الجمهورية الجزائرية الديمقراطية الشعبية

وزارة التعليم العالي والبحث العلمي

جامعة عمار ثليجي الأغواط

معهد علوم وتقنيات النشاطات البدنية والرياضية

مطبوعة بعنوان

الإحصاء الوصفي

المقياس: **الإحصاء الوصفي**

المستوى: **الثانية ليسانس**

القسم: **إدارة و تسيير المنشآت الرياضية**

التخصص: **كل التخصصات**

إسم ولقب المؤلف : **عادل عاشور**

السنة الجامعية : **2022/2021**

الجمهورية الجزائرية الديمقراطية الشعبية

وزارة التعليم العالي والبحث العلمي

جامعة عمار ثليجي الأغواط

معهد علوم وتقنيات النشاطات البدنية والرياضية

مطبوعة بعنوان

الإحصاء الوصفي

المقياس: **الإحصاء الوصفي**

المستوى: **الثانية ليسانس**

القسم: **إدارة و تسيير المنشآت الرياضية**

التخصص: **كل التخصصات**

إسم ولقب المؤلف : **عادل عاشور**

السنة الجامعية : **2022/2021**

تقديم

انطلاقا من التجربة المتواضعة و المقدرة بأكثر من عشر سنوات في تدريس مادة الإحصاء بمختلف تسمياته و أقسامه – وصفي، استدلالي ، تطبيقي و تقنيات المعالجة الإحصائية لطلبة علوم و تقنيات الأنشطة البدنية و الرياضية في أطوارها الثلاث، (ليسانس – ماستر – دكتوراه) بمختلف تخصصاتها ارتأيت في البداية أن أضع بين أيدي طلبة معهد STAPS بشكل خاص و طلبة كلية العلوم الإقتصادية و الإنسانية و الإجتماعية بشكل عام برنامج الإحصاء الوصفي المسطر من طرف الوزارة الوصية، بما يتوافق و المدة الزمنية المحددة للمقياس ، إذ تم إدراج المقياس، كمقياس سداسي للموسم الأول من السنة الثانية ليسانس، تتراوح مدته في أقصى حد بستة عشرة أسبوعا.

وروعي في تقديم هذه المواضيع – المحاضرات – الخلفية الرياضية لطلبة معهد علوم و تقنيات النشاطات البدنية و الرياضية و كذا العلوم الإنسانية و الاجتماعية، الذين في أغلبهم متحصلون على بكالوريا شعبة آداب هذا من جانب، ومن جانب آخر من جل التأسيس لعمليتي ترسيخ و توظيف أساليب الإحصاء الوصفي، و التي تعتبر الأساس لأي عملية بحث علمي في أي مجا، أما الرموز فقد تعمدنا إستخدام الرموز اللاتينية حتى يتسنى للطالب الاستفادة من المراجع باللغة العربية و الأجنبية على حد سواء. كما روعي في تقديم هذه المحاضرات تدعيمها بأمثلة تبسيطية للعلاقات الرياضية باستخدام مجموعة من التمارين التطبيقية والتي تنسجم مع طبيعة المادة المقدمة هدفها ترسيخ العلاقة الرياضية و تبسيط آليات التوظيف، لكن كل هذا يبقى مربوط بمدى مواظبة الطالب و حرصه على التعلم .

د.عادل عاشور

الفهرس

الفهرس

مقدمة

الفصل الأول : المفاهيم الأساسية لعلم الإحصاء

1-نبذة عن علم الإحصاء.....6
2-تعريف الإحصاء 9
3-وظائف علم الإحصاء.....10
4-أهمية الإحصاء 11
5-إستخدامات الإحصاء 12
6-بعض المصطلحات الإحصائية 12
7-المتغيرات 14
8-أنواع البيانات و طرق قياسها..... 15
9-مصادر جمع البيانات.....16
10- طرق جمع البيانات 17
11- أنواع العينات.....18
12- تحديد حجم العينة.....22

الفصل الثاني : عرض البيانات الإحصائية

1-أنواع الجداول التكرارية.....25
2-أنواع التكرارات.....26
3-عرض البيانات جدوليا.....27
1.3. عرض بيانات المتغير الوصفي في شكل جدول تكراري بسيط.....27
2.3. عرض بيانات المتغير الكمي في شكل جدول تكراري بسيط.....31

4- عرض البيانات بيانيا.....36

1.4. العرض البياني للبيانات الكيفية (الوصفية).....36

2.4. العرض البياني للبيانات الكمية.....38

الفصل الثالث : مقاييس النزعة المركزية

1- الوسط (المتوسط) الحسابي و خصائصه.....44

2-الوسط الهندسي و خصائصه.....51

3- الوسط التوافقي و خصائصه.....55

4-الوسط التربيعي وخصائصه.....57

5- الوسيط و خصائصه.....58

6- أشباه الوسيط 64

7-المنوال و خصائصه.....68

الفصل الرابع: مقاييس التشتت

1-المدى.....76

2- الإنحراف الربيعي.....78

3- الإنحراف المتوسط 82

4-التباين و الإنحراف المعياري وخصائصه.....85

5- مقاييس التشتت النسبي.....92

الفصل الخامس : مقاييس الشكل

1-الإلتواء.....95

2-التفرطح.....97

الفصل السادس : التوزيع الطبيعي

1- خصائص التوزيع الطبيعي.....98

2- استخدام خصائص المنحنى الطبيعي في تحليل البيانات.....104

الفصل السابع: الدلالة الإحصائية

1- الفروض وأنواعها.....110

2-مستويات الدلالة الإحصائية.....111

3-الخطأ من النوع الأول و النوع الثاني.....113

4-قوة الاختبار الإحصائي.....115

5-درجات الحرية.....115

مقدمة

مما لا شك فيه اليوم وفي عصر العولمة التي نعيشها وضخامة البيانات والمعلومات المتوفرة بين أيدينا حيث يتطلب منا العمل الجاد لاستغلال تلك المعلومات في عملية تنظيمية وتحليلية لها للوصول لقرارات أو التنبؤ على الأقل في وضع إستراتيجية للعمل خلال سنوات قادمة، فنحن في حاجة إلى مهارات خاصة للتعامل مع هذه المعلومات وهو ما يعرف بعلم الإحصاء فهو ذاك العلم الذي يبحث في كيفية جمع البيانات لظاهرة ما ومن ثم تصنيفها في جداول تمثل بيانياً ومن ثم استنتاج النتائج التي توصلنا لاتخاذ القرار المناسب، كما لا يغيب عن بالنا بأن علم الإحصاء هو مجموعة من الأساليب والعمليات الخاصة لمعالجة البيانات المتوفرة لدينا سواء الكمية منها أو الرقمية وللوصول لقيمة ذات دلالة للقيم موضوع البحث تتجانس معها القيم الأخرى وهذه القيمة ذات الدلالة تعرف بمقياس النزعة المركزية بالإضافة لمقاييس أخرى هي مقاييس التشتت .

الفصل الأول

المفاهيم الأساسية لعلم الإحصاء

إن مدلول كلمة إحصاء يراه البعض مجرد بيانات وأرقام بينما يراه آخرون طرق لجمع البيانات ووصفها وعرضها بصورة مبسطة في حين يراه البعض الآخر وسيلة لاتخاذ القرار حال عدم التأكد.

1/ نبذة عن علم الإحصاء:

اعتبر الإحصاء من الأمور القديمة المعروفة لدى المجتمعات، حيث يحرص القادة والزعماء والملوك على إحصاء عدد الجنود والأسلحة لخوض الحروب واستعراض القوة، كما تحرص الجماعات على إحصاء عدد أفرادها من أجل معرفة قوتها وكثرتها وقد أطلق على هذه الطرق علم الدولة أو علم الملوك ثم علم الإحصاء[1] . وقد وردت كلمة الإحصاء ومشتقاتها في القرآن الكريم إحدى عشرة مرة، كما وردت في السنة النبوية في مواضع متعددة، واستخدم الاحصاء في عصره الاول في جمع البيانات عن السكان و حصرهم لأهداف معينة تتمثل في استخدامهم في الجيوش أو لغرض فرض الضرائب أو لتوزيع الأراضي الزراعية على السكان بطريقة عادلة، و في القرن السابع عشر و الذي يمكن اعتباره العصر الإحصائي الثاني تم استخدام الطريقة الرقمية للدلالة على الظواهر موضوع البحث على اعتبار أن هذه الطريقة أدق و أقوى في التعبير عن هذه الظواهر و تركز الهدف من هده الطريقة في معرفة عدد السكان و المواليد و الوفيات ومقدار الثروة والدخل ومقدار الضرائب المحصلة و كمية الإنتاج و الاستهلاك، وباختصار نجد أن مجال الإحصاء قبل القرن الثامن عشر كان مرتبطا في الغالب بالمجالات الاقتصادية و الاجتماعية ، و كانت الأساليب الإحصائية المستخدمة تمتاز بالبساطة بحيث لم توفر للإحصاء الأسس و المقومات الكافية لأن يصبح علما، و يمكن تحديد بداية العصر الإحصائي الثالث مع تطور علوم الرياضيات في القرن الثامن عشر و ظهور بعض النظريات العلمية الهامة مثل نظرية الاحتمالات التي كان لها الدور الكبير في تطور الإحصاء على يد مجموعة من العلماء منهم فرانسيس يزدرو أيدجورث، وفرانسيس

[1] كلمة Statistics وهي كلمة لاتينية مشتقة من كلمة state و تعني الدولة .

غالتون، وكارل بيرسون، وجورج أودني بول، وآخرون، واكتسابه أهمية كبرى بحيث أصبح علما مستقلا وانتشر استخدامه وبدأ الاهتمام من قبل العلماء في تطبيق النظريات و الطرق و الأساليب الإحصائية في الكثير من فروع العلم الحديث كالهندسة والطب و الصيدلة و الزراعة و الصناعة و الجغرافيا و الفلك و علم النفس باعتباره الطريقة الصحيحة و الأسلوب الأمثل اتباعه في البحث العلمي و في القرن العشرين تطور علم الإحصاء وعزز من ذلك حاجة صناع القرار والقادة العسكريون في الحرب العالمية الثانية للخطط الإحصائية والمزيد من الأفكار الإحصائية، و أخيرا فقد أدى ظهور الحاسبات الآلية و تطورها في وقتنا الحالي بأنواعها المختلفة و قدرتها الفائقة و دقتها المتناهية إلى تمهيد الطريق لاستخدام و تطبيق الأساليب الإحصائية المختلفة في شتى المجالات و الميادين.

1.1. معنى كلمة إحصاء :

في القرآن الكريم:

ثُمَّ بَعَثْنَاهُمْ لِنَعْلَمَ أَيُّ الْحِزْبَيْنِ أَحْصَى لِمَا لَبِثُوا أَمَدًا (12) سورة الكهف .

إِنَّا نَحْنُ نُحْيِي الْمَوْتَى وَنَكْتُبُ مَا قَدَّمُوا وَآثَارَهُمْ وَكُلَّ شَيْءٍ أَحْصَيْنَاهُ فِي إِمَامٍ مُبِينٍ (12) سورة يس

لِيَعْلَمَ أَن قَدْ أَبْلَغُوا رِسَالَاتِ رَبِّهِمْ وَأَحَاطَ بِمَا لَدَيْهِمْ وَأَحْصَى كُلَّ شَيْءٍ عَدَدًا (28) سورة الجن وَكُلَّ شَيْءٍ أَحْصَيْنَاهُ كِتَابًا (29) سورة النبأ .

وَوُضِعَ الْكِتَابُ فَتَرَى الْمُجْرِمِينَ مُشْفِقِينَ مِمَّا فِيهِ وَيَقُولُونَ يَا وَيْلَتَنَا مَالِ هَذَا الْكِتَابِ لا يُغَادِرُ صَغِيرَةً وَلا كَبِيرَةً إِلا أَحْصَاهَا وَوَجَدُوا مَا عَمِلُوا حَاضِرًا وَلا يَظْلِمُ رَبُّكَ أَحَدًا (49) سورة الكهف

يَوْمَ يَبْعَثُهُمُ اللَّهُ جَمِيعًا فَيُنَبِّئُهُم بِمَا عَمِلُوا أَحْصَاهُ اللَّهُ وَنَسُوهُ وَاللَّهُ عَلَى كُلِّ شَيْءٍ شَهِيدٌ (6) سورة المجادلة .

لَقَدْ أَحْصَاهُمْ وَعَدَّهُمْ عَدًّا (94) سورة مريم .

وَآتَاكُم مِّن كُلِّ مَا سَأَلْتُمُوهُ وَإِن تَعُدُّواْ نِعْمَتَ اللّهِ لاَ تُحْصُوهَا إِنَّ الإِنسَانَ لَظَلُومٌ كَفَّارٌ (34) سورة إبراهيم .

يستفاد من مفهوم الكلمة إحصاء هنا بأنها بيانات

في اللغة:

عبر العرب عن كثرة الشيء وحجمه بالحصى، ويقال حصيت أي عددت وأحصيته أي ميزته بعضه عن بعض، والحصاة بمعنى العقل، واختار الإمام ابن القيم في البدائع (164/1) أن الإحصاء على ثلاثة مراتب هي:

1- إحصاء ألفاظها وعددها

2- فهم معانيها ومدلولها .

3- دعاؤه بها.

والإحصاء في الكلام : على ثلاث مراتب :

1- العدد: ومنه قوله تعالى {وأحصى كل شيء عدداً} [الجن :28] .

2- المفهم: ومنه يقال: رجل ذو حصاة أي: ذو لب وفهم، ومنه سمي العقل .

3- الإطاقة على العمل والقوة: ومنه قوله تعالى: { علم ألن تحصوه } [المزمل : 20] أي: لن تطيقوا العمل بذلك .

2.1. استخدام كلمة إحصاء:

تستخدم بصيغة الجمع أي إحصائيات (Statistics) ومرادفتها بيانات وتستخدم كمفرد إحصاء أو إحصاءه بمعنى النتيجة التي توصلنا إليها من إجراء الرياضيات على البيانات فمتوسط العمر لمجموعة من الناس يطلق عليه إحصاء أو إحصاءه.

علماً : فرع من فروع العلم، فالإحصاء مجرد بيانات تعتبر المادة الخام لهذا العلم والبيانات تجمع من مصادرها المطلوبة حسب المطلوب فتبوب وتلخص وتقييم ومن ثم يستنتج المراد منها.

2/ تعريف الإحصاء :

يحتل علم الإحصاء أهمية خاصة بين العلوم المختلفة سواء كانت طبيعية او إنسانية او اجتماعية وذلك لأنه يتعامل مع الحاجات الإنسانية والعلمية في مختلف مستوياتها وبأساليب متنوعة بدءا من البساطة التي تصف معلومات أولية (بيانات إحصائية)صعودا الى أساليب أكثر تعقيدا وتطورا لتحليل البيانات ذات الصلة بوصف وتفسير الظواهر والأحداث وعلاقتها مع بعضها والتنبؤ بها والسيطرة عليها في ضوء القوانين التي تحكمها ، لهذا اخذ ينظر للإحصاء بأنه ذلك العلم الذي يهتم بدراسة الظواهر الحياتية المختلفة باستخدام مجموعة من الوسائل الإحصائية لغرض وصفها وتفسيرها وإرساء التوقعات عن سلوك الظواهر مستقبلا واتجاهات تطورها وبالتالي يصبح هذا العلم أداة أساسية للوصول الى قرارات صائبة في وصف ودراسة الظواهر.

و يعرف علم الإحصاء بأنه:

- مجموعة من الطرق العلمية القياسية تهدف لجمع البيانات عن ظاهرة ومن ثم جدولتها وعرضها بصورة مختصرة وتقيمها للوصول لنتائج من خلال عينة صغيرة من مجتمع للتعرف على خصائص المجتمع بصورة تقريبية لخصائص العينة.
- هو العلم الذى يبحث فى الطرق والأساليب المناسبة لجمع وعرض وتبويب وتحليل البيانات حتى يمكن فهمها ، والعمل على الوصول إلى نتـائج وقرارات سليمة على ضوئها ، ثم تعميم النتـائج
- العلم الذي يهتم بطرق جمع البيانات وتبويبها و تلخيصها وتنظيمها وعرضها بشكل يمكن الإستفادة منها في وصف البيانات و تحليلها للوصول إلى قرارات سليمة في ظل ظروف عدم التأكد.
- العلم الذي يدرس مختلف طرق ووسائل جمع البيانات الكمية عن مختلف الظواهر الاقتصادية والاجتماعية وغيرها، وترتيب هذه البيانات وتبويبها وتحليلها وتفسيرها وتقديمها بأشكال و صور ملائمة بهدف تسهيل اتخاذ القرار على أساس سليم

- العلم الذي يبحث في الأساليب والطرق العلمية المناسبة لجمع البيانات، وتبويبها وتنظيمها وتحليلها وتفسيرها بهدف الوصول إلى النتائج اللازمة لزيادة المعرفة أو اتخاذ القرارات المناسبة وتعميمها.

و بناءا على التعاريف السابقة يستنتج أن الإحصاء هو طريقة تتضمن أربع خطوات هي:

جمع البيانات : و تشمل الحصول على القياسات أو القيم للمشاهدات و التجارب التي يجريها الباحث

تنظيم وعرض البيانات : تعني و ضع البيانات التي تم الحصول عليها في الخطوة الأولى في جداول معينة يتم تصميمها و عرضها بطرق مناسبة مثل الرسوم البيانية أو التوزيعات التكرارية.

تحليل البيانات: يعني استخدام الأساليب الإحصائية المختلفة في تحليل البيانات التي تم جمعها و عرضها .

استقراء النتائج و اتخاذ القرارات : و يقصد بها الاستنتاجات التي يتم التوصل إليها على شكل تقديرات أو تنبؤات . وإن كانت هذه الخطوة تتدرج ضمن الإحصاء الاستدلالي

3/ وظائف علم الإحصاء :

من التعريف السابق يمكن تحديد أهم وظائف علم الإحصاء في الآتي:

1.3. وصف البيانات

تعتبر طريقة جمع البيانات وتبويبها وتلخيصها من أهم وظائف علم الإحصاء، إذ لا يمكن الاستفادة من البيانات الخام، ووصف الظواهر المختلفة محل الاهتمام، إلا إذا تم جمع البيانات وعرضها في شكل جدولي، أو بياني من ناحية، وحساب بعض المؤشرات الإحصائية البسيطة التي تدلنا على طبيعة البيانات من ناحية أخرى.

2.3 الاستدلال الإحصائي

وهو أيضا من أهم الوظائف المستخدمة في مجال البحث العلمي، ويستند الاستدلال الإحصائي على فكرة اختيار جزء من المجتمع يسمى عينة بطريقة علمية مناسبة، بغرض

استخدام بيانات هذه العينة في التوصل إلى نتائج، يمكن تعميمها على مجتمع الدراسة، ومن ثم يهتم الاستدلال الإحصائي بموضوعين هما:

التقدير: وفيه يتم حساب مؤشرات من بيانات العينة تسمى إحصاء ، تستخدم كتقدير لمؤشرات المجتمع وتسمى معالم، ويطلق على المقاييس الإحصائية المحسوبة من بيانات العينة في هذه الحالة بالتقدير بنقطة، كما يمكن أيضا ، استخدام المقاييس الإحصائية المحسوبة من بيانات العينة في تقدير المدى الذي يمكن أن يقع داخله معلمة المجتمع باحتمال معين، ويسمى ذلك التقدير بفترة

اختبارات الفروض : وفيه يتم استخدام بيانات العينة للوصول إلى قرار علمي سليم بخصوص الفروض المحددة حول معالم المجتمع.

3.3. التنبؤ:

وفيه يتم استخدام نتائج الاستدلال الإحصائي، والتي تدلنا على سلوك الظاهرة في الماضي في معرفة ما يمكن أن يحدث لها في الحاضر والمستقبل .وهناك العديد من الأساليب الإحصائية المعروفة التي تستخدم في التنبؤ، ومن أبسطها أسلوب الاتجاه العام، وهي معادلة رياضية يتم تقدير معاملاتها باستخدام بيانات العينة، ثم بعد ذلك استخدام المعادلة المقدرة في التنبؤ بما يمكن أن يحدث للظاهرة في المستقبل.

4/ أهمية الإحصاء :

لقد احتل الإحصاء كعلم مركزاً مرموقاً بين العلوم الأخرى خلال القرنين الماضيين وقد برز الإحصاء في وقتنا الحاضر كأحد أهم العلوم التطبيقية التي لا يمكن الاستغناء عنها في أي فرع من فروع الحياة وبرز ذلك جلياً بانتشار الشعب الإحصائية ، والتي تسمى في المؤسسات الإنتاجية " مراقبة الجودة، أو التحكم النوعي، في كل المصانع والمؤسسات الإنتاجية و الخدمية للدول المتقدمة . والهدف من إنشاء هذه الشعب الإحصائية هو مراقبة الإنتاج وحساب معدلاته ومراقبة تطوراته والنسب المئوية، وكذلك ضمن اهتمام هذه الشعب إحصائيات عن العمال والموظفين والزبائن والمستهلكين بالإضافة لمراقبة

الإيرادات ورؤوس الأموال ...إلخ . فللإحصاء دور بارز في وضع الخطط المستقبلية عن طريق التنبؤ بالنتائج لكافة القطاعات حيث يمكن تطبيق علم الإحصاء في كل المجالات كالعلوم الإنسانية و الاجتماعية أو التجريبية (مثل الزراعة الصناعة ،الطب، الهندسة، الإدارة والاقتصاد...الخ)

5/ استخدامات الإحصاء :للإحصاء استخدامات عدة أهمها:

– التنبؤ :هو توقع حدوث شيء في المستقبل بناء على معطيات موجودة في الحاضر هذا بشكل عام، أما في الإحصاء فهو استخدام النتائج في تقدير رقمي – كمي – لبيان أشياء في المستقبل.

– اتخاذ القرار : هي عملية انتقالية يتم من خلالها اختيار أحد البدائل المناسبة من بين عدة بدائل، و ذلك بناء على معلومات متوفرة.

– التحقق: هي تلك العملية التي يتبعها الباحث من أجل التأكد من صحة ما يبحث عنه

– الرقابة :نعني بها آلية التأكد من جودة المنتوج أو الخدمة و تعتمد على الإحصاء كوسيلة لبلوغ الهدف.

6/ بعض المصطلحات الإحصائية :

1.6.الإحصاء الوصفي

يختص بالطرق المختلفة لجمع البيانات وتحليها ووصفها حتى تقدم بصورة بسيطة سهلة الفهم وذات دلالة دون أي تعميم. وسائل هذا النوع من الإحصاء العرض البياني والحساب فالعرض البياني يعني جدولة البيانات وعرضها بيانياً والحساب المقاييس الإحصائية لتلك البيانات كالنزعة المركزية من أمثلته الوسط الحسابي أو مقاييس التشتت أو عدم التجانس من أمثلته المدى ومقاييس أخرى.

2.6. الإحصاء الاستدلالي :

هو مجموعة الطرق للتعرف على خصائص المجتمع من خلال عينة من هذا المجتمع معتمدة طرق إحصائية محددة. و يشكل الإحصاء الاستدلالي مع الإحصاء

الوصفي فرعا علم الإحصاء الحديث وهما ضروريان لاتخاذ القرار.

الإحصاء الاستدلالي يتعامل مع التعميم والتقدير و التنبؤ إلا أنه يتسم في بعض الحالات بعدم التأكد مما يدعونا لمعالجة القياس في هذه الأحوال تحت باب علم الاحتمالات مما يعطي فكرة عن الخطأ المحتمل وقوعه من الباحث حال التعميم على المجتمع المسحوب منه العينة العشوائية محل الدراسة.

أما وسيلتي الإحصاء الاستدلالي فهما التقدير و اختبار الفروض .

3.6.المجتمع الاحصائي:

هو مجموعة من الأفراد أو ما شابه تعرف بمفردات المجتمع وهي محددة ومعروفة تماماً. مثلاً دراسة أعمار مدرسة ما فالمجتمع الإحصائي هنا هو طلاب المدرسة وقت الدراسة. فالمجتمع يمثل جميع المفردات الممكنة للظاهرة محل الدراسة، فالمجتمع قد يكون محدوداً (يمكن عد مفرداته ولو من الناحية النظرية) أو يكون مجتمع غير محدود (لا يمكن عد مفرداته)، وتعرف خصائص المجتمع التي يمكن قياسها كمياً بمعالم المجتمع وللمجتمع الغير محدود يستحيل الوصول للقيمة الحقيقية عند دراسة ظاهرة ما فلذا نستدل عليها بأخذ مفردات

قليلة العدد من المجتمع (العينة) للاستدلال على معالم المجتمع.

4.6. العينة :

هي مجموعة جزئية من المجتمع مكونة من عدة أفراد والدراسة عليها بدلاً من الدراسة على المجتمع حيث نوفر الوقت والجهد والمال وقد يستلزم الأمر ضرورة أخذ العينة من المجتمع فالعينة هي جزء من المجتمع ودراستها وما ينجم عنها من خصائص يمكن بواسطتها الاستدلال على خواص المجتمع ككل، وإحصائية العينة أي أحد خصائصها كالمتوسط الحسابي لأفرادها وهو قياس كمي أو جزء من أفرادها للكل كنسبة الناجحين بين أفراد العينة ويعتبر الوسط الحسابي أو النسبة (الاحتمال) من خصائص العينة ومنها يستدل على معالم المجتمع المسحوب منه العينة فرسوب ثلاثة طلاب من بين 20 طالباً

(العينة) أي 3 ÷ 50 = 0.06 هي من خواص العينة ويمكن الاستدلال منها حال عدد مفردات المجتمع (1000 مثلاً) من الطلاب المتقدمون للامتحان السابق في العينة يستدل برسوب 1000 × 0.06 = 60 طالب.

5.6.المعلمة و الإحصاءة:

المعلمة (Parameter) وهي القياس المستنتج من بيانات المجتمع مثل متوسط دخل الفرد في بلد ما يعتبر معلمة لأنه يعكس المستوى المعيشي لأفراد ذك البلد

والإحصاءة (Statistic) وهي المقياس المستخرج من بيانات العينة وهي محل التعامل مع بياناتها في معظم الأحوال.

وعليه يكون للعينة أهمية قصوى في معرفة خصائص مجتمع أو أكثر من خلال خصائص العينة محل الدراسة وباعتبار معلمة المجتمع مجهولة فتقدر بالمناسب من خلال حسابه من بيانات العينة.

6.6. الوصف الإحصائي : هو تلخيص لمجموعة القراءات محل الدراسة بطريقة مناسبة.

