جـامعــــة الجـزائـــر 3
كليـــة العلـوم الاقتصاديـــة والتجاريـــة وعلوم التسييـر

محاضرات في مقيـاس الإحصـاء 1

مطبوعة موجهة لطلبة السنة الأولى LMD جذع مشترك
العلوم الاقتصادية والتجارية وعلوم التسيير

من إعداد د: محمدي صبيحة

السنــة الجامعيــة 2022/2021

المقدمــة

بسم الله الرحمن الرحيم

الحمد لله رب العالمين فاتحة كل خير وتمام كل نعمة، والصلاة والسلام على أشرف المرسلين سيدنا وحبيبنا وشفيعنا محمد صلّ الله عليه وسلم وعلى آله وصحابته أجمعين ومن تبعه إلى يوم الدين.

المقدمـة

الحمد لله الذي أعانني على تقديم هذه المطبوعة لطلبة السنة الأولى جذع مشترك علوم اقتصادية وتجارية وعلوم التسيير، والتي جاءت بمحاور منسجمة مع البرنامج المعتمد من طرف وزارة التعليم العلي والبحث العلمي، حيث تشمل هذه المطبوعة على مجموعة من المحاضرات في الإحصاء الوصفي، هدفها اكتساب الطالب قاعدة أساسية لعلم الإحصاء الذي يجعله يعتمد على نفسه في الدراسة والبحث خاصة لما يدرس بعده فيما يتعلق الإحصاء الرياضي ثمّ التطبيقي.

يعدّ علم الإحصاء أحد الأساليب العلمية الشائعة الاستخدام في مختلف نواحي العلوم، إذ يُعتبر الأداة الرئيسية للتعبير الكمي عن مختلف الظواهر الاقتصادية والاجتماعية، ومن هنا جاءت أهمية هذه المطبوعة في تعريف الطالب استعمال المفاهيم الإحصائية والتقنيات الإحصائية الأساسية من خلال تعريف الإحصاء، البيانات الإحصائية وكيفية عرضها في جداول وتمثيلها بيانيا والتوزيعات التكرارية أو ما يعرف بالتلخيص الرقمي للبيانات، وكذا كيفية استعمال مقاييس النزعة المركزية بما فيها الوسط الحسابي، الوسيط، المنوال، ومتوسطات الأخرى، في الأخير التطرق إلى مقاييس التشتت والشكل بما فيها التباين والالتواء.

لقد حرصت على أن أتّبع كل تعريف أو مفهوم تعرضت إليه بأمثلة تساعد على فهمه وذلك تسهيلا للطلبة فهم وإدراك مضامين تلك التعاريف أو المفاهيم أو النظريات، مراعية التبسيط ما استطعت.

تتكون هذه المطبوعة من محاضرات ودروس موزعة عبر خمسة فصول، حسب برنامج وزارة التعليم العالي والبحث العلمي، حيث تناولت في الفصل الأول إلى علم الإحصاء ومفاهيمه العامة، حيث تطرقنا إلى نبذة تاريخية لعلم الإحصاء ثم تعريف علم الإحصاء وطبيعته وأهدافه وأقسامه مجالات استعمالاته، كما تطرقنا إلى بعض المصطلحات كالوحدة الإحصائية والمجتمع الإحصائي والعينة والمتغيرات العشوائية الكيفية والكمية ثمّ اتبعتها بفصل ثان يشمل جمع وعرض المعلومات الإحصائي والذي يبين مصادر وطرق جمع البيانات الإحصائية وكيفية عرضها، كما تضمن كيفية إعداد الجداول التكرارية، ويليه الفصل الثالث والذي يعدّ أطولهم على الإطلاق الذي خصّص لمقاييس النزعة المركزية، حيث تطرقنا لكل من الوسط الحسابي ومشابهاته من الوسط الهندسي والوسط التوافقي والوسط ألتربيعي ، والوسيط والمقاييس الشبيهة به، وكيفية تحديدها حسابيا في كلتا الحالتين في حالة البيانات المبوبة وغير المبوبة وكذا تحديدها بيانيا. أما الفصل الرابع فقد خصّص لمقاييس التشتت المطلقة والنسبية، ثم فصل أخير والذي خصّص بدوره لمقاييس الشكل بما فيها العزوم لما لها من علاقة بكل المقاييس سواء مقاييس النزعة المركزية أو مقاييس التشتت أو مقاييس الشكل. والحمد لله.

الفصل الأول

علم الإحصاء ومفاهيمه العامة

الفصل الأول: علم الإحصاء ومفاهيمه العامة

تمهيد:

إستخدم الإحصاء منذ العصور الأولى لنشوء الحضارات الإنسانية، وكان الهدف منه هو العدّ والحصر، إذ يُعد أداة رئيسية للتعبير الكمي عن مختلف الظواهر الاقتصادية والاجتماعية، يستعمل في القياس، التحليل والتنبؤ، ويحتلّ تدريس الإحصاء بمختلف مستوياته الوصفي، الرياضي والتطبيقي مكانة مميزة في مختلف برامج التكوين في مجال العلوم الاقتصادية والتجارية وعلوم التسيير. كما برزت الحاجة إلى استخدامه مع نشوء الحروب والأزمات المختلفة، فاستعمل في عدّ السكان وتحديد أو حصر الولادات والوفيات وكذا عدد المؤهلين لحمل الأسلحة، ولدفع الضرائب فضلا عن حصر الإنتاج الزراعي لغايات فرض الضرائب.

وفي القرن 20 ومع تطور الحياة البشرية تطور علم الإحصاء مفهوما واستخداما، حيث أصبح علما مستقلا بذاته وأداة مهمة في إجراء البحوث والتجارب في جميع العلوم سواء كانت طبيعية، تكنولوجية أو اجتماعية.

البوادر الأولى لهذا العلم ترجع إلى الزمن القديم عند الإنسان البدائي لما تحوّل من حياة التنقل إلى حياة الاستقرار، ومع هذا الاستقرار نتج مفهوم الاحتلال للمكان أي احتلال قطعة من الأرض واعتبارها مجالا خاصا، ثم اهتم بعدد الأشجار الموجودة في مجاله، عدد أفراد الخلية العائلية الموجودة في هذا المجال، وبعدد الحيوانات التي يتمكن من ترويضها لأغراض العمل.

أما البوادر العلمية للإحصاء كالنظرية لم تظهر إلا في القرن، 18 بحيث توجه الباحثون الرياضيون أمثال برنولي (bernouly) قوس (gauss) و لابلاس (laplace) نحو التحليل الإحصائي وإنشاء القوانين الاحتمالية، ولم يكتمل الإحصاء كعلم لجمع وعرض وتحليل واستخدام البيانات الإحصائية بغرض الاستدلال واتخاذ القرارات إلا في بداية القرن 20، الذي تحوّل إلى علم مستقل بذاته وأداة مهمة في إجراء البحوث والتجارب في جميع العلوم سواء كانت طبيعية أو تكنولوجية أو اجتماعية.

أولا: نبذة تاريخية لعلم الإحصاء

إن أول من استخدم الإحصاء هم قدماء المصريين من الفراعنة، إذ كانوا أول من قام باستخدام وتطبيق المفاهيم البدائية للإحصاء في بناء الأهرامات وتصميمها.

وفي عصر الخلافة العباسية، استخدم الخليفة العباسي المأمون فكرة الحصر والعدّ للوقوف على عدد السكان ومعرفة مقدار الزكاة.

وفي صدر الإسلام، يعتبر الأسلوب الذي كان الخليفة عمر ابن الخطاب رضي الله عنه يستخدمه لتقدير عدد المقاتلين من خلال معلومة عن عدد أرغفة الخبز المستهلكة إلا أفكارا تتلاغم كليا مع أسلوب (معلومات المتغير المساعد) والمستخدم حاليا في أساليب التقدير لمتغيرات يصعب أخذ المعلومات عنها في المعاينة، هذا بالإضافة لإجراء إحصاء عام وتدوين الدواوين في عصر خلافته[1].

أما في القرآن الكريم، فقد وردت كلمة الإحصاء في عدة آيات كريمات نذكر منها:

- في سورة مريم (الآية 94) قال سبحانه وتعالى: * لقد أحصاهم وعدهم عدا*. صدق الله العظيم بمعنى علو عددهم وأكد ذلك تأكيدا.
- في سورة إبراهيم (الآية 36)، يقول سبحانه وتعالى: *أتاكم من كل ما سألتموه وإن تعدوا نعمت الله لا تحصوها إن الإنسان لظلوم كفارٌ). صدق الله العظيم

فكلمة لا تحصوها تعني لا تطيقوا عدها ولا تقوموا بحصرها لكثرتها، فهذا دلالة على أن الإنسان عاجز عن إجراء المسح الشامل لمجتمع نعم الله له، لذا لا بد من دراسة عينة من هذا المجتمع.

- وفي سورة الكهف جاءت كلمة أحصاها بمعنى أحاط بها وجمعها وعلمها، أي أن الله سبحانه وتعالى قادر على حصر المخلوقات وعدها وعلم بكل صغيرة وكبيرة متعلقة بها.

أما في العصور الحديثة فملامحه الأولى برزت خلال القرن 16 الذي شهد بدايات معروفة في مجال الاحتمالات (الإحصاء التحليلي) بسبب انتشار لعب القمار بأوروبا حيث اتجه المقامرون

[1] سهيل أحمد سمحان، محمود حسين الوادي، مبادئ الإحصاء، دار الصفاء للنشر والتوزيع، الأردن، عمان، 2010، ص 18.

إلى علماء الرياضيات لتزويدهم بمعلومات خاصة حول فرص الربح والخسارة (ومن هؤلاء بسكال (pascal) و برنولي (bernouli) مؤسسي نظرية الاحتمالات).

كما قدم كردانو CARDANO (1501− 1571) بعض الأفكار في نظرية الاحتمالات المرتبطة برمي زهرة النرد (الطاولة)، كما ظهر بعض الطرق الإحصائية بأبعادها النظرية والتطبيقية.

وخلال القرن 17 ظهرت ولادة الإحصاء الحيوي (VITAL STATISTICS) من قبل العالم الإنجليزي JHON GRAUNT (1620− 1674) الذي كان أول من درس موضوع إحصاء الولادات والوفيات حسب توقعات البقاء للمرضى وبأعمار مختلفة الذي أطلق عليه اسم إحصاء السكان.

وخلال القرن 17 أيضا، وضع كل من PASCAL (1623 −1662) و JAMES BERNOULI (1654 −1705) حجر الأساس لأهم نظريتين هما:

1 −نظرية الاحتمالات(PROBABILITY THEORY) التي تعد العمود الفقري لعلم الإحصاء.

2 −نظرية الألعاب (المباريات) [2](GAMES THEORY).

أما في القرن 18، ظهرت نظرية قياس الخطورة عام 1735 على يد DANIEL BERNOULI (1700− 1782) المبنية على التوقع الرياضي.

وشهد القرن 19 ظهور طرق إحصائية متطورة في مجال علم النفس والاقتصاد، إذ قدم FISHER FECHNER (1801− 1877) إسهامات واسعة في مجال تصميم التجارب بعلم النفس، كما ظهر في نفس الفترة تطبيق جديد للطرق الإحصائية في العلوم الفيزيائية يتناول نظرية الغازات والتي تطورت فيما بعد إلى ميكانيكية.

وفي عام 1860، قام MAXWELL باستخدام منحنى التوزيع الطبيعي (DISTRIBUTION NORMAL)

[2] حسن ياسين طعمة، إيمان حسين حنوش، طرق الإحصاء الوصفي، دار صفاء للنشر والتوزيع، عمان، 2009، ص 22.

ولقد كتب العالم الرياضي GAUSS (1777– 1855) حول موضوعات الإحصاء خاصة تلك المبادئ الأساسية لطريقة المربعات الصغرى وكذلك المنحنى الطبيعي للأخطاء.

وفي عام 1833 اكتشف DE MOIVRE معادلة التوزيع الطبيعي (DISTRIBUTION NORAMAL) الذي يعتمد على الإحصاء التحليلي الذي يسمى هذا المنحنى بمنحنى (GAUSS) الذي اشتق معادلته أثناء دراسته للخطأ الناتج من القياسات المتكررة لنفس الكمية.

كما أن دراسات لابلاس (LAPLACE) أدت إلى نفس النتائج إضافة إلى تطبيقه للإحصاء في مجال الفلك.

كما تم تطبيق الإحصاء من قبل (CHARLES TYELL) في مجال الجيولوجيا، وفي المجال البيولوجي من قبل (CHARLES DARUIN) ومن قبل الأخصائي مندل (MANDEL).

كما تم تطبيق الإحصاء في علم الاجتماع وفي مجال الوراثة، في حين كان العالم الإنجليزي جالتون (F.GALTON) (1822– 1921) أول من كتب في موضوع الانحدار (REGRESSION) وتم استخدامه في التطبيقات البيولوجية الذي اشترك مع عالم الرياضيات كارل بيرسون (PEARSON) في إيجاد نظرية الارتباط (CORRELATION) والانحدار (REGRESSION) وإليه تعود أساسيات نظرية المعاينة.

أما أشهر علماء القرن 20 فهو العالم (FISHER) (1890– 1962) الذي طور علم الإحصاء وطبقه في علوم كثيرة كالزراعة والوراثة والاقتصاد ووضع أسس تصميم وتحليل التجارب، فهو يعد أبو الإحصاء ورائدا بارزا في تقديم المفاهيم الأساسية لموضوع التقدير النقطي وموضوع توزيعات المعاينة وموضوع اختيار الفروق باستخدام تحليل التباين.

والجدير بالذكر أن التطور في علم الإحصاء مازال مستمرا حيث بدأ يدرس كموضوع مستقل سنة 1950 في أمريكا وبدأت الجامعات الأمريكية بمنح الدرجات الجامعية العالية في الإحصاء سنة 1955، وبشكل عام فإن العوامل التي ساعدت على تطور علم الإحصاء هي:

- الحاجة المتزايدة للبيانات الإحصائية ومن قبل مختلف العلوم.
- قلة تكاليف الدراسات الإحصائية مقارنة بغيرها من الدراسات من العلوم الأخرى.

أما الاحتمالات فلم تشهد نفس التسارع في التطور كما في الإحصاء وذلك بسبب تعارضها مع المعتقدات الدينية.

وفي الأخير، وفي ضوء ما تقدم، يتضح بأن الإحصاء هو علم كسائر العلوم الأخرى، له نظرياته وقوانينه وأساليبه، ويعد من العلوم المهمة التي تهتم بوصف طرق متعددة في جمع البيانات وتخليصها(تصنيفها وتبويبها) وعرضها ومن تم تحليلها إحصائيا بهدف استخلاص النتائج المطلوبة وتفسيرها وتعميمها.

ثانيا: علم الإحصاء

1/ تعريف الإحصاء:

يعود الإحصاء إلى الكلمة اللاتينية (status) التي تعني الحالة أو الوضع السياسي للدولة بلغة الأرقام أي هو علم جمع وترتيب المعلومات لظاهرة معينة ثم تحليلها لإيجاد حلول مناسبة لمشاكل الحاضر أو للتنبؤات في المستقبل، إذن هو علم له نظرية وهدف ومنهجية، لذا له عدة تعاريف مختلفة اختلفت من حيث المضمون والشمولية، نذكر أهمها:

التعريف 1: الإحصاء هو علم التقديرات والاحتمالات كما أنه علم ابتكار وتطوير وتطبيق أساليب تمكننا من التقليل من عدم التأكد والعشوائية للاستدلال الإحصائي.

التعريف 2: الإحصاء هو العلم الذي يختص بجمع وتصنيف الحقائق العددية، تنظيمها وتلخيصها وعرضها وتحليلها بغية الوصول إلى نتائج مقبولة تؤدي إلى اتخاذ قرارات سليمة.

التعريف 3: الإحصاء هو ذلك العلم العملي الذي يسجل بطرق مختلفة واقع الأمر لظاهرة أو متغير معين، وتجسد هذا الواقع في قالب رقمي قياسي. كما أنه يختص بجمع وتحليل وتفسير البيانات العددية.

التعريف 4: الإحصاء هو العلم الذي يوفر جملة من المبادئ والقوانين والطرق العلمية المختلفة لجمع البيانات وتبويبها وتلخيصها وعرضها وتحليلها، ثم استخلاص الاستنتاجات للتوصل إلى

أحسن القرارات لحل المشاكل المختلفة على أساس من التحليل العلمي للبيانات المتاحة، نلاحظ أن الاحصاء يمتاز بالوراحل الاربعة وهي[3]:

1. المشاهدة والملاحظة؛
2. الفرضية: وهي التفسير الأولي للظاهرة؛
3. التنبؤ: وهي المعرفة الجديدة التي استنتجها الباحث؛
4. التحقيق: وهي مرحلة التأكد من صحة الفرضية والتي تم تفسير الظاهرة بها.

التعريف 5: هو مجموعة من الإجراءات والأساليب والتقنيات والقوانين النظرية المستخدمة بطريقة منهجية مضبوطة في جمع وعرض وتحليل البيانات الإحصائية حول ظاهرة ما، ثم استعمالها في اتخاذ القرارات أو القرارات المستقبلية حتى في ظروف عدم التأكد أو نقص المعلومات وذلك بفضل تطوير فرع الإحصاء الاستدلالي.

وتأسيسا على ما تقدم، يكون التعريف الأشمل والأمثل لعلم الإحصاء بأنه:

" الطريقة العليمة التي تحكم عملية جمع البيانات والحقائق عن ظاهرة أو فرضية معينة، وتنظيم وتبويب هذه البيانات والحقائق بالشكل الذي يسهل عملية تحليلها وتفسيرها، ومن ثم استخلاص النتائج واتخاذ القرارات في ضوء ذلك، ولتنبؤ بما ستؤول إليه الظاهرة في المستقبل"[4].

2/ طبيعة علم الإحصاء وأهدافه:

في الماضي، كانت كلمة الإحصاء تهدف إلى العدّ والحصر، كما أنها استخدمت للدلالة على أعمال وحسابات الدول في الحروب ولجباية الضرائب وتعداد السكان والقادرين على حمل السلاح، أما في العصر الحديث فقد تطور الإحصاء وأصبح علما قائما بحدّ ذاته يمرّ بأربعة مراحل والتي تتميّز بها كافة العلوم كالمشاهدة، الفرضية، التنبؤ والتحقيق. وتتمثل أهدافه العامة فيما يلي:

[3] نداء محمد الصوص، مبادئ الإحصاء، دار الأجنادين للنشر والتوزيع، ط 1، الأردن عمان، 2007، ص 09.

[4] حسين ياسين طعمه، إيمان حسين حنوش، طرق الإحصاء الوصفي، المرجع سبق ذكره، ص 25.

- تبسيط البيانات الكثيرة والمعقدة بعرضها في جدول أو رسوم بيانية والتعبير عنها ووصفها بأرقام مبسطة يسهل فهمها.
- وضع الحقائق في صورة عامة واضحة باستعمال الأرقام التي توضح الحقائق أكثر مما توضحه جملة عادية كنسبة النجاح في البكالوريا في الجزائر سنة 2021 هي (54%) وفي سنة 2020 كانت (64%).
- مقارنة المتغيرات المختلفة وإيجاد العلاقة بينهما، وتمثيل هذه العلاقة بنماذج رياضية.
- تمكين الباحثين في فروع العلوم المختلفة من اتخاذ القرارات المناسبة بقدر كبير من الصحة اعتمادا على البيانات المتاحة.
- يساعد في عملية التخطيط لأنه يستخدم في عملية التنبؤ ببيانات مستقبلية.
- تساعد الطرق الإحصائية في عملية التصنيف، كتصنيف الطلبة حسب أعمارهم ومعدلاتهم.
- تساعد الطرق الإحصائية في عملية التقييم الموضوعية لكل من الطالب والأستاذ.
- تعتبر الطرق الإحصائية من أدوات البحث الهامة والموضوعية وخاصة عند قياس الاتجاهات في بحوث الدراسات الوصفية، وكذلك عند إقامة التجارب الميدانية.
- تهدف الطرق الإحصائية إلى إيجاد أدوات مساعدة في عملية التنبؤ واتخاذ القرارات المناسبة.

3/ مراحل البحث الإحصائي:

من خلال التعريف السابق تبين أن الإحصاء يشمل عدة مراحل يجب على الباحث أن يتبعها في البحث الإحصائي، حيث يبدأ بالتحديد الدقيق للهدف الإحصائي الواضحة للبحث بما ينسجم وطبيعة البيانات المطلوب جمعها، ومن ثم القيام بتجميع هذه البيانات وفق الطريقة المناسبة والمصادر المتاحة لها، ثم عرض البيانات وتصنيفها بشكل يظهر العلاقة بينهما، ثم

تأتي مرحلة التحليل حيث يقوم الباحث بدراسة المعلومات الإحصائية وتحليلها ومن ثم تفسيرها واتخاذ القرارات المناسبة وفق النتائج المتحصل عليها[5].

4/ أقسام علم الإحصاء:

يميل بعض الإحصاء إلى تقسيم علم الإحصاء إلى قسمين رئيسين هما:

أ/ **الإحصاء الوصفي:** وهو يهتم بجمع البيانات وتصنيفها (تخليصها) وإمكانية عرضها جدوليا أو بيانيا، فيمكن القول بصورة عامة أن الإحصاء الوصفي يهدف إلى تصنيف البيانات وإعطاءها وصفا بسيطا للمقاييس والرسوم البيانية، بمعنى آخر هو اختصار عدد كبير جدا من البيانات الإحصائية في عدد محدود من الأرقام تسمى المقاييس الإحصائية (paramètres ou indicateurs)، أو في جدول إحصائي سهل للقراءة أو في رسوم بيانية، والغرض من كل ذلك هو إعطاء وصف أولي للظاهرة المدروسة بدون تحليل معمق[6].

ب / **الإحصاء الاستدلالي:** يتناول هذا النوع الطرق الإحصائية التي تستخدم في تحليل البيانات وتفسير النتائج، بهدف التوصل إلى استنتاج حول المصدر الذي جمعت منه البيانات من خلال الاعتماد على نظرية الاحتمالات، وينطلق من الجزء إلى الكل وفق طرق إحصائية محددة والذي يدخل ضمن الإحصاء التطبيقي. فغالبا ما تجرى دراسات إحصائية على عينات أي جزء من المجتمع بدلا من الدراسة الشاملة (المجتمع الإحصائي)، وذلك بسبب نقص الإمكانيات أو استحالة إجراء دراسة شاملة لأن الوقت غير كافي، وعليه فإن الإحصاء الاستدلالي يعني جمع البيانات حول العينة ثم تعميم النتائج إلى الكل من المجتمع، والشرط الأساسي لكي تكون عملية التعميم صحيحة يجب أن تكون العينة ممثلة تمثيلا جيدا للمجتمع أو تمثل صورة صغيرة أي مصغرة للمجتمع، فنقول هنا لقد استدللنا على خواص المجتمع على أساس خواص العينة، وهذا عكس الاستنباط الذي هو استخراج خواص من الجزء انطلاقا من خواص الكل.

---

[5] نداء محمد الصوص، مبادئ الإحصاء، المرجع سبق ذكره، ص 11.
[6] محمد جبر المغربي، الإحصاء الوصفي، المرجع سبق ذكره، ص 06.

وهنا يجب الإشارة إلى أن الإحصاء الوصفي و الإحصاء الاستدلالي هما فرعا علم الإحصاء الحديث، وكلاهما ضروريان في البحث واتخاذ القرارات السليمة، لذا ازدادت أهمية الإحصاء الاستدلالي على وجه التحديد.

5/ مجالات إستعمال الإحصاء:

نجد أن الإحصاء يتصف بالصفة التطبيقية لأنه يدرس الكمية للظواهر الاقتصادية والاجتماعية كونه علم قائم بحدّ ذاته، يتميز بمجموعة من الخصائص والمميزات التي تجعله مستقلا عن غيره من العلوم، في حين نجد علم الرياضيات يبحث في القوانين الموجودة بشكل عام، لذا نجد علم الإحصاء يحول مفاهيم مختلفة العلوم إلى قالب رقمي قياسي قابل للعد والملاحظة، ويعكس جوهر الظواهر والعمليات الخاصة بهذه العلوم ومدي تغيرها وتطورها، لذلك نلاحظ أن الإحصاء يستعمل في جميع المجالات، الفيزيائية، العلوم الطبيعية، العلوم الإنسانية، الإدارة، المحاسبة، علم الحياة، علم الوراثة، علم الاجتماع، العلوم الزراعية، وعلوم التربية وغيرها. لذا نجد مادة الإحصاء أصبحت مادة هامة تدرس في مختلف الكليات العملية والنظرية في كل جامعات مختلف الدول، وسوف نتناول بإيجاز استخداماته في المجالات التالية:

- **التحليل الاقتصادي:**

يستخدم علم الإحصاء في معظم الدراسات الاقتصادية التي تهدف عادة إلى التنبؤ والتخطيط، سواء أكانت على مستوى المؤسسة أم على مستوى الاقتصاد الوطني، حيث أصبحت المؤشرات والمقاييس الإحصائية من أهم الوسائل العلمية اللازمة في التحليل الاقتصادي، كالإحصائيات المتعلقة بالإنتاج الزراعي أو الصناعي، فنجد أن دراسة الأسعار والأجور والاستثمار والادخار والاستهلاك والتصدير والاستيراد، كلها أصبحت تعتمد على درجة كبيرة على الأسلوب الاستقرائي أو الاستدلالي الإحصائي والأسلوب الكمي في التحليل الاقتصادي وكذا حساب معدلات النمو، معرفة ومراقبة كمية الإنتاج في القطاعات المختلفة، دراسة السوق وغيرها.

- **في إدارة الأعمال:**

يلعب علم الإحصاء دورا هاما في مختلف النواحي الخاصة بإدارة الأعمال لما يقدمه من أساليب إحصائية تعتبر بمثابة منهج علمي، وطريقة البحث في إطار إدارة الأعمال، حيث يساهم الإحصاء في دراسة اتجاهات المبيعات وبيان الآثار الموسمية والدورية لها بهدف إعداد خطط مستقبلية للمؤسسات والشركات المدروسة، بالإضافة إلى تتبع حجم المخزون من المنتجات والمواد الخام والطاقة بهدف المحافظة على حجم معين من المخزون. كما أن الإحصاء يستخدم بشكل كبير في دراسة احتياجات المستهلكين ورغباتهم وأذواقهم بهدف توافق الرغبات مع الأذواق (دراسة السوق عن طريق بحوث التسويق).

للإحصاء استعمالات أخرى كتحديد درجة الجودة والكمية والرقابة على الجودة، حيث تلعب نظرية العينات والاختبارات المرتبطة بها دورا هاما في تنظيم الرقابة على الجودة وبالتالي تقدير المبيعات المستقبلية في إعداد الموازيات الخاصة بالإنتاج والتكاليف والمخزون في المواد المختلفة.

- **في الدراسات المحاسبة:**

يستخدم الإحصاء بشكل واسع في مجال الأعمال والدراسات المحاسبية مثل تحليل القوائم المالية، البحث عن المؤشرات ومعايير محاسبية، التنبؤ بالأرباح التجارية أو الصناعية، تحديد القيمة المضافة على صعيد المؤسسة أو القطاع الصناعي بشكل عام، ضف إلى ذلك تستخدم الطرائق الإحصائية، وبشكل خاص نظرية العينات في بحوث المراجعة التي تعتمد اعتمادا أساسيا على تطبيق طرق المعاينة في حالات التدقيق والبحث والاستكشاف والتقدير.

- **في دراسة الأسواق المالية:**

للإحصاء أهمية كبيرة في مجال العلوم المالية والمصرفية، إذ أنها تتميز بالتغيير الدائم الذي يجعل التنبؤ بها يتطلب أساليب إحصائية متكورة وعملية بالإضافة إلى السلوك البشري المعقد الذي يشكل أحد العوامل العامة أثناء دراسة الأسواق المالية والتغيرات التي تتعرض لها،

وإخلال دراسة للقروض المصرفية التي تمنحها المصارف في فترة زمنية معينة مراعاة للأغراض التي منحت من أجلها وآجالها مثل: النقد المتداول، أسعار السندات، المبيعات، أسعار فائدة الأسهم والسندات، المخزون، معدل العائد على الاستثمار، الودائع، القروض، رؤوس الأموال، أسعار الفائدة، أسعار الصرف، الرياضيات ودراسة التأمين وطرق حسابها، الصادرات والواردات، الدخل الوطني، أسواق البورصة، أسعار الأسهم، وحجم الأموال المستثمرة في المحافظة المالية.

- **في الدراسات الديمغرافية (السكانية):**

يعمل علم الإحصاء في الدراسات السكانية بهدف التعبير عن كثير من المؤشرات السكانية كالولادات، الوفيات، الهجرة، حيث أصبحت الأبحاث السكانية تعتمد بالدرجة الأولى بالدراسات الكمية سواء تعلق بالتعداد السكاني أو توزيعهم حسب العمر والجنس والحالة المدنية ومكان الولادة ودرجة التعليم، وكذا دراسة تجديد هذه المجتمعات السكانية عن طريق التحليل الإحصائي للعلاقات القائمة بين حوادث الولادة والوفاة والخصوبة والطلاق... .

