الجمهورية الجزائرية الديمقراطية الشعبية

وزارة التعليم العالي والبحث العلمي

جامعة ابي بكر بلقايد تلمسان

كلية العلوم الاقتصادية والتسيير والعلوم التجارية

مطبوعة في :

دروس في الاحصاء الوصفي (الاحصاء 1)

موجهة الى طلبة :

السنة أولى جذع مشترك السداسي الأول

من إعداد الأستاذ : بوصالح سفيان

استاذ محاضر "أ"

السنة الجامعية 2022- 2023

الفصل الأول : عموميات حول الاحصاء :

1- مقدمة في الاحصاء

1.1- نبذة تاريخية عن تطور الاحصاء :

الإحصاء، بمعنى العدّ والحصر، فكرة قديمة يرجع منشؤها إلى عهد بعيد في تاريخ المدنية الإنسانية، فالحاجة إلى الحصول على معلومات رقمية أو وصفية عن المجتمعات وظروفها المادية وشروط وجودها كانت حاجة ملحّة منذ أن وجدت المجتمعات البشرية المنظمة، وهنالك إحصاءات عند قدماء المصريين والصينيين والإغريق تخص مجتمعاتهم من حيث عدد السكان ومقدار الثروة الزراعية والمعدنية جُمعت للاهتداء بها في تصريف أمور الدولة ورسم سياستها.

1.1.1- الإحصاء عند العرب

ولقد كان العرب المسلمون من أوائل من استعان بلغة الأرقام في إحصاء مواردهم وحصر غنائمهم وجندهم وأعطياتهم وأسلحتهم، ومعرفة الثروات لتحصيل الزكاة عنها. ففي الإحصاء اللغوي نجد الكندي المتوفى سنة 260هـ، يصف في مؤلَّفه «رسالة في استخراج المعمَّى» عملية إحصاء تواتر الحروف في لغة ما، وذلك بأخذ عينة كافية من الكلام المنثور في تلك اللغة، وقد أحصى نصاً مؤلفاً من 3667 حرفاً ثم استعمل تلك النتائج بعد ترتيبها في استنباط نص معمًّى وينبه فيها على أمرٍ ذي بال، وهو أن النص المعمَّى ينبغي أن يكون ذا طول كافٍ يسمح بانطباق القواعد الإحصائية عليه، وهي فكرة رياضية في غاية الأهمية، هي فكرة قانون الأعداد الكبيرة.

ولعل الكندي هو أول من أجرى ذلك الإحصاء في تاريخ الدراسات الكمية على اللغة، ولا شك في أنه أفاد من إحصائيات حروف القرآن الكريم التي سبقت عصره (وهي تعود إلى القرن الهجري الأول، وينسب بعضها إلى صدر الإسلام.

كما كان للعرب في الإحصاء الاجتماعي أيضاً أثرٌ يجدر ذكره، وهو أن المفكر العربي ابن خلدون ربما كان أول من عالج قضايا السكان معالجة علمية، فبحث في عمران الدول واتساعها وتأخرها، وربط كل ذلك بنمو عدد السكان ونقصانهم.

2.1.1- القرن 17

عَرف القرن السابع عشر الميلادي مدرستين للإحصاء، أولاهما المدرسة الألمانية الوصفية، وكان على رأسها هرمان كنرنگ (1606–1681) H. Conring الذي بدأ بتدريس علم جديد سماه علم شؤون الحكومات Staatkunde يتناول دراسة الدولة وما يتعلق بها من أمور كثيرة كالأرض والسكان والثروة وغيرها، وكان ذلك في عام 1660، ولقد كان هذا العلم وصفياً لا يُعنى بالتعبير الرقمي كثيراً، وقد تبع كنْرنغ في هذا العلم أخنوول گوتفرايد1773–1719) G. Achenwall وأطلق على هذا العلم تسمية جديدة هي الإحصاءStatistic ، وهذه التسمية هي التي انتقلت من الألمانية إلى كثير من اللغات الأخرى. وقد نشر آخِنْوولعام 1749 كتاباً حول مبادئ الإحصاء في الدول الأوربية[1].

وكان على رأس المدرسة الثانية التي عرفت باسم مدرسةِ الحسبة السياسيين الإنجليزية» ج. گراونت (1640–1674) J. Grount الذي نشر سنة 1666 كتاباً درس فيه سجلات نفوس لندن وحسَب منها نسب الوفيات، ثم، و. بيتي (1623–1687) W. Petty الذي ألف عام 1683 كتاباً استخدم فيه الطرائق الكمية وسماه «الحساب السياسي»، ومن هذا الاسم أخذت المدرسة الثانية اسمها.[2]

3.1.1– القرن 19

وفي بداية القرن التاسع عشر الميلادي دخل الإحصاء مرحلة من مراحل تطوره على يد لاپلاس) Laplaceع. 1749–1827) الذي يجب أن يوضع في مقدمة أولئك الذين جعلوا من حساب الاحتمالات الأداة الأساسية لدراسة التحليل الإحصائي، وقد أوضح في كتابه «النظرية التحليلية للاحتمالات» عام 1812 الفوائد والميزات التي يمكن أن تستخلص من دراسة الظواهر الطبيعية التي أسبابها معقدة جداً إلى حد لا يمكن معه معرفتها جميعاً. ثم وسّع أ. كيتِلي A. Quetelet (1796–1874)حقل تطبيقات الإحصاء مسترشداً بأعمال لابلاس، فدرس الصفات الفيزيائية والفكرية والنفسية للكائنات البشرية، وأوجد بذلك نوعاً من «الفيزياء الاجتماعية» تتوزع وفقها هذه الصفات المتنوعة للجماهير على كائن اعتباري هو «الرجل الوسط». وبناءً على مبادرة كيتلي

[1] محمد بشير قابيل، سيف الدين بغدادي" .الإحصاء (علم"(- .الموسوعة العربية.Retrieved 2009-02-13 .

[2] David Freedman (2002) (pdf), From Association to Causation: Some Remarks on the History of Statistics, University of California,Berkeley http://www.stat.berkeley.edu/~census/521.pdf

انعقد في بروكسل عام 1853 المؤتمر الدولي الأول للإحصاء الذي بشر «بالمعهد الدولي للإحصاء» الذي أسس في لندن عام 1885.[3]

4.1.1- العصر الحديث

قد توسعت منذ نهاية القرن التاسع عشر وإلى يومنا هذا طرائق التحليل الإحصائي فوصلت إلى كل مجالات التحريات والدراسات العلمية، وأدت المسائل العلمية الحديثة المدروسة وفق هذا الأسلوب إلى تطوير سريع وكبير للنظرية الإحصائية.

فبعد أعمال كيتلي والسير ف. گالتون (1822-1911) F. Galton أنشأ ك. پيرسون K. Pearson (1857-1936) فرعاً جديداً للإحصاء يحمل اسم الإحصاء الحيوي «Biostatistics» الذي امتد حالياً إلى ميادين الاختبارات المتعلقة بعلم المداواة والطب العلاجي.

كذلك فقد توطدت الصلة بين الإحصاء والاقتصاد بابتداع فرع علمي جديد هو الاقتصاد القياسي Econometrics ويدعوه بعضهم الإحصاء الاقتصادي وكان رواده الأوائل أ.أ. كورنو A.A. Cournot (1801-1877)، وڤ. بارتو(1848-1929) V. Pareto ، و ل. ڤالراس L. Walras (1834-1910)، وف. ديڤيزيا(1889-1963) F. Divisia ، و ر.فريش (1895-1973) R. Frish الحائز جائزة نوبل عام 1969 في الاقتصاد السياسي وقد أسس مع ديفيزيا الجمعية الدولية للاقتصاد القياسي.

أما أعمال ج.ك. ماكسويل (1831-1879) J. C Maxwell التي أوصلت إلى النظرية الحركية للغازات فقد كانت نقطة البدء للميكانيك الإحصائي وللفيزياء النووية.

وقد امتدت استخدامات الإحصاء إلى الزراعة على يد ر.أ. فيشر (1890-1962) وإلى الصناعة على يد و. شيوهارت (1891-1967) W.Shewhart بدراسة المراقبة الإحصائية للجودة والوثوق، أما في ميادين العلوم الإنسانية فإن دراسة إ. سپيرمان) E. Spearman ع. 1863-1945) حول سلوك الأفراد، التي طُوِّرت بعد ذلك في علم النفس التطبيقي فقد أدت إلى وجود طرائق التحليل العاملي Factor Analysis الذي هو امتداد منطقي لدراسة الارتباط .Correlation وفي إدارة الأعمال والمشاريع غدت الطرائق الإحصائية عاملاً مساعداً لابدّ منه في دراسة حالة السوق ومراقبة

[3] Probability and Statistics on the Earliest Uses Pages (Univ. of Southampton)

الميزانية وإدارة المخزون الاحتياطي، وهي إضافة إلى نظرية الألعاب game theory ، ونظرية القرار Decision theory وإلى الطرائق الحديثة في الحساب قد مهدّت لولادة بحوث العمليات Operations research.

إن أعمال فيشر وإ.پيرسون (1895 -E. Pearson) وج. نيمان (1894-1981 J. Neyman) حول نظرية الاختبارات test theory ونظرية التقدير estimation theory التي تمخضت عن البحوث التجريبية المتعلقة بتطبيق طرق صبر الاراء، جعلت من الطرق الإحصائية أداة قوية وفعالة في البحث العلمي والتقني ومازال حقل استخدامها في نمو مستمر.[4]

2.1- الهدف من الاحصاء

لا يخفى على أحد الأهمية الكبيرة لعلم الإحصاء ودوره الكبير في خدمة العلوم الأخرى حيث يؤثر ويتأثر، أي يأخذ من العلوم ويوهب لها الأساليب الإحصائية التي يستخدمها الباحثون والمهتمون.

فإن أي تطور يحدث في النظرية الإحصائية لابد أن ينعكس بشكل مباشر على التطبيقات الميدانية والعملية في كافة مجالات الحياة، بدءاً من العلوم الاقتصادية والاجتماعية إلى العلوم الطبيعية.

والهدف من استخدام الأساليب الإحصائية، هو الوصول إلى النتائج الدقيقة، ومساعدة متخذي القرارات للحصول على أفضل قرار بأقل أخطاء ممكنة.

توجد حالة من عدم التأكد في حياتنـــا اليومية وفي حياتنا العامة بشكل لم يسبق له مثيل . فنحن نناقش كل ما توارثناه عن آبائنا وأجدادنا من أفكار وتعليم ,ونطلب إجابة علمية على كل ما نوجهه من أسئلة وحتى نستطيع الإجابة عن هذه الأسئلة وغيرها إجابة علمية صحيحة يجب أن نتبع المنهج الإحصائي .

ويعتبر الإحصاء أداة للتخطيط بالغة الأهمية ومن هنا فقد اهتمت الدول في عالمنا المعاصر بخلق اطارات ذى خلفية علمية متينة متفهم للأساليب الإحصائية المختلفة في التحليل بشكل يؤهله للمساهمة الفاعلة في عملية تخطيط وتنفيذ البرامج الإنمائية . وتتجلى أهمية الإحصاء في التعرف على المجالات الحيوية التي تعتمد على الأساليب الإحصائية في البحث والتحليل . وكذلك تبرز أهمية

- [4] Electronic Journ@l for History of Probability and Statistics/Journ@l Electronique d'Histoire des Probabilités et de la Statistique

علم الاحصاء في البحوث حيث إنه يساعد على تقديم أدق نوع ممكن من الوصف للمعطيات , إذ الوصف والموضوعية من سمات العلم الحديث . [5]

3.1- اهمية الاحصاء

وبغرض التوصل لفهم كيفية اتخاذ الشركات للقرارات، وكيفية اكتشاف علماء الفلك لأنواع جديدة من النجوم، وكيفية تحديد الباحثين في مجال الطب للجينات المرتبطة بمرض معين، وكيفية اتخاذ البنوك قرارا بمنح أو عدم منح شخص ما بطاقة ائتمان، وكيفية تحديد شركات التأمين تكلفة القسط، وما إلى ذلك، فلا يوجد جانب من جوانب الحياة الحديثة لا يمسه علم الإحصاء اين تعتمد الحكومات القديرة على التحليل الإحصائي الدقيق للبيانات في وصف الاقتصاد والمجتمع وبغرض الاستمتاع بهذا العلم سنتطرق الى بعض الامثلة التي يعالجها علم الاحصاء.

رضا العملاء

إن إدارة أي مؤسسة للبيع بالتجزئة على نحوفعال، لتحقق ربحا ونموا مع مرور الوقت، تتطلب إيلاء اهتمام دقيق للعملاء، ومنحهم السلعة أو الخدمة التي يريدونها. والفشل في القيام بذلك يعني أنهم سيتوجهون الى منافس يقدم ما هو مطلوب. ويمكننا محاولة تجنب ذلك من خلال جمع بيانات حول تطاعات العملاء قبل أن يبدءوا بانفاق أموالهم. ويمكننا تنفيذ دراسات مسحية لرضا العملاء،سائلين العملاء ما إذا كانوا سعداء بالسلعة أو الخدمة ، وعن الطرق التي يمكن من خلالها تحسسن ذلك.

للوهلة الأول قد يبدو أنه من الضروري منح الاستبيانات لجميع العملاء من أجل الحصول على نتائج موثوقة تعكس سلوك قاعدة العملاء بأكملها، لكن من الواضح أن هذه عملية تستغرق وقتا طويلا. ومع ذلك، توجد – لحسن الحظ – أساليب إحصائية نمكن من الحصول على نتائج دقيقة بما فيه الكفاية من عينة من العملاء فحسب. وفي الواقع، يمكن أن تكون النتائج أحيائا أكثر دقة من إشراك جميع العملاء. فمن الضروري أن نكون خذرين من بناء استنتاجات على عينة مشوهة؛ فربما ستكون النتائج غير مجدية في وصف كيفية تصرف العملاء عموما ، وعلى هذا الاساس طورت االاساليب الإحصائية التي تمكن من تجنب الأخطاء الناجمة عن احتيار العينة؛ ومن ثم الخلاص استنتاجات صحيحة.

التضخم

[5] علم الاحصاء مفدمة قصيرة جدا ديفيد جيه هند ترجمة احمد شكل مؤسسة هنداوي للتعليم والثقافة 2016 ص 11

إننا جميئا نألف فكرة أن الأشياء تزداد غلاءا بمرور الوقت. ولكن كيف يمكننا مقارنة تكاليف المعيشة اليوم بتكاليف المعيشة أمس؟ للقيام بذلك، نحتاج إل مقارنة الأشياء نفسها التي اشتريناها في اليومين. لكن للأسف، توجد تعقيدات؛ فالمحلات التجارية المختلفة تحدد أسعارا مختلفة للأشياء نفسها، والأشخاص المختلفون يشترون أشياء مختلفة، ويغير الآشخاص أنماط شرائهم، وثظفر منتجات جديدة في السوق وتختفي منتجات قديمة، وما شابه ذلك* كيف نضع مثل هذه التغسرات في الاعتبار عند تحديد ما إذا كانت الحياة أكثر تكلفة هذه الآيام أم لا؟ أنشأ الإحصائيون والاقتصاديون مؤشرات مثل *مؤشر أسعار التجزئة* و*مؤشر أسعار المستهلك* لقياس تكاليف المعيشة. وتستند هذه المؤشرات إلى سلة افتراضية للسلع (مئات منها) التي أشتراها الناس، إضافة إل دراسات استقصائية لاكتشاف الأسعار التى يباع بها كل عنصر في السلة. وتستخدم نماذج إحصائية متطورة لجمع أسعار العناصر المختلفة لتقدم رقما إجماليا واحدا يمكن مقارنته على مدار الزمن. وبالإضافة إلى كونها مؤشزا عن التضخم، تستخدم هذه المؤشرات أيضا لضبط حدود الإعفاء الضريبي والرواتب المرتبطة بالؤشر والمعاشات التقاعدية، وما إلى ذلك.

2- تعريف علم الإحصاء

احد التعريفات الجادة لعلم الإحصاء أنه تكنولوجية استخراج المعنى من البيانات.

هو ذلك الفرع من العلوم الذي يختص بالطرق العلمية التي لجمع وتلخيص وعرض وتحليل البيانات وذلك للوصول الى نتائج مقبولة و قرارات سليمة .[6]

يعرف علم الإحصاء بأنه العلم الذي يهتم بالطرق العلمية المتعلقة بتحديد مصادر البيانات وجمعها وتنظيمها وتلخيصها وعرضها، وتحديد الأساليب المناسبة لاستخراج وتحليل المقاييس والنتائج المساعدة في اتخاذ القرارات الملائمة.

ان الاستخدام التدريجي للبيانات الإحصائية لتمكنيالحكومات من إدارة بلدانها. هو الذي أدى إل ظهور كلمة statistic؛ بمعنى احصائيات؛ فهي بيانات عن الدولة stat. وتمتلك كل الدول المتقدمة الآن مكاتب إحصاء وطنية خاصة بها.

هناك فرق بين الاحصائيات والاحصاء ، فالإحصاء هو العلم الأساسي الشامل، أما الإحصائيات يقصد بها الكلمات الرقمية لعلم الإحصاء.

[6] مبادئ علم الاحصا ء ،وليد عبد الرحمن الفرا ، المملكة السعودية 1425 ه ص 3

وكلمة statistics مشتقة من كلمة status وتعني الدولة باللاتنية.

3- مراحل علم الإحصاء:

- ✓ المشاهدة والملاحظة
- ✓ الفرضية (تفسير الظاهرة)
- ✓ التنبؤ (استنتاجات الباحث)
- ✓ التحقق (التأكد من صحة الفرضية والمعتمدة على تفسير الظاهرة)

4- أنواع علم الإحصاء:

1.4 - الإحصاء الوصفي:(statistique descriptive)

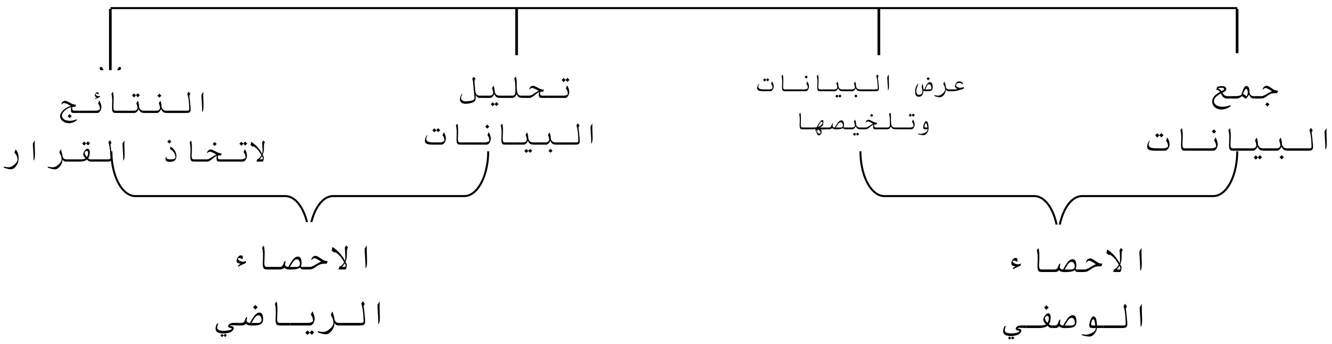
هو الإحصاء الذي يعتمد على وصف ظاهرة ما في فترة زمانية أو مكانية معينة دون الحاجة لتعميمها على ظواهر أخري من خلال الاهتمام بأساليب جمع البيانات وتبويبها وعرضها.

ومن أهم وسائلها:

I. الأشكال الهندسية للبيانات
II. الجداول الإحصائية للبيانات والتوزيعات التكرارية
III. الدراسة الرياضية للبيانات من خلال استخدام مقاييس النزعة المركزية و مقاييس التشتت ومقاييس الالتواء.

2.4- الاحصاء الاستدلالي (statistique mathématique)

ويقصد به مجموعة الطرق العلمية التي تستخدم للاستدلال على معالم مجتمع إحصائي من خلال بيانات إحصائية تم الحصول عليها من عينة سحبت من المجتمع نفسه وفقاً للأساليب الإحصائية المختلفة. وهو ما يبرز أهمية الدور الذي يلعبه هذا الفرع من الإحصاء عند إجراء الدراسات الإحصائية.



5- مصادر جمع البيانات :

تختلف طرق جمع البيانات حسب الهدف من الدراسة وأسلوب التحليل المتبع وبالتالي يوجد هناك عدة مصادر تتبع في حالة جمع البيانات ومن هذه المصادر .

1- مصادر رسمية : جمع البيانات من سجلات صادرة عن هيئات رسمية مثل الوزارات أو المؤسسات العامة (وزارات الصحة , التعليم , العمل , التخطيط , الإقتصاد , البترول, , الإتصلات , الكهرباء , ...) والمنظمات الدولية (صندوق النقد الدولي , البنك الدولي , منظمة الأوبك , منظمة الأغذية والزراعة للأمم المتحدة , ...)

2 – جمع البيانات بطرق غير رسمية : من خلال الحصول على بيانات غير متأكدين من درجة صحتها (الإنترنت , منظمات وشركات خاصة ذات مصالح متعددة !) .

3 – مصادر مباشرة (ميدانية) : جمع البيانات عن ظاهرة اثناء حدوثها في مجال ما على أنها دراسة قائمة على تجربة أو دراسة قائمة على مشاهدة (أسلوب الحصر الشامل , أسلوب المعاينة : الإستمارة الإحصائية ← الاتصال الهاتفي / المقابلة الشخصية / الاستبيان) .

4 – مصادر غير مباشرة (تاريخية) : جمع البيانات من خلال سجلات سبق نشرها مثل البيانات أو الإحصاءات أو النشرات الإحصائية التي تنشرها الوزارات أو الجهات الحكومية المختلفة كل في مجال عملها، بالإضافة إلى المؤسسات المتخصصة الأخرى

6- أسلوب جمع البيانات :

هناك أسلوبين لجمع البيانات

- أسلوب الحصر الشامل (المجتمع)

المجتمع : هو مجموعة من الافراد والوحدات والتي يتم دراستهم احصائيا مسل مجتمع الطلبة او السيارات.

يتناول هذا الأسلوب دراسة كافة مفردات المجتمع الإحصائي مثل (التعداد السكاني) .

مزاياه

1- دقة النتائج 2- عدم وجود أخطاء عشوائية

عيوبها

1- ارتفاع تكاليفها (جهد , المال)

2- عدم إمكانية تطبيقها على المجتمعات ذات المفردات الكبيرة الحجم

المجتمع

- أسلوب المسح العيني (العينة)

العينة : هي مجموعة من الافراد او الوحدات ماخوذة من المجنمع محال الدراسة اين يتم دراستهم احصائيا مثل طلبة جامعة من الجامعات او سيارات ولاية ما .

يتناول هذا الأسلوب اختيار جزء من المجتمع لتمثيله بطريقة العينة العشوائية بشرط أن تكون العينة ممثلة تمثيلا صادق دون تحيز

مزاياه

1- انخفاض تكاليفها 2- إمكانية تطبيقها مهما كان المجتمع

عيوبها

1- نتائجها تقريبية 2- إمكانية الخطأ نتيجة التحيز والاتساق

ويعتمد اختيار الأسلوبين على طبيعة الظاهرة والبيانات الإحصائية المطلوبة .

انواع العينات :

- العينة العشوائية : هو اختيار مجموعة من العناصر تكون فرصة ظهور كل عنصر من المجتمع متساوية مثلا اختيار طالب سواءا كان ذكر او انثى في مجتمع به عدد الاناث اكثر من الذكور هنا الطابة لديهم نفس الفرصة بينما الذكور لديهم اقل فرصة امام الاناث.

- العينة الفرضية : في هذه الحالة يتم اختيار عناصر العينة بما يتناسب اهداف الدراسة مثلا دراسة الاندماج في العمل للطلبة تتطلب دراسة الا الطلبة المتخرجين من الجامعة.

العينة الطبقية : في هذه الحالة يجب مراعاة اوزان الوان المجتمع فمثلا اذا اردنا دراسة نسبة المطالعة عند الطلبة في مجتمع نسبة الاناث 60% ونسبة الذكور 40% فيجب مراعاة هذه النسبة.

7- البيانات الإحصائية وانواعها :

المتغير الاحصائي : هو صفة موجودة وتغير نسبيا من فرد او وحدة لاخرى في العينة او المجتمع مثلا : نوع الهاتف ، تقدير الباكالوريا ، عدد افراد الاسرة ، معدل الباكالوريا

تقسم البيانات إلى مجموعتين :

- بيانات نوعية

وهي البيانات التي تقيس ظاهرة من الظواهر دون أن تأخذ قيما عددية وتقسم البيانات النوعية إلى

أ بيانات نوعية أسمية

تعتمد على التصنيف النوعي بغض النظر عن أهمية الترتيب

مثلا : تصنيف موظفي إحدى الشركات حسب الجنسية أو حسب التخصص .