7.6.الاستنباط الإحصائي : وهو أن يقوم الباحث باستخراج التصميمات والاستنتاجات المعتمدة على نظرية الاحتمال للحصول على مقدار الخطأ المحتمل وقوعه من قبل الباحث عند التعميم على المجتمع.

8.6. البيانات : هى مجموعة من المشاهدات أو الملاحظات التي تؤخذ أثناء دراسة معينة، وقد تكون بيانات كمية(رقمية) مثل أطوال وأوزان مجموعة من الطلاب أو علاماتهم أو بيانات وصفية (غير رقمية) مثل اللون والجنس ... إلخ.

7/ المتغيرات :

عبارة عن صفات و تعتبر الجزء الأساسي الذي يتعامل معه الإحصائي، فإذا اختلفت خاصية ما في مفردات مجموعة معينة كما أو كيفا، تكون هذه الخاصية هي المتغير، فالمتغير إذا له خصائص رقمية (كمية) وغير رقمية (وصفية) تتغير قيمته من عنصر

إلى آخر من عناصر المجتمع أو العينة .و بالتالي نستطيع أن نقول أن هناك نوعين من المتغيرات هما :

1.7. المتغيرات الكمية : (الرقمية) و تخص الجوانب المادية القابلة للقياس العددي أي يمكن التعبير عن وحداتها بأعداد او أرقام و هي نوعان هما :

المتغيرات الكمية المستمرة أو المتصلة: وهي التي تأخذ أي قيمة صحيحة أو كسرية كالطول والوزن والدخل و السرعة و غير ذلك.

المتغيرات المتقطعة أو المنفصلة: وهي التي تأخذ قيما صحيحة فقط كعدد أفراد الأسرة أو عدد طلاب مدرسة وغير ذلك

2.7. المتغيرات الكيفية : (الوصفية) تخص كل ما هو غير قابل للقياس العددي مثلا الحالة المدنية، المهنة، رأي،...إلخ

ملاحظة:

الصفة الكمية : هي جواب لسؤال كم ؟

الصفة الكيفية هي جواب لسؤال ماهو، ما هي ، ما؟

8/ أنواع البيانات و طرق قياسها :

يمكن تقسيم البيانات إلى مجموعتين هما:

1.8. البيانات الوصفية:

هي بيانات غير رقمية، أو بيانات رقمية مرتبة في شكل مستويات أو في شكل فئات رقمية، ومن ثم تقاس البيانات الوصفية بمعيارين هما:

أ – بيانات وصفية مقاسة بمعيار اسمي وهي بيانات غير رقمية تتكون من مجموعات متنافية، كل مجموعة لها خصائص تميزها عن المجموعة الأخرى، كما أن هذه المجموعات لا يمكن المفاضلة بينها، ومن الأمثلة على ذلك

النوع :متغير وصفي تقاس بياناته بمعيار اسمي (ذكر، انثى.)

الحالة الاجتماعية :متغير وصفي تقاس بياناته بمعيار اسمي (متزوج، أعزب، أرمل، مطلق.)

الجنسية :متغير وصفي تقاس بياناته بمعيار اسمي (جزائري، غير جزائري.)

ب – بيانات وصفية مقاسة بمعيار ترتيبي : وتتكون من مستويات، أو فئات يمكن ترتيبها تصاعديا أو تنازليا، ومن الأمثلة على ذلك:

تقديرات الطلاب : متغير وصفي تقاس بياناته بمعيار ترتيبي :ممتاز، جيد، مقبول ، متوسط، ضعيف.

المستول التعليمي :متغير وصفي تقاس بياناته بمعيار ترتيبي : أمي، يقرأ و يكتب، ابتدائي، متوسط، ثانوي، جامعي.

وهذا النوع من البيانات يمكن أن تعطى له أرقاما، ولكن هذه الأرقام لا يمكن أن تخضع للعمليات الحسابية المعروفة كالجمع، و الضرب ،...الخ.

2.8. البيانات الكمية

هي بيانات يعبر عنها بأرقام عددية تمثل القيمة الفعلية للظاهرة، وتنقسم إلى قسمين هما: بيانات فترة : وهي بيانات رقمية،تقاس بمقدار بعدها عن الصفر، أي أن للصفر دلالة على وجود الظاهرة، ومن أمثلة ذلك:

درجة الحرارة :متغير كمي تقاس بياناته بمعيار عددي، حيث أن درجة الحرارة 0 ليس معناه انعدام الظاهرة، ولكنه يدل على وجود الظاهرة.

بيانات نسبية : هي متغيرات كمية، تدل القيمة 0 على عدم وجود الظاهرة ومن الأمثلة على ذلك : المساحة ، الوزن ، الطول...إلخ

9/ مصادر جمع البيانات :

يحتاج كل باحث يريد تطبيق الطريقة الإحصائية المناسبة إلى جمع بيانات حول موضوع بحثه لغرض التحليل الإحصائي، كما أن هناك مصدرين للحصول منها على البيانات هما:

1.9. المصادر الأولية:

وهي المصادر التي نحصل منها على البيانات بشكل مباشر، حيث يقوم الباحث نفسه بجمع البيانات من المفردة محل البحث مباشرة، فعندما يهتم الباحث بجمع بيانات عن رياضة ما، يقوم بإجراء مقابلة مع النوادي و الرياضيين ، ويتم الحصول منهم مباشرة على بيانات خاصة بهذه الرياضة، مثل عدد الممارسين، و حالتهم الاجتماعية ، و خبرتهم، والحنسية، والمهنة، والدخل الشهري، وعدد أفراد الأسرة، والمستوى التعليمي، ... وهكذا. ويتميز هذا النوع من المصادر بالدقة والثقة في البيانات، لأن الباحث هو الذي يقوم بنفسه بجمع البيانات من المفردة محل البحث مباشرة، ولكن أهم ما يعاب عليها أنها تحتاج إلى وقت ومجهود كبير، ومن ناحية أخرى انها مكلفة من الناحية المادية.

2.9. المصادر الثانوية :

وهي المصادر التي نحصل منها على البيانات بشكل غير مباشر، بمعنى آخر يتم الحصول عليها بواسطة أشخاص آخرين، أو أجهزة، وهيئات رسمية متخصصة، مثل نشرات وزارة المالية، ونشرات مصلحة الإحصاء، ونشرات وزارة التعليم العالي وهكذا.

ومن مزايا هذا النوع من المصادر، توفير الوقت والجهد والمال، إلا أن درجة ثقة الباحث فيها ليست بنفس الدرجة في حالة المصادر الأولية.

10/ طرق جمع البيانات:

تعتبر طريقة جمع البيانات من أهم المراحل التي يعتمد عليها البحث الإحصائي، كما أن جمع البيانات بأسلوب علمي صحيح، يترتب عليه الوصول إلى نتائج دقيقة في التحليل، و يتحدد الأسلوب المستخدم في جمع البيانات، حسب الهدف من البحث، وحجم المجتمع محل البحث، وهناك أسلوبين لجمع البيانات هما:

1.10. أسلوب الحصر الشامل :

يستخدم هذا الأسلوب إذا كان الغرض من البحث هو حصر جميع مفردات المجتمع، وفي هذه الحالة يتم جمع بيانات عن كل مفردة من مفردات المجتمع بلا استثناء، كحصر ذوي

الاحتياجات الخاصة الممارسين لرياضة ما ، ويتميز أسلوب الحصر الشامل بالشمول وعدم التحيز، ودقة النتائج، ولكن يعاب عليه أنه يحتاج إلى الوقت والجهد والتكلفة العالية.

2.10. أسلوب المعاينة :

يهتم هذا الأسلوب على معاينة جزء من المجتمع محل الدراسة، يتم اختياره بطريقة علمية سليمة، ودراسته ثم تعميم نتائج العينة على المجتمع ، ومن ثم يتميز هذا الأسلوب بالآتي:

1-تقليل الوقت والجهد.

2-تقليل التكلفة.

3-الحصول على بيانات أكثر تفصيلا، وخاصة إذا جمعت البيانات من خلال استمارة استبيان.

4-كما أن أسلوب المعاينة يفضل في بعض الحالات التي يصعب فيها إجراء حصر شامل، مثل معاينة التلاميذ المتمدرسين ولكن يعاب على أسلوب المعاينة:أن النتائج التي تعتمد على هذا الأسلوب أقل دقة من نتائج أسلوب الحصر الشامل، وخاصة إذا كانت العينة المختارة لا تمثل المجتمع تمثيلا جيدا.

11/ أنواع العينات

تتعدد أنواع العينات، وتتوزع إلى أسلوبين، الأول، وهو أسلوب العينة العشوائية، والثاني، وهو أسلوب العينة غير العشوائية. ويتوقف اختيار أسلوب العينة المناسب على عنوان البحث، وأهدافه، ومنهجه المستخدم. وفيما يلي عرض مفصل عن أنواع العينات:

1.11. العينة العشوائية أو العينة الاحتمالية :

ويستخدمه الباحث إذا كان أفراد المجتمع الأصلي للدراسة معروفين، وفي هذه الحالة يتم الاختيار العشوائي على أساس تكافئ فرص الاختيار أمام جميع أفراد المجتمع دون تدخل من طرف الباحث. فمثلاً إذا كان مجتمع الدراسة، هو طلاب معهد التربية البدينة، ففي هذه الحالة الطلاب معروفين لأنهم مسجلين في المعهد، وبمقدور الباحث الحصول على

قوائم رسمية وحديثة بأعدادهم وبيانات أخرى عنهم، وبالتالي فرصة الاختيار العشوائي من هؤلاء تكون متاحة أمامهم دون تمييز أو تحيز من قبل الباحث، ومن أنواع الأسلوب العشوائي أو الاحتمالي:

أولا : العينة البسيطة:

يختار الباحث هذا النوع من العينات العشوائية إذا كان مجتمع الدراسة متجانساً ولهذا النوع خطوات، هي:

أ ـ إما استخدام القرعة، بحيث يتم تحديد أرقام لجميع أفراد المجتمع الأصلي للدراسة، ثم وضع هذه الأرقام في صندوق خاص وتحرك بعضها مع بعض، وبالتالي يتم سحب أرقام من الصندوق حتى يستوفي الباحث العدد المطلوب للعينة.

ب ـ وإما باستخدام جدول الأرقام العشوائية، بحيث يحدد الباحث أرقاماً من جدول الأرقام العشوائية بصورة طولية أم عرضية، وإذا استوفي العدد المحدد للعينة قام باختيار الأفراد الذين لهم الأرقام ذاتها في المجتمع الأصلي للدراسة، وبعدما ينتهي الباحث يكون هؤلاء الأفراد هم العينة المختارة.

ثانيا : العينة المنتظمة:

يختار الباحث هذا النوع من العينات إذا كان مجتمع الدراسة متجانساً، على غرار العينة البسيطة، لكن تختلف العينة المنتظمة عن العينة البسيطة في خطوات تكوينها. حيث تكون المسافة بين أرقام أفراد العينة متساوية، فمثلاً إذا كان مجتمع الدراسة يتألف من 200 فرداً، والعدد المطلوب للعينة، هو 20 فرداً، فالمسافة بين الرقم الأول للفرد والذي يليه هي 10، وهي عبارة عن حاصل القسمة: 200 ÷ 10.

إذ يبدأ الباحث باختيار الرقم الأول عشوائياً، وليكن مثلاً 4 وبالتالي تكون العينة المنتظمة مؤلفة من الأفراد الذين يحملون الأرقام التالية 4 ، 14 ، 24 ، 34 ، 44 ، 54 ، 64 ...،

ثالثا: العينة الطبقية:

يختار الباحث هذا النوع من العينات إذا كان مجتمع الدراسة غير متجانس؛ نظراً لأنه يتألف من فئات أو طبقات مختلفة بعضها عن بعض. ويتطلب هذا النوع مراعاة الخطوات التالية:

تحديد الفئات المتوافرة في مجتمع الدراسة.

تحديد أفراد كل فئة على حدة.

اختيار من كل فئة عينة عشوائية بسيطة تمثلها بحيث يتناسب عدد كل فئة في العينة مع عددها في المجتمع الأصلي للدراسة.

فمثلاً إذا كان عنوان الدراسة عن مشكلات طلاب معهد التربية البدينة، فإن الباحث أمام مجتمع مختلف في مشكلات الطلاب تبعاً لاختلافهم في الأعمار، والتخصصات الدراسية، والناحية الاجتماعية، والناحية الاقتصادية.

رابعا: العينة العنقودية:

يختار الباحث هذا النوع من العينات إذا كان مجتمع الدراسة على مستوى كبير، حيث يصعب عليه استخدام العينة البسيطة أو العينة المنتظمة أو العينة الطبقية، ويتبع الباحث في هذه الحالة تقسيم المجتمع إلى مناطق ثم إلى محافظات ثم إلى أجزاء صغيرة. حتى يصل إلى الأفراد المطلوبين للعينة، والصالحين لتمثيل مجتمع الدراسة، فمثلاً إذا أراد الباحث أن يتعرف على مدى استخدام أعضاء هيئة التدريس بالجامعات الجزائرية للتقنيات الحديثة في التدريس فإنه لا يلزم الباحث القيام بزيارة كل جامعة أو كلية أو معهد على حدة، بل يكتفي بعدد ممثل من هذه الجامعات.

2.11. العينة غير العشوائية: أو العينة غير الاحتمالية:

ويستخدمه الباحث إذا كان أفراد المجتمع الأصلي للدراسة غير معروفين، وفي هذه الحالة يتم الاختيار غير العشوائي، وذلك بتدخل من الباحث، بحيث يختار أفراداً ويترك أفراداً من مجتمع الدراسة على ضوء شروط حددها الباحث، فمثلاً إذا كان مجتمع

الدراسة، هو المدخنين مثلا، فأفراد المجتمع هنا لا يمثلون جميع المتعاطين لهذه السموم في المجتمع، بل هناك أفراد غير معروفين لدى الباحث وفي هذه الحالة يعمد الباحث إلى الأخذ بالأسلوب غير العشوائي، ومن أنواع هذا الأسلوب:

أولا: العينة الصدفية:

يختار الباحث أفراد هذه العينة بالصدفة، أي دون ترتيب سابق معهم. كأن يختار الباحث عدداً من المصلين عند خروجهم من المساجد، أو عدداً من الطلاب عند خروجهم من مدارسهم ويسألهم عن موقفهم حيال تأثير الفضائيات على التحصيل الدراسي للطلاب، ويعاب على هذا النوع من العينات أن أفرادها لا يمثلون مجتمع الدراسة بصورة دقيقة، وبالتالي فإنه من الصعب تعميم نتائج الدراسة على كل المجتمع الأصلي.

ثانيا: العينة الحصصية:

يقوم الباحث إذا أراد الأخذ بالعينة الحصصية بتقسيم مجتمع الدراسة إلى فئات، ثم يختار عدداً من الأفراد من كل فئة بما يتناسب وحجم الفئة في مجتمع الدراسة، وتشبه العينة الحصصية العينة الطبقية في هذا المعنى، لكن تختلف عنها في أن العينة الحصصية يتدخل الباحث في اختيار أفراد العينة، بينما في العينة الطبقية لا يتدخل مطلقاً في اختيار أفراد العينة، ويعاب على هذا النوع من العينات، هو أنه لا يمثل مجتمع الدراسة بصورة دقيقة .

ثالثا: العينة الغرضية:

يختار الباحث أفراد هذه العينة إذا أدرك أنهم يحققون أغراض دراسته، فمثلاً إذا كان الباحث يريد دراسة عن رياضيي النخبة في الجزائر، فإنه يختار الرياضيين الذين يعتقد أنهم يفيدونه في تحقيق أغراض بحثه.

12/ تحديد حجم العينة:

أورد الإحصائيون القواعد التالية التي يمكن الاسترشاد بها لتحديد حجم العينة المطلوب على النحو التالي:

من ثلاثين إلى خمسمائة مفردة يعتبر ملائمًا لمعظم الأبحاث والدراسات.

يجب ألا يقل عدد المفردات لكل طبقة عن ثلاثين مفردة في العينات الطبقية.

يفضل ألا تقل مفردات العينة عن عشرة أضعاف عدد متغيرات الدراسة.

قد يكون حجم عينة من عشر إلى عشرين مفردة مقبولاً إذا كان البحث تجريبيًّا وحجم الضبط والرقابة عالٍ ومبرر من الباحث.

كما صاغ إحصائيون آخرون تحديد الحجم الأمثل لاختيار العينة بصورة أخرى على النحو التالي:

في الدراسات الوصفية ينصح باستخدام ما نسبته 20 %من أفراد مجتمع صغير نسبيًّا (بضع مئات) ، و10% لمجتمع كبير(بضعة آلاف)، و5% لمجتمع كبير جدًّا (عشرات الآلاف) .

إذا كانت الدراسة تعتمد في تحليلها على العلاقات الارتباطية، يجب ألا يقل أفراد العينة عن عشرين مفردة، ويفضل أن يكون من خمسين إلى مائة مفردة.

يجب ألا يقل عدد عناصر المجموعة الواحدة في حالة الدراسات التجريبية أو شبه التجريبية ذات المجموعتين أو أكثر عن خمسة عشر عنصرًا.

كما اهتم أيضا باحثون آخرون بتحديد حجم العينة في ضوء تجانس أو عدم تجانس المجتمع، وذلك على النحو التالي:

إذا كان مجتمع الدراسة متجانسًا تقريبًا، وأراد الباحث درجة عالية من الدقة، فإن العينة تكون(عشوائية بسيطة) بحجم 23%.

أما إذا أراد الباحث تحقيق درجة مناسبة من الدقة، فإن العينة تكون(عشوائية بسيطة) بحجم 10%.

إذا كان مجتمع الدراسة غير متجانس وبه مجموعات متساوية الحجم تقريبًا، وأراد الباحث درجة عالية من الدقة، فإن العينة تكون(عشوائية بسيطة) بحجم.23% ، و(عشوائية طبقية) بحجم 10%.

أما إذا أراد الباحث تحقيق درجة مناسبة من الدقة، فإن العينة تكون(عشوائية بسيطة) بحجم 13%.

إذا كان مجتمع الدراسة غير متجانس، وكانت المجموعات فيه غير متساوية الحجم تقريبًا، وأراد الباحث تحقيق درجة عالية من الدقة، فإن العينة تكون(عشوائية طبقية) بحجم 10 %وعشوائية بسيطة بحجم 33%.

أما إذا أراد الباحث تحقيق درجة مناسبة من الدقة، فإن العينة تكون(عشوائية بسيطة) بحجم 23%.

إذا كان مجتمع الدراسة غير متجانس، وكانت المجموعات فيه صغيرة ومتطرفة، وأراد الباحث تحقيق درجة عالية من الدقة، فإن العينة تكون(عشوائية طبقية)بحجم 13%.

أما إذا أراد الباحث تحقيق درجة مناسبة من الدقة، فإن العينة تكون(عشوائية طبقية) بحجم 10%.

الفصل الثانـي

عرض البيانات الإحصائية

إن البيانات الخام التي يتم جمعها عن ظاهرة معينة لا يمكن وصفها وتفسيرها وهي في هيئتها الاولى (بيانات أولية) الأمر الذي يتطلب تلخيص ووضع هذه البيانات في جداول خاصة او رسوم بيانية هندسية بهدف تسهيل دراستها وإجراء التحليل الاحصائي عليها.

و لكن قبل التطرق إلى عرض البيانات جدوليا لابد من الإشارة إلى أنواع الجداول التكرارية و أنواع التكرارات.

1/ أنواع الجداول التكرارية: نميز بين نوعين من الجداول التكرارية هما:

1.1. الجدول التكراري : هو عبارة عن صورة تنقل المعلومات دون الإنقاص منها، من حالتها الأولى إلى حالة جديدة تتسم بالتنظيم و الرتيب و السهولة و الوضوح .و تختلف طرق ترتيب المعلومات في الجدول الإحصائي، باختلاف الأسلوب المستخدم و المنهج المتبع في الدراسة، كما تختلف الجداول الإحصائية باختلاف و تنوع المعطيات ، كأن تكون كمية أو كيفية، بسيطة أو مركبة .وعلى العموم فالكتابة النظرية للجدول الإحصائي تكون على النحو التالي :

2.1. الجدول التكراري المزدوج:

يستعمل الجدول المزدوج أو المركب عند دراسة خاصيتين في نفس الوقت في مجتمع ما، و توضع المعلومات الإحصائية كما يلي:

-الخاصية الأولى أفقيا.

-الخاصية الثانية عموديا.

مثال : الجدول التالي يوضح المستوى التعليمي و ممارسة الرياضة لـ 100 شخص

ممارسة الرياضة / المستوى التعليمي	يمارس	لا يمارس	المجموع
ابتدائي	10	02	12
متوسط	08	05	13
ثانوي	18	10	28
جامعي	32	18	50
المجموع	65	35	100

2/ أنواع التكرارات :

1.2. التكرار النسبي : يحسب التكرار النسبي بقسمة تكرار المجموعة على مجموع التكرارات، أي أن:

$$\text{التكرار النسبي} = \frac{\text{تكرار المجموعة}}{\text{مجموع التكرارات}} = \frac{f_i}{\sum f_i}$$

2.2. التكرار النسبي المئوي : يساوي التكرار النسبي ضرب 100.

3.2. التكرار التجميعي (المتجمع)الصاعد :وهو عبارة عن تكرار أية قيمة f_i أو فئة مضاف إليها مجموع تكرارات الفئات السابقة .بمعنى أن التكرار التجميعي للقيمة أو الفئة الأولى هو عبارة عن التكرار البسيط الأول n_1، التكرار التجميعي الصاعد للقيمة أو الفئة الثانية هو n_2+n_1 ، و في الأخير فإن التكرار التجميعي الصاعد للفئة الأخيرة يساوي $\sum f_i$.

4.2. التكرار التجميعي(المتجمع) النازل : وهو عبارة عن مجموع التكرارات مطروحا منه تكرارات الفئات السابقة بمعنى التكرار التجميعي النازل للفئة الأولى يساوي Σf_i و التكرار التجميعي النازل للفئة الثانية هو $\Sigma f_i - n_1$ أما التكرار التجميعي النازل للفئة الأخيرة يساوي التكرار المطلق للقيمة أو الفئة الأخيرة.

3/ عرض البيانات جدوليا:

يمكن عرض البيانات في صورة جدول تكراري، ويختلف شكل الجدول طبقا لنوع البيانات، وحسب عدد المتغيرات،وفيما يلي عرض بيانات متغير (وصفي أو كمي) في شكل جدول تكراري بسيط.

1.3. عرض بيانات المتغير الوصفي في شكل جدول تكراري بسيط

إذا كنا بصدد دراسة ظاهرة ما تحتوي على متغير وصفي واحد، فإنه يمكن عرض بياناته في شكل جدول تكراري بسيط، وهو جدول يتكون من عمودين، أحدهما به مستويات (مجموعات) المتغير، والثاني به عدد المفردات (التكرارات) لكل مستوى (مجموعة).

والمثال التالي يبين لنا كيف يمكن تبويب البيانات الوصفية الخام في شكل جدول تكراري.

مثال :

فيما يلي بيانات عينة مكونة من 40 طالب عن نوع الرياضة التي يمارسونها:

كرة قدم	سباحة	كرة اليد	سباحة	كرة اليد	سباحة	عدو	سباحة
كرة اليد	كرة قدم	كرة اليد	عدو	سباحة	كرة اليد	ملاكمة	كرة اليد
عدو	كرة اليد	كرة قدم	سباحة	كرة اليد	كرة اليد	عدو	سباحة
كرة اليد	سباحة	كرة اليد	كرة قدم	ملاكمة	عدو	ملاكمة	عدو
سباحة	كرة اليد	عدو	ملاكمة	كرة قدم	كرة اليد	عدو	سباحة

المطلوب:

- ما هو نوع المتغير؟ وما هو المعيار المستخدم في قياس البيانات؟
- اعرض البيانات في شكل جدول تكراري ثم كون التوزيع التكراري النسبي.
- علق على النتائج.

الحل : نوع الرياضة (كرة قدم – سباحة – كرة اليد – عدو– ملاكمة) متغير وصفي، تقاس بياناته بمعيار اسمي.

- لعرض البيانات في شكل جدول تكراري ، يتم إتباع الخطوات التالية:

- تكوين جدول تفريغ البيانات: وهو جدول يحتوي على علامات إحصائية، كل علامة تعبر عن تكرار للمجموعة التي تنتمي إليها نوع الرياضة ، وكل خمس علامات

تكون حزمة إحصائية، كما هو مبين بالجدول التالي:

جدول تفريغ البيانات

نوع الرياضة	العلامات الإحصائية	عدد الطلبة (التكرارات)
كرة قدم	𝍸	5
سباحة	𝍸 𝍸	10
كرة اليد	𝍸 𝍸 ///	13
عدو	𝍸 ///	8
ملاكمة	////	4
المجموع		40

- تكوين الجدول التكراري.

وهو نفس الجدول السابق، باستثناء العمود الثاني، ويأخذ الصورة التالية:

جدول التوزيع التكراري

نوع الرياضة	عدد الطلبة (التكرارات) *(fi)*	التوزيع التكراري النسبي
كرة قدم	5	$\left(\frac{5}{40}\right) = 0.125$
سباحة	10	$\left(\frac{10}{40}\right) = 0.25$
كرة اليد	13	$\left(\frac{13}{40}\right) = 0.325$

عدو	8	$\left(\frac{8}{40}\right)=0.20$
ملاكمة	4	$\left(\frac{4}{40}\right)=0.10$
المجموع	40	1.00

- التوزيع التكراري النسبي:

يحسب التكرار النسبي بقسمة تكرار المجموعة على مجموع التكرارات

أما التكرار النسبي المئوي فيساوي التكرار النسبي ضرب 100

والعمود الثالث في الجدول يعرض التكرار النسبي للطلبة الممارسين حسب نوع الرياضة .

التعليق:من الجدول أعلاه يلاحظ أن نسبة الطلبة الذين يمارسون رياضة "كرة اليد" في العينة هي 32.5 % وهي أكبر نسبة مما يدل على أن التوجه السائد في ممارسة الرياضة هو كرة اليد تليه السباحة، بينما نجد أن نسبة الطلبة الذين يمارسون رياضة "الملاكمة" حوالي 10.0 % وهي أقل نسبة.

مثال :

فيما يلي بيانات عن المستوى التعليمي لعينة من 50 فرد.

متوسط	أمي	ثانوي	متوسط	ثانوي	دراسات عليا	متوسط	ابتدائي
أمي	متوسط	ثانوي	ثانوي	متوسط	ثانوي	ابتدائي	متوسط
ابتدائي	ثانوي	أمي	جامعي	ثانوي	ابتدائي	أمي	ثانوي
متوسط	ابتدائي	متوسط	ثانوي	ابتدائي	متوسط	جامعي	متوسط
ثانوي	متوسط	ابتدائي	ثانوي	أمي	ابتدائي	ثانوي	ابتدائي
جامعي	ثانوي	جامعي	ابتدائي	جامعي	دراسات عليا	ثانوي	ثانوي
متوسط	أمي						

والمطلوب:

1- اعرض البيانات في شكل جدول تكراري.

