علم الإحصاء

تأليف

الدكتور : عماد توما كوش ولاء أحمد القزاز وفاء يونس حمودي

2014

جمهورية العراق
وزارة التعليم العالي والبحث العلمي
هيئة التعليم التقني

علم الأحصاء

تأليف
الدكتور: عماد توما كرش
المعهد التقني الموصل

ولاء احمد القزاز و وفاء يونس حمودي
المعهد التقني نينوى

2014

بسم الله الرحمن الرحيم

المقدمة

اشتقت كلمة الإحصاء من اللفظ اللاتيني " ستاتوس – Status" **بمعنى الدولة. وقد استعمل علم الإحصاء** قديما استعمالات مبكرة تضمنت تجميع البيانات والتخطيط ووصف مظاهرة متعددة **للدولة.**

في عام 1662 قام العالم " جون جرونت " بنشر معلومات إحصائية حول المواليد والوفيات، **ثم تلي على** جرونت بكثير من الدراسات حول الوفيات ونسب ومعدلات الأمراض وأحجام السكان والدخول ونسب **البطالة.**

تعتمد المجتمعات الكبيرة والحكومات على الدراسات الإحصائية كموجه أو مرشد في عمليات الدراسة وأخذ قرارات مستقبلية معينة وإصدار توجيهات خاصة ببعض **المشاكل على** سبيل **المثل،** تقدير حجم البطالة وتقدير حجم التضخم ومعدلات المواليد والوفيات وأسباب الضعف في العملية التعليمية والاقتصادية والعلمية والهندسية والصناعية والزراعية.

الإحصاء لا غنى عنه **لأي باحث في** شتى المجالات المختلفة إذ اعتمد **في** بحثه على الأسلوب **العلمي. أي أن** الإحصاء هو عصا الباحث **التي** تقوده إلى الطريق الصحيح، وهي **الأداة التي** تساعده على تفسير الظواهر **التي** يدرسها وتوضيح النتائج **التي** يحصل عليها ودلالات البيانات والأرقام **التي** يحصل عليها.

نتمنى ان يكون هذا الكتاب إضافة متواضعة الى المكتبة العربية، كتابا منهجيا لجميع الطلبة الذين يدرسون **مادة الإحصاء.**

والله ولي التوفيق .

المؤلفين

المحتويات

الفصل الثالث

الفصل الرابع

الفصل الخامس

الفصل السادس

الفصل الأول

مفهوم علم الإحصاء

1- مفهوم علم الاحصاء

1-1. المقـــــدمة

يعد علم الاحصاء(statistics) ، اليوم، من اهم العلوم التي تتوقف عليها التنمية السياسية والاقتصادية والثقافية الخ...، وللإحصاء حصة اساسية من عمل الدول والمؤسسات والمنظمات السياسية والاقتصادية والاجتماعية، عالميا ودوليا ومحليا، وكثيرا ما يرتهن مصير مشاريع او قرارات كبرى بالنتائج التي يقدمها الاحصاء في مجال معين.

وبصورة عامة، فان افتقاد الجهد الاحصائي، في مجال من المجالات، يمنع من التأكد وتحصيل الضمان في استجابة اي مشروع للواقع، كما يحول دون تحديد مدى نجاحه او اخفاقه، ويجعل في الاقدام عليه شيئا من المخاطرة.

جاء في موسوعة لالاند حول الاحصاء من الناحية الجوهرية انه مجموعة الوقائع التي يودي اليها اجتماع البشر في مجتمعات سياسية... لكن الكلمة عندنا سترتدي مفهوما اوسع، فنحن نعني بالإحصاء العلم الذي يكون موضوعه جمع وتنسيق وقائع كثيرة في كل صنف، بحيث يمكن الحصول على نسب عددية مستقلة استقلالا ملموسا عن المصادفة واستثناءاتها، وفي موسوعة المورد العربية يعرف بانه علم جمع وتصنيف وتعليل الوقائع او المعطيات الرقمية او العددية، يتخذ طريقة للتحليل في العلوم الدقيقة والعلوم الاجتماعية وفي المشروعات الاقتصادية على اختلافها. وهو يعنى، في آن معا، بوصف الوقائع وبالتنبؤ باحتمالات حدوث امر بعينه او حالة بعينها. وعلم الاحصاء علم حديث نشا في مطالع القرن العشرين، وتطور تطورا كبيرا بعد الحرب العالمية الثانية، وانما يعزى هذا التطور الكبير الى استحداث الحاسبات الالكترونية التي تتعامل مع كميات من الارقام ضخمة تعاملا سريعا. ومن تعريف علم الاحصاء انه منهج يتعاطى بالدرجة الاولى مع ظواهر رقمية وعددية معينة، ثم يقوم بتصنيفها وتحويلها الى نسب عددية خاصة، فيستطيع بالتالي تقديم وصف ميداني مرقم واكثر دقة للواقع، ويرفق ذلك الوصف بتقديم تصور علمي للعلل والاسباب التي ولدت الظاهرة المدروسة.

1-2. الهدف من دراسة مادة الاحصاء :

تهدف دراسة مادة الاحصاء الى:

1) تعريف الطالب بطبيعة علم الاحصاء وأهميته ومجالات استخدامه.
2) تعريف الطالب بأساليب جمع البيانات وطرق اخذ العينات ومن ثم طرق عرضها بالطرق البيانية.
3) تعريف الطالب باتباع الاساليب الاحصائية لغرض الوصول الى النتائج الدقيقة بأقصر طريقة واقل كلفة.

4) للإحصاء دور مهم واساسي في التخطيط للمشاريع سواء اكانت مشاريع فردية ام مشاريع تخص المجتمع استخدام الاساليب الاحصائية تساعد الطالب مستقبلا على اتخاذ قراراته سواء اكانت قرارات تجارية خاصة بالشراء وبيع المنتجات او تحديد اجور عمال .ام كانت قرارات صناعية او زراعية وغيرها .

5) توضيح للطالب كيف ان الاحصاء تغير من صورته القديمة الموجودة في اذهان الناس على انه علم العد وجمع البيانات وعرضها الى اعتباره الان علم يعمل على استخدام الاسلوب العلمي في طرق جمع البيانات وتبويبها وتحليلها وعرضها لاستخلاص النتائج واتخاذ القرار المناسب على ضوء ذلك .

3-1. طبيعة علم الاحصاء. Nature Of Statistics

كلمة (الاحصاء) في الماضي كانت تهدف الى العد والحصر حتى سمي الاحصاء بعلم العد (The Science of canting) . اما الاحصاء الأن فقد تطور كثيرا وخاصة في القرن العشرين واصبح علما مستقلا له اهميته كوسيلة واداة في البحث العلمي لجميع العلوم .

1-3-1. تعريف علم الحصاء :ـ

هناك تعاريف عديدة للإحصاء اختلفت وتباينت من حيث المضمون والشمول باختلاف مراحل تطوير هذا العلم والفوائد المتوخاة منه .

فقد عرف بأنه الطريقة العلمية التي تختص بجمع البيانات والحقائق بشكل يسهل عملية تحليلها وتفسيرها ومن ثم لستخلاص النتائج واتخاذ القرار على ضوء ذلك ، وهناك من عرفه بانه العلم الذي يعمل على استخدام الاسلوب العلمي في طرق جمع البيانات وتبويبها وتلخيصها وعرضها وتحليلها بهدف الوصول منها الى لستنتاجات وقرارات منلسبة.

ويمكن تقسيم علم الاحصاء الى قسمين هما :

1 ـ الاحصاء الوصفي Descriptive statistics : - ويتضمن الطرق الاحصائية المستخدمة في جمع البيانات والمعلومات عن ظاهرة معينة او مجموعة ظواهر وكيفية تنظيم وتصنيف وتبويب هذه البيانات مع امكانية عرضها في جداول ورسوم بيانية وحساب بعض المؤشرات الاحصائية.

2 ـ الاحصاء الاستنتاجي او الاستدلالي Statistical inference :- هو الشطر الاخر من علم الاحصاء الذي يهتم بموضوع التقديرات واختيار الفرضيات.

4-1 . اهمية علم الاحصاء وعلاقته بالعلوم الاخرى ومجالات تطبيقه

يعتبر علم الاحصاء احد الوسائل المهمة في البحث العلمي من خلال استخدام قواعده وقوانينه وطرقه في جمع المعلومات والبيانات اللازمة للبحث العلمي وتحليل هذه البيانات والمعلومات بهدف الوصول الى النتائج التي يهدف اليها البحث العلمي . وللإحصاء دورا بارزا في التخطيط ووضع الخطط المستقبلية عن طريق التنبؤ بالنتائج ففي قصة سيدنا يوسف عليه السلام مثال عظيم لدور الاحصاء في التخطيط فيبين ان هناك سنوات عجاف يقل فيها المحصول وسنوات سمان يزيد فيها المحصول ويبين انه يجب الاحتفاظ لسني القحط بادخار جزء من انتاج سني الرخاء . وفي مجال الاقتصاد والاجتماع يبرز دور الاحصاء في بحوث السكان متمثلا في تعدادات السكان فالتخطيط السليم لتنمية اقتصادية واجتماعية لا ينفصل ولا يمكن أن يتم بدون الدراسات الاحصائية للسكان ، فكيف نقرر اقامة مصانع ونحن لانعرف حجم قوة العمل المتوافرة والتي ستتوفر خلال فترة مقبلة وعلى اي اساس نقيم سياسة للإسكان ونحن لانعرف حجم قوة العمل المتوافرة والتي ستتوفر خلال فترة مقبلة وعلى اي اساس نقيم سياسة للسكان ونحن لانعرف معدلات الزواج والطلاق ..... وهكذا وفي مجال **الزراعة** يأتي دور الاحصاء في ان العلوم الزراعية تبدأ بالملاحظة وجمع بيانات عن الطبيعة في الحقل او المزرعة ثم يلي ذلك الدراسات العملية ويفيد الاحصاء في تنظيم وترتيب عملية الملاحظة والمشاهدة وجمع البيانات وتحليلها واستخلاص النتائج ولا يمكن أن يكون ذلك بغير دراسة كاملة بأساليب الاحصاء. وللإحصاء ايضا اهمية في **مجال الصناعة** من خلال استخدام النظرية الاحصائية في الانتاج الحربي وفي مجالات صناعات الفحم والحديد والغزل والمواد الكهربائية كما ان للإحصاء دور فعال في **مجال الطب والصحة العامة** في معرفة عدد المواليد وعدد الوفيات حيث تعتبر مؤشرات للمستوى الصحي العام ومؤشر لمدى تقدم البلد او تخلفه وأصبح للإحصاء أهمية كبرى في دراسة وتحليل العلاقات بين الامراض المختلفة وطرق العلاج واستخدام نظريات العروض الاحصائية اصبح الاساس في عمل شركات انتاج العقاقير والادوية . وعليه فأن الاحصاء بحد ذاته وسيلة وليس غاية فذاك يعني امكانية استخدامه اينما وجد في البحث العلمي.

5-1. الطريقة الاحصائية في البحث العلمي .

استخدام الاسلوب الاحصائي في البحث العلمي يعني توفير البيانات والمعلومات عن الظاهرة المطلوب دراستها في ذلك البحث وهذا يعني ان امكانية تطبيق الطريقة الاحصائية مرهونا بإمكانية التعبير عن هذه الظاهرة أو تلك تعبيرا كميا (رقميا) .وتمتاز الطريقة الاحصائية بكونها تهيئ اسلوبا موضوعيا محايدا للبحث له قواعده واصوله التي يجب أن يلتزم بها الباحث حتى يتجنب التحيز الشخصي والوقوع في بعض الاخطاء . كما يساعد استخدام الطريقة الإحصائية الى وصول الباحث الى النتائج الدقيقة بأقصر طريق وأقل كلفة .

1-6. المراحل الرئيسية للطريقة الاحصائية

1- تحديد مشكلة البحث

2 ـ جمع البيانات والمعلومات

3 ـ تصنيف البيانات وتبويبها

4 ـ عرض البيانات

5 ـ حساب المؤشرات أو المعالم للبيانات

6 ـ التفسير والتنبؤ

1-7. تصنيف وتبويب البيانات ، تكوين الجداول التكرارية البسيطة والمزدوجة

1-7-1. طبيعة البيانات الاحصائية

عند جمع البيانات حول ظاهرة ما نرمز للظاهرة بالرمز **y او x** او اي رمز اخر وكل مفردة او مشاهدة ترمز لها y_i او x_i فمثلا عند دراسة طول الطلبة لأحدى الجامعات فأننا نرمز لصفة الطول (الظاهرة) y وطول اي طالب (المفردة) بالرمز y_i . هذا وان قيمة y_i قد تختلف من طالب الى اخر لهذا نقول ان y متغير variables وعليه فان المتغير : هو اي ظاهرة تظهر اختلافات بين مفرداتها ويرمز له بالرمز y **او اي رمز اخر . وتقسم** المتغيرات الى قسمين :-

1 ـ **متغيرات وصفية او نوعية** :- هي الظواهر او الصفات التي لا يمكن قياسها مباشرة **بالأرقام** العددية مثل صفة لون العين (ازرق ،اسود ،بني) او الحالة الاجتماعية (غني ،متوسط الحال ، فقير) او الجنس (ذكر ،انثى)......الخ .

2 ـ **متغيرات كمية** : ـ هي الظواهر او الصفات التي يمكن قياسها مباشرة **بأرقام** عددية مثل صفة الطول ، الوزن ، العمر ، كمية المحصول .

وتنقسم المتغيرات الكمية الى قسمين هما : ـ

أ ـ متغيرات متصلة (مستمرة) : المتغير المستمر هو المتغير الذي **تأخذ** المشاهدة او المفردة فيه اية قيمة رقمية في مدى معين ، فلو فرضنا ان اطوال طلبة جامعة ما تتراوح بين 130,5**سم و**170 سم اي ان المتغير y يمكن

ان ياخذ اي قيمة بين 130,5 و170سم ومثال ذلك الوزن وكمية المحصول ودرجات الحرارة ، وبصورة عامة كل البيانات التي تقاس تعتبر بيانات لمتغير مستمر (تأخذ القيم عدد صحيح او كسر) .

ب ـ **متغيرات غير مستمرة (منفصلة)** : ـ تأخذ المشاهدة او المفردة فيها قيما متباعدة او منقطعة مثال ذلك عدد الثمار او عدد الوحدات الانتاجية او عدد الطلبة في الصفوف الاولى لجامعة ما ، فهي في الغالب تكون اعداد صحيحة . وبصورة عامة كل البيانات التي نحصل عليها من العد تعتبر بيانات لمتغير منفصل .

1-8. المجتمع والعينة :

المجتمع : - عبارة عن جميع القيم او المفردات التي يمكن ان يأخذها المتغير ، فلو كانت دراستنا حول لطول طلبة جامعة ما فان المجتمع في هذه الحالة هو اطوال جميع الطلبة في تلك الجامعة .والمجتمع اما ان يكون مجتمعا محدودا اي ممكن حصر عدد مفرداته . او يكون مجتمعا غير محدودا وهو المجتمع الذي من الصعب حصر عدد مفرداته مثل مجتمع نوع سمك معين في نهر دجلة او عدد البكتريا في حقل ما .

اما العينة : فهي جزء من المجتمع . فالعينة عبارة عن مجموعة من المشاهدات اختيرت بطريقة ما من المجتمع ففي بعض الاحيان دراسة المجتمع ككل قد يكون صعبا او يحتاج الى وقت وجهد ومال فيستعاض عن دراسة المجتمع بدراسة العينة وصفاتها ومنها يستنتج خواص المجتمع الاصلي الذي اخذت منه العينة .

1-9. اسلوب تصميم البحوث

هناك اعتبارات كثيرة يتوقف عليها تصميم البحث وعلى الباحث ان يأخذ بنظر الاعتبار مسالة الحصول على البيانات والمعلومات بأقصر وقت واقل جهد واوطئ كلفة وعليه يجب مراعاة مايلي عند تصميم البحث

1 ـ **تحديد الغرض من البحث** : من الضروري ان يكون الهدف محددا بشكل واضح ودقيق معروفة اهدافه واوجه الاستفادة من نتائجه.

2 ـ **امكانية التنفيذ العملي للبحث** : - من الضروري تحديد المتطلبات التي تلتزمها عملية تنفيذ البحث كالموارد المالية المطلوبة والامكانيات البشرية المتاحة في تحقيق بعض فقرات البحث وكذلك التأكد من مدى توفر البيانات والمعلومات الدقيقة عن مشكلة البحث.

3 ـ **تحديد اطار البحث** : - من المهم ان يحدد الباحث نوع وطبيعة مجال البحث او المجتمع الاحصائي والمجتمع الاحصائي عبارة عن مجموعة وحدات او مفردات ذات صفة او صفات مشتركة فمثلا اذا كان البحث يتعلق بأطوال طلبة جامعة بغداد فان المجتمع الاحصائي هو جميع الطلبة في جامعة بغداد والمفردة الطالب او الطالبة في هذا المجتمع. واذا كان البحث حول دخل العائلة الفلاحية في العراق فالمجتمع الاحصائي هو العوائل الفلاحية

الساكنة في العراق والوحدة الاحصائية او المفردة هي العائلة الواحدة والمجتمع يكون اما **مجتمع محدد** وهو **المجتمع الذي** يمكن الوصول الى كل مفردة فيه مثل مجتمع جامعة بغداد او يكون مجتمع غير محدد مثل كريات الدم البيضاء في دم الانسان ومجتمع الاسماك في نهر دجلة.

1-10. اسلوب جمع البيانات والمعلومات : ـ للوصول الى البيانات والمعلومات هناك اسلوبان يمكن من خلالهما جمع هذه البيانات والمعلومات كل منهما له ميزاته وعيوبه وهذان الاسلوب هما : ـ

1 ـ **اسلوب التسجيل الشامل** : ـ هو جمع البيانات من جميع المفردات التي يتكون منها المجتمع مجال البحث **ومثل ذلك التعداد العام للسكان من ميزات** هذا الاسلوب يعطي بيانات كاملة حول الظواهر التي يتم البحث عنها **اما عيوبه** فان هذا الاسلوب يحتاج الى وقت وجهد و**مال كما** لا يمكن استخدام هذا الاسلوب في المجتمعات غير **المحددة**.

2 ـ **اسلوب العينات** : ـ هو اخذ وحدات من المجتمع الاحصائي تسمى العينة sample **والغرض من اخذ** العينة ان تكون بديلا عن المجتمع الاحصائي وعن طريق صفاتها يتمكن **الباحث** ان يصف خواص المجتمع بتعميم النتائج **التي حصل** عليها من دراسة العينة . تفضل هذه الطريقة عن طريقة التسجيل الشامل **للأسباب** الاتية :

1 ـ توفر المال والجهد والوقت اللازم **لإجراء البحث**.

2 ـ صعوبة اجراء التسجيل الشامل بسبب طبيعة المجتمع فقد يكون المجتمع غير محدد او كبير **جدا. ومن عيوب** هذا الاسلوب فان محاولة التعرف على خواص المجتمع عن طريق دراسة جزء منه ينطوي عليه التضحية في دقة النتائج التي تستخرجها .

وتنقسم العينات الى قسمين رئيسيين هما :

1-11. العينات العشوائية والعينات الغير عشوائية

1-11-1. ـ **العينات العشوائية** : هي مجموعة المفردات المختارة من مجتمع الدراسة وليس للباحث دخل في اختيارها . وللعينات العشوائية انواع عديدة منها

أ ـ العينة العشوائية البسيطة Simple random sample

هي اختيار عينة عشوائية من مجتمع الدراسة بطريقة تعطي المفردات نفس الفرصة في الظهور .ويشترط **هنا ان** يكون المجتمع متجانس(مشترك في الصفات)فمثلا دراسة اسباب التدخين لدى الاناث نلاحظ ان **المجتمع** متجانس حيث ان كافة مفردات **هذا المجتمع** هم من الاناث والصفة المشتركة هي التدخين.

مثال (1)

أراد مدير شركة اختيار لجنة مكونة من 5 أشخاص من بين مجموعة موظفين عددهم 50 موظف، كيف يتم الاختيار بطريقة العينة العشوائية البسيطة؟

الحل:

يقوم المدير بكتابة أسماء جميع الموظفين على بطاقات كل اسم على بطاقة ثم يعمل على خلط البطاقات ووضعها في صندوق ثم يسحب 5 بطاقات بحيث يسحب بطاقة في كل مرة.

ب- العينة المنتظمة (الأسلوبية) Systematic sample

وهي العينة التي يتم اختيارها من مجتمع يكون موزعا على أساس معين، كأن يكون تصاعدياً أو تنازلياً.

ومن أمثلتها اختيار عدد من الصكوك المدفوعة من دفتر الصكوك المتسلسل أو اختيار عدد من المنازل المرقمة في محافظة ما.

وتتم عملية الاختيار بتحديد الزيادة المنتظمة (k) ثم تحديد مفردة البداية التي تكون عادة أقل من (k)، ومن ثم يتم إضافة هذه الزيادة المنتظمة بشكل متسلسل.

وتعطى الزيادة المنتظمة (k) بالقانون

$$\mathbf{k = N/n}$$

حيث **N** : حجم المجتمع   **n** : حجم العينة

مثال 2 :

إذا كان حجم مجتمع ما هو 1000 مفردة ويراد اختيار عينة حجمها 100 مفردة ، أوجد الزيادة المنتظمة k.

الحل:

$$\mathbf{k = N/n}$$

$$= 1000 / 100 = 10$$

مثل 3 :

يراد اختيار عينة حجمها 100 من مجتمع عدد أفراده 1260 أوجد مقدار الزيادة المنتظمة.

الحل:

$$k = N/n$$

$$= 1260 / 100 \quad = 12.6$$

ويتم التقريب للأسفل فيكون الجواب هو 13.

مثل4:

يراد اختيار عينة حجمها 200 مفردة من مجتمع حجمه 4000 مفردة ، كيف يتم ذلك بطريقة العينة المنتظمة؟

الحل:

1) نجد مقدار الزيادة المنتظمة k :

$$k = N/n$$

$$= 4000 / 200 = 20$$

2) نختار مفردة البداية وتكون أقل من 20 ولتكن 8.

3) نضيف k بشكل متسلسل

8 , 8 +20 , 8 + 20 + 20 , 8 + 20 + 20 + 20 + ...

8 , 28 , 48 , 68 , ... , 388.

مثل 5:

دائرة المحاسبة في شركة تود اجراء دراسة لعملاء الشركة من حيث طرق السداد وحجم الطلبات وكيفية الشحن، إذا تم اختيار العينات بطريقة العينة المنتظمة وكان عدد العملاء المسجلين 12500؛ وتم اختيار حجم العينة ليكون 350 أوجد العينة المطلوبة.

الحل:

$$k = N/n$$

$$= 12500 / 350$$

$$= 35.7$$

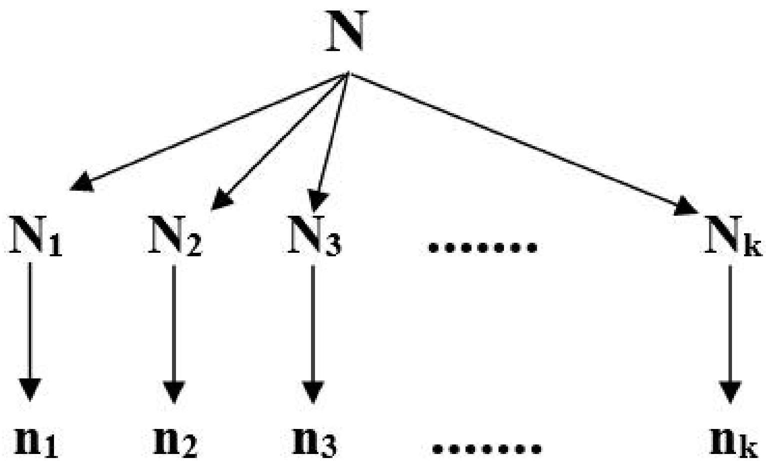
$$\approx 35$$

فتكون عينة العملاء المختارة هي كالآتي؛ إذا اعتبرنا أن مفردة البداية هي 18 :

18 , 53 , 88 , 123 , ...

ج ـ العينة الطبقية العشوائية : ـ يتم اختيار العينة عندما يكون المجتمع غير متجانس ،يقسم المجتمع الى طبقات كل طبقة تعتبر مجتمع متجانس ومن كل مجتمع يتم اختيار عينة عشوائية بسيطة يتناسب حجمها مع حجم الطبقة ثم تجمع هذه العينات ونحصل على الطبقة العشوائية.

مثلا لو كنا بصدد دراسة للمستوى العلمي لإحدى طلبة المعهد التقني نينوى هذا المجتمع غير متجانس من حيث التخصص العلمي فهناك اختصاص ادارة قانونية واختصاص محاسبة واختصاص تقنيات مالية ومصرفية واختصاص سياحة وهكذا.



حيث أن N : حجم المجتمع الطبقي

N_i : حجم الطبقة (i).

n_i : حجم العينة من الطبقة (i).

n : حجم العينة المطلوبة.

أي أن

$$N = N_1 + N_2 + N_3 + \dots + N_k$$

$$n = n_1 + n_2 + n_3 + \dots + n_k$$

ونجد حجم العينة من الطبقة (i) بالصيغة

$$n_i = \frac{N_i}{N} \times n$$

مثل 6 :

مجتمع حجمه 10000 مفردة ومكون من 4طبقات حجم كل طبقة على التوالي :

1000 , 3500 , 4000 , 1500 . يراد سحب عينة حجمها 400 مفردة من هذا المجتمع، كيف يتم ذلك بحيث تمثل هذه العينة تمثيلاً سليماً.

الحل:

$N = 10000 \qquad n = 400$

$N_1 = 1000 \quad N_2 = 3500 \quad N_3 = 4000 \quad N_4 = 1500$

$$n_1 = \frac{N_1}{N} \times n = \frac{1000}{10000} \times 400 = 40$$

$$n_2 = \frac{N_2}{N} \times n = \frac{3500}{10000} \times 400 = 140$$

$$n_3 = \frac{N_3}{N} \times n = \frac{4000}{10000} \times 400 = 160$$

$$n_4 = \frac{N_4}{N} \times n = \frac{1500}{10000} \times 400 = 60$$

نلاحظ أن $n = 40 + 140 + 160 + 60$

$= 400$

وهو حجم العينة المطلوبة.

مثل 7 :

أرادت مجلة متخصصة بحث مقدار الأرباح التي تحققها المنشآت التجارية كلٌ حسب نشاطها التجاري. فعملت بحث الدرلسة على 1000 منشأة. فإذا تم اختيار عينة مكونة من 100 منشأة. كيف يتم الاختيار بطريقة سليمة؟

الجدول التالي يبين عدد ونوع كل منشأة في مجتمع الدراسة.

Group	**No. Of companies**
Financial	300
Diversified	200
Commercial banking	200
Retailing	100
Transportation	100
Utilities	100
Total	**1000**

الحل:

$n = 100 \qquad N = 10000$

$N_6 = 100 \quad N_5 = 100 \quad N_4 = 100 \quad N_3 = 200 \quad N_2 = 200 \quad N_1 = 300$

$$n_1 = \frac{N_1}{N} \times n = \frac{300}{1000} \times 100 = 30$$

$$n_2 = \frac{N_2}{N} \times n = \frac{200}{1000} \times 100 = 20$$

$$n_3 = \frac{N_3}{N} \times n = \frac{200}{1000} \times 100 = 20$$

$$n_4 = \frac{N_4}{N} \times n = \frac{100}{1000} \times 100 = 10$$

$$n_5 = \frac{N_5}{N} \times n = \frac{100}{1000} \times 100 = 10$$

$$n_6 = \frac{N_6}{N} \times n = \frac{100}{1000} \times 100 = 10$$

نلاحظ أن

$$n = 30 + 20 + 20 + 10 + 10 + 10 = 100$$

د ـ العينة المتعددة المراحل : يتم تقسيم المجتمع الى وحدات اولية ثم يتم اختيار عينة عشوائية من هذه الوحدة الاولية ثم تقسم كل وحدة من الوحدات الاولية الى وحدات ثانوية ثم تؤخذ عينة كمرحلة ثانية ثم تقسم الى وحدات اصغر ونأخذ عينة منها الى ان نصل الى المفردة التي يتم جمع البيانات منها والتي تؤلف عينة البحث

مثال 8 :

يراد قياس المستوى التحصيلي لطلاب جامعة عمان الأهلية في سنة ما. كيف تتم عملية الاختيار بطريقة سليمة؟

الحل:

نظر على مجتمع الطلاب بالجامعة على أنه موزع على عدة كليات؛ ثم أن هذه الكليات مقسمة لعدة أقسام وهذه الأقسام بدورها تحتوي على عدة شعب دراسية. وبأخذ هذه المراحل بعين الاعتبار يتم اختيار العينة المطلوبة.

الجامعة ← كليات ← أقسام ← شعب دراسية

1-11-2. العينات غير العشوائية : ـ يقصد بها مجموعة من المفردات المختارة من مجتمع الدراسة بطريقة يكون للباحث دخل في اختيارها ومن هذه العينات .

أ ـالمعاينة الحصصية : تقسيم مجتمع الدراسة الى طبقات استنادا الى معايير تقسيم معينة تتعلق بطبيعة الدراسة ثم يتم اختيار عدد من المفردات من كل طبقة بشكل شخصي (غير عشوائي) بحيث ان عدد مفردات هذه العينات يشكل حجم العينة المطلوبة لتلك الدراسة .فلو كنا بصدد استطلاع راي الجمهور ببرامج التلفزيون فانه يمكن تقسيم مجتمع الدراسة الى ذكور واناث ثم يتم اختيار عينة من الذكور واخرى من الاناث تتناسب كل منهما مع عدد الذكور وعدد الاناث في مجتمع هذا الاستطلاع ومجموع مفردات هاتين العينتين تؤلفان حجم العينة المطلوب للاستطلاع .

ب ـ المعاينة العمدية : اختيار العينة بشكل متعمد يعتقد الباحث مسبقا بان مفردات هذه العينة هي خير من يمثل مجتمع الدراسة .

1-12. مصادر جمع البيانات :

يتم الحصول على البيانات و المعلومات من احد المصدرين الآتيين :

1 ـ **المصادر التاريخية** : هي البيانات المحفوظة لدى اجهزة الدولة المختلفة نتيجة الاستقصاءات او مسوحات قامت بها هذه الجهات لأغراض خاصة بها او تجمعت لديها بحكم وظائفها . مثال ذلك البيانات المتجمعة عن تعدادات السكان ،احصاءات الطلبة المتخرجين من الجامعات او احصاءات التجارة الداخلية والخارجية .

2- **مصادر الميدان** : بيانات ومعلومات يمكن الحصول عليها من مصادرها الاصلية بطريقة المراسلات (بالبريد) أو المواجهة (المقابلة الشخصية) أو عن طريق الهاتف أو اي وسيلة اتصال أخرى.

13-1. تصنيف وتبويب البيانات

لاحظنا ان عملية جمع البيانات تتم من خلال المصادر التاريخية او الميدانية باستخدام اسلوب التسجيل الشامل او اسلوب العينات حسب ما تتطلبه الدراسة ، ان البيانات المستحصل عليها بخصوص الظاهرة المعنية تسمى البيانات الاولية او البيانات غير المصنفة ،ان البيانات بشكلها الاولي تكون غير منظمة مما يتعذر على الباحث تكوين فكرة عن هذه الظاهرة او تلك التي جمعت البيانات ، كذلك يتعذر الاعتماد عليها بشكلها الغير المنظم لأغراض التحليل الاحصائي للوصول الى النتائج المطلوبة ، لذلك ان اولى الخطوات الهامة بعد عملية جمع البيانات هي عملية تصنيف وتبويب البيانات .

1- **مراجعة البيانات** : بعد اتمام عملية جمع البيانات وفق الوسيلة المناسبة لذلك البحث يتوجب الامر مراجعة وتدقيق البيانات لغرض التأكد من مطابقتها وتكاملها لمتطلبات الدراسة.
2- **تصنيف البيانات: بعد التأكد** من دقة البيانات التي تم الحصول عليها يتم عملية تصنيف البيانات على اساس الظواهر التي جمعت منها البيانات حيث يتم فرز بيانات كل ظاهرة على هيئة مجموعة فقد يكون التصنيف على ظاهرة العمر، الوزن، المهنة، الطول، الجنس .
3- **تبويب البيانات** : بعد اتمام عملية تصنيف البيانات تبدأ عملية التبويب , ويقصد بالتبويب عملية تفريغ البيانات المصنفة في جداول خاصة بحيث ان كل جزء من البيانات المصنفة عن الظاهرة المعنية يعود الى مستوى معين لتلك الظاهرة ، الهدف من عملية التبويب هو ابراز البيانات وتوضيحها في أضيق حيز ممكن كي يتمكن الباحث من تكوين فكرة عنها ويختلف اسلوب تبويب البيانات تبعا لطبيعتها . وفيما يلي عرض موجز لكل شكل من هذه الاشكال
   أ- **التبويب الزمني** : عبارة عن تجميع البيانات المصنفة وترتيبها في جداول على اساس ان كل مجموعة منها تعود لوحدة زمنية كاليوم ، الاسبوع ، الشهر ، السنة . والجدول التالي يوضح عدد الطلبة الخرجين لعدد من السنوات

السنوات	عدد الخريجين
2000	150
2001	180
2002	200
2003	250
مجموع	780

ب ـ التبويب الجغرافي : ـ تقسيم البيانات الى مجموعات كل منها خاص بوحدة جغرافية معينة او تقسيم اداري معين كالنواحي والاقضية والمحافظات والبلدان، القارات ، عدد الطلبة الخريجين حسب الجامعات العراقية .

