول لا سيما تلك التى تحتوى على أعمدة طويلة من الأرقام أن نترك سطراً أو فراغاً صغيراً بعد كل بضعة أرقام وذلك حتى لا تختلط

* فى سنة ١٩٣١ قررت الحكومة المصرية منح شركات غزل ونسج القطن إعانة قدرها ٢٠ قرشاً عن كل قنطار من القطن الخام تستعمله كمادة أولية .

الأرقام على العين وتسهل قراءتها . ويلاحظ أن معظم الجداول الرياضية التى تحتوى على أعمدة طويلة من الأرقام مثل جداول اللوغاريتمات والمربعات والجذور يراعى فى تركيبها هذه النقطة وذلك لكى يسهل استخراج الأرقام المطلوبة منها بدون إجهاد النظر فى عملية البحث عنها . والنموذج الآتى يوضح هذه النقطة :—

جدول رقم ٤

الأرقام القياسية لودائع صناديق التوفير وعدد المودعين بالقطر المصرى (١٩٠١ — ١٩١٢) «سنة الأساس ١٩٠١ = ١٠٠»

السنة	الودائع	عدد المودعين
١٩٠١	١٠٠	١٠٠
١٩٠٢	١٨١	٢١٠
١٩٠٣	٢٩٩	٣٠٠
١٩٠٤	٣٧٩	٤٣٣
١٩٠٥	٤٩٨	٦٤٤
١٩٠٦	٦٨٤	٨٧٦
١٩٠٧	٨٢١	١١٠٠
١٩٠٨	٨٤٦	١٢٨٧
١٩٠٩	٨٨٧	١٣٧٨
١٩١٠	٩٦١	١٥٤٤
١٩١١	١٠١٢	١٧٤٦
١٩١٢	١٢٠٠	٣٩٣١

المصدر : الأرقام القياسية محتسبة على أساس بيانات مصلحة البريد فى تقاريرها السنوية .

العرض البيانى

١ – طريقة الرسم البيانى

التمثيل البيانى للمعادلات الجبرية :

إذا أردنا تحديد موقع نقطة معينة فى فراغ ذى بعد واحد [١] مثل تحديد موقع قطار على خط حديدى ، فكل ما نحتاج إليه من بيانات هو معرفة المسافة التى يبعدها القطار عن نقطة ثابتة نتخذها كأساس لبدء القياس . فنقول مثلا أن القطار يبعد عن نقطة بدء القياس هذه أو « نقطة الأصل » عشرة كيلو مترات شرقاً أو غرباً (على فرض أن الخط يمتد من الشرق إلى الغرب) . ويمكن أن نتقدم خطوة أخرى فنستعيض عن وصف الاتجاه بتعبير رمزى فنعتبر جميع النقط الواقعة شرق نقطة الأصل موجبة وتلك الواقعة غرب هذه النقطة سالبة . وعندئذ يكفى لتحديد موقع أى نقطة فى هذا الفراغ أن نقول إنها تبعد (+ س) أو (– س) من الوحدات عن نقطة الأصـــل .

وإذا انتقلنا إلى الفراغ ذي البعدين – أى إذا انتقلنا من ميدان الخطوط المستقيمة إلى ميدان السطوح – نجد أن تحديد موقع نقطة على سطح معين (مثل فراغ ورقة الكتابة) يتطلب معرفة بيانات أخرى بالإضافة إلى تلك التى مكنتنا من تحديد موقع نقطة فى الفراغ ذى البعد الواحد . إذ فى هذه الحالة لو بدأنا القياس من نقطة ثابتة نعتبرها نقطة الأصل ، فلا يكفى أن نقول إن النقطة المراد تحديد موقعها تبعد عنها بمقدار (+ ١٠) أو (– ١٥) وحدة . فهناك بالإضافة إلى الإتجاهين الشرقى والغربى اتجاهان آخران شمالى وجنوبى

---

(١) الفراغ ذو البعد الواحد (One Dimensional Space) هو عبارة عن خط مستقيم يمتد إلى ما لا نهاية .

نقطة الأصل . وعلى ذلك فلتحديد نقطة معينة على سطح يجب أن نعلم كم تبعد هذه النقطة عن نقطة الأصل شرقاً أو غرباً وكم تبعد عنها شمالا أو جنوباً . وقياساً على ما فعلنا فى الحالة السابقة نعتبر الاتجاه الشمالى موجباً والجنوبى سالباً . وعندئذ يكفى لتحديد موقع أى نقطة فى الفراغ ذى البعدين أن نقول إنها تبعد (+س ، +ص) أو (+ س ، — ص) أو (— س ، + ص) أو (— س ، — ص) عن نقطة الأصل ، على فرض أن س تمثل عدد وحدات القياس التى نستخدمها فى الاتجاه الشرقى الغربى و ص تمثل عدداً من وحدات القياس فى الاتجاه الشمالى الجنوبى .

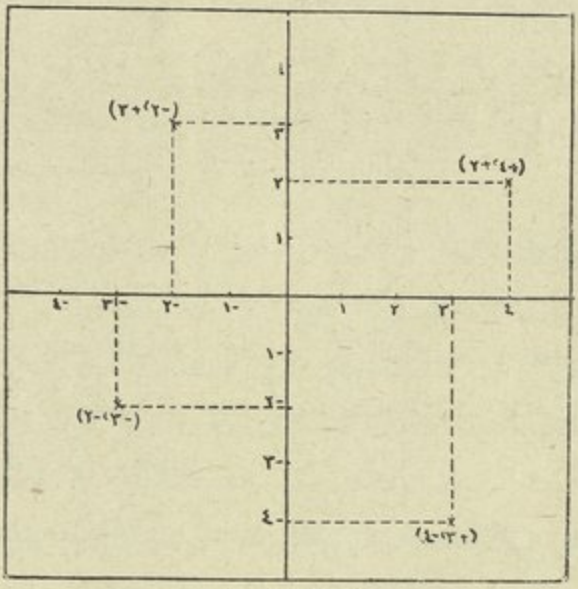
ولتسهيل عملية قياس الأبعاد فى الاتجاهات الأربعة السابق ذكرها ، نرسم من نقطة الأصل خطين متعامدين أحدهما يمتد أفقياً من الغرب إلى الشرق والآخر يمتد رأسياً من الجنوب إلى الشمال بحيث يقطع الأول فى نقطة الأصل . ثم نقسم كل منهما أقساماً متساوية يمثل كل منها وحدة من وحدات القياس التى نستعملها .

ومن المتفق عليه أن يسمى الخط الأفقى المار بنقطة الأصل المحور الأفقى أو المحور السينى والخط الرأسى المار بنفس النقطة المحور الرأسى أو المحور الصادى . كما أن بعد أى نقطة عن نقطة الأصل فى الاتجاه الأفقى (الشرقى أو الغربى) يسمى الأحداثى السينى وبعدها عن نقطة الأصل فى الاتجاه الرأسى (الشمالى أو الجنوبى) يسمى الأحداثى الصادى . وباستخدام هذه الطريقة فى التعبير نستطيع أن نقول إنه يمكن تحديد موقع أى نقطة إذا ما عرفنا أحداثييها أى بعديها عن نقطة الأصل . فلو كان لدينا نقطة إحداثياها (+ ٤ ، + ٢)[1] . نعلم بمجرد النظر إلى إشارتى الإحداثيين أن هذه النقطة تقع شرق نقطة الأصل وشمالها ، كذلك النقطة التى إحداثياها

---

(١) من المتفق عليه للتمييز بين الإحداثيين أن تكتب قيمة الاحداثى السينى دائماً فى الأول .

(+ ٣ ، — ٤) تقع شرق نقطة الأصل وجنوبها . كما أن النقطة التى إحداثياها (— ٢ ، + ٣) تقع غرب نقطة الأصل وشمالها والنقطة التى إحداثياها (— ٣ ، — ٢) تقع غرب نقطة الأصل وجنوبها . والرسم الآتى يبين مواقع هذه النقط الأربع بالنسبة لنقطة الأصل والمحورين .

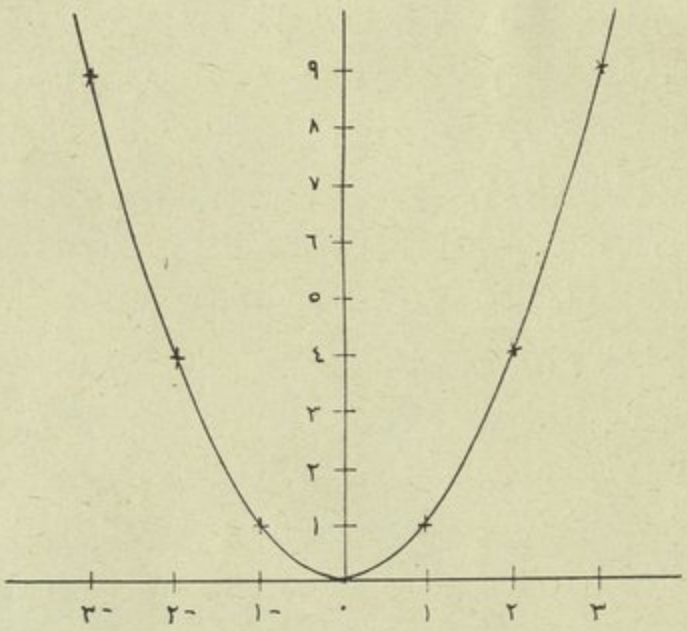


(شكل رقم ٦)
تحديد موقع النقط على الرسم البيانى

الآن وقد عرفنا كيف يمكن تحديد موقع أى نقطة بالنسبة للمحورين ، يمكننا أن ننتقل خطوة أخرى ونحدد خط سير هذه النقطة إذا عرفنا الإحداثيات التى تصف الأوضاع المختلفة التى تتخذها النقطة . فمثلا لو كانت قيم هذه الإحداثيات كما هو موضح فى الجدول الآتى :

قيم س	٣	٢	١	٠	١-	٢-	٣-
قيم ص	٩	٤	١	٠	١	٤	٩

فيمكننا تحديد خط سير هذه النقطة بالنسبة للمحورين على النحو المبين بالشكل الآتى :ـ



(شكل رقم ٧)
تمثيل المعادلة ص = س٢ بيانيا

غير أنه بالتأمل فى جدول الإحداثيات السابق نجد أن هناك علاقة واضحة بين قيم س وقيم ص . فقيمة ص هى دائماً مربع قيمة س بحيث يمكننا أن نصف العلاقة بينهما جبريا بالمعادلة ص = س٢. أما خط سير النقطة الموضح بالشكل رقم ٧ فهو عبارة عن التمثيل البيانى لهذه المعادلة . وكما أن القيم العددية للمتغيرين س ، ص مرتبطة ببعضها طبقاً لنظام معين تحدده العلاقة الجبرية فإن النقط التى تمثل هذه القيم فى الرسم البيانى مرتبطة ببعضها فى شكل هندسى له خواصه التى تميزه عن بقية الأشكال . فلو أخذنا أى نقطة على المنحنى الموضح بالشكل رقم ٧ وقسنا إحداثييها الأفقى والرأسى، نجد طول الأخير يعادل

دائماً طول مربع الأول . فكأن العلاقة الجبرية الثابتة بين المتغيرين منعكسة فى الخواص الهندسية المشتركة بين جميع النقط التى تمثل قيم المتغيرين والتى يجمعها خط بيانى معين . هناك إذن صلة وثيقة بين الجبر والهندسة أو بين التعبير الرقمى أو الرمزى والتعبير البيانى . فكل معادلة جبرية تربط المتغيرين س ، ص يمكن تمثيلها بيانيا على النحو السابق . ويختلف شكل الخط البيانى باختلاف درجة المعادلة . فمعادلات الدرجة الأولى (أى تلك التى تكون قوى المتغيرين فيها الواحد الصحيح) مثل ص = ٤ س + ٥ يمكن تمثيلها دائماً بمستقيمات . ومعادلات الدرجة الثانية (أى تلك التى يكون أحد المتغيرين فيها أو كلاهما مرفوعاً إلى القوة الثانية) مثل س^{٢} + ص^{٢} = ١٦ يمثلها إما قطع مكافىء أو قطع زائدى أو دائرة أو غير ذلك من المنحنيات حسب نوع المعادلة .

وعلى الرغم من بساطة الفكرة التى تقوم عليها العلاقة بين المعادلات الجبرية والمنحنيات إلا أن اكتشافها يعتبر حديث العهد نسبياً . فقد ظهر أول بحث عنها فى كتاب (Method of Science) الذى وضعه ديكارت (Descartes) فى سنة ١٦٣٧ وبه فتح آفاقاً جديدة وأوجد فرعاً جديداً فى العلوم الرياضية هو الهندسة التحليلية التى تسمى أحياناً هندسة ديكارت (Cartesian Geometry) نسبة إليه وتمييزاً لها عن هندسة يوقليدس (Euclidean Geometry) التى ترجع إلى عام ٣٠٠ قبل الميلاد . على أن فكرة استخدام المحورين لتحديد مواقع النقط فى الفراغ وهى الفكرة الأساسية فى هندسة Descartes كانت معروفة لقدماء المصريين . فقد استخدمها قدماء المصريين فى مسح الأراضى بنفس الطريقة التى نستعملها فى الغرض ذاته[١] . كذلك استخدم الإغريق والرومان هذه الطريقة فى مسح الأراضى وعمل الخرائط .

---

(١) راجع كتاب Mathematics in Human Affairs تأليف F. W. Kokomoor ص ٤٩٨ .

سوف لا نخوض هنا فى موضوع العلاقة بين المعادلات الجبرية والمنحنيات الذى يستطيع القارىء أن يتعمق فيه بالرجوع إلى أحد كتب الهندسة التحليلية(١). وإنما قصدنا من إيراد هذه المقدمة أن نشرح للقارىء المبتدىء فى شىء من التفصيل طريقة تحديد مواقع النقط بالنسبة للمحورين وأن نظهر له كيف يمكن تصوير العلاقة بين المتغيرات كما تحددها المعادلة الجبرية وكيف أن الخواص المميزة لهذه العلاقة تنعكس بدقة فى الرسم البيانى .

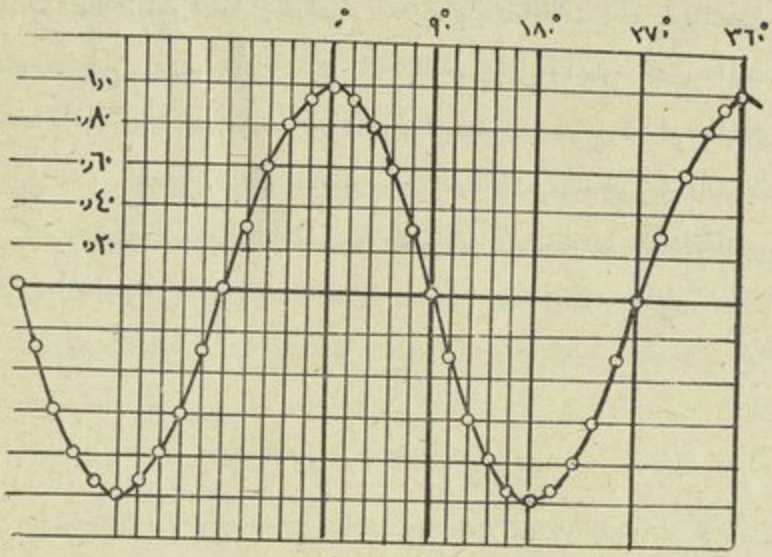
الرسم البيانى كإحدى طرق العرض الإحصائى

بنفس الطريقة التى أمكننا بها تصوير العلاقة بين المتغيرات الرياضية يمكننا تصوير العلاقة التى نشاهدها ونسجلها بين الظواهر المختلفة . غير أن ثمة فارق كبير بين العلاقة الرياضية والعلاقة المشاهدة . فالمتغيرات التى تحكم العلاقة بينها معادلة جبرية تظهر إذا صورت فى رسم بيانى على شكل منحن ممهد أى خال من الذبذبات والتعاريج وذلك لأنها تخضع فى تغيرها لنظام دقيق ثابت . فالمنحنى ما هو إلا خط سير نقطة معينة . وما دامت هذه النقطة تتبع فى سيرها نظاماً معيناً فإن المنحنى الناتج من تتبع سيرها يكون منظماً ممهداً حاملا لكل خواص المعادلة الجبرية . وللتثبت من صحة ذلك انظر مثلا إلى المنحنى الممهد الذى يوضح كيف تتغير قيمة جيب تمام الزاوية طبقاً لتغير قيمة الزاوية أو بعبارة أخرى المنحنى الذى يمثل المعادلة ص=جتاس. (شكل رقم ٨) .

أما المتغيرات التى نسجل قيمها من واقع المشاهدة والتى لا يحكم العلاقة بينها قانون رياضى فإنها إذا رصدت على الرسم البيانى تظهر عادة فى شكل

---

(١) انظر مثلا كتاب الهندسة التحليلية تأليف نصيف سعد وصادق بشارة أو كتاب R. G. D. Allen تأليف Mathematical Analysis for Economists

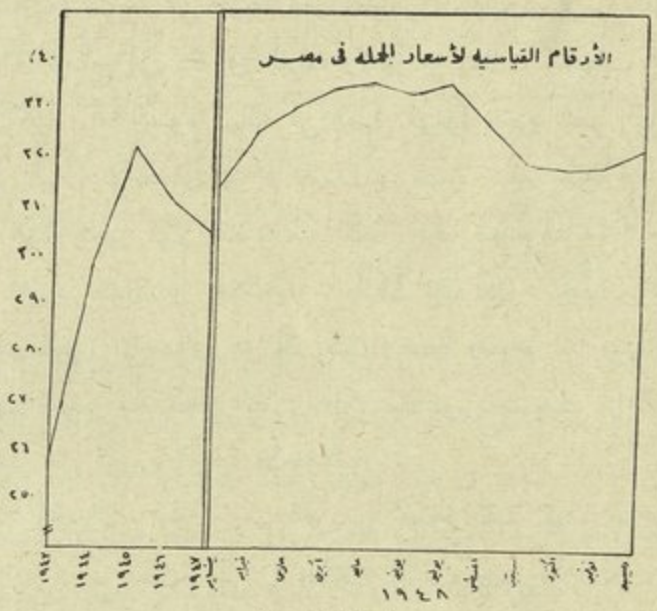


(شكل رقم ٨)

المنحنى الذى يمثل المعادلة ص = جتاس

خطوط منكسرة كثيرة الذبذبات والتعاريج . والسبب فى ذلك هو أن هذه الظواهر لا تخضع فى تغيرها لقانون دقيق وإنما تخضع عادة لعدة مؤثرات بعضها يحدث فى شىء من الانتظام وبعضها يحدث بطريقة عرضية فجائية فيؤثر على قيمة الظاهرة دون ما قاعدة على النحو الذى سنفصلة فى الباب الخاص بتحليل السلسلات الزمنية . والشكل رقم ٩ يبين تغير ظاهرة خاضعة لهذه المؤثرات .

على أن تغير الظواهر المشاهدة على النحو السابق بيانه لا يمنع من أن نلاحظ أحياناً فى الرسوم البيانية التى تمثلها شيئاً من النظام الذى ربما يلقى كثيراً من الضوء على وجود علاقة معينة بين هذه الظواهر التى نسجل قيمها من واقع المشاهدة . خذ مثلا ظاهرتين مثل ثمن سلعة معينة والطلب عليها فى إحدى الأسواق . إذا جمعنا أزواجاً من القيم المتناظرة التى يمثل كل زوج منها ثمن السلعة والطلب عليها كما نسجلها بإحدى طرق جمع البيانات الإحصائية ،



(شكل رقم ٩)

الأرقام القياسية لأسعار الجملة فى مصر ١٩٤٣ – ١٩٤٨

ثم رصدنا هذه القيم على رسم بيانى فإننا نحصل على منحن غير متصل وغير ممهد وإنما نلاحظ أنه يهبط من اليسار إلى اليمين وأنه محدب بالنسبة لنقطة الأصل. وإذا كررنا هذه العملية بالنسبة لسلعة أخرى نصل إلى نتيجة لاتختلف كثيراً عن السابقة . وتلك الخاصية التى تظهر فى منحنى الطلب تعكس فى الواقع علاقة معروفة لدى الاقتصاديين هى أنه كلما هبط الثمن زاد الطلب و بالعكس . والذى نود أن نستخلصه من هذا أنه على الرغم من أن الرسوم البيانية الإحصائية – أى تلك التى تسجل تغير الظواهر كما نشاهده – لا تظهر لنا علاقات دقيقة ثابتة بين الظواهر إلا أننا كثيراً ما نستطيع أن نقف منها على طريقة تغير الظواهر ، بل إنها قد توحى إلينا أحياناً بوجود علاقات بين الظواهر التى نسجلها .

والرسوم البيانية الإحصائية تفضل الجداول فى أنها تعطى الناظر إليها

فكرة صحيحة سريعة عن طريقة تغير الظواهر المسجلة فى الرسم . فالجداول الإحصائية تحتاج إلى دقة فى القراءة وكثيراً ما يتطلب استيعاب محتوياتها عناية وتركيزاً . أما الرسوم البيانية فمن السهل فهمها بمجرد النظر إليها . على أن الرسم البيانى الإحصائى إذا لم تصنعه يد خبيرة قد يكون من التعقيد والإبهام بحيث يتعذر فهم المقصود منه كلية . وهنا تنعدم الفائدة المرجوة منه ويضيع المجهود المبذول فى عمله هباء . ولذلك فإن أهم ما يجب مراعاته فى عمل الرسم البيانى الإحصائى هو أن يكون واضحاً وضوحاً تاماً وسهلا بحيث يمكن فهم الغرض منه بمجرد النظر . ولكى يتحقق ذلك يجب مراعاة النقط الآتية عند عمل الرسوم البيانية الإحصائية :—

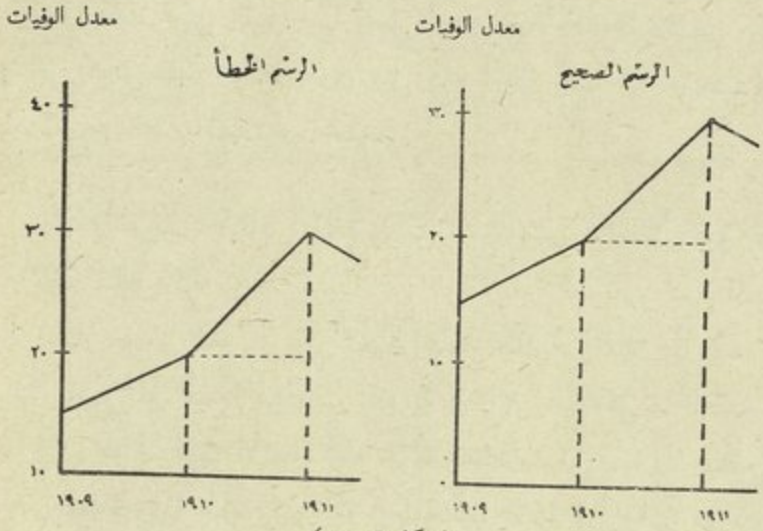
أولا — المحوران : يجب رسم المحورين بخطوط ثقيلة فهذا يساعد على وضوح الرسم . أما الخطوط الرأسية والأفقية الأخرى فيجب أن تكون خفيفة بل إنه يحسن رسمها بقلم رصاص خفيف ثم إزالتها بعد الانتهاء من الرسم فإن فائدتها قاصرة على المساعدة فى تحديد مواقع النقط ومتى تم ذلك فالأفضل ألا تظهر فى الرسم .

ثانياً — تخطيط المنحنيات : يختلف الرسم البيانى الإحصائى عن الرسم البيانى الرياضى اختلافاً بيناً من حيث طريقة تخطيط المنحنيات . فقد تعلمنا فى المدارس الثانوية أن نرسم المستقيمات والمنحنيات الرياضية بقلم رفيع جداً فتظهر خفيفة بحيث لا تكاد تراها إلا العين الفاحصة المدققة . ويرجع السبب فى ذلك إلى أننا قد نستعمل هذه الرسوم الرياضية فى قياس أبعاد وزوايا أو قد نحاول تحديد موقع نقطة تماس بين مستقيم ودائرة أو منحنى ، الأمر الذى يتطلب كثيراً من الدقة فى العمل . أما الرسوم البيانية الإحصائية فليس فيها ما يدعو إلى هذه الدقة الهندسية . وإنما الوضوح وسهولة الفهم مع شىء من الذوق السليم هو كل ما يجب توفره فيها . ولذلك يجب رسم المنحنيات فيها بخطوط ثقيلة واضحة تبرز للعين دون حاجة إلى فحص أو تدقيق . وفى

حالة وجود عدة منحنيات تمثل ظواهر مختلفة على نفس الرسم فلا بد أن يكون هناك فروق واضحة بينها ويكون ذلك باستعمال ألوان مختلفة لكل منها وإذا لم يتيسر ذلك يمكن التمييز بينها بجعل أحدها خطاً متصلا والاخر متقطعاً أو على شكل نقط وهكذا . وفى مثل هذه الحالة يجب أن يصحب الرسم بيان تفسيرى لتلك الألوان أو الخطوط .

ثالثاً — تقريب الأرقام : المقصود من الرسم البيانى الإحصائى هو إعطاء من يطلع عليه فكرة سريعة تقريبية عن طريقة تغير الظواهر وليس المقصود منه إظهار مقدار التغير فى قيم الظواهر بدقة حسابية . فمثلا إذا عملنا رسماً بيانياً يوضح تغير قيمة الصادرات المصرية فى ٢٠ عاماً فكل ما نقصده من هذا الرسم هو أن نظهر الاتجاه العام لحركة الصادرات من جهة وأن نوضح أى السنين حدث فيها هبوط كبير أو زيادة غير عادية من الجهة الأخرى . غير أننا لانقصد من مثل هذا الرسم أن يعلم من يطلع عليه أن قيمة الصادرات المصرية بلغت فى سنة ١٩٣٧ ٣٨,٦٦٥,١٠٤ جنيهاً . بينما كانت فى سنة ١٩٣٦ ٣٢,٩٧٩,٢٤٣ جنيهاً . فليس من المفروض ولا من المعقول أن نحاول إظهار مثل هذه الأرقام على الرسم البيانى العادى . وكل ما نستطيع عمله هو أن نقرب مثل هذه الأرقام إلى أقرب مائة ألف أو حتى إلى أقرب مليون من الجنيهات .

رابعاً — مقياس الرسم : من أهم قواعد الرسم البيانى الإحصائى هو بدء القياس الرأسى (المحور الصادى) بالصفر . أما القياس الأفقى الذى يمثل الزمن فى أغلب الحالات فيمكن بدؤه من أى سنة نختارها بحيث يتسع المجال لجميع السنين المطلوب عرض البيانات الخاصة بها . وكثيراً ما تهمل هذه القاعدة فتكون النتيجة تضليل الشخص الذى يطلع الرسم وتتضح أهمية هذه النقطة من مقارنة الرسمين الآتيين : —



(شكل رقم ١٠)

يوضح ضرورة بدء المحور الرأسى بالصفر

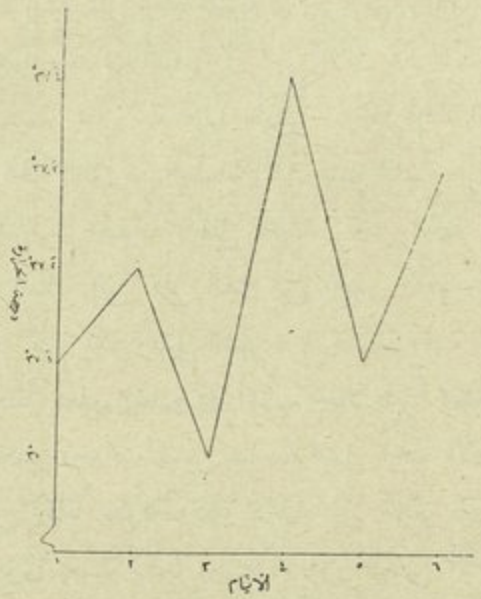
إذا نظرت إلى الرسم الخاطئ ظهر لك أن معدل الوفيات فى سنة ١٩١١ يعادل ضعف معدل الوفيات فى سنة ١٩١٠ والواقع غير ذلك كما تظهره الأرقام والرسم الصحيح ، فهو أكبر منه بمرة ونصف مرة فقط . وقليل ممن يطلعون على الرسم يتنبهون إلى هذه المغالطة الناشئة عن عدم بدء القياس الرأسى بالصفر . ولما كان من واجب الإحصائى جعل طريقة العرض سهلة يمكن تفهمها بمجرد النظر فعليه أن يتحاشى تضليل من يطلع على الرسم بهذه الكيفية .

أما من جهة وحدات القياس على الرسم فهذه مسألة لا يمكن وضع قاعدة ثابتة لها ، فهى تترك لذوق الرسام وحسن تصرفه . غير أنه يجب تحاشى

إظهار الحقائق بصورة مبالغ فيها أو طمس الحقائق عمداً عن طريق تكبير أو تصغير وحدات القياس إلى حد غير معقول . افرض مثلا أننا نرصد تقلبات درجة حرارة أحد المرضى على رسم بيانى ، وأن درجة حرارته فى ستة أيام متتالية كانت ثابتة تقريباً عند حد معين كالاتى :—

١' ٣٧° ، ٢' ٣٧° ، ٣٧° ، ٤' ٣٧° ، ١' ٣٧° ، ٣' ٣٧° .

من السهل جداً أن نغالى فى إظهار هذه التقلبات التافهة فى درجة الحرارة باختيار وحدات غير مناسبة فنظهرها من الحدة والعنف بحيث يقلق من ينظر إليها على صحة المريض . والشكل الاتى يوضح هذه النقطة :—



(شكل رقم ١١)

رسم بيانى مضلل يبالغ فى إظهار التقلبات فى قيمة الظاهرة

وبالمثل يمكن باختيار وحدات غير مناسبة للقياس إخفاء تقلبات يكون من الواجب إظهارها فى الرسم .

ويلاحظ القارىء إذا نظر إلى الرسم السابق (شكل رقم ١١) أنه نظراً لضرورة بدء القياس الرأسى بالصفر (دفعاً للتضليل كما سبق أن بينا) فى الوقت الذى يوجد لدينا فراغ كبير لا سبيل إلى استعماله (من . إلى ٣٧) لجأنا إلى طريقة « كسر » المحور الرأسى فى المسافة الواقعة بين ٠ ، ٣٧ لتنبيه الناظر إلى الرسم إلى أن هذه المسافة كان يجب أن تكون أطول مما هى عليه فى الرسم وإنما قصرت حتى لا يترك فراغ كبير لا لزوم له فى الرسم .

خامساً – تغيير مقياس الرسم على نفس المحور :

قد نلجأ أحياناً إلى تغيير وحدات القياس على المحور الأفقى ، كما لو أردنا مثلا رصد سلسلة زمنية تمثل أرقاماً سنوية فى جزء منها وأرقاماً شهرية فى الجزء الآخر . فإذا اخترنا وحدات القياس فى الجزء الأول من السلسلة الزمنية بحيث كان كل ١ سنتيمتر يمثل سنة فإننا نضطر ، إذا أردنا الاحتفاظ بنفس مقياس الرسم فى الجزء الثانى من السلسلة ، إلى تمثيل الشهر بـ $\frac{1}{12}$ من السنتيمتر الأمر الذى يفقد الرسم وضوحه ويقلل فائدته . ومن الجهة الأخرى ، إذا لجأنا إلى تكبير المسافة التى تمثل السنة على المحور الأفقى كثيراً فجعلناها ١٢ سنتيمتر مثلا كى يمكن تمثيل الشهر بـ ١ سنتيمتر فإن الرسم لا بد أن يشغل فراغاً ضخماً يتعذر إيجاده فى أغلب الحالات . ولتخلص من هذه الصعوبة ، نستخدم وحدات قياس مناسبة لتمثيل الجزء الأول من السلسلة الزمنية الذى يمثل أرقاماً سنوية وبعد الانتهاء من الرسم نقيم محوراً إضافياً عند آخر رقم سنوى بحيث يكون موازياً للمحور الرأسى الأصلى . ثم نبدأ بعد ذلك باختيار وحدات قياس جديدة على امتداد المحور الأفقى تمثل الجزء الثانى من السلسلة الزمنية الخاص بالأرقام الشهرية فيصبح الرسم فى النهاية

مكوناً من رسمين متلاصقين (ويحسن أن تفصلهما مسافة صغيرة جداً) الأول يمثل الأرقام السنوية والثانى يمثل الأرقام الشهرية على النحو المبين بالشكل رقم ٩ ص ٥٣ . ومن الواضح أننا نلجأ إلى عرض البيانات بهذه الكيفية إذا أردنا إعطاء الناظر إلى الرسم فكرة إجمالية عن تغير الظاهرة فى الماضى ثم فكرة تفصيلية عن طريقة تغيرها خلال السنة الأخيرة .

سادساً – العنوان ومذكرات المصادر .

كل رسم بيانى (ككل جدول إحصائى) يجب أن يحمل عنواناً يبين فى وضوح وإيجاز الغرض من الرسم . وقد يكتب العنوان فوق الرسم أو فى أسفله ولكنه يكتب عادة فى أسفل الرسم . وكما سبق أن بينا فى حالة الجداول الإحصائية لا بد من إيضاح المصادر التى استقيت منها البيانات المعروضة فى الرسم خصوصاً إذا استخدم الرسم البيانى فى إثبات وجهة نظر خاصة .

أوجزنا فيما تقدم القواعد العامة الواجب اتباعها فى عمل الرسوم البيانية الإحصائية ونورد فيما يلى أمثلة لبعض أنواع الرسوم البيانية الشائعة الاستعمال كى نعطى القارىء فكرة عملية عن طريقة عرض البيانات الإحصائية بالرسم البيانى .

١ – الرسوم البيانية التى تظهر مدى تقلبات الظاهرة :

لتوضيح هذا النوع من الرسوم البيانية نعرض البيانات الواردة فى الجدول رقم ٥ وهى عبارة عن السعر الأعلى والسعر الأدنى والسعر المتوسط لأسهم البنك العقارى المصرى خلال خمسة عشر عاماً .

جدول رقم ٥

البنك العقارى المصرى

أسعار الأسهم فى بورصة الأوراق المالية

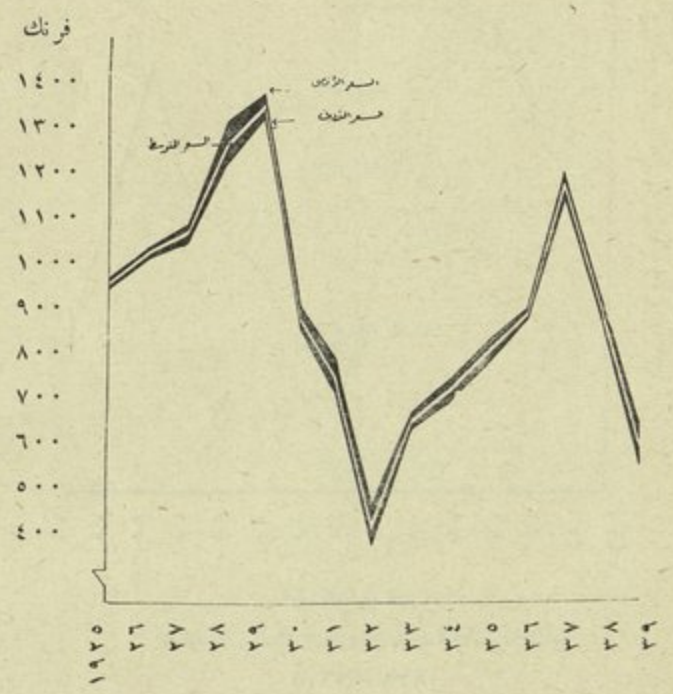
١٩٢٥ — ١٩٣٩

السنة	أسعار الأسهم فى بورصة الأوراق المالية		
	الحد الأقصى بالفرنكات	الحد الأدنى بالفرنكات	السعر المتوسط بالفرنكات
١٩٢٥	٩٣٠	٩٠٧	٩١٨,٥
١٩٢٦	٩٩٠	٩٦٣	٩٧٦,٥
١٩٢٧	١٠٢٩	٩٩٦	١٠١٢,٥
١٩٢٨	١٢٣١	١١٤٧,٥	١١٨٩,٢٥
١٩٢٩	١٢٨٥	١٢٤٠	١٢٦٧,٥
١٩٣٠	٨٦٩	٨٢٧	٨٤٨
١٩٣١	٧٧٠	٦٩٨	٧٢٤
١٩٣٢	٤٨٤	٤٠١	٤٤٢,٥
١٩٣٣	٦٦٢	٦٣٣	٦٤٧ ,٥
١٩٣٤	٧٢٩	٦٨٦	٧٠٧,٥
١٩٣٥	٨٠٣	٧٥٨	٧٨٠,٥
١٩٣٦	٨٦١	٨٤٦	٨٥٣,٥
١٩٣٧	١١٣٠	١٠٩٧	١١١٣,٥
١٩٣٨	٨٧٦	٨٤١	٨٥٨,٥
١٩٣٩	٦٣٥	٥٦٠	٥٩٧,٥

المصدر : إحصاء الشركات المساهمة التى يوجد عملها الرئيسى بمصر

لتصوير البيانات الموجودة بالجدول السابق بطريقة تظهر بوضوح مدى التقلبات التى طرأت على أسعار الأسهم خلال كل سنة ، نرسم الخطوط البيانية الثلاثة التى تمثل الحد الأقصى والحد الأدنى والسعر المتوسط . ويرسم خط السعر

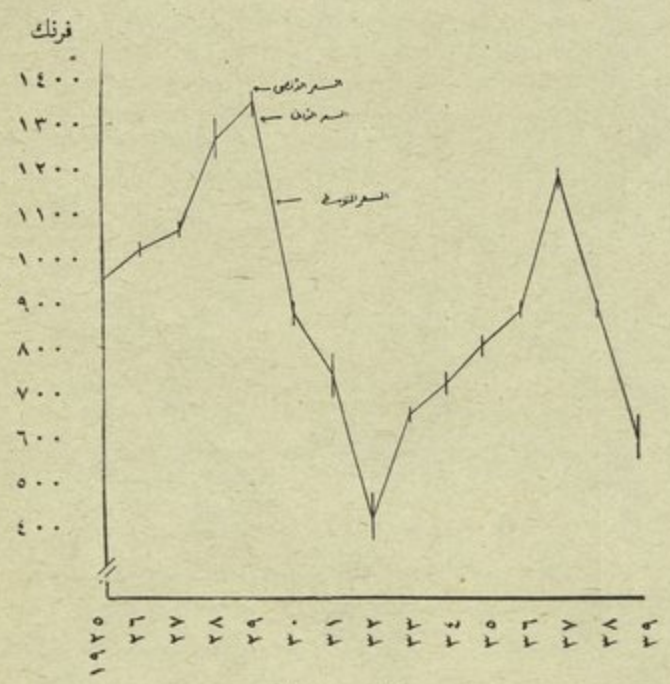
المتوسط بلون مخالف للون الخطين الاخرين أو يجعل أثقل منهما نوعاً كى يظهر بوضوح فى الرسم ، ثم تظلل المساحة المحصورة بين خطى الحد الأقصى والأدنى بحيث لا يقلل هذا التظليل من وضوح خط السعر المتوسط . والرسم الاتى يوضح مدى تقلبات أسعار أسهم البنك العقارى المبينة بالجدول رقم ٥ :



(شكل رقم ١٢)
أسعار أسهم البنك العقارى فى بورصة الأوراق المالية
١٩٢٥ – ١٩٣٩

وربما كان من الأفضل أن نكتفى برسم خط السعر المتوسط ثم نرصد النقط التى تمثل الأسعار العليا والدنيا ، وبدلا من أن نرسم خطاً بيانياً يمر بجميع النقط التى تمثل الأسعار العليا وآخر يمر بجميع النقط التى تمثل الأسعار الدنيا ، نرسم خطوطاً عمودية على المحور الأفقى يصل كل منها بين

زوج من القيم المتناظرة (الحد الأقصى والحد الأدنى لكل سنة) ماراً بخط السعر المتوسط فيكون طول كل عمود معبراً عن مدى تقلب السعر فى السنة المقابلة . والرسم الآتى يوضح هذه الطريقة : —

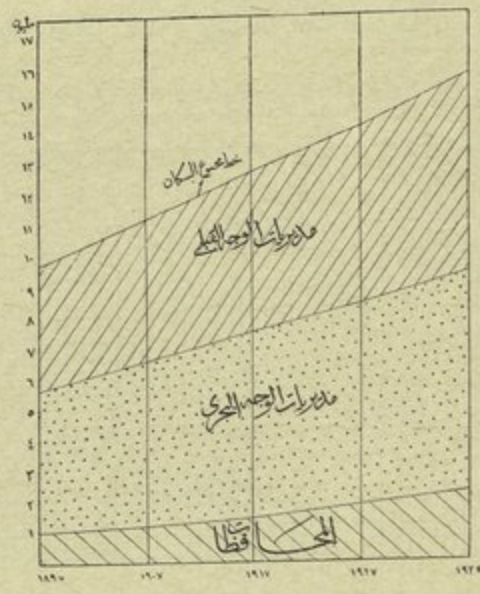


(شكل رقم ١٣)
أسعار أسهم البنك العقارى فى بورصة الأوراق المالية
١٩٢٥ — ١٩٣٩

٢ — الرسوم البيانية التى توضح تقسيم مجموع معين إلى فئات :

هذا النوع شائع الاستعمال جداً ويدل دلالة واضحة على أفضلية الرسوم البيانية على الجداول فى العرض الإحصائى . إفرض مثلا أننا أردنا أن نكون فكرة عامة عن توزيع سكان القطر المصرى على مديريات الوجه البحرى والوجه القبلى والمحافظات خلال الخمسين سنة الماضية ، يمكن أن نكون هذه الفكرة من دراسة الجداول الإحصائية التى تحتوى على البيانات الخاصة بذلك .

غير أن هذه الدراسة تتطلب دون شك مجهوداً ووقتاً أكبر بكثير مما يتطلبه مجرد النظر إلى رسم بيانى يعطينا نفس الفكرة فى سرعة وسهولة . (انظر شكل رقم ١٤) .



(شكل رقم ١٤)

رسم بيانى يوضح توزيع سكان القطر المصرى
على مديريات الوجه البحرى والوجه القبلى والمحافظات
١٨٩٧ – ١٩٣٧

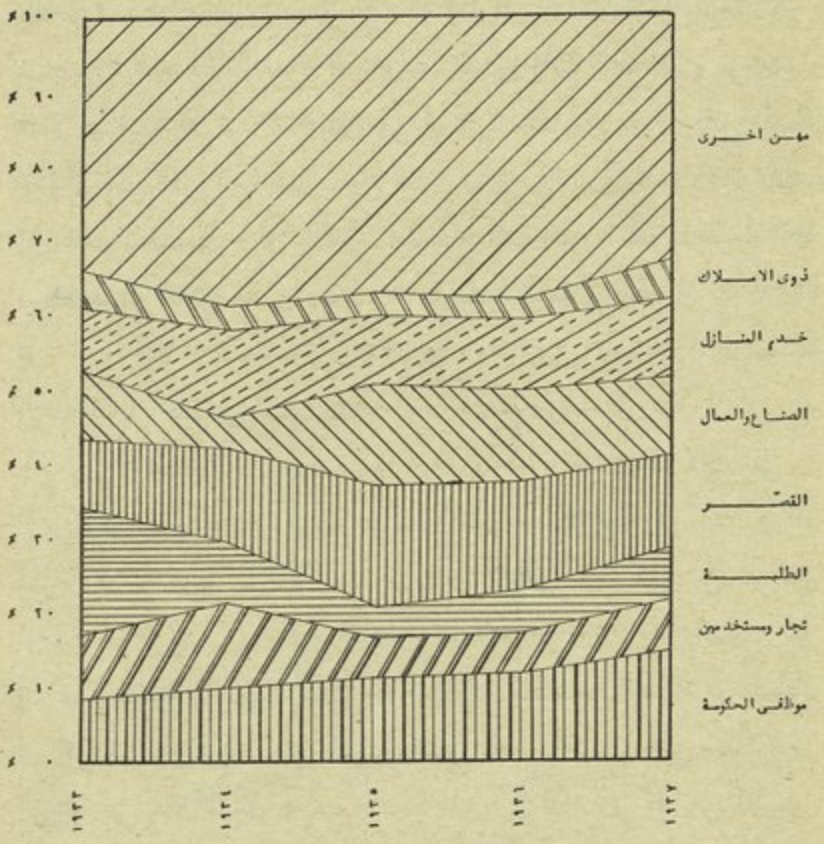
يلاحظ فى الرسم السابق أن المحور الرأسى لكل سنة مقسم إلى أقسام تتناسب مع عدد السكان فى كل فئة . ويمكن بمجرد النظر إلى الرسم أن نعرف ما إذا كان المجموع الكلى للسكان وعدد سكان كل فئة يتزايد أو يتناقص على مر السنين . فكلما اتسعت المساحة المظللة التى تمثل فئة معينة دل ذلك على زيادة عدد سكان هذه الفئة . وعلى الرغم من أن الرسم السابق يوضح ما إذا كانت أى فئة تزيد أو تنقص بطريقة مطلقة إلا أنه لا يوضح ما إذا كانت

نسبة الفئة إلى مجموع السكان تزيد أو تنقص . أو بعبارة أخرى فإنه لا يوضح تغير الأهمية النسبية لكل فئة . فإذا أردنا توضيح تغير الأهمية النسبية لكل فئة بيانياً ، نحسب عدد سكان كل فئة كنسبة مئوية من المجموع الكلى للسكان ثم نقسم المحور الرأسى لكل سنة طبقاً لهذه النسب . والرسمان الآتيان (شكل ١٥ وشكل ١٦) يوضحان هذه الطريقة .



(شكل رقم ١٥)
رسم بيانى يوضح تغير الأهمية النسبية لسكان المديريات والمحافظات
١٨٩٧ – ١٩٣٧

وهناك أنواع أخرى كثيرة من الرسوم البيانية ولكن مجالنا هذا لا يتسع للإشارة إليها ، ويستطيع الطالب الذى يرغب فى الاستزاده منها أن يطلع على أحد الكتب الخاصة بموضوع الرسم البيانى .



(شكل رقم ١٦)

رسم بيانى يوضح تقسيم ودائع صناديق التوفير فى القطر المصرى
حسب مهنة المودع (١٩٣٣ – ١٩٣٧)

الرسوم البيانية اللوغاريتمية

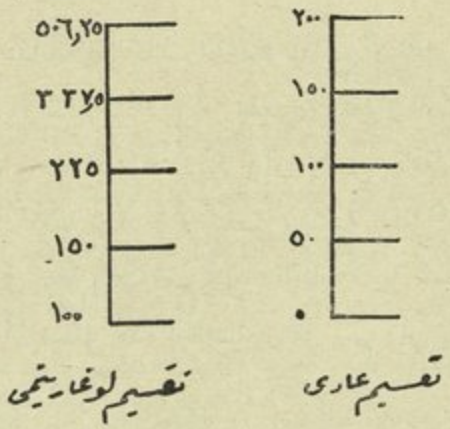
فى كثير من الحالات التى نراقب فيها تغير قيم الظواهر المختلفة يكون اهتمامنا منحصراً لا فى التغير المطلق فى قيمة الظاهرة (أى التغير مقدراً بدلالة الوحدات المطلقة) بل فى التغير النسبى (أى التغير منسوباً إلى أساس معين) فى هذه القيمة . افرض مثلا أننا نرغب فى مقارنة التغير فى عدد سكان مدينة القاهرة

بالتغير فى عدد سكان مدينة الإسكندرية لنعلم أى المدينتين تنمو بسرعة أكبر . نستطيع بمراجعة أرقام تعداد المدينتين فى السنين المختلفة أن نقف على مقدار التغير فى قيمة كل من الظاهرتين ومن ثم نقارن مقدار الزيادة فى كل منهما . غير أنه لكى تكون المقارنة مبنية على أساس سليم ، لا بد أن تنسب هذه الزيادة إلى عدد السكان فى بادئ الأمر . ولتوضيح أهمية هذه النقطة افرض أن عدد سكان مدينة القاهرة بلغ فى سنة ١٩٢٧ ١ مليون نسمة ثم ارتفع فى سنة ١٩٣٧ إلى ١,٢٥٠,٠٠٠ نسمة أى بزيادة قدرها ٢٥٠,٠٠٠ ، وأن عدد سكان مدينة الإسكندرية كان فى سنة ١٩٢٧ ٥٠٠,٠٠٠ وارتفع فى سنة ١٩٣٧ إلى ٧٠٠,٠٠٠ نسمة أى بزيادة قدرها ٢٠٠,٠٠٠ ، فهل يدل ذلك على أن الزيادة فى عدد سكان مدينة القاهرة أكبر من الزيادة فى عدد سكان مدينة الإسكندرية ؟ او قارنا الزيادة المطلقة فى كل من المدينتين لوجدنا أن الزيادة فى مدينة القاهرة أكبر بمقدار ٥٠,٠٠٠ وحدة . غير أن هذه المقارنة لا تفيدنا كثيراً إذا أردنا أن نقف على حقيقة نمو السكان فى كل من المدينتين . أى المدينتين تنمو بمعدل أكبر ؟ لكى نتمكن من الإجابة على هذا السؤال نحسب النسبة المئوية للزيادة فى كل من المدينتين فنجد أنها فى مدينة القاهرة (بناء على الأرقام الفرضية السابقة) ٢٥٪ فى حين أنها فى مدينة الإسكندرية ٤٠٪ . ومن هذا يتضح أنه على الرغم من أن الزيادة المطلقة فى عدد سكان مدينة القاهرة أكبر إلا أن الزيادة النسبية فى سكان القاهرة أقل منها فى الإسكندرية . هناك إذن فرق كبير بين التغير المطلق والتغير النسبى .

وكما أن الرسم البيانى العادى يوضح لنا مقدار التغير المطلق فى قيم الظواهر فإن الرسم البيانى اللوغاريتمى يوضح لنا مقدار التغير النسبى فى هذه القيم .

والفكرة الأساسية فى الرسم البيانى اللوغاريتمى هى أن يقسم المحوران بطريقة تجعل المسافات المتساوية على المحور تمثل نسباً متساوية وليس كميات

متساوية كما هو الحال فى الرسم العادى . فى الرسم البيانى العادى يبدأ القياس من الصفر عند نقطة الأصل ثم تتصاعد الأرقام الموضوعة أمام التقسيمات المتساوية على المحور على شكل متوالية عددية ١٠٠ ، ١٥٠ ، ٢٠٠ ، ٢٥٠ ، ٣٠٠ أما فى الرسم البيانى اللوغاريتمى فيبدأ القياس من أى رقم آخر خلاف الصفر[١] ثم تتصاعد الأرقام الموضوعة أمام التقسيمات المتساوية على المحور على شكل متوالية هندسية ١٠٠ ، ١٥٠ ، ٢٢٥ ، ٣٣٧,٥ ، ٥٠٦,٢٥ كما يتضح من مقارنة الرسمين الآتيين (شكل رقم ١٧) .



(شكل رقم ١٧)
التقسيم العادى والتقسيم اللوغاريتمى

يلاحظ فى التقسيم العادى أن المسافة بين ١٠٠ ، ١٥٠ (ولتكن ١ سنتيمتر مثلا) تعادل المسافة بين ١٥٠ ، ٢٠٠ وتعادل كذلك المسافة بين ٢٠٠ ، ٢٥٠ إلخ ، أى أن كل سنتيمتر يمثل ٥٠ وحدة . أما فى التقسيم اللوغاريتمى فإن المسافة بين ١٠٠ ، ١٥٠ (ولتكن ١ سنتيمتر مثلا) تعادل المسافة بين ١٥٠ ،

---

(١) من الواضح أنه لا يمكن بدء القياس اللوغاريتمى بالصفر لأنه قياس نسبى ، ووجود الصفر فى مقام أى نسبة معناه أن هذه النسبة تساوى ما لا نهاية . وبعبارة أخرى لا يمكن قياس التغير النسبى من أساس قدره صفر بل لا بد أن يكون الأساس عدد حقيقى (عدد صحيح عادة) .

٢٢٥ وتعادل كذلك المسافة بين ٢٢٥ ، ٣٣٧,٥ وهكذا أى أن كل سنتيمتر يمثل نسبة قدرها ٥٠٪ .

على أنه لو نظرنا إلى لوغاريتمات المتسلسلة الثانية (المتوالية الهندسية) نجد أنها تتخذ شكل متوالية عددية . فباستخراج هذه اللوغاريتمات من الجداول نجدها كالآتى :—

العدد	:	١٠٠	١٥٠	٢٢٥	٣٣٧,٥	٥٠٦,٢٥ ٥
اللوغاريتم	:	٢	٢,١٧٦	٢,٣٥٢	٢,٥٢٨	٢,٧٠٤ ٥

نستنتج من هذا أنه إذا كانت الأعداد تزداد (أو تنقص) بنسبة ثابتة فإن لوغاريتمات هذه الأعداد تزداد (أو تنقص) بمقدار ثابت . أو بالعكس إذا كانت لوغاريتمات الأعداد تزداد (أو تنقص) بمقدار ثابت نعلم فى الحال أن الأعداد ذاتها تزداد (أو تنقص) بنسبة ثابتة(١). أو بعبارة أخرى فإن التغير المطلق فى لوغاريتمات القيم يمثل بدقة تامة التغير النسبى فى القيم ذاتها .

يمكننا الآن أن نستغل هذه الخاصية فى عمل رسم بيانى يوضح التغير النسبى

---

(٥) لاحظ أن نسبة الزيادة ثابتة وقدرها ٥٠٪

(٥) لاحظ أن مقدار الزيادة ثابت وقدره ٠,١٧٦

(١) من المبادئ المعروفة فى الجبر أن لوغاريتمات المتوالية الهندسية تكون على شكل متوالية عددية ويمكن إثبات ذلك كالآتى :

إفرض أن $س_1$ ، $س_2$ ، $س_3$ فى متوالية هندسية .

$$\therefore \frac{س_2}{س_1} = \frac{س_3}{س_2}$$

$$\therefore \text{لو} \frac{س_2}{س_1} = \text{لو} \frac{س_3}{س_2}$$

$$\therefore \text{لو } س_2 - \text{لو } س_1 = \text{لو } س_3 - \text{لو } س_2$$

$\therefore$ لو $س_1$ ، لو $س_2$ ، لو $س_3$ فى متوالية عددية .

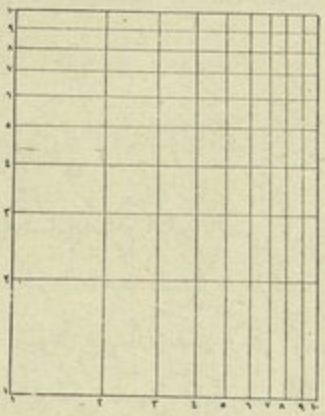
فى قيم الظواهر ، فبدلا من أن نرصد قيم الظاهرة ذاتها فيوضح لنا الرسم تغيرها المطلق ، نستعيض عن هذه القيم بلوغاريتماتها فيوضح لنا الرسم تغيرها النسبى . فإذا كانت هذه اللوغاريتمات تزداد بمقدار ثابت ، دل ذلك على أن الظاهرة نفسها تزيد بنسبة ثابتة . وإذا كانت اللوغاريتمات تزيد بمقادير متزايدة ، دل ذلك على أن الظاهرة نفسها تزداد بنسب متزايدة . وإذا كانت اللوغاريتمات تزيد بمقادير متناقصة ، دل ذلك على أن الظاهرة تزداد بنسب متناقصة .

طريقة عمل التقسيم اللوغاريتمى .

يظهر مما سبق أنه يمكننا إيضاح التغير النسبى فى قيمة الظاهرة على القياس العادى بشرط ألا نرصد قيم الظاهرة نفسها بل نرصد لوغاريتمات هذه القيم . على أن الطريقة الأكثر شيوعاً هى أن نقسم ورق الرسم تقسيماً لوغاريتمياً قبل البدء برصد النقط عليه وذلك لكى تسهل عملية رصد النقط على الرسم من جهة ، ولكى يعلم من يطلع عليه بمجرد النظر أنه رسم بيانى ذو مقياس نسبى (لوغاريتمى) من الجهة الأخرى . والواقع أن عمل هذا التقسيم للرسم البيانى اللوغاريتمى هو بمثابة استخدام ورق المربعات فى الرسم البيانى العادى . وتتلخص طريقة التقسيم فيما يأتى :—

نرسم المحورين كالمعتاد ونضع أمام نقطة الأصل ، أي فى مكان الصفر فى القياس العادي الرقم ١ (لاحظ أن لو ١ = ٠) ، ثم نأخذ على المحور بعداً يساوى لو ٢ ونكتب أمام هذا البعد الرقم ٢ . ومن المعروف أن لو ٢ = ٠,٣٠١ غير أن هذا لا يقيدنا بجعل لو ٢ على بعد ٠,٣٠١ سنتيمتر من من نقطة الأصل بل الواقع أنه لو فعلنا ذلك لأصبح الرسم من الصغر بحيث يتعذر رصد النقط عليه . الأفضل طبعاً أن نكبر المسافة بقدر الإمكان فنجعل لو ٢ على بعد ٠,٣٠١ بوصة أو ٠,٣٠١ ديسمتر (أى ٣,٠١

سنتيمتر) مثلا ، فلنا مطلق الحرية فى اختيار الوحدات المناسبة فى حدود الحيز الموجود لدينا على ورقة الرسم . بعد ذلك نأخذ من نقطة الأصل بعداً على المحور يساوى لو ٣ مستعملين نفس الوحدات السابقة ، ونكتب أمامة الرقم ٣ ، ثم نعود فنأخذ بعداً من نقطة الأصل يساوى لو ٤ ونكتب أمامه الرقم ٤ وهكذا حتى الرقم ١٠ فيكون تقسيم المحورين كالآتى :-



(شكل رقم ١٨)

تقسيم المحورين تقسيما لوغاريتميا

بعد أن ننتهى من عملية التقسيم على النحو السابق يمكننا أن نضع أمام التقسيمات أى سلسلة من الأرقام بحيث تكون متناسبة مع الأعداد الطبيعية من ١ إلى ١٠ فنستطيع مثلا أن نبدأ عند نقطة الأصل بالرقم ١٠٠ بدلا من ١ بحيث نضع ٢٠٠ فى مكان ٢ و ٣٠٠ فى مكان ٣ ، .... ، و ١٠٠٠ فى مكان ١٠ .

كذلك نستطيع أن نجعل أرقامنا تتسلسل على النحو الآتى :-

٥٠ ، ١٠٠ ، ١٥٠ ، ٢٠٠ ، .... ، ٥٠٠

أو ١٧ ، ٣٤ ، ٥١ ، ٦٨ ، .... ، ١٧٠

وهكذا يمكننا أن نضع أمام التقسيمات اللوغاريتمية أى سلسلة أخرى متناسبة مع الأعداد الطبيعية . ومن البديهى أننا نختار فى كل حالة سلسلة الأرقام التى تناسب قيم الظاهرة المطلوب عرض تغيرها النسبى بيانيا . فإذا كانت قيم الظاهرة تتغير فى مدى محصور بين ٦٠ و ٤٩٠ مثلا فإننا نختار سلسلة الأرقام من ٥٠ إلى ٥٠٠ . وإذا كان تغيرها ينحصر بين ١٢٠٠ و ٩٨٠٠ فإننا نختار سلسلة الأرقام من ١٠٠٠ إلى ١٠,٠٠٠ وهكذا ...

والمهم أن نلاحظ أن تغيير طريقة الترقيم لا يؤثر مطلقاً على طريقة التقسيم فالتقسيمات اللوغاريتمية على المحور تكون دائماً متناسبة مع لوغاريتمات الأعداد الطبيعية من ١ إلى ١٠ مهما كانت قيم الظاهرة المطلوب رصدها على الرسم . ويتضح السبب فى ذلك مما يأتى :—

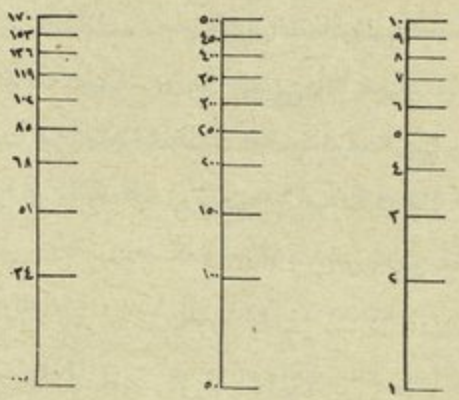
$$\text{لو } ١٠٠ - \text{لو } ٥٠ = \text{لو } \frac{١٠٠}{٥٠} = \text{لو } \frac{٢}{١} = \text{لو } ٢ - \text{لو } ١$$

$$\text{كذلك لو } ٣٤ - \text{لو } ١٧ = \text{لو } \frac{٣٤}{١٧} = \text{لو } \frac{٢}{١} = \text{لو } ٢ - \text{لو } ١$$

فالمسافة بين لو ١٠٠ و لو ٥٠ هى نفسها المسافة بين لو ٣٤ و لو ١٧ وكلاهما يساوى المسافة بين لو ٢ و لو ١

وبالمثل فإن المسافة بين لو ١٥٠ و لو ١٠٠ تساوى المسافة بين لو ٥١ و لو ٣٤ وتساوى كذلك المسافة بين لو ٣ و لو ٢ .

والخلاصة أنه بعد عمل التقسيم حسب لوغاريتمات الأعداد الطبيعية يمكننا وضع أى سلسلة من الأرقام أمام التقسيمات التى حددناها على المحور ما دامت متناسبة مع هذه الأعداد كما يتضح من الرسم الآتى :—



(شكل رقم ١٩)

طرق ترقيم المحور بعد عمل التقسيم اللوغاريتمى

تحديد موقع النقط بالنسبة للمحور اللوغاريتمى :

إذا كان ترقيم المحور اللوغاريتمى يبدأ من أى رقم خلاف الواحد الصحيح ، كما لو بدأنا القياس بالرقم ٥٠ مثلا ، نقسم الأرقام التى تمثل قيم الظاهرة على الرقم الذى يبدأ به القياس ثم نستخرج لوغاريتمات خوارج القسمة ونرصد هذه اللوغاريتمات على الرسم . وكأننا بعملية القسمة هذه قد رددنا هذه الأرقام إلى مكانها بين سلسلة الأعداد الطبيعية التى عملنا القياس على أساسها . فإذا أردنا مثلا رصد الرقم ٥٥ على قياس لوغاريتمى يبدأ بالرقم ٥٠ فإننا نرصد لوغاريتم $\frac{٥٥}{٥٠}$ أى ١,١ . وذلك لأنه ما دام الرقم ٥٠ يحتل مكان الرقم ١ فى سلسلة الأعداد الطبيعية والرقم ١٠٠ يحتل مكان الرقم ٢ ، فإن الرقم ٥٥ يحتل مكان الرقم ١,١ والرقم ٦٠ يحتل مكان الرقم ١,٢ وهكذا .

الدورات اللوغاريتمية :

من الواضح أنه لو بدأنا القياس بالرقم ٥٠ فإنه ينتهى بالرقم ٥٠٠ ، فما العمل إذا كانت قيم الظاهرة تتغير فى مدى واسع يمتد من ٥٠ إلى ٢٠٠٠ مثلا؟ فى مثل

هذه الحالات لا يكفى عمل قياس لوغاريتمى ذى دورة واحدة[1] (أى يمتد من ١ إلى ١٠ أو ٥٠ إلى ٥٠٠) بل نعمل قياساً مكوناً من دورتين أو أكثر بحيث يمكن رصد جميع قيم الظاهرة عليه . فإذا كانت الدورة الأولى تبدأ بالرقم ٥٠ فإن الدورة الثانية تبدأ بالرقم ٥٠٠ وتنتهى بالرقم ٥٠٠٠ بينما تبدأ الثالثة بالرقم ٥٠٠٠ وتنتهى بالرقم ٥٠٠٠٠ وهكذا . ويلاحظ أن تقسيم الدورات اللوغاريتمية الثانية والثالثة والرابعة .... إلخ يكون مماثلا تماماً لتقسيم الدورة الأولى . فالتقسيم يكون دائماً حسب لوغاريتمات الأعداد الطبيعية من ١ إلى ١٠ مهما كانت طريقة الترقيم ، ومهما كانت مرتبة الدورة . والشكل رقم ٢١ يوضح قياساً لوغاريتمياً مكوناً من عدة دورات .

التطبيقات العملية للرسوم البيانية اللوغاريتمية

للرسوم البيانية اللوغاريتمية تطبيقات عديدة فى ميدان الدراسات الإقتصادية ومن أهمها ما يأتى :—

أولا : توضيح العلاقة بين التغير النسبى فى قيم ظاهرتين بيانياً .

كثيراً ما يهمنا فى الدراسات الاقتصادية الوقوف على العلاقة بين التغير النسبى فى ظاهرة والتغير النسبى فى ظاهرة أخرى . ومن الأمثلة الواضحة على ذلك العلاقة بين التغير النسبى فى ثمن السلعة والتغير النسبى فى طلب المستهلكين لهذه السلعة . فمن المعروف أن مرونة الطلب $= \frac{\text{التغير النسبى فى الطلب}}{\text{التغير النسبى فى الثمن}}$

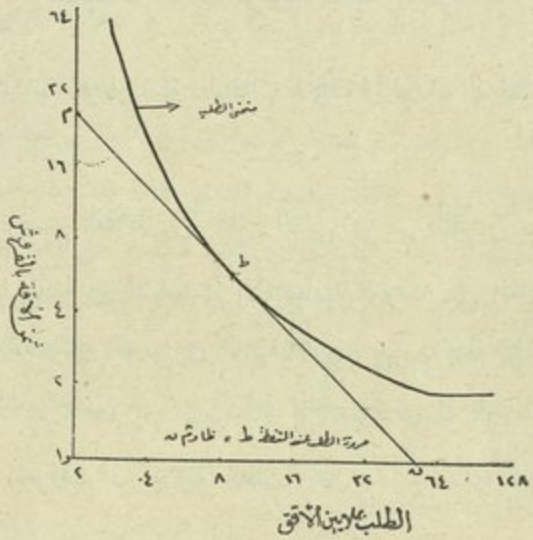
---

(١) المسافة بين لو ١ و لو ١٠ على الرسم تعتبر دورة لوغاريتمية كاملة . فاذا عملنا قياساً لوغاريتميا يمتد من لو ١ إلى لو ١٠٠ فإنه يكون مكوناً من دورتين الأولى تمتد من ١ إلى ١٠ والثانية تمتد من ١٠ إلى ١٠٠ وإذا كان القياس يمتد من ١ إلى ١٠٠٠ فإنه يكون مكوناً من ثلاث دورات وهكذا .

ويمكن تقدير قيمة مرونة الطلب بدقة بيانياً إذا ما رصدنا نقط الطلب على قياس لوغاريتمى مزدوج (حيث يكون كلا المحورين مقسما تقسيما لوغاريتمياً) . وتتلخص طريقة تقدير مرونة الطلب من الرسم البيانى اللوغاريتمى فيما يأتى :—

نرسم منحنى الطلب على ورقة مقسمة تقسيما لوغاريتمياً ، ثم نرسم مماساً لمنحنى الطلب عند النقطة المقابلة للثمن المراد تقدير المرونة عنده ، فتكون مرونة الطلب هى عبارة عن ميل المماس على محور الثمن أى ظل الزاوية التى يصنعها هذا المماس مع محور الثمن ، وذلك لأن ظل هذه الزاوية = $\frac{\text{التغير النسبى فى الطلب}}{\text{التغير النسبى فى الثمن}}$

وتتضح طريقة تقدير مرونة الطلب بيانياً من الرسم الآتى :—



(شكل رقم ٢٠)

تقدير مرونة الطلب بيانياً على قياس لوغاريتمى مزدوج

ومن الأمثلة الأخرى على استعمال الرسم البيانى اللوغاريتمى فى الدراسات

الاقتصادية ، تلك الحالة التى يستعمل فيها هذا الرسم لتوضيح العلاقة بين التغير النسبى فى الدخل والتغير النسبى فى عدد الأشخاص الذين يحصلون على هذا الدخل ، وهى العلاقة المعروفة بقانون باريتو (Pareto Income Law) . وينص هذا القانون على أنه كلما زادت قيمة الدخل بنسبة معينة كلما نقص عدد الأشخاص الذين يحصلون على هذا الدخل أو أكثر بنفس النسبة . وقد أثبتت الإحصاءات أن هذا القانون يصف حقيقة توزيع الدخول فى عدد كبير من البلاد . ولنأخذ كمثال توزيع الدخول فى الولايات المتحدة كما نشره مكتب الأبحاث الاقتصادية بنيويورك ثم نوضحه بياناً على قياس لوغاريتمى مزدوج .

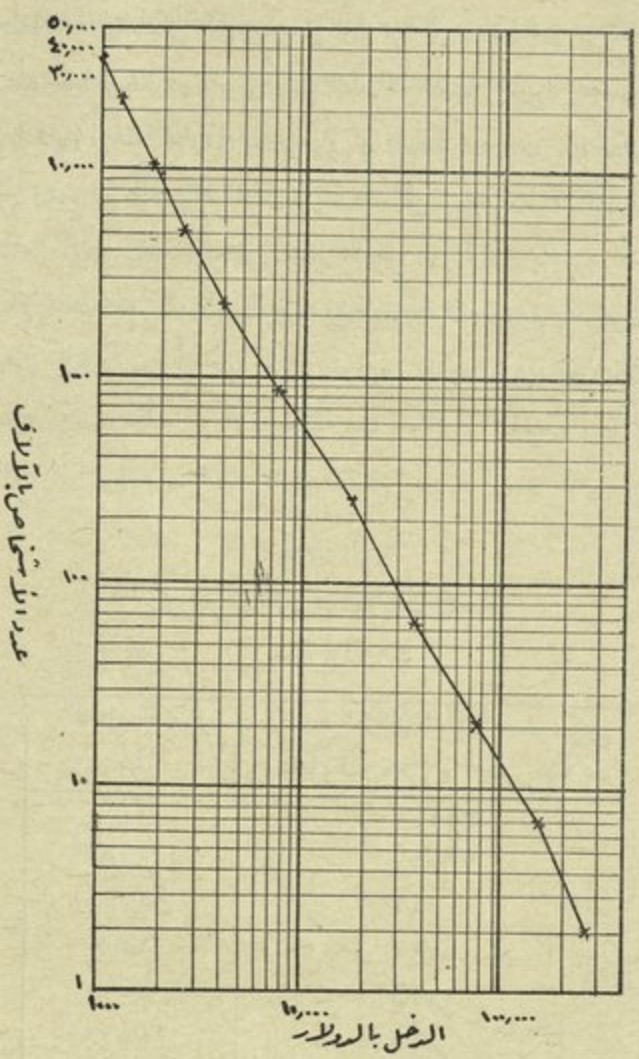
جدول رقم ٦

توزيع الدخول فى الولايات المتحدة

سنة ١٩١٩

الدخل بالدولار س	عدد السكان الذين يحصلون على دخل قدره س أو أكثر بالآلاف
١٠٠٠	٣٥٥٤١
١٢٥٠	٢٣٠١٠
١٧٥٠	١٠٥١٢
٢٥٠٠	٥٢٩٠
٤٠٠٠	٢٢٢٥
٧٥٠٠	٨٤٢
١٧٥٠٠	٢٥٤
٣٧٥٠٠	٦٢
٧٥٠٠٠	٢١
١٥٠٠٠٠	٧
٢٥٠٠٠٠	٢

المصدر : مكتب الأبحاث الاقتصادية بنيويورك .



(شكل رقم ٢١)

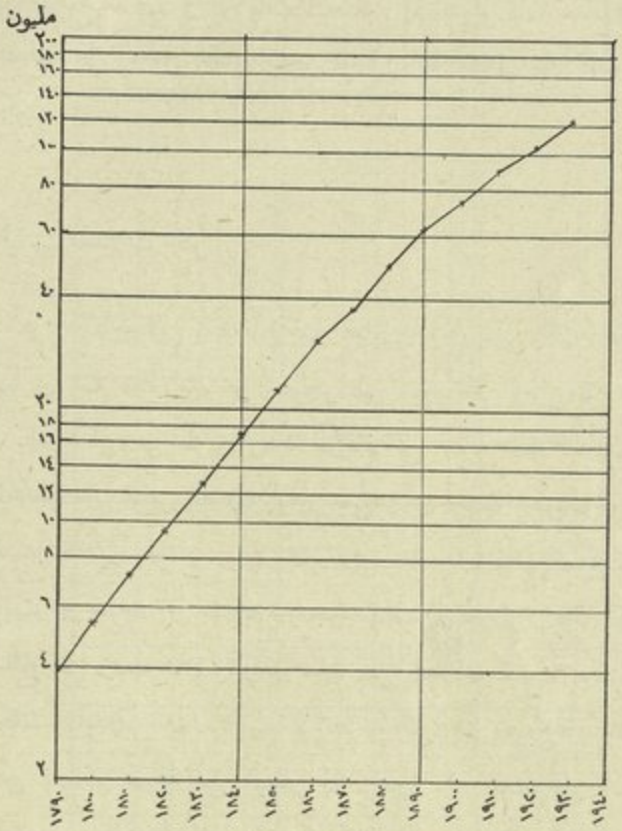
التمثيل البيانى لتوزيع الدخول فى الولايات المتحدة
على قياس لوغاريتمى مزدوج

يتضح من الرسم السابق أن الخط البيانى الذى يمثل العلامة بين التغير النسبى فى الدخل والتغير النسبى فى عدد الأشخاص الحائزين لهذا الدخل أو أكثر ،

قريب جداً من الخط المستقيم . وهذا يثبت أن قانون باريتو ينطبق إلى حد كبير على توزيع الدخول فى الولايات المتحدة . إذ أنه لو كان عدد الأشخاص الذين يحصلون على الدخل ينقص بنفس نسبة زيادة الدخل تماماً فإن هذا الخط البيانى يكون مستقيما .

ثانياً : قياس معدلات النمو .

من أهم الحالات التى يستخدم فيها القياس اللوغاريتمى تلك الحالات التى نرغب فيها فى تتبع معدلات نمو الظواهر التى تتغير مع الزمن مثل ظاهرة نمو السكان . إذ يمكن عن طريق رصد الأرقام التى تمثل عدد السكان فى سنى التعداد المختلفة على قياس نصف لوغاريتمى (أى حيث يكون المحور الرأسى فقط مقسماً تقسيماً لوغاريتمياً) أن نعرف ما إذا كان السكان يتزايدون بمعدل ثابت أو متزايد أو متناقص بمجرد النظر إلى الرسم . فإذا كان الخط البيانى الناتج من رصد هذه الأرقام عبارة عن خط مستقيم ، دل ذلك على أن السكان يتزايدون بمعدل ثابت (انظر شكل ٢٢) وإذا كان عبارة عن منحن يصعد بسرعة من اليسار إلى اليمين دل ذلك على أن السكان يتزايدون بمعدل مطرد الزيادة . أما إذاكان المنحنى الناتج يصعد ببطء ، دل ذلك على أن معدل الزيادة آخذ فى التناقص . والرسم الاتى يوضح نمو سكان الولايات المتحدة من سنة ١٧٩٠ إلى سنة ١٩٣٠ .



(شكل رقم ٢٢)

رسم بيانى نصف لوغاريتمى يوضح نمو سكان الولايات المتحدة

١٧٩٠ – ١٩٣٠

ثالثاً : التمثيل البيانى للظواهر التى تتغير فى مدى واسع :

بعض الظواهر تتغير فى مدى واسع جداً بحيث يكاد يكون من المستحيل عمليا تمثيلها على القياس العادى . افرض مثلا أن إحدى الظواهر تتغير فى مدى

ينحصر بين ١٠٠ و ١٠٠٠,٠٠٠,٠٠٠[1]. مهما اجتهدنا فى تصغير وحدات القياس فلا يمكن تمثيل تغير هذه الظاهرة على ورق الرسم البيانى العادى ، وذلك لأن فراغ ورقة الرسم العادية لا يكفى لهذا الغرض . وإذا استعملنا ورقة من الكبر بحيث يمكن أن تشمل مثل هذه القيم ، فإن الرسم يكاد يفقد معناه . أما باستعمال القياس اللوغاريتمى ، فيمكن توضيح تغير هذه الظواهر بيانيا بكل سهولة وفى حدود فراغ ورقة الرسم العادية . فالأرقام فى القياس اللوغاريتمى تتصاعد بسرعة على شكل متوالية هندسية كما بينا :—

١٠٠ و ١٠٠٠ و ١٠,٠٠٠ و ١٠٠,٠٠٠ و ١٠٠٠,٠٠٠ .... ، إلخ .

وإذا رجعنا إلى الشكل رقم ٢١ نجد أنه أمكن تمثيل ظاهرة تتغير من ١٠٠٠ إلى ٢٥٠٠٠٠ وأخرى تتغير من ٢ إلى ٣٥٥٤١ على محورين لوغاريتمين الأول يحتوى على ثلاث دورات والثانى يحتوى على خمس دورات . ومن الواضح أن تمثيل هاتين الظاهرتين على قياس عادى صعب جدا من الوجهة العملية.

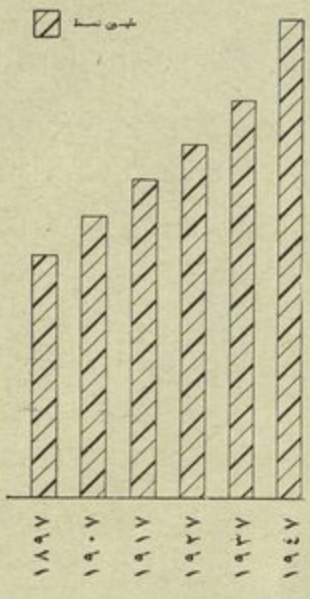
٢ — استخدام الأشكال الهندسية فى عرض البيانات الإحصائية :

إذا كان لدينا بعض الأرقام التى تمثل تغير ظاهرة معينة فى عدد قليل من السنين أو تمثل سلسلة زمنية ليست متصلة ، فالأفضل ألا نستخدم فى عرض هذه الأرقام طريقة الرسم البيانى المعتادة، بل نمثل قيم الظاهرة بأشكال هندسية — مستطيلات أو مربعات أو دوائر — بحيث يكون التغير فى مساحتها متناسباً مع التغير فى قيمة الظاهرة . والطريقة الأكثر شيوعاً هى استخدام المستطيلات الرأسية التى تكون قواعدها متساوية وأطوالها متناسبة مع القيم التى تتخذها الظاهرة فى السنين المختلفة . والواقع أن أطوال هذه المستطيلات تناظر

---

(١) ليست الظواهر التى تتغير فى مثل هذا المدى الواسع خيالية . فإن الرقم القياسى للاسعار فى ألمانيا تغير من ١٠٠ فى سنة ١٩١٤ إلى ١٠٠,٠٠٠,٠٠٠,٠٠٠,٠٠٠ فى سنة ١٩٢٣ وكان التغير فى كمية أوراق البنكنوت المتداولة فى نفس الوقت مماثلا لهذا .

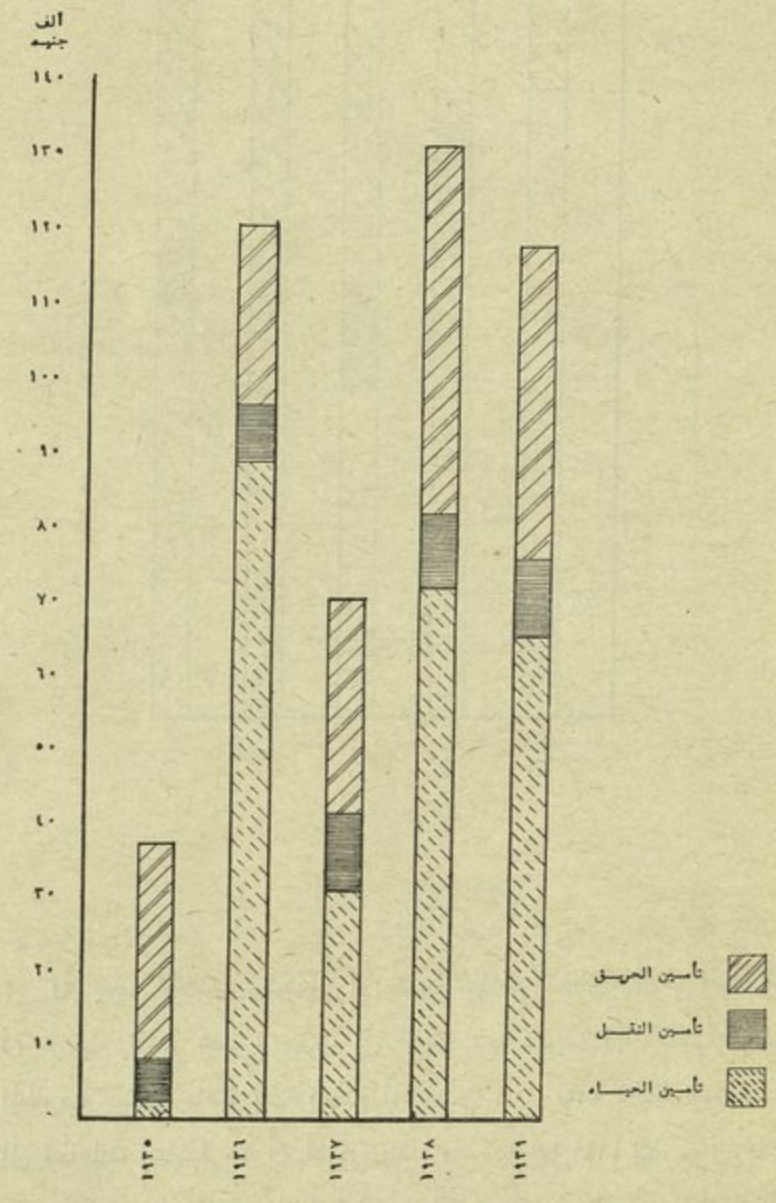
الإحداثيات الرأسية فى الرسم البيانى العادى ، فكلاهما يعبر عن قيم الظاهرة . والرسم الآتى يوضح أبسط الصور التى تتخذها هذه الطريقة .



(شكل رقم ٢٣)
سكان القطر المصرى فى سنى التعداد المختلفة
١٨٩٧ — ١٩٤٧

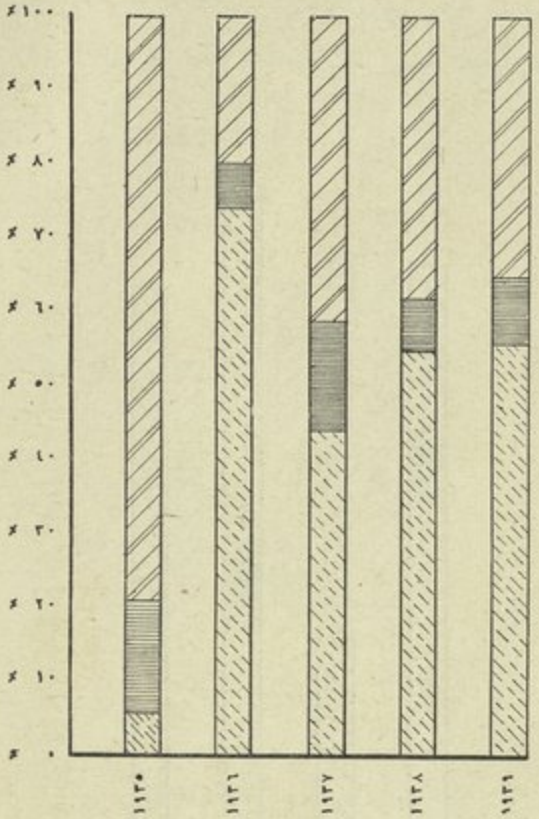
ويمكن كذلك استخدام طريقة المستطيلات الرأسية لإيضاح تقسيم مجموع معين إلى فئات بنفس الطريقة التى اتبعناها فى حالة الرسوم البيانية العادية .

والرسمان الآتيان يوضحان هذه الطريقة — شكل رقم ٢٤ يوضح التغير المطلق فى مجموع الرسوم التى حصلتها شركة مصر للتأمين عن كل نوع من أنواع التأمينات ، وشكل رقم ٢٥ يوضح تغير الأهمية النسبية لكل نوع من أنواع التأمينات من سنة ١٩٣٥ إلى ١٩٣٩ .



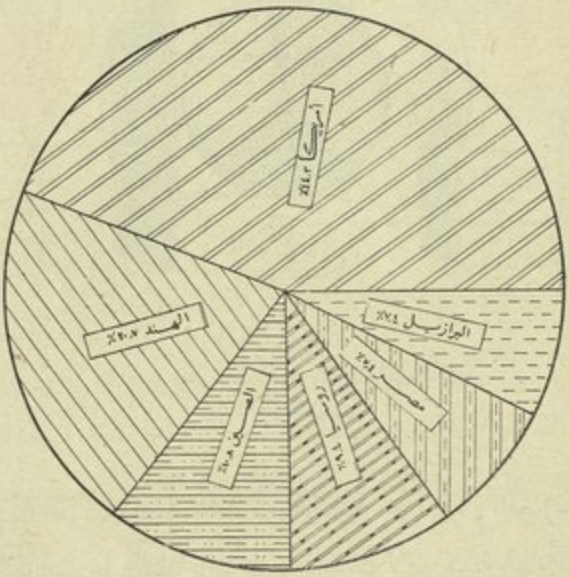
(شكل رقم ٢٤)

رسم يوضح التغير المطلق لكل نوع من أنواع التأمينات



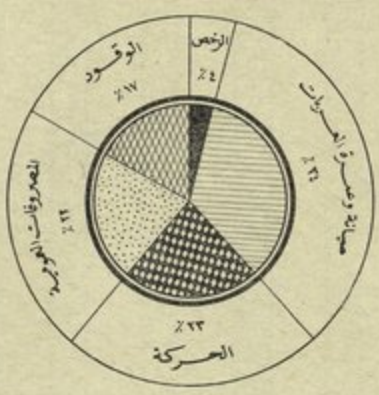
(شكل رقم ٢٥)
رسم يوضح تغير الأهمية النسبية
لكل نوع من أنواع التأمينات

وفى بعض الحالات يكون من الأنسب استخدام الدوائر بدلا من المستطيلات فى توضيح تقسيم مجموع معين إلى فئات ، خصوصاً إذا كانت البيانات المطلوب عرضها خاصة بسنة واحدة أو سنتين . وفى هذه الحالة تقسم الدائرة إلى قطاعات متناسبة مع أهمية كل فئة . ويمكننا عمل هذا التقسيم فى الدائرة عن طريق جعل قيمة الزاوية المركزية لكل قطاع متناسبة مع أهمية الفئة التى يمثلها هذا القطاع . والشكل الآتى يوضح هذه الطريقة .



(شكل رقم ٢٦)
محصول القطن العالمى موزعاً حسب الدول المنتجة
سنة ١٩٣٦

ويلاحظ أحياناً أنه عند عرض البيانات الإحصائية بهذه الطريقة تكون بعض قطاعات الدائرة من الصغر بحيث لا تتسع لوضع أى بيانات رقمية توضيحية تبين قيم الفئات المختلفة عليها . وفى هذه الحالة ، إذا لم نرغب فى تكبير الدائرة المظللة ذاتها ، يمكننا أن نرسم دائرة أخرى كبيرة حول الدائرة الأصلية (من نفس المركز) ثم نمد أنصاف الأقطار التى تفصل القطاعات المختلفة عن بعضها إلى محيط الدائرة الخارجية ونستعمل الفراغ الموجود بين محيطى الدائرتين لكتابة البيانات المطلوبة كما هو موضح بالشكل رقم ٢٧ .



(شكل رقم ٢٧)
تقسيم مصروفات أوتوبوس مدينة الإسكندرية
على الأبواب المختلفة

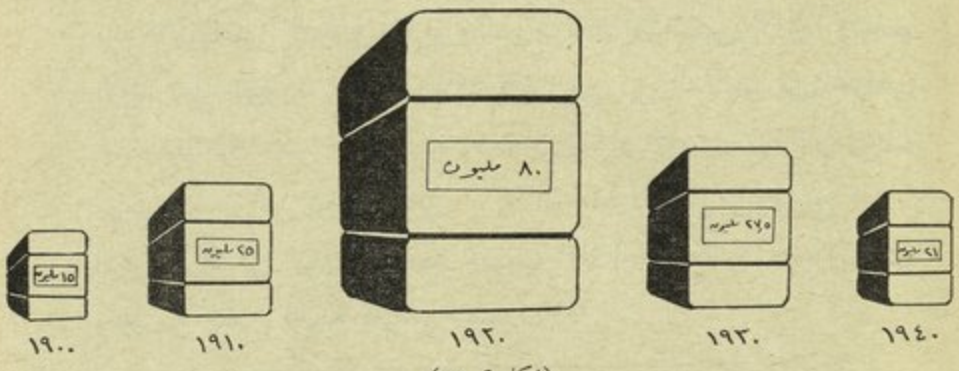
والواقع أن طريقة فصل البيانات الكتابية عن الرسم أفضل من طريقة كتابة هذه البيانات على القطاعات ذاتها . وذلك لأن كتابة البيانات على الرسم ذاته ربما تجعل الشخص الذى يطلع عليه يخطئ فى تقدير مساحة القطاع بالنسبة للمساحة الكلية للدائرة بمجرد النظر إلى الرسم . ويرجع خطأ التقدير هذا إلى ما يحتمل أن يسببه وجود البيانات الكتابية من خداع للبصر .

٣ ــ العرض التصويرى

إذا أردنا وضع البيانات الإحصائية فى صورة شيقة جذابة يمكن للشخص العادى أن يفهمها بسرعة وسهولة ، فيمكننا الاستعاضة عن الأشكال الهندسية بالصور التوضيحية . فمثلا إذا أردنا أن نعرض التغير فى صادرات القطن فى عدة سنوات ، يمكننا أن نعبر عن أرقام الصادرات برسم بالات قطن يكون

(١) هذا الرسم من عمل قسم الإحصاء بإدارة النقل المشترك لمنطقة الإسكندرية وقد تفضل حضرة صاحب العزة المدير العام وأذن لنا بنشره .

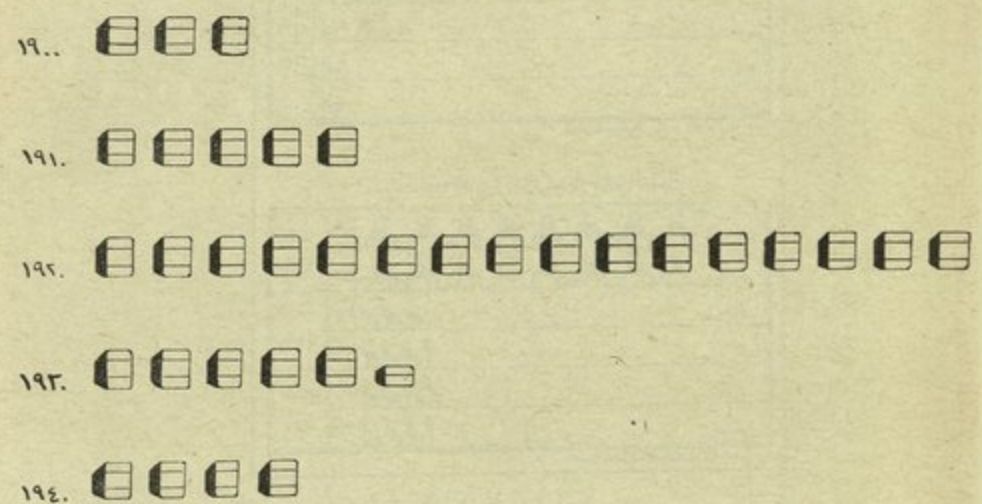
حجمها فى كل سنة متناسباً مع قيمة صادرات القطن فى تلك السنة كما هو موضح بالشكل الآتى :ـ



(شكل رقم ٢٨)

صادرات القطن المصرى بين سنة ١٩٠٠ وسنة ١٩٤٠

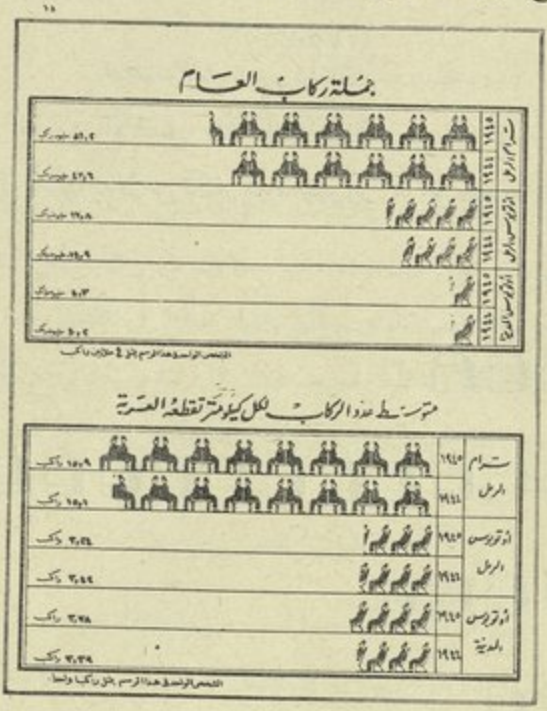
على أنه من الأفضل دائماً أن نعبر عن زيادة قيمة الظاهرة بتكرار الصورة التى تمثلها بدلا من تكبير هذه الصورة كما هو موضح بالشكل الآتى :ـ



(شكل رقم ٢٩)

صادرات القطن المصرى بين سنة ١٩٠٠ وسنة ١٩٤٠

والسبب فى تفضيل هذه الطريقة الأخيرة هو أن تكبير الصورة (أو تصغيرها) وإن كان يعطينا فكرة عامة عن زيادة قيمة الظاهرة (أو نقصها) فإنه لا يساعدنا كثيراً فى معرفة مقدار الزيادة أو النقص إلا إذا وضعت الأرقام على الصورة . وإن كنا قد ذكرنا من قبل أنه ليس من المفروض أن تعطينا طرق العرض البيانى فكرة دقيقة عن مقدار التغير فى قيم الظواهر ، وأن مهمتها الرئيسية هى توضيح التغير ، إلا أن هذا لا يمنع من توضيح مقدار التغير فى الشكل البيانى على وجه التقريب كلما أمكن ذلك . والشكل الآتى يعتبر من النماذج الجيدة للعرض التصويرى[1] .



(شكل رقم ٣٠)

عرض تصويرى لعدد ركاب الترام

والأوتوبوس فى خط الرمل ومدينة الإسكندرية

(١) هذا الرسم من عمل قسم الإحصاء بإدارة النقل المشترك لمنطقة الإسكندرية وقد تفضل حضرة صاحب العزة المدير العام فأذن لنا بنشره .

ومما تجب ملاحظته أن الإحصاءات التصويرية تمثل كميات أو قيما مقربة ولا تمثل تفاصيل . فلا يجوز استعمالها للتعبير عن أرقام يراد عرضها بدقة حسابية . كما أن هذه الطريقة تستعمل لغرض المقارنة بين قيم الظاهرة فى جهات مختلفة أو فى عدة سنوات (أو وحدات زمنية مختلفة) ولا يجوز استعمالها مطلقاً لذكر بيان عن سنة واحدة .

الباب الرابع

التوزيع التكرارى

فى أغلب الحالات التى نقوم فيها بعمل بحث إحصائى ترد إلينا البيانات الأولية فى شكل عدد كبير جداً من الأرقام التى يصعب استعمالها بحالتها الأصلية فى الوصول إلى النتائج التى نهدف إليها بعمل هذا البحث . افرض مثلا أننا أردنا تكوين فكرة عامة عن أعمار الطلبة المتقدمين للالتحاق بكلية التجارة فى العام الدراسى ١٩٤٧ — ١٩٤٨ . بالرجوع إلى شهادات ميلاد الطلبة المذكورين نجد أن أعمارهم موزعة كالآتى :—

سنة	شهر	سنة	شهر	سنة	شهر	سنة	شهر	سنة	شهر	سنة	شهر	سنة	شهر	سنة	شهر
١٩	١	٢١	١١	١٩	١٠	١٩	٢	٢٠	٦	٢١	٩	١٨	٣	١٨	٧	١٨	٢	١١	٣٦
٢٠	٨	١٧	٦	٢٢	١	٢١	١١	١٨	٥	١٩	٣	٢٢	٨	٢٠	٧	٢٠	١	١٧	
١٨	٢	١٨	١	٢٤	٣	١٨	١	٢١	٣	٢٢	١١	٢٠	١١	١٨	١	٢٤	٩	٢٤	
١٧	٤	٢٠	٣	٢٧	٢	٢٠	٣	١٨	٩	١٩	١	١٩	٣	١٨	١	٢١	٨	٢١	٣	٢١
٢٢	٥	٢٠	٣	٢٦	١١	٢٢	١	٢٢	٢	١٧	٧	١٨	٦	١٨	١	٢٠	١	٢٠	٧	٢٣
٢٤	١١	٢٢	١	٢٦	٧	٢٦	١	٢٢	٦	١٨	١١	٢١	١	٢٠	١١	٢٢	٦	٢٢	٩	٢٧
٢٣	٢	١٦	١٠	٣٠	٣	٢٠	١٠	٢٠	٤	٢٠	٢	٢٠	١	٢٠	٥	٢٠	٧	١٨	٣	١٧
٢٣	٧	١٦	٦	٢٩	١١	٢١	٣	٢٢	٩	٢١	٧	١٧	٧	٢٠	٥	١٩	١٠	١٧	٣	١٨
٣٤	٣	٢٢	١	٣٢	٦	١٨	٧	١٩	٦	١٨	٩	١٨	٧	٢٠	٩	١٨	٩	٢١	٨	١٨
٢١	٨	٢٠	١٠	٣٠	١	٢٠	١	٢٠	٥	١٩	٩	٢١	٧	٢٠	١	٢٠	٦	٢١	٥	٢١
١٦	١	٢٠	١٠	٣١	٢	٢١	٩	٢١	٨	٢٣	٥	٢٠	٧	٢٠	١١	٢٠	٩	١٧	١	١٩
٢٠	١	١٩	٦	٣١	١٠	٢٢	١	١٦	١٠	١٨	١١	٢٠	٩	١٨	٨	٢١	١	١٩
١٩	٢	٢١	٣	٣٥	٩	٢٦	١١	١٨	٨	١٨	١١	١٨	١	٢٠	٧	٢٤	١	٢٢
٢٠	١١	١٩	٤	٢٣	٩	٢٩	١١	٢٠	١	٢٠	٨	٢٢	٥	١٦	٧	١٦	٩	١٩	٤	١٩
١٩	١	٢٢	١٠	٢٣	١	٣١	٢	٢٠	٨	٢٠	٢	٢٠	١	٢١	١	٢٢	١	١٨	١	١٨	١١	١٨
٢٢	٦	١٨	١	٢٠	٨	٢٠	١١	٢٠	١	١٨	١	٢٠	١	٢٠	١	٢٠	٧	٢٠	٨	١٧	٤	١٩
٢٤	٩	١٨	٧	١٨	٢	١٨	٥	١٨	٨	١٧	٣	١٨	١١	٢٢	٨	٢٢	١	٢٠	١	١٨
٣٢	٢	١٩	٩	٢٤	١١	١٨	٣	١٦	٩	١٨	١٠	٢١	١	٢٠	٢	١٨	٧	٣٣	٧	١٦
٢٢	٩	٢٠	٢	٢٣	٧	١٨	٣	٢٥	١	١٧	١١	١٧	٧	١٧	٨	٢١	٥	١٩	٧	٢٠
٢٢	١١	٢١	٩	٣٤	٣	٢٣	٤	٢٢	١	١٩	١١	٢٠	٨	٣٤	١	٢١	٢	٢١	١	٢٣

لو نظرنا إلى هذا العدد الكبير من الأرقام لوجدنا أنه من الصعب جداً الوصول إلى نتائج ملموسة من مجرد قراءة الكشوف التى تحتوى على أعمار هؤلاء الطلبة . فالأرقام الأولية فى شكلها هذا وإن كانت تعطينا فكرة عن التفاوت الكبير بين أعمار بعض الطلبة إلا أنها لا تساعد كثيراً فى وصف المجموعة من حيث تركز مفرداتها حول سن معين هو السن المعتاد أو السن المتوسط أو السن النموذجى للطالب عند التحاقه بالجامعة . فإذا أردنا معرفة السن المتوسط بدقة من واقع هذه الأرقام ، يمكننا أن نصل إلى هذه النتيجة بجمع أعمار جميع الطلبة وقسمة الناتج على عددهم . غير أن صعوبة هذه العملية تتضح لنا من مجرد النظر إلى هذا الحشد الكبير من الأرقام ، وتزداد وضوحاً إذا كانت مجموعة الطلبة المراد تقدير العمر المتوسط لها تشمل بضعة آلاف من المفردات كما لو أردنا معرفة العمر المتوسط بين جميع طلبة الجامعة أو بضعة ملايين كما لو أردنا معرفة الأجر المتوسط للعامل المصرى فى جميع أنحاء القطر بهذه الطريقة المطولة العقيمة . من الواضح أنه لكى يسهل علينا معالجة مثل هذه المجموعات الكبيرة لا بد أن نضحى بشىء قليل من الدقة فى سبيل تسهيل العمل الحسابى الذى قد يكون مرهقاً ومملا خصوصاً إذا لم تتوفر لدينا بعض الآلات الحاسبة . وعلاوة على ذلك فإن توخى الدقة الحسابية التامة فى وصف مثل هذه المجموعات الكبيرة لا مبرر له إطلاقاً . فمن الوجهة العملية لا يهمنا كثيراً ما إذا كان عمر الطالب المتقدم للالتحاق بالكلية ١٧ سنة وخمسة أشهر أو ١٨ سنة وشهرين أو ١٨ سنة وخمسة أشهر . إذ لا داعى مطلقاً للتمييز بين هذا وذاك بناء على هذه الفروق الصغيرة فى السن . وبالمثل إذا كنا بصدد بحث عن مستوى الأجور فى مصر ، فلا يهمنا كثيراً من الوجهة العملية ما إذا كان العامل يتقاضى ٥ أو ٧ أو ٨ أو ١٠ قروش يومياً . ولو فرض أن كانت هناك مجموعة كبيرة من الأجور تتراوح بين ٥ و ١٠ قروش يومياً لأمكن اعتبارها جميعاً ٧,٥ قرش يومياً . على أن هذا التقريب فى الأرقام لا يترتب عليه عادة

تضحية كبيرة فى دقة العمل ، فما دامت الأجور موزعة بين ٥ و ١٠ بطريق المصادفة[١] فإن عدد الأجور المحصورة بين ٥ و ٧,٥ لا يختلف كثيراً عن عدد الأجور المحصورة بين ٧,٥ و ١٠ وعلى ذلك فإن الزيادة التى نضيفها إلى الأجور الأقل من ٧,٥ تعادل تقريباً النقص الذى نحدثه فى الأجور الأكبر من ٧,٥ . وهكذا يكون الخطأ فى نتائجها النهائية ضئيلا بحيث يمكن التجاوز عنه . ما دمنا إذن على استعداد للتغاضى عن مثل هذه الفروق الضئيلة ، فيمكننا اعتبار جميع الطلبة الواقعة أعمارهم بين ١٦ و ١٨ مجموعة أو فئة متجانسة من حيث السن بحيث يصبح الطالب الذى سنه ١٦,٥ سنة — من وجهة نظرنا هذه — مماثلا للطالب الذى سنه ١٧,٥ سنة ، إذ يمكننا من الوجهة العملية اعتبار سن كل منهم يعادل السن الواقع فى منتصف الفترة ما بين الحد الأعلى والحد الأدنى للفئة أو مركز الفئة وهو فى هذه الحالة ١٧ سنة . وبالمثل يمكننا اعتبار جميع الطلبة الواقعة أعمارهم بين ١٨ و ٢٠ فى مجموعة أخرى متجانسة يمثلها مركز هذه الفئة وهو السن ١٩ . ونستمر فى تقسيم مجموعتنا الأصلية على هذا النحو حتى يتم وضع كل طالب فى الفئة الخاصة به . وبهذه التضحية البسيطة فى الدقة التى لا تؤثر كثيراً على النتائج التى نحصل عليها يمكننا وضع هذه المجموعة الكبيرة من الأرقام فى صورة تسهل معالجتها وتوفر علينا وقتاً ومجهوداً كبيرين . فباتباع هذه الطريقة يمكننا تلخيص الأرقام الخاصة بأعمار الطلبة فى الجدول الآتى :—

---

(١) الذى نقصده من قولنا إن الأجور موزعة بين ٥ و ١٠ بطريق المصادفة أنه ليس هناك سبب يدعونا إلى الاعتقاد أن عدد الأجور القريبة من ٥ قروش يزيد (أو يقل) كثيراً عن عدد الأجور الأقرب إلى ١٠ قروش .

جدول رقم ٧

التوزيع التكرارى لأعمار الطلبة المتقدمين للالتحاق بكلية التجارة

١٩٤٧ – ١٩٤٨

فئات الأعمار	التكرار
١٦ وأقل من ١٨	٢٠
١٨ – ٢٠	٥٥
٢٠ – ٢٢	٦٥
٢٢ – ٢٤	٣٢
٢٤ – ٢٦	٨
٢٦ – ٢٨	٥
٢٨ – ٣٠	٢
٣٠ – ٣٢	٦
٣٢ – ٣٤	٣
٣٤ – ٣٦	٤
المجموع	٢٠٠

إن مجرد النظر إلى هذا الجدول يعطينا فكرة لا بأس بها عن مجموعة الطلبة هذه من حيث السن . إذ يظهر من الجدول بكل وضوح أن الأعمار مركزة فى الفئتين الثانية والثالثة ، أى أن أغلبية الطلبة المتقدمين للالتحاق بكلية التجارة تتراوح أعمارهم بين الثامنة عشرة والثانية والعشرين . وكلما ابتعدنا عن السن النموذجى – الذى يمكن اعتباره هنا سن العشرين على وجه التقريب – سواء أكان اتجاهنا نحو الفئات العليا أو الدنيا كلما قل عدد الطلبة الواقعة أعمارهم فى كل فئة ، حتى إذا ما وصلنا إلى الفئات المتطرفة وجدناها لا تحتوى إلا على طالبين أو ثلاثة . ولا شك أن توزيع الطلبة على الفئات المختلفة بهذه

الطريقة ، فضلا عن أنه يسهل العمل الحسابى ، يساعد كثيراً على وصف المجموعة بدقة . إذ يبرز لنا بوضوح الطابع الخاص الذى يميزها عن غيرها وهو ما لا نستطيع الوصول إليه بالتأمل ساعات طوال فى الأرقام الأصلية . ويسمى توزيع المفردات على الفئات المختلفة بهذه الكيفية التوزيع التكرارى كما يسمى الجدول الذي يشتمل على هذا التوزيع الجدول التكرارى . أما الفئات ذاتها فيعبر عنها بالفئات التكرارية .

ويلاحظ أنه بعد أن يتم توزيع المفردات على الفئات التكرارية المختلفة تضيع معالم الأرقام الأولية فلا يمكننا بالنظر إلى الجدول التكرارى أن نقرر ما إذا كان عمر الطالب فى أية فئة قريباً من الحد الأعلى أو الحد الأدنى للفئة . ذلك أن فكرة التوزيع التكرارى مبنية على افتراض أن جميع المفردات الواقعة داخل الفئة متساوية ومساوية لمركز الفئة .

اختيار الفئات .

من البديهى أن درجة الدقة التى نضحيها بوضع الأرقام الأولية فى شكل توزيع تكرارى تتوقف على اتساع الفئات . فقد تكون الفئة من الاتساع بحيث يمكن وضع جميع المفردات فى فئة أو اثنين كما لو جعلنا فئات أعمار الطلبة السابق ذكرهم ١٦ — ٢٦ و ٢٦ — ٣٦ . وفى هذه الحالة يكون عملنا بعيداً كل البعد عن الدقة رغم سهولته . وقد تكون الفئات من الضيق بحيث لا تشمل الفئة الواحدة سوى عدد قليل من المفردات كما لو جعلنا فئات أعمار الطلبة ١٦ — ١٦,٥ و ١٦,٥— ١٧ و ١٧ —١٧,٥ .... إلخ . وفى هذه الحالة يصبح التوزيع التكرارى غير منتظم فضلا عن أن درجة الدقة التى نكتسبها لا تتناسب مطلقاً مع زيادة مشقة العمل الحسابى . بل إنه من المحتمل أن نفقد شيئاً من دقة العمل إذا ما جعلنا طول الفئة من الصغر بحيث يصبح عدد الفئات كبيراً جداً . وذلك لأن دقة تمثيل مركز الفئة لبقية المفردات

الموجودة داخل الفئة تتوقف على وجود عدد كاف من المفردات فى هذه الفئة .

يتضح مما تقدم أنه يتعين علينا عند اختيار الفئات أن نوازن بين عاملين : الدقة من ناحية ومشقة العمل من ناحية أخرى . والقاعدة العامة فى تحديد طول الفئة هى ألا تكون من الكبر بحيث تضيع معالم التوزيع التكرارى فلا يكون له طابع مميز ، ولا من الصغر بحيث يشتمل التوزيع على فئات ليس بها عدد كاف من المفردات . هذا طبعاً مع استثناء الفئات الواقعة فى طرفى التوزيع التى يكون بها عادة عدد قليل جداً من المفردات . ولا شك أن عدد المفردات التى نعالجها وطبيعة النتائج التى نرغب فى الوصول إليها لها أثر كبير فى تحديد طول الفئة وبالتالى عدد الفئات . وقد حاول بعض الإحصائيين وضع علاقات رياضية تربط عدد فئات التوزيع التكرارى بعدد مفردات المجموعة الموضوعة تحت البحث وأهم هذه العلاقات تلك المعروفة بقاعدة ستيرجس Sturges' Rule التى تنص على أن :

$$ف = ١ + \frac{لو\ ن}{لو_{١٠}\ ٢}$$

حيث ف = عدد فئات التوزيع التكرارى

و ن = عدد مفردات المجموعة .

وحيث أن $\frac{١}{لو_{١٠}\ ٢} = ٣,٣٢١٩ = ٣,٣$ تقريباً

يمكن وضع قاعدة ستيرجس فى الصيغة الآتية :

$$ف = ١ + ٣,٣\ لو\ ن$$

وبتطبيق هذه القاعدة على مجموعة أعمار الطلبة المتقدمين للإلتحاق بكلية التجارة نجد أن :ـ

$$ف = ١ + ٣,٣\ لو\ ٢٠٠ = ٨,٥٩٣$$

أى أن عدد فئات التوزيع التكرارى يجب أن تكون ٨ أو ٩ حسب هذه القاعدة ، وهذا العدد لا يختلف كثيراً عن ذلك الذى وجدناه مناسباً لهذه المجموعة وقدره ١٠ فئات .

وقد توصل ستيرجس إلى قاعدته هذه من ملاحظة العلاقة بين عدد المفردات وعدد الفئات فى التوزيع التكرارى لذات الحدين الذى سنشرحه بالتفصيل فى الباب الخاص بالتوزيعات التكرارية النظرية .

ولتوضيح هذه العلاقة افرض مثلا أننا رمينا ٥ قطع من النقود كاملة التوازن ٣٢ مرة وأننا سجلنا عدد المرات التى تقع فيها القطع على الوجه فى كل مرة . الذى نتوقعه نظرياً هو أن يكون ظهور الوجه كالآتى :—

عدد مرات ظهور الوجه	التكرار
٠	١
١	٥
٢	١٠
٣	١٠
٤	٥
٥	١
المجموع	٣٢

كذلك إذا رمينا ٦ قطع ٦٤ مرة فالنتيجة التى نتوقعها نظرياً هى الآتية :—

عدد مرات ظهور الوجه	التكرار
٠	١
١	٦
٢	١٥
٣	٢٠
٤	١٥
٥	٦
٦	١
المجموع	٦٤

أى أنه عندما يكون عدد مفردات التوزيع التكرارى ٣٢ يكون عدد

$$\text{الفئات} = ١ + \frac{\text{لو } ٣٢}{\text{لو } ٢} = ٦ \text{ فئات}$$

وعندما يكون عدد مفردات التوزيع ٦٤ يكون عدد الفئات

$$= ١ + \frac{\text{لو } ٦٤}{\text{لو } ٢} = ٧ \text{ فئات وهكذا .}$$

هذه هى العلاقة التى لاحظها ستيرجس ورأى أن تطبق عملياً فى معرفة عدد فئات التوزيعات التكرارية بوجه عام .

على أنه ليس ثمة سبب يدعونا إلى التمسك بقاعدة ستيرجس أو أي علاقة رياضية أخرى عند تحديد الفئات . فاختيار الفئات يتوقف على طبيعة البيانات مثلما يتوقف على عدد المفردات . والواقع أن قاعدة Sturges لا تعطى نتائج مرضية إذا كان عدد المفردات كبيراً جداً (أكثر من ١٠٠٠ مثلا) أو صغيراً جداً (أقل من ١٠ مثلا) ، فهى تعطى عدداً من الفئات أقل من اللازم فى الحالة الأولى وأكثر من اللازم فى الحالة الثانية . ولبيان ذلك افرض أن عدد المفردات ١٠,٠٠٠ ، بتطبيق قاعدة Sturges يكون عدد الفئات يساوى :

١ + ٣,٣ لو ١٠,٠٠٠ = ١٤,٢ = ١٤ فئة تقريباً .

وهذا العدد بلاشك أقل من اللازم .

وإذا كان عدد المفردات ١٠ يكون عدد الفئات يساوى :

١ + ٣,٣ لو ١٠ = ٤ فئات تقريباً .

وبديهى أن أى إحصائى لا يفكر فى وضع عشر مفردات فى ٤ فئات .

وعلى ذلك فالأفضل أن يكون استعمال قاعدة Sturges مقصوراً على التوزيعات التكرارية التى يكون عدد مفرداتها متوسطاً (بين ١٠٠ و ١٠٠٠ مثلا) على ألا تطبق كلية إذا لاحظنا عدم اتفاقها مع القاعدة العامة .