2- كون التوزيع التكراري النسبي، ثم علق على النتائج.

الحل

1-عرض البيانات في شكل جدول تكراري:

المستوى التعليمي (أمي- ابتدائي_ متوسط- ثانوي- جامعي- دراسات عليا) متغير وصفي ترتيبي، ويمكن عرض البيانات أعلاه في شكل جدول تكراري بإتباع الآتي:

2-تكوين جدول تفريغ البيانات:

جدول تفريغ البيانات

المستوى التعليمي	العلامات الإحصائية	عدد الأفراد (التكرارات)
أمي	𝍸 /	6
ابتدائي	𝍸 𝍸	10
متوسط	𝍸 𝍸 //	12
ثانوي	𝍸 𝍸 𝍸	15
جامعي	𝍸	5
دراسات عليا	//	2
المجموع		50

- تكوين جدول التوزيع التكراري و التكرار النسبي:

جدول : التوزيع التكراري و التكرار النسبي

المستوى التعليمي	عدد الأفراد (التكرارات) *(fi)*	التوزيع التكراري النسبي
أمي	6	0.12
ابتدائي	10	0.20
متوسط	12	0.24
ثانوي	15	0.30
جامعي	5	0.10
دراسات عليا	2	0.04

المجموع	50	1.00

بتطبيق المعادلة السابقة يمكن حساب التكرارات النسبية، والعمود الثالث في الجدول اعلاه يبن هذا التوزيع، ومن التوزيع النسبي يلاحظ أن حوالي 30 % من أفراد العينة ممن لديهم مؤهل ثانوي، بينما يكون نسبة الأفراد ممن لديهم مؤهل اقل من الثانوي (متوسط، ابتدائي، أمي) أكثر من 56%، أما نسبة الأفراد الحاصلين على مؤهل دراسات عليا حوالي 4 % وهي أقل نسبة.

ملاحظات على الجدول

عند تكوين جدول ما لعرض البيانات، يجب مراعاة الآتي:

كتابة رقم للجدول.

كتابة عنوان للجدول.

لكل عمود من أعمدة الجدول عنوان يدل على محتواه.

يجب كتابة مصدر البيانات في الجدول.

2.3. عرض بيانات المتغير الكمي في شكل جدول تكراري بسيط

بنفس الأسلوب السابق المتبع في تكوين جدول تكراري، يمكن أيضا عرض بيانات المتغير الكمي في شكل جدول تكراري بسيط، ويتكون هذا الجدول من عمودين، الأول يحتوي على فئات تصاعدية للقراءات التي يأخذها المتغير، والثاني يشمل التكرارات أو عدد المفردات التي تنتمي قراءاتها للفئة المناسبة لها، فجدول التوزيع التكراري هو مجموعة من البيانات التي توضع بشكل منظم في جدول بهدف تلخيص تلك البيانات للوصول بسهولة اتخاذ القرار بإجراء ما والجدول هذا يشمل على عدد من الفئات المتساوية يقابل كل منها التكرار المناسب من البيانات حيث يتم حصر كل البيانات في الجدول والمعروف بجدول التوزيع التكراري.

الخطوات اللازمة لتكوين جدول التوزيع التكراري:

1-تحديد أكبر قيمة في البيانات وأصغر قيمة فيها

2-ايجاد المدى (Range) و يرمز له بالرمز R

3-المدى = أكبر قيمة – أصغر قيمة

4-تحديد عدد الفئات C: يجب ألا يكون عدد الفئات قليل لنفقد بعض البيانات أو كبيرة لنفقد صفة الاختصار وخلق فئات خالية من التكرار ولذا نهتم بمدى الاختلاف بين البيانات للوصول لعدد الفئات والذي يتراوح بين 5 ، 20 فئة ويمكن تطبيق القاعدة التالية :

$$C = 1 + (3.322 \log_{10}(n))$$

5-حساب طول الفئة L: حسب العلاقة التالية:

طول الفئة = المدى / عدد الفئات

$$L = \frac{R}{C}$$

6- تحديد الحد الأدنى لكل فئة والحد الأعلى لها كما يأتي:

- في حالة المتغير المتقطع: و يسمى بالتوزيع المغلق أي ان جميع الفئات لها حدود دنيا و حدود عليا و مثال ذلك عدد الأهداف المسجلة مثلا فالقيم تكون صحيحة موجبة ولا تقبل الكسور

الحد الأدنى = أصغر قيمة

الحد الأعلى = الحد الأدنى للفئة + (طول الفئة -1)

مثال : إذا كان طول الفئة = 10، و أصغر قيمة مسجلة هي 20 إذا تكون الفئات كالتالي:

الفئة الاولى: 20- 29 – الفئة الثانية : 30- 39 – الفئة الثالثة : 40-49 و هكذا

- في حالة المتغير المستمر: أي أن القيم مستمرة و تقبل الكسور مثل الأطوال و الأوزان

الحد الأدنى = أصغر قيمة

الحد الأعلى = الحد الأدنى للفئة + طول الفئة

مثال : إذا كان طول الفئة = 10، و أصغر قيمة مسجلة هي 20 إذا تكون الفئات كالتالي:

الفئة الاولى: 20– 30 – الفئة الثانية : 30– 40 – الفئة الثالثة : 40–50 و هكذا

6– تحديد تكرار الفئة وهو عبارة عن جميع المفردات التي تقع ضمن حدود الفئة من حيث القيمة العددية ويرمز له ب (fi) .

7– ايجاد مركز الفئة وهو القيمة التي تقع في منتصف الفئة ويرمز له ب (Mi) وفق الصيغة التالية:

مركز الفئة = (الحد الأعلى + الحد الأدنى) / 2

مثال: حالة المتغير المتقطع

البيانات التالية تمثل عدد الأهداف التي سجلها (20) رياضي في كرة السلة

Xi:65,56,89,83,75,60,70,65,55,67,45,65,74,72,62,48,69,65,49,49

المطلوب :

- تمثيل هذه البيانات في جدول توزيع تكراري
- التوزيع التكراري النسبي
- حساب مراكز الفئات

الحل :

باتباع الخطوات السابقة نجد :

1– أكبر قيمة = 89 و أصغر قيمة 45

2–المدى = أكبر قيمة – أصغر قيمة

المدى = 89– 45 = 44

3–تحديد عدد الفئات :

$C = 1 + (3.322 \log_{10}(n)) = 1 + (3.322*\log 20) = 5.32 \approx 5$

4– طول الفئة = المدى / عدد الفئات = 44/ 5 = 8.8 ≈ 9

والآن نضع البيانات في جدول تكراري كالتالي:

الفئات	العلامات الإحصائية	التكرار	التكرار النسبي	مراكز الفئات
53 – 45	////	4	0.20	(45+53)/2=49
62 – 54	/////	5	0.25	(54+62)/2=58
71 – 63	//// /	6	0.30	(63+71)/2=67
80 –72	///	3	0.15	(72+80)/2=76
89 –81	//	2	0.10	(81+89)/2=85
المجموع	n=20	20	1.00	

يتضح من الجدول أعلاه أن كل فئة تمثل عدد الأهداف وأن التكرارات تمثل عدد الرياضيين الذين سجلوها.

مثال: حالة المتغير المستمر

البيانات التالية تمثل تقييم (16) طالبا لأحد البرامج التكوينية ، حيث تعطى القيمة 100 لأعلى درجة و القيمة 0 لأضعف درجة و بناءا عليه نكون الجدول التكراري.

Xi:50, 82.5 ,75 ,60.5 ,76.6 ,80 ,74 ,66.5 ,78.5 ,72.5 ,84 ,71.5

,70.5 ,70 ,72 ,58

المطلوب :

- تمثيل هذه البيانات في جدول توزيع تكراري
- التوزيع التكراري النسبي
- حساب مراكز الفئات

<u>الحل :</u>

اكبر قيمة = 84 و أصغر قيمة = 50

المدى = 84 – 50 = 34

عدد الفئات = 1 +(3.322لو16) = 5

طول الفئة = المدى / عدد الفئات = 5/34 ≈ 7 والان نضع البيانات في جدول تكراري كالتالي:

الفئات	العلامات الإحصائية	التكرار	التكرار النسبي	مراكز الفئات
50– 57	/	1	0.0625	(50+57)/2=53.5
57 –64	//	2	0.1250	(57+64)/2=60.5
64 –71	///	3	0.1875	(64+71)/2=67.5
71 – 78	///// /	6	0.3750	(71+78)/2=74.5
78 - 85	////	4	0.2500	(78+85)/2=81.5
المجموع	n=16	16	1.00	

يتضح من الجدول أعلاه أن كل فئة تمثل درجة التقييم وان التكرارات تمثل عدد الطلاب.

كما يمكن أن نحتاج إلى معلومات إضافية عن البيانات، فمثلا :قد نحتاج إلى معرفة المفردات التي تقل قيمتها أو تزيد عن حد معيّن، وهذه المعلومات نحصل عليها من خلال إيجاد التكرارات المتجمعة الصاعدة والنازلة.

مثال : من المثال السابق أوجد التكرار المتجمع الصاعد و النازل

الحل :

الجدول التالي يوضح التكرار النسبي و التكرار النسبي المئوي و التكرار المتجمع الصاعد و النازل

الفئات	التكرار *fi*	التكرار النسبي	التكرار النسبي المئوي	التكرار المتجمع الصاعد	التكرار المتجمع النازل
57 −50	1	0.0625.	6.25	**1**	**16**
64 −57	2	0.1250	12.50	3=2+1	15=1−16
71 −64	3	0.1875	18.75	6=3+3	13=2−15
78 − 71	6	0.3750	37.50	12=6+6	10=3−13
85 − 78	4	0.2500	25.00	**16**=4+12	**4**=6−10
المجموع	16	1.00	100		

4/ عرض البيانات بيانيا :

العرض البياني للبيانات، هو أحد طرق التي يمكن استخدامها في وصف البيانات، من حيث شكل التوزيع ومدى تمركز البيانات، وفي كثير من النواحي التطبيقية يكون العرض البياني أسهل وأسرع في وصف الظاهرة محل الدراسة، وتختلف طرق عرض البيانات بيانيا حسب نوع البيانات المبوبة في شكل جدول تكراري، وفيما يلي عرض للأشكال البيانية المختلفة.

1.4. العرض البياني للبيانات الكيفية (الوصفية):

يمكن عرض البيانات الخاصة بمتغير وصفي في شكل دائرة بيانية أو أعمدة بيانية أو غيرهما، لنتمكن من خلاله وصف ومقارنة مجموعات أو مستويات هذا المتغير.

1- الدائرة البيانية

لعرض بيانات المتغير الوصفي في شكل دائرة، يتم توزيع الـ 360^{o} درجة حسب التكرار النسبي لمجموعات المتغير، حيث تحدد مقدار الزاوية الخاصة بالمجموعة رقم r بتطبيق المعادلة التالية:

التكرار النسبي للمجموعة $\times 360^{o}$ = مقدار الزاوية

مثال: لديك الجدول التكراري التالي و الذي يبين توزيع عينة مكونة من 1000 طالب حسب التخصص

التخصص	هندسة	أدب	تربية بدنية	إقتصاد	المجموع
عدد الطلبة	300	260	340	100	1000

و المطلوب منك تمثيل البيانات أعلاه في دائرة بيانية.

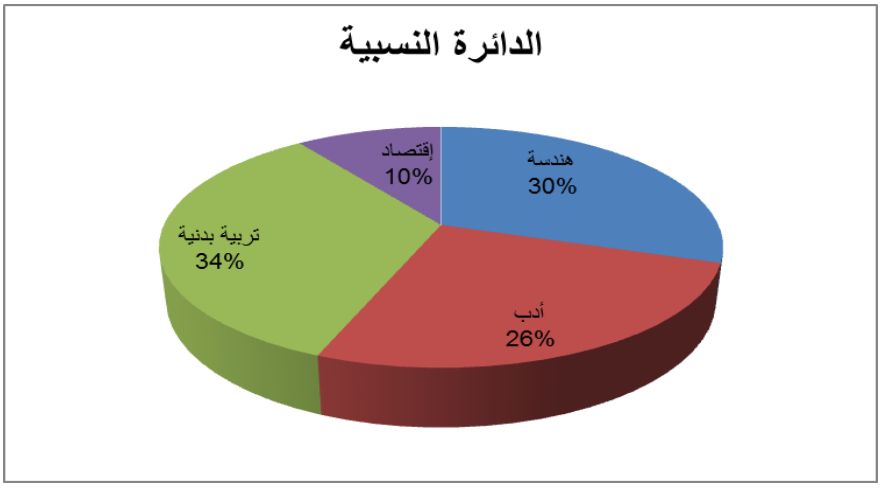
الحل :

- تحديد مقدار الزاوية المخصصة لكل تخصص، بتطبيق المعادلة:

التكرار النسبي × 360^o = مقدار الزاوية المخصص للتخصص

التخصص	عدد الطلبة	التكرار النسبي	مقدار الزاوية
هندسة	300	0.30	$360 \times 0.30 = 108^o$
أدب	260	0.26	$360 \times 0.26 = 93.6^o$
تربية بدنية	340	0.34	$360 \times 0.30 = 122.4^o$
إقتصاد	100	0.10	$360 \times 0.10 = 36^o$
المجموع	500	1.00	360^o

- رسم الدائرة : يتم رسم دائرة وتقسيمها إلى أربع أجزاء لكل تخصص جزء يتناسب مع مقدار الزاوية المخصصة له، كما هو مبين في الشكل التالي:

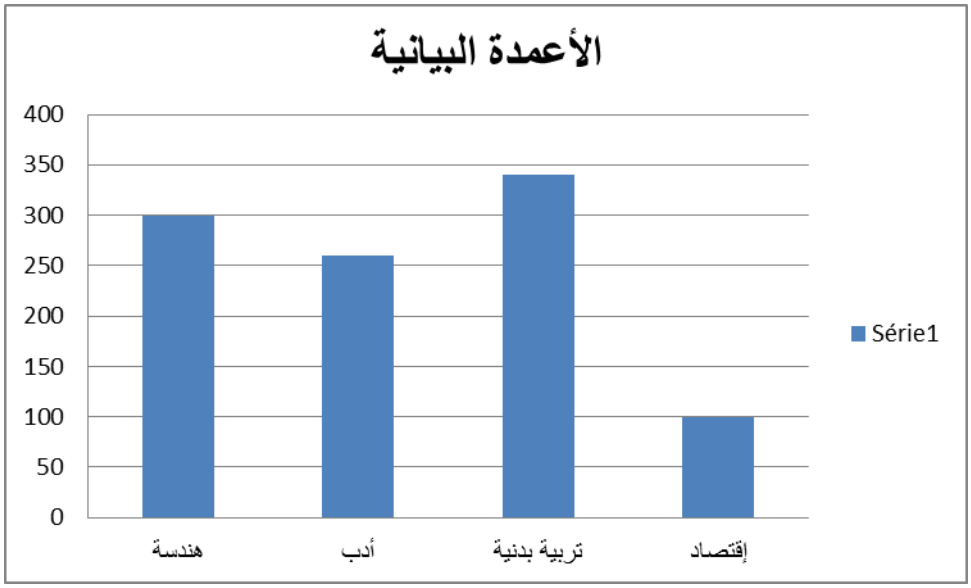


2- الأعمدة البيانية :

وهي عبارة عن مجموعة من المستطيلات الافقية او العمودية قواعدها متساوية وتمثل الصفة التي تم على اساسها التبويب (سنة ,محافظة , وهكذا) وارتفاعاتها تمثل البيانات المقابلة لتلك الصفة.

مثال :

استخدم بيانات الجدول التكراري السابق لرسم الأعمدة البيانية



2.4. العرض البياني للبيانات الكمية:

يمكن عرض البيانات الخاصة بمتغير كمي في شكل مدرج تكراري أو مضلع تكراري أو منحنى تكراري أو غيرهم، لنتمكن من خلاله وصف ومقارنة مجموعات أو مستويات هذا المتغير.

1- المدرج التكراري Histogram

لمدرج التكراري هو التمثيل البياني للجدول التكراري البسيط الخاص بالبيانات الكمية ، وهو عبارة عن أعمدة بيانية متلاصقة، حيث تمثل التكرارات على المحور الرأسي، بينما تمثل قيم المتغير (حدود الفئات) على المحور الأفقي، ويتم تمثيل كل فئة بعمود، ارتفاعه

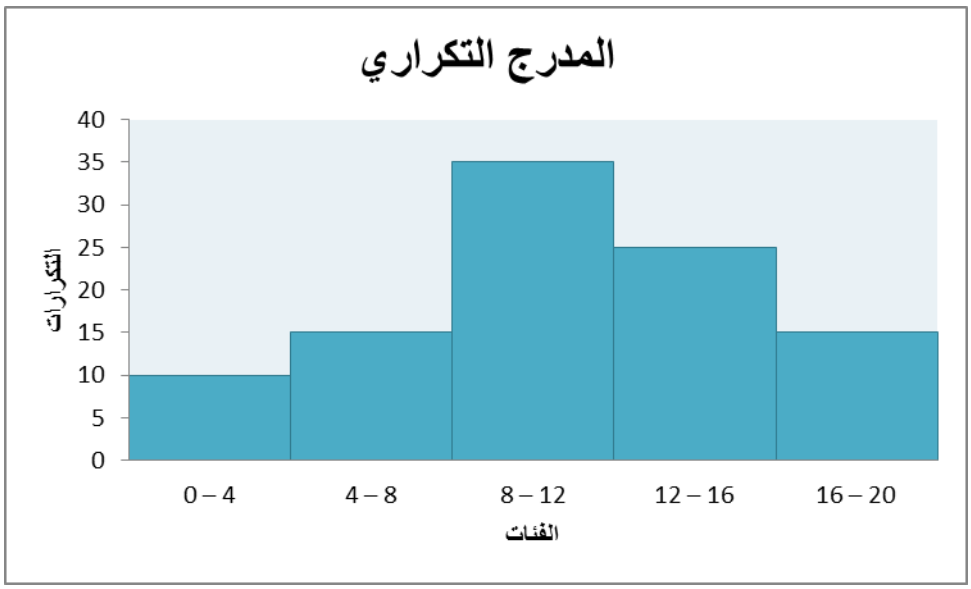
هو تكرار الفئة، وطول قاعدته هو طول الفئة.أما إذا كان التكرار نسبي فيعرف بالمدرج التكراري النسبي.

و لرسم المدرج التكراري يتم إتباع الخطوات التالية:

- رسم محوران متعامدان، الرأسي ويمثل التكرارات، الأفقي ويمثل الأوزان.
- كل فئة تمثل بعمود ارتفاعه هو تكرار الفئة، وطول قاعدته هو طول الفئة.
- كل عمود يبدأ من حيث انتهى به عمود الفئة السابقة.

مثال: لديك الجدول التالي الذي يبين درجات عينة مكونة من 100 طالب في مادة الإحصاء و المطلوب منك رسم المدرج التكراري

الفئات	التكرارات *fi*
0 – 4	10
4 – 8	15
8 – 12	35
12 – 16	25
16 – 20	15
المجموع	100



2- المضلع التكراري:

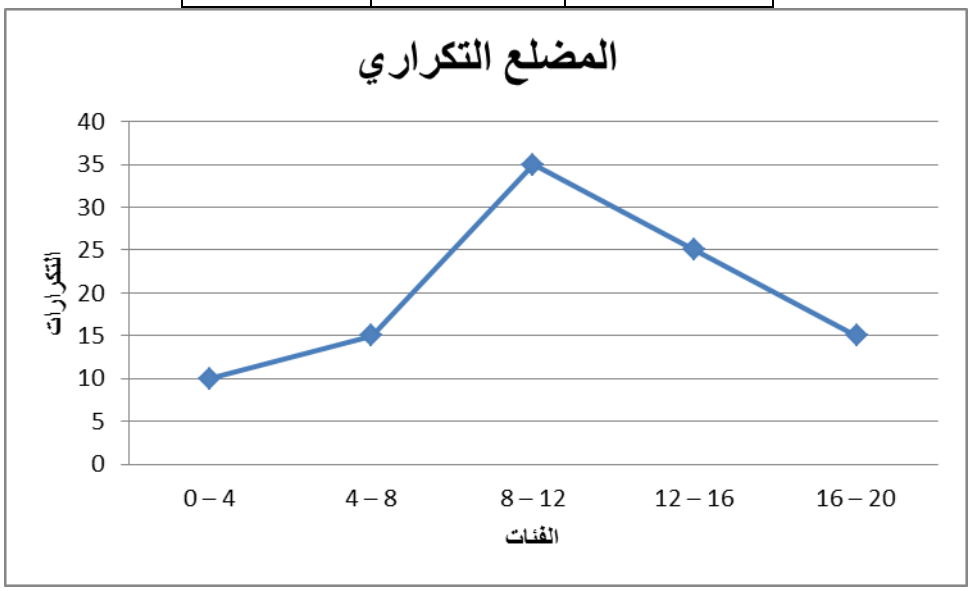
هو تمثيل بياني أيضا للجدول التكراري البسيط، حيث تمثل التكرارات على المحور الرأسي، ومراكز الفئات على المحور الأفقي، ثم التوصيل بين الإحداثيات بخطوط منكسرة، وبعد ذلك يتم توصيل طرفي المضلع بالمحور الأفقي.

مثال : استخدم بيانات الجدول التكراري السابق لرسم المضلع التكراري.

الحل : لرسم المضلع التكراري نتبع الخطوات التالية:

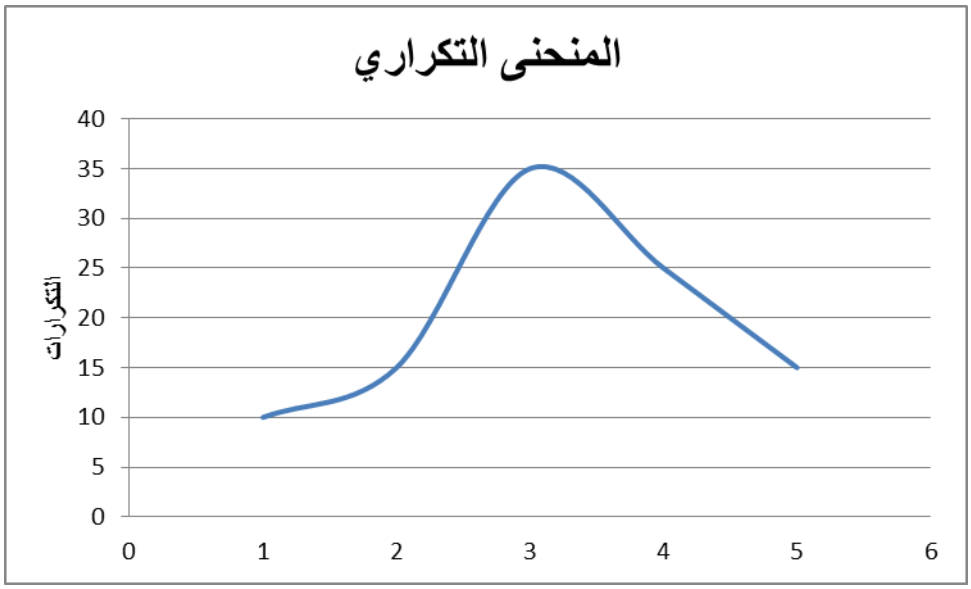
- حساب مراكز الفئات
- تلخيص نقط الاحداثيات
- التمثيل البياني لنقط الإحداثيات وتوصيلها بخطوط مستقيمة

الفئات	التكرارات *fi*	مراكز الفئات
4 – 0	10	2
8 – 4	15	6
12 – 8	35	10
16 – 12	25	14
20 – 16	15	18
المجموع	100	



3-المنحنى التكراري

بإتباع نفس الخطوات السابقة في رسم المضلع يمكن رسم المنحنى التكراري، ولكن يتم تمهيد الخطوط المنكسرة في شكل منحنى بحيث يمر بأكثر عدد من النقاط، وفي المثال السابق يمكن رسم المنحنى التكراري .

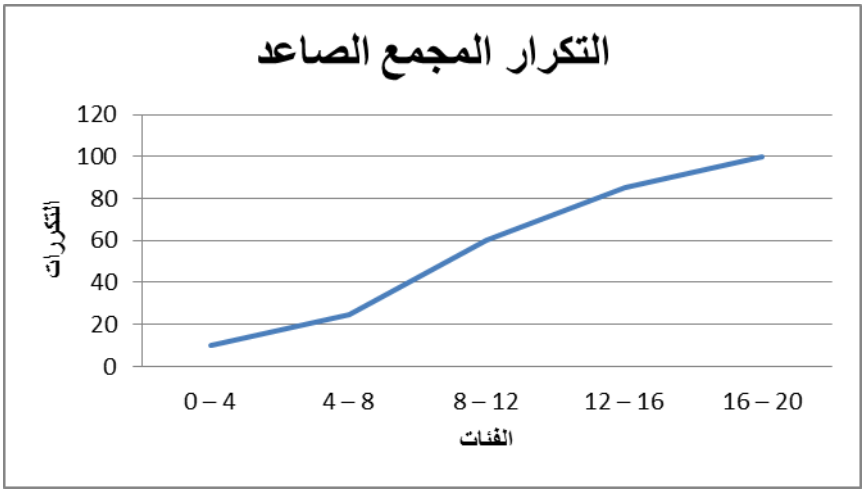


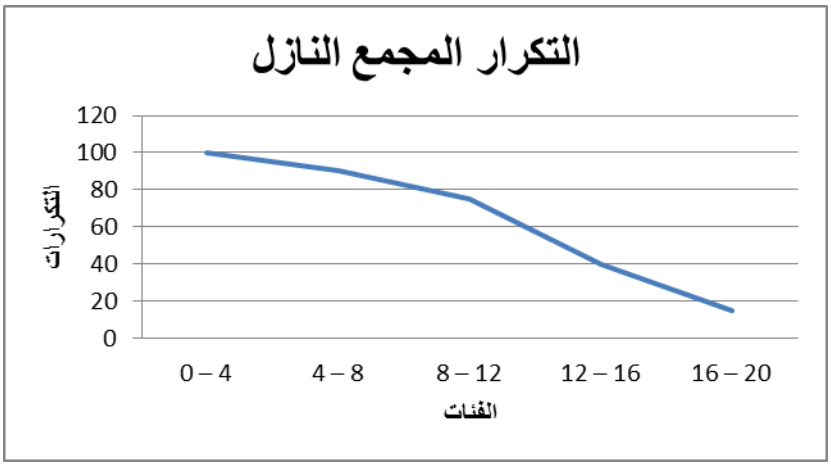
4- منحنى التكرارات المتجمعة الصاعدة و النازلة

يتم رسم منحنى التكرار المتجمع الصاعد عن طريق إيصال مجموعة النقاط ذات الإحداثيات التالية :الحدود العليا للفئات والتكرارات المتجمعة الصاعدة المقابل لها، أما منحنى التكرار المتجمع النازل فيتم رسمه بإيصال مجموعة النقاط ذات الإحداثيات : الحدود الدنيا للفئات والتكرارات المتجمعة النازلة المقابلة لها، وتسمى نقطة التقاطع بين المنحنيين بالوسيط، ولغرض رسم منحنى التكرار المتجمع الصاعد والنازل نأخذ المثال السابق.