اسم الجامعة	العدد
جامعة بغداد	2500
جامعة الموصل	2000
جامعة البصرة	2200
جامعة تكريت	1800
المجموع	8500

ج ـ التبويب الكمي : تقسيم البيانات الى مجموعات خاصة بوحدة معينة كوحدات الوزن والطول ، المساحة ،الحجم .....الخ

الجدول التالي يوضح توزيع الاجور اليومية لعمال احد المصانع

الاجرة اليومية بالدينار	عدد العمال
اقل من 3000دينار	185
اقل من3500	95
اقل من5000	70
من 5000 فاكثر	20
مجموع	370

دـ التبويب على اساس صفة معينة : ـ تجميع البيانات وترتيبها في جداول على مجموعة منها يشترك بصفة معينة كالجنس ، الحالة الاجتماعية ، عنوان الوظيفة والجدول التالي يوضح عدد الطلبة حسب الجنس.

الجنس	العدد
ذكور	123
اناث	77
مجموع	200

14-1. التوزيع التكراري Frequency Distribution

عبارة عن تلخيص وترتيب البيانات التي سبق ان جمعت وصنفت مقسمة الى عدد من المجاميع كل منها تسمى **الفئة** (class) هذه الفئات قد تكون مرتبة تصاعديا او تنازليا حسب طبيعة البيانات ويسمى توزيع عدد قيم x **حسب** الفئات بالتوزيع التكراري . وقد تكون فئات التوزيع التكراري متساوية في الطول ام غير متساوية وذلك يعتمد **على** طبيعة الدراسة ومتطلباتها. وفيما يلي توضيح لبعض المصطلحات.

البيانات غير المبوبة : هي البيانات الاولية التي جمعت ولم تبوب في جدول توزيع تكراري .

البيانات المبوبة : هي البيانات التي جمعت وبوبت ونظمت في جدول توزيع تكراري .

التوزيع التكراري : تقسيم البيانات او القيم الخاصة بظاهرة من الظواهر الاحصائية الى اصناف او فئات يطلق عليها بالتوزيع التكراري.

الفئة : هي المجاميع التي قسمت اليها قيم المتغير ، وكل فئة لها حدان ، حد ادنى ،وحد اعلى .

المدى :هو الفرق بين اكبر قيمة واصغر قيمة +1

مركز الفئة : هي القيمة الواقعة **عند منتصف الفئة** .

تكرار الفئة : عدد المفردات او القيم التي تقع في مدى تلك الفئة ويرمز له بـ f_i هذا وان مجموع **التكرارات** يجب ان يكون دائما مساوي للعدد الكلي لقيم الظاهرة .

طول الفئة : هو مقدار المدى بين حدي الفئة .

وطول الفئة يرمز له بـ L ويستخرج طول الفئة بـاستخدام احد القوانين الاتية :

1- يمكن إيجاد طول الفئة من العلاقة التالية

$$L = xL - xs$$

حيث **أن** :

طول الفئة : **L**

الحد الاعلى للفئة : **xL**

الحد الادنى للفئة : xs

2- كذلك يمكن إيجاد طول الفئة من العلاقة التالية

$$L = \frac{T.R}{m}$$

المدى = R.T ، **عدد الفئات** = m

3- **طول الفئة** = الفرق بين الحدين الادنى (او الحدين الاعلى) لفئتين متتاليتين

4- طول الفئة : الفرق بين مركزي فئتين متتاليتين

المدى : يرمز له T.R

$$T.R = XL - XS + 1$$

حيث ان:

LX = القيمة الأكبر في العينة

SX = القيمة الأصغر في العينة

مركز الفئة : يرمز له X

$$X = \frac{L.L + U.L}{2}$$

حيث ان :

الحد الاعلى للفئة L.L

الحد الادنى للفئة U.L

عدد الفئات: يرمز له بـ m هناك عدة طرق تقريبية لإيجاد عدد الفئات اهمها :

$$m = 1 + 3.322 \log n$$

او

$$m = 2.5 \sqrt[4]{n}$$

حيث ان : n = حجم العين

1-15. الجداول الاحصائية

هناك نوعان رئيسيان من الجداول الاحصائية

1- **الجدول البسيط** : هو الجدول الذي توزع فيه البيانات حسب صفة واحدة ويتألف عادة من عمودين . الاول يمثل تقسيمات صفة الظاهرة الى فئات او مجموعات والثاني يبين عدد المفردات **التابعة كل فئة او مجموعة**

مثل : الجدول التالي يمثل عدد من الطلبة حسب اوزانهم

فئات الوزن	عدد الطلبة
60-62	5
63-65	15
66-68	45
69-71	27
72-74	8
المجموع	100

2- **الجدول المركب (المزدوج)** : هو الجدول الذي توزع فيه البيانات حسب صفتين او ظاهرتين او **اكثر في نفس الوقت** ويتألف من :

الصفوف : تمثل فئات او مجاميع احدى الصفتين .

الاعمدة : تمثل فئات او مجاميع الصفة الاخرى.

مثل: الجدول التالي يبين توزيع عدد من الطلبة حسب صفتي الطول والوزن

الوزن / الطول	51-60	61-70	71-80	المجموع
121-140	20	6	4	30
141-160	2	40	10	52
161-180	2	6	10	18
المجموع	24	52	24	100

مثل (1)

لو اردنا عمل توزيع تكراري **للأعداد** الاتية التي تمثل الوزن بالكيلو غرامات لعشرين**طالبا في** المعهد **التقني** نينوى (67,55,65.70,75,60,89,83,65,56,49,65,49,48,69,62,72,45,56,74,)

نتبع الخطوات التالية

1- **نبحث** عن اكبر قيمة واصغر قيمة في المجموعة وهما (89-45) **وذلك للتوصل الى المدى الكلي**
2- **نجد عدد الفئات**
3- **نجدطول الفئة**
4- **نكتب حدود الفئات ونستخرج عدد التكرارات لكل فئة**

1- المدى

المدى الكلي = اكبر قيمة – اقل قيمة +1

$$T.R = XL - XS + 1 = 89 - 45 + 1 = 45$$

2- عدد الفئات

$$m = 1 + 3.322 \log n = 1 + 3.322 \log 20 = 5.32 \approx 5$$

او نستخدم القانون التالي:

$$5\ m = 2.5\ \sqrt[4]{n} = 2.5 \times \sqrt[4]{20} = 2.5 \times 2.114 = 5.28 \approx 5$$

3-طول الفئة

$$L = \frac{T.R}{m} = \frac{45}{9} = 9$$

كتابة حدود الفئات واستخراج عدد التكرارات لكل فئة

class	**f_i**	**f_i**
40-49	| | | |	4
50-59	| | |	3
60-69	| | | | | | |	7
70-79	| | | |	4
80-89	| |	2
Total	20	20

ملاحظة : هناك عدة طرق لكتابة حدود الفئات :

1- اما ان تكون الاعداد لمتغيرات منفصلة كما في المثال السابق
2- **او** تكون الاعداد لمتغيرات متصلة وهو الذي يمثل بعدد صحيح او كسر مثل الاوزان والاطوال وتكتب **الفئات كالاتي** :

من 40 **الى اقل من** 50

من 50 **الى اقل من** 60

من 60 **الى اقل من** 70

وللاختصار تكتب بالصيغة الاتية :

-40

-50

-60

- 70

تمتاز هذه الطريقة بالوضوح وتستخدم غالبا **لإعداد** التي تمثل متغيرات متصلة

3- وقد تكتب الفئات حسب الصيغة التالية :

اكبر من 40 **واقل من** 50

اكبر من 50 **واقل من** 60

اكبر من 60 **واقل من** 70

وللاختصار تكتب

-40

-50

-60

-70

قد يكون التوزيع في الجدول التكراري البسيط توزيعا منتظما كما في المثال السابق وذلك لتساوي طول الفئة ، او يكون التوزيع غير منتظم اذا كان طول الفئة غير متساوي ،او يكون الجدول مغلقا اذا كان الحد **الادنى** والحد الاعلى للفئة معروف ، او يكون الجدول مفتوحا في الحالات الاتية :

أ- يكون مفتوحا من الطرف الادنى فقط
ب- يكون مفتوحا من الطرف الاعلى فقط
ت- يكون مفتوحا من الطرفين (اذا كان الحد الانى والحد الاعلى للفئة غير معلوم)

1-16. الجداول التكرارية المزدوجة

مثل (2) : البيانات الاتية لظاهرتين x وy المطلوب تفريغها في جدول تكراري مزدوج

X 2 , 10 , 11 , 4 , 20 , 15 , 15 , 3

Y 3 , 2 , 5 , 6 , 8 , 10 , 2 , 10

X 25 , 25 , 20 , 22 , 15 , 20 , 30 , 30 ,

Y 2 , 6 , 5 , 9 , 15 , 12 , 3 , 9

X 35 , 30 , 35 , 31

Y 10 , 12 , 11 , 4

الحل : نستخرج معلومات لكل ظاهرة على حدى

1- المتغير X

المدى :

$$T.R = XL - XS + 1 = 35 - 2 + 1 = 34$$

عدد الفئات :

$$m = 2.5\sqrt[4]{n} = 2.5 \times \sqrt[4]{20} = 2.5 \times 2.114 = 5.28 \approx 5$$

طول الفئة :

$$L = \frac{T.R}{m} = \frac{34}{5} = 6.8 \approx 7$$

2- **المتغير** Y

3- **المدى** :

$$T.R = XL - XS + 1 = 15 - 2 + 1 = 14$$

4- **عدد الفئات** :

$$m = 1 + 3.322 \log n = 1 + 3.322 \log 20 = 5.31 \approx 5$$

5-طول الفئة :

$$L = \frac{T.R}{m} = \frac{14}{5} = 2.8 \approx 3$$

<table>
<tr><th>f_X</th><th>f_Y</th><th colspan="2">f_{iX}</th><th colspan="2">f_{iY}</th></tr>
<tr><td>2-8</td><td>2-4</td><td>|||</td><td>3</td><td>||||||</td><td>6</td></tr>
<tr><td>9-15</td><td>5-7</td><td>|||||</td><td>5</td><td>||||</td><td>4</td></tr>
<tr><td>16-22</td><td>8-10</td><td>||||</td><td>4</td><td>||||||</td><td>6</td></tr>
<tr><td>23-29</td><td>11-13</td><td>||</td><td>2</td><td>|||</td><td>3</td></tr>
<tr><td>30-36</td><td>14-16</td><td>||||||</td><td>6</td><td>|</td><td>1</td></tr>
<tr><td colspan="2">Total</td><td colspan="2">20</td><td colspan="2">20</td></tr>
</table>

بعد ذلك نضع المعلومات في جدول مزدوج وكالاتي :

<table>
<tr><th>X / Y</th><th>2-8</th><th>9-15</th><th>16-22</th><th>23-29</th><th>30-36</th><th>Total</th></tr>
<tr><td>2-4</td><td>|</td><td>||</td><td></td><td>|</td><td>||</td><td>6</td></tr>
<tr><td>5-7</td><td>|</td><td>|</td><td>|</td><td>|</td><td></td><td>4</td></tr>
<tr><td>8-10</td><td>|</td><td>|</td><td>||</td><td></td><td>||</td><td>6</td></tr>
<tr><td>11-13</td><td></td><td></td><td>|</td><td></td><td>||</td><td>3</td></tr>
<tr><td>14-16</td><td></td><td>|</td><td></td><td></td><td></td><td>1</td></tr>
<tr><td>Total</td><td>3</td><td>5</td><td>4</td><td>2</td><td>6</td><td>20</td></tr>
</table>

17-1. التوزيع التكراري المتجمع

التوزيع التكراري البسيط يعطينا عن عدد المفردات في كل فئة لكن في بعض الاحيان نرغب في معرفة عدد المفردات التي قيمتها أقل او أكثر من قيمة معينة في التوزيع التكراري . ويعرف التوزيع التكراري المتجمع: بانه التوزيع الذي يبين كمية التكرار المتجمع عند قيمة معينة من قيم المتغير العشوائي. وهناك **نوعان من الجداول** التكرارية المتجمعة . ويرمز له ب Fi

أ- **التوزيع التكراري المتجمع الصاعد** : وهو عبارة عن تجميع **التكرارات** من الفئة الاولى وانتهاء بالفئة الاخيرة منه ويتم حساب التكرارات المتجمعة على اساس الحدود العليا للفئات .

مثال (3) : لو اردنا عمل جدول التوزيع التكراري المتجمع الصاعد **للأعداد** الاتية التي تمثل الوزن بالكيلو غرامات لعشرين طالبا في المعهد التقني نينوى

(67,55,65.70,75,60,89,83,65,56,49,65,49,48,69,62,72,45,56,74,)

فالحل يكون كالاتي:

نجمع التكرارات من الفئة الدنيا الى الفئة العليا

الفئات	التكرارات f_i	الحدود العليا للفئات	المتجمع الصاعد F
40-49	4	اقل من 49	4
50-59	3	اقل من 59	7
60-69	7	اقل من 69	14
70-79	4	اقل من 79	18
80-89	2	اقل من 89	20
المجموع	**20**		

ب- **جدول التوزيع المتجمع النازل** : عبارة عن تجميع **التكرارات** ابتداء من الفئات العليا وانتهاء بالفئات الدنيا ،**بعبارة اخرى تناقص التكرارات** ابتداء بالفئة الاولى في التوزيع وانتهاء بالفئة الاخيرة منه . ويتم حساب التكرارات المتجمعة النازلة على اساس الحدود الدنيا للفئات .

فلو رجعنا الى المثل رقم (3) فان التوزيع التكراري المتجمع النازل يكون كالاتي :

الفئات	التكرارات fi	الحدود الدنيا **للفئات**	المتجمع النازل F
40-49	4	40فاكثر	20
50-59	3	50 فاكثر	16
60-69	7	60 فاكثر	13
70-79	4	70 فاكثر	6
80-89	2	80 فاكثر	2
المجموع	**20**		

الأسئلة

س1/ ما هو الهدف من دراسة مادة الإحصاء ؟

س2/ عرف علم الإحصاء وما هي اقسامه ؟

س3/ ما هي أهمية علم الإحصاء وما هي علاقته بالعلوم الأخرى ؟

س4/ ما المقصود بالمتغيرات المستمرة وغير المستمرة ؟

س5/ ما المقصود بالعينة والمجتمع ؟

س6/ ما هو أسلوب تصميم البحث ؟

س7/ ما هي أساليب جمع البيانات والمعلومات وضحها بالتفصيل ؟

س8/ عدد أنواع العينات ؟

س9/ ما هي مصادر جمع البيانات ؟

س10/ عرف كلا مما يأتي :

مراجعة البيانات ، تصنيف البيانات ، تبويب البيانات ، التبويب الزمني للبيانات ، التبويب الجغرافي للبيانات ، التبويب الكمي ، التبويب على أساس صفة معينة.

س11/ ما المقصود بما يلي :

التوزيع التكراري ، البيانات غير المبوبة ، البيانات المبوبة ، طول الفئة ، المدى ، مركز الفئة ، تكرار الفئة .

س12/ ما هي أنواع الجداول الإحصائية مع توضيح كل نوع ؟

الفصل الثاني

العرض البياني للبيانات المبوبة

2. العرض البياني للبيانات المبوبة

العرض البياني للبيانات المبوبة: ان الرسوم والاشكال الهندسية ما هي الا تعبير وتوضيح للبيانات بطريقة جذابة وسهلة وفعالة تساعد القارئ على فهم واستيعاب الظاهرة ومقارنتها مع بعضها . ووسائل التمثيل البياني كثيرة ومتنوعة وسنكتفي بشرح العرض البياني للتوزيعات التكرارية فقط . وعادة نخصص المحور الافقي او الاحداثي السيني لتمثيل قيم او فئات المتغير بينما نخصص المحور العمودي او الاحداثي الصادي لتمثل تكرارات هذا المتغير ويجب دائما ان يبدأ تدريج المحور العمودي من الصفر اما تدريج المحور الافقي فقد لا نبدأ بتدريجه من الصفر. ومن اشكال العرض البياني نذكر.

2-1. المدرج التكراري

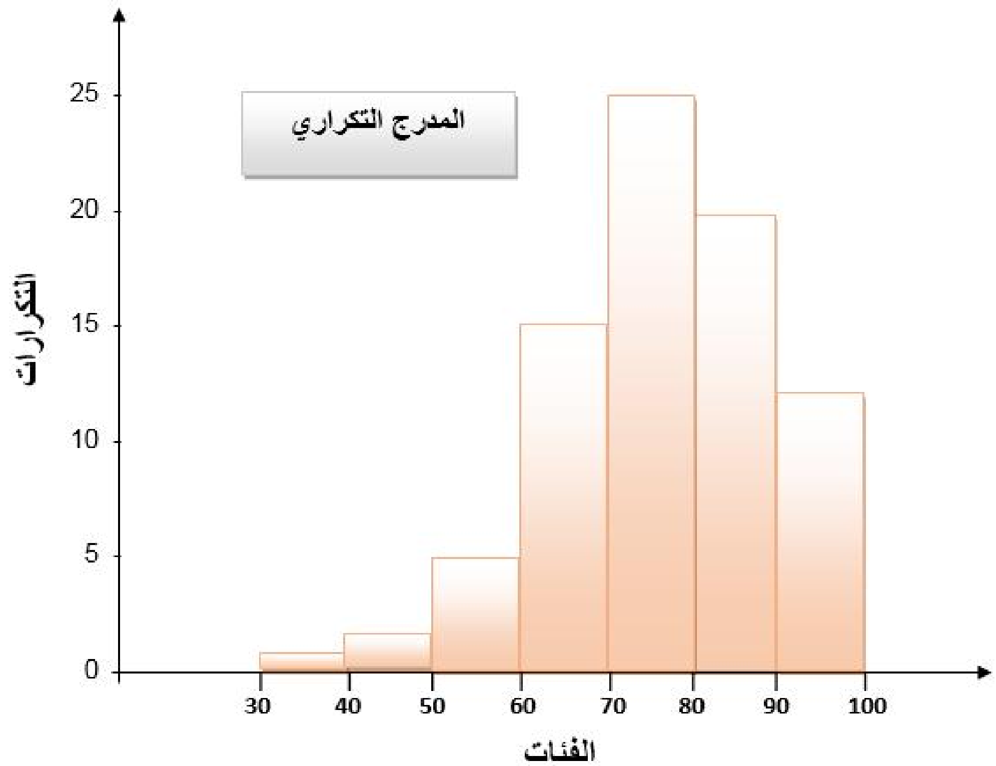
عبارة عن مستطيلات رأسية تمتد قواعدها على المحور الافقي لتمثل اطوال الفئات بينما ارتفاعاتها تمثل تكرارات الفئات. ولرسم مدرج تكراري نتبع الخطوات الاتية :

1- رسم المحور الافقي والمحور العمودي
2- تدريج المحور الافقي الى اقسام متساوية بحيث يشمل جميع حدود الفئات ويفضل ترك مسافة صغيرة بين نقطة الصفر والحد الادنى للفئة الاولى، ونقسم المحور العمودي الى اقسام متساوية بحيث تشمل اكبر التكرارات
3- يرسم على كل فئة مستطيلا رأسيا تمثل قاعدته طول تلك الفئة وارتفاعه يمثل تكرارها

مثال (1) :الجدول الاتي يمثل توزيع تكراري لأطول نباتات القطن /المطلوب : تمثيل التوزيع التكراري بمدرج تكراري

الفئات	30-40	40-50	50-60	60-70	70-80	80-90	90-100	المجموع
التكرارات	1	2	5	15	25	20	12	**80**

الحل:



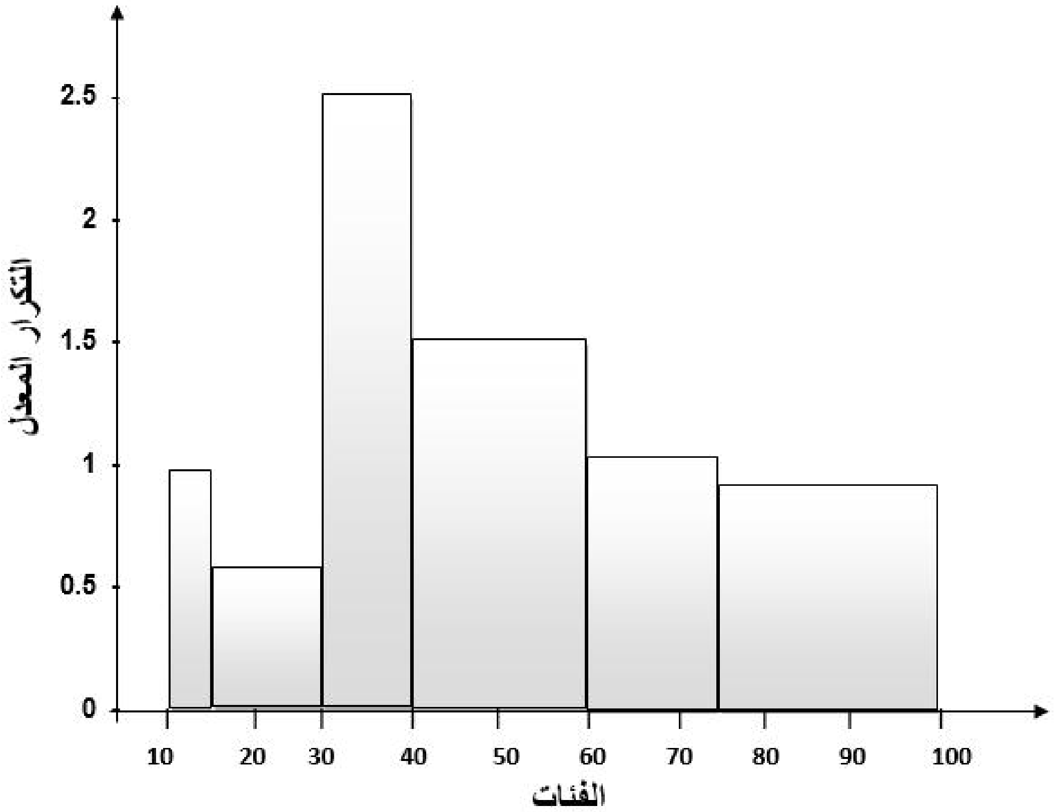
ملاحظة : اذا كانت الفئات غير متساوية عند رسم المدرج التكراري يتم استخراج التكرار المعدل حيث ان التكرار المعدل يساوي تكرار الفئة مقسوم على طول الفئة ويتم اعتماده في المحور العمودي.

مثال (2): التالي توزيع تكراري لفئات غير متساوية المطلوب رسم مدرج تكراري ؟

الفئات	**التكرارات** f_i	**طول الفئة** L	**التكرار المعدل** $f_i^* = f_i \,/\, L$
10-14	5	5	1
15-29	9	15	0.6
30-39	25	10	2.5
40-59	30	20	1.5
60-74	15	15	1
75-100	20	25	0.8
Total	**104**		

قانون التكرار المعدل يكتب بالصيغة التالية:

التكرار المعدل: $f_i^* = f_i \,/\, L$



2-2. المضلع التكراري :

هو وسيلة اخرى لتمثيل التوزيع التكراري بيانيا ويمكن رسمه **بإحدى** طريقتين اولهما :اذا كان المدرج التكراري **معلوم** . ويتم ذلك بتصنيف القواعد العليا لمستطيلات المدرج ثم نصل بين هذه النقط بمستقيمات ورسم فئة قبل الاولى تكرارها صفر وفئة بعد الاخيرة تكرارها صفر وبتصنيف هاتين الفئتين وتوصيل بقية الخط نحصل على ما يسمى **بالمضلع التكراري** .

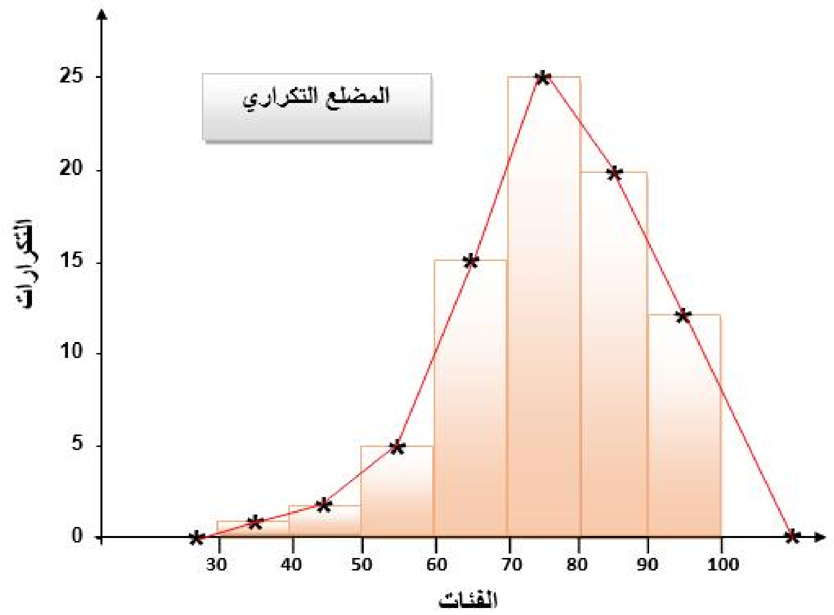
مثل (3) : **ارسم** ما يلي للبيانات التالية :

1- المضلع التكراري بطريقتين

2- المنحني التكراري

الفئات	30-40	40-50	50-60	60-70	70-80	80-90	90-100	**المجموع**
التكرارات	1	2	5	15	25	20	12	**80**

الطريقة الأولى:



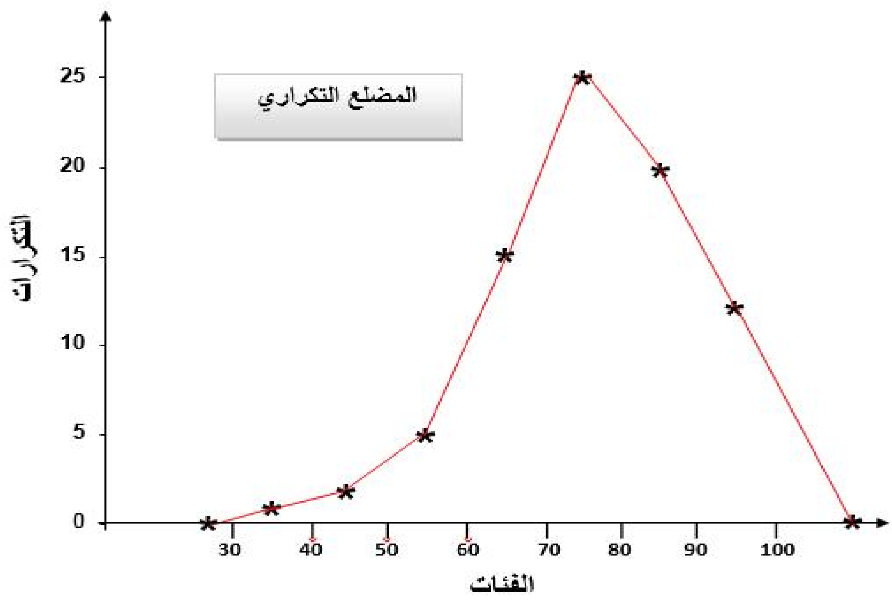
الطريقة الثانية : رسم المضلع التكراري على مراكز الفئات مباشرة دون ضرورة لرسم المدرج التكراري اولا وعليه فان المحور الافقي يمثل مراكز الفئات والمحور العمودي يمثل التكرارات ثم نصل النقاط بعضها ببعض وعليه فان خطوات رسم المضلع التكراري كما يأتي :

1- يجاد مراكز الفئات على المحور الافقي

2- تحديد النقطة التي تقابل مركز كل فئة على المحور الرأسي

3 – وصل مستقيمات بين النقط التي حددناها بعضها ببعض

وعليه فان المضلع التكراري يكون بالشكل التالي



2-3. المنحنى التكراري :

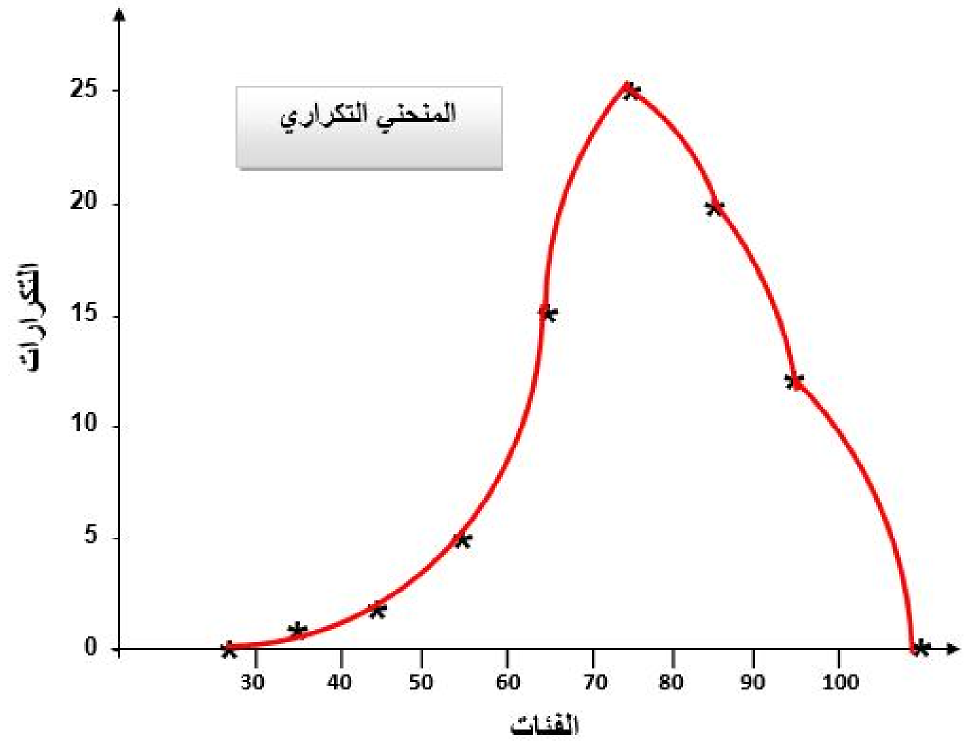
هو طريقة شائعة في الرسم البياني وهو عبارة عن منحنى يمر بمعظم النقاط الواقعة على مراكز الفئات والتي ارتفاعها يمثل تكرارات تلك الفئة

خطوات رسم المنحنى التكراري

1- نجد مراكز الفئات

2- نرسم الاحداثيين الافقي (الفئات) والعمودي (التكرارات) ثم نعين النقاط فوق مراكز الفئات ونصل بينها بمنحنى مستمر

ولرسم منحنى تكراري للمثال يكون كالاتي:



2-4. المنحنى التكراري المتجمع الصاعد:

لرسم هذا المنحنى نتبع الخطوات الاتية :

1- نكون جدولا تكراريا متجمعا صاعدا من الجدول التكراري البسيط

2- نرصد نقطا احداثياتها الافقية الحدود العليا للفئات واحداثياتها العمودية **التكرار المتجمع الصاعد** ونصل هذه النقاط ببعضها بخط منحني يكون هو المنحنى المتجمع الصاعد وتسري هذه الخطوات على الجداول الغير

منتظمة بدون ان نعدل التكرارات وذلك لان رسم المنحنى المتجمع الصاعد او النازل لتوزيع فئات غير متساوية لا يستدعي تعديل التكرارات .

مثال (4) : التوزيع الاتي يمثل ما تدفعه 150عائلة فلاحية للإيجار سنويا . **المطلوب**/ رسم منحنى متجمع صاعد لهذا التوزيع

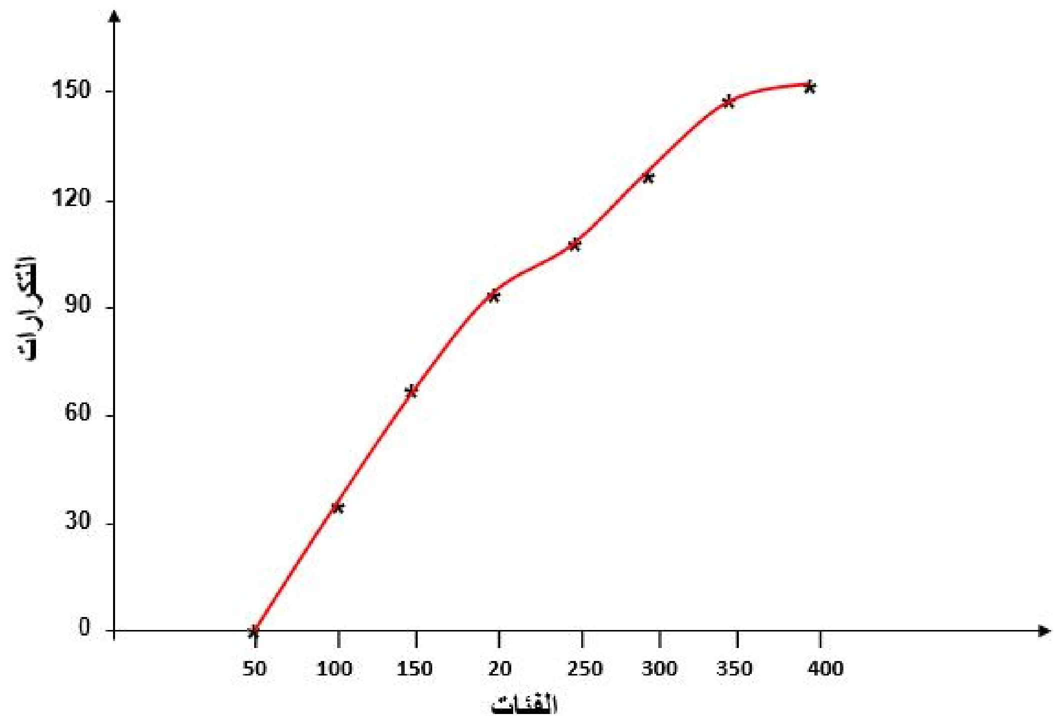
الفئات	50-100	100-150	150-200	200-250	250-300	300-350	350-400	**المجموع**
التكرارات	32	35	25	20	17	13	8	**150**

الحل :

نكون جدول التكرار المتجمع الصاعد وكما يلي :

التكرار المتجمع الصاعد	**الحدود العليا للفئات**
32	اقل من 100
67	اقل من 150
92	اقل من 200
112	اقل من 250
129	اقل من 300
142	اقل من 350
150	اقل من 400

نرسم منحني التكرار المتجمع الصاعد



المنحنى التكراري المتجمع الصاعد

5-2. المنحنى التكراري المتجمع النازل :

لرسم المنحنى من الجدول البسيط المنتظم وغير المنتظم نتبع الخطوات الاتية :

1- نكون جدولا تكراريا متجمعا نازلا من الجدول التكراري البسيط
2- نرصد نقاطا احداثياتها الافقية الحدود الدنيا للفئات واحداثياتها العمودية التكرارات المتجمعة النازلة ثم نصل هذه النقاط بعضها ببعض بخط منحني فيكون هو المنحنى التكراري المتجمع النازل .

