الجمهورية الجزائرية الديموقراطية الشعبية

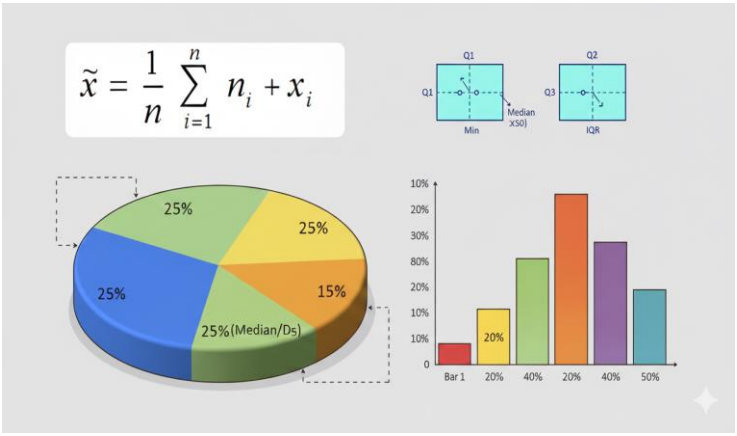
جامعة زيان عاشور بالجلفة

كلية العلوم الاجتماعية والإنسانية

قسم علم الاجتماع والديموغرافيا

مطبوعة بيداغوجية موجهة لطلبة السنة الأولى ليسانس (جذع مشترك علوم اجتماعية) ل م د في مقياس:

الإحصاء الوصفي



من إعداد:

د. شعلاني صلاح الدين

السنة الجامعية: 2026/2025

قائمة المحتويات

الموضوع	رقم الصفحة
المقدمة	
الفصل الأول: أساسيات علم الاحصاء	
1-1 ماهية علم الإحصاء وتطوره وعلاقته بالعلوم الأخرى	06
1-2 أهمية الإحصاء في العلوم الاجتماعية	07
1-3 مفاهيم عامة	09
1-4 أنواع الإحصاء	09
1-5 المجتمع الاحصائي	12
1-6 العينة	12
1-7 المتغيرات	13
1-7-1 تصنيف المتغيرات	13
1-7-2 تصنيف المتغيرات تبعا لعلاقتها السببية	17
1-8 أدوات جمع البيانات	17
1-8 مصادر جمع البيانات	17
الفصل الثاني: عرض البيانات الاحصائية	
2-1 جدول التوزيع التكراري لمتغير كمي متقطع	20
2-2 التمثيل البياني لمتغير كمي متقطع	21

2-3 جدول التوزيع التكراري لمتغير كمي مستمر	22
2-4 التمثيل البياني لمتغير كمي مستمر	22
2-5 جدول التوزيع التكراري لمتغير نوعي ترتيبي	23
2-6 التمثيل البياني لمتغير نوعي ترتيبي	24
2-7 جدول التوزيع التكراري لمتغير نوعي إسمي	26
2-8 التمثيل البياني لمتغير نوعي إسمي	27
الفصل الثالث: مقاييس التمركز	
3-1 مفهومها	28
3-2 المتوسط	28
3-2-1 تحديد قيمة المتوسط في حالة متغير متقطع	28
3-2-2 تحديد قيمة المتوسط في حالة متغير مستمر	29
3-2-3 خصائص المتوسط	30
3-4 الوسيط	31
3-4- 1 تحديد قيمة الوسيط في حالة متغير متقطع	31
3-4-2 تحديد قيمة الوسيط في حالة متغير مستمر	32
3-4-3 خصائص الوسيط	34
3-5 المنوال	34
3-5- 1 تحديد قيمة المنوال في حالة متغير متقطع	35

3-5-2 تحديد قيمة المنوال في حالة متغير مستمر	35
3-5-3 خصائص المنوال	36
الفصل الرابع: مقاييس التشتت	
4-1 مفهومها	38
4-2 المدى	38
4-3 الربيعيات والعشريات والمؤينات	38
4-3-1 الربيعيات	38
4-3-2 حساب الربيعيات في حالة البيانات غير المبوبة	39
4-3-3 حساب الربيعيات في حالة البيانات المبوبة	42
4-4 العشريات	45
4-4-1 حساب العشريات في حالة البيانات غير المبوبة	46
4-4-2 حساب العشريات في حالة البيانات المبوبة	46
4-5 المؤينات	49
4-5-1 حساب المؤينات في حالة البيانات غير المبوبة	49
4-5-2 حساب المؤينات في حالة البيانات المبوبة	50
4-6 التباين	52
4-7 الانحراف المعياري	53
4-8 الخطأ المعياري	53

الفصل الخامس: مقاييس العلاقة بين مقايس النزعة المركزية والتشتت	
5-1 معامل الاختلاف	55
الفصل السادس: مقاييس الشكل	
6-1 مفهومها	56
6-2 معامل الالتواء	56
6-2-1 معامل الالتواء لفيشر	56
6-2- 2 معامل الالتواء ليول	57
6-2- 3 معامل الالتواء لبيرسون	57
6-3 معامل التفرطح	58
6-3-1 معامل التفرطح لبيرسون	58
6-3-2 معامل التفرطح لفيشر	58
تمارين محلولة	60

مقدمة:

بسم الله، والصلاة والسلام على أشرف الأنبياء والرسل سيدنا محمّد عليه وعلى أهله أفضل الصلاة والتسليم.

أما بعد:

تضم هذه المطبوعة مجموعة من المحاضرات الموجهة لطلبة السنة الأولى ليسانس (جذع مشترك علوم اجتماعية) في مقياس الإحصاء الوصفي. وقد صيغت مادتها العلمية بناءً على الخبرة البيداغوجية لمؤلفها، الذي تولى تدريس هذا المقياس لعدة سنوات، مما أتاح له حصر الصعوبات التي تواجه الطلبة وتبسيط المفاهيم الأساسية بما يتوافق مع مقتضيات البرنامج الوزاري.

وفي الختام، نضع هذا العمل بين أيدي طلبتنا كمرجع مُبسط يرافقهم خلال مشوارهم الدراسي، آملين أن يسهم في تيسير استيعاب المادة العلمية وتحقيق أفضل النتائج، كما نؤكد أن هذا الجهد يبقى بشرياً يقبل النقد والتصويب، سائلين الله التوفيق والسداد لطلبتنا الأعزاء

الفصل الأول: أساسيات علم الاحصاء

1-1 ماهية علم الإحصاء وتطوره وعلاقته بالعلوم الأخرى:

يُعد علم الإحصاء من أقدم العلوم التي استخدمت منذ العصور القديمة، وتطوّر بشكل كبير عبر الزمن، ليصبح اليوم أداة أساسية في البحث العلمي وصنع القرار في مختلف المجالات.

أولا: البدايات القديمة:

تعود جذور الإحصاء إلى الحضارات القديمة مثل حضارة بابل ومصر القديمة، حيث استخدم الكهنة والمسؤولون تسجيل البيانات السكانية والزراعية والضريبية لأغراض إدارية وتنظيمية، في الصين القديمة، وُجدت سجلات سكانية تعود إلى 2000 سنة قبل الميلاد.

أما في الإمبراطورية الرومانية، فقد أجريت تعدادات سكانية دورية لغرض تحديد الضرائب والخدمة العسكرية.

ثانيا: العصور الوسطى:

في العصور الوسطى، ظل استخدام الإحصاء محدودًا وغالبًا ما كان يقتصر على التعدادات السكانية والضرائب في الدول الأوروبية، لم يكن هناك تطور منهجي في المفاهيم الإحصائية خلال هذه الفترة بسبب هيمنة الفكر الديني وانحسار النشاط العلمي.

ثالثا: القرن السابع عشر – بداية التأسيس العلمي:

شهد القرن 17 بداية تشكّل الأسس النظرية للإحصاء، خاصة من خلال أعمال العلماء مثل:

جون غرانت (John Graunt): أول من استخدم الإحصاء لتحليل سجلات الوفاة والمواليد في لندن، ويُعد من رواد الإحصاء الحيوي، بيير دي فيرما وبليز باسكال: أسهما في تطوير نظرية الاحتمالات، وهي حجر الأساس للكثير من المفاهيم الإحصائية.

رابعا: القرن الثامن عشر – نمو النظرية الإحصائية:

تم إدخال مصطلح "الإحصاء" (Statistics) لأول مرة للإشارة إلى البيانات التي تُجمع عن الدولة، حيث أصبح الإحصاء أداة تستخدم في فهم الظواهر السكانية والاقتصادية.

خامسا: القرن التاسع عشر – الرياضيات والإحصاء:

اندمج الإحصاء بشكل أعمق مع الرياضيات، وظهرت مفاهيم التوزيعات الاحتمالية، والانحدار الخطي، والانحراف المعياري، من أبرز الأسماء في هذه المرحلة: كارل بيرسون (Karl Pearson) الذي أسس المدرسة الحديثة للإحصاء الرياضي، وفرانسيس غالتون الذي أدخل مفهوم الارتباط.

سادسا: القرن العشرون – عصر الإحصاء التطبيقي:

شهد الإحصاء تطورًا مذهلًا مع ظهور الأساليب الإحصائية الحديثة مثل:

اختبار الفرضيات.

تحليل التباين (ANOVA).

النماذج الخطية واللوجستية.

تصميم التجارب.

من العلماء البارزين: رونالد فيشر (R.A. Fisher)، الذي يُعتبر مؤسس الإحصاء الحديث.

أصبح الإحصاء علمًا قائمًا بذاته يُدرّس في الجامعات ويُستخدم في مختلف التخصصات: الطب، الاقتصاد، الزراعة، النفس، الاجتماع...

سابعا: العصر الحديث – الإحصاء والبيانات الضخمة:

مع تطور التكنولوجيا وظهور الحواسيب، دخل الإحصاء مرحلة جديدة تعتمد على تحليل البيانات الضخمة (Big Data)، والتعلم الآلي، وبرامج التحليل الإحصائي.

أصبح علم الإحصاء أساسًا في مجالات الذكاء الاصطناعي، التسويق الرقمي، السياسة العامة، والتنبؤات الاقتصادية.

إذن كحوصلة عامة:

علم الإحصاء مرّ بمراحل تطورية طويلة، من تسجيل الأرقام في الحضارات القديمة إلى النماذج المعقدة والتحليل الحاسوبي في العصر الرقمي، واليوم يُعد الإحصاء أداة لا غنى عنها لفهم العالم واتخاذ قرارات مبنية على الأدلة والبيانات.

1-2 أهمية الإحصاء في العلوم الاجتماعية

إن الفائدة التي يقدمها لنا الاحصاء في العلوم الاجتماعية بشكل عام هو أنه يتيح لنا فهم وتفسير قدرا كبيرا من المعلومات بالنظر الى الحجم الهائل من البيانات التي تواجه الباحث خلال فترة معينة، مثلا ما هو معدل دخل الأسر؟ ما هو تباين الدخل؟ ما هي الأوقات التي يصل فيها النمو الديمغرافي الى الذروة؟ وباستخدام الإحصاء يمكننا الاجابة على هذه الأسئلة وتحليلها ووضع التفاسير للظاهرة المدروسة.

وفي علم السكان مثلا نواجه كميات هائلة من البيانات وكيف يؤثر تغير طفيف كان أم كبير (نقصد به متغير مستقل) في متغير تابع ما، هل هناك طريقة يمكن من خلالها

قياس هذه العلاقة؟ وما هي شـدة هذه العلاقة؟ وماذا تعني ذلك؟ .... فالإحصـاء يسـمح لنا للإجابة على هذه الأنواع من الأسئلة.

1-3 مفاهيم عامة:

تعريف الإحصاء:

هو أحد فروع الرياضـيات الهامة ذات التطبيقات الواسـعة بحيث يهتم بجمع وتلخيص وتمثيل وايجاد اسـتنتاجات من مجموعة البيانات المتوفرة، محاولا التغلب على مشـاكل مثل عدم تجانس البيانات وتباعدها، كل هذا يجعله ذا أهمية تطبيقية واسـعة في شـتى مجالات العلوم من الفيزياء إلى العلوم الاجتماعية وحتى الإنسانية، كما يلعب دورا في السياسة والأعمال.

تعريف:

هو العلم الذي يهتم بوصـف طرق متعددة لجمع البيانات والمشـاهدات ومن ثم يتم تنظيمها وعرضها باستخدام الأساليب العلمية لتحليلها واستخلاص النتائج منها.

يشـار للإحصـاء إلى أنه مجموعة من الأسـاليب العلمية القياسـية التي يمكن توظيفها لجمع المعطيات (البيانات والمعلومات) عن الظواهر وتبويبها وتلخيصـها وتقييمها والخروج منها باستنتاجات حول مجموعة وحدات المجتمع من خلال اعتماد جزء صغير من هذا المجتمع.

1-4 أنواع الإحصاء:

يمكن تقسيم علم الاحصـاء الى قسمين: احصـاء وصـفي وإحصـاء استدلالي (نقول عنه كذلك إحصـاء رياضـي)، الهدف من الاحصـاء الوصـفي هو تحويل عدة أرقام ومعطيات الى شكل موجز وواضح يعبر عن أهم ما يوجد في هذه المعطيات، أما الاحصاء الرياضي الغاية منه هو تحويل المشـاهدات الاحصـائية لاسـتخلاص بعض المعلومات عن مجتمع والذي هو عبارة عن جميع المفردات التي يمكن أن يأخذها المتغير.

وتتمثل المقارنة بينهما في أن:

- الإحصاء الوصفي يهتم بجمع البيانات ووصف الظاهرة من هذه البيانات.

– الإحصـــاء الاســـتدلالي يســـاعد في تقديم التوصـــيات واتخاذ القرارات التي تتعلق بالمجتمع في ضوء بيانات العينة.

كما تتمثل وظيفة الإحصاء الوصفي فيما يلي:

تلخيص البيانات وتقديم صـورة عامة وواضـحة حول الظواهر موضـوع البحث، وهو يتضـمن طرق جمع البيانات وعرضـها، ومقاييس النزعة المركزية، ومقاييس التشـتت، ومقاييس الشـكل بالإضافة الى دراسة معاملي الارتباط والانحدار.

يمكن تلخيصها فيما يلي:

– وظيفة العدِّ والحَصر، وتعتبر من أهمّ وظائف علمِ الإحصاء.

– استقطاب البيانات، وجمعها من مصادرها؛ حتى يتم دراسة المُجتمع وِفقاً للمعلومات المتوفرة حوله، والمعطيات المُشكلة التي تواجهه.

– وصـــف توزيع مجموعة كبيرة من القيم الخاصـــة بأحد المتغيرات (حســـاب مقاييس التمركز ومقاييس التشـتت)، ولا تقتصـر تلك الأسـاليب على دراسـة متغير واحد، لا تتعدى الى دراسة متغيرين (العلاقة بين المتغيرات).

