الجمهورية الجزائرية الديمقراطية الشعبية

وزارة التعليم العالي والبحث العلمي

جامعة محمد بوضياف المسيلة

كلية العلوم الاقتصادية والتجارية وعلوم التسيير

قسم العلوم التجارية

مطبوعة في مقياس

الإحصاء الوصفي

موجهة إلى طلبة السنة الأولى جذع مشترك

من إعداد الدكتورة:

نوال فرقش

فرحات عباس

أستاذ محاضر (أ)

السنة الجامعية: 2016-2017

فهرس المحتويات

فهرس المحتويات

مقدمة

مقدمة

يعتبر علم الإحصاء من أقدم العلوم حيث ظهر مع حاجة الإنسان الأولى للتعامل مع القيم والأعداد لتسيير الحياة اليومية. ومع التطور الهائل في كافة العلوم في أواخر القرن العشرين تطور ذلك العلم ليستفيد من تقنيات الحاسب الآلي بشكل يجعله العلم الأكثر تداخلا مع العلوم الأخرى المختلفة، ليصبح يستخدم في العلوم التجارية وعلوم الطب والهندسة والأدب وجميع العلوم الأخرى دون استثناء.

ويختلف مفهوم كلمة الإحصاء عند العامة، حيث يشيع لدى العديد من الناس بأن الإحصاء ما هو إلا عد أو حصر للأشياء والتعبير عنها بارقام، وهذا مفهوم محدود لذلك العلم، حيث أن الإحصاء كعلم، هو الذي يهتم بطرق جمع البيانات، وتبويبها، وتلخيصها بشكل يمكن الاستفادة منها في وصف البيانات وتحليلها للوصول إلى قرارات سليمة في ظل ظروف عدم التأكد.

وبذلك يمثل علم الإحصاء الأداة العلمية التي يتم من خلالها إبراز المعلومة المحتواة في البيانات والتي يصعب قراءتها من خلال البيانات مباشرة. وعليه، يعد هذا العلم من العلوم الضرورية التي على الطلبة بشكل عام وطلبة العلوم الاقتصادية بشكل خاص، اكتساب المفاهيم الأساسية منها بالشكل الذي يساعدهم على اتخاذ القرارات المناسبة في مجال تخصصهم.

واستفاءا لذلك الغرض، تم تصميم هذه المطبوعة التي تناولت نوعا من أنواع الإحصاء وهو الإحصاء الوصفي، وهي تضم في طياتها سلسلة من المحاضرات الموجهة لطلبة السنة الأولى نظام (ل.م.د) في العلوم الاقتصادية والتجارية وعلوم التسيير، حيث تتضمن دروس مبسطة ومختصرة وسهلة الفهم مدعمة بالعديد من الأمثلة والتمارين التي تعرض بحلول نموذجية بغية التخفيف من الصعوبات التي تواجه الطلبة في تحقيق فهم أفضل لهذا المقياس.

و بناء عليه، تم تقسيم هذه المطبوعة إلى الفصول الخمسة الموالية:

- المفاهيم الأساسية لعلم الإحصاء؛
- طرق عرض البيانات
- مقاييس النزعة المركزية؛
- مقاييس التشتت؛
- مقاييس الشكل؛

الفصل الأول

مفاهيم أساسية حول علم الإحصاء

- ✓ تعريف علم الإحصاء
- ✓ مراحل البحث الإحصائي
- ✓ بعض المفاهيم الأساسية في الإحصاء

الفصل الأول: مفاهيم أساسية حول علم الإحصاء

يعد الإحصاء من أهم الوسائل العلمية المستخدمة فى ميادين البحث العلمي بصفة عامة والعلوم الاقتصادية بصفة أخص، لذا ينبغى على الباحث الذى يريد القيام بإحصاء لبيانات بحثه أن يلم ببعض المفاهيم الإحصائية الهامة التي توضح له الطريق الصحيح لاستخدام تلك الوسيلة.

والانتشار الواسع لاستخدام الإحصاء دليل على قوة ومصداقية الطرق التي يوفرها هذا العلم، وإذا كان بطء تأقلم برامج الدراسة في التعليم الأساسي والثانوي قد حال دون تعليم كاف لتقنياته في هذه المراحل فإن الجامعات قد أدركت منذ فترة طويلة أهمية هذه الوسيلة ما جعل دراسة مقياس الإحصاء مدرج في مناهج جميع الشعب تقريبا.

1- تعريف علم الإحصاء

يعرف علم الإحصاء بأنه الطريقة الإحصائية التي تمكن من: [1]

- جمع الحقائق عن الظواهر المختلفة في شكل قياسي؛
- تسجيل بيانات تلك الحقائق في جداول تلخيصية؛
- عرض بيانات تلك الجداول بيانيا وتحليلها بهدف معرفة اتجاهات هذه الظواهر والعلاقات فيما بينها.

أي أن علم الإحصاء يختص بالطريقة العملية لجمع وتنظيم وتلخيص وعرض وتحليل البيانات بهدف الوصول إلى نتائج مقبولة وقرارات على ضوء هذا التحليل.

كما يقصد به "العلم الذي يدرس مختلف طرق ووسائل جمع البيانات الكمية عن مختلف الظواهر الاقتصادية والاجتماعية وغيرها، وترتيب هذه البيانات وتبويبها وتحليلها وتفسيرها وتقديمها بأشكال وصور ملائمة بهدف تسهيل اتخاذ القرار على أساس سليم".[2]

ويعرف في موضع آخر بأنه: "علم جمع وترتيب معلومات خاصة بظاهرة معينة وقياس الوقائع كأساس للاستقراء ".[1]

1- مصطفى الخواجة، **مقدمة في الإحصاء**، الدار الجامعية، الإسكندرية، 2002، ص 2.

[2]- محمد راتول، **الإحصاء الوصفي**، ديوان المطبوعات الجامعية، الجزائر، الطبعة الثانية، 2006، ص: 07.

ومنه، ومن خلال ما سبق من تعاريف، يمكن القول أن الإحصاء هو ذلك العلم الذى يهتم بجمع البيانات الكمية عن مختلف الظواهر وتنظيمها فى صورة جداول أو رسوم بيانية، ووصف تلك البيانات باستخدام مفاهيم إحصائية معينة ثم الاستدلال من تلك البيانات على نتائج معينة يراد الوصول إليها.

ومن هذا التعريف يمكن التوصل إلى جملة من النقاط، يتم إيجازها فيما يلي:

- البيانات هي المجال الرئيسي لمراحل علم الإحصاء.
- المراحل الأساسية للعملية الإحصائية تتمثل في أربع مراحل رئيسية هي جمع البيانات، تبويبها، العرض البياني لها ثم تحليلها.
- الهدف الأساسي من العملية الاحصائية هو تحليل البيانات وتفسيرها.
- يمكن تطبيق عملية الإحصاء في مختلف المجالات.

ويمكن تقسيم علم الإحصاء إلى قسمين وهما الإحصاء الوصفي والإحصاء التحليلي، حيث يشمل الأول الطرق المستخدمة في وصف وتنظيم وعرض البيانات وهو ما سيكون محور اهتمام هذه المطبوعة. أما النوع الثاني (الاستدلالي) فيتعلق باستخلاص النتائج من بيانات محدودة (العينة) وتعميمها على كل البيانات (المجتمع) حيز الدراسة.

2 – مراحل البحث الإحصائي:

من خلال التعريف السابق للإحصاء يتبين بأن البحوث الإحصائية في العموم تتضمن مراحل أساسية يجب على الباحث أن يتبعها، وهذه المراحل هي:

3- جيلالي جلاطو، **الإحصاء مع تمارين ومسائل محلولة**، ديوان المطبوعات الجامعية، الجزائر، 2002، ص: 03.

2-1- تحديد الظاهرة المدروسة

وذلك من خلال تحديد الإطار العام للظاهرة المدروسة والذي يشمل الهدف من الدراسة، المجتمع الإحصائي والمكان والوقت المناسبين لجمع البيانات، الصفات المطلوب معرفتها ووحدات القياس المستخدمة...إلخ.

2-2- جمع البيانات الإحصائية

يحتاج الباحث الذى يتعرض لدراسة ظاهرة ما من الظواهر إلى جمع بيانات حول طبيعة هذه الظاهرة من مصادر متعددة.

2-2-1-أنواع البيانات الإحصائية (المتغيرات الإحصائية)

البيانات قد تجمع وتوصف باستخدام الألفاظ فتسمى "بيانات كيفية"، أو تجمع بصورة عددية "رقمية" وتسمى فى هذه الحالة "بيانات كمية".

2-2-1-1- البيانات الكيفية

هي تلك المتغيرات أو الظواهر التي لا يمكن قياسها عدديا بل قياس تكرارها فقط، [1] مثل تقديرات النجاح، الجنسية، الحالة العائلية، اللون...إلخ. وتنقسم هذه البيانات إلى نوعين هما:

- بيانات كيفية قابلة للترتيب، حيث يمكن ترتيبها حسب رتبة معينة تصاعديا أو تنازليا مثل المستوى التعليمي، الرتب العسكرية، تقديرات النجاح وغيرها.

- بيانات كيفية غير قابلة للترتيب، مثل الجنسية، الحالة العائلية، الجنس، اللون.

2-2-1-2- البيانات الكمية

هي تلك الخصائص التي يمكن قياسها وهي أكثر المتغيرات انتشارا واستعمالا لأن لغة الإحصاء هي لغة الأرقام، مثال ذلك الإنتاج، الاستهلاك، الاستثمار، الوزن.[2] وتنقسم المتغيرات الكمية بدورها إلى قسمين هما:

---

[1]- وليد اسماعيل السيفو وآخرون : **أساسيات الأساليب الإحصائية للأعمال**، زمزم ناشرون وموزعون، الأردن، الطبعة الأولى 2010، ص: 29.