الفئات	التكرارات *fi*	ت م ص	ت م ن
4 – 0	10	10	100
8 – 4	15	25	90
12 – 8	35	60	75

16 – 12	25	85	40
20 – 16	15	100	15
المجموع	100		





الفصل الثالث

مقاييس النزعة المركزية

في كثير من النواحي التطبيقية يكون الباحث في حاجة إلى حساب بعض المؤشرات التي يمكن الاعتماد عليها في وصف الظاهرة من حيث القيمة التي تتوسط القيم أو تنزع إليها القيم ، ومن حيث التعرف على مدى تجانس القيم التي يأخذها المتغير ، و أيضا ما إذا كان هناك قيم شاذه أم لا. و الاعتماد على العرض البياني وحده لا يكفي ، و لذا يتناول هذا الفصل و الذي يليه عرض بعض المقاييس الإحصائية التي يمكن من خلالها التعرف على خصائص الظاهرة محل الدراسة و كذا امكانية مقارنة ظاهرتين أو أكثر و من أهم هذه المقاييس مقاييس النزعة المركزية و التي تسمى أيضا بمقاييس الموضع او المتوسطات و هي القيم التي تتركز القيم حولها، وفي ما يلي عرض لأهم هذه المقاييس.

1/ الوسط الحسابي أو المتوسط الحسابي (The Arithmetic Mean)

من أهم مقاييس النزعة المركزية و أكثرها استخداما في النواحي التطبيقية و يعتبر الوسط الحسابي أكثر المقاييس استخداماً ويعرف كالآتي:

الوسط الحسابي لمجموعة من القيم عددها n هو القيمة التي لو حلت محل كل قيمة في مجموعة القيم لكان مجموع القيم الجديدة مساويا لمجموع القيم الأصلية و يمكن حسابه للبيانات المبوبة و غير المبوبة كما يلي:

1.1. للبيانات غير المبوبة (Ungrouped Data)

ببساطة الوسط الحسابي هو مجموع القيم مقسوماً على عددها، فإذا كان لدينا n من القيم و يرمز لها بالرمز $x_1, x_2, x_3 \ldots\ldots x_n$ فإن الوسط الحسابي لهذه القيم و نرمز له بالرمز $\bar{x}$، وتقرأ أكس بار ونحن هنا بصدد العينة وفي حالة المجتمع يرمز للوسط الحسابي بالرمز μ ويقرأ ميو وعدد القيم N أي أنَّ: الوسط الحسابي للمجتمع

$$\mu = \frac{x_1+x_2+x_3+\cdots\ldots+x_n}{N} \quad \text{أو} \quad \mu = \frac{\sum_{i=1}^{n} x_i}{N}$$

الوسط الحسابي للعينة

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad \text{أو} \qquad \bar{x} = \frac{x_1+x_2+x_3+\cdots\ldots+x_n}{n}$$

حيث x_i تمثل قيم العينة و N عدد القيم

وفي حال وجود أكثر من مجموعة فيكون الوسط الحسابي هو:

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2 + n_3\bar{x}_3 + \cdots\ldots + n_n\bar{x}_n}{n_1 + n_2 + n_3 + \cdots\ldots + n_n}$$

حيث أن $\bar{x}_1$ هو المتوسط الحسابي لمجموعة القيم التي عددها n_1

حيث أن $\bar{x}_2$ هو المتوسط الحسابي لمجموعة القيم التي عددها n_2

حيث أن $\bar{x}_3$ هو المتوسط الحسابي لمجموعة القيم التي عددها n_3

حيث أن $\bar{x}_n$ هو المتوسط الحسابي لمجموعة القيم التي عددها n_n

مثال: أوجد المتوسط الحسابي لدرجات 5 طلاب في مادة الاحصاء و التي كانت بياناتهم كالتالي: 13، 16، 8، 11، 7

الحل: مجموع القيم = 13 + 16 + 8 + 11 + 7 = 55

عدد القيم = 5 إذن الوسط الحسابي = 55 ÷ 5 = 11

أو $\bar{x} = \frac{7+11+8+16+13}{5} = 11$

2.1. للبيانات المبوبة (Grouped Data) أو البيانات المجدولة

توجد ثلاث طرق لحساب الوسط الحسابى من البيانات المبوبة هى :

1.2.1. الوسط الحسابى بطريقة مراكز الفئات :

من أسهل الطرق و أكثرها استخداما ، فمن المعلوم أن القيم الأصلية لا يمكن معرفتها من

جدول التوزيع التكراري حيث أن هذه القيم موضوعة في شكل فئات ، و لذا يتم التعبير عن كل قيمة من القيم التي تقع داخل حدود الفئة بمركز هذه الفئة ومن ثم يؤخذ في الاعتبار ان مركز الفئة هو القيمة التقديرية لكل مفردة تقع في هذه الفئة.

فإذا كانت k هي عدد الفئات و كانت $x_1, x_2, x_3, \ldots\ldots, x_n$ هي مراكز هذه الفئات $f_1 . f_2 . f_3 \ldots\ldots f_k$ هي التكرارات فإن الوسط الحسابي يحسب بالمعادلة التالية :

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + x_3 f_3 + \cdots \ldots . + x_n f_n}{f_1 + f_2 + f_3 + \cdots \ldots + f_n}$$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i f_i}{\sum_{i=0}^{n} f_i}$$

حيث x_i هي مركز الفئة = $\frac{(\text{للفئة الأدنى الحد} + \text{للفئة الأعلى الحد})}{2}$

fi هي تكرار كل فئة

مثال : باستخدام البيانات الواردة في الجدول التالي لمعدلات ذكاء 60 تلميذ احسب الوسط الحسابي :

معدل الذكاء	10 – 19	20 – 29	30 – 39	40 – 49	50 – 59	60 – 69	70 – 79	المجموع
التكرار	4	6	8	12	14	9	7	60

الحل:

لحساب الوسط الحسابي نتبع الخطوات التالية:

- حساب مراكز الفئات x_i
- ضرب مركز كل فئة في التكرار المناظر لها $x_i.f_i$ ثم حساب المجموع $\sum x_i.f_i$
- حساب الوسط الحسابي بالمعادلة السابقة

و بعبارة أخرى نكون جدول تكراري يضم مراكز الفئات x_i وآخر يشمل $x_i.f_i$ بالشكل التالي :

الفئات	التكرار f_i	مركز الفئة x_i	$x_i.f_i$
10 – 19	4	14.5	58
20 – 29	6	24.5	147
30 – 39	8	34.5	276
40 – 49	12	44.5	534
50 – 59	14	54.5	763
60 – 69	9	64.5	580.5
70 – 79	7	74.5	521.5
المجموع	60		2880

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i} = \frac{2880}{60} = 48$$

الوسط الحسابي = 48

2.2.1. باستخدام الوسط الفرضي (بطريقة الانحرافات)

إذا كان A مركز فئة افتراضي يمكن أن يأخذ مركز الفئة التي يقابلها أكبر تكرار أو مركز أي فئة أخرى و $d_i = x_i - A$ فإننا نستخدم العلاقة التالية:

$$\bar{x} = \frac{\sum f_i d_i}{\sum f_i} + A$$

مثال : بنفس معطيات المثال السابق نكون جدول تكراري يضم مراكز الفئات x_i و انحرافات مراكز الفئات عن وسطها الفرضي و المعبر عنه بالرمز d على اعتبار أن A = 44.5 بالشكل التالي:

الحل :

الفئات	التكرار f_i.	مركز الفئة x_i	$d_i = x_i - A$	$f_i . d_i$
10 – 19	4	14.5	14.5 – 44.5 = – 30	– 120
20 – 29	6	24.5	– 20	– 120
30 – 39	8	34.5	– 10	– 80
40 – 49	12	44.5	0	0
50 – 59	14	54.5	10	140
60 – 69	9	64.5	20	180
70 – 79	7	74.5	30	210
المجموع	60		0	210

$$\overline{x} = \frac{\sum f_i d_i}{\sum f_i} + A = \frac{210}{60} + 44.5 = 48$$

3.2.1. باستخدام الوسط الفرضي مع الاختصار (الطريقة المختصرة)

يمكن القسمة على طول الفئة (10) لعمود الفروقd ووضع النواتج في عمود جديد كالتالي:

الفئات	التكرار f_i.	مركز الفئة x_i	$d = x_i - A$	d /10	f_i (d/10).
10 – 19	4	14.5	– 30	– 3	4 ×(– 3) = – 12
20 – 29	6	24.5	– 20	– 2	– 12
30 – 39	8	34.5	– 10	– 1	– 8
40 – 49	12	44.5	0	0	0
50 – 59	14	54.5	10	1	14

60 – 69	9	64.5	20	2	18
70 – 79	7	74.5	30	3	21
المجموع	60				21

$$\text{الوسط الحسابي} = 44.5 + \frac{21}{60} \times 10 = 48$$

1.3. الوسط الحسابي المرجح (المعدل)

في بعض الأحيان يكون لكل قيمة من قيم المتغير أهمية نسبية تسمى أوزان او ترجيحات و عند عدم أخذ هذه الأوزان في الاعتبار عند حساب الوسط الحسابي تكون النتيجة غير دقيقة ، لذلك نستخدم العلاقة التالية في حساب الوسط الحسابي المرجح و الذي نرمز له بالرمز $\bar{w}$

$$\bar{w} = \frac{\sum x_i w_i}{\sum w}$$

حيث x_i تمثل قيم العينة و w_i تمثل الأوزان النسبية لقيم العينة أما w فتمثل مجموع الأوزان.

مثال : أخذنا خمسة طلاب و سجلنا درجات هؤلاء الطلاب في مادة الإحصاء و عدد ساعات الاستذكار في الأسبوع كما في الجدول التالي و المطلوب حساب الوسط الحسابي المرجح.

	13	20	18	14	10
عدد ساعات الاستذكار	2	3	3	2	1

نجد أن الوسط الحسابي غير المرجح لدرجات الطلاب هي :

$$\bar{x} = \frac{(13 + 20 + 18 + 14 + 10)}{5} = 15$$

و إذا أردنا أن نحسب الوسط الحسابي للدرجات x المرجحة بعدد ساعات الاستذكار w نجد مايلي :

$$\overline{W} = \frac{\sum x_i w_i}{\sum w} = \frac{(10x1)+(14x2)+(18x3)+(20x3)+(13x2)}{11} = 16.18$$

و هذا الوسط المرجح أكثر دقة من الوسط الحسابي غير المرجح

4.1. خصائص الوسط الحسابي

- يعتمد على جميع القيم والمشاهدات محل الدراسة
- هو نقطة اتزان المشاهدات
- انحراف الدرجة عن المتوسط يساوي بعدها عنه.
- المجموع الجبري لانحرافات القيم عن وسطها الحسابي يساوي صفراً
- مربع الانحرافات اقل ما يمكن عن الوسط
- إضافة قيمة ثابتة إلى الدرجات أو طرحها منها أو ضربها فيه أو قسمتها عليه مما يجعل المتوسط الحسابي يزداد أو يقل بقيمة ثابتة.
- المتوسط الحسابي يتأثر بالدرجات القريبة منه تأثرا قليلا ، بينما يتأثر بالدرجات البعيدة عنه تأثرا كبيرا .
- أقل مقاييس النزعة المركزية تأثرا بالتقلبات العينية
- يتأثر بالقيم المتطرفة والقيم الشاذة لذا لا يصلح للتوزيعات الملتوية بمعنى أن الوسط الحسابي للقيم 20، 30 ، 40 ،50 ، 120هو 260 ÷ 5 = 52 أكبر من غالبية القيم الموجودة بسبب القيمة 120 الشاذة عن باقي القيم الأخرى المتقاربة فيما بينها وهذا يقودنا لعدم صلاحية الوسط الحسابي لتوزيعات تكرارية شديدة الالتواء.

5.1. الحالات التي لا يصلح المتوسط الحسابي فيها :

- عندما تكون قيم التوزيع متجمعة في طرف واحد أكثر من الطرف الأخر.

- إذا كانت هناك قيم شاذة تتخلل التوزيع.
- إذا كانت البيانات مبوبة في فئات وكان التوزيع مفتوحا في أحد طرفيه.
- إذا كانت الفئات متباعدة نسبيا.
- لا يمكن إيجاده من الرسم البياني كالوسيط والمنوال.
- لا يمكن إيجاده من التوزيع التكراري المفتوح إلا في حال تقدير مراكز الفئات المفتوحة.
- عدم صلاحية الوسط الحسابي لتوزيعات تكرارية شديدة الالتواء.
- يصعب حسابه في حالة البيانات الوصفية.

2/ الوسط الهندسي:The Geometric Mean

1.2. للبيانات غير المبوبة :

إذا كانت : xn ,....., x2 , x1 تمثل مفردات(قيم)ظاهرة معينة X فإن المتوسط الهندسي لهذه القيم ، هو عبارة عن الجذر النوني لجداء هذه القيم وهو معطى بالعلاقة التالية:

$$G = \sqrt[n]{x_1 \times x_2 \times x_3 \times \dots \dots \times x_n}$$

ولغرض تسهيل العملية الحسابية في حال كون البيانات كبيرة فإنه يفضل استخدام أخذ لوغاريتم البيانات لتصبح صيغة المتوسط الهندسي و هذا بعد تحويل القانون على شكل لوغاريتم كما يلي:

$$G = (x_1 \times x_2 \times \dots \dots \times x_n)^{\frac{1}{n}}$$

$$Log\ G = \frac{1}{n} Log\ (x_1 \times x_2 \times \dots \dots \times x_n)$$

$$\boldsymbol{Log\ G = \frac{1}{n} \sum Log\ x_i}$$

فالوسط الهندسي للقيمتين 3 ، 12 هو الجذر التربيعي لناتج ضربهما ويساوي 6

لأن 3 × 12 = 36 و الجذر التربيعي لـ 36 هو 6

والوسط الهندسي للكميات 2 ، 4 ، 8 هو الجذر التكعيبي لناتج ضربهم ويساوي 4 لأن

$2 \times 4 \times 8 = 64$ و الجذر التكعيبي للعدد 64 هو 4

$$G = \sqrt[3]{(2)(4)(8)} = \sqrt[3]{64} = 4$$

مثال : أوجد الوسط الهندسي للقيم 2، 4، 4، 8 بطريقتين

الحل:

حاصل ضرب القيم يساوي 256 فيكون:

$$G = \sqrt[4]{(2)(4)(4)(8)}$$

$$G = \sqrt[4]{256} = 4$$

يمكن استخدام اللوغاريتمات:

$$\text{Log}_{10}(G) = (1/4)[\log(2) + \text{Log}(4) + \text{Log}(4) + \text{Log}(8)]$$

$$= (1/4)[0.30103 + 0.50206 + 0.50206 + 0.9039]$$

$$= (1/4)(2.40824)$$

$$= 0.60206$$

$$G = (10)^{0.60206}$$

$$G = 4$$

باستخدام جدول للوغاريتمات أو الحاسبة أو الحاسوب $(10)i^{0.60206} = 4$

2.2. للبيانات المبوبة :

إذا كانت : $x_1 , x_2 , \ldots\ldots , x_n$ تمثل مفردات(قيم)ظاهرة معينة X و كانت $n_1 , n_2 , \ldots\ldots n_n$ تمثل التكرارات المقابلة لها فإن المتوسط الهندسي لهذه القيم معرف بالعلاقة التالية:

$$G = \sqrt[\Sigma f_i]{(x_1 f_1)(x_2 f_2)(x_3 f_3) \ldots\ldots (x_n f_n)}$$

أو بالطريقة الثانية و التي نتبع فيها الخطوات التالية :

- تعيين مركز الفئات
- تحديد لوغاريتم مراكز الفئات
- ضرب مركز كل فئة بتكرارها
- نقسم مجموع حواصل الضرب على مجموع التكرارات و تكون النتيجة هي لوغاريتم الوسط الهندسي

$$Log\ (G) = \frac{\sum f_i Log\ (x_i)}{N}$$

مثال :

أحسب الوسط الهندسي للبيانات المبينة في الجدول الآتي:

الفئات	التكرارات(f_i)
20 – 40	20
40 – 60	35
60 – 80	80
80 – 100	40
100 – 120	15
120 – 140	10
المجموع	200

الحـل:

باستخدام قانون اللوغاريتم:

الفئات (X)	التكرارات (f_i)	مراكز الفئات (x_i)	$\text{Log}(x_i)$	$f_i\ Log(x_i)$
20 – 40	20	30	1.4771	29.5424
40 – 60	35	50	1.6989	59.4639
60 – 80	80	70	1.8450	147.6078
80 – 100	40	90	1.9542	78.1697
100 – 120	15	110	2.0413	30.6208
120 – 140	10	130	2.1139	21.1394
المجموع	200			**366.544**

$$Log\ (G) = \frac{\sum f_i Log\ (x_i)}{N}$$

Log (G) = 366.544 ÷ 200

= 1.8327

G = 68.03 ومنه

أو باستخدام القانون الثاني

$$G = \sqrt[\sum f_i]{(x_1{}^{f_1})(x_2{}^{f_2})(x_3{}^{f_3}) \dots \dots (x_n{}^{f_n})}$$

$$G = \sqrt[200]{(30^{20})(50^{35})(70^{80})(90^{40})(110^{15})(130^{10})}$$

G = 68.03

3.2. خصائص المتوسط الهندسي :

- يدخل في حسابه جميع القيم ولكنه أقل تأثرا بالقيم المتطرفة من المتوسط الحسابي؛
- لا يمكن حسابه من جداول التوزيع التكراري المفتوحة
- ليس له معنى إذا كانت إحدى القيم سالبة أو تساوي الصفر و بالتالي فالمجال هو

] 0 ، ∞+[؛

- يستخدم بأكثر واقعية عند وصف الظواهر النسبية؛
- قيمة المتوسط الهندسي لأي ظاهرة أصغر دائما من قيمة المتوسط الحسابي
- يستخدم في حساب الأرقام القياسية لمناسيب الأسعار، وعند حساب المعدلات (معدلات الفائدة، معدل النمو الاقتصادي، معدل النمو السكاني...).

3/ الوسط التوافقي:The Harmonic Mean

المتوسط التوافقي لمجموعة من القيم هو مقلوب المتوسط الحسابي لمقاليب هاته القيم و هو من المقاييس الخاصة التي تستخدم لتحديد معدلات السرعة ومتوسط الأسعار ومتوسط الكثافة السكانية.

1.3. للبيانات المبوبة :

الوسط التوافقي لمجموعة من القيم هو مجموع مقلوبات تلك القيم مقسوماً على عددها أي:

$$\frac{1}{H} = \frac{1}{N}\left(\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \cdots \ldots + \frac{1}{x_n}\right)$$

$$\frac{1}{H} = \frac{1}{N}\sum \frac{1}{x_i}$$

$$\boldsymbol{H = \frac{N}{\sum \frac{1}{x_i}}}$$

مثال : أحسب الوسط التوافقي للقيم 2 ، 4 ، 8

الحل :

$$\frac{1}{H} = \frac{1}{3}\left(\frac{1}{2} + \frac{1}{4} + \frac{1}{8}\right)$$

H =3.43

2.3. للبيانات غير المبوبة

نستخدم القانون التالي لحساب الوسط التوافقي للبيانات غير المبوبة.

$$H = \frac{\sum f_i}{\sum(\frac{f_i}{x_i})}$$

مثال : أحسب الوسط التوافقي للبيانات المبينة في الجدول الآتي:

الفئات (X)	التكرارات (f_i)	مراكز الفئات (x_i)	$\frac{f_i}{x_i}$
20 – 40	20	30	**0.66**
40 – 60	35	50	**0.70**
60 – 80	80	70	**1.14**
80 – 100	40	90	**0.44**
100 – 120	15	110	**0.13**
120 – 140	10	130	**0.08**
المجموع	200		$\sum\left(\frac{f_i}{x_i}\right) = 3.15$

الحـل:

بتطبيق القانون السابق ذكره (أعلاه) للبيانات المبوبة:

$$H = \frac{\sum f_i}{\sum(\frac{f_i}{x_i})} = \frac{200}{3.15} = 63.49$$

3.3. خصائص الوسط التوافقي

الوسط الحسابي ≥ الوسط الهندسي ≥ الوسط التوافقي

لا يتأثر بالقيم المتطرفة وهي ميزة له

عيبه أنه لا يستخدم مع الصفر والقيم السالبة

4/ الوسط التربيعي:(Quadratic Mean)

المتوسط التربيعي لأي مجموعة من القيم هو الجذر التربيعي للمتوسط الحسابي لمربعات تلك القيم ، ويستخدم غالباً في الفيزياء والإلكترونيات ويتم حساب المتوسط التربيعي للبيانات الأولية والمبوبة كما يلي:

1.4. للبيانات غير المبوبة:

هو الجذر التربيعي لمربعات القيم مقسومة على عددها أي:

$$Q = \sqrt{\frac{x_1{}^2 + x_2{}^2 + x_3{}^2 + \cdots \ldots + x_n{}^2}{N}}$$

$$\boldsymbol{Q = \sqrt{\frac{\sum x_i{}^2}{N}}}$$

مثال : أحسب الوسط التربيعي للقيم 3 ،4 – ، 6 ، 5 ، 2

الحل :

$$Q = \sqrt{\frac{(-4)^2 + (3)^2 + (6)^2 + (5)^2 + (2)^2}{5}}$$

$$Q = \sqrt{\frac{90}{5}} = 4.2426$$

2.4. للبيانات المبوبة:

هو الجذر التربيعي لمربعات تكرارات الفئات مضروبة في مراكز فئاتها و يمثلها القانون التالي:

$$\boldsymbol{Q = \sqrt{\frac{\sum f_i x_i{}^2}{\sum f_i}}}$$

حيث القيم هي x_n , ... , x_2 , x_1 :تقابل كل منها مركز الفئة f_n , ... , f_2 , f_1 على الترتيب.

مثال:أوجد الوسط التربيعي للبيانات المبينة في الجدول الآتي:

الفئات	10 –20	20 –30	30 –40	40 –50	المجموع
التكرارات	4	6	7	3	20

الحل:

الفئات	التكرارات (f_i)	مراكز الفئات(x_i)	x_i^2	$f_i\ x_i^2$
20–10	4	15	225	900
30–20	6	25	625	3750
40–30	7	35	1225	8575
50–40	3	45	2025	6075
المجموع	20			19300

$$Q = \sqrt{\frac{\sum f_i {x_i}^2}{\sum f_i}} = \sqrt{\frac{19300}{20}}$$

$$\boldsymbol{Q = \sqrt{965} = 31.06}$$

5/ الوسيط Median

الوسيط هو أحد مقاييس النزعة المركزية و الذي يأخذ في الاعتبار رتب القيم ، و يعرف الوسيط بالقيمة المشاهدة التي تتوسط البيانات بعد ترتيبها تصاعدياً أو تنازلياً. وعليه تكون عدد القيم (المفردات) قبله مساوي لعدد القيم بعده، و نميز بين حالتين في حساب الوسيط هما :

1.5. للبيانات غير المبوبة : ونميز فيه حالتين هما :

1-عدد المفردات أو القيم (n) فردي :

رتبة الوسيط = $\frac{N+}{2}$

قيمة الوسيط = القيمة التي تقابل رتبة الوسيط

مثال : إذا كان لدينا المفردات التالية 3، 5 ، 9 ، 11 ، 13 و التي سبق و أن رتبناها ترتيبا تصاعديا

الحل :

رتبة الوسيط = (5+1)/2 = 3

قيمة الوسيط = 9

2-عدد المفردات أو القيم (n) زوجي : فالوسيط يقع بين القيمة رقم 2/n و القيمة ((2/n)+1) ومن ثم فإن :

رتبة الوسيط = $\frac{\left(1+\left(\frac{}{2}\right)\right)+\left(\frac{}{2}\right)}{2}$

قيمة الوسيط = $\frac{\left(1+\left(\frac{}{2}\right)\right)\text{القيمة رقم}\left(\frac{}{2}\right)+\text{القيمة رقم}}{2}$

مثال : أحسب الوسيط لهذه المفردات 3 ، 5 ، 9 ، 11 ، 13 ، 16 (مرتبة تصاعديا)

رتبة الوسيط = $\frac{\left(1+\left(\frac{6}{2}\right)\right)+\left(\frac{6}{2}\right)}{2}$ = 3.5

قيمة الوسيط = $\frac{\text{القيمة رقم}(3)+\text{القيمة رقم}(4)}{2}$ = $\frac{11+9}{2}$ = **10**

2.5. للبيانات المبوبة (من جدول توزيع تكراري) :

يتم اتباع الخطوات التالية :

- تكوين الجدول التكراري المتجمع الصاعد

- تحديد رتبة الوسيط = $\frac{\Sigma f_i}{2} = \frac{}{2}$

- تحديد فئة الوسيط كما في الشكل التالي:

التكرار المتجمع الصاعد السابق f_1	الحد الأدنى لفئة الوسيط A
رتبة الوسيط n/2 ⟵	الوسيط Med
التكرار المتجمع الصاعد اللاحق f_2	الحد الأعلى لفئة الوسيط

و يحسب الوسيط بتطبيق المعادلة التالية:

$$Med = A + \left(\frac{\frac{N}{2} - f_1}{f_2 - f_1} \times L\right)$$

حيث L هي طول فئة الوسيط و تحسب بالعلاقة التالية:

طول الفئة = الحد الأعلى للفئة – الحد الأدنى لنفس الفئة

<u>مثال :</u> فيما يلي بيان لأعمار ثلاثين تلميذ و عدد ممارستهم للرياضة خلال أسبوع و المطلوب حساب الوسيط حسابيا و بيانيا.

