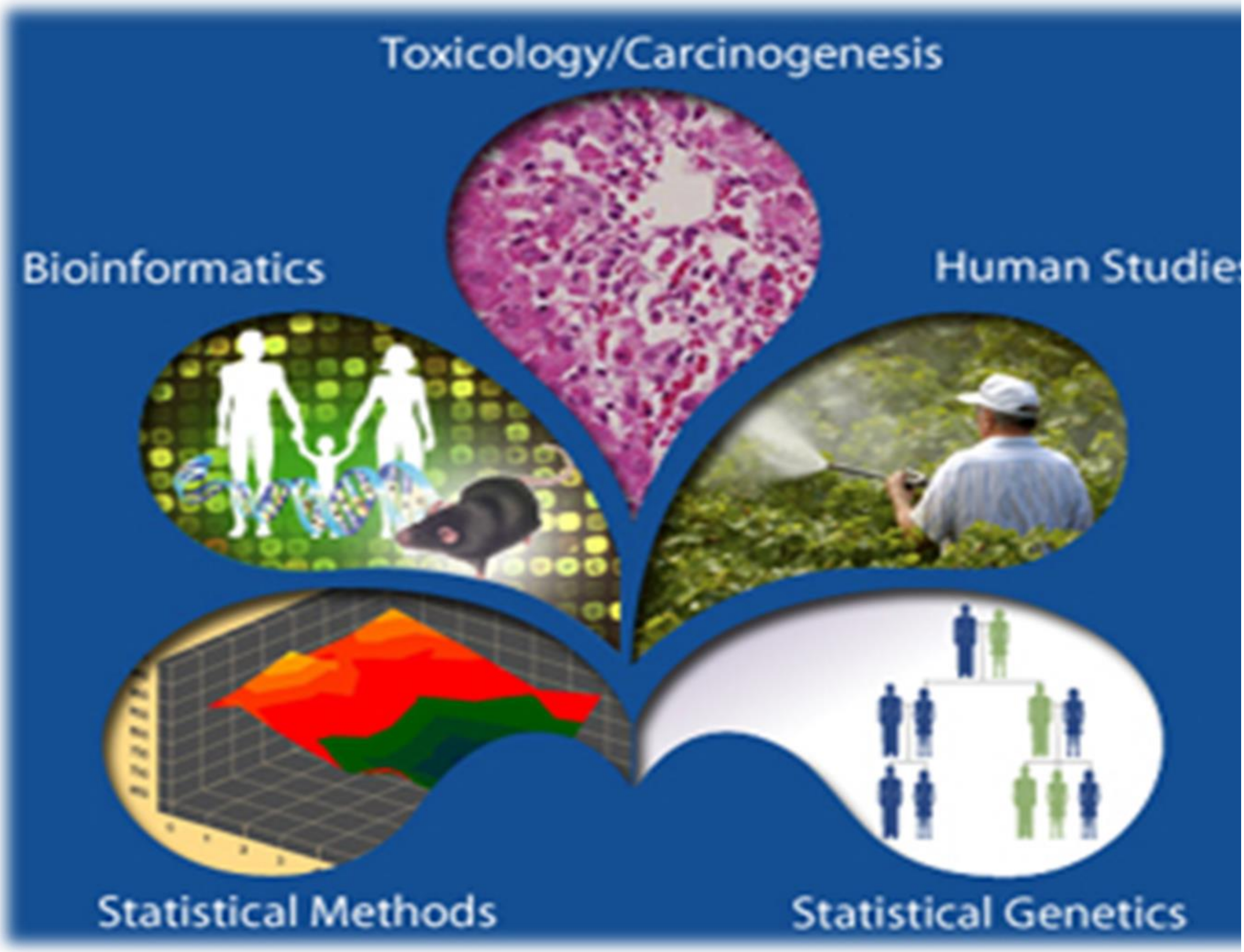
وزارة التعليم العالي والبحث العلمي العراقية
الجامعة التقنية الشمالية



الإحصاء البيولوجي
Biostatistics

الدكتور : عماد توما بني كرش

2018

مقدمة

يهتم علم الاحصاء بطرق جمع البيانات وتبويبها وتلخيصها بشكل يمكن الاستفادة منها في وصف البيانات وتحليلها للوصول الى قرارات سليمة في ظروف عدم التاكد ، فمفهوم علم الاحصاء ليس كما مفهومه شائع بين الناس ، بانه مجرد أرقام وبيانات رقمية فقط ، كاعداد السكان أو اعداد المواليد ، أواعداد الوفيات ، أواعداد المزارعين ، أواعداد المصانع ، أو هو عد أو حصر الاشياء والتعبير عنها بارقام ، فهذه كلها مفاهيم محدودة عن علم الاحصاء.

فالمفهوم الصحيح لعلم الاحصاء هو الطريقة العلمية التي تختص بجمع البيانات والحقائق عن ظاهرة أومجموعة ظواهر معينة وتنظيم وترتيب هذة البيانات والحقائق بطريقة يسهل عملية تحليلها وتقسيمها ومن ثم استخلاص النتائج واتخاذ القرار على ضوء ذلك.

وينقسم علم الاحصاء الى قسمين رئيسين هما: الاحصاء الوصفي والاحصاء الاستدلالي.

ومجالات تطبيق علم الاحصاء يكمن في البحوث الزراعية والبحوث الصناعية والبحوث النفسية والاجتماعية وبحوث الرياضة والشباب والبحوث الطبية والبحوث الاقتصادية والبحوث الهندسية وغيرها من مجالات العلم والمعرفة.

المؤلف

د. عماد توما بني كرش

الجامعة التقنية الشمالية

العراق - نينوى

المحتويات

الـفـصل الـثـالـث

مقاييس النزعة المركزية

الـفـصل الـرابـع

مقاييس التشتت

الفصل الخامس

الارتباط

الفصل الأول

مفهوم علم الإحصاء

1- مفهوم علم الاحصاء

1-1. المقـــــــــدمة

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 للاند 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 : -

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 استخلاص 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 استنتاجات وقرارات مناسبة.

ويمكن تقسيم علم الاحصاء الى قسمين هما :

1 - 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 - الاحصاء الاستنتاجي او الاستدلاليStatistical 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 - جمع البيانات والمعلومات
3 - تصنيف البيانات وتبويبها
4 - عرض البيانات
5 - حساب المؤشرات أو المعالم للبيانات
6 - 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 اطوال الطلبة لأحدى الجامعات فأننا نرمز لصفة الطول (الظاهرة) y ولطول 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 . وتنقسم المتغيرات الى قسمين :-

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

وتنقسم المتغيرات الكمية الى قسمين هما : -

أ - **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 و170 سم اي ان المتغير y يمكن ان ياخذ اي قيمة بين 130,5 و170 سم ومثال ذلك الوزن وكمية المحصول ودرجات الحرارة ، وبصورة عامة كل البيانات التي تقاس تعتبر بيانات لمتغير مستمر (تأخذ القيم عدد صحيح او كسر) .

ب - **متغيرات غير مستمرة (منفصلة)** : - 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 : ـ عبارة عن جميع القيم او المفردات التي يمكن ان يأخذها المتغير ، فلو كانت دراستنا حول اطوال 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** : ـ 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** : ـ 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 ومثال 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 ومال 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 .

1-11. 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 .

1- **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.

1-12. 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
2013	150
2014	180
2015	200
2016	250
مجموع	780

ب - **التبويب الجغرافي** : - 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 الانبار	1800
المجموع	8500

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

والجدول التالي يوضح توزيع الاجور اليومية لعمال احد المصانع

الاجرة اليومية بالدينار	عدد العمال
15000 اقل من	185
اقل من15500	95
اقل من 20000	70
من 20000 فاكثر	20
مجموع	370

د- **التبويب على اساس صفة معينة** : - 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

13-1. 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

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

m = عدد الفئات ، R.T = المدى

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

14-1. 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 يبن عدد المفردات التابعة كل فئة او مجموعة.

مثلا 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.

مثلا الجدول التالي يبين توزيع عدد من الطلبة حسب صفتي الطول والوزن

ا لوزن / 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

مثال (1): لو اردنا عمل توزيع تكراري للأعداد الاتية والتي تمثل الوزن بالكيلو غرامات لعشرين مريضا في مستشفى الموصل .

(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. نجد طول 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).

1-15. الجداول التكرارية المزدوجة

مثال (2) : البيانات الاتية لظاهرتين X و Y المطلوب تفريغها في جدول تكراري مزدوج

X	2	10	11	4	20	15	15	3	25	25	20	22	15	20	30	30	35	30	35	31
Y	3	2	5	6	8	10	2	10	2	6	5	9	15	12	3	9	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

المدى :

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

عدد الفئات :

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

طول الفئة :

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

f_X	f_Y	f_{iX}	f_{iY}

2-8	2-4	|||	3	||||||	6
9-15	5-7	|||||	5	||||	4
16-22	8-10	||||	4	||||||	6
23-29	11-13	||	2	|||	3
30-36	14-16	||||||	6			1
Total		**20**		**20**	

بعد ذلك نضع المعلومات في جدول مزدوج وكالاتي :

X / Y	2-8	9-15	16-22	23-29	30-36	Total
2-4			||				||	6
5-7										4
8-10					||		||	6
11-13									3
14-16							1
Total	**3**	**5**	**4**	**2**	**6**	**20**

16-1. 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 مريضا في احدى المستشفيات.

(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

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

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 المثال رقم (3) فان التوزيع التكراري المتجمع النازل يكون كالاتي :

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

1-17. التوزيع بدائرة كاملة

وهذة أخر طريقة من طرق عرض البيانات الإحصائية غير المبوبة ، وأهم استعمالات هذة الطريقة يتم بتقسيم الدائرة إلى أجزاء فيمثل المجموع الكلي بدائرة كاملة ، ويمثل كل جزء بقطاع دائري .

يكون قياس زاويتة يساوي :

(عدد التكرارات / المجموع الكلي) × 360°

وتكون النسبة المئوية لكل قطاع :

(عدد التكرارات / 100) × 100 %

مثال (4): استخدم شكل القطاعات الدائرية لتمثيل بيانات الجدول الاتي والذي يوضح توزيع عينة من (275) من المرضى النفسيين باحدى المستشفيات خلال سنة باحدى المدن الكبرى ، حسب نوع المرض النفسي .

نوع المرض النفسي	التوهان	الرهاب النفسي	الاكتئاب	الوسواس القهري
عدد المرضى	40	60	150	25

الحل :

مجموع المرضى = 40 + 60 + 150 + 25 = 275

1- نحسب قياس كل قطاع في الدائرة أوقياس زاوية كل قطاع

- قياس زاوية قطاع مرض التوهان = (40 / 275) × 360° = 52°
- قياس زاوية قطاع مرض الرهاب النفسي = (60 / 275) × 360° = 79°
- قياس زاوية قطاع مرض الاكتئاب = (150 / 275) × 360° = 196°
- قياس زاوية قطاع مرض الوسواس القهري = (25 / 275) × 360° = 33°

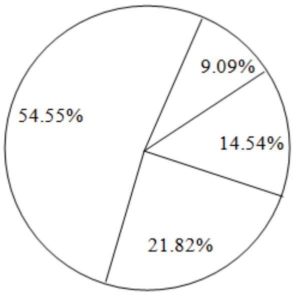
2- نحسب نسبة قياس كل قطاع

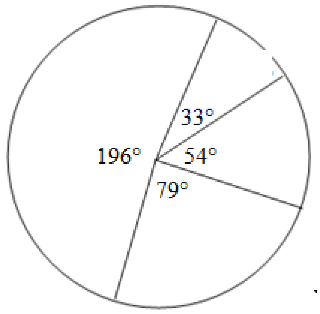
- نسبة قياس قطاع مرض التوهان = (40 / 275) × 100 % = 14.54 %
- نسبة قياس قطاع مرض الرهاب النفسي = (60 / 275) × 100 % = 21.82%
- نسبة قياس قطاع مرض الاكتئاب = (150 / 275) × 100 % = 54.55%
- نسبة قياس قطاع مرض الاكتئاب = (25 / 275) × 100 % = 9.09%

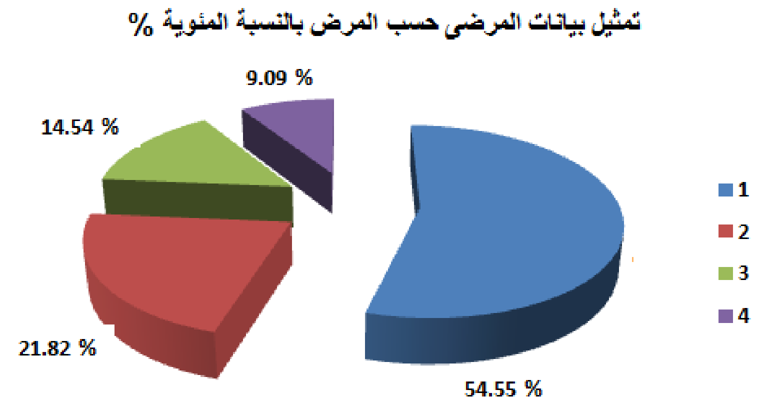
3- نكون جدول بالنتائج

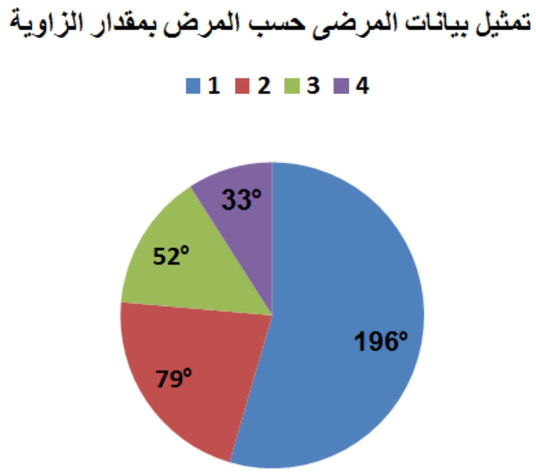
نوع المرض النفسي	التوهان	الرهاب النفسي	الاكتئاب	الوسواس القهري	المجموع
عدد المرضى	40	60	150	25	275
النسبة	14.54%	21.82%	54.55%	9.09%	100%
زاوية القطاع	52°	79°	196°	33°	360°

4- نرسم الدائرة ونوضح القطاعات حسب قياس الزوايا والنسب لكل قطاع كما في الشكل التالي :









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 ؟

س13 / عدد طرق عرض البيانات مع الامثلة ؟

س15 / رتب البيانات الاتية في جدول توزيع تكراري والتي تمثل كميات المواد الفاسدة بالطن لمواد غذائية مختلفة ، والتي تم الكشف عنها في مخازن احد اسواق المواد الغذائية في اربيل ؟

(26 ، 33 ، 21 ، 23 ، 55 ، 44 ، 52 ، 53 ، 36 ، 37 ، 27 ، 51 ، 29 ، 41 ، 48)

س16 / الجدول التالي يبين الضغط العالي والضغط الواطىء لمجموعة عددها 15 من المرضى في احدى المستشفيات ، والمطلوب إنشاء جدول تكراري مزدوج.

الضغط العالي	8	7	12	10	9	15	16	12	10	15	11	18	16	14	13
الضغط الواطىء	9	7	7	6	8	8	10	7	9	8	7	9	7	6	8

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.

1-2. المدرج التكراري

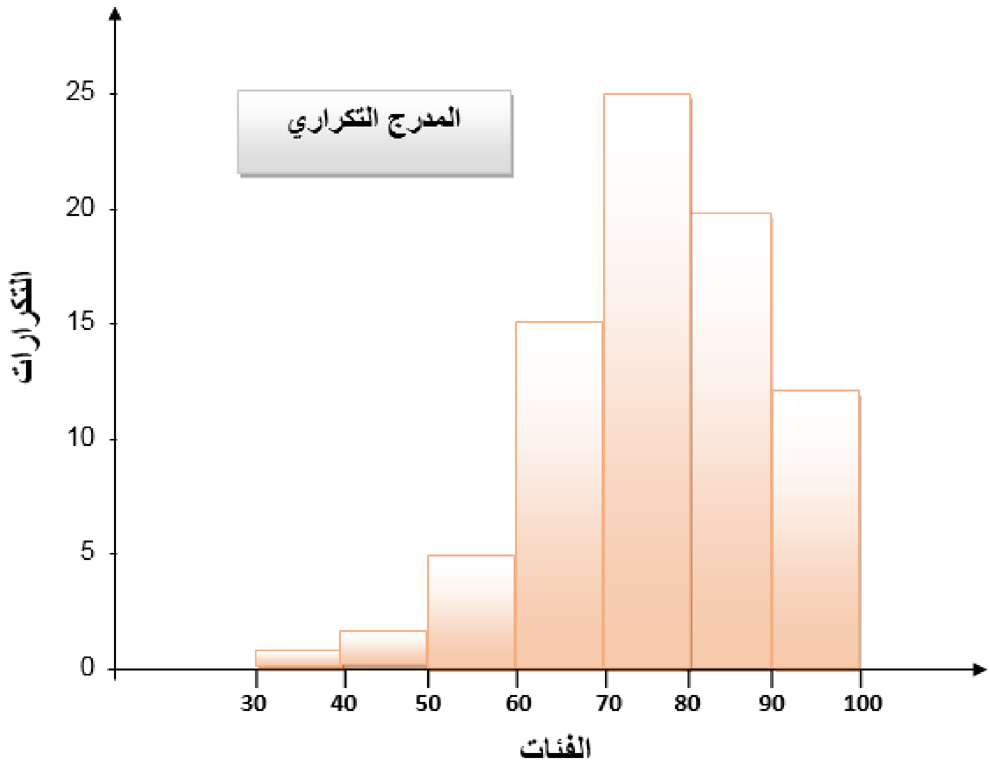
عبارة عن مستطيلات رأسية تمتد قواعدها على المحور الافقي لتمثل اطوال الفئات بينما ارتفاعاتها تمثل تكرارات الفئات. ولرسم مدرج تكراري نتبع الخطوات الاتية :

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
4-

مثال (1) :الجدول الاتي يمثل توزيع تكراري للزيادة في نسبة السكر بالمئة لعينة متكونة من 80 مريض 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:



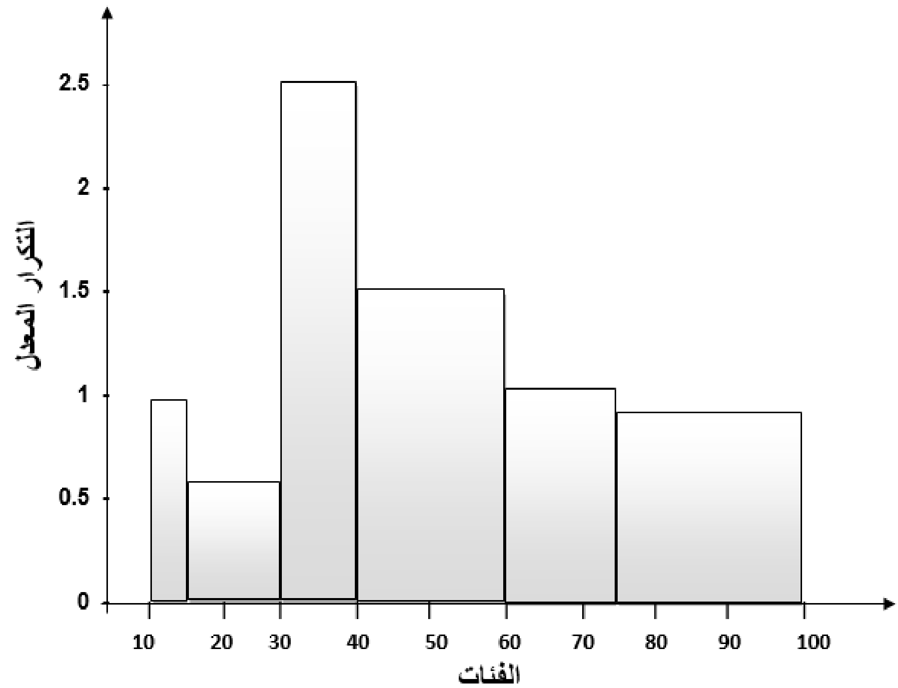
ملاحظة : اذا كانت الفئات غير متساوية عند رسم المدرج التكراري يتم استخراج التكرار المعدل حيث ان التكرار المعدل يساوي تكرار الفئة مقسوم على طول الفئة ويتم اعتماده في المحور العمودي.