بعد الاتفاق على تقسيم المدى بين أصغر قيمة وأكبر قيمة فى المجموعة إلى عدد مناسب من الفئات ، يجب مراعاة النقط الآتية فى اختيار أطوالها وحدودها :

أولا : يحسن أن نجعل طول الفترة ما بين الحد الأعلى والحد الأدنى واحداً فى جميع فئات التوزيع وذلك لتسهيل العمل الحسابى . وهناك بعض الحالات الخاصة التى لا مفر فيها من تغيير طول الفترة وذلك مراعاة للمحافظة على تجانس المفردات الموجودة فى كل فئة . افرض مثلا أننا أردنا عمل توزيع تكرارى لأعمار سكان القطر المصرى توطئة لمقارنة معدل الوفيات فى كل فئة . من الواضح أن توحيد طول الفترة فى جميع فئات الأعمار يكون غير مناسب لهذا الغرض . فليس المقصود هو مقارنة معدلات الوفيات فى كل سنة من العمر أو فى كل خمس أو عشر سنوات ، وإنما المقصود هو مقارنة الفئات التى تتميز بظروف خاصة عن غيرها أو بعبارة أخرى مقارنة وفيات الأطفال بوفيات الشباب ومتوسطى العمر والشيوخ وهكذا . وعلى ذلك يصح أن تكون فئات الأعمار كالآتى :—

أقل	من	١
١	—	٥
٥	—	١٥
١٥	—	٢٥
٢٥	—	٤٥
٤٥	—	٦٥
٦٥	فأكثر	

ويلاحظ أن كل فئة من الفئات السابقة تضم مجموعة متجانسة تقريباً من حيث ظروف التعرض للوفاة . كذلك إذا أردنا عمل جدول تكرارى يوضح توزيع الملكية الزراعية فى مصر ، فمن العبث أن نختار فئات متساوية فى الطول فهذه لا يمكن أن تكون ذات فائدة عملية . فليس المقصود من عمل

التوزيع التكرارى معرفة عدد الأشخاص الذين يملكون ثلاثة أفدنة وهؤلاء الذين يملكون عشرة أو عشرين أو أربعين فدانا . . . إلخ ، وإنما المقصود هو مقارنة صغار الملاك بذوى الملكيات المتوسطة وكبار الملاك . ونظراً لوجود فروق كبيرة بين الطبقات المختلفة من صغار الملاك بحيث لا يجوز اعتبارهم جميعاً مجموعة متجانسة يجب وضع الفئات الأولى بتفصيل أكثر هكذا :ـ

أقل من ١ فدان
١ ـ ٥
٥ ـ ١٠
١٠ ـ ٢٠
٢٠ ـ ٥٠
٥٠ فما فوق

ثانياً : يجب مراعاة الظروف التى جمعت فيها البيانات الأصلية ومقدار دقتها عند اختيار الفئات . فعند عمل التعداد العام للسكان مثلا ، يميل أشخاص كثيرون إلى ذكر أعمارهم فى شكل أعداد دائرية (مضاعفات للعشرة أو الخمسة) فالشخص الذى يكون عمره الحقيقى ٢٧ أو ٣٣ يكتفى عادة بأن يذكر أن عمره ٣٠ سنة . ولذلك يستحسن فى مثل هذه الحالة أن نجعل الفئات : ـ

١٥ ـ ٢٥
٢٥ ـ ٣٥
٣٥ ـ ٤٥

بدلا من

٢٠ ـ ٣٠
٣٠ ـ ٤٠
٤٠ ـ ٥٠

لأننا لو اتبعنا الطريقة الأولى لوضعنا الأشخاص الذين تتراوح أعمارهم بين ٢٧ ، ٣٣ فى الفئة ٢٥ ــ ٣٥ أى فى الفئة التى مركزها ٣٠ وهو الرقم الذى يذكره هؤلاء الأشخاص عادة . أما لو اتبعنا الطريقة الثانية فربما نضع شخصاً سنه الحقيقى ٢٧ أو ٢٨ (ويقربه هو إلى ٣٠) فى الفئة ٣٠ ــ ٤٠ التى مركزها ٣٥ . وما دمنا نعتبر جميع المفردات الموجودة داخل الفئة مساوية لمركز هذه الفئة، فإن هذا الشخص سيعد كأنه فى سن ٣٥ مع أن سنه فى الواقع ٢٧ أو ٢٨ .

ثالثاً : يحسن أن تكون الفئات متداخلة فتبدأ الفترة الثانية من حيث تنتهى الفترة الأولى هكذا .

٥ ــ ١٠
١٠ ــ ١٥
١٥ ــ ٢٠

والواقع أن وضع الفئات بهذه الصورة يفهم منه ضمناً أن المقصود هو :

٥ وأقل من ١٠
١٠ وأقل من ١٥ وهكذا

ولا يمكن أن يكون المقصود خلاف ذلك . ويفضل بعض الإحصائيين استخدام الفئات غير المتداخلة دفعاً لما قد يحدث من لبس عند وضع المفردات التى تتفق قيمها مع نهايات أو بدايات الفئات فتكتب الفئات بالطريقة الاتية :

٥ ــ ٩
١٠ ــ ١٤
١٥ ــ ١٩ وهكذا

غير أنه لا يمكن أن يكون المقصود هو ترك فراغ بين كل فئة وأخرى بحيث يتعذر وضع مفردة قيمتها ٩,٥ مثلا فى أى فئة . وإنما المقصود هو أن تمتد الفئة الأولى فتشمل جميع المفردات من :

٤,٥ — ٩,٥

والثانية من ٩,٥ — ١٤,٥

والثالثة من ١٤,٥ — ١٩,٥ وهكذا

ومن هذا يظهر أننا نعود دائماً إلى طريقة الفئات المتداخلة . ولهذا السبب لا نجد ما يبرر اعتراض بعض الكتاب على استعمالها[١].

ومهما يكن الغرض من وضع الفئات فى الصورة الأولى (غير المتداخلة) فيجب تعديلها لجعلها متداخلة قبل إجراء أية عملية حسابية على التوزيع مثل استخراج المتوسطات أو مقاييس التشتت وغير ذلك من المقاييس الإحصائية .

الطريقة العملية لوضع الأرقام الأولية فى الفئات التكرارية .

بعد أن ننتهى من اختيار الفئات نعمل جدولا يحتوى على عدد من الخانات بقدر عدد الفئات . ثم نرصد الأرقام الأولية واحداً تلو الاخر كلا فى الخانة الخاصة به وذلك بوضع علامة (/) للتعبير عن كل رقم بالطريقة التى سبق توضيحها عند شرح طرق التبويب . والواقع أن وضع الأرقام الأولية فى شكل توزيع تكرارى هو نوع من التبويب الكمى[٢].

والشكل الآتى يوضح طريقة رصد أعمار المائتى طالب المتقدمين للالتحاق بكلية التجارة فى سنة ١٩٤٧ :

---

(١) انظر مثلا مبادئ الإحصاء الدكتور عبد المنعم الشافعى ص ٩٦ من الجزء الأول .

(٢) راجع ص ٢٤ من هذا الكتاب .

فئات الأعمار	
١٦ – ١٨	٤
١٨ – ٢٠	٩
٢٠ – ٢٢	١٤
٢٢ – ٢٤	١٢
٢٤ – ٢٦	٨
٢٦ – ٢٨	٥
٢٨ – ٣٠	٢
٣٠ – ٣٢	٥
٣٢ – ٣٤	٣
٣٤ – ٣٦	٤

بعد ذلك نجمع الإشارات الموجودة أمام كل فئة فنحصل على التوزيع التكرارى لأعمار الطلبة الذى سبق بيانه فى الجدول رقم ٧ .

العرض البيانى للتوزيع التكرارى .

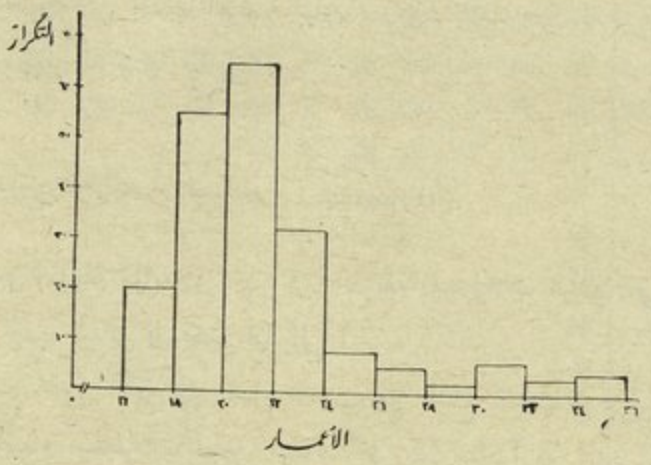
رأينا فى الباب السابق كيف أن العرض البيانى يفضل العرض الجدولى بوجه عام . ويمكننا أن نؤكد هذه النقطة مرة أخرى بأن نطلب إلى القارئ أن يقارن بين المجهود الذى يبذله فى استيعاب محتويات مائة جدول إحصائى وذلك الذى يبذله فى النظر إلى عدد مماثل من الرسوم البيانية التى تعرض نفس البيانات. من الواضح أن المجهود الأخير يكون أقل جداً من الأول . فالملل يتسرب إلى القارئ بعد دراسة خمسة أو عشرة جداول ، فى حين أن مجرد النظر إلى عدد كبير من الرسوم البيانية لا يبعث إلى نفسه الملل خصوصاً إذا كانت البيانات معروضة فى صورة واضحة جذابة . لهذا نجد أنه على الرغم من أن الجدول التكرارى يلخص الأرقام الأولية بطريقة تجعل من السهل تكوين فكرة عامة عن

المجموعة التى تمثلها هذه الأرقام ، إلا أنه لا يحقق هذا الغرض بنفس السرعة ونفس السهولة التى يحققها العرض البيانى . وهناك ثلاث طرق لعرض التوزيع التكرارى بيانيا نذكرها فيما يلى :—

١ — المدرج التكرارى أو الهستوجرام (Histogram) .

هذه الطريقة تشابه إلى حد كبير طريقة المستطيلات الرأسية التى أشرنا إليها فى الباب السابق وتتلخص فيما يأتى :—

نقسم المحور الأفقى إلى أقسام تتناسب مع اتساع الفئات . ثم نرسم فوق هذه الأقسام مستطيلات نجعل ارتفاع كل منها متناسباً مع عدد المفردات الموجودة داخل كل فئة وتكون قواعدها عبارة عن التقسيمات التى تمثل أطوال الفئات على المحور الأفقى . وحيث أن معظم التوزيعات التكرارية تكون مفرداتها مركزة فى الفئات الوسطى وتلك القريبة منها ، بحيث يقل عدد المفردات فى الفئات كلما ابتعدنا عن وسط التوزيع ، فإن المستطيلات التى تمثل عدد المفردات تتدرج فى الارتفاع حتى تصل إلى حد أعلى فى وسط التوزيع ثم تعود فتتدرج فى الانخفاض حتى نهاية التوزيع . ولهذا تظهر هذه المستطيلات فى الرسم على شكل مدرج يسمى المدرج التكرارى . والرسم الآتى يوضح المدرج التكرارى لأعمار مجموعة الطلبة المبينة بالجدول رقم (٧) .



(شكل رقم ٣١)

المدرج التكرارى لأعمار الطلبة المتقدمين للالتحاق
بكلية التجارة سنة ١٩٤٧ – ١٩٤٨

المدرج التكرارى لتوزيع ذي فئات غير متساوية

إذا كانت فئات التوزيع غير متساوية فى الطول فيجب إدخال تعديل بسيط على طريقة العمل لتجنب تضليل الشخص الذى ينظر إلى الرسم دون أن يتنبه إلى اختلاف اتساع الفئات . افرض مثلا أننا أردنا رسم مدرج تكرارى للتوزيع الآتى :–

جدول رقم ٨

فئات المتغير			التكرار
١٠	—	٢٠	٥٠
٢٠	—	٣٠	١٠٠
٣٠	—	٤٠	٢٠٠
٤٠	—	٦٠	٣٠٠
٦٠	—	١٠٠	٤٠٠
١٠٠	—	١٥٠	٢٥٠

إذا نظرنا إلى هذا التوزيع يتراءى لنا لأول وهلة أن تركز المفردات فى الفئة ٦٠ — ١٠٠ أكثر منه فى أية فئة أخرى . غير أنه بالتأمل فى الجدول نجد أن طول هذه الفئة يعادل ضعف طول الفئة التى مداها ٤٠ — ٦٠ ويعادل ضعفى طول الفئة التى مداها ٣٠ — ٤٠ . وعلى ذلك فإن اختلاف اتساع الفئات هو الذى يجعلنا نتوهم أن المفردات مركزة فى الفئة ٦٠ — ١٠٠ والحقيقة أن تركز المفردات فى الفئة ٣٠ — ٤٠ أكثر نسبياً منه فى الفئة ٦٠ — ١٠٠ أو أية فئة أخرى . وذلك لأنه لو أردنا مقارنة تركز المفردات فى هاتين الفئتين يجب توحيد أساس المقارنة فى الحالتين وإلا تكون مقارنتنا عديمة الجدوى إن لم تكن مضللة فعلا . فهذه المقارنة لا تختلف فى طبيعتها عن مقارنة الربح الذى حققه مشروع معين فى سنة واحدة بالأرباح التى حققها فى أربع سنوات مجتمعة . وبديهى أنه لكى تكون هذه المقارنة مبنية على أساس سليم يجب

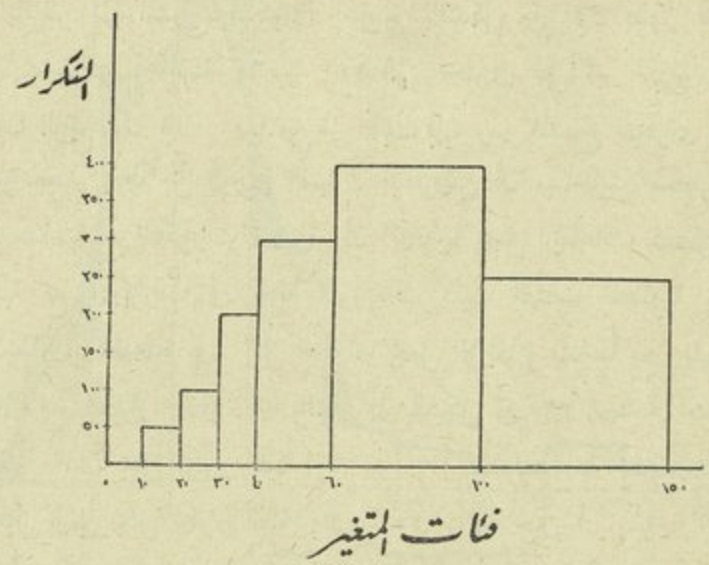
أن نقسم مجمل أرباح الأربع سنوات على عدد هذه السنوات فنحصل على رقم يناظر ربح السنة الواحدة . وبالمثل يمكن توحيد أساس المقارنة بين الفئات المختلفة بقسمة عدد التكرارات على طول الفئة فى كل حالة . وبعد هذا التعديل يصبح التوزيع التكرارى كالآتى :—

جدول رقم ٩

فئات المتغير	التكرار	التكرار مقدراً بوحدات طول الفئة
١٠ — ٢٠	٥٠	٥
٢٠ — ٣٠	١٠٠	١٠
٣٠ — ٤٠	٢٠٠	٢٠
٤٠ — ٦٠	٣٠٠	١٥
٦٠ — ١٠٠	٤٠٠	١٠
١٠٠ — ١٥٠	٢٥٠	٥

يظهر من الجدول السابق أنه بعد تعديل التكرارات لتأخذ فى الحسبان اختلاف اتساع الفئات يكون تركز المفردات فى الفئة ٣٠ — ٤٠ أكبر منه فى أية فئة أخرى .

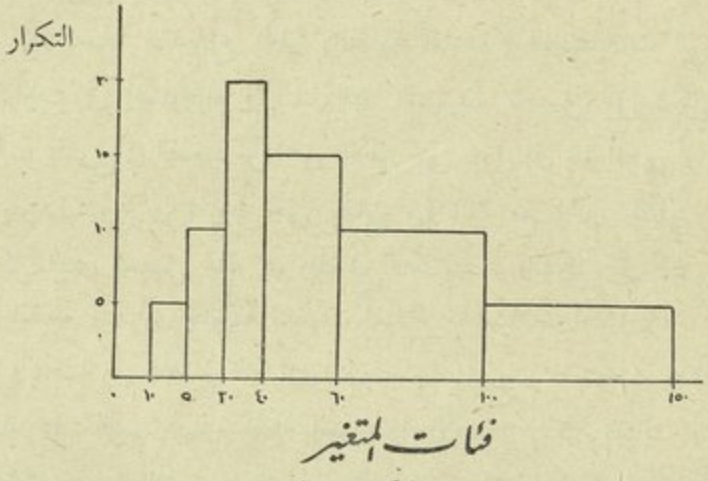
وتتضح أهمية عمل هذا التعديل من مقارنة المدرج التكرارى للتوزيع بحالته الأصلية بالمدرج التكرارى الذى يمثله بعد التعديل . (قارن شكلى ٣٢ و ٣٣) .



(شكل رقم ٣٢)

مدرج تكرارى ذو فئات غير متساوية

قبل تعديل التكرارات



(شكل رقم ٣٣)

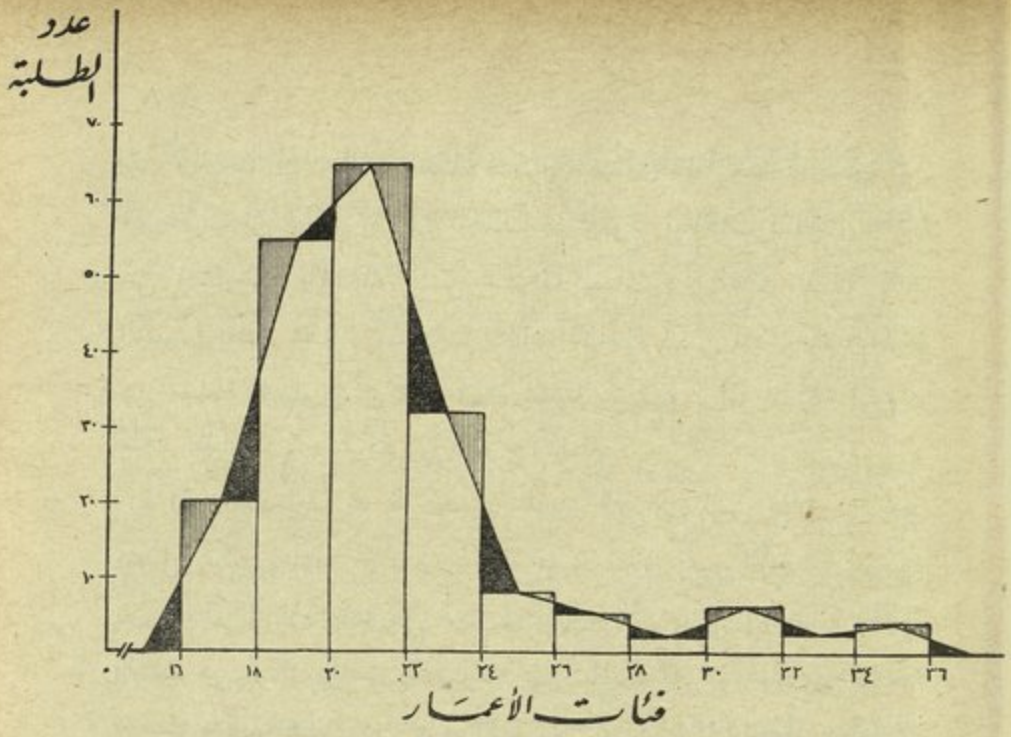
المدرج بعد تعديل التكرارات

بالنظر إلى المدرج التكرارى الموضح بالشكل رقم ٣٢ يظهر لنا أن المفردات مركزة فى الفئة ٦٠ — ١٠٠ التى تحتوى على أكبر تكرار وهذا خلاف الواقع كما بينا . وعلاوة على ذلك فإن رسم المدرج التكراري بهذه الكيفية يعتبر خطأ من الوجهة الفنية . فالمفروض أن مساحات المستطيلات تكون متناسبة مع التكرارات . والواقع أن اكتفاءنا بجعل ارتفاعات المستطيلات متناسبة مع التكرارات فى حالة التوزيعات ذات الفئات المتساوية يرجع إلى تساوى القاعدة فى كل حالة مما يجعل الإرتفاع متناسباً مع المساحة وبالتالى مع التكرار . غير أنه بالنظر إلى الشكل رقم ٣٢ نجد أن مساحة المستطيل الذى يمثل تكرار الفئة ٦٠ — ١٠٠ يعادل ثمانية أمثال مساحة المستطيل الذى يمثل تكرار الفئة ٣٠ — ٤٠ ، فى حين أن مساحة الأول يجب أن تكون ضعف مساحة الثانى فقط . أما الشكل رقم ٣٣ فيظهر لنا حقيقة موضع تركز المفردات كما أنه صحيح من الوجهة الفنية لأن مساحات المستطيلات فيه متناسبة مع تكرارات الفئات .

٢ — المضلع التكرارى . Freqency Polygon

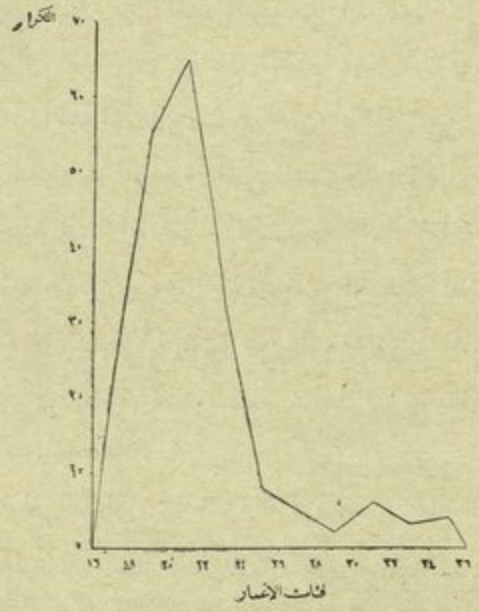
إذا نصفنا الاضلاع العليا (المقابلة للقاعدة) للمستطيلات التى تمثل التكرارات فى الهستوجرام ووصلنا نقط التنصيف نحصل على شكل خط منكسر يتدرج فى الصعود من المحور الأفقى حتى يصل إلى حد أقصى أو قمة ، ثم يعود فيتدرج فى الهبوط حتى يتقابل مرة ثانية مع المحور الأفقى . هذا الشكل الذى نحصل عليه هو المضلع التكرارى . والشكل رقم ٣٤ يوضح رسم المضلع التكرارى بطريقة تنصيف أضلاع مستطيلات الهستوجرام .

ويمكننا رسم المضلع التكرارى دون حاجة إلى رسم الهستوجرام وذلك بأن نرصد على الرسم البيانى نقطاً إحداثياتها الأفقية مراكز الفئات والرأسية تكرارات هذه الفئات ثم نصل هذه النقط ببعضها بخطوط مستقيمة كما هو موضح بالشكل رقم ٣٥ .



(شكل رقم ٣٤)

رسم المضلع التكرارى على الهستوجرام



(شكل رقم ٣٥)

المضلع التكرارى لأعمار الطلبة المتقدمين للالتحاق

بكلية التجارة ١٩٤٧ – ١٩٤٨

ويلاحظ أن مساحة المضلع التكرارى تساوى تماماً مساحة الهستوجرام . فلو نظرت إلى الشكل رقم ٣٤ لوجدت أن مجموع مساحات المثلثات الخارجة عن الهستوجرام والداخلة ضمن مساحة المضلع (أى تلك المظللة باللون الأسود) يعادل مجموع مساحات المثلثات الداخلة ضمن الهستوجرام والخارجة عن مساحة المضلع (أى تلك المظللة بخطوط رأسية) . أما المساحة المتروكة بدون تظليل فهى مشتركة بين المدرج والمضلع .

ومن المهم جداً أن نلاحظ أن المضلع التكرارى ليس عبارة عن خط متصل يمكن اعتباره خط سير نقطة معينة تتخذ جميع الأوضاع الموجودة بين النقط التى رصدناها على الرسم فى بادىء الأمر والتى إحداثياتها الأفقية والرأسية هى مراكز الفئات والتكرارات على التوالى ، بل هو عبارة عن مجموعة خطوط منكسرة تصل بين هذه النقط التى تحدد شكل المضلع . أما أية نقطة أخرى على هذه الخطوط فليس لها أى معنى خاص يتصل بالتوزيع التكرارى . وعلى ذلك فإذا أقمنا عموداً من أية نقطة على خط القاعدة — — خلاف مراكز الفئات — ليقاطع المضلع فإن طول هذا العمود لا يمثل التكرار عند هذه النقطة . والسبب فى ذلك أن كلا من الهستوجرام والمضلع التكرارى يمثل التكرارات الموجودة داخل الفئات ولا يمثل بأى حال من الأحوال التكرارات الموجودة عند أية نقطة كانت على خط القاعة (عند سن معين مثلا) . فالتوزيع التكرارى وإن كان يوضح لنا عدد المفردات الموجودة فى كل فئة إلا أنه لا يعطينا أية فكرة عن طريقة توزيع هذه المفردات داخل الفئات[1]. وستظهر أهمية هذه النقطة عندما نشرح الفرق بين المتغيرات المتصلة والمتغيرات المنفصلة .

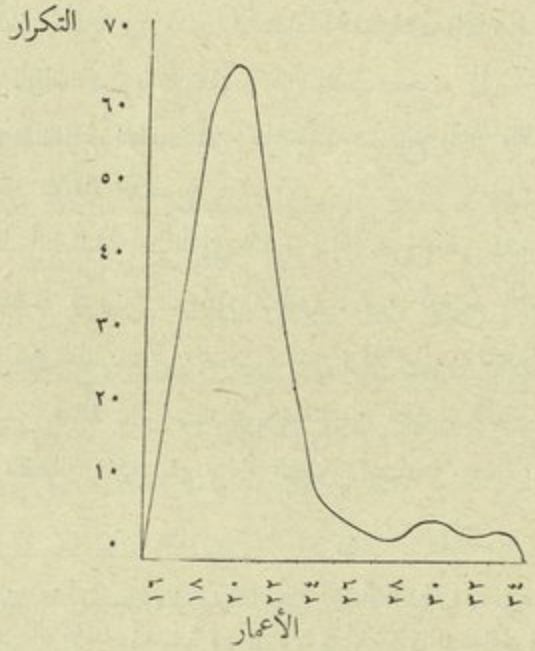
وليس للمضلع التكرارى فى حد ذاته أى تطبيق عملى يستحق الذكر ،

(١) نلجأ أحياناً إلى اعتبار جميع المفردات الموجودة داخل أية فئة مركزة عند مركز هذه الفئة . وأحياناً إلى اعتبار المفردات موزعة بانتظام على طول فترة الفئة . غير أن هذه فروض تقريبية نلجأ إليها لتسهيل العمل الحسابى ليس إلا .

وتكاد تكون فائدته مقصورة على استخدامه كخطوة أولية فى رسم المنحنى التكرارى الآتى الكلام عنه فيما يلى . ولا شك أن الهستوجرام يفضله كثيراً كطريقة من طرق العرض البيانى للتوزيع التكرارى ففكرته أقرب إلى الفهم ، كما أنه يظهر توزيع المفردات داخل الفئات بشكل واضح جلى ، الأمر الذى لا يحققه المضلع التكرارى.

٣ — المنحنى التكرارى .

إذا كان التوزيع التكرارى يمثل قيم متغير متصل ، أى متغير يصح أن تمتد القيم التى يتخذها على طول مدى الأعداد الحقيقية بين $+\infty$ ، $-\infty$ أو تمتد على طول فترة محدودة بين هاتين النهايتين دون أن يكون هناك فراغ بين قيمة وأخرى ، فيمكن تمهيد انكسارات المضلع التكرارى إما باليد أو بإحدى طرق توفيق المنحنيات ، فيتخذ شكل منحن متصل خال من الذبذبات والتعاريج يطلق عليه اسم المنحنى التكرارى والشكل الآتى يبين المنحنى التكرارى لتوزيع أعمار الطلبة المتقدمين للالتحاق بكلية التجارة :



(شكل رقم ٣٦)

المنحنى التكرارى لأعمار الطلبة المتقدمين للالتحاق بكلية التجارة ٤٧ — ٤٨

ومن الأمثلة على المتغيرات المتصلة تغير درجة الحرارة والزمن والطول والسن. فإذا ارتفعت درجة الحرارة من ٣٧° إلى ٤٠° فلا بد أن تمر بجميع القيم الواقعة بين هاتين الدرجتين . كذلك إذا زاد الطول من ١٦٠ سنتيمتراً إلى ١٧٠ سنتيمتراً فلا بد أن يكون قد زاد تدريجاً مارا بجميع القيم الواقعة فى هذه الفترة ، إذ لا يعقد أن يزيد الطول طفرة واحدة .

أما المتغير المنفصل فيتغير فى طفرات أو « قفزات » قد تكون كبيرة أو صغيرة ولا تشغل القيم التى يتخذها المتغير كل المدى الممتد بين نهايتى[1] الأعداد الحقيقية أو أية فترة واقعة بينهما . ومن الأمثلة على المتغيرات المنفصلة تغير الثمن والطلب والربح والتكاليف أو أى شىء لا يمكن التعبير عنه بدلالة أرقام كسرية كعدد الرجال أو عدد الحيوانات أو عدد السفن أو عدد أى شىء آخر لا يتغير إلا فى وحدات صحيحة . والسبب فى ذلك أن الأرقام الكسرية تشغل مكاناً فى مدى الأعداد الحقيقية . والظاهرة التى لا يمكن التعبير عنها بأرقام كسرية « تقفز » من عدد صحيح إلى آخر خلال هذا المدى متجاوزة الفراغ الموجود بين كل عدد صحيح وآخر والذى يشغله عدد لا نهاية له من الأرقام الكسرية .

ويلاحظ أننا نلجأ أحياناً إلى التعبير بيانياً عن بعض المتغيرات المنفصلة بمنحنيات متصلة لتسهيل التحليل النظرى . ومن أوضح الأمثلة على ذلك تمثيل العلاقة بين الطلب والثمن بمنحن متصل وكلاهما من المتغيرات المنفصلة . والواقع أنه ليس هناك أى سبب فنى يمنعنا من ذلك . فكثير من المتغيرات التى نعتبرها منفصلة يمكن بشىء من التصور اعتبارها متصلة . فثمن السلعة

---

(١) يلاحظ أننا نستعمل هنا عبارة « نهايتى الأعداد الحقيقية » من قبيل التجاوز فى التعبير . فالمفروض أن الأعداد الحقيقية تمتد إلى « ما لا نهاية » فى الاتجاهين الموجب والسالب . غير أن ذكر المدى يتضمن تصور وجود نهايتين يمتد بينهما هذا المدى حتى وإن كانت هاتين النهايتين (+ ∞ ، − ∞) .

مثلا لا يمكن التعبير عن تغيره إلا بدلالة الوحدات الصحيحة ولتكن الوحدة المليم أو نصف المليم . غير أنه ليس ثمة سبب يمنعنا من تصور وجود وحدات أصغر من ذلك بكثير . وإذا كانت العبرة بقيمة هذا المليم ، فهى الان لا تزيد عن قيمة كسر ضئيل من قيمة المليم فى سنة ١٩٠٠ . وقد تصل إلى أقل من ذلك فى المستقبل . والذى نود أن نستخلصه من ذلك أنه إذا كانت المناسبة العملية تجعلنا لا نتعامل بوحدات أقل من المليم فإن تجزئة الثمن إلى وحدات أقل من ذلك بكثير صحيح من الوجهة النظرية . كذلك يمكن تصور تغير الطلب بوحدات متناهية فى الصغر . وعلى ذلك فليس هناك ما يمنع من تمثيل العلاقة بين الثمن والطلب بمنحن متصل . وقياساً على ذلك يمكن أحياناً تمثيل بعض التوزيعات التكرارية التى تمثل متغيرات منفصلة بمنحنيات متصلة . غير أنه يجب توخى الحذر الشديد فى استعمال النتائج التى قد نصل إليها بهذه الطريقة .

التوزيع التكرارى المتجمع (Cumulative Frequency Distribution.) .

فى بعض حالات العرض الإحصائى نجد من المناسب أن نعرض البيانات التى يتضمنها التوزيع التكرارى فى شكل مخالف للشكل المعتاد . فقد يكون اهتمامنا محصوراً فى معرفة عدد المفردات الأقل (أو الأكبر) من قيمة معينة وليس عدد المفردات الواقعة بين حدود الفئات . فقد نرغب مثلا فى معرفة عدد ملاك الأراضى الزراعية الذين تقل ملكيتهم عن خمسة أفدنة ، وهؤلاء الذين تقل ملكيتهم عن عشرة أفدنة أو عشرين فداناً وهكذا . وفى هذه الحالة إذا كان لدينا توزيع تكرارى عادى يمثل توزيع ملكية الأراضى الزراعية فيمكن تعديله بضم مفردات كل فئة إلى جميع المفردات السابقة لها كى نحصل على عدد المفردات التى تقل جميعها عن الحد الأعلى للفئة . وتتضح هذه الطريقة من مقارنة الجدولين الآتيين :—

جدول رقم ١٠

توزيع تكرارى عادى لبعض فئات
الملكية الزراعية فى مصر
١٩٤٢

فئات الملكية	عدد الملاك
أقل من ١ فدان	١,٧٧٩,٧٨٠
١ — ٥	٥٨٢,٨٦٩
٥ — ١٠	٨٦,٣٦٤
١٠ — ٢٠	٤١,١٢٩
٢٠ — ٣٠	١١,٦٤٥
٣٠ — ٥٠	٩,٣٢٣
٥٠ — ١٠٠	٦,٧٩٥

جدول رقم ١١

توزيع تكرارى متجمع صاعد
لبعض فئات الملكية الزراعية فى مصر
١٩٤٢

الحد الأعلى للفئة	عدد الملاك
أقل من ١ فدان	١,٧٧٩,٧٨٠
» ٥	٢,٣٦٢,٦٤٩
» ١٠	٢,٤٤٩,٠١٣
» ٢٠	٢,٤٩٠,١٤٢
» ٣٠	٢,٥٠١,٧٨٧
» ٥٠	٢,٥١١,١١٠
» ١٠٠	٢,٥١٧,٩٠٥

المصدر : بيان معالى وزير المالية عن مشروع ميزانية الدولة لسنة ١٩٤٥ — ١٩٤٦[1]

يلاحظ من جدول التوزيع التكرارى المتجمع أنه لكى نحصل على عدد الملاك الذين يملك كل منهم أقل من خمسة أفدنة ، جمعنا عدد الملاك فى الفئة الثانية على عدد الملاك فى الفئة الأولى (فى جدول التوزيع العادى) . ولكى نحصل على عدد الملاك الذين يملك كل منهم أقل من عشرة أفدنة ، جمعنا عدد الملاك فى الفئة الثالثة على عدد الملاك فى الفئتين السابقتين وهكذا . وحيث أن عدد الملاك يتصاعد باستمرار عندما نضع التوزيع فى هذه الصورة

---

(١) يلاحظ أننا لم نجد حاجة إلى بيان جميع فئات الملكية الواردة بتقرير معالى وزير المالية .

فإنه يسمى فى هذه الحالة التوزيع التكرارى المتجمع الصاعد .

ويمكن عمل توزيع تكرارى متجمع فى صورة مخالفة للسابقة . وذلك بأن نسجل أولا عدد المفردات الأكبر من الحد الأدنى للفئة الأولى أي المجموع الكلى للتكرارات ، ثم نطرح من هذا المجموع عدد المفردات الواقعة فى الفئة الأولى وبذلك نحصل على التكرارات الأكبر من الحد الأدنى للفئة الثانية ثم نطرح بعد ذلك من المجموع ذاته عدد المفردات الواقعة فى الفئتين الأولى والثانية معا فنحصل على التكرارات الأكبر من الحد الأدنى للفئة الثالثة وهكذا . وتتضح هذه الطريقة من مقارنة الجدولين الآتيين :—

جدول رقم ١٢

توزيع تكرارى عادى لأعمار الطلبة المتقدمين للالتحاق بكلية التجارة سنة ١٩٤٧

فئات الأعمار	عدد الطلبة
١٦ — ١٨	٢٠
١٨ — ٢٠	٥٥
٢٠ — ٢٢	٦٥
٢٢ — ٢٤	٣٢
٢٤ — ٢٦	٨
٢٦ — ٢٨	٥
٢٨ — ٣٠	٢
٣٠ — ٣٢	٦
٣٢ — ٣٤	٣
٣٤ — ٣٦	٤
المجموع	٢٠٠

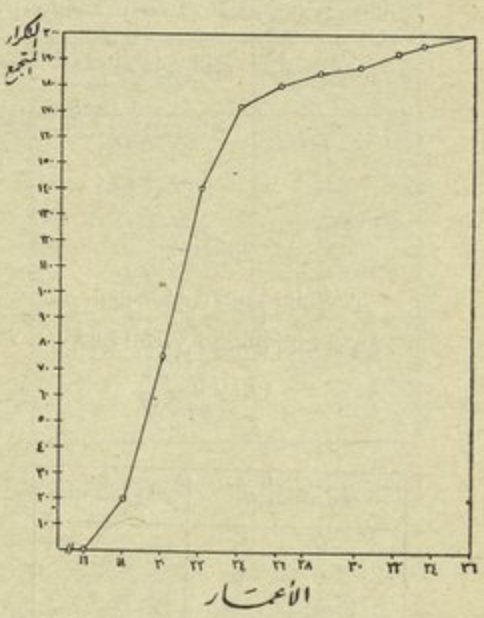
جدول رقم ١٣

توزيع تكرارى متجمع نازل لأعمار الطلبة المتقدمين للالتحاق بكلية التجارة سنة ١٩٤٧

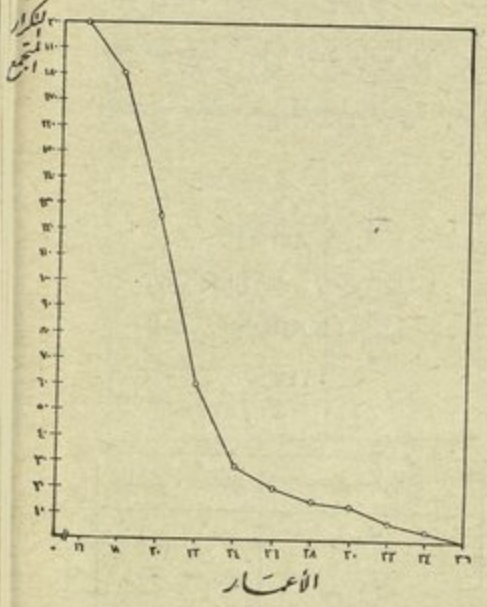
الحد الأدنى للفئة	عدد الطلبة
أكبر من ١٦	٢٠٠
» ١٨	١٨٠
» ٢٠	١٢٥
» ٢٢	٦٠
» ٢٣	٢٨
» ٢٦	٢٠
» ٢٨	١٥
» ٣٠	١٣
» ٣٢	٧
» ٣٤	٤
» ٣٦	٠

وحيث أن عدد الطلبة يتناقص باستمرار عندما نضع التوزيع فى هذه الصورة فإنه يسمى التوزيع التكرارى المتجمع النازل .

ويمكن توضيح التوزيع التكرارى المتجمع الصاعد والنازل بيانياً على النحو المبين فى الشكلين الآتيين :—



(شكل رقم ٣٧)
المنحنى التكرارى المتجمع الصاعد
لأعمار الطلبة المتقدمين للالتحاق بكلية التجارة
سنة ١٩٤٧



(شكل رقم ٣٨)
المنحنى التكرارى المتجمع النازل
لأعمار الطلبة المتقدمين للالتحاق بكلية التجارة
سنة ١٩٤٧

ولكل من المنحنيين التكراريين المتجمعين فوائد كثيرة . فسيجد القارئ أننا سنستعملها فى تعيين قيم بعض المقاييس الإحصائية . وإذا رصد المنحنيان على نفس الرسم فإن نقطة تقاطعهما تعتبر ذات أهمية خاصة ، إذ أنها تحدد لنا قيمة أحد مقاييس التوسط الهامة كما سيتضح من الباب التالى .

منحنى لورنز :

تستخدم فكرة التوزيع التكرارى المتجمع فى رسم منحنى لورنز الذى يبين درجة عدم المساواة فى توزيع الدخل أو الملكية أو بعبارة أخرى مدى انحراف التوزيع الفعلى للدخل أو الملكية عن التوزيع المثالى . والتوزيع المثالى هو ذلك الذى يحقق التناسب التام بين عدد مكتسبى الدخل أو عدد الملاك فى كل فئة وبين مجمل الدخول التى يكتسبونها أو مساحة الأراضى التى يمتلكونها . أى أن التوزيع المثالى للملكية الزراعية مثلا يتحقق إذا كان :

١٠٪ من عدد الملاك يمتلكون ١٠٪ من مساحة الأرض
٢٠٪ « « « « ٢٠٪ « « «
٣٠٪ « « « « ٣٠٪ « « «

أو بوجه عام :

س٪ « « « « س٪ « « «

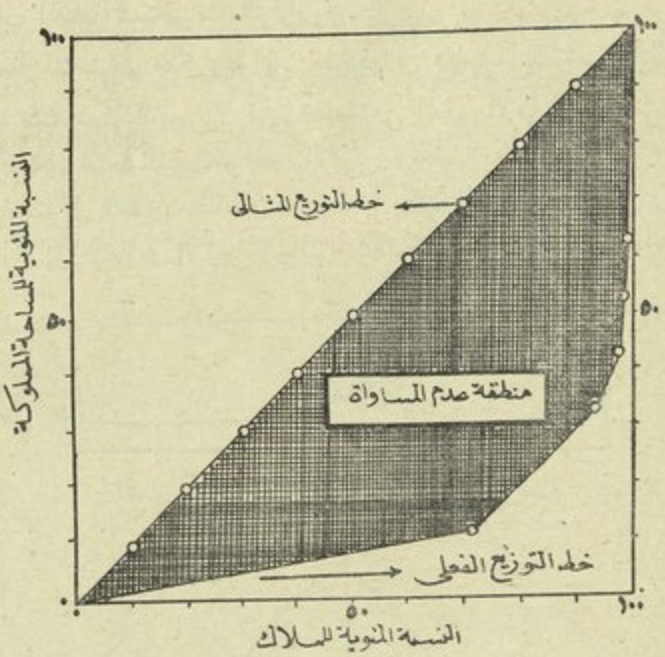
أما التوزيع الفعلى للملكية فإنه يختلف عن ذلك كثيراً . فإذا أخذنا على سبيل التمثيل توزيع الملكية الزراعية فى مصر طبقاً لإحصاءات سنة ١٩٤٢ نجد أنها كالاتى :—

فئات الملكية	عدد الملاك ٪	المساحة المملوكة ٪
أقل من فدان	٧٠,٥	١٢,٤
١ — ٥	٢٣,١	٢٠,٦
٥ — ١٠	٣,٤٥	٩,٧
١٠ — ٢٠	١,٦٥	٩,٥
٢٠ — ٥٠	٠,٨	١٠,٩
٥٠ فأكثر	٠,٥	٣٦,٩

ولكى يمكن مقارنة هذا التوزيع الفعلى بالتوزيع المثالى الذى سبق بيانه نضع الأرقام الخاصة بعد الملاك وتلك الخاصة بالمساحة المملوكة فى شكل توزيع تكرارى متجمع صاعد كما يأتى :—

فئات الملكية	عدد الملاك ٪	المساحة المملوكة ٪
أقل من فدان	٧٠,٥	١٢,٤
« « ٥	٩٣,٦	٣٣,٠
« « ١٠	٩٧,٠٥	٤٢,٧
« « ٢٠	٩٨,٧	٥٢,٢
« « ٥٠	٩٩,٥	٦٣,١
جميع الملاك	١٠٠	١٠٠

ولرسم المنحنى من واقع البيانات السابقة نقسم كلا من المحورين إلى ١٠ أقسام متساوية مستعملين نفس مقياس الرسم فى الحالتين ، ثم نكمل المربع بعمل محورين إضافين مناظريين للمحورين الأصليين . بعد ذلك نرصد خط التوزيع المثالى فنجد أنه عبارة عن قطر المربع الذى يمر بنقطة الأصل ثم نرصد خط التوزيع الفعلى وهو عبارة عن الخط الذى يصل بين النقط التى يكون الإحداثى السينى لها عبارة عن النسبة المئوية لعدد الملاك فى كل فئة والإحداثى الصادى لها عبارة عن النسبة المئوية للمساحة المملوكة فى فى نفس الفئة . ويستحسن أن تظلل المساحة الواقعة بين خط التوزيع الفعلى وخط التوزيع المثالى والتى يطلق عليها اسم « منطقة عدم المساواة » وذلك لإبرازها فى صورة واضحة . والرسم الاتى يوضح منحنى لورنز طبقاً لإحصاءات الملكية الزراعية فى سنة ١٩٤٢ .



(شكل رقم ٣٩)

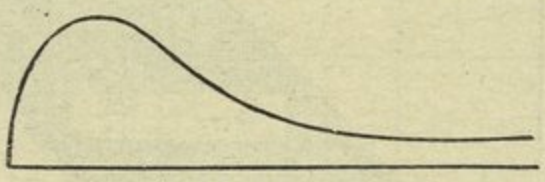
منحنى لورنز من واقع إحصاءات الملكية الزراعية فى مصر سنة ١٩٤٢

ومما تجب ملاحظته أنه إذا لم ترد الأرقام التى تمثل عدد ملاك كل فئة والمساحة التى يملكونها فى شكل نسب مئوية فيجب حساب هذه النسب المئوية أولا ثم اتباع الخطوات السالفة الذكر .

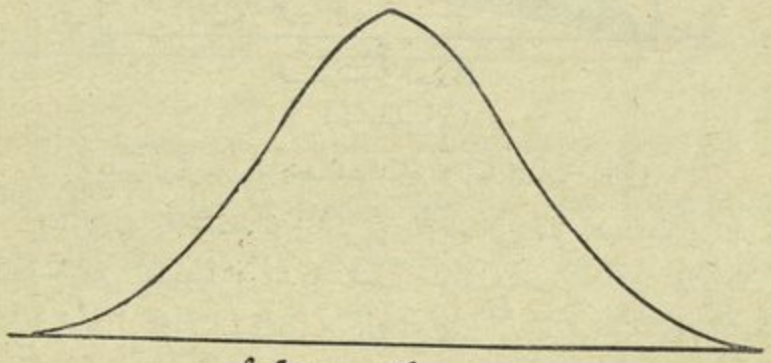
ومن الجدير بالذكر أنه بالرغم من أن منحنى لورنز لا يطبق عادة إلا لغرض مقارنة التوزيع الفعلى للدخل أو الملكية بالتوزيع المثالى إلا أنه يمكن تعميم تطبيقه لتوضيح ظاهرة التركز بوجه عام . فيمكن مثلا استخدامه لوصف درجة تركز العمل أو رأس المال فى الصناعة وذلك بأن نقسم المصانع إلى فئات ونحسب النسبة المئوية لعدد العمال أو مقدار رأس المال فى كل فئة ونقارنها بالنسبة المئوية لعدد المصانع فى نفس الفئة ومن الواضح أنه كلما اتسعت منطقة عدم المساواة كلما دل ذلك على وجود ظاهرة التركز ، أى تركز عدد كبير من العمال أو مقادير كبيرة من رؤوس الأموال فى عدد قليل من المصانع .

الأشكال المختلفة للمنحنيات التكرارية .

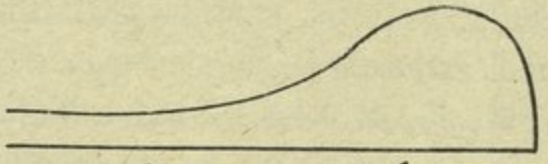
معظم المنحنيات التكرارية التى نصادفها فى دراستنا للظواهر المختلفة محدبة وشكلها يشبه شكل الناقوس . فهى تصعد من المحور الأفقى حتى تصل إلى حد أقصى ثم تعود فتهبط ثانية نحو المحور الأفقى . وبعضها يكون متماثلا ، أى له محور تماثل يقسمه إلى قسمين متطابقين ، وبعضها يكون ملتوياً إما إلى اليمين فى الاتجاه الموجب أو إلى اليسار فى الاتجاه السالب . والشكل الآتى يبين هذه الأنواع الثلاثة.



منحنى تكرارى ملتوى التواءً موجبا



منحنى تكرارى متماثل

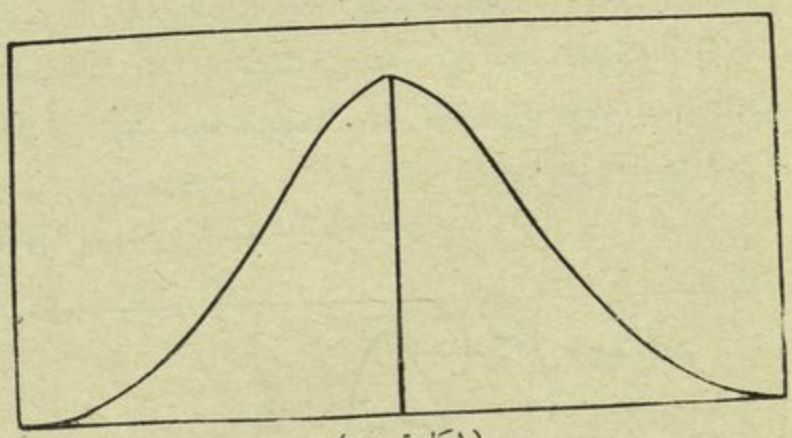


منحنى تكرارى ملتوى التواءً سالبا

(شكل رقم ٤٠)

بعض الأشكال المختلفة للمنحنيات التكرارية توضح التماثل والالتواء

وعلى وجه العموم فإننا نصادف المنحنيات القريبة من التماثل عادة فى دراسة الظواهر الطبيعية والبيولوجية كالطول والوزن ومحيط الرقبة . . . إلخ ، وأهم المنحنيات المتماثلة هو المنحنى الطبيعى أو منحنى الخطأ المتماثل (Normal Curve of Error) الذى يصف كثيراً من الظواهر الطبيعية والذى سيأتى الكلام عنه فيما بعد (انظر شكل ٤١) .

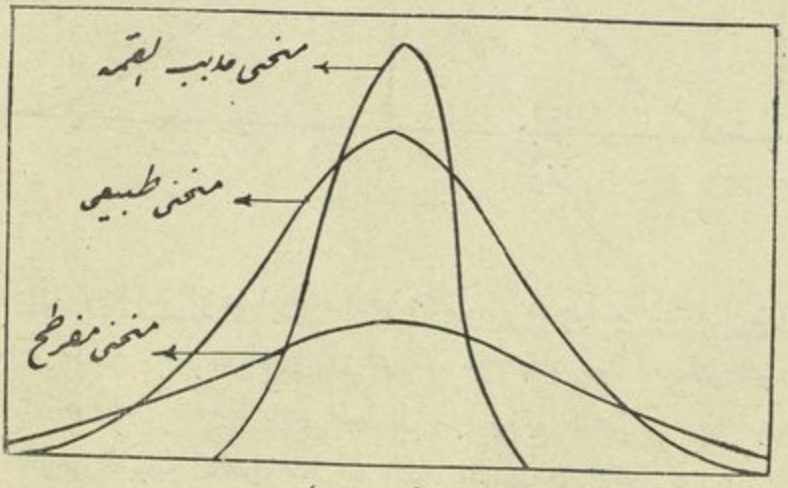


(شكل رقم ٤١)
منحنى الخطأ المتماثل

وفى الدراسات الإقتصادية نصادف كثيراً من المنحنيات الملتوية . فإذا رصدنا التوزيع التكرارى للدخل فى أي بلد من البلاد نجد أنه ملتو التواء ظاهراً نحو اليمين . وهذا يرجع طبعاً إلى سوء توزيع الدخول . فبينما نجد فى الظواهر الطبيعية كالطول مثلا أن الطول المتوسط هو الأكثر شيوعاً بين الناس وأن عدد الأشخاص الذين تزيد أطوالهم عن هذا المتوسط يعادل تقريبا عدد الأشخاص الذين تقل أطوالهم عنه ، نجد عند دراسة توزيع الدخول أن الدخل المتوسط ليس هو الدخل الأكثر شيوعاً بين الناس كما نجد أن عدد الأشخاص الذين يحصلون على دخل أقل من المتوسط أكبر بكثير جداً من عدد الأشخاص الذين يحصلون على دخل أكبر من المتوسط . والواقع أن هذه

الأقلية من الدخول الكبيرة هى التى تسبب رفع قيمة المتوسط إلى الحد الذى يجعله لا يمثل أغلبية الدخول تمثيلا صادقاً .

وبالإضافة إلى التماثل والالتواء يمكننا أن نقسم المنحنيات التكرارية من حيث تدبب قمتها أو انبساطها أو توسطها بين هذين الوضعين ، إلى ثلاثة أقسام : القسم الأول يشمل المنحنيات ذات القمم المدببة وتسمى (Leptokurtic) والقسم الثانى يشمل المنحنيات ذات القمم المنبسطة أو المفرطحة (platykurtic) والقسم الثالث يشمل المنحنيات التى تتوسط بين التدبب والتفرطح أو بعبارة أدق تلك التى تكون درجة تفرطحها معادلة تماماً لدرجة تفرطح المنحنى الطبيعى وتسمى (Mesokurtic) . والشكل الاتى يوضح المنحنى الطبيعى ومعه على نفس الرسم منحنى مدبب القمة وآخر مفرطح .



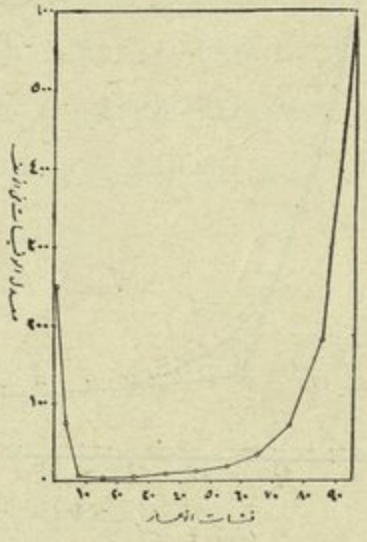
(شكل رقم ٤٢)

بعض الأشكال المختلفة للمنحنيات التكرارية توضح التدبب والتفرطح

---

(١) لم أطلع حتى الآن على مصطلحات عربية تناظر هذه الكلمات . غير أن كلمة KURTIC معناها محدب أو مقوس الظهر وكلمة PLATY تصف الشيء الضيق النحيل . وكلمة LEPTO معناها الانبساط والتفرطح . وكلمة MESO تفيد الاعتدال والتوسط ولذلك فإنى أقترح تسمية النوع الأول المنحنيات المدببة والثانى المنحنيات المنبسطة أو المفرطحة والثالث المنحنيات المعتدلة التفرطح . لشرح معانى هذه الكلمات انظر كتاب Applied General Statistics by Croxton and Cowden ص٢٥٨ .

وهناك بعض المنحنيات التكرارية التى يختلف شكلها كلية عن جميع الأشكال السابقة فهى ليست محدبة أو ناقوسية الشكل . فقد يصادفنا أحياناً توزيعات تكرارية إذا رصدت على الرسم البيانى تظهر على شكل منحن مقعر يشبه حرف (U) . فمثلا إذا رصدنا التوزيع التكرارى للوفيات حسب فئات الأعمار فى سنة ١٩٤٢ على الرسم البيانى نحصل على منحن ينحدر إلى أسفل ثم يصعد ثانية على شكل حرف (U) كما هو موضح بالرسم الآتى :-



(شكل رقم ٤٣)

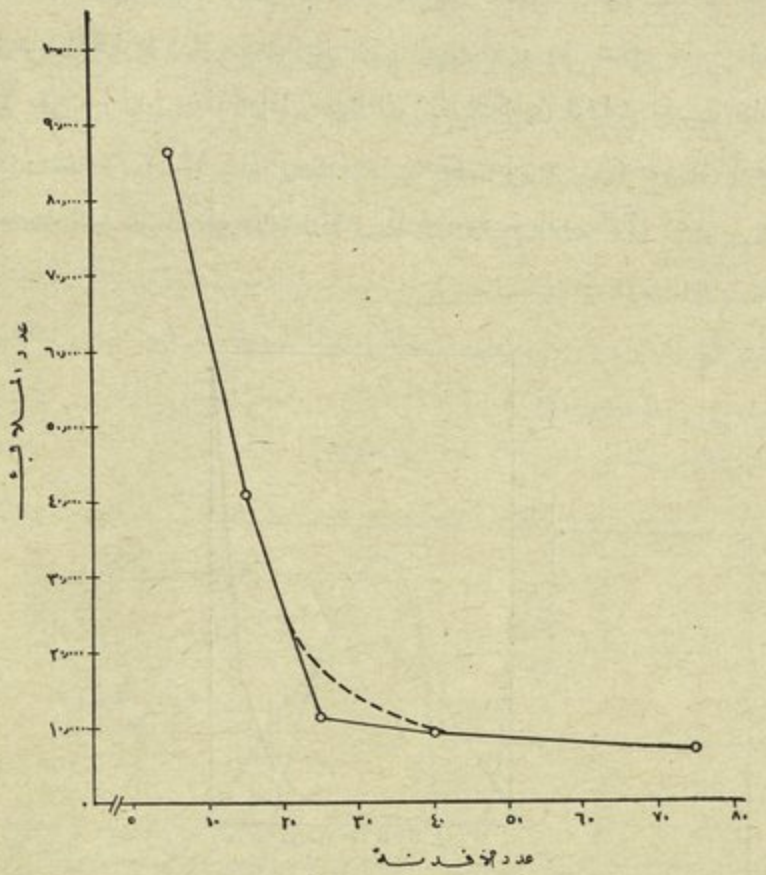
الوفيات حسب السن فى القطر المصرى سنة ١٩٤٢ (٥)

وفى بعض الحالات يصادفنا توزيع تكرارى إذا رصد على الرسم البيانى يظهر على شكل منحن ذى فرع واحد . فلو رصدنا التوزيع التكرارى للملكية

---

(٥) يمكن للقارئ الاطلاع على الأرقام التى استعملت فى هذا الرسم بالرجوع إلى الإحصاء السنوى للجيب لسنة ١٩٤٦ جدول رقم ٥ ص ٣٠ .

الزراعية فى مصر المبين بالجدول رقم ١٠ نحصل على الشكل الآتى :—



(شكل رقم ٤٤)

منحنى تكرارى ذو فرع واحد يمثل توزيع الملكية الزراعية بمصر (٥)

سنة ١٩٤٢

وصف التوزيعات التكرارية .

يمكن وصف التوزيعات التكرارية وتحديد أشكال المنحنيات التى تمثلها

---

(٥) يلاحظ أننا لم نرصد جميع فئات التوزيع لاستحالة ذلك عمليا على ورق الرسم العادى نظراً لأن عدد الملاك يتغير فى مدى واسع جداً . ولذلك اكتفينا ببيان فئات الملكية من ٥ — ١٠٠ فدان

والتمييز بينها عن طريق المقاييس الإحصائية الوصفية التى تعطينا قيم ثوابت (Constants) هذه التوزيعات أو المعالم المميزة لها . وأول هذه المقاييس الوصفية هى مجموعة مقاييس التوسط أو المتوسطات ، وهى التى تعين لنا قيمة الثابت الأول أو بعبارة أخرى تعطينا وصفاً أوليا للتوزيع . والمجموعة الثانية هى مجموعة مقاييس التشتت التى تنتقل بنا خطوة أخرى نحو وصف التوزيع فتحدد لنا قيمة الثابت الثانى . والمجموعة الثالثة هى مقاييس الالتواء التى تساعدنا على وصف التوزيع بتفصيل أكثر إذا ما أضيفت إلى المقاييس السابقة . وأخيراً هناك مقاييس التفرطح التى تتم لنا وصف التوزيع أو بالأحرى المنحنى الذى يمثله . فإذا ما حددنا هذه المقاييس الأربعة لمنحن تكرارى معين نكون قد وصفناه وصفاً دقيقاً كاملا يميزه عن المنحنيات الأخرى . وسنتكلم بشىء من التفصيل عن هذه المقاييس الوصفية بنفس هذا الترتيب فى الأبواب الثلاثة التالية .

الباب الخامس

مقاييس التوسط

رأينا فى الباب السابق كيف أن هناك أشكالا مختلفة للمنحنيات التكرارية وانتهينا إلى أنه يمكن تحديد شكل المنحنى وبالتالى وصف التوزيع التكرارى للمتغير عن طريق تحديد قيم ثوابت هذا التوزيع . ومقاييس التوسط هى المجموعة الأولى من المقاييس الإحصائية الوصفية ، وهى أهمها وأكثرها استعمالا وأقربها إلى الفهم . والمتوسط عبارة عن قيمة نموذجية نختارها أحياناً من بين المفردات التى تشملها المجموعة الموضوعة تحت البحث لتمثيل هذه المجموعة ، وأحياناً وهو الأغلب ، نلجأ إلى حسابها من واقع قيم مفردات المجموعة بطرق مختلفة سنخصص الجزء الأكبر من هذا الباب لشرحها . ومن هذا يتبين لنا أنه ليس من الضرورى أن تكون قيمة المتوسط هى إحدى القيم التى تتخذها الظاهرة فعلا فى تغيرها . فقد نقول إن متوسط أجر سائق الترام فى مدينة الإسكندرية هو ٢٣ ١/٢ قرشاً فى اليوم فى حين أنه ليس هناك سائق واحد يتقاضى هذا الأجر بالذات . وقد نقول إن متوسط ما تحمله عربة الترام من الركاب هو ٦٢ ١/٢ راكباً فى حين أنه لا يعقل طبعاً أن تحمل أى عربة هذا العدد من الركاب . غير أن صفة تمثيل المجموعة قد تكون متوفرة فى هذا المتوسط أكثر من توفرها فى أية قيمة أخرى من القيم التى تتخذها المفردات ذاتها . وبديهى أنه لكى يكون هذا التمثيل صادقاً يجب أن يحمل المتوسط صفات مفردات المجموعة التى ينوب عنها أو على الأقل صفات العدد الأكبر منها . وبالإضافة إلى ذلك فهناك عدة شروط أخرى يجب أن تتوفر فى المتوسط

باعتباره ممثلا للمجموعة ، وقد لخصها الأستاذ (Waugh) فيما يلى[١] :—

(١) يجب أن يحسب المتوسط على أساس قواعد محددة بدقة ولا يترك للتقدير الشخصى للباحث . وبعبارة أخرى يجب أن تكون قيمة المتوسط موضوعية بحيث لا يختلف إثنان فى تقديرها .

(٢) يجب أن يحسب المتوسط على أساس جميع القيم المشاهدة للظاهرة وإلا أصبح تمثيله غير دقيق .

(٣) يجب أن يكون للمتوسط خواص مميزة ملموسة تجعل من السهل تعريفه ، كما يجب ألا يبنى على فكرة رياضية معقدة .

(٤) يحسن بقدر الإمكان أن يكون المتوسط من القيم التى يمكن حسابها بسرعة وسهولة . غير أن هذا لا يعنى أن نضحى بدرجة كبيرة من الدقة فى سبيل سهولة العمل .

(٥) يحسن أن يكون المتوسط من النوع الذى لا تتأثر قيمته كثيراً بمصادفة اختيار العينة . ولتوضيح ذلك نذكر أنه إذا اخترنا عدة عينات اعتباطية من مجموع واحد وحسبنا قيم متوسطاتها فإن بعض المتوسطات الآتى ذكرها يعطينا قيما متقاربة ، والبعض الاخر يعطينا قيما متفاوتة ، ولا شك أن النوع الأول هو الأفضل .

(٦) من أهم الشروط الواجب توفرها فى المتوسط هو أن يكون من السهل معالجته جبرياً . وستتضح أهمية هذه النقطة عند شرح خواص الوسط الحسابى ، ونجمل فيما يلى المتوسطات المختلفة ، ثم ننتقل إلى شرح كل منها تفصيلا :—

(١) الوسط الحسابى Arithmetic Mean

---

(١) راجع كتاب Elements of Statistics تأليف A. E. Waugh ص ١٠٧ .

(٢) الوسط الهندسى Geometric Mean

(٣) الوسط التوافقى Harmonic Mean

(٤) الوسط التربيعى Quadratic Mean

(٥) الوسيط Median

(٦) المنوال Mode

١ ـ الوسط الحسابى

الوسط الحسابى هو أكثر مقاييس التوسط شيوعاً وأقربها إلى تحقيق الشروط السالفة الذكر ، كما أنه يعتبر أهمها من وجهة الدراسة الإحصائية النظرية والتطبيق العملى على السواء ، وكثيراً ما نشير إلى « المتوسط » دون تحديد قاصدين بذلك الوسط الحسابى . ويمكن الحصول على الوسط الحسابى لمجموعة من القيم بجمع هذه القيم وقسمة الناتج على عددها . فإذا كانت هذه القيم هى :

$$س_1 ، س_2 ، س_3 ..... س_ن$$

$$\text{فإن الوسط الحسابى} = \frac{س_1 + س_2 + س_3 + ... س_ن}{ن}$$

ويرمز عادة للوسط الحسابى بالرمز $\bar{س}$ كما يرمز لمجموع قيم س بالرمز $\Sigma س$[1] وعلى ذلك يكون الوسط الحسابى $= \frac{\Sigma س}{ن}$

---

(١) سنستعمل حرف Σ فى هذا الكتاب للدلالة على جمع قيم المتغيرات فمثلا

$$\Sigma س = س_1 + س_2 + س_3 + ...س_ن$$

وإذا أردنا أن نجمع قيم المتغير فى مدى محدود فيمكن إيضاح ذلك بوضع النهاية الصغرى للمدى تحت علامة الجمع والنهاية العظمى له فوق هذه العلامة هكذا

$$\sum_{1}^{4} س = س_1 + س_2 + س_3 + س_4 .$$

فلو فرض مثلا أن أسعار أقة اللبن فى خمسة أسواق مختلفة بمدينة الإسكندرية هى :-

٥ و ٦ و ٧ و ٨ و ٩ قرش على التوالى ، فإن الوسط الحسابى للأسعار $= \frac{٥+٦+٧+٨+٩}{٥} = ٧$ قروش للأقة ، غير أنه لكى تكون النتيجة التى نحصل عليها صحيحة ومنطقية ، يجب أن يكون السعر المتوسط لأقة اللبن يعادل مجمل ما دفعه المستهلكون لحيازة جميع أقق اللبن المبيعة فى هذه الأسواق مقسوماً على عدد هذه الأقق . ولاختبار صحة الرقم الذى حصلنا عليه ، نفرض أن الكميات المبيعة فى هذه الأسواق الخمسة فى وحدة زمنية معينة هى :- ١٠,٠٠٠ و ٥٠٠٠ و ٤٠٠٠ و ٣٠٠٠ و ٢٠٠٠ أقة على التوالى .

على هذا الأساس يكون مجمل ما أنفقه المستهلكون على اللبن فى هذه الوحدة الزمنية هو :-

$١٠٠٠٠ \times ٥ + ٥٠٠٠ \times ٦ + ٤٠٠٠ \times ٧ + ٣٠٠٠ \times ٨ + ٢٠٠٠ \times ٩$

$= ١٥٠,٠٠٠$ وبقسمة هذا المبلغ على عدد الأقق المبيعة فى جميع الأسواق يكون السعر المتوسط للأقة $= \frac{١٥٠,٠٠٠}{٢٤,٠٠٠} = ٦\frac{١}{٤}$ قرش

ومن هذا يتضح أن هناك اختلافاً بيناً بين النتيجتين . ومهمتنا الآن هى معرفة أى النتيجتين هى الصحيحة ؟

الواقع أن النتيجة الأخيرة هى الصحيحة ، أما النتيجة الأولى فلا تكون صحيحة إلا إذا كان عدد الأقق المبيعة فى جميع الأسواق واحداً . وذلك لأنه من تعريف الوسط الحسابى نجد أن السعر المتوسط $= \frac{\text{مجموع الأسعار}}{\text{عدد الأسعار (أو عدد الأقق المبيعة)}}$

وعلى الرغم من تساوى الأسعار فى السوق الواحدة إلا أنه لغرض حساب الوسط الحسابى بدقة يجب أن نعتبر بيع كل أقة عملية قائمة بذاتها بسعر مستقل . وعلى هذا الأساس يكون لدينا فى هذا المثال ٢٤٠٠٠ عملية لكل منها سعر

مستقل ويكون مجموع الأسعار فى جميع الأسواق :—

(٥ + ٥ + ٠٠٠٥ ، إلى ١٠,٠٠٠ مرة) + (٦ + ٦ + ٠٠٠٦ ، إلى ٥٠٠٠ مرة) + (٧ + ٧ + ٧ ٠٠٠ ، إلى ٤٠٠٠ مرة) + (٨ + ٨ + ٨ + ٠٠٠ ، إلى ٣٠٠٠ مرة) + (٩ + ٩ + ٩ ٠٠٠ ، إلى ٢٠٠٠ مرة)

وبوضع هذه الأرقام فى صورة أسهل وأكثر اختصارا نقول :—

مجموع الأسعار التى حصل تعامل بها فى جميع الأسواق

= ٥ × ١٠,٠٠٠ + ٦ × ٥٠٠٠ + ٧ × ٤٠٠٠ + ٨ × ٣٠٠٠ + ٩ × ٢٠٠٠ = ١٥٠,٠٠٠ وبالقسمة على عدد الأسعار « الذى هو عدد الأقق المبيعة » يكون متوسط السعر :

$$= \frac{١٥٠,٠٠٠}{٢٤,٠٠٠} = ٦\frac{١}{٤} \text{ قرش}$$

فكان التحليل الدقيق يظهر لنا أن النتيجة الثانية هى النتيجة الصحيحة .

أما إذا كانت الكميات المبيعة فى كل من الأسواق المختلفة متساوية ، فإن الطريقتين تؤديان إلى نفس النتيجة . ولإثبات ذلك إفرض أن الكمية المبيعة فى كل من الأسواق الخمسة هى ١٠٠٠ أقة ، فيكون الوسط الحسابى

$$= \frac{١٠٠٠ \times ٥ + ١٠٠٠ \times ٦ + ١٠٠٠ \times ٧ + ١٠٠٠ \times ٨ + ١٠٠٠ \times ٩}{١٠٠٠ + ١٠٠٠ + ١٠٠٠ + ١٠٠٠ + ١٠٠٠}$$

$$= \frac{١٠٠٠ (٥ + ٦ + ٧ + ٨ + ٩)}{٥٠٠٠} = \frac{ص}{٧}$$

وهى نفس النتيجة التى حصلنا عليها بالطريقة الأولى .

والسبب فى اتفاق نتيجتى الطريقتين إذا كان عدد الأقق المبيعة فى جميع الأسواق واحداً هو أننا نعطى كل سعر نفس الأهمية ، أى أننا لا نرجح أى سعر على الآخر . والخلاصة أنه ما دامت الكمية المبيعة فى جميع الأسواق

واحدة ، فلا يهمنا – لغرض تعيين الوسط الحسابى للأسعار – ما إذا كانت هذه الكمية أقة واحدة أو ألف أقة .

أما إذا اختلفت الكميات المبيعة بكل سعر فإن أهمية كل سعر تختلف تبعاً لذلك ويجب أن نرجح كل سعر حسب أهميته عند تعيين قيمة الوسط الحسابى . وتسمى هذه الطريقة ، طريقة الترجيح بالأوزان (أى بالأهمية) ويسمى المتوسط المستخرج بهذه الطريقة ، الوسط الحسابى المرجح بالأوزان . ولزيادة إيضاح فكرة الترجيح بالأوزان نسوق المثال الاتى :–

إفرض أننا أردنا أن نحسب متوسط التغير النسبى فى أسعار بعض الضروريات وأنه بعد جمع البيانات الإحصائية اللازمة وجدنا أن النسب المئوية لارتفاع الأسعار كانت كالآتى :–

اللحم	الخبز	السكر	الكبريت	الملح
٥٠	٥٠	٦٠	٣٤٠	٣٠٠

$$\text{الوسط الحسابى البسيط} = \frac{٥٠ + ٥٠ + ٦٠ + ٣٤٠ + ٣٠٠}{٥} = ١٦٠$$

ولا شك أنه لو وصفنا إرتفاع أسعار هذه الضروريات بأنه ١٦٠٪ فى المتوسط لكان وصفنا مضللا إلى أبعد الحدود . فمن البديهى أن غالبية المستهلكين لا تهتم بارتفاع سعر الكبريت أو الملح اهتمامها بارتفاع سعر اللحم أو الخبز . فالسلعتان الأخيرتان تستنفدان جزءاً من الدخل يعادل عشرات من المرات ما تستنفده السلعتان الأوليان . وعلى ذلك فالوسط الحسابى البسيط يعطينا صورة خاطئة تماماً عن مقدار ارتفاع الأسعار فهو يتأثر كثيراً بارتفاع سعرى الكبريت والملح مع ضآلة أهميتهما نسبياً . ولكى نتغلب على هذه الصعوبة نعطى لكل سلعة وزناً يتناسب مع أهميتها فى ميزانية المستهلك العادى . فلو

أعطينا اللحم وزناً قدره ٢٠ والخبز وزناً قدره ١٥ والسكر وزناً قدره ٥ وكلا من الملح والكبريت وزناً قدره ١ ، فإن الوسط الحسابى المرجح بالأوزان .

$$= \frac{٥٠ \times ٢٠ + ٥٠ \times ١٥ + ٦٠ \times ٥ + ٣٤٠ \times ١ + ٣٠٠ \times ١}{٢٠ + ١٥ + ٥ + ١ + ١} = ٦٤,٠٥$$

أى أن متوسط التغير النسبى فى أسعار هذه السلع هو ٦٤,٠٥٪ تقريباً . ولا شك أن هذا الرقم أصدق فى تعبيره عن حقيقة حالة إرتفاع الأسعار من الرقم الأول .

ويمكن أن نعبر عن الوسط الحسابى المرجح بالأوزان تعبيراً رمزياً كالآتى :—

$$\bar{س} = \frac{س_١ و_١ + س_٢ و_٢ + س_٣ و_٣ + \cdots\cdots + س_ن و_ن}{و_١ + و_٢ + و_٣ + \cdots\cdots + و_ن} = \frac{\Sigma\, س\, و}{\Sigma\, و}$$

حيث $س_١$ ، $س_٢$ ٠٠٠٠ $س_٣$ ، س هى قيم مفردات المجموعة أو قيم المتغير ، $و_١$ ، $و_٢$ ، $و_٣$ ٠٠٠٠ و هى الأوزان المتناسبة مع أهمية كل مفردة

ويلاحظ أن الوسط الحسابى للتوزيع التكرارى لا يختلف فى شىء عن الوسط الحسابى المرجح بالأوزان . فالتكرارات ما هى فى الواقع إلا أوزان تبين أهمية كل فئة من حيث عدد المفردات الواقعة فيها . وعلى ذلك فإن الوسط الحسابى للتوزيع التكرارى يساوى كذلك $\frac{\Sigma\, س\, و}{\Sigma\, و}$

حيث $س_١$ ، $س_٢$ ، $س_٣$ ٠٠٠٠ س تمثل مراكز الفئات

« $و_١$ ، $و_٢$ ، $و_٣$ ٠٠٠٠ $و_ن$ تمثل تكرارات الفئات .

خواص الوسط الحسابى .

للوسط الحسابى عدة خواص بعضها عظيم الفائدة من الناحية العملية

والبعض الآخر ذو أهمية خاصة من وجهة التحليل النظرى ونورد فيما يلى أهم هذه الخواص :—

الخاصية الأولى .

مجموع انحرافات (فروق) القيم عن وسطها الحسابى يساوى صفر . وهذه الخاصية ذات أهمية كبرى فى إرشادنا إلى الطرق المختصرة لحساب متوسط التوزيع التكرارى . ولإثبات هذه الخاصية ننظر إلى الأرقام الآتية التى سنفرض أن أوزانها أو تكراراتها متساوية :—

٤ ، ٧ ، ٩ ، ١٣ ، ١٧

$$\overline{س} = \frac{٤ + ٧ + ٩ + ١٣ + ١٧}{٥} = ١٠$$

والآن نحسب انحراف كل قية عن الوسط الحسابى مع العلم بأننا دائماً نحدد الإنحراف بأنه القيمة ناقص المتوسط أى (س — $\overline{س}$) وليس العكس .

القيمة	الانحراف عن المتوسط الحسابى
٤	— ٦
٧	— ٣
٩	— ١
١٣	+ ٣
١٧	+ ٧

من الواضح أن مجموع انحرافات القيم عن وسطها الحسابى يساوى صفر . وهذا صحيح مهما كانت هذه القيم . ويمكن إثبات ذلك ببرهان جبرى عام كالآتى :—

إفرض أن قيم المتغير هى س١ ، س٢ ، س٣ ٠٠٠٠ سن
وأن انحرافاتها عن الوسط الحسابى هى ح١ ، ح٢ ، ح٣ ٠٠٠٠ حن
والمطلوب إثباته أن ح١ + ح٢ + ح٣ + ٠٠٠٠ حن = صفر

عندنا $ح_1 = س_1 - \overline{س}$

$ح_2 = س_2 - \overline{س}$

$ح_3 = س_3 - \overline{س}$

$ح_ن = س_ن - \overline{س}$

بجمع الطرفين .

$\therefore \Sigma ح = \Sigma س - ن \overline{س}$

ولكن $\overline{س} = \frac{\Sigma س}{ن}$ .. $ن\overline{س} = \Sigma س$

$\therefore \Sigma ح = ن\overline{س} - ن\overline{س} =$ صفر

لو فرضنا أننا فى المثال السابق حسبنا انحرافات القيم (٤، ٧، ٩، ١٣، ١٧) عن أية قيمة أخرى خلاف الوسط الحسابى لوجدنا أن مجموع هذه الانحرافات لا يساوى صفراً . بل إنه لو كانت هذه القيمة أقل من الوسط الحسابى ، لأدى ذلك إلى زيادة الانحرافات الموجبة عن السالبة كما يتضح من الأرقام الآتية :-

القيمة	الانحراف عن وسط فرضى قيمته ٨
٤	— ٤
٧	— ١
٩	+ ١
١٣	+ ٥
١٧	+ ٩
مجموع الانحرافات	= + ١٠

ولو كان هذا الوسط الفرضى أكبر من الوسط الحسابى لأدى ذلك إلى زيادة الانحرافات السالبة عن الموجبة كما يتضح من الأرقام الآتية :ـ

القيمة	الانحرافات عن وسط فرضى قيمته ١٤
٤	ـ ١٠
٧	ـ ٧
٩	ـ ٥
١٣	ـ ١
١٧	+ ٣
مجموع الانحرافات =	ـ ٢٠

والآن نود أن نكون فكرة عن قيمة هذه الانحرافات عن الوسط الفرضى (سالبة كانت أم موجبة) وعلاقتها بالوسط الحسابى .

رأينا أنه إذا طرحنا الوسط الحسابى من كل قيمة كان مجموع بواقى الطرح (أى الانحرافات عن الوسط الحسابى) = صفر .

ولكن إذا طرحنا قيمة أخرى (وسط فرضى) خلاف الوسط الحسابى من هذه القيم كان مجموع بواقى الطرح (أى الانحرافات عن الوسط الفرضى) = كمية معينة موجبة أو سالبة .

هذه الكمية هى فى الواقع الفرق بين الوسط الحسابى والوسط الفرضى مكرراً عدة مرات حسب عدد القيم (٥ فى المثال المتقدم) .

وعلى ذلك فلو اخترنا وسطاً فرضياً وحسبنا مجموع انحرافات القيم عنه ثم قسمنا هذا المجموع على عدد القيم لكان الناتج هو الفرق بين الوسط الحسابى والوسط الفرضى ومن هذا يمكن معرفة الوسط الحسابى . وبتطبيق هذه فى النتيجة على المثال السابق نجد أن :ـ

مجموع الانحرافات عن وسط فرضى قدره ٨ = + ١٠

∴ الفرق بين الوسط الحسابى والوسط الفرضى = $\frac{١٠}{٥}$ = ٢

∴ الوسط الحسابى = ٨ + ٢ = ١٠ وهى نفس النتيجة التى نحصل عليها بالطريقة العادية أى بقسمة مجموع القيم على عددها . ويمكن الوصول إلى نفس هذه النتيجة باختيار أى وسط فرضى آخر .

على أنه يمكن تعميم العلاقة بين الوسط الحسابى والوسط الفرضى جبرياً كالآتى :

ليكن الوسط الفرضى هو $\overline{س_د}$ وانحرافات القيم عنه هى على التوالى ،

$ح_١ ، ح_٢ ، ح_٣ ..... ح_ن$

يمكن إثبات أن $\overline{س} = \overline{س_د} + \frac{\overline{ح}}{ن}$

حيث أن $\overline{س} = \frac{س_١ + س_٢ + ... + س_ن}{ن}$

∴ بإضافة $\overline{س_د}$ إلى الطرف الأيسر ثم طرحها ثانية .

$$\therefore \overline{س} = \overline{س_د} + \frac{س_١ + س_٢ + ... + س_ن}{ن} - \overline{س_د}$$

$$= \overline{س_د} + \frac{س_١ + س_٢ + ... س_ن - ن\overline{س_د}}{ن}$$

وبتوزيع $ن\overline{س_د}$ على جميع الحدود فى بسط الكسر .

$$\therefore \overline{س} = \overline{س_د} + \frac{(س_١ - \overline{س_د}) + (س_٢ - \overline{س_د}) + ... (س_ن - \overline{س_د})}{ن}$$

$$= \overline{س_د} + \frac{ح_١ + ح_٢ + ... ح_ن}{ن}$$

$$= \overline{س_د} + \frac{\overline{ح}}{ن}$$

وبنفس الطريقة يمكن إثبات أنه فى حالة التوزيعات التكرارية ،

فإن $\overline{س} = \overline{س_د} + \frac{\overline{ح و}}{\overline{و}}$

حيث و$_{١}$ ، و$_{٢}$ ، ... و$_{ن}$ تمثل تكرارات الفئات المختلفة .

ح$_{١}$ ، ح$_{٢}$ ، ... ح$_{ن}$ تمثل انحرافات مراكز الفئات عن الوسط الفرضى .

الخاصية الثانية .

مجموع مربعات إنحرافات القيم عن وسطها الحسابى أقل من مجموع مربعات الانحرافات عن أية قيمة أخرى . وسنرجئ الكتابة عن هذه الخاصية إلى الباب التالى الخاص بمقاييس التشتت نظراً لارتباطها بهذا الموضوع .

الخاصية الثالثة .

الوسط الحسابى لمجاميع (أو فروق) عدد من أزواج القيم المتناظرة التى تصف تغير ظاهرتين يساوى مجموع (أو الفرق بين) الوسطين الحسابيين للظاهرتين . فلو رمزنا لمجموع الظاهرتين بالرمز س وقيم الظاهرة الأولى بالرمز ص وقيم الظاهرة الثانية بالرمز ع بحيث تكون

$$س_{١} = ص_{١} + ع_{١}$$
$$س_{٢} = ص_{٢} + ع_{٢}$$
$$\vdots$$
$$\vdots$$
$$س_{ن} = ص_{ن} + ع_{ن}$$

فإنه بجمع الطرفين نجد أن

$$\Sigma س = \Sigma ص + \Sigma ع$$

وبالقسمة على ن

$$\therefore \bar{س} = \bar{ص} + \bar{ع}$$

وبديهى أنه إذا كانت قيم س تمثل الفروق بين الأزواج المتناظرة فإن

$$\bar{س} = \bar{ص} - \bar{ع}$$

ويمكن بنفس الطريقة تعميم النتيجة السابقة على العلاقة بين متوسط مجموع أى عدد من القيم المتناظرة والمتوسطات المستقلة لهذه القيم

فإذا كانت $س = ص \pm ع \pm ط \pm \ldots\ldots \pm هـ$

فإن $\bar{س} = \bar{ص} \pm \bar{ع} \pm \bar{ط} \pm \bar{هـ}$

الخاصية الرابعة

الوسط الحسابى لمجموعة من قيم المتغير مقسمة إلى مجموعات جزئية يساوى الوسط الحسابى المرجح بالأوزان لمتوسطات هذه المجموعات حيث تكون الأوزان متناسبة مع عدد التكرارات فى كل مجموعة (جزئية) . ولتوضيح ذلك رمزيا نقول إنه اذا كان

$$\Sigma س = \Sigma س_1 + \Sigma س_2$$

وكان عدد التكرارات فى المجموعة الجزئية الأولى $ن_1$ وعدد التكرارات فى المجموعة الجزئية الثانية $ن_2$ بحيث يكون عدد التكرارات فى المجموعة الأصلية $ن = ن_1 + ن_2$

$$\text{فإن } \bar{س} = \frac{ن_1 \bar{س}_1 + ن_2 \bar{س}_2}{ن_1 + ن_2} = \frac{ن_1 \bar{س}_1 + ن_2 \bar{س}_2}{ن} *$$

أو $ن \bar{س} = ن_1 \bar{س}_1 + ن_2 \bar{س}_2$

وبالمثل إذا كان هناك ر من المجموعات الجزئية تكراراتها $ن_1 ، ن_2 ، \ldots ن_ر$

فإن $ن \bar{س} = ن_1 \bar{س}_1 + ن_2 \bar{س}_2 + \ldots\ldots\ldots + ن_ر \bar{س}_ر$

وتظهر أهمية هذه الخاصية عند إيجاد الوسط الحسابى لتوزيع تكرارى ذى فئات غير متساوية .

---

* وذلك لأن $\bar{س} = \frac{\Sigma س}{ن} = \frac{\Sigma س_1 + \Sigma س_2}{ن} = \frac{ن_1 \bar{س}_1 + ن_2 \bar{س}_2}{ن}$

إيجاد الوسط الحسابى للتوزيع التكرارى

أولا : الطريقة المطولة .

لإيجاد الوسط الحسابى للتوزيع التكرارى نتبع خطوات العمل الاتية مع العلم بأننا نعتبر جميع المفردات الواقعة داخل الفئة الواحدة مساوية لمركز هذه الفئة.

(١) نعين مراكز الفئات

(٢) نضرب مراكز الفئات فى التكرارات المقابلة لها

(٣) نجمع حواصل الضرب الناتجة من الخطوة السابقة

(٤) نقسم حاصل الجمع على عدد المفردات (أى مجموع التكرارات) فنحصل على الوسط الحسابى .

ولتوضيح هذه الخطوات عملياً نوجد الوسط الحسابى للتوزيع التكرارى الذى يبين أعمار الطلبة المتقدمين للالتحاق بكلية التجارة سنة ١٩٤٧ المبين بالجدول رقم ١٢

جدول رقم ١٤

إيجاد الوسط الحسابى للتوزيع التكرارى بالطريقة المطولة

فئات الأعمار	التكرار و	مراكز الفئات س	مراكز الفئات × التكرار س و
١٦ — ١٨	٢٠	١٧	٣٤٠
١٨ — ٢٠	٥٥	١٩	١٠٤٥
٢٠ — ٢٢	٦٥	٢١	١٣٦٥
٢٢ — ٢٤	٣٢	٢٣	٧٣٦
٢٤ — ٢٦	٨	٢٥	٢٠٠
٢٦ — ٢٨	٥	٢٧	١٣٥
٢٨ — ٣٠	٢	٢٩	٥٨
٣٠ — ٣٢	٦	٣١	١٨٦
٣٢ — ٣٤	٣	٣٣	٩٩
٣٤ — ٣٦	٤	٣٥	١٤٠
المجموع	٢٠٠		٤٣٠٤

$$\bar{س} = \frac{\Sigma\, س\, و}{\Sigma\, و} \quad (\text{حيث } \Sigma\, و = \text{مجموع التكرارات})$$

$$= \frac{٤٣٠٤}{٢٠٠} = ٢١,٥٢$$

∴ الوسط الحسابى لأعمار الطلبة = ٢١,٥٢ سنة

ثانياً : إيجاد الوسط الحسابى للتوزيع التكرارى باستخدام وسط فرضى .

أثبتنا من قبل أنه عند استخدام الوسط الفرضى فى إيجاد الوسط الحسابى للتوزيع التكرارى فإن العلاقة التى تربط الوسطين هى .

$$\bar{س} = س_{ف} + \frac{\Sigma\, ء\, و}{\Sigma\, و} \quad \cdots\cdots (١)$$

والآن نطبق هذه العلاقة لإيجاد الوسط الحسابى للتوزيع التكرارى السابق بطريقة مختصرة تسهل العمل كثيراً وتقلل من احتمال ارتكاب الأخطاء فى العمليات الحسابية . ويلاحظ أنه يمكننا اختيار أى وسط فرضى حتى وإن كان خارجاً عن قيم التوزيع . غير أن الأفضل أن نجعل الوسط الفرضى مركز الفئة الوسطى أو تلك التى تحتوى على أكبر عدد من المفردات (أى التى يقابلها أكبر تكرار) . وتكون خطوات العمل كما يأتى :—

أولا : نعين مراكز الفئات .

ثانياً : نحسب انحرافات مراكز الفئات عن الوسط الفرضى .

ثالثاً : نضرب كل انحراف فى التكرار المقابل له .

رابعاً : نجمع حواصل الضرب ونقسم حاصل الجمع على عدد المفردات

أى مجموع التكرارات فيكون الناتج هو $\frac{\Sigma\, ء\, و}{\Sigma\, و}$ أى الفرق بين الوسط الحسابى والوسط الفرضى حسب العلاقة (١) ، ومن ثم نعين قيمة الوسط الحسابى . والجدول الآتى يوضح هذه الخطوات مع العلم بأن الوسط الفرضى الذى اخترناه هو ٢١

جدول رقم ١٥

إيجاد الوسط الحسابى للتوزيع التكرارى باستخدام وسط فرضى

فئات الأعمار	التكرار و	مراكز الفئات س	الانحراف عن الوسط الفرضى ء	الانحرافات × التكرار ء و
١٦ — ١٨	٢٠	١٧	—٤	—٨٠
١٨ — ٢٠	٥٥	١٩	—٢	—١١٠
٢٠ — ٢٢	٦٥	٢١	٠	٠
٢٢ — ٢٤	٣٢	٢٣	٢	٦٤
٢٤ — ٢٦	٨	٢٥	٤	٣٢
٢٦ — ٢٨	٥	٢٧	٦	٣٠
٢٨ — ٣٠	٢	٢٩	٨	١٦
٣٠ — ٣٢	٦	٣١	١٠	٦٠
٣٢ — ٣٤	٣	٣٣	١٢	٣٦
٣٤ — ٣٦	٤	٣٥	١٤	٥٦
المجموع ..	٢٠٠			١٠٤

$$\bar{س} = \bar{س}_{ء} + \frac{\Sigma ء و}{\Sigma و}$$

$$= ٢١ + \frac{١٠٤}{٢٠٠} = ٢١,٥٢$$

.. الوسط الحسابى لأعمار الطلبة = ٢١,٥٢ سنة وهى نفس النتيجة التى حصلنا عليها بالطريقة المطولة .

العمل بوحدات طول الفئة

إذا كانت فئات التوزيع التكراري متساوية كما هو الحال فى المثال المتقدم فيمكننا لزيادة تسهيل العمل الحسابى أن نحسب الانحرافات عن الوسط الفرضى بدلالة وحدات طول الفئة وذلك بأن نقسم كل انحراف على طول الفئة ثم نستمر فى العمل كالمعتاد . على أنه قبل أن نستخرج النتيجة النهائية يجب أن نرد قيمة Σ ء و إلى الوحدات الأصلية بضرب هذه القيمة فى طول الفئة فى النهاية . وقد يبدو لأول وهلة أن قسمة كل انحراف على طول الفئة ثم ضرب مجموع الانحرافات فى طول الفئة بعد ذلك لا يسهل العمل كثيراً بل إنه قد يسبب تعقيده نوعاً . والواقع خلاف ذلك ، فإن تقدير الانحرافات بدلالة وحدات طول الفئة لا يسهل العمل الحسابى فحسب بل إنه يساعد على اكتشاف الأخطاء فى حساب الانحرافات فى الحال . وذلك لأن الانحرافات المحسوبة بهذه الكيفية تكون دائماً مسلسلة حسب الأعداد الطبيعية ١، ٢، ٣، ٤، . . . إلخ فإذا لم نجدها على هذا النحو نعلم فى الحال أن خطأ قد ارتكب فى حسابها . والجدول رقم ١٦ يوضح خطوات العمل باستعمال هذه الطريقة .

تطبيق مراجعة شارلييه لاختبار صحة العمليات الحسابية .

العمل الحسابى الذي يتضمنه ضرب كل انحراف فى التكرار المقابل له ثم جمع حواصل الضرب معرض طبعاً لتسرب الخطأ إليه . ولذلك فمن الضرورى مراجعة هذا العمل بدقة وعناية . وقد وضع شارلييه (Charlier) طريقة سهلة لمراجعة العمليات الحسابية مبينة على العلاقة الآتية .

$$\Sigma \text{و} (\text{ء} + ١) = \Sigma \text{وء} + \Sigma \text{و}$$

وكل ما يلزم عمله لإجراء عملية المراجعة هذه هو أن نضيف إلى خطوات العمل الموضحة بالجدول رقم ١٥ خطوة أخرى هى حساب قيمة Σ و (ء + ١) كما يتضح من الجدول الآتى :—

جدول رقم ١٦

إيجاد الوسط الحسابى باستعمال وحدات طول الفئة مع تطبيق مراجعة شاراليه

مراكز الفئات س	التكرار و	الانحرافات عن الوسط الفرضى	الانحرافات بوحدات طول الفئة ء	الانحرافات × التكرار ء و	(ء+١)	و (ء+١)
١٧	٢٠	٤—	٢—	٤٠—	١—	٢٠—
١٩	٥٥	٢—	١—	٥٥—	٠	٠
٢١	٦٥	٠	٠	٠	١	٦٥
٢٣	٣٢	٢	١	٣٢	٢	٦٤
٢٥	٨	٤	٢	١٦	٣	٢٤
٢٧	٥	٦	٣	١٥	٤	٢٠
٢٩	٢	٨	٤	٨	٥	١٠
٣١	٦	١٠	٥	٣٠	٦	٣٦
٣٣	٣	١٢	٦	١٨	٧	٢١
٣٥	٤	١٤	٧	٢٨	٨	٣٢
المجموع	٢٠٠			٥٢		٢٥٢

مراجعة شاراليه :

$$\Sigma و (ء + ١) = \Sigma و ء + \Sigma و$$

وبالنظر إلى الجدول نجد أن $\Sigma و (ء + ١) = ٢٥٢$

، $\Sigma و ء + \Sigma و = ٥٢ + ٢٠٠ = ٢٥٢$

∴ الطرفان متساويان

ومن هذا نستنتج أن العمليات الحسابية صحيحة .

مجموع الانحرافات × التكرارات بوحدات طول الفئة = ٥٢

∴ مجموع الانحرافات × التكرارت بالوحدات الأصلية

= ٢ × ٥٢ = ١٠٤

وبتطبيق العلاقة $\bar{س} = س_{ء} + \frac{مج ء و}{مج و}$

∴ $\bar{س}$ = ٢١ + $\frac{١٠٤}{٢٠٠}$ = ٢١,٥٢

∴ الوسط الحسابى لأعمار الطلبة = ٢١,٥٢ سنة .

الوسط الحسابى لتوزيع تكرارى ذى فئات غير متساوية

يصادفنا أحياناً توزيعات تكرارية ذات فئات غير متساوية . ويمكن إيجاد الوسط الحسابى لمثل هذه التوزيعات بالطريقة المطولة كما هو الحال فى التوزيعات ذات الفئات المتساوية تماماً . غير أنه يمكن تسهيل العمل نوعاً باستخدام الخاصية الرابعة للوسط الحسابى[1] وذلك بتقسيم مجموعة المتغير الأصلية إلى مجموعات جزئية فئاتها متساوية ثم حساب متوسطات هذه المجموعات بالطريقة المختصرة وإيجاد الوسط الحسابى للمجموعة الأصلية منها باستخدام العلاقة :—

$$\bar{س} = \frac{ن_١ \bar{س}_١ + ن_٢ \bar{س}_٢ + \cdots + ن_ر \bar{س}_ر}{ن_١ + ن_٢ + \cdots + ن_ر}$$

والجدول الآتى يوضح هذه الطريقة :—

---

(١) راجع ص من هذا الكتاب .

جدول رقم ١٧

إيجاد الوسط الحسابي لتوزيع تكراري ذي فئات غير متساوية بالطريقة المختصرة

فئات المتغير	التكرار و	مراكز الفئات س	الانحرافات عن الوسط الفرضي بوحدات طول الفئة ى	ى و
٢ — ٤	١	٣	—٢	—٢
٤ — ٦	٢	٥	—١	—٢
٦ — ٨	٤	٧	٠	٠
٨ — ١٠	٣	٩	١	٣
المجموع	١٠			—١
١٠ — ١٥	٢	١٢,٥	—٢	—٤
١٥ — ٢٠	٥	١٧,٥	—١	—٥
٢٠ — ٢٥	١٠	٢٢,٥	٠	٠
٢٥ — ٣٠	٣	٢٧,٥	١	٣
المجموع	٢٠			—٦
٣٠ — ٤٠	٤	٣٥	—١	—٤
٤٠ — ٥٠	٦	٤٥	٠	٠
٥٠ — ٦٠	٣	٥٥	١	٣
٦٠ — ٧٠	٢	٦٥	٢	٤
المجموع	١٥			٣

الوسط الحسابى للمجموعة الجزئية الأولى = ٧ + ٢ (١/١٠) = ٦,٨

الوسط الحسابى للمجموعة الجزئية الثانية = ٢٢,٥ + ٥ (٦/٢٠) = ٢١

الوسط الحسابى للمجموعة الجزئية الثالثة = ٤٥ + ١٠ (٣/١٥) = ٤٧

... الوسط الحسابى للتوزيع التكرارى = $\frac{٦,٨ \times ١٠ + ٢١ \times ٢٠ + ٤٧ \times ١٥}{١٠ + ٢٠ + ١٥}$

= ٢٦,٥ تقريبا

٢ ــ الوسط الهندسى

الوسط الهندسى لمجموعة من القيم عدد مفرداتها ن هو الجذر النونى لحاصل ضرب هذه القيم . وعلى ذلك لو كان لدينا مجموعة من قيم المتغير س مفرداتها $س_١$ ، $س_٢$ ، $س_٣$ . . . . $س_ن$

فإن الوسط الهندسى لهذه القيم ويرمز إليه عاده بالحرف هـ

$$= \sqrt[ن]{س_١ \times س_٢ \times س_٣ \times \ldots س_ن}$$

$$= (س_١ \times س_٢ \times س_٣ \times \ldots س_ن)^{\frac{١}{ن}}$$

ويمكن تسهيل العمليات الحسابية الخاصة باستخراج الوسط الهندسى بأخذ لوغاريتمات الطرفين كالآتى :ــ

$$لو\ هـ = \frac{١}{ن} (لو س_١ + لو س_٢ + لو س_٣ + \ldots لو س_ن)$$

$$= \frac{\Sigma\ لو\ س}{ن}$$

ويتضح من المتساوية السابقة أن لوغاريتم الوسط الهندسى لمجموعة من القيم يساوى الوسط الحسابى للوغاريتمات هذه القيم .

وفى حالة وجود أوزان للقيم (أو فى حالة التوزيعات التكرارية) يكون الوسط الهندسى .

$$هـ = \sqrt[\Sigma و]{س_1^{و_1} \times س_2^{و_2} \times \ldots \ldots \times س_ن^{و_ن}}$$

$$= \left(س_1^{و_1} \times س_2^{و_2} \times \ldots \ldots \times س_ن^{و_ن} \right)^{\frac{1}{\Sigma و}}$$

$$\therefore \text{لو } هـ = \frac{1}{\Sigma و} (و_1 \text{ لوس}_1 + و_2 \text{ لوس}_2 + \ldots \ldots و_ن \text{ لوس}_ن)$$

$$= \frac{\Sigma و \text{ لوس}}{\Sigma و}$$

حيث $و_1$ ، $و_2$ ، . . . . $و_ن$ تمثل الأوزان (أو التكرارات)

وعند حساب الوسط الهندسى للتوزيع التكرارى تكون خطوات العمل كالآتى :

أولا : نعين مراكز الفئات .

ثانياً : نستخرج لوغاريتمات مراكز الفئات .

ثالثاً : نضرب لوغاريتمات مراكز الفئات فى التكرارات المقابلة لها .

رابعاً : نجمع حواصل الضرب فيكون الناتج هو Σ و لوس

وبقسمة هذا الناتج على مجموع التكرارات نحصل على لو هـ

وباستخدام جداول الأعداد المقابلة للوغاريتمات نحصل على قيمة الوسط الهندسى . والان نوضح هذه الخطوات عملياً بإيجاد الوسط الهندسى للتوزيع التكرارى لأعمار الطلبة المتقدمين للالتحاق بكلية التجارة المبين بالجدول رقم ١٨

جدول رقم ١٨

طريقة حساب الوسط الهندسى للتوزيع التكرارى عملياً

فئات الأعمار	التكرار و	مراكز الفئات و	لوس	و لوس
١٦ — ١٨	٢٠	١٧	١,٢٣٠٤	٢٤,٦٠٨٠
١٨ — ٢٠	٥٥	١٩	١,٢٧٨٨	٧٠,٣٣٤٠
٢٠ — ٢٢	٦٥	٢١	١,٣٢٢٢	٨٥,٩٤٣٠
٢٢ — ٢٤	٣٢	٢٣	١,٣٦١٧	٤٣,٥٧٤٤
٢٤ — ٢٦	٨	٢٥	١,٣٩٧٩	١١,١٨٣٢
٢٦ — ٢٨	٥	٢٧	١,٤٣١٤	٧,١٥٧٠
٢٨ — ٣٠	٢	٢٩	١,٤٦٢٤	٢,٩٢٤٨
٣٠ — ٣٢	٦	٣١	١,٤٩١٤	٨,٩٤٨٤
٣٢ — ٣٤	٣	٣٣	١,٥١٨٥	٤,٥٥٥٥
٣٤ — ٣٦	٤	٣٥	١,٥٤٤١	٦,١٧٦٤
المجموع	٢٠٠			٢٦٥,٤٠٤٧

$$\text{لو هـ} = \frac{\Sigma \text{ و لوس}}{\Sigma \text{ و}} = \frac{٢٦٥,٤٠٤٧}{٢٠٠} = ١,٣٢٧٠٢$$

... الوسط الهندسى = ٢١,٢٣ سنة

التطبيقات العملية للوسط الهندسى

أولا : إيجاد متوسط النسب : أهم الحالات التى يستخدم فيها الوسط الهندسى هى تلك التى نرغب فيها فى إيجاد متوسط نسب الزيادة (أو النقص) فى

قيم الظواهر . فالمتوسطات الأخرى تعطينا نتائج غير منطقية إذا استخدمت فى إيجاد النسبة المتوسطة للتغير . ولإيضاح ذلك افرض أن لدينا ظاهرة تتغير على النحو الآتى :—

السنة	قيمة الظاهرة	النسبة المئوية للزيادة
١٩٤٠	١٠٠	—
١٩٤١	١٥٠	٥٠
١٩٤٢	٣٠٠	١٠٠
١٩٤٣	٥٢٥	٧٥

إذ استخدمنا الوسط الحسابى لنسب الزيادة للتعبير عن نسبة الزيادة فى المتوسط فإننا نحصل على النتيجة الآتية :—

$$\bar{س} = \frac{٧٥ + ١٠٠ + ٥٠}{٣} = ٧٥$$

∴ نسبة الزيادة فى المتوسط = ٧٥٪

غير أنه لكى تكون هذه النتيجة منطقية يجب أن نحصل على قيمة الظاهرة فى السنة الأخيرة (١٩٤٣) إذا علينا نسبة الزيادة هذه لمدة ثلاث سنوات على قيمة الظاهرة فى السنة الأولى . أى أنه يجب أن تكون :

$$١٠٠ (١,٧٥)^{٣} = ٥٢٥$$

والواقع خلاف ذلك فإن $١٠٠ (١,٧٥)^{٣} = ٥٣٥,٩٣٧٥$

وعلى ذلك فالوسط الحسابى لا يصلح للتعبير بدقة عن متوسط النسب .

أما النسبة الصحيحة فهى النسبة التى إذا اعتبرت كسعر فائدة تجعل مبلغ ١٠٠ جنيه يزداد فى ثلاث سنوات إلى ٥٢٥ جنيه . أى أن النسبه الصحيحة هى قيمة ر فى المعادلة .

$١٠٠ (١ + ر)^{٣} = ٥٢٥$ حسب قانون الفائدة المركبة . وبحل هذه

المعادلة نجد أن قيمة ر = ٧٣,٨ ٪

أى أن النسبة المتوسطة للزيادة فى قيمة الظاهرة = ٧٣,٨ ٪

ويمكن الحصول على هذه النسبة المتوسطة باستخدام الوسط الهندسى كما يأتى :

قيمة الظاهرة فى سنة ١٩٤١ = ١٥٠ ٪ من قيمتها فى سنة ١٩٤٠

قيمة الظاهرة فى سنة ١٩٤٢ = ٢٠٠ ٪ من قيمتها فى سنة ١٩٤١

قيمة الظاهرة فى سنة ١٩٤٣ = ١٧٥ ٪ من قيمتها فى سنة ١٩٤٢

الوسط الهندسى لهذه النسب[1] = $\sqrt[3]{١٥٠ \times ٢٠٠ \times ١٧٥}$ = ١٧٣,٨

∴ النسبة المتوسطة للزيادة = ٧٣,٨ ٪ وهى النسبة الصحيحة كما سبق أن بينا . ولاختبار صحتها نحسب قيمة ١٠٠ (١,٧٣٨)^{٣} فنجد أنها = ٥٢٥

ثانياً — تقدير عدد السكان بين سنى التعداد

افرض أن عدد السكان فى سنة ١٩٢٧ = ٩ مليون نسمة

وأن عدد السكان فى سنة ١٩٣٧ = ١٦ مليون نسمة

وأردنا عمل تقدير تقريبى لعدد السكان فى سنة ١٩٣٢ على فرض أن السكان يتزايدون بنسبة ثابتة .

بتطبيق قانون الفائدة المركبة على زيادة السكان نجد أن :—

١٦ = ٩ (١ + ر)^{١٠} حيث ر تمثل نسبة الزيادة .

ولإيجاد قيمة ر نأخذ لوغاريتمات الطرفين كالاتى :—

لو ١٦ = لو ٩ + ١٠ لو (١ + ر)

---

(١) من المهم جداً أن نلاحظ أن الوسط الهندسى لنسب الزيادة ذاتها لا يعطينا النتيجة الصحيحة .

فإن $\sqrt[3]{٥٠ \times ١٠٠ \times ٧٥}$ = ٧٢,١١ ٪ وليست هذه هى النسبة المطلوبة . لأنه لو أخذنا متوسط النسب ذاتها فإننا نكون قد حسبنا الزيادة فى الظاهرة (السكان مثلا) على أساس قانون الفائدة البسيطة وليس قانون الفائدة المركبة . وبعبارة أخرى فإننا نكون قد تجاهلنا زيادة الزيادة .

∴ ١,٢٠٤١٢ = ٠,٩٥٤٢٤٣ + ١٠ لو (١ + ر)

∴ ١٠ لو (١ + ر) = ٠,٢٤٩٨٧٧

∴ لو (١ + ر) = ٠,٠٢٤٩٨٧٧

∴ ١ + ر = ١,٠٥٩١

أى أن نسبة الزيادة السنوية فى عدد السكان = ٥,٩١ ٪ تقريباً

وبتطبيق قانون الفائدة المركبة مرة أخرى نجد أن عدد السكان فى سنة ١٩٣٢ = ٩ (١,٠٥٩١)^{٥} = ١٢ مليون نسمة

غير أنه بدلا من إجراء هذه العمليات الحسابية المطولة والمعرضة للأخطاء يمكننا تقدير عدد السكان فى منتصف الفترة الواقعة بين سنتى التعداد مباشرة بإيجاد الوسط الهندسى لعدد السكان فى سنتى التعداد كالآتى :—

عدد السكان فى سنة ١٩٣٢ .

= $\sqrt{\text{عدد السكان فى سنة ١٩٢٧} \times \text{عدد السكان فى سنة ١٩٣٧}}$

= $\sqrt{٩ \times ١٦}$ = ١٢ مليون نسمة

وهى نفس النتيجة التى حصلنا عليها بتطبيق قانون الفائدة المركبة ، ويمكن إثبات هذه النتيجة بالبرهان الجبرى الاتى :—

افرض أن عدد السكان فى سنة التعداد الأولى = ع .

وأن عدد السكان فى سنة التعداد التالية (بعد ن سنة) = ع$_{ن}$

بتطبيق قانون الفائدة المركبة نجد أن ع$_{ن}$ = ع . (١ + ر)^{ن} حيث ر تمثل نسبة الزيادة .

وعدد السكان فى منتصف الفترة الواقعة بين سنتى التعداد = ع . (١+ر)$^{\frac{ن}{٢}}$

بتربيع هذه القيمة الأخيرة ثم أخذ جذرها التربيعى نجد أن :—

عدد السكان فى منتصف الفترة الواقعة بين سنتى التعداد .

= $\sqrt{\text{ع}^{٢} . (١ + \text{ر})^{ن}}$

$$= \sqrt{ع_{.} \times ع_{.}(١+ر)^{ن}} = \sqrt{ع_{.} \times ع_{ن}}$$

وهذه الكمية الأخيرة هى الوسط الهندسى لعدد السكان فى سنتى التعداد .

ثالثاً — عمل الأرقام القياسية .

يستخدم الوسط الهندسى كثيراً فى إيجاد متوسط التغير النسبى للأسعار عند عمل الأرقام القياسية . ولو أنه من الممكن استخدام أحد المتوسطات الأخرى فى إيجاد متوسط التغير النسبى للأسعار إلا أن الوسط الهندسى يفضل جميع المتوسطات الأخرى فى هذه الحالة بالذات لأنه يحقق شروطا خاصة يجب توافرها فى الرقم القياسى الجيد . وسنوضح ذلك بالتفصيل فى الباب الخاص بالأرقام القياسية .

٣ — الوسط التوافقى

الوسط التوافقى لمجموعة من القيم هو مقلوب الوسط الحسابى لمقلوبات هذه القيم .

فإذا كانت هذه القيم هى $س_{١}$ ، $س_{٢}$ ، $س_{٣}$ . . . $س_{ن}$ ورمزنا للوسط التوافقى بالرمز ق فإن

$$\frac{١}{ق} = \frac{\frac{١}{س_{١}} + \frac{١}{س_{٢}} + \cdots\cdots + \frac{١}{س_{ن}}}{ن} = \frac{\Sigma\left(\frac{١}{س}\right)}{ن}$$

$$\therefore \quad ق = \frac{ن}{\Sigma\left(\frac{١}{س}\right)}$$

وسيجد القارئ أن هذا المتوسط فكرته معقدة كما أن حسابه يتطلب مجهوداً كبيراً خصوصاً إذا كانت قيم المتغير كبيرة وكثيرة العدد .

وعلاوة على ذلك فإذا راجعنا الشروط الواجب توافرها فى المتوسط

الجيد نجد أن الوسط التوافقى لا يحقق معظمها . ولهذا فإن استعماله نادر ويكاد يكون تطبيقه العملى مقصوراً على تلك الحالات التى نعبر فيها عن أسعار السلع بدلالة عدد وحدات السلعة التى تشتريها وحدة النقود ، وليس بدلالة عدد وحدات النقود الواجب دفعها لحيازة وحدة السلعة كما هو المعتاد . فإذا قلنا مثلا أن سعر البيض فى أحد الأسواق ٣ بقرش وفى سوق أخرى ٥ بقرش وفى سوق ثالثة ٧ بقرش ، فإن السعر المتوسط (على فرض تساوى الأوزان) يكون دون شك ٥ بقرش . هذا السعر المتوسط (أى ٥ بيضات للقرش) هو فى الواقع الوسط التوافقى للأسعار الثلاثة . ولو حسبنا أحد المتوسطات الأخرى لهذه الأسعار لحصلنا على نتائج مخالفة كما يتضح من حساب الوسط الحسابى لها مع ملاحظة أن هذه الأسعار هى فى الحقيقة $\frac{١}{٣}$ قرش ، $\frac{١}{٥}$ قرش ، $\frac{١}{٧}$ قرش على التوالى .

$$\text{الوسط الحسابى للأسعار} = \frac{\frac{١}{٣} + \frac{١}{٥} + \frac{١}{٧}}{٣}$$

$$= \frac{١}{٣} \left(\frac{١٥ + ٢١ + ٣٥}{١٠٥} \right) = \frac{٧١}{٣١٥}$$

وهذا لا يساوى $\frac{١}{٥}$ وهو ما يجب أن يكون عليه السعر المتوسط بداهة .

أما الوسط التوافقى لهذه الأسعار فيمكن حسابه كالآتى :ـ

$$\frac{١}{ق} = \frac{٧ + ٥ + ٣}{٣} = ٥$$

... ق = $\frac{١}{٥}$ وهو السعر المتوسط

ويلاحظ أنه فى مثل هذه الحالات يكون الوسط التوافقى أسهل فى الحساب من أى وسط آخر . ويمكن ترجيح الوسط التوافقى بالأوزان بنفس الطريقة التى اتبعناها فى حالة الوسط الحسابى . فالوسط التوافقى المرجح بالأوزان هو :

$$\frac{و_١ + و_٢ + \cdots\cdots و_ن}{\left(\frac{١}{س_١} \times و_١\right) + \left(\frac{١}{س_٢} \times و_٢\right) + \cdots\cdots \left(\frac{١}{س_ن} \times و_ن\right)}$$

$$= \frac{\Sigma\, و}{\Sigma\left(\frac{و}{س}\right)}$$

وفى حالة التوزيعات التكرارية نتبع الخطوات الاتية فى حساب الوسط التوافقى :

أولا : نعين مراكز الفئات .

ثانياً : نحسب مقلوبات مراكز الفئات .

ثالثاً : نضرب مقلوبات مراكز الفئات فى التكرارات المقابلة لها ونجمع حواصل الضرب فيكون الناتج هو $\Sigma\left(\frac{و}{س}\right)$.

ثم نحسب قيمة الوسط التوافقى من العلاقة $\frac{١}{ق} = \frac{\Sigma\left(\frac{و}{س}\right)}{\Sigma\, و}$

حيث $\Sigma\, و$ تمثل مجموع التكرارات .

والجدول الآتى يوضح الطريقة العملية لحساب الوسط التوافقى :-

جدول رقم ١٩

حساب الوسط التوافقى للتوزيع التكرارى عملياً

فئات الأعمار	التكرار و	مراكز الفئات س	$\frac{١}{س}$	$\frac{١}{س} \times$ و
١٦ — ١٨	٢٠	١٧	٠,٠٥٨٨٢	١,١٧٦
١٨ — ٢٠	٥٥	١٩	٠,٠٥٢٦٣	٢,٨٩٥
٢٠ — ٢٢	٦٥	٢١	٠,٠٤٧٦٢	٣,٠٩٥
٢٢ — ٢٤	٣٢	٢٣	٠,٠٤٣٤٨	١,٣٩١
٢٤ — ٢٦	٨	٢٥	٠,٠٤٠٠٠	٠,٣٢٠
٢٦ — ٢٨	٥	٢٧	٠,٠٣٧٠٤	٠,١٨٥
٢٨ — ٣٠	٢	٢٩	٠,٠٣٤٤٨	٠,٠٦٩
٣٠ — ٣٢	٦	٣١	٠,٠٣٢٢٦	٠,١٩٤
٣٢ — ٣٤	٣	٣٣	٠,٠٣٠٣٠	٠,٠٩١
٣٤ — ٣٦	٤	٣٥	٠,٠٢٨٥٧	٠,١١٤
المجموع	٢٠٠			٩,٥٣٠

$$\frac{١}{ق} = \frac{\Sigma (\frac{و}{س})}{\Sigma و} = \frac{٩,٥٣٠}{٢٠٠}$$

∴ الوسط التوافقى = $\frac{٢٠٠}{٩,٥٣٠}$ = ٢٠,٩٨٦ سنة

٤ ــ الوسط التربيعى

الوسط التربيعى لمجموعة من القيم هو الجذر التربيعى لمتوسط مربعات هذه القيم . فإذا كانت هذه القيم س١ ، س٢ ، س٣ ... ، سن

فإن الوسط التربيعى = $\sqrt{\frac{\Sigma س^{٢}}{ن}}$

وفى حالة التوزيعات التكرارية يكون الوسط التربيعى = $\sqrt{\frac{\Sigma س^{٢} و}{\Sigma و}}$

وقد يبدو غريباً أن نختار مثل هذا المتوسط لوصف المجموعات مع أن هناك متوسطات أخرى أسهل فى الفكرة وأسهل فى الحساب .