تعتبر المقاييس والمؤشرات الإحصائية أساس الأبحاث السكانية المختلفة، وبالتالي تستخدم الطرق الإحصائية في دراسة الحوادث السكانية أخذا بالاعتبار الزمن كمتغير أساسي في قياس العمر وفترة الزواج وفترة الترمل ومدة الدراسة، سن الحمل، عدد وفيات الأجنة، عدد المواليد، توقيت الهجرة وغيرها. وبالتالي نستنتج أن الإحصاء يستخدم في جميع مراحل البحث السكاني باتخاذ أساليب إحصائية بكفاءة عالية.

- **في العلوم الإنسانية:**

إن مقاييس الإحصاء تسهل مهمة الباحث في العلوم الإنسانية وتزيل العقبات التي تتعرض له لذا أصبح علم الإحصاء فرعا من فروع العلوم الإنسانية وذلك نظرا لاستخداماته الكثيرة أثناء قيامه بالبحث وفي إيجاد حلول علمية ومنطقية، فسمي في علم الاجتماع بالإحصاء الاجتماعي، وفي علم الجغرافيا بالتحليل الكمي، وفي علم السكان بالإحصاء السكاني، وفي علم

المكتبات والإعلان بالإحصاء التطبيقي، وذلك بفضل استخدام وسائل إحصائية للتحليل التي تعتمد على وضع الفرضيات والاختبارات الإحصائية وتمنح للأبحاث في هذه العلوم سمة مميزة مرتبطة بتطبيق أساليب وأدوات إحصائية مختلفة.

- **في العلوم الدقيقة والتقنية:**

يستخدم الإحصاء بشكل واسع في فروع العلوم الدقيقة (العلوم البيولوجية) كالطب والصيدلة، علم الوراثة، وفي الهندسة وغيرها، لذا لا نجد أي مجال من مجالات البحث والدراسة لا يستعمل فيها علم الإحصاء، فالكثير من العلماء يحاولون إيجاد الأدوية للأمراض وذلك بتحليل عدة مشاهدات كلينيكية مستخلصة من العديد من المرضى[7].

ثالثا: تعريف بعض المصطلحات

ويشمل أهم المصطلحات والمفاهيم الإحصائية الأكثر شيوعا وتداولا على مستوى البحوث العلمية وهي على النحو التالي:

1) **الإحصائيات:** هي مجموعة من البيانات الإحصائية المنظمة في جدول وفي سلم بياني حول نشاط أو قطاع معين في الدولة، إذ تعتبر الإحصائيات مادة خام لعلم الإحصاء، مثل إحصائيات الإسكان ونعني بها مجموعة من البيانات الخاصة بالسكان في بلد ما، من بين هذه البيانات نجد: العدد الإجمالي للسكان، توزيع السكان حسب العمر، توزيع السكان حسب العدد أو حسب الجنس، التوزيع الجغرافي حسب الولايات، إحصائيات التجارة الداخلية والخارجية، إحصائيات الصناعة، إحصائيات التعليم العالي وغيرها.

2) **المجتمع الإحصائي:** هو مجموع الوحدات الإحصائية المراد دراستها والمعرفة بشكل دقيق والتي تشترك فيما بينها في الصفة الأساسية محل اهتمام الباحث، لذا يعرف المجتمع

[7] موساوي عبد النور، بركان يوسف، الإحصاء 1، دار العلوم للنشر والتوزيع، 2009، ص 06.

بأنه مجموعة من العناصر والأفراد التي تسند إليها الدراسة، يشترط في المجتمع الإحصائي أن يكون معرفا تعريفا جيدا مثل الأجور الشهرية لعمال المؤسسة (عمال المؤسسة = المجتمع الإحصائي)، إذن هي مجموعة المشاهدات والقياسات الخاصة بمجموعة من الوحدات الإحصائية والتي تخص ظاهرة من الظواهر القابلة للقياس مثل مجتمع من الطلبة، مجتمع من الأسر وغيرها[8].

3) **الوحدة الإحصائية:** هي العنصر أو الجزء الذي تجرى عليه الدراسة الإحصائية، فهو يمثل موضوع البحث، ويشترط في الوحدة أن تكون خاضعة لتعريف دقيق وواضح فقد تكون شيئا حيويا مثل شخص، سيارة، علبة، أو معنويا مثل فكرة...أو ماديا مثل مؤسسة، سيارة، علبة.

4) **العينة:** هي جزء من المجتمع الإحصائي يتم استخراجها بطرق إحصائية معينة حتى تكون ممثلة أحسن تمثيل للمجتمع الإحصائي، ويتم اعتماد العينة في الدراسة بدل المجتمع للأسباب التالية: كبر حجم المجتمع، ربحا للوقت والجهد والمال، الفحص قد يكون متلفا أو مؤذيا للوحدات، التسجيلات التاريخية قد لا تكون كاملة.

والعينة أنواع: وهي كما يلي[9]:

- **العينة العشوائية البسيطة:**

تعتبر العينة العشوائية البسيطة من أكثر العينات استعمالا، فهي تخص المجتمعات المتجانسة أي تلك التي لها خصائص متشابهة أو متقاربة مثل مجتمع الإطارات المتوسطة، مجتمع من الطلبة...إلخ إذن تسحب العينة العشوائية من مجتمع متجانس من حيث الخصائص ذات الصلة بموضوع البحث.

---

[8] جلاطو جيلالي، الإحصاء مع تمارين ومسائل محلولة، ديوان المطبوعات الجامعية، ط 7، 2007، ص05.
[9] حسين ياسين طعمه، إيمان حسين حنوش، طرق الإحصاء الوصفي، المرجع سبق ذكره، ص 40.

- **العينة الطبقية:**

ويقصد بالعينة الطبقية بأنها عملية تجزئة المجتمع الإحصائي إلى طبقات متجانسة، ومن ثم اختيار عينة عشوائية بسيطة من كل طبقة لتشكل في مجموعها حجم العينة المطلوبة، وتّعدّ العينة الطبقية من أفضل أنواع العينات الاحتمالية أو أكثرها دقة في تمثيل المجتمع[10].

- **العينة العنقودية:**

وهي الطريقة التي يتم بموجبها تقسيم المجتمع الإحصائي إلى عدد من المجموعات الجزئية التي يطلق على كل منها بالعنقود. وتعد العينة العنقودية الأسلوب البديل عن العينة العشوائية عندما يتعذر تحديد قائمة لعناصر المجتمع الإحصائي، كما تسمى أحيانا بعينة المجموعات التي تكون فيها وحدة المعاينة مكونة من عدة وحدات أولية، كما يطلق عليها أيضا بالعينة المساحية لأنها تقسم المجموعات على أساس جغرافي.

- **العينة المنتظمة:**

تعرف العينة العشوائية المنتظمة بأنها الطريقة التي يتم بموجبها تحديد نقطة الانطلاق ومدى السحب من خلال نسبة حجم المجتمع إلى حجم العينة التي تتحدد بشكل عشوائي (أي بالاستثناء إلى جداول والأرقام العشوائية) إذ تعد من أكثر أنواع العينات الاحتمالية استخداما في الأجهزة الإحصائية، نظرا لبساطتها وسهولة اختيار المفردات بموجبها وقلة التكاليف المترتبة على استخدامها[11].

---

[10] ابراهيم علي ابراهيم عبد ربه، مبادئ علم الإحصاء وتطبيقاته في استخدام Excel 2000/xp ، الدار الجامعية، الإسكندرية، مصر، 2005، ص 21.

[11] ابراهيم علي ابراهيم عبد ربه، مبادئ علم الإحصاء وتطبيقاته في استخدام Excel 2000/xp ، المرجع سبق ذكره، ص 24.

- **العينة ذات المراحل المتعددة:**

هي الطريقة التي تقسم المجتمع الإحصائي إلى وحدات أولية ثم اختيار عينة عشوائية من بين هذه الوحدات كمرحلة أولى، ثم تقسيم الوحدات الأولية إلى وحدات ثانوية ويتم اختيار عينة عشوائية من بين هذه الوحدات كمرحلة ثانية، ثم تقسيم الوحدات الثانوية إلى وحدات فرعية أصغر ويتم اختيار عينة عشوائية من بين هذه الوحدات وهكذا يتم الاستمرار بالتقسيم لحين انتهاء جمع البيانات في المرحلة الأخيرة والتي تؤلف عينة البحث بمعنى آخر أن العينة العشوائية المتعددة المراحل هي عبارة عن عينة يتم اختيارها على عدة مراحل حتى ننتهي من الحصول على عينة البحث في المرة الأخيرة.

وباختصار تستعمل هذه الطريقة عند دراسة ظاهرة إحصائية معينة على مستوى القطر ككل، وهذا بإتباع الخطوات التالية في المثال التالي[12]:

- ✓ تسحب ولاية بطريقة عشوائية من أصل مجموع الولايات.
- ✓ تسحب دائرة بطريقة عشوائية من أصل دوائر تلك الولاية.
- ✓ تسحب بلدية بطريقة عشوائية من أصل بلديات تلك الدائرة.
- ✓ تسحب مؤسسة واحدة بطريقة عشوائية من أصل مؤسسات تلك البلدية.

5) المتغير العشوائي:

هو العنصر المشترك لكل الوحدات الإحصائية التي تشكل المجتمع الإحصائي مثل الطول، السن، الإنتاج، المستوى العلمي... ، وعموما تنقسم المتغيرات الإحصائية إلى ما يلي:

أ/ المتغيرات الكيفية (وصفية- نوعية): variables qualitatives

هي تلك المتغيرات التي لا يمكن قياسها مباشرة بالأرقام فهي ليست عددية مثل الحالة المدنية، المهنة، الجنس (ذكر، أنثى)، لون الشعر، الجنسية، تقدير العلامات...، ويمكن أن تكون

[12] جلاطو جيلالي، الإحصاء مع تمارين ومسائل محلولة، المرجع سبق ذكره، ص06.

قابلة للترتيب (إنتاج جيد، مقبول أو ضعيف) أو تكون غير قابلة للترتيب (الجنس، الحالة المدنية، الذوق، الألوان).

ب/ المتغيرات الكمية (الرقمية): Variables quantitatives

هي تلك المتغيرات التي يمكن قياسها رقميا وعددها، وهي أكثر المتغيرات انتشارا واستعمالا لأن لغة الإحصاء هي لغة الأرقام، والتي تنقسم بدورها إلى قسمين وهما:

- **متغيرات كمية متقطعة (منفصلة): variables discrètes**

هي تلك المتغيرات التي تأخذ قيما صحيحة (قيم متباعدة ومتقطعة) ولا يمكن تجزئة وحدة قياسها، أو بمعنى آخر لا تأخذ قيما كسرية كعدد الأطفال في الأسرة، عدد الغرف، عدد قطع الغيار المنتجة...

- **متغيرات كمية مستمرة (متصلة): variables continues**

هي تلك التي تأخذ المشاهدة فيها كل القيم الممكنة لمجال الدراسة، أي قيمة رقمية في مدى معين، ونظرا للعدد غير المتناهي لهذه القيم نقسم مجال الدراسة إلى مجالات جزئية تسمى الفئات كالطول، الوزن، الوقت، الأجر الشهري وغيرها.

الفصل الثاني

جمع و عرض البيانات الإحصائية

تمهيد:

تعتبر عمليتي جمع المعلومات الإحصائية وعرضها من بين المراحل المهمة في عملية البحث الإحصائي، حيث يعتبر جمع البيانات من أهم العمليات الإحصائية وأكثرها دقة، فهي الأساس الذي يبنى عليها الإحصائي باقي دراسته وتحليله، وعلى صحته تتوقف صحة النتائج التي يتوصل إليها الباحث والتي تتم كما يلي:

أولا: جمع البيانات الإحصائية

تعتبر البيانات المادة الرئيسية التي تتوقف عليها بصورة اساسية دقة الوصف وتحليل الدراسة، لذا لا بد من الحرص الدقيق في الحصول على معلومات دقيقة وواقعية وسليمة حول الظاهرة محل الدراسة، حيث تعتبر البيانات العمود الفقري والحجر الاساسي في الاستنتاجات التي تبلغها من خلال إعتماد علم الإحصاء، فعلى قدر توفرها ودقتها وشمولها تتوقف دقة التحليل و الإستدلال الإحصائي وأهمية النتائج المتوصل إليها، وصحة القرارات المبنية عليها، ويتم جمع البيانات الإحصائية من المصادر و بالطرق التالية:

1)- مصادر جمع البيانات الإحصائية:

يعتمد الباحثون على مصدرين للحصول على البيانات الخاصة بظاهرة معينة، الأولى مباشرة و الثانية غير مباشرة وهما:

أ)- مصادر مباشرة (مصادر ميدانية):

هي الحصول على البيانات الإحصائية من مصادرها الأصلية مباشرة وتستخدم مصادر الميدان في حالة عدم توفر البيانات اللازمة من المصادر التاريخية وذلك بعدة طرق[13]:

[13] سهيل أحمد سمحان، محمود حسين الوادي، مبادئ الإحصاء للإقتصاد و العلوم الإدارية، المرجع سبق ذكره، ص 32.

* **المقابلة الشخصية:** وذلك من خلال طرح الأسئلة من قبل أشخاص مؤهلين ومديرين على الأشخاص الذين تتم مقابلتهم و يتم تدوين إجابتهم وتمتاز هذه الطريقة بما يلي:

– الحصول على المعلومات من مصادرها المباشرة.

– إمكانية توضيح أي غموض أو إلتباس بالأسئلة مما يزيد من دقة البيانات المدونة.

ولكن يعاب عليها أنه يمكن الوقوع في أخطاء خلال عملية تدوين الإجابات وإمكانية التحيز.

* **المقابلة الهاتفية:** تعد كوسيلة مباشرة لجمع البيانات من مصادرها وبتكاليف منخفضة، وهي مماثلة للمقابلة الشخصية، وأصبحت أكثر سهولة مع إدخال الهاتف النقال.

* **الإستبيان البريدي (المراسلة):**

يتم جمع البيانات عن طريق إرسال الإستبيان (إستمارة الأسئلة) إلى الأشخاص المبحوثين بالبريد ويطلب منهم الرّد عليها(ويشترط في ذلك التنظيم الجيد و الواضح لكل الأسئلة)، لكن ما يعاب عليها قلة الإستجابة وعدم إمكانية توضيح أي غموض في الأسئلة(وتمتاز بقلة تكاليفها).

* **الملاحظة المباشرة:** حيث يتم الحصول على البيانات الخاصة بالظاهرة المعينة عن طريق مراقبتها مباشرة وبصورة آلية أو يدوية.

ب)– مصادر غير مباشرة (التاريخية– مصادر رسمية):

وهي يتم جمع البيانات من مصادر تاريخية أي هي بيانات معدة مسبقا من قبل الجهات المعنية والتي غالبا ما تكون رسمية قد تتمثل في بيانات منشورة وجاهزة للإستخدام كالوثائق والمطبوعات والسجلات والبحوث والدراسات، كإمكانية الحصول على عدد اللاجئين وأماكن تواجدهم من هيئة الأمم المتحدة، أو عدد حوادث السيارات من دائرة السير[14].

[14] سهيل أحمد سمحان، مبادئ الاحصاء، المرجع سبق ذكره، ص 33.

2)- طرق جمع البيانات الإحصائية:

يتم جمع البيانات بطريقتين، إما طريقة الحصر الشامل(المسح الشامل) أو طريقة العينة، ويتم المفاضلة بين الطريقتين بناءا على عدة عوامل، إما لطبيعة المجتمع أو لطبيعة البيانات المطلوبة وأيضا الإمكانيات المادية المتاحة للباحث وهما:

أ)-المسح الشامل (الحصر الشامل): ويسمى " تعداد "، حيث يتم جمع البيانات من جميع عناصر المجتمع الإحصائي، اي دراسة كل وحدة في المجتمع، و بالتالي فإن النتائج تكون على مستوى عال من الدقة والوضوح و التفصيل، ونسبة الخطأ فيه بسيطة، ولكن يعاب عليه ارتفاع تكاليفه(أموال كثيرة) وحاجته للكثير من الوقت و الجهد، و لعدد كبير من الباحثين وأحيانا يصعب أو يستحيل إتباعها كما في فحص الدم.

ب)-العينة: وهي جزء من مجتمع الظاهرة قيد البحث ويجب اتباع أقصى درجات الحيطة والحذر عند أخذ العينة، وذلك لكي تمثّل المجتمع قيد الدراسة تمثيلا صادقا وسليما وصحيحا، أي تعميم نتائجها على المجتمع الذي سُحبت منه، ولكن النتائج هنا لن تكون بمستوى الدقة و الشمول كما في المسح الشامل.

يتطلب هذا الأسلوب دراسة جزء من المجتمع بحيث يراعي في الاختيار العينة أن تكون ممثلة للمجتمع تمثيلا صادقا، و يستخدم هذا الأسلوب إذا كان المجتمع أكبر مما تسمح به إمكانيات الباحث، أو إذا كان المجتمع متجانسا، بحيث دراسة عينة مأخوذة من مجتمع متجانس تؤدي إلى نفس النتائج التي نحصل عليها من دراسة نفس المجتمع وذلك بتلافي ضياع الوقت و الجهد.

ثانيا: عرض البيانات الإحصائية

تستخدم أساليب عرض البيانات الإحصائية شكل واسع في كل المجالات و العلوم، واكتسبت أساليب العرض أهمية أعظم في وقتنا الحاضر وذلك لقدرة التجهيزات الإلكترونية على القيام بالحسابات الإحصائية اللازمة ووضع أساليب العرض المختلفة، مما دعا إلى إعطاء هذا العنصر(الجانب) أهمية خاصة، حيث يتم عرض البيانات الإحصائية عرضا جدوليا وبيانيا وذلك

حسب طبيعة البيانات، ليتمكن الطالب من معرفة حالات استخدام كل أسلوب من هذه الأساليب حسب خصائصه وميزاته.

فبعد الإنتهاء من جمع البيانات سواء بأسلوب المسح الشامل أو العينة وسواء كانت البيانات كمية أو وصفية، فإنه يستلزم مراجعة تلك البيانات قبل تصنيفها ووضعها في جداول مناسبة، حيث يوجد عدة قواعد يجب إتباعها عند تصميم الجداول الإحصائية أهمها:

– أن يشمل الجدول الواحد وحدات متجانسة(متشابهة) أو توجد علاقة بينها؛

– لكل جدول رقم لتسهيل الرجوع إليه؛

– لكل جدول عنوان مختصر وواضح ومحدد لما يحتويه من معلومات مع ذكر وحدة القياس المستعملة فيه وذكر المصدر؛

– أن تكون عناوين الأعمدة و الصفوف مختصرة ودقيقة و مرتبة وفقا لتسلسل زمني أو حسب أهميتها؛

وتبعا لنوع المتغير يمكن تصنيف البيانات في الجداول المناسبة يطلق عليها إسم الجداول التكرارية وتسمى هذه العملية بالعرض الجدولي، ثم يتم توضيح هذه البيانات بشكل رسوم وأشكال بيانية وتسمى بالعرض البياني.

1)– العرض الجدولي للبيانات الإحصائية: Représentation par tableaux

بعد جمع البيانات لا بد من عرضها وتصنيفها وترتيبها، ثم تبويبها و وضعها في جداول مناسبة يسهل على الباحث التعامل معها و دراستها.

وفي ضوء ما تقدم، يُعرف تصنيف البيانات بأنه عملية فرز البيانات الخام حسب الصفة التي تشترك بها الظاهرة المدروسة، في حين يعرف التبويب بانه عملية تفريغ البيانات و تنظيمها في جداول خاصة على أساس اشتراكها بصفة معينة أو مجموعة من الصفات ثم تُحسب التكرارات[15].

[15] جلاطو جيلالي، الاحصاء، المرجع سبق ذكره، ص 11.

*** تفريغ البيانات وحساب التكرارات:**

نقوم بتشكيل جدول فيه(03) أعمدة، نضع في العمود الأول الفئات و العمود الثاني نقوم فيه بتفريغ البيانات، و ذلك بوضع كل قيمة في الفئة أو المجال الذي تنتمي إليه ويتم ذلك بتسجيل خطوط مائلة (مثل /)، أي بدل كتابة القيم نقوم بتسجيل في مكان كل قيمة خط مائل وعند وصول إلى خمس(05) قيم نقوم بغلق الأربع خطوط، الأولى بخط مائل عكسهم (////) (أي تكوين حزم من 4 خطوط يجمعها خط آخر)، وهذا لتسهيل الحساب لا غير(أي عملية عد الخطوط التي تمثل التكرارات المناظرة للفئات المختلفة)، ثم نقوم في العمود الثالث بحساب القيم في كل فئة و هو ما يسمى **بالتكرارات.**

أ)- العرض الجدولي للبيانات الكيفية (الوصفية، النوعية):

في حالة البيانات الكيفية يتكون جدول المعطيات على عمودين، يحتوي العمود الأول على رموز كتابية للخاصية المدروسة، أما العمود الثاني يحتوي على كل التكرارات، وهو ما يعرف بالتصنيف النوعي الذي ينظم البيانات وترتبها في جداول خاصة حسب اشتراكها بصفة معينة كالجنس(ذكر، أنثى) أو الحالة المدنية(متزوج، مطلق، أرمل) أو صنف الدم(AB,O^+,B,A^+)...

مثال 01: فيما يلي تقديرات 18 طالب في مقياس الرياضيات لسنة 2020 كما يلي:

جيد جدا- مقبول- ضعيف- جيد- جيد جدا- مقبول- ضعيف- ممتاز- جيد- ضعيف- جيد جدا- ضعيف- جيد- مقبول- ممتاز- ضعيف- مقبول- ممتاز.

المطلوب: تفريغ هذه البيانات في شكل جدول تكراري مناسب

الصفات (التقديرات)	التكرار بالاشارات (التفريغ)	عدد الطلبة(التكرارات)
ممتاز	///	3
جيد جدا	///	3
جيد	///	3
مقبول	////	4
ضعيف	/////	5
المجموع	/	**18**

ويأخذ العمود الأول(الصفات) و العمود الأخير(التكرارات)، فهنا نتحصل على ما يسمى بجدول توزيع تكراري بسيط للبيانات الوضعية حيث:

المجموع	ضعيف	مقبول	جيد	جيد جدا	ممتاز	التقديرات
18	5	4	3	3	3	التكرارات

ب)- العرض الجدولي للبيانات الكمية المنفصلة

المتغير الكمي(المنقطع) المنفصل هو ذلك المتغير الذي لا يكمن تجزئة قيمته، و الذي يأخذ قيم صحيحة، كعدد أفراد الأسرة، مقاسات الأحذية وغيرها.

مثال 02: البيانات التالية تمثل عدد أشجار الزيتون المملوكة من طرف 20 عائلة فلاحية في مدينة سطيف

03-07-08-10-05-03-07-09-08-03-10-05-10-10-08-05-07-10-09-03

المطلوب: تفريغ هذه البيانات في جدول تكراري بسيط.

الحل:

التكرارات	**تفريغ البيانات**	**عدد الأشجار**
4	////	3
3	///	5
3	///	7
3	///	8
2	//	9
5	~~////~~	10
20	/	**المجموع**

ج)- العرض الجدولي للبيانات الكمية المتصلة (المستمرة)

عند دراسة مغير كمي مستمر، فإن مجال الدراسة يضم ∞ من القيم، ولاستحالة أخذ كل هذه القيم ووضعها في جدول واحد، نقسم هذا المجال إلى مجالات جزئية تسمى **الفئات**، حيث يحدد عدد هذه الفئات حسب حجم العينة وحسب توزيع الوحدات الإحصائية على مجال الدراسة، ولكل فئة حد أدنى و حد أعلى، والفرق بينهما يسمى **بطول الفئة C**، حيث: **طول الفئة = الحد الأعلى للفئة – الحد الأدنى للفئة**

* **طول أو مدى الفئة:** هو المدى بين الحد الأدنى والحد الأعلى للفئة أو هو المدى الذي تتراوح بينه القيم التي تنتمي إلى تلك الفئة، ولا يوجد هناك طريقة محددة لحساب طول الفئة، فهناك من الباحثين من يقوم بتقسيم البيانات إلى عدد من الفئات ومن ثم يمكن تحديد طول الفئة وذلك بقسمة المدى العام على عدد الفئات.

– **الحد الأدنى:** هو بداية الفئة، ويرمز له بالرمز A_1

– **الحد الأعلى:** هو نهاية الفئة، ويرمز له بالرمز A_2

بحيث الفئة هي : $[A_1-A_2[$

– **طول الفئة:** هو مقدار سعة الفئة، أي المسافة(البعد) بين الحد الأعلى(A_2) للفئة و الحد الأدنى(A_1) للفئة، و يرمز له بـ(C)، أي: $C = A_2 - A_1$

لا توجد طريقة معينة لإيجاد عدد الفئات، وتختلف باختلاف أهمية الدراسة وحجم العينة، فالعالم الإحصائي **ستورجس STURGES** اقترح طريقة باستعمال اللوغارتمات وهي: $K = 1 + 3{,}322 \log N$

ومنه طول الفئة يساوي:

$$C = \frac{X_{max} - X_{min}}{1 + 3{,}322 \log N} \qquad \text{أي} \qquad C = \frac{R}{K}$$

مع :

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

حيث:

C: طول الفئة أو المدى

R: المدى العام

X_{max}: أكبر قيمة في البيانات

X_{min}: أصغر قيمة في البيانات

K: عدد الفئات

N: عدد قيم الظاهرة أو حجم العينة.

وإذا تم التصريح مباشرة بعدد الفئات يتم مباشرة تعويض K بتلك القيمة، أما في حالة عدم التصريح بذلك فتحسب بالقانون STURGES، و عند إيجاد طول الفئة C يُفضّل تقريبها إلى أقرب وحدة صحيحة.

* **عدد الفئات:** يحدد عدد الفئات بقيمة المدى العام على مدى الفئة (طول الفئة)، لذا يوجد عدّة قواعد يجب مراعاتها عند تحديد الفئات وهي:

– أن لا يكون عدد الفئات كبيرا جدا، و لا صغيرا جدا، تتراوح بين 5 و 20 أو 5 و 25 فئة للحفاظ على توازن التكرارات و لكي لا تفقد الجداول التكرارية أهميتها؛

– يستحسن أن تكون أطوال الفئات متساوية حتى يمكن مقارنة المتغير من مجال لآخر، فإذا كانت أطوال جميع الفئات لظاهرة معينة متساوية يسمى الجدول التكراري في هذه الحالة **بالجدول التكراري المنتظم**، أما إذا كانت الفئات غير متساوية فيقال أن الجدول التكراري **غير منتظم**؛

– إذا كان الحد الأدنى للفئة الأولى و الحد الأعلى للفئة الأخيرة معروفين، يقال أن الجدول التكراري **مغلق**، أما إذا كانا غير معرفين فيقال أن الجدول التكراري **مفتوحا**، كما يمكن أن يكون مفتوحا من طرفه الأدنى(من أوله) إذا كان الحد الأدنى للفئة الأولى غير معروف، ويكون مفتوحا من طرفه الأعلى إذا كان الحد الأعلى للفئة الأخيرة غير معروف، سُمى جدولا تكراريا **مفتوحا من آخره**، كما تبينه الجداول التالية[16].

* جدول مغلق من الطرفين:

التكرارات	الفئات
2	10−20
4	20−30
5	30−40

[16] حسين ياسين طعمه، إيمان حسين حموش، طرق الإحصاء الوصفي، المرجع سبق ذكره ص 62.

* جدول مفتوح من الطرفين:

التكرارات	الفئات
2	أقل من 20
4	20−30
5	30 فأكثر

* جدول مفتوح من الأسفل:

التكرارات	الفئات
2	أقل من 20
4	20−30
5	30−40

* جدول مفتوح من الأعلى:

التكرارات	الفئات
2	10−20
4	20−30
5	30 فأكثر

* الجدول التكراري المتصل: هو الجدول الذي يكون في الحد الأعلى للفئة السابقة نفسه الحد الأدنى للفئة اللاحقة، كما هو موضح في الجداول السابقة.

* الجدول التكراري المنقطع: هو عكس الجدول التكراري المتصل، أي لا يكون فيه الحد الأعلى للفئة السابقة هو نفسه الحد الأدنى للفئة اللاحقة، وبذلك تختلف كتابة حدود الفئات عنه، كما هو موضح في الجدول التالي:

التكرارات	الفئات
5	[45−53]
6	[54−62]
3	[63−71]
2	[72−80]
16	**المجموع**

تسمى حدود فئات الجدول المنقطع **بالحدود الظاهرية**، و لتحويله إلى جدول متصل يكفي أن نطرح 0,5 من الحدود الدنيا و إضافة 0,5 إلى الحدود العليا للفئات وتسمى حدود الفئات المتحصل عليها **بالحدود الفعلية** للفئات، فيصبح الجدول كالآتي:

التكرارات	الفئات
5	44,5−53,5
6	53,5−62,5
3	62,5−71,5
2	71,5−80,5
16	**المجموع**

* مركز الفئة: هو القيمة التي تتوسط المسافة بين الحد الأدنى و الحد الأعلى للفئة، و يرمز له بـ X_i، كما يمكن تعريفه أنه حاصل قسمة مجموع الحدين الأعلى والأدنى على 2، و الذي يمثل متوسط القيم الموجودة في كل فئة، بحيث: $X_i = (A_1 + A_2) / 2$

كما يمكن استنتاج القانون التالي لاستخراج حدود الفئة(القيمة العليا و الدنيا للفئة) باستعمال هذا القانون:

الحد الأدنى لفئة معينة = مركز الفئة لتلك الفئة – طول الفئة / 2

الحد الأعلى لفئة معينة = مركز الفئة لتلك الفئة + طول الفئة / 2

مثال 03: إذا كانت لدينا علامات 20 طالبا في مقياس ما كما يلي:

5−11−13−7−7−8−14−13−11−9−5−7−13−0−14−8−11−19−0−12

المطلوب: ضع هذه البيانات في جدول تكراري من فئات متساوية الطول ثم جد مراكز الفئات.