مثال (2): التالي توزيع تكراري لفئات غير متساوية من عينة مكونة 104 مريض بالماء الابيض في العين ، لبيان نسبة نجاح عملية زرع العدسات بعد سحب الماء الابيض في المستشفى الامريكي في اربيل ،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 .

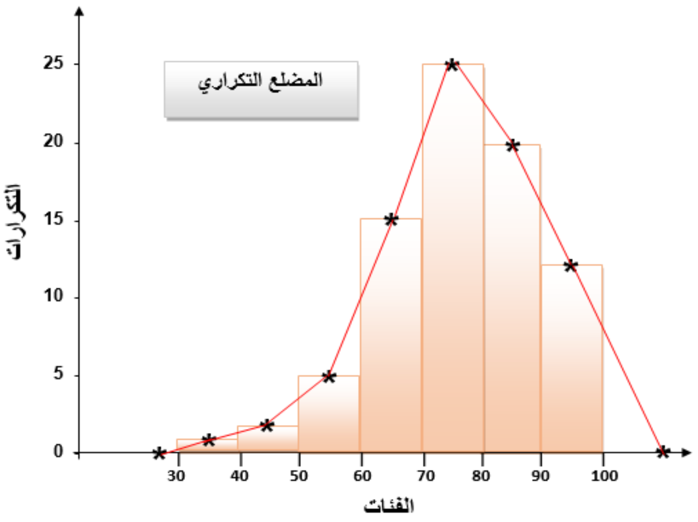
مثال (3) : ارسم ما يلي للبيانات التالية لعينة مكونة من 80 مريض في قسم الطوارء في مستشفى غرب اربيل ، حيث يبين الجدول نسبة خطورة المرضى بالمئة % .

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:



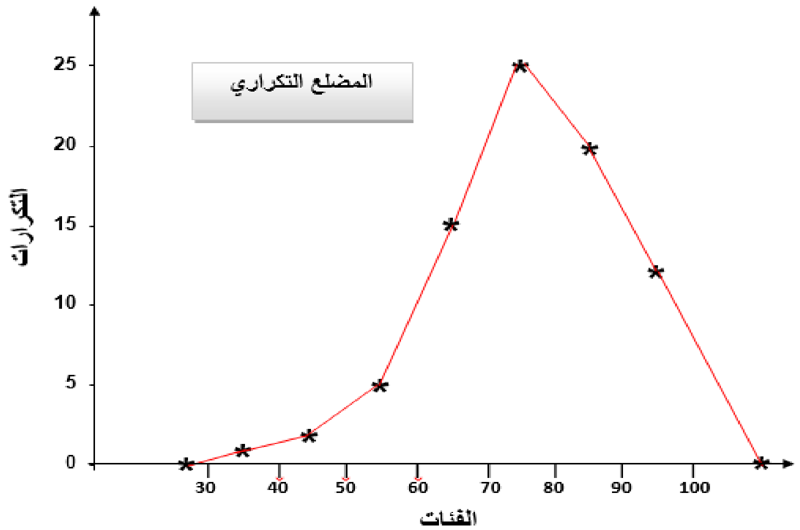
الطريقة الثانية : رسم المضلع التكراري على مراكز الفئات مباشرة دون ضرورة لرسم المدرج التكراري اولا وعليه فان المحور الافقي يمثل مراكز الفئات والمحور العمودي يمثل التكرارات ثم نصل النقاط بعضها ببعض وعليه فان خطوات رسم المضلع التكراري كما يأتي :

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 :

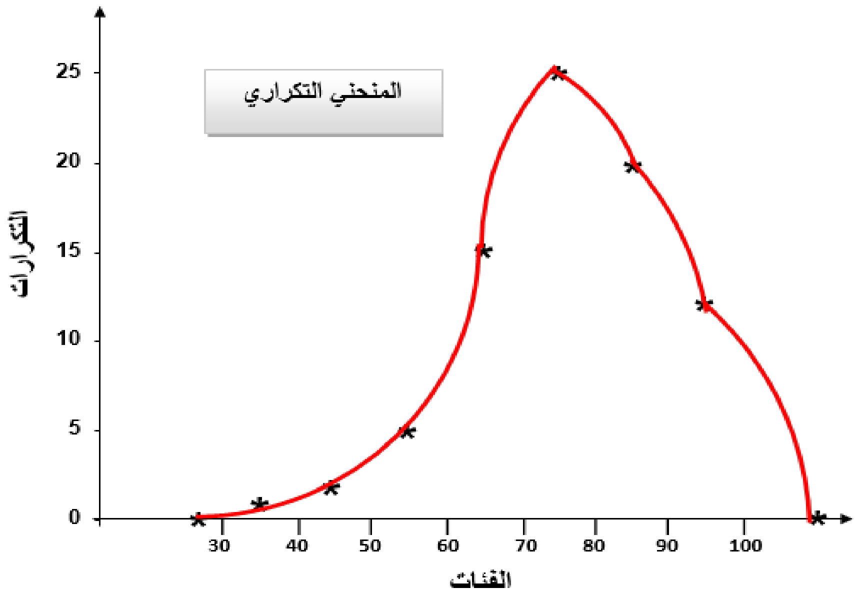
هو طريقة شائعة في الرسم البياني وهو عبارة عن منحنى يمر بمعظم النقاط الواقعة على مراكز الفئات والتي ارتفاعها يمثل تكرارات تلك الفئة

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:



4-2. 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 150 طبيبة لرعاية الحضانة لاطفالهم شهريا بالدولار ، المطلوب رسم منحنى متجمع صاعد لهذا التوزيع ؟

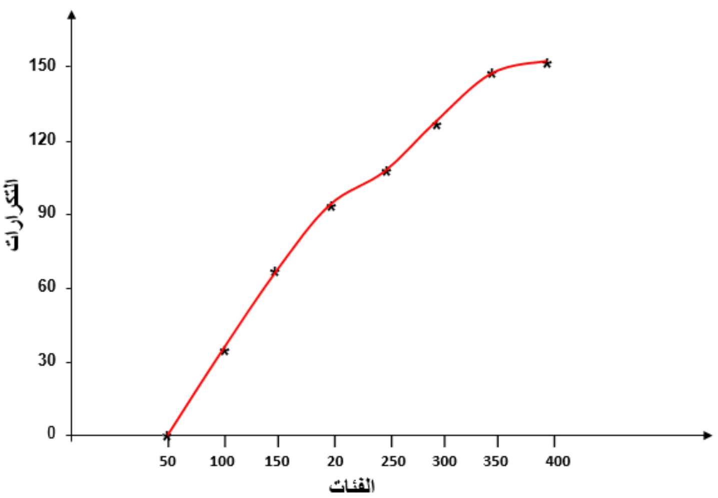
الفئات بالدولار	50-100	100-150	150-200	200-250	250-300	300-350	350-400	**المجموع**
التكرارات اطفال	32	35	25	20	17	13	8	**150**

الحل:

نكون جدول التكرار المتجمع الصاعد وكما يلي :

التكرار المتجمع الصاعد	الحدود العليا للفئات
32	اقل من 100
67	اقل من150
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 150 طبيبة لرعاية الحضانة لاطفالهم شهريا بالدولار، المطلوبرسم منحنى متجمع نازل لهذا التوزيع ؟

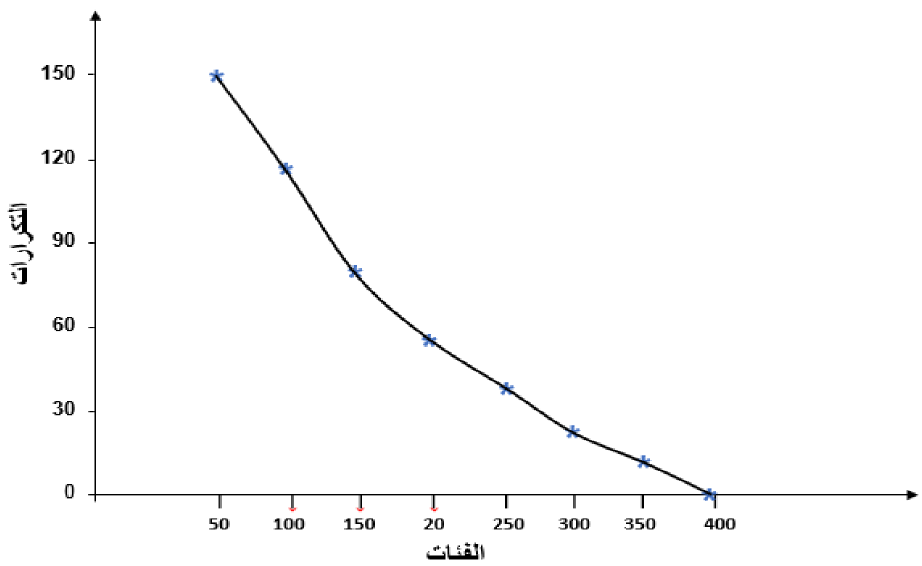
الفئات دولار	50-100	100-150	150-200	200-250	250-300	300-350	350-400	المجموع
التكرارات اطفال	32	35	25	20	17	13	8	**150**

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 150طبيبة لرعاية الحضانة لاطفالهم شهريا بالدولار، المطلوب، رسم منحنى متجمع صاعد ونازل لهذا التوزيع

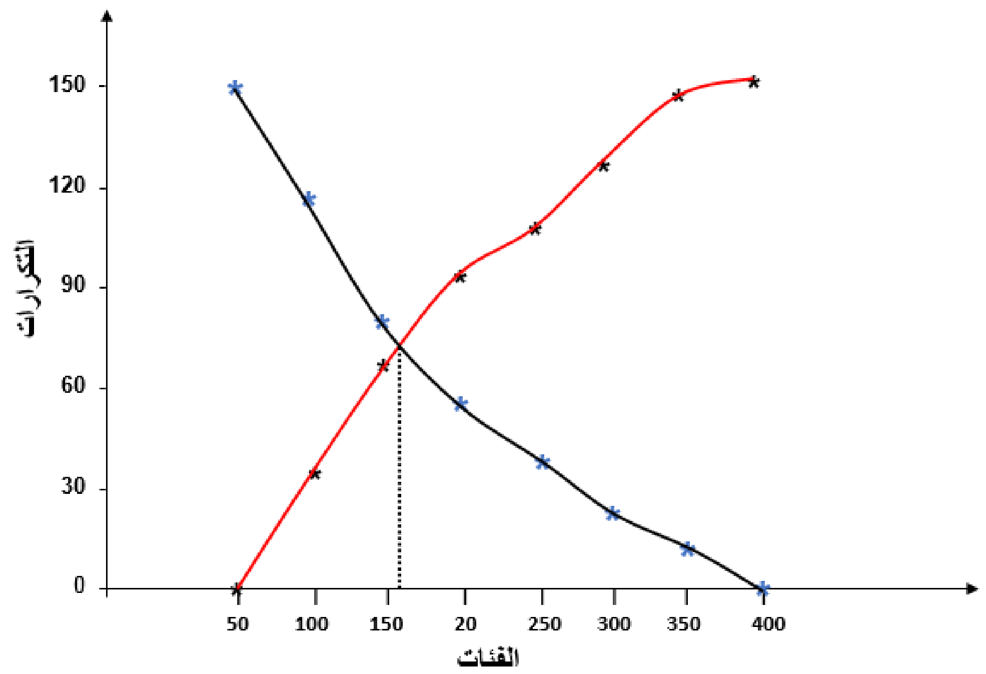
الفئات دولار	50-100	100-150	150-200	200-250	250-300	300-350	350-400	**المجموع**
التكرارات اطفال	32	35	25	20	17	13	8	**150**

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	افل 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	افل 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 اوزان عينة مكونة من 200 مريض في احدى المستشفيات 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 مكونة من 100 موظف صحي في مستشفى اربيل ، 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 الادوية بآلاف الدولارات لإحدى شركات الادوية في الـ70 يوم الأولى من العام 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

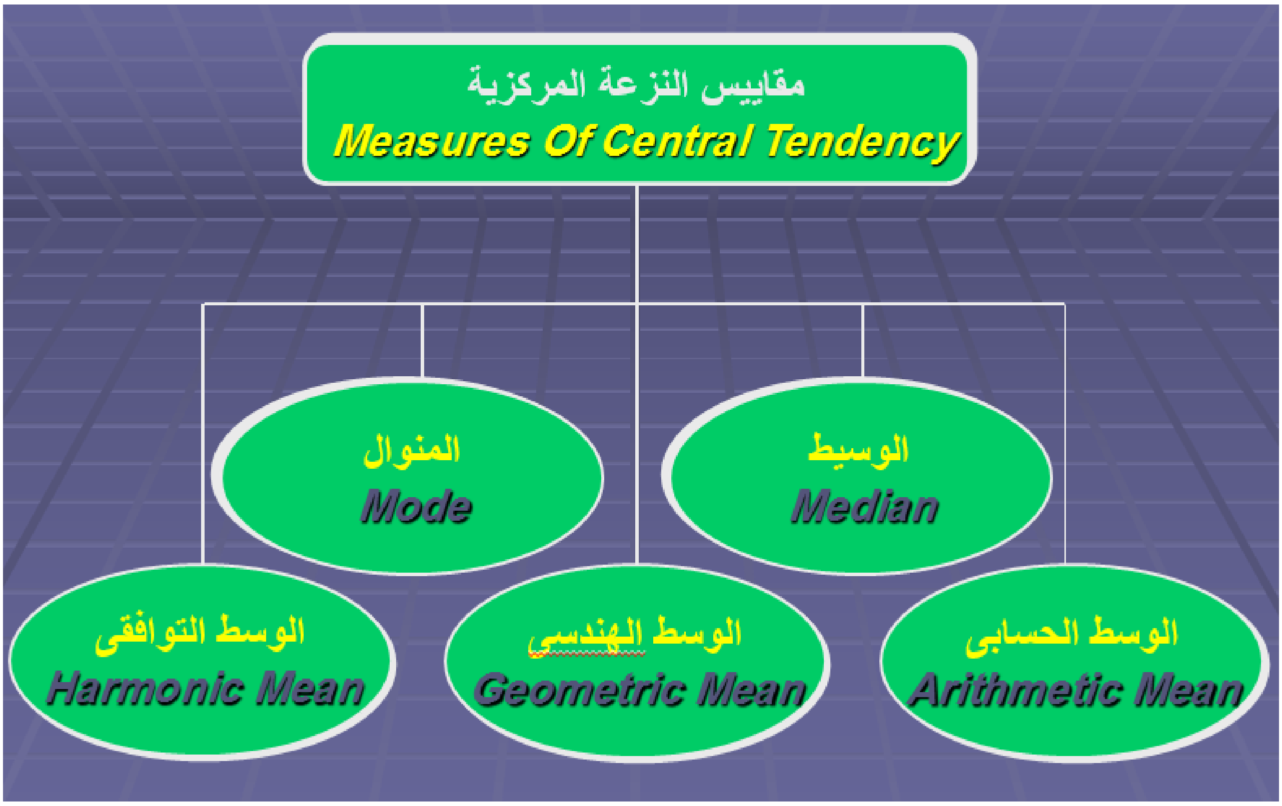
والمطلوب ما يلي :

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 $\ddot{X}$ 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 المرضى قوامها 15 مريض ، المطلوب ايجاد متوسط وزن المريض 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 $\ddot{X}$ 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 المرضى قوامها 15 مريض ، المطلوب ايجاد متوسط وزن المريض 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

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

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 تنظيف في احدى المستشفيات ،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)	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 المثال :

مثال (4) الجدول الاتي يبين توزيع الاجور الاسبوعية لـ 63 عامل في احدى المستشفيات ، 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	**<u>75</u>**	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

3-2. 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

3-2-1. 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 , **<u>134</u>** , 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 :

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

2-2-3. الوسيط لبيانات مبوبة :

1. **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 المدارس 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 الاسر ة ، 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:

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

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

اي الوسيط يقع بين التكرارات **27** و**47**

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

الاقرب الى **47** وان الوسيط يمثل مركز هذه الفئة

وعليه فالوسيط يساوي **12** وكالاتي :

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

2. 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 :

الوسيط = Me

الحد الادنى لفئة الوسيط = L_k

مجموع التكرارات الاصلية = $\sum f_i$

تكـــرار متجمع صاعد سابق = F_{k-1}

تكرار متجمع صاعد لاحق = F_k

طول فئة الوسيط = h_k

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 :

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

الحل:

ترتيب الوسيط :

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

الوسيط يقع بين التكرارات (18 و 34) $\longleftarrow$ 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 .