– تحليلُ البيانات والمعلومات (التَّحليل الكمي والنوعي للبيانات).

أي أن وظائف الاحصاء الوصفي هي:

1- جمع البيانات.

2- العرض الجدولي للبيانات.

3- العرض البياني في صورة رسم على هيئة مستطيلات أو منحنيات أو أعمدة أو دوائر.

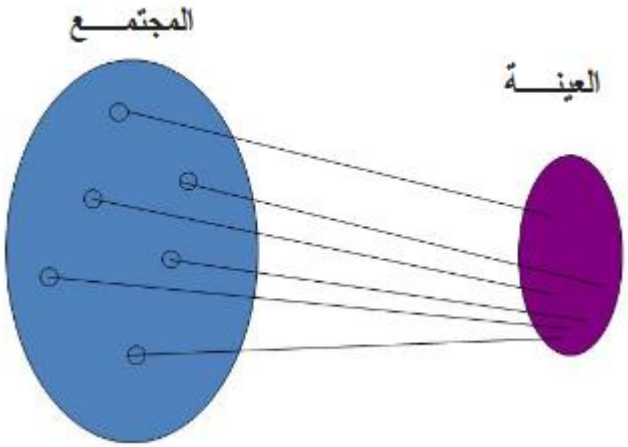
4- تلخيص البيانات بحساب بعض المقاييس الإحصائية الأساسية مثل

- مقاييس النزعة المركزية (المتوسط–الوسيط–المنوال....).
- مقاييس التشتت (المدى–التباين–الانحراف المعياري.....).

- مقاييس الالتواء والتفرطح.
- مقاييس الارتباط والانحدار.

أما **وظيفة الإحصاء الاستدلالي** تتمثل فيما يلي:

يهتم بأساليب التحليل التي تستدعي المقارنة والاستنتاج والحكم على ظواهر المجتمع من خلال بيانات العينة، أي تلك العملية التي تؤدي الى استخلاص النتائج العامة من النتائج الجزئية (عملية الاستقراء) وفقا لقوانين معينة يعني ذلك اجراء عملية التعميم النتائج، والسبب الذي يجعلنا نعتمد على أسلوب العينة بدل من اجراء عملية الحصر الشامل كثيرة منها توفير الجهد والمال والوقت.



1-5 المجتمع الاحصائي la population

تعريف 1: هو مجموعة عناصر متجانسة من الأشياء أو المفردات المراد إجراء دراستها. على سبيل المثال: طلبة الجامعة، المكلفين بالضريبة، عمال مؤسسة...، ويقسم الى مجتمع محدود وغير محدود[1].

تعريف 2: هو مجموعة كبيرة من الأفراد أو الأشياء أو الوحدات أو العناصر أو الأحداث التي تشترك فيما بينها في صفة أو أكثر ويرغب الباحث في دراستها وتحليل المعلومات المتوفرة عنها.

والمجتمع قد يكون محدد العدد مثل مجتمع طلاب الجامعة أو مجتمع عمال في مؤسسة ما وفي هذه الحالة يمكن دراسة جميع أفراد المجتمع أو أخذ عينة تكون ممثلة تمثيلا جيدا. وقد يكون المجتمع غير محدد مثل مجتمع الحشرات في غابة ما أو مجتمع الأسماك في المحيط أو مجتمع الأشجار، وهذه المجتمعات لا يمكن دراسة جميع أفرادها ولذلك نلجأ إلى استخدام أسلوب العينات لدراستها[2].

1-6 العينة L'échantillon

تعريف:

هي جزء من المجتمع الاحصائي، حيث تمثل صفاته وخصائصه، وهي عبارة عن عدد معين من المفردات (يسمى حجم العينة ويرمز له بالرمز n) مأخوذ من مجتمع ما، وتستخدم هذه العينة لعمل وصف وتحليل واستدلال حول خصائص المجتمع أي حول معالم المجتمع.

ويلجأ الباحث الى استخدام العينة عندما لا يمكن في بعض الأحيان دراسة جميع مفردات المجتمع الاحصائي نظر اً لاتساعه، فيتم اللجوء إلى دراسة عينة عشوائية منه.

---

[1] - James E. Burt, Gerald M. Barber, David L. Rigby , Elementary Statistics for Geographers, Third Edition , Guilford press , New York, 2009 ,p :5

[2] – محمد بداوي ، الإحصاء الوصفي ، دار هومة ، لجزائر ، 2017 ، ص : 17.

إن اختيار العينة بشكل دقيق ومناسب يعطي نتائج مشابهة إلى حد كبير للنتائج التي يمكن الحصول عليه عند دراسة كامل مجتمع الدراسة، وبشكل عام كلما كان حجم العينة أكبر كلما زاد تمثيلها لخصائص المجتمع موضوع الدراسة.

1-7 المتغيرات Les variables

لفظ "متغير" من الألفاظ التي يتكرر ورودها في علم الاجتماع وباقي العلوم الاجتماعية، ولا يمكن للباحث القيام بعملية التحليل دون استخدام هذا المصطلح، فالمتغيرات هي أساس البناء لأسئلة البحث وصياغة فرضياته[1].

تعريف:

هو مقدار له خصائص كمية ونوعية، تتغير قيمته من عنصر الى أخر من عناصر المجتمع الاحصائي أو العينة.

ملاحظة: نعتبر هنا متغير إحصائي وحيد نرمز له بـــــ X ، الهدف من هذا هو البحث عن أدوات أساسية تتكيف مع طبيعة X وتسمح بتمثيله.

1-7-1 تصنيف المتغيرات: يمكن تقسيمها الى قسمين متغيرات كمية ومتغيرات نوعية.

أولا: المتغير الكمي variable quantitative:

تعريف: نسمي صفة كمية عندما تكون أوضاعها المختلفة قابلة للقياس بأعداد تعطي شدتها، والمتغيرات الكمية تأخذ قيما عددية، لذا فإنه يطلق عليها اسم المتغيرات الكمية. ومن أمثلة المتغيرات الكمية: أعداد الطلبة في كل قسم علمي بكلية العلوم الاجتماعية في جامعة يان عاشور، الوزن (بالكيلوغرام) لعينة من الاطفال الدارسين بالابتدائي في مدينة

[1] – محمد بداوي ، مرجع سابق ، ص 22 .

معينة، العمر (بالسنوات) للموظفين العاملين بإحدى الشركات، وتنقسم المتغيرات الكمية الي قسمين هما[1]:

أ- المتغير الكمي المنفصل (المتقطع)variable quantitative discrète :

وهي المتغيرات الكمية التي تأخذ قيما عددية محددة صحيحة ولا تحتوي على قيم كسرية مثل: عدد المصانع في كل مدينة من مدن دولة ما، عدد حوادث السيارات التي وقعت في الربع الاول في سنة معينة في عدة مناطق مختلفة، وعدد المدرسين في كل مدرسة من المدارس الثانوية في مدينة ما خلال سنة معينة. والتمثيل المألوف لهذه المتغيرات هو على شكل أعمدة بيانية[2].

ب- المتغير الكمي المتصل:

وهو الذي يمكن قياسه بمعايير وحدوية والتي تأخذ أي قيمة يمكن تمثيلها في مدى معين من الأعداد الحقيقية. فمثلا دخل أسرة أو كميات الأمطار أو وزن المنتوج... إلخ.

في غالب الأحيان يتمثل المتغير الكمي المتقطع في القيم التي يمكن تعدادها وحصرها بعد عدد الوحدات الموجودة بحيث يكون معيار العد غير قابل للتجزئة أو التقسيم. أما المتغير الكمي المتصل يتمثل في القيم التي يمكن قياسها بواسطة معيار وحدوي متفق عليه قابل للتجزئة العشرية (مثل الكيلوغرام أو المتر)

ثانيا: المتغير النوعي variable qualitative:

نسمي الصفة الكيفية modalité لما تكون أوضاعها غير قابلة للتعيين إلا بنوعيتها، حيث أن هذه المتغيرات تصنف إلى مفردات (أو عناصر) المجتمع الاحصائي أو العينة في عدة

[1] - نفس المرجع ، ص: 23 .

[2] - Réjean Huot , Méthodes quantitatives pour les sciences humaines 2 édition , Les presses de l'université de laval, 2003 , p : 57

مجموعات، تشترك كل مجموعة في صفة معينة، ومن أمثلة المتغيرات النوعية أو الوصفية: متغير نوع العمل الذي يصنف المفردات الي أربع مجموعات هي (منتج، موظف، عمل حر، أو متقاعد)، ومتغير الجنس الذي يصـــنف مفردات المجتمع أو العينة إلى مجموعتين هما (ذكر، أنثى)، ومتغير الحالة الاجتماعية (أعزب، متزوج)، كذلك متغير مســتوى الدخل الذي يمكن تصنيفه إلى: عالي، متوسط، منخفض.

إذا كان هذا المتغير قابل للترتيب عندئذ يسـمى بمتغير ترتيبي، وإذا كان غير ذلك (أي غير قابل للترتيب) فيسمى بمتغير اسمي. ويمكن تقسيمه الى نوعين إسمي وترتيبي[1].

أ- المتغير الترتيبي variable ordinale:

تعريف:

يقال بأن المتغير النوعي متغير ترتيبي إذا أمكن وضـــع صـــفاته أو أقســـامه في ترتيب تصاعدي (أو تنازلي). ومن أمثلة المتغيرات الترتيبية: المستوى التعليمي والذي يمكن تصنيفه إلى [عالي، متوســـط، منخفض]، التقديرات المختلفة للنجاح في الثانوية [مقبول، جيد، جيد جدا، ممتاز]، تأييد قضية معينة والذي يمكن تصنيفه إلى [موافق، متردد، لا أوافق][2].

مثال:

بافتراض أننا قمنا بجمع تقديرات 30طالبًا في اختبار مقياس ما في قسم علم الاجتماع.
والتقدير هو المتغير الوصفي الترتيبي، وفئاته يمكن ترتيبها من الأفضل إلى الأدنى: ممتاز، جيد جداً، جيد، ومقبول.

البينات ترتب على الشكل التالي:

[1] - شيبات أحمد، الاحصاء الوصفي، دار الهدى، عين مليلة – الجزائر-، 2001، ص27.

[2] - Réjean Huot , op cit , p :58

- ممتاز 5 : طلاب
- جيد جداً 10 : طلاب
- جيد 12 : طالبًا
- مقبول 3 : طلاب

ب- المتغير الإسمي variable nominale:

تعريف:

يقال بأن المتغير النوعي متغير اسمي عندما لا تخضع صفاته لأي معيار من معاير الترتيب، وتكون الصفات او الانواع في المتغيرات الاسمية عبارة عن اسماء، ومن أمثلة المتغيرات الاسمية: لون الشعر: الذي يمكن تصنيفه إلى [أسود، بني، أصفر...] الجنس: الذي يمكن تقسيمة إلى [ذكر، أنثى][1].

مثال:

البيانات التالية تبين تخصصات 20 طالبا في إحدى الجامعات. والتخصص هنا هو المتغير الوصفي الإسمي، وفئاته هي: "الهندسة"، "الطب"، و"الآداب."

الهندسة	الآداب	الهندسة	الطب	الآداب
الآداب	الهندسة	الهندسة	الهندسة	الطب
الطب	الآداب	الطب	الآداب	الطب
الهندسة	الطب	الهندسة	الطب	الهندسة

[1] - محمد بداوي، مرجع سابق، ص 25.

1-7-2 تصنيف المتغيرات تبعا لعلاقتها السببية:

يسمى المتغير الذي يؤثر في متغير آخر أو أكثر بالمتغير المستقل (variable independent)، أما المتغير الذي يحدث فيه التغير أو الأثر فيسمى بالمتغير التابع (variable dependent)[1].

مثال:

هناك علاقة بين تحصيل الطلاب وعدد ساعات المراجعة أسبوعيا.

عدد ساعات المراجعة: متغير مستقل تحصيل الطلاب: متغير تابع

1-8 أدوات جمع البيانات

1-8-1 تعريف العناصر (les éléments, les individus): هي المفردات أو الوحدات المكونة للمجتمع.

1-8-2 تعريف (البيانات les données):

هي مجموعة المشاهدات المأخوذة أثناء دراسة معينة، وتقسم الى بيانات كمية quantitative مثل الطول والوزن والدخل ...، وبيانات نوعية qualitative مثل الجنس والمؤهل العلمي والحالة العائلية ...[2]

1-9 مصادر جمع البيانات

أما عن مصادر جمع البيانات نقسمها الى قسمين وهي:

1-9-1 المصدر الأول(تاريخي): وهو ما يؤخذ من السجلات المحفوظة مثل سجلات المواليد والوفيات، واحصائيات من هيئات حكومية وغير حكومية.

[1] - نفس المرجع، ص26.

[2] - نفس المرجع ، ص : 26 .

تعريف الوثيقة:

" كل وسيط يقدم معلومات أو حقيقة أو يساعد على تقديم حقيقة ما وتكون له الصفة القانونية".

فالوثيقة معلومة كتبت لحفظ الحقوق أو الهوية أو لتسيير النظام الإداري او لارتباطها بعلاقات الناس وتكون لها الصفة القانونية.

1-9-2 المصدر الثاني: وهو عبارة عن البينات التي جمعت من مفردات المجتمع الاحصائي كله (حصر شامل)، أو جزء منه (اختيار عينة)، ويكون ذلك على طريقتين: إما إجراء مقابلة شخصية مع المبحوثين، أو استخدام تقنية الاستبيان (Questionnaire).

يحدد الباحث في المرحلة الأولى شكل البيانات اللازمة لاختبار فروضه، بفحص ما يتيسر له من أدوات، ليختار أكثرها ملاءمة لتحقيق هدفه، وهو في هذا قد يحتاج إلى تعديل بعض أدواته، يبدأ الباحث بطرح إشكال ثم يضع فروضا تعتبر حلولا محتملة، بحيث أن نوع المشكلة وطبيعة الفروض هي التي تتحكم في اختيار الأدوات. وقد يتطلب بحث من الباحث عددا قليلا من الأدوات ويتطلب بحث آخر عددا أكبر. لذلك يجب أن يتوفر لدى الباحث مجموعة من الأدوات والتي هي مجموعة الوسائل والطرق والأساليب والإجراءات المختلفة التي يعتمد عليها في الحصول على المعلومات والبيانات اللازمة لإتمام وإنجاز البحث حول موضوع محدد أو مشكلة معينة.