الحل:

1)− حساب طول الفئة C بطريقة ستورجس:

$$C = \frac{X_{max} - X_{min}}{1 + 3{,}322 \log N} = \frac{19 - 0}{1 + 3{,}322 \log(20)} = \frac{19}{5{,}3186} = 3{,}5729 \approx 4$$

إذن طول الفئة يساوي: $C = 4$

2)- كتابة حدود الفئة: نعلم أننا وجدنا طول الفئات يساوي 4، وأن أصغر قيمة هي 0 ، وعليه يمكن أخذ هذه القيمة كحد أدنى للفئة الأولى، فتكتب حدود الفئات كما يلي:

0 وأقل من 4

4 وأقل من 8

8 وأقل من 12

12 وأقل من 16

16 وأقل من 20

كما يمكن كتابتها على الشكل التالي:

]4−0]

]8−4]

]12−8]

]16−12]

]20−16]

3)- تفريغ البيانات وحساب التكرارات: بعد كتابة حدود الفئات نقوم بتفريغ البيانات و حساب التكرارات كما يلي:

التكرارات	تفريغ البيانات	الفئات
2	0,0	0 وأقل من 4
5	7،7،7،5،5	4 وأقل من 8
6	11،11،11،9،8،8	8 وأقل من 12
6	14،14،13،13،13،12	12 وأقل من 16
1	19	16 وأقل من 20
20	/	**المجموع**

وإستناداً لما سبق، يمكن كتابة حدود الفئات وتفريغ البيانات بالشكل التالي:

الفئات	تفريغ البيانات	التكرارات	مراكز الفئات
[0−4[	II	2	2
[4−8[	IIII	5	6
[8−12[	IIII I	6	10
[12−16[	IIII I	6	14
[16−20[	I	1	18
المجموع		**20**	/

تتراوح قيمة الفئة الأولى بين 0 وأقل من 4، حيث تسمى 0 بالحد الأدنى للفئة الأولى و 4 بالحد الأعلى للفئة الأولى، وعليه تسمى القيم 0،4،8،12،16 بالحدود الدنيا للفئات، أما القيم 4،8،12،16،20 فتسمى بالحدود العليا للفئات، و الفرق بين الحد الأعلى والحد الأدنى للفئات يساوي طول الفئة.

إن حاصل قسمة مجموع الحدين الأعلى والأدنى لكل فئة على 2 يساوي **مركز الفئة** و الذي يمثل متوسط القيم الموجودة في كل فئة.

مثل مركز الفئة الأولى = الحد الأدنى + الحد الأعلى ÷ 2

يساوي(0+4) ÷ 2 = 2، كما يمكن إستنتاج القانون التالي:

− الحد الأدنى للفئة الأولى = مركز الفئة الأولى − طول الفئة ÷ 2

* الحد الأعلى للفئة الأولى = مركز الفئة الأولى + طول الفئة ÷ 2

إذا علمنا(من المثال السابق) ان مركز الفئة يساوي 2 و إن طول الفئة يساوي 4، وأردنا معرفة أو إيجاد الحد الأعلى و الحد الأدنى للفئة فإننا نحسب كما يلي:

* الحد الأدنى للفئة الأولى = 2 − 4 ÷ 2 = 0
* الحد الأعلى للفئة الأولى = 2 + 4 ÷ 2 = 4
* الحد الأدنى للفئة الثانية = 6 − 4 ÷ 2 = 4
* الحد الأعلى للفئة الثانية = 6 + 4 ÷ 2 = 8

وهكذا ...

*** التكرارات:**

هي عبارة عن جميع المفردات التي تقع ضمن حدود الفئة من حيث القيمة العددية[17]، ويرمز له بالرمز (n_i)

فإذا رمزنا إلى تكرار الفئة الأولى(n_1) وتكرار الفئة الثانية(n_2)، وتكرار الفئة الثالثة(n_3)، ......، وتكرار الفئة k بالرمز (n_k)، فإن مجموع التكرارات $\sum ni$

$$\sum n_i = n_1+n_2+n_3+ \ldots+n_i = N$$

إذ أن: N يمثل حجم العينة

*** التكرارات النسبية و النسبية المئوية:**

في بعض الأحيان قد يكون من المناسب أن تُعرض البيانات في شكل جدول توزيع تكراري نسبي بإظهار تكرار كل فئة كنسبة من المجموع الكلي للتكرارات، كما يمكن أن تكون عل شكل نسب مئوية، بحيث نجد:

- التكرارات النسبية:

هي عبارة عن حاصل قسمة التكرار المطلق للفئة (n_i) على مجموع التكرارات(N) ويرمز لها بـ(n_i ÷):

$$\text{التكرار النسبي} = \frac{\text{التكرار المطلق}}{\text{مجموع التكرارات}}$$

$$n_i \div = \frac{n_i}{N}$$

ويجب أن يكون مجموع التكرارات النسبية لكل جدول تكراري يساوي 1 صحيح.

مثال 04: فيما يلي جدول تكراري بسيط يمثل توزيع أوزان 100 شخص

المطلوب: اعداد جدول تكراري نسبي

المجموع	90 فأكثر	70 – 90	60 – 70	50 – 60	فئات الوزن(كغ)
100	15	20	25	40	التكرارات(n_i)

[17] حسين ياسين طعمة، إيمان حسين حموش، طرق الإحصاء الوصفي، المرجع سبق ذكره، ص 61.

الحل: يمكن الحصول على جدول تكراري نسبي وفقا للصيغة الآتية: $n_i\div = n_i / N$

فئات الوزن(كغ)	50 – 60	60 – 70	70 – 90	90 فأكثر	المجموع
التكرارات(n_i)	40	25	20	15	**100**
التكرار النسبي($n_i\div$)	0,4	0,25	0,20	0,15	**1**

من النتائج الواردة في الجدول السابق يتضح ما يلي:

1)- يكون مجموع التكرارات النسبية مساو للواحد صحيح دائما

$$\sum n_i\div = 1$$

2)- إن قيمة 0,2 مثلا تعني أن (20 %) من الأشخاص تتراوح أوزانهم بين 70 كغ إلى أقل من 90 كغ، وهكذا بالنسبة لبقية قيم الجدول.

التكرارات النسبية المئوية:

يمكن تحويل التكرارات النسبية إلى تكرارات مئوية وذلك بضرب التكرارات النسبية في 100، أي:

$$n\,\% = n \div x\ 100$$

بمعنى آخر يمكن كتابتها على الشكل التالي:

$$\text{التكرار النسبي المئوي} = \frac{\text{التكرار العادي}}{\text{مجموع التكرارات}} \times 100$$

وعليه تكون التكرارات المئوية للجدول السابق على النحو التالي:

الفئات	50–60	60–70	70–90	90 فأكثر	المجموع
التكرار النسبي المئوي	40 %	25 %	20 %	15 %	**100 %**

يتضح من الجدول السابق بأن مجموع التكرارات المئوية مساو إلى 100% دائما.

– **التكرار التجميعي الصاعد(ت.ت.ص):** و يرمز له بـ (n) نحصل عليه بجمع او تراكم التكرارات البسيطة المقابلة لكل فئة بدءا بتكرار الفئة الأولى وإنتهاءا بتكرار الأخيرة إلى أن نحصل على مجموع التكرارات، وذلك من بداية الجداول(أول فئة) إلى نهايتها.

عمليا نقوم بما يلي:

التكرار التجميعي الصاعد للفئة الأولى هو عبارة عن التكرار الأول(n_1)، التكرار التجميعي الصاعد للفئة الثانية هو n_1+n_2 والتكرار التجميعي الصاعد الفئة الثالثة هو $n_3+n_2+n_1$، و في الأخير فإن تكرار التجميعي الصاعد للفئة الأخير يساوي $\sum ni = N$.

– **التكرار التجميعي النازل(ت.ت.ن):** و يرمز له (n)، نتحصل عليه بطرح تكرارات الجدول الأصلي من مجموع التكرارات تعاقبا، إذن نجد أن:

التكرار التجميعي النازل للفئة الأولى يساوي: $\sum n_i = N$

والتكرار التجميعي النازل للفئة الثانية يساوي: $N - n_1$

والتكرار التجميعي النازل للفئة الثالث يساوي: $(N - n_1) - n_2$

أما التكرار التجميعي النازل للفئة الأخيرة يساوي التكرار البسيط للفئة الأخيرة.

كما يمكن حساب التكرار التجميعي الصاعد النسبي المئوي أو التكرار التجميع النازل المئوي من خلال قسمة التكرار المطلق أيا كان على مجموع التكرارات ضرب 100، و مجموع التكرارات النسبية المئوية لكل جدول تكراري يساوي 100، في حين ينتهي التكرار التجميعي الصاعد المئوي بـ 100 %، و يبدأ التكرار التجميعي النازل المئوي بـ 100%.

مثال 05:

إذا كان لدينا الجدول التكراري التالي، يبين أجور 50 عاملا(الوحدة 10^3).

الفئات	20-30	30-40	40-50	50-60	المجموع
التكرارات	25	30	40	5	100

المطلوب:

- إيجاد التكرارات النسبية و النسبية المئوية المطلقة.
- إيجاد التكرارات التجميعية الصاعدة و النازلة.
- إيجاد التكرارات التجميعية الصاعدة و النازلة النسبية.
- إيجاد التكرارات التجميعية الصاعدة و النازلة النسبية المئوية.

الحل:

الفئات	n_i	n_i ÷	n_i %	$n\nearrow$	$n\nearrow$ ÷	$n\nearrow$ %	$n\swarrow$	$n\swarrow$ ÷	$n\swarrow$ %
20–30	25	0,25	25%	25	0,25	25%	100	1	100%
30–40	30	0,30	30%	55	0,55	55%	75	0,75	75%
40–50	40	0,40	40%	95	0,95	95%	45	0,45	45%
50–60	5	0,05	5%	100	1	100%	5	0,05	5%
المجموع	100	1	100 %	/	/	/	/	/	/

ملاحظات حول تكوين الجداول: يشترط عند إنجاز الجداول توفير الشروط التالية[18]:

– وجود عنوان يوضح محتوى الجدول؛

– ابراز الوحدة المستعملة (دج، كغ، ...) بجانب العنوان أو أمام عنوان كل عمود و سطر إذا تعددت وحدات القياس؛

– يجب أن تكون عناوين السطور والأعمدة واضحة؛

– ذكر المصدر الذي استخرجت منه البيانات بالتفصيل في أسفل الجدول بهدف مراقبة صحة المعلومات الموجودة في هذا الجدول أو الإطلاع على معلومات أخرى؛

– تسجيل كل التوضيحات التي تخص العناوين أو المحتوى في أسفل الجدول و يشار إليها بعلامات خاصة.

2)–العرض البياني للبيانات الإحصائية:

وهي عبارة عن تمثيل ووصف البيانات التي يتم جمعها عن ظاهرة ما، بواسطة إشكال بيانية أو هندسية، بهدف إعطاء فكرة واضحة و سهلة وسريعة عن بيانات الظاهرة، فهي تحويل الأرقام المطلقة إلى أشكال بيانية أكثر حيوية، اي نستبدل طريقة القراءة باستخدام حاسة البصر التصويرية لأنه أقرب للفهم و المقارنة

[18] موساوي عبد النور و بركان يوسف، المرجع سبق ذكره، ص 13.

والإستنتاج من الطرق الأخرى، لذا تُستعمل عندما يُقصد منها إعطاء فكرة خاطفة وسريعة تُثبّت في الذهن و الذاكرة.

وقبل البدء في أي رسم بياني، هناك عدة نقاط يجب أخذها بعين الإعتبار وأهمها:

- تحديد الهدف من الرسم البياني؛
- ثم تحديد نوع الرسم المستخدم وحجمه وعنوانه؛
- وضع عنوان لكل رسم بياني وعادة ما يكون في أعلى الرسم.

وفيما يلي أهم طرق العرض البياني حسب مختلف أنواع المتغيرات الإحصائية.

أ)- العرض البياني في حالة متغير كيفي (وصفي، نوعي)

في حالة البيانات الكيفية يتكون الجدول المعطيات على عمودين، يحتوي العمود الأول على رموز كتابية للخاصية المدروسة، و العمود الثاني يحتوي على كل التكرارات، أما التمثيل البياني يتمثل في العروض البيانية التالية:

* **العرض الدائري:** ويتمثل في دائرة مقسمة إلى عدة أجزاء، كل جزء يقابل زاوية مركزية تتناسب مع التكرارات المقابلة لها، ولتحقيق ذلك نضيف عمودا إلى جدول المعطيات يحتوي على الزاوية المركزية المقابلة لكل تكرار، ويتم حساب هذه الزاوية المركزية عن طريق ضرب درجة الدائرة($360°$) في تكرار المشاهدة على مجموع التكرارات.

* **العمود المجزأ:** هو عبارة عن مستطيل مقسم إلى عدة أجزاء، كل جزء يقابل النسبة المئوية المقابلة لتكرار معين من تكرارات المتغير محل الدراسة.

* **الأعمدة المستطيلة:** وهي عبارة عن مستطيلات متباعدة بمسافات ثابتة ولها قواعد متساوية، وتتناسب أطوالها مع التكرارات المقابلة لها.

مثال 06: يبين الجدول التالي الإنتاج العالمي للذهب حسب القارات في سنة ما.

القارات	أوربا	آسيا	افريقيا	أمريكا	استراليا	**المجموع**
كمية الذهب	176	87	431	350	56	**1100**

<u>المطلوب</u>: اعرض هذه البيانات باستخدام الأشكال البيانية التالية: الدائرة البيانية– العمود المجزأ – ثم الأعمدة المستطيلة

<u>الحل</u>:

1)– الدائرة البيانية:

يتم إيجاد الزوايا المركزية للقارات وفقا للصيغة التالية:

– الزاوية المركزية لقارة أوربا:

$$\begin{array}{lcl} 360^\circ & \longrightarrow & 1100 \\ X & \longleftarrow & 176 \end{array} \quad \Longrightarrow \quad X = \frac{360^\circ \times 176}{1100} = 57{,}6^\circ$$

– الزاوية المركزية لقارة آسيا:

$$\begin{array}{lcl} 360^\circ & \longrightarrow & 1100 \\ X & \longleftarrow & 87 \end{array} \quad \Longrightarrow \quad X = \frac{360^\circ \times 87}{1100} = 28{,}47^\circ$$

– الزاوية المركزية لقارة إفريقيا:

$$\begin{array}{lcl} 360^\circ & \longrightarrow & 1100 \\ X & \longleftarrow & 431 \end{array} \quad \Longrightarrow \quad X = \frac{360^\circ \times 431}{1100} = 141{,}05^\circ$$

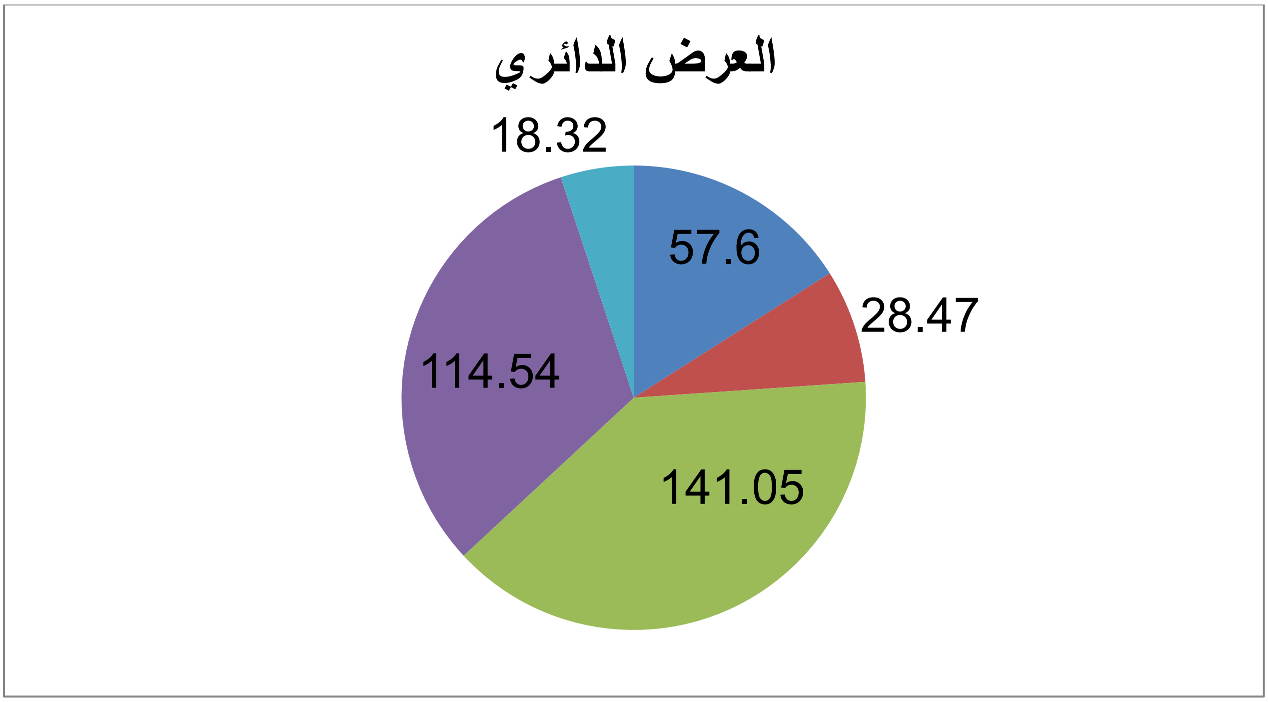
– الزاوية المركزية لقارة أمريكا:

$$\begin{array}{lcl} 360^\circ & \longrightarrow & 1100 \\ X & \longleftarrow & 350 \end{array} \quad \Longrightarrow \quad X = \frac{360^\circ \times 350}{1100} = 114{,}54^\circ$$

-الزاوية المركزية لقارة أستراليا:

$$\begin{aligned} 360° &\longrightarrow 1100 \\ X &\longleftarrow 56 \end{aligned} \Longrightarrow X = \frac{360° \times 56}{1100} = 18{,}32°$$

القارات	أوروبا	آسيا	إفريقيا	أمريكا	أستراليا	المجموع
كمية الذهب	176	87	431	350	56	**1100**
الزاوية المركزية	57,6°	28,47°	141,05°	114,54°	18,32°	**360°**



2)- العمود المجزأ

- قارة أوربا:

$$\left.\begin{aligned} 100\ \% &\longrightarrow 1100 \\ X &\longleftarrow 176 \end{aligned}\right\} \Longrightarrow X = \frac{100 \times 176}{1100} = 16\ \%$$

- قارة آسيا:

$$\left.\begin{aligned} 100\ \% &\longrightarrow 1100 \\ X &\longleftarrow 87 \end{aligned}\right\} \Longrightarrow X = \frac{100 \times 87}{1100} = 08\ \%$$

– قارة إفريقيا:

$$\left.\begin{array}{l} 100\,\% \longrightarrow 1100 \\ X \longleftarrow 431 \end{array}\right\} \Longrightarrow X = \frac{100 \times 431}{1100} = 39\,\%$$

– قارة أمريكا:

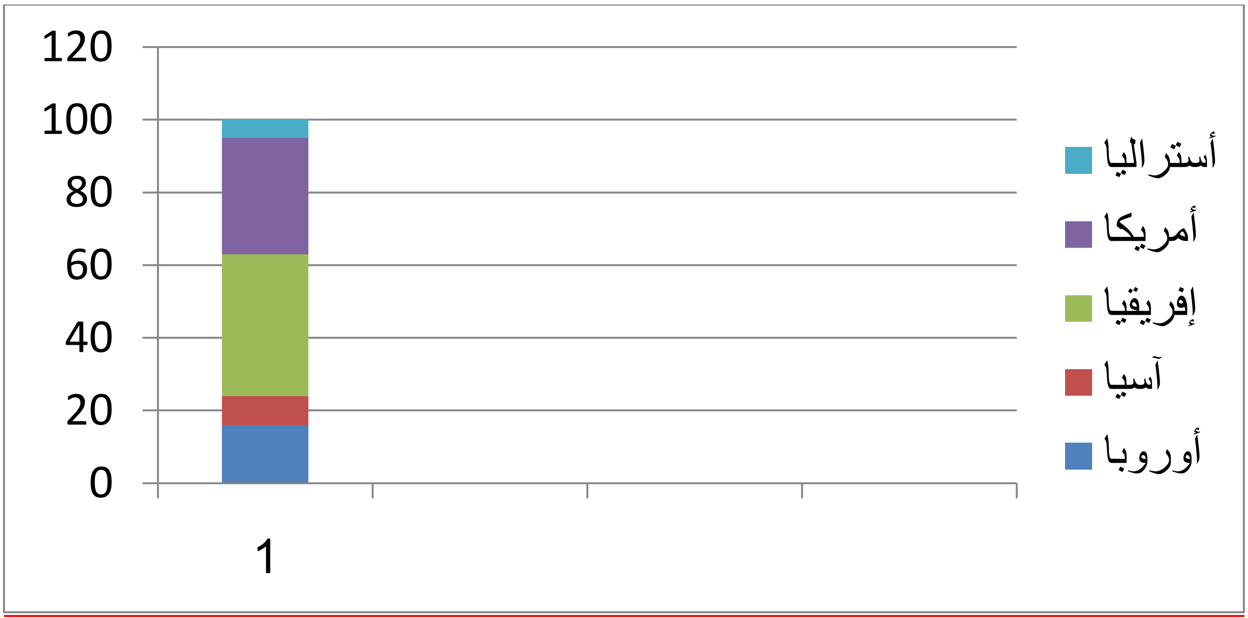
$$\left.\begin{array}{l} 100\,\% \longrightarrow 1100 \\ X \longleftarrow 350 \end{array}\right\} \Longrightarrow X = \frac{100 \times 350}{1100} = 32\,\%$$

– قارة أستراليا:

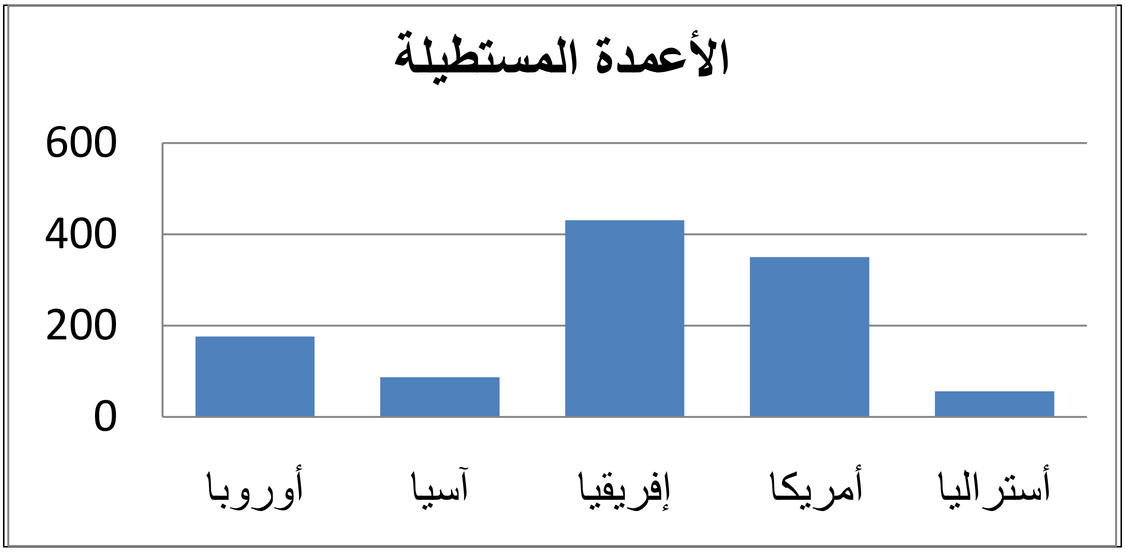
$$\left.\begin{array}{l} 100\,\% \longrightarrow 1100 \\ X \longleftarrow 56 \end{array}\right\} \Longrightarrow X = \frac{100 \times 56}{1100} = 5{,}09\,\%$$

القارات	أوروبا	آسيا	إفريقيا	أمريكا	أستراليا	المجموع
كمية الذهب	176	87	431	350	56	**1100**
النسب	16 %	8 %	39 %	32 %	5 %	**100 %**

العمود المجزأ:



3)- الأعمدة المستطيلة:



ب) العرض البياني في حالة متغير كمي منقطع (منفصل) :

يمكن تمثيل المتغيرات الكمية المنقطعة حسب العروض البيانية التالية:

* **الأعمدة البسيطة:** هي عبارة عن خطوط عمودية تتناسب أطوالها مع التكرارات المقابلة لها.

* **المنحنى أو خط بياني:** هو عبارة عن خط منكسر أو منحنى يصل بين النقاط التي إحداثياتها قيم المتغير والتكرارات المقابلة لها.

* **القطع المستقيمة للتكرارات التجميعية الصاعدة و النازلة:** هي عبارة عن قطع مستقيمة متصاعدة أو متنازلة حسب تصاعد أو تنازل التكرارات التجميعية الصاعدة أو النازلة المقابلة لكل قيمة من قيم المتغير الإحصائي المدروس.

