الجمهورية الجزائرية الديمقراطية الشعبية

وزارة التعليم العالي والبحث العلمي

جامعة سطيف1

كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير

الإحصاء 1

- مدعم بتمارين وإمتحانات محلولة -

- حسب المقرر الرسمي لوزارة التعليم العالي والبحث العلمي -

الأستاذ الدكتور:

ساعد بن فرحات

الدكتور:

عبد الحميد قطوش

السنة الجامعية : 2019/2018

المقدمــــــــة

يعتبر علم الإحصاء أحد أهم الأساليب الكمية العلمية الواسعة الاستخدام سواء في مجال البحث أو الإدارة أو اتخاذ القرار. يستعمل كأداة فعالة لجمع ومعالجة وتحليل البيانات حول المشكلة المطروحة ثم الجواب عليها بشكل موضوعي. ويعد أيضا أداة لخدمة متخذي القرار والعلماء في مختلف مجالات المعرفة عن طريق تزويدهم بالمؤشرات التحليلية التي تساعدهم على اتخاذ القرارات الرشيدة بشأن المشكلات قيد الدراسة. ولا غرابة إذا رأينا أن الإحصاء يدرس اليوم في جميع التخصصات ويوظف في جميع مجالات المعرفة العلمية.

بناءا على ما تقدم، وانطلاقا من أهمية هذه المادة وضرورتها بالنسبة للطلبة والباحثين فقد تم تأليف هذه المطبوعة لإثراء مكتباتنا بالمزيد من المؤلفات في هذا المجال.

توفر هذه المطبوعة لطلبة السنة الأولى ل.م.د، مسار العلوم الاقتصادية والتجارية وعلوم التسيير بصفة خاصة وحتى لطلبة مسارات أخرى حزمة متكاملة من المفاهيم والأدوات الأساسية في الإحصاء الوصفي أو الإحصاء1 حسب التسمية الرسمية للمقياس وفق المقرر الرسمي لوزارة التعليم العالي والبحث العلمي.

راعينا عند تأليف هذه المطبوعة العناصر التالية:

- عرض الأسس والقواعد العلمية التي تقوم عليها الأدوات الإحصائية موضوع الدراسة دون أن نبالغ في الغوص في جوانب نظرية غير ضرورية من الناحية التطبيقية.
- ضرورة الاستيعاب الجيد للمدلول الاحصائي والوظيفة التحليلية التي تحملها كل أداة.
- إعطاء أمثلة تطبيقية مع شرح النتيجة.
- تدعيم المفاهيم النظرية في آخر كل فصل بمسائل تطبيقية شاملة من الواقع.

وقد جاءت المادة العلمية لهذه المطبوعة في خمسة فصول. تناول الفصل الأول منه، مدخل عام إلى علم الإحصاء من خلال التعرض إلى أصل وتطور علم الإحصاء، تعريف الإحصاء، مراحله وفروعه، بعض المصطلحات والمفاهيم الإحصائية العامة: الوحدة الإحصائية، المجتمع الإحصائي، المتغير الإحصائي، العينة الإحصائية. أما الفصل الثاني فقد خصص لدراسة التوزيعات التكرارية، وذلك بدراسة التوزيع التكراري للمتغير الإحصائي المنفصل (المتقطع)، المتصل (المستمر) والكيفي، بالإضافة إلى دراسة قضية التمركز. وخصص الفصل الثالث لدراسة مقاييس النزعة المركزية بدءا بالمتوسط الحسابي ومشتقاته ثم الوسيط ومشتقاته وختاما بالمنوال. في حين خصص الفصل الرابع لدراسة مقاييس التشتت المطلقة والنسبية. أما الفصل الخامس والأخير فقد خصص لدراسة مقاييس الشكل. وقد تم عرض في كل فصل تمارين محلولة وأخرى مقترحة، لنختم هذا الكتاب بمجموعة من الامتحانات السابقة والتي أجريت بكلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير بجامعة سطيف1.

وفي الوقت الذي نضع هذا الجهد العلمي المتواضع بين أيدي زملائنا المدرسين والمتخصصين وأبنائنا الطلبة، فإنه يحذونا الأمل في أن تساهم هذه المطبوعة في الاستيعاب الجيد للمادة بالنسبة لفئة الطلبة الموجهة إليهم بصفة خاصة والمساهمة كذلك في تزويد الطلبة في طور التخرج والزملاء الأساتذة الباحثين ببعض الأدوات الإحصائية والمنهجية في إعداد المذكرات والرسائل والبحوث. ونستسمح القارئ الكريم العذر سلفا عن أي نقص ورد في هذه المطبوعة، فالنقص من سمات البشر والكمال لله وحده، ونسأل الله وندعوه سبحانه وتعالى أن يتقبل هذا العمل خالصا لوجهه الكريم.

والله الموفق والهادي إلى سواء السبيل.

الأستاذ الدكتور:

ساعد بن فرحات

الدكتور:

عبد الحميد قطوش

الفصل الأول

مدخل عام إلى علم الإحصاء

نتطرق من خلال هذا الفصل إلى العناصر التالية:

1- أصل وتطور علم الإحصاء.

2- تعريف الإحصاء، مراحله وفروعه.

3- بعض المصطلحات والمفاهيم الإحصائية العامة:

3-1- الوحدة الإحصائية.

3-2- المجتمع الإحصائي.

3-3- المتغير الإحصائي.

3-4- العينة الإحصائية.

تمارين محلولة وتمارين مقترحة

1- أصل وتطور علم الإحصاء :

إن كلمة الإحصاء *La statistique* مصطلح مشتق من كلمتين لاتينيتين:

Status: تعني الحالة أو الوضع، *Statos*: تعني الدولة.

ومن ذلك يمكن أن نفهم أن الإحصاء في تعريف بدائي له يعبر عن حالة أو وضع الدولة بلغة الأرقام، ولكن يبقى هذا المفهوم بسيطا لا يعبر عن الحقيقة العلمية لهذا الميدان من المعرفة.

إن المتتبع للبوادر الأولى لعلم الإحصاء يجد أنها ترجع إلى أزمنة قديمة جدا عند الإنسان البدائي لما تحول من حياة التنقل إلى حياة الاستقرار. مع هذا الاستقرار نتج مفهوم احتلال المجال، أي احتلال قطعة من الأرض واعتبارها مجالا خاصا. بعد ذلك أصبح يهمه أن يعبر على مساحة هذا المجال، عدد الأشجار المثمرة الموجودة في المجال، عدد أفراد الخلية العائلية، عدد الحيوانات التي تمكن من ترويضها. يعبر عن كل ذلك بعدد معين من الحصى. وهذه هي نفسها اهتمامات الدولة الحديثة ولكن بطريقة متطورة حيث أنشئت الدواوين والإدارات المتخصصة في جمع ونشر الإحصائيات في مختلف النشاطات الاجتماعية والاقتصادية لبلد ما. فمثلا في الجزائر الهيئة المكلفة بذلك هي الديوان الوطني للإحصائيات. أما البوادر العلمية للإحصاء كنظرية فلم تظهر إلا بداية من القرن الثامن عشر 18م، حيث توجه الباحثون الرياضيون وعلى رأسهم *Gauss* و *Laplace* و *Bernoulli* نحو التحليل الإحصائي وإنشاء القوانين الاحتمالية. ولم يكتمل الإحصاء كعلم لجمع وعرض وتحليل واستخدام البيانات الإحصائية بغرض الاستدلال واتخاذ القرارات إلا في بداية القرن العشرون 20م (الأربعينيات).

2- تعريف الإحصاء، مراحله وفروعه:

2-1- تعريف الإحصاء:

الإحصاء هو العلم الذي يبحث في طرق جمع البيانات الخاصة بمختلف الظواهر وعرضها وتحليلها للوصول إلى نتائج تساعد في اتخاذ القرارات المناسبة، فالإحصاء بهذا التعريف هو أسلوب منطقي منتظم موحد يعالج الموضوعات والخصائص التي يمكن أن يعبر عنها بصورة رقمية.

أما الإحصائيات فهي البيانات العددية المتعلقة بموضوع ما والمنظمة (في جداول أو رسوم بيانية) حول نشاط أو قطاع معين في الدولة، فمثلا نقول:

- إحصائيات السكان للتعبير عن مجموعة البيانات الخاصة بالسكان في بلد ما (العدد الإجمالي للسكان، توزيع السكان حسب العمر أو الجنس، التوزيع الجغرافي للسكان حسب الولايات)؛
- إحصائيات التجارة الخارجية؛
- إحصائيات التعليم العالي.

وبالتالي فإن الإحصائيات هي المادة الأولية التي تستخدم في علم الإحصاء.

2-2- مراحل (منهج) البحث الإحصائي:

من خلال التعريف السابق للإحصاء يتبين أن منهج البحث الإحصائي يتجسد في عدة مراحل على الباحث أن يتبعها. نوجز هذه المراحل فيما يلي:

2-2-1- التحديد الدقيق للهدف الإحصائي:

ونعني بذلك تحديد نوع المعلومات المراد جمعها، والتي تترجم إلى أسئلة تدرج في وثيقة خاصة تسمى استمارة. يشترط في ذلك التنظيم الجيد والوضوح الكامل للأسئلة، ويستنبط الهدف الإحصائي من الهدف العام من الدراسة الإحصائية.

مثال: نريد إجراء دراسة إحصائية حول مستوى المعيشة للأسرة في الجزائر (الهدف العام).

تحديد الهدف الإحصائي: دخل الأسرة – عدد الأفراد في الأسرة – نوع السكن– عدد الغرف.

2-2-2- جمع البيانات الإحصائية:

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

أ- الطريقة المباشرة والطريقة غير المباشرة:

أ-1- الطريقة المباشرة:

يقصد بهذه الطريقة قيام الباحث بجمع المعلومات الإحصائية بنفسه، من مصادرها الأولية، كأن يقوم بطرح الأسئلة مباشرة على الأسر.

أ-2- الطريقة غير المباشرة:

وتسمى أيضا طريقة البيانات الثانوية، وهي تشمل جميع البيانات والمعلومات الإحصائية المتوفرة من وثائق ومطبوعات ونشرات إحصائية التي تصدرها الهيئات والدواوين المختلفة، وكذلك الهيئات الدولية ومنظماتها المختلفة، ولهذه الطريقة فوائد متعددة أهمها أنها تؤدي إلى اقتصاد كبير في وقت الباحث ونفقاته، إلا أنها تشكو أيضا من عدد من العيوب منها:

- عدم التطابق في بعض الأحيان بين البيانات التي يوفرها المصدر الثانوي والبيانات التي يرغب الباحث في الحصول عليها؛
- نقص كمية البيانات ودرجة الدقة؛
- قد تكون الوحدة الإحصائية المستعملة لا تتطابق وخطة البحث.

ب- طريقة الحصر الشامل وطريقة العينة:

ب-1- طريقة الحصر الشامل:

حيث يتم حصر جميع الوحدات الإحصائية المكونة للمجتمع الإحصائي الخاضع للدراسة، ومن مزايا هذا الأسلوب أنه يعطينا صورة كاملة عن المجتمع الإحصائي، يتميز بالدقة المطلوبة، غير أن هذه الطريقة صعبة التنفيذ وتحتاج إلى تكاليف باهظة وجهاز إحصائي كبير ومتخصص.

ب-2- طريقة العينة الإحصائية:

حيث يتم دراسة جزء من المجتمع الإحصائي فقط، وذلك بأخذ عينة عشوائية من المجتمع ودراسة خواصها واستخلاص المعلومات اللازمة منها، ثم تعميم نتائجها على المجتمع الذي سحبت منه.

2-2-3- عرض البيانات الإحصائية:

بعد جمع البيانات الإحصائية لا بد من عرضها وتصنيفها بشكل يظهر العلاقة بينها، ويتم عرض البيانات بعدة طرق أهمها:

أ- العرض الكتابي:

هذه الطريقة تعني عرض البيانات الإحصائية في سياق فقرة نثرية، وهي معقولة فيما لو كانت المعلومات الإحصائية المعروضة تتألف من عدد قليل من الأرقام، إلا أن الإحصاءات في أغلب الأحيان تتألف من أعداد كثيرة يصعب ذكرها في مضمون النص المكتوب.

ب- العرض الجدولي:

تعرض البيانات في جداول، وذلك بتصنيف المعلومات وترتيبها وفقا لبعض خواصها، وأهم أساليب الترتيب هي:

الترتيب التاريخي، الترتيب الأبجدي، الترتيب الكمي، الترتيب الجغرافي. يشترط في الجدول المعلومات التالية كي يكون مقبولا علميا:

- العنوان الكامل والواضح للجدول (يحدد فيه الموضوع، المكان، الزمان)، ويكون عادة إما في أعلى الجدول أو أسفله ويرقم؛
- وحدة القياس: وتكون في أعلى الجدول إلى اليمين؛
- مصدر الجدول: أي تحديد مصدر البيانات الموجودة في الجدول، ويكون في أسفل الجدول.

وطريقة العرض الجدولي تمتاز بالدقة، ولذلك فهي أهم أسلوب متبع لعرض المعلومات، وما يؤخذ على هذه الطريقة عدم إعطاء فكرة سريعة بمجرد نظرة واحدة إلى الجدول.

مثال: ليكن لدينا الجدول التالي الذي يمثل صادرات دولة معينة خلال الفترة 1992- 1998.

الوحدة: 10^6 دولار　　الجدول (1−1): صادرات دولة معينة خلال الفترة 1992− 1998.

السنة	1992	1993	1994	1995	1996	1998	1999
قيمة الصادرات	98	110	130	120	140	135	145

المصدر: فرضي.

ج− العرض البياني:

يستعمل التمثيل البياني بهدف مقارنة قيم ظاهرة ما حسب المكان أو تطورها حسب الزمان، كما يتيح مقارنة عدة ظواهر في آن واحد. إن استخدام التمثيل البياني يجعل المعلومات الإحصائية أكثر وضوحا وفهما، مما يساعد على أخذ فكرة شاملة وسريعة عن الظاهرة المدروسة أي عكس العرض الجدولي. ومن بين أهم طرق العرض البياني نذكر:

− **الأعمدة:** تمثل البيانات بواسطة أعمدة يتناسب فيها طول العمود مع قيمة العدد (التكرار).