1-9-2-1 المقابلة

هي أسلوب من أساليب جمع البيانات، يقوم فيه الباحث بإجراء محادثة موجهة مع فرد أو مجموعة من الأفراد، بهدف الحصول على معلومات دقيقة وهادفة تتعلق بظاهرة معينة قيد الدراسة، وتتم المقابلة من خلال طرح الباحث لمجموعة من الأسئلة المعدّة مسبقًا، والتي يُنتظر من المبحوثين الإجابة عنها بشكل يعكس تجاربهم أو آرائهم أو معارفهم حول

موضوع البحث، وتُجرى المقابلة وفق منهجية علمية تضمن تحقيق أهداف البحث وتقليل التحيز[1].

1-9-2-2 الاستبيان

تعريف: هو مجموعة من الأسئلة المتنوعة والتي ترتبط مع بعضها البعض بشكل يحقق هي مجموعة مؤشرات يمكن عن طريقها اكتشاف أبعاد موضوع الدراسة عن طريق الاستقصاء التجربة أي إجراء بحث ميداني على جماعة محددة من الناس، وهي وسيلة الاتصال الرئيسية بين الباحث والمبحوث.[2]

[1] – نفس المرجع ، ص 27 .

[2] – سلاطنية بلقاسم، حسان الجيلالي، أسس المبحث العلمي، ديوان المطبوعات الجامعية، الجزائر، 2007، ص 77.

الفصل الثاني: عرض البيانات الاحصائية

2-1 جدول التوزيع التكراري لمتغير كمي متقطع:

نفرض أن البيانات تحوي على n مشاهدة، نقوم بترتيب هذه القياسات (المشاهدات) ولنفرض أن من بينها فقط r قياسا مختلفا هي القيم:

$$x_1, x_2, ..., x_r \qquad r \le n$$

في العمود الأول يتكون من r مشاهدة مختلفة للمتغير X ، هذه المشاهدات ترتب تصاعديا نرمز لها $\{x_i : i = 1,...,r\}$ ، في العمود الثاني نجد تكرار القيم x_i نرمز لها $\{n_i\}$ ، أحيانا n_i تستبدل بالمقادير $f_i = \frac{n_i}{n}$ (تدعى بالتكرارات النسبية)، n ترمز لعدد المشاهدات في المثال السابق نجد $n = \sum_{i=1}^{r} n_i$ ، التكرار المتراكم والتكرار المتراكم النسبي هما على التوالي[1]:

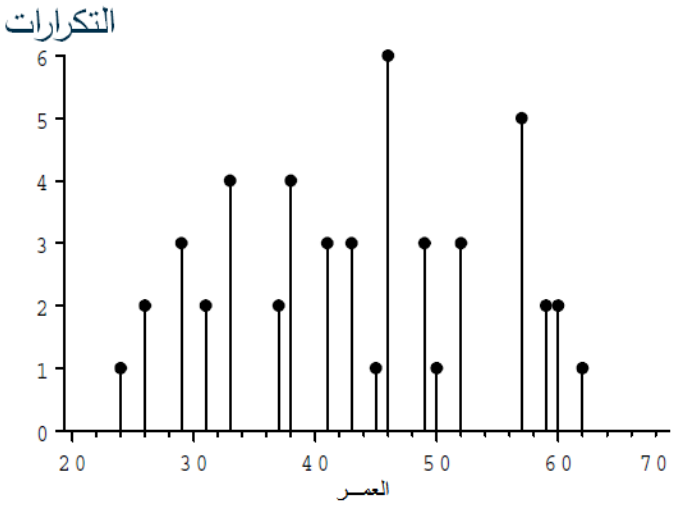
$N_i = \sum_{j=1}^{i} n_j$; $F_i = \sum_{j=1}^{i} f_j$. ويكون الجدول التكراري كما يلي:

x_i	n_i	N_i	$f_i(\%)$	$F_i(\%)$
24	1	1	2,08	2,08
26	2	3	4,17	6,25
29	3	6	6,25	12,50
31	2	8	4,17	16,67
33	4	12	8,33	25,00
37	2	14	4,17	29,17
38	4	18	8,33	37,50
41	3	21	6,25	43,75
43	3	24	6,25	50,00
45	1	25	2,08	52,08
46	6	31	12,50	64,58
49	3	34	6,25	70,83
50	1	35	2,08	72,91
52	3	38	6,25	79,16
57	5	43	10,42	89,58
59	2	45	4,17	93,75
60	2	47	4,17	97,92
62	1	48	2,08	100,00
المجمـــوع	48	—	100,00	—

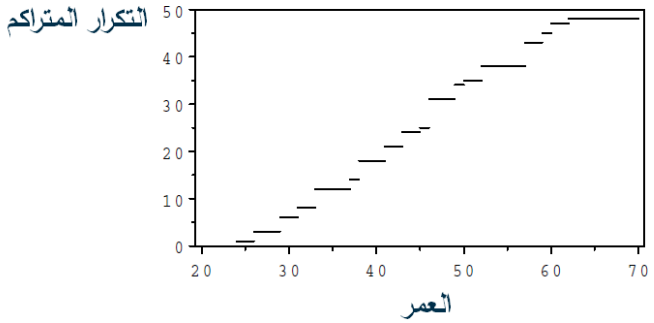
[1] عمر قاسم عزات ، مبادئ الاحتمالات والإحصاء، منشورات جامعة دمشق، 1995 ، ص 64.

2-2 التمثيل البياني لمتغير كمي متقطع:

وتمثيل قيم الجدول السابق على شكل أعمدة بيانية يكون كما يلي:



والشكل الاتي يبين عرض التكرار المتراكم



2-3 جدول التوزيع التكراري لمتغير كمي مستمر:

مثال: الجدول التالي يبين توزيع النفقات الأسبوعية لـ 100 طالب في دولة ما (الوحدة: وحدة نقدية)

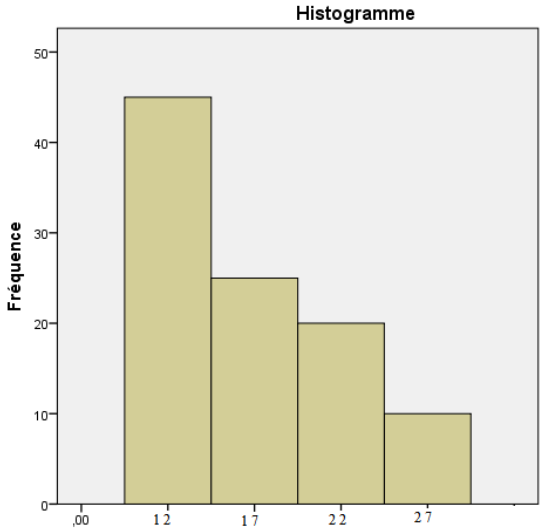
الفئات	14-10	19-15	24-20	29-25	المجموع
التكرار n_i	45	25	20	10	100

	التكرار n_i	التكرار النسبي $f_i(\%)$	التكرار النسبي المتراكم $F_i(\%)$
10-14	45	45,0	45,0
15-19	25	25,0	70,0
20-24	20	20,0	90,0
25-29	10	10,0	100,0
المجموع	100	100,0	

4-2 العرض البياني لمتغير كمي مستمر (مدرج تكراري):

نقوم بتمثيل بياني يمثل توزيع نفقات بواسطة مدرج تكراري، ويكون كما يلي:

والشكل الاتي يبين عرض المدرج التكراري



ملاحظة: يتم حساب طول الفئة انطلاقا من القانون الآتي:

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

المدى: أكبر قيمة – أصغر قيمة $R = x_r - x_1$ L'étendue :

K عدد الفئات تحسب من صيغة ستيرجس formule de sturges في حالة لم تكن كمعطى في البيانات.

$$K = 1 + 3.322 Log(n)$$

2-5 جدول التوزيع التكراري لمتغير نوعي ترتيبي:

مثال:

في إحدى الدراسات التي شملت 80 مبحوثا طلب منهم الإجابة عن سؤال المستوى التعليمي، فكانت النتائج كما يلي:

جامعي	جامعي	جامعي	جامعي	ثانوي	ثانوي	ثانوي	جامعي	جامعي	جامعي
جامعي	جامعي	جامعي	متوسط	متوسط	متوسط	متوسط	ثانوي	ثانوي	ثانوي
متوسط	جامعي	ابتدائي	ابتدائي	ابتدائي	ابتدائي	جامعي	جامعي	جامعي	جامعي
متوسط	جامعي	جامعي	جامعي	جامعي	ثانوي	ثانوي	ثانوي	ثانوي	ثانوي
متوسط	جامعي	جامعي	جامعي	جامعي	جامعي	متوسط	متوسط	متوسط	متوسط
متوسط	جامعي	جامعي	جامعي	جامعي	ثانوي	ثانوي	ثانوي	ثانوي	ثانوي
ثانوي	ثانوي	ثانوي	ثانوي	ثانوي	ابتدائي	ابتدائي	جامعي	جامعي	جامعي
ثانوي	ثانوي	متوسط	متوسط	متوسط	متوسط	ثانوي	ثانوي	ثانوي	ثانوي

المطلوب:

عرض بيانات الجدول بواسطة جدول تكراري؟

حل المثال:

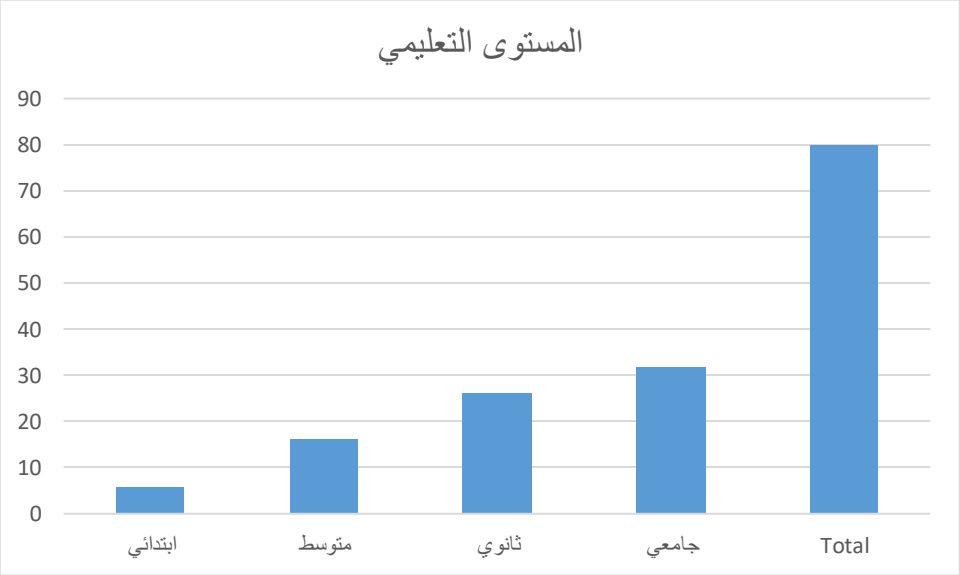
جدول تكراري لتقديرات لمتغير المستوى التعليمي

المستوى التعليمي	التكرارات	النسبة
ابتدائي	6	2,5
متوسط	16	18,8
ثانوي	26	32,5
جامعي	32	40,0
المجموع	80	100,0

6-2 التمثيل البياني لمتغير نوعي ترتيبي:

التمثيل البياني بالنسبة للمتغير الرتبي يتم بواسطة الأعمدة، وهو أكثر الطرق شيوعا.

أعمدة بيانية لمتغير المستوى التعليمي



2-7 جدول التوزيع التكراري لمتغير نوعي إسمي:

مثال:

في إحدى الدراسات التي شملت 80 مبحوثا طلب منهم الإجابة عن سؤال المهنة، فكانت النتائج كما يلي:

موظف	عامل	عمل حر	بدون عمل	عامل	موظف	موظف	موظف	موظف	موظف
بدون عمل	عامل	عمل حر	موظف	عمل حر	عامل	عامل	عامل	عامل	عامل
عمل حر	عمل حر	عمل حر	موظف	عامل	عمل حر	موظف	موظف	موظف	موظف
عامل	عامل	عامل	موظف	بدون عمل	موظف	موظف	موظف	موظف	موظف
موظف	موظف	موظف	موظف	موظف	موظف	موظف	موظف	موظف	موظف
موظف	موظف	موظف	موظف	عامل	موظف	موظف	موظف	عامل	عامل
بدون عمل	عامل	موظف	موظف	موظف	موظف	عامل	عامل	عامل	عامل
موظف	موظف	موظف	موظف	موظف	موظف	موظف	موظف	موظف	عامل

المطلوب: تقديم متغير نوع المهنة في شكل جدول تكراري؟

حل المثال:

جدول تكراري لمتغير المهنة

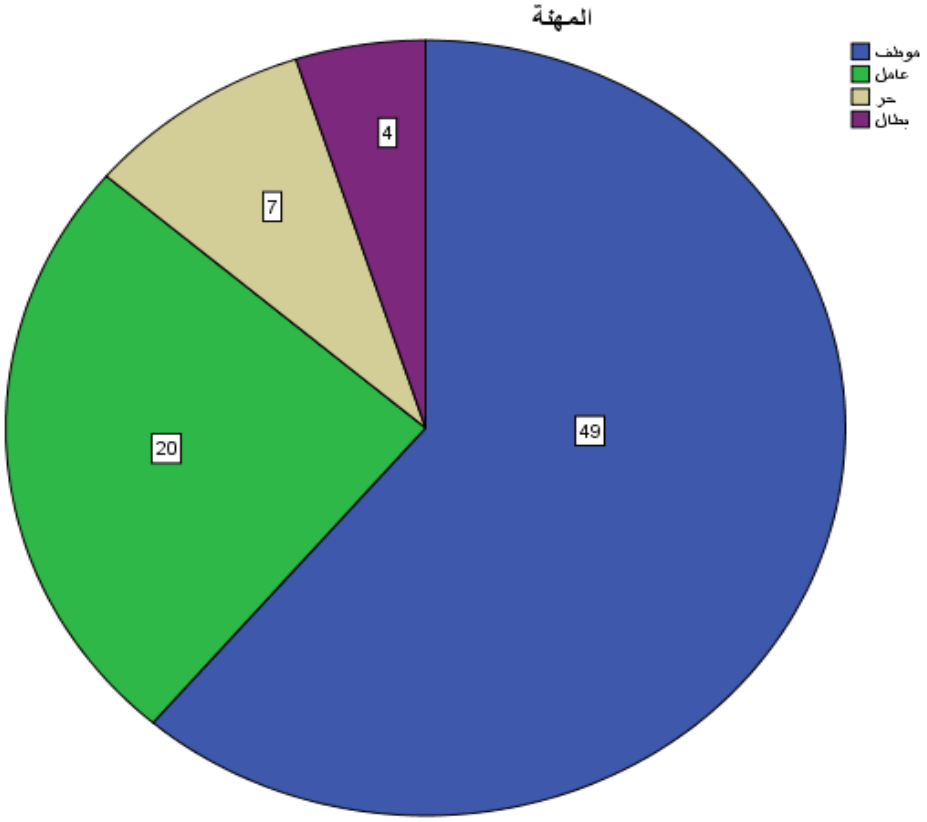
نوع المهنة	التكرارات	النسبة
موظف	49	61,3
عامل	20	25,0
عمل حر	7	8,8
بدون عمل	4	5,0
المجموع	80	100,0

2-8 التمثيل البياني لمتغير نوعي إسمي:

للتمثيل البياني في المتغير الاسمي تعتبر الدائرة النسبية من أفضل وأكثر الطرق شيوعاً لتمثيله بيانيا.