ب بيانات نوعية ترتيبية

يلعب الترتيب دورا أساسيا في تحديد معالم الظاهرة

مثلا : ترتيب موظفي إحدى الشركات حسب التأهيل

ثانوي – جامعة – ماجستير – دكتوراه

- بيانات كمية

وهي البيانات التي تأخذ قيما عددية صحيحة أو كسرية حسب ظروف الحالة حيث تقسم البيانات الكمية إلى:

أ بيانات كمية متقطعة

وهي البيانات التي تأخذ قيما عددية طبيعية

مثل عدد أفراد الأسرة

ب بيانات كمية مستمرة

هي البيانات التي تعتمد على وحدات القياس التي تأخذ قيم في مجال تغيراتها مثلا

وحدة قياس الطول إما تكون بالمتر أو السنتيمتر .

الفصل الثاني : عرض البيانات الإحصائية :

الخطوة التالية بعد جمع البيانات في مجال الإحصاء الوصفي، هو تبويب البيانات وعرضها بصورة يمكن الاستفادة منها في وصف الظاهرة محل الدراسة، من حيث تمركز البيانات، ودرجة تجانسها.
وهناك طريقتين لعرض البيانات هما:
1-عرض البيانات جدوليا.
2-عرض البيانات بيانيا.

1- جدول التوزيع التكراري :

يتم فيه تنظيم وتلخيص البيانات الوصفية او الكمية بما يسمى التوزيع التكراري حيث يتكون الجدول من عمودين أساسين الأول مخصص للمتغير المدروس ويرمز له ب x_i مع i تأخذ قيم من 1 إلى n وهو عدد الحالات التي لوحظ فيها المتغير . او c_i في حالة الفئات مع i تأخذ قيم من 1 إلى n وهو عدد فئات المتغير.

اما العمود الثاني فيخص التكرارات وهي عدد كل حالة للمتغير او عدد مشاهدات كل حالة ويرمز لها ب n_i مع i تاخذ قيم من 1 الى n وهو عدد الحالات التي لوحظ فيها المتغير.

1.1- عرض بيانات المتغير النوعي الاسمي في شكل جدول تكراري بسيط

مثال 01 :

لدينا البيانات التالية الخاصة بنوع الهاتف ل20 طالب ;E3(sumsung); E2(oppo)
E1(iphone); E10(sumsung); E9(iphone); E8(oppo); E7(oppo);
E6(huawei); E5(condor); E4(lg); E16(sumsung); E15(iphone);
E14(oppo); E13(condor); E12(sumsung); E11(iphone);
E20(iphone); E19(oppo); E18(iphone); E17(lg);

x_i	n_i	f_i	p_i
Iphone	6	$\frac{6}{20} = 0.3$	30%
Sumsung	4	$\frac{4}{20} = 0.2$	20%
Oppo	5	$\frac{5}{20} = 0.25$	25%
Lg	2	$\frac{2}{20} = 0.1$	10%
condor	2	$\frac{2}{20} = 0.1$	10%
Huawei	1	$\frac{1}{20} = 0.05$	5%
$\sum$	20	1	100%

حيث يخصص السطر الاخير في الجدول للمجاميع ويجب دائما ان يكون $\sum n_i$ هو حجم المجتمع او العينة .

عادة ما تكون الاعداد غير واضحة الفهم عند مقارنة حالات المتغير حاصة عندما يكون حجم العينة او المجتمع كبير فنظطر الى حساب التكرار النسبي لغرض توضيح توزيع حلالات المتغير ونرمز لهذا الاخير ب f_i ويحسب بالعلاقة التالية

$$f_i = \frac{n_i}{\sum n_i} = \frac{n_i}{N}$$

حيث n_i هو التكرار المطلق

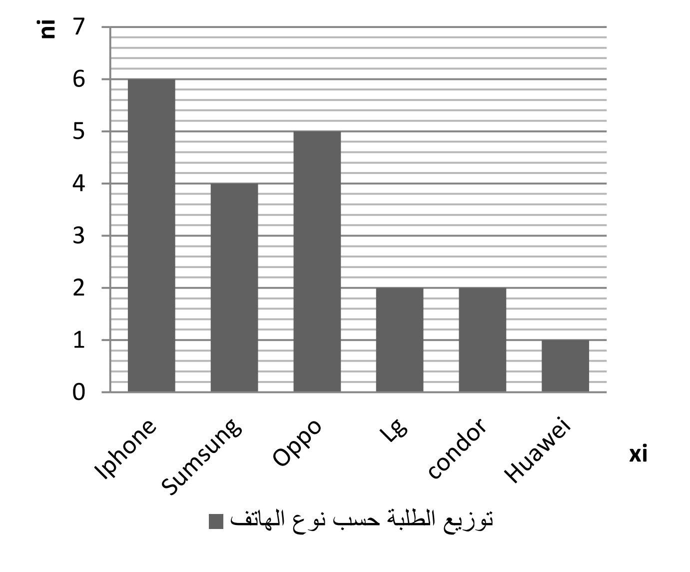
$\sum n_i$ هو حجم العينة او المجتمع

ملاحظة دائما $\sum n_i = 1$

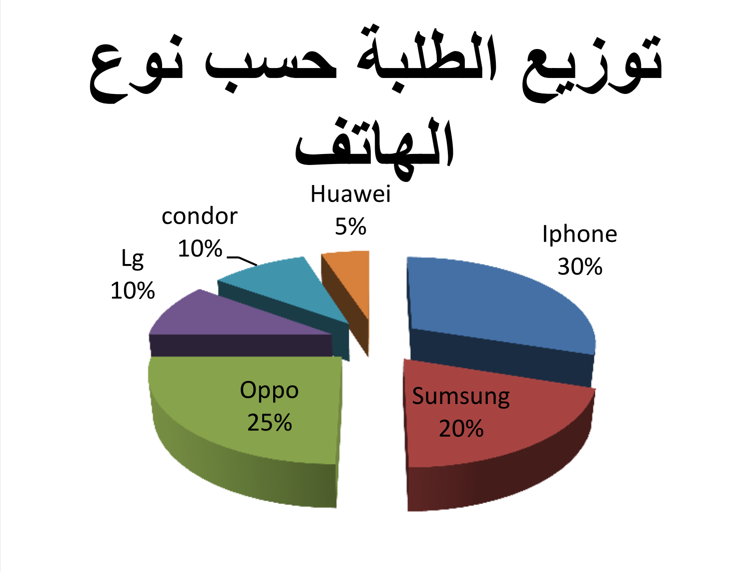
كما بمكن تحويل التكرار النسبي الى نسب مئوية بضريب التكرار النسبي f_i ب 100 ويرمز له ب p_i

التمثيلات البيانية :

الأشرطة البيانية



الدوائر النسبية



2.1- عرض بيانات المتغير النوعي الترتيبي في شكل جدول تكراري بسيط

مثال رقم 02 : لدينا تقديرات 30 طالب في شهادة الباكالوريا على النحو التالي :
مقبول ،قريب من الجيد ،جيد، جيد جدا ،ممتاز، مقبول ،قريب من الجيد، مقبول ،قريب من الجيد، مقبول، جيد، جيد جدا، ممتاز، مقبول ،قريب من الجيد، مقبول، جيد، مقبول ،قريب من الجيد، مقبول، جيد، مقبول، جيد جدا، جيد جدا، جيد، جيد، قريب من الجيد، قريب من الجيد، قريب من الجيد.

اذا كانت المشاهدات يمكن ترتيبها فيشترط ترتيبها تصاعديا أو تنازليا في العمود المخصص للمتغير غالبا ما ترتب تصاعديا من الأصغر إلى الأكبر

x_i	n_i	$n_i \uparrow$	$n_i \downarrow$	f_i	$f_i \uparrow$	$f_i \downarrow$
مقبول	10	10	30	0.33	0.33	1
قريب من الجيد	8	10+8=18	30-10=20	0.27	0.6	0.67
جيد	6	18+6=**24**	20-8=**12**	0.2	**0.8**	**0.4**
جيد جدا	4	24+4=28	12-6=6	0.13	0.93	0.2
ممتاز	2	28+2=30	6-4=2	0.07	1	0.07
$\sum$	30	/	/	1	/	/

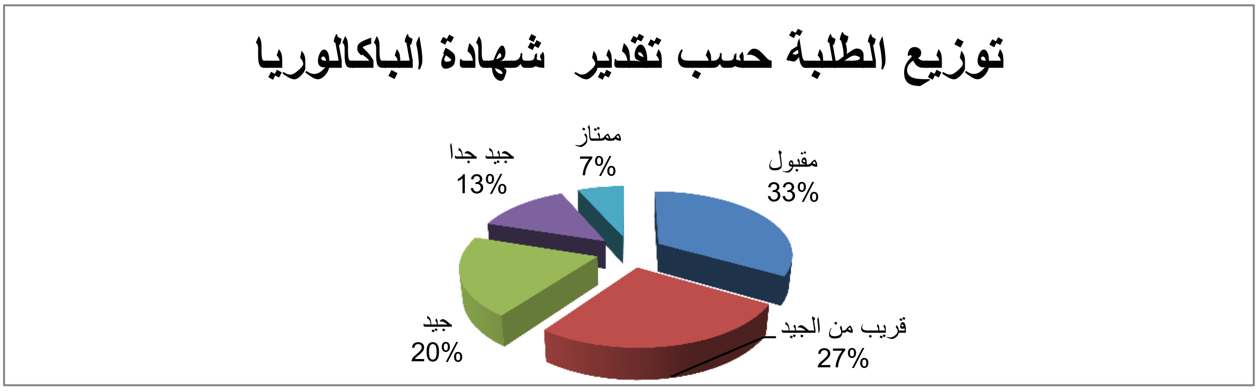
غالبا ما نريد معرفة عدد او نسبة الطلبة الذين تحصلوا على تقدير جيد فاكثر او اقل من الجيد وللإجابة على هذه الأسئلة من نوع اقل أو على الأكثر نجد $n_i \uparrow$ اذا كان عدد او $f_i \uparrow$ اذا كانت نسبة.

أما الإجابة على الأسئلة من نوع اكثر أو على الأقل نجد $n_i \downarrow$ اذا كان عدد او اذا $f_i \downarrow$ كانت نسبة.

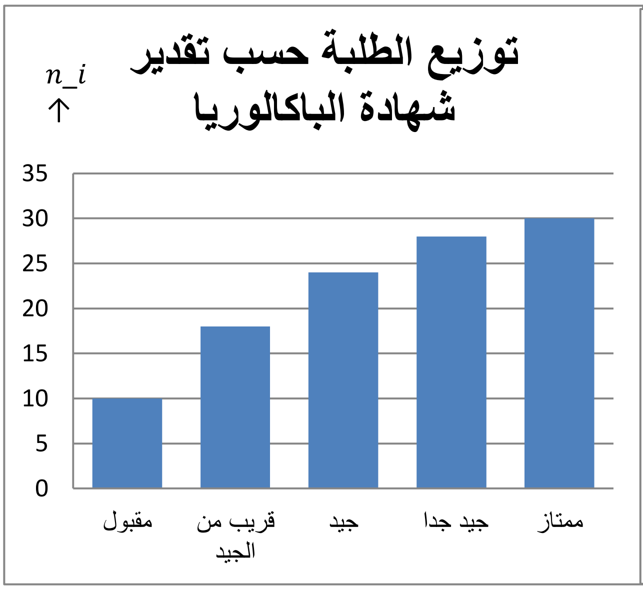
عدد الطلبة الذين تحصلوا على تقدير جيد فاكثر (أو اكثر او يساوي جيد) نلاحظ ان السؤال يتضمن كل الطلبة الذين تحصلوا على تقدير جيد وتقدير جيد جدا وتقدير مقبول 6+8+2 =12 او نذهب مباشرة الى $n_i \downarrow$ في الجدول ونقرأ القيمة التي توافق تقدير جيد .اما نسبة الطلبة الذين تحصلوا على تقدير جيد فاكثر هو حسب $f_i \downarrow$ من الجدول 0.4 اي 40 %.

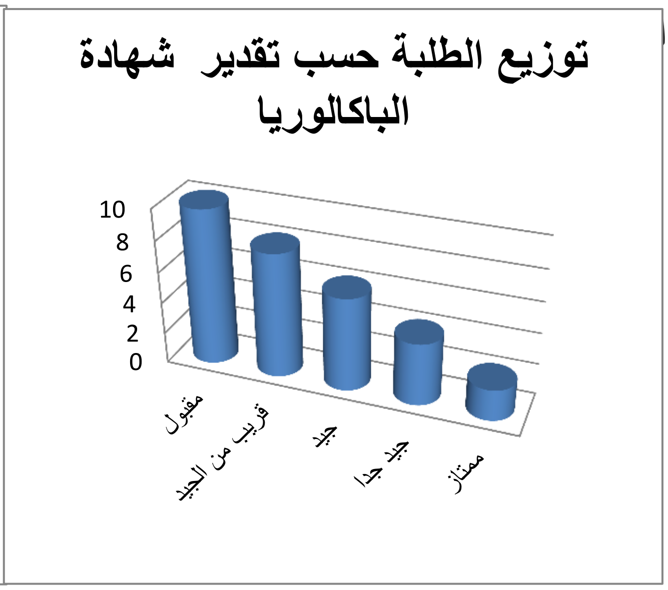
عدد الطلبة الذين تحصلوا على تقدير جيد فاقل (أو اكأقل او يساوي جيد) نلاحظ ان السؤال يتضمن كل الطلبة الذين تحصلوا على تقدير جيد وتقدير قريب من الجيد ا وتقدير ممتاز 6+8+10 =24 او نذهب مباشرة الى $n_i \uparrow$ في الجدول ونقرأ القيمة التي توافق تقدير جيد .اما نسبة الطلبة الذين تحصلوا على تقدير جيد فاكثر هو حسب $f_i \uparrow$ من الجدول 0.8 اي 80 %

الاشكال البيانية



أشرطة بيانية للتكرار الاعمدة البيانية





اما المضلعات او المنحنيات فليس لها معنى ويرجع ذلك لطبيعة المتغير الذي لا يمكن قياسه (لا يوجد سلم على محور xi)

3.1- عرض بيانات المتغير الكمي المتقطع في شكل جدول تكراري بسيط

مثال رقم 03 :

أحد ممثلي البلدية قام بدراسة 20 منزل بهدف معرفة عدد الاجهزة الكهرو منزلية لكل عائلة بغرض تقدير مستوى الرفاهية فتحصل على النتائج التالية :

11،7،8،6،5،9،7،8،9،8،7،9،9،12،10،10،6،7،8،8

حدد المصطلحات الإحصائية لهذه الظاهرة لخص هذه المعطيات في شكل جدول

المجتمع : العائلات العينة : 20 عائلة المتغير : عدد الاجهزة طبيعته : كمي متقطع

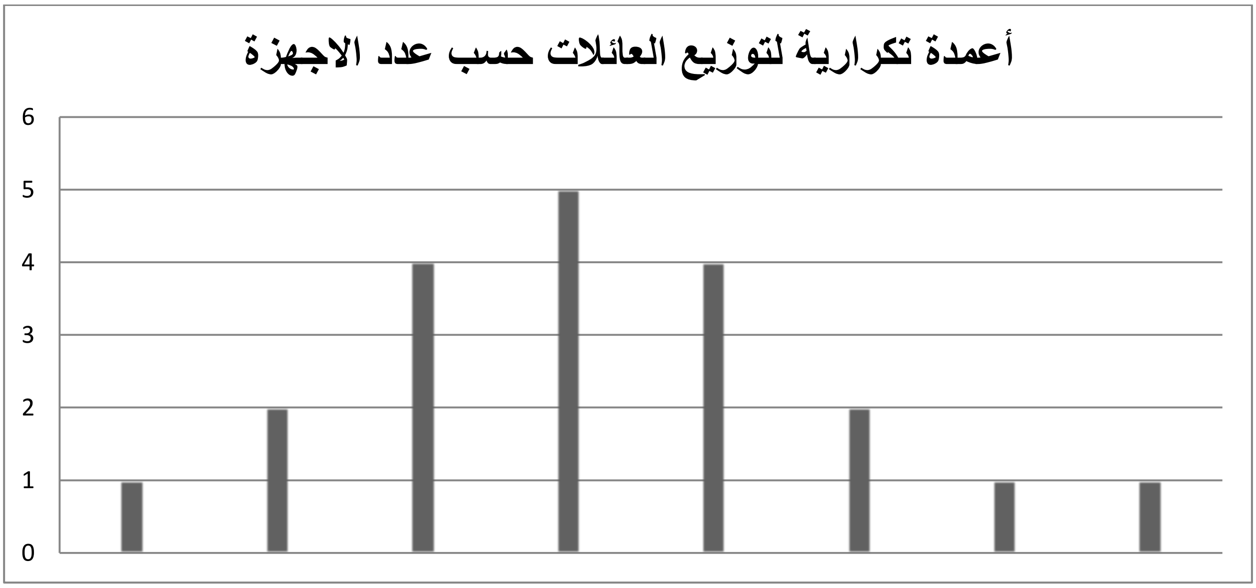
عدد الاجهزة	5	6	7	8	9	10	11	12	المجموع
التكرار	1	2	4	5	4	2	1	1	20

إن وضع البيانات بهذه الصورة أصبح أكثر وضوحا لمعرفة عدة معلومات كانت غير واضحة في الصورة الأولى. فمثلا من السهل الآن معرفة توزع عدد العائلات وكذا مستوى الرفاهية .

من جهة اخرى يمكن معرفة عدد او نسب العائلات التي عدد اجهزتها اقل او اكثر من x

x_i	n_i	$n_i \uparrow$	$n_i \downarrow$	f_i	$f_i \uparrow$	$f_i \downarrow$
5	1	1	20	0.05	0.05	1
6	2	3	19	0.1	0.15	0.95
7	4	7	17	0.2	0.35	0.85
8	5	12	13	0.25	0.6	0.65
9	4	16	8	0.2	0.8	0.4
10	2	18	4	0.1	0.9	0.2
11	1	19	2	0.05	0.95	0.1
12	1	20	1	0.05	1	0.05
$\sum$	20	/	/	1	/	/

التمثيلات البيانية



4.1- عرض بيانات المتغير الكمي المستمر في شكل جدول تكراري بسيط:

هناك طريقة أكثر اختصارا من السابقة يمكن بواسطتها وضع البيانات في جدول يبين ويوضح الخصائص العامة لهذه البيانات، يسمى هذا الجدول بجدول التوزيع التكراري ولتكوين مثل هذا الجدول نقوم بالتالي:

أولا: نحدد المجال (المدى) الذي تنتشر فيه البيانات، وهو الفرق بين أكبر قيمة للبيانات وأصغر قيمة لها، أي أن:

المدى = أكبر قيمة – أصغر قيمة.

$$E = x_{max} - x_{min}$$

ثانيا: نقسم المدى إلى فئات متساوية الطول بحيث يكون عددها مناسبا (ما بين 5 و 25 فئة) وهناك عدة طرق لحساب عدد الفئات نذكر منها:

1 – معادلة ستيرجسSturge التي تنص على أن عدد الفئات = 1 + 3.322 لغ عدد البيانات.

2 – معادلة يول yule التي تنص على عدد الفئات = $nc = 2.5 \times \sqrt[4]{n}$

حيث n عدد المفردات او حجم المجتمع او العينة

ثالثا: نحسب طول الفئة وهو يساوي المدى مقسوما على عدد الفئات

طول الفئة = $\frac{\text{المدى}}{\text{عدد الفئات}}$ $a = \frac{E}{nc}$

مما سبق نستخلص الطريقة المرنة في تحديد عدد الفئات وأطوالها والتي لا تعتمد على المعدلات الرياضية بل أن هذه الطريقة مرنة بطبيعتها وهي:

طول الفئة × عدد الفئات ≤ المدى

<u>مثال رقم 04 :</u>

البيانات التالية تمثل عدد ساعات العمل ل 50 عامل في مصنع خلال أسبوع :

المطلوب:

1) ما هو نوع المتغير؟ وطبيعته؟.

2) أي نوع من الفئات تستخدم في مثل هذه الحالة؟

3) تحديد عدد الفئات باستخدام معادلة Sturgers؟

4) تحديد عدد الفئات باستخدام معادلة Yule؟

5)انشاء جدول التوزيع التكراري مع حساب التكرار المتجمع الصاعد والنازل والتكرار النسبي؟

48	46	49	36	41	43	38	32	24	45
32	33	23	41	43	48	46	34	24	46
27	29	23	48	47	33	36	22	39	44
48	42	46	37	39	41	26	28	33	23
44	24	25	35	36	43	48	47	38	43

✓ ترتيب البيانات من الاصغر الى الأكبر

22	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	
1	3	3	1	1	1	1	1	0	
31	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	
0	2	3	1	1	3	1	2	2	
40	**41**	**42**	**43**	**44**	**45**	**46**	**47**	**48**	**49**
0	3	1	4	2	1	4	2	5	1

✓ حساب المدى

المدى = أكبر قيمة – أصغر قيمة.

$$E = x_{max} - x_{min} = 49 - 22 = 27$$

✓ حساب عدد الفئات او طول الفئة اذا كان عدد الفئات معلوم

حسب معادلة ستيرجسSturge

$$nc = 1 + 3.222 \log n = 1 + 3.322 \log 50 = 6.47 \ \cong 7 \ classes$$

حسب معادلة يول yule

$$nc = 2.5 \ \times \sqrt[4]{n} = 2.5 \ \times \sqrt[4]{50} = 6.64 \ \cong 7 \ classes$$

✓ نوع الفئات التي تستخدم في هذه الحالة هي الفئات المنتظمة (فئات متاوية الطول) والغرض منها هو تجميع القيم المتقاربة في مجموعات بحيث لا يكون عدد الفئات صغيرا فتضيع معالم التوزيع وتفقد كسيرا من التفاصيل واسضا لا يكون عدد الفئات كبيرا فتضيع فائدة التجميع للفئات.

✓ حساب مراكز الفئات

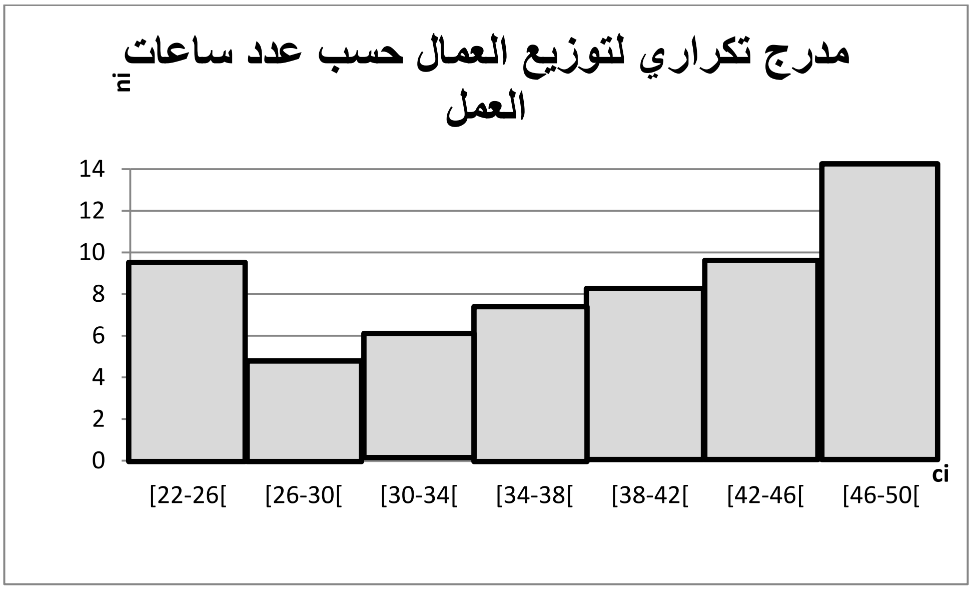
مركز الفئة الاولى هو :

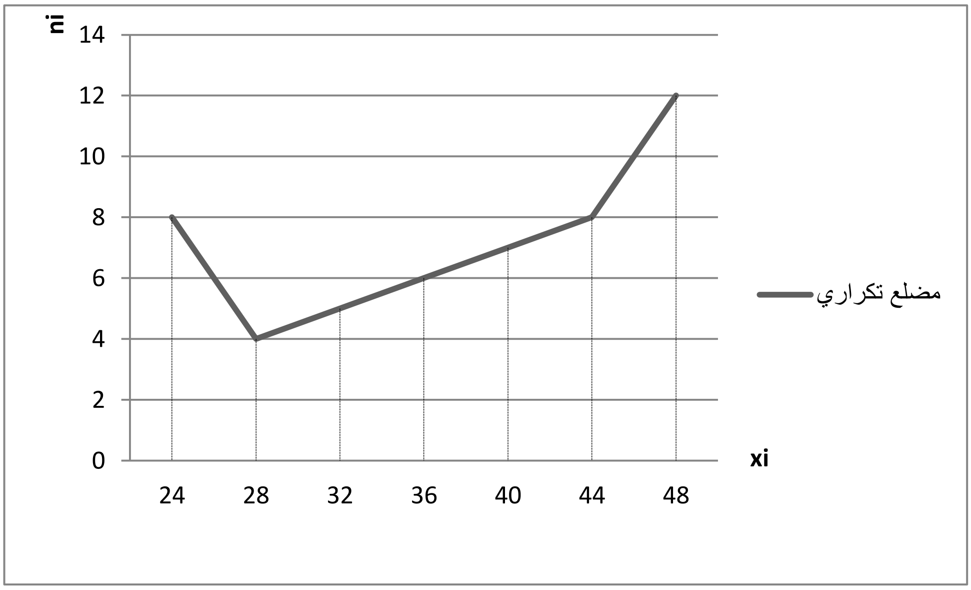
$$x_i = \frac{l_1 + l_2}{2} = \frac{22 + 26}{2} = 24$$

✓ جدول التوزيع التكراري :

c_i	x_i	n_i	$n_i \uparrow$	$n_i \downarrow$	f_i	$f_i \uparrow$	$f_i \downarrow$
[22-26[	24	8	8	50	0.16	0.16	1
[26-30[	28	4	12	42	0.08	0.24	0.84
[30-34[	32	5	17	38	0.1	0.34	0.76
[34-38[	36	6	23	33	0.12	0.46	0.66
[38-42[	40	7	30	27	0.14	0.6	0.54
[42-46[	44	8	38	20	0.16	0.76	0.4
[46-50[	48	12	50	12	0.24	1	0.24
$\sum$	/	50	/	/	1	/	/

✓ المدرج التكراري :





هو عبارة عن مستطيلات متلاصقة قاعدة كل منها هي طول الفئة وارتفاعها هو تكرار تلك الفئة

ملاحظة : المدرج التكراري صحيح اذا كانت مجموع مساحة المستطيلات مقسومة على طول الفئة مساوية لمجموع التكرارات .