مثال (5) : التوزيع الاتي يمثل ما تدفعه 150عائلة فلاحية للإيجار سنويا . المطلوب/ رسم منحنى متجمع صاعد لهذا التوزيع

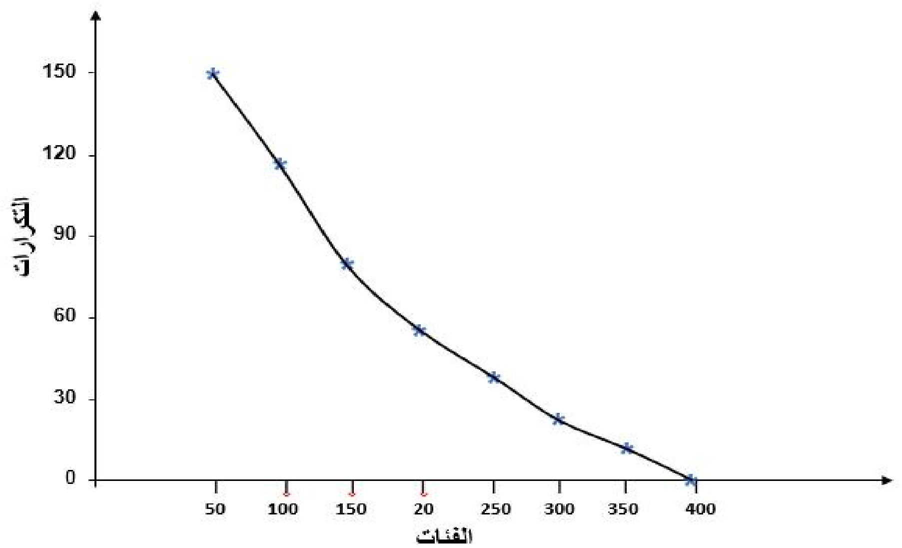
الفئات	50-100	100-150	150-200	200-250	250-300	300-350	350-400	**المجموع**
التكرارات	32	35	25	20	17	13	8	**150**

الحل :

نكون جدول التكرار المتجمع النازل وكما يلي :

التكرار المتجمع النازل	**الحدود الدنيا للفئات**
150	اكثر من 50
118	اكثر من 100
83	اكثر من 150
58	اكثر من 200
38	اكثر من250
21	اكثر من 300
8	اكثر من 350

نرسم منحني التكرار المتجمع النازل



المنحنى التكراري المتجمع النازل

2-6. المنحنى التكراري المتجمع الصاعد والنازل :

لرسم المنحنى من الجدول البسيط المنتظم وغير المنتظم نتبع الخطوات الاتية :

3- نكون جدولا تكراريا متجمعا نازلا من الجدول التكراري البسيط
4- نرصد نقاطا احداثياتها الافقية الحدود الدنيا للفئات واحداثياتها العمودية التكرارات المتجمعة النازلة ثم نصل هذه النقاط بعضها ببعض بخط منحني فيكون هو المنحنى التكراري المتجمع النازل .

3- نرصد نقطا احداثياتها الافقية الحدود العليا للفئات واحداثياتها العمودية التكرار المتجمع الصاعد ونصل هذه النقاط ببعضها بخط منحني يكون هو المنحنى المتجمع الصاعد وتسري هذه الخطوات على الجداول الغير منتظمة بدون ان نعدل التكرارات وذلك لان رسم المنحنى المتجمع الصاعد او النازل لتوزيع فئات غير متساوية لا يستدعي تعديل التكرارات .

مثال (6) : التوزيع الاتي يمثل ما تدفعه 150عائلة فلاحية للإيجار سنويا . المطلوب، **رسم منحنى متجمع صاعد ونازل** لهذا التوزيع

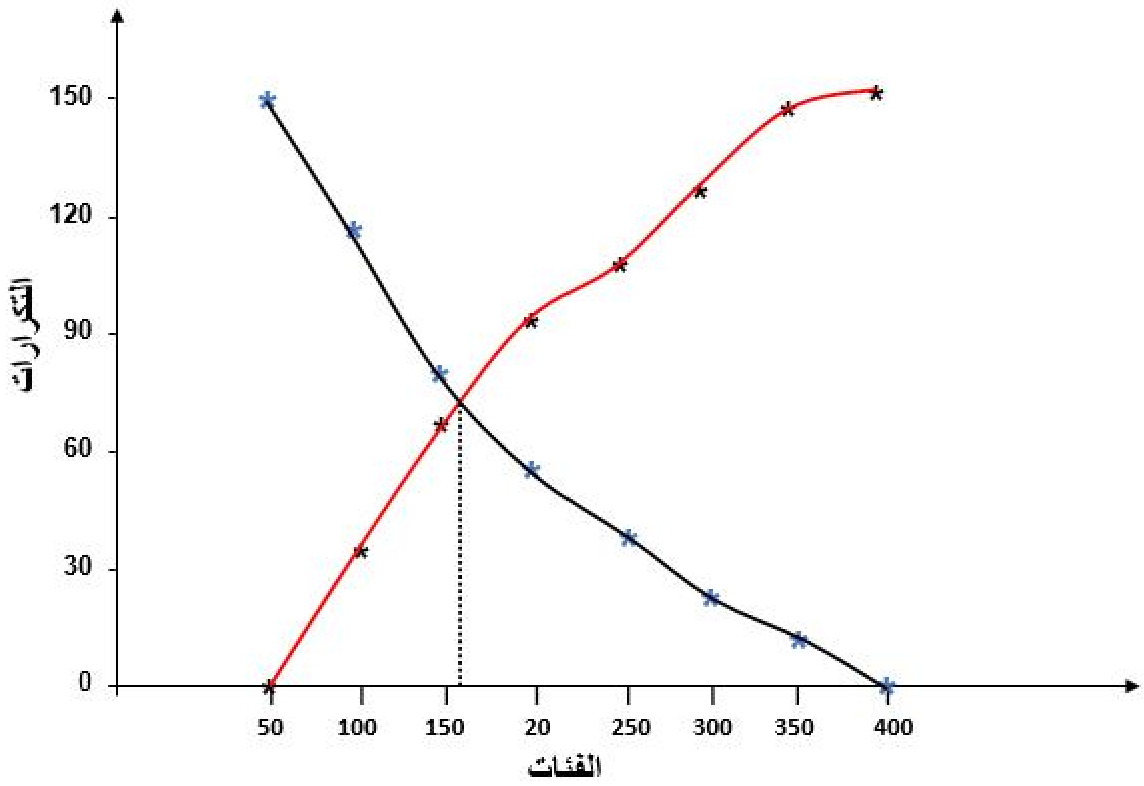
الفئات	50-100	100-150	150-200	200-250	250-300	300-350	350-400	**المجموع**
التكرارات	32	35	25	20	17	13	8	**150**

الحل :

نكون جدول التكرار المتجمع الصاعد النازل وكما يلي :

التكرار المتجمع النازل	الحدود العليا للفئات	**الحدود** الدنيا للفئات
150	اقل من 100	اكثر من 50
118	اقل من150	اكثر من 100
83	اقل من 200	اكثر من 150
58	اقل من 250	اكثر من 200
38	اقل من 300	اكثر من250
21	اقل من 350	اكثر من 300
8	اقل من 400	اكثر من 350

نرسم منحني التكرار المتجمع الصاعد و النازل معا وكما يلي



المنحنى التكراري المتجمع النازل والصاعد

الاسئلة

س1/ اكتب ما تعرفه عن ما يلي :

المدرج التكراري ، المضلع التكراري ، المنحني التكراري ، المنحني التكراري المتجمع الصاعد ، المنحني التكراري المتجمع النازل .

س2/ فرغ البيانات ادناه في جدول تكراري بسيط

48 , 38 , 51 , 45 , 56 , 39 , 50 , 71 , 65 , 34 , 51 , 66 , 27 , 73 , 69 , 34 , 43 , 34 , 53 , 54 , 71 ,
91 , 44 , 53 , 36 , 49 , 56 , 45 , 64 , 59 , 41 , 58 , 76 , 52 , 57 , 69 , 81 , 46 , 34 , 41 , 62 , 49 ,
43 , 55 , 79 , 66 , 87 , 81 , 70 , 67 , 55 , 53 , 84 , 52 , 56 , 44 , 51 , 65 , 76 , 52 , 54 , 33 , 95 ,
54 , 61 , 52 , 95 , 40 , 57 , 35 , 53 , 60 , 55 , 64 , 42 , 69 , 57 , 47 , 53 , 52 , 61 , 36 , 61 , 54 ,
57 , 80 , 46 , 61 , 54 , 94 , 55 , 85, 73 , 60 , 27 , 44 , 67 , 65 , 62 , 32 , 54

س3/ في تجربة لقياس السرع الاتية للمركبات على طريق خارجي اعطيت اليك البيانات كما في الجدول الاتي :
المطلوب عمل جدول توزيع متجمع صاعد وجدول توزيع متجمع نازل

الفئات	30-39	40-49	50-59	60-69	70-79	80-89	90-99	100-109	110-119	120-129	المجموع
التكرارات	3	6	24	64	50	29	14	6	3	1	200

س4/ الاتي توزيع تكراري لأوزان عينة من طلبة احدى الكليات قوامها مئة طالب المطلوب رسم مدرج تكراري لهذا التوزيع

الفئات	46-53	53-60	60-67	67-74	74-81	81-88	88-95	95-102	المجموع
التكرارات	7	15	27	21	14	8	5	3	100

س5/ البيانات التالية تمثل قيمة المنتجات المباعة بآلاف الدولارات لإحدى الشركات التجارية في الـ70 يوم الأولى من العلم الحالي.

56 75 70 66 60 **55** 65 70 65 56 66 71 62 67 71 61 67 61 70

60 75 69 71 57 69 72 68 57 72 68 65 63 73 66 63 58 73 67

62 72 58 74 60 81 80 74 76 74 73 58 72 **94** 78 91 85 77 83

77 82 76 62 78 88 64 87 55 79 57 64 79

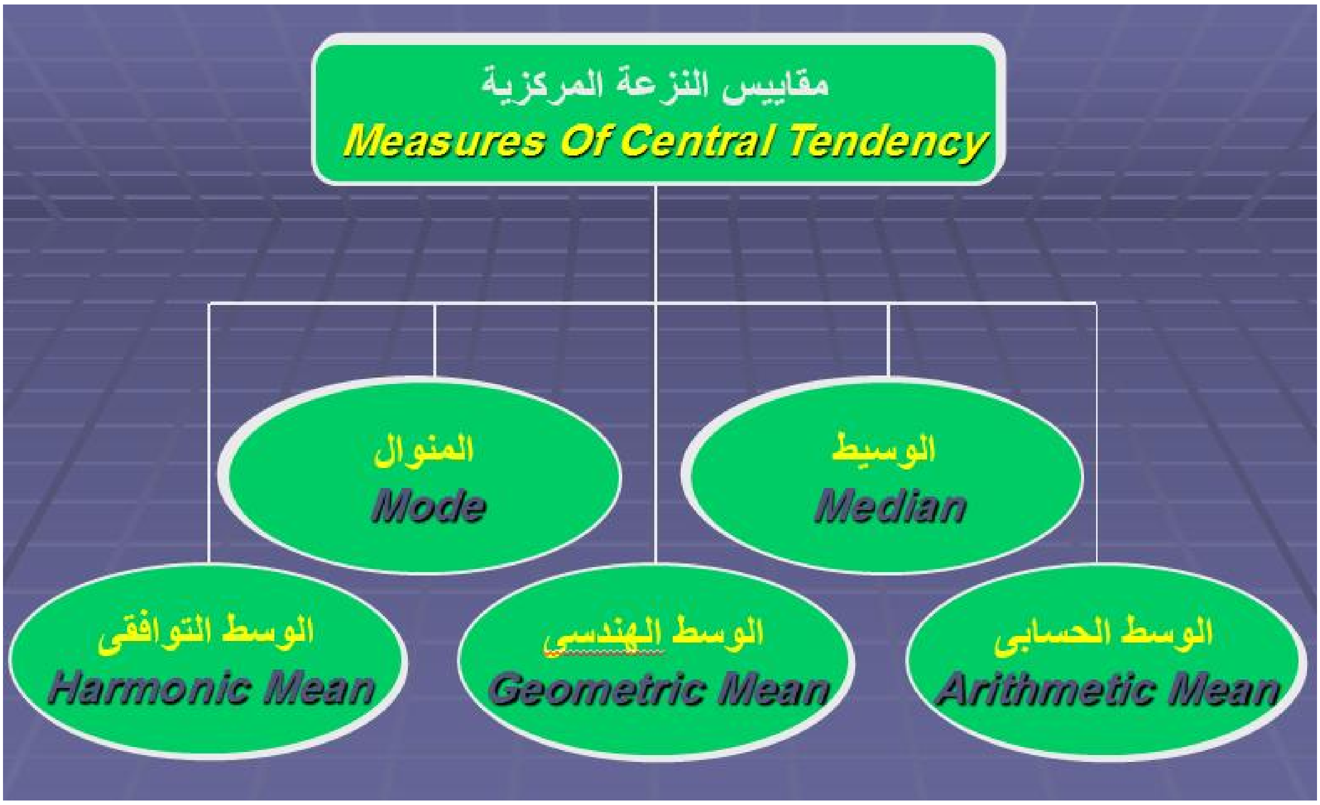
والمطلوب ما يلي :

1- كون التوزيع التكراري لقيمة المبيعات
2- رسم المدرج التكراري
3- رسم المضلع التكراري
4- رسم المنحني التكراري
5- رسم المنحني التكراري المتجمع الصاعد
6- رسم المنحني التكراري المتجمع النازل .

الفصل الثالث

مقاييس النزعة المركزية

3. مقاييس النزعة المركزية



مقاييس النزعة المركزية : رأينا في المحاضرات السابقة كيفية تمثيل البيانات بجداول ورسوم بغية تلخيصها وتوضيحها كذلك يمكن تمثيل البيانات بقيمة واحدة هي الوسط او المتوسط اي ان هذه البيانات تميل ان تقع في مركز البيانات المرتبة حسب الكبر لذلك تسمى مقاييس النزعة المركزية . والاوساط الاحصائية هي من اهم المقاييس الاحصائية الوصفية واكثرها استعمالا لدراسة البيانات ومقارنتها ، وهناك عدة انواع من المتوسطات واكثرها شيوعا واستعمالا هي :

الوسط الحسابي Arithmetic mean

الوسيط Median

المنوال Mode

الوسط الهندسي Geometric mean

الوسط التوافقي Harmonic mean

1-3. الوسط الحسابي :

يسمى في بعض الاحيان الوسط او المتوسط او المعدل الحسابي وهو من اهم مقاييس النزعة المركزية على الاطلاق لما يمتاز به من سهولة في استخراجه من جهة ولخضوعه للعمليات الحسابية من جهة اخرى . وهناك عدة طرق لاستخراجه وهي كالاتي :

1-1-3. الوسط الحسابي للبيانات غير المبوبة :

هناك طريقتين لحساب الوسط الحسابي

اولا : الطريقة المباشرة : الوسط الحسابي بموجب هذه الطريقة يمثل مجموع قياسات مفردات العينة مقسوما على عددها . يرمز للوسط الحسابي **بالرمز** Ẍ وعليه فان الوسط الحسابي بموجب هذه الطريقة

$$\ddot{X} = \frac{\sum_{i=1}^{n} x_i}{n}$$

مثال (1) : البيانات الاتية تمثل اوزان عينة من الطلبة قوامها 15**طالب** ، المطلوب ايجاد متوسط وزن **الطالب** في هذه العينة (متغيرات مستمرة)

50.2 , 60.9 , 68.3 , 59.2 , 58.1 , 62.3 , 65.3 , 52.9 , 61.5 , 63.2 , 59.1 , 69.3 , 64.2 , 65.2 , 56.6

الحل :

$$\ddot{X} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{50.1 + 60.9 + 59.2 + 58.1 + 62.3 + 65.3 + 52.9 + ... + 56.6}{15}$$

$$= \frac{916.3}{15} = 61.087 \ kg$$

لو قمنا بترتيب هذه البيانات تصاعديا لحصلنا على السلسلة التالية :

50.2 , 52.9 , 56.6 , 58.1 , 59.1 , 59.2, <u>60.9</u> , 61.9 , 62.3 , 63.2 , 64.2 , 65.2 , 68.3 , 69.3

نلاحظ تمركز قيمة Ẍ وسط هذه المجموعة هذا **ما نقصده** بمقاييس النزعة المركزية .

ثانيا : الطريقة المختصرة (طريقة الانحرافات)

تستخدم هذه الطريقة في حالة كون قياسات العينة اعداد كبيرة يصعب التعامل معها عند ايجاد الوسط الحسابي خصوصا عند عدم توفر حاسبات تفي بالغرض مما يفضل اختزال هذه الاعداد الى اعداد اصغر يسهل التعامل

معها . نختار وسط فرضي يكون قريب من الوسط الحسابي من نفس البيانات او خارج عنها ويرمز له A **ثم نجد** انحرافات القيم عن الوسط الفرضي ويرمز له di وعليه فان الوسط الحسابي يكون **كالاتي** :

$$\ddot{X} = A + \frac{\sum d_i}{n}$$

حيث ان :

تمثل الوسط الفرضي = A

تمثل انحرافات القيم عن الوسط الفرضي = di

وان $$d_i = x_i + A$$

ملاحظة قيمة A قيمة ثابتة نختارها من ضمن القيم ، نعتقد انها تمثل وسط القيم او مركزها اي تكون قريبة من **الوسط الحسابي**

مثال (2) : البيانات الاتية تمثل اوزان عينة من الطلبة قوامها 15 طالب ، المطلوب ايجاد متوسط وزن الطالب في هذه العينة (متغيرات مستمرة) بطريقة **الانحرافات** ؟

50.2 , 60.9 , 68.3 , 59.2 , 58.1 , 62.3 , 65.3 , 52.9 , 61.5 , 63.2 , 59.1 , 69.3 , 64.2 , 65.2 , 56.6

1- نحدد قيمة **A** بحيث تكون قريبة من الوسط الحسابي وفي قيمة **A تمثل 61.5** او اي قيمة اخرى نختارها
2- نجد قيمة **di والتي تساوي Xi-A** وعليه فان di للبيانات
3- **نجمع** Σ di

$d_i = x_i - A$	
$d_1 = 50.2 – 61.5 = - 11.3$	$d_9 = 61.5 - 61.5 = 0$
$d_2 = 60.9 – 61.5 = - 0.6$	$d_{10} = 63.2 – 61.5 = 1.7$
$d_3 = 68.3 – 61.5 = 6.8$	$d_{11} = 59.1 – 61.5 = - 2.4$
$d_4 = 59.2 – 61.5 = - 2.3$	$d_{12} = 69.3 – 61.5 = 7.8$
$d_5 = 58.1 – 61.5 = - 3.4$	$d_{13} = 64.2 – 61.5 = 2.7$
$d_6 = 62.3 – 61.5 = 0.8$	$d_{14} = 65.2 – 65.5 = - 0.3$
$d_7 = 65.3 – 61.5 = 3.8$	$d_{15} = 56.6 – 61.5 = - 4.9$
$d_8 = 52.9 – 61.5 = - 8.6$	
Σ di = - 5	

4- نطبق القانون

$$\ddot{X} = A + \frac{\sum d_i}{n}$$

$$\ddot{X} = A + \frac{\sum d_i}{n} = 61.5 + \frac{-5}{15} = 61.5 + 0.333 = 61.833$$

2-1-3. الوسط الحسابي لبيانات مبوبة :

اولا : الطريقة المباشرة

$$\ddot{X} = \frac{\sum f_i x_i}{\sum f_i}$$

حيث أن :

الوسط الحسابي = $\ddot{X}$

التكرارات = f_i

مراكز الفئات = X_i

خطوات ايجاد الوسط الحسابي لبيانات مبوبة :

1- تعيين مراكز الفئات

2- ضرب مركز كل فئة في التكرار المقابل له

3- **قسمة مجموع (حاصل ضرب مركز كل فئة** * تكرارها) **على مجموع التكرارات**

مثال (3) الجدول الاتي يبين توزيع الاجور الاسبوعية لـ 63 **عامل في احدى للشركات**

الفئات	50-60	60-70	70-80	80-90	90-100	100-110	المجموع
التكرارات	8	10	16	14	10	5	63

الحل :

الفئات	التكرارات (f_i)	مراكز الفئات (X_i)	$f_i\ X_i$
50-60	8	55	440
60-70	10	65	650
70-80	16	75	1200
80-90	14	85	1190
90-100	10	95	950
100-110	5	105	525
المجموع	63		4955

نطبق القانون

$$\ddot{X} = \frac{\sum f_i\ x_i}{\sum f_i}$$

$$\ddot{X} = \frac{\sum f_i\ x_i}{\sum f_i} = \frac{4950}{63} = 78.65$$

ثانيا : الطريقة المختصرة

نحتاج الى اختيار الوسط الفرضي ثم ايجاد انحرافات القيمة عن الوسط الفرضي حيث ان :

$$\mathbf{d_i = X_i - A}$$

وان X_i تمثل مركز الفئة ثم نستخدم القانون الاتي :

$$\ddot{X} = A + \frac{\sum f_i\ d_i}{\sum f_i}$$

خطوات ايجاد الوسط الحسابي لبيانات مبوبة بالطريقة المختصرة :

1- نجد مراكز الفئات $\mathbf{X_i}$
2- نحدد وسطا فرضيا $\mathbf{A}$
3- نجد انحرافات المراكز عن الوسط الفرضي $\mathbf{d_i}$
4- نضرب كل انحراف في التكرار المقابل له $\mathbf{d_i\ f_i}$
5- نجمع حاصل الضرب (الانحرافات * التكرارات)
6- نطبق القانون

ملاحظة : اختيار الوسط الفرضي يؤدي الى تبسيط العمليات الحسابية ويكون اختياره حسب القواعد الاتية :

1- ان يكون الوسط الفرضي مركز **لأحدى الفئات**
2- ان يكون قريبا من الوسط الحسابي
3- ان يكون امام اكبر تكرار **المثل** :

مثل (4) الجدول الاتي يبين توزيع الاجور الاسبوعية لـ 63 **عامل في احدى للشركات ، اوجد الوسط الحسابي**

الفئات	50-60	60-70	70-80	80-90	90-100	100-110	المجموع
التكرارات	8	10	16	14	10	5	63

الحل

الفئات	**التكرارات f_i**	**مركز الفئات (X_i)**	**$d_i = X_i - A$**	**$f_i d_i$**
50-60	8	55	55 – 75 = - 20	-160
60-70	10	65	65 – 75 = -10	-100
70-80	16	**75**	75 – 75 = 0	0
80-90	14	85	85 – 75 = 10	140
90-100	10	95	95 – 75 = 20	200
100-110	5	105	105 – 75 = 30	175
المجموع	**63**			**230**

نطبق القانون

$$\ddot{X} = A + \frac{\sum f_i d_i}{\sum f_i} = 75 + \frac{230}{63} = 78.65$$

ملاحظة: لقد تم اختيار 75 **كوسط فرضي** لأنه يقابل اكبر تكرار

3-1-3. مزايا وعيوب الوسط الحسابي

يتميز الوسط الحسابي بالمزايا التالية :

1. أنه سهل الحساب .
2. يأخذ في الاعتبار كل القيم .
3. أنه أكثر المقاييس استخداما وفهما .

ومن عيوبه .

1. أنه يتأثر بالقيم الشاذة والمتطرفة .
2. يصعب حسابه في حالة البيانات الوصفية .
3. يصعب حسابه في حالة الجداول التكرارية المفتوحة

2-3. الوسيط Median

الوسيط لمجموعة من القيم هو القيمة التي يقع ترتيبها وسط المجموعة عند ترتيب هذه القيم تصاعديا او تنازليا . **اي** تقسم المجموعة الى قسمين متساويين ، بحيث يساوي عدد الحدود التي اصغر من الوسيط عدد الحدود الاكبر منه ، فمثلا لو كانت لدينا القيم 2 , 5 ,9 , 6 ,17 واردنا ايجاد الوسيط لهذه المجموعة فإننا نرتب القيم تصاعديا فتصبح 17 , 9 , 6 , 5 , 2 فتكون القيمة التي تقع في الوسط هي الوسيط اي ان 6 هي الوسيط

1-2-3. الوسيط لبيانات غير مبوبة

أ- اذا كان عدد القيم فرديا فيكون ترتيب الوسيط كما في الصيغة الاتية :

$$T = \frac{n+1}{2}$$

حيث ان :

T = **تمثل** ترتيب الوسيط وان **n** تمثل عدد القيم

خطوات ايجاد الوسيط

1- ترتيب القيم اما تصاعديا او تنازليا

2- نجد ترتيب الوسيط حسب الصيغة الاتية :

$$T = \frac{n+1}{2}$$

3- تكون قيمة الوسيط هي القيمة الموجودة امام الترتيب الناتج في الخطوة (2)

مثال (5) - اوجد الوسيط للبيانات التالية

134 , 78 , 204 , 63 , 12 , 189 , 152 .

الحل :

نرتب القيم اما تصاعديا او تنازليا

ترتيب تصاعدي 12 , 63 , 78 , **134** , 152 , 189 , 204

او ترتيب تنازلي 204 , 189 , 152 , **134** , 78 , 63 , 12

ترتيب الوسيط

$$T = \frac{n+1}{2} = \frac{7+1}{2} = 4$$

الوسيط هو الترتيب الرابع اي ان :

$$Me = 134$$

ب- اذا كانت عدد القيم (n) زوجي فيكون الوسيط هو الوسط الحسابي لقيمتي الترتيبين اللذين تسلسلهما على التوالي

$$T = \frac{n}{2} \quad ; \quad T = \frac{n}{2} + 1$$

مثال (6) - المطلوب ايجاد الوسيط **للأعداد** الاتية :

152 , 189 , 12 , 63 , 204 , 78 , 134 , 7

الحل : نرتب القيم اما تصاعديا او تنازليا

الترتيب التصاعدي 7, 12 , 63 , 78 , 134 , 152 , 189 , 204 .

$$T = \frac{n}{2} = \frac{8}{2} = 4$$

العدد الذي ترتيبه الرابع هو 78

$$T = \frac{n}{2} + 1 = \frac{8}{2} + 1 = 5$$

العدد الذي ترتيبه الخامس هو 134

الوسيط يمثل الوسط الحسابي لهاتين القيمتين اي ان :

$$Me = \frac{78+134}{2} = \frac{212}{2} = 106$$

اما لو كان الترتيب تنازلي : 7 , 12 , 63 , 78 , 134 , 152 , 189 , 204

$$T = \frac{n}{2} = \frac{8}{2} = 4$$

العدد الذي ترتيبه الرابع هو 134

$$T = \frac{n}{2} + 1 = \frac{8}{2} + 1 = 5$$

العدد الذي ترتيبه الخامس هو 78

الوسيط يمثل الوسط الحسابي لهاتين القيمتين اي ان :

$$Me = \frac{134 + 78}{2} = \frac{212}{2} = 106$$

3-2-2. الوسيط لبيانات مبوبة :

أ- الوسيط لبيانات مبوبة (متغير متقطع)

المتغيرات المنفصلة او المتقطعة (Discrete Variables) : وهي المتغيرات الكمية التي **تأخذ** قيما عددية محددة صحيحة ولا تحتوي علي قيم كسرية مثل : عدد المصانع في كل مدينة من مدن دولة ما ، عدد حوادث السيارات التي وقعت في الربع الاول من هذه السنة في 5 مناطق مختلفة ، وعدد المدرسين **في كل مدرسة من الم**دارس الثانوية في مدينة **الموصل** خلال هذه السنة ، ....الخ.

ويمكننا ايجاد الوسيط من الجداول التكرارية البسيطة بتحويلها الى جداول تكرارية صاعدة او نازلة .

الوسيط في حالة التكرار المتجمع الصاعد

خطوات ايجاد الوسيط لبيانات مبوبة لمتغير متقطع

1- نجد التكرار المتجمع الصاعد

نجد ترتيب الوسيط والذي يساوي $T = \frac{\sum f_i}{2}$

ملاحظة: f_i هنا التكرارات الاصلية وليس التكرار المتجمع الصاعد

2- نحدد قيمة الوسيط وهي التي تقع بين التكرارين (يعني ترتيب الوسيط بين التكرارين)

3- نحدد فئة الوسيط مركز هذه الفئة يمثل الوسيط

مثال (7) :الاتي توزيع لعينة من الاسر **حسب عدد افراد الاسرة ، المطلوب** حساب الوسيط لعدد افراد الاسرة

الفئات	2-4	5-7	8-10	11-13	14-16	17-19	20-22	**المجموع**
التكرارات	6	9	12	20	14	11	8	**80**

الحل:

الفئات	**التكرارات** (fi)	الحدود العليا للفئات	**التكرار المتجمع الصاعد** (Fi)
2-4	6	4 **فأقل**	6
5-7	9	7 **فأقل**	15
8-10	12	10 **فأقل**	**27**
11-13	20	13 **فأقل**	**47**
14-16	14	16 **فأقل**	61
17-19	11	19 **فأقل**	72
20-22	8	22 **فأقل**	80
المجموع	**80**		

ترتيب الوسيط :

$$T = \frac{\sum f_i}{2} = \frac{80}{2} = 40$$

اي الوسيط يقع بين التكرارات **27 و47**

وعليه فان فئة الوسيط هي الفئة **(11-13)** لأنها

الاقرب الى 47 وان الوسيط يمثل مركز هذه الفئة

وعليه فالوسيط يساوي **12 وكالاتي** :

$$Me = \frac{11 + 13}{2} = 12$$

ب- الوسيط لبيانات مبوبة لمتغير مستمر

المتغيرات الكمية المتصلة (Continuous Variables): وهي المتغيرات الكمية التي **تأخذ** قيما تكون عددا صحيحا وكسرا من وحدة القياس مثل : متغير الدخل اليومي (بالدينار) لعينة من الاشخاص ،

الطول (مقاسا بالسنتمترات) ، الوزن (مقاسا بالكيلو جرام) ، العمر (مقاسا بالسنوات) درجة حرارة الجو (مقاسة بالدرجات) ، المساحة (مقاسة بالمتر المربع) وغيرها.

نستخدم القانون الاتي :

$$Me = L_K + \frac{\frac{\sum f_i}{2} - F_{K-1}}{F_K - F_{K-1}} \times h_K$$

حيث ان :

$\boldsymbol{Me}$ = الوسيط

$\boldsymbol{L_k}$ = الحد الادنى لفئة الوسيط

$\sum \mathbf{f}_i$ = مجموع التكرارات الاصلية

$\boldsymbol{F_{k-1}}$ = تكــــرار متجمع صاعدسابق

$\boldsymbol{F_k}$ = تكرار متجمع صاعد لاحق

$\boldsymbol{h_k}$ = طول فئة الوسيط

خطوات الحل :

1- نجد اما التكرار المتجمع الصاعد او التكرار المتجمع النازل
2- نجد ترتيب الوسيط من الصيغة

$$T = \frac{\sum f_i}{2}$$

3- نحدد فئة الوسيط والتي تقع بين التكرارين
4- نطبق صيغة القانون

$$Me = L_K + \frac{\frac{\sum f_i}{2} - F_{K-1}}{F_K - F_{K-1}} \times h_K$$

مثال (8) - اوجد الوسيط من التوزيع التكراري الاتي :

التكرار المتجمع الصاعد (Fi)	الحدود العليا للفئات	**التكرارات** (fi)	**الفئات**
8	60 فأقل	8	50-60
18	70 فأقل	10	60-70
34	80 فأقل	16	70-80
48	90 فأقل	14	80-90
58	100 فأقل	10	90-100
63	110 فأقل	5	100-110
65	110 فأقل	2	110-120
		65	**المجموع**

الحل :

ترتيب الوسيط :

$$T = \frac{\sum f_i}{2} = \frac{65}{2} = 32.5$$

الوسيط يقع بين التكرارات (18 و 34) ← اي ان فئة الوسيط هي (80 – 70)

نطبق صيغة القانون الاتية :

$$Me = L_K + \frac{\frac{\sum f_i}{2} - F_{K-1}}{F_K - F_{K-1}} \times h_K$$

$$L_K = 70 \;\; ; \;\; f_i = 65 \;\; ; \;\; F_K = 34 \;\; ; \;\; F_{K-1} = 18 \;\; ; \;\; h_K = 10$$

$$Me = L_K + \frac{\frac{\sum f_i}{2} - F_{K-1}}{F_K - F_{K-1}} \times h_K = 70 + \frac{\frac{65}{2} - 18}{34 - 18} \times 10 = 70 + \frac{145}{16} = 79.63$$

ملاحظة (1) : في حالة التكرار المتجمع الصاعد نختار الفئة الوسطية المقابلة للتكرار الاقرب

وفي حالة التكرار المتجمع النازل نختار الفئة الوسطية المقابلة للتكرار الابعد

ملاحظة (2) :**في حالة المتجمع** النازل نطبق الصيغة الاتية :

$$Me = L_K + \frac{F_{K-1} - T}{F_K - F_{K-1}} \times h_K$$

ملاحظة (3) : يمكن ايجاد الوسيط من بيانات مفتوحة كما يمكن ايجاده اذا كانت اطوال الفئات غير متساوية دون الحاجة الى تعديل التكرارات .