الفئات	11.5 – 14.5	14.5 – 17.5	17.5 – 20.5	20.5 – 23.5	23.5 – 26.5	26.5 – 29.5	**المجموع**
التكرار	8	4	7	6	2	3	**30**

<u>الحل</u>

أولا. حسابيا:

- رتبة الوسيط : n/2 = 2/30 = 15
- تكوين الجدول المتجمع الصاعد:

الفئات	التكرار (f_i)	التكرار المتجمع الصاعد
11.5 – 14.5	8	8
14.5 – 17.5	4	12
17.5 – 20.5	7	19
20.5 – 23.5	6	25
23.5 – 26.5	2	27
26.5 – 29.5	3	30
المجموع	**30**	

- تحديد فئة الوسيط: وهي الفئة التي تشمل قيمة الوسيط ، و يمكن معرفتها بتحديد التكرارين المتجمعين الصاعدين الذين يقع بينهما رتبة الوسيط (15) ، و في الجدول أعلاه نجد أن رتبة الوسيط تقع بين للتكرارين المتجمعين 12 و 19 ، و يكون الحد الأدنى لفئة الوسيط هو المناظر للتكرار المتجمع الصاعد السابق 12، و الحد الأعلى لفئة الوسيط هو المناظر للتكرار المتجمع الصاعد اللاحق 19، أي أن فئة الوسيط هي (20.5−17.5)

و بتطبيق معادلة الوسيط السابقة نجد أن :

A= 17.5 f_1=12 f_2=19 L=3

$$Med = A + \left(\frac{\frac{N}{2} - f_1}{f_2 - f_1} \times L\right)$$

$$Med = 17.5 + \left(\frac{15 - 12}{19 - 12} \times 3\right) = 18.78$$

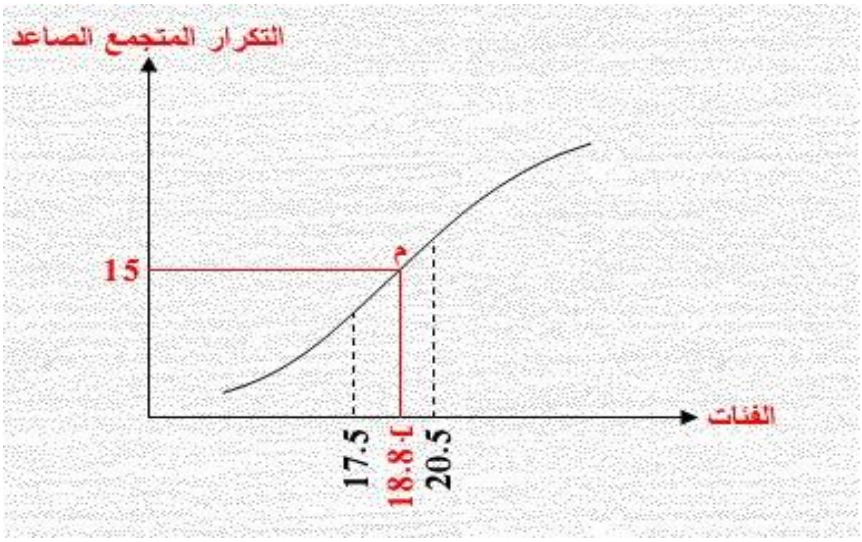
ثانيا. بيانيا:

كما بإمكاننا أن نحسب الوسيط من خلال الرسم البياني :

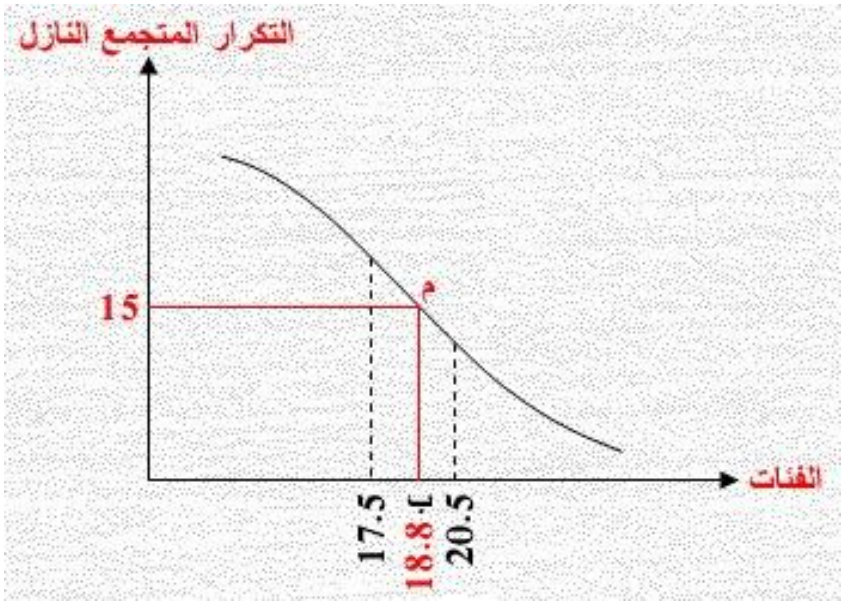
يمكن ذلك برسم محورين أفقي للفئات ورأسي للتكرار المتجمع ثم نعين النقاط في المستوى (14.5، 8)،(17.5، 12)،... ،(29.5 ، 30) نصل بين النقاط باليد نحصل على المنحنى التكراري المتجمع الصاعد كما مبين بالشكل.

من العدد 15 (ترتيب الوسيط) نرسم مستقيماً أفقياً يقطع المنحى في نقطة م مثلاً.

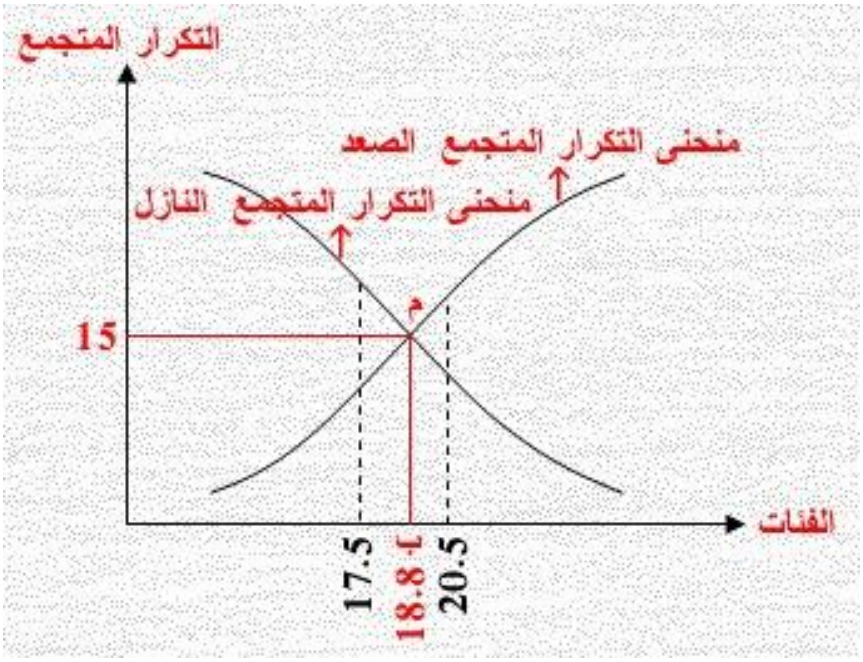
من نقطة م نسقط عموداً على المحور الأفقي يلاقيه في نقطة ولتكن ب العدد عند النقطة ب هو قيمة الوسيط والقيمة هنا هي 18.8



يمكن تكرار ذلك مع المنحنى التكراري المتجمع النازل ونصل لنفس الناتج لاحظ الشكل.



أو من خلال رسم المنحنيين الصاعد والنازل ومن نقطة تقاطعهم (م) نسقط العمود على المحور الأفقي (محور الفئات) لنحصل على قيمة الوسيط 18.8 (لاحظ الشكل).



3.5. خصائص الوسيط

- عدم تأثره بالقيم الشاذة
- تعتمد قيمته على التكرارات.
- يوجد في كل البيانات التي تتصف بالترتيب
- يهتم بالقيم الوسطى للبيانات فقط
- لا يصلح لإعطاء فكرة عن النزعة المركزية للبيانات في فئات متباعدة
- لا يسهل التعامل معه في التحاليل الإحصائية في دراسة ظاهرة ما في عدة عينات.

4.5. مدى تأثر الوسيط بالعمليات الحسابية

لتكن لدينا القيم التالية : 5 ، 8 ، 12 ، 9 ، 10 ، 15 ، 17 و المطلوب حساب الوسيط

نرتب القيم تصاعدياً تنازلياً (تنازلياً مثلاً) أي 5 ، 8 ، 9 ،10، 12 ، 15 ، 17

فالوسيط هو 10

لنقم بالعمليات الحسابية الأربع على البيانات الأصلية بعد ترتيبه تنازلياً كما مبين بالجدول الآتي:

ملاحظات	الوسيط	البيانات	العملية الحسابية
الوسيط الجديد = الوسيط السابق+3	13	8, 11, 12, 13, 15, 18, 20	الجمع: لنضف العدد 3 للمفردات (بعد الترتيب)
الوسيط الجديد = الوسيط السابق −4	6	1, 4, 5, 6, 8, 11, 13	الطرح: لننقص العدد 4 من كل مفردة
الوسيط الجديد = الوسيط السابق x 3	30	15, 24, 27, 30, 36, 45, 51	الضرب: نضرب كل مفردة في3
الوسيط الجديد = الوسيط السابق ÷ 2	5	2.5, 4, 4.5, 5, 6, 7.5, 8.5	القسمة : نقسم كل مفردة على العدد 2

بناء على ما ورد في الجدول أعلاه يتأثر الوسيط بالعمليات الحسابية الأربع وإن أجريت أكثر من عملية فمثلاً إذا أضفنا العدد 2 على كل مفردة ثم ضربنا النواتج في 3 فالوسيط الجديد للبيانات بعد الإضافة و الضرب (21, 30, 33, 36, 42, 51, 57) يكون مضافاً للوسيط السابق 2 ومضروباً في 3 أي:

$$(10 + 2) \times 3 = 36$$

6/ أشباه الوسيط

المشاهدات التي نقوم برصدها ضمن مفهوم التكرار في جدول التوزيع التكراري والتي سبق أن حسبنا من خلالها الوسيط والذي يقع ترتيبه في المنتصف (التكرار) أو التي

تكون القيمة التي يقل عنها 50%من المشاهدات ويزيد عنها 50%وبنفس الأسلوب يكون:

1.6. الربيعيات Quartiles:

هي القيم التي تقسم البيانات إلى أربع أجزاء متساوية و منها:

الربيع الأدنى هو القيمة التي يقل عنها 25%من المشاهدات أو يزيد عنها 75%من المشاهدات ويكون ترتيبه مجموع التكرار مقسوم على أربعة ويحسب كالوسيط.

الربيع الأعلى هو القيمة التي يقل عنها 75%من المشاهدات أو يزيد عنها 25%من المشاهدات ويكون ترتيبه مجموع التكرار مضروب في 75%ويحسب كالوسيط.

يرمز للربيع الأول أو الأدنى بالرمز Q_1 و يرمز للربيع الثالث أو الأعلى بالرمز Q_3 و الوسيط بالرمز Q_2 وهو ليس متوسط الربيعين أو لا يساوي ½($Q_3 + Q_1$).

2.6. العشيرات (Deciles):

مفردها العشير وهو ناتج التقسيم العشري ، 0.1هو العشير الأدنى (D_1) من القيم أي ما قبله عشر القيم وبعده تسعة أعشار القيم (0.9) وتحسب قيمته كحساب الوسيط.

3.6. المئينات (Percentile):

المئين الأول هو التقسيم 1من 100ورمزه C_1 أي C_1 = 0.01 من قيم المعطيات والمئين الستين C_{60} يقع عند 0.60وتحسب قيمته كحساب الوسيط.

<u>مثال :</u> فيما يلي الجدول التكراري لأعمار ثلاثين تلميذ و عدد ممارستهم للرياضة خلال أسبوع و المطلوب حساب المطلوب حساب كل من Q_1 , Q_3 , D_3 , C_{40}

الفئات	11.5 – 14.5	14.5 – 17.5	17.5 – 20.5	20.5 – 23.5	23.5 – 26.5	26.5 – 29.5	المجموع
التكرار	8	4	7	6	2	3	30

المطلوب حساب كل من Q_1 , Q_3 , D_3 , C_{40} حسابياً وبيانياً.

الحل:

أولا: حسابياً:

نكون الجدول التكراري المتجمع الصاعد وهو الأساس في حساب المطلوب

الفئات	التكرار (f_i)	التكرار المتجمع الصاعد
11.5 – 14.5	8	8
14.5 – 17.5	4	12
17.5 – 20.5	7	19
20.5 – 23.5	6	25
23.5 – 26.5	2	27
26.5 – 29.5	3	30
المجموع	**30**	

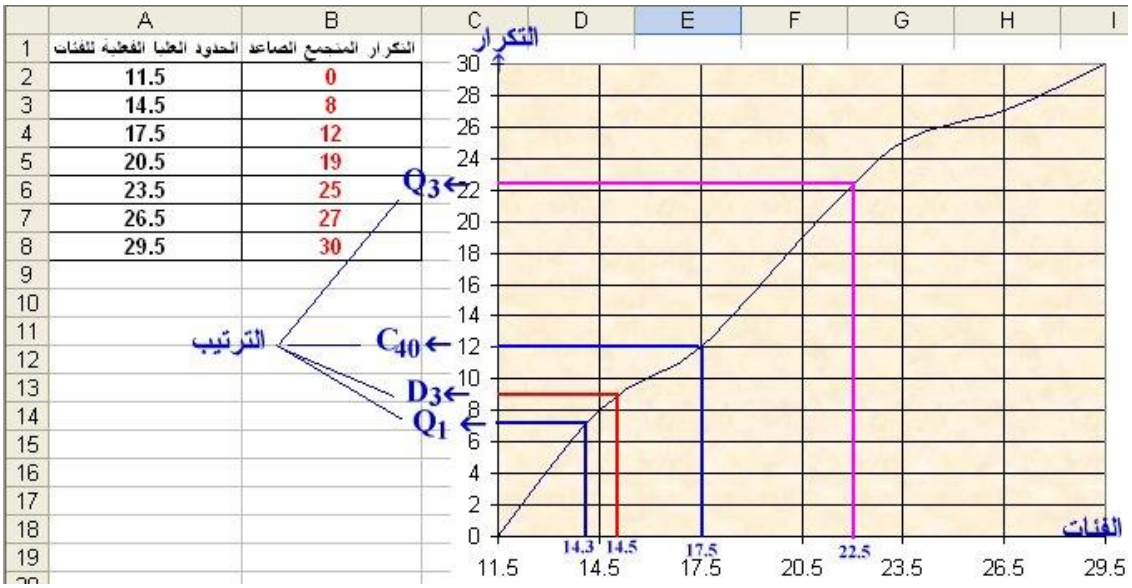
سنحدد ترتيب كل منهم ثم نعين القيمة بنفس الطريقة التي سبق ذكرها في حساب الوسيط أي باستخدام نفس القانون :

$$Med = A + \left(\frac{\text{الترتيب} - f_1}{f_2 - f_1} \times L\right)$$

Q_1	**الترتيب =** **30 ÷ 4 = 7.5**	طول الفئة L=3، الفئة (11.5–14.5) و بالتالي A=11.5 f1=0 ، f2=8 **الربيع الأول Q_1** وقيمته = 11.5 +(((0–8)/(0–7.5)).3)= **14.31** **الربيع الأول Q_1** وقيمته = 11.5 + 2.81 = **14.31**
Q_3	**الترتيب =** **(30 ÷ 4)3 =** **22.5**	طول الفئة L=3، الفئة (20.5–23.5) و بالتالي A=20.5 f1=19 ، f2=25 **الربيع الثالث Q_3** وقيمته = 20.5 +(((19–22.5)/(19–25)).3)= **22.25** الربيع الثالث Q_3 وقيمته = 20.5 + 1.75

		= 22.25
D_3	**الترتيب =** **9 = (10 ÷ 30)3**	طول الفئة L=3، الفئة (14.5−17.5) و بالتالي A=14.5 f1=8 ، f2=12 **العشير الثالث D_3** وقيمته = 14.5 +(((8−9)/(8−12)).3)= **15.25** **العشير الثالث D_3** وقيمته = 14.5 + 0.75 = **15.25**
C_{40}	**الترتيب =** **= (100 ÷ 30)40** **12**	طول الفئة L=3، الفئة (14.5−17.5) و بالتالي A=14.5 f1=8 ، f2=12 **المئين الأربعين C_{40}** وقيمته = 14.5 +(((8−12)/(12−8)).3)= **17.5** **المئين الأربعين C_{40}** وقيمته = 14.5 + 3 = **17.5** أو من خلال الجدول التكراري المتجمع الصاعد نلاحظ أن القيمة 12 موجودة وتقابلها 17.5 أي: C_{40} = 17.5

ثانياً: بيانياً:



4.6. المدى الربيعي المتوسط :

هو متوسط الربيعين الأول والثالث أي يساوي نصف مجموعهم وهو يختلف عن Q_2 الذي ترتيبه نصف التكرار ويندر استخدامه وخاصة للعينات التي حجمها أكبر من 5.

المدى الربيعي المتوسط = $\frac{Q_1+Q_3}{2}$ = ½ (22.5 + 14.31)

= 18.4 " سبق أن حسبنا قيمة الوسيط وكانت 18.8"

7/ المنوال Mode

يعرف المنوال لمجموعة من القيم بأنه القيمة الأكثر شيوعاً أو تكرارا بينها، وقد يوجد أكثر من منوال لمجموعة القيم، و يكثر استخدامه في حالة البيانات الوصفية لمعرفة النمط الشائع و يمكن حسابه للبيانات المبوبة و غير المبوبة كما يلي:

1.7.للبيانات غير المبوبة:

المنوال (Mod) = القيمة الأكثر تكرارا

مثال : لدينا مجموعة القيم: 12 ، 13، 8 ، 9 ، 12 ، 12، 23 لها منوال واحد هو القيمة 12

و مجموعة القيم : 12 ، 13، 8 ، 8 ، 12 ، 12، 23 يوجد أكثر من منوال 12, 8

و مجموعة القيم : 12 ، 13، 8 ، 7 ، 11 ، 15، 23 لا يوجد لها منوال .

- يمكن استخدام المنوال للقيم الكمية والنوعية ولكن يكثر استخدامه مع المعطيات النوعية.
- في حال وجود أكثر من منوال (متجاورة) فمتوسط هذه المنوالات يعتبر منوال التوزيع حال وجود تلك القيمة في التوزيع.
- يسمى التوزيع أُحادي المنوال إذا وجد منوال واحد، وثنائي المنوال إذا وجد منوالان، وهكذا.

- إذا وجد أكثر من منوال فالأكثر تكرار بينهم يدعى المنوال الرئيسي (Major Mode) والأخرى تعرف بالمنوال الفرعي أو الثانوي (Minor Mode)
- المنوال ببساطة يمثل بقيمة مسقط أعلى نقطة في المنحنى التكراري على محور الفئات.
- المنوال لا يتأثر بالقيم المتطرفة وهو يمثل غالبية القيم.
- تقل قيمته كمتوسط للقيم حال تباعد القيم عن بعضها البعض.

2.7. للبيانات المبوبة:

توجد أكثر من طريقة لحساب المنوال مثل الطريقة الحسابية ، الطريقة البيانية ، الحاسب الآلي.

أولا : الطريقة الحسابية : (تعرف بطريقة الفروق لبيرسون)

نستخدم فيها القانون التالي :

$$Mod = A + \left(\frac{d1}{d1 + d2} \times L \right)$$

حيث أن :

A : الحد الادنى لفئة المنوال

d1 : الفرق الأول = تكرار فئة المنوال – التكرار السابق لها

d2 : الفرق الثاني = تكرار فئة المنوال – التكرار الاحق لها

L : طول فئة المنوال

فئة المنوال = الفئة المناظرة لأكبر تكرار

مثال : فيما يلي جدول التوزيع التكراري يبين معدل ذكاء 50 طالب المطلوب حساب المنوال.

الفئات	15 – 12	18 – 15	21 – 18	24 – 21	27 – 24	30 – 27
التكرار	11	10	15	7	5	2

الحل:

يمكننا أن نلاحظ من الجدول أعلاه أن الفئة المنوالية هي الفئة (18−21) و التي يقابلها أكبر تكرار و بالتالي: الحد الأدنى لفئة المنوال A= 18

تكرار فئة المنوال هو القيمة 15

طول فئة المنوال L=3

التكرار السابق لفئة المنوال = 10

التكرار اللاحق لفئة المنوال = 7

ومنه d1 = 15− 10 = 5 و d2 = 15− 7 = 8

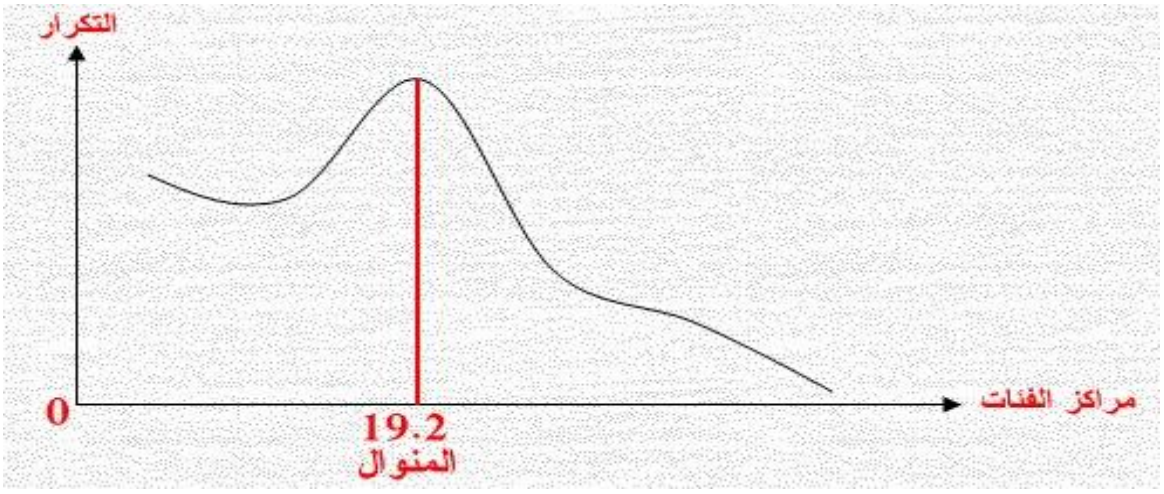
و بتطبيق قانون المنوال نجد ما يلي

$$Mod = A + \left(\frac{d1}{d1 + d2} \times L \right) = 18 + \left(\frac{5}{5 + 8} \times 3 \right)$$

$$Mod = 19.15$$

ثانيا: الطريقة البيانية:

برسم المنحنى التكراري باتخاذ مراكز الفئات كممثلة للتكرار أي نعين النقاط (13.5، 11) ، (16.5، 10) ، ... ، (28.5، 2) في المستوى ثم نصل بينها باليد فنحصل على المنحني التكراري كما مبين بالشكل.



من أعلى نقطة في المنحنى نسقط عموداً على المحور الأفقي ونقطة تقاطعه مع المحور تمثل قيمة المنوال كما مبين بالشكل.

أو من الجدول التكراري كما مبين بالشكل حيث م نقطة تقاطع المستقيان الواصلان من بداية الفئة المنوالية لبداية الفئة اللاحقة ، من نهاية الفئة المنوالية لنهاية الفئة السابقة، ومسقط م على المحور الأفقي يعطي قيمة المنوال.

أما في الحالة التي لا تتساوى فيها أطوال الفئات فيلزم هنا قسمة كل تكرار للفئة على طولها فنحصل على ما يعرف بالتكرار المعدل.

3.7. العلاقة بين الوسط والوسيط والمنوال:

يمكن استخدام الوسط الحسابي و الوسيط و المنوال في وصف المنحنى التكراري و لذي يعبر عن شكل توزيع البيانات كما يلي:

يكون المنحنى متماثل إذا كان : الوسط = الوسيط = المنوال – الشكل 3.

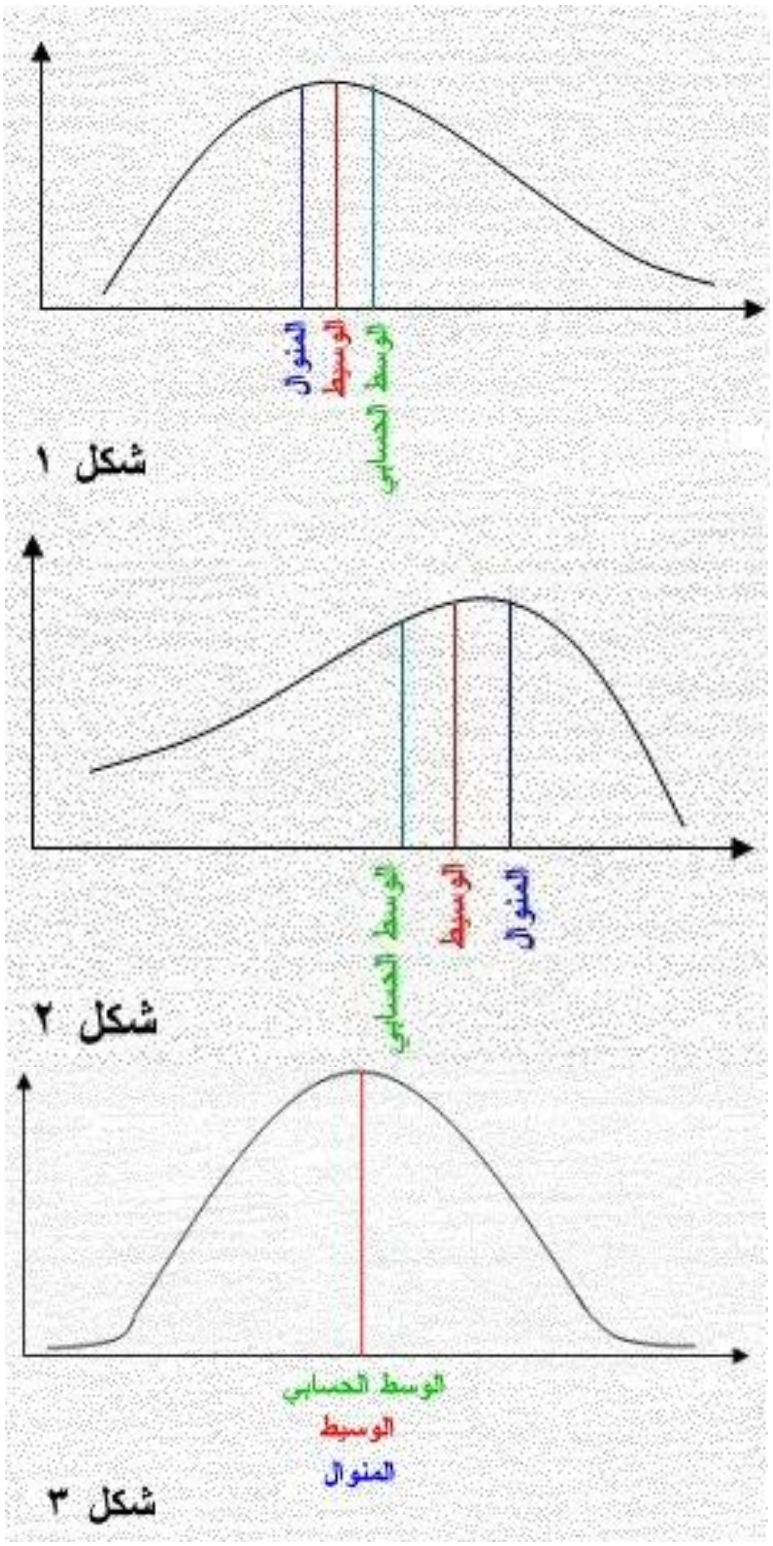
يكون المنحنى موجب الالتواء (ملتوي جهة اليمين)إذا كان:

الوسط الحسابي> الوسيط > المنوال – الشكل1

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

الوسط الحسابي< الوسيط < المنوال – الشكل 2

منحنى التوزيع يكون الالتواء فيه معتدلا إذا كان: الوسط الحسابي – المنوال = 3 (الوسط الحسابي – الوسيط).