[2]- جيلالي جلاطو، **مرجع سبق ذكره**، 2001، ص: 06.

– متغيرات كمية منقطعة، هي تلك المتغيرات التي تأخذ قيما صحيحة لا يمكن تجزئتها، مثل عدد الأطفال في الأسرة الواحدة، عدد الغرف بالبيت، عدد حوادث المرور، عدد أفراد الأسرة، عدد قطع الغيار المنتجة....إلخ.

– متغيرات كمية مستمرة، هي تلك المتغيرات التي تأخذ كل القيم الممكنة لمجال الدراسة، ونظرا للعدد غير المتناهي لهذه القيم نقسم مجال الدراسة إلى مجالات جزئية تسمى الفئات، مثل الطول، السن، الوزن،....إلخ.

2-2-2- طرق جمع البيانات

ويتم جمع البيانات الإحصائية بطرق مختلفة، وذلك حسب الهدف من الدراسة وأسلوب التحليل المتبع، ومن بين الطرق المتبعة في جمع البيانات ما يلي:

2-2-1- الطريقة غير المباشرة والطريقة المباشرة

تسمى الطريقة غير المباشرة أيضا بطريقة البيانات الثانوية، وهي تشمل جميع المصادر التي يتم الحصول منها على البيانات بشكل غير مباشر، بمعنى آخر يتم الحصول عليها بواسطة أشخاص آخرين، أو أجهزة وهيئات رسمية متخصصة مثل نشرات وزارة الزراعة ونشرات مصلحة الإحصاء ونشرات منظمة الأغذية" الفاو"....وهكذا. [1]

ولهذه الطريقة فوائد متعددة أهمها أنها تؤدي إلى اقتصاد كبير في وقت الباحث ونفقاته، إلا أنها تشكو أيضا من عدد من العيوب منها:

– قد تكون قديمة وغير متجددة.

– قد لا تفي تماما بغرض البحث.

– قد يكون بها بعض التحيز الذي يعيق من الاستفادة من البيانات بصورة كاملة.

أما الطريقة المباشرة في جمع البيانات فتسمى كذلك بطريقة البيانات الأولية ويقصد بها قيام الباحث بجمع المعلومات الإحصائية بنفسه، من المفردة محل البحث مباشرة. 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 وتكلفة ومجهود كبير.

[1]- شرف الدين خليل، **الإحصاء الوصفي**، شبكة الأبحاث والدراسات الاقتصادية، WWW.RR4EE.NET، ص: 11.

2-2-2-طريقة الحصر الشامل وطريقة العينة

تستخدم طريقة الحصر الشامل إذا كان الغرض من البحث هو حصر جميع مفردات المجتمع، حيث يتم في هذه الحالة جمع بيانات عن كل مفردة من مفردات المجتمع بلا استثناء كحصر جميع المزارع التي تنتج التمور، أو حصر البنوك الإسلامية في المنطقة. ويتم استخدام هذه الطريقة في الحالات الآتية:

- البيانات المطلوبة تخص كل مفردة من مفردات المجتمع.
- الحصول على نتائج تكون على مستوى عال من الدقة.
- عدم تجانس مفردات المجتمع وإذا ما كان المجتمع صغير نسبيا.

وتعتبر هذه الطريقة من أفضل طرق جمع البيانات وذلك لكونها تعط بيانات كاملة حول المشكلة التي تهم الباحث[1]، تتميز بالشمول وعدم التحيز، كما تمتاز بدقة النتائج لأنها تمس كل مفردة من مفردات المجتمع. ومن بين ما يعاب على هذه الطريقة أنها تحتاج إلى الوقت والمجهود، والتكلفة العالية.

أما الطريقة الثانية وهي طريقة العينة الإحصائية فتقوم على دراسة جزء من المجتمع الإحصائي فقط، وذلك بأخذ عينة عشوائية من المجتمع يتم اختيارها بطرق علمية سليمة ودراسة خواصها واستخلاص المعلومات اللازمة منها، ثم تعميم نتائجها على المجتمع الذي سحبت منه. ومن بين ما تتميز به هذه الطريقة:[2]

- توفير الجهد والتكلفة.
- توفير الوقت، فغالبا ما يكون الباحث مجبرا على جمع البيانات خلال فترة محددة.
- إمكانية الحصول على بيانات ومعلومات وفيرة لاقتصار الباحث على مجموعة جزئية من المجتمع.
- سيطرة الباحث على حجم محدد من أفراد المجتمع المتمثل بالعينة يؤدي إلى سيطرته على البيانات والمعلومات المجمعة وبالتالي الدقة في التعامل مع البيانات وتجميعها.
- صعوبة وفي بعض الأحيان استحالة الوصول إلى كل وحدات المجتمع.

ومن جانب آخر، يعاب على طريقة المعاينة أن النتائج التي تعتمد على هذا الأسلوب تكون أقل دقة من نتائج أسلوب الحصر الشامل، وخاصة إذا كانت العينة المختارة لا تمثل المجتمع تمثيلا جيدا.

---

[1]- خالد أحمد فرحان المشهداني ورائد عبد الخالق عبد الله العبيدي، **مبادئ الإحصاء متضمن التحليل الإحصائي SPSS**، دار الأيام، الأردن، 2013، ص:28.

[2]- عامر إبراهيم قنديلجي، **منهجية البحث العلمي**، دار اليازوري، عمان، الأردن، 2012، ص ص: 188-190.

2-3- عرض البيانات الإحصائية

الخطوة الموالية بعد جمع البيانات في مجال الإحصاء الوصفي، هو تبويب البيانات وعرضها بصورة يمكن الاستفادة منها في وصف الظاهرة محل الدراسة، من حيث تمركز البيانات ودرجة تجانسها. وهناك طريقتين لعرض البيانات هما العرض الجدولي والعرض البياني وهو ما سيتم التعرض له بشيء من التفصيل في الفصل الثاني من هذه المطبوعة.

2-4- تحليل البيانات الإحصائية

وتتضمن هذه المرحلة دراسة المعلومات الإحصائية وترتيبها وتحليلها إلى عناصرها الأولية وإظهار العلاقة بينها، ويتم تحليل المعلومات بإجراء الخطوات الآتية:

- ترتيب الإحصاءات وتصنيفها، ويمكن أن يكون الترتيب حسب النوع أو الكمية، كتصنيف السكان ما بين أعزب ومتزوج ومطلق وأرمل، كما يمكن أن يكون الترتيب جغرافيا، كأن يتم توزيع السكان في الجزائر حسب الولايات والدوائر والبلديات.
- حساب القيم المركزية لمجموعة البيانات ودراسة التشتت والالتواء فيها.
- دراسة علاقات الارتباط بين عوامل المجتمع الإحصائي.
- استنباط التقديرات أو التنبؤات التي تدل عليها الدراسة.

2-5- تفسير البيانات الإحصائية:

من المعروف أن الدراسات الإحصائية تتخذ أساسا في إعداد السياسات واتخاذ القرارات المتعلقة بالمواضيع الاقتصادية والاجتماعية وغير ذلك، وعليها تبنى اتجاهات الدولة أو الشركات أو المؤسسات العامة والخاصة، من هنا كان لزاما على الإحصائي باعتباره أكثر الناس دراية وخبرة في فهم مضمون الأعداد أن يفسر النتائج المتوصل إليها وأن يوضح بصراحة ما تعنيه.

3- بعض المفاهيم الأساسية في الإحصاء

يتميز علم الإحصاء مثل العلوم الأخرى بأنه مليء بالتعاريف والمصطلحات الجوهرية التي يستلزم فهم معانيها، منها:

3-1- المجتمع

هو مجموعة من المفردات التي تشترك في صفات وخصائص محددة. ومجتمع الدراسة هو الذي يشمل جميع مفردات الدراسة، أي هو الكل الذي نرغب دراسته،[1] مثل: مجتمع من الطلبة، مجتمع من الأسر، مجتمع من المؤسسات.

3-2- الوحدة الإحصائية

هي الوحدة الأساسية لتكوين المجتمع الإحصائي،[2] سواء أكانت هذه الوحدة إنسانا أو حيوانا أو شيئا ما.

3-3- الظاهرة الإحصائية

هي الخاصية المدروسة أو المتغير المدروس في المجتمع الإحصائي مثل طول القامة، السن، الوزن، العلامة المتحصل عليها في امتحان معين، الإنتاج ، الادخار ، الاستهلاك...إلخ.

3-4-العينة الإحصائية:

هي جزء من المجتمع الإحصائي يتم استخراجها بطرق إحصائية معينة حتى تكون ممثلة للمجتمع الإحصائي أحسن تمثيل. ويتوقف نجاح استخدام أسلوب العينة الإحصائية على عدة عوامل منها كيفية تحديد حجم العينة، طريقة اختيار مفردات العينة ونوع العينة المختارة. وفي العموم، يمكن تقسيم العينات وفقا لأسلوب اختيارها إلى نوعين هما العينات الاحتمالية والعينات غير الاحتمالية.

3-4-1- العينات الاحتمالية

هي العينات التي يتم اختيار مفرداتها وفقا لقواعد الاحتمالات، بمعنى آخر هي التي يتم اختيار مفرداتها من مجتمع الدراسة بطريقة عشوائية، بهدف تجنب التحيز الناتج عن اختيار المفردات. ومن أهم أنواع العينات الاحتمالية، ما يلي:

---

[1]- شرف الدين خليل، **مرجع سبق ذكره**، دون ذكر سنة النشر، ص: 12.

[2]- جيلالي جلاطو، **مرجع سبق ذكره**، 2001، ص: 05.