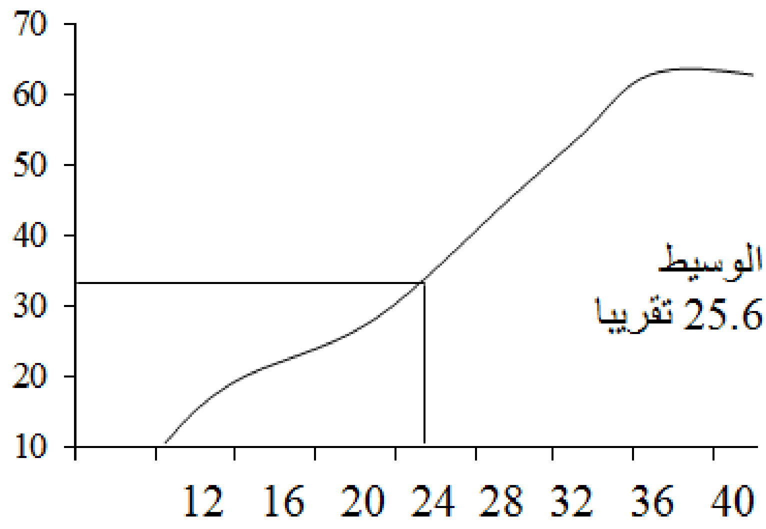
وإذا قمنا باستخدام إحدى الطريقتين الأولى والثانية فإننا :

1- نوجد ترتيب الوسيط .

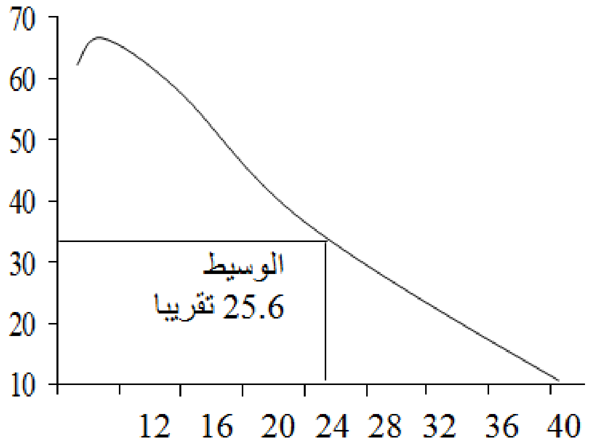
$$T = \frac{\sum f_i}{2}$$

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.

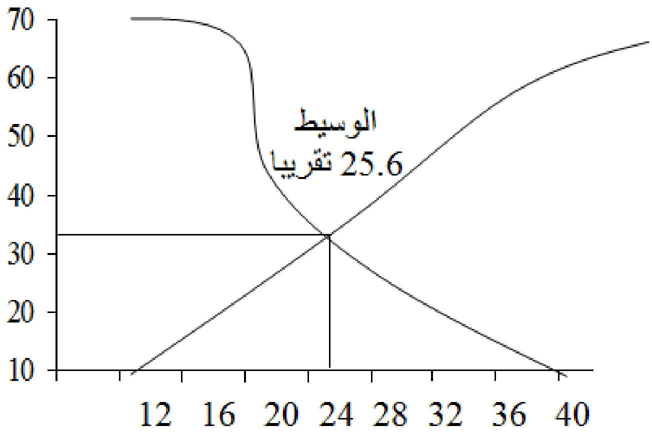
أما إذا استخدمنا الطريقة الثالثة (المنحنيان معا) فنقوم برسم المنحنين المتجمعين الصاعد والهابط معا فيتقابل المنحنيان معا في نقطة واحدة ، نسقط منها عمود يلاقى المحور السيني في نقطة هي تمثل قيمة الوسيط .



إيجاد الوسيط بيانياً باستخدام التكرار المتجمع الصاعد .



أيجاد الوسيط بيانياً باستخدام التكرار المتجمع الهابط .



استخدام المنحنى الصاعد والمنحنى الهابط

3-2-4. 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

مثال (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 منوال في المجموعة في حالة اكثر من رقم .

مثال (10) : احسب المنوال للبيانات التالية

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

لا يوجد منوال لهذه البيانات لأنه لا يوجد رقم متكرر

مثال (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** منوال 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 :

المنوال = 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 لعينة مكونة من اوزان 50 مريض 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 لعينة مكونة من 68 مريض والتي تمثل زيادة نسبة الكولسترول (Cholesterol) في الدم اوجد المنوال؟

الفئات نسبة زيادة الكولسترول	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

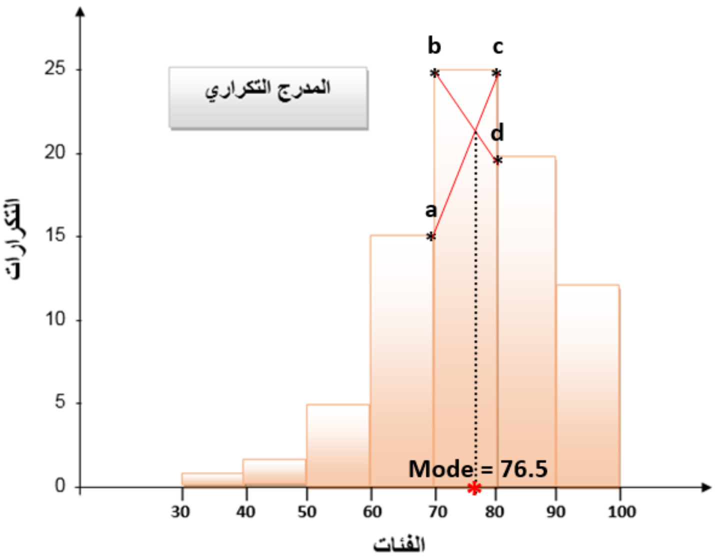
يمكن تقدير المنوال بطريقة الرسم البياني للمدرج التكراري وذلك باستعمال مستطيل الفئة المنوالية (وهو اعلى مستطيل لأنه يمثل اكثر التكرارات) والمستطيلان المجاوران له.

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

مثال (14): الجدول الاتي يمثل توزيع تكراري لاجور اسبوعية لمنتسبي احدى المستشفيات بالدولار، المطلوب ايجاد المنوال 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 حساب المنوال 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 - المنوال = 3 (الوسط الحسابي - الوسيط)

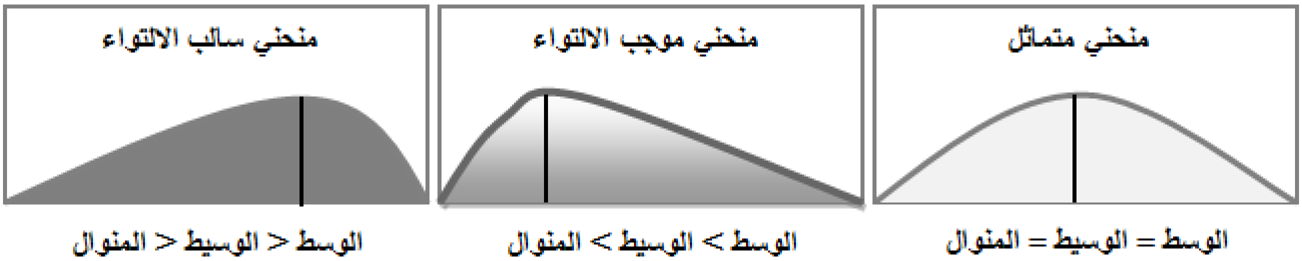
وقد وجد ان الوسيط تقع قيمته بين قيمتي الوسط الحسابي والمنوال .

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

الوسط الحسابي = الوسيط – المنوال

6-3-3. 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**) : في مختبر التحليلت المرضية تم سحب عينات دم من 10 مرضى في مستشفى اربيل ، ذات الحجم 5 مليلتر ، لفحص نسبة السكر المتراكم في الدم، وكانت نسبة السكر كالتالي :

115 123 119 123 124 119 123 121 123 121

والمطلوب : حساب الوسط الحسابي، والوسيط، والمنوال، ثم حدد شكل الالتواء لنسبة السكر في دم المرضى .

الحل

1- حساب الوسط الحسابي :

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

2- حساب الوسيط :

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

ترتيب القيم تصاعديا

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

عدد القيم = **10** ، وهو عدد زوجي.

الوسيط = الوسط الحسابي للقيمتين رقم **(5 , 6)**

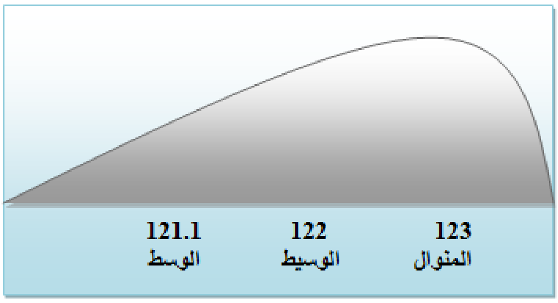
$$Med = \frac{121+123}{2} = \frac{244}{2} = 122$$

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 نسبة السكر المتراكم في الدم سالبة الالتواء.

مثال (16) **:** الجدول التكراري التالي يعرض توزيع 100 منتسب بعقد موقت في مستشفى اربيل ، حسب الأجر الاسبوعي 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 الحسابي 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

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

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	**23** $\leftarrow f_1$	(رتبة الوسيط 50)
أقل من 110	**51** $\leftarrow f_1$	
أقل من 130	**71**	
أقل من 150	**86**	
أقل من 170	**94**	
أقل من 190	**100**	

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 (90 - 110) .

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

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

إذا المنوال يحسب بتطبيق المعادلة التالية :

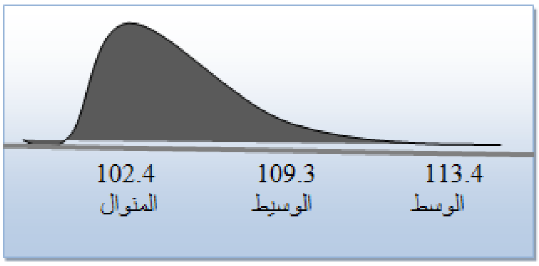
$$Mod = A + \frac{d_1}{d_1 + d_2} \times L = 90 + \frac{13}{13 + 8} \times 20 = 90 + \frac{260}{21} = 102.4 \ \ R.S$$

3-3-7. بيان شكل التوزيع .

من النتائج السابقة ، نجد أن :

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

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



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

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

1- الأدلة indexes
2- النسب ratios
3- المعَدَّلات 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

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

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

مثال(17): ازدادت أرباح مستشفى اهلي من 200 ألف دينار عراقي عام 1995 إلى 350 ألف دينار عراقي عام 2000، فإن المتوسط الهندسي في هذه الفترة هو:

$$G = \sqrt[n]{X_1F_1 \times X_2F_2 \times ... \times X_nF_n} = \sqrt{200000 \times 350000} = 264575$$

ثانيا : في حالة البيانات المبوبة-:

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

$$G = \sqrt[n]{X_1F_1 \times X_2F_2 \times ... \times X_nF_n}$$

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

مثال (18): اوجد الوسط الهندسي للبيانات التالية

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

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**

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

$$G=14.58$$

2-4-3. الوسط التوافقي ***Harmonic mean***

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

أولا: حساب 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: 8 , 8 , 7 , 4 , 3 , 5 , 2

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

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

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

H= 5.28

ثانيا: حساب 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

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

مثال (21) احسب الوسط التوافقي من الجدول التالي والذى يوضح التوزيع التكراري لسرعات 100مريض؟

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

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**

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

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 والمنوال 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

س12/ الاتي توزيع تكراري للكلسترول المتراكم لعينة من المرضى قوامها 80 اسرة ، المطلوب ايجاد الوسط الحسابي ، الوسيط والمنوال للكلسترول 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

س13/ الاتي توزيع تكراري لأطوال عينة من الاشخاص البالغين قوامها 50 شخص ، المطلوب حساب الوسط الحسابي ، الوسيط والمنوال 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 حدد شكل 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 الرابـع

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.

مثال(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

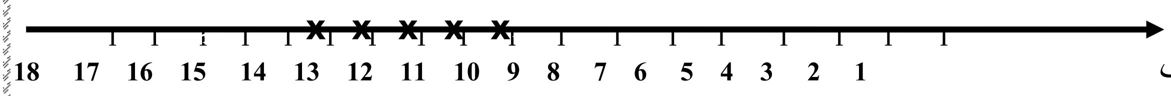
$$\ddot{X} = \frac{\sum x}{n} = \frac{3+8+9+10+11}{5} = \frac{45}{5} = 9$$

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

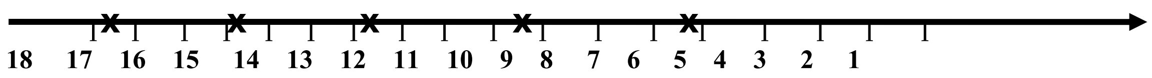
$$\ddot{X} = \frac{\sum x}{n} = \frac{1+5+9+13+17}{5} = \frac{45}{5} = 9$$

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

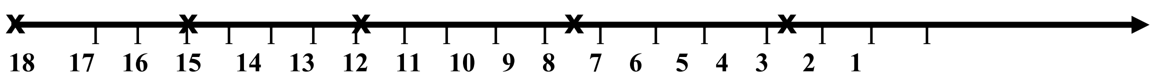
المجموعة الاولى



المجموعة الثانية



المجموعة الرابعة



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. المـــــدى :

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

$$R = U - L$$

مثال (1): الاتي بيانات غير مبوبة ، المطلوب ايجاد المدى لهذه البيانات

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

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

مثال (2): بيانات مبوبة من خلال جدول التوزيع التكراري التالي اوجد المدى؟

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

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 بالرمزQ_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 %75من بداية البيانات ويجب هنا ايضا ان يعاد ترتيب مجموعة القيم تصاعديا او تنازليا قبل حساب موقع او ترتيب الربيعين

مثال (3): 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 , **<u>3</u>** ,5 , 7 , 8 , **<u>10</u>** , 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: $\sum fi$ هي مجموع التكرارات

3. تحديد ترتيب الربيع الاعلى بالصيغة الاتية :

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

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 :

Q_1 الربيع الادنى

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(4) : احسب الانحراف الربيعي للتوزيع التكراري الاتي لنسبة شفاء 50 مريض من مرض السل الرئوي؟

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

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

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

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

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

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

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

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(4) : الجدول الاتي يبين رواتب موظفي احد المستشفيات بالدولار ، المطلوب : ايجاد الانحراف الربيعي بالرسم البياني؟

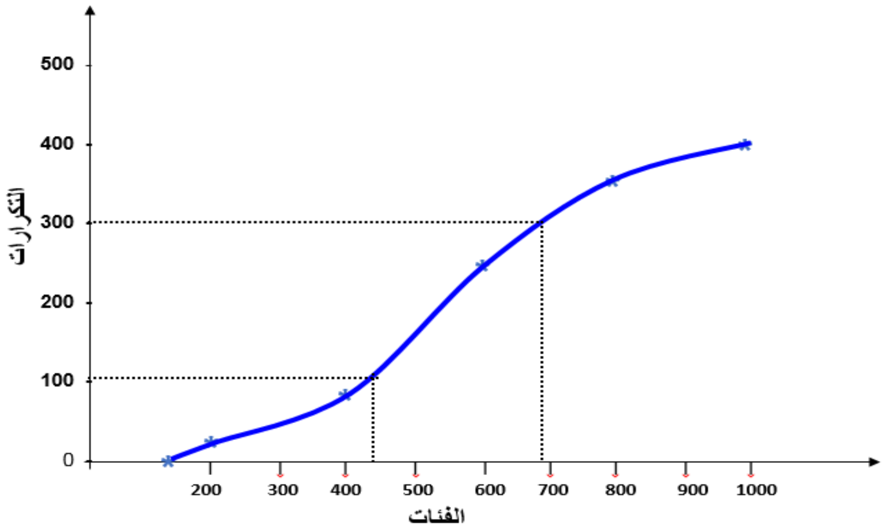
الفئات	0-200	200-400	400-600	600-800	800-1000	المجموع
التكرارات	18	72	154	111	45	400

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 700وان 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.

1-4-4. التباين في المجتمع (σ^2)

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

$$\sigma^2 = \frac{\sum (x-\mu)^2}{N}$$

حيث أن μ هو الوسط الحسابي في المجتمع .

أي أن : $\mu = \sum x / N$.

مثال(5) : مختبر للتحليلات المرضية ، يعمل به 15 موظف ، وكانت عدد سنوات الخبرة لهؤلاء الموظفين 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 .

الحـل

لحساب تباين سنوات الخبرة في المجتمع ، يتم استخدام المعادلة

$$\sigma^2 = \frac{\sum (x-\mu)^2}{N}$$

- الوسط الحسابي في المجتمع μ

$$\mu = \frac{1}{N}\sum x$$
$$= \frac{1}{15}(5+13++7+\ldots+12+10) = \frac{1}{15}(150) = 10$$

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

سنوات الخبرة x	$(x-\mu)$	$(x-\mu)^2$
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: $\sum(x-\mu)^2 = 130$

إذا تباين سنوات الخبرة للعمال في المصنع هو :

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

4-2-4. التباين في العينة (S^2)

في كثير من الحالات يكون تباين المجتمع σ^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 ، $x_1, x_2, ..., x_n$ ، 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 .

مثال(6) : إذا تم سحب عينة من اطباء في مستشفى اربيل حجمها 5 اطباء ، وسجل عدد سنوات الخبرة ، وكانت كالتالي .

8 13 10 5 9

احسب تباين سنوات الخبرة في العينة .

الحــل

لحساب التباين في العينة يتم تطبيق المعادلة ويتبع الآتـي :

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

- 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:

$$Standard \;\; Deviation = \sqrt{Variance}$$

مثال(7) : عيادة طبية ، يعمل بها 15 طبيب ، وكانت عدد سنوات الخبرة لهؤلاء الاطباء 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 للاطباء (المجتمع) ، ويرمز له بالرمز (σ) هو :

$$\sigma = \sqrt{\frac{1}{N}\sum x^2 - \mu^2}$$

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

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 .