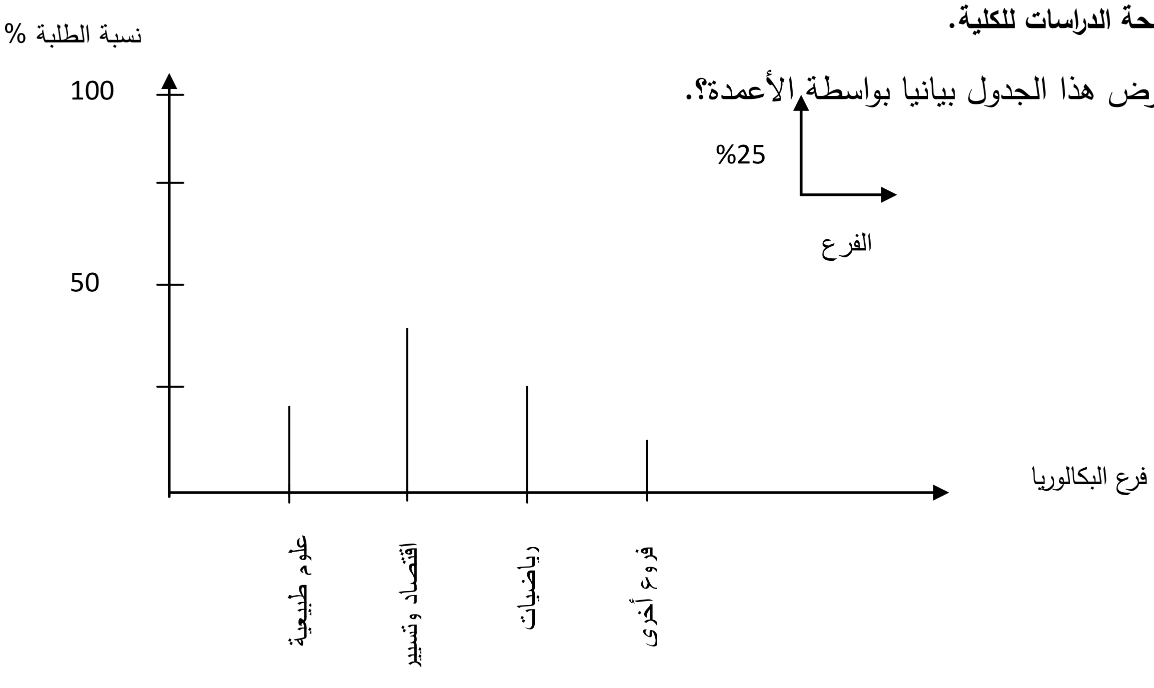
مثال(1−1): يمثل الجدول التالي توزيع طلبة السنة أولى LMD علوم اقتصادية بجامعة سطيف حسب فروع البكالوريا للسنة الجامعية 2007/2006:

الجدول (1−2): توزيع طلبة السنة أولى LMD علوم اقتصادية بجامعة سطيف حسب فروع البكالوريا للسنة الجامعية 2007/2006

فرع البكالوريا	علوم طبيعية	اقتصاد وتسيير	رياضيات	فروع اخرى	المجموع
نسبة الطلبة %	20	40	25	15	100

المصدر: مصلحة الدراسات للكلية.

المطلوب: أعرض هذا الجدول بيانيا بواسطة الأعمدة؟.

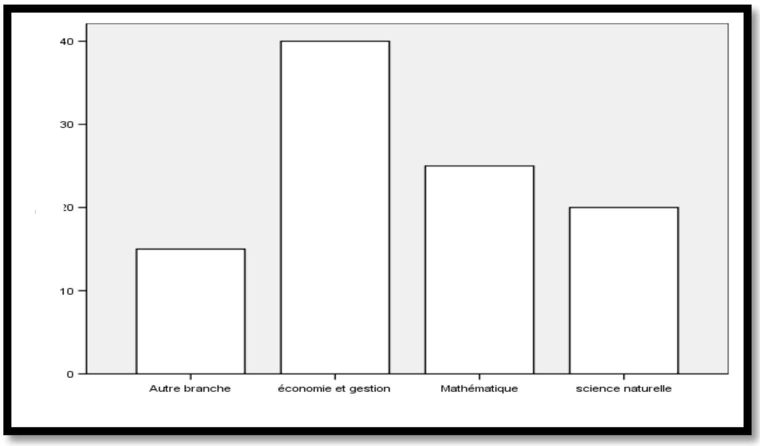


الشكل(1−1): توزيع طلبة السنة الأولى LMD علوم اقتصادية بسطيف حسب فرع البكالوريا لسنة 2007/2006

– **المستطيلات:** تمثل البيانات في هذه الحالة بواسطة المستطيلات، تكون في غالب الأحيان ذات عرض موحد، تتناسب فيها قيمة العدد (التكرار) مع مساحة المستطيل.

مثال(1−2): مثل بيانات الجدول (1−2) بواسطة المستطيلات؟.

نسبة الطلبة



فرع البكالوريا

الشكل(1−2): توزيع طلبة السنة الأولى LMD علوم اقتصادية بسطيف حسب فرع البكالوريا لسنة 2007/2006

– **طريقة الدوائر:** مبدأ هذه الطريقة مبني على ترجمة بيانات الجدول (الأعداد أو النسب) إلى زوايا، حيث يتناسب فيها التكرار مع قيس الزاوية، وذلك بتطبيق القاعدة الثلاثية ثم نقل نتائج الحسابات إلى شكل بياني متمثل في دائرة.

مثال(1−3): مثل بيانات الجدول (1−2) بواسطة الدائرة؟

القاعدة:

$$\mathbf{360^\circ \rightarrow 100\%}$$
$$\boldsymbol{\alpha_i^\circ \rightarrow f_i\% \Rightarrow \alpha_i^\circ = \frac{f_i\%\times 360}{100}}$$

التطبيق:

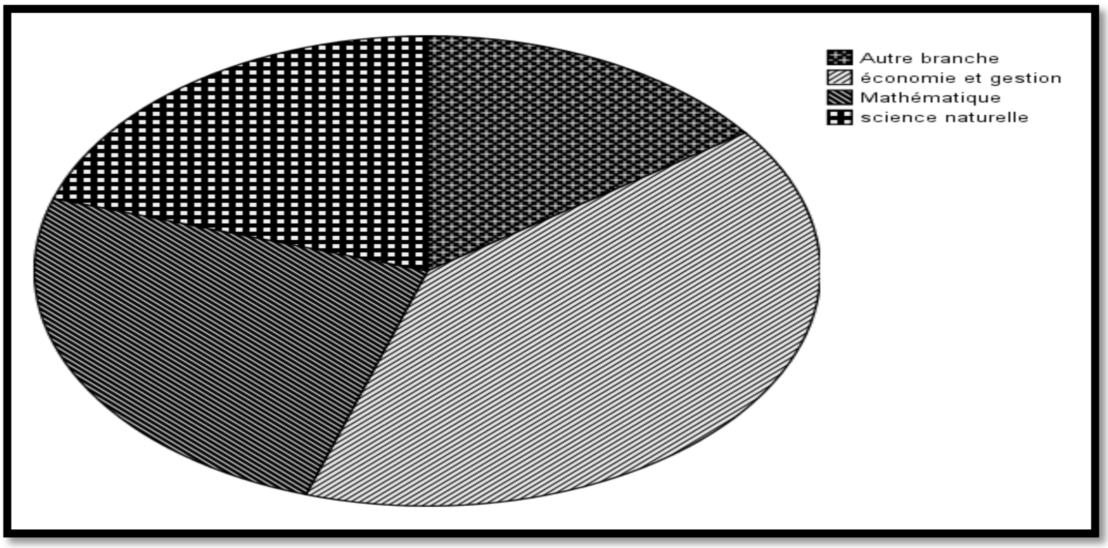
$$\mathbf{360^\circ \rightarrow 100\%}$$
$$\boldsymbol{\alpha_1^\circ \rightarrow 20\% \Rightarrow \alpha_1^\circ = \frac{20\times 360}{100} = 72^\circ}$$

$$\mathbf{360^\circ \rightarrow 100\%}$$
$$\boldsymbol{\alpha_2^\circ \rightarrow 40\% \Rightarrow \alpha_2^\circ = \frac{40\times 360}{100} = 144^\circ}$$

$$\mathbf{360^\circ \rightarrow 100\%}$$
$$\boldsymbol{\alpha_3^\circ \rightarrow 25\% \Rightarrow \alpha_3^\circ = \frac{25\times 360}{100} = 90^\circ}$$

$$\mathbf{360^\circ \rightarrow 100\%}$$
$$\boldsymbol{\alpha_4^\circ \rightarrow 15\% \Rightarrow \alpha_4^\circ = \frac{15\times 360}{100} = 54^\circ}$$

فرع البكالوريا	علوم طبيعية	اقتصاد وتسيير	رياضيات	فروع اخرى	المجموع
نسبة الطلبة %	20	40	25	15	100
الزوايا بالدرجات	72	144	90	54	360



الشكل(1-3): توزيع طلبة السنة الأولى LMD علوم اقتصادية بسطيف حسب فرع البكالوريا لسنة 2007/2006

❖ **ملاحظة:** إذا كان عدد بنود الجدول كثيرا جدا تصبح هذه الطريقة غير صالحة.

– **طريقة الخطوط المنكسرة:** تستعمل هذه الطريقة لتمثيل تطور متغير إحصائي على مدى فترة زمنية متوسطة أو طويلة (شهر بشهر – ثلاثي بثلاثي – سنة بسنة)، وهناك نوعان من الأشكال:

– الخطوط المنكسرة البسيطة: وهي تمثل تطور ظاهرة واحدة أو متغير واحد.

– الخطوط المنكسرة المركبة: وهي تمثل تطور ظاهرتين أو أكثر على الشكل نفسه، والغرض من هذه الطريقة هو مقارنة تطور هذه المتغيرات معا.

– **مثال(1-4):** يعطينا الجدول التالي تطور عدد الطلبة في كلية العلوم الاقتصادية حسب الجنس بين سنتي 1982 و 1989

الجدول (1-3): تطور عدد الطلبة في كلية العلوم الاقتصادية حسب الجنس بين سنتي 1982 و 1989

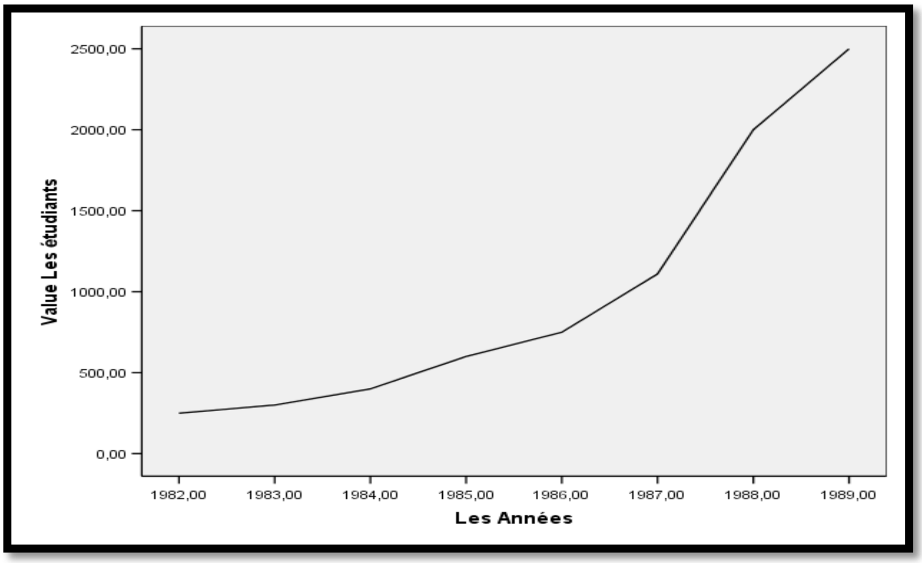
السنة	1982	1983	1984	1985	1986	1987	1988	1989
ذكور	200	240	300	400	450	660	1200	1500
إناث	50	60	100	200	300	450	800	1000
المجموع	250	300	400	600	750	1110	2000	2500

المصدر: مصلحة الدراسات للكلية.

المطلوب: 1- مثل بيانيا تطور العدد الإجمالي للطلبة من 1982 إلى 1989؟

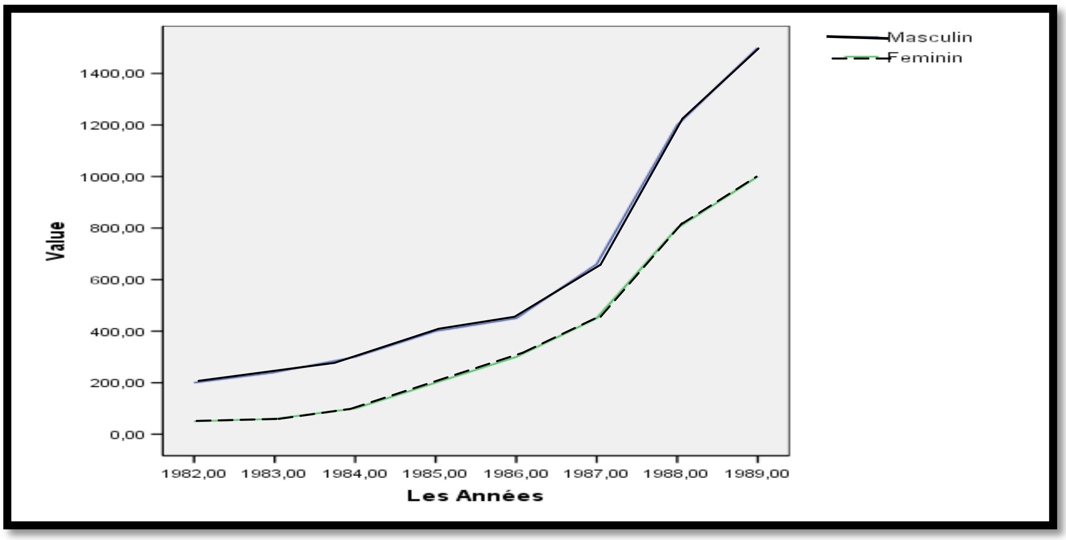
2- مثل بيانيا عدد الطلبة في الفترة نفسها حسب الجنس؟ ماذا تلاحظ؟.

الحل: 1- التمثيل البياني لتطور العدد الإجمالي للطلبة من 1982 إلى 1989:



الشكل(1-4): تطور العدد الإجمالي للطلبة من 1982 إلى 1989

2- التمثيل البياني لعدد الطلبة في الفترة نفسها حسب الجنس:



الشكل(1-5)- تطور عدد للطلبة حسب الجنس من 1982 إلى 1989

- **نلاحظ أن:** عدد الطلبة في تزايد مستمر للذكور والإناث، وأن تطور عدد الذكور أكبر من عدد الإناث في كل سنة.

2-2-4- تحليل البيانات الإحصائية:

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

أ- ترتيب الإحصاءات وتصنيفها، ويمكن أن يكون الترتيب حسب النوع أو الكمية، كتصنيف السكان ما بين أعزب ومتزوج ومطلق وأرمل، كما يمكن أن يكون الترتيب جغرافيا، كأن نوزع السكان في الجزائر حسب الولايات والدوائر والبلديات؛

ب- حساب القيم المركزية لمجموعة البيانات ودراسة التشتت والالتواء فيها؛

ج- دراسة علاقات الارتباط بين عوامل المجتمع الإحصائي؛

د- استنباط التقديرات أو التنبؤات التي تدل عليها الدراسة.