3-4-1-1- العينة العشوائية البسيطة

هي العينة التي تؤخذ بشكل يعطي لأي عنصر من عناصر المجتمع نفس الفرصة لأن يكون ضمن مفرداتها. يستخدم هذا النوع من العينات في حالة المجتمعات المتجانسة والمحدودة، ويتم اختيار أفراد العينة بكتابة أرقام أفراد المجتمع على بطاقات متشابهة وخلطها بشكل جيد ثم تتم عملية السحب عن طريق القرعة، أو باستخدام جداول الأرقام العشوائية أو باستخدام الحاسب الآلي.

3-4-1-2- العينة العشوائية الطبقية

في هذه الحالة ينبغي تقسيم المجتمع إلى أقسام أو طبقات مختلفة ثم يؤخذ من كل قسم أوطبقة عينة متجانسة بطريقة عشوائية، على أن يكون حجم الطبقة في العينة متناسب مع حجم الطبقة المناظرة لها في المجتمع الأصلي.

3-4-1-3- العينة العشوائية المنتظمة

يتم اختيارها من خلال تحديد مجتمع الدراسة ووضع أفراده في قائمة بشكل عشوائي وإعطاء كل منهم رقما، ثم يتم تحديد قاعدة الاختيار وفق قسمة حجم المجتمع على حجم العينة من أجل الحصول على طول الفترة، وبعد ذلك يتم انتقاء أحد الأرقام عشوائيا من بين الأرقام التي تساوي أو أقل من طول الفترة، ليتم اعتباره كعنصر أول من مفردات العينة ويشرع في إضافة طول الفترة له للحصول على المفردة الثانية، وهكذا نستمر في إضافة العدد الثابت إلى غاية الوصول إلى العدد الممثل لحجم العينة المطلوب.

وعليه، ووفقا لهذه الطريقة يتم الحصول على متتالية حسابية حدها الأول هو الرقم المختار عشوائيا في البداية، وأساسها هو طول الفترة، وعدد حدودها هو حجم العينة.[1]

3-4-1-4- العينة العشوائية العنقودية (المتعددة المراحل)

إذا كان المجتمع يتكون من أقسام متجانسة نبدأ باختيار بعض هذه الأقسام عشوائيا (كمرحلة أولى) ثم نختار عينة عشوائية بسيطة من كل قسم من الأقسام التي تم اختيارها (كمرحلة ثانية) وقد يحتاج الأمر إلى اختيار عينة عشوائية بسيطة من كل قسم من الأقسام التي تم اختيارها في المرحلة الثانية وهكذا...، والعينة التي تم اختيارها بهذه الطريقة تعرف بالعينة متعددة المراحل.

---

[1] -: فايز جمعة صالح النجار وآخرون، **أساليب البحث العلمي، منظور تطبيقي**، دار الحامد، عمان ، الأردن، 2009 ، ص 95.

3-4-2- العينات غير الاحتمالية

هي التي يتم اختيار مفرداتها بطريقة غير عشوائية، حيث يقوم الباحث باختيار مفردات العينة بالصورة التي تحقق الهدف من المعاينة.[1] و من أهم أنواع العينات غير الاحتمالية:

3-4-2-1- العينة القصدية

هي عينة يتم اختيار أفرادها بشكل مقصود ومستهدف لتوفر بعض الخصائص فيهم بما يخدم أهداف الدراسة، ويلجأ إلى هذا النوع من العينات عند توفر البيانات اللازمة للدراسة لدى فئة محددة من المجتمع الأصلي للدراسة. فعلى سبيل المثال عندما يريد الباحث أن يدرس سلوك المستهلكين لمنتجات مؤسسة ما، فإنه يحدد المستهلكين الحقيقيين لمنتجات تلك المؤسسة ويتجنب جمع البيانات من أشخاص لا يقبلون على انتقاء منتجاتها.

3-4-2-2- العينة العرضية

لا يخضع الباحث فيها عند اختياره لمفردات العينة لأي اعتبار وإنما يكون على سبيل المصادفة، كأن يشرع في توزيع الاستبيان على مجموعة من العاملين بمؤسسة ما وذلك بتسليم الاستمارة على من يلتقيه صدفة أثناء زيارته الميدانية أو كأن يقف أمام مدخل المؤسسة محل الدراسة ويختار الثلاثين الأوائل الذين يلتقي بهم عند دخولهم إليها.

وتمتاز هذه الطريقة بتوفير الوقت والتكاليف، كما يمكن من خلالها الحصول على معلومات موثوقة إذا ما كان المجتمع المستهدف بالدراسة على جانب كبير من التجانس، أما إذا كانت عناصر المجتمع غير متجانسة فذلك قد يؤدي إلى التحيز في اختيار العناصر المشكلة للعينة.

---

[1]- شرف الدين خليل، **مرجع سبق ذكره**، دون ذكر سنة النشر، ص: 13.

الفصل الثاني

طرق عرض البيانات

- ✓ عرض البيانات جدوليا
- ✓ الجداول التكرارية البسيطة
- ✓ الجداول التكرارية المزدوجة
- ✓ عرض البيانات بيانيا

الفصل الثاني: طرق عرض البيانات

سيتم التعرض في هذا الفصل إلى آلية عرض البيانات وذلك من خلال بيان آلية بناء الجداول التكرارية وطرق الحصول على التكرارات النسبية والتكرارات التراكمية. كما سيتم التعرض إلى عملية عرض البيانات بيانيا والأساليب المختلفة فيها، حيث يبقى الهدف من هذه العملية هو تقديم البيانات بطريقة مبسطة ومختصرة وهذا ليسهل فهمها واستنتاج بعض النتائج الأولية.

1- عرض البيانات جدوليا

عند توفر عدد كبير من البيانات يتطلب الأمر في كثير من الأحيان وضع القيم في جدول تكراري يلخص البيانات الإحصائية المدروسة بشكل يمكن من خلاله التعامل مع البيانات بقدرة وكفاءة أعلى. وذلك يتيح للباحث القدرة على التعمق في فهم البيانات الإحصائية بالإضافة إلى إمكانية إجراء تحليل إحصائي إستدلالي.

وتختلف الجداول الإحصائية باختلاف نوع البيانات من ناحية والغرض من الدراسة من ناحية أخرى لذلك يتم التمييز بين الحالات الموالية:

1-1- الجداول التكرارية البسيطة

يستخدم هذا النوع من الجداول لوصف وتلخيص البيانات التي تتعلق بظاهرة واحدة فقط سواء كانت كيفية أو كمية، حيث يتم تبويب البيانات من خلال تفريغها في جداول نهائية يحتوي كل منها على عمودين (سطرين). يبين العمود الأول قيم الظاهرة أو المتغير المدروس، وتكون هذه القيم على شكل قيم نقطية أو شكل مجالات، أما العمود الثاني فيحتوي على تكرارات هذه القيم أو المجالات.[1] كما يمكن تضمين أعمدة إضافية تحتوي معلومات تفصيلية عند الحاجة مثل بيان التوزيع النسبي أو التوزيع التراكمي.

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

---

[1]- جيلالي جلاطو، **مرجع سبق ذكره**، 2001، ص: 11.

2-1- 1- عرض بيانات المتغير الوصفي في الجداول التكرارية البسيطة

إذا كنا بصدد دراسة ظاهرة ما تحتوي على متغير وصفي واحد، فإنه يمكن عرض بياناته في شكل جدول تكراري بسيط، وهو جدول يتكون من ثلاث أعمدة، يخصص العمود الأول للصفات بعد ترتيبها إن كانت قابلة للترتيب والعمود الثاني يخصص لتفريغ البيانات فيما يخصص العمود الثالث للتكرارات، والمثال الآتي يوضح ذلك.

مثال (2-1): فيما يلي بيانات عن المستوى التعليمي لعينة من 50 فرد.

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

المطلوب: عرض البيانات في شكل جدول تكراري.

الحل:

المستوى التعليمي (ابتدائي- متوسط -ثانوي -جامعي -دراسات عليا) متغير وصفي ترتيبي، ويمكن عرض البيانات أعلاه في شكل جدول تكراري كالآتي:

الجدول رقم (2-1): توزيع الأفراد حسب المستوى التعليمي(متغيرة كيفية)

المستوى التعليمي	**العلامات**	التكرارات **ni**
ابتدائي	///// ///// ///	13
متوسط	///// ///// /////	15
ثانوي	///// ///// /////	15
جامعي	/////	5
دراسات عليا	//	2
المجيوع $\sum$	50	50

1-1-2- عرض بيانات المتغير الكمي المنفصل 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 ثلاثة أعمدة، الأول يحتوي على فئات تصاعدية للقراءات التي يأخذها المتغير، والثاني يخصص لتفريغ البيانات، في حين يشمل العمود الأخير التكرارات أو عدد المفردات التي تنتمي قراءاتها للفئة المناسبة لها. والمثال التالي يبين كيف يمكن عرض البيانات الكمية جدوليا.

مثال (2-2): يبين الجدول الآتي عدد الأطفال في العائلة لعينة مكونة من 42 عائلة

1	4	6	2	5	4	1
6	1	2	3	2	6	2
2	4	1	3	3	3	3
4	2	6	1	5	5	4
5	6	2	3	1	2	2
5	2	3	5	5	1	1

المطلوب: كون جدول التوزيع التكراري لعدد الاطفال في العائلة.