ولرسم المدرج التكراري اذا كانت الفئات غير متساوية الطول يجب تعديل التكرارات بقسمة التكرارات على طول الفئة

التكرار المعدل $n'_i = \frac{n_i}{a_i}$

المضلع التكراري هو عبارة عن خط منكسر يربط مراكز الفئات ويجب ان تكون المساحة الواقعة تحت المضلع التكراري مقسومة على طول الفئة مساوية لمجموع التكرارات

ولرسم المضلع التكراري اذا كانت الفئات غير متساوية الطول يجب تعديل التكرارات بقسمة التكرارات على طول الفئة.

3-2 الرسومات البيانية

تعتبر الرسومات البيانية أكثر الطرق الإحصائية استخداماً في وصف وتلخيص البيانات، وذلك لتميزها بالبساطة والسهولة والوضوح .

هنالك أنواع عديدة من الأشكال والرسوما ت يمكن استخدامها

لوصف البيانات لكل منها ميزاته وعيوبه من حيث درجة البساطة والوضوح ومن حيث إظهار معالم البيانات . وتبعاً لذلك يختلف استخدامها حسب نوع البيانات وأحجامها، عدد المتغيرات والغرض من وصف البيانات .

من أشهر انواع الاشكال والرسومات البيانية مرتبة حسي البساطة هي الاشكال المصورة ،الدوائر المجزأة ،الاعمدة الاشرطة وتشمل الاشرطة البسيطة والمتلاصقة (المدرجات) والاشرطة المركبة واخيرا المضلعات والمنحنيات .

- تعطي الاشكال المصورة انطباها بصريا جيدا عن مجموعة البيانات المبحوثة ولكن يعيبها عدم الدقة واخفاء التفاصيل .

مثال على ذلك :

- تستخدم الدوائر المجزأة عندما يكون الهدف مقارنة الاجزاء المختلفة بالنسبة للمجموع الكلي وعدد الأجزاء المقارنة قليل نسبيا (عدد حالات المتغير المدروس)

- تستخدم الأعمدة والاشرطة عندما تكون أجزاء الظاهرة المقارنة كثيرة العدد نسبيا وعندما نرغب في توضيح قيم الأجزاء المقارنة.

تستخدم الخطوط البيانية عندما يكون عدد المفردات كبير نسبيا او عندما يكون الغرض توضيح العلاقة بين المتغيرات لفترات زمنية متعاقبة كما في حالة السلاسل الزمنية.

الفصل الثالث : مقاييس النزعة المركزية

مقدمة :

ان الجداول التكرارية و الرسومات البيانية في تحليل ودراسة الظواهر لتحديد الخصائص والاتجاهات والعلاقات ، يعتمد على دقة التمثيل البياني نفسه ، وعليه فإنه من الأفضل اللجوء إلى طرق القياس الكمي، حيث يستخدم الباحث الطريقة الرياضية في القياس.

فالهدف الأساسي من استخدام مقاييس النزعة المركزية هو تلخيص البيانات في محاولة أخرى لوصفها عن طريق التعرف على مركزها وتموضعها ومن خلال هذين المؤشرات يتمكن الباحث من فهم أبعاد الظاهرة قيد الدراسة.

ومن أهم مقاييس النزعة المركزية التي سنتعرض إليها بالدراسة الوسط الحسابي والوسيط والمنوال ، كما سنتعرض بالدراسة لحساب كل منهم من البيانات (الغير مبوبة) ومن البيانات المبوبة

1- انواع البيانات الكمية :

بيانات غير مبوبة وغير مرجحة بيانات غير مبوبة مرجحة
بيانات مبوبة مرجحة

مبوبة = تأخذ شكل فئات غير مبوبة = بدون فئات مرجحة= بها تكرارات
غير مرجحة= لا يوجد تكرارات

ومن خلال الامثلة سنوضع انواع البيانات جيدا.

2- الوسيط la médiane:

هو أحد مقاييس النزعة المركزية الذي يأخذ بعين الاعتبار رتبة القيم و يعرف الوسيط على انه القيمة التي تقسم البيانات إلى جزئيين متساويين أي هناك 50% من القيم أقل من قيمة الوسيط و هناك 50% من القيم اكبر من قيمة الوسيط.

يعد الوسيط احد مقاييس النزعة المركزية المهمة في التطبيقات الإحصائية . ويعرف بأنه تلك القيمة التي يقع ترتيبها وسط المجموعة عند ترتيب هذه القيم ترتيبا تصاعديا او تنازليا أي تقسم المجموعة إلى قسمين متساويين بحيث تتساوى عدد الحدود التي اصغر من الوسيط مع عددا لحدود الأكبر منه.
ويرمز له بالرمز m_e يشترط في حساب الوسيط ان تكون البيانات مرتبة .(او تخضع الى ترتيب)

1.2- الوسيط في حالة بيانات نوعية اسمية:

لا يمكن حساب الوسيط في هذه الحالة نظرا لتعذر ترتيب البيانات وبالتالي لا يمكن التعبير باقل وأكثر.

2.3- الوسيط في حالة بيانات نوعية ترتيبية :

يمكن حساب الوسيط في هذه الحالة نظرا لتمكننا من ترتيب البيانات وبالتالي يمكن التعبير باقل وأكثر.

نأخذ معطيات المثال رقم 02 :

x_i	n_i	$n_i \uparrow$	$n_i \downarrow$	f_i	$f_i \uparrow$	$f_i \downarrow$
مقبول	10	10	30	0.33	0.33	1
قريب من الجيد	8	10+8=18	30-10=20	0.27	0.6	0.67
جيد	6	18+6=**24**	20-8=**12**	0.2	**0.8**	**0.4**
جيد جدا	4	24+4=28	12-6=6	0.13	0.93	0.2
ممتاز	2	28+2=30	6-4=2	0.07	1	0.07
$\sum$	30	/	/	1	/	/

✓ إيجاد رتبة الوسيط :

يمكن إيجاد رتبة الوسيط باستعمال عدة طرق فباستعمال $n_i \uparrow$ أو $n_i \downarrow$ نبحث عن الرتبة $Rg = \frac{\sum n_i}{2}$

$$Rg = \frac{\sum n_i}{2} = \frac{30}{2} = 15$$

نبحث في قيم $n_i \uparrow$ عن اقرب قيمة التي هي أكبر أو تساوي الرتبة Rg

فنجد أن 18 هي اقرب قيمة ل 15 .

نبحث في قيم $n_i \downarrow$ عن اقرب قيمة التي هي أكبر أو تساوي الرتبة Rg

فنجد أن20 هي اقرب قيمة ل 15 .

فباستعمال $f_i \uparrow$ أو $f_i \downarrow$ نبحث عن الرتبة $Rg = 0.5$

نبحث في قيم $f_i \uparrow$ عن اقرب قيمة التي هي أكبر أو تساوي الرتبة Rg

فنجد أن 0.6 هي اقرب قيمة ل0.5 .

نبحث في قيم $f_i \downarrow$ عن اقرب قيمة التي هي أكبر أو تساوي الرتبة Rg

فنجد أن0.67 هي اقرب قيمة ل 0.5 .

✓ إيجاد قيمة الوسيط :

الوسيط هو القيمة التي توافق الرتبة فنقرأ في العمود x_i عن الملاحظة التي قيمتها 18 حسب $n_i \uparrow$ أو قيمتها 20 حسب $n_i \downarrow$ أو 0.6 حسب $f_i \uparrow$ أو قيمتها 0.66 حسب $f_i \downarrow$

$$m_e = \text{قريب من الجيد}$$

15 طالب تحصلوا على تقدير يقل او يساوي عن قريب من الجيد و15 طالب تحصلوا على تقدير يزيد او يساوي عن قريب من الجيد. او بصيغة اخرى بما أن 15 هي 50 %من عدد الطلبة.

50% من الطلبة تحصلوا على تقدير يقل او يساوي عن قريب من الجيد و50% من الطلبة تحصلوا على تقدير يزيد او يساوي عن قريب من الجيد.

3.3- الوسيط في حالة بيانات كمية غير مرجحة وغير مبوبة :

احسب الوسيط للبيانات التالية:

المثال الأول : 1، 1، 2، 3، 3، 4، 4 ، 4، 6، 7،7، 9

يجب أولا ترتيب البيانات وبما ان عدد البيانات زوجي (12) ومنه

عندما يكون عدد القيم عدد زوجي يتعذر علينا إيجاد قيمة تتوسط السلسلة وبالتالي يكون الوسيط متوسط قيمتين بين القيمة التي رتبتها $\frac{n+1}{2}$ والقيمة التي رتبتها $\frac{n}{2}+1$

$$m_e = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2} = \frac{4+4}{2} = 4$$

المثال الثاني:0، 1، 1، 2.5، 3، 5، 5.5، 6، 7.5

عدد البيانات فردي (9) ومنه

عندما يكون عدد القيم عدد فردي يسهل علينا إيجاد قيمة تتوسط السلسلة

$$m_e = x_{\frac{n+1}{2}} = x_5 = 3$$

4.3- الوسيط في حالة بيانات كمية مرجحة وغير مبوبة :

نأخذ معطيات المثال رقم 03 :

x_i	n_i	$n_i \uparrow$	$n_i \downarrow$	f_i	$f_i \uparrow$	$f_i \downarrow$
5	1	1	20	0.05	0.05	1
6	2	3	19	0.1	0.15	0.95
7	4	7	17	0.2	0.35	0.85
8	5	12	13	0.25	0.6	0.65
9	4	16	8	0.2	0.8	0.4
10	2	18	4	0.1	0.9	0.2
11	1	19	2	0.05	0.95	0.1
12	1	20	1	0.05	1	0.05
$\sum$	20	/	/	1	/	/

✓ إيجاد رتبة الوسيط :

يمكن إيجاد رتبة الوسيط باستعمال عدة طرق فباستعمال $n_i \uparrow$ أو $n_i \downarrow$ نبحث عن الرتبة $Rg = \frac{\sum n_i}{2}$

$$Rg = \frac{\sum n_i}{2} = \frac{20}{2} = 10$$

نبحث في قيم $n_i \uparrow$ عن اقرب قيمة التي هي أكبر أو تساوي الرتبة Rg

فنجد أن 12 هي اقرب قيمة ل 10 .

نبحث في قيم $n_i \downarrow$ عن اقرب قيمة التي هي أكبر أو تساوي الرتبة Rg

فنجد أن 13 هي اقرب قيمة ل 10 .

فباستعمال $f_i \uparrow$ أو $f_i \downarrow$ نبحث عن الرتبة $Rg = 0.5$

نبحث في قيم $f_i \uparrow$ عن اقرب قيمة التي هي أكبر أو تساوي الرتبة Rg

فنجد أن 0.6 هي اقرب قيمة لـ0.5 .

نبحث في قيم $f_i \downarrow$ عن اقرب قيمة التي هي أكبر أو تساوي الرتبة Rg

فنجد أن 0.65 أي اقرب قيمة ل 0.5 .

✓ إيجاد قيمة الوسيط :

الوسيط هو القيمة التي توافق الرتبة فنقرأ في العمود x_i عن الملاحظة التي قيمتها 12 حسب $n_i \uparrow$ أو قيمتها 13 حسب $n_i \downarrow$ أو 0.6 حسب $f_i \uparrow$ أو قيمتها 0.65 حسب $f_i \downarrow$

$$m_e = 8$$

10 عائلات عدد الأجهزة الكهرومنزلية لديهم يقل او يساوي 8 أجهزة 10 عائلات عدد الأجهزة الكهرومنزلية لديهم يزيد او يساوي 8 أجهزة 10. او بصيغة اخرى بما أن 10 هي 50 %من عدد المنازل.

50% من العائلات عدد الأجهزة الكهرو منزلية لديهم يقل او يساوي 8 أجهزة 10 و 50% من العائلات عدد الأجهزة الكهرو منزلية لديهم يزيد او يساوي 8 أجهزة 10.

5.3- الوسيط في حالة بيانات كمية مبوبة :

نأخذ معطيات المثال رقم 04 :

c_i	n_i	$n_i \uparrow$	$n_i \downarrow$	f_i	$f_i \uparrow$	$f_i \downarrow$
[22-26[	8	8	50	0.16	0.16	1
[26-30[	4	12	42	0.08	0.24	0.84
[30-34[	5	17	38	0.1	0.34	0.76

[34-38[	6	23	33	0.12	0.46	0.66
[38-42[	7	30	27	0.14	0.6	0.54
[42-46[	8	38	20	0.16	0.76	0.4
[46-50[	12	50	12	0.24	1	0.24
$\sum$	50	/	/	1	/	/

✓ إيجاد رتبة الوسيط :

في هده الحالة لا يمكننا إيجاد قيمة توافق رتبة الوسيط ولكن سنجد فئة الوسيط ينتمي الى الفئة ذات الرتبة $Rg = \frac{\sum n_i}{2}$ أو $Rg = 0.5$

فباستعمال $n_i \uparrow$ أو $n_i \downarrow$

$$Rg = \frac{\sum n_i}{2} = \frac{50}{2} = 25$$

نبحث في قيم $n_i \uparrow$ عن اقرب قيمة التي هي أكبر أو تساوي الرتبة Rg

فنجد أن 30 هي اقرب قيمة ل 25 .

نبحث في قيم $n_i \downarrow$ عن اقرب قيمة التي هي أكبر أو تساوي الرتبة Rg

فنجد أن 27 هي اقرب قيمة ل 25 .

فباستعمال $f_i \uparrow$ أو $f_i \downarrow$ نبحث عن الرتبة $Rg = 0.5$

نبحث في قيم $f_i \uparrow$ عن اقرب قيمة التي هي أكبر أو تساوي الرتبة Rg

فنجد أن 0.6 هي اقرب قيمة ل0.5 .

نبحث في قيم $f_i \downarrow$ عن اقرب قيمة التي هي أكبر أو تساوي الرتبة Rg

فنجد أن0.54 هي اقرب قيمة ل 0.5 .

نكتفي بطريقة واحدة حسب المعطيات .

- ✓ إيجاد فئة الوسيط :

نجد ان الوسيط ينتمي الى الفئة]42-38]

$$m_e \in [42 - 38[$$

- ✓ إيجاد قيمة الوسيط :

باستعمال القانون التالي يمكننا إيجاد قيمة الوسيط :

$$m_e = l + \frac{Rg - n_{i-1} \uparrow}{n_i} \times a_i$$

حيث :

L : الحد الأدنى للفئة الوسيطية

Rg : رتبة الوسيط

$n_{i-1} \uparrow$ هو التكرار المتجمع الصاعد الذي يسبق التكرار المتجمع الصاعد للفئة الوسيطية

n_i هو تكرار المطلق للفئة الوسيطية

a_i هو طول الفئة الوسيطية.

$$m_e = l + \frac{Rg - n_{i-1} \uparrow}{n_i} \times a_i = 38 + \frac{25 - 23}{7} \times 4$$
$$= 39.14$$

25 عامل يتقاضى أقل من391.4 دج/سا (الأجر مضروب في 10دج/سا) و 25 عامل يتقاضى أكثر من391.4 دج/سا

فباستعمال f_i

$$m_e = l + \frac{Rg - f_{i-1}\uparrow}{f_i} \times a_i = 38 + \frac{0.5 - 0.46}{0.14} \times 4 = 39.14$$

50% من العمال يتقاضون أقل من391.4 دج/سا (الأجر مضروب في 10دج/سا) و 50% من العمال يتقاضون أكثر من391.4 دج/سا.

6.3- أشباه الوسيط :

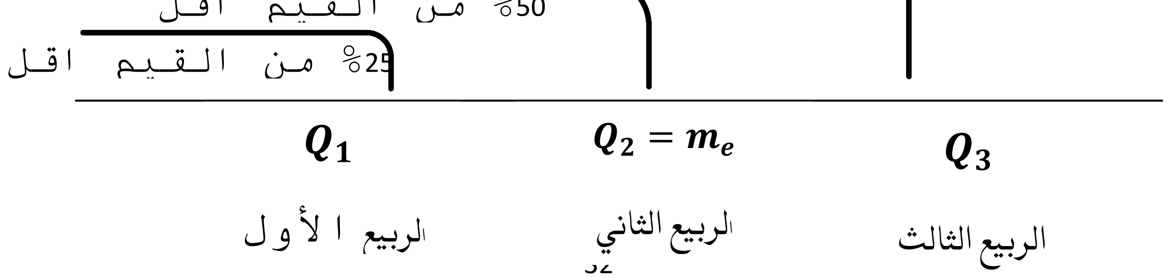
هناك مقاييس أخرى تشبه الوسيط كالربيعات و العشيرات و المؤينات وتخضع في عملية الحساب الى نفس طريقة حساب الوسيط.

✓ الربيعات :

عند تقسيم القيم إلى أربع أجزاء متساوية، يوجد ثلاث إحصاءات ترتيبي تسمى بالرباعيات، وهي:

- الربيع الأول: وهو القيمة التي يقل عنها ربع عدد القيم، أي يقل عنها 25% من القيم، ويرمز له بالرمز Q_1.
- الربيع الثاني: وهو القيمة التي يقل عنها نصف عدد القيم، أي يقل عنها 50% من القيم، ويرمز له بالرمز Q_2، ومن ثم يعبر هذا الربيع عن الوسيط.
- الربيع الثالث: وهو القيمة التي يقل عنها ثلاث أرباع عدد القيم، أي يقل عنها 75% من القيم، ويرمز له بالرمز Q_3.

والشكل (3-3) يبين أماكن الرباعيات الثلاث



ولحساب أي من الرباعيات الثلاث، يتم إتباع الآتي:

الربيع الأول رتيته $Rg = \frac{1\sum n_i}{4}$ أو $Rg = 0.25$

الربيع الثاني رتيته $Rg = \frac{2\sum n_i}{4}$ أو $Rg = 0.5$

الربيع الأول رتيته $Rg = \frac{3\sum n_i}{4}$ أو $Rg = 0.75$

نأخذ معطيات المثال السابق:

حساب الربيع الأول :

$$Rg = \frac{\sum n_i}{4} = \frac{50}{4} = 12.5$$

$$Q_1 \in [30 - 34[$$

$$Q_1 = l + \frac{Rg - n_{i-1} \uparrow}{n_i} \times a_i = 30 + \frac{12.5 - 12}{5} \times 4 = 30.4$$

13عامل يتقاضى أقل من304 دج/سا (الأجر مضروب في 10دج/سا) و 37 عامل يتقاضى أكثر من304 دج/سا

فباستعمالf_i

$$Q_1 = l + \frac{Rg - f_{i-1} \uparrow}{f_i} \times a_i = 38 + \frac{0.25 - 0.24}{0.1} \times 4$$
$$= 30.4$$

25% من العمال يتقاضون أقل من304 دج/سا (الأجر مضروب في 10دج/سا) و 75% من العمال يتقاضون أكثر من304 دج/سا.

حساب الربيع الثالث :

$$Rg = \frac{3\sum n_i}{4} = \frac{3 \times 50}{4} = 37.5$$

$$Q_3 \in [42 - 46[$$

$$Q_3 = l + \frac{Rg - n_{i-1} \uparrow}{n_i} \times a_i = 42 + \frac{37.5 - 30}{8} \times 4 = 45.75$$

37عامل يتقاضى أقل من457.5 دج/سا (الأجر مضروب في 10دج/سا) و 13 عامل يتقاضى أكثر من457.5 دج/سا

فباستعمال f_i

$$Q_3 = l + \frac{Rg - f_{i-1} \uparrow}{f_i} \times a_i = 42 + \frac{0.75 - 0.6}{0.16} \times 4$$
$$= 45.75$$

75% من العمال يتقاضون أقل من457.5 دج/سا (الأجر مضروب في 10دج/سا) و 25% من العمال يتقاضون أكثر من457.5 دج/سا.

ففي حالة بيانات كمية مبوبة وغير مرجحة فنكتفي بحساب الرتبة لكل ربيع ونرى ما يوافقها كقيمة في عمود x_i

✓ العشيرات:

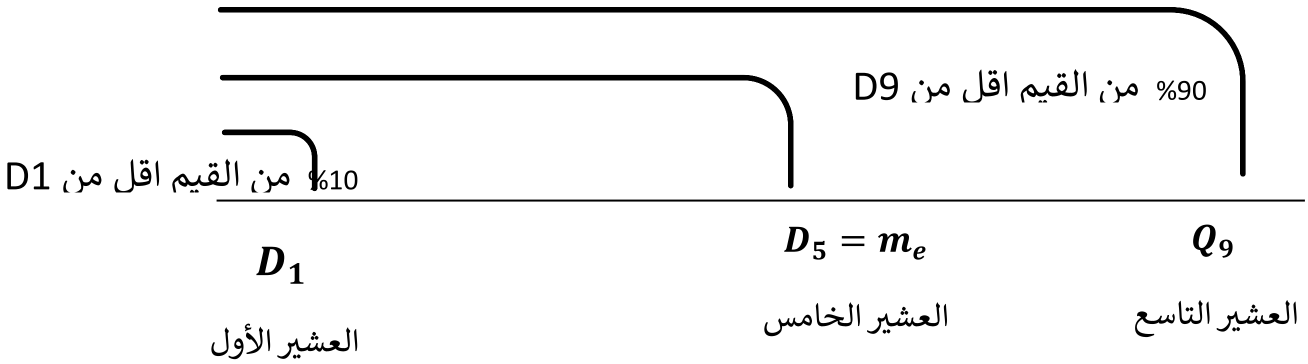
عند تقسيم القيم إلى عشرة أجزاء متساوية، يوجد تسعة إحصاءات ترتيبية تسمى بالعشيرات، وهي:

- العشير الأول: وهو القيمة التي يقل عنها عشر عدد القيم، أي يقل عنها 10% من القيم، ويرمز له بالرمز D_1.
- العشير الثاني: وهو القيمة التي يقل عنها عشري عدد القيم، أي يقل عنها 20%

من القيم، ويرمز له بالرمز D_2

- ....
- العشير الخامس: وهو القيمة التي يقل عنها خمسة أعشار عدد القيم، أي يقل عنها 50% من القيم، ويرمز له بالرمز D_5 وهو مساوي للوسيط
- ....
- العشير التاسع: وهو القيمة التي يقل عنها تسعة اعشار عدد القيم، أي يقل عنها 90% من القيم، ويرمز له بالرمز D_9

والشكل (3-3) يبين أماكن العشيرات الثلاث



ولحساب أي من العشيرات التسعة، يتم إتباع الآتي:

العشير الأول رتيته $Rg = \frac{1\sum n_i}{10}$ أو $Rg = 0.10$

العشير الثاني رتيته $Rg = \frac{2\sum n_i}{10}$ أو $Rg = 0.20$

الى غاية....................................

العشير التاسع رتيته $Rg = \frac{9\sum n_i}{10}$ أو $Rg = 0.90$

نأخذ معطيات المثال السابق:

حساب العشير الأول :

$$Rg = \frac{\sum n_i}{10} = \frac{50}{10} = 5$$

$$Q_1 \in [22 - 26[$$

$$D_1 = l + \frac{Rg - n_{i-1} \uparrow}{n_i} \times a_i = 22 + \frac{5 - 0}{8} \times 4 = 24.5$$

5عمال يتقاضون أجر أقل من345 دج/سا (الأجر مضروب في 10دج/سا) و 45 عامل يتقاضون اجر أكثر من245 دج/سا

فباستعمالf_i

$$D_1 = l + \frac{Rg - f_{i-1} \uparrow}{f_i} \times a_i = 22 + \frac{0.1 - 0}{0.16} \times 4 = 34.5$$

10% من العمال يتقاضون أقل من345 دج/سا (الأجر مضروب في 10دج/سا) و 90% من العمال يتقاضون أكثر من345 دج/سا.