2-2-5- تفسير البيانات الإحصائية:

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

2-3- فروع علم الإحصاء: ينقسم علم الإحصاء إلى:

أ- الإحصاء الوصفي:

هو ذلك الجزء من الإحصاء الذي يهتم بتلخيص البيانات الإحصائية إلى عدد محدود من الأرقام تسمى مقاييس إحصائية أو في جدول إحصائي يسهل القراءة أو في رسوم بيانية، والغرض من كل ذلك هو إعطاء وصفا أوليا للظاهرة المدروسة بدون تحليل معمق.

ب- الإحصاء الاستدلالي:

يستند على فكرة إجراء الدراسة الإحصائية (جمع البيانات) على جزء من المجتمع يسمى العينة، يتم إختيارها بطريقة علمية مناسبة، بغرض استخدام بيانات هذه العينة في التوصل إلى نتائج يمكن تعميمها على مجتمع الدراسة، فنقول لقد استدللنا على خواص المجتمع على أساس خواص العينة، وهذا عكس الاستنباط الذي يعني استخراج خواص الجزء إنطلاقا من خواص الكل.

3- بعض المصطلحات والمفاهيم الإحصائية العامة:

3-1- الوحدة الإحصائية

هي الكائن الواحد أو الخلية الأساسية التي تجرى عليه الدراسة الإحصائية، أي أن أسئلة الاستمارة تدور حوله، سواء أكان هذا الكائن إنسانا أو حيوانا أو شيئا، مثل: إنسان، بقرة، سيارة،.....إلخ.

أمثلة:

- دراسة إحصائية حول مستوى المعيشة للسكان في الجزائر، الوحدة الإحصائية هي الأسرة الواحدة.

- سبر الآراء حول الأوضاع السياسية والاجتماعية في الجزائر، الوحدة الإحصائية هي فرد في الجزائر (رجل أو إمرأة) عمره 18 سنة فأكثر.

- دراسة إحصائية حول إنتشار الدودة البيضاء في الأراضي الزراعية في ولاية ما، الوحدة الإحصائية هي 1م^2 من الأراضي الزراعية التي من الممكن أن تصيبها هذه الآفة في الولاية المعنية.

3-2- المجتمع الإحصائي:

هو مجموع الوحدات الإحصائية المراد دراستها والمعرفة بشكل دقيق والتي تشترك فيما بينها في الصفة الأساسية محل اهتمام الباحث، مثل: مجتمع من الطلبة، مجتمع من الأسر، مجتمع من المؤسسات.

أمثلة:

- دراسة إحصائية حول مستوى المعيشة للسكان في الجزائر، المجتمع الإحصائي هو جميع الأسرة الجزائرية في فترة الدراسة.

- سبر الآراء حول الأوضاع السياسية والاجتماعية في الجزائر، المجتمع الإحصائي هو جميع الأفرد في الجزائر (رجال أو نساء) أعمارهم 18 سنة فأكثر في فترة الدراسة.

- دراسة إحصائية حول إنتشار الدودة البيضاء في الأراضي الزراعية في ولاية ما، المجتمع الإحصائي هو جميع الأراضي الزراعية التي من الممكن أن تصيبها هذه الآفة في الولاية المعنية في فترة الدراسة.

3-3- المتغير الإحصائي:

3-3-1- تعريف المتغير الإحصائي:

هو الخاصة أو الصفة (النوعية أو الكمية) المشتركة لكل الوحدات الإحصائية التي تشكل المجتمع الإحصائي، مثل: الطول، السن، المستوى التعليمي، الإنتاج، ....إلخ.

3-3-2- أنواع المتغيرات الإحصائية:

تنقسم المتغيرات الإحصائية إلى قسمين:

أ- **متغيرات كيفية:** هي تلك المتغيرات التي لا يمكن قياسها كميا، إنما تأخذ أوصافا، وتنقسم بدورها إلى قسمين:

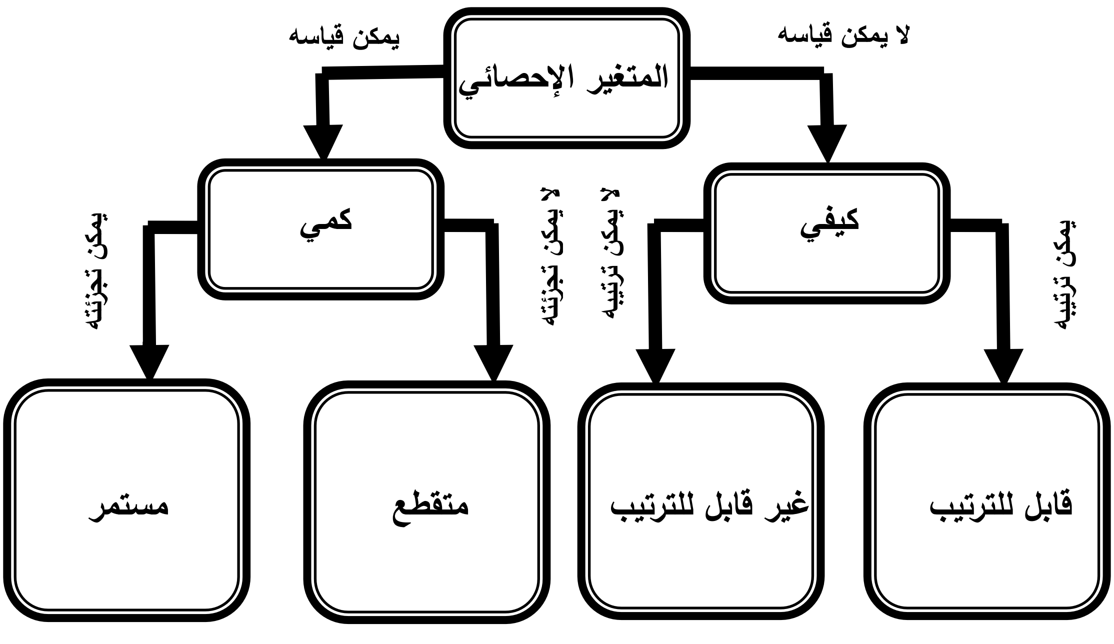
أ-1- متغيرات كيفية قابلة للترتيب: مثل المستوى التعليمي (إبتدائي، متوسط، ثانوي، جامعي)، ...إلخ.

أ-2- متغيرات كيفية غير قابلة للترتيب: مثل الجنسية (جزائري، تونسي،...)، الجنس (ذكور، إناث)، الحالة العائلية (أعزب، أرمل، متزوج، مطلق)، اللون (أسود، أبيض،...)،...إلخ.

ب- متغيرات كمية: هي تلك المتغيرات التي يمكن قياسها، وهي أكثر المتغيرات انتشارا واستعمالا لأن لغة الإحصاء هي لغة الأرقام، والمتغيرات الكمية تنقسم بدورها إلى قسمين:

ب-1- متغيرات كمية متقطعة: هي تلك المتغيرات التي تأخذ قيما صحيحة لا يمكن تجزئتها، مثل عدد الأطفال في الأسرة الواحدة، عدد قطع الغيار المنتجة....إلخ.

ب-2- متغيرات كمية مستمرة: هي تلك المتغيرات التي تأخذ كل القيم الممكنة لمجال الدراسة، ونظرا للعدد غير المتناهي لهذه القيم نقسم مجال القيم إلى مجالات جزئية تسمى الفئات، مثال الطول، السن، الوزن،....إلخ، وللمتغير الكمي المستمر وحدة قياس (متر مربع، سنتمتر، الدينار...إلخ)، والشكل (1-6) يلخص أنواع المتغيرات الإحصائية.



الشكل (1-6): مخطط يوضح أنواع المتغيرات الإحصائية

3-4- العينة الإحصائية:

هي جزء من المجتمع الإحصائي يتم استخراجها بطرق إحصائية معينة حتى تكون ممثلة للمجتمع الإحصائي أحسن تمثيل، ويتم الاعتماد عليها في الدراسة بدل المجتمع للأسباب التالية:

أ- كبر حجم المجتمع؛

ب- ربحا للوقت والجهد والمال؛

ج- الفحص قد يكون مؤذيا أو متلفا للوحدات؛

د- قد تكون الدراسة الشاملة مستحيلة في حالة حجم المجتمع غير محدود.

والعينات أنواع نذكر منها:

أ- العينة العشوائية البسيطة:

هي العينة التي تعطي فيها لجميع مفردات المجتمع المراد بحثه نفس الفرصة في الاختيار، وهذا يعني عدم الاهتمام ببعض المفردات أكثر من البعض الآخر وإتاحة الفرص المتكافئة أمام كل مفردة للظهور في العينة، ويمكن أن نحقق ذلك بأن نحضر عددا من البطاقات المتشابهة (في اللون والوزن والحجم وكل شيء) ويكتب على كل بطاقة رقما يمثل مفردة من مفردات المجتمع ونسحب العدد المطلوب من هذه البطاقات (بعد خلطها جيدا) فنجد أن الأرقام المسجلة عليها تعطي لنا المفردات التي تم اختيارها بطريقة عشوائية.

ب- العينة الطبقية:

إذا كان المجتمع يتكون من مجموعات من المفردات تتصف بالتجانس داخل كل مجموعة وبالتباين بين المجموعات المختلفة، ويراد أخذ عينة تكون ممثلة بقدر الإمكان لهذا المجتمع فلابد أن تكون هذه المجموعات ممثلة في العينة، وذلك بتقسيم المجتمع إلى أقسام تعرف بالطبقات، ثم تؤخذ عينة عشوائية من كل طبقة وبذلك نضمن تمثيل العينة لكل طبقات المجتمع.

ج- العينة متعددة المراحل:

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

تمارين محلولة للفصل الأول

التمرين الأول:

حدد كلا من: الوحدة الإحصائية، المجتمع الإحصائي، المتغير الإحصائي ونوعه فيما يلي:

1- توزيع عينة من 30 مؤسسة اقتصادية حسب رقم أعمالها السنوي بولاية سطيف.

2- عدد الغرف في المسكن الواحد لعينة من 50 مسكنا ببلدية سطيف.

3- توزيع 360 حالة زواج في إحدى البلديات حسب عمر الزوجة.

4- دراسة إحصائية حول الأجور الشهرية بالدينار لــ 65 عاملا في شركة متوسطة.

5- توزيع عينة من 40 فرد من الجالية المغاربية في فرنسا حسب البلد الأصلي.

6- توزيع عينة من 50 فرد حسب المستوى التعليمي في بلدية ما.

التمرين الثاني:

بغرض التعرف على احتياجات سكان ولاية سطيف من مادتي السميد والخبز (الجاهز لدى الخبازين)، قررت مؤسسة الصناعات الغذائية من الحبوب ومشتقاته بسطيف إجراء دراسة إحصائية حول الموضوع:

1- ماهو الهدف العام من الدراسة؟

2- ماهي المتغيرات الإحصائية المدروسة لكل نوع من الاستهلاك (السميد والخبز)؟

3- أذكر طبيعة كل متغير؟

4- ماهي الوحدة الإحصائية والمجتمع الإحصائي في هذه الدراسة؟

5- ماهي الطريقة الملائمة لجمع البيانات في مثل هذه الدراسة؟ علل ذلك؟

التمرين الثالث:

بهدف معرفة مدى أهمية الثروة الحيوانية للأبقار الحلوب في أحد بلديات ولاية سطيف أجريت دراسة على عينة من 42 مزرعة فكانت النتائج كالتالي:

1	3	2	4	10	9	5	3	2	1	0	11	6	6
8	8	1	1	0	0	5	7	9	7	2	4	5	8
2	4	5	7	1	0	2	4	6	7	5	8	5	3

المطلوب:

1- حدد في هذه المسألة: الوحدة الإحصائية، المجتمع الإحصائي، الخاصة الإحصائية ونوعها، أسلوب الدراسة الإحصائية (بواسطة العينة أو الشاملة، مباشرة أو غير المباشرة)؟

2- رتب هذه البيانات تصاعديا؟

3- أعرض هذه البيانات في جدول؟

4- مثل معطيات الجدول في رسم بياني بواسطة الأعمدة؟

5- ماهي أكبر وأصغر قيمة للتكرارات المطلقة؟ ماهي قيمة المتغير التي توافقها؟ اشرح معنى هذه التكرارات؟

6- ماهي نسبة المزارع التي تملك: بقرة واحدة؟ 8 بقرات؟

التمرين الرابع:

يتكون مجتمع من أربع فئات اجتماعية – مهنية، حجمه $N = 10000$، وحجم كل فئة هو كالتالي:

$$N_1 = 2000 \ , \ \mathrm{N}_2 = 3000 \ \ , \mathrm{N}_3 = 4000 \ , \ \mathrm{N}_4 = 1000$$

نريد سحب عينة حجمها $n = 180$:

1- ماهي طبيعة المجتمع المدروس؟

2- حدد عدد الوحدات الإحصائية التي يمكن سحبها من كل فئة؟

الحـــــــــلول

حل التمرين الأول:

تحديد كلا من: الوحدة الإحصائية، المجتمع الإحصائي، المتغير الإحصائي ونوعه:

المثال	المجتمع الإحصائي	الوحدة الإحصائية	المتغير الإحصائي	نوعه
01	جميع المؤسسات الاقتصادية بولاية سطيف	المؤسسة الواحدة	رقم الأعمال السنوي	كمي مستمر
02	جميع السكنات في بلدية سطيف	المسكن الواحد	عدد الغرف	كمي متقطع
03	360 حالة زواج في البلدية	حالة الزواج الواحدة	عمر الزوجة	كمي مستمر
04	65 عاملا في شركة متوسطة	العامل الواحد	الأجر الشهري	كمي مستمر
05	العدد الإجمالي للجالية المغاربية بفرنسا	الفرد الواحد	البلد الأصلي	كيفي غير قابل للترتيب
06	جميع سكان البلدية	الفرد الواحد	المستوى التعلمي	كيفي قابل للترتيب

حل التمرين الثاني:

1- الهدف العام من الدراسة: معرفة مدى إحتياجات سكان ولاية سطيف من مادتي السميد والخبز.

2- الخاصات أو المتغيرات الاحصائية المدروسة لكل نوع من الاستهلاك (السميد والخبز):

المتغير الأول: كمية السميد بالكيلوغرام المستهلكة يوميا المتغير الثاني: عدد الخبزات المشتراة يوميا.