1-5-4. 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(8) : البيانات التالية تمثل اوزان عينة من المرضى في مستشفى رزكاري قوامها 10 مرضى ، 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

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

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

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(9) : البيانات التالية تمثل اوزان عينة من الطلبة قوامها 10 طلاب ، المطلوب حساب قيمة الانحراف المعياري ، بالطريقة المختصرة : 56, 68, 72, 63, 65, 68, 71, 69, 62, 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
$\sum X=650$	$\sum X^2=42544$

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

2-5-4. حساب الانحراف المعياري لبيانات مبوبة

A- 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

$(Xi-\ddot{X})$ 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 $\left(X_i - \ddot{X}\right)$

4- نربع انحرافات مراكز الفئات عن الوسط الحسابي

5- نضرب التكرارات في مربعات الانحراف

6- نجمع حاصل الضرب

7- نطبق صيغة القانون

مثال(10) : الاتي جدول توزيع تكراري لنسبة الزيادة بالمئة في دهون الكبد لعينة مكونة من 50 مريض في مستشفى رزكاري ، 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 -Ẍ)	$(Xi - \ddot{X})^2$	$Fi(Xi - \ddot{X})^2$
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)

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

B - الطريقة المختصرة نستخدم الصيغة الاتية :

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

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

مثال(11) : الاتي جدول توزيع تكراري لنسبة الزيادة بالمئة في هرمون الغدة الدرقية لعينة مكونة من 50 مريض في مستشفى اربيل ، 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

الخطـــــــــــوات

الفئات	f_i	X_i (1)	$f_i X_i$ (2)	X_i^2 (3)	$f_i X_i^2$ (4)
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
$\sum$	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(12) : احسب معامل الاختلاف بطريقتين لدرجات مادة التشريح 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

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

$Q_1 = 58.39$; $Q_2 = 73.14$

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

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:

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

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

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

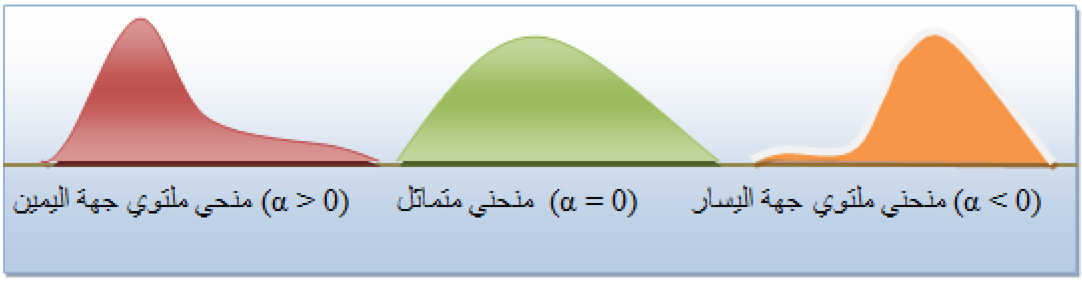
حيث أن α (ألفا) هو معامل الالتواء "لبيرسون"، $\overline{X}$ الوسط الحسابي، Med هو الوسيط، S هو الانحراف المعياري، ويمكن من خلال الإشارة التي يأخذها هذا المعامل الحكم على شكل الالتواء، كما يلي :

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

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

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

أشكال التواء البيانات



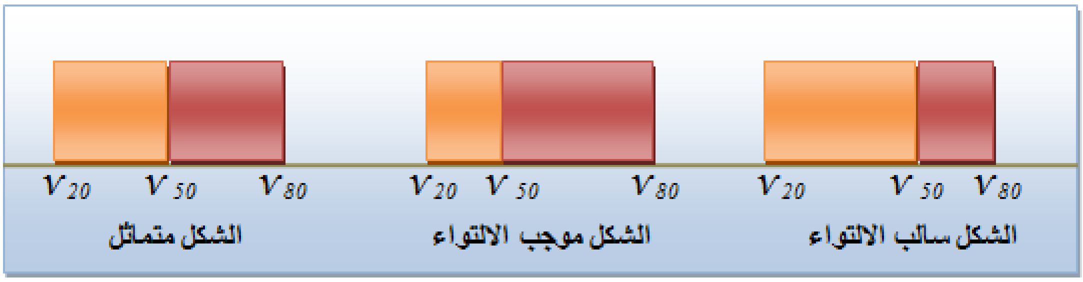
4-7-1-2. طريقة "المئين" في قياس الالتواء

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:

$$v_p = x_l + (R - l)(x_u - x_l)$$

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.

$$\alpha_{p,100-p} = \frac{(v_{100-p} - v_{50}) - (v_{50} - v_p)}{v_{100-p} - v_p}$$

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 :

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

مثال(12) : كان الضغط الواطىء لعينة مكونة من **8** مرضى كمايلي.

66 85 52 78 80 91 74 58

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

2- حساب معامل الالتواء الربيعي .

الحــل

1- حساب معامل الالتواء بطريقة "بيرسون" .

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

1- حساب الوسط الحسابي ، والانحراف المعياري :

ويكون :

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

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

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

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

$$= \sqrt{\frac{1258}{7}} = \sqrt{179.71428} = 13.406$$

2- حساب الوسيط **:**

موقع الوسيط : $(n+1)/2=(8+1)/2=4.5$

$$\begin{array}{cc|cc|cc|cc} 52 & 58 & 66 & 74 & 78 & 80 & 85 & 91 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array}$$

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- حساب 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 :

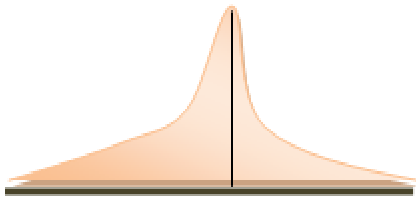
$$\alpha_q = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_1)} = \frac{(83.75 - 76) - (76 - 60)}{(83.75 - 60)}$$
$$= \frac{-8.25}{23.75} = -0.35$$

إذا توزيع درجات الطلاب ملتوي جهة اليسار.

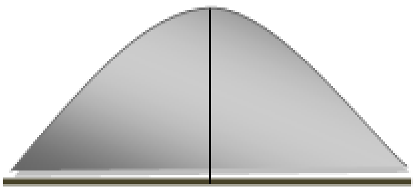
2-7-4. التفرطح Kurtosis

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) .

مثال(13) : كان الضغط الواطى لعينة مكونة من 8 مرضى في احدى المستشفيات ، كلاتي.

66 85 52 78 80 91 74 58

والمطلوب : إيجاد شكل التفرطح ؟

وبالتطبيق على بيانات المثال رقم (12) 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/ اذكر طرق حساب التباين للبينات المبوبة وغير المبوبة مع ذكر العلاقات الرياضية ؟

س6/ ماهي مزايا والانحراف المعياري؟

س7/ ماهي مزايا وعيوب الاختلاف؟

س8/ الاتي درجات الحرارة لمريض خلال يومين ، المطلوب ايجاد المدى لدرجات حرارة جسم المريض ثم قارن ايهما اكثر تجانسا خلال يومين

اليوم الاول 36 , 35.5 ,36.7 , 37 , 38 , 39 , 37.5 , 38.5

اليوم الثاني 37 , 36.5 , 38 ,37.5 , 36 , 39 , 39.5 , 37.4

س9/ ألجدول التكراري التالي يبين اوزان 50 مريض ، 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 المستشفيات وعدد المنتسبين الذين يعملون في كل مستشفى ، المطلوب ايجاد درجة التشتت في عدد المنتسبين الذين يعملون في كل مستشفى باستخدام:

1- الطريقة المطولة

2- الطريقة المختصرة؟

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

س14/ للجدول التكراري الآتي ، أوجد ما يلي :

1- المدى

2- الانحراف الربيعي

3- التباين

4- الانحراف المعياري

5- الاختلاف

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

الفصل الخامس

الأرتباط

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 ـــــــــ معامل ارتباط بيرسون

رتبيه- رتبيه ـــــــــ معامل سبيرمان

كمية- رتبيه ـــــــــ 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 او سالب) 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 :

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

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 1: تزوج رجل (X) ذوفصيلة دم A^+ مع امراة (Y) ذو فصيلة دم B^+ ، احسب معامل الارتباط بين الفصيلتين.

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 وللمتغير 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 وقوي بين فصيلة الدم A^+ و فصيلة الدم B^+ .

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 2 : البيانات الاتية تمثل كمية العلاج (X) والاستهلاك (y) بالدينار العراقي ، المطلوب حساب العلاقة بين كمية العلاج 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

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

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

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

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

العلاقة بين كمية العلاج والاستهلاك عالية جدا وقوية موجبة وقريبة من الواحد

3-5. 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.

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

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

حيث ان :

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

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

مثال 3 : الاتي تقديرات 6 طلبة في امتحان مادتي امراض الدم و الاحصاء الحياتي ، 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.

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

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 وy .

مثال 4 : الاتي تقديرات لكفاءة عشرة من الاطباء في احد المستشفيات من حيث مقدرتهم في علاج نوعين من الامراض ، المطلوب حساب معامل الارتباط البسيط بين المرضين:

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

تقديرات نوع **y** من المرض : 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 XوY من خلال احتساب التعديلات وذلك بإضافة الكمية $m(m^2-1)\backslash 12$
1. نجمع التعديلات الكلية للمتغيرين معا
2. نحسب معامل ارتباط الرتب لسبيرمان المعدل من الصيغة الاتية :

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

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 وعليه 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 = **7**

وبذلك معامل ارتباط الرتب لسبيرمان المعدل يكون كالاتي:

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

الارتباط متوسط القوة

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 fij Vj Ui 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 و y 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 5 : الاتي توزيع تكراري مزدوج لدرجات 100 طالب في مادتي الاحصاء X والتشريح Y ، المطلوب حساب معامل الارتباط البسيط ما بين X و Y

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
Total	6	16	31	21	26	100

خطوات الحل :

1. نجد مركز فئة X
2. نجد Ui حيث ان $U_i = \dfrac{X_i - A}{L_i}$
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 $V_i = \dfrac{Y_j - B}{L_j}$
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 نطبق صيغة القانون

X \ Y	*30-40*	*40-50*	*50-60*	*60-70*	*70-80*	*fi*	*Xi*	*Ui*	*Uifi*	*Ui²fi*	*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					
Vj²fj	24	16	0	21	104	∑156					
VjUifij	**14**	**11**	**0**	**26**	**92**	**∑143**					

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

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

الارتباط متوسط

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 $\mathbf{y_i}$ متغيرين من النوع الوصفي وعليه يمكن تصور جدول التوافق ذو المرتبة 2*2 وبالشكل التالي :

X \ Y	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 6 : من المعلوم ان عادة التدخين تؤثر تأثيرا سيئا على الصحة العامة للفرد ، المطلوب ايجاد العلاقة بين الحالة الصحية وعادة التدخين .

الحل :

عادة التدخين / الحالة الصحية	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

مثال 7 : الجدول التالي يبين نوع الدواء المستخدم لعلاج حالة مرضية معينة ونتائج العلاج ، المطلوب ايجاد معامل الاقتران لهذا التوزيع ؟

Y / X	نوع B	نوع A	مجموع

جيدة	25	12	37
متوسطة	10	50	60
سيئة	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

$$r = r_1 + r_2 + r_3$$

مثال 8 : 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
ضباب	20	45	40	105
المجموع	55	107	84	246

الحل :

1. نجد r_1 من الصيغة الاتية :

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

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

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

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

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

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

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

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

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

العلاقة 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 6 طلبة في مادتي الطيفيليات والتشريح ، المطلوب ايجاد معامل الارتباط بين تقديرات الطلبة

الطيفليات X	متوسط	جيد	مقبول	ضعيف	امتياز	جيد جدا
التشريح y	جيد	متوسط	ضعيف	مقبول	جيد جدا	امتياز

س4/ افرض ان اثناء وجود 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/ 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

س6/ 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)

الرائحةY / 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

العينات

1-6. العينة Sample

العينة هي جزء من المجتمع، أي هي جزء من الكل، على أن يكون هذا الجزء ممثلاً للكل. فالعينة يتم اختيارها ـ عادة ـ بهدف تعميم النتائج التي يحصل عليها الباحث منها على المجتمع بأكمله بعد ذلك ، ولذا يجب أن تكون العينة ممثلة للمجتمع المسحوبة منه تمثيلاً صادقاً حتى يتسنى للباحث استخدام بيانات ونتائج العينة في تقدير معالم المجتمع بشكل جيد. وسوف نتناول بالتفصيل أهم أنواع العينات ومتى وكيف يتم اختيار كل منها وحدود ذلك:

1- العينة الكبيرة Large Sample

تكون العينة كبيرة إذا كان حجمها أو عدد مفرداتها يساوي (30) مفردة أو أكثر. مع ملاحظة أنه كلما كانت العينة كبيرة بدرجة كافية كلما كان ذلك أفضل ، حيث يتمكن الباحث – في الغالب – من استخدام الكثير من أساليب التحليل الكمي والإحصائي ، أي كلما كان حجم العينة أكبر من 30 كلما كان ذلك أفضل.

2- العينة الصغيرة Small Sample

تكون العينة صغيرة إذا كان حجمها أو عدد مفرداتها أقل من (30) مفردة. وهناك أساليب تحليل إحصائي خاصة إذا كانت العينة صغيرة.

2-6. إحصاءات العينة Sample Statistics :

إحصاءات العينة هي المقاييس الإحصائية الخاصة بالعينة، أي التي تصف العينة بالاعتماد على القياس الكمي لمفردات العينة مثل:

1- الوسط الحسابي للعينة Mean .

2- تباين العينة (الانحراف المعياري لها) Standard Deviation .

3- النسبة في العينة (Ratio أو Proportion) .

وإحصاءات العينة هي التي تستخدم عادة في الاستدلال على معالم المجتمع. أي هي التي تستخدم في تقدير معالم المجتمع أو في اختبار الفروض حولها.

ونتيجة لما يقدمه أسلوب العينات من مزايا فإنه أصبح واسع الانتشار، وأصبحت له استخدامات كثيرة في مختلف المجالات. وفيما يلي نتنأول مزايا العينات، وأيضاً العيوب (أو المحاذير) التي يجب الانتباه إليها عند الدراسة بهذا الأسلوب.

3-6. مزايا العينات :

يمكن تلخيص مزايا الدراسة بأسلوب العينات فيما يلي:

1- توفير الوقت والجهد والتكاليف.

2- الحصول على معلومات أكثر تفصيلاً من تلك التي نحصل عليها من مفردات المجتمع لا سيما إذا كان حجم المجتمع كبيراً. فاختيار عينة يساعد الباحث على تركيز الجهد والدقة في جمع البيانات وبشكل أكثر تفصيلاً وبالتالي قد تكون نتائج العينة أكثر دقة من النتائج التي يحصل عليها الباحث باستخدام الحصر الشامل. كما أنه يمكن تقدير الأخطاء التي تتعرض لها العينات باستخدام النظريات الإحصائية الأمر الذي يمكن الباحث من تقدير دقة النتائج التي يحصل عليها باستخدام العينات.

3- هي الأسلوب الوحيد في حالة المجتمعات غير المحدودة حيث يستحيل حصر ودراسة جميع المفردات كدراسة الأسماك في البحر الأحمر مثلاً.

4- هي الأسلوب الوحيد المناسب في الحالات التي يترتب على الدراسة أهلاك وإتلاف المفردات. مثل، فحص دم مريض، اختبارات الجودة لكثير من المنتجات التي تؤدي إلى عدم صلاحيتها مرة أخرى مثل المعلبات، والمصابيح الكهربائية، إنتاج مزرعة من البيض.. الخ.

5- كذلك فإن العينات تناسب الظواهر التي تتغير بسرعة أي ذات الطبيعية المتغيرة باستمرار. في هذه الحالات يكون الأفضل استخدام العينات بدلاً من الحصر الشامل الذي يتطلب وقتاً أطول تكون الظاهرة أثناء هذا الوقت تغيرت أكثر من مرة وبالتالي تكون نتائج الحصر الشامل لا تعبر عن الظاهرة عند الانتهاء من الدراسة.

4-6. عيوب العينات :

تتلخص عيوب العينات فيما يلي :

1- تتوقف نتائج الدراسة بالعينات على مدى تمثيل العينة للمجتمع. فإذا كانت العينة غير ممثلة للمجتمع تمثيلاً صادقاً كانت النتائج غير دقيقة. لذلك ينصح دائماً بإتباع الأسلوب العلمي السليم عند اختيار العينة.

2- نتائج دراسة العينة قد تكون غير نهائية، وقد تكون في حاجة إلى تعميم. وسوف تختلف النتائج باختلاف نوع العينة حتى بإتباع الأسلوب العلمي في الاختيار.

3- تتعرض دراسة العينات عموماً لنوع من الأخطاء غير الأخطاء العادية تسمى " أخطاء المعاينة " Sampling Errors. وهي الأخطاء الناجمة عن الدراسة بالعينة وأن كانت هذه الأخطاء تتميز بأنه يمكن تقديرها والتحكم فيها بإتباع الأساليب الإحصائية السليمة عند اختيار العينة.

4- لا تصلح الدراسة للعينات في بعض الحالات التي لو تركت فيها بعض المفردات دون فحص أو دراسة يترتب عليها إلحاق الضرر بالمجتمع أو ببعض مفرداته (التطعيم، اسطوانات الغاز...).

5- في بعض الدراسات التي تحتاج دقة أكثر يتطلب الأمر زيادة حجم العينة الأمر الذي قد لا يكون متاحاً عملياً، أو لا يوفر كثيراً في الوقت والجهد والتكاليف وبالتالي لا تتم الاستفادة من مزايا العينات أو لا يكون هناك فرق كبير بين الدراسة بالعينة والدراسة بالحصر الشامل.

6-5. تقسيم العينات

وتنقسم العينات الى قسمين رئيسيين هما :

6-5-1. العينات العشوائية والعينات الغير عشوائية

6-5-1-1. 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 مديرمستشفى 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

$$K = \frac{N}{n}$$

حيث ان:

N (حجم المجتمع)

n (حجم العينة)

مثال 2 : إذا كان عدد المرضى في مستشفى الموصل هو 10000 مريض ويراد اختيار عينة حجمها 50 مريض ، أوجد الزيادة المنتظمة (k) ؟

الحل:

$$K = \frac{N}{n} = \frac{10000}{50} = 200$$

مثال 3 :يراد اختيار عينة حجمها 100 مريض ، بمرض السكر من مجتمع عدد أفراده 6051 مريض ، أوجد مقدار الزيادة المنتظمة؟

الحل:

$$K = \frac{N}{n} = \frac{6051}{100} = 60.51$$

ويتم التقريب للأسفل فيكون الجواب هو 60

مثال4: يراد اختيار عينة حجمها 100 مريض بمرض التهاب الامعاء من مجتمع حجمه 4000 مريض ، كيف يتم ذلك بطريقة العينة المنتظمة؟

الحل:

1) نجد مقدار الزيادة المنتظمة k :

$$K = \frac{N}{n} = \frac{4000}{400} = 10$$

2) نختار مفردة البداية وتكون أقل من 10 ولتكن 5

3) نضيف (k) بشكل متسلسل

10 , 10 + 5 , 10 + 5 + 5 , 10 + 5 + 5 + 5 , ….