مثال: مثل بيانيا على شكل دائرة نسبية متغير المهنة.

تمثيل متغير المهنة على شكل دائرة نسبية:



الفصل الثالث: مقاييس النزعة المركزية

Les mesures de position

3-1 مفهومها: هي قيمة تتمركز حولها البيانات، من أنواعها نجد: المتوسط والوسيط والمنوال[1].

3-2 المتوسط La Moyenne

تعريف: هو مجموع قيم المشاهدات على عدد المشاهدات ويرمز له بـ $\bar{x}$.

3-2-1 تحديد قيمة المتوسط في حالة متغير متقطع

حساب المتوسط لهذا النوع يكون كما يلي:

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

مثال:

نحسب متوسط السلسلة الاتية: 2،2،3،4،5،5،6،6،2،4

$$\bar{x} = \frac{1}{10}\sum_{i=1}^{10}(2+2+3+4+5+5+6+6+2+4) = 3.9$$

وهي محصلة جمع عشرة حدود مقسمة على 10، ولكننا نلاحظ أن هناك خمسة قيم فقط (2،3،4،5،6) كل قيمة مكررة عدد من المرات، في هذه الحالة نستخدم الصيغة الاتية:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} n_i \cdot x_i$$

[1] - CESAR PEREZ LOPEZ , DATA STATISTICAL ANALYSIS WITH R , Publié par CESAR PEREZ , 2012, p : 59

$$\overline{x} = \frac{1}{10}\sum_{i=1}^{10}(2\times 3+3\times 1+4\times 2+5\times 2+6\times 2) = 3.9$$

ويمكننا استخدام التكرارات النسبية في حسابنا للمتوسط، بحيث:

$$\sum_{i=1}^{n} f_i \cdot x_i$$

في مثالنا السابق نجد:

x_i	n_i	f_i	$f_i \cdot x_i$
2	3	0.3	0.6
3	1	0.1	0.3
4	2	0.2	0.8
5	2	0.2	1
6	2	0.2	1.2
المجموع			3.9

3-2-2 تحديد قيمة المتوسط في حالة متغير مستمر

حساب المتوسط لهذا النوع يكون كما يلي:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{k} n_i \cdot c_i$$

مثال:

نستخدم بيانات المثال () والذي يبين توزيع النفقات الأسبوعية لـ 100 طالب في دولة ما (الوحدة: وحدة نقدية)

الفئات	14-10	19-15	24-20	29-25	المجموع
التكرار n_i	45	25	20	10	100

المطلوب: حساب المتوسط الحسابي؟

حل المثال:

الفئات	التكرار n_i	c_i	$n_i \cdot c_i$
14-10	45	12	540
19-15	25	17	425
24-20	20	22	440
29-25	10	27	270
المجموع	100		1675

ويكون المتوسط كما يلي: $\bar{x} = \frac{1}{n}\sum_{i=1}^{k} n_i \cdot c_i = \frac{1675}{100} = 16.75$

3-2-3 خصائص المتوسط [1]:

يعتمد على جميع القيم والمشاهدات محل الدراسة

هو نقطة اتزان المشاهدات

انحراف المشاهدة عن المتوسط يساوي بعدها عنه.

المجموع الجبري لانحرافات القيم عن وسطها الحسابي يساوي صفراً.

يتأثر بالقيم المتطرفة والقيم الشاذة لذا لا يصلح للتوزيعات الملتوية ويقودنا لعدم صلاحية الوسط الحسابي لتوزيعات تكرارية شديدة الالتواء.

الحالات التي لا يصلح المتوسط الحسابي فيها:

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

2- إذا كانت هناك قيم شاذة تتخلل التوزيع.

3- إذا كانت البيانات مبوبة في فئات وكان التوزيع مفتوحا في أحد طرفيه.

4- إذا كانت الفئات متباعدة نسبيا.

[1] - نفس المرجع ، ص: 47 .

5- لا يمكن إيجاده من الرسم البياني كالوسيط والمنوال.

6. لا يمكن إيجاده من التوزيع التكراري المفتوح إلا في حال تقدير مراكز الفئات المفتوحة.

7. عدم صلاحية الوسط الحسابي لتوزيعات تكرارية شديدة الالتواء.

3-4 الوسيط La Médiane [1]

تعريف: هو أحد مقاييس التمركز، وهو القيمة التي تتوسط مجتمع وتقسمه الى نصفين متساويين.

3-4-1 تحديد قيمة الوسيط في حالة متغير متقطع

لتحديد قيمة الوسيط في هذه الحالة نقوم بترتيب القيم (n) ترتيب تصاعدي أو تنازلي، ونميز حالتين:

- n فردي (أي يكتب $n=2p+1$) عندئذ تعطى عبارة الوسيط بـ $Me=x_{p+1}$

- n زوجي ($n=2p$) عندئذ تعطى عبارة الوسيط بـ $Me=\frac{x_{p+1}+x_p}{2}$

مثال 1: اوجد الوسيط للقيم التالية: 20، 10، 15، 10، 25، 105، 30

بعد ترتيب هذه القيم تصاعديا، نجد: 10، 10، 15، 20، 25، 30، 105

عدد القيم فردي، إذن الحد الأوسط هو 20 ومنه الوسيط = 20

ملاحظة: رتبة الوسيط هي $\frac{n+1}{2}$ في هذه الحالة ومنه الرتبة هي $\frac{7+1}{2}=4$ الوسيط هو الحد الرابع.

مثال 2: لنحسب الوسيط للقيم التالية: 11، 65، 12، 13 ،15،4

بعد ترتيب هذه القيم تصاعديا: 4، 11 ،12، 13 ،15، 65

[1] - CESAR PEREZ LOPEZ , OP Cit , p :61

عدد القيم زوجي و نلاحظ أنه هناك حدين هما 12و13 فالوسيط يكون: $\frac{13+12}{2}=12.5$

3-4-2 تحديد قيمة الوسيط في حالة متغير مستمر

التكرارات المتراكمة تسمح لنا بمعرفة الفئة الوسيطية، من خلالها نجد الوسيط بواسطة الاستقطاب الخطي interpolation linéaire .

لكي نجد الوسيط نقوم برسم:

- المنحنى التكرار المتراكم (تكرار المتجمع الصاعد)، نحسب رتبة الوسيط وهي $\frac{n}{2}$ بغض النظر عما إذا كان n زوجي أو فردي.
- المنحنى التكرار المتراكم النسبي (تكرار المتجمع الصاعد النسبي)، نحسب رتبة الوسيط وهي 0.5

لحساب فاصلة هذه النقطة نستخدم إما خصائص ميل المستقيمات أو نستخدم مبرهنة طاليس، هذا بالنسبة للطريقة البيانية أما بالنسبة للطريقة الجبرية فنستخدم الصيغة الاتية:

$$Me = L_1 + \frac{\frac{n}{2} - N_0}{n_1} \times c$$

L_1 : الحد الأدنى للفئة الوسيطية، n_1 : تكرار الفئة الوسيطية، N_0 : التكرار المتجمع الصاعد السابق للفئة الوسيطية، c : طول الفئة الوسيطية.

مثال:

اليك جدول التوزيع التكراري الاتي:

الفئات	5 -0	10-5	15-10	20-15	25-20	المجموع
التكرار n_i	10	8	12	11	9	50

المطلوب: إيجاد قيمة الوسيط حسابيا وبيانيا؟

حل المثال:

أ- الطريقة الجبرية:

الفئات	التكرار n_i	التكرار المتراكم N_i
5-0	10	10
10-5	8	18
15-10	12	30
20-15	11	41
25-20	9	50
المجموع	50	

من عمود التكرار المتجمع الصاعد نلاحظ أن $\left(18<\frac{n}{2}<30\right)$ إذن الفئة الوسيطية هي]15-10]، ومن الصيغة الوسيط نحدد المقادير الاتية:

L_1 : 10 ، n_1 : 12 ، N_0 : 18 ، c : 5

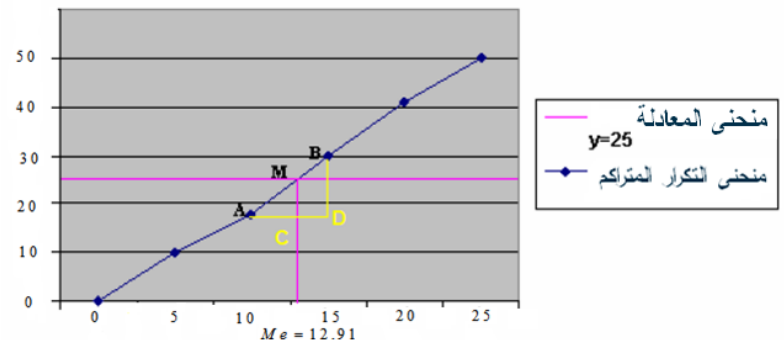
فقيمة الوسيط هي كما يلي: $$Me = L_1 + \frac{\frac{n}{2} - N_0}{n_1} \times c = 10 + \frac{25-18}{12} \times 5 = 12.91$$

ب- الطريقة البيانية:

الفئة الوسيطية هي]15-10] ، نرسم معادلة مستقيم $y = 25$ أو $y = 0.5$ عندما نود رسم التكرار المتراكم النسبي)، لتكن M نقطة تقاطع لهذا المستقيم، فاصلة M هي قيمة الوسيط.

سوف نرى كيف تم تحديد هذه القيمة بواسطة الطريقتين المشار إليها سابقا.

الشكل (): التكرار المتجمع الصاعد



3-4-3 خصائص الوسيط [1]:

3-4-4 مزايا الوسيط:

1-لا يتأثر بالقيم المتطرفة.

2-يمكن إيجاده بالرسم من خلال التكرار المتجمع الصاعد أو النازل.

3-سهل الحساب نوعا ما.

4- يمكن حسابه للتوزيعات التكرارية التي تختلف فيها أطوال الفئات دون حاجة إلى تعديل هذه الأطوال.

3-4-5 عيوب الوسيط:

لا يدخل في حسابه جميع القيم إذا يعتمد على قيمة واحدة أو قيمتين في المجموعة كلها طبقا لعدد البيانات إذا كانت فردية أو زوجية.

3-5 المنوال Le Mode [2]

3-5-1 تحديد قيمة المنوال في حالة متغير متقطع

تعريف: ليكن X متغير متقطع، نسمي منوال القيمة الأكثر شيوعا أي القيمة التي تقابل أكبر تكرار.

ملاحظة: قد يكون للسلسلة الاحصائية أكثر من منوالين تبعا لنوع الظاهرة المدروسة.

مثال: حدد المنوال لهذه السلسلة: $7,6,5,5,5,5,4,4,4,3,3,3,2,2,1,1,1$

حل المثال: المنوال لهذه السلسلة هو 5 لأنه تكرر أربع مرات.

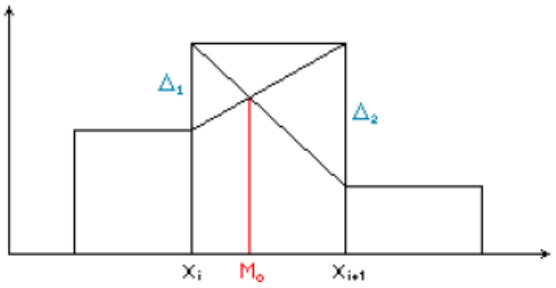
3-5-2 تحديد قيمة المنوال في حالة متغير مستمر

[1] - محمد بداوي، مرجع سابق، ص 51.

[2] - نفس المرجع، ص 52.

تعريف: ليكن X متغير مستمر، نسمي الفئة المنوالية **classe modale** الفئة التي لها أكبر تكرار. لحساب المنوال نستخدم تقنية المرجح Barycentre، بحيث عند تقدير هذه القيمة نبحث عن نسبة الانحرافات بين حدود الفئة المنوالية يجب أن تكون متساوية أي نجد الفرق بين تكرار الفئة المنوالية والتكرارين المجاورين لهذه الفئة ، أي نحسب المرجح.

بيانيا طريقة الأقطار diagonales تعطي لنا قيمة المنوال والشكل() يوضح ذلك.



الفئة المنوالية هي $[x_i , x_{i+1}[$، لحساب المنوال M_0 نلاحظ:

$$M_0 = x_i + \frac{\Delta_1}{\Delta_1 + \Delta_2} L .$$

حيث: x_i الحد الأدنى للفئة المنوالية

Δ_1 : الفرق بين التكرار الفئة المنوالية والتكرار السابق لها

Δ_2 : الفرق بين التكرار اللاحق وتكرار الفئة المنوالية

L : طول الفئة المنوالية.

مثال: نستخدم بيانات المثال السابق

الفئات	5 -0	10-5	15-10	20-15	25-20	المجموع
التكرار n_i	10	8	12	11	9	50

المطلوب: إيجاد قيمة المنوال؟

حل المثال:

الفئة المنوالية:]15 , 10]

$$x_i = 10 \ , \ L = 5 \ , \ \Delta_1 = 12 - 8 = 4 \ , \ \Delta_2 = 12 - 11 = 1$$

$$M_0 = x_i + \frac{\Delta_1}{\Delta_1 + \Delta_2} L = 10 + \frac{4}{4+1} \times 5 = 14$$

3-5-3 خصائص المنوال[1] :

1-يتأثر المنوال بالتكرار نفسه عندما يبلغ نهايته العظمى لذا يعتبر أكثر ثباتا واستقرارا من المتوسط الحسابي أو الوسيط

2-لا يتأثر المنوال بالقيم المتطرفة ولا بالقيم الوسطى في التوزيع التكراري.

3-يتأثر المنوال بعدد الفئات ومداها.

4-يمكن تعدد المنوال في جدول واحد.

3-5-4 مزايا المنوال:

1-سهل إيجاده حسابيا أو بيانيا.

2-لا يتأثر بالقيم المتطرفة.

3-تتوقف قيمته على أهمية المفردة من حيث الحجم أو الشيوع.