حساب العشير التاسع :

$$Rg = \frac{9 \sum n_i}{10} = \frac{9 \times 50}{10} = 45$$

$$D_9 \in [46 - 50[$$

$$D_9 = l + \frac{Rg - n_{i-1} \uparrow}{n_i} \times a_i = 46 + \frac{45 - 38}{12} \times 4 = 48.33$$

45 عامل يتقاضى أقل من483.33 دج/سا (الأجر مضروب في 10دج/سا) و 5 عمال يتقاضون أكثر من 483.33 دج/سا

فباستعمالf_i

$$Q_3 = l + \frac{Rg - f_{i-1} \uparrow}{f_i} \times a_i = 46 + \frac{0.9 - 0.76}{0.24} \times 4$$
$$= 48.33$$

90% من العمال يتقاضون أقل من483.33 دج/سا (الأجر مضروب في 10دج/سا)
و 10% من العمال يتقاضون أكثر من483.33 دج/سا.

ففي حالة بيانات كمية مبوبة وغير مرجحة فنكتفي بحساب الرتبة لكل عشير ونرى ما يوافقها كقيمة في عمود x_i

هناك مقاييس أخرى تسمى بالمؤينات لها نفس المفهوم وتتبع نفس الطريقة كالربيعات والعشيرات.

7.3- مزايا وعيوب الوسيط

من مزايا الوسيط

1- لا يتأثر بالقيم الشاذة أو المتطرفة .
2- كما أنه سهل في الحساب .
3- يمكن حسابه بيانيا

ومن عيوب الوسيط

1- أنه لا يأخذ عند حسابه كل القيم في الاعتبار، فهو يعتمد على قيمة أو قيمتين فقط .
2- يصعب حسابه في حالة البيانات الوصفية المقاسة بمعيار اسمي .

3- المنوال le mode :

هي القيمة التي تتكرر اكثر من غيرها من بين مجموعة من القيم او هي القيمة الشائعة من بين مجموعة من القيم وقد يكون للمجموعة اكثر من منوال حيث تتساوى تكرارات القيم اكثر من مرة.

يعرف المنوال بأنه القيمة الأكثر شيوعا أو تكرارا ، ويكثر استخدامه في حالة البيانات الوصفية ، لمعرفة النمط (المستوى) الشائع ويرمز له m_o

يشترط في حساب المنوال وجود تكرارات لتمييز القيمة ذات اكبر تكرار.

في حالة بيانات نوعية اسمية او ترتيبية او كمية غير مبوبة بشرط ان تكون مرجحة عن الملاحظة التي بها اكبر تكرار وتكون هي المنوال

المثال رقم 1 : المنوال هو :oppo نوع الهاتف الاكثر امتلاكا عند الطلبة هو oppo

المثال رقم 2 : المنوال هو : مقبول تقدير الباكالوريا الأكثر تحصيلا عند الطلبة هو مقبول

المثال رقم3 : المنوال هو : 8 أجهزة اكبر جل العائلات تمتلك 8 أجهزة كهرومنزلية

اما في حالة بيانات مبوبة فيختلف حساب المنوال .

1.3- حالة بيانات مبوبة وفئات متساوية الطول :

نأخذ معطيات المثال رقم 06 :

ساعات العمل	[22-26[	[26-30[	[30-34[	[34-38[	[38-42[	[42-46[	[46-50[
عدد العمال	24	28	32	36	40	44	48

المنوال هو القيمة ذات اكبر تكرار وبما أن لدينا فئات فان المنوال ينتمي الى الفئة ذات اكبر تكرار (بما ان الفئات متساوية الطول فيسهل تمييز الفئة)

$$m_o \in [46-50[$$

$$m_o = l + \frac{d_1}{d_1+d_2} \times a_i$$

l : الحد الأدنى لفئة المنوال (الفئة المناظرة لأكبر تكرار) .

d_1 : الفرق الأول = (تكرار فئة المنوال – تكرار سابق)

d_2 : الفرق الثاني = (تكرار فئة المنوال – تكرار لاحق)

a_i: طول فئة المنوال .

فئـــة المنوال = الفئة المناظرة لأكبر تكرار

$$m_o = 46 + \frac{(48-44)}{(48-44)+(48-0)} \times 4 = 46.30$$

أكبر عدد من العمال يعملون 46.30 ساعة

1.4- حالة بيانات مبوبة وفئات غير متساوية الطول :

الجدول التالي يبين توزيع العمال حسب السن

ساعات العمل	[20-24[	[24-36[	[26-30[	[30-38[	[38-42[	[42-50[	[50-60[
عدد العمال	22	18	26	20	10	8	6

في هذ الحالة يصعب إيجاد الفئة ذات اكبر تكرار وعل هذا الأساس يتم تقسيم التكرارات على طول الفئة لمعرفة لي الفئات تحصلت على اكسر حصة من التكرارات

التكرار المعدل $n'_i = \frac{n_i}{a_i}$

c_i	n_i	a_i	n'_i
[20-24[	22	4	5,5
[24-26[	18	2	9
[26-30[	26	4	6,5
[30-38[	20	8	2,5

[38-42[	10	4	2,5
[42-50[	8	8	1
[50-60[	6	10	0,6

يحسب d_1 و d_2 بالتكرارات المعدلة .

$$m_o \in [24-26[$$

$$m_o = l + \frac{d_1}{d_1+d_2} \times a_i = 24 + \frac{(9-5.5)}{(9-5.5)+(9-6.5)} \times 2 = 25.16$$

أكبر عدد من العمال سنهم هو 25.16 سنة.

2.5- مزايا وعيوب المنوال :

- لا يأخذ في عين الاعتبار جميع البيانات أي لا يتاثر بالقيم المتطرفة.

- يمكن حسابه بيانيا

-يمكن حسابه من الجداول الإحصائية المفتوحة

يعتبر افضل المقاييس لوصف الظواهر النوعية

ام عيوبه

- يفقد أهميته عندما يكون اكثر من منوال في السلسلة الإحصائية

4- المتوسط الحسابي: la moyenne arithmétiques

يعتبر المتوسط الحسابي من ابسط المقاييس المستخدمة في الاحصاء، حيث يقدم ملخصا للبيانات بتقديم قيمة افتراضية يتم حسابها لأي مجموعة من البيانات وليس من الضروري ان تكون موجودة في مجموعة من البيانات المشاهدة .[7]

[7] الاحصاء باستخدام spss ، لجنة الاعداد والترجمة ،شعاع للنشر والعلوم ، 2007 ص 10

هو من اهم مقاييس النزعة المركزية وأكثر استخداما في النواحي التطبيقية ويعرف عموما على أنه مجموع القيم مقسوما على عددها ويرمز له بالرمز $\bar{x}$

1.4- المتوسط الحسابي في حالة بيانات غير مبوبة وغير مرجحة:

لدينا نقاط مجموعة من الطلبة: 12، 5،8،13، 6،9،10

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

$$\bar{x} = \frac{6+9+10+13+8+5+12}{7} = \frac{63}{7} = 9$$

2.4- المتوسط الحسابي في حالة بيانات غير مبوبة و مرجحة:

مثال رقم 05: الجدول التالي يبين توزيع عدد العائلات حسب عدد افراد الاسرة

x_i	n_i	$x_i n_i$	f_i	$x_i f_i$
1	22	22	0.11	0.11
2	28	56	0.14	0.28
3	32	96	0.16	0.48
4	48	192	0.24	0.96
5	40	200	0.2	1
6	20	120	0.1	0.6
7	10	70	0.05	0.35
$\sum$	200	756	1	3.78

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i n_i}{\sum_{i=1}^{n} n_i}$$

$$\bar{x} = \frac{756}{200} = 3.78 \cong 4$$

اما القانون الثاني باستخدام النسب:

$$\bar{x} = \sum_{i=1}^{n} x_i f_i$$

متوسط افراد الاسرة هو 4 لهذه العائلات

3.4- المتوسط الحسابي في حالة بيانات مبوبة و مرجحة :

مثال رقم 06 : الجدول التالي يبين توزيع العمال حسب عدد ساعات العمل خلال اسبوع

ساعات العمل	[22-26[	[26-30[	[30-34[	[34-38[	[38-42[	[42-46[	[46-50[
عدد العمال	24	28	32	36	40	44	48

احسب المتوسط الحسابي بكلا الطريقتين

نلا حظ في هذه الحالة وجود فئات فلا بد من حساب مراكز الفئات للتمكن من تطبيق العددي للقانون المتوسط الحسابي

c_i	x_i	n_i	$x_i n_i$	f_i	$x_i f_i$
[22-26[	24	8	192	0.16	3.84
[26-30[	28	4	112	0.08	2.24
[30-34[	32	5	160	0.1	3.2
[34-38[	36	6	216	0.12	4.32
[38-42[	40	7	280	0.14	5.6
[42-46[	44	8	352	0.16	7.04
[46-50[	48	12	576	0.24	11.52
$\sum$	/	50	1888	1	37.76

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i n_i}{\sum_{i=1}^{n} n_i}$$

$$\bar{x} = \frac{1880}{50} = 37.76$$

اما القانون الثاني باستخدام النسب:

$$\bar{x} = \sum_{i=1}^{n} x_i f_i = 37.76$$

متوسط عدد ساعات العمل لهؤلاء العمال هو 37.76 ساعة

4.4 – خواص المتوسط الحسابي :

- خاصية الجمع والطرح :

تطبق هذه الخاصية عندما بتم إضافة او اقتطاع قيمة ثابتة من قيم السلسلة الإحصائية .

فإذا كانت القيم هي : $x_1, x_2, \ldots, x_n$ ، وتم إضافة مقدار ثابت (a) إلى كل قيمة من القيم ،

$$\overline{x+a} = \frac{\sum_{i=1}^{n}(x_i + a)n_i}{\sum_{i=1}^{n} n_i}$$

$$= \frac{\sum_{i=1}^{n}(x_i n_i + a n_i)}{\sum_{i=1}^{n} n_i} = \frac{\sum_{i=1}^{n} x_i n_i + a\sum_{i=1}^{n} n_i}{\sum_{i=1}^{n} n_i}$$

$$= \frac{\sum_{i=1}^{n} x_i n_i}{\sum_{i=1}^{n} n_i} + a\frac{\sum_{i=1}^{n} n_i}{\sum_{i=1}^{n} n_i}$$

$$\overline{x+a} = \bar{x} + a$$

$$\overline{x-a} = \bar{x} - a$$

إذا أضيف مقدار ثابت إلى كل قيمة من القيم ، فإن الوسط الحسابي للقيم المعدلة (بعد الإضافة) يساوى الوسط الحسابي للقيم الأصلية (قبل الإضافة) مضافا إليها هذا المقدار الثابت .

✓ خاصية الضرب والقسمة :

تطبق هذه الخاصية عندما بتم إضافة او اقتطاع قيمة غير ثابتة من قيم السلسلة الإحصائية.

فإذا كانت القيم هي : $x_1, x_2, \ldots, x_n$ ، وتم إضافة مقدار غير ثابت (a) إلى كل قيمة من القيم.

$$\overline{ax} = \frac{\sum_{i=1}^{n}(ax_i n_i)}{\sum_{i=1}^{n} n_i} = a\frac{\sum_{i=1}^{n} x_i n_i}{\sum_{i=1}^{n} n_i}$$

$$\overline{ax} = a\bar{x}$$

$$\overline{\left(\frac{x}{a}\right)} = \frac{\bar{x}}{a}$$

5.4- مزايا وعيوب الوسط الحسابي

يتميز الوسط الحسابي بالمزايا التالية :

- أنه سهل الحساب .
- يأخذ في الاعتبار كل القيم .
- أنه أكثر المقاييس استخداما وفهما .

ومن عيوبه .

- أنه يتأثر بالقيم الشاذة والمتطرفة .
- يصعب حسابه في حالة البيانات الوصفية . يصعب حسابه في حالة الجداول التكرارية المفتوحة

6.4- المتوسط الهندسي :

إنّ المتوسط الهندسي معرّف فقط لمجموعة أعداد فيها كل الحدود موجبة. كما ويستخدم غالبًا في الحالات التي تكون فيها المعطيات هي قيمًا من المفروض أن تضرب بعضها ببعض، أو تلك المعطيات ذات الطابع الأسي، كالنمو الأسي لمجموعات سكانية، أو لحساب نسبة الفائدة المعدلة على مر عدة سنين.

في حالة بيانات غير مبوبة وغير مرجحة

لتكن القيم التالية : $\mathrm{x}_1 ; \mathrm{x}_2 ; \dots \dots \dots ; \mathrm{x}_n$ فالمتوسط الهندسي لهذه القيم ونرمز له بــ H هو الجذر النوني

$$G_x = \sqrt[n]{\mathrm{x}_1 \times \mathrm{x}_2 \times \dots \dots \times . \mathrm{x}_n} = \left(\prod_1^n x_i\right)^{\frac{1}{n}}$$

مثال اوجد المتوسط الهندسي للقيم التالية : 0.15، 0.24، 0.35، 0.5

$$G_x = \sqrt[n]{\mathrm{x}_1 \times \mathrm{x}_2 \times \dots \dots \times . \mathrm{x}_n} = \left(\prod_1^n x_i\right)^{\frac{1}{n}} = (0.15 \times 0.24 \times 0.35 \times 0.5)^{\frac{1}{4}}$$
$$= 0.281$$

في حالة بيانات مرجحة :

$$G_x = \sqrt[\sum n_i]{\mathrm{x}_1{}^{\mathrm{n}_1} \times \mathrm{x}_2{}^{\mathrm{n}_2} \times \dots \dots \times . \mathrm{x}_n{}^{\mathrm{n}_n}} = \left(\prod_1^n x_i{}^{n_i}\right)^{\frac{1}{\sum n_i}}$$

مثال الجدول التالي يشير الى معدلات الزيادة في السكان لمدة 10 سنوات

معدلات الزيادة	0.05	0.075	0.1	0.15

عدد السنوات	3	4	2	1

المتوسط الهندسي هو

$$G_x = \sqrt[\sum n_i]{\mathrm{x}_1\, n_1 \ \times \mathrm{x}_2\, \mathrm{n}_2 \ \times \ldots \ldots \times . \mathrm{x}_n\, \mathrm{n}_n}$$
$$= \sqrt[10]{0.05^3 \times 0.075^4 \times 0.1^2 \times 0.15^1} = 0.075$$

مزايا وعيوب المتوسط الهندسي :

يمتاز المتوسط الهندسي بكونه لا يتأثر بالقيم المتطرفة أو الشاذة، وحتى في حال تأثره فيها فإنه يكون تأثيرًا ضعيفًا. كما أنه شامل لجميع قيم المتغير الأمر الذي يجعله من أكثر المقاييس تعبيرًا عن كافة البيانات. ولكن يشوب هذا المتوسط عيب ألا وهو عدم إمكانية استخدامه مع البيانات التي تحتوي قيمًا سالبة أو قيمة الصفر.

7.4- **المتوسط التوافقي :**

وهو احد مقاييس النزعة المركزية وهو مقلوب الوسط الحسابي لمقلوبات هذه القيم ويفضل استخدامه على باقي المتوسطات في حالة إيجاد معدل السرعات ومعدلات التغيير اي في حالة وجود علاقة عكسية بين xi و ni ولا يمكن استخدامه في حالة إذا كانت إحدى هذه القيم مساوية الى الصفر.

في حالة بيانات غير مبوبة وغير مرجحة

$$H_x = \frac{1}{\frac{1}{n}\sum \frac{1}{x_i}} = \frac{n}{\sum \frac{1}{x_i}}$$

مثال : اوجد المتوسط التوافقي للقيم التالية 100، 110، 90، 80

$$H_x = \frac{n}{\sum \frac{1}{x_i}} = \frac{4}{\frac{1}{80} + \frac{1}{90} + \frac{1}{100} + \frac{1}{110}} = 93.67$$

في حالة بيانات مرجحة :

$$H_x = \frac{\sum n_i}{\sum \frac{n_i}{x_i}} \qquad ; H_x = \frac{1}{\sum \frac{f_i}{x_i}}$$

مثال :قام محلل أسهم في احد البورصات بتحديد نسبة السعر إلى الربحية للمؤشر الذي يتتبع أسهم شركة W خلال 5 سنوات القيمة السوقية للشركة هي كالتالي 2 ، 9 ، 10 ، 12 ، 13 مليار اما المؤشر فكان كالتالي 0.5، 3 ،5 ،7 ،9 . احسب المتوسط التوافقي للمؤشر.

القيمة السوقية (السعر)	2	9	10	12	13
الارباح (الربح)					
نسبة السعر الى الربح	0.5	3	5	7	9
$\sum n_i$	46=13+12+10+9+2				
$\frac{n_i}{x_i}$	$\frac{2}{0.5} = 5$	$\frac{9}{3} = 3$	$\frac{10}{5} = 2$	$\frac{12}{7} = 1.71$	$\frac{13}{9} = 1.44$
$\sum \frac{n_i}{x_i}$	13.15=1.44+1.71+2+3+5				

$$H_x = \frac{\sum n_i}{\sum \frac{n_i}{x_i}} = \frac{46}{13.15} = 3.49$$

8.4- المتوسط التربيعي :

هو الجذر المربع للوسط الحسابي لمربعات القيم x_i

حالة بيانات غير مرجحة	حالة بيانات مرجحة	حالة بيانات مرجحة بتكرار نسبي
$Q = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}}$	$Q = \sqrt{\frac{\sum_{i=1}^{n} n_i x_i^2}{\sum n_i}}$	$Q = \sqrt{\sum_{i=1}^{n} f_i x_i^2}$

الفصل الرابع : مقاييس التشتت :

تقدم مقاييس النزعة المركزية ملخصات رقمية لمجموعة من القيم او الملاحظات ولكن وحدها لا تكفي لمعرفة شكل القيم الأخرى هل القيم هي بعيدة او قريبة من المتوسط ؟ما مدى اختلاف القيم عن بعضها البعض ؟

يقصد بالتشتت او الاختلاف :(هو التباعد أو التقارب الموجود بين قيم المشاهدات للعينة التابعة لمتغير ما) ومقاييس التشتت تحدد مدى تشتت قيم المشاهدات عن وسطها .

كلما كان مقياس التشتت كبيرا دل ذلك على عدم التجانس بين القيم , بينما يكون مقياس التشتت صغيرا عندما تكون الاختلافات بين قيم المشاهدات قليلة .

ان مقاييس التوسط (المتوسطات) تعطينا فكرة عن مكان تمركز قيم المشاهدات , أما مقاييس التشتت تعطينا فكرة عن مدى تجانس أو تباين هذه القيم حول مركزها أي درجة انتشارها .

ان أهمية مقاييس التشتت تتمثل في وصف التوزيعات ومقارنتها مع بعضها , حيث أن مقاييس التموضع وحدها لا تكفي لهذا الغرض . فقد يتساوى الوسط الحسابي لمجموعتين من القيم مثلا بينما يختلف مدى انتشار قيم لمجموعة الأولى عن انتشار قيم المجموعة الثانية .

1- المدى :

هو أبسط مقاييس التشتت ،ويعتبر مقياسا للتشتت إعطاء فكرة سريعة عن طبيعة البيانات عادة ما يستخدم وصف درجات الحرارة القصوى والدنيا في الأحوال الجوية ويحسب المدى في حالة البيانات غير المبوبة بتطبيق المعادلة التالية

المدى =أكبر قيمة – أصغر قيمة

اما في حالة البيانات المبوبة فهناك طريقتين :

- المدى= مركز الفئة الأخيرة - مركز الفئة الأولى
- المدى= الحد الثاني للفئة الأخيرة – الحد الأول للفئة الأولى

من مزايا وعيوب المدى أنه :

- أنه بسيط وسهل الحساب
- يكثر استخدامه عند الإعلان عن حالات الطقس، و المناخ الجوي، مثل درجات الحرارة، والرطوبة، والضغط الجوي.
- يستخدم في مراقبة الجودة .
- أنه يعتمد على قيمتين فقط ، ولا يأخذ جميع القيم في الحسبان .
- يتأثر بالقيم الشاذة .

2- الانحراف الربيعي :

يعتمد المدى على قيمتين متطرفتين ، هما أصغر قراءة ، وأكبر قراءة ، فإذا كان هناك قيم شاذة، ترتب على استخدامه كمقياس للتشتت نتائج غير دقيقة، من أجل ذلك لجأ الإحصائيون، إلى استخدام مقياس للتشتت يعتمد على نصف عدد القيم الوسطى، ويهمل نصف عدد القيم المتطرفة، ولذا لا يتأثر هذا المقياس بوجود قيم شاذة، ويسمى هذا المقياس بالانحراف الربيعي (Q)، ويحسب الانحراف الربيعي بتطبيق المعادلة التالية

$$DQ = \frac{Q_3 - Q_1}{2}$$

الانحراف الربيعي = (الربيع الثالث – الربيع الأول) / 2

الانحراف الربيعي = نصف المدى الربيعي

من مزايا وعيوب الانحراف الربيعي انه يفضل استخدامه كمقياس للتشتت في حالة

وجود قيم شاذة ، كما أنه بسيط وسهل في الحساب . ومن عيوبه ، أنه لا يأخذ كل القيم في الاعتبار.

3- التباين والانحراف المعياري :

هو أحد مقاييس التشتت ، وأكثرها استخداما في النواحي التطبيقية ، ويعبر عن متوسط مربعات انحرافات القيم عن وسطها الحسابي.

أولا: التباين في المجتمع $v(x)$ أو $\delta^2(x)$

إذا توافر لدينا قراءات عن كل مفردات المجتمع ، ولتكن: $x_1, x_2, \dots \dots . x_n$ ، فإن التباين في المجتمع ، ويرمز له بالرمز $v(x)$ يحسب باستخدام المعادلة التالية :

$$v(x) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{N}$$

اما الانحراف المعياري فهو جذر التباين ويرمز له ب δ (سيڤما)

$$\delta(x) = \sqrt{v(x)}$$

1.3- حالة بيانات كمية غير مرجحة وغير مبوبة :

$$v(x) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{N}$$

أو

$$v(x) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{N} = \frac{\sum_{i=1}^{n}(x_i^2 - 2x_i\bar{x} + \bar{x}^2)}{N}$$
$$= \frac{\sum_{i=1}^{n} x_i^2}{N} - 2\bar{x}\frac{\sum_{i=1}^{n} x_i}{N} + \frac{N\bar{x}^2}{N}$$

$$= \frac{\sum_{i=1}^{n} x_i^2}{N} - 2\bar{x}\bar{x} + \bar{x}^2 = \frac{\sum_{i=1}^{n} x_i^2}{N} - \bar{x}^2$$

$$v(x) = \frac{\sum_{i=1}^{n} x_i^2}{N} - \bar{x}^2$$

مثال رقم 07 :

لتكن القيم التالية عبارة عن أسعار سلع الخضر في سوق معين :

50 ، 67 ، 87، 92، 135، 120، 110، 80

نقوم بحساب المتوسط الحسابي :

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{N} = \frac{50+67+87+92+135+120+110+80}{8} = \frac{741}{8}$$
$$= 92.625\ da$$

متوسط سعر الخضر هو في السوق هو 92.625 دج

$$v(x) = \frac{\sum_{i=1}^{n} x_i^2}{N} - \bar{x}^2$$

$$= \frac{(50-92.625)^2 + (67-92.625)^2 + (87-92.625)^2 + (92-92.625)^2 + (135-92.625)^2}{8}$$

$$\frac{+(120-92.625)^2 + (110-92.625)^2 + (80-92.625)^2}{8} = \frac{1511.88}{8} = 688.98$$

$$\delta(x) = \sqrt{v(x)} = \sqrt{688.98} = 26.24$$

أسعار الحضر تبتعد عن متوسط السعر بمتوسط انحرافات 26.24 دج

2.3- حالة بيانات كمية مرجحة (مبوبة وغير مبوبة)

في هذه الحالة نطبق نفس القانون لان xi سوف تكون مكان ci .

$$v(x)=\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2 n_i}{\sum_{i=1}^{n} n_i}$$

أو

$$v(x)=\frac{\sum_{i=1}^{n} x_i^2 n_i}{\sum_{i=1}^{n} n_i}-\bar{x}^2$$

أو باستعمال f_i مكان n_i مع $f_i=\frac{n_i}{\sum n_i}$ $\quad n_i=f_i\times\sum n_i$

$$v(x)=\frac{\sum_{i=1}^{n} x_i^2 n_i}{\sum_{i=1}^{n} n_i}-\bar{x}^2=\frac{\sum_{i=1}^{n} x_i^2 f_i\times\sum n_i}{\sum_{i=1}^{n} n_i}-\bar{x}^2$$

$$=\frac{\sum n_i\sum_{i=1}^{n} x_i^2 f_i}{\sum_{i=1}^{n} n_i}-\bar{x}^2=\sum_{i=1}^{n} x_i^2 f_i-\bar{x}^2$$

$$v(x)=\sum_{i=1}^{n} x_i^2 f_i-\bar{x}^2$$

3.3- خواص التباين والانحراف المعياري :

- 1.3.3- خاصية الجمع والطرح :

تطبق هذه الخاصية عندما بتم إضافة او اقتطاع قيمة ثابتة من قيم السلسلة الإحصائية .