10 , 15 , 20 , 25 , … 3995

مثال 5: قسم المحاسبة في شركة بيع الادوية يرغب في اجراء دراسة لعملاء الشركة من حيث طرق السداد وحجم الطلبات وكيفية الشحن، إذا تم اختيار العينات بطريقة العينة المنتظمة وكان عدد العملاء المسجلين 2500؛ وتم اختيار حجم العينة ليكون 350 أوجد العينة المطلوبة.

الحل:

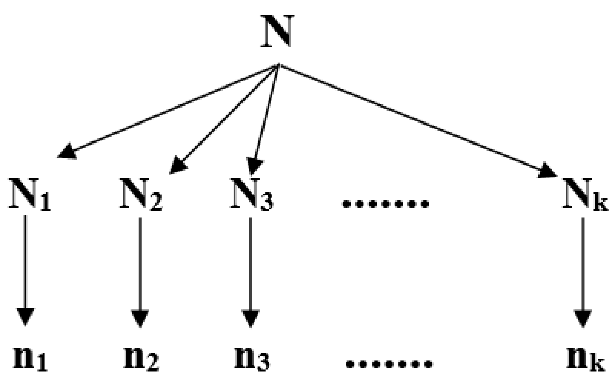
$$K=\frac{N}{n}=\frac{2500}{350}=7.14\approx 7$$

فتكون عينة العملاء المختارة هي كالآتي؛ إذا اعتبرنا أن مفردة البداية هي 20 :

18 , 27 , 34 , 41 , …

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 :

$$N = N_1 + N_2 + N_3 + ... + N_k$$

$$n = n_1 + n_2 + n_3 + ... + n_k$$

ونجد حجم العينة من الطبقة (i) بالصيغة :

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

مثال 6 : عدد المرضى الكلي الراقدين في احدى المستشفيات الكبرى 10000 مريض ومكون من 5 اقسام حجم كل قسم على التوالي وكما يلي :

(500 , 1000 , 3500 , 4000 , 5000). يراد سحب عينة حجمها 250 مفردة من هذا المجتمع، كيف يتم ذلك بحيث تمثل هذه العينة تمثيلاً سليماً.

الحل:

$$n = 250;\ N = 10000;\ N_1 = 500;\ N_2 = 1000;\ N_3 = 3500;\ N_4 = 4000;\ N_5 = 1500.$$

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

$$n_1 = \frac{N_1}{N} \times n = \frac{500}{10000} \times 250 = 12.5$$

$$n_2 = \frac{N_2}{N} \times n = \frac{1000}{10000} \times 250 = 25$$

$$n_3 = \frac{N_3}{N} \times n = \frac{3000}{10000} \times 250 = 75$$

$$n_4 = \frac{N_4}{N} \times n = \frac{4000}{10000} \times 250 = 100$$

$$n_5 = \frac{N_5}{N} \times n = \frac{1500}{10000} \times 250 = 37.5$$

نلاحظ أن:

$$n = n_1 + n_2 + n_3 + n_4 + n_5 = 12.5 + 25 + 75 + 100 + 37.5 = 250$$

وهو حجم العينة المطلوبة.

مثال 7 : أرادت مجلة طبية متخصصة دراسة مقدار الأرباح التي تحققها مستشفيات القطاع الخاص في

العراق حسب نشاطها الطبي . فعملت دراسة على 200 مستشفى . فإذا تم اختيار عينة مكونة من 50 مستشفى . كيف يتم الاختيار بطريقة سليمة؟

الجدول التالي يبين عدد ونوع كل مستشفيات في مجتمع الدراسة.

Group	**No. of Hospitals**
Operations	60
Nationals	40
Drugs	40
Consulting	20
Transportation	20
Services	20
Total	**200**

الحل:

$$n = 50;\ N = 200;\ N_1 = 60;\ N_2 = 40;\ N_3 = 40;\ N_4 = 20;\ N_5 = 20\ ;\ N_6 = 20.$$

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

$$n_1 = \frac{N_1}{N} \times n = \frac{60}{200} \times 50 = 15$$

$$n_2 = \frac{N_2}{N} \times n = \frac{40}{200} \times 50 = 10$$

$$n_3 = \frac{N_3}{N} \times n = \frac{40}{200} \times 50 = 10$$

$$n_4 = \frac{N_4}{N} \times n = \frac{20}{200} \times 50 = 5$$

$$n_5 = \frac{N_5}{N} \times n = \frac{20}{200} \times 50 = 5$$

$$n_6 = \frac{N_6}{N} \times n = \frac{20}{200} \times 50 = 5$$

نلاحظ أن:

$$n = n_1 + n_2 + n_3 + n_4 + n_5 + n_6 = 15 + 10 + 10 + 5 + 5 + 5 = 50$$

وهو حجم العينة المطلوبة.

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 العلمي لاطباء مستشفى اربيل في سنة 2016. كيف تتم عملية الاختيار بطريقة سليمة؟

الحل:

ننظر على مجتمع الاطباء بالمستشفى على أنه موزع على عدة تخصصات؛ ثم أن هذه التخصصات مقسمة لعدة أقسام وهذه الأقسام بدورها تحتوي على عدة شعب. وبأخذ هذه المراحل بعين الاعتبار يتم اختيار العينة المطلوبة.

المستشفى ← التخصصات ← الاقسام ← الشعب

6-5-1-2. العينات غير العشوائية : ـ يقصد بها مجموعة من المفردات المختارة من مجتمع الدراسة بطريقة يكون للباحث دخل في اختيارها ومن هذه العينات .

أ ـ **المعاينة 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 .

6 - 6 التوزيعات الاحتمالية

عند دراستنا للتوزيعات الاحتمالية ، نميز بين نوعين من هذه التوزيعات

1- التوزيعات الاحتمالية للمتغيرات المتقطعة .
2- التوزيعات الاحتمالية للمتغيرات المتصلة

. ويعد توزيع ذي الحدين مــن أهم التوزيعات الاحتمالية للمتغيرات المتقطعة ، كما يعتبر التوزيع الطبيعي وتوزيع ستيودنت (t) من اهم التوزيعات الاحتمالية المتصلة ولهذه التوزيعات أهمية خاصة عند دراسة العينات لاستخدامها عند تقدير معالم المجتمع .

6- 6 - 1 . التوزيعات الاحتمالية المتقطعة

6 – 6 – 1 – 1. توزيع ذى الحدين (Binomial Distribution)

كثير من التجارب العشوائية يكون لها نتيجتين فقط

النتيجة الأولى نجاح (S)

النتيجة الثانية فشل (F)

أمثلة على ذلك

1- عند اختيار عينة عشوائية من إنتاج أحد المصانع وكنا نبحث عن المنتج المعيب

النجاح هنا هو الحصول على المنتج المعيب

الفشل هنا هو الحصول على المنتج السليم

2- عند اختيار نتيجة طالبة من بين مجموعة من الطالبات فى مادة الإحصاء

النجاح هنا هو نجاح الطالبة فى مادة الإحصاء

الفشل هنا هو رسوب الطالبة فى مادة الإحصاء

3- عند اختيار نتيجة فحص مريض من بين مجموعة من المرضى

النجاح هنا هو الحصول على نتائج ضمن المعدلات القياسية

الفشل هنا هو الحصول على نتائج اكثر او اقل من المعدلات القياسية

6-6-1-1-1. تعريف محاولات برنولى (Bernoulli Trials)

هى سلسلة من المحاولات المكررة لتحقق الآتى:

1- كل محاولة لها نتيجتين فقط ونرمز لهما بالرمز (S) نجاح و (F) فشل
2- كل المحاولات مستقلة أى أن نتيجة أى محاولة ليس لها أى تأثير على نتائج المحاولات الأخرى
3- إحتمال النجاح واحتمال الفشل يبقى ثابتا فى كل محاولة وهو (**q** and **p**) حيث (**p + q = 1**)

إذا عرفنا المتغير العشوائى (X) بأنه عدد مرات النجاح فى (n) من محاولات برنولى فإن قيم (X) الممكنة هى (0, 1, 2, 3, 4, …, n) . اي انه (B(n, p)).

يكون للمتغير (X) توزيع يسمى توزيع ذى الحدين الاحتمالي وله الشكل العام ، كما في المعادلة التالية:

$$p\ (\mathrm{X} = \mathrm{r}) = \binom{n}{r} p^r q^{n-r} \quad , \quad r = 0, 1, 2, 3, \dots, n$$

أو

$$p\ (\mathrm{X} = \mathrm{r}) = \frac{\mathrm{n}!}{r!\,(n-r)!} p^r q^{n-r} \quad , \quad r = 0, 1, 2, 3, \dots, n$$

(!) تمثل المضروب . factorial function حيث :

$$n! = (n)(n-1)(n-2)(n-3)\dots(1)$$

(n) يمثل عدد مرات المحاولة في التجربة number of trials

(r) يمثل عدد مرات النجاح number of successes

(n-r) عدد مرات الفشل number of failures

(P) احتمال النجاح probability of success

(q) يمثل احتمال الفشل probability of failure ويساوي (p + q = 1)

$$(p+q)^n = \sum_{r=0}^{n} \binom{n}{r} p^r q^{n-r}$$

مثال 9:

إذا كان احتمال إصابة الهدف لشخص ما هو 1/7 واعطيت له فرصة الرماية فى 20 محاولة

1- ما هو احتمال إصابة الهدف مرتين على الأكثر

2- احتمال إصابة الهدف مرة واحدة

الحل

(X) متغير عشوائى يمثل عدد مرات النجاح فى إصابة الهدف فى 20 محاولة

n = 20

p = 1/7

p + q = 1

q = 6/7

r = 0, 1, 2, …, 20

$$p\ (\mathrm{X} = \mathrm{r}) = \frac{\mathrm{n}!}{r!\,(n-r)!} p^r q^{n-r} \quad , \quad r = 0, 1, 2, 3, \dots, 20$$

1- احتمال إصابة الهدف مرتين على الأكثر

أى إحتمال (x = 0 , x = 1 and x = 2)

$$\mathrm{P(r \le 2) = p(r = 0) + p(r = 1) + p(r = 2)}$$

$$p\ (\mathrm{r} = 2) = \frac{20!}{0!\,(20-0)!}(\frac{1}{7})^0(\frac{6}{7})^{20-0} + \frac{20!}{1!\,(20-1)!}(\frac{1}{7})^1(\frac{6}{7})^{20-1} + \frac{20!}{2!\,(20-2)!}(\frac{1}{7})^2(\frac{6}{7})^{20-2}$$

$$p\ (\mathrm{r} = 2) = (1)\left(\frac{1}{7}\right)^0\left(\frac{6}{7}\right)^{20} + (10)\left(\frac{1}{7}\right)^1\left(\frac{6}{7}\right)^{19} + (190)(\frac{1}{7})^2(\frac{6}{7})^{18}$$

$$p\ (\mathrm{r} = 2) = 0.00458 + 0.0764 + 0.24231 = 0.32529$$

2- احتمال إصابة الهدف مرة واحدة

أى إحتمال (x = 0 and x = 1)

$$\mathrm{P(r \le 1) = p(r = 0) + p(r = 1)}$$

$$p\ (\mathrm{r} = 1) = \frac{20!}{1!\,(20-1)!}(\frac{1}{7})^1(\frac{6}{7})^{20-1}$$

$$p\ (\mathrm{r} = 1) = (10)\left(\frac{1}{7}\right)^1\left(\frac{6}{7}\right)^{19}$$

$$p\ (\mathrm{r} = 1) = 0.0764$$

مثال 10 : في عائلة مكونة من ثلاثة أطفال ما هو احتمال أن يكون بنتين و ولد ، وما هو احتمال ثلاثة أولاد وما هو احتمال على الأقل ولد واحد .

الحل

1- احتمال أن يكون بنتين و ولد

$$(\ n = 3,\ \ r = 2,\ \ p = 0.5, \quad q = 0.5\ )$$

$$p\ (\mathrm{2G, 1B}) = \frac{3!}{2!\,(3-2)!}(0.5)^2(0.5)^{3-2}$$

$$p\ (\mathrm{2G, 1B}) = (3)(0.5)^2(0.5)^1 = 0.375$$

2- احتمال أن يكون بنتين و ولد

$$(\ n = 3,\ \ r = 3,\ \ p = 0.5, \quad q = 0.5\ )$$

$$p\ (3B) = \frac{3!}{3!\,(3-3)!}(0.5)^3(0.5)^{3-3}$$

$$p\ (3B) = (1)(0.5)^3(0.5)^0 = 0.125$$

3- احتمال على الأقل ولد واحد

$$p\ (\ 1B) = p(3B) + p(2B + 1G) + p(1B, 2G)$$

$$p\ (\ 1B) = \frac{3!}{3!\,(3-3)!}(0.5)^3(0.5)^{3-3} + \frac{3!}{2!\,(3-3)!}(0.5)^2(0.5)^{3-2} + \frac{3!}{2!\,(3-2)!}(0.5)^2(0.5)^{3-2}$$

$$p\ (\ 1B) = 0.125 + 0.375 + 0.375 = 0.875$$

مثال 11 : عولج 9 أشخاص جراحيًا ، نسبة النجاح 80% لكل شخص، علمًا أن عدد مرات الجراحة الناجحة يتبع للتوزيع ذو الحدين حيث (n = 9, p = 0.8) ما هو احتمال نجاح 7 عمليات جراحية.

الحل :

$$(\,n = 33,\ \ r = 7,\ \ p = 80\%\,,\quad q = 1.0 - 0.8 = 0.2)$$

$$p\ (r = 7) = \frac{9!}{7!\,(9-7)!}(0.8)^7(0.2)^{9-7}$$

$$p\ (r = 7) = \frac{9*8*7*6*5*4*3*2*1}{7*6*5*4*3*2*1(2*1)}(0.8)^7(0.2)^2$$

$$p\ (r = 7) = \frac{9*8}{(2*1)}(0.8)^7(0.2)^2 = 0.3099$$

مثال 12 : احتمال شفاء احد انواع مرضى الدم هو 0.4 ، اذا شخصَ 10 أفراد ووجد أنهم يحملون المرض ما هو احتمال :

A. أن يعيش 3 مرضى

B. أن يعيش على الأقل 8 مرضى

C. أن يعيش من 2 الى 5 مرضى

D. أن يعيش 10 مرضى

E. أن يعيش مريض واحد

الحل:

لنفترض أن X هو عدد الأفراد الذين سيعيشون ، وبالاعتبار لجدول التوزيع ذو الحدين

n = 10 and P = 0.4

A. أن يعيش 3 مرضى

$$p\ (\mathrm{X}=\mathrm{r})=\frac{\mathrm{n}!}{r!\,(n-r)!}p^{r}q^{n-r}\quad ,\quad r=0,1,2,3,\dots,10$$

$$p\ (r=3)=\frac{10!}{3!\,(10-3)!}(0.4)^{3}(0.6)^{10-3}$$

$$p\ (r=3)=\frac{10*9*8}{3*2*1}(0.4)^{3}(0.6)^{7}=120(0.064)(0.028)=0.21504$$

B- أن يعيش على الاقل 8 مرضى

$$p\ (\mathrm{X}=\mathrm{r})=\frac{\mathrm{n}!}{r!\,(n-r)!}p^{r}q^{n-r}\quad ,\quad r=0,1,2,3,\dots,10$$

$$p\ (r\geq 8)=p(x=8)+p(x=9)+p(x=10)$$

$$p\ (r\geq 8)=0.0106+0.0016+0.0001=0.0123$$

C- أن يعيش من 2 الى 5 مرضى

$$p\ (\mathrm{X}=\mathrm{r})=\frac{\mathrm{n}!}{r!\,(n-r)!}p^{r}q^{n-r}\quad ,\quad r=0,1,2,3,\dots,10$$

$$p\ (2\leq r\leq 5)=p(x=2)+p(x=3)+p(x=4)+p(x=5)$$

$$p\ (2\leq r\leq 5)=0.1209+0.2150+0.2508+0.2007=0.7874$$

D- أن يعيش 10 مرضى

$$p\ (\mathrm{X}=\mathrm{r})=\frac{\mathrm{n}!}{r!\,(n-r)!}p^{r}q^{n-r}\quad ,\quad r=0,1,2,3,\dots,10$$

$$p\ (r=10)=\frac{10!}{10!\,(10-10)!}(0.4)^{10}(0.6)^{10-10}$$

$$p\ (r=10)=(0.4)^{10}(0.6)^{0}=0.0001$$

E- أن يعيش مريض واحد

$$p\ (r=1)=\frac{10!}{1!\,(10-1)!}(0.4)^{1}(0.6)^{10-1}$$

$$p\ (r=1)=10*(0.4)^{1}(0.6)^{9}=0.0403$$

2-1-1-6-6 . استخدام الجداول في ايجاد التوزيع الاحتمالي ذو الحدين

ممكن ايجاد قيمة توزيع الاحتمالي ذو الحدين من الجداول كما في المثال التالي:

مثال 13: اوجد قيمة توزيع الاحتمالي ذو الحدين ن اذا علمت ان $(n=3 ,\quad p=5 ,\quad r = 2)$ من الجدول .

الحل

Table Portion of Binomial Probability Table

		p			
n	r	.10	.25	1/3	.50
	0	.7290	.4219	.2963	.1250
	1	.2430	.4219	.4414	.3750
3	2	.0270	.1406	.2222	**.3750**
	3	.0010	.0156	.0370	.1250

2-1-1-6-6 .3 الوسط الحسابي والتباين للتوزيع ذو الحدين

Mean and Standard Deviation of a Binomial Distribution

1. الوسط الحسابي (mean) يتم ايجاده حسب العلاقة التالية:

$(\mu = np)$

2. التباين (variance) يتم ايجاده حسب العلاقة التالية:

$$\sigma = \sqrt{np(1-p)} = \sqrt{npq}$$

مثال 14 : تكون نسبة الاستشفاء لدى مرضى الأورام عند علاجهم كيميائيا 60% في فترة 5 سنوات ، في مجموعة مكونة من 50 مريض بهذا النوع من الأورام احسب مايلي:

1- الوسط الحسابي للمتشافين

2- الانحراف المعياري

الحل

الوسط الحسابي

$$\mu = np = 50\ x\ 0.6 = 30$$

الانحراف المعياري

$$\sigma = \sqrt{np(1-p)} = \sqrt{npq} = \sqrt{50\ x\ 0.6x0.4} = 3.464$$

6-6 -2. توزيع بواسون (Poisson Distribution)

يستخدم توزيع بواسون عندما يكون لدينا متغير يحدث بشكل عشوائي مستقل خلال ظرف محدد ، بمعنى خلال وحدة قد تكون فترة زمنية ،عمر محدد، مساحة محددة ، حجم محدد ، طول معين.

