عنوان الكتاب : مبادئ الإحصاء

المؤلـــــف : عبد المنعم ناصر الشافعى

سنة النشــر : ١٩٣٩

رقم العهــدة : د ١٠٧٢٢

الـ ACC : ١٣٩٦٧

عدد الصفحات : ٣٨٠

رقم الفيلـــم : ١٩

مبادئ الإحصاء

الجزء الأول

تأليف

عبد المنعم ناصر الشافعي

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

بكالوريوس الشرف في العلوم الرياضية . بكالوريوس فى التجارة .
دكتوراه فى الفلسفة . زميل في الجمعية الملكية للاحصاء بلندن .

مدرس بكلية التجـــارة

يطلب من مكتبة النهضة المصرية بالقاهرة

القاهرة

مطبعة مصر شركة مساهمة مصرية

٤٠ شارع نوبار باشا (سابقا شارع البوليس)

١٩٣٩

بسم الله الرحمن الرحيم

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

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

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

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

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

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

كلية التجارة — المؤلف

أول يناير سنة ١٩٣٩ — **عبد المنعم ناصر الشافعى**

الفهرس

فهرس الأشكال

فهرس الجـــداول

مبادئ الاحصاء

الباب الأول

مقدمة

تعريف علم الاحصاء

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

منشأ فكرة الاحصاء

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

الاحصاء كعلم يبحث في جمع الحقائق الخاصة بشؤون الدولة

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

استخدام الاحصاء في الأبحاث العلمية والشئون الخاصة

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

الاحصاء الآن علم مستقل، له نظرياته وقواعده

٥ — وكان مما ساعد على سعة تطبيق هذا العلم ونشر تعاليمه أن توفر على دراسته عدد كبير من العلماء النابغين، فبحثوا نظرياته وبنوها على أسس علمية صحيحة، وهذبوا طرقه العملية على ضوء هذه النظريات، والخبرة العملية التى اكتسبوها من أبحاثهم. وكان من بين هؤلاء بعض فحول الرياضيين، مثل عائلة برنولى، وفردريك جاوس الألمانى، ولابلاس الفرنسى، وكيتليه البلجيكى، وجولتون الانجليزى؛ وفي وقتنا الحاضر لاينسى فضل كارل بيرسون، وآرثر بولى، وأدنى يول في انجلترا، وإرفنج فيشر في أمريكا[1]، وغيرهم في البلاد الأخرى. وهؤلاء الحديثون قد تفرغوا لدراسة علم الاحصاء، واستنباط نظرياته، وكشف غوامضه، وتطبيق هذه النظريات في العلوم الاقتصادية والاجتماعية والعلوم

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

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

كان الاحصاء في مبدأ نشأته لا يستخدم الطريقة الرقمية

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

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

٧ — بعد ذلك ظهرت بالتدريج أفضلية استخدام الطريقة الرقمية للدلالة على الظواهر التى كان يطلب مراقبتها وتسجيل أحوالها، لما فيها من تمام الوضوح ودقة التعبير — وبذلك عم استخدام الأرقام. وكانت الظواهر المختلفة تقاس كماً ويعبر عن مقاييسها بأعداد حسابية. ومن ثم أخذ الاحصاء شكلاً جديداً عبر عنه الانجليز في القرن السابع عشر بالحساب السياسى أى (Political Arithmetic)

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

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

استخدام النظريات الرياضية فى استنباط القوانين الاحصائية وتدعيم نتائجها

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

علم الاحصاء يطبق فى بحث مسائل علمية كثيرة

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

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

استخدام الاحصاء فى العلوم الطبيعية البحتة

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

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

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

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

وقد استخدم علماء النفس الطرق الإحصائية فى قياس ذكاء الأشخاص والتمييز بين المجموعات المختلفة من الأجناس البشرية ، وفى دراسة العلاقة بين ذكاء الشخص ومهارته فى بعض النواحى دون الأخرى ؛ وأمكنهم بواسطتها أيضاً دراسة تأثير البيئات المختلفة فى الأشخاص وسرعة نضوجهم العقلى . ومن البارزين فى هذا الميدان الأستاذ سبيرمان (Spearman) فى جامعة لندن .

الاحصاء فى العلوم الاقتصادية يرجع إلى زمن بعيد

١١ — وفى ميدان العلوم الاقتصادية يعتبر علم الإحصاء بمثابة أحد العناصر الأساسية التى يكوّن منها العلماء نظرياتهم ، والمحكّ الذى يختبرون به هذه النظريات ليتبينوا صلاحيتها لتفسير الظواهر الاقتصادية والاجتماعية المشاهدة فى الواقع . فبواسطة عمل إحصاءات التجارة الخارجية مثلا يمكن دراسة تأثير الضرائب الجمركية على الإنتاج الداخلى ، وعلى مستوى الأسعار . وبعمل إحصاءات عن كمية النقد المتداول وكمية الائتمان ، يمكن درس حالة الأسعار وما يتبعها من رواج فى التجارة ونشاط فى الأعمال .

الاحصاء فى دوائر الأعمال الخاصة

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

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

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

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

استخدام الاحصاء فى المسائل الاجتماعية

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

الاحصاء يساعد على تحديد النقط الأساسية التى تستلزم البحث والعلاج

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

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

المراجـــع

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

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

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

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

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

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

الباب الثانى

الطريقة الإحصائية

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

عملية الاحصاء تبدأ بملاحظة الظواهر التى نبحثها وجمع الحقائق عن أحوالها

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

عملية العـد في الاحصاء هي العملية الابتدائية : الاحصاء ليس مجرد علم عد

١٧ — والطريقة المتبعة فى الإحصاء هى — كما يدل عليها الاسم العربى — طريقة العدّ . ففى كل أبحاثنا الإحصائية نعبر عن مشاهداتنا للظواهر تعبيراً رقمياً ، ونسجل هذه الملاحظات فى صورة رقمية . فنى المثال الذى قدمناه فى البند السابق نعد لوزات القطن التى يحملها كل نبات ؛ كما أننا فى إحصاء السكان نعد الأشخاص المقيمين فى كل جهة من جهات المملكة ؛ وفى إحصاءات العمل نعد العمال المشتغلين فى كل صناعة وعدد العاطلين أيضاً ، وهكذا . وعملية العدّ هـذه هى خطوة أساسية فى الأعمـال الإحصائية ، مما جعل بعض الناس يعرف علم الإحصاء بأنه **علم العد** (Science of Counting) ؛ ولكن هذا تعريف قاصر لايفى بالمعنى ، وذلك لأن عملية العد ما هى إلا العملية الابتدائية . ومع أنها خطوة أساسـية فى العمل فإنها ليست كل شىء ، كما سيتضح لنا فيما يلى .

البيـانات الأوليـة

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

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

الطـرق التى نتبعها فى جمع البيـانات الأوليـة وفى استخراج المعلومات الثانوية تختلف عن بعضها

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

فإذا أردنا مثلا أن ندرس مستوى الأسعار العام فى بلد ما ، نبدأ بجمع البيانات الأولية عن الأسعار ؛ ولهذا يجب أولا أن نحدد أى السلع التى نجمع أسعارها ، ثم الوحدات التى نستخدمها فى بيان الأسعار ، ثم الهيئات أو الأشخاص الذين نلجأ إليهم فى معرفة هذه الأسـعار ، وكيفية إرسالهـا منهم إلينا لتصلنـا فى مواعيـد معينة ، وهكذا .

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

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

يجب أن نحدد المسائل التى نريد بحثها والظواهر التى تتصل بها ثم نجمع البيانات على هـذا الأسـاس

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

أحياناً يصعب تحديد العناصر المتصـلة بالموضـوع فنعمل دراسة تمهيدية لمعرفتها

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

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

عملية العدد تتطلب تعيين وحـــدة

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

الوحـدات الحسـابية مفروض أنها متطابقـة

فلاجراء عملية العد ، إذاً ، يجب أن نتفق أولا على وحدة ، ثم نبحث عن عـدد ما يحتويه الشىء المعدود من هذه الوحدات . والأصل فى الوحدات التى تضاف إلى بعضها لتكون شيئاً معيناً ، أن تكون كلها متماثلة متشابهة متساوية من جميع الوجوه ، أو كما يقول الرياضيون « متطابقة » ؛ فنحن مثـلا لا نضيف وحدات من البرتقال إلى وحدات من المنجه ، وإن كنا نتجاوز أحياناً ونضيف عدد صناديق البرتقال على عدد صناديق اليوسفى المصدرة إلى الخارج على اعتبار أنهما « صـادرات موالح » . ومع أن كثيراً منا يعتقد أن « للذكر مثـل حظ الأنثيين » ، فهم لا يغضبون كثيراً ولا قليلاً إذا قيل لهم إن تعداد القطر المصرى فى سنة ١٩٣٧ كان ١٥,٩٠٤,٥٢٥ نسمة ، حتى ولو كان من السكان ٧,٩٤٧,١٩٣ ذكوراً و ٧,٩٥٧,٣٣٢ إناثاً . والحقيقة أننا نكلف الأشياء ضد طباعها إذا تطلبنا أن تكون الوحدات متساوية تماما قبل أن نعدها سويا . فمن ذا الذى رأى شيئين

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

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

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

يجب تقسيم المجموعة إلى أجزاء متجانسة قبل عدها

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

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

تحديد معنى الوحدة يكون بتعيين الصفات المهمة التى يجب توافرها فى كل مفردة

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

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

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

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

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

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

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

وعلى هذا الاعتبار يمكننا أن نقول إن عدد سكان القطر المصرى فى سنة ١٩٣٧ كان ١٥ر٩٠٤ر٥٢٥ نسمة ؛ ونفهم أن الوحدة هنا رمز يدل فى نفس الوقت على الفلاح فى قريته ، والتاجر ساكن المدينة ، والمرأة فى منزلها ، واليافع فى مدرسته ، والصانع فى عمله ، وهكذا . والكل يشتركون فى صفة واحدة وهى أنهم كانوا جميعاً « على قيد الحياة فى منتصف الليلة الواقعة بين يومى ١٨ و ١٩ فبراير سنة ١٩٣٧ فى الأراضى المصرية » .

وكذلك نقول إن مقدار الوارد إلى مصر من أجهزة الراديو (للاستقبال) فى سنتى ١٩٣٦ و ١٩٣٧ كان على الترتيب ١٥ر٢٦٧ و ٢٠ر٠٨٠ جهازاً . ونعلم أن هذه الأجهزة تختلف بعضها عن بعض فى النوع والقيمة والحجم ؛ ولا يمنعنا هذا من أن نستنتج أن استعمال هذه الأجهزة فى مصر زاد زيادة محسوسة بين السنتين المذكورتين .

لدراسة الصفات نقيسها ونعبر عنها بأرقام

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

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

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

بعض الصفات يمكن قياسها والبعض يتعذر أو يستحيل قياسها مباشرة

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

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

دقة المقاييس تتوقف على الأجهزة المستعملة والغرض المقصود من البحث

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

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

قواعد جمع البيانات الاحصائية

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

ولتوضيح هذه القواعد نأخذ مثالا عملياً تظهر فيه الخطوات المهمة .

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

تعريف الوحدات التى يتناولها البحث

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

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

توحيد نظام وميعاد جمع البيانات عن الصناعات والجهات المختلفة

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

طبع كشوف وتوزيعها لجمع البيانات بواسطتها

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

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

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

اسم المصنع .... العنوان ....

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

يوم العطلة الأسبوعى .... مواعيد دفع الأجور ....

رقم مسلسل	العامل				الأجر				استقطاع غرامات ومواد تالفة	الأجر الاضافى		جملة أيام الشغل فى شهر مايو	جملة الكسب فى شهر مايو ١٩٣٨	ملاحظات
	جنسيته	جنسه	سنه	وظيفته	باليوم	بالأسبوع	بالشهر	بالقطعة		الساعات	الأجر			

المـــراجع

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

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

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

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

الباب الثالث

طرق عرض البيانات

توضيح البيانات الاحصائية وطرق عرضها

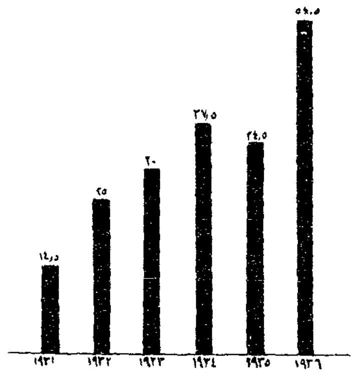
٣٥ — المهمة التي تلي عملية جمع البيانات التي تكلمنا عنها في الباب السابق هي ترتيب هذه البيانات وتنسيقها بطريقة تساعد على فهم مدلولها والاستفادة منها ، ثم عرضها وتوضيح معانيها بوسائل طريفة ومقبولة . وذلك لأن البيانات الإحصائية إذا سيقت في الصورة الرقمية الجافة ربما لا تشجع الشخص العادى على قراءتها والإقبال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

والوسائل التي نستخدمها لتوضيح هذه البيانات ، وكذلك طريقة عرضها ، تتوقف على نوع البيانات والغرض المقصود من هـذا الايضاح ، والحقائق التي نريد إبرازها بصفة خاصة ؛ وسنشرح في هذا الباب بعض هذه الطرق .

تمثيل الارقام هندسيا بواسطة خطوط ذات أطوال متناسبة

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

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



(شكل ١)

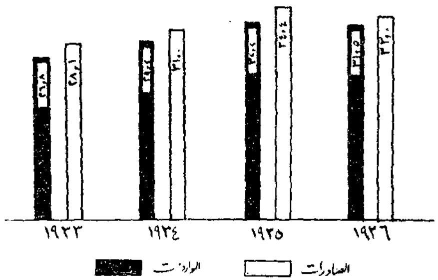
إنتاج المنسوجات فى مصر فى السنين ١٩٣١ — ١٩٣٦ بملايين الأمتار المربعة

أعمدة مزدوجة تمثل ظاهرتين فى شـكل واحـد

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

(١) الأرقام مأخوذة عن تقرير المستر سيلوس الملحق التجارى البريطانى فى مصر - انظر مجلة غرفة الاسكندرية التجارية ، عدد يونية ١٩٣٨ ص ٢٥ .

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



(شكل ٢)

يبين الصادرات والواردات المصرية بملايين الجنيهات

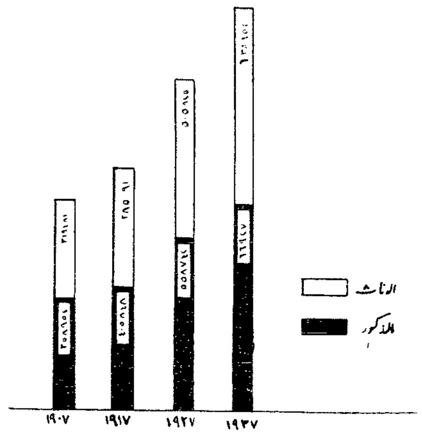
ولا نجد صعوبة فى رسم هذه الأعمدة إذا كانت وحدات القياس للظاهرتين متساوية ، كما فى هذه الحالة . فكل من الصادرات والواردات مقدرة قيمتها بملايين الجنيهات . أما إذا كانت الوحدات غير متساوية ، كأن تكون إحدى الظاهرتين محصول القطن مقدراً بآلاف القناطير ، والأخرى مساحة الأراضى المزروعة قطناً مقدرة بالأفدنة ،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

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

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

أعمدة مقسمة إلى أجزاء

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



(شكل ٣)

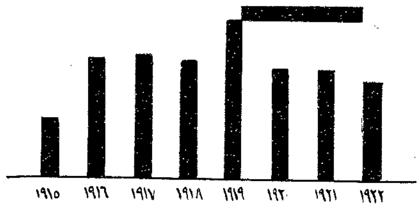
عدد سكان القاهرة من ذكور وإناث

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

يتكون كل عمود من أجزاء — مميزة عن بعضها بألوان مختلفة — كل منها يمثل رقماً من الأرقام الجزئية ؛ ومجموع هذه الأجزاء — وهو طول العمود الـكلى — يمثل رقم الجملة . ونرى (فى شكل ٣) تطبيق ذلك لتوضيح كيفية نمو عدد سكان القاهرة حسب التعدادات الأربعة الأخيرة : ١٩٠٧ و ١٩١٧ و١٩٢٧ و ١٩٣٧ .

كسر الأعمدة الطويـلة

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



(شكل ٤)

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

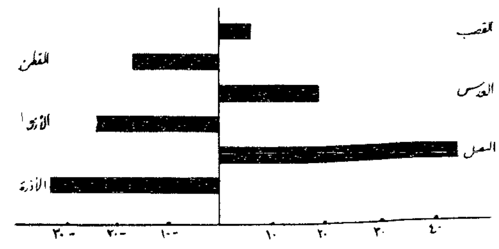
حركة أسعار القطن فى السنين ١٩١٥ — ١٩٢٢ بالنسبة إلى متوسط سعره فى سنة ١٩١٣ كأساس = ١٠٠ . [الأرقام مأخوذة من «الإحصاء السنوى العام» لسنة ١٩٣٥ — ١٩٣٦ صفحة ٥٠٤] .

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

خطوط أفقية

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

نرسم محورا رأسيا نبدأ منه القياس . ولـكل محصول نرسم خطا أفقيا يتناسب طوله مع مقدار الزيادة أو النقص (فى المائة) فى السعر . ونرسم كل الخطوط التى تمثل الزيادة على يمين المحور الرأسى والخطوط التى تمثل النقص على يساره . ويحسن أن نرسم فى أسفل الشـكل محورا أفقيا نبين عليه مقياس الرسم كما هو مبين فى شكل ٥ . [الأرقام مأخوذة من الاحصاء السنوى العام ١٩٣٥ — ١٩٣٦ ص ٥٠٤]



(شكل ٥)

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

المساحات قطاعات الدائرة

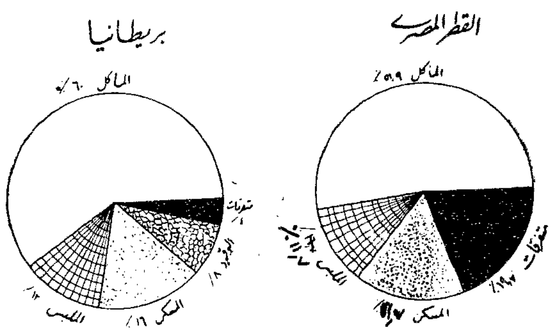
٤١ — يمكن أن نستخدم المســاحات بدل الخطوط أو الأعمدة لتمثيل البيانات . هنا تكون المســاحات متناسبة مع الأرقام التى تمثلها ، كما فى الخطوط أو الأعمدة .

وربما كانت الدائرة أبسط الأشكال وأحسنها بياناً فى حالة استعمال المساحات

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

فى هـذه الحالة نمثل الجملة العموميـة بالمساحة الـكلية للدائرة ، ونقسمها إلى قطاعات تتلاقى فى المركز بحيث تكون مساحاتها متناسبة مع المقادير الجزئية التى تـكون الجملة العموميـــة . وهذه القطاعات نميزها عن بعضها بألوان مختلفـة لزيادة الإيضاح .

خذ مثلا مصروفات الأسرة العادية وتوزيعها بين الأشياء المختلفة الضرورية للمعيشة . تبين من بحث عملته مصلحة الإحصاء المصرية فى سنة ١٩٢٠ أن الأسرة العـادية توزع مصروفاتها على الأبواب الرئيسية بنسبة ٥١٫٩ ٪ للأكل ، و١٦٫٧ ٪ للملابس ، و١١٫٧ ٪ للمسكن ، والباقى أى ١٩٫٧ ٪ للمتفرقات .



(شكل ٦)

تقسيم مصروفات نفقة المعيشة على الأبواب المختلفة فى مصر وبريطانيا

لتمثيل هذه البيانات بهذه الطريقة نرسم فى الدائرة أربع زوايا رأسها فى المركز بحيث تكون النسبة بين مقاديرها تساوى ٥١٫٩ : ١٦٫٧ : ١١٫٧ : ١٩٫٧

وبما أن مساحة قطاع الدائرة تتناسب مع زاوية رأسه ، ينتج أن النسبة بين مساحات القطاعات المرسومة بهذا الشكل تساوى نفس النسبة أى ٩ر٥١ : ٧ر١٦ : ٧ر١١ : ٧ر١٩ .

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

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

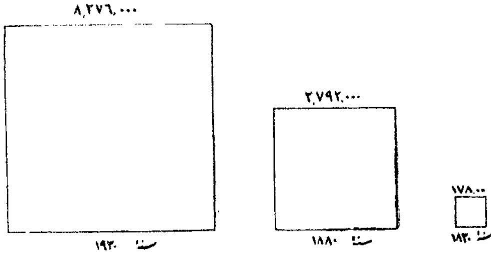
ونرى هذه القطاعات فى الشكل المرافق (رقم ٦) ، وهو يوضح هذا التقسيم فى مصر . وبجانبه أيضاً تقسيم المصروفات المعيشية فى بريطانيا حيث النسب هى ٦٠ ٪ للأكل و ١٢ ٪ للملابس و ١٦ ٪ لإيجار المسكن و ٨ ٪ للوقود والإنارة (وهذا باب جديد لا يوجد له نظير فى المصروفات المصرية ، وذلك نظراً لأهمية التدفئة فى انجلترا بسبب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

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

استخدام المربعات مساحاتها تتناسب مع مربعات الأضلاع

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

لنأخذ مثلا محصول القطن المصرى فى السنين ١٨٣٠ و ١٨٨٠ و ١٩٣٠ وهو ، على الترتيب ، ١٧٨٠٠٠ و ٢٧٩٢٠٠٠ و ٨٢٧٦٠٠٠ من القناطير[1] . ونرى فى شكل ٧ ثلاثة مربعات مساحاتها تتناسب مع هذه الأعداد ، أى أن الأضلاع تتناسب مع جذورها . وهذه الجذور بنسبة ١ : ٣ر٩٦ : ٦ر٧٨ تقريباً .



(شكل ٧)

مقدار محصول القطن المصرى بالقناطير فى سنى ١٨٣٠ و ١٨٨٠ و ١٩٣٠

استخدام الأشكال المجسمة لتمثيل البيانات

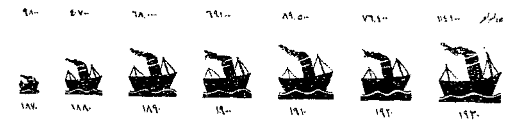
٤٣ — يمكننا أيضاً تمثيل البيانات الإحصائية بصورة مجسمة بواسطة استخدام الأشكال الهندسية المجسمة المعروفة ، مثل الكرة أو المكعب ؛ وهنا تكون النسبة بين أحجام الأجسام التى تمثل أعداداً معينة تساوى النسبة بين هذه الأعداد . ويلاحظ فى هذه الحالة أن حجم الجسم ، مثل المكعب أو الكرة ، يتناسب مع مكعب الضلع ، فى حالة المكعب ؛ أو نصف القطر ، فى حالة الكرة . فلو أردنا تمثيل الأرقام المذكورة فى البند السابق بواسطة ثلاثة مكعبات ، كانت أضلاعها متناسبة مع الجذور التكعيبية للأعداد ١٧٨٠٠٠ و ٢٧٩٢٠٠٠ و ٨٢٧٦٠٠٠ أى بنسبة ١ : ٢ر٥٠ : ٣ر٥٨ تقريباً . وكذلك لو أردنا تمثيلها

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

بكرات كانت النسبة بين أنصاف أقطارها تساوى هذه النسبة نفسها .

الصور والأشكال

٤٤ — يمكن أن نستبدل بهذه الأشكال الهندسية المسطحة أو المجسمة رسوماً أو صوراً معينة تكون لها دلالة خاصة ذات صلة متينة بالموضوع الذى نتكلم فيه . فلو حصلنا مثلا على أرقام لعدد البواخر التى مرت بقناة السويس فى عدة سنين متتالية ، وأردنا توضيح هذه البيانات بطريقة مشوقة ، يمكننا تمثيل الأرقام بصورة باخرة وتكبير هذه الصورة أو تصغيرها بنسبة الأرقام المختلفة التى لدينا . وكذلك إذا أردنا عمل مقارنة بين الممالك (أو التواريخ) المختلفة من حيث تعداد السكان أو قوة الجيش أو كمية الإنتاج أو الاستهلاك لسلعة معينة وهكذا ، نختار لكل من هذه الأشياء التى نقارنها رمزاً (مسطحاً أو مجسما ، حقيقياً أو خيالياً) يكون بسيطاً ما أمكن ، واضحاً ، سريع الدلالة على الشىء أو السلعة المقصودة بالمقارنة ؛ ثم نكبر أو نصغر هذا الرسم لكل مملكة (أو تاريخ) بقدر ما يناسب الرقم الخاص بهذه المملكة (أو التاريخ) .



(شكل ٨)

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

وتحديد المقاسات لكى تتناسب مع الأرقام التى تمثلها بهذه الرسوم والصور يكون بنفس القواعد السابق ذكرها فى البندين السابقين على حسب كون الرمز أو الصورة يدل على شكل مسطح أو مجسم . فاذا كان الأول أخذت النسب كما فى حالة المساحات المذكورة فى بند ٤٢ . وإذا كان الثانى اتبعت قاعدة تناسب الحجوم (بند ٤٣) .

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

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

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

تبويب البيانات وعمل الجداول

التقسيم إلى مجموعات متجانسة

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

ويراعى فى هـذا التبويب أن تشمل المجموعة الواحدة كل المفردات المتحدة فى صفة معينة من الصفات المهمة فى الموضوع ، أو عـدة صفات مرتبطة ببعضهـا . ولا بأس من تقسيم هـذه المجموعات الرئيسية إلى فروع ، وهـذه إلى أقسام إذا اقتضى الحال ، وهكذا .

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

وعلى كل حال فطريقة التبويب وتعيين الصفات أو الخواص التى تتخـذ أساساً لهـذا التبويب لا بد تتوقف على الغرض المقصود من عمل الإحصـاء ، والتفاصيل التى نحصل عليها عند جمع البيانات اللازمة .

عمل الجداول العـادية

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

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

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

(جدول ١) عدد المصانع بمدينة الإسكندرية فى سنة ١٩٢٧

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

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

والقاعدة العامة فى تصميم الجداول هى أن ننظر إلى تقسيم البيانات من

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

فرز البيانات عملية أساسية لعمل الجداول

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

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

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

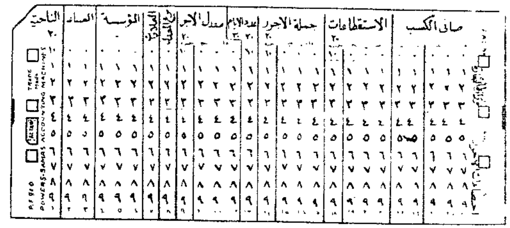
ولتسهيل عملية العدّ يحسن ، عند وضع الإشارات فى الخانات المناسبة ، أن نجعل كل خمس منها بجوار بعضها كوحدة للعدّ مفصولة عن غيرها من الوحدات بفراغ بسيط (مثل /// ///// ///// /////) . ويكون عدد الإشارات فى الخانة يساوى عدد هذه الوحدات مضروباً فى ٥ مضافاً إليه الباقى . والأحسن أن نشطب على كل أربعة خطوط أو شرط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

الوسائل الآلية للتبويب

طريقة البطاقات المثقوبة

٤٩ — قلنا إن هذه الطريقة اليدوية لا يمكن استخدامها فى الإحصاءات الكبيرة ، ولا بد من الاستعانة بالوسائل الآلية لتسهيلها . وأحسن الآلات المستخدمة لهذا الغرض هى الآلات المبنيـــة على نظام البطاقات المثقوبة (Punched-Card System) . ويوجد شركتان تصنعان هذه الآلات فى الوقت الحاضر ، وهما «هولريث» و «بورز — ساماس»[(١)] ؛ وهما تحتكران هذه الآلات فى جميع بلاد العالم .

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



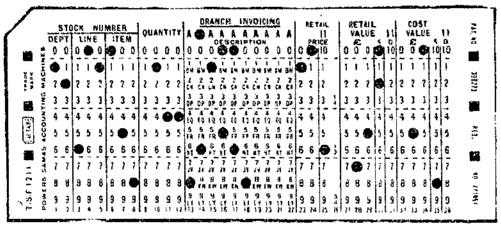
(شكل ٩)
بطاقة من ٢٦ عموداً مستعملة لإحصاء الأجور

تثقيب البطاقات حسب البيانات

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

(١) آلات (Powers-Samas, Hollerith) . الأشكال ٩ — ١٣ معارة من شركة بورز — مع الشكر .

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



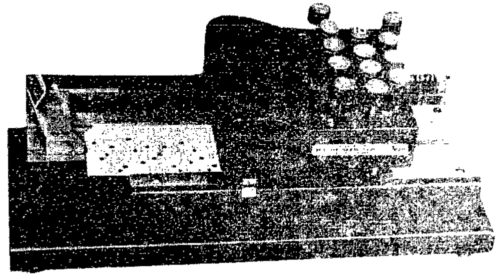
(شكل ١٠)
بطاقة مثقوبة معدة للفرز أو التبويب

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

الفرز الآلى

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

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

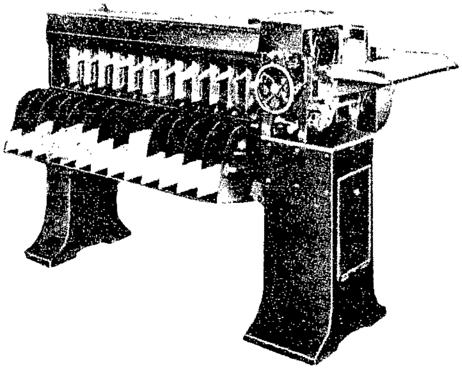


(شكل ١١)
آلة لتثقيب البطاقات

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

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

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



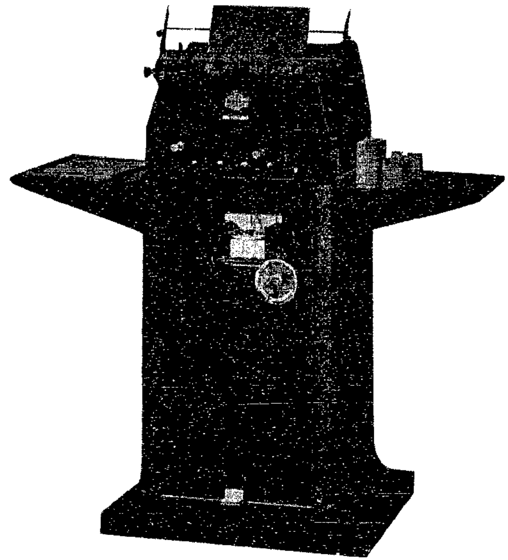
(شكل ١٢)
آلة كهربائية لفرز البطاقات

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

آلات لعمل الجداول والتبويب

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

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



(شكل ١٣)

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

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

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

وفى شكل ١٣ صورة لواحدة من هذه الآلات . وهى ، طبعاً ، فى غاية التعقيد والإتقان ، وأقل ما توصف به أنها قطعة ثمينة من ثمرات التقدم العلمى فى هذا العصر . وسرعتها فى الكتابة تبلغ حوالى ١٠٠ سطر فى الدقيقة .

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

المراجع

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

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

Connor, R. L., *Statistics in Theory and Practice*, Chapters V, VI

Secrist, H., *Statistical Methods*, Chapters VI, VII.

الباب الرابع

الرســـوم البيانية

تعريف الخط البيانى

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

وهذه الخطوط البيانية مستعملة كثيراً فى جميع العلوم ، خصوصا التى تبحث فى الظواهر عن طريق مشاهدتها وتقديرها رقميا . ولذلك سنشرح باختصـار طريقة رسمها وخواصها .

رسم الخطوط البيانية

٥٣ — لنأخذ مثلا كمية الانتاج المحلى فى مصر من المنسوجات القطنية فى السنين ١٩٣١ — ١٩٣٧ ، مقدرا بملايين الأمتار المربعة . وها هى الأرقام [١] :

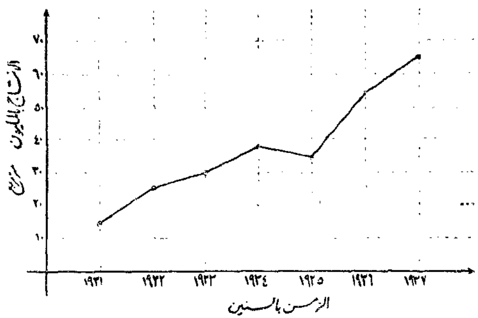
الســنين . . .	١٩٣١	١٩٣٢	١٩٣٣	١٩٣٤	١٩٣٥	١٩٣٦	١٩٣٧
الإنتـاج . . .	١٤٫٥	٢٥٫٠	٣٠٫٠	٣٧٫٥	٣٤٫٥	٥٤٫٥	٦٥٫٠٠

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

(١) مأخوذة عن تقرير الملحق البريطانى المذكور سابقا (رقم ١٩٣٧ تقريبى)

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

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



(شكل ١٤)

الانتاج المحلى من النسيج فى مصر فى السنين ١٩٣١ — ١٩٣٧

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

الخط الذى يصـل بين هذه النقط هو الخط البيـانى للانتاج فى هذه المدة . وهو يوضح كيفية تغير الإنتاج مع الزمن فى أثناء هذه المدة . ويكفى إلقاء نظرة

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

الخط البيانى المستوى يمثل العلاقة بين ظاهرتين فقط

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

يراعى أن يكون الشكل مقبولا

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

مقياس الرسم على كل محور يناسب أرقام الظاهرة التى تقاس عليه

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

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

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

مقياس الرسم يؤثر فى مظهر تغيرات الظاهرة

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

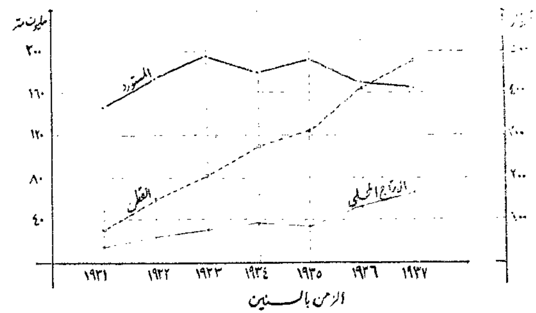
المقياس الكبير يبالغ فى شدة التغيرات والمقياس الصغير يضعف من أهميتها

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

رسم أكثر من خط بياني واحد في شكل

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

يمكن رسم خطين بيانيين أو أكثر في نفس الشكل ، كل منهـا يوضح العلاقة بين ظاهرة من هــذه ، والظاهرة المشتركة — وهي الزمن بالسنين في هذا المثـال .



(شكل ١٥)

الانتاج المحلي والمستورد من المنسوجات والقطن المستهلك محلياً

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

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

ونرى (في شكل ١٥) ثلاثة خطوط بيانية تصور في نفس الرسم كميـة الإنتـاج المحلي من المنسوجات القطنية ، وكمية المستورد من الخارج ، وكمية القطن المستهلك محلياً للغزل ، في المـــدة ١٩٣١ — ١٩٣٧ ؛ والأرقام هي كالآتي[1] : —

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

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

(١) عن تقرير الملحق التجاري البريطاني (أرقام ١٩٣٧ تقريبية) .

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

استعمال ورق المربعات للرسم البيانى العادى

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

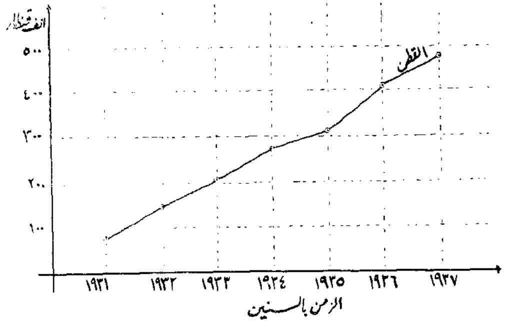
خط بيانى لقيم متغير مع لوغاريتمات قيم متغير آخر

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

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

الورق اللوغاريتمى

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

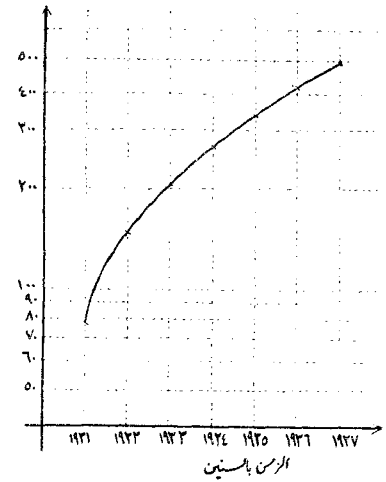


(شكل ١٦)

المستهلك من القطن فى المصانع المحلية فى مصر . تقسيم ورق عادى

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

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



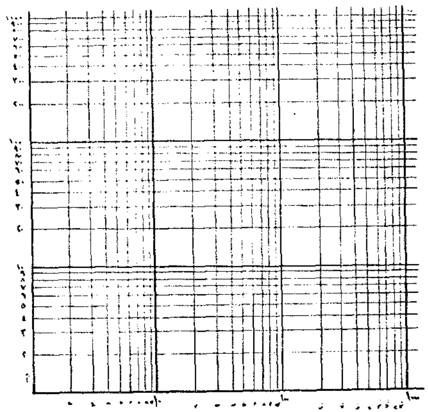
(شكل ١٧

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

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

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

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



(شكل ١٨)

تقسيم لوغاريثمى مزدوج

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

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

الخطوط البيانية تكون فى العـادة كثيـرة التعـاريج

٦١ — هذه الخطوط البيانية التى نرسمها — بناء على بيانات مأخوذة من التجارب العملية — تكون فى العادة كثيرة التذبذب وغير « ممهدة » ؛ وذلك لأن مشاهدات التجارب العملية كثيراً ما تكون معرضة لتقلبات المصادفة . وهذه التقلبات لا يمكن التغلب عليها لأنها ، بطبيعتها ، عرضية ولا يمكن التنبؤ بها . فمثلا قد يخطئ الباحث — بدون قصد طبعاً — فى قراءة مقياس الحرارة أو الضغط ، أو قد يخطئ الجهاز نفسه فى حالة ما لسبب غير معروف ، أو يطرأ ظرف طارئ على الظاهرة التى نبحثها — فيزيد مقدارها زيادة خارقة للقانون ، أو ينقص ، وهكذا . هذه التقلبات الناشئة عن مجرد المصادفة تكون « عشوائية »(١) أى بدون ضابط . فتارة تكون بالزيادة وأخرى تكون بالنقص ، وغير مقصودة فى أى اتجاه دون الآخر . وهى تسمى على العموم **أخطاء التجربة** ، أى (Experimental Errors) ؛ وأحياناً نسميها **تقلبات المصادفة** (Random Fluctuations) و **تغيرات عرضية أو فجائية** (Casual or Accidental Variations)

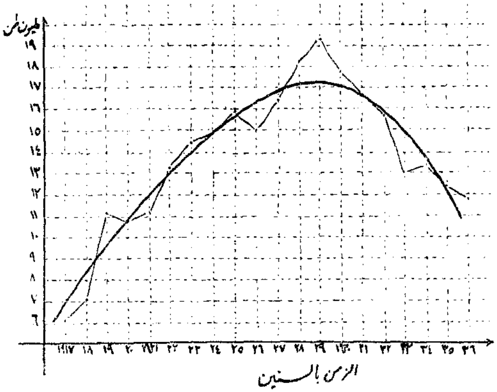
---

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

تمهيد الخطوط البيانية باليد

٦٢ — وللتخلص من هذه التذبذبات « نمهد » الخط البيانى الذى نحصل بواسطة البيانات الأصلية المأخوذة من التجربة والمشاهدات ؛ وذلك بأن نغض النظر عن هذه التذبذبات العرضية ، ونرسم خطاً ممهداً يتمشى مع الخط الأصلى ويهمل التفاصيل والتذبذبات الصغيرة .

وهذا يمكن عمله بالنظر بدون صعوبة . فبعد رصد النقط فى الشكل — بناء على الأرقام المعطاة لنا — نرسم باليد خطاً يتوسط بين هذه النقط بدون أن يمر بها جميعاً ، إذا كان المرور بها يسبب تعريجات فى الخط تفسد تمهيده بدون لزوم .



(شكل ١٩)

صافى حمولة السفن الانجليزية المارة بقناة السويس

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

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

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

معادلة الخط البيانى

وضع العلاقة فى صورة جبرية

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

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

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

ويمكننا وضع هذه العلاقة فى صورة مختصرة جدا وواضحة باستعمال بعض الرموز الرياضية البسيطة هكذا :

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

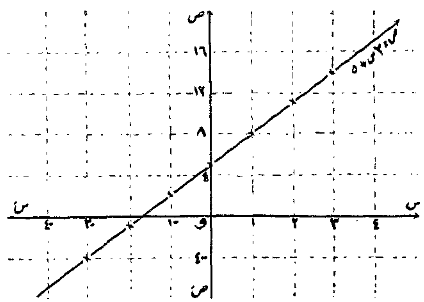
فاذا وضعنا

ص = ٣ س + ٥ . . . . . (١)

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

رسم الخط البيانى للعلاقة الجبرية

٦٤ — لكى نصوّر هذه العلاقة فى شكل خط بيانى نفرض عدة قيم مختلفة يأخذها المتغير المستقل س ؛ ولكل منها نحسب القيمة المناظرة التى يأخذها المتغير التابع ص بحكم هذه العلاقة المفروضة بينهما . ثم نضع هذه القيم المتناظرة فى



(شكل ٢٠)
الخط البيانى للعلاقة ص = ٣ س + ٥

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

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

س : −٣ ، −٢ ، −١ ، ٠ ، ١ ، ٢ ، ٣

ص : −٤ ، −١ ، ٢ ، ٥ ، ٨ ، ١١ ، ١٤ .

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

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

نحصل على نفس الخط إذا فرضنا قيما أخرى للتغير س

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

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

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

الخط البيانى هو مسار نقطة تتحرك بشرط معـــين

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

معادلة المستقيم

كل معادلة من الدرجة الأولى يمثلها خط مستقيم

٦٦ — نلاحظ أيضاً أن الخط المرســـوم فى (شكل ٢٠) خط مستقيم ، وأن معادلة هذا الخط ، وهى ص = ٣ س + ٥ ، من الدرجة الأولى بالنسبة إلى س و ص ، أى أنها لا تحتوى على قوى أعلى من الأولى لأى واحد من المتغيرين ، فلا نرى فيها حدوداً تحتوى على س^{٢} أو س^{٣} ... ولا ص^{٢} أو ص^{٣} ... وهذا التوافق ، بين صفة استقامة الخط البيانى وكون معادلته من الدرجة الأولى ، ليس مصادفة فى هذه المسألة فقط . ويمكن إثبات أن كل معادلة من الدرجة الأولى يمثلها خط مستقيم ، وأن كل خط مستقيم تمثله معادلة من الدرجة الأولى . ولكن المقام هنا لا يسمح بإيراد هذا البرهان ، ويجب أن يرجع القارىء إلى أحد الكتب فى الهندسة التحليلية [1] .

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

الصورة العامة لمعادلة المستقيم

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

ص = م س + ح

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

م = ٣ و ح = ٥ .

أى أن ميل هذا المستقيم على محور السينات يساوى ٣ ؛ وهو يساوى ظل الزاوية التى يصنعها هذا المستقيم مع محور السينات [١] ؛ وأن الكمية ح = ٥ تساوى الجزء الذى يقطعه هذا المستقيم من محور الصادات كما هو واضح فى الشكل ، حيث نجد المستقيم يقطع محور ص عند النقطة . ٥ .

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

ص = — ٥ س + ٤ .

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

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

معادلة القطع المكافئ

منحنى القطع المكافئ.

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

المطلوب رسم الخط البيانى الذى معادلته

ص = س^{٢} + ٢

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

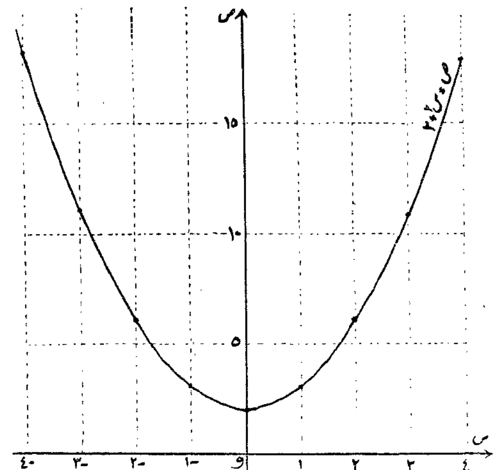
س : — ٤ ، — ٣ ، — ٢ ، — ١ ، ٠ ، ١ ، ٢ ، ٣ ، ٤ ،

ص : ١٨ ، ١١ ، ٦ ، ٣ ، ٢ ، ٣ ، ٦ ، ١١ ، ١٨ ،

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

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

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

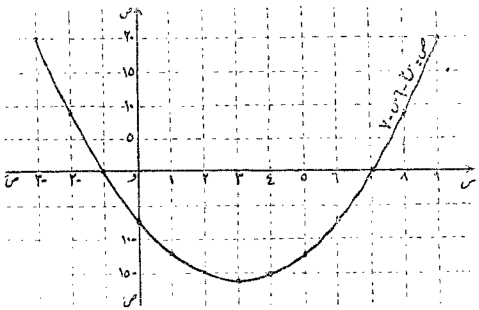


(شكل ٢١)
قطع مكافئ متماثل

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

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

وقد رأينا القطع المكافئ فى هذين الشكلين مقعراً إلى أعلى ورأسه فى أسفل . وذلك لأن معامل س^{٢} فى المعادلة موجب فى كلتا الحالتين . أما إذا كان معامل س^{٢} فى معادلة القطع المكافئ سالبا فإنه يظهر فى الشكل مقلوبا ، ويكون محدبا إلى أعلى وتكون رأسه أعلى نقطة فيه . وهذه خواص ثابتة لهذا المنحنى ،



(شكل ٢٢)
قطع مكافئ محور تماثله يوازى محور ص

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

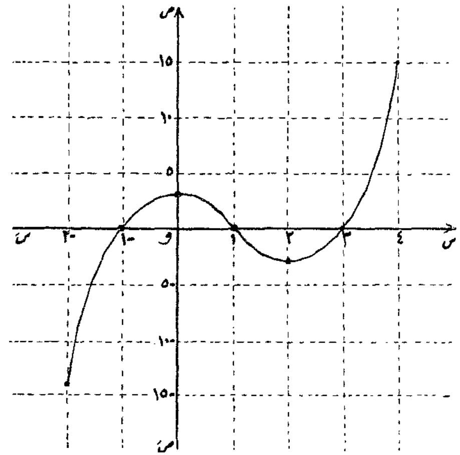
منحنى الدرجة الثالثة

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

المطلوب رسم المنحنى الذى معادلته هى :

$$ص = س^{3} - 3\,س^{2} - س + 3$$

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



(شكل ٢٣)

منحنى معادلة من الدرجة الثالثة

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

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

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

فى هذا المنحنى نجد النهاية العظمى أولا وتليها النهاية الصغرى إذا اتجهنا مع المنحنى فى الاتجاه سَ وس ، أى الاتجاه الذى تزيد فيه س ؛ والعكس يحصل إذا كانت إشارة $س^{3}$ سالبة فى معادلة المنحنى (أى مخالفة لإشارة ص) .

عدد نقط الرجوع فى أى منحنى

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

منحنى الخطأ المتماثل

٧١ — المطلوب رسم المنحنى الذى معادلته هى :

$$ص = ه^{-س^{2}}$$

حيث ه هى الكمية ٢٫٧١٨٢٨١٨٢٨٤ (وهى أساس اللوغاريتمات الطبيعية المستعملة فى الرياضة) [1] .

---

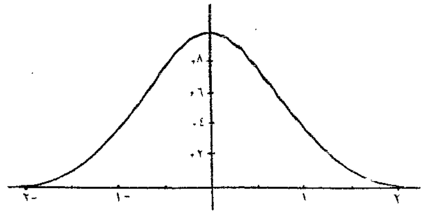
(١) هذه الكمية ه مشهورة جداً فى العلوم الرياضية :

$$ه = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \cdots \text{ إلى } \infty .$$

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

نفرض عدة قيم مناسبة للمتغير س ، ونحسب لكل منها القيمة المناظرة للمتغير ص . ويلاحظ أن س هنا مربعة وموجودة فى الأس ترفع إليها الكمية هـ . ويوجد جداول[١] للقوى المختلفة لهذه الكمية نستخرج منها القيم المطلوبة فنحصل على الجدول الآتى ، مع العلم بأن قيمة ص لا تتغير بتغير إشارة س :

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



(شكل ٢٤)

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

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

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

وهذا المنحنى مشهور جداً فى علم الإحصاء ؛ وهو أساسى جداً فى دراسة النظريات التى ينبنى عليها هذا العلم . وهذا المنحنى معروف بجملة أسماء نذكر منها «منحنى جاوس» (Gaussian Curve) ، نسبة إلى العالم الألمانى كارل ف. جاوس الذى استنبطه رياضيا ودرس خواصه . ويسمى أيضاً « منحنى الخطأ المتماثل » (Symmetrical Curve of Error) ، و « منحنى التكرار المتماثل » (Normal Frequency Curve) ، وهكذا . وسنعود إلى شرح بعض خواصه فى مناسبات أخرى .

ولو رسمنا المنحنى الذى معادلته

$$\text{ص} = \text{هـ}^{-(\text{س} - ٣)^{٢}}$$

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

المنحنى التكرارى غير المتماثل

٧٢ — المطلوب رسم المنحنى الذى معادلته هى

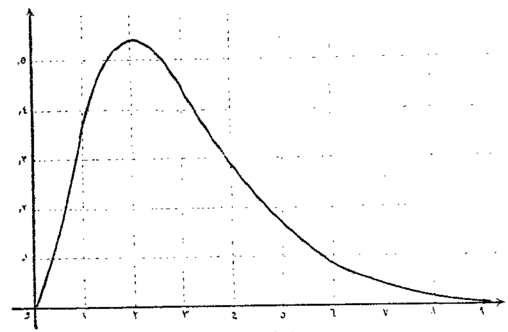
$$\text{ص} = \text{س}^{٢} . \text{هـ}^{-\text{س}}$$

ابتداء من س = ٠ إلى س = ∞ ، مع العلم بأن هـ لها نفس المعنى المذكور فى البند السابق .

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

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

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



(شكل ٢٥)

منحنى تكرارى غير متماثل

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

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

توفيق المنحنيات

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

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

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

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

(١) الكلمة الانجليزية هى (Curve Fitting)

يوجد طريقتان لتوفيق المنحنيات وهما[1] : طريقة المربعات الصغرى وطريقة العزوم . والطريقة الأولى هى الأسهل من الناحية العملية والنظرية ؛ وسنكتفى هنا بشرح هذه الطريقة . ويجب على القارئ أن يرجع إلى بعض الكتب الرياضية لمعرفة الطريقة الثانية[2] .

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

س$_1$ ، س$_2$ ، س$_3$ ، . . . . س$_ن$

و س$_1$ ، ص$_2$ ، ص$_3$ ، . . . ص$_ن$

نرسم محورين متعامدين ونرصد فى الشكل النقط التى إحداثياتها :

(س$_1$ ، ص$_1$) ، (س$_2$ ، ص$_2$) ، . . . . ، (س$_ن$ ، ص$_ن$)

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

لنفرض أن معادلة[3] هذا المستقيم على العموم هى :

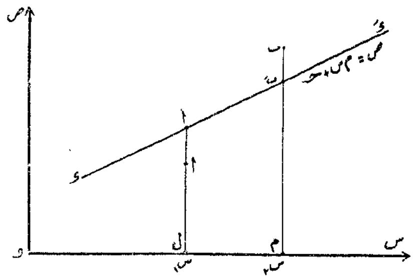
ص = م س + حـ . . . (١) ،

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

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

(٣) هذه المعادلة التى فرضنا، بناء على رأينا وخبرتنا بأنها توافق البيانات التى حصلنا عليها بالتجربة والمشاهدة ، نسميها « المعادلة الاعتبارية » (Empirical Equation)

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



(شكل ٢٦)

لنفرض أن النقطة ا (س$_1$ ، ص$_1$) مثلا هى إحدى هذه النقط فى الشكل ، وأن المستقيم ء ءَ هو المستقيم الذى معادلته ص = م س + حـ : ولنفرض أن هذه النقطة وقعت عفواً ، لسبب طرأ فى أثناء التجارب التى أجريت ، بعيدة نوعا عن هذا المستقيم . نرسم الخط الرأسى ل ا اَ مارا بالنقطة ا يقابل المستقيم فى النقطة اَ . فيكون البعد ا اَ هو مقدار « انحراف » هذه النقطة عن الوضع الذى كان يجب أن تكون فيه لو أنها تمشت تماما مع هذا القانون الاعتبارى الذى تمثله المعادلة (١) . ويصح أن نسميه أيضا « الخطأ التجريبى » .

نرى من الشكل أن الانحراف

ا اَ = ل اَ — ل ا

= ل اَ — ص$_1$

حيث ل اَ هو الإحداثي الرأسي للنقطة اَ الواقعة على الخط . والإحداثي الأفقي لهذه النقطة اَ يساوي الإحداثي الأفقي للنقطة ا أي يساوي س_1 . وبما أن النقطة اَ واقعة على المستقيم الذي معادلته $\text{ص} = \text{م س} + \text{حـ}$ ، فلا بد أن إحداثييها الرأسي والأفقي يحققان هذه المعادلة ؛ وبذلك يكون :

$$\text{ل اَ} = \text{م س}_1 + \text{حـ}$$

$$\therefore \text{الانحراف ا اَ} = \text{م س}_1 + \text{حـ} - \text{ص}_1 \text{ ، } \ldots \quad (٢).$$

وبالمثل إذا كانت ب $(\text{س}_2 \text{ ، ص}_2)$ نقطة أخرى في الشكل ، فإن انحرافها عن المستقيم يساوي البعد ب بَ ؛ ويكون

$$\text{ب بَ} = \text{م س}_2 + \text{حـ} - \text{ص}_2 \text{ .}$$

وهكذا مع جميع النقط الأخرى .

وطبيعي أن الخط $\text{ص} = \text{م س} + \text{حـ}$ يكون أوفق ما يمكن لتمثيل هذه النقط كلما كانت هذه الأخطاء أو الانحرافات صغيرة في المقدار ، بصرف النظر عن كونها موجبة أو سالبة . وبعبارة أخرى نقول إن هذا المستقيم يكون أفضل ما يمكن إذا كان مجموع مربعات هذه الأخطاء أصغر ما يمكن[١] ؛ أي أن المقدار

$$(\text{م س}_1 + \text{حـ} - \text{ص}_1)^2 + (\text{م س}_2 + \text{حـ} - \text{ص}_2)^2 + \ldots + (\text{م س}_{\text{ن}} + \text{حـ} - \text{ص}_{\text{ن}})^2$$
$$= \text{مجـ}(\text{م س} + \text{حـ} - \text{ص})^2$$

يكون أصغر ما يمكن .

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

ولكي يكون المقدار $\text{مجـ}(\text{م س} + \text{حـ} - \text{ص})^2$ أصغر ما يمكن ، يجب أن يكون مجموع الانحرافات نفسها يساوي صفراً ؛ وفي الوقت نفسه يجب أن يكون مجموع حواصل ضرب هذه الانحرافات ، كل في قيمة الإحداثي الأفقي للنقطة ، يساوي صفراً أيضاً[١] . أي أن :

$$(\text{م س}_1 + \text{حـ} - \text{ص}_1) + (\text{م س}_2 + \text{حـ} - \text{ص}_2) + \ldots + (\text{م س}_{\text{ن}} + \text{حـ} - \text{ص}_{\text{ن}}) = ٠ .$$

$$\text{و } \text{س}_1(\text{م س}_1 + \text{حـ} - \text{ص}_1) + \text{س}_2(\text{م س}_2 + \text{حـ} - \text{ص}_2) + \ldots + \text{س}_{\text{ن}}(\text{م س}_{\text{ن}} + \text{حـ} - \text{ص}_{\text{ن}}) = ٠ .$$

أو بالاختصار :

$$\text{م . مجـ س} + \text{ن حـ} = \text{مجـ ص} \quad \ldots \quad (٣) \text{ ،}$$

$$\text{و} \quad \text{م . مجـ س}^2 + \text{حـ . مجـ س} = \text{مجـ س ص} \quad \ldots \quad (٤) \text{ .}$$

وبما أن جميع قيم س وقيم ص معلومة ، وهي القيم المشاهدة في التجربة ، وكذلك ن ، يكون المجهول في هاتين المعادلتين الآنيتين (٣) و (٤) هما قيمتا م و حـ فقط . فنحل المعادلتين ونستخرج قيمتي م و حـ ونعوضهما في المعادلة (١) نحصل على معادلة المستقيم المطلوب .

مثال توفيق مستقيم

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

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

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

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

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

$$\text{مجـ ص} = \text{م . مجـ س} + \text{ن حـ}$$

$$\text{مجـ س ص} = \text{م . مجـ س}^2 + \text{حـ . مجـ س}$$

ولتسهيل العمل فى إيجاد قيم مجـ س ، مجـ ص ، مجـ س^2 ، مجـ س ص ، نرتب العمل فى جدول كالآتى :

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

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

∴ ٣٥٧ = ١١٠ م + ١١ حـ . . . . (١) ؛

و ٣٨٢٠ = ١٢١٠ م + ١١٠ حـ . . . . (٢) .

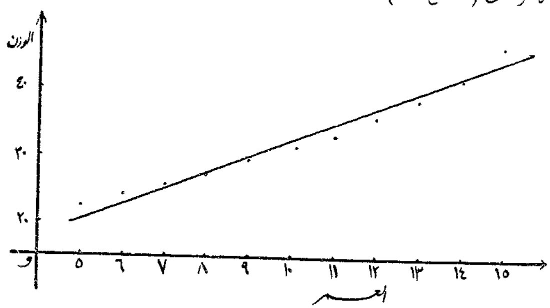
∴ م = $\frac{٣}{١١}$ ٢ = ٢ر٢٧٣ ؛

و حـ = $\frac{١٠٧}{١١}$ = ٩ر٧٢٧ .

∴ معادلة المستقيم المطلوب هي :

ص = ٢ر٢٧٣ س + ٩ر٧٢٧ . . . . (٣) ،

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



(شكل ٢٧)

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

العمر	س	ص	س^2	س . ص
٥	−٥	٢٢ر٥٠	٢٥	−١١٢ر٥٠
٦	−٤	٢٤ر٥٠	١٦	−٩٨ر٠٠
٧	−٣	٢٥ر٧٥	٩	−٧٧ر٢٥
٨	−٢	٢٧ر٢٥	٤	−٥٤ر٥٠
٩	−١	٢٩ر٥٠	١	−٢٩ر٥٠
١٠	٠	٣١ر٢٥	٠	٠
١١	١	٣٣ر٢٥	١	٣٣ر٢٥
١٢	٢	٣٥ر٧٥	٤	٧١ر٥٠
١٣	٣	٣٨ر٧٥	٩	١١٦ر٢٥
١٤	٤	٤١ر٧٥	١٦	١٦٧ر٠٠
١٥	٥	٤٦ر٧٥	٢٥	٢٣٣ر٧٥
المجموع	٠	٣٥٧ر٠٠	١١٠	٢٥٠ر٠٠

وبالتعويض في المعادلتين الآتيتين :

مج ص = م . مج س + ن . حـ ،

و مج س ص = م . مج س^2 + حـ . مج س ،

∴ ٣٥٧ = . + ١١ حـ ،

و ٢٥٠ = ١١٠ م + .

∴ م = ٢ر٢٧٣ ،

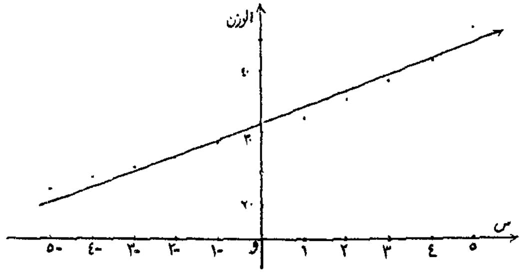
و حـ = ٣٢ر٤٥٥ .

∴ معادلة المستقيم في هذه الحالة هي :

ص = ٢ر٢٧٣ س + ٣٢ر٤٥٥ ،

حيث ص هي وزن الطفل بالكيلوجرام ، و س تدل على الفرق ، مقدراً

بالسنين ، بين العمر الحقيقي والعمر ١٠ سنوات . ونرى المستقيم مرسوماً في (شكل ٢٨) ويلاحظ أن ميل المستقيم واحد في هذين الشكلين . والفرق بين الشكلين هو أن المحور الرأسي نقل من مكانه فقط ، فتغير طول الجزء الذي يقطعه المستقيم من هذا المحور[1].



(شكل ٢٨)

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

ص = ا س^2 + ب س + حـ   . . . (١) ،

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

(س$_1$ ، ص$_1$) ، (س$_2$ ، ص$_2$) ، . . . ، (س$_ن$ ، ص$_ن$) .

---

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

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

(ا س$_{١}^{٢}$ + ب س$_{١}$ + حـ − ص$_{١}$)^{٢} + (ا س$_{٢}^{٢}$ + ب س$_{٢}$ + حـ − ص$_{٢}$)^{٢}

+ ... = مجـ (ا س^{٢} + ب س + حـ − ص)^{٢}

ولكى يكون هذا المجموع نهاية صغرى ، يجب أن يكون :

مجـ ص = ا . مجـ س^{٢} + ب . مجـ س + ن حـ . . . (٢)،

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

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

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

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

س : ٣ ٥ ٧ ٩ ١١ ١٣ ١٥ ١٧ ١٩

ص : ١٢ ١٦ ١٩ ٢١ ٢٣ ٢٠ ١٨ ١٥ ١٠

ولتكن المعادلة المطلوبة هى :

ص = ا س^{٢} + ب س + حـ (١)

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

س	سَ	ص	سَ^{٢}	سَ^{٣}	سَ^{٤}	سَ ص	سَ^{٢} ص
٣	−٨	١٢	٦٤	−٥١٢	٤٠٩٦	−٩٦	٧٦٨
٥	−٦	١٦	٣٦	−٢١٦	١٢٩٦	−٩٦	٥٧٦
٧	−٤	١٩	١٦	−٦٤	٢٥٦	−٧٦	٣٠٤
٩	−٢	٢١	٤	−٨	١٦	−٤٢	٨٤
١١	٠	٢٣	٠	٠	٠	٠	٠
١٣	٢	٢٠	٤	٨	١٦	٤٠	٨٠
١٥	٤	١٨	١٦	٦٤	٢٥٦	٧٢	٢٨٨
١٧	٦	١٥	٣٦	٢١٦	١٢٩٦	٩٠	٥٤٠
١٩	٨	١٠	٦٤	٥١٢	٤٠٩٦	٨٠	٦٤٠
المجموع	٠	١٥٤	٢٤٠	٠	١١٣٣٨	−٢٨	٣٢٨٠

وبالتعويض فى المعادلات

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

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

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

∴ ١٥٤ = ٢٤٠ ا + ٠ + ٩ حـ (٢) ،

و − ٢٨ = ٠ + ٢٤٠ ب + ٠ (٣) ،

و ٣٢٨٠ = ١١٣٣٨ ا + ٠ + ٢٤٠ حـ (٤) .

∴ ب = − ١١٦٧و٠

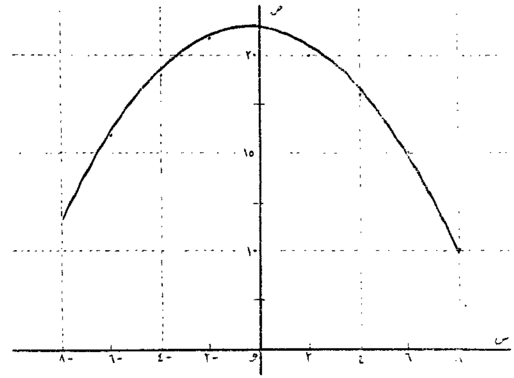
ومن المعادلتين (٢) و (٤)

∴ ا = − ١٦٧٧و٠ ، حـ = ٢١٫٥٨٣

وتكون معادلة المنحنى المطلوب هى :

ص = − ١٦٧٧ر. س^{٢} − ١١٦٧ر س + ٢١ر٥٨٣ .

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



(شكل ٢٩)

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

ص = ا س^{٣} + ب س^{٢} + حـ س + و

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

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

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

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

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

ولكن العمل بهذه الطريقة فى توفيق المنحنيات ذات الدرجات العليا

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

المراجع

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

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

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

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

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

BOWLEY, A.L., *Elements of Statistics*, Chapter VII.

KARSTEN, *Charts and Graphs*.

MILLS, F.C. *Statistical Methods*, Chapter II.

RIETZ, H. *Handbook of Mathematical Statistics*, Chapter IV.

WHITTAKER AND ROBINSON, *Calculus of Observations*, Chapter IX.

الباب الخامس

التوزيع التكرارى

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

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

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

ويصح أن نرتب مجموعة القيم التى نحصل عليها ترتيباً تصاعدياً أو تنازلياً تمهيداً لدراسة هذه الكمية المتغيرة . ولكن هذا الترتيب لن يساعدنا كثيراً خصوصاً إذا كان عدد القيم كبيراً حيث يتعذر على الإنسان أن يحصر ذهنه فى

مجموعة كبيرة من الأرقام المختلفة ولا يمكنه أن يستوعبها ليستنبط منها خواص المجموعة واتجاهاتها .

تقسيم مجموعة القيم إلى فئات

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

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

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

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

توزيع المفردات على الفئات

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

جدول ٣ — توزيع أطوال مجموعة من الرجال

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

نفرض أن مفردات الفئة تساوى مركزها

٨٣ — بعد توزيع المفردات على الفئات بهذا الشكل تضيع معالم القيم الأصلية التى وضعت المفردات فى الفئات بمقتضاها ؛ وحينئذ لا نعرف شيئاً عن أى مفردة سوى أنها واحدة فى فئة معينة محدودة بحدين معينين هما ١٦٠ سم و١٧٠ سم مثلا . وهذه المفردة تساوى جميع المفردات الأخرى التى معها فى نفس الفئة — وعددها جميعاً ٣١ فى هذا المثال . ولا يمكننا من واقع هذا الجدول أن نحدد طول أى شخص موجود فى أى فئة مثل الفئة ١٦٠ — ١٧٠ . ولا نعلم إذن إذا ما كان طوله قريباً من ١٦٠ سم أو هو أقرب إلى ١٧٠ سم . ولذلك نفرض عملياً أن جميع الأشخاص فى هذه الفئة طولهم = ١٦٥ سم ، أى يساوى مركز الفئة وهو منتصف المدى بين الحدين الأعلى والأدنى . ومع أن هذا الفرض ربما يحتاج إلى تبرير ، إذ أن الحكم فى هذه المسألة لابد يتوقف على كيفية توزيع المفردات عموماً ، ولكنه بالرغم من ذلك فرض مناسب وعملى ولا يبعد كثيراً عن الصواب فى أغلب الأحيان . أو قل إن الخطأ الناشىء عن هذا الفرض لا يؤثر كثيراً فى النتيجة التى نحصل عليها ، خصوصاً عندما يكون مدى الفئة قصيراً .

الجدول التكرارى . التوزيع التكرارى .

٨٤ — وعلى ذلك نجد فى هذه الفئة ٣١ شخصاً طولهم جميعاً ١٦٥ سم . أى أن الطول ١٦٥ سم « تكرر » ٣١ مرة فى هذه المجموعة . والجدول حينئذ يعطينا الأطوال المختلفة — مراكز الفئات — وعدد مرات « تكرار » كل واحد منها فى هذه المجموعة . أو ، بالاختصار ، يعطى الأطوال وتكراراتها . ولذلك نسمى هذا الجدول **الجدول التكرارى** [1] وهو يبين ما نسميه **التوزيع التكرارى** [2] لأطوال هذه المجموعة . وهذه الفئات التى تنقسم إليها المجموعة نسميها **فئات تكرارية** [3] .

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

تعيين عدد الفئات ومواقعها

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

عدد الفئات يناسب عدد المفردات وطول المدى

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

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

---

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

أن هذا العدد يكفي للتوزيع على الفئات بحيث تنال كل واحدة عدداً معقولا من المفردات ، وإلا أخذنا عدداً أصغر من الفئات .

طول مدى الفئة يسهل العمل ولكنه ليس دقيقا

ويلاحظ أنه كلما كثر عدد الفئات كان العمل الحسابي اللازم أصعب وأطول . وبالعكس إذا صغرنا عدد الفئات كانت العمليات الحسابية أقصر وأسهل ، ولكن على حساب الدقة فى النتيجة . وذلك لأنه كلما صغر عدد الفئات كان مدى الفئة طويلا . وينشأ عن هذا عدم تحقق الفرض بأن مركز الفئة يمثل جميع المفردات التى فيها (انظر بند ٨٣) ، إذ أن بعض القيم فى الفئة الطويلة المدى تكون بعيدة عن مركزها ، ونكون مخطئين كثيراً إذا اعتبرناها منطبقة عليه . فيجب إذن عند تحديد عدد الفئات التى ينقسم إليها المدى أن نأخذ فى الاعتبار درجة الدقة التى نبتغيها ، إذ أن عدد الفئات يعين طول الفترة فى كل فئة ، وبالتالى يعين مقدار كسور الفترة التى نهملها حينما نفرض أن كل القيم متجمعة عند مركز الفئة .

٨٦ — ولكن يجب ألا نغالى فى تضييق فترات الفئات طلباً للدقة . لأن هذا ينشأ عنه زيادة عدد الفئات أكثر من اللزوم ، فلا يكفى عدد المفردات للتوزيع على هذه الفئات الكثيرة الناتجة . ونحصل حينئذ على توزيع غير منتظم للتكرارات . ونرى هذا واضحاً من الجدول الآتى ، وهو يبين الدرجات التى حصل عليها عدد من التلاميذ (١٥٣٤) فى امتحان معين[1] :

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

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

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

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

جدول ٥ — التوزيع التكرارى للدرجات فى جدول ٤

فى فئات مدى فترتها درجة كاملة

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

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

(١) انظر شكل ٣٢ صفحة ١٠٣

جدول ٦ — التوزيع التكرارى السابق

فى فئات مدى فترتها درجتان

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

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

تعيين حدود الفئات

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

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

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

جدول ٧ — توزيع الملكية العقارية فى مصر فى سنة ١٩٣٦

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

الفئات غير المتساوية فى طول المدى

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

ويلاحظ أيضاً فى هذا الجدول أن الفئة الأخيرة ليس لها حد أعلى . وهذا هو ما نعبر عنه بقولنا أن الجدول **مفتوح من أعلى** (Open-end Table) وفى بعض الأحيان يكون الجدول مفتوح الطرفين . وعند ذلك تكون الفئة

الأولى فى الجدول معروفاً حدها الأعلى فقط ، كأن نقول عن الفئة الأولى فى الجدول المذكور فى بند ٨٢ إنها « أقل من ١٥٠ » ولانذكر ١٤٠ .

فئات خاصة للتفصيل فى بعض المناطق

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

تختار مبادئ الفئات بحيث تخفف أثر الخطأ فى جمع البيانات

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

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

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

ونجد مثالا آخر لهذه الأخطاء فى كيفية توزيع الدرجات فى الجداول المذكورة فى بند ٨٦ . حيث لاحظنا فى الجدول ٥ تضخما كبيراً فى تكرار الدرجة ٣ . وهذا بدون شك على حساب الفئتين المجاورتين لهذه — وعلى الأخص فئة الدرجة ٢٫٥ . والسبب فى ذلك واضح كما ذكرنا فى الحاشية فى صفحة ٨٨ . وهو أن كثيراً من التلاميذ الحاصلين فى الأصل على درجتين ونصف فقط يعطون نصف درجة للنجاح ، فيصبحوا عندنا فى فئة الدرجة ٣ . و يجوز أن يدخل هذه الفئة مفردات كانت فى الأصل فى فئة الدرجة ٢ ونقلت إلى ٣ لنفس السبب . و بذلك يتضخم تكرار الفئة ٣ كما نرى فى الجدول المذكور . (صفحة ٨٩) .

ويمكن تلافى أثر هذا الخطأ العملى بأن نختار حدود الفئات بحيث تكون الدرجة ٣ واقعة داخل فئة وليست فى طرفها . وأحسن نظام هو كالمقترح فى جدول ~~٧~~ ٦ حيث الفئة الثانية تمتد من ٢ إلى ٣٫٥ . وهذا يضمن أن كل الحالات التى كانت فى الأصل ٢ أو ٢٫٥ أو ٣ أو ٣٫٥ توضع فى المكان الصحيح ، حتى ولو كانت وضعت خطأ لسبب من الأسباب .

التعبير عن مبادئ ونهايات الفئات

٩٠ — وطريقة التعبير عن مبادئ الفئات ونهاياتها سهلة : و يحسن توحيد الطريقة لمنع الالتباس . وسنستعمل فى هذا الكتاب الصور الآتية لتدل على المعانى المبينة :

(١) — ٥ وأقل من ١٠ ، تكتب ٥ — للاختصار
١٠ » » ١٥ » ١٠ — »
١٥ » » ٢٠ » ١٥ — »

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

التى تليها ، وهكذا . و يكون مركز هذه الفئة المتوسط بين حديها أى ٧٫٥ للأولى و ١٢٫٥ للثانية و ١٧٫٥ للثالثة ، وهكذا .

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

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

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

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

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

و يصح أن يكون الجدول مفتوح الطرفين إذا ابتدأ وانتهى بمثلى هاتين الفئتين .

(٤) الصورة :

١٠

٢٠

٣٠

٤٠

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

(٥) الوضع

٤ — ٨

٨ — ١٢

١٢ — ١٦

بدون ذكر أي شيء يعين مبادئ الفئات أو نهاياتها ، وضع مبهم ويجب تحديده أو استبداله بإحدى الصور المحددة السابقة . إذ لا يمكن هنا تقرير الفئة التي توضع فيها القيمة ١٢ مثلا : هل توضع في الفئة الثانية أو الثالثة من هذه الفئات.

(٦) الوضع :

٥ — ٩

١٠ — ١٤

١٥ — ١٩

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

التغير المتصل

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

السلسلة المتصلة

ومجموعة القيم التي نحصل عليها لظاهرة مثل هذه نسميها **سلسلة متصلة**[1] . ومثل هذه المجموعة توضع في جدول تكراري « متصل » . أي أن الحد الأعلى لكل فئة يلاصق الحد الأدنى للفئة التي تليها مباشرة ؛ وكل قيمة بالقرب من هذين الحدين المتلاصقين تكون تابعة لفئة واحدة منها فقط .

ومثال ذلك الجدول التكراري للأعمار . إذ أن عمر الشخص يزيد بالتدريج بحيث إننا نجد أشخاصاً أعمارهم تساوي كل الأعمار التي بين ٢٥ و٣٠ سنة مثلا . وكذلك التوزيع التكراري للأطوال أو الأوزان وهكذا . وفي مثل هذه التوزيعات لا نجد صعوبة عملية أو نظرية في فرض قيم كل فئة تساوي مركز الفئة .

السلسلة المنفصلة

٩٢ — وبعكس ذلك نجد بعض الظواهر تتغير **تغيراً منفصلا** أي أن مقدار الظاهرة « يقفز » من قيمة إلى أخرى فجأة بدون أن يتدرج في القيم الواقعة

(١) Continuous Series

بينهما . مثال ذلك عدد ما عند الرجل من أطفال ، وعدد ما فى المصنع من عمال ، وعدد ما تحمله شجرة القطن من اللوز ، وهكذا . فبينما نجد أناساً كثيرين عند الواحد منهم ٣ أطفال ، وكثيرين غيرهم عند الواحد منهم ٤ أطفال ، لا نجد أحداً عنده ٣ر٧ من الأطفال مثلا . ونجد كثيراً من شجرات القطن تحمل كل واحدة ١٨ لوزة ، وكثيراً غيرها تحمل كل واحدة ١٩ لوزة مثلا ، ولا نجد شجرة واحدة تحمل ١٨ر٣ أو ١٨ر٦ من اللوزات .

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

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

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

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

(١) Discrete or Discontinuous Series

يلاحظ أيضاً أن الحد الأعلى لكل فئة منفصل عن الحد الأدنى للتى تليها ، وذلك لعدم وجود مصانع فيها عدد العمال واقع بين هذه الحدود : ٤ و ٥ أو ٩ و ١٠ مثلا .

ويجب أن نفرق بين هذه الحالة والحالة المذكورة فى (٦) من بند ٩٠ . والفرق هو أن المتغير هنا منفصل والفئة تمتد من ٥ إلى ٩ فقط ، على حين أن المتغير فى الوضع (٦) المذكور متصل والفئة تمتد من ٥ إلى ٩ وكسورها ؛ مثلا : ٩ سنين و ٣ أشهر أو ١٠ أو ١١ شهراً ، أو أى شىء أقل من ١٠ سنوات مهما كان .