3- طبيعة كل متغير:

كمية السميد بالكيلوغرام: متغير كمي مستمر عدد وحدات الخبز: متغير كمي متقطع.

4- الوحدة الاحصائية والمجتمع الاحصائي في هذه الدراسة:

الوحدة الإحصائية: الأسرة الواحدة.

المجتمع الإحصائي: جميع الأسر التي تقطن في ولاية سطيف خلال الفترة محل الدراسة.

5- الطريقة الملائمة لجمع البيانات في مثل هذه الدراسة:

- الطريقة المباشرة: لأنها تحتاج إلى نزول ميداني واستجواب بطريقة مباشرة مع وحدات الدراسة (الأسر).
- طريقة العينة: وهذا نظرا لكبر حجم المجتمع (صعوبة الحصر الشامل)، وكذلك ربحا للوقت والجهد والمال.

حل التمرين الثالث:

1- تحديد: الوحدة الاحصائية، المجتمع الاحصائي، الخاصة الاحصائية ونوعها، أسلوب الدراسة الاحصائية:

المجتمع الإحصائي	الوحدة الإحصائية	المتغير الإحصائي	نوعه	أسلوب الدراسة
جميع المزارع بالبلدية	المزرعة الواحدة	عدد الأبقار الحلوب	كمي متقطع	- الطريقة المباشرة - طريقة العينة

2- ترتيب البيانات تصاعديا:

0	0	0	0	1	1	1	1	1	2	2	2	2	2
3	3	3	4	4	4	4	5	5	5	5	5	5	6
6	6	7	7	7	7	8	8	8	8	9	9	10	11

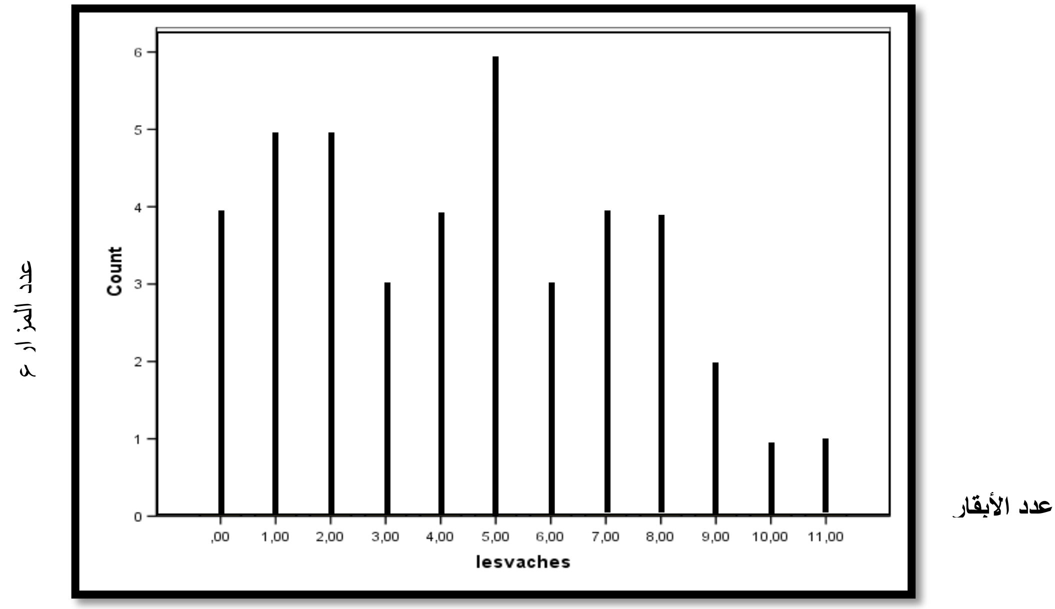
3- عرض هذه البيانات في جدول:

الجدول(1-4): توزيع 42 مزرعة حسب عدد الأبقار الحلوب.

عدد الأبقار (قيم المتغير) X_i	عدد المزارع (التكرار) n_i
0	4
1	5
2	5
3	3
4	4
5	6
6	3
7	4
8	4
9	2
10	1
11	1
المجموع $\sum n_i$	**42**

المصدر: دراسة ميدانية

4- تمثيل معطيات الجدول في رسم بياني بواسطة الأعمدة:



الشكل(1-7): توزيع 42 مزرعة حسب عدد الأبقار الحلوب

5- أكبر وأصغر قيمة للتكرارت المطلقة، قيمة المتغير التي توافقها، شرح معنى هذه التكرارات:

- أكبر قيمة للتكرارت المطلقة: 6 - أصغر قيمة للتكرارت المطلقة: 1
- القيمة المقابلة لأكبر تكرار: 5 - القيمة المقابلة لأقل تكرار: 10 و 11

الشرح:

- **القيمة 5 ذات التكرار 6:** عدد المزارع التي تحتوي على 5 بقرات هو 6 مزارع.
- **القيمة 10 ذات التكرار 1:** عدد المزارع التي تحتوي على 10 بقرات هو مزرعة واحدة.
- **القيمة 11 ذات التكرار 1:** عدد المزارع التي تحتوي على 11 بقرة هو مزرعة واحدة.

6- نسبة المزارع التي تملك:

- **بقرة واحدة:**

$$\begin{array}{l} \mathbf{42} \;\rightarrow\; \mathbf{100\%} \\ \mathbf{5} \;\rightarrow\; \boldsymbol{f}\% \;\Rightarrow \boldsymbol{f}\% = \mathbf{11,9\%} \end{array}$$

- **8 بقرات:**

$$\begin{array}{l} \mathbf{42} \;\rightarrow\; \mathbf{100\%} \\ \mathbf{4} \;\rightarrow\; \boldsymbol{f}\% \;\Rightarrow \boldsymbol{f}\% = \mathbf{9,52\%} \end{array}$$

حل التمرين الرابع:

1- طبيعة المجتمع المدروس: غير متجانس لأنه يتكون من فئات اجتماعية متباينة.

2- عدد الوحدات الإحصائية التي يمكن سحبها من كل فئة:

بتطبيق القاعدة الثلاثية:

$$\boldsymbol{N} \rightarrow \boldsymbol{n}$$

$$\boldsymbol{N_i} \rightarrow \boldsymbol{n_i} \;\Rightarrow\; \boldsymbol{n_i} = \frac{\boldsymbol{N_i}}{\boldsymbol{N}} \times \boldsymbol{n}$$

$$n_1 = \frac{\mathrm{N}_1}{\mathrm{N}} \times n = \frac{2000}{10000} \times 180 = 36$$

$$n_2 = \frac{\mathrm{N}_2}{\mathrm{N}} \times n = \frac{3000}{10000} \times 180 = 54$$

$$n_3 = \frac{\mathrm{N}_3}{\mathrm{N}} \times n = \frac{4000}{10000} \times 180 = 72$$

$$n_4 = \frac{\mathrm{N}_4}{\mathrm{N}} \times n = \frac{1000}{10000} \times 180 = 18$$

تمارين مقترحة

التمرين الأول:

1- عرف الإحصاء، ماهي فروعه، مع توضيح كل فرع؟

2- ما الفرق بين الإحصاء والإحصائيات؟

3- ماهي مراحل البحث الإحصائي؟

4- عرف المتغيرات الإحصائية، ماهي أنواعها، مع الشرح وإعطاء مثال لكل نوع؟

5- إليك التعريف التالي والذي يمثل طريقة من طرق جمع البيانات:"**هي جزء من المجتمع الإحصائي يتم استخراجها بطرق إحصائية معينة حتى تكون ممثلة للمجتمع الإحصائي أحسن تمثيل**".

أ- ماهي هذه الطريقة؟

ب- يتم الاعتماد عليها في الدراسة بدل طريقة أخرى:

- ماهي هذه الطريقة؟

- ماهي أسباب الاعتماد عليها؟

ج- أذكر بعض أنواع هذه الطريقة؟ مع الشرح؟

التمرين الثاني:

تتكون كلية الاقتصاد بجامعة فرحات عباس – سطيف – من خمس فئات طلابية موزعة كالتالي:

- حجم المجتمع $N = 10000$،

- حجم كل فئة هو كالتالي:

- عدد طلبة السنة الأولى LMD هو: 4200 طالب.

- عدد طلبة السنة الثانية LMD هو: 3100 طالب.

- عدد طلبة السنة الثالثة LMD هو: 2000 طالب.

- عدد طلبة الماستر 1 هو: 400 طالب.

- عدد طلبة الماستر 2 هو: 300 طالب.

نريد سحب عينة حجمها $n = 200$:

المطلوب:

1- ماهي طبيعة المجتمع المدروس؟ ما اسم هذه العينة؟

2- حدد عدد الوحدات الإحصائية التي يمكن سحبها من كل فئة؟

التمرين الثالث:

بهدف التعرف على الفئات الاجتماعية الأكثر فقرا وبطلب من الحكومة قرر الديوان الوطني للإحصائيات إجراء بحثا إحصائيا حول الموضوع في الجزائر.

1- ما هو الهدف العام من البحث؟

2- ما هو المتغير الإحصائي الذي يجب دراسته لتحقيق هذا الهدف؟

3- هل هذا المتغير من النوع المنفصل أو المتصل؟

4- حدد الوحدة الإحصائية والمجتمع الإحصائي في هذه الدراسة؟

5- ما هي الطريقة الملائمة لجمع البيانات في هذا البحث؟

التمرين الرابع:

الجدول التالي يبين توزيع اليد العاملة حسب قطاعات النشاط في دولة ما خلال الفترة 1995- 2000

الجدول(1-5): توزيع اليد العاملة حسب قطاعات النشاط في دولة ما خلال الفترة 1995- 2000

القطاعات	1995	1996	1997	1998	1999	2000
الفلاحة	969	963	960	960	960	990
الصناعة	431	458	468	475	495	510
النقل	142	148	152	160	166	170
المجموع	1542	1569	1580	1595	1621	1670

المصدر: فرضي

المطلوب:

1- مثل بيانيا تطور العدد الاجمالي للعمال؟ ما نوع هذا التمثيل؟ برر ذلك؟

2- مثل بيانيا تطور عدد للعمال حسب قطاع النشاط على نفس الشكل؟ ما نوع هذا التمثيل؟ برر ذلك؟

3- مثل تطور اليد العاملة في القطاعات المذكورة عن طريق الأعمدة؟

4- مثل تطور اليد العاملة في كل من الفلاحة والصناعة والنقل عن طريق الأعمدة؟

5- مثل بيانيا توزع اليد العاملة في مختلف القطاعات عن طريق الدائرة، وذلك خلال سنة 2000؟

الفصل الثاني

التوزيعات التكرارية

نتطرق من خلال هذا الفصل إلى العناصر التالية:

أولا- التوزيع التكراري للمتغير الإحصائي المنفصل (المتقطع).

ثانيا- التوزيع التكراري للمتغير الإحصائي المتصل (المستمر).

ثالثا- التوزيع التكراري للمتغير الإحصائي الكيفي.

رابعا- دراسة قضية التمركز.

تمارين محلولة وتمارين مقترحة

أولا- التوزيع التكراري للمتغير الإحصائي المنفصل (المتقطع):

1- التوزيع التكراري المطلق والنسبي وتمثيلهما البياني:

أ- التوزيع التكراري المطلق:

هو عبارة عن جدول يحتوي في صورته البسيطة على العناصر التالية:

أ-1- قيم المتغير الإحصائي:

وتتمثل في مختلف القيم التي يمكن أن يأخذها المتغير الإحصائي المدروس مرتبة ترتيبا تصاعديا وتظهر في العمود الأول ونرمز لها بالرمز X_{i} (i يشير إلى السطر في الجدول بحيث $i = 1,2,3 \dots, k$).

أ-2- التكرار المطلق:

وهو يمثل عدد المرات التي تتكرر فيها نفس القيمة ونرمز له بالرمز n_i.

مثال(2-1): البيانات التالية تمثل نتائج لدراسة إحصائية حول عدد الغرف في المسكن الواحد لعينة من 50 مسكن ببلدية سطيف.

5	2	4	3	3	6	3	2	4	4
2	2	4	3	7	5	4	8	7	4
3	4	7	3	5	2	8	4	3	6
4	5	2	4	6	3	6	3	4	3
2	4	3	5	1	4	5	3	3	2

المطلوب: أنشئ جدول التوزيع التكراري واشرح كل من n_2 و n_4؟

الحل: نقوم بترتيب البيانات تصاعديا ثم نعرضها في جدول توزيع تكراري كما يلي:

1، 22222222، 3333333333333، 4444444444444، 555555، 6666، 777، 88.

الجدول(2-1): توزيع 50 مسكن حسب عدد الغرف ببلدية سطيف

عدد الغرف (قيم المتغير) X_{i}	عدد المساكن (التكرار) n_i
1	1
2	8
3	13
4	13
5	6
6	4
7	3
8	2
المجموع	$\sum n_i = 50$

الشرح:

$n_2 = 8$: هناك 8 مساكن من بين 50 مسكنا عدد الغرف فيها يساوي 2.

$n_4 = 13$: هناك 13 مساكن من بين 50 مسكنا عدد الغرف فيها يساوي 4.

ملاحظة:

مجموع التكرارت n_i دائما يساوي حجم العينة n أي: $\boxed{\sum n_i = n}$

ب- التوزيع التكراري النسبي:

هو حاصل قسمة التكرار المطلق لكل قيمة من قيم المتغير الإحصائي المتقطع على مجموع التكرارات

$$\boxed{f_i = \frac{f_i}{\sum n_i}}$$

أما التكرار النسبي المئوي فهو عبارة عن التكرار النسبي مضروبا في مائة: $\boxed{f_{i\%} = \frac{f_i}{\sum n_i} \times 100}$

مثال(2-2): بالعودة إلى المثال(2-1)، أحسب التكرارات النسبية؟

الحل: الجدول(2-2): التوزيع النسبي للسكنات حسب عدد الغرف ببلدية سطيف

عدد الغرف X_i	عدد المساكن n_i	f_i	$f_{i\%}$
1	1	0,02	02
2	8	0,16	16
3	13	0,26	26
4	13	0,26	26
5	6	0,12	12
6	4	0,08	08
7	3	0,06	06
8	2	0,04	04
المجموع	$\sum n_i = 50$	**1**	**100**

الشرح:

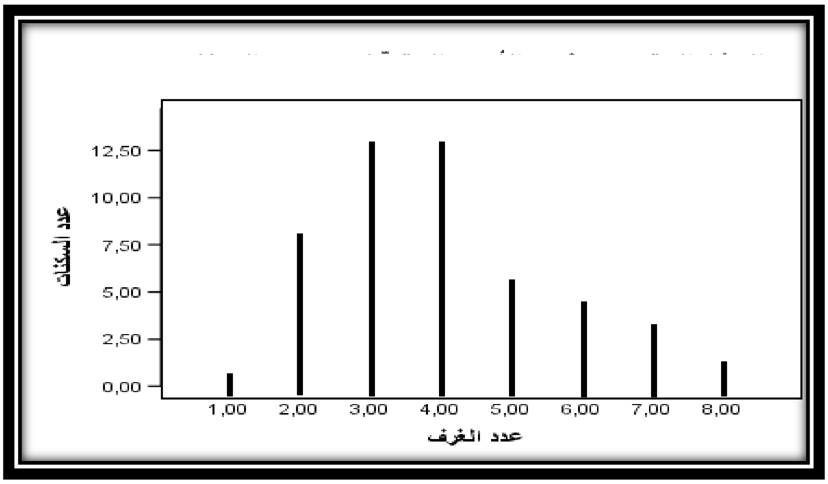
$f_2 = 0.16$: هناك 16 % من المساكن عدد الغرف فيها يساوي 2.