والواقع أن الوسط التربيعى ، كالوسط التوافقى ، نادر الاستعمال جداً . غير أن تلك الحالات النادرة التى يستعمل فيها ذات أهمية ولا يمكن فيها الاستعاضة عنه بأى متوسط آخر . افرض مثلا أننا أردنا حساب متوسط بعض الكميات التى تحتوى على قيم سالبة وكان اهتمامنا منصباً على معرفة متوسط القيم ذاتها بغض النظر عن إشاراتها السالبة أو الموجبة . من الواضح أنه فى مثل هذه الحالات لا يمكن استعمال الوسط الهندسى ، أما الوسط الحسابى فإنه يعطينا وصفاً بعيداً جداً عن الحقيقة لهذه القيم نظراً لأن القيم السالبة تلغى أثر قيم موجبة معادلة لها فى المجموع عند إجراء عملية الجمع الجبرى للقيم ، اللازمة لحساب الوسط الحسابى . فلو فرض أن كانت القيم السالبة تعادل فى مجموعها القيم الموجبة لكان الوسط الحسابى يساوى صفراً ، وهو لهذا السبب لا يصلح مطلقاً لوصف مجموعة من القيم بينها قيم سالبة . على أنه قد يرد على هذا بأنه من الممكن وصف هذه المجموعات عن طريق حساب الوسط الحسابى للقيم مجردة عن إرشاتها . وهذا الرد من الوجاهة بمكان ، غير أن هناك بعض الحالات التى

نجد فيها من الأنسب التخلص من إشارات القيم عن طريق تربيعها بدلا من مجرد إهمال هذة الإشارات ، وهذه هى الحالات التى يستعمل فيها الوسط التربيعى وسيجد القارئ استعمالا هاماً للوسط التربيعى فى الباب التالى الخاص بمقاييس التشتت . والواقع أن الوسط التربيعى ليس له أى تطبيقات عملية تستحق الذكر فى ميدان الدراسات الاقتصادية . وتكاد تطبيقاته تكون قاصرة على بعض فروع علم الطبيعة خصوصاً فيما يتعلق بدراسة التيارات الكهربائية (١).

ويمكن ترجيح الوسط التربيعى بالأوزان بنفس الطريقة التى اتبعناها فى حالة الوسط الحسابى ، كما أن الطريقة العملية لحساب الوسط التربيعى للتوزيع التكرارى لا تختلف عن الطريقة المطولة لإيجاد الوسط الحسابى . وفيما يلى خطوات العمل التى نتبعها فى إيجاد الوسط التربيعى للتوزيع التكرارى .

أولا : نعين مراكز الفئات .

ثانياً : نحسب مربعات مراكز الفئات .

ثالثاً : نضرب مربعات مراكز الفئات فى التكرارات المقابلة لها ونجمع حواصل الضرب فيكون الناتج هو Σ س^{٢} و .

رابعاً : نقسم Σ س^{٢} و على مجموع التكرارات .

خامساً : نستخرج الجذر التربيعى لخارج القسمة فيكون الناتج هو الوسط التربيعى . ولتوضيح هذه الخطوات عملياً نحسب الوسط التربيعى لأعمار الطلبة المتقدمين للالتحاق بكلية التجارة سنة ١٩٤٧ .

---

(١) انظر كتاب Calculus Made Easy by S. P. Thompson ص ٢٢٢ .

جدول رقم ٢٠

إيجاد الوسط التربيعى للتوزيع التكرارى

فئات الأعمار	التكرار و	مراكز الفئات س	س٢	س٢×و
١٦ — ١٨	٢٠	١٧	٢٨٩	٥٧٨٠
١٨ — ٢٠	٥٥	١٩	٣٦١	١٩٨٥٥
٢٠ — ٢٢	٦٥	٢١	٤٤١	٢٨٦٦٥
٢٢ — ٢٤	٣٢	٢٣	٥٢٩	١٦٩٢٨
٢٤ — ٢٦	٨	٢٥	٦٢٥	٥٠٠٠
٢٦ — ٢٨	٥	٢٧	٧٢٩	٣٦٤٥
٢٨ — ٣٠	٢	٢٩	٨٤١	١٦٨٢
٣٠ — ٣٢	٦	٣١	٩٦١	٥٧٦٦
٣٢ — ٣٤	٣	٣٣	١٠٨٩	٣٢٦٧
٣٤ — ٣٦	٤	٣٥	١٢٢٥	٤٩٠٠
المجموع	٢٠٠			٩٥٤٨٨

$$\text{الوسط التربيعى} = \sqrt{\frac{\Sigma\, س^2 و}{\Sigma\, و}}$$

$$= \sqrt{\frac{٩٥٤٨٨}{٢٠٠}}$$

$$= ٢١,٨٥ \text{ سنة}$$

العلاقات بين المتوسطات :

بمراجعة قيم الوسط الحسابى والهندسى والتوافقى والتربيعى للتوزيع التكرارى لأعمار الطلبة المتقدمين للالتحاق بكلية التجارة نجدها على التوالى :—

٢١,٥٢ ، ٢١,٢٣ ، ٢٠,٩٨٦ ، ٢١,٨٥ سنة

أى أن هذه النتائج تظهر أن :

الوسط التربيعى > الوسط الحسابى > الوسط الهندسى > الوسط التوافقى

والواقع أن هذه نتيجة عامة يمكن إثباتها جبريا كالآتى :-

أولا : لإثبات أن الوسط التربيعى أكبر من الوسط الحسابى افرض أن لدينا كميتين موجبتين هما ا ، ب

الوسط التربيعى لهاتين الكميتين = $\sqrt{\frac{ا^٢ + ب^٢}{٢}}$

والوسط الحسابى لهما = $\frac{ا + ب}{٢}$

فكأن المطلوب إثباته أن $\left(\sqrt{\frac{ا^٢ + ب^٢}{٢}}\right) > \left(\frac{ا + ب}{٢}\right)$

بتربيع الطرفين ... $\left(\frac{ا^٢ + ب^٢}{٢}\right) > \left(\frac{ا^٢ + ٢اب + ب^٢}{٤}\right)$

بتجزيء الحدود ... $\left(\frac{ا^٢ + ب^٢}{٤} + \frac{ا^٢ + ب^٢}{٤}\right) > \left(\frac{ا^٢ + ب^٢}{٤} + \frac{٢اب}{٤}\right)$

أى أن المطلوب إثباته هو أن $(ا^٢ + ب^٢) > ٢اب$

وحيث أن $ا^٢ + ب^٢ - ٢اب = (ا - ب)^٢$ = كمية موجبة

... $(ا^٢ + ب^٢) > ٢اب$

... الوسط التربيعى أكبر من الوسط الحسابى

ثانياً - لإثبات أن الوسط الحسابى أكبر من الوسط الهندسى افرض أن لدينا كميتين موجبتين ا ، ب

الوسط الحسابى لهاتين الكميتين = $\frac{ا + ب}{٢}$

والوسط الهندسى لهما = $\sqrt{اب}$

فكأن المطلوب إثباته أن $\left(\frac{ا + ب}{٢}\right) > \left(\sqrt{ا ب}\right)$

أو أن $(ا + ب) - (٢\sqrt{ا ب})$ = كمية موجبة

وحيث أن $ا + ب - ٢\sqrt{ا ب} = (\sqrt{ا} - \sqrt{ب})^{٢}$ = كمية موجبة

∴ الوسط الحسابى أكبر من الوسط الهندسى .

ثالثاً — لإثبات أن الوسط الهندسى أكبر من الوسط التوافقى نفرض كذلك أن لدينا كميتين موجبتين ا ، ب

الوسط الهندسى لهاتين الكميتين $= \sqrt{ا ب}$

والوسط التوافقى لهما $= \frac{٢ ا ب}{ا + ب}$

فكأن المطلوب إثباته أن $\sqrt{ا ب} > \frac{٢ ا ب}{ا + ب}$

أثبتنا من قبل أن $\frac{ا + ب}{٢} > \sqrt{ا ب}$

∴ $ا + ب > ٢\sqrt{ا ب}$

∴ $(ا + ب)\sqrt{ا ب} > ٢ ا ب$

∴ $\sqrt{ا ب} > \frac{٢ ا ب}{ا + ب}$

∴ الوسط الهندسى أكبر من الوسط التوافقى

ويلاحظ أن جميع النتائج السابقة تظل صحيحة مهما كان عدد الكميات التى نحسب متوسطاتها . وهناك علاقة أخرى تربط الوسط الحسابى والهندسى والتوافقى . إذا كانت هذه المتوسطات هى على التوالى ح ، هـ ، ق

فإن $هـ = \sqrt{ح ق}$

ولإثبات ذلك افرض أن لدينا أى كميتين موجبتين ا ، ب

$$\text{ه} = \sqrt{\text{أ ب}} = \sqrt{\frac{\text{١٢ أ ب}}{\text{ب + ١}} \times \frac{\text{أ + ب}}{\text{٢}}} = \sqrt{\text{ح ق}}$$

٥ – الوسيط

الوسيط هو القيمة الوسطى بين مفردات المجموعة . فإذا رتبت قيم المجموعة ترتيباً تصاعدياً (أو تنازلياً) يكون الوسيط هو تلك القيمة التى تقع فى وسط المفردات تماماً بحيث أن عدد المفردات التى تسبقها ، أى الأقل منها فى القيمة ، يعادل عدد المفردات التى تليها ، أى الأكبر منها فى القيمة . فإذا كانت مفردات المجموعة مرتبة كالآتى :–

٣ ، ٥ ، ٦ ، ٧ ، ٩ ، ١٣ ، ١٧

فإن قيمة الوسيط هى ٧ وذلك لأن عدد المفردات التى تسبقها (٣ ، ٥ ، ٦) يساوى عدد المفردات التى تليها (٩ ، ١٣ ، ١٧)

فإذا كان عدد المفردات زوجياً مثل :–

٢ ، ٥ ، ٦ ، ٧ ، ٩ ، ١٢

فمن الواضح أنه لا يوجد بينها قيمة واحدة ينطبق عليها تعريف الوسيط . غير أنه لو تصورنا وجود قيمة فى منتصف الفترة الواقعة بين القيمتين المتوسطتين (٦ ، ٧) فإنها تحقق هذا الشرط ،وعلى ذلك فإذا أردنا إيجاد الوسيط لمجموعة من المفردات عددها زوجى فإننا نعتبره الوسط الحسابى للقيمتين الواقعتين فى وسط المفردات تماماً . وبناء على هذه القاعدة يكون وسيط المفردات السابقة هو $٦\frac{١}{٢}$.

ومن أهم خواص الوسيط أنه لا يتأثر بقيم المفردات التى تسبقه أو تلك التى تليه . إذ أن المهم هو ترتيبه بين هذه المفردات . ومن هذا يتضح أن قيمة الوسيط ، بخلاف المتوسطات الأخرى السابق ذكرها ، لا تتأثر بوجود بعض القيم الشاذة بين المفردات . فمثلا إذا كانت القيم هى :–

٥ ، ٦ ، ٧ ، ٨ ، ١١ فإن الوسيط = ٧

وإذا كانت ١ ، ٦ ، ٧ ، ٩ ، ٣٥ فإن الوسيط = ٧

كما أنه لو زاد عدد هذه المفردات فإن قيمة الوسيط لا تتغير ما دام ترتيبه يظل فى الوسط . فإذا كانت القيم هى : ١، ٢، ٣، ٤، ٧، ١١، ١٣، ١٧، ١٩

فإن الوسيط = ٧ كذلك

وعلى وجه العموم إذا فرضنا أن عدد المفردات يساوى ن فإن ترتيب الوسيط = $\frac{ن + ١}{٢}$ *

فلو كان عدد القيم ٥ فإن الوسيط هو القيمة التى ترتيبها $\frac{٥ + ١}{٢}$ أى القيمة الثالثة.

ولو كان عدد القيم ٦ فإن الوسيط هو القيمة التى ترتيبها $\frac{٦ + ١}{٢} = ٣\frac{١}{٢}$

أى أننا فرضنا وجود قيمة وهمية بين القيمتين الثالثة والرابعة نعتبرها فى مركز الوسيط . ولكى يكون تعبيرنا عن مركز الوسيط أكثر دقة نقول إنه إذا كانت المفردات هى :—

$س_١$ ، $س_٢$ ، $س_٣$ ، .... ، $س_ن$

فإن قيمة الوسيط هى $س_{\frac{ن + ١}{٢}}$ أى القيمة السينية التى ترتيبها $\frac{ن + ١}{٢}$

على فرض أن عدد قيم س كان فردياً .

أما إذا كان عدد قيم س زوجياً فإن قيمة الوسيط

$= \frac{١}{٢} (س_{\frac{ن}{٢}} + س_{١ + \frac{ن}{٢}})$ أى أن الوسيط فى هذه الحالة هو متوسط القيمتين اللتين ترتيبهما $\frac{ن}{٢}$ ، $\frac{ن}{٢} + ١$

(*) يفضل بعض الإحصائيين اعتبار ترتيب الوسيط $\frac{ن + ١}{٢}$ بينما يفضل البعض الآخر اعتبار ترتيبه $\frac{ن}{٢}$. والواقع أن الفرق بين النتائج التى تؤدى إليها الطريقتان ضئيل جداً بحيث يمكن إهماله عملياً . غير أنى أعتقد أنه من الأفضل اعتبار ترتيب الوسيط $\frac{ن + ١}{٢}$ فى حالة ما إذا كانت مفردات المجموعة عبارة عن أرقام أولية ، $\frac{ن}{٢}$ إذا كانت مفردات المجموعة موضوعة فى شكل توزيع تكرارى خصوصاً إذا كان عدد مفردات التوزيع التكرارى زوجياً .

وعلى ذلك يكون الوسيط لسلسلة الأعداد الطبيعية من ١ إلى ٣٠

$$= \frac{١}{٢} (١٥ + ١٦) = ١٥ \frac{١}{٢}$$

إيجاد الوسيط للتوزيع التكرارى

يمكن حساب قيمة الوسيط بسهولة من التوزيع التكرارى المتجمع الصاعد أو النازل وذلك بعد تعيين ترتيبه بين المفردات . والجدل الآتى يوضح الطريقة العملية لإيجاد الوسيط باستخدام التكرار المتجمع الصاعد .

جدول رقم ٢١

إيجاد الوسيط لأعمار الطلبة المتقدمين للالتحاق بكلية التجارة سنة ١٩٤٧

فئات الأعمار	التكرار	الحدود العليا للفئات	التكرار المتجمع الصاعد
١٦ — ١٨	٢٠	أقل من ١٨	٢٠
١٨ — ٢٠	٥٥	« « ٢٠	٧٥
٢٠ — ٢٢	٦٥	« « ٢٢	١٤٠
٢٢ — ٢٤	٣٢	« « ٢٤	١٧٢
٢٤ — ٢٦	٨	« « ٢٦	١٨٠
٢٦ — ٢٨	٥	« « ٢٨	١٨٥
٢٨ — ٣٠	٢	« « ٣٠	١٨٧
٣٠ — ٣٢	٦	« « ٣٢	١٩٣
٢٤ — ٣٢	٣	« « ٣٤	١٩٦
٣٤ — ٣٦	٤	« « ٣٦	٢٠٠
المجموع ...	٢٠٠		

لإيجاد قيمة الوسيط نبدأ أولا بتعيين ترتيبه وهو فى هذه الحالة

$$\frac{٢٠٠}{٢} = ١٠٠$$

وبالنظر إلى الجدول رقم ٢١ نجد أن الـ ٧٥ قيمة الأولى جميعها أقل من

٢٠ سنة والـ ١٤٠ قيمة الأولى جميعها أقل من ٢٢ سنة . وحيث أن ترتيب الوسيط يقع بين هذين الحدين فلابد أن تكون قيمته محصورة بين ٢٠، ٢٢ سنة فلو فرضنا أن المفردات موزعة بانتظام داخل الفئات يمكننا بعملية استكمال بسيطة أن نحسب قيمة الوسيط فنقول :—

فرق فى التكرار قدره (١٤٠ — ٧٥) يناظره فرق فى الأعمار قدره (٢٢ — ٢٠) .

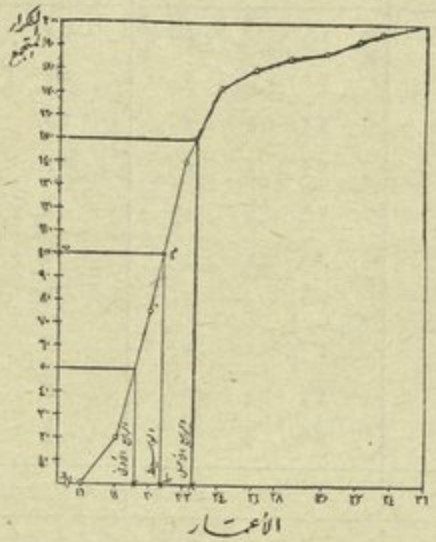
∴ فرق فى التكرار قدره (١٠٠ — ٧٥) يناظره فرق فى الأعمار قدره س

$$\text{س} = \frac{٢٥ \times ٢}{٦٥} = ٠,٧٦٩$$

∴ قيمة الوسيط = ٢٠ + ٠,٧٦٩ = ٢٠,٧٦٩ سنة

ويمكن بنفس الطريقة حساب الوسيط من التكرار المتجمع النازل .

إيجاد قيمة الوسيط بالرسم البيانى



(شكل رقم ٤٥)

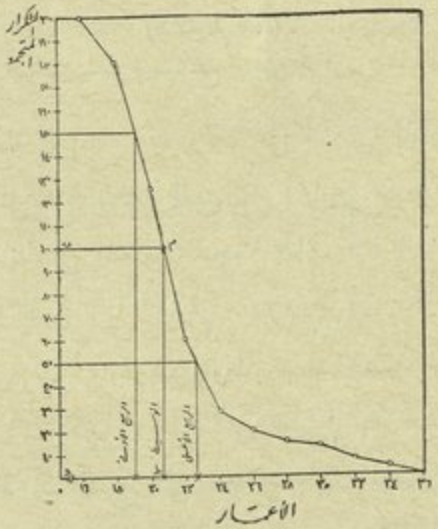
إيجاد قيمة الوسيط من المنحنى التكرارى المتجمع الصاعد

بما أن كل نقطة على المنحنى المتجمع الصاعد يكون إحداثيها الرأسى يساوى

عدد المفردات الأقل من القيمة السينية المقابلة لها ، فلو أخذنا على هذا المنحنى نقطة يكون إحداثيها الرأسى يساوى نصف مجموع التكرارات لكان عدد المفردات الأقل من القيمة السينية المقابلة لها يساوى نصف مجموع المفردات . ومن الواضح أنه ما دامت هذه القيمة السينية يسبقها نصف عدد المفردات فلا بد أن يليها النصف الآخر . وبناء على التعريف الذى أوردناه من قبل تكون هذه القيمة السينية هى قيمة الوسيط .

ومن هذا يتضح أنه من الممكن تعيين قيمة الوسيط بالرسم البيانى بالطريقة الآتية :-

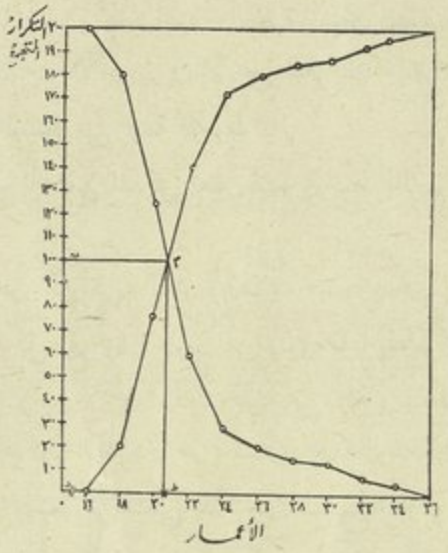
نأخذ على المحور الرأسى البعد (ون) بحيث يساوى نصف مجموع التكرارات (انظر شكل رقم ٤٥) ونرسم من (ن) عموداً يقطع المنحنى التكرارى المتجمع عند النقطة (م) . ثم نسقط عموداً من (م) فيقطع المحور الأفقى عند النقطة (ط) فيكون البعد (و ط) يساوى قيمة الوسيط . ويمكن بنفس الطريقة تعيين قيمة الوسيط من المنحنى التكرارى المتجمع النازل (شكل رقم ٤٦).



(شكل رقم ٤٦)

إيجاد قيمة الوسيط من المنحنى التكرارى المتجمع النازل

وإذا رسمنا المنحنيين التكراريين المتجمعين فى شكل واحد وأسقطنا من نقطة تقاطعهما عموداً على المحور الأفقى ، فإن هذا العمود يقطع المحور الأفقى عند قيمة الوسيط تماماً (شكل رقم ٤٧) .



(شكل رقم ٤٧)
إيجاد قيمة الوسيط من نقطة تقاطع المنحنيين التكراريين المتجمعين

حيث أن نقطة التقاطع هذه واقعة على المنحنى المتجمع الصاعد فإن إحداثيها الرأسى يمثل عدد الطلبة الذين تقل أعمارهم عن (و ط) . وبما أنها واقعة فى الوقت ذاته على المنحنى المتجمع النازل فإن إحداثيها الرأسى يمثل كذلك عدد الطلبة الذين تزيد أعمارهم عن (و ط) ، فكأن :

عددالطلبة الذين تقل أعمارهم عن (وط) = عدد الطلبة الذين تزيد أعمارهم عن (و ط) وعلى ذلك تكون (و ط) تمثل العمر الوسيط حسب التعريف الذى أوردناه . ويلاحظ أن النتائج التى نحصل عليها بالرسم البيانى تكون مماثلة تماماً للنتائج التى نحصل عليها بالحساب وإذا وجد اختلاف بسيط فإنه يرجع إلى عدم دقة الرسم .

التطبيقات العملية للوسيط :

من أهم خواص الوسيط أن قيمته لا تتأثر بوجود بعض المفردات الشاذة فى المجموعة التى نرغب فى حساب متوسطها . ولهذا السبب إذا وجدت مثل هذه المفردات فى مجموعة فالأفضل أن يكون المتوسط الذى يمثلها هو الوسيط .

ويستعمل الوسيط كذلك ، دون المتوسطات الأخرى ، فى التعبير عن متوسط الرتب . فلو رتبنا عدداً من الطلبة فى مرتبات تصاعدية حسب درجة الذكاء أو أية صفة أخرى فإن الوسيط هو أفضل قيمة للتعبير عن المرتبة المتوسطة .

وعلى الرغم من أنه إذا كانت المرتبات تتصاعد حسب الأعداد الطبيعية(١) كما هو المعتاد، فإن قيمة الوسط الحسابى تعادل قيمة الوسيط ، إلا أن استعمال هذا الأخير للتعبير عن المرتبة المتوسطة يكون أقرب إلى المنطق . فالوسط الحسابى للمرتبات يكاد يكون لا معنى له(٢) .

ويفضل الوسيط على الوسط الحسابى وغيره من المتوسطات فى حالة

---

(١) اذ كانت هذه المرتبات هى على التوالى

١ ، ٢ ، ٣ ، ٤ ، ٥ فإن الوسيط = ٣

والوسط الحسابى = $\frac{١+٢+٣+٤+٥}{٥}$ = ٣ كذلك

غير أن اتفاق النتيجتين لا يعنى أنه من الممكن استعمال الوسط الحسابى فى مثل هذه الحالات .

(٢) السبب فى ذلك أن الأعداد التى تعبر عن المرتبات تخالف فى طبيعتها الأعداد التى تعبر عن القياس . فالأعداد لها وظيفتان مختلفتان تمام الاختلاف .

الوظيفة الأولى هى التعبير عن التسلسل أو الترتيب (The Ordinal Aspect of Numbers)

والوظيفة الثانية هى التعبير عن القياس أو الكم (The Cardinal Aspect of Numbers)

وحيث أن حساب الوسط الحسابى يتضمن استخدام الوظيفة الثانية للأعداد فلا يجوز منطقياً استعماله للتعبير عن متوسط المرتبات .

انظر فى هذه النقطة كتاب Mathematical Analysis For Economists تأليف R. G. D. Allen. ص ٢ .

التوزيعات التكرارية المفتوحة أى تلك التى تكون فيها إحدى نهايتى فئة أو أكثر من فئات التوزيع غير محددة ، كما لو وضع التوزيع فى الصورة الآتية :-

فئات الدخل بالجنيهات	التكرار
أقل من ١٠٠	٥٠٠
١٠٠ ـ ٢٠٠	١٩٠٠
٢٠٠ ـ ٥٠٠	٢٦٠٠
٥٠٠ ـ ١٠٠٠	١٢٠٠
١٠٠٠ فما فوق	٨٠

وذلك لأن حساب الوسط الحسابى للتوزيع التكرارى يتطلب تعيين مراكز الفئات وما دامت نهاية الفئة غير محددة فلا يمكن تعيين مركز الفئة إلا إذا وضعنا لها نهاية تعسفية من عند ياتنا. أما حساب الوسيط فلا يتطب تعيين مراكز الفئات .

ومن أهم الحالات التى يستخدم فيها الوسيط ، حالة التجارب التى تسبب إتلاف المفردات المختبرة أثناء إجراء التجربة . فلو أردنا مثلا معرفة متوسط عمر مجموعة من « اللمبات » الكهربائية ، فبدلا من أن نحسب الوسط الحسابى ونتلف بذلك جميع اللمبات المختبرة ، نحسب الوسيط ونوقف التجربة عندما يحترق نصف عدد اللمبات فقط ، إذ من البديهى أن اللمبات التى لم تحترق حتى إيقاف التجربة تكون أعمارها أكبر من الوسيط وهو عمر آخر لمبة تحترق قبل إيقاف التجربة مباشرة . أو افرض مثلا أننا أردنا عمل تجربة على عدد من الحيوانات لمعرفة متوسط المدة التى تستطيع بقاءها فى هواء فاسد . نستطيع فى هذه الحالة أن نوقف التجربة عندما ينفق نصف عدد الحيوانات المختبرة ، إذ من الواضح أن النصف الباقى يكون طول مدة احتماله للهواء الفاسد أكبر من الوسيط .

الربيعان (QUARTILES) .

الربيع الأدنى (LOWER QUARTILE) هو القيمة التى يسبقها فى

الترتيب ربع القيم ويليها ثلاثة أرباع القيم على فرض أن هذه القيم مرتبة ترتيباً تصاعدياً .

والربيع الأعلى (UPPER QUARTILE) هو القيمة التى يسبقها ثلاثة أرباع القيم ويليها ربع القيم على فرض أن هذه القيم مرتبة ترتيباً تصاعدياً .

ويمكن تعيين قيمتى الربيعين إما بالحساب أو بالرسم البيانى بنفس الطريقة التى اتبعناها فى حالة الوسيط مع ملاحظة أن ترتيب الربيع الأدنى فى حالة التوزيعات التكرارية هو $\frac{ن}{٤}$ وترتيب الربيع الأعلى هو $\frac{٣ن}{٤}$.

فلإيجاد قيمة الربيعين للتوزيع التكرارى لأعمار الطلبة المتقدمين للالتحاق بكلية التجارة نستخدم التوزيع التكرارى المتجمع الموضح بالجدول رقم ٢١ كالاتى :—

ترتيب الربيع الأدنى = $\frac{ن}{٤} = \frac{٢٠٠}{٤} = ٥٠$

وبالرجوع إلى الجدول نجد أن الـ٢٠ قيمة الأولى جميعها أقل من ١٨ والـ٧٥ قيمة الأولى جميعها أقل من ٢٠ ، وحيث أن ترتيب الربيع الأدنى يقع بين هذين الحدين فلا بد أن تكون قيمته محصورة بين ١٨ و ٢٠ سنة . وبنفس طريقة الاستكمال التى اتبعناها فى إيجاد قيمة الوسيط نقول :—

فرق فى التكرار قدره (٧٥ — ٢٠) يناظره فرق فى الأعمار قدره (٢٠ — ١٨) .

∴ فرق فى التكرار قدره (٥٠ — ٢٠) يناظره فرق فى الأعمار قدره س

$$س = \frac{٣٠ \times ٢}{٥٥} = ١,٠٩١$$

∴ قيمة الربيع الأدنى = ١٨ + ١,٠٩١ = ١٩,٠٩١ سنة

ويمكن الحصول على نفس هذه النتيجة بالرسم البيانى باستخدام المنحنى التكرارى المتجمع الصاعد (شكل رقم ٤٥) أو باستخدام المنحنى التكرارى المتجمع النازل (شكل رقم ٤٦) .

وبالمثل يمكن حساب قيمة الربيع الأعلى :-

$$\text{ترتيب الربيع الأعلى} = \frac{٣ن}{٤} = \frac{٣ \times ٢٠٠}{٤} = ١٥٠$$

وبالرجوع إلى الجدول رقم ٢١ نجد أن قيمة الربيع الأعلى لا بد أن تكون محصورة بين ٢٢ و ٢٤ سنة وكما فعلنا فى حالة الوسيط والربيع الأدنى نقول :-

فرق فى التكرار قدره (١٧٢ — ١٤٠) يناظره فرق فى الأعمار قدره (٢٤ — ٢٢) .

∴ فرق فى التكرار قدره (١٥٠ — ١٤٠) يناظره فرق الأعمار قدره س

$$\text{س} = \frac{١٠ \times ٢}{٣٢} = ٠,٦٢٥$$

∴ قيمة الربيع الأعلى = ٢٢ + ٠,٦٢٥ = ٢٢,٦٢٥ سنة

ويمكن الحصول على نفس هذه النتيجة من المنحنى المتجمع الصاعد (شكل ٤٥) أو من المنحنى المتجمع النازل (شكل ٤٦) . ومما تجب ملاحظته أن الربيعين ليسا من مقاييس التوسط وإنما أوردناهما فى هذا الباب لارتباطهما بالوسيط . أما من جهة تطبيقاتها العملية فإنهما يستعملان فى قياس التشتت والالتواء على النحو الذى سنفصله فى البابين التاليين .

وهناك مقاييس أخرى مرتبطة بفكرة الوسيط منها العشير (Decile) وهو القيمة التى يسبقها $\frac{١}{١٠}$ القيم ويليها $\frac{٩}{١٠}$ القيم . والمئين (Percentile) وهو القيمة التى يسبقها $\frac{١}{١٠٠}$ من القيم ويليها $\frac{٩٩}{١٠٠}$ من القيم على فرض أن هذه القيم مرتبة ترتيباً تصاعدياً . غير أن هذه المقاييس الإحصائية قليلة الأهمية ونادرة الاستعمال ولذلك نكتفى بمجرد الإشارة إليها .

٦ ــ المنوال(١)

المنوال هو القيمة النموذجية التى يتركز حولها التكرار أو بعبارة أخرى فإن المنوال هو القيمة التى يكون تكرارها أعظم من تكرار أية قيمة أخرى . فلو فرض أن عشرة عمال يتقاضون الأجور اليومية الآتية :ــ

٩ ، ١٠ ، ١٢ ، ١٢ ، ١٢ ، ١٣ ، ١٤ ، ١٥ ، ١٩ ، ٢٠ قرشاً

فإن الأجر الأكثر شيوعاً أو الأجر النموذجى هو ١٢ قرشاً فى اليوم وهذه هى قيمة المنوال .

أما إذا كانت القيم كالآتى :ــ

٩ ، ١٠ ، ١٢ ، ١٣ ، ١٤ ، ١٩ ، ٢٠ فليس لها قيمة منوالية .

إيجاد المنوال للتوزيع التكرارى :

بالنظر إلى الجدول التكرارى الذى يمثل توزيع أعمار الطلبة المتقدمين للالتحاق بكلية التجارة (جدول رقم ٢١) نجد أن المنوال لا بد أن يكون واقعاً فى الفئة (٢٠ ــ ٢٢) لأنها الفئة التى يقابلها أكبر تكرار ، وتسمى الفئة المنوالية .

وإذا كانت فئات التوزيع متساوية يمكن تعيين الفئة المنوالية بمجرد النظر إلى الجدول التكرارى . أما إذا كانت الفئات مختلفة فى الطول فلا بد من تعديل

---

(١) يفضل البعض استعمال كلمة « النمط » بينما يفضل البعض الآخر استعمال كلمة « القمة » ونحن لا نجد فائدة من تعدد هذه المصطلحات بغض النظر عن أيهما أقرب إلى التعبير الأجنبى (Mode) خصوصاً وأن كلمة المنوال قد اكتسبت قوة الاستعمال . غير أننا لا نستطيع إلا أن نلاحظ أن استعمال كلمة « القمة » لا بد وأن يكون راجعاً إلى عدم فهم معنى المنوال . فالمنوال هو القيمة السينية الواقعة تحت قمة المنحنى التكرارى وليس القمة ذاتها .

التكرارات قبل تعيين الفئة المنوالية[١].

ومتى عينا الفئة المنوالية فيمكننا ، إذا أردنا الحصول على نتيجة تقريبية سريعة ، أن نعتبر قيمة المنوال تعادل مركز هذه الفئة . وعلى ذلك تكون قيمة المنوال للتوزيع التكرارى السابق الذكر هى ٢١ سنة ، غير أن هذه الطريقة تقريبية ولا يجوز استعمالها إذا أردنا الحصول على نتائج دقيقة .

فلو نظرنا إلى تكرار الفئة قبل المنوالية والفئة بعد المنوالية (جدول رقم ٢١) لاتضح لنا أن تركز التكرار يميل إلى اتجاه الفئة قبل المنوالية . أى أن موقع المنوال فى الفئة المنوالية يجب أن يكون أقرب إلى الحد الأدنى لهذه الفئة منه إلى الحد الأعلى .

وهناك طريقتان تقريبيتان أخريان[٢] لتعيين موقع المنوال داخل الفئة المنوالية كلتاهما تعتمد على مقارنة تكرارى الفئة قبل المنوالية والفئة بعد المنوالية ونلخصهما فيما يلى :—

أولا — طريقة الفروق

هذه الطريقة من وضع « كارل بيرسون » وتسمى أحياناً طريقة بيرسون وتتلخص فى أن يحدد موقع المنوال فى الفئة المنوالية بحيث يقسم هذه الفئة إلى قسمين متناسبين مع الفرق بين تكرار الفئة المنوالية والفئة قبل المنوالية من جهة

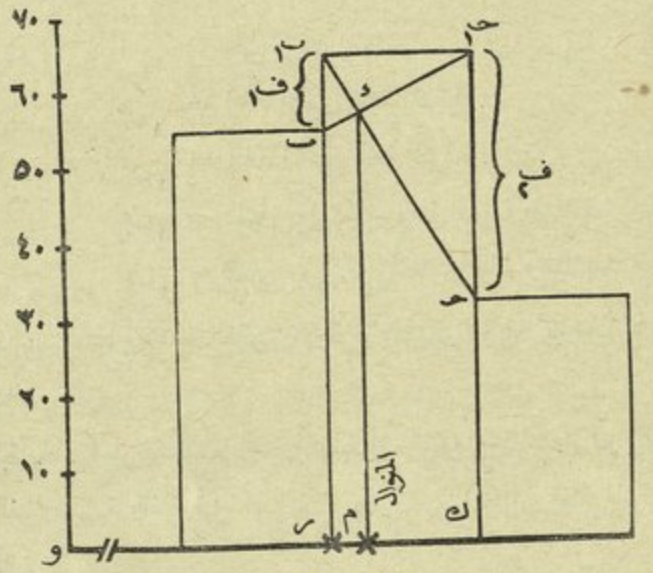
---

(١) راجع ص ٨٨ — ٩١ من هذا الكتاب .

(٢) الطرق الدقيقة لإيجاد المنوال تتطلب حساب عزوم التوزيع التكرارى وسنشير إليها فى الباب الخاص بالعزوم . ويلاحظ أنه إذا كان التوزيع التكرارى يمثله منحن متصل تصفه معادلة جبرية فيمكن إيجاد قيمة المنوال بدقة تامة باستخدام نظرية النهايات الكبرى والصغرى فى دراسة التفاضل . فمن الواضح أن قيمة المنوال هى قيمة س المقابلة للنهاية الكبرى لقيمة ص على فرض أن س تمثل المتغير و ص تمثل التكرار . وبالمثل فان الوسيط هو القيمة السينية المقابلة للاحداثى الصادى الذى يقسم المنحنى التكرارى إلى قسمين متساويين فى المساحة . أما قيمة الوسط الحسابى فهى القيمة السينية المقابلة للاحداثى الصادى الذى يمر بمركز المنحنى (Centroid)

والفرق بين تكرار الفئة المنوالية والفئة بعد المنوالية من الجهة الأخرى .

ولبيان ذلك نرسم ذلك الجزء من المدرج التكرارى الذى يبين تكرار الفئة المنوالية والفئة السابقة لها والفئة التى تليها فيكون موقع المنوال حسب طريقة بيرسون عند النقطة م (انظر شكل ٤٨) .



(شكل رقم ٤٨)

إيجاد المنوال بالرسم بطريقة الفروق

يتضح من الرسم أن المثلثين ء ح ح١ ، ء ب ب١ متشابهان

$$\therefore \frac{ب ب_1}{ح ح_1} = \frac{ب ء}{ح_1 ء} = \frac{م ر}{م ك}$$

$$\therefore \frac{م ر}{م ك} = \frac{ف_1}{ف_2}$$

أى أن الفئة المنوالية مقسمة بنسبة $\frac{ف_1}{ف_2}$

أى بنسبة $\frac{\text{الفرق بين تكرار الفئة المنوالية والفئة قبل المنوالية}}{\text{الفرق بين تكرار الفئة المنوالية والفئة بعد المنوالية}}$

ويمكننا تعيين قيمة المنوال بطريقة بيرسون دون حاجة إلى رسم المدرج التكرارى وذلك باستخدام القانون الآتى :-

$$\text{المنوال} = س_1 + \frac{ف_1}{ف_1 + ف_2} \times ا$$

حيث $ا$ = طول الفئة المنوالية

$س_1$ = الحد الأدنى للفئة المنوالية .

$س_2$ = الحد الأعلى للفئة المنوالية .

$ف_1$ = الفرق بين تكرار الفئة المنوالية والفئة قبل المنوالية .

$ف_2$ = الفرق بين تكرار الفئة المنوالية والفئة بعد المنوالية .

ويمكن استنتاج هذا القانون مباشرة بتطبيق طريقة بيرسون فى تقسيم الفئة المنوالية .

فطبقاً لهذه الطريقة نقسم الفئة المنوالية إلى قسمين متناسبين مع الفرق بين تكرار الفئة المنوالية والفئة قبل المنوالية ، والفرق بين تكرار الفئة المنوالية والفئة بعد المنوالية .

أى أن $\frac{م - س_1}{س_2 - م} = \frac{ف_1}{ف_2}$

$\therefore$ $م ف_2 - س_1 ف_2 = س_2 ف_1 - م ف_1$

$\therefore$ $م (ف_1 + ف_2) = س_2 ف_1 + س_1 ف_2$

ولكن $س_2 = س_1 + ا$

$\therefore$ $م (ف_1 + ف_2) = ف_1 (س_1 + ا) + س_1 ف_2$

$= ف_1 س_1 + س_1 ف_2 + ف_1 ا$

$= س_1 (ف_1 + ف_2) + ف_1 ا$

بقسمة الطرفين على $(ف_1 + ف_2)$.

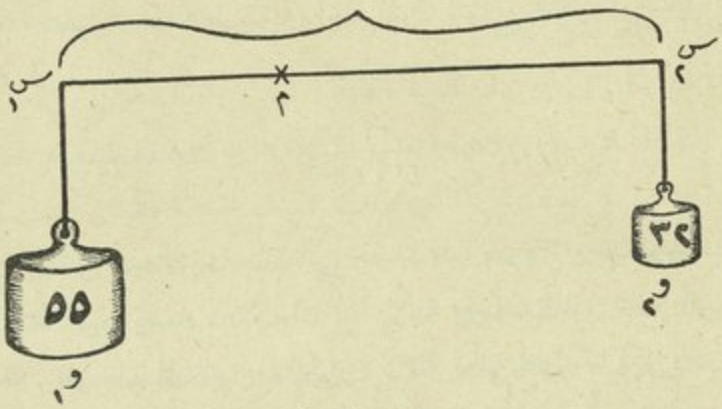
$$\therefore \quad م(*) = س_{١} + ا \times \frac{ف_{١}}{ف_{١} + ف_{٢}}$$

وبتطبيق هذا القانون على التوزيع التكرارى لأعمار الطلبة المتقدمين للالتحاق بكلية التجارة نجد أن :ـ

$$\text{المنوال} = ٢٠ + ٢ \times \frac{١٠}{١٠ + ٣٣} = ٢٠,٤٦٥ \text{ سنة}$$

ثانياً ـ طريقة الرافعة

الفكرة فى هذه الطريقة هى أن نتصور أن الفئة المنوالية هى عبارة عن رافعة من النوع الأول وأن تكرار الفئة قبل المنوالية عبارة عن القوة وتكرار الفئة بعد المنوالية عبارة عن المقاومة . أى أننا سنفرض أن كلا من تكرار الفئة قبل المنوالية (القوة) وتكرار الفئة بعد المنوالية (المقاومة) تتجاذب المنوال إليها وعلى هذا الأساس يكون المنوال عند محور الارتكاز كما يتضح من الشكل الآتى :



(شكل رقم ٤٩)
رسم يوضح طريقة الرافعة فى إيجاد قيمة المنوال

(*) لاحظ أن م وكلا من س١ ، س٢ تمثل أبعاداً (أو قيما) وليس مجرد نقط .

وحيث أن القوة × ذراعها = المقاومة × ذراعها

$\therefore$ $و_١ (م - س_١) = و_٢ (س_٢ - م)$

$\therefore$ $و_١ م - و_١ س_١ = و_٢ س_٢ - و_٢ م$

$\therefore$ $م (و_١ + و_٢) = و_٢ س_٢ + و_١ س_١$

ولكن $س_٢ = س_١ + ١$

$\therefore$ $م (و_١ + و_٢) = و_٢ (س_١ + ١) + و_١ س_١$

$= س_١ (و_١ + و_٢) + و_٢ ١$

وبالقسمة على $(و_١ + و_٢)$

$\therefore$ $م = س_١ + ١ \times \frac{و_٢}{و_١ + و_٢}$

وبتطبيق هذه الطريقة فى إيجاد المنوال للتوزيع التكرارى السابق الذكر نجد أن

$$م = ٢٠ + ٢ \times \frac{٣٢}{٣٢ + ٥٥} = ٢٠,٧٣٥ \text{ سنة}$$

وبمقارنة هذه النتيجة بتلك التى حصلنا عليها بطريقة بيرسون نجد أن هناك اختلافاً كبيراً بينهما . ولا شك أن طريقة بيرسون تعطينا نتائج أقرب إلى الدقة . فلو نظرنا إلى التوزيع التكرارى لأعمار الطلبة لظهر لنا بكل وضوح أن تركز الأعمار هو فى اتجاه الحد الأدنى للفئة المنوالية . وهذا يؤيد النتيجة التى حصلنا عليها بطريقة بيرسون . أما طريقة الرافعة فإنها تبتعد بقيمة المنوال عن الحد الأدنى للفئة المنوالية مما يخالف اتجاه تركز التكرار .

ولنأخذ مثالا آخر يوضح أفضلية طريقة بيرسون . افرض أننا أردنا إيجاد قيمة المنوال للتوزيع التكرارى الآتى :-

الفئات	التكرار
١٥ – ٢٠	٣٠
٢٠ – ٢٥	٧٠
٢٥ – ٣٠	٦٩
٣٠ – ٣٥	٢٨

يتضح بمجرد النظر إلى هذا التوزيع أن المنوال قريب جداً من الحد الأعلى للفئة المنوالية . وبتطبيق طريقة بيرسون نحصل على النتيجة الآتية :-

$$\text{المنوال} = ٢٠ + \frac{(٧٠ - ٣٠)}{(٧٠ - ٣٠) + (٧٠ - ٦٩)} \times ٥ = ٢٤,٨٨$$

وبتطبيق طريقة الرافعة نجد أن :-

$$\text{المنوال} = ٢٠ + \frac{٦٩}{٦٩ + ٣٠} \times ٥ = ٢٣,٤٨$$

ومن الواضح أن النتيجة الأولى أقرب إلى الصحة .

حالات استعمال المنوال .

ما من شك فى أنه إذا اكتفينا بالطرق التقريبية لحساب المنوال فإنه يكون أسهل المتوسطات جميعاً من وجهة العمل الحسابى . غير أن الطرق الدقيقة لحسابه يصعب فهمها على المبتدىء كما أنها تتطلب عمليات حسابية مطولة . ولذلك فإن اختيار المنوال دون المتوسطات الأخرى لتمثيل المجموعات يتوقف إلى حد كبير على درجة الدقة التى نتوخاها فى النتائج التى نحصل عليها . فإذا أردنا الحصول على نتائج تقريبية فالمنوال يعتبر أفضل المتوسطات الممثلة للمجموعة .

والمنوال كالوسيط لا يتأثر بوجود بعض القيم الشاذة بين مفردات المجموعة كما أن وجود فئات تكرارية مفتوحة لا يحول دون حسابه . والمنوال يفضل الوسط الحسابى كثيراً فى حالة التوزيعات التكرارية الملتوية التواء شديداً

إذ أن دقة تمثيل الوسط الحسابى للمجموعة تقل كلما زاد التواء المنحنى .

ويقترح بعض الإحصائيين تفضيل المنوال بوجه عام فى كل حالة تختلف فيها قيمته كثيراً عن الوسط الحسابى . والمنوال هو المتوسط الوحيد الذي يمكن استخدامه لإيجاد متوسط الظواهر التى لا يمكن التعبير عنها تعبيراً رقمياً مثل الصفات . إذ يمكن فى مثل هذه الحالات اعتبار القيمة المنوالية الصفة الأكثر شيوعاً بينما لا يمكن إيجاد الصفة المتوسطة بأية طريقة أخرى .

وحيث أن المنوال هو القيمة النموذجية التى يتركز حولها التكرار فإنه يحمل صفات عدد كبير من مفردات المجموعة التى يمثلها ولذلك فهو أقرب إلى معنى « المتوسط » من جميع المتوسطات الأخرى . ومع هذا فإن استعماله قليل ويرجع ذلك إلى الأسباب الآتية :—

١ — عدم قابليته للمعالجة الجبرية .

٢ — بعض التوزيعات التكرارية لها قيمة منوالية واحدة بينما البعض الآخر له قيمتان أو أكثر . فبعض المنحنيات التكرارية لها قمتان أو أكثر .

٣ — لا يمكن الاعتماد على قيمته إلا إذا كان عدد مفردات المجموعة كبيراً فالمجموعات الصغيرة ليس لها عادة قيمة منوالية يصح الاعتماد عليها .

٤ — يتأثر المنوال كثيراً بطريقة اختيار الفئات التكرارية ولذلك فإننا نحصل على قيم مختلفة للمنوال إذا اخترنا فئات مختلفة لنفس التوزيع .

على أنه إذا عينا قيمة المنوال بالطرق الدقيقة فإن معظم هذه الانتقادات تزول بينما يظل المنوال أقرب إلى تمثيل المجموعة من أى متوسط آخر .

العلاقة بين الوسط الحسابى والوسيط والمنوال .

من الواضح أنه إذا كان التوزيع التكرارى متماثلا فإن قيم كل من الوسط الحسابى والوسيط والمنوال تكون متساوية . وذلك لأنه فى التوزيعات المتماثلة تكون القيمة الوسطى هى أيضاً القيمة النموذجية التى يقابلها أكبر تكرار

ويقع الوسط الحسابى فى مركز الفئة الوسطى تماماً . كذلك يكون الربيعان على بعدين متساويين من الوسيط .

أما فى حالة التوزيعات الملتوية فإن قيم هذه المتوسطات تختلف ، كما أن الربيعين لا يكونان على بعدين متساويين من الوسيط . وقد وجد بيرسون بالمشاهدة المتكررة ، أنه فى حالة المنحنيات الملتوية التواء بسيطاً فإن العلاقة الاتية تربط المتوسطات الثلاثة :ـ

$$\overline{س} - م = ٣ (\overline{س} - و)$$

$$\text{أو } م = \overline{س} - ٣ (\overline{س} - و)$$

حيث $\overline{س}$ تمثل الوسط الحسابى

، م تمثل المنوال

، و تمثل الوسيط

ومن الواضح أنه يمكن استخدام هذه الطريقة فى إيجاد قيمة تقريبية للمنوال . وبتطبيقها على التوزيع التكرارى لأعمار الطلبة نجد أن :ـ

المنوال = ٢١,٥٢ ـ ٣ (٢١,٥٢ ـ ٢٠,٧٦٩) = ١٩,٤٦٧ سنة

ومن الواضح أن هذه الطريقة لا تصلح فى هذه الحالة لأن التوزيع ملتو التواء شديداً نحو اليمين .

الباب السادس

مقاييس التشتت

Variation

ذكرنا أن المتوسطات هى مجموعة المقاييس الوصفية الأولى التى تعطينا فكرة أولية عن التوزيع التكرارى. ومن الواضح أن وصف التوزيع التكرارى بأحد المتوسطات يعطينا فكرة غير تامة عن حقيقة المجموعة التى يمثلها التوزيع ، كما أن المقارنة بين المجموعات بناء على متوسطاتها فقط تكون ناقصة كذلك إن لم تكن مضللة فعلا . فقد يحدث أن يتساوى متوسطا مجموعتين ومع ذلك تكون مفرداتهما مختلفة كل الاختلاف . فربما تكون مفردات المجموعة الأولى قريبة فى القيمة من متوسطها أى مركزة حوله ، بينما تكون مفردات المجموعة الثانية متفاوتة فى القيمة وتختلف كثيراً عن متوسطها ، فيكون بعضها أكبر منه بكثير والآخر أقل منه بكثير .

افرض مثلا أننا أردنا مقارنة المجموعتين الآتيتين من درجات الحرارة :—

المدينة « ا » ٣٧° ، ٣٨° ، ٣٩° ، ٤٠° ، ٤١° .

المدينة « ب » ١٣° ، ٢٤° ، ٤٧° ، ٤٩° ، ٦٢° .

بإيجاد الوسط الحسابى لكل من هاتين المجموعتين نجد أنه يساوى ٣٩° فى الحالتين . غير أنه بالنظر إلى الأرقام ذاتها يتضح لنا أن المجموعتين مختلفتان اختلافاً بيناً من حيث مدى تقلبات درجة الحرارة . بحيث لا يجوز مطلقاً أن نعتبر أن طقس المدينتين واحد أو حتى متقارب . فدرجة الحرارة فى المدينة الأولى لا تختلف كثيراً عن الدرجة المتوسطة وهى ٣٩° فى حين أنها فى المدينة الثانية

تتقلب فى مدى واسع يمتد من ١٣° إلى ٦٢° . وبديهى أنه لو اكتفينا بمقارنة درجة الحرارة المتوسطة فى كل من المدينتين لرسمنا صورة خاطئة جداً لحقيقة الحالة .

كذلك إذا اختبرنا مجموعتين من اللمبات الكهربائية ووجدنا متوسط عمر اللمبة فى المجموعة الأولى ١٠٠٠ ساعة وفى المجموعة الثانية ١٠٠٠ ساعة كذلك ، فإن هذا لا يدل فى حد ذاته على أن لمبات المجموعتين متساوية فى الجودة . فقد تكون أعمار اللمبات فى المجموعة الأولى تتراوح بين ٧٥٠ و ١٢٠٠ فى حين أنها فى المجموعة الثانية تتغير فى مدى واسع يمتد بين ١٠٠ و ٣٠٠٠ .

لكى نجعل وصف المجموعات والمقارنة بينها أقرب إلى الدقة إذن ، نضيف إلى مقاييس التوسط التى تصف هذه المجموعات مقاييس أخرى تظهر درجة اختلاف مفردات هذه المجموعات عن بعضها أو عن متوسطاتها أو بعبارة أخرى تظهر طريقة توزيعها أو تشتتها إما فيما بينها أو حول أى مركز نحدده وهو عادة أحد المتوسطات .

وتسمى هذه المجموعة من المقاييس الوصفية ، مقاييس التشتت ونذكرها فيما يلى : —

١ — المدى Range

٢ — الانحراف الربيعى Quartile Deviation

٣ — الانحراف المتوسط Mean Deviation

٤ — الانحراف المعيارى Standard Deviation

---

(١) يفضل البعض تسمية هذا القياس « الانحراف القياسى » . وعلى الرغم من أنى أميل كذلك إلى تفضيل هذه التسمية إلا أنى سأستعمل فى هذا الكتاب اصطلاح الانحراف المعيارى الذى صاغه الدكتور عبد المنعم الشافعى وشاع استعماله كثيراً ، وذلك حرصاً على توحيد المصطلحات العربية .

١ ــ المدى

المدى هو أبسط مقاييس التشتت سواء من جهة الحساب أو من جهة الفكرة . وهو عبارة عن الفرق بين أكبر قيمة وأصغر قيمة فى المجموعة . فإذا رجعنا إلى المثال الخاص بمقارنة درجات الحرارة فى المدينتين « ا » ، « ب » نجد أن المدى فى المدينة الأولى = ٤١ ــ ٣٧ = ٤ درجات

والمدى فى المدينة الثانية = ٦٢ ــ ١٣ = ٤٩ درجة

وهذا المقياس مع بساطته وسهولة فهمه لا يعد بين مقاييس التشتت الهامة فهو مقياس تقريبى ولا يستعمل كثيراً للأسباب الاتية :ــ

أولا : يتأثر المدى بالقيم المتطرفة فقط ولا يأخذ فى الحسبان كافة القيم الأخرى . فإذا تصادف أن كانت أصغر قيمة وأكبر قيمة فى المجموعة قيمتين شاذتين أصبح المدى لا قيمة له إطلاقاً كمقياس للتشتت .

ثانياً : لا يمكن حساب المدى بدقة من الجداول التكرارية لا سيما إذا كانت تلك الجداول تحتوى على فئات مفتوحة . ولا بد لمعرفة المدى من الرجوع إلى الأرقام الأولية .

ثالثاً : فى حالة مقارنة مجموعتين من القيم عدد مفرداتهما مختلف اختلافاً كبيراً يكون المدى مقياساً مضللا للتشتت ، فإن زيادة عدد المفردات فى المجموعة يزيد من احتمال وجود قيم شاذة إما صغيرة جداً أو كبيرة جداً . وبالمثل إذا كانت المجموعتان مقدرتان بوحدات مختلفة فإن المدى يعطينا صورة خاطئة عن درجة تشتت كل مجموعة بالنسبة للأخرى .

المدى بين المئين العاشر والتسعين .

ينحصر اهتمامنا أحياناً فى معرفة المدى الذى تقع بينه نسبة معينة من المفردات وليس المفردات كلها (كما هو الحال فى المدى العادى) . فمثلا عند

عمل كثير من الإحصاءات التى تصف نواحى النشاط المدرسى كوصف مجموعة من الطلبة من حيث الذكاء أو المواظبة ، تهمل الـ ١٠٪ الأولى والـ ١٠٪ الأخيرة من المفردات باعتبارها حالات متطرفة أو شاذة . ويقاس التشتت على أساس المدى بين أصغر قيمة وأكبر قيمة من المفردات الباقية أى على أساس المدى بين المئين العاشر والتسعين .

٢ – الانحراف الربيعى

الانحراف الربيعى هو نصف المدى بين الربيعين[١]. فلو رمزنا للربيع الأعلى بالرمز $ع_{١}$ والربيع الأدنى بالرمز $ع_{٣}$ فإن :–

$$\text{الانحراف الربيعى} = \frac{ع_{٣} - ع_{١}}{٢}$$

وعلى ذلك يكون الانحراف الربيعى لأعمار الطلبة المتقدمين للالتحاق بكلية التجاره سنة ١٩٤٧ $= \frac{٢٢,٦٢٥ - ١٩,٠٩١}{٢} = ١,٧٦٧$

والانحراف الربيعى ، كالمدى ، لا يعتبر من مقاييس التشتت الدقيقة ، فهو مقياس تقريبى ولا يجوز استعماله إلا لتكوين فكرة مبدئية سريعة عن تشتت المفردات . ومن الواضح أن الانحراف الربيعى أفضل جداً من المدى حيث أنه لا يتأثر كثيراً بالقيم الشاذة التى تقع فى طرفى التوزيع .

٣ – الانحراف المتوسط

الانحراف المتوسط هو متوسط انحرافات قيم المفردات عن وسطها الحسابى بغض النظر عن إشارات الانحرافات[٢] .

---

(١) يلاحظ القارئ أن الانحراف الربيعى هو فى الواقع نوع من المدى ويسميه بعض الكتاب الأجانب Semi-Inter Quartile Range. بدلا من Quartile Deviation.

(٢) يذكر القارئ أننا أثبتنا من قبل أن المجموع الجبرى لانحرافات القيم عن وسطها الحسابى يساوى صفر وبالتالى يكون متوسط هذه الانحرافات أى $\frac{\Sigma ح}{ن}$ = صفر . غير أنه إذا حسبنا متوسط الانحرافات المجردة عن إشاراتها فان هذا المتوسط لا يساوى صفراً .

وعلى ذلك فإن الانحراف المتوسط = $\frac{\Sigma |ح|}{ن}$ (حيث ح تمثل الانحرافات)

فإذا كانت مفردات المجموعة هى

٣ ، ٦ ، ٧ ، ٩ ، ١٥

فإن وسطها الحسابى = ٨

وانحرافاتها عن الوسط الحسابى هى

— ٥ ، — ٢ ، — ١ ، ١ ، ٧

ومتوسط هذه الانحرافات المجردة عن إشاراتها (أى الانحراف المتوسط)

= $\frac{١٦}{٥}$ = ٣,٢

وفى حالة التوزيع التكرارى يكون الانحراف المتوسط = $\frac{\Sigma |ح و|}{ن}$

(حيث وتمثل التكرارات). ويتضح مما تقدم أن أول خطوة فى حساب قيمة الانحراف المتوسط هى إيجاد الوسط الحسابى. ويلى ذلك تعيين قيم الانحرافات (مع إهمال إشاراتها) ثم ضربها فى التكرارات المقابلة لها.

وبقسمة مجموع حواصل الضرب على عدد القيم يكون الناتج هو الانحراف المتوسط وتتضح هذه الخطوات من الجدول الآتى :—

---

(٥) الرمز |ح| يدل على أن إشارات قيم ح مهملة.

جدول رقم ٢٢

إيجاد الانحراف المتوسط لأعمار الطلبة المتقدمين للالتحاق بكلية التجارة سنة ١٩٤٧ (*)

فئات الأعمار	التكرار و	مراكز الفئات	ا ح ا	ا ح و ا
١٦ — ١٨	٢٠	١٧	٤,٥٢	٩٠,٤٠
١٨ — ٢٠	٥٥	١٩	٢,٥٢	١٣٨,٦٠
٢٠ — ٢٢	٦٥	٢١	٠,٥٢	٣٣,٨٠
٢٢ — ٢٤	٣٢	٢٣	١,٤٨	٤٧,٣٦
٢٤ — ٢٦	٨	٢٥	٣,٤٨	٢٧,٨٤
٢٦ — ٢٨	٥	٢٧	٥,٤٨	٢٧,٤٠
٢٨ — ٣٠	٢	٢٩	٧,٤٨	١٤,٩٦
٣٠ — ٣٢	٦	٣١	٩,٤٨	٥٦,٨٨
٣٢ — ٣٤	٣	٣٣	١١,٤٨	٣٤,٤٤
٣٤ — ٣٦	٤	٣٥	١٣,٤٨	٥٣,٩٢
المجموع	٢٠٠			٥٢٥,٦٠

$$\text{الانحراف المتوسط} = \frac{\Sigma\, ا\, ح\, و\, ا}{\Sigma\, و} = \frac{٥٢٥,٦}{٢٠٠} = ٢,٦٢٨ \text{ سنة}$$

ويمكن إيجاد الانحراف المتوسط للتوزيع التكرارى باستخدام وسط فرضى كما فعلنا فى حالة الوسط الحسابى ، غير أن هذه الطريقة معقدة نوعاً ولا تؤدى إلى تسهيل العمل الحسابى كثيراً . وكل ما يتصل بها من مزايا هو أنها تغنى عن ضرب الانحرافات عن الوسط الحسابى (التى تكون بها أرقام كسرية عادة) فى التكرارات المقابلة لها . غير أنها فى مقابل ذلك معرضة للأخطاء خصوصاً إذا كانت الفئة التى يقع فيها الوسط الفرضى مخالفة لتلك التى يقع

(*) لاحظ أن $\bar{س}$ = ٢١,٥٢ سنة .

فيها الوسط الحسابى . ولذلك فإننا لا نشير على الطالب المبتدىء باتباعها . والجدول الآتى يوضح هذه الطريقة :—

جدول رقم ٢٣

إيجاد الانحراف المتوسط للتوزيع التكرارى باستخدام وسط فرضى

فئات الأعمار	التكرار و	مراكز الفئات	الانحرافات عن الوسط الفرضى ء	| ء و |
١٦ — ١٨	٢٠	١٧	— ٤	٨٠
١٨ — ٢٠	٥٥	١٩	— ٢	١١٠
٢٠ — ٢٢	٦٥	٢١	٠	٠
٢٢ — ٢٤	٣٢	٢٣	٢	٦٤
٢٤ — ٢٦	٨	٢٥	٤	٣٢
٢٦ — ٢٨	٥	٢٧	٦	٣٠
٢٨ — ٣٠	٢	٢٩	٨	١٦
٣٠ — ٣٢	٦	٣١	١٠	٦٠
٣٢ — ٣٤	٣	٣٣	١٢	٣٦
٣٤ — ٣٦	٤	٣٥	١٤	٥٦
المجموع ...	٢٠٠			٤٨٤

يلاحظ أن $\frac{٤٨٤}{٢٠٠}$ أى $\left(\frac{\Sigma |ء و|}{\Sigma و}\right)$ هى عبارة عن الانحراف المتوسط عن وسط فرضى قدره ٢١ .

ولإيجاد الانحراف المتوسط عن الوسط الحسابى نجرى التعديل الآتى :— بالنظر إلى الجدول نجد أن مراكز الفئات الأقل من الوسط الحسابى هى ١٧ ، ١٩ ، ٢١ كما نجد أن مجموع المفردات الواقعة فى هذه

الفئات هو ٢٠ + ٥٥ + ٦٥ = ١٤٠

وحيث أنه عند حساب الانحرافات طرحنا مراكز الفئات من ٢١ بدلا من ٢١,٥٢ كما كان هو الواجب ، فكأننا أنقصنا كل انحراف بمقدار ٠,٥٢

وحيث أن عدد الانحرافات = عدد المفردات فكأننا أنقصنا جملة الانحرافات × التكرار بمقدار ٠,٥٢ × ١٤٠ = ٧٢, ٨٠

كذلك نجد بالنظر إلى الجدول أن مراكز الفئات الأكبر من الوسط الحسابى هى :

٢٣ ، ٢٥ ، ٢٧ ، ٢٩ ، ٣١ ، ٣٣ ، ٣٥

ومجموع المفردات الواقعة فيها هو :

٣٢ + ٨ + ٥ + ٢ + ٦ + ٣ + ٤ = ٦٠

وحيث أنه عند حساب الانحرافات طرحنا من مراكز الفئات هذه ٢١ بدلا من ٢١,٥٢ كما كان هو الواجب ، فكأننا غالينا فى تقدير كل انحراف بمقدار ٠,٥٢ وبناء عليه يكون الرقم الذى يمثل مجموع حواصل ضرب الانحرافات فى التكرار مغالى فيه بمقدار ٠,٥٢ × ٦٠ = ٣١,٢

وعلى ذلك فالرقم ٤٨٤ الذى يمثل مجموع حواصل ضرب الانحرافات فى التكرارات المقابلة لها أقل من اللازم بمقدار ٧٢,٨ وأكثر من اللازم بمقدار ٣١,٢ . ولكى نصل إلى الرقم الصحيح الذى يمثل مجموع حواصل ضرب الانحرافات عن الوسط الحسابى فى التكرارات المقابلة لها ، نضيف إلى ٤٨٤ المقدار ٧٢,٨ ونطرح منه المقدار ٣١,٢ .

فتكون $\Sigma |ح و|$ = ٤٨٤ + ٧٢,٨ − ٣١,٢ = ٥٢٥,٦

ومن هذا نستنتج أن الانحراف المتوسط = $\frac{٥٢٥,٦}{٢٠٠}$ = ٢,٦٢٨ سنة

وهى نفس النتيجة التى حصلنا عليها بالطريقة المباشرة أى بدون استخدام وسط فرضى .

وسيجد القارئ من مراجعة خطوات العمل فى هذه الطريقة ما يؤيد ما

ذكرناه عن كونها معقدة ومعرضة للخطأ . والواقع أن كل المجهود الذى وفرناه فى حساب انحرافات صحيحة بدلا من انحرافات كسرية أضعناه ثانية فى عمل تلك التعديلات التى يحتمل جداً أن يتسرب الخطأ إليها .

وربما يسهل تطبيق هذه الطريقة قليلا إذا ما وضعنا العلاقة بين مجموع الانحرافات عن الوسط الحسابى ومجموع الانحرافات عن الوسط الفرضى فى صورة معادلة يسهل تذكرها كالآتى :-

$$\Sigma |ح\ و| = \Sigma |ء\ و| + (\bar{س} - \bar{س}_ء)(و_١ - و_٢)$$

حيث $\Sigma |ح\ و|$ = مجموع الانحرافات المطلقة عن الوسط الحسابى (الانحرافات × التكرارات) .

، $\Sigma |ء\ و|$ = مجموع الانحرافات المطلقة عن الوسط الفرضى (الانحرافات × التكرارات) .

، $\bar{س}$ = الوسط الحسابى

، $\bar{س}_ء$ = الوسط الفرضى

، $و_١$ = عدد المفردات الأقل من $\bar{س}$

، $و_٢$ = عدد المفردات الأكبر من $\bar{س}$

وبتطبيق هذه العلاقة على المثال السابق نجد أن :-

$$\Sigma |ح\ و| = ٤٨٤ + (٢١,٥٢ - ٢١)(١٤٠ - ٦٠)$$

$$= ٤٨٤ + ٤١,٦$$

$$= ٥٢٥,٦$$

∴ الانحراف المتوسط = $\frac{٥٢٥,٦٠}{٢٠٠}$ = ٢,٦٢٨

٤ – الانحراف المعيارى

الانحراف المعيارى لمجموعة من القيم هو عبارة عن الوسط التربيعى لانحرافات هذه القيم عن وسطها الحسابى . وبعبارة أخرى فهو الجذر التربيعى لمتوسط مربعات الانحرافات عن الوسط الحسابى . فإذا كانت انحرافات ن من القيم عن وسطها الحسابى هى على التوالى :–

ح١ ، ح٢ ، ح٣ ، ..... حن

فإن الإنحراف المعيارى ويرمز له عادة بالحرف σ = $\sqrt{\frac{\Sigma ح^{٢}}{ن}}$ (*)

أما مربع هذه الكمية أى $\sigma^{٢}$ أو $\frac{\Sigma ح^{٢}}{ن}$ فيمسى التباين(١) (Variance)

وفى حالة التوزيع التكرارى فإن الانحراف المعيارى = $\sqrt{\frac{\Sigma ح^{٢} و}{\Sigma و}}$

(حيث و تمثل التكرار) .

والانحراف المعيارى هو أهم مقاييس التشتت وأكثرها استعمالا . وترجع أهميته إلى دخوله فى عدد كبير من المقاييس الإحصائية الأخرى كمعاملات الالتواء والتفرطح والارتباط فضلا عن كونه مقياساً دقيقاً للتشتت

ويلاحظ القارىء أن الفكرة التى أوحت إلينا بإهمال إشارات الانحرافات عن الوسط الحسابى فى حالة الانحراف المتوسط هى نفس الفكرة التى أوحت إلينا بتربيع هذه الانحرافات فى حالة الانحراف المعيارى . والسبب فى أخذ

(*) $\Sigma ح^{٢}$ معناها مجموع $ح^{٢}$ أى $ح_{١}^{٢} + ح_{٢}^{٢} + ح_{٣}^{٢} + ... + ح_{ن}^{٢}$

ويجب على الطالب أن يتنبه إلى التفرقة بين $\Sigma ح^{٢}$ حيث تربيع القيم يكون قبل الجمع وبين $(\Sigma ح)^{٢}$ حيث التربيع يكون بعد الجمع فالكمية الأولى هى مجموع المربعات والكمية الثانية هى مربع المجموع .

(١) هذه الكمية هى أيضاً العزم الثانى للقيم كما سيتضح من الباب التالى .

الجذر التربيعى لمتوسط مربعات الإنحرافات بدلا من متوسط المربعات ذاته كمقياس للتشتت ، هو أننا نراعى جعل هذا المقياس مقدراً بنفس الوحدات المقيسة بها قيم المتغير .

ويتضح من تعريف الانحراف المعيارى أن خطوات العمل التى نتبعها فى إيجاد قيمته هى الآتية :ـ

أولا : إيجاد الوسط الحسانى .

ثانياً : تعيين انحرافات قيم المتغير عن الوسط الحسابى .

ثالثاً : إيجاد مربعات هذه الانحرافات (وضربها فى التكرارات المقابلة لها فى حالة التوزيع التكرارى) .

رابعاً : جمع هذه المربعات (أو حواصل ضربها فى التكرارات) وقسمة حاصل الجمع على عدد المفردات .

خامساً : استخراج الجذر التربيعى لخارج القسمة الناتج من الخطوة السابقة فيكون هذا الجذر هو الانحراف المعيارى .

ولإيضاح هذه الخطوات نوجد أولا الانحراف المعيارى لمجموعة من الأرقام الأولية ثم ننتقل بعد ذلك إلى شرح الطريقة العملية لإيجاد الإنحراف المعيارى للتوزيع التكرارى .

افرض أننا أردنا إيجاد الانحراف المعيارى لمجموعة الأرقام الآتية :

٢٠ ، ٣٠ ، ٤٠ ، ٥٠ ، ٦٠ ، ٧٠ ، ٨٠

بعد أن نستخرج الوسط الحسابى ـ ومقداره ٥٠ فى هذه الحالة ـ نعمل جدولا كالآتى :ـ

جدول رقم ٢٤

إيجاد الانحراف المعيارى للأرقام الأولية

قيم المتغير	الانحرافات عن الوسط الحسابى ح	ح^{٢}
٢٠	٣٠—	٩٠٠
٣٠	٢٠—	٤٠٠
٤٠	١٠—	١٠٠
٥٠	٠	٠
٦٠	١٠+	١٠٠
٧٠	٢٠+	٤٠٠
٨٠	٣٠+	٩٠٠

$$\sigma = \sqrt{\frac{\Sigma ح^{٢}}{ن}} = \sqrt{\frac{٢٨٠٠}{٧}} = ٢٠$$

غير أنه يندر أن تصادفنا أرقام يمكن معالجتها بمثل هذه السهولة .

فالوسط الحسابى يكون عادة رقماً يحتوى على كسور ، وبالتالى فإن الانحرافات عنه تكون أرقاماً كسرية لا تسهل معالجتها . وللتخلص من هذه الصعوبة نلجأ إلى حساب متوسط مربعات الانحرافات عن أى وسط فرضى نختاره ، ومنه نحسب قيمة الانحراف المعيارى باستخدام العلاقة الآتية :—

$$\sqrt{\frac{\Sigma ح^{٢}}{ن}} = \sqrt{\frac{\Sigma ك^{٢}}{ن} - \left(\frac{\Sigma ك}{ن}\right)^{٢}}$$

حيث ح$_{١}$ ، ح$_{٢}$ ، ح$_{٣}$ ، ... ، ح$_{ن}$ تمثل الانحرافات عن الوسط الحسابى

و ك$_{١}$ ، ك$_{٢}$ ، ك$_{٣}$ ، ... ، ك$_{ن}$ تمثل الانحرافات عن الوسط الفرضى

ويمكن إثبات هذه العلاقة كالآتى :-

$$ح_{١} = س_{١} - \overline{س}$$
$$ح_{٢} = س_{٢} - \overline{س}$$
$$ح_{٣} = س_{٣} - \overline{س}$$
$$\vdots$$
$$\vdots$$
$$ح_{ن} = س_{ن} - \overline{س}$$

وقد أثبتنا من قبل أن $\overline{س} = \overline{س_{ف}} + \frac{مج ف}{ن}$ (حيث $\overline{س}$ تمثل الوسط الحسابى و $\overline{س_{ف}}$ تمثل الوسط الفرضى)

∴ بالتعويض عن قيمة $\overline{س}$ نجد أن :

$$ح_{١} = س_{١} - (\overline{س_{ف}} + \frac{مج ف}{ن}) = (س_{١} - \overline{س_{ف}}) - \frac{مج ف}{ن}$$

$$ح_{٢} = س_{٢} - (\overline{س_{ف}} + \frac{مج ف}{ن}) = (س_{٢} - \overline{س_{ف}}) - \frac{مج ف}{ن}$$

$$\vdots$$

$$ح_{ن} = س_{ن} - (\overline{س_{ف}} + \frac{مج ف}{ن}) = (س_{ن} - \overline{س_{ف}}) - \frac{مج ف}{ن}$$

ولكن $(س_{١} - \overline{س_{ف}}) = ف_{١}$ ، $(س_{٢} - \overline{س_{ف}}) = ف_{٢}$ وهكذا

$$\therefore ح_{١} = ف_{١} - \frac{مج ف}{ن}$$

$$ح_{٢} = ف_{٢} - \frac{مج ف}{ن}$$

$$\vdots$$

$$ح_{ن} = ف_{ن} - \frac{مج ف}{ن}$$

$$\therefore \Sigma ح = \Sigma (ى - \frac{\Sigma ى}{ن})$$

$$\therefore \Sigma ح^{٢} = \Sigma (ى - \frac{\Sigma ى}{ن})^{٢}$$

$$= \Sigma \left[ى^{٢} - ٢ ى \cdot \frac{\Sigma ى}{ن} + (\frac{\Sigma ى}{ن})^{٢} \right]$$

$$= \Sigma ى^{٢} - ٢ \Sigma ى \cdot \frac{\Sigma ى}{ن} + ن (\frac{\Sigma ى}{ن})^{٢}$$

$$\therefore \frac{\Sigma ح^{٢}}{ن} = \frac{\Sigma ى^{٢}}{ن} - ٢ \frac{\Sigma ى}{ن} \cdot \frac{\Sigma ى}{ن} + (\frac{\Sigma ى}{ن})^{٢}$$

$$\therefore \frac{\Sigma ح^{٢}}{ن} = \frac{\Sigma ى^{٢}}{ن} - (\frac{\Sigma ى}{ن})^{٢}$$

$$\therefore \sqrt{\frac{\Sigma ح^{٢}}{ن}} = \sqrt{\frac{\Sigma ى^{٢}}{ن} - (\frac{\Sigma ى}{ن})^{٢}}$$

وحيث أن العلاقة السابقة غير مقيدة بأية قيمة معينة للوسط الفرضى ، فلو اخترنا وسطاً فرضياً قدره صفراً لاتخذت هذه العلاقة الصورة الآتية :ـ

$$\sqrt{\frac{\Sigma ح^{٢}}{ن}} = \sqrt{\frac{\Sigma س^{٢}}{ن} - (\frac{\Sigma س}{ن})^{٢}} = \sqrt{\frac{\Sigma س^{٢}}{ن} - \bar{س}^{٢}}$$

حيث س$_{١}$ ، س$_{٢}$ ، س$_{٣}$ ، .... ، س$_{ن}$ تمثل انحرافات القيم عن وسط فرضى قدره صفر أى تمثل قيم المتغير ذاتها .

وفى حالة التوزيع التكرارى تكون هذه العلاقة كالآتى :ـ

$$\sqrt{\frac{\Sigma ح^{٢} و}{\Sigma و}} = \sqrt{\frac{\Sigma ى^{٢} و}{\Sigma و} - (\frac{\Sigma ى و}{\Sigma و})^{٢}} = \sqrt{\frac{\Sigma س^{٢} و}{\Sigma و} - (\frac{\Sigma س و}{\Sigma و})^{٢}}$$

والآن ننتقل إلى طرق حساب الانحراف المعيارى للتوزيع التكرارى عملياً مبتدئين أولا بالطريقة المطولة :ـ

جدول رقم ٢٥

إيجاد الانحراف المعيارى لأعمار الطلبة المتقدمين للالتحاق بكلية التجارة سنة ١٩٤٧
الطريقة المطولة

فئات الأعمار	التكرار و	مراكز الفئات	الانحرافات عن الوسط الحسابى *ح	ح٢	ح٢ × و
١٦ — ١٨	٢٠	١٧	—٤,٥٢	٢٠,٤٣٠	٤٠٨,٦٠٠
١٨ — ٢٠	٥٥	١٩	—٢,٥٢	٦,٣٥٠	٣٤٩,٢٥٠
٢٠ — ٢٢	٦٥	٢١	—٠,٥٢	٠,٢٧٠٤	١٧,٥٧٦
٢٢ — ٢٤	٣٢	٢٣	١,٤٨	٢,١٩٠	٧٠,٠٨٠
٢٤ — ٢٦	٨	٢٥	٣,٤٨	١٢,١١٠	٩٦,٨٨٠
٢٦ — ٢٨	٥	٢٧	٥,٤٨	٣٠,٠٣٠	١٥٠,١٥٠
٢٨ — ٣٠	٢	٢٩	٧,٤٨	٥٥,٥٩٠	١١١,٩٠٠
٣٠ — ٣٢	٦	٣١	٩,٤٨	٨٩,٨٧٠	٥٣٩,٢٢٠
٣٢ — ٣٤	٣	٣٣	١١,٤٨	١٣١,٧٩٠	٣٩٥,٣٧٠
٣٤ — ٣٦	٤	٣٥	١٣,٤٨	١٨١,٧١٠	٧٢٦,٨٤٠
المجموع	٢٠٠				٢٨٦٥,٨٦٦

$$\sigma = \sqrt{\frac{\Sigma\, ح^2\, و}{\Sigma\, و}} = \sqrt{\frac{٢٨٦٥,٨٦٦}{٢٠٠}} = ٣,٧٨٥$$

∴ الانحراف المعيارى للتوزيع التكرارى = ٣,٧٨٥ سنة

(*) لاحظ أن الوسط الحسابى للاعمار = ٢١,٥٢ سنة .

من الواضح أن الطريقة السابقة لحساب الانحراف المعيارى تكون مرهقة جداً فى العمل خصوصاً فيما يتعلق بتربيع الانحرافات (وهى عادة أرقام كسرية) وضربها فى التكرارات المقابلة .

ويمكن تسهيل العمل الحسابى نوعاً باستخدام العلاقة

$$\sigma = \sqrt{\frac{\Sigma س^{٢} و}{\Sigma و} - \left(\frac{\Sigma س و}{\Sigma و}\right)^{٢}}$$

كما يتضح من الجدول الآتى :-

جدول رقم ٢٦

إيجاد الانحراف المعيارى لأعمار الطلبة المتقدمين للالتحاق بكلية التجارة ١٩٤٧ - ١٩٤٨

فئات الأعمار	التكرار و	مراكز الفئات س	س × و	س^{٢} × و
١٦ - ١٨	٢٠	١٧	٣٤٠	٥٧٨٠
١٨ - ٢٠	٥٥	١٩	١٠٤٥	١٩٨٥٥
٢٠ - ٢٢	٦٥	٢١	١٣٦٥	٢٨٦٦٥
٢٢ - ٢٤	٣٢	٢٣	٧٣٦	١٦٩٢٨
٢٤ - ٢٦	٨	٢٥	٢٠٠	٥٠٠٠
٢٦ - ٢٨	٥	٢٧	١٣٥	٣٦٤٥
٢٨ - ٣٠	٢	٢٩	٥٨	١٦٨٢
٣٠ - ٣٢	٦	٣١	١٨٦	٥٧٦٦
٣٢ - ٣٤	٣	٣٣	٩٩	٣٢٦٧
٣٤ - ٣٦	٤	٣٥	١٤٠	٤٩٠٠
المجموع ...	٢٠٠		٤٣٠٤	٩٥٤٨٨

$$\sigma = \sqrt{\frac{\Sigma س^{٢} و}{\Sigma و} - \left(\frac{\Sigma س و}{\Sigma و}\right)^{٢}}$$

$$= \sqrt{\frac{٩٥٤٨٨}{٢٠٠} - \left(\frac{٤٣٠٤}{٢٠٠}\right)^{٢}}$$

$$= \sqrt{١٤٫٣٢٩٦}$$

$$= ٣٫٧٨٥$$

∴ الانحراف المعيارى = ٣٫٧٨٥ سنة

على أن أفضل وأسهل طريقة لحساب الانحراف المعيارى (وهى التى يجب استخدامها دائماً) هى أن نحسب الانحرافات عن وسط فرضى نختاره عند مركز الفئة التى تتركز فيها التكرارات ، ومن ثم نعين قيمة الانحراف المعيارى عن الوسط الحسابى بتطبيق العلاقة .

$$\sigma = \sqrt{\frac{\Sigma ى^{٢} و}{\Sigma و} - \left(\frac{\Sigma ى و}{\Sigma و}\right)^{٢}}$$

حيث (ى١ ، ى٢ ، ى٣ ..... ىن) تمثل الانحرافات عن الوسط الفرضى

وكما فعلنا فى حالة الوسط الحسابى نحسب الانحرافات عن الوسط الفرضى بوحدات طول الفئة على أن نعود إلى الوحدات الأصلية بعد استخراج النتيجة النهائية . ويمكن كذلك اختبار صحة العمليات الحسابية باستخدام مراجعة Charlier . والمراجعة فى هذه الحالة تنصب على قيمة $\Sigma ى^{٢}$ وحيث أن العمليات الحسابية التى تؤدى إلى هذا المجموع هى التى يحتمل أن يتسرب إليها الخطأ . كما أنه لو كانت هناك أخطاء فى العمليات السابقة فإنها تظهر فى هذا المجموع وبالتالى تظهرها مراجعة Charlier . وتتطلب هذه المراجعة عمل عمود إضافى يعطينا قيمة $\Sigma و(ى + ١)^{٢}$. والجدول الآتى يوضح خطوات العمل اللازمة لحساب الانحراف المعيارى بهذه الطريقة مع العلم بأن الوسط الفرضى هو ٢١ :-

جدول رقم ٢٧

إيجاد الانحراف المعيارى باستخدام وسط فرضى مع تطبيق مراجعة شارلييه

مراكز الفئات	التكرار و	الانحرافات عن الوسط الفرضى	الانحرافات بوحدات طول الفئة ى	ى و	ى٢ و	(ى+١)٢ و
١٧	٢٠	٤-	٢-	٤٠-	٨٠	٢٠
١٩	٥٥	٢-	١-	٥٥-	٥٥	٠
٢١	٦٥	٠	٠	٠	٠	٦٥
٢٣	٣٢	٢	١	٣٢	٣٢	١٢٨
٢٥	٨	٤	٢	١٦	٣٢	٧٢
٢٧	٥	٦	٣	١٥	٤٥	٨٠
٢٩	٢	٨	٤	٨	٣٢	٥٠
٣١	٦	١٠	٥	٣٠	١٥٠	٢١٦
٣٣	٣	١٢	٦	١٨	١٠٨	١٤٧
٣٥	٤	١٤	٧	٢٨	١٩٦	٢٥٦
	٢٠٠			٥٢	٧٣٠	١٠٣٤

مراجعة شارلييه : $\Sigma (ى + ١)^2 و = \Sigma ى^2 و + ٢\Sigma ى و + \Sigma و$

∴ ١٠٣٤ = ٧٣٠ + ٢ × ٥٢ + ٢٠٠ = ١٠٣٤

∴ العمليات الحسابية صحيحة .

$$\sigma = \sqrt{\frac{\Sigma ى^2 و}{\Sigma و} - \left(\frac{\Sigma ى و}{\Sigma و}\right)^2}$$

$$= \sqrt{\frac{٧٣٠}{٢٠٠} - \left(\frac{٥٢}{٢٠٠}\right)^2}$$

$$= \sqrt{٣,٥٨٢٤}$$

∴ σ = ١,٨٩٢٥ بوحدات طول الفئة وقدرها سنتان

∴ الانحراف المعيارى = ٢ × ١,٨٩٢٥ = ٣,٧٨٥ سنة

تصحيحات شبرد Sheppard's Corrections :

أشرنا من قبل إلى أن وضع الأرقام الأولية فى شكل توزيع تكرارى ينطوى على افتراض أن جميع المفردات الواقعة داخل الفئة متساوية ومركزة عند مركز الفئة بحيث يمكن اعتبار هذا المركز ممثلا لجميع المفردات الواقعة داخل الفئة . ومن الواضح أن هذا الفرض لا يمكن اعتباره صحيحاً فى جميع الحالات . وقد وجد Sheppard بعد تحليل رياضى عميق أنه يمكن تصحيح الخطأ المترتب على هذا الفرض فى قيمة التباين[1] (σ^2) بطرح كمية قدرها ($\frac{1}{12} \times$ مربع طول فئة التوزيع) من القيمة التى نحصل عليها بإحدى الطرق السابق توضيحها .

فإذا رمزنا لطول فئة التوزيع بالرمز ف فإن القيمة المعدلة للتباين .

$$= \sigma^2 - \frac{ف^2}{12}$$

وبناء عليه تكون القيمة المعدلة للانحراف المعيارى = $\sqrt{\sigma^2 - \frac{ف^2}{12}}$

على أنه من المهم أن نلاحظ أن تصحيحات Sheppard لا تطبق إلا إذا تحققت الشروط الآتية فى التوزيع التكرارى :-

أولا : يجب أن يكون المنحنى الذى يمثل التوزيع التكرارى من النوع المحدب الذى ينتهى طرفاه بملامسة المحور الأفقى .

ثانياً : يجب أن يكون المنحنى التكرارى متماثلا تقريباً ومعتدل التفرطح (Mesokurtic) أو على وجه أدق يجب أن يكون المنحنى قريباً من الطبيعى (Normal Curve) .

ثالثاً : يجب أن يكون عدد مفرادات التوزيع كبيراً (١٠٠٠ على الأقل)

رابعاً : يجب ألا يقل عدد فئات التوزيع عن ٢٠ فئة .

---

(١) وضع W. F. Sheppard تصحيحات خاصة بالعزوم الأخرى للتوزيعات التكرارية سنشير إليها فى الباب التالى .

وإزاء كل هذه القيود الموضوعة على تصحيحات Sheppard فإن مجال تطبيقها العملى محدود جداً . وحتى فى الحالات التى تتوفر فيها كل هذه الشروط فإن قيمة الانحراف المعيارى لا تتغير تغيراً ملموساً بعمل هذا التعديل . ولهذه الأسباب نستطيع أن نقول إنه على الرغم من روعة التحليل الرياضى الذى توصل Sheppard بمقتضاه إلى عمل هذه التعديلات فإن قيمتها العملية ليست كبيرة.

وبتطبيق تصحيحات Sheppard على قيمة الانحراف المعيارى للتوزيع التكرارى لأعمار الطلبة المتقدمين للالتحاق بكلية التجارة نحصل على النتيجة الآتية(١) :-

$$\text{التباين (بعد التصحيح)} = ١٤,٣٢٩٦ - \frac{٢^{٢}}{١٢} = ١٣,٩٩٦٦$$

∴ الانحراف المعيارى (بعد التصحيح) $= \sqrt{١٣,٩٩٦٦} = ٣,٧٤١$ سنة

وهذه النتيجة لا تختلف عن السابقة إلا بمقدار ٠,٠٤ سنة .

خواص الانحراف المعيارى وتطبيقاته :-

أولا : ذكرنا فى الباب السابق أن من خواص الوسط الحسابى أن مجموع مربعات الانحرافات عنه أقل من مجموع مربعات الانحراف عن أية قيمة أخرى . ومن الواضح أن هذه الخاصية تعنى أن قيمة الانحراف المعيارى تكون حداً أدنى إذا كانت الانحرافات مقيسة عن الوسط الحسابى ، أى أن :-

$$\frac{\Sigma ع^{٢}}{ن} < \frac{\Sigma ح^{٢}}{ن}$$

حيث $ح_{١}$ ، $ح_{٢}$ ، $ح_{٣}$ . . . $ح_{ن}$ تمثل الانحرافات عن الوسط الحسابى .

---

(١) لاحظ أن هذا التوزيع لا تتوفر فيه الشروط اللازمة لتطبيق تصحيحات Sheppard وإنما أجرينا هذا التصحيح لإيضاح طريقة عمله ليس غير .

و ء١ ، ء٢ ، ء٣ ... ءن تمثل الانحرافات عن أية قيمة أخرى .

ويمكن إثبات صحة ذلك مباشرة باستخدام العلاقة .

$$\frac{\Sigma ح^2}{ن} = \frac{\Sigma ء^2}{ن} - \left(\frac{\Sigma ء}{ن}\right)^2$$

التى سبق أن أثبتناها من قبل (١) .

إذ يظهر من هذه العلاقة أن : —

$$\frac{\Sigma ء^2}{ن} - \frac{\Sigma ح^2}{ن} = \left(\frac{\Sigma ء}{ن}\right)^2$$

وحيث أن $\left(\frac{\Sigma ء}{ن}\right)^2$ لا بد أن تكون كمية موجبة .

$$\therefore \quad \frac{\Sigma ء^2}{ن} > \frac{\Sigma ح^2}{ن}$$

ثانياً : فى التوزيعات التكرارية التى تمثلها منحنيات يقرب شكلها من شكل المنحنى الطبيعى يكون حوالى ثلثى التكرارات واقعاً فى مدى تحده القيمتان ($\bar{س} + \sigma$) ، ($\bar{س} - \sigma$) .

وبتطبيق هذه الخاصية على التوزيع التكرارى لأعمار الطلبة المتقدمين للالتحاق بكلية التجارة ، فإن الذى نتوقعه نظرياً أن يكون ثلثا الطلبة تقع أعمارهم بين (٢١,٥٢ + ٣,٧٨٥) ، (٢١,٥٢ — ٣,٧٨٥) أى أننا نتوقع أن تنحصر أعمار ١٣٣ طالباً بين ٢٥,٣٠٥ ، ١٧,٧٣٥

وبمراجعة التوزيع التكرارى ذاته نجد أن حوالى ١٥٠ طالباً تقع أعمارهم بين ١٧ ، ٢٦ والسبب فى اختلاف النتيجتين هو أن هذا التوزيع يختلف كثيراً عن التوزيع الطبيعى .

كذلك نجد فى التوزيعات القريبة من التوزيع الطبيعى أن حوالى ٩٥٪ من التكرارات تقع فى مدى تحده القيمتان ($\bar{س} + ٢\sigma$)، ($\bar{س} - ٢\sigma$)

(١) راجع ص ١٩١ من هذا الكتاب .

بينما تقع جميع التكرارات تقريباً (حوالى ٩٩,٧٪ منها) فى مدى تحده القيمتان ($\bar{س} + ٣\sigma$) و ($\bar{س} - ٣\sigma$)[١].

أما عن تطبيقات الانحراف المعيارى فلا تكاد تخلو أية طريقة إحصائية راقية من الاعتماد عليه ولذلك سوف لا نسرد هذه التطبيقات الآن تاركين للقارئ تتبعها بنفسه فى الأبواب التالية من هذا الكتاب. ويكفى أن نذكر هنا أن الانحراف المعيارى هو أدق مقاييس التشتت وأن مركزه بين مجموعة المقاييس الوصفية الثانية يناظر مركز الوسط الحسابى بين مجموعة المقاييس الوصفية الأولى .

مقاييس التشتت النسبى .

بمراجعة مقاييس التشتت التى ذكرناها حتى الآن نجد أن جميعها مقاييس مطلقة تقدر بدلالة الوحدات المستعملة فى قياس المتغيرات الموضوعة تحت البحث ، سواء أكانت هذه الوحدات عبارة عن سنوات أو جنيهات أو أرطال أو أمتار . وعلى ذلك إذا أردنا مقارنة درجة التشتت فى مجموعتين — أو فى صفتين مختلفتين متصلتين بنفس المجموعة — حال دون ذلك اختلاف وحدات القياس فى الحالتين . فلو فرض مثلا أننا أردنا مقارنة درجة تشتت مجموعة الطلبة المتقدمين للالتحاق بكلية التجارة من حيث السن والوزن ، وأنه بحساب قيمة الانحراف المعيارى للأوزان وجدناه ١٠ أرطال ، فهل يجوز مقارنة هذا الرقم بالانحراف المعيارى للأعمار الذى وجدنا أنه يساوى ٣,٧٨٥ سنة ؟ من الواضح أن مثل هذه المقارنة لا يستقيم لها معنى إلا إذا استعملنا مقياساً نسبياً للتشتت يخلصنا من هذه الوحدات المختلفة ، إذ لا يعقل أن نقارن سنوات بأرطال . ويمكن أن نجعل الانحراف المعيارى مقياساً نسبياً للتشتت

---

(١) الواقع أن هذا الخواص هى من خواص المنحنى الطبيعى وحده وتعميمها بالنسبة للمنحنيات الأخرى المشابهة له يكون من باب التقريب فقط .

بقسمته على الوسط الحسابى ، إذ أن خارج القسمة يكون عبارة عن نسبة مجردة من التمييز . أى أنه من الممكن استخدام النسبة $\frac{\sigma}{\bar{س}}$ كمقياس نسبى للتشتت . ومن المتفق عليه أن يوضع هذا المقياس فى شكل نسبة مئوية هكذا :—

$\frac{\sigma}{\bar{س}} \times ١٠٠$ وهو على هذه الصورة يسمى معامل الاختلاف (Coefficient of Variation) .

ولإيضاح طريقة استعماله نفرض أن الوسط الحسابى لأوزان الطلبة هو ١٦٠ رطلا ، وأن الوسط الحسابى لاعمار نفس المجموعة هو كما حسبناه من قبل ٢١,٥٢ سنة .

$$\text{معامل الاختلاف للأعمار} = \frac{٣,٧٨٥}{٢١,٥٢} \times ١٠٠ = ١٧,٦\ ٪$$

$$\text{معامل الاختلاف للأوزان} = \frac{١٠}{١٦٠} \times ١٠٠ = ٦,٢٥\ ٪$$

ومن هذا يتضح أن تشتت مجموعة الطلبة من حيث السن أكبر من تشتتها من حيث الوزن ، مع أنه لو صحت مقارنة الأرقام المطلقة لاعتبرنا تشتت المجموعة من حيث الوزن أكبر .

وليس استعمال مقاييس التشتت النسبية مقصوراً على تلك الحالات التى نرغب فيها فى مقارنة تشتت مجموعتين مقيستين بوحدات مختلفة . فقد نرغب فى مقارنة مجموعتين مقيستين بنفس الوحدات غير أن هناك فارقاً كبيراً بين متوسطيهما وفى مثل هذه الحالات فإن مقاييس التشتت المطلقة لا تصلح للمقارنة . افرض مثلا أننا أردنا مقارنة تشتت مجموعة من الطلبة من حيث الطول ومحيط الرقبة (وكلاهما مقدراً بالسنتيمترات) وكانت الأطوال ومحيطات الرقبة كالآتى :—

الأطوال	١٦٠ ،	١٦٤ ،	١٦٨ ،	١٧٢ ،	١٧٦
محيطات الرقبة	٣٠ ،	٣٢ ،	٣٤ ،	٣٦ ،	٣٨

أى أنه فى حالة الأطوال $\bar{س}$ = ١٦٨ سنتيمترا ، σ = ٤ $\sqrt{٢}$ سنتيمتر

وفى حالة محيطات الرقبة $\bar{س}$ = ٣٤ سنتيمترا ، σ = ٢ $\sqrt{٢}$ سنتيمتر

يبدو من مقارنة الانحراف المعيارى فى الحالتين أن تشتت الأطوال أكبر من تشتت محيطات الرقبة والواقع خلاف ذلك فإن تشتت محيطات الرقبة أكبر بكثير كما يتضح من مقارنة معامل الاختلاف فى الحالتين :—

$$\text{معامل الاختلاف للأطوال} = \frac{٤\sqrt{٢}}{١٦٨} \times ١٠٠ = ٢,٤\,\%$$

$$\text{معامل الاختلاف لمحيطات الرقبة} = \frac{٢\sqrt{٢}}{٣٤} \times ١٠٠ = ٨,٤\,\%$$

ومن هذا يظهر أنه إذا كان هناك فرق كبير بين متوسطى مجموعتين فان الانحراف المعيارى (أو غيره من مقاييس التشتت المطلقة) لا يصلح لمقارنة درجة التشتت فى كل منهما حتى وإن كانت مفردات كل من المجموعتين مقيسة بنفس الوحدات . ففى مثل هذه الحالات لا بد من استعمال مقياس نسبى للتشتت .

وبالإضافة إلى معامل الاختلاف هناك مقاييس أخرى للتشتت النسبى ، وهى عبارة عن مقاييس التشتت المطلقة السابق ذكرها معدلة بطريقة أو أخرى لتصبح نسبية . وفى أغلب الحالات يتم هذا التعديل عن طريق حساب مقياس التشتت المطلق كنسبة مئوية من الوسط الحسابى . فإذا أردنا مثلا تعديل الانحراف الربيعى لجعله مقياساً نسبياً للتشتت نضعه فى الصورة الآتية :—

$$\frac{١٠٠\,ع}{\bar{س}}$$ (حيث ع تمثل الانحراف الربيعى) .

على أن معامل الاختلاف هو أدق مقاييس التشتت النسبى وهو الذى يستعمله الإحصائيون عادة لمقارنة تشتت المجموعات .

الباب السابع

العزوم ومقاييس الالتواء والتفرطح (*)

أتينا حتى الآن بطرق وصف التوزيعات التكرارية من حيث القيمة المتوسطة التى تتركز حولها مفردات المجموعة ومن حيث درجة تشتت هذه المفردات حول تلك القيمة المتوسطة . ولكى نتم وصف التوزيع التكرارى يبقى أن نحدد شكل المنحنى الذى يمثله من حيث التماثل أو الالتواء من جهة وتدبب القمة أو تفرطحها من الجهة الأخرى . فهناك بالإضافة إلى المجموعتين الأولى والثانية من المقاييس الوصفية مجموعتان أخريان هما مقاييس الالتواء ومقاييس التفرطح . وقياس الالتواء والتفرطح يتطلب تعيين قيم ثوابت أخرى للتوزيعات التكرارية خلاف تلك التى درسناها حتى الان ، وهذه الثوابت هى عزوم التوزيع التكرارى التى نعرفها فيما يلى :-

١ - العزم الأول للتوزيع التكرارى حول وسطه الحسابى يساوى متوسط انحرافات قيم مفردات التوزيع عن هذا الوسط .

وباستخدام طريقة التعبير الرمزى التى اتبعناها حتى الآن يكون العزم الأول للتوزيع التكرارى حول وسطه الحسابى يساوى $\frac{\Sigma\, ح\, و}{\Sigma\, و}$ وسنعبر عنه

(*) الاصطلاح الأجنبى للالتواء هو Skewness وللتفرطح Kurtosis

بالرمز مـ١ . وحيث أن مجموع انحرافات قيم المفردات عن وسطها الحسابى يساوى صفراً فمن الواضح أن العزم الأول حول الوسط يساوى صفراً دائماً . على أنه من الممكن حساب العزم الأول (وغيره من العزوم) حول أية قيمة أخرى نختارها . غير أن الغالب أن نعين العزوم حول إحدى قيمتين : الوسط الحسابى أو الصفر . وإذا احتجنا أحياناً إلى تعيين العزوم حول أية قيمة أخرى فإنما يكون ذلك تمهيداً لتعيين قيمتها حول الوسط الحسابى .

ويلاحظ أنه إذا استخرجنا العزم الأول للتوزيع التكرارى حول الصفر لحصلنا على الوسط الحسابى ذاته : ففى هذه الحالة نستبدل الانحرافات ح١ ، ح٢ ، ... حن بالقيم س١ ، س٢ ، ... سن

$$\text{ويكون العزم الأول (حول الصفر)} = \frac{\Sigma\, س\, و}{\Sigma\, و} = \bar{س}$$

٢ ــ العزم الثانى للتوزيع التكرارى حول وسطه الحسابى يساوى متوسط مربعات انحرافات قيم مفردات التوزيع عن هذا الوسط . ونعبر عنه بالرمز $مـ_٢ = \frac{\Sigma\, ح^٢\, و}{\Sigma\, و}$.

ومن الواضح أن العزم الثانى حول الوسط الحسابى هو التباين أو مربع الانحراف المعيارى .

٣ ــ العزم الثالث للتوزيع التكرارى حول الوسط الحسابى هو متوسط مكعبات انحرافات قيم مفردات التوزيع عن هذا الوسط .

$$\text{والتعبير الرمزى عن العزم الثالث هو } مـ_٣ = \frac{\Sigma\, ح^٣\, و}{\Sigma\, و}$$

٤ ــ العزم الرابع للتوزيع التكرارى حول الحسابى هو متوسط الانحرافات المرفوعة إلى القوة الرابعة لقيم مفردات التوزيع عن وسطه الحسابى . والتعبير الرمزى عن العزم الرابع هو $مـ_٤ = \frac{\Sigma\, ح^٤ و}{\Sigma\, و}$

ويمكننا قياساً على ما تقدم إيجاد العزم الخامس والسادس والسابع . . إلخ . غير أن العزوم الأربعة الأولى هى فقط التى تعتبر ذات فائدة عملية . وعلى وجه العموم فإن العزم الرائى للتوزيع التكرارى هو :

$$م_{ر} = \frac{\Sigma\, ح^{ر} و}{\Sigma\, و} \quad (*)$$

وسيلاحظ القارئ أننا استعملنا فعلا كلا من العزم الأول والثانى فى الأبواب السابقة . فالعزم الأول حول الصفر هو كما ذكرنا عبارة عن الوسط الحسابى للتوزيع التكرارى والعزم الثانى حول الوسط الحسابى هو عبارة عن مربع الانحراف المعيارى . أما العزمان الثالث والرابع فسنستعملهما فى إيجاد معاملى الالتواء والتفرطح على التوالى .

العزم الثالث كمقياس للالتواء .

ذكرنا أن العزم الأول للتوزيع التكرارى حول وسطه الحسابى يساوى صفراً . وهذا صحيح دائماً سواء أكان التوزيع متماثلا أم غير متماثل أى ملتوياً .

أما العزم الثانى فيساوى دائماً كمية موجبة وذلك لأن تربيع الانحرافات يخلصنا من الإشارة السالبة .

أما العزم الثالث فقيمته تساوى صفراً فى التوزيعات المتماثلة فقط . أما فى التوزيعات الملتوية فقد تكون قيمته موجبة أو سالبة طبقاً لما إذا كان التوزيع ملتوياً التواء موجباً أو سالباً . وذلك لأنه على الرغم من أن تكعيب الانحرافات لا يخلصنا من الإشارة السالبة ، فتظل انحرافات القيم الأقل من المتوسط تحمل إشارة مضادة لإشارة انحرافات القيم الأكبر من

(*) لاحظ أنه فى حالة الأرقام الأولية غير الموضوعة فى شكل توزيع تكرارى يكون العزم الأول $= \frac{\Sigma\, ح}{ن}$ والثانى $= \frac{\Sigma\, ح^{٢}}{ن}$ والثالث $= \frac{\Sigma\, ح^{٣}}{ن}$ والرابع $= \frac{\Sigma\, ح^{٤}}{ن}$ والعزم الرائى $= \frac{\Sigma\, ح^{ر}}{ن}$

حيث ح تمثل انحرافات القيم عن وسطها الحسابى ، ن تمثل عدد المفردات .

المتوسط ، إلا أنه عند جمع مكعبات الانحرافات جمعاً جبرياً فإن النتيجة لا تكون صفراً . وذلك لأن عملية التكعيب تبالغ فى تكبير قيم الانحرافات الكبيرة ، فلو كانت الانحرافات الكبيرة سالبة يكون المجموع الجبرى لمكعبات الانحرافات سالباً ولو كانت الانحرافات الكبيرة موجبة فإن المجموع الجبرى لمكعبات الانحرافات يكون موجباً ولو كانت الانحرافات السالبة تماثل الانحرافات الموجبة تماماً (أى لو كان التوزيع متماثلا حول الوسط الحسابى) فإن المجموع الجبرى لمكعبات الانحرافات يساوى صفراً وبالتالى فإن قيمة العزم الثالث تكون صفراً كذلك . وتتضح هذه النقطة من مقارنة الجدولين الآتيين :-

جدول رقم ٢٨

العزم الأول والثالث لتوزيع متماثل

قيم المتغير	ح	ح٣
٣	-٦	-٢١٦
٦	-٣	-٢٧
٩ ($\bar{س}$)	٠	٠
١٢	+٣	+٢٧
١٥	+٦	+٢١٦
المجموع	٠٠	٠٠

جدول رقم ٢٩

العزم الأول والثالث لتوزيع غير متماثل

قيم المتغير	ح	ح٣
٣	-٣	-٢٧
٤	-٢	-٨
٦ ($\bar{س}$)	٠	٠
٧	+١	+١
١٠	+٤	+٦٤
المجموع	٠٠	+٣٠

$$م_{٣} = \frac{\Sigma ح^{٣}}{ن} = \text{صفر} \qquad م_{٣} = \frac{\Sigma ح^{٣}}{ن} = \frac{٣٠}{٥} = ٦$$

يتضح مما تقدم أنه يمكن استغلال خاصية العزم الثالث هذه فى قياس درجة التواء التوزيع أو بالأحرى درجة التواء المنحنى الذى يمثله . وحيث أنه

قلما نرغب فى قياس التواء المنحنيات إلا لغرض مقارنتها ببعضها فالأفضل أن نجعل مقياس الالتواء مقياساً نسبياً مستقلا عن الوحدات التى يقاس بها المتغير كما فعلنا فى مقاييس التشتت النسبى . وعلى ذلك فلكى نحصل على مقياس نسبى للالتواء نقسم العزم الثالث على مكعب الانحراف المعيارى[1] فيكون المقياس الأول للالتواء هو :—

$$\frac{\text{العزم الثالث}}{\text{مكعب الانحراف المعيارى}} = \frac{م_{٣}}{\sigma^{٣}} = \frac{م_{٣}}{\sqrt{(م_{٢})^{٣}}}$$

وقد اقترح كارل بيرسون استعمال مربع الكمية السابقة كمقياس للالتواء بدلا من الكمية ذاتها ونهج على منواله كثير من الإحصائيين . وعلى ذلك يكون مقياس الالتواء المستعمل عادة أو معامل الالتواء هو :—

$$\frac{\text{مربع العزم الثالث}}{\text{مكعب العزم الثانى}} = \frac{م_{٣}^{٢}}{م_{٢}^{٣}}$$ وسنعبر عنه رمزياً بالحرف $ل_{١}$ (*)

ويلاحظ أن قيمة $ل_{١}$ تساوى صفر إذا كان التوزيع التكرارى متماثلا .

طريقة حساب العزم الثالث للتوزيع التكرارى عملياً .

يتضح من تعريف العزم الثالث أن الخطوات الواجب اتباعها لتعيين قيمته هى الآتية :—

أولا : نعين الوسط الحسابى للتوزيع التكرارى .

ثانياً : نحسب انحرافات قيم المفردات عن الوسط الحسابى .

ثالثاً : نحسب مكعبات هذه الانحرافات ونضربها فى التكرارات المقابلة.

رابعاً : نجمع حواصل الضرب ونقسم حاصل الجمع على عدد المفردات فيكون خارج القسمة هو العزم الثالث .

---

(١) السبب فى القسمة على مكعب الانحراف المعيارى وليس الانحراف المعيارى نفسه أو مربعه أن وحدات القياس فى العزم الثالث تكون مرفوعة إلى القوة الثالثة وبالتالى يجب قسمته على كمية تكون الوحدات فيها مرفوعة إلى نفس القوة مثل $\sigma^{٣}$.

(*) التعبير الرمزى الأجنبى عن معامل الالتواء هو الحرف الاغريقى $\beta_{١}$ وينطق « بيتا » .

والجدول الآتى يوضح هذه الخطوات عملياً .

جدول رقم رقم ٣٠

إيجاد العزم الثالث للتوزيع التكرارى لأعمار الطلبة المتقدمين للالتحاق بكلية التجارة
(الطريقة المطولة)

فئات الأعمار	التكرار و	ح	ح٣	ح٣ و
١٦ — ١٨	٢٠	٤,٥٢—	٩٢,٣٤٥٠—	١٨٤٦,٩٠٠—
١٨ — ٢٠	٥٥	٢,٥٢—	١٦,٠٠٣٠—	٨٨٠,١٦٥—
٢٠ — ٢٢	٦٥	٠,٥٢—	٠,١٤٠٦—	٩,١٣٩—
٢٢ — ٢٤	٣٢	١,٤٨	٣,٢٤١٨	١٠٣,٧٣٧
٢٤ — ٢٦	٨	٣,٤٨	٤٢,١٤٤٢	٣٣٧,١٥٤
٢٦ — ٢٨	٥	٥,٤٨	١٦٤,٥٦٦٦	٨٢٢,٨٣٣
٢٨ — ٣٠	٢	٧,٤٨	٤١٨,٥٠٩٠	٨٣٧,٠١٨
٣٠ — ٣٢	٦	٩,٤٨	٨٥١,٩٧١٤	٥١١١,٨٢٨
٣٢ — ٣٤	٣	١١,٤٨	١٥١٢,٩٥٤٠	٤٥٣٨,٨٦١
٣٤ — ٣٦	٤	١٣,٤٨	٢٤٤٩,٤٥١٠	٩٧٩٧,٨٠١
المجموع	٢٠٠			١٨٨١٣,٠٣١

$$م_{٣} = \frac{\Sigma\, ح^{٣} و}{\Sigma\, و} = \frac{١٨٨١٣,٠٣١}{٢٠٠} = ٩٤,٠٦٥ \text{ سنة}$$

$$ل_{١} = \frac{(٩٤,٠٦٥)^{٢}}{(١٤,٣٢٩)^{٣}} = ٣,٠١ \text{ تقريباً}$$

من الواضح أن العمل الحسابى الذى يتطلبه إيجاد العزم الثالث بالطريقة السابقة يكون شاقاً جداً ومعرضاً لتسرب الأخطاء إليه . فعملية تكعيب

الانحرافات عن الوسط الحسابى – التى تحتوى عادة على كسور – ليست هينة . ولو فرض أن كان لدينا جداول مكعبات الأعداد فإنها لا تفيدنا كثيراً ، لأننا لا نستطيع أن نستخرج منها سوى مكعبات الأعداد المكونة من ثلاثة أرقام معنوية فقط كما أن النتائج المعطاة فى هذه الجداول تكون عادة مقربة إلى أقرب أربعة أرقام معنوية فقط . وعلى ذلك فلا بد أن نبحث عن طريقة تخلصنا من تكعيب الانحرافات الكسرية .. ومثلما فعلنا فى حالة الوسط الحسابى والانحراف المعيارى يمكننا أن نتخلص من الأرقام الكسرية كلية بحساب انحرافات القيم عن وسط فرضى مناسب ومن ثم نستخدم العلاقة بين العزم الثالث حول الوسط الحسابى والعزم الثالث حول الوسط الفرضى الآتى بيانها :–

$$\frac{\Sigma ح^{٣}}{ن} = \frac{\Sigma ى^{٣}}{ن} - ٣ \frac{\Sigma ى}{ن} \cdot \frac{\Sigma ى^{٢}}{ن} + ٢ \left(\frac{\Sigma ى}{ن}\right)^{٣}$$

حيث ح تمثل انحرافات القيم عن الوسط الحسابى

، ى تمثل انحرافات القيم عن الوسط الفرضى

، ن تمثل عدد المفردات

وطريقة إثبات هذه العلاقة لا تختلف عن طريقة إثبات العلاقة بين التباين حول الوسط الحسابى والتباين حول الوسط الفرضى التى أوردناها فى الباب السابق كما يتضح مما يأتى :–

أثبتنا من قبل أن $\Sigma ح = \Sigma (ى - \frac{\Sigma ى}{ن})$ (*)

$$\therefore \quad \frac{\Sigma ح^{٣}}{ن} = \frac{\Sigma (ى - \frac{\Sigma ى}{ن})^{٣}}{ن}$$

(*) راجع ص ١٩٢ من هذا الكتاب .

$$= \frac{\Sigma [ء^{٣} - ٣ء^{٢}(\frac{\Sigma ء}{ن}) + ٣ء(\frac{\Sigma ء}{ن})^{٢} - (\frac{\Sigma ء}{ن})^{٣}]}{ن}$$

$$= \frac{\Sigma ء^{٣} - ٣\frac{\Sigma ء}{ن}\Sigma ء^{٢} + ٣(\frac{\Sigma ء}{ن})^{٢}\Sigma ء - ن(\frac{\Sigma ء}{ن})^{٣}}{ن}$$

$$= \frac{\Sigma ء^{٣}}{ن} - ٣\frac{\Sigma ء}{ن} \cdot \frac{\Sigma ء^{٢}}{ن} + ٣(\frac{\Sigma ء}{ن})^{٣} - (\frac{\Sigma ء}{ن})^{٣}$$

$$= \frac{\Sigma ء^{٣}}{ن} - ٣\frac{\Sigma ء}{ن} \cdot \frac{\Sigma ء^{٢}}{ن} + ٢(\frac{\Sigma ء}{ن})^{٣}$$

ومن الواضح أنه فى حالة التوزيع التكرارى يكون :

$$م_{٣} = \frac{\Sigma ح^{٣} و}{\Sigma و} = \frac{\Sigma ء^{٣} و}{\Sigma و} - ٣\frac{\Sigma ء و}{\Sigma و} \cdot \frac{\Sigma ء^{٢} و}{\Sigma و} + ٢(\frac{\Sigma ء و}{\Sigma و})^{٣}$$

والجدول الآتى يوضح الطريقة العملية لإيجاد العزم الثالث حول الوسط الحسابى باستخدام وسط فرضى مع ملاحظة أن الانحرافات عن الوسط الفرضى (٢١) مقدرة بوحدات طول الفئة :ـ

جدول رقم ٣١

إيجاد العزم الثالث لأعمار الطلبة المتقدمين للالتحاق بكلية التجارة سنة ١٩٤٧

الطريقة المختصرة

مراكز الفئات	التكرار و	الانحرافات عن الوسط الفرضى ء	ءو	ء٢ و	ء٣ و
١٧	٢٠	—٢	—٤٠	٨٠	—١٦٠
١٩	٥٥	—١	—٥٥	٥٥	—٥٥
٢١	٦٥	٠	٠	٠	٠
٢٣	٣٢	١	٣٢	٣٢	٣٢
٢٥	٨	٢	١٦	٣٢	٦٤
٢٧	٥	٣	١٥	٤٥	١٣٥
٢٩	٢	٤	٨	٣٢	١٢٨
٣١	٦	٥	٣٠	١٥٠	٧٥٠
٣٣	٣	٦	١٨	١٠٨	٦٤٨
٣٥	٤	٧	٢٨	١٩٦	١٣٧٢
المجموع	٢٠٠		٥٢	٧٣٠	٢٩١٤

$$م_{٣} = \frac{\Sigma ء^{٣} و}{\Sigma و} - ٣ \frac{\Sigma ء و}{\Sigma و} \cdot \frac{\Sigma ء^{٢} و}{\Sigma و} + ٢ \left(\frac{\Sigma ء و}{\Sigma و}\right)^{٣}$$

$$= \frac{٢٩١٤}{٢٠٠} - ٣ \times \frac{٥٢}{٢٠٠} \times \frac{٧٣٠}{٢٠٠} + ٢ \left(\frac{٥٢}{٢٠٠}\right)^{٣}$$

$$= ١٤,٥٧ - ٢,٨٤٧ + ٠,٠٣٥١٥٢$$

$$= ١١,٧٥٨١٥٢$$ بوحدات طول الفئة .

وإذا أردنا أن نعود إلى الوحدات الأصلية نضرب هذا الرقم فى مكعب

طول الفئة فيكون العزم الثالث للتوزيع التكرارى :

$$= ١١,٧٥٨١٥٢ \times ٢^{٣} = ٩٤,٠٦٥٢١ \text{ سنة}$$

وهى نفس النتيجة التى حصلنا عليها بالطريقة المطولة .