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

لنأخذ الجدول التكرارى الآتى ، وهو يبين التوزيع التكرارى لأعمار

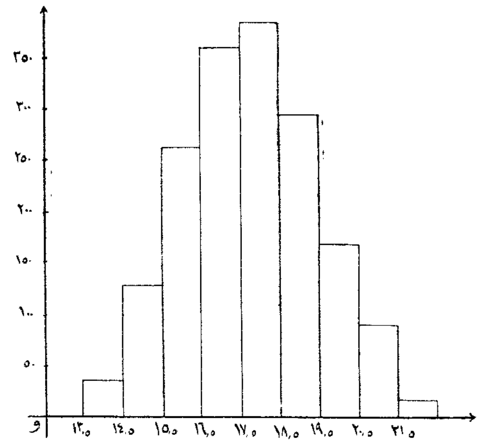
جدول ٩ — التوزيع التكرارى لأعمار ١٧٣٩ تلميذاً .

فئات العمر بالسنين	مراكز الفئات (العمر التقريبى)	عدد التلاميذ (التكرار)
١٣ر٥ وأقل من ١٤ر٥	١٤	٣٤
١٤ر٥ » » ١٥ر٥	١٥	١٢٨
١٥ر٥ » » ١٦ر٥	١٦	٢٦٢
١٦ر٥ » » ١٧ر٥	١٧	٣٦٠
١٧ر٥ » » ١٨ر٥	١٨	٣٨٦
١٨ر٥ » » ١٩ر٥	١٩	٢٩٤
١٩ر٥ » » ٢٠ر٥	٢٠	١٦٧
٢٠ر٥ » » ٢١ر٥	٢١	٩٢
٢١ر٥ » » ٢٢ر٥	٢٢	١٦
		المجموع ١٧٣٩

التلاميذ الناجحين من المدارس الأميرية فى امتحان شهادة الدراسة الثانوية (قسم أول) سنة ١٩٣٣ .

مدرج التكرار (الهستوجرام)

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



(شكل ٣٠) — مدرج تكرارى أو هيستوجرام

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

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

وإذا لم تكن الفئات متساوية المدى فتكون مساحات هذه المستطيلات (القاعدة × الارتفاع) هى التى تتناسب مع أرقام التكرار .

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

الفئات المفتوحة لا يرسم لها مستطيلات

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

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

مساحات المستطيلات تمثل التكرارات

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

(١) Histogram

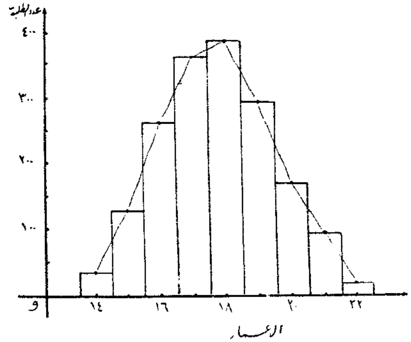
المفتوحة إذا وجدت ولم يرسم لها مستطيلات ؛ وحينئذ يجب استبعاد تكرارات هذه الفئات من حسابنا .

المضلع التكراري

٩٦ — يوجد طريقة أخرى لتوضيح التوزيع التكراري ، وهي رسم **المضلع التكراري**(١) .

لذلك نأخذ مراكز الفئات كإحداثيات أفقية ، والتكرارات المناظرة لها كإحداثيات رأسية . ثم نرصد في الشكل نقطاً بهذه الإحداثيات ، ونصل هذه النقط بخط منكسر فنحصل على ما نسميه المضلع التكراري .

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



(شكل ٣١)

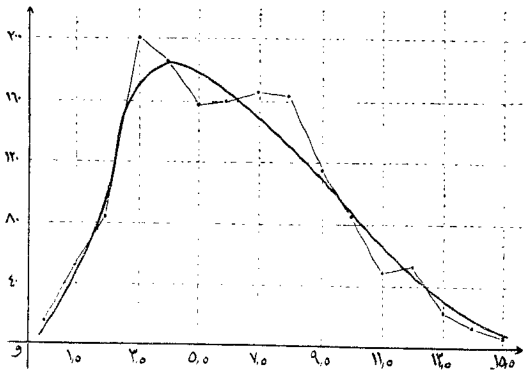
ويلاحظ أن المضلع التكراري يمكن رسمه من واقع المدرج التكراري . فهذه النقط التي نعينها في المضلع التكراري هي في الحقيقة منتصفات القواعد العليا

(١) Frequency Polygon

للمستطيلات في المدرج التكراري . فإذا وصلنا بين هذه المنتصفات بخط منكسر حصلنا على نفس المضلع التكراري السابق .

المضلع والمدرج متساويان في المساحة

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



(شكل ٣٢)

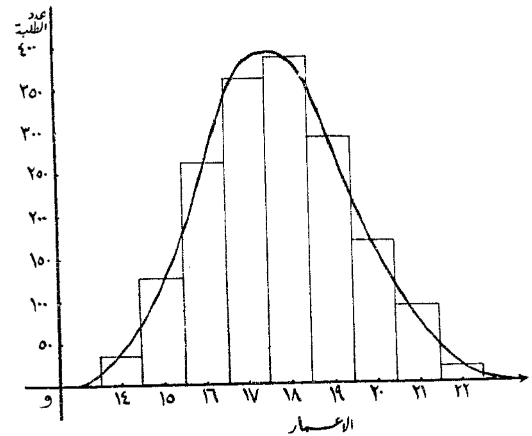
المضلع التكراري والمنحنى التكراري لتوزيع درجات بعض التلاميذ

المنحنى التكراري

٩٨ — إذا نحن رسمنا المضلع التكراري ثم مهدنا انكساراته نحصل على خط منحن خال من الذبذبات الفجائية ، ونسمي هذا الخط **المنحنى التكراري**(١)

(١) Frequency Curve

وهو يستعمل كطريقة أخرى لتمثيل التوزيعات التكرارية فى شكل هندسى واضح . وهي فى الحقيقة أحسن الطرق المستعملة لهذا الغرض بشرط أن يكون الرسم دقيقاً . ونرى فى شكل ٣٢ المضلع التكرارى والخط الممهد أو المنحنى التكرارى لتوزيع الدرجات المذكور فى جدول ٥ صفحة ٩٠ . وقد أخذنا مراكز الفئات هنا عند ½ و ½١ و ½٢ وهكذا .



(شكل ٣٣)

المدرج والمنحنى التكراريين لتوزيع أعمار بعض التلاميذ

مساحة المنحنى لاتساوى مساحة المضلع أو المدرج تماما

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

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

المدرج يؤول فى النهاية إلى المنحنى التكرارى

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

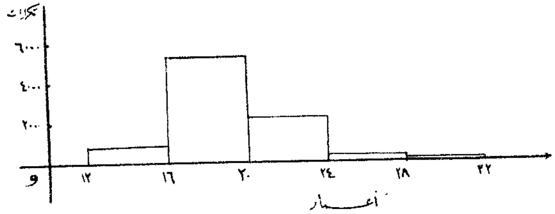
التدرج من الهستوجرام إلى المنحنى التكرارى

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

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

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

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



(شكل ٣٤)

مدرج تكرارى ، لتوزيع ذى فئات واسعة المدى

جدول ١٠ — توزيع أعمار ٨٤٤١ رجلا

فى فئات مداها ٤ سنين

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

نأخذ الآن نفس الأشخاص ونقسمهم إلى فئات ذات فترات أضيق .

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

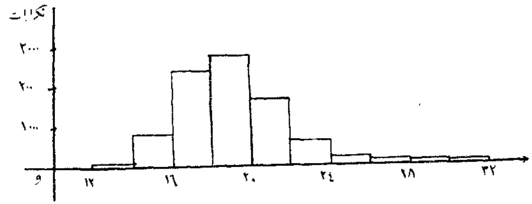
جدول ١١ — توزيع أعمار ٨٤٤١ شخصاً فى فئات مداها سنتان

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

وفى شكل ٣٥ نرى المدرج التكرارى الذى يمثل هذا التوزيع .

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

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

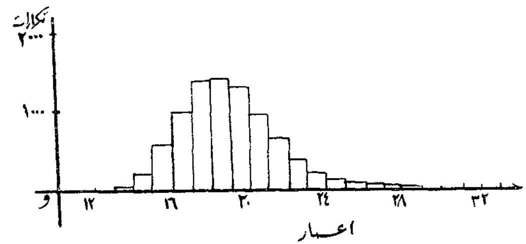


(شكل ٣٥)

مدرج تكرارى للتوزيع ذى فئات مداها سنتان

بدل سنتين . والجدول التكرارى الآتى (رقم ١٢) يبين توزيع أعمار هذه المجموعة [1] فى فئات طول الفترة فيها سنة واحدة .

وفى شكل ٣٦ نرى الهيستوجرام الذى يمثل هذا التوزيع .



(شكل ٣٦)

هيستوجرام ذو فئات طول فترتها سنة واحدة

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

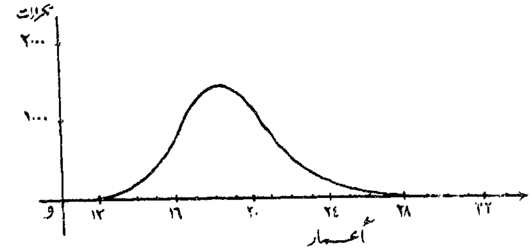
جدول ١٢ — التوزيع التكرارى لأعمار ٨٤٤١ شخصاً فى فئات طول فترتها سنة واحدة .

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

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

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

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



(شكل ٣٧)

منحنى تكرارى لتوزيع أعمار ٨٤٤١ شخصاً

الناتجة ، فنحصل على توزيع غير منتظم كما رأينا فى المثال المذكور فى بند ٨٦ . وإذا أمكن زيادة عدد المفردات كلما زاد عدد الفئات بسبب تجزئتها ، فلا بد أن نحصل فى النهاية على المنحنى الممهد .

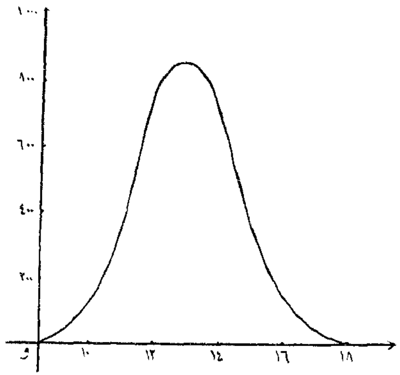
الأشكال المختلفة للمنحنيات التكرارية

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

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

١٠٢ — المنحنى التكرارى المتماثل ، وهو ، كما نرى فى شكل ٣٨ ، يشبه الناقوس العادى . وله محور تماثل رأسى يمر بنقطة النهاية العظمى له ، ويقسمه إلى

جزءين متطابقين . وقد سبقت الإشارة إليه فى بند ٧١ (شكل ٢٤) ، وسنعود إلى شرح بعض خواصه فى مناسبة أخرى . ونجد فى شكل ٣٨ مثالا عمليا لهذا المنحنى حيث نرى التوزيع التكرارى لأعمار هؤلاء التلاميذ متماثلا .



(شكل ٣٨)

منحنى تكرارى متماثل

وليست كل المنحنيات التكرارية من هذا النوع متشابهة . فقد يكون المنحنى ضيقاً ومرتفعاً أو قصيراً ومسطحاً وهكذا . وهذه الاختلافات تتوقف طبعاً على التوزيعات التكرارية وكيفية تغير الظواهر التى تمثلها هذه التوزيعات .

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

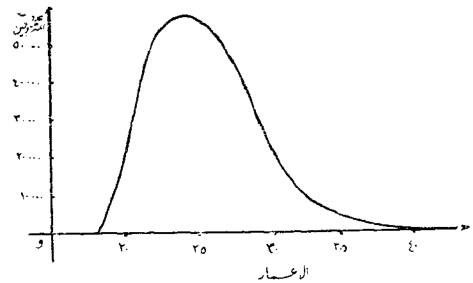
منحنى الخطأ المتماثل

ومن الأمثلة العملية لهذا المنحنى ، توزيع مقادير أخطاء التجارب العملية ، إذا كانت هذه الأخطاء بمحض المصادفة ولم تكن متحيزة فى ناحية معينة . فى هذه الحالة نجد أن تكرار خطأ قدره + ٠٫٠٣ مثلا فى قياس كمية معينة

(وزن جسم معلوم مثلا) يساوى تكرار الخطأ —٠٫٠٣ ، وهكذا . فعندما نرسم منحنياً تكرارياً لمقادير الخطأ نجده متماثلا . وهـذا هو السبب فى تسمية هـذا المنحنى **منحنى الخطأ المتماثل**[1] .

المنحنى الملتوى إلى اليســار

١٠٣ — المنحنى التكرارى غـير المتماثل ، وهو يشابه المنحنى السابق فى أن له قمـة واحدة ولـكن طرفيه غير متماثلين . فأحيـاناً يكون الطرف الأيمن ممتداً إلى مسافة أطول من اليسار ويسمى المنحنى حينئذ **ملتوياً الى اليسار** ؛ وأحياناً يكون الطرف الأيسر أطول من الأيمن ، ويسمى المنحنى حينئذ **ملتوياً الى اليمين** .



(شكل ٣٩)

المنحنى التكرارى لأعمار المتزوجين من الرجال (العزاب)

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

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

(1) Normal Curve of Error.

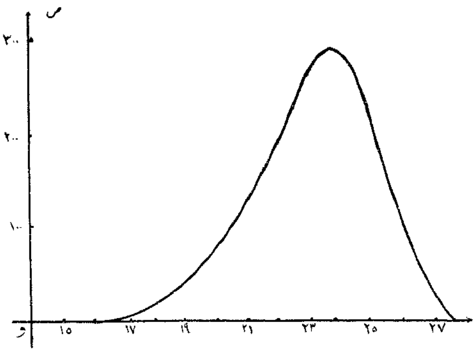
الذين تزوجوا من آنسات فى القطر المصرى فى سنة ١٩٣٥ . وتوزيع الأعمـار مبين فى جدول ١٣ .

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

فئات العمر	التـكرار	فئات العمر	التـكرار
١٨—	٧٨٥٧	٥٥—	٤٢
٢٠—	٥٤٣٣٣	٦٠—	١٩
٢٥—	٤٠٨٢١	٦٥—	٥
٣٠—	٩٥٧٨	٧٠—	٥
٣٥—	٢٤٩٣	٧٥—	٢
٤٠—	٤٥٤	٨٠ فأكثر	٣
٤٥—	١٦١		
٥٠—	٨٥	الجمـلة	١١٥٨٥٧

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

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



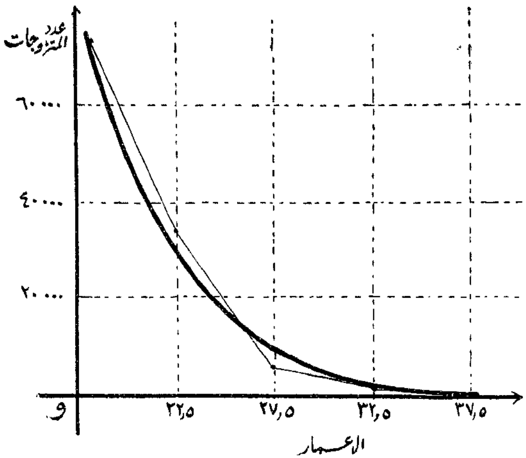
(شكل ٤٠)

منحن تكرارى ملتو الى اليمين

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

ونرى فى شكل ٤٠ منحنياً ذا فرع واحد نازلاً إلى اليمين ؛ وهو يبين التوزيع التكرارى لأعمار الآنسات اللاتى تزوجن من رجال (عزاب) فى القطر المصرى سنة ١٩٣٥ . ونرى الأرقام فى الجدول رقم ١٤ .

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

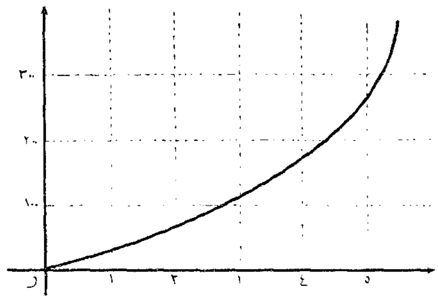


(شكل ٤١)

منحن تكرارى لتوزيع أعمار المتزوجات من الآنسات فى مصر سنة ١٩٣٥

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

ونرى فى شكل ٤٣ منحنياً تكرارياً ذا فرع واحد صاعداً من اليسار إلى اليمين .



شكل (٤٢)

منحن تـــكراري ذو فرع واحد — أيمن

جدول ١٤ — أعمار ١١٥٨٥٧ آنسة تزوجن من رجال عزاب

فى مصر سنة ١٩٣٥

فئات العمر	التـــكرار	فئات العمر	التـــكرار
١٦ —	٧٤٣٦١	٤٥ —	٤٨
٢٠ —	٣٤٣٢٣	٥٠ —	٢٧
٢٥ —	٦٠٣٠	٥٥ —	١٥
٣٠ —	٧٧٩	٦٠ —	٥
٣٥ —	١٩٧	٦٥ —	٢
٤٠ —	٧٥	٧٠ —	٥
		الجمـــلة	١١٥٨٥٧

منحـــنى تكراري ذو شعبتين

١٠٥ — المنحنى التكراري ذو النهاية الصغرى[1] وهو يخالف الأشكال السابق ذكرها . ويمثل ظاهرة تكون قيمها الصغرى شائعة وكذلك قيمها الكبرى . وأما القيم الوسطى فنادرة الحصول نسبياً . ومثال هـذه الظاهرة أعمار المتوفين من السكان . فعدد المتوفين من الأطفال (ذوى الأعمار الصغيرة) كبير جداً . وكذلك عدد المتوفين من الشيوخ (ذوى الأعمار الكبيرة) . وأما المتوفون من الشبان (وهم ذوو الأعمار المتوسطة بين الأطفال والشيوخ) فعددهم قليل نسبياً . والسبب فى ذلك واضح ؛ إذ أن الأطفال والشـيوخ أضعف من الشبان فى تحمل وطأة المرض ومقاومتـه ، فهم أكثر تعرضاً للوفاة .

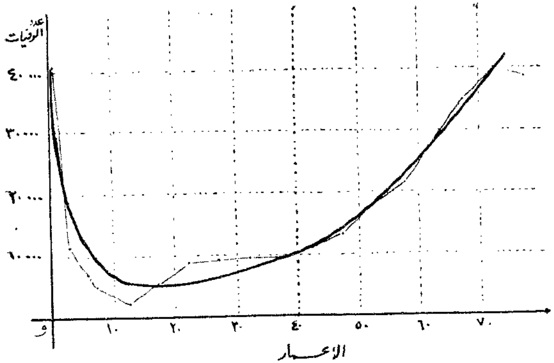
ونرى فى شـكل ٤٣ المنحنى التـكراري لتوزيع أعمـار المتوفيات من الإناث فى ألمانيا سنة ١٩٣٠ ؛ ونجد الأرقام فى جدول ١٥ .

جدول ١٥ — توزيع أعمار المتوفيات (بين عمر ٠ و ٧٥)

فى ألمانيا سنة ١٩٣٠

فئات السن	التكرار	فئات السن	التكرار	فئات السن	التكرار
أقل من ١ سنة	٤١٠٩١	٢٥—	٩٥٦٩	٥٥—	٢١٦٢٧
١ —	١٠٤٨٦	٣٠—	٩٦٥٨	٦٠—	٢٨٢٥٤
٥ —	٥٢٢٤	٣٥—	٩٩٧٦	٦٥—	٣٥٠٩٣
١٠—	٢٣٣٠	٤٠—	١١١٣٦	٧٠—	٤٠٢٧٣
١٥—	٥٧٠٩	٤٥—	١٣١٤٢	٧٥—	٣٧٨٣٦
٢٠—	٨٨٠٣	٥٠—	١٨٠٥٩	الجملة	٣٠٨٢٦٤

(١) U-Shaped Curve



(شكل ٤٣)
المنحنى التكرارى لأعمار المتوفيات (بين عمر ٠ و ٧٥) فى ألمانيا سنة ١٩٣٠

ويلاحظ أن النهاية الصغرى للمنحنى تقع فى فئة العمر ١٠ — ١٥ .

المنحنى التكرارى المتجمع

١٠٦ — المنحنى التكرارى المتجمع[1] وهو يبين تراكم التكرارات فى الفئات المتتالية . ولبيان طريقة رسمه نأخذ المثال المذكور فى بند ٩٣ (صفحة ٩٩) . وبواسطة الأرقام المعطاة فى جدول ٩ نكون جدول التكرار المتجمع المبين فى صفحة ١١٩ .

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

ونرسم المنحنى من واقع أرقام التكرارات المتجمعة والحدود العليا للفئات .

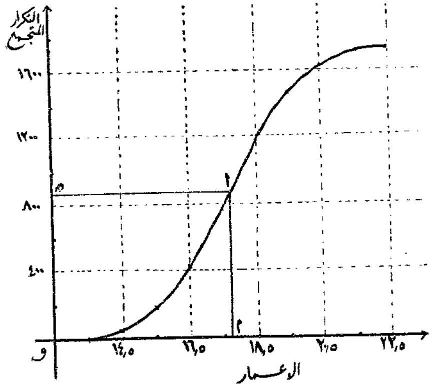
رسم المنحنى الصاعد

فنأخذ هـذه الحدود العليا كإحداثيات أفقية ، والتكرارات المتجمعة المناظرة لها كإحداثيات رأسية . والخط البيانى الذى يصل بين النقط المرصودة بهذه الإحداثيات هو المنحنى التكرارى المتجمع (الصاعد) المطلوب .

(١) Cumulative Frequency Curve = Ogive

جدول ١٦ — التكرارات المتجمعة لأعمار ١٧٣٩ تلميذاً

فئات العمر بالسنين	عدد التلاميذ (التكرار)	الحدود العليا للفئات	التكرار المتجمع (الصاعد)
		أقل من ١٣ر٥	٠
١٣ر٥ وأقل من ١٤ر٥	٣٤	» ١٤ر٥	٣٤
١٤ر٥ —	١٢٨	» ١٥ر٥	١٦٢
١٥ر٥ —	٢٦٢	» ١٦ر٥	٤٢٤
١٦ر٥ —	٣٦٠	» ١٧ر٥	٧٨٤
١٧ر٥ —	٣٨٦	» ١٨ر٥	١١٧٠
١٨ر٥ —	٢٩٤	» ١٩ر٥	١٤٦٤
١٩ر٥ —	١٦٧	» ٢٠ر٥	١٦٣١
٢٠ر٥ —	٩٢	» ٢١ر٥	١٧٢٣
٢١ر٥ وأقل من ٢٢ر٥	١٦	» ٢٢ر٥	١٧٣٩
الجملة	١٧٣٩		



(شكل ٤٤)
المنحنى التكرارى المتجمع الصاعد لأعمار ١٧٣٩ تلميذاً

نقط المنحنى ومعنى إحداثياتها

١٠٧ — وإذا أخذنا أية نقطة مثل ا على هذا المنحنى وأسقطنا منها عمودين على المحورين : ا م على المحور الأفقي و ا هـ على المحور الرأسي مثلا ، فإن البعد و م يمثل عمراً معيناً (= ١٧ر٧ سنة في الشكل) . أما البعد و هـ فهو يمثل عدداً من التلاميذ (= ٨٧٠ في الشكل) . والمهم أن كل هؤلاء التلاميذ عمرهم أقل من العمر ١٧ر٧ سنة الذي يمثله البعد و ا ، كما هو واضح من طريقة رسم المنحنى المتقدم شرحها .

وهكذا لو أخذنا نقطة ما على المحور الأفقي مثل ل ، وأقمنا منها عمودا يقابل المنحنى في نقطة مثل ں ، فإن البعد و ل ، مقيساً على المحور الأفقي (محور قياس الأعمار) ، يمثل عمراً معيناً ؛ والبعد ل ں ، مقيساً على المحور الرأسي (محور قياس التكرارات المتجمعة) ، يمثل عدداً من الأشخاص أعمارهم جميعا دون العمر و ل .

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

خواص المنحنى الصاعد لها فائدة عملية

١٠٨ — وهذه الخواص التي للمنحنى التكراري المتجمع ذات فائدة كبيرة : ونستخدمها كثيراً كما يتبين لنا في المستقبل . ويجب التنبه إلى نقطة مهمة عند رسم هذا المنحنى . وهي أن الإحداثيات الأفقية هي الحدود العليا للفئات التكرارية التي عندنا . وإغفال هذه الخاصة ينشأ عنه أخطاء في العمل .

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

الفئات جميعها ، فهو بالضرورة أكبر إحداثي رأسي . وإحداثيها الأفقي هو الحد الأعلى للفئة الكبرى ، فهو كذلك أكبر إحداثي أفقي .

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

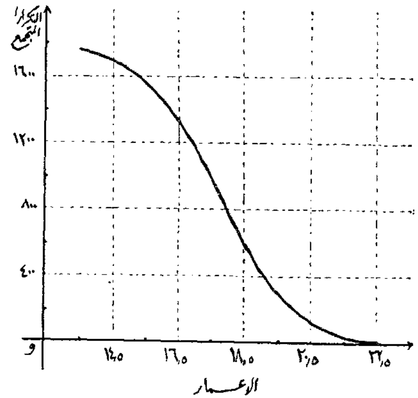
المنحنى التكراري المتجمع النازل

١٠٩ — ويمكن أن نمثل هذه الأعمار بمنحنى تكراري نازل (متجمع) أيضاً . ولبيان ذلك نأخذ نفس الفئات وتكراراتها ، ونكون منها الجدول التكراري المتجمع النازل كما يأتي :

جدول ١٧ — التكرار المتجمع النازل لأعمار ١٧٣٩ تلميذاً .

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

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



(شكل ٤٥)

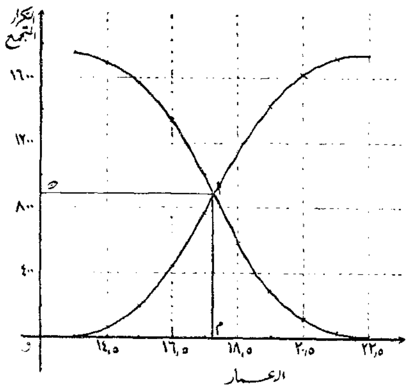
المنحنى التكرارى المتجمع النازل لأعمار ١٧٣٩ تلميذاً

رسم المنحنى

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

ولهذا المنحنى خواص مماثلة لخواص المنحنى الصاعد التى ذكرناها فى بند ١٠٧ . فكل نقطة عليه يمثل إحداثيها الرأسى عدداً من الأشخاص عمرهم جميعاً أكبر من العمر الذى يمثله إحداثيها الأفقى مقيساً على المحور الأفقى (وهو محور قياس الأعمار) ، وهذا ناتج طبعاً من تعريف المنحنى وطريقة رسمه .

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



(شكل ٤٦)

المنحنى التكرارى المتجمع الصاعد والمنحنى النازل لأعمار ١٧٣٩ تلميذاً

نقطة تقابل المنحنيين الصاعد والنازل

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

هـذا المثال) . والسبب فى ذلك واضح . فبما أن هـذه النقطة ا على المنحنى الصاعد (وليـكن إحـداثيها الأفقى و م كما فى شكل ٤٦) يكون إحداثيها الرأسى م ا يساوى عـدد الأشـخاص الذين عمرهم أقل من العمر و م ، كما قلنـا فى بند ١٠٧ . وبمـا أنها فى الوقت نفسه واقعة على المنحنى النـازل ، فإحداثيها الرأسى نفسه يمثـل عدداً من الأشـخاص أعمارهم جميعاً أكبر من العمر و م أيضاً ، كما ذكرنا فى بند ١٠٩ . وعلى ذلك فعدد الأشخاص الذين عمرهم أكبر من العمر و م ، يساوى عدد الأشخاص الذين عمرهم أقل من هذا العمر نفسه .

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

وكل نقطتين على هذين المنحنيين متحدتين فى الإحـداثى الأفقى يكون إحداثياهما الرأسيان مكملين لبعضهما ، أى أن مجموعهما يساوى المجموع الـكلى للتـكرارات . وهذا واضـح من طريقة رسمهما ومن طريقة تكوين التـكرار المتجمع الصاعد والنازل .

العلاقة بين المنحنى المتجمع والمنحنى التكرارى العادى

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

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

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

المنحنى التكرارى المتعدد القمم

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

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

خلط المجموعات المختلفة يسبب تعدد القمم أحيانا

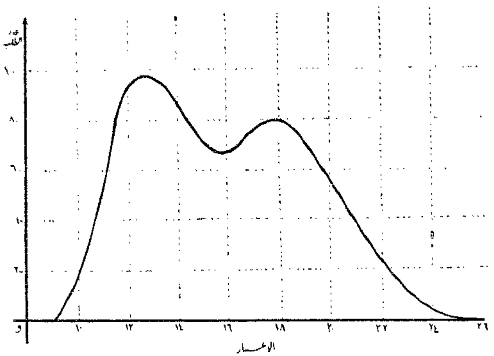
١١٣ — ومثل هذه المجموعات تكون مفرداتها غير متجانسة ، كأن تكون خليطا بين مجموعتين مختلفتين أو أكثر . وهذا الاختلاف بين المجموعات المركبة ينشأ عنه ازدياد فى تـكرارات الفئات الأولى حتى تصل إلى نهاية كبرى ثم تنقص . ولـكن لا تلبث أن تأخذ التـكرارات فى الزيادة مرة ثانية حتى تصل إلى نهاية كبرى ثانية — وربمـا ثالثة أو أكثر . وهـذا التذبذب يظهر أثره فى المنحنى التـكرارى فتتعدد القمم .

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

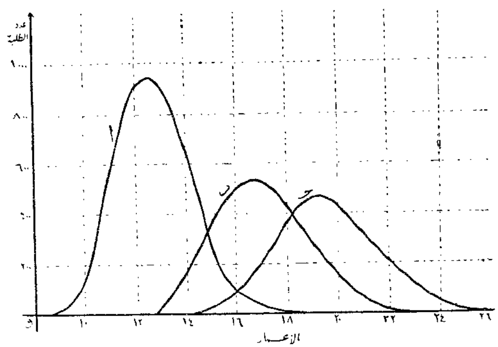
العمر	عدد التلاميذ	العمر	عدد التلاميذ
٩	١	١٩	٧٦٠
١٠	١٥٠	٢٠	٥٨٠
١١	٥٧٠	٢١	٣٧٧
١٢	٩٥٠	٢٢	٢٢٣
١٣	٩٨٧	٢٣	١١٣
١٤	٨٢٤	٢٤	٣٠
١٥	٦٦٨	٢٥	١١
١٦	٦٥٦	٢٦	٦
١٧	٧٧٠	٢٧	٢
١٨	٨٠٠	المجموع	٨٤٩٥

تعدد القمم دليل على عدم التجانس

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



(شكل ٤٧)
منحن تكرارى ذو قمتين

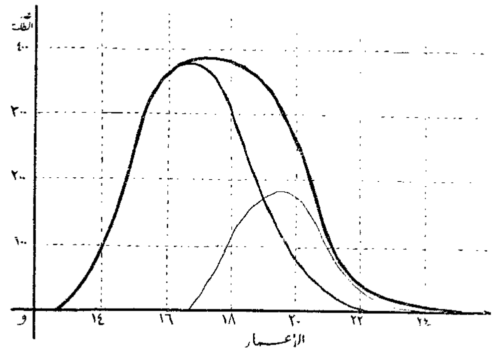


(شكل ٤٨)
المنحنيات التكرارية لثلاث مجموعات مختلفة

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

خلط المجموعات يسبب أحيانا اختلال التماثل أو تفرطح القمة

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



(شكل ٤٩)

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

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

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

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

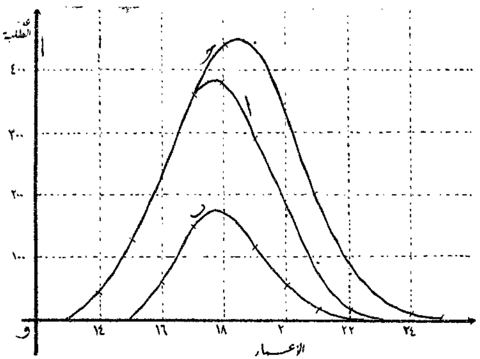
ونرى فى الشكل ثلاثة منحنيات : ا (الداخلى الأيسر) و ب (الداخلى الأيمن) و حـ (الخارجى) وهى تمثل المجموعة الأولى والثانية والثالثة على الترتيب .

شكل المنحنى التكرارى يدل على درجة التفاوت بين المفردات

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

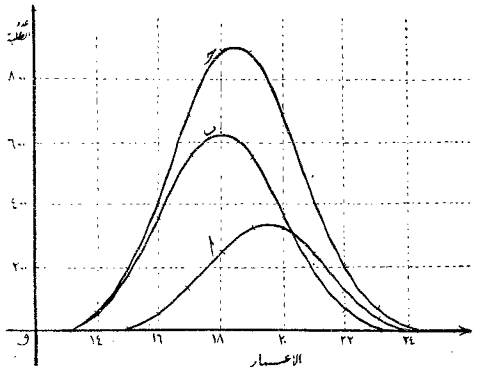
١١٦ — وكما أن خلط المجموعات يؤدى أحياناً إلى اختلال تماثل منحنى المجموعة المركبة منها ، يصح أن هذا الخلط بين مجموعتين يسوى ما فى أحدهما أو كليهما من التواء ، فتنتج منهما مجموعة متماثلة أو أقرب إلى التماثل من أيهما على حدة . ونرى مثالا لذلك فى شكل ٥٠ حيث نرى ثلاثة منحنيات منهما ا ذو التواء بسيط و ب ذو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

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



(شكل ٥٠)

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

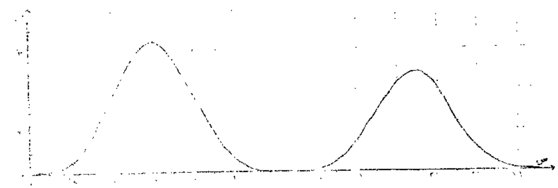


(شكل ٥١)

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

والمجموعات المتماثلة على العموم يتكون من خلطها مجموعات متماثلة ذوات منحنيات تكرارية متماثلة أيضاً . وفى حالات خاصة يتكون من جمع منحنيين متماثلين ، منحن ذو قمتين . ولتوضيح ذلك نأخذ منحنيين متماثلين ونخلطهما ونرسم المنحنى الناتج فى كل حالة . وتفصيل هذه الحالات هو كما يأتى :

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

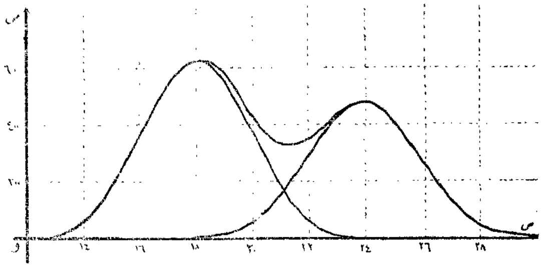


(شكل ٥٢)

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

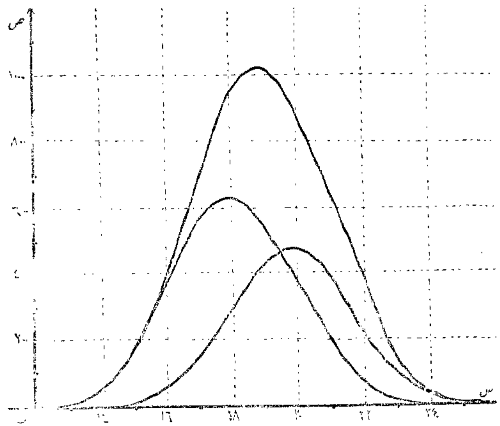
ثالثاً : حينما يشترك المنحنيان فى جزء كبير من قاعدتيهما .

فى هذه الحالة تمتزج المجموعتان و يتكون منهما مجموعة متماثلة لها منحن متماثل ذو قمة واحدة ، واقعة بين القمتين الأصليتين أيضاً كما فى شكل ٥٤ الذى يبين



(شكل ٥٣)

التوزيعين التكراريين لأعمار تلاميذ المدارس الأميرية المتقدمين إلى امتحان شهادة الدراسة الثانوية قسم أول وقسم ثان فى سنة ١٩٣٣ . ومن الرسم نرى أن



(شكل ٥٤)

الإحداثى الأفقى لقمة المنحنى الأول عند ١٨ سنة ، والثانى عند ٢٠ سنة . والمنحنى الناتج من المجموعتين متماثل أيضاً والإحداثى الأفقى لقمته عند ١٩ تقريباً .

رابعاً : حينما يكون البعد الأفقي بين القمتين صفراً .

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

---

(١) يمكن إثبات هذه الحالات جميعاً باستخدام نظرية النهايات الكبرى والصغرى فى حساب التفاضل والتكامل كما يأتى :

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

(١) : ص١ = هـ^-(س - ا)^٢

و (٢) : ص٢ = هـ^-(س - ب)^٢ ، ب > ا

وأن معادلة المنحنى المركب من جمعهما هى :

(٣) : ص = هـ^-(س - ا)^٢ + هـ^-(س - ب)^٢

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

∴ (س - ا) . هـ^-(س - ا)^٢ + (س - ب) . هـ^-(س - ب)^٢ = ٠

ومن الواضح أن س = ½ (ا + ب) تحقق هذه المعادلة ؛ وهى أيضاً نقطة رجوع لهذا المنحنى (٣) .

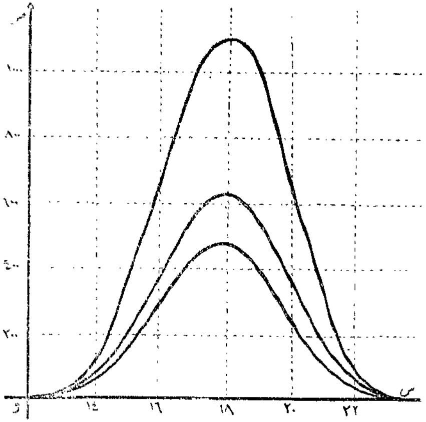
نوجد قيمتى ص عند ما س = ½ (ا + ب) ± حـ ، أى على بعد حـ على يمين ويسار نقطة الرجوع .

لو عوضنا بهاتين القيمتين عن س فى المعادلة (٣) نجد أن قيمة ص واحدة فى الحالتين . وينتج من ذلك أن المنحنى (٣) متماثل بالنسبة إلى محور رأسى يمر بنقطة الرجوع حيث س = ½ (ا + ب) .

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

صً = - ٢ هـ^-(س - ا)^٢ - ٢ هـ^-(س - ب)^٢ + ٤ (س - ا)^٢ × هـ^-(س - ا)^٢ + ٤ (س - ب)^٢ . هـ^-(س - ب)^٢ .

فلو عوضنا عن س بقيمتها عند نقطة الرجوع أى س = ½ (ا + ب) ، (*)



(شكل ٥٥)

(*) ∴ صً = [- ٤ + ٢ (ب - ا)^٢] × هـ مرفوعة إلى القوة - ((ب - ا)/٢)^٢

أى أن صً موجبة إذا كان (ب - ا)^٢ > ٢ ، وسالبة إذا كان (ب - ا)^٢ < ٢ .

∴ نقطة الرجوع تكون نهاية صغرى إذا كان ب - ا > √٢ ويكون للمنحنى (٣) حينئذ قمتان ، كما فى ثانياً (وأولا) .

وإذا كان ب - ا < √٢ ، تكون نقطة الرجوع نهاية كبرى ؛ ويكون للمنحنى نهاية كبرى ، أى قمة واحدة .

يلاحظ أننا أخذنا حالة خاصة فى هذا البرهان للسهولة ؛ ولكنه ممكن تعميمه .

المراجـــع

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

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

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

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

الباب السادس

المتوســـطات

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

يجب تعيين نموذج يمثل كل مجموعة

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

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

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

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

لا يمكن مقارنه المجموعات بدون معرفة نماذج لها

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

القيمة النموذجية لمجموعة تقع متوسطه بين طرفيها

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

علم الاحصاء ليس علم

١٢١ — والبحث فى هذه المتوسطات الإحصائية (Statistical Averages) وخواصها والأغراض التى تؤديها من أهم أبحاث هذا العلم ، ويشغل جزءاً كبيراً

المتوسطات فقط

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

تعريف المتوسط الاحصائى لمجموعة

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

على أى أساس نختار المتوسط ليمثل المجموعة

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

الوسط الحسابى يصلح لان يمثل المجموعة

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

---

(١) (Arithmetic Mean)

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

الوسط الهندسى يصلح أيضا كمتوسط

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

الوسط التوافقى

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

القيمة الأكثر شيوعا تصلح لأن تمثل المجموعة

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

القيمة الوسطى أى الوسيط تصلح أيضا

١٢٦ — يصح أيضاً أن نعتبر المتوسط الذى يمثل المجموعة هو القيمة الوسطى فيها ، بحيث إذا رتبت مفرداتها ، تصاعديا أو تنازليا ، كانت هى فى

---

(١) Geometric Mean

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

الوسط تماماً ، وعدد المفردات التى قبلها يساوى عدد المفردات التى بعدها . وبذلك يكون عدد المفردات التى أكبر من القيمة الوسطى يساوى تماماً عدد المفردات التى أصغر منها . ولاشك أن هذا اعتبار وجيه وله قيمته . والمتوسط بهذا المعنى يسمى بالانجليزية (The Median) ، وأقترح تسميته « **الوسيط** » (١)

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

طرق حساب المتوسطات

إيجاد الوسط الحسابى

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

٤٤ و ٤٢ و ٤١ و ٤٥ و ٤٨ و ٤٧ سنة ،

يساوى (٤٤ + ٤٢ + ٤١ + ٤٥ + ٤٨ + ٤٧) / ٦ = ٢٦٧ / ٦

= ٤٤ر٥ سنة .

---

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

ومن الواضح أنه يمكن اختصار هذا العمل نوعاً بأن نقول إنه

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

$$= ٥ر٤٤ \text{ سنة .}$$

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

٤ و ٢ و ١ و ٥ و ٨ و ٧ ؛

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

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

استخدام وسط فرضى

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

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

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

$$\bar{\text{س}} = \frac{\text{س}_{١} + \text{س}_{٢} + \text{س}_{٣} + \cdots + \text{س}_{\text{ن}}}{\text{ن}}$$

لتكن ا كمية اختيارية ، أيا كانت .

$$\therefore \bar{\text{س}} = \text{ا} + \frac{\text{س}_{١} + \text{س}_{٢} + \cdots + \text{س}_{\text{ن}}}{\text{ن}} - \text{ا}$$

$$= \text{ا} + \frac{\text{س}_{١} + \text{س}_{٢} + \cdots + \text{س}_{\text{ن}} - \text{ن ا}}{\text{ن}}$$

$$= \text{ا} + \frac{(\text{س}_{١} - \text{ا}) + (\text{س}_{٢} - \text{ا}) + (\text{س}_{٣} - \text{ا}) + \cdots + (\text{س}_{\text{ن}} - \text{ا})}{\text{ن}}$$

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

ويمكن وضع هذه النتيجة فى صورة مختصرة جدا إذا وضعنا

$$\text{مج س} = \text{س}_{١} + \text{س}_{٢} + \text{س}_{٣} + \cdots + \text{س}_{\text{ن}} \text{ ؛}$$

$$\text{مج}(\text{س} - \text{ا}) = (\text{س}_{١} - \text{ا}) + (\text{س}_{٢} - \text{ا}) + \cdots + (\text{س}_{\text{ن}} - \text{ا}) \text{ ؛}$$

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

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

إيجاد الوسط الحسابى للتوزيع التكرارى

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

نلاحظ مبدئياً أننا فى التوزيع التكرارى نعتبر كل المفردات التى فى فئة تكرارية واحدة جميعها متساوية ؛ وكل منها تساوى مركز الفئة (أنظر بند ٨٣).

وبناء على ذلك نعتبر مراكز الفئات هي الأعمار التى نريد إيجاد وسطها الحسابى ،

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

جدول ٢٠ — إيجاد الوسط الحسابى لأعمار ١٧٣٩ تلميذاً .

فئة العمر بالسنين	مركز الفئة (العمر = س)	عدد التلاميذ التكرار = ك	حاصل ضرب س × ك
١٣ر٥ وأقل من ١٤ر٥	١٤	٣٤	٤٧٦
١٤ر٥ « « ١٥ر٥	١٥	١٢٨	١٩٢٠
١٥ر٥ « « ١٦ر٥	١٦	٢٦٢	٤١٩٢
١٦ر٥ « « ١٧ر٥	١٧	٣٦٠	٦١٢٠
١٧ر٥ « « ١٨ر٥	١٨	٣٨٦	٦٩٤٨
١٨ر٥ « « ١٩ر٥	١٩	٢٩٤	٥٥٨٦
١٩ر٥ « « ٢٠ر٥	٢٠	١٦٧	٣٣٤٠
٢٠ر٥ « « ٢١ر٥	٢١	٩٢	١٩٣٢
٢١ر٥ « « ٢٢ر٥	٢٢	١٦	٣٥٢
		١٧٣٩	٣٠٨٦٦

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

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

∴ الوسط الحسابى = $\frac{٣٠٨٦٦}{١٧٣٩}$ = ١٧ر٧٤٩ سنة .

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

إيجاد الوسط الحسابى باستخدام وسط فرضى

١٣١ — الطريقة المتقدمة فى البند السابق متعبة نوعاً . ويمكن اختصارها وتسهيل العمليات الحسابية فيها باستخدام العلاقة التى أثبتناها فى بند ١٢٩ . وهى اختيار وسط فرضى ، وإيجاد الوسط الحسابى للفروق بين هذا الوسط الفرضى وبين الأعمار ، ثم إضافة هذا الوسط الحسابى للفروق إلى الوسط الفرضى — ينتج الوسط الحسابى للأعمار نفسها ، وهو الذى نرمى إليه .

ويمكن شرح خطوات العمل فى جدول كالآتى :

جدول ٢١ — إيجاد الوسط الحسابى بالطريقة المختصرة باستعمال وسط فرضى .

الفئات (بالسنين)	مركز الفئة (العمر = س)	عدد التلاميذ (التكرار = ك)	الانحرافات عن ١٨ (ح)	ضرب ح × ك
١٣ر٥ —	١٤	٣٤	—٤	—١٣٦
١٤ر٥ —	١٥	١٢٨	—٣	—٣٨٤
١٥ر٥ —	١٦	٢٦٢	—٢	—٥٢٤
١٦ر٥ —	١٧	٣٦٠	—١	—٣٦٠
١٧ر٥ —	١٨	٣٨٦	٠	٠
١٨ر٥ —	١٩	٢٩٤	١	٢٩٤
١٩ر٥ —	٢٠	١٦٧	٢	٣٣٤
٢٠ر٥ —	٢١	٩٢	٣	٢٧٦
٢١ر٥ —	٢٢	١٦	٤	٦٤
		١٧٣٩		—٤٣٦

نختار وسطاً فرضياً مناسباً ، وليكن العمر ١٨ ، أى مركز الفئة (١٧ر٥ و أقل من ١٨ر٥) . ويلاحظ أن تكرار هذه الفئة أكبر تكرار . وسنرى أن هذا الاختيار يساعدنا كثيراً فى تخفيف عمليات الضرب والعمليات الحسابية الأخرى . ولذلك يحسن دائماً أن نختار الوسط الفرضى عند مركز الفئة ذات التكرار الأكبر ، أو فئة قريبة منها .

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

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

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

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

= ١٨ — ر٢٥١

= ١٧و٧٤٩ سنة ؛

وهو نفس الناتج الذى حصلنا عليه من قبل .

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

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

طريقة تغيير الوحدات

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

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

فئات الأجور (بالقرش)	مركز الفئة (الأجر = س)	عدد العمال (التكرار = ك)	الانحراف عن ١٩ر٥ ع	الانحراف الجديد ب	ضرب ك × ب
١٢ وأقل من ١٥	١٣ر٥	٢٢٤	— ٦	— ٢	— ٤٤٨
١٥ » ١٨	١٦ر٥	١٩٧١	— ٣	— ١	— ١٩٧١
١٨ » ٢١	١٩ر٥	٣٧٥٥	٠	٠	٠
٢١ » ٢٤	٢٢ر٥	١٢٣٦	٣	١	١٢٣٦
٢٤ » ٢٧	٢٥ر٥	١٩٦	٦	٢	٣٩٢
٢٧ » ٣٠	٢٨ر٥	٤٠	٩	٣	١٢٠
٣٠ » ٣٣	٣١ر٥	١٠	١٢	٤	٤٠
		٧٤٣٢			— ٦٣١

لنأخذ فى هذا المثال الوسط الفرضى يساوى ١٩ر٥ ، وهو مركز الفئة الأكبر تكراراً . ثم نحسب انحرافات الأجور الأخرى عن هذا الأجر ، فنجدها فى العمود ع عبارة عن مكررات العدد ٣ ؛ وهو يساوى طول فترة الفئة . فبدل أن نضرب هذه الانحرافات فى التكرارات ، يمكننا اختزالها نوعاً بقسمتها جميعاً على العدد المشترك ٣ . فنحصل على أرقام أصغر وأسهل فى الضرب . وهذه نسميها

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

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

— ٦٣١ / ٧١٣٢ = — ٠٨٤٩ ر ٠ وحدة من وحدات ب

∴ الوسط الحسابى للأجور بالقروش يساوى

١٩٫٥ + (— ٠٨٤٩ ر × ٣) = ١٩٫٥ — ٢٥٤٧ ر

= ١٩٫٢٥ قرشاً (تقريباً)

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

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

١٣٣ — فى الجدول التكرارى المفتوح (أنظر بند ٩٠ «٣» صفحة ٩٥) لا يمكن إيجاد الوسط الحسابى . لأن الجدول المفتوح من أحد الطرفين أو كليهما يحتوى على فئة أو اثنتين لا يعرف مركزها . وعلى ذلك لا يمكن إجراء خطوة ضرب مركز الفئة فى تكرارها لإيجاد الوسط الحسابى المطلوب . ففى الجدول الآتى مثلا يمكننا ضرب تكرار الفئة الأولى فى صفر وتكرار الثانية فى مركزها وهو ٢٫٥ ، الوسط الحسابى لا يمكن إيجاده من الجداول المفتوحة

جدول ٢٣ — جدول تكرارى ، مفتوح من أعلى

لتوزيع عدد العمال فى مصانع القاهرة سنة ١٩٢٧

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

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

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

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

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

إيجاد المنوال

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

توزيع أعمار ١٧٣٩ تلميذاً

الفئات	مراكز الفئات	التكرارات
١٣ر٥ —	١٤	٣٤
١٤ر٥ —	١٥	١٢٨
١٥ر٥ —	١٦	٢٦٢
١٦ر٥ —	١٧	٣٦٠
١٧ر٥ —	١٨	٣٨٦
١٨ر٥ —	١٩	٢٩٤
١٩ر٥ —	٢٠	١٦٧
٢٠ر٥ —	٢١	٩٣
٢١ر٥ —	٢٢	١٦

المنوال يساوى مركز الفئة المنوالية تقريبا

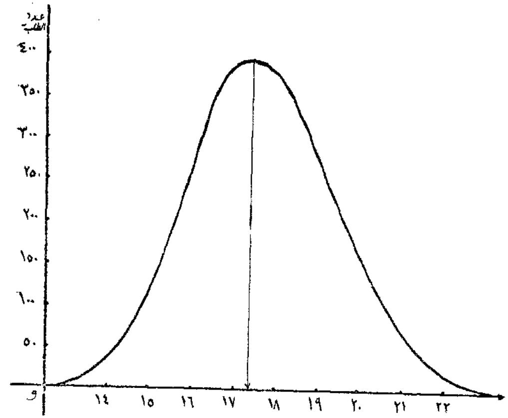
قلنا إن المنوال فى أى مجموعة هو القيمة الأكثر شيوعاً من غيرها ، أى القيمة التى تتكرر أكثر من غيرها . وبالنظر إلى التوزيع التكرارى الذى نحن بصدده ، نرى أن العمر ١٨ تكراره يساوى ٣٨٦ وهو أكبر من تكرار أى قيمة أخرى فى الجدول . وعلى ذلك يمكننا أن نعتبر هذه القيمة هى المنوال المطلوب . ويكون المنوال إذن يساوى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

ولكن هذا التقدير تقريبى ، إذ الفئة المنوالية التى تحتوى على التكرار الأكبر ، لها مدى واسع يبتدىء عند ١٧ر٥ وينتهى عند ١٨ر٥ . وحبذا لو علمنا على وجه التحديد أين يقع المنوال فى هذه الفترة : هل هو أقرب إلى الحد الأدنى أو الحد الأعلى ؟ وأين هو بالدقة ؟

ولتحديد موقع المنوال بالدقة نلجأ إلى طرق أخرى نشرحها فيما يلى .

إيجاد المنوال بالرسم

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



(شكل ٥٦)

تعيين المنوال بالرسم من المنحنى التكرارى

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

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

إيجاد المنوال بالحساب

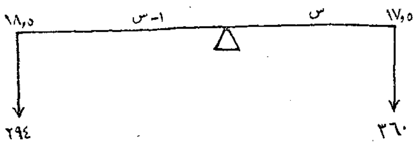
١٣٦ — توجد طريقة حسابية ، تقريبية نوعاً ، يمكن استخدامها لإيجاد المنوال . وهى تبنى على استخدام تكرارى الفئتين المجاورتين للفئة المنوالية ، من

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

الفئة	التكرار
١٦ر٥ وأقل من ١٧ر٥	٣٦٠
١٧ر٥ « ١٨ر٥	٣٨٦
١٨ر٥ « ١٩ر٥	٢٩٤

فالمنوال إذن يقع بين ١٧ر٥ و ١٨ر٥ . ليكن ١٧ر٥ + س . ويتعين المنوال تماماً إذا عرفنا قيمة هذا المجهول س . ولمعرفة قيمة س ، نقول إن التكرارين ٣٦٠ و ٢٩٤ أشبه بقوتين تؤثران فى المنوال وتتجاذبانه بحيث إن الأقوى منهما تجعله أدنى إلى حد الفئة المنوالية القريب منها ، مع وجوده دائماً داخل فترة الفئة المنوالية ولا يتعداها . ويكون موقع المنوال فى هذه الفترة إذن بحيث يتعادل تأثير القوتين مع بعضهما .

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



وعلى ذلك لحساب قيمة س نقول :

٣٦٠ × س = ٢٩٤ × (١ – س)   (١)

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

$$= ٤٥٠ر \quad \text{تقريباً}$$

$$\therefore \text{يكون المنوال} = ١٧ر٥ + ٠ر٤٥$$

$$= ١٧٩ر٥ \ \text{سنة} .$$

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

وعلى العموم ، إذا كان تكرار الفئة قبل المنوالية يساوى ك١ ، وتكرار الفئة بعد المنوالية يساوى ك٢ ، وكان طول الفترة فى الفئة المنوالية يساوى ل ، والحد الأدنى لها يساوى ا ، فان المنوال المطلوب

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

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

طريقة الفروق كارل بيرسون **١٣٧** — وقد اقترح بيرسون (Karl Pearson) طريقة أخرى لحساب المنوال ، تعتمد أيضاً على تكرارات الفئات الثلاثة ، المنوالية والمحيطتين بها .

والفكرة هنا أن نأخذ الفرق بين تكرارى الفئتين المنوالية والتى قبلها ، أمام الفرق بين تكرارى الفئتين المنوالية والتى بعدها ، كعاملين يؤثران فى المنوال ، بدل التكرارين الجانبيين ك١ ، ك٢ فى الطريقة السابقة .

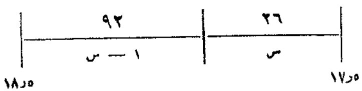
وهذا على اعتبار أن التغير فى التكرار من فئة إلى الفئة التى تليها ، يكون دليلا أدق على قرب أو بعد المنوال ، من نفس تكرارى الفئتين المجاورتين للمنوالية .

و بتطبيق هذه القاعدة على المثال الذى نحن بصدده نجد :

الفئـــــة	التكرار	الفرق
١٦ر٥ وأقل من ١٧ر٥	٣٦٠	
		٢٦
١٧ر٥ » ١٨ر٥	٣٨٦	
		٩٢
١٨ر٥ » ١٩ر٥	٢٩٤	

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

وهنا يكون المنوال واقعاً بين حدى الفئة المنوالية أيضاً ، وليكن ١٧ر٥ + س ، وتعيين قيمة س يكون بتقسيم المسافة بين ١٧ر٥ و ١٨ر٥ بنسبة ٢٦ : ٩٢ .



$$\therefore \quad س \times ٩٢ = ٢٦ (١ - س) ؛$$

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

$$= ٢٢ ر$$

$$\therefore \quad \text{المنوال} = ١٧ر٥ + ٠ر٢٢$$

$$= \cancel{١٧ر٧٢} \ ١٧٫٧٢ \ \text{سـنة} .$$

وعلى العموم : إذا كان تكرار الفئة المنوالية يساوى ك ، وتكرار التى قبلها ك١ ، وتكرار التى بعدها ك٢ ؛ وكان طول الفترة المنوالية يساوى ل ، والحد الأدنى لها يساوى ا ، فعلى أساس هذه الطريقة :

$$\text{المنوال} = ا + ل \times \frac{ك - ك_١}{(ك - ك_١) + (ك - ك_٢)}$$

$$= ا + ل \times \frac{ك - ك_١}{٢ك - ك_١ - ك_٢}$$

طريقة بيرسون أدق من السابقة

١٣٨ — ولو تأملنا فى هذا ، وجدنا أن الفكرة الأساسية فى طريقة بيرسون أقرب للصواب من فكرة الطريقة السابقة لها . وذلك لأنه إذا كان الفرق بين تكرارى فئتين متتاليتين صغيراً ، وكان المنوال واقعاً فى أحدها ، فانه يكون أقرب إلى الحد المشترك بين هاتين الفئتين مما لو كان الفرق بين تكراريها كبيراً .

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

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

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

إيجاد المنوال بتعديل الفترات

١٣٩ — إذا نحن غيرنا فترات الفئات فى أى توزيع تكرارى عادى ، وعدلنا مواقع حدودها ، فان هذا يحدث تغييراً فى التكرارات ؛ وفى الغالب يحدث تغييراً فى موقع الفئة المنوالية . ويمكن استخدام هــذه الظاهرة فى البحث عن المنوال . وذلك بأن نغير مواقع الفئات وفتراتهــا ، ونحصل على عــدة جداول تكرارية مختلفة لنفس المجموعة من القيم . وكل جدول يظهر فيه فئة ذات تكرار

أكبر يقع فيها المنوال . وبمقارنة الفئات المنواليّة وفتراتهــا التى نحصل عليها من الجداول التكرارية المختلفة يمكن الاهتداء إلى موقع المنوال ، ويكون إذ ذاك فى المنطقة المشتركة بين فترات الفئات المنواليّة للجداول التكرارية المختلفة .

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

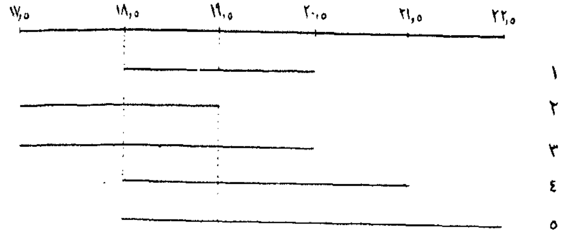
جدول ٢٤ — توزيع أعمار ٢٢٠٢ طالب

الأعمار	التكرارات فى فترات من :				
	٢ سنة	٢ سنة	٣ سنة	٣ سنة	٤ سنة
١٥	١١٤	١٥	٢٠٦	١٥	٨٠٤
١٦		٣٩١		٧٨٩	
١٧	٦٩٠				
١٨		٨٤٠	١١٧٣		
١٩	٧٧٥			١٠٣٥	١١٩٩
٢٠		٥٩٣			
٢١	٤٣٤		٥٠٢		
٢٢		٢٤٣		٢٩٨	
٢٣	١٣٤				١٧٦
٢٤		٨٠	٩٨		
٢٥	٤٣			٤٨	
٢٦		٢٤			
٢٧	١١		١٤		١٤
٢٨		٨		١٧	
٢٩	٣				
٣٠		٩	٩		
٣١	٩				٩
	٢٢٠٢	٢٢٠٢	٢٢٠٢	٢٢٠٢	٢٢٠٢

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

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

و إذا وضعنا هذه المناطق فوق بعضها فى نظام كالآتى :



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

تعدد المناويل

١٤٠ — إذا كان المنحنى التكرارى متعدد القمم[1]، فمعنى ذلك أن له أكثر من منوال واحد . و يمكننا إيجاد كل منها باستخدام تكرارى الفئتين المجاورتين لكل منوال بالطرق التى شرحناها فى بند ~~١٣٣ و ١٣٥~~ ١٢٤ — ١٢٧ . ولكن على كل حال فالمنوال لمثل هذا المنحنى لا يكون له فائدة كبيرة من حيث تمثيل المجموعة ، حيث

(١) يسمى بالانجليزية (Mutimodal) ; أو (Bimodal) إذا كان ذا قمتين

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

اختلاف قيم المنوال المحسوبة بالطرق المختلفة

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

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

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

١٤٢ – طريقة إيجاد الوسيط لمجموعة من القيم تنبني على ترتيبها تصاعدياً (أو تنازلياً) والقيمـة الوسطى ، التى يسبقها ويليها عددان متساويان من القيم ، هى الوسيط .

إيجاد الوسيط لمجموعـة من القيم

إذا كان عدد القيم صغيراً فمن السهل ترتيبها تصاعدياً ومعرفة الوسـيط ؛ وهو القيمة التى ترتيبها يساوى $\frac{١}{٢}$ (ن + ١) ، حيث ن هى عدد القيم جميعها .

فالعمر الوسيط للمجموعة الآتية مثلا :

١٧ ١٨ر٥ ٢٠ ٢١ ٢١ر٥ ٢٢ ٢٣ر٥

هو ٢١ سنة ؛ وهى القيمة الرابعة فى هـذه السلسلة التصاعدية ، أى التى ترتيبها $\frac{١}{٢}$ (٧ + ١) .

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

٧ر٥ ٩ ١٠ر٥ ١١ ١٢ ١٣ر٥ ١٤ ١٥ قرشاً ،

حيث القيمتان الرابعـة والخامسة فى السلسلة ، أى ١١ و ١٢ قرشاً ، يسبقهما ثلاث قيم أصغر منهما ، ويليهما ثلاث أخر أكبر منهما . ففى مثل هـذه الحالة نعتبر الوسط الحسابى بينهما هو الوسيط ، كما لو كانت هناك قيمة وهمية واقعة بين القيمتين الرابعة والخامسة ، وترتيبها $٤\frac{١}{٢}$ ، أى $\frac{١}{٢}$ (٨ + ١) ، بوضع ٨ بدل ن فى العبارة $\frac{١}{٢}$ (ن + ١) المذكورة أعلاه .

١٤٣ – إذا كان عدد القيم فى مجموعة ما كبيراً ، ونظمناها فى فئات على شـكل جدول تكرارى ، يمكن ترتيبها تصاعدياً فى جدول تكرارى متجمع يضع الفئات وراء بعضها ، ويبين تراتيب القيم فى السلسلة . وبواسطة هـذا التكرار

إيجاد الوسيط لتوزيـع تكرارى كبير

المتجمع يمكن معرفة قيمة الوسيط بعد معرفة ترتيبه . ولتوضيح ذلك نأخذ التوزيع التكرارى السابق ، ونوجد التكرار المتجمع كما شرحنا فى بند ٩٨

جدول ٢٥ — التكرارات المتجمعة لأعمار ١٧٣٩ تلميذاً .

فئات العمر بالسـنين	عـدد التلاميذ (التكرار)	الحدود العليا للفئات	التكرار المتجمع (الصاعد)
١٣ر٥ أقل من ١٤ر٥	٣٤	أقل من ١٤ر٥	٣٤
١٤ر٥ « « ١٥ر٥	١٢٨	« « ١٥ر٥	١٦٢
١٥ر٥ « « ١٦ر٥	٢٦٢	« « ١٦ر٥	٤٢٤
١٦ر٥ « « ١٧ر٥	٣٦٠	« « ١٧ر٥	٧٨٤
١٧ر٥ « « ١٨ر٥	٣٨٦	« « ١٨ر٥	١١٧٠
١٨ر٥ « « ١٩ر٥	٢٩٤	« « ١٩ر٥	١٤٦٤
١٩ر٥ « « ٢٠ر٥	١٦٧	« « ٢٠ر٥	١٦٣١
٢٠ر٥ « « ٢١ر٥	٩٢	« « ٢١ر٥	١٧٢٣
٢١ر٥ « « ٢٢ر٥	١٦	« « ٢٢ر٥	١٧٣٩
الجملة	١٧٣٩		

ومن هذا الجدول نرى أن ال ٣٤ قيمة الأولى جميعها أقل من ١٤ر٥ سنة . وال ١٦٢ قيمة الأولى كلها أقل من ١٥ر٥ . وهكذا .

وعلى حسب تعريف الوسـيط ، يكون ترتيبه فى هذه السلسلة يسـاوى $\frac{١}{٢}$ (١٧٣٩ + ١) = ٨٧٠ ، أى بوضع ن = ١٧٣٩ فى العبـارة $\frac{١}{٢}$ (ن + ١) .

الآن نبحث عن الوسيط وهو القيمة التى ترتيبها ٨٧٠ فى هذه السلسـلة .

وبالتأمل فى الجدول التكرارى المتجمع نجـد أن ال ٧٨٤ قيمة الأولى كلها أقل من ٥ر١٧ سنة ، فى حين أن ال ١١٧٠ قيمة الأولى جميعها أقل من ٥ر١٨ سنة .

∴ العمر الوسيط ، وترتيبه ٨٧٠ ، لابد واقع بين ٥ر١٧ و٥ر١٨ .

ليكن ٥ر١٧ + س . ويبقى إذن معرفة قيمة المجهول س .

نقول إن فرقاً فى الترتيب من ٧٨٤ إلى ١١٧٠ يقابله فرق فى القيمة من ٥ر١٧ إلى ٥ر١٨

∴ فرق فى الترتيب من ٧٨٤ إلى ٨٧٠ يقابله فرق فى القيمة قدره س .

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

$$\frac{\text{١١٧٠} - \text{٧٨٤}}{\text{٨٧٠} - \text{٧٨٤}} = \frac{\text{١٨ر٥} - \text{١٧ر٥}}{\text{س}} \qquad (\text{١})$$

$$\therefore \quad \text{س} = \text{١} \times \frac{\text{٨٧٠} - \text{٧٨٤}}{\text{١١٧٠} - \text{٧٨٤}} \qquad (\text{٢})$$

$$= \text{٢٢٣ر٠} \ \text{تقريبا}$$

$$\therefore \quad \text{الوسيط} = \text{١٧ر٥} + \text{٢٢٣ر}$$

$$= \text{١٧ر٧٢٣} \ \text{سنة .}$$

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

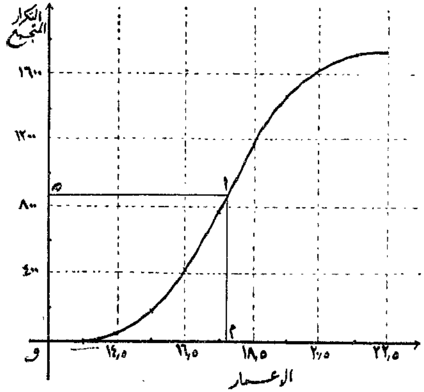
بالصدفة كان عدد القيم فى هذا المثال فردياً (= ١٧٣٩) ، ولذلك كان ترتيب الوسيط عدداً صحيحا (= ٨٧٠) . ولـكن إذا كان العدد الأصلى زوجياً فإن ترتيب الوسيط يكون عدداً كسرياً . مثلا ٥ر٨٧٠ إذا كان عدد المفردات

(١) هذا التناسب على فرض أن المفردات الموجودة فى الفئة ٥ر١٧ — ٥ر١٨ التى يقع فيها الوسـيط ، متدرجة بانتظام داخل هذه الفترة . وهـذا فرض عملى مناسب ، وفيه شىء من التقريب طبعاً . ولـكن الخطأ فى هذا صـغير ، خصوصاً إذا كانت فترة الفئة صغيرة .

١٧٤٠ بدل ١٧٣٩ . ولـكن هذا لا يؤثر فى الطريقة ، حيث نضع هذا الترتيب الـكسرى ٥ر٨٧٠ (وهو فرضى طبعاً أو وهمى) بدل الترتيب ٨٧٠ فى المعادلتين (١) و (٢) .

إيجاد الوسيط بالرسم

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



(شكل ٥٧)

تعيين الوسيط بالرسم

∴ نرسم المنحنى الصاعد لهذا التوزيع ، شكل ٥٧ ، ونعين نقطة عليه ، ا مثلا ، إحداثيها الرأسى يساوى نصف عدد القيم أى ½ × ١٧٣٩ .

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

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

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

ويتضح من شكل ٥٧ أن الوسيط = ١٧ر٧ سنة . ودرجة الدقة هنا تتوقف طبعاً على درجة الدقة فى الرسم ، ودقتنا فى قراءة البعد و م على المحور الأفقى . ولكن إذا كان الرسم دقيقاً ، وكان المنحنى ممهداً تمهيداً مضبوطاً ، ففي الإمكان أن نحصل على قيمة للوسيط أدق من القيمة التى نحصل عليها بالطريقة الحسابية المتقدم شرحها فى البند السابق . وعند ذلك لا نحتاج إلى الفرض التقريبي الذى اعتمدنا عليه (انظر الحاشية فى صفحة ١٦٢) فى تحرى موقع الوسيط داخل الفئة .

الربيعان الأدنى والأعلى

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

وأقترح تسميتهما **الربيع الأدنى**(١) و **الربيع الأعلى**(٢)

فلإيجاد الربيع الأدنى للأعمار فى هذا المثال نحسب ترتيبه وهو يساوى ¼ (١٧٣٩ + ١) ، أى بوضع ١٧٣٩ بدل ى فى العبارة ¼ (ى + ١) .

∴ ترتيب الربيع الأدنى = ٤٣٥

∴ نفس الربيع الأدنى يقع بين ١٦ر٥ و ١٧ر٥ . لأننا نرى من جدول التكرار المتجمع أن الـ ٤٢٤ قيمة الأولى كلها أصغر من ١٦ر٥ ، وأن الـ ٧٨٤ قيمة الأولى جميعها أصغر من ١٧ر٥ . فإذا فرضنا أن :

الربيع الأدنى = ١٦ر٥ + س

يكون $\frac{\text{س}}{١٧٫٥ - ١٦٫٥} = \frac{٤٣٥ - ٤٢٤}{٧٨٤ - ٤٢٤}$

∴ س = ١ × $\frac{١١}{٣٦٠}$

= ٠٣ و

∴ الربيع الأدنى = ١٦ر٥٣ سنة .

وكذلك نحسب الربيع الأعلى وترتيبه فى السلسلة يساوى ¾ (١٧٣٩ + ١) ، أى ١٣٠٥ ؛ وهو يقع فى الفئة ١٨ر٥ — ١٩ر٥ ؛ ونحسبه بنفس الطريقة . فينتج أنه يساوى ١٨ر٩٦ سنة .

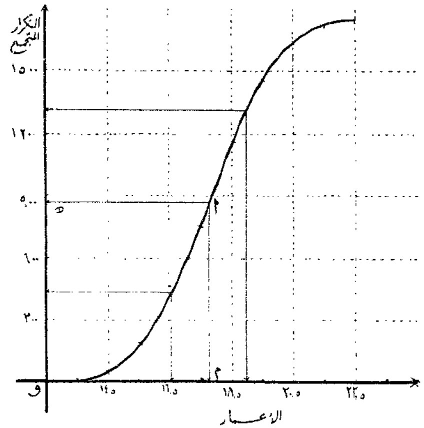
---

(١) الفعل ربع الشىء بمعنى كان فى ربعه أو قسمه الى أرباع . واسم الفاعل رابع ؛ والصفة المشبهة ربيع ، بمعنى الذى يقسم الشىء الى أرباع . وأظن أن هذه أنسب وأسهل من « الوسط الربعى » الذى اقترحها البعض .

(٢) بالانجليزية (Lower Quartile, Upper Quartile) على الترتيب

إيجاد الربيعين بالرسم

١٤٦ — يمكن الحصول على قيمتى الربيع الأدنى والأعلى من رسم المنحنى التكرارى المتجمع بنفس الطريقة التى استعملناها فى إيجاد الوسيط من نفس الرسم . فنأخذ على المحور الرأسى بعداً يساوى $\frac{1}{4}$ (١٧٣٩ + ١) ونرسم خطاً أفقياً يقابل المنحنى الصاعد فى نقطة . والإحداثى الأفقى لهذه النقطة يساوى الربيع الأدنى . وإذا أخذنا بعداً على المحور الرأسى يساوى $\frac{3}{4}$ (١٧٣٩ + ١) حصلنا على الربيع الأعلى بنفس الطريقة . ونرى ذلك موضحاً فى الشكل الآتى :



(شكل ٥٨) إيجاد الربيعين بالرسم

ويلاحظ أن إيجاد الربيعين غير ممكن بهذه السهولة من المنحنى التكرارى العادى ، لأن المنحنى التكرارى المتجمع يساعدنا فى هذه المسألة بالذات بمناسبة أن الإحداثيات الرأسية تمثل التكرارات المتجمعة ؛ فى حين أنها فى المنحنى التكرارى العادى تمثل الإحداثيات الرأسية تكرارات القيم منفردةً ، كما سبق أن أشرنا فى بند ١١١ (صفحة ١٢٥) .

العلاقة بين الربيعين والوسيط

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

الفرق بين الربيع الأعلى والوسيط لا يساوى الفرق بين الأخير والربيع الأدنى ، إلا إذا كان المنحنى التكرارى متماثلا ، بالمعنى الذى ذكرناه سابقاً (بند ١٠٢) . وهذه خاصة من خواص المنحنى التماثل ؛ وسنستعملها فيما بعد لاختبار درجة تماثل المنحنيات ، أو عدم تماثلها .

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

الربيع الأعلى — الوسيط = ١٨٫٩٦ — ١٧٫٧٢٣

= ١٫٢٣٧

الوسيط — الربيع الأدنى = ١٧٫٧٢٣ — ١٦٫٥٣

= ~~١٫٢٦٣~~ ١٫١٩٣

وهما متساويان تقريباً .

العشير والمئين

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

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

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

تعريف الوسط الهندسى

١٤٩ — الوسط الهندسى لـكميتين هو الجذر التربيعى لحاصل ضربهما . وقياساً على ذلك فالوسط الهندسى لمجموعة من الـكميات عددها ن يساوى الجذر النونى لحاصل ضرب هذه الـكميات جميعها . ويمكن وضع هـذا بصورة مختصرة كما يأتى :

نفرض أن القيم هى $\text{س}_1 ، \text{س}_2 ، \text{س}_3 ، \cdots ، \text{س}_{\text{ن}}$.

ونفرض أن الوسط الهندسى لها يساوى هـ مثلا .

$$\therefore \ \text{هـ} = \sqrt[\text{ن}]{\text{س}_1 \times \text{س}_2 \times \text{س}_3 \times \cdots \times \text{س}_{\text{ن}}}$$

$$= [\ \text{س}_1 \times \text{س}_2 \times \cdot\cdot \times \text{س}_{\text{ن}}\ ]^{\frac{1}{\text{ن}}}$$

كيفية إيجاده عمليا

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

$$\therefore \text{لو هـ} = \text{لو} [\ \text{س}_1 \times \text{س}_2 \times \cdot\cdot \times \text{س}_{\text{ن}}\ ]^{\frac{1}{\text{ن}}}$$

$$= \tfrac{1}{\text{ن}} \text{لو} [\text{س}_1 \times \text{س}_2 \times \cdot\cdot \times \text{س}_{\text{ن}}]$$

$$= \tfrac{1}{\text{ن}} [\text{لو س}_1 + \text{لو س}_2 + \text{لو س}_3 + \cdot\cdot + \text{لو س}_{\text{ن}}] ،$$

لأن لوغاريتم حاصل الضرب ، الذى فى الطرف الأيسر ، يساوى مجموع لوغاريتمات الـكميات $\text{س}_1 ، \text{س}_2 ، \cdots \text{س}_{\text{ن}}$.

وينتج إذن أن لوغاريتم الوسط الهندسى لأى عـدد من الـكميات يساوى الوسط الحسابى للوغاريتمات هذه الـكميات .

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

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

تعريف الوسط التوافقى

١٥٠ — يمكن تعريف الوسط التوافقي بين عدة كميات كما يأتى :

لتكن الـكميات الأصلية عددها ن وهى :

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

ولنرمز للوسط التوافقي بينها بالرمز ف .

العلاقة بين الوسط التوافقى ف والـكميات س هى :

مقلوب الوسط التوافقى يساوى الوسط الحسابى لمقلوبات الـكميات

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

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

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

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

٧ برتقالات بقرش كان هـذا هو الوسط التوافقى لسعر البرتقال وليس الوسط الحسابى .

