الجمهورية اليمنية
وزارة التعليم العالي والبحث العلمي
جامعة الأندلس للعلوم والتقنية
كليه العلوم الادارية

جامعة الأندلس
للعلوم والتقنية
Alandalus University For Science & Technology

الاحصاء

2018-2017م

مدرس المادة

د. محمود العزيزي

مع تحيات مركز للخدمات الطلابية بجامعة الأندلس

مبادئ الإحصاء للعلوم الإدارية

أ.د. عبدالله عباس مهدي المحزري د. محمود عبده حسن العزيزي

المحزري ، عبدالله عباس مهدي العزيزي، محمود عبده حسن .
مبادئ الإحصاء للعلوم الإدارية /

رقم الإيداع بدار الكتب / 2018م

مقدمة الكتاب

الحمد لله القائل " يَرْفَعِ اللَّهُ الَّذِينَ آمَنُوا مِنكُمْ وَالَّذِينَ أُوتُوا الْعِلْمَ دَرَجَاتٍ " (المجادلة، آية ١١). والصلاة والسلام على المبعوث رحمة للعالمين محمد بن عبد الله الصادق الأمين، وعلى آله وصحبه أجمعين، ومن تبعهم بإحسان إلى يوم الدين.

يشهد العالم تطوراً كبيراً في مجال البحث العلمي في مختلف الميادين، الطبية والاقتصادية والاجتماعية والثقافية ... ، ولعل من أهم العوامل التي أدت إلى هذا النشاط توجه معظم دول العالم إلى التنمية ، وحشد الطاقات العلمية المبدعة في كل ميدان لتحقيق أعلى مستوى لها في شتى مجالات الحياة ، كما أن التقدم السريع الناتج عن البحوث والدراسات العلمية في المجالات المختلفة قد انعكس على علم الإحصاء ، من خلال تطوير الأساليب الإحصائية لتلبية متطلبات العلوم المختلفة خصوصاً تلك التي تعتمد على الطابع الكمي ، فالبحث العلمي بأساليبه ووسائله طريق دقيق للوصول إلى المعلومات الصحيحة التي يمكن أن توظف في مجالات متعددة لرفع مستوى معيشة الإنسان وإيصاله إلى مستوى من الرفاهية ، والإحصاء هو أحد الوسائل المهمة التي تزود الباحثين بالمعلومات المعالجة لإلقاء الضوء على جوانب مهمة في دراساتهم التي تقودهم إلى تقدم العلم.

إن الهدف من هذا المقرر (الإحصاء في العلوم الطبية) يتمثل في تعريف الدارسين في العلوم الطبية ، والتخصصات الأخرى بعلم الإحصاء عن طريق تقديم عرض مبسط للمفاهيم الإحصائية الأساسية ، وطرق القياس وأدواته و قد رُوعي أثناء تقديم المادة العلمية لهذا الكتاب شرح المفاهيم بأسلوب واضح ولغة سليمة ومن ثم إعطاء المعادلات بشكل رياضي دقيق ، وتبع ذلك أمثلة محلولة ذات صلة بموضوع الكتاب، يتبع معظم هذه الأمثلة أنشطة تنفذ بعد شرح المفهوم والمثال ، ثم تدريبات تحل خارج وقت المحاضرة، وفي نهاية كل فصل مجموعة من التمارين التي تكسب الطالب المهارات اللازمة

التي تمكنه من التعامل مع البيانات الإحصائية على مستوى مادة هذا الكتاب ، يلي ذلك ملحق نهاية الكتاب يتضمن تمارين متنوعة وشاملة تقريباً لكل موضوعات الكتاب لِتُعين الطالب على التمكن من مادة الكتاب.

وقد احتوى هذا الكتاب على ستة فصول هي :

- الفصل الأول : مفاهيم إحصائية أساسية.
- الفصل الثاني : طرق عرض البيانات الإحصائية .
- الفصل الثالث : مقاييس النزعة المركزية .
- الفصل الرابع : مقاييس التشتت.
- الفصل الخامس : الارتباط والانحدار الخطي البسيط.
- الفصل السادس : السلاسل الزمنية.
- ملحق الكتاب : أسئلة تقويمية شاملة.

وقبل الختام نُرَحب بأي تغذية راجعة تُسهم في تحسين مادة هذا الكتاب.

ونسأل المولى عز وجل أن يكون هذا العمل خالصاً لوجهه الكريم وأن أكون قد وفقت بتزويد المكتبة العربية بكتاب أعتقد حاجتها إليه، وأن يكون هذا العمل من العلم الذي يُنْتَفع به.

المؤلف

د. محمود عبده حسن العزيزي

قائمة المحتويات

1

الفصل الأول

مفاهيم إحصائية أساسية

محتويات الفصل الأول

أهداف الفصل الأول

يتوقع من الطالب بعد الانتهاء من دراسة هذا الفصل أن يكون قادراً على:

- ✓ تعريف علم الإحصاء.
- ✓ التمييز بين الإحصاء الوصفي والإحصاء الاستدلالي.
- ✓ توضيح أهمية علم الإحصاء وميادينه.
- ✓ تحديد عناصر علم الإحصاء ووظائفه.
- ✓ التمييز بين البيانات الكمية والنوعية.
- ✓ معرفة مصادر جمع البيانات الإحصائية.
- ✓ معرفة طُرق جمع البيانات الإحصائية .
- ✓ معرفة أنواع المتغيرات .
- ✓ التمييز بين المتغيرات والثوابت .
- ✓ التمييز بين المتغير المتصل والمنفصل.
- ✓ معرفة أنواع المقاييس الإحصائية.
- ✓ التمييز بين المقاييس المختلفة.

مقدمة Introduction:

الحمدُ لله رب العالمين القائل في كتابه الكريم (وَوُضِعَ الْكِتَابُ فَتَرَى الْمُجْرِمِينَ مُشْفِقِينَ مِمَّا فِيهِ وَيَقُولُونَ يَا وَيْلَتَنَا مَالِ هَذَا الْكِتَابِ لَا يُغَادِرُ صَغِيرَةً وَلَا كَبِيرَةً إِلَّا أَحْصَاهَا وَوَجَدُوا مَا عَمِلُوا حَاضِراً وَلَا يَظْلِمُ رَبُّكَ أَحَداً) الكهف، اية 49.

أحصاها هنا بمعنى: أحاط بها وجمعها وعلمها .

والقائل سبحانه : (ثُمَّ بَعَثْنَاهُمْ لِنَعْلَمَ أَيُّ الْحِزْبَيْنِ أَحْصَى لِمَا لَبِثُوا أَمَداً) الكهف، اية 12

والقائل : (لِيَعْلَمَ أَن قَدْ أَبْلَغُوا رِسَالَاتِ رَبِّهِمْ وَأَحَاطَ بِمَا لَدَيْهِمْ وَأَحْصَى كُلَّ شَيْءٍ عَدَداً) الجن، اية 28

والقائل : (لَقَدْ أَحْصَاهُمْ وَعَدَّهُمْ عَدّاً) مريم، اية 94

والقائل : (يَوْمَ يَبْعَثُهُمُ اللَّهُ جَمِيعاً فَيُنَبِّئُهُم بِمَا عَمِلُوا أَحْصَاهُ اللَّهُ وَنَسُوهُ وَاللَّهُ عَلَى كُلِّ شَيْءٍ شَهِيدٌ) المجادلة، اية 6

و القائل سبحانه: (وَآتَاكُم مِّن كُلِّ مَا سَأَلْتُمُوهُ وَإِن تَعُدُّواْ نِعْمَتَ اللَّهِ لاَ تُحْصُوهَا إِنَّ الإِنسَانَ لَظَلُومٌ كَفَّارٌ) إبراهيم، اية 34

معنى لا تحصوها: لا تطيقوا عدها ولا تستطيعوا حصرها لكثرتها.

ومع أن الآيات السالفة الذكر لم يُقصد بها تعليمنا الإحصاء أو معناه أو الفائدة منه، ولكننا نفهم منها أن الخالق سبحانه وتعالى قادر على حصر المعلومات و عدها و العلم بكل صغيرة وكبيرة متعلقة بها.

أما الآية الأخيرة فتدل على أن الإنسان عاجز عن إجراء المسح الشامل لمجتمع نعم الله، ولذلك فلا بُد من دراسة عينة من هذا المجتمع عندما يريد الإنسان أن يتفكر في هذه النعم.

نتطلق من هنا لنلاحظ أن للإحصاء علاقة بدراسة المجتمع ودراسة أجزاء منه، ودراسة نشاطاته، وليس المقصود بالمجتمع هنا المجتمع الإنساني بالضرورة. ولكن المقصود أي مجموعة من العناصر التي تخضع للدراسة أو البحث مثل مجموعة المستشفيات أو المراكز الصحية في مدينة ما، أو مجموعة الأمراض المنتشرة في بلدة ما أو فترة ما، أو مجموعة الطلاب في جامعة ما .... و هكذا.

❖ نشأة علم الإحصاء وتطوره :

إن نشأة علم الإحصاء بمعنى الحصر والعد فكرة قديمة وربما كان قدماء المصريين أول من قام بتطبيقها و استخدامها حيثُ قام بُناة الأهرام بتعداد سكان مصر وثرواتها واستخدموا النتائج في تنظيم مشروع البناء.

وفي عصر الدولة الإسلامية استخدم الخليفة المأمون فكرة الحصر والعد لمعرفة عدد السكان، ومقدار الزكاة، بغرض التعرف على ما لديهم من الرجال وقدرتهم على الدفاع عن بلدانهم.

وكان استخدام الإحصاء في البداية مقصوراً على الأعمال الخاصة بشؤون الدولة، كما يدل على ذلك الأصل اللغوي في اسم هذا العلم، وهو باللغة الإنجليزية Statistics وهو مُشتق من كلمة State أي الدولة، ومعناه "مجموعة الحقائق الخاصة بشؤون الدولة"، ثم تطور علم الإحصاء بعد ذلك حتى أصبح يشمل معظم ميادين الحياة المعاصرة العلمية والعملية، وأصبحت له تطبيقات في مجالات البحث المختلفة وعلى رأسها البحوث في العلوم التطبيقية والطبية.

❖ تعريف علم الإحصاء:

أ- التعريف اللغوي:

الإحصاء لغة: بمعنى الحصر أو العد.

ويؤكد هذا المعنى الآيات القرآنية الآتية:

يقول تعالى: (وَأَحَاطَ بِمَا لَدَيْهِمْ وَأَحْصَى كُلَّ شَيْءٍ عَدَداً) الجن: 28

ويقول سبحانه وتعالى: (لَقَدْ أَحْصَاهُمْ وَعَدَّهُمْ عَدّاً) مريم: 94

ب- التعريف الاصطلاحي:

لقد وردت العديد من التعريفات لعلم الإحصاء ومن هذه التعريفات:

1. هو علم يدرس كيفية جمع وترتيب وتنظيم وتلخيص وتحليل البيانات واستخلاص النتائج وتفسيرها، بهدف الوصول إلى نتائج مقبولة و قرارات سليمة.
2. هو العلم الذي يهتم بجمع وتنظيم وتلخيص وعرض البيانات العددية أو الرقمية ثم اتخاذ القرار المبرر في ضوء تحليل النتائج.
3. هو مجموعة من الإجراءات والأساليب المستخدمة في جمع وعرض وتحليل البيانات التي تُبنى عليها القرارات في مواجهة عدم التأكد أو مواجهة معلومات ناقصة.

مما سبق يمكن القول بأن الإحصاء علم يهتم بجمع المعلومات أو البيانات وتنظيمها وتصنيفها وتبويبها وتلخيصها وتحليلها وعرضها عن طريق الجداول أو الرسوم البيانية أو غيرها بغرض الوصول منها إلى استنتاجات وقرارات مناسبة.

تدريب (1)

عرف علم الإحصاء لغةً واصطلاحاً؟

❖ عناصر علم الإحصاء:

لعلم الإحصاء عدد من العناصر أهمها:

1. جمع البيانات أو المعلومات Collection of Data or Information:

ويتم جمع البيانات بعدة طرق منها أسلوب الحصر الشامل (أي أخذ مجتمع الدراسة كاملاً) وأسلوب العينات (أي أخذ عينة من مجتمع الدراسة بشرط أن تكون ممثلة تمثيلاً جيداً لمجتمع البحث أو الدراسة) وسيتم تناول الموضوع بالتفصيل في فصل مستقل من هذا الكتاب.

2. تصنيف البيانات Classification of Data:

وتصنف البيانات إلى وصفية (نوعية) أو كمية (عددية).

3. التبويب:

ويقصد به وضع البيانات في جداول عادية أو مزدوجة أو مركبة وعمل التوزيعات التكرارية لها.

4. عرض البيانات Presentation of Data:

ويتم عرضها لسهولة التعرف عليها وتُعرض بعدة طرق منها:

الجداول بأنواعها – الرسومات البيانية بأنواعها.. وغيرها، وسيأتي توضيح ذلك لاحقاً.

5. تحليل البيانات Analysis of Data:

ويتم فيها دراسة الخصائص الأساسية للظاهرة المدروسة.

6. اختبار المؤشرات:

ويعني اختبارها من حيث كونها حقيقية أو معنوية وإمكانية التنبؤ والتقدير بناءً عليها.

7. استخلاص النتائج واتخاذ القرارات:

ويعني وصول الباحث إلى اتخاذ القرارات السليمة والوصول إلى حلول عامة لهذه المشكلات والمشكلات المشابهة لها.

❖ أهمية علم الإحصاء:

مما لا شك فيه أن علم الإحصاء يحتل مكاناً مرموقاً بين بقية العلوم في عصرنا الراهن، عصر المعلوماتية واقتصاد المعرفة، خاصة بعد ازدياد الطلب على البيانات من قبل الدول والمؤسسات والهيئات الدولية لرسم سياسات التنمية الاقتصادية والاجتماعية.

وتتجلى أهميته باعتباره يدخل في جميع مجالات الحياة العامة التي يستخدمها الأفراد على مختلف مستوياتهم العلمية والثقافية حتى أصبح التفكير الإحصائي ضرورة لكل فرد، كما أن علم الإحصاء أصبح من الدعائم الأساسية للبحوث العلمية في شتى المجالات .

وتتضح أهمية علم الإحصاء فيما يلي :

- يُساعد على تلخيص البيانات واستخلاص النافع منها .
- يُساعد على تخطيط وتصميم التجارب وعمل المسح الإحصائي.
- يُساعد على اختيار الأساليب البحثية المناسبة .
- يُساعد على كيفية استخدام نتائج البحث الإحصائي ، إذ تُستخدم النتائج فيما يلي:

أ . التنبؤ .

ب. اتخاذ القرار.

ج . التحقق : أي التثبت من صحة أو عدم صحة فرضية ما .

د . الرقابة على مدى الجودة والرقابة الكمية .

- يُساعد الباحث على تحليل الظاهرة قيد الدراسة بصورة علمية دقيقة .
- يُساعد الباحث على التنبؤ بالنتائج وبالتالي وضوح الرؤية المستقبلية للظاهرة .
- يُساعد الباحث على تعميم النتائج لموضوع الدراسة .
- يُساعد الباحث على معرفة تأثير كل عامل من العوامل المؤثرة على الظاهرة .

❖ ميادين علم الإحصاء:

نظراً لاستخدام الإحصاء في مجالات وميادين متعددة ومتغيرة فقد أصبح يُطلق على الإحصاء أسماء مختلفة تبعاً للمجال التطبيقي المنوط به كالإحصاء البيولوجي والإحصاء الجيني والإحصاء الطبي والإحصاء الزراعي والإحصاء الصناعي والإحصاء التربوي والإحصاء الجيوفيزيائي والإحصاء الفلكي وله العديد من التطبيقات الحياتية الأخرى.

فعلم الإحصاء قد طرق ميادين متعددة بعيدة عن المجال الأصلي الذي نشأ فيه هذا العلم ألا وهو شؤون الدولة، فاستُخدم في بحث المسائل العلمية المختلفة في شتى المجالات، ومن الميادين التي طرقها علم الإحصاء:

1 . العلوم الاقتصادية والإدارية والاجتماعية والتربوية:

يُعد علم الإحصاء المحك الذي يلجأ إليه رجال الاقتصاد لتفسير الظواهر الاقتصادية فعمل الإحصاءات عن التجارة الداخلية والخارجية تفيد في التعرف على الضرائب والجمارك وحجم التجارة. وكثيراً ما تستعمل المؤسسات والشركات الطرق الإحصائية لتحليل نماذج التغير والتنبؤ للنشاطات المستقبلية لتلك المؤسسات أو للاقتصاد بشكل عام. إن كثيراً من هذه التنبؤات يُرسي قواعد التخطيط والتحكم ، وبالإضافة للتنبؤ فإن حقولاً مثل التحكم في الإنتاج والتحكم في النوعية تستعمل الطرق الإحصائية كقاعدة أساسية.

وتستعمل المفاهيم الإحصائية في دراسات القوى العاملة واختيار الموظفين وأبحاث التسويق والإعلان والتحكم في المخزون والتجارب الصناعية والتحليل المالي وتدقيق الحسابات وتوظيف رؤوس الأموال والتنمية.

أما في الإدارة العامة والعلوم الاجتماعية فإن الطرق الإحصائية تُستخدم بكثرة وخاصة في تحليل القضايا السياسية والاجتماعية ، حيثُ يُعد الإحصاء في العلوم الاجتماعية وسيلة مهمة لقياس مدى رفاهية الناس وتقدم المجتمع و مستوى رقي الأفراد ثقافياً واجتماعياً وصحياً .

ويستخدم الإحصاء في انجاز الدراسات الكمية عن المجتمعات وظاهرة الفقر وحوادث المركبات وفي النماذج الانتخابية والتعداد السكاني والأمور التربوية وغيرها من الموضوعات.

2 . العلوم الطبيعية البحتة:

يُستخدم علم الإحصاء في علم الفلك في تحليل المشاهدات المتعلقة برصد النجوم والكواكب وفي التعرف على أحوال الجو وتحليل الظواهر الجوية والتنبؤ بأحوال الجو.

أما في علم الأحياء فقد استخدم الإحصاء في دراسة الأجناس والفصائل المختلفة للحيوان والنبات ومعرفة خصائص وصفات كل جنس، والاختلاف بين أفراد الجنس الواحد في تلك الخصائص التي تميز ذلك الجنس عن الأجناس الأخرى.

3 . العلوم الطبية:

للإحصاء دور مُهم جداً في مجال العلوم الطبية والصحية فيفيد الإحصاء الباحث في مجال الطب في التعرف على الظواهر الطبية وتفسيرها فعن طريق دراسة الرسوم البيانية والمنحنيات التي ترسلها الأجهزة الحديثة ومعرفة دلالاتها يستطيع المختص تشخيص حالة المريض وعلاجه. كما أنها

تفيد الباحثين في عمل إحصائيات عن المستشفيات والحالة الصحية والمرضية للمواطنين كالعلاقة بين ضغط الدم والعمر والوزن.

كما يلعب الإحصاء الطبي دوراً حيوياً في دراسة القضايا التي تُسهم في حل المشكلات المتعلقة بالأمراض وبناء النماذج الإحصائية لمعظم التطبيقات الطبية والتجارب العملية، وكذا التجارب المتعلقة بالأدوية ومدى إمكانية استخدامها في الطب العلاجي.

بناءً على ما سبق إيضاحه يُعد علم الإحصاء في عصرنا الحاضر هو الدعامة الأساسية التي تبني عليها كل تخطيط سليم في مختلف ميادين الحياة.

❖ وظائف علم الإحصاء :

توجد ثلاث وظائف رئيسية لعلم الإحصاء كما هي موضحة في الشكل الآتي:

أ. وظيفة وصفية : تعد من أهم الوظائف وتتمثل في وصف البيانات حيث يتم استكشاف الطبيعة الخاصة بالظواهر المختلفة أو المشكلات المطلوب دراستها وتحليلها من خلال جمع الحقائق أو البيانات عن الظاهرة المدروسة وتسجيلها في صورة رقمية أو وصفية وتصنيفها في جداول تكرارية ، وتمثيلها في أشكال بيانية ، بالإضافة إلى حساب بعض المؤشرات الإحصائية البسيطة للوقوف على طبيعة هذه البيانات وتحديد أهم خصائصها واتجاهاتها.

ب . وظيفة تحليلية (استدلالية) : من الوظائف المهمة التي تستخدم في مجال البحث العلمي ، وتعتمد على اختيار جزء من مجتمع الدراسة ليتم دراسته بهدف الوصول إلى نتائج يمكن تعميمها على المجتمع من خلال وضع فروض معينة يتم اختبار صحتها من عدمه واستخلاص النتائج واتخاذ القرارات المناسبة على أسس علمية .

ج - وظيفة تنبؤية : تعتمد هذه الوظيفة على دراسة سلوك المتغيرات خلال الفترة الماضية للظاهرة التي يُراد التنبؤ بسلوكها مستقبلاً ، كما تتطلب عملية التنبؤ مقارنة الظواهر بعضها ببعض ودراسة طبيعة العلاقة التي ترتبط بها تلك الظواهر لفهم واستيعاب المؤثرات ليكون التنبؤ ذو مصداقية .

> تدريب (2)
>
> وضح أهمية علم الإحصاء مبيناً مجالاته ووظائفه ؟

❖ أقسام علم الإحصاء:

ينقسم الإحصاء بصورة رئيسية إلى قسمين أساسيين يوضحهما الشكل الآتي:

أولاً / الإحصاء الوصفي (Descriptive Statistics).

هو ذلك الجزء من الإحصاء الذي يهتم بجمع البيانات وتنظيمها وتصنيفها وعرضها عن طريق الجداول أو الرسوم البيانية وغيرها.

يهتم علم الإحصاء الوصفي بالأساليب والطرق العلمية التي يُناط بها جمع وتبويب البيانات وعرضها في جداول أو رسوم بيانية وكذلك حساب المقاييس الإحصائية البسيطة مثل مقاييس النزعة المركزية ومقاييس التشتت وغيرها ، وهذه المقاييس تصف أو تشرح بيانات الظاهرة قيد الدراسة.

ويتعامل الإحصاء الوصفي مع البيانات من غير إضافة أو تعميم أو تنبؤ لسلوك الظاهرة .

فمثلاً، عندما يزور طبيب الأسنان إحدى المدارس ويفحص أسنان كل طالب يضع جدولاً يبين فيه عدد الطلبة الذين يعانون من تسوس في أسنانهم ثم يلخص هذا الجدول حسب عدد الأسنان التي أصابها التسوس ، فيذكر عدد الطلبة الذين يعانون تسوساً في سن واحدة ، وعدد الذين يعانون تسوساً في أكثر من سن ، فإنه يستعمل الإحصاء الوصفي في هذه الحالة.

ثانياً / الإحصاء الاستدلالي (التحليلي) (Inferential Statistics) :

هو ذلك الجزء من الإحصاء الذي يهتم بتحليل البيانات للتوصل إلى التنبؤ أو الاستنتاج ، واتخاذ القرارات المناسبة ، أو الوصول إلى تعميم عن خواص الكل (المجتمع) من واقع فحص جزء من هذا الكل (العينة).

أي أنه يهتم بالأساليب والطرق العلمية التي يُناط بها تحليل وتفسير واستخلاص النتائج وتعميمها واتخاذ قرار ، كما يهتم بالسلوك المستقبلي للظاهرة وذلك بناءً على المعلومات الجزئية (من العينة) والنتائج التي نحصل عليها باستخدام الأساليب والطرق الوصفية .

وينقسم الإحصاء التحليلي إلى قسمين هما : نظرية التقدير واختبار الفروض.

ويلعب هذا الجزء من الإحصاء دوراً مُهماً في تخطيط التجارب التي تُجمع منها البيانات وفي تصميمها

ويُمكننا القول إجمالاً أن الطرق التي تهتم بالبيانات المتوفرة فقط ولا تحاول التعميم من العينة المدروسة إلى مجتمع أكبر هي طرق الإحصاء الوصفي.

أما المعالجات التي تؤدي إلى تنبؤ أو استنتاج أو تعميم إلى مجموعة كبيرة من عينة منها فهي طُرق الإحصاء الاستدلالي (التحليلي).

> تدريب (3)
> ميز بين الإحصاء الوصفي والإحصاء الاستدلالي؟

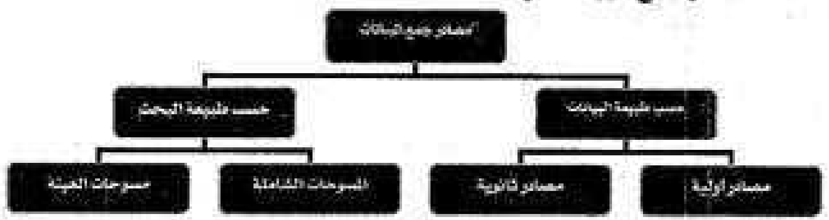
- البيانات الإحصائية Statistical data :

هي مجموعة الحقائق والمعلومات التي تتعلق بظاهرة ما وتُشكل المادة الخام لعلم الإحصاء .

وتنقسم البيانات الإحصائية إلى قسمين Types of Data :

1. البيانات الكمية Quantitative Data
2. البيانات النوعية (الوصفية) Qualitative Data

- مصادر جمع البيانات الإحصائية Sources of Data



1. مصادر البيانات حسب طبيعة البيانات .

وتنقسم إلى نوعين هما :

1. المصادر الأولية: وتحصل عليها من الأفراد والمجموعات ومقاييس تتبع الأثر .

أما البيانات الأولية: فهي بيانات تُجمع من قبل الباحث نفسه من مجتمع خاص لمشروع البحث ، وتكون تحت يديه ومن أدوات جمع البيانات الأولية :

1. الاستبيان Questionnaire ، وهو أداة لجمع البيانات الأولية تتمثل في مجموعة من الأسئلة المكتوبة تتعلق بظاهرة ما يُطلب من المستجيب الإجابة عنها.

2. المقابلة Interview : هي تفاعل لفظي يتم بين شخصين في موقف مواجهة حيث يحاول أحدهم وهو القائم بالمقابلة أن يستثير بعض المعلومات أو التعبيرات لدى المبحوث والتي تدور حول آرائه ومعتقداته.

3. الملاحظة Observation ، وتعني الانتباه المقصود والموجه نحو سلوك فردي أو جماعي مُعين بقصد متابعته ورصد تغيراته ليتمكن الباحث بذلك من : وصف السلوك فقط أو وصفه وتحليله أو وصفه وتقويمه.

ويعرفها البعض بأنها مشاهدة الظواهر من قبل الباحث أو من ينوب عنه بقصد تفسيرها واكتشاف أسبابها والتنبؤ بسلوك الظاهرة والوصول إلى القوانين التي تحكمها.

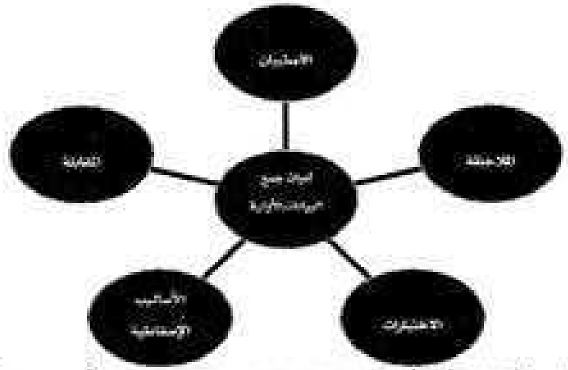
وقد يُراقب الباحث ظواهر يُمكن أن يُؤثر فيها كالتجارب في المختبرات، أو تجارب لا يستطيع التأثير فيها مثل علوم الفلك.

4. الاختبارات Tests : هي مجموعة من المثيرات (أسئلة شفوية أو كتابية أو صور أو رسوم) أعدت لتقيس بطريقة كمية أو كيفية سلوكاً ما، والاختبار يُعطي درجة ما أو قيمة ما أو رتبة ما للمفحوص . ويُمكن أن يكون الاختبار مجموعة من الأسئلة أو جهازاً معيناً.

وتُستخدم الاختبارات في القياس والكشف عن الفروق بين الأفراد والفروق بين الجماعات والفروق بين الأعمال.

5. طُرق الإسقاط Projective Method : هي أسلوب للحصول على الإجابات عن طريق استخدام المعاني المرتبطة بالكلمات ، وإكمال الجمل ، واختبارات الإسقاط ، والتي بدونها يصعب الحصول على تلك الإجابات .

والشكل الآتي يوضح أدوات جمع البيانات الأولية :



ب. **المصادر الثانوية (المكتبية)** : ويُقصد بالبيانات الثانوية البيانات التي جُمعت سابقاً من قبل باحث آخر أو مشروع آخر وغالباً ما يكون الاعتماد عليها أقل تكلفةً ووقتاً. ومن أمثلة مصادر البيانات الثانوية:

- الكتب والمؤلفات ذات العلاقة .
- المجلات العلمية والأبحاث المنشورة .
- الوثائق .
- الأطروحات الجامعية (أطروحات الدكتوراه ورسائل الماجستير).
- الانترنت وما يحويه من مصادر الكترونية مختلفة.
- قواعد البيانات المختلفة.
- بيانات دائرة الإحصاءات العامة ، وبيانات الوزارات والدوائر المختلفة.

وتفضل المصادر الأولية على غيرها من المصادر من قبل المهتمين بالبحث العلمي في حال توفرها لأنها تقود إلى معلومات أكثر دقة وواقعية مقارنة بالبيانات الثانوية.

2- **مصادر البيانات حسب طبيعة البحث.**

وتنقسم إلى نوعين هما :

أ- **بيانات المجتمع** Population Data، وهي البيانات التي تُجمع من مجموعة كلية أو كينونة كاملة يتشاركون في قواعد عامة من الخصائص وقد تكون مجموعة من الأفراد أو الظواهر أو المنشآت ، أي تشمل كافة وحدات المجتمع الإحصائي.

ب. **بيانات العينة** Sample Data، وهي البيانات التي تُجمع من مجموعة جزئية تُمثل جزء من المجتمع أو عدد من الحالات التي تؤخذ من المجتمع الأصلي ، أي استخدام بيانات تُمثل عدد من الوحدات أو جزء من المجتمع الكلي للوصول إلى استنتاجات عن المجتمع الكلي معتمدين على التحليل الإحصائي.

> تدريب (4)
> حدد مصادر جمع البيانات الإحصائية ؟

> تدريب (5)
> ما هي طُرق جمع البيانات الأولية ؟
> أي الطُرق السابقة مناسبة بصورة أفضل لجمع البيانات في البحوث الطبية ؟ مُبرراً لما تقول ؟

❖ المتغيرات الإحصائية (Statistical Variables).

المتغير Variable، هو خاصية قابلة للتغير من مشاهدة لأخرى في المجتمع الإحصائي مثل ، الطول ، الوزن ، الضغط ، العمر ، الحالة الاجتماعية ، الحالة التعليمية ، الدخل ، الجنس ، اللون ، نوع العمل، فصيلة الدم، الديانة ، الجنسية ، درجة الحرارة.

ويعرف أيضاً بأنه تلك السمة أو الصفة أو الكمية التي تتغير قيمتها من عنصر إلى آخر أو من مشاهدة إلى أخرى. مثلاً: لو أردت قياس أوزان مجموعة من المرضى لحصلت على مجموعة من الأوزان يُمثل كلاً منها وزن مريض، أي أن الوزن متغير.

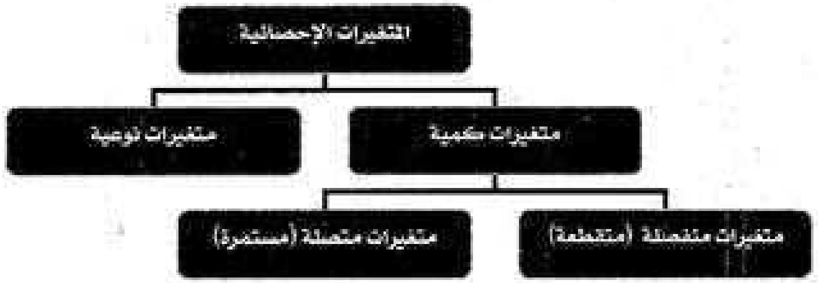
وكمثال آخر: لو سجلت درجات الحرارة في مدينة معينة كل يوم لمدة شهر، فإنك ستحصل على عدد من القيم التي تُمثل درجات الحرارة، وبالتالي فإن درجة الحرارة في تلك المدينة تُعد متغير.

وإليك أمثلة أخرى على المتغيرات:

- ضغط الدم لدى شخص عند قياسه يومياً لمدة شهر.
- مستوى السكر في الدم لدى مجموعة من المرضى.

والأمثلة على المتغيرات كثيرة جداً، اذكر بعضاً منها؟

وتُصنف المتغيرات الإحصائية إلى نوعين يوضحهما الشكل الآتي :



أولاً / المتغيرات النوعية (الوصفية) Qualitative Variables :

وهي المتغيرات التي تأخذ صفات معينة غير رقمية (متغيرات غير قابلة للقياس أو العد) ، مثل:

- الوظيفة (مدير عام، مدير إدارة، رئيس قسم، مشرف).
- الحالة التعليمية (أمي، يقرأ ويكتب، أساسي، ثانوي، جامعي فأكثر).
- الحالة الحضرية (ريف، حضر).
- الحالة الاجتماعية (متزوج، مطلق، أرمل، أعزب).
- الجنس (ذكر، أنثى).
- لون الشعر (أشقر، أحمر، أسود، ...).
- لون العيون (أزرق، عسلي، أسود، أخضر، ...).
- فصيلة الدم (O, A, B, AB,...).
- مستوى الدخل (عالي، متوسط، منخفض)

ثانياً/ المتغيرات الكمية (العددية) Quantitative Variables :

هي المتغيرات القابلة للقياس أو العد ، ويُعبر عنها بصور رقمية مثل :

- عدد المنظمات العاملة في دولة ما.
- عدد الموظفين في إحدى الشركات.
- عدد الطلبة الملتحقين بكلية العلوم الإدارية.
- عدد شركات القطاع الخاص في اليمن.

بمعنى عندما تكون السمة تحت الدراسة قابلة للقياس على مقياس عددي فإن المتغيرات التي تحصل عليها تتألف من مجموعة من الأعداد وتسمى متغيرات كمية أو عددية. والأمثلة على المتغيرات العددية كثيرة مثل:

- رواتب الموظفين في إحدى المؤسسات.
- درجة حرارة المريض.
- عدد أفراد الأسرة.
- درجات الطلاب في مقرر الإحصاء في العلوم الإدارية.

وتنقسم المتغيرات الكمية إلى قسمين هما :

أ. المتغيرات الكمية المنفصلة (المتقطعة) Discrete Variables :

هي تلك المتغيرات التي تأخذ قيماً قابلة للعد (قيم عددية صحيحة غير قابلة للتجزئة إلى قيم كسرية) ، أي أن قيم المشاهدات أو الملاحظات لهذه المتغيرات تكون متباعدة أو متقطعة.

ومن أمثلة المتغيرات المنفصلة :

- عدد الأطباء في أحد المستشفيات .
- عدد أفراد الأسرة .

ب. المتغيرات الكمية المتصلة (المستمرة) Continuous Variables :

هي المتغيرات التي يمكن أن يُعطى لها أي قيم رقمية ضمن مدى (فترة) محدد ومعلوم، وتكون قابلة للتجزئة إلى قيم كسرية، فعلى سبيل المثال نقول أن الفترة المحصورة بين الصفر والواحد بها عدد لا نهائي من الأعداد فهي بالتالي فترة مستمرة .

ومن أمثلة المتغيرات المتصلة: الأطوال والأوزان ودرجة الحرارة والأعمار، ويمكن قياس مثل هذه المتغيرات بأجزاء صغيرة جداً.

تدريب (6)

حدد نوع المتغيرات الآتية من حيث كونها متصلة أو منفصلة :

1- عدد الموظفين بإحدى الشركات.
2- عُمر أحد المرضى.
3- درجة حرارة مريض.

ما الفرق بين المتغيرات الكمية والمتغيرات النوعية ؟

الثوابت Constants:

الثابت Constant: هي تلك السمة أو الخاصية التي لا تتغير وهي تصف ماهية المواد في ظروف معينة، مثل: الكثافة النوعية لعنصر ما في ظرف محدد، فمثلاً الكثافة النوعية للماء النقي في درجة الحرارة العادية هي 1 جم لكل سم مكعب، كذلك معامل التمدد لعنصر الحديد النقي ثابت، معامل الاحتكاك بين مادتين محددتين ثابت، وهناك كثير من الثوابت الفيزيائية والرياضية وهناك ثوابت من نوع آخر تقع ضمن اهتمامات الإحصائي أو الباحث وهي تلك الثوابت التي تصنف المجتمعات مثل: معدل عمر المصابيح الكهربائية التي يُنتجها مصنع معين على مدى عام واحد، أو معدل الدخل السنوي للفرد في بلد معين خلال سنة معينة.

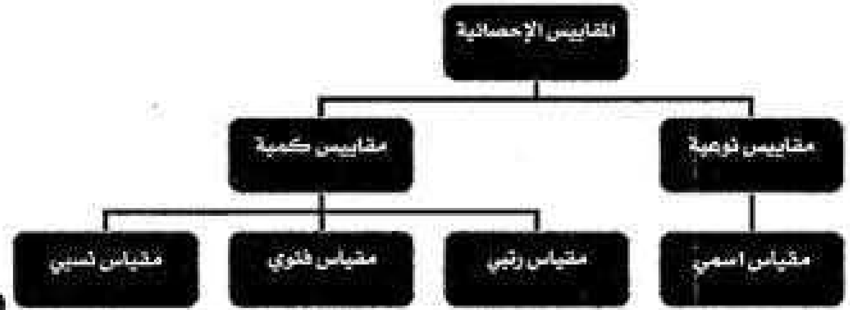
> تدريب (7)
> ما الفرق بين المتغير والثابت مع التوضيح بمثال لكل منهما؟

❖ مفهوم القياس Concept of the measurement.

لو أراد أحد الأطباء قياس تأثير أحد الأدوية على مرضاه فإنه يحتاج إلى ملاحظة التغير الحاصل على المرضى كالتغيرات التي تحدث في تحسن صحتهم خلال فترة زمنية محددة، بالإضافة إلى إمكانية استخدام هذا العلاج في حالات مشابهة، فإنه يحتاج إلى تدوين مشاهداته أو ملاحظاته بصورة تمكنه من التعميم إلى حالات أخرى، من هُنا قد يُعطي لهذه النتائج أي البيانات الملاحظة رموزاً رقمية للدلالة عليها، إن عملية إعطاء أرقام خاصة للأشياء هي ما يُعرف عادةً بالقياس.

إذاً القياس: هو نظام تصنيفي تُعطى فيه الأشياء أرقاماً خاصة بها لكي يسهل تسجيل وتلخيص الملاحظات ومعالجتها (تحليلها) إحصائياً.

وتوجد مقاييس مختلفة تنقسم حسب نوع المتغيرات إلى قسمين أساسيين والموضحة بالشكل الآتي:



أولاً / مقاييس المتغيرات النوعية وتتمثل في المقياس الاسمي Nominal Scale:

يُمثل هذا المستوى أدنى مستويات القياس أي أبسط أنواع المقاييس، حيث يستخدم الأرقام من أجل الدلالة على الأشياء أو مجموعات الأشياء ويُستخدم في معظم الأحوال مع المتغيرات النوعية، حيث يتولى هذا المقياس تصنيف الموضوعات أو الأشياء أو الأفراد إلى مجموعات مختلفة وفقاً لبعض الخصائص النوعية كتوزيع الأفراد حسب جنسهم (ذكر، أنثى) أو حسب الحالة الاجتماعية (متزوج، أعزب، أرمل، مطلق)، أو حسب الجنسية (يمني، سعودي، مصري، ....)، أو حسب محل الإقامة (شمال، شرق، جنوب، غرب) أو غيرها.

ولا تعمل هذه المقاييس بأكثر من تصنيف الأشياء من أجل التمييز بينها، اعتماداً على افتراض أن الأفراد يختلفون في صفة ما، ولتسهيل التعامل مع هذه المتغيرات وتخزينها ومن ثم عرضها في الحاسوب أصبح من الضروري تكميمها أي التعبير عنها رقمياً وذلك بإعطاء كل صفة رقماً أو مقداراً للتعرف عليه وتصنيفه فقط، وهذا الرقم لا يفيد في أكثر من التسمية أو التصنيف، إذ أن الأرقام في هذا المستوى أشبه بالأسماء والألقاب ولا تتضمن معنى للأفضلية (الأكبر والأصغر).

فمثلاً: إذا أعطي الرقم 1 للذكر والرقم 2 للأنثى فهذا لا يعني أن الذكر أقل من الأنثى أو أن الأنثى أفضل من الذكر، مع ملاحظة أن بداية العد والفروق بين الأرقام لا تؤثران في المقياس، كذلك قد نُعطي المتزوج الرقم 1 والأعزب الرقم 2 والأرمل الرقم 3 والمطلق الرقم 4 و هكذا، كما أن الأرقام في هذا المستوى غير قابلة لإجراء العمليات الحسابية الأربع (الجمع، الطرح، الضرب، القسمة) فلا يصح في هذا المقياس جمع الأرقام حيث لا معنى لذلك الجمع.

ثانياً / مقاييس المتغيرات الكمية وتنقسم إلى ثلاثة أقسام هي:

1. مقياس ترتيبي Ordinal Scale:

يُعد هذا النوع من المقاييس تالياً من حيث المستوى للمقاييس الاسمية، فهو أعلى منها، لأنه إضافة إلى تصنيف الأفراد أو الأشياء في مجموعات متمايزة، إلا أنه يرتب الأشياء أو الأفراد تصاعدياً أو تنازلياً بناءً على صفة أو خاصية معينة، ومع ذلك لا بد أن يتأثر - كمقياس- ببداية العد أو الترقيم على عكس مقياس التصنيف الاسمي الذي لا يتأثر ببداية العد، وعندما تُعطى الأرقام للأشياء والأفراد وفقاً لهذا المقياس فإن تلك الأرقام لا تُمثل كميات معينة، كما أن المسافات الفاصلة بين رقم وآخر لا يُشترط أن تكون متساوية، فمثلاً المعلم في الفصل يُمكنه ترتيب تلاميذه من الأعلى إلى الأدنى، أي أن البيانات تكون الأول (95)، الثاني (92)، الثالث (85)، الرابع (70)، ....الخ.

فليس شرطاً أن يكون الفرق بين الدرجات متساوياً ، بمعنى أنه ليس شرطاً أن يكون فرق الدرجات بين الطالبين الأول والثاني يساوي الفرق بين درجات الثاني والثالث ، فالمقياس الرُتبي لا يعطي صورة واضحة عن حجم الفروق الموجودة بين الأفراد المتجاورين بين أية مجموعة.