الحل:

الجدول رقم (2-2): توزيع الأسر حسب عدد الأطفال متغير كمي منفصل)

عدد الأطفال	**العلامات**	**عدد الأسر**
1	/// /////	8
2	///// /////	10
3	// /////	7
4	/////	5
5	// /////	7
6	/////	5
المجيوع $\sum$	/	42

1-1-3- عرض بيانات المتغير الكمي المتصل في شكل جدول تكراري بسيط

كما سبق الإشارة إليه، في المتغير الكمي المستمر يكون مجال الدراسة يضم مالا نهاية من القيم، ولتعذر وضع كل تلك القيم، يقسم مجال الدراسة إلى مجالات جزئية تسمى الفئات، حيث يحدد عدد هذه الفئات حسب حجم العينة وحسب توزيع الوحدات الإحصائية على مجال الدراسة. ولتكوين جدول التوزيع التكراري لهذا النوع من المتغيرات نتبع الخطوات الآتية:

1-1- 3-1- حساب المدى

ويمكن تحديد المدى R من خلال العلاقة الموالية:

المدى = أكبر قيمة – أصغر قيمة $R = X_{max} - X_{\min}$

1-1-3-2- حساب عدد الفئات

في هذا المجال وقصد تسهيل العملية وضع الإحصائي ستورجس (Sturges) قاعدة تجريبية لتحديد عدد الفئات وتعتمد هذه القاعدة على مجال الدراسة وحجم المجتمع.[1] وعليه، يتم حساب عدد الفئات من خلال العلاقة الآتية:

$$K = 1 + 3{,}322\log(n)$$

وينبغي الإشارة هنا إلى أن قاعدة ستورجس Sturges ليست إجبارية بل على الباحث اختيار طول الفئة المناسب والمبني على العديد من الاعتبارات منها رأي الباحث، الهدف من البحث، وحجم البيانات.[2]

1-1-3-3- حساب طول الفئات

يتم حساب طول الفئة من خلال العلاقة الموالية:

---

[1]- **نفس المرجع السابق**، ص، 12.

[2]- شرف الدين خليل، **مرجع سبق ذكره**، دون ذكر سنة النشر، ص: 18.

طول الفئة= $\frac{\text{المدى}}{\text{عددالفئات}}$ $L=\frac{R}{K}$

وبعد حساب طول الفئة يجب مراعاة تحقق المتباينة التالية:

طول الفئة × عدد الفئات ≤ المدى

وعلى أساس ما سبق من خطوات يتم تحديد حدود الفئات، حيث تكون بداية الفئة الأولى أصغر من أو تساوي أصغر قيمة في البيانات ونهاية الفئة الأخيرة أكبر من أكبر قيمة في البيانات. وفي الأخير يتم تحديد عدد القيم أو المشاهدات التي تقع في كل فئة على أن تكون لكل قيمة فئة واحدة فقط تنتمي إليها، **والمثال التالي** يبين كيفية التعامل مع البيانات الكمية المتصلة قصد إدراجها في جدول التوزيع التكراري.

مثال (2-3): يبين الجدول الآتي كمية الإنتاج اليومي لمدة 30 يوم في إحدى المؤسسات الصناعية:

35	34	44	33	45	36	34	33	33	35
45	44	33	37	34	40	41	43	35	34
39	33	41	42	41	40	37	34	38	36

المطلوب: تبويب البيانات في جدول تكراري حسب معادلة ستورجس (Sturges)

الحل:

كمية الإنتاج اليومي متغير كمي مستمر، ولكي يتم تبويب البيانات في شكل جدول تكراري، يتم اتباع الآتي:

- **حساب المدى:** المدى = أكبر قيمة – أصغر قيمة

$$R = X_{max} - X_{\min} = 45-33=\mathbf{12}$$

- حساب عدد الفئات حسب معادلة ستورجس (Sturges)

وذلك من خلال العلاقة الآتية:

$$K = 1 + 3{,}322 \log(n)$$
$$K = 1 + 3{,}322 \log(30) = 5.906 \approx 6$$

- حساب طول الفئة

وذلك من خلال العلاقة الموالية:

طول الفئة= $\frac{\text{المدى}}{\text{عددالفئات}}$ L= $\frac{R}{K}$ L= $\frac{12}{6}$ = 2

وعليه فإن طول الفئة هو 2

عند تحديد طول الفئة يجب مراعاة المتباينة التالية:

المدى ≤ عدد الفئات× طول الفئة

$$2 * 6 \leq 12$$

ومنه، صار بالإمكان تحديد الفئات، وذلك كما يلي:

- تبدأ الفئة بقيمة تسمى الحد الأدنى، وتنتهي بقيمة تسمي الحد الأعلى، ومن ثم نجد أن بداية الفئة الأولى تمثل أصغر من أو تساوي أصغر قيمة في البيانات. وعليه، تكون بداية الفئة الأولى هي 33.

- الحد الأعلى للفئة الأولى = الحد الأدنى + طول الفئة =33+2=35، إذا الفئة الأولى هي من 33 إلى 35 أي]33-35[.

- الحد الأدنى للفئة الثانية = الحد الأعلى للفئة الأولى 35+ طول الفئة 2=37. وبالتالي، **الفئة الثانية هي من 35 إلى 37 أي**]35-37[.

ونستمر بنفس الطريقة حتى يتم تكوين حدود الفئات الأخرى. وفي الأخير نقوم بتفريغ البيانات حسب توزيعها التكراري، مع التأكد من أن مجموع التكرارات يساوي عدد القيم.

الجدول رقم (2−3): توزيع المؤسسات حسب كمية الإنتاج اليومي (متغير كمي متصل)

كمية الإنتاج	العلامات	عدد المؤسسات
]35-33[	///// /////	10
]37-35[	/////	5
]39-37[	///	3
]41-39[	///	3
]43-41[	////	4
]45-43[	/////	5
المجيوع $\sum$	/	30

هذا، وتجدر الإشارة إلى أنه عند تكوين أي جدول للتوزيع التكراري، فإنه يمكن التمييز بين ثلاثة أنواع من التكرارات، وذلك كما يلي:

- التكرار المطلق وهو التكرار العادي.

- التكرار النسبي، الذي يستعمل للتعبير عن الأهمية النسبية لتكرار كل متغير أو فئة بالنسبة لإجمالي التكرارات، وهو يحسب بالصيغة الموالية:

التكرار النسبي= *تكرار الصفة* / *مجالتكارار ات*

- التكرارات التجميعية، وهذه الأخيرة تنقسم إلى قسمين التكرار المتجمع الصاعد، الذي يمثل مجموع الأفراد الذين تقل قيمتهم الإحصائية عن الحد الأعلى للفئة المقابلة، والتكرار المتجمع النازل، الذي يمثل مجموع الأفراد الذين تزيد قيمتهم الإحصائية عن الحد الأدنى للفئة المقابلة.

2-1- الجداول التكرارية المزدوجة

يستعمل جدول التوزيع التكراري المزدوج عند دراسة خاصيتين في آن واحد لمجتمع ما، حيث توضع البيانات الإحصائية في مثل هذه الجداول على الأشكال التي سنتعرض لها فيما بعد.

2- عرض البيانات بيانيا

بالإمكان وصف وتلخيص البيانات الإحصائية باستخدام الرسومات البيانية والأشكال الهندسية، إذ تمكن هذه الأخيرة من القيام بتحلي سريع للظاهرة المدروسة، وتستخدم أنواع مختلفة للعرض البياني حسب نوع المتغير المدروس.

2-1- العرض البياني في حالة متغير كيفي

2-1-1- العرض البياني للتكرارات البسيطة: هو عبارة عن أعمدة بسيطة تتناسب أطوالها مع التكرار المقابل لقيمة معينة للمتغير المدروس وتسمى الأعمدة البسيطة.

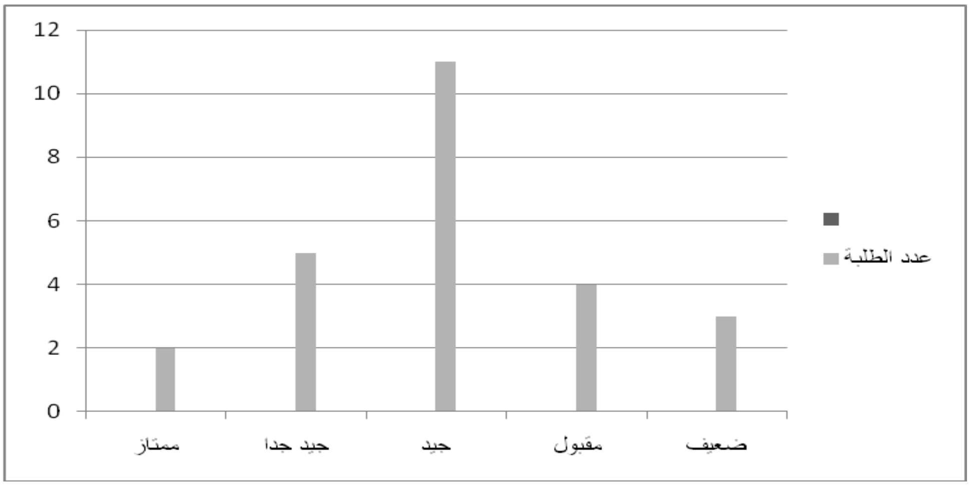
مثال: الجدول التالي يمثل تفريغ لبيانات معينة:

الرمز	cالتقدير	العلامات	F عدد الطلبة
A	ممتاز	\\	2
B	جيد جدا	\\\\\	5
C	جيد	\\\\\\ \\\\\	11
D	مقبول	\\\\	4
E	ضعيف	\\\	3
	$\sum$		25

- المطلوب: ما هي أفضل طريقة لعرض هذه البيانات؟

الحل:

يمكن عرض هذه البيانات بالأعمدة البيانية كما هو موضح أدناه:



كما يمكن عرض نفس البيانات في الدائرة النسبية، وذلك كما يلي:

حيث نستخرج نسب كل درجة (طول القوس) كما يلي:

- ممتاز= (2\25)× 360=28.8
- جيد جدا= (5\25)× 360=72
- جيد = (11\25)× 360=158.4
- مقبول = (4\25)× 360=57.6
- ضعيف = (3\25)× 360=43.2