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

اختلاف مفردات المجموعة فى الأهمية

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

وهـذا الإجراء نتبعه طبعاً فى كل الأحوال التى لا يكون فيها تفاضل بين المفردات فى الأهمية . ولـكن إذا أعطينا مجموعة وعلمنا أن بعض مفرداتهـا فعلا أهم من بعضها الآخر ، أو أن قيمة أرجح من أخرى لسبب من الأسباب ، وأردنا إيجاد المتوسط لهذه المجموعة ، لا يسعنا إلا نأخذ فى الاعتبار هـذا الاختلاف فى الأهمية بين المفردات . فهل نعتبر مثـلا سعر القمح فى ساحل روض الفرج بالقاهرة — وليكن ١٤٤ قرشاً للأردب مثلا — على قدم المساواة مع سعر القمح فى سوق ريفية صغيرة — وليكن ١٤٠ قرشاً — ونقول إن متوسط سعر القمح $= \frac{١٤٤+١٤٠}{٢}$ أى ١٤٢ قرشاً فى ذلك اليوم ؟ مع علمنا بأن ساحل روض الفرج يحصل فيه تعامل بآلاف الأرادب كل يوم ويجتمع فيه كبار التجار ،

وأن السوق الأخرى لا تعقد إلا يوماً كل أسبوع ولا يباع فيها إلا بعض عشرات الأرادب ؟

الوسط المرجح

١٥٢ — فى مثل هـذه الأحوال يجب أن نحسب **الوسط المرجح**(١) الذى يأخـذ فى الاعتبار الأهمية النسبية للمفردات ، حيث **نرجح** كل مفردة بما يتناسب وأهميتها بالنسبة إلى غيرها .

ويجب أن نفكر فى كيفية قياس الأهمية النسبية للمفردات المختلفة فى المجموعة ، كل حالة وما يناسبها . ومقـدار الأهميـة النسبية لكل مفردة نسميه **وزناً** (Weight) ، ونستعمله لترجيح هـذه المفردة بالنسبة إلى غيرها عنـد حساب الوسط المرجح .

تعريف الوسط المرجح

الوسط الحسابى المرجح يساوى مجموع حواصل ضرب القيم فى أوزانهـا ، مقسوماً على مجموع الأوزان . فاذا كانت القيم هى

$$\text{س}_١ ، \text{س}_٢ ، \text{س}_٣ ، \dots ، \text{س}_\text{ن}$$

وأوزانها $\text{و}_١ ، \text{و}_٢ ، \text{و}_٣ ، \dots ، \text{و}_\text{ن}$ على الترتيب ،

فان الوسط المرجح لهذه القيم حسب التعريف

$$= \frac{\text{س}_١ \cdot \text{و}_١ + \text{س}_٢ \cdot \text{و}_٢ + \dots + \text{س}_\text{ن} \cdot \text{و}_\text{ن}}{\text{و}_١ + \text{و}_٢ \qquad \dots + \text{و}_\text{ن}}$$

$$\text{أو} \quad = \frac{\sum(\text{س و})}{\sum \text{و}} .$$

لنفرض مثلا أن أسعار المواد الغذائية ارتفعت بنسبة ١٠٪ ، وأسعار

(١) (The Weighted Mean) بعض الناس يسمونه الوسط المعدل ؛ ولـكن هذه الترجمة غير مناسبة ، وتوقعنا فى التباس فى معنى التعديل .

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

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

$= ١٣٫٥٧$ ٪ .

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

اختيار الأوزان المناسبة

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

لنفرض مثلا أن أحد أحياء المدينة يحتوى على مائة منزل ، وأن كل منزل يسكن فيه عدد من الأسر ، وهذا العدد يختلف من منزل إلى آخر . والمطلوب

حساب متوسط عدد الأسر فى كل منزل .

إذا كنا فى صدد بحث المستهلك من مياه الشرب فى هذه المنازل فلا شك أن منزلا به خمس أسر مجموع أفرادها ٢٦ شخصاً أهم من هذه الناحية (مرتين تقريباً) من منزل آخر يسكن فيه خمس أسر مجموع أفرادها ١٤ شخصاً فقط . وعلى ذلك نختار الأوزان لحساب الوسط المرجح هنا تساوى عدد الأشخاص الساكنين فى كل منزل .

أما إذا كنا نبحث فيما تستهلكه الأسر من النور الكهربائى مثلا ، فإن العامل الذى له صلة أمتن بهذا الموضوع ، هو عدد الغرف فى كل منزل . حيث إن عدد المصابيح الكهربائية فى كل منزل يتناسب مع عدد غرفه ؛ واستهلاك الكهرباء يتناسب مع عدد الغرف ؛ وهو أشد ارتباطاً بهذا منه بأى عامل آخر . وعلى ذلك نختار الأوزان فى هذه الحالة تساوى عدد الغرف فى كل منزل .

الوسط المرجح يتغير بتغير الأوزان

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

لنأخذ القيم ١٥ و ١٢ و ٢٠

وأوزانها ٤ و ١ و ٢ على الترتيب .

∴ الوسط المرجح بهذه الأوزان

$$= \frac{٢ \times ٣٠ + ١ \times ١٢ + ٤ \times ١٥}{٢ + ١ + ٤} = ١٦ .$$

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

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

$$= ١٦ر١٧$$

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

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

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

$$= ١٤ر٢٨$$

ونرى هنا أن تغييراً قدره ٢٠ ٪ فى إحدى القيم غير الوسط المعدل من ١٦ إلى ١٤ر٢٧ ، أى بقدر ٨ر١٠ ٪ تقريباً . وهذا أكبر بكثير من تأثير التغيير فى الوزن .

والسبب فى ذلك واضح . فنحن إذا أنقصنا الوزن من ٤ إلى ٣ ينقص البسط تبعاً لذلك وينقص المقام أيضاً . وهذا النقص فى قيمة المقام يعوض شيئاً من النقص فى قيمة البسط ، ولا يتغير الكسر تغيراً كبيراً ؛ وكذلك الأمر إذا زاد الوزن . أما إذا أنقصنا القيمة ١٥ (أو أى قيمة أخرى) ، وتركنا الأوزان كما هى : فإن قيمة البسط تنقص وتبقى قيمة المقام كما هى . وبذلك تتغير قيمة الكسر تغيراً كبيراً . وكذلك فى حالة الزيادة .

الدقة فى تقدير الأوزان غير مهمة

والذى نستفيده من هذا أنه لا يلزم عند اختيار الأوزان أن نتحرى منتهى الدقة فى معرفة قيمها ، ولكن يكفى أن نأخذها مقربة بقدر الإمكان ؛ ولا يؤثر ذلك كثيراً فى النتائج التى نحصل عليها للسبب الذى ذكرناه . وهذا تسهيل كبير فى العمل إذ أنه يمكننا من اختصار عمليات الضرب والقسمة اللازمة بدون خطأ يذكر فى الوسط المرجح .

الأوزان عديمة الأثر

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

لنأخذ القيم ١٢ و ١٠ و ٢٠ مثلا ؛

وأوزانها ٤ ١ ٢ على الترتيب .

$$\text{الوسط الحسابى} = \frac{١٢ + ١٠ + ٢٠}{٣}$$

$$= ١٤ \text{ بدون أوزان .}$$

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

$$= ١٤ \text{ أيضاً .}$$

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

---

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

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

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

$\therefore \text{ا}(\text{ص}+\text{ع}-٢\text{س}) + \text{ب}(\text{ع}+\text{س}-٢\text{ص}) + \text{حـ}(\text{س}+\text{ص}-٢\text{ع}) = ٠$

أى $\text{ا ل} + \text{ب م} + \text{حـ ن} = ٠$ مثلا .

ولكن $\text{ل} + \text{م} + \text{ن} = ٠$

$\therefore \text{م}(\text{ب}-\text{ا}) + \text{ن}(\text{حـ}-\text{ا}) = ٠$. وبالمثل $\text{م}(\text{ب}-\text{حـ}) + \text{ل}(\text{ا}-\text{حـ}) = ٠$.

$\therefore \text{ل} : \text{م} : \text{ن} = (\text{ب}-\text{حـ}) : (\text{حـ}-\text{ا}) : (\text{ا}-\text{ب})$.

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

الوسط التوافقى المرجح

١٥٦ — لإيجاد الوسط التوافقى المرجح نضرب مقلوب كل قيمة فى وزن هذه القيمة ، ثم نقسم مجموع هذه المضروبات على مجموع الأوزان ؛ والناتج يساوى مقلوب الوسط التوافقى . فلو أخذنا القيم

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

وأوزانها

$$\text{و}_1 \text{ ، } \text{و}_2 \text{ ، } \text{و}_3 \quad \cdots \quad \text{و}_{\text{ن}} \text{ ،}$$

ورمزنا للوسط التوافقى المرجح بالحرف ق

يكون

$$\frac{1}{\text{ق}} = \frac{\frac{\text{و}_1}{\text{س}_1} + \frac{\text{و}_2}{\text{س}_2} + \cdots + \frac{\text{و}_{\text{ن}}}{\text{س}_{\text{ن}}}}{\text{و}_1 + \text{و}_2 + \cdots + \text{و}_{\text{ن}}}$$

الوسط الهندسى المرجح

١٥٧ — ولإيجاد الوسط الهندسى المرجح[1] نرفع كل قيمة إلى قوة تساوى وزنها ، ونضرب هذه القوى للقيم المختلفة ، ثم تأخذ الجذر الذى رتبته تساوى مجموع الأوزان ، لهذا الحاصل .

نفرض أن الوسط الهندسى المرجح يساوى هـ ، وأن القيم وأوزانها كما تقدم ، يكون

$$\text{هـ} = \left[(\text{س}_1)^{\text{و}_1} \times (\text{س}_2)^{\text{و}_2} \times \cdots \times (\text{س}_{\text{ن}})^{\text{و}_{\text{ن}}} \right]^{\frac{1}{\text{مجـ و}}}$$

$$\therefore \text{لو هـ} = \frac{\text{و}_1 . \text{لو س}_1 + \text{و}_2 . \text{لو س}_2 + \cdots + \text{و}_{\text{ن}} . \text{لو س}_{\text{ن}}}{\text{و}_1 + \text{و}_2 + \cdots + \text{و}_{\text{ن}}}$$

أى أن لوغاريتم الوسط الهندسى المرجح يساوى الوسط المرجح للوغاريتمات القيم الأصلية . وترجيح اللوغاريتمات يكون بضرب لوغاريتم كل قيمة فى وزن هذه القيمة ، ثم جمع هذه المضروبات وقسمة الناتج على مجموع الأوزان . ولكن هذا الوسط قليل الأهمية نظراً لكثرة تعقيده ، وصعوبة حسابه عملياً .

(١) (Weighted Geometric Mean.)

المنوال والوسيط المرجحان

١٥٨ — هذا ويمكن تعميم فكرة الترجيح لتشمل المنوال والوسيط فنحصل على « المنوال المرجح »[1] « والوسيط المرجح »[2] . فبعد اختيار الأوزان والاتفاق عليها يكون المنوال هو القيمة الأكبر وزناً من جميع القيم الأخرى . وهذا يتمشى مع الفكرة الأصلية للمنوال بمعنى أنه القيمة الأكثر شيوعاً من غيرها .

وأما الوسيط المرجح فنجده بسهولة بعد أن نرتب القيم تصاعديا ونضع أمام كل واحدة وزنها المتفق عليه . فالقيم فى هذا الوضع تشابه القيم فى الجدول التكرارى وأمامها تكراراتها . ونعتبر الأوزان هنا بمثابة التكرارات هناك ، ونجرى العمل بالطريقة المعروفة .

ولكن استعمال هذين الوسطين المرجحين نادر ، وهما لذلك قليلا الأهمية .

المقارنة بين المتوسطات المختلفة

١٥٩ — يحسن قبل أن نترك هذا الباب أن نبحث فى العلاقة بين هذه المتوسطات المختلفة والمقارنة بينها من حيث تأديتها للغرض المقصود منها : ألا وهو تمثيل المجموعات الإحصائية المختلفة .

الوسط الحسابى يتأثر بالقيم الشاذة

١٦٠ — رأينا أن الوسط الحسابى يساوى مجموع القيم مقسوما على عددها ، وأن القيم تؤخذ جميعها على علاتها بصرف النظر عما يكون بينها من تفاوت فى الأهمية . ولذلك فإن الوسط الحسابى يتأثر كثيراً بالقيم المتطرفة ، خصوصاً فى المجموعات الصغيرة العدد . وقد يعطى فكرة خطأ عن المجموعة إذا كانت إحدى قيمها بالصدفة كبيرة جداً (أو صغيرة جداً) بالنسبة لباقى القيم . ففى المجموعة الآتية للأجور مثلا :

(١) (Weighted Mode) (٢) (Weighted Median.)

٧٫٥ ؛ ٨ ؛ ٨٫٥ ؛ ٩ ، ١٠٫٥ ، ١١ ، ٢١٫٥ قرشاً ،

نلاحظ أن أغلبها حوالى ٩ قروش ، ولـكن أحد هـذه الأجور كبير جداً ، أى ٢١٫٥ قرشاً ؛ وربما كان وجوده فى هـذه المجموعة عن طريق المصادفة فقط ؛ والوسط الحسابى لهذه الأجور هو ١١ قرشاً تقريباً . وواضح أنه لا يمثل المجمـوعة تماماً ، إذ يظهر أن القيم تتجمع حول ٩ . والفرق بين ٩ و ١١ ناتج من شـذوذ الأجر ٢١٫٥ وبعده عن باقى الأجور الأخرى . ولو كانت المجموعة أكبر عدداً مما هى الآن ، ربما شملت أجراً منخفضاً جـداً (أقل من ٧٫٥ المذكور هنا) يعادل تأثير هذا الأجر الشاذ ، فيكون الوسط الحسابى أقرب إلى الحقيقة من ١١ .

عيوب الوسط الحسابى

١٦١ — والوسط الحسابى ، كما قلنا سابقاً ، لا يساعدنا فى التوزيعات التكرارية المفتوحة . وفى مثل هذه التوزيعات يجب أن نلجأ إلى الوسيط أو المنوال .

ويجب ألا نعتمد على الوسط الحسابى بمفرده فى تمثيل المجموعات غير المتجانسة ومتعددة المناويل . لأن الاقتصار على ذكر الوسط الحسابى فى مثل هـذه الأحوال — أو أى متوسط آخر — فيه تضليل ؛ ويجب أن نذكر هذه الحقيقة إلى جانب الوسط الحسابى إذا عرفنا قيمته . ولـكن الأحسن هو حساب هذه المناويل . وهـذا أوضح وأدق ، وأحسن دلالة على صفات المجموعة ، من الاقتصار على الوسط الحسابى الذى لا يمثلها تمثيلا صحيحاً .

والوسط الحسابى مضلل أيضاً فى حالة التوزيع التكرارى ذى النهاية الصغرى (بند ١٠٥ . شكل ٤٣ صفحة ١١٨) . لأنه فى هذه المجموعات يكون فى الوسط بين أكبر قيمة وأصغر قيمة . ويكون تكراره أصغر تكرار أو بالقرب منه . وعلى ذلك فلا يمكن أن نأخذه كنموذج للمجموعة .

ميزات الوسط الحسابى

١٦٢ — ومن أكبر ميزات الوسط الحسابى سهولته وبساطة الفكرة المبنى عليهـا ، ووضوحها . فمن السهل على كل إنسان فهم المعنى المقصود بالوسط الحسابى . بخلاف المتوسطات الأخرى ، فهى تجمع إلى التعقيد فى المعنى والفكرة ، صعوبة عملية فى حسابها ومعرفة مقاديرها بالدقة . ولذلك فالوسط الحسابى أكثر المتوسطات استعمالا ، رغم ما فيه من عيوب فى بعض الأحيان .

ومن الناحية النظرية أيضاً ، نجد فكرة الوسط الحسابى — أى مجموع القيم على عددها — تساعدنا كثيراً فى الأبحاث الرياضية الخاصة بهذه الموضوعات : فمن السهل استنباط خواص مهمة ومفيـدة للوسط الحسابى بطرق رياضـية يصعب تطبيقها فى حالة المتوسطات الأخرى . ولهذا نجد أن الوسط الحسابى قد تنـاولته أبحاث نظرية من نواح متعددة ، بخلاف المتوسطات الأخرى .

بعض ميزات المنوال وعيوبه

١٦٣ — نلاحظ أننا عند حساب المنوال فى أى توزيع تكرارى لا نهتم بالقيم أو الفئات المتطرفة فى الجدول . بل نعنى فقط بالثلاث فئات ذات التكرارات الـكبرى . وهذه تكون فى وسط الجدول فى العادة — إذا كان التوزيع ذا منوال واحد فقط . وهذه الخاصة التى يختص بها المنوال — والوسيط أيضاً — تساعدنا فى دراسة التوزيعات التكرارية المفتوحة من أحد الطرفين أو كليهما ، كما ذكرنا سابقاً . وهذه ميزة نحتاج إليها أحياناً .

والمنوال لا ينفعنا فى دراسة المنحنى التكرارى ذى النهاية الصغرى ، ولا فى المنحنى ذى الفرع الواحد — صاعداً كان أو هابطاً . والسبب فى ذلك واضح إذ أن هذه المنحنيات ليس لها منوال بالمعنى الذى نفهمه كما فى المنحنى التكرارى العادى . ومع أننا نجد فى كل من هذه التوزيعات التكرارية قيمة تكرارها أكبر من تكرار أى قيمـة أخرى ، إلا أنها تـكون قيمة طرفية ، لا يسبقها قيم

أو لا يليها قيم (كما نرى فى الأشكال ٤١ و ٤٢ و ٤٣) . ولكننا لا نعتمد كثيراً على القيم الطرفية فى تمثيل المجموعات .

إذا ضربنا المنوال فى عدد القيم كلها لا نحصل على مجموع هذه القيم كما فى الوسط الحسابى . فمثلا إذا عرفنا الوسط الحسابى لأجور العمال أمكننا معرفة جملة ما يكسبه هؤلاء العمال من أجور بضرب الوسط فى عددهم . ولكن لا يمكننا ذلك مع المنوال إلا إذا كان التوزيع التكرارى متماثلا ، حيث يكون المنوال يساوى الوسط الحسابى — والوسيط أيضاً .

هذا وقد ذكرنا سابقاً أن الصعوبة العملية فى حساب قيمة المنوال والأخطاء التى نتعرض لها فى ذلك ، تجعلنا لا نلجأ إليه كثيراً فى أعمالنا ؛ ولو أن الفكرة النظرية للمنوال فكرة صحيحة ووجيهة فى ذاتها .

أهم ميزات الوسيط

١٦٤ — من ميزات الوسيط أننا فى حساب قيمته لا نهتم بمقادير القيم مثل ما نهتم بترتيب هذه القيم بين بعضها : الأصغر فالأكبر . وهذه الخاصة يمكن الانتفاع بها تصحيح خطأ الوسط الحسابى عند تأثره بالقيم المتطرفة كما ذكرنا سابقاً (بند ١٦٠) . ففى مجموعة الأجور التى ذكرناها ، وهى

٧ر٥ ، ٨ ، ٨ر٥ ، ٩ ، ١٠ر٥ ، ١١ ، ٢١ر٥ قرشاً

لا نهتم بمقادير هذه الأجور عند حساب الوسيط ، بل نأخذ الأجر الأوسط منها وهو ٩ قروش هنا . وبذلك لا يتأثر المتوسط بالقيمة الشاذة ٢١ر٥ التى وجدت فى المجموعة صدفة ، وهكذا نأخذ فكرة صحيحة عن الأجور أحسن مما لو استعملنا الوسط الحسابى .

من خواص الوسيط أيضاً أن مجموع العروق (بدون إشارة) بين المفردات والوسيط يكون أقل من مجموع الفروق بين نفس المفردات وأى قيمة أخرى من قيم المجموعة ، غير الوسيط . فلو أخذنا المجموعة المذكورة أعلاه مثلا ، ووسيطها

٩ قروش ، نجد أن مجموع الفروق بدون إشارة ، بين هذه الأجور والعدد ٩ ، يساوى ١٩ . ولو أخذنا الأجر ١٠ر٥ مثلا بدل ٩ ، نجد هذا المجموع يساوى ٢٠ر٥ ؛ أو ٢٢ إذا أخذنا ٨ — ويمكن إثبات هذه الخاصة عموماً بدون صعوبة .

العلاقة بين الوسط والمنوال والوسيط

١٦٥ — فى التوزيع التكرارى المتماثل نجد أن الوسط الحسابى والوسيط والمنوال كلها متساوية . وهذه الخاصة تنتج من معنى التماثل . فهذا المنحنى ينقسم إلى نصفين متطابقين بواسطة الخط الرأسى الذى يمر بالقمة ، أى محور التماثل . وهذا يثبت أن المنوال يساوى الوسيط وهما عبارة عن مسقط قمة المنحنى على المحور الأفقى . ونظراً لتماثل توزيع التكرارات نرى أن القيم متساوية البعد عن المنوال (من أعلى ومن أسفل) متساوية التكرار أيضاً . وهذا يثبت أن المنوال هو الوسط الحسابى أيضاً .

وإذا كان الفرق بين هذه المتوسطات صغيراً فى منحنى تكرارى معين فهذا يدل على أن هذا المنحنى قريب من التماثل . ومثال ذلك المنحنى التكرارى فى شكل ٣٣ (صفحة ١٠٤ . التوزيع فى جدول ٩ صفحة ٩٩) — وجدنا أن الوسط الحسابى لهذا التوزيع التكرارى لأعمار ١٧٣٩ تلميذاً هو ١٧ر٧٤٩ سنة (صفحة ١٤٥) ، والوسيط ١٧ر٧٢٣ (صفحة ١٦٢) ، والمنوال ١٧ر٧٢ (صفحة ١٥٥) .

وعلى العموم يكون الوسيط واقعاً بين الوسط الحسابى والمنوال . وعندما يكون المنحنى التكرارى قريباً من التماثل جداً ، يكون الفرق بين الوسط والوسيط ثلث الفرق بين الوسط والمنوال تقريباً :

الوسط — الوسيط = $\frac{1}{3}$ (الوسط — المنوال) ، تقريباً

ويجب ملاحظة أن هذه العلاقة تقريبية وليس من اللازم تحققها بالضبط فى أى حالة معينة .

المراجع

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

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

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

Jones, D.C. *First Course in Statistics*, Chapters IV, V.

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

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

Secrist, H., *Statistical Methods*, Chapter IX.

الباب السابع

التشتت

معنى التشتت

١٦٦ — « التشتت »[١] فى أى مجموعة من القيم ، يقصد به التباعد بين مفرداتها أو التفاوت أو الاختلاف بينها . وهذا التشتت يكون صغيراً بالطبع إذا كان التفاوت بين مفردات القيم قليلا ، أى إذا كانت القيم قريبة من بعضها . ويكون التشتت كبيراً إذا كان التفاوت بينها كبيراً ، أى إذا كانت القيم بعيدة عن بعضها .

وعلى ذلك يمكننا أن نتخذ مقدار التشتت ، قليلا كان أو كبيراً ، كدليل على تجمع القيم وقربها من بعضها أو على تفرقها وتباعدها عن بعضها . وهكذا يكون لدينا مقياس لمقدار تجانس المجموعات الإحصائية ، أو عدم تجانسها .

لا شك أن صفة التجانس من عدمه . صفة تهمنا معرفتها فى كل مجموعة ندرسها . ومعرفة المتوسطات لا تغنينا عنها . فحبذا لو أمكننا قياس التشتت بطريقة تعبر — بدقة وبسهولة — عن هذه الصفة المهمة التى نقصدها .

مقاييس التشتت

١٦٧ — يوجد عدة طرق إحصائية لقياس التشتت تختلف فيما بينها فى الدقة والسهولة فى العمل ، وفى الأساس النظرى الذى تبنى عليه .

أسهل هذه الطرق عملياً هي طريقة قياس المدى[٢] بين أصغر وأكبر قيمة

---

(١) بالانجليزية (Dispersion)

(٢) بالانجليزية (Range)

فى المجموعة ، واعتبار طول هذا المدى كمقياس للتشتت فى هـذه المجموعة .

فالقيم الآتية مثلا :

٢٨ ، ٣١ ، ٣٢ ، ٣٢ر٥ ، ٣٣ ، ٣٥ ، ٣٥ر٥ ، ٣٦ ، ٣٧ر٥ ، ٤٠

تنتشر فى مـدى قدره من ٢٨ إلى ٤٠ ، أى ١٢ ، ونعتبرها أكثر تشتتاً من المجموعة الآتية :

٣٢ ، ٣٢ر٣ ، ٣٢ر٧ ، ٣٣ر٥ ، ٣٦ ، ٣٦ر٢ ، ٣٦ر٨ ، ٣٧ر٤ ، ٣٨ .

لأن هذه المجموعة الأخيرة تنتشر فى فترة أصغر ، وقدرها ٦ فقط ؛ وعلى ذلك فالقيم هنا أقرب إلى بعضها منها فى المجموعة الأولى .

هذه الطريقة غير دقيقة ومضللة

ولـكن هـذه الطريقة ، وإن كانت سهلة عملياً ، ليست دقيقة ؛ وأحياناً تـكون مضللة . حيث قد يسبب وجود قيمة متطرفة فى المجموعة زيادة كبيرة فى طول المدى ، يستدل منها — خطأ — على وجود تشتت كبير بين المفردات ، مع أن جميعها فى الواقع متجمعة بالقرب من بعضها ما عدا هـذه القيمة الشاذة ، كما يظهر فى المجموعة الآتية للأعمار مثلا :

١٨ ، ١٨ر٥ ، ١٩ ، ٢٠ ، ٢٠ر٥ ، ٢١ ، ٣١ سنة .

فوجود العمر الشاذ ٣١ فى المجموعة يجعل المدى ١٣ سنة ، وتبدو الأعمار كأنها مشتتة كثيراً مع أن أغلبيتها متجمعة حوالى ٢٠ سنة .

المدى بين الربيعين

١٦٨ — يمكن أن نتفادى هذا العيب بأن نحدد المـدى بحيث لا يتأثر بالقيم الشاذة والمتطرفة . لذلك نأخذ المدى بين الربيعين الأعلى والأدنى كمقياس للتشتت . وبهذه الطريقة نتخلص من تأثير القيم المتطرفة التى تكون أحياناً متأثرة بظروف خاصة لا تسرى على باقى القيم فى المجموعة .

ومن ناحية أخرى نعلم أن نصف عـدد المفردات فى المجموعة ينحصر بين

هذين الربيعين . وهذا هو الجزء المهم من المجموعة الذى يكون فيه ازدحام المفردات أكثر ما يمكن . وهو يحتوى على الأشياء المهمة فى المجموعة مثل الوسط الحسابى والمنوال والوسيط وباقى المتوسطات . وعلى ذلك فاتخاذ هذه المنطقة لقياس التشتت أحسن من الطريقة المتقدمة .

وفى العـادة نأخذ مقياس التشتت يساوى نصف المـدى بين الربيعين [1] ، بدلا من المدى كله .

وحساب هـذا المقياس سهل طبعاً ، حيث نحسب الربيعين الأعلى والأدنى بإحدى الطرق المعروفة ، ونأخـذ نصف الفرق بينهما . ففى التوزيع التـكرارى لأعمار ١٧٣٩ المذكور في جدول ٢٥ (صفحة ١٦١) مثلا ، وجدنا (ص ١٦٥) أن الربيع الأدنى للأعمار يسـاوى ١٦ر٥٣ سـنة والربيع الأعلى لها يساوى ١٨ر٩٦ سنة . ومقياس التشتت بهذه الطريقة يساوى إذن

$\frac{1}{2}$ (١٨ر٩٦ — ١٦ر٥٣) = ١ر٢١٥ سنة

أما بالطريقة السابقة ، فنجد من الجدول التكرارى (جدول ٢٥) أن أصغر قيمة فى المجموعة كلها هي ١٣ر٥ سنة ، وأكبر قيمة فى الجدول هى ٢٢ر٥ سنة . وعلى ذلك فالمدى بين أصـغر وأكبر قيمة يساوى ٩ سـنين ، وهو مقياس آخر للتشتت فى هذه المجموعة .

اختلاف مقاييس التشتت

١٦٩ — يلاحظ هنا اختلاف مقياسى التشتت لنفس المجموعة : أحـدهما يساوى ٩ سنين والآخر ١ر٢١٥ سـنة فقط . وهذا هو الواجب أن يكون رغم ما يمكن أن يقال بوجوب تطابق النتيجتين باعتبار أنهما مقياس لشىء واحد وهو تشتت مجموعة واحـدة . ولـكن السبب فى هـذا الاختلاف هو أننا نقيس هذا

---

(١) اسمه بالانجليزية (Semi-Inter-Quartile Range.)

الشىء الواحد على أسس مختلفة . كما لو أردنا مثلا قياس درجة النمو فى مجموعة من التلاميذ . فقد نأخذ طول قامة الشخص كدليل على درجة نموه ؛ وهذا يقاس بالسنتيمترات . وقد نأخذ وزنه كدليل على درجة النمو أيضاً ؛ ونقيس الوزن بالكيلو جرام أو الرطل مثلا . وطبعاً لا ننتظر أن يكون الرقم الدال على وزن الشخص هو نفس الرقم الدال على طوله ، باعتبار أنهما قياسان لشىء واحد وهو درجة النمو .

وهذا ينطبق أيضاً على الطرق التى سنذكرها لقياس التشتت . فهى طرق مبنية على أسس مختلفة ، ولا ينتظر أن تؤتى نفس النتائج .

والمهم فى هذه المسألة أننا إذا أردنا مقارنة التشتت فى مجموعتين ، يجب أن نقيسه فيهما بطريقة واحدة ثم نقارن النتيجتين . ولا نقارن مقياس التشتت فى إحداها محسوباً بطريقة معينة ، بمقياس التشتت فى المجموعة الثانية محسوباً بطريقة أخرى .

قياس التشتت حول قيمة معينة

١٧٠ — فى الطريقتين المتقدمتين نقيس تشتت القيم فيما بينها . ولكننا فى دراسة المتوسطات ، التى نستخدمها كنماذج تمثل المجموعات ، نحتاج إلى معرفة تشتت المجموعة حول هذه المتوسطات ، وخصوصاً الوسط الحسابى .

وقياس التشتت حول أى قيمة معينة نتخذ هذه القيمة كمركز ، ونبحث فى الفروق بينها وبين مفردات القيم كل على حدة . والتشتت حول هذه القيمة يكون كبيراً أو صغيراً حسب ما تكون هذه الفروق كبيرة أو صغيرة فى مجموعها . وهناك فكرتان للبحث فى هذه الفروق واستخدامها لقياس التشتت . والآن نشرح هاتين الفكرتين والطريقة العملية لتطبيق كل منهما .

الفرق بين أى قيمة فى مجموعة ومتوسط هذه المجموعة نسميه عادة **انحراف** هذه النقطة عن المتوسط [1]

(١) بالانجليزية (Deviation from the Mean.)

طريقة الانحراف المتوسط

١٧١ — الطريقة الأولى هى المسماة بطريقة **الانحراف المتوسط** [1] وفيها نوجد انحراف كل قيمة فى المجموعة عن الوسط الحسابى لها ؛ ونجرد هذه الانحرافات من الإشارات الجبرية ونجمعها ؛ ثم نقسم حاصل الجمع على عدد هذه الانحرافات . وبالطبع هذا العدد يساوى تماماً عدد القيم الأصلية فى المجموعة . وبذلك نحصل على الانحراف المتوسط المطلوب وهو مقياس لتشتت المجموعة حول وسطها الحسابى .

لنأخذ مثلا مجموعة الأجور الآتية (بالقروش) :

٩   ١٠ر٥   ١١   ١٢ر٥   ١٣   ١٣ر٥   ١٤ر٥

الوسط الحسابى لهذه الأجور ١٢ قرشاً وانحرافاتها عن هذا الوسط ، بدون إشارات ، هى على الترتيب :

٣   ١ر٥   ١   ٠ر٥   ١   ١ر٥   ٢ر٥

والوسط الحسابى لهذه الانحرافات (مجموعها مقسوماً على عددها) . يساوى ١ر٥٧ وهو الانحراف المتوسط المطلوب ، وهو التشتت حول الوسط الحسابى .

والسر فى أننا أهملنا إشارات هذه الانحرافات ، وهو أننا ننظر إلى الانحراف باعتباره مجرد فرق بين القيمة والمتوسط ، بصرف النظر عن كون هذا الفرق بالنقص أو بالزيادة ، لأن التشتت الذى نريد قياسه لا يميز بين النقص والزيادة عن المتوسط ، بل يهتم فقط بمقدار البعد عنه .

الانحراف المتوسط فى التوزيع التكرارى

١٧٢ — فى المثال السابق أخذنا مجموعة من القيم عددها صغير . والآن نشرح كيفية إيجاد الانحراف المتوسط لمجموعة كبيرة مقسمة إلى فئات فى جدول تكرارى .

(١) بالانجليزية (Mean Deviation)

لنأخذ نفس التوزيع التكرارى لأعمار ١٧٣٩ تلميذاً ، ونحسب الانحراف المتوسط للأعمار (عن الوسط الحسابى) .

جدول ٢٦ — إيجاد الانحراف المتوسط للأعمار

الأعمار	التكرارات ك	الانحرافات عن ١٨	انحرافات ح بدون إشارة	ضرب ح . ك
١٤	٣٤	—٤	٤	١٣٦
١٥	١٢٨	—٣	٣	٣٨٤
١٦	٢٦٣	—٢	٢	٥٢٤
١٧	٣٦٠	—١	١	٣٦٠
١٨	٣٨٦	٠	٠	٠
١٩	٢٩٤	١	١	٢٩٤
٢٠	١٦٧	٢	٢	٣٣٤
٢١	٩٢	٣	٣	٢٧٦
٢٢	١٦	٤	٤	٦٤
	١٧٣٩			٢٣٧٢

الوسط الحسابى لهذه الأعمار يساوى ١٧٫٧٤٩ سنة (كما رأينا فى بند ١٣٠ صفحة ١٤٥) . فلإيجاد الانحراف المتوسط يجب أن نطرح هذا الوسط الحسابى من كل من القيم ١٤ و ١٥ و ١٦ و ٠٠ و ٢١ و ٢٢ ، لكى نحصل على انحرافاتها عنه (وهي ، بدون إشارة ، = ٣٫٧٤٩ ، ٢٫٧٤٩ ، ١٫٧٤٩ ، ٠٠٠ ، ٣٫٢٥١ ، ٤٫٢٥١ على الترتيب) . ثم بعد ذلك نضرب كلا من هذه الانحرافات فى التكرار المناظر له ، ونجمع ونقسم على ١٧٣٩ ينتج الانحراف المتوسط .

ولكن هذه الطريقة تكون متعبة جداً لأن هذه الانحرافات كثيرة الأرقام

وعمليات ضربها فى تكراراتها تكون مرهقة . ولذلك نستخدم وسطا فرضياً كما فعلنا فى إيجاد الوسط الحسابي . نحسب الانحرافات عن هذا الوسط الفرضى ، ونجردها من إشاراتها ؛ ثم نضربها فى التكرارات المناظرة ، ونجمع الحواصل كما هو مبين فى الجدول السابق .

العدد الناتج ، وهو ٢٣٧٢ ، يساوى مجموع انحرافات القيم المعطاة عن الوسط الفرضى ١٨ . ولكن المطلوب هو مجموع الانحرافات عن الوسط الحسابى ١٧٫٧٤٩ . وعلى ذلك يجب تصحيح الرقم ٢٣٧٢ الذى حصلنا عليه . وهذا التصحيح يكون كما يأتى :

الأعمار التى أصغر من الوسط الحسابى ، وهى ١٤ و ١٥ و ١٦ و ١٧ (ومجموع تكراراتها = ٣٤ + ١٢٨ + ٢٦٢ + ٣٦٠ = ٧٨٤) طرحناها فى الجدول من العدد ١٨ . وكان يجب طرحها من ١٧٫٧٤٩ ، وعلى ذلك فكل من الانحرافات ح التى حصلنا عليها فى الجدول أمام هذه الأعمار أكبر من اللزوم بمقدار ٫٢٥١ (أى ١٨٫٠٠ — ١٧٫٧٤٩) . ومجموعها إذن أكبر من اللزوم بمقدار ٧٨٤ × ٠٫٢٥١ .

أما الأعمار التى أكبر من الوسط الحسابى ، وهى ١٨ و ١٩ و ٢٠ و ٢١ و ٢٢ (ومجموع تكراراتها = ٣٨٦ + ٢٩٤ + ١٦٧ + ٩٢ + ١٦ = ٩٥٥) فقد طرحنا منها العدد ١٨ فى الجدول . وكان يجب أن نطرح منها ١٧٫٧٤٩ . فالانحرافات ح التى حصلنا عليها أمام هذه الأعمار ، كل منها أقل من اللزوم بمقدار ٫٢٥١ (أى الفرق بين الوسطين الفرضى والحسابى) ومجموعها إذن أقل من اللزوم بمقدار ٩٥٥ × ٠٫٢٥١ .

ويكون إذن المجموع الذى حصلنا عليه أكبر من اللزوم بمقدار ٧٨٤ × ٠٫٢٥١ ، وأقل من اللزوم بمقدار ٩٥٥ × ٠٫٢٥١ .

∴ ٢٣٧٢ — ٧٨٤ × ٢٥١ر + ٩٥٥ × ٢٥١ر .

هو مجموع الانحرافات عن الوسط الحسابى المطلوب .

وعلى العموم يمكننا استنباط القاعدة الآتية من هذا المثال المتقدم وهى :

لنفرض أن :

س̄ = الوسط الحسابى لأى توزيع تكرارى	= ١٧ر٧٤٩	هنا
و = الوسط الفرضى لهذا التوزيع	= ١٨ر٠	«
ح = الوسط الحسابى — الوسط الفرضى ، أى س̄ — و	= —٢٥١ر٠	«
ك$_1$ = عدد القيم التى هى أصغر من س̄	= ٧٨٤	«
ك$_2$ = « « « « أكبر من س̄	= ٩٥٥	«
ﺣ = مجموع الانحرافات عن الوسط الحسابى (بدون إشارة)	= ؟	«
م = « « « « الفرضى	= ٢٣٧٢	«

∴ ﺣ = م + ح (ك$_1$ — ك$_2$) . . . . (١)

وبالتعويض عن م وح ، ك$_1$ وك$_2$ بقيمها فى هذا المثال ،

∴ ﺣ = ٢٣٧٢ — ٢٥١ر (٧٨٤ — ٩٥٥)

= ٢٣٧٢ + ٤٢ر٩٢١

= ٢٤١٤ر٩٢١

∴ الانحراف المتوسط = ٢٤١٤ر٩٢١ ÷ ١٧٣٩

= ١ر٣٩ سنة ؛

وهو مقياس التشتت المطلوب :

يجب أن نختار الوسط الفرضى فى نفس الفئة مع الحسابى

١٧٣ — عند تطبيق هذه العلاقة (١) المذكورة أعلاه ، يجب أن يكون الوسط الحسابى والوسط الفرضي كلاهما داخل حدود فئة واحدة ، وإلا كان هناك ثلاث مجموعات من القيم : مجموعة أصغر من الوسط الحسابى والفرضى معاً ،

ومجموعة أخرى أكبر من الوسط الحسابى والفرضى معاً ، ومجموعة ثالثة واقعة بين الوسطين أكبر من أحدهما وأقل من الآخر . وفى هذه المجموعات يكون التصحيح مخالفاً للنظام الذى ذكرناه فى هذا المثال واستنبطنا منه القاعدة المشار إليها .

الانحراف المتوسط عن الأوساط الأخرى

١٧٤ — يمكن بنفس هذه الطريقة طبعاً أن نحسب الانحراف المتوسط عن الوسيط أو المنوال أو أى متوسط آخر ، مع مراعاة وضع قيمة هذا المتوسط بدل قيمة الوسط الحسابى فى العلاقة (١) بند ١٧٢ ، ومراعاة الشرط المذكور فى بند ١٧٣ .

وإذا حسبنا الانحراف عن الوسيط فسنجده أقل من الانحراف المتوسط عن الوسط الحسابى أو أى متوسط آخر . وقد ذكرنا السبب فى ذلك فى آخر الباب السابق (بند ١٦٤) ، وهو أن مجموع أبعاد القيم فى مجموعة عن وسيطها (وهو مجموع الانحرافات بدون إشارة) أقل من مجموع أبعادها عن أى متوسط آخر . وبالتالى يكون الانحراف المتوسط عن الوسيط له نفس الصفة .

تعريف الانحراف المعيارى

١٧٥ — أهم مقاييس التشتت وأكثرها استعمالا هو المسمى **الانحراف المعيارى** (١) . وهو مقياس للتشتت حول الوسط الحسابى . وهو عبارة عن الجذر التربيعى لمتوسط مربعات انحرافات القيم عن الوسط الحسابى .

لنفرض على العموم أن عدد القيم فى مجموعة إحصائية هو ن ، وأن هذه القيم هى :

س$_1$ ، س$_2$ ، س$_3$ ، . . . . ، س$_ن$

(١) يسمى بالإنجليزية (Standard Deviation) . ويسمى أحياناً جذر متوسط مربع الخطأ (Root-Mean-Square Error)

وليكن $\bar{\text{س}}$ هو الوسط الحسابى لهذه المجموعة .

∴ انحرافات هذه القيم عن الوسط الحسابى تكون :

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

نربع هذه الانحرافات ونقسم مجموع هـذه المربعات على عددها ، وهو ن طبعاً ، فنحصل على متوسط مربعات الانحرافات عن الوسط الحسابى ، وهو

$$= \frac{1}{\text{ن}}\left[(\text{س}_1 - \bar{\text{س}})^2 + (\text{س}_2 - \bar{\text{س}})^2 + \cdots + (\text{س}_\text{ن} - \bar{\text{س}})^2\right]$$

$$= \frac{\sum (\text{س} - \bar{\text{س}})^2}{\text{ن}} \quad \ldots\ldots\ldots\ldots \quad (1)$$

∴ الانحراف المعيارى لقيم س ، ونرمز له بالحرف ع[1] ، هو

$$\text{ع} = \sqrt{\frac{\sum (\text{س} - \bar{\text{س}})^2}{\text{ن}}} \quad \ldots\ldots\ldots\ldots \quad (2)$$

الكمية ع^2 نسميها التباين[2].

١٧٦ — السر فى تربيع الانحرافات هنا هو أننا نريد التخلص من الإشارات السالبة فى بعض الانحرافات ، وهى التى أهملناها عند حساب الانحراف المتوسط ، للسبب الذى ذكرناه حينئذ (بند ١٦٠)

والسبب فى أخذ الجذر التربيعى فى المعادلة (٢) لإيجاد الانحراف المعيارى ، هو أننا نريد الرجوع إلى الوحدات الأصلية بعد تربيع الانحرافات . وذلك ليكون

---

(١) فى الكتب الأفرنجية يرمز للانحراف المعيارى بالحرف الأغريقى «σ» واسمه سيجما (Sigma) .

(٢) بالإنجليزية (Variance) ؛ ويرمز له بالحرف الاغريقى μ (ميو) .

التشتت مقيساً بنفس الوحدات المقيسة بها القيم المعطاة س_1 ، س_2 ، . . . ، $\text{س}_\text{ن}$. وقد لاحظنا فى المثال الذى أخذناه فى بند ١٧٢ أن مقياس تشتت الأعمار معبر عنه بالسنين ، وهى نفس الوحدات المقيسة بها الأعمار .

الانحراف المعيارى لتوزيع تكرارى

١٧٧ — المعادلة (٢) فى بند ١٦٤ تعبر عن الانحراف المعيارى لمجموعة من المفردات س_1 ، س_2 ، . . . ، $\text{س}_\text{ن}$ بصرف النظر عن كونها تحتوى على القيم المتساوية أى المتكررة ، أو أن جميع القيم مختلفة . وطبعاً إذا كان هناك عدد من القيم المتساوية ، أى قيمة متكررة عددا معيناً من المرات ، فإننا نضرب مربع انحراف هذه القيمة فى عدد مرات التكرار ؛ ومجموع حواصل الضرب يساوى مجموع مربعات انحرافات هذه القيم المتساوية .

وعلى ذلك فلإيجاد الانحراف المعيارى فى جدول تكرارى ، نضرب تكرار كل فئة فى مربع انحراف مركزها عن الوسط الحسابى ؛ ومجموع هذه الحواصل يساوى مجموع مربعات انحرافات القيم الأصلية عن الوسط . و بقسمة هذا على عدد المفردات كلها ، أى مجموع التكرارات ، واستخراج الجذر التربيعى للناتج ، نحصل على الانحراف المعيارى المطلوب .

ولتوضيح ذلك عملياً نحسب الانحراف المعيارى للتوزيع التكرارى الذى أخذناه سابقاً ، وهو توزيع أعمار ١٧٣٩ تلميذاً .

ترى الخطوات موضحة في جدول ٢٧ ، حيث نجد فى العمود (١) الحدود الدنيا لفئات الأعمار ؛ وفى العمود (٢) نجد مراكز هذه الفئات ؛ وفى العمود (٣) نجد التكرارات ك . وبما أن الوسط الحسابى لأعمار هؤلاء التلاميذ هو ١٧٫٧٤٩ سنة (كما وجدنا فى بند ١٣١ صفحة ١٤٦) ، فلإيجاد انحرافات القيم عن هذا الوسط ، نطرح ١٧٫٧٤٩ من القيم ١٤ ، ١٥ ، ١٦ ، . . . على التوالى ، فنحصل على الانحرافات المكتوبة فى عمود (٤) .

ولكننا نرى أن هذه الانحرافات عن الوسط الحسابى مكونة من أرقام كثيرة وتربيعها سيعطينا أرقاماً أكثر تعقيداً . هذا فضلا عن أن عمليات ضرب هذه الأعداد الكبيرة فى نفسها، ثم ضرب النواتج فى التكرارات، تكون شاقة وتستغرق وقتاً طويلا . ولذلك نختصر العمل ونترك هذه الانحرافات عن الوسط

جدول ٢٧ — حساب الانحراف المعيارى لأعمار ١٧٣٩ تلميذاً من الجدول التكرارى .

الفئات	مراكز الفئات س	التكرارات ك	انحرافات س عن الوسط الحسابى	انحرافات عن وسط فرضى : ع	(٣) × (٥) ع × ك	(٦) × (٥) ع ٢ . ك
(١)	(٢)	(٣)	(٤)	(٥)	(٦)	(٧)
١٣ر٥ —	١٤	٣٤	٣ر٧٤٩ —	٤ —	١٣٦ —	٥٤٤
١٤ر٥ —	١٥	١٢٨	٢ر٧٤٩ —	٣ —	٣٨٤ —	١١٥٢
١٥ر٥ —	١٦	٢٦٢	١ر٧٤٩ —	٢ —	٥٢٤ —	١٠٤٨
١٦ر٥ —	١٧	٣٦٠	ر٧٤٩ —	١ —	٣٦٠ —	٣٦٠
١٧ر٥ —	١٨	٣٨٦	ر٢٥١	٠	٠	٠
١٨ر٥ —	١٩	٢٩٤	١ر٢٥١	١	٢٩٤	٢٩٤
١٩ر٥ —	٢٠	١٦٧	٢ر٢٥١	٢	٣٣٤	٦٦٨
٢٠ر٥ —	٢١	٩٢	٣ر٢٥١	٣	٢٧٦	٨٢٨
٢١ر٥ -	٢٢	١٦	٤ر٢٥١	٤	٦٤	٢٥٦
		١٧٣٩			٤٣٦ —	٥١٥٠

الحسابى ، ونختار وسطاً فرضياً مناسباً . ونحسب الانحرافات عنه ، كما فعلنا عند

حساب الانحراف المتوسط ، ونجد فى العمود (٦) حواصل ضرب هذه الانحرافات فى التكرارات ؛ وفى عمود (٧) نجد حواصل ضرب مربعاتها فى التكرارات ؛ ومن هذا نحصل على مجموع مربعات الانحرافات عن هذا الوسط الفرضى ، وهو ٥١٥٠

ولمعرفة مجموع مربعات الانحرافات عن الوسط الحسابى ، نستخدم العلاقة[١] الآتية :

(١) لإثبات هذه العلاقة نفرض أن مفردات القيم هى : س_1 ، س_2 ، … $\text{س}_\text{ن}$ ؛ وأن الوسط الحسابى لها يساوى $\bar{\text{س}}$ ؛ وأن الوسط الفرضى يساوى و مثلا .

$\therefore \bar{\text{س}} = \frac{1}{\text{ن}} (\text{مج س})$ ؛ أى $\text{مج س} - \text{ن}\bar{\text{س}} = \cdot$.

نضع $\text{ح} = \bar{\text{س}} - \text{و}$ = الفرق بين الوسطين الحسابى والفرضى .

$\therefore$ انحراف القيمة س_1 عن الوسط الفرضى هو $\text{س}_1 - \text{و}$.

لكن $\text{س}_1 - \text{و} = \text{س}_1 - \bar{\text{س}} + \bar{\text{س}} - \text{و} = \text{س}_1 - \bar{\text{س}} + \text{ح}$

$\therefore (\text{س}_1 - \text{و})^2 = (\text{س}_1 - \bar{\text{س}})^2 + \text{ح}^2 + 2\text{ح}(\text{س}_1 - \bar{\text{س}})$ ؛

وبالمثل $(\text{س}_2 - \text{و})^2 = (\text{س}_2 - \bar{\text{س}})^2 + \text{ح}^2 + 2\text{ح}(\text{س}_2 - \bar{\text{س}})$ ؛

» . . . . .

» $(\text{س}_\text{ن} - \text{و})^2 = (\text{س}_\text{ن} - \bar{\text{س}})^2 + \text{ح}^2 + 2\text{ح}(\text{س}_\text{ن} - \bar{\text{س}})$

وبجمع هذه المتساويات

$\therefore \text{مج}(\text{س} - \text{و})^2 = \text{مج}(\text{س} - \bar{\text{س}})^2 + \text{ن}\text{ح}^2 + 2\text{ح}(\text{مج س} - \text{ن}\bar{\text{س}})$.

ولكن $\text{مج س} - \text{ن}\bar{\text{س}} = \cdot$. و $\text{مج}(\text{س} - \text{و})^2 = \text{ن}\text{ل}^2$

و $\text{مج}(\text{س} - \bar{\text{س}})^2 = \text{ن}\text{ع}^2$

$\therefore \text{ن}\text{ل}^2 = \text{ن}\text{ع}^2 + \text{ن}\text{ح}^2$

أو $\text{ل}^2 = \text{ع}^2 + \text{ح}^2$

وينتج من هذا أيضاً أن $\text{ن}\text{ل}^2 > \text{ن}\text{ع}^2$ دائماً ، إلا إذا كان $\text{ح} = \cdot$.

أى أن مجموع مربعات الانحرافات عن الوسط الحسابى دائماً أصغر من مجموع مربعات الإنحرافات عن أى وسط فرضى ، إلا إذا كان الوسطان متساويين . وهذه خاصة للوسط الحسابى تقابلها خاصة مثلها للوسيط ذكرناها فى بند ١٧٤ .

دل² = دع² + دح²

حيث دل² = مجموع مربعات انحرافات القيم عن الوسط الفرضى = ٥١٥٠ هنا

و دع² = « « « « « الحسابى = ؟ «

و ح = الفرق بين الوسطين (الحسابى — الفرضى) = — ر٢٥١ «

و د = عدد المفردات جميعها ، أى مجموع التكرارات = ١٧٣٩ «

∴ ٥١٥٠ = ١٧٣٩ع² + ١٧٣٩ (— ر٢٥١)²

∴ $\frac{٥١٥٠}{١٧٣٩}$ = ع² + ر٠٦٣٠٠١

∴ ع² = ٢و٨٩٨٥

∴ ع = ١و٧٠٢ سنة ؛

وهو الانحراف المعيارى للأعمار فى هذه المجموعة .

نلاحظ أن النسبة بين الانحراف المتوسط (بند ١٧٢) والانحراف المعيارى لهذه المجموعة ، تساوى ⅘ تقريباً . وهذه خاصة معروفة من خواص التوزيعات المتماثلة أو القريبة من التماثل .

طريقة تغيير الوحدات

١٧٨ — إذا كانت فترات الفئات فى التوزيع التكرارى غير الواحد الصحيح فيمكننا أن نستخدم وحدات جديدة للانحرافات ، كما فعلنا فى بند ١٣٢ (صفحة ١٤٧) . ولبيان ذلك نأخذ نفس التوزيع التكرارى للأجور المذكور فى جدول ٢٢ هناك .

فى هذا الجدول نجد فئات الأجور فى عمود (١) ومراكزها فى عمود (٢) وتكراراتها فى العمود (٣) . ونجد الانحرافات عن الوسط الفرضى ١٩ر٥ ، فى العمود (٤) . ونلاحظ هنا أن هذه الانحرافات تحتوى على العدد ٣ كعامل

مشترك (وهو يساوى طول فترة الفئات) ؛ فنقسم هذه الانحرافات على هذا العامل ، فنحصل على الانحرافات الجديدة فى العمود (٥) ، وهى مقيسة بوحدات كل منها تساوى ثلاثة قروش .

جدول ٢٧ — حساب الانحراف المعيارى لأجور

٧٤٣٣ عاملا ، بالطريقة المختصرة

فئات الأجر بالقرش	مراكز الفئات	التكرار ك	الانحرافات عن وسط فرضى : ح	الانحرافات الجديدة . ف	(٣)×(٥) ك × ف	(٥)×(٦) ك . ف²
(١)	(٢)	(٣)	(٤)	(٥)	(٦)	(٧)
١٢—	١٣ر٥	٢٢٤	—٦	—٢	—٤٤٨	٨٩٦
١٥—	١٦ر٥	١٩٧١	—٣	—١	—١٩٧١	١٩٧١
١٨—	١٩ر٥	٣٧٥٥	٠	٠	٠	٠
٢١—	٢٢ر٥	١٢٣٦	٣	١	١٢٣٦	١٢٣٦
٢٤—	٢٥ر٥	١٩٦	٦	٢	٣٩٢	٧٨٤
٢٧—	٢٨ر٥	٤٠	٩	٣	١٢٠	٣٦٠
٣٠—	٣١ر٥	١٠	١٢	٤	٤٠	١٦٠
		٧٤٣٢			—٦٣١	٥٤٠٧

ونجد فى العمود (٦) حواصل ضرب هذه الانحرافات فى التكرارات . ومجموع هذا العمود نستخدمه لإيجاد الفرق بين الوسط الفرضى والوسط الحسابى ، ولمعرفة هذا الأخير .

ونستعمل أرقام هذا العمود أيضاً كخطوة أولى للحصول على أرقام العمود (٧) ، وهى عبارة عن مربعات الانحرافات مضروبة فى التكرارات . ومجموع هذا العمود هو مجموع مربعات الانحرافات عن الوسط الفرضى ، بوحدات جديدة .

ولإيجاد الفرق بين الوسطين بالوحدات الجديدة ، نقسم مجموع العمود (٦) على مجموع التكرارات فى العمود (٣) .

$\therefore$ الفرق بين الوسطين = ح

$$= -\frac{\text{٦٣١}}{\text{٧٤٣٢}}$$

= − ٠٨٤٩ ر بالوحدات الجديدة .

وباستخدام العلاقة $\text{ن ل}^{2} = \text{ن ع}^{2} + \text{ن ح}^{2}$ ،

$\therefore \text{٥٤٠٧} = \text{٧٤٣٢ ع}^{2} + \text{٧٤٣٢}\,(-\,\text{٠٨٤٩ ر})^{2}$ بالوحدات الجديدة

$$\therefore \frac{\text{٥٤٠٧}}{\text{٧٤٣٢}} = \text{ع}^{2} + \text{٠٠٧٢ ر}$$

$$\therefore \text{ع}^{2} = \text{٧٢٧٦ ر} - \text{٠٠٧٢ ر}$$

$$= \text{٧٢٠٤ ر}$$

$\therefore$ ع = ٨٥ ر تقريباً ، بالوحدات الجديدة .

$\therefore$ = ٣ × ٨٥ ر « الأصلية

= ٥٥ ر ٢ قرشاً ؛

وهو الانحراف المعيارى للأجور .

مجموع مربعات القيم

١٧٩ — ويمكننا إيجاد الانحراف المعيارى لمجموعة من القيم إذا عرفنا مجموع مربعات هذه القيم ، والوسط الحسابى لهذه القيم . ولبيان ذلك نفرض أن القيم هى على العموم :

$$\text{س}_{1}\text{ ، س}_{2}\text{ ، س}_{3}\text{ ، . . . . . ، س}_{\text{ن}}\text{ ؛}$$

وأن وسطها الحسابى $\bar{\text{س}}$ ، حيث $\text{ن}\,\bar{\text{س}} = \text{مج س}$ ؛ وأن الانحراف المعيارى $\text{ع}_{\text{س}}$. انحراف القيمة س_1 عن الوسط الحسابى يساوى $\text{س}_1 - \bar{\text{س}}$.

$$\therefore (\text{س}_1 - \bar{\text{س}})^{2} = \text{س}_1^{2} + \bar{\text{س}}^{2} - \text{٢ س}_1\bar{\text{س}}$$

$$\therefore (\text{س}_2 - \bar{\text{س}})^{2} = \text{س}_2^{2} + \bar{\text{س}}^{2} - \text{٢ س}_2\bar{\text{س}}$$

. . . . .

$\therefore (\text{س}_{\text{ن}} - \bar{\text{س}})^{2} = \text{س}_{\text{ن}}^{2} + \bar{\text{س}}^{2} - \text{٢ س}_{\text{ن}}\bar{\text{س}}$ ؛ وبالجمع

$$\therefore \text{مج}(\text{س} - \bar{\text{س}})^{2} = \text{مج س}^{2} + \text{ن}\,\bar{\text{س}}^{2} - \text{٢}\,\bar{\text{س}}\,\text{مج س} .$$

$$\therefore \text{ن ع}^{2} = \text{مج س}^{2} + \text{ن}\,\bar{\text{س}}^{2} - \text{٢}\,\bar{\text{س}}\,(\text{ن}\,\bar{\text{س}}) .$$

$$\therefore \quad = \text{مج س}^{2} - \text{ن}\,\bar{\text{س}}^{2} \text{ ، . . . } (١) .$$

$$\text{أى}\quad \frac{1}{\text{ن}}\,\text{مج س}^{2} = \text{ع}^{2} + \bar{\text{س}}^{2} \text{ ، . . . } (٢) .$$

وهذه المعادلة (٢) مهمة ، وسنحتاج إلى استخدامها فى كثير من الأحيان . ويلاحظ أنه من الممكن استنباطها من المعادلة المذكورة فى البند السابق ، بوضع الوسط الفرضى و = ٠ ، وعليه يكون الفرق بين الوسطين ح = $\bar{\text{س}}$ − ٠ = $\bar{\text{س}}$.

ويتضح من هذا أننا لو علمنا مجموع مربعات القيم ومتوسطها الحسابى ، أمكننا إيجاد انحرافها المعيارى بسهولة .

الانحراف المعيارى لمجموعة كبيرة مركبة من مجموعات صغيرة

١٨٠ — إذا علمنا الانحرافات المعيارية لعدد من المجموعات ، أمكننا استنباط الانحراف المعيارى للمجموعة الكلية المكونة من هذه المجموعات . فمثلا لو تكونت مدرسة من أربع فرق ، وعلمنا الانحراف المعيارى لأعمار كل فرقة على حدة ، وكذلك الوسط الحسابى لأعمارها ، وعدد التلاميذ فيها ، يمكننا بسهولة استنباط الانحراف المعيارى للأعمار فى المدرسة كلها من هذه المعلومات . وذلك باستخدام العلاقة الآتية ، وهى تربط الانحراف المعيارى العمومى (للمجموعة الكلية) بالانحرافات المعيارية للمجموعات الجزئية ومتوسطاتها الحسابية . وهذه العلاقة هى :

$$\text{ه}\,\text{ع}_{\text{س}}^{\text{٢}} = \text{مجـ}\,\text{ن}\,\text{ع}^{\text{٢}} + \text{ه}\,\text{ع}_{\text{م}}^{\text{٢}} \quad . \; . \; . \; . \qquad (\text{١})$$

حيث ه = العدد الإجمالى للمفردات فى المجموعة الكلية ،

و $\text{ع}_{\text{س}}$ = الانحراف المعيارى للمجموعة الكلية ،

و $\text{ن}_{\text{١}}$ ، $\text{ن}_{\text{٢}}$ ، $\text{ن}_{\text{٣}}$ ، . . . ، $\text{ن}_{\text{ر}}$ عدد المفردات فى المجموعات الجزئية الأولى والثانية . . ،

و $\text{ع}_{\text{١}}$ ، $\text{ع}_{\text{٢}}$ ، $\text{ع}_{\text{٣}}$ ، . . . ، $\text{ع}_{\text{ر}}$ الانحرافات المعيارية لهذه المجموعات الجزئية ،

و $\text{ع}_{\text{م}}$ الانحراف المعيارى للمتوسطات الحسابية للمجموعات ،

$$\text{و}\;\text{ه} = \text{ن}_{\text{١}} + \text{ن}_{\text{٢}} + \text{ن}_{\text{٣}} \; . \; . \; . + \text{ن}_{\text{ر}} = \text{مجـ}\,\text{ن}$$

ولإثبات صحة هذه العلاقة نفرض أن مفردات المجموعات هى كما يأتى :

الأولى: $\text{ا}_{\text{١}}$ ، $\text{ا}_{\text{٢}}$ ، . . . ، $\text{ا}_{\text{ن}_{\text{١}}}$ ، عددها $\text{ن}_{\text{١}}$ ، ووسطها الحسابى $\text{م}_{\text{١}}$ ، وانحرافها المعيارى $\text{ع}_{\text{١}}$

الثانية : $\text{ب}_{\text{١}}$ ، $\text{ب}_{\text{٢}}$ ، . . . ، $\text{ب}_{\text{ن}_{\text{٢}}}$ « $\text{ن}_{\text{٢}}$ « « $\text{م}_{\text{٢}}$ « « $\text{ع}_{\text{٢}}$

الثالثة : $\text{حـ}_{\text{١}}$ ، $\text{حـ}_{\text{٢}}$ ، . . . ، $\text{حـ}_{\text{ن}_{\text{٣}}}$ « $\text{ن}_{\text{٣}}$ « « $\text{م}_{\text{٣}}$ « « $\text{ع}_{\text{٣}}$

. . . .

الأخيرة: $\text{ى}_{\text{١}}$ ، $\text{ى}_{\text{٢}}$ ، . . . ، $\text{ى}_{\text{ن}_{\text{ر}}}$ « $\text{ن}_{\text{ر}}$ « « $\text{م}_{\text{ر}}$ « « $\text{ع}_{\text{ر}}$

∴ $\text{مجـ}\,\text{ا} = \text{ن}_{\text{١}}\text{م}_{\text{١}}$ ، و $\text{مجـ}(\text{ا} - \text{م}_{\text{١}})^{\text{٢}} = \text{ن}_{\text{١}}\,\text{ع}_{\text{١}}^{\text{٢}}$ ؛ وبالمثل للمجموعات الأخرى .

الوسط الحسابى للمجموعة الكلية

$$= \frac{\text{مجـ ا} + \text{مجـ ب} + . . + \text{مجـ ى}}{\text{ن}_{\text{١}} + \text{ن}_{\text{٢}} + . . + \text{ن}_{\text{ر}}} = \bar{\text{س}} \text{ ، مثلا ،}$$

$$= \frac{\text{ن}_{\text{١}}\text{م}_{\text{١}} + \text{ن}_{\text{٢}}\text{م}_{\text{٢}} + . . + \text{ن}_{\text{ر}}\text{م}_{\text{ر}}}{\text{ن}_{\text{١}} + \text{ن}_{\text{٢}} + . . + \text{ن}_{\text{ر}}} \quad \text{وهو الوسط الحسابى للميات .}$$

∴ الوسط الحسابى للمجموعة الكلية هو فى الوقت نفسه الوسط الحسابى لمتوسطات المجموعات الجزئية .

∴ الانحراف المعيارى لهذه المتوسطات بين نفسها هو $\text{ع}_{\text{م}}$ حيث

$$\text{ه}\,\text{ع}_{\text{م}}^{\text{٢}} = \text{ن}_{\text{١}}(\text{م}_{\text{١}} - \bar{\text{س}})^{\text{٢}} + \text{ن}_{\text{٢}}(\text{م}_{\text{٢}} - \bar{\text{س}})^{\text{٢}} + . . . + \text{ن}_{\text{ر}}(\text{م}_{\text{ر}} - \bar{\text{س}})^{\text{٢}}$$

$$= \text{مجـ}\,\text{ن}(\text{م} - \bar{\text{س}})^{\text{٢}} \quad . \; . \; . \; . \qquad (\text{٢})$$

الانحراف المعيارى للمجموعة الأولى هو $\text{ع}_{\text{١}}$ ، حيث

$$\text{ن}_{\text{١}}\,\text{ع}_{\text{١}}^{\text{٢}} = \text{مجـ}(\text{ا} - \text{م}_{\text{١}})^{\text{٢}} .$$

وإذا أخذنا $\bar{\text{س}}$ كوسط فرضى لهذه المجموعة ، ووضعنا $\text{ح}_{\text{١}} = \text{م}_{\text{١}} - \bar{\text{س}}$ ، ورمزنا لمجموع مربعات الانحرافات عن الوسط الفرضى بالرمز $\text{ن}_{\text{١}}\,\text{ل}_{\text{١}}^{\text{٢}}$

$$\therefore \; \text{مجـ}(\text{ا} - \bar{\text{س}})^{\text{٢}} = \text{ن}_{\text{١}}\,\text{ل}_{\text{١}}^{\text{٢}} = \text{ن}_{\text{١}}\,\text{ع}_{\text{١}}^{\text{٢}} + \text{ن}_{\text{١}}\,\text{ح}_{\text{١}}^{\text{٢}} \; . . . (\text{بند ١٧٧})$$

$$= \text{ن}_{\text{١}}\,\text{ع}_{\text{١}}^{\text{٢}} + \text{ن}_{\text{١}}(\text{م}_{\text{١}} - \bar{\text{س}})^{\text{٢}}$$

$$\text{وكذلك}\; \text{مجـ}(\text{ب} - \bar{\text{س}})^{\text{٢}} = \text{ن}_{\text{٢}}\,\text{ع}_{\text{٢}}^{\text{٢}} + \text{ن}_{\text{٢}}(\text{م}_{\text{٢}} - \bar{\text{س}})^{\text{٢}}$$

$$\text{«}\; \text{مجـ}(\text{حـ} - \bar{\text{س}})^{\text{٢}} = \text{ن}_{\text{٣}}\,\text{ع}_{\text{٣}}^{\text{٢}} + \text{ن}_{\text{٣}}(\text{م}_{\text{٣}} - \bar{\text{س}})^{\text{٢}}$$

. . . . . .

$$\text{«}\; \text{مجـ}(\text{ى} - \bar{\text{س}})^{\text{٢}} = \text{ن}_{\text{ر}}\,\text{ع}_{\text{ر}}^{\text{٢}} + \text{ن}_{\text{ر}}(\text{م}_{\text{ر}} - \bar{\text{س}})^{\text{٢}} .$$

يلاحظ أن مجموع الأطراف اليمنى لهذه المتساويات يساوى مجموع مربعات انحرافات جميع المفردات ا ، ب ، . . . ، ى عن الوسط $\bar{\text{س}}$ ، ويساوى $\text{ه}\,\text{ع}_{\text{س}}^{\text{٢}}$.

∴ $\text{ه}\,\text{ع}_{\text{س}}^{\text{٢}} = \text{مجـ}\,\text{ن}\,\text{ع}^{\text{٢}} + \text{مجـ}\,\text{ن}(\text{م} - \bar{\text{س}})^{\text{٢}}$ ؛ وباستخدام العلاقة (٢)

∴ $= \text{مجـ}\,\text{ن}\,\text{ع}^{\text{٢}} + \text{ه}\,\text{ع}_{\text{م}}^{\text{٢}}$ ،

وهى النتيجة المطلوبة .

$$\therefore \; \text{مجـ}\,\text{ن}\,\text{ع}^{\text{٢}} = \text{ه}\,\text{ع}_{\text{س}}^{\text{٢}} - \text{ه}\,\text{ع}_{\text{م}}^{\text{٢}} \text{ ،} \quad . \; . \; . \qquad .(\text{٣})$$

وهذا معناه أن مجموع مربعات انحرافات المفردات ، كل واحدة عن الوسط الحسابى لمجموعتها ، يساوى مجموع مربعات انحرافاتها عن الوسط الحسابى العمومى ناقصاً مجموع مربعات انحرافات الأوساط الحسابية للمجموعات عن هذا الوسط العمومى نفسه . وهذه نتيجة مهمة سنحتاج إليها فيما بعد .

عزوم التوزيع التكرارى

١٨١ – لنفرض مجموعة من القيم في توزيع تكرارى ، وهى :

$س_1$ ، $س_2$ ، $س_3$ ، . . . .

وتكراراتها $ك_1$ ، $ك_2$ ، $ك_3$ ، . . . على الترتيب ، حيث مجـ ك = ن .

وليكن الوسط الحسابى لهذه القيم $\overline{س}$ والانحراف المعيارى $ع_س$.

الكمية $\frac{1}{ن}$ مجـ س . ك = العزم الأول للتوزيع التكرارى حول نقطة الأصل أو الصفر ،

$= م'_1$ . . . . . . (١) .

والكمية $\frac{1}{ن}$ مجـ $س^2$. ك = العزم الثانى للتوزيع حول نقطة الأصل ،

$= م'_2$ . . . . . (٢) .

وعلى العموم نقول :

$\frac{1}{ن}$ مجـ $س^ر$. ك = العزم الرائى للتوزيع حول نقطة الأصل

$= م'_ر$ ، . . . . . (٤) .

ويلاحظ أن $م'_1 = \overline{س}$ = الوسط الحسابى ،

وأن $م'_2 = ع^2 + \overline{س}^2$ [انظر معادلة (٢) بند ١٧٩] .

وإذا أخذنا الوسط الحسابى كمحور للعزوم ، وضربنا التكرارات فى انحرافات القيم عن الوسط الحسابى – أى بعدها عنه – نحصل على العزوم المختلفة حول الوسط الحسابى . وهذه نرمز لها بالحرف مـ .

∴ $\frac{1}{ن}$ مجـ $(س - \overline{س})$. ك = العزم الأول حول الوسط الحسابى .

$= م_1$ ، . . . . (٥) .

وكذلك $\frac{1}{ن}$ مجـ $(س - \overline{س})^2$. ك

= العزم الثانى حول الوسط الحسابى .

$= م_2$ ، . . . . (٦) .

وعموما : $\frac{1}{ن}$ مجـ $(س - \overline{س})^ر$. ك

= العزم الرائى حول الوسط الحسابى .

$= م_ر$ ، . . . . (٧) .

ويلاحظ أن $م_1$ = صفراً ، و $م_2 = ع^2$.

المقارنة بين تشتت المجموعات المختلفة

١٨٢ — مقاييس التشتت التى بحثناها فى هذا الباب يعبر كل منها عن التشتت بين مفردات مجموعة معينة . ولكنها لا تصلح على علاتها لمقارنة التشتت فى مجموعتين مختلفتين أو أكثر . لأنها معبر عنها بوحدات مطلقة هى نفس الوحدات المقيسة بها القيم الأصلية فى المجموعات ، كما رأينا مثلا فى بند ١٧٨ حيث وجدنا الانحراف المعيارى للأجور يساوى ٢ر٥٥ قرشاً ، وفى بند ١٧٢ وجدنا الانحراف المتوسط للأعمار يساوى ١ر١٣٩ سنة . ولا يمكننا طبعاً أن نقارن بين ٢ر٥٥ قرشاً و ١ر١٣٩ سنة ، ونقول إن تشتت الأجور أكبر أو أقل من تشتت الأعمار .

ومن جهة أخرى ، نعلم أن الانحراف المعيارى لأى مجموعة يقيس تشتتها حول وسطها الحسابى . فاذا كان لدينا مجموعتان لهما وسطان حسابيان مختلفان : مثلا فرقتان من التلاميذ فيهما الوسطان الحسابيان للأعمار ١٥ سنة فى الأولى و١٨ سنة

فى الثانية ، لايمكننا مقارنة الانحرافين المعياريين لهما مباشرة ، بدون عمل تصحيح يعادل اختلاف وسطيهما الحسابيين .

معامل الاختلاف

١٧٣ — معامل الاختلاف أى (Coe fficient of Variation) لأى مجموعة هو خارج قسمة الانحراف المعيارى على الوسط الحسابى ، مضروبا فى العدد ١٠٠ . فهو يعبر إذن عن الانحراف المعيارى (أو التشتت) فى صورة نسبة مئوية من الوسط الحسابى . وعلى ذلك يمكن استخدامه فى مقارنة التشتت للمجموعات المختلفة ، حتى ولو لم تتساو أوساطها الحسابية ، أو اختلف نوع الوحدات المستعملة فى قياس مفرداتها .

فى التوزيع التكرارى لأعمار ١٧٣٩ تلميذاً نرى أن معامل الاختلاف بين هذه الأعمار

$$= \frac{١٫٧٠٢}{١٧٫٧٤٩} \times ١٠٠$$

$$= ٩٫٦ \ .$$

وفى التوزيع التكرارى للأجور المذكورة فى بند ١٧٨ ، نجد معامل الاختلاف فى هذه المجموعة

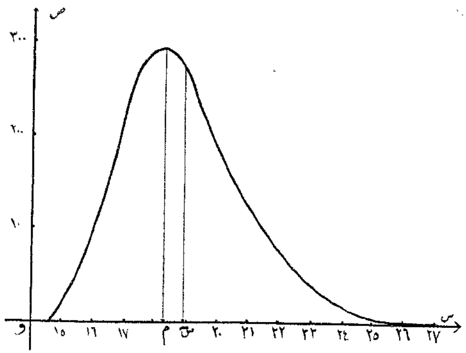
$$= \frac{٢٫٥٥}{١٩٫٢٥} \times ١٠٠$$ (انظر الوسط الحسابى فى بند ١٣٢)

$$= ١٣٫٢ \ .$$

∴ درجة الاختلاف بين الأجور (فى مجموعة ٧٤٣٢ عاملا) أكبر من درجة اختلاف الأعمار (فى مجموعة ١٧٣٩ تلميذاً) . ويلاحظ أن العددين ٩٫٦ و ١٣٫٢ لاتمييز لهما ، وهما فى الواقع نسبتان مئويتان .

١٨٤ — ذكرنا فى الباب الخامس أن المنحنيات التكرارية بعضها متماثل وبعضها غير متماثل أو ملتو . والآن نبحث فى كيفية قياس درجة الالتواء فى

المنحنيات التكرارية العادية ، أى التى يكون لها قمة واحدة ، كما فى الأشكال ٣٢ و ٣٣ و ٣٧ (صفحة ١٠٣ و ١٠٤ و ١١٠) . أما المنحنيات ذات الفرع الواحد (شكل ٤١ و ٤٢) أو ذات الفرعين (شكل ٤٣) فلا محل للبحث فيها من هذه الناحية .



(شكل ٥٩)

منحنى تكرارى ملتو إلى اليسار التواء موجباً

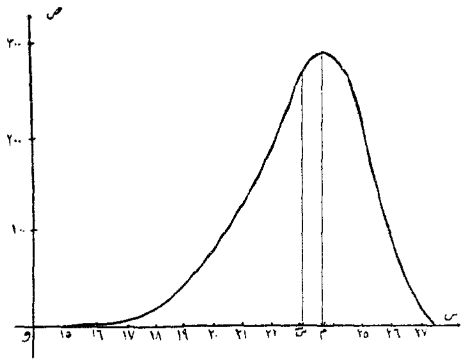
قلنا إن من خواص المنحنى المتماثل أن يكون الوسط الحسابى والمنوال متساويين . وهذه الخاصة لا توجد فى المنحنى غير المتماثل . يمكننا إذن اتخاذ هذه الخاصة أساساً لقياس الالتواء . فالفرق بين الوسط الحسابى والمنوال يساوى صفراً إذا كان المنحنى متماثلا ؛ ويكون قريباً من الصفر إذا كان المنحنى قريباً من التماثل ؛ ويكون الالتواء شديداً إذا كان هذا الفرق كبيراً . وهذا الفرق نقيسه بالنسبة إلى الانحراف المعيارى للتوزيع التكرارى الذى نبحثه ، وذلك لأننا نعتبر الانحراف المعيارى هو الانحراف « النموذجى » أو « العادى » لمفردات المجموعة عن الوسط الحسابى ؛ فننسب إليه انحراف المنوال عن الوسط الحسابى .

وبناء على ذلك يكون مقياس الالتواء[1] هو ى١ .

$$ى_1 = \frac{\text{الوسط الحسابى} - \text{المنوال}}{\text{الانحراف المعيارى}} = \frac{\bar{س} - م}{ع}$$

الالتواء الموجب

١٨٥ — إذا كان الوسط الحسابى أكبر من المنوال كان الالتواء موجباً كما نرى فى شكل ٥٩ حيث نجد المنحنى منحرفاً إلى اليسار ، ونرى المنوال م أصغر من الوسط الحسابى س̄ .