وعلى سبيل المثال أيضاً إذا أردنا ترتيب مجموعة الأبناء في إحدى الأسر حسب العمر فقد نحصل على البيانات الآتية :

الأفراد	العمر بالسنوات	الرتبة
أ	25	1
ب	23	2
ج	18	3
د	11	4
هـ	7	5

فإذا نظرنا إلى هذا المقياس وجدنا أن الفرد (أ) يحتل المرتبة الأولى ، ولا بد أن نبدأ المقياس من هذه النقطة ، أي من عند (أ) يليه (ب) ، ثم (ج) وهكذا ، أو بالعكس ، ولا يُمكن مثلاً أن نبدأ من عند الفرد (ج) أو (د) ، كما نلاحظ شيئاً آخر وهو أن عمر الفرد الأول 25 سنة والثاني 23 سنة أي أن الفرق بينهما سنتين ، في حين أن الفرق بين الثاني والثالث 5 سنوات ، والثالث والرابع 7 سنوات ، والرابع والخامس 4 سنوات ، أو بمعنى آخر أن المسافات بين الوحدات غير متساوية ، على الرغم من أن هذا التساوي يظهر في الرتب ، حيث نجد أن تنظيم هذه الرتب هو 1 ، 2 ، 3 ، 4 ، 5 ، ويُعد هذا مأخذاً على المقياس الرُتبي ، وهذا النوع من المقاييس كثيرة الاستخدام في ميدان العلوم الإنسانية والاجتماعية.

ومن أشهر الأمثلة على القياس الرُتبي ما يسمى بمقياس ليكرت للاتجاهات ، حيثُ يُطلب من أفراد الدراسة أن يعبروا عن درجة موافقتهم أو معارضتهم إزاء قضية ما ، وذلك بأن يختاروا إحدى هذه الاستجابات : موافق جداً ، موافق ، محايد ، معارض ، معارض جداً ، وكمثال آخر تقسيم أفراد الدراسة حسب المستوى الاقتصادي إلى : عالي ، متوسط ، منخفض ، ومثال ثالث مستوى الأداء في العمل (ممتاز ، جيد ، متوسط ، ضعيف) ، ويلاحظ في كل هذه الأمثلة إمكانية إعطاء أرقام للفئات تدريجياً من الأصغر إلى الأكبر والعكس ، ويكون لهذه الأرقام معنىً يتضمن الأفضلية (أي معنىً ترتيبياً) إلا أن الفروق والمسافات بين هذه الأفضليات لا يُمكن تحديدها ، ولا يمكن الزعم بأنها متساوية (أي لا توجد وحدة قياس بهذا المستوى من وحدات القياس).

س / لماذا لا نستطيع استخدام العمليات الحسابية في المقياس الترتيبي؟

ج / لأن المسافات بين الوحدات غير متساوية.

مثال على ذلك:

- ترتيب الطلبة الأوائل (الأول 95، الثاني 92 ، الثالث 82).
- ترتيب الأخوة من حيث العمر (الأول 15 سنة ، الثاني 11سنة ، الثالث 9 سنوات).

ملحوظة / يدور جدل حول تصنيف المقاييس على المستوى الرتبي تحت المقاييس الكمية، إذ يعتقد البعض أنها مقاييس نوعية والحقيقة أن القياس الرتبي كمي لأن علاقة الأكبر والأصغر صالحة فيه.

2. مقياس فئوي (فتري أو فاصلي) Interval Scale

يُعد هذا النوع من المقاييس أعلى من المقياسين السابقين ، ويقترب كثيراً إلى المعنى الكمي للقياس ، حيث تحمل الأرقام هنا معنىً كمياً ، وبالتالي يكون الحصول على وحدة القياس مُتاحاً ، وبعد الاطلاع على المثال التالي سوف نتيقن من أن لهذا المقياس وحدة قياس بالإضافة إلى سمتي التصنيف والترتيب اللتين يتمتع بهما المقياس الترتيبي ، كما أن نقطة الإسناد هنا وهي (الصفر) محض افتراض أي أن **الصفر افتراضي** (غير حقيقي) بمعنى أن الصفر في هذا المقياس لا يعني انعدام الخاصية ، وإنما هو صفر نسبي وليس مطلقاً.

فمثلاً بافتراض أن درجات الطلاب في مادة الإدارة الصحية تتوزع بين الصفر (صفر الكلية أو الجامعة وهي الدرجة 35) والمائة بوحدة الخمس درجات أي (35 ، 40 ، 45 ، 50 ، ... ، 100) فإنه يلاحظ ما يلي:

- الطلاب في هذه المادة مختلفون في تحصيلهم وهذا يمكن قياسه بالمقياس الإسمي.
- رتبة الطالب الذي درجته (90) أعلى من رتبة الطالب الذي درجته (85) وهذا يمكن قياسه بالمقياس الترتيبي.
- الطالب الذي درجته (90) أعلى من الطالب الذي درجته (85) بخمس درجات (وحدة قياس واحدة) وأعلى من الطالب الذي درجته (70) بعشرين درجة (أربع وحدات قياس) وهذا ما يوفره المقياس الفئوي.

ومن الأمثلة كذلك على هذا المقياس درجة الحرارة ، فدرجة الحرارة صفر درجة مئوية لا تعني انعدام درجة الحرارة وإنما تعني درجة تجمد الماء النقي عند مستوى سطح البحر، وتعادل درجة الصفر المئوي 32 درجة فهرنهايت و (273) بمقياس كلفن .

ومن الأمثلة على التدريج الفئوي كذلك مقياس الضغط الجوي (البارومتر) وقياس معامل الذكاء، والدرجات التي يحصل عليها الطلبة في الاختبارات التحصيلية، وأرقام السنوات على اعتبار أن السنة الأولى هي بداية التقويم الهجري لا تمثل صفراً حقيقياً بل هي نقطة مرجعية (الإسناد) .

3. مقياس نسبي (Ratio Scale)

وهو أرقى مستويات القياس الثلاثة السابقة جميعها حيث يتميز على مستوى القياس الفتري في أمرين هما :

1. أنه يمتلك خاصية الصفر المطلق (الحقيقي) الذي يدل على انعدام الخاصية أو السمة .
2. النسبة بين أي درجتين في القياس لا تعتمد على الوحدات المستخدمة.

كما تفوق القياس النسبي على المستويات السابقة باشتماله على جميع سماتها بالإضافة إلى سماته الخاصة به ، ويندر استخدام هذا النوع من المقاييس في قياس الظواهر السلوكية عامة، والنفسية خاصة ، حيث يصعب افتراض انعدام الظاهرة نهائياً ، بينما يُستخدم هذا المستوى من القياس في العلوم الطبيعية مثل: الدخل والوزن والطول والعمر وعدد الأبناء بالأسرة ..... إلى غير ذلك من المتغيرات التي يمكن التسليم بانعدام وجودها عند نقطة معينة .

ويمكن في هذا المقياس استخدام جميع العمليات الحسابية (جمع وطرح وضرب وقسمة)، ويمكننا في هذا المقياس المقارنة بين نسب القيم ، حيث يمكن في المتغيرات النسبية قسمة قيمها الرقمية بحيث تعطي نسباً ذات معنى . ومن الأمثلة على القياس النسبي ، الطول والمساحة والحجم والوزن والقوة .

فعلى سبيل المثال :

إذا قلنا أن طول النقطة = صفر فهذا يعني أن طولها معدوم .

وإذا قلنا : طول مستقيم = 30 سم وطول مستقيم آخر = 10 سم فهذا يعني أن طول القطعة الأولى ثلاثة أمثال طول القطعة الثانية.

وكمثال آخر : عندما نقول أن أسرة ما لديها صفر من الأبناء فهذا يعني أنها لا تمتلك أبناء أبداً ، وأن الأسرة التي لديها ستة أبناء يمثلون ضعف عدد الأبناء الذين لدى أسرة بها ثلاثة أبناء فقط .

تدريب (8)

حدد نوع المتغير والمقياس الإحصائي المناسب لما يلي:

الحالة الاجتماعية ، الوزن ، درجة الحرارة ، مستوى الدخل.

تمارين الفصل الأول

س1 / ضع دائرة حول رمز الإجابة الصحيحة فيما يلي:

1) الإحصاء الوصفي هو الذي يختص بـ :

أ. طُرق جمع البيانات وعرضها ووصفها . ب. اتخاذ القرار في ظل عدم التأكد .

ج. تحليل البيانات والوصول لاستنتاجات واتخاذ قرار. د. ليس ما سبق .

2) نوع المتغير " عدد الأطفال في أسرة " هو :

أ. متصل فئوي. ب. متصل ترتيبي. ج. منفصل نسبي . د. خلاف ذلك وهو:

3) مستوى دخل الفرد هو متغير:

أ. كمي فئوي. ب. نوعي اسمي. ج. كمي نسبي . د. خلاف ذلك وهو:

4) لا يدل الصفر على انعدام الخاصية في المقياس :

أ. الاسمي. ب. الترتيبي. ج. الفئوي. د. النسبي.

5) من وظائف علم الإحصاء :

أ. الوصف . ب. التحليل. ج. التنبؤ . د. جميع ما سبق .

س2 / عرف علم الإحصاء لغةً واصطلاحاً ، موضحاً أهميته .

س3 / ما هي مجالات علم الإحصاء ؟

س4 / ميز بين كل من الإحصاء الوصفي والإحصاء الاستدلالي.

س5 / ما هي أدوات جمع البيانات الأولية ؟

س6 / اذكر أنواع المتغيرات ، مع التوضيح بمثال لكل منها .

س7 / ما الفرق بين المتغير والثابت ؟

س8 / في أي مقياس من المقاييس يدل الصفر على انعدام الصفة أو الخاصية ؟

س9 / ما هي مصادر جمع البيانات ؟

س10 / ما هي مقاييس المتغيرات ؟ مع التوضيح بمثال لكل مقياس ؟

س11 / صنف البيانات التالية من حيث كونها (نوعية أو كمية) :

أ. لون العيون لأفراد عائلة عددهم (7) .

ب. عدد الأمراض التي أُصيب بها طفل خلال عامه الأول .

ج. شدة النزيف الدموي لمريض (بسيط، معتدل، حاد) .

د. فصيلة الدم عند الإنسان .

هـ. مستوى جودة دواء معين .

و. عدد أجهزة التعقيم في مختبر طبي.

ز. مستوى الدخل الشهري لطبيب.

ح. عدد العمليات الجراحية التي أجريت في السابق على مريض.

س12 / صنف المتغيرات حسب نوعها (منفصل أو متصل) :

أ. عدد المستشفيات في الجمهورية اليمنية.

ب. وزن طفل عند الولادة.

ج. مدة مكالمة هاتفية.

د. عدد المرضى المترددين على مستشفى خلال شهر.

هـ. ضغط الدم لدى مريض.

و. العدد اليومي للعمليات الجراحية التي تُجرى في أحد المستشفيات.

ز. عدد سيارات الإسعاف التي يمتلكها مستشفى الثورة.

ح. طول مريض بالسنتيمتر.

ك. عدد أيام غياب طبيب عن عمله خلال شهر.

2

الفصل الثاني

طرق عرض البيانات الإحصائية

محتويات الفصل الثاني

أهداف الفصل الثاني

يتوقع من الطالب بعد الانتهاء من دراسة هذا الفصل أن يكون قادراً على:

- ✓ تحديد طُرق عرض البيانات الإحصائية.
- ✓ عرض البيانات الإحصائية جدولياً.
- ✓ عرض البيانات الإحصائية المختلفة بيانياً .
- ✓ إنشاء جدول توزيع تكراري بفئات وبدون فئات .
- ✓ تكوين جدول التوزيع التكراري النسبي .
- ✓ تكوين جدول التوزيع التكراري المتجمع الصاعد .
- ✓ تكوين جدول التوزيع التكراري المتجمع الهابط .
- ✓ تمثيل منحنى التكرار المتجمع الصاعد والهابط بيانياً.
- ✓ تعريف المضلع والمنحنى والمدرج التكراري.
- ✓ تمثيل المضلع والمنحنى والمدرج التكراري بيانياً.
- ✓ التمييز بين المضلع التكراري والخط المتكسر.
- ✓ التمييز بين المدرج التكراري والأعمدة البيانية.

مقدمة:

هناك طُرق مختلفة تُستخدم عادةً لعرض البيانات الإحصائية بشكل مُنظم ومُركز وواضح والهدف الرئيس من استخدام هذه الوسائل هو تسهيل مهمة الباحث أو الإحصائي ومساعدته على إبراز ما يهدف إليه من البيانات وإعطاء صورة واضحة عن المشكلة قيد الدراسة أو البحث.

وفيما يلي استعراض لأهم هذه الطرق وأكثرها شيوعاً:

طُرق عرض البيانات الإحصائية

الجداول | الرسوم البيانية

أولاً / العرض الجدولي (Tables).

يُعد العرض الجدولي مُهم لترتيب البيانات وتلخيصها باستخدام الجداول التكرارية، ولكي يتم التعامل مع البيانات إحصائياً تمهيداً لاستخدام المقاييس الإحصائية يجب أولاً أن تبوب هذه البيانات حتى يتم دراستها .

وتصنف الجداول إلى قسمين هما:

يمكن تصنيف الجداول إلى

الجداول اليدوية | الجداول الآلية

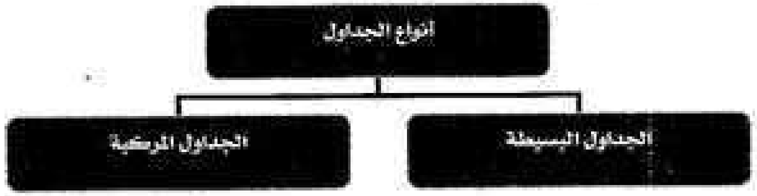
أ) الجداول اليدوية:

ويتم عمل الجداول يدوياً عندما يكون عدد الملاحظات (البيانات) معقولاً بحيث لا تزيد عن مئة قراءة، وعند عدم توفر الحاسبات، وأبسط هذه الجداول هو الجدول التكراري البسيط.

ب) الجداول الآلية:

ويتم عمل الجداول باستخدام الحاسوب (الكمبيوتر) وذلك عندما يكون عدد الملاحظات (البيانات) كبير بحيث يسهل حذف أو إضافة أو تعديل أي صف أو عمود.

ويوجد تصنيف آخر للجداول الإحصائية هو:



أ) الجداول البسيطة: وهي الجداول التي يكون لكل عمود فيها عنوان واحد فقط، ومن أمثلتها:

مثال (1): إذا كان عدد المرضى المترددين على بعض المستشفيات والمراكز الصحية في أمانة العاصمة لأحد الأيام كما في الجدول الآتي:

جدول (1)

أعداد المرضى المترددين على بعض المستشفيات والمراكز الصحية بأمانة العاصمة في أحد الأيام:

م	اسم المستشفى أو المركز الصحي	عدد المترددين (الزوار)
1	مستشفى الثورة العام	800
2	المستشفى الجمهوري	700
3	مستشفى السبعين	500
4	مركز الزهراوي الطبي	300
6	مستشفى جامعة العلوم والتكنولوجيا	650
7	مستشفى آزال التخصصي	250
8	مستشفى الأمانة التخصصي	150
9	مستشفى دار الرحمة التخصصي	100
10	المستشفى اليمني الألماني	320
11	المستشفى السعودي الألماني	600
	المجموع	4370

نلاحظ أن هذا الجدول من نوع الجداول البسيطة التي تعرض تغير ظاهرة مع مسميات، حيث أن عدد الزوار (المترددين) تمثل الظاهرة، والمسميات هي أسماء المستشفيات.

ب) الجداول المركبة: وهي الجداول التي يندرج تحت بعض عناوين أعمدتها عناوين فرعية.

مثال ذلك :

- الجداول الموجودة في سجلات التعداد العام للمساكن والسكان.
- الجداول الموجودة في سجلات التطعيم (التلقيح) للأطفال .

تدريب (1)

قم بإنشاء جدول إحصائي بسيط، وجدول مركب .

ويجب عند استعمال الطريقة الجدولية مراعاة ذكر ما يأتي:

- عنوان الجدول.
- الوحدات المستعملة.
- ذكر المصادر التي أُخذت منها البيانات.
- مذكرات تفسيرية لتفسر سبب شذوذ بعض البيانات إن وجدت.

ثانياً / الأشكال البيانية "طُرق العرض البياني" Graphical Diagrams

الأشكال والرسوم البيانية عامل مُهم لتسهيل عملية التحليل الإحصائي وجلب انتباه القارئ العادي الذي يهتم بالنظر إلى الأشكال والرسوم ولا يعير أي اهتمام إلى البيانات المعروضة بشكل جداول وأرقام مرصوصة على صفحات المطبوعات.

وتُعد الرسوم والأشكال البيانية من الطرق الفعالة في توضيح أي نقطة أو موضوع قد لا يُمكن توضيحه بمجرد النظر إلى الجداول الإحصائية مهما كانت مبسطة. فقد يفهمها الإحصائي لطبيعة تخصصه ولكن الرسوم والأشكال المنظمة والمبسطة تكون غالباً ممكنة الفهم بسهولة.

وللأشكال البيانية أنواع كثيرة نُشير إلى أهمها فيما يلي:

1. طريقة المستطيلات او الأعمدة البيانية (Bar Graph).

تتلخص هذه الطريقة بوضع المسميات على خط أفقي ورسم مستطيل على كل مُسمى يكون طول ارتفاعه مُمثلاً للقيمة المقابلة لذلك المسمى وذلك باستعمال مقياس رسم مناسب. ونستعمل هذه الطريقة لعرض تغير ظاهرة مع الزمن او مسميات اخرى.

مثال(2)

إذا كان عدد الطلبة في إحدى الكليات الطبية هو 560 طالباً وطالبة موزعين على الأقسام كما في الجدول الآتي:

القسم	عدد الطلاب		المجموع
	ذكور	إناث	
صيدلة	170	30	200
مختبرات	95	25	120
مساعدين	80	30	110
تمريض	50	30	80
اسنان	35	15	50
المجموع	430	130	560

المطلوب/

إيضاح هذه البيانات بيانياً بطريقة الأعمدة البيانية بطرق مختلفة بمقياس صحيح.

الحل:

أ. عرض البيانات حسب التخصص بغض النظر عن متغير الجنس .

ب. عرض البيانات بطريقة الأعمدة المتجاورة مع مراعاة متغير الجنس.

ج- عرض البيانات السابقة بطريقة الأعمدة المركبة مع مراعاة متغير الجنس.

2. طريقة الخط المتكسر (Broken Line Graph):

تستعمل هذه الطريقة لعرض البيانات الناتجة من تغير ظاهرة أو عدة ظواهر مع مسميات أو مع الزمن أو كليهما ، مثل: تغيُّر درجة حرارة مريض مع الزمن بالساعات. أو تغيُّر أعداد الطلبة في جامعة مع السنوات . أو تغيُّر أعداد الطلبة مع التخصصات على مدى فترة زمنية محددة.

مثال (3)

الجدول الآتي يبين عدد الصيدليات والمخازن في مدينة عدن خلال الفترة من 1997- 2003م :

السنة	1997	1998	1999	2000	2001	2002	2003
صيدليات	62	65	68	74	78	84	81

المطلوب/ عرض البيانات بطريقة الخط المتكسر.

الحل

3. طريقة الخط المنحني: Curve

وهذه الطريقة تماثل طريقة الخط المتكسر ونحصل عليها بتمهيد الخط المتكسر ليصبح على شكل منحنى بدون زوايا .

مثال (4)

اعرض البيانات في المثال (3) أعلاه بطريقة الخط المنحني.

الحل

طريقة الخط المنحني

4. طريقة الدائرة (Pie Chart).

تُعد الرسوم الدائرية من أشكال العرض البياني ، وفيها تعبر عن الظواهر بمساحات دائرية.

وتُستخدم عندما لا يكون الهدف متابعة تطور التغيرات التي تطرأ على ظاهرة معينة بل لإبراز الأجزاء التي تتكون منها تلك الظاهرة .

وتُمثل البيانات بهذه الطريقة بتقسيم الكل إلى أجزاء، فيُمثل المجموع الكلي بدائرة كاملة ويُمثل كل جزء بقطاع دائري يكون قياس زاويته مساوياً 360° مضروباً في نسبة الجزء للمجموع الكلي، اي ان :

مساحة الجزء المعني = (قيمة الجزء ÷ مجموع قيم الأجزاء) × 360°

مثال (5) يمثل الجدول الآتي عدد الأطباء في أحد المستشفيات خلال السنوات 1995 - 1996 / 1998 - 1999م

العام	عدد الأطباء
1995 - 1996م	90
1996 - 1997م	105
1997 - 1998م	120
1998 - 1999م	135
المجموع	450

المطلوب: عرض هذه البيانات بطريقة الدائرة.

الحـــل/

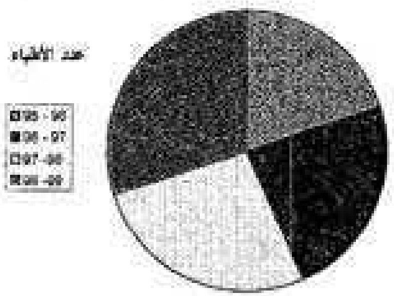
قياس زاوية قطاع (1995 - 1996م) هي $(90 \div 450) \times 360^\circ = 72^\circ$

قياس زاوية قطاع (1996 - 1997م) هي $(105 \div 450) \times 360^\circ = 84^\circ$

قياس زاوية قطاع (1997 - 1998م) هي $(120 \div 450) \times 360^\circ = 96^\circ$

قياس زاوية قطاع (1998 - 1999م) هي $(135 \div 450) \times 360^\circ = 108^\circ$

نرسم دائرة ونرسم القطاعات الأربعة التي تمثل السنوات حسب قياس زاوية كل قطاع كما يظهر في الشكل الآتي:



نشاط (1) :

البيانات الآتية تُمثل اعداد الطلبة المسجلين بإحدى الكليات الطبية :

المجموع	صحة مجتمع	أسنان	مختبرات	طب بشري	صيدلة	القسم
2500	400	600	300	500	700	عدد الطلبة

المطلوب :

عرض البيانات السابقة بطريقة:

أ. الأعمدة البيانية.   ب. الخط المنكسر.   ج. الخط المنحني.   د. الدائرة .

الحل:...

....

....

....

....

تدريب (2)

إذا كانت البيانات الآتية تمثل لون العيون لـ35 طالباً في إحدى التخصصات :

المجموع	اخرى	أزرق	بني	لون العيون
35	5	10	20	عدد الطلبة

المطلوب : عرض البيانات السابقة بطريقة

أ. الأعمدة البيانية .   ب. الخط المنكسر .

ج. الخط المنحني .   د. الدائرة .

التوزيع التكراري Frequency Distribution

لدى عرض كثير من الظواهر بالطرق السابقة لابد وأنك تلاحظ ان هذه الظواهر ربما تحتوي على أعداد كبيرة من البيانات ، ولذلك فعند عرضها بطريقة الجداول او المستطيلات مثلاً فإنك لا تستطيع المقارنة بين مفردات هذه البيانات ، ويصعب عليك فهمها ، لذلك لابد من تلخيصها وعرضها بطريقة مُبسطة تسهل عليك فهمها ، ومن هذه الطرق التوزيعات التكرارية.

فالتوزيعات التكرارية نتمكن بواسطتها من تنظيم البيانات الكثيرة بحيث لا تخسر هذه البيانات من اهميتها إلا الشيء اليسير، أو ربما لا تخسر شيئاً، والفكرة الأساسية لبناء التوزيع التكراري تتمثل في تقسيم مدى البيانات إلى فئات وحصر عدد البيانات الواقعة ضمن كل فئة.

♦ أولاً : التوزيع التكراري بدون فئات.

مثال(1)

توضح البيانات الآتية درجات 27 طالباً في اختبار قصير في مادة الإحصاء الطبي :

8	6	4	5	6	8	9	10	7
9	5	6	6	7	10	6	7	6
8	6	8	9	6	7	7	6	7

المطلوب / عرض هذه البيانات في جدول توزيع تكراري.

الحل

إذا كانت البيانات كمية نحسب أولاً المدى (Range)

مدى هذه البيانات = أعلى قيمة − اقل قيمة .

المدى $R = 10 - 4 = 6$

وحيث أن المدى صغير فيمكننا عرض البيانات في جدول توزيع تكراري بدون فئات كما يلي:

الدرجة	الفرز	التكرار F
4	/	1
5	//	2
6	//// ////	9
7	//// /	6
8	////	4
9	///	3
10	//	2
المجموع		27

ونلاحظ في بناء هذا الجدول أننا بدأنا من أصغر قيمة وهي 4 ثم رتبنا القيم ترتيباً تصاعدياً إلى أن وصلنا إلى أعلى قيمة 10 كما يظهر في العمود الأول. أما عناصر العمود الثاني فتمثل الفرز للبيانات، أما عناصر العمود الثالث فتمثل التكرار (Frequency).

يتضح لنا من الجدول أعلاه أن جدول التوزيع التكراري يتألف من :

1- فئات قيم المشاهدات أو القياسات.

2- التكرارات المقابلة لكل فئة أو قياس.

وهذا يعني أنه إذا اخترنا أي قيمة في البيانات فإننا سنتمكن من وضعها في فئة واحدة فقط.

وبهذا نتمكن من إفراغ جميع البيانات في فئات التوزيع التكراري وبدون التباس (غموض) وسيكون مجموع التكرارات مساوياً لعدد البيانات.

مثال (2)

البيانات الآتية تمثل فصيلة الدم ل(30) مريضاً :

AB	O	AB	B	AB	B	A	AB	B	A
O	B	A	B	A	B	A	B	A	B
A	B	B	B	A	B	A	B	AB	O

أ. ما نوع البيانات أعلاه ؟

ب. ما الفصيلة الأكثر تكراراً ؟ وماذا تمثل إحصائياً ؟

ج. كون جدول توزيع تكراري بسيط للبيانات السابقة.

د. اعرض البيانات السابقة بطريقة الأعمدة البيانية.

الحل

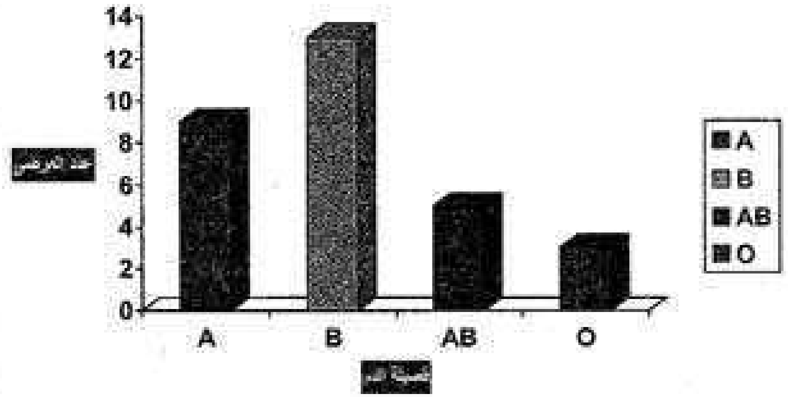
أ. نوع البيانات وصفية (نوعية).

ب. الفصيلة الأكثر تكرارا هي (B) وتمثل المنوال.

ج. جدول التوزيع التكراري .

فصيلة الدم X	التكرار F
A	9
B	13
AB	5
O	3
المجموع	30

د. عرض البيانات بطريق الأعمدة البيانية:



تدريب (3) :

إذا كانت البيانات الآتية تمثل فصيلة الدم لـ 50 شخصاً :

A	B	O	AB	O	B	AB	O	AB	O
B	B	A	B	A	A	O	A	O	A
A	A	B	A	A	A	AB	A	A	B
AB	A	O	A	O	A	B	A	A	A
O	AB	O	A	A	A	A	B	B	AB

المطلوب :

أ. ما نوع البيانات أعلاه.

ب. تكوين جدول توزيع تكراري بسيط للبيانات السابقة.

ج. ما هي الفصيلة الأكثر تكراراً ؟ وماذا تُمثل إحصائياً ؟

د. عرض البيانات السابقة بطريقة الأعمدة البيانية وطريقة الدائرة.

◆ ثانياً : التوزيع التكراري بفئات.

إذا كان مدى البيانات صغيراً أمكن بناء التوزيع التكراري مباشرة كما في المثال السابق ، أما إذا كان المدى كبيراً أو كان عدد البيانات كبيراً فإنه يجب في هذه الحالة أن تُقسم قيم البيانات إلى فئات يتراوح عددها في الغالب بين (5 - 15) فئة حسب كون عدد البيانات صغيراً أو كبيراً.

ويلاحظ أنه عند بناء جدول توزيع تكراري بفئات يجب مراعاة ما يلي:

- يجب أن تكون الفئات منفصلة عن بعضها البعض (غير متداخلة فيما بينها).
- أن تكون الفئات متساوية في الطول ما أمكن.
- يجب أن تكون الفئات كافية لاحتواء جميع البيانات.

أشكال التوزيع التكراري في فئات:

يمكن التمييز بين الأشكال الآتية للجداول التكرارية في فئات من حيث:

(أ) الانتظام في أطوال الفئات:

- جداول منتظمة حيث تكون الفئات فيها ذات أطوال متساوية.
- جداول غير منتظمة حيث لا تتساوى في أطوال الفئات.

(ب) تعيين حدود الفئتين الأولى والأخيرة:

- جدول تكراري مفتوح من أعلى حيث لا يعرف فيه الحد الأدنى للفئة الأولى.
- جدول تكراري مفتوح من أسفل حيث لا يعرف فيه الحد الأعلى للفئة الأخيرة.
- جدول تكراري مفتوح الطرفين حيث الحد الأدنى للفئة الأولى والحد الأعلى للفئة الأخيرة غير معلومين.

أشكال الفئات:

(أ) الصورة العامة للفئات:

مثل: 70 إلى اقل من 80 وتكتب 70 –

80 إلى اقل من 90 وتكتب 80 -

(ب) الفئات المحددة البداية والنهاية:

مثل: 10 - 6

11- 15

وفي هذا الشكل يحدد الحد الأدنى والأعلى لكل فئة، ويستخدم في حالة المتغيرات المتقطعة.

مثال(3) تمثل البيانات الآتية درجات 66 طالباً في الامتحان النهائي لمقرر الإدارة الصحية.

19	30	51	76	55	37	11	45	52	36	13
26	33	65	72	54	36	14	48	51	39	17
26	34	41	71	70	38	18	50	66	42	15
29	38	70	74	8	40	10	52	61	43	21
29	40	9	75	68	44	12	63	62	53	22
69	53	68	59	69	61	41	67	59	9	75

المطلوب / كون جدول توزيع تكراري للبيانات أعلاه.

ولعرض هذه البيانات في جدول توزيع تكراري نتبع الخطوات الآتية:

1. نحسب المدى للبيانات أعلاه من العلاقة ، المدى (Range) = أكبر قيمة – أقل قيمة.

للبيانات التي لدينا نجد أن المدى (R) $R = 76 - 8 = 68$

2. نحسب عدد الفئات (Classes) من العلاقة: $M = 1 + 3.31 \text{ Log } (n)$

حيث M ترمز لعدد الفئات، n ترمز لعدد القيم (المشاهدات)

$M = 1 + 3.31 \text{ Log}(66) = 7$

3. نوجد طول الفئة Length Of Class Interval (C) من العلاقة:

طول الفئة C = (المدى R ÷ عدد الفئات M) ثم نقرب الجواب دائماً إلى أعلى بحيث تكون دقة طول الفئة تساوي أو تقل عن دقة الأرقام المستعملة في البيانات.

$C = R \div M = 68 \div 7 = 9.7$

نقرب العدد 9.7 إلى أعلى فيصبح طول الفئة C = 10 .

4. نعين الحد الأدنى لأول فئة ويجب أن يكون هذا الحد مساوياً لأصغر قيمة في البيانات أو أقل منها بقليل، ويجب أن تكون درجة دقته نفس درجة دقة البيانات المستعملة.

إذاً الحد الأدنى للفئة الأولى = 8 ، ولحساب الحد الأدنى الحقيقي للفئة الأولى نطرح 0.5 من الحد الأدنى فيصبح (7.5) .

5. نعين الحد الأعلى للفئة الأولى وذلك بإضافة (طول الفئة مطروحاً منه واحد) إلى الحد الأدنى لتلك الفئة. إذاً الحد الأعلى للفئة الأولى هو 8 + 9 = 17 ، ولحساب الحد الأعلى الحقيقي للفئة الأولى نضيف 0.5 للحد الأعلى فيصبح 17.5 . وبهذا نكون قد حصلنا على الفئة الأولى وهي الفئة التي حدودها 17 – 8 وحدودها الحقيقية 17.5 – 7.5 .

6. نعين الحدود الدنيا والعليا لبقية الفئات بنفس الطريقة السابقة، ثم نعين الحدود الدنيا الحقيقية والعليا الحقيقية كما سبق.

7. نفرغ البيانات المعطاة لدينا على الفئات التي انشأناها وذلك باستعمال خط عمودي لكل قراءة وخط مائل للقراءة الخامسة في كل فئة وذلك لتسهيل جمع التكرارات في عمود مخصص في الجدول يُسمى عمود الفرز.

8. نجمع التكرارات المقابلة لكل فئة ونسجله في عمود التكرارات fi ثم نجمع التكرارت لجميع الفئات ونقارنه بعدد البيانات فإذا كان عدد البيانات n فإنه يجب أن يكون مساوياً لمجموع التكرارات أي أن :

$\sum_{i=1}^{h} f_i = n$ حيث عدد الفئات هو h. وبهذه الخطوة نكون قد انتهينا من تكوين جدول التوزيع التكراري.

9. إذا أردنا إيجاد مراكز الفئات باعتبار مركز الفئة يُمثل الفئة، فيمكن حساب ذلك من العلاقة: مركز الفئة = [(الحد الأدنى للفئة + الحد الأعلى للفئة) ÷ 2]، ولذلك فمركز الفئة الأولى هو $12.5 = [2 \div (8+17)]$. ويمكن إيجاد مركز أي من الفئات الأخرى بنفس الطريقة، أو بإضافة طول الفئة إلى مركز الفئة التي قبلها (إذا كان جدول التوزيع التكراري منتظم) فيكون مركز الفئة الثانية $12.5+10=22.5$ ومركز الفئة الثالثة $22.5+10=32.5$ وهكذا. والجدول الآتي يوضح ذلك.

Class Interval	Frequency			Class Point
حدود الفئة	(الفرز)	التكرار f	الحدود الحقيقية للفئة	مركز الفئة x
8 – 17	//// ////	10	7.5 – 17.5	12.5
18 – 27	//// /	6	17.5 – 27.5	22.5
28 – 37	//// ///	8	27.5 – 37.5	32.5
38 – 47	/ //// ////	11	37.5 – 47.5	42.5
48 – 57	//// ////	10	47.5 – 57.5	52.5
58 – 67	//// ////	9	57.5 – 67.5	62.5
68 – 77	// //// ////	12	67.5 – 77.5	72.5
المجموع				

جدول التوزيع التكراري لدرجات 66 طالباً في مقرر الإدارة الصحية

ملحوظة / عند عرض التوزيع التكراري لا يُكتب عمود إفراغ البيانات (الفرز) وفي بعض الأحيان لا يُكتب عمود الحدود الحقيقية .

◆ التوزيع التكراري النسبي Relative Frequency :

إن التكرار النسبي Relative Frequency لكل فئة هو نسبة تكرار تلك الفئة إلى مجموع التكرارات.
فإذا كان مجموع التكرارات n ، وكان تكرار فئة معينة f فإن تكرارها النسبي p يعطى من العلاقة:
التكرار النسبي للفئة = تكرار الفئة ÷ مجموع التكرارات   ورمزياً يُحسب من العلاقة $(p=f\div n)$
إن الجدول الذي يعطينا الفئات (أو مراكزها) مع تكراراتها النسبية يسمى التوزيع التكراري النسبي.
وعليه فإن الجدول الآتي يُمثل التوزيع التكراري النسبي للمثال السابق:

الدرجة	التكرار F	التكرار النسبي = التكرار ÷ مجموع التكرارات
8 – 17	10	0.15
18 – 27	6	0.09
28 – 37	8	0.12
38 – 47	11	0.17
48 – 57	10	0.15
58 – 67	9	0.14
68 – 77	12	0.18
المجموع	66	1

جدول التوزيع التكراري النسبي لدرجات 66 طالباً في الإدارة الصحية

ملحوظة / دائماً مجموع التكرارات النسبية = 1

نشاط (2) :

البيانات الآتية تمثل الأجر الشهري ل (30) موظفاً في القطاع الصحي بآلاف الريالات في أحد المستشفيات.

52	46	50	45	39	45	43	40	48	47
43	50	33	52	32	38	35	52	47	42
60	60	43	39	31	37	50	49	51	55

كون جدول توزيع تكراري للأجور ، ثم كون جدول التوزيع التكراري النسبي.

الحل:..

تدريب (3)

إذا كانت البيانات الآتية تمثل أعمار مجموعة من المرضى بأحد المستشفيات لأقرب سنة، المطلوب: كون جدول التكرار النسبي لهذه البيانات ثم مثله بيانياً.

العمر بالسنة	التكرار F	التكرار النسبي
25	7	
30	5	
35	9	
40	12	
45	11	
50	8	
المجموع	52	1

◆ **التكرار المتجمع Cumulative Frequency:**

نحتاج بعض الأحيان لمعرفة عدد المشاهدات (البيانات) التي تساوي قيمة معينة او تكون أصغر منها، أو أكبر منها. ولذلك سندرس نوعين من التوزيع التكراري المتجمع للتوضيح وهما:

أ. التوزيع التكراري المتجمع الصاعد.

ب. التوزيع التكراري المتجمع النازل (الهابط).

مثال(4)

يوضح الجدول التكراري الآتي أعمار مجموعة من المرضى لأقرب سنة:

العمر بالسنة	24	25	26	27	28	29	30	المجموع
التكرار	2	8	9	15	12	10	4	60

المطلوب / تكوين جدول التكرار المتجمع الصاعد والنازل (الهابط) للبيانات في الجدول.

الحـــل

أولاً: لتكوين الجدول التكراري المتجمع الصاعد نفتح سطر ثالث في الجدول السابق وذلك للتكرار المتجمع الصاعد كما يلي:

أول تكرار متجمع صاعد = أول تكرار في عمود التكرار =2 ، ثاني تكرار = 2 + 8 = 10 ، ثالث تكرار = 10 + 9 = 19 ، رابع تكرار = 19 + 15 = 34 ..... وهكذا.

فكل تكرار متجمع صاعد هو مجموع ذلك التكرار مع كل التكرارات السابقة.

فيكون جدول التكرار المتجمع الصاعد كما يلي:

العمر بالسنة	24	25	26	27	28	29	30	المجموع
التكرار	2	8	9	15	12	10	4	60
التكرار المتجمع الصاعد	2	10	19	34	46	56	60	

ملحوظة:

نلاحظ ان آخر تكرار متجمع صاعد = مجموع التكرارات، وأول تكرار متجمع صاعد = أول تكرار.

ثانياً: لتكوين جدول التكرار المتجمع النازل (الهابط) نفتح سطراً آخر للتكرار المتجمع النازل. نعلم ان مجموع التكرارات في الجدول السابق هو 60 ، ستكون هذه القيمة هي أول قيمة في التكرار المتجمع النازل (الهابط)، إذاً أول تكرار هابط = 60 ، ثاني تكرار متجمع هابط = أول تكرار هابط مطروحاً من التكرار المناظر له = 60 - 2 = 58، ثالث تكرار هابط = 58 - 8 = 50، ... وهكذا، فكل تكرار متجمع هابط هو مجموع التكرارات كلها ناقصاً مجموع التكرارات السابقة.

فيصبح جدول التكرار المتجمع النازل (الهابط) كما يلي:

العمر بالسنة	24	25	26	27	28	29	30	المجموع
التكرار	2	8	9	15	12	10	4	60
التكرار المتجمع النازل (الهابط)	60	58	50	41	26	14	4	

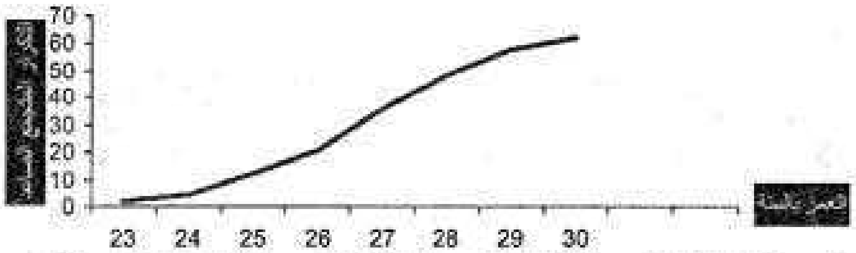
ملحوظة :

نلاحظ أن أول تكرار متجمع هابط = مجموع التكرارات، وآخر تكرار متجمع هابط = آخر تكرار.

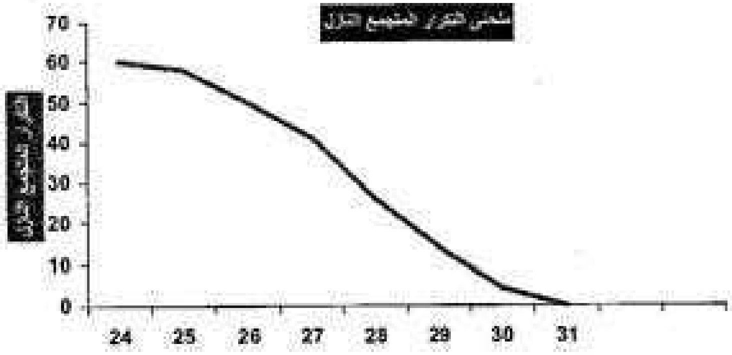
◆ **التمثيل البياني لجدولي التكرار المتجمع الصاعد والنازل أعلاه:**

لرسم منحنى التكرار المتجمع الصاعد نستخدم جدول التكرار المتجمع الصاعد بحيث يمثل المحور الأفقي الدرجات والمحور الرأسي التكرار المتجمع الصاعد، كما في الشكل الآتي :

منحنى التكرار المتجمع الصاعد



لرسم منحنى التكرار المتجمع النازل نستخدم جدول التكرار المتجمع النازل ، ونجعل المحور الأفقي يمثل الدرجات والمحور الرأسي يمثل التكرار المتجمع النازل، كما في الشكل:



ملحوظة:

يُمكن تمثيل كل جدول تكراري (صاعد او نازل) برسم بياني منفرد او ندمجهما في رسم واحد.

كما أنه يُمكن عمل جدول واحد يحتوي التكرارين المتجمع الصاعد والنازل معاً.