4-يعتبر من المقاييس الجيدة كأحد مقاييس النزعة المركزية خصوصا إذا كان التوزيع متماثلا.

3-5-5 عيوب المنوال:

1-يتم حسابه بطريقة تقريبية خاصة في التوزيعات التكرارية. إذا كان التوزيع به التواء فإن قيمة المنوال تبدو بعيدة عن مركز التوزيع أي بعيدة عن وسطه ويفقد المنوال جودته كأحد مقاييس النزعة المركزية.

[1] - نفس المرجع، ص: 54.

2-عادة يصعب تقديره إذا زادت عدد المفردات زيادة كبيرة وتساوت التكرارات الكبيرة في فئات متلاحقة.

3-لا يحسن استخدامه إذا كانت التوزيعات التكرارية مفتوحة لأن الحكم على كبر أي تكرار أو صغره يستلزم طول الفئة.

4-تتأثر قيمته بطريقة اختيار فئات التوزيع التكراري ومن ثم يمكن أن تختلف قيمته بين باحث وآخر تبعا لطول الفئة.

5-لا يصبح استخدامه ممثل للقيم في التوزيعات التكرارية حادة الالتواء[1].

[1] - نفس المرجع، ص: 55.

الفصل الرابع مقاييس التشتت

Les mesures de dispersion

4-1 مفهومها:

هي المقاييس التي تحدد قيمة انحراف البيانات عن القيمة الوسيطية[1].

4-2 المدى: L'étendue

تعريف: نسمي مدى لسلسلة إحصائية الفرق بين أكبر قيمة وأصغر قيمة أي:

$$R = x_r - x_1$$

ويعتبر المدى من أبسط مقاييس التشتت وأنه يتأثر بالقيم المتطرفة.

4-3 الربيعيات والعشريات والمؤينات:

4-3-1 الربيعيات:

الربيعيات (Quartiles) هي مقاييس إحصائية للموقع تستخدم لتقسيم مجموعة البيانات المرتّبة إلى أربعة أقسام متساوية.

4-3-1-1 مفهوم الربيعيات:

تقسم الربيعيات البيانات إلى أربع أرباع متساوية، حيث يضم كل ربع 25%من البيانات.

- الربيع الأول: هو القيمة التي تسبقها 25% من البيانات المرتبة، ويُطلق عليه أيضا "الربع الأدنى"
- الربيع الثاني: هو القيمة التي تسبقها 50% من البيانات المرتبة، ويُطلق عليه أيضا "الوسيط".
- الربيع الثالث: هو القيمة التي تسبقها 75% من البيانات المرتبة، ويُطلق عليه أيضا "الربع الأعلى".

[1] – عزت قاسم، مرجع سابق، ص:80.

ونرمز للربيعات بـ Q_i حيث $i=1,2,3$.

4-3-2 حساب الربيعيات في حالة البيانات غير المبوبة:

لدينا طريقتين لحساب الربيعات، الطريقة الأولى نقوم بترتيب القيم تصاعديا ثم نقوم بحساب رتبة الربيع، والطريقة الثانية لحساب الربيعيات نتبع خطوات مشابهة لحساب الوسيط، ولكننا نحدد موقع القيمة التي تقسم الربع. ويتم الحساب في الطريقة الثانية كالآتي:

الخطوة الأولى: نرتب البيانات تصاعدياً.

الخطوة الثانية: نقوم بتحديد رتبة الربيع (L_k)

نعتمد القانون التالي لتحديد رتبة الربيع (k) في البيانات:

$$L_k = \frac{\mathbf{k(n+1)}}{\mathbf{4}}$$

حيث:

- L_k : يمثل موقع الربيع المطلوب.
- k : يمثل رقم الربيع المطلوب (1 للربيع الأولQ_1، 2 للربيع الثاني Q_2، 3 للربيع الثالث Q_3)
- n : عدد المشاهدات الكلي.

ملاحظة:

إذا كانت النتيجة من الصيغة $L_k = \frac{\mathbf{k(n+1)}}{\mathbf{4}}$ عدد عشري نعتمد متوسط القيمتين.

مثال:

إليك البيانات التالية بعدما تم ترتيبها تصاعديا: 10، 12، 14، 15، 17، 18، 19، 20، 22، 23، 25.

n=11

المطلوب: حساب كل من الربيع الأول، الربيع الثاني، الربيع الثالث.

1- حساب الربيع الأول (Q_1): (k=1)

$$L_1= \frac{1(11+1)}{4} =\frac{12}{4} = 3$$

- النتيجة: Q_1 هي القيمة الثالثة في الترتيب وتساوي 14. ومعنى النتيجة أن 25% من البيانات هي أقل من 14 و75% من البيانات أكثر من 14.

2- حساب الربيع الثاني (Q_2): (k=2)

$$L_2= \frac{2(11+1)}{4} =\frac{24}{4} = 6$$

- النتيجة: Q_2 هي القيمة السادسة في الترتيب وتساوي 18. وهي قيمة الوسيط كما أشرنا أعلاه في المفهوم.

3- حساب الربيع الثالث (Q_3): (k=3)

$$L_3= \frac{3(11+1)}{4} =\frac{36}{4} = 9$$

- النتيجة: Q_3 هي القيمة التاسعة في الترتيب وتساوي 22. ومعنى النتيجة أن 75% من البيانات هي أقل من 22 و25% من البيانات أكثر من 22.

4-3-3 حساب الربيعيات في حالة البيانات المبوبة:

في حالة ما إذا كان المتغير الاحصائي المدروس كمي منقطع نتبع الطريقة التالية لحساب قيمة الربيع.

أولا: نقوم بحساب التكرار المتجمع الصاعد.

ثانيا: نقوم بتحديد رتبة الوسيط بالاعتماد على الصيغة التالية: $\frac{k.n}{4}$

ثالثا: نبحث عن القيمة المساوية للرتبة المطلوبة في عمود التكرار المتجمع الصاعد أو القيمة الأعلى منها مباشرة. والقيمة المقابلة لقيمة التكرار المتجمع الصاعد المحددة هي قيمة الربيع.

مثال:

البيانات التالية تمثل عدد الغرف في 30 منزل.

عدد الغرف	n_i	N↑
2	10	10
3	6	16
4	9	25
5	5	30
∑	30	/

المصدر: من إعداد الباحث اعتمادا على بيانات افتراضية.

المطلوب: المطلوب: حساب كل من الربيع الأول، الربيع الثاني، الربيع الثالث.

الحل:

أولا: حساب الربيع الأول: Q_1

- رتبة الربيع: $\frac{1\times30}{4} = 7.5$
- من الجدول نحدد قيمة الربيع الأول: القيمة المقابلة للتكرار المتجمع الصاعد الذي يساوي 7.5 أو أكبر منه مباشرة، وهي 10. إذن **$Q_1=2$**.

ثانيا: حساب الربيع الثاني: Q_2

- رتبة الربيع: $\frac{2\times30}{4} = 15$
- من الجدول نحدد قيمة الربيع الثاني: القيمة المقابلة للتكرار المتجمع الصاعد الذي يساوي 15 أو أكبر منه مباشرة، وهي 16. إذن **$Q_2=3$**.

- **ثالثا: حساب الربيع الثالث: Q_3**
- رتبة الربيع: $\frac{3\times30}{4} = 22.5$
- من الجدول نحدد قيمة الربيع الثاني: القيمة المقابلة للتكرار المتجمع الصاعد الذي يساوي 22.5 أو أكبر منه مباشرة، وهي 25. إذن **Q_3=4**.

في حالة البيانات المبوبة نستخدم الصيغة التالية:

$$Q_k = l_k + \left(\frac{\frac{k.N}{4} - n\uparrow - 1}{n}\right) \times c$$

$$Q_k = l_k + \left(\frac{\frac{k.N}{4} - n\uparrow - 1}{n_i}\right) \times c$$

حيث:

l_k : الحد الأدنى للفئة الربيعية.

N : مجموع التكرارات.

$n\uparrow -1$: التكرار المتجمع للفئة قبل الربيعية.

n_i: التكرار البسيط للفئة الربيعية.

c: طول الفئة.

مثال:

الجدول التالي يمثل بيانات الدخل الشهري (بالآلاف) لـ 110 أسرة.

الفئات	n_i	N↑
60-50	10	10
70-60	25	35
80-70	45	80
90-80	30	110
∑	110	/

المصدر: من إعداد الباحث اعتمادا على بيانات افتراضية.

المطلوب: المطلوب: حساب كل من الربيع الأول، الربيع الثاني، الربيع الثالث.

الحل:

أولا: حساب الربيع الأول: Q_1

- رتبة الربيع: $\frac{1\times110}{4}$ =27.5
- تحديد الفئة الربيعية: الفئة المقابلة للتكرار المتجمع الصاعد الذي يساوي 27.5 أو أكبر منه مباشرة، وهي 35. إذن الفئة الربيعية هي[60−70].

إذن بعد اكتمال معطيات الصيغة نقوم بالتطبيق كالآتي:

$$Q_1=60+\left(\frac{27.5-10}{25}\right)\times10$$

$$Q_1=60+\left(\frac{17.5}{25}\right)\times10$$

$$Q_1=60+7$$

$$Q_1=67$$

النتيجة: الربيع الأول يساوي 67 ألف.

التعليق: 25%من الأسر دخلها أقل من 67 ألف.

ثانيا: حساب الربيع الثاني: Q_2

- رتبة الربيع: $\frac{2\times110}{4}$ =55
- تحديد الفئة الربيعية: الفئة المقابلة للتكرار المتجمع الصاعد الذي يساوي 55 أو أكبر منه مباشرة، وهي 80. إذن الفئة الربيعية هي[70−80].

إذن بعد اكتمال معطيات الصيغة نقوم بالتطبيق كالآتي:

$$Q_2=70+\left(\frac{55-35}{45}\right)\times10$$

$$Q_2=70+\left(\frac{20}{45}\right)\times10$$

$$Q_2=70+4.44$$

$$Q_2=74.44$$

النتيجة: الربيع الثاني يساوي 74.44 ألف.

التعليق: 50% من الأسر دخلها أقل من 74.44 ألف (الوسيط).

ثالثا: حساب الربيع الثالث: Q_3

- رتبة الربيع: $\frac{3\times110}{4}$ =82.5
- تحديد الفئة الربيعية: الفئة المقابلة للتكرار المتجمع الصاعد الذي يساوي 82.5 أو أكبر منه مباشرة، وهي 110. إذن الفئة الربيعية هي[80−90].

إذن بعد اكتمال معطيات الصيغة نقوم بالتطبيق كالآتي:

$$Q_3=80+\left(\frac{82.5-80}{30}\right)\times10$$

$$Q_3=80+\left(\frac{2.5}{30}\right)\times 10$$

$$Q_3=80+\ 0.83$$

$$Q_3=80.83$$

النتيجة: الربيع الثالث يساوي 80.83ألف.

التعليق: 75% من الأسر دخلها أقل من 80.83 ألف.

4-4 العشريات:

العشريات وهي 9 قيم تقسم البيانات إلى 10 أجزاء متساوية، كل جزء يسمى العشير، ويمثل 10% من إجمالي المشاهدات.

- العشير الأول: هو القيمة التي تسبقها 10% من البيانات المرتبة.
- العشير الخامس: هو القيمة التي تسبقها 50% من البيانات المرتبة، ويُطلق عليه أيضا "الوسيط".
- العشير التاسع: هو القيمة التي تسبقها 90% من البيانات المرتبة.

ونرمز للعشريات بـ D_i حيث $i=1,2,3,4,5,6,7,8,9$.

1-4-4 حساب العشريات في حالة البيانات غير المبوبة:

نقوم بترتيب القيم تصاعديا ثم نقوم بحساب رتبة العشير، ويتم الحساب بالطريقة التالية:

الخطوة الأولى: نرتب البيانات تصاعدياً.

الخطوة الثانية: نقوم بتحديد رتبة العشير (R_k)

نعتمد القانون التالي لتحديد رتبة العشير (k) في البيانات:

$$R_k = \frac{\mathbf{k}(\boldsymbol{n}+\mathbf{1})}{\mathbf{10}}$$

حيث:

- R_k : يمثل رتبة العشير المطلوب.
- k : رقم العشير (من 1 إلى 9)
- n : عدد المشاهدات.
- **ملاحظة:**
- إذا كانت النتيجة من الصيغة $R_k = \frac{k(n+1)}{10}$ عدد عشري نعتمد متوسط القيمتين.

مثال:

لنستخدم البيانات التالية المرتبة تصاعديا: 2،5،6،9،11،13،15

المطلوب: أحسب العشير السادس.

الحل:

- تحديد رتبة العشير السادس.

$$R_k = \frac{6(7+1)}{10} = 4.8$$

بما أن رتبة العشير عدد عشري نعتمد متوسط القيمتين 9،11. ونحصل على نتيجة للعشير السادس 10. إذن $D_6=10$.

4-4-2 حساب العشريات في حالة البيانات المبوبة:

في حالة البيانات المبوبة نستخدم الصيغة التالية:

$$D_k = l_k + \left(\frac{\frac{k.N}{10} - n\uparrow - 1}{n}\right) \times c$$

$$D_k = l_k + \left(\frac{\frac{k.N}{10} - n\uparrow - 1}{n_i}\right) \times c$$

حيث:

l_k : الحد الأدنى للفئة العشيرية.

N : مجموع التكرارات.

$n \uparrow -1$: التكرار المتجمع للفئة قبل العشيرية.

n_i: التكرار البسيط للفئة العشيرية.

c: طول الفئة.

مثال:

سنعتمد على نفس المثال الذي اعتمدناه في حساب الربيعيات.

المطلوب: حساب كل من العشير الأول، العشير الخامس، العشير التاسع.

الحل:

أولا: حساب العشير الأول: D_1

- رتبة الربيع: $11= \frac{1\times110}{10}$
- تحديد الفئة العشيرية: الفئة المقابلة للتكرار المتجمع الصاعد الذي يساوي 11 أو أكبر منه مباشرة، وهي 35. إذن الفئة العشيرية هي $]70-60]$.

إذن بعد اكتمال معطيات الصيغة نقوم بالتطبيق كالآتي:

$$D_1=60+\left(\frac{11-10}{25}\right)\times10$$

$$D_1=60+\left(\frac{1}{25}\right)\times10$$

$$D_1=60+ 0.40$$

$$D_1=60.40$$

النتيجة: العشير الأول يساوي 60.40 ألف.

التعليق: 10% من الأسر دخلهم أقل من 60.40 ألف.

ثانيا: حساب العشير الخامس: D_5

- رتبة الربيع: $55=\frac{5\times110}{10}$
- تحديد الفئة العشيرية: الفئة المقابلة للتكرار المتجمع الصاعد الذي يساوي 55 أو أكبر منه مباشرة، وهي 80. إذن الفئة العشيرية هي]80-70[.