شكل (٦٠)
منحنى تكرار ملتوى إلى اليمين التواء سالبا

الالتواء السالب

١٨٦ — إذا كان الوسط الحسابى للتوزيع التكرارى أصغر من المنوال ، كان الالتواء سالباً . والمنحنى التكرارى فى هذه الحالة يكون منحرفاً إلى اليمين حيث تكون قمته بعيدة عن المحور الرأسى ، وبعدها ينزل المنحنى بسرعة . ونرى فى شكل ٦٠ منحنياً ذا التواء سالب . ونرى فى الشكل أن المنوال أبعد من الوسط الحسابى عن المحور الرأسى ، أى أن المنوال م أكبر من الوسط الحسابى س̄.

(١) معناها بالانجليزية (Skewness)

١٨٧ — يمكن أن نستعيض عن المقدار س̄ — م المذكور فى مقياس الالتواء ، بالمقدار ٣ (الوسط الحسابى — الوسيط) إذا كان المنحنى قريباً من التماثل . وقد ذكرنا سابقاً (بند ١٦٥ صفحة ١٨١) أن :

الوسط الحسابى — الوسيط = ⅓ (الوسط — المنوال) تقريباً .

وبناء على ذلك يكون لدينا مقياس تقريبى للالتواء نستعمله فى حالة المنحنيات القريبة من التماثل ؛ وهو :

$$ى'_1 = \frac{٣\,(\text{الوسط} - \text{الوسيط})}{\text{الانحراف المعيارى}}$$

عيوب هذين المقياسين

١٨٨ — والعيب الكبير فى هذين المقياسين ى١ و ى'١ أنهما يقتضيان معرفة الانحراف المعيارى للتوزيع التكرارى ؛ ونعلم أن حساب الانحراف المعيارى يحتاج إلى مجهود حسابى كبير ، خصوصاً فى المجموعات كثيرة العدد . وإذا كان التوزيع التكرارى مفتوحاً من أحد الطرفين أو من كليهما ، لا يمكن معرفة الوسط الحسابى ولا الانحراف المعيارى ، ولا يمكن استخدامهما — وهذا عيب آخر . وكلاهما تقريبى : فالأول ى يحتاج إلى معرفة المنوال ؛ وهذا صعب التحديد كما نعلم ، ولا نعرفه إلا بالتقريب . والمقياس الثانى ى' مقرب أيضاً كما ذكرنا ، ومجال استعماله مقصور على بعض المنحنيات فقط .

مقياس بولي للالتواء

١٨٩ — هناك مقياس آخر للالتواء . والفكرة الأساسية فى هذا المقياس هى أن الربيعين فى التوزيع التكرارى المتماثل متساويا البعد عن الوسيط .

لنفرض أن الوسيط ، فى أى توزيع تكرارى ، يساوى ط ، وأن الربيع الأدنى = ع١ ، والربيع الأعلى = ع٢ .

نضع $ب_2 = ع_2 - ط$ و $ب_1 = ط - ع_1$.

∴ فى المنحنى المتماثل يكون $\text{ں}_{2} = \text{ں}_{1}$.

وعلى ذلك فالـكمية $\text{ں}_{2} - \text{ں}_{1}$ تساوى صفراً فى المنحنيات المتماثلة ، وقريبة من صفر فى المنحنيات القريبة من التماثل . وتكون كبيرة إذا كان الالتواء شديداً.

وهنا ننسب الـكمية $\text{ں}_{2} - \text{ں}_{1}$ إلى الـكمية $\text{ں}_{2} + \text{ں}_{1}$ ، وهى تساوى الفرق بين الربيعين ($\text{ں}_{2} + \text{ں}_{1} = \text{ع}_{2} - \text{ع}_{1}$) . وقد سبق أن استخدمنا الفرق بين الربيعين كمقياس للتشتت (بند ١٦٨ صفحة ١٨٤) ، وعلى ذلك يكون مقياس[1] الالتواء بهذه الطريقة هو :

$$\text{ى}_{2} = \frac{\text{ں}_{2} - \text{ں}_{1}}{\text{ں}_{2} + \text{ں}_{1}}$$

وميزة هذا المقياس سهولة حساب الوسيط والربيعين بدقة أحسن من المنوال ، وأسهل من الانحراف المعيارى .

ويلاحظ أن الالتواء يكون سالباً إذا كان ں_{2} أصغر من ں_{1} ، أى إذا كان الربيع الأعلى أقرب إلى الوسيط من الربيع الأدنى . وهـذا يوافق تماماً المنحنى المنحرف بقمتـه إلى اليمين شكل ٦٠ . وعلى ذلك سيتفق المقياسان ى_{1} ، ى_{2} فى إشارة الالتواء السالب .

وكذلك يتفق المقياسان فى إشارة الالتواء الموجب حينما يكون ں_{2} أكبر من ں_{1} ، أى حينما يكون الربيع الأعلى أبعـد عن الوسيط من الربيع الأدنى ، (شكل ٥٩) .

(١) هذا المقياس اقترحه (A. L. Bowley) . أما المقياس الأول ى فينسب إلى (Karl Pearson) . انظر كتاب (F. C. Mills, Statistical Methods) صفحة ١٦٦ طبعة ١٩٣٤

١٩٠ — نأخذ التوزيع التكرارى الآتى ، ونحسب مقاييس الالتواء بهاتين الطريقتين . وهو توزيع أعمار الناجحين فى شهادة الدراسة الثانوية (قسم ثان) قسم العلوم سنة ١٩٣٦ .

جدول ٢٨ — إيجاد الوسط الحسابى والمنوال والوسيط والربيعين والانحراف المعيارى لأعمار ١٣٤٤ تلميذاً

الأعمار س	التكرار ك	انحرافات ح	ح . ك	ح² . ك	حدود عليا	تكرار متجمع
١٥	١١	— ٤	— ٤٤	١٧٦	أقل من ١٥ر٥	١١
١٦	٨٠	— ٣	— ٢٤٠	٧٢٠	« ١٦ر٥	٩١
١٧	٢٣٦	— ٢	— ٤٧٢	٩٤٤	« ١٧ر٥	٣٢٧
١٨	٢٩٢	— ١	— ٢٩٢	٢٩٢	« ١٨ر٥	٦١٩
١٩	٢٧٤	٠	٠	٠	« ١٩ر٥	٨٩٣
٢٠	١٨٥	١	١٨٥	١٨٥	« ٢٠ر٥	١٠٧٨
٢١	١٤٠	٢	٢٨٠	٥٦٠	« ٢١ر٥	١٢١٨
٢٢	٦٤	٣	١٩٢	٥٧٦	« ٢٢ر٥	١٢٨٢
٢٣	٣٢	٤	١٢٨	٥١٢	« ٢٣ر٥	١٣١٤
٢٤	٢٣	٥	١١٥	٥٧٥	« ٢٤ر٥	١٣٣٧
٢٥	٣	٦	١٨	١٠٨	« ٢٥ر٥	١٣٤٠
٢٦	٣	٧	٢١	١٤٧	« ٢٦ر٥	١٣٤٣
٢٧	١	٨	٨	٦٤	« ٢٧ر٥	١٣٤٤
	١٣٤٤		— ١٠١	٤٨٥٩		

الوسط الحسابى = ١٨ر٩٢٥ ، والمنوال = ١٨ر٢٥٧ ، والانحراف المعيارى = ١ر٩٠٠ الوسيط = ١٨ر٦٩٥ ، والربيع الأدنى = ١٧ر٥٣٢ ، والربيع الأعلى = ٢٠ر١٢٦ .

$$\therefore \quad \text{ى}_{1} = \frac{\text{١٨ر٩٢٥} - \text{١٨ر٢٥٧}}{\text{١ر٩٠٠}} = \text{٠ر٣٥٢}$$

$$\bar{\text{ى}}_{1} = \frac{\text{٣(١٨ر٩٢٥} - \text{١٨ر٦٩٥)}}{\text{١ر٩٠٠}} = \text{٠ر٣٦٣}$$

$$\text{ى}_{٢} = \frac{١ر٤٣١ - ١ر١٦٣}{٢ر٥٩٤} = ٠ر١٠٣$$

ويلاحظ أن المقياس الثانى يساوى الأول تقريباً . وأن الثلاثة متفقة فى الإشارة الموجبة ، دلالة على أن المنحنى ذو التواء موجب .

ولكن لا ينتظر أن يكون المقياس الثالث $\text{ى}_{٣}$ يساوى المقياس الأول أو الثانى ؛ لأن الفكرة الأساسية فيها تختلف . وعلى ذلك لا نتوقع أن نحصل من المقياسين على نفس النتيجة . ويجب أن نتنبه إلى هذه النقطة عند مقارنة الالتواء فى مجموعات مختلفة ، فنقيس الالتواء فى كل المجموعات بنفس الطريقة ، حتى تكون المقارنة على أساس مشترك ، وإلا كانت المقارنة خطأ .

١٩١ — يوجد مقياس ثالث للالتواء وهو

$$\text{ى}_{٣} = \frac{\sqrt[٣]{\text{م}_{٣}}}{\text{ع}}$$

حيث ع هى الانحراف المعيارى و $\text{م}_{٣}$ هى العزم الثالث حول الوسط الحسابى.

أى أن
$$\text{ى}_{٣} = \frac{\sqrt[٣]{\sum \text{ك} (\text{س} - \bar{\text{س}})^{٣}}}{\text{ع}}$$

والفكرة الأساسية فى هذا المقياس هى أن العزم الثالث للتوزيع المتماثل حول وسطه الحسابى يساوى صفراً . وكلما كان هذا العزم قريبا من الصفر ، كان المنحنى قريباً من التماثل . وبالعكس إذا كان العزم الثالث كمية كبيرة (موجبة أو سالبة) كان التواء المنحنى شديداً .

ولكن هذا المقياس لا يستعمل كثيراً ، لصعوبة حسابه عملياً .

المراجـــع

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

Connor, L. R., *Statistics in Theory and Practice*, Chapters XI, XII.

Jones, C. *First Course in Statistics*, Chapters VI, VII.

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

الباب الثامن

الارتباط

تعريف الارتباط بين متغيرين

١٩٢ — الارتباط[1] بين ظاهرتين أو كميتين متغيرتين معناه وجود علاقة بينهما ، بحيث إذا تغيرت إحداها فى اتجاه معين فإن الثانية تميل إلى التغير فى اتجاه معين أيضاً . ويصح أن يكون تغير الظاهرتين فى اتجاه واحد ، أو فى اتجاهين متضادين . وفى الحالة الأولى نسمى الارتباط « طرديا » ، حيث إذا زادت واحدة تميل الثانية إلى الزيادة أيضاً ؛ وإذا نقصت الأولى تميل الثانية إلى النقص أيضاً . وفى الحالة الثانية نسمى الارتباط « عكسياً » ، بمعنى أنه إذا تغيرت الكمية الأولى بالزيادة ، تميل الثانيـة فى تغيرها إلى النقص ؛ والعكس بالعكس .

١٩٣ — ولا يتحتم لوجود الارتباط أن كل زيادة تحصل فى أحد المتغيرين لابد أن يصحبها زيادة فى المتغير الآخر (أو نقص فى حالة الارتباط العكسى) ؛ أو أن يكون التغير فيهما بنسبة واحدة . على أن هذا إذا تحقق يكون دلالة على شدة الارتباط والعلاقة بين المتغيرين ؛ ولكن قد يتأتى أن يزيد المتغير الأول مثلا ؛ ونظراً لظروف طارئة ، فى حالة معينة ، ينقص المتغير الثانى ، على خلاف ما تقضى به العلاقة الطردية المفروضة بينهما . ولكن المهم فى الموضوع أنه فى أغلب الحالات نجد الزيادة فى المتغير الأول مصحوبة بزيادة فى الثانى ، فى حالة الارتباط الطردى (أو بنقص فى حالة الارتباط العكسى) ؛ ونجد النقص فى أحدهما مقروناً

(١) معناه بالانجليزية (Correlation) وقد سماه بعض الناس «علاقة مشتركة» ؛ ولكننا نرى كلمة ارتباط أحسن وأبسط .

بنقص فى الآخر (أو زيادة فى الحالة العكسية) . ولا تكون النسبة بين التغيرين ثابتة فى كل الأحوال التى تقع تحت ملاحظتنا ، ولكنها تتراوح حول مقدار معين .

وهذا هو السبب فى قولنا إن وجود الارتباط معناه أن أحد المتغيرين « يميل » إلى مصاحبة الثانى فى تغيره على وجه العموم .

ويلاحظ أن هذا التعريف أيضاً تتمثل فيه وجهة النظر الإحصائية . إذ أننا لا نرجع إلى الحالات الفردية عند استنباط القوانين أو القواعد التى تسير عليها الظواهر ، ولكننا نهتم فقط بالاتجاه العام الذى تأخذه المجموعات الكبيرة . وذلك لنتفادى أخطاء المصادفات وتقلبات الظروف الشاذة .

١٩٤ — يتضح لنا مما تقدم أن الارتباط بين متغيرين معناه أن التغير فى أحدهما يكون مصحوباً بتغير فى الآخر ، وأن هناك علاقة معينة بين اتجاهى التغير فيهما — طردية أو عكسية .

ولكن وجود ارتباط بين ظاهرتين متغيرتين ليس دليلا على أن إحداهما نتيجة للأخرى ، أو أن التغير فى واحدة تابع للتغير فى الأخرى ولا ينشأ إلا بسببه . بل هو يشير فقط إلى احتمال وجود هـذه العلاقة . لأن هذه العلاقة ما هى إلا نوع خاص من أنواع العلاقات التى يدل الارتباط على وجودها . وهى كما يأتى :

علاقة سببية مباشرة

الحالة الأولى — أن يكون أحد المتغيرين نتيجة مباشرة للثانى .

ومثال ذلك الارتباط بين سعر أى سلعة فى السوق وكمية المطلوب منها فى السوق . إذ أن ارتفاع الأسعار ينتج عنه مباشرة هبوط فى كمية المستهلك من هذه السلعة ؛ وانخفاضها يسبب ازدياداً فى الطلب .

علاقة سببية غير مباشرة

الحالة الثانية — يكون أحد المتغيرين سبباً غير مباشر للثانى ، يؤثر فيه بوساطة عامل ثالث أو أكثر . فارتفاع الرسوم الجمركية على المنسوجات القطنية مثلا ،

يسبب ارتفاعاً فى أسعارها الداخلية . وهذا الارتفاع يمكن أصحاب المصانع المحلية من رفع أجور عمالهم . فالارتباط بين مقدار الرسوم الجمركية فى هذه الحالة ومستوى الأجور ناتج من علاقة سببية غير مباشرة بينهما .

عامل واحد يؤثر فى المتغيرين

الحالة الثالثة — أن يكون كل من المتغيرين المرتبطين نتيجة لعامل ثالث ، مشترك بينهما ، يؤثر فيهما فى وقت واحد ، فيكون كل تغير فى أحدهما مصحوباً بتغير فى الآخر . مثال ذلك الارتباط بين أسعار سلعتين تستهلكهما طبقة معينة من السكان . فإن أسعارهما تكون مرهونة بالحالة الاقتصادية لهؤلاء السكان : فيرتفع سعر كل منهما إذا زادت القوة الشرائية لهم ، ويهبط السعران معاً إذا نقصت قوتهم الشرائية بسبب انتشار البطالة بينهم أو لأى سبب آخر . وكذلك الارتباط بين أسعار سلعتين تردان من بلد بعيدة جداً ، بحيث تكون نفقات إنتاجهما ضئيلة بالنسبة إلى نفقات النقل — فنجد أسعار هاتين السلعتين ترتفع وتنخفض معاً ، تبعاً لتغيرات نفقات النقل . وكذلك إذا كانتا تنتجان من مادة خام واحدة ، رئيسية فى كل منهما ، بحيث تكوّن الجزء الأكبر من نفقات الإنتاج فيهما — فأسعارهما ترتفع أو تنخفض معاً تبعاً لأسعار هذه المادة الرئيسية .

بعض العوامل مشتركة بين المتغيرين

الحالة الرابعة — أن يكون ضمن العوامل التى تؤثر فى أحد المتغيرين والعوامل التى تؤثر فى الآخر ، عامل مشترك أو أكثر . مثلا لو اختبرنا عدداً من التلاميذ فى مادتين مثل الجغرافيا والتاريخ ، نجد ارتباطاً شديداً بين درجات هاتين المادتين . والسبب فى ذلك أن مقدرة الشخص ونبوغه أو ضعفه فى مادة التاريخ يتوقف على استعداده العام أو ذكائه ، وعلى مقدرة أخرى نوعية خاصة بهذا العلم وطرائقه . ونبوغ أى شخص أو ضعفه فى علم الجغرافيا يتوقف أيضاً على استعداده العام أو ذكائه ، وبجانب ذلك على مقدرة خاصة بعلم الجغرافيا تخالف تلك المقدرة اللازمة

للنبوغ فى علم التاريخ . وعلى ذلك نجد هنا عاملا مشتركا — وهو الذكاء الشخصى أو الاستعداد العام — يؤثر فى الظاهرتين ، علاوة على عوامل أخرى خاصة بكل ظاهرة على حدة وليست مشتركة .

ومثال ذلك أيضاً سلعتان فى السوق تدخل فى إنتاجهما مادة خام أو أكثر بصفة رئيسية ، علاوة على مواد أخرى خاصة بكل سلعة ، ولا تدخل فى الأخرى . فالارتباط الذى نجده بين أسعار هاتين السلعتين ناتج من وجود عوامل مشتركة بينهما ضمن العوامل التى تؤثر فى كل واحدة .

١٩٥ — ومهما كان الارتباط بين المتغيرين شديداً ، فهو لا يكفى بمفرده لمعرفة نوع العلاقة بينهما . ولا بد لتحديد نوعها من الاستعانة بمعلوماتنا الخاصة ، وإلمامنا بظروف هذين المتغيرين . وعلى كل حال فنوع هذه العلاقة محصور فى الأنواع الأربعة التى ذكرناها . وفى كل نوع منها يصح أن يكون الارتباط شديداً أو ضعيفاً .

١٩٦ — أول خطوة فى دراسة الارتباط هى أن نبحث فى كيفية قياسه ، والتعبير عنه فى صورة رقمية تساعدنا فى عمل المقارنات بين الحالات المختلفة التى يظهر فيها الارتباط .

الارتباط والاقتران والتوافق

١٩٧ — ولو تأملنا فى الحالات المختلفة التى يمكن أن تعرض لنا عند دراسة الارتباط ، وجدناها تشمل عدة أنواع . وهذه يمكن تقسيمها إلى ثلاثة أنواع[1] كما يأتى :

---

(١) اقتران = Association · توافق = Contingency · ارتباط بسيط = Simple Correlation

أولا — العلاقة بين ظواهر يمكن أن تقاس و يعبر عنها فى صورة رقمية . وهذه العلاقة نسميها « ارتباطا » . ومثال ذلك العلاقة بين طول الشخص ووزنه ، و بين سعر السامة وكمية المطلوب منها ، و بين كمية المحصول وكمية السماد المستعمل فى حقل معين ، وهكذا .

ثانياً — العلاقة بين ظواهر لا يمكن قياسها رقمياً . وهذه العلاقة نسميها « الاقتران » . ومثال ذلك العلاقة بين جنسية الشخص وديانته ، و بين لون الزهرة ورائحتها ، و بين نوع الشخص (ذكر أو أنثى) ونوع العمل الذى يقوم به (صناعى أو تجارى) ، وهكذا .

ثالثاً — العلاقة بين ظواهر بعضها يقاس رقمياً و بعضها لا يقاس . وهذه نسميها « توافقاً » . ومثال ذلك العلاقة بين نوع القطن (سكلاريدس وأشمونى الخ) وطول تيلته بالسنتيمتر ، و بين نوع الحرفة التى يزاولها العامل وأجره بالقرش ، وهكذا .

الارتباط البسيط والمتعدد والجزئى

١٩٨ — لنأخذ أولا العلاقة بين الظواهر المقيسة ، أى الارتباط . ونجد هنا أن البحث يتفرع إلى فرعين(١) :

١ — الارتباط بين ظاهرتين أومتغيرين فقط . و يسمى « الارتباط البسيط » .

٢ — الارتباط بين ظاهرة وظاهرتين مجتمعتين أو منفردتين ، أو أكثر من ظاهرتين . و يسمى « الارتباط المتعدد » أو « الارتباط الجزئى » ، على الترتيب . وسنبدأ ببحث الارتباط البسيط وطريقة قياسه .

---

(١) ارتباط متعدد = (Multiple Correlation) . ارتباط جزئى (Partial Correlation)

لقياس الارتباط نحصل على قيم متناظرة للتغيرين

١٩٩ — عند دراسة الارتباط البسيط بين كميتين متغيرتين ، نبدأ بمشاهدة هاتين الكميتين فى عدد من الحالات ، وفى كل حالة من هذه تقيس كلا منهما ، وبدون القيم المتناظرة لها . وإذا كان عدد الحالات التى وقعت تحت ملاحظتنا ن مثلا ، حصلنا على قيم عددها ن للمتغير الأول ومثلها للمتغير الثانى ، تناظرها واحدة لواحدة . أى أننا نحصل على ن من أزواج القيم المتناظرة لهذين المتغيرين .

وللسهولة فى الكلام والتفكير ، نفرض أن المتغير الأول س ، والمتغير الثانى المرتبط به ص . ونفرض أن القيم التى حصلنا عليها للمتغير الأول فى هذه التجربة هى :

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

وأن القيم المناظرة لها للمتغير الثانى هى :

$$\text{ص}_1 ،\ \text{ص}_2 ،\ \text{ص}_3 ،\ .\ .\ .\ ،\ \text{ص}_{\text{ن}} .$$

ولنفرض أيضاً أن الوسطين الحسابيين لهاتين المجموعتين من القيم هما على الترتيب $\overline{\text{س}}$ و $\overline{\text{ص}}$ ؛ وأن الانحرافين المعياريين لهما هما ع ، عَ . وحسب تعريف الوسط الحسابى والانحراف المعيارى :

$$\text{يكون}\quad \text{ن}\,\overline{\text{س}} = \text{مجـ}\ \text{س} \qquad \text{و} \qquad \text{ن}\,\overline{\text{ص}} = \text{مجـ}\ \text{ص} ؛$$

$$\text{و}\quad \text{ن}\,\text{ع}^2 = \text{مجـ}(\text{س} - \overline{\text{س}})^2 \qquad \text{و} \qquad \text{ن}\,\text{عَ}^2 = \text{مجـ}(\text{ص} - \overline{\text{ص}})^2$$

مقياس الارتباط مشتق من قيم س وقيم ص على السواء

٢٠٠ — هذه القيم السينية والصادية هى كل ما لدينا من المعلومات فى هذا الموضوع ، وسنعتمد عليها كلياً فى قياس الارتباط بين هذين المتغيرين س و ص ؛ فلا بد إذن أن يكون مقياس الارتباط الذى نستخدمه مشتقاً من هذه القيم .

ولا بد أيضاً أن تدخل قيم س وقيم ص على قدم المساواة فى الصيغة العامة لهذا القياس . وبما أننا نستخدم جميع ما لدينا من المعلومات عن هذين المتغيرين لقياس الارتباط بينهما ، فسنجد أن جميع هذه القيم السينية والصادية تدخل فى حسابنا بدون تفضيل بين أى قيمة وأخرى .

مقياس الارتباط الذى نستعمله فى هذه الحالة نسميه « معامل الارتباط »[1] . والآن نتكلم فى كيفيه اشتقاقه من هذه القيم الموجودة بالتجربة والقياس ، طبقاً للتعريف الذى أوردناه (انظر بنـــد ١٩٢) ، والمعنى الذى نقصده من فكرة الارتباط .

الارتباط يقاس بواسطة التغيرات الحاصلة فى الظاهرتين

٢٠١ — قلنا إن الارتباط بين متغيرين معناه أن التغير فى أحدهما يكون — على العموم — مصحوباً بتغير فى الآخر ، بمعنى أن الزيادة فى أحدهما تكون — على العموم — مصحوبة مثلا[2] بزيادة فى الثـانى ، والنقص فى الأول يكون مصحوباً بنقص فى الثانى أيضاً .

وإذا تكلمنا عن التغير فى كمية مثل س (أو ص) ، فمقدار هـذا التغير يساوى الفرق بين القيمة التى تأخـذها س ومقدار معين يعتبر أساساً . وأحسن أساس نختـاره لقياس التغير فى قيم س هو بلا شك الوسط الحسابى لهـذه القيم السينية ، أى $\bar{\text{س}}$. وكذلك فى حالة ص نختار $\bar{\text{ص}}$.

وعلى ذلك تكون التغيرات فى قيم س وقيم ص هى :

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

$$\text{و } \text{ص}_1 - \bar{\text{ص}} \text{ ، } \text{ص}_2 - \bar{\text{ص}} \text{ ، } \text{ص}_3 - \bar{\text{ص}} \text{ ، } \ldots \text{ ، } \text{ص}_\text{ن} - \bar{\text{ص}} \text{ .}$$

---

(١) فى الانجليزية (Coefficient of Correlation) ويرمز له بالحرف r .

(٢) أو تـكون مصحوبة بنقص فى حالة الارتباط العكسى .

وهذه هى ، بعبارة أخرى ، انحرافات القيم عن وسطها الحسابى ؛ وهـذه الانحرافات إذن هى التى سنعتمد عليها فى قياس الارتباط .

تقسيم الانحرافات على الانحراف المعيارى

٢٠٢ — ولـكن هذه الانحرافات لا يمكن مقارنتها ببعضها كما هى على هذه الصورة . لأنه إذا فرضنا أن س تدل على عمر الشخص بالسنين مثلا و ص تدل على وزنه بالـكيلوجرام ، فان الانحرافات السينية تـكون مقيسة بالسنين ، فى حين أن الانحرافات الصادية تـكون مقيسة بالـكيلوجرامات . وعلاوة على ذلك فإن التشتت مختلف فى مجموعتى قيم س و ص ، مما يفسد المقارنة بين هذه الانحرافات على علاتها .

لذلك نقسم الانحرافات السينية على الانحراف المعيارى لقيم س ، والانحرافات الصادية على الانحراف المعيارى لقيم ص ، حيث إن الانحراف المعيارى هو ، كما سبق أن ذكرنا فى مناسبة أخرى ، الانحراف الذى تقاس جميع الانحرافات بالنسبة إليه حتى تمكن مقارنتها . وهكذا نحصل على انحرافات سينية وانحرافات صادية خالية من كل تميز ومعبر عنها بوحدات يمكن مقارنتها .

وهذه الانحرافات هى :

$$\frac{\text{س}_1 - \bar{\text{س}}}{\text{ع}} \text{ ، } \frac{\text{س}_2 - \bar{\text{س}}}{\text{ع}} \text{ ، } \frac{\text{س}_3 - \bar{\text{س}}}{\text{ع}} \text{ ، } \ldots \text{ ، } \frac{\text{س}_\text{ن} - \bar{\text{س}}}{\text{ع}} \text{ ؛}$$

$$\frac{\text{ص}_1 - \bar{\text{ص}}}{\bar{\text{ع}}} \text{ ، } \frac{\text{ص}_2 - \bar{\text{ص}}}{\bar{\text{ع}}} \text{ ، } \frac{\text{ص}_3 - \bar{\text{ص}}}{\bar{\text{ع}}} \text{ ، } \ldots \text{ ، } \frac{\text{ص}_\text{ن} - \bar{\text{ص}}}{\bar{\text{ع}}} \text{ .}$$

٢٠٣ — لنفرض أن هناك ارتباطاً شديداً بين س وص . إذن لو تغيرت س وزادت زيادة كبيرة (أو نقصت كثيراً) وانحرفت بذلك عن وسطها الحسابى انحرافا كبيراً ، كان ذلك مصحوبا بتغير كبير أيضـاً فى ص وانحراف كبير عن وسطها الحسابى أيضاً ؛ ونتج عن ذلك أن يكون حاصل ضرب هذين الانحرافين

كبيراً (موجباً إذا كان الارتباط طردياً ، أو سالباً إذا كان عكسيا) . وإذا تغيرت س بمقدار صغير فقط ، كان التغيير ص صغيراً أيضاً بحكم الارتباط الشديد بينهما ، وكان حاصل ضرب الانحرافين صغيراً ؛ ولكن هذا الحاصل الصغير يكون فقط فى الحالات التى تكون فيها قيم س و ص قريبة من وسطيهما الحسابيين . وعلى العموم يكون متوسط حواصل ضرب الانحرافات كبيراً فى حالة الارتباط الشديد .

أما إذا كان الارتباط ضعيفاً ، وكانت كل من س و ص تتغير مستقلة عن الأخرى ، فيصح أن تكون س كبيرة جدا ، وانحرافها عن وسطها كبيراً جداً ، دون أن يظهر لذلك أى أثر فى ص ، لضعف الصلة بينهما . وعلى ذلك يكون حاصل ضرب انحرافى س و ص صغيراً فى المتوسط .

وعلى ذلك نأخذ متوسط حاصل ضرب انحرافى س و ص عن وسطيهما الحسابيين ، كمقياس للارتباط بينهما ، كما نقول فى علم الطبيعة إن قوة الجاذبية بين جسمين تتناسب مع حاصل ضرب كتلتيهما ، و بين قطبى مغناطيسين تساوى حاصل ضرب شدتيهما . ويكون معامل الارتباط [1] إذن يساوى

$$\text{ر} = \frac{1}{\text{ن}}\left[\frac{\text{س}_1-\bar{\text{س}}}{\text{ع}}\times\frac{\text{ص}_1-\bar{\text{ص}}}{\text{ع}'}+\frac{\text{س}_2-\bar{\text{س}}}{\text{ع}}\times\frac{\text{ص}_2-\bar{\text{ص}}}{\text{ع}'}+\cdots+\frac{\text{س}_\text{ن}-\bar{\text{س}}}{\text{ع}}\times\frac{\text{ص}_\text{ن}-\bar{\text{ص}}}{\text{ع}'}\right]$$

$$= \frac{\sum(\text{س}-\bar{\text{س}})\cdot(\text{ص}-\bar{\text{ص}})}{\text{ن ع}\cdot\text{ع}'} \quad , \quad . \quad . \quad . \quad (1) .$$

وهذا هو أهم المقاييس المستعملة فى دراسة الارتباط .

---

(١) هذا العامل وضعه (Karl Pearson) وسماه « معامل حاصل ضرب العزوم للارتباط » (Product Moment Coefficient of Correlation)

حدود قيم المعامل ر وإشارته الجبرية

٢٠٤ — ويلاحظ أن جميع قيم س وقيم ص تدخل فى هذا المعامل على قدم المساواة كما قلنا من قبل (بند ٢٠٠) .

ومن المهم جداً أن نلاحظ أن هذا المعامل لا يزيد أبداً عن الواحد الصحيح ؛ وهو دائماً محصور[1] فى الفترة ما بين — ١ و + ١ . ويكون موجباً إذا كان الارتباط طردياً ، وسالباً إذا كان عكسياً . وإذا كان ر = ٠ كان الارتباط منعدماً ، أى أن الظاهرتين مستقلتان عن بعضهما تماماً .

لأنه فى حالة الارتباط الطردى تقترن قيم س الكبيرة ، على العموم ،

---

(١) يمكن إثبات ذلك بأن نبرهن على أن ر^2 دائماً أقل من ١ ، وذلك كما يأتى :

نكتب $\text{ل}_1 = \text{س}_1 - \bar{\text{س}}$ ، و $\text{ل}_2 = \text{س}_2 - \bar{\text{س}}$ ، . . للاختصار

و $\text{م}_1 = \text{ص}_1 - \bar{\text{ص}}$ ، و $\text{م}_2 = \text{ص}_2 - \bar{\text{ص}}$ ، . . . »

$\therefore$ $\text{ن ع}^2 = \text{ل}_1^2 + \text{ل}_2^2 + \cdots + \text{ل}_\text{ن}^2$ ؛ و $\text{ن ع}'^2 = \text{م}_1^2 + \text{م}_2^2 + \cdots + \text{م}_\text{ن}^2$.

$$\text{ر}^2 = [\text{ل}_1\text{م}_1 + \text{ل}_2\text{م}_2 + \cdots + \text{ل}_\text{ن}\text{م}_\text{ن}]^2 \div (\text{ل}_1^2 + \text{ل}_2^2 + \cdots + \text{ل}_\text{ن}^2)\ (\text{م}_1^2 + \text{م}_2^2 + \cdots + \text{م}_\text{ن}^2)$$

$$\text{البسط} = \text{ل}_1^2\text{م}_1^2 + \text{ل}_2^2\text{م}_2^2 + \text{ل}_3^2\text{م}_3^2 + \cdots + 2\,\text{ل}_1\text{م}_1(\text{ل}_2\text{م}_2 + \text{ل}_3\text{م}_3 + \cdots) + 2\,\text{ل}_2\text{م}_2(\text{ل}_3\text{م}_3 + \cdots) + \cdots$$

$$\text{المقام} = \text{ل}_1^2\text{م}_1^2 + \text{ل}_2^2\text{م}_2^2 + \text{ل}_3^2\text{م}_3^2 + \cdots + \text{ل}_1^2(\text{م}_2^2 + \text{م}_3^2 + \cdots) + \text{ل}_2^2(\text{م}_1^2 + \text{م}_3^2 + \cdots) + \cdots$$

$$\therefore \text{المقام} - \text{البسط} = \text{ل}_1^2\text{م}_2^2 + \text{ل}_2^2\text{م}_1^2 + \text{ل}_1^2\text{م}_3^2 + \text{ل}_3^2\text{م}_1^2 + \cdots - 2\,\text{ل}_1\text{م}_1\text{ل}_2\text{م}_2 - 2\,\text{ل}_1\text{م}_1\text{ل}_3\text{م}_3 - \cdots$$

$$= (\text{ل}_1\text{م}_2 - \text{ل}_2\text{م}_1)^2 + (\text{ل}_1\text{م}_3 - \text{ل}_3\text{م}_1)^2 + (\cdots)^2 + \cdots$$

= مجموع مربعات كاملة = كمية موجبة أو صفراً

$\therefore$ البسط $\leq$ من المقام . وعلى ذلك $\text{ر}^2 \leq 1$ ؛

ويكون ر محصوراً بين — ١ و + ١ ، ويصح أن ر = $\pm$ ١ .

بقيم ص الكبيرة أيضاً ، وكذلك تقترن قيم س الصغيرة بقيم ص الصغيرة . أى أن انحرافات س و ص عن وسطيهما الحسابيين تكون متحدة الاشارة . فتكون حواصل ضرب هذه الانحرافات موجبة ومجموع هذه الحواصل يكون موجباً ، وحينئذ تكون ر موجبة .

أما فى حالة الارتباط العكسى فتقترن قيم س الكبيرة بقيم ص الصغيرة على العموم ، والعكس بالعكس . وعلى ذلك تكون انحرافات س عن وسطها الحسابى مخالفة فى الإشارة لانحرافات ص . ويكون حاصل ضرب هذه الانحرافات سالباً ، وتكون ر سالبة بناء على ذلك .

وإذا كان بعض قيم س الكبيرة مقترنة بقيم كبيرة للمتغير ص و بعضها مقترناً بقيم صغيرة ، فإن بعض الحواصل يكون موجباً و بعضها يكون سالباً ، مما يجعل مجموع هذه الحواصل صغيراً أو صفراً ، فتكون ر صغيرة أو صفراً ، دلالة على ضعف الارتباط أو انعدامه .

صيغة أخرى لمعامل الارتباط

٢٠٥ — يمكننا ، إذا أردنا ، أن نستعيض[1] عن العبارة

---

(١) من الواضح أن $\text{مج}(\text{س}-\bar{\text{س}})(\text{ص}-\bar{\text{ص}}) = \text{مج}\,\text{س}\,\text{ص} - \text{ن}\,\bar{\text{س}}\,\bar{\text{ص}}$

لأن $(\text{س}_1-\bar{\text{س}})(\text{ص}_1-\bar{\text{ص}}) = \text{س}_1\text{ص}_1 + \bar{\text{س}}\,\bar{\text{ص}} - \bar{\text{س}}\,\text{ص}_1 - \bar{\text{ص}}\,\text{س}_1$ ،

و $(\text{س}_2-\bar{\text{س}})(\text{ص}_2-\bar{\text{ص}}) = \text{س}_2\text{ص}_2 + \bar{\text{س}}\,\bar{\text{ص}} - \bar{\text{س}}\,\text{ص}_2 - \bar{\text{ص}}\,\text{س}_2$ ،

. . . .

و $(\text{س}_\text{ن}-\bar{\text{س}})(\text{ص}_\text{ن}-\bar{\text{ص}}) = \text{س}_\text{ن}\text{ص}_\text{ن} + \bar{\text{س}}\,\bar{\text{ص}} - \bar{\text{س}}\,\text{ص}_\text{ن} - \bar{\text{ص}}\,\text{س}_\text{ن}$ ؛

وبالجمع

$$\therefore \text{مج}(\text{س}-\bar{\text{س}})(\text{ص}-\bar{\text{ص}}) = \text{مج}\,\text{س}\,\text{ص} + \text{ن}\,\bar{\text{س}}\,\bar{\text{ص}} - \bar{\text{س}}(\text{مج}\,\text{ص}) - \bar{\text{ص}}(\text{مج}\,\text{س})$$

$$= \text{مج}\,\text{س}\,\text{ص} + \text{ن}\,\bar{\text{س}}\,\bar{\text{ص}} - \bar{\text{س}}\,.\,(\text{ن}\,\bar{\text{ص}}) - \bar{\text{ص}}\,.\,(\text{ن}\,\bar{\text{س}})$$

$$= \text{مج}\,\text{س}\,\text{ص} - \text{ن}\,\bar{\text{س}}\,\bar{\text{ص}}\,.$$

$\text{مج}(\text{س}-\bar{\text{س}})(\text{ص}-\bar{\text{ص}})$ بالعبارة $\text{مج}\,\text{س}\,\text{ص} - \text{ن}\,\bar{\text{س}}\,\bar{\text{ص}}$.

وبذلك يصبح معامل الارتباط :

$$\text{ر} = \frac{\text{مج}\,\text{س}\,\text{ص} - \text{ن}\,\bar{\text{س}}\,\bar{\text{ص}}}{\text{ن}\,\text{ع}\cdot\text{ع}'} \quad \text{،} \quad \ldots \quad (1)\,.$$

وهذه الصورة أحياناً تكون أسهل فى الحساب من الصور المتقدمة ، لأن الوسط الحسابى $\bar{\text{س}}$ أو $\bar{\text{ص}}$ عادة يكون عدداً كسرياً ذا أرقام كثيرة ؛ وينتج من ذلك أيضاً أن الانحرافات $\text{س}-\bar{\text{س}}$ و $\text{ص}-\bar{\text{ص}}$ تكون أعداداً مكونة من أرقام كثيرة ، فتجعل العمليات الحسابية معقدة .

صيغة ثالثة للمعامل ر

٢٠٦ — ويمكننا أيضاً ، إذا أردنا ، أن نحسب الانحرافات السينية عن وسط فرضى مناسب بدل الوسط الحسابى ، وكذلك الانحرافات الصادية نحسبها عن وسط فرضى آخر بدل وسطها الحسابى . وهذه الطريقة نلجأ إليها فى بعض الأحيان إذا وجدنا أن عمليات الحساب فى إيجاد $\text{مج}(\text{س}-\bar{\text{س}})(\text{ص}-\bar{\text{ص}})$ ، أو $\text{مج}\,\text{س}\,\text{ص} - \text{ن}\,\bar{\text{س}}\,\bar{\text{ص}}$ ، صعبة أو طويلة . وحينئذ[1] يكون معامل الارتباط :

---

(١) هنا أيضاً

$$\text{مج}(\text{س}-\bar{\text{س}})(\text{ص}-\bar{\text{ص}}) = \text{مج}(\text{س}-\text{و})(\text{ص}-\text{و}') - \text{ن}\,\text{ح}\,\text{ح}'$$

لأن $(\text{س}_1-\bar{\text{س}})(\text{ص}_1-\bar{\text{ص}}) = (\text{س}_1-\text{و}+\text{و}-\bar{\text{س}})$

$$\times(\text{ص}_1-\text{و}'+\text{و}'-\bar{\text{ص}})$$

$$= (\text{س}_1-\text{و}-\text{ح})(\text{ص}_1-\text{و}'-\text{ح}')$$

$$= (\text{س}_1-\text{و})(\text{ص}_1-\text{و}') - \text{ح}(\text{ص}_1-\text{و}')$$

$$- \text{ح}'(\text{س}_1-\text{و}) + \text{ح}\,\text{ح}'$$

وكذلك $(\text{س}_2-\bar{\text{س}})(\text{ص}_2-\bar{\text{ص}}) = (\text{س}_2-\text{و})(\text{ص}_2-\text{و}') - \text{ح}(\text{ص}_2-\text{و}')$

$$- \text{ح}'(\text{س}_2-\text{و}) + \text{ح}\,\text{ح}'$$

و .. ..

و $(\text{س}_\text{ن}-\bar{\text{س}})(\text{ص}_\text{ن}-\bar{\text{ص}}) = (\text{س}_\text{ن}-\text{و})(\text{ص}_\text{ن}-\text{و}') - \text{ح}(\text{ص}_\text{ن}-\text{و}')$

$$- \text{ح}'(\text{س}_\text{ن}-\text{و}) + \text{ح}\,\text{ح}' \quad (*)$$

$$\text{ر} = \frac{\text{مج}(\text{س} - \text{و})(\text{ص} - \text{وَ}) - \text{ن ح حَ}}{\text{ن ع . عَ}} \quad \text{، . . . . . (٢) ،}$$

بفرض أن و = الوسط الفرضى السينى ، و ح = $\overline{\text{س}}$ − و ،

وأن وَ = « « الصادى ، و حَ = $\overline{\text{ص}}$ − وَ .

معادلة الفروق للعامل ر

٢٠٧ — ولو عوضنا فى المعادلة (١) عن العبارة مج س ص بالعبارة [1] :

$$\tfrac{1}{2}\left[\text{مج س}^2 + \text{مج ص}^2 - \text{مج}(\text{س} - \text{ص})^2\right]$$

نحصل على صيغة ثالثة لمعامل الارتباط ، نسميها **معادلة الفروق** [2] لحساب معامل الارتباط ، تمييزاً لها عن معادلات حاصل الضرب السابقة ، حيث إنها تحتوى على الفروق المربعة مج $(\text{س} - \text{ص})^2$ بدل حواصل الضرب مج(س ص) أو مج $(\text{س} - \overline{\text{س}})(\text{ص} - \overline{\text{ص}})$ ، أو مج(س − و)(ص − وَ) .

وهذه المعادلة هى :

(*) $\therefore\ \text{مج}(\text{س}-\overline{\text{س}})(\text{ص}-\overline{\text{ص}}) = \text{مج}(\text{س}-\text{و})(\text{ص}-\text{وَ}) - \text{ح}(\text{مج ص}-\text{ن وَ})$

$- \text{حَ}(\text{مج س} - \text{ن و}) + \text{ن ح حَ}$

$= \text{مج}(\text{س}-\text{و})(\text{ص}-\text{وَ}) - \text{ح}(\text{ن}\overline{\text{ص}}-\text{ن وَ})$

$- \text{حَ}(\text{ن}\overline{\text{س}} - \text{ن و}) + \text{ن ح حَ}$

$= \text{مج}(\text{س}-\text{و})(\text{ص}-\text{وَ}) - \text{ح . ن حَ}$

$- \text{حَ . ن ح} + \text{ن ح حَ}$

$= \text{مج}(\text{س}-\text{و})(\text{ص}-\text{وَ}) - \text{ن ح حَ}$

(١) واضح أن هاتين العبارتين متساويتان . لأن

$(\text{س} - \text{ص})^2 = \text{س}^2 + \text{ص}^2 - 2\,\text{س ص}$ ،

$\therefore\ \text{مج}(\text{س} - \text{ص})^2 = \text{مج س}^2 + \text{مج ص}^2 - 2\,\text{مج س ص}$ ،

أى أن $2\,\text{مج س ص} = \text{مج س}^2 + \text{مج ص}^2 - \text{مج}(\text{س} - \text{ص})^2$.

(٢) بالانجليزية (Difference Equation for the Correlation Coefficient)

$$\text{ر} = \frac{\text{مج س}^2 + \text{مج ص}^2 - \text{مج}(\text{س} - \text{ص})^2 - 2\,\text{ن}\,\overline{\text{س}}\,\overline{\text{ص}}}{2\,\text{ن ع عَ}} \quad \text{. . (٣) .}$$

وبما أن $\text{مج س}^2 = \text{ن ع}^2 + \text{ن}\,\overline{\text{س}}^2$ ،

و $\text{مج ص}^2 = \text{ن عَ}^2 + \text{ن}\,\overline{\text{ص}}^2$ ،

فلو عوضنا بهذه القيم عن مج س² و مج ص² فى المعادلة (٣) نحصل على صيغة خامسة لمعامل الارتباط وهى :

$$\text{ر} = \frac{1}{2\,\text{ع عَ}}\left[\text{ع}^2 + \text{عَ}^2 + (\overline{\text{س}} - \overline{\text{ص}})^2 - \frac{1}{\text{ن}}\,\text{مج}(\text{س} - \text{ص})^2\right] \quad \text{: (٤)}$$

حينما تكون قيم س هى قيم ص بترتيب مختلف

٢٠٨ — وفى بعض الأحيان يفضل استخدام إحدى هاتين المعادلتين فى حساب معامل الارتباط ، نظراً لسهولة حساب الفروق (س − ص) وتربيعها ، عن ضرب القيم نفسها أو ضرب انحرافاتها . فنرى مثلا فى الحالة الخاصة التى تكون فيها قيم ص هى نفس قيم س ولكن بترتيب مخالف ، أن هذه المعادلة الأخيرة تؤول إلى صورة بسيطة جداً . لأن الوسطين الحسابيين $\overline{\text{س}}$ و $\overline{\text{ص}}$ يكونان حينئذ متساويين ، وكذلك الانحرافان المعياريان . وتؤول المعادلة حينئذ إلى الصورة الآتية :

$$\text{رَ} = 1 - \frac{\text{مج}(\text{س} - \text{ص})^2}{2\,\text{ن . ع}^2} \qquad (٥)$$

حيث وضعنا رَ بدل ر فى هذه الحالة الخاصة ، ووضعنا ع لتقوم مقام الانحرافين المعياريين المتساويين .

الارتباط بين التراتيب . معامل سبيرمان

٢٠٩ — وإذا كانت قيم س تدل على تراتيب مجموعة من الأشخاص (عددهم ن) فى مسابقة معينة ، فى امتحان مادة ما مثلا ، وكانت فى الوقت نفسه قيم ص تدل على تراتيبهم فى مسابقة أخرى ، كانت قيم س ، وكذلك قيم ص ، عبارة عن الأعداد الطبيعية ١ ، ٢ ، ٣ ، . . . إلى ن ، مرتبة ترتيباً خاصاً .

أى أن الوسط الحسابى لقيم س ، أو ص ، يساوى

$$\bar{\text{س}} = \frac{١}{\text{ن}}\left(١+٢+٣\cdots+\text{ن}\right) = \frac{١}{٢}\left(\text{ن}+١\right)$$

والانحراف المعيارى لها ع حيث

$$\text{ن ع}^{٢} = \text{مج س}^{٢} - \text{ن}\,\bar{\text{س}}^{٢}$$

$$= ١^{٢}+٢^{٢}+٣^{٢}\cdots+\text{ن}^{٢} - \frac{\text{ن}}{٤}\left(\text{ن}+١\right)^{٢}$$

$$= \frac{\text{ن}\left(\text{ن}+١\right)\left(٢\text{ن}+١\right)}{٦} - \frac{\text{ن}}{٤}\left(\text{ن}+١\right)^{٢}$$

$$= \frac{١}{١٢}\,.\,\text{ن}\left(\text{ن}+١\right)\left(\text{ن}-١\right)$$

أى أن $\text{ع}^{٢} = \frac{١}{١٢}\,.\,\left(\text{ن}^{٢}-١\right)$

وبالتعويض بهذه القيمة عن ع^{٢} فى المعادلة (٥) نحصل على معامل الارتباط فى هذه الحالة الخاصة جداً ، وهو فى الحقيقة معامل الارتباط بين تراتيب الأشخاص فى المسابقتين . فإذا رمزنا له بالحرف الخاص ر تمييزاً له عن المعامل العام ر ، يكون

$$\text{ر} = ١ - \frac{٦\,\text{مج}\left(\text{س}-\text{ص}\right)^{٢}}{\text{ن}\left(\text{ن}^{٢}-١\right)} \quad . . . \quad (٦).$$

وقد استنبط اسبيرمان[1] هذا المعامل واستخدمه فى أبحاثه الخاصة فى علم النفس ، وسيأتى ذكره فى مناسبة أخرى .

---

(١) يسمى هذا المعامل بالانجليزية (Spearman's Rank-Coefficient of Correlation) ويرمز له بالحرف الاغريقى ρ (رو) تمييزاً له عن المعامل العادى الذى يرمز له بالحرف ر . انظر مقالة (W. Stephenson) فى صفحة ٣٣٥ فى عدد يناير سنة ١٩٣٤ من مجلة :

The British Journal of Psychology, General Section, vol, XXIV, Part, 3.

٢١٠ — هنالك معادلتان أخريان لحساب معامل الارتباط ر . وهما تعتمدان على معرفة الانحرافين المعياريين لقيم س وقيم ص ، والانحراف المعيارى للفرق بين القيم المتناظرة من س و ص ، أو الانحراف المعيارى للمجموع . | حساب ر بواسطة الانحراف المعيارى للفرق والمجموع

وهاتان المعادلتان هما[1] :

---

(١) لاثبات هاتين العلاقتين نفرض أن :

قيم س هى $\text{س}_{١}$ ، $\text{س}_{٢}$ ، . . . ، $\text{س}_{\text{ن}}$ ، وسطها الحسابى $\bar{\text{س}}$ ، وانحرافها المعيارى ع

« ص « $\text{ص}_{١}$ ، $\text{ص}_{٢}$ ، . . . $\text{ص}_{\text{ن}}$ « « $\bar{\text{ص}}$ « « ع'

الفروق « $\text{ف}_{١} = \text{س}_{١} - \text{ص}_{١}$ ، $\text{ف}_{٢} = \text{س}_{٢} - \text{ص}_{٢}$ ، . . .

وسطها الحسابى $\bar{\text{ف}} = \bar{\text{س}} - \bar{\text{ص}}$ ،

والمجاميع « $\text{ح}_{١} = \text{س}_{١} + \text{ص}_{١}$ ، $\text{ح}_{٢} = \text{س}_{٢} + \text{ص}_{٢}$ ، . . .

وسطها الحسابى $\bar{\text{ح}} = \bar{\text{س}} + \bar{\text{ص}}$

$\therefore\ \text{ن ع}_{\text{ف}}^{٢} = \text{مج}\left(\text{ف}-\bar{\text{ف}}\right)^{٢}$ ، $\text{ون ع}_{\text{ح}}^{٢} = \text{مج}\left(\text{ح}-\bar{\text{ح}}\right)^{٢}$

$\therefore\ \text{ف}_{١} - \bar{\text{ف}} = \left(\text{س}_{١} - \text{ص}_{١}\right) - \left(\bar{\text{س}} - \bar{\text{ص}}\right)$

$\text{و ح}_{١} - \bar{\text{ح}} = \left(\text{س}_{١} + \text{ص}_{١}\right) - \left(\bar{\text{س}} + \bar{\text{ص}}\right)$

$\text{ف}_{١} - \bar{\text{ف}} = \left(\text{س}_{١} - \bar{\text{س}}\right) - \left(\text{ص}_{١} - \bar{\text{ص}}\right)$

$\text{و ح}_{١} - \bar{\text{ح}} = \left(\text{س}_{١} - \bar{\text{س}}\right) + \left(\text{ص}_{١} - \bar{\text{ص}}\right)$

$\therefore\ \left(\text{ف}_{١}-\bar{\text{ف}}\right)^{٢} = \left(\text{س}_{١} - \bar{\text{س}}\right)^{٢} + \left(\text{ص}_{١} - \bar{\text{ص}}\right)^{٢}$

$- ٢\left(\text{س}_{١} - \bar{\text{س}}\right)\left(\text{ص}_{١} - \bar{\text{ص}}\right)$ ،

$\text{و}\left(\text{ف}_{٢}-\bar{\text{ف}}\right)^{٢} = \left(\text{س}_{٢} - \bar{\text{س}}\right)^{٢} + \left(\text{ص}_{٢} - \bar{\text{ص}}\right)^{٢}$

$- ٢\left(\text{س}_{٢} - \bar{\text{س}}\right)\left(\text{ص}_{٢} - \bar{\text{ص}}\right)$ ،

.. .. ،

$\text{و}\left(\text{ف}_{\text{ن}}-\bar{\text{ف}}\right)^{٢} = \left(\text{س}_{\text{ن}} - \bar{\text{س}}\right)^{٢} + \left(\text{ص}_{\text{ن}} - \bar{\text{ص}}\right)^{٢}$

$- ٢\left(\text{س}_{\text{ن}} - \bar{\text{س}}\right)\left(\text{ص}_{\text{ن}} - \bar{\text{ص}}\right)$ ؛ وبالجمع :

$\therefore\ \text{ع}_{\text{ف}}^{٢} = \text{ع}^{٢} + \text{ع}'^{٢} - ٢\,.\,\text{ر}\,.\,\text{ع}\,.\,\text{ع}'$ ،

وبالمثل بالنسبة إلى المجاميع ح :

$\therefore\ \text{ع}_{\text{ح}}^{٢} = \text{ع}^{٢} + \text{ع}'^{٢} + ٢\,.\,\text{ر}\,.\,\text{ع}\,.\,\text{ع}'$.

$$ر = \frac{\text{ع}^{2} + \text{عَ}^{2} - \text{ع}^{2}_{\text{ف}}}{\text{٢ ع . عَ}} \quad . \; . \; . \; . \quad (١) \;؛$$

أو
$$ر = \frac{\text{ع}^{2}_{\text{مج}} - \text{ع}^{2} - \text{عَ}^{2}}{\text{٢ ع . عَ}} \quad . \; . \; . \; . \quad (٢) \;؛$$

حيث ع وعَ هما الانحرافان المعياريان السينى والصادى على الترتيب ؛ وعف هو الانحراف المعيارى للفرق (س — ص) ؛ وعمج هو الانحراف المعيارى للمجموع (س + ص) .

وفى بعض الأحيان نجد حساب ر بواسطة إحدى هاتين المعادلتين — وخصوصاً الأولى منهما — أسهل منه بواسطة المعادلات السابقة .

حساب المعامل ر عمليا

٢١١ — لنأخذ مثالا عملياً ونحسب معامل الارتباط بهذه الطرق المختلفة . نأخذ مثلا أرقام[١] نسبة البطالة فى انجلترا وقيمة صادراتها بملايين الجنيهات فى العشر السنوات ١٩٢١—١٩٣٠ ، ونحسب معامل الارتباط بين هاتين الظاهرتين الاقتصاديتين . ونقتصر هنا على عشر قيم فقط لكل من المتغيرين ، حتى نرى خطوات العمل واضحة . ولكن فى المسائل العادية يجب ألا يقل عدد الحالات التى نبحثها عن ثلاثين حالة . وإلا كان تأثير الحالات الشاذة كبيراً جداً ، يحدث خطأ كبيراً فى النتيجة . ونجد الأرقام فى جدول ٢٩

نرى من الجدول أن س̄ = ١٢ر٤١ ، و ص̄ = ٧١٤ر١ ؛

وأن مج (س — س̄)٢ = ١٠ ع٢ = ٥٨ر٤٢٩ ؛

و مج (ص — ص̄)٢ = ١٠ عَ٢ = ٣٨٣٤٦ر٩ ؛

(١) الأرقام مأخوذة عن (Recueil Internationale du Statistique, 1919-1930) طبعة المعهد الدولى للاحصاء فى لاهاى سنة ١٩٣٤ (صفحة ١١٢ — ١١٣) .

و مج (س — س̄) (ص — ص̄) = — ٨٧٣ر٦١٠ ؛

∴ ع = ٢ر٤١٧٢ ، و عَ = ٦١ر٩٢٤٩ ؛

∴ ر = — $\frac{\text{٨٧٣ر٦١٠}}{\text{١٠ × ١٤٩ر٦٨٤٩}}$ ، حسب المعادلة (١) بند ٢٠٣

= — ٠ر٥٨٤ ؛

أى أنه ارتباط عكسى . وهذا هو المنتظر فى مثل هذه الحالة . إذ أن زيادة قيمة الصادرات تكون على العموم مصحوبة بزيادة فى كميتها ، أى بزيادة فى الإنتاج الصناعى ونقص فى عدد العمال العاطلين .

جدول ٢٩ — حساب معامل الارتباط بين نسبة البطالة وقيمة الصادرات فى انجلترا .

السنة	نسبة البطالة س	قيمة الصادرات ص	س — س̄	ص — ص̄	(س — س̄)٢	(ص — ص̄)٢	(س — س̄) × (ص — ص̄)
١٩٢١	١٧ر٠٪	٧٠٣	٤ر٥٩	— ١١ر١	٢١ر٠٦٨١	١٢٣ر٢١	— ٥٠ر٩٤٩
٢٢	١٤ر٣	٧٢٠	١ر٨٩	٥ر٩	٣ر٥٧٢١	٣٤ر٨١	١١ر١٥١
٢٣	١١ر٧	٧٦٧	— ر٧١	٥٢ر٩	ر٥٠٤١	٢٧٩٨ر٤١	— ٣٧ر٥٥٩
٢٤	١٠ر٣	٨٠١	— ٢ر١١	٨٦ر٩	٤ر٤٥٢١	٧٥٥١ر٦١	— ١٨٣ر٣٥٩
٢٥	١١ر٣	٧٧٣	— ١ر١١	٥٨ر٩	١ر٢٣٢١	٣٤٦٩ر٢١	— ٦٥ر٣٧٩
٢٦	١٢ر٥	٦٥٣	ر٠٩	— ٦١ر١	ر٠٠٨١	٣٧٣٣ر٢١	— ٥ر٤٩٩
٢٧	٩ر٧	٧٠٩	— ٢ر٧١	— ٥ر١	٧ر٣٤٤١	٢٦ر٠١	١٣ر٨٢١
٢٨	١٠ر٨	٧٢٤	— ١ر٦١	٩ر٩	٢ر٥٩٢١	٩٨ر٠١	— ١٥ر٩٣٩
٢٩	١٠ر٤	٧٢٠	— ٢ر٠١	٥ر٩	٤ر٠٤٠١	٣٤ر٨١	— ١٩ر٨٥٩
٣٠	١٦ر١	٥٧١	٣ر٦٩	— ١٤٣ر١	١٣ر٦١٦١	٢٠٤٧٧ر٦١	— ٥٢٨ر٠٣٩
المجموع	١٢٤ر١	٧١٤١			٥٨ر٤٢٩٠	٣٨٣٤٦ر٩٠	— ٨٧٣ر٦١٠
المتوسط	١٢ر٤١	٧١٤ر١			٥ر٨٤٢٩	٣٨٣٤ر٦٩	

طريقة أخرى لحساب ر

٢١٢ — يمكننا أن نتفادى الكسور الموجودة فى الوسطين الحسابيين والانحرافات عنهما، وعمليات ضربها الموجودة فى الأعمدة الثلاثة الأخيرة من جدول ٢٩ ، نسير فى العمل بالخطوات المبينة فى الجدول الآتى :

جدول ٣٠ — حساب معامل الارتباط بين نسبة البطالة وقيمة الصادرات فى انجلترا

السنة	نسبة البطالة س	قيمة الصادرات ص	س^{٢}	ص^{٢}	س × ص
١٩٢١	١٧ر٠	٧٠٣	٢٨٩ر٠٠	٤٩٤٢٠٩	١١٩٥١ر٠
٢٢	١٤ر٣	٧٢٠	٢٠٤ر٤٩	٥١٨٤٠٠	١٠٢٩٦ر٠
٢٣	١١ر٧	٧٦٧	١٣٦ر٨٩	٥٨٨٢٨٩	٨٩٧٣ر٩
٢٤	١٠ر٣	٨٠١	١٠٦ر٠٩	٦٤١٦٠١	٨٢٥٠ر٣
٢٥	١١ر٣	٧٧٣	١٢٧ر٦٩	٥٩٧٥٢٩	٨٧٣٤ر٩
٢٦	١٢ر٥	٦٥٣	١٥٦ر٢٥	٤٢٦٤٠٩	٨١٦٢ر٥
٢٧	٩ر٧	٧٠٩	٩٤ر٠٩	٥٠٢٦٨١	٦٨٧٧ر٣
٢٨	١٠ر٨	٧٢٤	١١٦ر٦٤	٥٢٤١٧٦	٧٨١٩ر٢
٢٩	١٠ر٤	٧٢٠	١٠٨ر١٦	٥١٨٤٠٠	٧٤٨٨ر٠
٣٠	١٦ر١	٥٧١	٢٥٩ر٢١	٣٢٦٠٤١	٩١٩٣ر١
مجموع . .	١٢٤ر١	٧١٤١	١٥٩٨ر٥١	٥١٣٧٧٣٥	٨٧٧٤٦ر٢
متوسط .	١٢ر٤١	٧١٤ر١			

∴ مج س^{٢} = ١٥٩٨ر٥١ = ١٠ع^{٢} + ١٠$\overline{\text{س}}^{٢}$ ،

∴ ١٠ ع^{٢} = ١٥٩٨ر٥١ — ١٠ (١٢ر٤١)^{٢} ،

= ٠٥٨ر٤٢٩

مج ص^{٢} = ٥١٣٧٧٣٥ = ١٠ع$'^{٢}$ + ١٠ $\overline{\text{ص}}^{٢}$ ،

∴ ١٠ ع$'^{٢}$ = ٥١٣٧٧٣٥ — ١٠ (٧١٤ر١)^{٢} ،

= ٠٣٨٣٦٥ر٩

مج س ص = ٨٧٧٤٦ر٢ و ن $\overline{\text{س}}$ $\overline{\text{ص}}$ = ١٠ × ٨٨٦١ر٩٨١

∴ مج س ص — ن $\overline{\text{س}}$ $\overline{\text{ص}}$ = — ٨٧٣ر٦١

= مج (س — $\overline{\text{س}}$) (ص — $\overline{\text{ص}}$) .

∴ ر = — ٥٨٤ر

ويلاحظ أن العمليات الحسابية هنا كبيره أيضا ، ولكننا نحصل على نفس القيم للانحرافين المعيار يين ومعامل الارتباط .

حساب ر باختيار وسطين فرضيين

٢١٣ — نشرح الآن طريقة حساب ر باختيار وسط فرضى مناسب لكل من س و ص . ويستحسن اختيار إحدى القيم المعطاة كوسط فرضى ، بحيث تكون قريبة من الوسط الحسابى . ولو اخترنا وسطى س و ص الفرضيين فى سطرين مختلفين فى الجدول ، حصلنا على انحرافين كل منهما يساوى صفراً ، ينتجان حاصلى ضرب كل منهما يساوى صفراً أيضاً . وهذا مما يسهل عمليات الضرب . لنأخذ الوسط الفرضى لقيم س يساوى ١٢ر٥ (وهى نسبة البطالة فى سنة ١٩٢٦) والوسط الفرضى لقيم ص يساوى ٧٠٩ (وهى قيمة الصادرات فى سنة ١٩٢٧) . وهذان الوسطان قريبان من الوسطين الحسابيين ، وهما ١٢ر٤١ و ٧١٤ر١ على الترتيب .

ونرى خطوات العمل موضحة فى الجدول رقم ٣١ :

مج (س — ١٢ر٥)^{٢} = ٥٨ر٥١ = ١٠ع^{٢} + ١٠($\overline{\text{س}}$ — ١٢ر٥)^{٢} .

١٠ ع^{٢} = ٥٨ر٥١ — ١٠ (— ٠٩ر)^{٢} ،

= ٥٨ر٤٢٩ .

و مج (ص − ٧٠٩)^٢ = ٣٨٦٠٧ = ١٠ عَ^٢ + ١٠ (صَ − ٧٠٩)^٢

∴ ١٠ عَ^٢ = ٣٨٦٠٧ − ١٠ (٥ر١)^٢

= ٣٨٣٤٦ر٩ .

و مج (س − ١٢ر٥) (ص − ٧٠٩) − ١٠ (١٢ر٤١ − ١٢ر٥) (٧١٤ر١ − ٧٠٩)

= − ٨٧٨ر٢ − ١٠ × − ٠ر٠٩ × ٥ر١ ،

= − ٨٧٣ر٦١ .

∴ ر = (مج (س − و) (ص − وَ) − ن ح حَ) / (ن ع . عَ)

= − ٨٧٣ر٦١ / ١٤٩٦ر٨٤٩

= − ٠ر٥٨٤ :

وهى نفس النتائج التى حصلنا عليها من قبل [١] .

واضح من هذا المثال أن الطريقة الأخيرة أسهل من الطريقتين المتقدمتين ، وأنه يمكن للانسان أن يختار أوساطاً فرضية مناسبة تختصر العمل الحسابى إلى حد كبير . والواقع أننا نستعمل هذه الطريقة فى أغلب الحالات إن لم يكن كلها ؛ وإنما أوردنا الطرق الثلاث لمعرفة خطوات العمل فى كل طريقة ؛ وليتبين أيضا أن النتائج التى تصل إليها واحدة مهما كانت الطريقة المتبعة . وعلى كل حال فالاختيار بين هذه الطرق الثلاث يتوقف على ظروف المسألة التى نحن بصددها وسهولة الأرقام المركبة منها أو صعوبتها .

(١) يمكن للقارىء أن يحسب ر بواسطة إحدى المعادلتين المذكورتين فى بند ٢١٠ فنأخذ بدل العمود الأخير فى جدول ٣١ قيم الفروق س — ص مثلا ، ثم نربعها ونحسب انحرافها المعيارى ، ونعوضه فى المعادلة (١) بند ٢١٠ .

جدول ٣١ — حساب معامل الارتباط بين نسبة البطالة وقيمة الصادرات فى انجلترا

السنة	نسبة البطالة س	قيمة الصادرات ص	س — ١٢ر٥	ص — ٧٠٩	(س-١٢٥)^٢	(ص — ٧٠٩)^٢	(س — ١٢ر٥) × (ص - ٧٠٩)
١٩٢١	١٧ر٠	٧٠٣	٤ر٥	— ٦	٢٠ر٢٥	٣٦	— ٢٧ر٠
٢٢	١٤ر٣	٧١٠	١ر٨	١١	٣ر٢٤	١٢١	١٩ر٨
٢٣	١١ر٧	٧٦٧	— ر٨	٥٨	ر٦٤	٣٣٦٤	— ٤٦ر٤
٢٤	١٠ر٣	٨٠١	— ٢ر٢	٩٢	٤ر٨٤	٨٤٦٤	— ٢٠٢ر٤
٢٥	١١ر٣	٧٧٣	— ١ر٢	٦٤	١ر٤٤	٤٠٩٦	— ٧٦ر٨
٢٦	١٢ر٥	٦٥٣	٠	— ٥٦	٠	٣١٣٦	٠
٢٧	٩ر٧	٧٠٩	— ٢ر٨	٠	٧ر٨٤	٠	٠
٢٨	١٠ر٨	٧٢٤	— ١ر٧	١٥	٢ر٨٩	٢٢٥	— ٢٥ر٥
٢٩	١٠ر٤	٧٢٠	— ٢ر١	١١	٤ر٤١	١٢١	— ٢٣ر١
٣٠	١٦ر١	٥٧١	٣ر٦	— ١٣٨	١٢ر٩٦	١٩٠٤٤	— ٤٩٦ر٨
المجموع	١٢٤ر١	٧١٤١			٥٨ر٥١	٣٨٦٠٧	— ٨٧٨ر٢
المتوسط	١٢ر٤١	٧١٤ر١					

٢١٤ — ونكرر هنا أن عدد الحالات التى أخذناها فى هذا المثال صغير جداً ، ويجعل النتيجة تحت رحمة أخطاء المصادفات إلى حد كبير . لأن زيادة كبيرة تحدث عن طريق المصادفة فى قيمة الصادرات فى أى سنة لسبب ما (لارتفاع الأسعار فجأة مثلا ، أو فى نسبة البطالة) ، قد يترتب عليها خطأ كبير فى النتيجة . ولذلك يجب أن يكون عدد الحالات التى نأخذها كبيراً — ٣٠ على الأقل — حتى يكون هناك فرصة لتعادل تأثير الحالات الشاذة مع بعضها .

يجب ألا يقل عدد الحالات التى تحسب منها ر عن ٣٠

ولكن إذا زاد عدد الحالات كثيراً ، فلا شك أن العمل الحسابى يكون مرهقاً للغاية لو احتفظنا بكل قيمة على انفرادها مع نظيرتها كما فعلنا فى المثال السابق . فلا بد إذن أن نفكر فى طريقة مختصرة نستخدمها عندما يكون لدينا

عــدد كبير من الحالات — ٢٠٠ أو ٣٠٠ مثلا — بحيث لا تكلفنا مجهوداً كبيراً زيادة عن اللزوم ، وفى الوقت نفسه تعطينا نتائج دقيقة .

جدول الارتباط

٢١٥ — لنفرض أننا نبحث فى العلاقة بين عمر الرجل وعدد ما عنده من الأطفال ؛ وأننا بحثنا حالة ٢٠٠ رجل متزوج ، وعرفنا عمر كل رجل وعــدد أطفاله ، فحصلنا على البيان الآتى :

رقم مسلسل	الأســم	العمر	عدد الأطفال
١	أحمد	٢٨	٣
٢	إبراهيم	٢٢	٠
٣	حسين	٤١	٤
..	..	..	..
..	..	..	..
٢٠٠	يوسف	٣٨	٦

نقسم هؤلاء الرجال إلى فئات مناسبة من حيث أعمارهم ولتكن هذه الفئات هى : ٢٠ — ٢٥ و ٢٥ — ٣٠ و ٠٠٠ و ٤٠ — ٤٥ سنة .

وكذلك نقسمهم إلى فئات من حيث عدد الأطفال . وليكن عدد الأطفال فى هــذه الفئات هو : ٠ و ١ و ٢ و ٠٠٠ و ٦ ثم نوزع هؤلاء الرجال فى « **جدول تكرارى مزدوج** » أو « **جدول ارتباط**[1] » كالآتى :

(١) اسمــه بالانجليزية (Double Frequency Table) أو (Correlation Table)

جدول ٣٢ — توزيع تكرارى مزدوج لأعمار وعدد أطفال ٢٠٠ رجل[1]

العمر / عدد الأطفال	٢٠ —	٢٥ —	٣٠ —	٣٥ —	٤٠ —	المجموع
٠	٦	٨	٩	١		٢٤
١	١	٢٥	١١	٧	٣	٤٧
٢	٤	١٣	١٥	٢٠	٦	٥٨
٣		٦	١٨	١٠	٧	٤١
٤		١	٦	٨	٥	٢٠
٥			١	٢	٥	٨
٦				٢		٢
المجموع	١١	٥٣	٦٠	٥٠	٣٦	٢٠٠

فنرى مثــلا أن أحمد يدخل فى الخانة ملتقى العمود الثــالث بالسطر الخامس لأن عمره ٢٨ ســنة (أى فى الفئة ٢٥ — ٣٠) وعدد أطفاله ٣ . وبالمثل يوضع إبراهيم فى الخانة ملتقى العمود الثانى بالسطر الثانى لأن عمره ٢٢ سنة وليس عنده أطفال بالمرة . وهكذا إلى آخر الكشف حيث يوضع يوسف فى الخانة الأخــيرة من العمود الخامس . وبعد توزيع كل الرجال على الخانات بهذه الطريقة نعد الحالات التى فى كل خانة ، فنحصل على التكرارات الموجودة فى الجدول أعلاه . وهو جدول تكرارى لأن العــدد ١٣ مثلا ، الذى نراه فى ملتقى العمود الثــالث

(١) البيــانات مأخوذة من بحث عمله المؤلف عن الحالة المعيشــية للعمال فى القاهرة سنة ١٩٣٥ .

بالسطر الرابع ، هو عدد الرجال الذين عمرهم بين ٢٠ و ٢٥ سنة وعند كل واحد منهم ٢ من الأطفال .

أما عمود المجاميع الأخير فهو عبارة عن التوزيع التكرارى لعدد الأطفال عند هؤلاء الرجال ، حيث منهم ٢٤ ليس عندهم أطفال بالمرة و ٤٧ عند كل منهم طفل واحد ، و ٥٨ عند كل منهم اثنان ، وهكذا .

وكذلك السطر الأخير ، فهو توزيع تكرارى لأعمار هؤلاء الرجال : فنرى منهم ١١ رجلا أعمارهم بين ٢٠ و ٢٥ سنة ، و ٥٣ رجلا أعمارهم بين ٢٥ و ٣٠ سنة . وهكذا .

٢١٦ — لحساب معامل الارتباط هنا نوجد أولا الوسطين الحسابيين والانحرافين المعياريين للأعمار وعدد الأطفال ، وذلك باستخدام التوزيعين التكراريين فى السطر الأخير والعمود الأخير .

جدول ٣٣ — إيجاد الوسط الحسابى والانحراف المعيارى لأعمار الرجال فى جدول ٣١

فئات العمر	مراكز الفئات س	التكرار ك	س — ٣٢ر٥ ح	ك . ح	ك . ح^{2}
٢٠ —	٢٢ر٥	١١	—١٠	—١١٠	١١٠٠
٢٥ —	٢٧ر٥	٥٣	—٥	—٢٦٥	١٣٢٥
٣٠ —	٣٢ر٥	٦٠	٠	٠	٠
٣٥ —	٣٧ر٥	٥٠	٥	٢٥٠	١٢٥٠
٤٠ —	٤٢ر٥	٢٦	١٠	٢٦٠	٢٦٠٠
		٢٠٠		١٣٥	٦٢٧٥

∴ الوسط الحسابى للأعمار هو :

س̄ = ٣٢ر٥ + $\frac{١٣٥}{٢٠٠}$

= ٣٣ر١٧٥ سنة

و ٦٢٧٥ = ٢٠٠ ع^{2} + ٢٠٠ (ر٦٧٥)^{2}

∴ ع^{2} = ٣١ر٣٧٥ — ر٤٥٥٦

= ٣٠ر٩١٩٤

∴ ع = ٥ر٥٦٠٥ سنة .

جدول ٣٤ — إيجاد الوسط الحسابى والانحراف المعيارى لعدد الأطفال عند الرجال فى جدول ٣٢

عدد الأطفال ص	التكرار كَ	ص — ٢ حَ	كَ . حَ	كَ . حَ^{2}
٠	٢٤	—٢	—٤٨	٩٦
١	٤٧	—١	—٤٧	٤٧
٢	٥٨	٠	٠	٠
٣	٤١	١	٤١	٤١
٤	٢٠	٢	٤٠	٨٠
٥	٨	٣	٢٤	٧٢
٦	٢	٤	٨	٣٢
	٢٠٠		١٨	٣٦٨

∴ الوسط الحسابى لعدد الأطفال هو :

صَ = ٢ + ١٨ / ٢٠٠

= ٢ر٠٩ طفلا

و ٣٦٨ = ٢٠٠ عَ٢ + ٢٠٠ (٠٩ر)٢

∴ عَ٢ = ١ر٨٤ — ٠٠٨١ر

= ١ر٨٣١٩

∴ عَ = ١ر٣٥٣٥ طفلا .

أحسن طريقة لحساب معامل الارتباط فى مثل هذه المسائل ، هى أن نأخذ وسطين فرضيين مناسبين للسينات والصادات ، ونحسب حواصل ضرب الانحرافات عن هذين الوسطين ، ونستخدم المعادلة

ر = [مج (س — و) (ص — وَ) — ن ح حَ] / [ن . ع . عَ]

ويستحسن عملياً أن نختار مركز إحدى الفئات السينية كوسط فرضى للسينات ، وكذلك فى الصادات .

لنأخذ الوسط الفرضى للأعمار و = ٣٢ر٥ سنة ، والوسط الفرضى لعدد الأطفال وَ = ٢ .

∴ ح = سَ — و = ٣٣ر١٧٥ — ٣٢ر٥ = ٠ر٦٧٥

، حَ = صَ — وَ = ٢ر٠٩ — ٢ر٠٠ = ٠٩ر

∴ ن ح حَ = ٢٠٠ × ٦٧٥ر × ٠٩ر = ١٢ر١٥٠ .

نحسب الانحرافات السينية عن ٣٢ر٥ ، والانحرافات الصادية عن ٢ . فنطرح ٣٢ر٥ من مراكز الفئات السينية المبينة فى رؤوس الأعمدة فى جدول الارتباط ، ونطرح ٢ من مراكز الفئات الصادية المبينة فى العمود الأيمن من نفس الجدول . وطبعاً يكون الانحراف السينى لأى خانة فى الجدول هو نفس الانحراف السينى للعمود الذى تقع فيه هذه الخانة ؛ ويكون الانحراف الصادى لها نفس الانحراف الصادى للسطر الذى تقع فيه أيضاً .