3-2-3. حساب الوسيط باستخدام الرسم

يمكن حساب الوسيط للبيانات المبوبة بطريقة آخري عن طريق رسم المنحنى المتجمع الصاعد أو الهابط أو كليهما .

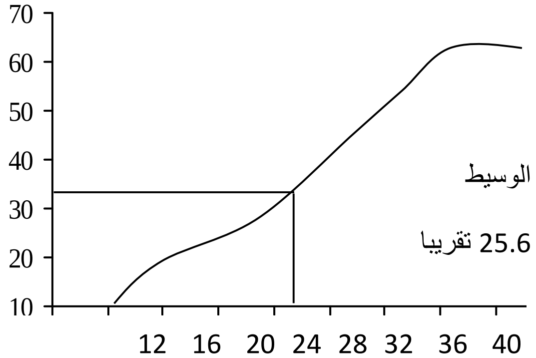
وإذا قمنا باستخدام إحدى الطريقتين الأولى والثانية فإننا :

1- نوجد ترتيب الوسيط .

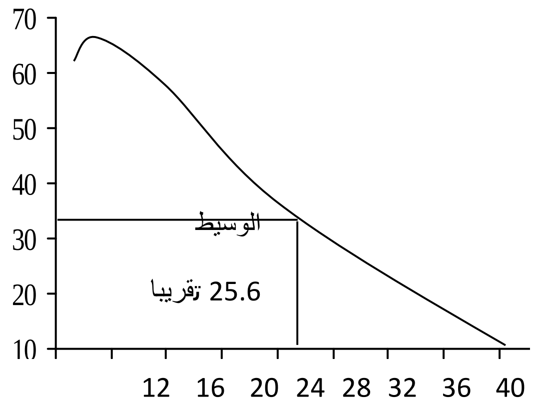
$$T = \frac{\sum f_i}{2}$$

2- نحددها على المحور الصادي ثم نرسم من هذه النقطة مستقيم يوازى المحور السيني ومن نقطة تقابله مع المنحنى نسقط عمود فتكون نقطة التقائه مع المحور السيني هي قيمة الوسيط.

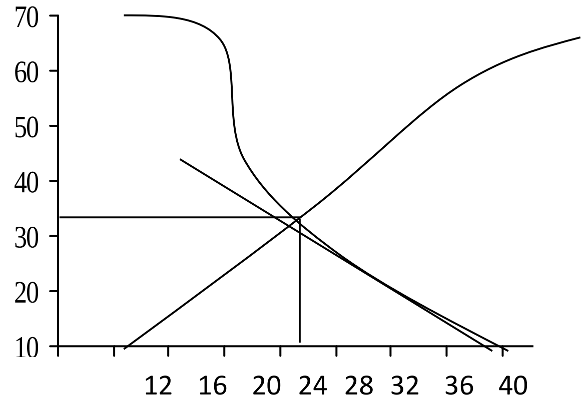
أما إذا استخدمنا الطريقة الثالثة (المنحنيان معا) فنقوم برسم المنحنين المتجمعين الصاعد والهابط معا فيتقابل المنحنيان معا في نقطة واحدة ، نسقط منها عمود يلاقى المحور السينى فى نقطة هي تمثل قيمة الوسيط .



إيجاد الوسيط بيانياً باستخدام التكرار المتجمع الصاعد .



أيجاد الوسيط بيانياً باستخدام التكرار المتجمع الهابط .



استخدام المنحنى الصاعد والمنحنى الهابط

3-2-3. مزايا وعيوب الوسيط

من مزايا الوسيط

1- لا يتأثر بالقيم الشاذة أو المتطرفة .

2- كما أنه سهل في الحساب .

3- مجموع قيم الانحرافات المطلقة عن الوسيط أقل من مجموع الانحرافات المطلقة عن أي قيم أخرى . أي أن : $\sum |x - Med| \leq \sum |x - a|\ ,\ \ a \neq Med$

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 nominal

3-3. المنوال Mode

المنوال ، مفهومه ،حسابه للبيانات غير المبوبة والبيانات المبوبة (طريقة بيرسون حسابيا وبيانيا)

هو أحد مقاييس النزعة المركزية المناسبة لمستويات القياس الأسمى ، ويعرف المنوال بأنه القيمة الأكثر شيوعاً أو القيمة **التي** تتكرر أكثر من غيرها .

1. يوصف التوزيع بأنه وحيد المنوال unimodal
2. وقد يكون للتوزيع منوا لين Bimodal
3. وأحيانا يكون له عدة منا ويل multi- modal

وفى مجموعة البيانات الصغيرة حيث لا تتكرر القيم لا يوجد منوال.

وعندما يكون للبيانات أكثر من منوال فلا يجوز حساب متوسطها لان ذلك يتنافى مع معنى المنوال (القيمة الأكثر تكراراً) ، كما أنه إذا حسب متوسط المنوالين مثلا فقد يكون لقيمة أقل تكراراً.

3-3-1. حساب المنوال من بيانات غير مبوبة

مثل (9) : اوجد المنوال في البيانات الاتية

3, 4 , 5 , 6 , 2 , 3

نرتب ترتيب تصاعدي

2 , 3 , **3** , **3** , 4 , 5 , 6

الحل : المنوال هو الرقم **3** لأنه تكرر اكثر من غيره من بين مفردات المجموعة .

ملاحظة : بعض القيم تكون عديمة المنوال اذا لم يوجد رقم متكرر اكثر من غيره كما قد يكون **هناك اكثر من منول في المجموعة في حالة اكثر من رقم** .

مثل (10) : احسب المنوال للبيانات التالية

2 , 4 , 6 , 8 , 10 , 12 , 14 , 19

لا يوجد منوال لهذه البيانات لأنه لا يوجد **رقم متكرر**

مثل (11) : جد المنوال للبيانات التالية

2 , 4 , 8 , 10 , 12 , 2 , 5 , 4 , 6

نرتب ترتيب تصاعدي

2 , **2** , **4** , **4** , 5 , 6 , 8 , 10 , 12

المنوال هنا **4 منول و 2** منوال لانهما تكررا بنفس المقدار

2-3-3. المنوال لبيانات مبوبة

يمكن ايجاد المنوال بعدة طرق بعد ايجاد الفئة المنوالية ، وتعرف الفئة المنوالية بانها : الفئة التي تحتوي **على** اكبر تكرار، وذلك لان المنوال حسب التعريف هو القيمة التي تتكرر اكثر من غيرها . ومن هذه الطرق

طريقة بيرسون وتسمى ايضا طريقة الفروق

ملخص هذه الطريقة (الخطوات)

1- نختار اكبر تكرار نفرضه f_k والفئة التي تقابله هي الفئة المنوالية (طولها h_k)
2- **نحدد التكرار** الذي قبله ويرمز له f_{k-1}
3- نحدد التكرار الذي بعده ويرمز له f_{k+1}
4- نحدد الحد الادنى للفئة المنوالية (بداية الفئة المنوالية) ويرمز لها L_k
5- نحسب قيمة المنوال بتطبيق صيغة القانون الاتي :

$$Mo = L_K + \frac{f_K - f_{k-1}}{(f_K - f_{K-1}) + (f_K - f_{K+1})} \times h_K$$

ولتسهيل الامر نرمز لـ الفرق بين تكرار الفئة المنوالية والفئة التي قبلها بـ Δ_1 ونرمز للفرق بين تكرار الفئة المنوالية والفئة التي بعدها بـ Δ_2 فيصبح القانون **كالاتي** :

$$\Delta_1 = f_K - f_{K-1} \qquad ; \qquad \Delta_2 = (f_K - f_{k-1}) + (f_K - f_{K+1})$$

$$Mo = L_K + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times h_K$$

حيث ان :

المنول = Mo

بداية الفئة المنوالية = L_k

تكرار الفئة المنوالية = f_k

التكرار السابق للفئة المنوالية = f_{k-1}

التكرار اللاحق للفئة المنوالية = f_{k+1}

طول الفئة المنوالية = h_k

مثال (12) : من الجدول الاتي اوجد المنوال بطريقة بيرسون

الفئات	60-50	70-60	80-70	90-80	100-90	110-100	120-110	**المجموع**
التكرارات	8	10	16	14	10	5	2	**65**

الحل :

الفئات	**التكرارات (f_i)**
50-60	8
60-70	10
70-80	**16**
80-90	14
90-100	10
100-110	5
110-120	2
المجموع	65

F_{k-1}

الفئة المنوالية تقابل اكبر تكرار f_k

F_{k+1}

1- اكبر التكرارات f_k هو 16 وعليه فان الفئة المنوالية هي (70-80)

2- F_{k-1} (التكرار السابق للفئة المنوالية هو 10

3- F_{k+1} (التكرار اللاحق للفئة المنوالية هو 14

4- L_k بداية الفئة المنوالية هو 70

5- h_k طول الفئة المنوالية هو 10

6- نطبق القانون

$$Mo = L_K + \frac{f_K - f_{k-1}}{(f_K - f_{K-1}) + (f_K - f_{K+1})} \times h_K$$

$$Mo = 70 + \frac{16 - 10}{(16 - 10) + (16 - 14)} \times 10 = 70 + \frac{60}{8} = 77.5$$

ملاحظة (1) : يمكن ايجاد المنوال من بيانات مفتوحة وهذه احد مزاياه

ملاحظة (2) : اذا كانت اطوال الفئات غير متساوية يستلزم تعديل التكرارات . والتكرار المعدل هو تكرار الفئة الاصلي مقسوم على طول الفئة ويرمز للتكرار المعدل بـ fi^*

مثال (13) : من التوزيع الاتي اوجد المنوال

الفئات	**التكرار** f_i	**طول الفئة** L	**التكرار المعدل** f_i^*
5-10	2	5	2 / 5 = 0.4
10-15	6	5	6 / 5 =1.2
15-25	10	10	10 / 10 = **1**
25-35	22	10	22 / 10 = **2.2**
35-50	27	15	27 / 15 = **1.8**
50-60	11	10	11 / 10 = 1.1
Total	**68**		

$$Mo = L_K + \frac{f_K - f_{k-1}}{(f_K - f_{K-1}) + (f_K - f_{K+1})} \times h_K$$

$$Mo = 25 + \frac{2.2 - 1}{(2.2 - 1) + (2.2 - 1.8)} \times 10 = 25 + \frac{1.2 \times 10}{1.2 + 0.4} = 25 + \frac{12}{1.6} = 32.5$$

3-3-3. أيجاد المنوال عن طريق الرسم

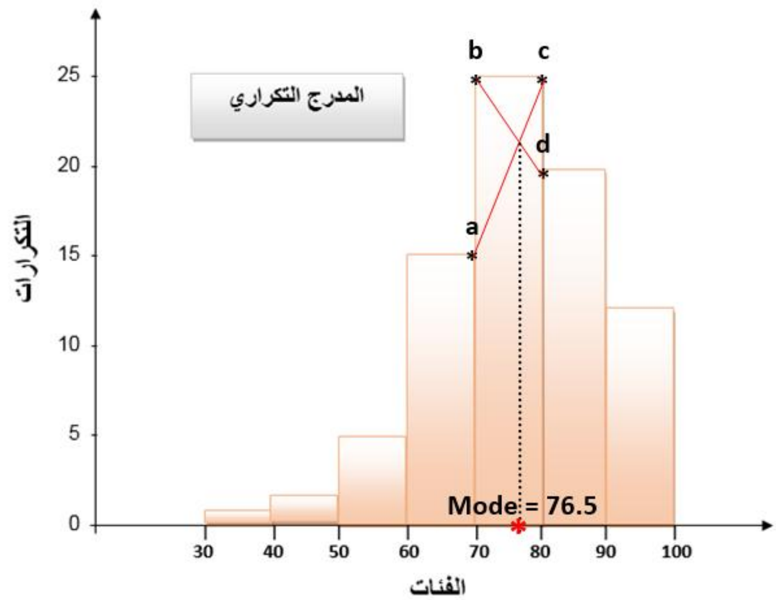
يمكن تقدير المنوال بطريقة الرسم البياني للمدرج التكراري وذلك باستعمال مستطيل الفئة المنوالية (وهو اعلى مستطيل لأنه يمثل اكثر التكرارات) والمستطيلان المجاوران له.

فان المنوال يتحدد بوصل النقطة **a** مع النقطة **c** والنقطة **b** مع النقطة **d** ومن نقطة تلاقيهما ننزل عمودا على المحور السيني يقطعه في نقطة هي قيمة المنوال .

مثال (14) :الجدول الاتي يمثل توزيع تكراري **لأطوال نباتات القطن ،المطلوب** ايجاد ا لمنوال بطريقة **الرسم** ؟

الفئات	30-40	40-50	50-60	60-70	70-80	80-90	90-100	**المجموع**
التكرارات	1	2	5	15	25	20	12	**80**

الحل:



3-3-4. مميزات وعيوب المنوال

مميزات المنوال :

1- مقياس سهل حسابه ولا يتأثر بالقيم الشاذة.

2- يمكن ايجاده للقيم الوصفية والتوزيعات التكرارية المفتوحة .

عيوب المنوال :

1- **عند حسب المنول لا تؤخذ** جميع قيم البيانات في الاعتبار .

2- قد يكون لبعض البيانات اكثر من منوال وبذلك لا يمكن تحديد قيمة وحيدة المنوال .

3-3-5. العلاقة بين الوسط الحسابي والوسيط والمنوال

توجد علاقة تجريبية بين المقاييس الثلاثة (الوسط الحسابي – الوسيط – المنوال) وذلك في حلة التوزيعات التكرارية أحادية المنوال وغير المتماثلة والمتماثلة وذات الالتواء البسيط ، وتعطى هذه العلاقة من خلال المعادلة التالية :

الوسط الحسابي - المنول = 3 (**الوسط الحسابي** - الوسيط)

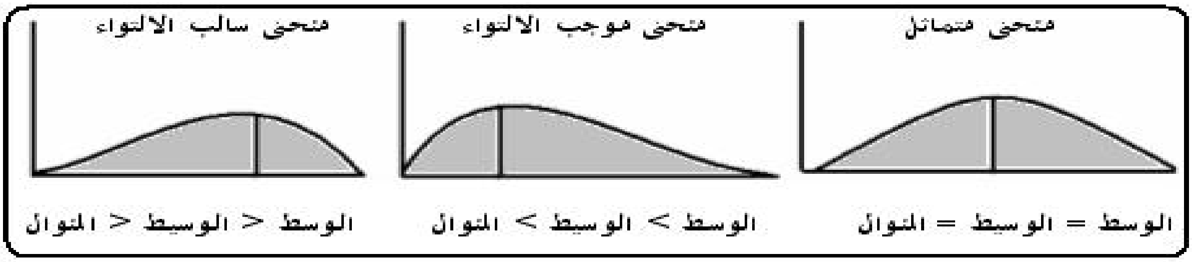
وقد وجد ان الوسيط تقع قيمته بين قيمتي الوسط الحسابي والمنوال .

وفي حالة التوزيعات التكرارية المتماثلة الوحيد المنوال فان قيمة الوسط الحسابي تكون مساوية لقيمة الوسيط تكون مساوية لقيمة المنوال أي ان :

الوسط الحسابي = الوسيط – **المنول**

3-3-6. استخدام مقاييس النزعة المركزية في تحديد شكل توزيع البيانات

يمكن استخدام الوسط الحسابي والوسيط والمنوال في وصف المنحنى التكراري، والذي يعبر عن شكل توزيع البيانات ، كما يلي :



- يكون المنحنى متماثل إذا كان :

الوسط = الوسيط = المنوال .

- يكون المنحنى موجب الالتواء (ملتوي جهة اليمين) إذا كان:

الوسط > الوسيط > المنوال

- يكون المنحنى سالب الالتواء (ملتوي جهة اليسار) إذا كان :

الوسط < الوسيط < المنوال

امثلة محلولة

مثال (15)

قام مدير مراقبة الإنتاج بسحب عينة من 10 عبوات من المياه المعبأة للشرب ، ذات الحجم 5 لتر ، والمنتجة بواسطة إحدى شركات تعبئة المياه لفحص كمية الأملاح الذائبة، وكانت كالتالي :

115 123 119 123 124 119 123 121 123 121

والمطلوب : حساب الوسط الحسابي، والوسيط، والمنوال، ثم حدد شكل الالتواء لهذه البيانات .

الحل

1- حساب الوسط الحسابي :

$$\bar{X} = \frac{\sum x}{n} = \frac{1211}{10} = 121.1$$

2- حساب الوسيط :

رتبة الوسيط : $(n+1)/2 = (10+1)/2 = 5.5$

ترتيب القيم تصاعديا

قيمة الوسيط: 122

الطاقة	115	119	119	121	121	123	123	123	123	124
الرتبة	1	2	3	4	5	6	7	8	9	10

رتبة الوسيط: 5.5

عدد القيم = **10** ، وهو عدد زوجي. الوسيط = الوسط الحسابي للقيمتين رقم **(5 , 6)**

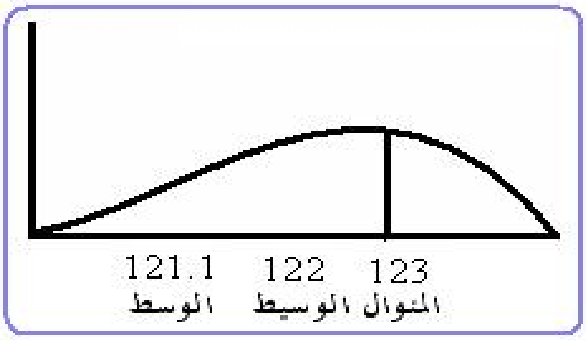
$$Med = \frac{121+123}{2} = \frac{244}{2} = 122$$

3- حساب المنوال :

المنوال يساوى القيمة الأكثر تكرارا: القيمة **123** تكررت أكثر من غيرها ، إذا

$$Mod = 123$$

وبمقارنة الوسط والوسيط و المنوال نجد أن :



نجد أن : الوسط < الوسيط < المنوال ، إذا توزيع بيانات كمية الأملاح سالبة الالتواء.

مثل (16)

الجدول التكراري التالي يعرض توزيع 100 عامل في مزرعة حسب الأجر اليومي بالدولار .

الأجر	50- 70	70-90	90-110	110-130	130-150	150-170	170 – 190	المجموع
عدد العمال	8	15	28	20	15	8	6	100

والمطلوب :

1- حساب الوسط والوسيط والمنوال .

2- بيان شكل توزيع الأجور في هذه المزرعة .

الحل

1- حساب الوسط والوسيط والمنوال .

أولا : الوسط الحسابي $\bar{x}$

فئات الأجر	التكرارات f	مراكز الفئات x	$f \cdot x$
50 – 70	8	60	480
70 – 90	15	80	1200
90 – 110	28	100	2800
110 - 130	20	120	2400
130 - 150	15	140	2100
150 – 170	8	160	1280
170 - 190	6	180	1080
المجموع	100		11340

$$\overline{X} = \frac{\sum fx}{\sum f} = \frac{11340}{100} = 113.4 \ \ R.S$$

ثانيا : الوسيط Med

رتبة الوسيط : (n/2 =100/2 =50)

تكوين التوزيع التكراري المتجمع الصاعد .

أقل من	تكرار متجمع صاعد
أقل من 50	**0**
أقل من 70	**8**
أقل من 90	$\leftarrow f_1$ **23**
أقل من 110	$\leftarrow f_1$ **51**
أقل من 130	**71**
أقل من 150	**86**
أقل من 170	**94**
أقل من 190	**100**

(رتبة الوسيط 50)

من الجدول أعلاه نجد أن :

$$\frac{n}{2}=50 \ , f_1=23 \ , f_2=51 \ , A=90 \ , L=110-90=20$$

إذا الوسيط قيمته هي :

$$Med = A+\frac{\frac{n}{2}-f_1}{f_2-f_1}\times L = 90+\frac{50-23}{51-23}\times 20$$

$$=90+\frac{27}{28}\times 20 = 90+\frac{540}{28}=90+19.286=109.3 \ \ R.S$$

ثالثا : المنوال Mod

الفئة المنوالية ، هي الفئة المناظرة لأكبر تكرار

أكبر تكرار = 28 ، وهو يناظر الفئة التقريبية (110 - 90) .

حساب الفروق : $d_2 = 28-20=8 \ , \ d_1=28-15=13$

الحد الأدنى للفئة : $A=90$ طول الفئة : $L=110-90=20$

إذا المنوال يحسب بتطبيق المعادلة التالية :

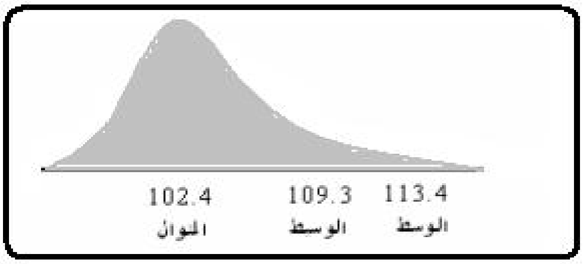
$$Mod = A+\frac{d_1}{d_1+d_2}\times L = 90+\frac{13}{13+8}\times 20 = 90+\frac{260}{21}=102.4 \ \ R.S$$

3-3-7. بيان شكل التوزيع .

من النتائج السابقة ، نجد أن :

الوسط الحسابي : $\overline{X} = 113.4$ الوسيط : $Med = 109.3$ المنوال : $Mod = 1024$

أي أن : الوسط > الوسيط > المنوال إذا توزيع بيانات الأجور موجب الالتواء. كما هو مبين في للشكل التالي:



3-4. المتوسط الهندسي geometric mean (G) والمتوسط التوافقي harmonic mean (H):

يفيد استخدام هذين المتوسطين في حالة الأرقام الموجبة، وإن استخداماتهما الأساسية تنحصر في حساب قيم نسبية مثل

1- الأدلة indexes
2- النسب ratios
3- المعَدلات rates
4- ويمكن حسابهما بالمعادلتين الآتيتين:

3-4-1. الوسط الهندسي *Geometric Mean*

أولا : في حالة البيانات المبوبة-:

إذا كانت قيم المتغير (x) هي $x_1 , x_2 , \ldots , x_n$ حيث (n) يمثل حجم المجموعة ؛ فإن الوسط الهندسي يمكن التعبير عنه على النحو التالي

$$G = \sqrt[n]{\left(X_1 . X_2 . X_3 . X_n\right)} = (X_1 + X_2 + X_3 +,...,+ X_n)^{\frac{1}{n}}$$

$$G = \frac{1}{n} Log\ (X_1 + X_2 + X_3 +,...,+ X_n) = \frac{1}{n} Log\ \sum_{i=1}^{n} X_i$$

مثال(17): ازدادت أرباح مبيعات شركة إنتاج الملابس من 200 ألف دينار عراقي عام 1995 **إلى** 350 ألف دينار **عراقي عام** 2000، ومن ثم فإن المتوسط الهندسي في هذه الفترة هو:

$$G = \sqrt[n]{X_1 F_1 \times X_2 F_2 \times \ldots \times X_n F_n} = \sqrt{200000 \times 350000} = 264575$$

ثانيا :في حالة البيانات المبوبة:-

الوسط الهندسي **لمراكز الفئات** $X_1, X_2, \ldots, X_n$ **والمرجح بالتكرارات المناظرة** يكون F_1, F_2, F_3

$$G = \sqrt[n]{X_1 F_1 \times X_2 F_2 \times \ldots \times X_n F_n}$$

$$LogG = \frac{\sum F\left(LogX\right)}{\sum F}$$

مثل (18) اوجد الوسط الهندسي للبيانات التالية

الفئات	10-0	20-10	30-20	40-30	المجموع
التكرارات	5	8	3	4	20

الحل

الفئات	التكرارات F	مراكز الفئات X_i	Log X	F . Log X
0 – 10	5	5	Log 5 = 0.699	5 x 0.699 = 3.495
10 – 20	8	15	Log 15 = 1.176	15 x 1.176 = 9.408
20 – 30	3	25	Log 25 = 1.397	25 x 1.397 = 4.191
30 – 40	4	35	Log 35 = 1.544	35 x 1.544 = 6.176
Total	**20**			**23.27**

$$LogG = \frac{\sum F\left(LogX\right)}{\sum F} = \frac{23.27}{20} = 1.16$$

$$G = 14.58$$

3-4-2. الوسط التوافقي *Harmonic mean*

هو مقلوب الوسط الحسابي لمقلوب القيم و يتم حسابه وفق الصيغة التالية:

أولا: حسب الوسط التوافقي في حالة البيانات غير المبوبة

$$H = \frac{n}{\sum_{i=1}^{n} \frac{1}{X_i}} \qquad or \qquad \frac{1}{H} = \frac{1}{n} \times \sum_{i=1}^{n} \frac{1}{X}$$

مثال (19) : أوجد الوسط التوافقي للبيانات التالية: 2 , 5 , 3 , 4 , 7 , 8 , 8

$$H = \frac{n}{\sum_{i=1}^{n} \frac{1}{X_i}} = \frac{7}{\frac{1}{2}+\frac{1}{5}+\frac{1}{3}+\frac{1}{4}+\frac{1}{7}+\frac{1}{8}+\frac{1}{8}} = \frac{7}{1.68} = 4.17$$

مثال (20) : المتوسط التوافقي للقيم: 12، 10، 7، 6، 6، 5، 3، هو:

$$\frac{1}{H} = \frac{1}{7}\left(\frac{1}{3} + \frac{1}{5} + \cdots + \frac{1}{12}\right)$$

H= 5.28

ثانيا: حسب الوسط التوافقي في حالة البيانات **المبوبة**

الوسط التوافقي لمراكز الفئات $X_1, X_2, \ldots, X_n$ والمرجح بالتكرارات المناظرة $f_1, f_2, \ldots, f_n$ يكون

$$H = \frac{\sum f_i}{\sum \frac{f_i}{X_i}}$$

مثل (21) احسب الوسط التوافقي من الجدول التالي والذي يوضح التوزيع التكراري لسرعات 100 متسابق

السرعات بالكيلو متر / الساعة	2.5 – 7.5	7.5 – 12.5	12.5 – 17.5	17.5 – 22.5	المجموع
عدد المتسابقين	20	50	20	10	100

الحل

الفئات	**التكرارات** ***f***	**مراكز الفئات** X_i	***1 / X***	***f * 1 / X***
2.5 – 7.5	20	5	1 / 5 = 0.2	20 * 0.2 = 4
7.5 – 12.5	50	10	1 / 10 = 0.1	50 * 0.1 = 5
12.5 – 17.5	20	15	1 / 15 = 0.0.67	20 * 0.067 = 1.34
17.5 – 22.5	10	20	1 / 20 = 0.05	10 * 0.05 = 0.5
Total	**100**			**10.84**

$$H = \frac{\Sigma f_i}{\Sigma \frac{f_i}{X_i}} = \frac{100}{10.84} = 9.25 \quad Km / h$$

الأسئلة

س1/ / ارسم مخطط موضحا فيه اهم مقاييس النزعة المركزية؟

س2/ ماهي مزايا وعيوب الوسط الحسابي ؟

س3/ ماهي مزايا وعيوب الوسيط ؟

س4/ ماهي مزايا وعيوب المنوال ؟

س5/ اكتب ما تعرفه عن الوسط الهندسي والوسط التوافقي ؟

س6/ ماهي العلاقة بين الوسط الحسابي والوسيط والمنوال ؟

س7/ وضح بالرسم طريقة استخدام مقاييس النزعة المركزية في تحديد شكل توزيع البيانات ؟

س8/ البيانات التالية تمثل عدد افراد عينة من الاسر قوامها 12 اسرة ، المطلوب ايجاد متوسط عدد افراد الاسرة (بالطريقة المختصرة والطريقة المباشرة) ، الوسيط ، المنوال

البيانات :

3 , 4 , 7 , 8 , 10 , 9 , 2 , 5 , 6 , 9 , 7 , 5 .

س9/ الاتي توزيع تكراري لعينة من الاسر قوامها 75 اسرة حسب عدد افراد الاسرة، المطلوب حساب متوسط عدد افراد الاسرة في هذه العينة (بالطريقة المباشرة والطريقة غير المباشرة) ، الوسيط والمنوال ثم حدد شكل الالتواء لهذه البيانات

الفئات	2-4	5-7	8-10	11-13	14-16	17-19	20-22	المجموع
التكرارات	8	12	20	13	10	8	4	75

س10/ اوجد الوسط الحسابي ، الوسيط والمنول لدرجات عينة من الطلبة قوامها 9 طلاب في امتحان معين الدرجات :

63 , 55 , 62 , 53 , 70 , 68 , 65 , 79 , 80

س11/ الاتي اعمار عينة من الافراد قوامها 12 فرد جد الوسط الحسابي ، الوسيط والمنوال لعمر الفرد في هذه العينة

20 , 22 , 19.5 , 26 , 24 .5 , 27 , 28 , 29 , 18 , 20 , 23 , 25

س**12**/ الاتي توزيع تكراري للدخل الشهري لعينة من الاسر قوامها 80 **لسرة** ، المطلوب ايجاد **الوسط الحسابي** ، الوسيط والمنوال للدخل الشهري **للأسرة** في هذه العينة (للمتجمع الصاعد و للمتجمع النازل)

الفئات	100-120	120-140	140-160	160-180	180-200	200-220	220-240	**المجموع**
التكرارات	3	7	14	20	18	12	6	80

س**13**/ الاتي توزيع تكراري **لأطول** عينة من الاشخاص البالغين قوامها 50**شخص ، المطلوب حسب الوسط** الحسابي ، الوسيط **و**المنوال لطول الشخص في هذه العينة بطريقة بيرسون

الفئات	150-160	160-170	170-180	180-190	190-200	**المجموع**
التكرارات	8	12	15	9	6	50

س14/ إذا كان لدينا التوزيع التكراري التالي اوجد الوسط الحسابي والوسيط والمنوال **، ثم حددشكل الالتواء** لهذه البيانات؟

الفئات	17-15	20-18	23-21	26-24	29-27	32-30	**المجموع**
التكرار	5	9	13	11	8	4	50

س15/احسب الوسط الهندسي و الوسط التوافقي للقيم التالية

2 , 12 , 5 , 10 , 4 , 3 , 7

س16/احسب الوسط الهندسي و الوسط التوافقي للقيم التالية

الفئات	20-14	26-20	32-26	38-32	44-38	50-44	**المجموع**
التكرار	7	9	3	10	7	4	40

الفصل الرابع

مقاييس التشتت

4- مقاييس التشتت

مقاييس التشتت : -هي تباعد او انتشار قيم مجموعة من المفردات عن بعضها البعض او عن قيمة معينة ثابتة (مثل الوسط الحسابي)وتستخدم لغرض اجراء المقارنة بين مجموعتين او اكثر من البيانات عن ظاهرة معينة.

مثل(1) الوسط الحسابي للمجموعات الثلاثة التالية يساوي 9

المجموعة الاولى :القيم 7, 8 , 9 , 10 , 11.

الوسط الحسابي

$$\ddot{X} = \frac{\sum x}{n} = \frac{7 + 8 + 9 + 10 + 11}{5} = \frac{45}{5} = 9$$

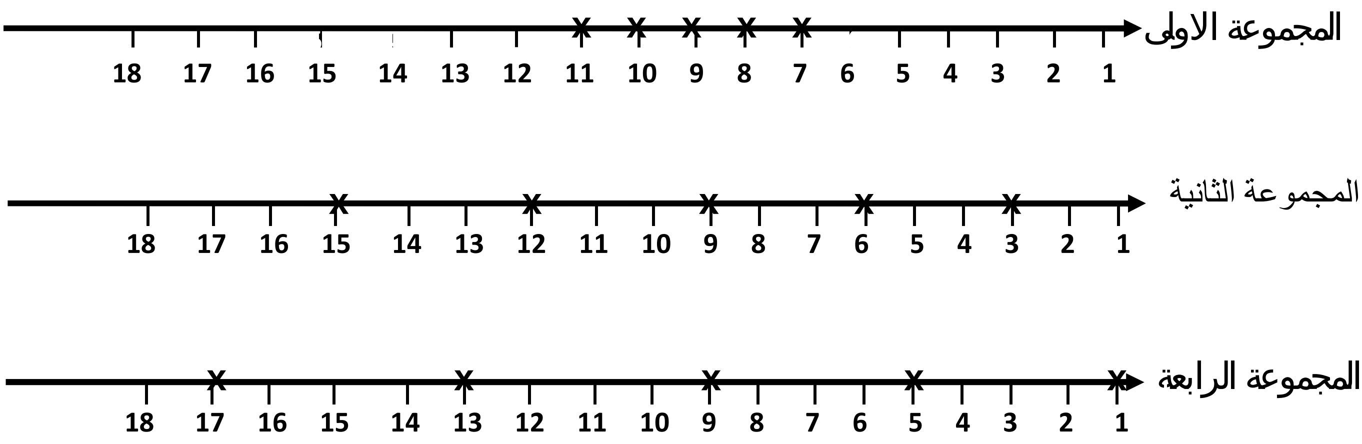
المجموعة الثانية : القيم 3, 6, 9, 12, 15

$$\ddot{X} = \frac{\sum x}{n} = \frac{3 + 8 + 9 + 10 + 11}{5} = \frac{45}{5} = 9$$

المجموعة الثالثة : القيم 1 , 5 , 9 , 13 , 17

$$\ddot{X} = \frac{\sum x}{n} = \frac{1 + 5 + 9 + 13 + 17}{5} = \frac{45}{5} = 9$$

عند المقارنة بين هذه القيم نلاحظ ان المجموعة الاولى اكثر تجانسا من المجاميع الاخرى وكما مبين بالرسم



4-1. مقاييس التشتت

وهناك نوعين من مقاييس التشتت وهي : -

1- مقاييس التشتت المطلقة
2- مقاييس التشتت النسبية

4-1-1. مقاييس التشتت المطلقة : - هي مقاييس تبين درجة تجانس قيم مجموعة من البيانات بشكل مطلق وتكون مقاسة بنفس وحدات قياس المتغير العشوائي (وحدات ،طول، وزن، عدد....الخ) .