إذن بعد اكتمال معطيات الصيغة نقوم بالتطبيق كالآتي:

$$D_5=70+\left(\frac{55-35}{45}\right)\times10$$

$$D_5=70+\left(\frac{20}{45}\right)\times10$$

$$D_5=70+4.44$$

$$D_5=74.44$$

النتيجة: العشير الخامس يساوي 74.44 ألف.

التعليق: نصف الأسر دخلهم أقل من 74.44 ألف ونصفهم أكثر، وهو الربيع الثاني والوسيط في نفس الوقت.

ثالثا: حساب العشير التاسع: D_9

- رتبة الربيع: $99=\frac{9\times110}{10}$
- تحديد الفئة العشيرية: الفئة المقابلة للتكرار المتجمع الصاعد الذي يساوي 99 أو أكبر منه مباشرة، وهي 110. إذن الفئة العشيرية هي]90-80[.

إذن بعد اكتمال معطيات الصيغة نقوم بالتطبيق كالآتي:

$$D_9=80+\left(\frac{99-80}{30}\right)\times10$$

$$D_9=80+\left(\frac{19}{30}\right)\times10$$

$$D_9=80+6.33$$

$$D_9=86.33$$

النتيجة: العشير التاسع يساوي 74.44 ألف.

التعليق: 90% من الأسر دخلهم أقل من 86.33 ألف.

4-5 المؤينات:

المؤينات هي 99 قيمة تقسم البيانات إلى 100 قسم بشكل متساو بعد ترتيبها تصاعدياً، بحيث يمثل كل مئين نسبة مئوية محددة من البيانات. ويرمز للمؤينات بـ P_i، حيث i=1,2,3,4,5,6,7,8,9,10,11…100

- المؤين الخامس العشرون: هو القيمة التي تسبقها 25% من البيانات المرتبة، وهو الربيع الأول.
- المؤين الخمسون: هو القيمة التي تسبقها 50% من البيانات المرتبة، وهو الوسيط وهو العشير الخامس.
- المؤين الخامس والسبعون: هو القيمة التي تسبقها 75% من البيانات المرتبة، وهو الربيع الثالث.

4-5-1 حساب المؤينات في حالة البيانات غير المبوبة:

نقوم بترتيب القيم تصاعديا ثم نقوم بحساب رتبة المؤين، ويتم الحساب بالطريقة التالية:

الخطوة الأولى: نرتب البيانات تصاعدياً.

الخطوة الثانية: نقوم بتحديد رتبة المؤين(R_k)

نعتمد القانون التالي لتحديد رتبة المؤين (k) في البيانات:

$$R_k = \frac{\mathbf{k}(n+1)}{\mathbf{100}}$$

حيث:

- R_k : يمثل رتبة المؤين المطلوب.
- k : رقم المؤين (من 1 إلى 99)
- n : عدد المشاهدات.

مثال:

لنستخدم البيانات التالية المرتبة تصاعديا: 2،3،4،5،5،6،7،8،9

المطلوب: أحسب المؤين العشرون.

الحل:

- تحديد رتبة المؤين العشرون.

$$\mathbf{R_{20}=\frac{20(9+1)}{100}= 2}$$

- رتبة المؤين هي 2، وعليه فإن المؤين العشرين هو القيمة الثانية في السلسلة المرتبة تصاعديا، إذن $P_{20}=3$
- إذا كانت النتيجة من الصيغة $R_k = \frac{k(n+1)}{100}$ عدد عشري نعتمد متوسط القيمتين.

4-5-2 حساب المؤينات في حالة البيانات المبوبة:

في هذه الحالة بما انه تتوفر لدينا بينات كمية متصلة فإن الصيغة تكون كالآتي:

$$P_k = l_k + \left(\frac{\frac{k.N}{100} - n\uparrow - 1}{n}\right) \times c$$

$$P_k = l_k + \left(\frac{\frac{k.N}{100} - n\uparrow - 1}{n_i}\right) \times c$$

حيث:

l_k : الحد الأدنى للفئة المؤينية.

N : مجموع التكرارات.

$n \uparrow -1$: التكرار المتجمع للفئة قبل المؤينية.

n_i: التكرار البسيط للفئة المؤينية.

c: طول الفئة.

سنعتمد على نفس المثال الذي اعتمدناه في حساب الربيعيات والعشريات.

المطلوب: حساب كل من المؤين العاشر، والمؤين الخامس والثمانون.

الحل:

أولا: حساب المؤين العاشر: P_{10}

- رتبة المؤين: $\frac{10\times110}{100}=11$
- تحديد الفئة المؤينية: الفئة المقابلة للتكرار المتجمع الصاعد الذي يساوي 11 أو أكبر منه مباشرة، وهي 35. إذن الفئة المؤينية هي $[60-70[$.

إذن بعد اكتمال معطيات الصيغة نقوم بالتطبيق كالآتي:

$$P_{10}=60+\left(\frac{11-10}{25}\right)\times10$$

$$P_{10}=60+\left(\frac{1}{25}\right)\times10$$

$$P_{10}=60+0.40$$

$$P_{10}=60.40$$

النتيجة: المؤين العاشر يساوي 60.40 ألف.

التعليق: 10% من الأسر دخلهم أقل أو يساوي 60.40 ألف.

ثانيا: حساب المؤين الخامس والثمانون: P_{85}

- رتبة المؤين: $\frac{85\times110}{100}=93.50$

- تحديد الفئة المؤينية: الفئة المقابلة للتكرار المتجمع الصاعد الذي يساوي 93.50 أو أكبر منه مباشرة، وهي 110. إذن الفئة المؤينية هي[80-90].

إذن بعد اكتمال معطيات الصيغة نقوم بالتطبيق كالآتي:

$$\mathbf{P_{85}}=80+\left(\frac{93.50-80}{30}\right)\times 10$$

$$\mathbf{P_{10}}=80+\left(\frac{13.50}{30}\right)\times 10$$

$$\mathbf{P_{85}}=80+4.50$$

$$\mathbf{P_{85}}=84.50$$

النتيجة: المؤين الخامس والثمانون يساوي 84.50ألف.

التعليق: 85% من الأسر دخلهم أقل أو يساوي 84.50 ألف.

4-6 التباين La variance

تعريف: هو مجموع مربعات انحرافات القيم عن متوسطها الحسابي مقسمة على عدد المشاهدات، ويحسب كما يلي[1]: $\sigma^2=\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2$.

ملاحظة: يعتبر التباين من أفضل مقاييس التشتت وهو الخاصية التي بنيت لكي تأخذ تشتت كل أفراد المجتمع.

مبرهنة: يمكننا أيضا كتابة التباين وفق الصيغة الاتية: $\sigma^2=\frac{1}{n}\sum_{i=1}^{n}x_i^2-\overline{x}^2$.

ملاحظة: عندما نريد إجراء عملية التقدير(estimation) للتباين لمتغير X انطلاقا من عينة عشوائية ذات حجم n، نستخدم صيغة التباين المصحح والمقسم على$n-1$(لأن تباين المجتمع يعتبر مقدر متحيز)، حيث تكون الصيغة المصححة كما يلي:

[1] - Timothy C. Urdan , Statistics in Plain English , Lwrence Erlbaum , association publications , London ,2005 , p :16.

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

علما أن أغلب البرامج الاحصائية تستخدم صيغة التباين المصحح.

4-7 الانحراف المعياري L'écart type

تعريف: هو الجذر التربيعي للتباين، ويحسب كما يلي: $\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$ ، أما صيغة الانحراف المصحح فتكون كما يلي[1]: $S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$

مثال: أحسب التباين للسلسلة الاحصائية الاتية: $2,3,4,5,6$

حل المثال: نستخدم الصيغة الاتية: $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}x_i^2 - \bar{x}^2$.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n}x_i = 4$$

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}x_i^2 - \bar{x}^2 = \frac{1}{5}\left(2^2+3^2+4^2+5^2+6^2\right) - 4^2 = 2$$

4-8 الخطأ المعياري:

الخطأ المعياري يقيس مدى دقة الوسط الحسابي للعينة في تمثيل الوسط الحسابي للمجتمع ككل. أي أن الخطأ المعياري هو مقياس إحصائي يُستخدم لتقدير مدى تباين أو تقلب القيم الإحصائية (مثل المتوسط) الناتجة عن العينة مقارنةً بالقيمة الحقيقية في مجتمع الدراسة، ولهذا يطلق عليه الخطأ المعياري للوسط الحسابي.

وصيغة القانون لحسابه هي بالشكل التالي:

$$\mathbf{SE} = \frac{\boldsymbol{s}}{\sqrt{\boldsymbol{n}}}$$

[1] -ibid , p :16.

مثال:

قمنا بتطبيق القانون على المثال رقم () وحصلنا النتائج الآتية:

$$SE=\frac{9.23}{\sqrt{110}}$$

$$SE=\frac{9.23}{10.48}\approx 0.88$$

ومعناه: أن قيمة الخطأ المعياري قليلة جدا، والذي يعني أن المتوسط الحسابي الذي تحصلنا عليه والذي يساوي 73.64 قريب جدا من المتوسط الحقيقي لدخل الأسر في المجتمع الأصلي.

الفصل الخامس: مقاييس العلاقة بين مقاييس النزعة المركزية والتشتت

5-1 معامل الاختلاف

يعتبر معامل الاختلاف من أهم المقاييس الإحصائية، لأنه مقياس تشتت نسبي (بنسبة مئوية)، مما يسمح لنا بمقارنة تشتت مجموعتين مختلفتين في وحدات القياس أو متباعدتين في الوسائط الحسابية.

وصيغة القانون لحسابه هي بالشكل التالي:

$$cv = \left(\frac{s}{\bar{x}}\right)\times 100$$

مثال:

قمنا بتطبيق القانون على المثال رقم () وحصلنا النتائج الآتية:

$$cv = \left(\frac{9.23}{73.64}\right)\times 100$$

$$cv = 12.53\%$$

ومعنى النتيجة والتي تعتبر منخفضة تبين أن توزيع الدخل بين الأسر يظهر عن تجانس كبير، والذي يعني أن دخل الأسر يقترب من بعضه البعض ولا يظهر تشتتا كبيرا حول المتوسط.

الفصل السادس: مقاييس الشكل

Les mesures de forme

6-1 مفهومها:

هي المقاييس التي تهتم بدراسة شكل توزيع البيانات، نجد معاملين رئيسين وهما الالتواء و التفرطح.

6-2 معامل الالتواء coefficient d'asymétrie

مفهومه: الالتواء هو من مقاييس الشكل مرتبط بتماثل البيانات حول قيمة مركزية في التوزيع الاحصائي، فكلما كان الشكل أكثر تماثلا فأن التوزيع يوصف بأن له توزيع طبيعي (جرسي)، وعندما يكون غير ذلك (التواء من جهة اليمين أو اليسار) فأن البيانات توصف بأن ليس لها توزيع طبيعي[1].

6-2-1 معامل الالتواء لفيشر coefficient d'asymétrie de Fisher

نعرف العزم المركزي من الرتبة الثالثة بـ : $m_3 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3$

فمعامل فيشر للالتواء يعطى كما يلي: $\gamma_1 = \frac{m_3}{\sigma^3}$

- إذا كان معامل الالتواء γ_1 موجبا ، فأن التوزيع يكون ملتويا من جهة اليمين.
- إذا كان معامل الالتواء γ_1 سالبا ، فأن التوزيع يكون ملتويا من جهة اليسار.
- إذا كان معامل الالتواء γ_1 معدوم ، فأن التوزيع يكون متماثلا.

[1] – فتحي عاروري ، المعاينة الإحصائية طرقها واستخداماتها ، دار الأكاديميون ، عمان ، ص: 28 .

6-2-2- معامل الالتواء ليول coefficient d'asymétrie de Yule

يرتكز معامل الالتواء ليول على تمركز الربيعات الثلاثة: الربيع الأول والثالث والوسيط ونظيم مسافة المدى الربيعي، ويعطى كما يلي: $A_Y = \frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1}$

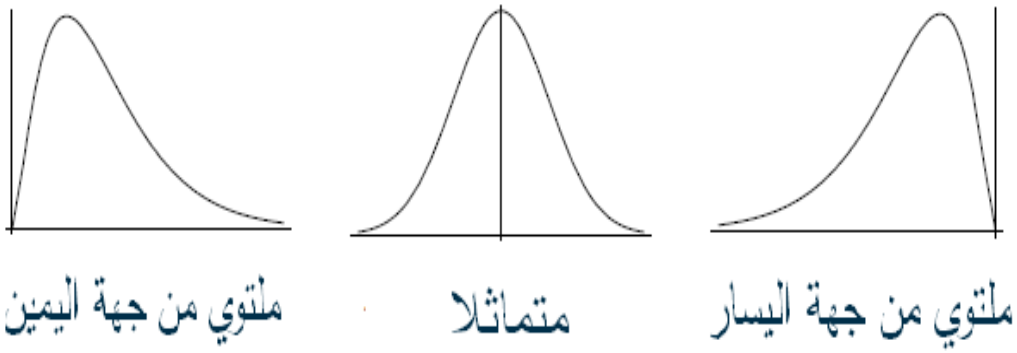
6-2-3- معامل الالتواء لبيرسون coefficient d'asymétrie de pearson

يرتكز معامل الالتواء لبيرسون على المقارنة بين المتوسط والمنوال من جهة والمتوسط والوسيط من جهة أخرى، مع معايرة الانحراف المعياري، ويعطى القانونين كما يلي:

- معامل بيرسون الأول A_{P_1} : $A_{P_1} = \frac{\bar{x} - M_0}{\sigma}$

- معامل بيرسون الثاني A_{P_2} : $A_{P_2} = \frac{3(\bar{x} - M_e)}{\sigma}$

جميع معاملات الالتواء لها نفس الخصائص التي ذكرناها سابقا ونوضح ذلك في الشكل البياني الاتي:



6-3- معامل التفرطح coefficient d'aplatissement

مفهومه: التفرطح هو من مقاييس الشكل مرتبط بتشتت البيانات حول قيمة مركزية في التوزيع الاحصائي، وممكن أن نجد شكلا متماثلا لكن لا يعني هذا أن له توزيع طبيعي،

فكلما كان الشكل أكثر استطالة (مدبب)، فأننا نطلق على التوزيع بأن له شكل مدبب ، وإذا كان الشكل أكثر تفرطحا ، فأننا نطلق على التوزيع بأن له شكل مفرطح، وإذا كان غير الاثنين (أقل استطالة وأقل تفرطحا) فان التوزيع يوصف بأن له توزيع طبيعي(جرسي)[1].