وعندما يكون لدينا جدول تكراري بفئات نعتبر مركز الفئة ممثلاً للمحور الأفقي والتكرارين المتجمع الصاعد والنازل ممثلين للمحور الرأسي.

تدريب (3)

البيانات الآتية تمثل أعمار مجموعة من مرضى فقر الدم :

47	65	21	76	83	60	10	70
70	38	68	82	53	45	64	49
61	68	79	56	47	23	54	32
61	38	48	79	75	77	7	25
48	41	57	83	41	55	59	15

المطلوب :

- إيجاد عدد ثم نسبة المرضى الذين أعمارهم 60 سنة فأكثر.
- إيجاد عدد ثم نسبة المرضى الذين أعمارهم أقل من 30 سنة .
- تكوين جدول تكراري بفئات للبيانات السابقة حيث طول الفئة 10 .
- تكوين جدول التكرار المتجمع الصاعد والهابط والنسبي للبيانات السابقة، و تمثيلهم بيانياً.

♦ تمثيل التوزيعات التكرارية بيانياً:

هناك ثلاث طرق لتمثيل التوزيعات التكرارية بيانياً وهي:

1. المدرج التكراري Frequency Histogram.
2. المضلع التكراري Frequency Polygon.
3. المنحنى التكراري Frequency Curve.

أولاً: المدرج التكراري (Frequency Histogram) .

وهو عبارة عن تمثيل تكرار كل فئة من فئات التوزيع التكراري بمستطيل بحيث يتناسب تكرار كل فئة مع مساحة المستطيل المقام على تلك الفئة .

أي يكون ارتفاع المستطيل = تكرار الفئة ÷ طول الفئة.

وإذا كان التوزيع التكراري ذا فئات متساوية فإننا نمثل كل فئة بمستطيل حدود قاعدته الحدود الفعلية لتلك الفئة ، وطول ارتفاعه يتناسب مع تكرارها . أي أننا نأخذ محورين متعامدين ، نرصد على المحور الأفقي الحدود الفعلية (الحقيقية) لكل فئة من فئات التوزيع التكراري، ونقيم على كل فئة مستطيلاً يتناسب طول ارتفاعه مع تكرار تلك الفئة.

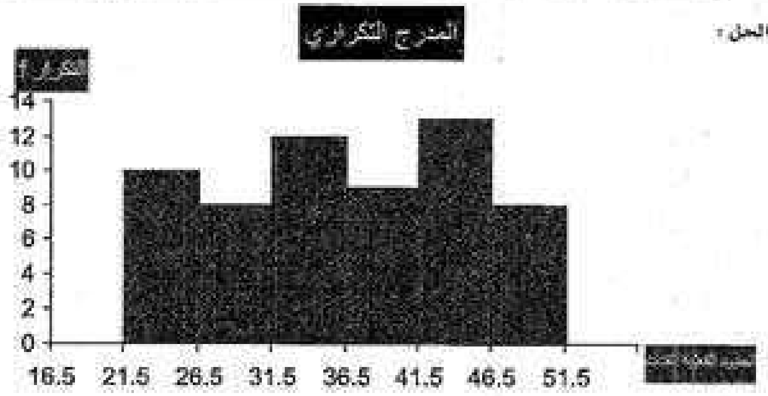
مثال/

إذا كان لديك جدول التوزيع التكراري الآتي :

التكرار f	مركز الفئة x	الحدود الحقيقية للفئة	حدود الفئة
10	24	21.5 - 26.5	22 - 26
8	29	26.5 – 31.5	27 - 31
12	34	31.5 – 36.5	32 - 36
9	39	36.5 – 41.5	37 - 41
13	44	41.5 – 46.5	42 - 46
8	49	46.5 – 51.5	47 - 51
60			المجموع

المطلوب / ارسم المدرج التكراري للبيانات الواردة في الجدول.

الحل :

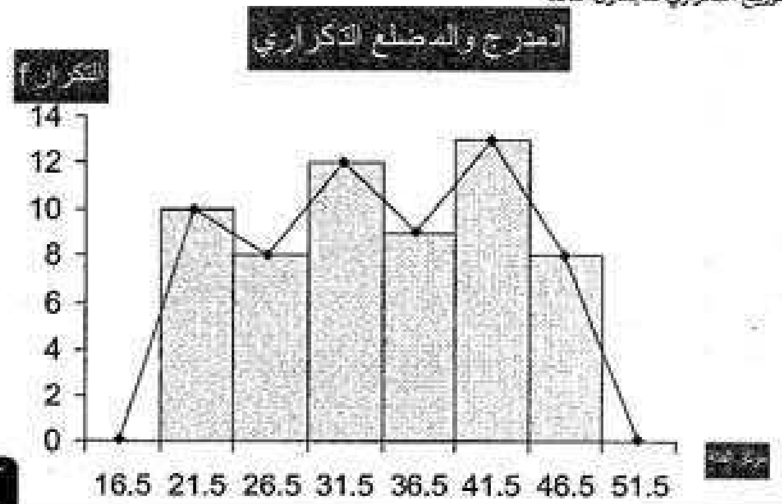


ثانياً: المضلع التكراري (Frequency Polygon).

المضلع التكراري هو مضلع مغلق نحصل عليه بتنصيف الأضلاع العلوية للمستطيلات في المدرج التكراري، ثم يوصل هذه النقاط بعضها مع بعض.

ولكي نغلق الخط المنكسر الذي حصلنا عليه نفترض أن هناك فئتين متطرفتين واحدة إلى أقصى اليسار والثانية إلى أقصى اليمين، وتكرار كل منهما صفر أي أن ارتفاع كل من المستطيلين المقامين على هاتين الفئتين صفر.

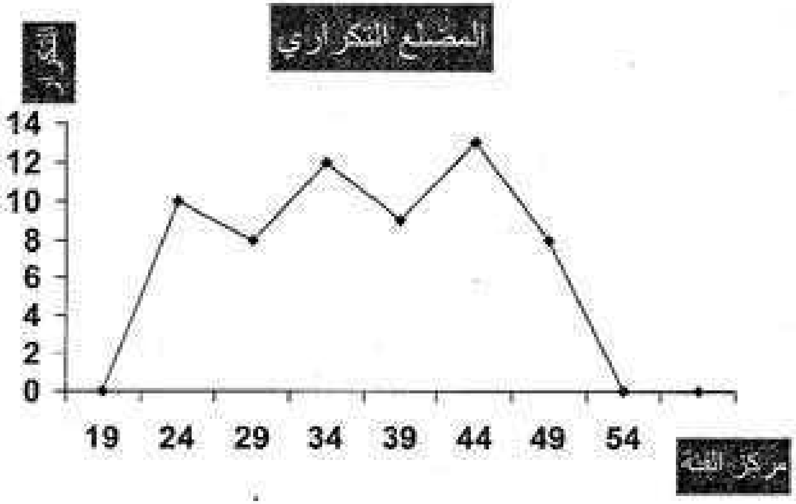
نأخذ مركز كل من هاتين الفئتين ونغلق المضلع كما في الشكل الآتي والذي يُمثل المضلع التكراري للتوزيع التكراري للجدول أعلاه



ويتضح من الشكل أن مجموع مساحات المستطيلات يساوي المساحة تحت المضلع التكراري.

وهناك طريقة أخرى لرسم المضلع التكراري وذلك بأخذ محورين متعامدين يمثل المحور الأفقي مركز الفئات ويمثل المحور الرأسي التكرارات.

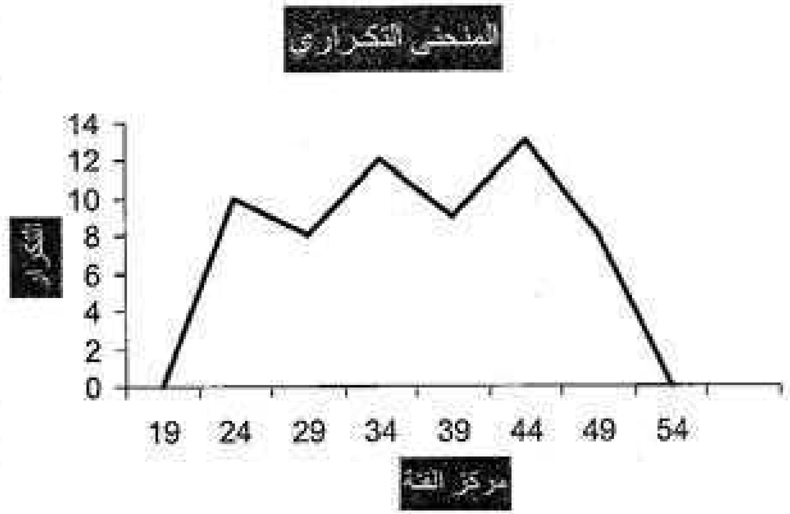
- نعتبر مركز كل فئة إحداثياً أفقياً لنقطة، ونعتبر تكرار هذه الفئة الإحداثي العمودي لتلك النقطة.
- نرصد جميع هذه النقاط ونوصل فيما بينها بخطوط مستقيمة .
- نُعين مركز فئة قبل الأولى ونعتبر تكرارها صفر ونُعين مركز فئة بعد الفئة الأخيرة مباشرة ونعتبر تكرارها صفر.
- نرصد هاتين النقطتين ونغلق بواسطتهما المضلع التكراري كما هو موضح بالشكل الآتي.



ثالثاً: المنحنى التكراري: Frequency Curve .

إذا مهدنا المضلع التكراري وجعلناه منحنى بدلاً من خطوط منكسرة فإننا نحصل على المنحنى التكراري. ويلاحظ أنه ينبغي عدم رسم المنحنى التكراري إلا إذا كانت الفئات كثيرة العدد وذات طول صغير، وكان عدد البيانات كبيراً، وكانت هذه البيانات من النوع المتصل مثل الزمن والوزن.

وللمنحنى التكراري والمساحة المحصورة بينه وبين المحور الأفقي أهمية كبيرة في دراسة الإحصاء، والشكل الآتي يوضح المنحنى التكراري للمضلع أعلاه:



وبنفس الطريقة التي مثلنا بها التوزيع التكراري بيانياً نستطيع تمثيل التوزيع التكراري النسبي، وذلك باستعمال التكرار النسبي على المحور العمودي بدلاً من التكرار.

نشاط (4)

إذا كانت مراكز الفئات لأعمار عدد معين من المرضى هي :

18 ، 21 ، 24 ، 27 ، 30 ، 33 ، 36

أوجد :

1. طول الفئة . 2. حدود الفئة لهذا التوزيع . 3. عدد الفئات .

الحل: ....

....

....

....

....

تدريب (3)

سُحبت عينة عشوائية من العاملين في إحدى شركات صناعة الأدوية وكان إنتاجهم الأسبوعي بالوحدة كما هو موضح في الجدول الآتي :

كمية الإنتاج	20 -24	25 -29	30 -34	35 -39	40 -44	45 -49	50 -54	55 -59	المجموع
عدد العاملين	14	22	32	55	44	36	27	20	250

المطلوب :

- رسم كل من المدرج التكراري و المضلع التكراري و المنحنى التكراري للبيانات في الجدول .
- رسم المدرج التكراري النسبي .
- تكوين جدول التكرار المتجمع الصاعد والهابط مع التمثيل بيانياً.
- احسب عدد العمال الذين ينتجون 30 وحدة فأكثر ، ثم نسبتهم.

تمارين الفصل الثالث

س1/ ضع خطاً تحت الإجابة الصحيحة فيما يلي :

1. عند تمثيل المدرج التكراري فإننا نمثل على المحور الأفقي:

أ. مراكز الفئات   ب. الحدود الفعلية للفئات   ج. الفئات   د. التكرار

2. مجموع التكرارات النسبية =

أ. 10   ب. 1   ج. مجموع التكرارات   د. صفر

3. للفئة (70- 60) فإن الحد الأعلى الفعلي:

أ. 61   ب. 70.5   ج. 71   د. 70

4. إذا كان حجم المجتمع الإحصائي 1000 مفردة وحجم فئة منه 100 مفردة فإن زاوية القطاع الدائري لها=

أ. 90°   ب. 36°   ج. 45°   د. 100°

5. عند عرض البيانات باستخدام الجداول يجب مراعاة ذكر ما يلي عدا:

أ- عنوان الجدول .   ب- الوحدات المستعملة.

ج- المصادر التي أخذت منها البيانات .   د- نوع الجدول .

6. من طرق عرض البيانات الإحصائية بيانياً ما يلي عدا :

أ- الأعمدة البيانية .   ب- الدائرة .   ج- المثلث .   د- الخط المنكسر.

7. طول الفئة =

أ- الحد الأعلى+ الحد الأدنى .   ب- (الحد الأعلى- الحد الأدنى) + 1.

ج- الحد الأعلى/ الحد الأدنى   د- (الحد الأعلى- الحد الأدنى)/2 .

8. من طرق تمثيل التوزيعات التكرارية بيانياً ما يلي عدا :

أ- المدرج التكراري .   ب- المضلع التكراري .

ج- المنحنى التكراري .   د- الخط المتكسر.

س2 / إذا كان عدد طلاب إحدى الكليات الطبية للعام 2012 – 2013م حسب التخصصات المختلفة كما يلي:

التخصص	العدد
صيدلة	540
مختبرات	180
أسنان	210
مساعدين	300
تمريض	270
المجموع	

المطلوب :

عرض هذه البيانات بطريقة المستطيلات وبطريقة الدائرة.

س3 / إذا كان لديك جدول التوزيع التكراري الآتي:

حدود الفئة	التكرار
8 -19	1
20 -31	2
32 -43	5
44 -55	8
56 -67	11
68 -79	14
80 -91	6
المجموع	

المطلوب /

أ - رسم المضلع والمدرج والمنحنى التكراري لهذا التوزيع.

ب- تكوين جدول التوزيع التكراري النسبي وتمثيله بيانياً.

ج- تكوين جدولي التكرار المتجمع الصاعد والهابط وتمثيلهما بيانياً.

س4 / إذا كانت درجات (84) طالباً في مادة إدارة مستشفيات كما يلي:

56	88	75	57	89	90	89	97	77	56	99	89
78	84	56	76	87	99	85	45	65	56	57	68
98	34	52	22	34	49	93	82	29	70	49	40
75	91	80	63	62	38	39	55	67	29	88	99
43	76	95	46	67	68	78	85	83	73	86	56
45	78	65	45	67	76	88	77	40	66	44	19
67	89	96	94	87	85	59	68	62	73	75	82

المطلوب :

أ- ضع البيانات السابقة في جدول توزيع تكراري ذي (7) فئات متساوية.

ب- ارسم المدرج التكراري والمضلع التكراري للبيانات في الجدول الذي كونته.

ج- كون جدول التوزيع التكراري النسبي والصاعد والنازل للبيانات في الجدول الذي كونته.

س5 / الجدول الآتي يوضح توزيعات التكرار المتجمع الصاعد والنازل لعدة بيانات.

الفئات	التكرار	مركز الفئة	التكرار المتجمع الصاعد	التكرار المتجمع الهابط
5-9	3			
10-14	9			
15-19	8			
20-24	10	22		
25-29	12			
المجموع	50			

المطلوب:

أ. أكمل الجدول أعلاه.

ب- ارسم منحنى التكرار المتجمع الصاعد، ومنحنى التكرار المتجمع الهابط على الشكل نفسه.

ج- ارسم عموداً من نقطة التقاطع على المحور الأفقي ما إحداثيات هذه النقطة ؟ وماذا تمثل ؟

3

الفصل الثالث

مقاييس النزعة المركزية

محتويات الفصل الثالث

أهداف الفصل الثالث

يتوقع من الطالب بعد الانتهاء من دراسة هذا الفصل أن يكون قادراً على:

- ✓ توضيح مفهوم النزعة المركزية .
- ✓ حساب مقاييس النزعة المركزية بطرق مختلفة.
- ✓ المقارنة بين مقاييس النزعة المركزية (المتوسط والوسيط والمنوال).
- ✓ حساب بعض مقاييس النزعة المركزية بيانياً .
- ✓ معرفة العلاقة بين كل من المتوسط الحسابي والوسيط والمنوال .
- ✓ معرفة مميزات وعيوب كل مقياس من مقاييس النزعة المركزية .
- ✓ حساب مقاييس الموضع (الربيعات والعشيرات والمئينات) .

❖مفهوم مقاييس النزعة المركزية (Measures of Central Tendency) .

تناولنا في الفصل السابق طُرق عرض البيانات جدولياً وبيانياً للتوزيعات التكرارية المختلفة ، ولا شك أن عرض البيانات جدولياً وبيانياً يساعد في وصف وتفسير الظواهر ويعطي فكرة عامة وسريعة عن طبيعة البيانات لكنها ليست كافية نظراً لاقتصارها على الوصف ، كما أن طُرق العرض السابقة قد تختلف باختلاف الباحثين، مما يعني تعدد الاستنتاجات حول ظاهرة واحدة ، لذلك يُفضل تلخيص البيانات باستخدام مؤشرات أو مقاييس رقمية تقيس خصائص الظاهرة وتساعد في مقارنة الظواهر المختلفة ، لذا سوف نتناول في هذا الفصل الأساليب الإحصائية التي تعطي قيم رقمية ذات دلالة معينة حول الظاهرة المدروسة .

يلاحظ أن التوزيعات التكرارية السابق الإشارة إليها قد أظهرت أن بيانات كل ظاهرة تميل للتجمع أو التمركز حول نقطة معينة ، لذلك نسعى إلى تحديد تلك النقطة من خلال عدة أساليب إحصائية تسمى مقاييس النزعة المركزية التي تعكس نقطة تمركز البيانات ، وهناك العديد من مقاييس النزعة المركزية تختلف فيما بينها من حيث القدرة في تلخيص أكبر قدر ممكن من المعلومات المستخرجة من بيانات العينة الممثلة للظاهرة.

وهي عدة مقاييس عددية مثل المتوسط الحسابي والوسيط والمنوال ، تُحدد موقع التوزيع ، فقد يكون هناك توزيعات تكرارية متشابهة في طبيعتها وشكلها ، لكنها تختلف في مواقعها ، وفي مثل هذه الحالة تكون معرفة مقاييس النزعة المركزية ذات فائدة في دراسة الفرق بين هذه التوزيعات التكرارية ، كما أن مقاييس النزعة المركزية تُعطي صورة جزئية عن المعلومات المحتواه في البيانات .

ومن أهم مقاييس النزعة المركزية ما يلي :

- الوسط الحسابي أو المتوسط الحسابي (Arithmetic Mean) .
- الوسيط (The Median) .
- المنوال (The Mode) .
- الوسط التوافقي (Harmonic Mean) .
- الوسط الهندسي (The Geometric Mean) .

وسنقتصر على المقاييس الأكثر استخداماً وهي المتوسط الحسابي والوسيط والمنوال.

أولاً : الوسط الحسابي / المتوسط الحسابي (Arithmetic Mean) .

تعريفه : المتوسط الحسابي لعدة قيم (ملاحظات او مشاهدات) هو مجموع هذه القيم مقسوماً على عددها، ويرمز له بالرمز ($\overline{X}$) ويقرأ إكس بار.

طرق حسابه :

الحالة الأولى: المتوسط الحسابي لبيانات خام (غير مبوبة).

القانون المستخدم في هذه الحالة:

المتوسط الحسابي لعدد من القيم = مجموع هذه القيم ÷ عدد القيم

ورمزياً : إذا كان لدينا عدد n من القيم (الملاحظات) $X_1 , X_2 , ..., X_n$ فإن المتوسط الحسابي لها يحسب من العلاقة الآتية :

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x_1 + x_2 + ... + x_n}{n}$$

مثال (1)

إذا كان عدد المترددين على عيادة خلال أسبوع كما يلي:

13 ، 15 ، 12 ، 10 ، 7 ، 5 ، 8

فأوجد المتوسط الحسابي لعدد المترددين على العيادة في اليوم الواحد.

الحل

المتوسط الحسابي ($\overline{X}$) = مجموع القيم ÷ عددها

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{8+5+7+10+12+15+13}{7} = 10$$

إذاً متوسط عدد الزوار في اليوم الواحد = 10

نشاط (1)

إذا كانت البيانات الآتية تمثل اعمار خمسة اشخاص (لأقرب سنة) :

17 ، 13 ، 8 ، 12 ، 10 فأوجد المتوسط الحسابي لهذه الأعمار .

الحل:...

....

....

تدريب (1)

إذا كانت أوزان 9 من الأطفال على النحو الآتي (لأقرب كجم):

36 ، 23 ، 28 ، 41 ، 40 ، 35 ، 27 ، 32 ، 25

فأوجد المتوسط الحسابي لأوزان الأطفال ؟

الحالة الثانية: المتوسط الحسابي لبيانات تكرارية بدون فئات .

القانون المستخدم في هذه الحالة :

المتوسط الحسابي = مجموع (القيم × تكراراتها) ÷ مجموع التكرارات .

ورمزياً :

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

حيث :

x_i : الملاحظة أو المشاهدة .

$\sum_{i=1}^{n} f_i$: مجموع التكرارات .

علماً بأن حجم العينة = مجموع التكرارات ، أي ان : $\sum_{i=1}^{n} f_i = n$

مثال (2)

البيانات الآتية تمثل عدد (51) فرداً في 9 أُسر:

رقم الأسرة	1	2	3	4	5	6	7	8	9	المجموع
التكرار (عدد افراد الأسرة)	10	7	9	7	4	4	5	3	2	51

أوجد المتوسط الحسابي .

الحل

نكون الجدول الآتي :

رقم الأسرة X	التكرار F (عدد افراد الأسرة)	X . F
1	10	10
2	7	14
3	9	27
4	7	28
5	4	20
6	4	24
7	5	35
8	3	24
9	2	18
المجموع	51	200

يُحسب المتوسط الحسابي كما يلي:

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i f_i}{n} = \frac{200}{51} = 3.92 \approx 4$$

نشاط (2)

إذا كانت البيانات الواردة في الجدول الآتي تمثل أطوال مجموعة من الأفراد (لأقرب سم) احسب المتوسط الحسابي لهذه البيانات.

الطول لأقرب سم X	عدد الأفراد (التكرار) F	X . F
120	9	
155	7	
170	11	
114	5	
125	4	
165	7	
132	3	
147	6	
المجموع		

الحل: ....

....

....

تدريب (2)

إذا كانت البيانات الواردة في الجدول التالي تمثل أوزان مجموعة من الأفراد (لأقرب كجم) :

الوزن X	50	62	75	85	82	55	59
التكرار F	4	9	8	11	5	7	6

المطلوب :

أوجد المتوسط الحسابي للبيانات في الجدول اعلاه .

الحالة الثالثة: المتوسط الحسابي لبيانات تكرارية بفئات .

القانون المستخدم في هذه الحالة :

المتوسط الحسابي = مجموع (مركز الفئة × التكرار) ÷ مجموع التكرارات .

ورمزياً:

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

حيث X_i مركز الفئة

مثال (3)

من جدول التوزيع التكراري الآتي اوجد المتوسط الحسابي :

الفئة	مركز الفئة X	التكرار F	X . F
30 – 34	32	3	96
35 – 39	37	6	222
40 – 44	42	10	420
45 – 49	47	4	188
50 – 54	52	5	260
55 – 59	57	2	114
المجموع		30	1300

الحل

يحسب المتوسط الحسابي من العلاقة السابقة كما يلي :

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i f_i}{n} = \frac{1300}{30} = 43.333$$

نشاط (3)

إذا كانت البيانات الآتية تمثل اعمار مجموعة من المرضى لأقرب سنة ، احسب المتوسط الحسابي لهذه الأعمار :

العمر بالسنة	التكرار F	الفئة الفعلية	مركز الفئة X	X . F
5 – 9	18			
10 – 11	58			
12 – 13	62			
14 – 16	72			
17 – 19	57			
20 – 22	42			
23 – 26	36			
27 – 36	55			
المجموع	400			7112.5

الحل:

تدريب (3)

إذا كانت لديك البيانات الآتية :

23	6	25	14	9	5	7
23	12	19	15	7	8	14
5	3	21	17	18	9	4
13	16	14	24	22	11	2

- كون جدول توزيع تكراري بفئات بحيث طول الفئة يساوي 4 .
- أحسب المتوسط الحسابي من الجدول الذي تم تكوينه .

الحالة الرابعة : طريقة الوسط الفرضي .

تستخدم هذه الطريقة إذا كانت الأرقام التي سيتم معالجتها كبيرة بهدف تبسيط إجراء العمليات الحسابية حسب الخطوات الآتية :

1. تحديد مراكز الفئات وتكراراتها.
2. تحديد وسط فرضي من بين مراكز الفئات ويفضل أن يكون مركز الفئة ذات التكرار الأكبر .
3. تحسب انحراف مراكز الفئات عن الوسط الفرضي وذلك بطرح الوسط الفرضي من مركز كل فئة ووضع النتائج في عمود جديد.
4. توجد حاصل ضرب كل انحراف بتكرار الفئة التي يعود إليها وتضع النتائج في عمود اخر.
5. يحسب الوسط الحسابي من العلاقة الآتية :

$$\overline{X} = L + \frac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i}$$

حيث L : الوسط الفرضي .

d : انحراف مركز الفئة عن الوسط الفرضي .

مثال (4)

من البيانات الواردة في الجدول الآتي أحسب الوسط الحسابي بطريقة الوسط الفرضي.

الفئة	F التكرار	X مركز الفئة	D	f . d
385 – 399	1	392	-90	-90
400 – 414	2	407	-75	-150
415 – 429	3	422	-60	-180
430 – 444	4	437	-45	-180
445 – 459	5	452	-30	-150
460 – 474	6	367	-15	-90
475 – 489	8	482	0	0
490 – 504	6	497	15	90
505 – 519	5	512	30	150
520 – 534	4	527	45	180
535 – 549	3	542	60	180
550 – 564	2	557	75	150
565 – 579	1	572	90	90
المجموع	50			0

الحل

نلاحظ أن أكبر تكرار هو 8 وهو تكرار الفئة 475 – 489 وبذلك يمكن تحديد مركز هذه الفئة (482) ليكون وسطاً فرضياً ونتائج الحل كما يلي :

$$\overline{X} = L + \frac{\sum_{i=1}^{n} f_i d_i}{\sum_{i=1}^{n} f_i} = 482 + \frac{0}{50} = 482$$

تدريب (4)

من التدريب رقم (3) السابق احسب الوسط الحسابي بطريقة الوسط الفرضي .

الحالة الخامسة : طريقة الانحرافات المختصرة :

إن الهدف من استخدام هذه الطريقة هو تبسيط العمليات الحسابية ويُفضل استخدام هذه الطريقة في حالة الجداول المنتظمة وتتم العمليات الحسابية حسب الخطوات الآتية:

1. تحديد مراكز الفئات وتكراراتها .
2. تحديد وسط فرضي من بين مراكز الفئات .
3. تحسب انحراف مراكز الفئات عن الوسط الفرضي وذلك بطرح الوسط الفرضي من مركز كل فئة وذلك بوضع النتائج في عمود جديد.
4. تقسم كل قيمة في عمود الانحرافات على طول الفئة (C) فنحصل على الانحرافات المختصرة.
5. توجد حاصل ضرب كل انحراف في عمود الانحرافات المختصرة بتكرار الفئات التي يعود إليها وتضع النتائج في عمود آخر ، ويحسب الوسط الحسابي بطريقة الانحرافات المختصرة من العلاقة التالية :

$$\overline{X} = L + \frac{\sum_{i=1}^{n} f_i d_m}{\sum_{i=1}^{n} f_i} . C$$

حيث L : الوسط الفرضي .

d_m : الانحرافات المختصرة .

C : طول الفئة .

مثال (5)

من البيانات الواردة في الجدول في (مثال 4) احسب الوسط الحسابي بطريقة الانحرافات المختصرة.

الفئة	التكرار F	مركز الفئة X	d	dm = d/c	(dm) (f)
385 – 399	1	392	- 90		
400 – 414	2	407	- 75		
415 – 429	3	422	- 60		
430 – 444	4	437	- 45		
445 – 459	5	452	- 30		
460 – 474	6	367	- 15		
475 – 489	8	482	0		
490 – 504	6	497	15		
505 – 519	5	512	30		
520 – 534	4	527	45		
535 – 549	3	542	60		
550 – 564	2	557	75		
565 – 579	1	572	90		
المجموع	50				

الحل

1. نكمل الجدول أعلاه ثم نحسب الوسيط من العلاقة الآتية :

$$\overline{X} = L + \frac{\sum_{i=1}^{n} f_i d_m}{\sum_{i=1}^{n} f_i} . C = 482 + \frac{0}{50} . 15 = 482$$

❖ **بعض خواص الوسط الحسابي:**

1. قيمة الوسط الحسابي لا تختلف عند حسابها بأكثر من طريقة لنفس البيانات.
2. لا يمكن حساب الوسط الحسابي للبيانات النوعية.
3. لا يمكن حساب الوسط الحسابي من الجداول التكرارية المفتوحة.
4. لا يمكن إيجاد الوسط الحسابي بيانياً، إلا في حالة التوزيع الطبيعي.
5. مجموع انحرافات القيم عن وسطها الحسابي يساوي صفر أي أن :

$$\sum_{i=1}^{n}(X_i-\overline{X})=0$$

فمثلاً: إذا كان لدينا البيانات الآتية (6،4،2) فإن وسطها الحسابي = 4 ، وبالتالي فإن

$$\sum_{i=1}^{n}(X_i-\bar{x})=(2-4)+(4-4)+(6-4)=0$$

6. تتأثر قيمة الوسط الحسابي كثيراً بالقيم المتطرفة (القيم الصغيرة جداً أو الكبيرة جداً).

مثال ذلك: إذا كانت رواتب (5) موظفين بالدولار كما يلي: (200,300,250,350, 9000) فإن المتوسط الحسابي للرواتب

$$\overline{X}=\frac{\sum_{i=1}^{n}x_i}{n}=\frac{200+300+250+350+9000}{5}=\frac{10100}{5}=2020$$

فنلاحظ من خلال المثال أن القيمة 9000 كبيرة مقارنة ببقية القيم، كما نلاحظ أنها أثرت على المتوسط الحسابي لرواتب الموظفين.

7. يأخذ الوسط الحسابي جميع القيم بعين الاعتبار.
8. الوسط الحسابي محدد جبرياً بدقة ويتمتع بخواص جبرية لا يتمتع بها غيره من المقاييس.
9. لتكن $X_1, X_2, X_3,\dots, X_n$ مجموعة من القيم، وسطها الحسابي X فإذا اضفنا لكل قيمة من القيم الأصلية عدداً ثابتاً C فإن الوسط الحسابي الجديد يساوي الوسط الحسابي للقيم الأصلية مضافاً إليه نفس المقدار الثابت أي Y يصبح Y = X+C.

وبالمثل في حالة طرح مقدار ثابت من جميع القيم فإن الوسط الحسابي الجديد Y يكون على الصورة: Y = X - C .

مثال توضيحي للخاصية (9)

أوجد الوسط الحسابي للبيانات التالية : 3، 2، 6، 5، 4

الحل

الوسط الحساب يحسب من العلاقة

$$\overline{X}=\frac{\sum_{i=1}^{n}x_i}{n}=\frac{3+2+6+5+4}{5}=\frac{20}{5}=4$$

إذا تم إضافة قيمة ثابتة ولتكن (C = 7) لجميع القيم فإن القيم الجديدة سوف تكون :

10، 9، 13، 12، 11

والوسط الحسابي الجديد لهذه البيانات هو :

$$\overline{X}=\frac{\sum_{i=1}^{n}x_i}{n}=\frac{10+9+13+11+12}{5}=\frac{55}{5}=11$$

وهذا يعني أنه عند إضافة القيمة 7 للمتوسط السابق (4) نجد ان

المتوسط الحسابي الجديد = المتوسط للبيانات الأصلية (4) + القيمة المضافة للبيانات (7)= 11

وهي نفس القيمة التي حصلنا عليها أعلاه.

وعند طرح عدد ثابت من جميع القيم في المثال أعلاه وليكن مثلاً 4 ستحصل على القيم الآتية :

-2، -1، 2، 1، 0

وعند حساب المتوسط الحسابي لهذه القيم فإنه يساوي صفر

ولو طرحنا القيمة 4 من المتوسط الحسابي للبيانات الأصلية لحصلنا على نفس المتوسط = صفر.

10. لتكن $X_1, X_2, X_3,\dots, X_n$ مجموعة من القيم وسطها الحسابي X فإذا ضربنا كل قيمة من القيم الأصلية بعدد ثابت C فإن الوسط الحسابي الجديد Y يساوي الوسط الحسابي للقيم الأصلية مضروباً في المقدار الثابت نفسه Y = X . C وبالمثل نجد Y = X/C في حالة قسمة كل مفردة على عدد ثابت.

مثال توضيحي للخاصية (10)

لو اخذنا نفس القيم في المثال (6) والتي متوسطها = 4

لو ضربنا كل قيمة من هذه القيم في العدد 9 مثلاً فستحصل على القيم الجديدة الآتية :

36 ، 45 ، 54 ، 27، 18

نحسب المتوسط الحسابي للقيم الجديدة باستخدام القانون الآتي :

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{36 + 45 + 54 + 27 + 18}{5} = \frac{180}{5} = 36$$

نلاحظ انه لو ضربنا المتوسط للقيم الأصلية والذي يساوي 4 في العدد 9 لحصلنا على المتوسط الجديد والذي = 36

كذلك في حالة القسمة :

لو قسمنا جمع القيم السابقة في المثال (6) على العدد 2 لحصلنا على القيم الآتية :

1 ، 1.5 ، 3 ، 2.5 ، 2

نحسب المتوسط الحسابي للقيم الجديدة كما يلي :

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{1 + 1.5 + 3 + 2.5 + 2}{5} = \frac{10}{5} = 2$$

نلاحظ انه لو قسمنا المتوسط الحسابي للقيم الأصلية والذي = 4 على العدد 2 لحصلنا على المتوسط الجديد والذي = 2 .

> تدريب (5)
>
> اذكر خواص المتوسط الحسابي، مع التوضيح بأمثلة ما امكن .

ثانياً / الوسيط (The Median) .

يُعد الوسيط من مقاييس النزعة المركزية المهمة التي لا تتأثر بالقيم الشاذة كالوسط الحسابي، ولمعالجة العيوب التي وجهت للوسط الحسابي يُستخدم الوسيط كمقياس للنزعة المركزية ، ويرمز له بالرمز ($\tilde{X}$) وتقرأ (إكس تلدا) **ويُفضل استخدامه في حالة وجود قيم متطرفة.**

تعريفه :

الوسيط / هو القيمة التي تتوسط البيانات بعد ترتيبها تصاعدياً أو تنازلياً.

طُرق حسابه :

الحالة الأولى : حساب الوسيط في حالة البيانات غير المبوبة .

في حالة وجود بيانات غير مبوبة ، فيجب أولاً إعادة ترتيبها تصاعدياً أو تنازلياً ثم نحسب قيمة الوسيط كما يلي :

أ. إذا كان عدد المشاهدات (n) فردياً، فإن قيمة الوسيط عبارة عن قيمة المشاهدة التي ترتيبها $\frac{n+1}{2}$.

مثال (1)

احسب الوسيط للبيانات الآتية:

.12 ،9 ،10 ،11 ،8 ،15 ،12 ،9 ،10 ،13 ،8

الحل:

اولاً: نرتب البيانات تصاعدياً أو تنازلياً، (هنا سنرتبها تصاعدياً)

8، 8، 9، 9، 10، 10، 11، 12، 12، 13، 15

بما أن عدد المفردات فردياً (n =11)

إذاً ترتيب الوسيط = $\frac{n+1}{2} = \frac{11+1}{2} = 6$

ومنه قيمة الوسيط عبارة عن المفردة رقم (6) في البيانات بعد ترتيبها

إذاً $\tilde{X} = 10$

ب. عندما يكون عدد المشاهدات (n) **زوجياً**، فإن قيمة الوسيط هو متوسط المفردتين اللتين ترتيبهما $\frac{n}{2}$ ، $\frac{n}{2}+1$.

اي أن قيمة الوسيط في هذه الحالة تساوي المتوسط الحسابي للقيمتين الوسيطتين.

مثال (2)

احسب الوسيط للبيانات الآتية:

120، 50، 55، 45، 53، 50، 54، 52

الحل

أولاً : نرتب البيانات تصاعدياً أو تنازلياً كما يلي:

45، 50، 50، 52، 53، 54، 55، 120

بما أن عدد المفردات زوجياً (n = 8)

إذاً ترتيب الوسيط يقع بين المفردتين $\frac{n}{2}, \frac{n}{2}+1$

رتبة القيمة الوسيطية الأولى $\frac{n}{2}=\frac{8}{2}=4$

رتبة القيمة الوسيطية الثانية $\frac{n}{2}+1=\frac{8}{2}+1=5$

ومنه قيمة الوسيط تساوي المتوسط للقيمتين الوسيطيتين (الرابعة والخامسة).

$$\tilde{X} = \frac{52 + 53}{2} = 52.5$$

يلاحظ من هذا المثال أن الوسيط لم يتأثر بوجود قيم شاذة بين البيانات وهي المشاهدة الأخيرة وقيمتها 120

نشاط (1)

احسب الوسيط للبيانات الآتية:

أ. 5، 8، 9، 6، 14، 8، 6، 4، 7، 3، 5، 6، 9

ب. 8، 4، 13، 5، 16، 45، 7، 6، 7، 26

الحل: ....

....

....

....

....

....

تدريب (1)

أوجد الوسيط للبيانات الآتية:

أ- 26، 25، 34، 48، 76، 100، 69، 99، 56، 11، 8

ب- 5، 46، 68، 48، 97، 85، 91، 37، 82، 94، 46، 71

الحالة الثانية : حساب الوسيط لبيانات مبوبة بدون فئات .

وفي هذه الحالة يُمكن حساب الوسيط باتباع عدة طُرق هي :

أ. حساب الوسيط باستخدام التكرار المتجمع الصاعد .

ب. حساب الوسيط باستخدام التكرار المتجمع الهابط .

وسنتناول كل طريقة من هذه الطُرق بالتفصيل :

أ. حساب الوسيط باستخدام التكرار المتجمع الصاعد .

لحساب الوسيط باستخدام التكرار المتجمع الصاعد نتبع الخطوات الآتية :

- نكون جدولاً تكرارياً متجمعاً صاعداً .
- نحدد ترتيب الوسيط من العلاقة : ترتيب الوسيط = $\frac{n}{2}$ حيث n مجموع التكرارات.
- نحدد الفئة الوسيطية التي تحتوي على ترتيب الوسيط .
- يتم حساب الوسيط للتكرار المتجمع الصاعد من العلاقة :

$$\tilde{X} = L + \frac{\frac{n}{2} - f_1}{f_2} . C$$

حيث:

L : الحد الأدنى للفئة الوسيطية.

$\frac{n}{2}$: ترتيب الوسيط .

C : طول الفئة الوسيطية.

f_1 : التكرار المتجمع الصاعد للفئة السابقة للفئة الوسيطية.

f_2 : تكرار الفئة الوسيطية .

مثال (3)

من جدول التوزيع التكراري الآتي أوجد الوسيط باستخدام التكرار المتجمع الصاعد :

الملاحظة X	2	5	8	6	4	3	المجموع
التكرار F	7	6	11	5	2	1	32

الحل

1. نكون جدول توزيع تكراري متجمع صاعد .

الملاحظة X	التكرار F	التكرار المتجمع الصاعد ت.م.ص
2	7	7
3	1	8
4	2	10
5	6	16
6	5	21
8	11	32
المجموع	32	

2. نحدد ترتيب الوسيط من العلاقة : ترتيب الوسيط = $\frac{n}{2}$ حيث n مجموع التكرارات.

إذاً ترتيب الوسيط $\frac{32}{2} = 16$

3. نحدد الفئة الوسيطية بالبحث عن القيمة 16 في عمود التكرار المتجمع الصاعد من أعلى إلى أسفل وهي التي تناظر القيمة 5 ، نفترض ان القيمة 5 فئة بطول = 1 .

الحدود الفعلية لهذه الفئة هي 4.5 - 5.5

4. نحسب الوسيط من العلاقة الآتية :

$$\tilde{X} = L + \frac{\frac{n}{2} - f_1}{f_2}.C = 4.5 + \frac{16 - 10}{6}.1 = 5.5$$

ب- حساب الوسيط باستخدام التكرار المتجمع الهابط .

لحساب الوسيط باستخدام التكرار المتجمع الهابط (النازل) نتبع الخطوات الآتية :

- نكون جدولاً تكرارياً متجمعاً هابطاً .
- نحدد ترتيب الوسيط من العلاقة: ترتيب الوسيط = $\frac{n}{2}$ حيث n مجموع التكرارات.
- نحدد الفئة الوسيطية التي تحتوي على ترتيب الوسيط .
- يتم حساب الوسيط للتكرار المتجمع الهابط من العلاقة :

$$\tilde{X} = h - \frac{\frac{n}{2} - f_1}{f_2}.C$$

حيث:

h : الحد الأعلى للفئة الوسيطية.

n/2 : ترتيب الوسيط .

C : طول الفئة الوسيطية.

f_1 : التكرار المتجمع الهابط للفئة اللاحقة للفئة الوسيطية.

f_2 : تكرار الفئة الوسيطية .

مثال (4)

من المثال (3) السابق أوجد الوسيط باستخدام التكرار المتجمع الهابط

الحل:

1. نكون جدول توزيع تكراري متجمع هابط (ت . م .هـ).

الملاحظة X	التكرار F	ت.م.هـ
2	7	32
3	1	25
4	2	24
5	6	22
6	5	16
8	11	11
المجموع	32	

2. تحدد ترتيب الوسيط من العلاقة :

ترتيب الوسيط = $\frac{n}{2}$ حيث n مجموع التكرارات.

إذاً ترتيب الوسيط $\frac{32}{2} = 16$

3. تحدد الفئة الوسيطية بالبحث عن القيمة 16 في عمود التكرار المتجمع الهابط من أسفل إلى أعلى وهي التي تناظر القيمة 6

الحدود الفعلية للفئة 6 هي 5.5 - 6.5

4. تحسب الوسيط من العلاقة الآتية :

$$\bar{x} = h - \frac{\frac{n}{2} - f_1}{f_2}.C = 6.5 - \frac{16 - 11}{5}.1 = 5.5$$

تدريب (2)

من جدول التوزيع التكراري احسب الوسيط .

الملاحظة X	2	4	6	8	9	10	المجموع
التكرار F	5	7	2	4	3	6	27

الحالة الثالثة : حساب الوسيط لبيانات مبوبة بفئات .

في هذه الحالة ستتبع نفس الخطوات السابقة في حالة حساب الوسيط لبيانات مبوبة بدون فئات.