أمثلة

1. عدد عمليات القسطرة في مستشفى خلال اسبوع.
2. عدد مرات حدوث حالات مرضية بين اطباء مستشفى ما خلال سنة.
3. عدد خلايا الفيروسات النامية في حجم معين من وسط النمو.
4. عدد كريات الدم البيضاء في عدد معين من مربعات شريحة العد .
5. عدد المرضى في مستشفى ما خلال شهر.
6. عدد مرات اصابة طفل عمره سنة بالاسهال.

ملاحظة : توزيع بواسون لا يستخدم في حالة الأمراض المعدية Infectous diseases ، لأن الأحداث (الأشخاص المصابين بالعدوى) لا يكونوا مستقلين عن بعضهم البعض.

صيغة توزيع بواسون

$$\boldsymbol{p(r) = \frac{\lambda^r e^{-\lambda}}{r!}} \qquad\qquad \boldsymbol{for \quad r = 0, 1, 2, \ldots \ldots}$$

λ: حرف يوناني يلفظ لامدا (Greek letter lambda)

وهي المتوسط الحسابي للأحداث خلال فترة معينة .(مسافة ، وزن ، حجم ، مساحة ، زمن ... الخ).

$\boldsymbol{e}$: العدد النيبيري وقيمته (2.718).

$\boldsymbol{r}$: عدد الأحداث (number of events) خلال فترة معينة (مسافة ، وزن ، حجم ، مساحة ، زمن ... الخ).

مثال 15 : ماهو احتمال نتيجة عد كرات الدم البيضاء في شريحة عد الخلايا الدموية (haemocytometer) وكانت النتائج حسب الجدول التالي:

عدد كرات الدم البيضاء (x_i)	2	5	7	3	8	9	10	1	11	6
التكرار (f_i)	3	4	5	8	3	7	3	9	3	4

في هذه الحالة المساحة المعطاة هي حجرات شريحة العد والأحداث هي كرات الدم البيضاء

لذا فالعدد الكلي لكرات الدم البيضاء في شريحة يمكن حسابه كالاتي:

$\sum f_i x_i$ =2*3+5*4+7*5+3*8+8*3+9*7+10*3+1*9+11*3+6*4 = 268

عدد الحجرات في شريحة العد هي التكرارات حسبما هو معطى ويساوي

$$\sum f_i = 3 + 4 + 5 + 8 + 3 + 7 + 3 + 9 + 3 + 4 = 49$$

لذا الوسط الحسابي لعدد كرات الدم الحمراء في كل حجرة

يساوي

$$\ddot{X} = \frac{\sum f_i x_i}{\sum f_i} = \frac{268}{49} = 5.47 = \lambda$$

$$r = 10 \,, \qquad \lambda = 5.47 \,, \qquad e = 2.718$$

$$p(r) = \frac{\lambda^r e^{-\lambda}}{r!} \qquad for \quad r = 0, 1, 2, \ldots\ldots, 10$$

$$\mathrm{p(r = 10)} = \frac{5.47^{10} \times 2.718^{-5.47}}{10!} = \frac{23981329.9 \times 0.00421}{3628800} = 0.0278$$

لذا يمكن ايجاد احتمال ظهور أي عدد من الخلايا في أي حجرة عشوائيًا.

مثال 16: مركز رعاية أولية مخصص لاستقبال الحالات الطارئة ، كان المتوسط الحسابي لعدد المرضى 7.0 مريض لكل ساعة من الساعة الخامسة صباحا حتى الساعة الثانية عشر مساءا ، ماهو احتمال أن يكون عدد المرضى 5 عند اختيار أي ساعة عشوائيًا.

الحل

$$r = 5 \,, \qquad \lambda = 7 \,, \qquad e = 2.718$$

$$p(r) = \frac{\lambda^r e^{-\lambda}}{r!} \qquad for \quad r = 0, 1, 2, \ldots\ldots, 10$$

$$\mathrm{p(r = 5)} = \frac{7^5 \times 2.718^{-7}}{5!} = \frac{16807 \times 0.00091}{120} = 0.1278$$

مثال 17: أجريت دراسة في مستشفى الموصل على عدد حالات لدغات الأفاعي خلال العام 2016 ، باعتبار عدد الحالات في هذا المستشفى يخضع لتوزيع بواسون ($P(\lambda = 9)$ أوجد مايلي:

A. اذا اخترنا سنة عشوائية ، ما هو احتمال أن يكون عدد الحالات الملدوغة (4) ؟

B- ما هو احتمال أن يكون عدد تلك الحالات أقل من (2) ؟

B- ما هو العدد المتوقع لعدد الحالات المصابة باللدغات خلال العام ؟

الحل

A- اذا اخترنا سنة عشوائية ، ما هو احتمال أن يكون عدد الحالات الملدوغة (4) ؟

$$r = 4\,, \qquad \lambda = 9\,, \qquad e = 2.718$$

$$p(r) = \frac{\lambda^r e^{-\lambda}}{r!} \qquad for \quad r = 0, 1, 2, \dots\dots, 4$$

$$\mathrm{p(r = 4)} = \frac{9^4 \times 2.718^{-9}}{4!} = 0.03377$$

B- ما هو احتمال أن يكون عدد تلك الحالات أقل من (2) ؟

$$p(r < 2) = p(r = 1) + p(r = 0)$$

$$p(r < 2) = \frac{9^1 \times 2.718^{-9}}{1!} + \frac{9^0 \times 2.718^{-9}}{0!} = 0.00111 + 0.00012 = 0.00123$$

C- ما هو العدد المتوقع لعدد الحالات المصابة باللدغات خلال العام ؟

$$\mu = \lambda = 9\ Cases$$

مثال 18 : لنفترض أن احتمال الموت عند الاصابة باللوكيميا يخضع لتوزيع بواسون حيث أن معدل الموت هو 3 لكل 10000 طفل. ما هو احتمال أن يموت 7 من 10000 ؟

الحل

$$r = 7\,, \qquad \lambda = 33\,, \qquad e = 2.718$$

$$p(r) = \frac{\lambda^r e^{-\lambda}}{r!} \qquad for \quad r = 0, 1, 2, \dots\dots, 7$$

$$\mathrm{p(r = 7)} = \frac{3^7 \times 2.718^{-3}}{7!} = 0.0216 = 2.16\ \%$$

6-6-2-1. استخدام الجداول لإيجاد قيمة توزيع بواسون

ممكن ايجاد قيمة توزيع بواسون من الجداول كما في المثال التالي:

مثال 19: اوجد قيمة توزيع بواسون ن اذا علمت ان $p\ (r = 2)$ عندما $\lambda = 0.3$ من الجدول .

الحل

	λ				
r	0.1	0.2	0.3	0.4	0.5
0	.9048	.8187	.7408	.6703	.6065
1	.0905	.1637	.2222	.2681	.3033
2	.0045	.0164	.0333		
3	.0002	.0011	.0033	.0072	.0126
4	.0000	.0001	.0003	.0007	.0016

2-2-6-6 . الوسط الحسابي والتباين لتوزيع بواسون

Mean and Standard Deviation of a Poisson Random variable

1. الوسط الحسابي أو العدد المتوقع يساوي (λ)

$$Mean(Expected\ number)\mu = \lambda$$

2. صفة مميزة لتوزيع بواسون أن الوسط الحسابي يساوي التباين ويساوي (λ)

$$Variance \quad \sigma^2 = \lambda$$

3. الانحراف المعياري يساوي جذر التباين

$$Standard\ Deviation \quad \sigma = \sqrt{\sigma^2} \quad = \sqrt{\lambda}$$

مثال 20: تتوزع البكتيريا بشكل مستقل عن بعضها البعض في محلول ما ، وحيث أن عدد خلايا البكتيريا بالملميتر يخضع لتوزيع بواسون بمعدل mean يساوي 2 .

ما هو احتمال أن تحتوي عينة بمقدار 1 ملميتر من المحلول على :

1. خلية بكتيريا واحدة أو أقل ؟
2. أكثر من 5 خلايا بكتيريا ؟

الحل

1. خلية بكتيريا واحدة أو أقل ؟

$$p(r \leq 1) = p(r = 1) + p(r = 0)$$

$$p(r < 2) = \frac{2^1 \times 2.718^{-2}}{1!} + \frac{2^0 \times 2.718^{-2}}{0!} = 0.2707 + 0.1353 = 0.406$$

2 . أكثر من 4 خلايا بكتيريا ؟

$$p(r > 5) = 1 - P(r \leq 5)$$
$$= 1 - [p(r=5) + p(r=4) + p(r=3) + p(r=2) + p(r=1) + p(r=0)]$$

$$p(r > 5) = 1 - [\frac{2^5 \times 2.718^{-2}}{5!} + \frac{2^4 \times 2.718^{-2}}{4!} + \frac{2^3 \times 2.718^{-2}}{3!} + \frac{2^2 \times 2.718^{-2}}{2!} + \frac{2^1 \times 2.718^{-2}}{1!} + \frac{2^0 \times 2.718^{-2}}{0!}$$
$$= 1 - \ 0.0361 + 0.0902 + 0.1804 + 0.2707 + 0.135$$
$$= 1 - 0.9834 = 0.0166$$

ملاحظة:

من الممكن استخدام توزيع بواسون بدلا من التوزيع ذو الحدين لإيجاد الاحتمالات المنفصلة للمتغيرات العشوائية ، عندما تكون قيمة (p) (احتمال النجاح) منخفضة ، وعدد مرات المحاولة (n) كبيرة.

تختلف المصادر العلمية في ذلك ، هناك ثلاثة مصادر تشير الى أنه يجب استخدام توزيع بواسون بدلا من التوزيع ذو الحدين عندما:

$$n \geq 20 \quad and \quad p \leq 0.05$$
$$n \geq 100, \quad \leq 0.1 \quad and \quad np \leq 10$$
$$n \geq 100 \quad and \quad p \leq 0.01$$

مثال 21 : لنفترض أن احتمال نمو بكتريا (سودوموناس) في صالة عمليات هو(0.33) ، ضمن مجموعة من (50) صالة عمليات، ما هو احتمال أن تحتوي 5 مجموعات على هذه البكتريا؟

الحل

عند استخدام التوزيع ذو الحدين حيث أن:

$$n = 50, \ r = 5, \quad p = 0.33, \quad and \quad q = 0.67$$

احتمال وجود الكائن في 5 مجموعات من 50 مجموعة

$$p\ (X = r) = \frac{n!}{r!\,(n-r)!} p^r q^{n-r} \quad , \quad r = 0, 1, 2, 3, \dots, 50$$
$$p\ (r = 5) = \frac{50!}{5!\,(50-5)!} (0.33)^5 (0.67)^{50-5}$$

$$p\ (r=5)=2118760\ \times(0.33)^5(0.67)^{45}=0.000123$$

عند استخدام توزيع بواسون ، نحسب μ حسب العلاقة التالية :

$$\mu=\lambda=np=50\ \times0.33=16.5$$

ومن ثم نحسب احتمال وجود البكتريا في 5 مجموعات من 50 مجموعة

$$r=5\,,\qquad \lambda=16.5\,,\qquad e=2.718$$

$$\mathrm{p(r=5)}=\frac{\lambda^r e^{-\lambda}}{r!}=\frac{16.5^5\ \times\ 2.718^{-16.5}}{5!}=0.000696$$

نلاحظ أن قيمة توزيع بواسون $[p(r=5)=0.000696]$ قريبة جداً من قيمة التوزيع ذو الحدين $[p(r=5)=0.000696]$

3-6-6. التوزيعات الاحتمالية المتصلة

المتغير العشوائي المتصل (x) هو ذلك المتغير الذي يمكن أن يأخذ عدداً لا نهائياً من القيم المعلومة، واحتمال أن تقع (x) داخل أى فترة يمثلها مساحة التوزيع الاحتمالي (ويسمى أيضاً دالة الكثافة) داخل هذه الفترة، والمساحة الكلية تحت المنحنى (الاحتمال) تساوى (1).

هناك بعض التوزيعات الاحتمالية المتصلة لها دوال كثافة احتمال محددة ومنها:

1- التوزيع الطبيعي
2- التوزيع الطبيعي (القياسي) المعياري
3- توزيع (t)

1-3-6-6. التوزيع الطبيعي (Normal Distribution)

هو أفضل وأكثر التوزيعات الاحتمالية المتصلة استخداماً في النواحي التطبيقية، ومنها الاستدلال الإحصائي شاملا التقدير، واختبارات الفروض، كما أن معظم التوزيعات يمكن تقريبها إلى هذا التوزيع.

والتوزيع الطبيعي هو توزيع احتمالي متصل، وهو جرسي الشكل ومتماثل حول الوسط الحسابي، ويمتد إلى مالا نهاية فى الاتجاهين، ولكن معظم المساحة (الاحتمال) تتركز حول الوسط الحسابى .

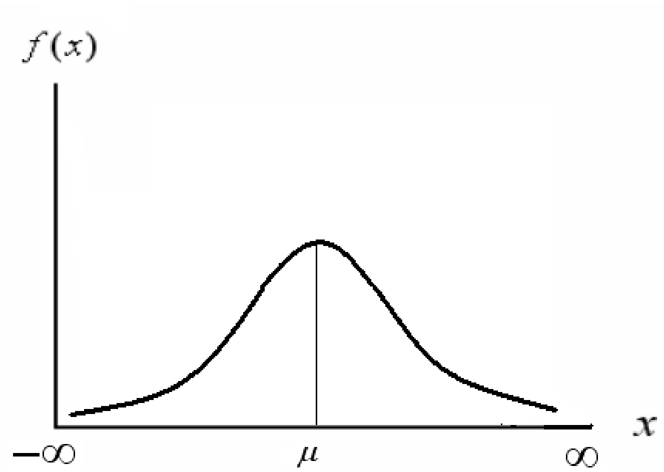
ويعتبر التوزيع الطبيعي من أهم أنواع التوزيعات الاحصائية المتصلة **ومن خصائصه**:

1. توزيع جرسي أي يشبه الجرس.
2. توزيع متصل
3. توزيع متماثل حول الوسط

4. الالتواء (الاطراف) والتفلطح (القمة) يساوي صفر.
5. يحوي منوال ووسط ووسيط واحد وذات قيم متساوية بمعنى أن الجزء الذي على يمين الوسط مطابق للجزء الايسر
6. الذيلين الايمن والايسر يقتربان من الخط الافقي ولكن لا تلامسه
7. المساحة الكلية تحت المنحنى تساوي واحد صحيح .

- منحنى دالة الاحتمال للتوزيع الطبيعي له خاصية شكل الجرس. ويتحدد شكل الجرس تماماً لأي توزيع طبيعي خاصة إذا علمنا الوسط الحسابي (μ) والانحراف المعياري (σ) لهذا التوزيع.
- تدل قيمة الوسط الحسابي على مكان مركز الجرس، كما تدل قيمة الانحراف المعياري على كيفية الانتشار.
- القيمـــــة الصغـــيرة لقيمة الانحراف المعياري تعني أن لدينا جرس طويل مدبب، والقيمة الكبيرة لقيمة الانحراف المعياري تعني أن الجرس قصير ومفرطح.

والشكل التالي يوضح ذلك



وتوجد معلمتين لهذا التوزيع هما :

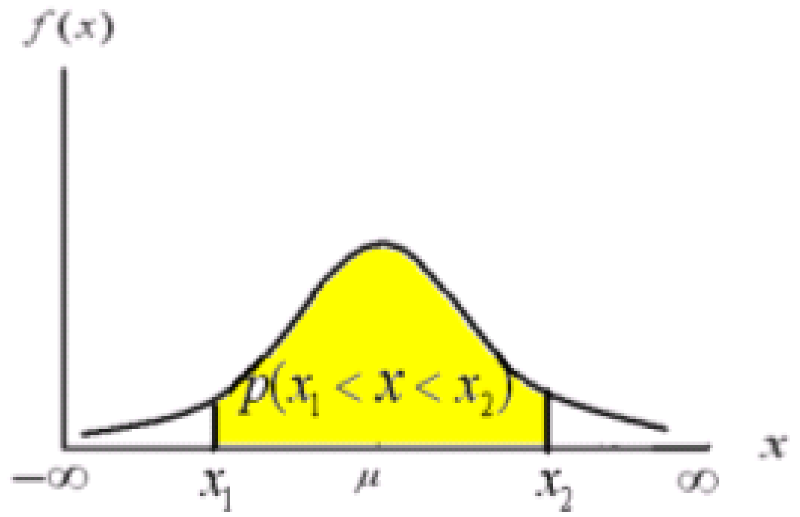
1- الوسط الحسابي ($E(x) = \mu$)
2- التباين : ($Var(x) = \sigma^2$)

ومن ثم يعبر عن توزيع المتغير بالرموز : $x \sim N(\mu, \sigma^2)$ ويعني ذلك أن المتغير العشوائي يتبع التوزيع الطبيعي بمتوسط (μ) ، وتباين (σ^2) .

وإذا كان لدينا توزيع طبيعي ذو وسط حسابي وانحراف معياري فإن معادلة منحنى دالة كثافة الاحتمال تكون على الصورة التالية:

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\; e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} , \; -\infty < x < \infty \;,\; \pi = 22/7$$

ويتم حساب الاحتمالات بفرض أن الاحتمال المطلوب حسابه هو $[p(x_1 < x < x_2)]$ وهذا الاحتمال يحدد بالمساحة التالية:



وحيث أن هذا التوزيع من التوزيعات المستمرة، فإن هذه المساحة (الاحتمال) تحسب بإيجاد التكامل التالي:

$$p(x_1 < X < x_2) = \int_{x_1}^{x_2} f(x)dx = \int_{x_1}^{x_2} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

وهذا التكامل يصعب حسابه، ومن ثم لجأ الإحصائيين إلى عمل تحويلة رياضية Transform، يمكن استخدام توزيعها الاحتمالي في حساب مثل هذه الاحتمالات، وهذه التحويلة هي:

$$z = \left(\frac{x - \mu}{\sigma}\right)$$

ويعرف المتغير الجديد بـ (Z) وهو المتغير الطبيعي القياسي Standard Normal Variable ، أو المعياري

الاسئلة

س1: ما العينة العشوائية المنتظمة إذا أراد مدير مستشفى تشكيلها على أن يكون حجمها 20 . مستخدماً جدول الأعداد العشوائية ابتداءً من الصف الثامن و العمود السابع ؟

س2: يبلغ عدد موظفي مستشفى الموصل 500 موظف مرقمين من 1 إلى 500 أراد مدير المستشفى إرسال 15 موظف لحضور ندوة صحية في الوزارة ، المطلوب : سحب عينة عشوائية منتظمة حجمها 15 موظف باستخدام جدول الأعداد العشوائية ابتداءً من الصف الثاني والعشرون و العمود الثالث.