مثال 07: لدراسة متوسط عدد الأطفال في الأسرة في عائلة ما، أُخذت عينة متكونة من 25 أسرة، فكانت النتائج كما يلي:

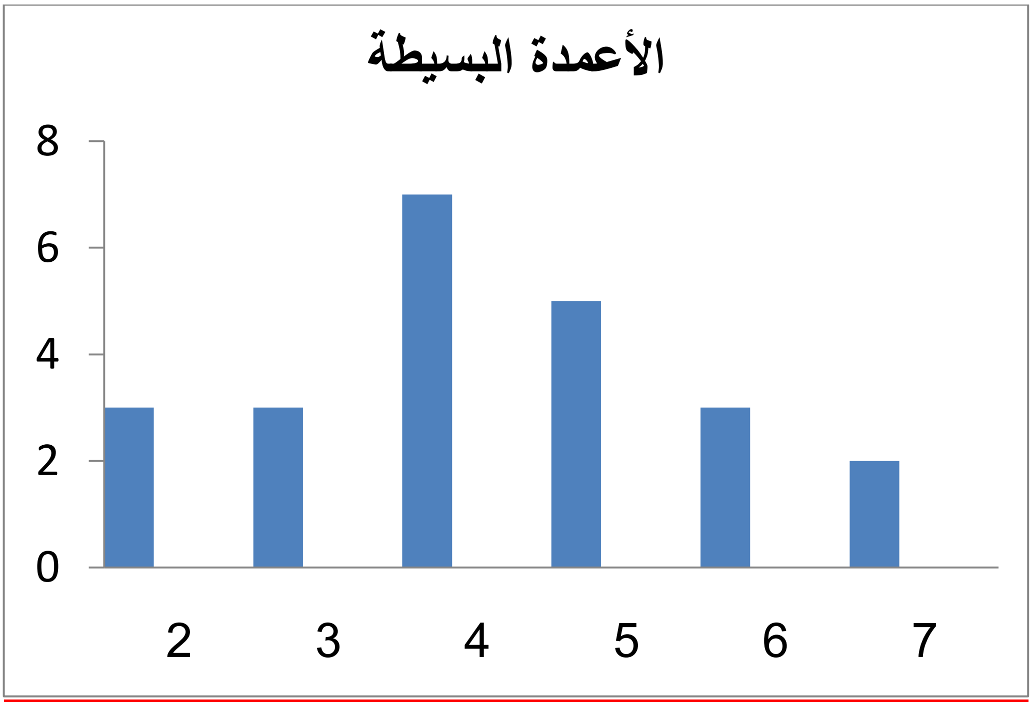
5، 4، 4، 5، 4، 3، 5، 6، 3، 4، 6، 7، 2، 8، 6، 2، 5، 4، 7، 4، 2، 5، 8، 3، 4

المطلوب: إعداد هذه البيانات في جدول تكراري ثم تمثيل هذا الجدول حسب العروض البيانية المناسبة

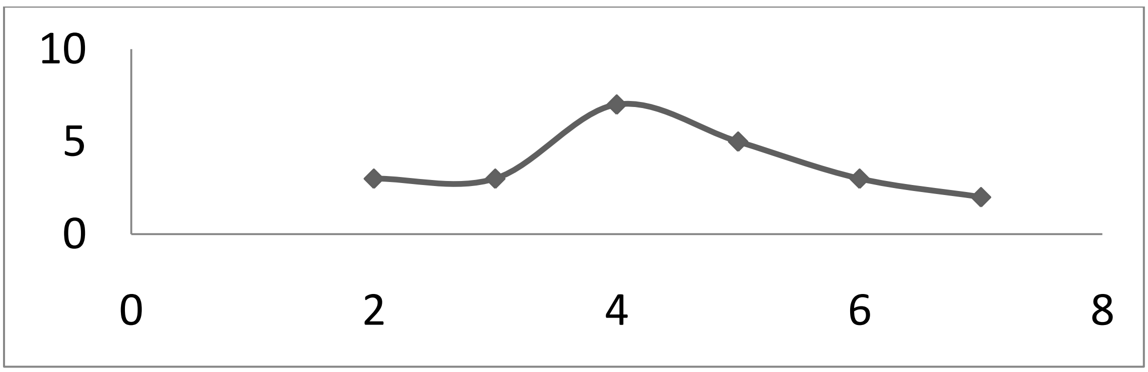
الحل: يجب أن نحدد التكرار التجميعي الصاعد و النازل كما يلي:

عدد الأطفال X_i	2	3	4	5	6	7	8	المجموع
التكرارات n_i	3	3	7	5	3	2	2	**25**
ت ت ص $n\nearrow$	3	6	13	18	21	23	25	/
ت ت ن $n\swarrow$	25	22	19	12	7	4	2	/

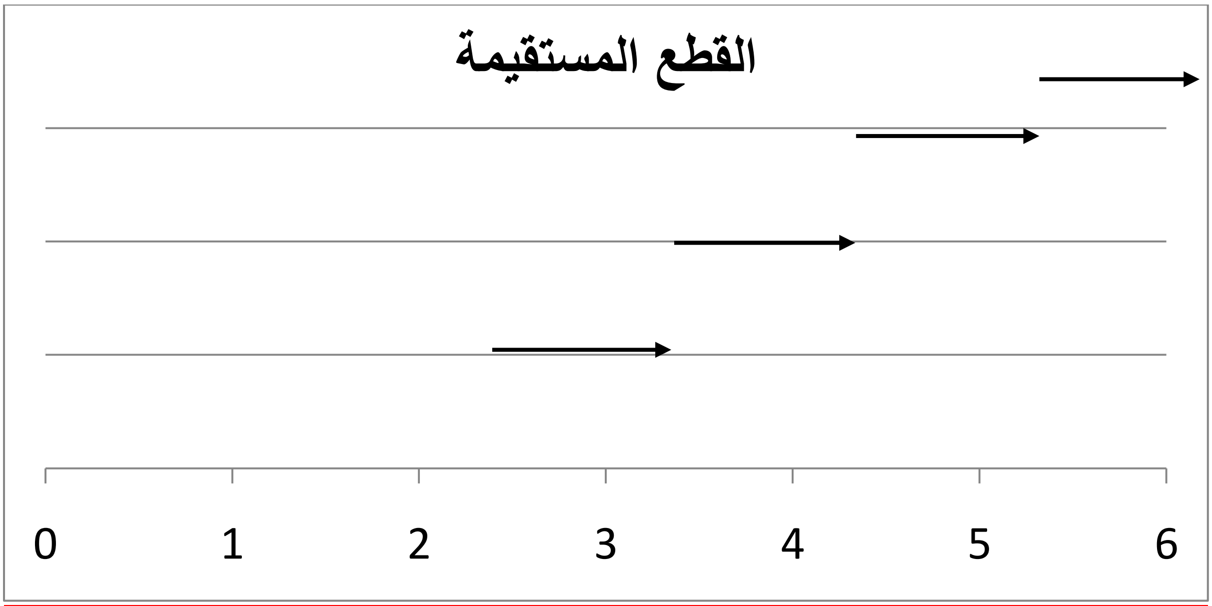
1)- الأعمدة البسيطة



2)- المنحنى البياني



3)- القطع المستقيمة الصاعدة والنازلة



ملاحظـة: لرسم القطعة المستقيمة المقابلة للقيمة 3 ، و الذي تكرارها الصاعد يساوي 3، نرسم قطعة مستقيمة طولها 1 سم، عند إحداثيات النقطة(3،3)، تبدأ من مستوى القيمة 2، وهكذا بالنسبة لبقية القطع المستقيمة.

ج)- العرض البياني في حالة متغير كمي مستمر(متصل):

تعتبر العروض البيانية للمتغير الكمي المستمر أكثر العروض البيانية استعمالا، ومن أهمها ما يلي:

***- المدرج التكراري:**

هو عبارة عن مجموعة من المستطيلات العمودية المتجاورة ومتلاصقة طول كل مستطيل منها يتناسب مع التكرار المقابل له، وقاعدة كل منها تساوي طول الفئة المقابل لها، حيث توضع الفئات على المحور الأفقي و التكرارات على المحور العمودي[19].

وعند رسم المدرج التكراري يمكن أن نميز بين حالتين:

[19] تيلولت سامية، مبادئ في علم الإحصاء، ديوان المطبوعات الجامعية، 2016، ص 41.

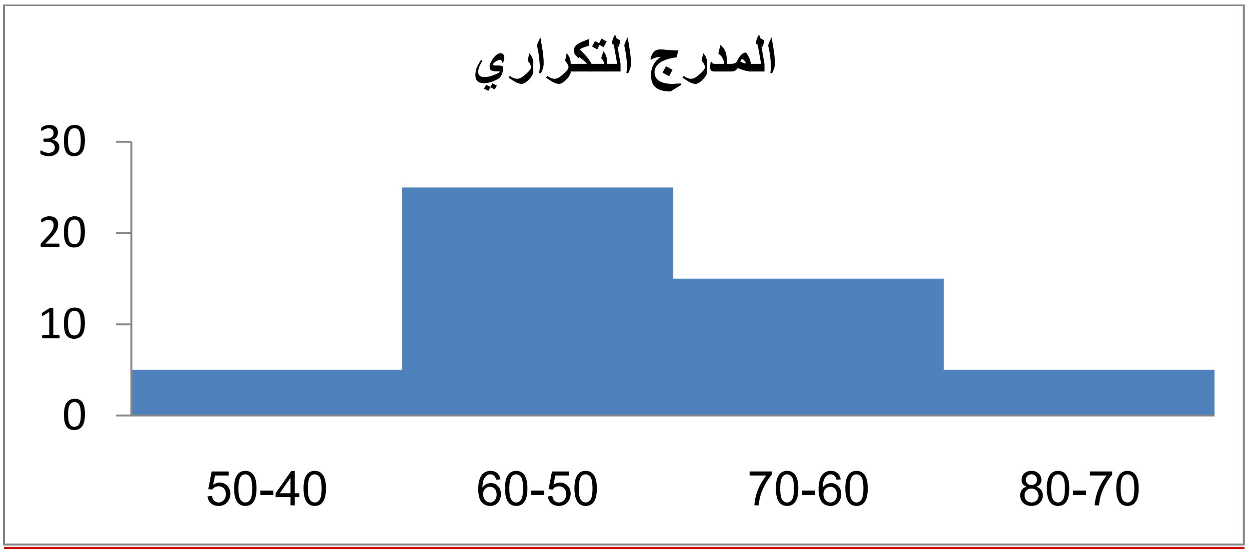
– **الحالـة الأولـى:**عندما تكون الفئات متساوية، في هذه الحالـة تكون قاعدة المقارنـة ثابتـة ومتساوية، ومن ثم لا نجري أي تعديل على جدول التوزيع التكراري.

مثال 08: يمثل الجدول التكراري التالي أوزان 50 شخص

المطلوب: تمثيل بيانيا التوزيع باستخدام المدرج التكراري

الأوزان	40–50	50–60	60–70	70–80	**المجموع**
التكرارات	5	25	15	5	**50**

الحل: بما أن الجدول منتظم، نقوم بالرسم مباشرة دون الحاجة إلى تعديل التكرارات، وهي كالآتي:



– **الحالة الثانية:** في حالة أطوال الفئات غير متساوية(الجدول غر منتظم) هنا يجب تعديل التكرارات، لآن القاعدة المقارنة غير ثابتة حتى يكون هناك تناسب بين طول الفئة والتكرار المقابل لها، أي إيجاد عدد الوحدات الاحصائية الموزعة على وحدة قياس معينـة، ويتم هذا عن طريق التكرارات **المعدلة.**

و التكرار المعدل هو عبارة عن النسبة بين التكرار البسيط وطول الفئة المقابلة له.

أي : $ER = n_i / c_i$

مثال 09: يمثل الجدول التالي أوزان 50 شخص

الأوزان	50-60	60-70	70-90	90-100	المجموع
التكرارات	5	15	20	10	**50**

المطلوب: تمثيل بيانيا التوزيع باستخدام المدرج التكراري

الحل: بما أن الجدول التكراري غير منتظم(أطوال فئات غير متساوية)، لذا يجب القيام بتعديل التكرارات، اي إيجاد التكرار المعدل(ER) كالآتي: أي : $ER_i = n_i / c_i$

بحيث:

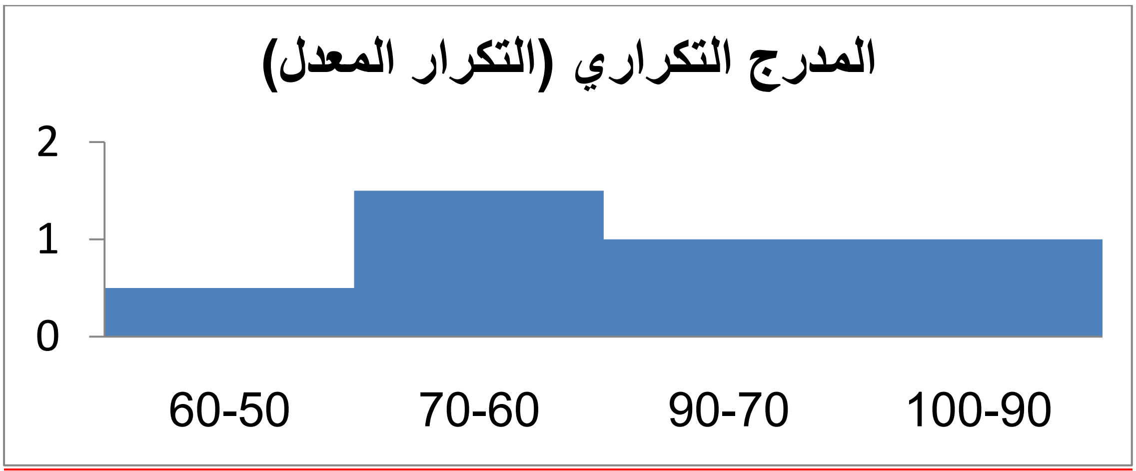
n_i: تكرار كل فئة.

c_i: يمثل طول الفئة المقابل.

ER_i: التكرار المعدل للفئة.

الأوزان	50-60	60-70	70-90	90-100	المجموع
التكرارات ni	5	15	20	10	**50**
طول الفئة ci	10	10	20	10	/
التكرار المعدل ERi	5/10 = 0,5	15/10 = 1,5	20/10 = 1	10/10 = 1	/

التمثيل البياني:



ملاحظة: نقوم بتعديل التكرارات عندما تكون الفئات غير متساوية في حاليتين:

- عند رسم المدرج التكراري.
- و عند تحديد الفئة المنوالية وحساب المنوال.

*** المضلع التكراري:**

هو مجموعة من قطع مستقيمة متصلة ومنكسرة، بمعنى آخر هو عبارة عن خط بياني منكسر يتحدد إحداثياتها عن طريق مراكز الفئات و التكرارات المقابلة لها، حيث نرسم محورين متعامدين، على المحور الأفقي مراكز الفئات، و المحور العمودي التكرارات، ثم نضع نقاط تتناسب مع كل مركز وتكرارها ونصل بين تلك النقاط بخط منكسر.

أما بدارية الخط فنصله بمركز فئة وهي سابق للفئة الأولى وتكرارها يساوي الصفر، ونفس الشيء بالنسبة لنهاية الخط نصله بمركز فئة وهمي لاحق للفئة الأخيرة وتكرارها يساوي الصفر.

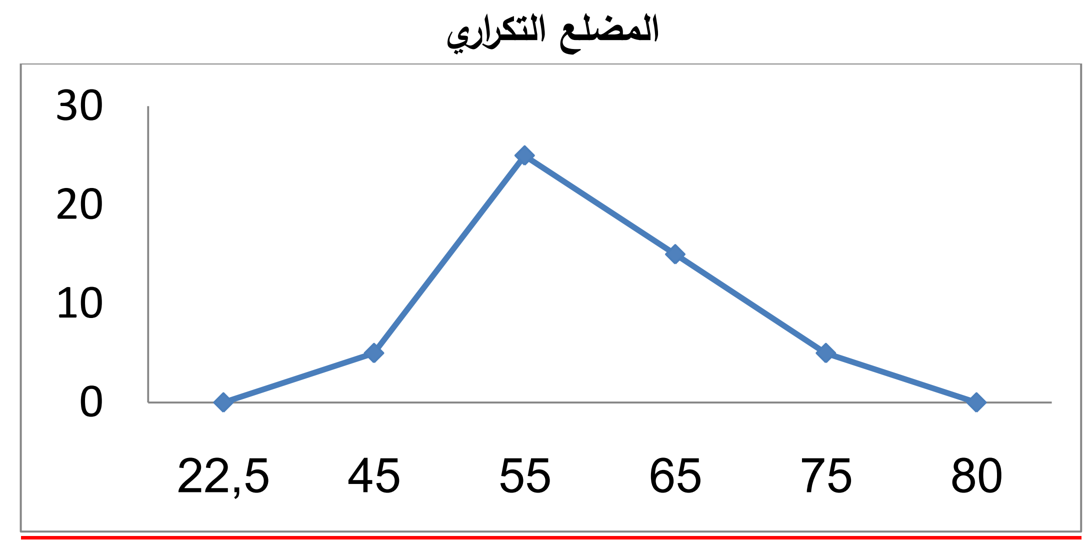
يمتاز المضلع التكراري عن المدرج التكراري بأنه يسمح لنا بمقارنة أكثر من ظاهرتين عند رسمهما في نفس المعلم، وانطلاقا من المثال السابق(وزن 50 شخص) يوضح الشكل التالي كيفية رسمه كما يلي:

الأوزان	40-50	50-60	60-70	70-80	**المجموع**
التكرارات	5	25	15	5	**50**
مراكز الفئات X_i	45	66	65	75	/

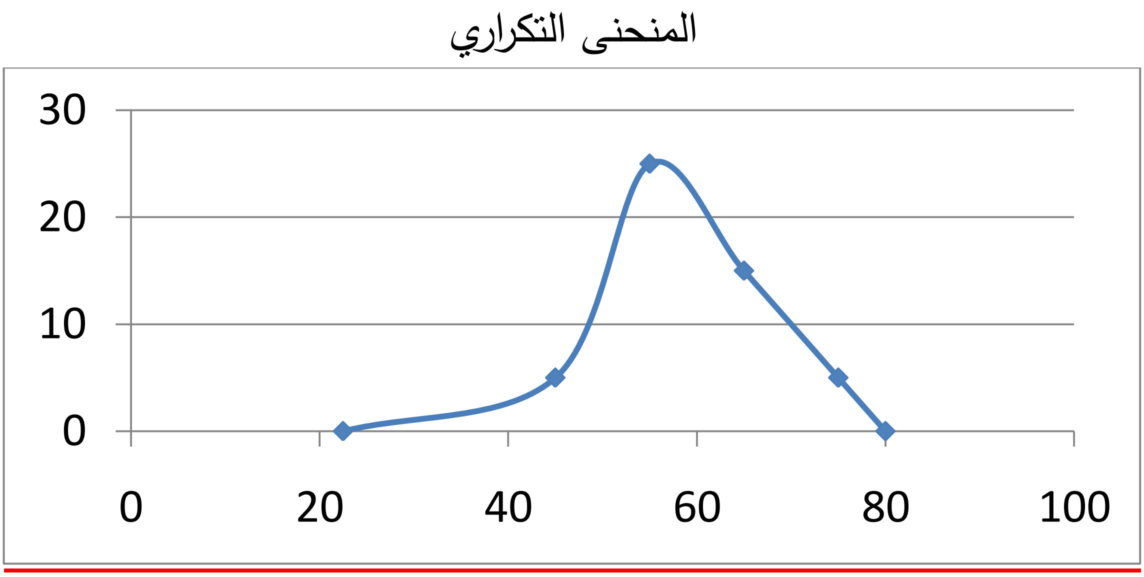
بحيث:

$X_1 = (40 + 50) / 2 = 45$ $X_2 = (50 + 60) / 2 = 55$ $X_3 = (60 + 70) / 2 = 65$

$X_4 = (70 + 80) / 2 = 75$



المنحنى التكراري: باتباع نفس الخطرات التي قمنا بها في رسم المضلع التكراري نقوم برسم المضلع التكراري، ولكن نقوم بتوصيل النقاط بخط ممهد باليد بدلا من الخط المنكسر، كما هو موضح في الشكل التالي:



ملاحظة: يمكن رسم المنحنى التكراري في نفس المعلم مع المدرج التكراري وذلك من خلال التوصيل بين منصفات قسم المستطيلات بخط منكسر لآن منتصف القمة يقع عند مركز الفئة.

*** المنحنى التجميعي الصاعد و النازل:**

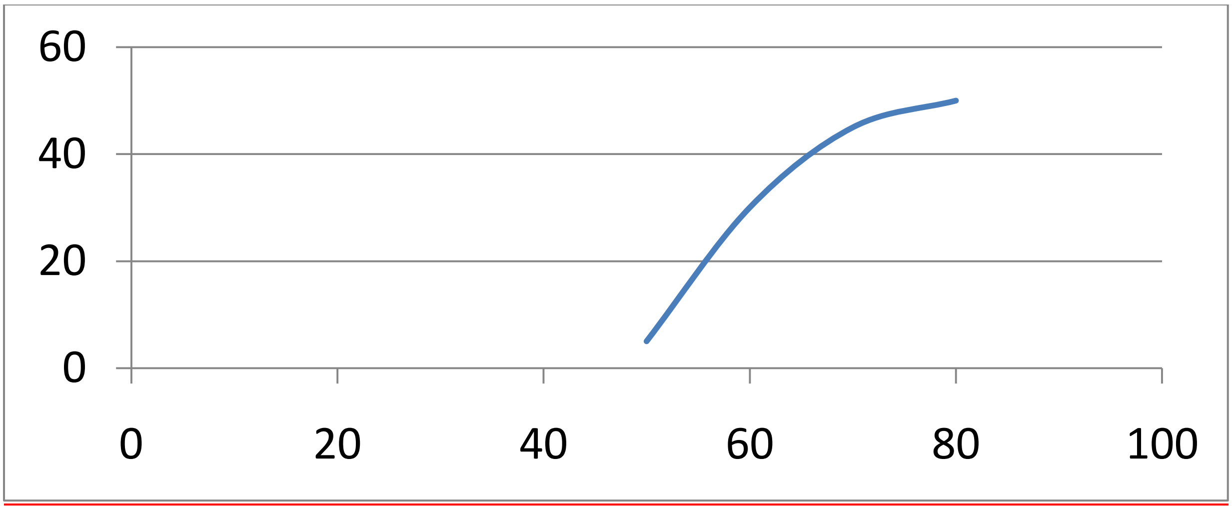
يمكن تمثيل المتغير الكمي المستمر بمنحنين يمثلان التكرار التجميعي الصاعد و النازل في نفس الرسم حيث:

- يرسم **المنحنى التجميعي الصاعد** عن طريق إيصال مجموعة من النقاط التي إحداثياتها الحدود العليا للفئات والتكرارات التجميعية الصاعدة المقابلة لها.

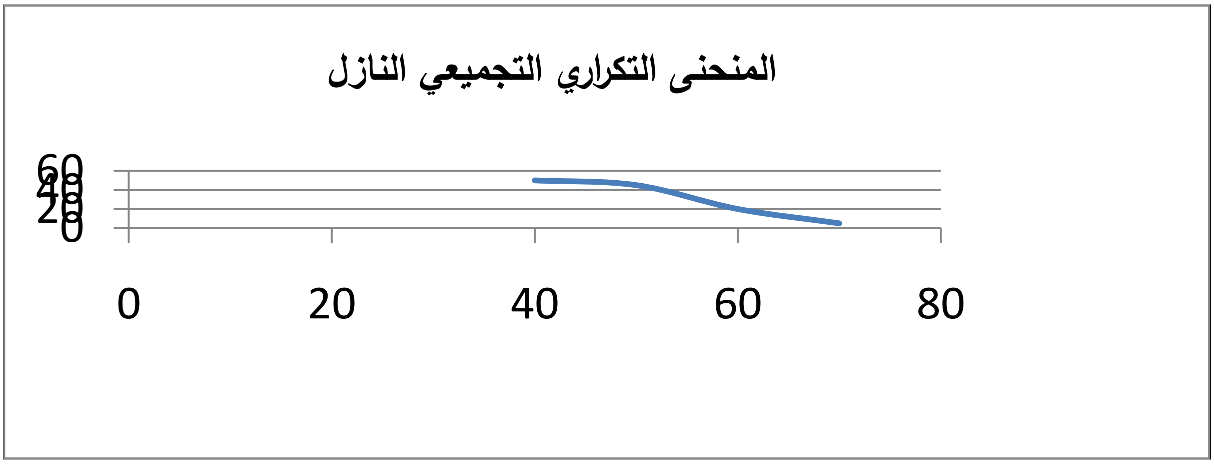
مثال 10: انطلاقا من المثال السابق(وزن 50 شخص)، يوضح الشكل التالي كيفية رسمه كما يلي:

الفئات	40-50	50-60	60-70	70-80	المجموع
n_i	2	25	15	5	**50**
n ↗	5	30	45	50	/
n ↙	50	45	20	5	/

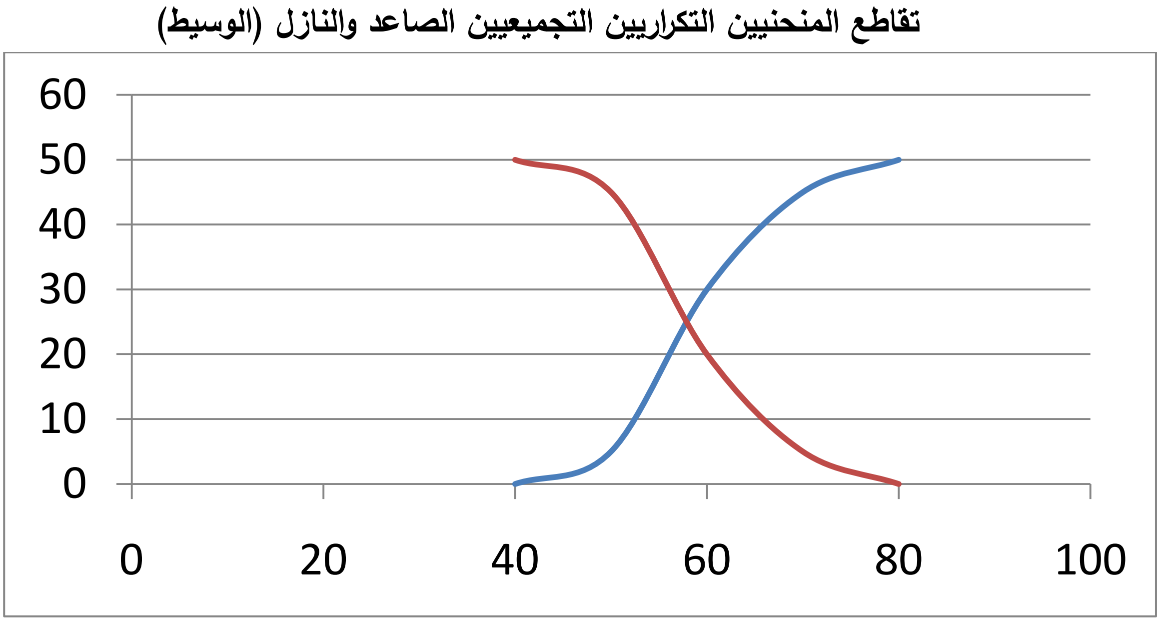
منحنى التكرار التجميعي الصاعد



- **كما يمكن رسم المنحنى التجميعي النازل** عن طريق إيصال مجموعة نقاط التي إحداثياتها الحدود الدنيا للفئات والتكرارات التجميعية النازلة المقابلة لها، كما هو موضح في الشكل التالي:



ملاحظة: فاصلة نقطة تقاطع المنحنيين التجميعي الصاعد و النازل تسمى " بالوسيط "، أما ترتيبها يساوي: $\sum n_i/2$، هذا ما سيبينه الشكل الموالي:



الفصل الثالث

مقاييس النزعة المركزية

الفصل الثالث: مقاييس النزعة المركزية

تمهيد:

لقد تطرقنا في الفصل السابق إلى تصنيف البيانات الإحصائية و تبويبها في جداول تكرارية، وتمثيلها باستخدام أشكال بيانية متعددة بهدف وصف هذه البيانات فقط، لكن الدراسة الإحصائية لا تكتفي بمثل هذه الأمور، بل تحاول أن تلخص أهم صفات هذه البيانات الرقمية في عدد واحد يُرمز لها و يدل عليها، هذا العدد قد يوضح نزعتها للتجمع أو نزعتها للتشتت، فعندما تميل القيم إلى التجمع نحو قيمة معينة تسمى بمتوسط هذه القيم، فيزداد عدد القيم كلما قربت من المتوسط، و يقل عددها كلما بعدت عنه، أي أن للقيم نزعة أو ميلا لقيمة موجودة في مركزها، ويطلق على خاصية تجمع القيم حول نقطة معينة بخاصية **النزعة المركزية** لأنها هي القيمة التي تتجمع حولها باقي القيم.

تعد مقاييس النزعة المركزية من المؤشرات الإحصائية الوصفية التي تُستخدم لوصف بيانات مجموعة ما أو توزيع تكراري معين، من خلال إيجاد قيمة وحيدة تلخص مجموعة من القيم التي تمثلها أفضل تمثيل وتعبر عنها أفضل تعبير، وسميت هذه المؤشرات بمقاييس النزعة المركزية كونها تتمركز في الوسط لهذا أطلق عليها أحيانا مصطلح **المتوسطات** أي هو عملية اختيار قيمة معينة أو التعبير عن مجموعة من القيم بقيمة واحدة، حيث تُستخدم هذه المتوسطات بشكل واسع في موضوع الإستدلال الإحصائي لأهميتها، ولكل متوسط طريقة مختلفة في الحساب، كما له استخدامات مختلفة في الحياة العملية.

وتتمثل مقاييس النزعة المركزية في المقاييس التالية:

- الوسط الحسابي Moyenne arithmétique
- الوسيط Médiane
- المنوال Mode
- الوسط الهندسي Moyenne Géométrique
- الوسط التوافقي Moyenne Harmonique

– الوسط التربيعي Moyenne Quadriennale

وفيما يلي نقوم بشرح مفصل للمقاييس الواردة سالفا على النحو الآتي:

أولا: الوسط الحسابي (Moyenne Arithmétique)

يعتبر الوسط الحسابي من أهم مقاييس النزعة المركزية وأكثرها شيوعا واستخداما لسهولته وفائدته وامتيازه بخصائص جيدة يجعله في مقدمة مقاييس النزعة المركزية، ويسمى أيضا بالمتوسط، يرمز له بـ $\bar{X}$، و يمكن حسابه بعدة طرق حسب اختلاف حالات المتغيرات الإحصائية وخصائصها، إذ يمكن أن نفرق بين حالتين[20]:

1)– الوسط الحسابي في حالة بيانات غير المبوبة: أي في حالة قيم بدون تكرار.

أ)– الطريقة المباشرة (الوسط الحسابي البسيط): يستعمل الوسط الحسابي البسيط عندما يكون لقيم المتغير الإحصائي نفس المستوى من اهمية.

بافتراض لدينا عينة من القيم (المشاهدات) N و هي: $(X_1, X_2, X_3, \ldots\ldots X_n)$

عندئذ يعرف الوسط الحسابي $\bar{X}$ بأنه مجموع قيم المشاهدات مقسما على عددها، أي

$$\bar{X} = \frac{X_1, X_2, X_3, \ldots\ldots X_n}{N}$$

بصورة عامة: $\bar{X} = \dfrac{\sum X_i}{N}$

حيث:

$\bar{X}$: الوسط الحسابي

مجموع X_i: مجموع قيم المتغير الإحصائي

N: عدد القيم

مثال 01 : تمثل القيم التالية كمية الأمطار المتساقطة سنويا(مم) في ولاية الشرق الجزائري للفترة(2007–2011)

.520 , 350 , 410 , 365, 460

المطلوب: إيجاد متوسط كمية الأمطار المتساقطة

[20] جلاطو جيلالي، الاحصاء،المرجع سبق ذكره، ص 30.

الحل:

$$\bar{X} = \frac{\sum X_i}{N} = \frac{520 + 350 + 410 + 365 + 410}{5} = 421 \text{ مم}$$

ب) الطريقة المختصرة (الوسط الحسابي بواسطة الوسط الفرضي)

تستخدم هذه الطريقة عندما يكون لدينا كم هائل من القيم وأعداد كبيرة جدا، يصعب التعامل معها عند إيجاد الوسط الحسابي البسيط لها.