حيث يتم حساب الوسيط بطريقتين هما :

أ. حساب الوسيط باستخدام التكرار المتجمع الصاعد .

ب. حساب الوسيط باستخدام التكرار المتجمع الهابط .

وسنتناول كل طريقة من هذه الطرق بالتفصيل :

أ. حساب الوسيط باستخدام التكرار المتجمع الصاعد .

لحساب الوسيط باستخدام التكرار المتجمع الصاعد نتبع الخطوات الآتية:

- نكون جدولاً تكرارياً متجمعاً صاعداً .

- تحدد ترتيب الوسيط من العلاقة: ترتيب الوسيط = $\frac{n}{2}$ حيث n مجموع التكرارات.
- تحدد الفئة الوسيطية التي تحتوي على ترتيب الوسيط .
- يتم حساب الوسيط للتكرار المتجمع الصاعد من العلاقة التالية :

$$\tilde{X} = L + \frac{\frac{n}{2} - f_1}{f_2}.C$$

حيث

L : الحد الأدنى للفئة الوسيطية.

n/2 : ترتيب الوسيط .

C : طول الفئة الوسيطية.

f_1 : التكرار المتجمع الصاعد للفئة السابقة للفئة الوسيطية.

f_2 : تكرار الفئة الوسيطية .

مثال (5) :

من جدول التوزيع التكراري الآتي أوجد الوسيط باستخدام التكرار المتجمع الصاعد :

الفئة	3 - 7	8 - 12	13 - 17	18 - 22	23 - 27	المجموع
التكرار F	2	3	10	5	4	24

الحل

1. نكون جدول توزيع تكراري متجمع صاعد (ت . م . ص).

الفئة	التكرار F	ت.م.ص
3 – 7	2	2
8 – 12	3	5
13 – 17	10	15
18 – 22	5	20
23 - 27	4	24
المجموع	24	

2. تحدد ترتيب الوسيط من العلاقة : ترتيب الوسيط = $\frac{n}{2}$ حيث n مجموع التكرارات.

إذاً ترتيب الوسيط $\frac{24}{2} = 12$ ، ويقع هذا الترتيب ضمن التكرار المتجمع الصاعد 15 (من اعلى إلى اسفل) ، ويناظر الفئة 17 – 13 .

3. تحدد الفئة الوسيطية وهي 17 – 13

الحدود الفعلية للفئة الوسيطية هي 17.5 – 12.5

4. تحسب الوسيط من العلاقة الآتية:

$$\tilde{X} = L + \frac{\frac{n}{2} - f_1}{f_2}.C = 12.5 + \frac{12-5}{10}.5 = 12.5 + 3.5 = 16$$

ب. حساب الوسيط باستخدام التكرار المتجمع الهابط .

لحساب الوسيط باستخدام التكرار المتجمع الهابط (النازل) نتبع الخطوات الآتية :

- تكون جدولاً تكرارياً متجمعاً هابطاً .
- تحدد ترتيب الوسيط من العلاقة : ترتيب الوسيط = $\frac{n}{2}$ حيث n مجموع التكرارات.
- تحدد الفئة الوسيطية التي تحتوي على ترتيب الوسيط .
- يتم حساب الوسيط للتكرار المتجمع الهابط من العلاقة الآتية :

$$\tilde{X} = h - \frac{\frac{n}{2} - f_3}{f_2}.C$$

حيث:

h : الحد الأعلى للفئة الوسيطية.

n/2 : ترتيب الوسيط .

C : طول الفئة الوسيطية.

f_3 : التكرار المتجمع الهابط للفئة اللاحقة للفئة الوسيطة.

f_2 : تكرار الفئة الوسيطية .

مثال (6)

من المثال (5) اعلاه أوجد الوسيط باستخدام التكرار المتجمع الهابط

الحل

1. تكون جدولاً توزيعاً تكرارياً متجمعاً هابطاً .

الملاحظة X	التكرار F	التكرار المتجمع الهابط ت.م.هـ
3 – 7	2	24
8 – 12	3	22
13 – 17	10	19
18 – 22	5	9
23 – 27	4	4
المجموع	24	

2. نحدد ترتيب الوسيط من العلاقة :

ترتيب الوسيط = $\frac{n}{2}$ حيث n مجموع التكرارات.

إذاً ترتيب الوسيط $\frac{24}{2} = 12$ ، ويقع هذا الترتيب ضمن التكرار المتجمع الهابط 19، (من أسفل إلى أعلى) ويناظر الفئة 17 – 13 .

3. نحدد الفئة الوسيطية وهي 17 – 13

الحدود الفعلية للفئة الوسيطية هي 17.5 – 12.5

4. نحسب الوسيط من العلاقة الآتية :

$$\tilde{X} = h - \frac{\frac{n}{2} - f_1}{f_2}.C = 17.5 - \frac{12-9}{10}.5 = 17.5 - 1.5 = 16$$

نشاط (2)

من الجدول الآتي احسب الوسيط .

الملاحظة X	التكرار F	التكرار المتجمع الصاعد ت.م.ص
3 – 7	2	
8 – 12	3	
13 – 17	10	
18 – 22	5	
23 - 27	13	
28 - 32	4	
المجموع	37	

الحل:..

....

....

....

....

الحالة الرابعة: حساب الوسيط بيانياً .

ويتم حساب الوسيط بيانياً بطريقتين هما:

أ . برسم منحنيي التكرار المتجمع الصاعد والهابط على شكل واحد.

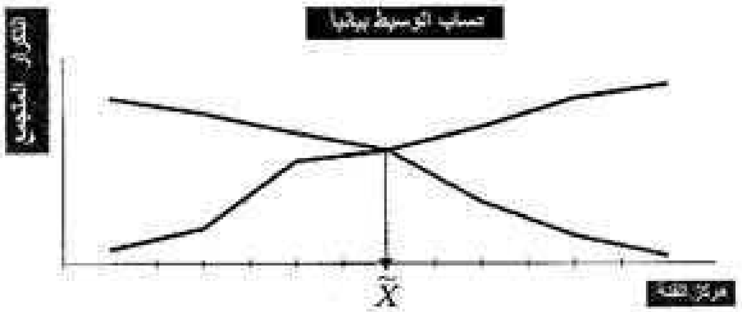
ب. برسم منحنى التكرار المتجمع الصاعد فقط .

وفيما يلي توضيح ذلك :

أ . برسم منحنيي التكرار المتجمع الصاعد والهابط على شكل واحد.

يمكننا إيجاد الوسيط بهذه الطريقة بإتباع الآتي :

- ترسم منحنى التكرار المتجمع الصاعد مع منحنى التكرار المتجمع الهابط على مستوى واحد ، ومن نقطة تقاطع هذين المنحنيين ننزل عموداً على المحور الأفقي .
- نقطة تقاطع العمود مع المحور الأفقي تمثل الوسيط بيانياً .

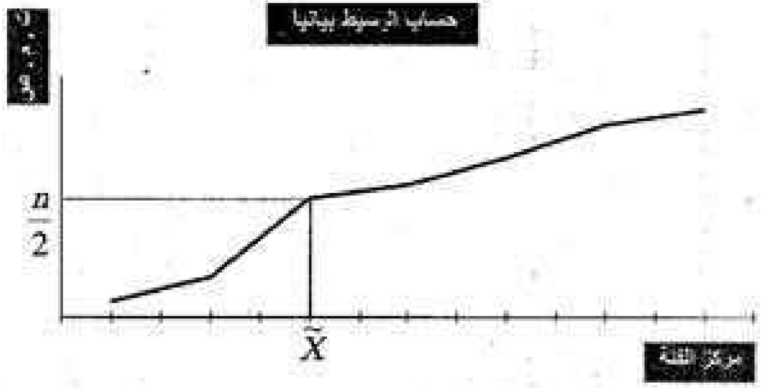


ب. برسم منحنى التكرار المتجمع الصاعد فقط .

يمكننا إيجاد الوسيط بهذه الطريقة بإتباع الآتي :

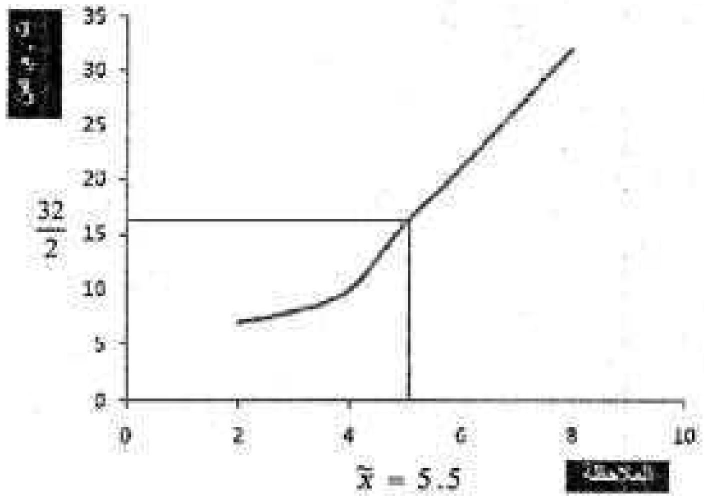
- ترسم منحنى التكرار المتجمع الصاعد .
- نحدد ترتيب الوسيط على المحور الرأسي (محور التكرار المتجمع الصاعد).
- من النقطة التي تمثل ترتيب الوسيط على المحور الرأسي نرسم مستقيم موازي للمحور الأفقي .

- من نقطة تقاطع هذا المستقيم مع منحنى التكرار المتجمع الصاعد نسقط عمود على المحور الأفقي .
- نقطة تقاطع هذا العمود مع المحور الأفقي تمثل الوسيط بيانياً.



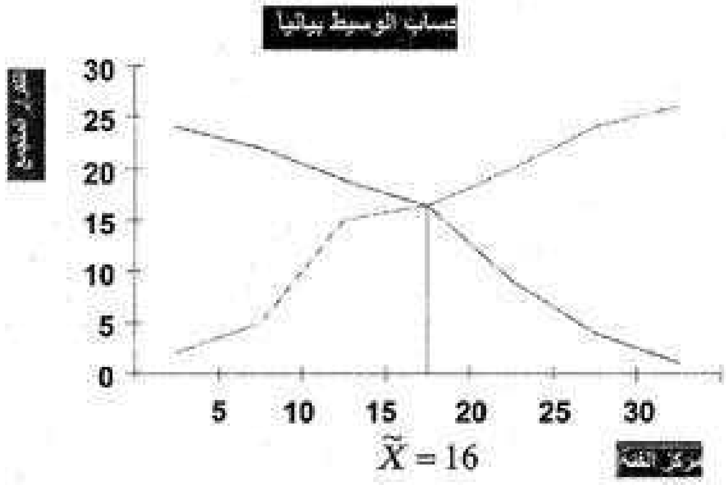
مثال (7)

من جدول التوزيع التكراري في مثال (3) اوجد الوسيط بيانياً .



مثال (8)

من جدول التوزيع التكراري في مثال (5) اوجد الوسيط بيانياً .



نشاط (2)

من الجدول الآتي أوجد الوسيط بيانياً .

الملاحظة X	التكرار F	التكرار المتجمع الصاعد ت.م.ص
4 - 10	7	
11 - 17	9	
18 - 24	8	
25 - 31	11	
المجموع		

الحل:...

....

....

- **خواص الوسيط:**

1- يقع الوسيط في أي توزيع تكراري غير متماثل بين الوسط الحسابي والمنوال.
2- يتأثر الوسيط بالقيم الوسطى ولا يتأثر بالقيم المتطرفة.
3- سهولة حساب قيمته.
4- يُمكن حسابه إذا كانت إحدى الفئات مفتوحة من أعلى أو من أسفل الجدول.
5- يُمكن حسابه إذا كانت الفئات غير متساوية الطول.
6- مقياس غير محدد بدقة في الجداول التكرارية إذ أن طريقة حسابه تعطي قيمة تقريبية.
7- لا يتمتع بأي خواص جبرية بحيث يُمكن الاستفادة منه في حسابات لاحقة.
8- لا يُمكن حسابه للبيانات النوعية.

تدريب (5)

اذكر خواص الوسيط، مع التوضيح بأمثلة ما أمكن .

ثالثاً / المنوال (The Mode)

تعريفه:

هو الملاحظة (القيمة) الأكثر تكراراً في البيانات الإحصائية. ويرمز له بالرمز $\hat{X}$ ويُقرأ (إكس هات).

طُرق حسابه :

الحالة الأولى : حساب المنوال لبيانات غير مبوبة .

مثال(1)

احسب المنوال لكل مما يأتي:

أ) 2، 2، 5، 5، 5، 4، 4، 4، 4، 5، 5، 5

الحل:

المنوال هو 5 لأنها تكررت ست مرات أكثر من القيم الأخرى.

ب) 3، 5، 7، 2، 9، 4، 6

الحل:

تكرار هذه البيانات متساوية لذلك فهي عديمة المنوال.

ج) 4، 4، 4، 5، 5، 6، 6، 6، 6، 7، 7، 7، 10، 10، 8، 8، 8، 8

الحل:

لدينا منوالان هما 8 ، 6 . إذا تساوت قيمتان من حيث تكرارهما يكون للقيم منوالان.

نشاط (1)

أوجد المنوال للبيانات الآتية :

أ- 8 ، 8 ، 5 ، 7 ، 4 ، 4 ، 4 ، 6 ، 6 ، 8 ، 9 ، 7 ، 8 ، 10

ب- 5 ، 6 ، 4 ، 8 ، 9 ، 1 ، 7

ج- 3 ، 5 ، 2 ، 2 ، 3 ، 5 ، 6 ، 3 ، 2 ، 9 ، 7

الحل:...

....

....

....

تدريب (1)

أوجد المنوال للبيانات الآتية :

أ- 78 ، 45 ، 69 ، 45 ، 76 ، 12 ، 32 ، 99

ب- 98 ، 87 ، 65 ، 54 ، 32 ، 21 ، 57

ج- 10 ، 5 ، 4 ، 5 ، 10 ، 5 ، 5 ، 5 ، 8 ، 10 ، 9 ، 10 ، 6

الحالة الثانية : حساب المنوال لبيانات مبوبة بدون فئات .

المنوال = الملاحظة (القيمة) الأكثر تكراراً .

مثال (2)

أوجد المنوال للبيانات الآتية :

أ)

التكرار F	الملاحظة X
3	8
8	10
5	12
7	14
9	16
6	18
	المجموع

الحل

المنوال يساوي 16 لأنها الملاحظة الأكثر تكراراً .

ب)

التكرار F	الملاحظة X
4	8
3	10
8	12
6	14
8	16
7	18
	المجموع

الحل:

لدينا منوالان هما 12 ، 16

ج)

التكرار F	الملاحظة X
5	8
5	10
5	12
5	14
5	16
5	18
	المجموع

الحل:

لا يوجد منوال (عديمة المنوال) لأن القيم لها نفس التكرار.

الحالة الثالثة : حساب المنوال لبيانات مبوبة بفئات.

في التوزيعات التكرارية ذات الفئات يُمكن حساب المنوال بطريقتين هما :

الطريقة الأولى:

المنوال = مركز الفئة المنوالية .

والفئة المنوالية : هي الفئة ذات التكرار الأكبر.

مثال (3)

اوجد المنوال لكل من جداول التوزيعات التكرارية الآتية:

أ)

الفئات	التكرار f
10 – 14	4
15 – 19	7
20 – 24	8
25 – 29	5
30 – 34	3
المجموع	

الحل: أعلى تكرار يساوي (8)

الفئة التي تقابل أعلى تكرار هي (20 – 24)

مركز الفئة ذات التكرار الأكبر 22 = 2 ÷ 44 = 2 ÷ (24 + 20)

إذاً المنوال يساوي (22).

ب)

الفئات	التكرار f
10 – 14	3
15 – 19	9
20 – 24	6
25 – 29	9
30 – 34	2
المجموع	

الحل:

نلاحظ أن أعلى تكرار يساوي (9)

ولدينا تكرارين تقابلهما الفئتين (15 – 19) ، (25 – 29)

إذاً لدينا منوالان هما مراكز الفئتين 17 = 2 ÷ 34 = 2 ÷ (19 + 15)

27 = 2 ÷ 54 = 2 ÷ (29 + 25) ، إذاً المنوالين هما 27،17.

ج)

الفئات	التكرار f
10 – 14	6
15 – 19	6
20 – 24	6
25 – 29	6
30 – 34	6
المجموع	

الحل: بما أن التكرارات متساوية إذاً البيانات عديمة المنوال.

2. طريقة الفروق لبيرسون .

ويحسب المنوال حسب هذه الطريقة من العلاقة الآتية :

$$\hat{x} = L + \frac{d_1}{d_1 + d_2} . C$$

حيث $\hat{x}$: المنوال

L : الحد الأدنى للفئة المنوالية .

d_1 : الفرق بين تكرار الفئة المنوالية وتكرار الفئة السابقة لها.

d_2 : الفرق بين تكرار الفئة المنوالية وتكرار الفئة اللاحقة لها .

C : طول الفئة المنوالية .

مثال (4)

من جدول التوزيع التكراري الآتي أوجد المنوال بطريقة الفروق لبيرسون :

الفئات	التكرار f
10 – 14	4
15 – 19	7
20 – 24	8
25 – 29	5
30 – 34	3

الحل

من الجدول نلاحظ ان الفئة المقابلة لأكبر تكرار (الفئة المنوالية) هي 24 – 20 وان تكرار الفئة قبل المنوالية = 7 ، وتكرار الفئة بعد المنوالية = 5 ، وطول الفئة المنوالية = 5 ، وتكرار الفئة المنوالية = 8

وبالتعويض في القانون سنحصل على المنوال -

$$\hat{x} = L + \frac{d_1}{d_1 + d_2}.C = 20 + \frac{1}{1+3} \times 5 = 21.25$$

نشاط (2)

من جدول التوزيع التكراري الآتي أوجد المنوال بطريقة الفروق لبيرسون:

التكرار f	الفئات
7	10 – 14
5	15 – 19
8	20 – 24
3	25 – 29
9	30 – 34
	المجموع

الحل:...

....

....

....

الحالة الرابعة : حساب المنوال بيانياً.

يُمكن حساب المنوال لأي بيانات برسم منحنى التوزيع التكراري لهذه البيانات، ومن أعلى قمة في المنحنى (التي تناظر أعلى تكرار) نُسقط عموداً على المحور الأفقي، نقطة التقاء هذا العمود مع المحور الأفقي تُمثل المنوال بيانياً.

كما يُمكن حساب المنوال للبيانات المبوبة بيانياً من خلال رسم المدرج التكراري للفئة المنوالية والفئة السابقة لها والفئة اللاحقة لها ، ثم يتم إيصال خط من الزاوية العلوية اليمنى للفئة السابقة للفئة المنوالية مع الزاوية العلوية اليمنى للفئة المنوالية، وإيصال خط من الزاوية العليا اليسرى للفئة اللاحقة للفئة المنوالية مع الزاوية العلوية اليسرى للفئة المنوالية .

من نقطة تقاطع الخطين نُسقط عموداً على المحور الأفقي .

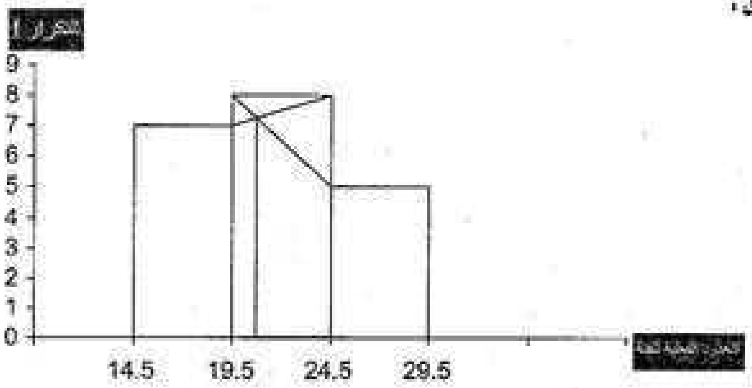
نُقطة تقاطع هذا العمود مع المحور الأفقي تُمثل المنوال بيانياً .

والمثال الآتي يوضح ذلك .

مثال (5)

من مثال (4) السابق أوجد المنوال بيانياً .

الحل :



$\hat{x} = 21.25$

نلاحظ من الشكل أن المنوال يساوي تقريباً 21.25

تدريب (2)

من جدول التوزيع التكراري التالي أوجد :

الفئة	20 - 29	30 - 39	40 - 49	50 - 59	60 - 69	المجموع
التكرارات	180	250	210	190	350	1280

1. المنوال بطريقة الفروق لبيرسون ؟
2. المنوال بيانياً ؟

❖ خواص المنوال:

1- لا يتأثر المنوال بالقيم المتطرفة والوسطى في التوزيع التكراري بل يتأثر بالتكرارات.
2- يتأثر بعدد فئات التوزيع التكراري ومداها.
3- يُمكن حسابه من الجداول التكرارية المفتوحة.
4- ليس مناسباً للعمليات الجبرية.
5- يعتمد على جزء من البيانات.
6- قد يوجد أكثر من منوال للبيانات وفي هذه الحالة يقال ان البيانات غير متجانسة.
7- يُمكن استخدامه في حالة البيانات النوعية.

❖ العلاقة بين المتوسط الحسابي والوسيط والمنوال :

مما سبق يُلاحظ أن الوسط الحسابي يتأثر بالقيم الشاذة ، وبالتالي فإن قيمته تتجه نحو ذيل التوزيع التكراري ، بينما المنوال يتأثر بالقيم الأكثر شيوعاً وبالتالي يتجه نحو قمة التوزيع ، اما الوسيط فإنه يقع بينهما نظراً لأن تأثره بالقيم الشاذة والقيم الشائعة أقل من الوسط الحسابي والمنوال على التوالي.

إذا كان لدينا توزيع تكراري به منوال واحد تكون العلاقة بين المتوسط والوسيط والمنوال كما يلي:

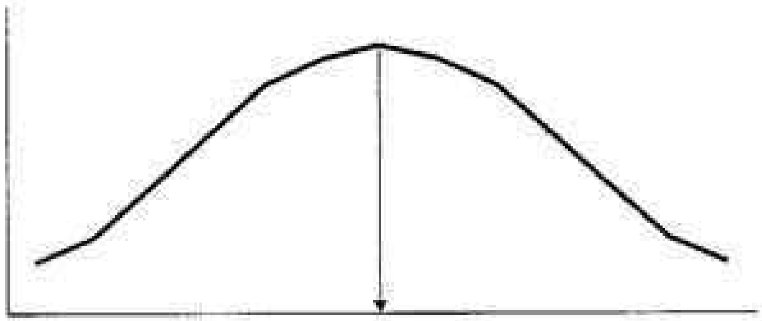
1. في حالة التوزيع المتماثل .

المتوسط الحسابي = الوسيط = المنوال .

ورمزياً : $\overline{X} = \hat{X} = \widetilde{X}$

والشكل الآتي يوضح ذلك :

التوزيع المتماثل



$\widetilde{X} = \hat{X} = \overline{X}$

تدريب (1)

إذا كان لديك توزيع تكراري متماثل وسيطه = 7 ، أوجد المنوال والوسط الحسابي للتوزيع المذكور .

2. في حالة المنحنيات الملتوية التواء بسيطاً أي قريبة من التماثل فإن :

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

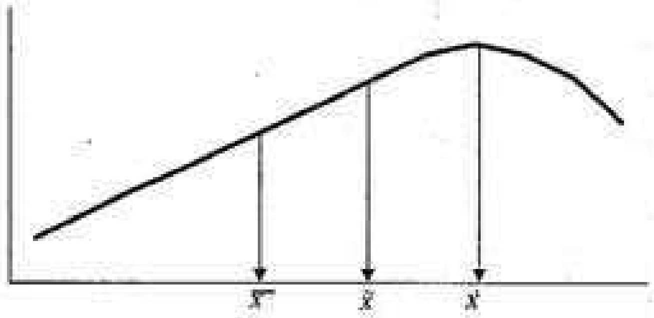
ويُعبر عن ذلك رمزياً بالعلاقة : $\bar{X} - \hat{x} = 3(\bar{x} - \tilde{X})$

ولدينا حالتين هما :

أ. في حالة التوزيع السالب الالتواء يكون المنوال أكبر من كل من الوسيط والوسط الحسابي.

والشكل الآتي يوضح ذلك :

التوزيع التكراري السالب الالتواء

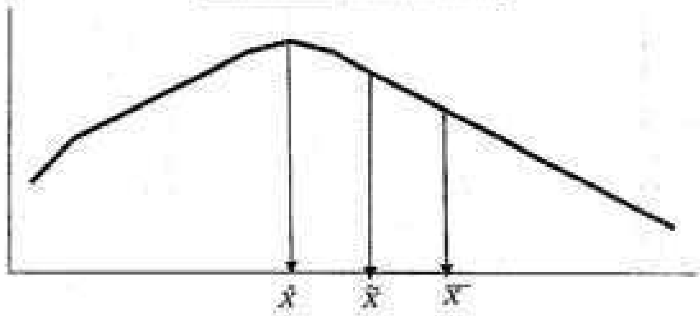


نلاحظ من الشكل أن : $\hat{X} > \tilde{X} > \bar{X}$

ب. في حالة التوزيع الموجب الالتواء يكون المنوال أصغر من كل من الوسيط والوسط الحسابي.

والشكل الآتي يوضح ذلك :

التوزيع التكراري الموجب الالتواء



نلاحظ من الشكل أن : $\hat{X} < \tilde{X} < \bar{X}$

مثال (1)

إذا كان متوسط التوزيع = 5 و وسيطه = 4 احسب المنوال ؟ ثم حدد نوع الالتواء ؟

الحل

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

$$\bar{x} - \hat{x} = 3(\bar{x} - \tilde{x})$$
$$5 - \hat{x} = 3(5 - 4)$$
$$5 - \hat{x} = 3$$
$$\hat{x} = 5 - 3$$
$$\hat{x} = 2$$

إذاً المنوال = 2

وحيث أن المنوال أصغر من الوسيط وأصغر من الوسط الحسابي فيكون الالتواء موجب باتجاه اليمين .

ملحوظات /

1. في التوزيعات المتماثلة التي لها أكثر من منوال فإن الوسط الحسابي أو الوسيط يكون هو المقياس المناسب.
2. عندما يكون التوزيع شديد الالتواء فإن الوسيط غالباً ما يكون أفضل مقاييس النزعة لأنه يقع بين الوسط الحسابي والمنوال.
3. في التوزيعات التي تحتوي قيماً شاذة أو متطرفة فإن الوسيط أو المنوال هما المقياسان المناسبان وإذا كان حجم البيانات كبيرة فإن الوسط الحسابي هو الأسهل لكنه ليس الأفضل .

نشاط (1)

احسب المتوسط الحسابي لجدول تكراري إذا علمت أن :

المنوال = 7 والوسيط = 5 ؟ ثم حدد نوع الالتواء .

الحل: ....

....

....

....

....

❖ مقارنة بين صفات الوسط الحسابي والوسيط والمنوال :

م	الوسط الحسابي	الوسيط	المنوال
1	أكثر مقاييس النزعة المركزية استعمالاً	متوسط الاستعمال	قليل الاستعمال
2	سهولة حسابه	سهولة حسابه	سهولة حسابه.
3	يتأثر بالقيم المتطرفة والشاذة	لا يتأثر بالقيم المتطرفة والشاذة	لا يتأثر بالقيم الشاذة لأنه يعتمد على التكرار
4	نستطيع معرفة مجموع التوزيع إذا علم الوسط الحسابي وعدد المشاهدات أو مجموع التكرارات	لا نستطيع إيجاد المجموع	لا نستطيع إيجاد المجموع
5	مجموع انحرافات القيم عن وسطها الحسابي = صفر	مجموع انحرافات القيم عن وسيطها ليس بالضرورة أن يساوي صفر	مجموع انحرافات القيم عن منوالها ليس بالضرورة أن يساوي صفر
6	يعتمد على جميع المشاهدات	لا يعتمد على جميع المشاهدات	لا يعتمد على جميع المشاهدات
7	يتأثر الوسط الحسابي إذا حذفنا أية مشاهدة من المشاهدات	لا يعتمد على جميع القيم دائماً	لا يعتمد على جميع القيم دائماً
8	لا يعتمد على ترتيب القيم	يعتمد على ترتيب القيم	لا يعتمد على ترتيب القيم
9	يتأثر بالعمليات الحسابية الأربع	يتأثر بالعمليات الحسابية الأربع	يتأثر بالعمليات الحسابية الأربع

رابعاً / الوسط التوافقي (Harmonic Mean).

تعريفه :

هو عبارة عن مقلوب الوسط الحسابي لمقلوب القيم .

ويرمز له بالرمز $\overline{X}_h$

طُرق حسابه .

الحالة الأولى : الوسط التوافقي لبيانات غير مبوبة .

يُحسب من العلاقة الآتية :

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

مثال (1)

اوجد الوسط التوافقي للقيم . 2 ، 4 ، 7 ، 6 ، 3 ، 9

الحل:

$$\overline{X}_h = \frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}} = \frac{6}{\frac{1}{9}+\frac{1}{3}+\frac{1}{6}+\frac{1}{7}+\frac{1}{4}+\frac{1}{2}} = \frac{6}{1.5} = 4$$

نشاط (1) :

إذا كانت أعمار مجموعة من الأطفال المصابين بإحدى الأمراض على النحو الآتي:

2 ، 3 ، 6 ، 1 ، 5 ، 4 ، 2 ، 3 ، 2

احسب الوسط التوافقي لهذه البيانات .

الحل:..

....

....

....

....

....

تدريب (1)

اوجد الوسط التوافقي للبيانات الآتية :

8،5،7،7،10،11،9،6

الحالة الثانية : الوسط التوافقي لبيانات مبوبة .

يحسب الوسط التوافقي للبيانات المبوبة من العلاقة الآتية :

$$\overline{X}_h = \frac{\sum_{i=1}^{n} f_i}{\sum_{i=1}^{n} \left(\frac{f_i}{x_i}\right)}$$

حيث x_i تعبر عن مراكز الفئات .

مثال (2)

من جدول التوزيع التكراري الآتي أوجد الوسط التوافقي :

الفئة	10 - 13	14 - 17	18 - 21	22 - 25	26 - 29	المجموع
التكرار F	3	10	20	15	2	50

الحل

تتطلب العلاقة الرياضية للوسط التوافقي إنشاء عمودين ، نوجد فيهما مراكز الفئات ، وكذا حاصل قسمة التكرارات على مراكز الفئات ، كما يوضحها الجدول الآتي :

الفئات	التكرارات F	مركز الفئة X	F / X
10 – 13	3	11.5	0.261
14 – 17	10	15.5	0.645
18 – 21	20	19.5	1.026
22 – 25	15	23.5	0.638
26 – 29	2	27.5	0.073
المجموع	50		2.643

$$\overline{X}_h = \frac{\sum_{i=1}^{n} f_i}{\sum_{i=1}^{n} \left(\frac{f_i}{x_i}\right)} = \frac{50}{2.643} = 18.92$$

نشاط (2)

من البيانات الواردة في الجدول الآتي اوجد الوسط التوافقي :

الملاحظة X	3 - 7	8 - 12	13 - 17	18 - 22	23 - 27	المجموع
التكرار F	2	3	10	5	4	24

الحل:..

....

....

....

....

....

....

تدريب (2)

من جدول التوزيع التكراري الآتي أوجد الوسط التوافقي.

الملاحظة	2	3	4	5	6	7	8	9
التكرار F	7	9	7	4	4	5	3	2

❖ خواص الوسط التوافقي :

1. يعتمد على جميع الملاحظات .
2. قابل للمعالجات الرياضية .
3. لا يمكن حسابه من الجداول التكرارية مفتوحة الأطراف .
4. لا يمكن حسابه إذا كانت إحدى القيم تساوي صفر .
5. تستحوذ القيم الصغيرة على أوزان أكبر من القيم الكبيرة بسبب المقلوب .

خامساً / الوسط الهندسي (The Geometric Mean) .

تعريفه :

لتكن $x_1, x_2, x_3, \ldots, x_n$ مجموعة من القيم ، فإن وسطها الهندسي هو الجذر النوني لحاصل ضربها.

ويستخدم عندما تكون البيانات في شكل نسبي أو معدلات نمو .

طُرق حسابه :

الحالة الأولى : الوسط الهندسي لبيانات غير مبوبة .

يُحسب من العلاقة الآتية :

$$G = \sqrt[n]{x_1 . x_2 . x_3 \ldots . x_n}$$

مثال (1)

احسب الوسط الهندسي للبيانات الآتية:

50 ، 20 ، 2 ، 5

الحل:

$$G = \sqrt[n]{x_1 . x_2 . x_3 \ldots . x_n} = \sqrt[4]{(50)(20)(2)(5)} = \sqrt[4]{10000} = 10$$

ويُمكن استخدام اللوغاريتمات لطريق العلاقة السابقة عندما يزيد عدد المشاهدات (n) حيث تُستخدم الصيغة الآتية لحساب الوسط الهندسي .

أي أن :

$$\log G = \frac{\sum_{i=1}^{n} \log x_i}{n}$$

$$\log G = \frac{\log x_1 + \log x_2 + \log x_3 + \ldots + \log x_n}{n}$$

من العلاقة السابقة يتضح أن لوغاريتم الوسط الهندسي هو متوسط لوغاريتمات القيم .

يُمكن حساب المثال (1) بالعلاقة الثانية كما يلي :

$$\log G = \frac{1.699 + 1.301 + 0.301 + 0.699}{4} = \frac{4}{4} = 1$$

ولإيجاد قيمة G نأخذ معكوس الدالة اللوغاريتمية (الدالة الأُسية للأساس 10) للطرفين كما يلي:

$10^{\log G} = 10^1 = 10$ وهي نفس القيمة السابقة .

نشاط (1)

إذا كانت البيانات الآتية تمثل أوزان خمسة أطفال لأقرب كيلو جرام :

23 ، 20 ، 25 ، 14 ، 12

أوجد الوسط الهندسي لأوزان الأطفال.

الحل:...

....

....

....

....

تدريب (1)

احسب الوسط الهندسي للبيانات الآتية :

4 ، 7 ، 15 ، 6 ، 5 ، 6

الحالة الثانية : الوسط الهندسي لبيانات مبوبة .

لحساب الوسط الهندسي في حالة البيانات المبوبة نستخدم العلاقة الآتية :

$$G = \sqrt[\sum f]{x_1^{f_1} . x_2^{f_2} . x_3^{f_3} \ldots . x_n^{f_n}}$$

وعند أخذ لوغاريتمات مراكز الفئات نستخدم العلاقة الآتية :

$$\log G = \frac{\sum_{i=1}^{n} f_i \log x_i}{n}$$

مثال (2)

من بيانات الجدول الآتي أوجد الوسط الهندسي :

الفئة	10 - 13	14 - 17	18 - 21	22 - 25	26 - 29	المجموع
التكرار F	3	4	2	3	1	13

الحل:

باستخدام العلاقة :

$$G = \sqrt[\sum f]{x_1^{f_1}.x_2^{f_2}.x_3^{f_3}.....x_n^{f_n}}$$

نكون جدول التوزيع التكراري كما يلي :

الفئات	التكرارات	مركز الفئة	
10 - 13	3	11.5	1520.875
14 - 17	4	15.5	57720.063
18 - 21	2	19.5	380.25
22 - 25	3	23.5	12977.875
26 - 29	1	27.5	27.5
المجموع	13	-	-

$$G = \sqrt[13]{(1520875)(57720063)(380.25)(12977875)(27.5)} = 17.243$$

نشاط (2)

من جدول التوزيع التكراري الآتي أوجد الوسط الهندسي :

الفئة	2 - 4	5 - 7	8 - 10	11 - 13	المجموع
التكرار F	3	4	2	3	13

الحل:...

....

....

....

....

تدريب (2)

إذا كان المتوسط الحسابي لعددين هو 10 وكان وسطهما الهندسي هو 8 أوجد العددين.

◆ من مزايا الوسط الهندسي:

1- يعتمد على جميع القيم.

2- محدد جبرياً بدقة.

3- قليل التأثر بالقيم المتطرفة نحو اليمين بعكس الوسط الحسابي، لذلك إذا كان التوزيع **ملتوٍ نحو اليمين** فإن الوسط الهندسي يكون **أفضل** المقاييس لتمثيل البيانات.

4- يُعد ذا أهمية كبيرة عند إنشاء جداول الأرقام القياسية.

◆ من عيوب الوسط الهندسي:

1- لا يُمكن حسابه للجداول المفتوحة.

2- لا يُمكن حسابه إذا كانت إحدى القيم سالبة أو صفر.

3- هو قيمة مجردة وغير واضحة في بعض الحالات.

4- صعوبة حسابه وخاصة إذا كانت البيانات كبيرة جداً.

- الربيعات والعشيرات والمئينات (Deciles Quartiles and Percentiles) :

التعريف :

1. الربيعات Quartiles : النظام الذي يقسم المساحة تحت المنحنى التكراري أو المضلع التكراري الذي يُمثل أي توزيع تكراري إلى أربعة أقسام متساوية من حيث المساحة يسمى بالربيعات .

وتقسم الربيعات إلى ثلاثة أقسام هي :

(أ) الربيع الأول First Quartile (Q_1): هو القيمة التي يسبقها ربع المشاهدات أو التكرارات ويليها ثلاثة أرباع التكرارات أو المشاهدات .

(ب) الربيع الثاني Second Quartile (Q_2): هو القيمة التي يسبقها نصف المشاهدات أو التكرارات ويليها النصف الأخر.

(ج) الربيع الثالث Third Quartile (Q_3): هو القيمة التي يسبقها ثلاثة أرباع المشاهدات أو التكرارات ويليها الربع الأخير.

ملحوظة : الربيع الثاني = الوسيط

2. العشيرات Deciles: النظام الذي يقسم المساحة تحت المنحنى التكراري أو المضلع التكراري الذي يُمثل أي توزيع تكراري إلى عشرة أقسام متساوية من حيث المساحة يسمى بالعشيرات .

وتُقسم العشيرات إلى تسعة أقسام تبدأ بالعُشير الأول ثم الثاني وتنتهي بالعشير التاسع ويرمز لكل منها بالرمز $D_i(i = 1,2,3,....,9)$ وتعرف كما عرفنا الربيعات ، وستعرف العشير الأول والثاني ونترك للطالب تعريف بقية العشيرات .

(أ) العشير الأول First Decile (D_1): هو القيمة التي يسبقها (0.1)من المشاهدات أو التكرارات ويليها تسعة أعشار من التكرارات أو المشاهدات .

(ب) العشير الثاني Second Decile (D_2): هو القيمة التي يسبقها (0.2) من المشاهدات أو التكرارات ويليها ثمانية أعشار من التكرارات أو المشاهدات.

ملحوظة : العشير الخامس = الربيع الثاني = الوسيط

3.المئينات Percentiles: النظام الذي يقسم المساحة تحت المنحنى التكراري أو المضلع التكراري الذي يمثل أي توزيع تكراري إلى مائة قسم متساوية من حيث المساحة يُسمى بالمئينات .

وتقسم المئينات إلى تسعة وتسعين قسم تبدأ بالمئين الأول والثاني وتنتهي بالمئين التاسع والتسعين ويُرمز لكل منها بالرمز $P_i(i = 1,2,3,....,99)$ وتعرف كما عرفنا الربيعات والعشيرات ، وسنعرف المئين الأول ونترك للطالب تعريف بقية المئينات .

(أ) المئين الأول (P_1) First Percentile: هو القيمة التي يسبقها واحد من مائة من المشاهدات أو التكرارات ويليها تسعة وتسعون من مائة من التكرارات أو المشاهدات .

ملحوظة : المئين الخمسين = الوسيط = الربيع الثاني = العشير الخامس

إذا كانت مقاييس النزعة المركزية تستهدف قياس أو تحديد المركز الذي تتمحور حوله البيانات وأن جميعها تتمثل بقيمة مفردة واحدة تكون ممثلة لبيانات العينة ، فإن هناك مقاييس أخرى تشبه الوسيط في أنها مقاييس للموقع في طريقة حسابها ، ولكنها ليست من المتوسطات ، وهذه المقاييس هي الربيعات التي تقسم المفردات إلى أربعة أجزاء كل منها يشمل على عدد متساو من المفردات ويرمز له بالرمز $Q_i(i = 1,2,3)$ ويلاحظ أننا لا نحتاج لحساب الربيع الرابع لأنها ستكون مكملة للربيعات Q_1, Q_2, Q_3 .

كذلك يمكن تقسيم البيانات إلى عشرة أقسام متساوية تسمى بالعشيرات ويرمز لها بالرمز $D_i(i = 1,2,3,...,9)$.

كما يمكن تقسيم البيانات إلى مئة قسم متساو تسمى بالمئينات ويرمز لها بالرمز $P_i(i = 1,2,3,....,99)$ ولحساب هذه المقاييس لدينا الحالات الآتية :

الحالة الأولى : في حالة البيانات غير المبوبة .

نقوم بتحديد موقع الربيع أو العشير أو المئين من العلاقات الآتية :

موقع الربيع $$\frac{i(n+1)}{4}, i = 1,2,3$$

موقع العشير $$\frac{i(n+1)}{10}, i = 1,2,3,...,9$$

موقع المئين $$\frac{i(n+1)}{100}, i = 1,2,3,....,99$$

مثال (1)

للبيانات التالية 6، 18، 15، 9، 11، 8، 14 احسب $Q_1 , Q_3 , \tilde{X} , D_5 , P_{30}$

الحل:

نرتب البيانات تصاعدياً 18، 15، 14، 11، 9، 8، 6

أولاً : حساب Q_1

بما أن عدد المفردات فردياً نقوم بما يلي :

- تحديد موقع الربيع الأول من العلاقة : $\frac{i(n+1)}{4} = \frac{1(7+1)}{4} = \frac{8}{4} = 2$

إذاً العنصر الثاني يمثل الربيع الأول أي : $Q_1 = 8$

ثانياً : حساب Q_3

- تحدد موقع الربيع الثالث من العلاقة : $\frac{i(n+1)}{4} = \frac{3(7+1)}{4} = \frac{24}{4} = 6$

فيكون العنصر الذي ترتيبه 6 يمثل الربيع الثالث أي أن $Q_3 = 15$

ثالثاً : حساب الوسيط $\tilde{X}$ ...... وهو القيمة التي تتوسط البيانات بعد ترتيبها تصاعدياً أو تنازلياً :

وهو العنصر الرابع $\tilde{X} = 11$

رابعاً : حساب العشير الخامس D_5

تحدد موقع العشير الخامس من العلاقة : $\frac{i(n+1)}{10} = \frac{5(7+1)}{10} = \frac{40}{10} = 4$

فيكون العنصر الرابع يمثل العشير الخامس أي أن $D_5 = 11$

نلاحظ أن قيمة العشير الخامس = قيمة الوسيط

خامساً : حساب قيمة P_{30}

تحدد قيمة المئين الثلاثين من العلاقة : $\frac{i(n+1)}{100} = \frac{30(7+1)}{100} = \frac{240}{100} = 2.4$

وبعد الحصول على الموقع 2.4 نجد أن قيمة المئين الثلاثين هي العنصر الثاني مضافاً إليه حاصل ضرب 0.4 في حاصل الفرق بين القيمتين الثالثة والثانية أي أن :

$p_{30} = 8 + (9-8)\times 0.4 = 8.4$

نشاط (1)

للبيانات الآتية 9، 4، 12، 7، 2، 25، 13، 22، 18 احسب

$Q_1 , Q_3 , \tilde{X} , D_5 , P_{30} , D_7 , P_{50}$

ثم قارن بين كل من P_{50} ، D_5 ، $\tilde{X}$ ماذا تلاحظ ؟

الحل:..