4-1-2. مقاييس التشتت النسبية: إذا كنا بصدد إجراء مقارنة بين توزيعات تكرارية ليست لها نفس وحدة القياس أو ليس لهما نفس المتوسط، فمن الضروري هنا استخدام مقاييس التشتت النسبية .

4-2. اهم مقاييس التشتت

ا - المدى Range

2 – الانحراف الربيعي Quartile Deviation

3 – التباين Variance

4 – الانحراف المعياري Standard Deviation

5 - معامل الاختلاف Coefficient of Variation

4-2-1. المـــدى :

يسمى احيانا بمجال التغير وهو من ابسط مقاييس التشتت المطلق ويعرف بأنه الفرق بين اعلى قيمة واقل قيمة من مجموعة البيانات غير المبوبة. فإذا كانت x_L تمثل اعلى قيمة وان x_s تمثل ادنى قيمة فان المدى يحسب وفق الصيغة التالية :

المدى المطلق = أكبر قيمة – أصغر قيمة

$$R = X_L - X_S$$

اما في حالة البيانات المبوبة فان المدى : عبارة عن الفرق بين الحد الادنى للفئة الاولى والحد الاعلى للفئة الاخيرة فاذا كان الحد الاعلى U والحد الادنى L فان المدى

المدى = الحد الاعلى لأكبر الفئات – الحد الأدنى لأصغر الفئات

$$R = U - L$$

مثل (1)**الاتي** بيانات غير مبوبة ، المطلوب ايجاد المدى لهذه البيانات

2 , 5 , 3 , 8 , 7 , 10 , 9 , 12 , 15

$$R = X_L - X_S = 15 - 2 = 13$$

مثل (2) بيانات مبوبة

من خلال جدول التوزيع التكراري التالي اوجد المدى

المجموع	40-48	32-40	24-32	16-24	8-16	الفئات
62	3	6	8	25	20	التكرارات

الحل

$$R = U - L = 48 - 8 = 40$$

لا يستخدم المدى كثيرا لأنه يستند الى قيمتين الاولى والاخيرة ويهمل باقي القيم وهذا يعني انه مقياس حساس جدا لأي خطأ قد يحصل في قياس احدى هاتين القيمتين او كليهما كما لا يمكن حسابه في حالة الجداول التكرارية المفتوحة.

4-2-1-1. مزايا وعيوب المدى

أولا:- مزايا المدى

1 - أبسط وأسهل طريقة لحساب التشتت

2 - مقياس سريع لمدى التشتت المفردات أو حينما يكون للمفردات المتطرفة أهمية خاصة.

ثانيا :عيوب المدى :

1- ليس للمدى أهمية كبيرة في البحوث العلمية نظرا لأنه لا يأخذ في الاعتبار تشتت كل المفردات في حسابه.

2 - مقياس تقريبي غير دقيق

3 - يتأثر تأثرا كبيرا بالقيم المتطرفة

4 - يصعب تقدير قيمته من الجداول التكرارية المفتوحة.

4-3. الانحراف الربيعي

من اهم عيوب المدى هو اعتماده على القيمتين الاولى والاخيرة التي غالبا ما تكون شاذة (متطرفة) وبهدف التغلب على هذا العيب نقوم بحذف بعض القيم الشاذة فاذا اهملنا الربع الاول والربع الاخير من هذه القيم فانه يمكن

الحصول على مقياس تشتت يعتبر افضل من المدى ويعتمد في حسابه على كل من الربعين الادنى (الاول) والاعلى (الثالث) ويسمى بالانحراف الربيعي (نصف المدى الربيعي) .

ويعرف **الانحراف الربيعي** بانه متوسط الفرق بين الربيع الثالث والربيع الاول لمجموعة من البيانات سواء كانت مبوبة او غير مبوبة . فإذا رمزنا للربيع الاول(الادنى) Q_1 وللربيع الثالث بالرمز Q_3 وللانحراف الربيعي $Q.D$ وعليه فان الانحراف الربيعي:

$$Q.D = \frac{Q_3 - Q_1}{2}$$

وطريقة حساب الربيعين الادنى والاعلى هي تماما كطريقة حساب الوسيط

4-3-1. حساب الربيعين من بيانات غير مبوبة

لحساب قيمة الربيع الاعلى والربيع الادنى يجب تحديد ترتيب (موقع) كل منهما مسبقا مثلما فعلنا عند حساب قيمة الوسيط .

ترتيب (موقع) الربيع الادنى = عدد القيم / 4

وعليه فان :

$$T.Q_1 = \frac{n}{4}$$

حيث ان $T.Q_1$ تمثل ترتيب الربيع الاول وان n هي عدد القيم

وترتيب الربيع الثالث (الاعلى) = (عدد القيم / 4) * 3

$$T.Q_3 = \frac{n}{4} \times 3$$

اي على بعد 75% من بداية البيانات ويجب هنا ايضا ان يعاد ترتيب

مجموعة القيم تصاعديا او تنازليا قبل حساب موقع او ترتيب الربيعين

مثل (1) احسب الربيعين الاعلى والادنى والانحراف الربيعي لمجموعة البيانات الاتية :

15 , 10 , 12 , 8 , 2 , 5 , ,7 , 3

الحل :

1- نرتب القيم ترتيبا تصاعديا او ترتيبا تنازليا

ترتيب تصاعدي

2 , **3** ,5 , 7 , 8 , **10** , 12 , 15

2- نجد ترتيب الربيع الادنى (الاول)

أي ان ترتيب الربيع الاول هو الترتيب

$$T.Q_1 = \frac{n}{4} = \frac{8}{4} = 2$$

الثاني من بين البيانات ويساوي (3) **اي ان**

$$Q_1 = 3$$

3- نحدد ترتيب الربيع الاعلى (الثالث)

$$T.Q_3 = \frac{n}{4} \times 3 = \frac{8}{4} \times 3 = 6$$

بمعنى ان موقع الربيع الثالث هو الترتيب السادس من البيانات ويساوي 10 **اي ان**

$$Q_3 = 10$$

4- نجد الانحراف الربيعي والذي يمثل متوسط الفرق بين الربيع الاعلى والربيع الادنى

$$Q.D = \frac{Q_3 - Q_1}{2} = \frac{10 - 3}{2} = 3.5$$

4-3-2. الانحراف الربيعي للبيانات المبوبة حسابيا وبيانيا

طريقة حساب الربيع الاعلى والادنى من بيانات مبوبة (جدول تكراري) تماثل تماما طريقة حساب الوسيط السابق شرحها .

خطوات ايجاد الربيعين لبيانات مبوبة كالاتي:

1. **اعداد جدول تكراري تجمعي صاعد او نازل**
2. تحديد ترتيب الربيع الادنى وذلك بتطبيق الصيغة الاتية :

$T.Q_1 = \frac{\sum f_i}{4}$ حيث ان **∑fi** هي مجموع التكرارات

3. تحديد ترتيب الربيع الاعلى بالصيغة الاتية :

$$T.Q_3 = \left(\frac{\sum f_i}{4}\right) \times 3$$

4. حساب قيمة كل من الربيعين من جدول تكراري متجمع صاعد او متجمع نازل وباستخدام **القانون الاتي:**

أ- قيمة الربيع الادنى

$$Q_1 = L_1 + \frac{T.Q_1 - f_{K-1}}{f_K} \times h_k$$

حيث ان

الربيع الادنى Q_1

بداية فئة الربيع الادنى L_1

ترتيب الربيع الادنى $T.Q_1$ ويساوي $T.Q_1 = \frac{\sum f_i}{4}$

التكرار المتجمع السابق الادنى f_{k-1}

الفرق بين التكرارين الصاعد السابق واللاحق f_k وهو نفسه التكرار الاصلي لفئة الربيع الادنى

طول فئة الربيع الادنى h_k

ب- قيمة الربيع الاعلى

والقانون يكون بالصيغة الاتية :

$$Q_3 = L_3 + \frac{T.Q_3 - f_{K-1}}{f_K} \times h_k$$

حيث ان

الربيع الاعلى Q_3

بداية فئة الربيع الاعلى L_3

ترتيب الربيع الاعلى $T.Q_3$ ويساوي $T.Q_3 = \frac{\sum f_i}{4} \times 3$

التكرار المتجمع السابق الاعلى f_{k-1}

الفرق بين التكرارين الصاعد f_k ويعتبر التكرار الاصلي لفئة الربيع **الأعلى (السابق واللاحق)**

طول فئة الربيع الاعلى h_k

5. حساب الانحراف الربيعي وبالصيغة الاتية :

$$Q.D = \frac{Q_3 - Q_1}{2}$$

مثال : احسب الانحراف الربيعي للتوزيع التكراري الاتي لدرجات 50 من الطلبة في احدى السنوات

الفئات	0-10	10-20	20-30	30-40	40-50	المجموع
التكرارات	5	8	15	16	6	50

الحل :

لحساب قيمتي الربيع الاعلى والادنى يلزم اعداد جدول متجمع صاعد او متجمع نازل

نفرض اننا نعد جدول متجمع صاعد

الفئات	التكرارات	الحدود العليا للفئات	التكرار المتجمع الصاعد
0-10	5	اقل من 10	5
10-20	8	اقل من 20	13
20-30	15	اقل من 30	28
30-40	16	اقل من 40	44
40-50	6	اقل من 50	50
المجموع	50		

F_{k-1}

$T.Q_1$

$T.Q_3$

نجد ترتيب الربيع الادنى

$$T.Q_1 = \frac{\sum f_i}{4} = \frac{50}{4} = 12.5$$

1- نجد قيمة الربيع الأدنى

$$L_1 = 10 \quad ; \quad T.Q_1 = 12.5 \quad ; \quad f_{k-1} = 5 \quad ; \quad h_k = 10 \quad , \quad f_k = 13$$

$$Q_1 = L_1 + \frac{T.Q_1 - f_{K-1}}{f_K} \times h_k$$

$$Q_1 = 10 + \frac{12.5 - 5}{13 - 5} \times 10 = 10 + \frac{7.5 \times 10}{8} = 19.375$$

2- ثم نجد ترتيب الربيع الاعلى

$$T.Q_3 = \frac{\sum f_i}{4} \times 3 = \frac{50}{4} \times 3 = 37.5$$

ايجاد قيمة الربع الأعلى

$$Q_3 = L_3 + \frac{T.Q_3 - f_{K-1}}{f_K} \times h_k$$

$$Q_3 = 30 + \frac{37.5 - 28}{44 - 28} \times 10 = 30 + \frac{9.5 \times 10}{16} = 35.937$$

ايجاد الانحراف الربيعي

$$Q.D = \frac{Q_3 - Q_1}{2} = \frac{35.937 - 19.375}{2} - 8.2812$$

4-3-3. استخراج الانحراف الربيعي الادنى والاعلى بالرسم البياني

من الممكن ايجاد قيمة الانحراف الربيعي بالرسم وذلك باستخراج قيمتي الربيعين بيانيا وكالاتي:

1. نستخرج من الجدول الاصلي جدولا تكراريا متجمعا صاعدا او نازلا
2. نرسم المنحنى التكراري للمتجمع الصاعد او النازل
3. نعين نقطتي ترتيب كل من الربيع الادنى والربيع الاعلى على المحور العمودي
4. نرسم من كل من هاتين النقطتين مستقيما موازيا للمحور الافقي ثم نسقط من نقطتي التقائهما مع المنحنى عمودين على المحور الافقي فتكون نقطتا تلاقي هذين العمودين مع المحور الافقي مساويتين لقيمتي الربيع الادنى والربيع الاعلى على التوالي

مثال :

الجدول الاتي يبين توزيع الاجور في احد المصانع ، المطلوب : ايجاد الانحراف الربيعي بالرسم البياني

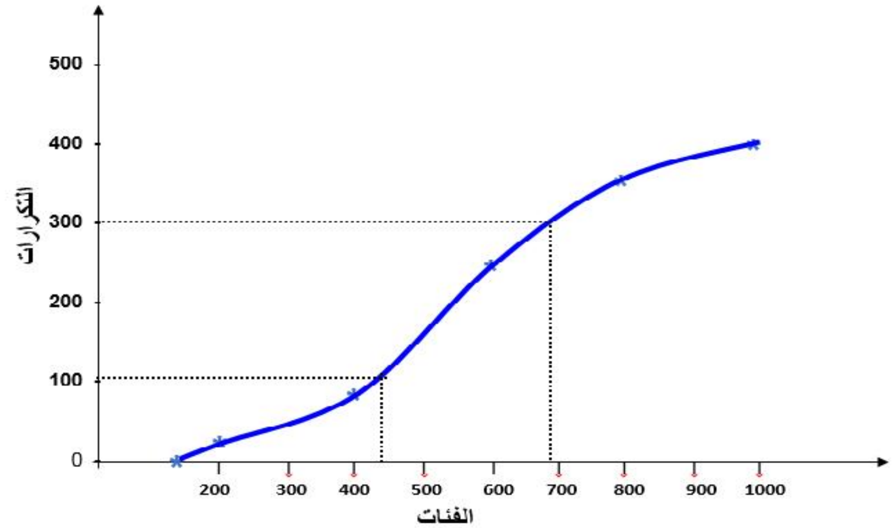
الفئات	0-200	200-400	400-600	600-800	800-1000	المجموع
التكرارات	18	72	154	111	45	400

الحل :

نعمل جدولا تكراريا متجمعا صاعدا

نرسم المنحنى المتجمع الصاعد بأخذ الحدود العليا والتكرارات المتجمعة

الحدود العليا للفئات	التكرار المتجمع الصاعد
اقل من 100	0
اقل من 200	18
اقل من 400	90
اقل من 600	244
اقل من 800	355
اقل من 1000	400



من الرسم نلاحظ ان قيمة الربيع الادنى (الاول) هو تقريبا 410 ، وان قيمة الربيع الاعلى (الثالث) هو 700 وان الانحراف الربيعي

$$Q.D = \frac{Q_3 - Q_1}{2} = \frac{700 - 410}{2} = 145$$

والانحراف الربيعي يسمى ايضا بنصف المدى الربيعي لأنه يساوي نصف المدى بين الربيع الثالث والربيع الاول وكذلك هذا المقياس يتوقف على قيمتين فقط من قيم التوزيع هما قيمتين الربيعين الاول والثالث ولهذا فانه يتأثر بتغير العينة ولكنه افضل من المدى لأنه **لا** يتأثر بالقيم المتطرفة ويمكن استخراجه من الجداول المفتوحة عندما يراد معرفة درجة تركز القيم حول الوسيط.

4-3-4. مزايا وعيوب الانحراف الربيعي

من مزايا الانحراف الربيعي، يفضل استخدامه كمقياس للتشتت في حالة وجود قيم شاذة ، كما أنه بسيط وسهل في الحساب . ومن عيوبه ، أنه لا يأخذ كل القيم في الاعتبار.

4-4. التباين Variance

هو أحد مقاييس التشتت ، وأكثرها استخداما في النواحي التطبيقية ، **وهو** مجموع مربعات انحرافات القيم عن وسطها الحسابي مقسوما على عددها , والذي يرمز له **بالرمز** S^2.

4-4-1. التباين في المجتمع ($\pm^2$)

إذا توافر لدينا قراءات عن كل مفردات المجتمع ، ولتكن: $x_1, x_2, \ldots, x_N$ ، فإن التباين في المجتمع ، ويرمز له بالرمز $\pm^2$(سيكما) يحسب باستخدام المعادلة التالية :

$$\pm^{2} = \frac{\sum \left(x - «\right)^{2}}{N}$$

حيث أن » هو الوسط الحسابي في المجتمع ، أي أن : $« = \sum x/N$.

مثال(1)

مصنع لتعبئة المواد الغذائية ، يعمل به 15 عامل ، وكانت عدد سنوات الخبرة لهؤلاء العمال كما يلي :

5 13 7 14 12 9 6 8 10 13 14 6 11 12 10

بفرض أن هذه البيانات تم جمعها عن كل مفردات المجتمع ، فأوجد التباين لعدد سنوات الخبرة .

الحل

لحساب تباين سنوات الخبرة في المجتمع ، يتم استخدام المعادلة

$$\pm^{2} = \frac{\sum \left(x - «\right)^{2}}{N}$$

- الوسط الحسابي في المجتمع »

$$\begin{aligned} « &= \frac{1}{N}\sum x \\ &= \frac{1}{15}(5+13++7+\ldots+12+10) = \frac{1}{15}(150) = 10 \end{aligned}$$

- حساب مربعات الانحرافات $\sum(x-«)^2$

سنوات الخبرة x	$(x-\text{«})$	$(x-\text{«})^2$
5	5-10 = -5	25
13	3	9
7	-3	9
14	4	16
12	2	4
9	-1	1
6	-4	16
8	-2	4
10	0	0
13	3	9
14	4	16
6	-4	16
11	1	1
12	2	4
10	0	0
150	0	130

بما أن: $\sum(x-\text{«})^2 = 130$

إذا تباين سنوات الخبرة للعمال في المصنع هو :

$$\pm\ ^2 = \frac{\sum (x-u)^2}{N} = \frac{130}{15} = 8.67$$

4-4-2. التباين في العينة (s^2)

في كثير من الحالات يكون تباين المجتمع $\pm\ ^2$ غير معلوم، وعندئذ يتم سحب عينة من هذا المجتمع ، ويحسب التباين من بيانات العينة كتقدير لتباين المجتمع ، فإذا كانت قراءات عينة عشوائية حجمها n هي ، $x_1, x_2, \ldots, x_n$ ، فإن تباين العينة ويرمز له بالرمز s^2 هو:

$$S^2 = \frac{\sum (x-\bar{x})^2}{n-1}$$

حيث أن $\bar{x}$ هو الوسط الحسابي لقراءات العينة ، أي أن : $\bar{x} = \sum x/n$ ، وتباين العينة المبين بالمعادلة اعلاه هو التقدير غير المتحيز لتباين المجتمع .

مثل(2)

إذا تم سحب عينة من عمال مصنع حجمها 5 **عمل** ، **وسجل عددسنوات الخبرة** ، **وكانت كالتالي** .

8 13 10 5 9

احسب تباين سنوات الخبرة في العينة .

الحل

لحساب التباين في العينة يتم تطبيق المعادلة ويتبع الآتي :

$$S^2 = \frac{\sum (x-\bar{x})^2}{n-1}$$

- الوسط الحسابي في العينة :

$$\bar{x} = \frac{1}{n}\sum x = \frac{1}{5}(8+13+10+5+9) = \frac{1}{5}(45) = 9$$

- **حساب مربعات الانحرافات** $\sum (x-\bar{x})^2$

سنوات الخبرة x	8	13	10	5	9	**45**
$(x-\bar{x})$	-1	4	1	-4	0	**0**
$(x-\bar{x})^2$	1	16	1	16	0	**34**

أي أن : $\sum (x-\bar{x})^2 = 34$ ،

- إذا تباين سنوات الخبرة في العينة قيمته هي :

$$s^2 = \frac{\sum (x-\bar{x})^2}{n-1} = \frac{34}{(5-1)} = \frac{34}{4} = 8.5$$

- في هذه الحالة يمكن القول بأن تباين العينة 8.5 ، وهو في نفس الوقت تقدير غير متحيز لتباين المجتمع .

5-4. الانحراف المعياري Standard Deviation

عند استخدام التباين كمقياس من مقاييس التشتت، نجد أنه يعتمد على **مجموع مربعات الانحرافات**، ومن ثم لا يتمشى هذا المقياس مع وحدات قياس المتغير محل الدراسة ، ففي المثال السابق ، نجد أن تباين سنوات الخبرة في العينة 8.5، فليس من المنطق عند تفسير هذه النتيجة أن نقول ، " تباين سنوات الخبرة هو 8.5 سنة تربيع "، لأن وحدات قياس المتغير هو عدد السنوات، من أجل ذلك لجأ الإحصائيين إلى مقياس منطقي يأخذ في الاعتبار الجذر التربيعي للتباين ، لكي يناسب وحدات قياس المتغير، وهذا المقياس هو الانحراف المعياري.

إذا الانحراف المعياري ، هو الجذر التربيعي الموجب للتباين ، أي أن:

$$S\tan dard \quad Deviation = \sqrt{Variance}$$

مثل(1)

مصنع لتعبئة المواد الغذائية ، يعمل به 15 عامل ، وكانت عدد سنوات الخبرة لهؤلاء العمال كما يلي :

5 13 7 14 12 9 6 8 10 13 14 6 11 12 10

بفرض أن هذه البيانات تم جمعها عن كل مفردات المجتمع ، فأوجد الانحراف المعياري لسنوات الخبرة لعمال المصنع (المجتمع) ، ويرمز له بالرمز (±) هو :

$$\pm = \sqrt{\frac{1}{N}\sum x^2 - «^2}$$

$$= \sqrt{\frac{1}{15}1630 - 10^2} = \sqrt{8.67} = 2.94$$

في هذه الحالة ، يكون الانحراف المعياري لسنوات الخبرة في المجتمع هو 2.94 **سنة** .

والسبب في اخذ الجذر التربيعي هو **لأجل** ان يكون مقياس التشتت (الانحراف المعياري) مقاسا بنفس وحدات القيم الاصلية فقد قمنا بتربيع الانحرافات ولكي نرجع الى الوحدات الاصلية بعد التربيع لابد ان **نأخذ الجذر** التربيعي .

4-5-1. الانحراف المعياري لبيانات غير مبوبة :- هناك طرقيتين :

أ - الطريقة المطولة باستخدام الانحرافات عن الوسط الحسابي وحسب الصيغة الاتية :

$$S = \sqrt{\frac{\left(X - \ddot{X}\right)^2}{n}}$$

مثال (1) :

البيانات التالية **تمثل** اوزان عينة من الطلبة قوامها 10**طلاب ،** المطلوب حساب قيمة الانحراف المعياري

البيانات بالطريقة المطولة : 56, 68, 72, 63, 65, 68, 71, 69, 62, 56,

الحل :

1- **نحسب** الوسط الحسابي لهذه البيانات

$$\ddot{X} = \frac{56 + 68 + 72 + 63 + 65 + 68 + 71 + 69 + 56}{10} = 65$$

2- نجد انحرافات القيم عن الوسط الحسابي ثم نجد مجموع مربعات الانحرافات

X	$\left(X - \ddot{X}\right)$	$\left(X - \ddot{X}\right)^2$
56	56 – 65 = -9	81
68	68 – 65 = 3	9
72	72 – 65 = 7	49
63	63 – 65 = -2	4
65	**65 – 65 = 0**	**0**
68	68 – 65 = 3	9
71	71 – 65 = 6	36
69	69 – 65 = 4	16
62	62 – 65 = -3	9
56	56 – 65 = -9	81
Total		294

3 - نطبق صيغة القانون

$$S = \sqrt{\frac{\left(X - \dot{\ddot{X}}\right)^2}{n}}$$

$$S = \sqrt{\frac{\left(X - \dot{\ddot{X}}\right)^2}{n}} = \sqrt{\frac{294}{10}} = \sqrt{29.4} = 5.422$$

ب – الطريقة المختصرة (بدون استخدام الوسط الحسابي) وذلك باستخدام الصيغة الاتية :

$$S = \sqrt{\frac{\sum X^2 - \frac{\left(\sum X\right)^2}{n}}{n}}$$

مثال (2) :

البيانات التالية **تمثل** اوزان عينة من الطلبة قوامها 10**طلاب** ، المطلوب حساب قيمة الانحراف المعياري بالطريقة المختصرة : 56, 62, 69, 71, 68, 65, 63, 72, 68, 56,

خطوات الحل : 1- **نجد مجموع القيم** ، 2- نجد مجموع مربعات القيم ، 3- نطبق صيغة القانون

X	X^2
56	3136
68	4624
72	5184
63	3969
65	4225
68	4624
71	5041
69	4761
62	3844
56	3136
$\sum X = 650$	$\sum X^2 = 42544$

$$S=\sqrt{\frac{\sum X^2-\frac{\left(\sum X\right)^2}{n}}{n}}=\sqrt{\frac{42544-\frac{(650)^2}{10}}{10}}=\sqrt{\frac{294}{10}}=\sqrt{29.4}=5.422$$

4-5-2. حساب الانحراف المعياري لبيانات مبوبة

1- **الطريقة المطولة** : (حساب انحرافات مراكز الفئات عن الوسط الحسابي) باستخدام الصيغة الاتية

$$S=\sqrt{\frac{\sum f_i\times\left(X_i-\ddot{X}\right)^2}{\sum f_i}}$$

حيث ان $\sum$**fi** تعني مجموع التكرارات

(Xi-Ẍ) انحرافات القيم عن الوسط الحسابي

خطوات الحل :

1- نجد مراكز الفئات ***Xi***

2- نجد الوسط الحسابي باستخدام الصيغة $\ddot{X}=\frac{\sum f_i X_i}{\sum f_i}$

3- نجد انحرافات مراكز الفئات عن الوسط الحسابي $\left(X_i-\ddot{X}\right)$

2- نربع انحرافات مراكز الفئات عن الوسط الحسابي

3- نضرب التكرارات في مربعات الانحراف

4- نجمع حاصل الضرب

5- نطبق صيغة القانون

مثال (1)

الاتي جدول توزيع تكراري لدرجات مجموعة من الطلبة في امتحان فصلي ، المطلوب ايجاد الانحراف المعياري بالطريقة المطولة

الفئات	0-2	2-4	4-6	6-8	8-10	المجموع
التكرارات	5	10	20	10	5	50

الحل :

الخطوة (2) نجد الوسط الحسابي ويساوي

$$\ddot{X} = \frac{\sum f_i X_i}{\sum f_i} = \frac{250}{50} = 5$$

نطبق الخطوات السابقة عدا الخطوة الخامسة

الفئات	Fi	Xi	Fi Xi	$(Xi - \ddot{X})$	$(Xi - \ddot{X})^2$	$Fi(Xi - \ddot{X})^2$
0 -2	5	1	5*1=5	1 -5= - 4	4^2=16	5*16=80
2 -4	10	3	30	-2	4	40
4 -6	20	5	100	0	0	0
6 -8	10	7	70	2	4	40
8 -10	5	9	45	4	16	80
$\sum$	50		250		40	240

الخطوة (5)

$$S = \sqrt{\frac{\sum f_i \times \left(X_i - \ddot{X}\right)^2}{\sum f_i}} = \sqrt{\frac{240}{50}} = \sqrt{4.8} = 2.19$$

ب - الطريقة المختصرة نستخدم الصيغة الاتية :

$$S = \sqrt{\frac{\sum f_i X_i^2}{\sum f_i} - \frac{\left(\sum f_i X_i\right)^2}{\sum f_i}} = \sqrt{\frac{\sum f_i X_i^2 - \frac{\left(\sum f_i X_i\right)^2}{\sum f_i}}{\sum f_i}}$$

يلاحظ ان هذه الصيغة هي نفس الصيغة لبيانات غير مبوبة (الطريقة المختصرة) الا اننا ادخلنا التكرارات

مثال (2)

الاتي جدول توزيع تكراري لدرجات مجموعة من الطلبة في امتحان فصلي ، المطلوب ايجاد الانحراف المعياري بالطريقة المختصرة

الفئات	0-2	2-4	4-6	6-8	8-10	المجموع
التكرارات	5	10	20	10	5	50

الحل :

1- نجد مراكز الفئات
2- ضرب مركز كل فئة بالتكرار المقابل له
3- تربيع مراكز الفئات
4- ضرب التكرار في مربع مركز الفئة المقابل له

5- تطبيق صيغة القانون

الحل: الخطـــــــــوات

الفئات	f_i	(1) X_i	(2) $f_i X_i$	(3) X_i^2	(4) $f_i X_i^2$
0 -2	5	1	5*1=5	$1^2=1$	5*1=5
2-4	10	3	30	9	10*9=90
4 -6	20	5	100	25	500
6 -8	10	7	70	49	490
8 -10	5	9	45	81	405
∑	50		250		1490

الخطوة (5)

$$S = \sqrt{\frac{\sum f_i X_i^2 - \frac{\left(\sum f_i X_i\right)^2}{\sum f_i}}{\sum f_i}} = \sqrt{\frac{1490 - \frac{(250)^2}{50}}{50}}$$

$$= \sqrt{\frac{1490 - 1250}{50}} = \sqrt{\frac{240}{50}} = \sqrt{4.8} = 2.19$$

4-5-3. مميزات وعيوب الانحراف المعياري

1- مميزات الانحراف المعياري

1- إن حسابه يعتمد على كافة البيانات المتاحة

2- انه مقياس سهل الفهم والحساب

3- خضوعه للعمليات الجبرية

4- قابليته للتجزئة والاندماج

2- عيوب الانحراف المعياري

1- لا يمكن حساب قيمته في حالة التوزيعات التكرارية المفتوحة من طرف واحد او طرفين

2- لا يمكن حسابه في حالة البيانات الوصفية

3- تتأثر قيمته في حالة وجود قيم شاذة أو متطرفة **وكذلك** يتأثر وعلى نحو كبير بأخطاء المعاين.

6-4. **معامل الاختلاف** المعياري **Coefficient of Variation**

وهو من أدق مقاييس التشتت النسبية ، لذلك فهو يصلح للمقارنة بين التوزيعات المختلفة ، ويتم حسابه على أساس (**قسمة الانحراف** المعياري **على الوسط الحسابي**)×100 ويرمز له بالرمز (C .V)

وهو مقياس لا يعتمد على الوحدات ويعطى بالعلاقة التالية :

$$C.V = \frac{S}{\ddot{X}}$$

او

$$C.V = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

مثال

احسب معامل الاختلاف بطريقتين **لدرجات الطلاب الموضحة في الجدول لآتي**

الفئات	40 - 49	50 - 59	60 - 69	70 - 79	80 - 89	90 - 99	**المجموع**
التكرارات	2	9	15	11	2	1	40

الحل :

الطريقة الأولى

$$C.V = \frac{S}{\ddot{X}}$$

$$S = 10.67 \qquad ; \qquad \ddot{X} = 65.75$$

$$C.V = \frac{S}{\ddot{X}} = \frac{10.67}{65.75} = 0.167$$

الطريقة الثانية

$$C.V = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

$$Q_1 = 58.39 \quad ; \quad Q_3 = 73.14$$

$$C.V = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{73.14 - 58.39}{73.14 + 58.39} = 0.111$$

4-7. مقاييس الشكل Measures of Shape

ومن أهم هذه المقاييس ما يلى:

4-7-1. الالتواء Skewness

وهو مقياس تشتت نسبى، ويحدد طبيعة البيانات من حيث التماثل Symmetric **أو الالتواء**

4-7-1-1. طريقة "بيرسون Person" في قياس الالتواء

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:

المنول = الوسط الحسابي – 3 (الوسط الحسابي – الوسيط)

ومن ثم فإن طريقة "بيرسون" في قياس الالتواء ، تتحدد بالمعادلة التالية.

$$\sim = \frac{3(Mean - Median)}{S\tan dard\ Deviation} = \frac{3(\overline{X} - Med)}{S}$$

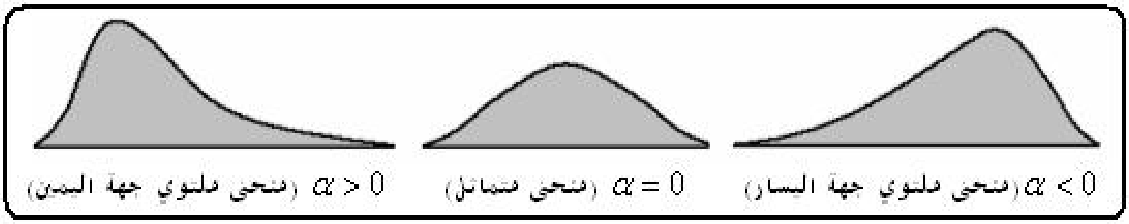
حيث أن ~ (**ألفا**) هو معامل الالتواء "لبيرسون"، $\overline{X}$ **الوسط الحسابي**، Med هو الوسيط، S هو الانحراف المعياري، ويمكن من خلال الإشارة التي يأخذها هذا المعامل الحكم على شكل الالتواء، كما يلي :

1- إذا كان (الوسط الحسابي = الوسيط) كان قيمة المعامل $(\sim = 0)$، ويدل ذلك على أن منحنى التوزيع **التكراري متماثل.**

2- إذا كان (الوسط الحسابي > الوسيط) كان قيمة المعامل $(\sim > 0)$، ويدل ذلك على أن منحنى التوزيع التكراري ملتوي جهة اليمين .

3- إذا كان (الوسط الحسابي < الوسيط) كان قيمة المعامل $(\sim < 0)$، ويدل ذلك على أن منحنى التوزيع التكراري ملتوي جهة اليسار.