نرسم الدائرة النسبية:

2-2- العرض البياني في حالة متغير كمي منفصل

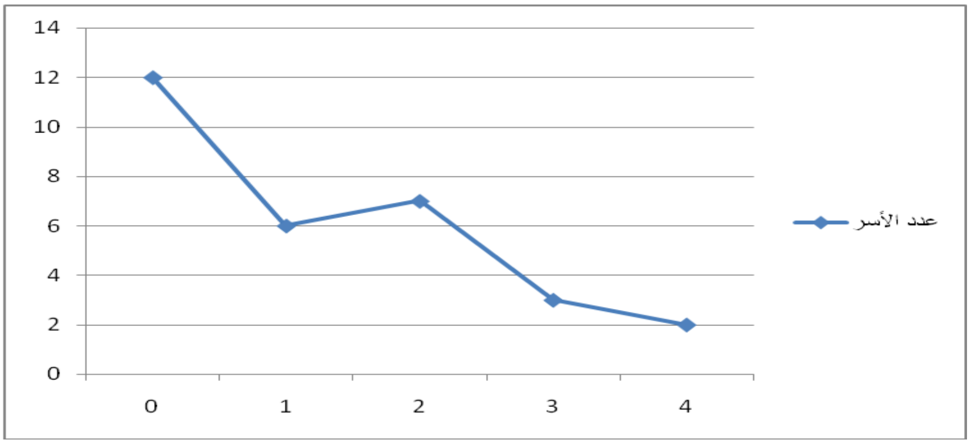
يمكن عرض هذا النوع من البيانات عن طريق الأعمدة البيانية، المنحنى التكراري، وذلك كما في المثال التالي:

مثال: الجدول التالي يمثل تفريغ لبيانات معينة:

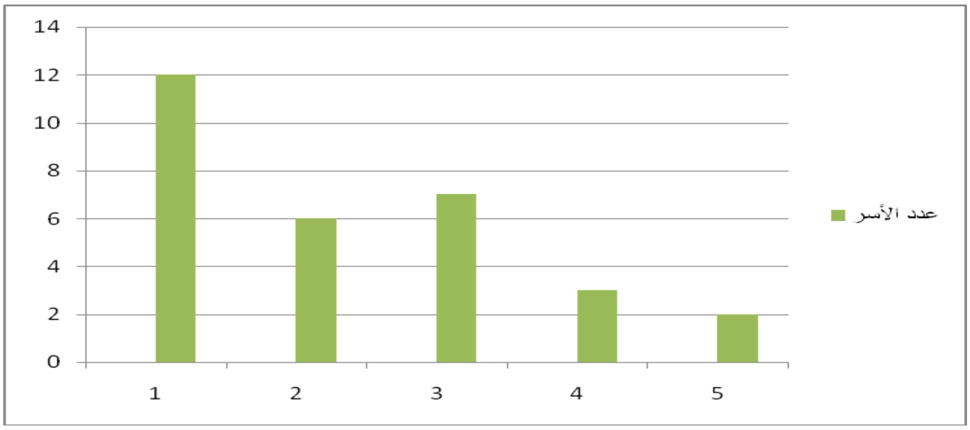
عدد الأسر f	العلامات	عدد الأطفال X
12	\\\\\ \\\\\ \\	0
6	\\\\\ \	1
7	\\\\\ \\	2
3	\\\	3
2	\\	4
30		$\sum$

- المطلوب: عرض البيانات في منحنى بياني وأعمدة بيانية.

الحل: المنحنى البياني يكون بالشكل الآتي:



أما الأعمدة البيانية فتأخذ الشكل الموالي:



2-3- العرض البياني في حالة المتغير الكمي المتصل

إن العروض البيانية للمتغير الكي المتصل من أكثر العروض البيانية استعمالا ومن أهمها:

2-3-1- المدرج التكراري

وهو عبارة عن مستطيلات(أعمدة)متلاصقة تمثل تكرارات أو قيم فئة من الفئات، حيث أن طول كل منها يتناسب مع التكرار المقابل، وقاعدة كل منها تساوي طول الفئة المقابلة، حيث توضع الفئات على محور السينات، بينما توضع التكرارات على محور العينات.

2 - 2 -المضلع التكراري

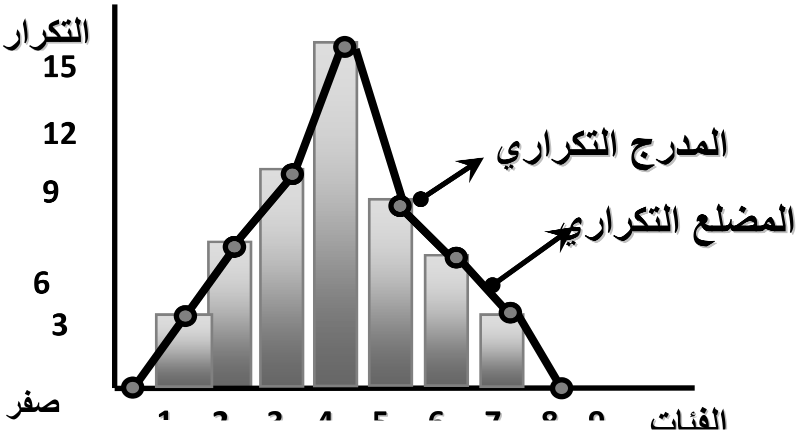
هو مجموعة من القطع المستقيمة المتصلة والمنكسرة تتحدد بنقاط إحداثياتها بمراكز الفئات والتكرارات المقابلة لها.

مثال: الجدول التالي يمثل تفريغ لبيانات معينة:

الفئات (c)	العلامات	التكرارf
10--20	\\\	**3**
20--30	\\\\\ \	**6**
30--40	\\\\\ \\\\\	**10**
40--50	\\\\\ \\\\\ \\\\\	**15**
50--60	\\\\\ \\\	**8**
60--70	\\\\\	**5**
70--80	\\\	**3**
Σ		50

- المطلوب: عرض البيانات ضمن منحنى تكراري، مدرج تكراري.

الحل:


التكرار
15
12
9
6
3
صفر
المدرج التكراري
المضلع التكراري
1 2 3 4 5 6 7 8 9
الفئات

الفصل الثالث

مقاييس النزعة المركزية

- ✓ الوسط الحسابي
- ✓ الوسط الهندسي
- ✓ الوسط التوافي
- ✓ الوسيط
- ✓ المنوال

الفصل الثالث: مقاييس النزعة المركزية

تسمى مقاييس الترعة المركزية بمقاييس الموضع أو المتوسطات، وهى القيم التى تتركز القيم حولها ومن هذه المقاييس، الوسط الحسابي ، والمنوال ، والوسيط ، والوسط الهندسي ، والوسط التوافقي والرباعيات ، والمئينات ، وفيما يلي عرض لأهم هذه المقاييس.

1- الوسط الحسابي

من أهم مقاييس الترعة المركزية ، وأكثرها استخداما في النواحي التطبيقية،[1] ويمكن حسابه للبيانات المبوبة وغير المبوبة ، كما يلي:

1-1- حالة البيانات غير المبوبة: (الغير مدرجة ضمن جدول تكراري)

يحسب من خلال العلاقة: $\bar{X} = \frac{\sum X}{n}$

حيث: x تمثل بيانات الظاهرة و n عدد قيم الظاهرة.

مثال: أحسب المتوسط الحسابي لعدد زبائن أحد المحلات التجارية لخمسة أيام متوالية: 50 ، 70 ، 60 ، 80 ، 90

المتوسط الحسابي: $\bar{X} = \frac{\sum X}{n} = \frac{50+70+60+80+90}{5} = 70$ بيستقبل المحل يوميا تقريبا 70 زبونا.

1-2- حالة البيانات المبوبة: (مدرجة ضمن جدول تكراري)

يتم التعبير عن كل قيمة من القيم التي تقع داخل حدود الفئة بمركز هذه الفئة (نعتبرها تمثل Xi) مع توفر التكرارات fi ، وبذلك يكون المتوسط الحسابي يمثل مجموع ضرب مراكز الفئات في التكرارات مقسوما على مجموع التكرارات.

[1]- شرف الدين خليل، **مرجع سبق ذكره**، دون ذكر سنة النشر، ص: 31.

مثال:الجدول التالي يبين توزيع دخل 40 موظف حسب مداخيلهم الشهرية ، أحسب متوسط دخل الموظف لهذه العينة

فئات الدخل	32-34	34-36	36-38	38-40	40-42	42-44
عدد الموظفين	4	7	13	10	5	1

الحل:

الفئاتC	التكراراتf	مراكز الفئاتX	$Xi.fi$
32-34	4	(32+34)/2=33	33 x 4=132
34-36	7	35	245
36-38	13	37	481
38-40	10	39	390
40-42	5	41	205
42-44	1	43	43
$\sum$	40	/	1496

2- المتوسط الهندسي

في حالات عدّة تكون قيم الظاهرة المدروسة عبارة عن نسب أو معدلات، وهي الحالات التي نرغب فيها بدراسة معدل تغيّر ظاهرة ما، هنا المتوسط الحسابي لا يصف هذه الظاهرة الوصف السليم ولا يعط أي فكرة صحيحة عن مثل هذه الظواهر، لهذا دعت الضرورة إيجاد متوسط آخر يصلح لوصف مثل هذه الظواهر يسمى بالمتوسط الهندسي.

المتوسط الهندس.