س3: في أحد المستشفيات حيث عدد الاطباء 50 طبيبا مرقمين من 1 إلى 50 أراد مدير المستشفى مناقشة هؤلاء الاطباء حول كيفية تحسين الأداء و زيادة الخدمات الطبية في المستشفى. المطلوب سحب عينة عشوائية منتظمة حجمها 15 طبيب .مستخدماً جدول الأعداد العشوائية ابتداءً من الصف الثامن و العمود العاشر.

س4: يبين الجدول التالي توزيع الموظفين في إحدى المستشفيات:

أطباء	عمال	ممرضون	موظفي مختبر	إداريون	المجموع
40	30	85	25	20	200

المطلوب : سحب عينة عشوائية طبقية مكونة من 20 موظفا لدراسة كفاءة العاملين في المستشفى وذلك بتكوين عينات عشوائية بسيطة .

س5: لدراسة الأداء الوظيفي و الكفاءة عند الموظفين في إحدى المستشفيات ، تم سحب عينة عشوائية طبقية مكونة من 7 أفراد من 35 موظف موزعين كما يبين الجدول التالي:

اطباء	ممرضين	ادريون	**المجموع**
11	17	7	35

ما حجم كل عينة عشوائية بسيطة مسحوبة من كل طبقة ؟

الفصل السابع

الاحصاءات الحيوية

7 -1. الإحصاءات الحيوية

ان الحوادث الحياتية تحدد بشكل أو بآخر وجود الرباط البشري في مجتمعنا هذا ، وهي تعتبر بمثابة السلسلة الزمنية ترتبط حلقاتها بعضها مع البعض الآخر بحيث لا يستطيع أي كائن بشري متحضر الاستغناء عنها في حياته اليومية . والحوادث الحياتية تشمل :

1. الولادات (الخصوبة).
2. الوفيات
3. الزواج
4. الطلاق

ويعتبر تسجيل هذه الحوادث من الأمور المهمة والضرورية لكل بلد متحضر من اجل التعرف على المؤشرات الرئيسية إلى التخطيط السليم .

7 -1 – 1. إحصاءات الوفيات :

تتوقف نسبة نمو السكان في أي مجتمع وفي أي فترة زمنية على الفرق بين عدد الوفيات والمهاجرين من ذلك المجتمع وعدد الولادات والوافدين إلى ذلك المجتمع في نفس الفترة الزمنية . حيث ان الموت لا يحدث من غير وجود سبب له سواء أكان نتيجة لحادثة معينة أو لمرض معين ، وان هذه الأسباب تختلف باختلاف الأشخاص والظروف المحيطة ، ولمعرفة المستوى الصحي ومدى تأثير الأمراض يجب معرفة الأسباب المؤدية إلى الوفاة لسنوات للتوصل إلى الأسباب الحقيقية لتلك الوفيات .

وفيما يلي أهم المقاييس للتعرف على معدلات الوفاة .

$$1-\ \text{معدل الوفاة الخام} = \frac{\text{عدد المتوفين خلال السنة}}{\text{عدد السكان في منتصف السنة}} \times 1000$$

مثال 1: من الجدول التالي:

السنة	عدد السكان	عدد الوفيات
2013	201921	1928
2014	203211	1930
2015	206212	1937

احسب مايلي :

- معدل الوفاة الخام لعام 2015.

- معدل الوفاة الخام للسنوات الثلاث .

الحل:

معدل الوفاة الخام للعام 1980 = (203211 ÷ 1930) x 1000= 9.44 لكل 1000 شخص.

معدل الوفاة الخام للسنوات الثلاث = $\frac{1937+1930+1928}{206212+203211+201921} \times 1000 = 9.479$ لكل الف شخص.

1 – معدل الوفاة النوعي : ويمكن تقسيمه إلى :

-معدل الوفاة حسب الجنس = $\frac{\text{عدد المتوفين من الذكور او الاناث في السنة}}{\text{عدد الذكور او الاناث من السكان في منتصف السنة}} \times 1000$

-معدل الوفاة حسب العمر= $\frac{\text{عدد المتوفين في عمر معين خلال السنة}}{\text{عدد السكان في ذلك العمر في منتصف السنة}} \times 1000$

-معدل الوفاة حسب المنطقة = $\frac{\text{عدد الوفيات في الريف او الحضر خلال السنة}}{\text{عدد سكان الريف او الحضر في منتصف السن}} \times 1000$

-معدل الوفاة حسب السبب المرض = $\frac{\text{عدد الوفيات من مرض معين خلال السنة}}{\text{عدد السكان في منتصف تلك السنة}} \times 1000$

مثال 2: بلغ عدد وفيات الذكور عام 2010 في العراق 14592 كما بلغ عدد السكان الذكور في بداية السنة ونهايتها (3077408 ، 3185117) على التوالي ، احسب معدل الوفيات الخام للذكور .

الحل :

عدد الذكور في منتصف السنة = (3185117+3077408) ÷2=3131263

معدل الوفيات الخام للذكور = 14592 ÷3131263 x 1000=4.7

مثال 3: بلغ عدد وفيات العراق عام 2009 (10767) للأشخاص الذين أعمارهم 61 سنة فما فوق . بينما بلغ عدد الأشخاص للأعمار المذكورة في بداية السنة ونهايتها (356900 ، 369035) احسب معدل الوفيات الخام للأعمار أعلاه .

الحل:

عدد السكان في منتصف السنة =(369035+356900)÷ 2= 362967.5

معدل الوفيات الخام = 10767÷ 362967.5 x 1000= 29.66

2- معدل الوفاة المعياري (المصحح) :

من الخطأ مقارنة معدلات الوفيات في أكثر من مجتمع كما لاحظنا في استخدام معدلات الوفيات الخام في حالة وجود متغير يؤثر على الوفيات مثل العمر والجنس أو الحالة المرضية ، ولكي تصبح المقارنة صحيحة فإنها تجري عن طريق مقارنة ما يسمى بالمعدلات المعيارية (وهي إدخال أسلوب في عملية الترجيح وإعطاء وزن لكل معدل قبل إجراء الجمع والاختبار العقلاني للأوزان ، والتصحيح هو عملية إزالة الاختلافات بين المجتمعات السكانية).

$$\text{معدل الوفاة المعياري}\left[\text{لفئة او لمدينة او لمجتمع}\right] = \frac{\text{مجموع عدد الوفيات المتوقعة}\left[\text{لفئة او لمدينة او لمجتمع}\right]}{\text{مجموع عدد السكان}} \times 1000$$

وهناك طريقتين لتصحيح معدلات الوفيات :

- الطريقة المباشرة .
- الطريقة غير المباشرة

مثال 4: استعمل البيانات التالية لحساب معدل الوفاة المعياري لكل من مدينتي الموصل واربيل .

العمر	الموصل		اربيل	
	السكان	الوفيات	السكان	الوفيات
20- 29	45000	90	30000	45
30- 39	40000	120	30000	105
40-49	35000	140	40000	180
50-59	35000	150	50000	255
المجموع	**150000**	**500**	**150000**	**555**

عدد السكان المعياري = عدد سكان الموصل +عدد سكان اربيل

= 45000 + 30000= 75000 للسنة الأولى .... وهكذا

معدل الوفاة المتوقع للموصل للسنة الأولى = عدد الوفيات ÷ عدد السكان x عدد السكان المعياري لتلك السنة

= 90 ÷ 45000 x 75000 = 150 وهكذا لبقية السنين .

العمر	عدد السكان	عدد الوفيات المتوقع	
		موصل	اربيل
20- 29	75000	150	112.5
30- 39	70000	210	245
40-49	75000	300	337.5
50-59	85000	400	360
المجموع	**300000**	**1060**	**1055**

معدل الوفاة المعياري = عدد الوفيات المتوقع ÷عدد السكان المعياريx 1000

لمدينة الموصل = 1060 ÷300000 x 1000= 3.53 ويساوي تقريبا 4

لمدينة اربيل = 1055 ÷ 300000 x 1000=3.52 ويساوي تقريبا 4

معدل الوفاة الخام لمدينة الموصل = 500÷150000 x 1000= 3.33

معدل الوفاة الخام لمدينة اربيل = 555÷150000 x1000= 3.7

2-1-7. إحصاءات الولادات

$$\text{معدل الولادات الحية الخام} = \frac{\text{عدد المواليد الاحياء في السنة}}{\text{عدد السكان في منتصف السنة}} \times 1000$$

مثال 5: بلغ عدد المواليد في العراق في سنة 2011 (64758 مولود) ، بينما كان عدد السكان في بداية السنة 6131489 و في نهايتها 6339960 . ما هو معدل المواليد الخام خلال السنة ؟

الحل:

عدد السكان في منتصف السنة = (6131489 + 6339960) ÷2= 6235725

معدل المواليد = (64758 ÷ 6235725) x 1000= 10.4

$$\text{معدل الولادات الحية النوعية} = \frac{\text{عدد الولادات الحية للفئة}}{\text{عدد السكان من تلك الفئة}} \times 1000$$

$$\text{معدل التوالد العام} = \frac{\text{عدد المواليد الاحياء}}{\text{عدد المتزوجات في منتصف السنة}} \times 1000$$

مثال 6: بلغ عدد المواليد الأحياء في العراق 64758 في سنة 2013 وبلغ عدد المتزوجات في بداية السنة ونهايتها 911234 و 939482 . والمطلوب ما هو معدل التوالد العام ؟

الحل :

عدد المتزوجات في منتصف السنة = (911234 + 939482) ÷ 2= 925358

معدل التوالد العام = (64758 ÷ 925358) x 1000 = 7 مولود .

$$\text{معدل التوالد الخاص} = \frac{\text{عدد المواليد لامهات لفئة من العمر}}{\text{عددالمتزوجات في منتصف السنة لنفس فئة العمر}} \times 1000$$

مثال 7: بلغ عدد المواليد الأحياء لأمهات أعمارهن 15-20 سنة (9467) في سنة 1957 وبلغ عدد النساء المتزوجات بهذا العمر في بداية ونهاية السنة 84935 و 87605 . ما هو معدل التوالد لهذه الفئة ؟

الحل:

عدد المتزوجات في منتصف السنة = (84935 + 87605) ÷ 2= 86290

معدل التوالد الخاص = (9467 ÷ 86290) x 1000= 109.7

3-1-7. إحصاءات الخصوبة

ان ارتفاع الخصوبة وانخفاضها بين الشعوب المختلفة يعني ارتفاع في عدد المواليد أو انخفاضها ومن هنا وجدت هذه المقاييس لغرض المقارنة بين الدول المختلفة وبين مجموعات اجتماعية في الشعب الواحد .

$$\text{- معدل الخصوبة العام} = \frac{\text{عدد الولادات الحية خلال سنة معينة}}{\text{عدد النساء في السكان لمنتصف تلك السنة}} \times 1000$$

$$\text{- معدل الخصوبة النوعي} = \frac{\text{عدد الولادات الحية لعمر معين خلال سنة معينة}}{\text{عدد النساء بنفس العمر في منتصف تلك السنة}} \times 1000$$

$$\text{-معدل الخصوبة القياسية} = \frac{\text{عدد المواليد في سنة معينة}}{\text{مجموع الاناث في سن (15-49) لنفس السنة}} \times 1000$$

4-1-7. إحصاءات الزواج والطلاق

$$\text{-معدل الزواج} = \frac{\text{عدد الاشخاص المتزوجين خلال السنة}}{\text{عددالاشخاص الذين هم في سن الزواج ولم يتزوجوا}} \times 1000$$

$$\text{- معدل الطلاق} = \frac{\text{عدد شهادات الطلاق التي تمت خلال السنة}}{\text{عددالمتزوجين من السكان لنفس السنة}} \times 1000$$

5-1-7. جداول الحياة

1. جداول الحياة : عبارة عن مجموعة من القيم الرقمية توضح كيف ان مجموعة فرضية من السكان قد ولدت ثم بدأت بالتناقص عن طريق الوفاة بصورة تدريجية إلى أن تتلاشى .

2. استخدامات جداول الحياة في مجال الدراسات السكانية:

1. تسهيل العمليات المختلفة في الدراسات السكانية وخاصة خطر الموت عند كل فئة عمرية .
2. مقارنة ظروف الوفيات في مختلف الدول بصورة دقيقة .
3. تصوير النماذج الرياضية للسكان .
4. التنبؤ بعدد السكان وخصائصهم وحساب معدل التوالد
5. استخراج المقاييس التي تقيس بها كيفية تناقص وتغير السكان
6. التكهن بمستقبل السكان.
7. تتيح لشركات التأمين أسلوب سهل لتحيدي أقساط التأمين .
8. تفيد في المجالات التجارية بالنسبة للمخزون السلعي .

3. أنواع جداول الحياة :

هناك نوعان من جداول الحياة شاع استخدامها هي :

1- جدول الحياة لحياة جيل واحد .

2 – جدول الحياة الجاري أو المستمر .

4. جدول الحياة لجيل فعلي :

وتسمى أيضا الطريقة المباشرة : و تتطلب معرفة عدد المواليد في المجتمع خلال سنة معينة ومن ثم تتبعهم على مدار السنوات المختلفة المتعاقبة إلى أن ينقرضوا عن آخرهم ، ويتم ذلك عمليا بعمل بطاقة لكل مولود في سنة معينة ثم تسحب بطاقات المتوفين من هؤلاء المواليد على مدار السنوات المتعاقبة حتى تنتهي كل هذه البطاقات . وبذلك يتسنى معرفة عدد الباقين على قيد الحياة في مراحل العمر المختلفة ومعدلات البقاء على قيد الحياة لهذه الأعمار ، ويطلق على هذه الطريقة المباشرة أحيانا طريقة تعقب الأجيال.

وتعترض الطريقة المباشرة في عمل جداول الحياة الصعاب التالية:

A- نقص تسجيل المواليد والوفيات في معظم البلدان.

B- مقارنة بطاقات المواليد والوفيات عملية مجهدة وباهظة النفقات وغير عملية .

C- لا يمكن عمل جدول ويظل معلقا إلى أن يموت آخر فرد من المواليد وهذا يستغرق قرنا من الزمن .

D- أن هذا الانتظار يجعل الصورة التي تعترضها هذه الجداول غير حقيقية .

ومع هذا فان جدول الحياة للجيل له بعض الاستعمالات العملية منها **:**

1- دراسة تعداد الحيوانات .
2- دراسة طول حياة الآلات والمصابيح الكهربائية والفترة التي يمكن ان تستمر في الإنارة ......الخ .
3- تستخدم في تحليل احتمال بقاء المرضى الراقدين على قيد الحياة عند دراسة التأثيرات العلاجية لهم .

5. جدول الحياة المستمر :

ان هذا النوع من جداول الحياة يوفر معلومات عن الوفيات وتوقعاتها وعدد الباقين على قيد الحياة لكافة السكان ولكافة الأعمار خلال فترة محددة .

فمثلا لو أخذنا نفوس العراق لعام 1968 لوجب علينا ان نعتمد على معدلات الوفاة حسب الأعمار المفردة لهذه السنة لكي نحصل على تخمينات عن توقعات الحياة لكل فرد من أفراد المجموعة الفرضية بالنسبة لمعدلات الوفيات الحقيقية للسكان كافة .

ومن الممكن – في بعض الأحيان – الاعتماد على عدد الوفيات الحادثة خلال 3 سنوات بدلا من سنة واحدة ، أي أننا سنسحب الوفيات للأعوام 1967 ، 1968 ، 1969 ونقوم بتقسيم عدد الوفيات لكل فئة عمرية على عدد السكان لمنتصف السنوات الثلاث وبهذا تكون سنة 1968 هي السنة المعينة .

إن جداول الحياة سواء كانت من النوع الأول أو من النوع الثاني إما أن تكون كاملة أو أن تكون مختصرة . فالكاملة تعني إعداد جدول الحياة لكافة مفردات السنين ، أي لكل سنة من سنوات العمر مبتدئة من الصفر إلى آخر عمر يبلغه الفرد .

أما الجداول المختصرة فتنتظم على شكل فئات عمرية (عادة تكون كل خمس سنوات) عدا الفئة العمرية الأولى إذ تقسم إلى قسمين :

الفئة الأولى اقل من سنة والثانية من سنة إلى 4 سنين . ومن الجدير بالذكر بيان ضرورة بناء جداول منفردة للحياة لكل من الإناث والذكور وكذلك تنظيم جدول موحد للقطر وعدم توزيعه على المناطق .

6. الأسس والفروض التي يستند إليها بناء جداول الحياة :

نستند في عملية بناء جداول الحياة إلى فكرة المجتمع الثابت أو الساكن وهو مجتمع افتراضي تتحدد معالمه بما يلي :

1- إن عدد مواليده السنوية = عدد وفياته السنوية .
2- إن معدل الوفيات لكل فئة عمرية معدل ثابت لا تتغير بتغير الزمن .
3- إن عدد الأشخاص الباقين على قيد الحياة عند كل عمر أو كل فئة عمرية يتحدد بالمعدلات العمرية الثابتة وليس مثلا بالهيكل الكلي لأي سكان حقيقيين .
4- ان عدد المواليد السنوية مفترض ثباته عند نفس الرقم الذي بدأنا به الجدول وهو عادة رقم افتراضي = 1000 أو 10000 أو 100000 أو مليون .
5- إن حجم السكان في كل فئة عمرية ثابت وان التوزيع العمري لسكان جدول الحياة يكون ثابتا .

7. تكوين جداول الحياة :

يتكون جدول الحياة من الحقول التالية :

العمر x	عدد الأحياء lx	عدد الوفيات dx	الاحتمال السنوي للوفيات 1000 qx	عدد السنين التي عاشتها المجموعة Lx	مجموع السنوات التي عاشتها المجموعة بعد العمر Tx	توقع الحياة Ex
1	2	3	4	5	6	7

8. طرق إيجاد الأعمدة:

العمود الأول (العمر) : يعطى كله أو بعضه في السؤال .

العمود الثاني : عدد الأحياء في العمر السابق – عدد المتوفين للعمر السابق .

العمود الثالث : عدد الأحياء لذلك العمر – عدد الأحياء للعمر الذي يليه .