على أنه ليس هناك أية حاجة إلى العودة إلى الوحدات الأصلية . فلو أردنا حساب قيمة معامل الالتواء مثلا نستطيع أن نقسم العزم الثالث المقدر بوحدات طول الفئة على مكعب الانحراف المعيارى المقدر بوحدات طول الفئة كذلك ، وبهذه الطريقة يكون مقياس الالتواء :

$$\frac{م_{٣}}{\sigma^{٣}} = \frac{١١,٧٥٨١٥٢}{٦,٧٧٨٠٩٥}$$

{ حيث كل من $م_{٣}$ ، σ مقدراً بوحدات طول الفئة }

$$= ١,٧٣٤٧$$

ومربع هذا المقياس أى $ل_{١} = (١,٧٣٤٧)^{٢} = ٣,٠٠٩$ وهى نفس النتيجة التى حصلنا عليها بالطريقة المطولة .

ويلاحظ أن قيمة $ل_{١}$ فى هذه الحالة موجبة وهى النتيجة التى نتوقعها ، فبالرجوع إلى الشكل رقم ٣٦ نجد أن المنحنى التكرارى الذى يمثل هذا التوزيع ملتو التواء موجباً (نحو اليمين) .

المقاييس الأخرى للالتواء .

بالإضافة إلى مقياس الالتواء المبنى على العزم الثالث للتوزيع التكرارى هناك مقاييس أخرى بعضها ينبنى على مقارنة قيمتى الوسط الحسابى والمنوال والبعض الآخر ينبنى على مقارنة قيمتى الربيعين . ونورد هذه المقاييس فيما يلى :—

١ — قياس الالتواء بمقارنة الوسط الحسابى والمنوال .

عرفنا من دراسة المنوال أنه لا يتأثر بوجود بعض القيم الشاذة بين المفردات فى حين أن الوسط الحسابى يتأثر بهذه القيم ، ووجدنا أن قيمة المنوال تعادل قيمة الوسط الحسابى تماماً فى حالة التوزيعات التكرارية المتماثلة . وعلى ذلك يمكننا

استخدام العلاقة بين الوسط الحسابى والمنوال فى قياس درجة الالتواء . فمن الواضح أن التواء المنحنى التكرارى ينشأ عن وجود بعض القيم الشاذة فى أحد طرفى التوزيع . وبناء على ذلك لو قسنا الفرق بين المتوسط الذى يتأثر بالقيم الشاذة وذلك الذى لا يتأثر بهذه القيم لأخذنا فكرة عامة عن درجة التواء المنحنى الذى يمثل التوزيع التكرارى . وعلى ذلك يمكننا اعتبار الكمية .

($\overline{س} - م$) مقياساً للالتواء (حيث $\overline{س}$ تمثل الوسط الحسابى ، م تمثل المنوال) . ولكى نجعل هذا المقياس نسبياً للأسباب التى أوضحناها من قبل ، نقسمه على الانحراف المعيارى فيكون المقياس النسبى للالتواء هو :—

$$\frac{\overline{س} - م}{\sigma}$$

ويلاحظ أنه إذا كان هذا المقياس موجباً فإن المنحنى يكون ملتوياً التواء موجباً (نحو اليمين) وإذا كان سالباً فإن المنحنى يكون ملتوياً التواء سالباً (نحو اليسار) وذلك لأن قيمة المنوال تكون دائماً أقل من الوسط الحسابى فى حالة المنحنيات الملتوية نحو اليمين وبالعكس . وباستخدام هذه الطريقة فى قياس الالتواء يمكننا أن نعين درجة التواء المنحنى التكرارى لأعمار الطلبة المتقدمين للالتحاق بكلية التجارة كالآتى :—

$$\frac{\overline{س} - م}{\sigma} = \frac{٢١,٥٢ - ٢٠,٤٦٥}{٣,٧٨٧٥} = ٠,٢٧٩$$

وحيث أنه فى حالة المنحنيات الملتوية التواء بسيطاً تكون :

$\overline{س} - م = ٣ (\overline{س} - و)$ حيث وتمثل الوسيط ، فيمكن اعتبار الكمية $\frac{٣ (\overline{س} - و)}{\sigma}$ مقياساً آخر للالتواء على ألا يستخدم إذا كان المنحنى التكرارى شديد الالتواء .

---

(٥) راجع ص ١٧٨ من هذا الكتاب .

٢ – قياس الالتواء بمقارنة الربيعين .

رأينا أنه فى حالة التوزيعات التكرارية المتماثلة يقع الربيعان على بعدين متساويين من الوسيط وإذا كان المنحنى ملتوياً نحو اليمين كان الربيع الأعلى أكثر بعداً عن الوسيط من الربيع الأدنى . واو كان المنحنى ملتوياً نحو اليسار كان الربيع الأدنى أكثر بعداً عن الوسيط .

وعلى هذا الأساس يمكننا استخدام الفرق بين بعدى الربيعين عن الوسيط فى قياس درجة التواء المنحنيات . ويكون مقياس الالتواء هو :–

$(ع_٣ - و) - (و - ع_١)$ حيث ($ع_٣$ تمثل الربيع الأعلى ، $ع_١$ تمثل الربيع الأدنى) .

فإذا كانت هذه القيمة موجبة كان المنحنى ملتوياً التواء موجباً وبالعكس .

غير أنه يلاحظ أن هذا المقياس مطلق ولا يصلح لأغراض المقارنة .

ولكى نجعله مقياساً نسبياً مستقلا عن وحدات القياس نقسمه على الانحراف الربيعى فيكون مقياس الالتواء النسبى هو :–

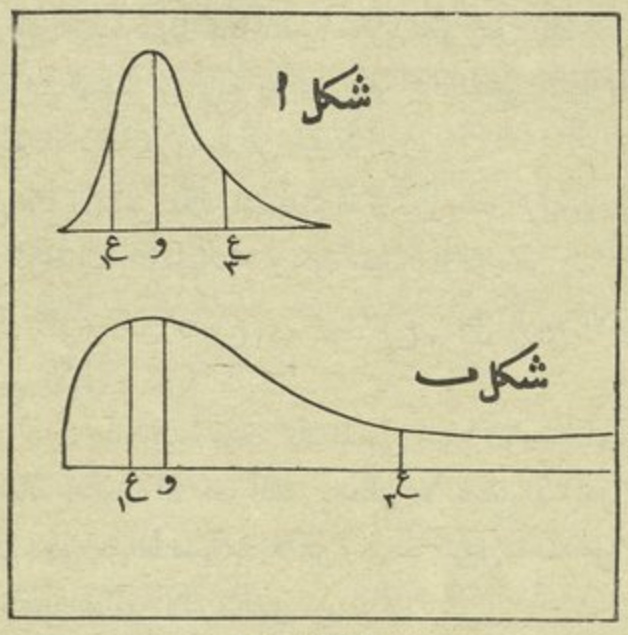
$$\frac{(ع_٣ - و) - (و - ع_١)}{\frac{ع_٣ - ع_١}{٢}} = \frac{٢(ع_١ + ع_٣ - ٢و)}{ع_٣ - ع_١}$$

وحيث أن هذا المقياس فى صورته هذه لا يمكن أن يتغير إلا فى حدود + ٢ فقد جرت العادة على أن يقسم على ٢ كى يصبح مدى تغيره ± ١ . وعلى ذلك يكون مقياس أو معامل الالتواء الربيعى هو : $\frac{ع_١ + ع_٣ - ٢و}{ع_٣ - ع_١}$

ولإثبات أن الكمية $\frac{٢(ع_١ + ع_٣ - ٢و)}{ع_٣ - ع_١}$ لايمكن أن تتغير إلا فى حدود ± ٢

نفرض أولا أن المنحنى ملتو التواء شديداً نحو اليمين . كما زاد هذا

الالتواء كلما بعدت ع٣ عن « و » فى الوقت الذى تقترب فيه ع١ من « و » كما يتضح من مقارنة ا ، ب بشكل رقم ٥٠ .



(شكل رقم ٥٠)

رسم يوضح تغير مركز الربيعين بالنسبة للوسيط حسب درجة التواء المنحنى

والآن إفرض أن الالتواء إلى اليمين زاد إلى أقصى حد ممكن حتى اقتربت ع١ من « و » بدرجة تسمح لنا بفرض تساوى هاتين القيمتين ، أى أنه فى أقصى حالات الالتواء إلى اليمين سنعتبر أن ع١ = و

∴ بالتعويض عن قيمة ع١ فى الكمية التى تمثل مقياس الالتواء الربيعى نجد أن هذه الكمية = $\frac{٢ (و + ع_٣ - ٢ و)}{(ع_٣ - و)}$ = + ٢

وبالمثل يمكننا فى أقصى حالات الالتواء إلى اليسار أن نفرض أن ع٣ = و فيكون مقياس الالتواء الربيعى = $\frac{٢ (ع_١ + و - ٢ و)}{و - ع_١}$

$$= \frac{٢ (ع_١ - و)}{(و - ع_١)}$$

$$= \frac{-٢ (و - ع_١)}{(و - ع_١)} = -٢$$

أى أن المقياس لا يمكن أن يتغير إلا فى حدود $\pm$ ٢

وعليه فإن هذا المقياس مقسوماً على ٢ أى $\frac{ع_١ + ع_٣ - ٢ و}{ع_٣ - ع_١}$

لا يمكن أن يتغير إلا فى حدود $\pm$ ١

وباستخدام هذا المقياس فى إيجاد درجة التواء المنحنى التكرارى الذى يمثل أعمار الطلبة المتقدمين للالتحاق بكلية التجارة نجد أن :-

$$\text{معامل الالتواء الربيعى} = \frac{(٢٢,٦٢٥ + ١٩,٠٩١ - ٢ \times ٢٠,٧٦٩)}{٢٢,٦٢٥ - ١٩,٠٩١}$$

$$= ٠,٠٥ \text{ تقريباً}$$

ويلاحظ أن اختلاف النتائج التى حصلنا عليها باستخدام مقاييس الالتواء الثلاثة لايدل على خطأ أحدها وصحة الآخر . فالواقع أن جميعها صحيحة غير أن كلا منهايقيس درجة الالتواء بمستوى أو معيار خاص .

مقاييس التفرطح :

أشرنا من قبل إلى أنه يمكن التمييز بين المنحنيات التكرارية على أساس التدبب والتفرطح بالإضافة إلى التماثل والالتواء . وكما أن قيمة العزم الثالث تبين لنا درجة التواء المنحنى فإن قيمة العزم الرابع تستخدم لقياس درجة التفرطح .

ومقياس التفرطح أو معامل التفرطح هو $\frac{م_٤}{\sigma^٤}$ ونرمز إلى هذا المعامل بالحرف ل$_٢$ (١) .

(١) التعبير الرمزى الأجنبى عن معامل التفرطح هو β_2 .

ويلاحظ أن قسمة العزم الرابع على σ^4 ترجع إلى نفس السبب الذى من أجله قسمنا العزم الثالث على σ^3 عند تعيين معامل الالتواء .

والكمية $ل_2 = 3$ فى حالة المنحنى الطبيعى أو المنحنيات المعادلة له فى درجة التفرطح (Mesokurtic curves) وتكون أكبر من ٣ فى حالة المنحنيات المدببة القمة (Leptokurtic curves) بينما تكون أقل من ٣ فى حالة المنحنيات المنبسطة أو المفرطحة (Platykurtic curves) ولهذا السبب فإن بعض الإحصائيين يفضل استعمال الكمية ($ل_2 - 3$) كمقياس للتفرطح حتى تكون قيمتها موجبة فى حالة المنحنيات المدببة القمة وسالبة فى حالة المنحنيات المفرطحة وتساوى صفراً فى حالة المنحنيات المعتدلة التفرطح .

ومن الواضح أنه إذا حاولنا إيجاد قيمة العزم الرابع حول الوسط الحسابى عن طريق استخراج قيمة $\Sigma ح^4$ فإن العمل الحسابى يكون مرهقاً فضلا عن تعرضه للأخطاء ولذلك نلجأ إلى إيجاد قيمة العزم الرابع حول أى وسط فرضى نختاره ثم نستخدم العلاقة :—

$$\frac{\Sigma ح^4}{ن} = \frac{\Sigma ى^4}{ن} - 4\frac{\Sigma ى}{ن} \cdot \frac{\Sigma ى^3}{ن} + 6\left(\frac{\Sigma ى}{ن}\right)^2 \cdot \frac{\Sigma ى^2}{ن} - 3\left(\frac{\Sigma ى}{ن}\right)^4$$

(حيث ح تمثل الانحرافات عن الوسط الحسابى ، ى تمثل الانحرافات عن الوسط الفرضى) . وفى حالة التوزيع التكرارى تتخذ هذه العلاقة الصورة الآتية :—

$$\frac{\Sigma و ح^4}{\Sigma و} = \frac{\Sigma و ى^4}{\Sigma و} - 4\frac{\Sigma و ى}{\Sigma و} \cdot \frac{\Sigma و ى^3}{\Sigma و} + 6\left(\frac{\Sigma و ى}{\Sigma و}\right)^2 \cdot \frac{\Sigma و ى^2}{\Sigma و} - 3\left(\frac{\Sigma و ى}{\Sigma و}\right)^4$$

ويمكن استنتاج هذه العلاقة بنفس الطريقة التى اتبعناها فى إيجاد العلاقة بين العزم الثالث حول الوسط الحسابى والعزم الثالث حول الوسط الفرضى . والجدول الآتى يوضح إيجاد العزم الرابع بهذه الطريقة .

جدول رقم ٣٢

إيجاد العزم الرابع التكرارى لأعمار الطلبة المتقدمين للالتحاق بكلية التجارة ١٩٤٧ – ١٩٤٨

مراكز الفئات	التكرار و	ء بوحدات طول الفئة	ء و	ء٢ و	ء٣ و	ء٤ و
١٧	٢٠	٢–	٤٠–	٨٠	١٦٠–	٣٢٠
١٩	٥٥	١–	٥٥–	٥٥	٥٥–	٥٥
٢١	٦٥	٠	٠	٠	٠	٠
٢٣	٣٢	١	٣٢	٣٢	٣٢	٣٢
٢٥	٨	٢	١٦	٣٢	٦٤	١٢٨
٢٧	٥	٣	١٥	٤٥	١٣٥	٤٠٥
٢٩	٢	٤	٨	٣٢	١٢٨	٥١٢
٣١	٦	٥	٣٠	١٥٠	٧٥٠	٣٧٥٠
٣٣	٣	٦	١٨	١٠٨	٦٤٨	٣٨٨٨
٣٥	٤	٧	٢٨	١٩٦	١٣٧٢	٩٦٠٤
المجموع	٢٠٠		٥٢	٧٣٠	٢٩١٤	١٨٦٩٤

$$م_٤ = \frac{\Sigma ء^٤ و}{\Sigma و} - ٤ \frac{\Sigma ء و}{\Sigma و} \cdot \frac{\Sigma ء^٣ و}{\Sigma و} + ٦ \left(\frac{\Sigma ء و}{\Sigma و}\right)^٢ \frac{\Sigma ء^٢ و}{\Sigma و} - ٣ \left(\frac{\Sigma ء و}{\Sigma و}\right)^٤$$

$$= \frac{١٨٦٩٤}{٢٠٠} - ٤ \times \frac{٥٢}{٢٠٠} \times \frac{٢٩١٤}{٢٠٠} + ٦ \times \left(\frac{٥٢}{٢٠٠}\right)^٢ \times \frac{٧٣٠}{٢٠٠} - ٣ \left(\frac{٥٢}{٢٠٠}\right)^٤$$

$$= ٧٩,٧٨٣٩٣٠٧٢$$

$$\therefore \text{ل}_{٢} = \frac{\text{م}_{٤}}{\sigma^{٤}} = \frac{٧٩,٧٨٣٩٣٠٧٢}{(١,٨٩٢٥)^{٤}} = ٦,٢٣$$

ويلاحظ أن قيمة ل$_{٢}$ فى هذه الحالة $>$ ٣ مما يدل على أن المنحنى الذى يمثل التوزيع التكرارى لأعمار الطلبة لا بد وأن يكون مدبب القمة . وبالرجوع إلى شكل ٣٦ نجد ما يؤيد ذلك .

التشتت والتفرطح :

يبدو لأول وهلة بالنظر إلى بعض الأشكال المختلفة للمنحنيات التكراريه أنه إذا كان المنحنى التكرارى مدبب القمة فإن تشتت المفردات حول متوسطها يكون صغيراً بينما إذا كان المنحنى مفرطحاً فإن تشتت المفردات يكون كبيراً . وعلى الرغم من أن كثيراً من المنحنيات المدببة تصف توزيعات درجة تشتتها صغيرة وأن كثيراً من المنحنيات المفرطحة تصف توزيعات درجة تشتتها كبيرة إلا أنه ليس من الضرورى أن تكون هناك صلة ما بين درجة التشتت ودرجة التفرطح فكلا منها يصف خاصية مستقلة تماماً فى التوزيع التكرارى .

فالذى نقصده بالتفرطح هو تفرطح قمة المنحنى وليس انفراج المنحنى من أسفل أى اتساع المدى الممتد بين طرفيه على المحور الأفقى . ولزيادة إيضاح الفرق بين التشتت والتفرطح نأخذ المثال الآتى الذى يوضح المقارنة بين توزيعين تكراريين متماثلين ودرجة تشتتهما واحدة غير أن درجة تفرطح كل منهما مختلفة :ـ

جدول رقم ٣٣

التوزيع الأول

حساب العزم الثانى والعزم الرابع لتوزيع متماثل ومدبب القمة

قيم المتغير	التكرار	ح بوحدات طول الفئة	ح^{٢}	ح^{٢}	ح^{٤} و
١٠	١	—٤	١٦	١٦	٢٥٦
٢٠	٢	—٣	٩	١٨	١٦٢
٣٠	٥	—٢	٤	٢٠	٨٠
٤٠	١٦	—١	١	١٦	١٦
٥٠	٥٢	٠	٠	٠	٠
٦٠	١٦	١	١	١٦	١٦
٧٠	٥	٢	٤	٢٠	٨٠
٨٠	٢	٣	٩	١٨	١٦٢
٩٠	١	٤	١٦	١٦	٢٥٦
المجموع	١٠٠			١٤٠	١٠٢٨

$$م_{٢} = \frac{١٤٠}{١٠٠} = ١,٤$$

$$، م_{٤} = \frac{١٠٢٨}{١٠٠} = ١٠,٢٨$$

$$\text{معامل الاختلاف} = \frac{\sqrt{١,٤}}{٥٠} \times ١٠٠ = ٢,٣٦٦\ ٪$$

$$، \quad ل_{٢} = \frac{١٠,٢٨}{(١,٤)^{٢}} = ٥,٢٤$$

وحيث إن ل$_{٢}$ > ٣ ∴ المنحنى التكرارى الذى يمثل هذا التوزيع مدبب القمة .

جدول رقم ٣٤

التوزيع الثانى

حساب العزم الثانى والعزم الرابع لتوزيع معتدل التفرطح

قيم المتغير	التكرار	ح	ح^{٢}	ح^{٢}	ح^{٤}و
١٠	٠	٤—	١٦	٠	٠
٢٠	٢	٣—	٩	١٨	١٦٢
٣٠	٨	٢—	٤	٣٢	١٢٨
٤٠	٢٠	١—	١	٢٠	٢٠
٥٠	٤٠	٠	٠	٠	٠
٦٠	٢٠	١	١	٢٠	٢٠
٧٠	٨	٢	٤	٣٢	١٢٨
٨٠	٢	٣	٩	١٨	١٦٢
٩٠	٠	٤	١٦	٠	٠
المجموع	١٠٠			١٤٠	٦٢٠

$$م_{٢} = \frac{١٤٠}{١٠٠} = ١,٤$$

$$م_{٤} = \frac{٦٢٠}{١٠٠} = ٦,٢$$

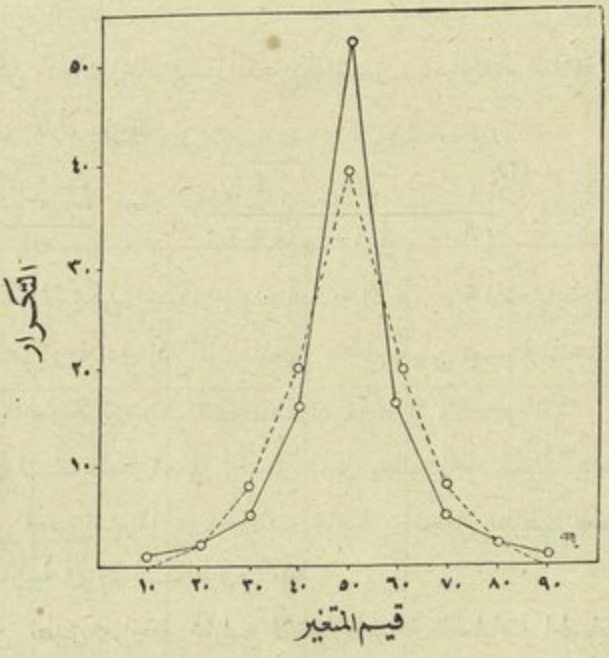
$$\text{معامل الاختلاف} = \frac{\sqrt{١,٤}}{٥٠} \times ١٠٠ = ٢,٣٦٦ \%$$

$$\text{، } ل_{٢} = \frac{٦,٢}{(١,٤)^{٢}} = ٣,١٦$$

وحيث إن قيمة ل$_{٢}$ لاتختلف كثيراً عن ٣ فإن التوزيع معتدل التفرطح .

وبمقارنة النتائج التى حصلنا عليها فى الحالتين نجد أن درجة التشتت واحدة فى

التوزيعين سواء استخدمنا الانحراف المعيارى أو معامل الاختلاف لمقارنة درجة التشتت فى كل منهما . أما درجة التفرطح فتختلف اختلافا كبيراً فى التوزيعين كما يتضح من مقارنة قيمة ل$_{٢}$ فى الحالتين ومن الشكل الآتى الذى يوضح المنحنيين التكراريين لهذين التوزيعين :—



(شكل رقم ٥١)

مضلعان تكراريان متساويان فى التشتت ومتفاوتان فى التفرطح

إيجاد قيمة المنوال باستخدام معاملى الالتواء والتفرطح :

من المشاهد أنه فى حالة التوزيعات التكرارية الملتوية التواء بسيطاً فإن قيمة $\frac{م - ٣}{٣\sigma}$ أى $\sqrt{ل_{١}}$ وهى أحد مقاييس الالتواء تعادل ضعف قيمة $\frac{\bar{س} - م}{\sigma}$ وهى مقياس آخر للالتواء كما بينا .

$$\text{أى أن} = \frac{(\bar{س} - م)}{\sigma} = \frac{\sqrt{ل_١}}{٢} \quad (١)$$

وباستخدام هذه العلاقة يمكن إيجاد قيمة للمنوال أدق من تلك التى حصلنا عليها باتباع الطرق التقريبية التى ذكرناها فى الباب الخامس بشرط أن يكون التواء التوزيع بسيطاً .

ويمكن كذلك إيجاد قيمة أكثر دقة للمنوال باستخدام العلاقة الآتية التى تنسب إلى كارل بيرسون :-

$$\frac{\bar{س} - م}{\sigma} = \frac{\sqrt{ل_١}\ (ل_٢ + ٣)}{٢\ (٥\ ل_٢ - ٦\ ل_١ - ٩)} \quad (٢)$$

على أنه لا يمكن استخدام هذه العلاقة إلا فى حالة المنحنيات الملتوية التواء بسيطاً أو على وجه أدق إذا كان المنحنى يدخل ضمن مجموعة المنحنيات الملتوية المعروفة بمجموعة بيرسون . « Pearson's System of Curves » .

ولا شك أن مشقة العمل الحسابى الذى يتطلبه إيجاد المنوال بهذه الطرق ، علاوة على قصر استعمالها على حالات خاصة ، تجعل تطبيقها العملى نادراً .

مراجعة شارلييه فى حساب العزوم العليا :

يمكن تطبيق مراجعة شارلييه لاختبار صحة العمليات الحسابية الخاصة بتعيين قيم العزوم العليا للتوزيع التكرارى كما فعلنا فى حالة الوسط الحسابى والانحراف المعيارى . وتطبيق مراجعة شارلييه فى حالة العزم الثالث يتطلب إيجاد قيمة Σ و $(ى + ١)^{٣}$ ثم اختبار صحة المعادلة :

$$\Sigma \text{ و } (ى + ١)^{٣} = \Sigma \text{ و } ى^{٣} + ٣\ \Sigma \text{ و } ى^{٢} + ٣\ \Sigma \text{ و } ى + \Sigma \text{ و}$$

من واقع الأرقام التى نحصل عليها .

وفى حالة العزم الرابع تتطلب المراجعة إيجاد قيمة Σ و $(ى + ١)^{٤}$ ثم

---

(١) انظر كتاب Elements of statistical method by Albert waugh ص ٢٠٥ وكتاب Elementary statistics by Camp ص ٤٧ .

(٢) انظر كتاب The methods of statistics by L. H. C. Tippet ص ٣٤ .

اختبار صحة المعادلة .

$$\Sigma و (ى + ١)^{٤} = \Sigma و ى^{٤} + ٤ \Sigma و ى^{٣} + ٦ \Sigma و ى^{٢} + ٤ \Sigma و ى + \Sigma و$$

فإذا كانت هذه المعادلة صحيحة طبقاً للأرقام التى نحصل عليها كانت العمليات الحسابية صحيحة . والجدول الآتى يوضح الخطوات اللازمة لتطبيق مراجعة شارلييه فى حساب العزوم العليا .

جدول رقم ٣٥

مراجعة شارلييه فى حساب العزوم العليا

للتوزيع التكرارى لأعمار الطلبة المتقدمين للالتحاق بكلية التجارة سنة ١٩٤٧

مراكز الفئات	التكرار	ى	ى + ١	(ى + ١)^{٣}	و(ى + ١)^{٣}	و(ى + ١)^{٤}
١٧	٢٠	−٢	−١	−٢٠	−٢٠	٢٠
١٩	٥٥	−١	٠	٠	٠	٠
٢١	٦٥	٠	١	١	٦٥	٦٥
٢٣	٣٢	١	٢	٨	٢٥٦	٥١٢
٢٥	٨	٢	٣	٢٧	٢١٦	٦٤٨
٢٧	٥	٣	٤	٦٤	٣٢٠	١٢٨٠
٢٩	٢	٤	٥	١٢٥	٢٥٠	١٢٥٠
٣١	٦	٥	٦	٢١٦	١٢٩٦	٧٧٧٦
٣٣	٣	٦	٧	٣٤٣	١٠٢٩	٧٢٠٣
٣٥	٤	٧	٨	٥١٢	٢٠٤٨	١٦٣٨٤
المجموع	٢٠٠				٥٤٦٠	٣٥١٣٨

فى حالة العزم الثالث نجد أن

$$٥٤٦٠ = ٢٩١٤ + ٣ \times ٧٣٠ + ٣ \times ٥٢ + ٢٠٠$$

$$= ٥٤٦٠ \quad \therefore \text{ العمليات الحسابية صحيحة}$$

فى حالة العزم الرابع نجد أن

٣٥١٣٨ = ١٨٦٩٤ + ٤ × ٢٩١٤ + ٦ × ٧٣٠ + ٤ × ٥٢ + ٢٠٠

= ٣٥١٣٨ ∴ العمليات الحسابية صحية (*)

تصحيحات شبرد فى حساب العزوم العليا :

ذكرنا فى الباب السابق أن الفرض الخاص باعتبار جميع المفردات مركزة عند مركز الفئة لا يمكن اعتباره صحيحاً فى كل الحالات وأوردنا طريقة تصحيح الخطأ المترتب على هذا الفرض فى حساب الانحراف المعيارى وهى الطريقة التى اقترحها W. F. Sheppard . وقد وضع « شبرد » كذلك تصحيحاً للخطأ المترتب على هذا الفرض عند حساب قيمة العزم الرابع للتوزيع التكرارى . أما العزم الثالث فقيمته لا تحتاج إلى تعديل(١).

والتصحيح فى حالة العزم الرابع هو كالآتى :—

$$م_٤ \text{ (بعد التصحيح)} = م_٤ - \frac{١}{٢} ف^٢ م_٢ + \frac{٧}{٢٤٠} ف^٤$$

(حيث ف تمثل طول فئة التوزيع التكرارى)

وحيث أن العمل الحسابى فى حالة العزم الرابع يكون غالباً على أساس اعتبار طول الفئة يساوى الوحدة فإن التصحيح يكون كالآتى :—

$$م_٤ \text{ (بعد التصحيح)} = م_٤ - \frac{١}{٢} م_٢ + \frac{٧}{٢٤٠}$$

(لأن ف = ١) .

ويلاحظ أن عمل هذا التصحيح مقيد بنفس الشروط التى ذكرهانا فيما يتعلق بتصحيح قيمة الانحراف المعيارى(٢) .

(*) يلاحظ أننا استعملنا قيم $ح_٤ و$ ، $ح_٣ و$ ، $ح_٢ و$ ، $ح_١ و$ المبينة بالجدول رقم ٣٢ .

(١) يلاحظ أن قيمة العزم الأول لا تحتاج إلى تعديل كذلك .

(٢) راجع ص ١٩٩ من هذا الكتاب .

الباب الثامن

مقدمة لنظرية الاحتمالات والعينات

التوزيعات التكرارية النظرية

التوزيعات التكرارية التى درسناها حتى الآن هى عبارة عن توزيعات تكرارية حقيقية مستمدة من واقع مشاهدة أو قياس الظواهر المختلفة كالسن والوزن والدخل وتوزيع الملكية العقارية . . . إلخ . هناك بجانب هذه التوزيعات نوع آخر من التوزيعات التكرارية التى يمكن الحصول عليها أو استنتاجها نظرياً دون حاجة إلى قياس الظواهر أو مشاهدتها وتسجيل قيمتها . وسنخصص هذا الباب لدراسة بعض هذه التوزيعات التكرارية النظرية التى ستكون عظيمة الفائدة فى موضوع العينات . وسيؤدى بنا ذلك إلى دراسة أولية فى نظرية الاحتمالات التى ينبنى عليها البحث الخاص بالعينات . ونورد فيما يلى هذه التوزيعات النظرية :—

١ — التوزيع التكرارى لأحجام جميع العينات الممكنة (حالة خاصة من التوزيع التكرارى لذات الحدين) .

٢ — توزيع برنوللى « Bernoulli » (الحالة العامة للتوزيع التكرارى لذات الحدين) .

٣ — التوزيع الطبيعى أو قانون الأعداد الكبيرة(١) .

---

(١) التعبير الأجنبى المقابل للتوزيع الطبيعى هو (Normal Distribution) والتعبير المقابل لقانون الأعداد الكبيرة هو (Law of Large Numbers) والتعبير الأخير قليل الاستعمال .

٤ – توزيع بواسون « Poisson » أو قانون الاحتمالات الصغيرة(١).

التوزيع التكرارى لأحجام جميع العينات الممكنة

(حالة خاصة من التوزيع التكرارى لذات الحدين)

إفرض أن لدينا مجموعاً معيناً عدد مفرداته (ن) وأردنا أن نختار منه جميع العينات الممكنة من أى حجم نريده (من . إلى ن) ، فما هو عدد جميع الطرق الممكن بها اختيار عينة يتراوح حجمها بين (٠ ، ن) . أو على وجه التفصيل ، ما عدد الطرق التى يمكن بها اختيار عينة عدد مفرداتها (٠) من مجموع عدد مفرداته (ن) وما عدد الطرق التى يمكن بها اختيار عينة عدد مفرداتها (١) من مجموع عدد مفرداته (ن) ، وماعدد الطرق التى يمكن بها اختيار عينة عدد مفرداتها (٢) من مجموع عدد مفرداته (ن) ، وهكذا . . .

من البديهى أن اختيار عينة عدد مفرداتها (٠) هو بمثابة عدم اختيار أية عينة إطلاقاً . وهذا يمكننا أن نفعله بطريقة واحدة . فإذا كان من الممكن الاختيار بعدة طرق فلا يمكن الامتناع عن الاختيار إلا بطريقة واحدة فقط . ويمكن الحصول على هذه النتيجة البديهية بتطبيق قانون التوافيق المعروف الذى ينص على أن عدد الطرق الممكن بها اختيار (م) مفردة من مجموع عدد مفرداته (ن) هو $ق_{م}^{ن}$

---

(١) التعبير الأجنبى المقابل لما اسميناه قانون الاحتمالات الصغيرة هو (Law of Small Numbers) والواقع أن هذا التعبير الأجنبى غير دقيق وقد اعترض على هذه التسمية كثير من الكتاب الأجانب حتى أصبح استعمالها نادراً ولذلك لم نتقيد بترجمتها فعبارة « قانون الأعداد الصغيرة » لا تدل على شىء ما يتصل بهذا التوزيع . أما عبارة « قانون الاحتمالات الصغيرة » فهى تدل حقيقة على طبيعة هذا التوزيع كما سيجد القارىء من دراسته .

بوضع م = .

∴ عدد الطرق التى يمكن بها الحشول على عينة عدد مفرداتها (٠) من مجموع عدد مفرداته (ن) .

$= \text{ن}_{٠}$ = ١

كذلك فإن عدد الطرق التى يمكن بها اختيار عينة عدد مفرداتها (١) من مجموع عدد مقرداتها (ن)

$= \text{ن}_{١}$ = ن

وعدد الطرق التى يمكن بها اختيار عينة مفرداتها (٢) من مجموع عدد مفرداته (ن)

$$= \text{ن}_{٢} = \frac{\text{ن} (\text{ن} - ١)}{٢ !} = \frac{\text{ن} !}{٢ ! (\text{ن} - ٢) !} \quad (*)$$

وعدد الطرق التى يمكن بها اختيار عينة عدد مفرداتها (٣) من مجموع عدد مفرداته (ن)

$$= \text{ن}_{٣} = \frac{\text{ن} (\text{ن} - ١) (\text{ن} - ٢)}{٣ !} = \frac{\text{ن} !}{٣ ! (\text{ن} - ٣) !} \quad (*)$$

وعدد الطرق التى يمكن بها اختيار عينة عدد مفرداتها (ن) من مجموع عدد مفرداته (ن)

$= \text{ن}_{\text{ن}}$ = ١

(*) سنستعمل علامة التعجب « ! » فى هذا الكتاب للدلالة على التباديل تمشياً مع الاتجاه الحديث فى التعبير الرمزى الأجنبى . وعلى ذلك فإن ن ! تعنى « تباديل ن » .

(*) هذه قاعدة معروفة فى دراسة التباديل والتوافيق .

وهذا أيضاً بديهى لأن اختيار عينة عدد مفرداتها (ن) هو بمثابة اختيار المجموع كله وهذا لا يمكن أن نفعله إلا بطريقة واحدة .

يمكننا الآن أن نلخص النتائج السابقة فى شكل جدول تكرارى حيث يكون المتغير هو حجم العينة (أو عدد مفردات العينة) والتكرار هو عدد الطرق الممكن بها اختيار عينة من هذا الحجم من مجموع عدد مفرداته (ن)

التوزيع التكرارى لذات الحدين (حالة خاصة)

حجم العينة	عدد العينات الممكن اختيارها (التكرار)		
٠	$\text{ت}_{\text{ن}}^{\text{ق}}$	=	١
١	$\text{ت}_{\text{ن}}^{\text{ق}_{١}}$	=	ن
٢	$\text{ت}_{\text{ن}}^{\text{ق}_{٢}}$	=	$\frac{\text{ن} (\text{ن} - ١)}{٢ !}$
٣	$\text{ت}_{\text{ن}}^{\text{ق}_{٣}}$	=	$\frac{\text{ن} (\text{ن} - ١) (\text{ن} - ٢)}{٣ !}$
⋮	⋮		⋮
(ن − ٣)	$\text{ت}_{\text{ن}}^{\text{ق}_{\text{ن}-٣}} = \text{ت}_{\text{ن}}^{\text{ق}_{٣}}$	=	$\frac{\text{ن} (\text{ن} - ١) (\text{ن} - ٢)}{٣ !}$
(ن − ٢)	$\text{ت}_{\text{ن}}^{\text{ق}_{\text{ن}-٢}} = \text{ت}_{\text{ن}}^{\text{ق}_{٢}}$	=	$\frac{\text{ن} (\text{ن} - ١)}{٢ !}$
(ن − ١)	$\text{ت}_{\text{ن}}^{\text{ق}_{\text{ن}-١}} = \text{ت}_{\text{ن}}^{\text{ق}_{١}}$	=	ن
ن	$\text{ت}_{\text{ن}}^{\text{ق}_{\text{ن}}} = \text{ت}_{\text{ن}}^{\text{ق}_{٠}}$	=	١

لكى نحصل على عدد جميع العينات الممكن اختيارها من أى حجم فما علينا إلا أن نجمع التكرارات المبينة بالجدول السابق . أى أن عدد جميع العينات الممكن اختيارها من مجموع قدره (ن)

$$= \text{ت}_{\text{ن}}\text{ق}_{\text{ن}} + \text{ت}_{\text{ن}-1}\text{ق}_{\text{ن}-1} + \cdots + \text{ت}_{2}\text{ق}_{2} + \text{ت}_{1}\text{ق}_{1} + \text{ت}_{0}\text{ق}_{0}$$

وبالنظر إلى هذه التكرارات نجد أنها تمثل معاملات س وقواها التصاعدية فى مفكوك ذات الحدين $(1 + \text{س})^{\text{ن}}$. إذ أن :

$$(1 + \text{س})^{\text{ن}} = \text{ت}_{\text{ن}}\text{ق}_{\text{ن}} + \text{ت}_{\text{ن}-1}\text{ق}_{\text{ن}-1}\text{س} + \text{ت}_{\text{ن}-2}\text{ق}_{\text{ن}-2}\text{س}^{2} + \cdots + \text{ت}_{1}\text{ق}_{1}\text{س}^{\text{ن}-1} + \text{ت}_{0}\text{ق}_{0}\text{س}^{\text{ن}}$$

بوضع س = 1 فى الكمية $(1 + \text{س})^{\text{ن}}$

$$\therefore (1 + 1)^{\text{ن}} = \text{ت}_{\text{ن}}\text{ق}_{\text{ن}} + \text{ت}_{\text{ن}-1}\text{ق}_{\text{ن}-1} + \text{ت}_{\text{ن}-2}\text{ق}_{\text{ن}-2} + \cdots + \text{ت}_{1}\text{ق}_{1} + \text{ت}_{0}\text{ق}_{0}$$

ويلاحظ أن حدود الطرف الأيسر فى المعادلة هى نفسها تكرارات العينات المبينة بالجدول السابق .

$$\therefore \text{مجموع التكرارات} = (1 + 1)^{\text{ن}} = 2^{\text{ن}}$$

تماثل التوزيع التكرارى لأحجام العينات

بالنظر إلى الجدول الذى وضح تكرارات العينات ذات الأحجام المختلفة يتضح فى الحال أن هذا التوزيع التكرارى متماثل . وذلك لأن :ـ

تكرار العينة التى حجمها (٠) تكرار العينة التى حجمها ن

» » » (١) = » » » » (ن — ١)

» » » (٢) = » » » » (ن — ٢)

» » » (٣) = » » » » (ن — ٣)

فمن البدهى أن كل طريقة نختار بها عينة عدد مفرداتها (١) هى أيضا طريقة نترك بها عينة عدد مفرداتها (ن — ١) وكل طريقة نختار بها عينة عدد مفرداتها (٢) هى أيضاً طريقة نترك بها عينة عدد مفرداتها (ن — ٢) .

والواقع أنه يمكن استنتاج ظاهرة تماثل هذا التوزيع مباشرة من قانون التوافق المعروف .

$$\text{ن ق ر} = \text{ن ق ن-ر}$$

ومن الواضح أن التماثل هنا يكون حول الحجم المتوسط للعينة أى حول القيمة $\frac{\text{ن}}{٢}$ التى تقع فى النصف تماماً .

وعلى الرغم من أنه يظهر بمجرد النظر إلى الجدول أن الوسط الحسابى لهذا التوزيع التكرارى المتماثل هو $\frac{\text{ن}}{٢}$ فسنؤيد هذه النتيجة باستخراج الوسط الحسابى للتوزيع بالطريقة المعتادة .

الوسط الحسابى للتوزيع التكرارى لأحجام العينات (التوزيع المتماثل لذات الحدين) .

لإيجاد الوسط الحسابى لهذا التوزيع نضرب كل قيمة فى التكرار المقابل لها ونقسم على مجموع التكرارات وقدره $٢^{\text{ن}}$

$$\text{الوسط الحسابى} = \frac{١}{٢^{ن}} \left\{ (٠) \times \text{ن ق}_{٠} + (١) \times \text{ن ق}_{١} \right.$$
$$\left. + (٢) \times \text{ن ق}_{٢} + \cdots\cdots + (\text{ن}) \times \text{ن ق}_{\text{ن}} \right\}$$

ولإيجاد مجموع الحدود المحصورة بين القوسين الكبيرين نتبع الطريقة الآتية :-

$$(١ + \text{س})^{\text{ن}} = \text{ن ق}_{٠} + \text{ن ق}_{١}\text{س} + \text{ن ق}_{٢}\text{س}^{٢} + \text{ن ق}_{٣}\text{س}^{٣}$$
$$+ \cdots\cdots + \text{ن ق}_{\text{ن}}\text{س}^{\text{ن}}$$

بمفاضلة طرفى المعادلة بالنسبة للمتغير س

$$\therefore \text{ن}(١ + \text{س})^{\text{ن}-١} = ٠ + \text{ن ق}_{١} + (٢) \times \text{ن ق}_{٢}\text{س}$$
$$+ (٣) \times \text{ن ق}_{٣}\text{س}^{٢} + \cdots\cdots + (\text{ن}) \text{ن ق}_{\text{ن}}\text{س}^{\text{ن}-١}$$

بوضع س = ١ فى المعادلة الأخيرة

$$\therefore \text{ن}(١ + ١)^{\text{ن}-١} = ٠ + (١) \times \text{ن ق}_{١} + (٢) \times \text{ن ق}_{٢}$$
$$+ (٣) \times \text{ن ق}_{٣} + \cdots\cdots + (\text{ن}) \text{ن ق}_{\text{ن}}$$

ومن الواضح أن الطرف الأيسر فى هذه المعادلة هو عبارة عن الحدود المطلوب إيجاد مجموعها وهى المحصورة بين القوسين $\{\ \}$

$$\therefore \text{الوسط الحسابى للتوزيع التكرارى} = \frac{\text{ن}(١+١)^{\text{ن}-١}}{٢^{\text{ن}}} = \frac{\text{ن} \times ٢^{\text{ن}-١}}{٢^{\text{ن}}} = \frac{\text{ن}}{٢}$$

الانحراف المعيارى للتوزيع

لإيجاد الانحراف المعيارى للتوزيع التكرارى المتماثل لذات الحدين ، نحسب أولا مجموع مربعات الانحرافات عن وسط فرضى قيمته (٠) ثم نحصل على الانحراف المعيارى باستخدام العلاقة :-

$$\frac{\Sigma ح^{٢} و}{\Sigma و} = \frac{\Sigma س^{٢} و}{\Sigma و} - (\bar{س})^{٢} \quad (*)$$

$$\Sigma س^{٢} و = (٠)^{٢} \times ق_{٠} + (١)^{٢} \times ق_{١} + (٢)^{٢} \times ق_{٢} + \cdots$$
$$+ (ن)^{٢} \times ق_{ن}$$

للحصول على مجموع حدود الطرف الأيسر نتبع نفس الطريقة التى اتبعناها فى إيجاد الوسط الحسابى

$$(١+س)^{ن} = ق_{٠} + ق_{١} س + ق_{٢} س^{٢} + ق_{٣} س^{٣} + \cdots$$
$$+ ق_{ن} س^{ن}$$

بمفاضلة طرفى المعادلة بالنسبة للمتغير س

$$\therefore ن (١ + س)^{ن-١} = ق_{١} + (٢) ق_{٢} س + (٣) ق_{٣} س^{٢} + \cdots$$
$$+ (ن) ق_{ن} س^{ن-١}$$

بضرب طرفى هذه المعادلة الأخيرة فى س

(*) راجع ص ١٩٣ من هذا الكتاب .

$$\therefore \text{ن } (١+\text{س})^{\text{ن}-١} \text{ س} = \text{ن}\,\text{ق}_{١} \text{ س} + (٢)\,\text{ن}\,\text{ق}_{٢} \text{ س} +$$

$$(٣)\,\text{ن}\,\text{ق}_{٣} \text{س}^{٣} + \cdots\cdots + (\text{ن})\,\text{ن}\,\text{ق}_{\text{ن}} \text{س}^{\text{ن}}$$

بمفاضلة طرفى هذه المعادلة بالنسبة ل س نحصل فى الطرف الأيمن على :

$$\text{س} \times \text{ن } (\text{ن}-١)\ (١+\text{س})^{\text{ن}-٢} + \text{ن } (١+\text{س})^{\text{ن}-١} \times ١$$

وفى الطرف الأيسر على :

$$\text{ن}\,\text{ق}_{١} + (٢)^{٢}\,\text{ن}\,\text{ق}_{٢}\,\text{س} + (٣)^{٢}\,\text{ن}\,\text{ق}_{٣}\,\text{س}^{٢} + \cdots$$
$$+ (\text{ن})^{٢}\,\text{ن}\,\text{ق}_{\text{ن}}\,\text{س}^{\text{ن}-١}$$

بوضع س = ١

يصبح الطرف الأيمن = ن (ن − ١) ٢$^{\text{ن}-٢}$ + ن × ٢$^{\text{ن}-١}$

والطرف الأيسر = $\text{ن}\,\text{ق}_{١} + (٢)^{٢}\,\text{ن}\,\text{ق}_{٢} + (٣)^{٢}\,\text{ن}\,\text{ق}_{٣} + \cdots$

$$+ (\text{ن})^{٢}\,\text{ن}\,\text{ق}_{\text{ن}}$$

ومن الواضح أن حدود الطرف الأيسر هى عبارة عن $\Sigma\,\text{س}^{٢}\,\text{و}$

أى أن $\Sigma\,\text{س}^{٢}\,\text{و}$ = ن (ن − ١) ٢$^{\text{ن}-٢}$ + ن × ٢$^{\text{ن}-١}$

$$\therefore \frac{\Sigma\,\text{س}^{٢}\,\text{و}}{\Sigma\,\text{و}} = \frac{١}{\text{ن}\,\text{ق}_{٢}} \left\{ \text{ن } (\text{ن}-١)\,٢^{\text{ن}-٢} + \text{ن} \times ٢^{\text{ن}-١} \right\}$$

$$= \frac{\text{ن } (\text{ن}-١)}{٤} + \frac{\text{ن}}{٢}$$

$$\therefore \frac{\Sigma\,\text{ح}^{٢}\,\text{و}}{\Sigma\,\text{و}} = \left\{ \frac{\text{ن } (\text{ن}-١)}{٤} + \frac{\text{ن}}{٢} \right\} - \left(\frac{\text{ن}}{٢}\right)^{٢}$$

$$\left(\text{لاحظ أن } \bar{\text{س}} = \frac{\text{ن}}{٢} \right)$$

∴ مربع الانحراف المعيارى حول الوسط الحسابى $= \frac{ن(ن-١)}{٤} + \frac{ن}{٢} - \frac{ن^{٢}}{٤}$

$$= \frac{ن^{٢} - ن + ٢ن - ن^{٢}}{٤} = \frac{ن}{٤}$$

∴ الانحراف المعيارى للتوزيع التكرارى $= \sqrt{\frac{ن}{٤}} = \frac{١}{٢}\sqrt{ن}$

ويمكن الحصول على الوسط الحسابى والانحراف المعيارى للتوزيع التكرارى لذات الحدين بدون استخدام التفاضل كالآتى :-

$$\text{الوسط الحسابى} = \frac{\Sigma س و}{\Sigma و}$$

$$= \frac{١}{٢^{ن}} \left\{ ٠ \times ق_{ن} + (١) \times ق_{ن-١} + (٢) \times ق_{ن-٢} + \cdots + (ن) \times ق_{ن} \right\}$$

$$= \frac{١}{٢^{ن}} \left\{ (١) \times ن + (٢) \times \frac{ن(ن-١)}{!٢} + (٣) \times \frac{ن(ن-١)(ن-٢)}{!٣} + \cdots + (ن) \times ١ \right\}$$

$$= \frac{ن}{٢^{ن}} \left\{ ١ + (ن-١) + \frac{(ن-١)(ن-٢)}{!٢} + \cdots\cdots + ١ \right\}$$

وبالنظر إلى الحدود بين القوسين الكبيرين نجد أنها عبارة عن مفكوك ذات الحدين $(١+١)^{ن-١}$

∴ الوسط الحسابى $= \frac{ن}{٢^{ن}} (١+١)^{ن-١}$

$$= \frac{ن \times ٢^{ن-١}}{٢^{ن}} = \frac{ن}{٢}$$

وبطريقة مماثلة لهذه يمكن الحصول على الانحراف المعيارى للتوزيع التكرارى المتماثل لذات الحدين بدون استخدام التفاضل

مبادئ أولية فى نظرية الاحتمالات

المعنى العادى والمعنى الرياضى للاحتمال .

الاحتمال بالمعنى العادى المألوف يفيد الإمكان المقرون ببعض الثقة وبعض الشك فى وقوع حدث معين . فقد تكون ثقتنا فى وقوع حدث معين معدومة ومع ذلك فإن تمسكنا بأهداب الناحية المأمونة يدفعنا إلى القول بأن الحدث محتمل الوقوع .

ومن الناحية الأخرى فقد تكون ثقتنا فى وقوع الحدث تكاد تبلغ حد التأكيد ومع ذلك فلكى لا نتهم بأننا نلقى الكلام جزافاً نكتفى بالقول بأن الحدث محتمل الوقوع . وهكذا نرى أن المعنى العادى للاحتمال يشمل الاحتمالات الضعيفة والقوية على السواء وليس فيه أى أثر للقياس الدقيق . غير أننا نسمع أحياناً فى المحادثات العادية شخصاً يبلغ به الحماس إلى حد القول أن احتمال وقوع حدث معين (كاحتمال فوز جواد معين فى حلبة السبق) هو ١٠٠٪ أو ٩٩٪ إظهاراً لدرجة ثقته فى وقوع الحدث . هذا القياس هو طبعاً مجرد تقدير شخصى بحت غير مستند إلى أى أساس منطقى . وقد يقول شخص آخر بنفس الحماس أن احتمال وقوع الحدث ذاته هو ٥٪ حيث تقديره هو أيضاً شخصى بحت .

أما الاحتمال الرياضى فيتميز عن الاحتمال فى معناه العادى فى نقطتين جوهريتين :—

أولا : أنه يمكن قياسه وتحديده بدقة إذا توفرت لدينا معلومات معينة عن الحدث المراد تقدير احتمال وقوعه .

ثانياً : أنه مقياس موضوعى مستقل عن التقدير الشخصى . بمعنى أنه لا يمكن أن يختلف اثنان فى قياس الاحتمال الرياضى لوقوع حدث معين إذا كان هذا القياس ممكنا .

وقبل أن نعطى التعريف العام للاحتمال الرياضى نسوق بعض الأمثلة التى تساعد على تفهم معناه . إذا رمينا قطعة من النقود كاملة التوازن إلى أعلى بدون أى تحيز فما هو احتمال ظهور الوجه عندما تستقر ؟

من الواضح أنه إذا رمينا القطعة وتركناها تستقر فلا يمكن أن يحدث إلا أحد أمرين :—

إما أن تسقط على الوجه أو تسقط على الظهر . وما دام ليس هناك أى سبب معروف يدعونا إلى الاعتقاد بأن القطعة ستسقط على أحد الوجهين دون الآخر ، فلا يمكننا إلا أن نقول إن الاحتمالين متعادلان . أى أن احتمال سقوط قطعة النقود على الظهر يعادل فى القوة احتمال سقوطها على الوجه لأنه ليس ثمة سبب يدعو القطعة إلى الاستقرار على وجه دون الآخر إلا تلك المجموعة من العوامل غير القابلة للتحديد التى نطلق عليها اسم المصادفة .

وإذا كانت المصادفة هى التى جعلت القطعة تسقط على الظهر فى إحدى المرات فهى وحدها كفيلة بأن تجعلها تسقط على الوجه فى مرة أخرى .

وعلى ذلك فإن جعلنا الواحد الصحيح مقياساً للتأكيد فإن هذا « الواحد » يكون موزعاً بالتساوى بين الاحتمالين المتعادلين فى القوة — أى سقوط قطعة النقود على الوجه وسقوطها على الظهر .

∴ احتمال سقوط قطعة النقود على الوجه = احتمال سقوطها على الظهر = ومجموع الاحتمالين = $\frac{١}{٢} + \frac{١}{٢} = ١$ وهو مقياس التأكيد .

على أنه من المهم جداً أن نلاحظ أن قولنا إن احتمال سقوط قطعة النقود على الوجه = $\frac{١}{٢}$ لا يعنى بالضرورة أنه لو رمينا القطعة مرتين أو حتى عشرة مرات فلا بد أن نحصل على الوجه فى إحداها . إذ أنه من الممكن تصور رمى قطعة النقود ٥٠ مرة دون أن يظهر الوجه إطلاقاً ولو أن احتمال حدوث ذلك ضئيل جداً وسنرى كيف نحسبه بدقة فيما بعد . أما المعنى الدقيق لقولنا إن احتمال سقوط قطعة النقود على الوجه يساوى ($\frac{١}{٢}$) فهو أنه إذا رمينا هذه القطعة عدداً كبيراً جداً من المرات فإننا نحصل على الوجه فى نصف هذه المرات تقريباً . وسنعود إلى هذه النقطه بتفصيل أكثر عند شرح المعنى الإحصائى للاحتمال .

كذلك إذا رمينا زهرة نرد كاملة التوازن فإن احتمالات سقوطها على أى وجه من الأوجة الستة متعادلة وذلك لأنه ليس هناك سبب يحملنا على الاعتقاد بأنها ستسقط على وجه دون الآخر . وعلى ذلك فمقياس التأكيد وهو الواحد الصحيح يكون موزعاً بالتساوى بين الاحتمالات الستة أى أن كل احتمال منها يساوى ($\frac{١}{٦}$) . فاحتمال ظهور « الجهار » عندما نرمى زهرة النرد = احتمال ظهور « الشيش » = احتمال ظهور أى وجه آخر = $\frac{١}{٦}$

كذلك لو أخذنا مجموعة جديدة كاملة من ورق اللعب وأحكمنا خلطها بحيث أصبح لكل ورقة نفس الفرصة فى الظهور عندما نسحب منها واحدة ، فإن احتمال ظهور ورقة معينة عند السحب = $\frac{١}{٥٢}$ ، كما أن احتمال ظهور أية صورة = $\frac{١٢}{٥٢} = \frac{٣}{١٣}$

على ضوء هذه المعلومات يمكننا أن نعرف الاحتمال الرياضى بالطريقة الآتية :—

إذا كان حدث معين يمكن أن يقع بعدد من الطرق قدره (ن) كل منها له نفس الفرصة فى الوقوع وبين هذه الطرق عدد قدره (ا) ميزاً عن الباقى بأية وسيلة كانت ونعتبر وقوعه بمثابة النجاح فإن احتمال النجاح = $\frac{ا}{ن}$

وحيث إن مقياس التأكيد = ١

فإن احتمال الفشل = $١ - \frac{ا}{ن} = \frac{ن - ا}{ن}$

فإذا رمزنا لاحتمال النجاح بالرمز ح واحتمال الفشل بالرمز ل فمن الواضح أن (ح + ل) = ١

فمثلا إذا رمينا زهرة النرد واعتبرنا ظهور « الشيش » أو « الجهار » بمثابة النجاح فإن احتمال النجاح = $\frac{ا}{ن} = \frac{٢}{٦} = \frac{١}{٣}$

واحتمال الفشل (عدم ظهور الشيش أو الجهار)

$$= \frac{ن - ا}{ن} = \frac{٦ - ٢}{٦} = \frac{٢}{٣}$$

ومن هذا يتضح أنه عند تقدير أى احتمال رياضى نبحث عن عدد طرق وقوع الحدث التى نعتبرها بمثابة نجاح ونقسمه على عدد جميع الطرق الممكنة لوقوع الحدث أو بعبارة أخرى فإن احتمال النجاح هو عبارة عن :-

$$\frac{\text{عدد طرق النجاح الممكنة}}{\text{عدد جميع الطرق الممكنة لوقوع الحدث}}$$

مثال - ما هو احتمال سحب أية ورقة من نوع « السباتى » من مجموعة كاملة جديدة ومحكمة الخلط من ورق اللعب ؟

هناك ١٣ ورقة من هذا النوع ، أى هناك ١٣ طريقة للسحب نعتبرها بمثابة نجاح بينما هناك ٥٢ طريقة للسحب عموماً .

∴ احتمال سحب ورقة من نوع « السباتى » .

$$= \frac{\text{عدد طرق النجاح الممكنة}}{\text{عدد جميع الطرق الممكنة}} = \frac{١٣}{٥٢} = \frac{١}{٤}$$

ونظرية الاحتمالات مع تعقيدها تنبنى على قاعدتين أساسيتين ، الأولى خاصة بجمع الاحتمالات والثانية خاصة بضرب الاحتمالات .

جمع الاحتمالات .

إذا كان احتمال وقوع حدث معين = $ح_١$ واحتمال وقوع حدث آخر = $ح_٢$ وكان حدوث أيهما مانعاً من حدوث الآخر(*) واعتبرنا وقوع أى الحدثين بمثابة نجاح فإن احتمال النجاح = $ح_١ + ح_٢$

وعلى وجه التعميم إذا كانت احتمالات وقوع (ن) من الأحداث هى على التوالى $ح_١$ ، $ح_٢$ ، $ح_٣$ .... $ح_ن$ فإن احتمال وقوع أيها .

$= ح_١ + ح_٢ + ح_٣ + ... ح_ن$

هذا على شرط أن يكون وقوع أيها مانعاً من حدوث كل الآخرين .

ويجدر بالطالب المبتدئ أن يتيقظ جيداً لاشتراط كون كل حدث مانعاً للآخر عند جمع الاحتمالات وإلا وقع فى أخطاء جسيمة . فإذا كان من الممكن لحدثين أن يقعا فى نفس الوقت فلا يجوز مطلقاً جمع احتماليهما .

(*) المقصود من قولنا إن « حدوث أحدهما مانع من حدوث الآخر » أنه لا يمكن أن يقع الحدثان فى نفس الوقت بل إذا وقع أحدهما امتنع الآخر . فإذا رمينا قطعة من النقود إلى أعلى واستقرت والوجه إلى أعلى فمن البديهى أن ظهور الوجه يعتبر مانعاً من ظهور الظهر .

ولإيضاح أهمية هذا الشرط إفرض أننا رمينا قطعة من النقود إلى أعلى مرتين

احتمال ظهور الوجه فى المرة الأولى = $\frac{١}{٢}$ كما بينا من قبل

احتمال ظهور الوجه فى المرة الثانية = $\frac{١}{٢}$ كذلك

لو كان من الجائز جمع هذين الاحتمالين لكان احتمال ظهور الوجه إما فى المرة الأولى أو فى المرة الثانية = $\frac{١}{٢} + \frac{١}{٢} = ١$ = مقياس التأكيد وهذا خطأ واضح ويمكن للقارئ أن يثبته فى الحال عن طريق رمى قطعة نقود مرتين إلى أعلى فسيتضح له من تكرار التجربة أنه من الجائز جداً أن يرمى قطعة النقود مرتين متتاليتين دون أن يحصل على الوجه فى أيهما . والسبب فى هذا الخطأ هو أن وقوع الحدث الأول (أى ظهور الوجه فى الرمية الأولى) لا يعتبر مانعاً من وقوع الحدث الثانى (أى ظهور الوجه فى الرمية الثانية) ولهذا السبب لا يجوز جمع الاحتمالين . وبالمثل إذا اشتريت عشرة تذاكر يانصيب كل منها خاصة بجمعية مختلفة وكان عدد التذاكر التى تصدرها كل جمعية عشرة فقط بينها واحدة تفوز بالجائزة فلا يمكنك أن تقول إن احتمال فوزك بإحدى الجوائز العشرة = $\frac{١}{١٠} + \frac{١}{١٠} + \frac{١}{١٠} + \frac{١}{١٠} + \cdots$ إلى عشر مرات = ١ = مقياس التأكيد . فمن الجائز جداً أن تشترى هذه التذاكر العشرة دون أن تفوز بأية جائزة . أما إذا اشتريت التذاكر العشرة الخاصة بجمعية واحداة ففى هذه الحالة فقط تستطيع أن تطمئن إلى فوزك بالجائزة . إذ فى هذه الحالة يكون وقوع أى حدث (فوز إحدى التذاكر) مانعاً لحدوث الآخرين وبالتالى يجوز جمع الاحتمالات .

ضرب الاحتمالات .

افرض أن حدثاً معيناً يمكن وقوعه بعدد من الطرق المختلفة قدره (ن) منها (ا) من الطرق نعتبر وقوعه بأية طريقة منها بمثابة نجاح وأن حدثاً آخر يمكن وقوعه بـ (ن$_١$) طريقة مختلفة منها عدد قدره (ا$_١$) نعده بمثابة نجاح . فما هو

احتمال النجاح المزدوج (أى نجاح الحدثين الأول والثانى معاً) مع العلم بأن الحدثين مستقلين تماماً بحيث لا يؤثر أحدهما فى الآخر بأية طريقة كانت ؟

للإجابة على هذا السؤال نستخدم القاعدة العامة التى أوردناها لتقدير الاحتمالات فيكون احتمال النجاح المزدوج .

$$= \frac{\text{عدد طرق النجاح المزدوج الممكنة}}{\text{عدد جميع الطرق الممكنة لوقوع الحدثين معاً}}$$

ومن الواضح أن عدد طرق النجاح المزدوج الممكنة = ا × ا$_{١}$ لأن أى واحدة من الـ (ا) طريقة لنجاح الحدث الأول يمكن أن تتفق مع أى واحدة من الـ (ا$_{١}$) طريقة لنجاح الحدث الثانى . وبالمثل فإن عدد جميع الطرق الممكنة لوقوع الحدثين معاً = ن × ن$_{١}$

$$\therefore \text{احتمال النجاح المزدوج} = \frac{\text{ا ا}_{١}}{\text{ن ن}_{١}} = \frac{\text{ا}}{\text{ن}} \times \frac{\text{ا}_{١}}{\text{ن}_{١}}$$

أى أن احتمال النجاح المزدوج لحدثين مستقلين عن بعضهما يساوى حاصل ضرب احتمال نجاح الحدث الأول فى احتمال نجاح الحدث الثانى(*) .

ولتوضيح قاعدة ضرب الاحتمالات افرض أننا رمينا زهرتى النرد فى نفس الوقت ، فما احتمال حصولنا على « الشيش » فى كلتيهما ؟

احتمال النجاح فى الزهرة الأولى = $\frac{١}{٦}$

« « « « = $\frac{١}{٦}$

$$\therefore \text{احتمال النجاح المزدوج} = \frac{١}{٦} \times \frac{١}{٦} = \frac{١}{٣٦}$$

(*) بديهى أننا لا نشترط فى ضرب الاحتمالات أن يكون أحدهما مانعاً للآخر كما فعلنا فى حالة جمع الاحتمالات فنحن هنا نبحث فى احتمال وقوع الحدثين معاً .

وكما أن احتمال النجاح المزدوج

$$= \frac{\text{عدد طرق النجاح المزدوج الممكنة}}{\text{عدد جميع الطرق الممكنة لوقوع الحدثين معاً}}$$

فإن احتمال الفشل المزدوج

$$= \frac{\text{عدد طرق الفشل المزدوج الممكنة}}{\text{عدد جميع الطرق الممكنة لوقوع الحدثين معا}} = \frac{(\text{ن}-\text{ا})\ (\text{ن}_{\text{ا}}-\text{ا}_{\text{ا}})}{\text{ن}\ \text{ن}_{\text{ا}}}$$

أما احتمال نجاح الحدث الأول وفشل الثانى $= \frac{\text{ا}}{\text{ن}} \times \frac{\text{ن}_{\text{ا}}-\text{ا}_{\text{ا}}}{\text{ن}_{\text{ا}}}$

واحتمال فشل الحدث الأول ونجاح الثانى $= \frac{\text{ن}-\text{ا}}{\text{ن}} \times \frac{\text{ا}_{\text{ا}}}{\text{ن}_{\text{ا}}}$

وبما أن هذه هى كل الاحتمالات الممكنة فلا بد أن يكون مجموعها يساوى مقياس التأكيد . وللتثبت من صحة ذلك نجمع هذه الاحتمالات مع العلم بأن وقوع أيها مانع من وقوع الآخرين :

$$\frac{\text{ا}_{\text{ا}}\ \text{ا}_{\text{ا}}}{\text{ن}\ \text{ن}_{\text{ا}}} + \frac{(\text{ن}-\text{ا})\ (\text{ن}_{\text{ا}}-\text{ا}_{\text{ا}})}{\text{ن}\ \text{ن}_{\text{ا}}} + \frac{\text{ا}\ (\text{ن}_{\text{ا}}-\text{ا}_{\text{ا}})}{\text{ن}\ \text{ن}_{\text{ا}}} + \frac{\text{ا}_{\text{ا}}\ (\text{ن}-\text{ا})}{\text{ن}\ \text{ن}_{\text{ا}}} = \text{ا}$$

وقياساً على ما تقدم يتضح أنه إذا كان هناك (ن) من الأحداث المستقلة واحتمالات نجاحها هى على التوالى $\text{ح}_{\text{ا}}$ ، $\text{ح}_{\text{٢}}$ ، $\text{ح}_{\text{٣}}$ ... $\text{ح}_{\text{ن}}$ فإن احتمال نجاحها جميعاً $= \text{ح}_{\text{ا}} \times \text{ح}_{\text{٢}} \times \text{ح}_{\text{٣}} \ldots \times \text{ح}_{\text{ن}}$

الآن وقد عرفنا كيف نقدر احتمال وقوع حدث معين ننتقل خطوة أخرى لتقدير احتمال حصولنا على عدد معين من مرات النجاح إذا أجرينا تجربة عدداً محدوداً من المرات .

عرفنا أن احتمال ظهور « الشيش » إذا رمينا زهرة النرد هو $\frac{\text{ا}}{\text{٦}}$ فما هو احتمال

حصولنا على « الشيش » ٣ مرات إذا رمينا الزهرة ٥ مرات ؟ بتطبيق قاعدة ضرب الاحتمالات نقول :

احتمال نجاح الرمية الأولى والثانية والثالثة معاً = $\frac{١}{٦} \times \frac{١}{٦} \times \frac{١}{٦} = (\frac{١}{٦})^{٣}$

غير أننا نرغب فى معرفة احتمال حصولنا على « الشيش » الذى نعتبره نجاحاً ثلاث مرات فقط وليس أكثر . ولكى يتحقق ذلك يجب ، إذا نجحت الرميات الأولى والثانية والثالثة معاً ، أن تفشل جميع الرميات التالية أى الرابعة والخامسة وحيث أن احتمال فشل أى رمية = $\frac{٥}{٦}$

∴ احتمال نجاح الرميات الأولى والثانية والثالثة وفشل الرابعة والخامسة = $(\frac{١}{٦})^{٣} (\frac{٥}{٦})^{٢}$.

والآن نتساءل عما إذا كانت هذه هى الإجابة على السؤال الذى وضعناه . أو بعبارة أخرى ، هل هذا العدد $(\frac{١}{٦})^{٣} (\frac{٥}{٦})^{٢}$ يمثل احتمال حصولنا على « الشيش » ٣ مرات إذا رمينا الزهرة ٥ مرات ؟ بالتأمل قليلا نجد أن هذا العدد لا يعطينا الإجابة الصحيحة . فنحن لم نتساءل عن احتمال نجاح الرميات الأولى والثانية والثالثة بالذات وفشل الرميتين الرابعة والخامسة بالذات ، وإنما تساءلنا عن احتمال نجاح أى ثلاث رميات وفشل أى رميتين . ومن البديهى أن احتمال نجاح أى ثلاث رميات مهما كان ترتيبها أكبر بكثير من احتمال نجاح ثلاث رميات محددة بالذات . فهذا الاحتمال الأخير ما هو إلا حالة خاصة من الاحتمال الأول . ولذلك فللحصول على احتمال نجاح أى ثلاث رميات وفشل أى رميتين مهما كان ترتيبها يجب أن نبحث عن جميع التوافيق الممكنة لنجاح ثلاث رميات وفشل رميتين ثم نجمع الاحتمالات الخاصة بكل هذه التوافيق . وللحصول على عدد الطرق الممكن بها اختيار ثلاث رميات من خمسة (لغرض اعتبارها ناجحة) نطبق قانون التوافيق المعروف $ق_{م}^{ن}$

فيكون عدد التوافيق الممكنة فى هذه الحالة $= ق_{٣}^{٥} = ١٠$

وحيث إن احتمال حصول كل من هذه التوافيق العشرة = $(\frac{1}{6})^3 (\frac{5}{6})^2$

∴ احتمال حصول أى واحد منها أو بعبارة أخرى احتمال نجاح أى ثلاث رميات (وفشل الرميتين الباقيتين) = $(\frac{1}{6})^3(\frac{5}{6})^2 + (\frac{1}{6})^3(\frac{5}{6})^2 + \cdots +$ إلى $^5ق_3$ مرة أى إلى ١٠ مرات = $^5ق_3 (\frac{1}{6})^3 (\frac{5}{6})^2$

وقياساً على النتيجة السابقة نقول إن احتمال الحصول على « الشيش » $ر$ مرة إذا رمينا زهرة النرد (ن) مرة

$= {}^{ن}ق_{ر} (\frac{1}{6})^{ر} (\frac{5}{6})^{ن-ر}$

وعلى وجه التعميم إذا كان احتمال نجاح حدث معين هو ح واحتمال فشله هو ل = (١ — ح) فإن احتمال النجاح $ر$ مرة فى عدد من المحاولات قدره (ن)

$${}^{ن}ق_{ر} \, ح^{ر} \, ل^{ن-ر}$$

وهذه النتيجة الهامة معروفة بقانون Bernoulli نسبة إلى واضعها James Bernoulli

المعنى الإحصائى للاحتمال .

بمراجعة تعريف الاحتمال الرياضى الوارد فى (ص ٢٣٩) نجد أنه يعطينا الاحتمال على شكل نسبة تحدد درجة ثقتنا فى وقوع حدث معين . فنقول مثلا إن احتمال ظهور « الشيش » عند رمى زهرة النرد هو $(\frac{1}{6})$ واحتمال ظهور الوجه عند رمى قطعة النقود هو $(\frac{1}{2})$. غير أنه بالتأمل فى هذا المعنى نجد أنه لا يحتمل الاختبار والتحليل الدقيقين . فإذا أردنا أن نعطى معنى دقيقاً لهذه النسبة لوجدنا أنفسنا فى حيرة شديدة . ذلك أن الحدث إما أن يقع أو لا يقع ولا يمكن أن يقع نصفه أو سدسه . فإذا رمينا قطعة النقود واعتبرنا ظهور الوجه بمثابة نجاح فإما أن يظهر الوجه أولا يظهر . وعلى ذلك فبعد انتهاء التجربة نجد

أن الواقع يخالف الاحتمال وسواء أكنا قد قدرنا هذا الاحتمال بـ (١/٦) أو (١/٣) أو (١/٢) قبل إجراء التجربة فالنتيجة واحدة بعد إجراء التجربة . . . . . إما أن يقع الحدث أو لا يقع بغض النظر عن التقدير الذى كنا قد أعطيناه للاحتمال من قبل .

أما المعنى الإحصائى للاحتمال فإنه يخلصنا من هذه المشكلة . فالاحتمال بالمعنى الإحصائى هو عبارة عن تكرار نسبى ، فإذا كان التقدير الرياضى لوقوع حدث معين هو (١/٢) فالمعنى الإحصائى لذلك هو أنه كررنا التجربة عدداً كبيراً جداً من المرات (يقرب من ∞) فإن الحدث يقع فى نصف هذه المرات ولا يقع فى النصف الآخر . واستناداً إلى هذا الاتجاه فى تفسير معنى الاحتمال نستطيع أن نعطى معنى دقيقاً محدداً للنسبة التى تبين قيمة الاحتمال . فبدلا من أن نقول إنها تمثل درجة ثقتنا فى وقوع الحدث كما هو الحال فى المعنى الرياضى نقول إنها تمثل التكرار النسبى لوقوع الحدث . فإذا كانت هذه النسبة تساوى (١/٢) فمعناها الإحصائى هو ٥٠٪ أو ٥٠٠ فى الألف من مجموع التكرارات . وإذا رمينا قطعة من النقود ١٠٠٠٠ مرة فإن النتيجة التى نحصل عليها لا يمكن أن تختلف كثيراً (*) عن ٥٠٠٠ وجه ، ٥٠٠٠ ظهر إلا إذا كانت القطعة ليست كاملة التوازن وكأن الواقع لا يخالف الاحتمال بالمعنى الإحصائى . وعلى ذلك فما من شك فى أن المعنى الإحصائى للاحتمال أقرب إلى منطق الأشياء من المعنى الرياضى ولهذا نجد أن الاتجاه بين الرياضيين أنفسهم هو نحو الأخذ بهذا التفسير الإحصائى للاحتمال .

(*) يمكن أن تؤثر عوامل المصادفة فى أية تجربة من التجارب التى نجريها فتجعل العدد الذى نحصل عليه فعلا يختلف قليلا عن ٥٠٠٠ وجه ، ٥٠٠٠ ظهر . غير أن هناك حدوداً لا يمكن أن يتجاوزها الفرق بين العدد الذى نتوقعه نظرياً والعدد الذى نحصل عليه فعلا بالتجربة كما سيتضح فيما بعد .

على ضوء هذا المعنى الإحصائى للاحتمال نعود إلى قانون Bernoulli فنعطى ر قيماً مختلفة كى نحصل على احتمال حدوث أى عدد من مرات النجاح (من .إلى ن) إذا أجرينا تجربة معينة (ن) مرة . فمثلا إذا رمينا ٤ قطع نقود إلى أعلى واعتبرنا ظهور الوجه بمثابة النجاح ، فبإعطاء ر (فى قانون برنوللى) قيماً مختلفة من (٠ إلى ٤) نحصل على التوزيع التكرارى الآتى لعدد مرات النجاح .

عدد مرات النجاح (ر)	احتمال النجاح	(التكرار النسبى)
٠	$ن ق_{٠} (\frac{١}{٢})^{٠} (\frac{١}{٢})^{٤} = ن ق_{٠} (\frac{١}{٢})^{٤}$	
١	$ن ق_{١} (\frac{١}{٢})^{١} (\frac{١}{٢})^{٣} = ن ق_{١} (\frac{١}{٢})^{٤}$	
٢	$ن ق_{٢} (\frac{١}{٢})^{٢} (\frac{١}{٢})^{٢} = ن ق_{٢} (\frac{١}{٢})^{٤}$	
٣	$ن ق_{٣} (\frac{١}{٢})^{٣} (\frac{١}{٢})^{١} = ن ق_{٣} (\frac{١}{٢})^{٤}$	
٤	$ن ق_{٤} (\frac{١}{٢})^{٤} (\frac{١}{٢})^{٠} = ن ق_{٤} (\frac{١}{٢})^{٤}$	

بما أن هذه هى جميع الاحتمالات الممكنة فلا بد أن يكون مجموعها يساوى مقياس التأكيد كما يتضح مما يأتى :

$$\text{مجموع الاحتمالات} = (\frac{١}{٢})^{٤} [ن ق_{٠} + ن ق_{١} + ن ق_{٢} + ن ق_{٣} + ن ق_{٤}]$$

$$= (\frac{١}{٢})^{٤} (١ + ١)^{٤} = ١ = \text{مقياس التأكيد}$$

ومما تجب ملاحظته أن قانون برنوللى لا يقتصر على الحالة التى يكون احتمال النجاح فيها معادلا لاحتمال الفشل كما فى حالة رمى قطع النقود . وقد تعمدنا استنتاجه من حالة زهرة النرد حيث احتمال النجاح مخالف لاحتمال الفشل دفعاً لما قد يتطرق إلى الذهن من انطباقه على حالة تعادل احتمال النجاح مع احتمال الفشل فقط . كما أنه من الواضح أن تطبيق هذا القانون غير مقيد باتخاذ

(ن) أية قيمة معينة . وبناء على ذلك إذا فرضنا أن احتمال نجاح أى حدث هو (ح) واحتمال فشله هو (ل) وأن عدد المحاولات هو (ن) نستطيع بتطبيق قانون برنولى أن نكتب الحالة العامة للتوزيع التكرارى لذات الحدين كالآتى —

الحالة العامة للتوزيع التكرارى لذات الحدين

عدد مرات النجاح (س)	التكرار النسبى (و)
٠	$ن_{ق_٠} ح^{٠} ل^{ن} = ل^{ن}$
١	$ن_{ق_١} ح^{١} ل^{ن-١} = ن ح ل^{ن-١}$
٢	$ن_{ق_٢} ح^{٢} ل^{ن-٢} = \frac{ن(ن-١)}{!٢} ح^{٢} ل^{ن-٢}$
٣	$ن_{ق_٣} ح^{٣} ل^{ن-٣} = \frac{ن(ن-١)(ن-٢)}{!٣} ح^{٣} ل^{ن-٣}$
⋮	⋮
ن	$ن_{ق_ن} ح^{ن} ل^{٠} = ح^{ن}$

ويلاحظ هنا أيضاً أن مجموع التكرارات النسبية (أى الاحتمالات) يساوى الوحدة أى يساوى مقياس التأكيد . إذ أن هذه التكرارات هى على التوالى حدود مفكوك ذات الحدين $(ح + ل)^{ن}$. وحيث أن $ل = ١ - ح$

∴ مجموع التكرارات $= (ح + ١ - ح)^{ن} = ١$

وبمقارنة التوزيع التكرارى لأحجام جميع العينات الممكنة(٥) بهذا التوزيع نجد تشابهاً كبيراً بينهما . إذ الواقع أن التوزيع الأول ما هو إلا حالة خاصة من هذا التوزيع الأخير فإذا جعلنا قيمة (ح) معادلة لقيمة (ل) فى هذا التوزيع وضربنا كل تكرار فى $\left(\frac{١}{ح}\right)^{ن}$ لحصلنا على التوزيع الأول ويلاحظ أن هذا التوزيع لا يكون متماثلا إلا إذا كانت قيمة (ح) معادلة لقيمة (ل) غير أنه يقترب من التماثل كلما زادت قيمة (ن) كما سيتضح فيما بعد .

الوسط الحسابى والانحراف المعيارى للتوزيع التكرارى لذات الحدين .

بنفس الطريقة التى حصلنا بها على الوسط الحسابى والانحراف المعيارى للتوزيع التكرارى لأحجام العينات يمكن أن نحصل على الوسط الحسابى والانحراف المعيارى للتوزيع التكرارى لذات الحدين . فللحصول على الوسط الحسابى نحسب Σ س$_{ن}$ و مع ملاحظة أن Σ و = ١

$$\Sigma \text{س}_{ن} \text{و} = (٠)\,ل^{ن} + (١)\,ن\,ح\,ل^{ن-١} + (٢)\,\frac{ن(ن-١)}{!٢}\,ح^{٢}\,ل^{ن-٢}$$

$$+ (٣)\,\frac{ن(ن-١)(ن-٢)}{!٣}\,ح^{٣}\,ل^{ن-٣} + \cdots\cdots + (ن)\,ح^{ن}$$

وللحصول على مجموع حدود الطرف الأيسر نتبع نفس الطريقة التى

---

(٥) راجع صفحة ٢٢٩ من هذا الكتاب .

اتبعناها فى حالة التوزيع التكرارى لأحجام العينات فنقول(*):

$$(ل + ح)^{ن} = ل^{ن} + ن ح ل^{ن-١} + \frac{ن(ن-١)}{!٢} ح^{٢} ل^{ن-٢}$$

$$+ \frac{ن(ن-١)(ن-٢)}{!٣} ح^{٣} ل^{ن-٣} + \cdots + ح^{ن}$$

بمفاضلة الطرفين بالنسبة لـ ح

$$\therefore ن(ل + ح)^{ن-١} = ٠ + ن ل^{ن-١} + \frac{ن(ن-١)}{!٢} ٢ ح ل^{ن-٢}$$

$$+ \frac{ن(ن-١)(ن-٢)}{!٣} ٣ ح^{٢} ل^{ن-٣} + \cdots + ن ح^{ن-١}$$

بضرب الطرفين فى ح

$$\therefore ن ح (ل + ح)^{ن-١} = ن ح ل^{ن-١} + (٢) \frac{ن(ن-١)}{!٢} ح^{٢} ل^{ن-٢}$$

$$+ (٣) \frac{ن(ن-١)(ن-٢)}{!٣} ح^{٣} ل^{ن-٣} + \cdots + (ن) ح^{ن} \cdots (١)$$

(*) الفضل فى إيجاد قيمة الوسط الحسابى والانحراف المعيارى للتوزيع التكرارى لأحجام العينات باستخدام التفاضل يرجع إلى الأستاذ Hyman Levy أستاذ الرياضة بالكلية الإمبراطورية للعلوم بجامعة لندن . وقد وجدنا أنه من السهل تعميم هذه الطريقة لتشمل الحالة العامة للتوزيع التكرارى لذات الحدين . أما الطريقة المتبعة عادة فهى إيجاد قيم هذه الثوابت بالجبر . وهى طريقة معقدة خصوصاً فى حالة الانحراف المعيارى ويمكن للطالب الغير الملم بطرق التفاضل الرجوع إليها فى كتاب « مقدمة لنظرية الإحصاء » تأليف يول وكندال ص ١٧٣ .

وبالنظر إلى حدود الطرف الأيسر نجد أنها هى بالضبط Σ س^{ن} و

$\therefore$ Σ س^{ن} و = ن ح (لاحظ أن (ح + ل) $^{ن-١}$ = ١)

وحيث أن مجموع التكرارات = ١

$\therefore$ الوسط الحسابى للتوزيع التكرارى لذات الحدين = ن ح مهما كانت قيمة ح . وبوضع ح = ل = $\frac{١}{٢}$ نجد أن الوسط الحسابى للتوزيع التكرارى = $\frac{ن}{٢}$ وهى نفس النتيجة التى حصلنا عليها فى الحالة الخاصة لهذا التوزيع .

ولإيجاد قيمة الإنحراف المعيارى نبدأ من المعادلة (١) حيث وجدنا أن :

$$(ل + ح)^{ن} = ن ح\, ل^{ن-١} + (٢)\frac{ن(ن-١)}{!٢} ح^{٢} ل^{ن-٢}$$

$$+ (٣)\frac{ن(ن-١)(ن-٢)}{!٣} ح^{٣} ل^{ن-٣} + ن(ح)^{ن}$$

فاضل الطرفين بالنسبة ل (ح)

$$\therefore ن ح (ن-١)(ل + ح)^{ن-٢} + (ل + ح)^{ن-١} ن = ن ل^{ن-١}$$

$$+ (٢)^{٢}\frac{ن(ن-١)}{!٢} ح ل^{ن-٢} + (٣)^{٢}\frac{ن(ن-١)(ن-٢)}{!٣} ح^{٢} ل^{ن-٣}$$

$$+ \cdots\cdots + (ن)^{٢} ح^{ن-١}$$

بالضرب فى ح مع العلم بأن كلا من $(ل + ح)^{ن-٢}$ و $(ل+ح)^{ن-١}$=١

$$ح[ن ح (ن - ١) + ن] = ن ح ل^{ن-١} + (٢)^{٢} \frac{ن(ن-١)}{!٢} ح^{٢} ل^{ن-٢}$$

$$+ (٣)^{٢} \frac{ن(ن-١)(ن-٢)}{!٣} ح^{٣} ل^{-٣} + ... + (ن)^{٢} ح^{ن}$$

وبالنظر إلى حدود الطرف الأيسر نجد أنها عبارة عن Σ س^{٢}و أى مجموع حواصل ضرب مربعات القيم فى التكرارات المقابلة لها .

أى أن Σ س^{٢}و = ح [ن ح (ن - ١) + ن] = ن^{٢} ح^{٢} - ن ح^{٢} + ن ح

وللحصول على مربع الانحراف المعيارى نطرح من هذه الكمية مربع الوسط الحسابى فيكون

$$\sigma^{٢} = ن^{٢} ح^{٢} - ن ح^{٢} + ن ح - ن^{٢} ح^{٢}$$

$$= ن ح - ن ح^{٢} = ن ح (١ - ح) = ن ح ل$$

$$\sigma = \sqrt{ن ح ل}$$

وبوضع ح = ل = $\frac{١}{٢}$ تكون قيمة $\sigma = \sqrt{\frac{ن}{٤}}$

وهى نفس الطريقة التى حصلنا عليها فى الحالة الخاصة .

معنى الوسط الحسابى والانحراف المعيارى للتوزيع التكرارى لذات الحدين وتطبيقاتهما .

لتوضيح معنى الوسط الحسابى (ن ح) افرض أننا رمينا قطعة نقود كاملة التوازن ٩٠٠ مرة . النتيجة التى نتوقعها نظرياً هى أن نحصل على الوجه فى نصف عدد المرات تقريباً حيث أنه ليس هناك أى سبب معروف من شأنه أن

يجعل عدد مرات ظهور الوجه تزيد أو تقل عن عدد مرات ظهور الظهر . والوسط الحسابى (ن ح) هو الذى يعطينا تلك القيمة التى نتوقعها نظرياً . وعلى ذلك يكون عدد مرات ظهور الوجه الذى نتوقعه نظرياً فى هذه التجربة هو ٩٠٠ × $\frac{١}{٢}$ = ٤٥٠

غير أنه إذا لم يكن هناك سبب معروف يجعل التكرار يميل نحو الوجه أو نحو الظهر فى كل مره نجرى فيها التجربة ، إلا أن هناك تلك العوامل المجهولة التى نطلق عليها اسم عوامل المصادفة والتى من شأنها أن تجعل النتائج التى نحصل عليها فعلا مخالفة قليلا لتلك التى نتوقعها نظرياً . فقد نرمى قطعة النقود ٩٠٠ مرة فنحصل على ٤٥٥ وجه وإذا كررنا التجربة فقد نحصل على ٤٤٥ وجه وإذا كررنا التجربة مرات أخرى فقد نحصل على ٤٥٢ ، ٤٤٧ ، ٤٤٦ ، ٤٥٣ ، ٤٥١ ، ٤٤٨ أو على أعداد مقاربة لهذه . ولو فرض أن كررنا هذه التجربة عدداً كبيراً جداً من المرات وحسبنا متوسط النتائج التى نحصل عليها فى كل مرة فإن هذا المتوسط يعادل (ن ح) أى ٤٥٠ أو على الأقل فإن هذا هو المفروض نظرياً إذا ما كررنا التجربة عدداً لا نهاية له من المرات .

غير أن عوامل المصادفة لا يمكن أن تعتبر مسئولة عن كل خطأ فى التجربة أو كل فرق بين القيمة التى نحصل عليها فعلا وتلك التى نتوقعها نظرياً . فقد تكون قطعة النقود ليست كاملة التوازن الأمر الذى من شأنه أن يجعل النتائج متحيزة إما نحو الوجه أو نحو الظهر . وقد يتسرب عنصر التحيز أيضاً إلى طريقة رمى قطعة النقود إلى أعلى وهكذا . ومن الواضح أنه إذا لم تكن ثمة وسيلة للتمييز بين الأخطاء (أو الفروق) التى يمكن أن تعزى لعوامل المصادفة وتلك التى ترجع إلى عامل التحيز أو أىّ عامل إرادى آخر يؤثر فى نتيجة التجربة ، فإن النتائج التى نحصل عليها تكون موضعاً للشك بل أن قيمتها تكاد تكون معدومة . فإذا كنا سنلقى مسئولية أى خطأ يحدث فى التجارب التى نجريها

على عوامل المصادفة فأى ثمرة نجنيها إذن من إجراء هذه التجارب ؟ على أنه من السهل جداً تمييز أخطاء المصادفة عن الأخطاء الراجعة إلى عوامل أخرى باستخدام بعض الطرق الإحصائية . وقيمة الانحراف المعيارى للتوزيع التكرارى لذات الحدين هى التى تعين لنا حدود الأخطاء (أو الفروق) التى يمكن أن تعزى إلى عوامل المصادفة . فإذا كان الفرق بين النتيجة التى نحصل عليها فعلا بالتجربة وتلك التى نتوقعها نظرياً يقل عن قيمة الانحراف المعيارى ($\sqrt{\text{ن ح ل}}$) فمن المحتمل أن يكون هذا الفرق راجعاً إلى عوامل المصادفة . وإذا كان الفرق يعادل أو يزيد عن ضعف الانحراف المعيارى فإن احتمال حدوثه بطريق المصادفة يكون ضئيلا . أما إذا كان هذا الفرق يزيد عن ثلاثة أمثال الانحراف المعيارى فيكاد يكون من المستحيل عمليا أن يكون راجعاً إلى عوامل المصادفة . وذلك لأن احتمال حدوث فرق مثل هذا بطريق المصادفة من الضآلة بحيث يمكن إهماله من الوجهة العملية .

وعلى وجه التحديد فإن احتمال الحصول على قيمة تختلف عن الوسط الحسابى (ن ح) بمقدار $\pm \sqrt{\text{ن ح ل}}$ أى $\pm \sigma$ = ٣١,٧٪ تقريباً واحتمال الحصول على قيمة تختلف عن (ن ح) بمقدار $\pm ٢ \sqrt{\text{ن ح ل}}$ = ٤,٥٥٪ تقريباً واحتمال الحصول على قيمة تختلف عن (ن ح) بمقدار

$\pm ٣ \sqrt{\text{ن ح ل}}$ = ٠,٢٧٪ تقريباً .

ومن الواضح أن هذا الاحتمال الأخير من الضآلة بحيث يمكن إهماله .

وسيتضح لنا سبب تعيين هذه الحدود بالذات لأخطاء المصادفة من دراسة منحنى الخطأ المتماثل أو المنحنى الطبيعى وهو الوضع النهائى الذى يتخذه التوزيع التكرارى لذات الحدين عندما تكبر قيمة ن حتى تقترب من ∞ .

ولنقبل هذه النتائج مؤقتاً وننتقل إلى تطبيقها العملى مرجئين شرح الأساس

الذى تستند إليه إلى حيث نتكلم عن خواص المنحنى الطبيعى .

ولتوضيح التطبيق العملى لكل من الوسط الحسابى والانحراف المعيارى لذات الحدين افرض أننا رمينا قطعة نقود كاملة التوازن ٩٠٠ مرة واعتبرنا ظهور الوجه بمثابة نجاح .

عدد مرات النجاح التى نتوقعها نظرياً = ن ح = ٩٠٠ × $\frac{١}{٢}$ = ٤٥٠

الانحراف المعيارى = $\sqrt{\text{ن ح ل}}$ = $\sqrt{٩٠٠ \times \frac{١}{٢} \times \frac{١}{٢}}$ = ١٥

إذا حصلنا على عدد من مرات النجاح ينحصر بين (٤٥٠ + ١٥) ، (٤٥٠ — ١٥) أى بين ٤٦٥ ، ٤٣٥ فيمكن اعتبار الفرق بين العدد الذى حصلنا عليه فعلا وذلك الذى نتوقعه نظرياً راجعاً إلى عوامل المصادفة حيث أن احتمال حدوث مثل هذا الفرق بطريق المصادفة = ٣١,٧ ٪

أما إذا حصلنا على عدد من مرات النجاح يختلف عن ذلك الذى نتوقعه نظرياً بمقدار ٢ $\sqrt{\text{ن ح ل}}$ (بالزيادة أو النقص) أى إذا حصلنا على عدد من مرات النجاح ينحصر بين ٤٨٠ ، ٤٢٠ فإن احتمال حدوث مثل هذا الاختلاف بطريق المصادفة يكون ضئيلا فهو يعادل ٤,٥٥ ٪

وإذا حصلنا على عدد من مرات النجاح يختلف عن ذلك الذى نتوقعه نظرياً بمقدار ٣ $\sqrt{\text{ن ح ل}}$ (بالزيادة أو النقص) أى إذا حصلنا على عدد من مرات النجاح ينحصر بين ٤٩٥ ، ٤٠٥ فلا يمكن أن نعزوا مثل هذا الاختلاف إلى عوامل المصادفة حيث أن احتمال حدوثه بهذه الكيفية يعادل ٠,٢٧٪ وهو احتمال ضئيل يمكن إهماله ، فيكاد يكون من المؤكد أن تكون عوامل أخرى غير المصادفة هى التى تسببت فى حدوث مثل هذا الاختلاف .

والخلاصة أنه إذا كان الفرق بين العدد الذى نتوقعه نظرياً وذلك الذى نحصل عليه بالتجربة لا يتجاوز $\pm$ ٢ σ فإنه يعتبر فرقاً ظاهرياً .

(Insignificant Difference) يمكن أن يعزى إلى عوامل المصادفة .

أما الفروق التى تتجاوز $\pm 2\sigma$ (أو $\pm 3\sigma$ فى نظر بعض الإحصائيين) فإنها تعتبر فروقاً جوهرية (Significant Differences) لا يمكن أن تعزى إلى عوامل المصادفة (*) .

أمثلة على استعمال الوسط الحسابى والانحراف المعيارى لذات الحدين :

(١) أنتجت إحدى مزارع البرتقال ١٠,٠٠٠ ثمرة وجد أن بينها ٥٧٠ من الثمار العطبة بينما كانت النسبة المعتادة للعطب هى ١٠٪ من مجموع الثمار . هل يدل الفرق بين عدد الثمار العطبة الذى نتوقعه نظرياً وذلك الذى حصلنا عليه فعلا على حدوث تغير جوهرى فى ظروف الزراعة ، أم يمكن أن نعزوه إلى عوامل المصادفة ؟

الحل : احتمال الحصول على ثمرة عطبة = ١٠٪

$$\therefore ح = \frac{١}{١٠} \quad ول = \frac{٩}{١٠}$$

عدد الثمار العطبة الذى نتوقعه نظرياً = ن ح = ١٠٠٠٠ × $\frac{١}{١٠}$ = ١٠٠٠

الفرق بين العدد الذى نتوقعه نظرياً وذلك الذى حصلنا عليه فعلا

= ١٠٠٠ − ٥٧٠ = ٤٣٠

الانحراف المعيارى = $\sqrt{ن ح ل}$ = $\sqrt{١٠٠٠٠ \times \frac{١}{١٠} \times \frac{٩}{١٠}}$ = $\sqrt{٩٠٠}$ = ٣٠

الفرق بدلالة وحدات الانحراف المعيارى = $\frac{٤٣٠}{٣٠}$ = ١٤,٣٣ σ

(*) يتضح مما تقدم أن هناك خلاف بين الإحصائيين فيما يتعلق بالتمييز بين الفروق الظاهرية والفروق الجوهرية . فبعضهم يعتبر الفرق الذى يعادل ٢ $\sqrt{ن ح ل}$ فرقاً جوهرياً لا يمكن أن يعزى إلى عوامل المصادفة ، والبعض الآخر يشترط لاعتبار الفرق جوهرياً أن يكون مساوياً ٣ $\sqrt{ن ح ل}$ على الأقل .

وحيث أن الفرق $< ٣\sigma$

∴ الفرق جوهرى ولا يمكن أن نعزوه إلى عوامل المصادفة .

أى أن هذا الفرق يدل على تحسن فعلى فى ظروف الزراعة .

(٢) أنبت أحد أنواع البسلة ١٤٣٠٠ من البذور الصفراء و ٤٩٠٠ من البذور الخضراء فى حين أن عدد البذور الصفراء الذى نتوقعه حسب قانون مندل (Mendel) يعادل ٧٥٪ من مجموع البذور . هل يمكن أن نعزو الفرق بين العدد الذى حصلنا عليه فعلا وذلك الذى نتوقعه نظرياً إلى عوامل المصادفة ؟

الحل : احتمال الحصول على بذرة صفراء = ٧٥ ٪

∴ ح = $\frac{٣}{٤}$ و ل = $\frac{١}{٤}$

عدد البذور الصفراء الذى نتوقعه نظرياً = ن ح = ١٩٢٠٠ × $\frac{٣}{٤}$ = ١٤٤٠٠

∴ الفرق بين العدد الذى نتوقعه والعدد الذى حصلنا عليه .

= ١٤٤٠٠ − ١٤٣٠٠ = ١٠٠

$\sigma = \sqrt{ن ح ل} = \sqrt{١٩٢٠٠ \times \frac{٣}{٤} \times \frac{١}{٤}} = \sqrt{٣٦٠٠} = ٦٠$

الفرق بدلالة وحدات الانحراف المعيارى (σ) = $\frac{١٠٠}{٦٠}$ = ١,٦٦ σ

وحيث أن الفرق $< ٣\sigma$

∴ الفرق ظاهرى ويمكن أن نعزوه إلى عوامل المصادفة .

أى أن مثل هذه النتيجة لا تعتبر مخالفة لقانون مندل .

الوسط الحسابى والانحراف المعيارى لنسبة النجاح :

رأينا أن الوسط الحسابى لعدد مرات النجاح يساوى ن ح وأن الانحراف المعيارى لهذا العدد = $\sqrt{ن ح ل}$. غير أننا نجد من الأفضل أحياناً أن نعبر

عن مرات النجاح فى شكل نسبة مئوية بدلا من التعبير عنها فى شكل أعداد مطلقة . وفى هذه الحالة فإن الوسط الحسابى لنسبة النجاح = $\frac{\text{ن ح}}{\text{ن}} = \text{ح}$

كما أن الانحراف المعيارى للنسبة = $\frac{\sqrt{\text{ن ح ل}}}{\text{ن}} = \sqrt{\frac{\text{ح ل}}{\text{ن}}}$

ويجب أن نتنبه جيداً عند استعمال هاتين النتيجتين فى المسائل العملية إلا أن كلا من ح ، $\sqrt{\frac{\text{ح ل}}{\text{ن}}}$ لا يشير إلى نسب مئوية وإنما إلى نسب منسوبة للواحد الصحيح . ونظراً لأن معظم المسائل التى نصادفها تعطى النسب المتوسطة للنجاح فى شكل نسب مئوية فيجب رد هذه النسب إلى أساس الواحد الصحيح قبل استخدام النتيجتين السابقتين . ولتسهيل المقارنة بين كيفية استخدام الوسط الحسابى والانحراف المعيارى لذات الحدين فى حالة الأعداد المطلقة وحالة النسب ، نأخذ نفس الأمثلة التى عالجناها بالطريقة الأولى وذلك لحلها بهذه الطريقة الجديدة .

حل المثال الأول :

النسبة المتوسطة للنجاح أو نسبة الثمار العطبة التى نتوقعها نظرياً

= ح = ٠,١ ، ل = ٠,٩

نسبة الثمار العطبة التى حصلنا عليها فعلا = $\frac{٥٧٠}{١٠٠٠}$ = ٠,٠٥٧

الفرق بين النسبتين = ٠,١ — ٠,٠٥٧ = ٠,٠٤٣

الانحراف المعيارى للنسبة المتوسطة = $\sqrt{\frac{\text{ح ل}}{\text{ن}}} = \sqrt{\frac{٠,٩ \times ٠,١}{١٠٠٠٠}}$

= $\sqrt{٠,٠٠٠٠٠٩}$ = ٠,٠٠٣

الفرق بدلالة وحدات الانحراف المعيارى = $\frac{٠٫٤٣}{٠٫٠٠٣}$ = ١٤,٣٣ σ

وحيث أن الفرق > ٣ σ

∴ الفرق جوهرى ولا يمكن أن يعزى إلى عوامل المصادفة

وهى نفس النتيجة التى حصلنا عليها بالطريقة السابقة .

حل المثال الثانى :

النسبة المتوسطة للنجاح أو نسبة البذور الصفراء التى نتوقعها نظرياً

= ٧٥٪ = $\frac{٣}{٤}$ ، ل = $\frac{١}{٤}$

نسبة البذور الصفراء التى حصلنا عليها فعلا = $\frac{١٤٣٠٠}{١٩٢٠٠}$ = $\frac{١٤٣}{١٩٢}$

الفرق بين النسبتين = $\frac{٣}{٤}$ − $\frac{١٤٣}{١٩٢}$ = $\frac{١}{١٩٢}$

الانحراف المعيارى = $\sqrt{\frac{ح ل}{ن}}$ = $\sqrt{\frac{\frac{١}{٤} \times \frac{٣}{٤}}{١٩٢٠٠}}$ = $\frac{١}{٣٢٠}$

الفرق بدلالة وحدات الانحراف المعيارى = $\frac{١}{١٩٢} \div \frac{١}{٣٢٠}$ = $\frac{٣٢٠}{١٩٢}$ = ١,٦٦ σ

وحيث أن الفرق أقل من ٣ σ ∴ الفرق ظاهرى

وهى نفس النتيجة التى حصلنا عليها بالطريقة السابقة .

العزوم العليا للتوزيع التكرارى لذات الحدين :

يمكن بطريقة مماثلة لتلك التى اتبعناها فى إيجاد قيمة الوسط الحسابى والانحراف المعيارى أن نعين قيمتى العزم الثالث والرابع للتوزيع التكرارى لذات الحدين .

والعزم الثالث = ن ح ل (ل − ح)

فإذا كانت ح = ل كما هو الحال فى التوزيع التكرارى المتماثل لذات الحدين فمن الواضح أن ن ح ل (ل − ح) = صفراً .

والعزم الرابع = ٣ ح^{٢} ل^{٢} ن^{٢} + ح ل ن (١ − ٦ ح ل) .

ومن هذا يتضح أن معامل الالتواء ل$_{١}$ = $\frac{(م_{٣})^{٢}}{(م_{٢})^{٣}}$ = $\frac{(ل-ح)^{٢}}{ن\ ح\ ل}$

ومعامل التفرطح ل$_{٢}$ = $\frac{م_{٤}}{(م_{٢})^{٢}}$ = ٣ + $\frac{١ - ٦\ ح\ ل}{ن\ ح\ ل}$

ويلاحظ أن قيمة ل$_{١}$ تقل كلما زادت قيمة ن أى أن درجة الالتواء تقل كلما زادت قيمة ن . والوضع النهائى للتوزيع التكرارى لذات الحدين عندما تقترب ن من ∞ هو منحن متماثل تماماً . إذ من الواضح أن قيمة ل$_{١}$ فى هذه الحالة تساوى صفراً .

أما قيمة ل$_{٢}$ فإنها تقترب من ٣ كلما زادت قيمة ن وتعادل ٣ تماماً إذا كانت ن = ∞ . أى أن درجة تفرطح التوزيع التكرارى لذات الحدين تكون فى هذه الحالة معادلة تماماً لدرجة تفرطح المتحنى الطبيعى . والواقع أن الوضع النهائى للتوزيع التكرارى لذات الحدين عندما تقترب ن من ∞ هو المنحنى الطبيعى ذاته كما سيتضح فيما يلى . ويستطيع القارئ أن يتبين بنفسه اتجاه التوزيع التكرارى لذات الحدين نحو التماثل من جهة والاعتدال فى التفرطح من الجهة الأخرى عند ما تزيد قيمة ن بإعطاء ن قيماً كبيرة ثم إيجاد قيمة ل$_{١}$ و ل$_{٢}$.

فلو فرض أن ح = ٠,٠٢ ، ل = ٠,٩٨ ، ن = ١٠٠٠

فإن ل$_{١}$ = ٠,٠٤٧٠

و ل$_{٢}$ = ٣,٠٤٥٠

أى أنه على الرغم من الاختلاف الكبير بين قيمتى ل ، ح فإن التوزيع التكرارى لذات الحدين يكون قريباً جداً من التوزيع الطبيعى إذا كانت ن = ١٠٠٠

الوضع النهائى للتوزيع التكرارى لذات الحدين عند ما تتخذ (ن) قيماً كبيرة

التوزيع الطبيعى

إذا زادت قيمة ن فى التوزيع التكرارى لذات الحدين (ح + ل)^ن إلى ما لا نهاية ، فإن عدد فئات التوزيع تزيد أيضاً إلى ما لا نهاية ويتخذ التوزيع التكرارى لذات الحدين شكل التوزيع الطبيعى Normal Distribution وتصبح أضلاع المضلع التكرارى متناهية فى الصغر فيتخذ المضلع شكل منحن متصل هو المنحنى الطبيعى أو منحنى الخطأ المتماثل (انظر شكل رقم ٤١) . ومعادلة هذا المنحنى هى :-

$$ص = \frac{١}{\sigma \sqrt{٢ ط}} \cdot هـ^{-\frac{س^٢}{٢\sigma^٢}}$$

(على فرض أن المساحة الواقعة تحت المنحنى = ١)

حيث ص = الإحداثى الرأسى المناظر لقيمة س

س = الإحداثى الأفقى مقيساً من نقطة الأصل التى تقع عند الوسط الحسابى للتوزيع وبعبارة أخرى فإن قيم س تمثل الانحرافات عن الوسط الحسابى للتوزيع .

σ = الانحراف المعيارى للتوزيع

ط = النسبة التقريبية (٣,١٤١٦)

هـ = أساس لوغاريتمات نابيير (٢,٧١٨٢٨)

ويمكن استنتاج هذه المعادلة من التوزيع التكرارى لذات الحدين (ح + ل)^ن عندما تقرب ن من ∞ وعلى فرض أن ح = ل (أى على فرض أن التوزيع متماثل) بالطريقة الآتية :-

احتمال النجاح ر مرة إذا كان عدد المحاولات ن هو

$$ن ق_ر \, ح^{ر} \, ل^{ن-ر}$$

وحيث أن ح = ل = $\frac{١}{٢}$

∴ احتمال النجاح ر مرة

$$= ن ق_ر \, (\frac{١}{٢})^{ن} = \frac{ن!}{ر!(ن-ر)!} (\frac{١}{٢})^{ن}$$

ولتسهيل العمل نفرض أن ن عدد زوجى ويساوى ٢ م ، فيكون احتمال النجاح م مرة أى $ن ق_م \, (\frac{١}{٢})^{ن}$ هو أكبر احتمال .

وحيث أن الاحتمال هو عبارة عن تكرار نسبى كما بينا ، فإن التكرار المقابل لعدد من مرات النجاح قدره م يكون أكبر تكرار . (*)

وبعبارة أخرى فإن الإحدائى الرأسى المقابل لعدد من مرات النجاح قدره م هو الإحدائى المقابل للقيمة المنوالية فى التوزيع وينحدر المنحنى بانتظام نحو المحور الأفقى من جانبى هذا الإحدائى . فإذا رمزنا إلى قيمة هذا الإحدائى بالرمز (ص.) فإن :

$$ص. \, (\text{احتمال النجاح م مرة}) = ن ق_م \, (\frac{١}{٢})^{ن}$$

وحيث أن ن = ٢ م

$$\therefore \text{ص.} = \frac{(٢م)!}{م!\,م!} (\frac{١}{٢})^{٢م} \quad \cdots\cdots (١)$$

وإذا رمزنا للإحدائى المقابل لعدد من مرات النجاح قدره (م + س)

(*) من الواضح أن م = $\frac{ن}{٢}$ = الوسط الحسابى = المنوال (نظراً لتماثل التوزيع) وهذا يقابله أكبر تكرار كما نعلم .

بالرمز ص$_{س}$ فإن ص$_{س}$ (احتمال النجاح م + س مرة)

$$= \frac{(٢م)!}{(م+س)!\,(م-س)!} \left(\frac{١}{٢}\right)^{٢م} \cdots\cdots (٢)$$

وبقسمة (٢) على (١)

$$\therefore \frac{ص_{س}}{ص_{.}} = \frac{م\,(م-١)\,(م-٢) \cdots (م-س+١)}{(م+س)\,(م+س-١) \cdots (م+٣)\,(م+٢)\,(م+١)}$$

وبعكس وضع حدود المقام

$$= \frac{م\,(م-١)\,(م-٢) \cdots (م-س+١)}{(م+١)\,(م+٢)\,(م+٣) \cdots (م+س-١)\,(م+س)}$$

بالنظر إلى هذا الكسر نجد أن كلا من البسط والمقام يحتوى على س من الحدود .

وبقسمة كل حد من هذه الحدود فى البسط والمقام على م فإن

$$\frac{ص_{س}}{ص_{.}} = \frac{(١-\frac{١}{م})\,(١-\frac{٢}{م}) \cdots (١-\frac{س-١}{م})}{(١+\frac{١}{م})\,(١+\frac{٢}{م}) \cdots (١+\frac{س-١}{م})\,(١+\frac{س}{م})}$$

والآن افرض أن قيمة م كبيرة جداً بحيث تكون $\frac{س}{م}$ كمية صغيرة وبحيث تكون $(\frac{س}{م})^{٢}$ من الصغر بحيث يمكن إهمالها ، ثم خذ لوغاريتمات الطرفين .

$$\text{لو} \left(\frac{ص_{س}}{ص_{.}}\right) = \left[\text{لو}\,(١-\frac{١}{م}) + \text{لو}\,(١-\frac{٢}{م}) + \cdots\cdots\right]$$

$$- \left[\text{لو}\,(١+\frac{١}{م}) + \text{لو}\,(١+\frac{٢}{م}) + \cdots\right]$$

وبإيجاد مفكوك كل قوس من الأقواس الموجودة فى الطرف الأيسر بتطبيق قانون المتسلسلة اللوغاريتمية مع إهمال جميع الحدود التى تشمل $\left(\frac{س^{٢}}{م^{٢}}\right)$ أو قوى أكبر من ذلك أو بعبارة أخرى مع إهمال جميع الحدود ما عدا الحد الأول من مفكوك كل قوس (*)

$$\therefore \text{ لو } \left(\frac{ص_{س}}{ص_{.}}\right) = \left[-\frac{١}{م} - \frac{٢}{م} - \frac{٣}{م} - \cdots - \frac{س-١}{م}\right]$$

$$- \left[\frac{١}{م} + \frac{٢}{م} + \frac{٣}{م} + \cdots + \frac{س-١}{م} + \frac{س}{م}\right]$$

$$= -\frac{٢}{م}\left[١ + ٢ + ٣ + ٤ + \cdots (س-١)\right] - \frac{س}{م}$$

وحيث أن مجموع الأعداد الطبيعية من ١ إلى (س − ١) = $\frac{(س-١)(س)}{٢}$

$$\therefore \text{ لو } \left(\frac{ص_{س}}{ص_{.}}\right) = -\frac{٢}{م}\left[\frac{س(س-١)}{٢}\right] - \frac{س}{م}$$

$$= -\frac{س(س-١)}{م} - \frac{س}{م} = -\frac{س^{٢}}{م}$$

$$\therefore \frac{ص_{س}}{ص_{.}} = هـ^{-\frac{س^{٢}}{م}} \text{ (حيث هـ أساس لوغاريتمات نابيير)}$$

(*) قانون المتسلسلة اللوغاريتمية هو لو(١ + س) = $س - \frac{س^{٢}}{٢} + \frac{س^{٣}}{٣} - \frac{س^{٤}}{٤} + \ldots$

ولو(١ − س) = $-س - \frac{س^{٢}}{٢} - \frac{س^{٣}}{٣} - \frac{س^{٤}}{٤} - \ldots$

ويلاحظ أننا نكتفى فى هذه الحالة بالحد الأول فقط فى مفكوك كل قوس حيث أن الحدود الأخرى صغيرة جداً ويمكن إهمالها.

ولكن $\sigma = \sqrt{\text{ن ع ل}} = \sqrt{\text{ن} \times \frac{1}{2} \times \frac{1}{2}} = \sqrt{\frac{\text{م}^2}{4}} = \frac{\text{م}}{2}$

$\therefore \quad \sigma^2 = \frac{\text{م}}{2} \quad \therefore \text{م} = 2\sigma^2$

$$\therefore \quad \frac{\text{ص}_{\text{س}}}{\text{ص}_{\cdot}} = \text{هـ}^{-\frac{\text{س}^2}{2\sigma^2}}$$

$$\therefore \text{ص}_{\text{س}} = \text{ص}_{\cdot} \, \text{هـ}^{-\frac{\text{س}^2}{2\sigma^2}} \quad \ldots\ldots\ldots (٣)$$

ويمكن إيجاد قيمة ص. (احتمال النجاح م مرة أو أكبر تكرار) بالطريقة الآتية :-

$$\text{ص}_{\cdot} = \frac{\text{ن}!}{\text{م}!\,(\text{ن}-\text{م})!} \, \text{ح}^{\text{م}} \, \text{ل}^{\text{ن}-\text{م}}$$

وقد سبق أن بينا أن م هى القيمة المنوالية وهذه تكون معادلة لقيمة الوسط الحسابى فى التوزيعات المتماثلة أى أن م = ن ح

$$\therefore \text{ص}_{\cdot} = \frac{\text{ن}!}{(\text{ن ح})!\,(\text{ن}-\text{ن ح})!} \, \text{ح}^{\text{ن ح}} \, \text{ل}^{\text{ن}-\text{ن ح}}$$

وحيث أن (١ – ح) = ل

$$\therefore \text{ص}_{\cdot} = \frac{\text{ن}!}{(\text{ن ح})!\,(\text{ن ل})!} \, \text{ح}^{\text{ن ح}} \, \text{ل}^{\text{ن ل}}$$

وبتطبيق قانون سترلنج (Stirling's Theorem) لإيجاد قيمة تقريبية لـ ن !(*)

$$\therefore ص = \frac{\sqrt{٢ط}\, ن^{ن+\frac{١}{٢}}\, هـ^{-ن}}{\sqrt{٢ط}\, ح^{ح+\frac{١}{٢}}\, هـ^{-ح} \times \sqrt{٢ط}\, ل^{ل+\frac{١}{٢}}\, هـ^{-ل}}$$

$$= \frac{\sqrt{٢ط}\, ن^{ن+\frac{١}{٢}}\, هـ^{-ن}}{\sqrt{٢ط}\, ح^{ح+\frac{١}{٢}}\, ل^{ل+\frac{١}{٢}}\, هـ^{-(ح+ل)}}$$

وحيث أن $هـ^{-ح} \times هـ^{-ل} = هـ^{-(ح+ل)} = هـ^{-ن}$

$$\therefore ص = \frac{ن^{ن+\frac{١}{٢}}}{\sqrt{٢ط}\, ح^{ح+\frac{١}{٢}}\, ل^{ل+\frac{١}{٢}}}$$

حيث أن $ن^{ن} = ن^{(ح+ل)} = ن^{ح+ل} = ن^{ح} \times ن^{ل}$

$$\therefore ص = \frac{ن^{\frac{١}{٢}}}{\sqrt{٢ط}\, ح^{ح+\frac{١}{٢}}\, ل^{ل+\frac{١}{٢}}}$$

$$= \frac{ن^{\frac{١}{٢}}}{\sqrt{٢ط}\, ح^{\frac{١}{٢}}\, ل^{\frac{١}{٢}}}$$

$$= \frac{١}{\sqrt{٢ط}\, ح^{\frac{١}{٢}}} = \frac{١}{\sqrt{٢ط}\, \sqrt{ن ح ل}}$$

(*) قانون سترلنج هو :

$$ن! = \sqrt{٢ط}\, ن^{ن+\frac{١}{٢}}\, هـ^{-ن}$$

$$= \frac{١}{\sigma \sqrt{٢ط}}$$

وبالتعويض عن قيمة ص. فى (٣) تكون معادلة المنحنى الطبيعى

$$ص_{س} = \frac{١}{\sigma \sqrt{٢ط}} \; هـ^{-\frac{س^{٢}}{٢\sigma^{٢}}}$$

هذا على فرض أن المساحة الواقعة تحت المنحنى تساوى ١ أى على فرض أن مجموع التكرارات = مجموع الاحتمالات = ١

فإذا كان مجموع التكرارات = ن تكون معادلة المنحنى الطبيعى هى

$$ص = \frac{ن}{\sigma \sqrt{٢ط}} \; هـ^{-\frac{س^{٢}}{٢\sigma^{٢}}}$$

ويمكن بطريقة مماثلة استنتاج هذه المعادلة من التوزيع التكرارى لذات الحدين إذا كانت (ح) لا تساوى (ل).