يعرف المتوسط الهندسي لمجموعة من القيم أو الأرقام $x1, x2$,.........xn والتي عددها n بالجذر النوني لحاصل ضرب هذه القيم.[1]

[1] – علي عبد السلام العياري و علي حسين العجيل ، **مرجع سبق ذكره**، ص5

2-1-المتوسط الهندسي للبيانات الأولية(غير المبوبة)

الوسط الهندسي هو عبارة عن الجذر من الرتبة n لنواتج ضرب المشاهدات ببعضها البعض، أي:

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

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

الحل:

$$G = \sqrt[4]{2 \times 4 \times 4 \times 8}$$
$$= \sqrt[4]{256}$$
$$= 4$$

:

2-2- المتوسط الهندسي للبيانات المبوبة

في البيانات المبوبة نستخدم المعادلة التالية:

$$G = \sqrt[n]{x_1^{f1} . x_2^{f2} \dots\dots\dots x_n^{fn}}$$

أو من معادلة اللوغاريتم :

$$\text{Log G.M} = (1/n)(\sum f_i \log x_i)$$

n= مجموع التكرارات f_i

f_i = تكرار كل فئة

x_i = مركز الفئة

مثال:

أوجد الوسط الهندسي لقيم جدول التوزيع التكراري :

حدود الفئة	مراكز الفئاتx_i	التكرارf_i	$\log x_i$	$f_i \log x_i$
50 - 59	54.5	3	1.7363	5.2091
60 – 69	64.5	5	1.8095	9.0477
70 – 79	74.5	2	1.8721	3.7443

80 - 89	84.5	1	1.9268	1.9268
90 – 99	94.5	4	1.9754	7.9017
المجموع		15		27.8296

الحل :

من الجدول السابق باستخدام معادلة اللوغاريتم :

$$\text{Log G.M} = (1/n)(\sum f_i \log x_i)$$

$$= (1/15)(27.8296) = 1.8553$$

وكما وضحنا في المثال السابق فإن :

G.M=71.66

3− المتوسط التوافقي

المتوسط التوافقي لمجيوعة من القيم هو مقلوب المتوسط الحسابي لمقاليب هذه القيم، وهو من المقاييس الخاصة التي تستخدم لتحديد معدلات السرعة ومتوسط الأسعار ومتوسط الكثافة السكانية.

4− الوسيط

هو أحد مقاييس الترعة المركزية، والذي يأخذ في الاعتبار رتب القيم ، ويعرف الوسيط بأنه القيمة التي يقل عنها نصف عدد القيم (50% من القيم أقل منه، 50 % من القيم أعلى منه).وفيما يلي كيفية حساب الوسيط في حالة البيانات غير مبوبة ، والبيانات المبوبة.

4−1− <u>حالة البيانات غير المبوبة:</u> (الغير مدرجة ضمن جدول تكراري)

لبيان كيف يمكن حساب الوسيط للبيانات غير المبوبة ، نتبع الخطوات التالية:

- نرتب القيم ترتيبا تصاعديا .

- في حالة عدد البيانات فردي رتبة الوسيط توافق القيمة $\frac{N}{2}$ ،

مثال : لدينا قيم المثال السابق: 50 ، 70 ، 60 ، 80 ، 90

بعد ترتيب القيم (50 60 70 80 90) الوسيط يوافق الرتبة الثالثة أي الوسيط هو :Me= 70

في حالة عدد البيانات زوجي رتبة الوسيط تقع بين القيمة $\frac{N}{2}$ و القيمة $\frac{N}{2}+1$، أي أن الوسيط يحدد من خلال العلاقة:

الوسيط = (القيمة رقم (n/2) + القيمة رقم (n/2)+1))) \ 2

مثال : مع اضافة اليوم السابع في المثال السابق: 50 ، 70 ، 60 ، 80 ، 90 ، 100

بعد ترتيب القيم (50 60 70 80 90 100)

الوسيط = (القيمة رقم (n/2)=70 + القيمة رقم (n/2)+1)=80) \ 2=70+80\2=75=Me

4-2- <u>حالة البيانات المبوبة:</u> يتم تحديد الوسيط باتباع الخطوات الأتية:

– نكون جدول التكرار المتجمع الصاعد و النازل ، ثم نحدد ترتيب الوسيط من العلاقة : $C1=\frac{\sum fi}{2}$

– نوجد قيمة الوسيط من العلاقة الآتية : $Me = L + \frac{C1-C2}{C3}.h$ (هذه العلاقة تستخدم في حالة الجدول الصاعد والذي سنكتفي به في حالة الحساب) .

حيث : L الحد الأدنى للفئة الوسيطية ، h : طول الفئة، C1 :ترتيب الوسيط المحدد سابقا.

C2: التكرار المتجمع الصاعد السابق للفئة الوسيطية، C3: تكرار الفئة الوسيطية.

مثال : من معطيات المثال الخاص بدخل الموظفين ، نحدد الوسيط كما يلي:

ت م الصاعد	مراكز الفئات X	التكرارات f	الفئات C
4	33	4	32-34
11	35	7	34-36
24	37	13	36-38
34	39	10	38-40
39	41	5	40-42
40	43	1	42-44
	/	40	$\sum$

نحدد ترتيب الوسيط من العلاقة : $C1 = \frac{\sum fi}{2} = 40/2 = 20$

الوسيط موجود ضمن الفئة $[36-38]$، L الحد الأدنى للفئة الوسيطية=36 ، h : طول الفئة=2.

C2: التكرار المتجمع الصاعد السابق للفئة الوسيطية=11، C3: تكرار الفئة الوسيطية=13.

- نوجد قيمة الوسيط من العلاقة الآتية : $Me = L + \frac{C1-C2}{C3}.h = 36 + \frac{20-11}{13}.2 = 37.38$

ويلاحظ أن قيمة الوسيط لا بد وأن تقع داخل حدود فئة الوسيط ، أي لا تقل عن بداية فئة الوسيط ولا تزيد عن نهايتها .

5- المنوال

تعبر قيمة المنوال عن المشاهدة الأكثر تكرارا فهو بمثابة القيمة الشائعة، وقد يكون للبيانات منوال واحد ويمكن أن يكون لها أكثر من منوال، كما يمكن أن لا يوجد منوال لمجموعة من البيانات. ويعتبر المنوال أفضل مقياس لوصف البيانات النوعية.

5-1- طرق حساب المنوال

تختلف طرق حساب المنوال في حالة البيانات الأولية منها في حالة البيانات المبوبة:

- في البيانات الأولية، المنوال هو القيمة أو الصفة الأكثر تكرارا مقارنة ببقية القيم أو الصفات، وعلى ضوء هذا التعريف فإن المنوال لمجموعة من البيانات قد لا يكون قيمة أو صفة واحدة.

مثال: البيانات التالية تمثل التقديرات التي تحصل عليها 10 طلاب:

ممتاز، جيد، جيد جدا، جيد، متوسط، فوق المتوسط، جيد، ضعيف، جيد جدا، جيد.

المطلوب :إيجاد المنوال لهذه البيانات ؟

الحل:

نلاحظ أن الصفة جيد هي الصفة الأكثر تكرار من بين الصفات وعليه فإن المنوال هو الصفة جيد

مثال:

متغير إحصائي يمثل عدد الغرف في البيت لعدد من العائلات القاطنة بحي معين، أوجد المنوال؟

3 2 5 3 3 4 5 2 4 3

الحل:

المنوال هو القيمة الأكثر تكرارا مقارنة ببقية القيم، وعليه نلاحظ أن القيبة 3 هي الأكثر تكرارا(العدد 3 يتكرر 4 مرات)، وبالتالي فإن المنوال هنا هو3 .

5-2- المنوال للبيانات المبوبة:

يمكن تحديده من خلال العلاقة: $D = L + \frac{\Delta 1}{\Delta 1 + \Delta 2}.h$

حيث : L الحد الأدنى للفئة المنوالية ، h : طول الفئة،

$\Delta 1$:الفرق بين أكبر تكرار و السابق له ، $\Delta 2$:الفرق بين أكبر تكرار و اللاحق له.

مثال : من معطيات المثال الخاص بدخل الموظفين ، نحدد المنوال كما يلي:

الفئات C	التكرارات f	مراكز الفئات X
32−34	4	33
34−36	7	35
36−38	13	37
38−40	10	39
40−42	5	41
42−44	1	43
$\sum$	40	/

الفئة ذات الأكبر تكرارا $[36-38]$،L الحد الأدنى للفئة المنوالية=36 ، h : طول الفئة=2.

$\Delta 1$:الفرق بين أكبر تكرار و السابق له=13−7 =6 ، $\Delta 2$:الفرق بين أكبر تكرار و اللاحق له=13−10=3

المنوال = $D = L + \frac{\Delta 1}{\Delta 1 + \Delta 2}.h = 36 + \frac{6}{6+3}.2 = 37.33$

الفصل الرابع

مقاييس التشتت

- ✓ المدى
- ✓ الانحراف المتوسط
- ✓ التباين
- ✓ الانحراف المعياري

الفصل الرابع: مقاييس التشتت

عند مقارنة مجموعتين من البيانات ، يمكن استخدام شكل التوزيع التكراري، أو المنحنى التكراري ، وكذلك بعض مقاييس النزعة المركزية ، مثل الوسط الحسابي والوسيط ، والمنوال ، والإحصاءات الترتيبية ، ولكن استخدام هذه الطرق وحدها لا يكفي عند المقارنة ، فقد يكون مقياس النزعة المركزية للمجموعتين متساوي ، وربما يوجد اختلاف كبير بين المجموعتين من حيث مدى تقارب وتباعد البيانات من بعضها البعض ، أو مدى تباعد أو تقارب القيم عن مقياس النزعة المركزية.