فإذا كانت القيم هي : $x_1, x_2, \dots, x_n$ ، وتم إضافة مقدار ثابت (a) إلى كل قيمة من القيم ،

$$v(x+a)=\frac{\sum_{i=1}^{n}(x_i+a)^2 n_i}{\sum_{i=1}^{n} n_i}-\left(\overline{(x+a)}\right)^2$$

$$=\frac{\sum_{i=1}^{n}({x_i}^2 n_i+2ax_i n_i+a^2 n_i)}{\sum_{i=1}^{n} n_i}-$$

$$=\frac{\sum_{i=1}^{n}({x_i}^2 n_i)}{\sum_{i=1}^{n} n_i}+\frac{\sum_{i=1}^{n}(2ax_i n_i)}{\sum_{i=1}^{n} n_i}+\frac{\sum_{i=1}^{n}(a^2 n_i)}{\sum_{i=1}^{n} n_i}-(\bar{x}^2+2\bar{x}a+a^2)$$

$$=\frac{\sum_{i=1}^{n}{x_i}^2 n_i}{\sum_{i=1}^{n} n_i}+2a\frac{\sum_{i=1}^{n}x_i n_i}{\sum_{i=1}^{n} n_i}+a^2\frac{\sum_{i=1}^{n} n_i}{\sum_{i=1}^{n} n_i}-(\bar{x}^2+2\bar{x}a+a^2)$$

$$\frac{\sum_{i=1}^{n}{x_i}^2 n_i}{\sum_{i=1}^{n} n_i}+2a\bar{x}+a^2-\bar{x}^2-2a\bar{x}-a^2$$

$$v(x\pm a)=\frac{\sum_{i=1}^{n}{x_i}^2 n_i}{\sum_{i=1}^{n} n_i}-\bar{x}^2$$

إذا أضيف مقدار ثابت إلى كل قيمة من القيم ، فإن التباين للقيم المعدلة (بعد الإضافة) لا يتغير .

✓ خاصية الضرب والقسمة :

تطبق هذه الخاصية عندما يتم إضافة او اقتطاع قيمة غير ثابتة من قيم السلسلة الإحصائية.

فإذا كانت القيم هي : $x_1, x_2, \dots, x_n$ ، وتم إضافة مقدار غير ثابت (a) إلى كل قيمة من القيم.

$$v(ax)=\frac{\sum_{i=1}^{n}(ax_i)^2 n_i}{\sum_{i=1}^{n} n_i}-\left(\overline{(ax)}\right)^2$$

$$=\frac{\sum_{i=1}^{n}a^2{x_i}^2 n_i}{\sum_{i=1}^{n} n_i}-a^2\bar{x}^2=a^2\left(\frac{\sum_{i=1}^{n}{x_i}^2 n_i}{\sum_{i=1}^{n} n_i}-\bar{x}^2\right)$$

$$v(ax)=a^2v(x)$$

$$v\left(\frac{x}{a}\right)=\frac{1}{a^2}v(x)$$

مثال رقم 08 :

الجدول التالي يبين توزيع العمال حسب الأجور (x 1000 دج)

الاجر	[20-24[	[24-26[	[26-30[	[30-38[	[38-42[	[42-50[
عدد العمال	22	18	26	16	10	8

- احسب المتوسط الحسابي والانحراف المعياري
- ادا ارادت المؤسسة تقديم علاوة ثابتة في الأجور قدرها 1500 دج ما هو المتوسط و الانحراف المعياري الجديدين .
- قامت إدارة المؤسسة بعملية تحفيض الأجور بنسبة 5 % ما هو المتوسط و الانحراف المعياري الجديدين

c_i	n_i	f_i	x_i	x_in_i	$(x_i-\bar{x})$	$(x_i-\bar{x})^2n_i$	${x_i}^2n_i$	x_if_i	${x_i}^2f_i$
[20-24[	22	0,22	22	484	-7,74	1317,97	10648	4,84	106,48
[24-26[	18	0,18	25	450	-4,74	404,42	11250	4,5	112,5
[26-30[	26	0,26	28	728	-1,74	78,72	20384	7,28	203,84
[30-38[	16	0,16	34	544	4,26	290,36	18496	5,44	184,96
[38-42[	10	0,1	40	400	10,26	1052,68	16000	4	160
[42-50[	8	0,08	46	368	16,26	2115,10	16928	3,68	169,28
	100			2974		5259,24	93706	29,74	937,06

حساب المتوسط الحسابي :

$$\bar{x}=\frac{\sum_{i=1}^{n}x_in_i}{\sum_{i=1}^{n}n_i}=\frac{2974}{100}=29.7$$

اما القانون الثاني باستخدام النسب:

$$\bar{x} = \sum_{i=1}^{n} x_i f_i = 29.74$$

متوسط الاجر هو 2974 دج

حساب التباين والانحراف المعياري :

$$v(x) = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2 n_i}{\sum_{i=1}^{n} n_i} = \frac{5259.24}{100} = 52.5924$$

$$\delta(x) = \sqrt{v(x)} = \sqrt{52.5924} = 7.25$$

حساب المتوسط والانحراف المعيار الجديدين بعد إضافة قيمة ثابتة .

$$\overline{x+a} = \bar{x} + a = 2974 + 1500 = 4474\ da$$

$$\delta(x+a) = \sqrt{v(x+a)} = \sqrt{v(x)}$$

التباين والانحراف المعياري لا يتغير.

حساب المتوسط والانحراف المعيار الجديدين بعد إضافة قيمة غير ثابتة

تحفيض الأجور بنسبة 5 % أي :

الاجر الجديد =(الأجر القديم – الاجر القديم مضروب في نسبة التخفيض)

الاجر الجديد =الأجر القديم (1-0.05)

الاجر الجديد =الأجر القديم مضروب 0.95

$$\overline{ax} = a \times \bar{x} = 0.95 \times 2974 = 2825.3\ da$$

$$\delta(ax) = \sqrt{v(ax)} = \sqrt{a^2 v(x)} = \sqrt{0.95^2 v(x)} = 0.95\ \delta(x)$$
$$= 0.95 * 7.25 = 6.8875$$

4- معامل الاختلاف *Coefficient de Variation*

هو عبارة عن قسمة مقاييس التشتت المطلق على مقاييس النزعة المركزية , ويستخدم لغرض المقارنة بين تشتت مجتمعين أو ظاهرتين مختلفتين . والصيغة الرياضية له :

$$cv = \frac{\delta(x)}{\bar{x}}$$

مجال cv	$0 < cv < 0.2$	$0.2 \leq cv < 0.4$	$0.4 \leq cv < 0.6$	$0.6 \leq cv < 0.8$	$0.8 \leq cv \leq 1$
التشتت	تشتت ضعيف جدا	تشتت ضعيف	تشتت ضعيف نوعا ما	تشتت قوي	تشتت قوي جدا

Cv هو محصور بين 0 و 1

مثال : تم اختيار فندقين من بين مجموعة من الفنادق , لمقارنة أيهما أفضل من ناحية الايرادات السنوية , وكان الوسط الحسابي للايرادات السنوية للفندق الاول *(78)* والانحراف المعياري لها *(80)* , اما الفندق الثاني فكان الوسط الحسابي لايراداته السنوية *(73)* والانحراف المعياري له *(76)* , أي الفندقين كان ايراداته أفضل ؟

الحل :

$$C.V1 = \frac{80}{78} * 100 = 102.56\%$$

$$C.V2 = \frac{76}{73} * 100 = 104.11\%$$

اذا التشتت في الفندق الاول أقل أي انه هو الأفضل.

الفصل الرابع مقاييس الشكل (الالتواء والتفلطح):

تخبرنا مقاييس التشتت مهدى انحراف القيم المفردة بعضها عن بعض ولكنها لا تخبرنا بطريقة انحرافها. وبالتحديد لا توضح ما اذا كانت الانحرافات الأكبر تميل الى ان تكون نحو القيم الكبرى ام القيم الصغرى في السلسلة الإحصائية.

فنحتاج عادة الى مقاييس تضع في الاعتبار وتقيس عدم التناظر في توزيع القيم ويسمى احد هذه الأنواع الالتواء حيث نجد الالتواء موجب او على اليمين او نحو القيم الكبرى مثل توزيع الثروة اين يمتلك العديد من الافراد مبالغ صغيرة ومتقاربة بينما يمتلك القليل من الافراد مبالغ كبيرة.

1-مقاييس الالتواء:

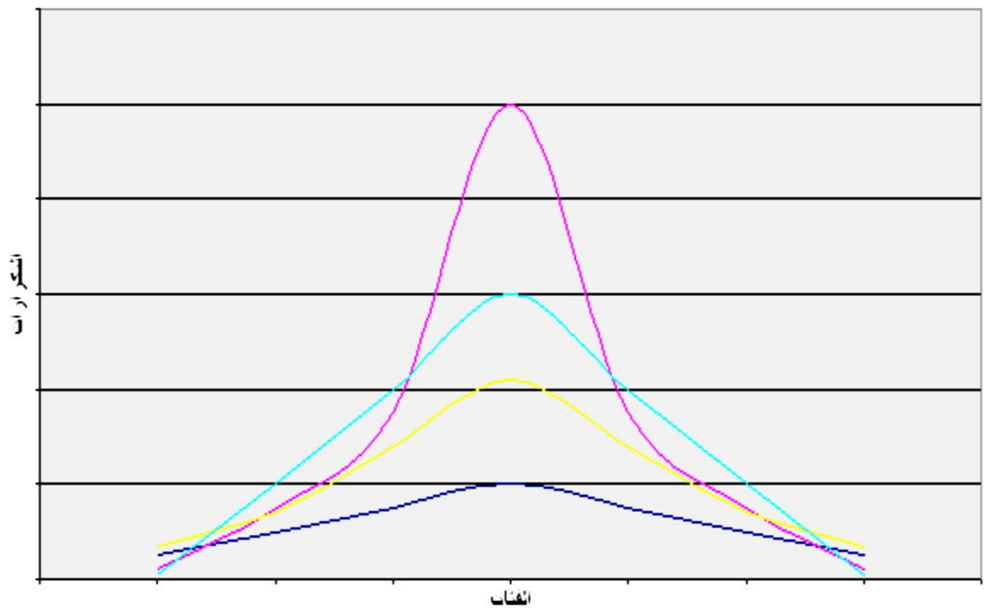
يقصد بالالتواء: عدم تماثل منحنى التوزيع التكراري حول نقطة المركز المتوسط $\bar{}$.

فيكون منحنى التوزيع التكراري متماثلا حول نقطة المركز (المتوسط) إذا أسقطنا عمود من قمة المنحنى التكراري وقسمه إلى قسمين متطابقين ، أما عكس ذلك فيكون التوزيع غير متماثل أي ملتو إما إلى جهة اليمين أو إلى جهة اليسار.

المنحنى التكراري لأي توزيع يأخذ أحد الأشكال الآتية:

- المنحنى المتماثل هو المنحنى الذي إذا قسمناه إلى نصفين أنطبق هذان النصفان على بعضهما البعض تماما

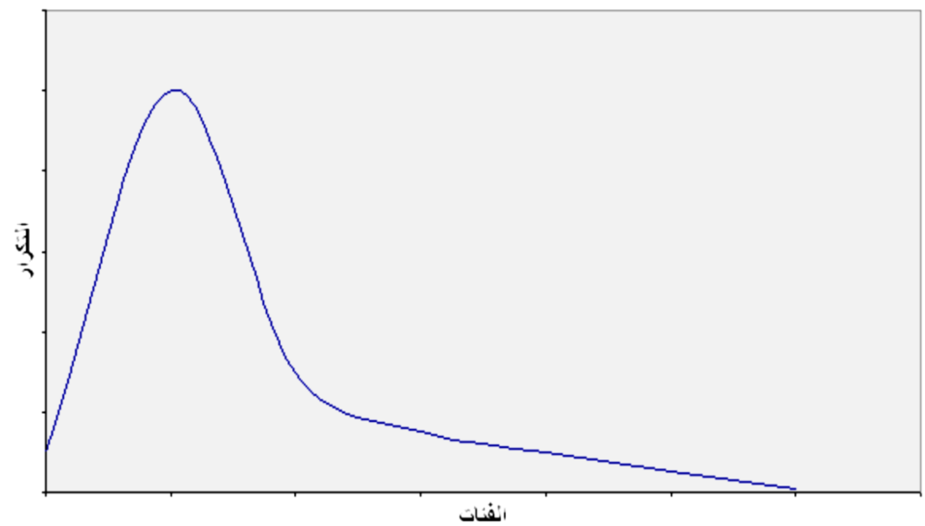
شكل يوضح منحنيات التوزيع المتماثل



المنحنى المتماثل (الوسط الحسابي = الوسيط = المنوال)

- إذا كانت التكرارات تتركز عند أصغر القيم يصبح المنحنى ملتويا التواء موجب جهة اليمين كما يظهر في الشكل التالي:

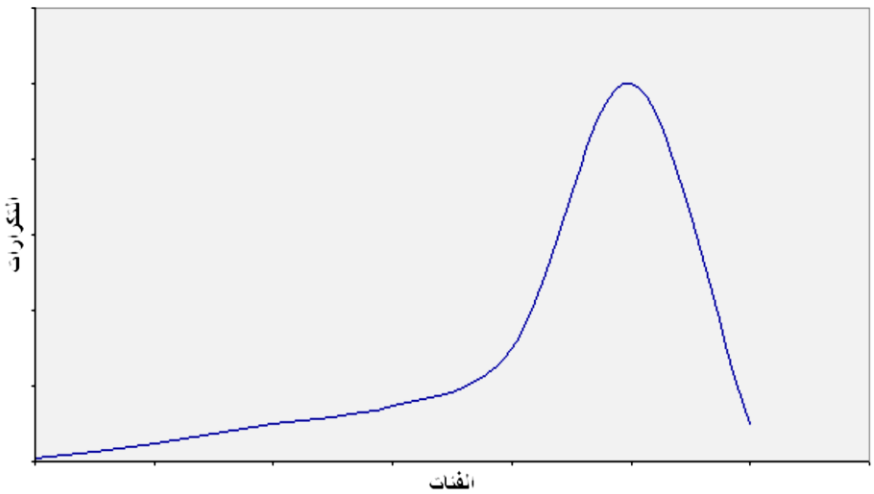
شكل يوضح منحنى ملتوى جهة اليمين



التوزيع (المنحنى) ملتوي التواء موجب جهة اليمين (الوسط الحسابي > الوسيط >المنوال)

- أما في حالة تركز التكرارات عند أكبر القيم فيكون المنحنى في تلك الحالة ملتوي التواء سالب جهة اليسار يظهر من الشكل التالي:

شكل يوضح منحنى ملتوى جهة اليسار



التوزيع ملتوي جهة اليسار (الوسط الحسابي <الوسيط< المنوال)

1.1- قياس باستعمال مقاييس النزعة المركزية :

يمكن قياس الالتواء بمعامل الالتواء والذي يفيدنا في الحكم على مدى تماثل أو التواء التوزيع .

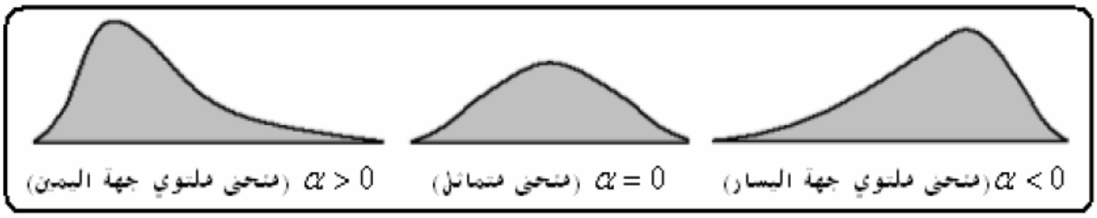
حيث تأخذ هذه الطريقة في الاعتبار العلاقة بين الوسط والوسيط والمنوال، في حالة ما إذا كان التوزيع قريب من التماثل وليس شديد الالتواء ، وهذه العلاقة هي:
المنوال = المتوسط الحسابي – 3 (المتوسط الحسابي – الوسيط)
ومن ثم فإن طريقة "بيرسون" في قياس الالتواء ، تتحدد بالمعادلة التالية

$$\alpha = \frac{3(\bar{x} - M_e\ )}{\sigma} = \frac{(\bar{x} - M_o\ )}{\sigma}$$

ألفا هو معامل "، $\bar{x}$ الوسط الحسابي M_e هو الوسيط σ هو الانحراف المعياري حيث أن الالتواء "لبيرسون
ويمكن من خلال الإشارة التي يأخذها هذا المعامل الحكم على شكل الالتواء، كما يلي:

- $0 = \alpha$ التوزيع متماثل
- $0 < \alpha$ التوزيع ملتوي نحو اليمين او التواء موجب او غير متناظر من جهة اليمين

- $\alpha < 0$ التوزيع ملتوي نحو اليسار او التواء سالب او غير متناظر من جهة اليسار



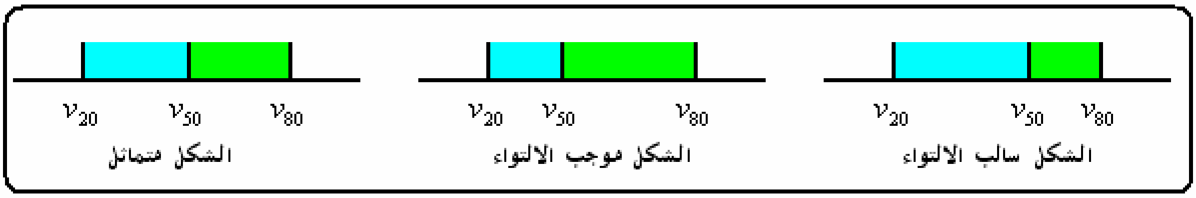
2.1- طريقة " المئين "في قياس الالتواء

ويحسب بالمسافة الموجودة بين مئينين متقابلين حول المئين 50 حسب المعادلة التالية :

ليكن الفا معامل الاتواء المؤيني

$$\alpha = \frac{(c_{100-p} - c_{50}) - (c_{50} - c_p)}{c_{100-p} - c_p}$$

ويمكن ان يكون p المؤين 99 او98 ......او 60



- $\alpha = 0$ التوزيع متماثل
- $\alpha > 0$ التوزيع ملتوي نحو اليمين او التواء موجب او غير متناظر من جهة اليمين

- $0 > \alpha$ التوزيع ملتوي نحو اليسار او التواء سالب او غير متناظر من جهة اليسار

اذ ما استبدلنا p ب25 سنجد معامل الالتواء الربيعي

$$\alpha = \frac{(Q_3 - Q_2) - (Q_2 - Q_{25})}{Q_3 - Q_1}$$

ويعاب على معامل التواء بيرسون بالصيغتين السابقتين :

- أنه لا يمكن حسابه من الجداول التكرارية المفتوحة من أحد الطرفين أو من كليهما.
- كما لا يمكن حسابه في حالة المنحنيات شديدة التواء.

3.1- معامل الالتواء باستعمال العزوم :

يعرف العزم بصفة عامة من خلال المعادلة التالية :

$$\mu_r = \frac{\sum n_i (x_i - \bar{x})^r}{\sum n_i} \qquad \mu_r = \frac{\sum (x_i - \bar{x})^r}{n}$$

R=1 هوالعزم من الدرجة الاولى ويساوي 0

r=2هوالعزم من الدرجة الثانية

R=3هوالعزم من الدرجة الثالثة

R=4 هوالعزم من الدرجة الرابعة

معامل فيشر للالتواء gamma

$$\boldsymbol{\gamma_1 = \frac{\mu_3}{\mu_2^{\frac{3}{2}}}} \qquad \text{او} \qquad \boldsymbol{\gamma_1 = \frac{\mu_3}{\sigma^3}}$$

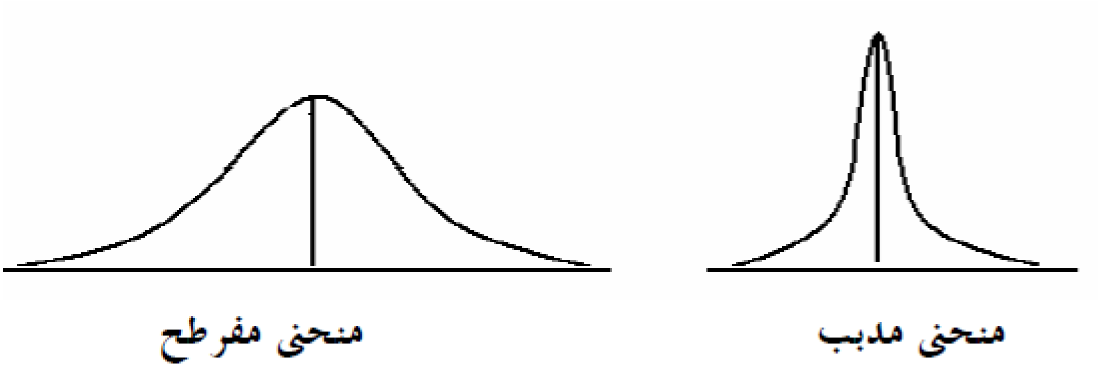
ɣ1=0 متناظر　ɣ1 اصغر من الصفر التواء سالب　　ɣ1 اكبر من الصفر التواء موجب

2-التفلطح :

يقصد بالتفرطح مقدار التدبب (الانخفاض أو الارتفاع) في قمة المنحنى مقارنة بقمة منحنى التوزيع الطبيعي :

وتكون قيمة معامل التفرطح تساوى 3 في حالة التوزيع الطبيعي المعتدل

عند تمثيل التوزيع التكراري في شكل منحنى تكراري ، قد يكون هذا المنحنى منبسط ، أو متطاول ، فعندما يتركز عدد أكبر من القيم بالقرب من منتصف المنحنى، ويقل في طرفيه، يكون المنحنى متطاولا ، وعندما يتركز عدد أكبر على طرفي المنحنى ، ويقل بالقرب من المنتصف يكون المنحنى مفرطحا ، أو منبسطا، ويظهر ذلك من الشكل التالي:



ويمكن قياس التفرطح باستخدام عدد من الطرق،

معامل kelley معامل التفرطح المؤيني = $\frac{\text{نصف المدى الربيعي}}{\text{المدى المؤيني}}$

$$A = \frac{1}{2} \times \frac{Q_3 - Q_1}{D_9 - D_1}$$

وهذا المعامل يساوي 0.26 اذل كان منحنى طبيعي.

$A = 0.26$ منحنى معتدل

$A > 0.26$ منحنى متطاول

$A < 0.26$ منحنى متفلطح

ومنها طريقة العزوم ، حيث يحسب معامل

بتطبيق المعادلة التالية :

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3$$

ɣ2 =3 توزيع معتدل   ɣ2اصغر من الصفر مفلطح   ɣ2 اكبر من الصفر متطاول

مثال رقم 09 :

ليكن الجدول التالي عبارة عن توزيع مجموعة من الأمتعة حسب الوزن كغ

الوزن	[10-14[	[14-18[	[18-22[	[22-26[	[26-30[	[30-34[
عدد الامتعة	10	20	30	25	10	5

احسب معامل الالتواء والتفلطح بطريقتين .

c_i	n_i	$n_i \uparrow$	x_i	$x_i n_i$	$(x_i-\bar{x})$	$(x_i-\bar{x})^2 n_i$	$n_i \uparrow$	$(x_i-\bar{x})^3 n_i$	$(x_i-\bar{x})^4 n_i$
[10-14[	10	10	12	120	-8,8	774,4	10	-6814,72	59969,536
[14-18[	20	30	16	320	-4,8	460,8	30	-2211,84	10616,832
[18-22[	30	60	20	600	-0,8	19,2	60	-15,36	12,288
[22-26[	25	85	24	600	3,2	256	85	819,2	2621,44
[26-30[	10	95	28	280	7,2	518,4	95	3732,48	26873,856
[30-34[	5	100	32	160	11,2	627,2	100	7024,64	78675,968
Σ	100			2080		2656		2534,4	178769,92

- الالتواء باستعمال مقاييس النزعة المركزية أي باستعمال العلاقة

$$\alpha = \frac{3(\bar{x}-M_e)}{\sigma} = \frac{(\bar{x}-M_o)}{\sigma}$$

تطبق هذه العلاقة اذا كان التوزيع قريب التناظر وسنتأكد من ذلك = $3(\bar{x}-M_e)$
$(\bar{x}-M_o)$

المتوسط الحسابي $\bar{x}$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i n_i}{\sum_{i=1}^{n} n_i} = \frac{2080}{100} = 20.8$$

متوسط وزن الامتعة هو 20.8 كغ

الوسيط :

الوسيط ينتمي الى الفئة ذات الرتبة $Rg = \frac{\sum n_i}{2}$ أو $Rg = 0.5$

فباستعمال $n_i \uparrow$

$$Rg = \frac{\sum n_i}{2} = \frac{100}{2} = 50$$

نبحث في قيم $n_i \uparrow$ عن اقرب قيمة التي هي أكبر أو تساوي الرتبة Rg

فنجد أن 60 هي اقرب قيمة ل 50 .

- ✓ إيجاد فئة الوسيط :

نجد ان الوسيط ينتمي الى الفئة]22-18]

$$m_e \in [18 - 22[$$

- ✓ إيجاد قيمة الوسيط :

باستعمال القانون التالي يمكننا إيجاد قيمة الوسيط :

$$m_e = l + \frac{Rg - n_{i-1} \uparrow}{n_i} \times a_i = 18 + \frac{50 - 30}{30} \times 4 = 20.66$$

50% من الأمتعة وزنها أقل من 20.66 كغ و 50% من الأمتعة وزنها أكثر من 20.66 كغ.