لكل خانة نحسب حاصل الضرب :

تكرار الخانة × انحرافها السينى × انحرافها الصادى

ومجموع هذه الحواصل لجميع الخانات يساوى الكمية التى نبحث عنها ، وهى مج (س — و) (ص — وَ) .

إذا اخترنا الوسطين الفرضيين عند مركزى فئتين ، فسنجد عملياً أن الحواصل فى عمود الوسط الفرضى السينى ، وفى سطر الوسط الفرضى الصادى ، كلها تساوى صفرا . فيمكننا إذن أن نهمل هذا العمود وهذا السطر بالمرة ، ونشطبهما من الجدول .

ولتسهيل عمليات الضرب وضمان الدقة وعدم السهو فى العمل ، نكتب الانحراف السينى لكل خانة فى ركن معين منها ، والانحراف الصادى لها فى ركن آخر معين أيضاً ، وحاصل ضرب الاثنين فى التكرار فى ركن ثالث من الخانة نفسها . ويجب تمييز الكتابة فى هذه الأركان المختلفة فتستعمل ألوان مختلفة من الحبر أو الرصاص مثلا ، منعاً للالتباس . ونرى هذه الخطوات موضحة فى الجدول الآتى :

ويلاحظ أننا هنا كتبنا الانحرافات السينية فى الركن الجنوبى الغربى من كل خانة ، والانحراف الصادى فى الركن الجنوبى الشرقى . وكتبنا الحاصل النهائى فى الركن الشمالى الشرقى .

جدول ٣٥ — حساب معامل الارتباط من جدول تكرارى مزدوج باختيار وسطين فرضيين

ص \ س	٢٢٫٥	٢٧٫٥	٣٢٫٥	٣٧٫٥	٤٢٫٥	المجموع	الحواصل
٠	١٢٠ ٦ ٢- (١٠-)	٨٠ ٨ ٢- (٥-)	٩	١٠- ١ ٢- (٥)		٢٤	٢٠٠ ١٠-
١	١٠ ١ ١- (١٠-)	١٢٥ ٢٥ ١- (٥-)	١١	٢٥- ٧ ١- (٥)	٣٠- ٣ ١- (١٠)	٤٧	١٢٥ ٦٥-
٢	٤	١٣	١٥	٢٠	٦	٥٨	
٣		٣٠- ٦ ١ (٥-)	١٨	٥٠ ١٠ ١ (٥)	٧٠ ٧ ١ (١٠)	٤١	١٢٠ ٣٠-
٤		١٠- ١ ٢ (٥-)	٦	٨٠ ٨ ٢ (٥)	١٠٠ ٥ ٢ (١٠)	٢٠	١٨٠ ١٠-
٥			١	٣٠ ٢ ٣ (٥)	١٥٠ ٥ ٣ (١٠)	٨	١٨٠
٦				٤٠ ٢ ٤ (٥)		٢	٤٠
المجموع	١١	٥٣	٦٠	٥٠	٢٦	٢٠٠	٨٥٥ ١١٥-
الحواصل	١٣٠	٢٠٥ ٤٠-		٢٠٠ ٤٥-	٣٢٠ ٣٠-	٨٥٥ ١١٥-	٧٤٠

ويلاحظ أن مجموع الحواصل واحد ، سواء حسبناه أفقياً بالسطور أو رأسياً بالأعمدة . ويحسن مراجعة المجموعين على بعضهما منعاً للخطأ . والمجموع النهائى يساوى ، كما قلنا من قبل ،

$$مج (س - و) (ص - وَ) = ٧٤٠$$

$$ر = \frac{٧٤٠ - ١٢ر١٥}{٢٠٠ \times ٥ر٥٦٠٥ \times ١ر٣٥٣٥}$$

$$= \frac{٧٢٧ر٨٥}{١٥٠٥ر٢٢٧}$$

$$= ٠ر٤٨٤$$

وهذا المعامل يدل على وجود علاقة طردية بين سن الرجل وعدد الأطفال الذين يعولهم ؛ وهـذا هو المنتظر بطبيعة الحال . ومقدار المعامل ر فى هذه الحالة يدل على أن الارتباط شديد نوعا .

طريقة المجموعات القطرية .

٢١٧ — نشرح الآن طريقة أسهل لحساب ر من الجدول التكرارى المزدوج ، وهى مبنية على الفكرة التى شرحناها فى بند ٢١٠ . وهى كما يأتى :

أقطار الفروق المتساوية

فى جدول ٣٢ صفحة ٢٣٥ ، نأخذ الأقطار النازلة من اليمين إلى اليسار[1] ، وهى التى تبدأ من الخانات ذات القيم الصغرى لكل من س و ص وتنتهى بالقيم الكبرى لها . ونجمع تكرارات الخانات الواقعة على كل قطر من هـذه الأقطار المتوازية . وطبعاً ستكون الخانات المتتالية على أى قطر ، مشتركة فى ركن واحد دائماً . ونأخذ حواصل جمع التكرارات لكل قطر على حدة ، ونكتبها بالترتيب حسب هذه الأقطار .

وفيما يلى بيان بتكرارات الفئات الواقعة على الأقطار بالترتيب من أعلى إلى أسفل . وسنكتب تكرارات كل قطر حسب مواقعها على نفس القطر فى جدول ٣٢

(١) هذا طبعاً على فرض أن ترتيب فئات س فى الجدول تصاعدى من اليمين إلى اليسار ، وفئات ص متصاعدة من فوق إلى تحت . وهذا هو النظام المعتاد . إلا أنه أحياناً قد يعكس ترتيب فئات س أو ص أو هما معاً . وعلى كل حال فالمقصود هو الأقطار التى تبدأ من ناحية القيم الصغيرة للمتغيرين معاً ، وتنتهى فى ناحية القيم الكبيرة لهما .

القطر الأول ١ + ٣ = ٤
« الثانى ٩ + ٧ + ٦ = ٢٢
« الثالث ٨ + ١١ + ٢٠ + ٧ = ٤٦
« الرابع ٦ + ٢٥ + ١٥ + ١٠ + ٥ = ٦١
« الخامس ١ + ١٣ + ١٨ + ٨ + ٥ = ٤٥
« السادس ٤ + ٦ + ٦ + ٢ = ١٨
« السابع ١ + ١ + ٢ = ٤
٢٠٠

نعتبر فئات س كأنها مراتب لقيم س، وفئات ص كأنها مراتب متتالية لقيمها، فيكون لدينا خمس مراتب سينية وسبع صادية. وبالتأمل فى الخانات التى على القطر الأول، نجد فى كل منها أن الفرق بين مرتبتى س و ص يساوى + ٣، وفى القطر الثانى الفرق يساوى + ٢، ...، وهكذا إلى القطر السابع حيث الفرق بين مرتبتى س وص فى كل خانة يساوى —٣. وهكذا نسمى هذه الأقطار **أقطار الفروق المتساوية**

٢١٨ — يمكننا إذن اعتبار حواصل جمع التكرارات على هذه الأقطار كأنها تكرارات لهذه الفروق المختلفة بين مراتب س و ص، ثم نطبق الفكرة التى أوردناها فى بند ٢١٠، بخصوص الانحراف المعيارى للفرق بين قيم س و ص. ونسير فى العمل كما فى الجدول الآتى:

$$\therefore ٢٠٠\ ع_{ف}^{٢} = ٣٢٣ - \frac{(٩)^{٢}}{٢٠٠}$$
$$= ٣٢٢٫٥٩٥$$
$$= حـ \text{ مثلا.}$$

جدول ٣٦ — حساب الانحراف المعيارى للفرق بين مرتبتى س و ص.

الفروق ف	التكرارات ك	ك . ف	ك . ف^{٢}
٣	٤	١٢	٣٦
٢	٢٢	٤٤	٨٨
١	٤٦	٤٦	٤٦
٠	٦١	٠	٠
—١	٤٥	—٤٥	٤٥
—٢	١٨	—٣٦	٧٢
—٣	٤	—١٢	٣٦
	٢٠٠	٩	٣٢٣

يلاحظ أنه كان من الممكن الاستغناء عن كتابة العمود الأول من هذا الجدول، وكتابة صفر أمام التكرار الأوسط، أو أى تكرار قريب منه؛ وكتابة —١ و —٢ — و٣ إلى أعلى و +١ و +٢ و +٣ إلى أسفل، أو العكس، بدون تأثير فى النتيجة النهائية.

٢١٩ — ونعمل مثل ذلك فى فئات س باعتبارها مراتب، تكراراتها هى الأرقام الموجودة فى السطر الأخير من جدول ٣٢. وكذلك فى فئات ص باعتبارها مراتب وتكراراتها موجودة فى العمود الأخير فى نفس الجدول. فنرى من جدول ٣٧ أن:

$$٢٠٠\ ع^{٢} = ٢٥١ - \frac{(٢٧)^{٢}}{٢٠٠}$$
$$= ٢٤٧٫٣٥٥$$
$$= أ \text{ مثلا.}$$

جدول ٣٧ — حساب الانحراف المعيارى لمراتب س

مراتب س	تكرارات ك	ع	ع . ك	ع² . ك
١	١١	—٢	—٢٢	٤٤
٢	٥٣	—١	—٥٣	٥٣
٣	٦٠	٠	٠	٠
٤	٥٠	١	٥٠	٥٠
٥	٢٦	٢	٥٢	١٠٤
	٢٠٠		٢٧	٢٥١

جدول ٣٨ — حساب الانحراف المعيارى لمراتب ص

مراتب ص	تكرارات ك	ع	ع . ك	ع² . ك
١	٢٤	—٢	—٤٨	٩٦
٢	٤٧	—١	—٤٧	٤٧
٣	٥٨	٠	٠	٠
٣	٤١	١	٤١	٤١
٥	٢٠	٢	٤٠	٨٠
٦	٨	٣	٢٤	٧٢
٧	٢	٤	٨	٣٢
	٢٠٠		١٨	٣٦٨

وبالمثل نرى من جدول ٣٨ أن :

$$\text{٢٠٠ } \bar{\text{ع}}^{٢} = \text{٣٦٨} - \frac{(\text{١٨})^{٢}}{\text{٢٠٠}}$$

$$= \text{٣٦٦ر٣٨}$$

$$= \text{ب .}$$

لكن $\text{ر} = \dfrac{\text{ع}^{٢} + \bar{\text{ع}}^{٢} - \text{ع}_{\text{ن}}^{٢}}{\text{٢ ع . }\bar{\text{ع}}}$ ، بند ٢١٠ (١)

$$\therefore \quad \text{ر} = \frac{\text{أ} + \text{ب} - \text{هـ}}{\text{٢} \sqrt{\text{أ ب}}}$$

$$= \frac{\text{٢٤٧ر٣٥٥} + \text{٣٦٦ر٣٨} - \text{٣٢٢ر٥٩٥}}{\text{٢} \sqrt{\text{٩٠٦٢٥ر٩٢٤٩}}}$$

$$= \frac{\text{٢٩١ر١٤٠}}{\text{٦٠٢ر٠٨٢٨}}$$

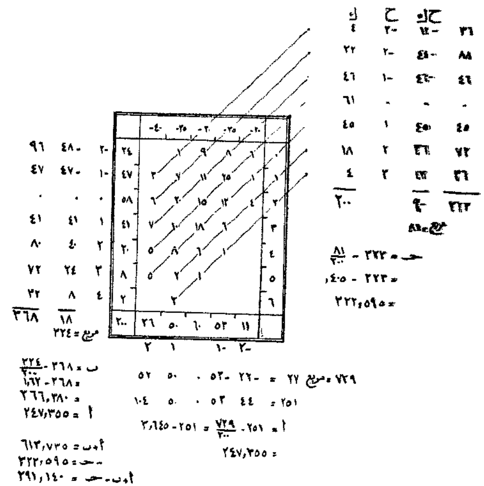
$$= \text{ر٤٨٤}$$

وهى نفس النتيجة التى حصلنا عليها من قبل . ولا شك أن هذه الطريقة أسهل بكثير من الطريقة السابقة فى هذا المثال ، ولكنها فى بعض الأحيان ربما تكون أطول . ويمكننا اختصار العمل إلى حد كبير لو استعملنا ورقة مقسمة إلى مربعات وسرنا فى العمل كما فى جدول ٣٩ .

أقطار المجاميع المتساوية

٢٢٠ — يمكننا أيضاً أن نحسب ر من المعادلة (٢) بند ٢١٠ ، أى باستخدام فكرة الانحراف المعيارى لمجموع مرتبتى س و ص ، بدل الفرق بينهما كما فعلنا فى الطريقة السابقة. فاذا أخذنا الأقطار المتعامدة مع الأقطار السابقة ، وجدنا أن مجموع مرتبتى س و ص فى كل الخانات التى على قطر واحد ثابت . وهذه الأقطار إذن هى **أقطار المجاميع المتساوية** ؛ وهذه بيانها كما يأتى :

جدول (٣٩) لحساب معامل الارتباط



القطر	مجموع المرتبتين	تكرارات الخانات		جملة التكرار
الأول	٢	٦	=	٦
الثاني	٣	١ + ٨	=	٩
الثالث	٤	٤ + ٢٥ + ٩	=	٣٨
الرابع	٥	١٣ + ١١ + ١	=	٢٥
الخامس	٦	٦ + ١٥ + ٧	=	٢٨
السادس	٧	١ + ١٨ + ٢٠ + ٣	=	٤٢
السابع	٨	٦ + ١٠ + ٦	=	٢٢
الثامن	٩	١ + ٨ + ٧	=	١٦
التاسع	١٠	٢ + ٥	=	٧
العاشر	١١	٢ + ٥	=	٧
		مجموع التكرارات	=	٢٠٠

٢٢١ – نحسب الانحراف المعيارى المجموع من هذا التوزيع التكرارى كما حسبنا الانحراف المعيارى للفرق فى بند ٢١٨ . **و بيان ذلك كما يأتى :**

الانحراف المعيارى لمجموع المرتبتين

جدول ٤٠ — حساب الانحراف المعيارى لمجموع مرتبتى س ، ص

مجموع المرتبتين	التكرار ك	الانحراف ع	ك . ع	ك . ع ^{٢}
٢	٦	٤—	٢٤—	٩٦
٣	٩	٣—	٢٧—	٨١
٤	٣٨	٢ —	٧٦—	١٥٢
٥	٢٥	١—	٢٥—	٢٥
٦	٢٨	•	•	•
٧	٤٢	١	٤٢	٤٢
٨	٢٢	٢	٤٤	٨٨
٩	١٦	٣	٤٨	١٤٤
١٠	٧	٤	٢٨	١١٢
١١	٧	٥	٣٥	١٧٥
	٢٠٠		٤٥+	٩١٥

$$\therefore ٢٠٠ \ ع_م^{٢} = ٩١٥ - \frac{(٤٥)^{٢}}{٢٠٠}$$

$$= ٩٠٤{,}٨٧٥$$

$$= م \quad \text{مثلا}$$

وبالمثل نحسب الكميتين ا و ب كما فعلنا فى بند ٢١٩ ، ونعوض فى المعادلة :

$$\text{ر} = \frac{ع^{2}_{+} - ع^{2} - ع^{2}}{2\,ع \cdot ع} \quad \text{... بند ٢١٠ (٢) ،}$$

$$= \frac{\text{م} - \text{ا} - \text{ب}}{2\sqrt{\text{ا ب}}}$$

$$= \frac{\text{٩٠٤ر٨٧٥} - \text{٦١٣ر٧٣٥}}{2\sqrt{\text{٩٢٤٩ر٩٠٦٢٥}}}$$

$$= \frac{\text{٢٩١ر١٤٠}}{\text{٦٠٢ر٨٣٨}}$$

$$= \text{٠ر٤٨٤}$$

وهى نفس النتيجة التى حصلنا عليها من قبل . ولـكن هذه الطريقة فى الغالب تكون أطول من الطريقة المتقدمة ، إذ من الواضح أن الجـدول التـكرارى لمجاميع مرتبتى س و ص يكون أطول من الجدول التـكرارى للفرق بينهما ؛ فقد رأينا هنـا أن مجموع المرتبتين يتغير من ٢ إلى ١١ ، فى حين أن الفرق بينهما يتغير من — ٣ إلى + ٣ .

٢٢٢ — فى الطريقتين الأخيرتين لحساب معامل الارتباط ، تكلمنا فقط عن مرتبتى س و ص والفرق بينهما أو مجموعهما ؛ ولم نتكلم أبداً عن القيم الفعلية التى تأخذها س أو ص . وما كان ينبغى لنا أن نهتم بهـذه القيم هنا ؛ لأن مجموع قيمتى س و ص ، أو الفرق بينهما ، عدد ليس له معنى ، إذ لا معنى لمجموع

عمر رجل (مقدراً بالسنين) وعدد أطفاله ؛ ولا معنى للفرق بينهما أيضاً . وقد تخلصنا من هذه الصعوبة بأن قسمنا القيم إلى مراتب ، يمكن طرحها أو جمعها ، وحصلنا بسهولة على نفس النتيجة التى حصلنا عليها باستخدام القيم نفسها . فهل من الممكن تعميم هذه الفـكرة واستخدامها فى كل المسـائل ؟ إن أمكن هذا فلا شك أننـا نختصر كثيراً من العمل الحسابى الذى يستلزمه حساب معامل الارتبـاط من القيم نفسها .

يجب أن نلاحظ هنـا أن مراتب س و ص منتظمة ، بمعنى أن الفرق بين المرتبة الأولى والثانية ، مثلا يساوى الفرق بين الثالثة أو الرابعة ، أو أى مرتبتين متتاليتين . وهو يساوى ٥ سنين فى حالة س و ١ طفل فى حالة ص . ويجب أن نميز بين هـذه الحالة والحالات الأخرى التى تـكون فيهـا الفروق بين المراتب غير متساوية ، التى نسميها حينئذ « تراتيب » . بدل مراتب .

ومن الواضح طبعاً أنه من الممكن حساب معامل الارتباط باستخدام مراتب القيم ، إذا كانت هذه القيم موزعة فى فئات ذات فترات متساوية ، سواء بالطريقة الأساسية التى شرحناها أولا ، أو بطريقة المجموعات القطرية .

٢٢٣ — أما فى حالة عـدم تساوى الفترات بين القيم المتتالية . فنستخدم معامل الارتبـاط بين الترتيب الذى وضعه اسـبيرمان (صفحة ٢٣٦) حيث نستعمل تراتيب قيم س وتراتيب قيم ص بدل القيم نفسها . معامل الارتباط بين تراتيب س و ص

وهذه المعادله هى كما رأينا فى بند ٢٠٩ :

$$\text{ر} = 1 - \frac{6\,\text{مجـ ف}^{2}}{\text{ن} (\text{ن}^{2} - 1)}$$

حيث ر هى معامل الارتباط (وهو هنا تقريبى بالنسبة إلى المعامل ر) ؛

و ن هى عدد الحالات ؛

و ف هى الفرق بين ترتيب س و ص ، كل فى مجموعتها .

كيفية حساب ر عمليا

٢٢٤ — لنحسب هذا المعامل ر بين تراتيب س و ص فى المثال الوارد فى بند ٢١٠ (صفحة ٢٢٨) ، حيث س تدل على نسبة البطالة بين العمال فى انجلترا ، و ص تدل على قيمة الصادرات ، فى السنين ١٩٢١ — ١٩٣٠ .

نرتب أولا قيم س وقيم ص تصاعدياً (أو تنازلياً) لنعرف ترتيب كل قيمة بين أخواتها . ومن الضرورى أن يكون نظام الترتيب واحداً فى مجموعتى س و ص . إما تصاعدى فى الاثنين معاً أو تنازلى ، حتى نميز بين الارتباط الطردى والارتباط العكسى .

القيم السينية مرتبة تصاعدياً ، هى وتراتيبها كما يأتى :

٩ر٧ ، ١٠ر٣ ، ١٠ر٤ ، ١٠ر٨ ، ١١ر٣ ، ١١ر٧ ، ١٢ر٥ ، ١٤ر٣ ، ١٦ر١ ، ١٧ر٠٠ ؛

١ ، ٢ ، ٣ ، ٤ ، ٥ ، ٦ ، ٧ ، ٨ ، ٩ ، ١٠ ؛

وكذلك القيم الصادية وتراتيبها ، تصاعدياً أيضاً ، هى :

٥٧١ ، ٦٥٣ ، ٧٠٣ ، ٧٠٩ ، ٧٢٠ ، ٧٢٠ ، ٧٢٤ ، ٧٦٧ ، ٧٧٣ ، ٨٠١ ؛

١ ، ٢ ، ٣ ، ٤ ، ٥½ ، ٥½ ، ٧ ، ٨ ، ٩ ، ١٠ .

يلاحظ هنا كيف عالجنا القيمة المتكررة ٧٢٠ ، فأعطينا كل مفردة منهما نفس الترتيب .

نضع هذه التراتيب بدل القيم نفسها الواردة فى جدول ٢٩ (صفحة ٢٢٩) ونحسب ر . وها هى تلك الخطوات موضحة فى الجدول الآتى :

جدول ٤١ — حساب معامل الارتباط الترتيبى ر

بين نسبة البطالة وقيمة الصادرات فى انجلترا

السنة	نسبة البطالة س	قيمة الصادرات ص	تراتيب س	تراتيب ص	الفرق بين الترتيبين ف	$ف^٢$
١٩٢١	١٧ر٠٠	٧٠٣	١٠	٣	٧	٤٩
٢٢	١٤ر٣	٧٢٠	٨	٥½	٢½	٦ر٢٥
٢٣	١١ر٧	٧٦٧	٦	٨	— ٢	٤
٢٤	١٠ر٣	٨٠١	٢	١٠	— ٨	٦٤
٢٥	١١ر٣	٧٧٣	٥	٩	— ٤	١٦
٢٦	١٢ر٥	٦٥٣	٧	٢	٥	٢٥
٢٧	٩ر٧	٧٠٩	١	٤	— ٣	٩
٢٨	١٠ر٨	٧٢٤	٤	٧	— ٣	٩
٢٩	١٠ر٤	٧٢٠	٣	٥½	— ٢½	٦ر٢٥
٣٠	١٦ر١	٥٧١	٩	١	٨	٦٤
						٢٥٢ر٥

وبالتعويض فى المعادلة مجـ $ف^٢$ = ٢٥٢ر٥ ، ن = ١٠

$$\therefore \quad ر = ١ - \frac{٦ \times ٢٥٢ر٥}{١٠ \left(١٠٠ - ١ \right)}$$

$$= ١ - \frac{١٥١٥}{٩٩٠} ،$$

= — ٥٣٠ر ،

بينما ر = — ٥٨٤ر .

ولاشك أن فى هذه الطريقة تسهيلاً كبيراً واختصاراً للعمل الحسابى، حيث نستعيض عن القيم الأصلية ذات الأرقام الكثيرة ، تراتيبها ذات الأرقام المختصرة ، علاوة على أننا غير محتاجين إلى حساب الوسطين الحسابيين أو الانحرافين المعياريين. وقد رأينا أن تأثير هـذا الاختصار على النتيجة النهائية ليس كبير جداً ؛ وربما لا يزيد عن ١٠ أو ١٥ ٪ ، كما هو واضح من مقارنة قيمتى ر ى ر .

معامل الائتلاف

٢٢٥ — يوجد طريقة أخرى أكثر اختصاراً من السابقة لقياس العلاقة أو الارتباط بين ظاهرتين مقيستين ، وذلك باستخدام معادلة وضعها المستر ج . أ . يول . وهذه المعادلة تعطى مقياس العلاقة الذى نسميه [1] **معامل الائتلاف** ونرمز له بالحرف ل وهو :

بول

$$\text{ل} = \frac{\sqrt{\text{أ د}} - \sqrt{\text{ب ح}}}{\sqrt{\text{أ د}} + \sqrt{\text{ب ح}}}$$

حيث أ = عدد الحالات التى فيها كل من س و ص فوق المتوسط

« د = « « « « « س و ص تحت «

« ب = « « « « قيمة س تحت و ص فوق «

« ح = « « « « « س فوق و ص تحت «

(١) هـذا العامل يسمى (G. U. Yule's "Coefficient of Colligation") ويرمز له بالحرف الأغريق (ω) « أوميجا » — انظر مقالته فى : (*Journal of the Royal Statistical Society*, (Vol. 75 (1911 - 1912), pp 579 - 652) وانظر أيضاً انتقادات (K. Pearson and Heron) فى مجلة (*Biometrika*, Vol. 9 (1913) pp 159-315) على هذا العامل ، وعلى معامل الاقتران الذى سيأتى ذكره فى بند ٢٢٨ .

ولحساب هذا المعامل نحسب المتوسط لقيم س وقيم ص ثم نوزع قيم س و ص فى جدول مزدوج من أربع خانات فقط مثل الجدول المبين ، ثم نحسب عدد الحالات التى تقع فى كل خانة . ففى المسألة التى بأيدينا مثلا ، نعلم أن الوسط الحسابى لقيم س (أى نسبة البطالة) هو ١٢ر٤١ ، والوسط الحسابى لقيم ص هو ١ر٧١٤ .

قيم ص	قيم س	فوق المتوسط	تحت المتوسط
	فوق المتوسط	أ	ب
	تحت المتوسط	ح	د

.·. عدد الحالات التى فيها س فوق المتوسط يساوى ٤ ، كما يتضح من جدول ٤١. من هـذه الحالات واحدة فيها ص فوق متوسطها ، وثلاث فيهـا ص تحت متوسطها .

وعدد الحالات التى فيها س أقل من المتوسط يساوى ٦ طبعاً .

ومن هذه الحالات نجد خمساً فيها ص فوق المتوسط ، وواحدة فقط فيها ص تحت المتوسط .

وعلى ذلك يكون توزيع الحالات العشر كما يأتى :

قيم ص	قيم س	+	—
	+	١	٥
	—	٣	١

حيث تدل العلامة + على فوق المتوسط ، والعلامة — علي تحت المتوسط .

$$\therefore \quad ل = \frac{\sqrt{١ \times ١} - \sqrt{٥ \times ٣}}{\sqrt{١ \times ١} + \sqrt{٥ \times ٣}}$$

$$= \frac{١ - ٣ر٨٧٣٠}{١ + ٣ر٨٧٣٠}$$

$$= - ٥٩٠ر$$ ؛

بينما ر $= - ٥٣٠ر$ ،

و ں $= - ٥٨٤ر$.

ونلاحظ أن هـذه المقاييس الثلاثة كلها متقاربة ، وكلها متحدة فى الإشارة السالبة ، دلالة على أن العلاقة عكسية .

٢٢٦ — ويمكن تطبيق هذه الطريقة الأخـيرة بسهولة فى حالة جـدول الارتباط ، حيث يكون عدد الحالات كبيراً. وذلك بأن نرسم خطاً رأسياً فى الجدول يقسم الفئات السينية إلى جزءين : أقل من الوسط الحسابى وأكبر منه . وكذلك نرسم خطاً أفقياً يقسم الفئات الصادية إلى قسمين أيضاً : تحت المتوسـط وفوقه . معامـل الائتـلاف من جـدول الارتبـاط

وهذان المحوران يقسمان الجدول إلى أربعة أقسام ؛ ومجموع تكرارات الخانات الموجودة فى كل قسم تعطينا الأعداد ا ، ب ، حـ ، ى المطلوبة لحسـاب معامل الائتلاف .

ففى جدول ٣٢ مثلا ، نعلم أن $\overline{س}$ = ٣٣ر١٧٥ سنة ، و $\overline{ص}$ = ٢ر٠٠٩ طفلا . وعلى ذلك فالخط الرأسى يقع بين الفئة السينية الثالثة (وهى ٣٠ — ٣٥ ، التى مركزها ٣٢ر٥ أقل من $\overline{س}$) والرابعة . والخط الأفقى يقع بين الفئة الصادية الثالثة (فيها قيم ص = ٢ أقل من $\overline{ص}$) والرابعة .

مجاميع تكرارات الخانات فى الأقسام الأربعة هى كما يأتى :

ص \ س	+	—
+	٣٩	٣٢
—	٣٧	٩٢

$$\therefore \quad ل = \frac{\sqrt{٣٩ \times ٩٢} - \sqrt{٣٢ \times ٣٧}}{\sqrt{٣٩ \times ٩٢} + \sqrt{٣٢ \times ٣٧}}$$

$$= \frac{٥٩ر٨٩٩٩ - ٣٤ر٤٠٩٣}{٥٩ر٨٩٩٩ + ٣٤ر٤٠٩٣}$$

$$= ٢٧٠ر$$ ،

ولـكن ں $= ٤٨٤ر$ ، (انظر صفحة ٢٤١) .

٢٢٧ — و يلاحظ أن الفرق بين المعاملين ل و ں فى هذه الحالة كبير نوعاً .ولو أن فئات الأعمار كانت أضيق مدىً مما هى فى جدول ٣٢ ، لأمكن تقسيم الأعمار فوق المتوسط وتحته ، تقسيما أدق مما فعلنا هنا ، وحصلنا بذلك على قيمة للمعامل ل أدق من ٢٧٠ و ٠ وأقرب إلى ں منها .

و يلاحظ أن المعاملين متحدان فى الإشـارة الموجبة ، حيث يدل كل منهما على أن العلاقة طردية بين السن وعدد الأطفال ، كما هو منتظر .

٢٢٨ — تكلمنا فى البنود السابقة عن طرق قيـاس العلاقة بين الظواهر المتغيرة التى يمكن قياسها رقمياً . و يبقى إذن أن نبحث فى طرق قياس العلاقة بين الصفات التى لا يمكن قياسها . وقد سبق أن ذكرنا أننا نسمى هـذه العلاقة « الاقتران » بين الصفات . ومقياس هذه العلاقة نسميه « معامل الاقتران » . الاقتران بين الصفات غير المقيسة

لنفرض أننا نبحث فى العلاقة بين جنسية التاجر (فى القاهرة مثلا) ونوع العمل الذى يقوم به . ولنفرض أن البيانات التى لدينا تقسم التجار من حيث الجنسية إلى صنفين فقط : مصريين وأجانب ؛ وتقسم الأعمال التجارية نوعين أيضاً : أعمالا مالية صرفة ، وأعمالاً تجارية للبيع والشراء . ففى القاهرة مثلا نجد عدد المشتغلين بالأعمال التجارية فى سنة ١٩٢٧ يساوى ٩٣٧٠٠ . من هؤلاء ٨٢٢٠٠ مصريون و ١١٥٠٠ أجانب . ومن المصريين ٦٧٠٠ يشتغلون بالأعمال المالية و ٧٥٥٠٠ بالمتاجرة . ومن الأجانب ٢٠٠٠ يشتغلون بالأعمال المالية و ٩٥٠٠ بأعمال المتاجرة . فهل هناك علاقة بين جنسية الشخص ونوع العمل الذى يقوم به ، وما مقياس هذه العلاقة ؟

هنا نستخدم « معامل الاقتران » الذى وضعه ج . أ . يول[١] لقياس العلاقة بين الصفات التى لا تقاس . ولذلك نرتب الأرقام المعطاة فى « جدول الاقتران » كما يأتى :

جدول الاقتران بين الجنسية ونوع العمل

		الجنسية	
		مصريون	أجانب
نوع العمل	مالي	٦٧٠٠	٢٠٠٠
	متاجرة	٧٥٥٠٠	٩٥٠٠

(١) انظر كتاب :
G. U. Yule "*Introduction to the Theory of Statistics*" (1937)
وانظر أيضاً مقالته فى مجلة
Phil. Trans. Roy Soc., Series A, vol. 194 (1900), p. 257

$$\text{معامل الاقتران} = \frac{\text{٧٥٥٠٠} \times \text{٢٠٠٠} - \text{٩٥٠٠} \times \text{٦٧٠٠}}{\text{٧٥٥٠٠} \times \text{٢٠٠٠} + \text{٩٥٠٠} \times \text{٦٧٠٠}}$$

= ٤٠٧ر

ولا ضرورة هنا للاشارة الجبرية ؛ لأن المعامل إذا دل على اقتران بين الجنسية المصرية وأعمال المتاجرة ، فهو يدل فى الوقت نفسه على « تنافر » بين الجنسية المصرية والأعمال المالية . والمفهوم من هذا المعامل أن هناك علاقة أو ارتباطاً بين صفة الجنسية المصرية وصفة العمل التجارى ، ومثلها بين صفة الجنسية الأجنبية وصفة العمل المالى .

وعلى العموم إذا كان توزيع الأعداد فى جدول الاقتران هو :

ا	ب
حـ	د

يكون معامل الاقتران

$$\text{ن} = \frac{\text{ا د} - \text{ب حـ}}{\text{ا د} + \text{ب حـ}} \text{ ، . . . (١) ؛}$$

مع صرف النظر عن الإشارة الجبرية . وبديهىّ أن هذا المعامل دائما أقل من ١ .

معامل التوافق

٣٢٩ — إذا كانت إحدى الظاهرتين اللتين نبحث العلاقة بينهما ، أو كلتيهما ، تنقسم إلى أكثر من نوعين ، فان معامل الاقتران لا يساعدنا فى هذه الحالة . وعندئذ نستخدم « معامل التوافق » الذى وضعه كارل بيرسون[١] لقياس العلاقة بين الصفات غير المقيسة ، أو بين صفات بعضها يقاس بالأرقام وبعضها لا يقاس .

(١) اسمه بالانجليزية Karl Pearson's "Coefficient of Contingency" ؛ انظر كتاب Whittaker and Robinson "*Calculus of Observations*," (1929), p. 338.
وهناك يرمز لمعامل التوافق بالحرف الأغريقى Φ ، الذى نستبدل به الحرف العربى قَ .

ولتعريف هذا المعامل نأخذ المثال الآتى :

فى سنة ١٩٣٤ تقدم إلى امتحان شهادة الدراسة الثانوية (قسم أول) طلبة من ١٤٤ مدرسة . وكانت هذه المدارس مقسمة (من حيث الادارة الفنية) إلى ثلاث أنواع : ا (مدارس أميرية)، وب (مدارس خاضعة لتفتيش وزارة المعارف)، و حـ (مدارس غير خاضعة لهذا التفتيش) . ثم قسمت هذه المدارس حسب نسبة النجاح إلى خمس رتب : ا ، ب ، حـ ، ى ، هـ ، فيها نسب النجاح كما يأتى :

ا : ٨٠ وأقل من ١٠٠ ٪   ى : ٢٠ وأقل من ٤٠ ٪

ب : ٦٠ « ٨٠ ٪   هـ : ٠ « ٢٠ ٪

حـ : ٤٠ « ٦٠ ٪

وكان توزيع المدارس ال ١٤٤ على هذه المجموعات المزدوجة كما هو مبين فى « جدول التوافق » الآتى :

جدول ٤٢ — توزيع ١٤٤ مدرسة حسب النوع والرتبة

		رتب المدارس					المجموع
		ا	ب	حـ	ى	هـ	
نوع المدارس	ا	١	١٨	٩	١	٠	٢٩
	ب	٢	٠	٢	٢٧	١٦	٤٧
	حـ	٢	٢	٨	٢٢	٣٤	٦٨
المجموع		٥	٢٠	١٩	٥٠	٥٠	١٤٤

حيث العدد الموجود فى كل خانة من هذا الجدول يدل على عدد المدارس (أو التكرار) التى تجتمع فيها الصفتان المبينتان .

٢٣٠ — ولحساب معامل التوافق ، نربع تكرار كل خانة ونقسمه على حاصل ضرب المجموعين الرأسى والأفقى للعمود والصف الملتقيين فى هذه الخانة ، كما هو مبين فى الجدول الآتى :   حساب معامل التوافق

جدول ٤٣ — حساب معامل التوافق بين نوع المدرسة ورتبتها

		رتب المدارس					المجموع
		ا	ب	حـ	ى	هـ	
نوع المدارس	ا	١ / (٥ × ٢٩)	٣٢٤ / (٢٠ × ٢٩)	٨١ / (١٩ × ٢٩)	١ / (٥٠ × ٢٩)	٠	ر٧١٣١
	ب	٤ / (٥ × ٤٧)	٠	٤ / (١٩ × ٤٧)	٧٢٩ / (٥٠ × ٤٧)	٢٥٦ / (٥٠ × ٤٧)	ر٤٤٠٦
	حـ	٤ / (٥ × ٦٨)	٤ / (٢٠ × ٦٨)	٦٤ / (١٩ × ٦٨)	٤٨٤ / (٥٠ × ٦٨)	١١٥٦ / (٥٠ × ٦٨)	ر٥٤٦٥
المجموع		ر٠٣٥٦	ر٥٦١٥	ر٢٠١٠	ر٤٥٣٢	ر٤٤٨٩	١ر٧٠٠٢

فلو رمزنا للمجموع الكلى ١ر٧٠٠٢ بالحرف حـ ، يكون معامل التوافق الذى يسميه بيرسون جذر متوسط مربع التوافق[(١)]

ت = √(حـ − ١) . . . . (١)

---

(١) بالأنجليزية (Root-Mean Square Contingency)

ولـكن أغلب الإحصائيين[1] يستخدمون معادلة أخرى لمعامل التوافق وهى :

$$\text{ن} = \sqrt{\frac{\text{هـ} - ١}{\text{هـ}}} \quad \ldots \quad (٢)$$

كل من هذين المعاملين ن و نَ موجب[2] ؛ ولـكنهما غير متساويين ، لأن :

$$\text{ن}^{٢} = \frac{[illegible]}{\text{نَ}^{٢} + ١} \quad \ldots \quad (٣)$$

و

$$\text{نَ}^{٢} = \frac{\text{ن}^{٢}}{١ - \text{ن}^{٢}} \quad \ldots \quad (٤)$$

وكل منهما يمكن استخدامه كقياس للعلاقة بين الصفات المبينة فى الجدول .

وبناء على ذلك يكون معامل التوافق بين نوع المدارس ورتبتها فى المسألة التى نحن بصددها هو :

$$\text{نَ} = \sqrt{٧٠٠٢ر} \text{ ،}$$

$$= ٨٤ر \text{ ،}$$

فى حين أن $\text{ن} = \sqrt{\frac{٧٠٠٢ر}{١ر٧٠٠٢}}$ ،

$$= ٦٤ر \text{ ،}$$

وهذا معناه أن هناك علاقة بين نوع المدرسة ونسبة النجاح بين طلبتها ، وإلا كان كل من المعاملين صفراً ، أو ما يقرب من الصفر .

(١) واضح أن المعامل ن يكون دائماً أقل من الواحد . وهو أكثر استعمالاً من المعامل نَ .

(٢) انظر كتاب G. Thompson *"How To Measure Correlation."* (1924), p. 20 حيث يستعمل الحرف C بدل ن . أو كتاب Sargant Florence *"Statistical Methods"*, (1929). p. 505.) ويلاحظ أن فلورنس يذكر المعادله (٢) فى شكل مخالف . ولـكن يمكن تحويرها بسهولة حتى تأخذ الشكل المذكور أعلاه .

المراجـــع

BOWLEY, A.L. *Elements of Statistics*, Chapter VI Part II.

CONNOR, L.R., *Statistics in Theory and Practice*, Chapter XV.

FLORENCE, S.P., *Statistical Methods*, Chapter IX.

JONES, C., *First Course in Statistics*, Chapter X.

KING, W. *Statistical Method*, Chapter XVII.

MILLS, F.C., *Statistical Methods*, Chapter X.

SECRIST, H., *Statistical Methods*, Chapter XIII.

THOMPSON, G. *How to Measure Correlation*.

WHITTAKER AND ROBINSON, *Calculus of Observations*, Chapter XII.

الباب السابع

الارتباط

خطوط الانحدار المستقيمة والمنحنية

٢٣١ — تكلمنا فى الباب السابق عن الارتباط ومعناه وكيفية قياسه فى الحالات المختلفة . ومقاييس الارتباط التى نحصل عليها بهذه الطرق تعرفنا درجة العلاقة بين الكميتين أو الظاهرتين المتغيرتين . ولكن هذه المعرفة لا تفيدنا كثيراً فى دراسة هاتين الظاهرتين إذا اقتصرت فائدتها على إثبات وجود هذه العلاقة أو عدم وجودها ، وقياسها إن وجدت . لأن المفهوم عادة من وجود علاقة أو ارتباط بين متغيرين أننا إذا علمنا قيمة أحدهما فى حالة ما ، أمكننا — بناء على وجود هذه العلاقة أو الارتباط — أن نقدر ، ولو بالتقريب ، قيمة المتغير الثانى ؛ وأن تقديرنا هذا يكون أقرب إلى الحقيقة كلما كان الارتباط شديداً . والآن نبحث فى هذه الناحية من موضوع الارتباط .

ازدياد عدد الأطفال مع العمر

٢٣٢ — رأينا فى المثال العملى الذى أخذناه فى بند ٢١٥ (جدول ٣٢ صفحة ٢٣٥) أن هناك علاقة طردية بين عمر الرجل وعدد أطفاله . أى أن عدد الأطفال يزيد ، على وجه العموم ، بزيادة العمر . ومعنى ذلك أن رجلا عمره ٣٠ سنة مثلا يكون عدد أطفاله فى العادة أكبر من عدد أطفال رجل عمره ٢٥ سنة فقط ؛ وأن رجلا عمره ٣٥ سنة يكون عدد أطفاله فى ، العادة ، أكبر من عدد أطفال رجل عمره ٣٠ سنة فقط ، وهكذا . على أن هذا لا يمنع طبعاً من أن رجلا

معيناً عمره ٢٥ سنة مثلا عدد أطفاله ، لسبب ما ، أكبر من عدد أطفال رجل آخر عمره ٣٠ سنة . ولكن هذا إن تحقق فلن يكون إلا نادرا .

إذن يكون متوسط عدد الأطفال عند مجموعة الرجال الذين سنهم بين ٢٥ و٣٠ سنة مثلا ، أكبر من متوسط عدد الأطفال عند مجموعة الرجال الذين تنحصر أعمارهم بين ٢٠ و٢٥ سنة فقط . وهكذا فى فئات الأعمار المتتالية ، التى نراها فى الجدول رقم ٣٢ . ويمكننا[1] بسهولة حساب متوسط عدد الأطفال لكل مجموعة من هذه المجموعات الخمسة وهى .

مجموعة الرجال الذين أعمارهم تساوى ٢٢ر٥ سنة ، متوسط عدد أطفالهم = ر٧٣

» » » » » ٢٧ر٥ » » » » = ١ر٣٨

» » » » » ٣٢ر٥ » » » » = ٢ر٠٧

» » » » » ٣٧ر٥ » » » » = ٢ر٦٢

» » » » » ٤٢ر٥ » » » » = ٣ر١٢

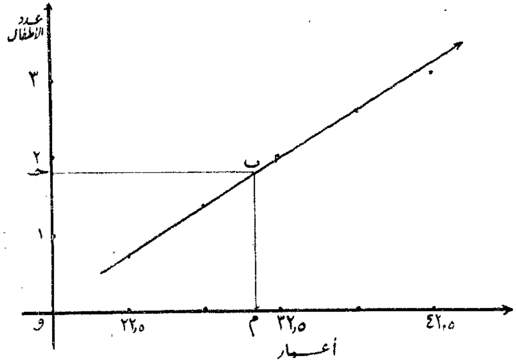
ونرى هنا بوضوح كيف يزداد متوسط عدد الأطفال بالتدريج مع زيادة عمر الرجل .

٢٣٣ — لنرسم محورين متعامدين ، ونقيس على المحور الأفقى الأعمار مثلا ، رسم
ونقيس عدد الأطفال على المحور الرأسى . ثم نرصد النقط التى إحداثياتها الأفقية هى

---

(١) كل عمود فى جدول ٣٢ عبارة عن توزيع تكرارى لعدد الأطفال المبينة فئاته فى العمود الأيمن من الجدول . ففى كل خانة فى أى عمود نضرب تكرارها فى عدد الأطفال المبين أمام هذه الخانة ، ونجمع هذه الحواصل ونقسم على المجموع الكلى للعمود ، فنحصل على متوسط عدد الأطفال لهذا العمود : مثلا العمود الثالث نجد المتوسط = (٨×٠ + ٢٥×١ + ١٣×٢ + ٦×٣ + ١×٤) ÷ ٥٣ = ١ر٣٨ طفلا .

٢٢٫٥ ، ٢٧٫٥ ، ٣٢٫٥ ، ٣٧٫٥ ، ٤٢٫٥ ؛ وإحداثياتها الرأسية هى على الترتيب ٠٫٧٣ ، ١٫٣٨ ، ٢٫٠٧ ، ٢٫٦٢ ، ٣٫١٢ .



(شكل ٦١)
متوسط عدد الأطفال والعمر

ثم نصل بين هذه النقط بخط ممهد على قدر الامكان ؛ ويظهر من شكل ٦١ أن الخط البيانى فى هذه الحالة مستقيم تقريباً .

هذا الخط إذن يبين العلاقة بين عمر الرجل ومتوسط ما عنده من الأطفال . وهو وإن كان مرسوماً من واقع خمس نقط معينة فقط ، ألا وهى النقط المقابلة للأعمار ٢٢٫٥ ، ٢٧٫٥ ، . . . ، ٤٢٫٥ ، فهو يمكننا من معرفة ماذا يكون متوسط عدد الأطفال عند ما يكون عمر الرجل ٣١ أو ٣٩ سنة مثلا ، أو أى عمر آخر . وذلك بأن نحدد على المحور الأفقي بعداً ، مثل و م يمثل ٣١ سنة (أو ٣٩ أو أي عمر نريده) ثم نقيم من م عموداً على المحور الأفقى فيقابل الخط البيانى فى نقطة مثل ب . والإحداثى الرأسى لهذه النقطة ، وهو م ب = و حـ فى الشكل ، هو متوسط عدد الأطفال عندما يكون العمر يساوى ٣١ سنة (أو ٣٩ أو الخ) .

خط انحدار عدد الأطفال على العمر

٢٣٤ — هذا الخط البيانى نسميه [1] . **خط انحدار عدد الأطفال على العمر** . وهو كما رأينا يصور العلاقة بين العمر ومتوسط عدد الأطفال ، وبواسطته يمكن تقدير عدد الأطفال عند أى رجل إذا علم عمر هذا الرجل . وهذا التقدير يعطى متوسط ما عند أمثال هذا الرجل من الأطفال ؛ وهو تقدير يقرب من الحقيقة كلما كان الارتباط شديداً بين العمر وعدد الأطفال .

فائدة خط الانحدار فى دراسة الارتباط

٢٣٥ — خط الانحدار إذن يحقق لنا الفائدة التى نرجوها من دراسة العلاقة بين كميتين متغيرتين ؛ فهو يصور لنا هذه العلاقة فى شكل هندسى منظور ، نرى فيه كيف تميل إحدى الكميتين إلى متابعة الأخرى فى تغيرها — إما طردياً وإما عكسياً . وحبذا لو أمكننا أيضاً تصوير هذه العلاقة أو الارتباط فى صورة جبرية أو تحليلية ، إذ أن هذا يكون بلاشك أوفى وأتم ، من الوجهتين النظرية والعملية .

وهذا الأمر ميسور لنا ، حيث قد علمنا (فى الباب الرابع) أن الأشكال البيانية يمكن الدلالة عليها بمعادلات جبرية . فلنبحث إذن فى استنباط معادلة خط الانحدار ، مستقيما كان أو منحنياً ؛ وإذا حصلنا عليها فقد حصلنا على الصورة الجبرية أو التحليلية المطلوبة للعلاقة بين المتغيرين تحت البحث .

رسم شكل الانتشار

٢٣٦ — المقصود من خط الانحدار هو كما قلنا تصوير العلاقة بين المتغيرين فى صورة جبرية تحليلية . لنفرض أن قيم المتغيرين هى كالمعتاد :

---

(١) يسمى بالانجليزية (Regression Line of the Number of Children on the age of the father) وأول من وضع هذه التسمية هو فرانسيس جولتون (Francis Galton) وترجمتها بالعربية « ارتداد » . وسيظهر فيما يلى السبب الذى جعله يستخدم هذا المعنى ، والسبب الذى جعلنا نفضل كلمة « انحدار » . على أن المعنى الذى تؤديه كلمة ارتداد أو (Regression) يقتصر على ناحية واحدة فقط للفكرة التى يصورها خط الانحدار .

$س_1 ، س_2 ، س_3 ، . . . . . ، س_ن$

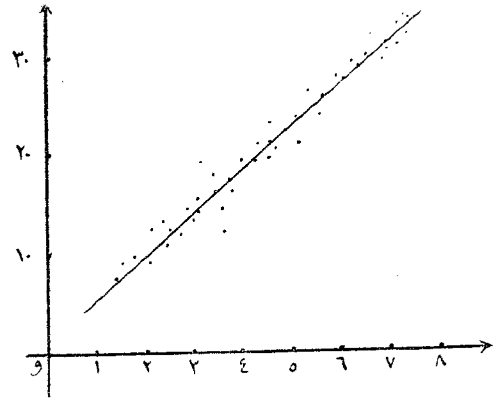
و $ص_1 ، ص_2 ، ص_3 ، . . . . . ، ص_ن$ ؛

وأن الوسطين الحسابيين هما $\bar{س}$ و $\bar{ص}$ ، والانحرافين المعياريين هما ع و عَ .

$\therefore$ ن $\bar{س}$ = مج س ، ن $\bar{ص}$ = مج ص ؛

و ن $ع^2$ = مج$(س - \bar{س})^2$ ، ن $عَ^2$ = مج$(ص - \bar{ص})^2$.

نرسم محورين متعامدين ونرصد في الشكل النقط التي إحداثياتها الأفقية تساوي



(شكل ٦٢)

شكل انتشار نقط على خط مستقيم

قيم س ، وإحداثياتها الرأسية تساوي قيم ص المناظرة لها، أي النقط التي إحداثياتها هي:

$(س_1 ، ص_1) ، (س_2 ، ص_2) ، (س_3 ، ص_3) ، . . . . ، (س_ن ، ص_ن)$.

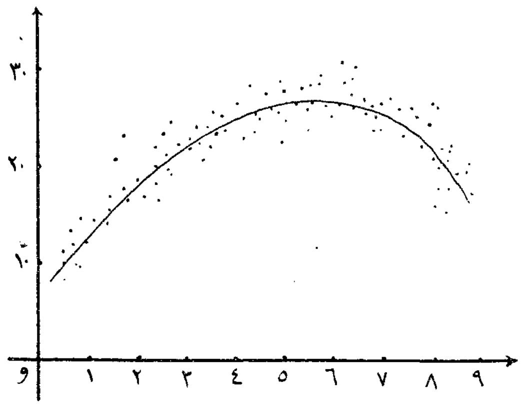
يتكون من هذه النقط البيانية **شكل الانتشار**[1] كالذي نراه في (شكل ٦٢) حيث نجد النقط منتشرة في الشكل ومبعثرة نوعاً . وبطبيعة الحال

(١) يسمى بالانجليزية : (Scatter Diagram)

إذا كانت هناك علاقة تربط المتغيرين س و ص ، فسيكون أثرها أن تنتشر هذه النقط بشكل منتظم يسير على قاعدة معينة . أما إذا كانت النقط في هذا الشكل مبعثرة حيثما اتفق وبدون أي نظام ملحوظ ، فهذا دليل على أن العلاقة بين المتغيرين معدومة أو ضعيفة جداً .

خط الانتشار إما مستقيم أو غير مستقيم

٢٣٧ — في بعض الأحيان نجد أن النقط في شكل الانتشار تنتظم في خط مستقيم ، أو ما يقرب من خط مستقيم كما نرى في شكل ٦٢ ؛ وأحياناً تنتشر هذه النقط على خط غير مستقيم ، ولكنه خط منتظم . وهو إما منحن ذو نهاية واحدة



(شكل ٦٣)

شكل انتشار نقط على خط منحن من الدرجة الثانية

(عظمى أو صغرى) ، كما نرى في (شكل ٦٣) ؛ وتكون إذ ذاك معادلته من الدرجة الثانية كما رأينا في بند ٧٨ من الباب الرابع . وإما أن يكون منحنياً ذا نهايتين أو أكثر ، فتكون معادلته من الدرجة الثالثة أو من درجة أعلى .

تشتت النقط حول خط الانتشار

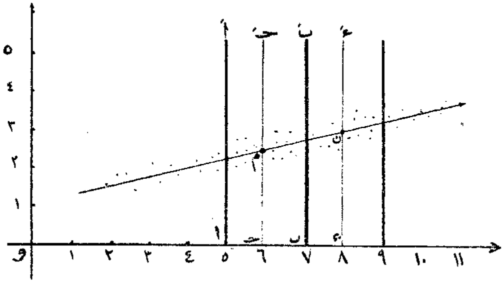
٢٣٨ — ولا يلزم أن تكون جميع النقط البيانية فى شكل الانتشار واقعة على خط الانتشار أو ملاصقة له ، إذ يصح أن تشذ نقطة معينة ، لسبب ما ، وتقع بعيدة عن هذا الخط . على أنه إذا كان عدد كبير من النقط واقعاً على هذا الخط أو ملاصقاً له ، وكانت النقط الباقية على مقربة منه ، كان ذلك دليلا واضحاً على شدة الارتباط بين المتغيرين . وبالعكس إذا كان كثيراً ما تشذ النقط وتبعد عن هذا الخط ، ولم يكن هناك خط آخر أقرب إلى النقط من هذا ، دل ذلك على ضعف الارتباط بين المتغيرين .

ويمكننا التعبير عن هذا باختصار ، فنقول إن « تشتت » النقط حول خط الانتشار يكون صغيراً إذا كان الارتباط بين المتغيرين شديداً ، ويكون كبيراً إذا كان الارتباط ضعيفاً . وربما أمكننا استخدام هذه الفكرة كأساس لابتكار مقياس للارتباط . وسنرى فيما بعد كيف نعبر عن معامل الإرتباط العادى بدلالة التشتت حول هذا الخط ، الذى يصح أن نسميه مبدئياً (١) **خط العلاقة المتوسطة** بين المتغيرين .

خط الانتشار هو نفس خط الانحدار

٢٣٩ — لنأخذ قيمتين من قيم س ، مثلاً ٥ و ٧ ؛ ونرسم فى شكل الانتشار خطين رأسيين يقابلان المحور الأفقى عند ٥ وعند ٧ مثلا أ أَ و ب بَ فى شكل ٦٤ . جميع النقط البيانية الواقعة بين هذين الخطين لها إحداثيات أفقية بين ٥ و ٧ . أى أنها تكون فئة من فئات س ، حدها الأدنى ٥ والأعلى ٧ ، ومركزها ٦ طبعاً . ويمكن اعتبار أن الإحداثيات الأفقية لجميع هذه النقط تساوى ٦ تقريبا ، كالمعتاد فى التوزيعات التكرارية . نرسم خطاً رأسياً يمر بمركز هذه الفئة مثل حـ حـَ ، يقابل خط الانتشار أو خط العلاقة المتوسطة فى نقطة م

(١) بالانجليزية (Line of Average Relationship.)



(شكل ٦٤)

خط الانتشار هو نفس خط الانحدار

مثلا . فمن الواضح أن الإحداثى الرأسى لهذه النقطة م ، وهو = حـ م ، يساوى متوسط الإحداثيات الرأسية لجميع النقط البيانية الواقعة بين الخطين أ أَ و ب بَ .

لأن خط الانتشار يمر دائماً متوسطاً بين النقط التى حوله — وهذه هي وظيفته طبعاً — ولأن الإحداثيات الأفقية لجميع النقط الواقعة بين هذين الخطين ، أصبحت تساوى ٦ ، وبذلك تركزت جميع هذه النقط على الخط الرأس حـ حـَ .

وكذلك إذا أخذنا الفترة ٧ — ٩ لقيم س ، ورسمنا الخط ء ءَ ليقابل خط الانتشار عند ن ، يكون ء ن مساوياً لمتوسط الإحداثيات الرأسية للنقط الواقعة فى الفترة ٧ — ٩ .

وهكذا فكل نقطة من النقط التى يتألف منها خط الانتشار ، يمثل إحداثيها الأفقى قيمة سينية ، ويمثل إحداثيها الرأسى متوسطاً صاديا مناظراً لهذه القيمة السينية ؛ وعلى ذلك يكون هذا الخط هو خط انحدار ص على س ، حسب خواص خط الانحدار المذكورة فى بندى ٢٣٤ و ٢٣٥ .

٢٤٠ — معادلة خط الانحدار هى إذن معادلة خط الانتشار ، أو خط العلاقة المتوسطة الذى يسير متوسطاً بين النقط البيانية :

$$(\text{س}_1 ، \text{ص}_1) ، (\text{س}_2 ، \text{ص}_2) ، \ldots ، (\text{س}_\text{ن} ، \text{ص}_\text{ن})$$

المرسومة فى شكل الانتشار . فلنبحث الآن فى كيفية استنباط هذه المعادلة . هنا ينقسم البحث إلى حالتين :

الحالة الأولى : وفيها يكون خط الانحدار مستقيما ، ذا معادلة من الدرجة الأولى بالنسبة إلى س و ص . ويسمى الارتباط حينئذ « الارتباط ذا الانحدار المستقيم » أو « الارتباط المستقيم » .

الحالة الثانية : وفيها يكون خط الانحدار غير مستقيم ، وتكون معادلته من الدرجة الثانية ، أو من درجة أعلى . ويسمى الارتباط حينئذ « الارتباط غير المستقيم »[1] .

٢٤١ — فى الحالة الأولى ، حالة الارتباط المستقيم[2] ، تكون معادلة خط انحدار ص على س هى معادلة أفضل خط مستقيم يوافق النقط البيانية المعطاة وهى :

---

(١) الارتباط المستقيم يسمى (Linear Correlation) أو (Normal Correlation) وغير المستقيم (Non-Linear Correlation)

(٢) لإثبات ذلك نفرض أن هذه المعادلة (من الدرجة الأولى) هى :

$$\text{ص} = \text{م}\,\text{س} + \text{هـ} \quad . \quad . \qquad (1)$$

ونوجد قيمتى م و هـ من المعادلتين الآتيتين :

$$\text{مج ص} = \text{م} . \text{مج س} + \text{ن}\,\text{هـ} \qquad (2)$$

$$\text{و} \quad \text{مج س ص} = \text{م} . \text{مج س}^2 + \text{هـ} . \text{مج س} \qquad (3)$$

ولكن $\text{مج س} = \text{ن}\bar{\text{س}}$ و $\text{مج ص} = \text{ن}\bar{\text{ص}}$ و $\text{مج س}^2 = \text{ن ع}^2 + \text{ن}\bar{\text{س}}^2$ ، بند ١٧٧ (*)

$$(\text{س}_1 ، \text{ص}_1) ، (\text{س}_2 ، \text{ص}_2) ، \ldots ، (\text{س}_\text{ن} ، \text{ص}_\text{ن}) .$$

ولإيجاد هذه المعادلة ، نوفق خطاً مستقيما لهذه النقط بنفس الطريقة التى شرحناها فى بند ٧٥ (صفحة ٧٥) . ونجد بسهولة أن معادلة خط انحدار ص على س هى :

$$\text{ص} - \bar{\text{ص}} = \text{ر}\,\frac{\text{عَ}}{\text{ع}}\,(\text{س} - \bar{\text{س}}) \quad . \quad . \qquad (1) ،$$

حيث $\bar{\text{س}}$ = الوسط الحسابى لقيم س المعطاة فى المسألة ،

و $\bar{\text{ص}}$ = « « « ص « « ،

و ر = معامل الارتباط بين المتغيرين س و ص ،

و ع = الانحراف المعيارى السينى ،

و عَ = « « الصادى .

والكمية $\text{ر}\,\frac{\text{عَ}}{\text{ع}}$ تسمى[1] ***معامل الانحدار للمتغير ص على المتغير س***

وهذا الخط كما قلنا فى بند ٢٣٩ يعبر عن العلاقة بين قيم المتغير س ومتوسطات القيم الصادية التى تقترن بكل منها . أى أن الإحداثى الأفقى لأى نقطة

(*) فينتج من (٢) أن $\bar{\text{ص}} = \text{م}\,\bar{\text{س}} + \text{هـ}$ أى أن $\text{هـ} = \bar{\text{ص}} - \text{م}\,\bar{\text{س}}$

وينتج من (٣) أن $\text{مج س ص} = \text{م ن}(\text{ع}^2 + \bar{\text{س}}^2) + \text{هـ} . \text{ن}\bar{\text{س}}$

$$= \text{ن م ع}^2 + \text{ن م} . \bar{\text{س}}^2 + \text{ن}\,\bar{\text{س}}\,\bar{\text{ص}} - \text{ن م}\,\bar{\text{س}}^2 \quad \ldots \quad (4)$$

ولكن $\text{ن} . \text{ر} . \text{ع عَ} = \text{مج س ص} - \text{ن}\,\bar{\text{س}}\,\bar{\text{ص}}$ معادلة (١) بند ٢٠٥

$\therefore \qquad = \text{ن م ع}^2$

$\therefore \quad \text{م} = \text{ر}\,\frac{\text{عَ}}{\text{ع}}$ وبالثانى $\text{هـ} = \bar{\text{ص}} - \text{ر}\,\frac{\text{عَ}}{\text{ع}}\,\bar{\text{س}}$.

وبالتعويض عن م و هـ فى المعادلة المفروضة ص = م س + هـ تنتج المعادلة المطلوبة .

(١) بالانجليزية (Coefficient of Regression of Y upon X) وتسمى أحياناً نسبة الانحدار أى (Regression Ratio of Y on X)

على هذا المستقيم ، يمثل قيمة معينة للمتغير س ، والإحداثى الرأسى يمثل متوسط قيم التغير ص التى تقترن بهذه القيمة السينية الخاصة

ففي مسألة الارتباط بين أعمار الرجال وأعداد أطفالهم (جدول ٣٢) لو أخذنا س تمثل العمر و ص تمثل عدد الأطفال ، نجد أن معادلة خط انحدار عدد الأطفال على العمر هى :

ص – ٢٫٠٩ = ٫٤٨٤ × (١٫٣٥٣٥ / ٥٫٥٦٠٥) (س – ٣٣٫١٧٥) ،

أى ص = ٫١١٧٨ س – ١٫٨١٨٠ . . . . (٢) ،

حيث س = عمر الرجل

و ص = متوسط عدد الأطفال عند الرجال الذين عمرهم يساوى س .

فلو أخذنا رجلا عمره ٣١ سنة مثلا ، نجد أن عدد أطفاله فى المتوسط يكون

ص = ٫١١٧٨ × ٣١ – ١٫٨١٨٠ ،

= ١٫٨٣٣٨ طفلا .

ويمكن الوصول إلى نفس النتيجة بالرسم ؛ حيث نرسم المستقيم الذى معادلته ص = ٫١١٧٨ س – ١٫٨١٨ ونأخذ على المحور الأفقى البعد و ا = ٣١ سنة ونرسم ا ب رأسيا ليقابل المستقيم فى نقطة ب يكون إحداثيها الرأسى ا ب = و حـ = ١٫٨٣٣٨ طفلا (شكل ٦١ صفحة ٢٦٤)

خط انحدار س على ص

٢٤٢ — أما معادلة خط انحدار س على ص فنحصل عليها بنفس الطريقة بوضع س بدل ص ووضع ص بدل س فى كل خطوة من البرهان السابق . وهذه المعادلة هى :

$$\text{س} - \bar{\text{س}} = \text{ر}\,\frac{\text{ع}}{\text{ع}}\,(\text{ص} - \bar{\text{ص}}) \quad . . . \quad (١) .$$

هنا ص تمثل أى قيمة معينة للمتغير ص ، فى حين أن س تمثل متوسط القيم التى تقترن بهذه القيمة الصادية ؛ والكمية ر ع/ع هى معامل انحدار س على ص ، السينية أو نسبة انحدار س على ص .

وفى مسألة الأعمار وعدد الأطفال ، نجد معادلة خط انحدار الأعمار على عدد الأطفال هى :

س – ٣٣٫١٧٥ = ٫٤٨٤ × (٥٫٥٦٠٥ / ١٫٣٥٣٥) (ص – ٢٫٠٩)

أى س = ١٫٩٨٨٤ ص + ٢٩٫٠١٩٢ . . (١) .

فلو أردنا مثلا معرفة ماذا يكون متوسط عمر الرجل الذى عنده ٣ من الأطفال ، نعوض فى هذه المعادلة ص = ٣ ،

∴ س = ١٫٩٨٨٤ × ٣ + ٢٩٫٠١٩٢

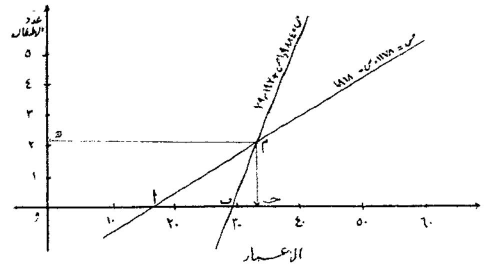
= ٣٤٫٩٨٤٤ سنة .

وهذا معناه أن الرجال الذين عندهم ٣ أطفال يكون متوسط أعمارهم ٣٥ سنة تقريباً .

خواص خطوط الانحدار

٢٤٣ — إذا كان الارتباط بين المتغيرين طرديا ، نجد أن خطى الانحدار يميلان على المحور الأفقى ميلا موجباً ؛ أى أن كلا منهما يصنع زاوية حادة مع محور س ، ظلها يساوى كمية موجبة . أما إذا كان الارتباط عكسياً (ر سالبة) ، فكل منهما يصنع زاوية منفرجة مع محور س ، وظلها يساوى كمية سالبة . وهذا واضح من صورة معادلتى خط الانحدار حيث تظهر الكمية ر مضروبة فى س (أو فى ص) ، حيث تكون إشارتها الجبرية هى إشارة الميل أو ظل الزاوية التى يصنعها المستقيم مع المحور الأفقى .

وإذا رسمنا خطى الانحدار فى شكل واحد من واقع معادلتيها ، نجد أنهما يتقابلان فى النقطة التى إحداثيها الأفقى يساوى الوسط الحسابى السينى ، وإحداثيها الرأسى يساوى الوسط الحسابى الصادى . والسبب فى ذلك واضح ، إذ نرى أن النقطة $(\overline{\text{س}} ، \overline{\text{ص}})$ تحقق معادلتى خطى الانحدار فى وقت واحد . فى مسألة الأعمار وعدد الأطفال نجد الخطين يتقابلان فى النقطة م (٣١٫١٧٥ ، ٢٫٠٩) كما نرى فى شكل ٥٥ .



(شكل ٦٥)
خطا انحدار الأعمار وعدد الأطفال

وهذان الإحداثيان هما كما نعلم الوسطان الحسابيان للأعمار وأعداد الأطفال .

٢٤٤ – بالنظر إلى معادلتى خط الانحدار وهما : حاصل ضرب معاملى الانحدار يساوى ر^2

$$\text{ص} = \text{ر}\frac{\text{ع}_\text{ص}}{\text{ع}_\text{س}}\text{س} + \left(\overline{\text{ص}} - \text{ر}\frac{\text{ع}_\text{ص}}{\text{ع}_\text{س}}\overline{\text{س}}\right) \text{، (وهو خط انحدار ص على س)}$$

$$\text{س} = \text{ر}\frac{\text{ع}_\text{س}}{\text{ع}_\text{ص}}\text{ص} + \left(\overline{\text{س}} - \text{ر}\frac{\text{ع}_\text{س}}{\text{ع}_\text{ص}}\overline{\text{ص}}\right) \text{، (« « « س على ص)}$$

نجد أن حاصل ضرب معامل س فى الأولى ، ومعامل ص فى الثانية (أى معاملى الانحدار) يساوى ر^2 ، وهو مربع معامل الارتباط .

∴ فى مسألة الاعمار والأطفال ، حيث المعادلتان هما

$$\text{ص} = ٠٫١١٧٨ \text{ س} - ١٫٨١٨٠ \text{ ،}$$

$$\text{و} \quad \text{س} = ١٫٩٨٨٤ \text{ ص} + ٢٧٫٠١٩٢ \text{ ،}$$

$$\text{يكون} \quad \text{ر}^2 = ٠٫١١٧٨ \times ١٫٩٨٨٤ = ٠٫٢٣٤٢ \text{ تقريباً}$$

$$\text{ومن ثم ر} = ٠٫٤٨٤$$

وفى شكل ٦٥ نرى أن حاصل ضرب هذين المعاملين هو

$$\text{ر}^2 = \frac{\text{حـ م}}{\text{ا حـ}} \times \frac{\text{ب حـ}}{\text{حـ م}} \text{ ،}$$

$$= \frac{\text{ب حـ}}{\text{ا حـ}} \text{ ،}$$

ومن هذه النسبة بين البعدين ب حـ و ا حـ نحسب قيمة ر .

٢٤٥ – لنبحث[1] الآن فى تشتت النقط المعطاة تشتت النقط حول المستقيم

$$(\text{س}_١ ، \text{ص}_١) ، (\text{س}_٢ ، \text{ص}_٢) ، \ldots ، (\text{س}_\text{ن} ، \text{ص}_\text{ن}) \text{ ،}$$

حول خط الانتشار أو خط الانحدار المستقيم الذى معادلته :

$$\text{ص} = \text{م س} + \text{حـ} \text{ ،}$$

تلك الفكرة التى نوهنا عنها فى آخر بند ٢٣٨ ؛ ولنبحث فيما إذا كان هناك أى علاقة بين مقدار هذا التشتت ومعامل الارتباط بين المتغيرين س و ص .

قلنا فى الباب السابع ، إن التشتت حول الوسط الحسابى يقاس بالجذر التربيعى لمتوسط مربعات الانحرافات عن هذا الوسط . وكذلك نقول هنا إن تشتت هذه النقط حول الخط ص = م س + حـ يقاس بالجذر التربيعى لمتوسط مربعات انحرافات النقط عن الخط .

عند توفيق المستقيم ص = م س + حـ ، قلنا (بند ٧٥ صفحة ٧٥) إن انحراف النقطة ا $(\text{س}_١ ، \text{ص}_١)$ مثلا عند هذا المستقيم ، وليكن $\text{ع}_١$ مثلا ، هو

(١) يحسن للقارئ المبتدئ أن يترك هذا البند وما يليه فى هذا الباب .

$$\text{ع}_1 = \text{م}\,\text{س}_1 + \text{ﺣ} - \text{ص}_1 \quad \ldots (١)\,؛$$

وكذلك النقطة ب (س_2 ، ص_2) يكون انحرافها عن المستقيم ،

وليكن ح_2 ، هو

$$\text{ع}_2 = \text{م}\,\text{س}_2 + \text{ﺣ} - \text{ص}_2 \qquad (٢)\,؛$$

وهكذا مع جميع النقط .

∴ مجموع هذه الانحرافات

$$\text{مج}\,\text{ع} = \text{مج}\,(\text{م}\,\text{س} + \text{ﺣ} - \text{ص})\,،$$

$$= \text{صفراً} \qquad .(٣)$$

لأن $\text{مج}\,\text{ص} = \text{م}\,\text{مج}\,\text{س} + \text{ن}\,\text{ﺣ}$ من معادلات توفيق المستقيم

$$\text{ص} = \text{م}\,\text{س} + \text{ﺣ}$$

لو ضربنا طرف المعادلة (١) فى ع_1 والمعادلة (٢) فى ع_2 وهكذا

$$\therefore\ \text{ع}_1^2 = \text{م}\,.\,\text{ع}_1\text{س}_1 + \text{ﺣ}\,.\,\text{ع}_1 - \text{ع}_1\text{ص}_1\,،$$

$$\text{و}\quad \text{ع}_{\text{ن}}^2 = \text{م}\,.\,\text{ع}_2\text{س}_2 + \text{ﺣ}\,.\,\text{ع}_2 - \text{ع}_2\text{ص}_2\,،$$

$$\ldots \qquad \ldots \qquad \ldots \qquad .$$

$$\text{و}\quad \text{ع}_2^2 = \text{م}\ \text{ع}_{\text{ن}}\,\text{س}_{\text{ن}} + \text{ﺣ}\,.\,\text{ع}_{\text{ن}} - \text{ع}_{\text{ن}}\,\text{ص}_{\text{ن}}\ \text{؛ وبالجمع}$$

$$\therefore\ \text{مج}\,\text{ع}^2 = \text{م}\,.\,\text{مج}\,\text{ع}\,\text{س} + \text{ﺣ}\,.\,\text{مج}\,\text{ع} - \text{مج}\,\text{ع}\,\text{ص}\ ،\ldots (٤).$$

ولكن $\text{مج}\,\text{ع}\,\text{س}. = \text{مج}\,\text{س}\,(\text{م}\,\text{س} + \text{ﺣ} - \text{ص})$

$$= \text{مج}\,(\text{م}\,\text{س}^2 + \text{ﺣ}\,\text{س} - \text{س}\,\text{ص})$$

$$= \text{م}\,.\,\text{مج}\,\text{س}^2 + \text{ﺣ}\,.\,\text{مج}\,\text{س} - \text{مج}\,\text{س}\,\text{ص}$$

$$= \text{صفراً}\,.$$

لأن $\text{مج}\,\text{س}\,\text{ص} = \text{م}\,\text{مج}\,\text{س}^2 + \text{ﺣ}\,\text{مج}\,\text{س}$

وكذلك $\text{ﺣ}\,.\,\text{مج}\,\text{ع} = \text{صفراً}$.

∴ $\text{مج}\,\text{ع}^2 = -\,\text{مج}\,\text{ع}\,\text{ص}$ ، وبضرب طرفى معادلة (١) فى ص_1 وهكذا

$$= -\,\text{مج}\,(\text{م}\,.\,\text{س}\,\text{ص} + \text{ﺣ}\,\text{ص} - \text{ص}^2)$$

$$= \text{مج}\,\text{ص}^2 - \text{م}\,\text{مج}\,\text{س}\,\text{ص} - \text{ﺣ}\,.\,\text{مج}\,\text{ص} \quad \ldots (٥)$$

ولكن $\text{مج}\,\text{ص}^2 = \text{ن}\,\text{ع}'^2 + \text{ن}\,\overline{\text{ص}}^2$ ؛

$$\text{و}\ \ \text{ﺣ}\,.\,\text{مج}\,\text{ص} = \text{ﺣ}\,.\,\text{ن}\,\overline{\text{ص}} = (\overline{\text{ص}} - \text{م}\,\overline{\text{س}})\,.\,\text{ن}\,\overline{\text{ص}}\,،$$

$$= \text{ن}\,\overline{\text{ص}}^2 - \text{م}\,.\,\text{ن}\,\overline{\text{س}}\,\overline{\text{ص}}\,؛$$

$$\text{و}\qquad \text{م} = \text{ر}\,\frac{\text{ع}'}{\text{ع}}$$

فلو كتبنا $\text{ن}\,\text{سـ}'^2$ بدل $\text{مج}\,\text{ع}^2$ فى معادلة (٥) وعوضنا هذه القيم فيها

$$\therefore\ \text{ن}\,\text{سـ}'^2 = \text{ن}\,\text{ع}'^2 + \text{ن}\,\overline{\text{ص}}^2 - \text{م}\,(\text{مج}\,\text{س}\,\text{ص} - \text{ن}\,\overline{\text{س}}\,\overline{\text{ص}}) - \text{ن}\,\overline{\text{ص}}^2$$

$$= \text{ن}\,\text{ع}'^2 - \text{ر}\,.\,\frac{\text{ع}'}{\text{ع}}\,(\text{ن}\,.\,\text{ر}\,.\,\text{ع}\,.\,\text{ع}')\ ،$$ [انظر معادلة (١)

(بند ٢٠٥)

$$\therefore\ \text{سـ}'^2 = \text{ع}'^2\,(1 - \text{ر}^2) \quad \ldots\ldots \qquad (٦)$$

$$\therefore\ \text{ر} = \sqrt{1 - \frac{\text{سـ}'^2}{\text{ع}'^2}} \quad \ldots \qquad (٧)$$

وهكذا قد أثبتنا أن هناك علاقة جبرية بين معامل الارتباط ر والتشتت سـَ ، للنقط المعطاة حول خط الانحدار أو خط الانتشار ، على فرض أن هذا الخط مستقيم . ومن الواضح أن ر تكون كبيرة إذا كانت سـَ صغيرة « أى التشتت صغير » ، والعكس بالعكس . وهذا هو ما أشرنا إليه فى بند ٢٣٨ .