خواص المنحنى الطبيعى :

(١) المنحنى الطبيعى ناقوسى الشكل ومتماثل تماماً حول محور تماثل يقطع المحور الأفقى عند النقطة التى تحدد مركز كل من الوسط الحسابى والوسيط والمنوال وهى النقطة التى تعتبر عادة نقطة الأصل عند تمثيل هذا المنحنى بيانياً ومحور التماثل هو عبارة عن الإحداثى الرأسى الذى يمثل أكبر تكرار. ودلالة ذلك هو أن قيم المتغيرات الموزعة توزيعاً طبيعياً (مثل الطول والوزن وغير ذلك من الظواهر الطبيعية) تتركز عند المتوسط. وكلما ابتعدنا عن

المتوسط سواء فى اتجاه موجب أو فى اتجاه سالب كلما قل تركز القيم حتى إذا ما وصلنا إلى طرفى المنحنى تقل التكرارات جداً حتى تتلاشى كلية . وهذه التكرارات القليلة تمثل القيم الشاذة التى لا تكاد تخلو منها أية ظاهرة طبيعية .

(٢) مفروض نظرياً أن طرفى المنحنى يمتدان فى الاتجاهين الموجب والسالب إلى + ∞ ، − ∞ دون أن يقطعا المحور الأفقى . غير أن ذلك الجزء من المنحنى الذى يعتبر ذا فائدة عملية هو المساحة المحصورة بين القيمتين ، « $\bar{س}$ + ٣ σ » ، ($\bar{س}$ − ٣ σ » حيث أن هاتين القيمتين تحصران بينهما ٩٩,٧٣ ٪ من المساحة الكلية أى أن جميع التكرارات تقريباً تكون واقعة فى هذا الجزء من المنحنى

(٣) أهم خواص المنحنى الطبيعى هو أنه يمكن تعيين مساحة أجزائه المحصورة بين أى قيمتين عن طريق قياس انحراف (أو فرق) كل من القيمتين عن الوسط الحسابى والتعبير عن هذا الفرق بدلالة وحدات الانحراف المعيارى .

فمثلا المساحة المحصورة بين ($\bar{س}$ + σ) ، ($\bar{س}$ − σ) = ٦٨,٢٧ ٪ من المساحة الكلية .

والمساحة المحصورة بين ($\bar{س}$ + ٢ σ) ، ($\bar{س}$ − ٢ σ) = ٩٥,٤٥ ٪ من المساحة الكلية .

والمساحة المحصورة بين ($\bar{س}$ + ٣ σ) ، ($\bar{س}$ − ٣ σ) = ٩٩,٧٣ ٪ من المساحة الكلية .

والمساحة المحصورة بين ($\bar{س}$ + ٠,٦٧٤٥ σ) ، ($\bar{س}$ − ٠,٦٧٤٥ σ) = ٥٠ ٪ من المساحة الكلية .

ومن هذا نستنتج أن :

٦٨,٢٧ ٪ من التكرارات تكون واقعة بين ($\bar{س}$ + σ) ، ($\bar{س}$ − σ)

∴ احتمال وقوع قيمة من قيم المتغير خارج هذين الحدين = ٣١,٧٣ ٪

و ٩٥,٤٥٪ من التكرارات تكون واقعة بين (س+٢ σ) ، (س−٢ σ)

∴ احتمال وقوع قيمة من قيم المتغير خارج هذين الحدين = ٤,٥٥ ٪
و٩٩,٧٣٪ من التكرارات تكون واقعة بين ($\bar{س}$+٣σ) ، ($\bar{س}$−٣σ)

∴ احتمال وقوع قيمة من قيم المتغير خارج هذين الحدين = ٠,٢٧ ٪
و٥٠ ٪ من التكرارات تكون واقعة بين ($\bar{س}$ + ٠,٦٧٤٥ σ) ،
($\bar{س}$ − ٠,٦٧٤٥ σ) .

∴ احتمال وقوع قيمة من قيم المتغير خارج هذين الحدين = ٥٠ ٪
ويمكن تعيين المساحة المحصورة بين ($\bar{س}$ + ر σ) ، ($\bar{س}$ − ر σ)
حيث ر تمثل أية قيمة كانت باستعمال الجداول المعدة لذلك والواردة فى نهاية هذا الكتاب . وهذه الخاصية هى التى سنستند إليها فى توفيق المنحنى الطبيعى للظواهر المختلفة .

عزوم المنحنى الطبيعى :

يمكن إيجاد عزوم المنحنى الطبيعى بالتكامل بنفس الطريقة التى أوضحناها فى نهاية الباب السابق . وحيث أن هذا المنحنى متماثل فإن جميع العزوم الفردية تساوى صفراً وبالتالى تكون قيمة ($ل_١$) أى معامل الالتواء تساوى صفراً .

$$\text{والعزم الثانى } م_٢ = \frac{١}{\sigma \sqrt{٢ط}} \int_{-\infty}^{+\infty} س^٢ هـ^{-\frac{س^٢}{٢\sigma^٢}} \, ى \, س$$

وباستخدام طريقة التكامل بالتجزيء نجد أن $م_٢ = \sigma^٢$

أى أن σ المبينة فى معادلة المنحنى هى عبارة عن الانحراف المعيارى للتوزيع .

---

(١) للوقوف على خطوات العمل التفصيلية الخاصة بإيجاد عزوم المنحنى الطبيعى بالتكامل انظر كتاب A first course in statistics by D. C. Jones ص ٢٠٧ .

والعزم النونى (على فرض أن $ن$ عدد زوجى)

$$م_{ن} = \frac{١}{\sigma \sqrt{٢ط}} \int_{-\infty}^{+\infty} س^{ن} هـ^{-\frac{س^{٢}}{٢\sigma^{٢}}} . ى س$$

$$= \sigma^{ن} \frac{ن!}{! \left(\frac{ن}{٢}\right) ٢^{\frac{ن}{٢}}}$$

وبوضع قيمة $ن = ٤$ نجد أن العزم الرابع للمنحنى الطبيعى

$$م_{٤} = \frac{٤!}{! (٢) ٢^{٢}} \sigma^{٤} = ٣\sigma^{٤}$$

ومن ثم نستنتج أن معامل التفرطح فى المنحنى الطبيعى

$$ل_{٢} = \frac{م_{٤}}{(م_{٢})^{٢}} = \frac{٣\sigma^{٤}}{\sigma^{٤}} = ٣$$

ومن هذه النتيجة يتضح لنا السبب فى اعتبار ($ل_{٢} = ٣$) الحد الفاصل بين المنحنيات المدببة القمة (Platykurtic) والمنحنيات المفرطحة (Leptokurtic) إذ يظهر من ذلك أن المنحنى الطبيعى هو الذى يحدد مستوى الاعتدال فى التفرطح (Mesokurtosis) .

الوضع النهائى للتوزيع التكرارى لذات الحدين عندما تتخذ (ح) قيماً صغيرة :

توزيع بواسون Poisson Distribution .

افرض أن حدثاً يقع فى المتوسط م مرة إذا كان عدد المحاولات ن

بحيث يكون احتمال وقوع الحدث ح = $\frac{م}{ن}$ (أو م = ن ح = $\overline{س}$)

ثم افرض أن هذا الحدث نادر الوقوع وأن ح صغيرة جداً بحيث لو زادت ن حتى قاربت ∞ تظل م تساوى عدداً محدوداً .

احتمال النجاح ر مرة فى عدد من المحاوت قدره ن

$$= \frac{ن!}{ر!\,(ن-ر)!}\, ح^{ر}\,(١-ح)^{ن-ر} \quad (\text{بتطبيق قانون برنوللى}) .$$

$$= \frac{ن\,(ن-١)\,(ن-٢)\,(ن-٣)\cdots(ن-ر+١)}{ر!}\, ح^{ر}\,(١-ح)^{ن-ر}$$

بتوزيع ح^{ر} على جميع حدود البسط وعددها ر وذلك بضرب كل حد فى ح

$$= \frac{ن ح\,(ن ح-ح)(ن ح-٢ح)\,(ن ح-٣ح)\cdots[ن ح-(ر-١)\,ح]}{ر!}\,(١-ح)^{ن-ر}$$

$$= \frac{ن ح(ن ح-ح)\,(ن ح-٢ح)\,(ن ح-٣ح)\cdots[ن ح-(ر-١)ح]}{ر!\,(١-ح)^{ر}}\,(١-ح)^{ن}$$

ولكن م = ن ح

... احتمال النجاح ر مرة

$$= \frac{م\,(م-ح)\,(م-٢ح)\,(م-٣ح)\cdots[م-(ر-١)\,ح]}{ر!\,(١-ح)^{ر}}\,(١-ح)^{\frac{ن}{ح}}$$

$$= \frac{م\,(م-ح)\,(م-٢ح)\,(م-٣ح)\cdots[م-(ر-١)\,ح]}{ر!\,(١-ح)^{ر}}\left\{(١-ح)^{\frac{١}{ح}}\right\}^{-م}$$

والآن إذا كانت ح صغيرة جداً كما فرضنا فإن

$(١-ح)^{-\frac{١}{ح}}$ تساوى هـ أساس لوغاريتمات نابيير وما دامت ح صغيرة جداً فيمكن اعتبار (م − ح) = م و (م − ٢ ح) = م

وهكذا بالنسبة لبقية حدود بسط الكسر

كذلك فإن الكمية $(١-ح)^{ر}$ تقرب جداً من ١ .

وعلى هذا الأساس يكون احتمال النجاح ر مرة فى ن من المحاولات

$$= \frac{م^{ر}}{ر!}\,هـ^{-م}$$

وهذا هو قانون بواسون أو قانون الاحتمالات الصغيرة . وبإعطاء ر قيماً مختلفة (من . إلى ن) نجد أن حدود التوزيع التكرارى لذات الحدين تتخذ الوضع المبين بالجدول الآتى وهو المعروف بتوزيع بواسون .

توزيع بواسون

عدد مرات النجاح	احتمال النجاح (التكرار النسبى)
٠	$هـ^{-م}$
١	$م\ هـ^{-م}$
٢	$\frac{م^{٢}}{٢!}\ هـ^{-م}$
٣	$\frac{م^{٣}}{٣!}\ هـ^{-م}$
:	:
:	:
:	:
$ن$	$\frac{م^{ن}}{ن!}\ هـ^{-م}$

ويلاحظ أن مجموع الاحتمالات (أى مجموع التكرارات النسبية) .

$$= هـ^{-م} (١ + م + \frac{م^{٢}}{!٢} + \frac{م^{٣}}{!٣} + \cdots \frac{م^{ن}}{ن!}) .$$

$$= هـ^{-م} \times هـ^{م} = ١$$

عزوم توزيع بواسون :

لإيجاد الوسط الحسابى لتوزيع بواسون نضرب قيم المتغير فى التكرارات المقابلة لها (الموضحة بالجدول السابق) للحصول على $\Sigma\ س_{ن}\ و$.

$$\Sigma\, س_{ن}\, و = (٠) \times هـ^{-م} + (١) \times م هـ^{-م} + (٢) \times \frac{م^{٢}}{٢!} هـ^{-م}$$

$$+ \cdots + (ن) \times \frac{م^{ن}}{ن!} هـ^{-م}$$

وبأخذ الكمية م هـ$^{-م}$ مضروباً مشتركاً فى حدود الطرف الأيسر

$$\therefore \Sigma\, س_{ن}\, و = م هـ^{-م} \left(١ + م + \frac{م^{٢}}{٢!} + \frac{م^{٣}}{٣!} + \cdots + \frac{م^{ن-١}}{(ن-١)!} \right)$$

$$= م\, هـ^{-م} \times هـ^{م} = م$$

وحيث أن مجموع التكرارات أى $\Sigma\, و = ١$

فإن الوسط الحسابى لتوزيع بواسون = $\Sigma\, س_{ن}\, و = م$

ولإيجاد الانحراف المعيارى نستخرج أولا قيمة $\Sigma\, س_{ن}^{٢}\, و$

$$\Sigma\, س_{ن}^{٢}\, و = (٠) \times هـ^{-م} + (١) \times م\, هـ^{-م} + (٢)^{٢} \times \frac{م^{٢}}{٢!} هـ^{-م}$$

$$+ \cdots + (ن)^{٢} \times \frac{م^{ن}}{ن!} هـ^{-م}$$

$$= م هـ^{-م} \left(١ + (٢) م + (٣) \frac{م^{٢}}{٢!} + \cdots + (ن) \frac{م^{ن-١}}{(ن-١)!} \right)$$

(١٨)

$$= م\,هـ^{-م}\left(1 + (1+1)م + (2+1)\frac{م^2}{!2} + \cdots + \left[(ن-1)+1\right]\frac{م^{ن-1}}{!(ن-1)}\right)$$

بفصل حدود الطرف الأيسر المضروبة فى (١) على حدة :

$$\therefore \Sigma\, س^2 و = م\,هـ^{-م}\left\{\left(1 + م + \frac{م^2}{!2} + \cdots + \frac{م^{ن-1}}{!(ن-1)}\right) + \left\{\left(م + م^2 + \frac{م^3}{!2} + \cdots + \frac{م^{ن-2}}{!(ن-2)}\right)\right.\right.$$

$$\left.\left. = م\,هـ^{-م}\left\{\left(هـ^{م} + م(1 + م + \frac{م^2}{!2} + \cdots + \frac{م^{ن-2}}{(ن-2)!}\right)\right.\right.$$

$$= م\,هـ^{-م}(هـ^{م} + م\,هـ^{م})$$

$$= م + م^2$$

وحيث أن $\sigma^2 = \Sigma\, س^2 و - (\bar{س})^2$ (لاحظ أن $\Sigma\, و = 1$)

$$\therefore \sigma^2 = م + م^2 - م^2 = م \text{ وهذه هى قيمة العزم الثانى } مـ_2$$

$$\therefore \sigma = \sqrt{م}$$

وبالمثل فإن العزم الثالث $مـ_3 = م$

والعزم الرابع $مـ_4 = 3م^2 + م$

ومن هذا نستنتج أن قيمة معامل الالتواء :

$$ل_1 = \frac{(مـ_3)^2}{(مـ_2)^3} = \frac{م^2}{م^3} = \frac{1}{م}$$

ومعامل التفرطح $ل_2 = \frac{مـ_4}{(مـ_2)^2} = \frac{3م^2 + م}{م^2} = 3 + \frac{1}{م}$

أى أن المنحنى مدبب القمة (Leptokurtic) ما دامت قيمة م موجبة

الباب التاسع

توفيق المنحنيات واختبار حسن المطابقة

الوضع المشاهد والوضع المثالى لتغير الظواهر :

ذكرنا عند شرح طرق العرض البيانى أن الرسوم البيانية الإحصائية التى تسجل تغير الظواهر التى نحصل على قيمها من واقع المشاهدة تظهر على شكل خطوط منكسرة كثيرة الذبذبات والتعاريج فى حين أن الرسوم البيانية الرياضية تظهر على شكل خطوط مستقيمة أو منحنيات ممهدة متصلة . وذكرنا أن سبب هذا الاختلاف هو أن الرسوم البيانية الإحصائية تسجل قيم ظواهر تخضع فى تغيرها لعدة مؤثرات بعضها يحدث فى شىء من الانتظام والبعض الاخر يحدث عرضاً فيؤثر على قيمة الظاهرة بالصعود أحياناً وبالهبوط أحياناً أخرى ، فى حين أن الرسوم البيانية الرياضية هى عبارة عن خطوط سير نقط تخضع فى تغيرها لنظام ثابت تحدده علاقة جبرية . ثم أشرنا إلى أن التأمل فى بعض الرسوم البيانية الإحصائية يوحى إلينا أحياناً بوجود شىء من الانتظام فى طريقة تغير الظواهر المشاهدة . والواقع أن معظم الظواهر التى نسجل قيمها من واقع المشاهدة تتغير طبقاً لنظام معين ، غير أن التقلبات العرضية التى تطرأ عليها تحجب أحياناً هذا النظام عن الظهور بوضوح فى الرسم البيانى فتبدو لنا هذه الظواهر وكأنها تتغير دون ما قاعدة . فى كثير من الحالات نجد أن الخطوط البيانية التى تمثل الظواهر المشاهدة ، مع كثرة ذبذباتها وتعاريجها ، تقارب فى شكلها العام شكلا هندسياً معروفاً الأمر الذى يحملنا على الاعتقاد بأنه لولا وجود تلك التقلبات العرضية لأصبح الخط البيانى الذى يمثل الظاهرة

المشاهدة مطابقاً للشكل الهندسى الممهد . وبعبارة أخرى فإن فى هذا ما يحملنا على الاعتقاد بأنه إذا كان الخط البيانى الكثير التعاريج يمثل الوضع المشاهد لتغير الظاهرة فإن الشكل الهندسى المقارب له يمثل الوضع المثالى لهذا التغير ، أى التغير الحقيقى فى قيمة الظاهرة مجرداً عن المؤثرات العرضية . وحيث أنه من المهم جداً أن نقف على الوضع المثالى لتغير الظواهر بعد استبعاد المؤثرات العرضية فلا بد أن نبحث عن طريقة لمعرفة أقرب شكل هندسى يصف تغيرها وتعيين المعادلة الجبرية التى تمثله . وتوفيق المنحنيات هو عبارة عن اختيار هذا الشكل الهندسى وتعيين معادلته الجبرية .

كيفية اختيار المنحنى الذى يناسب تغير الظاهرة :

فى بعض الحالات يكفى مجرد النظر إلى الرسم البيانى الذى يسجل قيم الظاهرة كما نشاهدها لتعيين نوع المنحنى الذى يمثلها . فقد نجد أن الخط البيانى الذى يمثل تغير الظاهرة قريب جداً من المستقيم كما هو الحال لو رصدنا على الرسم قيما ظاهرة تزداد أو تنقص بكمية ثابتة تقريباً سنة بعد أخرى . ومن الواضح أنه فى مثل هذه الحالة نوفق للظاهرة مستقيماً .

ومن المعروف أن معادلة المستقيم هى ص = ا + ب س حيث (س) هى المتغير المستقل و (ص) المتغير التابع و (ا) هى قيمة ص عند ما تكون(س=٠) أو بعبارة أخرى فهى المسافة بين نقطة الأصل ونقطة تقاطع المستقيم مع المحور الصادى . أما (ب) فهى عبارة عن ميل المستقيم بالنسبة للمحور السينى أو ظل زاوية تقاطعه مع هذا المحور .

ويكفى لرسم المستقيم أن نحدد قيمتى ا ، ب . والطريقة المتبعة عادة فى توفيق المستقيمات هى أن نختار المستقيم الذى يكون مجموع مربعات انحرافات

القيم عنه أقل من مجموع مربعات انحرافات القيم عن أى مستقيم آخر . وتسمى هذه الطريقة طريقة المربعات الدنيا (Method of Least Squares) وسنشرحها فيما يلى .

وقد نجد بمجرد النظر إلى الخط البيانى الذى يمثل قيم الظاهرة أنه أقرب إلى المنحنى منه إلى المستقيم . وفى هذه الحالة نوفق للظاهرة منحنياً إما من الدرجة الثانية أى قطع مكافئ (انظر شكل ٧ ص ٤٩) أو من الدرجة الثالثة أو من درجات أعلى بطريقة المربعات الدنيا كذلك(١) .

على أنه كثيراً ما نجد أنه من الصعب تعيين درجة معادلة المنحنى الذى يناسب تغير الظاهرة من مجرد النظر إلى الرسم . وفى مثل هذه الحالات نستعين بجدول كالآتى يسمى جدول الفروق .

جدول رقم ٣٦

يوضح تساوى الفروق الثالثة فى دوال الدرجة الثالثة(٢)

قيم س	قيم ص	الفروق الأولى $\triangle^1$	الفروق الثانية $\triangle^2$	الفروق الثالثة $\triangle^3$	الفروق الرابعة $\triangle^4$
٠	٠				
		١			
١	١				
		٧	٦		
٢	٨			٦	
		١٩	١٢		٠
٣	٢٧			٦	
		٣٧	١٨		٠
٤	٦٤			٦	
		٦١	٢٤		
٥	١٢٥				

(١) هناك طريقة أخرى لتوفيق المنحنيات هى طريقة العزوم (Method of Moments) ولكن مجالنا هذا لا يسمح بشرحها ويستطيع الطالب أن يرجع إليها فى كتاب A First Course in Statistics By C. Jones ص ١٩٤

(٢) دوال الدرجة الثالثة هى تلك التى تكون فيها (س) مرفوعة إلى القوة الثالثة مثل الدالة ص = س^3 المبينة بعض قيمها العددية بالجدول .

بعد عمل جدول كهذا نوضح فيه قيم س وقيم ص المناظرة لها والفروق بين قيم ص المتتالية ننظر إلى هذه الفروق ، فإذا كانت الفروق الأولى متساوية تقريباً (أى إذا كانت الفروق الثانية صغيرة جداً وتقرب من الصفر أو تساويه) دل ذلك على أن أنسب شكل هندسى لتمثيل تغير الظاهرة هو الخط المستقيم .

وإذا كانت الفروق الثانية متساوية تقريباً (أى إذا كانت الفروق الثالثة تقرب من الصفر أو تساويه) دل ذلك على أن أنسب شكل هندسى لتمثيل تغير الظاهرة هو منحن معادلته من الدرجة الثانية . أما إذا كانت الفروق الثالثة متساوية (أى إذا كانت الفروق الرابعة تساوى صفراً) كما هو الحال فى المثال الموضح بالجدول السابق دل ذلك على أن المنحنى الذى يمثل الظاهرة معادلته من الدرجة الثالثة . وبالمثل إذا كانت الفروق «النونية» متساوية دل ذلك على أن أنسب منحن لوصف تغير الظاهرة هو منحن معادلته من الدرجة النونية أى تكون قيمة س فيها مرفوعة إلى القوة النونية . والصيغة العامة لهذه المعادلة هى :

$$ص = ا_{١} + ا_{٢} س + ا_{٣} س^{٢} + \cdots + ا_{ن} س^{ن-١}$$

ومن الواضح أن توفيق منحن من هذه الدرجة يتطلب تعيين قيم

$ا_{١} ، ا_{٢} ، ا_{٣} ، \cdots ، ا_{ن}$ وهذه يمكن تعيينها بطريقة المربعات الدنيا التى ننتقل الان إلى شرحها .

يتطلب توفيق مستقيم بطريقة المربعات الدنيا تعيين قيمتى ا ، ب فى المعادلة $ص = ا + ب س$

ويمكن استنتاج هاتين القيمتين من المعادلتين اللتين تسميان بالمعادلتين الطبيعيتين :—

$$\Sigma ص = ن ا + ب \Sigma س$$

$$\Sigma س ص = ا \Sigma س + ب \Sigma س^{٢}$$

ومن الواضح أن هاتين المعادلتين تحتويان على مجهولين فقط هما ا ، ب إذ أن كلا من Σ ص ، Σ سـ ، Σ سـ ص ، Σ سـ^{٢} تحسب من واقع القيم المشاهدة للظواهر المطلوب توفيق مستقيمات لوصف تغيرها . أما (ن) فهى عدد القيم المشاهدة .

وفيما يلى طريقة الحصول على هاتين المعادلتين :

نفرض أن القيمة المشاهدة للظاهرة هى (ص) وأن قيمة الظاهرة المحسوبة من المعادلة هى ص$_{م}$.

أفضل مستقيم ، طبقاً لطريقة المربعات الدنيا ، هو ذلك الذى يجعل مجموع مربعات انحرافات القيم المشاهدة عن تلك التى نستنتجها من المعادلة حداً أدنى أى أن أفضل مستقيم هو ذلك الذى يجعل قيمة Σ (ص − ص$_{م}$)^{٢} أقل ما يمكن .

حيث أن ص$_{م}$ = ا + ب سـ .

∴ Σ (ص − ص$_{م}$)^{٢} = Σ [(ص − (ا + ب سـ)]^{٢} = Σ (ص − ا − ب سـ)^{٢}

= Σ (ص^{٢} − ٢ا ص − ٢ب سـ ص + ا^{٢} + ٢ا ب سـ + ب^{٢} سـ^{٢})

= Σ ص^{٢} − ٢ا Σ ص − ٢ب Σ سـ ص + ن ا^{٢} + ٢ا ب Σ سـ + ب^{٢} Σ سـ^{٢}

بتسمية هذه الكمية (ط) ثم مفاضلتها جزئياً بالنسبة لـ ا (١)

∴ $\frac{\text{ء ط}}{\text{ء ا}}$ = − ٢ Σ ص + ٢ن ا + ٢ ب Σ س ... (١)

وباستخدام نظرية النهايات الكبرى والصغرى نجعل هذه الكمية الأخيرة تساوى صفراً (١) .

∴ − ٢ Σ ص + ٢ن ا + ٢ب Σ س = .

---

(١) من المعروف أن النهاية الصغرى (أو الكبرى) للمنحنى توجد حيث يكون ميل المنحنى يساوى صفراً . وميل المنحنى يساوى المشتقة الأولى (أو معامل التفاضل) التى استخرجنا قيمتها فى المعادلة (١) .

∴ ∑ ص = ن ا + ب ∑ س وهذه هى المعادلة الطبيعية الأولى .

بمفاضلة الكمية « ط » بالنسبة ل « ب »

$$\therefore \frac{\text{د ط}}{\text{د ب}} = -2 \sum \text{س ص} + 2 \text{ ا} \sum \text{س} + 2 \text{ ب} \sum \text{س}^2$$

بوضع هذه الكمية تساوى صفراً .

∴ − ٢ ∑ س ص + ٢ ا ∑ س + ٢ ب ∑ س^2 = ٠

∴ ∑ س ص = ا ∑ س + ب ∑ س^2 وهذه هى المعادلة الطبيعية الثانية .

وبنفس الطريقة يمكن استنتاج المعادلات الطبيعية الخاصة بمنحنيات الدرجة الثانية أو الدرجات العليا .

فمعادلة الدرجة الثانية هى ص = ا + ب س + ح س^2

والمعادلات الطبيعية اللازمة لتعيين قيم ا ، ب ، ح هى

$$\sum \text{ص} = \text{ن ا} + \text{ب} \sum \text{س} + \text{ح} \sum \text{س}^2$$

$$\sum \text{س ص} = \text{ا} \sum \text{س} + \text{ب} \sum \text{س}^2 + \text{ح} \sum \text{س}^3$$

$$\sum \text{س}^2\text{ص} = \text{ا} \sum \text{س}^2 + \text{ب} \sum \text{س}^3 + \text{ح} \sum \text{س}^4$$

وبحل هذه المعادلات باستخدام القيم المشاهدة لكل من س ، ص يمكن الحصول على قيم ا ، ب ، ح .

مثال على توفيق مستقيم بطريقة المربعات الدنيا :

الجدول الآتى يوضح مجمل ودائع صندوق التوفير بالقطر المصرى محسوبة كنسبة مئوية من قيمة الودائع فى سنة ١٩٠١ . والمطلوب توفيق مستقيم يصف تغير هذه الظاهرة بطريقة المربعات الدنيا .

جدول رقم ٣٧

ودائع صندوق التوفير بالقطر المصرى محسوبة كنسبة مئوية من قيمتها فى سنة ١٩٠١

السنة	الودائع	السنة	الودائع	السنة	الودائع
١٩١٩	٢٢١٥	١٩٢٣	٣٣٧١	١٩٢٧	٤٣٤٣
١٩٢٠	٢٤٢٦	١٩٢٤	٣٧٠٠	١٩٢٨	٤٦٦٠
١٩٢١	٢٦٦٩	١٩٢٥	٣٨٢٥	١٩٢٩	٤٨٣٥
١٩٢٢	٢٠٩٢	١٩٢٦	٣٩٧٩	١٩٣٠	٤٩٣٩

المصدر : هذه الأرقام محسوبة على أساس البيانات التى تنشرها مصلحة البريد فى تقاريرها السنوية

معادلة المستقيم هى ص$_م$ = ا + ب س$_م$.

حيث قيمتا ا ، ب تحددهما المعادلتان الطبيعتان الآتيتان :ـ

$$\Sigma ص = ن ا + ب \Sigma س$$

$$\Sigma س ص = ا \Sigma س + ب \Sigma س^{٢}$$

ولحل هاتين المعادلتين نحسب قيم كل من Σ ص ، Σ س ، Σ س ص ، Σ س^{٢} من واقع القيم المشاهدة للظاهرة باستخدام جدول كالآتى ، مع العلم بأن قيم س مقيسة من نقطة الأصل الواقعة عند سنة ١٩١٩* .

* من الممكن اختصار العمل الحسابى كثيراً باختيار نقطة الأصل عند السنة الوسطى إذا كان عدد السنين فردياً وفى منتصف الفترة الواقعة بين السنتين المتوسطتين إذا كان عدد السنين زوجياً . وسنوضح هذه الطرق المختصرة فى الباب الخاص بتحليل السلسلات الزمنية .

جدول رقم ٣٨

توفيق مستقيم لودائع صندوق التوفير بالقطر المصرى بطريقة المربعات الدنيا

السنة	القيم المشاهدة للظاهرة ص	س	س^{٢}	س ص	قيم الظاهرة من المعادلة ص$_{م}$
١٩١٩	٢٢١٥	٠	٠	٠	٢٢١٨,٠٧٣
١٩٢٠	٢٤٢٦	١	١	٢٤٢٦	٢٤٧٩,٣٩٣
١٩٢١	٢٦٦٩	٢	٤	٥٣٣٨	٢٧٤٠,٧١٣
١٩٢٢	٢٩٠٢	٣	٩	٨٧٠٦	٣٠٠٢,٠٣٣
١٩٢٣	٣٣٧١	٤	١٦	١٣٤٨٤	٣٢٦٣,٣٥٣
١٩٢٤	٣٧٠٠	٥	٢٥	١٨٥٠٠	٣٥٢٤,٦٧٣
١٩٢٥	٣٨٢٥	٦	٣٦	٢٢٩٥٠	٣٧٨٥,٩٩٣
١٩٢٦	٣٩٧٩	٧	٤٩	٢٧٨٥٣	٤٠٤٧,٣١٣
١٩٢٧	٤٣٤٣	٨	٦٤	٣٤٧٤٤	٤٣٠٨,٦٣٣
١٩٢٨	٤٦٦٠	٩	٨١	٤١٩٤٠	٤٥٦٩,٩٥٣
١٩٢٩	٤٨٣٥	١٠	١٠٠	٤٨٣٥٠	٤٨٣١,٢٧٣
١٩٣٠	٤٩٣٩	١١	١٢١	٥٤٣٢٩	٥٠٩٢,٥٩٣
المجموع	٤٣٨٦٤	٦٦	٥٠٦	٢٧٨٦٢٠	٤٣٨٦٣,٩٩٦

بالتعويض فى المعادلتين الطبيعيتين نجد أن :—

$$٤٣٨٦٤ = ١٢\ ا + ٦٦\ ب$$

و

$$٢٧٨٦٢٠ = ٦٦\ ا + ٥٠٦\ ب$$

وبحل هاتين المعادلتين نجد أن :—

$$ا = ٢٢١٨,٠٧٣$$

، $$ب = ٢٦١,٣٢$$

∴ معادلة المستقيم هى ص = ٢٢١٨,٠٧٣ = ٢٦١,٣٢ س

ومن هذه المعادلة يمكن حساب القيم النظرية للظاهرة بالتعويض عن قيمة س

فمثلا قيمة الظاهرة فى سنة ١٩١٩ = ٢٢١٨,٠٧٣ + ٠ × ٢٦١,٣٢ = ٢٢١٨,٠٧٣

وقيمة الظاهرة فى سنة ١٩٢٥ = ٢٢١٨,٠٧٣ + ٦ × ٢٦١,٣٢ = ٣٧٨٥,٩٩٣

ويلاحظ أن هذه القيم المحسوبة من المعادلة قريبة جداً من القيم المشاهدة المناظرة لها مما يدل على أن المستقيم الذى وفقناه يصلح لوصف تغير الظاهرة .

ويلاحظ كذلك أن طريقة العمل اللازمة لتوفيق قطع مكافىء لا تختلف عن هذه الطريقة إلا فى أننا نحتاج إلى حساب قيم $\Sigma س^3$ ، $\Sigma س^4$ ، $\Sigma س^2 ص$ علاوة على القيم التى استعملناها فى توفيق المستقيم . وسنشرح الطريقة العملية لتوفيق القطع المكافئ فى الباب الخاص بتحليل السلسلات الزمنية .

وبالإضافة إلى المستقيم والقطع المكافئ ، هناك مجموعة أخرى كبيرة من المنحنيات التى يمكن توفيقها للظواهر المختلفة حسب طبيعة تغير هذه الظواهر ونذكر منها ما يلى :-

١ - المنحنى اللوجيستى (Logistic Cuvre) .

$$\text{ومعادلته ص} = \frac{ك}{١ + هـ^{ا + ب س}}$$

حيث هـ أساس لوغاريتمات نابيير و ك ، ا ، ب هى الثوابت التى تحدد شكل المنحنى . وهذا المنحنى هو أنسب المنحنيات لوصف نمو الظواهر ، وعلى وجه الخصوص نمو السكان سنة بعد أخرى(١) .

---

(١) للوقوف على الطريقة العملية لتوفيق هذا المنحنى انظر كتاب Applied General Statistics By Croxton & Cowden ص ٤٥٣ .

٢ ــ المنحنى الذى معادلته ص = ا س^{ب} (Simple Parabolic Cuvre)

ويناسب الظواهر التى إذا رصدت على قياس لوغاريتمى مزدوج تظهر على شكل خط مستقيم أو قريب من المستقيم مثل ظاهرة توزيع الدخول التى أشرنا إليها من قبل[1].

وذلك لأنه إذا أخذنا لوغاريتمات الطرفين تصبح معادلة المنحنى .

لو ص = لو ا + ب لو س

ومن هذا يظهر أن هذا المنحنى يصبح خطاً مستقيما معادلته

ص = ا + ب س إذا رصدنا لوغاريتمات القيم بدلا من القيم ذاتها على الرسم البيانى .

٣ ــ المنحنى الطبيعى . (Normal Cuvre)

وقد سبق أن أشرنا إليه وذكرنا أنه يصف طائفة كبيرة من الظواهر الطبيعية. ونظراً لأهمية هذا المنحنى فى الدراسات الإحصائية سنشرح طريقة توفيقه للتوزيعات التكرارية بالتفصيل .

٤ ــ منحنى بواسون (Poisson Cuvre) .

وقد سبق أن أشرنا إليه ونضيف هنا أنه يصلح لوصف الظواهر والأحداث النادرة الوقوع . وسنشرح طريقة توفيقية بالتفصيل .

توفيق المنحنى الطبيعى للتوزيعات التكرارية .

كثيراً ما يكون لدينا من الأسباب ما يدعو إلى الاعتقاد بأن قيم مفردات مجموع معين موزعة توزيعاً طبيعياً .. فقد نختار من هذا المجموع عدة عينات

---

(١) راجع ص ٧٦ من هذا الكتاب . وللوقوف على الطريقة العملية لتوفيق هذا المنحنى انظر كتاب Elements of Statistics By Davis and Nelson ص ٢٢٥ .

ونجد فى كل مرة أن قيم مفردات هذه العينات موزعة توزيعاً يقرب من الطبيعى أو قد نجد أنه كلما كبر حجم العينة التى نختارها من هذا المجموع كلما قرب التوزيع التكرارى لمفردات العينة من الوضع الطبيعى الأمر الذى يؤيد أنه لو كان من الممكن عملياً جمع بيانات عن جميع مفردات المجموع لوجدنا أنها موزعة توزيعاً طبيعياً . وما دامت دراسة العينات هى مجرد وسيلة لدراسة المجموع الأصلى الذى انتخبت منه ، فيمكننا أن نكتفى بدراسة العينة على أن نمهد المضلع التكرارى الذى يمثل توزيع مفرداتها فنجعله يتخذ الوضع المثالى الذى هو المنحنى الطبيعى . وكما أن المضطلع التكرارى يمثل توزيع مفردات العينة المحدودة فإن المنحنى الطبيعى يمثل توزيع مفردات المجموع اللانهائى. ومتى وقفنا على طريقة توزيع مفردات المجموع فمن السهل وصف هذا المجموع والتعرف على خواصه وكأننا بذلك قد تمكننا من دراسة المجموع الأصلى ذاته عن طريق دراسة عينة مختارة من بين مفرداته . وعلى ضوء هذه المقدمة يمكننا أن نلخص المقصود بتوفيق المنحنى الطبيعى فى العبارة الآتية :—

لو فرض أن كان لدينا توزيع تكرارى قريب من الطبيعى لقيم ظاهرة معينة سجلناها من واقع المشاهدة أو القياس المباشر ، فكيف يكون توزيع هذه القيم فى وضعه المثالى ، أى فى الوضع الذى يتخذ فيه المضلع التكرارى للتوزيع شكل المنحنى الطبيعى ؟

الطريقة العملية لمعرفة هذا الوضع المثالى هى أن نعين التوزيع الطبيعى المشترك مع التوزيع التكرارى الذى نبحثه فى قيم المقاييس الوصفية الأساسية التى تحدده ، أى المشترك معه فى قيمتى الوسط الحسابى والانحراف المعيارى ، على فرض أن مجموعى مفردات التوزيعين متساويان .

يلاحظ أولا أن المساحة الواقعة تحت المنحنى الطبيعى الذى سنوفقه تعادل

المساحة الواقعة تحت المضلع التكرارى الذى يمثل التوزيع المشاهد لأن كلا من هاتين المساحتين يمثل عدد المفردات الذى فرضنا تساويه فى التوزيعين .

ويلاحظ كذلك أنه يمكننا تعيين نسبة مساحة المنحنى الطبيعى المحصورة بين أى قيمتين من قيم المتغير إلى المساحة الكلية للمنحنى متى عرفنا انحراف كل من القيمتين عن الوسط الحسابى مقدراً بوحدات الانحراف المعيارى*

فيمكننا مثلا تعيين نسبة المساحة الواقعة بين $(\bar{س} + \sigma)$ ، $(\bar{س} - \sigma)$ والتى سبق أن ذكرنا أنها تعادل ٦٨,٢٧٪ من المساحة الكلية للمنحنى . ويمكننا كذلك تعيين نسبة المساحة المحصورة بين الحد الأدنى والحد الأعلى لأى فئة من فئات التوزيع وبالتالى عدد المفردات الواقعة داخل هذه الفئة متى عينا انحرافى هذين الحدين عن الوسط الحسابى بدلالة وحدات الانحراف المعيارى ، وذلك باستخدام جدول مساحات المنحنى الطبيعى الوارد فى نهاية هذا الكتاب . ومتى عينا عدد المفردات الواقعة داخل كل فئة نكون قد حددنا التوزيع الطبيعى المشترك مع التوزيع المشاهد فى المقاييس الوصفية .

وتسمى هذه الطريقة طريقة التوفيق بالمساحات (Fitting By Areas)

وهناك طريقة أخرى لتوفيق المنحنى الطبيعى هى طريقة التوفيق بالإحداثيات (Fitting By Ordinates) . وليس هناك ما يبرر تفضيل أى الطريقتين على الاخرى إذا أنهما تؤديان إلى نفس النتائج تقريباً كما أن العمل الحسابى يكاد يكون متعادلا فى كل منهما . وحيث أن الطريقة الأولى هى الأكثر شيوعاً فسنكتفى بتوضيحها تفصيلا . وفيما يلى خطوات العمل اللازمة لتوفيق المنحنى الطبيعى للتوزيعات التكرارية بطريقة المساحات :-

---

(*) هذه إحدى خواص المنحنى الطبيعى التى أشرنا إليها من قبل فى ص ٢٦٧ .

(١) طريقة التوفيق بالاحداثيات لا تعدو استخراج قيمة ص من معادلة المنحنى الطبيعى مباشرة ويمكن تسهيل العمل كثيراً باستعمال جدول إحداثيات المنحنى الطبيعى .

أولا : نعين الوسط الحسابى والانحراف المعيارى للتوزيع التكرارى المشاهد ونعتبر أنهما يساويان الوسط الحسابى والانحراف المعيارى للمنحنى الطبيعى المطلوب توفيقه .

ثانياً : نعين انحرافات الحد الأدنى لكل فئة عن الوسط الحسابى مع مراعاة أنه إذا لم تكن الفئات متداخلة فلا بد من تعديلها لجعلها متداخلة .

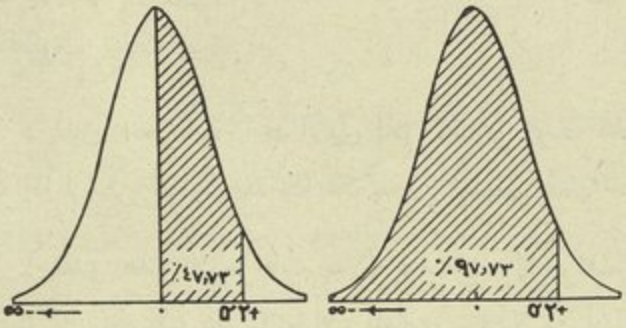
ثالثاً : نقدر هذه الانحرافات بدلالة وحدات الانحراف المعيارى وذلك بقسمة كل انحراف على الانحراف المعيارى . فبدلا من أن نقول إن الانحراف عن الوسط الحسابى = ١٢ وحدة من وحدات القياس نقول إنه = ٣ σ (على فرض أن σ = ٤) .

رابعاً : نستخرج من جدول مساحات المنحنى الطبيعى نسبة المساحة الواقعة بين كل انحراف وبين الطرف الأيسر (السالب) للمنحنى ، أو بعبارة أخرى نستخرج عدد المفردات الأقل من الحد الأدنى لكل فئة . فلو فرض أن كان انحراف الحد الأدنى للفئة الأولى قدره ٢σ نجد بالنظر إلى جدول المساحات أن المساحة المحصورة بين ٢σ ومحور التماثل = ٤٧,٧٣٪

ومن ذلك نستنتج أن المساحة المحصورة بين ٢σ والطرف الأيسر للمنحنى = ٥٠٪ + ٤٧,٧٣٪ = ٩٧,٧٣٪ من المساحة الكلية (انظر شكل رقم ٥٢)[1] .

---

(١) لاحظ أن محور التماثل يقسم المنحنى الطبيعى إلى قسمين متساويين فى المساحة أى أن المساحة الواقعة بين هذا المحور والطرف الأيسر للمنحنى تساوى ٥٠ ٪ من المساحة الكلية .



(شكل رقم ٥٢)

رسم يوضح نسبة المساحة المحصورة بين + ٥٢ و — ∞ ، + ٥٢ ومحور التماثل

وبالمثل نستخرج المساحة الواقعة بين الحد الأدنى للفئة الثانية والطرف الأيسر للمنحنى . ونستمر على هذا النحو حتى نحصل على المساحة الواقعة بين الحد الأدنى لكل فئة والطرف الأيسر للمنحنى .

خامساً: بطرح كل مساحة من المساحة التى تليها نحصل على المساحة المحصورة بين الحدد الأدنى لكل فئة والحد الادنى للفئة التى تليها . ومن هذا نستنتج نسبة عد المفردات الواقعة داخل كل فئة إلى مجموع مفردات التوزيع أو بعبارة أخرى نحصل على التكرار النسبى فى كل فئة .

سادساً: نضرب هذه التكرارات النسبية فى مجموع التكرارات فنحصل على عدد المفردات الواقعة داخل كل فئة على فرض أن قيم الظاهرة موزعة توزيعاً طبيعياً وهذه هى النتيجة التى نرغب فى الوصول إليها .

ويلاحظ أنه إذا رسمنا المنحنى التكرارى الذى يمثل التوزيع النظرى الذى نحصل عليه بهذه الطريقة نجد أنه قريب جداً من المنحنى الطبيعى . وإذا اختلف عنه قليلا فما ذلك إلا لأننا نضطر إلى جعل عدد مفردات التوزيع محدودا

والمفروض أن المضلع التكرارى لا يتخذ شكل المنحنى الطبيعى إلا إذا كان عدد مفردات التوزيع يقرب من ما لا نهاية(١).

ولإيضاح خطوات العمل السالفة الذكر نوفق منحنياً طبيعياً للتوزيع التكرارى لنسب ذكاء ٢٧١٠ تلميذاً وهو المبين بالجدول رقم ٣٩

جدول رقم ٣٩

التوزيع التكرارى لنسب ذكاء ٢٧١٠ تلميذاً (٥)

نسب الذكاء	التكرار	نسب الذكاء	التكرار
٤٠ — ٥٠	٦	١٠٠ — ١١٠	٥٩٦
٥٠ — ٦٠	١٥	١١٠ — ١٢٠	٤٠٠
٦٠ — ٧٠	٨٣	١٢٠ — ١٣٠	١٨٩
٧٠ — ٨٠	٢٢٦	١٣٠ — ١٤٠	٦٥
٨٠ — ٩٠	٤٧٥	١٤٠ — ١٥٠	١١
٩٠ — ١٠٠	٦٤٤		
		المجموع	٢٧١٠

لتوفيق المنحنى الطبيعى لهذا التوزيع نبدأ أولا بحساب كل من الوسط الحسابى والانحراف المعيارى فنجد أن $\bar{س}$ = ٩٨,٧٣٤ وσ = ١٦,٤٧ ثم نعمل جدولا كالآتى :—

(١) راجع ص ٢٥٩ من هذا الكتاب .

(٥) مدنا بهذه البيانات حضرة الأستاذ مترى أمين من بحث له عن ضعفاء العقول لتوفيق منحن طبيعى لها .

جدول رقم ٤٠

توفيق منحنى طبيعى لنسب ذكاء ٢٧١٠ تلميذاً

الحد الأدنى للفئة	انحراف الحد الأدنى عن $\bar{س}$	الانحراف بدلالة وحدات σ	النسبة المئوية للمساحة		التكرار النظرى	التكرار المشاهد
			تحت الحد الأدنى للفئة	داخل الفئة		
٤٠	— ٥٨,٧٣٤	— ٣,٥٦٧	٠,٠٢	٠,١٣	٤	٦
٥٠	— ٤٨,٧٣٤	— ٢,٩٥٩	٠,١٥	٠,٧٩	٢٢	١٥
٦٠	— ٣٨,٧٣٤	— ٢,٣٥٢	٠,٩٤	٣,٠٧	٨٣	٨٣
٧٠	— ٢٨,٧٣٤	— ١,٧٤٥	٤,٠١	٨,٧٠	٢٣٦	٢٢٦
٨٠	— ١٨,٧٣٤	— ١,١٣٧	١٢,٧١	١٧,١٠	٤٦٤	٤٧٥
٩٠	— ٨,٧٣٤	— ٠,٥٣٠٤	٢٩,٨١	٢٣,٣٨	٦٣٤	٦٤٤
١٠٠	+ ١,٢٦٦	+ ٠,٠٧٦٩	٥٣,١٩	٢١,٩٩	٥٩٦	٥٩٦
١١٠	+ ١١,٢٦٦	+ ٠,٦٨٤٢	٧٥,١٨	١٤,٩٧	٤٠٦	٤٠٠
١٢٠	+ ٢١,٢٦٦	+ ١,٢٩١	٩٠,١٥	٦,٩٨	١٨٩	١٨٩
١٣٠	+ ٣١,٢٦٦	+ ١,٩٠٠	٩٧,١٣	٢,٢٧	٦٢	٦٥
١٤٠	+ ٤١,٢٦٦	+ ٢,٥٠٦	٩٩,٤٠	٠,٥١	١٤	١١
١٥٠	+ ٥١,٢٦٦	+ ٣,١١٤	٩٩,٩١			
المجموع					٢٧١٠	٢٧١٠

بالنظر إلى الجدول السابق نجد أن التكرارات المشاهدة لا تختلف كثيراً عن التكرارات النظرية التى حصلنا عليها بتوفيق المنحنى الطبيعى الأمر الذى يؤيد أن الوضع المثالى لتوزيع نسب الذكاء هو المنحنى الطبيعى . وسنورد فيما بعد كيفية اختبار حسن مطابقة المنحنى الطبيعى وغيره من المنحنيات للظواهر المختلفة .

توفيق منحنى بواسون للتوزيعات التكرارية :

إذا كان لدينا توزيع تكرارى يمثل حدثاً نادر الوقوع غير أنه يتكرر فى

شيء من الانتظام من آن لاخر مثل حدث ميلاد ثلاثة توائم أو أكثر أو حادث خروج قطار ترام عن القضبان أو أى حدث آخر احتمال وقوعه ضئيل جداً فإن أنسب منحن لوصف التوزيع التكرارى فى وضعه المثالى هو منحنى بواسون ويمكن توفيق منحنى بواسون بالتعويض مباشرة فى معادلته عن قيمة الوسط الحسابى التى نستخرجها من التوزيع المشاهد . ويذكر القارئ أن معادلة هذا المنحنى هى :

$$ص = \frac{م^{ر}}{ر!} هـ^{-م}$$ حيث ص تمثل التكرار ، م الوسط الحسابى للتوزيع و ر عدد مرات النجاح (أى قيمة المتغير) ، و هـ أساس لوغاريتمات نابيير .

ومن الواضح أنه يمكن تعيين التكرارات النظرية بمجرد تعيين قيمة م

ونورد فيما يلى التوزيع التكرارى لحوادث خروج قطارات ترام رمل الإسكندرية عن القضبان خلال عام ١٩٤٥ ثم نوضح طريقة توفيق منحنى بواسون لهذا التوزيع :

جدول رقم ٤١

التوزيع التكرارى لحوادث خروج قطارات ترام الرمل عن القضبان خلال عام ١٩٤٥[1]

عدد الحوادث ر	التكرار المشاهد	التكرار النظرى .
٠	٣١٢	٣٠٩
١	٤٦	٥٢
٢	٦	٤
٣	١	٠,٢٥
٤	٠	٠
المجموع	٣٦٥	٣٦٥,٢٥

لتوفيق منحنى بواسون للتوزيع التكرارى الموضح بالجدول السابق نبدأ أولا بإيجاد قيمة الوسط الحسابى للتوزيع (م) فنجد أنه يساوى ٠,١٦٧

وبالتعويض عن قيمة م فى المعادلة يمكننا أن نحسب التكرارات النظرية مع العلم بأن $\frac{م^{ر}}{ر!}$ هـ$^{-م}$ تعطينا التكرارات النسبية ، ولكى نحصل على التكرارات المطلقة نضرب هذه فى مجموع التكرارات الذى سنفرض

---

(١) حصلنا على هذه البيانات بمساعدة حضرة الأستاذ محمد توفيق إسماعيل رئيس قسم الإحصاء بإدارة النقل المشترك لمنطقة الإسكندرية بإذن من حضرة صاحب العزة المدير العام .

(٢) التكرار المشاهد هو عبارة عن عدد أيام السنة التى وقع فيها العدد المقابل من الحوادث (٠، ١، ٢، ٣، ٤) أما التكرار النظرى فهو الذى نحصل عليه بتوفيق منحنى بواسون .

تساويه فى كل من التوزيع المشاهد والتوزيع النظرى وفيما يلى طريقة العمل :-

$$ص_{.} = ٣٦٥ \ (هـ^{-٠,١٦٧})$$

$$\therefore \ لو\ ص_{.} = لو\ ٣٦٥ - ٠,١٦٧\ لو\ هـ \ (*)$$

$$= ٢,٥٦٢٣ - ٠,١٦٧ \times ٠,٤٣٤٣ \quad (حيث\ لو_{١٠}\ هـ = ٠,٤٣٤٣)$$

$$= ٢,٤٨٩٨$$

$$\therefore \quad ص_{.} = ٣٠٩ \ تقريباً .$$

وبالمثل فإن $ص_{١} = ٣٦٥ \ (هـ^{-٠,١٦٧} \times ٠,١٦٧) = ٥٢$ تقريباً

$$ص_{٢} = ٣٦٥ \left\{ هـ^{-٠,١٦٧} \times \frac{(٠,١٦٧)^{٢}}{!\ ٢} \right\} = ٤ \ تقريباً$$

$$ص_{٣} = ٣٦٥ \left\{ هـ^{-٠,١٦٧} \times \frac{(٠,١٦٧)^{٣}}{!\ ٣} \right\} = ٠,٢٥ \ تقريباً$$

اختبار حسن المطابقة

بمقارنة التكرارات النظرية التى نحصل عليها بتوفيق المنحنى الطبيعى أو منحنى بواسون بالتكرارات الفعلية التى نحصل عليها من واقع المشاهدة أو القياس المباشر نجد أن هناك فروقا بين التكرارات المتناظرة . وهذه الفروق لا بد من حدوثها لأن التوزيع التكرارى المشاهد يمثل عينة محدودة فى حين أن التوزيع التكرارى النظرى يصف المجموع ، ولا يمكن أن تكون العينة مطابقة تماماً للمجموع الذى اختيرت منه ، غير أن حجم هذه الفروق قد يكون من الكبر بحيث يثبت عدم صلاحية المنحنى الطبيعى (أو غيره من المنحنيات التى وفقناها للتوزيع) لوصف المجموع الأصلى . أو إذا كان لدينا من الأسباب ما يدعو إلى الجزم بان المجموع الذى اختيرت منه العينة موزع

(*) استعملنا هنا اللوغاريتمات للأساس ١٠ نظراً لتعود الطلبة على استعمالها غير أنه من الواضح أن العمل الحسابى يسهل كثيراً إذا استعملنا لوغاريتمات Napier التى أساسها الكمية هـ .

توزيعاً طبيعياً فإن مثل هذه الفروق الكبيرة تثبت أن العينة نفسها متحيزة ولا تمثل المجموع تمثيلا دقيقاً .

وأول طريقة تتبادر إلى الذهن لاختبار حسن مطابقة التوزيع النظرى الذى حصلنا عليه بتوفيق أحد المنحنيات للتوزيع المشاهد هى أن نعين الفروق بين التكرارات المشاهدة والتكرارات النظرية المقابلة لها ثم نحسب مجموعها وكلما كان هذا المجموع كبيراً كلما دل ذلك على عدم صلاحية المنحنى الذى وفقناه لوصف تغير الظاهرة تحت البحث . ومن الواضح أنه لكى نكون فكرة صحيحة عن حجم هذه الفروق فلا بد أن نهمل إشاراتهاعند حساب المجموع وذلك لأنه إذا اكتفينا بجمع هذه الفروق جمعاً جبرياً فقد نحصل على مجموع صغير فى بعض الحالات التى تكون فيها الفروق ذاتها كبيرة جداً ولكنها متضادة فى الإشارة . غير أن مجرد جمع الفروق مع إهمال إشاراتها لا يكفى لإعطائنا مقياساً دقيقاً لاختبار حسن مطابقة المنحنى الذى وفقناه للتوزيع المشاهد . إذ من الواضح أن هذا المقياس يتوقف إلى حد كبير على عدد التكرارات . فإذا كان هذا العدد كبيراً كان مجموع الفروق المطلقة بين التكرارات المشاهدة والنظرية كبيراً حتى إذا كانت المطابقة جيدة جداً . ولكى نتغلب على هذه الصعوبة نحسب مجموع الفروق كنسبة مئوية من مجموع التكرارات .

ويسمى المقياس فى صورته هذه « النسبة المئوية لسوء المطابقة » (Percentage Misfit) ويمكن أن نعبر عنه رمزياً كالآتى :—

$$\text{النسبة المئوية لسوء المطابقة} = \frac{\Sigma |و - وَ|}{\Sigma و} \times ١٠٠$$

حيث و تمثل التكرارات المشاهدة ، وَ تمثل التكرارات النظرية . على أن هذا المقياس معرض لنقد واضح هو أنه لا يعطينا حداً فاصلا بين النسبة التى

نعتبرها كبيرة (أى دالة على عدم جودة المطابقة) وتلك التى نعتبرها صغيرة (أى دالة على جودة المطابقة) ، بل أن الأمر متروك للتقدير الشخصى للباحث .

وعلى ذلك فإن النسبة المئوية لسوء المطابقة ليست مقياساً موضوعياً يصح الاعتماد عليه وتكاد تكون فائدتها مقصورة على مقارنة حسن مطابقة منحنيين مختلفين لنفس التوزيع أو مقارنة حسن مطابقة توزيعين مختلفين لنفس المنحنى .

وعلى الرغم من هذا العيب الواضح فى هذا المقياس فقد ظل مستعملا مدة غير قصيرة فى اختبار حسن مطابقة التوزيعات النظرية للتوزيعات المشاهده حتى وضع الأستاذ كارل بيرسون طريقة أخرى أكثر دقة هى تلك المعروفة باختبار χ^2 لحسن المطابقة . والأساس الرياضى الذى بنى عليه بيرسون هذا الاختبار يخرج بنا عن نطاق هذا الكتاب(١) ولذلك سنكتفى بشرح طريقة تطبيقه مع الإشارة بإيجاز إلى الأسس النظرية التى ينبنى عليها .

واختبار χ^2 مبنى على نفس فكرة « النسبة المئوية لسوء المطابقة » . والواقع أن العلاقة بينهما تشابه إلى حد ما العلاقة بين الانحراف المتوسط والانحراف المعيارى . غير أن اختبار χ^2 يمتاز على « النسبة المئوية لسوء المطابقة » فى أنه يحدد لنا مقدار الفروق التى يمكن أن تحدث بطريق المصادفة والتى لا تدل بالتالى على سوء المطابقة ، أو بعبارة أخرى فإنه يحدد لنا مستوى موضوعياً نحكم بمقتضاه على ما إذا كانت الفروق التى نجدها بين التوزيع المشاهد والتوزيع النظرى فروقاً جوهرية تدل على عدم صلاحية المنحنى الموفق لوصف المجموع الذى انتخبت منه القيم المشاهدة ،

---

(١) يستطيع القارىء الذى يرغب فى الوقوف على الأساس الرياضى الذى ينبنى عليه اختبار χ^2 أن يرجع إلى البحث الأصلى الذى نشره بيرسون فى Philosophical Magazine عدد ٣٠٢ يوليه سنة ١٩٠٢ ص ١٥٧ – ١٧٥ .

أو فروقاً ظاهرية يمكن أن تعزى إلى عوامل المصادفة فلا تدل على وجود اختلاف أساسى بين التوزيع النظرى والتوزيع المشاهد ويتوقف هذا الاختبار على حساب قيمة χ^2 التى تساوى

$$\chi^2 = \frac{(و - وَ)^2}{وَ}$$

حيث وتمثل التكرارات المشاهدة ، وَ تمثل التكرارات النظرية .

ونذكر فيما يلى خطوات العمل اللازمة لحساب قيمة χ^2

أولا : نعين انحرافات التكرارات المشاهدة عن التكرارات النظرية بطرح الثانية من الأولى ، أى نوجد قيم (و — وَ) .

ثانياً : حيث أن جمع هذه الانحرافات جمعاً جبرياً لا يعطينا فكرة صحيحة عن حجم الفروق بين التكرارات المشاهدة والتكرارات النظرية كما سبق أن بينا فإننا نتخلص من إشارة الانحرافات بتربيعها بدلا من مجرد إهمالها كما فعلنا فى إيجاد النسبة المئوية لسوء المطابقة . وعلى ذلك تكون الخطوة الثانية هى تعيين قيم (و — وَ)^2 .

ثالثاً : نقسم مربع كل انحراف على التكرار النظرى المقابل له لبيان أهميته النسبية . وذلك لأن الانحراف قد يكون كبيراً لمجرد أن التكرار نفسه كبير . وقد يكون صغيرا لمجرد كون التكرار صغيراً . ولتوضيح أهمية هذه الخطوة افرض أننا رمينا قطعة من النقود ١٠٠٠٠ مرة وحصلنا على ٥١٠٠ وجه و ٤٩٠٠ ظهر .

انحراف التكرار المشاهد عن التكرار النظرى = ٥١٠٠ — ٥٠٠٠ = ١٠٠

ثم افرض أننا رمينا قطعة أخرى ١٠٠ مرة وحصلنا على ٨٠ وجه و ٢٠ ظهر

انحراف التكرار المشاهد عن التكرار النظرى = ٨٠ — ٥٠ = ٣٠

على الرغم من أن الانحراف الأول أكبر بكثير من الانحراف الثانى بدلالة

الوحدات المطلقة إلا أنه من الواضح أنه أصغر من الثانى نسبياً .

وعلى ذلك فإن الخطوة الثالثة هى إيجاد قيم $\frac{(و - وَ)^٢}{وَ}$

رابعاً : نجمع قيم $\frac{(و - وَ)^٢}{وَ}$ المجموع الذى نحصل عليه هو قيمة $\chi^٢$

ومن الواضح أنه كلما كانت قيمة $\chi^٢$ كبيرة كلما دل ذلك على أن الفروق بين التكرارات النظرية والمشاهدة الكبيرة . والمشكلة الآن هى معرفة قيمة $\chi^٢$ المسموح بها أى تلك التى يمكن أن تحدث بطريق المصادفة .

بمراجعة خطوات العمل السابقة الذكر نجد أن قيمة $\chi^٢$ لا تتوقف على حجم الفروق فحسب بل على عددها أيضاً ، أو بعبارة أخرى فهى تتوقف على عدد الفئات التكرارية فى التوزيع . فإذا كان عدد الفئات كبيراً فالذى نتوقعه طبعاً هو أن تكون قيمة $\chi^٢$ كبيرة

وقد وضع كارل بيرسون جداول تعطينا قيم $\chi^٢$ التى يمكن أن تحدث بطريق المصادفة حسب عدد فئات التوزيع التكرارى .[١] غير أنه يجب ملاحظة النقط الآتية عند استعمال هذه الجداول :—

أولا : يجب أن تضم الفئات التى لا تحتوى على تكرارات كافية إلى بعضها لتكوين فئات تحتوى على عدد كاف من المفردات (٥ على الأقل[٢]) والسبب فى ذلك أنه إذا كان عدد التكرارات فى الفئة صغيراً جداً ١ أو ٢ أو ٣ أو ٤ فإن دلالة الفرق بين التكرار النظرى والفعلى (المشاهد) تكون ضعيفة . فلو رمينا قطعة من النقود ٤ مرات مثلا وحصلنا على الوجه مرة

---

(١) انظر جدول قيم $\chi^٢$ فى نهاية هذا الكتاب .

(٢) ٥ هو الحد الأدنى الذى لا يجوز بأى حال من الأحوال أن يقل عنه تكرار أى فئة عند تطبيق اختبار $\chi^٢$. غير أن الأفضل ألا يقل تكرار أى فئة عن ١٠ .

واحدة فقط . فإن الفرق بين عدد مرات النجاح الفعلى والنظرى فى مثل هذه الحالة لا يمكن الاعتماد عليه فى الوصول إلى أية نتيجة لأن عدد المحاولات ليس كافياً

ثانياً : يلاحظ أن المنحنيات التى نوفقها لوصف التوزيعات التكرارية هى فى الواقع « مقيدة » ببعض خواص هذه التوزيعات فالمنحنى الطبيعى الذى وفقناه لنسب ذكاء مجموعة التلاميذ ليس أى منحن حيثما اتفق ، فقد اشترطنا أن يكون هذا المنحنى الطبيعى متفقاً مع التوزيع المشاهد فى قيم كل من الوسط الحسابى والانحراف المعيارى وعدد المفردات . ولذلك يجب أن نطرح عدد هذه القيود (أو درجات التقييد) من عدد الفئات التكرارية لنحصل على ما يسمى درجات الحرية (Degrees of Freedom) وكأننا بذلك قد أزلنا أثر هذا التقييد

متى حصلنا على عدد درجات الحرية هذه ، يمكننا فى الحال معرفة قيمة χ^2 التى يمكن أن نتوقع حدوثها بطريق المصادفة من الجداول المعدة لذلك . إذ أن هذه الجداول تعطينا قيم χ^2 المقابلة لأى عدد من درجات الحرية

ولإيضاح استعمال هذه الجداول افرض أن قيمة χ^2 التى حصلنا عليها هى ٦ وأن عدد فئات التوزيع ١١ وأن المنحنى الطبيعى الذى وفقناه للتوزيع مقيد بكل من الوسط الحسابى والانحراف المعيارى ومجموع التكرارات الخاصة بالتوزيع المشاهد . حيث أن هناك ثلاثة قيود على المنحنى الموفق فإن :—

عدد درجات الحرية = ١١ — ٣ = ٨

وبالنظر إلى الجدول نجد أن احتمال الحصول بطريق المصادفة على قيمة لـ χ^2 قدرها ٦ لدرجات حرية عددها ٨ هو بين ٥٠ ٪ و ٧٠ ٪ . وما دام هذا الاحتمال كبيراً فإن التوزيع فى هذه الحالة يعتبر مطابقاً للتوزيع الطبيعى . فلو فرض أننا حصلنا على قيمة χ^2 قدرها ٢٠ ، نجد بالرجوع إلى الجدول

أن احتمال الحصول على مثل هذه القيمة بطريق المصادفة لعدد من درجات الحرية قدره ٨ = ١٪ وهذا احتمال ضئيل نستنتج منه أن المنحنى الموفق لا يصلح لوصف التوزيع المشاهد . أما إذا حصلنا على قيمة لـ χ^2 قدرها ١٥ فبالنظر إلى الجدول نجد أن احتمال الحصول على مثل هذه القيمة بطريق المصادفة لعدد من درجات الحرية قدره ٨ = ٥٪ وهذا هو الحد الأقصى لقيمة χ^2 التى يجوز أن نعزوها إلى عوامل المصادفة . فإذا كانت قيمة χ^2 أكبر من ذلك كان احتمال الحصول عليها بطريق المصادفة أقل من ٥٪ وهو احتمال ضعيف لا يمكن أن نعتمد عليه لاعتبار المصادفة وحدها مسئولة عن ظهور مثل هذه الفروق بين التكرارات النظرية والتكرارات المشاهدة (١) .

الآن وقد كونا فكرة عامة عن اختبار χ^2 ننتقل إلى تطبيقه العملى فنختبر حسن مطابقة كل من المنحنى الطبيعى ومنحنى سواسون للتوزيعين المذكورين بالجدولين رقم ٤٠ ، ٤١ .

---

(١) يميل معظم الإحصائيين إلى اعتبار الاحتمال الذى قيمته ٥٪ الحد الفاصل بين الاحتمالات الصغيرة التى يمكن إهمالها عملياً والاحتمالات الكبيرة التى لا يمكن التغاضى عنها . فإذا كان احتمال الحصول على قيمة معينة لـ χ^2 بطريق المصادفة أقل من ٥٪ فإن هذا يدل على وجود اختلاف جوهرى بين التوزيع النظرى . وإذا كان احتمال الحصول على قيمة معينة لـ χ^2 بطريق المصادفة أكبر من ٥٪ دل ذلك على أن الاختلاف بين التوزيع المشاهد والتوزيع النظرى اختلاف ظاهرى يمكن أن يعزى إلى عوامل المصادفة .

جدول رقم ٤٢

اختبار حسن مطابقة التوزيع الطبيعى للتوزيع التكرارى لنسب الذكاء

نسب الذكاء	التكرار المشاهد و	التكرار النظرى و	و — وَ	(و — وَ)٢	(و — وَ)٢ / وَ
٤٠ — ٥٠	٦	٤	—٥	٢٥	٠,٩٦
٥٠ — ٦٠	١٥	٢٢			
٦٠ — ٧٠	٨٣	٨٣	٠٠	٠٠	٠٠
٧٠ — ٨٠	٢٢٦	٢٣٦	—١٠	١٠٠	٠,٤٢
٨٠ — ٩٠	٤٧٥	٤٦٤	١١	١٢١	٠,٢٦
٩٠ — ١٠٠	٦٤٤	٦٣٤	١٠	١٠٠	٠,١٦
١٠٠ — ١١٠	٥٩٦	٥٩٦	٠٠	٠٠	٠٠
١١٠ — ١٢٠	٤٠٠	٤٠٦	—٦	٣٦	٠,٠٩
١٢٠ — ١٣٠	١٨٩	١٨٩	٠٠	٠٠	٠٠
١٣٠ — ١٤٠	٦٥	٦٢	٣	٩	٠,١٥
١٤٠ — ١٥٠	١١	١٤	—٣	٩	٠,٦٤
المجموع ...	٢٧١٠	٢٧١٠			٢,٦٨

يتضح من الجدول أن قيمة χ^2 = ٢,٢٨

وعدد درجات الحرية = عدد فئات التوزيع — عدد درجات التقييد

= ١٠ — ٣ = ٧(*) .

(*) يلاحظ أنه نظراً لعدم وجود تكرارات كافية فى الفئة الأولى من التوزيع لجأنا إلى ضمها إلى الفئة الثانية وبذلك أصبح عدد فئات التوزيع ١٠ بدلا من ١١ ويلاحظ كذلك أن التوزيع الطبيعى مقيد بكل من الوسط الحسابى والانحراف المعيارى ومجموع التكرارات الخاصة بالتوزيع المشاهد فتكون عدد درجات التقييد ٣ .

وبالرجوع إلى جدول χ^2 نجد أن احتمال الحصول بطريق المصادفة على قيمة ل χ^2 قدرها ٢,٦٨ لعدد من درجات الحرية قدره ٧ = ٩٠٪ تقريباً ومن هذا نستنتج أن مطابقة التوزيع الطبيعى للتوزيع المشاهد جيدة جداً مما يدل على أن نسب الذكاء موزعة توزيعاً طبيعياً .

جدول رقم ٤٣

اختبار حسن مطابقة منحنى بواسون لحوادث خروج قطارات ترام الرمل عن قضبان سنة ١٩٤٥

عدد الحوادث	التكرار المشاهد و	التكرار النظرى وَ	(و — وَ)	(و — وَ)^2	$\frac{(و - وَ)^2}{وَ}$
٠	٣١٢	٣٠٩	+٣	٩	٠,٠٢٩
١	٤٦	٥٢	—٦	٣٦	٠,٦٩٢
٢ فأكثر	٧	٤,٢٥	+٢,٧٥	٧,٥٦	١,٧٨١
المجموع ...	٣٦٥	٣٦٥,٢٥			٢,٥٠٢

حيث أننا قيدنا منحنى بواسون بكل من الوسط الحسابى ومجموع تكرارات التوزيع المشاهد فتكون عدد درجات التقييد = ٢

... عدد درجات الحرية = ٣ — ٢ = ١

وبالنظر إلى الجدول نجد أن احتمال الحصول بطريق المصادفة على قيمة ل χ^2 قدرها ٢,٥٠٢ لعدد من درجات الحرية قدرة ١ = ١٢٪ تقريباً

... منحنى بواسون يوافق هذا التوزيع التكرارى .

الباب العاشر

الارتباط

المقاييس الإحصائية التى أوردناها حتى الان جميعها متعلق بوصف الظواهر ومعرفة طريقة تغيرها إما فى وضعها المشاهد أو فى وضعها المثالى .

والمرحلة الثانية فى دراستنا تنتقل بنا إلى بحث العلاقة بين الظواهر المختلفة لمعرفة ما إذا كانت هناك صلة بين تغيرها ، أى لمعرفة ما إذا كان تغير إحداها مرتبطاً بتغير الأخرى . وقد تكون علاقة التبعية بين الظواهر كاملة بحيث يكون تغير إحداها متوقفاً كلية على تغير الأخرى بمعنى أنه لو علمنا قيمة الظاهرة الأولى نستطيع أن نستنتج منها قيمة الظاهرة الثانية بدقة وفى هذه الحالة تكون العلاقة بين الظاهرتين علاقة دالية (Functional Relationship) يمكن تمثيلها بمعادلة رياضية . فمثلامساحة الدائرة تتوقف كلية على طول نصف قطرها أو بعبارة أخرى فإن المساحة دالة نصف القطر ويمكن تحديدها بدقة إذا ما عرفنا طول نصف القطر دون حاجة إلى قياسها . غير أنه فى معظم الأبحاث العلمية التى نقوم بها فى شتى فروع النشاط الفكرى تعرض لنا فروض خاصة بوجود علاقات معينة بين الظواهر المختلفة ليست بقوة العلاقات الدالية . فقد يعرض للباحث الاجتماعى أو النفسانى فرض خاص بوجود علاقة بين البطالة والإجرام . وقد يعرض للباحث فى علم الوراثة فرض خاص بوجود علاقة بين طول الأب وطول الابن أو بين صفات الأب وصفات الابن بوجه عام . وقد يعرض للطبيب فرض خاص بوجود علاقة بين الحقن بلقاح خاص واكتساب مناعة ضد مرض معين وقد تنشأ هذه الفروض إما من مجرد الاستنتاج المنطقى أو من مشاهدة تغير ظاهرتين (أو أكثر) وملاحظة أن تغير إحداها

مرتبط بتغير الأخرى . وما من شك فى أن الباحث العملى لا يمكن أن يتخذ كل فرض يعن له أساساً لتصرفاته مهما بلغت قوة اعتقاده الشخصى فى صحته . ذلك أن بعض هذه الفروض التى تبدو صحيحة ومنطقية لأول وهلة يظهر خطؤها بالاختبار الدقيق . هذا الاختبار هو الذى تهيؤه لنا دراسة الارتباط . وأهمية هذا الموضوع لا تحتاج إلى ترديد ويكفى أن نذكر أنه يعطينا معياراً نقدر به قيمة الفروض التى تعرض لنا فى دراساتنا المختلفة فإما أن نضفى عليها صفة الحقائق العلمية الثابتة أو ننفيها كمزاعم باطلة(١). وعلى ذلك يمكن أن نعتبر أن الفائدة الأولى لدارسة الارتباط هى اختبار صحة الفروض العلمية كاختبار صحة الفرض القائل بأن هناك علاقة بين البطالة والإجرام أو بين صفات الأب وصفات الابن . على أن هناك فائدة أخرى عظيمة لدراسة الارتباط ، هى أن اكتشاف ارتباط إحصائى شديد بين ظاهرتين غير متصلتين فى الظاهر قد يرشد الباحث العلمى إلى احتمال وجود علاقة حقيقية بينهما . والذى نقصده بالارتباط الإحصائى بين ظاهرتين هو هو مصاحبة القيم الكبيرة للظاهرة الأولى للقيم الكبيرة للظاهرة الثانية حيث يكون الارتباط موجباً أو طردياً أو مصاحبة القيم الكبيرة للظاهرة الأولى للقيم الصغيرة للظاهرة الثانية حيث يكون الارتباط سالباً أو عكسياً .

ومن المهم جداً أن نلاحظ أن وجود ارتباط إحصائى شديد بين ظاهرتين لا يعتبر دليلا كافياً على أن إحداهما تسبب الأخرى . فربما يكون هناك عامل أو عدة عوامل تؤثر فى الظاهرتين معاً فتجعل تغير إحداهما يبدو كنتيجة لتغير الأخرى كما أن الارتباط الإحصائى قد يكون راجعاً لعوامل المصادفة ليس غير . على أنه وإن كان الارتباط الإحصائى لا يكفى وحده لإثبات

---

(١) يذكر القارئ ذلك الجدل الخطير الذى أثير حول علاقة درجة الرطوبة بتكاثر ميكروب الكوليرا . ولا شك أن التصرف فى مثل هذه المسائل طبقاً لفروض لم تثبت صحتها بصفة قاطعة قد يعرض حياة شعب بأكمله لأخطار جسيمة .

علاقة السببية بين الظواهر إلا أن فائدته تتجلى فى تأييد هذه العلاقة أو نفيها إذا كانت هناك أدلة أخرى يعتمد عليها الباحث فى إثباتها خصوصاً إذا كان هناك رأيان متعارضان . خذ مثلا علاقة السببية بين البطالة والإجرام . عندما تعرض هذه المسألة للباحث الاجتماعى أو النفسانى فإنه يحاول استخلاص السبب والنتيجة بطريقة تشبه الآتية :—

« البطالة معناها انقطاع مورد الرزق وانقطاع مورد الرزق يؤدى إلى العجز عن إشباع الحاجات الضرورية الملحة . وهذا العجز يلحق بالإنسان ألما قد يضطره إلى الانحراف عن المسلك المستقيم فى سبيل الخلاص منه فيندفع نحو السرقة وما إلى ذلك » .

غير أن مثل هذا الاستنتاج ، مع قوة الحجة التى يستند إليها ، قد لا يجد قبولا عاماً . فقد يكون فى رأى آخر أن الألم الذى يلحق بالإنسان لعجزه عن إشباع حاجاته الضرورية يوقظ حواسه ويشحذ عزيمته فيدفعه إلى الابتكار والتجديد وليس إلى الإجرام . وهكذا تظل المسألة جدلية يعوزها التأييد على وجه أو آخر . فإذا ما وضع هذا الفرض تحت الاختبار ووجد أن هناك ارتباط إحصائى بين البطالة والإجرا م فلا شك أن الرأى الأول يكون هو الأرجح .

قياس الارتباط بين الظواهر القابلة للقياس العددى

ذكرنا أن معنى وجود ارتباط بين ظاهرتين هو أن تكون قيم هاتين الظاهرتين متمشية مع بعضها سواء فى اتجاه طردى أو عكسى ، أى أن تغير إحداهما يكون مصحوباً بتغير مناظر له فى الأخرى . ومن الواضح أنه لكى نقف على مقدار التغير فى قيمة الظاهرة فلا بد أن نتفق على أساس معين نقيس منه هذا التغير . وأفضل طريقة للتعبير عن مقدار التغير فى ظاهرة ما هى قياس القيم التى تتخذها بالنسبة لإحدى قيمها المتوسطة (الوسط الحسابى مثلا) . وبمقارنة انحرافات قيم كل من الظاهرتين عن وسطيهما

الحسابيين يمكننا الحكم على ما إذا كانت هناك علاقة أو ارتباط بين تغير الظاهرتين . غير أننا رأينا عند دراسة مقاييس التشتت أن المقارنة قد تكون أحياناً غير ممكنة أو غير مجدية إما لاختلاف وحدات القياس أو لوجود فرق كبير بين متوسطى الظاهرتين ولذلك استعملنا مقاييس التشتت النسبى بدلا من المقاييس المطلقة لغرض المقارنة . وهنا أيضاً نتخلص من اختلاف الوحدات بقسمة الانحرافات عن الوسط الحسابى على الانحراف المعيارى فتكون القيم التى نقارنها هى :—

الظاهرة الأولى $\frac{س_1 - \overline{س}}{\sigma_س}$ ، $\frac{س_2 - \overline{س}}{\sigma_س}$ ، ... ، $\frac{س_ن - \overline{س}}{\sigma_س}$

الظاهرة الثانية $\frac{ص_1 - \overline{ص}}{\sigma_ص}$ ، $\frac{ص_2 - \overline{ص}}{\sigma_ص}$ ، ... ، $\frac{ص_ن - \overline{ص}}{\sigma_ص}$

والآن نلاحظ أن انحرافات قيم س الواقعة فى طرفى التوزيع (وهى القيم الكبيرة جداً والصغيرة جداً) تكون كبيرة ، وكلما اتجهنا نحو وسط التوزيع كلما قلت قيم هذه الانحرافات . وكذلك الحال بالنسبة لانحرافات قيم ص عن وسطها الحسابى . وعلى ذلك لو كانت قيم س الكبيرة تناظر قيم ص الكبيرة أو كانت قيم س الكبيرة تناظر قيم ص الصغيرة فإن حاصل ضرب الانحرافات السينية فى الانحرافات الصادية يكون كبيراً . وبعبارة أخرى فإن حاصل ضرب الانحرافات هذا يكون كبيراً إذا كانت هناك علاقة بين تغير الظاهرتين سواء أكانت هذه العلاقة طردية أو عكسية . وكل ما

هنالك من فرق بين الحالتين هو أن حاصل ضرب الانحرافات السينية فى الصادية يكون موجباً إذا كانت العلاقة طردية وسالباً إذا كانت العلاقة عكسية . وبناء على ذلك يمكننا اتخاذ متوسط حواصل ضرب الانحرافات السينية فى الانحرافات الصادية كمقياس لدرجة الارتباط بين س ، ص فإن كان هذا المتوسط كبيراً – سواء أكان سالباً أم موجباً – كان الارتباط بين المتغيرين شديداً . وعلى هذا الأساس يكون معامل الارتباط بين س ، ص ويرمز له عادة بالحرف ر هو (*) .

$$ر = \frac{١}{ن} \left\{ \frac{س_١ - \overline{س}}{\sigma_س} \times \frac{ص_١ - \overline{ص}}{\sigma_ص} + \frac{س_٢ - \overline{س}}{\sigma_س} \times \frac{ص_٢ - \overline{ص}}{\sigma_ص} \right.$$

$$\left. + \cdots \frac{س_ن - \overline{س}}{\sigma_س} \times \frac{ص_ن - \overline{ص}}{\sigma_ص} \right\}$$

$$= \frac{١}{ن} \left\{ \frac{\Sigma (س - \overline{س})(ص - \overline{ص})}{\sigma_س \quad \sigma_ص} \right\}$$

وهذا المعامل هو أهم مقاييس الارتباط ويمكن بسهولة إثبات أن قيمته لا تتغير إلا فى حدود + ١ . فإذا كانت قيمة ر = ١ فإن العلاقة بين المتغيرين تكون عبارة عن علاقة دالية . وإذا كانت ر = صفر دل ذلك على أن الارتباط بين المتغيرين معدوم أى أن كلا منهما مستقل تماماً عن الاخر . وإذا كانت ر تساوى أية قيمة أخرى بين هذين الحدين فإنها تدل على وجود ارتباط شديد أو ضعيف بين الظاهرتين طبقاً لما إذا كانت قريبة من

(*) هذا المعامل من وضع كارل بيرسون ويسمى (Product-Moment Correlation Coefficient) أى معامل حاصل ضرب العزوم .

الواحد أو الصفر (١) .

ويمكن وضع معامل الارتباط فى صيغة أخرى تسهل العمل الحسابى نوعاً ما ، فعندنا :

$$\Sigma (س - \bar{س})(ص - \bar{ص}) = \Sigma (س ص - س\bar{ص} - ص\bar{س} + \bar{س}\,\bar{ص})$$

$$= \Sigma س ص - \bar{ص}\Sigma س - \bar{س}\Sigma ص + ن\bar{س}\,\bar{ص}$$

$$= \Sigma س ص - ن\bar{س}\,\bar{ص} - ن\bar{س}\,\bar{ص} + ن\bar{س}\,\bar{ص}$$

$$= \Sigma س ص - ن\bar{س}\,\bar{ص}$$

$$\therefore ر = \frac{\Sigma س ص - ن\bar{س}\,\bar{ص}}{ن\sigma_س \sigma_ص} = \frac{\Sigma س ص - ن\bar{س}\,\bar{ص}}{\sqrt{\frac{\Sigma س^2}{ن} - \bar{س}^2} \times \sqrt{\frac{\Sigma ص^2}{ن} - \bar{ص}^2}}$$

ولتوضيح طريقة حساب قيمة ر عملياً نحسب معامل الارتباط بين إيرادات الحكومة المصرية والرقم القياسى لأسعار الجملة خلال العشرة سنوات المنتهية فى ٣١ ديسمبر سنة سنة ١٩٤٤ .

---

(١) سيجد القارئ من متابعة قراءة هذا الباب أن دلالة قيمة ر تتوقف على عدد القيم المستعملة فى حسابها إذ يمكن أن نحصل على قيمة كبيرة لمعامل الارتباط بطريق المصادفة إذا كان عدد قيم المتغيرين صغيراً . وفى هذه الحالة فإنها لا تدل على وجود ارتباط بين المتغيرين حتى إذا كانت قريبة من الواحد الصحيح .

جدول رقم ٤٤

إيجاد معامل الارتباط بين إيرادات الحكومة المصرية والرقم القياسى لأسعار الجملة

السنة	إيرادات الحكومة س	الرقم القياسى لأسعار الجملة ص	س٢	ص٢	س ص
	مليون جنيه	٪			
١٩٣٥	٤١,١	٩٥	١٦٨٩,٢١	٩٠٢٥	٣٩٠٤,٥
١٩٣٦	٤١,٨	٨٦	١٧٤٧,٢٤	٧٣٩٦	٣٥٩٤,٨
١٩٣٧	٤٣,٧	٨٧	١٩٠٩,٦٩	٧٥٦٩	٣٨٠١,٩
١٩٣٨	٤٤,٢	٩٠	١٩٥٣,٦٤	٨١٠٠	٣٩٧٨
١٩٣٩	٤٦,١	٨٨	٢١٢٥,٢١	٧٧٤٤	٤٠٥٦,٨
١٩٤٠	٤٣,٧	١٠٠	١٩٠٩,٦٩	١٠٠٠٠	٤٣٧٠
١٩٤١	٥٦,٣	١٣٠	٣١٦٩,٦٩	١٦٩٠٠	٧٣١٩
١٩٤٢	٦٧,١	١٨١	٤٥٠٢,٤١	٣٢٧٦١	١٢١٤٥,١
١٩٤٣	٧٧,٨	٢٢٧	٦٠٥٢,٨٤	٥١٥٢٩	١٧٦٦٠,٦
١٩٤٤	٨٧,٧	٢٦٨	٧٦٩١,٢٩	٧١٨٢٤	٢٣٥٠٣,٦
المجموع ...	٥٤٩,٥	١٣٥٢	٣٢٧٥٠,٩١	٢٢٢٨٤٨	٨٤٣٣٤,٣

$$ر = \frac{\Sigma س ص - ن \bar{س}\,\bar{ص}}{ن\, ع_س\, ع_ص}$$

$$= \frac{٨٤٣٣٤,٣ - ١٠ \times ٥٤,٩٥ \times ١٣٥,٢}{\sqrt{١٠} \sqrt{\frac{٣٢٧٥٠,٩١}{١٠} - (٥٤,٩٥)^٢} \times \sqrt{\frac{٢٢٢٨٤٨}{١٠} - (١٣٥,٢)^٢}}$$

$$= ٠,٩٩$$

يظهر من هذه النتيجة أن هناك ارتباطاً شديداً بين إيرادات الحكومة المصرية ومستوى الأسعار الأمر الذى يؤيد الرأى القائل بأن الزيادة فى الإيرادات ليست زيادة حقيقية وإنما هى مجرد تضخم فى الميزانية يرجع إلى حالة التضخم النقدى التى نشأت عن الحرب .

إيجاد معامل الارتباط باستخدام وسط فرضى

يمكن تسهيل العمل الحسابى الخاص بإيجاد معامل الارتباط كثيراً باستخدام وسطين فرضيين لقيم س وقيم ص . فلو فرض أن انحرافات قيم س عن الوسط الفرضى هى $ى_1$ ، $ى_2$ ، $ى_3$ ، ..... ، $ى_ن$

وأن انحرافات قيم ص عن الوسط الفرضى هى :

$ى'_1$ ، $ى'_2$ ، $ى'_3$ ، .... ، $ى'_ن$

$$\text{فإن } ر = \frac{\Sigma ى ى' - ن\left(\frac{\Sigma ى}{ن} \cdot \frac{\Sigma ى'}{ن}\right)}{ن\, ع_س\, ع_ص}$$

ويمكن استنتاج هذه الصيغة من الصيغة الأصلية لمعامل الارتباط كالآتى :

$$\Sigma (س - \bar{س})(ص - \bar{ص}) = \Sigma (ى - \frac{\Sigma ى}{ن})(ى' - \frac{\Sigma ى'}{ن})$$

$$= \Sigma [ى ى' - (ى \cdot \frac{\Sigma ى'}{ن}) - (ى' \cdot \frac{\Sigma ى}{ن}) + (\frac{\Sigma ى}{ن} \cdot \frac{\Sigma ى'}{ن})]$$

$$= \Sigma ى ى' - (\frac{\Sigma ى'}{ن} \cdot \Sigma ى) - (\frac{\Sigma ى}{ن} \cdot \Sigma ى') + ن (\frac{\Sigma ى}{ن} \cdot \frac{\Sigma ى'}{ن})$$

$$= \Sigma ى ى' - ن (\frac{\Sigma ى}{ن} \cdot \frac{\Sigma ى'}{ن})$$

$$\therefore \quad ر = \frac{\Sigma ء ى - ن (\frac{\Sigma ء}{ن} \cdot \frac{\Sigma ى}{ن})}{ن \quad ع_س \quad ع_ص}$$

وبقسمة البسط والمقام على ن

$$\text{فإن} \quad ر = \frac{\frac{\Sigma ء ى}{ن} - (\frac{\Sigma ء}{ن} \cdot \frac{\Sigma ى}{ن})}{ع_س \quad ع_ص}$$

وفى حالة التوزيعات التكرارية فإن

$$ر = \frac{\frac{\Sigma ء و ى}{\Sigma و} - (\frac{\Sigma ء و}{\Sigma و} \cdot \frac{\Sigma ى و}{\Sigma و})}{ع_س \quad ع_ص}$$

الارتباط بين التوزيعات التكرارية :

إذا كان لدينا عدد كبير من القيم المتناظرة للمتغيرين فإن حساب معامل الارتباط بالطريقة السابقة ، أى على أساس الاحتفاظ بكل قيمتين متناظرتين على انفراد ، يتطلب عمليات حسابية مطولة ومعرضة للأخطاء . على أنه يمكن تسهيل العمل الحسابى كثيراً بوضع كل من المتغيرين فى فئات مناسبة ثم رصد البيانات فى جدول تكرارى مزدوج . ولإيضاح طريقة تركيب الجدول التكرارى المزدوج واستخدامه فى حساب معامل الارتباط افرض أننا أردنا إيجاد معامل الا تباط بين وزن الأم ووزن النتاج فى مجموعة الأبقار التى تمتلكها كلية الزراعة بجامعة فاروق الأول . بالرجوع إلى الأرقام الأولية نجد أن هذه الأوزان موزعة كما هو مبين بالجدول الآتى :—

جدول رقم ٤٥

وزن الأم ووزن النتاج فى مجموعة الأبقار التى تمتلكها كلية الزراعة

وزن الأم بالكيلوجرام	وزن النتاج بالكيلوجوام	وزن الأم بالكيلوجرام	وزن النتاج بالكيلوجرام	وزن الأم بالكيلوجرام	وزن النتاج بالكيلوجرام	وزن الأم بالكيلوجرام	وزن النتاج بالكيلوجرام
٣٧٢	٢٢	٣٧٧	٣٦	٣٦٥	٢٧	٣٦٤	١٧
٤٠٣	٢٧	٣٩٠	٢٨	٣٢٧	٢٩	٣٩٧	٢٢
٤١١	٢٧	٤٨٦	٢٨	٣٥١	٢٩	٣٦٨	٢٧
٤٣٥	٢٢	٤٧٧	٢٥	٣٥٦	٣٠	٣٥٩	٢٨
٣٨٥	٣٢	٥١٥	٣٢	٥١٤	٢٧	٤٧٥	٢٩
٤٤٤	٢٥	٤٨٢	٣٩	٣٠٨	٣٢	٣٨٢	٢٤
٤٠٠	٢١	٤٨٠	٢٨	٤٣٢	٣١	٣٧٦	٢٥
٤٨٣	٣٢	٣٥٦	٢٢	٤٣٣	٢٧	٤١٠	٢٦
٣٧٩	٢٤	٣٢٦	٣٢	٣٨٢	٢٤	٣٧٩	٢٦
٣٤٠	٢٦	٣٥٩	٢٤	٤٠٩	٢٦	٤٧٣	٢٤
٤٢٣	٢٥	٣٨٨	٢٤	٤٢٦	٢٨	٣٢٧	١٨
٣٨٥	٢٢	٤٤٥	٣٢	٤٣٤	١٩	٤١٢	٢٠
٤٦٣	٢٢	٤٢٨	٣١	٤٢١	٣٤	٤٤٢	٢٣
٤٢٧	٢٣	٤٢٥	٢٦	٤٢٠	٢٥	٣٢٩	٢١
٥٣٤	٢٨	٤٥٢	٢٥	٤٠٥	٢٣	٥١٠	٢٥
٤٣٨	٢٥	٤٤٨	٢٨	٣٦٧	٣٤	٤٢٧	٢٤
٤٣٢	٢٧	٤٨٠	٢٥	٣٢٥	٢٧	٤٩٧	٣٠
٤٤٢	٢٧	٥١٧	٢٧	٤٥٤	٢٦	٥٠٦	٢٥
٣٧٩	٢٨	٥١٧	٢٩	٤٢٠	٢٩	٣٢٧	٢٢
٣٦٤	٢٠	٤٥٩	٣٢	٤٥٧	٢٤	٣٥٩	٢٣

المصدر : قسم الإنتاج الحيوانى بكلية الزراعة بجامعة فاروق الأول .

بالنظر إلى هذه الأرقام يتضح أن وزن الأم يتغير فى مدى محصور بين ٣٠٦ ، ٥٣٤ كيلوجرام ووزن النتاج يتغير فى مدى محصور بين ١٥ ، ٤٠ كيلوجرام . وعلى ذلك نستطيع أن نقسم وزن الأم إلى ١٠ فئات طول كل منها ٢٥ كيلوجرام .

(٣٠٠ — ٣٢٥ ، ٣٢٥ — ٣٥٠ ، ٠٠٠٠٠ إلخ) ووزن النتاج إلى ٥ فئات طول كل منها ٥ كيلوجرام (١٥ — ٢٠ ، ٢٠ — ٢٥ ، ٠٠٠٠٠ إلخ) ثم نعمل جدولا مقسماً أفقياً ورأسياً كما يأتى :—

جدول رقم ٤٦

الجدول التكرارى المزدوج لوزن الأم ووزن النتاج فى مجموعة أبقار كلية الزراعة

وزن النتاج س / وزن الأم ص	١٥—٢٠	٢٠—٢٥	٢٥—٣٠	٣٠—٣٥	٣٥—٤٠	المجموع
٣٠٠ — ٣٢٥				١		١
٣٢٥ — ٣٥٠	١	٢	٣	١		٧
٣٥٠ — ٣٧٥	١	٥	٤	٢		١٢
٣٧٥ — ٤٠٠		٦	٤	١	١	١٢
٤٠٠ — ٤٢٥		٣	٧	١		١١
٤٢٥ — ٤٥٠	١	٤	٨	٣		١٦
٤٥٠ — ٤٧٥		٣	٢	١		٦
٤٧٥ — ٥٠٠			٥	٢	١	٨
٥٠٠ — ٥٢٥			٥	١		٦
٥٢٥ — ٥٥٠			١			١
المجموع ...	٣	٢٣	٣٩	١٣	٢	٨٠

لوضع الأرقام الأولية فى الجدول التكرارى المزدوج نأخذ أزواج القيم المتناظرة تباعاً ونرصد كلا منها فى المربع الخاص به فى الجدول . فمثلا نرصد الزوج الأول من الأوزان (٣٧٢ ، ٢٢) فى المربع المقابل للصف الثالث (أفقياً) والعمود الثانى (رأسياً) ، أى فى المربع المقابل للصف الذى عنوانه (٣٥٠ — ٣٧٥) والعمود الذى عنوانه (٢٠ — ٢٥) والزوج الثانى (٤٠٣ ، ٢٧) فى المربع المقابل للصف الخامس الذى عنوانه (٤٠٠ — ٤٢٥) والعمود الثالث الذى عنوانه (٢٥ — ٣٠) ونستمر على هذا النحو حتى ننتهى من رصد جميع أزواج القيم المتناظرة .

ويلاحظ أننا بهذه الطريقة نحصل على توزيعين تكراريين فى نفس الجدول الأول هو التوزيع التكرارى لقيم س (رأسياً) والثانى هو التوزيع التكرارى لقيم ص (أفقياً) ويمكن من هذين التوزيعين إيجاد الوسط الحسابى والانحراف المعيارى لكل من المتغيرين بسهولة . غير أن الفائدة الأساسية للجدول التكرارى المزدوج هى أنه يسهل لنا الحصول على حواصل ضرب كل انحرافين متناظرين فى بعضهما ثم فى التكرارات المقابلة لهما . أى أنه يسهل لنا الحصول على قيمة Σ ءءَ و .

ولإيضاح طريقة حساب معامل الارتباط من الجدول التكرارى المزدوج عملياً افرض أننا اخترنا ٢٧,٥ كوسط فرضى للسينات و ٤٣٧,٥ كوسط فرضى للصادات .

نأخذ كل مربع تباعاً ونضرب انحرافه السينى فى انحرافه الصادى فى تكراره . ويلاحظ أن حاصل الضرب هذا يكون صفراً فى كل مربع لا يحتوى على تكرارات ولذلك نهمل هذه المربعات . ويلاحظ كذلك أن الانحرافات السينية فى جميع المربعات الواقعة فى العمود المقابل للوسط الفرضى للسينات تساوى صفراً ولذلك نهمل مربعات هذا العمود . وبالمثل فإن الانحرافات الصادية الواقعة فى الصف المقابل للوسط الفرضى للصادات تساوى صفراً ولذلك نهمل جميع مربعات هذا الصف . أما المربعات الباقية فإننا نحسب حواصل

ضرب انحرافها السينى × انحرافها الصادى × تكرارها .

ولتسهيل جمع هذه الحواصل نكتب الحاصل الخاص بكل مربع فى أحد أركان هذا المربع وليكن الركن الجنوبى الغربى . والجدول الآتى يوضح طريقة العمل.

جدول رقم ٤٧

حساب معامل الارتباط بين وزن الأم ووزن النتاج فى مجموعة أبقار كلية الزراعة باستخدام الجدول التكرارى المزدوج

ى / سَ	−٢	−١	٠	١	٢	المجموع	الحواصل
−٥				١ (−٥)		١	−٥
−٤	١ (٨)	٢ (٨)	٣	١ (−٤)		٧	١٢
−٣	١ (٦)	٥ (١٥)	٤	٢ (−٦)		١٢	١٥
−٢		٦ (١٢)	٤	١ (−٢)	١ (−٤)	١٢	٦
−١		٣ (٣)	٧	١ (−١)		١١	٢
٠	١	٤	٨	٣		١٦	٠
١		٣ (−٣)	٢	١ (١)		٦	−٢
٢			٥	٢ (٤)	١ (٤)	٨	٨
٣			٥	١ (٣)		٦	٣
٤			١			١	٠
المجموع ...	٣	٢٣	٣٩	١٣	٢	٨٠	
الحواصل ...	١٤	٣٥	٠	−١٠	٠		٣٩

يتضح من الجدول أن مجموع حواصل ضرب الانحرافات السينية × الانحرافات الصادية × التكرارات = ٣٩ أى أن Σ ء$_{س}$ ء$_{ص}$ و = ٣٩

ويبقى لإيجاد معامل الارتباط أن نحسب الإنحراف المعيارى لكل من التوزيعين التكراريين لقيم س وقيم ص ، ويلاحظ أن كلا التوزيعين موضحاً بالجدول التكرارى المزدوج . وفيما يلى طريقة حساب الانحراف المعيارى لكل منهما عملياً .

جدول رقم ٤٨

إيجاد الانحراف المعيارى لوزن النتاج فى مجموعة أبقار كلية الزراعة

فئات الوزن	مراكز الفئات	التكرار و	الانحرافات عن الوسط الفرضى ء	ء و	ء^{٢} و
١٥ — ٢٠	١٧,٥	٣	٢—	٦—	١٢
٢٠ — ٢٥	٢٢,٥	٢٣	١—	٢٣—	٢٣
٢٥ — ٣٠	٢٧,٥	٣٩	٠	٠	٠
٣٠ — ٣٥	٣٢,٥	١٣	١	١٣	١٣
٣٥ — ٤٠	٣٧,٥	٢	٢	٤	٨
المجموع ...		٨٠		١٢—	٥٦

$$\sigma_{س} = \sqrt{\frac{٥٦}{٨٠} - \left(\frac{١٢-}{٨٠}\right)^{٢}}$$

= ٠,٨٢٣ بوحدات طول الفئة

جدول رقم ٤٩

إيجاد الانحراف المعيارى لوزن الأم فى مجموعة أبقار كلية الزراعة

فئات الوزن	مراكز الفئات	التكرار و	الانحرافات عن الوسط الفرضى ى	ى و	ى٢ و
٣٠٠–٣٢٥	٣١٢,٥	١	–٥	–٥	٢٥
٣٢٥–٣٥٠	٣٣٧,٥	٧	–٤	–٢٨	١١٢
٣٥٠–٣٧٥	٣٦٢,٥	١٢	–٣	–٣٦	١٠٨
٣٧٥–٤٠٠	٣٨٧,٥	١٢	–٢	–٢٤	٤٨
٤٠٠–٤٢٥	٤١٢,٥	١١	–١	–١١	١١
٤٢٥–٤٥٠	٤٣٧,٥	١٦	٠	٠	٠
٤٥٠–٤٧٥	٤٦٢,٥	٦	١	٦	٦
٤٧٥–٥٠٠	٤٨٧,٥	٨	٢	١٦	٣٢
٥٠٠–٥٢٥	٥١٢,٥	٦	٣	١٨	٥٤
٥٢٥–٥٥٠	٥٣٧,٥	١	٤	٤	١٦
المجموع		٨٠		–٦٠	٤١٢

$$\sigma_{ص} = \sqrt{\frac{٤١٢}{٨٠} - \left(-\frac{٦٠}{٨٠}\right)^{٢}}$$

= ٢,١٤١ بوحدات طول الفئة

$$\therefore ر = \frac{\left(-٠,٧٥ - \times ٠,١٥ - \right) - \frac{٣٩}{٨٠}}{٠,٨٢٣ \times ٢,١٤١}$$

$$= \frac{٠,٣٧٥}{١,٧٦٢٠٤} = ٠,٢١٣ \quad (*)$$

ومن هذا يتضح أن الارتباط بين وزن الأم ووزن النتاج فى هذه المجموعة من الأبقار ضئيل .

إيجاد معامل الارتباط بطريقة الأقطار ذات الفروق المتساوية[١].

إذا رجعنا إلى الجدول التكرارى المزدوج (رقم ٤٦) ورسمنا أقطاراً متوازية تنزل من اليمين إلى اليسار[٢] بحيث يمر كل منها بعدد من المربعات المتقابلة بالرأس فإننا نلاحظ أن الفرق بين مرتبة الفئة السينية ومرتبة الفئة الصادية لجميع المربعات الواقعة على نفس القطر يكون ثابتاً . ولذلك تسمى هذه الأقطار أقطار الفروق المتساوية فإذا اعتبرنا حواصل جمع التكرارات الموجودة داخل المربعات الواقعة على نفس القطر تكراراً لهذا الفرق الثابت بين مرتبة س ومرتبة ص يمكننا أن نحسب الانحراف المعيارى لهذه الفروق بالطريقة المعتادة ولنرمز له بالرمز $\sigma_{ن}$ بعد ذلك نحسب الانحرافين المعيارى لكل من قيم س و قيم ص ولتكن قيمتهما بوحدات طول الفئة $\sigma_{س}$ ، $\sigma_{ص}$ على التوالى .

(*) يلاحظ أنه إذا كنا نعمل بوحدات طول الفئة عند حساب معامل الارتباط كما هو الحال فى هذا المثال فليس هناك داع للرجوع إلى الوحدات الأصلية مطلقاً . وذلك لأنه عند استخراج النتيجة النهائية يكون كل من البسط والمقام مقسوماً على كمية هى عبارة عن حاصل ضرب طول الفئة السينية × طول الفئة الصادية .

(١) تسمى هذه الطريقة بالإنجليزية Diagonal Summation Method

(٢) هذه الأقطار تكون موازية للقطر الذى يمتد من الركن الشمالى الشرقى إلى الركن الجنوبى الغربى للجدول على فرض أن الفئات السينية مرتبة ترتيباً تصاعدياً من اليمين إلى اليسار وأن الفئات الصادية مرتبة ترتيباً تصاعدياً من أعلى الجدول إلى أسفله . فإذا كان ترتيب الفئات مخالفاً لذلك فإن الأقطار تبدأ من ناحية الفئات الصغيرة وتتجه نحو الفئات الكبيرة لها . انظر فى هذه النقطة كتاب Statistical Calculation By E. G. Chambers ص ٤٩ .

فإذا ما حسبنا هذه الانحرافات المعيارية الثلاثة نحسب قيمة ر من الصيغة الآتية وهى مشتقة من الصيغة الأصلية لمعامل الارتباط .

$$ر = \frac{\sigma^{2}_{س} + \sigma^{2}_{ص} - \sigma^{2}_{ف}}{2\sigma_{س}\,\sigma_{ص}}$$

ولإيجاد معامل الارتباط بهذه الطريقة نبدأ أولا بحساب تكرارات الأقطار (أى تكرارات الفروق بين مرتبات قيم س وقيم ص) وبالرجوع إلى الجدول رقم ٤٦ نجدها كالآتى :-

تكرار القطر الأول		= ١
« « الثانى		= ١
« « الثالث	٣ + ٢ + ١	= ٦
« « الرابع	٢ + ٤ + ١	= ٧
« « الخامس	١ + ٥ + ٤ + ١	= ١١
« « السادس	١ + ٦ + ٧ + ٣	= ١٧
« « السابع	٣ + ٨ + ١ + ١	= ١٣
« « الثامن	٤ + ٢ + ٢	= ٨
« « التاسع	١ + ٣ + ٥ + ١	= ١٠
« « العاشر		= ٥
« « الحادى عشر		= ١

والآن نوجد الانحراف المعيارى للفروق بين المرتبات متخذين الفرق السادس كوسط فرضى :-

جدول رقم ٥٠

إيجاد الانحراف المعيارى للفروق بين مرتبات قيم س وقيم ص

مرتبة الفرق	التكرارات و	الانحرافات عن الوسط الفرضى ء	ء و	ء٢ و
١	١	٥—	٥—	٢٥
٢	١	٤—	٤—	١٦
٣	٦	٣—	١٨—	٥٤
٤	٧	٢—	١٤—	٢٨
٥	١١	١—	١١—	١١
٦	١٧	٠	٠	٠
٧	١٣	١	١٣	١٣
٨	٨	٢	١٦	٣٢
٩	١٠	٣	٣٠	٩٠
١٠	٥	٤	٢٠	٨٠
١١	١	٥	٥	٢٥
المجموع ...	٨٠		٣٢	٣٧٤

$$\sigma^2_{ف} = \frac{٣٧٤}{٨٠} - \left(\frac{٣٢}{٨٠}\right)^2$$

$$= ٤,٥١٥$$

وحيث أن $\sigma_{س}$ = ٠,٨٢٣ ، $\sigma_{ص}$ = ٢,١٤١ بوحدات طول الفئة(١)

---

(١) يفضل البعض حساب الانحراف المعيارى لمرتبات الفئات السينية والانحراف المعيارى لمرتبات الفئات الصادية بدلا من الانحراف المعيارى للقيم ذاتها . والواقع أنه ما دمنا نعمل بوحدات طول الفئة فإن الطريقتين تؤديان إلى نفس النتيجة تماماً .

$$\therefore \text{ر} = \frac{(٠,٨٢٣)^{٢} + (٢,١٤١)^{٢} - ٤,٥١٥}{٢ \times ٠,٨٢٣ \times ٢,١٤١} = ٠,٢١٣$$

وهى نفس النتيجة التى حصلنا عليها بالطريقة السابقة .

معامل سبيرمان للارتباط بين المرتبات[١].

يحدث أحياناً أن تكون الظاهرتان المطلوب إيجاد معامل الارتباط بينهما من الظواهر التى لا يمكن قياس القيم التى تتخذها بدقة غير أنه يمكن وضعها فى شكل مرتبات مسلسلة حسب الأعداد الطبيعية ١ ، ٢ ، ٣ ، ٤ ، .... ، ن .

فإذا أردنا مثلا معرفة ما إذا كانت هنا علاقة بين الذكاء وحسن السلوك فى مجموعة من الطلبة فلا شك أنه من الأسهل والأفضل أن نضع هؤلاء الطلبة فى مرتبات حسب هاتين الصفتين بدلا من أن نحاول إعطائهم درجات مطلقة قلما تخلو من عنصر التعسف . وقد اشتق « سبيرمان » من الصيغة الأصلية لمعامل الارتباط صيغة سهلة لقياس الارتباط بين المرتبات وهى :—

$$\text{ر} = ١ - \frac{٦ \Sigma \text{ف}^{٢}}{\text{ن}(\text{ن}^{٣} - ١)} \quad (*)$$

حيث ف$_{١}$ ، ف$_{٢}$ ، ف$_{٣}$ ... ، ف$_{\text{ن}}$ تمثل الفروق بين المرتبات المتناظرة للمتغيرين . و (ن) عدد الأزواج المتناظرة للمتغيرين .

وطريقة الوصول إلى هذه الصيغة لمعامل الارتباط تتلخص فيما يأتى :—

افرض أن مرتبات المتغير الأول (س) هى :—

١ ، ٢ ، ٣ ، ٤ ، ..... ، ن

---

(١) يسمى هذا المعامل بالإنجليزية Spearman's Rank Correlation Coefficient

(*) سنرمز إلى معامل الارتباط بين المرتبات بالحرف (ر) تمييزاً له عن معامل الارتباط العادى الذى نرمز له بالحرف (ر) .

وأن المرتبات المناظرة فى المتغير الثانى (ص) هى :-

٣ ، ٤ ، ١ ، ٥ ، ٧ ، ..... ، (وهى نفس الأرقام فى وضع مختلف) .

يمكننا أن نحصل على الوسط الحسابى والانحراف المعيارى لقيم س بسهولة كما يأتى :-

$$\text{س} = \frac{\text{١ + ٢ + ٣ + ... ن}}{\text{ن}} = \frac{\frac{\text{ن (ن + ١)}}{\text{٢}}}{\text{ن}} \quad \text{(نظرية جبرية)}$$

$$= \frac{\text{ن (ن + ١)}}{\text{٢ن}} = \frac{\text{ن + ١}}{\text{٢}}$$

وحيث أن قيم ص هى نفس قيم س وإنما فى وضع مختلف فإن :

$$\overline{\text{س}} = \overline{\text{ص}}$$

$$\text{كذلك } \sigma^{٢}_{\text{س}} = \frac{\Sigma \text{س}^{٢}}{\text{ن}} - \overline{\text{س}}^{٢}$$

$$\text{ولكن } \Sigma \text{ س}^{٢} = ١^{٢} + ٢^{٢} + ٣^{٢} + ... \text{ ن}^{٢}$$

$$= \frac{\text{ن (ن + ١) (٢ن + ١)}}{٦} \quad \text{(نظرية جبرية)}$$

$$\therefore \sigma^{٢}_{\text{س}} = \frac{\text{ن (ن + ١) (٢ن + ١)}}{\text{٦ن}} - \left(\frac{\text{ن + ١}}{\text{٢}}\right)^{٢}$$

$$= \frac{\text{ن}^{٢} - ١}{١٢}$$

وبالمثل فإن $\sigma^2_{ص} = \frac{ن^2 - ١}{١٢}$

أما الكمية Σ ف^2. فنحصل عليها بدلالة مرتبات المتغيرين بالطريقة الآتية :-

$$\Sigma \text{ ف}^2 = \Sigma (\text{س} - \text{ص})^2$$

$$= \Sigma \text{ س}^2 - ٢ \Sigma \text{ س ص} + \Sigma \text{ ص}^2$$

ولكن Σ س$_ن^2 = \Sigma$ ص$^2 = \frac{ن (ن + ١) (٢ن + ١)}{٦}$

$$\therefore \Sigma \text{ ف}^2 = \frac{٢ن (ن + ١) (٢ن + ١)}{٦} - ٢ \Sigma \text{ س ص}$$

$$\therefore \Sigma \text{ س ص} = \frac{ن (ن + ١) (٢ن + ١)}{٦} - \frac{١}{٢} \Sigma \text{ ف}^2 .$$

وحيث أن $ر = \frac{\Sigma \text{ س ص} - ن \overline{\text{س}}\ \overline{\text{ص}}}{ن \ \sigma_{\text{س}} \ \sigma_{\text{ص}}}$

$\therefore$ بالتعويض عن قيم Σ س ص ، $\overline{\text{س}}$ ، $\overline{\text{ص}}$ ، $\sigma_{\text{س}}$ ، $\sigma_{\text{ص}}$ نجد أن :-

$$ر = \frac{\frac{ن(ن+١)(٢ن+١)}{٦} - \frac{١}{٢}\Sigma \text{ ف}^2 - ن\left(\frac{ن+١}{٢}\right)^2}{\frac{ن (ن^2 - ١)}{١٢}}$$

$$= \frac{\frac{\text{ن}(\text{ن}+١)(٢\text{ن}+١)}{٦} - \frac{\text{ن}(\text{ن}^{٢}+٢\text{ن}+١)}{٤} - \frac{١}{٢}\Sigma \text{ف}^{٢}}{\frac{\text{ن}(\text{ن}^{٢}-١)}{١٢}} - \frac{\text{ن}(\text{ن}^{٢}-١)}{١٢}$$

$$= ١ - \frac{٦\Sigma \text{ف}^{٢}}{\text{ن}(\text{ن}^{٢}-١)}$$

وفيما يلى بيان خطوات العمل الواجب اتباعها فى حساب معامل الارتباط بين المرتبات :-

أولا : رتب قيم المتغير الأول فى مرتبات تصاعدية حسب الأعداد الطبيعية .

ثانياً : إوجد مرتبات المتغير الثانى المناظرة لمرتبات المتغير الأول .

ثالثاً : احسب الفروق بين المرتبات المتناظرة للمتغيرين .

رابعاً : إوجد مربعات هذه الفروق ثم عين قيمة معامل الارتباط بين المرتبات من القانون .

$$\text{ر} = ١ - \frac{٦\Sigma \text{ف}^{٢}}{\text{ن}(\text{ن}^{٢}-١)}$$

ولإيضاح هذه الخطوات عملياً نحسب معامل الارتباط بين مرتبات سبعة تلاميذ فى اختبارين مختلفين كما هو موضح بالجدول الآتى :-

جدول رقم ٥١

إيجاد معامل الارتباط بين مرتبات سبعة تلاميذ فى اختبارين

المرتبة فى الاختبار الأول	المرتبة فى الاختبار الثانى	الفروق بين المرتبات المتناظرة ف	مربعات الفروق ف٢
١	١	٠	٠
٢	٢	٠	٠
٣	٤	١-	١
٤	٣	١+	١
٥	٥	٠	٠
٦	٧	١-	١
٧	٦	١+	١
المجموع			٤

$$ر = ١ - \frac{٦ \times ٤}{٧ (٤٩ - ١)} = ٠,٩٣$$

ويلاحظ أنه من الممكن استخدام معامل « سبيرمان » لتقدير درجة ارتباط المتغيرات التى يمكن قياسها بدقة والتى يمكن بالتالى إيجاد معامل الارتباط العادى بينها . أى أنه يمكن فى جميع الأحوال الاستعاضة عن معامل الارتباط العادى بمعامل « سبيرمان » على أن تستبدل قيم المتغيرين بمرتبات طبعاً . وعلى الرغم من أن معامل « سبيرمان » هو مقياس دقيق للارتباط بين المرتبات إلا أنه يعتبر مقياساً تقريبياً إذا استعمل فى مكان الارتباط العادى . وعلى وجه

العموم فإن الاختلاف بين النتيجتين اللتين نحصل عليهما باستخدام المعاملين لا تزيد عادة عن ٠,٠٢ وعلى ذلك فإن معامل « سبيرمان » يعطينا تقريباً لا بأس به لمعامل الارتباط العادى .