6-3-1 معامل التفرطح لبيرسون coefficient d'aplatissement de pearson

يحسب كالاتي: $\beta_2 = \frac{m_4}{\sigma^4}$

- إذا كان $\beta_2 = 3$ ، فالتوزيع طبيعي.
- إذا كان $\beta_2 < 3$ ، فالتوزيع مفرطح.
- إذا كان $\beta_2 > 3$ ، فالتوزيع مدبب.

6-3-2 معامل التفرطح لفيشر coefficient d'aplatissement de Fisher

يحسب كالاتي: $\gamma_2 = \beta_2 - 3 = \frac{m_4}{\sigma^4} - 3$

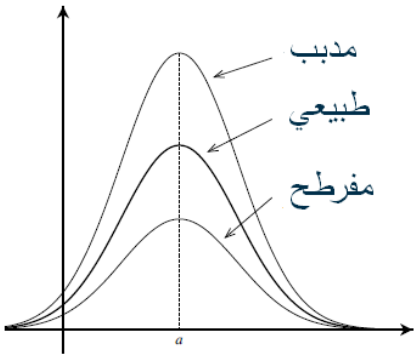
ملاحظة: معظم البرامج الاحصائية تستخدم صيغة فيشر .

- إذا كان معامل التفرطح $\gamma_2 > 0$ ، فأن التوزيع يكون مدببا.
- إذا كان معامل التفرطح $\gamma_2 < 0$ ، فأن التوزيع يكون مفرطحا.

[1] - نفس المرجع ، ص: 28.

– إذا كان معامل التفرطح $\gamma_2 = 0$ ، فأن التوزيع يكون طبيعيا.

ونوضح حالات التفرطح في الشكل البياني الاتي:



تمارين محلولة

تمرين1:

يمثل الجدول التالي توزيع قياسات لسرعة 100 مركبة بـ (كلم/ سا) عبر طريق سريع ، وكانت النتائج مبينة في الجدول التوزيع التكراري الاتي:

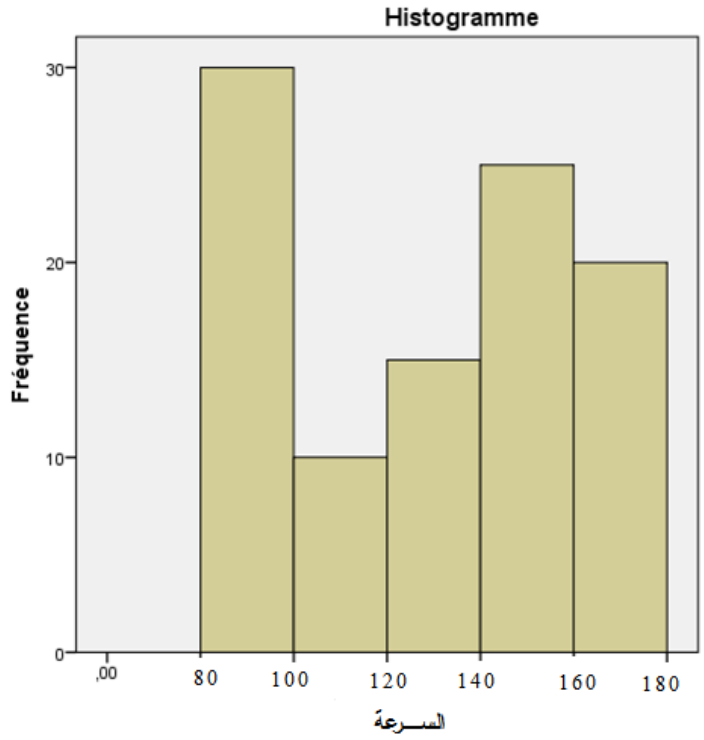
السرعة	التكرار	التكرار النسبي	التكرار المتراكم
80-100	30	30,0	30,0
100-120	10	10,0	10,0
120-140	15	15,0	15,0
140-160	25	25,0	25,0
160-180	20	20,0	20,0
Total	100	100,0	100,0

المطلوب: قدم عرض بياني في شكل مدرج تكراري وفي شكل دائرة نسبية؟

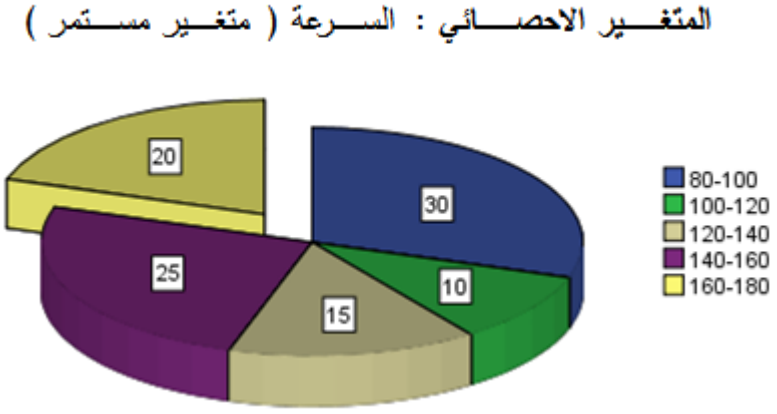
حل التمرين1:

1- المدرج التكراري

المتغير الاحصائي : السرعة (متغير مستمر)



1- الدائرة النسبية:



تمرين2:

تم احصاء عدد الأفراد داخل الأسرة النووية في ناحية من مدينة ما (الاحصائيات شملت 1000 أسرة)، وأظهرت النتائج ما يلي:

عدد الأفراد	1	2	3	4	5	6	7	8	9
عدد الأسر	2	4	15	116	446	56	278	62	21

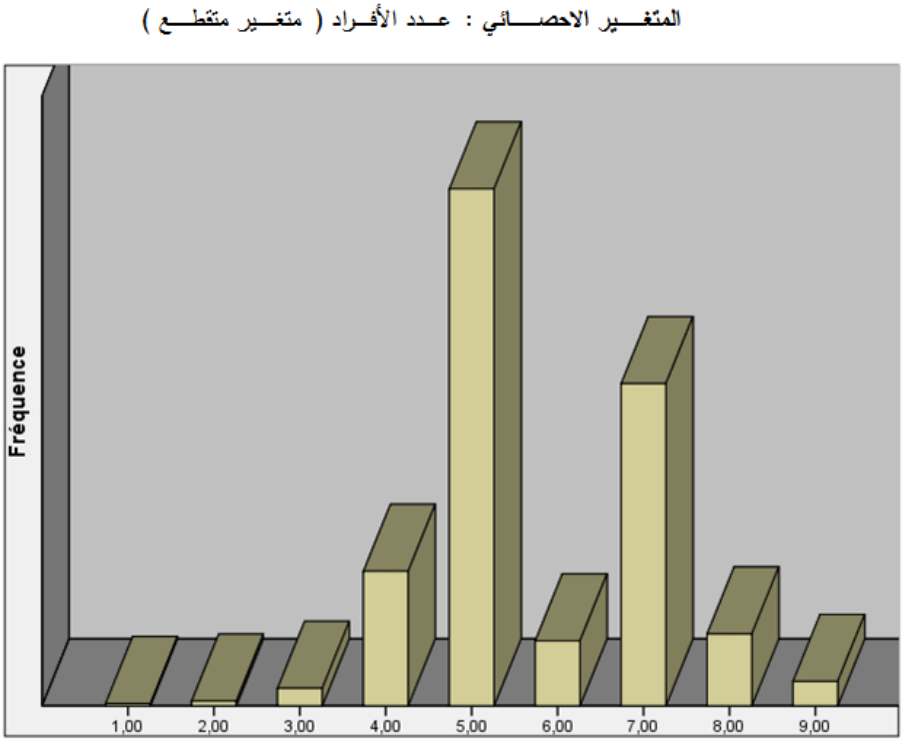
المطلوب: تمثيل البيانات على شكل أعمدة بيانية بسيطة؟

حل التمرين 2:

المتغير الاحصائي : عدد الأفراد (متغير متقطع)

عدد الأفراد	التكرار n_i	التكرار النسبي f_i (%)	ت. النسبي المتراكم F_i (%)
1,00	2	,2	,2
2,00	4	,4	,6
3,00	15	1,5	2,1
4,00	116	11,6	13,7
5,00	446	44,6	58,3
6,00	56	5,6	63,9
7,00	278	27,8	91,7
8,00	62	6,2	97,9
9,00	21	2,1	100,0
المجموع	1000	100,0	

الشكل البياني: الأعمدة البيانية



تمرين3:

يقدم الجدول التالي عدد الوفيات خلال سنة في بلد ما.

العمر	0	40	50	60	70	80	90	110
الوفيات التراكمية	0	438	1312	3563	8716	10284	11848	12106

المطلوب:

1- ما هو المتغير الاحصائي؟

2- قدم شكل بياني على شكل مدرج تكراري؟

حل التمرين3:

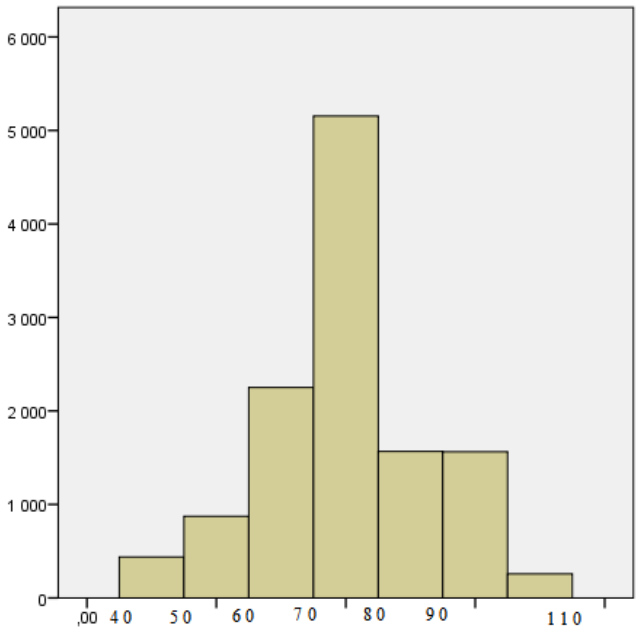
1- يمثل المتغير الاحصائي مدة الحياة وهو متغير مستمر.

2- المدرج التكراري

أولا يجب تقديم جدول توزيع تكراري

الفئات (العمر)	]0 – 40]	]40 – 50]	]50 – 60]	]60 – 70]	]70 – 80]	]80 – 90]	]90 – 110]
التكرار n_i	438	874	2251	5153	1568	1564	258

ثانيا: المدرج التكراري



تمرين4:

اليك السلسلة الاحصائية التالية:

x_i	22	23	24	25	30	31	32	33	34	35	36	37	38	39
n_i	1	2	3	3	5	2	3	3	5	4	6	7	6	5

المطلوب: – حساب كل مقاييس التمركز والتشتت؟

حل التمرين 4:

أولا: نحسب التكرار المتراكم

39	38	37	36	35	34	33	32	31	30	25	24	23	22	x_i
5	6	7	6	4	5	3	3	2	5	3	3	2	1	n_i
55	50	44	37	31	27	22	19	16	14	9	6	3	1	N_i

ثانيا: مقاييس التمركز

1- المتوسط الحسابي:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} n_i \cdot x_i = \frac{1830}{55} = 33.27$$

2- الوسيط:

نلاحظ أن: $55 \times 0.50 = 27.5$ إذن: $M_e = 35$

3- المنوال:

نلاحظ أن القيمة لها أكبر تكرار إذن: $M_0 = 37$

ثالثا: مقاييس التشتت

1- المدى

$$R = x_r - x_1 = 39 - 22 = 17$$

2- الانحراف الربيعي:

نلاحظ أن: $55 \times 0.25 = 13.75$ إذن: $Q_1 = 30$.

$55 \times 0.75 = 41.25$ إذن: $Q_3 = 37$.

أما قيمة الانحراف الربيعي فهي:

$$[\ Q_1\ ;\ Q_3\ ] = [\ 30\ ;\ 37\ ] \Rightarrow ET = Q_3 - Q_1 = 37 - 30 = 7$$

3- التباين والانحراف المعياري:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = 24.53 \ , \quad S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} = 4.95$$

4- معامل الاختلاف:

$$\frac{S}{\bar{x}} \times 100 = \frac{4.95}{33.27} \times 100 = 14.87\%$$

تمرين5:

نسمي متغير الوزن (القيم مدورة الى أقرب عدد) لـ 50 تلميذ في ابتدائية ' س':

30	25	20	22	32	23	36	25	24	23
36	36	36	25	25	25	25	20	32	32
36	36	36	30	30	30	30	20	32	32
32	32	32	32	25	25	23	23	23	23
30	25	20	22	32	23	36	25	24	23

المطلوب: – ايجاد توزيع تكراري؟

– حساب كل مقاييس التمركز والتشتت والشكل؟

حل التمرين5:

أولا: نحسب جدول التوزيع التكراري:

– حساب طول الفئة انطلاقا من القانون الاتي:

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

المدى: أكبر قيمة – أصغر قيمة $R = x_r - x_1$ **L'étendue :**

$$R = x_{\max} - x_{\min} = 36 - 20 = 16$$

K عدد الفئات تحسب من صيغة ستيرجس formule de sturges

$$K = 1 + 3.22Log(n) = 1 + 3.22Log(50) = 6.47$$

ويمكن إيجادها من صيغة يول formule de Yule

$$K = 2.5\sqrt[4]{n} = 2.5\sqrt[4]{50} = 6.64$$

إذن: $L = \frac{R}{K} \approx 3$

فالجدول التكراري يكون كما يلي:

الفئات	n_i	c_i	$n_i.c_i$	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$n_i(x_i - \bar{x})^2$	$(x_i - \bar{x})^3$	$n_i(x_i - \bar{x})^3$	$(x_i - \bar{x})^4$	$n_i(x_i - \bar{x})^4$	$n_i \uparrow$
20-23	6	21,5	129	-7,08	50,1264	300,7584	-354,894912	-2129,36947	2512,65598	15075,9359	6
23-26	20	24,5	490	-4,08	16,6464	332,928	-67,917312	-1358,34624	277,102633	5542,05266	26
29-32	6	30,5	183	1,92	3,6864	22,1184	7,077888	42,467328	13,589545	81,5372698	32
32-35	10	33,5	335	4,92	24,2064	242,064	119,095488	1190,95488	585,949801	5859,49801	42
35-38	8	36,5	292	7,92	62,7264	501,8112	496,793088	3974,3447	3934,60126	31476,8101	50
المجموع	50		1429		0	1399,68	0	1720,0512	0	58035,8339	