المنوال :

المنوال هو القيمة ذات اكبر تكرار وبما أن لدينا فئات متساوية الطول فان المنوال ينتمي الى الفئة ذات اكبر تكرار مطلق $m_o \in [18 - 22[$

فئـــة المنوال = الفئة المناظرة لأكبر تكرار

$$m_o = l + \frac{d_1}{d_1+d_2} \times a_i = 18 + \frac{(30-20)}{(30-20)+(30-25)} \times 4 = 20.66$$

اغلب الأمتعة وزنها 20.66 كغ.

$$3(20.8 - 20.66) = (20.8 - 20.66) \leftrightarrow 0.42 = 0.14$$

العلاقة لم تتحقق وبالتالي لا يمكن حساب الالتواء باستعمال مقاييس النزعة المركزية ولهذا سنحسبها باستعمال معامل الالتواء العشيري او الموين او الربيعي سوف نحتار معامل الالتواء العشيري .

$$\alpha = \frac{(c_{100-p} - c_{50}) - (c_{50} - c_p)}{c_{100-p} - c_p}$$

عند p=10 ستصبح العلاقة بالشكل التالي :

$$\alpha = \frac{(c_{90} - c_{50}) - (c_{50} - c_{10})}{c_{90} - c_{10}} = \frac{(D_9 - M_e) - (M_e - D_1)}{D_9 - D_1}$$

سوف نحسب العشير الأول و التاسع :

العشير الأول ينتمي الى الفئة ذات الرتبة $Rg = \frac{\sum n_i}{10}$ أو $Rg = 0.1$

فباستعمال $n_i \uparrow$

$$Rg = \frac{\sum n_i}{10} = \frac{100}{10} = 10$$

نبحث في قيم $n_i \uparrow$ عن اقرب قيمة التي هي أكبر أو تساوي الرتبة Rg

فنجد أن 10 هي اقرب قيمة ل 10 .

$$D_1 = 14$$

10% من الأمتعة وزنها أقل من 14 كغ و 90% من الأمتعة وزنها أكثر من 14 كغ.

العشير التاسع ينتمي الى الفئة ذات الرتبة $Rg = \frac{9\sum n_i}{10}$ أو $Rg = 0.9$

فباستعمال $n_i \uparrow$

$$Rg = \frac{9\sum n_i}{10} = \frac{900}{10} = 90$$

نبحث في قيم $n_i \uparrow$ عن اقرب قيمة التي هي أكبر أو تساوي الرتبة Rg

فنجد أن 95 هي اقرب قيمة ل 90 .

- ✓ إيجاد فئة العشير التاسع :

نجد ان العشير التاسع ينتمي الى الفئة [26-30[

$$D_9 \in [26 - 30[$$

- ✓ إيجاد قيمة العشير التاسع :

باستعمال القانون التالي يمكننا إيجاد قيمة العشير التاسع :

$$D_9 = l + \frac{Rg - n_{i-1}\uparrow}{n_i} \times a_i = 26 + \frac{90 - 85}{10} \times 4 = 28$$

90% من الأمتعة وزنها أقل من 28 كغ و 10% من الأمتعة وزنها أكثر من 28 كغ.

$$\alpha = \frac{(28 - 20.66) - (20.66 - 14)}{28 - 14} = \frac{7.34 - 6.66}{14} = 0.05$$

التوزيع ملتوي نحو اليمين او التواء موجب او غير متناظر من جهة اليمين ضعيف جدا

الالتواء باستعمال العزوم :

الانحراف المعياري $\delta(x)$

$$\delta(x) = \sqrt{v(x)} = \sqrt{\frac{\sum (x_i - \bar{x})^2 \times n_i}{\sum n_i}} = \sqrt{\frac{2656}{100}} = \sqrt{26.56} = 5.15$$

حساب العزم من الدرجة الثالثة :

$$\mu_3 = \frac{\sum (x_i - \bar{x})^3 \times n_i}{\sum n_i} = \frac{2534.4}{100} = 25.344$$

حساب معامل الالتواء :

$$\gamma_1 = \frac{\mu_3}{\sigma^3} = \frac{25.344}{5.15^3} = 0.80$$

اذن الالتواء موجب ضعيف موعا ما

- التفلطح

التفلطح باستعمال مقاييس النزعة المركزية أي باستعمال معامل التفرطح المؤيني

$$A = \frac{1}{2} \times \frac{Q_3 - Q_1}{D_9 - D_1}$$

سوف نحسب الربيع الثالث والربيع الأول

الربيع الأول ينتمي الى الفئة ذات الرتبة $Rg = \frac{1\sum n_i}{4}$ أو $Rg = 0.25$

فباستعمال $n_i \uparrow$

$$Rg = \frac{\sum n_i}{4} = \frac{100}{4} = 25$$

نبحث في قيم $n_i \uparrow$ عن اقرب قيمة التي هي أكبر أو تساوي الرتبة Rg

فنجد أن 30 هي اقرب قيمة ل 25 .

- ✓ إيجاد فئة الربيع الأول:

نجد ان الربيع الثالث ينتمي الى الفئة [14-18[

$$Q_1 \in [14 - 18[$$

- ✓ إيجاد قيمة الربيع الأول:

باستعمال القانون التالي يمكننا إيجاد قيمة الربيع الأول:

$$Q_1 = l + \frac{Rg - n_{i-1} \uparrow}{n_i} \times a_i = 14 + \frac{25 - 10}{20} \times 4 = 17$$

25% من الأمتعة وزنها أقل من 17 كغ و 75% من الأمتعة وزنها أكثر من 17 كغ.

الربيع الثالث ينتمي الى الفئة ذات الرتبة $Rg = \frac{3\sum n_i}{4}$ أو $Rg = 0.75$

فباستعمال $n_i \uparrow$

$$Rg = \frac{3\sum n_i}{4} = \frac{300}{4} = 75$$

نبحث في قيم $n_i \uparrow$ عن اقرب قيمة التي هي أكبر أو تساوي الرتبة Rg

فنجد أن 85 هي اقرب قيمة ل 75 .

- ✓ إيجاد فئة الربيع الثالث:

نجد ان الربيع الثالث ينتمي الى الفئة [22-26[

$$Q_3 \in [22 - 26[$$

- ✓ إيجاد قيمة الربيع الثالث:

باستعمال القانون التالي يمكننا إيجاد قيمة الربيع الثالث:

$$Q_3 = l + \frac{Rg - n_{i-1}\uparrow}{n_i} \times a_i = 22 + \frac{75 - 60}{25} \times 4 = 24.4$$

75% من الأمتعة وزنها أقل من 24.4 كغ و 25% من الأمتعة وزنها أكثر من 24.4 كغ.

$$A = \frac{1}{2} \times \frac{Q_3 - Q_1}{D_9 - D_1} = \frac{1}{2} \times \frac{24.4 - 17}{28 - 14} = 0.264$$

هذا المعامل يساوي 0.26 اذل كان منحنى طبيعي.

$A = 0.26$ منحنى معتدل

$A > 0.26$ منحنى متطاول

$A < 0.26$ منحنى متفلطح

بالتالي 0.264 اكبر من 0.26 وبالتالي المنحنى فريب الاعتدال ضعيف التطاول

- ➢ التفلطح باستعمال العزوم :

حساب العزم من الدرجة الرابعة :

$$\mu_4 = \frac{\sum (x_i - \bar{x})^4 \times n_i}{\sum n_i} = \frac{178770}{100} = 1787.7$$

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3 = \frac{1787.7}{26.56^2} - \mathbf{3} = \mathbf{2.53} - \mathbf{3} = \mathbf{0.42}$$

الشكل ضعيف التفلطح.

الفصل الخامس : مقاييس التمركز

مقدمة :

1- تعريف : تعرف مقاييس التمركز على انها المقاييس التي تقيس كثافة قيم السلسلة حول قيمة وسطية مثل قياس تمركز الاجور او الدخل .

ولقياس التمركز يجب معرفة القيم الاجمالية و الوسيط النوعي و منحنى لورانز ومعامل جيني .

2- القيم الاجمالية

ليكن لدينا سلسلة احصائية متكونة من n ملاحظة ولكل ملاحظة لديها حجم او كتلة ni فيكون :

$x_i n_i$ تسمى بالتكرارات الاجمالية

${x_i n_i}^{\uparrow}$ تسمى القيم الاجمالية المطلقة الصاعدة

$q_i = \frac{x_i n_i}{\sum x_i n_i}$ وتسمى التكرارات الاجمالية النسبية

${q_i}^{\uparrow}$ تسمى بالتكرارات الاجمالية الصاعدة

3- الوسيط النوعي :

الوسيط النوعي للقيم الاجمالية هو القيمة التي تقسم السلسلة موزونة بالقيم الاجمالية الى قسمين 50 % من القيم الاجمالية اقل من الوسيط النوعي و 50 % من القيم الاجمالية اكثر من الوسيط النوعي. لديه نفس طريقة حساب الوسيط. ويرمز له ب ML

كيفية الحساب :

اولا نحدد رتبة الوسيط النوعي Rg_{ml}

$\mathrm{Rg_{ml}} = \frac{\sum \mathrm{x_i n_i}}{2}$ باستخدام القيم الاجمالية المطلقة

$Rg_{ml} = 0.5$ باستخدام القيم الاجمالية النسبية

تحديد فئة التي ينتمي اليها الوسيط النوعي

نجد الفئة بالبحث عن القيمة في التكرارات الاجمالية المطلقة التي اكبر او تساوي الرتبة (حالة ترار اجمالي مطلق)

او نجد الفئة بالبحث عن القيمة في التكرارات الاجمالية النسبية التي اكبر او تساوي الرتبة (حالة ترار اجمالي نسبي)

ونكتب حسب ${x_i n_i}^{\uparrow}$ او حسب ${q_i}^{\uparrow}$

حساب القيمة :

$ml = l + \frac{Rg_{ml} - (x_{i-1} n_{i-1})^{\uparrow}}{x_i n_i} * a_i$ في حالة تكرارات اجمالية مطلقة

$ml = l + \frac{Rg_{ml} - q_{i-1}^{\uparrow}}{q_i} * a_i$ في حالة تكرارات اجمالية نسبية

حيث :

ml الوسيط النوعي

l الحد الادنى للفئة الوسيط النوعي

Rg_{ml} رتبة الوسيط النوعي

$(x_{i-1} n_{i-1})^{\uparrow}$ التكرار الاجمالي المطلق الصاعد الذي يسبق التكرار الاجمالي المطلق الصاعد لفئة الوسيط النوعي

او $q_{i-1}^{\uparrow}$ التكرار الاجمالي النسبي الصاعد الذي يسبق التكرار الاجمالي النسبي الصاعد لفئة الوسيط النوعي

$x_i n_i$ التكرار الاجمالي المطلق لفئة الوسيط النوعي او q_i التكرار الاجمالي النسبي لفئة الوسيط النوعي

a_i طول فئة الوسيط النوعي

مثال :الجدول التالي يبين مبيعات الشهرية للخضر والفواكه لتجار سوق الجملة للخضر والفواكه المبيعات (بمليون دينار)

المبيعات	[50-100[	[100-150[	[150-200[	[200-250[	[250-300[	[300-350[
عدد التجار	10	50	25	10	3	2

احسب الوسيط والوسيط النوعي

الحل :

ci	x_i	n_i	$n_i^{\uparrow}$	f_i	$f_i^{\uparrow}$	$x_i n_i$	$x_i n_i^{\uparrow}$	q_i	$q_i^{\uparrow}$
[50-100[	75	10	10	0.1	0.1	750	750	0.0497	0.0497
[100-150[	125	50	60	0.5	0.6	6250	7000	0.4139	0.4636
[150-200[	175	25	85	0.25	0.85	4375	11375	0.2897	0.7533
[200-250[	225	10	95	0.1	0.95	2250	13625	0.1491	0.9024
[250-300[	275	3	98	0.03	0.98	825	14450	0.0546	0.957
[300-350[	325	2	100	0.02	1	650	15100	0.043	1
المجموع		100		1		15100	/	1	/

حساب الوسيط :

✓ إيجاد رتبة الوسيط :

الوسيط ينتمي الى الفئة ذات الرتبة $Rg_{me} = \frac{\sum n_i}{2}$ أو $Rg_{me} = 0.5$

فباستعمال $n_i \uparrow$

$$Rg_{me} = \frac{\sum n_i}{2} = \frac{100}{2} = 50$$

نبحث في قيم $n_i \uparrow$ عن اقرب قيمة التي هي أكبر أو تساوي الرتبة Rg_{me}

فنجد أن 60 هي اقرب قيمة ل 50 .

إيجاد فئة الوسيط :

نجد ان الوسيط ينتمي الى الفئة]100-150[

$$m_e \in [100 - 150[$$

✓ إيجاد قيمة الوسيط :

باستعمال القانون التالي يمكننا إيجاد قيمة الوسيط :

$$m_e = l + \frac{Rg - n_{i-1}\uparrow}{n_i} \times a_i$$

$$m_e = l + \frac{Rg - n_{i-1}\uparrow}{n_i} \times a_i = 100 + \frac{50 - 10}{50} \times 50$$
$$= 140$$

50% من التجار يبيعون شهريا أقل من140 مليون دج و 50% من يبيعون شهريا أكثر من140 مليون دج

حساب الوسيط النوعي :

✓ إيجاد رتبة الوسيط النوعي :

الوسيط النوعي ينتمي الى الفئة ذات الرتبة $Rg_{ml} = \frac{\sum x_i n_i}{2}$ أو

$Rg_{ml} = 0.5$

فباستعمال $x_i n_i \uparrow$

$$Rg_{ml} = \frac{\sum n_i}{2} = \frac{15100}{2} = 7550$$

نبحث في قيم $x_i n_i \uparrow$ عن اقرب قيمة التي هي أكبر أو تساوي الرتبة Rg_{ml}

فنجد أن 11375 هي اقرب قيمة ل 7550 .

نبحث في قيم $q_i \uparrow$ عن اقرب قيمة التي هي أكبر أو تساوي الرتبة Rg_{ml}

فنجد أن 0.7533 هي اقرب قيمة ل 0.5 .

إيجاد فئة الوسيط النوعي :

نجد ان الوسيط النوعي ينتمي الى الفئة [150-200[

$$\mathrm{ml} \in [150 - 200[$$

✓ إيجاد قيمة الوسيط النوعي :

باستعمال القانون التالي يمكننا إيجاد قيمة الوسيط النوعي :

$$\mathrm{ml} = l + \frac{Rg_{ml} - (x_{i-1} n_{i-1}) \uparrow}{x_i n_i} \times a_i$$

$$= 150 + \frac{7550 - 7000}{4375} \times 50 = 156.28$$

او باستعمال التكرارات الاجمالية النسبية نجد :

$$ml = l + \frac{Rg_{ml} - q_{i-1}^{\uparrow}}{q_i} * a_i \quad ; \quad ml = 150 + \frac{0.5 - 0.4636}{0.2897} * 50 = 156.28$$

50% من المبيعات الاجمالية الشهرية او رقم الاعمال الشهري للسوق انجزت من طرف التجار الذين يبيعون شهريا أقل من156.28 مليون دج و 50% من المبيعات الاجمالية الشهرية او رقم الاعمال الشهري للسوق انجزت من طرف التجار الذين يبيعون شهريا أكثر من156.28 مليون دج.

الفرق بين الوسيط والوسيط النوعي :

يرمز للفرق بين الوسيط والوسيط النوعي ب ΔM ويحسب كالتالي

$\Delta M = ml - me$ وهو دائما موجب لان ml دائما اكبر منme

يمكننا معرفة درجة التمركز بقسمة الفرق بين الوس النوعي والوسيط على المدى

$$\frac{\Delta M}{E} = \frac{\Delta M}{x_{max} - x_{min}} = \frac{156.28 - 140}{350 - 50} = \frac{16.28}{300} = 0.05$$

وتشير الى تمركز ضعيف

التفسير :

عندما تكون النسبة اقل من 40 % التمركز يكون ضعيف

عندما تكون النسبة بين 40 % و 60 % التمركز يكون متوسط

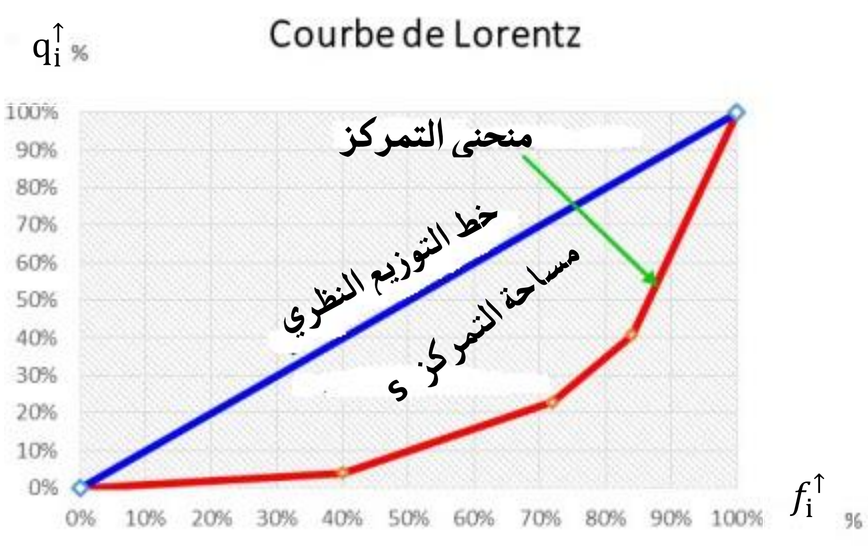
عندما تكون النسبة اكثر من 60 % التمركز يكون قوي

4- منحنى التمركز او منحنى لورانز:

هو تمثيل تخطيطي (بالرسم البياني) لعدم المساواة في الدخل أو الثروة، وضعه الاقتصادي الأمريكي ماكس لورينز (Max Lorenz) في سنة 1905.

يرسم هذا المنحنى في مربع طول كل ضلعه هو 1 او 100 % وهو منحنى الدالة $Q = g(f)$ التي تمثل التكرارات الاجمالية النسبية المتجمعة الصاعدة $q_i^{\uparrow}$ بدلالة التكرارات النسبية المتجمعة الصاعدة $f_i^{\uparrow}$

ولرسم هذا المنحنى نضع على محور الفواصل التكرارات النسبية المتجمعة الصاعدة $f_i^{\uparrow}$ وعلى محور التراتيب التكرارات الاجمالية النسبية المتجمعة الصاعدة $q_i^{\uparrow}$



ايجاد قيمة الوسيط والوسيط النوعي من الرسم يتم بقلب المحاور لمنحنى لورنز ورسم منحنى التكرار المتجمع الصاعد النسبي على يمين الرسم ونقوم باسقاط رتبة الوسيط في منحنى لورنز على منحنى الترار المتجمع الصاعد النسبي على اليمين بخط افقي

5- مؤشر التمركز :او معامل جيني

نسمي معامل جيني نسبة الى الاحصائي الايطالي corrado gini وهو يساوي ضعف المساحة المحصورة بين منحنى التمركز وخط التوزيع النظري وهو محصور بين 0 و 1 فاذا كان معامل جيني قريب من 0 اي ان منحنى التمركز قريب من خط التوزيع النظري فهذا يدل على توزيع ضعيف التمركز اما في حالة معامل جيني قريب من 1 اي ان منحنى التمركز بعيد من خط التوزيع النظري فهذا يدل على توزيع قوي التمركز.

حساب معامل جيني : ضعف مساحة التمركز i=2s هذا اذا كنا نعلم المساحة

اما اذا لم نكن نعلم المساحة فهناك طريقتين :

طريق التكامل اي ان مساحة التمركز هي : $s = \int_0^1 [x - g(x)]dx$

ومعامل جيني هو $i = 2\int_0^1 [x - g(x)]dx$

اما في حالة عدم معرفة دالة التمركز :

سيتم تقسيم مساحة التمرز الى قطع ويتم حساب مجموع مساحات القطع :

- طريقة المثلثات :

  نحسب مؤشر التمركز باستعمال العلاقة التالية :

$$i = \sum_{j=1}^{r-1} (f_j{}^{\uparrow} q_{j+1}{}^{\uparrow} - f_{j+1}{}^{\uparrow} q_j{}^{\uparrow})$$

حيث :

r هو عدد الفئات

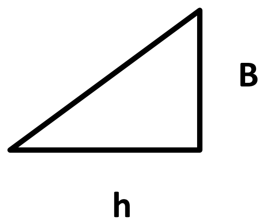
$f_j{}^{\uparrow}$ التكرارات النسبية المتجمعة الصاعدة

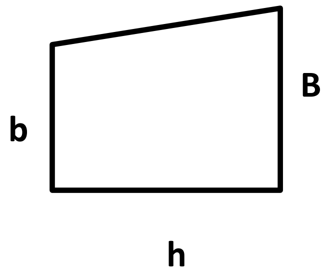
$q_i^{\uparrow}$ التكرارات الاجمالية النسبية المتجمعة الصاعدة

مساحة المثلث

$s = \frac{B * h}{2}$

- طريقة شبه المنحرف :





مساحة شبه المنحرف

$$s = \frac{(B+b)*h}{2}$$

نحسب مؤشر التمركز باستعمال العلاقة التالية :

$$i = 1 - \left[f_1 {q_1}^{\uparrow} + \sum_{j=2}^{r} (f_j ({q_{j-1}}^{\uparrow} + {q_j}^{\uparrow}) \right]$$

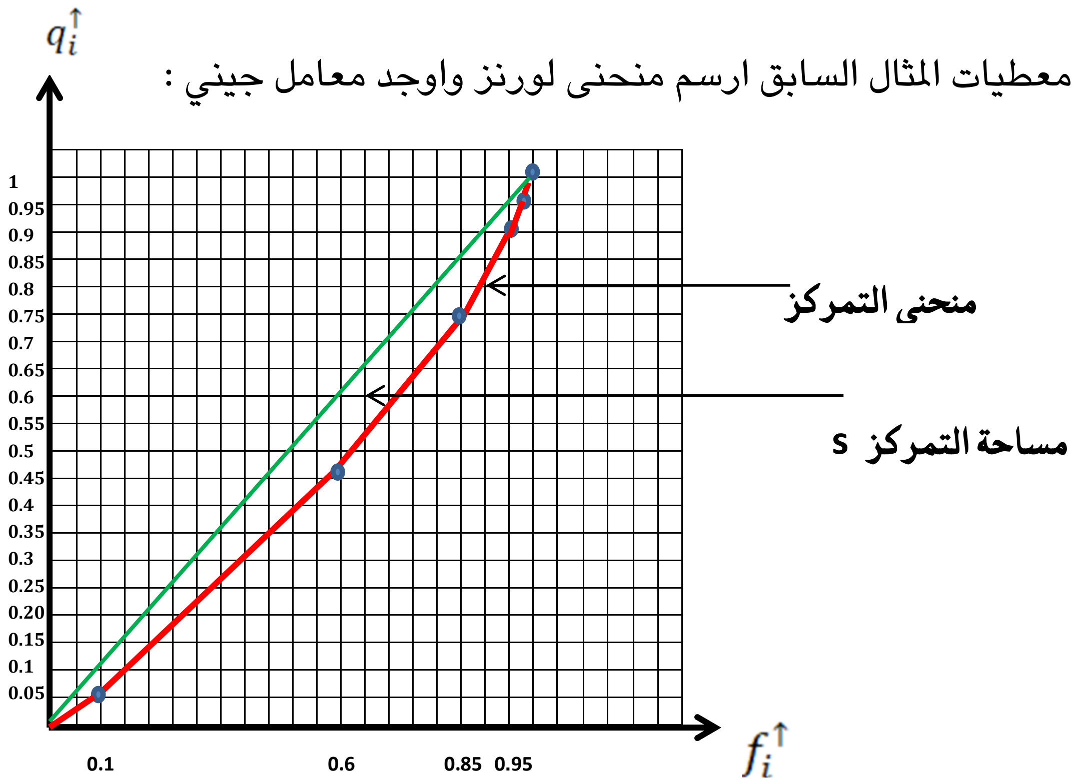
حيث :

r هو عدد الفئات

f_j التكرارات النسبية

$q_i^{\uparrow}$ التكرارات الاجمالية النسبية المتجمعة الصاعدة $q_i^{\uparrow}$

مثال :ناخد معطيات المثال السابق ارسم منحنى لورنز واوجد معامل جيني :



حساب معامل جيني :

باستعمال طريقة المثلثات $i = \sum_{j=1}^{r-1}(f_j{}^{\uparrow} q_{j+1}{}^{\uparrow} - f_{j+1}{}^{\uparrow} q_j{}^{\uparrow})$

ci	$f_j{}^{\uparrow}$	$q_j{}^{\uparrow}$	$f_j{}^{\uparrow} q_{j+1}{}^{\uparrow}$	$f_{j+1}{}^{\uparrow} q_j{}^{\uparrow}$	$f_j{}^{\uparrow} q_{j+1}{}^{\uparrow} - f_{j+1}{}^{\uparrow} q_j{}^{\uparrow}$
[50-100[	0,1	0,0497	0.1∗0.4636=0.04636	0.6∗0.497=0.02982	0.01654
[100-150[	0,6	0,4636	0.6∗0.7533=0.45198	0.85∗0.4636=0.39406	0.05792
[150-200[	0,85	0,7533	0.85∗0.9024=0.76704	0.95∗0.7533=0.71536	0.05168
[200-250[	0,95	0,9024	0.95∗0.957=0.90915	0.98∗0.9124=0.88435	0.0248
[250-300[	0,98	0,957	0.98∗1=0.98	1∗0.957=0.957	0.023
[300-350[	1	1	0	0	0
∑		/			0.173663

باستعمال طريقة شبه المنحرف :

$$i = 1 - \left[f_1 q_1{}^{\uparrow} + \sum_{j=2}^{r} (f_j (q_{j-1}{}^{\uparrow} + q_j{}^{\uparrow}) \right]$$

$$i = 1 - [f_1 q_1{}^{\uparrow} + ((f_2(q_{2-1}{}^{\uparrow} + q_2{}^{\uparrow}) + (f_3(q_{3-1}{}^{\uparrow} + q_3{}^{\uparrow}) + (f_4(q_{4-1}{}^{\uparrow} + q_4{}^{\uparrow}) + (f_5(q_{5-1}{}^{\uparrow} + q_5{}^{\uparrow}) + (f_6(q_{6-1}{}^{\uparrow} + q_6{}^{\uparrow}))]$$

$$i = 1 - [0.1 * 0.0497 + ((0.5(0.0497 + 0.4636) + (0.25(0.4636 + 0.7533) + (0.1(0.7533 + 0.9024) + (0.03(0.9024 + 0.957) + (0.02(0.957 + 1))]$$

$$i = 1 - 0.826337 = 0.173663$$

معامل التمركز هو 0.17 وبالتالي نقول ان التوزيع هو ضعيف التمركز .

ci	f_i	$f_j{}^{\uparrow}$	q_i	$q_j{}^{\uparrow}$	مساحة شبه المنحرف
[50-100[	0,1	0,1	0,0497	0,0497	0,00249
[100-150[	0,5	0,6	0,4139	0,4636	0,12833
[150-200[	0,25	0,85	0,2897	0,7533	0,15211
[200-250[	0,1	0,95	0,1491	0,9024	0,08279
[250-300[	0,03	0,98	0,0546	0,957	0,02789
[300-350[	0,02	1	0,043	1	0,01957
∑	1		1		0,41317
	معامل جيني =2 * (0.5 – (مجموع مساحة شه المنحرفات))				

❋شبه منحرف هو مجموع المثلث والمستطيل

$$i = 2s\ ;\ s = 0.5 - (somme\ des\ surfaces\ de\ trapezes)$$

$$i = 2 * (0.5 - 0.41317) = 0.173663$$

الفصل السادس : الارقام القياسية

مقدمة

تستعمل الارقام بصفة دائمة للتعبير عن الظواهر ا لإحصائية خاصة اذا تعلق الامر بالمقارنة او استنتاج التطور (ارتفاع او تراجع) وتستعمل في ميادين شتى مثل الرقم القياسي لأسعار الجملة واسعار التجزئة والرقم القياسي للواردات والصادرات ويختلف تركيبة كل نوع من هذه الارقام باختلاف الاهمية النسبية للسلع التي تدخل في تركيبة هذا الرقم .