ولحساب الوسط الحسابي الفرضي يجب اتباع الخطوات التالية:

– اختيار أي عدد بصورة فرضية من بين قيم المجموعة X_i، وينصح تجنب القيم المتطرفة و نرمز له بـ X_o

– ايجاد انحرافات قيم المجموعة X عن الوسط الفرضي X_o، ونرمز لها بـ d_i ونطبق القانون التالي:

$$\bar{X} = X_o + \frac{\sum d_i}{N}$$

حيث:

X_o: وسط فرضي يتم اختياره من بين قيم المجموعةX_i ، و يفضل اختيار القيمة التي تتكرر أكثر من غيرها.

d_i: انحرافات قيم المجموعةX_i عن الوسط الفرضي X_o حيث: $d_i = Xi - X_o$

N: عدد القيم

$$d_1 = X_1 - X_o$$
$$d_2 = X_2 - X_o$$
$$d_3 = X_3 - X_o$$
.
$$d_n = X_n - X_o$$

مثال 02: بالعودة إلى المثال السابق

لايجاد متوسط كمية الأمطار المتساقطة باستعمال الطريقة المختصرة، نقوم باختيار $X_o = 410$ كوسط فرضي حيث نجد:

$d_1 = X_1 – X_o= 520 - 410 = + 110$
$d_2 = X_2 – X_o= 350 - 410 = - 60$
$d_3 = X_3 – X_o= 410 - 410 = 0$
$d_4 = X_4 – X_o= 365 - 410 = - 45$
$d_5 = X_5 – X_o= 460 - 410 = + 50$

ومنه: $\sum d_i = 55$

و المتوسط الحسابي يساوي:

$$\bar{X} = X_o + \frac{\sum d_i}{N} = 410 + \frac{55}{5} = 410 + 11 = 421$$ مم

وهي نفس النتيجة التي تحصلنا عليها بالطريقة المباشرة

ج)- طريقة الوسط الحسابي المرجح Moyenne Arithmétique Pondérée

في بعض الأحيان نجد مشاهدات لها اثقال (معاملات) وأوزان مختلفة ونرمز له بـ W_i:، وإذا لم نأخذ بعين الاعتبار هذا الترجيح فلا يمكن للوسط الحسابي أن يعطي صورة تمثيلية صحيحة لهذا المقياس الإحصائي، لذا يجب أن يحسب الوسط الحسابي المرجح كما يلي[21]:

$$\bar{X} = \frac{(X_1 . W_1) + (X_2.W_2) + \ldots + (X_n.W_n)}{W_1 + W_2 + \ldots + W_n}$$

مثال 03: فيما يلي علامات طالب في 5 مقاييس: 16, 18 ,12, 08, 02

فكانت معاملاتها على التوالي: 3 ,4 ,2 ,1,1

المطلوب: احسب الوسط الحسابي لعلامات هذا الطالب

الحل: نلاحظ أن لكل علامة وزن مختلف عن العلامات الأخرى، ولحساب الوسط الحسابي لا بد من ترجيح كل علامة بوزنها ثم نقسم المجموع على مجموع الأوزان كما يلي:

$$\bar{X} = \frac{(X_1 . W_1) + (X_2.W_2) + \ldots + (X_5.W_5)}{W_1 + W_2 + \ldots + W_5}$$

$$\bar{X} = \frac{(16x3)+(18x4)+(12x2)+(08x1)+(2x2)}{3 + 4 + 2 + 1 + 2}$$

[21] نفس المرجع السابق، ص 31.

$$\bar{X} = \frac{48 + 72 + 24 + 8 + 4}{12} = \frac{156}{12} = 13$$

كما يمكن استعمال الوسط الفرضي عند حساب الوسط الحسابي المرجح بفرض $X_o = 10$ نجد أن:

$$\bar{X} = \frac{10 + (16-10)x3 + (18-10)x4 + (12-10)x2 + (8-10)x1 + (2-12)x2}{12}$$

$$\bar{X} = 10 + \frac{18 + 32 + 4 - 2 - 16}{12} = 10 + 3 = 13$$

نفس النتيجة التي تحصلنا عليها سابقا.

2)- الوسط الحسابي في حالة بيانات مبوبة:

لقد رأينا في الفصل السابق كيفية إعداد الجداول التكرارية ووضع البيانات في فئات، كل فئة تشمل مجموعة من القيم تنتمي إلى مجال معين لا نعلم القيم الفعلية للبيانات وإنما نعلم فقط المجال الذي تتراوح فيه القيم عندها.

نعلم أن لكل فئة حد أعلى A_2 وحد أدنى A_1، لذا يمكن حساب متوسط القيم في كل فئة حسب العلاقة التالية:

$$Xi = \frac{A1 + A\ 2}{2}$$

و يعتبر المتوسط المتحصل عليه بمثابة متوسط القيم الموجودة في تلك الفئة و نسميه **بمركز الفئة**، فإذا ضربنا كل مركز بالتكرار المقابل له نحصل على مجموع القيم في تلك الفئة، وبجمع كل هذه المجاميع نتحصل على المجموع الكلي للقيم، وبقسمة هذا الأخير على مجموع التكرارات نتحصل على الوسط الحسابي $\bar{X}$ للجدول ويتم حساب ذلك حسب الطرق التالية:

أ)– الطريقة المباشرة:

إذا كـان لـدينا توزيـع تكـراري معـين، ومراكـز فئاتـه هـي: $(X_1,X_2,\ldots.X_k)$، والتكـرارات المقابلـة لهـذه المراكـز هـي: $(N_1,N_2,\ldots.N_k)$، علـى الترتيـب، عندئـذ يمكـن حسـاب الوسـط الحسابي$(\bar{X})$ بموجب هذه الطريقة اتباع الخطوات التالية:

– إيجاد مراكز الفئات X_i بحيث:

$$Xi = \frac{A_1 + A_2}{2}$$

مع A_1: الحد الأدنى للفئة

A_2: الحد الأعلى للفئة

– إيجاد مجموع حاصل ضرب مراكز الفئات في التكرارات المقابلة لها، اي: $\sum x_i \,.\, n_i$

– تطبيق العلاقة التالية:

$$\bar{X} = \frac{\sum x_i \,.\, n_i}{\sum n_i}$$

ب)– طريقة الانحرافات عن الوسط الفرضي:

تستعمل هذه الطريقـة فقـط فـي حالـة الجـداول المنتظمـة أي طـول الفئـات متسـاوية، فقد تصادفنا أحيانا بيانات كبيرة من حيث الحجم و كثيرة من حيث العدد، ولحساب الوسط الحسابي بالطريقـة المباشـرة تنـتج لـدينا أرقـام كبيـرة جـدا، و لتبسـيط الحسـاب و تسـهيل العمليـات الحسـابية نستعمل **طريقـة الوسـط الفرضـي**، حيث نفرض وسطا حسابيا يفضل ان يكون من أحد مراكز الفئات و هذا لتسهيل العمليات الحسابية لا غير.

فالصيغة الأولى المستعملة سابقا في حالة البيانات غير المبوبة كانت كما يلي:

$$\bar{X} = x_o + \frac{\sum d_i}{N}$$

اما في هذه الحالة بالنسبة للبيانات المبوبة فتصبح كما يلي:

$$\bar{X} = x_o + \frac{\sum n_i \,.\, d_i}{\sum ni}$$

أي ضربنا كل إنحراف بالتكرار المقابل له.

ويجب أن تتوفر في الوسط الفرضي الشروط التالية:

– يجب أن يكون سهلا بحيث يسهل طرحه من قيم X_i.

– يجب أن يكون عدد القيم التي اصغر منه لا يختلف كثيرا عن عدد القيم التي أكبر منه، و بالتالي يكون هناك انحرافات سالبة وانحرافات موجبة.

– يجب ان يكون مساويا لإحدى قيم مراكز الفئات و بالتالي فإحدى الإنحرافات تكون مساوية للصفر.

ولإيجاد قيمة الوسط الحسابي حسب هذه الطريقة نطبق العلاقة التالية:

$$\bar{X} = x_o + \frac{\sum n_i (X_i - X_o)}{\sum ni}$$

ج)- الطريقة المختصرة:

تُطبق هذه الطريقة في الجداول التكرارية المنتظمة فقط، وسميت بطريقة الإنحرافات المختصرة لأنها تقوم على فكرة إختصار أو قسمة الإنحرافات الناتجة عن الوسط الفرضي على طول الفئة، وحيث يجب اتبع الخطوات التالية:

– يجب إيجاد مراكز الفئات.

– نختار وسطا فرضيا مناسبا X_o، وعادة ما يكون الأكبر تكرارا، ثم نبحث عن انحرافات مراكز الفئات عن الوسط الفرضي $X_i - X_o$.

– إذا كانت أطوال الفئات متساوية، في هذه الحالة نأخذ طول الفئة C كعامل مشترك، وإذا لم يكن هناك عامل مشترك فلا يمكن تطبيق طريقة الإنحرافات المختصرة، ويلجأ إلى الطريقة المباشرة أو طريقة الانحرافات فقط.

– نطبق العلاقة الإحصائية التالية:

$$\bar{X} = x_o + \frac{\sum n_i (X_i - X_o) / C}{\sum ni} \times C$$

مثال 04: إذا كان لدينا الجدول التكراري التالي بحيث طول الفئات يساوي 2، و X_O يساوي 07.

احسب الوسط الحسابي بطرقه الثلاث

الفئات	N_i	X_i	$N_i . X_i$	$X_i - X_O$	$N_i (X_i - X_O)$	$N_i (X_i - X_O)/C$
0−2	3	1	3	1−7= −6	2 x (−6) = −18	− 18/2 = − 9
2−4	6	3	18	3−7= −4	− 24	− 24/2 = − 12
4−6	9	5	45	5−7= −2	− 18	− 18/2 = − 9
6−8	10	7	70	7−7=0	0	0
8−10	8	9	72	9−7=2	16	16/2 = 8
10−12	7	11	77	11−7=4	28	28/2 = 14
12−14	2	13	26	13−7=6	12	12/2 = 6
المجموع	**45**	/	**311**	/	**− 4**	**− 2**

الحل

الطريقة الأولى: المباشرة

هو مجموع مركز كل فئة مضروبا في تكرارها ثم نقسم عدد التكرارات فنتحصل على الوسط الحسابي، والجدول أعلاه يبين الحسابات اللازمة لذلك.

إذن الوسط الحسابي يساوي:

$$\bar{X} = \frac{\sum x_i . n_i}{\sum n_i} = \frac{311}{45} = 6,91$$

نلاحظ أن قيمة الوسط الحسابي محصورة بين مجال القيم المعطاة فهي محصورة بين الحد الأدنى للفئة الأولى أي القيمة 0، و الحد الأعلى للفئة الأخيرة أي القيمة 14.

الطريقة الثانية: طريقة انحرافات في الوسط الفرضي

نفرض أن الوسط الفرضي هو 7 (كما هو مبين في المعطيات للتمرين)، إذن الوسط الحسابي يساوي

$$\bar{X} = x_o + \frac{\sum n_i (X_i - x_o)}{\sum ni} = 7 + \frac{(-4)}{45} = 6,91$$

وهي نفس النتيجة المتحصل عليها في الطريقة المباشرة.

3)- الطريقة الثالثة: الطريقة المختصرة

لحساب الوسط الحسابي في هذه الحالة، نقوم بقسمة الانحرافات التي قمنا بحسابها على طول الفئة فتحصل على النتائج الموجودة في الجدول أعلاه، و بالتالي فالوسط الحسابي يساوي:

$$\bar{X} = X_o + \frac{\sum n_i (X_i - X_o) / C}{\sum ni} \times C$$

$$\bar{X} = 7 + \frac{(-2) \times 2}{45} = 6{,}91$$

وهي نفس النتيجة المتحصل عليها في الطريقتين السابقتين.

خصائص الوسط الحسابي:

* يعتبر الأكثر المتوسطات إستخداما لأنه سهل وبسيط ويخضع لجميع العمليات الجبرية.

* يتأثر الوسيط الحسابي بالقيم الشاذة أو المتطرفة (القيم التي تقع في طرفي مجال الدراسة).

* لا يمكن إيجاده في الجداول التكرارية المفتوحة، كما لا يمكن تحديده بيانيا.

* يستعمل الوسط الحسابي في حالة المتغيرات الكمية فقط أي القابلة للقياس.

* لا يمكن أن يكون لأي توزيع تكراري أكثر من وسط حسابي، إذن يوجد وسط حسابي وحيد بالنسبة لتوزيع تكراري معين أو بالنسبة لسلسلة إحصائية معينة.

* أساس حساب الوسط الحسابي هو الحساب التجميعي .

* لا تكون للوسط الحسابي قيمة من بين قيم المتغير الإحصائي المستعملة في حسابه إلا نادرا.

* نستنتج أن من علاقة $\bar{X}$ أن $\sum Xi = n\,\bar{X}$

* إن مجموع انحرافات قيم المتغير الإحصائي بالنسبة للوسط الحسابي تساوي الصفر، ونرمز لهذه الإنحرافات بالرمز e_i

حيث: $\sum e_i = \sum (X_i - \bar{X}) = O$.

مثال 05: إذا كانت لدينا السلسلة الإحصائية التالية: 50-35-30-45

$$\bar{X} = \frac{45 + 30 + 35 + 50}{4} = \frac{160}{4} = 40$$

$$\sum e_i = \sum (X_i - \bar{X}) = O$$

$$= (45-40) + (30-40) + (35-40) + (50-40)$$

$= 5 + (-10) + (-5) + 10 = 0$ صفر

* عند إضافة كمية ثابتة و لتكن a إلى كل قيمة من قيم المتغير X_i، أي أن: $Y_i = X_i + a$

فإن الوسط الحسابي $\bar{Y}$ للقيم الجديدة عبارة عن الوسط الحسابي للقيم الأصلية مضافا له قيمة الثابت a بحيث:

$$\bar{Y} = \bar{X} + a$$

مثال 06: لتكن لدينا بيانات المتغير X_i الآتية: 10، 9، 12، 11، 8

نريد إثبات صحة العلاقة: $\bar{Y} = \bar{X} + a$ بعد افتراض قيمة $a = 7$

الحل:

– نقوم أولا بإضافة قيمة الثابت $a = 7$ إلى كل قيمة من قيم المتغير X_i أي أن $X_i + 7 = Y_i$، عندئذ تكون قيم المتغير الجديدة Y_i كالآتي: 17، 16، 19، 18، 15 $= Y_i$

$$\bar{Y} = \frac{\sum y_i}{N} = \frac{15 + 18 + 19 + 16 + 17}{5} = 17$$

ـ ثم نقوم بإيجاد الوسط الحسابي $\bar{X}$ للقيم الأصلية X_i نتحصل على القيم التالية:

$$\bar{X} = \frac{\sum x_i}{N} = \frac{8 + 11 + 12 + 9 + 10}{5} = \frac{50}{5} = 10$$

وعليه فإن الوسط الحسابي $\bar{Y}$ للقيم الجديدة Y_i هو في الواقع عبارة عن:

$$\bar{Y} = \bar{X} + 7 = 10 + 7 = 17$$

* عند طرح كمية ثابتة، و لتكن(b) من كل قيمة من قيم المتغير X_i، أي أن: $Y_i = X_i - b$.

فإن الوسط الحسابي $\bar{Y}$ للقيم الجديدة عبارة عن الوسط الحسابي للقيم الأصلية مطروحا منه قيمة الثابت b بحيث:

$$\bar{Y} = \bar{X} - b$$

مثال 07: لتكن لدينا بيانات المتغير X_i الآتية: 7، 10، 13، 9، 6

نريد إثبات صحة العلاقة $\bar{Y} = \bar{X} - b$، بعد افتراض قيمة $b = 5$

الحل:

– نقوم أولا بطرح قيمة الثابت $b=5$ من كل قيمة من قيم المتغير X_i

أي أن $Y_i = X_i - 5$، عندئذ تكون قيم المتغير الجديد Y_i كالآتي:

Y_i : 1 ،4 ،8 ،5 ،2

$$\bar{Y} = \frac{\sum Yi}{N} = \frac{1 + 4 + 8 + 5 + 2}{5} = \frac{20}{5} = 4$$

ثم نقوم بإيجاد الوسط الحسابي $\bar{X}$ للقيم الأصلية X_i كالآتي:

$$\bar{X} = \frac{\sum Xi}{N} = \frac{6 + 9 + 13 + 10 + 7}{5} = \frac{45}{5} = 9$$

$$\bar{Y} = \bar{X} - 5 = 9 - 5 = 4$$

* عند ضرب كل قيمة من قيم المتغير X_i بكمية ثابتة C، أي أن: $Y_i = C . X_i$

فإن الوسط الحسابي $\bar{Y}$ للقيم الجديدة عبارة عن الوسط الحسابي للقيم الأصلية $\bar{X}$ مضروبا في قيمة الثابت C أي:

$$\bar{Y} = C\,\bar{X}$$

مثال 08: إذا كان لدينا البيانات التالية: 20، 25، 10، 12، 13

الحل: نريد إثبات صحة العلاقة $\bar{Y} = C\,\bar{X}$ بعد افتراض قيمة $C = 2$.

– نقوم أولا بضرب كل قيمة من قيم المتغير X_i بقيمة الثابت $C = 2$، أي أن $Y_i = 2X_i$،

عندئذ تكون قيم المتغير الجديد Y_i تساوي: (2x20)، (2x25)، (2x10)، (2x12)، (2x13)

Y_i =

Y_i = 26 ،24 ،20 ،50 ،40

ومنه:

$$\bar{Y} = \frac{\sum Yi}{N} = \frac{160}{5} = 32$$

– ثم نقوم بإيجاد الوسط الحسابي $\bar{X}$ للقيم الأصلية X_i كما يلي:

$$\bar{X} = \frac{\sum Xi}{5} = \frac{80}{5} = 16$$

– وعليه فإن الوسط الحسابي $\bar{Y}$ للقيم الجديدة هو:

$$\bar{Y} = C\,\bar{X} = 2 \times 16 = 32$$

* عند قسمة كل قيمة من قيم المتغير X_i على كمية ثابتة d أي:

$$Yi = X_I \div d$$

فإن الوسط الحسابي $\bar{Y}$ للقيم الجديدة عبارة عن الوسط الحسابي $\bar{X}$ للقيم الأصلية مقسوما على قيمة الثابت d، أي:

$$\bar{Y} = \bar{X} \div d$$

ثانيا: الوسيط LA MEDIANE

يعد الوسيط من أفضل مقاييس النزعة المركزية وأكثرها استخداما في وصف بيانات المجموعات أو التوزيعات التكرارية التي تحتوي على قيم شاذة أو متطرفة، كونه لا يتأثر بهذا النوع من القيم[22].

و يعرف بأنه القيمة التي تفصل السلسلة الإحصائية إلى قسمين متساويين بعد ترتيب القيم تصاعديا أو تنازليا و يرمز له ب، Me، و يتم حسابه كالآتي:

1)-الوسيط في حالة بيانات غير مبوبة:

إذا كان لدينا عدد معين من المعطيات N، لحساب الوسيط نقوم بترتيب القيم ترتيبا تصاعديا أو تنازليا، ثم يتم تحديد ترتيب الوسيط، وهنا نُميّز بين حالتين:

أ)- عندما يكون عدد القيم فردي، فإن الوسيط هو قيمة المتغير الإحصائي الذي يشغل الرتبة: $\frac{N+1}{2}$

مثال 09: أوجد الوسيط للبيانات التالية: 4، 2، 3، 10، 8، 6، 11

الحل: نقوم أولا بترتيب القيم ترتيبا تصاعديا كما يلي: 2، 3، 4، 6، 8، 10، 11.

ثم نحسب ترتيب الوسيط أي موقعه و الذي يساوي: $\frac{N+1}{2} = \frac{7+1}{2} = \frac{8}{2} = 4$

بحيث: N = 7 ، إذن رتبة الوسيط هي 4.

نلاحظ أن القيمة التي رتبتها 4 بين القيم هي القيمة 6 و هي قيمة الوسيط.

ب)- عندما يكون عدد القيم زوجي، فإن الوسيط في هذه الحالة يكون لدينا " مجال وسيط " و ليس قيمة منفردة كما في الحالة السابقة.

وتكون قيمة الوسيط هنا عبارة عن الوسط الحسابي ذات الترتيب : N/2 و $\frac{N}{2}+1$ على التوالي،

[22] حسن ياسين طعمة ، طرق الإحصاء الوصفي، المرجع سبق ذكره، ص 129.

حيث قيمة الوسيط تساوي: $\frac{N/2 + (N/2 + 1)}{2}$

مثال 10: البيانات التالية تمثل أعمار 8 طلبة. 23، 17، 22، 19، 18، 24، 21، 16

المطلوب: أحسب العمر الوسيط Me

الحل: لحساب العمر الوسيط، نتبع الخطوات التالية:

1)- ترتيب القيم ترتيبا تصاعديا كالآتي: $16 < 17 < 18 < 19 < 21 < 22 < 23 < 24$

Me

2)- بما أن عدد القيم زوجي (N = 8)، إذن توجد قيمتان وسيطتان حيث ترتيبهما هو:

* ترتيب قيمة الوسيط الأول: $\frac{N}{2} = \frac{8}{2} = 4$ الرتبة الرابعة

* ترتيب قيمة الوسيط الثاني: $\frac{N}{2} + 1 = \frac{8}{2} + 1 = 5$ الرتبة الخامسة

3)-تحديد القيمتين اللتين يقع بينهما الوسط وهما: 19 و 21

4)- حساب قيمة الوسيط: $Me = \frac{19 + 21}{2} = 20$

ملاحظة: عند وجود قيم السلسلة المنفصلة و المعبر عنها بجدول تكراري، في هذه الحالة حساب الوسيط من الجدول الإحصائي يكون كالآتي:

– نقوم بقسمة N على 2 أي حساب N/2

– ثم نقارن النتيجة المتحصل عليها مع قيم N↗ و نأخذ القيمة التي توافق نتيجة N / 2 أو أكبر منها.

مع رتبة الوسيط Me هي $\sum n_i / 2$

مثال11 : لدينا التوزيع التكراري التالي:

المطلوب: أحسب الوسيط

Xi	1	2	3	4	**المجموع**
ni	20	16	10	4	**50**
N↗	20	36	46	50	/

$N / 2 = 50 / 2 = 25$

القيمة 25 ليست موجودة ضمن التكرارات المتجمعة الصاعدة، إذن الوسيط يقابل القيمة 2 في الجدول ومنه: $Me=2$

بيانيا: تقاطع المنحنيين للتكرار التجميعي الصاعد و النازل يؤدي إلى نقطة واسقاطها على محور القيم نستنتج الوسيط

2) الوسيط في حالة بيانات مبوبة أو حالة توزيع تكراري

يتم حساب الوسيط في حالتين هما:

أ)- الوسيط في حالة التكرار التجميعي الصاعد:

حسب هذه الحالة يجب اتباع الخطوات التالية:

- تحديد التكرار التجميعي الصاعد.
- تحديد رتبة الوسيط، وهو عبارة عن نصف مجموع التكرارات أي: $\sum n_i / 2$
- تحديد فئة الوسيط، اي الفئة التي يقع فيها الوسيط، وهي الفئة التي تقابل التكرار التجميعي الصاعد الذي يساوي ترتيب الوسيط او أكبر منه مباشرة.
- تحديد وحساب الوسيط بتطبيق العلاقة الإحصائية التالية:

$$Me = A_1 + \frac{\sum n_i / 2 - N\nearrow_{(n-1)}}{N_{me}} \times C$$

حيث:

A_1: الحد الأدنى للفئة الوسيطية

$\sum n_i / 2$: ترتيب الوسيط

$N\nearrow_{(n-1)}$: التكرار التجميعي الصاعد الذي يسبق التكرار التجميعي الصاعد للفئة الوسيطية.

N_{me} : التكرار الأصلي أو العادي المقابل للفئة الوسيطة.

C: طول الفئة الوسيطية

ملاحظة: يستخدم هذا القانون في حساب الوسيط في حالة التوزيعات التكرارية المنتظمة و غير المنتظمة على حد سواء، بغض النظر كون الجداول مغلقة أو مفتوحة

كما يمكن حساب الوسيط عن طريق التكرارات النسبية المئوية، ويصبح القانون كالآتي:

$$Me = A_1 + \frac{\sum n_i / 2 - N\nearrow_{(n-1)} \%}{N_{me} \%} \times C$$

حيث:

A_1: الحد الأدنى للفئة الوسيطية

$N\nearrow_{(n-1)} \%$: التكرار التجميعي الصاعد بقيم مئوية للفئة التي تسبق فئة الوسيط.

$N_{me} \%$: التكرار النسبي لفئة الوسيط بقيم مئوية.

C: طول الفئة.

مثال 12: ليكن الجدول التكراري التالي:

الفئات	20−40	40−50	50−60	60−70	70−80	المجموع
التكرارات	25	30	54	17	12	**140**
ت ت ص	25	57	111	128	140	/
ت ت ن	140	115	83	29	12	/

المطلوب: تحديد قيمة الوسيط

الحل:

1)− تحديد التكرار التجميعي الصاعد في العمود 3 من الجدول

2)− ترتيب الوسيط : $\sum n_i / 2 = 140 / 2 = 70$

3)− تحديد الفئة الوسيطية: و هي الفئة التي تقابل التكرار التجميعي الصاعد الذي يساوي 70 أو أكبر من 70 مباشرة، نلاحظ أن العدد 70 غير موجود من بين قيم التكرار التجميعي الصاعد، ولكن العدد الأكبر منه مباشرة هو 111. إذن القيمة التي تقابل التكرار التجميعي الصاعد قيمة 111 هي الفئة]50−60[.

4)− نحسب الوسيط بتطبيق المعادلة الإحصائية التالية:

$$Me = A_1 + \frac{\sum n_i / 2 - N\nearrow_{(n-1)}}{N_{me}} \times C$$

$$Me = 50 + \frac{70 - 57}{54} \times 10 = 52{,}40$$

ب)− الوسيط في حالة التكرار التجميعي النازل:

يتم حساب الوسيط في هذه الحالة حسب الخطوات التالية:

– تحديد التكرار التجميعي النازل.

– تحديد ترتيب الوسيط و هو: $\sum n_i / 2$

– تحديد قيمة الوسيط، أي الفئة التي تقابل التكرار التجميعي النازل الذي يساوي ترتيب الوسيط أو أكبر منه مباشرة.

– حساب الوسيط بتطبيق العلاقة الإحصائية التالية:

$$Me = A_2 - \frac{\sum n_i / 2 - N\swarrow_{(n+1)}}{N_{me}} \times C$$

حيث:

A_2: الحد الأعلى للفئة الوسيطية.

$N\swarrow(n+1)$: التكرار التجميعي النازل اللاحق للفئة الوسيطية

مثال 13: بالرجوع للمثال السابق نجد:

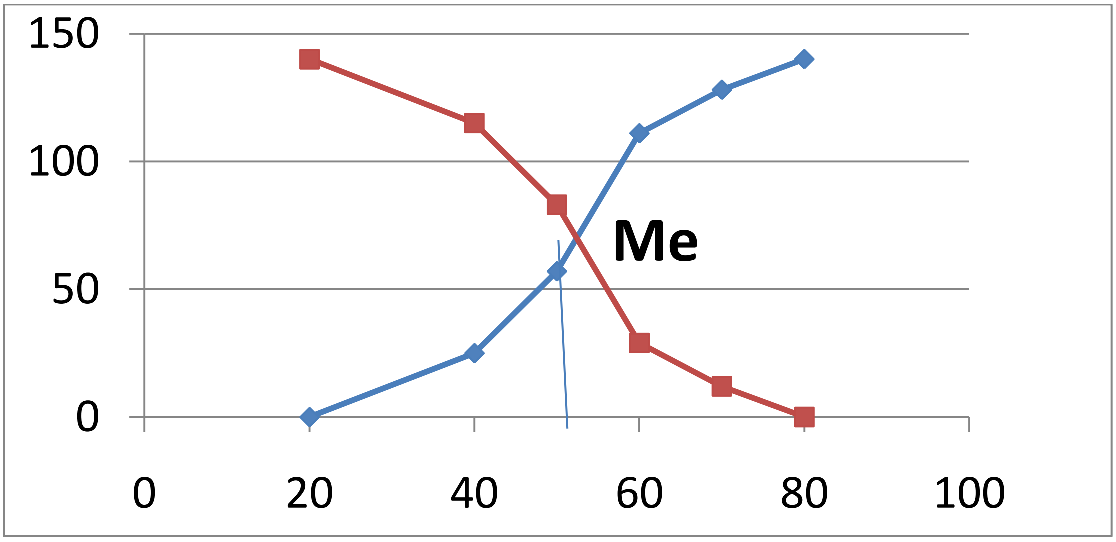
$$Me = 60 - \frac{70 - 29}{54} \times 10 = 52{,}40$$

ج)– إيجاد الوسيط بيانيا:

يمكن إيجاد قيمة الوسيط باستخدام الطريقة البيانية وفقا للخطوات التالية:

– رسم منحنيي التكرار التجميعي الصاعد و النازل، وتمثل إحداثيات نقطة تقاطع المنحنيين قيمة و ترتيب الوسيط

وهذا ما نجده بيانيا استنادا للمثال السابق.