ومثال على ذلك ، إذا كان لدينا مجموعتين من الطلاب ، وكان درجات المجموعتين كالتالي:

88	67	85	81	78	70	63	**المجموعة الأولى**
77	74	75	78	77	78	73	**المجموعة الثانية**

لو قمنا بحساب الوسط الحسابي لكل مجموعة ، نجد أن الوسط الحسابي لكل منهما يساوي **76** درجة ، ومع ذلك درجات المجموعة الثانية أكثر تجانسا من درجات المجموعة الأولى . من أجل ذلك لجأ الإحصائيون إلى استخدام مقاييس أخرى لقياس مدى تجانس البيانات، أو مدى انتشار البيانات حول مقياس النزعة المركزية، ويمكن استخدامها في المقارنة بين مجموعتين أو أكثر من البيانات، ومن هذه المقاييس ، مقاييس التشتت ، والالتواء ، و التفرطح ، وسوف نركز في هذا الفصل على مقاييس التشتت .

1- **المدى:** هو أبسط مقاييس التشتت ، ويحسب المدى في حالة البيانات غير المبوبة بتطبيق المعادلة التالية .

المدى = أكبر قيمة – أصغر قيمة

$$R = X_{max} - X_{\min}$$

وأما المدى في حالة البيانات المبوبة له أكثر من صيغة، ومنها المعادلة التالية:

المدى = (الحد الأعلى للفئة الأخيرة) – (الحد الأدنى للفئة الأولى)

مثال

الجدول التكراري التالي يبين توزيع 60 مزرعة حسب المساحة المنزرعة بالذرة بالألف دونم .

المساحة	15-20	21-26	27-32	33-38	39-44	45-50
عدد المزارع	3	9	15	18	12	3

والمطلوب حساب المدى للمساحة المنزرعة بالذرة.

الحل:

المدى = (الحد الأعلى للفئة الأخيرة) – (الحد الأدنى للفئة الأولى)

المدى = 50 – 15 = 35 دونم

R = 50 – 15 = 35

2- الانحراف المتوسط

مقاييس التشتت هي مقاييس لقوة تجمع البيانات حول بعضها، ومن حيث أن التجمع يكون حول قيمة متوسطة، فإذا كان مقدار الاختلاف(الانحراف)بين القيم ومتوسطها كبيرا دلّ ذلك على عدم تجانسها والعكس صحيح.

يعرف الانحراف المتوسط بأنه المتوسط الحسابي للقيم المطلقة لانحرافات القيم عن متوسطها الحسابي ونرمز له بالرمز MD

$$MD = \frac{\sum |Xi - \overline{X}|}{n}$$

وهو يحسب وفقا لحالتين:

2-1-حالة البيانات غير المبوبة: (الغير مدرجة ضمن جدول تكراري)

بعد حساب المتوسط الحسابي يمكن حساب الانحراف المتوسط مباشرة.

مثال :مثلا ضمن القيم : 50 ، 70 ، 60 ، 80 ، 90 نحدد الانحراف المتوسط

نحدد أولا المتوسط الحسابي: $\overline{X} = \frac{\sum Xi}{n} = \frac{50+70+60+80+90}{5} = 70$

الانحراف المتوسط: $MD = \frac{\sum |Xi - \overline{X}|}{n} = \frac{|50-70|+|70-70|+|60-70|+|80-70|+|90-70|}{5} = 12$

2-2-حالة البيانات المبوبة:

في هذه الحالة يحسب الانحراف المتوسط من خلال العلاقة :

$$MD = \frac{\sum |Xi - \overline{X}|.fi}{fi}$$

حيث : Xi مراكز الفئات، fi التكرارات .

مثال:

يبين الجدول التكراري التالي توزيع 40 أسرة حسب الإنفاق الشهري بالألف الدولارات .

الإنفاق	2 - 5	6 - 9	10 - 13	14 – 17	18 – 21
عدد الأسرة	1	8	13	10	8

المطلوب: أوجد الانحراف المتوسط

<u>الحل:</u>

لحساب الانحراف المتوسط ، يتم تطبيق المعادلة الخاصة بالانحراف المتوسط في حالة البيانات المبوبة، ويتبع الآتي

- **تكوين جدول لحساب مكونات المعادلة:**

حدود الإنفاق	عدد الأسر f_i	مركز الفئة x_i	$x_i f_i$	الوسط الحسابي $\bar{x}$	$\lvert x_i - \bar{x} \rvert$	$\lvert x_i - \bar{x} \rvert \; f$
2 - 5	1	3.5	3.5	$\bar{x} = \frac{\sum x}{n} = \frac{524}{40} = 13.1$	**9.6**	**9.6**
6 - 9	8	**7.5**	**60**		**5.6**	**44.8**
10 - 13	13	**11.5**	**149.5**		**1.6**	**20.8**
14 – 17	10	**15.5**	**155**		**2.4**	**24**
18 – 21	8	**19.5**	**156**		**6.4**	**51.2**
sum	40		**524**			**150.4**

إذا الانحراف المتوسط هو :

$$MD = \frac{\sum \lvert x_i - \bar{x} \rvert \; f_i}{\sum f_i} = \frac{150.4}{40} = 3.76$$

3- التباين

هو عبارة عن المتوسط الحسابي لمربعات الفروق بين قيم المتغير الإحصائي ومتوسطها الحسابي،[1] ونستخدم مربعات الفروق هنا تفاديا لاستخدام القيم المطلقة كما هو الشأن في الانحراف المتوسط،

يرمز للتباين بالرمز $V(X)$ في حالة بيانات المجتمع وبالرمز s^2 في حالة بيانات العينة ويحسب حسب أنواع البيانات إن كانت مبوبة أو غير مبوبة وذلك كالآتي:

3-1- حسابه في حالة البيانات غير المبوبة

يحسب من خلال العلاقة: $V(x)=\frac{\sum (Xi-\overline{X})^2}{n}$ في حالة بيانات المجتمع، أما في حالة بيانات العينة فيحسب من خلال العلاقة الموالية: $V(x)=\frac{\sum (Xi-\overline{X})^2}{n-1}$

مثال: مثلا ضمن القيم : 50 ، 70 ، 60 ، 80 ، 90 نحدد التباين كما يلي:

نحدد أولا المتوسط الحسابي:

$$\overline{X}=\frac{\sum Xi}{n}=\frac{50+70+60+80+90}{5}=70$$

التباين:

$$V(x)=\frac{\sum (Xi-\overline{X})^2}{n}=\frac{(50-70)^2+(70-70)^2+(60-70)^2+(80-70)^2+(90-70)^2}{5}=\frac{1000}{5}=200$$

3-2- حسابه في حالة البيانات المبوبة

في هذه الحالة يحسب التباين من خلال العلاقة $V(X)=\frac{\sum (Xi-\overline{X})^2.fi}{fi}$

أما في حالة البيانات التي تخص العينة فتحسب من خلال العلاقة:

$$V(X)=\frac{\sum (Xi-\overline{X})^2.fi}{n-1}$$

---

[1]- مصطفى يوسف كافي وآخرون، **مرجع سبق ذكره**، ص:127 .

مثال : من معطيات المثال الخاص بدخل الموظفين ، نحدد التباين كما يلي:

الفئاتC	fi	Xi	$fi \times Xi$	$(Xi-\overline{X})^2$	$(Xi-\overline{X})^2 \times fi$
32-34	4	33	132	19,36	77,44
34-36	7	35	245	5,76	40,32
36-38	13	37	481	0,16	2,08
38-40	10	39	390	2,56	25,6
40-42	5	41	205	12,96	64,8
42-44	1	43	43	31,36	31,36
$\sum$	**40**	/	**1496**	**72,16**	**241,6**

المتوسط الحسابي : $\overline{X} = \dfrac{\sum_{i=1}^{n} Xi.fi}{\sum_{1}^{n} fi} = \dfrac{1496}{40} = 37.4$

التباين : $V(X) = \dfrac{\sum (Xi-\overline{X})^2.fi}{fi} = \dfrac{241.6}{40} = 6.04$

4- الانحراف المعياري

الانحراف المعياري هو الجذر التربيعي للتباين، ويعتبر من أهم المقاييس الإحصائية للتشتت وأكثرها استخداما في النظريات والقوانين الإحصائية، ويرمز له بالرمز ويرمز له بالرمز **S** أو σ .

و يحسب وفقا لحالتين

4-1- حالة البيانات غير المبوبة: (الغير مدرجة ضمن جدول تكراري)

من خلال التعريف السابق ، يحسب الانحراف المعياري بتجذير التباين:

$$\sigma = \sqrt{V(x)} = \sqrt{\frac{\sum (Xi-\overline{X})^2}{n}}$$

مثال: ضمن القيم : 50 ، 70 ، 60 ، 80 ، 90 نحدد الانحراف المعياري كما يلي:

التباين: $V(x) = 200$ الانحراف المعياري : $\sigma(x) = \sqrt{V(x)} = \sqrt{200} = 14.14$

4-2- حالة البيانات المبوبة

من خلال التعريف السابق ، يحسب الانحراف المعياري بتجذير التباين:

$$\sigma = \sqrt{V(x)} = \sqrt{\frac{\sum (Xi - \overline{X})^2 . fi}{\sum fi}}$$

مثال:من خلال المثال الخاص بمداخيل الموظفين ، الانحراف المعياري :

$$\sigma = \sqrt{V(x)} = \sqrt{\frac{241.6}{40}} = \sqrt{6.04} = 2.45$$

5- مقاييس التشتت النسبية

قد يتطلب الأمر في بعض الأحيان إجراء مقارنة بين تشتت مجموعتين أو أكثر من القيم المختلفة من حيث الوسط الحسابي أو أن قيم مفردات كل مجموعة مقاسه بوحدات قياس تختلف عن الأخر . و عندئذ فان مقاييس التشتت أيا كان سوف لن يكون نافعا لوحده في إجراء مقارنات من هذا النوع ، إنما يستوجب الأمر إيجاد مقياس تشتت أخر أكثر ملائمة لهذه الحالات ، هذا النوع من المقاييس تسمى بمقاييس التشتت النسبي وهي:

5-1- معامل التشتت المستند إلى الانحراف المتوسط

افرض ان **M.D(A)** تمثل الانحراف المتوسط المحتسبة على أساس نقطة اختيارية **(A)** وقد تكون (الوسط الحسابي أو الوسيط أو المنوال) عندئذ يعرف معامل التشتت للمتوسط على نحو الأتي:

$$C.D(A) = \frac{M.D(A)}{A}$$

5-2- معامل التشتت المستند إلى الانحراف المعياري (معامل الاختلاف)

افرض ان $\overline{X}$ يمثل الوسط الحسابي لمجموعة من القيم ، و أ، **(S)** يمثل الانحراف المعياري لها . عندئذ يعرف معامل الاختلاف و الذي نرمز له بالرمز **(C.V)** على النحو الأتي:

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

إن معامل الاختلاف يعتبر بحق أفضل معاملات التشتت الأنفة الذكر كونه يعتمد على أفضل مقاييس النزعة المركزية و أفضل مقاييس التشتت . إن هذا المعامل يوضح نسبة حصة كل وحدة من وحدات الوسط الحسابي و الانحراف المعياري و عليه عند إجراء مقارنة بين قيم مجموعتين ثم مقارنة معامل الاختلاف للمجموعة الأولى مع معامل الاختلاف للمجموعة الثانية ، عندئذ يقال عن المجموعة بأنها أكثر تجانسا إذا كان معامل الاختلاف اقل من الأخر.

الفصل الخامس

مقاييس الشكل

- ✓ مقاييس الالتواء
- ✓ العزم الأول
- ✓ العزم الثاني
- ✓ معامل التفرطح

الفصل الخامس: مقاييس الشكل

إذا كانت مقاييس النزعة المرمزية ومقاييس التشتت تسمح بتلخيص بيانات أي ظاهرة في صورة أرقام تعط فكرة عن خصائص توزيع هذه البيانات ودرجة تجانسها أو اختلافها، فإن هذا الوصف تبقى تنقصه الدقة المطلوبة للتعرف على خواص التوزيع خاصة فيما يخص انتشار البيانات على المنحنى البياني الممثل لها من حيث التوائه وتفرطحه عن الوضع الطبيعي، لذلك دعت الحاجة لاستخدام مقاييس أخرى لتحقيق هذا الغرض سميت مقاييس الالتواء والتفرطح.

1- مقاييس الالتواء

هناك طرق كثيرة لقياس الالتواء ومنها ما يلي:

1-1- العزم الأول

ويوصف من خلال العلاقة التالية :

$$\alpha_1 = \frac{3(\bar{X} - m_d)^3}{s}$$

مثال: حسب مقياس الالتواء (العزم الاول) للبيانات التالية مع تفسير النتائج

x_i	4	12	6	15	8	7	1

الحل : ان المتوسط يبلغ $(\bar{X} = 7.57)$ والوسيط $(m_d = 7)$ والتباين $(s^2 = 22.29)$ والانحرف المعياري يمثل جذر التباين وقيمته $(s = \sqrt{s^2} = \sqrt{(22.29)} = 4.72)$ وهي ادوات لحساب مقياس الالتواء التالي

$$sk_1 = \frac{3(\bar{X} - m_d)^3}{s} = \frac{3(7.57 - 7)^3}{4.72} = 0.1177$$

ومنه البيانات ذات التواء موجب

2-2- العزم الثاني

ويوصف من خلال العلاقة التالية

$$\alpha_2 = \frac{m_3}{s^3}$$

حيث ان s الانحراف المعياري و m_3 **العزم الثالث**، حول الوسط الحسابي ويحسب بالعلاقة التالية

$$m_3 = \frac{\sum_{i=1}^{n} f_i \ (X_i - \bar{X})}{\sum_{i=1}^{n} f_i}$$

2-3- العزم الثالث

العزم الثالث القياسي حول المتوسط يدعى معامل الالتواء لتوزيع المتغير العشوائي X ويحسب من خلال العلاقة التالية:

$$\alpha_3 = \frac{E(X-\mu)^3}{\sigma^3} = \frac{\mu_3}{\mu_2^{3/2}}$$

ويمثل مدى التماثل (symmtric_لدالة الكثافة الاحتمالية للتوزيع فاذا كانت قيمته مساوية للصفر فان دالة الكثافة الاحتمالية تعتبر متماثلة ، اما القيمة السالبة تشير الى التواء الى جهة اليسار اما القيمة الموجبة فتشير الى التواء نحو جهة اليمين ويقدر من خلال بيانات العينة بالصيغة التالية ولحالة البيانات غير المبوبة :

مثال: احسب معامل الالتواء للبيانات (10,24,34,46,50) .

الحل :

$$\bar{X} = \sum_{i=1}^{n} X_i/n = \frac{10+24+34+46+50}{5} = 32.8$$

$$s_x^2 = \frac{1}{n-1}\left\{\sum_{i=1}^{n} {x_i}^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2\right\} = \frac{1}{4}(6448-(146)^2) = 267.322$$

$$s_x = \sqrt{s_x^2} = \sqrt{267.322} = 16.35, \quad n = 5$$

ومن النتائج نكون الجدول التالي لايجاد بسط القانون

xi	xi-32.8	(xi-32.8)/16.35	((xi-32.8)/16.35)**3
10	-22.8	-1.39450	-2.71176
24	-8.8	-0.53823	-0.15592
34	1.2	0.07339	0.00040
46	13.2	0.80734	0.52622
50	17.2	1.05199	1.16421
			-1.17685

ومن ذلك فان معامل الالتواء يحسب بالاتي

$$sk = \frac{n}{(n-1)(n-2)}\sum\left[\frac{(X_i-\bar{X}\,)}{s}\right]^3 = \frac{5}{4\times 3}\,(-1.17685) = -0.49$$

وفي حالة البيانات المبوبة يحسب بالعلاقة التالية

$$sk = \frac{\sum_{\mathrm{i}=1}^{n}(m_i-\bar{X}\,)^3 \mathrm{f}_i}{\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{f}_i}$$

2- معامل التفرطح ($\boldsymbol{kurtosis}$)

ان العزم الرابع القياسي حول المتوسط يدعى معامل التفرطح ويقيس مدى اختلاف شكل التوزيع للبيانات عن التوزيع الطبيعي حيث ان القيمة الموجبة للمعامل تدعى $sharper\ peak$ للتوزيع الطبيعي والقيمة السالبة للمعامل تعني ان توزيع البيانات يمتلك $flatter\ peak$ عن التوزيع الطبيعي وصيغته

ويقيس ذيل دالة كثافة الاحتمال ، ويقدر من بيانات العينة ذات الحجم n بالصيغة التالية

ku

اوبالصيغة المعدلة التالية

$kum =$

مثال: لبيانات المثال السابق احسب معامل التفرطح .

الحل:

حيث ان المتوسط يمكن ان يحسب بسهولة ويساوي

وبالاعتماد عليها وعلى صيغة الحساب للمعامل نكون الجدول التالي

xi	xi-32.8	(xi-32.8)/16.35	((xi-32.8)/16.35)**3	((xi-32.8)/16.35)**4
10	-22.8	-1.39450	-2.71176	3.78154
24	-8.8	-0.53823	-0.15592	0.08392
34	1.2	0.07339	0.00040	0.00003
46	13.2	0.80734	0.52622	0.42484
50	17.2	1.05199	1.16421	1.22474
			-1.17685	5.51500

ان صيغة المعامل المعدلة

$kum =$

وان المعامل وفق الصيغة غير المعدلة يبلغ $(ku = -1.621)$ والقيمة السالبة للمعامل تعني ان توزيع البيانات تمتلك قمة ($flatter\ peak$) بالمقارنة مع التوزيع الطبيعي .

وللبيانات المبوبة يحسب بالعلاقة

حيث ان m_i مركز الفئة i .

قائمة المراجع

قائمة المراجع

1- مصطفى الخواجة، **مقدمة في الإحصاء**، الدار الجامعية، الإسكندرية، 2002.

2- محمد راتول، **الإحصاء الوصفي**، ديوان المطبوعات الجامعية، الجزائر، الطبعة الثانية، 2006.

3- جيلالي جلاطو، **الإحصاء مع تمارين ومسائل محلولة**، ديوان المطبوعات الجامعية، الجزائر، 2002 .

4- وليد اسماعيل السيفو وآخرون : **أساسيات الأساليب الإحصائية للأعمال**، زمزم ناشرون وموزعون، الأردن، الطبعة الأولى 2010.

5- شرف الدين خليل، **الإحصاء الوصفي**، شبكة الأبحاث والدراسات الاقتصادية، WWW.RR4EE.NET.

6- خالد أحمد فرحان المشهداني ورائد عبد الخالق عبد الله العبيدي، **مبادئ الإحصاء متضمن التحليل الإحصائي SPSS** ، دار الأيام، الأردن، 2013.

7- عامر إبراهيم قنديلجي، **منهجية البحث العلمي**، دار اليازوري، عمان ، الأردن، 2012 .

8- :فايز جمعة صالح النجار وآخرون، **أساليب البحث العلمي، منظور تطبيقي**، دار الحامد، عمان ، الأردن، 2009 .

9- علي عبد السلام العياري وعلي حسين العجيل ، **الإحصاء والاحتمالات النظرية والتطبيق**"، منشورات ELGA، مالطا، 2000.

10- مصطفى يوسف كافي وآخرون **الإحصاء في الإدارة والاقتصاد**، مكتبة المجتيع العربي، الأردن، الطبعة الأولى ، 2012.