1- تعريف :

تعرف الارقام القياسية على انها مؤشرات احصائية تقيس التغير النسبي الواقع في الزمان والمكان لظاهرة ما او مجموعة من الظواهر.

تفيد الارقام القياسية في استنتاج مدى الاختلاف بين فترتين سواء بالزيادة أو النقصان وغالبا ما يعبر عن ذلك من خلال القسمة اي نسبة قيمة في فترة او مكان حالي الى قيمة في فترة سابقة او مكان آخر.

اذا كان الاختلاف بين فترتين او مكانين يأخذان قيمتين وحيدتين يسمى في هذه الحالة الرقم القياسي البسيط اما إذا كان الاختلاف بين فترتين او مكانين يأخذان عدة قيم مترابطة يسمى في هذه الحالة الرقم القياسي المرجح

التغير النسبي عامة هو قيمة في سنة المقارنة على قيمة في سنة الاساس

2- الرقم القياسي البسيط :

ليكن x متغير كمي يأخذ قيما مختلفة في زمنين مختلفين او مكانين ولتكن x_t قيمة المتغير x في حالة او ظرف t و x_0 قيمة المتغير x في الحالة 0.

نسمي الرقم القياسي البسيط ونرمز له ب $I_{1/0} = \frac{x_1}{x_0}$ ويساوي قيمة x في الحالة t على قيمة x في حالة 0

$$I_{\frac{t}{0}}(x) = \frac{x_t}{x_0} * 100$$

فاذا كان الاختلاف في الزمن نقول ان القيمة x_0 في سنة او شهر او يوم الاساس

والقيمة x_1 في سنة او شهر او يوم المقارنة

مثال : اذا كان سعر اللحوم الحمراء فئة البقر في رمضان 2020 هو 1200 دج وسعر اللحوم الحمراء من نفس الفصيلة في رمضان 2021 هو 1500 دج اوجد الرقم القياسي البسيط لسعر هذا النوع من اللحم

لدينا $x_{2020} = 1200$ و $x_{2021} = 1500$

$$I_{21/20}(x) = \frac{x_{21}}{x_{20}} * 100 = \frac{1500}{1200} * 100 = 125$$

اذا كانت الزيادة تساوي 100 يدل على تساوي الاسعار وبالتالي نقول ان السعر بقي ثابت

اذا كانت الزيادة اكبر من 100 يدل على زيادة الاسعار ب(النتيجة - 100) وبالتالي نقول ان السعر ارتفع

اذا كانت الزيادة اصغر من 100 يدل على تراجع الاسعار ب(100- النتيجة) وبالتالي نقول ان السعر تراجع

العلاقة بين الرقم القياسي و معدل النمو :

ليكن معدل النمو هو r نجد ان معدل النمو مساو لــ $r = \frac{x_{2021}-x_{2020}}{x_{2020}} =$

$\frac{1500-1200}{1200} = 0.25$

نجد ان معدل النمو هو الرقم القياسي -1

الرقم القياسي يساوي معدل النمو + 1

2-1 خواص الارقام القياسية البسيطة :

- خاصية المطابقة :

$$I_{0/0}(x) = \frac{x_0}{x_0} * 100 = 100 \, ; I_{t/t}(x) = \frac{x_t}{x_t} * 100 = 100$$

الرقم القياسي لنفس القيمة هو 1

- خاصية المعكوسية :

$$I_{t/0}(x) \times I_{0/t}(x) = 1 \, ; I_{0/t}(x) = \frac{1}{I_{t/0}(x)}$$

الرقم القياسي للحالة 0 بالنسبة للحالة 1 هو مقلوب للرقم القياسي للحالة 1 بالنسبة للحالة 0

- خاصية الدوران :

$$I_{a/b}(x) \times I_{b/c}(x) \times I_{c/a}(x) = 1$$

- خاصية التحويل :

$$I_{c/b}(x) = \frac{I_{c/a}(x)}{I_{b/a}(x)}$$

- جداء رقمين قياسيين :

اذا كان x مساو ل y مضروب في z اي ان $x = y \times z$ فان الرقم القياسي ل x هو :

$$I_{\frac{t}{0}}(x) = I_{\frac{t}{0}}(y) \times I_{\frac{t}{0}}(z)$$

- مجموع رقمين قياسيين :
- اذا كان x مساو ل y مضروب في z اي ان $x = y + z$ فان الرقم القياسي ل x هو

$$I_{\frac{t}{0}}(x) = I_{\frac{t}{0}}(y) + I_{\frac{t}{0}}(z)$$

اذا كان متوفر معلومات عن سعر سلعة في فترتين او مكانين فيكون $V = P \times Q$

اي ان قيمة سلعة V تتحدد بسعرها P وكميتها Q

$I_{\frac{t}{0}}(p) = \frac{p_t}{p_0} \times 100$ ويسمى الرقم القياسي للأسعار

$I_{\frac{t}{0}}(Q) = \frac{Q_t}{Q_0} \times 100$ ويسمى الرقم القياسي للكميات

$I_{\frac{t}{0}}(V) = \frac{V_t}{V_0} \times 100$ ويسمى الرقم القياسي للقيم

بتطبيق خاصية الجداء نجد ان :

$$V = P \times Q, I_{\frac{t}{0}}(V) = I_{\frac{t}{0}}(p) \times I_{\frac{t}{0}}(Q) * 100; I_{\frac{t}{0}}(V) = \frac{p_t}{p_0} \times \frac{Q_t}{Q_0} \times 100$$

مثال : اشترى احد الاشخاص من سوق الجملة للخضر والفواكه 50 قنطار من البطاطا بسعر 45 دج واشترى في اليوم التالي من نفس السوق 60 قنطار بسعر 44 دج

احسب الرقم القياسي البسيط للسعر والكميات والمبلغ الشراء

$I_{\frac{t}{0}}(p) = \frac{p_t}{p_0} \times 100 = \frac{44}{45} \times 100 = 97.77$ انخفاض في السعر ب 2.23%

$I_{\frac{t}{0}}(p) = \frac{Q_t}{Q_0} \times 100 = \frac{60}{50} \times 100 = 120$ ارتفاع في الكمية ب 20 %

$I_{\frac{t}{0}}(V) = \frac{V_t}{V_0} \times 100 = \frac{44*60}{45*50} \times 100 = 117.33$ ارتفاع في مبلغ الشراء ب 17.33

$$I_{\frac{t}{0}}(V) = I_{\frac{t}{0}}(p) \times I_{\frac{t}{0}}(Q) * 100 = 0.9777 * 1.2 * 100 = 117.33$$

2-2 الارقام القياسية البسيطة المتسلسلة :

اذا توفرت لدينا معلومات $x_1; x_2; x_3; \ldots; x_n$ في فترات او اماكن متتالية 1 ، 2، 3، ...، n ولتكن ارقام القياسية $I_{\frac{2}{1}}(x); I_{\frac{3}{2}}(x); I_{\frac{4}{3}}(x); \quad I_{\frac{5}{4}}(x) \ldots\ldots\ldots\ldots$

فاذا اردنا حساب الرقم القياسي المتسلسل للفترة 6 نسبة الى الفترة 3 تكون مساوية لجداء الارقام القياسية البسيطة داخل هذه الفترة

$$I_{\frac{6}{3}}(x) = I_{\frac{6}{5}}(x); I_{\frac{5}{4}}(x); I_{\frac{4}{3}}(x)$$

3- الارقام القياسية التجميعية :

1.3- الرقم القياسي التجميعي البسيط :

الرقم القياسي البسيط يسمح لنا بمقارنة الاسعار والكميات او القيم لسلعة واحدة فماذا لو كانت لدينا مجموعة السلع فهنا يصبح الرقم القياسي التجميعي البسيط لكل هذه السلع هو مجموع الارقام القياسية مقسومة على n: لنرمز للأرقام القياسية البسيطة ب J

$$I_{\frac{t}{o}} = \frac{J_1 + J_2 + J_3 + \cdots + J_n}{n} ; I_{\frac{t}{o}} = \frac{\sum J_i}{n}$$

مثال : الجدول التالي يبين اسعار اللحوم في رمضان 2021 ورمضان 2022

	لحم الدجاج	لحم الديك الرومي	لحم العجل	لحم الغنم	لحم الماعز
رمضان 2021	260	400	1200	1400	700
رمضان 2022	340	500	1500	2000	900

اوجد الرقم القياسي البسيط لمجموع هذه السلع

	لحم الدجاج	لحم الديك الرومي	لحم العجل	لحم الغنم	لحم الماعز	
رمضان 2021	260	400	1200	1400	700	
رمضان 2022	340	500	1500	2000	900	مجموع
J_i	130.76	125	125	142.85	128.57	652.18

مع

$$J_{21/20} = \frac{p_{21}}{p_{22}} \times 100 \ ; I_{\frac{t}{o}} = \frac{\sum J_i}{n} = \frac{652.18}{5} = 130.436$$

متوسط ارتفاع الاسعار هو 30.436 %

2.3- الرقم القياسي التجميعي المرجح :

نعتبر لو كانت هذه الارقام القياسية البسيطة J_i مرجحة بقيم α_i يصبح الرقم القياسي التجميعي المرجح كالتالي :

$$I_{\frac{t}{o}} = \frac{(J_1 \times \alpha_1) + (J_2 \times \alpha_2) + \cdots + (J_n \times \alpha_n)}{\alpha_1 + \alpha_2 + \cdots + \alpha_n} \ ; I_{\frac{t}{o}} = \frac{\sum J_i \times \alpha_i}{\sum \alpha_i}$$

اذا كانت β_i الثقل النسبي للرقم القياسي حيث $\beta_i = \frac{\alpha_i}{\sum \alpha_i}$ يصبح الرقم القياسي المرجح هو

$$I_{\frac{t}{o}} = \sum J_i \times \beta_i$$

مثال ناخذ معطيات المثال السابق ونفترض ان اللحون كانت لها نسب الطلب التالية

لحم الدجاج 50 % ، لحم الداند 17 %، لحم البقر 10 % ،لحم الغنم 20 %، لحم الماعز 3 %

اوجد الرقم القياسي التجميعي المرجح :

$$I_{\frac{t}{o}} = \sum J_i \times \beta_i = 130.76 * 0.5 + 125 * 0.17 + 125 * 0.10 + 142.85 * 0.20 + 128.57 * 0.03 = 131.55$$

متوسط ارتفاع الاسعار هو 31.55%

4-الارقام القياسية المركبة :

1.4- الرقم القياسي لاسبير :

هو عبارة عن وسط حسابي لأرقام قياسية متسلسلة مرجحة بأثقال الاساس

لاسبير للاسعار	لاسبير للكميات	لا سبير للقيم
$L_{\frac{t}{0}}(p) = \frac{\sum P_t \times Q_0}{\sum P_0 \times Q_0} * 100$	$L_{\frac{t}{0}}(Q) = \frac{\sum P_0 \times Q_t}{\sum P_0 \times Q_0} * 100$	$L_{\frac{t}{0}}(V) = \frac{\sum P_t \times Q_t}{\sum P_0 \times Q_0} * 100$

2.4- الرقم القياسي باش :

هو عبارة عن وسط حسابي لأرقام قياسية متسلسلة مرجحة بأثقال المقارنة

باش للاسعار	باش للكميات	باش للقيم
$P_{\frac{t}{0}}(p) = \frac{\sum P_t \times Q_t}{\sum P_0 \times Q_t} * 100$	$P_{\frac{t}{0}}(Q) = \frac{\sum P_t \times Q_t}{\sum P_t \times Q_0} * 100$	$P_{\frac{t}{0}}(V) = \frac{\sum P_t \times Q_t}{\sum P_t \times Q_t} * 100$

3.4- الرقم القياسي فيشر :

هو الوسط الهندسي لرقم لاسبير وباش

$$F(p) = \sqrt{L(p) \times P(p)}\,; F(p) = \sqrt{L(Q) \times P(Q)} \quad ; F(V) = \sqrt{L(V) \times P(V)}$$

مثال : الجدول التالي يبين اسعار وكميات 3 سلع a وb وc في 2020 والارقام القياسية البسيطة للكميات والسعر في 2022 بالنسبة ل 2020

السلع	p_{2020}	Q_{2020}	$I_{\frac{22}{20}}(p)$	$I_{\frac{22}{20}}(Q)$
a	4	5	200	25
B	10	4	120	125
c	8	5	75	100

1- احسب حصة كل سلعة من ميزانية 2020 (معامل الميزانية) لكل سلعة لسنة الاساس
2- بمساعد ة النتائج السابقة احسب رقم لاسبير للاسعار ل2022 نسبة الى 2020
3- احسب معامل الميزانية الخاص ببكل سلعة لسنة المقارنة 2022
4- بين الر رقم باش للكميات هو متوسط التوافقي للأرقام القياسية البسيطة ثم استنتج صيغة هذا الرقم بدلالة الاسعار والكميات

5-احسب رقم باش للكميات ل2022 نسبة الى 2020

6-احسب رقم فيشر للاسعار ل2022 نسبة الى 2020

الحل :

لينا $I_{\frac{22}{20}}(p) = \frac{p_{22}}{p_{20}} \times 100 \ ; p_{22} = \frac{I_{\frac{22}{20}}(p) \times p_{20}}{100}$

$$I_{\frac{22}{20}}(Q) = \frac{Q_{22}}{Q_{20}} \times 100 \ ; Q_{22} = \frac{(I_{\frac{22}{20}}(Q) \times Q_{20})}{100}$$

ميزانية كل سلعة هو السعر مضروب في الكمية لتلك السلعة

سنحصل على الجدول التالي :

السلع	p_{20}	Q_{20}	p_{22}	Q_{22}	$I_{\frac{22}{20}}(p)$	$I_{\frac{22}{20}}(Q)$	ميزانية كل سلعة في 2020	معامل الميزانية ل2020	ميزانية كل سلعة في 2020	معامل الميزانية ل2020
a	4	5	8	1,25	200	25	20	0,2	10	0,1
B	10	4	12	5	120	125	40	0,4	60	0,6
c	8	5	6	5	75	100	40	0,4	30	0,3
المجموع							100	1	100	1

❋حساب رقم لاسبير للأسعار 2022 اساس 2020

رقم لا سبير تعريفا هو المتوسط الحسابي الارقام القياسية البسيطة مرجحة بسنة الاساس بما اننا نبحث عن لاسبير للسعر فسيكون مرجح بمعاملات الميزانية لسن الاساس .

$$I_{\frac{t}{o}} = \sum J_i \times \beta_i$$

يصبح لدينا $I_{\frac{t}{o}} = L_{22/20}$ ، $J_i = I_{\frac{22}{20}}(p)$ ، β_i معاملات الميزانية لسنة الاساس

$$L_{\frac{22}{20}}(p) = \sum I_{\frac{22}{20}i}(p) \times \beta_{i20}\ ; L_{\frac{22}{20}}(p) = I_{\frac{22}{20}}(p_a) \times \beta_{a20} + I_{\frac{22}{20}}(p_b) \times \beta_{b20} + I_{\frac{22}{20}}(p_c) \times \beta_{c20}$$

$$L_{\frac{22}{20}}(p) = 200 * 0.2 + 120 * 0.4 + 75 * 0.4 = 118$$

الاسعار ارتفعت بنسبة 18 %

$$L_{\frac{22}{20}}(p) = I_{\frac{22}{20}}(p_a) \times \frac{p_{a20} * Q_{a20}}{\sum p_{20i} \times Q_{20i}} + I_{\frac{22}{20}}(p_b) \times \frac{p_{b20} * Q_{b20}}{\sum p_{20i} \times Q_{20i}} + I_{\frac{22}{20}}(p_c) \times \frac{p_{c20} * Q_{c20}}{\sum p_{20i} \times Q_{20i}}$$

$$L_{\frac{22}{20}}(p) = \frac{p_{a22}}{p_{a20}} \times \frac{p_{a20} * Q_{a20}}{\sum p_{20i} \times Q_{20i}} + \frac{p_{b22}}{p_{b20}} \times \frac{p_{b20} * Q_{b20}}{\sum p_{20i} \times Q_{20i}} + \frac{p_{c22}}{p_{c20}} \times \frac{p_{c20} * Q_{c20}}{\sum p_{20i} \times Q_{20i}}$$

نختزل القيم باللون الاصفر والمقام يصبح عامل مشترك نجد

$$L_{\frac{22}{20}}(p) = \frac{p_{a22} * Q_{a20} + p_{b22} * Q_{b20} + p_{c22} * Q_{c20}}{\sum p_{20i} \times Q_{20i}}; L_{\frac{22}{20}}(p) = \frac{\sum p_{22i} \times Q_{20i}}{\sum p_{20i} \times Q_{20i}}$$

بين ان رقم باش للكميات هو المتوسط التوافقي للأرقام القياسية البسيطة مرجح بفترة المقارنة

$$P_{\frac{22}{20}}(Q) = \frac{1}{\sum \frac{\beta_{i22}}{I_{\frac{22}{20}i}(Q)}}$$

$$P_{\frac{22}{20}}(Q) = \frac{1}{\sum \frac{\beta_{i22}}{I_{\frac{22}{20}i}(Q)}} \ ; P_{\frac{22}{20}}(Q) = \frac{1}{\frac{\beta_{a22}}{\frac{Q_{a22}}{Q_{a20}}} + \frac{\beta_{b22}}{\frac{Q_{b22}}{Q_{b20}}} + \frac{\beta_{c22}}{\frac{Q_{c22}}{Q_{c20}}}}$$

$$P_{\frac{22}{20}}(Q) = \frac{1}{\frac{\frac{P_{a22} \times Q_{a22}}{\sum P_{22} \times Q_{22}}}{\frac{Q_{a22}}{Q_{a20}}} + \frac{\frac{P_{b22} \times Q_{b22}}{\sum P_{22} \times Q_{22}}}{\frac{Q_{b22}}{Q_{b20}}} + \frac{\frac{P_{c22} \times Q_{c22}}{\sum P_{22} \times Q_{22}}}{\frac{Q_{c22}}{Q_{c20}}}}$$

$$P_{\frac{22}{20}}(Q) = \frac{1}{\frac{P_{a22} \times Q_{a20}}{\sum P_{22} \times Q_{22}} + \frac{P_{b22} \times Q_{b20}}{\sum P_{22} \times Q_{22}} + \frac{P_{c22} \times Q_{c20}}{\sum P_{22} \times Q_{22}}}$$

$$P_{\frac{22}{20}}(Q) = \frac{\sum P_{22} \times Q_{22}}{\sum P_{22} \times Q_{20}} * 100$$

حساب رقم باش :

$$P_{\frac{22}{20}}(Q) = \frac{\sum P_{22} \times Q_{22}}{\sum P_{22} \times Q_{20}} * 100 \ ; P_{\frac{22}{20}}(Q) = \frac{8 * 1.25 + 12 * 5 + 6 * 5}{8 * 5 + 12 * 4 + 6 * 5} * 100 = 0.84$$

انخفاض الكمية ب 15.25%

حساب رقم فيشر للاسعار

$$F(p) = \sqrt{L(p) \times P(p)}$$

$$P_{\frac{22}{20}}(p) = \frac{\sum P_{22} \times Q_{22}}{\sum P_{20} \times Q_{22}} * 100 \; ; P_{\frac{22}{20}}(Q) = \frac{8 * 1.25 + 12 * 5 + 6 * 5}{4 * 1.25 + 10 * 5 + 8 * 5} * 100 = 1.05$$

$$F(p) = \sqrt{1.18 \times 1.05} = 1.11$$

الاسعار ارتفعت ب 11.35 %

الفصل السابع : الانحدار والارتباط

مقدمة

لطالما نتساءل عند حصولنا على نتائج الظاهرة عن اسباب التي ادت الى حدوث هذه النتائج بشكل معين ونريد ان نعرف مدى تأثير تلك الاسباب على الظاهرة محل الدراسة مثلا نتساءل هل الادخار هو تابع تمما للدخل او نسبيا او يتأثر بظواهر اخرى كعدد افراد العائلة و نسق الحياة ومكان العيش الخ

من هنا استدعت الضرورة البحث عن علاقة رياضية بين المتغيرات تسمح بوضع المتغيرات في نسق يسمى نموذج رياضي في شكل معادلات نستفيد منها لإجراء التنبؤ والبحث عن مقياس يقيس قوة العلاقة بين المتغيرات تساعد في ترتيب مسببات الظاهرة.