الخطأ المعيارى لقيم ص

٢٤٦ — هذا المقدار سـ يشابه تماماً الانحراف المعيارى لمجموعة من القيم حول متوسطها الحسابى ، إذ أن الانحراف المعيارى يساوى الجذر التربيعى

لمجموع مربعات الانحرافات عن الوسط مقسوماً على عـدد القيم ن . وكذلك سہَ تساوى الجـذر التربيعى لمجموع انحرافات النقط عن الخط (الذى سميناه خط العلاقة المتوسـطة) مقسوماً على عدد النقط وهو ن . وللتمييز نسميه[1] **الخطأ المعيارى** للقيم الصادية ، ونرمز له بالرمز سہص أو للسهولة بالحرف سہَ ، وذلك لأن هذه الانحرافات هى فى الحقيقـة انحرافات القيم الفعلية التى يأخذها المتغير ص فى التجربة ، عن القيم النظرية التى كان يجب أن يأخذها هذا المتغير ، لو أنه اتبع القانون الاعتبارى الذى تمثله المعادلة المفروضة :

ص = م س + حـ

حساب سہ من الخطأ المعيارى لقسيم ص

٢٤٧ — هكذا نرى أنه من الممكن حساب معامل الارتباط بين متغيرين مثل س و ص ، بأن نوفق خط العلاقة المتوسطة بينهما ، ثم نحسب الفرق بين قيمة ص المشاهدة فى التجربة وقيمتها النظرية ، المحسوبة على أساس هذه المعادلة التى وفقناها . ثم نحسب الخطأ المعيارى سہَ^{2} بتربيع هذه الفروق وقسمة مجموع هذه المربعات على عددها . ونحسب أيضـاً الانحراف المعيارى عَ للقيم الصادية نفسها المشاهدة من التجربة . ثم نحسب ر من المعادلة

سہَ^{2} = عَ^{2} (١ — ر^{2}) .

وهذه الطريقة أفضل وأعم ، وأحياناً تـكون أسهل من الطرق العـادية ؛ والحقيقة أننا لا نحتاج عملياً لحساب القيم النظرية من المعادلة التى نوفقها ثم نطرحها من القيم المشاهدة ؛ ولـكننا نحصل على سہَ^{2} مباشرة من المعادلة (٥) بند ٢٣٨ وهى

ن سہَ^{2} = مح ع^{2}

= مح ص^{2} — م مح س ص — حـ مح ص .

---

(١) فى الانجليزية (Standard Error, *Sy.*)

ويلاحظ أن الـكميات مح ص و مح س ص معلومة لنا حيث استخدمناها فى توفيق المعادلة الأصلية ص = م س + حـ . ومتى عرفنا الانحراف المعيارى للقيم الصادية ، وهو عَ ، أمكننا بسرعة حساب مح ص^{2} من المعادلة

مح ص^{2} = ن عَ^{2} + ن ص̄^{2} .

خط انحدار غير مستقيم

٢٤٨ — ننتقل الآن إلى البحث فى الارتباط غير المستقيم الذى فيه يكون خط الانحدار غير مستقيم ، بل منحنياً من الدرجة الثانية أو أعلى .

لنفرض أولا أن معادلة خط الانحدار من الدرجة الثانيـة بالنسبة إلى س ، وأنها على الصورة

ص = ا س^{2} + ب س + حـ ، . . . . . ، (١)

حيث ا ، ب ، حـ ثلاث كميات ثابتة ، نختارها بحيث يكون هذا المنحنى (١) أوفق المنحنيات لتمثيل القيم المشاهدة المعطاة وهى :

(س$_{1}$ ، ص$_{1}$) ، (س$_{2}$ ، ص$_{2}$) ، . . . . ، (س$_{ن}$ ، ص$_{ن}$) .

و بطبيعة الحال تـكون قيم ا ، ب ، حـ هذه تحقق المعادلات الثلاث :

مح ص = ا . مح س^{2} + ب . مح س + ن حـ . . . (٢) ،

مح س ص = ا . مح س^{3} + ب . مح س^{2} + حـ . مح س . . . (٣) ،

مح س^{2} ص = ا . مح س^{4} + ب . مح س^{3} + حـ . مح س^{2} . . . (٤) ،

وهذه المعادلات الثلاث كما نعلم ، كافية لتعيين قيم المجاهيل الثلاثة ا ، ب ، حـ . وهكذا نحصل على معادلة خط الانحدار للمتغير ص على المتغير س ؛ ويبقى أن نعرف كيف نستخدم هذه المعادلة لقياس الارتباط بين هذين المتغيرين كما فعلنا فى حالة الارتباط المستقيم .

تشتت حول خط منحن

٢٤٩ — لنبحث فى تشتت النقط (س$_1$ ، ص$_1$) ، . . . ، (س$_ن$ ، ص$_ن$) حول هذا الخط الاعتبارى : ص = ا س^2 + ب س + حـ .

لنفرض أن ع$_1$ هو انحراف النقطة (س$_1$ ، ص$_1$) عن هذا المنحنى ،

∴ ع$_1$ = ا س$_1^2$ + ب س$_1$ + حـ − ص$_1$ . . . . (١) ،

ولو ضربنا هذا الانحراف فى س$_1$ و س$_1^2$ و ص$_1$ و ع$_1$ على التوالى ،

∴ ع$_1$ س$_1$ = ا س$_1^3$ + ب س$_1^2$ + حـ س$_1$ − س$_1$ ص$_1$. (٢) ،

و ع$_1$ س$_1^2$ = ا س$_1^4$ + ب س$_1^3$ + حـ س$_1^2$ − س$_1^2$ ص$_1$ . . (٣) ،

و ع$_1$ ص$_1$ = ا . س$_1^2$ ص$_1$ + ب . س$_1$ ص$_1$ + حـ ص$_1$ − ص$_1^2$. (٤) ،

و ع$_1^2$ = ا . ع$_1$ س$_1^2$ + ب ع$_1$ س$_1$ + حـ ع$_1$ − ع$_1$ ص$_1$. (٥) ،

وهكذا بالنسبة إلى جميع النقط الأخرى (س$_2$ ، ص$_2$) ، . . . ، (س$_ن$ ، ص$_ن$) .

فيكون مجموع مربعات انحرافات النقط جميعها عن هذا الخط يساوى :

مجـ ع^2 = ن سَ^{2}

= ا مجـ ع س^2 + ب مجـ ع س + حـ . مجـ ع − مجـ ع ص ، . . (٦) .

ولكن مجـ ع = ا مجـ س^2 + ب مجـ س + ن حـ − مجـ ص معادلة (١) أعلاه ،

= صفراً « (٢) بند ٢٤٨ .

و مجـ ع س = ا . مجـ س^3 + ب مجـ س^2 + حـ . مجـ س − مجـ س ص « (٢) أعلاه ،

= صفراً « (٣) بند ٢٤٨ .

و مجـ ع س^2 = ا . مجـ س^4 + ب . مجـ س^3 + حـ . مجـ س^2 − مجـ س^2 ص « (٣) أعلاه ،

= صفراً « (٤) بند ٢٤٨ .

∴ ن سَ^{2} = مجـ ع^2

= − مجـ ع ص ؛ ومن معادلة (٤) أعلاه :

= مجـ ص^2 − ا مجـ س^2 ص − ب مجـ س ص − حـ مجـ ص ، . (٧) ،

ويلاحظ أن الكميات مجـ س^2 ص و مجـ س ص و مجـ ص هى نفس الكميات التى استخدمناها فى توفيق المنحنى وحساب قيم الكميات ا و ب و حـ ، ولسنا محتاجين لحسابها من جديد . أما الكمية مجـ ص^2 فهى معروفة بدلالة الوسط الحسابى والانحراف المعيارى لقيم المتغير ص ، حيث مجـ ص^2 = ن عَ^{2} + ن $\overline{\text{ص}}^2$

وبناء على ذلك يمكننا بسهولة معرفة مقدار تشتت النقط المعطاة حول خط الانحدار ، أو خط العلاقة المتوسطة ، الذى معادلته التحليلية هى :

ص = ا س^2 + ب س + حـ .

والكمية سَ نطلق عليها أيضاً اسم **الخطأ المعيارى** لقيم ص كما قلنا فى بند ٢٤٦ .

٢٥٠ — فى حالة الارتباط المستقيم وجدنا أن العلاقة الجبرية بين الخطأ المعيارى سَ ومعامل الارتباط ر هى : دليل الارتباط

ر^2 = ١ − $\dfrac{\text{سَ}^2}{\text{عَ}^2}$ ، (١) ،

حيث عَ هى الانحراف المعيارى لقيم ص .

وقياساً على ذلك ، إذا كان ط يدل على معامل الارتباط فى هذه الحالة ، حالة الارتباط غير المستقيم ، تكون العلاقة الجبرية بين ط و سَ و عَ هى :

ط^2 = ١ − $\dfrac{\text{سَ}^2}{\text{عَ}^2}$ ، (٢) .

ويكون ط مقياس الارتباط غير المستقيم . وللتمييز بينه وبين ر ، معامل الارتباط المستقيم ، نسميه[١] **دليل الارتباط** .

وبوضع مجـ ص^{٢} = ن ع$'^{٢}$ + ن $\overline{\text{ص}}^{٢}$ فى المعادلة (٧) بند ٢٤٩ ، واستبدال سه$'^{٢}$ فى المعادلة (٢) أعلاه بقيمتها

$$\therefore\ \text{ط}^{٢} = \frac{١}{\text{ن ع}'^{٢}}\left[\text{ا مجـ س}^{٢}\text{ص} + \text{ب . مجـ س ص} + \text{حـ . مجـ ص} - \text{ن}\,\overline{\text{ص}}^{٢}\right]$$

خط الانحدار من الدرجة الثالثة

٢٥١ — وكذلك لو كان خط الانحدار ذا معادلة من الدرجة الثالثة مثل

$$\text{ص} = \text{ا س}^{٣} + \text{ب س}^{٢} + \text{حـ س} + \text{ى} \quad . \quad . \quad (١) .$$

يمكننا بنفس البرهان السابق إثبات أن الخطأ المعيارى فى هذه الحالة أيضاً هو سه$'$ حيث

$$\text{ن سه}'^{٢} = \text{مجـ ص}^{٢} - \text{ا مجـ س}^{٣}\text{ص} - \text{ب . مجـ س}^{٢}\text{ص} - \text{حـ . مجـ س ص} - \text{ى . مجـ ص} ؛$$

ويكون دليل الارتباط هو ط أيضاً ، حيث

$$\text{ط}^{٢} = ١ - \frac{\text{سه}'^{٢}}{\text{ع}'^{٢}} .$$

خط الانحدار من أى درجة

٢٥٢ — وهكذا لو كان خط الانحدار من أى درجة فوق الثالثة . وعلى العموم نفرض أن معادلة خط الانحدار من أى درجة ، ونفرض أن هذه المعادلة هى

$$\text{ص} = \text{ك} + \text{ل س} + \text{م س}^{٢} + \text{و س}^{٣} + \text{هـ س}^{٤} + \ldots \quad (٢) ،$$

حيث ك ، ل ، م ، و ، هـ ، . . . كميات ثابتة مستقلة عن كل من س و ص ، ويصح أن يكون بعضها سالباً أو موجباً أو صفراً .

(١) بالانجليزية (Index of Correlation) ؛ ويرمز له بالحرف الأغريقى ρ (رُوه)

وباتباع نفس الخطوات كما فى بند ٢٤٩ ، ويمكننا بسهولة إثبات أن الخطأ المعيارى (وهو مقياس التشتت) هو سه$'$ حيث :

$$\text{ن سه}'^{٢} = \text{مجـ ص}^{٢} - \text{ك . مجـ ص} - \text{ل . مجـ س ص} - \text{م . مجـ س}^{٢}\text{ص} - \text{و . مجـ س}^{٣}\text{ص} - \text{هـ . مجـ س}^{٤}\text{ص} - \ldots \quad (٣) .$$

وفى حالة ما يكون خط الانحدار مستقيما ، تكون المعادلة (٢) من الدرجة الأولى فقط وتكون م = و = هـ = . . . صفراً . وإذا كان خط الانحدار من الدرجة الثانية كانت و = هـ = . . . = صفراً . وهكذا . وفى كل هذه الأحوال يكون دليل الارتباط ، وهو مقياس العلاقة بين المتغيرين س و ص ، هو ط حيث :

$$\text{ط}^{٢} = ١ - \frac{\text{سه}'^{٢}}{\text{ع}'^{٢}}$$

ط = ر فى حالة الانحدار المستقيم فقط

٢٥٣ — ويلاحظ أن دليل الارتباط ط يساوى معامل الارتباط ر فى حالة واحدة فقط ، ألا وهى حالة الارتباط المستقيم حيث يكون خط الانحدار مستقيما ، معادلته من الدرجة الأولى ، وحيث تؤول المعادلة (٣) فى البند السابق إلى الصورة .

$$\text{ن سه}^{٢} = \text{مجـ ص}^{٢} - \text{ك . مجـ ص} - \text{ل مجـ س ص} ،$$

حيث ك هنا بدل حـ ، ول بدل م المستعملتين في بند ٢٤٥ ، ومنه ينتج أن $\text{سه}'^{٢} = \text{ع}^{٢}\,(١ - \text{ر}^{٢})$ ، كما فى معادلة (٦) بند ٢٤٥

$$\text{ط}^{٢} = ١ - (١ - \text{ر}^{٢})$$
$$= \text{ر}^{٢} .$$

ولكن هذه العلاقة بين ط و ر ليست صحيحة فى حالة الانحدار غير المستقيم

على اختلاف أنواعه ، حيث تكون معادلة الخطأ المعيارى معقدة ، كما رأينا فى (٣) بند ٢٥٢ ؛ ولا يمكن اختصارها كما فعلنا هنا فى حالة الانحدار المستقيم .

ط أعم وأدق من ر فى قياس الارتباط

٢٥٤ — وكذلك العلاقة بين سَ و ر ، وهى

$$\text{سَ}^{2} = \text{عَ}^{2}\,(1 - \text{ر}^{2})$$

ليست صحيحة إلا فى حالة الانحدار المستقيم فقط . وعلى ذلك لا يمكن الاعتماد على معامل الارتباط ر كمقياس دقيق للارتباط فى حالات الانحدار غير المستقيم بل يجب فى هذه الأحوال استخدام ط . لأن العلاقة بين سَ وط هى دائماً

$$\text{سَ}^{2} = \text{عَ}^{2}\,(1 - \text{ط}^{2})$$

مهما كانت درجة معادلة خط الانحدار . ولذلك يعتبر دليل الارتباط أعم وأدق من معامل الارتباط فى قياس العلاقة بين متغيرين .

ر هى معامل الاستقامة لخط الانحدار

٢٥٥ — وهذا هو السبب فى أننا نرى حديثاً أن بعض الإحصائيين ينصح بعدم استعمال ر بتاتاً لقياس الارتباط ، ويقترح استخدامه فقط كمقياس لدرجة استقامة الانحدار ، وتسميته **معامل الاستقامة**[١] . وذلك لأنه إذا كان ر كبيراً كان مقدار سَ صغيراً ، وبالتالى يكون تشتت النقط حول الخط المستقيم صغيراً ، مما يدل على أن هذا المستقيم يوافق جيدا النقط المعطاة . أى أن خط الانحدار قريب من المستقيم . وإذا كان ر صغيراً ، كان التشتت سَ حول هذا المستقيم كبيراً ، مما يدل على أن هذا المستقيم لا يوافق هذه النقط جيداً . أى أن خط الانحدار بعيد عن الاستقامة .

(١) بالإنجليزية (Coefficient of Linearity) ؛ انظر بحث الأستاذ M. Frèchet فى مجلة Revue de l'Institut Internationale de Statistique, 1936, Part 4 p. 16–23.

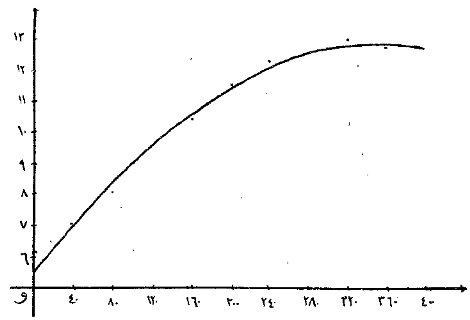
المقارنة بين ط و ر عملياً

٢٥٦ — نأخذ الآن مثالا ونحسب معامل الارتباط ودليل الارتباط ، ونقارن بينهما فى قياس العلاقة بين المتغيرين .

عملت تجربة فى عدد من الحقول بقصد دراسة العلاقة بين كمية السماد المستعمل وكمية المحصول الناتج ، فكانت الأرقام كما يأتى[١] :

كمية السماد للفدان بالكيلوا جرام	٠	٤٠	٨٠	١٢٠	١٦٠	٢٠٠	٢٤٠	٢٨٠	٣٢٠	٣٦٠	٤٠٠
محصول الفدان بالأردب	٦٫٣	٧٫٠	٨٫١	٩٫٣	١٠٫٥	١١٫٦	١٢٫٣	١٣٫١	١٣٫٠	١٢٫٨	١٢٫٥

لنرمز لكمية السماد بالحرف س وللمحصول بالحرف ص . ثم نرسم محورين متعامدين ونرصد النقط التى إحداثياتها الأفقية تساوى قيم س ، وإحداثياتها الرأسية تساوى قيم ص على الترتيب ، كما فى شكل ٦٦ . ويتضح



(شكل ٦٦)

العلاقة بين كمية السماد والمحصول

(١) فى هذا المثال النظرى أخذنا ١١ حقلا فقط وذلك لتوضيح خطوات العمل . ولكن عملياً يجب ألا يقل العدد عن حوالى ٣٠ ؛ وذلك لكى نتفادى أخطاء المصادفات .

من مواقع هذه الإحدى عشرة نقطة أن أحسن خط يوافقها تكون معادلته من الدرجة الثانية . لنفرض هذه المعادلة هى :

ص = ا س^{٢} + ب س + هـ (١)

حيث ص تساوى كمية المحصول و س تمثل كمية السماد . ونوجد قيم ا ، ب ، هـ كالمعتاد من المعادلات

مجـ ص = ا . مجـ س^{٢} + ب . مجـ س + ١١ هـ (٢)

و مجـ س ص = ا . مجـ س^{٣} + ب . مجـ س^{٢} + هـ . مجـ س (٣) ،

و مجـ س^{٢} ص = ا . مجـ س^{٤} + ب . مجـ س^{٣} + هـ . مجـ س^{٢} (٤) .

ولتسهيل حل هذه المعادلات لإيجاد قيم ا ، ب ، هـ ، نبدأ قياس السينات من الوسط حتى يكون بعضها سالباً والبعض موجباً فتختصر ، ولاسيما أن قيم س المعطاة واقعة على مسافات متساوية من بعضها . ونجد خطوات العمل موضحة فى الجدول الآتى :

جدول ٤٤ — توفيق منحن من الدرجة الثانية وحساب دليل الارتباط .

كمية السماد	س	سَ	المحصول ص	سَ^{٢}	سَ^{٤}	سَ ص	سَ^{٢} ص	ص^{٢}
٠	−٢٠٠	−٥	٦٫٢	٢٥	٦٢٥	−٣١٫٠	١٥٥٫٠	٣٨٫٤٤
٤٠	−١٦٠	−٤	٧٫٠	١٦	٢٥٦	−٢٨٫٠	١١٢٫٠	٤٩٫٠٠
٨٠	−١٢٠	−٣	٨٫١	٩	٨١	−٢٤٫٣	٧٢٫٩	٦٥٫٦١
١٢٠	−٨٠	−٢	٩٫٣	٤	١٦	−١٨٫٦	٣٧٫٢	٨٦٫٤٩
١٦٠	−٤٠	−١	١٠٫٥	١	١	−١٠٫٥	١٠٫٥	١١٠٫٢٥
٢٠٠	٠	٠	١١٫٦	٠	٠	٠	٠	١٣٤٫٥٦
٢٤٠	٤٠	١	١٢٫٣	١	١	١٢٫٣	١٢٫٣	١٥١٫٢٩
٢٨٠	٨٠	٢	١٣٫١	٤	١٦	٢٦٫٢	٥٢٫٤	١٧١٫٦١
٣٢٠	١٢٠	٣	١٣٫٠	٩	٨١	٣٩٫٠	١١٧٫٠	١٦٩٫٠٠
٣٦٠	١٦٠	٤	١٢٫٨	١٦	٢٥٦	٥١٫٢	٢٠٤٫٨	١٦٣٫٨٤
٤٠٠	٢٠٠	٥	١٢٫٥	٢٥	٦٢٥	٦٢٫٥	٣١٢٫٥	١٥٦٫٢٥
المجموعات		٠	١١٦٫٤	١١٠	١٩٥٨	٧٨٫٨	١٠٨٦٫٦	١٢٩٦٫٣٤

وبالتعويض فى المعادلات (٢) ، (٣) ، (٤) أعلاه

∴ ١١٦٫٤ = ١١٠ ا + ٠ + ١١ هـ (٤) ،

و ٧٨٫٨ = ٠ + ١١٠ ب + ٠ (٥) ،

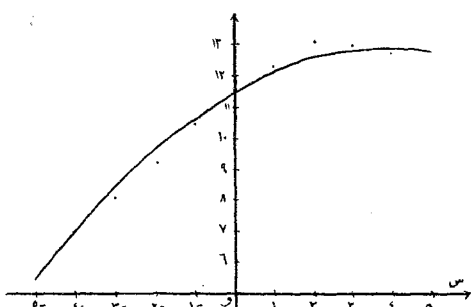
و ١٠٨٦٫٦ = ١٩٥٨ ا + ٠ + ١١٠ هـ (٦) ،

∴ ا = − ٠٫٠٩٠٢ ، ب = ٠٫٧١٦٤ ، هـ = ١١٫٤٨٣٩

∴ معادلة خط الانحدار هى

ص = − ٠٫٠٩٠٢ سَ^{٢} + ٠٫٧١٦٤ سَ + ١١٫٤٨٣٩ (٧)

حيث ص تساوى المحصول بالأرادب ، فى حين أن سَ تدل على كمية السماد محسوبة ابتداء من ٢٠٠ كيلو جرام ، بوحدات كل منها تساوى ٤٠ كيلو جراماً . ونرى هذا المنحنى مرسوماً فى شكل ٦٧ .



(شكل ٦٧)

ولكن ١١ سهـ^{٢} = مجـ ص^{٢} — ا مجـ س^{٢} ص — ب مجـ س ص — هـ مجـ ص

= ١٢٩٦٫٣٤ + ٩٨٫٠١١٣ — ٥٦٫٤٥٢٣ — ١٣٣٦٫٧٢٦٠

= ١٫١٧٣٠

$$\text{سَ}^{2} = \text{١٠٦٦ر}$$

$$\text{مجـ ص}^{2} = \text{١١}\,\text{عَ}^{2} + \text{١١}\,\overline{\text{ص}}^{2} \quad \text{و} \quad \text{١١}\,\overline{\text{ص}} = \text{١١٦ر٤}$$

$$\text{عَ}^{2} = \frac{\text{١}}{\text{١١}} \times \text{١٢٩٦ر٣٤} - (\text{١٠ر٥٨١٨})^{2} \text{ ،}$$

$$= \text{٥ر٨٧٤٢} \quad .$$

$$\text{ط}^{2} = \text{١} - \frac{\text{١٠٦٦ر}}{\text{٥ر٨٧٤٢}}$$

$$= \text{٩٨١٨ر}$$

$$\text{ط} = \text{٩٩ر}$$

ولحساب معامل الارتباط ر ، تستخدم المعادلة

$$\text{ر} = \frac{\text{مجـ س ص} - \text{ن}\,\overline{\text{س}}\,\overline{\text{ص}}}{\text{ن . ع . عَ}} \text{ ،}$$

حيث $\text{ن} = \text{١١}$ ، $\overline{\text{س}} = \text{٠}$ ، $\overline{\text{ص}} = \text{١٠ر٥٨١٨}$ ، و $\text{مجـ سَ ص} = \text{٧٨ر٨}$

و $\text{١١ع}^{2} = \text{مجـ سَ}^{2} = \text{١١٠}$ ، و $\text{عَ}^{2} = \text{٥ر٨٧٤٢}$

$$\therefore \quad \text{ر} = \text{٩٧ر}$$

قيمة ط تتوقف على نوع المنحنى المفروض كخط الانحدار

٢٥٧ — نرى من هذا المثال أن ط أكبر من ر ؛ والسبب فى ذلك هو أن القيم المعطاة يوافقها منحن من الدرجة الثانية أفضل من خط مستقيم . ولذلك كان تشتتها حول الخـط من الدرجة الثانية ، الذى استخدمناه فى حساب ط ، أقل من تشتتها حول الخط المستقيم الذى افترضنا موافقته لتمثيل هذه القيم ، فى حساب المعامل ر .

ولو أننا رأينا فى بادئ الأمر — خطأً — أن القيم المعطاة يوافقها منحن من الدرجة الثالثة ، بدلا من الدرجة الثانية ، وحسبنا دليل الارتباط على هـذا الأساس ، وليكن طَ مثلا . لوجدنا أن طَ أصـغر من ط . والسبب فى ذلك أن

تشتت القيم حول هذا الخط الأخير أكبر من تشتتها حول الخط الأصلى ذى الدرجة الثانية ، الذى يوافقها فى الحقيقة أفضل من الخط ذى الدرجة الثالثة .

وبالمثل ، لو أن هـذه القيم المعطاة يوافقها خط مستقيم أحسن من أى خط آخر ، وأننا رأينا (خطأ ، لسبب ما) أن نوفق لها منحنياً من الدرجة الثانية ، فحسبنا ط على هذا الأساس ، لوجدنا أن معامل الارتباط ر أكبر من الدليل ط . والسبب فى ذلك هو نفس السبب المتقـدم : ألا وهو أن تشتت القيم حول الخط الجيد أقل من تشتتها حول أى خط آخر أقل موافقة من الأول . والقاعدة إذن هى أن مقياس الارتباط (ر أو ط) المحسوب على أساس الخط الجيد ، أكبر من المقياس المحسوب على أساس أى خط آخر أقل منه توفيقاً .

قياس الارتباط فى النهاية مسألة اعتبارية يدخل فيها العامل الشخصى

٢٥٨ — مقياس الارتباط الذى نحصل عليه فى أى مسألة معينة يتوقف إذن على نوع المنحنى الذى نعتبره صالحاً لتمثيل العلاقة المتوسطة بين المتغيرين . ومسألة اختيار المنحنى الأصلح ، وإن كانت تبدو سهلة فى ظاهرها ، حيث يسهل فى بعض الأحيان الاتفاق على نوع المنحنى الذى يوافق مجموعة من القيم المشاهدة لمتغيرين أحسن من غيره ، كثيراً ما تكون صعبة أو ملتبسة لا يمكن الوصول فيها إلى رأى قاطع . وفى مثل هذه الأحوال لا بد أن يترك أمر اختيار المنحنى الأصلح إلى تقديرنا واعتبارنا . وبهذه الطريقة يدخل العامل الشخصى فى أبحاثنا ويتحكم فى المقياس الذى نحصل عليه لدرجة العلاقة بين المتغيرين تحت البحث .

وهكذا نرى أن مسألة قياس الارتباط بين متغيرين علمت لنا مجموعـة من القيم المشاهدة لها ، تستند فى النهاية على أمور اعتبارية — إلى حدٍ ما على الأقل . وكنا نود لو أن هذا العامل الشخصى لم يتدخل فى حكمنا على الأشياء ، ولـكن للأسف لا توجـد قاعدة مضبوطة نرجع إليها فى معرفة أى نوع من المنحنيات

يصلح لتمثيل مجموعة من القيم المعينة أفضل من غيره . وهذه حقيقة لا نستطيع فى الوقت الحاضر أن نكذبها أو ننكرها .

نسبة الارتباط

٢٥٩ — إذا كان لدينا جدول تكرارى مزدوج لمتغيرين مثل س وص ، ورأينا من توزيع قيمها فى هذا الجدول أن خط الانحدار بينهما لا يمكن أن يكون مستقيما ، فلا يمكننا والحالة هذه الاعتماد على معامل الارتباط ر كمقياس دقيق للعلاقة بين المتغيرين . ولا يمكننا أيضاً حساب دليل الارتباط ط ، لأن هذا لا بد له من توفيق منحن لقيم ص بالنسبة إلى س ؛ وهذا غير ميسور ، إذ أن قيم س وقيم ص ليست عندنا منفردة ، بل مقسمة إلى مجموعات فى فئات تكرارية مزدوجة . وقد وضع كارل بيرسون مقياساً سماه[1] **نسبة الارتباط** ، لاستخدامه فى مثل هذه الأحوال . ولنرمز لهذا المقياس بالحرف ى . ونحسبه كما يأتى :

أولا : نحسب المتوسط الصادى لكل عمود (أى فى كل فئة من فئات س) ؛

ثانياً : فى كل خانة من أى عمود ، نحسب مربع انحراف قيمة ص عن الوسط الحسابى الصادى فى هذا العمود ؛

ثالثاً : نجمع مربعات الانحرافات الصادية لكل عمود ، فنحصل على تشتت الصادات فى هذا العمود حول وسطها الحسابى

رابعاً : نجمع هذه التشتتات للأعمدة كلها فنحصل على مجموع مربعات انحرافات القيم الصادية كل عن الوسط الحسابى فى عمودها . لنفرض أن هذا المجموع يساوى $ن\ سم^{٢}_{م\,ص}$ ، حيث ن يساوى الحالات فى الجدول كله .

(١) اسمه بالانجليزية (Correlation Ratio) ، ويرمز له بالحرف الأغريقى η (إيتا) .
انظر كتاب Whittaker and Robinson : *Calculus of Observations*, (1929), p. 336.

خامساً : نحسب الانحراف المعيارى لقيم ص ؛ وليكن هذا يساوى عَ كالمعتاد

سادساً : نحسب ى من المعادلة

$$ى^{٢} = ١ - \frac{سم^{٢}_{م\,ص}}{عَ^{٢}} \qquad (١)$$

ويلاحظ أن هذه المعادلة على صورة معادلات ط و ر ، ما عدا التشتت $سم_{م\,ص}$ هنا بدل التشتت البسيط سمَ هناك .

حساب التشتت $سم_{م\,ص}$ عمليا

٢٦٠ — ولكن حساب التشتت $سم_{م\,ص}$ بالطريقة المتقدمة متعب ويستلزم مجهوداً كبيراً ، ولذلك نحسبه بطريقة مختصرة غير مباشرة ، بواسطة حساب الانحراف المعيارى لنفس المتوسطات الحسابية للأعمدة التى أوجدناها فى الخطوة الأولى أعلاه . لنفرض أن هذا الانحراف المعيارى يساوى $ع_{م}$ مثلا .

وبما أن مجموع مربعات انحرافات عدد من القيم كل واحدة عن الوسط الحسابى لمجموعتها ، يساوى مجموع مربعات انحرافاتها عن الوسط الحسابى العمومى ، ناقصاً مجموع مربعات انحرافات الأوساط الحسابية للمجموعات عن هذا الوسط العمومى نفسه[1] ،

$$\therefore\ ن\,.\,سم^{٢}_{م\,ص} = ن\,عَ^{٢} - ن\,ع^{٢}_{م} \qquad (٢) ،$$

$$\therefore\ ى^{٢} = ١ - \frac{عَ^{٢} - ع^{٢}_{م}}{عَ^{٢}} \qquad (٣) ،$$

$$\therefore\ ى = \frac{ع_{م}}{عَ} \qquad (٤) ،$$

وهذه العلاقة أبسط من المعادلة (١) ، من ناحية الشكل ومن الناحية

(١) انظر المعادلة (٣) فى بند ١٨٠ صفحة ٢٠١ ، حيث نضع $ن\ سم^{٢}_{م\,ص}$ بدل $ن\ ع^{٢}$ هناك .

العملية الحسابية أيضاً ، حيث تؤول المسألة إلى حساب المتوسطات الحسابية للأعمدة ، وحساب انحرافها المعيارى ، والانحراف المعيارى لعموم الصادرات ؛ ونسبة الارتباط هى خارج قسمة الأخيرين .

نسبة الارتباط تساوى دليل الارتباط فى النهاية

٢٦١ — وإذا تأملنا فى الفكرة الأساسية التى ينبنى عليها تعريف نسبة الارتباط ، نجد أن هذه النسبة تساوى (تقريباً) دليل الارتباط . فقد قلنا إن $\sum$ سه²ص ، المستعملة فى حساب ى ، هى مجموع مربعات انحرافات القيم الصادية ، كل عن الوسط الحسابى لمجموعتها ، فى الجدول التكرارى المزدوج ؛ وكل مجموعة من هذه القيم الصادية تقترن بقيمة معينة للمتغير س ، ألا وهى مركز الفئة السينية المبينة أعلى كل عمود من أعمدة هذا الجدول التكرارى . انظر مثلا فى جدول ٣٢ (صفحة ٢٣٥) حيث نجد فى العمود الثالث منه ٥٣ رجلا عمرهم جميعاً يساوى ٢٧٫٥ سنة . ولكن عدد ما عندهم من الأطفال يختلف : فمنهم ثمانية ليس عندهم أطفال بالمرة ، ومنهم ٢٥ عندهم طفل واحد ، وهكذا . ومتوسط ما عند هؤلاء من الأطفال يساوى ١٫٣٨ طفل .

وقلنا إن $\sum$ سه'² المستعملة فى حساب ط ، تساوى مجموع مربعات انحرافات القيم الصادية ، كل عن القيمة النظرية المحسوبة على أساس معادلة خط الانتشار ، أو خط العلاقة المتوسطة (وهو فى النهاية خط الانحدار) ، الذى فرضنا أنه يمثل العلاقة بين المتغيرين س ، ص . ولو رجعنا إلى شكل ٦٤ (صفحة ٢٦٩) ، وما قلناه فى بند ٢٣٩ ، وجدنا أن هذه القيمة النظرية هى فى النهاية متوسط القيم الصادية الواقعة فى هذه الفئة . ومن ذلك يتضح أن سه' و سهم ، وبالتالى ط ، ى ، متساويتان فى النهاية ؛ وأن الفرق بينهما عملياً يكون صغيراً ، خصوصاً لو كانت الفئات السينية فى جدول الارتباط ضيقة المدى ، وكذلك فئات ص ؛

الفرق بين دليل الارتباط ط ونسبة الارتباط ى ، لم يخرج إذن عن كونه تقريباً فى حدود الدقة ، بقصد تسهيل العمل الحسابى . فنستعمل ط حينما يكون لدينا قيم منفردة متناظرة للمتغيرين س و ص ، عددها صغير لا يستلزم مجهوداً حسابياً كبيراً . أما إذا كان لدينا عدد كبير جداً من قيم س و ص فقسمناها إلى فئات ووضعناها فى جدول تكرارى مزدوج ، تعذر علينا استخدام ط ، ولجأنا حينئذ إلى حساب نسبة الارتباط ى من واقع هذا الجدول .

ويلاحظ بهذه المناسبة أن استخدام ى يغنينا عن عملية توفيق المنحنى اللازم لحساب ط ؛ ويعفينا من مواجهة المشكلة المرتبطة بهذه العملية والتى تكلمنا عنها فى بند ٢٥٧ ، ألا وهى اختيار أصلح المنحنيات لتمثيل البيانات المعطاة لنا .

حساب نسبة الارتباط عملياً

٢٦٢ — نأخذ الآن مثالا مختصراً لتوضيح خطوات العمل فى حساب ى . فى جدول ٤٥ نجد التوزيع التكرارى المزدوج لأعمار الرجال المتزوجين (من آنسات) وأعمار زوجاتهم يوم الزواج . والملاحظ عادة أن الآنسات يتزوجن فى مقتبل العمر ، ولا يبقى منهن لسن متأخرة إلا القليل ، ومهما بقين فهن لا يتزوجن بعد سن معينة ، إلا نادراً جداً . أما الرجال فكثيراً ما يتزوجون فى السنين المتأخرة من العمر .

نحسب المتوسط الصادى لكل عمود، وهو متوسط أعمار زوجات الرجال الذين من سن واحدة. ويلاحظ أن هذه المتوسطات تزيد بسرعة فى بادىء الأمر مع أعمار الرجال ، ولكن معدل هذه الزيادة لا يلبث أن يهبط ويتلاشى تدريجياً ، مما يدل على أن أعمار الزوجات من الآنسات يزيد تبعاً لأعمار أزواجهن إلى حد معين ثم يقف . أى أن خط انحدار ص على س ليس مستقيما . ومن هذا يتضح أن معامل الارتباط ر لا يصلح فى هذه الحالة ، وينبغى إذن حساب نسبة الارتباط ى .

جدول ٤٥ — توزيع تكرارى مزدوج لأعمار الرجال = س ،
وأعمار زوجاتهم = ص

ص \ س	٢٠ —	٣٠ —	٤٠ —	٥٠ —	٦٠ —	٧٠ —	٨٠ —	المجموع
١٥ —	٣٥	٦٣	٢٢	٢				١٢٢
٢٠ —	٢٥	٧٢	٣٣	١٣	٥	١		١٤٩
٢٥ —	١٧	٥١	٣٠	١٦	٨	٢	١	١٢٥
٣٠ —	٣	٣٣	٢٨	١٩	١٥	٧	٣	١٠٧
٣٥ —	١	١٨	٢٠	١٥	٧	٤	٠	٦٥
٤٠ —		٣	١٢	١٠	٥	١	١	٣٢
المجموع	٨٠	٢٤٠	١٤٥	٧٥	٤٠	١٥	٥	٦٠٠
متوسط ص	٢١ر٨١٢٥	٢٥ر٠٠٠	٢٨ر٤٣١٠	٣١ر٦٣٣٣	٣٢ر٣٧٥٠	٣٣ر١٦٦٦	٣٣ر٥٠٠	

لذلك نحسب الانحراف المعيارى $ع_م$ لهذه المتوسطات الحسابية الصادية ؛ وكذلك الانحراف المعيارى لقيم ص ؛ وتكون ى تساوى خارج قسمة الأول على الثانى .

لايجاد الانحراف المعيارى لقيم ص ، نستخدم التوزيع التكرارى لها ، كما هو مبين بالجدول السابق :

جدول ٤٦ — حساب الانحراف المعيارى لأعمار الزوجات

الفئات	مراكزها ص	التكرار ك	الانحرافات ع	ك . ع	ك . $ع^٢$
١٥ —	١٧ر٥	١٢٢	— ١٠	— ١٢٢٠	١٢٢٠٠
٢٠ —	٢٢ر٥	١٤٩	— ٥	— ٧٤٥	٣٧٢٥
٢٥ —	٢٧ر٥	١٢٥	٠	٠	٠
٣٠ —	٣٢ر٥	١٠٧	٥	٥٣٥	٢٦٧٥
٣٥ —	٣٧ر٥	٦٥	١٠	٦٥٠	٦٥٠٠
٤٠ —	٤٢ر٥	٣٢	١٥	٤٨٠	٧٢٠٠
		٦٠٠		— ٣٠٠	٣٢٣٠٠

∴ الوسط الحسابي العمومى لقيم ص يساوى

$$\overline{ص} = ٢٧ر٥ - \frac{٣٠٠}{٦٠٠} ،$$

$$= ٢٧ \text{ سنة .}$$

$$٣٢٣٠٠ = ٦٠٠ \, ع'^٢ + ٦٠٠ \, (٠ر٥)^٢$$

$$ع'^٢ = ٥٣ر٥٨٣٣$$

$$ع' = ٧ر٣٢٠٦ \text{ سنة}$$

ولحساب الانحراف المعيارى $ع_م$ للمتوسطات الصادية حول الوسط الحسابى العمومي للصادرات وهو ٢٧ سنة ، نربع انحراف كل منها عن ٢٧ ، ونضرب هذا المربع فى عدد المفردات التى يمثلها المتوسط . ونجد الخطوات موضحة فى الجدول الآتى :

جدول ٤٧ — حساب الانحراف المعيارى للمتوسطات

المتوسط الصادى	تكرار العمود ك	الانحراف عن ٢٧ ، أى ع	ع^2	ك . ع^2
٢١ر٨١٢٥	٨٠	— ٥ر١٨٧٥	٢٦ر٩١٠٢	٢١٥٢ر٨١٢٥
٢٥ر٠٠٠٠	٢٤٠	— ٢ر٠٠	٤ر٠٠٠٠	٩٦٠ر٠٠٠٠
٢٨ر٤٣١٠	١٤٥	١ر٤٣١٠	٢ر٠٤٧٨	٢٩٦ر٩٢٥٣
٣١ر٦٣٣٣	٧٥	٤ر٦٣٣٣	٢١ر٤٦٧٥	١٦١٠ر٠٦٠١
٣٢ر٣٧٥٠	٤٠	٥ر٣٧٥٠	٢٨ر٨٩٠٦	١١٥٥ر٦٢٥٠
٣٣ر١٦٦٦	١٥	٦ر١٦٦٦	٣٨ر٠٢٧٠	٥٧٠ر٤٠٤٣
٣٣ر٥٠٠٠	٥	٦ر٥٠٠٠	٤٢ر٢٥٠٠	٢١١ر٢٥٠٠
	٦٠٠			٦٩٥٧ر٠٧٧٧

∴ ٦٠٠ ع$^2_م$ = ٦٩٥٧ر٠٧٧٧ ،

و ع$_م$ = ٣ر٤٠٥٢ .

∴ ى = $\frac{٣ر٤٠٥٢}{٧ر٣٢٠٦}$ ،

= ٠ر٤٦٥

ولو حسبنا معامل الارتباط بين س و ص نجد أنه يساوى ٠ر٤٥٠ ومن ذلك يظهر لنا أن ى أكبر من ر . وهذا هو الواجب أن يكون ، حيث قد رأينا أن خط الانحدار فى هذه المسألة غير مستقيم . أى أننا لو كنا وفقنا خطين للقيم المعطاة فى المسألة ، أحدهما مستقيم معادلته من الدرجة الأولى ، والآخر منحن معادلته من الدرجة الثانية ، كنا نجد تشتت النقط حول المستقيم أكبر من تشتتها

حول الخط الآخر . وعلى ذلك فمقياس الارتباط المبنى على فرض أن الانحدار مستقيم ، وهو معامل الارتباط ر ، يكون أصغر من المقياس الآخر الذى لا يستند على هذا الفرض الخطأ .

نسبة الارتباط أعم من المعامل ر

٢٦٣ — ولو كان الانحدار فى هذه المسألة مستقيما ، كنا نجد أن نسبة الارتباط تساوى معامل الارتباط . وعلى ذلك نقول إن نسبة الارتباط ى مقياس أعم من معامل الارتباط ر ، وتعطينا نتائج أدق فى الأحوال التى يخفق فيها معامل الارتباط فى قياس العلاقة بين المتغيرين وتصويرها على حقيقتها . ويمكننا إذن الاعتماد على نسبة الارتباط ى فى جميع الأحوال ؛ بخلاف معامل الارتباط لا نستعمله إلا إذا علمنا أن الانحدار بين المتغيرين مستقيم ، أو قريب جداً من الاستقامة .

خط انحدار س على ص

٢٦٤ — يلاحظ أننا قصرنا الكلام فيما سبق على خط انحدار المتغير ص على المتغير س ، وبحثنا فى توفيق معادلات خطوط الانحدار على صورة تعطى المتغير ص منفرداً فى أحد طرفى المعادلة ، وقوى س وحدها فى الطرف الآخر ، كما فى بند ٢٥٦ مثلا حيث :

ص = — ٠٩٠٢ر س^2 + ٧١٦٤ر س + ١١ر٤٨٣٩

أى أننا اعتبرنا س كمتغير مستقل واعتبرنا ص كمتغير تابع له . وحسبنا المقاييس ر أو ط أو ى على هذا الأساس ، فحسبنا التشتت حول خط انحدار ص على س .

أما إذا كانت معادلة خط الانحدار تعطى س فى صورة متغير تابع إلى ص ، أى خط انحدار س على ص ، فإننا نحسب سه^2 بدل سه$'^2$ ، أى الخطأ المعيارى لقيم س

بدل الخطأ المعيارى لقيم ص ؛ ونستعمل ع بدل عَ ، وع$_{\text{مس}}$ بدل ع$_{\text{مس}}$: فنحسب مقياس الارتباط من أحدى المعادلات الآتية :

$$\text{ر}^{٢} \text{ أو ط}^{٢} = ١ - \frac{\text{سـ}^{٢}}{\text{ع}^{٢}}$$

و

$$\text{ى}^{٢} = ١ - \frac{\text{سـ}^{٢}_{\text{مس}}}{\text{ع}^{٢}}$$

أى

$$\text{ى} = \frac{\text{ع}_{\text{مس}}}{\text{ع}}$$

حيث ع$_{\text{مس}}$ هى الانحراف المعيارى للمتوسطات السينية .

كل مقاييس الارتباط تبنى على أساس التشتت حول خط الانحدار

٢٦٥ — والخلاصة التى نخرج بها من هذا الباب والباب السابق ، هي أن العلاقة بين كميتين متغيرتين يمكن قياسها بعدة مقاييس أهمها ثلاثة ، وهى معامل الارتباط ر ، ودليل الارتباط ط ، ونسبة الارتباط ى . وأن الاختيار بين هذه الثلاثة يتوقف على نوع العلاقة بين المتغيرين ، أو بالأحرى على شكل خط الانحدار . والنقطة المهمة فى هذا الموضوع هى أن الفكرة الأساسية فى هذه المقاييس الثلاثة أصلها واحد فى كل الأحوال، وهو مقدار تشتت القيم المعطاة للمتغيرين حول خط العلاقة المتوسطة بينهما ، أى خط الانحدار .

المراجع

Bowley, A. L., *Elements of Statistics*, Part II. Chapters VI., VII.

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

الباب العاشر

الارتباط المتعدد والارتباط الجزئى

---

٢٦٦ — فى البابين الثامن والتاسع درسنا الارتباط بين ظاهرتين فقط ، واستنبطنا مقاييس مختلفة لقياس العلاقة بين المتغيرين ، على فرض أنهما لا يتأثران بأى عوامل أخرى خارجية عنهما .

ولكن هذا الفرض قليلا ما يتحقق عملياً ، خصوصاً فى المسائل الاقتصادية والاجتماعية ، وفى المسائل العلمية البحتة أيضاً ، حيث نرى الظاهرة التى نبحثها تتأثر بعوامل كثيرة بدرجات مختلفة . ولدراسة العلاقة بين هذه الظاهرة والظواهر الأخرى مجتمعةً أو منفردةً ، يجب علينا التفكير فى مقاييس للارتباط تأخذ فى الحسبان هذه الظروف الجديدة ، التى لم نعرها اهتماماً من قبل .

وسنشرح فى هذا الباب باختصار مقاييس الارتباط التى يمكن استعمالها فى مثل هذه المسائل . وسنقتصر على دراسة الارتباط بين الظواهر التى يمكن قياسها رقمياً .

الارتباط المتعدد والارتباط الجزئي

٢٦٧ — لنفرض أننا ندرس العلاقة بين ظاهرة مثل كمية محصول الفدان (من القمح أو الأذرة أو . . .) وظواهر أخرى مثل كمية السماد المستعمل وكمية مياه الرى ومتوسط درجة حرارة الجو أثناء النمو . كل واحد من هذه العوامل الأخيرة يؤثر فى كمية المحصول ، وكل منها يتغير مستقلا عن العوامل الأخرى ،

وكمية المحصول تتغير تبعاً لهذه العوامل . أى أن كمية السماد (س) وكمية مياه الرى (ص) ودرجة الحرارة (ع) كلها متغيرات مستقلة ، وكمية المحصول (م) متغير تابع لها جميعاً .

إذا أردنا قياس العلاقة بين كمية المحصول م والظواهر الأخرى س وص وع مجتمعةً ، أى أنها تتغير جميعاً فى وقت واحد بدون قيد ولا شرط ، فنستخدم لذلك مانسميه [1] **معامل الارتباط المتعدد**. أما العلاقة بين كمية المحصول م وأحد العوامل الأخرى — مثل س — منفرداً ، أى بفرض أن العاملين الآخرين ص وع يبقيان ثابتين أثناء تغير س وم ، فتسمى [2] **الارتباط الجزئى** ؛ وهذا نقيسه باستخدام **معامل الارتباط الجزئى** .

عمل من التجربة على قيم متناظرة للتغيرات

٣٦٨ — لدراسة الارتباط المتعدد وقياسه عملياً ، نبدأ كالمعتاد بمشاهدة الظواهر المتغيرة التى نبحثها فى عدة ظروف أو حالات مختلفة . وفى كل حالة ندون قيمة كل من هذه المتغيرات الأربعة م و س و ص و ع ؛ فنحصل بذلك على جدول به أربع سلسلات من القيم المتناظرة لهذه المتغيرات . فإذا أجرينا عدة تجارب فى حقول مختلفة (أو سنين مختلفة) على هذا المحصول ، وكانت قيم كل من س و ص و ع تختلف بعضها عن بعض من حقل إلى آخر ، وكذلك تتغير م تبعاً لذلك ، حصلنا لكل من هذه المتغيرات على قيم بعدد حقول التجارب التى لدينا ، وليكن ن مثلا .

توفق معادلة تربط هذه المتغيرات

من هذه القيم المشاهدة للمتغيرات م و س و ص و ع ، نوفق معادلة جبرية تمثل العلاقة بين م كمتغير تابع وبين س و ص و ع كمتغيرات مستقلة ، أحسن

---

(١) بالانجليزية (Coefficient of Multiple Correlation)

(٢) بالانجليزية (Coefficient of Partial Correlation)

ما يمكن . وذلك بطريقة المربعات الصغرى أو بأى طريقة أخرى .

مثال معادلة من الدرجة الأولى

٣٦٩ — نأخذ على سبيل التمثيل الحالة البسيطة التى فيها تكون المعادلة بين م و س و ص و ع من الدرجة الأولى . ولنفرض أننا وفقنا هذه المعادلة بطريقة المربعات الصغرى فكانت كما يأتى :

م = ا س + ب ص + حـ ع + ى ، (١)

حيث ا و ب و حـ و ى كميات عددية ثابتة مستقلة عن المتغيرات م و س و ص و ع . وهذه الكميات نختارها بحيث تجعل هذه المعادلة أصلح من أى معادلة أخرى لتمثيل القيم المعطاة . وهذه الكميات نستخرجها ، كما نعلم فى طريقة توفيق المنحنيات ، من المعادلات[1] الآتية الآتية :

مجـ م = ا . مجـ س + ب . مجـ ص + حـ . مجـ ع + ن ى (٢) ،

مجـ م س = ا . مجـ س^{٢} + ب . مجـ س ص + حـ . مجـ ع س + ى . مجـ س (٣) ،

مجـ م ص = ا . مجـ س ص + ب . مجـ ص^{٢} + حـ . مجـ ع ص + ى . مجـ ص (٤) ،

مجـ م ع = ا . مجـ س ع + ب . مجـ ص ع + حـ . مجـ ع^{٢} + ى . مجـ ع (٥) .

من هذه المعادلات نستخرج قيم ا و ب و حـ و ى ؛ وبذلك نحصل على المعادلة (١) التى تربط المتغير م بالمتغيرات س و ص و ع . وهذه المعادلة هى معادلة « خط » العلاقة المتوسطة بين م والثلاثة المتغيرات مجتمعةً .

---

(١) يلاحظ أن كيفية تركيب هذه المعادلات هى كما يأتى : نضرب المعادلة (١) فى معامل ى فيها ثم نجمعها لجميع الحالات تنتج المعادلة (٢) ؛ ثم نضرب (١) فى معامل ا أى فى س ، ثم نجمع فنحصل على المعادلة (٣) . ثم نضرب (١) فى معامل ب أى ص ونجمع فنحصل على (٤) . وأخيراً نضرب (١) فى معامل حـ ، أى فى ع ، ونجمع فنحصل على (٥) . وهذه الطريقة تتمشى مع الطريقة التى شرحناها سابقا .

حساب الخطأ المعيارى لمعادلة العلاقة المتوسطة

٢٧٠ — نحسب تشتت القيم المعطاة ، أو الخطأ المعيارى لقيم م ، المشاهدة حول هذا « الخط » كما فعلنا فى الارتباط العادى فى الباب السابق . لذلك نحسب انحرافات قيم م عن هذه المعادلة . فإذا فرضنا أن ح_1 هو خطأ القيمة م_1 أو انحرافها عن هذه المعادلة :

$$\therefore \text{ح}_1 = \text{ا}\,\text{س}_1 + \text{ب}\,\text{ص}_1 + \text{حـ}\,\text{ع}_1 + \text{ء} - \text{م}_1 \qquad (٦)$$

ونضرب طرفى هذه المعادلة فى الكميات م_1 و س_1 و ص_1 و ع_1 و ح_1 على التوالى .

$$\therefore \text{ح}_1\text{م}_1 = \text{ا}\,.\,\text{م}_1\text{س}_1 + \text{ب}\,.\,\text{م}_1\text{ص}_1 + \text{حـ}\,.\,\text{م}_1\text{ع}_1 + \text{ء}\,.\,\text{م}_1 - \text{م}_1^2 \quad ،\ (٧)$$

$$،\ \text{ح}_1\text{س}_1 = \text{ا}\,.\,\text{س}_1^2 + \text{ب}\,.\,\text{س}_1\text{ص}_1 + \text{حـ}\,.\,\text{ع}_1\text{س}_1 + \text{ء}\,.\,\text{س}_1 - \text{م}\,\text{س}_1 \quad ،\ (٨)$$

$$،\ \text{ح}_1\text{ص}_1 = \text{ا}\,.\,\text{س}_1\text{ص}_1 + \text{ب}\,.\,\text{ص}_1^2 + \text{حـ}\,.\,\text{ع}_1\text{ص}_1 + \text{ء}\,.\,\text{ص}_1 - \text{م}_1\text{ص}_1 \quad ،\ (٩)$$

$$،\ \text{ح}_1\text{ع}_1 = \text{ا}\,.\,\text{س}_1\text{ع}_1 + \text{ب}\,.\,\text{ص}_1\text{ع}_1 + \text{حـ}\,\text{ع}_1^2 + \text{ء}\,.\,\text{ع}_1 - \text{م}_1\text{ع}_1 \quad ،\ (١٠)$$

$$،\ \text{ح}_1^2 = \text{ا}\,.\,\text{ح}_1\text{س}_1 + \text{ب}\,.\,\text{ح}_1\text{ص}_1 + \text{حـ}\,.\,\text{ح}_1\text{ع}_1 + \text{ء}\,\text{ح}_1 - \text{م}_1\text{ح}_1 \quad ،\ (١١)$$

ولو جمعنا هذه المعادلات بالنسبة إلى جميع قيم المتغير م

$$\therefore \text{مجـ}\,\text{ح} = \text{ا}\,.\,\text{مجـ}\,\text{س} + \text{ب}\,.\,\text{مجـ}\,\text{ص} + \text{حـ}\,.\,\text{مجـ}\,\text{ع} + \text{ن}\,\text{ء} - \text{مجـ}\,\text{م}\ ،$$

$$= \text{صفراً}\ . \qquad \text{انظر المعادلة (٢) أعلاه ؛}$$

$$،\ \text{مجـ}\,\text{ح}\,\text{م} = \text{ا}\,.\,\text{مجـ}\,\text{م}\,\text{س} + \text{ب}\,.\,\text{مجـ}\,\text{م}\,\text{ص} + \text{حـ}\,.\,\text{مجـ}\,\text{م}\,\text{ع} + \text{ء}\,.\,\text{مجـ}\,\text{م} - \text{مجـ}\,\text{م}^2 \quad ،\ (١٢)$$

$$\text{و}\ \text{مجـ}\,\text{ح}\,\text{س} = \text{ا}\,.\,\text{مجـ}\,\text{س}^2 + \text{ب}\,.\,\text{مجـ}\,\text{س}\,\text{ص} + \text{حـ}\,.\,\text{مجـ}\,\text{ع}\,\text{س} + \text{ء}\,.\,\text{مجـ}\,\text{س} - \text{مجـ}\,\text{م}\,\text{س}$$

$$= \text{صفراً} \qquad \text{انظر معادلة (٣).}$$

وبالمثل $\text{مجـ}\,\text{ح}\,\text{ص} = \text{مجـ}\,\text{ح}\,\text{ع} = \text{صفراً}$.

$$\therefore \text{مجـ}\,\text{ح}^2 = \text{ا}\,.\,\text{مجـ}\,\text{ح}\,\text{س} + \text{ب}\,.\,\text{مجـ}\,\text{ح}\,\text{ص} + \text{حـ}\,.\,\text{مجـ}\,\text{ح}\,\text{ع} + \text{ء}\,.\,\text{مجـ}\,\text{ح} - \text{مجـ}\,\text{م}\,\text{ح}\ .$$

$$= -\,\text{مجـ}\,\text{م}\,\text{ح}\ ، \qquad (١٣)\ ،$$

ولو وضعنا $\text{مجـ}\,\text{ح}^2 = \text{ن}\,\text{سه}_{\text{م}}^2$ ، ووضعنا $\text{ع}_{\text{م}}$ للانحراف المعيارى لقيم المتغير م ، ورمزنا للوسط الحسابى لها بالحرف $\bar{\text{م}}$

$$\therefore \text{مجـ}\,\text{م}^2 = \text{ن}\,\text{ع}_{\text{م}}^2 + \text{ن}\,\bar{\text{م}}^2$$

$$\therefore \text{ن}\,\text{سه}_{\text{م}}^2 = -\,\text{مجـ}\,\text{م}\,\text{ح}$$

$$= \text{مجـ}\,\text{م}^2 - \text{ا}\,.\,\text{مجـ}\,\text{م}\,\text{س} - \text{ب}\,.\,\text{مجـ}\,\text{م}\,\text{ص} - \text{حـ}\,.\,\text{مجـ}\,\text{م}\,\text{ع} - \text{ء}\,.\,\text{مجـ}\,\text{م}\ .$$

معامل الارتباط المتعدد

٢٧١ — وقياساً على تعريف الارتباط بدلالة الخطأ المعيارى سه الذى استخدمناه فى الباب السابق ، يكون معامل الارتباط المتعدد بين م والمتغيرات س و ص و ع مجتمعة هو $\text{ر}_{\text{م.س ص ع}}$ ، حيث

$$\text{ر}^2_{\text{م.س ص ع}} = ١ - \frac{\text{سه}_{\text{م}}^2}{\text{ع}_{\text{م}}^2}$$

$$\therefore \text{ر}^2_{\text{م.س ص ع}} = \frac{\text{ا}\,.\,\text{مجـ}\,\text{م}\,\text{س} + \text{ب}\,.\,\text{مجـ}\,\text{م}\,\text{ص} + \text{حـ}\,.\,\text{مجـ}\,\text{م}\,\text{ع} + \text{ء}\,.\,\text{مجـ}\,\text{م} - \text{ن}\,\bar{\text{م}}^2}{\text{ن}\,.\,\text{ع}_{\text{م}}^2}$$

وبذا نحصل على معامل الارتباط المتعدد المطلوب ؛ وهو يقيس العلاقة بين

المتغير م والمتغيرات الأخرى كلها مجتمعةً . أى أنه يقيس لنا درجة تأثير المتغيرات س و ص و ع مع بعضها فى كمية المحصول م .

تعريف معامل الارتباط الجزئى

٢٧٢ — بقى أمامنا أن نعالج المسألة الأخرى ، وهى قياس العلاقة بين المتغير م وأحد المتغيرات الأخرى ، على فرض بقاء المتغيرين الآخرين ثابتين . فنريد مثلا معرفة معامل الارتباط بين كمية المحصول وكمية السماد المستعمل ، على فرض أن كمية مياه الرى ودرجة الحرارة لا تتغيران . هذا المعامل يسمى **معامل الارتباط الجزئى أو الصافى** [١]

طريقة استبعاد العوامل الغريبة

٢٧٣ — يمكننا إيجاد معامل الارتباط الجزئى بأن ننتقى من نتائج التجارب التى نجريها جميع الحالات التى فيها كمية مياه الرى ثابتة ، وكذلك درجة حرارة الجو ، ولا يتغير إلا كمية السماد المستعمل وكمية المحصول . ومن هذه الحالات نحسب معامل الارتباط الجزئى من واقع الأرقام التى تدل على قيم م و س فى هذه الحالات دون غيرها .

وهذا الإجراء سهل فى بعض المسائل ؛ والواقع أن هذا هو المتبع عادة فى التجارب الزراعية . فى هذه المسألة مثلا يمكننا تثبيت عاملى الرى ودرجة الحرارة بغاية السهولة بأن نأخذ حقول التجارب كلها فى منطقة زراعية واحدة ، فنضمن بذلك تساوى درجة الحرارة فى كل منهـا ؛ ونروى كل الحقول عدداً معيناً من المرات وفى نفس المواعيد وبنفس المقدار ، فنضمن بذلك تساوى كمية مياه الرى فى كل الحقول . ونغير كمية السماد الموضوع فى هذه الحقول ، ونلاحظ المحصول فى كل منها . وبذلك يمكننا استبعاد تأثير جميع العوامل ويبقى عامل السماد وحده

(١) بالانجليزية (Coefficient of Partial or Net Correlation)

الذى يؤثر فى كمية المحصول . ونضمن حينئذ أن كل تغيير فى كمية المحصول يكون ناشئاً عن تغير مناظر له فى كمية السماد . فنحسب معامل الارتباط بين م و س من واقع الأرقام التى نحصل عليها من هذه الحقول ؛ ويكون هذا المعامل هو معامل الارتباط الجزئى بين كمية المحصول وكمية السماد ، بفرض أن الرى ودرجة الحرارة لم يتغيرا .

استبعاد العوامل الغريبة ممكن فى العلوم الطبيعية ، ومستحيل فى المسائل الاجتماعية

٢٧٤ — وهذه الطريقة — طريقة عزل العامل الذى نريد دراسته واستبعاد العوامل الأخرى من المجال — هى الطريقة العلمية المتبعة فى أبحاث العلوم البحتة مثل الكيمياء والطبيعة وعلوم الحياة وغيرها . وهى سليمة طبعاً من الناحية المنطقية والنظرية ؛ ولا غبار عليها من الناحية العملية فهى سهلة وميسورة فى تلك العلوم ، ولو أن استبعاد العوامل الغريبة ومحو أثرها غالباً ما يكلف الباحث مجهوداً كبيراً ونفقات طائلة . ولكن هذة الطريقة غير ميسورة بل مستحيلة فى دراسة الظواهر الاقتصادية والاجتماعية ، حيث إن هذه الظواهر تكون دائماً متأثرة بعوامل شتى ، متشابكة مع بعضها لا يمكن عزلها أو استبعاد بعضها . ولذلك يتعين على الباحث الاقتصادى أن يفكر فى استنباط طريقة أخرى لا تعتمد على استبعاد أثر العوامل الغريبة بتثبيتها أثناء إجراء التجربة .

حاصل ضرب معاملى الانحدار

٢٧٥ — رأينا فى بند ٢٤٤ (صفحة ٢٧٤) أن حاصل ضرب معامل انحدار ص على س[١] ، فى معامل انحدار س على ص يساوى $\text{ر}^{٢}$ ، أى مربع معامل الارتباط بينهما . وقياساً على هذا يمكننا إيجاد معامل الارتباط الجزئى الذى نحن بصدده .

(١) يلاحظ أنه فى تعريف معامل انحدار ص على س فى صفحة ٢٧١ سقط حرف وصحته :

$$\text{ر}\,\frac{\bar{\text{ع}}}{\text{ع}} = \text{معامل انحدار ص على س}$$

نفرض أن لدينا المتغير م والمتغيرات الثلاثة س و ص و ع كما فى المسألة السابقة . ولتكن معادلة « خط » العلاقة المتوسطة بين هذه المتغيرات هى :

$$م = ا_{٢١}س + ب_{٣١}ص + ح_{٤١}ع + ء_{١} \quad \ldots \quad (١)$$

حيث $ا_{٢١}$ ، $ب_{٣١}$ ، $ح_{٤١}$ ، $ء_{١}$ كميات ثابتة معينة .

هذه المعادلة هى فى الواقع معادلة انحدار بين م والمتغيرات الأخرى ، كما سبق أن قلنا من قبل ؛ فلو ثبتنا ص و ع وجعلنا س فقط تتغير و م تبعاً لها ، كما هو مفروض فى تعريف الارتباط الجزئى ، آلت هذه المعادلة إلى الصورة :

$$م = ا_{٢١}س + \text{كمية ثابتة} \quad \ldots \quad (٢)$$

وهذه هى معادلة خط العلاقة المتوسطة بين م و س ، أى خط انحدار م على س . وتكون $ا_{٢١}$ إذن هى معامل انحدار المتغير م على المتغير س ، باعتبار أن س متغير مستقل و م متغير تابع له .

نوفق معادلة أخرى تعبر عن المتغير س بدلالة المتغيرات م ، ص ، ع ، من واقع القيم المعطاة لنا من التجربة ؛ وذلك بنفس الطريقة التى اتبعناها فى توفيق المعادلة (١) بند ٢٦٩ . ولتكن هذه المعادلة الجديدة هى :

$$س = ا_{١٢}م + ب_{٣٢}ص + ح_{٤٢}ع + ء_{٢} \quad \ldots \quad (٣) ؛$$

حيث كل من $ا_{١٢}$ ، $ب_{٣٢}$ ، $ح_{٤٢}$ ، $ء_{٢}$ كمية ثابتة مستقلة عن المتغيرات س ، م ، ص ، ع .

إذا ثبتنا ص ، ع وجعلنا م ، س فقط تتغيران ، تؤول هذه المعادلة إلى معادلة خط انحدار س على م ، وتكون $ا_{١٢}$ هى أيضاً معامل انحدار س على م .

وهكذا نحصل على معاملى الانحدار بين المتغيرين م و س ، كل على الآخر ، باعتبار المتغيرين الآخرين ص و ع ثابتين . وهذان العاملان هما $ا_{٢١}$ و $ا_{١٢}$ ،

ويكون معامل الارتباط الجزئى بين م و س ، بفرض ص و ع ثابتين ، هو[١] $ر_{٢١.٣٤}$ حيث :

$$ر^{٢}_{٢١.٣٤} = ا_{٢١} . ا_{١٢} \quad ، \ldots \quad (٤) .$$

وكذلك لإيجاد معامل الارتباط الجزئى بين م و ص بفرض س و ع ثابتين ، نوفق معادله تعطى ص بدلالة م و س و ع ولتكن هى :

$$ص = ا_{٣٢}س + ب_{١٣}م + ح_{٤٣}ع + ء_{٣} \quad . \quad (٥) .$$

ويكون إذن معامل الارتباط الجزئى بين م و ص هو $ر_{٣١.٢٤}$ حيث :

$$ر^{٢}_{٣١.٢٤} = ب_{٣١} . ب_{١٣} \quad (٦) .$$

وأخيراً نوفق معادلة بين ع والمتغيرات الأخرى ولتكن هى

$$ع = ا_{٣٤}س + ب_{٣٤}ص + ح_{١٤}م + ء_{٤} \quad (٧) ،$$

ويكون معامل الارتباط الجزئى بين م و ع هو $ر_{٤١.٢٣}$ ، حيث :

$$ر^{٢}_{٤١.٢٣} = ح_{١٤} \, ح_{٤١} \quad (٨) .$$

استنباط المعاملات ر فى حالة خاصة

٢٧٦ — نأخذ حاله خاصة ونوجد معامل الارتباط الجزئى لتوضيح الخطوات :

لنفرض للسهولة أن لدينا متغيرين مستقلين فقط وهما س و ص ، ومتغيراً ثالثاً تابعاً لهما وهو م . وأن لدينا ن من القيم لكل من هذه المتغيرات الثلاثة ، وهى على التوالى :

---

(١) يدل الرمز المرقوم $ر_{٢١.٣٤}$ على معامل الارتباط الجزئي بين المتغيرين الأول والثاني (م و س) بفرض أن الثالث والرابع (ص و ع) ثابتان ، وهكذا .

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

$$\text{و} \quad \text{ص}_{١} \quad \text{ص}_{٢} \quad \text{ص}_{٣} \quad . \; . \; . \quad \text{ص}_{\text{ن}} \text{ ،}$$

$$\text{و} \quad \text{م}_{١} \quad \text{م}_{٢} \quad \text{م}_{٣} \quad . \; . \; . \quad \text{م}_{\text{ن}} \text{ .}$$

ولنفرض ، للسهولة أيضاً ، أن هذه القيم مقيسة ابتداء من الوسط الحسابي لكل منها ؛ أي أن

$$\text{مجـ س} = \text{مجـ ص} = \text{مجـ م} = \text{صفراً ؛}$$

$$\text{وكذلك } \bar{\text{س}} = \bar{\text{ص}} = \bar{\text{م}} = \text{صفراً ،}$$

∴ تكون الانحرافات المعيارية لهذه المتغيرات هي $\text{ع}_{١}$ و $\text{ع}_{٢}$ و $\text{ع}_{٣}$ حيث

$$\text{ن ع}_{١}^{٢} = \text{مجـ م}^{٢} \text{ ، } \text{ن ع}_{٢}^{٢} = \text{مجـ س}^{٢} \text{ ، } \text{ن ع}_{٣}^{٢} = \text{مجـ ص}^{٢} \text{ .}$$

ويكون أيضاً معامل الارتباط بين م و س مثلا يساوي $\text{ر}_{٢١}$ ، حيث

$$\text{ر}_{٢١} = \frac{\text{مجـ م س}}{\text{ن ع}_{١} \text{ع}_{٢}} \quad . \; . \; . \; . \quad \text{معادلة (١) بند ٢٠٣ ،}$$

نوفق معادلة بين م و س و ص من الدرجة الأولى ، من واقع هذه القيم المعطاة . ولتكن هذه المعادلة هي .

$$\text{م} = \text{ا س} + \text{ب ص} + \text{حـ} \quad . \; . \; . \quad \text{(١) ،}$$

حيث كل من ا و ب و حـ كمية ثابتة ، تستخرج من المعادلات المعروفة في توفيق المنحنيات وهي :

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

$$\text{و مجـ س م} = \text{ا مجـ س}^{٢} + \text{ب مجـ س ص} + \text{حـ . مجـ س} \quad \text{(٣) ،}$$

$$\text{و مجـ ص م} = \text{ا . مجـ س ص} + \text{ب مجـ ص}^{٢} + \text{حـ . مجـ ص} \quad \text{(٤) ،}$$

وبوضع $\text{مجـ س} = \text{مجـ ص} = \text{مجـ م} = .$ ، ينتج أن $\text{حـ} = .$ ؛

وبالتعويض عن $\text{مجـ س}^{٢}$ و $\text{مجـ ص}^{٢}$ و $\text{مجـ م}^{٢}$ بقيمها ، ووضع :

$$\text{مجـ م س} = \text{ن ر}_{٢١} \text{ . ع}_{١} \text{ع}_{٢}$$

$$\text{و} \quad \text{مجـ م ص} = \text{ن ر}_{٣١} \text{ . ع}_{١} \text{ع}_{٣}$$

$$\text{و} \quad \text{مجـ س ص} = \text{ن ر}_{٣٢} \text{ . ع}_{٢} \text{ع}_{٣}$$

$$\therefore \quad \text{ر}_{٢١} \text{ . ع}_{١} = \text{ا . ع}_{٢} + \text{ب . ر}_{٣٢} \text{ع}_{٣} \quad \text{(٥) ،}$$

$$\text{و} \quad \text{ر}_{٣١} \text{ . ع}_{١} = \text{ا . ر}_{٣٢} \text{ع}_{٢} + \text{ب . ع}_{٣} \quad \text{(٦) ،}$$

$$\therefore \quad \text{ع}_{١} (\text{ر}_{٢١} - \text{ر}_{٣١} \text{ . ر}_{٣٢}) = \text{ا} (١ - \text{ر}_{٣٢}^{٢}) \text{ع}_{٢}$$

$$\text{و} \quad \text{ع}_{١} (\text{ر}_{٣١} - \text{ر}_{٢١} \text{ . ر}_{٣٢}) = \text{ب} (١ - \text{ر}_{٣٢}^{٢}) \text{ع}_{٣}$$

∴ معادلة انحدار م على س و ص ، هي

$$\frac{\text{م}}{\text{ع}_{١}} = \frac{\text{ر}_{٢١} - \text{ر}_{٣١} \text{ . ر}_{٣٢}}{١ - \text{ر}_{٣٢}^{٢}} \text{ . } \frac{\text{س}}{\text{ع}_{٢}} + \frac{\text{ر}_{٣١} - \text{ر}_{٢١} \text{ . ر}_{٣٢}}{١ - \text{ر}_{٣٢}^{٢}} \text{ . } \frac{\text{ص}}{\text{ع}_{٣}} \text{ ، } . \; . \; . \quad \text{(٧) .}$$

وبمثل ذلك نوفق معادلة انحدار س على م و ص ، وهي :

$$\frac{\text{س}}{\text{ع}_{٢}} = \frac{\text{ر}_{٢١} - \text{ر}_{٣١} \text{ . ر}_{٣٢}}{١ - \text{ر}_{٣١}^{٢}} \text{ . } \frac{\text{م}}{\text{ع}_{١}} + \frac{\text{ر}_{٣١} - \text{ر}_{٢١} \text{ . ر}_{٣٢}}{١ - \text{ر}_{٣١}^{٢}} \text{ . } \frac{\text{ص}}{\text{ع}_{٣}} \quad . \; . \; . \quad \text{(٨) .}$$

ومن هاتين المعادلتين نرى أن معاملي انحدار م على س ، س على م هما على التوالي :

$$\frac{\text{ع}_{١}}{\text{ع}_{٢}} \text{ . } \frac{\text{ر}_{٢١} - \text{ر}_{٣١} \text{ . ر}_{٣٢}}{١ - \text{ر}_{٣٢}^{٢}} \text{ ، } \frac{\text{ع}_{٢}}{\text{ع}_{١}} \text{ . } \frac{\text{ر}_{٢١} - \text{ر}_{٣١} \text{ . ر}_{٣٢}}{١ - \text{ر}_{٣١}^{٢}} \text{ .}$$

وينتج إذن أن معامل الارتباط الجزئي بين م و س هو الوسط الهندسي لهذين المعاملين

$$\therefore \quad \text{ر}_{٢١.٣} = \frac{\text{ر}_{٢١} - \text{ر}_{٣١} \text{ر}_{٣٢}}{\sqrt{(١ - \text{ر}_{٣١}^{٢})(١ - \text{ر}_{٣٢}^{٢})}} \quad . \; . \; . \; . \; . \quad \text{(٩) ،}$$

حيث $\text{ر}_{\text{٢١}}$ ، $\text{ر}_{\text{٣١}}$ ، $\text{ر}_{\text{٣٢}}$ هى معاملات الارتباط العادية بين م و س ، م و ص ، س و ص على التوالى .

وهكذا أمكننا الحصول على معامل الارتباط الجزئى بين متغيرين (م و س) بفرض ثبات الثالث (ص) ، بدلالة معاملات الارتباط العادية بين أزواج بسيطة من المتغيرات .

معامل الارتباط الجزئى من رتب عليا

٢٧٧ — وإذا كان لدينا عدد من المتغيرات أكبر من ثلاثة ، وأردنا معامل الارتباط الجزئى بين اثنين منها مع تثبيت المتغيرات الباقية ، نستخرج أولا معاملات الارتباط العادية بين أزواج بسيطة من المتغيرات . ومن هذه نستنبط معامل الارتباط بين متغيرين مع تثبيت واحد غيرهما ، باستخدام العلاقة (٩) من البند السابق . نسمى هذا معاملاً من الرتبة الأولى مثلا $\text{ر}_{\text{٢١.٣}}$ ، حيث يوجد متغير واحد ثابت ، وهو الثالث المعبر عنه بالرقم ٣ بعد النقطة .

وبواسطة معاملات الرتبة الأولى نستنبط معاملات الرتبة الثانية ، وهى معاملات الارتباط الجزئى مع تثبيت متغيرين . مثلا $\text{ر}_{\text{٢١.٣٤}}$ هو معامل الارتباط الجزئى بين المتغيرين الأول والثانى ، بفرض تثبيت الثالث والرابع .

ومن هذه نستنبط معاملات الرتبة الثالثة وهكذا . والعلاقة بين معاملات الرتب المتتالية يمكن وضعها فى الصورة الآتية [١] :

$$\text{ر}_{\text{٢١.٣٤}} = \frac{\text{ر}_{\text{٢١.٣}} - \text{ر}_{\text{٤١.٣}} \cdot \text{ر}_{\text{٤٢.٣}}}{\sqrt{(١-\text{ر}^{٢}_{\text{٤١.٣}})(١-\text{ر}^{٢}_{\text{٤٢.٣}})}} ،$$

$$\text{ر}_{\text{٢١.٣٤٥}} = \frac{\text{ر}_{\text{٢١.٣٤}} - \text{ر}_{\text{١٥.٣٤}} \cdot \text{ر}_{\text{٢٥.٣٤}}}{\sqrt{(١-\text{ر}^{٢}_{\text{١٥.٣٤}})(١-\text{ر}^{٢}_{\text{٢٥.٣٤}})}} ،$$

(١) انظر كتاب G.U.Yule, *Introduction to the Theory of Statistics* (1937) p. 261

وكتاب Mills, C.F. *Statistical Methods*, (1924) p. 508

وعلى العموم يكون :

$$\text{ر}_{\text{٢١.٣٤٥...ن}} = \frac{\text{ر}_{\text{٢١.٣٤...(ن-١)}} - \text{ر}_{\text{ن١.٣٤...(ن-١)}} \cdot \text{ر}_{\text{ن٢.٣٤...(ن-١)}}}{\sqrt{(١-\text{ر}^{٢}_{\text{ن١.٣٤...(ن-١)}})(١-\text{ر}^{٢}_{\text{ن٢.٣٤...(ن-١)}})}}$$

فى حالة علاقة غير خطية نأخذ متغيراً جديداً

٢٧٨ — يلاحظ أننا فرضنا فى هذا الباب أن العلاقة بين المتغير م والمتغيرات س ، ص ، ع علاقة خطية ، أى أنها من الدرجة الأولى . ومن الممكن تعميم هذه الطريقة لتشمل الحالات الأخرى التى تكون فيها العلاقة بين م وبعض المتغيرات ، أو كلها ، غير خطيــة . فلنفرض مثلا أن المتغير س يظهر فى شكل $\text{س}^{٣}$ فى معادلة العلاقة المتوسطة [المعادلة رقم (١) بند ٢٦٩] ؛ أى أنه يظهر فى الدرجة الثالثة بدلاً من الدرجة الأولى . فى هذه الحالة يمكننا أن نستبدل المتغير $\text{س}^{٣}$ بمتغير جديد مثل ل ؛ وتكون قيم ل حينئذ تساوى تكعيب قيم س التى حصلنا عليها من التجربة . ونوفق المعادلة بين المتغير التابع م والمتغيرات المستقلة ل ، ص ، ع بنفس الطريقة المستخدمة فى توفيق المعادلة العادية .