دلالة معامل الارتباط :

أشرنا من قبل إلى أنه قد يحدث تلازم بين قيم ظاهرتين بطريق المصادفة فتكون قيمة معامل الارتباط كبيرة دون أن يدل ذلك على وجود أية علاقة بين الظاهرتين . ومن الواضح أنه كلما كان عدد أزواج القيم التى نختبر العلاقة بينها صغيراً كلما كان احتمال وجود ارتباط شديد بينها بطريق المصادفة كبيراً وكلما كانت دلالة معامل الارتباط ضعيفة . فلو فرض أن كان لدينا زوجان فقط من قيم المتغيرين س ، ص وحسبنا معامل الارتباط بينهما فإن قيمته = ١ مهما كانت قيم المتغيرين . ولإثبات ذلك افرض أن قيم س كانت س١ ، س٢ وأن قيم ص كانت ص١ ، ص٢

$$ر = \frac{\frac{\Sigma\, س ص}{ن} - \bar{س}\,\bar{ص}}{ع_س \times ع_ص}$$

$$= \frac{\frac{1}{2}(س_1 ص_1 + س_2 ص_2) - \left(\frac{س_1 + س_2}{2}\right)\left(\frac{ص_1 + ص_2}{2}\right)}{\sqrt{\frac{س_1^2 + س_2^2}{2} - \left(\frac{س_1 + س_2}{2}\right)^2} \times \sqrt{\frac{ص_1^2 + ص_2^2}{2} - \left(\frac{ص_1 + ص_2}{2}\right)^2}}$$

(١) الذى نقصده بقولنا أن دلالة معامل الارتباط تكون ضعيفة هو أنه لا يمكن عليها فى استنتاج وجود علاقة حقيقية بين الظاهرتين .

$$= \frac{\frac{1}{4} س_1 ص_1 + \frac{1}{4} س_2 ص_2 - \frac{1}{4} س_1 ص_2 - \frac{1}{4} س_2 ص_1}{\sqrt{\frac{1}{4}(س_1 - س_2)^2} \sqrt{\frac{1}{4}(ص_1 - ص_2)^2}}$$

$$= \frac{\frac{1}{4}(س_1 ص_1 + س_2 ص_2 - س_1 ص_2 - س_2 ص_1)}{\frac{1}{2}(س_1 - س_2) \times \frac{1}{2}(ص_1 - ص_2)} = ١$$

والآن افرض أن لدينا ثلاثة أزواج من قيم س ، ص المتناظرة وافرض أن قيم س هى ١ ، ٢ ، ٣ وأن ص تتخذ نفس هذه القيم ولكن بترتيبات مختلفة . ما هى قيم معاملات الارتباط التى نحصل عليها بين قيم س وقيم ص فى كل ترتيب من هذه الترتيبات ؟ يمكن وضع جميع الترتيبات الممكنة لقيم ص المناظرة لقيم س (التى لا يتغير ترتيبها) فى الجدول الآتى :—

س	ص	س	ص	س	ص	س	ص	س	ص	س	ص
١	١	١	١	١	٢	١	٢	١	٣	١	٣
٢	٢	٢	٣	٢	٣	٢	١	٢	١	٢	٢
٣	٣	٣	٢	٣	١	٣	٢	٣	٢	٣	١

ويمكننا من الجدول السابق أن نعين معامل الارتباط الخاص بكل ترتيب فنجد أن هذه المعاملات هى كالاتى :—

١ ، $\frac{1}{2}$ ، $-\frac{1}{2}$ ، $\frac{1}{2}$ ، $-\frac{1}{2}$ ، $-$ ١ (حسب ترتيب أعمدة الجدول من اليمين إلى اليسار) وبالتأمل فى هذه النتائج نجد ما يأتى :—

فى حالة واحدة من ٦ حالات وجدنا أن قيمة ر = ١

فى حالتين » » » » » ر = $\frac{1}{2}$

فى » » » » » ر = $-\frac{1}{2}$

فى حالة واحدة » » » » » ر = $-$ ١

وعلى ذلك فإذا تغاضينا عن طريقة ترتيب القيم وتركنا ذلك لمحض المصادفة فإن احتمال الحصول على قيمة لـ ر قدرها ١ أو — ١ بطريق المصادفة = $\frac{١}{٦}$ واحتمال الحصول على قيمة لـ ر قدرها $\frac{١}{٢}$ أو — $\frac{١}{٢}$ بطريق المصادفة = $\frac{٢}{٦} = \frac{١}{٣}$ وما دام احتمال الحصول على قيمة لمعامل الارتباط قدرها ١ أو $\frac{١}{٢}$ بطريق المصادفة كبيراً ($\frac{١}{٦}$ ، $\frac{١}{٣}$ على التوالى) فلا يمكن أن يدل معامل الارتباط هذا على وجود علاقة حقيقية بين الظاهرتين أى أن دلالته تكون ضعيفة جداً . على أنه من الواضح أنه كلما زاد عدد أزواج القيم المتناظرة التى نحسب معامل الارتباط بينها كلما قل احتمال الحصول على قيمة كبيرة لمعامل الارتباط بطريق المصادفة. والواقع أن معامل ارتباط قدره ٠,٣ إذا كان عدد أزواج القيم ١٠٠ تكون دلالته فى إثبات وجود علاقة حقيقية بين ظاهرتين أقوى من معامل ارتباط قدره ٠,٦ إذا كان عدد أزواج القيم المتناظرة لا يزيد عن ١٠ ، بل إن هذا الأخير ليس له أية دلالة مطلقاً وذلك لأن احتمال حدوثه بطريق المصادفة أكبر من ٠,٠٥ . والجدول الآتى يبين احتمالات الحصول عل قيم مختلفة لمعامل الارتباط (ر) بطريق المصادفة حسب عدد أزواج القيم المتناظرة :—

جدول رقم ٥٢

دلالة معامل الارتباط

عدد أزواج القيم المتناظرة	ح = احتمال الحصول على قيمة ر المبينة بالجدول بطريق المصادفة			
	ح = ٠,١	ح = ٠,٠٥	ح = ٠,٠٢	ح = ٠,٠١
١٠	ر = ٠,٥٥	ر = ٠,٦٣	ر = ٠,٧٢	ر = ٠,٧٦
١٥	٠,٤٤	٠,٥١	٠,٥٩	٠,٦٤
٢٠	٠,٣٨	٠,٤٤	٠,٥٢	٠,٥٦
٢٧	٠,٣٢	٠,٣٨	٠,٤٥	٠,٤٩
٣٢	٠,٣٠	٠,٣٥	٠,٤١	٠,٤٥
٣٧	٠,٢٨	٠,٣٢	٠,٣٨	٠,٤٢
٤٢	٠,٢٦	٠,٣٠	٠,٣٦	٠,٣٩
٥٠	٠,٢٣	٠,٢٨	٠,٣٣	٠,٣٥
٦٠	٠,٢١	٠,٢٥	٠,٣٠	٠,٣٣
٧٠	٠,٢٠	٠,٢٣	٠,٢٨	٠,٣٠
٨٠	٠,١٨	٠,٢٢	٠,٢٦	٠,٢٨
٩٠	٠,١٧	٠,٢٠	٠,٢٤	٠,٢٧
١٠٠	٠,١٦	٠,٢٠	٠,٢٣	٠,٢٦

الارتباط بين الصفات والظواهر غير القابلة للقياس العددى

كثيراً ما تعرض لنا فى الأبحاث العلمية ، خصوصاً فى أبحاث الوراثة والطب وعلم النفس ، مسألة دراسة العلاقة بين الصفات مثل العلاقة بين لون عين الأب ولون عين الابن أو بين الحقن بلقاح واق واكتساب المناعة ضد المرض أو بين إدمان المخدرات والإجرام أو بين الصمم والبكم وضعف القوة العقلية وعلى

الرغم من أنه لا يمكن قياس هذه الصفات بدلالة أى وحدات معروفة إلا أنه يمكن البحث فى الناحية العددية منها عن طريق العد أو التكرار وليس عن طريق القياس بالوحدات . وعلى ذلك فمن الممكن استخدام الطريقة الإحصائية فى دراسة العلاقة بينها . وهناك عدة مقاييس للارتباط بين الصفات أهمها ما يأتى :—

(١) معامل الاقتران Coefficient of Association

(٢) معامل التوفيق Coefficient of Contingency

١ — معامل الاقتران

افرض أننا أردنا أن نختبر تأثير الحقن باللقاح الواقى من الكوليرا فى اكتساب الجسم مناعة ضد هذا المرض وأنه بالرجوع إلى الإحصاءات الخاصة بهذا المرض وجدنا البيانات الآتية :—

عدد الأشخاص الذين حقنوا باللقاح الواقى ولم يصابوا بالكوليرا = ١٩٢

» » » » » » وأصيبوا بالكوليرا = ٤

» » » لم يحقنوا » ولم يصابوا بالكوليرا = ١١٣

» » » » » » وأصيبوا بالكوليرا = ٣٤

مجرد النظر إلى هذه الأرقام يكفى لإعطائنا فكرة عامة عن وجود علاقة إيجابية بين الحقن باللقاح الواقى واكتساب المناعة ضد الكوليرا . وتزداد هذه العلاقة وضوحاً إذا حسبنا النسبة المئوية للإصابات بين الأشخاص الذين حقنوا باللقاح وقارناها بالنسبة المئوية للإصابات بين الأشخاص الذين لم يحقنوا .

(*) هذه البيانات واقعية مستمدة من

« The Statistics of anti-typhoid & anti-cholera inoculations. » By M. Greenwood and G. U. Yule. Proceedings of The Royal Society of Medecine, vol. 8, 1916.

وبحساب هاتين النسبتين نجدهما ٢,٠٤١ ٪ ، ٢٣,١٣ ٪ على التوالى غير أنه لكى نكون فكرة دقيقة عن قوة هذه العلاقة نضع الأرقام السابقة فى جدول كالآتى ومنه نحسب معامل الاقتران الذى وضعه الأستاذ (Yule) .

جدول رقم ٥٣

الاقتران بين الحقن باللقاح الواقى والمناعة ضد الكوليرا

	لم يصب بالكوليرا	أصيب بالكوليرا
حقن باللقاح الواقى	ا ١٩٢	ب ٤
لم يحقن باللقاح الواقى	حـ ١١٣	ء ٣٤

$$\text{معامل الاقتران} = \frac{\text{ا} \times \text{ء} - \text{ب} \times \text{حـ}}{\text{ا} \times \text{ء} + \text{ب} \times \text{حـ}}$$

$$= \frac{١٩٢ \times ٣٤ - ٤ \times ١١٣}{١٩٢ \times ٣٤ + ٤ \times ١١٣} = ٠,٨٧$$

وهذه النتيجة تدل على وجود علاقة إيجابية قوية بين الحقن باللقاح الواقى واكتساب المناعة ضد الكوليرا ، إذ من الواضح أن قيمة هذا المعامل تكون دائماً أقل من الواحد الصحيح ما دامت ب حـ > صفر . وتبلغ الواحد الصحيح تماماً إذا كانت ب حـ = صفر أى إذا كان الاقتران بين الظاهرتين كاملا . فلو فرض فى المثال السابق أن أحداً من الذين حقنوا باللقاح الواقى لم يصب بالكوليرا أى إذا كانت ب = صفر

$$\text{فإن معامل الاقتران} = \frac{\text{ا} \times \text{ى} - \text{حـ} \times \text{ب}}{\text{ا} \times \text{ى} + \text{حـ} \times \text{ب}} = ١$$

٢ – معامل التوافق

يلاحظ أن معامل الاقتران لا يمكن استخدامه إلا فى إيجاد العلاقة بين الصفات أو الظواهر التى تنقسم إلى قسمين فقط . فبمراجعة جدول الاقتران نجد أنه مقسم إلى ٤ خانات ا ، ب ، حـ ، ى طبقاً لتقسيم كل من الصفتين المقترنتين إلى قسمين ولذلك فكثيراً ما يسمى « جدول ٢ × ٢ » . فلو فرض أن كانت إحدى الصفتين (أو كلتيهما) مقسمة إلى أكثر من قسمين فلا يمكن حساب معامل الاقتران بينهما وإنما يستعاض عنه بمقياس آخر وضعه كارل بيرسون اسمه معامل التوافق .

ولإيضاح معامل التوافق وطريقة حسابه عملياً افرض أننا أردنا اختبار العلاقة بين لون الشعر ولون العين لمعرفة ما إذا كان الشعر الداكن يصحبه عادة عيون قاتمة اللون . ثم إفرض أننا قسمنا ألوان الشعر إلى أربعة أقسام (أصفر – بنى – أسود – أحمر) وألوان العين إلى ثلاثة أقسام (زرقاء – رمادية أو خضراء – بنية أو سوداء) .

نبدأ أولا بعمل جدول مقسم أفقياً ورأسياً على طريقة الجدول التكرارى المزدوج ثم نرصد المفردات المشاهدة فى هذا الجدول كالآتى :–

جدول رقم ٥٤

التوافق بين لون العين ولون الشعر فى ٦٨٠٠ من الذكور بمدينة بادن(١)

لون العين	لون الشعر				المجموع
	أصفر (س١)	بنى (س٢)	أسود (س٣)	أحمر (س٤)	
أزرق (ص١)	١٧٦٨	٨٠٧	١٨٩	٤٧	٢٨١١ (ن ص١)
رمادى أو أخضر (ص٢)	٩٤٦	١٣٨٧	٧٤٦	٥٣	٣١٣٢ (ن ص٢)
بنى أو أسود (ص٣)	١١٥	٤٣٨	٢٨٨	١٦	٨٥٧ (ن ص٣)
المجموع	٢٨٢٩ (ن س١)	٢٦٣٢ (ن س٢)	١٢٢٣ (ن س٣)	١١٦ (ن س٤)	٦٨٠٠ ن

لحساب معامل التوافق من الجدول السابق نحسب التكرارات التى نتوقعها نظرياً فى كل خانة من خانات الجدول على فرض أن الظاهرتين مستقلتان تماماً عن بعضهما . وهذه التكرارات النظرية يمكن إيجادها بسهولة كالآتى :—

إذا كان مجموع التكرارات المشاهدة فى العمود (الرأسى) الذى عنوانه س$_{ح}$ هو ن س$_{ح}$ ومجموع التكرارات المشاهدة فى الصف (الأفقى) الذى عنوانه ص$_{ر}$ هو ن ص$_{ر}$.

(١) هذه البيانات مأخوذة من كتاب «مقدمة لنظرية الإحصاء» تأليف يول وكندال نقلا عن By Ammon «Zur Anthropologie Der Badener»

فإن التكرارى النظرى الذى نتوقعه فى الخانة التى يلتقى عندها كل من العمود س٢ والصف ص١ على فرض استقلال الظاهرتين عن بعضهما هو :

$$\frac{\text{ن س}_{٢} \times \text{ن ص}_{١}}{\text{ن}}$$

أى أن التكرار النظرى فى الخانة التى يلتقى عندها العمود س٣ والصف ص٢ (أى الخانة التى تبين توافق الشعر الأسود مع العيون الرمادية أو الخضراء والتى تكرارها المشاهد ٧٤٦) .

$$= \frac{\text{ن س}_{٣} \times \text{ن ص}_{٢}}{\text{ن}} = \frac{٣١٣٢ \times ١٢٢٣}{٦٨٠٠} = ٥٦٣^{*}$$

وبنفس الكيفية نحصل على التكرارات النظرية فى جميع خانات الجدول ثم نسجلها فى جدول آخر كالآتى : —

(*) حيث أن مجموع الأشخاص ذوى الشعر الأسود فى المجموعة كلها = ١٢٢٣

∴ احتمال الحصول على شخص شعره أسود من بين أفراد المجموعة مهما كان لون عينيه $= \frac{١٢٢٣}{٦٨٠٠} =$ ١٨٪ تقريباً

فإذا اعتبرنا أن الـ ٣١٣٢ شخصاً من ذوى العيون الرمادية أو الخضراء عينة اعتباطية مأخوذة من المجموع من غير تحيز فإن نسبة ذوى الشعر الأسود الموجودين بينهم لا بد أن تكون هى نفس نسبة ذوى الشعر الأسود فى المجموع كله أى ١٨٪ ، هذا على فرض عدم وجود أية علاقة بين الشعر الأسود والعيون الرمادية أو الخضراء . وبعبارة أخرى فعلى فرض استقلال الظاهرتين تماماً عن بعضهما لا بد أن يكون ١٨٪ من ذوى العيون الرمادية أو الخضراء شعرهم أسود .

∴ عدد الأشخاص ذوى الشعر الأسود المتوقع وجودهم بطريق المصادفة بين ذوى العيون الرمادية أو الخضراء على أساس استقلال الظاهرتين

$$= \text{١٨٪ من ٣١٣٢} \quad \text{أو} \quad \frac{١٢٢٣}{٦٨٠٠} \times ٣١٣٢ = ٥٦٣$$

جدول رقم ٥٥

التكرارات النظرية لتوافق لون الشعر ولون العين على أساس استقلال الظاهرتين

لون العين	لون الشعر				المجموع
	أصفر	بنى	أسود	أحمر	
أزرق	١١٦٩	١٠٨٨	٥٠٦	٤٨	٢٨١١
رمادى أو أخضر	١٣٠٣	١٢١٢	٥٦٣	٥٣,٤	٣١٣١,٤
بنى	٣٥٧	٣٣٢	١٥٤	١٤,٦	٨٥٧,٦
المجموع	٢٨٢٩	٢٦٣٢	١٢٢٣	١١٦	٦٨٠٠,٠

وبمقارنة مجاميع الأعمدة فى هذا الجدول بمجاميع الأعمدة فى الجدول رقم ٥٤ نجد أن المجموع الكلى للتكرارات النظرية يعادل المجموع الكلى للتكرارات المشاهدة وكذلك نجد أن مجموع كل عمود فى هذا الجدول يعادل مجموع العمود المناظر له فى الجدول السابق .

بعد إيجاد التكرارات النظرية نحسب الفروق بينها وبين التكرارات المشاهدة المناظرة لها ، ثم نوجد مربعات هذه الفروق ومن ثم نستخرج قيمة χ^2 بنفس الطريقة التى أوضحناها فى الباب السابق (١).

فإذا ما استخرجنا قيمة χ^2 نحسب معامل التوافق أو جذر متوسط مربع

---

(١) راجع ص ٢٩٦ من هذا الكتاب .

التوافق[1] من القانون الآتى : —

$$\text{معامل التوافق} = \sqrt{\frac{{}^{2}x}{{}^{2}x + ن}}$$

(حيث $ن$ = المجموع الكلى للتكرارات) .

والجدول الآتى يوضح خطوات العمل المؤدية إلى حساب قيمة ${}^{2}x$.

جدول رقم ٥٦

حساب قيمة ${}^{2}x$ من جدول التوافق

رقم الخانة	و — وَ	(و — وَ)^{2}	$\frac{(و - وَ)^{2}}{وَ}$
س$_{1}$ ص$_{1}$	٥٩٩	٣٥٨٨٠١	٣٠٧
س$_{1}$ ص$_{2}$	—٣٥٧	١٢٧٤٤٩	٩٨
س$_{1}$ ص$_{3}$	—٢٤٢	٥٨٥٦٤	١٦٤
س$_{2}$ ص$_{1}$	—٢٨١	٧٨٩٦١	٧٢٠
س$_{2}$ ص$_{2}$	١٧٥	٣٠٦٢٥	٢٥
س$_{2}$ ص$_{3}$	١٠٦	١١٢٣٦	٣٤
س$_{3}$ ص$_{1}$	—٣١٧	١٠٠٤٨٩	١٩٩
س$_{3}$ ص$_{2}$	١٨٣	٣٣٤٨٩	٦٠
س$_{3}$ ص$_{3}$	١٣٤	١٧٩٥٦	١١٧
س$_{4}$ ص$_{1}$	—١	١	٠٠
س$_{4}$ ص$_{2}$	—٠,٤	٠,١٦	٠٠ (٢)
س$_{4}$ ص$_{3}$	١,٤	١,٩٦	٠٠
المجموع			١٠٧٦

(١) هذا المعامل يسمى بالإنجليزية Root Mean Square Contingency .

(٢) الأرقام ضئيلة ويمكن إهمالها .

$$\chi^2 = 1076$$

$$\therefore \text{معامل التوافق} = \sqrt{\frac{\chi^2}{\chi^2 + ن}} = \sqrt{\frac{1076}{1076 + 6800}}$$

$$= \sqrt{0{,}1366} = 0{,}37$$

ويمكن اختصار العمل الحسابى الذى يتطلبه تعيين قيمة معامل التوافق بالطريقة الآتية :—

$$\chi^2 = \Sigma \frac{(و - وَ)^2}{وَ} = \Sigma \frac{(و^2 - 2 و وَ + وَ^2)}{وَ}$$

$$= \Sigma \frac{و^2}{وَ} - 2 \Sigma و + \Sigma وَ$$

وحيث أن مجموع التكرارات المشاهدة يساوى مجموع التكرارات النظرية

$$\therefore \Sigma و = \Sigma وَ = ن$$

$$\therefore \chi^2 = \Sigma \frac{و^2}{وَ} - ن$$

فإذا رمزنا للكمية $\Sigma \frac{و^2}{وَ}$ بالرمز هـ

فإن $\chi^2 = هـ - ن$

$$\therefore \text{معامل التوافق} = \sqrt{\frac{هـ - ن}{هـ - ن + ن}} = \sqrt{\frac{هـ - ن}{هـ}}$$

ومن الواضح أن حساب الكمية هـ أسهل من حساب χ^2 .

وبحساب قيمة و٢ فى المثال السابق نجد أنها = ٧٨٧٥,٢

$$\therefore \text{معامل التوافق} = \sqrt{\frac{٧٨٧٥,٢ - ٦٨٠٠}{٧٨٧٥,٢}} = ٠,٣٧$$

وهى نفس النتيجة التى حصلنا عليها بالطريقة المطوّلة .

وتجدر الإشارة إلى عيب كبير فى معامل التوافق هو أن قيمته تتوقف على عدد خانات جدول التوافق . ولذلك يجب ذكر عدد الخانات مع النتيجة النهائية. والأرقام الآتية تبين الحد الأقصى الذى يمكن أن يتخذه معامل التوافق حسب عدد خانات الجدول(١) : —

إذا كان الجدول مكوناً من ٢ × ٢ خانة فإن أقصى قيمة لمعامل التوافق = ٠,٧٠٧

» » » » ٣ × ٣ » » » » » » = ٠,٨١٦

» » » » ٤ × ٤ » » » » » » = ٠,٨٦٦

» » » » ٥ × ٥ » » » » » » = ٠,٨٩٤

» » » » ٦ × ٦ » » » » » » = ٠,٩١٣

» » » » ٧ × ٧ » » » » » » = ٠,٩٢٦

» » » » ٨ × ٨ » » » » » » = ٠,٩٣٥

» » » » ٩ × ٩ » » » » » » = ٠,٩٤٣

» » » » ١٠ × ١٠ » » » » » » = ٠,٩٤٩

ومن الواضح أن قيمة معامل التوافق تقترب من الواحد الصحيح كلما زاد عدد خانات الجدول .

---

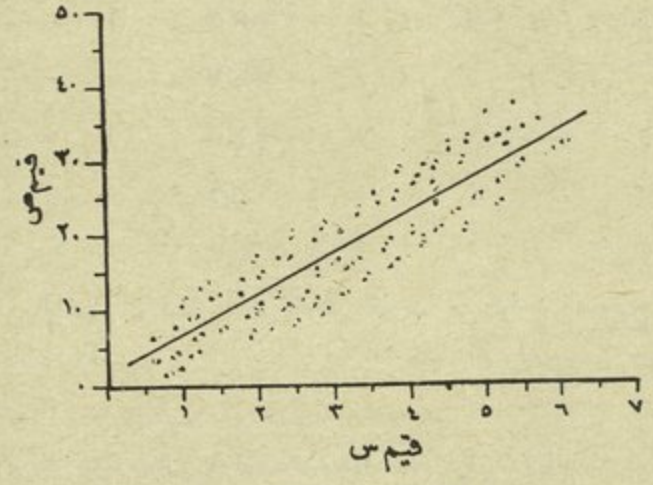
(١) يمكن معالجة مسائل الاقتران والتوافق بتطبيق اختبار χ^2 مباشرة بعد استخراج التكرارات النظرية على أساس استقلال الظاهرتين . فإذا كان الجدول مكونا من س × ص من الخانات فإن عدد درجات الحرية تكون (س − ١) (ص − ١) فمثلا فى جدول ٢ × ٢ يكون عدد درجات الحرية ١ وفى جدول ٣ × ٤ يكون عدد درجات الحرية ٢ × ٣ = ٦

الباب الحادى عشر

التوضيح البيانى للارتباط

خطوط الانحدار ومعادلاتها

افرض أن لدينا ن من أزواج القيم المتناظرة (س١ ، ص١) ، (س٢ ، ص٢) ، .... (سن ، صن) وأننا رصدنا النقط التى إحداثياتها هذه القيم على رسم بيانى . إذا كانت هناك علاقة قوية (أو ارتباط) بين قيم س وقيم ص المناظرة لها لوجدنا أن هذه النقط تنتشر حول خط منتظم إما مستقيم أو منحن على النحو المبين فى الشكل الآتى الذى يسمى شكل الانتشار .



(شكل رقم ٥٣)
شكل الانتشار فى حالة الارتباط المستقيم

ويمكن توفيق هذا الخط المنتظم وإيجاد معادلته بطريقة المربعات الدنيا التى أشرنا إليها من قبل . ومتى عرفنا معادلة المستقيم (أو المنحنى) أمكننا تعيين قيمة ص التى تقترن بأية قيمة من قيم س وبالعكس .

غير أنه فى حالات كثيرة تكون القيمة الواحدة للمتغير س مقترنة بعدة قيم للمتغير ص كما يتضح بالنظر إلى أى جدول تكرارى مزدوج مثل ذلك المبين فيما يلى :—

ص / س	س١	س٢	س ن	متوسطات س
ص١				$\bar{س}$
ص٢				$\bar{س}_{٢}$
ص ن				$\bar{س}_{ن}$
متوسطات ص	$\bar{ص}_{١}$	$\bar{ص}_{٢}$	$\bar{ص}_{ن}$	

من الواضح أن أية قيمة من قيم س (س١ مثلا) يقترن بها أكثر من قيمة واحدة من قيم ص . وبالمثل فإن أية قيمة من قيم ص (ص١ مثلا) يقترن بها أكثر من قيمة واحدة من قيم س . فإذا أردنا توضيح العلاقة بين س ، ص بيانياً من واقع المعلومات المبينة فى الجدول التكرارى المزدوج فإنه يبدو أن أقرب طريقة إلى الدقة هى أن نرصد قيمة س١ مع متوسط قيم ص المقترنة بها ، وقيمة س٢ مع متوسط قيم ص المقترنة بها وهكذا . وبعبارة أخرى فإننا نرصد على الرسم البيانى النقط التى إحداثياتها هى على التوالى

$(س_1 ، \overline{ص}_1)$ ، $(س_2 ، \overline{ص}_2)$ ، $(س_3 ، \overline{ص}_3)$ ، ... $(س_ن ، \overline{ص}_ن)$.

ومن ثم نوجد معادلة أحسن مستقيم يصف اتجاه هذه النقط بطريقة المربعات الدنيا . وسنجد أن معادلته هى[1]:-

$$ص - \overline{ص} = ر \cdot \frac{\sigma_ص}{\sigma_س} (س - \overline{س})$$

ويسمى هذا المستقيم خط انحدار ص على س كما تسمى الكمية $ر \frac{\sigma_ص}{\sigma_س}$ معامل انحدار ص على س[2].

ويمكن أن نستنتج من معادلة خط الانحدار أحسن قيمة لـ (ص) تقترن بأية قيمة نعينها للمتغير (س) .

ومن الجهة الأخرى يمكننا أن نرصد النقط التى إحداثياتها هى على التوالى $(\overline{س}_1 ، ص_1)$ ، $(\overline{س}_2 ، ص_2)$ ، $(\overline{س}_3 ، ص_3)$ ، ... $(\overline{س}_ن ، ص_ن)$ فتكون معادلة المستقيم الذى يصف إتجاه هذه النقط هى

$$س - \overline{س} = ر \frac{\sigma_س}{\sigma_ص} (ص - \overline{ص}) .$$

ويسمى هذا المستقيم خط انحدار س على ص كما تسمى الكمية $ر \frac{\sigma_س}{\sigma_ص}$

---

(١) يمكن للطالب أن يستنتج هذه المعادلة بنفسه باتباع نفس الطريقة التى شرحناها فى الباب الخاص بتوفيق المنحنيات ص ٢٧٨ والتى لا نجد فائدة من تكرارها هنا .

(٢) يسمى خط الانحدار بالإنجليزية Regression line ومعامل الانحدار Coefficient of Regression.

معامل انحدار س على ص . ويمكن أن نستنتج من معادلة هذا المستقيم أحسن قيمة لـ (س) تقترن بأية قيمة نعينها لـ (ص) .

ومن الواضح أنه إذا كان الارتباط بين س ، ص كاملا بحيث تكون ر = ١ فإن معادلتى خطى الانحدار تصبحان متماثلتين . وبعبارة أخرى فإن الخطين ينطبقان على بعضهما . وذلك لأنه إذا كانت ر = ١ فإن معادلة خط انحدار ص على س تكون :

$$ص - \overline{ص} = \frac{\sigma_{ص}}{\sigma_{س}} (س - \overline{س}) .$$

أى أن $س - \overline{س} = \frac{\sigma_{س}}{\sigma_{ص}} (ص - \overline{ص})$ وهى نفس معادلة خط انحدار س على ص .

طريقة المساحات الدنيا لإيجاد معادلتى خطى الانحدار .

على الرغم من أن طريقة المربعات الدنيا ما زالت معتبرة بوجه عام أفضل طريقة لإيجاد معادلتى خطى الانحدار إلا أن الاختبار الدقيق يظهر لنا أن هذه الطريقة تعطينا نتائج متناقضة فى جميع الحالات التى لا يكون فيها الارتباط بين الظاهرتين كاملا . ولبيان ذلك نكتب هاتين المعادلتين ثم ننتقل إلى شرح وجه النقص فيهما :-

$$ص_م = ر \frac{\sigma_{ص}}{\sigma_{س}} س \quad \cdots\cdots \quad (١)$$

$$س_م = ر \frac{\sigma_{س}}{\sigma_{ص}} ص \quad \cdots\cdots \quad (٢)$$

هاتان المعادلتان هما معادلتا خطى الانحدار على فرض أن قيم س ، ص مقيسة من وسطيهما الحسابيين .

والان افرض أن ص فى المعادلة (٢) تساوى صم فى المعادلة (١)

أى أن صم = ص = ر $\frac{\sigma_{ص}}{\sigma_{س}}$ س

وبالتعويض عن قيمة ص فى المعادلة (٢)

∴ سم = ر $\frac{\sigma_{س}}{\sigma_{ص}}$ × ر $\frac{\sigma_{ص}}{\sigma_{س}}$ س

∴ سم = ر^{٢} س

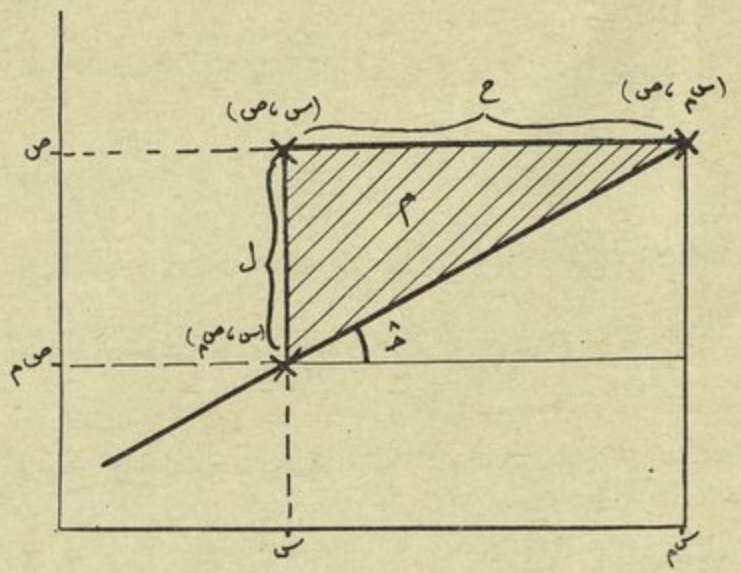
أى أن (سم) لاتساوى (س) فى حين أننا فرضنا أن (صم) = (ص)* ويلاحظ أن هذا التناقض يزول فى حالة واحدة فقط هى تلك التى تكون فيها ر = ١

أما طريقة المساحات الدنيا فتعطينا معادلتين ليس بينهما أى تناقض سواء أكانت ر = ١ أو لا تساويه .

وأساس هذه الطريقة هو أن نعتبر أن خط الانحدار هو الخط الذى يكون مجموع « مساحات الانحرافات » عنه حداً أدنى(١). والمقصود بمساحات الانحرافات تلك المثلثات الناتجة من تقاطع خط الانحدار مع إحداثيات النقط المشاهدة مثل المساحة المظللة م المبينة بالشكل رقم ٥٤

(*) يلاحظ أن (صم) هى قيمة ص النظرية أى تلك التى نحسبها من المعادلة . أما (ص) فهى القيمة المشاهدة . وكذلك الحال بالنسبة لـ (س) .

(١) لاحظ أن خط الانحدار الذى نعينه بطريقة المربعات الدنيا هو الخط الذى يجعل مجموع « مربعات الانحرافات » عنه حداً أدنى .



(شكل رقم ٥٤)

على فرض أن معادلة المستقيم الذى يحقق هذا الشرط هى كمعادلة أى مستقيم آخر صم = ا + ب س

فإن المطلوب هو تعيين قيمتى ا ، ب اللتين تجعلان مجموع « مساحات الانحرافات » حداً أدنى . ولتعيين هاتين القيمتين افرض أن س ، ص عبارة عن أى زوج من قيم المتغيرين وأن سم ، صم هما القيمتان المناظرتان المحسوبتان من المعادلة . ثم افرض أن :—

ح = س — سم (الانحراف الأفقى عن خط الانحدار)

ل = ص—صم (الانحراف الرأسى عن خط الانحدار)

$\hat{حـ}$ = زاوية تقاطع خط الانحدار مع المحور الأفقى .

م = « مساحة الانحراف » الموضحة بالشكل رقم ٥٤

$\hat{ب}$ = ميل خط الانحدار على المحور الأفقى = ظا $\hat{حـ}$

$$\text{المساحة م} = \frac{\text{ح} \times \text{ل}}{٢}$$

$$\text{ولكن ب} = \text{ظا} \hat{\text{ح}} = \frac{\text{ل}}{\text{ح}} \quad \therefore \text{ح} = \frac{\text{ل}}{\text{ب}}$$

أى أنه لأى نقطة من النقط المشاهدة مثل (س ، ص) فإن :

$$\text{مساحة الانحراف م} = \frac{\text{ل}^٢}{٢\text{ب}} \quad \cdots \quad (١)$$

وبجمع هذه المساحات لجميع القيم المشاهدة نجد أن

$$\Sigma \text{م} = \frac{\Sigma \text{ل}^٢}{٢\text{ب}} \quad \cdots \quad (٢)$$

ولكن ل = ص − صم = ص − (ا + ب س)

$$\therefore \Sigma \text{م} = \frac{\Sigma (\text{ص} - \text{ا} - \text{ب س})^٢}{٢\text{ب}} \quad \cdots \quad (٣)$$

والآن نوجد قيمتى ا ، ب اللتين تجعلان Σ م حداً أدنى . وللحصول على هاتين القيمتين نفاضل المعادلة (٣) بالنسبة لـ ا ، ب على التوالى ونساوى المشتقة الجزئية بالصفر فى الحالتين كما هو موضح فيما يأتى :—

بمفاضلة المعادلة (٣) بالنسبة لـ ا

$$\therefore \frac{\text{د} \Sigma \text{م}}{\text{د ا}} = \frac{-٢ \Sigma (\text{ص} - \text{ا} - \text{ب س})}{٢\text{ب}} = \text{صفر}$$

$$\therefore \Sigma (\text{ص} - \text{ا} - \text{ب س}) = \text{صفر}$$

أى أن Σ (ص − صم) = صفر

$$\therefore \Sigma \text{ص} = \Sigma \text{صم}$$

∴ $\Sigma ص = ن ا + ب \Sigma س$ (حيث أن $\Sigma ص_م = ن ا + ب \Sigma س$)

∴ $ا = \frac{\Sigma ص - ب \Sigma س}{ن}$ ٠ ٠ ٠ ٠ (٤)

وبمفاضلة المعادلة (٣) بالنسبة لـ ب

∴ $\frac{د \Sigma م}{د ب} = \frac{\Sigma}{٢} \left\{ \frac{-٢ب س (ص - ا - ب س) - (ص - ا - ب س)^٢}{ب^٢} \right\}$

= صفر

وبالضرب فى $- ٢$ فإن

$$\frac{\Sigma [٢ ب س (ص - ا - ب س) + (ص - ا - ب س)^٢]}{ب^٢} = \text{صفر}$$

أى أن $\Sigma [٢ ب س (ص - ا - ب س) + (ص - ا - ب س)^٢] =$ صفر

وبتكملة المربع بإضافة $ب^٢ س^٢$ ثم طرحها ثانية .

∴ $\Sigma \{ [(ص - ا - ب س) + ب س]^٢ - ب^٢ س^٢ \} =$ صفر

$= \Sigma \{ [ص - ا]^٢ - ب^٢ س^٢ \} =$ صفر

وحيث أن الكمية بين القوسين $\{ \}$ عبارة عن فرق بين مربعين

∴ هذه الكمية $= \Sigma \{ (ص - ا - ب س) (ص - ا + ب س) \} =$ صفر

∴ $\Sigma (ص - ا - ب س) \{ (ص - ا) + ب س \} =$ صفر

∴ $\Sigma (ص - ا) (ص - ا - ب س) + \Sigma ب س (ص - ا - ب س) =$ صفر

∴ $\Sigma ب س (ص - ا - ب س)$

$= - \Sigma [ص (ص - ا - ب س) - ا (ص - ا - ب س)]$

$\therefore \Sigma ب س (ص - ص_م) = - \Sigma ص (ص - ص_م) + ا \Sigma (ص - ص_م)$

وحيث أن $\Sigma (ص - ص_م) =$ صفر

$\therefore \Sigma ب س (ص - ص_م) = - \Sigma ص (ص - ص_م)$

$\therefore \Sigma ب س (ص - ا - ب س) = - \Sigma ص (ص - ا - ب س)$

$\therefore ب \Sigma س ص - ا ب \Sigma س - ب^{٢} \Sigma س^{٢} = - \Sigma ص^{٢} + ا \Sigma ص + ب \Sigma س ص$

بالتعويض عن قيمة ا بما تساويه طبقاً للمعادلة (٤)

$$\therefore ب \Sigma س ص - ب \Sigma س \left(\frac{\Sigma ص - ب \Sigma س}{ن} \right) - ب^{٢} \Sigma س^{٢}$$

$$= - \Sigma ص^{٢} + \Sigma ص \left(\frac{\Sigma ص - ب \Sigma س}{ن} \right) + ب \Sigma س ص$$

$$\therefore ب^{٢} \frac{(\Sigma س)^{٢}}{ن} - \frac{ب \Sigma س \Sigma ص}{ن} - ب^{٢} \Sigma س^{٢}$$

$$= - \Sigma ص^{٢} + \frac{(\Sigma ص)^{٢}}{ن} - \frac{ب \Sigma س \Sigma ص}{ن}$$

وبضرب الطرفين فى $- \frac{١}{ن}$

$$\therefore ب^{٢} \left[\frac{\Sigma س^{٢}}{ن} - \frac{(\Sigma س)^{٢}}{ن^{٢}} \right] = \frac{\Sigma ص^{٢}}{ن} - \frac{(\Sigma ص)^{٢}}{ن^{٢}}$$

$$\therefore ب^{٢} \sigma^{٢}_{س} = \sigma^{٢}_{ص}$$

$$\therefore ب = \pm \frac{\sigma_{ص}}{\sigma_{س}}$$

وتتوقف إشارة ب على إشارة معامل الارتباط . والآن فإن صيغة المعادلة الأصلية ص م = ا + ب س تصبح

$$\text{ص م} = \pm \frac{\sigma_{ص}}{\sigma_{س}} \text{س م} \quad \left\{ \begin{array}{l} \text{حيث كل من قيم س ، قيم ص مقيسة} \\ \text{من متوسطها وبالتالى فإن ا} = ٠ \end{array} \right.$$

وهذه هى معادلة خط انحدار ص على س

وبالمثل يمكن إثبات أن معادلة خط انحدار س على ص بطريقة المساحات الدنيا هى س م = $\pm \frac{\sigma_{س}}{\sigma_{ص}}$ ص .

ويلاحظ أنه ليس بين هاتين المعادلتين أى تناقض سواء أكانت ر = ١ أم لا . ولهذا السبب فإن طريقة المساحات الدنيا تعتبر بلاشك أفضل من طريقة المربعات الدنيا .

خطوط الانحدار المنحنية ونسبة الارتباط .

إذا كان خطا انحدار المتغيرين غير مستقيمين فلا يمكن قياس العلاقة بينهما بمعامل الارتباط العادى . فمن الممكن أن يكون هناك علاقة حقيقية بين الظاهرتين ومع ذلك يكون معامل الارتباط العادى من الضآلة بحيث لا يظهر هذه العلاقة . فلو فرض مثلا أن اتخذ خطا الانحدار شكلا نصف دائرى متماثل ، فإن معامل الارتباط العادى يساوى صفراً على الرغم من وجود علاقة حقيقية بين الظاهرتين . وفى مثل هذه الحالات — أى حالات خطوط الانحدار المنحنية — لا بد أن نستعيض عن معامل الارتباط العادى بمقياس آخر يسمى نسبة الارتباط (Correlation Ratio) وهناك نسبتان للارتباط (ى س ، ى ص) لكل جدول تكرارى مزدوج ويمكن حسابهما كالآتى : —

$$ى^{٢}_{س ص} = \frac{\sigma^{٢}_{\bar{ص}_{س}}}{\sigma^{٢}_{ص}}$$

$$،ى^{٢}_{ص س} = \frac{\sigma^{٢}_{\bar{س}_{ص}}}{\sigma^{٢}_{س}}$$

حيث $\sigma^{٢}_{\bar{ص}_{س}}$ تمثل مربع الانحراف المعيارى لمتوسطات قيم ص فى كل عمود (رأسى) من أعمدة الجدول أو بعبارة أخرى فهى تمثل مربع الانحراف المعيارى للمتوسطات الصادية التى يقابل كل منها فئة من الفئات السينية .

، $\sigma^{٢}_{ص}$ تمثل مربع الانحراف المعيارى العادى لجميع قيم ص فى الجدول

وبالمثل فإن $\sigma^{٢}_{\bar{س}_{ص}}$ هى عبارة عن مربع الانحراف المعيارى لمتوسطات قيم س فى كل صف (أفقى) من الجدول أى مربع الانحراف المعيارى للمتوسطات السينية التى يقابل كل منها فئة من الفئات الصادية .

، $\sigma^{٢}_{س}$ هى عبارة عن مربع الانحراف المعيارى العادى لجميع قيم س فى الجدول . وخطوات العمل اللازمة لحساب نسبة الارتباط الأولى هى :-

أولا : نحسب متوسط قيم ص لكل فئة من الفئات السينية أى متوسط قيم ص لكل عمود .

ثانياً : نحسب الوسط الحسابى العام لجميع قيم ص فى الجدول ، ويلاحظ أنه يساوى الوسط الحسابى لمتوسطات الأعمدة .

ثالثاً : نحسب مربع الانحراف المعيارى لمتوسطات الأعمدة حول وسطها الحسابى ، أى نستخرج قيمة $\sigma^{٢}_{\bar{ص}_{س}}$.

رابعاً : نحسب مربع الانحراف لجميع قيم ص من واقع التوزيع التكرارى للصادات المبين بالجدول ، أى نستخرج قيمة $\sigma^{٢}_{ص}$

ولإيضاح هذه الخطوات عملياً نحسب نسبة الارتباط (ىص س) للجدول التكرارى المزدوج الذى يوضح العلاقة بين وزن الأم ووزن النتاج فى مجموعة أبقار كلية الزراعة .

جدول رقم ٥٧

الجدول التكرارى المزدوج لوزن الأم (ص) ووزن النتاج (س) فى مجموعة أبقار كلية الزراعة

ص \ س	١٥ – ٢٠	٢٠–٢٥	٢٥–٣٠	٣٠–٣٥	٣٥–٤٠	المجموع
٣٠٠ – ٣٢٥				١		١
٣٢٥ – ٣٥٠	١	٢	٣	١		٧
٣٥٠ – ٣٧٥	١	٥	٤	٢		١٢
٣٧٥ – ٤٠٠		٦	٤	١	١	١٢
٤٠٠ – ٤٢٥		٣	٧	١		١١
٤٢٥ – ٤٥٠	١	٤	٨	٣		١٦
٤٥٠ – ٤٧٥		٣	٢	١		٦
٤٧٥ – ٥٠٠			٥	٢	١	٨
٥٠٠ – ٥٢٥			٥	١		٦
٥٢٥ – ٥٥٠			١			١
المجموع ...	٣	٢٣	٣٩	١٣	٢	٨٠
متوسطات الأعمدة ص س	٣٧٩,١٦	٣٩٩,٤٦	٤٣٢,٣٧	٤١٨,٢٧	٤٣٧,٥٠	

جدول رقم ٥٨

حساب الوسط الحسابى والانحراف المعيارى لقيم ص

فئات الوزن	مراكز الفئات	التكرار و	ء	ء و	ء٢ و
٣٠٠ــ٣٢٥	٣١٢,٥	١	ــ٥	ــ٥	٢٥
٣٢٥ــ٣٥٠	٣٣٧,٥	٧	ــ٤	ــ٢٨	١١٢
٣٥٠ــ٣٧٥	٣٦٢,٥	١٢	ــ٣	ــ٣٦	١٠٨
٣٧٥ــ٤٠٠	٣٨٧,٥	١٢	ــ٢	ــ٢٤	٤٨
٤٠٠ــ٤٢٥	٤١٢,٥	١١	ــ١	ــ١١	١١
٤٢٥ــ٤٥٠	٤٣٧,٥	١٦	٠	٠	٠
٤٥٠ــ٤٧٥	٤٦٢,٥	٦	١	٦	٦
٤٧٥ــ٥٠٠	٤٨٧,٥	٨	٢	١٦	٣٢
٥٠٠ــ٥٢٥	٥١٢,٥	٦	٣	١٨	٥٤
٥٢٥ــ٥٥٠	٥٣٧,٥	١	٤	٤	١٦
المجموع		٨٠		ــ٦٠	٤١٢

الوسط الحسابى العام للصادات ($\overline{ص}$) = ٤٣٧,٥ + $\frac{٢٥ \times ٦٠-}{٨٠}$

= ٤١٨,٧٥

$$= \sigma^2_{س} = \frac{٤١٢}{٨٠} - \left(\frac{٦٠-}{٨٠}\right)^2$$

= ٤,٥٨٧٥ بوحدات طول الفئة

= ٤,٥٨٧٥ × ٢٥٢ = ٢٨٦٧,١٨٧٥ بالوحدات الأصلية

أما قيمة $\sigma^2_{\bar{ص}_س}$ فنحصل عليها من الجدول الآتى :-

جدول رقم ٥٩

حساب الانحراف المعيارى للمتوسط

$\bar{ص}_س$	و س	$\bar{ص}_س - \bar{ص}$	$(\bar{ص}_س - \bar{ص})^2$	و $(\bar{ص}_س - \bar{ص})^2$
٣٧٩,١٦	٣	-٣٩,٥٩	١٥٦٧,٣٦٨١	٤٧٠٢,١٠٤٣
٣٩٩,٤٦	٢٣	-١٩,٢٩	٣٧٢,١٠٤١	٨٥٥٨,٣٩٤٣
٤٣٢,٣٧	٣٩	+١٣,٦٢	١٨٥,٥٠٤٤	٧٢٣٤,٦٧١٦
٤١٨,٢٧	١٣	-٠,٤٨	٠,٢٣٠٤	٢,٩٩٥٢
٤٣٧,٥٠	٢	+١٨,٧٥	٣٥١,٥٦٢٥	٧٠٣,١٢٥٠
المجموع	٨٠			٢١٢٠١,٢٩٠٤

$$\sigma^2_{\bar{ص}_س} = \frac{٢١٢٠١,٢٩٠٤}{٨٠} = ٢٦٥,٠١٦١٣$$

$$\therefore ى^2_{س\bar{ص}} = \frac{٢٦٥,٠١٦١٣}{٢٨٦٧,١٨٧٥} = ٠,٠٩٢٤$$

$$\therefore ى_{س\bar{ص}} = ٠,٣٠٤$$

وبنفس الطريقة يمكن حساب قيمة ى س ص .

وبمقارنة هذه النتيجة بمعامل الإرتباط لنفس الجدول التكرارى وقدره ٠,٢١٣ نجد أن نسبة الارتباط أكبر . والسبب فى ذلك أن خط انحدار ص

على س فى المثال السابق ليس مستقيماً وبالتالى فإن نسبة الارتباط أفضل من معامل الارتباط فى وصف العلاقة بين هذين المتغيرين .

والواقع أن نسبة الارتباط تفضل معامل الارتباط بوجه عام . فلو فرض مثلا أن حسبنا نسبة الارتباط فى حالة خطوط الانحدار المستقيمة فإننا نحصل على قيمة لها تعادل قيمة معامل الارتباط . ومن هذا يتضح أنها مقياس أعم من معامل الارتباط لأنها تصلح لقياس العلاقة بين الظواهر سواء أكانت خطوط الانحدار مستقيمة أم منحنية .

الباب الثانى عشر

نظرية العينات

من أهم المسائل التى تواجه الباحث العلمى عند شروعه فى القيام بأى بحث هو تحديد نطاق العمل . وعلى الرغم من أن لكل بحث ظروفه الخاصة التى تحدد نطاق العمل اللازم له والمجهود والمال الممكن بذلهما فى سبيل إنجازه إلا أنه من الواضح أنه كلما زاد عدد المفردات التى يشملها البحث أيا كان نوعه كلما أصبحت النتائج التى يتوصل إليها الباحث مستندة إلى أساس أقوى . غير أن مشقة العمل من جهة وإمكان الحصول على نتائج لا بأس بها بإجراء التجارب على عدد محدود من المفردات من الجهة الأخرى هو الذى جعل البحوث العلمية تقوم على أساس دراسة عينات محدودة مختارة من المجموعات التى نرغب فى التعرف على خواصها بدلا من دراسة المجموعات الأصلية ذاتها . والواقع أنه فى معظم الحالات يكاد يكون من المستحيل عملياً وصف المجموعات الكبيرة إلا عن طريق دراسة عينات تمثلها . على أنه من الواضح أن محاولة التعرف على خواص المجموع عن طريق دراسة عينة مختارة منه تنطوى على بعض التضحية فى دقة النتائج التى نحصل عليها . ويتوقف مقدار هذه التضحية إلى حد كبير على طريقة اختيار العينة التى سنجعلها موضوع البحث . وهناك ثلاثة أنواع من العينات يستطيع الباحث أن يختار أحدها تبعاً لظروف البحث الذى يجريه وهى :—

أولا : العينة الاعتباطية ويسميها البعض العينة العشوائية RANDOM SAMPLE

ثانياً : العينة المقصودة PURPOSIVE SAMPLE

ثالثاً : العينة الطبقية STRATIFIED SAMPLE

وفيما يلى تعريف وحالات استعمال كل منها .

أولا : العينة الاعتباطية :

على الرغم من أن اسم العينة الاعتباطية يحمل على الظن بأن مفرداتها تختار اعتباطاً دون ما قاعدة إلا أن الواقع خلاف ذلك . فالعينة الاعتباطية هى تلك التى تختار من المجموع بطريقة تعطى كل مفردة من مفرداته نفس الفرصة فى الظهور . فإذا أردنا مثلا اختيار عينة اعتباطية عدد مفرداتها ١٠٠ من بين عمال مصنع معين يضم ٥٠٠٠ عامل ، فلكى نضمن إعطاء جميع العمال نفس فرصة الاختيار ، نعطى كل عامل رقماً معيناً ونسجل هذه الأرقام على بطاقات متساوية الحجم تماماً ثم نحكم خلطها جيداً ونختار منها المائة المطلوبة . ومما تجب ملاحظته أنه على الرغم من أن الغرض من إعطاء كل مفردة نفس الفرصة فى الظهور هو عدم التحيز نحو أى صفة من صفات المجموع دون الأخرى ، إلا أنه ليس هناك ما يضمن أن تأتى نتيجة الاختيار محققة لهذا الغرض . وبعبارة أخرى فإن إعطاء كل مفردة نفس الفرصة فى الظهور لا يعتبر فى حد ذاته ضماناً كافياً لأن تكون العينة خالية تماماً من التحيز وممثلة للمجموع تمثيلا دقيقاً . فمن الممكن أن نختار عينة اعتباطية ثم نجد جميع مفرداتها شاذة لا تمثل المجموع ولو أن احتمال حدوث مثل هذا الشذود يتضاءل جداً إذا كان عدد مفردات العينة كبيراً . والواقع أن دقة تمثيل العينة الاعتباطية للمجموع الذى اختيرت منه تتوقف على عدد مفردات العينة ، فكلما كان هذا العدد كبيراً زادت دقة تمثيل العينة للمجموع . ولهذا السبب يجب أن يقتصر استعمال العينة الاعتباطية على الحالات التى تسمح فيها ظروف البحث باختيار عينات كبيرة حيث تكون عوامل المصادفة كفيلة بتحقيق الاختيار المحايد ما دامت لكل مفردة من مفردات المجموع

نفس الفرصة فى الظهور . وسنعود إلى دراسة هذه النقطة بتفصيل أكثر فى القسم التالى من هذا الباب الخاص بحجم العينة . ونظراً لأن العينة الاعتباطية هى أهم أنواع العينات نورد فيما يلى أمثلة توضح بعض الطرق المتبعة فى اختيارها . لكى نختار عينة اعتباطية من بالة من القطن لغرض اختبار طول تيلة الشعيرات مثلا ، نتبع الطريقة الآتية :ـ

نأخذ من أجزاء مختلفة من البالة كتلا كبيرة نوعاً من شعر القطن ثم نقسم كل كتلة منها إلى نصفين نحتفظ بأحدهما ونستبعد الآخر . بعد ذلك نقسم النصف الذى احتفظنا به إلى نصفين آخرين ونستبعد أحدهما . ونستمر فى عملية التقسيم إلى نصفين واستبعاد أحدهما حتى لا يتبقى لدينا من كل كتلة اخترناها فى بادئ الأمر سوى عدد قليل من الشعيرات . فإذا خلطنا الشعيرات القليلة المتبقية من كل كتلة فإننا نحصل على عينة اعتباطية خالية تماماً من التحيز(١).

وثمة طريقة أخرى لاختيار العينات الاعتباطية تصلح للتطبيق فى حالة كون المجموع الذى نختاره منه كبيراً ولكنه محدود . ولإيضاح هذه الطريقة افرض أننا أردنا اختيار عدد من المنازل الموجودة بإحدى المدن الكبيرة بطريقة اعتباطية . يمكننا أن نقسم هذه المدينة إلى عدد من الأحياء ونعطى كلا منها رقماً خاصاً نسجله على بطاقة . ثم نقسم كل حى إلى عدد من الشوارع ونعطى كلا منها رقماً خاصاً كذلك . ثم نعطى كل منزل من المنازل الموجودة بكل شارع رقماً . فإذا ما انتهينا من عملية الترقيم هذه نتناول البطاقات المسجل عليها أرقام الأحياء ونحكم خلطها ثم نختار منها واحدة اعتباطاً . وبنفس الطريقة نختار أحد شوارع الحى ثم أحد المنازل الموجودة به . ونكرر هذه العملية حتى نحصل على العدد المطلوب من المنازل .

---

(١) انظر كتاب The Methods of Statistics, By L.H.C. Tippet, ص ٦٨.

ثانياً : العينة المقصودة :

إذا كانت ظروف البحث لا تسمح إلا باختيار عدد قليل من المفردات فإن صفة تمثيل المجموع تمثيلا دقيقاً لا تتوافر فى العينة الاعتباطية كما أوضحنا . والأفضل أن نختار هذه المفردات القلائل بطريقة يراعى فيها أن تكون جميع المفردات المختارة قريبة من المتوسط كى تتوافر فيها صفة التمثيل المطلوبة . وتسمى العينة المختارة بهذه الكيفية عينة مقصودة . وأهم عيوب العينة المقصودة أنها لا تختار بطريقة موضوعية تستبعد أثر العنصر الشخصى من الاختيار كما هو الحال فى العينة الاعتباطية ولذلك فمن السهل أن تأتى متحيزة نحو صفة ما من صفات المجموع . وبالإضافة إلى ذلك فإن العينة المقصودة لا تظهر درجة تشتت المفردات حول متوسطها نظراً لأن جميع مفرداتها تختار بالقرب من هذا المتوسط كما ذكرنا . ولهذه الأسباب يجب ألا نلجأ إلى استعمالها إلا إذا اضطرتنا ظروف البحث إلى اختيار عدد قليل جداً من المفردات .

ثالثاً : العينة الطبقية :

. فى كثير من الحالات يكون المجموع الذى نختار منه إما منقسماً بطبيعته إلى طبقات أو يمكن بسهولة تقسيمه إلى طبقات يتصف كل منها بصفة خاصة تميزها عن غيرها . فإذا كنا نحرص على إبراز هذه الصفات جميعاً فى العينة فإننا نعتبر كل طبقة من طبقات المجموع مجموعاً مستقلا ونختار من كل منها عينة اعتباطية . فإذا جمعنا هذه العينات معاً لنكون من بينها عينة واحدة فإننا نحصل على ما يسمى عينة طبقية . ويلاحظ أن طريقة العينة الطبقية تعتبر وسطاً بين الطريقتين السابقتين فهى تجمع بعض مزايا كل منهما ، ففيها عنصر الاختيار المتعمد الذى يضمن عدم خلو العينة من بعض الصفات الهامة الموجودة فى المجموع وفيها كذلك عنصر الاختيار المحايد الذى يقلل احتمال التحيز نحو مفردة أو أخرى داخل

كل طبقة . ولتوضيح طريقة اختيار العينة الطبقية ، افرض مثلا أننا أردنا اختيار عينة طبقية من الملاك الزراعيين بالقطر المصرى . للحصول على هذه العينة نبدأ بتقسيم الملاك إلى طبقات حسب المساحة التى يملكونها (أقل من فدان ، ١ — ٥ ، ٥ — ١٠ ، ٠٠٠٠) ثم نختار من كل طبقة عينة اعتباطية . وبذلك نضمن أن تكون العينة الطبقية المكونة من هذه العينات شاملة لكافة طبقات الملاك الزراعيين . ولا شك أن العينة الطبقية تفضل العينات الأخرى فى تلك الحالات التى يكون المجموع فيها منقسما بطبيعته إلى أقسام مميزة عن بعضها ويمثل كل قسم منها وحدة متجانسة المفردات تقريباً . أما إذا كان المجموع كله متجانساً أو قريباً من التجانس فلا محل لاستخدام العينة الطبقية .

حجم العينة :

ذكرنا أن مشقة العمل تضطرنا دائماً إلى الاستعاضة عن دراسة المجموعات الكبيرة بدراسة عينات مختارة منها . غير أن مشقة العمل يجب ألا تتخذ عذراً لجعل العينات من الصغر بحيث لا تتوافر فيها صفة تمثيل المجموع بدرجة معقولة من الدقة . إذا أنه بدون توافر هذا الشرط يصبح البحث الذى نجريه لا قيمة له مطلقاً بل ربما أصبح مضللا وخطيراً إذا بنينا آراءنا وتصرفاتنا عليه . فمن الخطأ مثلا أن نعلن نظافة قطر يضم عشرين مليوناً من السكان من وباء حل به على أساس فحص عشرات أو حتى مئات من الأفراد والتثبت من خلوهم من هذا الوباء . وبالمثل فلا يجوز أن نقول إن متوسط أجر العامل فى مدينة كبيرة كمدينة الإسكندرية خمسة جنيهات فى الشهر إذا كان عدد العمال الذين سجلنا أجورهم ١٠ عمال أو ٥٠ عاملا أو حتى ٥٠٠ عامل . وإنما تكون نتائجنا أقرب إلى الدقة إذا كان بحثنا شاملا لخمسة أو عشرة آلاف أو خمسين ألف عامل . من الواضح أن درجة

الدقة تزداد كلما زاد عدد المفردات التى يشملها البحث أو بعبارة أخرى كلما زاد حجم العينة . وسنجد فيما يلى أن هناك علاقة رياضية محددة بين حجم العينة ودقة النتائج التى نتوصل إليها من دراسة هذه العينة . ولإيضاح ذلك افرض أن لدينا مجموعة كبيرة من المفردات — عمال القطر المصرى مثلا — وأنه لكى نصف أجور هذه المجموعة عن طريق حساب الوسط الحسابى لها ، أخذنا عدداً من العينات من مختلف جهات القطر بحيث تحتوى كل عينة على نفس العدد من المفردات . إذا حسبنا متوسطات هذه العينات ووجدنا أنها لا تختلف كثيراً عن بعضها ، أو بعبارة أخرى لو كان الانحراف المعيارى لهذه المتوسطات صغيراً ، لكان احتمال تسرب الخطأ إلى نتائجنا عن طريق إحلال متوسط العينة محل متوسط المجموع الأصلى ضئيلا . ولو كان الانحراف المعيارى لهذه المتوسطات كبيراً لدل ذلك على أن المتوسطات تختلف كثيراً عن متوسطها (وهو يساوى متوسط المجموع الأصلى ذاته) . ولكان احتمال الخطأ كبيراً .

وسنثبت فيما يلى أنه كلما زاد حجم العينات كلما قل الانحراف المعيارى لمتوسطاتها وكلما أصبحت النتائج التى نصل إليها بإحلال الوسط الحسابى للعينة محل الوسط الحسابى للمجموع الأصلى أقرب إلى الدقة . فإذا رمزنا للانحراف المعيارى للمتوسطات بالرمز ($\sigma_م$) والانحراف المعيارى للمجموع الأصلى بالرمز (σ) وكان عدد مفردات العينة (م) فإن :—

$$\sigma_م = \frac{\sigma}{\sqrt{م}}$$

ولكى نثبت صحة هذه العلاقة سنثبت أولا أن الوسط الحسابى لمتوسطات العينات يساوى الوسط الحسابى للمجموع الأصلى ومن ثم ننتقل إلى إيجاد

(*) سيجد القارئ مما يلى أنه إذا أخذنا جميع العينات الممكنة التى تحتوى كل منها على نفس العدد من المفردات من مجموع معين فان الوسط الحسابى لمتوسطات هذه العينات يساوى الوسط الحسابى للمجموع الأصلى ذاته . وبناء على ذلك نستطيع أن نقول من باب التقريب أنه إذا أخذنا عدداً كبيراً من العينات من مجموع معين فإن الوسط الحسابى لمتوسطات هذه العينات يساوى الوسط الحسابى للمجموع الأصلى .

الانحراف المعيارى للمتوسطات وسنجد أنه يساوى $\frac{\sigma}{\sqrt{م}}$

الوسط الحسابى لمتوسطات العينات :

لإثبات أن الوسط الحسابى لمتوسطات العينات يساوى الوسط الحسابى للمجموع الأصلى افرض أن عدد مفردات هذا المجموع هو ($ن$) وأن هذه المفردات هى على التوالى $س_1$ ، $س_2$ ، $س_3$ .... $س_ن$

ثم افرض أننا اخترنا من هذا المجموع جميع العينات الممكنة التى يحتوى كل منها على (م) من المفردات .

عدد جميع العينات الممكن اختيارها هو عدد الطرق التى يمكن بها أن نختار (م) مفردة من عددٍ من المفردات قدره ($ن$) .

ومن المعروف أن هذا العدد = $ن_م = \frac{ن!}{م!\,(ن-م)!}$

∴ متوسط متوسطات العينات :

$$= \frac{\left(\frac{س_1 + س_2 + ..}{م}\right) + \left(\frac{س_2 + س_3 + ..}{م}\right) + ... \text{ إلى } ن_م \text{ من الحدود}}{ن_م}$$

$$= \frac{\left[\frac{1}{م} س_1 + \frac{1}{م} س_2 + ..\right] + \left[\frac{1}{م} س_2 + \frac{1}{م} س_3 + ..\right] + ...}{ن_م} \quad ... (1)$$

إذا توصلنا إلى طريقة تمكننا من الحصول على عدد المرات التى تظهر فيها $س_1$ فى العينات المختلفة (أو عدد العينات التى تحتوى على $س_1$) لكان هذا العدد معادلا تماماً لعدد المرات التى تظهر فيها $س_2$ أو $س_3$ أو $س_ن$ وذلك لأن كل مفردة لها نفس الفرصة فى الظهور . وإذا حصلنا على عدد المرات التى تظهر فيها $س_1$ أمكننا الحصول على معامل $س_1$ فى المعادلة (١) ومن الواضح أنه هو أيضاً معامل $س_2$ ، $س_3$ ، ... $س_ن$

وللحصول على هذا العدد افرض أننا استبعدنا $س_1$ من المجموع الأصلى واخترنا من المفردات الباقية وعددها $(ن - ١)$ جميع العينات الممكنة التى يحتوى كل منها على $(م - ١)$ من المفردات .

$$\text{عدد جميع العينات الممكنة (بدون } س_1) = {}^{ن-١}م_{-١} = \frac{(ن-١)!}{(م-١)!\,(ن-م)!}$$

من البدهى أنه إذا أضفنا $(س_1)$ إلى كل من هذه العينات الناقصة لما غيرنا عددها وإن كان عدد المفردات فى كل عينة سيصبح (م) وهو العدد المطلوب فى كل عينة ، كما أن عدد مفردات المجموع الذى نختار منه سيعود إلى ما كان عليه فيصبح (ن) بعد إضافة $س_1$. ومن هذا يتضح أن عدد العينات التى تحتوى على (م) مفردة بينها $(س_1)$ الممكن اختيارها من مجموع يحتوى على (ن) مفردة هو :-

$$ق_{م-١}^{ن-١} = \frac{(ن-١)!}{(م-١)!\,(ن-م)!}$$

ومن الواضح أن نفس هذه الكمية تمثل عدد العينات التى تحتوى على $س_2$ أو $س_3$ ... أو $س_ن$ وذلك لأنه كان من الممكن أن نستبعد $س_2$ أو $س_3$ أو $س_ن$ بدلا من $س_1$ ثم نضيفها إلى كل عينة تحتوى على $(م - ١)$ مفردة كما فعلنا مع $س_1$. الآن وقد حصلنا على عدد العينات التى تحتوى

على س$_1$ يمكننا بسهولة أن نحصل على معامل س$_1$ فى الكسر الذى يمثل متوسط متوسطات العينات .

معامل س$_1$ فى المعادلة (١) .

$$= \frac{(ن-١)!}{(م-١)!\,(ن-م)!} \times \frac{١}{م} \times \frac{م!\,(ن-م)!}{ن!} \quad (٥)$$

$$= \frac{١}{ن} \text{ وهذا كما ذكرنا هو معامل جميع قيم س.}$$

∴ متوسط متوسطات العينات

$$= \frac{١}{ن} س_1 + \frac{١}{ن} س_2 + \frac{١}{ن} س_3 + \cdots \frac{١}{ن} س_ن$$

$$= \frac{س_1 + س_2 + س_3 + \cdots + س_ن}{ن} = \text{متوسط المجموع الأصلى .}$$

الانحراف المعيارى لمتوسطات العينات :

لحساب الانحراف المعيارى لمتوسطات جميع العينات التى يحتوى كل منها على (م) مفردة الممكن اختيارها من مجموع عدد مفرداته (ن) نستخدم العلاقة $\sigma^2_م = \frac{\Sigma س^2_م}{و} - \overline{س}^2_م$

(حيث قيم س$_م$ فى هذه العلاقة تمثل متوسطات العينات ، (و) تمثل عددها ، $\overline{س}_م$ تمثل متوسطها[1]) .

---

(٥) قسمنا عدد المرات التى تظهر فيها (س$_1$) على (م) للحصول على معامل (س$_1$) فى بسط الكسر ثم قسمنا مرة أخرى على المقام للحصول على معامل (س$_1$) فى الكسر كله .

(١) يجب على الطالب ألا يخلط بين قيم س$_م$ فى هذه العلاقة وقيم س التى تمثل مفردات المجموع الأصلى . وقد استعلمنا (و) بدلا من (ن) التى اعتدنا استعمالها فى هذه العلاقة للدلالة على عدد قيم المتغير حتى لا تختلط مع (ن) التى تمثل عدد مفردات المجموع .

$$\Sigma س_م^{2} = \left(\frac{س_1 + س_2 + \cdots}{م}\right)^{2} + \left(\frac{س_2 + س_3 + \cdots}{م}\right)^{2}$$

$$+ \left(\frac{س_3 + س_4 + \cdots}{م}\right)^{2} + \cdots \text{ إلى } ن \text{ من الحدود}$$

$$= \frac{1}{م^{2}} \left\{ (س_1 + س_2 + \cdots)^{2} + (س_2 + س_3 + \cdots)^{2} + \cdots \right\} \cdots (1)$$

بفك هذه الأقواس سيظهر عندنا نوعان من الحدود :

الأول عبارة عن $س_1^{2}$ ، $س_2^{2}$ ، $س_3^{2}$ ، $\cdots س_ن^{2}$

والثانى عبارة عن $س_1 س_2$ ، $س_1 س_3$ ، $س_2 س_3$ ، ... ، إلى جميع التوافيق الممكنة لقيمتين من قيم س ويلاحظ أن معامل $س_1^{2}$ يساوى معامل $س_2^{2}$ أو أية قيمة أخرى من قيم $س^{2}$ وبالمثل فإن معامل $س_1 س_2$ هو أيضاً معامل $س_1 س_3$ أو أى توفيق آخر لقيمتين من قيم س وعلى ذلك يمكن وضع مفكوك الأقواس فى الصيغة الآتية :—

$$\Sigma س_م^{2} = \frac{1}{م^{2}} (س_1^{2} + س_2^{2} + \cdots + س_ن^{2})$$

$$+ \frac{ب}{م^{2}} (2 س_1 س_2 + 2 س_1 س_3 + \cdots) \cdots (2)$$

حيث ا تمثل عدد المرات التى تظهر فيها $س_1^{2}$ فى المفكوك وهو عبارة عن عدد العينات التى تحتوى على $س_1$

، ب تمثل عدد المرات التى تظهر فيها $2 س_1 س_2$ فى المفكوك وهو عبارة عن عدد العينات التى تحتوى على $س_1$ ، $س_2$ معاً .

وقد سبق أن بينا أن عدد العينات التى تحتوى على $س_1$ $= ن^{-1} ق_{م-1}$

$$\therefore ا = ن^{-1} ق_{م-1} = \frac{(ن - 1)!}{(م - 1)!\,(ن - م)!}$$

$$\therefore \text{معامل } س_1^{\,2} \text{ فى (٢)} = \frac{1}{م^2} \times \frac{(ن - 1)!}{(م - 1)!\,(ن - م)!}$$

ولإيجاد قيمة ب نستبعد $س_1$ ، $س_2$ من المجموع الأصلى ونختار من هذا المجموع الناقص جميع العينات الممكنة التى يحتوى كل منها على (م – ٢) مفردة ومن الواضح أن عدد هذه العينات يساوى عدد العينات التى تحتوى على $س_1$ ، $س_2$ معاً إذ أنه من الممكن إضافة $س_1$ ، $س_2$ إلى كل من هذه العينات الناقصة بدون تغيير عددها .

$$\therefore ب = ن^{-2} ق_{م-2} = \frac{(ن - 2)!}{(م - 2)!\,(ن - م)!}$$

∴ معامل $س_1$ $س_2$ (أو أى توفيق آخر لقيمتين من قيم س) فى المعادلة (٢)

$$= \frac{2}{م^2} \times \frac{(ن - 2)!}{(م - 2)!\,(ن - م)!}$$

وبالتعويض عن قيمة ا ، ب فى المعادلة (٢)

$$\therefore \Sigma س_م^2 = \frac{1}{م^2} \times \frac{(ن-1)!}{(م-1)!\,(ن-م)!} (س_1^2 + س_2^2 + \ldots + س_ن^2)$$

$$+ \frac{2}{م^2} \times \frac{(ن-2)!}{(م-2)!\,(ن-م)!} (س_1 س_2 + س_2 س_3 + \ldots)$$

ولإيجاد قيمة $\frac{\Sigma س_م^2}{و}$ نقسم الطرف الأيسر على $\frac{ن!}{م!\,(ن-م)!}$

وهى الكمية التى تمثل عدد المتوسطات

$$\therefore \frac{\Sigma س_م^2}{و} = \frac{1}{م ن} (س_1^2 + س_2^2 + \ldots + س_ن^2)$$

$$+ \frac{2(م-1)}{م ن (ن-1)} (س_1 س_2 + س_2 س_3 + \ldots)$$

ولإيجاد الانحراف المعيارى لمتوسطات العينات نطرح من هذا المقدار $\overline{س_م}^2$

ولكن $\overline{س_م}^2 = \overline{س}^2 = \left(\frac{س_1 + س_2 + \ldots + س_ن}{ن}\right)^2$ (*)

وبإيجاد مفكوك هذا القوس

(*) $\overline{س_م}$ تمثل متوسط متوسطات العينات ، $\overline{س}$ تمثل متوسط المجموع الأصلى وقد أثبتنا من قبل تساوى هاتين القيمتين .

$$\therefore \overline{س_م}^2 = \frac{1}{ن^2} (س_1^2 + س_2^2 + \cdots + س_ن^2)$$

$$+ \frac{1}{ن^2} (2 س_1 س_2 + 2 س_2 س_3 + \cdots)$$

وحيث أن $ع_م^2 = \frac{\Sigma س_م^2}{و} - \overline{س_م}^2$

$$\therefore ع_م^2 = (س_1^2 + س_2^2 + \cdots + س_ن^2) \left(\frac{1}{م ن} - \frac{1}{ن^2}\right)$$

$$+ (س_1 س_2 + س_2 س_3 + \cdots) \left(\frac{2(م-1)}{م ن (ن-1)} - \frac{2}{ن^2}\right)$$

$$= (س_1^2 + س_2^2 + \cdots + س_ن^2) \left(\frac{ن-م}{م ن^2}\right)$$

$$+ 2 (س_1 س_2 + س_2 س_3 + \cdots) \left(\frac{م-ن}{م ن^2 (ن-1)}\right)$$

$$= (س_1^2 + س_2^2 + \cdots + س_ن^2) \left(\frac{ن-م}{م ن^2}\right)$$

$$- 2 (س_1 س_2 + س_2 س_3 + \cdots) \left(\frac{ن-م}{م ن^2 (ن-1)}\right)$$

والآن نقارن الانحراف المعيارى للمتوسطات بالانحراف المعيارى للمجموع الأصلى الذى نحصل عليه من العلاقة $ع^2 = \frac{\Sigma س^2}{ن} - \overline{س}^2$

$$\sigma_{\bar{س}}^{2} = \frac{س_1^{2} + س_2^{2} + \cdots + س_ن^{2}}{ن} - \left(\frac{س_1 + س_2 + \cdots + س_ن}{ن}\right)^{2}$$

$$= \frac{س_1^{2} + س_2^{2} + \cdots + س_ن^{2}}{ن}$$

$$- \left[\frac{1}{ن^{2}} (س_1^{2} + س_2^{2} + \cdots + س_ن^{2}) + \frac{2}{ن^{2}} (س_1 س_2 + س_2 س_3 + \cdots)\right.$$

$$= (س_1^{2} + س_2^{2} + \cdots + س_ن^{2}) \left(\frac{1}{ن} - \frac{1}{ن^{2}}\right)$$

$$- \frac{2}{ن^{2}} (س_1 س_2 + س_2 س_3 + \cdots)$$

ويلاحظ أنه إذا ضربنا هذه الكمية الأخيرة فى $\frac{ن - م}{م (ن - 1)}$ نحصل على كمية تساوى $\sigma_{م}^{2}$

$$\therefore \sigma_{م} = \frac{ن - م}{م (ن - 1)} \times \sigma^{2} = \frac{\sigma^{2}}{م} \times \frac{ن - م}{ن - 1}$$

وحيث أن (ن) تمثل عدد مفردات المجموع الأصلى وهذا دائماً كبير جداً بالنسبة لعدد مفردات العينة (م) فيمكن اعتبار $\frac{ن - م}{ن - 1} = 1$

$$\therefore \sigma_{\bar{س}}^{2} = \frac{\sigma^{2}}{م}$$

$$\therefore \sigma_{م} = \frac{\sigma}{\sqrt{م}}$$

وحيث أن الانحراف المعيارى للمجموع الأصلى كمية ثابتة لا تتأثر بقيمة (م) فإن $\sigma_{م}$ تتغير تبعاً لقيمة (م) .

وكلما كانت (م) كبيرة كانت $\sigma_{م}$ صغيرة أى كلما كان مقدار الخطأ الذى يتسرب إلى نتائجنا نتيجة لإحلال الوسط الحسابى لعينة عدد مفرداتها (م) محل الوسط الحسابى للمجموع الأصلى صغيرا .

والكمية $\sigma_{م}$ التى تمثل الانحراف المعيارى لمتوسطات العينات تسمى الخطأ المعيارى للوسط الحسابى (Standard Error of The Mean) . ولما كان التوزيع التكرارى لمتوسطات العينات يقرب من التوزيع الطبيعى[1] فإننا نحصل على النتائج الهامة الآتية : —

٩٥,٤٥ ٪ من متوسطات العينات تنحصر بين

$$(\bar{س} + ٢\ \sigma_{م}) \text{ ، } (\bar{س} - ٢\ \sigma_{م})$$

, ٩٩,٧٣ ٪ من متوسطات العينات تنحصر بين

$$(\bar{س} + ٣\ \sigma_{م}) \quad (\bar{س} - ٣\ \sigma_{م})$$

, ٥٠ ٪ من متوسطات العينات تنحصر بين

$$(\bar{س} + ٠,٦٧٤٥\ \sigma_{م}) \text{ ، } (\bar{س} - ٠,٦٧٤٥\ \sigma_{م})$$[1]

ومن هذا نستنتج أن احتمال اختلاف الوسط الحسابى للعينة عن الوسط

---

(١) هذا على فرض أن العينات كبيرة . أما متوسطات العينات الصغيرة (التى تحتوى على أقل من ٣٠ مفردة) فإنها تكون موزعة طبقاً لمنحنى آخر يعرف باسم Student's Curve وسنقصر بحثنا هذا على حالة العينات الكبيرة .

الحسابى للمجموع الأصلى بمقدار + ٢ $\sigma_م$ = ٤,٥٥ ٪ وهو احتمال ضئيل وأن احتمال اختلافه عنه بمقدار + ٣ $\sigma_م$ = ٠,٢٧ ٪ وهو احتمال ضئيل جداً يمكن إهماله عملياً . أما احتمال اختلاف الوسط الحسابى للعينة عن الوسط الحسابى للمجموع الأصلى بمقدار + ٠,٦٧٤٥ $\sigma_م$ فإنه يساوى ٥٠ ٪ تماماً ولذلك فإن الكمية ٠,٦٧٤٥ $\sigma_م$ تسمى :

الخطأ المحتمل للوسط الحسابى (Probable Error of The Mean) ويسميها البعض الخطأ المتعادل للوسط الحسابى وهذه التسمية الأخيرة أقرب إلى الدقة .

والذى نستخلصه من هذه النتائج أنه إذا استخرجنا وسطاً حسابياً لعينة عدد مفرداتها (م) واعتبرناه يمثل الوسط الحسابى للمجموع الذى انتخبت منه فإن الخطأ الناشىء عن ذلك لا يمكن عملياً أن يتجاوز + ٣ $\sigma_م$

وبعبارة أخرى فإن الخطأ لا يمكن أن يتجاوز $\frac{\sigma ٣}{\sqrt{م}}$ (حيث σ تمثل الانحراف المعيارى للمجموع) . وحيث أنه يكاد يكون من المستحيل عملياً حساب الانحراف المعيارى للمجموع الأصلى ذاته فإن الطريقة المتبعة عادة تتلخص فى الاستعاضة عن الانحراف المعيارى للمجموع الأصلى بالانحراف المعيارى للعينة نفسها . وعلى هذا الأساس يكون الخطأ المعيارى للوسط الحسابى

$$= \frac{\sigma}{\sqrt{م}}$$ (حيث σ تمثل الانحراف المعيارى للعينة)*

(*) يميل معظم الإحصائيين إلى عمل تعديل بسيط فى قيمة الانحراف المعيارى للعينة عند إحلاله محل الانحراف المعيارى للمجموع الأصلى ، فبدلا من اعتباره

$$= \sqrt{\frac{\Sigma ح^٢}{ن}}$$ يفضل هؤلاء اعتباره $\sqrt{\frac{\Sigma ح^٢}{ن-١}}$ ومهما تكن الأسباب النظرية التى تدعو إلى عمل هذا التعديل فمن الواضح أن قيمته العملية ضئيلة جداً فى حالة العينات الكبيرة أى عند ما تكون قيمة ن كبيرة .

تطبيقات عملية للخطأ المعيارى للوسط الحسابى

مراقبة جودة الإنتاج

مراقبة جودة السلع التى تنتجها المصانع من المسائل العظيمة الأهمية فى الحياة العملية خصوصاً فى حالات الإنتاج النمطى على نطاق كبير . فمن المسلم به أنه مهما بلغت دقة الآلات التى تستخدم فى المصانع ومهما بلغت مهارة العمال فإن بعض المنتجات تنحرف قليلا أو كثيراً عن مستوى الجودة المطلوب . فالصانع الماهر قد يخطىء أحياناً كما أن وسائل الإنتاج الآلية مهما بلغت من الدقة والإحكام فقد تخرج أحياناً بعض المنتجات المعيبة . وموضوع مراقبة جودة الإنتاج يبحث فى إيجاد طريقة لوضع حدود لا يجوز أن تزيد عنها نسبة المنتجات المعيبة على فرض أن وسائل الإنتاج سائرة بانتظام . أو بعبارة أخرى إذا تجاوزت نسبة المنتجات المعيبة هذه الحدود لدل ذلك على وجود خلل فى عملية الإنتاج يجب إصلاحه . ولإيضاح هذا الموضوع افرض أن مصنعاً ينتج أسطوانات من الصلب قطر كل منها خمسة سنتيمترات . مهما بلغت دقة الآلات ومهارة العمال فلا يمكن عملياً أن يكون قطر كل أسطوانة خمسة سنتيمترات بالضبط بل إن بعضها قد يزيد قطره عن هذا المستوى بمقدار معين (نصف ملليمتر أو ربع ملليمتر مثلا) بينما قد يقل قطر البعض الآخر عن هذا المستوى قليلا . فلو فرض أن الوسط الحسابى للأقطار ($\bar{س}$) والانحراف المعيارى لها (σ) يمكننا بمساعدة هذه البيانات أن نختبر عملية الإنتاج لمعرفة ما إذا كان قد طرأ عليها أى خلل من شأنه أن يقلل مستوى الدقة المطلوب . ولغرض الاختبار نأخذ عدداً كبيراً من العينات الاعتباطية يحتوى كل منها على عدد معين من المفردات (م مثلا) ثم نحسب متوسطات هذه العينات . فإذا كان ٩٩,٧٣٪ من هذه المتوسطات (أى جميعها تقريباً) لا يختلف عن الوسط الحسابى للأقطار ($\bar{س}$)

بأكثر من $\frac{٣\sigma}{\sqrt{م}}$ (أى ثلاثة أمثال الخطأ المعيارى للوسط الحسابى) دل ذلك على أن المجموع الأصلى الذى انتخبت منه العينات لم يتغير وبالتالى يمكن اعتبار عملية الإنتاج سائرة بانتظام . على أنه مما تجب ملاحظته أن هذا الاختبار لا يعدو كونه اختباراً مبدئياً . وهناك عدة اختبارات أخرى يجب إجراؤها كى تكون عملية المراقبة محكمة(١) .

اختبار دلالة الفرق بين متوسطى عينتين :

فى كثير من التجارب العلمية خصوصاً تلك المتصلة بالأبحاث البيولوجية وأبحاث علم النفس تعرض لنا مسألة اختبار دلالة الفرق بين متوسطى عينتين لمعرفة ما إذا كان هناك اختلاف جوهرى بين المجموعين اللذين انتخبت منهما العينتان .

ولاختبار دلالة هذا الفرق نعين الخطأ المعيارى للفرق بين المتوسطين الذى نعبر عنه بالرمز ($\sigma_{ف}$) . فإذا كان الفرق بين المتوسطين أكبر من $٣\sigma_{ف}$ دل ذلك على وجود اختلاف جوهرى بين المجموعين اللذين انتخبنا منهما العينتين . وإذا كان الفرق أقل من $٣\sigma_{ف}$ فإنه لا يدل على وجود اختلاف جوهرى بين المجموعين بل يمكن أن نعزوه إلى عوامل المصادفة(٢) .

---

(١) يستطيع الطالب الذى يرغب فى التوسع فى موضوع مراقبة جودة الإنتاج أن يرجع إليه فى كتاب Economic Control of Quality of Manufactured Product. By W.A. Shewhart.

(٢) يفضل بعض الإحصائيين اعتبار الفرق الذى يزيد عن $٢\sigma_{ف}$ فرقاً جوهرياً لا يجوز أن يعزى إلى عوامل المصادفة . وذلك لأن احتمال حدوث هذا الفرق بطريق المصادفة = ٤,٥٥٪ وهو احتمال ضئيل . أما احتمال حدوث فرق قدره $٣\sigma_{ف}$ فلا يزيد عن ٠,٢٧٪ . وسيجد القارئ بمراجعة خواص المنحنى الطبيعى أن هذه الاحتمالات مقدرة على أساس كون الفروق موزعة توزيعاً طبيعياً .

يبقى الآن أن نحدد قيمة الخطأ المعيارى للفرق بين المتوسطين ، وسنثبت فيما يلى أن :

$$\sigma_{ف} = \sqrt{\frac{\sigma_{1}^{2}}{ن_{1}} + \frac{\sigma_{2}^{2}}{ن_{2}}}$$

حيث σ_{1} ، σ_{2} تمثلان الانحرافين المعياريين للعينتين ، $ن_{1}$ ، $ن_{2}$ تمثلان عدد مفردات العينتين .

افرض أن لدينا (ن) من أزواج القيم المتناظرة فى العينتين س ، ص هى على التوالى :

(س$_{1}$ ، ص$_{1}$) ، (س$_{2}$ ، ص$_{2}$) ، ..... (س$_{ن}$ ، ص$_{ن}$)

الخطأ المعيارى للفروق بين متوسطات العينات هو عبارة عن الانحراف المعيارى لهذه الفروق وهذا يمكن الحصول عليه بالطريقة المعتادة . ولتبسيط طريقة الإثبات نحسب أولا الانحراف المعيارى للفروق بين القيم المتناظرة ذاتها ومن ثم نحسب الانحراف المعيارى للفروق بين المتوسطات :

العينة (س)	العينة (ص)	الفروق
س$_{1}$	ص$_{1}$	س$_{1}$ — ص$_{1}$
س$_{2}$	ص$_{2}$	س$_{2}$ — ص$_{2}$
س$_{3}$	ص$_{3}$	س$_{3}$ — ص$_{3}$
.	.	.
.	.	.
.	.	.
س$_{ن}$	ص$_{ن}$	س$_{ن}$ — ص$_{ن}$

لإيجاد الانحراف المعيارى للفروق بين أزواج القيم المتناظرة نبدأ بإيجاد

الوسط الحسابى للفروق وبعد ذلك نحسب الانحرافات عنه ومربعات هذه الانحرافات .

$$\text{الوسط الحسابى للفروق} = \frac{\text{مجموع الفروق}}{\text{عددها}} = \frac{\Sigma\,(س - ص)}{ن}$$

$$= \frac{\Sigma\,س - \Sigma\,ص}{ن} = \overline{س} - \overline{ص}$$

والآن نحسب انحرافات الفروق عن متوسطها ثم مربعات هذه الانحرافات كالآتى :

انحرافات الفروق عن متوسطها	مربعات الانحرافات
$(س_1 - ص_1) - (\overline{س} - \overline{ص})$	$[(س_1 - ص_1) - (\overline{س} - \overline{ص})]^2$
$(س_2 - ص_2) - (\overline{س} - \overline{ص})$	$[(س_2 - ص_2) - (\overline{س} - \overline{ص})]^2$
$(س_3 - ص_3) - (\overline{س} - \overline{ص})$	$[(س_3 - ص_3) - (\overline{س} - \overline{ص})]^2$
. . . .	. . . .
. . . .	. . . .
. . . .	. . . .
. . . .	. . . .
$(س_ن - ص_ن) - (\overline{س} - \overline{ص})$	$[(س_ن - ص_ن) - (\overline{س} - \overline{ص})]^2$

∴ مجموع مربعات انحرافات الفروق عن متوسطها $= \Sigma\,[(س - ص) - (\overline{س} - \overline{ص})]^2$

$= \Sigma\,[(س - \overline{س}) - (ص - \overline{ص})]^2$

$(س - \overline{س})$ تمثل انحرافات قيم س عن وسطها الحسابى ونرمز لها بالرمز (ح)

، (ص − $\overline{\text{ص}}$) تمثل انحرافات قيم ص عن وسطها الحسابى ونرمز لها بالرمز (ل)

∴ مجموع مربعات انحرافات الفروق عن متوسطها = Σ (ح − ل)^{٢}

∴ مربع الانحراف المعيارى للفروق $= \frac{\Sigma (ح - ل)^{٢}}{ن}$

$$= \frac{\Sigma (ح^{٢} - ٢ ح ل + ل^{٢})}{ن} = \frac{\Sigma ح^{٢}}{ن} - \frac{٢ \Sigma ح ل}{ن} + \frac{\Sigma ل^{٢}}{ن}$$

ولكن $ر_{س ص}$ (معامل الارتباط بين س ، ص) $= \frac{\Sigma ح ل}{ن \sigma_{س} \sigma_{ص}}$

∴ مربع الانحراف المعيارى للفروق $= \frac{\Sigma ح^{٢}}{ن} - ٢ ر_{س ص} \sigma_{س} \sigma_{ص} + \frac{\Sigma ل^{٢}}{ن}$

ولكن $\frac{\Sigma ح^{٢}}{ن} = \sigma^{٢}_{س}$ ، $\frac{\Sigma ل^{٢}}{ن} = \sigma^{٢}_{ص}$

∴ مربع الانحراف المعيارى للفروق $= \sigma^{٢}_{س} - ٢ ر_{س ص} \sigma_{س} \sigma_{ص} + \sigma^{٢}_{ص}$

∴ الانحراف المعيارى للفروق $= \sqrt{\sigma^{٢}_{س} - ٢ ر_{س ص} \sigma_{س} \sigma_{ص} + \sigma^{٢}_{ص}}$

وعلى فرض أن مفردات العينتين مستقلة تماماً عن بعضها فإن ر = صفر

∴ الانحراف المعيارى للفروق $= \sqrt{\sigma^{٢}_{س} + \sigma^{٢}_{ص}}$

وللحصول على الانحراف المعيارى للفروق بين المتوسطات نعتبر أن كلا

من (س١ ، ص١) ، (س٢ ، ص٢) ... إلخ هى متوسطات عينات مختلفة بدلا من كونها مفردات . وفى هذه الحالة يكون الانحراف المعيارى للفروق بين المتوسطات (أى الخطأ المعيارى) هو :

$$\sigma_{ف} = \sqrt{\sigma^{2}_{\bar{س}} + \sigma^{2}_{\bar{ص}}}$$

على فرض أن المتوسطات غير مرتبطة ببعضها .

غير أن الانحراف المعيارى (أو الخطأ المعيارى) للوسط الحسابى = $\frac{\sigma}{\sqrt{ن}}$

$$\therefore \sigma^{2}_{\bar{س}} = \frac{\sigma^{2}_{س}}{ن} \quad \text{وبالمثل فإن} \quad \sigma^{2}_{\bar{ص}} = \frac{\sigma^{2}_{ص}}{ن} \quad (٤)$$

$$\therefore \sigma_{ف} \text{ (الخطأ المعيارى للفرق بين أى متوسطين)} = \sqrt{\frac{\sigma^{2}_{س}}{ن} + \frac{\sigma^{2}_{ص}}{ن}}$$

ويلاحظ أنه إذا اختلف عدد مفردات العينتين اللتين نختبر دلالة الفرق بين متوسطيهما وكان عددهما $ن_{١}$ ، $ن_{٢}$ على التوالى

فإن الخطأ المعيارى للفرق بين المتوسطين = $\sqrt{\frac{\sigma^{2}_{س}}{ن_{١}} + \frac{\sigma^{2}_{ص}}{ن_{٢}}}$

مثال : أجرى اختبار للذكاء على مجموعة من التلاميذ ومجموعة من التلميذات وكان الوسط الحسابى والانحراف المعيارى لدرجات كل من المجموعتين كالآتى :

(*) نلفت نظر القارئ إلى أن $\sigma_{س}$ معناها الانحراف المعيارى لقيم س ، $\sigma_{\bar{س}}$ معناها الانحراف المعيارى للمتوسطات السينية .

	التلاميذ	التلميذات
الوسط الحسابى	١٢٤	١٢١
الانحراف المعيارى	١٢	١٠
عدد المفردات	٧٢	٥٠

هل الفرق بين متوسطى المجموعتين يدل على وجود اختلاف جوهرى بينهما أم يمكن أن نعزوه إلى عوامل المصادفة ؟

الفرق بين المتوسطين = ١٢٤ — ١٢١ = ٣ .

$$\text{الخطأ المعيارى للفرق} = \sqrt{\frac{(١٢)^{٢}}{٧٢} + \frac{(١٠)^{٢}}{٥٠}} = \sqrt{٤} = ٢$$

وحيث أن الفرق بين المتوسطين أقل بكثير من ثلاثة أمثال الخطأ المعيارى للفرق .

∴ هذا الفرق لا يدل على وجود اختلاف جوهرى بين ذكاء التلاميذ وذكاء التلميذات ويمكن أن نعزوه إلى عوامل المصادفة .

ومما تجب ملاحظته أن هذا الاختبار لا يطبق إلا فى حالة العينات الكبيرة أو على وجه التحديد تلك التى لا تقل عدد مفردات كل منها عن ٣٠(١) .

وكما أوجدنا الخطأ المعيارى للوسط الحسابى والخطأ المعيارى للفرق بين متوسطى عينتين يمكننا إيجاد الأخطاء المعيارية لكافة المقاييس الإحصائية الأخرى ونذكر أهمها فيما يلى :

---

(١) لاختبار دلالة الفرق بين متوسطى عينتين صغيرين نطبق اختبار معروف به (« t » Test) وضعه « Student » ويمكن للطالب الرجوع إليه فى كتاب :

Statistical methods for reeasrch workers, by R.A. Fisher.

الأخطاء المعيارية لبعض المقاييس الإحصائية(١)

المقياس الإحصائى	الخطأ المعيارى
١ ــ الوسط الحسابى لأى توزيع تكرارى	$\frac{\sigma}{\sqrt{ن}}$
٢ ــ الانحراف المعيارى لتوزيع طبيعى	$\frac{\sigma}{\sqrt{٢ن}}$
٣ ــ العزم الثانى حول الوسط الحسابى لتوزيع طبيعى	$\sqrt{\frac{٢}{ن}}\sigma^{٢}$
٤ ــ العزم الثالث » » » » »	$\sqrt{\frac{٦}{ن}}\sigma^{٣}$
٥ ــ العزم الرابع » » » » »	$\sqrt{\frac{٩٦}{ن}}\sigma^{٤}$
٦ ــ معامل الالتواء ($ل_{١}$) » »	
٧ ــ معامل التفرطح ($ل_{٢}$) » »	$\sqrt{\frac{٢٤}{ن}}$
٨ ــ معامل الارتباط ($ر$) » »	$\frac{١ - ر^{٢}}{\sqrt{ن}}$
٩ ــ نسبة الارتباط (ى) » »	$\frac{١ - ى^{٢}}{\sqrt{ن}}$
١٠ ــ معامل الاختلاف (ف) » »	$\frac{ف}{\sqrt{٢ن}}\left[١ + ٢\left(\frac{ف}{١٠٠}\right)^{٢}\right]^{\frac{١}{٢}}$

(١) انظر كتاب A First Course in Statistics By D.C. Jones P. 163.

ولكى نقدر أهمية هذه الأخطاء المعيارية يجب أن نتذكر أنه عند ما نحسب أى مقياس إحصائى لعينة ما فإننا نفعل ذلك لتعميم النتيجة التى نحصل عليها على المجموع الذى انتخبت منه هذه العينة . وهذا التعميم يتضمن طبعاً قبولنا لنتائج تقريبية . إذ لا يعقل مثلا أن يكون الوسط الحسابى للعينة هو نفسه الوسط الحسابى للمجموع الذى انتخبت منه أو أن معامل الارتباط كما نستخرجه من واقع قيم عينتين يساوى معامل الارتباط بين المجموعين اللذين انتخبت منهما هاتان العينتان . غير أن الأخطاء المعيارية تعين لنا حدود الأخطاء التى يمكن أن تتسرب إلى نتائجنا من جراء إحلال العينة محل المجموع الذى انتخبت منه . فهذه الأخطاء لا يمكن أن تتجاوز ثلاثة أمثال الخطأ المعيارى . ولهذا يحسن دائماً عند استخراج قيمة أى مقياس إحصائى أن نذكر بجانبه قيمة الخطأ المعيارى لهذا المقياس كى يعلم من يطلع عليه حدود الخطأ المحتمل تسربه إلى النتيجة .

الباب الثالث عشر

تحليل السلسلات الزمنية

إذا بحثت فى ظاهرة معينة خلال فترة طويلة نسبيا من الزمن وجدت أن قيم هذه الظاهرة تتغير بمضى الزمن . وهى فى تغيرها هذا خاضعة لعدة مؤثرات بعضها منتظم أى ناشئ عن عوامل يمكن حصرها ومعرفة درجة تأثيرها على الظاهرة على وجه التقريب والبعض الآخر عبارة عن مؤثرات عرضية فجائية تؤثر على الظاهرة حيثما اتفق أحياناً بالصعود وأحياناً بالهبوط دون أن تتبع أى قاعدة يمكن جعلها موضع دراسة علمية مفيدة . ويمكن تمييز ثلاثة أنواع من التغيرات التى تطرأ على قيم الظواهر بانتظام :

١ — التغيرات الاتجاهية أو الاتجاه العام Secular Trend

٢ — التغيرات أو التقلبات الموسمية Seasonal Variations

٣ — التغيرات أو التقلبات الدورية Cyclical Variations

أما التغيرات غير المنتظمة فسنشير إليها بالتغيرات أو التقلبات العرضية Accidental Variations . وسنشرح طبيعة كل من هذه التغيرات فى إيجاز ثم ننتقل إلى عرض طرق تحليل السلسلات الزمنية :

١ ــ التغيرات الاتجاهية

أول ما يلفت النظر فى التغيرات التى تطرأ على قيمة غالبية الظواهر هو النمو الطبيعى المتدرج عاماً بعد عام . فالعالم ليس فى حالة سكون بل هو دائم التطور وإن كانت درجة التطور لا تلاحظ فى الفترات القصيرة إلا أنها تكون واضحة فى الفترات الطويلة . فلو بحثت فى ظاهرة استهلاك السكر مثلا فى أية دولة من دول العالم خلال فترة طويلة من الزمن لوجدت أن الاستهلاك الآن أضعافاً مضاعفة لما كان عليه منذ خمسين سنة أو مائة سنة . ويرجع ذلك من جهة ، إلى زيادة الإنتاج العالمى منه بتقدم الطرق الفنية فى إنتاجه وزيادة المساحة المخصصة لزراعة قصب السكر والبنجر ومن الجهة الأخرى إلى زيادة السكان وتحسين مستوى المعيشة الذى يتمشى مع درجة التقدم . فبعد أن كان السكر يعتبر من الكماليات وكان استهلاكه قاصراً على الطبقات الغنية أصبح الآن من ضروريات الحياة لكافة الطبقات .

على أنه هناك ظواهر أخرى تأخذ فى الانكماش التدريجى عاماً بعد عام . وهى فى ذلك أيضاً خاضعة لنظام التطور الذى يطرأ على حياة الإنسان فيغير الطرق التى يتبعها فى إشباع حاجاته . ومن الأمثلة على ذلك تناقص استهلاك الحرير الطبيعى بظهور الحرير الصناعى وانتشار استعماله بين سكان العالم . ومنها أيضاً الاختفاء التدريجى للعربات التى تجرها الخيل من بعض المدن الكبيرة وظهور السيارات محلها كوسيلة للنقل . كذلك يمكن ملاحظة هبوط استهلاك « الكيروسين » فى معظم الدول المتحضرة وذلك بإحلال الكهرباء والغاز محله كوقود للمنازل والمصانع وغيرها .

ومهمة الإحصائى فى دراسة الاتجاه العام للظواهر هى محاولة الوصول إلى قاعدة منظمة يمكن بواسطتها وصف سير الظاهرة فى الظروف العادية وقياس مقدار الانحرافات عن هذا الاتجاه العام توطئة لمعرفة أسبابها .

التغيرات الموسمية

المقصود بالتغيرات الموسمية تلك التى تحدث بانتظام فى وحدات زمنية متعاقبة كشهر معين من أشهر السنة أو يوم معين من كل شهر أو كل أسبوع أو حتى ساعة معينة من كل يوم . فلو بحثت مثلا فى ظاهرة استهلاك اللحوم فى مصر لوجدت بلا شك أن هناك تغيراً موسمياً يتخذ شكل زيادة كبيرة فى الاستهلاك فى عيد الأضحى المبارك ويتكرر هذا التغير بانتظام فى نفس الموعد من كل سنة . وإذا بحثت فى حركة سحب الودائع من البنوك لوجدتها تزداد فى أول كل شهر وهو الميعاد الذى يحل فيه دفع مهايا الموظفين وإيجارات المنازل والدفعات الشهرية المختلفة . كذلك تشتد حركة السحب فى البلاد التى تطبق نظام الأجر الأسبوعى كإنجلترا مثلا فى أيام الجمعة حيث تسحب الشركات مبالغ كبيرة من ودائعها فى البنوك لاستعمالها فى دفع الأجور . ومتى تسلم العمال أجورهم أنفقوها فى دفع إيجارات منازلهم (وهى عادة أسبوعية فى إنجلترا) وسداد حساباتهم المختلفة وهكذا تنتقل النقود من أيديهم إلى صاحب المنزل والبدال والقصاب فتجد طريقها مرة أخرى إلى خزائن البنوك فى يوم الإثنين عادة . وعلى ذلك فإذا نقصت الودائع نقصاً كبيراً يوم الجمعة فليس فى هذا ما يزعج البنك مطلقاً إذ أنه يعلم تماماً أن هذا النقص هو ظاهرة مؤقتة تتكرر بانتظام وأنه فى خلال يوم أو اثنين تنعكس الآية فتعود الودائع إلى التدفق على البنك مرة أخرى .

التغيرات الدورية

من المشاهد أن النشاط الاقتصادى المسير بنظام الرأسمالية تحل به فترات من الرخاء والكساد تتعاقب فى شىء من الانتظام كل عشر سنوات تقريباً ويطلق عليها الاقتصاديون اسم الدورة التجارية . ولسنا هنا فى معرض الكلام عن الدورة التجارية من حيث أسباب حدوثها أو نتائجها الاقتصادية . وإنما الذى

يهمنا من الوجهة الإحصائية هو تأثيرها على قيم الظواهر التى تكون موضع بحثنا .

فمن المعروف أن فترة الكساد هى الفترة التى يقل خلالها الطلب الإجمالى على كافة السلع والخدمات . فإذا كنا نبحث فى ظاهرة استهلاك السكر فى مصر مثلا ووجدنا أنه على الرغم من أن اتجاهها العام هو نحو الصعود إلا أنه طرأ عليها هبوط كبير فى سنة ١٩٣١ فمن السهل طبعاً أن نستنتج أن هذا الهبوط هو من تأثير الكساد التجارى الذى حل بالبلاد فى تلك السنة . هذا النوع من التغير فى قيم الظواهر هو المقصود بالتغيرات الدورية . ومما تجب ملاحظته أن التغيرات الدورية لا تحدث فى فترات متساوية تماماً مثل التغيرات الموسمية وإن كان حدوثها يتكرر بانتظام ولكن فى فترات أطول كثيراً من فترة الدورة الموسمية .

التغيرات العرضية

التغيرات العرضية هى تلك التى تطرأ على قيم الظواهر إما نتيجة لتلك العوامل غير القابلة للتحديد التى نطلق عليها اسم عوامل المصادفة أو لحدوث طارئ فجائى كالحرب أو الإضراب العام أو الثورة أو الفيضانات الخطيرة أو أى عامل من العوامل الطبيعية التى تؤثر على الإنتاج .

إذا تأملنا الآن فى قيمة أية ظاهرة فى كليتها لوجدنا أن هذا الكل هو عبارة عن « محصلة » مجموعة المؤثرات السابق ذكرها التى تشترك جميعها فى تكوين هذه القيمة .

فإذا رمزنا للقيمة الاتجاهية بالرمز ح

وللتغيرات الموسمية بالرمز م

وللتغيرات الدورية بالرمز ى

وللتغيرات العرضية بالرمز ع

فإن قيمة الظاهرة كما نشاهدها = ح × م × و × ع

ويمكن كذلك أن نعتبر أن القيمة المشاهدة للظاهرة = ح + م + و + ع

حيث تكون قيمة كل مؤثر مقدرة بدلالة الوحدات المطلقة . غير أن الطريقة الأولى أفضل لأنها تعطى لكل مؤثر أهميته النسبية .

على ضوء هذه المعلومات نستطيع أن نقول إن المقصود بتحليل السلسلة الزمنية هو رد القيمة الكلية للظاهرة إلى العناصر المكونة لها ، أو بعبارة أخرى عزل كل من المؤثرات الاتجاهية والموسمية والدورية والعرضية على حدة لمعرفة مدى تأثير كل منها على قيمة الظاهرة . وسنتناول فيما يلى الطرق المختلفة لحساب هذه المؤثرات .

طرق تعيين الاتجاه العام

أبسط طريقة لتعيين الاتجاه العام هى أن نرصد قيم الظاهرة فى الوحدات الزمنية المتتالية على رسم بيانى ثم نمهد الخط البيانى الناتج بطريقة تقديرية (أى نمهده باليد حسب تقديرنا) . ذلك أن هذا التمهيد يستبعد أثر التقلبات الدورية والموسمية والعرضية ، التى تظهر فى الرسم على شكل ذبذبات وتعاريج ، ويجعل الخط البيانى الذى يمثل الظاهرة عبارة عن شكل هندسى منتظم (إما مستقيم أو منحنى) يصف الاتجاه العام لتغير الظاهرة . وعلى الرغم من أن هذه الطريقة تمتاز بالسهولة والسرعة إلا أنها تتوقف كلية على التقدير الشخصى كما أنها لا تستند إلى أى أساس علمى ولذلك نكتفى بمجرد الإشارة إليها . أما الطرق الدقيقة لإيجاد الاتجاه العام . فسنشرحها فى شىء من التفصيل فيما يلى :

أولا : طريقة المتوسطات المتحركة Moving Averages

التقلبات الموسمية والدورية والعرضية التى تطرأ على قيمة الظاهر تكون أحياناً بالصعود وأحياناً بالهبوط . ولذلك فإذا أخذنا متوسط قيم الظاهرة فى عدة سنين

فإن التقلبات المتضادة الاتجاه تميل إلى التلاشى بتعادل الانحرافات السالبة والموجبة عن هذا المتوسط الذى نعتبره يمثل القيمة الاتجاهية للظاهرة . وأساس طريقة المتوسطات المتحركة هو اختيار فترة زمنية معينة يعتبر الإحصائى أن متوسط قيم الظاهرة خلالها يكون خالياً من التقلبات بأنواعها . ويتوقف اختيار هذه الفترة فى الغالب على طول فترة التقلبات الدورية كما تظهر فى الرسم البيانى(١) فإذا كان الخط البيانى الناتج من رصد قيم الظاهرة يتكون من موجات تمتد كل منها على فترة طولها ثلاث سنوات فإننا نختار متوسطا فترته ثلاث سنوات وإذا كانت الموجات تمتد على فترات طول كل منها خمس سنوات فإننا نختار متوسطاً فترته خمس سنوات وهكذا . فإذا ما انتهينا من تحديد فترة المتوسط (ولنفرض أنها ثلاث سنوات) نحسب المتوسطات المتحركة بالطريقة الآتية :

أولا : نجمع قيم الظاهرة فى الثلاثة سنوات الأولى ونكتب هذا المجموع أمام السنة الوسطى أى السنة الثانية فى هذه الحالة ، ثم نترك السنة الأولى ونأخذ مجموع السنوات الثلاث التالية (أى الثانية والثالثة والرابعة) ونكتب هذا المجموع أمام السنة الوسطى لهذه المجموعة أى أمام السنة الثالثة . ونستمر فى هذه العملية حتى يكون لدينا مجموع متحرك أمام كل سنة ما عدا السنة الأولى والأخيرة (انظر جدول رقم ٦٠)

ثانياً : نقسم كل مجموع متحرك على ثلاثة فنحصل على متوسطات متحركة لكل ثلاث سنوات . ونضع رقم المتوسط المتحرك أمام المجموع المتحرك كما يتضح من الجدول الآتى :

---

(١) التقلبات الموسمية لا تزيد فترة تكرارها عادة عن السنة ، فهى تحدث عادة فى شهر أو فصل معين من كل سنة أو فى فترة أقل من ذلك . ولذلك إذا أخذنا القيم السنوية للظاهرة فإن أثر التقلبات الموسمية لا يظهر فيها . وحيث أنه عند حساب الاتجاه العام نأخذ عادة القيم السنوية للظاهرة فإن التقلبات التى يتعين استبعادها هى التقلبات الدورية والعرضية . وحيث أن التقلبات العرضية تحدث خبط عشواء دون ما قاعدة فإن متوسط أى عدد من السنين كفيل بإزالتها .

جدول رقم ٦٠

استهلاك الحرير الصناعى فى الولايات المتحدة من سنة ١٩٢٨ إلى سنة ١٩٣٧
مقدراً بآلاف الأطنان

السنة	قيمة الظاهرة	مجموع متحرك فترة ٣ سنوات	متوسط متحرك فترته ٣ سنوات
١٩٢٨	١٠٠١٠١	٠٠٠٠٠٠	٠٠٠٠٠٠
١٩٢٩	١٣١٤٤٨	٣٤٩٥١٧	١١٦٥٠٦
١٩٣٠	١١٧٩٦٨	٤٠٦٧٧٦	١٣٥٥٩٢
١٩٣١	١٥٧٣٦٠	٤٢٧٣٦٩	١٤٢٤٥٦
١٩٣٢	١٥٢٠٤١	٥٢١٢٨٤	١٧٣٧٦١
١٩٣٣	٢١١٨٨٣	٥٥٨٦٩٥	١٨٦٢٣٢
١٩٣٤	١٩٤٧٧١	٦٥٩٣٣٠	٢١٩٧٧٧
١٩٣٥	٢٢٥٦٧٦	٧٤٥٠٤١	٢٤٨٣٤٧
١٩٣٦	٢٩٧٥٩٤	٨١١٤٦٥	٢٧٠٤٨٨
١٩٣٧	٢٦١١٩٥	٠٠٠٠٠٠	٠٠٠٠٠٠

المتوسطات المتحركة التى تشمل فترتها على عدد زوجى من السنين :

إذا أردنا حساب متوسط متحرك فترته سنتان نحسب المجاميع المتحركة لكل سنتين كالمعتاد . ولكن بما أنه فى هذه الحالة لا تكون هناك سنة وسطى نضع أمامها المجموع فإننا نلجأ إلى وضع المجموع المتحرك بين السنة الأولى والثانية وبين السنة الثانية والثالثة وهكذا . بعد ذلك نحسب المتوسطات المتحركة بقسمة كل مجموع على ٢ ونضع كل متوسط أمام المجموع المقابل له . وحيث أن هذه المتوسطات لا تكون مقابلة للسنوات التى تكون السلسلة الزمنية فإنها تكون قليلة الفائدة . وذلك لأن المقصود من حساب هذه المتوسطات المتحركة هو إيجاد القيم الاتجاهية المقابلة لكل سنة . وللتخلص من هذه الصعوبة نكرر العملية السابقة مرة أخرى فنأخذ مجموعاً متحركاً للمتوسطين الأول والثانى ثم

الثانى والثالث وهكذا . ونرصد هذه المجاميع الجديدة فى منتصف المسافة بين كل متوسطين من المتوسطات الأولى ثم نحسب متوسطات هذه المجاميع الجديدة (بالقسمة على ٢) ونضع كل منها أمام المجموع المقابل وبهذه الطريقة يكون كل متوسط موضوعاً أمام سنة معينة كما يتضح من الجدول الآتى :

جدول رقم ٦١

إيجاد الاتجاه العام للإيرادات الإجمالية لترام الرمل (*)
عن طريق حساب متوسط متحرك فترته سنتان

السنة	الإيرادات	مجموع متحرك فترته سنتان	متوسط متحرك فترته سنتان	مجموع متحرك فترته سنتان للمتوسط السابق	متوسط متحرك معدل فترته سنتان
	جنيه				
١٩٢٩	٢٢١٤٩٥	٤٣٦٤٣٤	٢١٨٢١٧		
١٩٣٠	٢١٤٩٣٩	٤٠٥٦٦٦	٢٠٢٨٣٣	٤٢١٠٥٠	٢١٠٥٢٥
١٩٣١	١٩٠٧٢٧	٣٦٨١٢٢	١٨٤٠٦١	٣٨٦٨٩٤	١٩٣٤٤٧
١٩٣٢	١٧٧٣٩٥	٣٤٨٧٩٥	١٧٤٣٩٧,٥	٣٥٨٤٥٨,٥	١٧٩٢٢٩,٢٥
١٩٣٣	١٧١٤٠٠	٣١٩٧٧٧	١٥٩٨٨٨,٥	٣٣٤٢٨٦	١٦٧١٤٣
١٩٣٤	١٤٨٣٧٧	٢٧٩٥٣٥	١٣٩٧٦٧,٥	٢٩٩٦٥٦	١٤٩٨٢٨,
١٩٣٥	١٣١١٥٨	٢٦٠٠٤٦	١٣٠٠٢٣	٢٦٩٧٩٠,٥	١٣٤٨٩٥,٢٥
١٩٣٦	١٢٨٨٨٨	٢٧٢١٨٥	١٣٦٠٩٢,٥	٢٦٦١١٥,٥	١٣٣٠٥٧,٧٥
١٩٣٧	١٤٣٢٩٧				

ومن الممكن طبعاً اختصار خطوات العمل السابقة بعدم حساب المتوسط المتحرك الأول والاكتفاء باستخراج المجموع المتحرك الأول ثم حساب مجموع متحرك فترته سنتان لهذا المجموع ثم قسمة المجاميع الأخيرة على ٤ كما يتضح من الجدول الآتى :

(*) حصلنا على هذه البيانات من قسم الإحصاء بإدارة النقل المشترك لمنطقة الاسكندرية .