1- تعريف التوزيعات الثنائية :

تعرف التوزيعات الثنائية على انها التوزيعات التي تخضع الى متغيرين x وy بتكرار n ويمكن ان تكون x وy متغيرين من نفس النوع (كميان او نوعيان) او يختلفان (نوعي كمي او كمي نوعي)

يرمز للمتغير x ب x_i ويأخذ قيم من 1 الى p حيث p عدد مشاهدات المتغير x

يرمز للمتغير x ب y_i ويأخذ قيم من 1 الى q حيث q عدد مشاهدات المتغير y

جدول التوزيعات الثنائية :

يعبر جدول التوزيعات الثنائية عن توزيع عدد مشاهدات المتغير ين x وy معا ويأخذ الشكل التالي :

$x_i \backslash y_i$	y_1	y_2	*****	y_q	$\sum$
x_1	n_{11}	n_{11}	*****	n_{11}	n_{1j}
x_2	n_{11}	n_{11}	*****	n_{11}	n_{2j}
*****			*****		$n_{..j}$
x_p	n_{11}	n_{11}	*****	n_{11}	n_{pj}
$\sum$	n_{i1}	n_{i2}	$n_{i..}$	n_{iq}	n

حيث :

$x_i = x_1, x_2, \dots, x_p$ هي مشاهدات المتغير x و $i = 1,2,..;p$ هي حالات المتغير x

$y_i = y_1, y_2, \dots, y_q$ هي مشاهدات المتغير y و $j = 1,2,..;q$ هي حالات المتغير y

n_{ij} هي تكرارات المطلقة المشتركة بين x وy

n_{11} هو التكرار المشترك بين x_1 و y_1 عندما i=1 و j = 1

n_{21} هو التكرار المشترك بين x_2 و y_1 عندما i=2 و j = 1

n_{63} هو التكرار المشترك بين x_6 و y_3 عندما i=6 و j = 3

$\sum n_{i1}$ هو مجموع التكرارات المشتركة لـ x_i عند y_1 عندما **i=1,2,…,p** و j = 1

ونجدها في الخانة على اقصى يمين الجدول ويسمى بالتوزيع الهامشي لـ x

$\sum n_{3j}$ هو مجموع التكرارات المشتركة لـ x_3 و y_j عندما i=3 و j = 1,2,…,q

ونجدها في الخانة اسفل الجدول ويسمى بالتوزيع الهامشي لـ y

$$\sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij} = n$$

1.2- التكرارات النسبية المشتركة :

تستخدم التكرارات النسبية عندما نريد تبسيط الاعداد في الجدول وتسهيل اجراء المقارنة ويرمز لها بـ f_{ij} وتساوي $f_{ij} = \frac{n_{ij}}{n}$ ومجموعها يساوي

$$\sum_{i=1}^{p}\sum_{j=1}^{q} f_{ij} = 1$$

توزيع التكرارات النسبية المشتركة له نفس شكل توزيع التكرارات المطلقة المشتركة

y_i / x_i	y_1	y_2	*****	y_q	$\sum$
x_1	f_{11}	f_{11}	*****	f_{11}	f_{1j}
x_2	f_{11}	f_{11}	*****	f_{11}	f_{2j}
*****	*****	*****	*****		$f_{..j}$
x_p	f_{11}	f_{11}	*****	f_{11}	f_{pj}
$\sum$	f_{i1}	f_{i2}	$n_{i..}$	f_{iq}	n

f_{ij} هي تكرارات النسبية المشتركة بين x وy

f_{11} هو التكرار المشترك النسبي بين x_1 و y_1 عندما i=1 و j = 1

f_{21} هو التكرار المشترك النسبي بين x_2 و y_1 عندما i=2 و j = 1

f_{42} هو التكرار المشترك النسبي بين x_4 و y_2 عندما i=4 و j = 2

$\sum f_{i1}$ هو مجموع التكرارات المشتركة النسبية لــ x_i عند y_1 عندما i=1,2,...,p و j = 1

ونجدها في الخانة على اقصى يمين الجدول ويسمى بالتوزيع النسبي الهامشي لــ x

$\sum f_{2j}$ هو مجموع التكرارات المشتركة لــ x_2 و y_j عندما i=2 و j = 1,2,...,q

ونجدها في الخانة اسفل الجدول ويسمى بالتوزيع النسبي الهامشي لــ y

2.2-- التوزيعات الهامشية :

هو عبارة عن توزيع لمتغير لوحده ويستخرج توزيع المتغير x من جدول التوزيعات المشتركة بأخذ قيم x و مجموع التكرارات المشتركة المطلقة او النسبية الافقية الموافقة لكل مشاهدة من x.اما في حالة توزيع المتغير y يستخرج من جدول التوزيعات المشتركة بأخذ قيم y و مجموع التكرارات المشتركة المطلقة او النسبية العمودية الموافقة لكل مشاهدة من y

التوزيع الهامشي لــ x

x_i	x_1	x_2	......	x_p	$\sum$
n_i	$\sum n_{1j}$	$\sum n_{2j}$	....	$\sum n_{pj}$	n

التوزيع الهامشي لــ y

y_j	y_1	y_2	......	y_q	$\sum$
n_j	$\sum n_{i1}$	$\sum n_{i2}$	....	$\sum n_{iq}$	n

3.2- التوزيعات الشرطية :

هو عبارة عن توزيع لمتغير عند ثبات المتغير الاخر

توزيع المتغير x_i عند y_1 هي كل الثنائيات $(x_i; n_{i1})$

التوزيع الشرطي لــ x عند y =1

x_i	x_1	x_2	......	x_p	$\sum$
n_{i1}	n_{i1}	n_{i1}	....	n_{p1}	$n_{11}+n_{21}+\cdots+n_{p1}$

التوزيع الشرطي لــ y عند x =5

y_i	y_1	y_2	......	y_q	$\sum$
n_{5j}	n_{51}	n_{52}	....	n_{5q}	$n_{51}+n_{52}+\cdots+n_{5q}$

4.2– متوسط وتباين التوزيعات الثنائية :

الحالة	المتوسط الحسابي	التباين المشترك	التباين
توزيع ثنائي بتكرارات مطلقة	/	$cov(x,y)=\frac{\sum x_i y_j n_{ij}}{\sum n_{ij}}-\bar{x}^2\bar{y}^2$	/
توزيع ثنائي بتكرارات مطلقة	/	$cov(x,y)=\sum x_i y_j f_{ij}-\bar{x}^2\bar{y}^2$	/
التوزيع الهامشي لx	$\bar{x}=\frac{\sum x_i n_i}{\sum n_i}$; $\bar{x}=\sum x_i f_i$	/	$v(x)=\frac{\sum {x_i}^2 n_i}{\sum n_i}-\bar{x}^2$
التوزيع الهامشي لy	$\bar{y}=\frac{\sum y_j n_j}{\sum n_j}$; $\bar{y}=\sum y_j f_j$	/	$v(y)=\frac{\sum {y_j}^2 n_j}{\sum n_j}-\bar{x}^2$
التوزيع الشرطي ل x عند y =1	$\overline{x_{i1}}=\frac{\sum x_{i1} n_{i1}}{\sum n_{i1}}$; $\overline{x_{i1}}=\sum x_{i1} f_{i1}$	/	$v(x)=\frac{\sum {x_{i1}}^2 n_{i1}}{\sum n_{i1}}-\overline{x_{i1}}^2$
التوزيع الشرطي لــ y عند x =5	$\overline{y_{5j}}=\frac{\sum y_{5j} n_{5j}}{\sum n_{5j}}$; $\overline{y_{5j}}=\sum y_{5j} f_{5j}$	/	$v(y)=\frac{\sum {y_{5j}}^2 n_{5j}}{\sum n_{5j}}-\overline{y_{5j}}^2$

مثال : الجدول التالي يبين توزيع الاسر حسب عدد الافراد وعدد الغرف في حي 100 سكن :

ليكن xi هو عدد الافراد و yi هو عدد الغرف

$x_i \backslash y_i$	1	2	3	4	5	$\sum$
2	9	10	6	2	0	27
3	3	12	5	3	1	24
4	0	8	13	7	4	32
5	0	1	4	8	4	17
$\sum$	12	31	28	20	9	100

1- كون جدول التوزيعات المشتركة النسبية

اوجد توزيع الاسر حسب عدد الغرف عندما يكون عدد الافراد هو 3 واحسب متوسط وتباين عدد الغرف

اوجد توزيع الاسر حسب عدد الافراد عندما يكون عدد الغرف هو 4 واحسب متوسط وتباين عدد الافراد

اوجد متوسط عدد الافراد وتباينه ومتوسط عدد الغرف وتباينه

الحل : لدينا $\boldsymbol{f}_{ij} = \frac{\boldsymbol{n}_{ij}}{n}$

$x_i \backslash y_i$	1	2	3	4	5	$\sum$
2	0,09	0,10	0,06	0,02	0,00	0,27
3	0,03	0,12	0,05	0,03	0,01	0,24
4	0,00	0,08	0,13	0,07	0,04	0,32
5	0,00	0,01	0,04	0,08	0,04	0,17
$\sum$	0,12	0,31	0,28	0,2	0,09	1

توزيع الاسر حسب عدد الغرف عندما يكون عدد الافراد هو 3

Y3j	1	2	3	4	5	$\sum$
N3j	3	12	5	3	1	24
Y3j*n3j	3	24	15	12	5	59
Y3j²*n3j	3	48	45	48	25	169

حساب المتوسط الحسابي

$\overline{y_{3J}} = \frac{\sum y_{3j} n_{3j}}{\sum n_{3j}} = \frac{59}{24} = 2.45 \cong 2$ متوسط عدد الغرف للأسر ذات 3 افراد هو غرفين

$$v(y) = \frac{\sum {y_{5j}}^2 n_{5j}}{\sum n_{5j}} - \overline{y_{5j}}^2 = \frac{169}{24} - 2.45^2 = 1.04$$

توزيع الاسر حسب عدد الافراد عندما يكون عدد الغرف هو 4

Xi4	2	3	4	5	∑
Ni4	8	7	3	2	20
Xi4*ni4	16	21	12	10	59
$Xi4^2$*ni4	32	63	48	50	193

حساب المتوسط الحسابي

$\overline{x_{i4}} = \frac{\sum x_{i4} n_{i4}}{\sum n_{i4}} = \frac{59}{20} = 2.95 \cong 3$. متوسط عدد الافراد في الاسر للمساكن ذات 4 غرف هو 3 أفراد

حساب التباين

$$v(x) = \frac{\sum {x_{i4}}^2 n_{i4}}{\sum n_{i4}} - \overline{x_{i1}}^2 = \frac{193}{20} - 2.95^2 = 0.9475$$

التوزيع الهامشي للمتغير x

xi	2	3	4	5	∑
ni	27	24	32	17	100
xi*ni	54	72	128	85	339
xi^2*ni	108	216	512	425	1261

حساب المتوسط الحسابي

$\overline{x_i} = \frac{\sum x_i n_i}{\sum n_i} = \frac{339}{100} = 3.39 \cong 3$. متوسط عدد الافراد في الاسر هو 3 أفراد

حساب التباين

$$v(x) = \frac{\sum {x_{i4}}^2 n_{i4}}{\sum n_{i1}} - \overline{x_{i1}}^2 = \frac{1261}{100} - 3.39^2 = 1.1179$$

التوزيع الهامشي للمتغير y

Yj	1	2	3	4	5	∑
Nj	12	31	28	20	9	100

Yj*nj	12	62	84	80	45	283
Yj²*nj	12	124	252	320	225	933

حساب المتوسط الحسابي

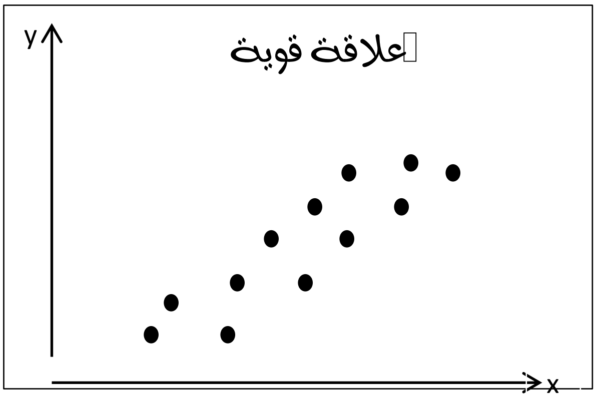
$$\overline{y_j} = \frac{\sum y_j n_j}{\sum n_j} = \frac{283}{100} = 2.83 \cong 3$$ متوسط عدد الغرف للأسر ذات هو 3غرف

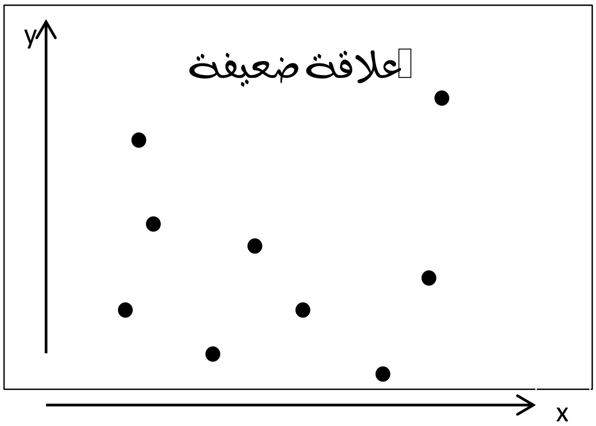
$$v(y) = \frac{\sum {y_j}^2 n_j}{\sum n_j} - \overline{y_j}^2 = \frac{933}{100} - 2.83^2 = 1.32$$

1- الارتباط بين متغيرين :

1.3- سحابة النقاط :

هي عبارة عن مجموعة احداثيات النقاط الممثلة على معلم الذي محوره الافقي هو المتغير المستقل x ومحوره العمودي المتغير y وتسمى النقاط المنتشرة او المتبعثرة على هذا المعلم بسحابة النقاط ويساعد على تكوين استنتاج مبدئي حول امكانية وجود علاقة قوية اوضعيفة بين المتغير x و y





2.3- الانحدار الخطي البسيط :

يعتبر الانحدار الخطي البسيط من ابسط انواع نماذج الانحدار و يعبر عن هذا النموذج بمعادلة مستقيم من الشكل

$y = ax + b$ حيث : y هي قيمة المتغير التابع المقدرة و x هي قيمة المتغير المستقل

و a هو ميل المستقيم ويعكس العلاقة الموجودة بين x وy اما b فهو القيمة التي توافق تقاطع المستقيم مع محور x او عندما تكون x=0 في النموذج.

يفيد مستقيم الانحدار او يسمى بنموذج التعديل في الاقتراب ما يمكن من كل نقاط سحابة النقاط ويمكن الحصول على تعديل جيد برسم مستقيم يدويا في الرسم وايجاد ميل (a) وتقاطع المستقيم مع محور x (b)

او تعديل مثالي باستعمال طريقة المربعات الصغرى حيث ان :

$$y = ax + b \quad ; a = \frac{cov(x;y)}{v(x)} \quad ; b = \bar{y} - a\,\bar{x}$$

مع

$$cov(x;y) = \frac{\sum x_i y_i}{n} - \bar{x} * \bar{y} \quad ; v(x) = \frac{\sum x_i}{n} - \bar{x}^2$$

يسمى الفرق بين قيم y الحقيقية وقيم y المقدرة بخطأ التقدير الخطأ العشوائي ونرمز له بــ e وتزيد جودة التعديل كلما اقترب مجموع الاخطاء من الصفر ان y المقدرة

$$\hat{y} = ax + b + e$$

يرمز لمعامل التحديد بــ R^2 ويقيس جودة التعديل او نسبة تفسير النموذج لتغير الظاهرة ويحسب كالاتي :

$$R^2 = \frac{\text{التغير المفسر}}{\text{التغير الكلي}} = \frac{\sum(\hat{y}_i - \bar{y})}{\sum(y_i - \bar{y})}$$

كلما اقترب معامل التحديد من 1 كلما ذادت جودة التعديل

3.3– معامل الارتباط :

هو معامل يتم اللجوء اليه عندما نريد معرفة قوة العلاقة الموجودة بين المتغيرات او ترتيب ها حسب مدى تأثيرها على الظاهرة ويتصف معامل الارتباط بصفتين الاشارة والقيمة فاذا كانت الاشارة موجبة تشير الى علاقة طردية اما الاشارة السالبة فتدل على العلاقة العكسية بين المتغيرين اما القيمة فكلما اقتربت من الصفر ضعف الارتباط وكلما اقترب من 1 زاد الارتباط ويحسب بالشكل التالي :

مع $0 \leq r \leq 1$ $r = \dfrac{cov(x;y)}{\sigma_x * \sigma_y}$

معامل الارتباط	$0 \leq r \leq 0.2$	$0.2 < r \leq 0.4$	$0.4 < r \leq 0.6$	$0.6 < r \leq 0.8$	$0.8 < r \leq 1$
المجال	ارتباط ضعيف جدا	ارتباط ضعيف	ارتباط متوسط القوة	ارتباط قوي	ارتباط قوي جدا

مثال :

نريد معرفة العلاقة الموجودة بين سعر صرف الدينار مقابل الدولار و حجم الصادرات مليار دولار

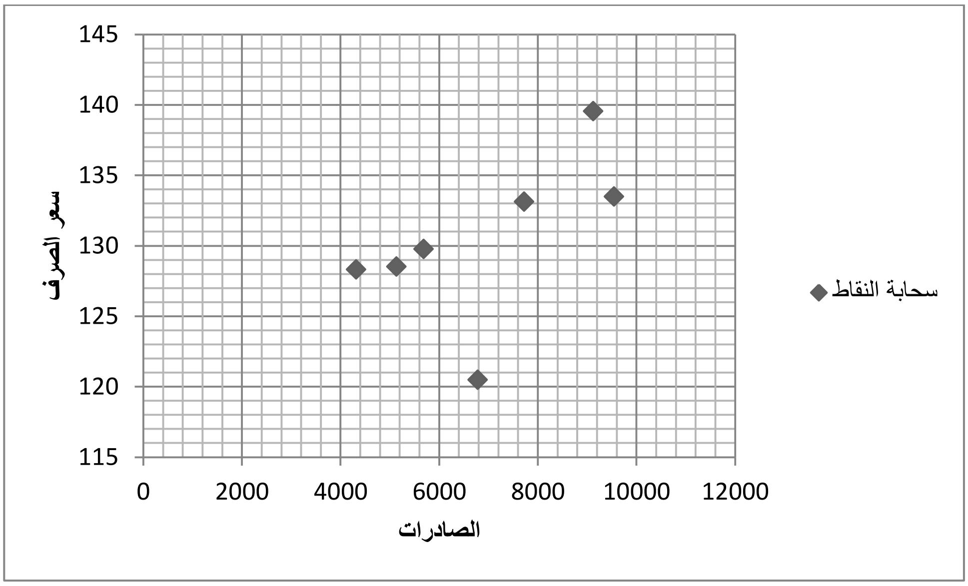
الفصول	فصل 1 2020	فصل 2 2020	فصل 3 2020	فصل 4 2020	فصل 1 2021	فصل 2 2021	فصل 3 2021
سعر الصرف دج/$	120.5	128.33	128.52	129.77	133.14	133.49	139.56
الصادرات مليون $	6782	4321	5136	5686	7725	9550	9127

1- ارسم سحابة النقاط 2- اوجد معادلة الانحدار وارسم المستقيم على نفس الشكل 3- توقع سعر الصرف اذا بلغت الصادرات 10 مليار دولار

2- هل التعديل الخطي هو افضل تعديل فسر ذلك

3- ما هي نسبة تأثير الصادرات على سعر الصرف .

الحل :



معادلة الانحدار :

لدينا تعديل خطي من الشكل

$$y = ax + b \quad ; a = \frac{cov(x;y)}{v(x)} \quad ; b = \bar{y} - a\,\bar{x}$$

x	y	xi yi	xi²
6782	120,5	817231	45995524
4321	128,33	554513,93	18671041
5136	128,52	660078,72	26378496
5686	129,77	737872,22	32330596
7725	133,14	1028506,5	59675625
9550	133,49	1274829,5	91202500
9127	139,56	1273764,12	83302129
48327	913,31	6346795,99	357555911

$$\bar{x} = \frac{\sum x_i}{n} = \frac{48327}{7} = 6903.8$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{913.31}{7} = 130.47$$

$$cov(x;y) = \frac{\sum x_i y_i}{n} - \bar{x} * \bar{y}$$

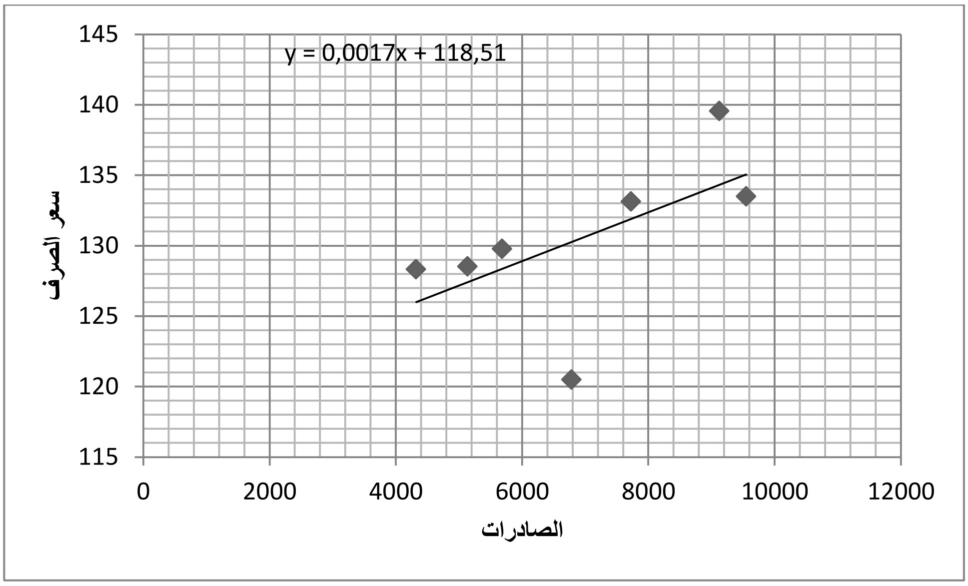
$$cov(x;y) = \frac{\mathbf{6346795{,}99}}{7} - 6903.8 * 130.47 = 5919.17$$

$$v(x) = \frac{\sum {x_i}^2}{n} - \bar{x}^2 \; ; v(x) = \frac{\mathbf{357555911}}{7} - 6903.8^2 = 3416172{,}41$$

$$a = \frac{cov(x;y)}{v(x)} \quad ; a = \frac{5919.17}{3416172{,}41} = 0.0017$$

$$b = \bar{y} - a\,\bar{x} \; ; b = 130.47 - 0.0017 * 6903.8 = 118.51$$

$$y = 0.0017\,x + 118.51$$



توقع سعر الصرف اذا بلغت الصادرات 10 مليار دولار :

$$y = 0.0017\,x + 118.51\ ; y = 0.0017 * 10000 + 118.51\ ; y = 135.83$$

معامل التحديد :

$$R^2 = \frac{\text{التغير المفسر}}{\text{التغير الكلي}} = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2}$$

xi	yi	$\hat{y}$	$(\hat{y}-\bar{y})^2$	$(y-\bar{y})^2$
6782	120,5	130,261716	0,04	99,4578796
4321	128,33	125,997561	20,03	4,59183673
5136	128,52	127,409705	9,38	3,81365102
5686	129,77	128,362686	4,45	0,49400816
7725	133,14	131,895645	2,02	7,11365102
9550	133,49	135,057808	21,02	9,10315102
9127	139,56	134,324879	14,84	82,5761653
48327	913,31		71,79	207,15

$$R^2 = \frac{\sum(\hat{y}_i - \bar{y})^2}{\sum(y_i - \bar{y})^2} = \frac{71.79}{207.15} = 0.34$$

التعديل الخطي ليس افضل تعديل نلاحظ ان جودة التعديل 34% اي ان النموذج فسر الظاهرة بنسبة 34 %

معامل الارتباط :

$$r = \frac{cov(x;y)}{\sigma_x * \sigma_y}$$

x	y	xi yi	xi²	yi²
6782	120,5	817231	45995524	14520,25
4321	128,33	554513,93	18671041	16468,5889
5136	128,52	660078,72	26378496	16517,3904
5686	129,77	737872,22	32330596	16840,2529
7725	133,14	1028506,5	59675625	17726,2596
9550	133,49	1274829,5	91202500	17819,5801
9127	139,56	1273764,12	83302129	19476,9936
48327	913,31	6346795,99	357555911	119369,316

$$v(x) = \frac{\sum {x_i}^2}{n} - \bar{x}^2 \ ; v(x) = \frac{\mathbf{357555911}}{7} - 6903.8^2$$
$$= 3416172{,}41; \sigma(x) = \sqrt{v(x)} = \sqrt{3416172{,}41}$$
$$= 1848.29$$

$$v(y) = \frac{\sum {y_i}^2}{n} - \bar{y}^2 \ ; v(y) = \frac{\mathbf{3119369{,}316}}{7} - 130.47^2$$
$$= 29.59; \sigma(y) = \sqrt{v(y)} = \sqrt{29.59} = 5.44$$

$$r = \frac{cov(x;y)}{\sigma_x * \sigma_y} = \frac{5919.17}{1848.29 * 5.44} = 0.59$$

حجم الصادرات يؤثر على سعر الصرف طرديا بنسبة 59 %

قائمة المراجع :

1- محمد بشير قابيل، سيف الدين بغدادي علم الإحصاء .الموسوعة العربية .
Retrieved 2009-02-13

2- David Freedman (2002) (pdf), From Association to Causation: Some Remarks on the History of Statistics, University of California,Berkeley http://www.stat.berkeley.edu/~census/521.pdf

3- Probability and Statistics on the Earliest Uses Pages (Univ. of Southampton)

4- Electronic Journ@l for History of Probability and Statistics/Journ@l Electronique d'Histoire des Probabilités et de la Statistique

5- علم الاحصاء مقدمة قصيرة جدا ديفيد جيه هند ترجمة احمد شكل مؤسسة هنداوي للتعليم والثقافة 2016

6- مبادئ علم الاحصا ء ،وليد عبد الرحمن الفرا ، المملكة السعودية 1425

7- Probabilites et statistiques de A à Z ,françois dress ,dunod 2004

8- Element de statistiques et de probabilites offices des publications universitaires chamoun chamoun;2010

9- Statistiques descriptives ; offices des publications universitaires,hamdani hocine ;2006

10- Statistique et calcul des probabilites ;walder masierie dalloz ;2004

11- Economie et statistique appliquees ,dominick salvatore ,serie schaum 1984

12- الكامل في الاحصاء عبد الرزاق عزوز ديوان المطبوعات الجامعية 2010

13- الاحصاء للاداريين والاقتصاديين ، دلال القاضي ،سهيلة عبد الله ، محمود البياتي دار حامد للنشر 2003

14- أساسيات الاحصاء و الاحتمالات ،محمد يوسف أشقر ، عبد اللطيف يوسف الصديقي ، دار الراتب الجامعية 2001