ويجوز طبعاً أن تكون العلاقة بين المتغير الجديد ل والمتغير المستبدل ، بشكل آخر غير ل = $\text{س}^{٣}$ ، مثلا ل = $\text{ش}^{\text{٥د٤}}$ ، أو لو س ، أو $\text{أ}^{\text{س}}$ ، أو . . . ؛ ولكن البحث فى هذه المسائل يخرجنا عن نطاق هذا الكتاب ، حيث يستلزم إلماماً ببعض النظريات الرياضية الصعبة والمعقدة التى لا محل لها هنا .

٢٧٩ — وقد فرضنا ضمناً فى هذا البحث أيضاً ، أن المتغيرات التى نبحثها ذات توزيع تكرارى متماثل . وهذا أيضاً فرض لا يتحقق إلا فى حالات خاصة فقط ، وإنما اخترناه هنا للسهولة .

المراجـــع

Bowley, A. L., *Elements of Statistics*, Part II. Chapter VIII.

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

Rietz., H., *Handbook of Mathematical Statistics*, Chapter IX.

Whittaker and Robinson., *Calculus of Observations*, Chapter XII.

Yule, G. U., *Introduction to the Theory of Statistics*. Chapter XIV.

الباب الحادى عشر

الأرقام القياسية

تعريف الرقم القياسى

٢٨٠ — الرقم القياسى هو عبارة عن رقم نسبى ، أو ملخص لعدة أرقام نسبية ، ينشأ لبيان وقياس الحركة أو التغير فى أى ظاهرة معينة بالنسبة إلى أساس معين . ولتركيب الرقم القياسى لأى ظاهرة ، نكوّن نسبة مئوية بين القيمة المقارنة لهذه الظاهرة والقيمة الأخرى لها ، المعتبرة أساساً للمقارنة . فالرقم القياسى مثلا لسعر القمح هذا العام بالنسبة إلى سعره فى سنة ١٩٢٠ كأساس ، يساوى خارج قسمة السعر الحالى على السعر فى سنة ١٩٢٠ (سنة الأساس) مضروباً فى العدد ١٠٠ . وكذلك الرقم القياسى لسعر القطن فى بورصة مينا البصل هذا الأسبوع بالنسبة إلى سعره فى بورصة ليفربول كأساس ، يساوى خارج قسمة السعر فى مينا البصل على السعر فى ليفربول ، مضروباً فى العدد ١٠٠ أيضاً . وهكذا .

استخدام الأرقام القياسية فى الاقتصاد

٢٨١ — والأرقام القياسية كثيرة الاستعمال فى الأبحاث الاقتصادية جميعها ، وفى غيرها من المسائل العلمية أيضاً . وهى أداة نافعة جداً فى تصوير التغيرات التى تطرأ على الظواهر الاقتصادية المختلفة ، خصوصاً تلك الظواهر المركبة من عدة عوامل متغيرة فى وقت واحد . ومثال ذلك المستوى العمومى للأسعار فهو عبارة عن ملخص لأسعار جميع السلع ؛ وكل سلعة تحيط بها ظروف خاصة بها تعمل على تغيير أسعارها ، وظروف مشتركة تجعلها تتبع الحركة العامة للسوق . فالرقم القياسى لمستوى الأسعار يكوّن لنا فكرة ملخصة ودقيقة وواضحة عن

تغيرات أسعار هذه السلع فى مجموعها . وبدون هذا الرقم لا يمكننا دراسة الحالة العامة للسوق ، ومستوى الأسعار وتأثيرها فى الحالة الاقتصادية للبلد .

والحقيقة أن مسائل الأسعار هى أهم النواحى التى نستخدم فيها الأرقام القياسية ؛ إلا أن هناك عدة مسائل اقتصادية واجتماعية نحتاج فى دراستها إلى استخدام الأرقام القاسية . ومثال ذلك إنشاء رقم قياسى للنشاط الصناعى أو التجارى ، ورقم قياسى للأجور ، ورقم قياسى لمستوى المعيشة ونفقتها ، وغير ذلك من المسائل الهامة . ولكننا سنقتصر الكلام فيما يلى على تركيب الأرقام القياسية للأسعار ؛ والمفهوم طبعاً أن الطرق المستعملة فى مسائل الأسعار تنطبق بذاتها ، مع تغيير بسيط فى معنى الرموز ، على المسائل الأخرى مثل الأجور أو الانتاج ونحوها .

تعيين سنة الأساس

٢٨٢ — يمكننا تركيب أرقام قياسية — لمستوى الأسعار مثلا — على عدة صور ؛ وسنشرح الآن بعض الطرق المختلفة لتركيب الأرقام القياسية . ويجب أن نلاحظ فى مبدأ الأمر أن النتائج التى نحصل عليها بهذه الطرق لن تكون متساوية ، ولو أنها مشتقة من نفس البيانات .

وعند إنشاء أى رقم قياسى لا بد أن نتفق على الأساس الذى سنتخذه لتركيب هذا الرقم القياسى ، فنأخذ مثلا[١] سنة معينة (أو فترة أخرى أكبر أو أصغر من سنة) ونعتبرها أساساً . وهذه نسميها سنة الأساس أو القاعدة . وفى العادة تكون سنة (أو فترة) الأساس سابقة للسنة التى نريد مقارنتها ، ولكن أحياناً يكون المطلوب رقماً قياسياً للأسعار فى سنة ١٩٣٥ مثلا بالنسبة إلى سنة ١٩٣٨ كأساس . ولذلك سنتكلم عن السنة (أو البلد) الأساسية والسنة المقارنة ، وعن الأسعار الأساسية والأسعار المقارنة ، وهكذا .

(١) يصح طبعاً أن يكون الأساس مكاناً معيناً ، كما لو أردنا تكوين الرقم القياسى لمستوى الأسعار فى الأسكندرية بالنسبة إلى القاهرة كأساس . هنا الأساس هو القاهرة ، ولا ذكر للزمن أو لسنة الأساس التى تسمى بالانجليزية (Base Year)

الرموز المستعملة

٢٨٣ — لنفرض أن لدينا عدة سلع يتركب منها الرقم القياسى ، ولتكن الأسعار الأساسية لهذه السلع هى :

$ع_٠ ، ع'_٠ ، ع''_٠ ، ع'''_٠$ ... بعدد السلع الموجودة .

ولتكن الكميات الأساسية لهذه السلع هى على التوالى :

$ك_٠ ، ك'_٠ ، ك''_٠ ، ك'''_٠$ ...

ولنفرض أن الأسعار والكميات المقارنة لهذه السلع هى على التوالى :

$ع_١ ، ع'_١ ، ع''_١ ، ع'''_١$ ...

و $ك_١ ، ك'_١ ، ك''_١ ، ك'''_١$ ...

إذا قسمنا السعر المقارن لأى سلعة على سعرها الأساسى ، وضربنا الناتج فى ١٠٠ ، حصلنا على ما نسميه **منسوب السعر**[١] لهذه السلعة . وإذا رمزنا للمنسوب بالحرف س

$\therefore$ $س = \frac{ع_١}{ع_٠} \times ١٠٠$ ، $س' = \frac{ع'_١}{ع'_٠} \times ١٠٠$ ، وهكذا مع باقي السلع .

ونشرح الآن بعض الأسس التى يمكن أن يبنى عليها تركيب أرقام قياسية للأسعار بالمعنى المقصود فى تعريف الرقم القياسى . ولن نتعرض هنا إلى المفاضلة بين هذه الأسس ، ولا إلى كيفية اختيار أصلحها .

(١) اسمه بالانجليزية (Price Relative) . بعض الناس يستعمل كلمة « السعر النسبي » . انظر الاحصاء السنوى العام لسنة ١٩٣٥ — ١٩٣٦ صفحة ٤٩٠ . ولكن هذه الكلمة فى رأيي لا تؤدى المعنى هنا . لأن هذا ليس سعراً ، بل نسبة .

الرقم التجميعي البسيط

٢٨٤ — الرقم التجميعي البسيط للأسعار[1] . ولتركيب هـذا الرقم نقسم حاصل جمع الأسعار المقارنة على حاصل جمع الأسعار الأساسية لـكل السلع كما هى وبدون مفاضلة بينها أو ترجيح البعض عن البعض . وهذا الرقم هو إذن

$$\frac{ع_{١} + عَ_{١} + عً_{١} + \dots}{ع_{٠} + عَ_{٠} + عً_{٠} + \dots} = \frac{\text{مجـ } ع_{١}}{\text{مجـ } ع_{٠}} \text{ ، } \dots \quad (١)$$

لنأخذ مثلا سلع القطن والقمح والفول والشعير فى سنتى ١٩٣١ و ١٩٣٥ ، ونحسب الرقم القياسى لأسعار هذه المحاصيل فى سنة ١٩٣٥ بالنسبة إلى سنة ١٩٣١ كأساس :

جدول ٤٨ — أسعار[2] محاصيل القطن والقمح والفول والشعير فى سنتى ١٩٣١ و ١٩٣٥

السـنة	القطن بالقنطار	القمح بالأردب	الفول بالأردب	الشعير بالأردب
	جـ م	جـ م	جـ م	جـ م
١٩٣١	٣ر٦٢٥	١ر٢١٩	١ر٣٩٩	ر٦٤٣
١٩٣٥	٣ر٩٦٠	١ر٥٣١	١ر٣٤٣	ر٨٠٢

$$\therefore \frac{\text{مجـ } ع_{١}}{\text{مجـ } ع_{٠}} = \frac{٣ر٦٩٠ + ١ر٥٣١ + ١ر٣٤٣ + ر٨٠٢}{٣ر٦٢٥ + ١ر٢١٩ + ١ر٣٩٩ + ر٦٤٣}$$

$$= \frac{٧ر٦٣٦}{٦ر٨٨٦} \text{ .}$$

(١) بالانجليزية (Simple Aggregative Index)

(٢) انظر الاحصاء السنوى العام ١٩٣٥ — ١٩٣٦ صفحتى ٤٠٩ و ٤١٠ . هذه الأرقام خاصة بمحاصيل زراعة مصلحة الأملاك الأميرية .

$$\therefore \frac{\text{مجـ } ع_{١}}{\text{مجـ } ع_{٠}} \times ١٠٠ = ١١٠ر٨٠ \text{ ،}$$

أى أن مستوى أسـعار هذه المحاصيل ارتفع ٨ر١٠٪ فى سنة ١٩٣٥ عما كان عليه فى سنة ١٩٣١ .

الرقم التجميعى سهل الحساب ولـكنه لا يعطي للسلع أهميتها الحقيقية

٢٨٥ — وهـذا الرقم التجميعى البسيط هو أسهل الأرقام القياسية عملا وتركيباً ، إذ نضع الأسعار على علاتها وبدون تحريف أو تعديل مهما كانت . ولـكن هذه البساطة والسهولة هما فى نفس الوقت عيب يؤخذ عليه . لأن جميع السلع هنا تعامل نفس المعاملة ، بدون ترجيح أو تعزيز بعضها بما يتناسب وأهميتها بجانب غيرها .

هذا فضلا عن أن اختلاف الوحدات المستعملة فى تسعيرات السلع المختلفة ، وما ينشأ عنه من تـكبير أو تصغير السعر $ع_{٠}$ أو $ع_{١}$ ، يعطى بعض السلع أهمية مفتعلة ليست لها . فمثلا إذا كانت ع سعر الخـبز ، وهو ١٥ مليما للأقة ، وكانت عَ سعر الفحم ، وهو ١٨٠ قرشاً للطن ، وجدنا أن ع صغيرة جداً بالنسبة إلى عَ ؛ وبذلك نعطى لسعر الفحم وتغيراته وزناً وأهمية أكبر من سعر الخبز ، الذى هو فى الحقيقة أولى بهذه الأهمية .

الرقم التجميعى المرجـح

٢٨٦ — ولتصحيح هذا العيب فى الرقم التجميعى البسيط ، نرجح أسعار السلع المختلفة بأوزان تتناسب وأهمية هذه السلع . فنستخدم الـكميات المنتجة (أو المستهلكة) من هذه السلع كأوزان . ويصح أن نستعمل الـكميات الأساسية أو الـكميات المقارنة كأوزان ، وبذلك نحصل على صيغتين للرقم التجميعى المرجح ، وهما:

١ — الرقم التجميعى المرجح بالـكميات الأساسية ، وهو :

$$\frac{ع_{١} ك_{٠} + عَ_{١} كَ_{٠} + عً_{١} كً_{٠} + \dots}{ع_{٠} ك_{٠} + عَ_{٠} كَ_{٠} + عً_{٠} كً_{٠} + \dots} = \frac{\text{مجـ } ع_{١} ك_{٠}}{\text{مجـ } ع_{٠} ك_{٠}} \dots (١٢) \text{ ؛}$$

ب — الرقم التجميعى المرجح بالـكميات المقارنة ، وهو :

$$\frac{\text{ع}_1\text{ك}_1 + \text{ع}'_1\text{ك}'_1 + \text{ع}''_1\text{ك}''_1 + \cdots}{\text{ع}_0\text{ك}_1 + \text{ع}'_0\text{ك}'_1 + \text{ع}''_0\text{ك}''_1 + \cdots} = \frac{\text{مج ع}_1\text{ك}_1}{\text{مج ع}_0\text{ك}_1} \quad \ldots (\text{٢ ب}).$$

ويلاحظ فى كلتا الحالتين ا و ب ، أن الأوزان المستعملة فى البسط هي نفسها المستعملة فى المقام .

ولتطبيق هذين الرقمين نأخذ كميات هذه المحاصيل الأربعة (فى أراضى مصلحة الأملاك الأميرية أيضاً) وهى المبينة فى الجدول الآتى :

جدول ٤٩ — كميات محاصيل القطن والقمح والفول والشعير
فى سنتى ١٩٣١ و ١٩٣٥

السنة	القطن	القمح	الفول	الشعير
	قناطير	أرادب	أرادب	أرادب
١٩٣١	٨٢٥٨	٣٨٠٩	٢٥٨٩	٣١٧٩
١٩٣٥	٩٥٥٩	٤٥٧٦	٣٠٠٦	٤١٩٩

هنا نجد

$$\frac{\text{مج ع}_1\text{ك}_0}{\text{مج ع}_0\text{ك}_0} = \frac{\text{٣٫٩٦٠} \times \text{٨٢٥٨} + \text{١٫٥٣١} \times \text{٣٨٠٩} + \text{١٫٣٤٣} \times \text{٢٥٨٩} + \text{٠٫٨٠٢} \times \text{٣١٧٩}}{\text{٣٫٦٢٥} \times \text{٨٢٥٨} + \text{١٫٢١٩} \times \text{٣٨٠٩} + \text{١٫٣٩٩} \times \text{٢٥٨٩} + \text{٠٫٦٤٣} \times \text{٣١٧٩}}$$

$$= \frac{\text{٤٤٥٥٩٫٨٤٤}}{\text{٤٠٢٤٤٫٥٢٩}}$$

$$\therefore \quad \text{١٠٠} \times \frac{\text{مج ع}_1\text{ك}_0}{\text{مج ع}_0\text{ك}_0} = \text{١١٠٫٧٢} .$$

وهذا الرقم يساوى ، بالتقريب ، الرقم الذى حصلنا عليه بدون أوزان . وبالمثل نجد أن الرقم القياسي بالأوزان ب هو : ا

$$\text{١٠٠} \times \frac{\text{مج ع}_1\text{ك}_1}{\text{مج ع}_0\text{ك}_1} = \text{١١٠٫٨٨} ؛$$

وهى أيضاً نتيجة مخالفة لـكل من النتيجتين السابقتين ، ولو أن الفرق صغير . ولـكنه يصح أن يكون أكبر من ذلك إذا زاد عدد السلع ، أو إذا كان الاختلاف بين الـكميات الأساسية والـكميات المقارنة أكبر مما هو فى هذا المثال .

ويلاحظ أن الرقم البسيط يقع فى الوسط بين الرقمين (٢ ا) و (٢ ب) .

الرقم القياسى الأمثل

٢٨٧ — يصح أن نجمع بين نظام الأوزان المستعمل فى الرقم (٢ ا) والنظام المستعمل فى (٢ ب) ، فنحصل على رقم جديد يكون أكثر اعتدالا وأقل تحيزاً من كل منهما . فمثلا لو أخذنا الوسط الهندسى لهذين الرقمين نحصل على ما يسميه الأستاذ إرفنج فيشر **الرقم القياسى الامثل**(١) ، وهو

$$\sqrt{\frac{\text{مج ع}_1\text{ك}_0}{\text{مج ع}_0\text{ك}_0} \times \frac{\text{مج ع}_1\text{ك}_1}{\text{مج ع}_0\text{ك}_1}} \quad \ldots \quad (\text{٣})$$

وسنرى فيما بعد أن هذا الرقم يستحق هذه التسمية حقيقة ، حيث تجتمع فيه كل الصفات المطلوبة فى الرقم القياسى الصحيح ، ويخلو من العيوب التى تشوب الأرقام القياسـية الأخرى . وهذا ما يجعله فى المـكان الأول بين جميع الأرقام القياسية والمثل الأعلى لها .

ولو حسبنا الرقم القياسى فى المثال الذى بأيدينا على أساس هذه المعادلة نجد أنه يساوى $\sqrt{\text{١٢٢٧٦٫٦٣٣٦}} = \text{١١٠٫٨٠٠}$

رقم آخر يجمع بين الرقمين ٢ ا و ٢ ب

٢٨٨ — ويصح أن نجمع بين نظامى الأوزان فى (٢ ا) و (٢ ب) فى صورة أخرى ، فنأخذ الوزن لـكل سلعة يساوى الوسـط الحسابى (أو أى وسط آخر) للـكميتين ك_0 و ك_1 ، فنحصل على الصورة الآتية :

(١) يعرف هذا الرقم باسم (Irving Fisher's "Ideal Index.")

$$\frac{\text{مج ع}_1 (\text{ك}_0 + \text{ك}_1)}{\text{مج ع}_0 (\text{ك}_0 + \text{ك}_1)} \qquad (١٣) \; .$$

وهذه الصيغة (١٣) أسهل فى العمل الحسابى من الصيغة (٣) ، ولكن هذه الأخيرة تفضلها من عدة وجوه نشرحها فى مناسبة أخرى .

المتوسط البسيط للمناسيب

٢٨٩ — يمكننا تركيب أرقام قياسية أيضاً باستخدام مناسيب الأسعار بدل الأسعار نفسها ، فنحسب الوسط الحسابى ، أو الهندسى أو التوافقى (بسيطاً أو مرجحاً بأوزان مناسبة) لمناسيب الأسعار التى سبق أن عرفناها فى بند ٢٨٣ . فاذا فرضنا أن هذه المناسيب للسلع المختلفة ، وعددها ن ، هى :

$$\text{س} \text{ ، } \text{سَ} \text{ ، } \text{سً} \text{ ، } \ldots$$

يكون الوسط الحسابى البسيط لهذه المناسيب هو

$$\frac{\text{س} + \text{سَ} + \text{سً} + \cdots}{\text{ن}} = \frac{\text{مج س}}{\text{ن}} \qquad (٤) \; .$$

والوسط الهندسى البسيط المناسيب

$$= \sqrt[\text{ن}]{\text{س} \times \text{سَ} \times \text{سً} \times \cdots} \qquad (٥) \; .$$

والوسط التوافقى للمناسيب هو ف مثلا ، حيث

$$\frac{\text{ن}}{\text{ف}} = \frac{1}{\text{س}} + \frac{1}{\text{سَ}} + \frac{1}{\text{سً}} + \cdots \qquad (٦) \; .$$

وفى المثال الذى بأيدينا نرى أن هذه المناسيب هى على التوالى :

منسوب سعر القطن $= \frac{٣ر٩٦٠}{١ر٢١٩} \times ١٠٠ = ١٠٩ر٢٤$ ؛

« « القمح $= \frac{١ر٥٣١}{١ر٢١٩} \times ١٠٠ = ١٢٥ر٥٩$ ؛

منسوب سعر الفول $= \frac{١ر٣٤٣}{١ر٣٩٩} \times ١٠٠ = ٩٦ر٠٠$ ؛

« « الشعير $= \frac{٠ر٨٠٢}{٠ر٦٤٣} \times ١٠٠ = ١٢٤ر٧٣$.

∴ الوسط الحسابى البسيط لها $= ١١٣ر٨٩$ ؛

و « التوافقى « « $= ١١٣ر٠٦$ ؛

و « الهندسى « « $= ١١٣ر٢١$.

وبناء على ذلك يكون مستوى أسعار هذه المحاصيل ارتفع فى سنة ١٩٣٥ عنه فى سنة ١٩٣١ بمقدار ١٣ر٨٩ أو ١٣ر٠٦ أو ١٣ر٢١ فى المائة ، حسب الصيغة التى نختارها للرقم القياسى : الوسط الحسابى أو التوافقى أو الهندسى (البسيط) . وهذه النتائج مختلفة فيما بينها ، وتخالف أيضاً النتائج التى حصلنا عليها من قبل . وهذا بالرغم من أن كلها ترمى إلى تصوير شىء واحد : ألا وهو مستوى أسعار هذه المحاصيل فى سنة ١٩٣٥ بالنسبة إلى سنة ١٩٣١ . ونكتفى هنا بالإشارة إلى هذا الاختلاف وسنتكلم عن منشئه فيما بعد .

ترجيح المناسيب حسب أهمية السلع

٢٩٠ — هذه المتوسطات البسيطة لا تفرق بين مناسيب السلع المختلفة بل تعاملها جميعاً نفس المعاملة ، مع العلم بأن بعض السلع يزيد فى الأهمية عن البعض الآخر . وعلى ذلك فالأرقام القياسية التى نحصل عليها بالمتوسطات البسيطة لا تصور الحالة على حقيقتها فتعطى نتائج مضللة أو خاطئة . ولذلك يستحسن تعديل هذه المتوسطات باستخدام أوزان تتناسب مع أهمية السلع ، نرجح بها المناسيب الخاصة بها .

وأحسن شىء نقيس به أهمية السلعة هو قيمتها ، أى حاصل ضرب سعرها فى كميتها . ولكن أى كمية وأى سعر ؟ فقد عرفنا أن لكل سلعة سعراً أساسياً

وسعر مقارن ، وكذلك كمية أساسية وكمية مقارنة . فمن الممكن إذن أن نختار أحد التوافيق الآتية وهى :

ا — السعر الأساسى × الكمية الأساسية $= \text{ع}_{٠} \times \text{ك}_{٠} = \text{م}_{٠}$ مثلا ؛

ب — « » × « المقارنة $= \text{ع}_{٠} \times \text{ك}_{١} = \text{مـ}_{٠}$ « ؛

حـ — « المقارن × « الأساسية $= \text{ع}_{١} \times \text{ك}_{٠} = \text{مـ}_{١}$ « ؛

ى — « » × « المقارنة $= \text{ع}_{١} \times \text{ك}_{١} = \text{م}_{١}$ « .

وعلى ذلك يمكن تركيب رقم قياسى من المناسيب على الصور الآتية :

وسط حسابى مرجح بأوزان ا وهو $\frac{\text{مج س م}_{٠}}{\text{مج م}_{٠}}$ ... (٤ ا) ،

» » » » ب » $\frac{\text{مج س مـ}_{٠}}{\text{مج مـ}_{٠}}$ ... (٤ ب) ،

» » » » حـ » $\frac{\text{مج س مـ}_{١}}{\text{مج مـ}_{١}}$ ... (٤ حـ) ،

» » » » ى » $\frac{\text{مج س م}_{١}}{\text{مج م}_{١}}$ ... (٤ ى) .

يلاحظ أنه بوضع $\text{س} = \frac{\text{ع}_{١}}{\text{ع}_{٠}}$ يتضح أن (٤ ا) هو نفس (١٢ ا) المذكور فى بند ٢٨٦ ، وأن (٤ ب) هو نفس (٢ب) .

الوسط التوافقى المرجح

٢٩١ — وبالمثل نحصل على الوسط التوافقى المرجح ، بهذه الأوزان الأربعة ؛ ويكون مقلوب الوسط التوافقى يساوى :

$\frac{١}{\text{مج م}_{٠}} \times \text{مج} \left(\frac{\text{م}_{٠}}{\text{س}} \right)$ ... (٥ ا) ،

أو $\frac{١}{\text{مج مـ}_{٠}} \times \text{مج} \left(\frac{\text{مـ}_{٠}}{\text{س}} \right)$ ... (٥ ب) ،

أو $\frac{١}{\text{مج مـ}_{١}} \times \text{مج} \left(\frac{\text{مـ}_{١}}{\text{س}} \right)$ ... (٥ حـ) ،

أو $\frac{١}{\text{مج م}_{١}} \times \text{مج} \left(\frac{\text{م}_{١}}{\text{س}} \right)$ ... (٥ ى) .

الوسط الهندسى المرجح

٢٩٢ — أما فى الوسط الهندسى المرجح فتظهر الأوزان كأسس ترفع إليها المناسيب ، حيث قد قلنا فى باب المتوسطات إن لوغاريتم الوسط الهندسى لأى مجموعة من القيم ، يساوى الوسط الحسابى للوغاريتمات هـذه القيم . وعلى ذلك فالأرقام القياسية المركبة على أساس الوسط الهندسى المرجح (بأوزان ا و ب و حـ و ى) هى :

ا — بأوزان م٠ هو $\sqrt[\text{مج م}_{٠}]{(\text{س})^{\text{م}_{٠}} \times (\text{س}')^{\text{م}'_{٠}} \times (\text{س}'')^{\text{م}''_{٠}} \times \dots}$ (٦ ا) ؛

ب — « مـ٠ هو $\sqrt[\text{مج مـ}_{٠}]{(\text{س})^{\text{مـ}_{٠}} \times (\text{س}')^{\text{مـ}'_{٠}} \times (\text{س}'')^{\text{مـ}''_{٠}} \times \dots}$ (٦ ب) ؛

حـ — « مـ١ هو $\sqrt[\text{مج مـ}_{١}]{(\text{س})^{\text{مـ}_{١}} \times (\text{س}')^{\text{مـ}'_{١}} \times (\text{س}'')^{\text{مـ}''_{١}} \times \dots}$ (٦ حـ) ؛

ى — « م١ هو $\sqrt[\text{مج م}_{١}]{(\text{س})^{\text{م}_{١}} \times (\text{س}')^{\text{م}'_{١}} \times (\text{س}'')^{\text{م}''_{١}} \times \dots}$ (٦ ى) .

يوجد ستة أسس للارقام القياسية

٢٩٣ — وهكذا يكون لدينا ستة أسس لتركيب أرقام قياسية وهى : (١) تجميع الأسعار نفسها ؛ (٢) الوسط الحسابى للمناسيب ؛ (٣) الوسط التوافقى ؛ (٤) الوسط الهندسى . ونذكر على سبيل الحصر أيضاً (٥) الوسيط و (٦) المنوال لهذه المناسيب ، ولكن استعمالها نادر جداً .

وفى كل من هـذه الحالات يصح أن نستخدم أوزاناً (نختارها بعدة طرق) لترجيح السلع المختلفة بما يتناسب وأهميتها ، أو لا نستخدم أوزاناً بالمرة .

دخول عنصر الزمن فى المقارنة على أساس قديم

٢٩٤ — عندما نستخدم الأرقام القياسية لبيان حركة الأسعار — أو أى ظاهرة أخرى — أثناء مدة طويلة ، يدخل فى المسألة عنصر جديد وهو الزمن ؛ حيث إذا طالت المدة بين السنة أو الفترة المعتبرة أساساً والسنة المقارنة ، فإن الزمن

كفيل بأن يحدث تغييراً كبيراً فى الظروف المحيطة بالسلع التى نبحثها والتى يتركب منها الرقم القياسى . فقد يحدث مثلا أن بعض السلع التى كانت شائعة الاستهلاك فى سنة الأساس قل استهلاكها أو انعدم فى الأزمنة الحديثة ؛ أو بالعكس قد توجد سلع لم تكن معروفة من قبل ، كما نرى فى حالة سلعة مثل الحرير الصناعى ومنسوجاته . أو على الأقل تتغير الأهمية النسبية بين السلع التى تدخل فى تركيب الرقم القياسى . وهذا إذا حصل — ولا بد أن شيئاً من هذا يحصل إذا طالت الفترة — لا يمكن الاطمئنان إلى صحة المقارنة ، وتنعدم بذلك الفائدة التى من أجلها فكرنا فى إنشاء الرقم القياسى . ويجب إذن عند البحث فى الرقم القياسى للأسعار أو خلافها ، بالنسبة إلى سنة بعيدة ، أن نتأكد من استمرار الظروف على حالها — ولو تقريباً — أثناء فترة المقارنة .

نستعمل طريقة السلسلة إذا تغيرت الظروف

٢٩٥ — وإذا لم نضمن ثبات هذه الظروف الاقتصادية ، فيحسن أن نركب سلسلة من الأرقام القياسية كل سنة بالنسبة إلى سابقتها كأساس ، وهذه السلسلة يمكن استخدامها لمراقبة حركة الأسعار من سنة إلى أخرى تسبقها أو تليها .

ولتوضيح طريقة تركيب هذه السلسلة[1] ، نفرض أن لدينا أسعار السلع فى عدة سنين متتالية مثل ١٩٢٠ و ١٩٢١ و ١٩٢٢ و .. ؛ ولتكن هذه الأسعار كما يأتى :

$ع_{٠}$ ، $ع_{١}$ ، $ع_{٢}$ ، . . . للسلعة الأولى ؛

و $ع'_{٠}$ ، $ع'_{١}$ ، $ع'_{٢}$ ، . . . « الثانية ؛

و $ع''_{٠}$ ، $ع''_{١}$ ، $ع''_{٢}$ ، . . . « الثالثة ؛

وهكذا لبقى السلع .

نحسب الرقم القياسى للأسعار فى سنة ١٩٢١ بالنسبة إلى سنة ١٩٢٠ ، طبقاً لأى معادلة من المعادلات التى شرحناها فى هذا الباب . ولنرمز إلى هذا الرقم

(١) هذه الطريقة تسمى بالإنجليزية (Chain System)

للاختصار بالرمز[1] $ع_{٠١}$. ثم نحسب الرقم القياسى لأسعار سنة ١٩٢٢ بالنسبة إلى سنة ١٩٢١ كأساس ؛ وليكن هذا الرقم هو $ع_{١٢}$. ونحسب أيضاً رقم سنة ١٩٢٣ بالنسبة إلى سنة ١٩٢٢ وليكن $ع_{٢٣}$ ؛ وهكذا مع باقى السنين . فنحصل على سلسلة الأرقام القياسية الآتية :

$ع_{٠١}$ ، $ع_{١٢}$ ، $ع_{٢٣}$ ، $ع_{٣٤}$ ، . . . .

وهى تعبر عن مستوى الأسعار فى كل سنة بالنسبة إلى التى قبلها مباشرة كأساس . وبذلك يكون الأساس فى هذه السلسلة **متحركاً**[2] ، وليس ثابتاً كما لو اتخذنا سنة ١٩٢٠ أساساً لكل السنين .

ميزة هذه الطريقة على السابقة

٢٩٦ — وعند حساب أى واحد من هذه الأرقام يمكننا أن ندخل فى الاعتبار أى تغيير يطرأ على الظروف المحيطة بهذه السلع ، فنكيف تركيب الرقم بما يلائم الحالة فى كل سنة . فمثلا ندخل فى الرقم أى سلعة عظم شأنها فى السوق بعد أن كانت قليلة الأهمية ؛ أو نهمل سلعة أخرى ضعفت حركتها وفقدت أهميتها فى السوق بعد أن كانت تلعب دوراً خطيراً بين السلع من قبل . وبالجملة نعدل الأهمية النسبية بين السلع الداخلة فى الرقم بما يتناسب مع الظروف .

وهذه المرونة غير موجودة فى الرقم القياسى ذى القاعدة الثابتة ، الذى تنسب فيه الأسعار فى السنة المقارنة إلى السنة الأساسية مباشرة ، وبذلك ترجح هى أو مناسيبها بأوزان ثابتة على طول السنين ، فلا تتمشى مع الظروف الاقتصادية التى تحيط بالسلع الداخلة فى تركيب الرقم .

(١) هذا الرمز ينطق هكذا : عين صفر واحد ، أى مستوى الأسعار للسنة ١ ، بالنسبة إلى السنة ٠ كأساس ؛ وكذلك $ع_{١٢}$ عين واحد اثنين ، أى أسعار السنة ٢ بالنسبة إلى السنة ١ وهكذا .

(٢) يسمى بالإنجليزية (Shifting or Moving Base)

تثبيت الأساس في طريقة السلسلة

٢٩٧ — وإذا أردنا الرجوع إلى أساس ثابت ، فيمكننا بسهولة أن نعبر عن مستوى الأسعار في هذه السنين بالنسبة إلى سنة ١٩٢٠ (السنة الأساسية) بطريق غير مباشر كما يأتي :

∵ $\text{ع}_{\text{١٠}}$ هو مستوى الأسعار في سنة ١٩٢١ بالنسبة إلى سنة ١٩٢٠

و $\text{ع}_{\text{٢١}}$ « « « « ١٩٢٢ « « ١٩٢١

∴ حاصل الضرب

$$\text{ع}_{\text{١٠}} \times \text{ع}_{\text{٢١}}$$

يعبر عن مستوى الأسعار في سنة ١٩٢٢ بالنسبة إلى سنة ١٩٢٠ (عن طريق سنة ١٩٢١) . وكذلك حاصل الضرب

$$(\text{ع}_{\text{١٠}} \times \text{ع}_{\text{٢١}}) \times \text{ع}_{\text{٣٢}}$$

يعبر عن مستوى الأسعار في سنة ١٩٢٣ بالنسبة إلى سنة ١٩٢٠ ، (عن طريق سنتي ١٩٢١ و ١٩٢٢) .

وهكذا نحصل على الرقم القياسي لمستوى الأسعار في كل سنة بواسطة الرقم القياسي للسنة التي قبلها مباشرة ؛ فتتكون لدينا سلسلة من الأرقام القياسية تعبر عن مستوى الأسعار منسوباً إلى أساس ثابت (وهو سنة ١٩٢٠ هنا) . وبالطبع تكون هذه الأرقام على العموم مخالفة (١) للأرقام القياسية المحسوبة بالطريقة المباشرة ، أي بنسبة كل سنة إلى سنة الأساس مباشرة .

مثال لحساب الأرقام القياسية عملياً

٢٩٨ — نأخذ مثلا أسعار أربعة المحاصيل المصرية : القمح والأذرة والفول والشعير في السنين ١٩٣٠ ـ ١٩٣٥ ، ونحسب سلسلة الأرقام القياسية لمستوى أسعار هذه المحاصيل بالطرق المختلفة . وللسهولة نستعمل الوسط الحسابي البسيط للمناسيب .

(١) إلا في حالة استعمال الوسط التجميعي البسيط للأسعار ، وفي حالة ما تتغير أسعار جميع السلع بنسبة واحدة .

جدول ٥٠ — أسعار بعض الحبوب المصرية(١)
في السنين ١٩٣٠ — ١٩٣٥

	الأسعار بالقرش للأردب					
	١٩٣٠	١٩٣١	١٩٣٢	١٩٣٣	١٩٣٤	١٩٣٥
القمح	١٢٩	١٣٢	١٠٧	١١٢	١٥٠	١٤٣
الأذرة	٩٥	٧٧	٥٩	١٠٤	١٠٦	٧٣
الفول	١٨٣	١٥٩	٨٥	٨٨	١٥٥	١٥٣
الشعير	٧٠	٨٨	٥٦	٤٦	٩٠	٨٥

الوسط الحسابي للمناسيب بأساس ثابت

أولا : توجد الأرقام القياسية للأسعار بالنسبة إلى سنة ١٩٣٠ كأساس ثابت ؛ فنحسب في كل سنة منسوب السعر لكل سلعة بالنسبة إلى السنة الأساسية ، ثم نوجد الوسط الحسابي للمناسيب في كل سنة ينتج الرقم القياسي للأسعار . وهذه العمليات مبينة في الجدول الآتي :

جدول ٥١ — مناسيب أسعار أربعة محاصيل
في السنين ١٩٣١ — ١٩٣٥ بالنسبة إلى سنة ١٩٣٠ كأساس

	١٩٣٠	١٩٣١	١٩٣٢	١٩٣٣	١٩٣٤	١٩٣٥
القمح	١٠٠	١٠٢ر٣٣	٨٢ر٩٥	٨٦ر٨٢	١١٦ر٢٨	١١٠ر٨٥
الأذرة	١٠٠	٨١ر٠٥	٦٢ر١١	١٠٩ر٤٧	١١١ر٥٨	٧٦ر٨٤
الفول	١٠٠	٨٦ر٨٩	٤٦ر٤٥	٤٨ر٠٩	٨٤ر٧٠	٨٣ر٦١
الشعير	١٠٠	١٢٥ر٧١	٨٠ر٠٠	٦٥ر٧١	١٢٨ر٥٧	١٢١ر٤٣
متوسط المناسيب	١٠٠	٩٩ر٠٠	٦٧ر٨٨	٧٧ر٥٢	١١٠ر٢٨	٩٨ر١٨

(١) انظر الاحصاء السنوي العام ١٩٣٥ — ١٩٣٦ . صفحة ٥٠٤ .

فالأرقام القياسية للأسعار فى السنين ١٩٣١ — ١٩٣٥ بالنسبة إلى سنة ١٩٣٠ كأساس ثابت ، هى على الترتيب :

٩٩ر٠٠ ، ٦٧ر٨ ، ٧٧ر٥ ، ١١٠ر٣ ، ٩٨ر٢ .

الرقم القياسى بأساس متحرك

ثانيا : توجد الأرقام القياسية على نظام السلسلة ذات الأساس المتحرك ؛ فنحسب منسوب سعر السلعة فى كل سنة بالنسبة إلى سابقتها مباشرة . والمنسوب بهذا المعنى الخاص نسميه[1] **النسيب** . فيكون نسيب سعر القمح فى سنة ١٩٣٥ مثلا ، هو خارج قسمة السعر فى سنة ١٩٣٥ على السعر فى سنة ١٩٣٤ ، مضروبا فى ١٠٠ . ونرمز له بالرمز $س_{٥٤}$ مثلا .

والرقم القياسى المطلوب يكون إذن هو الوسط الحسابى لأنسبة السلع فى كل سنة . أى أن سنة ١٩٣٥ يكون رقمها القياسى بطريقة السلسلة هو :

$$ع_{٥٤} = \frac{١}{٤} (مج\ س_{٥٤})$$

جدول ٥٢ — أنسبة أسعار أربعة المحاصيل فى السنين ١٩٣١ — ١٩٣٥ . أساس متحرك

	١٩٣١	١٩٣٢	١٩٣٣	١٩٣٤	١٩٣٥
القمح	١٠٢ر٣٣	٨١ر٠٦	١٠٤ر٦٧	١٣٣ر٩٣	٩٥ر٣٣
الأذرة	٨١ر٠٥	٧٦ر٦٣	١٧٦ر٢٥	١٠١ر٩٣	٦٨ر٨٧
الفول	٨٦ر٨٩	٥٣ر٤٦	١٠٣ر٥٣	١٧٦ر١٣	٩٨ر٧١
الشعير	١٢٥ر٧١	٦٣ر٦٣	٨٢ر١٤	١٩٥ر٦٦	٩٤ر٤٥
متوسط الأنسبة	٩٩ر٠٠	٦٨ر٧٠	١١٦ر٦٥	١٥١ر٩١	٨٩ر٣٤

وعلى ذلك فالأرقام القياسية للأسعار فى السنين ١٩٣١ — ١٩٣٥ ، كل بالنسبة إلى سابقتها ، هى على الترتيب :

(١) بالانجليزية (Link Relative)

٩٩ و ٦٨ر٧ و ١١٦ر٥ و ١٥١ر٩ و ٨٩ر٣ ؛

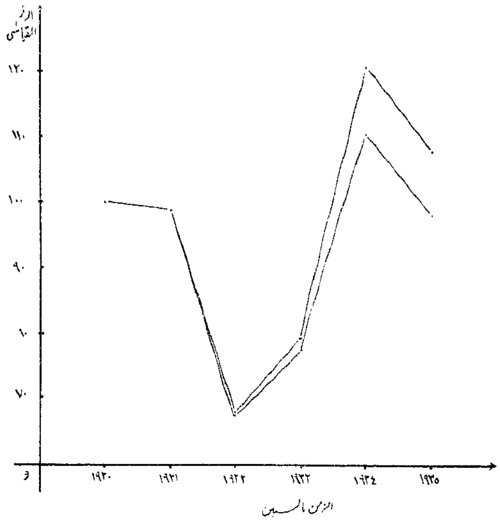
أى $ع_{١٠}$ و $ع_{٢١}$ و $ع_{٣٢}$ و $ع_{٤٣}$ و $ع_{٥٤}$

و إذا أردنا الرجوع إلى أساس ثابت (١٩٣٠ مثلا) ، نحسب الحواصل

$ع_{١٠}$ ، $ع_{١٠} \times ع_{٢١}$ ، $ع_{١٠} \times ع_{٢١} \times ع_{٣٢}$ ، $ع_{١٠} \times ع_{٢١} \times ع_{٣٢} \times ع_{٤٣}$ ، ...

فينتج الرقم القياسى لكل سنة بالنسبة إلى سنة ١٩٣٠ كأساس (غير مباشر) . وهذه الأرقام هى :

٩٩ر٠٠ ؛ ٦٨ر٠ ؛ ٧٩ر٣ . ١٢٠ر٥ ؛ ١٠٧ر٧



(شكل ٦٨)
الأرقام القياسية لأسعار الحبوب بالنسبة إلى سنة ١٩٣٠

ونلاحظ أن هذه الأرقام مخالفة للتى حصلنا عليها فى جدول ٥١ ، حيث نسبنا كل سنة إلى سنة الأساس مباشرة . ويلاحظ أن الاختلاف يزيد كلما بعدنا عن سنة الأساس ، ونرى ذلك موضحاً فى شكل ٦٨ .

بعض الأرقام القياسية المهمة

الرقم القياسى لأسعار الجملة

٢٩٩ — **الرقم القياسى لأسعار الجملة**[1] من أهم الأرقام القياسية المستعملة فى المسائل الاقتصادية على العموم . ونجد الحكومات فى كل البلاد المتمدنة تقوم بعمل ونشر هذا الرقم للدلالة على حركة التجارة ، وعلى الحالة الاقتصادية العامة فى البلد .

سنة ١٩١٣ كانت تؤخذ كأساس

٣٠٠ — وقد كانت أغلب الحكومات إلى عهد قريب تنشئ الرقم القياسى لأسعار الجملة بالنسبة إلى الأسعار فى سنة ١٩١٣ كأساس ؛ لأن هذه السنة تعتبر عادية وخالية من التقلبات الاقتصادية العنيفة ؛ ولأنها تمثل الظروف الاقتصادية قبل الحرب العظمى ١٩١٤—١٩١٨ . ولكن نظراً لطول المدة من سنة ١٩١٣ إلى الآن ، وتغير الظروف الاقتصادية تغيراً كبيراً أثناء هذه المدة ، بدأت بعض الدول فى عمل الأرقام القياسية لأسعار الجملة على أساس أحدث مثل سنة ١٩٣٠ أو بعدها . والرقم المصرى الجديد[2] أساسه سنة ١٩٣٥ .

الطريقة المتبعة فى رقم أسعار الجملة

٣٠١ — والطريقة المتبعة عادة هى الوسط الهندسى (البسيط) لمناسيب الأسعار ؛ وفى مصر تؤخذ الأنسبة . كل شهر بالنسبة إلى سابقه ، على نظام السلسلة الذى شرحناه فى بند ٢٩٧ .

وعدد السلع التى يتكون منها الرقم يكون عادة حوالى ١٠٠ سلعة ، تختار بحيث تمثل السوق تمثيلا صحيحاً ، بحيث لا يتحيز إلى بعض السلع دون الأخرى ، مثل السلع المحلية والمستوردة ، والسلع الجاهزة والنصف مصنوعة والخام ، والسلع الزراعية والصناعية ؛ وهكذا .

(١) بالانجليزية (Whole Sale Prices Index Number)

(٢) انظر الاحصاء السنوى العام ١٩٣٥ —١٩٣٦ صفحة ٤٩٠ .

ونكتفى هنا بهذا الوصف البسيط ؛ وسنعود إلى شرح هذا الرقم بالتفصيل فى المستقبل .

الرقم القياسى لنفقة المعيشة

٣٠٢ — **الرقم القياسى لنفقة المعيشة**[1] ، وهو يقوم مقام الرقم القياسى لأسعار التجزئة فى أغلب الأبحاث ، من أهم الأرقام القياسية المستعملة أيضاً ، حيث إنه يمس النواحى الاجتماعية والتجارية والصناعية للدولة فى نفس الوقت . وربما يعتبر أهم من الرقم القياسى لأسعار الجملة الذى يختص بناحية واحدة من نواحى النشاط الاقتصادى للبلد ، ألا وهى تجارة الجملة .

والمقصود من إنشاء هذا الرقم القياسى هو قياس التغيرات التى تطرأ على نفقة المعيشة بسبب تغيرات أسعار الحاجيات الضرورية للحياة ، والتى يستهلكها السواد الأعظم من السكان ، لكى نكون دائماً على بينة من الحالة المعيشية للشعب ودرجة رفاهيته ؛ ولنعلم مبلغ كفاية الأجور فى توفير أسباب المعيشة ؛ ومقدرة الناس على شراء السلع والحاجيات التى يستهلكونها .

والأساس المتفق عليه فى إنشاء هذا الرقم ، فى كل الدول تقريباً ، هو أسعار هذه الحاجيات قبل الحرب العظمى أى سنة ١٩١٣ ، لنفس السبب الذى ذكرناه .

تستخدم طريقة الوسط المرجح

٣٠٣ — والطريقة المتبعة فى تركيب الرقم القياسى لنفقة المعيشة ، هى الوسط المرجح لمناسيب أسعار هذه الحاجيات فى الشهر أو السنة المقارنة ، على أساس أسعارها فى سنة ١٩١٣ .

والأوزان المستعملة فى ترجيح مناسيب الأسعار تتناسب مع أهمية السلع . وتقاس أهمية السلعة بمقدار ما يخصص للإنفاق عليها من الدخل الكلى .

(١) اسمه بالإنجليزية (Cost of Living Index Number)

تقسيم السلع فى أربع مجموعات أو خمس

٣٠٤ — وفى العادة تقسّم الحاجيات التى تدخل فى تركيب الرقم القياسى لنفقة المعيشة (من سلع وخدمات وغيرها) إلى مجموعات ؛ ويكوّن لكل منها رقم قياسى خاص يبين التغيرات فى أسعارها على حدة ومن هذه الأرقام الجزئية يكوّن الرقم القياسى العام الذى يشمل كل المجموعات . وهذه المجموعات خمس فى أغلب البلاد ؛ وهى تمثل البنود الرئيسية للانفاق وهى : المأكل ، والملبس والمسكن ، والإضاءة ، والتدفئة (للبلاد الباردة) ، والمصروفات النثرية التى لا تدخل تحت أى واحد من هذه البنود .

وتقسيم الدخل بين هذه المجموعات ، و بين المفردات الداخلة فى كل مجموعة مبنى على أساس ميزانية الأسرة النموذجية ، وكيفية توزيع دخلها فى شراء ما كانت تستهلكه من الحاجيات قبل الحرب — أوفى سنة ١٩٢٠ بالنسبة لمصر [1] ، وفى سنة ١٩٢٧ — ١٩٢٨ بالنسبة لألمانيا .

معنى الرقم القياسى لنفقة المعيشة

٣٠٥ — و بناء على ذلك يكون معنى الرقم القياسى لنفقة المعيشة ، وليكن ١٣٢ مثلاً ، كما كان فى شهر مايو سنة ١٩٣٨ بالنسبة إلى سنة ١٩١٣ ، هو أن المعيشة التى كانت تتكلف ١٠٠ قرش فى سنة ١٩١٣ أصبحت تتكلف ١٣٢ قرشاً فى مايو سنة ١٩٣٨ . وهذا طبعاً لا يقرر أن الناس لا زالوا يعيشون كما عاشوا فى سنة ١٩١٣ ، أو أنهم يستهلكون الآن ما كانوا يستهلكونه فى تلك السنة . وهذه فى الحقيقة نقطة ضعف فى طريقة تركيب هذا الرقم ، وفى غيره أيضاً ، نظراً لطول المدة التى مضت من سنة ١٩١٣ إلى الآن .

وقد رأينا فى بند ٤١ (صفحة ٣٢) النسب المئوية لتقسيم الدخل بين البنود

(١) انظر الإحصاء السنوى العام ١٩٣١ — ١٩٣٢ صفحة ٤٣٩ . أو انظر تقرير نفقات المعيشة الملحق « بالإحصائية الشهرية الزراعية » شهر نوفمبر سنة ١٩٢٠ .

الأربعة المهمة فى مصر وانجلترا (خمسة بنود) ، وهى الأرقام المستعملة فى تركيب الرقم القياسى لنفقة المعيشة فى البلدين .

الرقم القياسى للانتاج

٣٠٦ — يلى هذين الرقمين فى الأهمية **الرقم القياسى للإنتاج الصناعى** [1] (أو الزراعى) ؛ ونجد كثيراً من البلاد الصناعية فى الوقت الحاضر تعنى بعمل ونشر أرقام قياسية للإنتاج الصناعى بأقسامه المختلفة ، وخصوصاً منذ سنة ١٩٣٠ .

وهذا الرقم يستخدم للدلالة على حالة النشاط الصناعى فى الدولة . وهو يبين كمية المنتج من الصناعات المختلفة ، بصرف النظر عن القيمة النقدية لهذه المنتجات. والمقصود بالكمية ، مجردة عن القيمة ، هو عدد الوحدات المنتجة فى كل صناعة؛ لأن هذا العدد يدل بدقة على درجة النشاط أو الركود فى الأعمال الصناعية ؛ و يدل أيضاً على حالة العمل والبطالة بين جمهور العمال ؛ و يدل على مقدار الكفاية الإنتاجية للصناعات المختلفة والعمال المشتغلين فيها ؛ وغير ذلك من الظواهر الاقتصادية المهمة .

ومثل هذا يقال عن الإنتاج الزراعى . وفى مصر يوجد رقم قياسى لهذا الإنتاج الزراعى تنشره مصلحة عموم الإحصاء [2] ، ولكن لا يوجد رقم قياسى للإنتاج الصناعى .

طريقة إنشاء الرقم القياسى للانتاج

٣٠٧ — الطريقة المتبعة فى تركيب هذا الرقم القياسى تشابه طريقة الرقم القياسى لنفقة المعيشة ، حيث نتفق على سنة عادية نعتبرها أساساً ؛ ونكوّن لكل صناعة منسوباً لكمية الانتاج فيها فى السنة المقارنة بالنسبة للسنة الأساسية .

(١) اسمه بالانجليزية (Index Number of (the Physical Volume of) Production)

(٢) انظر الإحصاء السنوى العام ١٩٣٥ — ١٩٣٦ صفحة ٣٨٤ .

وهـذه المناسيب ترجح بأوزان تتناسب وأهمية الصناعات ، ويستخرج الوسط المرجح لهذه المناسيب ، وهو الرقم القياسى المطلوب .

وتقاس أهميـة كل صناعة بمقدار صافى إنتاجها [1] أثنـاء فترة زمنية معينة تعتبر عادية وهادئة (تؤخذ سنة الأساس عادة) ؛ والأوزان المستعملة تتناسب مع هذه المقادير .

بعض الصعوبات العملية

٣٠٨ — الصعوبات التى نلاقيها فى عمل هذا الرقم القياسى هى أولا من ناحية الحصول على بيانات عن الصناعات المختلفة يمكن مقارنتها بسهولة على مرور السنين . لأن المنتجات الصناعية تتغير فى شكلها وماهيتها وفى طريقة صنعها ، على مرور الزمن . وعلى ذلك فمن الصعب مقارنة عدد الوحدات المنتجة هـذا العـام بعدد الوحدات المنتجة فى نفس المصنع ، أو نفس الصناعة ، من مدة خمس سنين مثلا .

وثانياً نلاقي صعوبة أخرى فى جمع بيانات يعتمد عليها من عـدد كبير من الصناعات المؤسسات ، أو فى مدة قصيرة تسمح بعمـل الرقم القياسى للإنتـاج ونشره بسرعة ، قبل مرور مدة طويلة على الفترة المشمولة فى تركيب الرقم .

رغم هذه الصعوبات يجب عمل هذا الرقم

٣٠٩ — ولـكن هذه الصعوبات ، وإن لم يمكن التغلب عليها تماماً ، يجب ألا تقعدنا عن تركيب هذا الرقم بأى طريقة كانت ، لما له من الفائدة العظيمة فى تصوير حركة النشاط الصـناعى ، ومعرفة مواطن الركود ومعالجتها بسرعة ، ومعرفة نواحى النشاط واستغلالها . ونكتفى هنا بالإشارة إلى هـذه الصعوبات وسوف نعود إلى شرحها بالتفصيل والطرق المستعملة فى بعض البلاد للتغلب على هذه الصعوبات العملية .

(١) أو بعدد العمال المشتغلين فيها .

أرقام قياسية أخرى

٣١٠ — خلاف هـذه الأرقام القياسية ، نجد بعض البلاد تكوّن أرقاماً قياسية لأسعار الأوراق المـالية وللأجور وغيرها من الظواهر الاقتصادية والاجتماعية المهمة . وسيأتى الـكلام على هذه الأرقام وكيفية تركيبها فيما بعد .

المـراجع

Bowley, A. L.	*Elements of Statistics*	Chapter	IX
Conner, L.R.	*Statistics in Theory and Practice*	"	XV
Mills, F.C.	*Statistical Methods*	"	VI
Secrist, H.	*Statistical Methods*	"	XV

الباب الثاني عشر

اختبار الأرقام القياسية

توجد عدة صيغ مختلفة للرقم القياسي

٣١١ — تكلمنا فى الباب السابق عن الصيغ المختلفة المستعملة فى تركيب الأرقام القياسية ؛ وشرحنا طريقة بنائها من البيانات التى يمكننا الحصول عليها ، مثل أسعار وكميات السلع . وقد أوردنا هذه المعادلات أو الصيغ المختلفة على اعتبار أن كلا منها يعبر بطريقة ما عن المعنى أو الغرض المقصود من الرقم القياسى : ألا وهو رقم نسبى ملخص يقيس مقدار التغير فى ظاهرة أو عدة ظواهر بين لحظة وأخرى ، أو بين بلد وأخرى .

النتائج تختلف حسب الصيغ المستعملة

٣١٢ — وقد رأينا أن الأرقام القياسية المحسوبة على أساس هذه المعادلات المختلفة تؤدى إلى نتائج مختلفة أيضاً ، أى أنها تصور الأحوال تصويراً مختلفاً على حسب المعادلة أو الصيغة التى نستعملها . وهذا بطبيعة الحال يؤدى بنا إلى التساؤل :

أى هذه الصيغ أو المعادلات صحيح ، وأيها خطأ ؟

وكيف نختبر أى واحدة منها لنتبين صحتها أو خطأها ، ومقدار هذا الخطأ ؟

وكيف نختار الأصلح من بينها جميعاً ؟

٣١٣ — والمفاضلة بين صيغ الأرقام القياسية المختلفة لا بد تتناول الناحيتين النظرية والعملية . وقد بحث الأستاذ إرفنج فيشر[1] فى الأسس النظرية لكيفية

(١) انظر كتابه (I. Fisher. *The Making of Index Numbers*)

اختبار الأرقام القياسية وتصحيحها ، وإليه يرجع الفضل الأكبر فى هذا الموضوع ؛ وسنشرح هنا باختصار الاختبارات التى يقترحها وكيفية تطبيقها .

الرموز المستعملة فى البحث

٣١٤ — لنفرض ، كما سبق ، أن أسعار السلع وكمياتها فى الفترة (أو البلد) الأساسية كما هى يلى على الترتيب :

$$\text{ع}_{٠} ، \text{ع}'_{٠} ، \text{ع}''_{٠} ، \ldots ،$$

$$\text{و} \qquad \text{ك}_{٠} ، \text{ك}'_{٠} ، \text{ك}''_{٠} ، \ldots ؛$$

وأن هذه الأسعار والكميات فى السنة (أو البلد) المقارنة هى :

$$\text{ع}_{١} ، \text{ع}'_{١} ، \text{ع}''_{١} ، \ldots ،$$

$$\text{و} \qquad \text{ك}_{١} ، \text{ك}'_{١} ، \text{ك}''_{١} ، \ldots .$$

وإذا كان هناك سنة مقارنة أخرى نضع الرقم ٢ بدل الرقم ١ بجانب هذه الحروف . وعلى العموم ندل على السنة الأساسية بالرقم ٠ ، والسنين الأخرى بالأرقام ١ و ٢ و ٣ و ٠٠ ، إذا وجدت . فاذا استعملنا الحرف س للدلالة على مناسيب الأسعار ، كان $\text{س}_{٠١}$ يدل على منسوب سعر السنة ١ على أساس السنة ٠ ، وكان $\text{س}_{٠٢}$ يدل على منسوب السعر فى السنة ٢ بالنسبة إلى السنة ٠ كأساس . وبالمثل $\text{س}_{١٢}$ يدل على منسوب السعر للسنة ٢ على أساس السنة ١ ، وهكذا . أى أن الرقم الأول يدل على السنة المعتبرة أساساً ، والرقم الثانى يعين السنة المقارنة .

اختبار ترتيب وضع السلع

٣١٥ — من البدهى أن الرقم القياسى الصحيح يجب ألا تتغير قيمته إذا تغير ترتيب أسعار السلع فى المعادلة الجبرية للرقم ؛ فمثلا فى الرقم التجميعى البسيط ، وهو المذكور فى المعادلة (١) بند ٢٨٥ ، :

$$\frac{\text{ع}_{١} + \text{ع}'_{١} + \text{ع}''_{١} + \cdots}{\text{ع}_{٠} + \text{ع}'_{٠} + \text{ع}''_{٠} + \cdots} ،$$

لا تتغير قيمة هذا الكسر إذا غيرنا ترتيب وضع الأسعار ع ، عَ ، عً ، . . فى البسط أو فى المقام . وعلى ذلك تكون هذه المعادلة ، أو الصيغة مستوفية لهذا الشرط . وفى الواقع نجد أن كل الصيغ التى ذكرناها فى الباب السابق تستوفى هذا الشرط .

الانعكاس فى الزمن

الانعكاس فى الزمن أو فى المكان

٣١٦ — إذا أخذنا سلعة معينة وحسبنا الرقم القياسى لسعرها فى سنة ١٩٣٨ بالنسبة إلى سعرها فى سنة ١٩٣٠ مثلا ، فمن البدهى أن هذا الرقم القياسى يساوى مقلوب الرقم القياسى لسعر هذه السلعة فى سنة ١٩٣٠ بالنسبة إلى سعرها فى سنة ١٩٣٨ كأساس . وإلا كان هذا الرقم القياسى خاطئاً ، ولا يؤدى المعنى أو الغرض الذى نرمي إليه من استخدام الأرقام القياسية . وكذلك لو كنا نبحث فى الأسعار فى مكانين مختلفين بدل سنتين .

وعلى ذلك يجب أن ينقلب الرقم القياسى عندما نعكس ترتيب السنين ، أو الأماكن المنسوبة والمنسوب إليها . فلو كان الرقم القياسى للأسعار فى القاهرة ، مثلا ، بالنسبة إلى الاسكندرية يساوى ١٢٥ ٪ ، يجب أن يكون الرقم القياسي للاسكندرية بالنسبة إلى القاهرة ٨٠ ٪ . وهذا واضح لأنه إذا كانت النسبة بين سعر سلعة فى القاهرة وسعرها فى الاسكندرية هى ١٢٥ ٪ ، أى كنسبة ٥ : ٤ ، كانت النسبة بين سعرها فى الاسكندرية وسعرها فى القاهرة تساوى ٤ : ٥ ، أى ٨٠ ٪ . وما يقال عن سلعة واحدة هنا يقال عن عدة سلع .

وبالاختصار يجب أن يكون حاصل ضرب الرقمين القياسيين المتبادلين فى الزمن (أو المكان) يساوى ١ .

إذن لو أخذنا صيغة من الصيغ التى نعرفها ، وحسبنا بواسطتها الرقم القياسى للأسعار فى سنة ١٩٣٨ بالنسبة إلى سنة ١٩٣٠ مثلا ، ثم عكسنا الترتيب الزمنى وحسبنا ، بواسطتها أيضاً ، الرقم القياسى للأسعار فى سنة ١٩٣٠ بالنسبة إلى سنة ١٩٣٨ ، يجب أن يكون حاصل ضرب النتيجة يساوى الواحد . وإلا فتكون هذه الصيغة غير صالحة للاستعمال كرقم قياسى دقيق ، ويجب نبذها .

البديل الزمنى

٣١٧ — ***البديل الزمنى*** [1] لأى رقم قياسى هو نفس الرقم محسوباً بطريقة عكسية ، بحيث تعتبر السنة المقارنة هى السنة الأساسية ، والسنة الأساسية هى السنة المقارنة . أى أننا نضع الأسعار $ع_1$ بدل $ع_0$ والعكس بالعكس ، فى معادلة الرقم القياسى ؛ وكذلك $ك_1$ بدل $ك_0$ إذا وجدت .

∴ الرقم $\frac{\text{مج }ع_1}{\text{مج }ع_0}$ يكون بديله الزمنى $\frac{\text{مج }ع_0}{\text{مج }ع_1}$

وهذا البديل يعبر عن أسعار السنة . بالنسبة إلى أسعار السنة ١ .

وكذلك الصيغة $\frac{\text{مج }ع_1 ك_0}{\text{مج }ع_0 ك_0}$ يكون بديلها الزمنى $\frac{\text{مج }ع_0 ك_1}{\text{مج }ع_1 ك_1}$.

» » $\frac{\text{مج }ع_1 ك_1}{\text{مج }ع_0 ك_1}$ » » » $\frac{\text{مج }ع_0 ك_0}{\text{مج }ع_1 ك_0}$.

شرط الانعكاس فى الزمن

٣١٨ — ولكى يكون أى رقم مستوفياً شرط ***الانعكاس فى الزمن*** [2] ، يجب إذن أن يكون حاصل ضرب هذا الرقم فى بديله الزمنى يساوى الواحد تماماً . فلنختبر إذن بعض المعادلات التى حصلنا عليها فى الباب السابق لنرى إذا ما كانت تفى بهذا الشرط أو لا .

---

(١) يسمى بالانجليزية (Time Reciprocal)

(٢) اسمه بالإنجليزية (Time Reversal Test)

اختبار الأرقام التجميعية

٣١٩ — من الواضح أن الرقم التجميعى البسيط يستوفي هذا الشرط ، لأن :

$$\frac{\text{مجـ}\ \text{ع}_{1}}{\text{مجـ}\ \text{ع}_{0}} \times \frac{\text{مجـ}\ \text{ع}_{0}}{\text{مجـ}\ \text{ع}_{1}} \ (\text{وهو البديل الزمنى}) = 1$$

أما الرقم التجميعى المرجح ، سواء بكميات السنة الأساسية أو بكميات السنة المقارنة ، فلا يستوفى هذا الشرط . لأن

$$\frac{\text{مجـ}\ \text{ع}_{1}\text{ك}_{0}}{\text{مجـ}\ \text{ع}_{0}\text{ك}_{0}} \times \text{بديله الزمنى وهو}\ \frac{\text{مجـ}\ \text{ع}_{0}\text{ك}_{1}}{\text{مجـ}\ \text{ع}_{1}\text{ك}_{1}} \neq 1 \ ؛$$

$$\text{وكذلك}\quad \frac{\text{مجـ}\ \text{ع}_{1}\text{ك}_{1}}{\text{مجـ}\ \text{ع}_{0}\text{ك}_{1}} \times \text{»}\ \text{»}\ \text{»}\ \frac{\text{مجـ}\ \text{ع}_{0}\text{ك}_{0}}{\text{مجـ}\ \text{ع}_{1}\text{ك}_{0}} \neq 1\ \text{أيضاً} .$$

الرقم الأمثل ينعكس فى الزمن

٣٢٠ — أما الصيغة المشتقة من هاتين الأخيرتين ، والتى يسميها فيشر الرقم القياسى الأمثل كما ذكرنا فى بند ٢٨٧ ، فهى تستوفى هذا الشرط . لأن هذه الصيغة وهى :

$$\sqrt{\frac{\text{مجـ}\ \text{ع}_{1}\text{ك}_{0}}{\text{مجـ}\ \text{ع}_{0}\text{ك}_{0}} \times \frac{\text{مجـ}\ \text{ع}_{1}\text{ك}_{1}}{\text{مجـ}\ \text{ع}_{0}\text{ك}_{1}}}\ \text{، بديلها الزمنى هو}\ \sqrt{\frac{\text{مجـ}\ \text{ع}_{0}\text{ك}_{1}}{\text{مجـ}\ \text{ع}_{1}\text{ك}_{1}} \times \frac{\text{مجـ}\ \text{ع}_{0}\text{ك}_{0}}{\text{مجـ}\ \text{ع}_{1}\text{ك}_{0}}}$$

وواضح أن حاصل ضرب هاتين الكميتين = ١ تماماً .

وكذلك الصيغة التى أوردناها فى بند ٢٨٨ ، وهى

$$\frac{\text{مجـ}\ \text{ع}_{1}\ (\text{ك}_{0} + \text{ك}_{1})}{\text{مجـ}\ \text{ع}_{0}\ (\text{ك}_{0} + \text{ك}_{1})}$$

تنعكس فى الزمن أيضاً .

الوسط الحسابى للمناسيب ، بسيط أو المرجح ، لا ينعكس فى الزمن

٣٢١ — الوسط الحسابى للمناسيب ، سواء كان بسيطاً أو مرجحاً بأى الأوزان التى ذكرناها فى بند ٢٩٠ ، لا ينعكس فى الزمن . فالوسط البسيط مثلا ، وهو $\frac{1}{ن}$ (مجـ س) حيث $\text{س} = \text{ع}_{1} \div \text{ع}_{0}$ ، لا ينعكس لأن

$$\frac{1}{ن}\left(\frac{\text{ع}_{1}}{\text{ع}_{0}} + \frac{\text{ع}'_{1}}{\text{ع}'_{0}} + \frac{\text{ع}''_{1}}{\text{ع}''_{0}} + \cdots\right) \times \frac{1}{ن}\left(\frac{\text{ع}_{0}}{\text{ع}_{1}} + \frac{\text{ع}'_{0}}{\text{ع}'_{1}} + \frac{\text{ع}''_{0}}{\text{ع}''_{1}} + \cdots\right) \neq 1 .$$

ويلاحظ هنا أن البديل الزمنى للوسط الحسابى للمناسيب يساوى مقلوب وسطها التوافقى .

وكذلك الوسط المرجح بالأوزان ا المذكورة فى بند ٢٩٠ ، وهو

$$\frac{\text{مجـ}\ \text{س م}_{0}}{\text{مجـ}\ \text{م}_{0}} = \frac{\text{مجـ}\ \text{ع}_{1}\text{ك}_{0}}{\text{مجـ}\ \text{ع}_{0}\text{ك}_{0}}\ \text{، لأن}\ \text{س} = \frac{\text{ع}_{1}}{\text{ع}_{0}}\ \text{،}\ \text{م}_{0} = \text{ع}_{0}\text{ك}_{0}\ ؛$$

وقد أثبتنا ، فى البند السابق ، أن هذا الأخير لا ينعكس . وبالمثل نجد $\frac{\text{مجـ}\ \text{س م}_{1}}{\text{مجـ}\ \text{م}_{1}}$ لاينعكس أيضاً . وكذلك الوسطان المرجحان بالأوزان حـ ، ى المذكوران فى بند ٢٩٠ ، لا ينعكسان فى الزمن .

الوسط التوافقى لا ينعكس فى الزمن

٣٢٢ — والوسط التوافقى للمناسيب ، بسيطاً كان أو مرجحاً بأى أوزان ، لا ينعكس أيضاً ؛ بدليل أن مقلوبه يساوى البديل الزمنى للوسط الحسابى ، كما لاحظنا فى البند السابق . أى أن الوسط الحسابى هو البديل الزمنى لمقلوب الوسط التوافقى . وعدم انعكاس أحدهما يقتضى عدم انعكاس الآخر .

الوسط الهندسى البسيط ينعكس فى الزمن

٣٢٣ — أما الوسط الهندسى للمناسيب فلا ينعكس فى الزمن إلا إذا كان بسيطاً غير مرجح بأى نوع من الأوزان . بدليل أن الوسط الهندسى البسيط ، وهو

$$\sqrt[ن]{\frac{\text{ع}_{1}}{\text{ع}_{0}} \times \frac{\text{ع}'_{1}}{\text{ع}'_{0}} \times \frac{\text{ع}''_{1}}{\text{ع}''_{0}} \times \cdots}\ \text{، بديله الزمنى هو}\ \sqrt[ن]{\frac{\text{ع}_{0}}{\text{ع}_{1}} \times \frac{\text{ع}'_{0}}{\text{ع}'_{1}} \times \frac{\text{ع}''_{0}}{\text{ع}''_{1}} \times \cdots}$$

ومن الواضح أن حاصل ضرب هاتين الكميتين يساوى ١ تماماً .

وإذا أخذنا أى واحد من الأوساط الهندسية المرجحة المذكورة فى بند ٢٩٢ نجد أنه لا ينعكس . فمثلا نأخذ :

$$\sqrt[\text{مج م}_0]{(\text{س})^{\text{م}_0} \times (\bar{\text{س}})^{\bar{\text{م}}_0} \times (\bar{\bar{\text{س}}})^{\bar{\bar{\text{م}}}_0} \times \cdots}$$

أى $\left[\left(\frac{\text{ع}_1}{\text{ع}_0}\right)^{\text{ع}_0 \text{ك}_0} \times \left(\frac{\bar{\text{ع}}_1}{\bar{\text{ع}}_0}\right)^{\bar{\text{ع}}_0 \bar{\text{ك}}_0} \times \left(\frac{\bar{\bar{\text{ع}}}_1}{\bar{\bar{\text{ع}}}_0}\right)^{\bar{\bar{\text{ع}}}_0 \bar{\bar{\text{ك}}}_0} \cdots\right]^{\frac{1}{\text{مج ك}_0 \text{ع}_0}}$.

ولو وضعنا ع_1 بدل ع_0 و ع_0 بدل ع_1 فى هذه المعادلة ، وكذلك بالنسبة إلى ك_0 و ك_1 ، نحصل على البديل الزمنى لهذه المعادلة . ومن الواضح أن حاصل ضرب هذه المعادلة فى هذا البديل الزمنى لا يساوى الواحد أبداً .

تقييد المعادلات التى لا تنعكس فى الزمن

٣٢٤ — هذا الاختيار الزمنى إذن أداة نافعة نختبر بها الأرقام القياسية فى أشكالها المختلفة ، ونستبعد منها ما لا يصلح فلا نستعمله ، ونستبقى منها ما كان مستوفياً لشروط هذا الاختبار . وبذلك نضمن صحة المقارنة التى نجريها بواسطة هذه الأرقام القياسية . أو على الأقل نضمن سلامتها من الأخطاء المترتبة على عدم استيفاء الرقم القياسى لشروط هذا الاختبار .

ولكن هذا الاختبار لا يمس إلا ناحية واحدة فقط ، وهى نسبة الأسعار فى تاريخين (أو مكانين) مختلفين . وهناك ناحية أخرى لا تقل عن هذه فى الأهمية . فالأرقام القياسية ، كما قلنا سابقاً ، لا تستخدم فقط لبيان حركة أسعار السلع ، بل تستخدم أيضاً فى بيان التغيرات فى كميات السلع ، وفى غيرها من الظواهر مثل الأجور والانتاج وهكذا .

منسوب سعر سلعة فى منسوب كميتها يساوى منسوب القيمة

٣٢٥ — ولو حسبنا الرقم القياسى (منسوب) لسعر سلعة واحدة فى سنتين مختلفتين واحدة بالنسبة الى الأخرى ، وحسبنا أيضاً الرقم القياسى (منسوب) لكمية هذه السلعة فى نفس السنتين ، فمن البدهى أن حاصل ضرب هذين الرقمين يساوى النسبة بين قيمتى هذه السلعة فى نفس السنتين ، واحدة بالنسبة إلى الأخرى . وإلا كان الرقم القياسى المستعمل خاطئاً فى تصويره للأحوال .

شرط الانعكاس فى العامل

٣٢٦ — وقياساً على ذلك إذا استعملنا أى صيغة من الصيغ التى نعرفها للأرقام القياسية فى حساب الرقم القياسى لأسعار عدة سلع فى سنتين مختلفتين واحدة بالنسبة إلى الأخرى ، واستعملنا نفس الصيغة فى حساب الرقم القياسى لكميات هذه السلع فى نفس السنتين ، يجب أن يكون حاصل ضرب هذين الرقمين القياسيين ، للأسعار والكميات ، مساويا للنسبة بين قيم هذه السلع (القيمة = الكمية × السعر) فى نفس السنتين تحت البحث . وخلاف ذلك تكون الصيغة المستعملة في حساب هذين الرقمين القياسيين صيغة خاطئة ولا تصلح للاستعمال . وهكذا يكون لدينا اختبار آخر نمتحن بواسطته ما عندنا من الأرقام القياسية ، ونستبعد منها ما لا يستوفى هذا الشرط .

هذا الشرط يسمى شرط **الانعكاس فى العامل** أو **الانعكاس المعاملى**[(١)] ويمكن وضعه فى صورة مختصرة كما يأتى :

نفرض أن الأسعار الأساسية والمقارنة للسلع هى ع_0 و ع_1 ؛ وكذلك الكميات الأساسية والمقارنة ك_0 و ك_1 ، فيكون مجموع قيم هذه السلع فى السنة الأساسية والسنة المقارنة هو على الترتيب مج ع_0 ك_0 و مج ع_1 ك_1 .

∴ يكون منسوب القيم $= \frac{\text{مج ع}_1 \text{ك}_1}{\text{مج ع}_0 \text{ك}_0}$.

ونفرض أن **البديل المعاملى** لأى معادلة من معادلات الأرقام القياسية ، هو نفس المعادلة موضوعاً فيها ع بدل ك و ك بدل ع ، أى نضع الكمية بدل السعر والعكس بالعكس . فاذا كانت المعادلة الأصلية تعبر عن رقم قياسى للأسعار ، يكون البديل المعاملى هو الرقم القياسى للكميات ، مركباً بنفس الطريقة ، والعكس بالعكس .

(١) بالإنجليزية (Factor Reversal Test)

وبذلك يكون الشرط الواجب تحققه فى كل رقم قياسى ليكون صالحاً للاستعمال هو :

$$\text{الرقم القياسى} \times \text{البديل العاملى له} = \frac{\sum ع_1 ك_1}{\sum ع_0 ك_0} ،$$

وهذا هو شرط الانعكاس العاملى .

تطبيق الاختبارين على كل الأرقام القياسية

٣٢٧ — وهكذا يكون لدينا اختباران أساسيان ، نستخدمهما فى انتقاء الأرقام القياسية الصالحة واستبعاد غيرها . والرقم الصالح للاستعمال إذن هو الرقم الذى يستوفى هذين الاختبارين فى وقت واحد . فنختبر كل الأرقام التى عرفناها بالنسبة للانعكاس الزمنى ونستبعد مالا ينعكس منها ؛ ثم نختبر الباقى بالنسبة إلى الانعكاس العاملى . وسواء اختبرنا الأرقام بهذا الترتيب أو بالعكس ، فالنتيجة واحدة فى الحالتين . وبذلك نحصل على الأرقام القياسية التى تستوفى الاختبارين معاً . فتكون صالحة للاستعمال .