ثانيا: مقاييس التمركز

1- المتوسط الحسابي:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} n_i \cdot c_i = \frac{1429}{50} = 28.58$$

4- الوسيط:

- تحديد قيمة الوسيط: $(50 \times 0.50 = 25)$ ، إذن الفئة الوسيطية هي $[23-26[$ ، ومن صيغة الوسيط نحدد المقادير الاتية:

$L_1 : 23$ ، $n_1 : 20$ ، $N_0 : 6$ ، $c : 3$

فقيمة الوسيط هي كما يلي: $Me = L_1 + \frac{\frac{n}{2} - N_0}{n_1} \times c = 23 + \frac{25-6}{20} \times 3 = 25.85$

5- المنوال:

نلاحظ أن أكبر تكرار هو 20، إذن الفئة المنوالية هي $]23-26]$ ، ومن صيغة المنوال نحدد المقادير الاتية:

$$\Delta_1 = 20-6=14 \ ; \ \Delta_2 = 20-6=14$$
$$x_i = 23 \ , \ L=3$$

فقيمة المنوال هي كما يلي: $M_0 = x_i + \frac{\Delta_1}{\Delta_1+\Delta_2} L = 23 + \frac{14}{14+14} \times 3 = 24.5$

ثالثا: مقاييس التشتت

1- المدى

$$R = x_{\max} - x_{\min} = 36-20=16$$

2- الربيعيات:

- تحديد قيمة الربيع الأول: $(50 \times 0.25 = 12.5)$ ، إذن فئة الربيع الأول هي $]23-26]$ ، ومن صيغة الربيع الأول نحدد المقادير الاتية:

$L_1 : 23$ ، $n_{1/4} : 20$ ، $N_{01} : 6$ ، $c : 3$

فقيمة الربيع الأول هي كما يلي: $Q_1 = L_1 + \frac{\frac{n}{4} - N_{01}}{n_{1/4}} \times c = 23 + \frac{12.5-6}{20} \times 3 = 23.975$

- تحديد قيمة الربيع الثالث: $(50 \times 0.75 = 37.5)$ ، إذن فئة الربيع الثالث هي $]32-35]$ ، ومن صيغة الربيع الثالث نحدد المقادير الاتية:

$L_1 : 32$ ، $n_{3/4} : 10$ ، $N_{03} : 32$ ، $c : 3$

فقيمة الربيع الثالث هي كما يلي: $Q_3 = L_1 + \frac{\frac{3n}{4} - N_{03}}{n_{3/4}} \times c = 32 + \frac{37.5-32}{10} \times 3 = 33.65$

3- التباين والانحراف المعياري:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n} n_i \left(x_i - \bar{x}\right)^2 = \frac{1399.68}{49} = 28.56 \ ,$$

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n} n_i \left(x_i - \bar{x}\right)^2} = \sqrt{28.56} = 5.34$$

رابعا: مقاييس الشكل

- معامل الالتواء

معامل الالتواء لفيشر coefficient d'asymétrie de Fisher

نعرف العزم المركزي من الرتبة الثالثة بـ : $m_3 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^3 = \frac{1720.0512}{50} = 34.40$

نعرف العزم المركزي من الرتبة الرابعة بـ : $m_4 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^4 = \frac{58035.833}{50} = 1160.71$

فمعامل فيشر للالتواء يعطى كما يلي: $\gamma_1 = \frac{m_3}{S^3} = \frac{34.4}{(5.34)^3} = 0.226$

معامل الالتواء ليول coefficient d'asymétrie de Yule

يرتكز معامل الالتواء ليول على تمركز الربيعات الثلاثة: الربيع الأول والثالث والوسيط ونظيم مسافة المدى الربيعي، ويعطى كما يلي:

$$A_Y = \frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1} = \frac{33.65 - 2(25.85) + 23.975}{33.65 - 23.975} = 0.612$$

معامل الالتواء لبيرسون coefficient d'asymétrie de pearson

يرتكز معامل الالتواء لبيرسون على المقارنة بين المتوسط والمنوال من جهة والمتوسط والوسيط من جهة أخرى، مع معايرة الانحراف المعياري، ويعطى القانونين كما يلي:

- معامل بيرسون الأول A_{P_1} : $A_{P_1} = \frac{\bar{x} - M_0}{S} = \frac{28.58 - 24.5}{5.34} = 0.764$

- معامل بيرسون الثاني A_{P_2} : $A_{P_2} = \frac{3(\bar{x} - M_e)}{S} = \frac{3(28.58 - 25.85)}{5.34} = 1.533$

2- معامل التفرطح

معامل التفرطح لبيرسون coefficient d'aplatissement de pearson

يحسب كالاتي: $\beta_2 = \frac{m_4}{S^4} = \frac{1160.71}{(5.34)^4} = 1.427$

معامل التفرطح لفيشر coefficient d'aplatissement de Fisher

يحسب كالاتي: $\gamma_2 = \beta_2 - 3 = \frac{m_4}{S^4} - 3 = 1.427 - 3 = -1.572$

تمرين6:

في محل لبيع الأجهزة الكهرو منزلية سجل خلال يوم بيع 100 جهاز، حيث تراوحت الأسعار ما بين 200 و1200 وحدة نقدية، حيث كانت المعطيات ممثلة في الجدول التكراري الاتي:

الأسعار(و.نقدية)	400-200	600-400	800-600	1000-800	1200-1000
الوحدات المباعة	15	25	35	10	15

المطلوب:

- إيجاد قيمة الوسيط حسابيا وبيانيا ؟
- إيجاد قيمة المنوال حسابيا وبيانيا ؟
- إيجاد الربيعات وقدمهما في علبة whiskers (Boite à moustaches)، ثم أحسب الانحراف الربيعي؟

حل التمرين6:

إيجاد قيمة الوسيط:

1- الطريقة الجبرية:

الفئات	التكرار n_i	التكرار المتراكم N_i
400-200	15	15
600-400	25	40
800-600	35	75
1000-800	10	85
1200-1000	15	100
المجموع	100	

- تحديد قيمة الوسيط : $(100\times 0.50=50)$ ، إذن الفئة الوسيطية هي $]800-600]$ ، ومن صيغة الوسيط نحدد المقادير الاتية:

L_1 : 600 ، n_1 : 35 ، N_0 : 40 ، c : 200

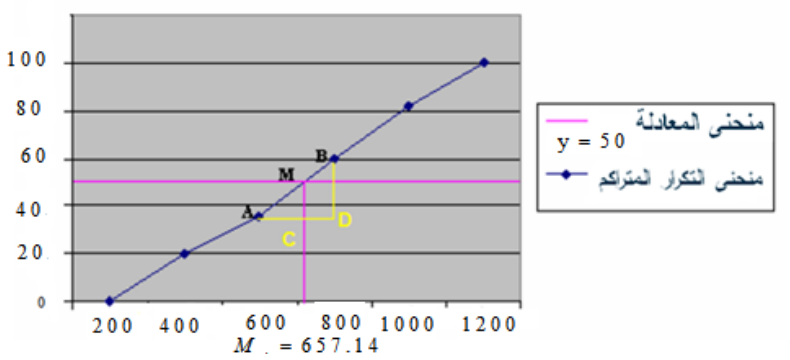
فقيمة الوسيط هي كما يلي: $$Me=L_1+\frac{\frac{n}{2}-N_0}{n_1}\times c=600+\frac{50-40}{35}\times 200=657.14$$

2- الطريقة البيانية:

الفئة الوسيطية هي $]800-600]$ ، نرسم معادلة مستقيم $y=50$ أو $y=0.5$ عندما نود رسم التكرار المتراكم النسبي)، لتكن M نقطة تقاطع لهذا المستقيم ، فاصلة M هي قيمة الوسيط.

سوف نرى كيف تم تحديد هذه القيمة بواسطة طريقتي معمل التوجيه و مبرهة طاليس.

الشكل(): التكرار المتجمع الصاعد



إيجاد قيمة المنوال:

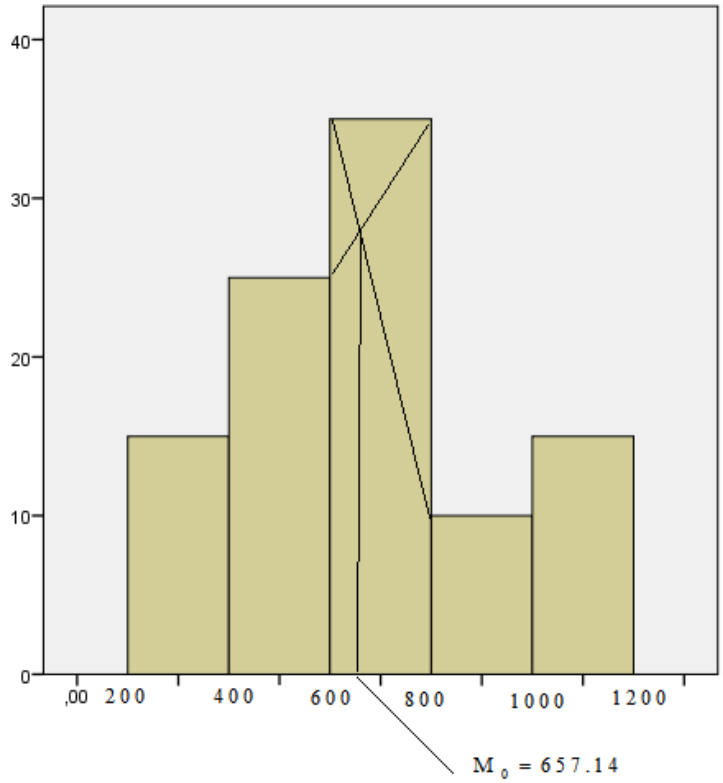
1- حسابيا:

نلاحظ أن أكبر تكرار هو 35 ، إذن الفئة المنوالية هي]800-600[، ومن صيغة المنوال نحدد المقادير الاتية:

$$\Delta_1 = 35-25 = 10 \;\; ; \;\; \Delta_2 = 35-10 = 25$$
$$x_i = 600 \;\; , \;\; L = 200$$

فقيمة المنوال هي كما يلي: $M_0 = x_i + \frac{\Delta_1}{\Delta_1 + \Delta_2} L = M_0 = 600 + \frac{10}{10+25} 200 = 657.14$

2- بيانيا:



إيجاد الربيعات:

– تحديد قيمة الربيع الأول : $(100\times 0.25=25)$ ، إذن فئة الربيع الأول هي $]600-400[$ ، ومن صيغة الربيع الأول نحدد المقادير الاتية:

L_1 : 400 ، $n_{1/4}$: 25 ، N_{01} : 15 ، c : 200

فقيمة الربيع الأول هي كما يلي: $Q_1 = L_1 + \frac{\frac{n}{4} - N_{01}}{n_{1/4}} \times c = 400 + \frac{25-15}{25} \times 200 = 480$

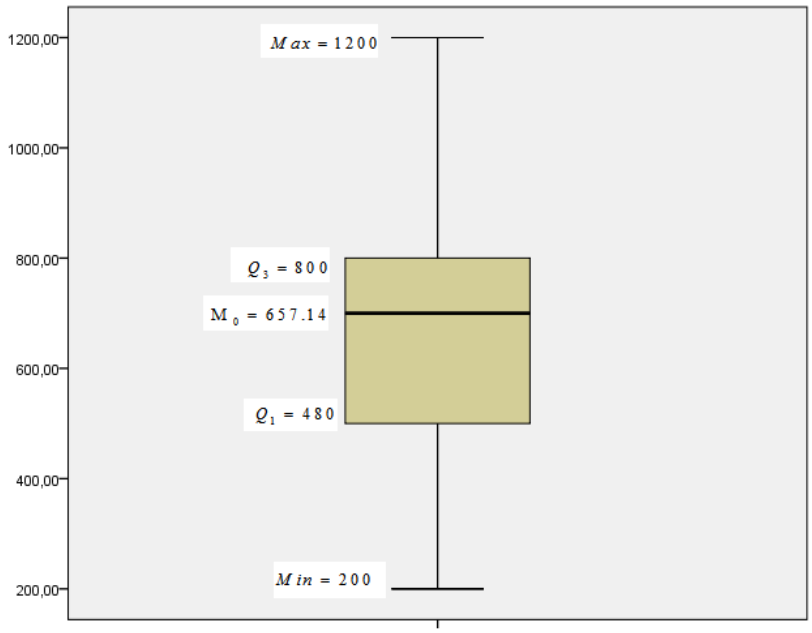
– تحديد قيمة الربيع الثالث : $(100\times 0.75=75)$ ، إذن فئة الربيع الثالث هي $]800-600[$ ، ومن صيغة الربيع الثالث نحدد المقادير الاتية:

L_1 : 600 ، $n_{3/4}$: 35 ، N_{03} : 40 ، c : 200

فقيمة الربيع الثالث هي كما يلي: $Q_3 = L_1 + \frac{\frac{3n}{4} - N_{03}}{n_{3/4}} \times c = 600 + \frac{75-40}{35} \times 200 = 800$

إذن قيمة الانحراف الربيعي هي كما يلي: $ET = Q_3 - Q_1 = 800 - 480 = 320$

تمثيل الربيعات في مخطط صندوق Box plot (Boite à moustaches) يكون كما يلي:



المراجع:

I) الكتب باللغة العربية:

1-سلاطنية بلقاسم، حسان الجيلالي، أسس المبحث العلمي، ديوان المطبوعات الجامعية، الجزائر، 2007.

2-شيبات أحمد، الاحصاء الوصفي، دار الهدى، عين مليلة – الجزائر-، 2001.

3-عمر قاسم عزات، مبادئ الاحتمالات والإحصاء، منشورات جامعة دمشق، 1995.

4-محمد بداوي، الإحصاء الوصفي، دار هومة، لجزائر، 2017

II) الكتب باللغة الانجليزية :

1- CESAR PEREZ LOPEZ, DATA STATISTICAL ANALYSIS WITH R, Publié par CESAR PEREZ, 2012
2- Gupta, Statistics, SBPD Publications, New Delhi, 2021
3- Isabel Willemse , Statistical Methods and Calculation Skills 3 rd, Cape town ,2009
4- Réjean Huot, Méthodes quantitatives pour les sciences humaines 2 édition, Les presses de l'université de laval , 2003
5- James E. Burt, Gerald M. Barber, David L. Rigby, Elementary Statistics for Geographers, Third Edition, Guilford press, New York, 2009
6- Jaap de Gruijter, Dick J. Brus, Marc F.P. Bierkens, Sampling for Natural Resource Monitoring, springer, New York
7- Steven K. Thompson, Sampling 3rd, Wiley, New jersey, 2012
8- Timothy C. Urdan, Statistics in Plain English, Lwrence Erlbaum , association publications, London ,2005.