$f_{4\%} = 26\%$: هناك 26% من المساكن عدد الغرف فيها يساوي 4.

ج- التمثيل البياني للتوزيع التكراري المطلق والنسبي:

يمثل التكرار المطلق أو النسبي للمتغير الإحصائي المتقطع عن طريق **الأعمدة**، حيث يتناسب طول العمود مع التكرار المطلق أو النسبي الموافق له.

مثال(2-3): مثل بيانيا معطيات المثال(2-1)؟



الشكل(2-1): توزيع 50 مسكن حسب عدد الغرف ببلدية سطيف

2- التوزيع التكراري التجميعي الصاعد والنازل وتمثيلهما البياني:

أ- التوزيع التكراري التجميعي الصاعد:

أ-1- التوزيع التكراري التجميعي الصاعد المطلق $N_i^{\uparrow}$:

يحسب كالتالي:

$$N_1^{\uparrow} = n_1$$

$$N_2^{\uparrow} = n_1 + n_2 \Rightarrow N_2^{\uparrow} = N_1^{\uparrow} + n_2$$

$$N_3^{\uparrow} = n_1 + n_2 + n_3 \Rightarrow N_3^{\uparrow} = N_2^{\uparrow} + n_3$$

$$\boxed{N_i^{\uparrow} = n_1 + n_2 + \cdots + n_i = N_{i-1}^{\uparrow} + n_i}$$

أ-2- التوزيع التكراري التجميعي الصاعد النسبي $F_i^{\uparrow}$:

يحسب كالتالي:

$$\boxed{F_i^{\uparrow} = f_1 + f_2 + \cdots + f_i = F_{i-1}^{\uparrow} + f_i}$$

أو:

$$\boxed{F_i^{\uparrow} = \frac{N_i^{\uparrow}}{\sum n_i}}$$

أ-3- التوزيع التكراري التجميعي الصاعد النسبي المئوي $F_i^{\uparrow}\%$: $\boxed{F_i^{\uparrow}\% = F_i^{\uparrow} \times 100}$

ملاحظة: التكرار المتجمع الصاعد المطلق الأول يساوي دائما التكرار المطلق الأول($N_i^{\uparrow} = n_1$)، والتكرار المتجمع الصاعد المطلق الأخير يساوي دائما مجموع التكرارات ($N_k^{\uparrow} = \sum n_i$).

ب- التوزيع التكراري التجميعي النازل:

ب-1- التوزيع التكراري التجميعي النازل المطلق $N_i^{\downarrow}$

يحسب كالتالي:

$$N_1^{\downarrow} = n$$

$$N_2^{\downarrow} = n - n_1 \Rightarrow N_2^{\downarrow} = N_1^{\downarrow} - n_1$$

$$N_3^{\downarrow} = n - n_1 - n_2 \Rightarrow N_3^{\downarrow} = N_2^{\downarrow} - n_2$$

$$\boxed{N_i^{\downarrow} = n - n_1 - \cdots - n_{i-1} = N_{i-1}^{\downarrow} - n_{i-1}}$$

ب-2- التوزيع التكراري التجميعي النازل النسبي $F_i^{\downarrow}$:

يحسب كالتالي:

$$\boxed{F_i^{\downarrow} = 1 - f_1 - \cdots - f_{i-1} = F_{i-1}^{\downarrow} - f_{i-1}}$$

أو:

$$\boxed{F_i^{\downarrow} = \frac{N_i^{\downarrow}}{\sum n_i}}$$

ب-3- التوزيع التكراري التجميعي النازل النسبي المئوي $F_i^{\downarrow}\%$: $\boxed{F_i^{\downarrow}\% = F_i^{\downarrow} \times 100}$

ملاحظة: التكرار المتجمع النازل المطلق الأول يساوي دائما مجموع التكرارات ($N_1^{\downarrow} = \sum n_i$)، والتكرار المتجمع النازل الأخير يساوي دائما التكرار المطلق الأخير($N_k^{\downarrow} = n_k$).

مثال(2-4): بالعودة إلى بيانات المثال(2-1)، أحسب كلا من:

$N_i^{\uparrow}$، $N_i^{\downarrow}$، $F_i^{\uparrow}$، $F_i^{\downarrow}$ ، $F_i^{\uparrow}\%$، $F_i^{\downarrow}\%$ ، ثم إشرح كلا من: $N_2^{\uparrow}$، $N_5^{\downarrow}$، $F_2^{\uparrow}\%$، $F_5^{\downarrow}\%$

الجدول(2-3): التكرارات المطلقة والنسبية التجميعية الصاعدة والنازلة لتوزيع السكنات ببلدية سطيف حسب عدد الغرف

عدد الغرف X_i	عدد المساكن n_i	$N_i^{\uparrow}$	$N_i^{\downarrow}$	$F_i^{\uparrow}$	$F_i^{\downarrow}$	$F_i^{\uparrow}\%$	$F_i^{\downarrow}\%$
1	1	1	50	0,02	1	2	100
2	8	9	49	0,18	0,98	18	98
3	13	22	41	0,44	0,82	44	82
4	13	35	28	0,70	0,56	70	56
5	6	41	15	0,82	0,30	82	30
6	4	45	9	0,90	0,18	90	18
7	3	48	5	0,96	0,10	96	10
8	2	50	2	1	0,04	100	4
المجموع	$\sum n_i = 50$	/	/	/	/	/	/

الشرح:

$N_2^{\uparrow} = 9$: هناك 9 مساكن من بين 50 مسكنا عدد الغرف فيها أقل أو يساوي 2.

$N_5^{\downarrow} = 15$: هناك 15 مسكنا من بين 50 مسكنا عدد الغرف فيها أكبر أو يساوي 5.

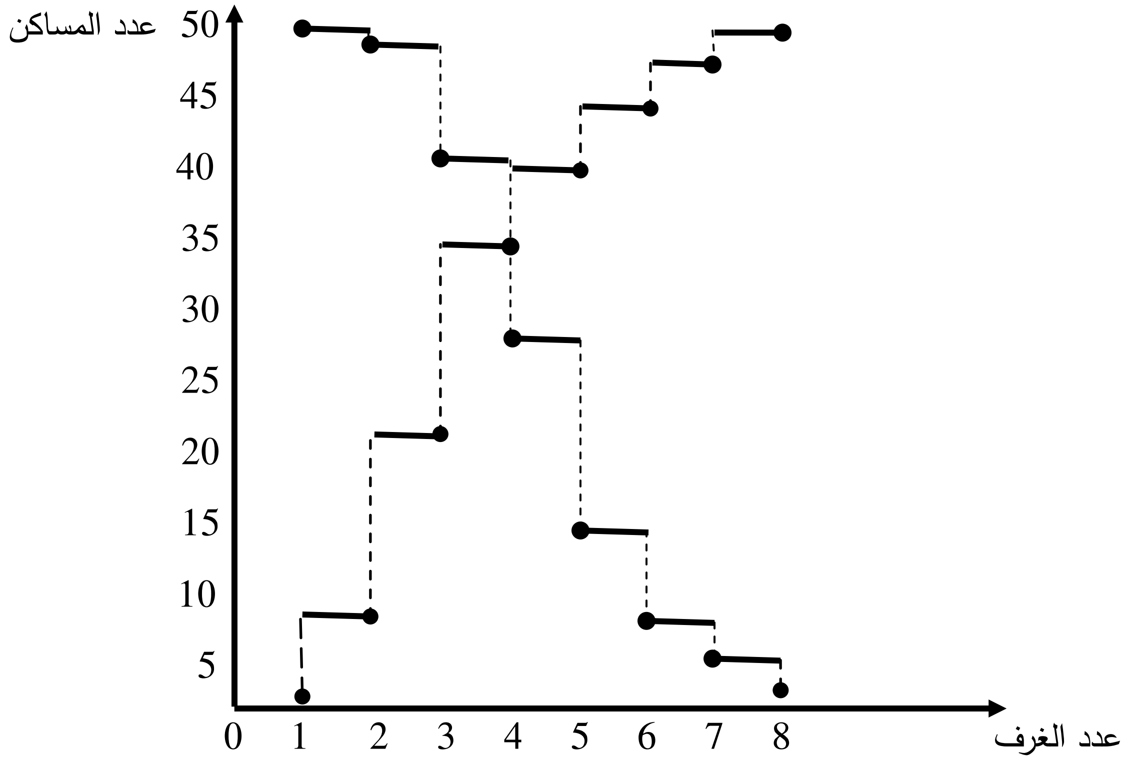
$F_2^{\uparrow}\% = 18\%$: هناك 18% من المساكن عدد الغرف فيها أقل أو يساوي 2.

$F_5^{\downarrow}\% = 30\%$: هناك 30% من المساكن عدد الغرف فيها أكبر أو يساوي 5.

ج- التمثيل البياني للتوزيع التكراري التجميعي الصاعد والنازل:

يمثل التكرار التجميعي الصاعد والنازل المطلق أو النسبي للمتغير الإحصائي المتقطع عن طريق **المنحنى المتجمع الصاعد والنازل،** حيث نلاحظ أنه يأخذ الشكل السلمي إما صاعدا أو نازلا، فنسميه **منحنى سلمي.** كما أنه يظهر على شكل أجزاء متقطعة دلالة على أن المتغير من النوع المنفصل أو المتقطع.

مثال(2-5): التمثيل البياني عن طريق المنحنى المتجمع الصاعد والنازل للمثال(2-1)



الشكل(2−2):التكرارات المطلقة الصاعدة والنازلة لتوزيع 50 مسكن حسب عدد الغرف ببلدية سطيف

ثانيا− التوزيع التكراري للمتغير الإحصائي المتصل (المستمر).

1− جدول التوزيع التكراري لمتغير إحصائي مستمر:

إذا كان المتغير الإحصائي من النوع المتصل رأينا أنه يقبل عددا غير متناهي من القيم الممكنة، المحصورة بين أصغر X_{min} قيمة وأكبر قيمة X_{max} وعليه يستحيل أن نمثله بجدول على شكل قيم فردية كما هو الحال في المتغير المنفصل، فنلجأ في هذه الحالة إلى تجميع أو تكثيف البيانات في مجموعات جزئية نسميها " **فئات** ". فماهو عدد الفئات التي يمكن تحديدها وكيف؟

ليس هناك قاعدة نظرية لذلك، وإنما يشترط أن لا يكون عدد الفئات كثير جدا (يفوق 15 فئة) فيصبح الجدول ضخما يصعب تحليله وقراءته، أو يكون عدد الفئات قليل جدا (أقل من 5 فئات) فيصبح الجدول مبسط جدا أين يفقد حينها دقة وتفاصيل البيانات.

رغم ذلك فقد اجتهد بعض العلماء في تحديد قاعدة نظرية لإيجاد عدد الفئات، ومنهم العالم ستورجس (Sturges) الذي وضع قاعدة تجريبية لحساب طول الفئات، حيث تعتمد هذه القاعدة على مجال الدراسة وحجم المجتمع أو العينة.

أ− تحديد عدد الفئات K:

$$\boxed{K = 1 + 1,322\ ln(n)} \quad \text{أو} \quad \boxed{K = 1 + 3,322\ log(n)}$$

حيث: K تمثل عدد الفئات و n تمثل حجم العينة أو المجتمع

ب- تحديد أطوال الفئات C:

$$\boxed{C = \frac{E}{K}}$$

C : تمثل طول الفئة.

E: المدى العام وهو الفرق بين أكبر قيمة وأقل قيمة، أي: $\boxed{E = X_{max} - X_{min}}$

وعلى العموم فإن كل فئة تتميز بما يلي:

- كل فئة تتميز بحدين، حد أدنى وحد أقصى، هذه الحدود وبالخصوص الحد الأقصى يمكن أن يكون حدا فعليا أو غير فعلي، كأن نقول الفئة من أ إلى **ب** أي:[أ – ب] ⇐ ب يعتبر حدا فعليا أي ينتمي للفئة، أو من أ إلى أقل من **ب** أي:[أ – ب[⇐ ب يعتبر حدا غير فعلي أي لا ينتمي للفئة.
- يمكن أن نلاحظ في بعض الجداول أن الفئة الأولى في الجدول غير محددة الحد الأدنى، كأن نقول مثلا 100 فما أقل أي (≥ 100)، كما أن الفئة الأخيرة قد تكون غير محددة الحد الأقصى كأن نقول مثلا 10000 فما أكثر (≤ 10000)، وهذا النوع من الفئات يطرح إشكالا في حساب مراكز الفئات.
- كل فئة تتميز بمجال أو طول وهو : $\boxed{\Delta x_i = Lim_{sup} - Lim_{inf}}$
- كل فئة تتميز بمركز يحسب كالتالي: $\boxed{C_i = \frac{Li_{sup} + Lim_{inf}}{2}}$
- إذا كانت أطوال الفئات متساوية فإن الفروق بين المراكز تساوي أطوال الفئات.

ملاحظة هامة: مهما تكن الطريقة المستعملة في تحديد عدد الفئات وأطوالها فإن المهم في كل ذلك هو أن لا يكون هناك تداخل بين الفئات، بحيث كل قيمة في السلسلة لا يمكن وضعها إلا في فئة واحدة.

كما أن جدول التوزيع التكراري لمتغير إحصائي مستمر قد يحتوي على التكرارات التالية:

- التكرار النسبي f_i والتكرار النسبي المئوي$f_i\%$.
- التكرار التجميعي الصاعد المطلق $N_i^{\uparrow}$ والنازل المطلق $N_i^{\downarrow}$
- التكرار التجميعي الصاعد النسبي $F_i^{\uparrow}$، والصاعد النسبي المئوي $F_i^{\uparrow}\%$.
- التكرار التجميعي النازل النسبي $F_i^{\downarrow}$، والنازل النسبي المئوي $F_i^{\downarrow}\%$.

ملاحظة: يتم حساب التكرارات السابقة بنفس الطريقة المذكورة في المتغير الكمي المتقطع.