4.7. خصائص المنوال

يعتبر المنوال أفضل مقياس من الوسط والوسيط للقياسات النوعية الصرفة مثل متغير النوع والجنسية وما شابه لتمثيله الظاهرة محل الدراسة.

- قيمته تمثل الأكثر تجمع للظاهرة.
- سهولة حسابه كما بينا ذلك.
- توجد عدة طرقه لحسابه وكلها تقريبية مع اختلاف القيمة فقط المدرج التكراري والفروق الطريقتان اللتان تتفق في الإجابة.
- يستخدم المنوال لحساب مقياس سريع عند عدم الاهتمام بالدقة للمقياس.

الفصل الرابع

مقاييس التشتت

إن مقاييس النزعة المركزية الوسط والوسيط والمنوال قيم تتمركز حولها القيم المشاهدة وقد بينا بيانياً بأن هذه القيم تمثل بنقط على محور الفئات وكان من أهمها الوسط الحسابي حيث بينا أن مجموع انحرافات المشاهدات عنه يساوي صفر ، فعند مقارنة مجموعتين من البيانات يمكن استخدام شكل التوزيع التكراري أو المنحنى التكراري و كذلك بعض مقاييس النزعة المركزية مثل الوسط و الوسيط و المنوال و الاحصاءات الترتيبية ، و لكن استخدام هذه الطرق وحدها لا يكفي عند المقارنة، فقد يكون مقياس النزعة المركزية للمجموعتين متساوي و ربما يوجد اختلاف كبير بين المجموعتين من حيث مدى تقارب و تباعد البيانات من بعضها البعض ، أو مدى تقارب و تباعد القيم عن مقياس النزعة المركزية.

و مثال ذلك إذا كان لدينا مجموعتين من الطلاب و كانت نتائجهم كالتالي :

المجموعة الأولى : 42 ، 35 ، 45 ، 56 ، 32 متوسطها الحسابي 42

المجموعة الثانية : 30 ، 40 ، 60 ، 20 ، 60 متوسطها الحسابي 42

و مع ذلك درجات المجموعة الأولى أكثر تجانسا من المجموعة الثانية رغم تساوي وسطهم الحسابي.

لذلك فمقاييس النزعة المركزية غير كافية لوصف البيانات من حيث تفاوت البيانات عن وسطها (تشتتها) فالحاجة استدعت مقاييس أخرى تعرف بمقاييس التشتت أو مقاييس الانتشار كالمدى ونصف المدى الربيعي والانحراف المتوسط والتباين والانحراف المعياري ومعامل الاختلاف وأخرى.

و الهدف من دراسة الإحصاء هو وجود هذا التشتت بين القيم فلو اتفق بأن أوزان الإنتاج لسلعة واحدة متساوية في المواصفات أو كان كل الناس لهم نفس الطول أو لهم نفس فصيلة الدم أو ... لما دعت الحاجة لوجود هذا العلم فالتفاوت بين المخلوقات في خصائصها أو التفاوت في التحصيل الدراسي وما إلى ذلك دعتنا لإيجاد مقاييس التشتت.

1/ المدى: Rang

هو أبسط مقاييس التشتت و يعرف بأنه الفرق بين أكبر قراءة وأصغر قراءة في مجموعة القراءات و نميز بين حالتين هما :

1.1. للبيانات غير المبوبة :

بتطبيق المعادلة التالية : المدى = الفرق بين أكبر قيمة وأصغر قيمة.

مثال : أحسب المدى لمجموعة القيم 23 ، 34 ، 33 ، 12 ، 40

نلاحظ أن اكبر قيمة هي 40 و أصغر قيمة هي 12 و بالتالي فالمدى يساوي

$40 - 12 = 28$

من الواضح أن المدى يهتم فقط بتلك القيمتين ولا يتأثر بالقيم الأخرى ويعتبر من أبسط مقاييس التشتت ولا يعتبر مقياس مهم للتشتت، وكلما صغرت قيمة المدى قل تشتت المجموعة فمجموعة القيم 17 ، 24 ، 13 ، 33 ، 22 المدى المطلق لها

$33 - 13 = 20$ أقل تشتت من المجموعة السابقة التي مداها 28 ويصح القول بأن المجموعة الأولى (المدى 28) أكثر تشتت من المجموعة الثانية (المدى 20)

2.1. للبيانات المبوبة :

يعرف المدى بأكثر من طريقة و له أكثر من صيغة :

- المدى يساوي الفرق بين قيمتي الحدين الأعلى للفئة العليا والحد الأدنى للفئة الدنيا

المدى = قيمة الحد الأعلى للفئة العليا – قيمة الحد الأدنى للفئة الدنيا

- المدى يساوي الفرق بين قيمتي مركزي الفئتين العليا والدنيا أي:

المدى = قيمة مركز الفئة العليا – قيمة مركز الفئة الدنيا

- المدى يساوي الفرق بين قيمتي الحدين الفعليين للفئتين العليا والدنيا أي:

المدى = قيمة الحد الفعلي للفئة العليا – قيمة الحد الفعلي للفئة الدنيا

لكن غالبا ما نستخدم الطريقة الأولى أو الثانية في حساب المدى.

مثال : الجدول التالي يبين أعمار 30 طالب في والمطلوب حساب المدى لهذه الأعمار بمختلف الصيغ.

الفئات	24 – 28	29 – 33	34 – 38	39 – 43	44 – 48	48 – 52
التكرار	3	4	7	6	8	2

الحل :

الحد الأعلى للفئة العليا = 52

الحد الأدنى للفئة الدنيا = 24

المدى = 52 – 24 = 28

مركز الفئة العليا (الأخيرة) (48 + 52) ÷ 2 = 50

مركز الفئة السفلى (الأولى) (24 + 28) ÷ 2 = 26

المدى = 50 – 26 = 24

الحد الأعلى الفعلي للفئة العليا = 52.5

الحد الأدنى الفعلي للفئة الدنيا = 23.5

المدى = 52.5 – 23.5 = 29

3.1. مميزات المدى :

- سهولة حساب المدى.
- يعطى فكرة سريعة عن طبيعة البيانات ويستخدم كثيرا في مراقبة جودة الإنتاج و أسعار الأوراق المالية في البورصة و كذلك في وصف طبيعة الأحوال الجوية.
- ذو علاقة مع الانحراف المعياري (سيدرس لاحقاً) كتأكيد على صحة الانحراف المعياري حيث أن الانحراف المعياري لا يزيد أو لا يقل عن سبعة أمثال المدى فإن تحقق ذلك يعني صحة القيمة المحسوبة وإلا احتمال الخطأ في القيمة المحسوبة للانحراف المعياري.

4.1 . عيوب المدى:

- يعتمد فى حسابه على قيمتين فقط من البيانات مع إهمال باقي القيم.
- يتأثر كثيرا بالقيم الشاذة (المتطرفة) لذلك فهو مقياس تقريبي لا يعتمد عليه.
- لا يمكن حسابه من توزيع تكراري مفتوح.
- لا يتفق مع المفهوم القائل بأن زيادة مفردات العينة لظاهرة ما يؤدي لزيادة التجانس لتوزيع الظاهرة، فالقيم 45 ، 47 ، 49، 44، 43 ، 70 فالمدى 27 وبحذف القيمة الأخيرة يكون المدى 6 والقيم هنا متجانسة وغير متباينة ولكن وجود القيمة 70 عكس الصورة بعدم التجانس بين أفراد المجموعة.
- قد يعطي نتيجة خاطئة للمقارنة بين مجموعتين مختلفتان في الحجم.

كما يمكن استخدام المدى كمؤشر إضافي وهذا في حال تساوى المتوسطات. فإذا تقدم مجموعة من الأفراد للامتحان الذي يتضمن الإجابة على أربعة اختبارات كل منها من 100درجة وكانت النتائج للأفراد الأربع مبينة بالجدول التالي حيث في حالة تساوي المتوسطات نأخذ أقلهم مدى لقلة التشتت:

الإسم	D1	D2	D3	D4	المتوسط	المدى
محمد	82	90	72	68	78	8
علي	88	80	84	60	78	28
خالد	78	84	70	80	78	14
منصور	80	86	66	80	78	20

فيكون النجاح من نصيب محمد لكون المدى بين درجاته يساوي 8 وهو الأقل عن مدى الآخرين.

2/ الانحراف الربيعي: (Quartile deviation)

من أهم عيوب المدى أنه يتأثر بالقيم الشاذة و بالتالي فهو لا يعطى صورة صادقة عن

طبيعة البيانات .لذلك دعت الحاجة إلى إيجاد مقياس آخر يتم من خلاله التخلص من تأثير القيم الشاذة وهو ما يسمى بنصف المدى الربيعي.أو الانحراف الربيعي ويعرف كما يلي:

هو نصف الفرق بين المدى الربيعي الثالث والمدى الربيعي الأولويرمز له بالرمز Q أي أن: $Q = \frac{1}{2}(Q_1 - Q_3)$ حيث أن Q_1 الربيع الأول, Q_3 الربيع الثالث.

يفضل للباحث استخدامه حال استخدام الوسيط ويسمى الانحراف الربيعي أيضاً بنصف المدى الربيعي للقانون أعلاه، ويسمى كذلك الربيع الثاني أسوة بالربيع الأول والثالث وهو أفضل من المدى لأنه لا يتأثر بالقيم المتطرفة مستبعد القيم المتطرفة من الأعلى والأسفل.

1.2. للبيانات غير مبوبة:

يحسب من خلال الخطوات التالية :

- ترتب البيانات تصاعدياً
- نحدد رتبتي الربيعين الأول $(n \div 4)$ والثالث $[3(n \div 4)]$ حال n تقبل القسمة على 4 وإلا نأخذ متوسط القيمتين اللتين تقع بينهم قيمة $n \div 4$ ، $3(n \div 4)$ والمثال التالي يبين ذلك المعنى:

مثال:

أوجد الانحراف الربيعي لكل من درجات المجموعتين الآتيتين.

المجموعة الأولى: 22 ، 24 ، 36، 21 ، 25 ، 30 ، 20 ، 28

المجموعة الثاني: 21 ، 20 ، 25 ، 17 ، 19 ، 15، 22 ، 18، 20

الحل:

نرتب المجموعة الأولى تصاعدياً: 20 ، 21 ، 22 ، 24 ، 25 ، 28 ، 30 ، 36

بما أن ترتيب الربيع الأول = $8 \div 4 = 2$ فإن قيمة الربيع الأول هي القيمة الثاني في الترتيب وهي 21 أي الربيع الأول = 21

بما أن ترتيب الربيع الثالث = 3(8 ÷ 4) = 3 × 2 = 6 فإن قيمة الربيع الثالث هي القيمة السادسة في الترتيب وهي 28 أي الربيع الثالث = 28

الإنحراف الربيعي = (28 – 21) ÷ 2 = 7 ÷ 2 = **3.5**

نرتب المجموعة الثانية تصاعدياً: 15 ،17 ،18 ، 19 ، 20 ، 20 ، 21 ، 22 ، 25

بما أن ترتيب الربيع الأول = 9 ÷ 4 = 2.25 فإن قيمة الربيع الأول هي متوسط القيمتين الثانية والثالثة = (17 + 18) ÷ 2 = 35 ÷ 2 = 17.5

بما أن ترتيب الربيع الثالث = 3(9÷ 4) =3× 2.25= 6.75 فإن قيمة الربيع الثالث هي متوسط القيمتين السادسة والسابعة = (20 + 21) ÷ 2= 41 ÷ 2= 20.5

نصف المدى الربيعي = (20.5 – 17.5) ÷ 2 = 3 ÷ 2 = **1.5**

2.2. للبيانات المبوبة:

يحسب نصف المدى الربيعي لهذه البيانات بنفس الطريقة التي سبق شرحها لحساب الوسيط وهو طريقة الفروق ويحسب الربيع الأول 1Q بوضع n/4 بدلا من n/2 في قانون الوسيط ويحسب الربيع الثالث 3Q بوضع3n/4 بدلا من n/2 في قانون الوسيط وبعد ذلك نحسب نصف المدى الربيعي بعد حساب 1Q و 3Q من العلاقتين التاليتين:

$$Q1 = A + \left(\frac{\frac{N}{4} - f_1}{f_2 - f_1} \times L\right)$$

$$Q3 = A + \left(\frac{\frac{3N}{4} - f_1}{f_2 - f_1} \times L\right)$$

مثال: الجدول التكراري الآتي يبين بيانات أعمار 30 شخص حسب عدد ممارستهم للرياضة والمطلوب حساب نصف المدى الربيعي بكل الطرق الممكنة.

الفئات	12 – 14	15 – 17	18 – 20	21 – 23	24 – 26	27 – 29
التكرار	8	4	7	6	2	3

الحل:

بنفس الطريقة في حساب الوسيط نحسب قيمة كل من الربيع الأول والثالث من التكرار المتجمع الصاعد :

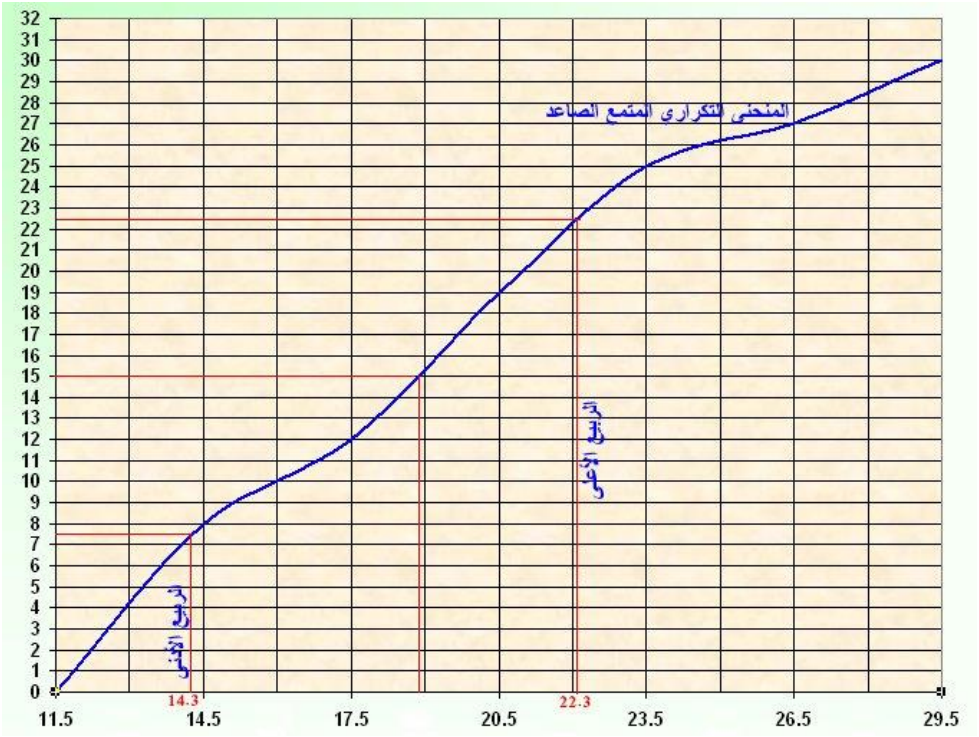
نكون الجدول الآتي:

الحدود الفعلية للفئات	التكرار (fi)	التكرار التراكمي الصاعد
11.5 – 14.5	8	8
14.5 – 17.5	4	12
17.5 – 20.5	7	19
20.5 – 23.5	6	25
23.5 – 26.5	2	27
26.5 – 29.5	3	30

Q_1	**الترتيب =** **7.5 = 4 ÷ 30**	طول الفئة L=3، الفئة (11.5–14.5) و بالتالي A=11.5 f1=0 ، f2=8 **الربيع الأول Q_1 وقيمته =** **14.31** =(3.((0−8)/(0−7.5)))+ 11.5 **الربيع الأول Q_1 وقيمته = 11.5 + 2.81 = 14.31**
Q_3	**الترتيب =** **22.5 = (4 ÷ 30)3**	طول الفئة L=3، الفئة (20.5–23.5) و بالتالي A=20.5 f1=19 ، f2=25 **الربيع الثالث Q_3 وقيمته =** **22.25** =(3.((19 −25)/(19−22.5)))+ 20.5 الربيع الثالث Q_3 وقيمته = 22.25 = 1.75 + 20.5

نصف المدى الربيعي = (22.25 – 14.3) ÷ 2 = 3.97

الطريقة البيانية:



3.2. مميزات نصف المدى الربيعي:

- يفضل استخدامه كمقياس للتشتت في حالة وجود قيم شاذة نحو الكبر أو الصغر.
- يمكن حسابه من التوزيعات التكرارية المفتوحة من الطرفين.

4.2. عيوب نصف المدى الربيعي:

- لا يأخذ جميع القيم في الاعتبار.
- لا يسهل التعامل معه في التحليل الإحصائي

3/ الانحراف المتوسط: Mean Deviation Or The Average Deviation

الانحراف المتوسط أو متوسط الانحرافات هو ناتج مجموع القيمة المطلقة (الناتج الموجب) لانحرافات القيم عن وسطها الحسابي مقسوما على عددها، ويرمز له بالرمز MDو نميز فيه بين حالتين هما:

1.3. للبيانات غير المبوبة :

نستخدم القانون التالي :

$$MD = \frac{\sum|x_i - \bar{x}|}{N}$$

تنبيه: القيمة المطلقة للعدد X تكتب | X | وتقرأ مقياس X أو مطلق X

وإن |5| = 5 ، | – 5 | = 5

مثال: أوجد الانحراف المتوسط للقيم 14 ، 16 ، 10 ، 8 ، 2

الحل:

مجموع قيم المشاهدات = 14 + 16 + 10 + 8 + 2 = 50

الوسط الحسابي = 50 ÷ 5 = 10

الانحرافات عن الوسط الحسابي هي |14 –10|،|16 – 10| ، |10 – 10| ، |8 – 10| ، |2 – 10| أي 4 ، 6 ، 0 ، 2 ، 8

مجموع الانحرافات = 4+ 6 + 0 + 2 + 8 = 20

الانحراف المتوسط = 20 ÷ 5 = **4**

2.3. للبيانات المبوبة:

نحسب الوسط الحسابي بعد ضرب مركز كل فئة في تكرارها ثم حساب الانحرافات عن الوسط و هذا باستخدام القانون التالي :

$$MD = \frac{\sum f_i|x_i - \bar{x}|}{N}$$

مثال:

احسب الانحراف المتوسط من جدول التوزيع التكراري الآتي لعلامات 40 طالب في مادة الإحصاء

الحل: نتبع الخطوات التالية:

- نكون جدول تكراري يضم مراكز الفئات x_i وآخر يشمل $x_i.f_i$ لحساب الوسط الحسابي
- حساب القيم المطلقة للفروق بين مراكز الفئات و وسطها الحسابي مضروبا في تكرار كل فئة

أي أننا نكون جدول لحساب مكونات معادلة الإنحراف المتوسط.

الحل:

الفئات	التكرار f_i	مركز الفئة x_i	$f_i\ x_i$	$\lvert x_i - \bar{x} \rvert$	$f_i \lvert x_i - \bar{x} \rvert$
5−2	1	3.5	3.5	| 3.5 − 10.7 | = 7.2	7.2
8−5	8	6.5	52	| 6.5 −10.7 | = 4.2	33.6
11−8	13	9.5	123.5	| 9.5 −10.7 | = 1.2	15.6
14−11	10	12.5	125	| 12.5 −10.7 | = 1.8	18
17−14	8	15.5	124	| 15.5 −10.7 | = 4.8	38.4
المجموع	40		428		112.8

1−حساب الوسط الحسابي

$$X = \frac{\sum xi.fi}{\sum fi} = \frac{428}{40} = \mathbf{10.7}$$

2−الإنحراف المتوسط = 40/112.8 = **2.82**

3.3. مزايا وعيوب الانحراف المتوسط

الانحراف المتوسط أكثر دقة من المدى والانحراف الربيعي لشموله كل القيم ولكنه محدود الاستخدام لتأثره بالقيم الشاذة وتجاهله الإشارة السالبة، كما يمكن حساب الفرق عن طريق الوسيط بدلاً من الوسط الحسابي حيث يكون المجموع أصغر ما يمكن إلا أن الحساب عن طريق الوسط الحسابي هو الأكثر شيوعاً، ومع ذلك تظل أهمية هذا المقياس محدودة.

4/ التباين والانحراف المعياري:

سبق أن ذكرنا بأن مجموع انحرافات القيم عن وسطها الحسابي يساوي صفر وقولنا المجموع يساوي الصفر يعني وجود فروق سالبة وأخرى موجبة وللتخلص من الفروق السالبة قمنا بأخذ الانحراف المطلق أي بضرب الفرق السالب بسالب 1 وعرفنا ذلك بالانحراف المتوسط وتوجد طريقة أخرى للتخلص من الفروق السالبة هذه وذلك بتربيعها لتصبح موجبة ومجموع مربعات الانحرافات للقيم عن وسطها الحسابي يعرف بالتباين (Variance) في حين

الجذر التربيعي لهذا المجموع (مجموع مربعات الانحرافات) يعرف بالانحراف المعياري (Standard Deviation)،

ولهذا فإن الانحراف المعيارى يعرف عن طريق التباين والذى يعرف على أنه متوسط مربع انحرافات القيم عن وسطها الحسابي و يرمز له بالرمز σ^2 والجذر التربيعى للتباين ينتج عنه مقياس من أهم وأدق مقاييس التشتت وهو ما يسمى بالانحراف المعيارى ويرمز له بالرمز σ وسوف نتناول طريقة حسابه فى حالة البيانات غير المبوبة أو المباشرة والبيانات المبوبة كما يلي :

1.4. حالة البيانات غير المبوبة

إذا كان لدينا مجموعة من البيانات من مجتمع إحصائى عدد مفرداته N و لتكن : $x_1, x_2, x_3, \ldots\ldots, x_n$ و متوسط هذه البيانات هو $\bar{x}$ فإن مربعات انحرافات هذه القيم

عن وسطها الحسابى يكون :

$$, (x_2 - \bar{x})^2, (x_3 - \bar{x})^2, \ldots\ldots\ldots (x_n - \bar{x})^2 (x_1 - \bar{x})^2$$

و بالتالي يعرف التباين كالتالي :

$$\sigma^2 = \frac{\sum(x_i - \mu)^2}{N} \qquad \text{أو} \qquad \sigma^2 = \frac{1}{N}\sum(x_i - \mu)^2$$

حيث μ هو الوسط الحسابي للمجتمع أي أن

$$\mu = \frac{\sum x_i}{N} = \bar{x}$$

كما يمكن صياغته حسب المعادلة التالية بالإضافة إلى صيغ أخرى:

$$\sigma^2 = \frac{\sum x_i^2}{N} - \bar{x}^2$$

و بما أن الانحراف المعياري هو الجذر التربيعي للتباين أي :

$$\text{الإنحراف المعياري} = \sqrt{\text{التباين}}$$

فالإنحراف المعياري للمجتمع يساوي:

$$\sigma = \sqrt{\frac{\sum x_i^2}{N} - \bar{x}^2}$$

كما يمكن صياغته حسب المعادلة التالية بالإضافة إلى صيغ أخرى:

$$\sigma = \sqrt{\frac{\sum(x_i - \bar{x})^2}{N}} \qquad \text{أو} \qquad \sigma = \sqrt{\frac{1}{N}\sum(x_i - \bar{x})^2}$$

وأحيانا يطرح العدد واحد من المفردات (n−1) أو من مجموع التكرارات وذلك عند تقدير مقياس التباين والانحراف المعياري للعينات الصغيرة الحجم وأن العدد واحد يمثل درجات الحرية (تعتبر العينة كبيرة إذا زاد حجمها عن 30 مشاهدة) ويكتب كما يلي:

التباين :

$$S^2 = \frac{\sum(x_i - \mu)^2}{N-1} \qquad \text{أو} \qquad S^2 = \frac{1}{N-1}\sum(x_i - \mu)^2$$

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

الإنحراف المعياري :

$$= \sqrt{\frac{\sum(x_i - \bar{x})^2}{N-1}}\ S \qquad \text{أو} \qquad = \sqrt{\frac{1}{N-1}\sum(x_i - \bar{x})^2}\ S$$

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

مثال : احسب كلاً من التباين والانحراف المعياري للقيم 12 ، 15 ، 11 ، 17 ، 18 ، 20 ، 19

الحل: نكون جدول المعلومات التالي:

x_i	x_i^2	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
12	144	12 – 16 = –4	16
15	225	15 – 16 = – 1	1
11	121	11 – 16 = – 5	25
17	289	17 – 16 = 1	1
18	324	18 – 16 = 2	4
20	400	20 – 16 = 4	16
19	361	19 – 16 = 3	9
$\sum x_i = 112$	$\sum x_i^2 = 1864$	$= 112/7 = 16\bar{x}$	$\sum (x_i - \bar{x})^2 = 72$

1-إذا اعتبرنا أن القيم السابقة خاصة بالمجتمع يكون لدينا تباين المجتمع 10.28 و هذا بعد تطبيق إحدى القوانين السابقة:

$$\sigma^2 = \frac{72}{7} = 10.28$$

أو

$$\sigma^2 = \frac{1864}{7} - 16^2 = 10.28$$

أما الإنحراف المعياري للمجتمع فيساوي الجذر التربيعي لتباين المجتمع أي الجذر التربيعي للقيمة 10.28 و يساوي σ = **3.20**