العمود الرابع : عدد المتوفين ÷ عدد الأحياء x 1000 .

العمود الخامس : (عدد الأحياء لذلك العمر + عدد الأحياء للعمر الذي يليه) ÷ 2

العمود السادس: يتأتى من جمع الأرقام في العمود بعضها مع بعض على شكل متجمع صاعد مبتدئين من اكبر عمر إلى أن نصل إلى أول عمر (اصغر عمر) .

العمود السابع : $\dfrac{\text{مجموع السنوات التي عاشتها المجموعة بعد العمر}}{\text{عدد الاحياء في العمر}}$

مثال 8 : نفترض ان عدد الأحياء في سنة العمر صفر 100000 وعدد الوفيات كما يلي :

السن	0	1	2	3	4	5
عدد الوفيات	240	249	259	268	277	286

والمطلوب أوجد جدول الحياة بشكل كامل .

الحل :

العمر x	**عدد الأحياء lx**	**عدد الوفيات dx**	**الاحتمال السنوي للوفيات 1000 qx**	**عدد السنين التي عاشتها المجموعة Lx**	**مجموع السنوات التي عاشتها المجموعة بعد العمر Tx**	**توقع الحياة Ex**
صفر	100000	240	2.4	99880	497105	4.97
1	99760	249	2.496	99880	397225	3.98
2	99511	259	2.6027	99381.5	297345	2.98
3	99252	268	2.7002	99118	197963.5	1.999
4	98984	277	2.7984	98845.5		
5	98707	286	2.8975			

عدد الأحياء في العمر 1 = 100000- 240 = 99780

احتمال الوفاة لكل ألف للعمر صفر = 240 ÷ 100000 x 1000=2.4

عدد السنين التي عاشتها المجموعة للعمر 1 = (9970 + 99511) ÷ 2=99880

توقع الحياة للعمر 1 = 497105 ÷ 100000 = 4.97

مثال 9 : من المعلومات الآتية أكمل جدول الحياة :

العمر x	عدد الأحياء lx	عدد الوفيات dx	الاحتمال السنوي للوفيات 1000 qx	عدد السنين التي عاشتها المجموعة Lx	مجموع السنوات التي عاشتها المجموعة بعد العمر Tx	توقع الحياة Ex
30	924600					
31	921317					
32	917880					
33	914282					
34	910516					
35	906554					
36	902393					
37	898007					

أسئلة متنوعة

س1 : أوجد قيمة الوسط الحسابي والوسيط والمنوال ، ثم حدد شكل الالتواء لمستوى الهيموكلوبين في الدم لعينة مكونة من خمسين شخصا تم تلخيص مستوى الهيموكلوبين لهم كما في الجدول التالي؟

الفئات	12.95 – 13.95	13.95 – 14.95	14.95 – 15.95	15.95 – 16.95	16.95 – 17.95	17.95 – 18.95	المجموع
التكرارات	3	5	15	16	10	1	50

س2 : احسب معامل الاختلاف لدرجات حرارة اجسام 40 طفل ، الموضحة في الجدول لآتي

الفئات	36 – 35.5	36.5 –36	37 – 36.5	37.5 – 37	38 – 37.5	38.5 – 38	المجموع
التكرارات	2	9	15	11	2	1	40

س3: أ- التالي توزيع تكراري لفئات غير متساوية المطلوب رسم مدرج تكراري ؟

التكرارات (f_i)	الفئات
5	10-14
9	15-29
25	30-39
30	40-59
15	60-74
20	75-100
104	**Total**

ب- الجدول التالي يبين احصاءات الوفيات لثلاث سنوات في مدينة زاخو ، والمطلوب حساب :

1- معدل الوفاة الخام لعام 2014.

2- معدل الوفاة الخام للسنوات الثلاث .

السنة	عدد السكان	عدد الوفيات
2013	201921	1928
2014	203211	1930
2015	206212	1937

س4: كانت معدلات الحالة الصحية لسبعة مرضى في احدى المستشفيات نهارا وليلا كالتالي والمطلوب حساب معامل سبيرمان للارتباط.

حالة المريض نهارا (X)	سيئة	مقبولة	جيدة	جيدة جدا	ممتازة	جيدة	سيئة جدا	المجموع
حالة المريض ليلا (Y)	ممتازة	ممتازة	جيدة	مقبولة	سيئة	سيئة	سيئة جدا	**80**

س5: أوجد قيمة معامل الاقتران بين النوع X (ذكر/ أنثى) والإصابة بمرض الاكتئاب Y (مصاب/ غير مصاب) حسب البيانات التالية :

الاكتئاب / النوع	مصاب	غير مصاب	المجموع
ذكر	12	8	
أنثى	4	6	
المجموع			

س6: أراد باحث أن يقيس مدى التوافق بين عيون الاباء وعيون الابناء فقام بجمع البيانات من مجموعة من الاباء والابناء فكانت النتائج كما في الجدول الاتي ،المطلوب : ايجاد معامل التوافق لهذا التوزيع ؟

عيون الاباء / عيون الابناء	سوداء	خضراء	بنية	المجموع
سوداء	2	4	4	
خضراء	6	1	3	
بنية	3	2	5	
المجموع				

س7 : أوجد قيمة الوسط الحسابي والوسيط والمنوال ، ثم حدد شكل الالتواء لمستوى الهيموكلوبين في الدم لعينة مكونة من سبعين شخصا تـم تلخيص مستوى الهيموكلوبين لهم كما في الجدول التالي؟

الفئات	11 - 12	12 - 13	13 - 14	14 - 15	15 – 16	16 - 17	المجموع
التكرارات	9	17	15	8	10	11	70

س8 : احسب معامل الاختلاف للضغط الواطىء لسبعين مريض ، الموضحة في الجدول لآتي

الفئات	9 – 9.5	9.5 – 10	10 – 10.5	10.5 – 11	11 – 11.5	11.5 – 12	المجموع
التكرارات	12	9	15	11	14	9	70

س9: أ- نفترض ان عدد الأحياء في سنة العمر صفر 100000 وعدد الوفيات كما يلي :

السن	0	1	2	3	4	5
عدد الوفيات	240	249	259	268	277	286

والمطلوب ايجاد جدول الحياة بشكل كامل .

ب- الجدول الاتي يمثل توزيع تكراري لاجور اسبوعية لمنتسبي احدى المستشفيات ،المطلوب ايجاد المنوال 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

س10: كانت معدلات درجات الحرارة لسبعة مرضى في احدى المستشفيات نهارا وليلا كالتالي والمطلوب حساب معامل سبيرمان للارتباط.

درجة حرارة المريض نهارا (X)	ممتازة	مقبولة	متوسطة	جيدة جدا	مقبولة	متوسطة	سيئة جدا	المجموع
درجة حرارة المريض ليلا (Y)	ممتازة	جيدة	جيدة	متوسطة	مقبولة	سيئة	سيئة جدا	80

س11: من المعلوم ان عادة التدخين تؤثر تأثيرا سيئا على الصحة العامة للفرد ، المطلوب ايجاد العلاقة بين الحالة الصحية (X) وعادة التدخين (Y) .

الحالة الصحية / عادة التدخين	جيدة	غير جيدة	
يدخن	4	19	
لايدخن	16	8	
المجموع			

س12: أراد باحث أن يقيس مدى التوافق بين شعر الاباء وشعر الابناء فقام بجمع البيانات من مجموعة من الاباء والابناء فكانت النتائج كما في الجدول الاتي ،المطلوب : ايجاد معامل التوافق لهذا التوزيع ؟

شعر الاباء / شعر الابناء	اسود	بني	اصفر	المجموع
اسود	5	4	3	
بني	9	6	5	
اصفر	8	5	7	
المجموع				

س13 : قام مدير المراقبة الصحية في مدينة بغداد 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 .

س14 : احسب معامل الاختلاف لاوزان 50 طفل حديثي الولادة ، الموضحة في الجدول لآتي:

اوزان الاطفال (بالكيلوغرام)	1.5 – 2	2– 2.5	2.5 – 3	3 – 3.5	3.5 – 4	4 – 4.5	المجموع
التكرارات	2	9	15	11	2	11	50

س15: أ- بلغ عدد وفيات الذكور عام 2015 في العراق 245992 كما بلغ عدد السكان الذكور في بداية السنة ونهايتها (14077408 ، 14585117) على التوالي ، احسب معدل الوفيات الخام للذكور ؟

ب- استعمل البيانات التالية لحساب معدل الوفاة المعياري لكل من مدينتي الموصل وكركوك .

العمر	**الموصل**		**كركوك**	
	السكان	**الوفيات**	**السكان**	**الوفيات**
29 -20	45000	90	30000	45
39 -30	40000	120	30000	105
49-40	35000	140	40000	180
59-50	35000	150	50000	255
المجموع	**150000**	**500**	**150000**	**555**

س16: كانت معدلات الحالة الصحية لسبعة مرضى في احدى المستشفيات نهارا وليلا كالتالي والمطلوب حساب معامل سبيرمان للارتباط.

حالة المريض نهارا (X)	سيئة	مقبولة	جيدة	متوسطة	ممتازة	جيدة	مقبولة	**المجموع**
حالة المريض ليلا (Y)	ممتازة	ممتازة	جيدة	مقبولة	متوسطة	سيئة	سيئة جدا	**80**

س17: أوجد قيمة معامل الاقتران بين النوع X (ذكر/ أنثى) والإصابة بمرض النكاف Y (مصاب/ غير مصاب) حسب البيانات التالية :

النكاف / **النوع**	**مصاب**	**غير مصاب**	**المجموع**
ذكر	**60**	**45**	
أنثى	**17**	**3**	
المجموع			

س18: 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 ؟

عيون الاباء / عيون الابناء	دهس	اصطدام	انقلاب	المجموع
صحو	7	3	9	
ممطر	9	5	7	
ضباب	3	7	3	
المجموع				

المصادر

[1]. Bigos, S. J., Battie, M. C., Fisher, L. D., Hansson, T. H., Spengler, D. M., and Nachemson, A. L. [1992]. A prospective evaluation of commonly used pre-employment screening tools for acute industrial back pain. Spine, 17: 922–926.

[2]. Borer, J. S. [1987]. t-PA and the principles of drug approval (editorial). New England Journal of Medicine, 317: 1659–1661.

[3]. Borgen, K. T. [1990]. Of apples, alcohol, and unacceptable risks. Risk Analysis, 10: 199–200.

[4]. Chaitman, B. R., Ryan, T. J., Kronmal, R. A., Foster, E. D., Frommer, P. L., Killip, T., and the CASS Investigators [1990]. Coronary Artery Surgery Study (CASS): comparability of 10 year survival in randomized and randomizable patients. Journal of the American College of Cardiology, 16: 1071–1078.

[5]. Crouch, E. A. C., and Wilson, R. [1982]. Risk Benefit Analysis. Ballinger, Cambridge, MA.

[6]. Fisher, L. D., Giardina, E.-G., Kowy, P. R., Leier, C. V., Lowenthal, D. T., Messerli, F. H., Pratt, C. M., and Ruskin, J. [1987]. The FDA Cardio-Renal Committee replies (letter to the editor). Wall Street Journal, Wed., Aug. 12, p. 19.

[7]. Fisher, L. D., Kaiser, G. C., Davis, K. B., and Mock, M. [1989]. Crossovers in coronary bypass grafting trials: desirable, undesirable, or both? Annals of Thoracic Surgery, 48: 465–466.

[8]. Fisher, L. D., Dixon, D. O., Herson, J., and Frankowski, R. F. [1990]. Analysis of randomized clinical trials: intention to treat. In Statistical Issues in Drug Research and Development, K. E. Peace (ed.).Marcel Dekker, New York, pp. 331–344.

[9]. Fisher, L. D., and Zeh, J. [1991]. An information theory approach to presenting predictive value in the Cox proportional hazards regression model (unpublished).

[10]. International Conference on Harmonisation [2000]. ICH Harmonised Tripartite Guideline: E10. Choice of Control Group and Related Issues in Clinical Trials. http://www.ich.org

[11]. Kaiser, G. C., Davis, K. B., Fisher, L. D., Myers, W. O., Foster, E. D., Passamani, E. R., and Gillespie,

[12]. M. J. [1985]. Survival following coronary artery bypass grafting in patients with severe angina pectoris (CASS) (with discussion). Journal of Thoracic and Cardiovascular Surgery, 89: 513–524.

[13]. Khachaturian, Z. S. [1985]. Diagnosis of Alzheimer's disease. Archives of Neurology, 42: 1097–1105.

[14]. Kowey, P. R., Fisher, L. D., Giardina, E.-G., Leier, C. V., Lowenthal, D. T., Messerli, F. H., and Pratt, C. M. [1988]. The TPA controversy and the drug approval process: the view of the Cardiovascular and Renal Drugs Advisory Committee. Journal of the American Medical Association,260: 2250–2252.

[15]. Lin, L. I. [1989]. A concordance correlation coefficient to evaluate reproducibility. Biometrics, 45: 255–268.

[16]. Yuh, L. [1995],"Statistical Considerations for Bioavailability/Bioequivalence Studies." In Pharmacokinetics: Regulatory, Industrial, Academic Perspectives, 2nd ed. (P.G. Welling and F.L.S. Tse, eds.). New York: Marcel Dekker. pp.479–502

[17]. Zazgonik, J., Huang, M.L., Van Peer, A., et al. [1993],"Pharmacokinetics of Orally Administered Levocabastine in Patients with Renal Insufficiency,"Journal of Clinical Pharmacology, 33:1214–1218.

[18]. Zhi, J., Melia, A.T., Guerciolini, R., et al. [1994],"Retrospective Population–Based Analysis of the Dose–Response (Fecal Fat Excretion) Relationship of Orlistat in Normal and Obese Volunteers,"Clinical Pharmacology & Therapeutics, 56:82–85.

[19]. Zhi, J., Melia, A.T., Guerciolini, R., et al. [1996],"The Effect of Orlistat on the Pharmacokinetics and Pharmacodynamics of Warfarin in Healthy Volunteers,"Journal of Clinical Pharmacology,36:659–666.

[20]. Larry Winner., [2004], " Introduction to Biostatistics" , Department of Statistics University of Florida, Book.

[21]. Gerald V. B., Lloyd D. F., Patrick J. H. et al., [2004]" Biostatistics A Methodology for the Health Sciences" Department of Biostatistics and Department of Environmental and Occupational Health Sciences University of Washington Seattle, Washington, A john wiley & sons, inc., publication, Published by John Wiley & Sons, Inc., Hoboken, New Jersey. Published simultaneously in Canadam Book.

[22]. Marc M. Triola., [2006], " Biostatistics for the Biological and Health Sciences with Statdisk" , Mario F. Triola, Dutchess Community College, Book.

[23]. Mahajan B. K., [2010], " Methods in Biostatistics" , 7th Edition, Book.

[24]. Jay S. K. and Rcnald J. D.,[2008], "Biostatistics for Oral Healthcare", All Free Medical Books.

[25]. B. Burt Gerstman. , [2017], " Basic Biostatics", Copyright 2017 Jones and Bartlett Publishers, Book.

[26]. Kazmi Waqar H., [2015], " Basics In Epidemiology & Biostatistics" , Paperback , Book.

[27]. Steve Selvin., [2015], " A Biostatistics Toolbox for Data Analysis ", R. Relevant R code, data sets, and links to public data sets are available from www.cambridge.org/9781107113084.

[28]. Robert R. Sokal and F. James Rohlf., [2009], " introduction to Biostatistics", Manufactured in the United Stales of America Dover Publications, Inc., 31 Fast 2nd Street, Mineola , N.Y. 1 1501.

[29]. Daryl S. P., [2015], " Applied statistical designs for the researcher" , BioScience Laboratories, Inc. Bozeman, Montana, U.S.A.

[30]. Wayne W. Daniel., [2008], " Biostatistics: A Foundation for Analysis in the Health Sciences", https://www.amazon.in/Biostatistics-Foundation-Analysis-Probability-Statistics.

[32]. Bernard Rosner., [2010], " Fundamentals of Biostatistics", https://www.amazon.in/Fundamentals-Biostatistics-Rosner Biostatics/dp/0538733497.

[33]. Emad Toma Bane , Walaa Ahmed Qazzaz and Wafa Younis Hamoudi, [2015] , " Statistics Science" , Mathematics in Morocco, http://www.mathmaroc.com/2015/03/statistique.html.

دوال في الاكسل (Excel)

=MAX(A2:A22).	القيمة العظمى	Maximum
=MIN(A2:A22) .	القيمة الصغرى	Minimum
=2.5*POWER(n;0.25) .	عدد الفئات	Number of Class
-------	عدد الفئات التقريبي	Approximate Number of Class
Rang Length	طول المدى	Rang Length
Class Length	طول الفئة	Class Length
-------	طول الفئة التقريبي	Approximate Class Length
=SUM(A2:A22) .	المجموع	Summation
=AVERAGE(A2:A22) .	المتوسط الحسابي	Median
=SUMPRODUCT(D33:D37;E33:E37) .	مجموع حاصل الضرب	Sum Product
=ZTEST(A2:A22;4) .	اخبار Z	Z - Test
=AVEDEV(A2:A22) .	الانحراف المتوسط	Standard Mean
=CORREL((A33:A37);(B33:B37)) .	معامل الارتباط	Correlation coefficient
=DEVSQ(A2:A22) .	مجموع مربعات الانحراف	Sum Deviation Boxes
=FREQUENCY(A2:A22;B33:B37) .	التكرارات	Frequency
=GEOMEAN(A2:A22) .	المتوسط الهندسي	Geometric Mean
=HARMEAN(A2:A22) .	المتوسط التوافقي	Harmonic Mean
=MEDIAN(A2:A22) .	الوسيط	Median
=MODE(A2:A22) .	المنوال	Mode
=PERCENTILE((A2:A22);K) .	المئين (K = 0.3)	Percentile
=PERMUT(n;k) .	تباديل (n = 10 , k = 3)	Permute
QUARTILE(A2:A22;1) .	الربيعات(الاول و الثالث)	First Quartile
QUARTILE(A2:A22;3) .	الربيعات(الاول و الثالث)	Third Quartile
=STDEV(A2:A22) .	الانحراف المعياري	Standard deviation
=STDEV(A2:A22)/AVERAGE(A2:A22).	معامل الاختلاف	C . V
=STANDARDIZE(A2;AVERAGE(A2:A22);STDEV(A2:A22)).	القيمة المعيارية	Standard Value
=SKEW(A2:A22).	الالتواء	Sprains
=KURT(A2:A22).	التفلطح	Splaying