مثال(2-6):

البيانات التالية تمثل أوزان 60 طالبا بالكيلوغرام في أحد أقسام الـــ: LMD بكلية العلوم الاقتصادية وعلوم التسيير بجامعة سطيف:

67	64	68	73	73	54	61	74	60	78
80	74	65	63	60	69	72	66	77	65
74	50	76	69	68	66	78	63	70	55
67	67	64	76	61	72	72	57	65	77
59	71	79	78	58	63	74	66	73	67
61	71	69	68	73	81	64	61	84	55

المطلوب: أنشئ جدول التوزيع التكراري واحسب كلا من: $N_i^{\uparrow}$، $N_i^{\downarrow}$، f_i ، $f_i\%$، $F_i^{\uparrow}\%$، $F_i^{\downarrow}\%$ ، ثم إشرح كلا من: n_2 ، $f_2\%$ ، $N_2^{\uparrow}$ ، $N_5^{\downarrow}$، $F_2^{\uparrow}\%$، $F_5^{\downarrow}\%$

الحل: أول خطوة نقوم بها هي ترتيب البيانات السابقة ترتيبا تصاعديا :

61	60	60	59	58	57	55	55	54	50
65	64	64	64	63	63	63	61	61	61
68	67	67	67	67	66	66	66	65	65
72	72	71	71	70	69	69	69	68	68
76	74	74	74	74	73	73	73	73	72
84	81	80	79	78	78	78	77	77	76

لدينا: $C = \frac{E}{K}$

- **حساب المدى:** $E = X_{max} - X_{min} = 84 - 50 = 34$

- **حساب عدد الفئات:** $K = 1 + 3{,}322\ log(n) = 6{,}9 \approx 7$ ، أي 7 فئات.

- **حساب طول الفئة:** $K = \frac{E}{K} = \frac{34}{7} = 4{,}92\ \approx 5$

أي طول كل فئة يساوي 5 كلغ، ومنه تكون الفئات هي: [50 – 55[، [55 – 60[، ....، [80 – 85[.

نقوم بتفريغ البيانات في الفئات المذكورة فنحصل على الجدول الموالي:

الجدول(2–4): توزيع 60 طالبا حسب الوزن بالكيلوغرام في أحد أقسام الـ: LMD بكلية العلوم الاقتصادية – سطيف

$F_i^{\downarrow}\%$	$F_i^{\uparrow}\%$	$N_i^{\downarrow}$	$N_i^{\uparrow}$	$f_i\%$	f_i	C_i	n_iعدد الطلبة	X_iأوزان الطلبة
100	3,3	60	2	3,3	0,033	52,5	2	]55 – 50]
96,7	11,6	58	7	8,3	0,083	57,5	5	]60 – 55]
88,4	31,6	53	19	20	0,2	62,5	12	]65 – 60]
68,4	58,4	41	35	26,8	0,268	67,5	16	]70 – 65]
41,6	81,7	25	49	23,3	0,233	72,5	14	]75 – 70]
18,3	95	11	57	13,3	0,133	77,5	8	]80 – 75]
5	100	3	60	5	0,05	82,5	3	]85 – 80]
/	/	/	/	**100**	**1**	/	$\sum n_i = 60$	**المجموع**

الشرح:

$n_2 = 5$: هناك 5 طلبة من بين 60 طالبا أوزانهم تتراوح بين 55 و 60 كلغ.

$f_2\% = 8{,}3\%$: هناك 8,3% من الطلبة أوزانهم تتراوح مابين 55 و 60 كلغ.

$N_2^{\uparrow} = 7$: هناك 7 طلبة من بين 60 طالبا أوزانهم أقل من 60 كلغ.

$N_5^{\downarrow} = 25$: هناك 25 طالبا من بين 60 طالبا أوزانهم أكبر أو يساوي 70 كلغ.

$F_2^{\uparrow}\% = 11{,}6\%$: هناك 11,6% من الطلبة أوزانهم أقل من 60 كلغ.

$F_5^{\downarrow}\% = 41{,}6\%$: هناك 41,6% من الطلبة أوزانهم أكبر أو تساوي 70 كلغ.

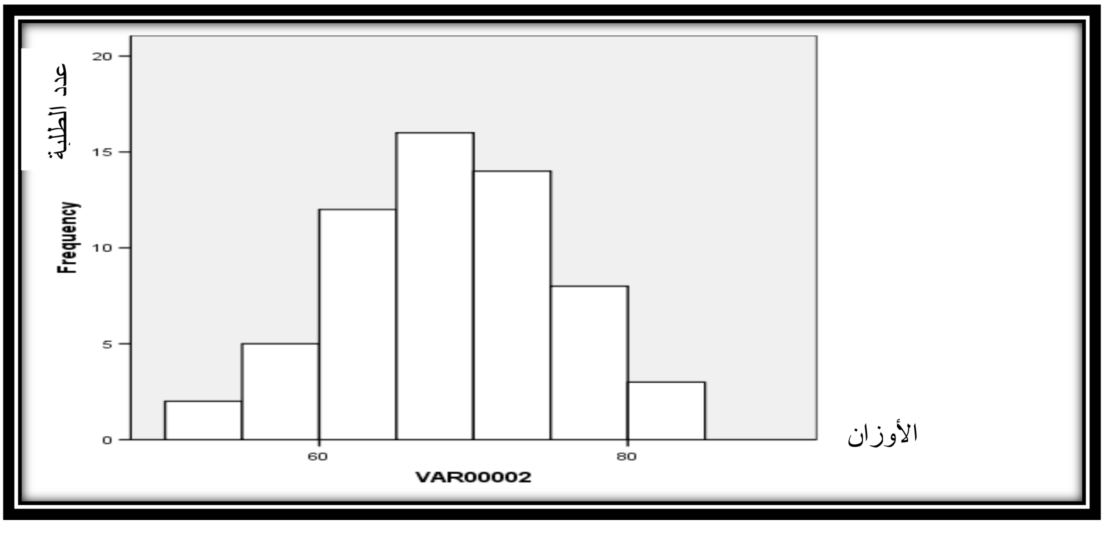
2- التمثيل البياني للمتغير الإحصائي المستمر:

أ- يمثل التكرار المطلق والنسبي للمتغير الإحصائي المستمر عن طريق **المدرج التكراري** حيث تتناسب مساحة المستطيل مع التكرار المطلق أو النسبي الموافق له، إذا ربطنا مراكز الفئات بواسطة خطوط مستقيمة مع بعضها البعض نحصل على **المضلع التكراري**، وإذا رسمنا منحنى بجوار المضلع التكراري نحصل على **المنحنى التكراري**.

مثال(2-7):

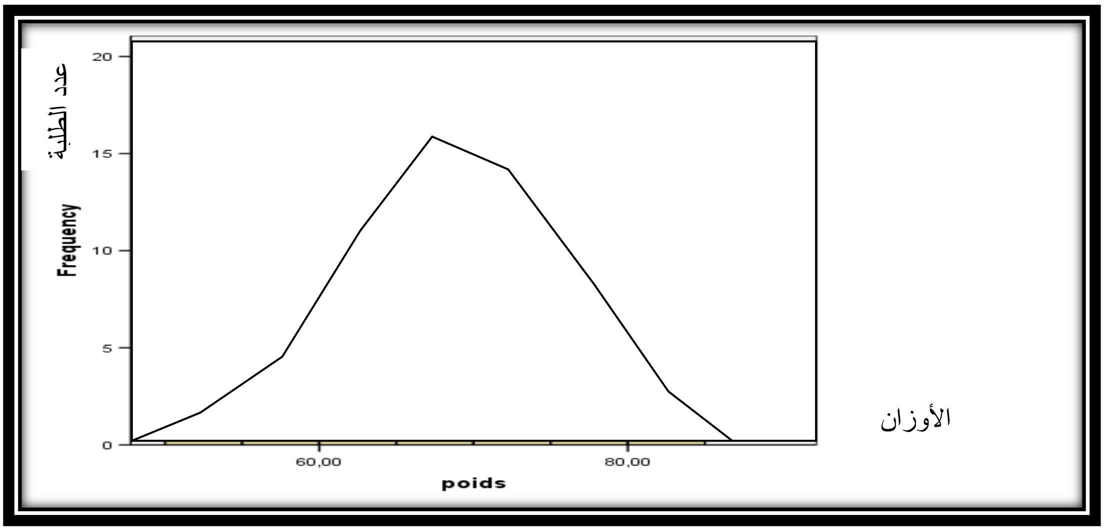
التمثيل البياني للمثال رقم (2-6):

- بواسطة المدرج التكراري:



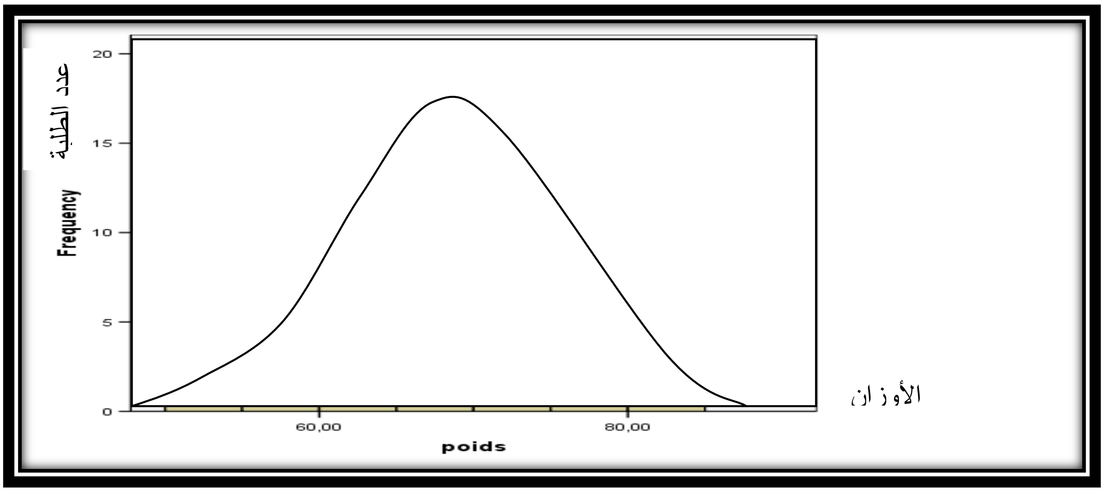
الشكل(2-3): تمثيل بياني بواسطة المدرج التكراري لتوزيع الطلبة الجامعة حسب أوزانهم

- بواسطة المظلع التكراري:



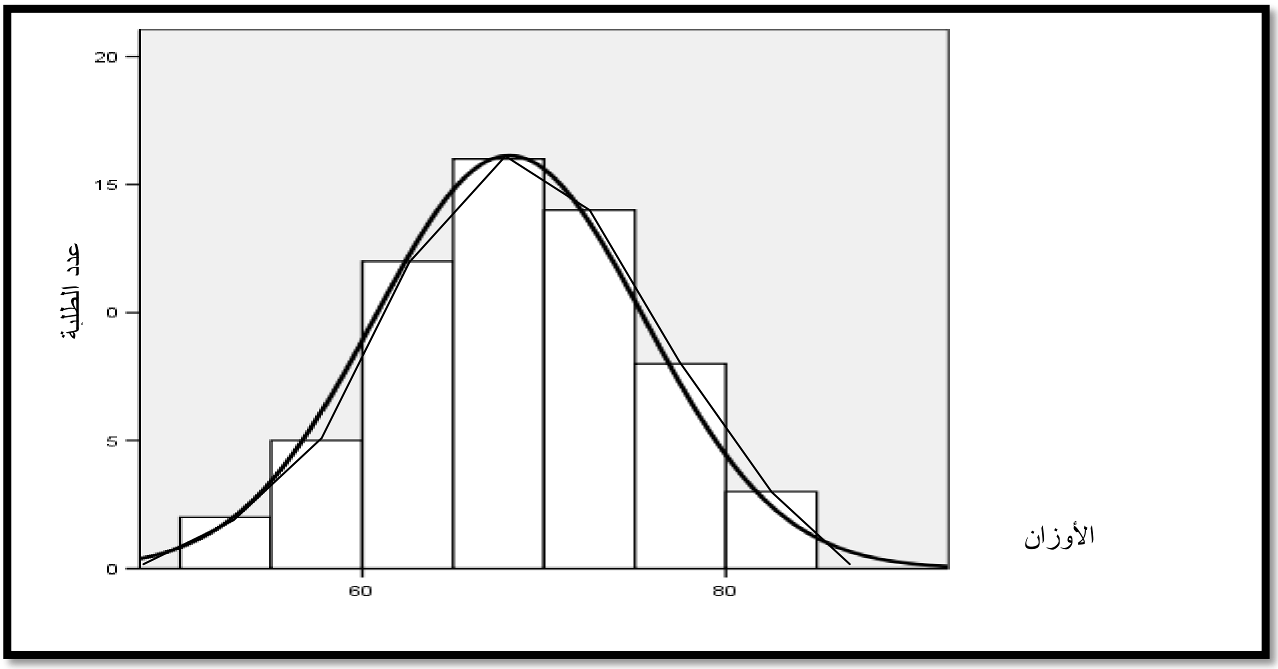
الشكل(2-4): تمثيل بياني بواسطة المظلع التكراري لتوزيع الطلبة الجامعة حسب أوزانهم

- بواسطة المنحنى التكراري:



الشكل(2-5): تمثيل بياني بواسطة المنحنى التكراري لتوزيع الطلبة الجامعة حسب أوزانهم