2-إذا اعتبرنا أن القيم السابقة خاصة بالعينة يكون لدينا تباين العينة = 12 و هذا بعد تطبيق إحدى القوانين السابقة

$$S^2 = \frac{72}{7-1} = 12$$

أو

$$S^2 = \frac{1864 - (7 * 16^2)}{7-1} = \frac{1864 - 1792}{6} = 12$$

أما الإنحراف المعياري للعينة فيساوي الجذر التربيعي لتباين العينة أي الجذر التربيعي للقيمة 12

و يساوي S = **3.46**

2.4. حالة البيانات المبوبة :

إذا كان لدينا عدد k من الفئات ذات المراكز $x_1,x_2,x_3,.....x_k$ ولها التكرارات $f_1,f_2,f_3,.....,f_k$ على الترتيب فإن المعادلات السابقة الخاصة بالتباين و الانحراف المعياري تصبح كالتالي :

تباين المجتمع :

$$\sigma^2 = \frac{\sum f_i {x_i}^2}{N} - \bar{x}^2$$

كما يمكن صياغته حسب المعادلة التالية بالإضافة إلى صيغ أخرى:

$$\sigma^2 = \frac{1}{N}\sum f_i(x_i - \mu)^2 \qquad \text{أو} \qquad \sigma^2 = \frac{\sum f_i(x_i-\mu)^2}{N}$$

الإنحراف المعياري $$\sigma = \sqrt{\frac{\sum f_i x_i^2}{N} - \bar{x}^2}$$

$$\sigma = \sqrt{\frac{\sum f_i(x_i-\bar{x})^2}{N}} \quad \text{أو} \quad \sigma = \sqrt{\frac{1}{N}\sum f_i(x_i - \bar{x})^2}$$

أما بالنسبة للعينة فيكون :

تباين العينة: $$S^2 = \frac{1}{N-1}\sum f_i(x_i - \bar{x})^2$$

الإنحراف المعياري : $$S = \sqrt{\frac{1}{N-1}\sum f_i(x_i - \bar{x})^2}$$

بالإضافة إلى قوانين أخرى

مثال :

يمثل الجدول التالي توزيع تكرار ي لأوزان 50 رياضيا

الفئات	54-50	58-54	62-58	66-62	70-66	74-70	78-74
التكرارات	5	8	11	12	7	4	3

المطلوب حساب التباين و الانحراف المعياري لهذا التوزيع

الحل:

نكون جدول المعلومات التالي :

الفئات	التكرار f_i	مركز الفئة x_i	$f_i.x_i$	x_i^2	$f_i.x_i^2$
54-50	5	52	260	2704	13520
58-54	8	56	448	3136	25088
62-58	11	60	660	3600	39600
66-62	12	64	768	4096	49152
70-66	7	68	476	4624	32368

74-70	4	72	288	5184	20736
78-74	3	76	228	5776	17328
المجموع	50		3128	29120	197792

1-إذا اعتبرنا أن القيم السابقة خاصة بالمجتمع يكون لدينا تباين المجتمع 42.08 و هذا بعد تطبيق إحدى القوانين السابقة:

حساب المتوسط الحسابي :

$$\bar{x} = \frac{\sum f_i x_i}{N}$$

$$\bar{x} = \frac{3128}{50} = 62.56$$

حساب التباين :

$$\sigma^2 = \frac{\sum f_i {x_i}^2}{N} - \bar{x}^2$$

$$\sigma^2 = \frac{197792}{50} - 62.56 = 42.08$$

الإنحراف المعياري :

$$\sigma = \sqrt{42.08} = 6.48$$

3.4. خصائص الإنحراف المعياري:

من خصائص الانحراف المعياري ، ما يلي:

- الانحراف المعياري للمقدار الثابت يساوي صفرا ، أي أنه إذا كان لدينا القراءات التالية:

$x: a, a, a,\ldots,a$ حيث أن a مقدار ثابت فإن $sx = 0$ حيث أن sx تعبر عن الانحراف المعياري لقيم x

- إذا أضيف مقدار ثابت إلى كل قيمة من قيم المفردات ، فإن الانحراف المعياري للقيم الجديدة)القيم بعد الإضافة) تساوي الانحراف المعياري للقيم الأصلية (القيم بعد الإضافة) ، فإذا كانت القيم الأصلية هي: $x1, x2, \ldots, xn$ وتم إضافة مقدار

ثابت a إلى كل قيمة من قيم x فإن الانحراف المعياري للقيم الجديدة $x1 + a, x2 + a, \ldots, xn + a$ هي : $(y = x + a)$: $s_y = s_x$

- إذا ضرب كل قيمة من قيم المفردات في مقدار ثابت ، فإن الانحراف المعياري للقيم الجديدة يساوي الانحراف المعياري للقيم الأصلية مضروبا في الثابت ، أى أن إذا كان قيم x هي القيم الأصلية ، وكانت القيم الجديدة هي : $y = ax$ حيث أن a مقدار ثابت فإن : $s_y = as_x$
- إذا كان لدينا التوليفة الخطية $y = ax + b$ فإن الانحراف المعياري للمتغير y هو أيضا $s_y = as_x$.

4.4. مزايا و عيوب الإنحراف المعياري

- هو أكثر مقاييس التشتت استخداما
- يسهل التعامل معه رياضيا
- يأخذ كل القيم في الاعتبار
- يتأثر بالقيم الشاذة
- قابل للعمليات الجبرية لذلك فهو كثير الاستخدام في القوانين والنظريات الإحصائية
- يمكن الاعتماد عليه للمقارنة بين تشتت توزيعين إحصائيين من نفس النوعية ولهما نفس المتوسط الحسابي
- يأخذ نفس وحدة القياس للمتغير الأصلي

5.4. تقدير مدى الانحراف المعياري

يمكن قياس درجة تشتت البيانات من خلال تقدير المدى الذي يقع داخله الانحراف المعياري و هو:

$$\frac{\text{المدى}}{6} < \text{الإنحراف المعياري} < \frac{\text{المدى}}{4}$$

و إذا كان المدى الذي يقع فيه الانحراف المعياري صغير دل ذلك على أن تشتت البيانات صغير و العكس صحيح، و إذا و قع الإنحراف المعياري خارج المدى دل ذلك على وجود قيم شاذة.

أما إذا أردنا الحصول على قيمة تقديرية للانحراف المعياري فإننا غالباً ما نأخذ القيمة :

$$S = \frac{\text{المدى}}{4}$$

5/ مقاييس التشتت النسبي

هي مقاييس تستخدم للمقارنة بين مجموعتين من البيانات المختلفة في وحدات القياس، فهي مقاييس خالية من وحدات القياس، ويمكن وصفها بأنها مقاييس نسبة، ومن أهمها :

التشتت المطلق أو النسبي و معامل الاختلاف : فالتغير الفعلي أو التشتت كما نحصل عليه من الانحراف المعياري أو غيره من مقاييس التشتت يسمى بالتشتت المطلق ، ولكن تغير أو تشتت 1 متر عن مسافة 1000 متر يختلف في تأثيره عن تغير 1 متر في مسافة 20 متر. ومقياس هذا التأثير نحصل عليه بالتشتت النسبي ويعرف بما يلي :

$$\text{التشتت النسبي} = \frac{\text{التشتت المطلق}}{\text{الوسط}}$$

1.5. معامل الاختلاف C.V (Coefficient of Variation)

يستخدم لمقارنة التشتت بين مجموعتين وذلك للاختلاف الواضح في الوسط الحسابي لمجموعتين من حيث القيمة فصغر الوسط الحسابي في المجموعة الأولى في مقابل كبره في المجموعة الثانية وهو النسبة المئوية بين الانحراف المعياري والوسط الحسابي وبالتالي لا يعتمد على وحدات المتغير الأصلي وبالتالي يمكن استخدامه لمجموعتين مختلفتين في الوحدات، فإذا كان التشتت المطلق هو الانحراف المعياري والمتوسط هو فإن التشتت النسبي يسمى بمعامل الاختلاف أو معامل التشتت ويعرف كالآتي :

معامل الاختلاف = (الانحراف المعياري ÷ الوسط الحسابي) X 100

$$CV = \frac{\sigma}{\bar{x}}.100$$

مثال : أوجد معامل الاختلاف لمجموعتين من درجات الطلبة في مادة الإحصاء حسب المعطيات أدناه:

	المجموعة الأولى	المجموعة الثانية
الوسط الحسابي	40	33
الإنحراف المعياري	24	16

الحل :

- معامل اختلاف المجموعة الأولى :

$$CV = \frac{\sigma}{\bar{x}}.100 = \frac{24}{40} X\ 100 = 60\%$$

- معامل اختلاف المجموعة الثانية :

$$CV = \frac{\sigma}{\bar{x}}.100 = \frac{16}{33} X\ 100 = 48\%$$

بالمقارنة نجد أن تشتت المجموعة الأولى أكبر من تشتت المجموعة الثانية التي تعتبر أكثر تجانسا و أقل ابتعادا عن متوسطها و هذا يعني أن معامل الاختلاف لا يؤكد دائما قيمة التشتت المحسوبة عن طريق الانحراف المعياري لأنه قد يعطي العكس.

2.5. خصائص معامل الاختلاف :

– يقيس الاختلاف النسبي دون وحدة تمييز؛

– ليس له معنى إذا كانت المتوسطات الحسابية معدومة؛

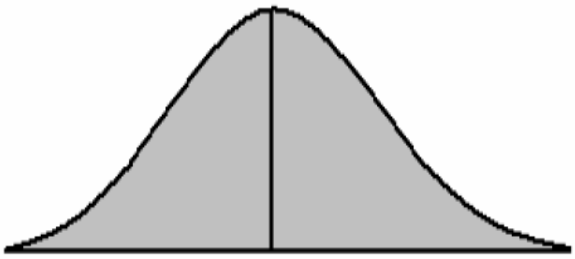
– يستخدم لمقارنة مجموعتين أو أكثر من البيانات من حيث التشتت، خاصة إذا اختلفت متوسطاتها الحسابية.

– يأخذ بعين الاعتبار كل القيم .

الفصل الخامس

مقاييس الشكل

عند تمثيل بيانات الظاهرة في شكل منحنى تكراري فإن هذا المنحنى يأخذ أشكالا مختلفة ، فقد يكون هذا المنحنى متماثل بمعنى أن له قمة في المنتصف ، و لو أسقطنا عمودا من قمته على المحور الأفقي لشطره نصفين متماثلين. فالتوزيع التكراري المتماثل ما كانت تكراراته موزعة توزيعاً متماثلاً حول وسطه الحسابي بما يعني قيم المتغير المتساوية البعد عن الوسط الحسابي لها نفس التكرار.



منحنى التوزيع الطبيعي (منحنى متماثل)

1/الالتواءSKEWNESS

لقد سبق وتحدثنا عن طرق عرض البيانات جدوليا وبيانيا ثم مقاييس النزعة المركزية ومقاييس التشتت ولم نتعرض لأى مقياس يوضح درجة التواء المنحنى التكراري .ويقصد بكلمة التواء هو بعد المنحنى عن التماثل، ويكون منحنى التوزيع التكراري ملتويا نحو اليمين إذا كانت القيم المتطرفة نحو اليمين تؤثر على الوسط الحسابي وتتجه به نحو اليمين، وبذلك يكون الوسط الحسابي أكبر من الوسيط. أما إذا كان التوزيع ملتويا نحو اليسار فإن الوسط الحسابي يكون أصغر من الوسيط.

وهو مقياس نسبي يمكن استخدامه للمقارنة بين التواء التوزيعات كما أن قيمة مقياس الالتواء محصورة بين3+ ، 3 – فالقيمة السالبة تعني الالتواء جهة اليسار والقيمة الموجبة تعني الالتواء جهة اليمين والقيمة صفر تعني عدم وجود التواء أو التماثل (التوزيع متماثل).

و هناك طرق كثيرة لقياس الالتواء منها :

1/ طريقة بيرسون في قياس الالتواء

يقاس الالتواء بأحد المقياسين الآتيين:

$$r = \frac{(\bar{x} - Mod)}{\sigma} \qquad \text{أو} \qquad r = \frac{3(\bar{x} - med)}{\sigma}$$

حيث σ تمثل الإنحراف المعياري و med هو الوسيط و mod هو المنوال

2/ طريقة العزوم في قياس الالتواء

$$r = \frac{{m_3}^2}{\sigma^3}$$

حيث : $m_3 = \frac{\sum(x_i - \bar{x})^3}{n}$ وهذا في حالة البيانات غير المبوبة أو المباشرة

$m_3 = \frac{\sum fi(x_i - \bar{x})^3}{n}$ وهذا في حالة البيانات المبوبة

مثال :

أوجد معامل الإلتواء لدرجات الطلاب إذا علمت أن

$\bar{x} = 65.7$, $med = 65.41$, $mod = 64.83$, $\sigma = 10.67$

الحل :

$$r = \frac{3(\overline{65.7} - 65.41)}{10.67} = 0.08$$

أو

$$r = \frac{(\overline{65.7} - 64.83)}{10.67} = 0.08$$

إذن المنحنى عديم الإلتواء أي أن متماثل

2/ التفرطح(KURTOSIS or PEAKEDNESS)

هو مقياس يقيس درجة علو أو انخفاض أي منحنى توزيع تكراري بالنسبة للمنحنى الطبيعي وهو منحنى متماثل حول الرأس يمر بالمتوسط و للتفرطح ثلاثة أنواع

- متوسط التفرطح (معتدل)
- مدبب : معظم القيم بالقرب من الوسط الحسابي والذيلين
- مفرطح: معظم القيم بعيدة عن الوسط والذيلين

يحسب من الصيغة الرياضية التالية :

$$k = \frac{m_4}{\sigma^4} - 3$$

حيث : $m_4 = \frac{\sum(x_i - \bar{x})^4}{n}$ وهذا في حالة البيانات غير المبوبة أو المباشرة

$m_4 = \frac{\sum fi(x_i - \bar{x})^4}{n}$ وهذا في حالة البيانات المبوبة

الفصل السادس

التوزيع الطبيعي

التوزيع الطبيعي(المعتدل) Normal Distribution

إن من أهم التوزيعات الاحتمالية المتصلة هو التوزيع الطبيعي، أو المنحنى الطبيعي (NORMAL CURVE) أو توزيع جاوس (GAUSSIAN DISTRIBUTION).

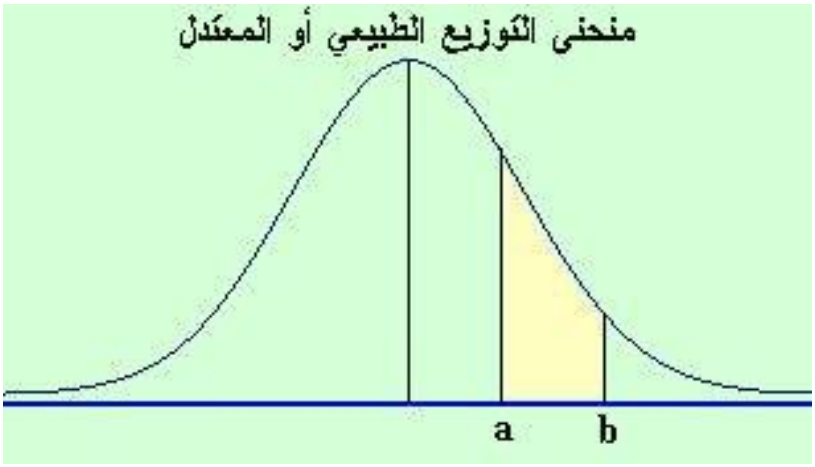
يطلق عادة على التوزيع الاعتدالي اسم المنحنى الاعتدالي وهو من المنحنيات المتصلة التي تعتبر من أهم المنحنيات المستخدمة في البحوث النفسية والتربوية.

يرتبط هذا التوزيع بمتغير عشوائي متصل وهو دالة في المتغير العشوائي ويمكن تمثيله بيانياً وهو من أهم التوزيعات الاحتمالية لتمثيله العديد من الظواهر وهو المناسب لها سواء كانت القيم التي تحدث في الظاهرة كبيرة جداً أو صغيرة جداً باحتمالات صغيرة.

هو توزيع مستمر يعرف أيضاً بتوزيع جاوس (كارل جاوس) حيث جرى نشره سنة 1733م ويعتبر المتغير المعتدل عشوائي مستمر لكونه يتكون من عدد لانهائي من القيم الحقيقية والتي يمكن ترتيبها على مقياس متصل، وهو من أهم التوزيعات في علم الإحصاء بل يعتبر أساساً لكثير من النظريات الإحصائية الرياضية ويلعب دوراً أساسياً في اختبارات الفروض الإحصائية وفترات الثقة وغير ذلك وأن الكثير من الصفات كالطول والوزن ومستوى الذكاء والزواج وما إلى ذلك إذا قيست ولعدد كبير من المشاهدات فإن توزيعها يقترب من التوزيع الطبيعي إن لم يكن يأخذ صورة التوزيع الطبيعي، ويعرف بأسماء مختلفة منها التوزيع ألجرسي لكون شكله يشبه الجرس.

1/ خصائص التوزيع المعتدل:

- منحنى التوزيع المعتدل متصل (مستمر) بمعنى أنه توجد لكل قيمة من قيم X قيمة مناظرة من قيم Y بما في ذلك القيم الكسرية مهما صغرت ، منحناه (Normal Curve) يشبه شكل الجرس ويمتد ذراعه من $-\infty$ إلى ∞ لذلك يمكن أن يسمى المنحنى الجرسي أو منحنى الصدفة أو منحنى الخطأ كما هو موضح بالشكل



معادلته الرياضية في الفترة] – ∞ ، +∞ [هي:

$$Y = \frac{e^{-\frac{(X-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}} \quad , -\infty < X < \infty$$

بإجراء تكامل Y على الفترة] – ∞ ، ∞ [نحصل على المساحة تحت المنحنى وفوق المحور الأفقي، والتمثيل البياني له كما مبين بالشكل المقابل وكل نقطة من نقاط المنحنى تمثل قيمة لدالة تعرف بدالة كثافة الاحتمال (Probability density function) عند هذه النقطة والاحتمال هنا أي في التوزيع المستمر هو قيمة المساحة تحت منحنى دالة الكثافة المناظرة لفترة وليس لنقطة فالمساحة باللون الأصفر والمحصورة بين المنحنى والمحور الأفقي والمستقيمان $x_1 = a$ ، $x_2 = b$ تساوي احتمال المتغير العشوائي المستمر x أي قيمة في الفترة [a , b] هذا وأن المساحة الكلية الواقعة بين منحى التوزيع المعتدل والخط الأفقي تساوي الواحد الصحيح وهي ما تعرف بالمساحة تحت المنحنى= 1 ولمعرفة احتمال وقوع x بين x_1 و x_2 نحسب تكامل الدالة السابقة من x_1 إلى x_2 ، مع ملاحظة أن احتمال أي حدث P(A) يقع بين الصفر والواحد الصحيح أي أن: $0 < P(A) < 1$

- المنحنى متماثل حول الخط الرأسي (العمود النازل من أعلى نقطة للمنحنى على الخط الأفقي) وإن التماثل يعني بأن صورة الشكل على أحد جانبي محور التماثل هي الجزء الواقع على الجانب الأخر وموقع العمود على الخط الأفقي يمثل قيمة الوسط

الحسابي أي أن المنحنى متماثل حول وسطه الحسابي أو حول المستقيم $x = \mu$ ، وان μ هي القيمة المتوقعة ويصل المنحنى لقيمته العظمى عند $X = \mu$

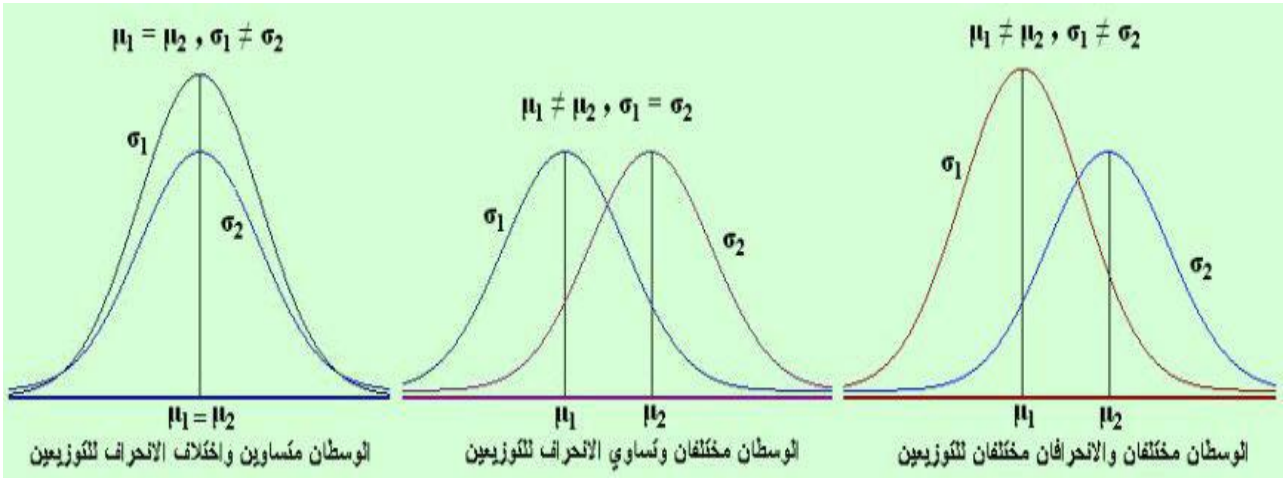
- المنحنى ممتد من $-\infty$ إلى $+\infty$ ولا يلتقي بالمحور الأفقي

أي أنه يمتد من كلتا الجهتين إلى اللانهاية. أي أن المنحنى يقترب تدريجياً من المحور الأفقي ولكنه لا يمسه مهما مددناه من كلتا الجهتين، ولا نحتاج عادة إلى مد طرفي المنحنى بعيداً إلى أقصى اليمين أو أقصى اليسار..

- نقط انقلاب المنحنى وهى النقط التي يتغير فيها اتجاه انحناء المنحنى تحدث عند الانحرافين المعياريين ± 1 على جانبي المتوسط.
- المساحة الكلية تحت المنحنى، أي المساحة المحصورة بين المنحنى والمحور الأفقي تساوي الواحد الصحيح.

والمساحة المحصورة بين أجزاء من المنحنى الاعتدالي المعياري والمحور الأفقي يمكن اعتبارها تكرارات نسبية، أي أننا إذا رسمنا من أي نقطتين على المحور الأفقي مستقيمين موازيين للمحور ارأسي ومددناهما حتى يقابلا المنحنى فإنه يمكننا تحديد علاقة ثابتة للمنحنى الواحد مهما اختلف شكله بين المسافة على المحور الأفقي مقاسه بوحدات انحرافات معيارية والمساحة تحت هذا المنحنى.

- في المنحنى الاعتدالي (في نقطة التقاطع بين المحور الأفقي والمحور الرأسي) نجد أن قيمة الوسط الحسابي والوسيط والمنوال هي نفسها
- للمنحنى المعتدل معلمتين هما الوسط الحسابي والانحراف المعياري معتمد كلياً عليهم فاختلاف الوسط أو الانحراف المعياري لتوزيعين معتدلين يعني اختلاف في الشكل أو اختلاف في المركز كما هو مبين بالشكل الآتي ولكل زوج (μ ، σ) للوسط والانحراف المعياري منحنى توزيع مختلف وبالتالي تختلف المساحة تحت المنحنى لكل منحنى ولذا أخذنا (0 ، 1) كتوزيع معياري يسمى التوزيع الطبيعي المعياري متغيره العشوائي هو Z .



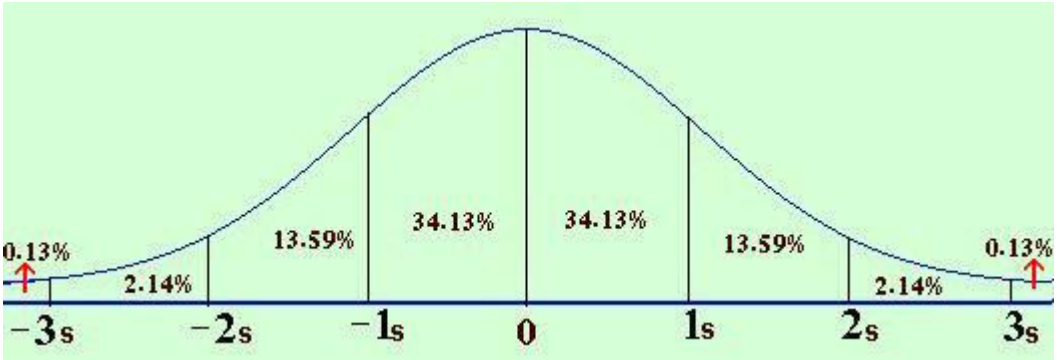
- للمنحنى قمة واحدة أي له منوال واحد وبالتالي فالمنحني وحيد المنوال
- المساحة الواقعة تحت المنحنى والمحصورة بالمستقيمين:

$x = \mu - \sigma$ و $x = \mu + \sigma$ تساوي 68.26% تقريباً من المساحة الكلية تحت المنحنى أي 68.26% من قيم المتغير العشوائي المعتاد تقع في $[\mu + \sigma \ ، \ \mu - \sigma]$

$x = \mu - 2\sigma$ و $x = \mu + 2\sigma$ تساوي 95.45% تقريباً من المساحة الكلية تحت المنحنى أي 95.45% من قيم المتغير العشوائي المعتاد تقع في $[\mu + 2\sigma \ ، \ \mu - 2\sigma]$

$x = \mu - 3\sigma$ و $x = \mu + 3\sigma$ تساوي 99.73% تقريباً من المساحة الكلية تحت المنحنى أي 99.73% من قيم المتغير العشوائي المعتاد تقع في $[\mu + 2\sigma \ ، \ \mu - 2\sigma]$

أي أن وقوع أي مفردة على بعد 1، 2، 3 انحرافات معيارية (1s, 2s, 3s) من الوسط الحسابي هي القيم السابقة كما مبين بالشكل الآتي:



لاحظ أن 34.19% من المساحة تحت المنحنى التي تساوي الواحد الصحيح أي 0.3413 ، وبجمع القيم المبينة في الرسم أعلاه نجد أنها تساوي الواحد الصحيح تقريباً.

إن هذه القيم ما هي إلا احتمالات للقيم كمساحة تحت المنحنى ولأي دالة احتمال يكون مجموع احتمالاتها البسيطة يساوي الواحد الصحيح ونقصد في الأصل المساحة هنا لمساحة الأعمدة للقيم ولكن من الصعب رسم كل الأعمدة وعرض احتمال كل منها ولذا استعضنا عنها باحتمالاتها.