....

....

....

تدريب (1)

للبيانات الآتية 11، 4، 12، 7، 25، 13، 22، 8 احسب

$Q_2 , \tilde{X} , D_5 , P_{50}$

ثم قارن بين النواتج التي حصلت عليها ماذا تلاحظ ؟

ملحوظة :

المئين الخامس والعشرين = الربيع الأول .

المئين الخامس والسبعين = الربيع الثالث .

الحالة الثانية : في حالة البيانات المبوبة .

في حالة البيانات المبوبة سوف نحدد موقع كل مقياس بنفس الطريقة للبيانات غير المبوبة ، إلا أننا سوف نستخدم $\sum f_i$ بدلاً من n وتحسب الربيعات والعشيرات والمئينات من العلاقات الآتية :

1. الربيعات تحسب من العلاقة الآتية :

$$Q_i = L + \frac{\frac{(i)\sum f_i}{4} - f_1}{f_2}.C \quad , i = 1,2,3$$

2. العشيرات تحسب من العلاقة الآتية :

$$D_i = L + \frac{\frac{(i)\sum f_i}{10} - f_1}{f_2}.C \quad , i = 1,2,3,...,9$$

3. المئينات تحسب من العلاقة الآتية :

$$P_i = L + \frac{\frac{(i)\sum f_i}{100} - f_1}{f_2}.C \quad , i = 1,2,3,...,99$$

حيث :

L : الحد الأدنى للفئة الربيعية أو العشيرية أو المئينية .

f_1 : التكرار المتجمع الصاعد للفئة السابقة لموقع المقياس .

f_2 : تكرار الفئة الربيعية أو العشيرية أو المئينية .

C : طول الفئة الربيعية أو العشيرية أو المئينية .

مثال (2)

من بيانات الجدول الآتي:

الفئات	التكرارات f	التكرار المتجمع الصاعد ت.م.ص
80 – 109	26	26
110 – 139	78	104
140 – 169	122	226
170 – 199	34	260
200 – 229	14	274
230 – 259	8	282

أوجد :

1. العشير الأول والثاني .
2. الربيع الأول والثالث .
3. المئين الخامس والسبعين.
4. قارن بين قيمة الربيع الثالث والمئين الخامس والسبعين ماذا تلاحظ ؟

الحل:

1. نحدد موقع D_1 , D_2

موقع العشير الأول $\frac{1(282)}{10} = 28.2$

موقع العشير الثاني $\frac{2(282)}{10} = 56.4$

فتكون قيمة العشير الأول $D_1 = 110 + \frac{28.2 - 26}{78} \times 30 = 110.846$

وقيمة العشير الثاني $D_2 = 110 + \frac{56.4 - 26}{78} \times 30 = 121.69$

2. نحدد موقع الربيع الأول $\frac{1(282)}{4} = 70.5$

نحدد موقع الربيع الثالث $\frac{3(282)}{4} = 211.5$

فتكون قيمة الربيع الأول $Q_1 = 110 + \frac{70.5 - 26}{78} \times 30 = 127.115$

وقيمة الربيع الثالث : $Q_3 = 140 + \frac{211.5 - 104}{122} \times 30 = 166.434$

3. نحدد موقع المئين الخامس والسبعين (P_{75}) : $\frac{75(282)}{100} = 211.5$

قيمة المئين الخامس والسبعين هي : $P_{75} = 140 + \frac{211.5 - 104}{122} \times 30 = 166.434$

4. نلاحظ أن قيمة الربيع الثالث = قيمة المئين الخامس والسبعين.

ملحوظة :

المدى الربيعي = الربيع الثالث – الربيع الأول

المدى الربيعي = المئين75 – المئين25

نشاط (2)

من البيانات الواردة في جدول التوزيع التكراري الآتي :

الفئة	التكرار	التكرار المتجمع الصاعد ت.م.ص
15 – 19	6	
20 – 24	8	
25 – 29	2	
30 – 34	9	
35 – 39	5	
40 – 44	4	

1. أكمل الجدول .
2. احسب العشير الثامن ، والربيع الثالث ، والمئين الثمانون.

الحل:..

....

....

تدريب (2)

من البيانات الواردة في جدول التوزيع التكراري الآتي أوجد :

الملاحظة	7	5	11	4	6
التكرار	25	13	18	9	12

1. الربيع الأول والمئين الخامس والعشرين، وقارن النتيجة التي حصلت عليها.
2. الربيع الثالث والمئين الخامس والسبعين ، وقارن النتيجة التي حصلت عليها.
3. هل يمكن إيجاد الربيعات والعشيرات والمئينات بيانياً، وضح ذلك بمثال .

تمارين الفصل الرابع

س1 / ضع دائرة حول رمز الإجابة الصحيحة فيما يلي :

1. الوسيط لمجموعة القيم 30, 38, 35, 44, 40, 35 هو:

أ. 3.5 ب. 100 ج. 150 د. خلاف ذلك وهو:

2. إذا كان مجموع انحرافات 100 مفردة عن وسط فرضي قدره 110 يلغ 400- فإن الوسط الحسابي هو:

أ. 4- ب. 4 ج. 115 د. خلاف ذلك وهو:

3. مجموع انحرافات 50 مفردة عن وسط حسابي قدره 30 هو:

أ. 100 ب. 30 ج. 3 د. خلاف ذلك وهو:

4. إذا كان الوسط الحسابي لعلامات 8 طلاب هو 60، والوسط الحسابي لعلامات 5 طلاب آخرين هو 70، فإن الوسط الحسابي لعلامات جميع الطلاب هو:

أ. 5.5 ب. 8.87 ج. 65 د. خلاف ذلك وهو:

5. الوسط التوافقي للقيم التالية (3، 5، 7، 6، 6، 12، 10) يساوي:

أ. 5 ب. 8.57 ج. 7.58 د. لا شيء مما سبق.

6. المنوال للبيانات التالية (7، 5، 5، 5، 3، 3، 2، 7، 7، 6، 7، 6) يساوي:

أ. لدينا منوالان هما 7 و5 ب. عديمة المنوال ج. 5 د. 6

7. من عيوب الوسط الهندسي ما يلي عدا:

أ. يعتمد على جميع القيم.

ب. لا يمكن حسابه وخاصة إذا كانت البيانات كبيرة جداً.

ج. صعوبة حسابه وخاصة إذا كانت البيانات كبيرة جداً.

د. لا يمكن حسابه إذا كانت إحدى القيم سالبة أو صفر.

8. الوسيط للبيانات التالية (2، 7، 4، 6، 5، 4، 9، 10) يساوي:

أ. 5.5 ب. 6 ج. 5 د. 5.875

9. إذا كان المتوسط الحسابي = الوسيط = المنوال فإن منحنى التوزيع يكون:

أ. متماثل ب. ملتو نحو اليمين ج. ملتو نحو اليسار د. موجب الالتواء

10. يحسب الوسط الحسابي بطريقة الانحرافات المختصرة من العلاقة:

أ. $\bar{x} = L + \frac{\sum fd_m}{\sum f}.C$ ب. $\bar{x} = L + \frac{\sum fd}{\sum f}$ ج. $\bar{x} = \frac{\sum x}{n}$ د. لا شيء مما سبق

11. المئين الثلاثين (P_{30}) للبيانات الآتية (6، 18، 15، 9، 11، 8، 14) يساوي:

أ. 8.4 ب. 15 ج. 4 د. 11

س2 / البيانات التالية تمثل تصنيف عدد من مرضى ضغط الدم المرتفع حسب مستوى الكوليسترول في الدم :

مستوى الكوليسترول	عدد المرضى
195 -	1
200 -	3
205 -	4
210 -	7
215 -	4
220 – 225	1

المطلوب :

- اوجد المتوسط الحسابي والوسيط والمنوال جبرياً .
- اوجد الوسيط بيانياً .
- اوجد المنوال بيانياً .
- اوجد $Q_1, Q_2, P_{30}, D_3, \bar{x}, D_7, P_{90}$

س 3 / عرف كلاً من :

- المتوسط الحسابي .
- الوسيط .
- المنوال .
- الوسط الهندسي .
- الوسط التوافقي .

س4 / اذكر مزايا وعيوب كل مقياس من مقاييس النزعة المركزية ؟

س5 / قارن بين خصائص كل من المتوسط الحسابي والوسيط والمنوال ؟

س6/ البيانات الآتية تمثل أوزان عينة من الأطفال مقاسة لأقرب كيلو جرام بعد سنة من الولادة:

10، 7، 10 ، 8 ، 9 ، 10، 11، 7 ، 9، 8

المطلوب:

حساب المتوسط والوسيط والمنوال لأوزان هؤلاء الأطفال.

4

الفصل الرابع

مقاييس التشتت

محتويات الفصل الرابع

أهداف الفصل الرابع

يتوقع من الطالب بعد الانتهاء من دراسة هذا الفصل أن يكون قادراً على:

- ✓ توضيح مفهوم التشتت.
- ✓ التمييز بين مقاييس النزعة المركزية ومقاييس التشتت.
- ✓ معرفة صفات المقياس الجيد للتشتت .
- ✓ معرفة الاستخدامات الرئيسية للتشتت .
- ✓ تعريف كل مقياس من مقاييس التشتت .
- ✓ حساب مقاييس التشتت بطرق صحيحة .
- ✓ معرفة خواص كل مقياس من مقاييس التشتت .
- ✓ حساب وتفسير معامل الاختلاف والقيمة المعيارية .
- ✓ وصف التماثل والالتواء والتفرطح للتوزيعات التكرارية بعد تمثيلها بيانياً.

مقدمة Introduction:

رغم أهمية مقاييس النزعة المركزية (المتوسط والوسيط والمنوال) لإعطاء وصف أو مؤشر للظواهر أو البيانات من خلال نقطة واحدة تتمركز حولها القيم تعبر عن تلك المشاهدات أو المفردات ضمن مجموعة من القيم إلا أن هذه المقاييس وحدها لا تكفي؛ فمثلاً إذا كان لدينا مجموعتان من القيم:

قيم المجموعة الأولى هي: 43   80   121   80

قيم المجموعة الثانية هي: 77   80   87   80

نلاحظ أن المجموعتين لهما نفس المتوسط الحسابي وهو (81) وكذلك لهما الوسيط نفسه وهو (80) ولهما المنوال نفسه هو (80).

وعلى الرغم من ذلك فلا نستطيع القول بأن المجموعتين متكافئتان فقيم المجموعة الأولى متباعدة، بينما قيم المجموعة الثانية متقاربة، وفي هذه الحالة فإن مقاييس النزعة المركزية لا تعطي وصفاً دقيقاً، لذا لابد من استخدام مقاييس أخرى تسمى مقاييس التشتت .

◆ مفهوم التشتت The Variance Concept .

يُقصد بالتشتت مدى تباعد أو تبعثر القيم في التوزيع عن بعضها البعض .

أي أنه كلما تباعدت القيم عن بعضها يكون التوزيع أكثر تشتتاً ، وكلما تقاربت من بعضها يكون التوزيع أقل تشتتاً . وتوجد مقاييس عدة للتشتت منها :

- المدى The Range.
- المدى الربيعي quartile Range .
- الانحراف المطلق (الانحراف المتوسط) Mean Deviation.
- التباين The Variance .
- الانحراف المعياري Standard Deviation.

◆ المقياس الجيد للتشتت يتصف بما يلي :

- يجب أن يكون معتمداً على جميع البيانات.
- يجب أن تكون وحدات قياسه هي نفسها وحدات قياس المفردات.
- يجب أن يكون معرفاً بوضوح تام.
- يجب أن يتوافق مع القواعد الرياضية.
- يجب أن لا يخضع إلى حسابات معقدة ومملة.

* الاستخدامات الرئيسية للتشتت يمكن تلخيصها فيما يلي:

- يعطي فكرة عن مصداقية قيمة مركزية ما .
- يجعل المقارنة بين مجموعتين من البيانات ممكنة، أخذاً في الاعتبار مقدار تغيرهما.
- يعمل على تزويد المختص أو الباحث بالأساس (القاعدة) التي بواسطتها يمكن التحكم بالتغير.
- له استخدامات واسعة في المجالات التطبيقية في كافة حقول المعرفة.

* مقاييس التشتت .

سنتناول في دراستنا المقاييس التالية للتشتت .

أولاً : المدى (Range)

تعريفه :

يُعد المدى من أبسط مقاييس التشتت من حيث المفهوم، وهو عبارة عن الفرق بين أعلى قيمة (مفردة) وأصغر قيمة (مفردة) في التوزيع ، ويرمز له بالرمز(R) .

كلما كان المدى صغيراً كلما كان دليلاً على تجانس البيانات أي انخفاض التشتت .

طُرق حسابه :

الحالة الأولى: المدى لبيانات غير مبوبة.

المدى (R) = أعلى قيمة في البيانات – أقل قيمة في البيانات.

مثال (1)

أوجد المدى للبيانات الآتية : 2، 4، 7، 8، 9، 19

الحل:

المدى (R) = أعلى قيمة – أقل قيمة

$R= 19 - 2 = 17$ (المدى)

نشاط (1)

إذا كانت البيانات الآتية تُمثل أوزان مجموعة من الأشخاص أوجد المدى لهذه البيانات :

65 ، 98 ، 11 ، 65 ، 47 ، 98 ، 65 ، 48 ، 12 ، 15

الحل:..

تدريب (1)

إذا كانت أعمار مجموعة من الأطفال (لأقرب سنة) كما يلي:

11 ، 4 ، 8 ، 2 ، 7 ، 6 ، 4 ، 9 ، 5

أوجد المدى لهذه البيانات ؟

الحالة الثانية : المدى لبيانات مبوبة بدون فئات .

المدى (R) = أعلى قيمة – أقل قيمة

مثال (2)

من جدول التوزيع التكراري التالي أوجد المدى :

التكرار F	المشاهدة X
9	3
5	5
12	7
6	9
7	11
3	13
4	15
	المجموع

الحل

المدى (R) = أعلى قيمة – أقل قيمة

$R = 15 - 3 = 12$ (المدى)

نشاط (2)

من البيانات الواردة في جدول التوزيع التكراري الآتي أوجد المدى :

الملاحظة	10	9	8	7	6	5	المجموع
التكرار	5	4	8	6	5	3	

الحل:..

تدريب (2)

من جدول التوزيع التكراري الآتي أوجد المدى :

الملاحظة	10	20	30	40	المجموع
التكرار	3	12	8	6	

الحالة الثالثة : المدى لبيانات مبوبة بفئات .

وفي حالة الجداول التكرارية بفئات :

المدى = الحد الأعلى لآخر فئة – الحد الأدنى لأول فئة

مثال (3)

من جدول التوزيع التكراري الآتي أوجد المدى :

الفئة	6 - 8	9 - 11	12 - 14	15 - 17	المجموع
التكرار F	12	17	15	10	

الحل:

$R = 17 - 6 = 11$ (المدى)

نشاط (3)

من البيانات الواردة في جدول التوزيع التكراري الآتي أوجد المدى :

الفئة	5 - 9	10 - 14	15 - 19	20 - 24	المجموع
التكرار	8	13	15	6	

الحل: ....

....

تدريب (3)

أوجد المدى للبيانات الآتية:

الفئة	5 -9	10 -14	15 -19	20 -24	25 -29	30 -34	المجموع
التكرار	8	13	15	11	9	6	

◆ خصائص المدى:

1- يتأثر بالقيم الحدية فقط ولا يأخذ بالحسبان جميع القيم، لذلك لا يكشف عن انتشار البيانات الواقعة بين القيم الحدية.

2- لا يمكن حساب المدى بدقة في التوزيعات التكرارية المفتوحة.

3- سهولة حسابه مقارنة بمقاييس التشتت الأخرى.

4- لا يمكن حسابه لبيانات نوعية.

5- لا يعتمد في حسابه على جميع القيم، وبالتالي يُعد مقياس غير جيد للتشتت.

◆ المدى الربيعي Inter quartile Range و نصف المدى الربيعي

في حالة وجود قيم متطرفة في البيانات، أو إذا كانت البيانات توحي بوجود التواء شديد، فيكون مناسباً استخدام المدى الربيعي، كما انه إذا كانت البيانات مبوبة في جدول تكراري مفتوح، فإنه لا مناص من استخدام المدى الربيعي.

تعريف المدى الربيعي :

يعرف المدى الربيعي بأنه الفرق بين الربيع الأعلى Q_3 (الربيع الثالث) والربيع الأدنى Q_1 (الربيع الأول) ، أي ان : المدى الربيعي = $Q_3 - Q_1$.

تعريف نصف المدى الربيعي :

هو عبارة عن المدى الربيعي مقسوماً على 2 ، أي ان : نصف المدى الربيعي = $\frac{Q_3 - Q_1}{2}$

طرق حسابه:

الحالة الأولى: حسابه لبيانات غير مبوبة

مثال (1)

احسب المدى الربيعي ونصف المدى الربيعي للبيانات الآتية:

10، 80، 60، 5، 100، 30، 15

الحل

من تعريف المدى الربيعي يتضح أن حسابه يتطلب إيجاد كل من الربيع الأعلى والربيع الأدنى، وحساب الربيعات يتطلب أولاً ترتيب البيانات تصاعدياً.

أولاً: نرتب البيانات تصاعدياً كما يلي:

5، 10، 15، 30، 60، 80، 100

ثانياً: نوجد قيمة الربيع الأدنى Q_1

ترتيب الربيع الأدنى: $O=\frac{n+1}{4}=\frac{7+1}{4}=\frac{8}{4}=2$

أي أن قيمة الربيع الأدنى تساوي المفردة التي ترتيبها 2

إذاً قيمة الربيع الأدنى $Q_1=10$

ثالثاً: نوجد قيمة الربيع الأعلى Q_3

ترتيب الربيع الأعلى: $O=\frac{3(n+1)}{4}=\frac{3(7+1)}{4}=\frac{3(8)}{4}=\frac{24}{4}=6$

أي أن قيمة الربيع الأعلى يساوي المفردة التي ترتيبها 6

إذاً قيمة الربيع الأعلى $Q_3=80$

وبناءً عليه فإن المدى الربيعي يساوي: $Q_3-Q_1=80-10=70$ = المدى الربيعي

كذلك $\frac{Q_3-Q_1}{2}=\frac{80-10}{2}=\frac{70}{2}=35$ = نصف المدى الربيعي

نشاط (1)

إذا كانت البيانات الآتية تمثل أجور الموظفين في إحدى المراكز الصحية بآلاف الريالات:

47، 36، 42، 50، 45، 38، 37، 40، 35، 32، 30، 27، 24، 25

احسب المدى الربيعي ونصف المدى الربيعي لهذه البيانات.

الحل: ....

....

تدريب (1)

إذا كانت البيانات الآتية تمثل أعمار عدد من المرضى:

40، 65، 63، 62، 60، 42، 52، 55، 50، 45، 40، 37، 35، 30

احسب المدى الربيعي ونصف المدى الربيعي لهذه البيانات.

الحالة الثانية: حسابه لبيانات مبوبة.

خطوات حسابه:

1. تحدد رتبتي الربيع الأول والثالث.

ترتيب الربيع الأول = $\frac{n}{4}$، وترتيب الربيع الثالث = $\frac{3n}{4}$، حيث n عدد المفردات.

نحسب قيمة الربيعين الأول والثالث بنفس طريقة حساب الوسيط من التوزيعات التكرارية باستخدام المعادلة الآتية:

$$Q_i = L + \frac{\frac{(i)\sum f_i}{4} - f_1}{f_2}.C \quad ,i=1,2,3$$

مثال (2)

احسب المدى الربيعي ونصف المدى الربيعي للبيانات الواردة في الجدول الآتي :

الفئات	التكرار f	التكرار المتجمع الصاعد
10 -	4	4
15 -	6	10
20 -	12	22
25 -	20	42
30 -	18	60
35 -	14	74
40 -	14	88
45 - 50	2	90
المجموع		

الحل

ترتيب الربيع الأول = $\frac{100}{4} = 25$

ترتيب الربيع الثالث = $\frac{100(3)}{4} = 75$

قيمة الربيع الأول : $Q_1 = 25 + \frac{25-22}{20} \times 5 = 25.75$

قيمة الربيع الثالث : $Q_3 = 40 + \frac{75-74}{14} \times 5 = 40.36$

وبناءً عليه فإن المدى الربيعي يساوي : المدى الربيعي = $Q_3 - Q_1 = 40.36 - 25.75 = 14.61$

كذلك : نصف المدى الربيعي = $\frac{Q_3 - Q_1}{2} = \frac{40.36 - 25.75}{2} = \frac{14.61}{2} = 7.305$

نشاط (2)

من جدول التوزيع التكراري الآتي أوجد المدى الربيعي ونصف المدى الربيعي ؟

الملاحظة	4	9	6	4	7	5	المجموع
التكرار	7	3	6	5	4	1	

الحل...

....

....

....

....

تدريب (2)

من جدول التوزيع التكراري التالي أوجد المدى الربيعي ونصف المدى الربيعي ؟

الفئة	6 -10	11 -15	16 -20	21 -25	26 -30	31 -35	36 -40
التكرار	6	4	11	12	15	9	7

مميزات نصف المدى الربيعي :

أ. لا يأخذ القيم المتطرفة في الحسبان.

ب. إمكانية حسابه من الجداول التكرارية المفتوحة .

عيوب نصف المدى الربيعي :

أ. الاكتفاء به لقياس التشتت غير دقيق ، لأنه يقتصر على نصف المشاهدات وهي التي تقع بين الربيع الأدنى والربيع الأعلى.

ب. صعوبة التعامل معه في التحليل الإحصائي المتقدم.

• **ثانياً: الانحراف المتوسط Mean Deviation**

تعريفه:

يُعرف الانحراف المتوسط بأنه متوسط الانحرافات المطلقة للمشاهدات عن وسطها الحسابي، ويسمى كذلك الانحراف المطلق، ويرمز له بالرمز (M . D).

طرق حسابه:

الحالة الأولى: حساب الانحراف المتوسط لبيانات غير مبوبة.

يحسب من العلاقة الآتية:

$$M.D = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n}$$

حيث: n عدد القيم (الملاحظات).

$\bar{x}$: المتوسط الحسابي للقيم.

x_i: القيم (الملاحظات).

مثال (1)

إذا كانت البيانات الآتية تُمثل أعمار 5 أفراد مصابين بمرض ما:

40، 45، 50، 55، 60 فأوجد الانحراف المتوسط لهذه الأعمار؟

الحل:

أولاً / نوجد المتوسط الحسابي لهذه البيانات من العلاقة الآتية:

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{40 + 45 + 50 + 55 + 60}{5} = \frac{250}{5} = 50$$

ثانياً / نحسب الانحراف المتوسط من العلاقة التالية:

$$M.D = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n}$$

$$= \frac{|40-50| + |45-50| + |50-50| + |55-50| + |60-50|}{5}$$

$$= \frac{10+5+0+5+10}{5} = \frac{30}{5} = 6$$

نشاط (1)

اوجد الإنحراف المطلق (المتوسط) للبيانات الآتية:

3، 4، 8، 7، 9، 6، 4

الحل: ....

....

....

....

....

الحالة الثانية: حساب الانحراف المتوسط لبيانات مبوبة بدون فئات.

يحسب من العلاقة الآتية:

$$M.D = \frac{\sum_{i=1}^{n} |x_i - \bar{x}| f_i}{n}$$

حيث: n عدد القيم (الملاحظات).

$\bar{x}$: المتوسط الحسابي للقيم.

x_i: القيم (الملاحظات).

f_i: التكرار.

مثال (2)

احسب الانحراف المتوسط للتوزيع التكراري في الجدول الآتي:

الملاحظة x	2	5	8	11	المجموع
التكرار f	5	3	4	3	

الحل:

نرتب البيانات كما في الجدول الآتي:

x_i	f	x . f			
2	5	10	-4	4	20
5	3	15	-1	1	3
8	4	32	2	2	8
11	3	33	5	5	15
المجموع	15	90			46

أولاً / نحسب المتوسط الحسابي

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i f_i}{\sum_{i=1}^{n} f_i} = \frac{90}{15} = 6$$

ثانياً / نحسب الانحراف المتوسط من العلاقة.

$$M.D = \frac{\sum_{i=1}^{n} |x_i - \overline{x}| f_i}{n} = \frac{46}{15} = 3.07$$

نشاط (2)

من جدول التوزيع التكراري الآتي أوجد الانحراف المتوسط

الملاحظة	4	5	8	7	6	المجموع
التكرار	8	4	6	15	11	

الحل: ....

....

تدريب (2)

من جدول التوزيع التكراري الآتي احسب الانحراف المتوسط :

الملاحظة	2	4	8	5	6	المجموع
التكرار	3	4	6	3	2	

الحالة الثالثة : حساب الانحراف المتوسط لبيانات مبوبة بفئات .

يحسب من العلاقة الآتية :

$$M.D = \frac{\sum_{i=1}^{n} |x_i - \overline{x}| f_i}{n}$$

حيث : n عدد القيم (مجموع التكرارات).

$\overline{X}$: المتوسط الحسابي .

x_i : مراكز الفئات .

f_i : التكرار.

مثال (3)

من جدول التوزيع التكراري الآتي احسب الانحراف المتوسط :

الفئة	1 -30	31 -60	61 -90	91 -120	121 -150	151 -180	المجموع
التكرار	3	9	20	22	13	8	75

الحل:

نكون الجدول التالي

الفئة	التكرار f	مركز الفئة x	x . f			
1 -30	3	15.5	46.5	-82.8	82.8	248.4
31 -60	9	45.5	409.5	-52.8	52.8	475.2
61 -90	20	75.5	1510	-22.8	22.8	456
91 -120	22	105.5	2321	7.2	7.2	158.4
121 -150	13	135.5	1761.5	37.2	37.2	483.6
151 -180	8	165.5	1324	67.2	67.2	537.6
المجموع	75		7372.5			2359.2

أولاً / توجد المتوسط الحسابي

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i f_i}{\sum_{i=1}^{n} f_i} = \frac{7372.5}{75} = 98.3$$

ثانياً / تحسب الانحراف المتوسط من العلاقة التالية

$$M.D = \frac{\sum_{i=1}^{n} \left| x_i - \overline{x} \right| f_i}{n} = \frac{2359.2}{75} = 31.5$$

نشاط (3)

من جدول التوزيع التكراري الآتي احسب الانحراف المتوسط :

الفئة	6 -10	11 -15	16 -20	21 -25	26 -30	المجموع
التكرار	2	6	9	5	3	25

الحل:...

....

....

....

....

تدريب (3)

من جدول التوزيع التكراري الآتي احسب الانحراف المتوسط :

الفئة	5 -9	10 -14	15 -19	20 -24	25 -29	30 -34	المجموع
التكرار	8	13	15	11	9	6	62

- خواص الانحراف المتوسط:

1- إهمال الإشارة الجبرية عند حساب الانحرافات.

2- يعتبر أكثر شمولاً من المدى لأنه يأخذ جميع قيم التوزيع في حساب التشتت .

3- من الممكن حساب متوسط الانحراف بدلالة الوسيط بدلاً من الوسط الحسابي.

4- مفهوم وسطي يأخذ قيم كافة المفردات بعين الاعتبار عند حسابه.

5- لا يمكن حسابه لبيانات نوعية.

6- لا يمكن حسابه من جداول التوزيع التكراري المفتوحة.

ثالثاً : التباين والانحراف المعياري Variance and Standard Deviation

1. التباين (The Variance) :

إذا كانت نقطة الضعف الأساسية في الانحراف المتوسط هي أنه يُهمل الإشارة الجبرية ، ويأخذ كل الانحرافات على أنها موجبة ، وللتغلب على هذا العيب نقوم بتربيع انحرافات القيم عن وسطها الحسابي ، وبهذا ستتوصل لقياس آخر من مقاييس التشتت ويسمى بالتباين .

تعريفه :

متوسط مجموع مربعات انحرافات القيم عن وسطها الحسابي ويرمز له بالرمز (S^2) ، أما في حالة حسابه للمجتمع فترمز له بالرمز (σ^2) وينطق سيجما تربيع .

طُرق حسابه :

الحالة الأولى : حساب التباين لبيانات غير مبوبة.

يحسب من العلاقة الآتية :

$$\sigma^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n} \qquad \ldots\ldots\ldots (1)$$

حيث : n عدد القيم ،

$\bar{X}$: المتوسط الحسابي .

X_i : القيم (الملاحظات)

ونستخدم العلاقة (1) عند أخذ المجتمع كاملاً .

أما عند أخذ عينة أقل من 30 فيتم طرح 1 من مجموع العينة (n) للتقليل من خطأ المعاينة الذي قد يقع فيه الباحث أثناء اختيار العينة وتصبح العلاقة كما يلي:

$$S^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1} \qquad \ldots\ldots\ldots (2)$$

وعندما تكون العينة n أكبر من أو تساوي 30 فإنها تأخذ حكم المجتمع أي أننا نقسم في هذه الحالة على n بدلاً من n-1 .

مثال (1)

إذا كانت البيانات الآتية تمثل أوزان مجموعة من الأطفال لأقرب كجم ، احسب التباين لهذه الأوزان.

2 ، 7 ، 9 ، 4 ، 8 ، 5 ، 7 ، 6 ، 8 ، 4

الحل:

- نوجد أولاً المتوسط الحسابي:

$$\bar{X} = \frac{\sum x}{n} = \frac{4+8+6+7+5+8+4+9+7+2}{10} = \frac{60}{10} = 6$$

- نوجد فروق البيانات عن المتوسط الحسابي ثم نربعها (الفروق) ونجمعها ونقسم على (n-1) فنحصل على التباين (S^2) ، أي نستخدم العلاقة الآتية

$$S^2 = \frac{\sum_{i=1}^{n} (x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2}{n-1}$$

$$= \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1} = \frac{44}{9} = 4.89$$

نشاط (1)

أوجد التباين للبيانات الآتية :

30 ، 25 ، 20 ، 15 ، 10 ، 5

الحل:...

....

....

....

تدريب (1)

أوجد التباين للبيانات الآتية:

23 ، 16 ، 8 ، 12 ، 4 ، 2 ، 6 ، 3

ثانياً : حساب التباين لبيانات مبوبة بدون فئات :

يحسب التباين في حالة البيانات المبوبة بدون فئات من العلاقة التالية :

$$S^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2 f_i}{n-1}$$

حيث : n عدد القيم (مجموع التكرارات)

$\overline{x}$: المتوسط الحسابي .

x_i : القيم (الملاحظات) .

f_i : التكرار.

مثال (2)

من جدول التوزيع التكراري الآتي أوجد التباين :

الملاحظة	4	5	8	7	6	المجموع
التكرار	8	4	6	15	11	

الحل:

نكون جدول توزيع تكراري كما يلي :

الملاحظة	التكرار				
4	8	32	-2	4	32
5	4	20	-1	1	4
8	6	48	2	4	24
7	15	105	1	1	15
6	11	66	0	0	0
المجموع	44	271			75

أولاً / تحسب المتوسط الحسابي للبيانات في الجدول :

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i f_i}{\sum_{i=1}^{n} f_i} = \frac{271}{44} = 6.16 \approx 6$$

ثانياً / تحسب التباين من العلاقة الآتية:

$$S^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2 f_i}{n} = \frac{75}{44} = 1.7$$

تدريب (2)

من جدول التوزيع التكراري الآتي أوجد التباين :

الوزن X	50	75	85	82	55	59
التكرار F	4	8	11	5	7	6

ثالثاً : حساب التباين لبيانات مبوبة بفئات .

تعريف: إذا كانت مراكز توزيع فئات جدول تكراري هي $X_1, X_2, \ldots, X_n$ وكانت التكرارات المقابلة لها هي $f_1, f_2, \ldots, f_n$ فالتباين S^2 يحسب من العلاقة

$$S^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2 f_i}{n-1}$$

حيث $\overline{x}$: المتوسط الحسابي .

n : عدد القيم (مجموع التكرارات)

X_i : مراكز الفئات .

f_i : التكرار.

مثال (3)

احسب التباين للتوزيع التكراري في الجدول الآتي:

حدود الفئة	30 -34	35 -39	40 -44	45 -49	50 -54
التكرار	6	5	12	9	8

الحل:

نوجد مراكز الفئات، ونضرب كل مركز بالتكرار المقابل له لكي نحسب الوسط الحسابي؛ ثم نجد انحرافات مراكز الفئات عن الوسط الحسابي، ونربع هذه الانحرافات ونضرب كلاً منها بالتكرار المقابل لها كما يظهر في الجدول الآتي:

الفئة	مركز الفئة	التكرار				
30 -34	32	6	192	-11	121	726
35 -39	37	5	185	-6	36	180
40 -44	42	12	504	1	1	12
45 -49	47	9	423	4	16	144
50 -54	52	8	416	9	81	648
المجموع		40	1720			1710

الحل:

أولاً / نوجد الوسط الحسابي من العلاقة الآتية :

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i f_i}{\sum_{i=1}^{n} f_i} = \frac{1720}{40} = 43$$

ثانياً / نوجد التباين من العلاقة الآتية:

$$S^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2 f_i}{n} = \frac{1710}{40} = 42.75$$

نشاط (3)

الجدول الآتي يتضمن نتائج 33 شخصاً من المتقدمين للحصول على عمل في إحدى شركات الأدوية ، والمطلوب حساب التباين لهذه البيانات ؟

الفئات X	22 -32	33 -43	44 -54	55 -65	66 -76	المجموع
التكرار F	5	6	10	9	4	34

الحل:..

....

....

....

....

....

تدريب (3)

من جدول التوزيع التكراري الآتي أوجد التباين :

الفئات X	22 -32	33 -43	44 -54	55 -65	66 -76	المجموع
التكرار F	8	9	10	7	4	38

◆ توجد علاقات أخرى لحساب التباين منها :

أولاً : في حالة البيانات غير المبوبة نستخدم العلاقة الآتية:

$$S^2 = \frac{\sum_{i=1}^{n} x_i^2 - n\overline{x}^2}{n-1}$$

مثال (4)

من مثال (1) في حساب التباين السابق أحسب التباين بهذه العلاقة وقارن النتيجة التي حصلت عليها مع النتيجة السابقة .

الحل

القيم هي : 4، 8، 6، 7، 5، 8، 4، 9، 7، 2

نحسب المتوسط الحسابي :

$$\overline{X}=\frac{\sum x}{n}=\frac{4+8+6+7+5+8+4+9+7+2}{10}=\frac{60}{10}=6$$

$$\overline{X}^2=36$$

$$\sum_{i=1}^{n} x_i^2=16+64+36+49+25+64+16+81+49+4=404$$

وبالتعويض في العلاقة :

$$S^2=\frac{\sum_{i=1}^{n} x_i^2-n\overline{x}^2}{n-1}=\frac{404-(10)(36)}{9}=4.89$$

وهي قيمة مساوية للقيمة السابقة .

ثانياً : في حالة البيانات المبوبة نستخدم العلاقة الآتية :

$$S^2=\frac{\sum_{i=1}^{n} x_i^2 f_i-n\overline{x}^2}{n-1}$$

مثال (5)

من مثال (2) في حساب التباين السابق أحسب التباين بهذه العلاقة وقارن النتيجة التي حصلت عليها مع النتيجة السابقة .

الحل

نكون جدول التوزيع التكراري كما يلي :

الملاحظة	التكرار			
4	8	32	16	128
5	4	20	25	100
8	6	48	64	384
7	15	105	49	735
6	11	66	36	396
المجموع	44	271		1743

من الجدول أعلاه نعوض في العلاقة الآتية :

$$S^2=\frac{\sum_{i=1}^{n} x_i^2 f_i-n\overline{x}^2}{n-1}=\frac{1743-(44)(6.16)^2}{43}=1.71$$

وهي مساوية تقريباً للقيمة التي حصلنا عليها سابقاً والفرق في الناتج سببه التقريب في حساب المتوسط الحسابي في المثال السابق .

مثال (6)

من مثال (3) في حساب التباين السابق أحسب التباين بهذه العلاقة وقارن النتيجة التي حصلت عليها مع النتيجة السابقة .

الحل

نكون جدول توزيع تكراري كما يلي:

الفئة	مركز الفئة	التكرار			
30-34	32	6	192	1024	6144
35-39	37	5	185	1369	6845
40-44	42	12	504	1764	21168
45-49	47	9	423	2209	19881
50-54	52	8	416	2704	21632
المجموع		40	1720		75670

لإيجاد التباين نعوض في العلاقة الآتية:

$$S^2 = \frac{\sum_{i=1}^{k} x_i^2 f_i - n\bar{x}^2}{n} = \frac{75670 - (40)(43)^2}{40} = 42.75$$

وهي قيمة مساوية للقيمة التي حصلنا عليها سابقاً.

نشاط (4)

احسب التباين بالعلاقات أعلاه للأنشطة (1، 2، 3) في درس التباين، وقارن النتائج التي ستحصل عليها مع النتائج التي حصلت عليها سابقاً؟ ماذا تلاحظ؟

الحل:...
....
....
....
....

تدريب (4)

باستخدام العلاقات السابقة لحساب التباين أوجد التباين للبيانات الآتية:

7، 3، 5، 8، 2، 9، 6

2. الانحراف المعياري (Standard Deviation)

تعريفه:

الانحراف المعياري هو الجذر التربيعي للتباين.

أي أن: $S = \sqrt{S^2}$

طرق حسابه:

الحالة الأولى: حساب الانحراف المعياري لبيانات غير مبوبة.

يُحسب من العلاقة التالية لبيانات المجتمع:

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

حيث: n: عدد القيم.

$\bar{x}$: المتوسط الحسابي.

x_i: القيم (الملاحظات).

أو العلاقة الآتية وتستخدم للعينات:

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

وتطرح (1) من حجم العينة في هذه الحالة للتقليل من خطأ المعاينة الذي قد يقع فيه الباحث أثناء اختيار العينة وتُطبق هذه العلاقة إذا كان حجم العينة أقل من (30) عنصر، أما إذا زاد حجم العينة عن (30) عنصر فيصبح حكمها حكم المجتمع أي تقسم على (n) بدلاً من (n-1).

مثال (1)

من مثال (1) في موضوع التباين وجدنا أن التباين = 4.89

لحساب الانحراف المعياري يمكننا حساب التباين ثم نأخذ الجذر التربيعي للناتج النهائي للتباين فنحصل على الانحراف المعياري.

إذاً الانحراف المعياري لهذا المثال هو: $S = \sqrt{S^2} = \sqrt{4.89} = 2.21$

الحالة الثانية: حساب الانحراف المعياري لبيانات مبوبة بدون فئات.

نستخدم العلاقة الآتية:

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

مثال (2)

من مثال (2) في موضوع التباين وجدنا أن التباين = 1.7

لحساب الانحراف المعياري يمكننا حساب التباين ثم نأخذ الجذر التربيعي للناتج النهائي للتباين فنحصل على الانحراف المعياري.

إذا الانحراف المعياري لهذا المثال هو

$$S = \sqrt{S^2} = \sqrt{1.7} = 1.3$$

الحالة الثالثة: حساب الانحراف المعياري لبيانات مبوبة بفئات.

نستخدم العلاقة الآتية:

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

مثال (3)

من مثال (3) في موضوع التباين وجدنا أن التباين = 42.75

لحساب الانحراف المعياري يمكننا حساب التباين ثم نأخذ الجذر التربيعي للناتج النهائي للتباين فنحصل على الانحراف المعياري.

إذا الانحراف المعياري لهذا المثال هو:

$$S = \sqrt{S^2} = \sqrt{42.75} = 6.5$$

نشاط (1)

احسب الانحراف المعياري بالعلاقات أعلاه للأنشطة (1، 2، 3) في درس التباين؟

الحل: ....

....

....

....

نشاط (2)

مجموعة مكونة من 28 مفردة، فإذا كان $\sum x = 1893$ وكان $\sum x^2 = 127994$، احسب المتوسط الحسابي والتباين والانحراف المعياري؟

الحل: ....

....

....

....

....

تدريب (1)

احسب الانحراف المعياري بالعلاقات أعلاه للتدريبات (1، 2، 3) في درس التباين؟

◆ **شروط استخدام الانحراف المعياري للمقارنة بين ظاهرتين:**

- تساوي الأوساط الحسابية للظاهرتين.
- تساوي أحجام العينات.
- وحدات القياس واحدة للظاهرتين.

❖ معامل (شبرد) لتصحيح التباين :

عند حساب الانحراف المعياري S فإنه يكون مُعرضاً لبعض الأخطاء الناتجة من تجميع البيانات في فئات وتسمى بأخطاء التجميع .

ولتعديل هذا الخطأ يمكن استخدام ما يُعرف بمعامل (شبرد) ويعطى من العلاقة الآتية :

التباين المعدل = التباين من المفردات المجمعة − $\frac{C^2}{12}$

حيث

C : طول الفئة .

$\frac{C^2}{12}$: معامل التصحيح .

ولحساب معامل شبرد للمثال (3) في درس التباين نتبع ما يلي :

$$S^2 = 43.85 - \frac{5^2}{12} = 43.85 - 2.08 = 41.77$$ (التباين المعدل)

الانحراف المعياري المعدل او المصحح = 6.46

بعض خصائص الانحراف المعياري:

1- يعتمد الانحراف المعياري في حسابه على جميع القيم.

2- انه مقياس محدد جبرياً ويتمتع ببعض المزايا الجبرية.

3- لا يمكن حسابه للجداول التكرارية المفتوحة.

4- يتأثر بالقيم الشاذة.

5- لا يمكن حسابه للبيانات النوعية.

6- إذا اضفنا أو طرحنا مقداراً ثابتاً C إلى كل قيم Xi فإن قيمة الانحراف المعياري لا تتغير.