أشكال التواء البيانات



2-1-7-4. طريقة "المئين" في قياس الالتواء

المئين ينتج من ترتيب البيانات تصاعديا، ثم تقسيمها البيانات إلى 100 جزء، يفصل بينها قيم تسمى المئين، وعلى سبيل المثال يعرف المئين 15 ويرمز له بالرمز (v_{15}) على أنه القيمة التي يقل عنها 15% من القيم، ولحساب قيمة المئين p ، ونرمز له بالرمز (v_p) ، يتبع نفس الفكرة المستخدمة في حساب الربيع كما يلي:

1. ترتب القيم تصاعديا: $x_{(1)} < x_{(2)} < \ldots < x_{(n)}$
2. رتبة المئين: $R = (n+1)\left(\frac{p}{100}\right)$.
3. **إذا كانت الرتبة** R عدد صحيح فإن $(v_{15} = x_{(R)})$.
4. **أما إذا كانت الرتبة** R عدد كسري فإن قيمة المئين (v_p) تحسب بالمعادلة التالية:

$$\neg_{p} = x_l + (R-l)(x_u - x_l)$$

وتعتمد فكرة المئين في قياس الالتواء على مدى قرب المئين v_p ، والمئين v_{100-p} ، من المئين v_{50} ، وكمثال على ذلك ، عند قياس الالتواء باستخدام المئين 20 ، والمئين 80 ، يلاحظ على الرسم التالي حالات الالتواء :

ومن الشكل أعلاه يلاحظ الآتي :

1. إذا كان بعد المئين (v_{80}) عن المئين (v_{50}) يساوي بعد المئين (v_{20}) عن المئين (v_{50}) كان

التوزيع متماثلا .

2. إذا كان بعد المئين (v_{80}) عن المئين (v_{50}) أكبر من بعد المئين (v_{20}) عن المئين (v_{50}) **كان** التوزيع موجب الالتواء .

3. إذا كان بعد المئين (v_{80}) عن المئين (v_{50}) أقل من بعد المئين (v_{20}) عن المئين (v_{50}) **كان** التوزيع سالب الالتواء .

وبشكل عام يمكن الحكم على شكل التوزيع باستخدام معامل الالتواء المئيني، ويأخذ المعادلة التالية.

$$\sim_{p,100-p} = \frac{(\neg_{100-p} - \neg_{50}) - (\neg_{50} - \neg_{p})}{\neg_{100-p} - \neg_{p}}$$

حيث أن : $v_p < v_{50} < v_{100-p}$ ويفضل استخدام هذا المعامل في حالة البيانات التي تحتوي على قيم شاذة ، وأيضا البيانات التي لا نعرف لها توزيع محدد، وعندما نستخدم المئين 25 $(v_{25} = Q_1)$، المئين 75 $(v_{75} = Q_3)$ نحصل على معامل الالتواء الربيعي ، وهو :

$$\neg_q = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{Q_3 - Q_1}$$

مثل

كانت درجات 8 طلاب في الاختبار النهائي في **مادة إحصاء ، كالتالي.**

66 85 52 78 80 91 74 58

والمطلوب : 1- حساب معامل الالتواء بطريقة " بيرسون " .

2- حساب معامل الالتواء الربيعي .

الحل

1- حساب معامل الالتواء بطريقة "بيرسون" .

في هذه الحالة يتم تطبيق المعادلة رقم (5-2) كما يلي:

1- حساب الوسط الحسابي ، والانحراف المعياري :

الدرجة x	x^2
66	4356
85	7225
52	2704
78	6084
80	6400
91	8281
74	5476
58	3364
584	43890

$$\sum x = 584 \;,\; \sum x^2 = 43890$$

ويكون :

$$\bar{x} = \frac{\sum x}{n} = \frac{584}{8} = 73$$

$$s = \sqrt{\frac{\sum x^2 - (\sum x)^2 / n}{n-1}} = \sqrt{\frac{43890 - (584)^2 / 8}{8-1}}$$
$$= \sqrt{\frac{1258}{7}} = \sqrt{179.71428} = 13.406$$

2- حساب الوسيط :

موقع الوسيط : $(n+1)/2=(8+1)/2=4.5$

52	58	66	74	78	80	85	91
1	2	3	4	5	6	7	8

2.25 4.5 6.75

$$Med = 74 + 0.5(78-74) = 76$$

3- معامل الالتواء "بيرسون"

$$s.c = \frac{3(\bar{x} - Med)}{S} = \frac{3(73-76)}{13.406} = -0.67$$

إذا منحنى توزيع درجات الطلاب ملتوي جهة اليسار .

2- معامل الالتواء الربيعي .

لحساب معامل الالتواء الربيعي ، يتم تطبيق الاتي :

1- حساب الربيع الأدنى .

موقع الرباعي : $(n+1)/4=(8+1)(1/4)=2.25$

إذا

$$Q_1 = 58 + (2.25 - 2)(66 - 58) = 60$$

2- حسب الرباعي الأعلى.

موقع الرباعي : $(n+1)/(3/4)=(8+1)\ (3/4)=6.75$

إذا

$$Q_3 = 80 + (6.75 - 6)(85 - 80) = 83.75$$

3- الوسيط (الربيع الثاني)

$$Med(Q_2) = 76$$

إذا معامل الالتواء الربيعي هو :

$$\sim_q = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_1)} = \frac{(83.75 - 76) - (76 - 60)}{(83.75 - 60)}$$

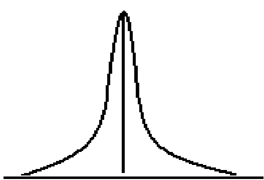
$$= \frac{-8.25}{23.75} = -0.35$$

إذا توزيع درجات الطلاب ملتوي جهة اليسار.

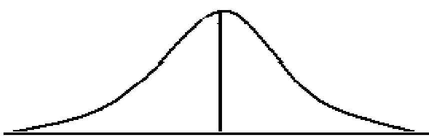
2-7-4. التفرطح Kurtosis

وهو مقياس يصف ارتفاع قمة المنحنى من حيث الاعتدال أو التدبب أو التفرطح .

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



منحنى مدبب



منحنى مفرطح

ويمكن قياس التفرطح باستخدام عدد من الطرق، ومنها طريقة العزوم ، حيث يحسب معامل التفرطح

(K) بتطبيق المعادلة التالية :

$$K = \frac{\frac{1}{n}\sum (x-\bar{x})^4}{S^4}$$

حيث أن المقدار $\sum(x-\bar{x})^4 / n$ هو العزم الرابع حول الوسط ، s هو الانحراف المعياري . ومعامل التفرطح في التوزيع الطبيعي يساوي 3 ، ومن ثم يمكن وصف منحنى التوزيع من حيث التفرطح ، والتدبب كما يلي :

1. **إذا كان** k=3 كان منحنى التوزيع معتدلا .
2. **إذا كان** k>3 كان منحنى التوزيع مدببا .
3. **إذا كان** k<3 كان منحنى التوزيع منبسطا (مفرطحا) .

مثل

كانت درجات 8 طلاب في الاختبار النهائي في **مادة** الرياضيات ، **كالتالي.**

66 85 52 78 80 91 74 58

والمطلوب : إيجاد شكل التفرطح ؟

وبالتطبيق على بيانات المثال رقم (5-1) **نجد أن**: $\bar{x} = 73$

x	66	85	52	78	80	91	74	58	584
$(x-\bar{x})$	-7	12	-21	5	7	18	1	-15	0
$(x-\bar{x})^2$	49	144	441	25	49	324	1	225	1258
$(x-\bar{x})^4$	2401	20736	194481	625	2401	104976	1	50625	376246

ومن البيانات أعلاه نجد أن:

$$s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}} = \sqrt{\frac{1258}{7}} = 13.406$$

$$\frac{1}{n}\sum(x-\bar{x})^2 = \frac{1}{8}(376246) = 47030.75$$

إذا معامل التفرطح هو:

$$K = \frac{47030.75}{(13.406)^4} = \frac{47030.75}{(32299.58)} = 1.456$$

إذا شكل توزيع بيانات الدرجات مفرطح.

الأسئلة

س1/ عرف مقاييس التشتت واذكر اهم أنواعها |؟

س2/ عرف ما يلي :

المدى ، الانحراف الربيعي ، التباين ، الانحراف المعياري ، الاختلاف .

س3/ ماهي مزايا وعيوب المدى ؟

س4/ ماهي مزايا وعيوب الانحراف الربيعي ؟

س5/ اذكرطرق حسب التباين للبينات المبوبة وغير المبوبة مع ذكر العلاقات الرياضية؟

س6/ ماهي مزايا والانحراف المعياري؟

س7/ ماهي مزايا وعيوب الاختلاف؟

س8/ الاتي درجات الحرارة لأحدى المدن خلال يومين ، المطلوب اوجد المدى لدرجات الحرارة ثم قارن ايهما اكثر تجانسا خلال يومين

اليوم الاول 7 , 8 , 11 , 10 , 4 , 3 , 5 , 6

اليوم الثاني 7 , 9 , 12 , 16 , 8, 4 , 5 , 7

س9/ الجدول التكراري التالي يبين اوزان 50 طالب ، المطلوب حساب المدى لهذا التوزيع

فئات الوزن	50-55	55-60	60-65	65-70	70-75	75-80	80-85	المجموع
التكرارات	8	5	11	16	4	4	2	50

س10/ جد الانحراف الربيعي للقيم التالية؟

2 , 7, 9 , 3, 10 , 12 , 22 , 4 , 8 , 20 , 19 , 18 , 17 , 5 , 21 , 23

س11/ احسب الانحراف الربيعي للتوزيع التكراري الاتي

الفئات	0-10	10-30	30-50	50-65	65-90	90-100	100-110	مجموع
التكرارات	3	6	13	15	12	9	2	60

س12/ اوجد الانحراف المعياري للقيم الاتية بالطريقة المطولة (الانحرافات) و بالطريقة المختصرة؟

القيم 19 , 13 , 15 , 12 , 10 , 9

س13/ الجدول الاتي يبين عدد المعامل وعد العمال الذين يشتغلون في كل معمل ، المطلوب ايجاد درجة التشتت في عدد العمال الذين يشتغلون في كل معمل باستخدام الطريقة المطولة والطريقة المختصرة

عدد المعامل	10-20	20-30	30-40	40-50	50-60	60-70	70-80	المجموع
عدد العمل	25	30	42	68	85	35	15	300

س14/ أوجد ما يلي :

1- المدى

2- الانحراف الربيعي

3- التباين

4- الانحراف المعياري

5- الاختلاف للجدول التكراري الآتي

الفئات	3 - 9	9 - 15	15 - 21	21 - 27	27 - 33	33 - 39	المجموع
التكرار	10	12	8	6	3	1	40

الفصل الخامس

الارتباط

5- الارتباط Correlation

هو العلاقة بين ظاهرتين مثل العلاقة بين طول الشخص (سم) ووزنه (كغم) او العلاقة بين نسبة الشفاء من مرض معين وكمية الجرعة من الدواء المخصص للمريض وعمر المريض ، او العلاقة بين الدخل والاستهلاك والعلاقة بين درجات الطلبة وعدد ساعات الدراسة .

5-1. تحديد أسلوب قياس الارتباط المناسب وفقا لنوع البيانات:

كمية- كمية ــــــــــ معامل ارتباط بيرسون

رتبيه- رتبيه ــــــــــ معامل سبيرمان

كمية- رتبيه ــــــــــ معامل سبيرمان

5-2. معامل الارتباط الخطي البسيط :هو المقياس الذي نقيس به درجة الارتباط .

والارتباط بين الظواهر اما ان يكون موجب او سالب وهذا لا يعكس **قوة الارتباط ، فقوة الارتباط تكون محصورة** بين الصفر **و** +1 **او** -1 وكلما اقترب من الواحد يكون الارتباط قويا وكلما اقترب من الصفر يكون ضعيفا **أما اذا كان الارتباط** يساوي صفر يعني لا يوجد ارتباط واذا كان الارتباط يساوي واحد (**موجب اوسالب**) يعني ارتباط تام . ويرمز للارتباط بالرمز $r \times y$.

5-2-1. الارتباط لبيانات غير مبوبة

1. الطريقة المختصرة (طريقة انحرافات القيم عن الوسط الحسابي)وذلك باستخدام الصيغة الاتية :

$$r \times y = \frac{\sum \left(X_i - \ddot{X}\right).\left(y - \ddot{y}\right)}{\sqrt{\sum \left(X_i - \ddot{X}\right)^2.\left(y - \ddot{y}\right)^2}}$$

حيث ان $r \times y$ تعني الارتباط بين $X \times Y$

Xi قيم مشاهدات المجموعة الاولى وyi قيم مشاهدات المجموعة الثانية

$\ddot{X}$ **الوسط الحسابي للمجموعة الاولى وان** $\ddot{y}$ الوسط الحسابي للمجموعة الثانية

مثال : احسب معامل الارتباط بين المتغيرين **(X) و (Y)**

قيم X 2, 2 , 5 , 4 , 5 , 6 , 3 , 5 , 4 ,

قيم Y 3, 5 , 7 , 8 , 9 , 11 , 6 , 8 , 6 ,

الحل :

1 – نجد الوسط الحسابي لقيم X **والوسط الحسابي** لقيم Y

2 – نجد انحرافات القيم عن الوسط الحسابي للمتغير $\ddot{X}$ **و** للمتغير $\ddot{Y}$

3 – نربع انحرافات القيم عن الوسط الحسابي للمتغير X **والمتغير** Y

$$\ddot{X} = \frac{\sum X_i}{n} = \frac{36}{9} = 4 \quad ; \quad \ddot{Y} = \frac{\sum Y_i}{n} = \frac{63}{9} = 7$$

X_i	$(X_i - \ddot{X})$	$(X_i - \ddot{X})^2$	Y_i	$(Y_i - \ddot{Y})$	$(Y_i - \ddot{Y})^2$	$(X_i - \ddot{X}) . (Y_i - \ddot{Y})$
2	2 – 4 = -2	4	3	-4	16	- 2 . - 4 = 8
2	-2	4	5	-2	4	4
5	1	1	7	0	0	0
4	0	0	8	1	1	0
5	1	1	9	2	4	2
6	2	4	11	4	16	8
3	-1	1	6	-1	1	1
5	1	1	8	1	1	1
4	0	0	6	-1	1	0
36	0	16	63		44	24

$$r \times y = \frac{\sum \left(X_i - \ddot{X}\right).\left(Y - \ddot{Y}\right)}{\sqrt{\sum \left(X_i - \ddot{X}\right)^2.\left(Y - \ddot{Y}\right)^2}} = \frac{24}{\sqrt{16 \times 44}} = 0.905$$

الارتباط موجب

2. **الطريقة المطولة** باستخدام القيم الاصلية وحسب الصيغة الاتية :

$$r = \frac{n.\sum X_i y_i - \sum X_i \sum y_i}{\sqrt{[n.\sum X_i^{\,2} - (\sum X_i)^2].[n.\sum y_i^2 - \left(\sum y_i\right)^2]}}$$

مثال : البيانات الاتية تمثل الدخل (X) والاستهلاك (y) بالدينار العراقي ، **المطلوب حساب العلاقة** بين الدخل والاستهلاك

X 200 , 300 , 400 , 600 , 900

Y 180 , 270 , 320 , 480 , 700

الحل :

1- نجد مجموع X

2- نجد مجموع y

3- نضرب X في y

4- نجد مجموع x^2 ومجموع y^2

X_i	y_i	$X_i\, y_i$	X_i^2	y_i^2
200	180	36000	40000	32400
300	720	81000	90000	72900
400	320	128000	160000	102400
600	480	288000	360000	230400
900	700	630000	810000	490000
2400	1950	1163000	1460000	928100

$$r = \frac{n.\sum X_i y_i - \sum X_i \sum y_i}{\sqrt{[n.\sum X_i^{\,2} - (\sum X_i)^2].[n.\sum y_i^2 - \left(\sum y_i\right)^2]}}$$

$$r = \frac{5(1163000) - (2400)(1950)}{\sqrt{\left[5 \times 1460000 - (2400)^2\right] \cdot \left[5 \times 928100 - (1950)^2\right]}}$$

$$r = \frac{5815000 - 4680000}{\sqrt{[7300000 - 5760000] \cdot [4640500 - 3802500]}}$$

$$r = \frac{1135000}{\sqrt{[1540000] \cdot [838000]}} = \frac{1135000}{1136010.5} = 0.999$$

العلاقة بين الدخل والاستهلاك عالية جدا وقوية موجبة وقريبة من الواحد

5-3. ارتباط الرتب (سبيرمان وسبيرمان المعدل)

ان الصيغ السابقة الخاصة لحساب معامل الارتباط البسيط تستند بالحقيقة على اعتبار ان المتغيرات المعتمدة في الحساب هي متغيرات من النوع الكمي . الا انه من الناحية العملية هناك الكثير من الحالات التي تكون فيها المتغيرات من النوع الوصفي (اي متغيرات غير قابلة للقياس كالمهنة ، الحالة الاجتماعية ، تقديرات درجات وغيرها). وبهدف حساب الارتباط بين متغيرات من هذا النوع فانه لا يمكن استخدام الصيغ السابقة مباشرة لهذا الغرض بل اجراء بعض التحويرات عليها بالشكل الذي يجعلها ممكنة الاستخدام . ان المعامل الذي يقيس درجة الترابط ما بين صفتين يدعى بمعامل ارتباط الرتب لسبيرمان.

5-3-1. معامل ارتباط الرتب : هو الذي يمثل درجات الارتباط بين المتغيرات من النوع الوصفي (متغيرات وصفية)

$$r \times y = 1 - \frac{6.\sum d_i^2}{n.(n^2 - 1)}$$

حيث ان :

Di = Xi-yi اي انحراف قيم الرتب Xi عن قيم الرتب yi ويعني هذا ان معامل الارتباط البسيط يساوي

$$r \times y = 1 - \frac{6.\sum (X_i - y_i)^2}{n.(n^2 - 1)}$$

مثل : الاتي تقديرات 6طلبة في امتحان الرياضيات والاحصاء ، المطلوب حساب العلاقة بين تقديرات المادتين

تقديرات X ضعيف ، امتياز ، جيد ، متوسط ، مقبول ، جيد جدا

تقديرات y مقبول ، جيد جدا ، جيد ، ضعيف ، متوسط ، امتياز

الحل : الخطوات

1- لتحويل التقديرات الى ارقام نعطي لهذه التقديرات ارقام متسلسلة ونرتبها اما تصاعديا او تنازليا وكما يلي :

ضعيف	مقبول	متوسط	جيد	جيد جدا	امتياز
1	2	3	4	5	6

1- نرتب تقديرات X وتقديرات y ترتيبا تصاعديا او ترتيبا تنازليا

2- نجد di من العلاقة (Xi – yi)

3-نجد di من العلاقة (Xi – yi)

4-نجد تربيع di ثم نجمع di^2

5-نطبق صيغة القانون.

ترتيب التقديرات	تسلسل الرتب	ترتيبX	ترتيب y	$d_i = (X_i - y_i)$	d_i^2
ضعيف	1	1	2	-1	1
مقبول	2	6	5	1	1
متوسط	3	4	4	0	0
جيد	4	3	1	2	4
جيد جدا	5	2	3	-1	1
امتياز	6	5	6	-1	1
Total					8

$$r \times y = 1 - \frac{6.\sum d_i^2}{n.(n^2 - 1)} = 1 - \frac{6 \times 8}{6.(36 - 1)} = 1 - \frac{48}{210} = 1 - 0.228 = 0.772$$

5-3-2. ارتباط سبيرمان المعدل

في حالة تكرار بعض قيم احد المتغيرين او كليهما عندئذ لا يمكن استخدام صيغة معامل ارتباط الرتب لسبيرمان بل يستوجب اجراء بعض التعديلات عليها وعلى النحو التالي :

بعد ترتيب قيم المتغير (الصفات) على نحو تصاعدي او تنازلي يتم تخصيص قيم سلسلة الاعداد الطبيعية **كرتب** لهذه الصفات ، ومن ثم يتم حساب معدل القيم المخصصة واعادة تخصيصه لتلك الصفات المكررة ، بعد ذلك يتم التعديل من خلال اضافة الكمية $m(m^2-1)\backslash 12$ **الى** $\sum di^2$ ،حيث m تمثل عدد مرات تكرار الصفة ،هذه الكمية **تضاف الى** $\sum di^2$ مقابل كل صفة مكررة .هذا الاجراء يتم لكلا المتغيرين X **و** y .

مثال : الاتي تقديرات لكفاءة عشرة من العاملين في احد المصانع من حيث ادارتهم في تشغيل نوعين من المكائن الحديثة , المطلوب حساب معامل الارتباط البسيط

تقديرات نوع X **من المكائن** : جيد ، متوسط ، جيد جدا ، متوسط ، امتياز ، ضعيف ، **مقبول** ، متوسط ، جيد ، جيد.

تقديرات نوع **y** من المكائن : متوسط ، جيد ، جيد ، مقبول ، جيد جدا ، مقبول ، ضعيف ، متوسط ، متوسط ، امتياز .

الحل:

1. نرتب تقديرات المتغير X تصاعديا او تنازليا ، ثم نضع لها ارقام متسلسلة
2. **نجد معدل الرتب** (**رتب** X)
3. نرتب تقديرات المتغير الاخر y اما تصاعديا او تنازليا ونضع لها ارقام متسلسلة
4. **نجد معدل رتب** y
5. **نجد** di **من العلاقة** (Xi – yi)
6. **نربع** di **ثم نجد مجموع** di^2
7. **نحسب عدد التكرارات** للتقديرات المتكررة لكل متغير X**و**y من خلال احتساب التعديلات وذلك **بإضافة** الكمية $m(m^2-1)\backslash 12$

1. نجمع التعديلات الكلية للمتغيرين معا
2. **نحسب معامل ارتباط الرتب لس**بيرمان المعدل من الصيغة الاتية :

$$r \times y = 1 - \frac{6 \,.(\sum d_i^2 - t)}{n.\left(n^2 - 1\right)}$$

تقديرات X	ترتيب X تصاعدي	معدل رتب X	تقديرات y	ترتيب y تصاعدي	معدل رتب y	d_i	d_i^2
جيد	1 ضعيف	7	متوسط	1 ضعيف	5	2	4
متوسط	2 مقبول	4	جيد	2 مقبول	7.5	-3.5	12.25
جيد جدا	3 متوسط	9	جيد	3 مقبول	7.5	1.5	2.25
متوسط	4 متوسط	4	مقبول	4 متوسط	2.5	1.5	2.25
امتياز	5 متوسط	10	جيد جدا	5 متوسط	9	1	1
ضعيف	6 جيد	1	مقبول	6 متوسط	2.5	-1.5	2.25
مقبول	7 جيد	2	ضعيف	7 جيد	1	1	1
متوسط	8 جيد	4	متوسط	8 جيد	5	-1	1
جيد	9 جيد جدا	7	متوسط	9 جيد جدا	5	2	4
جيد	10 امتياز	7	امتياز	10 امتياز	10	-3	9
المجموع							39

نحسب عدد التكرارات وكالاتي :

تكرارات المتغير X : لقد تكرر التقدير المتوسط 3 مرات فيعني ان $m = 3$ وعليه

$m(m^2-1)\backslash 12 = 3(3^2-1)\backslash 12 = 24\backslash 12 = \mathbf{2}$

عدد تكرارات التقدير جيد هو 3 فأذن

$m(m^2-1)\backslash 12 = 3(9-1)\backslash 12 = 24\backslash 12 = \mathbf{2}$

تكرارات المتغير y : التقدير مقبول تكر مرتين و عليه فان $m = 2$

$m\,(m^2-1)\backslash 12 = 2(2^2-1)\backslash 12 = 2(4-1)\backslash 12 = 6\backslash 12 = \mathbf{0.5}$

المتوسط تكرر 3 مرات فإذن m = 3

$$3(3^2-1)\backslash 12 = 3*8 \backslash 12 = 24\backslash 12 = \mathbf{2}$$

التقدير جيد تكرر مرتين فان m = 2

$$2(2^2-1)\backslash 12 = 2*3 \backslash 12 = 6\backslash 12 = \mathbf{0.5}$$

وعليه فان التعديل الكلي يساوي

$$2 + 2 + 0.5 + 2 + 0.5 = \mathbf{7}$$

وبذلك معامل ارتباط الرتب لسبيرمان المعدل يكون كالاتي:

$$r \times y = 1 - \frac{6\,.(\sum d_i^2 - t)}{n.\left(n^2 - 1\right)} = 1 - \frac{6.\left(39 + 7\right)}{10.\left(10^2 - 1\right)} = 1 - \frac{6 \times 46}{990} = 1 - 0.2787 = 0.721$$

4-5. ارتباط البيانات المبوبة (للصفات)

افرض وجود توزيع تكراري مزدوج عدد صفوفه (فئات المتغير الاول)ويرمز له K وعدد اعمدته (فئات المتغير الثاني) ويرمز له m ، وهذا يعني ان عدد خلايا هذا التوزيع Km. **وان** fi تمثل تكرار الخلية المقابلة لفئة المتغير **الاول** (X) **وان** fj يمثل تكرار الخلية المقابلة للمتغير الثاني (y) وافرض ايضا ما يلي :

Xi مراكز فئات المتغير X (**الصفوف**)

Yi مراكز فئات المتغير y (**الاعمدة**)

fi مجاميع **التكرارات المقابلة لفئات** X

fj مجاميع **التكرارات المقابلة لفئات** y

Li**طول كل فئة من فئات** X

Lj**طول كل فئة من فئات** y

A يمثل وسط فرضي اختير من بين مراكز فئات X

B يمثل وسط فرضي اختير من بين مراكز فئات y

وان

$$U_i = \frac{X_i - A}{L_i}$$

وان

$$U_j = \frac{y_j - B}{L_j}$$

وان Ui fi هي حاصل ضرب Ui في fi

وان Uj fj هي حاصل ضرب Uj في fj

وان Ui^2 fi هي حاصل ضرب مربع Ui في fi

وان Uj^2 fj هي حاصل ضرب مربع Uj في fj

وان Ui Vj fij هي حاصل ضرب Ui في Vj في تكرار الخلية المقابلة للفئة i من X وللفئة j من y

وان n تمثل المجموع الكلي للتكرارات في التوزيع المزدوج وان $n = \sum fj = \sum fi$ عندئذ وفق هذه المعطيات يمكن حساب معامل الارتباط بين X وy وفق الصيغة الاتية :

$$r \times y = \frac{n\sum^{k}\sum^{m} U_i V_j f_{ij} - \left(\sum U_i f_i\right).\left(\sum V_j f_j\right)}{\sqrt{\left[n\sum^{k} U_i^2 f_i - \left(\sum^{k} U_i f_i\right)^2\right].\left[n\sum^{m} V_j^2 f_j - \left(\sum^{m} V_j f_j\right)^2\right]}}$$

مثال :

الاتي توزيع تكراري مزدوج لدرجات 100 طالب في مادتي الاحصاء X والرياضيات y ، المطلوب حسب معامل الارتباط البسيط ما بين X وy

y / X	30- 40	40-50	50-60	60-70	70-80	f_i
40-50	2	4	6			12
50-60	3	6	7	1		17
60-70	1	3	10	4	1	19
70-80		3	5	8	9	25
80-90			2	5	11	18
90-100			1	3	5	9
f_j	6	16	31	21	26	100

خطوات الحل :

1. نجد مركز فئة X
2. نجد Ui حيث ان $U_i = \frac{X_i - A}{L_i}$
3. نجد حاصل ضرب $Ui\ fi$ ثم نجمع حاصل الضرب
4. نجد حاصل ضرب مربع Ui في fi ثم نجمع حاصل الضرب
5. نجد مركز فئة y
6. نجد V_j حيث ان $V_i = \frac{Y_j - B}{L_j}$
7. نجد حاصل ضرب $V_j\ f_j$ ثم نجمع حاصل الضرب
8. نجد حاصل ضرب مربع V_j في f_i ونجمع حاصل الضرب
9. نجد حاصل ضرب $U_i . V_j . f_{ij}$
10. نجد حاصل ضرب $V_i . U_i . f_{ij}$ ثم نطابق صيغة القانون

Y / X	30-40	40-50	50-60	60-70	70-80	fi	Xi	Ui	Uifi	Ui^2fi	UiVjfij
40-50	2 **8** 8	4 **8** 8	6 **0** 0	— **0** 0	— 0	12	45	-2	-24	48	**16**
50-60	3 **6**	6 **6**	7 **0**	1 **-1**	—	17	55	-1	17	17	**11**
60-70	1 **0**	3 **0**	10 **0**	4 **0**	1 **0**	19	**65**	0	0	0	**0**
70-80	— **0**	3 **-3**	5 **0**	8 **8**	9 **18**	25	75	1	25	25	**23**
80-90	— **0**	— **0**	2 **0**	5 **10**	11 **44**	18	85	2	36	72	**54**
90-100	— **0**	— **0**	1 **0**	3 **9**	5 **30**	9	95	3	27	81	**39**
fj	6	16	31	21	26	100			∑47	∑234	**∑143**
Yj	35	45	**55**	65	75	—					
Vj	-2	-1	0	1	2	—					
Vjfj	-12	-16	0	21	25	∑45					
Vj^2fj	24	16	0	21	104	∑156					
VjUifij	**14**	**11**	**0**	**26**	**92**	**∑143**					

$$r \times y = \frac{n\sum^{k}\sum^{m} U_i V_j f_{ij} - \left(\sum U_i f_i\right).\left(\sum V_j f_j\right)}{\sqrt{\left[n\sum^{k} U_i^2 f_i - \left(\sum^{k} U_i f_i\right)^2\right].\left[n\sum^{m} V_j^2 f_j - \left(\sum^{m} V_j f_j\right)^2\right]}}$$

$$r \times y = \frac{100 \times (143) - (47).(45)}{\sqrt{\left[100 \times (243) - (47)^2\right].\left[100 \times (156) - (45)^2\right]}} = 0.657$$

5-5. معامل الاقتران Coefficient of Association

يعرف **معامل الاقتران** بانه :مقياس لقياس العلاقة بين متغيرين وصفيين مفرغة بياناتهما في جدول **توافق ذو مرتبة** 2*2 لا يمكن اخضاعهما للقياس الكمي . افرض ان X_i و y_i متغيرين من النوع الوصفي وعليه يمكن تصور جدول التوافق ذو المرتبة 2*2 وبالشكل التالي :

y / X	Y_1	Y_2	مجموع
X_1	f_{11}	f_{12}	f_1
X_2	f_{21}	f_{22}	f_2
مجموع	$f_{.1}$	$f_{.2}$	n

وعليه فان معامل الاقتران بين المتغيرين X و y يستخرج وفق الصيغة الاتية :

$$C\,.\,A = \frac{f_{11}\,f_{22} - f_{12}\,f_{21}}{f_{11}\,f_{22} + f_{12}\,f_{21}}$$

بمعنى ان معامل الاقتران =
حاصل ضرب العناصر القطرية الرئيسية – حاصل ضرب العناصر الثانوية
____
حاصل ضرب العناصر القطرية الرئيسية + حاصل ضرب العناصر الثانوية

وان قيمة معامل الاقتران تتراوح بين 1+ و 1-

مثال (1)

من المعلوم ان عادة التدخين تؤثر تأثيرا سيئا على الصحة العامة للفرد ، المطلوب ايجاد العلاقة بين الحالة الصحية وعادة التدخين .