جدول رقم ٦٢

إيجاد متوسط متحرك فترته سنتان بالطريقة المختصرة

السنة	الإيرادات	مجموع متحرك فترته سنتان	مجموع متحرك فترته سنتان للمجموع السابق	متوسط متحرك معدل فترته سنتان
	جنيه			
١٩٢٩	٢٢١٤٩٥			
		٤٣٦٤٣٤		
١٩٣٠	٢١٤٩٣٩		٨٤٢١٠٠	٢١٠٥٢٥
		٤٠٥٦٦٦		
١٩٣١	١٩٠٧٢٧		٧٧٣٧٨٨	١٩٣٤٤٧
		٣٦٨١٢٢		
١٩٣٢	١٧٧٣٩٥		٧١٦٩١٧	١٧٩٢٢٩,٢٥
		٣٤٨٧٩٥		
١٩٣٣	١٧١٤٠٠		٦٦٨٥٧٢	١٦٧١٤٣
		٣١٩٧٧٧		
١٩٣٤	١٤٨٣٧٧		٥٩٩٣١٢	١٤٩٨٢٨
		٢٧٩٥٣٥		
١٩٣٥	١٣١١٥٨		٥٣٩٥٨١	١٣٤٨٩٥,٢٥
		٢٦٠٠٤٦		
١٩٣٦	١٢٨٨٨٨		٥٣٢٢٣١	١٣٣٠٥٧,٧٥
		٢٧٢١٨٥		
١٩٣٧	١٤٣٢٩٧			

ويلاحظ أنه عند حساب متوسط متحرك فترته ٤ سنوات أو ٨ سنوات نتبع نفس الطريقة السابقة فنحسب أولا مجموعاً متحركاً فترته ٤ سنوات (أو ٨ سنوات) ثم نحسب لهذا المجموع مجموعاً متحركاً آخر فترته سنتان ونقسم الناتج على ٨ (أو على ١٦) لنحصل على المتوسط المتحرك المعدل . ومن هذا يتضح أن المجموع المتحرك الأول تكون فترته دائماً مماثلة لفترة المتوسط المتحرك المراد حسابه ، أما المجموع المتحرك الثانى الذى يراد به تعديل مراكز المجاميع الأولى وجعلها مقابلة للسنوات ، فإن فترته تكون سنتين فى كل حالة . والجدول الآتى يوضح طريقة حساب متوسط متحرك معدل فترته ٤ سنوات .

جدول رقم ٦٣

إيجاد متوسط متحرك فترته ٤ سنوات بالطريقة المختصرة

السنة	قيم الظاهرة	مجموع متحرك فترته ٤ سنوات	مجموع متحرك فترته سنتان للمجاميع السابقة	متوسط متحرك معدل فترته ٤ سنوات
١٩١٩	٩٢٩١			
١٩٢٠	٨٧١٨	٦٢٥٠٧		
١٩٢١	١٩٧٥١	٨٥٧٧٤	١٤٨٢٨١	١٨٥٣٥
١٩٢٢	٢٤٧٤٧	١١٩٢٩٩	٢٠٥٠٧٣	٢٥٦٣٤
١٩٢٣	٣٢٥٥٨	١٥٧٨٢٥	٢٧٧١٢٤	٣٤٦٤٠
١٩٢٤	٤٢٢٤٣			
١٩٢٥	٥٨٢٧٧			

ومما تجب ملاحظته أنه كلما كان عدد السنين التى ندرسها كبيراً كلما كانت القيم الاتجاهية أقرب إلى الدقة . وذلك لأن الانحرافات السالبة عن الاتجاه العام تميل فى المدة الطويلة إلى التعادل مع الانحرافات الموجبة . كذلك كلما كانت المدة طويلة كان عدد القيم الاتجاهية التى نحصل عليها كبيراً . وذلك لأن طريقة المتوسطات المتحركة لا تعطينا قيما اتجاهية لجميع السنين . فلو كنا مثلا نحسب متوسطاً متحركاً فترته ٣ سنوات لاضطررنا إلى ترك السنة الأولى والأخيرة بدون قيمة اتجاهية . وعلى وجه العموم إذا كانت فترة المتوسط عددا فردياً ($ن$ مثلا) كان عدد القيم الاتجاهية الناقصة فى طرفى السلسلة الزمنية $= \frac{ن - ١}{٢}$. وإذا كان الفترة زوجية كان عدد القيم الناقصة فى طرفى السلسلة $= \frac{ن}{٢}$

وقبل أن نترك طريقة المتوسطات المتحركة لإيجاد الاتجاه العام لا بد أن نناقش الأساس الذى تستند إليه لنختبر مدى صحته . هل حساب المتوسطات

المتحركة يستبعد فعلا أثر جميع التقلبات التى تطرأ على قيمة الظاهرة ؟

لوكان هذا صحيحاً فإذا أخذنا سلسلة زمنية خالية تماماً من التقلبات وحسبنا لها متوسطاً متحركاً أيا كانت فترته يجب أن نحصل على نفس القيم الأصلية للسلسلة الزمنية (ما دام ليس هناك تقلبات يزيلها حساب المتوسط) .

ومن الجهة الأخرى إذا أخذنا سلسلة زمنية مكونة من تقلبات منتظمة فقط وحسبنا لها متوسطاً متحركاً تتفق فترته مع الدورة التى تحدث بها هذه التقلبات فإن المتوسطات الناتجة لا بد أن تكون جميعها تساوى صفراً .

والواقع أن كلا الاختبارين يؤيد الفرض القائل بأن حساب المتوسطات المتحركة يزيل أثر التقلبات . والتثبت من صحة ذلك نأخذ أولا سلسلة زمنية خالية تماماً من التقلبات ونحسب لها متوسطات متحركة من فترات مختلفة على النحو المبين بالجدول الآتى :

جدول رقم ٦٤

السنة	الظاهرة	مجموع متحرك فترته ٥ سنوات	متوسط متحرك فترته ٥ سنوات	مجموع متحرك فترته ٧ سنوات	متوسط متحرك فترته ٧ سنوات
١٩٢٠	١				
١٩٢١	٣				
١٩٢٢	٥	٢٥	٥		
١٩٢٣	٧	٣٥	٧	٤٩	٧
١٩٢٤	٩	٤٥	٩	٦٣	٩
١٩٢٥	١١	٥٥	١١	٧٧	١١
١٩٢٦	١٣	٦٥	١٣	٩١	١٣
١٩٢٧	١٥	٧٥	١٥		
١٩٢٨	١٧				
١٩٢٩	١٩				

يتضح من الجدول السابق أنه إذا أخذنا سلسلة زمنية خالية تماماً من التقلبات وحسبنا لها متوسطاً متحركاً فإننا نحصل على نفس قيم السلسلة مرة ثانية[1] . والآن نأخذ سلسلة زمنية أخرى مكونة من تقلبات منتظمة فقط ثم نحسب لها متوسطاً متحركاً فترته تناسب هذه التقلبات كما هو مبين بالجدول الآتى :

جدول رقم ٦٥

السنة	الظاهرة	مجموع متحرك فترته ٧ سنوات	متوسط متحرك فترته ٧ سنوات
١٩٢٠	+٢		
١٩٢١	+٢		
١٩٢٢	٠		
١٩٢٣	—١	٠	٠
١٩٢٤	—٢	٠	٠
١٩٢٥	—٢	٠	٠
١٩٢٦	+١	٠	٠
١٩٢٧	+٢	٠	٠
١٩٢٨	+٢	٠	٠
١٩٢٩	٠	٠	٠
١٩٣٠	—١	٠	٠
١٩٣١	—٢		
١٩٣٢	—٢		
١٩٣٣	+١		

(١) هذه النتيجة تكون صحيحة إذا كانت قيم الظاهرة يمثلها مستقيم كما هو الحال فى المثال السابق . أما إذا كان الخط البيانى الذى يمثل الظاهرة منحنياً فإن المتوسطات المتحركة تقلل من درجة انحناء المنحنى . وعلى وجه العموم فإن طريقة المتوسطات المتحركة لا تصلح لإيجاد الاتجاه العام إذا كان هذا الاتجاه منحنياً .

يتضح مما تقدم أن طريقة المتوسطات المتحركة تزيل فعلا أثر التقلبات التى تطرأ على قيمة الظاهرة .

ثانياً : إيجاد الاتجاه العام بتوفيق مستقيم أو منحن :

يمكن إيجاد الاتجاه العام للسلسلة الزمنية بتوفيق مستقيم أو منحن لها بطريقة المربعات الدنيا التى شرحناها فى الباب الخاص بتوفيق المنحنيات[1] . فإذا نظرنا إلى الرسم الذى يمثل مصروفات إدارة النقل المشترك لمنطقة الإسكندرية



(شكل رقم ٥٥)

مصروفات إدارة النقل المشترك لمنطقة الإسكندرية

(شكل ٥٥) لتبين لنا فى الحال أن الاتجاه العام لهذه الظاهرة يصفه منحن من الدرجة الثانية . وبتوفيق منحن من الدرجة الثانية لهذه الظاهرة نجد توافقاً كبيراً بين القيم المشاهدة والقيم النظرية (التى تمثل الاتجاه العام) .

---

(١) راجع صفحة ٢٨٠ من هذا الكتاب .

والجدول الآتى يوضح طريقة توفيق المنحنى الذى يصف الاتجاه العام لمصروفات إدارة النقل المشترك بمنطقة الإسكندرية .

جدول رقم ٦٦

توفيق منحن من الدرجة الثانية لمصروفات إدارة النقل المشترك بالإسكندرية

السنة	س	المصروفات بآلاف الجنيهات	س ص	س٢ص	س٤	القيم الاتجاهية
١٩٢٩	ـ ٨	١٣٠,٣	ـ١٠٤٢,٤	٨٣٣٩,٢	٤٠٩٦	١٤٨,٨٨
١٩٣٠	ـ ٧	١٣٢,٥	ـ٩٢٧,٥	٦٤٩٢,٥	٢٤٠١	١٣٢,٩٢
١٩٣١	ـ ٦	١٣١,٦	ـ٧٨٩,٦	٤٧٣٧,٦	١٢٩٦	١٢٢,٥٤
١٩٣٢	ـ ٥	١٢٢,٠	ـ٦١٠,٠	٣٠٥٠,٠	٦٢٥	١١٧,٧٤
١٩٣٣	ـ ٤	١٢٧,٥	ـ٥١٠,٠	٢٠٤٠,٠	٢٥٦	١١٨,٥٢
١٩٣٤	ـ ٣	١١٩,٤	ـ٣٤٨,٢	١٠٤٤,٦	٨١	١٢٤,٩٢
١٩٣٥	ـ ٢	١٤٩,٠	ـ٢٩٨,٠	٥٩٦,٠	١٦	١٣٦,٨٢
١٩٣٦	ـ ١	١٥٤,٣	ـ١٥٤,٠	١٥٤,٣	١	١٥٤,٣٤
١٩٣٧	٠	١٩٧,٨	٠٠٠٠	٠٠٠٠	٠٠٠٠	١٧٧,٤٤
١٩٣٨	١	٢٠٦,٧	٢٠٦,٧	٢٠٦,٧	١	٢٠٦,١٢
١٩٣٩	٢	٢٤٨,٥	٤٩٧,٠	٩٩٤,٠	١٦	٢٤٠,٣٨
١٩٤٠	٣	٢٧٣,٨	٨٢١,٤	٢٤٦٤,٢	٨١	٢٨٠,٢٢
١٩٤١	٤	٢٩٢,٦	١١٧٠,٤	٤٦٨١,٦	٢٥٦	٣٢٥,٦٤
١٩٤٢	٥	٢٥٠,٩	١٧٥٤,٥	٨٧٧٢,٥	٦٢٥	٣٧٦,٦٤
١٩٤٣	٦	٣٤٧,٨	٢٥٦٦,٨	١٥٤٠٠,٨	١٢٩٦	٤٣٣,٢٢
١٩٤٤	٧	٢٩٣,٠	٣٤٥١,٠	٢٤١٥٧,٠	٢٤٠١	٤٩٥,٣٨
١٩٤٥	٨	٥٩٧,١	٤٧٧٦,٨	٣٨٢١٤,٤	٤٠٩٦	٥٦٣,١٢
المجموع	٠٠	٤١٥٤,٨	١٠٥٦٤,٦	١٢١٣٤٥,٤	١٧٥٤٤	٤١٥٤,٨٤

معادلة المنحنى هى ص = ا + ب س + حـ س^{٢}

ويمكن إيجاد قيم ا ، ب ، حـ من المعادلات الطبيعية الآتية :

$$\Sigma ص = ن ا + ب \Sigma س + حـ \Sigma س^{٢}$$

$$\Sigma س ص = ا \Sigma س + ب \Sigma س^{٢} + حـ \Sigma س^{٣}$$

$$\Sigma س^{٢} ص = ا \Sigma س^{٢} + ب \Sigma س^{٣} + حـ \Sigma س^{٤}$$

ويلاحظ بالنظر إلى الجدول رقم ٦٦ أننا اختصرنا خطوات العمل كثيراً بجعل نقطة الأصل عند السنة الوسطى (١٩٣٧) فأصبح كل من :

Σ س ، Σ س^{٣} = صفرا .

كما يلاحظ أننا لم نعمل عموداً خاصاً بقيم س^{٢} وذلك لأنه ما دامت قيم س قبل وبعد نقطة الأصل عبارة عن الأعداد الطبيعية من ١ إلى ٨ فيمكن إيجاد قيمة Σ س^{٢} باستخدام القانون الآتى :

$$\sum_{١}^{ن} س^{٢} = \frac{ن (ن + ١) (٢ن + ١)}{٦}$$

$$\therefore \sum_{١}^{٨} س^{٢} = \frac{٨ \times ٩ \times ١٧}{٦} = ٢٠٤$$

$\therefore$ Σ س^{٢} فى المعادلات الطبيعية = ٢٠٤ × ٢ = ٤٠٨

وبالتعويض فى المعادلات الطبيعية الثلاثة نجد أن :

$$٤١٥٤,٨ = ١٧ ا + ٤٠٨ حـ \quad (١)$$

$$١٠٥٦٤,٦ = ٤٠٨ ب \quad (٢)$$

$$١٢١٣٤٥,٤ = ٤٠٨ ا + ١٧٥٤٤ حـ \quad (٣)$$

وبضرب المعادلة (١) فى ٢٤

$$\therefore \quad ٩٩٧١٥,٢ = ٤٠٨ ا + ٩٧٩٢ حـ \quad (٤)$$

وبطرح المعادلة (٤) من المعادلة (٣)

∴ ٢١٦٣٠,٢ = ٧٧٥٢ حـ

∴ حـ = ٢,٧٩

وبالتعويض عن قيمة حـ فى المعادلة (١)

∴ ٤١٥٤,٨ = ١٧ ا + ٤٠٨ × ٢,٧٩

∴ ا = ١٧٧,٤٤

∴ معادلة المنحنى هى ص = ١٧٧,٤٤ + ٢٥,٨٩ س × ٢,٧٩ س^{٢}

حيث نقطة الأصل عند سنة ١٩٣٧

وبالتعويض عن قيم س فى هذه المعادلة نحصل على القيم الاتجاهية الموضحة بالعمود الأخير من الجدول رقم ٦٦

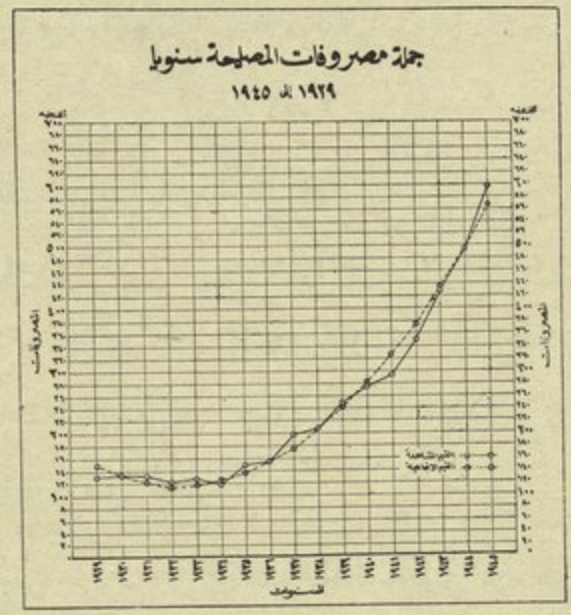
فمثلا للحصول على القيمة الاتجاهية للظاهرة فى سنة ١٩٣٥ نجرى العمل كالآتى :

ص$_{١٩٣٥}$ = ١٧٧,٤٤ − ٢ × ٢٥,٨٩ + ٢,٧٩ (− ٢)^{٢}

= ١٣٦,٨٢

وهى القيمة الاتجاهية المبينة بالجدول أمام سنة ١٩٣٥

والرسم الآتى يوضح نتيجة توفيق المنحنى لمصروفات إدارة النقل المشترك بمنطقة الإسكندرية ومنه يظهر أن هناك توافقاً كبيراً بين القيم المشاهدة والقيم المحسوبة من معادلة المنحنى مما يدل على أن هذا المنحنى يصف الاتجاه العام للمصروفات بدقة .



(شكل رقم ٥٦)

توفيق منحن لمصروفات إدارة النقل المشترك بمنطقة الإسكندرية

ويلاحظ أنه فى حالة كون السلسلة الزمنية تشتمل على عدد زوجى من السنين فلا يمكن اختصار خطوات العمل بالطريقة السابقة وذلك لعدم وجود سنة وسطى نعتبرها نقطة الأصل . وللتغلب على هذه الصعوبة نعتبر نقطة الأصل فى منتصف المسافة الواقعة بين السنتين المتوسطتين ونجعل وحدات س تمثل نصف سنة كما هو مبين بالجدول الآتى الذى يوضح خطوات العمل فى توفيق مستقيم بطريقة المربعات الدنيا فى حالة كون عدد السنين زوجياً .

جدول رقم ٦٧

توفيق مستقيم بطريقة المربعات الدنيا للأرقام القياسية لودائع صندوق التوفير
بالقطر المصرى (الطريقة المختصرة)

السنة	الرقم القياسى للودائع ص	س	س٢	س ص	القيم الاتجاهية
١٩١٩	٢٢١٥	ـ ١١	١٢١	ـ ٢٤٣٦٥	٢٢١٨,٠٧
١٩٢٠	٢٤٢٦	ـ ٩	٨١	ـ ٢١٨٣٤	٢٤٧٩,٣٩
١٩٢١	٢٦٦٩	ـ ٧	٤٩	ـ ١٨٦٨٣	٢٧٤٠,٧١
١٩٢٢	٢٩٠٢	ـ ٥	٢٥	ـ ١٤٥١٠	٣٠٠٢,٠٣
١٩٢٣	٣٣٧١	ـ ٣	٩	ـ ١٠١١٣	٣٢٦٣,٣٥
١٩٢٤	٣٧٠٠	ـ ١	١	ـ ٣٧٠٠	٣٥٢٤,٦٧
نقطة الأصل	٠٠	٠٠	٠٠	٠٠	٣٦٥٥,٣٣
١٩٢٥	٣٨٢٥	١	١	٣٨٢٥	٣٧٨٥,٩٩
١٩٢٦	٣٩٧٩	٣	٩	١١٩٣٧	٤٠٤٧,٣١
١٩٢٧	٤٣٤٣	٥	٢٥	٢١٧١٥	٤٣٠٨,٦٣
١٩٢٨	٤٦٦٠	٧	٤٩	٣٢٦٢٠	٤٥٦٩,٥٩
١٩٢٩	٤٨٣٥	٩	٨١	٤٣٥١٥	٤٨٣١,٢٧
١٩٣٠	٤٩٣٩	١١	١٢١	٥٤٣٢٩	٥٠٩٢,٥٩
المجموع	٤٣٨٦٤	٠٠	٥٧٢	٧٤٧٣٦	٤٣٨٦٣,٩٦

المعادلات الطبيعية اللازمة لتوفيق مستقيم بطريقة المربعات الدنيا هى :

$$\Sigma ص = ن ا + ب \Sigma س$$

$$\Sigma س ص = ا \Sigma س + ب \Sigma س^{٢}$$

وحيث أن $\Sigma س = صفر$

$$\therefore \text{ا} = \frac{\Sigma \text{ص}}{\text{ن}} = \frac{٤٣٨٦٤}{١٢} = ٣٦٥٥,٣٣$$

$$\text{، ب} = \frac{\Sigma \text{س ص}}{\Sigma \text{س}^{٢}} = \frac{٧٤٧٣٦}{٥٧٢} = ١٣٠,٦٦$$

∴ معادلة خط الاتجاه هى :

ص = ٣٦٥٥,٣٣ + ١٣٠,٦٦ س

حيث نقطة الأصل واقعة بين سنتى ١٩٢٤ ، ١٩٢٥ ووحدات س تمثل نصف سنة .

فمثلا للحصول على القيمة الاتجاهية لسنة ١٩٢٧ نجرى العمل كالآتى :

$$\text{ص}_{١٩٢٧} = ٣٦٥٥,٣٣ + ٥ \times ١٣٠,٦٦ = ٤٣٠٨,٦٣$$

وبمقارنة القيم الاتجاهية التى حصلنا عليها بهذه الطريقة المختصرة نجد أنها هى نفس القيم التى حصلنا عليها بتوفيق مستقيم لهذه الظاهرة بالطريقة المطولة (راجع صفحة ٢٨٢) . وإذا كانت معادلة هذا المستقيم تختلف عن المعادلة التى حصلنا عليها بالطريقة المطولة فإن ذلك يرجع إلى اختلاف نقطة الأصل واختلاف وحدات س فى الحالتين .

ويمكن أن نستخدم معادلة المستقيم أو المنحنى الذى يصف الاتجاه العام للظاهرة فى التنبؤ بما ستكون عليه قيمة الظاهرة فى المستقبل على وجه التقريب وعلى فرض ثبات الظروف المحيطة بهذه الظاهرة .

فإذا أردنا أن نحصل على القيمة الاتجاهية للرقم القياسى لودائع صندوق التوفير فى سنة ١٩٣٥ نستخرج قيمة س المقابلة لسنة ١٩٣٥ فنجد أنها تساوى ٢١ (لاحظ أن نقطة الأصل بين سنتى ١٩٢٤ ، ١٩٢٥ ووحدات س تمثل نصف سنة)

وبوضع س = ٢١ فى معادلة المستقيم نجد أن :

$$\text{ص}_{١٩٣٥} = ٣٦٥٥,٣٣ + ٢١ \times ١٣٠,٦٦ = ٦٣٩٩,١٩$$

وما من شك فى أن طريقة إيجاد الاتجاه العام بتوفيق مستقيم أو منحن أفضل وأعم من طريقة المتوسطات المتحركة . فهذه الطريقة الأخيرة لا تصلح لإيجاد الاتجاه العام إذا كان تغير الظاهرة فيه انحناء ظاهر . أما طريقة توفيق المنحنيات فإنها تصلح لوصف اتجاه تغير الظواهر أيا كان شكل هذا الاتجاه . وسيجد القارئ من مراجعة الباب الخاص بتوفيق المنحنيات أن هناك مجموعة كبيرة من المنحنيات التى يمكن اختيار أحدها لوصف الاتجاه العام لأية ظاهرة . على أن دقة النتائج التى نحصل عليها بهذه الطريقة تتوقف كلية على اختيار نوع المنحنى الذى يناسب تغير الظاهرة ولذلك فإن تطبيقها يتطلب إلماماً تاماً بعدد كبير من المنحنيات التى يصعب فهم معادلاتها على الطالب المبتدئ .

طرق حساب التغيرات الموسمية

هناك عدة طرق لحساب التغيرات الموسمية غير أننا سنكتفى هنا بشرح طريقتين منها فقط . والطريقة الأولى تعطينا التغيرات الموسمية على شكل نسب مئوية تسمى النسب الموسمية تظهر بمجرد النظر إليها درجة تأثر قيمة الظاهرة بالموسم .

والطريقة الثانية تعطينا التغيرات الموسمية مقدرة بوحدات القياس العادية . وفيما يلى شرح موجز لهاتين الطريقتين :

١ — إيجاد النسب الموسمية بطريقة المتوسطات البسيطة :

لشرح هذه الطريقة افرض أننا أردنا أن نعرف أى الأيام الأسبوع تزداد فيه عادة حركة النقل على خط ترام الرمل عن المتوسط وأى الأيام تقل فيه هذه الحركة عادة عن المتوسط . وافرض أننا اعتبرنا أن جملة الإيرادات المتحصلة يومياً من بيع التذاكر تعبر بدقة عن حركة النقل بالترام[1] . أول طريقة تتبادر

(١) من الواضح أن هذا الفرض لا يمكن اعتباره صحيحاً على وجه الإطلاق نظراً لوجود عدد كبير من الركاب الذين يحملون اشتراكات أو كارنيهات . غير أنه إذا اعتبرنا أن استعمال هؤلاء لمركبات الترام لا يتغير كثيراً من يوم لآخر فإن فرضنا لا يكون بعيداً عن الصحة .

إلى الذهن لمقارنة حركة النقل فى الأيام المختلفة هى أن نحسب متوسط الإيرادات المتحصلة فى كل يوم من أيام الأسبوع على حدة خلال فترة طويلة نسبياً من الزمن ثم نقارن هذه المتوسطات اليومية . غير أنه لكى تكون المقارنة سهلة وواضحة نتخذ أساساً نعتبره يمثل حركة النقل العادية . فإذا كان المتوسط اليومى أكبر من هذا الأساس ظهر لنا فى الحال أن تأثير الموسم صعودياً فى هذا اليوم . وبالمثل إذا كان المتوسط اليومى أقل من هذا الأساس دل ذلك على أن قيمة الظاهرة تقل بتأثير الموسم فى هذا اليوم . ومن الواضح أن أفضل طريقة لاختيار هذا الأساس هى أن نحسب الوسط الحسابى للمتوسطات اليومية ونعتبره يمثل حركة النقل العادية . كما أن المقارنة تكون أسهل وأكثر وضوحاً إذا حسبنا متوسط كل يوم كنسبة مئوية من هذا المتوسط العام (الوسط الحسابى للمتوسطات اليومية) الذى نعتبره أساساً للمقارنة . وهذه الطريقة السابقة هى طريقة المتوسطات البسيطة لإيجاد النسب الموسمية . ويتضح مما تقدم أن خطوات العمل اللازمة لإيجاد النسب الموسمية بهذه الطريقة هى الآتية :

أولا : نحسب المتوسطات اليومية إذا كان التأثير الموسمى يظهر فى يوم معين من كل أسبوع أو المتوسطات الشهرية إذا كان التأثير الموسمى يظهر فى شهر معين من كل سنة وهكذا بالنسبة لأية وحدة زمنية أخرى يظهر فيها التأثير الموسمى .

ثانياً : نحسب الوسط الحسابى للمتوسطات اليومية (أو الشهرية) .

ثالثا : نحسب كل متوسط من المتوسطات اليومية (أو الشهرية) كنسبة مئوية من الوسط الحسابى لهذه المتوسطات فتكون هذه النسب المئوية هى النسب الموسمية المطلوبة . .

ولإيضاح هذه الخطوات عملياً نحسب النسب الموسمية اليومية لإيرادات التذاكر فى خط ترام الرمل بمدينة الإسكندرية .

جدول رقم ٦٨

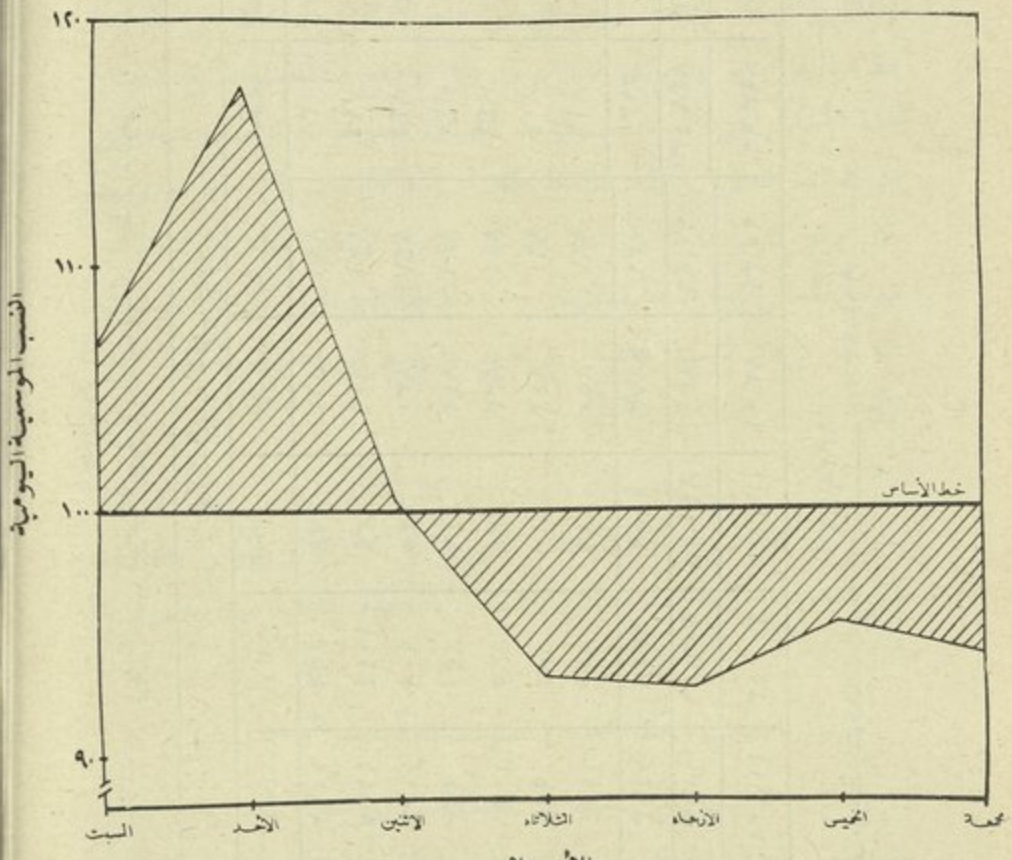
إيجاد النسب الموسمية اليومية لإيرادات تذاكر خط ترام الرمل

رقم الأسبوع	الإيرادات اليومية لأقرب جنيه							المجموع الأسبوعى
	الأحد	الاثنين	الثلاثاء	الأربعاء	الخميس	الجمعة	السبت	
١	١٠٤٢	٨٦٢	٧٧٨	٨٦٧	٨٨٦	١٠٣١	١١٥١	٦٦١٧
٢	١٢٣٧	١٠٤٤	٨٣٩	٨٠٩	٨٢١	٧٩٤	٩٣٠	٦٤٧٤
٣	١٠٣٥	٩٠٩	٨٢٥	٨٣٠	٨٤٧	٨٢١	٨٥٩	٦١٢٦
٤	٩٤٣	٨٩٠	٨٤٦	٨٤٢	٨٨١	٨٢٧	٩٣٨	٦١٦٧
٥	١٠٥٨	٨٩٥	٨٣٥	٨٣٣	٨٤٤	٨٤٦	٩٧٥	٦٢٨٦
٦	٩٩٤	٨٥٥	٨٨٠	٨٤٢	٨٦٠	٨٠٤	٩٢٥	٦١٦٠
٧	١١٠٤	٨٩٣	٨٨٦	٨٥٠	٨٨٧	٨١٦	٩٨٠	٦٤١٦
المجموع	٧٤١٣	٦٣٤٨	٥٨٨٩	٥٨٧٣	٦٠٢٦	٥٩٣٩	٦٧٥٨	٤٤٢٤٦
المتوسطات اليومية	١٠٥٩	٩٠٦,٨٦	٨٤١,٢٨	٨٣٩	٨٦٠,٨٦	٨٤٨,٤٣	٩٦٥,٤٣	٦٣٢٠,٨٦ ٥
النسب الموسمية اليومية	١١٧,٢٨	١٠٠,٤٣	٩٣,١٧	٩٢,٩١	٩٥,٣٤	٩٣,٩٦	١٠٦,٩١	٧٠٠

٥ لاحظ أن متوسط المتوسطات اليومية = $\frac{٦٣٢٠,٨٦}{٧}$ = ٩٠٢,٩٤

يتضح من الجدول السابق أن يوم الأحد هو اليوم الذى تبلغ فيه إيرادات التذاكر فى خط ترام الرمل حدها الأقصى . أما يوم الإثنين فإن حركة النقل فيه عادية لأن الإيرادات التى تحصل عادة فى هذا اليوم قريبة جداً من المتوسط أما بقية أيام الأسبوع (فيما عدا يوم السبت) فإن الحركة فيها أقل من المتوسط .

والرسم الآتى يوضح هذه النسب الموسمية اليومية بيانا :



(شكل رقم ٥٧)

النسب الموسمية اليومية لإيرادات التذاكر فى خط ترام الرمل بمدينة الإسكندرية

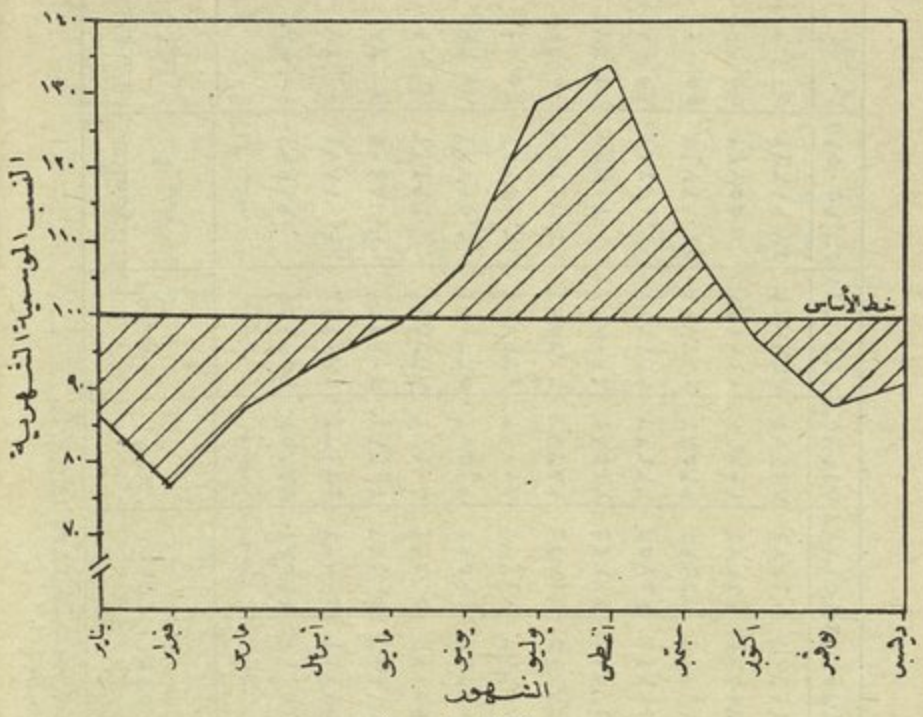
ولزيادة إيضاح هذه الطريقة نوجد النسب الموسمية الشهرية لإيرادات التذاكر بخط ترام الرمل .

جدول رقم ٦٩

إيجاد النسب الموسمية الشهرية لإيرادات تذاكر خط ترام الرمل

الأشهر	١٩٤٠	١٩٤١	١٩٤٢	١٩٤٣	١٩٤٤	١٩٤٥	١٩٤٦	١٩٤٧	المجموع	المتوسطات الشهرية	النسب الموسمية الشهرية
	جنيه	جنيه	جنيه	جنيه	جنيه	جنيه	جنيه	جنيه	جنيه	جنيه	٪
يناير	٧٥٢٣	١٠٠٥٢	١٥٣٠٥	١٧٥٨٩	٢٥٦٩٠	٢٩٤١٤	٢٨٥٩٢	٢٥١٨٧	١٥٩٣٥٢	١٩٩١٩	٨١,٠١
فبراير	٦٧٥٨	٩٢٢٤	١٤٢٠٦	١٥٥٩٥	٢٣٩٠٤	٢٦١١٩	٢٣٢٤٧	٢٣٤٣٤	١٤٢٤٨٧	١٧٨١٠,٩	٧٢,٤٤
مارس	٧٧١١	١١٠٤٩	١٦٣٦٠	١٧٥٩٣	٢٧٤٩٠	٣٠٦١٣	٢٦٤٨٥	٢٦٦٠٢	١٦٣٩٠٣	٢٠٤٨٧,٩	٨٣,٣٣
أبريل	٧٩٤١	١٢٣٣٣	١٧٧١٤	٢١٩٣٥	٢٩٨٥٦	٣٢٠٥٩	٢٩٤٢٥	٢٧١٠٥	١٧٨٣٦٨	٢٢٢٩٦	٩٠,٦٨
مايو	٨٠٤١	١٣٢٥٠	١٩٥٤٩	٢٤٩٥٠	٣١١٧٢	٣٦٢٧٠	٢٩٥٩١	٢٧٥٤٢	١٩٠٣٦٥	٢٣٧٩٥,٦	٩٦,٧٨
يونيه	٨١٢٣	١٦٤٤٦	٢٠٠٠٨	٢٦٣٣٧	٣٥٠٤٧	٣٨٦٥٨	٣٣٢٤٠	٢٩٧٧٨	٢٠٧٦٣٧	٢٥٩٥٤,٦	١٠٥,٥٦
يوليه	٨٠٦٣	١٦٢١٢	١٨٨٨٠	٣٤٤٢٦	٤٦٤٤٤	٥٠٠٨٥	٤٣٨٠٧	٣٧٣٨٩	٢٥٥٣٠٦	٣١٩١٣,٢٥	١٢٩,٨٠
أغسطس	٨٥٤٤	١٧٣٤٩	٢٢١٩٢	٣٩٩١١	٧٤٣٣٧	٤٨٥٣٤	٤٢٦٢٢	٤١٥٢٧	٢٦٨٠١٦	٣٣٥٠٢	١٣٦,٢٦
سبتمبر	٨٤٧٣	١٧٢٩٥	١٩٤٣٨	٣٠٨٢٢	٣٩٥١٢	٤٢٢١٤	٣٥٧٤٩	٣٢٦٦٦	٢٢٦١٦٩	٢٨٢٧١,١	١١٤,٩٨
أكتوبر	٨٦٨٢	١٥٥٤٥	١٩٦٦١	٢٩٥٢٦	٣٣٦١٣	٣٢٧٤١	٢٩٤٣٣	٢٨٥٦٩	١٩٧٧٧٠	٢٤٧٢١,٢٥	١٠٠,٥٥
نوفمبر	٨٧١٦	١٤٤٠٠	١٧٨٩٤	٢٦٤٤٠	٣٠٥٥٥	٢٩٥١٣	٢٧٦٢١	٢٦٩٠١	١٨٢٠٤٠	٢٢٧٥٥	٩٢,٥٥
ديسمبر	٩٦٠٩	١٥١٣٩	١٩٣٢٢	٢٩٣٢١	٣٠٧٦٧	٣٠٣١٢	٢٧٢٤٠	٢٧٢٤٨	١٨٨٩٥٨	٢٣٦١٩,٧٥	٩١,٠٧
المجموع	٩٨١٨٤	١٦٨٢٩٤	٢٢٠٥٢٩	٣١٤٤٤٥	٤٠١٣٨٧	٤٢٦٥٣٢	٣٧٧٠٥٢	٣٥٣٩٤٨	٢٣٦٠٣٧١	٢٩٥٠٤٦,٣٥	١٢٠٠,٠١

(٣٤)



(شكل رقم ٥٨)

النسب الموسمية الشهرية لإيرادات تذاكر خط ترام الرمل بمدينة الإسكندرية بعد استبعاد تأثير الاتجاه العام منها

يتضح بالنظر إلى النسب الموسمية المبينة بالجدول رقم ٦٩ أو إلى الرسم البيانى الذى يوضح هذه النسب (شكل رقم ٥٨) أن التأثير الموسمى يظهر بشكل واضح جداً فى أشهر الصيف (يوليه وأغسطس وسبتمبر) وهذه هى النتيجة التى نتوقعها طبعاً نظراً لوقوع موسم الاصطياف فى هذه الأشهر .

استبعاد الاتجاه العام من النسب الموسمية

إذا كانت قيم الظاهرة خاضعة لتأثير الاتجاه العام كما هو المعتاد فلا بد من استبعاد هذا الاتجاه عند حساب النسب الموسمية . وذلك لأنه إذا كان الاتجاه العام للظاهرة هو نحو الصعود فإن بعض النسب الموسمية تكون أكثر مما يجب أن تكون . وبالعكس إذا كان اتجاه الظاهرة نحو الهبوط فإن بعض

هذه النسب تكون أقل مما يجب أن تكون . ولتوضيح ذلك افرض أننا نعالج إحصاءات شهرية مثل إيرادات ترام الرمل الموضحة بالجدول رقم ٦٩ . إذا كان الاتجاه العام صعودياً كما هو الحال فى هذه الظاهرة فإن تأثيره يجعل قيمة شهر ديسمبر من أى سنة أعلى من قيمة شهر يناير من السنة نفسها بمقدار $\frac{١١}{١٢}$ من المعدل السنوى للاتجاه (أو بمقدار ١١ مرة المعدل الشهرى للاتجاه) وتكون قيمة نوفمبر أعلى من قيمة يناير من السنة نفسها بمقدار $\frac{١٠}{١٢}$ من المعدل السنوى للاتجاه (أو بمقدار ١٠ أمثال المعدل الشهرى للاتجاه) . وهكذا تكون قيمة كل شهر أعلى من قيمة الشهر السابق بمقدار المعدل الشهرى للاتجاه . ولذلك إذا لم نستبعد أثر الاتجاه العام فقد يبدو أن هناك تقلبات موسمية حتى إذا لم يكن هناك مؤثر موسمى حقيقى . فالنسبة الموسمية لشهر ديسمبر تكون دائماً مرتفعة بتأثير الاتجاه العام مع أن الظاهرة قد لا تكون خاضعة لأى مؤثر موسمى فى هذا الوقت من السنة ، أو قد يكون تأثير الاتجاه العام مختلطاً بتأثير الموسم مع أن هذا الأخير هو الذى نود دراسته على حدة . ويمكن استبعاد تأثير الاتجاه العام من النسب الموسمية بإحدى الطريقتين الآتيتين :

أولا : استبعاد الاتجاه قبل حساب المتوسطات

تتلخص هذه الطريقة فى استبعاد الاتجاه العام من القيم الفعلية (المشاهدة) للظاهرة قبل اتخاذ أية خطوة لحساب النسب الموسمية . فإذا كنا نبحث فى القيم الشهرية لظاهرة معينة خلال ثمانى سنوات ، أى لو كان عدد القيم الفعلية للظاهرة ٩٦ كما هو الحال فى الأرقام الخاصة بإيرادات ترام الرمل الموضحة بالجدول رقم ٦٩ ، نعين ٩٦ قيمة اتجاهية بإحدى الطرق التى أوردناها من قبل لإيجاد الاتجاه العام ثم نحسب كل قيمة فعلية كنسبة مئوية من القيمة الاتجاهية المناظرة لها ونكون بذلك قد استبعدنا أثر الاتجاه العام من القيم الفعلية للظاهرة .

بعد ذلك نعين المتوسطات الشهرية ومن ثم نستخرج النسب الموسمية كالمعتاد . ومن الواضح أن العمل الحسابى الذى تتطلبه هذه الطريقة يكون مرهقاً جداً خصوصاً إذا كنا نبحث فى قيم الظاهرة خلال عدد كبير من السنين .

ثانياً : استبعاد الاتجاه بعد حساب المتوسطات

تتلخص هذه الطريقة فى حساب المتوسطات الشهرية ثم قسمة هذه المتوسطات على القيم الاتجاهية الشهرية للظاهرة خلال السنة التى تقع فى وسط الفترة الزمنية التى نبحثها والتى نعتبرها بمثابة سنة نموذجية . ولإيجاد القيم الاتجاهية الشهرية خلال السنة النموذجية ليس من الضرورى أن نستخدم جميع القيم الشهرية للظاهرة (٩٦ فى المثال السابق) فى حساب معادلة خط الاتجاه بل نستخرج هذه المعادلة باستخدام متوسط القيمة الشهرية للظاهرة فى كل سنة (أى نستخدم ٨ قيم فقط فى المثال السابق) . وهذه المعادلة تعطينا معدل الاتجاه السنوى (أو النصف السنوى) ومنه نستطيع أن نحسب المعدل الشهرى للاتجاه بالقسمة على ١٢ (أو على ٦ إذا كان المعدل نصف سنوى) .

ولتوضيح هذه الطريقة نستبعد أثر الاتجاه العام من النسب الموسمية الشهرية لإيرادات ترام الرمل المبينة بالجدول رقم ٦٩ .

نبدأ أولا بإيجاد معادلة خط الاتجاه العام بطريقة المربعات الدنيا على أساس متوسط الإيرادات الشهرية فى كل سنة كما هو موضح بالجدول الآتى :

جدول رقم ٧٠

إيجاد معادلة خط الاتجاه لإيرادات ترام الرمل

السنة	متوسط الإيرادات الشهرية ص	س	س ص
١٩٤٠	٨١٨٢	— ٧	— ٥٧٢٧٤,
١٩٤١	١٤٠٢٤,٥	— ٥	— ٧٠١٢٢,٥
١٩٤٢	١٨٣٧٧,٤٢	— ٣	— ٥٥١٣٢,٢٦
١٩٤٣	٢٦٢٠٣,٧٥	— ١	— ٢٦٢٠٣,٧٥
١٩٤٤	٣٣٤٤٨,٩٢	+ ١	٣٣٤٤٨,٩٢
١٩٤٥	٣٥٥٤٤,٣٣	+ ٣	١٠٦٦٣٢,٩٩
١٩٤٦	٣١٤٢١	+ ٥	١٥٧١٠٥
١٩٤٧	٢٩٤٩٥,٦٧	+ ٧	٢٠٦٤٦٩,٦٩
المجموع	١٩٦٦٩٧,٥٩	٠٠	٢٩٤٩٢٤,٠٩

$$أ = \frac{\Sigma ص}{ن} = \frac{١٩٦٦٩٧,٥٩}{٨} = ٢٤٥٨٧,٢$$

$$ب = \frac{\Sigma س ص}{\Sigma س^{٢}} = \frac{٢٩٤٩٢٤,٠٩}{١٦٨} = ١٧٥٥,٥$$

∴ معادلة خط الاتجاه هى :

ص = ٢٤٥٨٧,٢ + ١٧٥٥,٥ س

حيث نقطة الأصل بين سنتى ١٩٤٣ ، ١٩٤٤ ووحدات س تمثل نصف سنة

وبقسمة (ب) على ٦ نحصل على المعدل الشهرى للاتجاه فتكون معادلة خط الاتجاه هى :

ص = ٢٤٥٨٧,٢ + ٢٩٢,٦ س (حيث وحدات س تمثل أشهراً)

والآن ننتقل إلى تعيين القيم الاتجاهية فى السنة الوسطى التى سنعتبرها بمثابة سنة نموذجية[1] .

يلاحظ أن قيمة (ا) فى المعادلة هى عبارة عن القيمة الاتجاهية للظاهرة عند نقطة الأصل أى القيمة الاتجاهية فى منتصف المسافة بين شهرى يونيه ويوليه من السنة الوسطى النموذجية . وللحصول على القيمة الاتجاهية فى شهر يونية نطرح من هذه القيمة نصف المعدل الشهرى للاتجاه فتكون القيمة الاتجاهية النموذجية لشهر يونيه = ٢٤٥٨٧,٢ − $\frac{١}{٢}$ × ٢٩٢,٦ = ٢٤٤٤٠,٩

وللحصول على القيمة الاتجاهية فى شهر يوليه نضيف إلى القيمة الاتجاهية عند نقطة الأصل نصف المعدل الشهرى للاتجاه فتكون القيمة الاتجاهية النموذجية لشهر يوليو = ٢٤٥٨٧,٢ + $\frac{١}{٢}$ × ٢٩٢,٦ = ٢٤٧٣٣,٥

ومتى عينا القيمتين الاتجاهيتين لشهرى يونيه ويوليه فمن السهل تعيين القيم الإتجاهية فى بقية الأشهر بحيث تكون القيمة ، فى كل شهر أكبر من القيمة فى الشهر السابق له بمقدار المعدل الشهرى للاتجاه . وبعد أن ننتهى من تعيين القيم الإتجاهية النموذجية نحسب المتوسطات الشهرية كنسب مئوية من القيم الاتجاهية النموذجية المناظرة لها كما هو موضح بالجدول الآتى :

---

(١) يلاحظ أن هذه السنة النموذجية هى عبارة عن الاثنى عشر شهراً التى تقع فى وسط السلسلة الزمنية تماماً أى التى تكون نقطة الأصل فى منتصفها .

جدول رقم ٧١

استبعاد الاتجاه العام من النسب الموسمية الشهرية لإيرادات ترام الرمل

الأشهر	المتوسطات الشهرية	القيم الاتجاهية النموذجية	المتوسطات الشهرية / القيم الاتجاهية × ١٠٠	النسب الموسمية بعد استبعاد الاتجاه
			٪	٪
يناير	١٩٩١٩	٢٢٩٧٧,٩	٨٦,٦٩	٨٦,٨٧
فبراير	١٧٨١٠,٩	٢٣٢٧٠,٥	٧٦,٥٤	٧٦,٧٠
مارس	٢٠٤٨٧,٩	٢٣٥٦٣,١	٨٦,٩٥	٨٧,١٣
أبريل	٢٢٢٩٦	٢٣٨٥٥,٧	٩٣,٤٦	٩٣,٦٦
مايو	٢٣٧٩٥,٦	٢٤١٤٨,٣	٩٨,٥٤	٩٨,٧٥
يونيه	٢٥٩٥٤,٦	٢٤٤٤٠,٩	١٠٦,١٩	١٠٦,٤١
يوليه	٣١٩١٣,٢٥	٢٤٧٣٣,٥	١٢٩,٠٣	١٢٩,٣١
أغسطس	٣٣٥٠٢	٢٥٠٢٦,١	١٣٣,٨٧	١٣٤,١٦
سبتمبر	٢٨٢٧١,١	٢٥٣١٨,٧	١١١,٦٦	١١١,٩١
أكتوبر	٢٤٧٢١,٢٥	٢٥٦١١,٣	٩٦,٥٢	٩٦,٧٢
نوفمبر	٢٢٧٥٥	٢٥٩٠٣,٩	٨٧,٨٤	٨٨,٠٣
ديسمبر	٢٣٦١٩,٧٥	٢٦١٩٦,٥	٩٠,١٦	٩٠,٣٥
المجموع			١١٩٧,٤٥	١٢٠٠,٠٠

يلاحظ فى الجدول السابق أنه بعد أن نحسب المتوسطات الشهرية كنسب مئوية من القيم الاتجاهية النموذجية المناظرة لها فإن مجموع هذه النسب المئوية لا يساوى ١٢٠٠ أى أن متوسطها لا يساوى ١٠٠ .

ولتعيين النسب الموسمية الموضحة بالعمود الأخير من هذا الجدول نضرب

كل نسبة من هذه النسب المئوية فى $\frac{١٢٠٠}{١١٩٧,٤٥}$. فالنسبة الموسمية لشهر

يناير مثلا = ٨٦,٦٩ × $\frac{١٢٠٠}{١١٩٧,٤٥}$ = ٨٦,٨٧ .

٢ — إيجاد التقلبات الموسمية بدلالة الوحدات المطلقة

على الرغم من أن إيجاد النسب الموسمية هو أفضل طريقة لمعرفة مدى تأثر الظاهرة بالموسم إلا أن مشقة العمل الحسابى اللازم لاستخراجها يجعلنا نكتفى أحياناً بالتعبير عن التقلبات الموسمية بدلالة الوحدات المطلقة . كذلك تكون هذه الطريقة الأخيرة لا بد منها إذا أردنا لسبب ما معرفة القيم الفعلية للتقلبات الموسمية . وأول خطوة فى عزل التقلبات الموسمية على هذا النحو هى إيجاد القيم الاتجاهية ثم حساب انحرافات القيم المشاهدة للظاهرة عن الاتجاه العام . ويلاحظ أن هذه الانحرافات تكون عبارة عن مزيج من التقلبات الموسمية والعرضية (على فرض أن الظاهرة ليست خاضعة لمؤثر دورى) .

ولاستبعاد التقلبات العرضية نأخذ المتوسط الشهرى (أو اليومى أو متوسط كل ثلاثة أشهر حسب فترة تكرار التقلبات الموسمية) للانحرافات عن الاتجاه العام . ذلك أن هذه التقلبات العرضية تحدث حيثما اتفق فتسبب ارتفاع قيمة الظاهرة أحياناً وانخفاضها أحياناً ، أى أن بعضها يكون سالباً والبعض الآخر موجباً . وعلى ذلك فإن أثرها يتلاشى تقريباً عند حساب المتوسطات الشهرية (أو اليومية أو الثلاث شهرية) للانحرافات عن الاتجاه العام . ويمكن بناء على ذلك اعتبار هذه المتوسطات الشهرية تقلبات موسمية بحتة . هذا طبعاً على فرض عدم خضوع الظاهرة لمؤثر دورى . وفيما يلى خطوات العمل اللازمة لتعيين التقلبات الموسمية بدلالة الوحدات المطلقة :

أولا : نحسب القيم الاتجاهية للظاهرة بإيجاد متوسط متحرك تتوقف فترته على طول الدورة الموسمية .

ثانياً : نحسب انحرافات القيم المشاهدة للظاهرة عن القيم الإتجاهية .

ثالثاً : نرتب هذه الانحرافات فى شكل جدولى يسمح باستخراج المتوسطات الشهرية (*) لها بسرعة وذلك بوضع انحرافات كل شهر فى عمود واحد .

رابعاً : نحسب المتوسطات الشهرية للانحرافات فتكون هى القيم الفعلية للتقلبات الموسمية .

ومن الواضح أنه إذا أردنا إيجاد قيمة التقلبات العرضية نطرح التقلبات الموسمية من الانحرافات عن الاتجاه العام التى ذكرنا أنها عبارة عن مزيج من التقلبات الموسمية والعرضية . ولتوضيح هذه الخطوات عملياً نحسب التقلبات الموسمية فى إنتاج الفحم فى إنجلترا كما هو مبين بالجدولين رقم ٧٢ ، رقم ٧٣

(*) أو متوسطات أية فترة زمنية أخرى حسب الحالة .

جدول رقم ٧٢

إنتاج الفحم فى بريطانيا كل ثلاثة أشهر من سنة ١٩٢٧ إلى سنة ١٩٣١(*)

السنة		إنتاج الفحم بملايين الأطنان	مجموع متحرك فترته ٤ وحدات زمنية	مجموع متحرك فترته وحدتان للمجموع السابق	متوسط متحرك معدل فترته ٤ وحدات زمنية	الانحرافات عن الاتجاه العام
١٩٢٧	الربع الأول	٦٨,٣				
	« الثانى	٦٢,٦	٢٥٥,٣			
	« الثالث	٦١,١	٢٥٢,٤	٥٠٧,٧	٦٣,٥	— ٢,٤
	« الرابع	٦٣,٣	٢٤٧,٧	٥٠٠,١	٦٢,٥	+ ٠,٨
١٩٢٨	« الأول	٦٥,٤	٢٤٣,٠	٤٩٠,٧	٦١,٣	+ ٤,١
	« الثانى	٥٧,٩	٢٤١,٢	٤٨٤,٢	٦٠,٥	— ٢,٦
	« الثالث	٥٦,٤	٢٤٣,٩	٤٨٥,١	٦٠,٦	— ٤,٢
	« الرابع	٦١,٥	٢٤٨,٧	٤٩٢,٦	٦١,٦	— ٠,١
١٩٢٩	« الأول	٦٨,١	٢٥٥,١	٥٠٣,٨	٦٣,٠	+ ٥,١
	« الثانى	٦٢,٧	٢٦٠,٦	٥١٥,٧	٦٤,٥	— ١,٨
	« الثالث	٦٢,٨	٢٦٢,٦	٥٢٣,٢	٦٥,٤	— ٢,٦
	« الرابع	٦٧,٠	٢٥٩,٠	٥٢١,٦	٦٥,٢	+ ١,٨
١٩٣٠	« الأول	٧٠,١	٢٥٢,٥	٥١١,٥	٦٣,٩	+ ٦,٢
	« الثانى	٥٩,١	٢٤٧,١	٤٩٩,٦	٦٢,٤	— ٣,٣
	« الثالث	٥٦,٣	٢٣٦,٥	٤٨٣,٦	٦٠,٤	— ٤,١
	« الرابع	٦١,٦	٢٣٢,٢	٤٦٨,٧	٥٨,٦	+ ٣,٠
١٩٣١	« الأول	٥٩,٥	٢٢٧,٠	٤٥٩,٢	٥٧,٤	+ ٢,١
	« الثانى	٥٤,٨	٢٢٣,٤	٤٥٠,٤	٥٦,٣	— ١,٥
	« الثالث	٥١,١				
	« الرابع	٥٨,٠				

(*) هـذه الأرقام مستمدة من كتاب : Elementary Statical Methods by : E.C. Rhodes

يلاحظ بالنظر إلى الجدول السابق أننا اخترنا متوسطاً متحركاً فترته أربع وحدات زمنية (حيث كل وحدة عبارة عن ثلاثة أشهر) لتمثيل الاتجاه العام للظاهرة. وبطرح القيم الاتجاهية من القيم المشاهدة للظاهرة نحصل على الانحرافات عن الاتجاه العام المبينة قيمها بالعمود الأخير من الجدول. وقد ذكرنا من قبل أن هذه الانحرافات هى عبارة عن مزيج من التقلبات الموسمية والعرضية. ولكى نعزل التقلبات الموسمية على حدة نأخذ متوسطات هذه الانحرافات فى كل ربع من أرباع السنة. ولتسهيل حساب هذه المتوسطات نرتب الأنحرافات عن الاتجاه العام فى شكل جدولى كما يأتى :

جدول رقم ٧٣

إيجاد التقلبات الموسمية لإنتاج الفحم فى بريطانيا

السنة	الربع الأول	الربع الثانى	الربع الثالث	الربع الرابع
١٩٢٧			—٢,٤	+٠,٨
١٩٢٨	+٤,١	—٢,٦	—٤,٢	—٠,١
١٩٢٩	+٥,١	—١,٨	—٢,٦	+١,٨
١٩٣٠	+٦,٢	—٣,٣	—٤,١	+٣,٠
١٩٣١	+٢,١	—١,٥		
المجموع	+١٧,٥	—٩,٢	—١٣,٣	+٥,٥
المتوسط	+٤,٤	—٢,٣	—٣,٣	+١,٤

وهذه المتوسطات المبينة بالسطر الأخير من الجدول تمثل التقلبات الموسمية. ويلاحظ أنه كلما كان عدد السنوات التى تمتد خلالها السلسلة الزمنية كبيراً كلما كانت النتائج التى نحصل عليها أقرب إلى الدقة. فقد افترضنا أن حساب

متوسطات الانحرافات يستبعد أثر التقلبات العرضية . وهذا الفرض لا يكون صحيحاً إلا فى المدة الطويلة حيث تميل التقلبات العرضية التى تسبب زيادة قيمة الظاهرة إلى التعادل مع تلك التى تحدث نقصاً فيها .

الآن وقد حسبنا كلا من الاتجاه العام والتقلبات الموسمية يمكننا أن نحلل القيم المشاهدة للظاهرة إلى العناصر التى تكونها . وتتضح نتيجة هذا التحليل فى السنتين الأوليين من الجدول الآتى :

جدول رقم ٧٤

تحليل السلسلة الزمنية إلى العناصر المكونة لها

السنة		الاتجاه العام	التقلبات الموسمية	التقلبات العرضية	القيم المشاهدة للظاهرة
١٩٢٧	الربع الثالث	٦٣,٥	—٣,٣	+٠,٩	٦١,١
	» الرابع	٦٢,٥	+١,٤	—٠,٦	٦٣,٣
١٩٢٨	» الأول	٦١,٣	+٤,٤	—٠,٣	٦٥,٤
	» الثانى	٦٠,٥	—٢,٣	—٠,٣	٥٧,٩
	» الثالث	٦٠,٦	—٣,٣	—٠,٩	٥٦,٤
	» الرابع	٦١,٦	+١,٤	—١,٥	٦١,٥

يلاحظ بالنظر إلى الجدول السابق أنه لكى نحصل على قيم التقلبات العرضية طرحنا مجموع القيمة الاتجاهية والتقلبات الموسمية فى كل وحدة زمنية من القيمة المشاهدة المناظرة لها . فمثلا قيمة التقلبات العرضية فى الربع الثالث من سنة ١٩٢٧ تساوى :

$$٦١,١ - [٦٣,٥ + (-٣,٣)] = +٠,٩$$

وقيمتها فى الربع الأول من سنة ١٩٢٨ تساوى :

$$٦٥,٤ - [٦١,٣ + ٤,٤] = -٠,٣$$

استبعاد تأثير الموسم من القيم المشاهدة للظاهرة

الغرض الرئيسى من حساب النسب الموسمية هو دراسة تأثير الموسم على قيم الظاهرة وهذه الدراسة ذات أهمية كبرى بالنسبة للمنتجين بوجه عام وأولئك الذين يتغير الطلب على منتجاتهم كثيراً من وقت لآخر بنوع خاص . فحساب النسب الموسمية يجعل من السهل على هؤلاء المنتجين إعداد العدة لمواجهة الزيادة (أو النقص) فى الطلب على منتجاتهم . فمثلا شركات السكك الحديدية أو شركات الترام تجد فائدة كبيرة فى دراسة تأثير الموسم على حركة النقل . إذ أن ذلك يمكنها من إعداد قطارات إضافية فى أيام السنة (أو ساعات اليوم) التى تشتد فيها حركة النقل وتقليل عدد القطارات فى الأوقات التى تقل فيها هذه الحركة . وبالإضافة إلى ذلك فإن المنتج يهمه أن يقف على الحالة الحقيقية للطلب على منتجاته . فقد تزداد إيراداته كثيراً فى وقت معين من أوقات السنة ومع ذلك تكون هذه الزيادة مؤقتة تزول بزوال تأثير الموسم على الإيرادات . ولكى يقف المنتج على الحالة الحقيقية لإيراداته التى يستطيع أن يتخذها أساساً لقراراته وتصرفاته المتعلقة بمستقبل المشروع الذى يشرف عليه فلا بد أن يستبعد تأثير الموسم من هذه الإيرادات . ويمكن استبعاد تأثير الموسم بقسمة القيم المشاهدة للظاهرة على النسب الموسمية . فلو فرض أن كانت إيرادات محل تجارى فى شهر معين هى ٥٠٠٠ جنيه وكانت النسبة الموسمية للإيرادات فى هذا الشهر هى ١٢٥ ٪ . فإن الإيرادات بعد استبعاد تأثير الموسم :

$$= ٥٠٠٠ \times \frac{١٠٠}{١٢٥} = ٤٠٠٠ \text{ جنيه}$$

وإذا كانت النسبة الموسمية فى هذا الشهر = ٥٠ ٪ فإن الإيرادات بعد استبعاد تأثير الموسم :

$$= ٥٠٠٠ \times \frac{١٠٠}{٥٠} = ١٠٠٠٠ \text{ جنيه}$$

أما إذا كانت قيم التقلبات الموسمية مقدرة بدلالة الوحدات المطلقة، فيمكن استبعادها بطرحها مباشرة من القيم المشاهدة للظاهرة . وبديهى أننا لا نقصد مما تقدم أن المنتج لا يشعر بتأثير الموسم على حالة الطلب على منتجاته إلا عن طريق حساب النسب الموسمية . فكل منتج يعلم جيداً متى يزداد الطلب على منتجاته ومتى يقل هذا الطلب . غير أن معلوماته عن تلك الزيادة أو ذلك النقص تكون تقريبية كما أنه إذا اقترنت الزيادة الموسمية بزيادة دائمة فى الطلب فلا يمكن التمييز بينهما إلا عن طريق حساب النسب الموسمية .

التغيرات الدورية

الطريقة المعتادة لقياس تأثير التقلبات الدورية هى استبعاد الاتجاه العام والتأثيرات الموسمية على التوالى من قيم الظواهر وذلك بقسمة القيمة المشاهدة للظاهرة على القيمة الاتجاهية ثم على النسبة الموسمية المناظرة لها . فقد ذكرنا من قبل أنه إذا كانت ح ، م ، و ، ع تمثل القيم الاتجاهية والتقلبات الموسمية والدورية والعرضية على التوالى فإن القيمة المشاهدة للظاهرة = ح × م × و × ع وبالقسمة على ح نحصل على م × و × ع .

وبالقسمة مرة أخرى على م نحصل على و × ع .

أما التغيرات العرضية فهى بطبيعتها لا يمكن حصرها أو استبعادها . غير أنه إذا فرضنا أنها تحدث فى اتجاهات متضادة كما هو الغالب فيمكن استبعادها بأخذ متوسطات متحركة للناتج الأخير بعد استبعاد تأثير الاتجاه العام وتأثير الموسم . والجدول الآتى يوضح طريقة حساب النسب الدورية[1] عملياً :

---

(١) تسمى هذه النسب بالإنجليزية Cyclical Relatives

جدول رقم ٧٥

إيجاد النسب الدورية باستبعاد الاتجاه العام والتقلبات الموسمية

الزمن	القيم المشاهدة ح × م × و × ع	القيمة الاتجاهية ح	$\frac{\text{ح} \times \text{م} \times \text{و} \times \text{ع}}{\text{ح}} \times ١٠٠$	النسب الموسمية م	النسب الدورية $\frac{\text{م} \times \text{و} \times \text{ع}}{\text{ح}}$
			٪	٪	٪
١٩٢٠ يناير	١٦٩٦	٣٩٤٩	٤٢,٩	٧٥,١	٥٧,١
فبراير	٢١٢٨	٣٩٥٩	٥٣,٨	٩٦,٢	٥٥,٩
مارس	٢٥١١	٣٩٦٩	٦٣,٣	١٠٧,٥	٥٨,٩
أبريل	٢٨٦٠	٣٩٧٩	٧١,٩	١٢٣,٦	٥٨,٢
مايو	٢٨٥٢	٣٩٨٩	٧١,٥	١٢٢	٥٨,٦
يونيه	٢٦٣٧	٣٩٩٩	٦٥,٩	١١١,١	٥٩,٣
يوليه	١٩٦٧	٤٠٠٩	٤٩,١	٨٣,٣	٥٨,٩
أغسطس	١٦٩٥	٤٠١٨	٤٢,٢	٧١,٧	٥٨,٩
سبتمبر	٢٠٨٤	٤٠٢٨	٥١,٧	٨٧,٧	٥٩
أكتوبر	٢٦٣٧	٤٠٣٨	٦٥,٣	١٠٧,٩	٦٠,٥
نوفمبر	٢٧٣٦	٤٠٤٨	٦٧,٦	١١٠,٥	٦١,٢
ديسمبر	٢٧٣١	٤٠٥٨	٦٧,٣	١٠٣,٣	٦٥,٢
١٩٢١	....	....	....	....	....
....	....	....	....	....	....
....	....	....	....	....	....

الارتباط بين السلسلات الزمنية

رأينا أن السلسلات الزمنية تكون خاضعة فى تغيراتها لأثر الاتجاه العام والتقلبات الموسمية والدورية والعرضية . فإذا أردنا حساب الارتباط بين سلسلتين زمنيتين فمن المهم جداً أن نحدد أولا طبيعة العلاقة المطلوب إيجادها بين الظاهرتين. فقد يكون المطلوب هو مجرد معرفة ما إذا كان الاتجاه العام للظاهرة الأولى يتمشى مع الاتجاه العام للظاهرة الثانية صعوداً أو هبوطاً . وفى هذه الحالة يكفى أن نحسب القيم الاتجاهية لكل من الظاهرتين بإحدى الطرق التى سبق شرحها ثم نرصد خطى الاتجاه على ورق الرسم البيانى مستخدمين نفس المحور الزمنى فى الحالتين إذ سيتضح من الرسم مباشرة ما إذا كان هناك اتفاق بين اتجاه الظاهرتين أو ما إذا كانتا تتغيران فى اتجاهين مختلفين . وليس من الضرورى فى مثل هذه الحالة أن نوجد معامل الارتباط بين القيم الاتجاهية للظاهرتين لأن العلاقة بين خطى الاتجاه تظهر من الرسم بطريقة لا تحتاج إلى تأييد . بل الواقع أن حساب معامل الارتباط بين القيم الاتجاهية ليس له ما يبرره إطلاقاً نظراً لأن مجرد اتفاق القيم الاتجاهية للظاهرتين فى الصعود أو فى الهبوط سيترتب عليه وجود ارتباط إحصائى شديد جداً بينهما مهما اختلف معدل الصعود أو الهبوط فى كل منهما . فلو فرض مثلا أن حسبنا معامل سبيرمان للارتباط بين ظاهرتين متحدتين فى الاتجاه الصعودى ولكن مختلفتين اختلافاً كبيراً فى معدل الصعود فإن النتيجة تكون الواحد الصحيح مهما كانت طبيعة العلاقة بين الظاهرتين . ونحن نعلم أن النتائج التى نحصل عليها باستخدام معامل الارتباط العادى لا تختلف كثيراً عن تلك التى نحصل عليها باستخدام معامل سبيرمان .

على أنه فى أغلب الحالات التى نبحث فيها فى العلاقة بين السلسلات الزمنية يكون من الأجدى بحث هذه العلاقة بعد استبعاد تأثير الاتجاه العام من الظاهرتين . ذلك أن اتفاق الاتجاه العام للظاهرتين فى الصعود أو الهبوط

يوجد دائماً نوعاً من الارتباط الظاهرى أو الارتباط الوهمى كما بينا ، وهذا لا يمكن الاعتماد عليه فى وصف العلاقة بين الظواهر . فإذا كانت السلسلتان الزمنيتان المطلوب بحث العلاقة بينهما تمثلان أرقاماً سنوية لا يظهر فيها الأثر الموسمى (الذى يظهر فى فترات زمنية أقل من السنة) ثم استبعدنا منها أثر الاتجاه العام وحسبنا بعد ذلك معامل الارتباط بينهما ووجدناه شديداً فلا شك أن دلالته تكون كبيرة فى إثبات وجود العلاقة بين الظاهرتين . ذلك أنه بعد استبعاد تأثير الاتجاه العام من السلسلتين الزمنيتين لا يبقى إلا المؤثرات الدورية وما اختلط بها من مؤثرات عرضية . وليس من شك فى أن اتحاد المؤثرات الدورية أقل احتمالا من اتحاد الاتجاه العام للظاهرتين بنفس الكيفية . والذى نستخلصه مما تقدم أنه عند حساب معامل الارتباط بين سلسلتين زمنيتين لغرض الوقوف على طبيعة العلاقة بينهما ، يجب أولا وقبل كل شىء أن نستبعد أثر الاتجاه العام من كل منهما . والمثال الآتى يوضح طريقة حساب معامل الارتباط بين السلسلات الزمنية .

مثال : احسب معامل الارتباط بين الرقم القياسى للإنتاج الصناعى والرقم القياسى لأسعار الأسهم من واقع الأرقام المبينة بالجدول الآتى :

جدول رقم ٧٦

السنة	الرقم القياسى للإنتاج الصناعى س	الرقم القياسى لأسعار الأسهم ص
١٩٤٤	١٠١	١٠٠,٦
١٩٤٥	٩٥	٩٨,١
١٩٤٦	١٠٤	١٠٣,٥
١٩٤٧	١٠٨	١٠٠,٠
١٩٤٨	١٠٦	٩٥,٤
١٩٤٩	١١١	٩٦,٧
١٩٥٠	١١٩	٩٥,٣
	٧٤٤	٦٨٩,٦

خطوات العمل :

أولا : نحسب القيم الاتجاهية لكل من الظاهرتين بطريقة المربعات الدنيا .

ثانياً : نحسب انحراف كل قيمة من قيم الظاهرتين عن القيمة الاتجاهية المناظرة لها وذلك لاستبعاد أثر الاتجاه العام* .

ثالثاً : نحسب الارتباط بين انحرافات الظاهرتين عن اتجاههما العام بالطريقة المعتادة .

والجدول الآتى يوضح خطوات العمل بعد حساب القيم الاتجاهية للظاهرتين بطريقة المربعات الدنيا .

جدول رقم ٧٧

---

* يفضل بعض الإحصائيين استبعاد الاتجاه العام بقسمة قيمة الظاهرة على القيمة الاتجاهية بدلا من طرح الاتجاه العام من قيمة الظاهرة ، غير أن هذه الطريقة تتطلب عملا حسابياً مطولا وتؤدى إلى نفس النتائج تقريباً .

جدول رقم ٧٧

حساب معامل الارتباط بين سلسلتين زمنيتين

السنة	القيم الاتجاهية للإنتاج الصناعى ح	القيم الاتجاهية لأسعار الأسهم حَ	انحرافات قيم الظاهرتين عن اتجاههما العام		(س — ح)٢	(ص — حَ)٢	(س — ح) × (ص — حَ)
			س — ح	ص — حَ			
١٩٤٤	٩٦,٨٦	١٠١,٣٩	+ ٤,١٤	— ,٧٩	١٧,١٣٥٦	,٦٢٤١	— ٣,٢٧٠٦
١٩٤٥	١٠٠,٠٠	١٠٠,٤٣	— ٥,٠٠	— ٢,٣٣	٢٥,٠٠٠٠	٥,٤٢٨٩	+ ١١,٦٥٠٠
١٩٤٦	١٠٣,١٤	٩٩,٤٧	+ ,٨٦	+ ٤,٠٣	,٧٣٩٦	١٦,٢٤٠٩	+ ٣,٤٦٥٨
١٩٤٧	١٠٦,٢٩	٩٨,٥١	+ ١,٧١	+ ١,٤٩	٢,٩٢٤١	٢,٢٢٠١	+ ٢,٥٤٧٩
١٩٤٨	١٠٩,٤٣	٩٧,٥٦	— ٣,٤٣	— ٢,١٦	١١,٧٦٤٩	٤,٦٦٥٦	+ ٧,٤٠٨٨
١٩٤٩	١١٢,٥٧	٩٦,٦٠	— ١,٥٧	+ ,١٠	٢,٤٦٤٩	,٠١٠٠	— ,١٥٧٠
١٩٥٠	١١٥,٧١	٩٥,٦٤	+ ٣,٢٩	— ,٣٤	١٠,٨٢٤١	,١١٥٦	— ١,١١٨٦
المجموع	٧٤٤,٠٠	٦٨٩,٦٠	٠٠	٠٠	٧٠,٨٥٣٢	٢٩,٣٠٥٢	٢٠,٥٦٣٢

الانحراف المعيارى لانحرافات قيم س عن الاتجاه العام* $= \sqrt{\frac{٧٠,٨٥٣٢}{٧}} = ٣,١٨١٦$

الانحراف المعيارى لانحرافات قيم ص عن الاتجاه العام $= \sqrt{\frac{٢٩,٣٠٥٢}{٧}} = ٢,٠٤٦١$

∴ معامل الارتباط $= \frac{٢٠,٥٢٦٣}{٧ \times ٣,١٨١٦ \times ٢,٠٤٦١} = ٠,٤٥$

* لاحظ أن الوسط الحسابى لكل من الانحرافات السينية والصادية عن الاتجاه العام يساوى صفرا .

الباب الرابع عشر

الأرقام القياسية

الرقم القياسى هو أداة لقياس التغير النسبى فى قيم الظواهر من وقت إلى آخر أو من مكان إلى آخر . وأبسط شكل يتخذه هذا الرقم هو عبارة عن نسبة مئوية تعبر عن قيمة الظاهرة فى سنة معينة (أو مكان معين) بدلالة قيمتها فى سنة أخرى تعتبر أساساً للقياس أو قيمتها فى مكان آخر يعتبر أساساً للقياس كذلك . فإذا كان سعر القطن مثلا فى سنة ١٩٣٧ هو ١٥ جنيهاً للقنطار وكان سعره فى سنة ١٩١٤ خمسة جنيهات للقنطار فإن الرقم القياسى لسعر القطن فى سنة ١٩٣٧ هو $\frac{١٥}{٥} \times ١٠٠ = ٣٠٠$ ٪ على فرض أن سنة ١٩١٤ اعتبرت أساساً للقياس. غير أن تركيب الأرقام القياسية لا يكون عادة بمثل تلك السهولة. فقد نحتاج أحياناً إلى مقارنة ظاهرتين تكون قيمة كل منهما عبارة عن ملخص لعدة قيم أخرى تشترك جميعها فى تكوينها . ومن الأمثلة الواضحة على هذه الظواهر المركبة التى نحتاج إلى مقارنتها ، نفقات المعيشة أو المستوى العام للأسعار فى سنتين مختلفتين أو فى مكانين مختلفين ، إذ من الواضح أن مقارنة مستوى الأسعار بين مكان وآخر أو بين وقت وآخر تتطلب حساب متوسط عام يمثل جميع الأسعار فى كل من المكانين أو الزمنين المطلوب عمل المقارنة بينهما . ذلك أن أسعار جميع السلع لا تتغير بنفس النسبة أو فى نفس الاتجاه . فقد يرتفع بعضها بنسبة كبيرة والبعض الآخر بنسبة صغيرة بينما يظل بعضها ثابتاً وينخفض البعض الآخر . فإذا ذكرنا أن مستوى الأسعار ارتفع فى سنة ١٩٣٧ بمقدار ٢٠ ٪ عن مستوى الأسعار فى سنة ١٩١٤ فإن هذا لا يعنى طبعاً أن أسعار جميع السلع فى سنة ١٩٣٧ تزيد بمقدار ٢٠ ٪ عن الأسعار المناظرة لها فى

سنة ١٩١٤ وإنما المقصود من ذلك أن المتوسط العام للأسعار فى سنة ١٩٣٧ يزيد بمقدار ٢٠ ٪ عن المتوسط العام للأسعار فى سنة ١٩١٤ . أى أنه لو فرض أن كان متوسط الأسعار فى سنة ١٩١٤ يساوى ١٠٠ فإن متوسط الأسعار فى سنة ١٩٣٧ يكون ١٢٠ . وهذا الرقم الأخير هو عبارة عن الرقم القياسى لأسعار سنة ١٩٣٧ منسوبة إلى أسعار ١٩١٤ التى نعتبرها كأساس يقاس التغير بالنسبة إليه .

وبالإضافة إلى الرقم القياسى للأسعار يمكن عمل أرقام قياسية لأية ظاهرة أخرى يراد مقارنة القيم التى تتخذها بين وقت وآخر أو مكان وآخر . فيمكن مثلا عمل رقم قياسى للإنتاج ورقم قياسى للصادرات ورقم قياسى للأجور غير أن أننا سنكتفى هنا بشرح طرق تركيب الأرقام القياسية للأسعار فى شىء من التفصيل نظراً لأنها أهم الأرقام القياسية وأكثرها استعمالا على أن تعتبر هذه الطرق بمثابة نماذج يمكن أن ننسج على منوالها فى تركيب أى رقم قياسى آخر . ومن الواضح أنه ليس ثمة صعوبة فى عمل أرقام قياسية للظواهر التى يمكن تمثيلها بقيمة واحدة فى كل من السنتين المطلوب عمل المقارنة بينهما . فالرقم القياسى فى هذه الحالة لا يعدو كونه نسبة مئوية بين قيمة الظاهرة فى السنة المقارنة وقيمتها فى السنة التى اختيرت كأساس كما سبق أن بينا .

ومن الأمثلة على هذه الأرقام القياسية البسيطة الرقم القياسى لسعر سلعة واحدة فى سوق معينة .

وإنما تبدأ الصعوبة إذا أردنا مقارنة أسعار جميع السلع فى سنة معينة بأسعار هذه السلع فى سنة الأساس إذ لا بد فى هذه الحالة أن نلخص التغيرات فى أسعار جميع السلع فى شكل رقم واحد يعبر عن ارتفاع أو انخفاض المستوى العام للأسعار فى السنة المقارنة بالنسبة للمستوى العام للأسعار فى سنة الأساس . وهناك عدة طرق لتلخيص البيانات الخاصة بالأسعار لتسهيل مقارنتها بين وقت

وآخر وتتوقف طريقة تركيب الرقم القياسى على اختيار إحداها . غير أنه مهما اختلفت طريقة تركيب الأرقام القياسية للأسعار فجميعها تشترك فى الخطوات الآتيـــة :

١ — اختيار السلع التى تدخل فى تركيب الرقم القياسى واختيار المصادر التى تستقى منها أسعار هذه السلع .

٢ — اختيار سنة الأساس .

٣ — إعطاء كل سلعة « وزناً » يتناسب مع أهميتها

٤ — تلخيص البيانات فى شكل رقم واحد يمثل متوسط التغير النسبى فى الأسعار .

١ — اختيار السلع والمصادر التى تستقى منها الأسعار

إذا تأملنا فى قائمة السلع التى يعرضها محل تجارى متوسط الحجم لاتضح لنا فى الحال أنه يكاد يكون من المستحيل عملياً حساب الرقم القياسى للأسعار على أساس متوسط التغير النسبى فى أسعار جميع السلع . ولو فرض جدلا أننا قبلنا تحمل مشقة العمل الحسابى الذى يتطلبه إدخال كل سلعة فى تركيب الرقم القياسى فإننا نواجه مشكلة الأنواع المتعددة للسلعة الواحدة التى يكون من الصعب جداً حصرها . على أنه لو أمعنا النظر فى هذه المسألة لتبين لنا أن هناك عدداً محدوداً من السلع الهامة التى تستنفد الجزء الأعظم مما ينفقه الفرد العادى على الاستهلاك . وهذه هى السلع التى نختارها لغرض حساب متوسط التغير النسبى فى أسعارها ونعتبرها بمثابة « عينة » تمثل جميع السلع . أما عن المصادر التى تستقى منها أسعار هذه السلع الهامة فهذه مسألة لا تنطوى على صعوبة كبيرة ، فالمفروض أن أسعار السلعة الواحدة تتساوى فى السوق الواحدة بفعل المنافسة الحرة . أما فروق الأسعار التى تشاهد بين محل تجارى وآخر نظراً لاختلاف الموقع أو نوع الخدمة وما إلى ذلك فإنها تكون عادة فروقاً صغيرة يمكن

إزالة أثرها بأخذ متوسط لعدد من أسعار كل سلعة نحصل عليه من المحلات التجارية النموذجية أو من نشرات الأسعار الرسمية إذا وجدت .

٢ — اختيار سنة الأساس

يراعى فى اختيار سنة الأساس أن تكون سنة عادية خالية من الشذوذ . فلا يصح مثلا اختيار إحدى سنوات الحرب كقاعدة للمقارنة كما لا يجوز اختيار سنة تقع فى أوج الرخاء أو فى أعماق الكساد التجارى نظراً لارتفاع الأسعار أو انخفاضها بدرجة غير عادية تفسد المقارنة فى مثل هذه الظروف .

وإلى عهد قريب كانت سنة ١٩١٣ — ١٩١٤ تستخدم كأساس فى معظم الأرقام القياسية للأسعار . وعلى الرغم من أنها كانت تقع فى فترة من الرخاء التجارى إلا أن وقوعها قبل الحرب العالمية الأولى مباشرة واعتبارها نهاية مرحلة من مراحل التاريخ جعلها تصلح كأساس للمقارنة . وأخيراً استخدمت سنة ١٩٣٥ كأساس لأرقام قياسية كثيرة باعتبارها سنة عادية استقرت فيها الحياة الاقتصادية بعد الكساد الشديد الذى حل بالعالم فى سنة ١٩٣٠ وامتد بضع سنوات . والسبب فى تغيير سنة الأساس من وقت لآخر هو أنه لكى تكون المقارنة صحيحة فلا بد أن تكون الظروف المحيطة بإنتاج واستهلاك السلع فى السنة المقارنة لا تختلف كثيراً عن هذه الظروف فى سنة الأساس . فإذا ابتعدت سنة الأساس كثيراً عن السنة المقارنة لوجدنا أن هناك بعض السلع الهامة التى لم تكن موجودة فى سنة الأساس أو على الأقل لم تكن موجودة بشكلها الحالى فإذا أردنا مثلا عمل رقم قياسى لتكاليف النقل من القاهرة إلى الإسكندرية فلا يصح طبعاً اتخاذ سنة ١٨٥٠ كأساس إذ أن وسائل النقل فى تلك السنة تختلف اختلافاً كلياً عن وسائل النقل فى الوقت الحاضر .

وللتغلب على صعوبة تغير الظروف الحالية عن الظروف المناظرة لها فى سنة

الأساس تستخدم أحياناً طريقة الأساس المتحرك (Chain Base) . وفى هذه الحالة تحسب أسعار كل سنة كنسب مئوية من أسعار السنة السابقة لها مباشرة . ويمكن بعد ذلك رد هذه النسب المئوية إلى أساس ثابت كما يأتى :

إذا كان سعر سنة ١٩٣٦ بالنسبة إلى سعر سنة ١٩٣٥ = ١٢٥ ٪

وسعر سنة ١٩٣٧ » » » » ١٩٣٦ = ١١٠ ٪

وسعر سنة ١٩٣٨ » » » » ١٩٣٧ = ١٠٥ ٪

فإن سعر سنة ١٩٣٧ بالنسبة إلى سعر سنة ١٩٣٥

$$= ١١١ \times \frac{١٢٥}{١٠٠} = ١٣٧,٥ \text{ ٪}$$

وسعر سنة ١٩٣٨ بالنسبة إلى سعر سنة ١٩٣٥

$$= ١٠٥ \times \frac{١١٠}{١٠٠} \times \frac{١٢٥}{١٠٠} = ١٤٤,٤ \text{ ٪}$$

وبناء على ذلك يكون لدينا الأرقام القياسية الآتية على فرض أن سنة ١٩٣٥ هى سنة الأساس .

١٩٣٥	١٩٣٦	١٩٣٧	١٩٣٨
١٠٠	١٢٥	١٣٧,٥	١٤٤,٤

٣ ــ اختيار الأوزان المناسبة لكل سلعة

سبق أن بينا فى الباب الخاص بالمتوسطات أنه عند حساب متوسط عدد من القيم فلا بد أن تعطى كل قيمة وزناً يتناسب مع أهميتها[1] . وحيث أن

---

(١) راجع صفحة ١٢٩ من هذا الكتاب .

تركيب الرقم القياسى للأسعار يتطلب حساب متوسط عام لأسعار عدد كبير من السلع المختلفة الأهمية فلا بد من ترجيح هذا المتوسط بالأوزان كى يكون الرقم القياسى معبراً بدقة عن حقيقة التغير فى المستوى العام للأسعار .

والقاعدة العامة فى اختيار الأوزان هى أن تعطى كل سلعة أهمية تتناسب مع قيمة ما يستهلك منها . فإذا كنا بصدد تركيب رقم قياسى لأسعار التجزئة ووجدنا أن الشخص المتوسط ينفق على اللحم ضعف ما ينفق على الخبز فلا بد أن يكون الوزن الذى نعطيه للحم ضعف الوزن الذى نعطيه للخبز . وهناك عدة طرق لاختيار الأوزان تختلف باختلاف طريقة تركيب الرقم وسيأتى الكلام عن كل منها عند شرح هذه الطرق .

٤ — تلخيص البيانات

هناك طريقتان رئيسيتان لتلخيص البيانات فى شكل رقم واحد يمثل متوسط التغير النسبى . ويتفرع من هاتين الطريقتين عدة طرق أخرى تختلف باختلاف الطريقة المتبعة لاختيار الأوزان . وهاتان الطريقتان هما :

أولا : الأرقام القياسية التجميعية Aggregative Index Numbers

ثانياً : الأرقام القياسيةالنسبية Relative Index Numbers

أولا : الأرقام التجميعية

١ — الرقم التجميعى البسيط للأسعار

فى هذه الطريقة نأخذ حاصل جمع الأسعار فى السنة المقارنة كنسبة مئوية من حاصل جمع الأسعار فى سنة الأساس . فلو فرض أن كانت أسعار السلع فى السنة المقارنة :

$ع_{ن_1}$ ، $ع_{ن_2}$ ، $ع_{ن_3}$ ، . . . . . $ع_{ن_ر}$

وكانت أسعارها فى سنة الأساس

$ع_{س_1}$ ، $ع_{س_2}$ ، $ع_{س_3}$ ، . . . . . $ع_{س_ر}$

$$\text{فإن الرقم التجميعى البسيط} = \frac{ع_{ن_1} + ع_{ن_2} + ع_{ن_3} + \cdots + ع_{ن_ر}}{ع_{س_1} + ع_{س_2} + ع_{س_3} + \cdots + ع_{س_ر}}$$

$$= \frac{\Sigma\, ع_{ن}}{\Sigma\, ع_{س}} \times ١٠٠$$

وعلى الرغم من أن هذه هى أسهل الطرق فى تركيب الأرقام القياسية إلا أنها لا تستند إلى أساس صحيح . ذلك أن أسعار السلع التى تدخل فى تركيب الرقم القياسى تكون محسوبة على أساس وحدات مختلفة بعضها كبير وبعضها صغير ولذلك يكون الرقم القياسى متأثراً تأثراً عظيما بأسعار السلع ذات الوحدات الكبيرة . ولإيضاح ذلك افرض أن أسعار السلع فى السنة المقارنة وسنة الأساس كانت كالآتى :

	القطن بالقنطار	السكر بالأقة	اللحم بالأقة
السنة المقارنة	$\frac{ص}{١٠٠٠}$	$\frac{ص}{١٠}$	$\frac{ص}{٥٠}$
سنة الأساس	٢٠٠	١	٥

$$\text{الرقم القياسى التجميعى البسيط} = \frac{١٠٠٠+١٠+٥٠}{٢٠٠+١+٥} \times ١٠٠ = ٥١٤ \text{ ٪}$$

وهذا الرقم يكاد يعادل الرقم القياسى للقطن بمفرده وهو ٥٠٠ ٪ فى حين أن الرقم القياسى لكل من السلعتين الأخريين هو ١٠٠٠ ٪ . وهذه النتيجة تظهر أن هذا الرقم القياسى يكون مضللا إذا كان هناك اختلاف كبير بين وحدات السلع التى تدخل فى تركيبه .

٢ – الرقم التجميعى المرجح بالأوزان

لكى نراعى أهمية كل سلعة عند حساب الرقم القياسى التجميعى ، نعطى كل سلعة وزناً يمثل الكميات المنتجة أو المستهلكة منها . وهنا تعرض لنا مشكلة المفاضلة بين اختيار كميات سنة الأساس أو كميات السنة المقارنة أو متوسط الأولى والثانية . وفيما يلى عرض موجز لهذه الطرق المختلفة .

(ا) استخدام كميات سنة الأساس كأوزان

على فرض أن الكميات المستهلكة من السلع فى سنة الأساس هى على التوالى :

$ك_{س_1}$ ، $ك_{س_2}$ ، $ك_{س_3}$ ، . . . . $ك_{س_ر}$

فإن الرقم التجميعى المرجح بكميات سنة الأساس :

$$= \frac{ع_{ن_1} ك_{س_1} + ع_{ن_2} ك_{س_2} + \ldots . ع_{ن_ر} ك_{س_ر}}{ع_{س_1} ك_{س_1} + ع_{س_2} ك_{س_2} + \ldots . ع_{س_ر} ك_{س_ر}} \times ١٠٠$$

$$= \frac{\Sigma ع_{ن} ك_{س}}{\Sigma ع_{س} ك_{س}} \times ١٠٠$$

وبالتأمل فى هذا الرقم نجد أنه متحيزاً إلى أعلى . ذلك أن الكميات المستهلكة

من السلع المختلفة لا تتغير بنفس النسبة كما أن الأسعار لا تتغير بنفس النسبة . فلو فرض أن ارتفعت أسعار بعض السلع بينما انخفضت أسعار البعض الآخر فإنه يترتب على ذلك نقص الكميات المستهلكة من الفئة الأولى وزيادة الكميات المستهلكة من الفئة الثانية طبقاً للعلاقة المعروفة بين الثمن والطلب . وعلى ذلك فإن السلع التى ارتفعت فى السعر تعطى وزناً أكبر من أهميتها الحقيقية بينما تعطى السلع التى انخفضت فى السعر وزناً أقل من أهميتها الحقيقية فى وقت تركيب الرقم القياسى . وحيث أننا نعطى وزناً أكبر مما يجب للسلع التى ارتفعت فى الثمن ووزناً أقل مما يجب للسلع التى انخفضت فى الثمن فإن الرقم القياسى يكون متحيزاً إلى أعلى .

(ب) استخدام كميات السنة المقارنة كأوزان

على فرض أن كميات السنة المقارنة هى على التوالى

$ك_{ن_1}$ ، $ك_{ن_2}$ ، $ك_{ن_3}$ ، . . . . $ك_{ن_ر}$

فإن الرقم التجميعى المرجح بهذه الكميات يكون $\frac{\Sigma ع_{ن} \; ك_{ن}}{\Sigma ع_{س} \; ك_{ن}} \times ١٠٠$

وتركيب الرقم القياسى بهذه الطريقة يتطلب من الإحصائى وقتاً ومجهوداً أكبر كثيراً مما تتطلبه الطريقة السابقة . إذ لا بد فى هذه الطريقة من تعيين أوزان جديدة للسلع فى كل سنة نظراً لتغير الكميات المستهلكة من سنة إلى أخرى . وكما وجدنا أن الطريقة السابقة تجعل الرقم القياسى متحيزاً إلى أعلى فإن هذه الطريقة تجعله متحيزاً إلى أسفل .

ذلك أن السلع التى انخفضت أثمانها تزداد الكمية المستهلكة منها عادة ولذلك فهى تعطى أهمية أكثر مما يجب لمجرد أن ثمنها قد انخفض ، وكأن تغير الثمن يقرر إلى حد بعيد الأوزان التى نستخدمها لقياس تغير الثمن نفسه .

(ح) استخدام متوسط كميات السنة المقارنة وسنة الأساس كأوزان

رأينا أن استخدام كميات سنة الأساس يجعل الرقم القياسى متحيزاً إلى أعلى وأن استخدام كميات السنة المقارنة يجعل الرقم متحيزاً إلى أسفل . وقد اقترح الأستاذ مارشال (Marshall) للتخلص من هذه الصعوبة استخدام متوسط كميات السنة المقارنة وسنة الأساس كأوزان . وعلى هذا الأساس يكون الرقم

$$\text{القياسى} = \frac{\Sigma ع_{ن} \quad (ك_{ن} + ك_{س})}{\Sigma ع_{س} \quad (ك_{ن} + ك_{س})} \times ١٠٠$$

وهذه الطريقة تواجه نفس الصعوبات العملية التى تواجهها الطريقة السابقة والتى تنشأ من ضرورة تغيير الأوزان من سنة إلى أخرى .

الرقم القياسى الأمثل Fisher's Ideal Index Number

تتلخص هذه الطريقة فى تركيب رقمين قياسيين مستقلين ، نستخدم فى أحدهما أوزان سنة الأساس كما فعلنا فى الطريقة « ا » ونستخدم فى الثانى أوزان السنة المقارنة كما فعلنا فى الطريقة « ب » والرقم القياسى الأمثل هو الوسط الهندسى لهذين الرقمين . أى أنه يساوى :

$$١٠٠ \sqrt{\frac{\Sigma ع_{ن} \quad ك_{س}}{\Sigma ع_{س} \quad ك_{س}} \times \frac{\Sigma ع_{ن} \quad ك_{ن}}{\Sigma ع_{س} \quad ك_{ن}}}$$

وسنرى مزايا هذا الرقم عند شرح طرق اختبار صحة الأرقام القياسية .

ثانياً : الأرقام النسبية

أساس هذه الطريقة هو حساب سعر كل سلعة فى السنة المقارنة كنسبة

مئوية من سعرها فى سنة الأساس . وتسمى هذه النسب المئوية مناسيب الأسعار (Price Relatives) وبعد أن نحسب منسوب سعر كل سلعة نعين متوسط المناسيب عن طريق استخراج وسطها الحسابى أو الهندسى أو التوافقى فيكون هذا المتوسط هو الرقم القياسى المطلوب . وفيما يلى الطرق المختلفة لحساب متوسط المناسيب :

(ا) الوسط الحسابى للمناسيب

إذا فرضنا أن لدينا (ن) من السلع وأن أسعارها فى السنة المقارنة محتسبة كنسب مئوية من أسعارها فى سنة الأساس هى على التوالى :

س١ ، س٢ ، س٣ ، . . . . ، سن

فإن الرقم القياسى للأسعار = $\frac{س_1 + س_2 + س_3 + \cdots س_ن}{ن} = \frac{\Sigma س}{ن}$

هذا على فرض أننا اعتبرنا جميع السلع متعادلة فى الأهمية. أما إذا أعطينا كل سلعة وزناً يتناسب مع أهميتها وكانت هذه الأوزان :

و١ ، و٢ ، و٣ ، . . . . ، ون

فإن الرقم القياسى للأسعار = $\frac{\Sigma س و}{\Sigma و}$

(ب) الوسط الهندسى للمناسيب

باستخدام نفس الرموز السابقة يكون الوسط الهندسى البسيط لمناسيب

الأسعار = $\sqrt[ن]{س_1 \times س_2 \times \cdots س_ن}$

والوسط الهندسى المرجح بالأوزان

$$= \sqrt[\Sigma و]{(س_1)^{و_1} \times (س_2)^{و_2} \times \cdots (س_ن)^{و_ن}}$$

ويمكن طبعاً تبسيط هذه الكمية بأخذ لوغاريتمات الطرفين كالآتى :

$$\text{لو (الوسط الهندسى)} = \frac{1}{\Sigma و}(و_1 \text{لو} س_1 + و_2 \text{لو} س_2 + \cdots + و_ن \text{لو} س_ن)$$

ومن الممكن كذلك تركيب الرقم القياسى عن طريق تعيين الوسط التوافقى لمناسيب الأسعار غير أن هذه الطريقة نادرة الاستعمال وليس هناك ما يبرر تفضيلها لا من الوجهة النظرية ولا من وجهة المناسبة العملية .

اختبار صحة الأرقام القياسية

رأينا أن هناك عدداً كبيراً من الطرق المختلفة لتركيب الأرقام القياسية . ومن الواضح أن كل طريقة من هذه الطرق تعطينا نتيجة تختلف عن تلك التى نحصل عليها بأية طريقة أخرى . ولذلك فلا بد أن نبحث عن وسيلة للمفاضلة بين هذه الطرق المختلفة كى نعلم أى النتائج التى نحصل عليها أقرب إلى الدقة . وقد وضع الأستاذ « إرفنج فيشر » بحثاً تفصيليا فى هذا الموضوع ضمنه بعض الاختبارات التى اقترح تطبيقها على الصيغ المختلفة لتركيب الأرقام القياسية كما أوضح فيه طرق تعديل هذه الصيغ لجعلها أقرب إلى الدقة . وأهم الاختبارات التى اقترح تطبيقها هى :

أولا : اختبار الانعكاس الزمنى Time Reversal Test

ثانياً : اختبار الانعكاس المعاملى Factor Reversal Test

وفيما يلى شرح موجز لهذين الاختبارين[١] .

اختبار الانعكاس الزمنى

إذا كان سعر سلعة معينة فى سنة ١٩٣٧ يساوى ٢٠٠ ٪ من سعرها فى سنة ١٩١٤ فمن البديهى أن سعر السلعة فى سنة ١٩١٤ يساوى ٥٠ ٪ من سعرها فى سنة ١٩٣٧ وقياساً على ذلك إذا كان الرقم القياسى للأسعار فى سنة ١٩٣٧ يساوى ٢٠٠ ٪ على اعتبار أن سنة الأساس هى سنة ١٩١٤ فلا شك أنه إذا عكسنا وضع هاتين السنتين فجعلنا سنة ١٩٣٧ سنة الأساس وسنة ١٩١٤ السنة المقارنة فلا بد أن يكون الرقم القياسى الذى نحصل عليه هو ٥٠ ٪ أو بعبارة أخرى إذا عكسنا وضع سنة الأساس والسنة المقارنة واستخرجنا رقماً قياسيا جديداً فلا بد أن يكون حاصل ضرب الرقم الأصلى فى الرقم الجديد = ١

هذا هو اختبار الانعكاس الزمنى الذى اقترحه الأستاذ « فيشر » . وعلى الرغم من أن تحقق شرط الانعكاس الزمنى يبدو كبديهية واضحة لا تحتاج إلى اختبار فإن كثيراً من الصيغ التى ذكرناها آنفاً لا تحقق هذا الشرط كما سيتضح من تطبيق الاختبار عليها . ولتطبيق هذا الاختبار نضرب كل رقم فى بديله الزمنى[٢] الذى نحصل عليه من تبديل وضع سنة الأساس والسنة المقارنة لنرى ما إذا كان حاصل الضرب يكون الواحد الصحيح .

ولنتناول الآن بعض الصيغ السالفة الذكر لنرى أيها يحقق اختبار الانعكاس فى الزمن .

---

(١) هناك اختبار ثالث هو الاختبار الدائرى (Circular Test) غير أن هذا الاختبار قليل الأهمية من الناحية العمليه كما أن الأساس النظرى الذى ينبنى عليه خاطىء كما أوضح ذلك الأستاذ Fisher فى كتابه «The Making of Index Numbers» ص ٢٧١

(٢) البديل الزمنى معناه بالإنجليزية (Time Reciprocal)

أولا : الرقم التجميعى البسيط

حيث أن هذا الرقم (*) $= \frac{\Sigma ع_{ن}}{\Sigma ع_{س}}$

(حيث $ع_{ن}$ تمثل أسعار السنة المقارنة ، $ع_{س}$ تمثل أسعار سنة الأساس)

فإن بديله الزمنى $= \frac{\Sigma ع_{س}}{\Sigma ع_{ن}}$

وحيث أن $\frac{\Sigma ع_{ن}}{\Sigma ع_{س}} \times \frac{\Sigma ع_{س}}{\Sigma ع_{ن}} = ١$

∴ هذا الرقم يحقق شرط الانعكاس الزمنى

ومما تجب ملاحظته أن مجرد تحقيق شرط الانعكاس الزمنى لا يجعل هذا الرقم كاملا . فقد ذكرنا من قبل أن هذا الرقم لا يصلح مطلقاً إذا كان هناك تفاوت كبير بين وحدات السلع الداخلة فى تركيبه .

ثانياً : الرقم التجميعى المرجح بالأوزان

(ا) الرقم التجميعى المرجح بأوزان سنة الأساس $= \frac{\Sigma ع_{ن}\ ك_{س}}{\Sigma ع_{س}\ ك_{س}}$

∴ البديل الزمنى $= \frac{\Sigma ع_{س}\ ك_{ن}}{\Sigma ع_{ن}\ ك_{ن}}$

(*) يلاحظ أننا سنهمل ضرب الصيغ المختلفة للأرقام القياسية فى ١٠٠ لتسهيل الكتابة وعلى الطالب أن يتذكر أن الرقم القياسى يكون دائما على شكل نسبة مئوية .

ولكن $\frac{\Sigma ع_{ن} ك_{س}}{\Sigma ع_{س} ك_{س}} \times \frac{\Sigma ع_{س} ك_{ن}}{\Sigma ع_{ن} ك_{ن}}$ لا يساوى ١

∴ الرقم لايحقق شرط الانعكاس فى الزمن . وبنفس الطريقة يمكن إثبات أن الرقم التجميعى المرجح بأوزان السنة المقارنة لا يحقق هذا الشرط .

(ب) الرقم التجميعى المرجح بمتوسط أوزان سنة الأساس والسنة المقارنة

$$= \frac{\Sigma ع_{ن} (ك_{س} + ك_{ن})}{\Sigma ع_{س} (ك_{س} + ك_{ن})}$$

وحيث أن $\frac{\Sigma ع_{ن} (ك_{س} + ك_{ن})}{\Sigma ع_{س} (ك_{س} + ك_{ن})} \times \frac{\Sigma ع_{س} (ك_{ن} + ك_{س})}{\Sigma ع_{ن} (ك_{ن} + ك_{س})} = ١$

∴ هذا الرقم يحقق شرط الانعكاس الزمنى .

(حـ) الرقم القياسى الأمثل $= \sqrt{\frac{\Sigma ع_{ن} ك_{س}}{\Sigma ع_{س} ك_{س}} \times \frac{\Sigma ع_{ن} ك_{ن}}{\Sigma ع_{س} ك_{ن}}}$

وبما أن حاصل ضرب هذا الرقم فى بديله الزمنى = ١

∴ الرقم القياسى الأمثل يحقق شرط الإنعكاس فى الزمن

وبتطبيق اختبار الانعكاس الزمنى على الأرقام النسبية السابق بيانها نجد أن الصيغة الوحيدة التى تحقق هذا الشرط هى الوسط الهندسى البسيط للمناسيب .

أما الوسط الهندسى المرجح بالأوزان فلا يحقق شرط الانعكاس الزمنى .

٢ ــ اختبار الانعكاس المعاملى

إذا حسبنا رقماً قياسيـــا لسعر سلعة واحدة ثم حسبنا رقماً قياسياً آخر

للكمية المستهلكة من هذه السلعة فلا بد أن يكون حاصل ضرب هذين الرقمين يمثل الرقم القياسى لقيمة السلعة . فإذا كان منسوب سعر السلعة = $\frac{ع_ن}{ع_س}$

وكان منسوب الكمية = $\frac{ك_ن}{ك_س}$ فمن الواضح أن

$$\frac{ع_ن}{ع_س} \times \frac{ك_ن}{ك_س} = \frac{\text{قيمة السلعة فى السنة المقارنة}}{\text{قيمة السلعة فى سنة الأساس}} = \text{منسوب القيمة}(*)$$

ولاشك أنه ما دام هذا الاختبار صحيحاً بالنسبة لمنسوب سلعة واحدة فلا بد أن يكون صحيحاً بالنسبة للرقم القياسى لعدة سلع وإلا كان الرقم القياسى غير منطقى . وبعبارة أخرى فإذا حسبنا رقماً قياسيا لأسعار بعض السلع باستعمال صيغة معينة ثم حسبنا رقماً قياسياً لكميات هذه السلع باستعمال نفس الصيغة فلا بد أن يكون حاصل ضرب الرقمين

$$= \frac{\Sigma\, ع_ن ك_ن}{\Sigma\, ع_س ك_س} = \text{الرقم القياسى للقيمة}$$

هذا هو اختبار الانعكاس المعاملى . ولتوضيح دلالة هذا الاختبار وأهميته افرض أن لدينا سلعتين كالقطن والقمح . وافرض أن سعر القطن فى السنة المقارنة ضعف ما كان عليه فى سنة الأساس وأن الكمية المستهلكة من القطن نصف ما كانت عليه فى سنة الأساس . وافرض كذلك أنه لم يطرأ أى تغيير على سعر القمح أو الكمية المستهلكة منه بين سنة الأساس والسنة المقارنة.

بناء على هذه البيانات نستطيع أن نحصل على النتائج الآتية :

منسوب سعر القطن = ٢٠٠ ٪ ومنسوب كميته = ٥٠ ٪

(*) لاحظ أن منسوب سعر سلعة واحدة هو نفسه الرقم القياسى لسعر هذه السلعة .

منسوب سعر القمح = ١٠٠ ٪ ومنسوب كميته = ١٠٠ ٪

∴ الرقم القياسى للأسعار باستخدام الوسط الحسابى البسيط للمناسيب

$$= \frac{١٠٠+٢٠٠}{٢} = ١٥٠ \text{ ٪}$$

والرقم القياسى للكميات باستخدام الوسط الحسابى البسيط للمناسيب

$$= \frac{١٠٠+٥٠}{٢} = ٧٥ \text{ ٪}$$

حاصل ضرب الرقمين = $١٠٠ \left(\frac{١٥٠}{١٠٠} \times \frac{٧٥}{١٠٠}\right) = ١١٢,٥$ ٪

فكأن الرقم القياسى للقيمة = ١١٢,٥ ٪

وهذا خطأ طبعاً فإن قيمة القطن فى السنة المقارنة لا تختلف مطلقاً عن قيمة القطن فى سنة الأساس . فقد ارتفع السعر إلى الضعف بينما انخفضت الكمية المستهلكة إلى النصف وعلى ذلك تظل القيمة الإجمالية للمستهلك من القطن ثابتة .

أما القمح فقد ذكرنا أن كلا من سعره وكميته لم يتغير وعلى ذلك فإن قيمته لم تتغير . وحيث أنه لم يطرأ أى تغير على قيمة السلعتين بين سنة الأساس والسنة المقارنة فلا بد أن يكون الرقم القياسى للقيمة ١٠٠ ٪ وليس ١١٢,٥ ٪ .

ومن هذا يتضح أن الوسط الحسانى البسيط للمناسيب يعطينا نتائج خاطئة لأنه لا يحقق شرط الانعكاس المعاملى .

ولنأخذ مثالا آخر يوضح خطأ طريقة الوسط الحسابى للمناسيب .

افرض أن سعر القطن فى السنة المقارنة ارتفع إلى ضعف ما كان عليه فى سنة الأساس وأن كميته فى السنة المقارنة ازدادت إلى ثلاثة أمثال ما كانت عليه فى سنة الأساس . وافرض أن سعر القمح فى السنة المقارنة ارتفع إلى ثلاثة أمثال ما كان عليه فى سنة الأساس وأن كميته أصبحت ضعف ما كانت عليه فى سنة الأساس .

من الواضح أن القيمة الإجمالية للقطن والقمح معا ازدادت إلى ستة أمثال

ما كانت عليه فى سنة الأساس أى أن الرقم القياسى للقيمة = ٦٠٠ ٪

غير أنه بحساب الرقم القياسى للأسعار والرقم القياسى للكميات بطريقة الوسط الحسابى للمناسيب ، نجد أن حاصل ضربهما لا يساوى ٦٠٠ ٪

فالرقم القياسى للأسعار باستخدام الوسط الحسابى للمناسيب

$$= \frac{٣٠٠ + ٢٠٠}{٢} = ٢٥٠ \text{ ٪}$$

والرقم القياسى للكميات باستخدام الوسط الحسابى للمناسيب

$$= \frac{٢٠٠ + ٣٠٠}{٢} = ٢٥٠ \text{ ٪}$$

وبناء على هذه النتائج يكون الرقم القياسى للقيمة = ١٠٠ $\left(\frac{٢٥٠}{١٠٠} \times \frac{٢٥٠}{١٠٠}\right)$

= ٦٢٥ ٪

وهذه نتيجة خاطئة طبعاً إذ الواقع أن القيمة فى السنة المقارنة زادت إلى ستة أمثال ما كانت عليه فى سنة الأساس كما سبق أن بينا .

ولا شك أن مصدر هذا الخطأ هو طريقة تركيب الرقم القياسى ا

ولكى تكون طريقة تركيب الرقم القياسى صحيحة فلابد أن يكون حاصل ضرب الرقم القياسى للأسعار فى الرقم القياسى للكميات (المستخرج بنفس الطريقة يساوى الرقم القياسى للقيمة . إذ بهذا يتحقق شرط الانعكاس المعاملى فى الرقم القياسى .

وللحصول على الرقم القياسى للكميات بنفس الطريقة التى حصلنا بها على الرقم القياسى للأسعار نرجع إلى الصيغة التى حسبنا بها الرقم القياسى للأسعار ونبدل فيها كل (ع) بـ (ك) وكل (ك) بـ (ع) والصيغة الجديدة بعد هذا هذا التبديل تسمى البديل المعاملى[1] ، ولتحقيق شرط الانعكاس المعاملى لا بد أن يكون :

الرقم القياسى × بديله المعاملى = الرقم القياسى للقيمة

---

(١) البديل المعامل معناه بالإنجليزية Factor Reciprocal

تطبيق اختبار الانعكاس المعاملى على بعض الأرقام القياسية

(١) الرقم التجميعى البسيط = $\frac{\Sigma ع_ن}{\Sigma ع_س}$ .. البديل المعاملى = $\frac{\Sigma ك_ن}{\Sigma ك_س}$

وحيث أن $\frac{\Sigma ع_ن}{\Sigma ع_س} \times \frac{\Sigma ك_ن}{\Sigma ك_س}$ لا يساوى $\frac{\Sigma ع_ن \, ك_ن}{\Sigma ع_س \, ك_س}$

.. هذا الرقم لا يحقق شرط الانعكاس المعاملى.

(٢) الرقم التجميعى المرجح بأوزان سنة الأساس = $\frac{\Sigma ع_ن \, ك_س}{\Sigma ع_س \, ك_س}$

.. البديل المعاملى = $\frac{\Sigma ك_ن \, ع_س}{\Sigma ك_س \, ع_س}$

ومن الواضح أن حاصل ضرب هذين الرقمين لا يساوى الرقم القياسى للقيمة وبتطبيق هذا الاختبار على الرقم التجميعى المرجح بأوزان السنة المقارنة نجد أنه لا ينعكس فى المعامل.

(٣) الرقم القياسى الأمثل = $\sqrt{\frac{\Sigma ع_ن \, ك_س}{\Sigma ع_س \, ك_س} \times \frac{\Sigma ع_ن \, ك_ن}{\Sigma ع_س \, ك_ن}}$

.. البديل المعاملى = $\sqrt{\frac{\Sigma ك_ن \, ع_س}{\Sigma ك_س \, ع_س} \times \frac{\Sigma ك_ن \, ع_ن}{\Sigma ك_س \, ع_ن}}$

ومن الواضح أن حاصل ضرب هذين الرقمين

$= \sqrt{\left(\frac{\Sigma ع_ن \, ك_ن}{\Sigma ع_س \, ك_س}\right)^2} = \frac{\Sigma ع_ن \, ك_ن}{\Sigma ع_س \, ك_س}$

ومن هذا يتضح أن الرقم القياسى الأمثل يحقق شرط الانعكاس المعاملى علاوة على تحقيقه لشرط الانعكاس الزمنى .

أما الأرقام النسبية فليس بينها رقم واحد يحقق شرط الانعكاس المعاملى . وعلى ذلك يكون الرقم القياسى الأمثل هو الرقم الوحيد الذى يحقق شرطى الانعكاس الزمنى والانعكاس المعاملى معاً . ومن هنا كانت تسميته بالرقم الأمثل .

تعديل الأرقام القياسية لجعلها تنعكس فى الزمن وفى المعامل

يمكن تعديل أى صيغة تستعمل لتركيب الرقم القياسى لجعلها تنعكس فى الزمن والمعامل . ولنبدأ بتعديل الأرقام لجعلها تحقق الاختبار الأول .

يمكن تعديل أى رقم قياسى لجعله يحقق شرط الانعكاس الزمنى بأن نضرب هذا الرقم فى مقلوب بديله الزمنى ثم نستخرج الجذر التربيعى لحاصل الضرب . وبعبارة أخرى فإن الرقم القياسى المعدل يساوى الوسط الهندسى للرقم الأصلى ومقلوب بديله الزمنى . وللاختصار فى التعبير نسمى مقلوب البديل الزمنى المقلوب الزمنى (Time Antithesis)

ويلاحظ أنه إذا كان الرقم الأصلى ينعكس فى الزمن من تلقاء نفسه (أى بدون حاجة إلى تعديله) فإن المقلوب الزمنى يساوى الرقم نفسه .

ولإيضاح طريقة التعديل نأخذ الرقم التجميعى المرجح بأوزان سنة الأساس وهو كما نعلم من الأرقام التى لا تنعكس فى الزمن ثم نجرى عليه التعديل المذكور .

$$\text{الرقم الأصلى} = \frac{\Sigma\, ع_{ن}\, ك_{س}}{\Sigma\, ع_{س}\, ك_{س}} \quad \therefore \text{ بديله الزمنى} = \frac{\Sigma\, ع_{س}\, ك_{ن}}{\Sigma\, ع_{ن}\, ك_{ن}}$$

$$\text{ومقلوبه الزمنى (أى مقلوب البديل الزمنى)} = \frac{\text{ع ن ك ن}}{\text{ع س ك ن}}$$

$$\therefore \text{الرقم المعدل} = \sqrt{\frac{\text{ع ن ك س}}{\text{ع س ك س}} \times \frac{\text{ع ن ك ن}}{\text{ع س ك ن}}}$$

ويلاحظ أن هذا الرقم المعدل هو نفس الرقم القياسى الأمثل الذى رأينا أنه ينعكس فى الزمن . ومن الواضح أنه لو كان الرقم الأصلى مرجحاً بكميات السنة المقارنة وعدلناه بهذه الكيفية فإننا نحصل على الرقم الأمثل كذلك وبنفس هذه الطريقة يمكن تعديل الأرقام النسبية .