7- إذا ضربنا كافة القيم بعدد ثابت وليكن C فإن التباين يساوي التباين للقيم الأصلية مضروباً في مربع المقدار الثابت والانحراف المعياري الجديد يساوي الانحراف المعياري للقيم الأصلية مضروباً في المقدار الثابت $S_y = CS_x$.

8- إذا قسمنا جميع القيم على مقدار ثابت C فإن الانحراف المعياري الناتج هو الانحراف المعياري للقيم الأصلية مقسوماً على المقدار الثابت C .

❖ مقاييس التشتت النسبي .

1. معامل الاختلاف (Coefficient of Variation):

تعريفه :

هو احد مقاييس التشتت النسبي ويستخدم للمقارنة بين مجموعتين من البيانات تختلفان في وحدات القياس.

بمراجعة مقاييس التشتت التي سبق لنا مناقشتها (المدى، الانحراف المتوسط، التباين، الانحراف المعياري) نجد انها مقاييس مطلقة محسوبة بدلالة وحدات القياس الأصلية المستقلة في قياس المتغيرات التي ندرسها، سواءً أكانت هذه الوحدات عبارة عن درجات أو سنوات او غير ذلك.

وإذا أردنا مقارنة درجة التشتت لمجموعتين أو صنفين متصلين بنفس المجموعة حال دون ذلك اختلاف وحدات القياس المستخدمة في الحالتين، فمثلاً إذا أردنا مقارنة التشتت في أوزان مجموعة بالتشتت في أعمار نفس المجموعة أو مجموعة اخرى، فنجد ان مقياس التشتت في المجموعة الأولى يكون بالكيلو جرام بينما يكون في المجموعة الثانية بالسنوات، ولا يعقل أن نقارن بين الكيلو جرام والسنوات، ونقول أن الكيلو جرام أكبر او أقل من تشتت السنوات. ولإجراء هذه المقارنة لابد من التخلص من وحدات القياس، وذلك باستخدام مقياس نسبي يخلصنا من هذه الوحدات المختلفة، ويمكن أن نجعل الانحراف المعياري مقياساً للتشتت بقسمته على الوسط الحسابي، فيكون خارج القسمة عبارة عن نسبة مجردة من التمييز يُطلق عليها معامل الاختلاف أو معامل التغير ويحسب من العلاقة $C.V = \frac{s}{\bar{x}}(100\%)$

مثال (1)

إذا كان لديك بيانات إحصائية انحرافها المعياري = 13 ، ومتوسطها الحسابي = 50 فاوجد معامل الاختلاف لهذه البيانات ؟

الحل

$$C.V = \frac{s}{\bar{x}}(100\%) = \frac{13}{50}(100\%) = 26\%$$

مثال (2)

إذا كان الوسط الحسابي لأطوال مجموعة من الأشخاص = 165 سم بانحراف معياري = 8 وكان متوسط أوزانهم = 70 كجم بانحراف معياري = 7 .

المطلوب

احسب معامل الاختلاف للطول والوزن ثم حدد أيهما أكثر تشتتاً ؟

الحل

معامل الاختلاف للطول $C.V=\frac{s}{\bar{x}}(100\%)=\frac{8}{165}(100\%)=4.85\%$

معامل الاختلاف للوزن $C.V=\frac{s}{\bar{x}}(100\%)=\frac{7}{70}(100\%)=10\%$

أوضحت نتائج التحليل أن متغير الوزن أكثر تشتتاً من متغير الطول ، بمعنى أن الأشخاص أكثر تجانساً في أطوالهم عما هو في أوزانهم.

نشاط (1)

إذا كان الوسط الحسابي لإنتاج العامل في المعمل A = 1400 وحدة وبانحراف معياري = 400، وكان الوسط الحسابي لإنتاج العامل في المعمل B = 1300 وحدة وبانحراف معياري = 300 خلال شهر ، أي من المعملين أقل تشتتاً في الإنتاج.

الحل: ....

....

....

....

....

....

تدريب (1)

إذا كان لديك بيانات إحصائية متوسطها الحسابي = 15 والتباين لها = 25 أوجد معامل الاختلاف لهذه البيانات .

2. القيم المعيارية

تعريف :

هي إحدى مقاييس التشتت النسبي وتعرف بأنها خارج قسمة الفرق بين القيمة الفعلية والمتوسط الحسابي للبيانات على الانحراف المعياري .

ويمكن تعريفها رمزياً كما يلي :

إذا كانت $X_1, X_2, X_3, \dots, X_n$ هي مجموعة من القيم متوسطها الحسابي $\bar{X}$ وانحرافها المعياري S

فنرمز للقيمة المعيارية بالرمز Z وتحسب من العلاقة التالية في حالة العينة : $Z=\frac{x_i-\bar{x}}{s}$

وفي المعادلة أعلاه يتم تحويل كل قيمة من قيم المتغير العشوائي X إلى قيمة معيارية Z وهي التي تستخدم كأساس لمقارنة القيم ببعضها على أساس الوحدات المعيارية .

مثال (1)

إذا كانت لدينا درجات طلاب في مادتي البيولوجي والكيمياء ونرغب في مقارنة درجتي أحد الطلبة في المادتين ، فإذا كانت درجته في البيولوجي 80 وفي الكيمياء 70 ، وإذا كان متوسط درجات الطلبة في البيولوجي 85 بانحراف معياري 5 ، وان متوسط درجات الطلاب في الكيمياء = 62 بانحراف معياري = 6 ، فهل يعني أن تحصيل الطالب في البيولوجي أفضل من تحصيله في الكيمياء ؟

الحل

نحسب الدرجة المعيارية للمادتين :

درجة الطالب المعيارية في البيولوجي Z_1

$$Z_1=\frac{x_i-\bar{x}}{s}=\frac{80-85}{5}=-1$$

درجة الطالب المعيارية في الكيمياء Z_2

$$Z_2=\frac{x_i-\bar{x}}{s}=\frac{70-62}{6}=1.33$$

يلاحظ أن تحصيل الطالب في الكيمياء أفضل من تحصيله في البيولوجي على الرغم من أن درجته في البيولوجي أكبر من درجته في الكيمياء لأن درجته في الكيمياء تبعد بمقدار 1.33 عن المتوسط .

تدريب (1)

إذا كانت لديك البيانات التالية : 4 ، 6 ، 8 ، 10 ، 12 ، 14

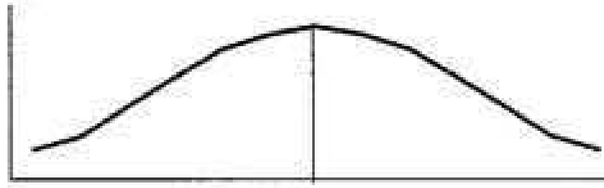
احسب كلاً مما يأتي :

1. المتوسط الحسابي لهذه البيانات .
2. الانحراف المعياري لهذه البيانات .
3. معامل الاختلاف لهذه البيانات .
4. القيمة المعيارية للقيمتين 4 ، 12 .

◆ مقاييس التماثل والالتواء :

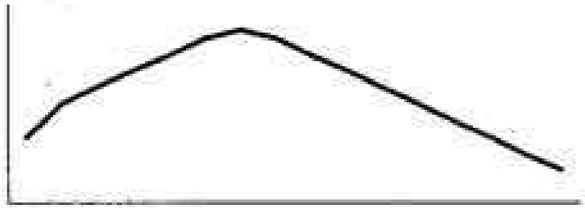
يقال أن التوزيع متماثل عندما يتطابق نصفيه حول محور عمودي كما بالشكل التالي :

التوزيع التكراري المتماثل



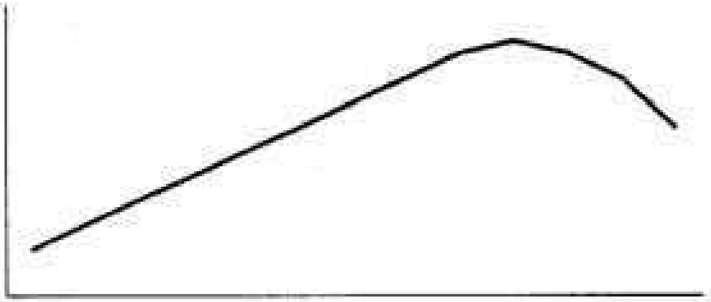
وعندما لا يتطابق جانبي التوزيع يقال عنه توزيع ملتو ، فعندما يكون الالتواء باتجاه اليمين يقال عنه توزيع موجب الالتواء حيث يصبح طرفه الأيمن أطول من طرفه الأيسر (ملتو نحو اليمين) كما بالشكل التالي :

التوزيع التكراري الموجب الالتواء



وعندما يكون الالتواء باتجاه اليسار يقال عنه توزيع سالب الالتواء حيث يصبح طرفه الأيسر أطول من طرفه الأيمن (ملتو نحو اليسار) كما بالشكل التالي :

التوزيع التكراري السالب الالتواء



وهناك أكثر من طريقة لحساب معامل الالتواء منها :

1. معامل (بيرسون) للالتواء:

يعد معامل بيرسون للالتواء من أهم مقاييس الالتواء والتماثل ويرمز له بالرمز S_k ويحسب من العلاقة التالية :

معامل الالتواء = (الوسط الحسابي – المنوال) ÷ الانحراف المعياري

ورمزياً من العلاقة :

$$S_k = \frac{\overline{X} - \hat{X}}{S}$$

حيث :

$\overline{X}$: المتوسط الحسابي

$\hat{X}$: المنوال

S : الانحراف المعياري

ومن عيوبه: شموله على المنوال الذي قد لا يكون موجوداً بين البيانات أو قد يكون للبيانات أكثر من منوال، بالإضافة إلى صعوبة تحديده في حالة البيانات المبوبة، ولمعالجة هذا العيب نلاحظ أن الوسيط في التوزيعات الملتوية يبعد حوالي الضعف عن المنوال لذا تم اشتقاق صيغة بديلة لمعامل بيرسون للالتواء هي:

$$S_k = \frac{3(\bar{x} - \tilde{x})}{S}$$

حيث: $\tilde{x}$ ترمز للوسيط

ويكون معامل بيرسون للالتواء موجباً باتجاه اليمين إذا كان الوسط الحسابي أكبر من الوسيط والمنوال، ويكون سالباً باتجاه اليسار إذا كان الوسط الحسابي أقل من الوسيط والمنوال وعند تطابق المتوسطات الثلاثة (حالة التماثل) فإن $S_K = 0$

وعموماً فإن قيمة S_K تقع بين 3- ، 3.

مثال (1)

احسب معامل بيرسون لالتواء التوزيع الذي فيه $\bar{x} = 5$ ، $\tilde{x} = 4$ ، $S = 1.2$ ثم حدد درجة ونوع الالتواء.

الحل

نعوض بالقيم في العلاقة الآتية:

$$S_k = \frac{3(\bar{x} - \tilde{x})}{S} = \frac{3(5-4)}{1.2} = \frac{3}{1.2} = 2.5$$

ومن هذه النتيجة نستدل على أن درجة الالتواء في شكل التوزيع كبيرة نسبياً، ويكون التوزيع موجباً باتجاه اليمين.

نشاط (1)

علل: من عيوب معامل بيرسون للالتواء احتوائه على المنوال؟ وكيف تعالج هذا العيب؟

الحل: ....

....

....

نشاط (2)

إذا توافرت لديك البيانات التالية والتي تمثل خلاصة مختصرة لنتائج الطلاب والطالبات الناجحين في مادة الإحصاء الطبي في إحدى السنوات

المقياس	الطلاب	الطالبات
الوسط الحسابي للدرجات	70	80
الوسيط للدرجات	75	75
المنوال للدرجات	85	65
الانحراف المعياري للدرجات	21	20

المطلوب:

حساب معامل الالتواء لكل من الطلاب والطالبات بالصيغتين السابقتين، وقارن النتائج التي حصلت عليها.

الحل: ....

....

....

....

....

....

....

تدريب (1)

إذا كانت الرواتب المحددة لخريجي كلية الطب عند بداية العمل في المستشفيات الحكومية يبلغ متوسطها 1250 دولار وبانحراف معياري مقداره 350، فهل يُعد الراتب 2000 دولار راتباً اعتيادياً يقع ضمن هذا المدى؟

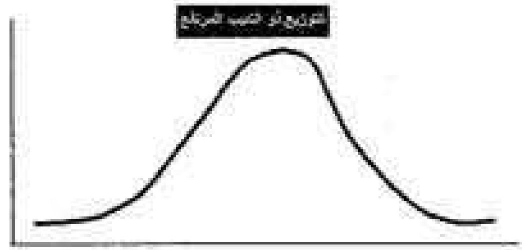
❖ معامل التفرطح والتدبب :

تعريف :

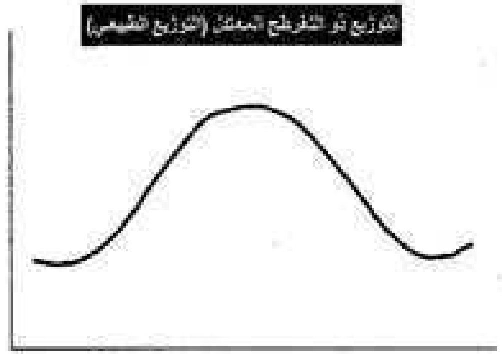
التفرطح هو درجة تدبب قمة منحنى التوزيع التكراري ، أي يقيس مدى تدبب قمة المنحنى التكراري .

سندرس فيما بعد التوزيع الإعتدالي الذي من صفاته ان تفرطحه = 3 ويعتبر هذا التوزيع أساس المقارنة عند قياس التفرطح وعليه فإن شكل التفرطح يأخذ حالات عدة منها :

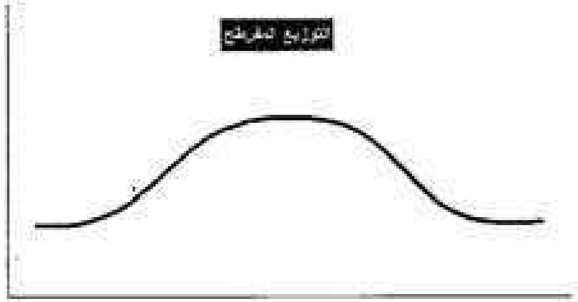
1 . التوزيع ذو التدبب المرتفع : وتكون أطرافه واسعة نسبياً وقمته ضيقة ، ويكون تفرطحه أكبر من (3)، ويأخذ الشكل البياني الآتي :



2 . التوزيع المعتدل التفرطح : ويكون مقياس تفرطحه يساوي 3 ويوضحه الشكل التالي :



3 . التوزيع المفرطح : وتكون قمة المنحنى مسطحاً واطرافه ضيقة ، ويكون تفرطحه اقل من 3 ، ويأخذ الشكل التالي :



ويحسب مقياس التفرطح من العلاقة التالية : $k_1 = \frac{m_4}{m_2^2}$

حيث :

m_4 : العزم الرابع حول المتوسط ويحسب من العلاقة $m_4 = \frac{\sum (x_i - \bar{x})^4 f_i}{n}$

m_2^2 : هو العزم الثاني حول المتوسط (التباين) أو هو القوة الرابعة للانحراف المعياري .

مثال (1)

إذا كان لديك توزيع تكراري بحيث $m_4 = 87.933$ ، $m_2 = 6.0336$ فأحسب مقياس التفرطح لهذا التوزيع ؟

الحل:

لحساب مقياس التفرطح للبيانات السابقة نعوض في العلاقة التالية :

$$k_1 = \frac{m_4}{m_2^2} = \frac{87.933}{(6.0336)^2} = 2.415$$

تدريب (2)

1 . عرف التفرطح ، ومتى يكون التوزيع معتدل ؟

تمارين الفصل الرابع

س1 / اكمل الجدول التالي ثم اعتمد عليه للإجابة من الأسئلة (1 – 30)

الفئة	F	ت.م.ص	ت.م.هـ	مركز الفئة X	X.F					
2 -4	1									
5 -7	2									
8 -10	4									
11 -13										
14 -16	1									
المجموع	10									

1. هذا الجدول هو :

أ. منتظم مفتوح . ب. منتظم مغلق . ج. غير منتظم مفتوح . د. غير منتظم مغلق.

2. مركز الفئة الثالثة =

أ. 9 . ب. 12 . ج. 6 . د. خلاف ذلك وهو :

3. التكرار النسبي للفئة الثانية :

أ. 0.10 . ب. 0.20 . ج. 0.40 د. خلاف ذلك وهو :

4. المدى لهذه البيانات =

أ. 14 . ب. 15 ج. 16 د. خلاف ذلك وهو :

5. تكرار الفئة الرابعة =

أ. 5 ب. 1 ج. 0 د. خلاف ذلك وهو :

6. زاوية القطاع للفئة الثالثة =

أ. 72 ب. 108 ج. 144 د. خلاف ذلك وهو :

7. التكرار المتجمع الصاعد للفئة الرابعة =

أ. 20 ب. 80 ج. 90 د. خلاف ذلك وهو :

8. التكرار المتجمع الهابط للفئة الثالثة =

أ. 70 ب. 90 ج. 30 د. خلاف ذلك وهو :

9. التكرار النسبي للفئة الخامسة =

أ. 0.10 ب. 0.20 ج. 0.30 د. خلاف ذلك وهو :

10. الوسط الحسابي لهذه البيانات =

أ. 10 ب. 9 ج. 9.5 د. خلاف ذلك وهو :

11. الوسيط لهذا التوزيع =

أ. 10 ب. 10.5 ج. 9 د. خلاف ذلك وهو :

12. المنوال لهذا التوزيع =

أ. 9 ب. 9.5 ج. 10 د. خلاف ذلك وهو :

13. هذا التوزيع من حيث الإلتواء هو:

أ. سالب. ب. موجب. ج. معتدل د. خلاف ذلك وهو :

14. إذا عدلت جميع قيم التوزيع بحيث طرح منها العدد 6 فإن الوسط الحسابي الجديد هو:

أ. 3 ب. 6 ج. 9 د. خلاف ذلك وهو :

15. التباين لهذه البيانات هو:

أ. 10 ب. 9 ج. 9.8 د. خلاف ذلك وهو :

16. الانحراف المعياري لهذا التوزيع هو:

أ. 10.8 ب. 9.8 ج. 3.28 د. خلاف ذلك وهو :

17. مجموع انحرافات القيم عن العدد 9 =

أ. 10 ب. 9 ج. 0 د. خلاف ذلك وهو :

18. الانحراف المتوسط لهذه البيانات هو:

أ. 7.7 ب. 2.4 ج. 4.2 د. خلاف ذلك وهو :

19. الربيع الثاني لهذا التوزيع =

أ. 9 ب. 10 ج. 9.9 د. خلاف ذلك وهو :

20. العشير الخامس لهذه البيانات =

أ. 9 ب. 10 ج. 9.9 د. خلاف ذلك وهو :

21. المئين 50 لهذا الجدول =

أ. 9 ب. 10 ج. 9.9 د. خلاف ذلك وهو :

22. العلامة المعيارية لمركز الفئة 9 في هذا التوزيع =

أ. صفر ب. 1 ج. 2 د. خلاف ذلك وهو :

23. الحد الأدنى الفعلي للفئة الرابعة =

أ. 10.5 ب. 11.5 ج. 11 د. خلاف ذلك وهو :

24. الحد الأعلى الفعلي للفئة الرابعة =

أ. 13 ب. 13.5 ج. 12.5 د. خلاف ذلك وهو :

25. المدى الربيعي لهذه البيانات =

أ. 6.5 ب. 5.5 ج. 4.5 د. خلاف ذلك وهو :

26. نصف المدى الربيعي لهذه البيانات =

أ. 3.25 ب. 2.75 ج. 2.25 د. خلاف ذلك وهو :

27. طول الفئة لهذا التوزيع =

أ. 4 ب. 6 ج. 3 د. خلاف ذلك وهو :

28. إذا عدلت جميع قيم التوزيع بحيث ضربت بالرقم 2 فإن التباين الجديد =

أ. 43.2 ب. 56.2 ج. 10.8 د. خلاف ذلك وهو :

29. . إذا عدلت جميع قيم التوزيع بحيث أضيف لها الرقم 2 فإن التباين الجديد =

أ. 12.8 ب. 8.8 ج. 10.8 د. خلاف ذلك وهو :

30. . إذا عدلت جميع قيم التوزيع بحيث ضربت بالرقم 2 فإن الانحراف المعياري =

أ. 6.56 ب. 7.5 ج. 10.8 د. خلاف ذلك وهو :

س2/ احسب الانحراف المتوسط لكل من القياسات التالية:

a) 2, 3, 4, 3, 5, 6, 2.
b) 45, 42, 45, 43.

س3/ اوجد المدى والتباين والانحراف المعياري للقيم التالية:

a) 5, 6, 8, 10, 12, 13.
b) 52, 56, 55, 56, 52.

س4/ الجدول التالي يمثل التوزيع التكراري لوزن 30 طفلاً لأقرب كيلو جرام:

الوزن	2 - 4	5 - 7	8 - 10	11 - 13
التكرار	5	8	12	5

أوجد :

أ) المدى ب) الانحراف المطلق ج) التباين والانحراف المعياري لأوزان الأطفال

س5/ أوجد المدى والتباين والانحراف المعياري لدرجات 103 طالباً موزعة كما يلي:

الدرجة	25	27	31	33	38	40	42	45	46	50
التكرار	5	10	12	15	19	15	11	8	6	2

س6/ البيانات الآتية درجات 60 طالباً في مادة الإحصاء الطبي (الدرجة العظمى 30 درجة):

12	19	16	16	18	25	20	18	20	19	20	22
19	25	20	23	22	27	15	18	17	11	18	22
18	19	20	20	26	20	23	12	10	15	16	15
28	16	15	17	22	16	21	25	29	24	19	17
13	12	28	16	17	27	12	14	15	21	22	23

كون جدول تكراري للبيانات السابقة بفئات بحيث يكون طول الفئة يساوي 5، ثم أوجد المدى والانحراف المتوسط والانحراف المعياري والتباين لدرجات هؤلاء الطلاب.

5

الفصل الخامس

الارتباط والانحدار الخطي البسيط

محتويات الفصل الخامس

أهداف الفصل الخامس

يتوقع من الطالب بعد الانتهاء من دراسة هذا الفصل أن يكون قادراً على:

- ✓ توضيح مفهوم الارتباط.
- ✓ استخدام لوحة الانتشار لمعرفة العلاقة بين متغيرين من حيث كونها طردية أو عكسية.
- ✓ حساب معامل الارتباط الخطي البسيط بين متغيرين وتفسير نتيجته.
- ✓ حساب معامل ارتباط الرتب بين متغيرين وصفيين.
- ✓ التمييز بين معامل ارتباط بيرسون ومعامل ارتباط سبيرمان .
- ✓ المقارنة بين مقاييس الارتباط المختلفة ومتى يُستخدم كل مقياس .
- ✓ حساب معامل الاقتران.
- ✓ حساب معامل التوافق.
- ✓ توضيح مفهوم الانحدار .
- ✓ توضيح المقصود بطريقة المربعات الصغرى.
- ✓ حساب معاملات خط انحدار متغير تابع على متغير مستقل.
- ✓ تقدير قيمة المتغير التابع عند معرفة قيمة المتغير المستقل.

مقدمة Introduction:

غالباً ما يكون من المهم دراسة العلاقة بين ظاهرتين أو أكثر وتحديد نوع وقوة تلك العلاقة، كالعلاقة بين ظاهرتي التدخين والإصابة بالسرطان ، التغذية والصحة ، الإنفاق والدخل، أو العلاقة بين ظاهرتي زيادة الإنتاج والقوى العاملة، أو العلاقة بين الطول والوزن وغيرها.

نظرية الارتباط عادة تظهر شدة أو قوة العلاقة بين الظاهرتين أو بين المتغيرين X و Y. أما دراسة هذه العلاقة من خلال التمثيل البياني بأفضل علاقة اقتران ممكنة بالشكل $Y = f(X)$ فتسمى بدراسة الانحدار ويسمى المستقيم أو المنحنى الذي يمثل هذه الدالة بمستقيم أو منحنى الانحدار.

الانحدار يعد أحد الأساليب الإحصائية المهمة والتي تستخدم بشكل واسع جداً ومنذ فترات طويلة لتحديد التأثيرات بين المتغيرات المستقلة X والمتغير التابع Y، ويمكن أن توضع هذه المتغيرات على شكل معادلات خطية بحيث يمكن استخدامها للتنبؤ عن قيمة المتغير التابع Y بدلالة المتغيرات المستقلة X فإذا كان المتغير التابع Y يعتمد على متغير مستقل واحد X فتسمي الانحدار عندئذ الانحدار الخطي البسيط Simple Linear Regression، أما إذا كان المتغير التابع Y يعتمد على عدد من المتغيرات المستقلة فيسمى الانحدار بالانحدار المتعدد Multiple Regression .

أولاً : الارتباط Correlation

من خلال دراستنا للمقاييس السابقة (مقاييس النزعة المركزية ومقاييس التشتت) والتي تستخدم لتحليل متغير واحد مثل أعمار مجموعة من الأشخاص أو أوزانهم أو أطوالهم، وفي هذا الفصل سوف ندرس أساليب إحصائية جديدة تستخدم لتحليل متغير ما من خلال علاقته بمتغير آخر، أو بعدة متغيرات مثل دراسة العلاقة بين الدخل والإنفاق أو بين الدخل والحالة التعليمية،... ومن خلال دراسة العلاقة بين متغير وآخر نتمكن من تحديد وجود العلاقة بين المتغيرات ونوعها ومتانتها وهذا الأسلوب يدعى بتحليل الارتباط.

ويمكن تقسيم العلاقات بين الظواهر إلى قسمين هما :

أ . علاقات تابعية، وذلك عندما يكون مقابل كل قيمة للمتغير المستقل قيمة للمتغير التابع مثل العلاقة بين مساحة المربع وطول ضلعه وعلاقة مساحة الدائرة بنصف قطرها.

ب. علاقات ارتباطية، وذلك عندما يكون مقابل كل قيمة للمتغير المستقل قيمة تقريبية أو احتمالية للمتغير التابع ومن أمثلة العلاقات الارتباطية: العُمر والإصابة بمرض السكري ، الدخل والإنفاق، مؤهل العامل وأجره وعدد الوحدات المنتجة ونصيب الوحدة الواحدة من التكاليف الثابتة. وسوف

نركز على النوع الثاني لأن النوع الأول سبق دراسته في الثانوية. مع ملاحظة أن العلاقة بين المتغيرات أو الظواهر قد تكون إما خطية وتمثل بخط مستقيم، أو منحنية (غير خطية) وتمثل بمنحنى غير خطي، وسنكرس دراستنا لدراسة الحالة التي تكون فيها علاقة الارتباط خطية بين الظاهرتين، حتى نتمكن فيما بعد أن تحدد قوة العلاقة واتجاهها.

والارتباط: هو ظاهرة إحصائية تطلق على العلاقة بين متغيرين، وهناك الكثير من المسائل العملية التي تكون فيها معرفة العلاقة بين متغيرين أو أكثر مهمة، ومن الأمثلة على ذلك:

1. دراسة العلاقة بين دخل الفرد وعدد ساعات العمل.
2. دراسة العلاقة بين أوزان وأطوال مجموعة من الأطفال في عمر معين.
3. معرفة العلاقة بين عدد الساعات التي يدرسها الطالب ومعدله التراكمي في الجامعة.

*** معامل الارتباط (Coefficient Correlation)**

يعرف معامل الارتباط بأنه: مقياس لدرجة العلاقة بين متغيرين.

كما يقصد بمعامل الارتباط أنه: مقياس إحصائي يستخدم لبيان نوع العلاقة بين المتغيرات سواءً كانت هذه العلاقة طردية أو عكسية.

ويأخذ معامل الارتباط قيمة من القيم المحصورة بين (1+ ، 1-) فمعامل ارتباط (1+) بين متغيرين يُعبر عن علاقة موجبة تامة (طردية) بين المتغيرين، أما معامل ارتباط (1-) فيُعبر عن علاقة تامة ولكنها سالبة (عكسية) بين المتغيرين. ومعاملات الارتباط بين أزواج من المتغيرات والتي تتراوح قيمها بين (1+ و 1-) تُعبر عن علاقة غير تامة فمعامل ارتباط (0.90 ، 0.90 -) يُعبر عن علاقة قوية بين المتغيرات، ومعامل ارتباط (0.50 ، 0.50-) يُعبر عن علاقة متوسطة، أما معامل ارتباط (0.20 ، 0.20 -) فيعبر عن علاقة ضعيفة بين المتغيرين والشكل الآتي يوضح ذلك.

• أسباب الارتباط:

هناك أسباب مختلفة لوجود ارتباط بين أي متغيرين يُمكن تلخيصها فيما يلي:

أ. وجود علاقة سببية مباشرة: في هذه الحالة فإن أحد المتغيرين يؤثر في الآخر مباشرة ، بحيث يُمكن تحديد أي منهما يُعد متغير مؤثر وأي منهما يُعد متغير متأثر ، فمثلاً من المعلوم أن هناك علاقة ارتباط موجبة وقوية بين وفرة المياه واتساع الرقعة الزراعية في المناطق الجغرافية المختلفة ، فهنا تكون وفرة المياه هي التي تسبب التوسع في الرقعة الزراعية وليس العكس كما توجد أيضاً علاقة ارتباط بين الدخل والإنفاق الاستهلاكي للأسرة.

ب. وجود سبب مشترك يؤثر في المتغيرين: وفي هذه الحالة يُعد كل من المتغيرين متغيراً مُتأثراً لهذا السبب المشترك ، وبالتالي لا يُفسر وجود علاقة الارتباط بينهما على أساس أن حدوث أحدهما يسبب في حدوث الآخر . مثال ذلك: وجود علاقة ارتباط قوية وموجبة بين الإنفاق على بند الطعام والإنفاق على بند الرعاية الطبية للأسرة ومستوى دخل الأسرة الذي يؤثر في بند الإنفاق.

ج. الصدفة: قد يحدث أحياناً أن تكون هناك علاقة ارتباط بين متغيرين نتيجة للصدفة ، ويُسمى الارتباط في هذه الحالة بالارتباط الوهمي أو الكاذب وعادةً ما يحدث هذا في العينات صغيرة الحجم.

يترتب على ما سبق أن وجود معامل ارتباط قوي بين متغيرين لا يعني بالضرورة وجود علاقة سببية بينهما دائماً.

• شكل الانتشار (Scatter Diagram)

يمكننا تقدير أو تحديد قيمة المتغير التابع من قيمة المتغير المستقل من خلال العلاقة الرياضية بين المتغيرات المستقلة والتابعة ، أما إذا أردنا تحديد قوة واتجاه العلاقة بين متغير وآخر فإننا نحتاج لعرض شكل الانتشار لتمثيل العلاقة بين ظاهرتين باستخدام الرسم البياني للمتغير المستقل X والذي يمثله المحور الأفقي ، والمتغير التابع Y والذي يمثل المحور العمودي . وتُمثل كل ثنائية (X,Y) بنقطة في المستوى الإحداثي فنحصل على شكل الانتشار والذي يتوقف على نوع العلاقة بين المتغيرين والأشكال التالية تعرض بعض الأشكال المختلفة للانتشار والتي نلاحظ فيها وجود الارتباط من عدمه ، أي تحديد اتجاه وقوة العلاقة .

شكل الانتشار

من خلال الأشكال السابقة ومن قراءة هذه الأشكال يُمكن تحديد طبيعة ونوع ومتانة العلاقة بين المتغيرات . فإذا كانت نقاط الانتشار تقع في مجال متقارب وعلى امتداد خط مستقيم واحد من أسفل إلى أعلى فإننا نستنتج وجود علاقة طردية قوية بين المتغير X والمتغير Y كما في الشكل a أعلاه.

اما إذا كان شكل الانتشار في حدود خط مستقيم يتجه من أعلى إلى أسفل وبدرجة تركيز عالية حول الخط فإن هذا يدل على وجود علاقة عكسية قوية بين المتغيرين Y , X كما في الشكل b .

اما إذا اظهر شكل الانتشار تركيزاً للنقاط ولكن حول منحنى وليس خطاً مستقيماً فإننا نستنتج من ذلك وجود علاقة غير خطية بين المتغيرين Y , X كما في الشكل e .

اما إذا كان شكل الانتشار أكثر تشتتاً للنقط بحيث يتعذر معه أن تقع في شكل خط مستقيم أو منحنى فإن ذلك يدل على عدم وجود علاقة ارتباطية بين المتغيرين Y , X كما في الشكل f .

وفي حالة وقوع جميع النقاط على خط مستقيم فإن العلاقة تامة موجبة أو سالبة كما في الشكلين c,d

ونخلص من ذلك إلى أن الأشكال السابقة هي عبارة عن أمثلة لأشكال الانتشار والتي تبين مدى قوة أو ضعف الارتباط بين المتغيرين Y , X ، أي أنها تعطي فكرة واضحة فيما إذا كانت Y تزداد بزيادة قيمة X أو تنقص بزيادة X أو أنها لا تتأثر.

مثال (1)

سجل أحد الباحثين عدد الساعات التي قضاها ثمانية طلاب تم اختيارهم عشوائياً من قسم صحة مجتمع في التحضير لاختبار مادة الإحصاء الطبي ، ودرجاتهم في ذلك الاختبار فوجدها كما في الجدول الآتي :

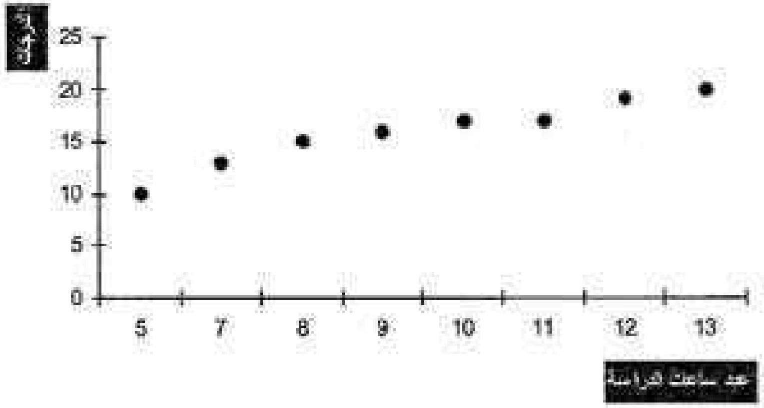
درجة الطالب في اختبار مادة الإحصاء الطبي Y	عدد الساعات التي قضاها الطالب للتحضير للاختبار X
15	8
17	11
16	9
13	7
10	5
17	10
19	12
20	13

ارسم لوحة الانتشار واذكر العلاقة بين درجة امتحان الإحصاء وعدد ساعات الدراسة.

الحل:

نرسم مستوى إحداثي بحيث يمثل محور السينات عدد ساعات الدراسة ، ومحور الصادات درجات الطالب في الاختبار ، ثم نرصد النقاط التي إحداثياتها الأزواج المرتبة الآتية:

(8,15),(11,17),(9,16),(7,13),(5,10),(10,17),(12,19),(13,20)



يظهر من لوحة الانتشار أنه يُمكن اعتبار العلاقة بين عدد ساعات الدراسة والدرجة في امتحان مادة الإحصاء الطبي علاقة خطية طردية ، أي كلما زادت X زادت Y .

- **أساليب قياس معامل الارتباط الخطي البسيط .**

يمكن حساب معامل الارتباط بين متغيرين بأكثر من طريقة سنتناول منها :

1 . معامل ارتباط بيرسون (Pearson Coefficient Correlation) :

ويستخدم لقياس العلاقة بين متغيرين تم قياس كل منهما إما بمستوى قياس فاصلي أو نسبي.

تعريف :

نفرض أن $x_1, x_2, \dots, x_n$ تُمثل قياسات ظاهرة ما ولتكن X ، ونفرض أن $y_1, y_2, \dots, y_n$ تُمثل قياسات ظاهرة أخرى ولتكن Y ، فإنه يُمكن حساب معامل ارتباط بيرسون بإحدى العلاقات الآتية:

العلاقة (1) :

$$r = \frac{n\sum x_i y_i - (\sum x_i)(\sum y_i)}{\sqrt{[n\sum x_i^2 - (\sum x_i)^2][n\sum y_i^2 - (\sum y_i)^2]}}$$

العلاقة (2) :

$$r = \frac{\sum x_i y_i - n(\bar{x})(\bar{y})}{\sqrt{[\sum x_i^2 - n(\bar{x})^2][\sum y_i^2 - n(\bar{y})^2]}}$$

مثال (1)

تُمثل البيانات الواردة في الجدول الآتي الطول X والوزن Y لخمسة طلاب في إحدى الكليات:

الطول X (سم)	164	169	165	170	172
الوزن Y (كجم)	58	67	74	71	70

احسب معامل ارتباط بيرسون الخطي بين Y , X ، وحدد نوع العلاقة وقوتها.

الحل

نرتب خطوات الحل كما في الجدول الآتي:

N	X	Y	X . Y		
1	172	70	12040	29584	4900
2	170	71	12070	28900	5041
3	165	74	12210	27225	5476
4	169	67	11323	28561	4489
5	164	58	9512	26896	3364
المجموع	840	340	57155	141166	23270

1. نحسب المتوسط لقيم X وقيم Y، كما يلي:

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{840}{5} = 168$$

$$\overline{y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{340}{5} = 68$$

وبالتعويض في المعادلة الآتية نجد معامل ارتباط بيرسون:

$$r = \frac{\sum x_i y_i - n(\overline{x})(\overline{y})}{\sqrt{\left[\sum x_i^2 - n(\overline{x})^2\right]\left[\sum y_i^2 - n(\overline{y})^2\right]}} = \frac{57155-(5)(168)(68)}{\sqrt{\left[141166-5(168)^2\right]\left[23270-5(68)^2\right]}} = \frac{35}{\sqrt{(46)(150)}} = 0.42$$

يتضح ان العلاقة بين الطول والوزن طردية (موجبة) ، وبدرجة متوسطة.

نشاط: احسب معامل الارتباط للبيانات في المثال(1) اعلاه باستخدام العلاقة (1) وقارن بين الناتجين.

مثال (2)

طُبق مقياس للاتجاهات نحو العمل ومقياس للشعور بالانتماء للمهنة على عشرة من مديري المستشفيات فحصلوا على الدرجات المثبتة في الجدول الآتي:

N	X	Y	X . Y		
1	16	14	224	256	196
2	20	13	260	400	169
3	24	18	432	576	324
4	28	19	532	487	361
5	32	22	704	1024	484
6	36	23	828	1296	529
7	40	24	960	1600	576
8	44	28	1232	1936	784
9	48	31	1488	2304	961
10	52	30	1560	2704	900
	340	222	8220	12880	5284

المطلوب:

احسب العلاقة الإرتباطية بين متغيري الاتجاهات نحو العمل والشعور بالانتماء للمهنة.

الحل: سنستخدم العلاقة (1) في الحل كما يلي:

$$r = \frac{n\sum x_i y_i - (\sum x_i)(\sum y_i)}{\sqrt{\left[n\sum x_i^2 - (\sum x_i)^2\right]\left[n\sum y_i^2 - (\sum y_i)^2\right]}} = \frac{(10)(8220)-(340)(222)}{\sqrt{\left[10(12880)-(340)^2\right]\left[10(5284)-(222)^2\right]}}$$

$$= \frac{6270}{\sqrt{(13200)(3556)}} = \frac{6720}{6851.219} = 0.981$$

نلاحظ ان العلاقة قوية موجبة بين مقياس الاتجاه نحو العمل ومقياس الشعور بالانتماء للمهنة.

نشاط: احسب معامل الارتباط للبيانات في مثال(2) اعلاه باستخدام العلاقة (2) وقارن بين الناتجين.

ماذا تلاحظ؟

مثال (3)

احسب معامل الارتباط بين قيم المتغيرين (X,Y) المثبتة في الجدول الآتي

N	X	Y	X . Y		
1	5	10	50	25	100
2	2	4	8	4	16
3	4	8	32	16	64
4	6	12	72	36	144
5	3	6	18	9	36
6	8	16	128	64	256
7	7	14	98	49	196
8	9	18	162	81	324
9	1	2	2	1	4
	45	90	570	285	1140

الحل

$$r = \frac{n\sum x_i y_i - (\sum x_i)(\sum y_i)}{\sqrt{\left[n\sum x_i^2 - (\sum x_i)^2\right]\left[n\sum y_i^2 - (\sum y_i)^2\right]}} = \frac{(9)(570) - (45)(90)}{\sqrt{\left[9(285) - (45)^2\right]\left[9(1140) - (90)^2\right]}}$$

$$= \frac{1080}{\sqrt{(540)(2160)}} = \frac{1080}{1080} = 1$$

نلاحظ في هذا المثال أن العلاقة تامة موجبة.

تدريب :

احسب معامل الارتباط للبيانات في المثال(3) اعلام باستخدام العلاقة (2) وقارن بين الناتجين. ماذا تلاحظ؟

نشاط (1)

احسب معامل الارتباط بين درجات الإحصاء الطبي X ومناهج البحث العلمي Y لسبعة طلاب والموضحة في الجدول الآتي محدداً نوع العلاقة وقوتها :

n	X	Y	X . Y		
1	7	11	77	49	121
2	6	12	72	36	144
3	8	13	104	64	169
4	5	10	50	25	100
5	6	13	78	36	169
6	7	15	105	49	225
7	3	9	27	9	81
	42	83	513	268	1009

الحل...

....

....

....

....

....

....

....

....

....

....

تدريب (1)

يوضح الجدول الآتي العمر (X) وضغط الدم (Y) لثمان أشخاص:

n	X	Y	X.Y		
1	56	147			
2	42	125			
3	72	160			
4	36	118			
5	63	149			
6	47	128			
7	55	150			
8	49	145			

احسب معامل الارتباط بين المتغيرين ، وحدد نوع العلاقة وقوتها .

- خصائص معامل ارتباط بيرسون :

1. تتراوح قيمة معامل الارتباط بين (1- ، 1+).
2. تكون قيمة معامل الارتباط = 1+ أو 1- عندما تكون العلاقة بين المتغيرين تامة ، وعند أي قيم أخرى محصورة بينهما تكون العلاقة غير تامة .
3. عند عدم وجود علاقة بين المتغيرات فإن قيمة معامل الارتباط = صفر .
4. تكون العلاقة طردية عندما يكون معامل الارتباط موجب ، وتكون العلاقة عكسية عندما يكون معامل الارتباط سالب .

- من مزايا معامل الارتباط أن قيمته مجردة من وحدات القياس.
- من قيود معامل الارتباط لبيرسون انه لا يُعبر عن متانة العلاقة بشكل صحيح إلا إذا كانت خطية، كما انه لا يُمكن استخدامه إذا كانت المتغيرات نوعية.