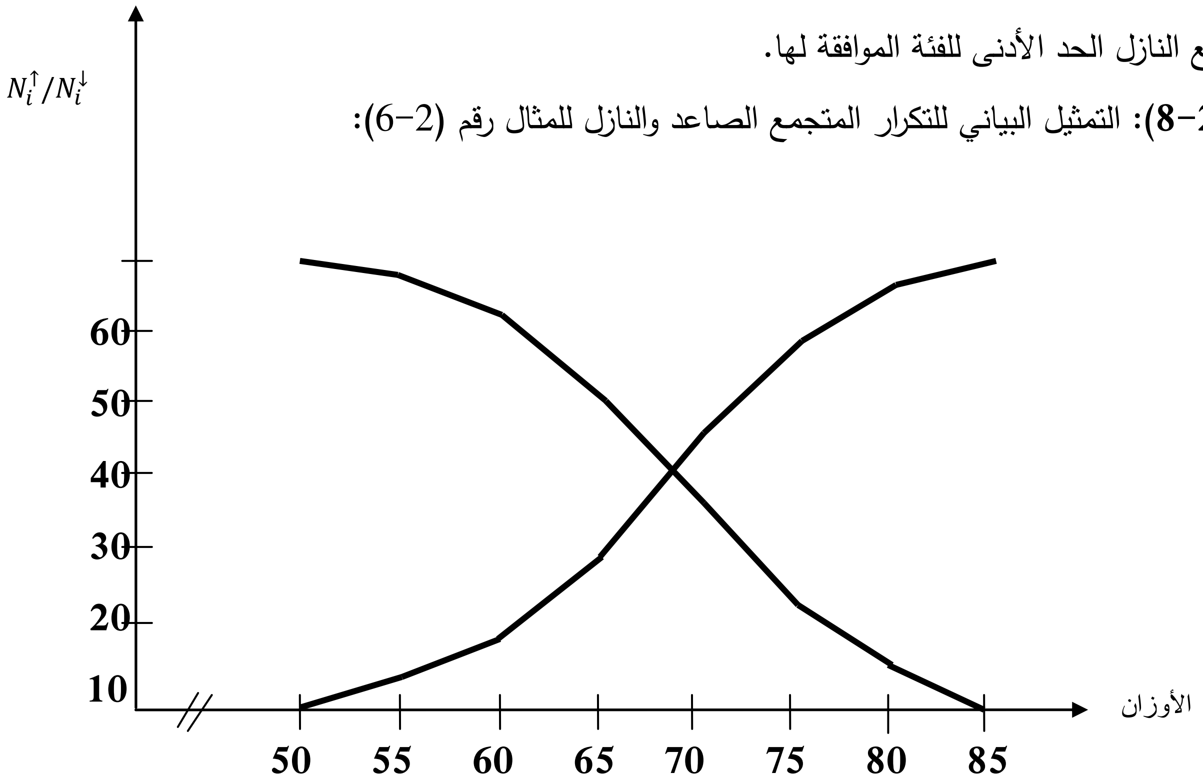
كما يمكن عرض الأشكال الثلاثة على نفس الرسم:



الشكل(2–6): تمثيل بياني بواسطة المدرج التكراري، المظلع التكراري، المنحنى التكراري لتوزيع الطلبة الجامعة حسب أوزانهم

ب– يمثل التكرار المتجمع الصاعد والنازل للمتغير الإحصائي المستمر عن طريق **المنحنى المتجمع الصاعد والنازل**، حيث نرفق بكل قيمة للتكرار المتجمع الصاعد الحد الأعلى للفئة الموافقة لها ونرفق بكل قيمة للتكرار المتجمع النازل الحد الأدنى للفئة الموافقة لها.

مثال(2–8): التمثيل البياني للتكرار المتجمع الصاعد والنازل للمثال رقم (2–6):



الشكل(2–7): المنحنى المتجمع الصاعد والنازل المطلق لتوزيع الطلبة الجامعة حسب أوزانهم

3- التمثيل البياني للتوزيع التكراري لمتغير إحصائي متصل في حالة عدم تساوي أطوال الفئات (طريقة التصحيح):

إذا كانت الفئات غير متساوية، نقوم بتعديل التكرارات لأن قاعدة المقارنة غير ثابتة، حتى يكون هناك تناسب بين طول الفئة والتكرار المقابل لها، أي إيجاد عدد الوحدات الإحصائية الموزعة على وحدة قياس معينة.

التكرار المعدل: هو عبارة عن النسبة بين التكرار البسيط وطول الفئة المقابلة مضروبا في الطول الشائع.

$$n_i' = \frac{n_i}{a_i} \times a$$ أو $$f_i' = \frac{f_i}{a_i} \times a$$ أو $$f_i'\% = \frac{f_i\%}{a_i} \times a$$

حيث:

n_i' : تمثل التكرار المطلق المعدل. f_i' : تمثل التكرار النسبي المعدل.

$f_i'\%$: تمثل التكرار النسبي المئوي المعدل. a_i : تمثل طول كل فئة.

a : تمثل الطول الشائع.

n_i ، f_i ، $f_i\%$: تمثل بهذا الترتيب التكرار المطلق، التكرار النسبي، التكرار النسبي المئوي.

مثال(2-9): يبين التوزيع التالي توزيع 100 عامل حسب الأجر اليومي، مثله بيانيا؟

الجدول(2-5): توزيع 100 عامل حسب الأجر اليومي

فئات الأجور	التكرار البسيط n_i
$[20 - 25[$	5
$[25 - 35[$	15
$[35 - 40[$	20
$[40 - 55[$	25
$[55 - 75[$	30
$[75 - 80[$	5
مجموع التكرارات	**100**

الحل: نلاحظ أن أطوال الفئات غير متساوية، وعليه يجب إجراء التصحيح على التكرار n_i قبل التمثيل البياني:

فئات الأجور	التكرار البسيط n_i	طول الفئة a_i	$\frac{n_i}{a_i}$	$n_i' = \frac{n_i}{a_i} \times a$
$[20 - 25[$	5	5	1	5
$[25 - 35[$	15	10	1,5	7,5
$[35 - 40[$	20	5	4	20
$[40 - 55[$	25	15	1,66	8,33
$[55 - 75[$	30	20	1,5	7,5
$[75 - 80[$	5	5	1	5
مجموع التكرارات	**100**	/	/	/

الطول الشائع هو 5، أي $a = 5$

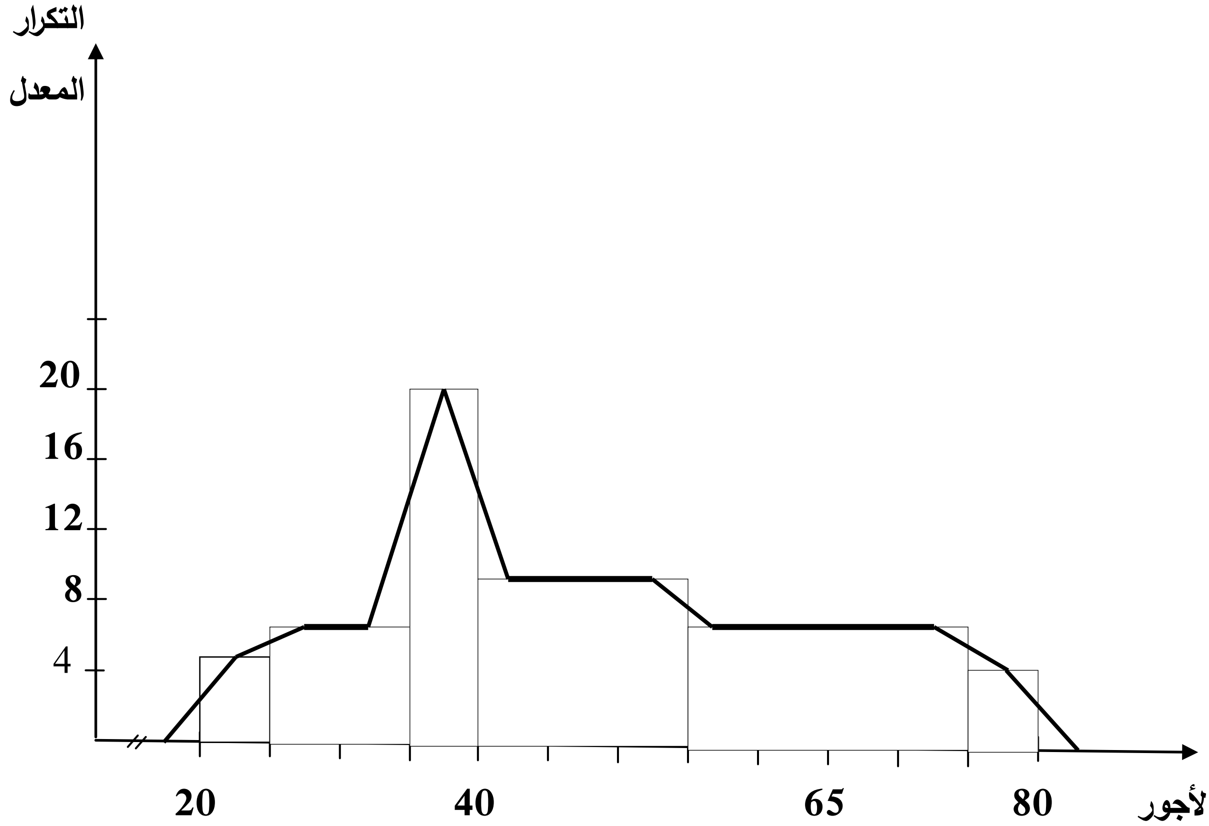
ملاحظة:

نقوم بتعديل التكرارات في حالة الفئات غير المتساوية في حالتين:

1- عند رسم المدرج التكراري.

2- عند تحديد الفئة المنوالية وحساب المنوال.

التمثيل البياني:



الشكل(2-8): توزيع 100 عامل حسب الأجر اليومي

ثالثا- التوزيع التكراري للمتغير الإحصائي الكيفي.

1- التوزيع التكراري للمتغير الإحصائي الكيفي القابل للترتيب وتمثيله البياني:

1-1- جدول التوزيع التكراري:

إذا كان المتغير المدروس كيفيا قابلا للترتيب فإن جدول التوزيع التكراري يحتوي على أنواع المتغير في العمود الأول والتكرار المطلق n_i وكذلك التكرار النسبي والنسبي المئوي، وكذلك التكرار المتجمع الصاعد والنازل المطلق والنسبي.

1-2- التمثيل البياني للمتغير الكيفي القابل للترتيب

تمثل التكرارات المطلقة للمتغير الكيفي القابل للترتيب عن طريق **المستطيل المئوي**، حيث تتاسب مساحة كل جزء من المستطيل التكرار المطلق أو النسبي الموافق له.

مثال(2-10):

يمثل الجدول التالي توزيع عينة من 50 فرد حسب المستوى التعليمي، والمطلوب حساب كلا من: $N_i^{\uparrow}$، $N_i^{\downarrow}$، f_i ، $f_i\%$، $F_i^{\uparrow}\%$، $F_i^{\downarrow}\%$ ، ثم إشرح كلا من: n_2 ، $N_2^{\uparrow}$ ، $N_4^{\downarrow}$، $F_2^{\uparrow}\%$، $F_4^{\downarrow}\%$ ، مع التمثيل البياني لهذا التوزيع؟

الجدول(2-6): توزيع عينة من 50 فرد حسب المستوى التعليمي

المستوى التعليمي	n_i
إبتدائي	4
متوسط	14
ثانوي	20
جامعي (تدرج)	10
دراسات عليا	2
المجموع	50

الحل:

المستوى التعليمي	n_i	f_i	$f_i\%$	$N_i^{\uparrow}$	$N_i^{\downarrow}$	$F_i^{\uparrow}$	$F_i^{\downarrow}\%$
إبتدائي	4	0,08	8	4	50	8	100
متوسط	14	0,28	28	18	46	36	92
ثانوي	20	0,40	40	38	32	76	64
جامعي (تدرج)	10	0,20	20	48	12	96	34
دراسات عليا	2	0,04	4	50	2	100	4
المجموع	50	1,00	100	/	/	/	/

الشرح:

$n_2 = 14$: هناك 14 فردا من بين 50 فردا مستواهم العلمي هو المستوى المتوسط.

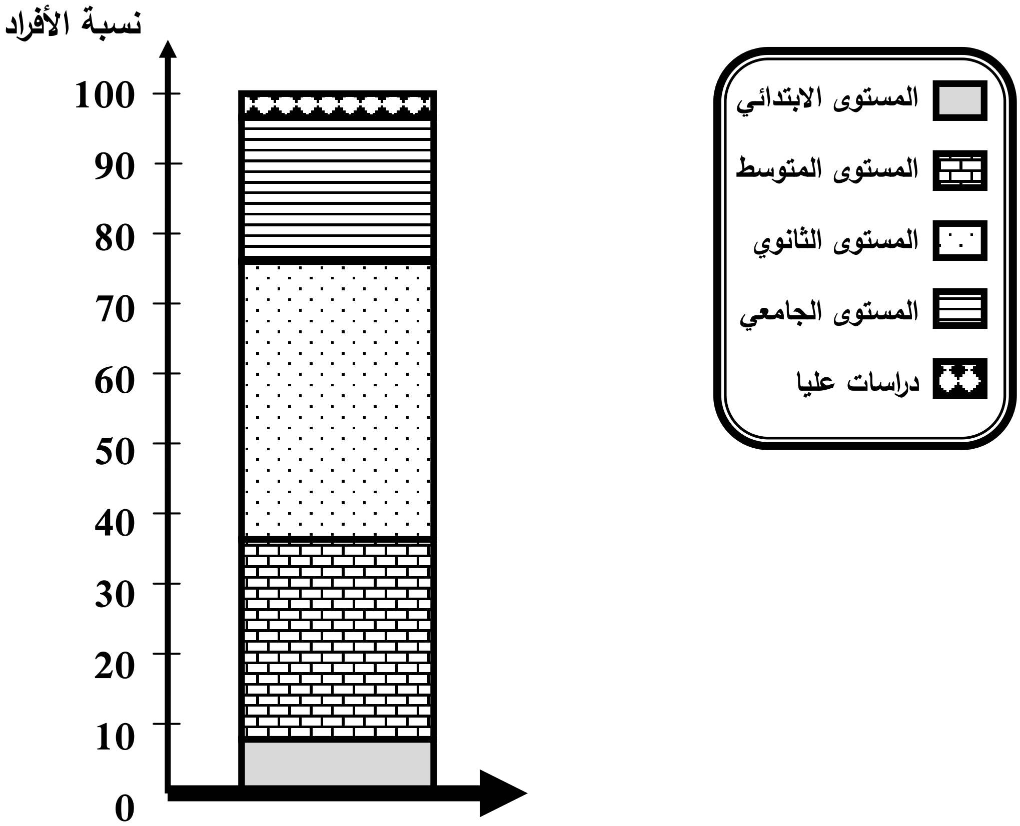
$N_2^{\uparrow} = 18$: هناك 18 فردا من بين 50 فردا مستواهم العلمي على الأكثر هو المستوى المتوسط.

$N_4^{\downarrow} = 12$: هناك 12 فردا من بين 50 فردا مستواهم العلمي على الأقل هو المستوى الجامعي.

$F_2^{\uparrow}\% = 36\%$: هناك 36% من الأفراد مستواهم العلمي على الأكثر هو المستوى المتوسط.

$F_4^{\downarrow}\% = 34\%$: هناك 34% من الأفراد مستواهم العلمي على الأقل هو المستوى الجامعي.

- التمثيل البياني



الشكل(2-9): توزيع عينة من 50 فرد حسب المستوى التعليمي

2- التوزيع التكراري للمتغير الإحصائي الكيفي الغير قابل للترتيب وتمثيله البياني

2-1- جدول التوزيع التكراري:

إذا كان المتغير المدروس كيفيا غير قابل للترتيب فإن جدول التوزيع التكراري يحتوي على أنواع المتغير في العمود الأول والتكرار المطلق n_i وكذلك التكرار النسبي والنسبي المئوي، أما التكرار المتجمع الصاعد والنازل المطلق والنسبي فليس له معنى.