اختبار بعض الأرقام

٣٢٨ — الآن نختبر بعض الأرقام القياسية التى عرفناها ، بالنسبة إلى الانعكاس العاملى . وسنكتفى هنا باختبار تلك الأرقام التى ثبت لنا أنها تنعكس فى الزمن . وطريقة اختبار أى رقم هى أن نضرب به فى بديله العاملى ؛ فاذا كان حاصل الضرب مساوياً منسوب القيم ، كان الرقم منعكساً .

٣٢٩ — الرقم التجميعى البسيط (ينعكس فى الزمن) لا ينعكس فى العامل ، لأنه يساوى

$$\frac{\sum ع_1}{\sum ع_0} \text{ ، وبديله العاملى يساوى } \frac{\sum ك_1}{\sum ك_0} .$$

$$\text{ولكن} \qquad \frac{\sum ع_1}{\sum ع_0} \times \frac{\sum ك_1}{\sum ك_0} \neq \frac{\sum ع_1 ك_1}{\sum ع_0 ك_0} .$$

$\therefore$ هذا الرقم لا ينعكس فى العامل .

الرقم الأمثل ينعكس فى العامل أيضاً

٣٣٠ — نأخذ الرقم الأمثل الذى وضعه فيشر ، وهو

$$\sqrt{\frac{\sum ع_1 ك_0}{\sum ع_0 ك_0} \times \frac{\sum ع_1 ك_1}{\sum ع_0 ك_1}} \text{ ، وبديله العاملى وهو } \sqrt{\frac{\sum ك_1 ع_0}{\sum ك_0 ع_0} \times \frac{\sum ك_1 ع_1}{\sum ك_0 ع_1}} .$$

وواضح أن حاصل ضرب هاتين الكميتين $= \frac{\sum ع_1 ك_1}{\sum ع_0 ك_0}$.

أى أن هذا الرقم الأمثل ينعكس فى العامل وفى الزمن أيضا . والواقع أنه الرقم الوحيد ، بين الأرقام التى ذكرناها كلها ، الذى يستوفى شرطى الانعكاس معاً . وهو لهذا السبب أحسن الأرقام القياسية وأصحها من الناحية النظرية — والعملية أيضاً ، إلى حد ما . ومن ثم كانت تسميته بالرقم الأمثل (Ideal) .

الوسط الهندسى لا ينعكس

٣٣١ — الوسط الهندسى البسيط لا ينعكس فى العامل ، لأنه يساوى :

$$\sqrt[ن]{\frac{ع_1}{ع_0} \times \frac{\bar{ع}_1}{\bar{ع}_0} \times \cdots} \text{ ؛ وبديله العاملى يساوى } \sqrt[ن]{\frac{ك_1}{ك_0} \times \frac{\bar{ك}_1}{\bar{ك}_0} \times \cdots} \text{ ؛}$$

وحاصل ضرب هاتين الكميتين $\neq \frac{\sum ع_1 ك_1}{\sum ع_0 ك_0}$ منسوب القيم .

ومن السهل إثبات أن باقى الأرقام القياسية التى ذكرناها فى هذا الباب وفى الباب السابق ، بسيطة كانت أو مرجحة بأى نوع من الأوزان ، لا تنعكس فى العامل .

ومن هذا يظهر لنا أن الاختبار الثانى — الانعكاس فى العامل — هو فى الواقع اختبار « أصعب » من الاختبار الأول أو الزمنى ، حيث بعض الأرقام تمر فى الاختبار الأول ولا تمر فى الثانى .

أخطاء الأرقام القياسية

الوسط الحسابى فى

٣٣٢ — لنبحث الآن فيما يترتب على عدم استيفاء شروط هذين الاختبارين .

لنأخذ مثلا الوسط الحسابى (البسيط) للمناسيب . فقد رأينا أن بديله الزمنى يساوى مقلوب الوسط التوافقى ؛ وأنه لا ينعكس فى الزمن . وعلى ذلك فخارج قسمة الوسط الحسابى على الوسط التوافقى لا يساوى ١ . فهل هو أكبر أوأقل من ١ ؟ الزمنى من ١

وجدنا فى المثال الذى أخذناه فى الباب السابق (صفحة ٣٢١) أن الوسط الحسابى لمناسيب أسعار أربعة محاصيل هو ١١٣٫٨٩٪ ، وأن الوسط التوافقى لها ١١٣٫٠٦ . خارج قسمة هذين (الأول على الثانى) يساوى ١٫٠٠٧ . أى أن حاصل الضرب (الوسط × بديله الزمنى) فى هذا المثال أكبر من ١ .

والحقيقة أن حاصل ضرب الوسط الحسابى للمناسيب فى بديله الزمنى دائماً يكون أكبر من ١ . وهذا يمكن إثباته جبرياً بسهولة [١] . وبالعكس نجد أن الوسط التوافقى مضروباً فى بديله الزمنى يعطى ناتجاً أقل من ١ . ويمكننا أيضاً إثبات مثل هذا بالنسبة للوسط المرجح بأوزان ا ، ب ، ح ، أو ء (بند ٢٩٠) ، وللوسط التوافقى أيضاً [٢] .

---

(١) لاثبات ذلك نضرب الوسط الحسابى فى بديله الزمنى ، مع العلم بأن البديل الزمنى للمنسوب س هو ١ ÷ س

$$\therefore \frac{1}{\text{ن}}(\text{س} + \text{س}' + \text{س}'' + \cdots) \times \frac{1}{\text{ن}}\left(\frac{1}{\text{س}} + \frac{1}{\text{س}'} + \frac{1}{\text{س}''} + \cdots\right)$$

$$= \frac{1}{\text{ن}^2}\left[\text{ن} + \frac{\text{س}}{\text{س}'} + \frac{\text{س}'}{\text{س}} + \frac{\text{س}}{\text{س}''} + \frac{\text{س}''}{\text{س}} + \cdots\right]$$

$$\therefore \text{حاصل الضرب} = \frac{1}{\text{ن}^2}\left[\text{ن} + \frac{\text{ن}(\text{ن}-1)}{2}\ \text{حداً مثل}\ \left(\frac{\text{س}}{\text{س}'} + \frac{\text{س}'}{\text{س}}\right)\right.$$

بعدد توافيق الحروف س ، سَ ، سً . . مثنى مثنى . وكل من هذه الحدود أكبر من ٢ . لأن $\text{ح} + \frac{1}{\text{ح}} > 2$ ، لأن $\text{ح}^2 + 1 > 2\,\text{ح}$ لأن $(\text{ح} - 1)^2 > 0$]

$$\therefore \text{حاصل الضرب} > \frac{1}{\text{ن}^2}[\text{ن} + \text{ن}^2 - \text{ن}]\text{؛ أى} > 1\ \text{وهو المطلوب .}$$

(٢) البرهان فى هذه الحالة أكثر تعقيداً منه فى الحالة المذكورة . انظر ص ٢٨٦ من كتاب : (I. Fisher, *The Making of Index Numbers*.)

٣٣٣ — وينتج من ذلك أن الوسط الحسابى للمناسيب ، البسيط أو المرجح ، **متحيز إلى أعلى** [١] . والوسط التوافقى **متحيز إلى أسفل** [٢] ، ومعنى ذلك أن الوسط الحسابى للمناسيب ، فى دلالته على مستوى الأسعار فى سنة معينة بالنسبة إلى أخرى ، يميل إلى إظهار هذا المستوى أعلى مما هو فى الحقيقة . وبالعكس مع الوسط التوافقى ، فهو يميل إلى تصوير مستوى الأسعار أقل مما هو فى الحقيقة . تحيز الوسطين الحسابى والتوافقى

فنرى إذن أن عدم انعكاس الرقم القياسى فى الزمن ، وهو ناتج من خطأ فى حقيقته أو كيفية تركيبه ، يترتب عليه تحيز الرقم ، وعدم دقته فى تصوير الأشياء على حقيقتها . ونخطئ إذا نحن اعتمدنا على مثل هذا الرقم .

وكذلك عدم الانعكاس فى المعامل يترتب عليه تحيز فى الأرقام القياسية .

٣٣٤ — وقد ينشأ التحيز أيضاً عن نوع الأوزان المستعملة فى ترجيح الرقم القياسى ، وكيفية إدخالها أو استعمالها فى المعادلة التى تمثله . فلنأخذ مثلا الوسط الحسابى للمناسيب ، وهو كما نعلم بدون أوزان $= \frac{\text{مجـ س}}{\text{ن}}$ تحيز الأوزان

$$= \frac{1}{\text{ن}}\left[\frac{\text{ع}_1}{\text{ع}_.} + \frac{\text{ع}'_1}{\text{ع}'_.} + \frac{\text{ع}''_1}{\text{ع}''_.} + \cdots\right]$$

فاذا رجحناه بالأوزان $\text{م}_1 = \text{ع}_1\,\text{ك}_.$ يصبح

$$\frac{\text{مجـ س م}_1}{\text{مجـ م}_1} = \frac{1}{\text{مجـ م}_1}\left[\left(\frac{\text{ع}_1}{\text{ع}_.}\right)\text{ع}_1\,\text{ك}_. + \left(\frac{\text{ع}'_1}{\text{ع}'_.}\right)\text{ع}'_1\,\text{ك}'_. + \cdots\right]\text{،}$$

أى أننا نرجح المناسيب الكبيرة ، التى فيها البسط ع_1 كبيراً بالنسبة إلى $\text{ع}_.$ ، بأوزان كبيرة أيضاً ؛ وفى الوقت نفسه نرجح المناسيب الصغيرة ، التى فيها

---

(١) بالانجليزية (Baised Upwards). (٢) (Baised Downwards)

$ع_1$ صغيرة بالنسبة إلى $ع_0$ ، بأوزان صغيرة أيضاً . وبذلك نتحيز للمناسيب الكبيرة ضد المناسيب الصغيرة . وينتج إذن أن هذا الرقم متحيز إلى أعلى بسبب الأوزان .

وكذلك الوسط المرجح بالأوزان $م_1 = ع_1\, ك_1$ ، أى الأوزان ء بند ٢٩٠ ، يكون متحيزاً إلى أعلى أيضاً .

ومن هذا يتبين لنا أن الوسط الحسابى لمناسيب الأسعار المرجح بأوزان $م_0$ أو $م_1$ ، يكون بطبيعة تركيبه متحيزاً إلى أعلى حتماً ، سواء فى ذلك أكانت الأسعار فى ارتفاع أو فى هبوط .

المتوسط المرجح ا أو ب ، أصغر من المرجح حـ أو ء

٣٣٥ — وإذا قارنا بين متوسط المناسيب المرجح ا أو ب (أى بأوزان $م_0$ أو $م_0$) ومتوسطها المرجح حـ أو ء (أوزان $م_1$ أو $م_1$) ، نجد أن المتوسط الأول دائماً أصغر من المتوسط الثانى . وذلك سواء كانت حركة الأسعار فى هبوط أو صعود .

لأنه عند استعمال الأوزان $م_0$ أو $م_0$ نضرب المناسيب $\frac{ع_1}{ع_0}$ فى $ع_0\, ك_0$ أو فى $ع_0\, ك_1$ ، على الترتيب . وبذلك تختفي الأسعار $ع_0$ وتبقى الحواصل $ع_1\, ك_0$ أو $ع_1\, ك_1$. أما عند استعمال الأوزان $م_1$ أو $م_1$ فنضرب نفس المناسيب $\frac{ع_1}{ع_0}$ فى $ع_1\, ك_0$ أو $ع_1\, ك_1$ على الترتيب . وتكون النتيجة ، كما قلنا فى البند السابق ، أن المناسيب الكبيرة تضرب فى أوزان كبيرة أيضاً ، والمناسيب الصغيرة تضرب فى أوزان صغيرة ، وبذلك يكون المتوسط الناتج أكبر من المتوسط الأول حتماً ، سواء ارتفعت الأسعار أو هبطت .

أما العلاقة بين المتوسطين المرجحين بأوزان ا و ب ، أو بين المتوسطين المرجحين حـ و ء ، فلا يمكن تحديدها بطريقة جازمة كما فعلنا هنا ، إذ يصح أن يكون المتوسط ا أكبر أو أصغر من المتوسط ب . وكذلك المتوسطان حـ و ء .

تحيز الوسط الهندسى المرجح

٣٣٦ — لنبحث الآن فى الوسط الهندسى المرجح وما يترتب على عدم انعكاسه فى الزمن .

نرمز إلى الوسط الهندسى البسيط بالحرف هـ ، وللوسط الهندسى المرجح بالأوزان ا و ب و حـ و ء ، بالرموز الآتية[1] على الترتيب :

ها ، هب ، هج ، هد .

ونرمز للبديلات الزمنية لهذه المتوسطات بالرموز

$ها_1$ ، $هب_1$ ، $هج_1$ ، $هد_1$.

فنعلم أن

$$ها = \sqrt[\text{مج}\,م_0]{(س)^{م_0} \times (\bar{س})^{\bar{م}_0} \times (\bar{\bar{س}})^{\bar{\bar{م}}_0} \times \cdots} \quad (1) ;$$

وأن

$$ها_1 = \sqrt[\text{مج}\,م_1]{(س)^{-م_1} \times (\bar{س})^{-\bar{م}_1} \times (\bar{\bar{س}})^{-\bar{\bar{م}}_1} \times \cdots} \quad (2) ;$$

لأن $م_0 = ع_0\, ك_0$ ، وبديلها الزمنى $م_1 = ع_1\, ك_1$ ؛

و $س = \frac{ع_1}{ع_0}$ » $= \frac{ع_0}{ع_1} = س^{-1}$.

فلكى ينعكس ها فى الزمن يجب أن يكون $ها \,.\, ها_1 = 1$

أو بعبارة أخرى ، يجب أن يكون لوغاريتم هذا الحاصل يساوى صفراً .

وإذا كان $ها \,.\, ها_1 = 1$ ، فإن $لو\; ها \,.\, ها_1 = 0$.

» » $ها \,.\, ها_1 > 1$ » $لو\; ها \,.\, ها_1 > 0$.

» » $ها \,.\, ها_1 < 1$ » $لو\; ها \,.\, ها_1 < 0$.

(١) تنطق هذه الرموز هكذا : هـ ألف ، هـ باء ، هـ جيم ، هـ دال .

ولكن

$$\text{لو ها} = \frac{1}{\text{مجـ}_{\cdot}}\left(\text{م}_{\cdot}\,\text{لو س} + \text{مَ}_{\cdot}\,\text{لو سَ} + \text{مً}_{\cdot}\,\text{لو سً} + \cdot\cdot\right) \;..\;(٣)؛$$

$$\text{و} \; -\text{لو ها}_1 = \frac{1}{\text{مجـ}_1}\left(\text{م}_1\,\text{لو س} + \text{مَ}_1\,\text{لو سَ} + \text{مً}_1\,\text{لو سً} + \dots\right) \;...\;(٤).$$

والطرف الأيسر فى كل من هاتين المتساويتين عبارة عن وسط مرجح للوغاريتمات المناسيب ، بأوزان $\text{م}_{\cdot}$ و م_1 على الترتيب . ونرى فى (٤) ، كما فى البندين السابقين ، أن المناسيب الكبيرة مرجحة بأوزان كبيرة أيضاً ، والمناسيب الصغيرة مرجحة بأوزان صغيرة أيضاً .

وينتج من ذلك أن الطرف الأيسر من المتساوية (٤) أكبر من الطرف الأيسر فى (٣) . وسواء فى ذلك ارتفعت الأسعار أو هبطت .

$$\therefore \; \text{لو ها . ها}_1 = \text{لو ها} + \text{لو ها}_1$$
$$= \text{كمية سالبة .}$$

$$\therefore \; \text{ها . ها}_1 < ١ .$$

وعلى ذلك يكون الوسط الهندسى المرجح بأوزان ا متحيزاً إلى أسفل دائماً .

٣٣٧ — نأخذ الوسط المرجح **هب** ، ونثبت بنفس الطريقة أنه متحيز إلى أسفل أيضاً . المعادلة الجبرية لهذا الوسط هى كما نعلم : الوسط هب متحيز لأسفل

$$\text{هب} = \sqrt[\text{مجـ}_{\cdot}]{(\text{س})^{\text{مـ}_{\cdot}} \times (\text{سَ})^{\text{مـَ}_{\cdot}} \times (\text{سً})^{\text{مـً}_{\cdot}} \times \dots} \qquad (١)$$

وبديله الزمنى هب_1

$$\text{هب}_1 = \sqrt[\text{مجـ}_1]{(\text{س})^{\text{مـ}_1} \times (\text{سَ})^{\text{مـَ}_1} \times (\text{سً})^{\text{مـً}_1} \times \dots} \qquad (٢)$$

حيث $\text{مـ}_{\cdot} = \overset{\text{ع}_{\cdot}\,\text{ك}_1}{\cancel{\text{ع}_1\,\text{ك}_{\cdot}}}$ ، وبديلها الزمنى $\text{مـ}_1 = \text{ع}_1\,\text{ك}_{\cdot}$ ؛

و $\text{س} = \frac{\text{ع}_1}{\text{ع}_{\cdot}}$ و « $\text{س}^{-1} = \frac{\text{ع}_{\cdot}}{\text{ع}_1}$.

$$\therefore \; \text{لو هب} = \frac{1}{\text{مجـ}_{\cdot}}\left(\text{مـ}_{\cdot}\,\text{لو س} + \text{مـَ}_{\cdot}\,\text{لو سَ} + \text{مـً}_{\cdot}\,\text{لو سً} + \cdot\cdot\right) \;..\;(٣)؛$$

$$\text{و} \; -\text{لو هب}_1 = \frac{1}{\text{مجـ}_1}\left(\text{مـ}_1\,\text{لو س} + \text{مـَ}_1\,\text{لو سَ} + \text{مـً}_1\,\text{لو سً} + \cdot\cdot\right) \;..\;(٤).$$

وهنا أيضاً نرى أن المناسيب الكبيرة مرجحة بأوزان كبيرة ، والعكس بالعكس . وينتج من ذلك أن الطرف الأيسر من (٤) أكبر من نظيره فى (٣) .

$$\therefore \; \text{لو هب . هب}_1 = \text{كمية سالبة}$$

ويكون $\text{هب . هب}_1 < ١$.

أى أن الوسط المرجح **هب** متحيز إلى أسفل .

٣٣٨ — وبنفس البرهان تثبت أن الوسطين **هج** و **هر** متحيزان إلى أعلى ، أى أن الوسطان هج و هب متحيزان إلى أعلى

$$\text{هج} \times \text{هج}_1 > ١ \; ؛$$

$$\text{و} \quad \text{هر} \times \text{هر}_1 > ١ \; ،$$

حيث
$$\text{هج} = \sqrt[\text{مجـ}_1]{(\text{س})^{-\text{مـ}_1} \times (\text{سَ})^{-\text{مـَ}_1} \times \dots}$$

وبديله الزمنى
$$\text{هج}_1 = \sqrt[\text{مجـ}_{\cdot}]{(\text{س})^{-\text{مـ}_{\cdot}} \times (\text{سَ})^{-\text{مـَ}_{\cdot}} \times \dots}$$

وكذلك
$$\text{هر} = \sqrt[\text{مجـ}_1]{(\text{س})^{\text{م}_1} \times (\text{سَ})^{\text{مَ}_1} \times \dots}$$

وبديله الزمنى $هـ_{١} = \sqrt[ع م٠]{(س)^{-م٠} \times (\bar{س})^{-\bar{م}٠} \times ...}$

٣٣٩ — واضح من هذه البراهين أن تحيز الوسط الهندسى ناتج عن الأوزان المستعملة ، ولاسيما وأننا نعرف أن الوسط الهندسى من الأصل ينعكس فى الزمن إذا لم يرجح بأى نوع من الأوزان . وخاصة الانعكاس فى الزمن التى يمتاز بها الوسط الهندسى البسيط ، هى السبب فى تفضيله على غيره من الأرقام القياسية فى كثير من الأبحاث الخاصة بالأسعار . وقد ذكرنا سابقاً أنه الرقم القياسى المطبق عملياً فى كثير من الدول لبيان حركة أسعار **الجملة** .

عدم الانعكاس فى المعامل أكثر تعقيداً

٣٤٠ — تكلمنا هنا عن الأخطاء التى تنشأ عن عدم انعكاس الرقم القياسى فى الزمن . ولم نبحث فى الأخطاء الأخرى الناشئة عن عدم الانعكاس فى المعامل . والبحث فى هذه الأخطاء أكثر تعقيداً من البحث الذى أوردناه فى البنود الأخيرة من هذا الباب .

الاختبار الدورى

٣٤١ — عدا الاختبارين اللذين بحثناهما فى هذا الباب ، يوجد اختبار ثالث يسمى [1] **الاختبار الدورى** . والفكرة الأساسية فى هذا الاختبار تتلخص فيما يأتى : إذا كان لدينا أسعار سلعة واحدة فى ثلاث سنين (أو أماكن) معينة ، السنين ١ و ٢ و ٣ مثلا ، وكانت مناسيب أسعارها لهذه السنين هى :

$$س_{٢١} \text{ ، } س_{٣٢} \text{ ، } س_{٣١}$$

حيث الرقم الأول يدل على السنة المعتبرة أساساً ، فإننا نجد دائماً أن :

$$س_{٢١} \times س_{٣٢} = س_{٣١} \text{ ؛}$$

أى أن $\quad س_{٢١} \times س_{٣٢} \times س_{١٣} = ١ \text{ .}$

(١) بالانجليزية (Circular Test)

والاختبار الدورى يرمى إلى أن الرقم القياسى لأسعار عدة سلع يجب أن يستوفى هذا الشرط . أى أنه ، بفرض ع هى الرقم القياسى ، يجب أن يكون

$$ع_{٢١} \times ع_{٣٢} \times ع_{١٣} = ١$$

الاختبار الدورى معقول شكلا

٣٤٢ — هذا الاختبار يبدو معقولا فى ظاهره ، ولكنه من الناحية النظرية خطأ ولا يمكن أن يستوفى هذا الشرط بأى رقم قياسى صالح . بل الواجب أن الرقم القياسى الصحيح لا يستوفى هذا الاختبار . وذلك لأن أوزان السلع التى تدخل فى تركيب الرقم القياسى فى السنة ٢ بالنسبة إلى السنة ١ مثلا ، تخالف الأوزان المستعملة لنفس السلع بين السنتين ٢ و ٣ ، أو السنتين ٣ و ١ . هذا فضلا عن أن السلع نفسها قد تختلف ، حيث تدخل سلع جديدة وتسقط سلع قديمة . خصوصاً عند مقارنة الأسعار بين الممالك المختلفة (بدل السنين) ، حيث يجوز أن تكون السلع المهمة فى المقارنة بين البلدين ١ و ٢ مثلا ، والتى تدخل فى تركيب الرقم القياسى بينهما ، غير مهمة ، أو غير موجودة بالمرة ، فى المقارنة بين المملكتين ٢ و ٣ أو بين ٣ و ١ . وعلى ذلك لا يصح نظرياً أن يتحقق الاختبار الدورى للأرقام القياسية بين هذه البلاد . والواجب أن هذا الاختبار لا يتحقق فى مثل هذه الأحوال . وأى رقم قياسى يتحقق فيه هذا الاختبار يكون **غير صحيح** ، ويجب رفضه وعدم الاعتماد عليه بالمرة .

الاختبار الدورى خطأ نظرياً . وهو صحيح فقط إذا كانت الأوزان واحدة

أما إذا كانت الأوزان المستعملة واحدة فى الحالتين فالشرط يتحقق ، ولكن هذا نادر جداً لا فائدة منه عملياً .

الأرقام القياسية الجيدة قريبة من تحقيق الاختبار

٣٤٣ — وبالرغم من هذا ، نجد فى الواقع أن أحسن الأرقام القياسية وأصلحها من الناحية النظرية ، أقرب من غيرها تحقيقاً لهذا الاختبار الخاطئ ، ولكنها لا تحققه تماماً بالطبع . فمثلا نجد [1] أن الرقم الأمثل ، الذى سبقت

(١) انظر كتاب I. Fisher *The Making of Index Numbers*, 1927, p. 277

الإشارة إليه — وهو أحسن الأرقام القياسية المعروفة وأصحها وأدقها — أدنى[١] إلى تحقيق هذا الاختبار من بعض الأرقام القياسية الأخرى . ولكنه لا يحقق الاختبار تماماً . ويلاحظ أيضاً أن الأرقام القياسية الرديئة بعيدة عن تحقيق هذا الاختبار أى أن حاصل الضرب

$$ع_{٢١} \times ع_{٣٢} \times ع_{١٣}$$

يختلف كثيراً عن ١ ، وهو القيمة المطلوبة حسب شروط الاختبار .

والخلاصة أن وجود فرق كبير بين هذا الحاصل والواحد الصحيح ، أو عدم وجود الفرق بالمرة ، يدل على خطأ الرقم القياسى وعدم صلاحيته للاستعمال .

(١) أى أن حاصل الضرب $ع_{٢١} \times ع_{٣٢} \times ع_{١٣}$ أقرب إلى ١ فى حالة الرقم الأمثل منه فى حالة الأرقام الأخرى .

تعديل الأرقام القياسية

٣٤٤ — عرفنا فى هذا الباب الطرق التى نستعملها لاختبار الأرقام القياسية من حيث جودتها وصلاحيتها للتعبير بدقة عن الفكرة التى نرمى إليها باستخدام الرقم القياسى . و بتطبيق اختبارى الانعكاس فى الزمن وفى العامل على الأرقام القياسية المعروفة لنا ، أمكننا أن نحكم على كل منها ، فرفضنا بعضها واستبقينا البعض الآخر .

ولكن هذين الاختبارين كانا فى الواقع بمثابة امتحان قاس لهذه الأرقام ، فكانت النتيجة أن رفضنا جميعها ، ما عدا واحداً فقط ، ألا وهو الرقم الأمثل الذى يقترحه فيشر . فهل من سبيل إلى تصحيح هذه الأرقام المرفوضة حتى تستوفى هذين الاختبارين ؟

جميع الأرقام القياسية يمكن تعديلها حتى تنعكس فى الزمن أو فى العامل . وسنشرح هنا طرق التعديل ونطبقها على بعض الأرقام القياسية المعروفة .

تعريف المقلوب الزمنى

٣٤٥ — **المقلوب الزمنى**[١] لأى رقم قياسى هو عبارة عن خارج قسمة الواحد الصحيح على بديله الزمنى . فمثلا الرقم التجميعى بأوزان ك. نعلم أنه لا ينعكس فى الزمن ؛ ومعادلته هى كما نعلم :

$$\frac{\sum ع_{١} ك_{٠}}{\sum ع_{٠} ك_{٠}}$$

ومعادلة بديله الزمنى هى

$$\frac{\sum ع_{٠} ك_{١}}{\sum ع_{١} ك_{١}}$$

(١) اسمه بالإنجليزية (Time Antithesis)

فينتج أن المقلوب الزمنى معادلته هى :

$$\frac{\text{مج ع}_1\,\text{ك}_1}{\text{مج ع}_0\,\text{ك}_1}\,.$$

وكذلك الوسط الحسابي للمناسيب المرجح بأوزان م_0 مثلا ، وبديله الزمنى ، معادلتاهما على الترتيب هما :

$$\frac{\text{مج س م}_0}{\text{مج م}_0} \quad \text{و} \quad \frac{\text{مج}\left(\frac{\text{م}_1}{\text{س}}\right)}{\text{مج م}_1}\,.$$

∴ معادلة المقلوب الزمنى لهذا الوسط هى :

$$\frac{\text{مج م}_1}{\text{مج}\left(\frac{\text{م}_1}{\text{س}}\right)}\,.$$

مقلوب الزمنى يساوى الرقم نفسه إذا كان منعكسا

وإذا كان الرقم الأصلى ينعكس فى الزمن ، كان المقلوب الزمنى مساوياً للرقم نفسه . ويكون شرط الانعكاس فى الزمن هو أن الرقم يساوى مقلوبه الزمنى ، أى ١ ÷ البديل الزمنى .

فالرقم التجميعى البسيط مثلا ، نعلم أنه ينعكس فى الزمن ، ومعادلته هى :

$\frac{\text{مج ع}_1}{\text{مج ع}_0}$ ، وبديله الزمنى معادلته هى $\frac{\text{مج ع}_0}{\text{مج ع}_1}$

∴ مقلوبه الزمنى يساوى $\frac{\text{مج ع}_1}{\text{مج ع}_0}$ = الرقم الأصلى نفسه .

٣٤٦ — الوسط الهندسى بين أى رقم قياسى ومقلوبه الزمنى ، ينعكس فى الزمن حتى ولو كان الرقم الأصلى لا ينعكس . فمثلا :

الرقم $\frac{\text{مج ع}_1\,\text{ك}_0}{\text{مج ع}_0\,\text{ك}_0}$ ، ومقلوبه الزمنى يساوى $\frac{\text{مج ع}_1\,\text{ك}_1}{\text{مج ع}_0\,\text{ك}_1}$

والوسط الهندسى لهذين هو :

$$\sqrt{\frac{\text{مج ع}_1\,\text{ك}_0}{\text{مج ع}_0\,\text{ك}_0} \times \frac{\text{مج ع}_1\,\text{ك}_1}{\text{مج ع}_0\,\text{ك}_1}}$$

الانعكاس فى الزمن خاصة لازمة للمقلوب الزمنى

وهو الرقم الأمثل المعروف لنا . ونعلم أنه ينعكس فى الزمن . وهكذا فى أى رقم آخر نختاره ، نجد أن الجذر التربيعى لحاصل ضربه فى مقلوبه الزمنى ينعكس فى الزمن . والحقيقة أن هذه الخاصة ناتجة مباشرة من تعريف المقلوب الزمني . إذ لو كان ا أى رقم قياسى ، وكان مقلوبه الزمنى ب ، فإن المقلوب الزمنى لهذا الأخير هو ا نفسه . لأنه حسب تعريف المقلوب الزمنى يكون

$$\text{ب} = \frac{1}{\text{بديل ا الزمنى}}\ ،$$

$$\therefore\ \text{بديل ب الزمنى} = \frac{1}{\text{ا}}\ ،$$

$$\text{أى أن} \quad \text{ا} = \frac{1}{\text{بديل ب الزمنى}}\ ،$$

$$= \text{مقلوب ب الزمنى}\ ،$$

وبناء على ذلك يكون المقلوب الزمنى لحاصل الضرب ا ب يساوى ب ا . ويكون إذن $\sqrt{\text{ا ب}}$ رقماً ينعكس فى الزمن ، حيث إن مقلوبه الزمنى يساوى نفسه .

وهكذا يمكننا تعديل أى رقم قياسى لكى ينعكس فى الزمن ، بأن نوجد الوسط الهندسى بينه وبين مقلوبه الزمنى . وهذا الوسط الهندسى نفسه يكون هو الرقم القياسى المعدل .

المقلوب المعاملى

٣٤٧ — وبمثل هـذه الطريقة نعدل الأرقام القياسية لـكى تنعكس فى المعامل . فنوجد **المقلوب المعاملى**[1] للرقم ؛ فيكون الوسط الهندسى بين الرقم ومقلوبه المعاملى مستوفياً لشرط الانعكاس فى المعامل . والمقلوب المعاملى لأى رقم هو خارج قسمة منسوب القيم $\frac{\text{مج ع}_1\ \text{ك}_1}{\text{مج ع}_0\ \text{ك}_0}$ على البديل المعاملى للرقم نفسه .

لنأخذ مثلا الرقم التجميعى المرجح بأوزان ك_0 ؛ وهو كما نعلم لا ينعكس فى المعامل . هذا الرقم وبديله المعاملى هما على الترتيب :

$$\frac{\text{مج ع}_1\ \text{ك}_0}{\text{مج ع}_0\ \text{ك}_0} \quad \text{و} \quad \frac{\text{مج ك}_1\ \text{ع}_0}{\text{مج ك}_0\ \text{ع}_0} \ .$$

∴ المقلوب المعاملى ، على حسب التعريف ، يكون

$$= \frac{\text{مج ع}_1\ \text{ك}_1}{\text{مج ع}_0\ \text{ك}_0} \div \frac{\text{مج ك}_1\ \text{ع}_0}{\text{مج ك}_0\ \text{ع}_0}$$

$$= \frac{\text{مج ع}_1\ \text{ك}_1}{\text{مج ع}_0\ \text{ك}_1}$$

ويكون الوسط الهندسى بين هذا الأخير والرقم الأصلى ، يساوى

$$\sqrt{\frac{\text{مج ع}_1\ \text{ك}_0}{\text{مج ع}_0\ \text{ك}_0} \times \frac{\text{مج ع}_1\ \text{ك}_1}{\text{مج ع}_0\ \text{ك}_1}}$$

وهذا هو الرقم الأمثل الذى نعرف أنه ينعكس فى المعامل .

الانعكاس فى المعامل نتيجة تعريف المقلوب

٣٤٨ — وهذه الخاصة أيضاً ، مثل خاصـة الانعكاس فى الزمن ، نتيجة مباشرة لتعريف المقلوب المعاملى . فلنفرض مثلا أن ا هو أى رقم قياسى وأن حـ هو مقلوبه المعاملى .

$$\therefore\ \text{حـ} = \frac{\text{مج ع}_1\ \text{ك}_1}{\text{مج ع}_0\ \text{ك}_0} \div (\text{البديل المعاملى للرقم الأصلى ا}) \ \dots \ (1)$$

(١) يسمى بالانجليزية (Factor Antithesis)

∴ البديل المعاملى للرقم حـ

$$= \frac{\text{مج ك}_1\ \text{ع}_1}{\text{مج ك}_0\ \text{ع}_0} \div (\text{الرقم ا نفسه}) \qquad (2)$$

$$\therefore\ \text{ا} = \frac{\text{مج ع}_1\ \text{ك}_1}{\text{مج ع}_0\ \text{ك}_0} \div (\text{البديل المعاملى للرمز حـ}) \qquad (3)$$

= المقلوب المعـاملى للرمز حـ .

أى أن ا و حـ كل منهما المقلوب المعاملى للآخر . وينتج من ذلك أن المقلوب المعاملى لحاصل الضرب ا حـ هو حـ ا . إذن يكون الوسط الهندسى $\sqrt{\text{ا حـ}}$ رقماً قياسياً ينعكس فى المعامل ، لأن :

$$\sqrt{\text{حـ ا}} \times \text{بديله المعاملى}$$

$$= \sqrt{\text{ا حـ}} \times \sqrt{\text{بديل ا المعاملى} \times \text{بديل حـ المعاملى}}$$

$$= \frac{\text{مج ع}_1\ \text{ك}_1}{\text{مج ع}_0\ \text{ك}_0} \ ،$$ انظر المعادلتين (١) و (٢) أعلاه ؛

= منسوب القيم .

تعديل الرقم القياسى من حيث الزمن والمعامل

٣٤٩ — بينا أن الوسط الهنـدسى بين أى رقم ومقلوبه الزمنى يسـتوفى شرط الانعكاس فى الزمن ، حتى ولو كان الرقم الأصلى لا يستوفى هذا الشرط . وكذلك بينا أن الوسط الهنـدسى بين أى رقم قياسى ومقلوبه المعاملى ينعكس فى المعامل ، ولو كان الرقم الأصلى لا ينعكس . وبذلك توصلنا إلى طريقة لتعديل أى رقم لا ينعكس فى الزمن أو فى المعـامل . أما إذا كان الرقم لا ينعكس فى الزمن ولا فى المعامل ، فتلزم معالجته مرتين : الأولى بالنسبة إلى الزمن ، والثانية بالنسبة للمعامل . وهـذا يكون بإيجاد الوسط الهنـدسي بين هذا الرقم ومقلوبه الزمنى ، فينتج لنا رقم ينعكس فى الزمن . ثم نعالج هذا بالنسبة للمعامل ، فنوجد

الوسط الهندسى بينه و بين مقلوبه المعاملى . وهذا الأخير ينعكس فى المعامل وفى الزمن أيضاً . أى أن المعالجة الأخيرة بخصوص الانعكاس فى المعامل لا تؤثر فى صفة الانعكاس فى الزمن المكتسبة من المعالجة الأولى .

ولإثبات ذلك نفرض أن ا أى رقم قياسى لا ينعكس فى الزمن ولا فى المعامل ، ونفرض أن مقلوبه الزمنى هو ـ ، فينتج أن $\sqrt{\text{ا ـ}}$ ينعكس فى الزمن .

ونفرض أن اَ هو بديل معاملى ا ؛ وأن ـَ هو بديل معاملى ـ . فينتج أن ـَ هو مقلوب زمنى اَ ؛ لأن

$$\text{ـَ} = \text{بديل معاملى ـ ،}$$

$$= \text{بديل معاملى للمعادلة } \left(\frac{1}{\text{بديل زمنى ا}} \right) .$$

$$= \text{بديل زمنى للمعادلة } \left(\frac{1}{\text{بديل معاملى ا}} \right)$$

$$= \text{بديل زمنى } \left(\frac{1}{\text{اَ}} \right) = \frac{1}{\text{بديل زمنى اَ}}$$

$$= \text{مقلوب زمنى اَ}$$

∴ $\sqrt{\text{اَ ـَ}}$ ينعكس فى الزمن ؛ وهو بديل معاملى $\sqrt{\text{ا ـ}}$.

∴ مقلوب معاملى $\sqrt{\text{ا ـ}}$ هو خارج قسمة منسوب القيم على $\sqrt{\text{اَ ـَ}}$.

∴ الوسط الهندسى بين $\sqrt{\text{ا ـ}}$ ومقلوبه المعاملى

$$= \left[\sqrt{\frac{\text{ا ـ}}{\text{اَ ـَ}}} \times \frac{\text{ح ع}_1\text{ ك}_1}{\text{ح ع. ك.}} \right]^{\frac{1}{2}} .$$

هذا ينعكس فى المعامل ، طبقا لبند ٣٤٧ . وهو ينعكس أيضاً فى الزمن ، لأن كل جزء من أجزائه : $\sqrt{\text{ا ـ}}$ و $\sqrt{\text{اَ ـَ}}$ ومنسوب القيم ، ينعكس فى الزمن على حدته . وهو إذن ينعكس فى الزمن وفى المعامل فى نفس الوقت .

ترتيب التعديل بالنسبة إلى الزمن أو المعامل ، غير مهم

٣٥٠ — وسواء فى ذلك إذا أجرينا التعديل بالنسبة إلى الزمن أولا ثم بالنسبة إلى المعامل ، أو العكس . والنتيجة التى نصل إليها فى الحالتين واحدة كما يتضح مما يأتى :

لنفرض أيضاً أن ا أى رقم قياسى لا ينعكس فى الزمن ولا فى المعامل ، وأن اَ بديله المعاملى ؛ وأن ـ ، ـَ هما المقلوبان الزمنيان لهما على الترتيب .

∴ الوسط الهندسى بين ا ومقلوبه المعاملى

$$\left[\frac{\text{ا}}{\text{اَ}} \times \frac{\text{ح ع}_1\text{ ك}_1}{\text{ح ع. ك.}} \right]^{\frac{1}{2}} ;$$

وهذا ينعكس فى المعامل . والمقلوب الزمنى لهذا الأخير هو

$$\left[\frac{\text{ـ}}{\text{ـَ}} \times \frac{\text{ح ع}_1\text{ ك}_1}{\text{ح ع. ك.}} \right]^{\frac{1}{2}} .$$

∴ الوسط الهندسى بين هذين هو

$$\left[\sqrt{\frac{\text{ا ـ}}{\text{اَ ـَ}}} \times \frac{\text{ح ع}_1\text{ ك}_1}{\text{ح ع. ك.}} \right]^{\frac{1}{2}} ،$$

وهذا الرقم ينعكس فى المعامل وفى الزمن معاً ، وقد حصلنا عليه بتعديل ا بالنسبة للمعامل أولا ، ثم تعديل الناتج بالنسبة إلى الزمن . ولو عالجنا ا بالنسبة إلى الزمن أولا لحصلنا على $\sqrt{\text{ا ـ}}$. و بتعديل هذا بالنسبة إلى المعامل نحصل على

$$\left[\sqrt{\frac{\text{ا ـ}}{\text{اَ ـَ}}} \times \frac{\text{ح ع}_1\text{ ك}_1}{\text{ح ع. ك.}} \right]^{\frac{1}{2}} \text{، وهى نفس النتيجة السابقة .}$$

تعديل صيغ الأرقام بسبب تعقيدها وصعوبتها

٣٥١ — يمكننا إذن أن نعدل أى رقم قياسى حتى ينعكس فى الزمن وفى المعامل أيضاً ، وهذا التعديل هو من حيث الصيغة أو التركيب . وبذلك يمكننا أن نجعل جميع صيغ الأرقام القياسية التى عرفناها صالحة للاستعمال .

ولـكن هذا التعديل — مع الأسف — يعطينا فى جميع الأحوال صيغا أكثر تعقيداً فى التركيب ، وأشد إرهاقاً فى الحساب ، من الصيغ الأصلية المقصود تعديلها. وهذا التعقيد وماينشأ عنه من زيادة فى أعباء العمل الحسابى واحتمال الخطأ فيه ، ينال من فائدة هذا التعديل و يقلل من أهميته من الناحية العملية . ولا يسلم من هـذا التعقيد إلا الصيغ التجميعية التى ذكرناها فى بند ٢٨٦ ، حيث ينتج من تعديلها الرقم الأمثل ، أو

$$\sqrt{\frac{\sum ع_{١} ك_{٠}}{\sum ع_{٠} ك_{٠}} \times \frac{\sum ع_{١} ك_{١}}{\sum ع_{٠} ك_{١}}} \ .$$

أما من حيث السرعة فى العمل الحسابى فهذا الرقم أفضل من أغلبية الأرقام الأرقام كلها (٧٥ ./. منها تقريباً)(١) .

تعديل الأرقام بتهجين أوزانها

٣٥٢ — نبحث الآن فى تعديل صيغ الأرقام القياسـية بمعالجة الأوزان المستعملة لترجيح الأسعار الداخلة فى تركيب هذه الأرقام .

وجدنا فى الصيغ التجميعية (بند ٢٨٥ وبند ٢٨٦) أننا نرجح الأسعار إما بالـكميات $ك_{٠}$ أو بالـكميات $ك_{١}$ ؛ ووجـدنا فى كل حالة أن الصيغة الناتجة من استعمال واحد أو الآخر من هذين النظامين لاتنعكس فى الزمن ولا فى العامل . ويمكننا تحسين هذه النتيجة لو استعملنا « هجناً » من هذين النوعين من الأوزان . فقد رأينا مثلا (بند ٣٢٠) أن الصيغة

$$\frac{\sum ع_{١} (ك_{٠} + ك_{١})}{\sum ع_{٠} (ك_{٠} + ك_{١})}$$

تنعكس فى الزمن ؛ فضلا عن أنها تأخذ فى الاعتبار ظروف السنتين المقارنة والأساسية من حيث تحديد الأوزان . ولو أنها لا تنعكس فى العامل ولـكنهـا أسهل فى الحساب من الرقم الأمثل .

(١) أنظر I. Fisher *The Making of Index Numbers* (1327).

التهجين الهندسى والتوافقى

٣٥٣ — ويمكننا تهجين الأوزان بطريقـة أخرى ، كأن نأخذ الوسط الهندسى أو التوافقى بين $ك_{٠}$ و $ك_{١}$ بدل الوسط الحسابى الذى أخذناه هنـا . ولـكن العمل الحسابى اللازم هنـا يكون أصعب منه لو أخذنا الوسط الحسابى . وفى كل هذه الأحوال نحصل على صيغ تنعكس فى الزمن ، مثل الصيغة المذكورة .

وكذلك فى الصيغ الأخرى غير التجميعية ، مثل الوسط الحسابى أو الهندسى (البسيط أو المرجح) لمناسيب الأسعار ، يمكننا تهجين الأوزان . فنعرف أن الأوزان المستعملة فى هذه الصيغ هى :

$$م_{٠} = ع_{٠} ك_{٠} \quad و \quad م_{١} = ع_{١} ك_{١} \ ،$$

$$و \quad مـ_{٠} = ع_{٠} ك_{١} \quad و \quad مـ_{١} = ع_{١} ك_{٠} \ .$$

والتهجين بين الأوزان $م_{٠}$ و $م_{١}$ ، أو بين $مـ_{٠}$ و $مـ_{١}$ ، يعطى نتـائج تنعكس فى الزمن ، سواء كان التهجين على نظام الوسط الحسابى أو التوافقى أو الهندسى . وأما التهجين بين م و مـ فينتج صيغاً لا تنعكس فى الزمن . وهذا واضـح ، حيث أن حاصـل جمع (أو ضرب) $ع_{٠} ك_{٠}$ و $ع_{١} ك_{١}$ — أو $ع_{٠} ك_{١}$ و $ع_{١} ك_{٠}$ — عبارة عن كمية متماثلة بالنسـبة إلى الرقمين ٠ و ١ ، يوجد أحدهما حيثما وكيفما يوجـد الآخر . فلا تتغير هذه الـكمية إذا ما أخذ كل منهما مكان الآخر عند تبديل سنة الأساس بسنة المقارنة . بخلاف حاصل جمع (أو ضرب) $ع_{٠} ك_{١}$ و $ع_{٠} ك_{٠}$ — أى $مـ_{٠}$ مع $م_{١}$ أو $مـ_{١}$ مع $م_{٠}$ — فهو كمية غير متماثلة بالنسبة للرقمين ٠ و ١ ، فنرى هنا مثلا أن صفراً موجود ثلاث مرات اثنتين مع ع وثالثه مع ك ، بينما ١ يوجد مرة فقط مع ك ولا يوجـد مع ع أبداً . وهذه الـكمية تتغير عند تبديل السنة الأساسية بالسنة المقارنة ، فلا ينعكس الرقم فى الزمن .

تعديل الأرقام بتهجين معادلاتها

٣٥٤ — ويمكننا تعديل الأرقام القياسية لكى تنعكس فى الزمن ، وفى الوقت نفسه نجمع بين ميزات الأوزان المختلفة ، بإجراء التهجين بين الصيغ نفسها المستعملة فيها هذه الأوزان ، بدلا من تهجين الأوزان . والتهجين هنا أيضاً يكون إما حسابياً ، أو توافقياً ، أو هندسياً ، حيث نوجد الوسط الحسابى أو التوافقى أو الهندسى بين الصيغتين المراد تهجينهما ، للجمع بين ميزات الأوزان فيهما . وأحسن مثال لذلك هو تهجين الرقمين التجميعيين المرجحين بالكميات ك_0 و ك_1 وهما

$$\frac{\text{مجـ ع}_1\,\text{ك}_0}{\text{مجـ ع}_0\,\text{ك}_0} \quad \text{و} \quad \frac{\text{مجـ ع}_1\,\text{ك}_1}{\text{مجـ ع}_0\,\text{ك}_1}$$

حيث نتج من تهجينهما هندسياً الرقم الأمثل :

$$\sqrt{\frac{\text{مجـ ع}_1\,\text{ك}_0}{\text{مجـ ع}_0\,\text{ك}_0} \times \frac{\text{مجـ ع}_1\,\text{ك}_1}{\text{مجـ ع}_0\,\text{ك}_1}}$$

تهجين المعادلات هندسيا أفضل

٣٥٥ — التهجين الهندسى هنا أفضل من الحسابى أو التوافقى . لأن عمليات الجمع المستعملة فى هذين الأخيرين تكون عقبة كبيرة فى اختصار الكسور التى تقابلها فى القسمة على البديلات الزمنية أو المعاملية ، لاستخراج المقلوبات الزمنية والمعاملية للأرقام . وهذا مما يمنع تحقيق شرط الانعكاس المنشود .

٣٥٦ — يمكننا إذن أن نعدل صيغ الأرقام القياسية لكى تنعكس فى الزمن بتهجين الأوزان أو الصيغ نفسها . وهذه الطريقة تخالف طريقة المقلوب الزمنى ، ولكننا نحتاج فيها إلى انتقاء الصيغ أو الأوزان المناسبة لكى تنتج من التهجين صيغ تنعكس فى الزمن ، بخلاف الطريقة الأولى فهى صريحة تؤدى إلى النتيجة المطلوبة مباشرة ، حيث نوجد البديل الزمنى للرقم المراد تعديله ، ونقسم هذا الأخير على البديل الزمنى ، ونستخرج الجذر التربيعى ، فنحصل على الرقم الجديد المعدل .

المراجع

FISHER, I., *The Making of Index Numbers*, Chapters IV, V, VII, VIII, XIII.

MILLS, F.C., *Statistical Methods*, Chapter VI.

RIETZ, H. *Handbook of Mathematical Statistics*, Chapter XII.

الباب الثالث عشر

المفاضلة بين الأرقام القياسية

٣٥٧ — عرفنا فى الباب الحادى عشر الصيغ المختلفة لتركيب الأرقام القياسية ؛ وفى الباب السابق بحثنا فى صلاحية كل من هذه الصيغ لتأدية الغرض المقصود من إنشاء الرقم القياسي . وعند تطبيق الاختبارات التى أجريناهـا على هـــذه الأرقام القياسية حصلنا على مجموعة جديدة من الصيغ : المقلوب الزمنى والمقلوب المعاملى لـكل الصيغ التى اختبرناها أو حاولنا تعديلها .

وكل واحدة من هذه الصيغ أمكننا تعديلها وتصحيحها لـكى تستوفى شروط الانعكاس فى الزمن وفى العامل ، وبذلك أصبحت جميعها صالحة للاستعمال — نظريا . وتبقى الآن مسألة المفاضلة بين هذه الأرقام ، واختيار أحسنها من الوجهة النظرية والعملية أيضاً .

الصيغ الأصلية للأرقام القيـــاسية

٣٥٨ — يمكن تقسيم صيغ الأرقام القياسـية التى حصلنا عليها إلى ست مجموعات رئيسية حسب نوع المتوسط المستعمل فى تركيبها . وهذه المجموعات هى :

١ — مجموعة الوسط الحسابى لمناسيب الأسعار :

(١) بسيط ، (٢) مرجح بأوزان م$_{٠}$ ؛ و (٣) مرجح بأوزان مـ$_{٠}$ ؛ و (٤) مرجح بأوزان مـ$_{١}$ ؛ و (٥) مرجح بأوزان م$_{٠١}$.

ولـكل من هذه الخمسة مقلوب معاملى ؛ فتـكون جملتها عشرة .

٢ — مجموعة الوسط التوافقى لمناسيب الأسعار :

بسيط أو مرجح بأوزان م$_{٠}$ أو مـ$_{٠}$ أو مـ$_{١}$ أو م$_{٠١}$.

ولـكل من هذه الخمسة مقلوب معاملى ، وجملتها عشرة أيضاً :

٣ — مجموعة الوسط الهندسى للمناسيب .

بسيط ومرجح بأوزان م$_{٠}$ أو مـ$_{٠}$ أو مـ$_{١}$ أو م$_{٠١}$.

ولـكل منها مقلوب معاملى .

٤ — مجموعة الوسيط للمناسيب :

بسيط أو مرجح[1] بأى نوع من الأوزان المذكورة ومقلوباتها المعاملية .

٥ — مجموعة المنوال للمناسيب :

بسيط أو مرجح[1] ، ومقلوباتها المعاملية .

٦ — الأرقام التجميعية للأسعار .

(١) بسيط ؛ و (٢) مرجح بأوزان ك$_{٠}$ ، و (٣) مرجح بأوزان ك$_{١}$.

ولـكل من هذه مقلوب معاملى أيضاً .

عدد الصيـغ الأصـــلية

٣٥٩ — ومن هذا البيان يظهر أن عدد الصيغ الأصلية يساوى ٢٨ ، يشتق منهـا ٢٨ أخرى كمقلوبات معاملية لها ، فيكون المجموع ٥٦ . ولـكن يجب أن نلاحظ أن بعض الصيغ مشتركة ، فقد لاحظنا مثلا فى بند ٢٩٠ أن الصيغة (٢) من المجموعة ١ ، تطابق الصيغة (٢) من المجموعة ٦ . ومثال ذلك كثير بين المجموعات المذكورة . ولذلك فهذه الصيغ ليست كلها مختلفة .

الصيغ الناتجة من التعـديل

٣٦٠ — وعند ما نعدل هــذه الصيغ لتنعكس فى الزمن — ما عـدا التى تنعكس من نفسها مثل الهندسى البسيط والتجميعى البسيط — نحصل على مجموعة أخرى من الأرقام القياسية عـددها حوالى الخمسين . وعند ما نعدلها بالنسبة إلى

(١) انظر معنى الوسيط المرجح والمنوال المرجح فى باب التوسطات بند ١٥٨

المعامل نحصل على مجموعة ثالثة ؛ وعند تعديلها بالنسبة للزمن والمعامل معاً نحصل على مجموعة رابعة من الصيغ . وهكذا عند ما نجرى التعديلات الآخرى مثل تهجين الأوزان أو تهجين الصيغ مع بعضها ، هندسياً أو حسابياً أو توافقياً ، يزداد العدد إلى ١٧٠ وهذا ينقص إلى ١٣٤ بعد استبعاد المكررات[١] .

٣٦١ — من بين هذه المجموعة الكبيرة نريد اختيار رقم واحد يجمع بين السهولة فى الحساب والسلامة من التعقيد فى الصورة والخطأ النظرى ، بحيث يكون حساساً يظهر التغيرات التى تطرأ على الأسعار من وقت لآخر ، ويقيسها باعتدال وبدون تحيز .

رفض الوسيط والمنوال ومشتقاتهما

٣٦٢ — وبالتأمل فى المجموعات الست التى ذكرناها فى بند ٣٥٨ وما يشتق منها ، لا نتردد فى أن نرفض مجموعتى الوسيط والمنوال بسبب ضعف حساسيتهما . لأن وسيط مجموعة من مناسيب الأسعار (أو المنوال) هو فى الواقع منسوب سلعة واحدة فى وسط السلسلة . وقد علمنا أن الوسيط لا يتأثر مهما تغيرت القيم التى تحته ، ما دامت فى تغيرها لا تزال أقل منه : ومهما تغيرت القيم التى فوقه ، ما دامت لا تزال أكبر منه . فمجموعة المناسيب الآتية مثلا :

٨٥ ، ٩٢ ، ٩٩ ، ١٠٦ ، ١١١ ، ١١٩ ، ١٢٠ ، ١٢١ ، ١٢٥

وسيطها ١١١ . وهذا الوسيط يبقى كما هو لو تغيرت المناسيب الأخرى وأصبحت كما يأتى مثلا :

٩٣ ، ٩٤ ، ١٠٢ ، ١٠٣ ، ١١١ ، ١١٢ ، ١٢٥ ، ١٣١ ، ١٤٥

ومن الواضح أن مستوى الأسعار فى المجموعة الثانية أرفع بكثير منه فى الأولى . وبالرغم من ذلك فان الوسيط لا « يحس » بهذا الفرق الكبير ولا

(١) انظر كتاب فيشر المشار إليه فى ص٢٠٢ .

يتأثر به . فيجب إذن أن نرفض الوسيط ولا نعتمد عليه فى قياس تغيرات مستوى الأسعار . وكذلك نرفض المنوال وكل ما يشتق منه ومن الوسيط .

رفض الصيغ غير المرجحة

٣٦٣ — نرفض أيضاً الأرقام ذات الصيغ البسيطة غير المرجحة ، حتى ولو كانت معدلة لتستوفى الاختبارين . لأنها لا تفاضل بين السلع المختلفة بما يتناسب وأهميتها . وبديهي أننا لو عرفنا الكميات $ك_{٠}$ و $ك_{١}$ للسلع فالأولى أن نستخدمها كأوزان لانشاء رقم قياسى مرجح يأخذ فى الاعتبار الأهمية النسبية للسلع المختلفة ، بدل استخدامها فى تعديل رقم بسيط لينعكس فى الزمن أو فى المعامل ، وهو لا يفرق بين السلع الهمة وغيرها .

تطبيق الاختبارين

٣٦٤ — الخطوة التالية فى عملية الاختيار واضحة ، ألا وهى تطبيق اختبارى الانعكاس فى الزمن وفى المعامل . فنرفض جميع الصيغ التى لا تحقق شروط هذين الاختبارين معاً . وبذلك يبقى عدد صغير جداً من الأرقام القياسية — ١٣ فقط[١] . وهذه الثلاثة عشر هى نتيجة تعديل الأرقام الآتية فى الزمن وفى المعامل معاً :

الوسط الحسابى للمناسيب مرجح بأوزان $م_{٠}$ أو $م_{١}$ أو $\sqrt{م_{٠} م_{١}}$ ؛ وهجن بين الأول والثانى ؛ والوسط الهندسى بأوزان $م_{٠}$ أو $م_{١}$ أو $\sqrt{م_{٠} م_{١}}$ وهجن بين الأول والثانى ؛ والرقم التجميعى بأوزان $ك_{٠}$ أو $(ك_{٠} + ك_{١})$ أو الوسط التوافقى بين $ك_{٠}$ و $ك_{١}$ ، أو الهندسى $\sqrt{ك_{٠} ك_{١}}$ ؛ أو بأوزان $\frac{م_{٠} + م_{١}}{ع_{٠} + ع_{١}}$.

وهذه الأرقام غير متساوية طبعاً ؛ ولكنها تعطى نتائج قريبة جداً من

(١) انظر كتاب فيشر صفحة ٢١٩ ، وصفحات ٤٦٥ — ٤٨٥ حيث توجد معادلات هذه الأرقام وغيرها . ويرى القارئ معادلات هذه الأرقام الثلاثة عشر فى آخر الكتاب صفحة ٣٧٥ .

بعضها ، بحيث يمكننا اعتبارها متساوية عمليا لو أهملنا الفروق الصغيرة فى المنازل العشرية البعيدة .

الرقم الأمثل افضل من الجميع

٣٦٥ — وبما أن جميع هذه الأرقام تستوفى اختبارى الانعكاس فى الزمن، فالمفاضلة بينها تكون على أساس البساطة فى التركيب والسهولة فى العمل والحساب. ولو تأملنا فى هذه الصيغ (صفحة ٣٧٥) وجدنا أن أبسطها تركيباً وأسهلها جميعاً هى الصيغة التجميعية التى عرفناها بالرقم الأمثل ، وهى

$$\sqrt{\frac{\text{مـ ع}_{1}\text{ ك}_{0}}{\text{مـ ع}_{0}\text{ ك}_{0}} \times \frac{\text{مـ ع}_{1}\text{ ك}_{1}}{\text{مـ ع}_{0}\text{ ك}_{1}}} \;.$$

ومن حيث السرعة فى الحساب فهى أسرع من الجميع .

بعض الاعتبارات العملية

٣٦٦ — هكذا نختار بين الأرقام القياسية من حيث الصيغة أو طريقة التركيب . ولكن صيغة الرقم القياسى ليست كل شىء فيه ، بل توجد اعتبارات أخرى وشروط يجب توافرها فيه قبل أن نعتمد عليه اعتماداً كلياً فى تصوير حركة مستوى الأسعار (أو الإنتاج أو أى ظاهرة أخرى) . وأهم هذه الاعتبارات الأخرى من الناحية العملية هى :

١ — ما هى السلع التى تدخل فى تركيب الرقم ، وما هى السلع التى يمكن إهمالها .

٢ — عدد السلع التى نأخذها .

٣ — دقة البيانات الخاصة بالأسعار والأوزان .

٤ — السرعة فى الحساب .

نوع السلع التى يدخل فى الرقم

٣٦٧ — وبخصوص نوع السلع التى ندخلها فى تركيب الرقم ، يجب أن نأخذ فى الحسبان الغرض الذى من أجله ننشىء الرقم القياسى ، ونوع الرقم الذى

نريده فإذا كنا مثلا ننشىء رقماً قياسياً لأسعار الجملة ، لاستخدامه للدلالة على حركة السوق العامة والنشاط التجارى على العموم ، فيجب أن تمثل السلع التى نأخذها جميع السلع المتداولة ذات الأهمية فى السوق ، فلا تتحيز للسلع الصناعية مثلا ضد السلع الاستهلاكية ، أو للسلع المستوردة ضد السلع الوطنية أو العكس ، أو المنتجات الزراعية ضد المنتجات الصناعية ، وهكذا .

أما إذا أردنا رقماً قياسياً لأسعار التجزئة ، فلا يدخل فيه إلا السلع التى تباع بالتجزئة أى السلع الاستهلاكية . وهذه يكون أغلبها سلعاً تامة الصنع جاهزة للاستهلاك ، أو سلعاً زراعية استهلاكية مثل القمح والبطاطس واللحوم وغيرها .

تقسيم السلع إلى مجموعات متجانسة

٣٦٨ — ويحسن تقسيم السلع هنا إلى مجموعات ، تمتاز المفردات فى كل منها بصفات خاصة ، ذات أهمية فى الناحية التى يتناولها الرقم القياسى المراد إنشاؤه . ففى الرقم القياسى العام لأسعار الجملة مثلا ، تقسم السلع إلى مجموعات المواد الغذائية والمواد الخام المستعملة فى الصناعة ، والمواد النصف المصنوعة ، والمنتجات الجاهزة ، ومجموعة منتجات التعدين . ويصح أن تقسم أى واحدة من هذه المجموعات إلى مجموعات فرعية زيادة فى التفصيل .

وعلى كل حال فالأساس الذى نبنى عليه التقسيم إلى المجموعات يختلف حسب الغرض الذى يستعمل فيه الرقم القياسى . فيصح أن يكون التقسيم مبنياً على أساس نوع المادة الرئيسية فى السلع ؛ كأن نقول مثلا مجموعة منتجات الحديد والصلب ، ومجموعة لمنتجات المعادن الأخرى ، ومجموعة لمنتجات القطن ، أو الصوف أو غيره . أو أن يكون على أساس طوائف المستهلكين من أفراد أو صناعات ؛ أو أن يكون التقسيم على أساس جغرافى ، كما نفعل أحياناً حيث تقسم السلع إلى مستوردة ووطنية .

في كل مجموعة نأخذ السلع غير المتشابهة في الحركة

٣٦٩ — وعند اختيار السلع من أي مجموعة، يجب أن تكون السلع المختارة تمثل الحركات أو الاتجاهات المختلفة للأسعار في هذه المجموعة. فاذا كان هناك عدة سلع، تسير أسعارها دائماً في اتجاه واحد هبوطاً أو صعوداً، فيكفي أن نأخذ منها واحدة أو اثنتين لتمثيلها، بدل أن نأخذها جميعها وندخلها في الرقم القياسي. لأن إدخال الكل لا يؤثر في النتيجة أي تأثير يذكر بجانب المشقة الحسابية الناتجة عن ذلك.

أما السلع التي تتغير أسعارها في اتجاهات مختلفة أو بنسب مختلفة، فهذه يجب إدخالها في الرقم القياسي، حتى تتمثل فيه هذه الاتجاهات المتباينة، فيكون أدق في تصوير الحالة على حقيقتها.

الترجيح غير المباشر بأخذ عدة تسعيرات لنفس السلعة

٣٧٠ — ومن الواضح أن كثرة السلع المأخوذة من أي مجموعة تعمل على تعزيز هذه المجموعة وترجيح التغيرات التي تحصل في أسعار سلعها. وبالمثل إذا أخذنا جملة تسعيرات لنفس السلعة وأدخلنا هذه التسعيرات في الرقم القياسي، كما لو كانت تسعيرات لسلع مختلفة، فإن هذا يكون بمثابة **ترجيح غير مباشر** (١) لهذه السلعة وتعزيز لأهميتها على غيرها.

عدد السلع في الرقم القياسي

٣٧١ — وبدهي أن عدد السلع يزيد في دقة الرقم القياسي، ويجعله أقرب إلى الصحة في تمثيل الحركة العامة لمستوى الأسعار. ويجب إذن أن يكون عدد السلع التي ندخلها في الرقم القياسي كبيراً، بحيث لا يقل عن حوالى ٣٠ سلعة. ومعلوم طبعاً أن زيادة السلع ينشأ عنها زيادة كبيرة في العمليات الحسابية

(١) هذا هو المتبع في إنشاء الرقم القياسي الجديد لأسعار الجملة في مصر. انظر الاحصاء السنوي العام سنة ١٩٣٥ — ١٩٣٦ صفحة ٤٩٠، حيث يسمونه «تثقيلاً غير مباشر»؛ وهي في رأيي ترجمة ركيكة للعبارة الانجليزية (Indirect Weighting).

خصوصاً في الأرقام القياسية ذات الصيغ المعقدة. على أن الزيادة في دقة الرقم القياسي الناشئة عن زيادة عدد السلع لا تتناسب معها تماماً. حيث إذا زاد عدد السلع إلى أربعة أمثاله، نقص الخطأ المحتمل للرقم القياسي إلى النصف فقط. أي أن دقته تزيد إلى الضعف فقط. وكذلك إذا زاد عدد السلع إلى تسعة أمثاله، زادت الدقة إلى ثلاثة أمثال فقط. والسبب في ذلك أن الخطأ في الرقم القياسي — حسب نظرية الاحتمالات — يتناسب مع $\frac{1}{\sqrt{\text{ن}}}$، حيث ن يساوي عدد السلع.

وعلى ذلك لا يجدي كثيراً أن نزيد عدد السلع طلباً للدقة؛ ويكفي في المسائل العادية أن نأخذ حوالى ٥٠ سلعة، على ألا يكون العدد أقل من ٣٠. وليس هناك أي فائدة — عملياً — من أخذ عدد يزيد على ٢٠٠ سلعة.

الخطأ في بيانات الأسعار أعظم خطراً من الخطأ في تقدير الأوزان

٣٧٢ — إذا حصل خطأ في جمع البيانات الخاصة بالأسعار فلا بد أن يظهر تأثير في النتيجة النهائية للرقم القياسي المحسوب من هذه البيانات. أما إذا كان هناك خطأ في الأوزان المستعملة فانه لا يؤثر كثيراً في النتيجة. وذلك لأن الوزن المضروب في السعر أو في منسوبه، يوجد في البسط وفي المقام أيضاً. فاذا تغير بالزيادة مثلا، زاد البسط والمقام معاً — زيادة مختلفة على العموم — وكان تأثير ذلك في قيمة الكسر صغيراً نسبياً؛ وكذلك إذا نقص الوزن. أما إذا أخطأنا في السعر أو منسوب السعر، فإن البسط يتغير وحده دون المقام، فيكون التغير في قيمة الكسر أكبر منه في الحالة الأولى.

وعلى ذلك يجب أن نعتني كل العناية في جمع بيانات الأسعار، وحساب مناسيبها بدقة؛ ولا بأس من تقريب الأوزان إلى أعداد صحيحة لتسهيل العمليات الحسابية.

الأرقام التجميعية أسرع فى الحساب من غيرها

٣٧٣ — أما من حيث السرعة فى العمليات الحسابية فنجد على العموم أن الأرقام القياسية ذات الصيغ التجميعية أسهل فى حسابها من الأرقام الأخرى . وأسهلها جميعاً وأسرعها فى الحساب هو الرقم التجميعى البسيط ؛ ويليه فى ذلك الرقم الذى معادلته :

$$\frac{\sum ع_1 (ك_0 + ك_1)}{\sum ع_0 (ك_0 + ك_1)} \text{ ؛}$$

وهو يجمع بين السرعة والانعكاس فى الزمن . وكذلك يعطى للسلع أهميتها حسب أوزان معتدلة ، تأخذ فى الاعتبار ظروف السنة الأساسية وظروف السنة المقارنة ، ولكنه لا ينعكس فى المعامل .

ويلى ذلك فى السرعة الرقم الأمثل، وهو كما نعلم يستوفى شروط الانعكاس ، ولكنه أكثر تعقيداً من هذا .

جدول معادلات الأرقام القياسية التى تنعكس فى الزمن وفى المعامل المذكورة فى صفحة ٣٦٩ ، وعددها ١٣ .

فى هذه المعادلات قد وضعنا الرموز الآتية للاختصار :—

$$س = \frac{ع_1}{ع_0} = \text{منسوب السعر ، } ص = \frac{ك_1}{ك_0} = \text{منسوب الكمية}$$

$$\text{و}\quad ن = \frac{\sum ع_1 ك_1}{\sum ع_0 ك_0} = \text{منسوب القيم ،}$$

$$\text{و}\quad م_0 = ع_0 ك_0 \text{ ، } مـ_0 = ع_0 ك_1 \text{ ، } مـ_1 = ع_1 ك_0 \text{ ، } م_{11} = ع_1 ك_1$$

المجموعة الأولى : الوسط الحسابى مرجحاً بأوزان $مـ_1$ أو $م_{11}$ أو $\sqrt{م_0 م_{11}}$ ، وهجن بين الأول والثانى ، كل منها معدلا فى الزمن والمعامل .

وهذه هى على الترتيب :

$$(1) - \left[\frac{\sum س\, مـ_1}{\sum مـ_1} \times \left(ن \div \frac{\sum ص\, مـ_0}{\sum مـ_0}\right) \times \frac{\sum مـ_0}{\sum \left(\frac{مـ_0}{س}\right)} \times \left(ن \div \frac{\sum مـ_1}{\sum \left(\frac{مـ_1}{ص}\right)}\right)\right]^{\frac{1}{4}}$$

$$(2) - \left[\frac{\sum س\, م_{11}}{\sum م_{11}} \times \left(ن \div \frac{\sum ص\, م_{11}}{\sum م_{11}}\right) \times \frac{\sum م_0}{\sum \left(\frac{م_0}{س}\right)} \times \left(ن \div \frac{\sum م_0}{\sum \left(\frac{م_0}{ص}\right)}\right)\right]^{\frac{1}{4}}$$

$$(3) - \left[\frac{\sum س \sqrt{م_0 م_{11}}}{\sum \sqrt{م_0 م_{11}}} \times \left(ن \div \frac{\sum ص \sqrt{م_0 م_{11}}}{\sum \sqrt{م_0 م_{11}}}\right) \times \frac{\sum \sqrt{م_0 م_{11}}}{\sum \left(\frac{\sqrt{م_0 م_{11}}}{س}\right)} \times \left(ن \div \frac{\sum \sqrt{م_0 م_{11}}}{\sum \left(\frac{\sqrt{م_0 م_{11}}}{ص}\right)}\right)\right]^{\frac{1}{4}}$$

(٤) - $\sqrt{(1)\times(2)}$ السابق ذكرها .

المجموعة الثانية : الوسط الهندسى مرجحاً بأوزان $م_0$ أو $م_1$ أو $\sqrt{م_0 م_1}$ ، وهجن بين الأول والثانى ، كل منها معدلا فى الزمن والمعامل .

وهى على الترتيب :

$$(5)-\left[\sqrt[\text{مج}\,م_0]{(س)^{م_0}\times(\bar{س})^{\bar{م}_0}\times\dots}\times ن\Big\{\div\sqrt[\text{مج}\,م_0]{(ص)^{م_0}\times(\bar{ص})^{\bar{م}_0}\times\dots}\Big\}\right.$$

$$\left.\times\sqrt[\text{مج}\,م_1]{(س)^{م_1}\times(\bar{س})^{\bar{م}_1}\times\dots}\times ن\Big\{\div\sqrt[\text{مج}\,م_1]{(ص)^{م_1}\times(\bar{ص})^{\bar{م}_1}\times\dots}\Big\}\right]^{\frac{1}{4}}$$

$$(6)-\left[\sqrt[\text{مج}\,\text{ـ}_0]{(س)^{\text{ـ}_0}\times(\bar{س})^{\bar{\text{ـ}}_0}\times\dots}\times ن\Big\{\div\sqrt[\text{مج}\,\text{ـ}_0]{(ص)^{\text{ـ}_0}\times(\bar{ص})^{\bar{\text{ـ}}_0}\times\dots}\Big\}\right.$$

$$\left.\times\sqrt[\text{مج}\,\text{ـ}_1]{(س)^{\text{ـ}_1}\times(\bar{س})^{\bar{\text{ـ}}_1}\times\dots}\times ن\Big\{\div\sqrt[\text{مج}\,\text{ـ}_1]{(ص)^{\text{ـ}_1}\times(\bar{ص})^{\bar{\text{ـ}}_1}\times\dots}\Big\}\right]^{\frac{1}{4}}$$

$$(7)-\left[\sqrt[\text{مج}\sqrt{م_0م_1}]{(س)^{\sqrt{م_0م_1}}\times\dots\times}\; ن\Big\{\div\sqrt[\text{مج}\sqrt{م_0م_1}]{ص^{\sqrt{م_0م_1}}\times\dots}\Big\}\right]^{\frac{1}{2}}$$

(٨) - $\sqrt{(5)\times \cancel{(5)}\,(6)}$ السابق ذكرها .

المجموعة الثالثة : الرقم التجميعى مرجحاً بأوزان $ل_0$ ، أو بالوسط الحسابى بين $ل_0$ و $ل_1$ ، أو بالوسط الهندسى ، أو التوافقى بينهما ، أو بالأوزان $\frac{م_0+م_1}{ع_0+ع_1}$ ، كل منها معدلا فى الزمن والعامل .

وهى على الترتيب :

(٩) — $\sqrt{\frac{\text{مج}\,ع_1ك_0}{\text{مج}\,ع_0ك_0}\times\frac{\text{مج}\,ع_1ك_1}{\text{مج}\,ع_0ك_1}}$ ، الرقم الأمثل

$$(10)-\left[\frac{\text{مج}\,ع_1(ك_0+ك_1)}{\text{مج}\,ع_0(ك_0+ك_1)}\times\Big\{ن\div\frac{\text{مج}\,ك_1(ع_0+ع_1)}{\text{مج}\,ك_0(ع_0+ع_1)}\Big\}\right]^{\frac{1}{2}}$$

$$(11)-\left[\frac{\text{مج}\,ع_1\sqrt{ك_0ك_1}}{\text{مج}\,ع_0\sqrt{ك_0ك_1}}\times\Big\{ن\div\frac{\text{مج}\,ك_1\sqrt{ع_0ع_1}}{\text{مج}\,ك_0\sqrt{ع_0ع_1}}\Big\}\right]^{\frac{1}{2}}$$

$$(12)-\left[\frac{\text{مج}\,ع_1\left(\frac{2}{\frac{1}{ك_0}+\frac{1}{ك_1}}\right)}{\text{مج}\,ع_0\left(\frac{2}{\frac{1}{ك_0}+\frac{1}{ك_1}}\right)}\times\left\{ن\div\frac{\text{مج}\,ك_1\left(\frac{2}{\frac{1}{ع_0}+\frac{1}{ع_1}}\right)}{\text{مج}\,ك_0\left(\frac{2}{\frac{1}{ع_0}+\frac{1}{ع_1}}\right)}\right\}\right]^{\frac{1}{2}}$$

$$(13)-\left[\frac{\text{مج}\,ع_1\left(\frac{م_0+م_1}{ع_0+ع_1}\right)}{\text{مج}\,ع_0\left(\frac{م_0+م_1}{ع_0+ع_1}\right)}\times\left\{ن\div\frac{\text{مج}\,ك_1\left(\frac{م_0+م_1}{ك_0+ك_1}\right)}{\text{مج}\,ك_0\left(\frac{م_0+م_1}{ك_0+ك_1}\right)}\right\}\right]^{\frac{1}{2}}$$

المراجـــع

CONNOR, L.R., *Statistics in Theory and Practice*, Chapter XX.

FISHER, I., *The Making of Index Numbers*, Chapters XV-XVII.

MILLS, F.C., *Statistical Methods*, Chapter VI.

دليـــل

بأسماء الأعلام والاصطلاحات الافرنجية

تصحيح أخطاء مطبعية

الصفحة	السطر	خطأ	صواب
٧	٣ هامش	الطبيعة	الطبيعية
١٤	٨	بترك	نترك
١٥	٢	هو مستوى	هى مستوى
٩٧	٢ من أسفل	**مفصلا**	**منفصلا**
١٥٢	٦ « «	المعامله	المعادلة
١٥٤	٤	١٧٩ر٥	١٧ر٩٥
١٥٥	٢ « «	ك ك$_{١}$ (فى بسط الكسر)	ك — ك$_{١}$
١٥٧	٦	طالب	طالباً
١٥٨	٦	بينهما	بينها
١٩٣	٤	بند ١٦٤	بند ١٧٥
١٩٥	٨	مجـ س — ن $\overline{\overline{\text{س}}}$ = ٠	مجـ س — ن $\overline{\text{س}}$ = ٠
١٩٥	١٣	(س$_{٢}$ — و) =^{٢} (٠٠٠)	(س$_{٢}$ — و)^{٢} = (٠٠٠)
٢٠٤	٣	١٧٣	١٨٣
٢١٧	٣	وبدون	وندون
٢٢٠	٢	التغيير ص	التغير فى ص
٢٣٥	١	٢٠٠ رجل	٢٠٠ رجلا
٢٣٦	١٣	فى حدول ٣١	فى جدول ٣٢
٢٦٠	٤	$\bar{\text{ن}}^{٢}$	ن^{٢}
٢٧١	١١	ر$\bar{\text{ع}}$	ر $\frac{\bar{\text{ع}}}{\text{ع}}$
٢٧٣	١	متوسط القيم	متوسط القيم السينية
٢٧٣	٣	السينية أو نسبة	أو نسبة
٢٧٦	١٣	ح$^{٢}_{ن}$	ح$^{٢}_{٢}$
٢٧٦	١٥	ح$^{٢}_{٢}$	ح$^{٢}_{ن}$
٢٨١	١١	نطق	نطلق
٢٨٤	١ « «	1936	1933