الحل :

الحالة الصحية \ عادة التدخين	يدخن	لا يدخن	**المجموع**
جيدة	f_{11} 40	f_{12} 50	90
غير جيدة	f_{21} 50	f_{22} 60	110
المجموع	90	110	200

$$C.A = \frac{f_{11}\,f_{22} - f_{12}\,f_{21}}{f_{11}\,f_{22} + f_{12}\,f_{21}} = \frac{(40).(60) - (50).(50)}{(40).(60) + (50).(50)}$$

$$= \frac{2400 - 2500}{2400 + 2500} = \frac{-100}{4900} = -0.02$$

العلاقة عكسية

مثل (2)

الجدول التالي يبين عدد الحوادث التي تعرضت لها شاحنات لنقل البضائع خلال فترة زمنية معينة **موزعة حسب نوع الحادث وحالة الجو** ، المطلوب ايجاد معامل الاقتران لهذا التوزيع

X \ Y	دهس	اصطدام	مجموع
صحو	25	12	37
ممطر	10	50	60
مجموع	35	69	97

الحل :

$$C.A=\frac{f_{11}\,f_{22}-f_{12}\,f_{21}}{f_{11}\,f_{22}+f_{12}\,f_{21}}=\frac{(25).(50)-(12).(10)}{(25).(50)+(12).(10)}$$

$$=\frac{1250-120}{1250+120}=\frac{1130}{1370}=0.824$$

العلاقة قوية جدا

6-5. معامل التوافق Coefficient of contingency

درسنا في الفقرة السابقة ان استخدام معامل الاقتران مقصورا على الظواهر التي تنقسم الى مجموعتين ، اما اذا كانت احد الظاهرتين اللتين نبحث العلاقة بينهما او كلتيهما تنقسم الى اكثر من نوعين ، فان معامل الاقتران لا يساعدنا في هذه الحالة وعندئذ نستخدم معامل التوافق الذي وضعه بيرسون لقياس العلاقة بين الصفات غير المقاسة او بين صفات بعضها تقاس **بالأرقام** وبعضها لا يقاس .ومعامل التوافق حسب الصيغة الآتية

$$C = \sqrt{\frac{r - 1}{r}}$$

$$r_1 = \frac{f_{11}^2}{T_1.T_1} + \frac{f_{12}^2}{T_1.T_2} + \frac{f_{13}^2}{T_1.T_3} + \frac{f_1 m^2}{T_1.T.m} = \frac{1}{T_1}.\frac{\sum f_{11}^2}{T_i}$$

$$r_2 = \frac{f_{12}^2}{T_2.T_1} + \frac{f_{22}^2}{T_2.T_2} + \frac{f_{23}^2}{T_2.T_3} + \frac{f_1 m^2}{T_2.T.m} = \frac{1}{T_2}.\frac{\sum f_{21}^2}{T_i}$$

وهكذا $\mathbf{r}_3$ و$\mathbf{r}_4$ لجميع الصفوف اي ان

$$r = r_1 + r_2 + r_3$$

مثل : الجدول الاتي يبين عدد حوادث الطرق التي تعرضت لها شاحنات لنقل البضائع خلال فترة زمنية موزعة حسب نوع الحادث وحالة الجو ،المطلوب : ايجاد معامل التوافق لهذا التوزيع

نوع الحادث / حالة الجو	دهس	اصطدام	انقلاب	المجموع
صحو	25	12	9	46
ممطر	10	50	35	95
ضبابي	20	45	40	105
المجموع	55	107	84	246

الحل :

1. **نجد** r_1 من الصيغة الاتية :

$$r_1 = \frac{f_{11}^{\ 2}}{T_1 . T_1} + \frac{f_{12}^{\ 2}}{T_1 . T_2} + \frac{f_{13}^{\ 2}}{T_1 . T_3} + \frac{f_1 m^{\ 2}}{T_1 . T . m} = \frac{1}{T_1} . \frac{\sum f_{11}^{\ 2}}{T_i}$$

$$r_1 = \frac{25^{\ 2}}{(46).(55)} + \frac{12^{\ 2}}{(46).(107)} + \frac{9^{\ 2}}{(46).(84)}$$

$$= \frac{625}{2530} + \frac{144}{4922} + \frac{81}{3864}$$

$$= 0.247 + 0.02 + 0.02 = 0.296$$

$$r_2 = \frac{f_{12}^{\ 2}}{T_2 . T_1} + \frac{f_{22}^{\ 2}}{T_2 . T_2} + \frac{f_{23}^{\ 2}}{T_2 . T_3} + \frac{f_1 m^{\ 2}}{T_2 . T . m} = \frac{1}{T_2} . \frac{\sum f_{21}^{\ 2}}{T_i}$$

$$r_2 = \frac{10^{\ 2}}{(95).(55)} + \frac{50^{\ 2}}{(95).(107)} + \frac{35^{\ 2}}{(95).(84)} = 0.419$$

$$r_3 = \frac{20^{\ 2}}{(105).(55)} + \frac{45^{\ 2}}{(105).(107)} + \frac{40^{\ 2}}{(105).(84)} = 0.431$$

$$r = r_1 + r_2 + r_3 = 0.296 + 0.419 + 0.431 = 1.143$$

$$C = \sqrt{\frac{r-1}{r}} = \sqrt{\frac{1.143 - 1}{1.143}} = \sqrt{\frac{0.143}{1.143}} = \sqrt{0.125} = 0.354$$

العلاق إيجابية ضعيفة

الأسئلة

س1/ احسب معامل الارتباط البسيط بين y وX للبيانات الاتية :

قيم X   2 , 3 , 4 , 5 , 6 , 7 ,

قيم y   0 , 3 , 4 , 4 , 6 , 11 ,

س2/ احسب معامل الارتباط للمتغرين y وX باستخدام الطريقة المطولة

X   2 , 2 , 5 , 4 , 5 , 6 , 3 , 5 , 4 ,

y   3 , 5 , 7 , 8 , 9 , 11 , 6 , 8 , 6 ,

س3/ الاتي تقديرات 6طلبة في مادتي الرياضيات والاحصاء ، المطلوب ايجاد معامل الارتباط بين تقديرات الطلبة

الرياضيات X   متوسط   جيد   مقبول   ضعيف   امتياز   جيد جدا

الاحصاء y   جيد   متوسط   ضعيف   مقبول   جيد جدا   امتياز

س4/ افرض ان اثناء وباء التيفوئيد مثلا اجرى احد الاطباء تجربة مصل جديد على عينة من الافراد حجمها (343) وكانت النتائج كما في الجدول التالي ، المطلوب حساب معامل الاقتران بين حالة التلقيح بالمصل والاصابة بالمرض

التلقيح / الاصابة	لقح بالمصل	لم يلقح بالمصل	المجموع
لم يصب بالمرض	192	113	305
اصيب بالمرض	34	4	38
المجموع	226	117	343

س5/ في دراسة لمعرفة العلاقة بين عدد حقول النفط المكتشفة و طول الأنابيب بالكيلومتر الناقلة للنفط خلال عدة سنوات ،سجلت 6 قراءات . و المطلوب **ايجاد معامل الارتباط**. (**الحل** : 0.886 **الارتباط طردي قوي**)

عدد حقول النفط X	55	54	56	61	62	63
طول الانابيب Y	21906	22300	23100	23203	23203	23600

س6/ **كانت نتائج مجموعة من الطلبة في الامتحانات** النهائية كما في الجدول التالي ، المطلوب تقدير العلاقة بين تحصيل الطالب والمواظبة على حضور المحاضرات

المواظبة / التحصيل الدراسي	جيد	متوسط	رديء	المجموع
الناجحين	80	20	5	105
المكملين	30	40	40	110
الراسبين	10	40	85	135
المجموع	120	100	130	350

س7/ عند دراسة العلاقة بين رائحة ولون الزهرة لعينة مكونة من 30 زهرة كانت لدينا النتائج التالية :

احسب معامل التوافق C بين اللون ورائحة الزهور ؟ (الحل : 0.22 علاقة ضعيفة)

الرائحة Y / اللون X	بدون رائحة	له رائحة	المجموع
اصفر	6	4	10
ابيض	7	2	9
احمر	6	5	11
المجموع	19	11	30

الفصل السادس

السلاسل الزمنية

6- السلاسل الزمنية Time Senes

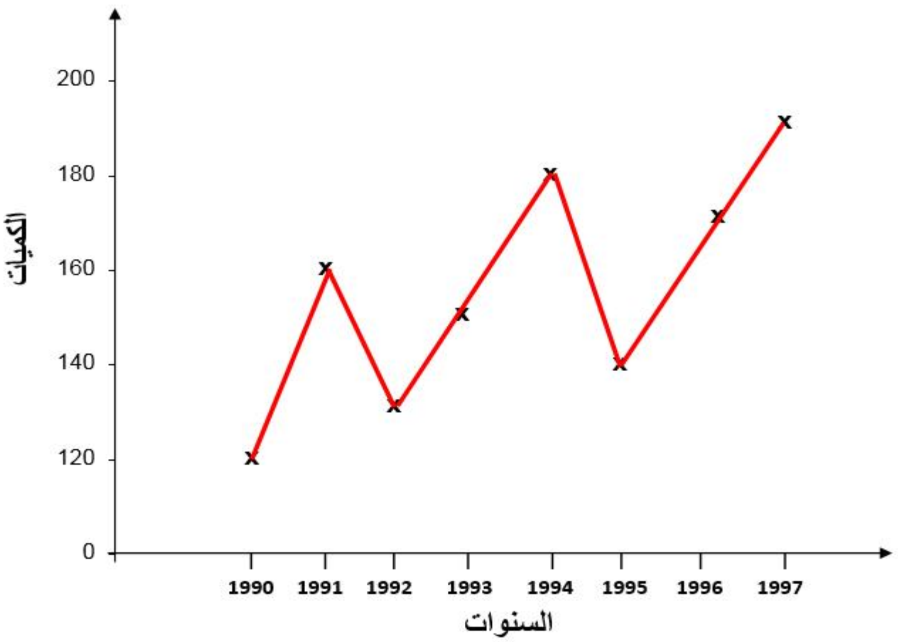
6-1. السلسلة الزمنية : هي مجموعة من القيم المشاهدة لظاهرة معينة لمدة من الزمن (عدة سنوات) في فترات زمنية متساوية . ان مجموعة القيم الدالة على صادرات العراق السنوية من التمور من سنة 1924 ــ 1970 هي سلسلة زمنية . وان المبالغ الدالة على مقدار ما يبيعه احد المخازن في اليوم ولمدة اسبوع هي سلسلة زمنية ايضا ، وكذلك اعتبار الجدول الذي يشير الى درجات الحرارة في بغداد في نهاية كل ساعة ولمدة يوم كامل هو سلسلة زمنية وهكذا . لذلك يمكن اعتبار السلسلة الزمنية علاقة بين متغيرين هما قيمة الظاهرة والزمن .

تهتم كثير من الدراسات ولاسيما الاقتصادية والاجتماعية بدراسة السلسلة الزمنية وذلك لان كثيرا من الظواهر الاقتصادية كالصادرات السنوية مثلا استعرضت وبحثت لعدد من السنين فانه يمكن معرفة طبيعة التغيرات التي تطرأ على قيم الظاهرة مع الزمن وتحديد الاسباب والنتائج وتفسير العلاقات المشاهدة بينها والتنبؤ بما سيحدث من تغير على قيم الظاهرة في المستقبل على ضوء ما حدث لها في الماضي .

مثل : البيانات التالية تمثل الكميات المستوردة من الحبوب خلال ثمانية سنوات بالطن , المطلوب تمثيل هذه السلسة بيانيا

السنوات	1990	1991	1992	1993	1994	1995	1996	1997
الكميات	120	160	130	150	180	140	170	190

الرسم البياني



من الرسم البياني يتبين بان رغم التذبذبات فان المنحنى الى ارتفاع بمرور الزمن

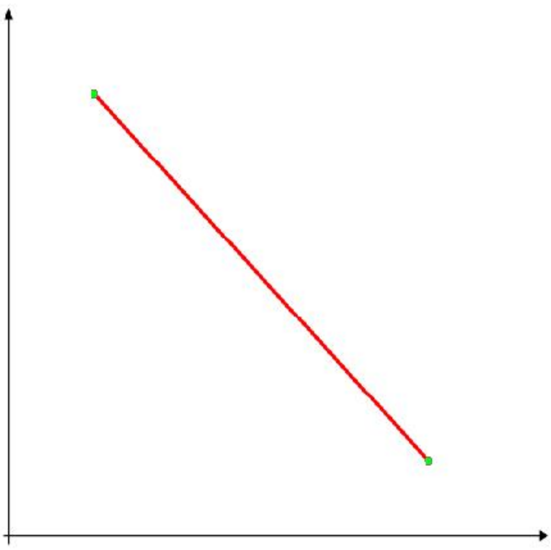
6-2. تحليل السلسلة الزمنية : ان دراسة السلسلة الزمنية يقتضي تحليلها الى العوامل المؤثرة فيها وقد وجد ان السلسلة الزمنية **تتأثر** بالعوامل الاتية :

1. الاتجاه العام
2. التغيرات الموسمية
3. التغيرات الدورية
4. التغيرات العرضية

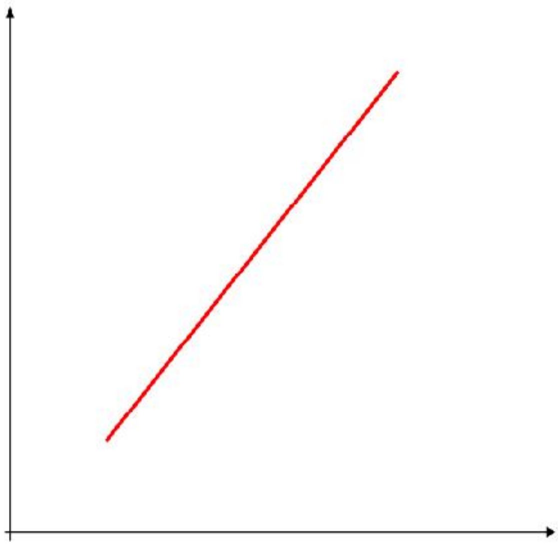
1-الاتجاه العام :هو العامل الاكثر تأثيرا على القيم الظاهرة في المدى الطويل فمثلا عدد سكان **العراق خلال** الاربعين سنة الماضية يمثل سلسلة زمنية **متأثرة** بدرجة كبيرة بالاتجاه العام . نظرا لميل جميع الاعداد المكونة لهذه السلسلة الى التزايد بصورة عامة . يكون الاتجاه العام موجب عندما تتزايد قيمة الظاهرة على مرور الزمن فمثلا نمو السكان يمثل سلسلة زمنية اتجاهها العام موجب .

ويكون التجاه العام سالبا عندما تتناقص قيم الظاهرة على مرور الزمن فمثلا الوفيات الناتجة عن الاصابة بمرض الجدري تكون سلسلة زمنية اتجاهها سالب .

والاتجاه العام يكون مستقيما اذا كان التغير في قيم الظاهرة كمية ثابتة موجبة او سالبة ويكون التجاه العام غير مستقيم (منحني) اذا كان التغير في قيم الظاهرة غير ثابت خلال المدة ويمكن توضيح ذلك بالشكلين الآتيين .



حالة الانكماش(القص)
الاتجاه العام سالب



حالة النمو المستمر
الاتجاه العام موجب

2 - **التغيرات الموسمية** : تشير التغيرات الموسمية الى متوسط التغير المنظم الذي يحدث خلال سنة واحدة او فصل واحد او شهر واحد .... الخ ان التغيرات الموسمية تتكرر خلال فترات منتظمة تتكون نتيجة اختلاف المناخ او عادات اجتماعية او مناسبات دينية او ما شابه **ذلك. فمثلا** التغيرات الموسمية في زيادة الاستهلاك من وقود

التدفئة في فصل الشتاء وزيادة الطلب على المراوح في فصل الصيف لذلك يمكن التنبؤ بمقدار الظاهرة في **المستقبل حسب فصول** او اشهر السنة .

3 – **التغيرات الدورية** :هي التغيرات التي تتكرر خلال فترة زمنية تزيد عن سنة مثل تعاقب الدورات الاقتصادية (الرخاء والانكماش) وتسمى بالتذبذبات الدورية وانها اقل انتظاما من التغيرات الموسمية.

4 – **التغيرات العرضية** : هي التغيرات التي تحدث **لأسباب** غير متوقعة مثل الحروب والكوارث الطبيعية ، ويصعب التنبؤ بالفترات التي يمكن ان تحدث فيها هذه التغيرات .

6-2-1. طرق ايجاد خط الاتجاه العام

1. طريقة متوسطي نصفي السلسلة
2. طريقة المتوسطات المتحركة
3. طريقة المربعات الصغرى

6-2-2. طريقة متوسطي نصفي السلسلة

بموجب هذه الطريقة تقسم السلسلة الى قسمين يفضل ان يكونا متساويين ثم نوجد الوسط الحسابي للقيم في كل قسم فنحصل على نقطتين على خط السلسلة الزمنية . ثم نرسم مستقيما بين النقطتين فيكون هو خط الاتجاه العام . ان هذه الطريقة بسيطة ولكن النتائج التي نحصل عليها قد لا تكون دقيقة كذلك فان العمل بهذه الطريقة يقتصر على الحالات التي يكون فيها الاتجاه العام مستقيما او قريبا من الاستقامة

مثال :

اوجد خط التجاه العام للسلسلة الزمنية الاتية باستخدام طريقة متوسطي نصفي السلسلة

السنوات : 1955 , 1956 1957 , 1958 , 1959 , 1960 , 1961 , 1962 , 1963 , 1964

السكان: 5.93 , 6.09, 6.26 , 6.45 , 6.67 , 6.89 , 7.13 , 7.37 , 7.62 , 7.88

(بالملايين)

الحل :

نقسم السلسلة الى قسمين متساويين الاول من سنة (1955 – 1959) والسنة الوسطى فيه هي سنة 1957 **والقسم الثاني من السلسلة من سنة** (1960 -1964) والسنة الوسطى فيه هي سنة 1962 **ومن ثم نجد الوسط الحسابي لكل قسم وكالاتي** :

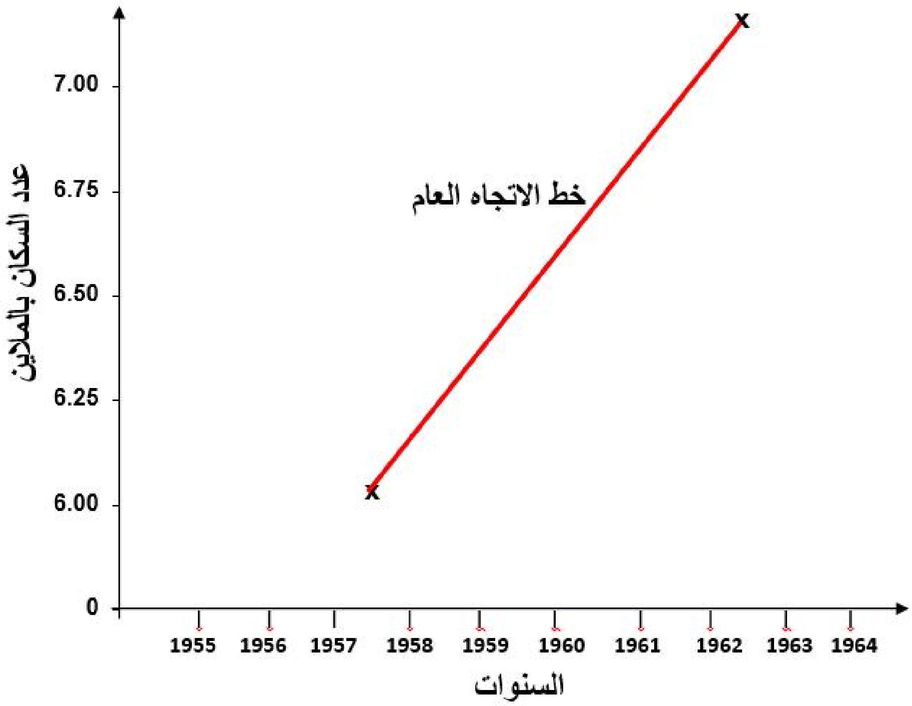
الوسط الحسابي للقسم الاول يساوي

$$\ddot{X} = \frac{\sum X_i}{n} = \frac{5.93+6.09+6.26+6.45+6.67}{5} = 6.28$$

الوسط الحسابي للقسم الثاني

$$\ddot{X} = \frac{\sum X_i}{n} = \frac{6.89+7.13+7.37+7.62+7.88}{5} = 7.38$$

نرسم الاحداثي السيني (المحور الافقي) و**يمثل السنوات والاحداثي الصادي (المحور العمودي)** ويمثل عدد السكان بالملايين ونحدد النقطتين (الوسط الحسابي لكل قسم) ثم نرسم خط مستقيم يصل بين النقطتين فنحصل على **خط الاتجاه العام**



6-2-3. طريقة المتوسطات المتحركة :

بموجب هذه الطريقة نختار عدد من السنين ولتكن ثلاثة سنوات او خمس سنوات ونحسب متوسط قيم الظاهرة لهذه السنين ثم نترك القيمة الاولى من القيم التي اخذناها **ونأخذ** بدلها القيمة التالية للمجموعة السابق اخذها فنحصل على قيم بنفس عدد القيم السابقة **ونأخذ** متوسطها ثم نترك القيمة الاولى من المجموعة الجديدة وناخذ بدلها القيمة التالية للمجموعة ثم نستخرج متوسطها وهكذا نحصل على المتوسطات المتحركة لقيم الظاهرة ثم نضع المتوسط لكل مجموعة امام القيمة الوسطى من قيمها ان كان عددها فرديا وامام احدى القيمتين الوسطيتين ان كان عددا زوجيا ثم نثبت هذا المتوسط على شكل نقاط ونصل بينها بخط مستقيم فيكون خط الاتجاه العام .

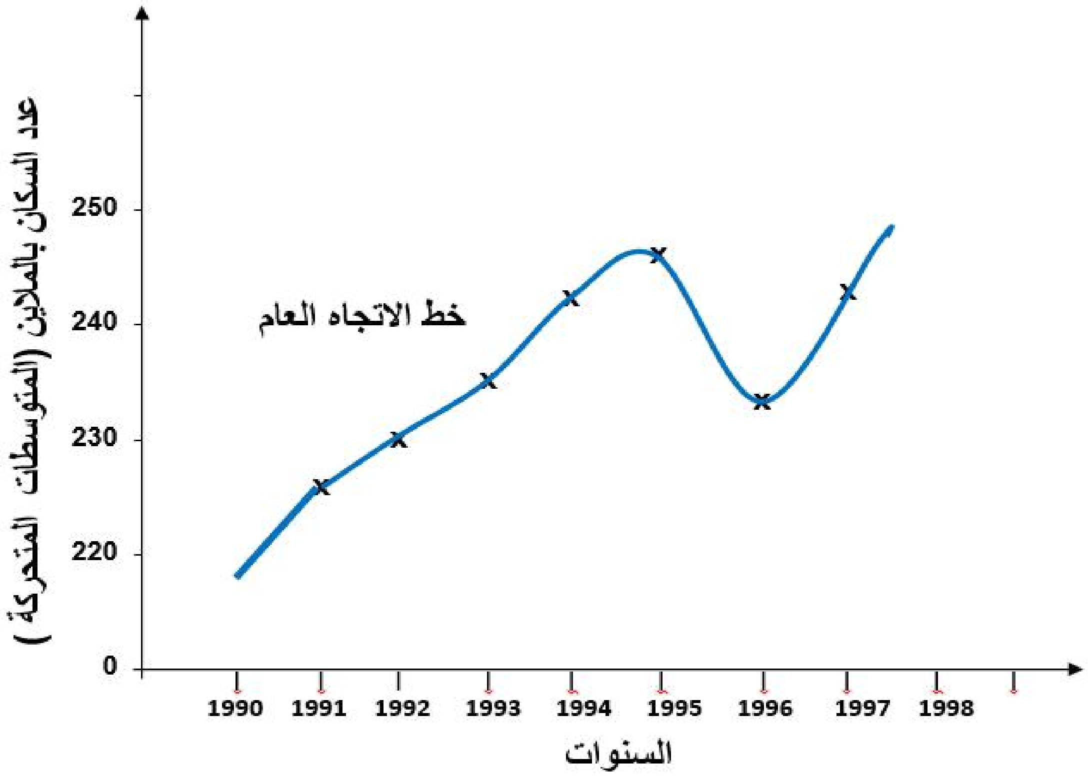
مثال : الجدول الاتي يبين قيم المبيعات الفعلية yi لاحد مصانع السمنت وللسنوات من 1990ـ 1998 ، المطلوب رسم خط الاتجاه العام باستخدام طريقة المتوسطات المتحركة .

السنوات : 1990 , 1991 , 1992 , 1993 , 1994 , 1995 , 1996 , 1997 , 1998

المبيعات بالملايين : 210 , 240 , 230 , 220 , 260 , 250 , 230 , 220 , 280

الحل : نفرض ان عدد السنين المناسب هو ثلاث سنوات

السنوات	المبيعات بالملاين y_i	**مجموع ثلاث سنوات**	**متوسط ثلاث سنوات (المتوسطات المتحركة)**
1990	210		
1991	240	680	226.67
1992	230	690	230
1993	220	710	236.67
1994	260	730	243.33
1995	250	740	246.67
1996	230	700	233.33
1997	220	730	243.33
1998	280		



6-2-4. طريقة المربعات الصغرى

تعتبر هذه الطريقة من ادق طرق تعيين خط الاتجاه العام

$$Y_i = a + b.E$$

حيث ان a و b قيمتان ثابتتان

وان b تمثل

$$b = \frac{\sum y_i E_i - n\bar{Y}\ddot{E}}{\sum E^2 - n\ddot{E}^2}$$

وa تمثل

$$a = \bar{Y} - b\ddot{E}$$

$$\bar{E} = \frac{\sum Ei}{n} \quad ; \quad \bar{Y} = \frac{\sum y_i}{n}$$

مثل : الجدول الاتي يبين قيم المبيعات الفعلية (بالملايين) لأحد المصانع وللسنوات 96-98 وللمواسم الاتية:

المبيعات yi 280 , 220 , 230 , 250 , 260 , 220 , 230 , 240 , 210

الموسم Ei 9 , 8 , 7 , 6 , 5 ,4 , 3 , 2 , 1 ,

المطلوب : استخراج معادلة خط الاتجاه العام بطريقة المربعات الصغرى ثم قدر قيمة المبيعات للسنوات 1999 و 2000

E_i	y_i	$y_i E_i$	E_i^2
1	210	210	1
2	240	480	4
3	230	690	9
4	220	880	16
5	260	1300	25
6	250	1500	36
7	230	1610	49
8	220	1760	64
9	280	2520	81
45	2140	10950	285

ثم نجد $\bar{Y}$ و $\bar{E}$ اي الوسط الحسابي للمبيعات y_i **وللموسم** E_i

$$\bar{E} = \frac{\sum Ei}{n} = \frac{45}{9} = 5$$

$$\bar{Y} = \frac{\sum y_i}{n} = \frac{2140}{9} = 237.78$$

$$b=\frac{\sum y_i E_i - n\bar{Y}\ddot{E}}{\sum E^2 - n\ddot{E}^2}=\frac{10950 \;-9\times 5\times 237.78}{285\;-9\times 5^2}=\frac{350}{60}=4.17$$

$$a=\bar{Y}-b\ddot{E}=237.78-(4.17).(5)=237.78-20.85=216.93$$

$$Y_i=a+b.E=216.93-4.17\,E$$

ويمكن الاستفادة من معادلة خط الاتجاه العام للتنبؤ بقيم المبيعات لكل ثلاثة اشهر **من عام** 1999 **و**2000 **وذلك** بالتعويض عن قيمة E_i في معادلة خط الاتجاه العام وكما يلي :

$$1999 \qquad \Rightarrow \qquad E=10$$

$$Y_{1999}=a+b.E=216.93+4.17\times 10=258.63$$

$$2000 \qquad \Rightarrow \qquad E=11$$

$$Y_{2000}=a+b.E=216.93+4.17\times 11=262.8$$

وهكذا لبقية المواسم كما يمكن التعويض، **عن** E_I بالأشهر **او بالسنوات ، مثلا لسنة** 2004 ولشهر أيار

$$5\;/\;2005 \qquad \Rightarrow \qquad E=15.417$$

$$Y_{2000}=a+b.E=216.93+4.17\times 15.417=281.22$$

الأسئلة

س1/ الجدول التالي يبين انتاج البيض في احدى محطات الدواجن عن الفترة من 1991ـ 2000 بالميون بيضة،

المطلوب : رسم خط الاتجاه العام بطريقة المتوسطات المتحركة

الانتاج السنوي : 16.2, 15.2 , 15.1 16.9 ,16.8 ,12.9 ,13.8 , 16.1 , 20.4 , 17.8

السنوات : 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000

س2/ البيانات التالية تمثل الادخارات في احد المصارف مقدرة بالملايين ، **المطلوب/ اوجد معادلة خط الاتجاه** العام باستخدام طريقة المربعات الصغرى ثم قدر **لعام** 1990

المدخرات y_I	31	46	55	60	77	69	78	81
السنوات E	1982	1983	1984	1985	1986	1987	1988	1989

الفصل السابع

الأرقام القياسية

7- الارقام القياسية Index number

ان دراسة الارقام القياسية على جانب كبير من الاهمية لجميع الباحثين في المجالات الاقتصادية والاجتماعية والتجارية فالرقم القياسي يشير الى الانخفاض والارتفاع الذي يطرأ على الظاهرة موضوعة البحث وبواسطة الرقم القياسي يتمكن الباحث الاقتصادي ان يعطي رأيه في ارتفاع وانخفاض مستوى المعيشة في قطر من الاقطار او مدينة من المدن خلال فترة معينة من الزمن . كذلك يتمكن هذا الباحث عند وقوفه على التغير الحقيقي في الحالة الاقتصادية ان يضع سياسة اقتصادية انمائية مرنة من واقع التغيرات التي توصفها هذه الارقام .

وعن طريق دراسة الارقام القياسية للأجور مثلا يتمكن الباحث من ان يقف على الظروف الاجتماعية والاقتصادية للعمال في الصناعات المختلفة في السنوات المختلفة واثر التحسينات في وسائل الانتاج وتخفيض ساعات العمل وغير ذلك مما يتعلق بالأجور ومستواها.

الرقم القياسي : هو قياس احصائي يبين التغير في قيمة ظاهرة معينة او مجموعة من الظواهر ذات العلاقة بالنسبة الى قيمتها في زمان او مكان معين .

فترة الاساس : هي الفترة التي تنسب اليها عادة اسعار وكميات الفترات الاخرى وهذه الفترة قد تكون شهرا او سنة او غيرها ولكن الشائع ان تكون سنة واحدة وتسمى (سنة الاساس) كما يشترط ان تكون سنة طبيعية خالية من الشذوذ كالحروب والازمات او الكوارث.

فترة المقارنة : هي الفترة التي تنسب اسعارها او كمياتها الى اسعار او كميات فترة الاساس .

7-1. انواع الارقام القياسية المستخدمة في التحليل الاحصائي :

اولا : الارقام القياسية البسيطة Simple index number

ثانيا : الارقام القياسية المرجحة Weighted index number

تكوين الارقام القياسية : من الممكن ايجاد ارقام قياسية لأية ظاهرة لغرض مقارنة قيمتها بين فترة واخرى . فمثلا نستطيع عمل رقم قياسي للإنتاج وكذلك للصادرات والواردات والاجور وغير ذلك ولكن من الارقام القياسية الاكثر شيوعا المهمة والاكثر شيوعا هي الارقام القياسية للأسعار.

7-1-1. الارقام القياسية البسيطة Simple index number

يحسب الرقم القياسي البسيط للأسعار بموجب هذه الطريقة بإيجاد الوسط الحسابي للأسعار خلال سنتي الاساس والمقارنة ثم نسب الوسط الحسابي للأسعار لسنة المقارنة الى الوسط الحسابي لسنة الاساس ونضرب الناتج *100للحصول على نسبة مئوية .

الوسط الحسابي للأسعار في سنة المقارنة = $\dfrac{\text{مجموع الاسعار في سنة المقارنة}}{\text{عدد الاسعار}}$

$$\overline{P}_i = \frac{\sum P_i}{n}$$

حيث ان P_i هي اسعار سنة المقارنة

اما الوسط الحسابي للأسعار في سنة الاساس = $\dfrac{\text{مجموع الاسعار في سنة الاساس}}{\text{عدد الأسعار}}$

$$\overline{P}_\circ = \frac{\sum P_\circ}{n}$$

حيث ان P_O هي اسعار سنة الاساس

وعليه فان الرقم القياسي = $\dfrac{\text{الوسط الحسابي للأسعار في سنة المقارنة}}{\text{الوسط الحسابي للأسعار في سنة الأساس}}$ * 100 %

$$P.I = \frac{\sum \frac{P_i}{n}}{\sum \frac{P_\circ}{n}} \times 100\ \%$$

وعليه فان الرقم القياسي للأسعار ياخذ الصيغة الاتية :

$$P.I = \frac{\sum P_i}{\sum P_i} \times 100\ \%$$

مثال:

الجدول الاتي يبين اسعار بعض السلع في سنتي (1969) (1993) ، المطلوب ايجاد الرقم القياسي البسيط للأسعار لسنة 1993 علما بان 1969 هي سنة الاساس

السلعة	السعر	
	1969	1993
قمح	600	4000
سكر	30	700
لحم	32	320
المجموع	662	4390

$$P.I = \frac{\sum P_i}{\sum P_i} \times 100\ \% = \frac{4390}{662} \times 100\ \% = 663\ \%$$

اي ان الاسعار ارتفعت بمقدار 663 دينار عما كانت عليه في سنة الاساس 1969

7-1-1-1. استخدام الرقم القياسي البسيط للمقارنة بين كميات السلع وقيمتها

يستخدم الرقم القياسي البسيط في حالة المقارنة بين كميات السلع وذلك بقسمة كمية السلعة في سنة المقارنة qi على كمية السلعة في سنة الاساس qo **وضرب الناتج** *100

اي ان الرقم القياسي البسيط للكمية ياخذ الصيغة الاتية :

$$Q.I = \frac{q_i}{q_\circ} \times 100$$

اما اذا توفرت بيانات عن سعر السلعة والكمية المباعة او المنتجة في سنتي الاساس والمقارنة فيمكننا استخراج الرقم القياسي لقيمة الانتاج وذلك بقسمة قيمة السلعة في سنة المقارنة على قيمتها في سنة الاساس وضرب **الناتج** * 100

ملاحظة : القيمة = الكمية * السعر ويرمز للقيمة بـ V وعليه فان $V = q \times p$

وعليه فان الرقم القياسي للقيمة ياخذ الصيغة الاتية :

$$Q.I = \frac{p_i\ q_i}{p_\circ\ q_\circ}$$

$p_i\, q_i$ القيمة في سنة المقارنة

$p_o\, q_o$ هي القيمة في سنة الاساس

مثال :

اذا كانت الكمية المباعة من احدى السلع في سنة 1991 (150) **وحدة بسعر** (20) دينار للوحدة ، وقد ارتفعت الكمية المباعة من هذه **للسلعة في سنة** 1997 **الى** (250) **وحدة وبسعر** (22) دينار للوحدة .

المطلوب : **احسب أ**- الرقم القياسي البسيط لسعر السلعة عام 1997 **باعتبار ان سنة** 1991 هي سنة الاساس

ب – الرقم القياسي البسيط لكمية الانتاج لعام 1997 **باعتبار ان سنة** 1991 هي سنة **الاساس**

ج – الرقم القياسي البسيط للقيمة لعام 1997

الحل :

أ- نحسب الرقم القياسي البسيط **للأسعار لعام** 1991

$$P.I = \frac{P_{i\,1997}}{P_{\circ\,1991}} \times 100\ \% = \frac{22}{20} \times 100\ \% = 110\ \%$$

بمعنى ان السعر ارتفع بنسبة 10% **لعام** 1997 **مقارنة بسنة الاساس**

ب- الرقم القياسي البسيط لكمية الانتاج

$$Q.I = \frac{q_{i\,1997}}{q_{\circ\,1991}} \times 100\ \% = \frac{250}{150} \times 100\ \% = 166.67\ \%$$

بمعنى ان الكمية المنتجة ارتفعت بنسبة 66.67% **لعام** 1997

ث- **الرقم** القياسي البسيط للقيمة

$$V.I = \frac{P.1 \times Q.I}{100} \times 100\ \% = \frac{110 \times 166.67}{100} \times 100\ \% = 183.3\ \%$$

كما يمكن استخراج نفس النتيجة بقسمة حاصل ضرب الرقمين القياسيين للسعر والكمية المستخرجة سابقا على

$$V.I = \frac{P.1 \times Q.I}{100} \times 100\ \% = \frac{110 \times 166.67}{100} \times 100\ \% = 183.3\ \%$$

2-7. الارقام القياسية المرجحة :

ان ما يعيب الارقام القياسية البسيطة كونها تعطي اهمية متساوية للسلع المختلفة لدى استخراج ارقامها القياسية اي انها تعطي صورة غير دقيقة للتغيرات الحاصلة في مستويات الاسعار او الكميات وغيرها .