نلاحظ من الرسم البياني أن فاصلة نقطة المنحنيين هي 52,40 وتمثل قيمة الوسيط، أما ترتيب نقطة التقاطع هي: 70 وتمثل ترتيب الوسيط و التي تساوي $\sum n_i / 2$

د)- مزايا وعيوب الوسيط

* **مزايا الوسيط:** يتميز الوسيط بعدد من المزايا تتمثل بما يلي[23]:

- يعد الوسيط من أفضل مقاييس النزعة المركزية وأكثرها تمثيلا بالنسبة للتوزيعات التي تحتوي على قيم شاذة (متطرفة)، لأنه لا يتأثر إطلاقا بهذا النوع من القيم .
- بساطة فكرة الوسيط لأنه سهل الفهم و الحساب.
- يمكن إيجاد بيانا باستخدام المنحنيين التجميعيين الصاعد و النازل.
- يمكن تقدير قيمة الوسيط عن طريق التخمين والتأمل.
- يمكن إيجاده في الجداول التكرارية المغلقة أو المفتوحة من طرف واحد أو طرفين على حد سواء.
- يمكن إيجاد الوسيط للبيانات الكيفية بشرط إمكانية ترتيبها تصاعديا أو تنازليًا وأن يكون عددها فرديًا.

* **عيوب الوسيط:** نذكر منها:

- لا يمكن إيجاد الوسيط للتوزيعات التكرارية التي يكون فيها تكرار الفئة الأولى(n_1) أكبر من ترتيب الوسيط ($\underline{\sum n_i / 2}$)، لكن يمكن إيجاده بطريقة غير مباشرة من خلال العلاقة التقريبية بين المتوسطات الثلاثة: $\bar{X} - M_o = 3(\bar{X} - M_e)$
- لا يخضع الوسيط للعمليات الجبرية بشكل مطلق عند احتسابه.
- لا يستند الوسيط عند احتسابه على كافة البيانات المتاحة في حالة البيانات الغير المبوبة وكذلك في حالة التوزيعات التكرارية.

3)- مقاييس أخرى من عائلة الوسيط:

وهي مقاييس أخرى شبيهة بالوسيط، فإذا كان الوسيط يقسم السلسلة الإحصائية إلى قسمين متساويين فهناك مقاييس أخرى تقسم المعطيات إلى أربعة أجزاء متساوية، أو إلى عشرة أو إلى مائة، كما يلي[24]:

[23] حسن ياسين طعمة، طرق الإحصاء الوصفي، المرجع سبق ذكره، ص 142.

[24] محمد جبر المغربي، الإحصاء الوصفي، المرجع سبق ذكره ، ص 171.

أ)- الربيعات LES QUARTILES

هي التي تقسم المعطيات إلى 04 أجزاء متساوية، كل جزء يمثل 25 % من المعطيات أو ربع المعطيات وتسمى بالربيعي الأول، الثاني، الثالث، ويرمز لهم بـ: Q_1، Q_2، Q_3، و الربيعي الثاني Q_2 هو نفسه الوسيط Me.

بحيث:

- الربيعي الأول Q_1، ترتيبه في البيانات غير المبوبة $(N+1)/4$ و البيانات المبوبة $\sum n_i/4$ يقسم البيانات إلى قسمين 25 % أقل منه 75 % أكبر منه.

- أما الربيعي الثالث Q_3، و الذي ترتيبه البيانات غير المبوبة $3(N+1)/4$،

وفي البيانات المبوبة $3 \sum n_i/4$ يقسم البيانات إلى قسمين 75 % أقل منه 25 % أكبر منه.

- و الربيعي الثاني Q_2 ، هو نفسه الوسيط

ب)-العشيرات LES DECILES

كما تسمى أيضا بالعشريات التي تقسم البيانات إلى 10 أجزاء متساوية، كل جزء يمثل 10 % من المعطيات أو عُشر المعلومات وتسمى بالعشير الأول و الثاني و ... التاسع الذي يرمز لهم بـ D_1، D_2، D_3، ...،D_9، و العشير الخامس D_5 هو نفسه الوسيط Me و الربيعي الثاني Q_2،

بحيث:

- العشير الأول D_1، ترتيبه في البيانات غير المبوبة $(N+1)/10$ و في البيانات المبوبة $\sum n_i/10$ ، يقسم البيانات إلى قسمين 10 %أقل منه، و 90 % أكبر منه.

- أما العشير الثاني D_2، ترتيبه في البيانات غير المبوبة $2(N+1)/10$ و في البيانات المبوبة $2 \sum n_i/10$، يقسم البيانات إلى قسمين 20 % أقل منه، و 80 % أكبر منه.

- وهكذا حتى نصل إلى العشير التاسع أو الأخير D_9، الذي ترتيبه في البيانات غير المبوبة $9(N+1)/10$

و في البيانات المبوبة $9 \sum n_i/10$ ،حيث يقسم البيانات إلى قسمين 90 % أقل منه و 10% أكبر منه.

ج)-المئويات LES PERCENTILES

كما تسمى أيضا المئينات التي تقسم البيانات إلى 100 أجزاء متساوية، لكل جزء يمثل 1 % من المعطيات و تسمى بالمئين الأول، الثاني، ... حتى المئين التاسع و التسعين، و الذي يرمز لهم بـ: P_1، P_2، P_3، ....،P_{99}، و المئين الخمسين P_{50} و هو نفسه الربيعي الثاني Q_2 و الوسيط M_e.

بحيث:

–المئين الأول P_1 ترتيبه في البيانات غير المبوبة، $(N+1)/100$ و في البيانات المبوبة $\sum n_i/100$ ، يقسم البيانات إلى قسمين: 1 % أقل منه، و 99 % أكبر منه.

– أما المئين الثاني P2 ترتيبه في البيانات غير المبوبة $2(N+1)/100$

وفي البيانات المبوبة $2\sum n_i/100$، يقسم البيانات إلى قسمين 2 % أقل منه، و 98 % أكبر منه.

– وهكذا حتى نصل إلى المئين التاسع و التسعين أو الأخير P_{99}، الذي ترتيبه في البيانات غير المبوبة $99(N+1)/100$ و في البيانات المبوبة $99\sum n_i/100$ يقسم البيانات إلى 99 % أقل منه، و 1% أكبر منه.

كما نجد المئين الخامس و العشرين P_{25} هو نفسه الربيعي الأول Q_1، و المئين P_{75} هو نفسه الربيعي الثالث Q_3.

$P_{25} =$ (

$P_{75} =$ (

أما في حالة البيانات المبوبة يمكن حساب كل من الربيعيات أو العُشيرات أو المئيات بتعميم طريقة حساب الوسيط، حيث يمكن تحديد وحساب قيمة المتغير الإحصائي إذا عرفت رتبته و هي إستخدام الصيغة التالية:

مثال 14: حالة عدد البيانات فردي

أحسب الربيعي الأول و الثالث للبيانات التالية التي تمثل أعمار 7 أشخاص:

17، 21، 19، 18، 24، 20، 16

الحل: ترتيب الأعمار ترتيبا تصاعديا

24، 21، 20، 19، 18، 17، 16

– إيجاد ترتيب Q_1: $\frac{N+1}{4} = \frac{7+1}{4} = 2$ إذن قيمة Q_1 هي 17

– إيجاد ترتيب Q_3: $3\left(\frac{N+1}{4}\right) = 3\left(\frac{7+1}{4}\right) = 6$ إذن قيمة Q_3 هي 21

مثال 15: في حالة عدد البيانات زوجي

أحسب كل من كل من Q_3، D_6، P_{65} للبيانات التالية التي تمثل علامات 8 طلبة في مقياس الإحصاء:

4، 2، 9، 12، 15، 6، 8، 17

المطلوب: حساب كل من Q_3، D_6، P_{65}

<u>الحل</u>: نقوم أولا بالترتيب: 15، 12، 9، 8، 6، 4، 2، 17

<u>1)- حساب الربيعي الثالث Q_3</u>

- ترتيب Q_3: $3\frac{(N+1)}{4} = 3\frac{(8+1)}{4} = 6{,}75$ إذن رتبة Q_3 محصورة بين 06 و 07

أي بين العلامة 12 و 15 بحيث يبعد عن القيمة السادسة بـ 75 % من المسافة الفاصلة بين القيمة السادسة و السابعة، إذن قيمة الربيعي الثالث Q_3:

$$Q_3 = 12 + 0{,}75\ (15-12) = 12 + 2{,}25 = 14{,}25$$

2)-حساب العشري السادس D_6:

- ترتيب D_6: $6\frac{(N+1)}{10} = 6\frac{(8+1)}{10} = 5{,}4$

ومنه قيمة D_6 هي واقعة بين القيمة الخامسة و السادسة أي بين 9 و 12 على بعد 0,4 من المسافة الواقعة بينهما أي:

$$D_6 = 9 + 0{,}4\ (12 - 9) = 9 + 1{,}2 = 10{,}2$$

3)- حساب المئين الخامس و الستون P_{65}:

- ترتيب P_{65} : $\frac{65(N+1)}{100} = \frac{65(8+1)}{100} = 5{,}85$

ومنه قيمة P_{65} محصورة بين الرتبتين الخامسة و السادسة على بعد 0,85 من المسافة الواقعة بينهما أي:

$$P_{65} = 9 + 0{,}85\ (12 - 9) = 9 + 2{,}55 = 11{,}55$$

<u>ثالثا</u>: المنوال

يمثل المنوال القيمة التي تقابل أكبر تكرار مطلق أو اكبر تكرار نسبي في السلسة الإحصائية ، و بالتالي هو قيمة المتغير الإحصائي الأكثر إنتشاراً أو شيوعا أو القيمة المسيطرة أو بمعنى أدقّ هو النقطة التي تدلّ على أكثر قيم التوزيع تكرارا، أي القيم التي تكرر أكثر من

غيرها في المجموعة أو التوزيع[25]. ويرمز له بـ Mo، و حسابه يفضل على غيره كونه سهل الحساب و لا يتأثر بالقيم الشاذة و المتطرفة.

يمكن إيجاد المنوال في حالة البيانات المبوبة و غير المبوبة، المغلقة و المفتوحة على حد سواء، و على النحو التالي:

1)- المنوال في حالة البيانات غير المبوبة:

حسب هذه الحالة ووفقا للتعريف السابق، نجد 4 حالات ممكنة لقيم المنوال و هي[26]:

<u>مثال</u>: أحسب المنوال لكل مجموعة

المجموعة 01: 10، 20، 30، 40 50، 70، 90

المجموعة 02: 10، 20، 15، 18، 20، 20، 17

المجموعة 03: 19، 19، 18، 19، 20، 18، 15

المجموعة 04: 09، 12، 17، 15، 12، 09، 15

الحل:

المجموعة الأولى: لا يوجد منوال، وذلك لعدم وجود قيم تتكرر أكثر من غيرها، في هذه الحالة كل قيم المتغير لها نفس الأهمية (فهو توزيع عديم المنوال).

المجموعة الثانية: يود منوال واحد لأن القيمة 20 هي القيم الأكثر تكرارا (فهو توزيع وحيد المنوال)

المجموعة الثالثة: قيمة المنوال التي تساوي 19 لأنها القيمة التي تكررت أكثر من غيرها، كما يوجد منوال آخر يساوي 18 التي تكررت هي الأخرى في المجموعة، إذن هناك منوالان ويسمى (توزيع ثنائي المنوال).

المجموعة الرابعة: ونجد القيم 9، 12، 15 لهم نفس التكرار، إذن هناك أكثر من منوالين لها يسمى (توزيع بمتعدد المنوال)

---

25 حسن طعمة، طرق الإحصاء الوصفي، المرجع سبق ذكره ص 143.

26 جلاطو جيلالي، الإحصاء، المرجع سبق ذكره، ص 50.

2)- حساب المنول في حالة البيانات المبوبة:

يمكن إيجاد المنوال من البيانات المبوبة بعدّة طرق وهي[27]:

أ)- طريقة مركز الفئة المنوالية:

تعدّ هذه الطريقة من أبسط طرق إيجاد المنوال، وأقلّها دقّة، لأن قيمة المنوال المحسوبة تعدّ قيمة تقريبية لأنها تتغيّر بتغيّر مدى وحدود الفئات.

ولإيجاد المنوال للتوزيعات وفقا لهذه الطريقة يجب إيجاد الفئة المنوالية وهي التي يقابلها أكبر تكرار في التوزيع، ثم حساب المنوال الذي هو مركز هذه الفئة، وفقا للصيغة التالية:

$$M_o = \frac{A1+A2}{2}$$

حيث:

A_1: يمثل الحد الأدنى للفئة المنوالية

A_2: يمثل الحد الأعلى للفئة المنوالية

مثال16: يمثل الجدول التالي توزيع 100 عائلة فلاحية حسب ملكيتها من أشجار الزيتون في ولاية جيجل.

التكرارات	الفئات
30	50-60
20	60-70
40	70-80
10	90-100
100	**المجموع**

المطلوب:

بما أن الجدول التكراري منتظم، فلا داع إلى التعديل، و بالتالي الفئة المنوالية هي الفئة الثالثة(80-70) كونها تقابل اكبر تكرار هو(40)، إذن المنوال يساوي:

$$Mo = \frac{A_1 + A_2}{2} = \frac{70+80}{2} = 75$$

[27]. جلاطو جيلالي، الإحصاء، المرجع سبق ذكره، ص 52.

ملاحظة:

عند حساب المنوال للتوزيعات التكرارية غير المنتظمة أي أطوال فئاتها غير متساوية، ينبغي أولا تعديل التكرارات قبل القيام بتحديد الفئة المنوالية.

ب)- طريقة الفروق (لبيرسون Pearson)

تُعدّ هذه الطريقة من أدقّ طرق إيجاد المنوال وأكثرها استخداما، ولإيجاد المنوال حسب هذه الطريقة نتبع الخطوات التالية:

- تحديد الفئة المنوالية وهي الفئة التي تقابل أكبر تكرارا.
- إيجاد الفرق الأول Δ_1 وهو الفرق بين تكرار الفئة المنوالية وتكرار الفئة السابقة لها.
- إيجاد الفرق الثاني Δ_2 وهو الفرق بين تكرار الفئة المنوالية وتكرار الفئة اللاحقة لها أي التي تليها.
- حساب المنوال حسب العلاقة التالية:

$$M_o = A1 + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times C$$

حيث:

A_1: الحد الأدنى للفئة المنوالية

C: طول الفئة

مثال17: تمثل البيانات التالية الدخل الشهري لـ 50 عائلة موزعة كما يلي:

فئات الدخل	20-25	25-30	30-35	35-40	**المجموع**
عدد الأسر	12	22	10	6	**50**

المطلوب: إيجاد المنوال بطريقة الفروق لبيرسون

الحل:

- نجد أن الفئة المنوالية هي الفئة:]25-30] المقابلة لأكبر تكرار 22.
- نجد $\Delta_1 = 22 - 12 = 10$
- نجد $\Delta_2 = 22 - 10 = 12$
- نطبق العلاقة:

$$Mo = A1 + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times C$$

$$Mo = 25 + \frac{10}{10+12} \times 5 = 27,27$$

ج)- طريقة العزوم (الرافعة)

تعدّ هذه الطريقة أدقّ من طريقة مركز الفئة المنوالية، ولإيجاد المنوال حسب هذه الطريقة يجب إتباع الخطوات التالية:

- تحديد الفئة المنوالية وهي التي تقابل أكبر تكرارا أو أكبر تكرار معدل.
- إيجاد(n–1): التكرار السابق لتكرار الفئة المنوالية.
- إيجاد(n+1): التكرار اللاحق لتكرار الفئة المنوالية.
- حساب المنوال وفقا للعلاقة التالية:

$$Mo = A_1 + \frac{(n-1)}{(n-1)+(n+1)} \times C$$

<u>مثال18</u> : نأخذ المثال السابق لحساب المنوال نجد:

*]25-30[هي الفئة المنوالية
* 10 =(n + 1)
* 10 =(n + 2)
* $25 = A_1$

$$Mo = 25 + \frac{12}{12+10} \times 5 = 27{,}27$$

ملاحظة:

- يُعاب على طريقة الرافعة اعتمادها على تكراري الفئتين السابقة و اللاحقة وإهمالها لأكبر تكرار وهو تكرار الفئة المنوالية.
- تعتمد طريقة بيرسون على الفروق، بينما تعتمد طريقة الرافعة على التكرارات الأصلية، لذا نجد أن قيمة المنوال تختلفان و لكنهما متقاربان على العموم.

د)- الطريقة البيانية:

لتحديد المنوال بيانيا نتبع مجموعة من الخطوات التالية:

- نرسم محورين متعامدين: المحور الأفقي للفئات و المحور العمودي للتكرارات؛

– نرسم المدرج التكراري ونكتفي برسم ثلاث مستطيلات فقط، تمثل تكرار الفئة المنوالية وتكرار الفئة قبل المنوالية و الفئة بعد المنوالية (وفي حالة جدول غير منتظم لا بد من رسمه باستخدام التكرارات المعدلة بدل التكرارات الأصلية)؛

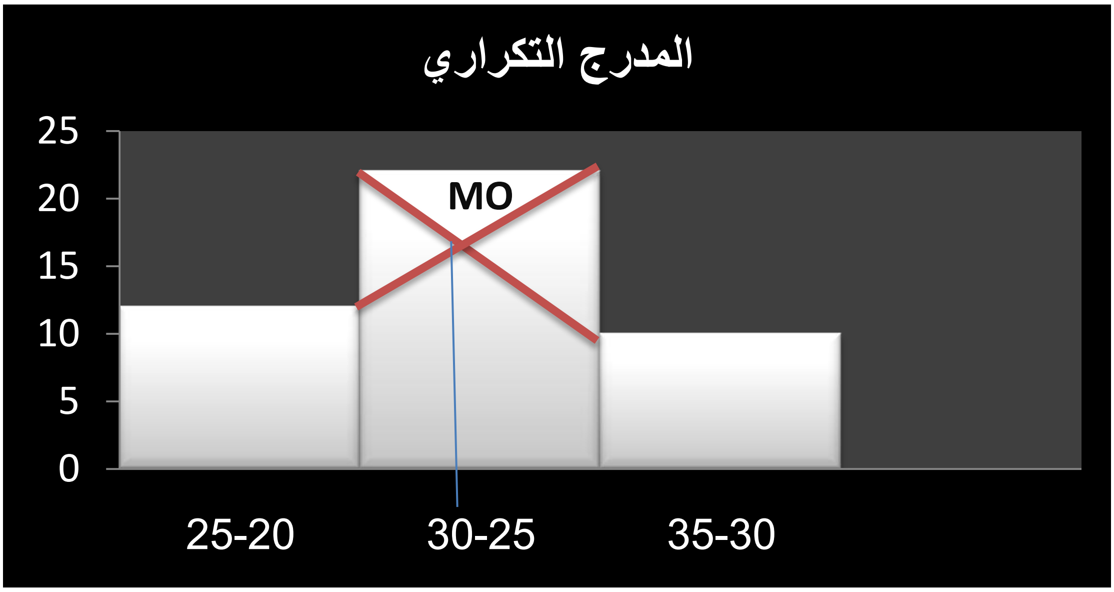
– نصل بخط مستقيم رأس الحد الأعلى للفئة المنوالية برأس الحد الأعلى للفئة التي تسبق الفئة المنوالية؛

– نصل بخط مستقيم رأس الحد الأدنى للفئة المنوالية برأس الحد الأدنى للفئة التي تلي الفئة المنوالية؛

– من نقطة تقاطع المستقيمين نُسقط عمودا على المحور الأفقي تمثل تقديرا لقيمة المنوال بيانيا.

مثال19: بالرجوع للمثال السابق، استنتج قيمة المنوال بيانيا.

الحل:



هـ)– الطريقة غير المباشرة(العلاقة بين المتوسطات الثلاثة):

يمكن إيجاد قيمة المنوال بطريقة غير مباشرة من خلال العلاقة بين المتوسطات الثلاثة:

$$\bar{X} - Mo = 3\,(\bar{X} - Me)$$

وبتبسيط العلاقة السابقة نتحصل على الصيغة غير المباشرة للمنوال كالآتي:

$$Mo = 3\,Me - 2\,\bar{X}$$

ملاحظة:

في كل الحالات يقع Me بين Mo و$\bar{X}$ وذلك حسب الحالات التالي:

- عندما يكون التوزيع التكراري متماثل أو متناظر نجد: $\bar{X} = Me = Mo$
- عندما يكون التوزيع التكراري غير متماثل من اليمين نجد: $\bar{X} > Me > Mo$
- عندما يكون التوزيع التكراري غير متماثل من اليسار نجد: $\bar{X} < Me < Mo$
- عندما يكون التوزيع التكراري قريب من التماثل أو التناظر : $\bar{X} - Mo \approx 3(\bar{X} - Me)$

مثال20: إذا كان لدينا توزيع قريب من التماثل، و قيمة المنوال 6 و قيمة الوسيط 12، جد الوسيط الحسابي ؟

$$\bar{X} - Mo = 3(\bar{X} - Me)$$

$$\bar{X} - 6 = 3(\bar{X} - 12)$$

$$\bar{X} - 6 = 3(\bar{X} - 36)$$

$$\bar{X} = 15$$

3)- <u>المنوال للجدول التكراري غير المنتظم:</u>

يتم حساب المنوال ورسم المدرج التكراري للجدول غير المنتظم من خلال التكرار المعدل ER.

<u>مثال20</u>: لدينا التوزيع التكراري التالي، احسب المنوال.

الفئات	Ni	C	ER
2-7	5	5	1
7-12	15	5	3
12-18	12	6	2
18-23	10	5	2
23-27	4	4	1
المجموع	**46**	/	/

$$Mo = A1 + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times C$$

$$Mo = 7 + \frac{(3-1)}{(3-1) + (3-2)} \times 5 = 7 + 3,33 = 10,33$$

4)- خصائص المنوال:

- يمكن تحديد المنوال بسهولة لأنّه يمثّل القيمة الأكثر إنتشارا؛

- لا يتأثر بالقيم الشاذة و المتطرفة؛
- يمكن حسابه في الجداول المفتوحة شرط أن لا تكون الفئة المفتوحة هي الفئة المنوالية؛
- يمكن إيجاده في حالة بيانات كيفية، حيث يأخذ الصفة الأكثر تكرارا؛
- مقياس غير دقيق لأنه لا يأخذ بعين الإعتبار جميع القيم أثناء حسابه؛
- لا يمكن حسابه في التوزيعات التكرارية التي تحتوي على فئتين منواليتين، إلا أنّه بالإمكان إيجاد قيمته بطريقة غير مباشرة من خلال العلاقة التقريبية بين المتوسطات الثلاثة.

رابعا: المتوسطات الأخرى

1)- الوسط الهندسي Moyenne Géométrique

الوسط الهندسي لـ N قيمته من القيم المتغير الإحصائي هو عبارة عن الجذر النوني لجداء هذه القيم، و يرمز له بـ G أو MG أو $\bar{X}G$

يُستخدم الوسط الهندسي لإيجاد معدلات الزيادة أو النقصان لمتغير معيّن مثل:

- تزايد عدد السكان(معدل النموّ السكاني)؛
- معدل النموّ والتطّور؛
- معدل الناتج الوطني؛
- معدل التشغيل أو البطالة أو التضخم؛
- معل تغيير الأسعار والأرقام القياسية؛
- معدل تطور متغير ما معبرا عنه بسلاسل زمنية.

أ)- الوسط الهندسي للبيانات غير المبوبة (الوسط الهندسي البسيط)

إذا كان لدينا أعداد الموجبة $X_1, X_2 \dots X_N$، فإن وسطها الهندسي يعرف بالمعادلة التالية:

$$\sqrt{x_1 . x_2} .$$

ولتسهيل العمليات الحسابية نقوم بإدخال اللوغاريتم لكل الطرفين على الصيغة السابقة، لتصبح بالشكل التالي:

$$\text{Log } G = \log [(X_1, X_2, X_3 \ldots X_n)^{1/n}]$$

$$\text{Log } G = 1/n\, [\log X_1 + \log X_2 + \log X_3 + \ldots + \log X_n.]$$

$$\text{Log } G = 1/n \sum \log X_i$$

$$\boxed{G = 10^{\log G}}$$

مثال21: أوجد الوسط الهندسي للقيم التالية بطريقتين: 2،4،6.

الحل: الطريقة الأولى:

$$G = \sqrt[3]{2x4x5} = \sqrt[3]{48} = 3{,}63$$

$$G = 48^{(1/3)} = 3{,}63$$

أو:

$$G = \sqrt[3]{48}$$

$$\text{Log } G = 1/3 \log (48) = \frac{\log (48)}{3} = \frac{1{,}6812}{3} = 0{,}5604$$

$$\text{Log } G = 0{,}5603 \qquad G = 10^{0{,}5603} = 3{,}63$$

الطريقة الثانية:

$$\text{Log } G = 1/3 \sum \log Xi$$

$$\text{Log } G = 1/3\, [\log 2 + \log 4 + \log 6]$$

$$\text{Log } G = 1/3\, [1{,}6811] = 0{,}5603$$

$$\text{Log } G = 0{,}5603 \qquad G = 10^{0{,}5603} = 3{,}63$$

نلاحظ وجود نفس النتيجة في الطريقتين

ب)- الوسيط الهندسي للبيانات المبوبة:_ MOYENNE GEOMETRIQUE PONDEREE

في حالة التوزيعات التكرارية حيث القيم Xi لها أوزان أو تكرارات ni، ولتحديد الوسط الهندسي يجب إتباع الخطوات التالية:

- تحديد مراكز الفئات ثم قيمة لوغارتماتها؛
- ضرب لوغاريتم مركز كل فئة في تكرارها أي: $ni \log x_i$،
- إيجاد مجموع $n_i \log x_i$،

– جمع حاصل الضرب نتحصل على مجموع التكرارات ولوغارتمات مراكز الفئات

– قسمة حاصل الضرب على مجموع التكرارات

– نطبق العلاقة، ثم نحسب G الذي يساوي $G=10^{log}$

تصبح العلاقة كما يلي:

$$\text{Log } G = \frac{\sum ni \log xi}{\sum ni}$$

مثال22: أحسب الوسيط الهندسي للتوزيع التكراري التالي:

المجموع	120−130	110−120	100−110	90−100	**الفئات**
40	4	12	14	10	**ni**

الحل:

$n_i \log x_i$	$\text{Log } x_i$	x_i	n_i	**الفئات**
19,777	1,9777	95	10	90−100
28,2954	2,0211	105	14	100−110
24,7272	2,0606	115	12	110−120
8,3876	2,0969	125	4	120−130
81,1872	/	/	**40**	**المجموع**

$$\text{Log } G = \frac{\sum ni \log xi}{\sum ni} = \frac{81.1872}{40} = 2.0296$$

$$G = 10^{2.0296} = 107.05$$

2)– الوسط التوافقي Moyenne Harmonique

الوسط التوافقي هو مقلوب الوسط الحسابي لمقلوب قيم المتغير الإحصائي، و يرمز له بـ: H أو GH أو $\bar{X}_H$، و يستخدم عامة في حالة وجود علاقة عكسية بين المتوسطات[28] .

كما يستخدم لإيجاد متوسط معدلات زمنية مثل كلم/الساعة، عدد الوحدات المنتجة يوميا، عدد الصفقات المبرمجة في السنة، متوسطات الأسعار.

[28] تيلولت سامية، مبادئ في الإحصاء، المرجع سبق ذكره،ص 81.

أ)- الوسط التوافقي في حالة بيانات غير مبوبة(الوسط التوافقي البسيط):

ليكن $X_1, X_2, \ldots, X_n$ قيم المتغير الإحصائي، يحسب الوسط التوافقي على الشكل التالي:

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

مثال23: أحسب H للقيم التالية: 2 , 8 , 13 , 22

$$H = \frac{4}{\frac{1}{22} + \frac{1}{13} + \frac{1}{8} + \frac{1}{2}} = \frac{1}{0,746} = 5,36$$

ب)- الوسط التوافقي في حالة بيانات مبوبة:

في حالة البيانات المبوبة والجداول التكرارية يصبح القانون كالآتي:

$$H = \frac{\sum ni}{\sum \frac{ni}{xi}}$$

مثال23: حسب المثال السابق نحسب الوسط التوافقي:

الفئات	90−100	100−110	110−120	120−130	المجموع
ni	10	14	12	4	40
xi	95	105	115	125	/
Ni/xi	0,111	0,133	0,104	0,032	0,38

$$H = \frac{\sum ni}{\sum \frac{ni}{xi}} = \frac{40}{0,38} = 105,26$$

3)- الوسط التربيعي: MOYENNE QUADRATIQUE

وهو الجذر التربيعي لمتوسط مربعات قيم المتغير الإحصائي، ويرمز له بـ Q أو MQ أو $\cancel{X}Q$

ويمكن حسابه في الحالتين المبوبة وغير المبوبة كما يلي[29]:

أ)- الوسط التربيعي في حالة بيانات غير مبوبة:

ويعطي العلاقة التالية:

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

مثال24: أحسب الوسط التربيعي للقيم التالية: 4 ، 6، 3، 1

$$Q = \sqrt{\frac{1^2 + 3^2 + 6^2 + 4^2}{4}} = \sqrt{15,5} = 3,39$$

ب)-الوسط التربيعي في حالة بيانات مبوبة:

$$Q = \sqrt{\frac{\sum ni\, xi^2}{\sum ni}}$$

مثال25: من المثال السابق جد الوسط التربيعي :

الحل:

الفئات	90−100	100−110	110−120	120−130	المجموع
ni	10	14	12	4	**40**
Xi	95	105	115	125	/
Xi^2	9025	11025	13225	15625	/
Xi^2Ni	90250	154350	158700	62500	**465800**

$$\sqrt{\ } Q = \sqrt{\frac{465800}{40}} = 11645 = 107.91$$

[29] جلاطو جيلالي، الاحصاء، المرجع سبق ذكره، ص 39.