ولتعديل الأرقام القياسية لجعلها تحقق شرط الانعكاس المعاملى نحسب الوسط الهندسى للرقم الأصلى ومقلوبه المعاملى(Factor Antithesis) كما فعلنا فى الحالة السابقة .

والمقلوب المعاملى لأى رقم هو خارج قسمة الرقم القياسى للقيمة على البديل المعاملى للرقم المطلوب تعديله . ويلاحظ أن الرقم القياسى للقيمة هو دائماً

$$\frac{\text{ع ن ك ن}}{\text{ع س ك س}}$$

ولتوضيح عمل هذا التعديل نأخذ على سبيل التمثيل الرقم التجميعى المرجح بكميات سنة الأساس الذى نعلم أنه لا ينعكس فى المعامل .

$$\text{الرقم الأصلى} = \frac{\text{ع ن ك س}}{\text{ع س ك س}} \quad \therefore \text{ بديله المعاملى} = \frac{\text{ك ن ع س}}{\text{ك س ع س}}$$

$$\therefore \text{مقلوبه المعاملى} = \frac{\text{ع ن ك ن}}{\text{ع س ك س}} \div \frac{\text{ك ن ع س}}{\text{ك س ع س}} = \frac{\text{ع ن ك ن}}{\text{ع س ك ن}}$$

... الوسط الهندسى للرقم الأصلى ومقلوبه المعاملى :

$$= \sqrt{\frac{\text{ع ن ك ص}}{\text{ع س ك ص}} \times \frac{\text{ع ن ك ن}}{\text{ع س ك ن}}}$$

وهذا هو الرقم الأمثل الذى نعرف أنه ينعكس فى المعامل .

ويلاحظ أنه إذا كان الرقم لا ينعكس فى الزمن ولا فى المعامل وجب تعديله مرتين ، فنعدله أولا لنجعله يحقق شرط الإنعكاس فى الزمن ثم نتناول الصيغة المعدلة فنعدلها مرة أخرى لنجعلها تحقق شرط الانعكاس فى المعامل . على أن تعديل الأرقام القياسية بهذه الكيفية لا يجعلها تبلغ حد الكمال . فكثيراً ما تكون الصيغة المعدلة من التعقيد بحيث يصعب استعمالها . كما أن الصيغة النهائية المعدلة كثيراً ما تأخذنا بعيداً جداً عن الغرض الأصلى الذى نرمى إليه من تركيب الرقم القياسى وهو قياس متوسط التغير النسبى فى قيم الظواهر . ولا شك أن الرقم القياسى الأمثل أفضل الأرقام السابقة فهو يحقق شرطى الانعكاس الزمنى والمعاملى معاً فضلا عن أن صيغته خالية من التعقيد .

الأرقام القياسية المصرية

١ — الرقم القياسى لأسعار الجملة[1]

أولا : الرقم القديم

بدأت مصلحة الإحصاء بنشر رقم قياسى شهرى لأسعار الجملة منذ سنة ١٩١٤ ويدخل فى تركيب هذا الرقم ٢٦ سلعة تمثل أهم المحاصيل الزراعية وتشمل أيضاً بعض منتجات الصناعة كالصابون والسكر وزيت البترول . وتحصل المصلحة على بيانات خاصة بأسعار هذه السلع عن طريق إرسال

---

(١) يستطيع الطالب أن يطلع على مذكرة مصلحة الإحصاء بخصوص طريقة تركيب هذا الرقم فى الإحصاء السنوى العام لسنة ١٩٣٥ (ص ٤٨٧) هى السنة التى عدلت فيها المصلحة هذه الطريقة .

كشوف للتجار يوضحون بها أسعار السلع وانتداب بعض موظفى المصلحة لجمع بيانات عن الأسعار من أسواق القاهرة والإسكندرية . إذ كان يعمل رقم خاص للقاهرة وآخر للإسكندرية . وفى بعض الحالات كانت المصلحة تجمع بيانات عن عدة أصناف من نفس السلعة لغرض تركيب الرقم . فكانت تأخذ من السكر مثلا خمسة أصناف ومن الفحم صنفين فكان عدد الأصناف ٣٩ بينما كان عدد السلع ٢٦ فقط . على أن تعدد الأسعار الخاصة بالسلعة الواحدة لم يكن المقصود منه ترجيح أى سلعة فى الأهمية على غيرها (كما هو الحال فى الرقم الجديد) وإنما كان المقصود منه حساب متوسط للأسعار المختلفة للسلعة الواحدة وكان هذا المتوسط وحده هو الذى يدخل فى تركيب الرقم القياسى أما سنة الأساس فقد اعتبرت الفترة الواقعة بين يناير سنة ١٩١٣ ويوليه سنة ١٩١٤ أى أن سعر كل سلعة فى الشهر المقارن ينسب إلى متوسط سعر السلعة فى هذه الفترة . وبعد حساب مناسيب الأسعار بهذه الطريقة يكون الرقم القياسى العام لأسعار الجملة هو الوسط الهندسى البسيط للمناسيب . أى أن جميع السلع اعتبرت متساوية فى الأهمية وهذا طبعاً عيب كبير فى هذا الرقم .

ثانياً : الرقم الجديد

فى سنة ١٩٣٥ شعرت المصلحة بأن هذا الرقم الجديد أصبح لا يصف اتجاهات الأسعار فى مصر وصفاً دقيقاً . فمن جهة ابتعدت سنة الأساس كثيراً عن السنوات المقارنة وبذلك أصبحت المقارنة لا تقوم على أساس صحيح نظراً لاختلاف الظروف المحيطة بانتاج واستهلاك السلع فى الحالتين . ومن الجهة الأخرى فإن زيادة أهمية السلع المصنوعة فى مصر على أثر النهضة الصناعية التى أعقبت الحرب العالمية الأولى والتى دعمت باتباع مبدأ الحماية الجمركية المعتدلة والأخذ بسياسة منح الإعانات الحكومية لبعض الصناعات الأساسية جعلت الرقم القياسى القديم فى حاجة إلى تعديل كى يكون متمشياً مع التطور الاقتصادى للبلاد . وقد انتهزت المصلحة فرصة تركيب رقم جديد لتلافى العيوب التى كانت موجودة فى طريقة تركيب الرقم القديم . فزاد عدد السلع التى تدخل فى تركيبه من ٢٦ إلى ٨٧ بينها حاصلات زراعية وسلع

مصنوعة منتجة محلياً أو مستوردة . كذلك اتسعت رقعة المصادر التى تستقى منها المصلحة بياناتها عن أسعار السلع فأصبح عدد التجار الذين يمدونها بالبيانات نحواً من ٣٠٠ تاجر كما زاد عدد المندوبين الذين يجمعون البيانات الإحصائية من الأسواق مباشرة . واكتفت المصلحة بأن تنشر رقماً قياسياً واحداً يمثل القطر كله بدلا من عمل رقمين مستقلين أحدهما للقاهرة والآخر للإسكندرية . أما عن طريقة تركيب الرقم فأهم التعديلات التى أدخلت عليها كانت كالآتى :

أولا : اعتبر أساس المقارنة متوسط السعر فى سنة ١٩٣٥ بدلا من سنة ١٩١٣ — ١٩١٤ .

ثانياً : اتبعت طريقة الأساس المتحرك — أى أن السعر فى الشهر الأول من سنة ١٩٣٦ ينسب إلى متوسط السعر فى سنة ١٩٣٥ وبعد ذلك ينسب السعر فى كل شهر إلى السعر فى الشهر السابق له مباشرة . وهذه الطريقة أفضل من طريقة الأساس الثابت لأنها طريقة مرنة تسمح بإدخال بعض السلع الجديدة فى تركيب الرقم وإهمال تلك التى تختفى من الأسواق أو تقل أهميتها كثيراً.

ثالثاً : أعطى لكل سلعة وزن يتناسب مع أهميتها عن طريق زيادة عدد أصناف السلع المهمة التى تدخل فى تركيب الرقم . فمثلا كان يؤخذ من القطن ٣٩ صنفاً يحسب لكل منها منسوب خاص عند حساب الرقم القياسى العام . كذلك كان يحسب للقمح ١٥ منسوباً تدخل جميعها فى تركيب الرقم النهائى ولذلك نجد أن عدد المناسيب ١٩٢ فى حين أن عدد السلع ٨٧ فقط . أما طريقة حساب متوسط المناسيب فلم تتغير بل ظلت الوسط الهندسى البسيط للمناسيب . وما دامت السلع الهامة قد رجحت بزيادة عدد المناسيب التى تمثلها فإن أهميتها تظهر طبعاً فى الرقم النهائى ولا داعى لترجيحها بالأوزان مرة أخرى .

٢ — الرقم القياسى لنفقات المعيشة

يقاس مستوى المعيشة بكمية ما يستهلكه الفرد من السلع والخدمات فى وحدة الزمن . فإذا قلنا إن مستوى المعيشة فى أمريكا أعلى منه فى الهند فإنما نقصد بذلك أن ما يستهلكه الفرد المتوسط فى الدولة الأولى أكثر مما يستهلكه الفرد

المتوسط فى الثانية . وكثيراً ما يؤدى التهاون فى التعبير إلى الخلط بين مستوى المعيشة (Standard of Living) ونفقات المعيشة (Cost of Living) مع أن الفكرتين مختلفتان تماما . فقد يكون مستوى المعيشة فى جهة معينة أعلى منه فى جهة أخرى ومع ذلك فإن نفقات المعيشة فى الجهة الأولى قد تكون أقل منها فى الثانية نظراً لاختلاف مستوى الأسعار فى الحالتين . والتمييز بين الاصطلاحين نقول إن ارتفاع مستوى المعيشة معناه ارتفاع الدخل الحقيقى للفرد أو زيادة مقدرته على إشباع حاجاته فى حين أن ارتفاع نفقات المعيشة يكون مصحوباً عادة بهبوط فى الدخل الحقيقى نظراً لأنه ينشأ عن هبوط القوة الشرائية لوحدة النقود . والرقم القياسى لنفقات المعيشة يقيس التغير فى الأسعار التى يدفعها المستهلك مباشرة أى أسعار التجزئة والسلع التى تدخل فى تركيبه هى السلع الشائعة الاستهلاك ومعظمها من الضروريات القصوى للحياة كالخبز واللحم غير أن عدداً كبيراً منها يعتبر من الكماليات التقليدية كالدخان . وبالإضافة إلى هذه السلع يدخل فى تركيب الرقم إيجارات المنازل باعتبارها تستنفد جزءاً كبيراً من ميزانية الفرد المتوسط .

والطريقة المتبعة عادة فى تركيب الرقم القياسى لنفقات المعيشة هى أن تعمل أرقام قياسية مستقلة لكل من أسعار المواد الغذائية والملابس وإيجارات المساكن وغير ذلك ثم تجمع هذه الأرقام لتركيب الرقم القياسى العام لنفقات المعيشة بطريقة الوسط الحسابى المرجح بالأوزان . وهذه الأوزان تمثل أهمية كل باب من أبواب الإنفاق بالنسبة لغيره .

وفى مصر كانت مصلحة الإحصاء حتى الحرب العالمية الأخيرة تقوم بنشر أرقام قياسية شهرية لنفقات المعيشة على أساس متوسط الأسعار السائدة فى المدة الواقعة بين يناير سنة ١٩١٣ ويوليه سنة ١٩١٤ . وكانت الأوزان المستعملة لتقدير أهمية كل باب من أبواب الإنفاق عند حساب المتوسط العام للتغير فى نفقات المعيشة مستمدة من بحث أجرته المصلحة فى سنة ١٩٢٠ على عدد كبير من أسر صغار الموظفين لمعرفة كيفية توزيع الدخل على أبواب الإنفاق المختلفة .

غير أنه منذ الحرب العالمية الأخيرة رأت المصلحة نظراً لابتعاد سنة الأساس كثيراً عن السنوات المقارنة من جهة ونظراً لتغير طريقة توزيع الدخل من الجهة الأخرى أن تنشر بالإضافة إلى الرقم القديم ، رقماً قياسيا جديداً لنفقات المعيشة بأساس جديد وأوزان جديدة . فاختارت متوسط الأسعار فى الثلاثة أشهر يونيه ويوليه وأغسطس من سنة ١٩٣٩ أساساً جديداً للرقم القياسى كما أجرت بحثاً جديداً للوقوف على التغير فى طريقة توزيع الدخل . والأرقام الآتية توضح مدى اختلاف الأوزان المستخدمة فى تركيب الرقم القديم والرقم الجديد .

أبواب الإنفاق	الأهمية النسبية لكل باب	
	فى الرقم القديم	فى الرقم الجديد
	٪	٪
الطعام (ويشمل الصابون والوقود)	٥١,٩	٤٥,٠
الملابس	١٦,٧	١٦,٧
إيجار المسكن	١١,٧	١٦,٠
أجور الانتقال	١,٤	٣,٠
مصروفات مدرسية	٦,١	٦,٥
سجاير ومصروفات نثرية	٥,٨	٥,٨
متنوعات	٦,٤	٧,٠
	١٠٠	١٠٠

وكما تختلف طريقة توزيع الدخل من وقت إلى آخر فإنها تختلف كذلك من دولة إلى أخرى . ويمكن القول بوجه عام أنه كلما ارتفع مستوى المعيشة كلما قل ذلك الجزء من الدخل المخصص للإنفاق على الغذاء . فى الولايات المتحدة حيث مستوى المعيشة أعلى منه فى أية جهة أخرى نجد أن الإنفاق على الغذاء يستنفد ٣٨,٢ ٪ من الدخل (حسب إحصاء سنة ١٩١٨) فى حين أن هذه النسبة فى الهند تبلغ ٨١,٧ ٪ (حسب إحصاء سنة ١٩١٤) .

جدول مساحات منحنى الخطأ المتماثل

ع / σ	٠,٠٠	٠,٠١	٠,٠٢	٠,٠٣	٠,٠٤	٠,٠٥	٠,٠٦	٠,٠٧	٠,٠٨	٠,٠٩
٠,٠	٠,٠٠٠٠	٠,٠٠٤٠	٠,٠٠٨٠	٠,٠١٢٠	٠,٠١٥٩	٠,٠١٩٩	٠,٠٢٣٩	٠,٠٢٧٩	٠,٠٣١٩	٠,٠٣٥٩
٠,١	٠,٠٣٩٨	٠,٠٤٣٨	٠,٠٤٧٨	٠,٠٥١٧	٠,٠٥٥٧	٠,٠٥٩٦	٠,٠٦٣٦	٠,٠٦٧٥	٠,٠٧١٤	٠,٠٧٥٣
٠,٢	٠,٠٧٩٣	٠,٠٨٣٢	٠,٠٨٧١	٠,٠٩١٠	٠,٠٩٤٨	٠,٠٩٨٧	٠,١٠٢٦	٠,١٠٦٤	٠,١١٠٣	٠,١١٤١
٠,٣	٠,١١٧٩	٠,١٢١٧	٠,١٢٥٥	٠,١٢٩٣	٠,١٣٣١	٠,١٣٦٨	٠,١٤٠٦	٠,١٤٤٣	٠,١٤٨٠	٠,١٥١٧
٠,٤	٠,١٥٥٤	٠,١٥٩١	٠,١٦٢٨	٠,١٦٦٤	٠,١٧٠٠	٠,١٧٣٦	٠,١٧٧٢	٠,١٨٠٨	٠,١٨٤٤	٠,١٨٧٩
٠,٥	٠,١٩١٥	٠,١٩٥٠	٠,١٩٨٥	٠,٢٠١٩	٠,٢٠٥٤	٠,٢٠٨٨	٠,٢١٢٣	٠,٢١٥٧	٠,٢١٩٠	٠,٢٢٢٤
٠,٦	٠,٢٢٥٧	٠,٢٢٩١	٠,٢٣٢٤	٠,٢٣٥٧	٠,٢٣٨٩	٠,٢٤٢٢	٠,٢٤٥٤	٠,٢٤٨٦	٠,٢٥١٨	٠,٢٥٤٩
٠,٧	٠,٢٥٨٠	٠,٢٦١٢	٠,٢٦٤٢	٠,٢٦٧٣	٠,٢٧٠٤	٠,٢٧٣٤	٠,٢٧٦٤	٠,٢٧٩٤	٠,٢٨٢٣	٠,٢٨٥٢
٠,٨	٠,٢٨٨١	٠,٢٩١٠	٠,٢٩٣٩	٠,٢٩٦٧	٠,٢٩٩٥	٠,٣٠٢٣	٠,٣٠٥١	٠,٣٠٧٨	٠,٣١٠٦	٠,٣١٣٣
٠,٩	٠,٣١٥٩	٠,٣١٨٦	٠,٣٢١٢	٠,٣٢٣٨	٠,٣٢٦٤	٠,٣٢٨٩	٠,٣٣١٥	٠,٣٣٤٠	٠,٣٣٦٥	٠,٣٣٨٩
١,٠	٠,٣٤١٣	٠,٣٤٣٨	٠,٣٤٦١	٠,٣٤٨٥	٠,٣٥٠٨	٠,٣٥٣١	٠,٣٥٥٤	٠,٣٥٧٧	٠,٣٥٩٩	٠,٣٦٢١
١,١	٠,٣٦٤٣	٠,٣٦٦٥	٠,٣٦٨٦	٠,٣٧١٨	٠,٣٧٢٩	٠,٣٧٤٩	٠,٣٧٧٠	٠,٣٧٩٠	٠,٣٨١٠	٠,٣٨٣٠
١,٢	٠,٣٨٤٩	٠,٣٨٦٩	٠,٣٨٨٨	٠,٣٩٠٧	٠,٣٩٢٥	٠,٣٩٤٤	٠,٣٩٦٢	٠,٣٩٨٠	٠,٣٩٩٧	٠,٤٠١٥
١,٣	٠,٤٠٣٢	٠,٤٠٤٩	٠,٤٠٦٦	٠,٤٠٨٣	٠,٤٠٩٩	٠,٤١١٥	٠,٤١٣١	٠,٤١٤٧	٠,٤١٦٢	٠,٤١٧٧

١,٢	٠,٣٨٤٩	٠,٣٨٦٩	٠,٣٨٨٨	٠,٣٩٠٧	٠,٣٩٢٥	٠,٣٩٤٤	٠,٣٩٦٢	٠,٣٩٨٠	٠,٣٩٩٧	٠,٤٠١٥
١,٣	٠,٤٠٣٢	٠,٤٠٤٩	٠,٤٠٦٦	٠,٤٠٨٣	٠,٤٠٩٩	٠,٤١١٥	٠,٤١٣١	٠,٤١٤٧	٠,٤١٦٢	٠,٤١٧٧
١,٤	٠,٤١٩٢	٠,٤٢٠٧	٠,٤٢٢٢	٠,٤٢٣٦	٠,٤٢٥١	٠,٤٢٦٥	٠,٤٢٧٩	٠,٤٢٩٢	٠,٤٣٠٦	٠,٤٣١٩
١,٥	٠,٤٣٣٢	٠,٤٣٤٥	٠,٤٣٥٧	٠,٤٣٧٠	٠,٤٣٨٢	٠,٤٣٩٤	٠,٤٤٠٦	٠,٤٤١٨	٠,٤٤٣٠	٠,٤٤٤١
١,٦	٠,٤٤٥٢	٠,٤٤٦٣	٠,٤٤٧٤	٠,٤٤٨٥	٠,٤٤٩٥	٠,٤٥٠٥	٠,٤٥١٥	٠,٤٥٢٥	٠,٤٥٣٥	٠,٤٥٤٥
١,٧	٠,٤٥٥٤	٠,٤٥٦٤	٠,٤٥٧٣	٠,٤٥٨٢	٠,٤٥٩١	٠,٤٥٩٩	٠,٤٦٠٨	٠,٤٦١٦	٠,٤٦٢٥	٠,٤٦٣٣
١,٨	٠,٣٦٤١	٠,٤٦٤٩	٠,٤٦٥٦	٠,٤٦٦٤	٠,٤٦٧١	٠,٤٦٧٨	٠,٤٦٨٦	٠,٤٦٩٣	٠,٤٦٩٩	٠,٤٧٠٦
١,٩	٠,٤٧١٣	٠,٤٧١٩	٠,٤٧٢٦	٠,٤٧٣٢	٠,٤٧٣٨	٠,٤٧٤٤	٠,٤٧٥٠	٠,٤٧٥٨	٠,٤٧٦٢	٠,٤٧٦٧
٢,٠	٠,٤٧٧٣	٠,٤٧٧٨	٠,٤٧٨٣	٠,٤٧٨٨	٠,٤٧٩٣	٠,٤٧٩٨	٠,٤٨٠٣	٠,٤٨٠٨	٠,٤٨١٢	٠,٤٨١٧
٢,١	٠,٤٨٢١	٠,٤٨٢٦	٠,٤٨٣٠	٠,٤٨٣٤	٠,٤٨٣٨	٠,٤٨٤٢	٠,٤٨٤٦	٠,٤٨٥٠	٠,٤٨٥٤	٠,٤٨٥٧
٢,٢	٠,٤٨٦١	٠,٤٨٦٥	٠,٤٨٦٨	٠,٤٨٧١	٠,٤٨٧٥	٠,٤٨٧٨	٠,٤٨٨١	٠,٤٨٨٤	٠,٤٨٨٧	٠,٤٨٩٠
٢,٣	٠,٤٨٩٣	٠,٤٨٩٦	٠,٤٨٩٨	٠,٤٩٠١	٠,٤٩٠٤	٠,٤٩٠٦	٠,٤٩٠٩	٠,٤٩١١	٠,٤٩١٣	٠,٤٩١٦
٢,٤	٠,٤٩١٨	٠,٤٩٢٠	٠,٤٩٢٢	٠,٤٩٢٥	٠,٤٩٢٧	٠,٤٩٢٩	٠,٤٩٣١	٠,٤٩٣٢	٠,٤٩٣٤	٠,٤٩٣٦
٢,٥	٠,٤٩٣٨	٠,٤٩٤٠	٠,٤٩٤١	٠,٤٩٤٣	٠,٤٩٤٥	٠,٤٩٤٦	٠,٤٩٤٨	٠,٤٩٤٩	٠,٤٩٥١	٠,٤٩٥٢
٢,٦	٠,٤٩٥٣	٠,٤٩٥٥	٠,٤٩٥٦	٠,٤٩٥٧	٠,٤٩٥٩	٠,٤٩٦٠	٠,٤٩٦١	٠,٤٩٦٢	٠,٤٩٦٣	٠,٤٩٦٤
٢,٧	٠,٤٩٦٥	٠,٤٩٦٦	٠,٤٩٦٧	٠,٤٩٦٨	٠,٤٩٦٩	٠,٤٩٧٠	٠,٤٩٧١	٠,٤٩٧٢	٠,٤٩٧٣	٠,٤٩٧٤
٢,٨	٠,٤٩٧٤	٠,٤٩٧٥	٠,٤٩٧٦	٠,٤٩٧٧	٠,٤٩٧٧	٠,٤٩٧٨	٠,٤٩٧٩	٠,٤٩٨٠	٠,٤٩٨٠	٠,٤٩٨١
٢,٩	٠,٤٩٨١	٠,٤٩٨٢	٠,٤٩٨٣	٠,٤٩٨٤	٠,٤٩٨٤	٠,٤٩٨٤	٠,٤٩٨٥	٠,٤٩٨٥	٠,٤٩٨٦	٠,٤٩٨٦
٣,٠	٠,٤٩٨٦٥	٠,٤٩٨٧	٠,٤٩٨٧	٠,٤٩٨٨	٠,٤٩٨٨	٠,٤٩٨٨	٠,٤٩٨٩	٠,٤٩٨٩	٠,٤٩٨٩	٠,٤٩٩٠

جدول x^2

درجات الحرية	احتمال الحصول على قيمة x^2 المبينة بالجدول بطريق المصادفة						
	٠,٠٠١	٠,٠١	٠,٠٥	٠,١٠	٠,٥٠	٠,٩٠	٠,٩٩
١	١٠,٨٢٧	٦,٦٣٥	٣,٨٤١	٢,٧٠٦	٠,٤٥٥	٠,٠١٥٨	٠,٠٠٠١٥٧
٢	١٣,٨١٥	٩,٢١٠	٥,٩٩١	٤,٦٠٥	١,٣٨٦	٠,٢١١	٠,٠٢٠١
٣	١٦,٢٦٨	١١,٣٤١	٧,٨١٥	٦,٢٥١	٢,٣٦٦	٠,٥٨٤	٠,١١٥
٤	١٨,٤٦٥	١٣,٢٧٧	٩,٤٨٨	٧,٧٧٩	٣,٣٥٧	١,٠٦٤	٠,٢٩٧
٥	٢٠,٥١٧	١٥,٠٨٦	١١,٠٧٠	٩,٢٣٦	٤,٣٥١	١,٦١٠	٠,٥٥٤
٦	٢٢,٤٥٧	١٦,٨١٢	١٢,٥٩٢	١٠,٦٤٥	٥,٣٤٨	٢,٢٠٤	٠,٨٧٢
٧	٢٤,٣٢	١٨,٤٧٥	١٤,٠٦٧	١٢,٠١٧	٦,٣٤٦	٢,٨٣٣	١,٢٣٩
٨	٢٦,١٢٥	٢٠,٠٩٠	١٥,٥٠٧	١٣,٣٦٢	٧,٣٤٤	٣,٤٩٠	١,٦٤٦
٩	٢٧,٨٧٧	٢١,٦٦٦	١٦,٩١٩	١٤,٦٨٤	٨,٣٤٣	٤,١٦٨	٢,٠٨٨
١٠	٢٩,٥٨٨	٢٣,٢٠٩	١٨,٣٠٧	١٥,٩٨٧	٩,٣٤٢	٤,٨٦٥	٢,٥٥٨
١١	٣١,٢٦٤	٢٤,٧٢٥	١٩,٦٧٥	١٧,٢٧٥	١٠,٣٤١	٥,٥٧٨	٣,٠٥٣
١٢	٣٢,٩٠٩	٢٦,٢١٧	٢١,٠٢٦	١٨,٥٤٩	١١,٣٤٠	٦,٣٠٤	٣,٥٧١
١٣	٣٤,٥٢٨	٢٧,٦٨٨	٢٢,٣٦٢	١٩,٨١٢	١٢,٣٤٠	٧,٠٢٤	٤,١٠٧
١٤	٣٦,١٣٣	٢٩,١٤١	٢٣,٦٨٥	٢١,٠٦٤	١٣,٣٣٩	٧,٧٩٠	٤,٦٦٠
١٥	٣٧,٦٩٧	٣٠,٥٧٨	٢٤,٩٩٦	٢٢,٣٠٧	١٤,٣٣٩	٨,٥٤٧	٥,٢٢٩
١٦	٣٩,٢٥٢	٣٢,٠٠٠	٢٦,٢٩٦	٢٣,٥٤٢	١٥,٣٣٨	٩,٣١٢	٥,٨١٢
١٧	٤٠,٧٩٠	٣٣,٤٠٩	٢٧,٥٨٧	٢٤,٧٦٩	١٦,٣٣٨	١٠,٠٨٥	٦,٤٠٨
١٨	٤٢,٣١٢	٣٤,٨٠٥	٢٨,٨٦٩	٢٥,٩٨٩	١٧,٣٣٨	١٠,٨٦٥	٧,٠١٥
١٩	٤٣,٨٢٠	٣٦,١٩١	٣٠,١٤٤	٢٧,٢٠٤	١٨,٣٣٨	١١,٦٥١	٧,٦٣٣
٢٠	٤٥,٣١٥	٣٧,٥٦٦	٣١,٤١٠	٢٨,٤١٢	١٩,٣٣٧	١٢,٤٤٣	٨,٢٦٠
٢١	٤٦,٧٩٧	٣٨,٩٣٢	٣٢,٦٧١	٢٩,٦١٥	٢٠,٣٣٧	١٣,٢٤٠	٨,٨٩٧
٢٢	٤٨,٢٦٨	٤٠,٢٨٩	٣٣,٩٢٤	٣٠,٨١٣	٢١,٣٣٢	١٤,٠٤١	٩,٥٤٢
٢٣	٤٩,٧٢٨	٤١,٦٣٨	٣٥,١٧٢	٣٢,٠٠٧	٢٢,٣٣٧	١٤,٨٤٨	١٠,١٩٦
٢٤	٥١,١٧٩	٤٢,٩٨٠	٣٦,٤١٥	٣٣,١٩٦	٢٣,٣٣٧	١٥,٦٥٩	١٠,٨٥٦
٢٥	٥٢,٦٢٠	٤٤,٣١٤	٣٧,٦٥٢	٣٤,٣٨٢	٢٤,٣٣٧	١٦,٤٧٣	١١,٥٢٤

دليل بالرموز المستعملة فى الكتاب

الرمز	البيان
Σ (•)	علامة جمع المتغيرات مثلا Σ س=س١ + س٢+٠٠٠٠س ن ويناظر فى بعض الكتب العربية الرمز (مج)
σ	الانحراف المعيارى ويناظر فى بعض الكتب العربية الرمز (ع)
$\bar{س}$ •	الوسط الحسابى
$\bar{س}_ء$	الوسط الفرضى
ح	الانحراف عن الوسط الحسابى
ء	الانحراف عن الوسط الفرضى
ر	معامل الارتباط
ر	معامل الارتباط بين المرتبات
ى	نسبة الارتباط
!	علامة التباديل فمثلا ن != ن (ن—١) ×٠٠٠٠×٣×٢×١
Σ |س|	مجموع قيم س مع إهمال إشاراتها
χ	هذه لها معنى خاص موضح تفصيلا فى صلب الكتاب ص ٢٩٥

• يستعمل بعض المؤلفين الرمز (مج) للدلالة على جمع المتغير . وقد فضلنا استعمال الرمز Σ نظراً لوضوحه واختلافه عن بقية الحروف التى تستعمل للتعبير عن قيم المتغيرات ذاتها.

تمارين مختارة من الامتحانات التى وضعها المؤلف فى كلية التجارة بجامعة فاروق الأول

أولا ـ امتحانات السنة الثالثة

يونية سنة ١٩٤٥ (الورقة الأولى)

١ـ وضح كيف يمكن استخدام بعض خواص المنحنى المتماثل كأساس لقياس درجة التواء المنحنيات ، ثم بين أن مقياس الالتواء الربيعى لا يمكن أن يتغير إلا فى حدود ± 1 .

٢ـ اشرح طريقة الأنسبة الشهرية لحساب النسب الموسمية ، ثم اذكر بعض الفوائد العملية لهذه النسب .

٣ـ اشرح طريقة تعديل الأرقام القياسية النسبية والتجميعية لجعلها تحقق شرطى الانعكاس الزمنى والانعكاس المعاملى .

٤ـ (ا) إذا كانت ح ، هـ ، ق تمثل على التوالى الوسط الحسابى والهندسى والتوافقى لأى كميتين موجبتين .

اثبت أن $ح > هـ > ق$

(ب) وضح العلاقة بين العزم الثالث للتوزيع التكرارى حول الوسط الحسابى والعزم الثالث حول أى وسط فرضى آخر .

يونيه سنة ١٩٤٥ (الورقة الثانية)

أجب على الأسئلة الآتية :

١ــ الجدول الآتى يبين الأرقام القياسية لودائع صناديق التوفير فى القطر المصرى على اعتبار سنة الأساس ١٩٠١ .

السنة	الرقم القياسى للودائع	السنة	الرقم القياسى للودائع
١٩١٩	٢٢١٥	١٩٢٥	٣٨٢٥
١٩٢٠	٢٤٢٦	١٩٢٦	٣٩٧٩
١٩٢١	٢٦٦٩	١٩٢٧	٤٣٤٣
١٩٢٢	٢٩٠٢	١٩٢٨	٤٦٦٠
١٩٢٣	٣٣٧١	١٩٢٩	٤٨٣٥
١٩٢٤	٣٧٠٠	١٩٣٠	٤٩٣٩

أوجد معادلة المستقيم الذى يصف اتجاه هذه الودائع بطريقة المربعات الدنيا ومنها استنتج القيم الاتجاهية للظاهرة فى السنوات الآتية :

١٩١٥ ، ١٩٢٥ ، ١٩٣٥

٢ــ الجدول الآتى يبين أوزان مجموعتين من الجنود « ا » ، « ب »

الوزن بالرطل	تكرار المجموعة « ا »	تكرار المجموعة « ب »	الوزن بالرطل	تكرار المجموعة « ا »	تكرار المجموعة « ب »
١٠٠ وأقل من ١١٠	٧	١٢	١٦٠ وأقل من ١٧٠	٥٩	١٧٥
١١٠ « ١٢٠	٣٩	٧٩	١٧٠ « ١٨٠	١٩	٨٩
١٢٠ « ١٣٠	١٢٣	٢٥٤	١٨٠ « ١٩٠	٥	٣٧
١٣٠ « ١٤٠	١٨١	٤٣٦	١٩٠ « ٢٠٠	٥	١٩
١٤٠ « ١٥٠	١٨٣	٤٠٤	٢٠٠ فأكثر	٣	٨
١٥٠ « ١٦٠	١٢٢	٣٠٨			

بإيجاد معامل الاختلاف لكل من المجموعتين بين أيهما تظهر درجة أكبر من التشتت .

٣ ــ الجدول الآتى يبين التوزيع التكرارى لأجور مجموعة من العمال :

الأجر بالقروش	التكرار	الأجر بالقروش	التكرار
١٢,٥ وأقل من ١٧,٥	٢	٣٧,٥ وأقل من ٤٢,٥	٤
١٧,٥ » ٢٢,٥	٢٢	٤٢,٥ » ٤٧,٥	٦
٢٢,٥ » ٢٧,٥	١٩	٤٧,٥ » ٥٢,٥	١
٢٧,٥ » ٣٢,٥	١٤	٥٢,٥ » ٥٧,٥	١
٣٢,٥ » ٣٧,٥	٣		

والمطلوب :

أولا : رسم المدرج التكرارى (الهستوجرام) لتوزيع الأجور وتعيين قيمة المنوال من الرسم ثم تحقيق النتيجة بطريقة الفروق .

ثانياً : إيجاد قيم الوسيط والربيع الأعلى والربيع الأدنى ثم تعيين درجة التواء المنحنى الذى يمثل هذا التوزيع التكرارى .

٤ ــ الجدول الآتى يبين نسب عدد حالات الأمراض المعدية إلى مجموع السكان فى عدد من المدن :

نسبة الحالات لكل ١٠٠٠ من السكان	عدد المدن	نسبة الحالات لكل ١٠٠٠ من السكان	عدد المدن
٠ وأقل من ٢	٥	١٢ وأقل من ١٤	١٦
٢ » ٤	٣٩	١٤ » ١٦	٧
٤ » ٦	٦٩	١٦ » ١٨	٥
٦ » ٨	٤١	١٨ » ٢٠	٣
٨ » ١٠	٢٩	٢٠ » ٢٢	٤
١٠ » ١٢	٢٢	٢٦ » ٢٨	١

والمطلوب :

إيجاد الانحراف المتوسط للتوزيع التكرارى .

سبتمبر سنة ١٩٤٥

أجب على الأسئلة الآتية :

١ — الأرقام الآتية تمثل الأنسبة النموذجية لكمية أوراق البنكنوت المتداول فى مصر فى الفترة الواقعة بين سنة ١٩٢٦ وسنة ١٩٣٠

الأشهر	الأنسبة النموذجية	الأشهر	الأنسبة النموذجية
يناير	٩٥,١٠٣	يولية	٩٣,٨٧٤
فبراير	٩٩,٠٠٤	أغسطس	٩٧,٠٤٦
مارس	٩٩,٥٤٥	سبتمبر	١١٨,٧٣١
أبريل	٩٩,٠٥٤	أكتوبر	١٠٩,٨٥٦
مايو	٩٥,٤٥٠	نوفمبر	٩٧,٠٤٠
يونية	٩٣,٢٤٢	ديسمبر	٩٦,٥١٧

والمطلوب إيجاد النسب الموسمية لتداول أوراق البنكنوت فى مصر.

٢ — الجدول الآتى يبين خصوبة الزواج فى إنجلترا فى الفترة الواقعة بين سنة ١٩٠٣ وسنة ١٩١٣ :

السنة	خصوبة الزواج	السنة	خصوبة الزواج
١٩٠٣	٣,٨٨	١٩٠٩	٣,٥٠
١٩٠٤	٣,٨٣	١٩١٠	٣,٤٢
١٩٠٥	٣,٧٢	١٩١١	٣,٣١
١٩٠٦	٣,٧٠	١٩١٢	٣,٢٥
١٩٠٧	٣,٥٩	١٩١٣	٣,٢٦
١٩٠٨	٣,٦٣		

أوجد القيم الاتجاهية لهذه الظاهرة بطريقة المربعات الدنيا .

٣ ـ الجدول الآتى يبين نتيجة اختبار عدد من اللمبات الكهربائية لمعرفة أعمارها :

عمر اللمبة بالساعات	نسبة التكرار ٪	عمر اللمبة بالساعات	نسبة التكرار ٪	عمر اللمبة بالساعات	نسبة التكرار ٪	عمر اللمبة بالساعات	نسبة التكرار ٪
أقل من ١٠٠	١	٥٠٠ ـ ٦٠٠	٥,٧	١٠٠٠ ـ ١١٠٠	٧,٨	١٥٠٠ ـ ١٦٠٠	٥
١٠٠ ـ ٢٠٠	١,٥	٦٠٠ ـ ٧٠٠	٦,٦	١١٠٠ ـ ١٢٠٠	٧,٦	١٦٠٠ ـ ١٧٠٠	٤,٤
٢٠٠ ـ ٣٠٠	٣,١	٧٠٠ ـ ٨٠٠	٧,٣	١٢٠٠ ـ ١٣٠٠	٧,٣	١٧٠٠ ـ ١٨٠٠	٣,١
٣٠٠ ـ ٤٠٠	٤,٤	٨٠٠ ـ ٩٠٠	٧,٦	١٣٠٠ ـ ١٤٠٠	٦,٦	١٨٠٠ ـ ١٩٠٠	١,٥
٤٠٠ ـ ٥٠٠	٥	٩٠٠ ـ ١٠٠٠	٧,٨	١٤٠٠ ـ ١٥٠٠	٥,٧	١٩٠٠ ـ ٢٠٠٠	١

والمطلوب :

أولا : إيجاد متوسط عمر اللمبة الكهربائية بالساعات .

ثانياً : إيجاد العزم الثالث للتوزيع التكرارى .

٤ – الجدول الآتى يبين الدرجات التى حصل عليها عدد من الطلبة فى امتحان معين :

فئات الدرجات	عدد الطلبة	فئات الدرجات	عدد الطلبة
١ إلى ٥	٥	٣٦ إلى ٤٠	٧٩
٦ » ١٠	٩	٤١ » ٤٥	٥٠
١١ » ١٥	٢٨	٤٦ » ٥٠	٣٧
١٦ » ٢٠	٤٩	٥١ » ٥٥	٢١
٢١ » ٢٥	٥٨	٥٦ » ٦٠	٦
٢٦ » ٣٠	٨٢	٦١ » ٦٥	٣
٣١ » ٣٥	٨٧		

والمطلوب :

أولا : تعيين الوسيط والربيعين وتحقيق النتائج بالرسم البيانى .

ثانياً : إيجاد الدرجة المتوسطة (المتوسط الحسابى للدرجات) وعدد الطلبة الذين حصلوا على هذه الدرجة أو أقل .

اكتوبر سنة ١٩٤٦

أجب عن الأسئلة الآتية :

السؤال الأول : الجدول الآتى يبين الأرقام القياسية الشهرية للمبيعات فى عدد من المتاجر الكبيرة فى الفترة الواقعة بين سنة ١٩٢٥ وسنة ١٩٢٩ بعد استبعاد تأثير الاتجاه العام منها :

الزمن	١٩٢٥	١٩٢٦	١١٩٢٧	١٩٢٨	١٩٢٩
يناير	٨٧,٦٥	٨٦,١	٨٥,٣	٨٥,٢	٨٢,٠
فبراير	٨٣,١٥	٨٣,٠	٨٣,٣	٨٢,٤	٨٢,٩
مارس	٨٧,٩٥	٩٢,٠	٨٨,٩	٩٠,٧	٩٧,٢
أبريل	٩٥,٩٥	٩٦,٦	١٠٢,٢	٩٧,٧	٩٣,٣
مايو	٩٥,٥٥	١٠٣,٢	٩٨,٤	٩٩,٥	٩٨,٦
يونية	٩١,٩٥	٩٤,٣	٩٤,٧	٩٤,٦	٩٧,٧
يولية	٧٢,٦	٧٢,٤	٧١,٣	٧٣,٩	٧٢,٣
أغسطس	٧٣,٣	٧٧,٥	٧٩,٨	٧٤,٨	٧٤,٦
سبتمبر	٩٣,٤	٩٧,٧	٩٦,٦	١٠٣,٩	٩٤,١٥
أكتوبر	١١٧,٤	١١٢,٥	١٠٩,٩	١٠٨,١	١٠٨,١
نوفمبر	١١٧,٣	١١٦,١	١١٨,٤	١١٤,٥	١١٧,٢
ديسمبر	١٦٨,٧	١٧٣,٤	١٧٠,٩	١٧٥,٣	١٧٠,٩٥

والمطلوب :

أولا : إيجاد النسب الموسمية لحركة المبيعات

ثانياً : توضيح هذه النسب برسم بيانى مظلل يظهر بمجرد النظر أى المواسم تزداد فيها حركة المبيعات عن المتوسط ، وأيها تقل فيها عن المتوسط .

السؤال الثانى : الجدول الآتى يبين عدد سكان القطر المصرى فى سنى التعداد المختلفة :

سنة التعداد	عدد السكان
١٨٩٧	٩٧١٤٥٢٥
١٩٠٧	١١٢٨٧٣٥٩
١٩١٧	١٢٧٥٠٩١٨
١٩٢٧	١٤٢١٧٨٦٤
١٩٣٧	١٥٩٠٤٥٢٥

أوجد معادلة المستقيم الذى يصف الاتجاه العام لنمو السكان ، ومنها استنتج قيمة تقريبية لما سيكون عليه عدد سكان القطر المصرى فى التعداد المقبل (سنة ١٩٤٧) . إوجد كذلك قيمة تقريبية لعدد السكان فى سنة ١٩٧٠ .

نوفمبر سنة ١٩٤٦

أجب عن الأسئلة الآتية :

السؤال الأول : اذكر مزايا وعيوب صيغة الرقم القياسى الآتية التى اقترحها الأستاذ مارشال لقياس متوسط التغير النسبى فى مستوى الأسعار :

$$\frac{\Sigma\, ع_{ن}\ (ك_{.} + ك_{ن})}{\Sigma\, ع_{.}\ (ك_{.} + ك_{ن})}$$

(حيث $ع_{.}$ ، $ك_{.}$ تمثلان على التوالى السعر والكمية فى سنة الأساس، $ع_{ن}$ ، $ك_{ن}$ تمثلان السعر والكمية فى السنة المقارنة .

وضح كيف يمكن تعديل هذه الصيغة لجعلها أقرب إلى الدقة .

السؤال الثانى : أوجد الوسط الحسابى والانحراف المعيارى لمتوالية الأعداد الطبيعية من ١ إلى ن .

السؤال الثالث : الجدول الآتى يبين توزيع عمال الولايات المتحدة على الحرف المختلفة فى سنتى ١٨٧٠ ، ١٩٣٠

الحرفة (أو الصناعة)	عدد العمال بالآلاف	
	١٨٧٠	١٩٣٠
الزراعة والحرف الملحقة بها	٦٤٢٨	١٠٢٤٢
المناجم	١٧٢	٩٨٣
صناعة الآلات والبضائع	٢٦٧٤	١٣٧٩٠
التجارة والنقل	١١٠٤	٩٩٦٣
المهن الكتابية	٢٠٦	٣٩٣٥
الخدمات المنزلية والشخصية	١١٦٨	٥٤٤٨
المهن الفنية (الطب والمحاماه إلخ)	٣٣٨	٣١١٠
خدمات عامة متنوعة	٧٣	٦٩٢

استخدم إحدى طرق العرض البيانى لإيضاح تغير الأهمية النسبية للحرف المختلفة بين السنتين المذكورتين .

السؤال الرابع :

(ا) اكتب مقارنة موجزة بين مقاييس التوسط المختلفة .

(ب) أثبت أن الوسط الهندسى لأى كميتين موجبتين هو أيضاً الوسط الهندسى للوسط الحسابى والوسط التوافقى لهاتين الكميتين .

يونية سنة ١٩٤٧

أجب عن الأسئلة الآتية :

السؤال الأول : على فرض أن سكان دولة ما يزيدون بنسبة ثابتة ، إثبت أن عدد السكان فى منتصف الفترة الواقعة بين سنتى تعداد متتاليتين يساوى الوسط الهندسى لعدد السكان فى هاتين السنتين ، طبق النتيجة السابقة لإيجاد قيمة تقريبية لعدد سكان مصر فى سنة ١٩٤٢ مع العلم بأن عدد السكان حسب تعداد سنة ١٩٣٧ هو ١٥٩٠٥٠٠٠ وعدد السكان حسب تعداد سنة ١٩٤٧ هو ١٩٢٥٠٠٠٠ .

السؤال الثانى : الجدول الآتى يبين التوزيع التكرارى المتجمع لبعض فئات الدخل فى إنجلترا فى سنة ١٩٢٨ ، والمطلوب رصد هذا التوزيع على قياس لوغاريتمى مزدوج لمعرفة مدى مطابقته لقانون باريتو (Pareto Income Law)

الدخل بالجنيهات (س)	عدد السكان الذين يحصلون على دخل قدره (س) أو أكثر
٥٠٠٠	٢٧٧١١
٦٠٠٠	٢٠٩٧٥
٧٠٠٠	١٦٥٤٤
٨٠٠٠	١٣٣١٧
١٠٠٠٠	٩١٦٣
١٥٠٠٠	٤٥٩٥
٢٠٠٠٠	٢٧٨١
٢٥٠٠٠	١٨٥١
٣٠٠٠٠	١٣٢٤
٤٠٠٠٠	٧٥٣
٥٠٠٠٠	٥٠٠

ضع لوغاريتمات الأعداد التى تستخرجها فى شكل جدولى قبل رصدها على ورق الرسم البيانى

السؤال الثالث : ما معنى تحليل السلسلة الزمنية ؟ وما أغراض هذا التحليل ؟

السؤال الرابع : إذا كانت $ح_1$ ، $ح_2$ ، $ح_3$ ..... $ح_ن$ تمثل انحرافات «ن» من القيم عن متوسطها الحسابى $\overline{س}$ ، وكانت $د_1$ ، $د_2$ ، $د_3$ ..... $د_ن$ تمثل انحرافات نفس القيم عن أى وسط فرضى آخر ، أثبت أن :

أولا : $\frac{\Sigma د}{ن} = \overline{س} - س_د$ (حيث $س_د$ تمثل الوسط الفرضى)

ثانياً: $\frac{\text{ح}_{د}^{2}}{\text{ن}} = \frac{\text{ح}^{2}}{\text{ن}} + \left(\frac{\text{ح}_{د}}{\text{ن}}\right)^{2}$

يونيه سنة ١٩٤٧

السؤال الأول : الجدول الآتى يبين أوزان مجموعة من طلبة إحدى الجامعات مقدرة بالأرطال .

فئات الوزن	التكرار	فئات الوزن	التكرار
٨٩,٥ — ٩٩,٥	١	١٥٩,٥ — ١٦٩,٥	٦٢
٩٩,٥ — ١٠٩,٥	١٩	١٦٩,٥ — ١٧٩,٥	٢٨
١٠٩,٥ — ١١٩,٥	٨٠	١٧٩,٥ — ١٨٩,٥	١١
١١٩,٥ — ١٢٩,٥	١٧٩	١٨٩,٥ — ١٩٩,٥	٤
١٢٩,٥ — ١٣٩,٥	٢٦٠	١٩٩,٥ — ٢٠٩,٥	٤
١٣٩,٥ — ١٤٩,٥	٢١٠	٢٠٩,٥ — ٢١٩,٥	٢
١٤٩,٥ — ١٥٩,٥	١٤٠		١٠٠٠

والمطلوب إيجاد الوسط الحسابى والانحراف المعيارى للأوزان مع تطبيق اختبار شارليه للمراجعة

السؤال الثانى : الجدول الآتى يبين عدد سكان كندا فى سنى التعداد المختلفة

السنة	عدد السكان	السنة	عدد السكان
١٨٨١	٤,٢٦٦,٠٤١	١٩١١	٧,١٦٩,٩٦٠
١٨٩١	٤,٧٧٠,١٢٣	١٩٢١	٨,٧٦٧,٢٠٦
١٩٠١	٥,٣٢٢,٢٣٨	١٩٣١	١٠,٣٧٦,٧٨٦

إوجد نسب زيادة السكان بين سنى التعداد المختلفة ، وبإيجاد الوسط الهندسى لهذه النسب ، إوجد معدل الزيادة العشر سنوية فى المتوسط بين سنة ١٨٨١ ، ١٩٣١ (استخرج نسب الزيادة فى السكان لأقرب ثلاثة أرقام عشرية)

السؤال الثالث : وفق مستقيم بطريقة المربعات الدنيا للظاهرة المبينة بالجدول الآتى : ومنه استنتج قيمة الظاهرة فى سنتى ١٩٣٥ ، ١٩٤٦

السنة	القيمة المشاهدة للظاهرة
١٩٣٦	٣٦٦٤,٥
١٩٣٧	٣٩١٥,٩
١٩٣٨	٢٠٠١
١٩٣٩	٢٨٦٦,٨
١٩٤٠	٣٦٩٣,١
١٩٤١	٣٥٧٤,٣

السؤال الرابع : أوجد قيمة المنوال للتوزيع التكرارى المبين بالجدول الآتى وحقق النتيجة بالرسم البيانى

فئات المتغير	التكرار
١٠ — ١٢	٣
١٢ — ١٤	١٥
١٤ — ١٦	٢٠
١٦ — ١٨	١٠
١٨ — ٢٠	٢

سبتمبر سنة ١٩٤٧

أجب عن الأسئلة الآتية :

السؤال الأول : الجدول الآتى يبين إنتاج القطن مقدراً بالقناطير فى إحدى مديريات القطر المصرى .

السنة	اشمونى	جيزه ٧	سكلاريدس	الجملة
١٩٣٠	٦٩٧١٣	١٩٠٩٩	٥٥٤٩٣	١٤٤٣٠٥
١٩٣١	٨٥٦٧٩	١٣٠٠٤	٨٠٦٧٦	١٨٩٣٥٩
١٩٣٢	١٠٣٢٨٨	٣٨٨٩٤	١٠١٦٧٨	٢٤٣٨٦٠
١٩٣٣	١٤٦٦١٩	٣٢٩٠٤	١٤٥٢٥٦	٣٢٤٧٧٩

استخدم إحدى طرق العرض البيانى لوضع البيانات المبينة فى الجدول السابق فى صورة سهلة يمكن أن يفهم منها بمجرد النظر تغير الأهمية النسبية لكل نوع من أنواع القطن فى هذه المديرية .

السؤال الثانى : إذا كانت ح ، هـ ، ق تمثل على التوالى الوسط الحسابى والهندسى والتوافقى لأى كميتين موجبتين أثبت أن : أولا — ح > هـ > ق

ثانياً — هـ = $\sqrt{\text{ح ق}}$

السؤال الثالث : وضح بمثال عددى طريقة وضع البيانات الإحصائية الأولية فى شكل توزيع تكرارى ثم اذكر الفائدة العملية لوضع البيانات فى هذه الصورة .

السؤال الرابع : الجدول الآتى يبين عدد سكان الولايات المتحدة حسب الإحصاءات الرسمية بين سنتى ١٧٩٠ و ١٩٣٠

السنة	عدد السكان	السنة	عدد السكان	السنة	عدد السكان
١٧٩٠	٣٩٢٩٢١٤	١٨٤٠	١٧٠٦٩٤٥٣	١٨٩٠	٦٢٩٤٧٧١٤
١٨٠٠	٥٣٠٨٤٨٣	١٨٥٠	٢٣١٩١٨٧٦	١٩٠٠	٧٥٩٩٤٥٧٥
١٨١٠	٧٢٩٣٨٨١	١٨٦٠	٣١٤٤٣٣٢١	١٩١٠	٩١٩٧٢٢٦٦
١٨٢٠	٩٦٣٨٤٥٣	١٨٧٠	٣٨٥٥٨٣٧١	١٩٢٠	١٠٥٧١٠٦٢٠
١٨٣٠	١٢٨٦٦٠٢٠	١٨٨٠	٥٠١٥٥٧٨٣	١٩٣٠	١٢٢٧٧٥٠٤٦

ارصد البيانات السابقة على قياس نصف لوغاريتمى واذكر بإيجاز ما تستنتجه من الرسم .

سبتمبر سنة ١٩٤٧

السؤال الأول : الجدول الآتى يبين النقص فى قيمة أحد الأصول خلال خمس سنوات .

السنة	قيمة الأصل بالجنيهات
١٩٣٠	٧٢٠
١٩٣١	٦٠٩
١٩٣٢	٥٠٠
١٩٣٣	٣٧٥
١٩٣٤	٢٢٧
١٩٣٥	١٠٧

والمطلوب إيجاد المعدل المتوسط للنقص فى قيمة الأصل خلال هذه المدة .

(عن طريق حساب الوسط الهندسى لنسب التناقص)

السؤال الثانى : الجدول الآتى يبين الأرقام القياسية لإنتاج السيارات فى الولايات المتحدة فى الفترة الواقعة بين سنة ١٩١٩ وسنة ١٩٣٠ .

السنة	الرقم القياسى
١٩١٩	٥٠
١٩٢٠	٥٨
١٩٢١	٤١
١٩٢٢	٦٦
١٩٢٣	١٠٢
١٩٢٤	٩١
١٩٢٥	١٠٧
١٩٢٦	١٠٨
١٩٢٧	٨٦
١٩٢٨	١١٠
١٩٢٩	١٣٥
١٩٣٠	٨٥

إوجد القيم الاتجاهية للظاهرة عن طريق حساب متوسط متحرك فترته ٣ سنوات وآخر فترته ٤ سنوات ، ثم ارصد هذه المتوسطات مع القيم الأصلية للظاهرة على ورق الرسم البيانى .

السؤال الثالث : إوجد الوسط الحسابى والانحراف المعيارى للتوزيع التكرارى المبين بالجدول الآتى ، وحقق النتائج التى تحصل عليها بتطبيق مراجعة شارلييه .

مراكز الفئات	التكرار
١٧٣,٥	٣
١٧٧,٥	٩
١٨١,٥	٢٩
١٨٥,٥	٧٦
١٨٩,٥	١٠٤
١٩٣,٥	١١٠
١٩٧,٥	٨٨
٢٠١,٥	٣٠
٢٠٥,٥	٦
٢٠٩,٥	٤
٢١٣,٥	٢
٢١٧,٥	١

السؤال الرابع : إوجد قيم الوسيط والربيعين للتوزيع التكرارى المبين بالسؤال الثالث ، وحقق النتائج بالرسم البيانى .

يونية سنة ١٩٤٨

أجب عن الأسئلة الآتية :

السؤال الأول : تتبع بإيجاز المراحل التى تمر بها البيانات الإحصائية الخاصة بتعداد سكان القطر المصرى من وقت جمعها من المنازل حتى يتم تبويبها . وضح كذلك الأسس التى تنبنى عليها عملية التبويب .

السؤال الثانى : وضح معنى التشتت مستعيناً ببعض الأمثلة . فى أى الحالات يجب استعمال مقاييس التشتت النسبية بدلا من المقاييس المطلقة ؟

السؤال الثالث : اشرح بإيجاز طريقة المتوسطات البسيطة فى إيجاد النسب الموسمية . ثم اذكر الفوائد العملية للحصول على هذه النسب .

السؤال الرابع : قسم ورقة الرسم البيانى المعطاه لك تقسيما لوغاريتمياً ، ثم ارصد عليها قيم الظاهرة المبينة بالجدول الآتى :

واردات المملكة المتحدة : ١٨٢٠ — ١٨٩٧

السنة الوسطى	الواردات بملايين الجنيهات (متوسط ثلاث سنوات)	السنة الوسطى	الواردات بملايين الجنيهات (متوسط ثلاث سنوات)
١٨٢١	٣١	١٨٦٠	٢٠٢
١٨٢٤	٣٩	١٨٦٣	٢٥٠
١٨٢٧	٤٣	١٨٦٦	٢٨٠
١٨٣٠	٤٧	١٨٦٩	٢٩٨
١٨٣٣	٤٧	١٨٧٢	٣٥٢
١٨٣٦	٥٤	١٨٧٥	٣٧٣
١٨٣٩	٦٣	١٨٧٨	٣٧٥
١٨٤٢	٦٥	١٨٨١	٤٠٧
١٨٤٥	٧٧	١٨٨٤	٣٩٦
١٨٤٨	٩١	١٨٨٧	٣٦٧
١٨٥١	١٠٧	١٨٩٠	٤٢٨
١٨٥٤	١٤٠	١٨٩٣	٤١٠
١٨٥٧	١٧٢	١٨٩٦	٤٣٧

يونية سنة ١٩٤٨

أجب عن الأسئلة الآتية :

السؤال الأول : الجدول الآتى يبين إنتاج الحديد الخام فى الولايات المتحدة فى الفترة الواقعة بين سنة ١٩٠٣ وسنة ١٩١٦ .

السنة	الإنتاج الشهرى بآلاف الأطنان	السنة	الإنتاج الشهرى بآلاف الأطنان
١٩٠٣	١٤٥٢	١٩١٠	٢٢٣٧
١٩٠٤	١٣٤٤	١٩١١	١٩٤٤
١٩٠٥	١٨٨٢	١٩١٢	٢٤٤٨
١٩٠٦	٢٠٦٦	١٩١٣	٢٥٦٠
١٩٠٧	٢١٠٩	١٩١٤	١٩٢١
١٩٠٨	١٣٠٢	١٩١٥	٢٤٧٢
١٩٠٩	٢١١٦	١٩١٦	٣٢٥٢

إوجد معادلة المستقيم الذى يصف الاتجاه العام لهذه الظاهرة بطريقة المربعات الدنيا ، ثم ارصد القيم المشاهدة للظاهرة على ورق الرسم البيانى ومعها على نفس الرسم المستقيم الذى يمثل الاتجاه العام .

السؤال الثانى : إوجد الوسط الهندسى للتوزيع التكرارى المبين بالجدول الآتى :

فئات المتغير	التكرار	فئات المتغير	التكرار
٣٥ — ٤٥	١	٨٥ — ٩٥	٢٢
٤٥ — ٥٥	٦	٩٥ — ١٠٥	٢٧
٥٥ — ٦٥	٨	١٠٥ — ١١٥	١٨
٦٥ — ٧٥	٥	١١٥ — ١٢٥	١٤
٧٥ — ٨٥	١٤	المجموع	١١٥

السؤال الثالث : الجدول الآتى يبين أسعار صرف الجنيه الإسترلينى بالفرنكات فى الفترة الواقعة بين سنة ١٨٨٢ وسنة ١٩١٣ .

سعر الصرف	التكرار	سعر الصرف	التكرار
٢٥,٠٧ — ٢٥,٠٩	١	٢٥,٢٧ — ٢٥,٢٩	٢٦
٢٥,٠٩ — ٢٥,١١	٤	٢٥,٢٩ — ٢٥,٣١	٢١
٢٥,١١ — ٢٥,١٣	١٤	٢٥,٣١ — ٢٥,٣٣	٢٠
٢٥,١٣ — ٢٥,١٥	٢٠	٢٥,٣٣ — ٢٥,٣٥	٤
٢٥,١٥ — ٢٥,١٧	٤٥	٢٥,٣٥ — ٢٥,٣٧	٦
٢٥,١٧ — ٢٥,١٩	٦٠	٢٥,٣٧ — ٢٥,٣٩	٢
٢٥,١٩ — ٢٥,٢١	٤٠	٢٥,٣٩ — ٢٥,٤١	٢
٢٥,٢١ — ٢٥,٢٣	٤٣	٢٥,٤١ — ٢٥,٤٣	٢
٢٥,٢٣ — ٢٥,٢٥	٤٢		
٢٥,٢٥ — ٢٥,٢٧	٣٢	المجموع	٣٨٤

والمطلوب إيجاد الوسط الحسابى والانحراف المعيارى لهذا التوزيع التكرارى بالطريقة المختصرة على أساس اختيار ٢٥,٢٤ كوسط فرضى لأسعار الصرف .

يونيو سنة ١٩٤٩

أجب عن الأسئلة الآتية :

السؤال الأول : اذكر باختصار الفوائد العملية التى يحتمل أن تجنيها الإدارة العامة لإحدى شركات النقل بالمدن من إنشاء قسم فنى للإحصاء .

السؤال الثانى : فى أول يناير سنة ١٩٢٩ أودع مبلغ ١٢١٠٠ جنيه بفائدة مركبة فصارت جملته فى أول يناير سنة ١٩٤٩ ١٦٩٠٠ جنيه والمطلوب :

أولا : استخدام الوسط الهندسى لإيجاد جملة المبلغ فى أول يناير من سنة ١٩٣٤ ، سنة ١٩٣٩ ، وسنة ١٩٤٤

ثانياً : إيجاد سعر الفائدة المركبة خلال مدة الإيداع .

السؤال الثالث : إشرح أهمية الترجيح بالأوزان عند حساب المتوسطات موضحاً إجابتك ببعض الأمثلة العددية .

السؤال الرابع : وضح الأسس التى ينبنى عليها اختيار فئات التوزيعات التكرارية ، وبتطبيق هذه الأسس ضع البيانات الآتية فى شكل توزيع تكرارى ثم اعرض هذا التوزيع بيانياً :

عدد وحدات التيار الكهربائى المستهلكة فى ٧٥ مسكن خاص فى شهر معين

٨	٣٨	٥٤	٦١	٧١	٧٦	٨٣	٨٩	٩٤	١١٨	١٣٦
٩	٤٠	٥٦	٦٢	٧٢	٧٧	٨٣	٩٠	٩٥	١٢١	١٤٤
١٠	٤١	٥٧	٦٤	٧٣	٧٨	٨٤	٩٠	٩٦	١٢٥	١٤٨
١٩	٥٠	٥٨	٦٦	٧٤	٧٩	٨٤	٩١	٩٨	١٢٦	١٥٧
٢٨	٥١	٥٨	٦٧	٧٥	٨٠	٨٦	٩٢	١٠٥	١٢٨	١٥٨
٣٦	٥٢	٥٩	٦٨	٧٥	٨١	٨٧	٩٣	١١٤	١٣١	
٣٧	٥٣	٦٠	٧٠	٧٥	٨٢	٨٨	٩٤	١١٥	١٣٥	

يونيو سنة ١٩٤٩

أجب عن الأسئلة الآتية :

السؤال الأول : الجدول الآتى يبين الأرباح الصافية التى حققها أحد المصارف فى الفترة الواقعة بين سنة ١٩٢٠ وسنة ١٩٤١ والمطلوب توفيق مستقيم بطريقة المربعات الدنيا لتعيين القيم الاتجاهية لهذه الظاهرة :

السنة	الأرباح بآلاف الجنيهات	السنة	الأرباح بآلاف الجنيهات
١٩٢٠	٨٠٠	١٩٣١	١٢٦٦
١٩٢١	٥١٠	١٩٣٢	١٢٢٥
١٩٢٢	٦١٨	١٩٣٣	١١١٦
١٩٢٣	٨٠٥	١٩٣٤	١٢١٨
١٩٢٤	٨٥١	١٩٣٥	١٤٤٦
١٩٢٥	٩٦٧	١٩٣٦	١٧٧٥
١٩٢٦	١١٣٣	١٩٣٧	٢٠١٤
١٩٢٧	١٢٤٨	١٩٣٨	١٥٨٤
١٩٢٨	١٢٨٤	١٩٣٩	١٩٨٢
١٩٢٩	١٤٧٣	١٩٤٠	٢١٧١
١٩٣٠	١٦٤٤	١٩٤١	٢٦٦١

السؤال الثانى : الجدول الآتى يوضح العلاقة بين رأس المال والربح فى إحدى المؤسسات التجارية خلال عشر سنوات :

السنة	رأس المال بمئات الجنيهات	الربح بمئات الجنيهات	السنة	رأس المال بمئات الجنيهات	الربح بمئات الجنيهات
١٩٣٢	١٠	٣	١٩٣٧	٦٠	١٥
١٩٣٣	٢٠	٤	١٩٣٨	٧٠	١٥
١٩٣٤	٣٠	٨	١٩٣٩	٨٠	٢٥
١٩٣٥	٤٠	٨	١٩٤٠	٩٠	٢٢
١٩٣٦	٥٠	١٠	١٩٤١	١٠٠	٣٠

والمطلوب إيجاد معامل الارتباط بين رأس المال والربح .

السؤال الثالث : إوجد الانحراف المعيارى للتوزيع التكرارى المبين بالجدول الآتى :

مراكز الفئات	التكرار
٩,٥	٤
١١,٥	١١
١٣,٥	٢٠
١٥,٥	٢٩
١٧,٥	٣٠
١٩,٥	٢٥
٢١,٥	٢١
٢٣,٥	٩
٢٥,٥	٦
٢٧,٥	٣
٢٩,٥	١
٣١,٥	١

(اعتبر ١٧,٥ وسطاً فرضياً)

سبتمبر سنة ١٩٤٩

السؤال الأول : إوجد معادلة المستقيم الذى يصف الاتجاه العام للظاهرة المبينة بالجدول الآتى ثم ارصد القيم المشاهدة والقيم الاتجاهية على ورق الرسم البيانى :

السنة	قيمة الظاهرة
١٩٣١	٢٦٤
١٩٣٢	١٠٦
١٩٣٣	٩٥
١٩٣٤	١٥٦
١٩٣٥	١٧٤
١٩٣٦	١٨٨
١٩٣٧	٢٩٦
١٩٣٨	٢٣٩
١٩٣٩	١٩٨
١٩٤٠	٢٣٤
١٩٤١	٣١٧

السؤال الثانى : إوجد الوسط الهندسى للتوزيع التكرارى المبين بالجدول الآتى :ـ

فئات المتغير	التكرار	فئات المتغير	التكرار
١٥ ـ ٢٠	٨٨	٤٥ ـ ٥٠	١٠
٢٠ ـ ٢٥	٢٢٩	٥٠ ـ ٥٥	١٥
٢٥ ـ ٣٠	٣٧٢	٥٥ ـ ٦٠	١
٣٠ ـ ٣٥	١٣٩	٦٠ ـ ٦٥	٢
٣٥ ـ ٤٠	٥٤	٦٥ ـ ٧٠	١
٤٠ ـ ٤٥	١٤	٧٠ ـ ٧٥	١

السؤال الثالث : الجدول الآتى يبين الأرقام القياسية ربع السنوية لمستوى العمل فى كندا على أساس كون سنة ١٩٢٦ سنة أساس .

السنة	الربع الأول	الربع الثانى	الربع الثالث	الربع الرابع
١٩٣٦	٩٨	١٠٢	١٠٧	١١٠
١٩٣٧	١٠٣	١١٤	١٢٣	١٢٢
١٩٣٨	١٠٨	١١٢	١١٥	١١٤
١٩٣٩	١٠٦	١١٣	١٢٠	١٢٣

والمطلوب استخدام هذه البيانات لإيجاد النسب الموسمية لمستوى العمل ثم توضيح هذه النسب بيانياً .

مايو سنة ١٩٥٠

السؤال الأول

الجدول الآتى يبين كمية البنكنوت المصدر فى مصر فى آخر كل شهر من شهور السنوات الخمس الأخيرة مقدراً بملايين الجنيهات

الشهر	سنة ١٩٤٥	سنة ١٩٤٦	سنة ١٩٤٧	سنة ١٩٤٨	سنة ١٩٤٩
يناير	١٢٢	١٤٦	١٤١	١٤١	١٥٩
فبراير	١٢٣	١٤٥	١٤١	١٤٠	١٥٨
مارس	١٢٥	١٤٤	١٤٠	١٤٠	١٥٨
أبريل	١٢٦	١٤٣	١٤٠	١٣٩	١٥٨
مايو	١٢٦,٥	١٤٢	١٣٩	١٤٤	١٥٥
يونية	١٢٩	١٣٩	١٣٥	١٤٢	١٥١
يولية	١٢٩,٥	١٣٩	١٣١	١٤٠	١٥٠
أغسطس	١٣٠	١٣٩	١٣٠	١٣٧	١٤٦
سبتمبر	١٣٣	١٣٩	١٤٥	١٤٤	١٥١
أكتوبر	١٤٥	١٤٧	١٤٤	١٥٥	١٦٦
نوفمبر	١٤٧,٥	١٤٨	١٤٥	١٦١	١٧٤
ديسمبر	١٤٨	١٤٥	١٤٥	١٦٤	١٧٤

والمطلوب إيجاد النسب الموسمية لكمية البنكنوت المصدر ثم تمثيلها بيانياً .

السؤال الثانى

الجدول الآتى يبين التوزيع التكرارى للمزارع الأمريكية حسب المساحة كما يبين القيمة الإجمالية للمزارع الموجودة بكل فئة تكرارية مقدرة بملايين الدولارات .

مساحة المزرعة بالفدان	عدد المزارع	النسبة المئوية لعدد المزارع	القيمة الإجمالية بملايين الدولارات
أقل من ٢٠	٧٩٦,٥٣٥	١٢,٤	٢,٤٥٣
٢٠ — ٥٠	١,٥٠٣,٧٣٢	٢٣,٣	٥,٨٦٤
٥٠ — ١٠٠	١,٤٧٤,٧٤٥	٢٢,٩	١١,١٨٣
١٠٠ — ١٧٥	١,٤٤٩,٦٣٠	٢٢,٥	٢٠,٩٠٢
١٧٥ — ٥٠٠	١,٠٠٦,٤٧٧	١٥,٦	٢٦,٣٩٠
٥٠٠ — ١٠٠٠	١٤٩,٨١٩	٢,٣	٥,٥٣٤
١٠٠٠ فأكثر	٦٧,٤٠٥	١,٠	٥,٥٩٨
المجمـــوع	٦,٤٤٨,٣٤٣	١٠٠,٠	٧٧,٩٢٤

ارسم منحنى لورنز من واقع البيانات السابقة مع إيضاح خطوات العمل بالتفصيل .

السؤال الثالث

إعرض الفئات الخمس الأولى للتوزيع التكرارى لحجم المزرعة المبين بالسؤال الثانى بيانياً بطريقة الهستوجرام .

السؤال الرابع

استخدم متوسط مناسب لإيجاد المساحة المتوسطة للمزرعة الأمريكية من واقع البيانات الموضحة بالسؤال الثانى . وضح سبب اختيارك للمتوسط الذى تستخدمه .

سبتمبر سنة ١٩٥٠

أجب عن الأسئلة الآتية :

السؤال الأول

الجدول الآتى يبين توزيع الدخول فى الولايات المتحدة فى سنة ١٩١٩ :

الدخل بالدولار (س)	عدد السكان الذين يحصلون على دخل قدره (س) أو أكثر بالآلاف
٧٥٠	٣٥٥٤١
١٢٥٠	٢٣٠١٠
١٧٥٠	١٠٥١٢
٢٥٠٠	٥٢٩٠
٤٠٠٠	٢٢٢٥
٧٥٠٠	٨٤٢
١٧٥٠٠	٢٥٤
٣٧٥٠٠	٦.٢
٧٥٠٠٠	٢١
١٥٠٠٠٠	٧
٢٥٠٠٠٠	٢

والمطلوب رصد هذا التوزيع على قياس لوغاريتمى مزدوج لمعرفة مدى مطابقته لقانون باريتو فى توزيع الدخل .

السؤال الثانى

إوجد معامل الارتباط بين المتغيرين س و ص من واقع البيانات الموضحة بالجدول الآتى :

قيم س / قيم ص	٣	٥	٧	٩	١١	المجموع
٣٥				١	١	٢
٣٠			٢	١	١	٤
٢٥		١	٣	٢		٦
٢٠		٤	٢			٦
١٥	١	١				٢
المجموع	١	٦	٧	٤	٢	٢٠

السؤال الثالث

الجدول الآتى يبين العمر المتوسط للطالب وقت التحاقه بجامعة هارفارد بأمريكا من سنة ١٨٥٦ إلى سنة ١٩٠٠ :

الفترة الزمنيـــة	متوسط عمر الطالب عند التحاقه بالجامعة
١٨٥٦ — ١٨٦٠	١٧,٩
١٨٦١ — ١٨٦٥	١٨,٤
١٨٦٦ — ١٨٧٠	١٨,٤
١٨٧١ — ١٨٧٥	١٨,٦
١٨٧٦ — ١٨٨٠	١٨,٨
١٨٨١ — ١٨٨٥	١٩,٠
١٨٨٦ — ١٨٩٠	١٩,٣
١٨٩١ — ١٨٩٥	١٩,١
١٨٩٦ — ١٩٠٠	١٩,٠

إوجد معادلة المستقيم الذى يصف الاتجاه العام لهذه الظاهرة بطريقة المربعات الدنيا — ثم ارصد قيم الظاهرة على ورق الرسم البيانى ومعها على نفس الرسم المستقيم الذى يمثل الاتجاه العام .

السؤال الرابع

اشرح بالتفصيل طريقة وضع البيانات الإحصائية الأولية فى شكل توزيعات تكرارية .

مايو سنة ١٩٥١

السؤال الأول

الجدول الآتى يبين الإنتاج السنوى للأسمنت فى مصر مقدراً بآلاف الأطنان بين سنة ١٩٣٧ وسنة ١٩٥٠ :

السنة	إنتاج الاسمنت بآلاف الأطنان	السنة	إنتاج الاسمنت بآلاف الأطنان
١٩٣٧	٣٢٠	١٩٤٤	٤٢٠
١٩٣٨	٣٧٠	١٩٤٥	٤٣٠
١٩٣٩	٣٥٠	١٩٤٦	٥٩٠
١٩٤٠	٣٦٠	١٩٤٧	٦٥٠
١٩٤١	٤٢٠	١٩٤٨	٧٧٠
١٩٤٢	٣٨٠	١٩٤٩	٧٨٠
١٩٤٣	٣٢٠	١٩٥٠	٨٣٠

والمطلوب إيجاد معادلة المستقيم الذى يصف الاتجاه العام لهذه الظاهرة بطريقة المربعات الدنيا ثم رصد الظاهرة على ورق الرسم البيانى ومعها على نفس الرسم خط الاتجاه العام .

السؤال الثانى

الجدول الآتى يبين سكان المملكة المصرية وسكان مديرية أسوان فى سنى التعداد المختلفة :

سنة التعداد	سكان المملكة المصرية	سكان مديرية أسوان
١٨٨٢	٦٨٠٤٠٢١	١٤١٨٩٤
١٨٩٧	٩٧١٤٥٢٥	٢٠٥٩٦٠
١٩٠٧	١١٢٨٧٣٥٩	٢٣٢٨١٣
١٩١٧	١٢٧٥٠٩١٨	٢٥٣٣٤٠
١٩٢٧	١٤٢١٧٨٦٤	٢٦٧٣٥٧
١٩٣٧	١٥٩٣٢٦٩٤	٣٠٥٠٩٦
١٩٤٧	١٩٠٩٢٣٨٩	٢٩٠٨٤٢

والمطلوب رصد هاتين الظاهرتين على مقياس نصف لوغاريتمى (فى رسم واحد) بطريقة تسهل المقارنة بين معدلات النمو فى الحالتين . علق بإيجاز على ما يظهره الرسم من اتجاهات .

تنبيه : ضع لوغاريتمات الأعداد فى شكل جدولى قبل رصدها على الرسم .

السؤال الثالث

إوجد معامل الارتباط بين المتغيرين « س » و « ص » بطريقة الأقطار ذات الفروق المتساوية من واقع القيم المبينة بالجدول الآتى :

قيم س / قيم ص	٤٠ – ٤٤	٤٤ – ٤٨	٤٨ – ٥٢	٥٢ – ٥٦	٥٦ – ٦٠	٦٠ – ٦٤	المجموع
٣٥ – ٥٥	١	١٠	١٥				٢٦
٥٥ – ٧٥			٦	٢٢	٣		٣١
٧٥ – ٩٥				٢	٨	٢	١٢
٩٥ – ١١٥					١	٢	٣
١١٥ – ١٣٥				١			١
١٣٥ – ١٥٥					١	١	٢
١٥٥ – ١٧٥						٥	٥
المجمـــوع	١	١٠	٢١	٢٥	١٣	١٠	٨٠

سبتمبر سنة ١٩٥١

أجب عن الأسئلة الآتية :

السؤال الأول

الجدول الآتي يبين كمية القطن الصادر شهرياً مقدرة بآلاف القناطير بين سنة ١٩٤٤ و سنة ١٩٤٨ :

الشهور	٤٤ – ٤٥	٤٥ – ٤٦	٤٦ – ٤٧	٤٧ – ٤٨
سبتمبر	١٠	١٤٥	٤٩٨	٤٥٤
أكتوبر	١٩٣	٢٤٨	٤٤٧	١٨٤
نوفمبر	١٨٦	٣٢٢	٤٨٣	٦٣٩
ديسمبر	٥٧٢	١٤١	٧١٨	١٣٨٣
يناير	١١٢	١٨٩	٦٦٦	٦٣٢
فبراير	٤٢٦	٢٩٢	٤١٩	٦٥٤
مارس	٢٢٢	٣٢٢	٦٧٨	٦٨٠
أبريل	٥٠٧	٤٥٨	٧٥٠	٥٠٠
مايو	٣٢٤	٥٣٠	٧٩٧	٨٨٤
يونية	٦١٤	٦٢٠	٥١٤	٧٤٨
يولية	٦٠٤	٤٢٧	٤٧٤	٣٧٩
أغسطس	٤٥٩	٧٢٧	٥١٢	٦٦٢

والمطلوب إيجاد النسب الموسمية الشهرية لحركة تصدير القطن ثم تمثيلها بيانياً .

السؤال الثانى

الأرقام الآتية تمثل الدرجات التى حصل عليها خمسة عشر طالباً فى اختبارين مختلفين :

الاختبار الأول	٩٣	٨٩	٧٢	٦٩	٦٤	٥٧	٥٣	٥٢	٤٩	٤٦	٤٣	٣٨	٣٧	٣٢	٢٩
الاختبار الثانى	٢٢	١٩	٦	١٨	٢٠	١٦	١١	٩	١٥	١٢	١٠	٧	١٣	١٢	٨

والمطلوب إيجاد معامل سبيرمان للارتباط بين مرتبات الطلبة فى هذين الاختبارين .

السؤال الثالث

إوجد معامل الاختلاف ومعامل الالتواء للتوزيع التكرارى المبين بالجدول الآتى :

فئات المتغير	التكرار
١٦٠ — ١٨٠	٢٠٠
١٨٠ — ٢٠٠	٥٥٠
٢٠٠ — ٢٢٠	٦٥٠
٢٢٠ — ٢٤٠	٣٢٠
٢٤٠ — ٢٦٠	٨٠
٢٦٠ — ٢٨٠	٥٠
٢٨٠ — ٣٠٠	٢٠
٣٠٠ — ٣٢٠	٦٠
٣٢٠ — ٣٤٠	٣٠
٣٤٠ — ٣٦٠	٤٠

ثانياً ـ امتحانات البكالوريوس

(يونيه سنة ١٩٤٦)

السؤال الأول

الجدول الآتى يبين العلاقة بين النسبة المئوية للتوظف بين العمال المشتركين فى النقابات (س) وبين معدل الزواج لكل ألف من السكان (ص) فى إنجلترا فى الفترة الواقعة بين سنة ١٨٦٠ وسنة ١٩١٤ . أوجد معامل الارتباط بين هاتين الظاهرتين وعلق بإيجاز على النتيجة :

قيم س / قيم ص	٨٩ـ٩٠	٩٠ـ٩١	٩١ـ٩٢	٩٢ـ٩٣	٩٣ـ٩٤	٩٤ـ٩٥	٩٥ـ٩٦	٩٦ـ٩٧	٩٧ـ٩٨	٩٨ـ٩٩	٩٩ـ١٠٠	المجموع
١٤ ـ ١٤,٥	١	١		١			١					٤
١٤,٥ ـ ١٥			٢	١		١						٤
١٥ ـ ١٥,٥			١	٢	٣	٣		٢	١			١٢
١٥,٥ ـ ١٦						١	٣	٧	٤			١٥
١٦ ـ ١٦,٥					٢			٣	٢			٧
١٦,٥ ـ ١٧					١		١	١	٢	١		٦
١٧ ـ ١٧,٥										٣	١	٤
١٧,٥ ـ ١٨									١	٢		٣
المجموع	١	١	٣	٤	٦	٥	٥	١٣	١٠	٦	١	٥٥

السؤال الثانى

إوجد الوسط الحسابى والانحراف المعيارى للتوزيع التكرارى لذات الحدين $(ح + ل)^{ن}$ حيث ح احتمال نجاح حدث معين و ل احتمال فشل هذا الحدث و ن عدد المحاولات .

السؤال الثالث

اشرح طريقة تركيب جداول الحياة ووضح بعض الفوائد العملية لهذه الجداول . اذكر أهم المعلومات التى تستوقف النظر فى جدول الحياة المصرى لسنة ١٩٣٧ .

السؤال الرابع

على فرض أن لديك مجموع معين يحتوى على ن من المفردات واخترت منه جميع العينات الممكنة التى يحتوى كل منها على م من المفردات ، أثبت أن المتوسط الحسابى لمتوسطات هذه العينات يساوى المتوسط الحسابى للمجموع الأصلى .

السؤال الخامس

باختيار قيم موجبة مناسبة للمتغير س ، ارسم المنحنى الذى يمثل الدالة ص = $س^{٢} + ٥$ س + ٢٠٠ على ورق الرسم البيانى . وبإيجاد معامل التفاضل عند نفس قيم س التى اخترتها ارسم المستقيم الذى يمثل الدالة المشتقة (الأولى) . وعلى فرض أن المعادلة الأصلية تمثل العلاقة بين التكاليف الكلية (ص) لأحد مصانع السكر مقدرة بالجنيهات وبين الإنتاج الأسبوعى (س) لهذا المصنع مقدراً بالأطنان ، فماذا تمثل الدالة المشتقة ؟ استنتج من المعادلة الأصلية معادلة ترى تمثل العلاقة بين التكاليف المتوسطة والإنتاج الأسبوعى للمصنع ، ومنها احسب تكاليف الطن الواحد فى المتوسط إذا أنتج المصنع ١٠ ، ٤٠ ، ٥٠ ، ٨٠ طن فى الأسبوع على التوالى . علق على هذه النتائج الأخيرة .

السؤال السادس

لماذا لا يعتبر المعدل الأولى للوفيات مقياساً صالحاً لمقارنة الحالة الصحية فى البلاد المختلفة ؟ وما هى طرق تعديله لجعله صالحاً لهذا الغرض ؟ كيف تعلل اتجاه معدل الوفيات فى مصر نحو الصعود خلال العشرين سنة الأخيرة على الرغم من تحسن الحالة الصحية فى البلاد خلال هذه الفترة ؟

سبتمبر سنة ١٩٤٦

السؤال الأول

استعمل البيانات الآتية فى إيجاد قيمة تقريبية لكل من المعدل الإجمالى والمعدل الصافى للتوالد Gross and Net Reproduction Rates

فئات الأعمار	عدد سكان الإناث بالآلاف	عدد المواليد	عدد الباقين على قيد الحياة من كل ألف من المواليد الإناث
١٥ — ١٩	١٥٥٨	٣٨٧٠٠	٩١٤
٢٠ — ٢٤	١١١٢	١٤٦٢٠٠	٨٩٩
٢٥ — ٢٩	١٥٩٥	١٩٩٣٠٠	٨٨٤
٣٠ — ٣٤	١٦٢٩	١٣٢٢٠٠	٨٦٨
٣٥ — ٣٩	١٦٢٧	٧١٩٠٠	٨٥٢
٤٠ — ٤٤	١٥٢٢	٢٢٢٠٠	٨٣٤
٤٥ — ٤٩	١٤٠١	١٧٠٠	٨١٣

مع العلم بأن نسبة المواليد الإناث إلى مجموع المواليد هى ٤٨,٦ ٪

السؤال الثانى

ما عدد جميع العينات الممكن أخذها من مجموع مكون من ن من المفردات (مهما كان عدد مفردات العينة) ؟ وما متوسط عدد مفردات العينة ؟ .

السؤال الثالث

استنتج قانون Spearman لإيجاد معامل الارتباط بين تراتيب الظواهر ثم طبق هذا القانون لإيجاد معامل الارتباط بين تراتيب الظاهرتين الآتيتين :

تراتيب س	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	١١	١٢	١٣	١٤	١٥	١٦
تراتيب ص	١٣	١٦	١٤	١٥	١٠	١٢	٤	١١	٥	٩	٨	٣	١	٦	٧	٢

السؤال الرابع

ما أوجه النقص فى الإحصاءات الحيوية المصرية ؟ إلى أى حد يمكننا معرفة الاتجاه العام لخصوبة السكان من واقع الإحصاءات المنشورة ، وما البيانات التى تقترح إضافتها لكى يكون قياس خصوبة السكان دقيقاً ؟ .

السؤال الخامس

حدّد بدقة المعنى الرياضى للاحتمال ، ثم اثبت أنه إذا كان احتمال نجاح حدث معين هو ح واحتمال فشله هو ل — فإن احتمال النجاح ر مرة إذا كررنا التجربة ن مرة هو :

$$\frac{\text{ن !}}{\text{ر ! (ن — ر) !}} \quad \text{ح}^{\text{ر}} \, \text{ل}^{\text{ن — ر}}$$

السؤال السادس

اشرح بالتفصيل كيف يمكن استخدام فكرة الخطأ المعيارى للمتوسط الحسابى فى مراقبة جودة المنتجات (Quality Control)

يونيو سنة ١٩٤٧

أجب عن الأسئلة الآتية :

السؤال الأول

البيانات الآتية مستخرجة من جدول الحياة المصرى للذكور .

e_x	T_x	L_x	q_x	P_x	d_x	l_x	x
؟	؟	٩١٢٦٥	؟	؟	؟	١٠٠٠٠٠	٠
؟	٣٤٧٣٤٧٤	؟	؟	؟	؟	؟	١
؟	؟	؟	؟	٠,٩٢٧٦٥	؟	٧٣٢٢١	٢

والمطلوب تكملة هذا الجدول بوضع البيانات المستعاض عنها بعلامات الاستفهام ، ثم إيجاد تعريف دقيق لكل من l_x ، d_x ، P_x ، q_x ، L_x ، T_x ، e_x وضح خطوات العمل بالتفصيل .

السؤال الثانى

أوجد المتوسط الحسابى والانحراف المعيارى لتوزيع بواسون
للاحتمالات الصغيرة "Poisson Distribution"

السؤال الثالث

(ا) أظهرت التجارب فى الماضى أن نسبة اللمبات الكهربائية المعيبة فى منتجات أحد المصانع هى ١٠٪ فما احتمال الحصول على ٤ لمبات معيبة إذا اخترنا من منتجات هذا المصنع عينة عدد مفرداتها ٢٠ على فرض أنه لم يطرأ أى تغيير على ظروف الإنتاج ؟

(ب) إذا اخترنا الآن عينة تحتوى على ١٠٠٠٠ لمبة من منتجات المصنع ذاته ووجدنا بينها ٩٠٠ لمبة معيبة ، فهل يدل ذلك على حدوث تغير جوهرى فى ظروف الإنتاج ؟

السؤال الرابع

أثبت أن معامل الارتباط بين تراتيب ن من أزواج القيم المتناظرة

$$1 - \frac{6 \Sigma د^2}{ن^3 - ن}$$

حيث $د_1$ ، $د_2$ ، $د_3$ ، . . . . . $د_ن$ تمثل الفروق بين التراتيب المتناظرة .

يونيو سنة ١٩٤٧ (الورقة الثانية)

أجب عن الأسئلة الآتية :

السؤال الأول

استعمل البيانات الآتية فى إيجاد قيمة تقريبية لكل من المعدل الإجمالى والمعدل الصافى للتوالد Gross and Net reproduction rates

فئات الأعمار	عدد السكان الإناث بالآلاف ١٩٣٧	عدد المواليد سنة ١٩٣٧	عدد الباقين على قيد الحياة من كل ألف من المواليد الإناث
١٥ — ١٩	٧٢٩	٨٧٤٨٠	٦٢٠
٢٠ — ٢٤	٦١٨	١١١٢٤٠	٦٠٨
٢٥ — ٢٩	٨٤٣	١٦٠١٧٠	٥٩٣
٣٠ — ٣٤	٩٥١	١٢٣٦٣٠	٥٧٥
٣٥ — ٣٩	٩١٣	٧٣٠٤٠	٥٥٤
٤٠ — ٤٤	٧٨٦	١٥٧٢٠	٥٣٠
٤٥ — ٤٩	٧٥٠	١٥٠٠	٥٠٦

مع العلم بأن نسبة المواليد الإناث تعادل ٤٩,٢ ٪ من مجموع المواليد .

أوجد كذلك متوسط عمر الأم عند ميلاد طفلها فى سنة ١٩٣٧ .

السؤال الثانى

الأرقام الآتية تبين رأس المال المستثمر فى متجر معين فى عشرة سنوات متتالية وكذلك الأرباح التى حققها المتجر فى نفس السنوات .

رأس المال بمئات الجنيهات	١٠	٢٠	٣٠	٤٠	٥٠	٦٠	٧٠	٨٠	٩٠	١٠٠
الربح بمئات الجنيهات	١	٤	٨	٥	١٠	١٥	١٤	٢٠	٢٢	٣٠

أوجد معامل الارتباط بين هذه المفردات ثم أوجد معامل الارتباط بين تراتيبها وقارن بين النتيجتين .

أوجد كذلك قيمة الربح كنسبة مئوية من رأس المال فى المتوسط .

السؤال الثالث

وفق منحنى طبيعى (Normal Curve of Error) للتوزيع التكرارى المبين بالجدول الآتى :

فئات المتغير	التكرار
١٧٢ — ١٧٥	٣
١٧٦ — ١٧٩	٩
١٨٠ — ١٨٣	٢٩
١٨٤ — ١٨٧	٧٦
١٨٨ — ١٩١	١٠٤
١٩٢ — ١٩٥	١١٠
١٩٦ — ١٩٩	٨٨
٢٠٠ — ٢٠٣	٣٠
٢٠٤ — ٢٠٧	٦
٢٠٨ — ٢١١	٤
٢١٢ — ٢١٥	٢
٢١٦ — ٢١٩	١

مع العلم بأن المتوسط الحسابى للتوزيع = ١٩١,٨ والانحراف المعيارى = ٦,٤٨ .

ارسم المضلع التكرارى للتوزيع الأصلى على ورق الرسم البيانى ثم وضح على نفس الرسم المنحنى الطبيعى الذى وفقته .

سبتمبر سنة ١٩٤٧

أجب عن الأسئلة الآتية :

السؤال الأول

الجدول الآتى يوضح العلاقة بين أطوال مجموعة من تلاميذ المدارس مقدرة بالبوصات (س) وأوزان هذه المجموعة مقدرة بالأرطال (ص) .

أوجد معامل الارتباط بين هاتين الظاهرتين واختبر دلالته .

قيم س / قيم ص	٤٠ – ٤٣	٤٤ – ٤٧	٤٨ – ٥١	٥٢ – ٥٥	٥٦ – ٥٩	٦٠ – ٦٣	المجموع
٣٥ – ٥٤	١	١٠	١٥				٢٦
٥٥ – ٧٤			٦	٢٢	٣		٣١
٧٥ – ٩٤				٢	٨	٢	١٢
٩٥ – ١١٤					١	٢	٣
١١٥ – ١٣٤				١			١
١٣٥ – ١٥٤					١	١	١
المجمـــوع	١	١٠	٢١	٢٥	١٣	٥	٧٥

السؤال الثانى

الجدول الآتى يبين الحوادث التى وقعت لـ ٦٤٧ عامل فى أحد مصانع الذخيرة خلال سنة معينة . والمطلوب توفيق منحنى بواسون Poisson Curve لهذه البيانات ثم اختبار حسن المطابقة باستعمال جداول χ^2

(لو هـ = ٠,٤٣٤٣)

١٠

عدد الحوادث	التكرار المشاهد
٠	٤٤٧
١	١٣٢
٢	٤٢
٣	٢١
٤	٣
٥	٢

السؤال الثالث

الجدول الآتى يبين نتيجة تجربة رمى عشر قطع من النقود ١٠٢٤ مرة . والمطلوب إيجاد التكرارات النظرية المقابلة للتكرارات المشاهدة المبينة بالجدول بتطبيق قانون برنوللى للاحتمالات (اعتبر ظهور الوجه بمثابة نجاح فى هذه التجربة) .

عدد مرات النجاح	التكرار المشاهد
٠	١
١	٩
٢	٤٦
٣	١٢٣
٤	٢٠٧
٥	٢٥٣
٦	٢٠٩
٧	١٢٢
٨	٤٤
٩	٩
١٠	١

أوجد كذلك المتوسط الحسابى والانحراف المعيارى للتوزيع التكرارى النظرى .

يونية سنة ١٩٤٨

أجب عن الأسئلة الآتية :

السؤال الأول

الجدول الآتى يوضح العلاقة بين سعر الخصم (س) والاحتياطى النقدى للبنوك التجارية محتسباً كنسبة مئوية من مجمل ودائعها (ص) خلال ٥٠٠ أسبوع

والمطلوب إيجاد معامل الارتباط بين المتغيرين واختبار دلالته .

قيم س / قيم ص	١– ٢	٢– ٣	٣– ٤	٤– ٥	٥– ٦	٦– ٧	٧– ٨	٨– ٩	٩– ١٠	المجموع
٢١ – ٢٣	٢								١	٣
٢٣ – ٢٥	٢								١	٣
٢٥ – ٢٧		١	٣							٤
٢٧ – ٢٩				٤٠	٧٠	٤٠	٩	١		١٦٠
٢٩ – ٣١			٥	٨	٢١٠	١٧				٢٤٠
٣١ – ٣٣					٧	٤٣				٥٠
٣٣ – ٣٥					٨	١٢				٢٠
٣٥ – ٣٧		٣				٨	١			١٢
٣٧ – ٣٩					٧			١		٨
المجمـــوع	٤	٤	٨	٤٨	٣٠٢	١٢٠	١٠	٢	٢	٥٠٠

السؤال الثانى

الجدول الآتى يبين التوزيع التكرارى لوقوع حادث معين بأحد المصانع . والمطلوب توفيق منحنى بواسون للاحتمالات الصغيرة لهذا التوزيع وإيجاد التكرارات النظرية المقابلة للتكرارات المشاهدة .

عدد الحوادث	التكرار	عدد الحوادث	التكرار	عدد الحوادث	التكرار	عدد الحوادث	التكرار
٠	١٤	٣	٧٠	٦	٣١	٩	٩
١	٣٧	٤	٦٤	٧	١٩	١٠	٥
٢	٧٦	٥	٥٣	٨	١٤	١١	٥
						١٢	١

السؤال الثالث

اختبر حسن مطابقة التوزيع التكرارى المبين بالسؤال الثانى لتوزيع بواسون باستخدام جداول χ^2 .

سبتمبر سنة ١٩٤٨

السؤال الأول

وضح معنى خطى الانحدار للمتغيرين س ، ص مستعيناً بالرسم البيانى . ثم قارن بإيجاز بين طريقة المربعات الدنيا وطريقة المساحات الدنيا فى إيجاد معادلتيهما . (استنتاج هذه المعادلات رياضياً غير مطلوب) .

السؤال الثانى

اذكر بالتفصيل خواص المنحنى الطبيعى "Normal Curve of Error" ثم وضح كيف يمكن استغلال بعض هذه الخواص فى الشئون التجارية (صناعة الملابس الجاهزة مثلا) .

السؤال الثالث

وضح كيف يمكن تعديل المعدل الأولى للوفيات لجعله صالحاً لمقارنة الحالة الصحية بين دولة وأخرى ، أو بين وقت وآخر فى نفس الدولة .

السؤال الرابع

استخدم البيانات الآتية فى تركيب رقمين قياسيين لنفقات المعيشة ، أحدهما خاص بطبقة العمال والآخر خاص بالطبقة المتوسطة ، مستعملا فى

الحالتين طريقة الوسط الحسابى المرجح بالأوزان لمناسيب الأسعار .

أبواب الإنفاق	النسبة المئوية من الدخل المخصصة للانفاق على كل باب		النسبة المئوية للزيادة فى الأسعار	
	طبقة العمال	الطبقة المتوسطة	طبقة العمال	الطبقة المتوسطة
الطعام	٦٠	٣٦	٥٩	٦٤
إيجار المسكن	١٦	١٥	٥١	٨٥
الملابس	١٢	٩	١١٥	١١٥
الإضاءة	٨	٤,٥	٧٠	٧٠
متنوعات	٤	٣٥,٥	٨٠	٨٥
	١٠٠	١٠٠		

يونيو سنة ١٩٤٩

أجب عن الأسئلة الآتية :

السؤال الأول

أوجد نسبة الارتباط (ى ص س) من واقع البيانات الموضحة بالجدول الآتى :

عدد الأشخاص بالمسكن (ص)	عدد حجرات المسكن (س)									المجموع
	٢	٣	٤	٥	٦	٧	٨	٩	١٠	
١	٢٦	٩	٨	٥	١	٤	١	١	١	٥٦
٢	١٤	٢٤	٦٠	٤٠	٢٩	٩	٨	٣	٤	١٩١
٣	٥	٢٤	٦١	٥٧	٣٤	١١	٨	٣	٥	٢٠٨
٤	٤	١٦	٥٧	٤٤	٢٧	١٦	٣	٢	١٠	١٧٩
٥	٣	١٨	٣٦	٤٢	٢١	٦	٥	٥	١٢	١٤٨
٦	٢	٨	٢١	٤٣	١٩	٤	٥	٣	٧	١١٢
٧	٠٠	٥	١٨	١٤	١١	٥	٣	٢	٤	٦٢
٨	٠٠	١	١٤	١١	١٠	٠٠	٢	٢	٤	٤٤
المجمـــوع	٥٤	١٠٥	٢٧٥	٢٥٦	١٥٢	٥٥	٣٥	٢١	٤٧	١٠٠٠

مع العلم بأن الانحراف المعيارى لقيم ص = ١,٨٣

السؤال الثانى

بإيجاد معامل الاختلاف (Goefficient of Variation) لكل من عدد حجرات المسكن وعدد الأشخاص بالمسكن ، بين أيهما تظهر درجة أكبر من التشتت .

السؤال الثالث

وفق منحنى طبيعى للتوزيع التكرارى المبين بالجدول الآتى ، ثم ارسم المدرج التكرارى ومعه على نفس الرسم المنحنى الذى وفقته :

مراكز الفئات	٥٩,٥	٦٢,٥	٦٥,٥	٦٨,٥	٧١,٥	٧٤,٥	٧٧,٥	٨٠,٥
التكرار	٤	٢٤	١٨٨	٤٠٤	٣٨٧	١٥٨	٣٠	٥

يونية سنة ١٩٥٠

السؤال الأول

استخدم البيانات الآتية فى إيجاد قيمة تقريبية لكل من المعدل الإجمالى والمعدل الصافى للتوالد

فئات الأعمار	عدد السكان الإناث بالآلاف سنة ١٩٤٠	عدد المواليد سنة ١٩٤٠	عدد الباقين على قيد الحياة من كل ١٠٠٠ من المواليد الإناث
١٥ — ٢٠	١٤٥٨	١٧٤٩٦٠	٦٢٠
٢٠ — ٢٥	١٢٣٦	٢٢٢٤٨٠	٦٠٨
٢٥ — ٣٠	١٦٨٦	٣٢٠٣٤٠	٥٩٣
٣٠ — ٣٥	١٩٠٢	٢٤٧٢٦٠	٥٧٥
٣٥ — ٤٠	١٨٢٦	١٤٦٠٨٠	٥٥٤
٤٠ — ٤٥	١٥٧٢	٣١٤٤٠	٥٣٠
٤٥ — ٥٠	١٥٠٠	٣٠٠٠	٥٠٦

مع العلم بأن نسبة المواليد الإناث إلى مجموع المواليد تعادل ٤٩,٢ ٪ .
أوجد كذلك عدد الأمهات اللاتى تقل أعمارهن عن ٢٨ سنة فى سنة ١٩٤٠ .

السؤال الثانى

الجدول الآتى يبين المرتبات التى حصل عليها ٤٨ طالباً فى اختبارين مختلفين :

الاختبار الأول / الاختبار الثانى	مقبـــول	جيـــد	جيد جـــداً	المجموع
مقبول	١٠	٢	١	١٣
جيـــد	٤	١٧	٣	٢٤
جيد جداً	٢	٣	٦	١١
المجموع	١٦	٢٢	١٠	٤٨

أوجدمعامل التوافق (Root - Mean - Square Contingency) بين نتائج هذين الاختبارين .

السؤال الثالث

الجدول الآتى يبين التوزيع التكرارى لعدد الحوادث التى وقعت لـ ١٠٠٠ عامل فى أحد المصانع خلال سنة معينة . والمطلوب توفيق منحنى بواسون للاحتمالات الصغيرة لهذا التوزيع ثم إيجاد النسبة المئوية لسوء المطابقة

عدد الحوادث	التكرار المشاهد
٠	٣٥٠
١	٣٦٠
٢	١٩٠
٣	٧٠
٤	٢٠
٥	١٠

(لو هـ — ٠,٤٣٤٣)

١٠

السؤال الرابع

اشرح بالتفصيل طريقة تركيب جداول الحياة (Life Tables) مع الإشارة بإيجاز إلى الأهمية العملية لهذه الجداول .

مايو سنة ١٩٥١

أجب عن السؤالين الأول والثانى واثنين فقط من الأسئلة الباقية :

السؤال الأول

الجدول الآتى يبين العلاقة بين أعمار مجموعة من الرجال (س) وأعمار زوجاتهم (ص) .

ص / س	٢٠–	٣٠–	٤٠–	٥٠–	٦٠–	٧٠–	٨٠– ٩٠	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		٦٥
٤٠ – ٤٥		٣	١٢	١٠	٥	١	١	٣٢
المجموع	٨٠	٢٤٠	١٤٥	٧٥	٤٠	١٥	٥	٦٠٠

والمطلوب إيجاد نسبة الارتباط (ى ص.س) بين هذين المتغيرين .

السؤال الثانى

إذا كان احتمال فقد سفينة فى رحلة معينة هو ٠,٠٢ وكانت إحدى شركات الملاحة التى تقطع سفنها هذه الرحلة تملك ستة سفن فما احتمال :

(ا) فقد سفينة واحدة من سفن هذه الشركة

(ب) فقد سفينتين على الأكثر .

(حـ) عدم فقد أية سفينة .

السؤال الثالث

اشرح فى شىء من التفصيل طريقة حساب المعدل الصافى للتوالد ثم وضح أهمية هذا المقياس فى تقدير خصوبة السكان .

السؤال الرابع

أوجد الوسط الحسابى والانحراف المعيارى لتوزيع بواسون للاحتمالات الصغيرة .

السؤال الخامس

ما أهم البيانات التى تجمع عن الإنتاج الصناعى فى مصر ؟

لو كنت بصدد عمل رقم قياسى للانتاج الصناعى بين كيف تستفيد من هذه البيانات فى تكوين هذا الرقم .

السؤال السادس

أثبت أن

$$\frac{\Sigma (س - \bar{س}) (ص - \bar{ص})}{ن \; ع_{س} \; ع_{ص}} = \frac{ع^{٢}_{س} + ع^{٢}_{ص} - ع^{٢}_{ف}}{٢ \; ع_{س} \; ع_{ص}}$$

حيث $\sigma_س$ و $\sigma_ص$ يمثلان على التوالى الانحرافين المعياريين لقيم س وقيم ص .

و $\sigma_ف$ تمثل الانحراف المعيارى للفروق بين قيم «س» و «ص»

وعلى فرض أن كانت كل من قيم « س » وقيم « ص » تمثل مرتبات مسلسلة حسب الأعداد الطبيعية إثبت أن الصيغة الأخيرة تتخذ الشكل الآتى :

$$١ - \frac{٦ \Sigma ف^{٢}}{ن (ن^{٢} - ١)}$$ حيث ف$_{١}$ ، ف$_{٢}$ ، ..... ف$_{ن}$

تمثل الفروق بين المرتبات المتناظرة لـ « س » و « ص »

و « ن » عدد أزواج القيم .

فهرس

صفحة

الباب الأول

الطريقة الإحصائية

تعريف الطريقة الإحصائية ٧
الأخطاء الشائعة فى استخدام الطريقة الإحصائية ١١
تطور الطريقة الإحصائية وظهور الإحصاء كعلم مستقل ١٣

الباب الثانى

جمع وتبويب البيانات الإحصائية

طرق جمع البيانات الإحصائية ١٧
مراجعة البيانات الإحصائية ٢٢
طرق تبويب البيانات الإحصائية ٢٣
التبويب الآلى ٢٦
أهم مصادر الإحصاءات المصرية ٣٦

الباب الثالث

طرق عرض البيانات الإحصائية

(١) العرض الجدولى ٤٠
(٢) العرض البيانى ٤٦
طريقة الرسم البيانى ٤٦
الرسم البيانى كإحدى طرق العرض الإحصائى ٥١

صفحة

الرسوم البيانية اللوغاريتمية ٦٥
طريقة عمل التقسيم اللوغاريتمى ٦٩
تحديد موقع النقط بالنسبة للمحور اللوغاريتمى ٧٢
التطبيقات العملية للرسوم البيانية اللوغاريتمية ٧٣
استخدام الأشكال الهندسية فى عرض البيانات ٧٩
(٣) العرض التصويرى ٨٤

الباب الرابع
التوزيع التكرارى

التوزيع التكرارى ٨٨
اختيار فئات التوزيع التكرارى ٩٢
طريقة وضع الأرقام فى الفئات التكرارية ٩٩
المدرج التكرارى أو الهستوجرام ١٠١
المدرج التكرارى لتوزيع ذى فئات غير متساوية ١٠٢
المضلع التكرارى ١٠٦
المنحنى التكرارى ١٠٩
المتغيرات المتصلة والمتغيرات المنفصلة ١١٠
التوزيع التكرارى المتجمع ١١١
منحنى لورنز ١١٥
الأشكال المختلفة للمنحنيات التكرارية ١١٨
وصف التوزيعات التكرارية ١٢٢

الباب الخامس
مقاييس التوسط

معنى المتوسط ١٢٤
الشروط الواجب توافرها فى المتوسط الجيد ١٢٥
الوسط الحسابى ١٢٦
الوسط الحسابى المرجح بالأوزان ١٢٩

صفحة

خواص الوسط الحسابى ١٣٠
إيجاد الوسط الحسابى للتوزيع التكرارى عملياً ١٣٧
مراجعة شارلييه لاختبار صحة العمليات الحسابية ١٤٠
الوسط الحسابى لتوزيع تكرارى ذى فئات غير متساوية ١٤٢
الوسط الهندسى ١٤٤
حساب الوسط الهندسى للتوزيع التكرارى عملياً ١٤٥
التطبيقات العملية للوسط الهندسى ١٤٦
الوسط التوافقى ١٥٠
حساب الوسط التوافقى للتوزيع التكرارى ١٥٣
الوسط التربيعى ١٥٤
حساب الوسط التربيعى للتوزيع التكرارى ١٥٦
العلاقة بين المتوسطات ١٥٧
الوسيط ١٥٩
إيجاد الوسيط للتوزيع التكرارى ١٦١
إيجاد الوسيط بالرسم البيانى ١٦٢
التطبيقات العملية للوسيط ١٦٥
الربيع الأعلى والربيع الأدنى ١٦٦
المنوال ١٦٩
إيجاد المنوال بطريقة الفروق ١٧٠
إيجاد المنوال بطريقة الرافعة ١٧٣
حالات استعمال المنوال ١٧٥
العلاقة بين الوسط الحسابى والوسيط والمنوال ١٧٦

الباب السادس

مقاييس التشتت

معنى التشتت ١٧٨
المدى ١٨٠
المدى بين المئين العاشر والتسعين ١٨٠

صفحة

الانحراف الربيعى ١٨١
الانحراف المتوسط ١٨١
إيجاد الانحراف المتوسط للتوزيع التكرارى ١٨٣
الانحراف المعيارى ١٨٧
الطرق المختلفة لحساب الانحراف المعيارى ١٩٢
مراجعة شاربيه فى حساب الانحراف المعيارى ١٩٤
تصحيحات شبرد ١٩٦
خواص الانحراف المعيارى وتطبيقاته ١٩٧
مقاييس التشتت النسبى ١٩٩
معامل الاختلاف ٢٠٠

الباب السابع

العزوم ومقاييس الالتواء والتفرطح

تعريف العزوم ٢٠٢
العزم الثالث كقياس للالتواء ٢٠٤
حساب العزم الثالث للتوزيع التكرارى ٢٠٧
المقاييس الأخرى للالتواء ٢١١
قياس الالتواء بمقارنة الوسط الحسابى والمنوال ٢١١
قياس الالتواء بمقارنة الربيعين ٢١٣
مقاييس التفرطح ٢١٥
حساب العزم الرابع للتوزيع التكرارى ٢١٧
العلاقة بين التشتت والتفرطح ٢١٨
إيجاد المنوال باستخدام معاملى الالتواء والتفرطح ٢٢١
مراجعة شاربيه فى حساب العزوم العليا ٢٢٣
تصحيحات شبرد فى حساب العزوم العليا ٢٢٤

صفحة

الباب الثامن

مقدمة لنظرية الاحتمالات والعينات

التوزيعات التكرارية النظرية ٢٢٥
التوزيع التكرارى لذات الحدين (حالة خاصة) ٢٢٦
مبادئ أولية فى نظرية الاحتمالات ٢٣٥
المعنى العادى والمعنى الرياضى للاحتمال ٢٣٥
جمع الاحتمالات ٢٣٩
ضرب الاحتمالات ٢٤٠
قانون برنوللى ٢٤٤
المعنى الإحصائى للاحتمال ٢٤٤
الحالة العامة للتوزيع التكرارى لذات الحدين ٢٤٧
الوسط الحسابى والانحراف المعيارى للتوزيع التكرارى لذات الحدين ٢٤٨
الوسط الحسابى والانحراف المعيارى لنسبة النجاح ٢٥٦
التوزيع الطبيعى ٢٦٠
خواص المنحنى الطبيعى ٢٦٦
عزوم المنحنى الطبيعى ٢٦٨
توزيع بواسون ٢٧٠
عزوم توزيع بواسون ٢٧٢

الباب التاسع

توفيق المنحنيات واختبار حسن المطابقة

الوضع المشاهد والوضع المثالى لتغير الظواهر ٢٧٥
كيفية اختيار المنحنى الذى يناسب تغير الظاهرة ٢٧٦
طريقة المربعات الدنيا ٢٧٧
توفيق المنحنى الطبيعى للتوزيعات التكرارية ٢٨٤
توفيق منحنى بواسون للتوزيعات التكرارية ٢٩٠
اختبار حسن المطابقة ٢٩٣
النسبة المئوية لسوء المطابقة ٢٩٤
اختبار χ^2 لحسن المطابقة ٢٩٥

صفحة

الباب العاشر

الارتباط

معنى الارتباط وأهميته ٣٠٢

قياس الارتباط بين الظواهر القابلة للقياس العددى ٣٠٤

إيجاد معامل الارتباط عملياً ٣٠٨

الارتباط بين التوزيعات التكرارية ٣١٠

طريقة الأقطار ذات الفروق المتساوية ٣١٧

معامل سبيرمان للارتباط بين المرتبات ٣٢٠

دلالة معامل الارتباط ٣٢٥

الارتباط بين الظواهر غير القابلة للقياس العددى ٣٢٨

معامل الاقتران ٣٢٩

معامل التوافق ٣٣١

الباب الحادى عشر

التوضيح البيانى للارتباط

خطوط الانحدار ومعادلاتها ٣٣٨

طريقة المربعات الدنيا لإيجاد معادلتى خطى الانحدار ٣٣٩

طريقة المساحات الدنيا لإيجاد معادلتى خطى الانحدار ٣٤١

خطوط الانحدار المنحنية ونسبة الارتباط ٣٤٧

حساب نسبة الارتباط عملياً ٣٤٨

الباب الثانى عشر

نظرية العينات

— أنواع العينات ٣٥٣

— حجم العينة وعلاقته بدقة النتائج ٣٥٧

الوسط الحسابى لمتوسطات العينات ٣٥٩
الانحراف المعيارى لمتوسطات العينات (الخطأ المعيارى) ٣٦١
الخطأ المحتمل للوسط الحسابى ٣٦٨
تطبيقات عملية للخطأ المعيارى للوسط الحسابى ٣٦٩
اختبار دلالة الفرق بين متوسطى عينتين ٣٧٠
الأخطاء المعيارية لأهم المقاييس الإحصائية ٣٧٦

الباب الثالث عشر

تحليل السلسلات الزمنية

التغيرات الاتجاهية أو الاتجاه العام ٣٧٩
التغيرات الموسمية ٣٨٠
التغيرات الدورية ٣٨٠
التغيرات العرضية ٣٨١
طرق تعيين الاتجاه العام ٣٨٢
أولا — طريقة المتوسطات المتحركة ٣٨٢
ثانياً — إيجاد الاتجاه العام بتوفيق مستقيم أو منحن ٣٩٠
طرق حساب التغيرات الموسمية ٣٩٧
أولا — إيجاد النسب الموسمية بطريقة المتوسطات البسيطة ٣٩٧
استبعاد الاتجاه العام من النسب الموسمية ٤٠٢
ثانياً — إيجاد التقلبات الموسمية بدلالة الوحدات المطلقة ٤٠٨
استبعاد تأثير الموسم من القيم المشاهدة للظاهرة ٤١٣
إيجاد النسب الدورية عمليا ٤١٥
الارتباط بين السلسلات الزمنية ٤١٦

الباب الرابع عشر

الأرقام القياسية

تعريف الرقم القياسى ٤٢٠

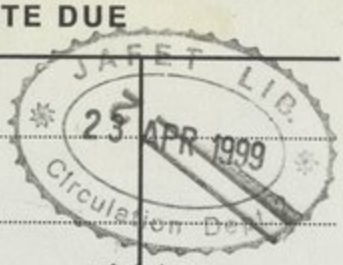
	صفحة
اختيار السلع والمصادر التى تستقى منها الأسعار	٤٢٢
اختيار سنة الأساس	٤٢٣
اختيار الأوزان	٤٢٤
تلخيص البيانات	٤٢٥
الأرقام التجميعية	٤٢٥
الرقم القياسى الأمثل	٤٢٩
الأرقام النسبية	٤٢٩
اختبار صحة الأرقام القياسية	٤٣١
اختبار الانعكاس الزمنى	٤٣٢
اختبار الانعكاس المعاملى	٤٣٤
تعديل الأرقام القياسية	٤٣٩
الأرقام القياسية المصرية	٤٤١
الرقم القياسى لأسعار الجملة	٤٤١
الرقم القياسى لنفقات المعيشة	٤٤٣

° ° °

جدول مساحات منحنى الخطأ المتماثل	٤٤٦
جدول X^2	٤٤٨
دليل بالرموز المستعملة فى الكتاب	٤٤٩
أسئلة مختارة من الامتحانات التى وضعها المؤلف	٤٥٠

١٩٥٢/٣٣٥٢

DATE DUE


JAFET LIB.
23 APR 1999
Circulation Dept.


JAFET LIB.
11 MAY 1999
Circulation Dept.2

311.2:H21tA:c.1
حمدى ،محمد مظلوم
طرق الاحصاء
AMERICAN UNIVERSITY OF BEIRUT LIBRARIES
01012969