لذلك نلجأ الى استخدام الارقام القياسية المرجحة اي التي تعطي لكل سلعة الوزن الحقيقي الخاص بأهميتها وذلك من خلال ترجيح الاسعار وحيث يوجد لدينا نوعية من الاسعار (اسعار سنة الاساس واسعار سنة المقارنة) كذلك نوعين من الكميات (كميات سنة الاساس وكميات سنة المقارنة) لذلك يمكن ان نحصل على ثلاثة انواع من الارقام القياسية المرجحة تعرف بأسماء العلماء الذين توصلوا ليها وهي : -

1- رقم لاسبير **Laspeyers Index**
2- رقم باش **Paasche**
3- رقم فيشر **Fisher(الامثل)**

1-2-7. رقم لاسبير : يعتمد على الترجيح بكميات سنة الاساس

$$\text{رقم لاسبير} = \frac{\text{مجموع اسعار سنة المقارنة *كميات سنة الاساس}}{\text{مجموع اسعار سنة الاساس *كميات سنة الاساس}} * 100$$

اي ان رقم لاسبير يساوي

$$P.I(Las.) = \frac{\sum P_i q_\circ}{\sum P_\circ q_\circ} \times 100 \%$$

مثال: الجدول الاتي يبين كميات السلع واسعارها في سنتين 1939 , 1957 ، المطلوب : ايجاد الرقم القياسي للأسعار في سنة 1957 باعتبار ان سنة 1939 هي سنة الاساس بطريقة لاسبير

نوع السلعة	السعر		الكمية	
	1939	1957	1939	1957
قمح	2	4	18	24
باقلاء	4	9	6	12
ذرة	1	3	20	32
رز	1	3	250	400

الحل :

نضرب اسعار سنة المقارنة * كميات سنة الاساس ثم نجمع حاصل الضرب

نضرب أسعار سنة الأساس * كميات سنة الأساس ثم نجمع حاصل الضرب

$p_{i\ (1957)} \cdot q_{\circ\ (1939)}$	$p_{\circ\ (1939)} \cdot q_{\circ\ (1939)}$
$4 \times 18 = 72$	$2 \times 18 = 36$
$9 \times 6 = 54$	$4 \times 6 = 24$
$3 \times 20 = 60$	$1 \times 20 = 20$
$250 \times 3 = 750$	$1 \times 250 = 250$
936	330

$$P.I(Las.) = \frac{\sum P_i q_\circ}{\sum P_\circ q_\circ} \times 100\ \% = \frac{936}{330} \times 100\ \% = 284\ \%$$

يلاحظ ان هذا الرقم القياسي يمكن استخراجه في حالة عدم توفر البيانات الخاصة بكميات سنة المقارنة (لأنه يعتمد على كميات سنة الاساس)

2-2-7. رقم باش (الترجيح بكميات سنة المقارنة)

الرقم القياسي لباش = $\frac{\text{مجموع اسعار سنة المقارنة * كميات سنة المقارنة}}{\text{مجموع اسعار سنة الاساس * كميات سنة المقارنة}}$ *100

ويأخذ الرقم القياسي لباش الصيغة الاتية:

$$P.I(Paa.) = \frac{\sum P_i q_i}{\sum P_\circ q_i} \times 100\ \%$$

مثال : الجدول الاتي يبين كميات السلع واسعارها في سنتين 1939 ، 1957 ، المطلوب : ايجاد الرقم القياسي **للأسعار في سنة** 1957 **باعتبار ان سنة** 1939 هي سنة الاساس بطريقة باش

نوع السلعة	السعر		الكمية	
	1939	1957	1939	1957
قمح	2	4	18	24
باقلاء	4	9	6	12
ذرة	1	3	20	32
رز	1	3	250	400

الحل :

1. نضرب الاسعار في سنة المقارنة * كميات سنة المقارنة ونجمع حاصل الضرب
2. نضرب لسعار سنة الاساس * كميات سنة المقارنة ونجمع حاصل الضرب

$p_{i\,(1957)} \cdot q_{i\,(1957)}$	$p_{\circ\,(1939)} \cdot q_{i\,(1957)}$
$24 \times 4 = 96$	$2 \times 24 = 48$
$12 \times 9 = 108$	$12 \times 4 = 48$
$32 \times 3 = 96$	$32 \times 1 = 32$
$400 \times 3 = 1200$	$400 \times 1 = 400$
1500	528

$$P.I(Paa.) = \frac{\sum P_i q_i}{\sum P_\circ q_i} \times 100\% = \frac{1500}{528} \times 100\% = 284\%$$

3-2-7. الرقم القياسي الامثل

ويسمى بمعادلة فيشر وبموجب هذه المعادلة يمكن احتساب الرقم القياسي للاسعار من الجذر التربيعي لحاصل ضرب الناتج من معادلة باش * الرقم القياسي الناتج من معادلة لاسبير

وعليه فان الرقم القياسي الامثل يساوي

$$P.I(Fis.) = \sqrt{\left(\frac{\sum p_i q_\circ}{\sum p_\circ q_\circ}\right).\left(\frac{\sum p_i q_i}{\sum p_\circ q_i}\right)} \times 100\%$$

وبالرجوع الى المثال السابق للاسبير وباش فان الرقم القياسي لفيشر يساوي

$$P.I(Fis.) = \sqrt{\left(\frac{\sum p_i q_\circ}{\sum p_\circ q_\circ}\right).\left(\frac{\sum p_i q_i}{\sum p_\circ q_i}\right)} \times 100\%$$

$$= \sqrt{\left(\frac{936}{330}\right).\left(\frac{1500}{528}\right)} \times 100\% = \sqrt{2.84 + 2.84} \times 100\% = 284\%$$

مثل : الجدول الاتي يبين الاسعار والكميات المقابلة لها

البيانات / السلعة	P_o لسعار لساس	P_i لسعار مقارنة	q_o كميات اساس	q_i كميات مقارنة
1	4	10	20	25
2	2	5	40	30
3	6	9	15	20
4	8	15	10	15

المطلوب :

أ- رقم لاسبير القياسي

ب- رقم باش القياسي

ج- رقم فيشر القياسي (الامثل)

الحل :

أ- رقم لاسبير

$p_\circ \cdot q_\circ$	$p_i \cdot q_\circ$
$4 \times 20 = 80$	$10 \times 20 = 200$
$2 \times 40 = 80$	$5 \times 40 = 200$
$6 \times 15 = 90$	$9 \times 15 = 135$
$8 \times 10 = 80$	$15 \times 10 = 150$
330	685

$$P.I\left(Las.\right) = \frac{\sum P_i\, q_\circ}{\sum P_\circ\, q_\circ} \times 100\ \% = \frac{685}{330} \times 100\ \% \approx 207\ \%$$

ب- رقم باش

$p_i \cdot q_i$	$p_\circ \cdot q_i$
$10 \times 25 = 250$	$4 \times 25 = 100$
$5 \times 30 = 150$	$2 \times 30 = 60$
$9 \times 20 = 180$	$6 \times 20 = 120$
$15 \times 15 = 225$	$8 \times 15 = 120$
805	400

$$P.I(Paa.) = \frac{\sum P_i q_i}{\sum P_\circ q_i} \times 100\ \% = \frac{805}{400} \times 100\ \% \approx 201\ \%$$

ج- رقم فيشر

$$P.I(Fis.) = \sqrt{\left(\frac{\sum p_i q_\circ}{\sum p_\circ q_\circ}\right) \cdot \left(\frac{\sum p_i q_i}{\sum p_\circ q_i}\right)} \times 100\ \%$$

$$= \sqrt{\left(\frac{685}{330}\right) \cdot \left(\frac{805}{400}\right)} \times 100\ \% = \sqrt{2.07 + 2.01} \times 100\ \% \approx 202\ \%$$

الأسئلة

س1/ البيانات التالية توضح الكميات المصدرة من مجموعة من السلع في عامي 2000 و 2009 :

السلعة	الوحدة	2000	2009
أسماك	طن	70	200
لحوم	طن	20	80
مواد بترولية	الف برميل	400	900

س2/ يبين الجدول التالي الأسعار والكميات لثلاث منتجات :

المنتجات \ السنوات	سنة 2010		سنة 2000	
	السعر	الكمية	السعر	الكمية
الأول	9	6000	10	2000
الثاني	10	8000	15	3000
الثالث	8	7500	20	5000

المطلوب : 1- الرقم القياسي التجميعي البسيط للأسعار

2- الرقم القياسي التجميعي (رقم لاسبير)

3- الرقم القياسي التجميعي (رقم باش)

4- الرقم القياسي التجميعي (رقم فيشر)

س3/ الجدول الآتي يبين أسعار وكميات السلع A، B، C في عامي 2005، 2013.

السلعة	السعر		الكمية	
	(P_0) 2005	(P_1) 2013	(Q_0)2005	(Q_1)2013
A	13	18	8	14
B	16	21	6	12
C	22	28	10	18

المطلوب تقدير كل من : (1) رقم لاسبير للكميات (2) رقم فيشر للأسعار

الفصل الثامن

بعض المواضيع التطبيقية

8- بعض المواضيع التطبيقية

بعد دراستنا لمراحل الطريقة الاحصائية والمام الطالب ببعض المراحل الاساسية أعتقد ان من الضروري ان نطلع على بعض المواضيع التطبيقية التي تعالج بعض النواحي الاقتصادية والاجتماعية وغيرها ، والمواد التطبيقية كثيرة ومتشعبة فهناك الاحصاءات الحيوية والتجارية والهندسية والصناعية والزراعية واحصاءات الدخل القومي والخ ....

8-1. الاحصاءات الحيوية :

الاحصاءات الحيوية هي الاحصاءات التي تبحث وتحلل المظاهر المختلفة لحياة الانسان منذ ولادته الى وفاته . فهي تهتم بالولادات والوفيات والزواج والطلاق والامراض والعاهات والحرف والصناعات والهجرة . وهكذا نرى أهمية هذا النوع من الدراسة للباحث الاقتصادي والاجتماعي .

8-2. تعداد السكان

اهتمت الدول قديما بمعرفة عدد السكان وذلك لتقدير قوتها البشرية في الحروب وكذلك لجباية الضرائب . وكانت عملية العد نجري بطرق غير علمية وبتواريخ غير محددة . وفي الوقت الحاضر تطور اسلوب تعداد السكان واصبح يجري بفترات منتظمة ويشمل عد جميع سكان البلد في فترة معينة وبأساليب احصائية حديثة . وترمي عملية التعداد في الوقت الحاضر الى التعرف على الصفات المختلفة للسكان وتوزيع السكان جغرافيا ،توزيع السكان حسب العمر والنوع ، الحالة المدنية والعلمية والدينية ،توزيع السكان حسب الحرف وغير ذلك من الاوصاف التي تساعد على اعطاء صورة واضحة عن احوال السكان الاجتماعية والاقتصادية وهذه الصورة تساعد المسؤولين على وضع تخطيط سليم شامل . ونظرا لأهمية احصاء السكان لبرامج التخطيط ، فقد فكرت بعض الدول في جعل عملية تعداد السكان تجري بفترات قصيرة (كل خمس سنوات مثلا) الا ان العمل المرهق والوقت والمال اللازم لإجراء التعداد تحول دون تقصير فترة التعداد .

وفي العراق يجري التعداد كل عشر سنوات ، فقد قامت الحكومة العراقية بإجراء تسجيل عام لسكان العراق لمعرفة حالة السكان في سنة 1977 وقد سبق ان قامت الحكومة العراقية بعدة تسجيلات عامة قبل هذا التسجيل الاخير . وذلك في السنوات 1927 – 1934 – 1947 - 1957- 1965 . أما تعداد 1927 فلم يكن سوى محاولة فاشلة فلم تمضي بضعة شهور حتى تبين للحكومة فشل العملية فألغتها . اما تسجيل 1934 فقد اتبع نظام تعيين لجان استقرت في الاماكن على ان يستدعي المختار رؤساء العائلات للأدلاء بالبيانات المطلوبة وقد شمل هذا التسجيل جميع انحاء العراق وكان الغرض منه مركزا لخدمة أغراض التجنيد والانتخابات ، وقد استعمل اساسا لمنح (دفتر الجنسية) التي استبدلت فيما بعد بدفاتر النفوس لاستعمالها كمستند رسمي . أما تسجيل عام 1947 فيعتبر اول تسجيل جرى بواسطة العدادين والهيئات حيث قاموا بزيارة المساكن لاستقاء المعلومات

المطلوبة . وقد شملت عملية التسجيل جميع انحاء العراق في المدن والقصبات في يوم واحد . اما في القرى والارياف فقد بدأت قبل يوم التسجيل وانتهت في نفس يوم التسجيل . ولم تنجح العملية في القرى والارياف نجاحها في المدن والقصبات . وقد تخلف عدد كبير من السكان عن التسجيل وقد تبين ذلك عام 1957 . حيث كان التسجيل عام 1957 اكثر دقة وشمولا . اما التسجيل في عام 1965 فقد كان من المفروض ان يجري عام 1967 لكن لضرورة الاعداد للانتخابات فقد قدم موعد التعداد . وامتاز هذا التعداد عن التعدادات السابقة بكون الاستمارة احتوت على اسئلة اكثر فقد اضيف الى الاستمارة اسئلة تتعلق بالحرف والقومية والخدمة العسكرية وغيرها . وهناك طريقتان للتعداد .

8-2-1. التعداد الفعلي

وفي هذه الطريقة تعلن الدولة منع التجول ويحصر السكان كما هم في الواقع وقت التعداد ، ففي كل محل (بيت ، فندق ...الخ) يعتبر جميع الاشخاص الموجودين فيه ساعة التعداد من سكان البلدة سواء أكانوا زائرين أو من أصل السكان . كما ان أحد أفراد العائلة المتغيب عنها في يوم التعداد لا يعد من عائلته بل يعد مع سكان المدينة التي هو فيها . وهذه الطريقة سهلة وقليلة الاخطاء . اذ ان العداد لا يحتاج الى عد كل شخص في أي مكان يوجد فيه الا ان التعداد بهذه الطريقة لا يصور الاشياء على حقيقتها ويعطي معلومات غير صحيحة اذا كانت الدولة واسعة المساحة . والعراق من الدول التي تستخدم هذه الطريقة .

8-2-2. التعداد النظري

وهو حصر السكان حسب اقامتهم المعتادة . فأفراد العائلة المتغيبون عن البيت يحسبون مع عائلاتهم ولا يعدون مع سكان المدن الاخرى التي هم فيها وقت التعداد . وهذه الطريقة صعبة من الناحية العملية اذ تتطلب وضع أسئلة اضافية في استمارة التعداد لمعرفة محل الاقامة الحقيقية لكل شخص فيحصل الالتباس الذي يؤدي الى تسرب أخطاء كبيرة ولأجل ان يكون التعداد دقيقا يجب تهيئة جهاز منظم من العدادين والفنيين كما ان درجة ثقافة الشعب تؤثر في ذلك . ومن الدول التي تطبق هذه الطريقة الولايات المتحدة والمانيا .

والتعداد يجري عن طريق طبع استمارات وتوزيعها على العائلات والمحلات لملئها من قبلهم أو بواسطة العدادين . وتعتبر البيانات الموجودة في الاستمارات سرية يعاقب من يشي بها ويستعملها لغير أغراضها الاحصائية .

8-3. تقدير عدد السكان :

تحتاج برامج التخطيط للدولة الى معرفة عدد السكان سنويا ولا يمكن معرفته بواسطة التعداد العام الذي يجري مرة واحدة كل عشر سنوات بسبب صعوبة العمل والتكاليف الكثيرة لذلك تلجأ الى عمل تقديرات سنوية بعدد السكان وهناك ثلاث طرق لتقدير عدد السكان وهي (1) الزيادة الطبيعية (2) المتوالية العددية (3) المتوالية الهندسية .

8-3-1. الزيادة الطبيعية

وهذه الطريقة مبنية على اساس أنه اذا عرف التعداد لسنة 1957 مثلا وأريد معرفة عدد السكان التقديري للسنة 1958 فيجب ان تعرف عدد كل من المواليد والوفيات والداخلين والراحلين عنه في الفترة من نهاية 1957 **حتى** نهاية 1958 .

والزيادة الطبيعية للسكان = عدد المواليد + عدد المهاجرين للداخل – (عدد الوفيات +عدد المهاجرين للخارج)

ويمكن استخدام هذا المقياس لتقدير السكان في أي وقت اذا ضبطت احصاءات الهجرة وكان تسجيل المواليد والوفيات دقيقا .

8-3-2. المتوالية العددية

في حالة التقدير بطريقة المتوالية العددية يجب أن يكون معروفا لدينا **تعددان** سابقان للتقدير وذلك لاستخراج **الاساس** d اي اننا نطبق صيغة الحد الاخير L

$$L = A + (n-1)\, d$$

تطبق هذه الصيغة مرتين في المرة الاولى يكون

L_1 = التعداد الاخير قبل سنة التقدير

A التعداد السابق له

N_1 = (عدد السنين بين العدادين + l)

d = الزيادة السنوية للسكان . وفي المرة الثانية يكون

L_2 عدد النفوس في سنة التقدير للتعداد الاخير

N = (عدد السنين بين سنة التقدير والتعداد الاخير +1)

A_2 التعداد الاخير .

8-3-3. المتوالية الهندسية

في حالة التقدير بطريقة المتوالية الهندسية نطبق القانون

$$L = A * R^{n-1}$$

يطبق هذا القانون مرتين في المرة الاولى لاستخراج قيمة R معدل الزيادة السنوية للسكان A_1 , L_1 كعدادين الاخير والذي قبله ، n_1 (عدد السنين بين التعددين +1)، ويطبق القانون مرة ثانية لاستخراج عدد السكان في سنة التقدير حيث

L_2 عدد السكان في سنة التقدير

A_2 التعداد الاخير

N_2 عدد السنين بين سنة التقدير والتعداد الاخير .

مثال :

كان تعداد سكان العراق 4816185 **نسمة سنة** 1947 **وفي سنة** 1957 **كان التعداد** 6339960 والمطلوب تقدير **عدد السكان سنة** 1964 **باستخدام** :

أولا : المتوالية العددية

ثانيا : المتوالية الهندسية

الحل : بطريقة المتوالية العددية

$$L_1= A +(n-1)d$$

$$6339960= 486185 + (11-1)\ d$$

الزيادة السنوية $d= 633990-4816185\ /10 = 152378$

$$L_2 = 6339960+(8-1)*\ 152378$$

$L= 6339960 +1066346 = 7406306$ **نسمة عدد السكان سنة** 1964

الحل : بطريقة المتوالية الهندسية

$$L_1= A *R^{n1-1}$$

$$6339960= 4816185 * R^{11*1}$$

$$R= (6339960 / 4816185)^{1/10}$$

$$\text{Log } R= 1/10 (\text{Log } 6339960- \text{Log } 4816185)$$

$$Log\ R = 1/10\ (6.8021 - 6.6826)$$

$$LogR = 0.01195$$

ونقف الى هذه المرحلة بدون ان نجد العدد المقابل لأننا نحتاج لو R في الخطوة التالية

والان نفرض ان عدد السكان في سنة 1964 = L_2

$$L_2 = 6339960 * R^{8-1}$$

$$Log\ L_2 = Log\ 6339960 + 7 * Log\ R$$

$$Log\ L_2 = 6.8021 + 0.01195$$

حيث $Log\ R = 0.01195$ من السابق

$$Log\ L_2 = 6.8021 + 0.08365 = 6.88575$$

$LogL_2 = 7688000$ عدد نفوس العراق سنة 1964 .

الجداول الإحصائية

جدول كا^2

درجة الحرية	مستوى الدلالة أو الثقة		
	0.05	0.01	0.001
1	3.84	6.64	10.83
2	5.99	9.21	13.82
3	7.82	11.35	16.27
4	9.49	13.28	18.47
5	11.07	15.09	20.52
6	12.59	16.81	22.46
7	14.07	18.48	24.32
8	15.51	20.09	26.13
9	16.92	21.67	27.88
10	18.31	23.21	29.59
11	19.68	24.73	31.26
12	21.03	26.22	32.91
13	22.36	27.69	34.53
14	23.69	29.14	36.12
15	25.00	30.58	37.70
16	26.30	32.00	39.25
17	27.59	33.41	40.79
18	28.87	34.81	42.31
19	30.14	36.19	43.82
20	31.41	37.57	45.32
21	32.67	38.93	46.80
22	33.92	40.29	48.27
23	35.17	41.64	49.73
24	36.42	42.98	51.18
25	37.65	44.31	52.62
26	38.89	45.64	54.05

27	40.11	46.96	55.48
28	41.34	48.28	56.89
29	42.56	49.59	58.30
30	43.77	50.89	59.70
31	44.99	52.19	61.10
32	46.19	53.49	62.49
33	47.40	54.78	63.87
34	48.60	56.06	65.25
35	49.80	57.34	66.62
36	51.00	58.62	67.99
37	52.19	59.89	69.35
38	53.38	61.16	70.71
39	54.57	62.43	72.06
40	55.76	63.69	73.41
41	56.94	64.95	74.75
42	58.12	66.21	76.09
43	59.30	67.46	77.42
44	60.48	68.71	78.75
45	61.66	69.96	80.08
46	62.83	71.20	81.40
47	64.00	72.44	82.72
48	65.17	73.68	84.03
49	66.34	74.92	85.35
50	67.51	76.15	86.66
51	68.67	77.39	87.97
52	69.83	78.62	89.27
53	70.99	79.84	90.57
54	72.15	81.07	91.88
55	73.31	82.29	93.17
56	74.47	83.52	94.47
57	75.62	84.73	95.75
58	76.78	85.95	97.03

59	**77.93**	**87.17**	**98.34**
60	**79.08**	**88.38**	**99.62**
61	**80.23**	**89.59**	**100.88**
62	**81.38**	**90.80**	**102.15**
63	**82.53**	**92.01**	**103.46**
64	**83.68**	**93.22**	**104.72**
65	**84.82**	**94.42**	**105.97**
66	**85.97**	**95.63**	**107.26**
67	**87.11**	**96.83**	**108.54**
68	**88.25**	**98.03**	**109.79**
69	**89.39**	**99.23**	**111.06**
70	**90.53**	**100.42**	**112.31**
71	**91.67**	**101.62**	**113.56**
72	**92.81**	**102.82**	**114.84**
73	**93.95**	**104.01**	**116.08**
74	**95.08**	**105.20**	**117.35**
75	**96.22**	**106.39**	**118.60**
76	**97.35**	**107.58**	**119.85**
77	**98.49**	**108.77**	**121.11**
78	**99.62**	**109.96**	**122.36**
79	**100.75**	**111.15**	**123.60**
80	**101.88**	**112.33**	**124.84**
81	**103.01**	**113.51**	**126.09**
82	**104.14**	**114.70**	**127.33**
83	**105.27**	**115.88**	**128.57**
84	**106.40**	**117.06**	**129.80**
85	**107.52**	**118.24**	**131.04**
86	**108.65**	**119.41**	**132.28**
87	**109.77**	**120.59**	**133.51**
88	**110.90**	**121.77**	**134.74**
89	**112.02**	**122.94**	**135.96**
90	**113.15**	**124.12**	**137.19**

91	114.27	125.29	138.45
92	115.39	126.46	139.66
93	116.51	127.63	140.90
94	117.63	128.80	142.12
95	118.75	129.97	143.32
96	119.87	131.14	144.55
97	120.99	132.31	145.78
98	122.11	133.47	146.99
99	123.23	134.64	148.21
100	124.34	135.81	149

جدول ت

درجة الحرية	مستوى الدلالة							
طرف واحد	0.1	0.05	0.01	0.005	0.001	0.0005	0.0001	0.00005
طرفين	0.2	0.1	0.05	0.01	0.005	0.001	0.0005	0.0001
2	1.89	2.92	4.30	9.92	14.09	31.60	44.70	100.14
3	1.64	2.35	3.18	5.84	7.45	12.92	16.33	28.01
4	1.53	2.13	2.78	4.60	5.60	8.61	10.31	15.53
5	1.48	2.02	2.57	4.03	4.77	6.87	7.98	11.18
6	1.44	1.94	2.45	3.71	4.32	5.96	6.79	9.08
7	1.41	1.89	2.36	3.50	4.03	5.41	6.08	7.89
8	1.40	1.86	2.31	3.36	3.83	5.04	5.62	7.12
9	1.38	1.83	2.26	3.25	3.69	4.78	5.29	6.59
10	1.37	1.81	2.23	3.17	3.58	4.59	5.05	6.21
11	1.36	1.80	2.20	3.11	3.50	4.44	4.86	5.92
12	1.36	1.78	2.18	3.05	3.43	4.32	4.72	5.70
13	1.35	1.77	2.16	3.01	3.37	4.22	4.60	5.51
14	1.35	1.76	2.14	2.98	3.33	4.14	4.50	5.36
15	1.34	1.75	2.13	2.95	3.29	4.07	4.42	5.24
16	1.34	1.75	2.12	2.92	3.25	4.01	4.35	5.13
17	1.33	1.74	2.11	2.90	3.22	3.97	4.29	5.04

18	1.33	1.73	2.10	2.88	3.20	3.92	4.23	4.97
19	1.33	1.73	2.09	2.86	3.17	3.88	4.19	4.90
20	1.33	1.72	2.09	2.85	3.15	3.85	4.15	4.84
21	1.32	1.72	2.08	2.83	3.14	3.82	4.11	4.78
22	1.32	1.72	2.07	2.82	3.12	3.79	4.08	4.74
23	1.32	1.71	2.07	2.81	3.10	3.77	4.05	4.69
24	1.32	1.71	2.06	2.80	3.09	3.75	4.02	4.65
25	1.32	1.71	2.06	2.79	3.08	3.73	4.00	4.62
26	1.31	1.71	2.06	2.78	3.07	3.71	3.97	4.59
27	1.31	1.70	2.05	2.77	3.06	3.69	3.95	4.56
28	1.31	1.70	2.05	2.76	3.05	3.67	3.93	4.53
29	1.31	1.70	2.05	2.76	3.04	3.66	3.92	4.51
30	1.31	1.70	2.04	2.75	3.03	3.65	3.90	4.48
35	1.31	1.69	2.03	2.72	3.00	3.59	3.84	4.39
40	1.30	1.68	2.02	2.70	2.97	3.55	3.79	4.32
45	1.30	1.68	2.01	2.69	2.95	3.52	3.75	4.27
50	1.30	1.68	2.01	2.68	2.94	3.50	3.72	4.23
55	1.30	1.67	2.00	2.67	2.92	3.48	3.70	4.20
60	1.30	1.67	2.00	2.66	2.91	3.46	3.68	4.17
65	1.29	1.67	2.00	2.65	2.91	3.45	3.66	4.15
70	1.29	1.67	1.99	2.65	2.90	3.43	3.65	4.13
75	1.29	1.67	1.99	2.64	2.89	3.42	3.64	4.11
80	1.29	1.66	1.99	2.64	2.89	3.42	3.63	4.10
85	1.29	1.66	1.99	2.63	2.88	3.41	3.62	4.08
90	1.29	1.66	1.99	2.63	2.88	3.40	3.61	4.07
95	1.29	1.66	1.99	2.63	2.87	3.40	3.60	4.06
100	1.29	1.66	1.98	2.63	2.87	3.39	3.60	4.05
200	1.29	1.65	1.97	2.60	2.84	3.34	3.54	3.97
500	1.28	1.65	1.96	2.59	2.82	3.31	3.50	3.92
1000	1.28	1.65	1.96	2.58	2.81	3.30	3.49	3.91
∞	1.28	1.64	1.96	2.58	2.81	3.29	3.48	3.89

جدول ف

درجة حرية التباين الصغير	درجة حرية التباين الكبير								
	1	2	3	4	5	6	8	12	ã
1	161	200	216	225	230	234	239	244	254
2	18.5	19.0	19.2	19.3	19.3	19.3	19.4	19.4	19.5
3	10.1	9.6	9.3	9.1	9.0	8.9	8.8	8.7	8.5
4	7.7	6.9	6.6	6.4	6.3	6.2	6.0	5.9	5.6
5	6.6	5.8	5.4	5.2	5.1	5.0	4.8	4.7	4.4
6	6.0	5.1	4.8	4.5	4.4	4.3	4.2	4.0	3.7
7	5.6	4.7	4.4	4.1	4.0	3.9	3.7	3.6	3.2
8	5.3	4.5	4.1	3.8	3.7	3.6	3.4	3.3	2.9
9	5.1	4.3	3.9	3.6	3.5	3.4	3.2	3.1	2.7
10	5.0	4.1	3.7	3.5	3.3	3.2	3.1	2.9	2.5
11	4.8	4.0	3.6	3.4	3.2	3.1	3.0	2.8	2.4
12	4.8	3.9	3.5	3.3	3.1	3.0	2.9	2.7	2.3
13	4.7	3.8	3.4	3.2	3.0	2.9	2.8	2.6	2.2
14	4.6	3.7	3.3	3.1	3.0	2.9	2.7	2.5	2.1
15	4.5	3,7	3.3	3.1	2.9	2.8	2.6	2.5	2.1
16	4.5	3.6	3.2	3.0	2.9	2.7	2.6	2.4	2.0
17	4.5	3.6	3.2	3.0	2.8	2.7	2.6	2.4	2.0
18	4.4	3.6	3.2	2.9	2.8	2.7	2.5	2.3	1.9
19	4.4	3.5	3.1	2.9	2.7	2.6	2.5	2.3	1.9
20	4.4	3.5	3.1	2.9	2.7	2.6	2.5	2.3	1.8
21	4.3	3.5	3.1	2.8	2.7	2.6	2.4	2.3	1.8
22	4.3	3.4	3.1	2.8	2.7	2.6	2.4	2.2	1.8
23	4.3	3.4	3.0	2.8	2.6	2.5	2.4	2.2	1.8
24	4.3	3.4	3.0	2.8	2.6	2.5	2.4	2.2	1.7
25	4.2	3.4	3.0	2.8	2.6	2.5	2.3	2.2	1.7
26	4.2	3.4	3.0	2.7	2.6	2.5	2.3	2.2	1.7
27	4.2	3.4	3.0	2.7	2.6	2.5	2.3	2.1	1.7
28	4.2	3.3	3.0	2.7	2.6	2.4	2.3	2.1	1.7
29	4.2	3.3	2.9	2.7	2.5	2.4	2.3	2.1	1.6

30	**4.2**	**3.3**	**2.9**	**2.7**	**2.5**	**2.4**	**2.3**	**2.1**	**1.6**
40	**4.1**	**3.2**	**2.8**	**2.6**	**2.5**	**2.3**	**2.2**	**2.0**	**1.5**
60	**4.0**	**3.2**	**2.8**	**2.5**	**2.4**	**2.3**	**2.1**	**1.9**	**1.4**
120	**3.9**	**3.1**	**2.7**	**2.5**	**2.3**	**2.2**	**2.0**	**1.8**	**1.3**
ã	**3.8**	**3.0**	**2.6**	**2.4**	**2.2**	**2.1**	**1.9**	**1.8**	**1.0**

المصادر

1- حسن محمد حسن ، أساسيات الإحصاء وتطبيقاته ، دار المعرفة الجامعية ، 1992 .

2- حسن محمد حسن ، مبادئ الإحصاء الاجتماعي ، دار المعرفة الجامعية ، 2000 .

3- عبد الله عبد الحليم وآخرون ، الإحصاء مفاهيم أساسية ، 2003 .

4- محمد بهجت كشك ، مبادئ الإحصاء الاجتماعي ، دار المعرفة الجامعية ، 1996 .

5- مصطفى زايد ، الإحصاء ووصف البيانات ، 1989 .

6- احمد عبادة سرحان ، العينات ، مكتبة النهضة المصرية .1970

7- مقدمة الطرق الإحصائية ، دار المعرفة بمصر ، 1992.

8- احمد عبادة سرحان وصلاح الدين طلبة ، أسس الإحصاء ، دار الكتب الجامعية ، 1968 .

-

9- D. M. Bates and D. G. Watts (1988), Nonlinear Regression Analysis and Its Applications. John Wiley & Sons, New York.

10- Richard A. Becker, John M. Chambers and Allan R. Wilks (1988), The New S Language. Chapman & Hall, New York. This book is often called the “Blue Book”.

11- John M. Chambers and Trevor J. Hastie eds. (1992), Statistical Models in S. Chapman & Hall, New York. This is also called the “White Book”.

12- John M. Chambers (1998) Programming with Data. Springer, New York. This is also called the “Green Book”.

13- A. C. Davison and D. V. Hinkley (1997), Bootstrap Methods and Their Applications, Cambridge University Press.

14- Annette J. Dobson (1990), An Introduction to Generalized Linear Models, Chapman and Hall, London.

15- Peter McCullagh and John A. Nelder (1989), Generalized Linear Models. Second edition, Chapman and Hall, London.

16- John A. Rice (1995), Mathematical Statistics and Data Analysis. Second edition. Duxbury Press, Belmont, CA.

17- S. D. Silvey (1970), Statistical Inference. Penguin, London.