ملاحظة: هناك علاقة تربط كل من Q,H,G,X في كل الحالات وبالنسبة لأي سلسلة إحصائية أو توزيع تكراري فإن:

$H < G < X < Q$ نلاحظ من المثال السابق أن:

$$105{,}26 < 107{,}05 < 107{,}50 < 107{,}91$$

الفصل الرابع

التشتت

الفصل الرابع: مقاييس التشتت

تمهيد:

بعدما تطرقنا إلى مقاييس النزعة المركزية التي تبيّن مدى تمركز البيانات الإحصائية حول نقطة معينة، ولكنها غير كافية لتحديد خواص الظاهرة بشكل جيد ولاسيما في مجال المقارنة بين عدّة مجموعات من الظواهر المدروسة.

فإن مقاييس التشتت تبيّن مدى انتشار قيم السلسلة الإحصائية حول قيمة مركزية، أي تباعدها أو تقارب بعضها من البعض الآخر، وهذا ما يسمى بالتشتت، يمكن أن يكون لتوزيعين إحصائيين نفس الوسط الحسابي، ولكنهما يختلفان تماما من حيث توزيع القيم حول هذا الوسط، مثلا إذا كانت لدينا مجموعتين هما:

المجموعة الأولى: 1500، 3500، 1000 وسطها الحسابي يساوي 2000

المجموعة الثانية: 2000، 2000، 2000 وسطها الحسابي يساوي 2000

نجد الوسط الحسابي للمجموعتين متساو وهو 2000، إلا أن قيم المجموعة الثانية متجانسة أي انتشارها أو تباعدها يساوي الصفر (لا يوجد تشتت)، في حين أن قيم المجموعة الأولى ليست متجانسة وانتشارها واسع، والفرق بين أكبر وأصغر قيمة يساوي 2500 = 1000-3500، وعليه كلما تباعدت قيم المجموعة عن بعضها البعض كلما فقد المتوسط قيمته، لذا جاءت أهمية دراسة مقاييس التشتت التي تعدّ مقاييس بديلة عن مقاييس النزعة المركزية وذلك عندما يتعذر علينا عملية الوصف أو المقارنة بين نتيجة لتساوي المتوسطات المحسوبة، وهي نوعان:

- مقاييس التشتت المطلقة: وهي التي تقيس درجة التشتت مُقدّرة بالوحدات، والتي تقيس قيم الظاهرة المدروسة نفسها كالكيلوغرام، السنتيمترات، الدينار... وهي المدى، الانحراف الربيعي، الانحراف المتوسط، والانحراف المعياري.
- مقاييس التشتت النسبية: وهي التي تقيس التشتت على شكل نسب، وهي المدى النسبي، الانحراف الربيعي النسبي، والانحراف المعياري النسبي، ومعامل الاختلاف.

أولا: مقاييس التشتت المطلقة:

وهي مؤشرات إحصائية وصفية تُستخدم لقياس التشتت، أي تباعد المطلق فيما بين قيم المجموعة أو التوزيع التكراري، وتكون هذه المقاييس على أنواع عدّة نذكر منها ما يلي:

1/ المدى العام: Range أو Etendue

يسمى بالمدى أو المدى العام وهو من أبسط مقاييس التشتت مفهوما وتطبيقا، فهو يعبّر عن مدى تغيّر الظاهرة الإحصائية، ويعرف بأنه الفرق بين أكبر وأصغر قيمة في البيانات المعطاة، ما يعاب عليه أنه غير دقيق لا يعتمد إلا على حساب قيمتين في السلسلة ، وبالتالي لا يُعير اهتماما للقيم الاخرى في المجموعة، ولا يعطي تصورا واضحا عن مدى انتشار وتوزيع القيم داخل المجموعة، كما لا يمكن إيجاده في التوزيعات المفتوحة من طرف واحد أو من الطرفين[30]، ويرمز له بــ R أو E حيث: $\mathbf{R=X_{max} - X_{min}}$

مثال 1: أحسب المدى العام للقيم التالية: 5، 8، 12، 10، 15، 6.

$R=15 - 5 =10$

مثال 2: لدينا التوزيعين التاليين:

المجموعة A : 6، 8، 9، 10، 12.

المجموعة B : 14، 4، 8، 1، 3.

المدى العام المجموعة A يساوي: $E_A= 12-6=6$

المدى العام المجموعة B يساوي: $E_B= 14-1=13$

نقول أن المجموعة B أكثر تشتتا من المجموعة A، وأن المجموعة A أكثر تجانسا من المجموعة B ولكن المجموعة A أحسن من المجموعة B لأن التشتت الضعيف أحسن من التشتت القوي.

1 جلاطو جيلال، الاحصاء، المرجع سبق ذكره، ص 70.

وبالنسبة للبيانات المبوبة يتم حساب المدى العام من خلال الفرق بين الحد الأعلى للفئة الأخيرة والحد الأدنى للفئة الأولى، أو من خلال أخر مركز الفئة مطروحا منها أول مركز الفئة، لذا نجد هذه الطريقة المفضلة لأنها تلغي أثر القيم المتطرفة.

مثال 3: الجدول التكراري التالي يمثل الأجور (بالدينار) لعينة 50 موظفا في مؤسسة جزائرية.

الفئـات	70−80	80−90	90−100	100−110	110−120	Σ
التكرار	8	13	15	6	8	**50**

المطلوب: حساب المدى العام

$$R=120-70=50$$

2/ المدى الربيعي:

يسمح لنا المدى الربيعي التخلص من أثر القيم المتطرفة التي لا تعتمد على أكبر وأصغر قيمة، وهو يعرف بأنه الفرق بين الربيعي الثالث والربيعي الأول ونرمز له بـ RQ أو $Q=Q_3-Q_1$ حيث:

ويتميز بما يلي[31]:

- يضم 50% من المجتمع مهما كان التوزيع الإحصائي؛
- يتغير طوله مقارنة بالمدى العام حسب طبيعة التوزيع؛
- استعمالاته محدودة نظرا لبساطته، غير أنه أحسن من المدى العام.

3/ الإنحراف الربيعي: interquartile

نظرا لندرة استخدام المدى الربيعي في الدراسات الإحصائية، يستخدم بدلا منه مقياس الانحراف الربيعي الذي لا يتأثر بالقيم الشّادة والمتطرفة ويسمى أيضا نصف المدى الربيعي، والذي يساوي نصف المجال ما بين الربيعيات وهو قريب جدا من الوسيط، ولكنه لا يعتمد في

[31] موساوي عبد النور، بركان يوسف، الإحصاء، المرجع سبق ذكره، ، ص 70.

حسابه على جميع الوحدات، كما أن حسابه يتعلق بترتيب الوحدات وليس بقيمتها[32]، ونرمز له بـ EQ أو ΔQ

$$EQ = (Q_3 - Q_1)/ 2 \quad \text{أو} \quad EQ = IQ / 2$$

مثال 4: أحسب المدى الربيعي والانحراف الربيعي للقيم التالية:29 ، 25، 24، 21، 18، 17، 14.

- أولا نقوم بالترتيب لـ Q_1 هو: $(N + 1) / 4$ ومنه قيمة Q_1 هي القيمة التي تقع في المرتبة الثانية في السلسلة وهي القيمة 17.
- ثمّ الترتيب لـ Q_3 هو: $3(N+1)/4$ ومنه قيمة Q_3 هي القيمة التي تقع في المرتبة السادسة في السلسلة وهي القيمة 25.
- إذن المدى الربيعي يساوي: $RQ = 25-17=8$
- والانحراف الربيعي يساوي: $\Delta Q = (25-17)/2 = 8/2= 4$

كما يمكن حساب المدى والانحراف الربيعيين في حالة البيانات المبوبة (الجداول التكرارية) حسب العلاقة التي دارسناها في الفصل السابق والمتعلقة بالوسيط، وهذا ما سوف نراه في المثال الأتي:

مثال 5: الجدول التكراري التالي يمثل توزيع 130 وموظف، جد كل من المدى والانحراف الربيعيين.

الفئات	90-100	100-110	110-120	120-130	130-140	Σ
Ni	10	25	45	30	20	**130**
N↗	10	35	80	110	130	//

الحل: نحسب أولا قيمة الربيعي الأول Q_1 وفقا للعلاقة التالية:

$$Q_1 = A1 + \frac{\sum N/4 - N\nearrow(n-1)}{N_Q} \times C$$

رتبة الربيعي الأول Q_1 هي: $\frac{N}{4} = \frac{130}{4} = 32{,}5$

[32] جلاطو جيلالي، الإحصاء، المرجع سبق ذكره، ص 71.

4 4

وبالتالي الفئة التي تحتوي على الربيعي الأول هي الفئة: 110 – 100

ومنه قيمة الربيعي الأول تساوي :

$$Q_1 = 100 + \frac{32,5 - 10}{25} \times 10 = 109$$

رتبة الربيعي الثالث هي: $\frac{3N}{4} = \frac{3x130}{4} = 97,5$

وبالتالي الفئة التي تحتوي على الربيعي الثالث هي الفئة: 130 – 120

ومنه قيمة الربيعي الثالث تساوي :

$$Q_3 = 120 + \frac{97,5 - 80}{30} \times 10 = 125,83$$

المدى الربيعي يساوي $RQ = 125,83 - 109 = 16,83$

الإنحراف الربيعي يساوي :

$$EQ = \frac{Q_3 - Q_1}{2} = \frac{16,83}{2} = 8,41$$

4/ الانحراف المتوسط : Ecart moyen

يعرف الانحراف المتوسط بأنه مجموع الانحرافات بالقيمة المطلقة مقسومة على عدد القيم، مع إهمال الإشارات الجبرية لهذه الانحرافات، حيث يتم استخدام القيمة المطلقة لأنه في حال استخدام انحرافات القيم عن الوسط الحسابي، فإن المجموع يساوي الصفر وذلك لوجود قيم سالبة وأخرى موجبة.

لذا نجد الانحراف المتوسط يأخذ بعين الاعتبار جميع قيم المجموعة عند الحساب ويتأثر بها على عكس المقاييس السابقة (المدى العام والمدى الربيعي)، إذن هو أكثر دقة وأكثر استعمالا ويمكن حسابه بالنسبة للوسط الحسابي □ ، أو بالنسبة للوسيط Me ونرمز له بـ $E\bar{X}$[33].

---

[33] نداء محمد الصوص، مبادئ الإحصاء، المرجع سبق ذكره، ص 53.

➢ **الانحراف المتوسط بالنسبة للوسط الحسابي :** هو الوسط الحسابي للقيم المطلقة لفوارق القيم Xi بالنسبة للوسط الحسابي وهو يساوي :

$$E\bar{X} = \frac{\sum |Xi - \bar{X}|}{N}$$

وإذا كانت كذلك القيّم Xi لها تكرارات n_i الانحراف المتوسط يساوي:

$$E\bar{X} = \frac{\sum n_i |Xi - \bar{X}|}{\sum n_i}$$

ويستعمل كذلك هذا القانون في حالة البيانات المبوبة المعبّر عنها بالفئات حيث Xi نعوّضها بمراكز الفئات.

➢ **الانحراف المتوسط بالنسبة للوسيط:**

نستبدل في هذه الحالة فقط الوسط الحسابي بالوسيط ويكون الانحراف المتوسط بالنسبة للوسيط كما يلي:

$$E_{\bar{X}} = \frac{\sum |Xi - Me|}{N}$$

وإذا كانت القيم Xi لها تكرارات ni ، فإن الإنحراف المتوسط يساوي:

$$E_{\bar{X}} = \frac{\sum n_i |Xi - Me|}{\sum n_i}$$

وفي حالة التوزيعات التكرارية المعبّر عنها بالفئات، نعوّض Xi بمراكز الفئات.

أ/ الإنحراف المتوسط في حالة البيانات غير المبوبة:

لإيجاد الإنحراف المتوسط لهذا النوع من البيانات يجب إتباع الخطوات التالية:

- إيجاد المتوسط الحسابي $\bar{X}$ وفقا للقانون التالي: $\bar{X} = \sum Xi / N$
- إيجاد انحرافات القيم Xi عن وسطها الحسابي الحقيقي $\bar{X}$ ، أي $(xi - \bar{X})$
- إيجاد الانحرافات المطلقة $|Xi - \bar{X}|$، واخذ المجموع للناتج $\sum |Xi - \bar{X}|$
- نطبق القانون التالي: $EX = \sum |Xi - \bar{X}| / N$

مثال 06: لدينا السلسلة الإحصائية التالية: 3، 1، 4، 5، 2.

المطلوب : تحديد الانحراف المتوسط.

الحل:

– تحديد أولا الوسط الحسابي: $\bar{X} = \frac{3 + 1 + 4 + 5 + 2}{5} = \frac{15}{5} = 3$

– إيجاد مجموع القيم المطلقة للفروق: $\sum |xi - \bar{X}| = |3-3|+|1-3|+ |4-3| + |5-3| + |2-3|$

$= |0| + |-2| + |+1| + |-2| + |-1|$

$= 0 + 2 + 1 + 2 + 1 = 6$

– حساب الإنحراف المتوسط بالصيغة التالية: $E\bar{X} = \frac{\sum |xi - \bar{X}|}{N}$

$E\bar{X} = 6 / 5 = 1,2$

ب/ الإنحراف المتوسط في حالة البيانات المبوبة:

مثال 06: احسب الانحراف المتوسط للتوزيع التكراري التالي:

الفئات	ni	Xi	Xini	$\lvert xi-\bar{X}\rvert$	$ni\lvert xi-\bar{X}\rvert$
2–4	2	3	6	11	22
4–12	3	8	24	6	18
12–14	7	13	91	1	7
14–22	5	18	90	4	20
22–24	3	23	69	9	27
المجموع	20	/	280	/	94

$\bar{X} = 280 / 20 = 14$

$E\bar{X} = \frac{\sum ni|xi - \bar{X}|}{\sum ni} = \frac{94}{20} = 4,7$

ومن المثال السابق يتضح لنا أن حساب الإنحراف المتوسط ليس سهلا لذلك يقل استخدامه إحصائياً.

5/ التباين والانحراف المعياري:

أ/التباين: هو عبارة عن الوسط الحسابي لمربعات الفوارق بين قيم المتغير الإحصائي ووسطها الحسابي، ويدل على مدى تشتت هذه القيم حول أحد المقاييس المركزية الوسط الحسابي.

إن حساب التباين يكتسي أهمية كبيرة في علم الإحصاء لأن التحليل الإحصائي يعتمد في كثير من الحالات على تحليل التباين ويرمز له بـ $v(x)$.

التباين في حالة البيانات غير المبوبة:

حسب القانون التالي: $v(x)=\frac{\sum (xi-\bar{X})^2}{N}$

كما يحسب التباين بتطبيق قانون كوينغ Formule de Koeing وهي تسمى كذلك بالطريقة المختصرة وهي كما يلي:

$$v(x)=\frac{\sum (xi-\bar{X})^2}{N}=\frac{\sum (xi^2-2xi\bar{X}+\bar{X}^2)}{N}$$

$$=\frac{\sum xi2}{N}-\frac{2\sum xi\bar{X}}{N}+\frac{\sum \bar{X}2}{N}$$

$$=\frac{1}{N}\sum xi^2-2\bar{X}\frac{\sum xi}{N}+\frac{N\bar{X}^2}{N}$$

$$=\frac{\sum xi^2}{N}-2\bar{X}\bar{X}+\bar{X}^2$$

$$=\frac{\sum xi^2}{N}-2\bar{X}^2+\bar{X}^2=\frac{\sum xi^2}{N}-\bar{X}^2$$

مثال7: اليك السلسلة الإحصائية التالية

1، 10، 12، 5، 6، 4، 8، 2.

المطلوب: احسب التباين بطريقتين:

الطريقة الأولى البسيطة:

$$\bar{X} = \frac{2+8+4+6+5+12+10+1}{8} = \frac{48}{8} = 6$$

$$V(X) = \frac{\sum (xi-\bar{X})^2}{N} = \frac{(2-6)^2+(8-6)^2+(4-6)^2+(5-6)^2+(6-6)^2+(12-6)^2+(10-6)^2+(1-6)^2}{8}$$

$$= 102/8 = 12.75$$

الطريقة الثانية المختصرة:

$$V(X) = \frac{\sum xi^2}{N} - \bar{X}^2 \qquad V(X) = \frac{22+82+42+62+52+122+102+12}{8} - (6)^2$$

$$V(X) = \frac{390}{8} - 36 = 48.75 - 36 = 12.75$$

نلاحظ نفس النتيجة

ب/التباين في حالة البيانات المبوبة:

في حالة التوزيعات الإحصائية المعبرة عنها بالفئات تكون هي مراكز الفئات.

$$v(x) = \frac{\sum ni(xi-\bar{X})^2}{\sum ni}$$

وبالنسبة للبيانات المبوبة، صيغة التباين تصبح كما يلي:

$$V(X) = \frac{\sum nixi^2}{\sum ni} - \bar{X}^2$$

مثال8 : إليك الجدول التكراري التالي، احسب التباين بطريقتين.

الفئات	ni	xi	Xi^2	ni . xi	Xi^2 . ni	$(xi-\bar{X})^2$	$ni\,(xi-\bar{X})^2$
20−30	2	25	625	50	1250	380.25	760.5
30−40	4	35	1225	140	4900	90.25	361
40−50	8	45	2025	360	16200	0.25	02
50−60	5	55	3025	275	15125	110.25	551.25
60−70	1	65	4225	65	4225	420.25	420.25
Σ	20	/	/	890	41700	/	2095

الطريقة الأولى:

$$v(x)=\frac{\sum ni(xi-\bar{X})^2}{\sum ni} = \frac{2095}{20} = 104.75$$

الطريقة الثانية:

$$V(X)=\frac{\sum nixi^2}{\sum ni} - \bar{X}^2 = \frac{41700}{20} - (44.5)2 = 2085-1980.25 = 104.75$$

ب/الانحراف المعياري:

الإنحراف المعياري هو عبارة عن الجذر التربيعي الموجب للتباين، هو من أكثر مقاييس التشتت أهمية وإستعمالا في الدراسات الإحصائية، ويرمز له بالرمز SD ويعطي بالعلاقات التالية:

➢ **الانحراف المعياري في حالة البيانات غير المبوبة:**

$$SD = \sqrt{V(X)} = \sqrt{\frac{\sum xi^2 - \bar{X}^2}{N}} = \sqrt{\frac{\sum (xi-\bar{X})^2}{N}}$$

من المثال السابق نجد الانحراف المعياري: $SD = 12,75 \sqrt{\;} = 3,57$

➢ **الانحراف المعياري في حالة البيانات المبوبة:**

$$SD = \sqrt{\underline{\sum ni(xi-\bar{X})^2}} = \sqrt{\underline{\sum nixi^2} - \bar{X}^2}$$

$\sum ni$ $\sum ni$

من المثال السابق، الانحراف المعياري يساوي:

$$SD = \sqrt{104.75} = 10.23$$

ج/ أهمية الانحراف المعياري وخصائصه:

يُعد الانحراف المعياري من أفضل مقاييس التشتت المطلق وأدقها على الإطلاق، ويسمى أحيانا بالانحراف القياسي، وانه يتأثر بالقيم المتطرفة كما هو الحال بالنسبة للوسط الحسابي، كما يستخدم في قياس الإرتباط بين المتغيرات العشوائية وفي السلاسل الزمنية لقياس علاقة الارتباط بين الظاهرة المدروسة والزمن أو في دراسة الدورات الاقتصادية، ومن خصائصه:

– يعتبر الانحراف المعياري من اهم مقاييس التشتت لأنه يستعمل في حساب عدة مؤشرات منها معامل الارتباط، وفي تحديد أشكال التوزيعات الإحصائية والاحتمالية.

– يتأثر الانحراف المعياري بالقيمة المتطرفة كما هو الحال النسبة للوسط الحسابي فهو أكثر مقاييس التشتت أهمية وأكثرها دقة، يعتبر كثير الاستخدام في القوانين والنظريات الإحصائية.

– الانحراف المعياري لقيمة ثابتة يساوي الصفر.

$$V(a) = 0$$

$$V(ax) = a^2V(x)$$

$$V(a+x) = V(a) + V(x) = V(x)$$

– يتمتع الانحراف المعياري بخاصية هامة، وهي أنه في حالة التوزيع الطبيعي يقع الوسط الحسابي في وسط المنحنى، والمدى بين الوسط الحسابي والانحراف المعياري بحصر نسب معينة من قيم التوزيع على الشكل التالي:

1/ $(2/3\ SD \pm \bar{X})$.............$(2/3SD+\bar{X};\ 2/3SD-\bar{X})$..........50%.

معناه أن المجال يحصر 50% من قيم التوزيع.

2/ (1SD $\pm\bar{X}$)....................(1SD+$\bar{X}$, 1SD-$\bar{X}$)..................68.27%

معناه أن المجال يحصر 68.27% من قيم التوزيع.

3/ (2SD ± $\bar{X}$)................(2SD+$\bar{X}$; 2SD-$\bar{X}$)...............95%

معناه أن المجال يحصر 95% من قيم التوزيع

4/ (3SD ± $\bar{X}$)................(3SD+$\bar{X}$; 3SD-$\bar{X}$)..................99%

معناه أن المجال يحصر 99% من قيم التوزيع

مثال9: أحسب النسبة [$\bar{X}$±SD] إذا كان لدينا:

$\bar{X}$ = 44.5 SD = 10.23

[$\bar{X}$± SD] = [44.5+10.23 ; 44.5−10.23] = [54.73 ; 34.27]

[$\bar{X}$± SD] = [55 ; 35]

]55-50]+] 50-40]+] 40-30] =] 54-34]

X_2 + 8 + X_1

[30-40 [..........C=10............ni=4 .... X_1= (4x4)/10 = 16/10 = 1.6 ≈2

[34-40 [..........C=4..............ni= X_1

[50-60[..........C=10.............ni=5 .... X_2= (5x5)/10 = 25/10 = 2.5 ≈ 3

[55-50[..........C=5..............ni= X_2

N = 2 + 8 + 3 = 13

→← (13/21) / 100 = 65 % ← هذه هي نسبة المجال: [55-34]

نلاحظ أن هذه النسبة قريبة جدا من النسبة النظرية التي تساروي % 65 ، ما يدل على أن هذا التوزيع قريب من التناظر.

-هناك علاقة بين SD و E و IQ لتحديد العلاقة نفرض أن التوزيع المدروس يكون قريبا الى التماثل أو التناظر بالنسبة للقيمة المركزية لذا نجد[34]:

$Q_3 = \bar{X} + 0.67\ SD$..$Q_1 = \bar{X} - 0.67\ SD$

ومن هاتين العلاقتين نستنتج أن:

$Q_3 - Q_1 = 1.34\ SD$

- العلاقة بين الانحراف المعياري والانحراف المتوسط: تعتبر هذه العلاقة تجريبية، فكلما كان التوزيع الإحصائي أقرب إلى التماثل أو التناظر كلما كانت العلاقة بين المقياسين صحيحة: $EX = 4/5\ SD$

ثانيا: مقاييس التشتت النسبي

إن مقاييس التشتت المطلقة لا تسمح لنا بمقارنة التشتت بين ظاهرتين مختلفتين، لذلك نلجأ إلى مقاييس التشتت النسبية و هي مقاييس التشتت المطلقة قسمة المتوسط الذي يُحسب حوله، وعليه يتم حساب مقاييس التشتت النسبي كما يلي:

1)- المدى النسبي: هو المدى العام على الوسط الحسابي، أي:

$$E\bar{X} = \frac{Xmax - Xmin}{\bar{X}} \cdot 100$$

2)- معامل الإختلاف النسبي Coefficient de Variation

ويسمى كذلك بالإنحراف المعياري النسبي، ويستعمل لقياس التشتت النسبي لمجموعة ما أو توزيع إحصائي مغلق من الجهتين، فهو النسبة بين الإنحراف المعياري المتوسط الحسابي، ونعتمد عليه للمقارنة بين التوزيعات الإحصائية غير المتجانسة لها وحدات قياس مختلفة، وهو الأكثر إستعمالا لأنه أكثرهم دقة ومؤشر قياسي للتشتت، فهو يربط بين أحد مقاييس النزعة المركزية X وأحد مقاييس التشتت SD، ويعطى بالعلاقة التالية:

$$CV = \frac{SD}{\bar{X}} \cdot 100$$

ملاحظة:

- كلما كان CV كبير كلما دلّ ذلك على قوة التشتت بين مفردات الظاهرة.

[34] جلاطو جيلالي، الإحصاء، المرجع سبق ذكره، ص 76.

– كلما كان CV صغير كلما دلّ ذلك على تجانس مفردات الظاهرة.

– إذا كان الوسطان الحسابيان متساويان، فإن المتغير ذو الإنحراف المعياري الأكبر يكون أكثر تشتتا.

– أما إذا كان الإنحرافان المعياريان متساويان فإن المتغير ذو الوسط الحسابي الأكبر يكون أقل تشتتا.

3)– معامل الإختلاف الربيعي:

كما يعرف كذلك باسم معامل الانحراف الربيعي النسبي أو معامل التغير النسبي، فهو يستعمل لقياس درجة التشتت النسبي لمجموعة ما أو توزيع معين في كلتا الحالتين للبيانات المبوبة و غير المبوبة على حدّ سواء، ويستعمل لحساب تشتت الجداول المفتوحة ويعتمد على الربيعي الأول و الثالث، ويرمز له بـ CVQ.

يعد معامل الإختلاف الربيعي من أفضل معاملات الإختلاف لقياس التشتت النسبي للمجموعات أو التوزيعات للمقارنة بين مجموعتين أو أكثر تختلف في وحدات قياسها لا سيما تلك التوزيعات التي تكون غير محدودة المعالم أين لا يمكن حساب $\bar{X}$، CVQ ويرمز له بـ :

$$CVQ = \frac{Q_3 - Q_1}{Q_3 + Q_1} \cdot 100$$

مثال 09: إذا كانت لدينا قيم الربيعيين الأول و الثالث لبيانات أعمار وأوزان عينة من الأشخاص:

– المجموعة الأولى: $Q_1 = 22$ عام $Q_3 = 37$ عام

– المجموعة الثانية: $Q_1 = 56$ كغ $Q_3 = 80$ كغ

المطلوب: قارن بين تشتت مجموعة الأعمار و مجموعة الأوزان

الحل: للمقارنة بين المجموعتين، نقوم بحساب معامل الإختلاف الربيعي لكلا المجموعتين وفقا للصيغة التالية:

للأعمار: $$CVQ = \frac{Q_3 - Q_1}{Q_3 + Q_1} = \frac{37 - 22}{37 + 22} \cdot 100 = 25{,}4\ \%$$

للأوزان: $$CVQ = \frac{80 - 65}{80 + 65} \cdot 100 = 10{,}34\ \%$$

نجد أن CVQ للأعمار أكبر من CVQ للأوزان، إذن فمجموعة الأوزان هي الأكثر تجانسا من مجموعة الأعمار.

4)- المتوسط الربيعي أو المدى الربيعي النسبي MeQ

وهو النسبة بين المدى الربيعي و الوسيط X 100، يعد أحسن مقاييس التشتت في حالة الجداول المفتوحة:

$$MeQ = \frac{Q_3 - Q_1}{m_e} \cdot 100$$

الإنحراف الربيعي النسبي (Wr)

$$Wr = \frac{Q_3 - Q_1}{2me} \cdot 100$$

5)- النسبة بين المدى الربيعي و المدى العام

يبين هذا المقياس تشتت 50 % من الوحدات الإحصائية المركزية (أي التي تقع في مركز التوزيع) حول الوسيط مقارنة بالمدى العام، ونكتب علاقة هذا المقياس بالشكل التالي:

$$R = \frac{Q_3 - Q_1}{E} \cdot 100$$

ومن مميزاته:

- يُستعمل لقياس تشتت 50 % من الوحدات الإحصائية التي تقع حول القيمة المركزية بالنسبة لنفس التوزيع.
- إذا كان $R = 50\ \%$ يكون التوزيع الإحصائي متماثل أو متناظر.
- إذا كان $R > 50\ \%$ يكون التوزيع الإحصائي قوي التشتت بالنسبة للقيمة المركزية
- إذا كان $R < 50\ \%$ يكون التوزيع الإحصائي قليل التشتت بالنسبة للقيمة المركزية(الوسيط)

مثال 10: إذا كان لدينا

$$Q_3 = 1200 \qquad Q_1 = 800$$
$$X_{max} = 1300 \qquad X_{min} = 600$$

$$R = \frac{1200 - 800}{1300 - 600} \times 100 = \frac{400}{700} \times 100 = 57{,}14\ \%$$

نلاحظ أن المدى الربيعي يمثل 57,14 % من المدى العام و يحتوي على 50 % من الوحدات الإحصائية، إذن يوجد تشتت نسبي وأقرب إلى التماثل.