1 ≈ 0.9998 = 0.0013 + 0.0214 + 0.359 + 0.3413 + 0.3143 +0.1359 + 0.0214 + 0.0013

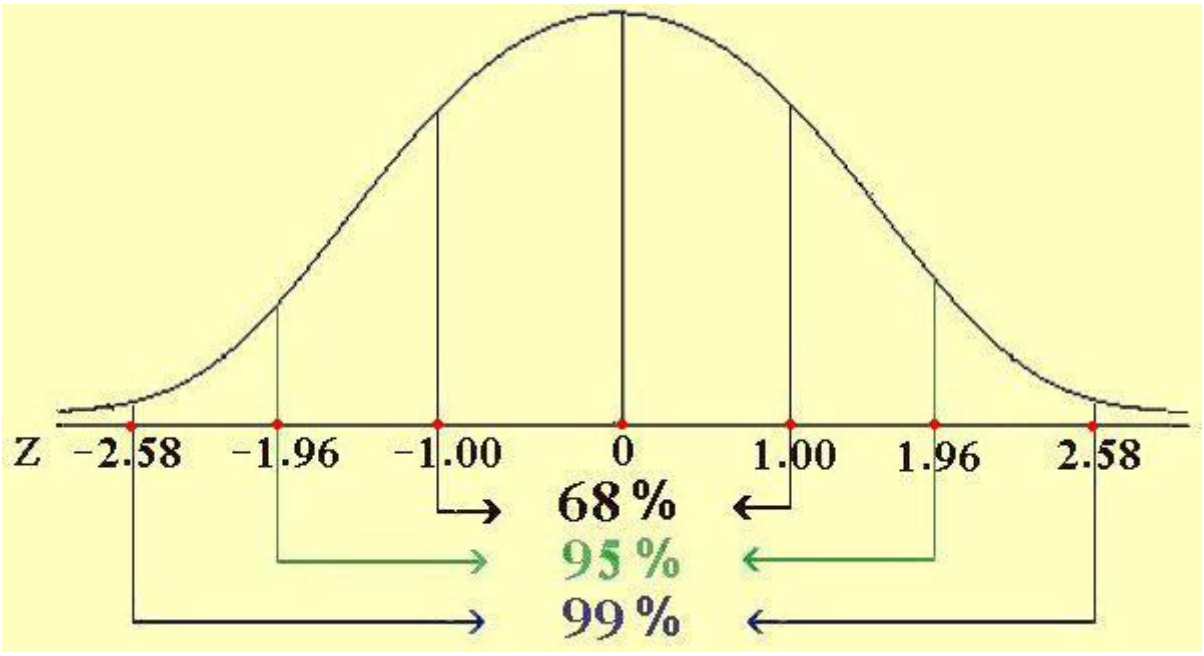
والتوزيع الطبيعي المعياري (Standard Normal Distribution) الذي وسطه صفر وانحرافه المعياري 1 متغيره العشوائي المعياري Z بالصيغة السابق ذكرها، ومنحناه كما مبين أعلاه ويمكن حذف s من القيم على الخط الأفقي وقد نضع قيم x والمناظرة لها Z على الخط الأفقي إن دعت الحاجة.

وقد أمكن إيجاد جدول لتوزيع معتدل معياري لقياس جميع التوزيعات المعتادة (اعتدالي) من خلال النظرية التالية:

إذا كانت x متغير عشوائي له توزيع اعتدالي بتوقع μ وتباين σ^2 حيث σ الانحراف المعياري فإن: $Z = (x - \mu) / \sigma$ له توزيع اعتدالي وسطه صفر وانحرافه المعياري واحد صحيح ويعرف بالتوزيع المعتدل المعياري وله جداول لقيم دالة التوزيع الاحتمالي ولقيم دالة كثافة الاحتمال لأي قيمة من قيم المتغير العشوائي، ويعرف المنحنى هنا بالمنحنى الطبيعي القياسي خطه الأفقي مقسم لدرجات Z كما مبين بالشكل التالي والذي يبين أيضاً المساحة تحت المنحنى وقد قسمت لدرجات معيارية حسب قيم Z حيث قيم x تناظرها قيم Z تحسب من الصيغة الرياضية السابقة أو من الصيغة : $Z = (x - X) / s$ حيث s الانحراف المعياري وأن القيمة Z قيمة معيارية وهي الفرق بين القيمة المشاهدة والوسط الحسابي معبراً عنها بوحدات معيارية بمعنى أكثر دقة قيمة Z عبارة عن عدد الوحدات المعيارية (الانحراف المعياري) التي تفصل بين قيمة x والوسط الحسابي.

وهناك خواص أخرى من بينها إذا كان Ln(x) توزيع طبيعي فإن x توزيع طبيعي يمكن صياغة معادلة المنحنى بدلالة Z على الصورة الآتية حيث أن Y تمثل كثافة قيم المتغير الطبيعي المعياري أو التكرارات للمنحنى.

$$Y = \frac{e^{-Z^2/2}}{\sqrt{2\pi\sigma^2}}$$



يمكن تحويل قيمة المتغير المعتدل x لمتغير معتدل معياري Z من الصيغة السابق فمثلاً إذا كان لدينا توزيع اعتدالي وسطه 150 درجة وانحرافه المعياري 90 درجة فيمكن باستخدام الصيغة السابقة حساب قيمة 2 = 70x نستخدم الصيغة السابقة أي أن:

$$Z = (270 - 90) \div 90 = 2$$

بالرجوع لجدول Z نجد أن المساحة تحت المنحنى التي تقابل Z = 2 تساوي 0.9772 (المساحة التي تقع على يسار العدد 2 (الشكل كل السابق)، وتحسب بطريقتين:

الأولى : المساحة = 1 – (0.0214 + 0.0013) = 1 – 0.0227 = 0.9773

الثانية : المساحة = 0.0013 + 0.0214 + 0.1359 + 0.3413 + 0.3413 + 0.9771 = 0.1359

2/ استخدام خصائص المنحنى الطبيعي في تحليل البيانات:

سنعرض فيما يلي بعض الأسئلة ونجيب عليها باستخدام مجموعة افتراضية من البيانات حتى يتسنى للطالب أو الباحث ملاحظة كيفية استخدام جدول المساحات في الإجابة عن هذه هذه الأسئلة.

مثال:

عندنا بيانات حصلنا عليها في اختبار ما من مجتمع إحصائي متماثل لعينة حجمها 900 ووسطها الحسابي 60 وانحرافها المعياري 15 .

1- ما هي النسبة المئوية للحالات التي تقع بين المتوسط والدرجة 80 ؟ وما هي الرتبة المئينية المقابلة لهذه الدرجة في المجتمع الأصلي؟

الحل:

أولا : نحول الدرجة الخام إلى درجة معيارية

$$Z = \frac{X - \overline{X}}{S_X} = \frac{80 - 60}{15} = 1,33$$

ثانيا : نذهب للجدول ونبحث في العمود الأول عن الدرجة المعيارية 1,33 ، فنجد أن المساحة المحصورة بين المتوسط وهذه الدرجة من العمود الثاني هي 0,40824، وبذلك تكون الرتبة المئينية المقابلة للدرجة 80 هي 0,90824 = 0,40824 + 0,50 يعنى % 90,82

2- ما هي النسبة المئوية التي تقع بين الدرجتين 51 & 68 ؟

الحل:

أولا: نحول كل درجة خام إلى درجة معيارية على حده، ثم نبحث في الجدول عن المساحات لكل منهما.

$$Z = \frac{X - \overline{X}}{S_X} = \frac{68 - 60}{15} = 0,53 \rightarrow 0,2019$$

$$Z = \frac{X - X}{SX} = \frac{51 - 60}{15} = -0,6 \rightarrow 0,2257$$

ثانيا: نجمع المساحتين معاً، أي أن المساحة المحصورة بين الدرجتين 51 & 68 هي:

$$0,2019 + 0,2257 = 0,4276$$

أي ان النسبة المئوية هي : 42.76%

3- ما هي النسبة المئوية للحالات التي نتوقع أن نحصل على درجة أعلى من 80 ؟

الحل:

أولا: نحول الدرجة الخام 80 إلى درجة معيارية، ثم نبحث في الجدول عن المساحة.

$$Z = \frac{X - X}{Sx} = \frac{80 - 60}{15} = 1,33 \rightarrow 0,40824$$

ثانيا: نطرح هذه المساحة من 0,50 هي:

$$0,50 - 0,40824 = 0,09176$$

أي أن النسبة المئوية المطلوبة هي 9,176 %

4-ما هي النسبة المئوية للحالات التي تقع بين الدرجتين 77 & 90 ؟

الحل:

أولا: نحول كل درجة خام إلى درجة معيارية على حده، ثم نبحث في الجدول عن المساحات لكل منهما.

$$Z = \frac{X - X}{Sx} = \frac{90 - 60}{15} = 2 \rightarrow 0,47725$$

$$Z = \frac{X - X}{Sx} = \frac{77 - 60}{15} = 1,13 \rightarrow 0,37076$$

ثانيا : نطرح من المساحة الكبيرة المساحة الصغيرة لكي نحصل على المساحة المحصورة بين الدرجتين.

0,47725 - 0,37076 = 0,10649

أي أن النسبة المئوية بين الدرجتين هي 10,649%

5- ما هو احتمال أن يحصل شخص اختير بطريقة عشوائية من المجتمع الاصلى على درجة 40 أو أكثر ؟

الحل:

أولا: نحول الدرجة الخام 40 إلى درجة معيارية، ثم نبحث في الجدول عن المساحة.

$$Z = \frac{X - X}{SX} = \frac{40 - 60}{15} = -1,33 \rightarrow 0,40824$$

ثانيا: لإيجاد النسبة المئوية للحالات التي تفوق الدرجة المعيارية 1,33- يجب أن نضيف 0,50 إلى المساحة السابقة.

0,50 + 0,0,40824 = 0,90824

وللتعبير عن هذه المساحة باستخدام الاحتمالات يجب أن نحول هذه النسبة المئوية إلى كسر عشري فتصبح 90,824% . يعني أن هناك احتمالا كبيراً أن يحصل شخص اختير بطريقة عشوائية من المجتمع الأصلي على درجة 40 أو أكثر.

6- أراد باحث أن يختار المجموعة المرتفعة في هذا الاختبار من المجتمع الأصلي وهم الذين يمثلون 5% العليا من الدرجات. ما هي الدرجة التي يجب أن يقبلهاالباحث لتكون بمثابة حد فاصل يعتمد عليه في اختيار هؤلاء الأشخاص؟.

الحل :

أولا: إذا كان 5% أعلى من الخط الفاصل، فإن 45% (5% - 50%) تنحصر بين المتوسط وهذا الخط.

ثانيا: نذهب إلى الجدول ونبحث عن الدرجة المعيارية المقابلة للنسبة 0,45 فنجد أن الدرجة 1,645 تناظر الكسر 0,45 .

ثالثا: نحول الدرجة المعيارية السابقة إلى درجة خام باستخدام القانون.

$$Z = \frac{X - X}{SX}$$

$$1{,}645 = \frac{X - 60}{15}$$

$$X = 84{,}675$$

أي أن الشخص الذي حصل على درجة 84,675 أو أكثر في هذا الاختبار يتم اختياره ضمن المجموعة المرتفعة دون سواه.

الفصل السابع

الدلالة الإحصائية

1/ الفروض وأنواعها

الفروض Hypotheses هي علاقات متوقعة بين متغيرين أو أكثر ، أو هى توقعات الباحث لنتائج دراسته ، وتعد الفروض حلولاً محتملة للمشكلة موضع الدراسة ، وتعتمد صياغة الفروض على النظريات أو البحوث السابقة أو كليهما، كما أنها تستخدم المصطلحات والمتغيرات التي حددها الباحث ، والفرض هو حل للمشكلة تؤيده بعض المعلومات أو الحقائق أو الأدلة النظرية أو الدراسات السابقة ، ولكن صحته تعتمد على مدى تأييد الأدلة والشواهد والبيانات الفعلية للفرض .

وتوجد ثلاثة أنواع من الفروض وهى :

1.1. الفرض البحثي Research Hypothesis :

يشتق الفرض البحثي عادة اشتقاقاً مباشراً من إطار نظري معين ، وهو يربط بين الظاهرة المراد تفسيرها وبين المتغير أو المتغيرات التي استخدمناها في هذا التفسير ،ومن أمثلة الفروض البحثية :

- توجد علاقة بين الرضا عن رئيس القسم وطلبته في قسم التربية الخاصة .
- تختلف طالبات المرحلة الثانية عن الطلاب في مستوى القدرة اللفظية .

وبالنظر إلى هذه الفروض نجد أن كلاً منها يتناول ظاهرة معينة واستند إلى إطار نظري في تحديد المتغيرات التفسيرية لهذه الظاهرة .

2.1. الفرض الصفري Null Hypothesis :

يظن البعض أن الفرض الصفري عكس الفرض البحثي ، لكن هذا غير صحيح ، فالفرض الصفري يعبر عن قضية إذا أمكن رفض صحتها فإن ذلك يؤدى إلى الإبقاء على فرض بحثى معين .

وهو يعنى أيضاً عدم وجود علاقة بين المتغيرات أو عدم وجود فروق بين المجموعات ، ولذلك فهو يسمى فرض العدم ، ومعنى ذلك أنه فرض العلاقة الصفرية أو الفروق الصفرية بين المتوسطات " تساوى المتوسطات " ، ويلجأ الباحث للفرض الصفري

فى حال تعارض الدراسات السابقة أو فى حال عدم وجود دراسات سابقة فى موضوع بحثه.

ومن أمثلته : لا توجد فروق بين طريقتي العلاج (أ & ب) فى تعديل السلوك المرضى .

3.1. الفرض الإحصائى Statistical Hypothesis :

عندما نعبر عن الفروض البحثية والصفرية بصيغة رمزية وعددية ، فإنها تسمى عادة الفروض الإحصائية ، فالفرض الإحصائي الصفري يعد بمثابة قضية تتعلق بحدث مستقبلي أو بحدث نواتجه غير معلومة حين التنبؤ ، ولكنه يصاغ صياغة رمزية تسمح بإمكانية رفضه ، وهو ما نلجأ بالفعل إلى اختباره بالأساليب الإحصائية.

وقد يكون الفرض الإحصائي "فرض موجه Directed " وهو صياغة للفرض مع تحديد اتجاه العلاقة " موجبة أو سالبة " ، أو تحديد اتجاه للفروق بين المجموعات فى المتغير التابع ،ومن أمثلته :

– توجد علاقة موجبة بين درجات التحصيل والابتكار لدى طلاب الجامعة.

– يوجد فرق دال إحصائياً بين متوسطي درجات المجموعتين التجريبية والضابطة فى التحصيل ،لصالح المجموعة التجريبية .

وقد يكون الفرض الإحصائي " فرض غير موجه " وهو صياغة للفرض دون تحديد اتجاه للعلاقة أو الفروق ، ومن أمثلته : توجد علاقة بين درجات التحصيل والابتكار لدى طلبة كلية التربية الأساسية .

– يوجد فرق دال إحصائياً بين متوسطي درجات المجموعتين التجريبية والضابطة في التحصيل الدراسي .

2/ مستويات الدلالة الإحصائية Level of Significance :

إن القرار الذى يتخذه الباحث فيما يتعلق بالفرض الصفري الذى يود اختباره أو التحقق من صحته يتطلب وجود قاعدة يستند إليها فى هذا الشأن ،فالباحث يحاول التوصل إلى أدلة من البيانات التي قام بجمعها تمكنه من رفض الفرض الصفري وقبول

أو تأييد الفرض البحثي الذى يشتق من إطار نظري يتبناه ويرى انه يفسر الظاهرة تفسيراً منطقياً ، لذلك ينبغي أن يحدد الباحث قبل عملية جمع البيانات قيمة احتمالية معينة تبين مقدار الخطأ الذى يقبل أن يقع فيه نتيجة رفضه للفرض الصفري ،وبعبارة أخرى إذا قرر الباحث على أساس البيانات التجريبية التي حصل عليها رفض الفرض الصفري ، فإن احتمال خطأ هذا القرار يكون أقل من أو مساوياً هذه القيمة التي يطلق عليها مستوى الدلالة الإحصائية أو ألفا .

وطبقاً لإجراءات اختبار الفرض الصفري فإننا نرفض الفرض إذا كانت إحصاءة العينة " كالفرق بين المتوسطات ، أو معامل الارتباط " أكبر أو أصغر مما يمكن توقعه طبقاً لعوامل الصدفة وحدها ، ونستخلص أن هناك فرقاً دالاً أو علاقة دالة بين المتغيرات ، إلا أن هناك خطأ شائعاً هو الخلط بين الدلالة الإحصائية والفائدة العملية للنتائج ، فالنتائج الدالة إحصائياً لا تنطوي بالضرورة على قيمة عملية أو نظرية .

ومن الأخطاء الشائعة أيضاً الخلط بين الدلالة الإحصائية والدلالة النفسية أو التربوية ، إن الدلالة النفسية أو التربوية تعنى القدر الذى يمكن لنتيجة ما أن تضيف للمعرفة ، وتتضمن الدلالة النفسية أو التربوية ثلاثة عناصر :

- ✓ قيمة الفروض التي وضعها الباحث والأفكار النظرية التي استمدت منها هذه الفروض ، وقدرتها على تفسير البيانات التي يحصل عليها الباحث .
- ✓ كفاية الدراسة كاختبار للفروض ، بما في ذلك مدى جودة تصميمها ، واستخدام أدوات حديثة صادقة في جمع البيانات .
- ✓ وضوح نتائج الدراسة ،فالنتيجة الدالة إحصائياً لا تضيف دائماً لفهمنا للسلوك الإنساني ، ومع ذلك فقد يكون لدى البعض نزعة للتركيز على الدلالة الإحصائية ، رغم ما قد يكون بالنتائج من ضعف ، لا يساعد على تفسير سليم له معنى لهذه النتائج .

ومن المتفق عليه استخدام مستويات الدلالة التالية فى البحوث النفسية والتربوية والاجتماعية :

مستويات الدلالة" الشك فى القرار "	0.01	0.02	0.03	0.04	0.05
مستوى الثقة فى القرار أو النتائج	0.99	0.98	0.97	0.96	0.95

كما يستخدم مستويات الدلالة (0.001 ، 0.002 ، 0.003 ، 0.004 ،0.005) أو (0.000) وذلك لتقليل الخطأ في رفض الفرض الصفري الصحيح ، فكلما صغرت قيمة مستوى الدلالة كلما زاد خطأ النوع الثاني " بيتا " .

ويكتفى الباحثون دائماً بمستويي الدلالة (0.05 &0.01) وهو أمر متفق عليه وليس له دليل علمي أو منطقي.

3/ خطأ النوع الأول & والثاني Type 1 Error Type 2 Error :

المنطق في اختبار الفروض هو أن الباحث يفترض صحة الفرض الذى يرغب فى اختباره، ثم يفحص نتائج هذا الفرض في ضوء توزيع العينة الذى يعتمد على صحة الفرض، وإذا تحدد من توزيع العينة أن البيانات الملاحظة احتمال حدوثها كبير فانه يتخذ قراراً بان البيانات لا تتعارض مع الفرض ، ومن ناحية أخرى إذا كان احتمال مجموعة البيانات الملاحظة ضعيف في حالة الفرض الصحيح ، فإن قراره يكون بأن البيانات تتعارض مع الفرض.

وإن صدق النتائج التي نحصل عليها من العينة يتوقف على درجة تمثيلها للمجتمع الأصلي الذى سحبت منه ، وحيث إننا نرتضى عينة لبحثنا فإننا مضطرون لقبول ما تأتى به العينة ، لأننا لا نملك إلا أن نأخذ بصحة المعلومات والبيانات التي وفرتها لنا ونستخدم ذلك فى الحكم على الفرض الخاص بالمجتمع ككل.

ومن ثم يتضح أن أي حكم أو قرار نتخذه بصدد الفرض الصفري يحتمل الصحة أو الخطأ ، ونكون بذلك أمام أربعة بدائل :

أ- أن يكون الفرض الصفري صحيحاً ، وتأتى نتائج العينة تؤيد صحته فإننا نقبله ويكون القرار سليماً ، أو الحكم صائباً.

ب- أن يكون الفرض الصفري خاطئاً ، وتأتى نتائج العينة تثبت صحته، فإننا نقبله ويكون القرار خاطئاً أو الحكم غير صائب ويسمى خطأ بيتا أو نمط "2" ويعنى قبول الفرض الصفري بينما هو في واقع الأمر خاطئ.

جـ- أن يكون الفرض الصفري صحيحاً ، وتأتى النتائج من العينة لا تؤيده ، فإننا نرفضه ويكون القرار خاطئاً ، والحكم غير صائب ويسمى خطأ ألفا أو نمط "1" ويعنى رفض الفرض الصفري بينما هو في واقع الأمر صحيح .

د- أن يكون الفرض الصفري خاطئاً ، وتأتى نتائج العينة تؤيد خطئه فإننا نرفضه ويكون القرار صائباً أو الحكم سليماً .

ويمكن تلخيص الحالات السابقة على النحو التالي :

القرار	الفرض الصفري	
	صحيح	خطأ
قبول الفرض الصفري	قرار صائب	خطأ النوع الثاني بيتا
رفض الفرض الصفري	خطأ النوع الأول ألفا	قرار صائب

ويمكن توضيح نوعى الخطأ بالمثالين الآتيين :

مثال 1 :

نفترض أن التغذية الراجعة ليس لها تأثير بالفعل على سلوك حل المشكلة ، ولكننا لاحظنا عن طريق الصدفة أن سلوك حل المشكلة كان أفضل في وجود التغذية الراجعة ، فإننا ربما نستنتج أن التغذية الراجعة تؤدى إلى تحسين سلوك حل المشكلة في حين أن الأمر ليس كذلك ، فعندئذ نكون قد وقعنا في خطأ من النوع الأول " ألفا" .

مثال 2 : عند محاكمة متهم يمكن الوقوع في أي من نوعى الخطأ ، فتجريم شخص برئ يعد خطأ من النوع الأول "الفا" ، وتبرئة شخص مذنب يعد خطأ من النوع الثاني "بيتا" ، وبالطبع ينبغي التقليل بقدر الإمكان من كلا النوعين من الأخطاء.

4/ قوة الاختبار الإحصائي Test Power:

تعتمد قوة الاختبار على كل من مستوى الدلالة ألفا وخطأ النوع الثاني بيتا وحجم العينة ، وهى احتمال قرار رفض فرض العدم عندما يكون البديل صحيحاً قوة الاختبار الإحصائي = 1– بيتا

ويمكن زيادة قوة الاختبار عن طريق مستوى الدلالة وتباين الدرجات وحجم العينة ، فإذا كان مستوى الدلالة ثابتاً وكذلك التباين فإن زيادة حجم العينة يزيد من قوة الاختبار ، وليس معنى هذا أن حجم العينة هو السبب فى زيادة قوة الاختبار ، وإنما قيمتي مستوى الدلالة "ألفا" وخطأ النوع الثاني "بيتا" وكذلك تباين المجتمع لهما أثر كبير على قوة الاختبار بجانب حجم العينة ، فإذا كانت قيمة ألفا ثابتة وكذلك حجم العينة ،فإن قيمة بيتا تقل بزيادة الفرق بين المتوسطين ، ومعنى هذا أنه كلما كان الفرق بين المتوسطين كبيراً ، فإن احتمال قبول فرض العدم يقل ، أما إذا كان الفرق بين المتوسطين ثابتاً وكذلك حجم العينة ، فإن قيمة بيتا تزداد كلما نقصت قيمة ألفا ، أى أنه إذا كانت ألفا صغيرة فقد نفشل فى رفض فرض العدم بالرغم من وجود فرق بين المتوسطين .

وإذا كانت قيمة ألفا ثابتة وكذلك الفرق بين المتوسطين ، فإن حجم العينة يحدد قيمة بيتا ،فكلما صغرت العينة تزداد قيمة بيتا ومن ثم تنقص قوة الاختبار ، وكلما زاد حجم العينة فإن قيمة بيتا تنقص وتزداد قوة الاختبار .

5/ درجات الحرية Degrees of Freedom:

ويقصد بها عدد أفراد العينة ناقصاً عدد القيود ، فإذا رمزنا لحجم العينة بالرمز (ن) فإن الحرية في اختيار أفراد العينة هي (ن –1) وتسمى بدرجات الحرية ، وتختلف وفقاً للاختبارات الإحصائية المستخدمة ، أو القيود التي يتم وضعها للمقارنة

قائمة المراجع

المراجع :

1-أمانى موسى محمد، التحليل الإحصائى للبيانات، معهد الدراسات والبحوث الإحصائية – جامعة القاهرة، 2007.

2-إبراهيم مراد الدعمة ومازن حسن الباشا " :أساسيات في علم الإحصاء مع تطبيقات spss دار المناهج، الأردن، الطبعة الأولى2013

3-خالد أحمد فرحان المشهداني ورائد عبد الخالق عبد الله العبيدي " :مبادئ الإحصاء" ، دار الأيام، .الأردن،2013

4-عبد الرزاق عزوز " :الكامل في الإحصاء "الجزء الأول، ديوان المطبوعات الجامعية، الجزائر،2010

5-عبد الرزاق عزوز " :الكامل في الإحصاء "الجزء الثاني، ديوان المطبوعات الجامعية، الجزائر،2011

6-محمد بوهزة " :محاضرات في الإحصاء الوصفي"، دار المحمدية العامة، الجزائر، 2011

7-وليد إسماعيل السيفو " :أساسيات الأساليب الإحصائية للأعمال"، منشورات زمزم، الأردن، الطبعة الأولى 2010.

8-إياد محمد الهوبي، الإحصاء التطبيقي ، الكلية الجامعية للعلوم و التكنولوجيا ، فلسطين، 2014.

9-مروان عبد المجيد ابراهيم، الأسس العلمية والطرق الاحصائية للاختبارات، والقياس في التربية الرياضية، الجامعة الهولندية، 1998.

10- عزام عبد الرحمن صبري، الإحصاء التطبيقي بنظام spss، الدار المنهجية للنشر و التوزيع، الطبعة الاولى 2015.

11- Amany Mousa," Statistical Data Analysis", Center for Advancement of Postgraduate Studies and Research, Faculty of Engineering, Cairo University , Cairo 2005.

12- Bernard Py : « Statistique Descriptive », 4 ème Edition Economica, Paris, 1996.

13- Nick, Todd G.. "Descriptive Statistics". *Topics in Biostatistics*. New York: Springer. 2007