2. معامل ارتباط سبيرمان (Coefficient of Rank Correlation) :

من أهم معاملات الارتباط للرُتب معامل سبيرمان (Coefficient of Rank Correlation) ويستخدم لقياس العلاقة الإرتباطية بين متغيرين تم قياسهما بمقياس رتبي ويُحسب من العلاقة الآتية:

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

حيث

n : عدد أزواج البيانات (X , Y)

d : الفرق بين رتب X ورتب Y

طرق حسابه :

الحالة الأولى : في حالة عدم وجود رتب مكررة.

مثال (1)

احسب معامل سبيرمان للارتباط بالرتب بين المعدلات الآتية لعشرة طلاب في شهادة الثانوية والفصل الجامعي الأول الموضحة بالجدول الآتي:

معدل الطالب في الثانوية X	89	87	90	94	85	93	88	79	85	77
معدل الطالب للفصل الجامعي الأول Y	78	76	81	82	74	80	71	65	72	61

الحل

نرتب معدلات X بحيث نعطي الرتبة 1 لأعلى معدل وهكذا ، ونرتب معدلات Y بالمثل ثم نجد الفرق بين رتبتي كل طالب كما هو موضح بالجدول التالي :

ملحوظة / يُمكننا أن نعكس الترتيب بحيث نعطي الرتبة 1 لأقل معدل ولكن للمتغيرين معاً .

قيم X	قيم Y	رتب X	رتب Y	D	
77	61	10	10	0	0
85	72	7	7	0	0
79	65	9	9	0	0
88	71	5	8	-3	9
93	80	2	3	-1	1
82	74	8	6	2	4
94	82	1	1	0	0
90	81	3	2	1	1
87	76	6	5	1	1
89	78	4	4	0	0
					16

نعوض في المعادلة الآتية :

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6(16)}{10(99)} = 1 - 0.1 = 0.90$$

نلاحظ أن العلاقة قوية موجبة .

مثال (2)

من البيانات الواردة في الجدول والتي تمثل تقديرات احد طلبة المختبرات في فصلين دراسيين أوجد معامل الارتباط بين نتيجة الفصلين:

تقديرات الفصل الأول X	تقديرات الفصل الثاني Y	رتب X	رتب Y	d	
جيد جدا	ممتاز	2	1	1	1
جيد	جيد	3	3	0	0
ممتاز	جيد جدا	1	2	-1	1
مقبول	ضعيف	4	5	-1	1
ضعيف	مقبول	5	4	1	1
					4

الحل:

1. نعطي الرتبة 1 للتقدير ممتاز والرتبة 2 للتقدير جيد جداً ، ..... للمتغيرين(X,Y).

2. توجد الفرق بين رتب المتغيرين (X,Y) كما هو موضح بالجدول في العمود 5 ، ثم تربيع الفرق بين الرتب كما هو موضح في العمود 6 .

3. تحسب معامل الارتباط من العلاقة الآتية :

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6(4)}{5(24)} = 1 - \frac{24}{120} = 1 - 0.20 = 0.80$$

نلاحظ أن العلاقة قوية موجبة .

مثال (3)

إذا كانت لدينا البيانات الآتية والتي توضح متوسط عدد الأطفال في الأسرة والمستوى التعليمي للأم:

المستوى التعليمي للأم	دكتوراه	ماجستير	جامعية	ثانوية	أساسية	تقرأ وتكتب	أمية
متوسط عدد الأطفال	2	4	5	7	9	8	10

المطلوب :

تحديد نوع وقوة العلاقة الإرتباطية بين المستوى التعليمي للأم ومتوسط عدد الأطفال في الأسرة؟

الحل

نرتب المتغيرين كما هو موضح بالجدول:

No.	المستوى التعليمي للأم	متوسط عدد الأطفال	رتب المستوى التعليمي للأم	رتب متوسط عدد الأطفال	D	
1	أمية	10	1	7	-6	36
2	تقرا وتكتب	8	2	5	-3	9
3	أساسية	9	3	6	-3	9
4	ثانوية	7	4	4	0	0
5	جامعية	5	5	3	2	4
6	ماجستير	4	6	2	4	16
7	دكتوراه	2	7	1	6	36
						110

من الجدول نعوض في العلاقة:

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6(110)}{7(48)} = 1 - 1.964 = -0.96$$

تشير النتيجة إلى وجود علاقة ارتباط متينة بين المستوى التعليمي للأم ومتوسط عدد الأطفال في الأسرة، وهي علاقة عكسية قوية جداً، وهذا يدل على أنه كلما زاد المستوى التعليمي للأم كلما قل متوسط عدد الأفراد في الأسرة.

مقارنة بين معاملي ارتباط بيرسون وسبيرمان:

- معامل ارتباط سبيرمان أقل دقة وكفاءة من معامل ارتباط بيرسون، ويُعزى ذلك إلى عدم استخدام القيم الأصلية لكل من المتغيرين في معامل ارتباط سبيرمان، بينما اكتسب معامل ارتباط بيرسون الدقة العالية والكفاءة من استخدامه للقيم الأصلية.
- يكثر استعمال معامل ارتباط بيرسون في الدراسات الكمية، بينما يعد معامل ارتباط الرتب هو الأنسب والأكثر استعمالاً في حالة ما إذا كانت البيانات المتاحة وصفية.

نشاط (1)

احسب معامل الارتباط بين رُتب 7 طلاب في مادتين دراسيتين الواردة في الجدول:

NO.	X	Y	D	
1	3	6	-3	9
2	2	1	1	1
3	1	5	-4	16
4	4	4	0	0
5	5	7	-2	4
6	6	2	4	16
7	7	3	4	16
				62

الحل: ....

....

....

....

....

تدريب (1)

الجدول الآتي يوضح رُتب 5 طلاب في مادتي تثقيف صحي وصحة مجتمع، احسب معامل الارتباط بين رُتب المادتين:

NO.	X	Y	d	
1	4	6		
2	5	3		
3	10	7		
4	6	5		
5	8	8		

الحالة الثانية: في حالة وجود رتب مكررة.

مثال (3)

أوجد معامل الارتباط لدرجات أحد الطلبة عند التخرج من الثانوية العامة وتقديرات الفصل الجامعي الأول والتي كانت كما يلي:

الثانوية X	56	70	85	80	90	82	70	66
الجامعة Y	مقبول	متوسط	ممتاز	جيدجدا	ممتاز	جيد	متوسط	مقبول

الحل

1. نعط أعلى درجة الرتبة 1 ، .....
2. نعط أعلى تقدير الرتبة 1 ، ..... كما هو موضح بالجدول.
3. في حالة الرتب المكررة نأخذ المتوسط الحسابي لهذه الرتب ويعد هو الرتبة الجديدة

x	Y	رتب X	رتب Y	d	
56	مقبول	8	7.5 7	0.5	0.25
70	متوسط	5.5 5	5.5 6	0	0
85	ممتاز	2	1.5 1	0.5	0.25
80	جيد جدا	4	3	1	1
90	ممتاز	1	1.5 2	-0.5	0.25
82	جيد	3	4	-1	1
70	متوسط	5.5 6	5.5 5	0	0
66	مقبول	7	7.5 8	-0.5	0.25
$\sum$					3

4. نحسب معامل الارتباط من العلاقة الآتية:

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6(3)}{8(63)} = 1 - \frac{18}{504} = 1 - 0.036 = 0.964$$

نشاط (2)

من البيانات الواردة في الجدول احسب معامل ارتباط سبيرمان للرُتب:

قيم X	قيم Y	رتب X	رتب Y	d	
36	24				
25	13				
48	38				
36	24				
36	37				
92	80				
75	70				

الحل:...

....

....

....

....

....

تدريب (2)

من البيانات الواردة في الجدول احسب معامل ارتباط سبيرمان للرُتب:

التقدير X	ل	م	جـ	جـ جـ	ل	م	جـ جـ	ض
قيم Y	18	24	13	38	24	37	80	70

حيث الرموز في الجدول تعني:

(م) ممتاز، (جـ جـ) جيدجداً ، (جـ) جيد ، (ل) مقبول ، (ض) ضعيف.

• ارتباط الصفات (Correlation of Attributes)

عرفنا سابقاً أن معامل ارتباط بيرسون يعطي قوة العلاقة في حالة البيانات الكمية سواءً كانت مبوبة أو غير مبوبة ، وكذلك معامل ارتباط سبيرمان الذي يستخدم لحساب قوة الارتباط في حالة البيانات الكمية والوصفية التي لها صفة الترتيب ، إلا أنه توجد بيانات وصفية مميزة ولا نستطيع ترتيبها مثل الحالة الاجتماعية (متزوج ، أعزب ، مطلق ، أرمل) وكذلك لون العينين (زرقاء ، سوداء ، بنية) ولقياس العلاقة (قوة الارتباط) لهذه البيانات نشأت الحاجة إلى إيجاد مقاييس تقيس الارتباط بين هذه الصفات منها :

• معامل الاقتران (Coefficient of Association) :

يستخدم لقياس قوة الارتباط عندما يكون لدينا متغيرين نوعيين ولكل متغير صفتين فقط . أي يكون لدينا جدول اقتران 2 * 2 ، كأن ندرس العلاقة بين الإصابة بالسرطان والتدخين لمجموعة أشخاص مدخنين وغير مدخنين فالمتغيرين هما :

المرض (مصاب ، غير مصاب) ، والتدخين (مدخن ، غير مدخن) ولا يستخدم معامل الاقتران عندما يكون كل من المتغيرين أو أحدهما له أكثر من صفتين .

ومعامل الاقتران كمعامل الارتباط قيمته تكون محصورة بين (1- ، 1+) .

ولحساب معامل الاقتران تقسم المفردات حسب المتغيرات والصفات كما في الجدول الآتي :

المتغير A / المتغير B	القيمة (1)	القيمة (2)	المجموع
القيمة (1)	A	B	A + B
القيمة (2)	C	D	C + D
المجموع	A + C	B + D	N= A + B + C + D

ويحسب معامل الاقتران من العلاقة الآتية :

$$r_c = \frac{AD-BC}{AD+BC}$$

وتكون هناك علاقة قوية بين المتغيرات كلما اقتربت قيمة معامل الاقتران من الواحد ، وبصورة عامة نقول أن هناك علاقة عندما تكون قيمة r أكبر من 0.5

مثال (1)

قام أحد الباحثين بدراسة العلاقة بين التدخين والتهاب اللثة ، فجمع بيانات عن 400 شخص فكان توزيعهم حسب الصفات كما يلي :

A / B	مدخن	غير مدخن	المجموع
مصاب	140	20	160
غير مصاب	40	200	240
المجموع	180	220	400

المطلوب :

احسب معامل الاقتران.

الحل

$$r_c = \frac{AD-BC}{AD+BC} = \frac{(140)(200)-(20)(40)}{(140)(200)+(20)(40)} = \frac{28000-800}{28000+800} = \frac{27200}{28800} = 0.944$$

وهذا يدل على أن هناك علاقة قوية جداً بين التدخين ومرض التهاب اللثة ، وتجدر الإشارة إلى أنه ليس هناك أي مدلول لإشارة معامل الاقتران ، وذلك لأن إشارته تختلف حسب ترتيب تكرار الصفات في الجدول لذلك تؤخذ قيمته المطلقة للدلالة على وجود العلاقة من عدمها ؛ وقيمة هذا المعامل دائماً تكون أقل من الواحد الصحيح ما دامت $B . C > 0$ ولا تساوي الواحد إلا إذا كانت $B . C = 0$

نشاط (1)

قام أحد الباحثين بدراسة العلاقة بين تناول القات والتهاب اللثة، فجمع بيانات عن 500 شخص فكان توزيعهم حسب الصفات كما يلي:

A / B	متعاطي القات	غير متعاطي للقات	المجموع
مصاب	170	60	230
غير مصاب	50	220	270
المجموع	220	280	500

المطلوب:

احسب معامل الاقتران. وحدد قوة العلاقة بين المتغيرات.

الحل:..

....

....

....

....

تدريب (1)

قام أحد الباحثين بدراسة العلاقة بين التدخين والسرطان، فجمع بيانات عن 1200 شخص فكان توزيعهم حسب الصفات كما يلي:

A / B	مدخن	غير مدخن	المجموع
مصاب	500	130	630
غير مصاب	100	470	570
المجموع	600	600	1200

المطلوب:

احسب معامل الاقتران. وحدد قوة العلاقة بين المتغيرات.

• معامل التوافق Contingency Coefficient

يُستخدم معامل التوافق لقياس الارتباط في حالة البيانات الوصفية لظاهرتين، أو عندما تكون بيانات إحدى الظواهر وصفية والأخرى كمية.

والفرق بين هذا القياس ومعامل الاقتران أن الظواهر هنا قد تقسم إلى أكثر من فئتين، حيث تبوب البيانات في جدول يُسمى جدول التوافق، و يُمكن توضيح كيفية حساب معامل التوافق من خلال المثال الآتي:

مثال (1):

باستخدام بيانات الجدول الآتي أوجد مدى الارتباط بين المستوى التعليمي والتدخين.

التدخين / المستوى التعليمي	نعم	لا	
غير متعلم	6	4	10
متوسط	4	4	8
جامعي	5	7	12
	15	15	30

الحل:

يُحسب معامل التوافق باستخدام العلاقة الآتية:

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

حيثُ تحسب B من كافة خلايا جدول التوافق، وذلك من خلال الجمع لكافة الخلايا لناتج قسمة مربع كل خلية على حاصل ضرب مجموع الصف في مجموع العمود الواقعة فيه الخلية كما يلي:

$$B = \frac{(6)^2}{(15)(10)} + \frac{(4)^2}{(15)(10)} + \frac{(4)^2}{(15)(8)} + \frac{(4)^2}{(15)(8)} + \frac{(5)^2}{(15)(12)} + \frac{(7)^2}{(15)(12)} = 1.0244$$

$$C = \sqrt{\frac{1.0244-1}{1.0244}} = 0.15$$

وهذا يعني أن الارتباط بين المستوى التعليمي وظاهرة التدخين ضعيفة، وقيمة معامل التوافق تكون موجبة بصورة دائمة، لذلك فهو لا يحدد اتجاه العلاقة هل هي عكسية أم طردية، وبالتالي فإنه يجيب فقط على التساؤل حول مدى وجود العلاقة بين الظواهر المدروسة.

تدريب (1)

البيانات الواردة في الجدول الآتي توضح العلاقة بين متغيري الحالة الاجتماعية والتدخين:

الحالة الاجتماعية / التدخين	أعزب	متزوج	مطلق	أرمل	المجموع
يدخن	30	80	20	50	180
لا يدخن	25	15	30	50	120
المجموع	55	95	50	100	300

المطلوب :

احسب معامل التوافق.

- الانحدار Regression:

- مفهوم الانحدار Concept of Regression

من أهم التطبيقات الإحصائية في الاقتصاد والإدارة والعلوم الطبية والعلوم التربوية والعلوم البحتة وغيرها مسألة التنبؤ ، وهو يبنى على وجود علاقة بين متغيرين كأن تعرف العلاقة التي تربط المتغيرين X,Y مثلاً ونريد ان نقدر او نتنبأ القيمة Y التي تقابل قيمة معينة للمتغير X .

لقد سبق دراسة مفهوم الارتباط والارتباط الخطي البسيط بين متغيرين ، وبذلك تعرفت على طريقة تقيس بها قوة العلاقة الخطية بين متغيرين ، هل هذه العلاقة طردية أم عكسية ،وقد تم توضيح ذلك في لوحة الانتشار.

وفي دراستنا للانحدار سنحاول تكوين معادلة تُعبر عن العلاقة بين المتغيرين ، وتسمى المعادلة الخطية التي تعطينا أحد المتغيرين بدلالة المتغير الأخر معادلة الانحدار الخطي ، والتنبؤ واحد من أهم أغراض دراسة الانحدار ، وهو التنبؤ بقيمة متغير ما إذا عرفنا قيمة متغير آخر او متغيرات أخرى.

سنقتصر الشرح على التنبؤ بقيمة متغير واحد إذا عرفنا قيمة متغير آخر .

- معادلة الانحدار الخطي البسيط Y على X:

إذا كان لدينا عينة من الأزواج المرتبة: $(x_1, y_1)(x_2, y_2)......(x_n, y_n)$

حيث $x_1, x_2,.....x_n$ قيم المتغير X ويسمى متغير مستقل ، وحيث $y_1, y_2,.....y_n$ القيم المقابلة للمتغير Y ويسمى متغير تابع وإذا رصدنا هذه النقاط على المستوى (X Y) تحصل على لوحة الانتشار، ونتمكن من الحكم فيما إذا كان من المعقول تطبيق خط مستقيم على شكل الانتشار أم لا .

فإذا فرضنا ان هناك علاقة خطية بين المتغيرين X, Y امكن التعبير عن هذه العلاقة بالمعادلة:

$$y = a + b x + e$$

ويقابل هذه المعادلة للمشاهدة التي رقمها(i) المعادلة: $y_i = a + bx_i + e_i$

حيث :

a معامل ثابت (الجزء المقطوع من محور الصادات y) ويصبح مساوياً لقيمة y عندما x = 0 .

b يُدعى بميل الانحدار ويُمثل مقدار التغير في y عند زيادة المتغير x بمقدار وحدة واحدة ، كما يُطلق عليه بمعامل الانحدار.

Y المتغير التابع .

X المتغير المستقل .

e_i هي الأخطاء العشوائية $i = 1,2,...n$ وتشير إلى المشاهدات المختلفة للمتغيرات.

ويكون المطلوب هو تقدير b , a لكي نستطيع تقدير y المقابلة للمتغير x ، ويُعاد كتابة العلاقة السابقة بالشكل التالي : $\hat{y} = \hat{a} + \hat{b}x$

ويُطلق على هذه العلاقة معادلة خط انحدار y على x حيث :

$\hat{y}$: القيمة التقديرية للمتغير y .

$\hat{a}$: القيمة التقديرية لـ a .

$\hat{b}$: القيمة التقديرية لـ b .

وبناءً عليه يتم البحث عن معادلة أفضل خط مستقيم يمثل البيانات المدروسة ويمر بأغلب نقاط شكل انتشار هذه البيانات.

*** تقدير معادلة خط الانحدار السابقة:**

لتقدير معادلة خط الانحدار السابقة توجد عدة طرق لذلك ، من أهمها طريقة المربعات الصغرى ، وتتلخص هذه الطريقة في إيجاد قيم $(\hat{a}, \hat{b})$ التي تجعل قيم مربعات الأخطاء أقل ما يمكن ، وهذا المبدأ يعتمد على تخفيض الفروق بين القيم المقدرة والمشاهدات الأصلية إلى أقل ما يمكن .

ولتقدير المعلمتين $(\hat{a}, \hat{b})$ يُمكن استخدام العديد من العلاقات الرياضية أبسطها العلاقات الآتية:

$$\hat{b} = \frac{n\sum x_i y_i - (\sum x_i)(\sum y_i)}{n\sum x_i^2 - (\sum x_i)^2}$$

أو العلاقة المكافئة

$$\hat{b} = \frac{\sum x_i y_i - n(\bar{x})(\bar{y})}{\sum x_i^2 - n\bar{x}^2}$$

$$\hat{a} = \bar{y} - \hat{b}\bar{x}$$

أو العلاقة المكافئة

$$\hat{a} = \frac{\sum y_i - \hat{b}\sum x_i}{n}$$

حيث $\bar{x}$ هي الوسط الحسابي للمشاهدات $x_1, x_2, ..., x_n$ ، $\bar{y}$ هي الوسط الحسابي للمشاهدات $y_1, y_2, ..., y_n$ ، وبالتعويض عن قيمتي $\hat{a}, \hat{b}$ في المعادلة $\hat{y} = \hat{a} + \hat{b}x$ نحصل على معادلة خط انحدار y على x .

مثال (1)

يوضح الجدول الآتي درجات 9 طلاب في اختبار مقرر مايكرو بيولوجي (X) واختبار صيدلانيات (Y)

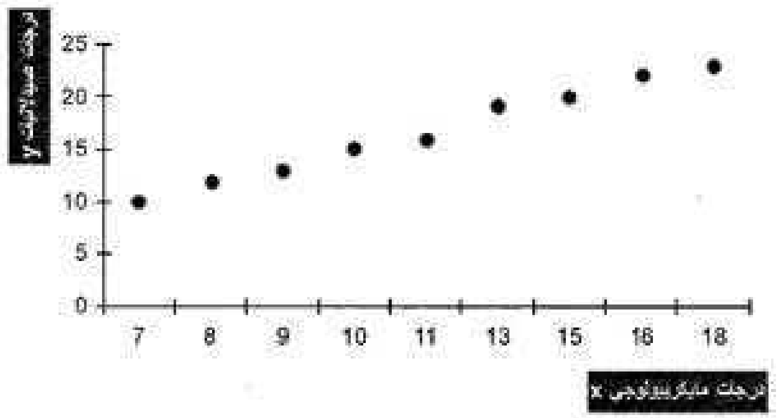
درجات مايكرو بيولوجي (x)	7	8	9	10	11	13	15	16	18
درجات صيدلانيات (y)	10	12	13	15	16	19	20	22	23

المطلوب :

1. ارسم لوحة الانتشار.
2. اوجد معادلة خط انحدار Y على X .
3. أخذ الطالب في اختبار مايكرو بيولوجي الدرجة 18 وغاب عن اختبار صيدلانيات، ما هي الدرجة التقديرية التي يحصل عليها الطالب في اختبار صيدلانيات .
4. ما هو الخطأ في التقدير لدرجة الطالب في اختبار صيدلانيات إذا حصل على 10 درجات في اختبار مايكرو بيولوجي؟

الحل

1. نرسم لوحة الانتشار برصد النقاط (x,y) المعطاة في الجدول كما يظهر في الشكل الآتي :



2. نجري الحسابات كما في الجدول الآتي :

N	X	Y	X . Y	
1	7	10	70	49
2	8	12	96	64
3	9	13	117	81
4	10	15	150	100
5	11	16	176	121
6	13	19	247	169
7	15	20	300	225
8	16	22	352	256
9	18	23	414	324
المجموع	107	150	1922	1389

من الجدول نجد :

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{107}{9} = 11.9$$

$$\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{150}{9} = 16.7$$

لإيجاد قيمة $\hat{b}$ نعوض في المعادلة :

$$\hat{b} = \frac{\sum x_i y_i - n(\bar{x})(\bar{y})}{\sum x_i^2 - n\bar{x}^2} = \frac{1922-(9)(11.9)(16.7)}{1389-(9)(11.9)^2} = \frac{133.43}{114.51} = 1.165$$

لإيجاد قيمة $\hat{a}$ نعوض في المعادلة :

$$\hat{a} = \bar{y} - \hat{b}\bar{x} = 16.7-(1.165)(11.9) = 2.837$$

وبالتعويض بقيمة $\hat{a}$ ، $\hat{b}$ في معادلة خط الانحدار

$$\hat{y} = \hat{a} + \hat{b}x$$
$$\hat{y} = 2.837 + 1.165x$$

وهذه هي معادلة خط انحدار Y على X .

3. نعوض قيمة x = 18 في معادلة خط الانحدار فنجد العلامة التقديرية في الامتحان الثاني :

$$\hat{y} = 2.837 + 1.165x = 2.837 + (1.165)(18) = 23.807$$

4. نعوض قيمة x = 10 في معادلة الانحدار فنجد

$$\hat{y} = 2.837 + 1.165x = 2.837 + (1.165)(10) = 14.487$$

ويكون الخطأ في التقدير

$$\hat{e} = y_i - \hat{y} = 15 - 14.487 = 0.513$$

❖ خواص معادلة الانحدار:

أ. يمر خط الانحدار المقدر بالنقطتين $(\bar{x}, \bar{y})$.

ب. مجموع انحرافات المشاهدات عن خط الانحدار = صفر ، أي أن ،

$$\sum e_i = \sum (y_i - \hat{y}) = 0$$

ج. مجموع مربعات انحرافات المشاهدات عن خط الانحدار المقدر أقل ما يمكن ،

أي أن ، $\min \sum e_i^2 = \sum (y_i - \hat{y})^2$

مثال (2)

يبين الجدول الآتي الأرباح والإنفاق على الاستثمار لعدد خمس شركات أدوية (بملايين الريالات)

الأرباح X	10	8	5	9	8
الإنفاق الاستثماري Y	3	2	1	3	2

باستخدام معادلة المربعات الصغرى ، أوجد معادلة الانحدار الخطي البسيط Y على X

ثم أوجد قيمة كل من : $\sum e_i$ ، $\sum e_i^2$

الحل

نكون الجدول الآتي :

No.	الأرباح x	الإنفاق y	x .y				
1	10	3	30	100	2.86	0.14	0.0196
2	8	2	16	64	2	0	0
3	5	1	5	25	0.71	0.29	0.0841
4	9	3	27	81	2.43	0.57	0.3249
5	8	1	8	64	2	-1	1
	40	10	86	334	10	0	1.43

أ. إيجاد معادلة الانحدار الخطي البسيط :

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{40}{5} = 8 \quad , \quad \overline{y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{10}{5} = 2$$

$$\hat{b} = \frac{\sum x_i y_i - n(\overline{x})(\overline{y})}{\sum x_i^2 - n\overline{x}} = \frac{86-(5)(8)(2)}{334-5(8)^2} = 0.43$$

$$\hat{a} = \overline{y} - \hat{b}\overline{x} = 2-(0.43)(8) = -1.44$$

وتكون معادلة الانحدار الخطي البسيط هي :

$$\hat{y} = -1.44 + 0.43x$$

ب. من الجدول السابق يتضح أن :

$$\sum e_i = \sum (y_i - \hat{y}) = 0$$

لأنها تمثل انحرافات مجموعة من القيم حول متوسطها ، ممثلاً في خط الانحدار .

اما مجموع مربع الأخطاء فهي :

$$\sum e_i^2 = \sum (y_i - \hat{y})^2 = 1.43$$

نشاط (1)

من البيانات الواردة في الجدول الآتي اوجد :

No.	X	Y	X . Y	
1	4	12		
2	6	15		
3	8	11		
4	5	13		
5	9	14		
6	2	8		
المجموع				

أ. معادلة انحدار y على x

ب. قدر قيمة y عندما تكون x = 5

ج. احسب الخطأ في التقدير عندما x = 5

الحل:...

تدريب (1)

من البيانات الواردة في الجدول الآتي اوجد:

No.	X	Y	X . Y	
1	3	8		
2	4	7		
3	5	6		
4	6	5		
5	7	4		
6	8	3		
المجموع				

ا. معادلة انحدار y على x

ب. قدر قيمة y عندما تكون x = 7

ج. احسب الخطأ في التقدير عندما x = 7

تمارين الفصل الخامس

س1 / ضع دائرة حول رمز الإجابة الصحيحة فيما يلي:

1. معامل ارتباط بيرسون يقع بين:

ا. [1,0]   ب. [-1,1]   ج. [-1,0]   د. خلاف ذلك وهو:

2. إذا كانت العلاقة بين Y,X علاقة طردية تامة فإن معامل الارتباط بين Y,X هو:

ا. -1   ب. 1   ج. 0   د. خلاف ذلك وهو:

3. قام أحد الباحثين بدراسة العلاقة بين التدخين والتهاب اللثة فجمع بيانات عن 200 شخص فكان توزيعهم حسب الصفات كما يلي:

A / B	مدخن	غير مدخن	المجموع
مصاب	700	100	800
غير مصاب	200	1000	1200
المجموع	900	1100	2000

معامل الاقتران للبيانات في الجدول اعلاه يساوي:

ا. 0.944   ب. 0.49   ج. 1   د. جميع الإجابات خطأ

4. إذا انطبق خطا معادلتي الانحدار على بعضهما بعضا فإن العلاقة بينهما تكون:

ا. صفرية   ب. لا علاقة بينهما   ج. ناقصة   د. تامة

5. إذا كانت قيم X (1، 2، 3، 4، 5) وقيم y (3، 6، 9، 12، 15) فإن الارتباط بين X و y هو:

ا. -1   ب. 0.12   ج. 0.10   د. خلاف ذلك وهو:

6. عندما تكون قيمة معامل الارتباط تساوي 1 فإن العلاقة بين المتغيرين تكون:

ا. طردية تامة   ب. طردية غير تامة   ج. ناقصة   د. صفرية

* في دراسة لمعرفة العلاقة بين النفقات على الشركات الدوائية (X)، وحجم الإنتاج الدوائي (Y)، إذا توفرت لديك البيانات الآتية:

$$\sum Xi = 77, \overline{X} = 11, \sum Yi = 35, \sum XiYi = 436, \sum Y_i^2 = 203, \sum X^2 = 941$$

1. يكون عدد المشاهدات n هو:

أ. 7   ب. 9   ج. 10   د. خلاف ذلك وهو:

2. تكون قيمة الوسط الحسابي $\overline{y}$ هي:

أ. 4   ب. 5   ج. 6   د. خلاف ذلك وهو:

3. تكون قيمة معامل الارتباط r بين المتغيرين هو:

أ. 0.60   ب. 0.90   ج. 0.99   د. خلاف ذلك وهو:

4. تشير قيمة r أعلاه إلى أن العلاقة بين النفقات والإنتاج هي:

أ. عكسية ضعيفة   ب. عكسية قوية   ج. طردية قوية   د. خلاف ذلك وهو:

س2 / الجدول الآتي يوضح أن 10 طلاب مرتبين ترتيباً أبجدياً حسب مستوى أدائهم في كل من (الجانب العملي والمحاضرات) في مادة البيولوجي.

العملي (X)	7	3	5	8	2	9	8	5	4	6
المحاضرات (Y)	9	4	10	6	7	5	3	7	2	10

احسب معامل الارتباط المناسب.

س3 / إذا كانت درجات الحرارة والضغط الجوي لخمس مواسم ساعة الظهيرة في أحد الأيام كما يلي:

درجات الحرارة (X)	22	18	15	10	7
الضغط الجوي (Y)	65	70	72	75	78

- اوجد معامل ارتباط سبيرمان بين المتغيرين.
- اوجد معامل ارتباط بيرسون بين المتغيرين.
- حدد نوع العلاقة وقوتها بين المتغيرين.

س4 / من البيانات الواردة في الجدول الآتي احسب معامل الارتباط بين المتغيرين X , Y

n	X	Y	X y	
1	89	95		
2	90	88		
3	78	76		
4	67	70		
5	77	75		
6	98	97		
المجموع				

س5 / من دراسة العلاقة بين المتغيرين X, Y وجدنا البيانات الآتية:

$$n = 4, \quad \sum x = 20, \quad \sum x^2 = 120, \quad \sum y = 16,$$
$$\sum y^2 = 84, \quad \sum xy = 100, \quad \sum D^2 = 0$$

1. اوجد معامل الارتباط الخطي بين x و y (بيرسون).
2. اوجد معامل ارتباط سبيرمان للرتب بين x و y.
3. اوجد معادلة الانحدار الخطي بين x و y.
4. اوجد القيمة التقديرية لـ y عندما x=5.
5. اوجد معامل الاختلاف لقيم كل من x و y.
6. اي البيانات أكثر تشتتاً قيم X او قيم Y ؟ ولماذا ؟

أسئلة تقييمية شاملة

س1 / ضع علامة (✓) أو (×) أمام العبارات الآتية مع تصحيح الخطأ:

1. ( ) الحالة الاجتماعية متغير نوعي ويقاس بمقياس ترتيبي.
2. ( ) مركز الفئة = الحد الأعلى للفئة – الحد الأدنى للفئة.
3. ( ) يتأثر المنوال بالقيم المتطرفة.
4. ( ) في التوزيع السالب للالتواء يكون المنوال أكبر من الوسيط والمتوسط الحسابي.
5. ( ) المئين الخامس والسبعين يدعى بالربيع الثالث.
6. ( ) إذا كان معامل الارتباط (r = 0) فإنه لا يوجد ارتباط.
7. ( ) معامل التوافق يقيس العلاقة بين متغيرين لكل منهما تصنيفين أو أكثر.
8. ( ) معدل الولادة = عدد المواليد أحياء ÷ العدد الكلي للسكان .
9. ( ) الإحصاءات الحيوية: هي مجموعة الحوادث والأحداث التي تصيب الإنسان منذ لحظة ولادته حتى لحظة وفاته.
10. ( ) في التوزيع المتماثل يمكن حساب الوسط الحسابي بيانياً.
11. ( ) يدل الصفر على انعدام الخاصية في المقياس الفئوي .
12. ( ) يستخدم معامل بيرسون لقياس العلاقة بين متغيرين تم قياسهما بمقياس رتبي.
13. ( ) تكون العلاقة طردية عندما يكون معامل الارتباط سالب.
14. ( ) يُمكن استخدام معامل ارتباط بيرسون إذا كانت المتغيرات نوعية.
15. ( ) المدى يعتمد في حسابه على جميع البيانات.
16. ( ) إذا كان للبيانات منوالين فإن التوزيع احادي القمة.
17. ( ) يُمكن حساب الانحراف المعياري للجداول التكرارية المفتوحة.
18. ( ) يُمكن حساب المنوال للبيانات النوعية.
19. ( ) يُمكن حساب المدى من جداول التوزيع التكراري المفتوحة.
20. ( ) في التوزيع غير المتماثل يُمكن حساب الوسط الحسابي بيانياً.
21. ( ) يتأثر المنوال بالقيم الوسطى ولا يتأثر بالقيم المتطرفة.
22. ( ) التباين = جذر الانحراف المعياري.
23. ( ) في الإحصاء الاستدلالي يُمكن تعميم النتائج.
24. ( ) يُمكن حساب الوسيط لبيانات نوعية.

س2 / من جدول التوزيع التكراري الآتي أجب عن الفقرات التي تليه:

الفئة	التكرار f	مركز الفئة x	f . x				ت.م.ص
5 – 9	9	7	63	-13.068	117.612	1536.95	9
10 – 14	7	12	84	-8.068	56.476	455.65	16
15 – 19	5	17	85	-3.068	15.340	47.06	21
20 – 24	10	22	220	1.932	19.320	37.33	31
25 – 29	3	27	81	6.932	20.796	144.16	34
30 – 34	4	32	128	11.932	47.728	569.49	38
35 – 39	6	37	222	16.932	101.592	1720.16	44
المجموع	44		883		378.864	4510.8	

1. المنوال للبيانات في الجدول السابق يساوي:

أ. 20   ب. 8.6   ج. 102.52   د. 22

2. الانحراف المتوسط للبيانات في الجدول السابق يساوي:

أ. 20   ب. 22   ج. 102.52   د. 8.6

3. التباين للبيانات في الجدول السابق يساوي:

أ. 22   ب. 20   ج. 10.125   د. 102.52

4. الوسيط للبيانات في الجدول السابق يساوي:

أ. 20   ب. 22   ج. 10.125   د. 102.52

5. الانحراف المعياري للبيانات في الجدول السابق يساوي:

أ. 10.125   ب. 20   ج. 102.52   د. 34

6. المدى للبيانات في الجدول السابق يساوي:

أ. 34   ب. 20   ج. 10.125   د. 0.19-

7. معامل بيرسون للالتواء (SK) من البيانات السابقة في الجدول يساوي:

أ. 19-   ب. 34   ج. 20   د. خلاف ذلك وهو.

8. المتوسط الحسابي للبيانات في الجدول تساوي:

أ. 8.6   ب. 102.52   ج. 22   د. 20.068

9. درجة تدبب قمة منحنى التوزيع التكراري تعرف ب:

أ. التفلطح   ب. الانحدار   ج. الارتباط   د. الالتواء

س3 / من الجدول السابق في (السؤال الثاني)

1. أوجد الوسيط بيانياً للبيانات الواردة وقارن النتيجة التي حصلت عليها مع قيمة الوسيط جبرياً..

ماذا تلاحظ؟

2. أوجد المنوال بيانياً.

3. علل.. من عيوب معامل (بيرسون) للالتواء: احتواءه على المنوال.

س4 / ضع خط تحت الإجابة الصحيحة لكل من الأسئلة الآتية:

1) من ميادين علم الإحصاء العلوم:

أ- الطبيعية .   ب- الطبية .   ج- الاقتصادية .   د- جميع ما سبق .

2) جزء من الإحصاء يهتم بجمع البيانات وتنظيمها وتصنيفها وعرضها عن طريق الجداول أو الرسوم البيانية وغيرها يعرف بالإحصاء:

أ- الوصفي .   ب- الاستدلالي .   ج- الطبي .   د- جميع الإجابات خاطئة .

3) فصيلة الدم في جسم الإنسان تعد من البيانات:

أ- الكمية .   ب- النوعية .   ج- الكمية والنوعية معا .   د- ليس ما سبق .

4) كل الخصائص الإحصائية التي تحصل عليها من العينة الإحصائية تعرف ب:

أ- المعاينة .   ب- المتغيرات .   ج- المعالم .   د- التقديرات .

5) عدد المرضى الذين يدخلون المستشفى في اليوم الواحد متغير:

أ- منفصل .   ب- متصل .   ج- (أ+ب) .   د- جميع ما سبق خطأ .

6) درجة حرارة المريض متغير:

أ- منفصل .   ب- متصل .   ج- (أ+ب) .   د – ليس ما سبق .

7) مجموع الملاحظات (البيانات) مقسوماً على عددها يعرف ب:

أ- المتوسط الحسابي   ب- الوسيط   ج- المنوال   د- المدى .

8) القيمة الأكثر تكراراً في البيانات الإحصائية تعرف ب:

أ- المتوسط الحسابي   ب- الوسيط   ج- المنوال   د- المدى .

9) القيمة التي تتوسط البيانات بعد ترتيبها تصاعدياً أو تنازلياً تعرف ب:

أ- المتوسط الحسابي   ب- الوسيط   ج- المنوال   د- المدى .

10) تباعد أو تبعثر القيم في التوزيع عن بعضها البعض يعرف ب :

أ- التمركز   ب- النزعة المركزية   ج- التشتت   د- الانحراف المعياري .

11) من مقاييس النزعة المركزية ما يلي عدا:

أ- المنوال .   ب- المدى .   ج- الوسيط .   د- المتوسط الحسابي .

12) من مقاييس التشتت التالي عدا واحد :

أ- المنوال .   ب- المدى .   ج- التباين .   د- الانحراف المتوسط .

13) إذا كان عدد المترددين على عيادة خلال أسبوع كما يلي :

(7 ، 9 ، 15 ، 14 ، 18 ، 5 ، 2) فإن المتوسط الحسابي لهم يساوي :

أ- 5   ب- 7   ج- 9   د- 10 .

14) المدى لهذه البيانات (3 ، 4 ، 5 ، 85 ، 12 ، 8 ، 9 ، 25) يساوي :

أ- 3   ب- 19   ج- 8.5   د- 82 .

15) الوسيط لهذه البيانات (9 ، 8 ، 10 ، 3 ، 15 ، 6 ، 7 ، 3 ، 11) يساوي:

أ- 3   ب- 8   ج- 9   د- 7 .

16) المنوال لهذه البيانات (10.9.7.8.6.5.2.4.3) يساوي :

أ- 6   ب- 3   ج- 9   د- عديمة المنوال .

17) الوسيط للبيانات التالية:(4.3.11.3.7.6.5.3.10.8.5.4) يساوي:

أ- 3   ب- 8   ج- 5   د- 4 .

18) المنوال للبيانات التالية :(10.9.7.2.2.2.6.8.8.8.6.5.2.4.3) يساوي :

أ- 8   ب- 2   ج- 4.4   د- عديمة المنوال .

19) المنوال للبيانات التالية :(10.3.3.3.3.10.10.10.10.3) يساوي :

أ- 10   ب- 3   ج- لدينا منوالان 3 ، 10   د- عديمة المنوال .

20) المنوال للبيانات التالية :(10.9.7.8.6.6.6.7.8.7.9.10) يساوي :

أ- 6   ب- 7   ج- 6.5   د- لدينا منوالان 6، 7 .

21) إذا كان لدينا مجموعة من القيم وسطها الحسابي (20) ومجموعها (160) فإن عدد عناصرها يكون :

أ- 16   ب- 8   ج- 32   د- 80

22) مركز الفئة يساوي :

أ- الحد الأعلى – الحد الأدنى   ب- الحد الأعلى + الحد الأدنى/2

ج- الحد الأعلى + الحد الأدنى   د- الحد الأعلى – الحد الأدنى/2

23) إذا كان الانحراف المعياري لمجموعة من القيم يساوي (9) فإن التباين للقيم المذكورة يكون :

أ- 3   ب- 18   ج- 81   د- 9

24) إذا كانت مجموعة من القيم وسطها الحسابي (16) والتباين لها (3) فإن الانحراف المعياري لها يساوي :

أ- 1.73   ب- 4   ج- 1.5 ،   د- 9

25) السمة أو الصفة أو الكمية التي تتغير قيمتها من عنصر إلى آخر ومن مشاهدة إلى أخرى تعرف ب:

أ- الثابت   ب المتغير   ج- المتغير المنفصل   د- ليس ما سبق .

26) كيفية اختيار عينة البحث من مجتمع البحث تعرف ب :

أ- المعالم   ب- التقديرات   ج المعاينة   د- المتغير المنفصل

27) المتغير الذي يأخذ قيما قابلة للعد (محدودة) يعرف ب :

أ- المتغير المتصل   ب- المتغير المنفصل   ج- المتغير   د- لا شئ ما سبق.

28) كل الخصائص الإحصائية التي نحصل عليها من المجتمع الإحصائي تعرف ب :

أ- المعاينة   ب- المعالم   ج- التقديرات ،   د- مجتمع البحث .

29) مجموع التكرار النسبي يساوي :

أ- تكرار الفئة / مجموع التكرارات   ب- مجموع التكرارات   ج- 1   د- 100

30) المقياس الجيد للتشتت يجب أن يتضمن الآتي عدا واحدة :

أ- يجب أن يكون معرفاً بوضوح تام .

ب- يجب أن يتوافق مع القواعد العامة الرياضية .

ج- يجب أن يكون معتمداً على جميع البيانات .

د- يجب أن يخضع لحسابات معقدة ومملة .

31) من خصائص المدى :

أ- يتأثر بالقيم الحدية.   ب- يهمل الإشارات الجبرية .

ج- يمكن حسابه في التوزيعات التكرارية المفتوحة .   د- صعوبة حسابه .