الفصل الخامس

مقـاييس الشكــل

الفصل الخامس: مقــاييس الشكــل

تمهيـد

إضافة إلى المقاييس السابقة سنتطرق إلى مقاييس جديدة تبين لنا شكل التوزيع الإحصائي، الشكل الأول يقيس إلتواء منحنى التوزيع والثاني يقيس ارتفاع قمة منحنى التوزيع أو اعتداله إذا ما كانت حادة أو مفرطحة أو متطاولة، وتدعى مقاييس الشكل.

وقبل التطرق إلى هذه النقطة يجب تعريف مصطلح يتداول كثيرا ألا وهو العزوم التي تعتمد عليها مقاييس الشكل.

أولا: العزوم Moments

تُعرف كلمة العزم بالقوة أي مقدار العمل الذي تُحدثه، ويتوقف هذا العمل على القوة نفسها أو المسافة بين هذه القوة والنقطة التي عندها تحدث أثرها، فالمفهوم الإحصائي للعزم يتطلب تحديد النقطة التي يحسب عندها العزم، فقد يحسب حول الصفر أو حول المتوسط الحسابي أو حول أي متوسط فرضي أخر.

تكون العزوم إما أصلية أي تتمحور حول الأصل وإما تكون مركزية وتتمحور حول المتوسط الحسابي وهما نوعان: العزوم البسيطة أو الإبتدائية و العزوم المركزية.

1)- العزم حول النقطة الاختيارية:

ويسمى أيضا بالعزم البسيط، الذي يعرف بأنه الوسط الحسابي لمجموع انحرافات القيم حول الوسط الفرضي Xo مرفوع إلى قوة r أي درجة r [35].

فمثلا إذا كان لدينا المجموعة التالية من الأعداد التالية: 14، 10، 6، 5.

فإن العدد الأول هو 5 يعني قيمة تزيد عن الصفر بمقدار 5، والعدد الثاني وهو 6 يعني قيمة تزيد عن الصفر بمقدار6، ... وهنا يقال أن نقطة الأساس المستخدمة بالنسبة لهذه المجموعة هي الصفر.

[35] نداء محمد الصوص، مبادئ الإحصاء، المرجع سبق ذكره، ص 66.

أما إذا تغيرت نقطة الأساس إلى نقطة أخرى A ولتكن 2، فإن العدد الأول يعني 5 يزيد عن النقطة الاختيارية A بـ 3، والعدد الثاني يعني 6 يزيد عن A بـ 4، وهكذا... إلى أخر عدد، فإذا رمزنا للقيم بـ Xi ولعددها بـ N وللعزم حول نقطة الاختيار من المرتبة r بـ m_r فإن:

$$m_r = \frac{(x_1-A)^r + (x_2-A)^r + (x_3-A)^r + \ldots + (x_i-A)^n}{N} = \frac{\sum(xi-A)^r}{N}$$

$$m_1 = \frac{(5-2)+(6-2)+(10-2)+(14-2)}{4} = \frac{27}{4} = 6{,}75$$

يسمى المقدار 6.75 بالعزم الأول حول النقطة A = 2، أما العزم الثاني حول النقطة A لنفس القيم تساوي:

$$m_2 = \frac{\sum(xi-A)^r}{N} = \frac{(5-2)^2+(6-2)^2+(10-2)^2+(14-2)^2}{4} = \frac{233}{4} = 58{,}25$$

2)- العزم حول نقطة الصفر:

يعرف العزم حول نقطة الصفر لمجموع من القيم بالشكل التالي في حالة انعدام النقطة الإختيارية A = 0 فيصبح كما يلي:

$$m_r = \frac{(x_1-0)^r + (x_2-0)^r + (x_3-0)^r + \ldots + (x_i-0)^n}{N} = \frac{\sum(xi)^r}{N}$$

إذن العزم الأول حول الصفر يساوي الوسط الحسابي □ ، حيث r يساوي 1 ، وحسب المثال السابق

نجد:

$$m_1 = \frac{(5-0)^1+(6-0)^1+(10-0)^1+(14-0)^1}{4} = \frac{35}{4} = 8{,}75$$

حيث نجد :

$$m_0 = \frac{1}{N}\sum(x_i-0)^r$$

$$m_1 = \frac{1}{N}\sum(x_i)^1 \longrightarrow m_1 = \square$$

$$m_2 = \frac{1}{N}\sum(x_i)^2 \longrightarrow m_2 = Q^2$$

$$m_3 = \frac{1}{N}\sum(x_i)^3$$

$$M_4 = \frac{1}{N}\sum(x_i)^4$$

إذن العزم الثاني من الدرجة الثانية هو الوسط التربيعي.

أما في حالة البيانات المبوبة، العزم البسيط ذو الدرجة r يساوي:

$$m_r = \frac{\sum n_i x_i^r}{\sum n_i}$$

إذا كان r يساوي 1، فإن: $m_1 = \frac{\sum n_i x_i^1}{\sum n_i}$

إذن العزم البسيط الأول هو الوسيط الحسابي

إذا كان r يساوي 2، فإن: $m_2 = \frac{\sum n_i x_i^2}{\sum n_i}$

3)- العزم حول الوسط الحسابي **Moment Centré**

ويسمى كذلك بالعزم المركزي، ويرمز له بـ μ_r ، و هو يعرف بانه الوسط الحسابي لمجموع انحرافات القيم حول الوسط الحسابي مرفوع إلى قوة r، إذا كان $Xo = \bar{X}$ فتصبح المعادلة كما يلي:

$$\mu_r = \frac{\sum (x_i - \bar{X})^r}{N}$$

مثال 01: أحسب العزم المركزي للقيم السابقة، 14، 10، 6، 5.

الحل: نقوم أولا بحساب الوسط الحسابي $\bar{X} = \frac{5 + 6 + 10 + 14}{4} = \frac{35}{4} = 8,75$

$$\mu_r = \frac{\sum (x_i - \bar{X})^r}{N} = \frac{(5 – 8,75) + (6 – 8,75) + (10 – 8,75) + (14 – 8,75)}{4} = \frac{0}{4} = 0$$

إذا كان r يساوي 2، فإن العزم المركزي يساوي : $\mu_2 = \frac{\sum (xi - \bar{X})^2}{N}$

إذن نجد أن العزم المركزي من الدرجة الثانية يساوي التباين، أي: $\mu_2 = V(x) = SD^2$

أما في حالة البيانات المبوبة للعزم المركزي ذو الدرجة r يساوي:

$$\mu_r = \frac{n_1 (x_1 - \bar{X})^r + n_2 (x2 - \bar{X})^r + ... + ni(xi - \bar{X})^n}{n_1 + n_2 + ... + n_i}$$

إذن:

$$\mu_r = \frac{\sum ni (x_i - \bar{X})^r}{\sum ni}$$

أما في حالة البيانات المبوبة فتصبح المعادلة كما يلي:

$$\mu_5 = \frac{\sum n_i(x_i - \bar{X})^r}{\sum n_i}$$

مع xi مراكز الفئات

– إذا كان $r = 0$ فإن: $\mu_0 = 1$ $\qquad \mu_0 = \frac{\sum n_i(x_i - \bar{X})^0}{\sum n_i} = 1$

– إذا كان $r = 1$ فإن: $\mu_1 = 0$ $\qquad \mu_1 = \frac{\sum n_i(x_i - \bar{X})^1}{\sum n_i} = 1$

$$\mu_1 = \frac{1}{\sum n_i} . \sum n_i(x_i - \bar{X})^1 = \frac{1}{\sum n_i} \sum n_i x_i - \frac{1}{\sum n_i} . \sum n_i \bar{X}$$

$$= \bar{X} - \bar{X} = 0$$

ومنه: $\mu_1 = 0$

– إذا كان $r = 2$ فإن: $\mu_2 = v(x)$ $\qquad \mu_2 = \frac{\sum n_i(x_i - \bar{X})^2}{\sum ni}$

العزم المركزي الثاني يساوي التباين $V(X)$ ويساوي مربع الإنحراف المعياري SD^2.

إذن: $\mu_o = 1$

$\mu_1 = 0$

$\mu_2 = V(X)$

إذا كان $r = 0$ فإن: $\mu_0 = 1$ و $m_o = 1$

إذا كان $r = 1$ فإن: $\mu_1 = 0$ و $m_1 = \bar{X}$

إذا كان $r = 2$ فإن: $\mu_2 = v(x)$ و $m_2 = Q^2$

حيث: $Q=$: الوسط التربيعي

$v(x)$: التباين مع $v(x) = SD^2$

4)- العلاقة بين العزوم البسيطة و المركزية:

لاحظنا أن $\mu_0 = m_o = 1$ إذن يمكن تحويل العزوم المركزية حسب العزوم البسيطة معناه[36]:

- إذا كان $r = 0$ فإن: $\mu_o = 1$
- إذا كان $r = 1$ فإن: $\mu_1 = 0$
- إذا كان $r = 2$ فإن: $\mu_2 = \frac{\sum x_i^2}{N} - \bar{X}^2$:

$$= m_2 - m_1^2$$

- إذا كان $r = 3$ فإن: $\mu_3 = \frac{\sum (x_i - \bar{X})^3}{N} = m_3 - 3m_2\ m_1 + 2m_1^3$
- إذا كان $r = 4$ فإن: $\mu_4 = m_4 - 4m_1\ m_3 + 6m_1^2 m_2 - 3m_1^4$

مثال 02: إذا كان لدينا البيانات التالية، احسب العزوم حول الوسط الحسابي ثم حول الصفر.

1، 2، 4، 6، 7

الحل:

1)- حساب العزوم حول الوسط الحسابي (العزم المركزي)

$$m_r = \frac{\sum (x_i - \bar{X})^r}{N}$$

نستخرج أولا قيمة الوسط الحسابي $\bar{X}$:

$$\mu_1 = \frac{(1-4) + (2-4) + (4-4) + (6-4) + (7-4)}{5} = \frac{-3 -2 +0 +2 +3}{5} = \frac{0}{5} = 0$$

$$\mu_1 = 0$$

$$\mu_2 = \frac{(1-4)^2 + (2-4)^2 + (4-4)^2 + (6-4)^2 + (7-4)^2}{5} = \frac{9 + 4 + 0 + 4 + 9}{5}$$

---

[36] جلاطو جيلالي، الإحصاء مع تمارين، المرجع سبق ذكره، ص 78.

$\mu_2 = \frac{26}{5} = 5{,}2 \qquad \mu_2 = 5{,}2$

$$\mu_3 = \frac{(1-4)^3 + (2-4)^3 + (4-4)^3 + (6-4)^3 + (7-4)^3}{5}$$

$$= \frac{(-3)^3 + (-2)^3 + 0^3 + 2^3 + 3^3}{5} = \frac{-27 -8 +0 +8 +27}{5} = \frac{0}{5} = 0$$

التوزيع متماثل $\mu_3 = 0$

$$\mu_4 = \frac{(-3)^4 + (-2)^4 + 0^4 + 2^4 + 3^4}{5}$$

$$= \frac{81 + 16 + 0 + 16 + 81}{5} = \frac{194}{5} = 38{,}8$$

$\mu_4 = 38{,}8$

2)– حساب العزوم حول الصفر

$$m_1 = \frac{1 + 2 + 4 + 6 + 7}{5} = \frac{20}{5} = 4$$

$$m_2 = \frac{1^2 + 2^2 + 4^2 + 6^2 + 7^2}{5} = \frac{1+4+16+36+49}{5} = \frac{106}{5} = 21{,}2$$

$$m_3 = \frac{1^3 + 2^3 + 4^3 + 6^3 + 7^3}{5} = \frac{1+8+64+216+343}{5} = \frac{632}{5} = 126{,}4$$

$$m_4 = \frac{1^4 + 2^4 + 4^4 + 6^4 + 7^4}{5} = \frac{1+16+256+1296+2401}{5} = \frac{3970}{5} = 794$$

$m_4 = 794$

إذن بتطبق العلاقات السابقة نجد ما يلي:

$\mu_2 = m_2 - m^2_1$

$\mu_3 = m_3 - 3m_2m_1 + 2\, m^3_1$

$\mu_4 = m_4 - 4m_3\, m_1 + 6\, m_2 m^2_1 - 3m^4$

من المثال السابق توصلنا إلى النتائج التالية:

$m_1 = 4 \qquad \mu_1 = 0$

$m_2 = 21{,}2 \qquad \mu_2 = 5{,}2$

$m_3 = 126,4$ $\mu_3 = 0$

$m_4 = 794$ $\mu_4 = 38,8$

وهنا نتحقق من النتائج باستخدام العلاقات بين العزوم السابقة و هي كالآتي:

1)- $\mu_2 = m_2 - m^2_1$

$\mu_2 = 21,2 - (4)^2 = 5,2$

وهي نفس النتيجة في المثال $\mu_2 = 5,2$

2)- $\mu_3 = m_3 - 3(m_2 m_1) + 2\, m^3_1$

$\mu_3 = 126,4 - 3(21,2 \text{ x } 4) + 2(4)^3$

$\mu_3 = 126,4 - 3(84,8) + 128$

$126,4 - 254,4 + 128 = 0$

وهي نفس النتيجة : $\mu_3 = 0$

3)- $\mu_4 = m_4 - 4(m_3\, m_1) + 6\,(m_2\, m^2_1) - 3m^4_1$

$\mu_4 = 794 - 4(126,4 \text{ x } 4) + 6(21,2 \text{ x } 4^2) - 3(4)^4$

$\mu_4 = 794 - 2022,4 + 2035,2 + 768$

$\mu_4 = 38,8$

نلاحظ من المثال السابق أن العلاقة بين العزوم محققة.

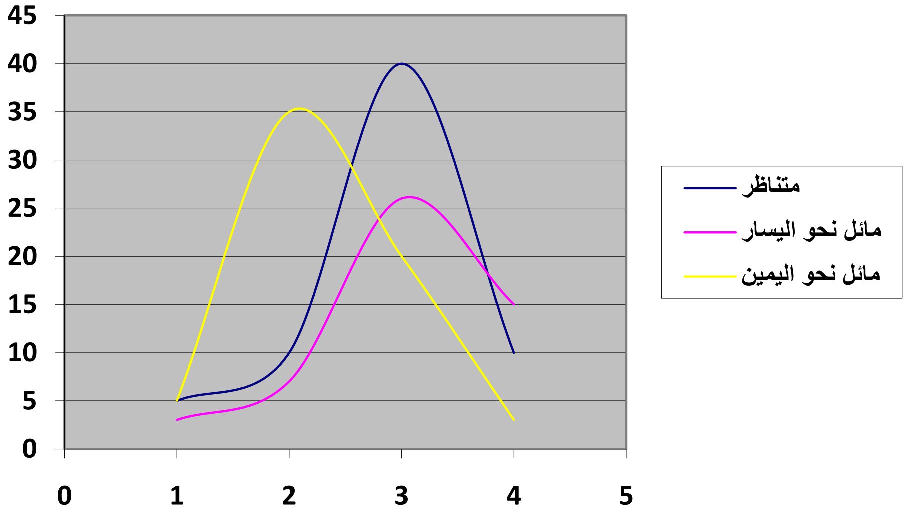
ثانيا: مقاييس الشكل

1)- الإلتـواء Dissymétries

يقيس الإلتواء درجة تماثل البيانات حول وسطها الحسابي، ونقول أن التوزيع الإحصائي يكون متناظر إذا كان لكل قيمة X_i تقابلها قيمة X_j موجودة على نفس البعد بالنسبة للوسط الحسابي $\bar{X}$، أي $\bar{X}$ هو مركز المجال $[X_i، X_j]$. و ينتج عن ذلك منحنى بياني متماثل أي طبيعي وهو يشبه شكل الجرس، حيث نجد أن المقارنة بين المنوال و الوسط الحسابي تسمح بقياس التماثل أو الإلتواء.

هذا المنحنى لبعض البيانات يبين لنا أن التوزيع قد يكون ملتوي إلى اليمين أو إلى اليسار كما يلي:

هذا المنحنى لبعض البيانات يبين لنا أن التوزيع قد يكون ملتوي إلى اليمين أو إلى اليسار كما هو في الشكل التالي:



مقاييس الإلتـواء

لقد تعرضنا سابقا لشكل التوزيع التكراري، بحيث قد يكون:

– متناظرا أو متماثلا إذا كانت: $\bar{X} = Me = Mo$

– غير متناظر من اليمين أي ملتويا جهة اليمين عندما تكون القيم المتطرفة متمركزة جهة اليمين

أي $\bar{X} > Me > Mo$.

– غير متناظر من اليسار أي ملتويا جهة اليسار عندما تكون القيم المتطرفة متمركزة جهة

اليسار $\bar{X} < Me < Mo$.

لذا يبرز هنا دور مقياس الإلتواء الذي يقيس درجة اتجاه الإلتواء، ولتحديد شكل التوزيع يجب

حساب معاملات الإلتواء و هي:

أ)– معامل الإلتواء لبيرسون Karl Pearson

قلنا سابقا أن إذا كان التوزيع متماثل يعني لدينا المتوسطات الثلاث متساوية أي $\bar{X} = Me = Mo$

أما إذا كان التوزيع غير متماثل تصبح هذه المتوسطات غير متساوية، لذا نلجأ إلى العلاقة التقريبية، و هي: $\bar{X} - Mo = 3(\bar{X} - Me)$

ومن هنا نستنتج أول مقياس يدل على الإلتواء حيث يُحسب بطريقتين[37]:

معامل بيرسون الأول: $P_1 = \frac{\bar{X} - Mo}{SD}$

معامل بيرسون الثاني: $P_1 = 3\frac{(\bar{X} - Mo)}{SD}$

وتنحصر قيمة P_1 في كلتا الحالتين بين $+1$ و -1 .

فإذا كانت $P_1 = 0$ .... التوزيع متماثل.

فإذا كانت $P_1 > 0$ .... التوزيع ملتوي نحو اليمين.

فإذا كانت $P_1 < 0$ .... التوزيع ملتوي نحو اليسار.

ب)- معامل الإلتواء لفيشر Ronald Aylmer Fisher

يقيس هذا المعامل درجة إلتواء شكل التوزيع الإحصائي، ونعتمد في ذلك على قيمة العزم المركزي من المرتبة الثالثة، ولاستبعاد وحدة القياس، نقسمه على الإنحراف المعياري من نفس المرتبة أي الثالثة و نرمز له بـ F_1، واختير العزم المركزي من المرتبة الثالثة لأن قيمته في حالة التوزيع المتناظر تساوي إلى صفر.

ويعطى بالصيغة التالية[38]: $F1 = \frac{U3}{SD3}$

مع: $\mu_3 = \frac{\Sigma n_i (X_i - \bar{X})^3}{\Sigma n_i}$

ملاحظة:

- إذا كان $F1 = 0$ فإن .... التوزيع متماثل.
- إذا كان $F_1 > 0$ فإن.............. التوزيع ملتوي نحو اليمين.
- إذا كان $F_1 < 0$ فإن............. التوزيع ملتوي نحو اليسار.

ج)- معامل يول و كندال Yule and Kendall

---

[37] تيلولت سامية، مبادئ في الإحصاء، المرجع سبق ذكره، ص 137.

[38] جلاطو جيلالي، الإحصاء مع تمارين، المرجع سبق ذكره، ص 81.

ويعرف كذلك بمعامل الإلتواء الربيعي، ويستخدم في حالة التوزيعات المفتوحة، أين لا يمكن حساب الوسط الحسابي $\bar{X}$ والإنحراف المعياري SD، وهو عبارة عن نسبة الربيعيات و الوسيط على المدى الربيعي، ويرمز له بـ C_{yk} ويعطى بالعلاقة التالية مع $(-1 < C_{yk} < +1)$:

$$C_{yk} = \frac{(Q_3 - Me) - (Me - Q_1)}{Q_3 - Q_1}$$

$$= \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{Q_3 - Q_1} = \frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1}$$

ملاحظة:

- إذا كان $C_{yk} = 0$ فإن .... التوزيع متناظر.
- إذا كان $C_{yk} > 0$ فإن .... التوزيع ملتوي نحو اليمين (إلتواء موجب).

فإذا كانت $C_{yk} < 0$ فإن .... التوزيع ملتوي نحو اليسار (إلتواء سالب).

بمعنى كلما ابتعدت قيمة معامل الإلتواء عن الصفر، زادت حدة الإلتواء يمينا أو يسارا و العكس صحيح.

مثال 03: إذا كان لدينا المعلومات التالية:

$\bar{X} = 56$ $\quad$ $Q_1 + 40,67$

$Me = 57,3$ $\quad$ $Q_3 = 72,35$

$Mo = 62,85$ $\quad$ $SD = 22\ ,36$

المطلوب: إيجاد كل من معامل الإلتواء لبيرسون بطريقتين ثم يول.

الحل:

1)- طريقة بيرسون

الطريقة الأولى:

$$P_1 = \frac{\bar{X} - Mo}{SD}$$

$$P_1 = \frac{56 - 62,85}{22,36} = -\ 0,306$$

الطريقة الثانية:

$$P_2 = \frac{3(\bar{X} - Me)}{SD}$$

$$P_2 = \frac{3(56 - 57,3)}{22,36} = -\ 0,174$$

نلاحظ أن النتيجة سالبة وضعيفة و بالتالي نقول أن التوزيع مائل نحو اليسار.

1)- طريقة يــول

$$C_{yk} = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{Q_3 - Q_1} = \frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1}$$

$$= \frac{72{,}35 - 2(57{,}3) + (40{,}67)}{72{,}35 - 40{,}67} = -0{,}05$$

النتيجة سالبة $C_{yk} < 0$ إذن إلتواء سالب فالتوزيع ملتوي نحو اليسار.

كما أن هناك مقاييس أخرى للإلتواء مُعرّفة بدلالة المئينات و التي تعرف بمعامل الإلتواء المئيني، و هي كالآتي:

$$C_{yK} = \frac{(P_{90} - P_{50}) - (P_{50} - P_{10})}{P_{90} - P_{10}} = \frac{P_{90} - 2P_{50} + P_{10}}{P_{90} - P_{10}}$$

2 -التفرطـح

يقيس التفرطح درجة تحدب التوزيع التكراري مقارنة مع شكل التوزيع الطبيعي، إذا كان التوزيع ذو قمة عالية أي محدبا يقال أنه حاد القمة، اما المنحنى الذي قمته مسطحة فيسمى منحنى مفرطح، أما منحنى التوزيع الطبيعي على هيئة جرس فهو منحنى معتدل[39]، كما هو موضح في الشكل التالي:

أ)- معامل بيرســون للتفرطـح

هو عبارة عن النسبة بين العزم المركزي من الدرجة الرابعة ومربع العزم المركزي من الدرجة الثانية، ولمعرفة إذا ما كان المنحنى مدبدبا(حاد) أو مفرطح لا بد من مقارنة معامل تفرطح المنحنى بمعامل منحنى التوزيع الطبيعي، لذا اعتمد كل من فيشر و بيرسون على العزم المركزي من الدرجة الرابعة(التي تتراوح قيمة المعامل من 1 إلى ∞ +) لآنه يساوي 3 في حالة التوزيع الطبيعي(أي توزيع على شكل جرس)، و يرمز للمعامل بالرمز P_2 ، و يتم حسابه وفقا للصيغة التالية:

[39] نداء محمد الصوص، مبادئ الإحصاء، المرجع سبق ذكره، 65.

$$P_2 = \frac{\mu_4}{SD^4} = \frac{\mu_4}{\mu_2^2}$$

علما أن $SD^2 = \mu_2$، و يمكننا أن نميز بين هذه الحالات:

– إذا كان $P_2 = 3$ فإن ....منحنى التوزيع معتدل.

– إذا كان $P_2 > 3$ فإن ....منحنى التوزيع مدبدب أو متطاول(تشتت ضعيف).

– إذا كان $P_2 < 3$ فإن ....منحنى التوزيع مفرطح(تشتت قوي بالنسبة لمركز التوزيع).

و في حالة الجداول التكرارية المفتوحة يمكن قياس التفرطح كما يلي:

بالإعتماد على الربيعيات و المئينيات و تسمى هذه الطريقة بمعامل التفرطح المئيني و صيغته كما يلي:

$$P_2 = \frac{IQ}{P_{90} - P_{10}} = \frac{(Q_3 - Q_1) / 2}{P_{90} - P_{10}}$$

في هذه الحالة معامل تفرطح منحنى التوزيع الطبيعي يساوي 0,263 بدلا من 3 في طريقة العزوم و بالتالي نميز بين 3 حالات:

– إذا كان $P_2 = 0,263$ فإن .... منحنى التوزيع متماثلا.

– إذا كان $P_2 > 0,263$ فإن .... منحنى التوزيع مدبدبا.

– إذا كان $P_2 < 0,263$ فإن .... منحنى التوزيع مفرطحا.

ب)– معامل فيشر للتفرطح

هو عبارة عن معامل بيرسون مطروحا منه 3، ويرمز له بـ $\mathbf{F_2}$ و يعطى بالصيغة التالية:

$$\mathbf{F_2} = P_2 - 3$$

بحيث يدرس حسب الحالات التالية:

– إذا كان$F_2 = 0$ فإن .... التوزيع متماثل طبعي.

– إذا كان $F_2 > 0$ فإن .... التوزيع متطاول(حاد).

– إذا كان $F_2 < 0$ فإن .... التوزيع مفرطح.

مثال 04: إليك الجدول التالي:

أحسب كل من معامل الإلتواء لفيشــر ثم معامل التفرطح

الفئات	n_i	x_i	$n_i x_i$	$(x_i-\bar{X})$	$n_i (x_i-\bar{X})^2$	$(x_i-\bar{X})^3$	$n_i(x_i-\bar{X})^3$	$(x_i-\bar{X})^4$	$n_i(x_i-\bar{X})^4$
3–7	5	5	25	− 6	180	− 216	− 1080	1296	6480
7–9	6	8	48	− 3	54	− 27	− 162	81	386
9–13	6	11	66	0	0	0	0	0	0
13–15	4	14	56	3	36	27	108	81	324
15–19	3	17	51	6	108	216	648	1296	3888
19–21	2	20	40	9	162	729	1458	6561	13122
المجموع	26	/	286	/	540	/	972	/	24300

الحل:

1)– حساب معامل الالتـواء لفيشــر

$$F_1 = \frac{\mu_3}{SD^3} \qquad \mu_3 = \frac{\sum n_i (X_i - \bar{X})^3}{\sum n_i}$$

يجب حساب أولا الوسط الحسابي الذي يساوي:

$$\bar{X} = \frac{\sum X_i n_i}{\sum n_i} = \frac{286}{26} = 11$$

ثم نحسب التباين الذي يساوي:

$$V(X) = \frac{\sum n_i (X_i - \bar{X})^2}{\sum n_i} = \frac{540}{26} = 20,769 \approx 20,77$$

ومنه قيمة الإنحراف المعياري يساوي:

$$SD = \sqrt{V(X)} = \sqrt{20,77} = 4,56$$

إذا نحسب معامل الإلتواء لفيشر الذي يساوي:

$$F_1 = \frac{\mu_3}{SD^3}$$

$$\mu_3 = \frac{\sum n_i (X_i - \bar{X})^3}{\sum n_i} = \frac{972}{26} = 37,38$$

$$F_1 = \frac{37,38}{(4,56)^3} = \frac{37,38}{94,81} = 0,39$$

$F_1 > 0$

إلتواء موجب إذن التوزيع مائل نحو اليمين.

2)- حساب معامل التفرطـح لبيرسـون

$$P_2 = \frac{\mu_4}{SD^4}$$

$$\mu_4 = \frac{\sum n_i (X_i - \bar{X})^4}{\sum n_i} = \frac{24300}{26} = 934,615$$

$$P_2 = \frac{934,62}{(4,56)^4} = \frac{934,62}{432,37} = 2,16$$

نلاحظ أن $P_2 < 3$ ومنه التوزيع مفرطح أي قوي التشتت

د.محمدي ص.

الجزائر 2022-06-14

و الحمـــــــــد لله

قائمة المراجع

قائمة المراجع:

1. إبراهيم علي إبراهيم عبد ربه، مبادئ علم الإحصاء وتطبيقاته باستخدام Excel 2000/XP، الدار الجامعية، الإسكندرية، مصر 2005.
2. أحمد شيبات، الاحصاء الوصفي، الرياضيات في الجامعة، جامعة منتوري، قسنطينة، بدون سنة.
3. تيلولت سامية، مبادئ في الاحصاء، دار الحديث للكتاب، القبة، الجزائر، ط3، 2016.
4. جلاطو جيلالي، الاحصاء مع التمارين ومسائل محلولة، ديوان المطبوعات الجامعية، 2007.
5. حسن ياسين طعمة، طرق الاحصاء الوصفي، دار صفاء، للنشر والتوزيع، عمان 2009.
6. سهيل أحمد سمحان، محمود حسين الوادي، مبادئ الاحصاء للاقتصاء والعلوم الادارية، دار صفاء للنشر والتوزيع، عمان،2010.
7. فتحي حمدان، كامل فليفل، مبادئ الاحصاء، دار المناهج للنشر والتوزيع، عمان، 2009.
8. محمد جبر المغربي، الاحصاء الوصفي، المكتبة المصرية للنشر والتوزيع، ط1، عملن، 2007.
9. موساوي عبد النور، بركان يوسف، دار العلوم للنشر والتوزيع، الجزائر 2009.
10. نداء محمد الصوص، مبادئ الإحصاء، دار أجنادين للنشر والتوزيع، ط1، عمان، 2007.

فهرس المحتويات

فهـرس المحتويـــات