س5/ من جدول التوزيع التكراري الآتي :

الفئة	الفئة الفعلية	مركز الفئة X	F التكرار	X .F						ت.م.ص
11 - 15			3							
16 – 20			7							
21 – 25			15							
26 – 30			5							
المجموع			30							

◆ اكمل الجدول أعلاه ثم أجب عن الفقرات الآتية :

1) المتوسط الحسابي للبيانات في الجدول يساوي :

أ- 23 ب- 21.7 ج- 22.17 د- 19

2) المنوال للبيانات في الجدول يساوي :

أ- 19 ب- 22.17 ج- 23 د- 3.4

3) الوسيط للبيانات في الجدول يساوي:

أ- 22.17 ب- 19 ج- 22.17 د- 23

4) المدى للبيانات في الجدول يساوي:

أ- 18.85 ب- 3.4 ج- 23 د- 19

5) الانحراف المتوسط للبيانات في الجدول يساوي :

أ- 19 ب- 3.4 ج- 18.85 د- 4.43

6) التباين للبيانات في الجدول يساوي :

أ- 4.43 ب- 3.4 ج- 23 د- 18.85

7) الانحراف المعياري للبيانات في الجدول يساوي :

أ- 18.85 ب- 4.34 ج- 19 د- 23

◆ من الجدول أعلاه :

1- ارسم المدرج التكراري والمضلع التكراري والمنحنى التكراري.

2- اوجد الوسيط بيانياً وقارن النتيجة مع النتيجة التي حصلت عليها جبرياً. ماذا تلاحظ ؟

س6 / من جدول التوزيع التكراري التالي :

الملاحظةX	2	4	6	8	10	المجموع
التكرار f	5	3	4	6	2	20

1) المتوسط الحسابي للبيانات في الجدول يساوي :

أ- 7.7 ب- 5.7 ج- 8 د- 2.8

2) المنوال للبيانات في الجدول يساوي :

أ- 6 ب- 8 ج- 7.3 د- 3.4

3) الوسيط للبيانات في الجدول يساوي :

أ- 6 ب- 9 ج- 7.3 د- 8

4) المدى للبيانات في الجدول يساوي :

أ- 4 ب- 6 ج- 8 د- 7.3

5) الانحراف المتوسط للبيانات في الجدول يساوي :

أ- 7.3 ب- 4 ج- 5 د- جميع الإجابات خطأ.

6) التباين للبيانات في الجدول يساوي :

أ- 8 ب- 7.3 ج- 4 د- جميع الإجابات خطأ.

7) الانحراف المعياري للبيانات في الجدول يساوي :

أ- 7.3 ب- 2.7 ج- 8 د- لا شيء مما سبق

◆ من جدول التوزيع التكراري أعلاه :

أ. اوجد الوسيط بيانياً للبيانات الواردة وقارن النتيجة التي حصلت عليها مع قيمة الوسيط جبرياً.. ماذا تلاحظ؟

ب. اوجد المنوال بيانياً.

ج . مثل البيانات بالأعمدة البيانية.

س7 / ضع دائرة حول رمز الإجابة الصحيحة فيما يلي:

1. الجدول المنتظم:

أ. أطوال فئاته غير متساوية   ب. حدوده العليا والدنيا معلومة

ج. أطوال فئاته متساوية   د. جميع ما سبق.

2. للفئة (6- 4) فإن الحدود الفعلية:

أ. 6.5- 4.6   ب. 6.5- 3.5   ج. 5.5- 3.5   د. 5.5- 4.5

3. للفئة (11- 9) فإن مركز الفئة:

أ. 9   ب. 10   ج. 11   د. 20

4. إذا كان تكرار الفئة = 2 ومجموع التكرارات = 10 فإن التكرار المئوي لها هو:

أ. 20٪   ب. 30٪   ج. 40٪   د. 50٪

5. إذا كانت جامعة تحتوي عدداً من الكليات وأردنا سحب عينة ممثلة لهذه الكليات فأفضل عينة هي:

أ. المنتظمة   ب. الطبقية   ج. البسيطة   د. لاشيء مما ذكر

6. من عيوب المسح الشامل:

أ. تلف الوحدات المستخدمة في التجربة في بعض الحالات

ب. ارتفاع التكلفة

ج. عدم إمكانية تطبيقه في بعض الحالات

د. جميع ما ذكر

7. إذا كان عدد مفردات المجتمع الإحصائي 1000، وعدد مفردات طبقة الإناث منهم 360 مفردة، وأردنا سحب عينة حجمها 100 مفردة ممثلة للمجتمع بما يتناسب مع عدد مفردات كل طبقة فإن عدد من يمثلون طبقة الإناث في العينة هو:

أ. 100   ب. 36   ج. 360   د. 1000

8. أرقام السيارات متغير:

أ. كمي نسبي   ب. كمي متصل   ج. اسمي منفصل   د. كمي فئوي

9. الرتب العسكرية متغير:

أ. كمي متصل   ب. كمي فئوي   ج. ترتيبي   د. نسبي

10. إذا كان المدى = 20، وعدد الفئات = 5، فإن طول الفئة:

أ. 2   ب. 10   ج. 3   د. 4

11. إذا كان عدد مفردات المجتمع الإحصائي = 1000 مفردة، وحجم الطبقة أ = 250 مفردة، فإن زاوية القطاع لهذه الطبقة هي:

أ. °250   ب. °90   ج. °100   د. °360

12. إذا كانت لديك البيانات التالية (5، 4، 3، 2، 1) فإن الوسط الحسابي لها =

أ. 4   ب. 2   ج. 3   د. 2.5

13. للبيانات (1، 2، 4، 7، 6) فإن الوسيط هو:

أ. 2   ب. 4   ج. 6   د. 5

14. للبيانات (12، 10، 8، 6، 4، 2) فإن الوسيط =

أ. 8   ب. 6   ج. 7   د. 7.5

15. إذا كان الوسط الحسابي لمجموعة من البيانات = 10، وأضفنا إلى جميع قيم التوزيع العدد 5 فإن الوسط الحسابي الجديد =

أ. 15   ب. 5   ج. 10   د. 25

16. مجموع انحرافات القيم عن وسطها الحسابي =

أ. 0   ب. 5   ج. 3   د. 1

17. إذا كان الوسط الحسابي = 5، والوسيط = 4، فإن المنوال =

أ. 3   ب. 4   ج. 2   د. لا شيء مما ذكر

18. إذا كان التوزيع ثنائي القمة فهذا يعني أن له:

أ. منوال واحد   ب. منوالان   ج. 3 منوالات   د. لا يوجد منوال

19. إذا علمت أن قيمة الوسط الحسابي = 5، وقيمة الوسيط = 10، فإن قيمة الربيع الثاني =

أ. 5   ب. 15   ج. 10   د. 50

20. في حالة الفئات المفتوحة واحد من هذه لا يصلح كمقياس من مقاييس النزعة المركزية:

أ. الوسيط   ب. الوسط الحسابي   ج. المنوال   د. التباين

21. إذا كانت قيمة الوسط الحسابي 25، وقيمة الوسيط = 10، فإن قيمة العشير الخامس =

أ. 7   ب. 3   ج. 4   د. 10

22. إذا كان عدد مفردات المجتمع = 500، وأردنا سحب عينة منتظمة منه عدد مفرداتها 10، وكان رقم المفردة الأولى = 2، فإن رقم المفردة الرابعة =

أ. 2   ب. 152   ج. 22   د. 52

23. عند تمثيل المدرج التكراري فإننا نمثل على المحور الأفقي:

أ. مراكز الفئات   ب. الحدود الفعلية للفئات   ج. الفئات   د. التكرار

24. مجموع التكرارات النسبية =

أ. 10   ب. 1   ج. مجموع التكرارات   د. 0

25. للفئة (60,5 - 70,5) فإن الحد الأعلى الفعلي:

أ. 61   ب. 70.5   ج. 71   د. 70

26. إذا أضفنا إلى جميع قيم توزيع وسطه = 50 العدد (7 -) فإن الوسط الحسابي الجديد له =

أ. 57   ب. 50   ج. 43   د. لا شيء مما ذكر

27. إذا كان حجم المجتمع الإحصائي 1000 مفردة، وحجم فئة منه 100 مفردة، فإن زاوية القطاع لها=

أ. 90°   ب. 36°   ج. 45°   د. 100°

28. للبيانات (6، 5، 4، 3، 2، 1) فإن المنوال =

أ. 3   ب. 4   ج. 3.5   د. لا يوجد منوال

29. إذا علمت أن المنحنى متماثل وأن الوسط الحسابي = 5، فإن المنوال =

أ. 4   ب. 5   ج. 6   د. لا يمكن تحديده

30. إذا علمت أن المنحنى سالب الالتواء وأن الوسط الحسابي = 5، فإن المنوال =

أ. 4   ب. 5   ج. 6   د. لا يمكن تحديده

31. الربيع الثالث هو القيمة التي يسبقها:

أ. 25٪ من البيانات   ب. 50٪ من البيانات   ج. 75٪ من البيانات   د. لا شيء مما ذكر

32. إذا كان الوسيط = 20، والمنوال = 15، فإن المئين 50 =

أ. 20   ب. 15   ج. 32   د. لا يمكن تحديده

33. إذا كان لديك البيانات (0، 4، 3، 1) فإن الانحراف المتوسط =

أ. 1   ب. 1.5   ج. 2   د. 4

34. إذا كان الانحراف المعياري = 2، الانحراف المتوسط = 5، فإن التباين =

أ. 4   ب. 25   ج. 1   د. 2.5

35. واحد من هذه ليس من مقاييس التشتت:

أ. المدى الربيعي   ب. الانحراف المتوسط   ج. التباين   د. الربيع الثاني

36. إذا كان تباين توزيع ما = 16، وأضفنا إلى جميع قيم التوزيع العدد 10، فإن الانحراف المعياري الجديد =

أ. 4   ب. 14   ج. 26   د. 116

37. التباين للبيانات (5، 5، 5، 5، 5، 5، 5) هو:

أ. 5   ب. 0   ج. 5-   د. 1

38. أرقام قاعات التدريس هي متغير:

أ. اسمي منفصل   ب. اسمي متصل   ج. كمي نسبي   د. فئوي

39. كلما كان حجم العينة أكبر كلما كان الخطأ المعياري:

أ. أصغر   ب. أكبر   ج. ثابت   د. لا علاقة بينهما

40. روضة مختلطة عدد طلابها 180 طالباً وطالبة، عدد الذكور فيها 30 طالباً، فإن زاوية القطاع الدائري التي تمثل الطالبات هي:

أ. 300°   ب. 120°   ج. 90°   د. 60°

41. الوسط الفرضي لأي جدول تكراري يساوي:

أ. مركز الفئة   ب. أكبر تكرار   ج. مركز الفئة المقابل لأكبر تكرار   د. مجموع التكرارات

42. إذا كان الوسط الحسابي لتوزيع أحادي المنوال = 60، والوسيط = 50، فإن المنوال يساوي:

أ. 50   ب. 40   ج. 20   د. 30

43. عندما حسبنا معامل الارتباط بين متغيرين (x,y) وجدناه يساوي - 0.90 فهذه العلاقة بين متغيرين هي علاقة:

أ. طردية قوية   ب. طردية ضعيفة   ج. عكسية قوية   د. عكسية ضعيفة

44. إذا كان الانحراف المعياري لمجموعة من المشاهدات = 1.3، فإن التباين يساوي:

أ. 1.69   ب. 1.3   ج. 16.9   د. 0.12

45. إذا كان المتوسط الحسابي لخمس قيم = 4 فإن مجموع هذه القيم يساوي:

أ. 18 ب. 19 ج. 20 د. 21

46. إذا كان مجموع مربعات 8 قيم هي 256، وكان مجموع هذه القيم يساوي 32، فإن الانحراف المعياري لهذه القيم يساوي:

أ. 2 ب. 4 ج. 6 د. 8

47. الإحصائي الذي يسبقه ويليه قسمان من القيم يسمى:

أ. وسيطاً ب. منوالاً ج. وسطاً حسابياً د. وسطاً افتراضياً

48. أحد الاختيارات التالية لا يعتبر من مقاييس التشتت:

أ. الانحراف المتوسط ب. الانحراف المعياري ج. الوسيط د. المدى

49. لدينا مجموعة من القيم المجهولة لعناصر متوسطها يساوي 5، وعددها يساوي 10، فإن مجموع عناصر هذه القيم يساوي:

أ. 15 ب. 50 ج. 2 د. 7.5

50. رتبت علامات ثمانية طلاب تصاعدياً (9، 10، 12، 13، 14، 15، 17، 18)، فالوسيط لهذه الدرجات هو:

أ. 13.5 ب. 13 ج. 14 د. 27

51. قيم عددها (7)، ومجموعها (35)، ومجموع مربعات الانحراف عن وسطها الحسابي يساوي (28)، فإن الانحراف المعياري لهذه القيم يساوي:

أ. 20 ب. 0 ج. 2 د. جميع الإجابات خطأ.

52. مجموع المتقدمين للامتحان الشامل تموز يساوي 16000 طالباً، ومجموع طلبة البرنامج الصحي 4000 طالباً، فإن زاوية القطاع الدائري للبرنامج الصحي تساوي:

أ. °63 ب. °27 ج. °90 د. °270

53. المدى هو أبسط مقاييس:

أ. الاحتمالات ب. التشتت ج. النزعة المركزية د. العينات

54. يصنف الوسيط ضمن مقاييس:

أ. الاحتمالات ب. التشتت ج. النزعة المركزية د. العينات.

55. إذا وصل معامل الارتباط بين متغيرين إلى 0.80 فيكون هذا الارتباط:

أ. مقبولاً ب. معدوماً ج. مرفوضاً د. عالياً

56. في دراسة لأثر التدخين على أعصاب الجراحين والتحكم في العمليات الجراحية، فالعينة التي يتم اختيارها من هذا المجتمع تكون:

أ. طبقية انتقائية ب. عشوائية طبقية ج. عشوائية انتقائية د. منتظمة

57. كلية طبية فيها 6000 طالب، منهم 2400 طالب في سنة أولى، و 1800 طالب في سنة ثانية، و 1400 طالب في سنة ثالثة، و 400 طالب في سنة رابعة، أردنا سحب عينة عشوائية بنسبة 2 % من مجموع الطلبة، فإن حجم عينة السنة الثالثة يساوي:

أ. 48 ب. 28 ج. 36 د. 8

58. مجموع الأفراد الذين نعمم عليهم نتائج الدراسة يمثلون:

أ. المتغير ب. الفئة ج. العينة د. مجتمع الدراسة

59. الحد الأدنى الفعلي للفئة 7– 3 هو:

أ. 3.5 ب. 7.5 ج. 2.5 د. 6.5

60. إذا كان الانحراف المعياري لمجموعة من المشاهدات يساوي 9، فإن التباين يساوي:

أ. 3 ب. 18 ج. 81 د. 9

61. عندما يكون معامل الارتباط ‎-0.85‎ ، فهذا يعني أن درجة العلاقة بين المتغيرين تكون:

أ. معدومة ب. متوسطة ج. قوية د. ضعيف

62. الانحراف المعياري للعلامات التالية: (80، 86، 90، 94، 100) يساوي:

أ. 46.4 ب. 40 ج. 6.32 د. 6.81

63. إذا كان مجموع انحرافات 6 قيم عن الوسط الفرضي 72 يساوي 18، فإن الوسط الحسابي لهذه القيم يساوي:

أ. 69 ب. 78 ج. 75 د. 66

64. الانحراف المتوسط للقيم (1، 2، 3، 4) يساوي:

أ. ‎-1‎ ب. 1.5 ج. ‎-1.5‎ د. 1

65. عند اختيار عينة الدراسة يؤخذ بعين الاعتبار:

أ. انتقاء أفراد عينة الدراسة بدقة ب. اختيار الأفراد المناسبين

ج. تكافؤ الفرص لجميع عناصر مجتمع الدراسة د. اختيار عدد كبير من الأفراد

66. يطلق على العينة التي تسمح لكل فرد من أفراد المجتمع نفس الفرصة في الاختيار ليكون أحد أفراد العينة تلك هي العينة:

أ. العشوائية ب. المنتظمة ج. الطبقية د. متعددة المراحل

67. مستوى القياس الذي تستخدم فيه الأرقام لغايات التصنيف فقط المستوى:

أ. الاسمي ب. الرتبي ج. الفئوي د. النسبي

68. من خلال الجدول التالي أجب عن الأسئلة التي تليه:

التكرار	الفئات
2	50 – 54
3	55 – 59
1	60 – 64
1	65 – 69
2	70 – 74
1	75 – 79

1. الحد الأعلى الفعلي للفئة الثالثة هو:

أ. 65 ب. 69 ج. 64.5 د. 69.5

2. الحد الفعلي الأعلى للفئة الخامسة هو:

أ. 70 ب. 74 ج. 69.5 د. 74.5

3. مركز الفئة الرابعة هو:

أ. 65 ب. 66 ج. 67 د. 68

4. التكرار المتجمع الصاعد للفئة السادسة هو:

أ. 10 ب. 11 ج. 12 د. 13

5. طول الفئة من (59 – 55):

أ. 4 ب. 5 ج. 6 د. 3

69. الوسط الحسابي لعلامات أحد الطلبة الآتية (60، 70، 50، 40، 30) هو:

أ. 40 ب. 50 ج. 55 د. 60

70. إذا كان معدل درجات الحرارة للبيانات التالية (30، 20، 15، -15 ، 0، -10 ، -5 ، x) ووسطها الحسابي يساوي 5، فإن قيمة x تساوي:

أ. -15 ب. 0 ج. 15 د. -10

71. المقياس الذي تعتمد قيمته على قيم البيانات جميعها هو:

أ. الوسط الحسابي ب. الوسيط ج. المنوال د. لا شيء مما ذكر

72. إذا كان مقدار الربيع الأعلى 36.5، والربيع الأدنى 20، فإن نصف المدى الربيعي يساوي:

أ. 5.53 ب. 4.25 ج. 6.25 د. 8.25

73. إذا بلغ مجموع 40 مشاهدة 600، وأضيف العدد 5 لكل مشاهدة، فإن الوسط الحسابي بعد التعديل والزيادة يساوي:

أ. 20 ب. 295 ج. 15 د. 100

74. إذا كانت علامة أحد الطلبة في مادة الإحصاء 80، بانحراف معياري مقداره 5 لكل مشاهدة، ووسط حسابي 75، فإن العلامة المعيارية z لمادة الإحصاء هي:

أ. 4 ب. 3 ج. 2 د. 1

75. الانحراف المعياري للأعداد التالية (2، 3، 4، 7) يساوي:

أ. $\sqrt{14}$ ب. $\frac{3}{2}$ ج. $\frac{\sqrt{7}}{2}$ د. $\sqrt{\frac{7}{2}}$

76. إذا كانت قيمة التباين لمجموعة من المشاهدات 49، فإن قيمة الانحراف المعياري لها هي:

أ. 14 ب. 7 ج. 24.5 د. $\sqrt{7}$

77. إذا كان المتوسط الحسابي للقيم (3، 7، 8، x) يساوي 7 ، فإن قيمة x تساوي:

أ. 4 ب. 8 ج. 9 د. 10

78. إذا كان المتوسط الحسابي لعشرة قيم يساوي 15، فإن مجموع القيم يساوي:

أ. 150 ب. 1.5 ج. 75 د. 30

79. إذا كان مجموع علامات 12 طالبا يساوي 216 علامة، وزاد المعلم علامتين لكل طالب فإن الوسط الحسابي لعلاماتهم بعد الزيادة يساوي:

أ. 18   ب. 16   ج. 24   د. 20

80. لحساب معامل ارتباط سبيرمان (الرتب) بين متغيرين لـ 8 أشخاص، وجد أن $\sum d = 20$ ، $\sum d^2 = 50$ ، حيث d الفرق بين الرتب المتناظرة، وبذلك يكون معامل الارتباط مساوياً لـ:

أ. $\frac{25}{42}$ ب. $\frac{17}{42}$ ج. $\frac{8}{42}$ د. $\frac{-8}{42}$

81. العينة التي تمنح كل فرد من أفراد المجتمع نفس الفرصة في الاختيار ليكون أحد أفرادها هي العينة:

أ. الطبقية   ب. العشوائية البسيطة   ج. المنتظمة   د. متعددة المراحل

82. مستوى القياس الذي تكون وحداته متساوية، وليس له صفر مطلق هو:

أ. النسبي   ب. الرتبي   ج. الفئوي   د. الاسمي

83. إذا كان مقدار الربيع الأعلى 75، ومقدار الربيع الأدنى 35، فإن قيمة نصف المدى الربيعي هي:

أ. 40   ب. 55   ج. 25   د. 20

84. المحور الأفقي للتكرار المتجمع الصاعد يتكون من:

أ. التكرار المتجمع الصاعد   ب. الحدود العادية للفئات

ج. مراكز الفئات   د. الحدود الفعلية العليا

85. إذا كان الوسط الحسابي لعلامات طلبة صف في مادة الإحصاء 75، بانحراف معياري مقداره 3، وكانت علامة أحد الطلبة في هذه المادة 60، فإن قيمة العلامة المعيارية لهذا الطالب هي:

أ. 5   ب. 3   ج. 5-   د. 3-

86. من المقدار $\sum_{i=1}^{10} d^2 = 100$ فإن قيمة معامل ارتباط الرتب لسبيرمان هو:

أ. 0.90-   ب. 0.99   ج. 0.40-   د. خلاف ذلك وهو:

87. يتم حساب معامل انحدار X/Y باستخدام الصيغة التالية:

أ. $\frac{\sum XY - n\overline{XY}}{\sum Y^2 - n\overline{Y}^2}$ ب. $\frac{\sum (xi-\bar{x})Fi}{\sum X^2 - n\overline{X}^2}$ ج. $\frac{\sum fix^2}{\sum fi}$ د. خلاف ذلك هو:

88. يتم حساب الوسيط بيانياً من خلال:

أ. المنحنى التكراري   ب. المضلع التكراري   ج. المدرج التكراري   د. خلاف ذلك هو:

89. البيانات التالية تمثل أجور عينة من العمال في إحدى المستشفيات خلال أسبوع:

Xi = 10, 8, 16, 11, 8, 7

1. قيمة الوسط الحسابي للأجور هو:

أ. 18   ب. 6   ج. 16   د. خلاف ذلك وهو:

2. قيمة المنوال للأجور هي:

أ. 10   ب. 11   ج. 12   د. خلاف ذلك وهو:

3. قيمة الوسيط للأجور هو:

أ. 9   ب. 10   ج. 11   د. خلاف ذلك وهو:

4. قيمة الانحراف المعياري S للأجور هو:

أ. 4   ب. 5   ج. 6   د. خلاف ذلك وهو:

5. قيمة معامل الاختلاف C.V للأجور هو:

أ. 28%   ب. 30%   ج. 35%   د. خلاف ذلك وهو:

90. في دراسة لمعرفة العلاقة بين الكمية المنتجة Yi من محصول القطن، وكمية السماد Xi المستخدمة، وقد توفرت لديك المعلومات التالية:

$$\sum Xi = 84, \bar{Y} = 3, \sum Yi = 36, \sum XiYi = 465, \sum Y_i^2 = 203, \sum X_i^2 = 941$$

1. يكون عدد المشاهدات n هو:

أ. 11   ب. 12   ج. 14   د. خلاف ذلك وهو:

2. تكون قيمة الوسط الحسابي $\overline{X}$ هي:

أ. 10   ب. 15   ج. 12   د. خلاف ذلك وهو:

3. تكون قيمة معامل الارتباط r هو:

أ. 1.50   ب. 1.00   ج. 0.99-   د. خلاف ذلك وهو:

4. تشير قيمة r أعلاه إلى أن العلاقة بين الكمية المنتجة Y والسماد Xi هي:

أ. عكسية ضعيفة   ب. عكسية قوية   ج. طردية موجبة   د. خلاف ذلك وهو:

91. الجدول الآتي يوضح التوزيع التكراري لقيم الإنفاق الاستهلاكي الشهري لعينة من الأسر:

فئات الاستهلاك	120	-140	-160	-180	200- 220	المجموع
عدد الأسر Fi	4	12	18	10	6	50

ضع دائرة حول رمز الإجابة الصحيحة:

1. الحد الأدنى للفئة المنوالية هو:

أ. 120 ب. 220 ج. 180 د. خلاف ذلك وهو:

2. قيمة المنوال للإنفاق الاستهلاكي هو:

أ. 170 ب. 175 ج. 160 د. خلاف ذلك وهو:

3. قيمة الوسط الحسابي للإنفاق الاستهلاكي $\overline{X}$ هو:

أ. 180 ب. 165 ج. 175 د. خلاف ذلك وهو:

4. موقع (ترتيب) الوسيط للإنفاق الاستهلاكي هو:

أ. 16 ب. 15 ج. 18 د. خلاف ذلك وهو:

5. قيمة الانحراف المعياري S هو:

أ. 20 ب. 10 ج. 15 د. خلاف ذلك وهو:

6. قيمة معامل الإلتواء لبيرسون Y_1 هو:

أ. 0 ب. 0.5- ج. 0.5 د. خلاف ذلك وهو:

7. قيمة معامل الاختلاف C.V هو:

أ. 27% ب. 28% ج. 30% د. خلاف ذلك وهو:

92. المنوال للبيانات غير المبوبة يمثل القيمة:

أ. الأقل قيمة ب. الأكثر قيمة ج. الأكثر تكراراً د. خلاف ذلك وهو:

93. الوسط الحسابي للبيانات غير المبوبة يحسب من العلاقة التالية:

أ. $\overline{X}=\frac{\sum fi}{\sum xi}$ ب. $\overline{X}=\frac{\sum xi}{n}$ ج. $\overline{X}=\frac{\sum xifi}{\sum xi}$ د. خلاف ذلك وهو:

94. من مقاييس النزعة المركزية ما يلي عدا:

أ. المنوال ب. الربيع الثاني ج. المدى د. الوسط الحسابي

95. إذا كانت الفئة (160 – x) إحدى فئات جدول تكراري ومركز هذه الفئة هو 150، فإن قيمة X هي:

أ. 50 ب. 152 ج. 165 د. خلاف ذلك وهو:

96. إذا كان الوسط الحسابي لأعمار 5 طلاب هو 20، والوسط الحسابي لأعمار 3 طلاب آخرين هو 18، فإن الوسط الحسابي لأعمار جميع الطلاب هو:

أ. 20 ب. 21 ج. 19 د. خلاف ذلك وهو:

97. إذا كان المدى يساوي (42)، وأردنا بناء جدول تكراري من (6) فئات، فإن الحد الأعلى لفئة حدها الأدنى 10 هو:

أ. 17 ب. 8 ج. 18 د. خلاف ذلك وهو:

98. قيمة الوسط الحسابي هو:

أ. قيمة تقع في منتصف المفردات ب. قيمة تمثل مركز ثقل المجموعة

ج. قيمة تكرارات أكثر من غيرها من القيم د. خلاف ذلك وهو:

99. التباين يتم احتسابه بالعلاقة التالية:

أ. $\frac{\sum\left(xi^2-fix\right)^2}{\sum xi}$ ب. $\frac{\sum\left(xi-\bar{x}\right)^2 fi}{\sum fi-1}$ ج. $\frac{\sum\left(fi^2-xi\right)^2}{\sum fi-1}$ د. خلاف ذلك وهو:

100. إذا كان حجم المجتمع 800 مفردة، وأردنا اختيار عينة منتظمة بنسبة 5%، فإن كسر المعاينة هو:

أ. 4 ب. 40 ج. 400 د. خلاف ذلك وهو:

101. إذا كانت الفئة (50-X) إحدى فئات جدول تكراري، ومركز هذه الفئة هو 100، فإن قيمة X هي:

أ. 50 ب. 90 ج. 150 د. خلاف ذلك وهو:

102. إذا كان الوسط الحسابي لعلامات 8 طلاب هو 60، والوسط الحسابي لعلامات 5 طلاب آخرين هو 70، فإن الوسط الحسابي لعلامات جميع الطلاب هو:

أ. 5.5 ب. 8.87 ج. 65 د. خلاف ذلك وهو:

103. إذا كان المدى يساوي 70، وأردنا بناء جدول تكراري من 7 فئات فإن الحد الأعلى لفئة حدها الأدنى 20 هو:

أ. 20 ب. 40 ج. 60 د. خلاف ذلك وهو:

104. الوسيط لمجموعة القيم 30, 38, 35, 44, 40, 35 = Xi هو:

أ. 3.5 — ب. 100 — ج. 150 — د. خلاف ذلك وهو:

105. إذا كان مجموع انحرافات 100 مفردة عن وسط فرضي قدره 110 يبلغ -400 فإن الوسط الحسابي هو:

أ. -4 — ب. 4 — ج. 115 — د. خلاف ذلك وهو:

106. مجموع انحرافات 50 مفردة عن وسط حسابي قدره 30 هو:

أ. 100 — ب. 30 — ج. 3 — د. خلاف ذلك وهو:

107. معامل الاختلاف C.V لمجموعة من القيم 6، 3، 0 & 3-، 4- = Xi هو:

أ. 0% — ب. 3% — ج. 30% — د. خلاف ذلك وهو:

108. البيانات التالية لكميات معروضة من علاج ما X والأسعار المناظرة Y:

الكميات X	10	7	5	8	7	5
الأسعار Y	26	18	10	20	15	7

* ضع دائرة حول رمز الإجابة الصحيحة:

1. الوسط الحسابي للكميات $\overline{X}$ هو:

أ. 5 — ب. 4 — ج. 8 — د. خلاف ذلك وهو:

2. الوسط الحسابي للأسعار $\overline{Y}$ هو:

أ. 8 — ب. 5 — ج. 10 — د. خلاف ذلك وهو:

3. معامل الارتباط الخطي لبيرسون r هو:

أ. 1 — ب. -1 — ج. 1.8 — د: خلاف ذلك وهو:

4. تشير قيمة معامل الارتباط إلى ان العلاقة بين الكميات المعروضة X والأسعار Y علاقة:

أ. طردية ضعيفة — ب. عكسية ضعيفة — ج. عكسية قوية — د. خلاف ذلك وهو:

5. معامل الانحدار Y/X هو:

أ. 2 — ب. 1 — ج. 4.5 — د. خلاف ذلك وهو:

6. تقدر الكمية المعروضة X عندما يكون السعر 50 = Y بالقيمة:

أ. 20 — ب. 30 — ج. 40 — د. خلاف ذلك وهو:

7. تقدر الكمية المعروضة Y عندما يكون السعر 30 = X بالقيمة:

أ. 50 — ب. 60 — ج. 70 — د. خلاف ذلك وهو:

109. الوسط الحسابي لمجموعة القيم (9، 16، 8، 10، 13، 10) هو:

أ. 12 — ب. 13 — ج. 14 — د. خلاف ذلك وهو:

110. الوسيط لمجموعة القيم (5، 49، 30، 10، 14، 25، 45) هو:

أ. 25 — ب. 45 — ج. 28 — د. خلاف ذلك وهو:

111. المنوال لمجموعة القيم (35، 10، 28، 19، 28، 14) هو:

أ. 28 — ب. 35 — ج. 14 — د. خلاف ذلك وهو:

112. المدى للبيانات التالية (18، 17، 16، 24، 4، 22) هو:

أ. 20 — ب. 8 — ج. 23 — د. خلاف ذلك وهو:

113. عند رسم منحنى تكراري فإننا نعين على المحور العمودي:

أ. حدود الفئات — ب. التكرار — ج. مراكز الفئات — د. خلاف ذلك وهو:

114. إذا كان الوسط الحسابي لـ10 قيم هو 20، فإن مجموع القيم هو:

أ. 20 — ب. 200 — ج. 100 — د. خلاف ذلك وهو:

115. مجموع انحرافات القيم عن وسطها الحسابي يساوي:

أ. 1 — ب. 0 — ج. رقم موجب — د. خلاف ذلك وهو:

116. إذا رسب 40 طالباً في مادة الإحصاء التي بها 500 طالب فإن نسبة الناجحين هي:

أ. 8% — ب. 4% — ج. 92% — د. خلاف ذلك وهو:

117. يمكن استخدام العينة العشوائية البسيطة إذا كانت مفردات المجتمع:

أ. متباينة — ب. متجانسة — ج. عددها صغير — د. خلاف ذلك وهو:

118. المقاييس المذكورة أدناه جميعها تعتبر ممثلة لمقاييس التشتت عدا:

أ. الانحراف المطلق — ب. التباين — ج. الربيع الثالث — د. معامل الاختلاف

119. إذا كان الحد الأدنى للفئة الأولى لتوزيع تكراري (10) وطول الفئة = 5 فإن مركز الفئة الخامسة =

أ. 30 — ب. 32 — ج. 33 — د. خلاف ذلك وهو:

120. إذا كان مركز الفئة في توزيع تكراري منتظم (19) وكان طولها (6) فإن مركز الفئة التي تسبقها:

أ. 13.   ب. 16   ج. 22   د. 25.

121. باستعمال المنحنى المتجمع الصاعد أو النازل أو كليهما معاً يمكن إيجاد ما يلي بيانياً:

أ. المتوسط   ب. الوسيط   ج. المنوال   د. جميع الإجابات خاطئة.

122. باستعمال المدرج التكراري يمكن إيجاد ما يلي بيانياً:

أ. المتوسط   ب. الوسيط   ج. المنوال   د. جميع الإجابات خاطئة.

123. في توزيع تكراري منتظم ذو فئات منفصلة إذا كان مركز الفئة الثالثة (17) وطول الفئة (5) فإن الحد الأدنى للفئة الثالثة:

أ. 12   ب. 22   ج. 15   د. جميع الإجابات خاطئة.

124. أكثر مقاييس النزعة المركزية استعمالاً هو:

أ. المتوسط   ب. الوسيط   ج. المنوال   د. الوسط الهندسي.

125. إذا كان المتوسط الحسابي لمجموعة من البيانات =60 ، والوسيط= 50، فإن المنوال=

أ. 30   ب. 40   ج. 50   د. جميع الإجابات خاطئة.

126. إحدى القيم الآتية يمكن أن تمثل معامل ارتباط عكسي بين المتغيرين (x,y):

أ. 0.8 -   ب. 1.7 -   ج. 0.3   د. 0.6 -

127. إذا كانت كل قيمة من قيم x تساوي 10 وكل قيمة من قيم y تساوي 20 وكان عدد قيم x ثمان قيم فإن معامل الارتباط بين المتغيرين =

أ. 1-   ب. 1   ج. صفر   د. جميع الإجابات خاطئة.

128. إذا كانت معادلة خط الانحدار لـ y على x هي y=0.8 x + 10 وكانت معادلة خط انحدار x على y هي x = 0.6 y + 8 فإن معامل الارتباط بين المتغيرين (x,y) =

أ. 0.5   ب. 0.7   ج. 0.7-   د. جميع الإجابات خاطئة.

129. إذا كان عدد المصابين الجدد بمرض الكوليرا في مدينة ما عام 2007م = 24 حالة ومعدل الإصابة بنفس المرض = 2.5 فإن عدد سكان المدينة في منتصف العام =

أ. 16800   ب. 19800   ج. 12600   د. خلاف ذلك وهو :

130. إذا كان نسبة الهلاك بمرض السرطان في مدينة ما = 9 ، وعدد حالات الوفاة = 7 فإن عدد حالات الإصابة =

أ. 125   ب. 1024   ج. 625   د. خلاف ذلك وهو :

131. إذا كان عدد الإصابات القديمة والحديثة بمرض السل في مدينة ما ، في لحظة معينة = a ، وإذا كان معامل الانتشار = 10 وعدد سكان المدينة في نفس اللحظة = 825 نسمة فإن قيمة a =

أ. 625   ب. 425   ج. 825   د. خلاف ذلك وهو :

132. بينت إحصائية عن حالات الإصابة بمرض الايدز في إحدى المدن أن عدد المصابين بذلك المرض وقت التعداد هو 5642 حالة وكان عدد سكان المدينة 315678 نسمة فإن معدل انتشار المرض:

أ. 16.5 لكل ألف   ب. 12.4 لكل ألف   ج. 22.6 لكل ألف   د. خلاف ذلك وهو :

قائمة المراجع

- أحمد عبدالسميع طيبة (2008). مبادئ الإحصاء. دار البداية ، عمان ، الأردن.
- أحمد الرفاعي غنيم، و نصر محمود صبري (2000م). تعلم بنفسك التحليل الإحصائي للبيانات باستخدام SPSS ، دار قباء للطباعة والنشر والتوزيع، القاهرة، مصر.
- حمزة محمد دودين (2010م) . التحليل الإحصائي المتقدم باستخدام SPSS ، دار المسيرة ، ط1 ، عمان ، الأردن .
- خالد قاسم سمور (2007م) . الإحصاء ، دار الفكر ، عمان ، الأردن .
- خليفة عبد السميع خليفة. الإحصاء التربوي، دار النهضة العربية، القاهرة، مصر.
- دلال القاضي، سهيلة عبد الله و محمود البياتي (2003م). الإحصاء للإداريين والاقتصاديين، دار الحامد للنشر والتوزيع، عمان، الأردن.
- دومينيك سالفاتور (1982م) . سلسلة ملخصات شوم نظريات ومسائل في الإحصاء والإقتصاد القياسي ، ترجمة سعدية حافظ منتصر، دار ماكجروهيل للنشر .
- زياد رمضان (1997م). مبادئ الإحصاء الوصفي والتطبيقي والحيوي. (ط4)، دار وائل للنشر، عمان، الأردن.
- شحادة مصطفى (1998م). مبادئ الإحصاء الوصفي والحيوي والتطبيقي وتطبيقات من البيئة الفلسطينية، دار الفاروق للثقافة والنشر، نابلس، فلسطين.
- صلاح الدين محمود علام (2003م). تحليل بيانات البحوث النفسية والتربوية والاجتماعية ، دار الفكر العربي ، القاهرة ، مصر .
- عبد الجبار توفيق (1983م) . التحليل الإحصائي في البحوث التربوية والنفسية والاجتماعية ، مؤسسة الكويت للتقدم العلمي.
- عبدالله فلاح المنيزل (ــ) . الإحصاء الاستدلالي وتطبيقاته في الحاسوب باستخدام الرزم الاحصائية (SPSS) ، دار وائل للنشر، عمان ، الأردن.
- عبدالله زيد الكيلاني و نضال كمال الشريفين (2007م) . مدخل إلى البحث في العلوم التربوية والإجتماعية ، دار المسيرة ، عمان ، الأردن .

- فايز جمعة صالح النجار ، نبيل جمعة النجار و ماجد راضي الزعبي (2008م) . أساليب البحث العلمي منظور تطبيقي دار الحامد للنشر والتوزيع ، عمان ، الأردن.
- فريد خليل الجاعوني، حبيب علي إسماعيل و عدنان محمود غانم (1998م) . مبادئ الإحصاء، ط1 ، مكتبة مراد، صنعاء، اليمن.
- محمد حسين محمد رشيد (2008م) . الإحصاء الوصفي والتطبيقي والحيوي ، دار صفا للنشر والتوزيع ، عمان ، الأردن .
- محمد عبد الكريم المنصوب (1998م) . مفاهيم أساسية في الإحصاء ، دار الخيرة، اليمن، صنعاء.
- محمد عبد المجيد المصري، الإحصاء التربوي لطلبة الدراسات العليا والباحثين في العلوم التربوية والنفسية (الجزء الأول)، مركز التربية للطباعة والنشر والخدمات الطلابية، صنعاء، اليمن.
- محمد عبد المجيد المصري، الإحصاء التربوي لطلبة الدراسات العليا والباحثين في العلوم التربوية والنفسية (الجزء الثاني)، مركز التربية للطباعة والنشر والخدمات الطلابية، صنعاء، اليمن.
- محمد عوض ومحمود ابو صالح (1983م) . مقدمة في الإحصاء ، مؤسسة دار جون وايلي وابنائه.
- محمد صبحي ابو صالح وآخرون (2000م) . مقدمة في الطرق الإحصائية ، ط1 ، دار اليازوري العلمية ، عمان ، الأردن .
- محمد صبحي ابو صالح ، عدنان محمد عوض (2010م) . مقدمة في الإحصاء مبادئ وتحليل باستخدام SpSS ، ط5 ، دار المسيرة ، عمان ، الأردن .
- محمد الصيرفي (2007م) . الحاسوب والإحصاء الإجتماعي ، دار الوفاء ، الإسكندرية ، مصر.
- محمد شامل بهاء الدين فهمي (2005م) . الإحصاء بلا معاناة ، المفاهيم مع التطبيقات باستخدام SPSS ، معهد الإدارة العامة بالرياض .
- محمد مفرح العسالي، عبد الحكيم العبيد، سعيد أحمد حسن (2008م). مبادئ الإحصاء، ط1، مركز الأمين، صنعاء، اليمن.
- محمد محمد المزاح (2008م) . مبادئ الإحصاء والإحتمالات للعلوم الإدارية والتطبيقية ، مركز التعليم المفتوح والتعلم عن بُعد ، جامعة العلوم والتكنولوجيا، اليمن .
- مصطفى خلف عبد الجواد (2009م) . الإحصاء الإجتماعي ، المبادئ والتطبيقات ، دار المسيرة ، عمان ، الأردن .

- منصور علي عطاء (1999م). مقدمة في الإحصاء و نظرية الاحتمالات، ط1، مركز الصلاحي لخدمات الطالب، صنعاء، اليمن.

- منصور محمد إسماعيل العريقي (2009م). طرق البحث للباحثين في العلوم الإدارية والتسويقية والمالية والمصرفية ، قسم إنتاج المقررات - عمادة التعليم المفتوح والتعلم عن بُعد ، جامعة العلوم والتكنولوجيا ، اليمن .

- نبيل جمعة صالح النجار (2010م) . الإحصاء في التربية والعلوم الاجتماعية مع تطبيقات برمجية SPSS ، دار الحامد ، عمان ، الأردن .

- ندى محمد الصوص (2007م) . مبادئ الإحصاء ، دار أجنادين ، الرياض ، السعودية .

- نورث جميل سليم ، شيلز ثامر نور الدين (2007م) . مبادئ الإحصاء لغير الإحصائيين ، الوراق للنشر والتوزيع ، عمان ، الأردن .

- الوديان، عادل مفلح، والعساوي، أحمد أحمد، والخريجي، عبدالله بن علي (2007). **مقدمة في الإحصاء والاحتمالات**، مكتبة الرُّشد، المملكة العربية السعودية، الرياض.

- وليد اسماعيل السيفو ، عيد أحمد أبو بكر ، غالب عوض الرفاعي (2010م) ، أساسيات الأساليب الإحصائية للأعمال وتطبيقاتها في العلوم المالية والإدارية والإقتصادية ، زمزم ناشرون وموزعون ، عمان، الأردن .

- Martin Bland ، ترجمة أحمد دشاش ديب ، شفيق ياسين ، وائل الإمام (2000). المدخل إلى الإحصاء الطبي ، المركز العربي للتعريب والترجمة والتأليف والنشر ، دمشق ، سوريا .

- James A. Twaite, Jane A. Monroe(1979). Introductory Sta s cs. London, England.

- Richard I.Levin , David S. Rubin (1994). Sta s cs For Management , sixth Edition, New Delhi.

- S.P.GUPTA(2006). Sta s cal Methods. Sultan chand & Sons, New Delhi.