2-2- التمثيل البياني للمتغير الكيفي الغير قابل للترتيب:

تمثل التكرارات المطلقة للمتغير الكيفي الغير قابل للترتيب عن طريق **الدائرة النسبية**، حيث يتناسب قيس كل زاوية مع التكرار المطلق أو النسبي الموافق له.

مثال(2-11): يمثل الجدول التالي توزيع عينة من 40 فرد من الجالية المغاربية في فرنسا حسب البلد الأصلي والمطلوب حساب كلا من $\boldsymbol{f_i}$ ، $\boldsymbol{f_i}\%$ مع شرح n_2 و n_4 و $f_2\%$ ؟.

الجدول(2-7): توزيع عينة من 40 فرد من الجالية المغاربية في فرنسا حسب البلد الأصلي

البلد الأصلي	المغرب	الجزائر	تونس	ليبيا	المجموع
عدد الأفراد	12	16	9	3	**40**

الحل:

البلد الأصلي	n_i	f_i	$f_i\%$	الزاوية المركزية
المغرب	12	0,3	30	°108
الجزائر	16	0,4	40	°144
تونس	9	0,225	22,5	°81
ليبيا	3	0,075	7,5	°27
المجموع	40	1,00	100	°360

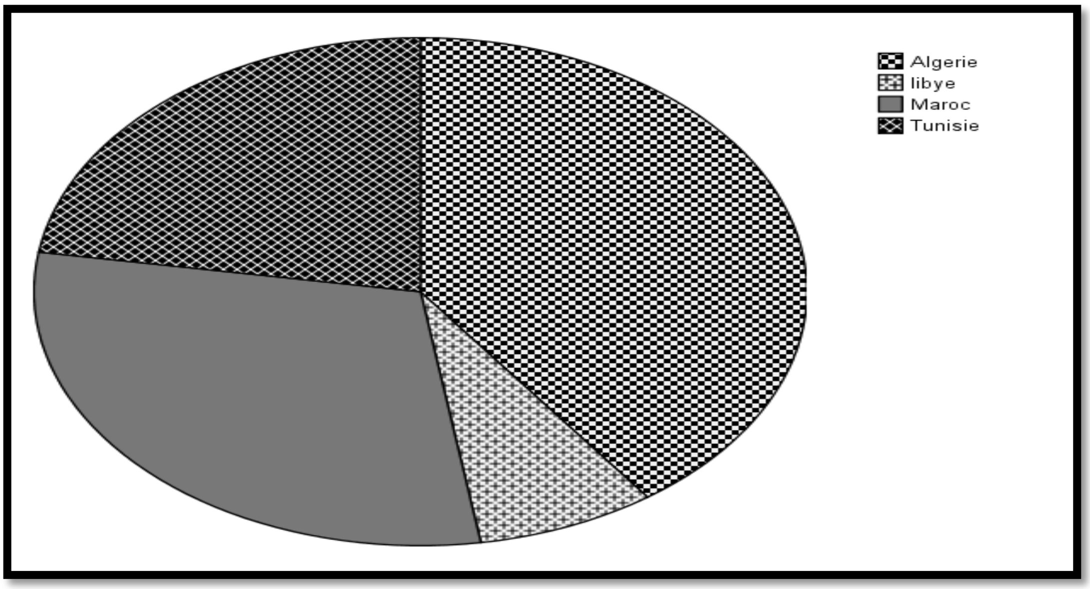
الشرح:

$n_2 = 16$: هناك 16 فردا من بين 40 فردا من الجالية المغاربية في فرنسا من أصول جزائرية.

$n_4 = 3$: هناك 3 أفراد من بين 40 فردا من الجالية المغاربية في فرنسا من أصول ليبية.

$f_2\% = 40\%$: هناك 40% من أفراد الجالية المغاربية في فرنسا من أصول جزائرية.

- التمثيل البياني:



الشكل(2−10): توزيع عينة من 40 فرد من الجالية المغاربية في فرنسا حسب البلد الأصلي

رابعا- دراسة قضية التمركز:

دراسة قضية التمركز هي طريقة إحصائية تطبق على التوزيعات التكرارية (الجداول) التي يكون موضوعها توزيع ثروة معينة (دخل، أجور، أراضي...إلخ)، على فئات متفاوتة (فئات من الأسر، فئات العمال، فئات من المزارعين...إلخ)، لكي نعرف هل هناك تمركز في توزيع هذه الثروة أي هل هناك توزيع عادل أو غير عادل لهذه الثروة على مختلف هذه الفئات، فهي إذن تجيب على سؤالين:

1- هل هناك فوارق أو تمركز في توزيع هذه الثروة على مختلف فئات المجتمع الإحصائي؛

2- إذا كانت هناك فوارق فبكم تقدر، أي إعطاء تقييما كميا لهذه الفوارق والحكم عليها هل هي فوارق كبيرة، كبيرة جدا أو معتدلة (يتم ذلك بواسطة نسبة كما سنبين لاحقا).

وقضية التمركز لا تطبق إلا على المتغيرات الإحصائية المستمرة ذات القيم الموجبة، لذا تبدو أهميته واضحة في الظواهر الاقتصادية كلها، ويعتبر توزيع الدخل القومي على المواطنين من أكثر التطبيقات العملية لظاهرة التمركز.

تدرس قضية التمركز بطريقتين:

- طريقة بيانية: بواسطة مربع جيني للتمركز؛
- طريقة حسابية بواسطة معامل جيني للتمركز I_{Gini}

نعرض الطريقتين من خلال المثال التالي:

مثال(2-12):

الجدول التالي يبين الدخل الشهري لــ 43 أسرة بالدينار الجزائري:

الجدول(2-8): الدخل الشهري لــــ 43 أسرة بالدينار الجزائري

الدخل الشهري	عدد الأسر
]4000 – 2000]	5
]6000 – 4000]	8
]8000 – 6000]	12
]10000 – 8000]	10
]12000 – 10000]	8
المجموع	43

المطلوب: دراسة قضية التمركز؟

الحــــــل:

لدراسة التمركز نتبع الخطوات التالية:

الخطوة الأولى: إنشاء جدول التوزيع التكراري الخاص بالتمركز(تحضير الحسابات).

1- حساب التكرار النسبي f_i

2- حساب التكرار النسبي المتجمع الصاعد $F_i^{\uparrow}$

3- حساب مراكز الفئات C_i

4- حساب المقادير : $n_i \times c_i$

5- حساب النسب: $s_i = \frac{n_i\, c_i}{\sum n_i c_i}$

6- حساب النسبة المتجمعة الصاعدة $S_i^{\uparrow}$ بنفس طريقة حساب التكرار النسبي المتجمع الصاعد $F_i^{\uparrow}$

7- مقارنة بعض القيم لـ: s_i و f_i ، وإعطاء قراءتها الإقتصادية (بالخصوص الفئتين الأولى والأخيرة).

بالرجوع إلى المثال(2-12):

الدخل الشهري	c_i	n_i	f_i	$.F_i^{\uparrow}$	$n_i\, c_i$	s_i	$.S_i^{\uparrow}$
]4000 – 2000]	3000	5	0,1163	**0,1163**	15000	0,0473	**0,0473**
]6000 – 4000]	5000	8	0,1860	**0,3023**	40000	0,1262	**0,1735**
]8000 – 6000]	7000	12	0,2791	**0,5814**	84000	0,2650	**0,4380**
]10000 – 8000]	9000	10	0,2326	**0,814**	90000	0,2839	**0,7319**
]12000 – 10000]	11000	8	0,1860	**1**	88000	0,2776	**1**
المجموع	/	43	1	/	317000	1	/

مقارنة بعض القيم لـ: s_i و f_i :

الفئة الأولى: 11,63% من الأسر يحصلون فقط على 4,73% من الدخل الشهري الإجمالي.

الفئة الأخيرة: 18,6% من الأسر يحصلون لوحدهم على 27,76% من الدخل الشهري الإجمالي.

* نلاحظ أن هناك فوارق في توزيع الدخل الشهري الإجمالي.

الخطوة الثانية: دراسة التمركز بيانيا (مربع جيني)

لرسم مربع جيني نتبع الخطوات التالية:

1- نرسم مربع طول ضلعه 100 ملم موجه على معلم متعامد ومتجانس.

2- تمثل على محور الفواصل $F_i^{\uparrow}$ و على محور التراتيب $S_i^{\uparrow}$.

3- تعيين المثلث ABC على المربع (أنظر الشكل (2-11).

4- ننقل النقاط ($S_i^{\uparrow}$ ،$F_i^{\uparrow}$) إلى الشكل، مثلا النقطة الأولى (0,1163 ، 0,0473).

5- نربط بين النقاط بمنحنى منكسر.

6- نعرف بعض المفاهيم على الشكل (خط العدالة المطلقة، منحنى لورونز، منطقة التمركز، المساحات خارج منطقة التمركز).

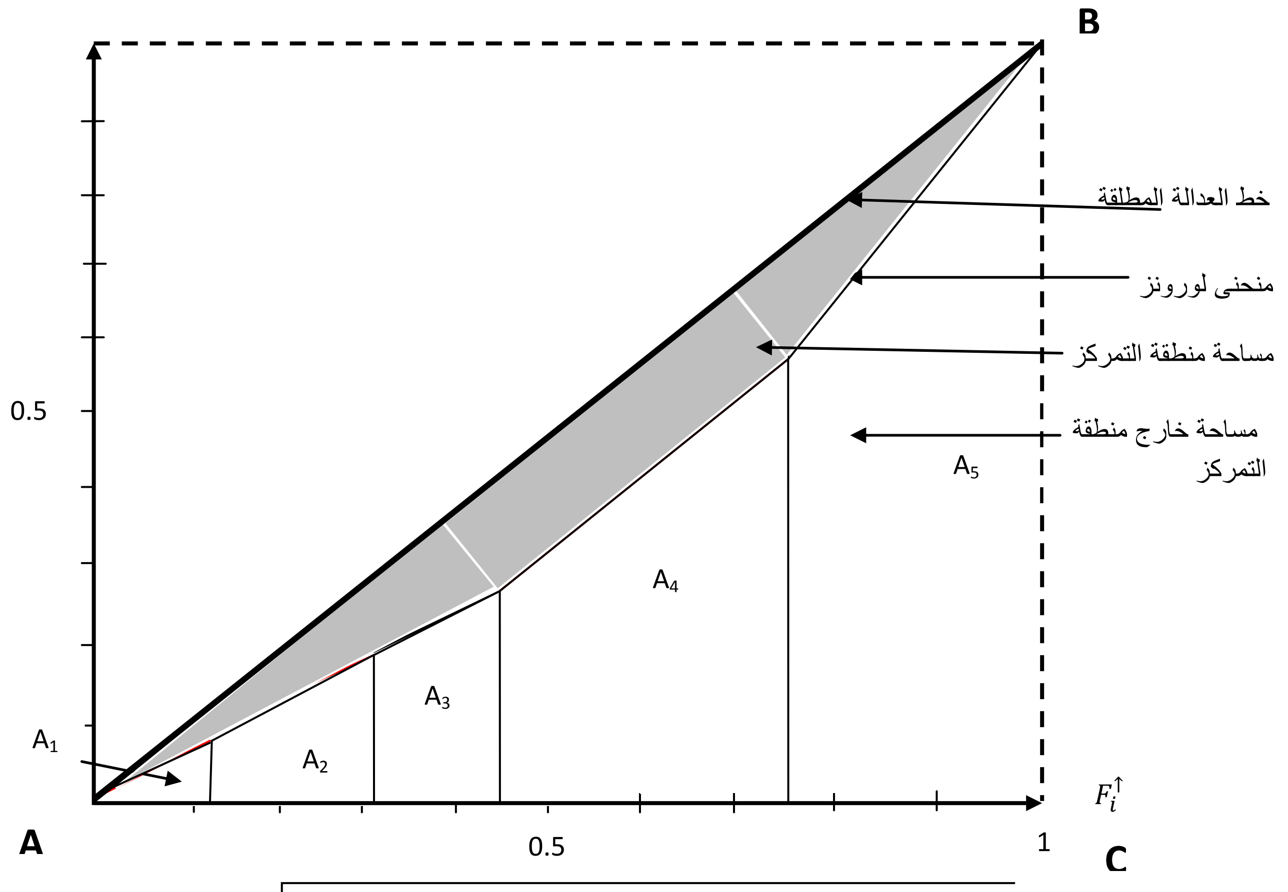
إن شكل منحنى لورونز يأخذ الحالات التالية:

- إذا انطبق منحنى لورونز على المنصف (AB) - خط العدالة (تنعدم مساحة التمركز) - فإننا نقول أن توزيع الثروة عادل تماما(إنعدام التمركز).

- إذا انطبق منحنى لورونز على المثلث (ABC)(مساحة التمركز تساوي مساحة المثلث) فإننا نقول أن توزيع الثروة جائز تماما(تمركز كلي).

- كلما اقترب منحنى لورونز من خط العدالة (AB)(كانت مساحة التمركز صغيرة مقارنة بمساحة المثلث ABC) كلما كان التوزيع التكراري أكثر عدالة، وكلما ابتعد منحنى لورونز من خط العدالة (AB) (كانت مساحة التمركز كبيرة مقارنة بمساحة المثلث ABC) كلما كان التوزيع التكراري أقل عدالة.

بالرجوع إلى المثال(2-12) نقوم برسم مربع جيني كما يلي:



الشكل(2-11): مربع جيني لتوزيع 43 عامل حسب الأجر الشهري

***التعليق على الشكل:**

نراقب الشكل المحصل عليه بالعين المجردة، نلاحظ أن منحنى لورونز يقترب من خط العدالة وبالتالي فإن مساحة التمركز تبدو صغيرة مقارنة بمساحة المثلث ABC ومنه فإن التمركز يبدو ضعيفا أي أكثر عدالة.

الخطوة الثالثة: دراسة التمركز حسابيا (حساب معامل جيني للتمركز)

إن معامل جيني للتمركز هو عبارة عن حاصل قسمة مساحة التمركز (S) المحصورة بين منحنى لورونز وخط العدالة (AB) على مساحة المثلث (ABC)، ومنه:

– إذا كانت النسب نسبية فإن:

$$\boxed{I_{Gini} = \frac{S}{\text{مساحة المثلث } ABC} = \frac{S}{1/2} = 2S \ldots\ldots\ldots\ldots (1)}$$

$$S = \frac{1}{2} - (A_1 + A_2 + \cdots + A_K)$$

حيث A_1، A_2، ...، A_K تمثل مساحات أشباه المنحرف تحسب كلا منها وفق القاعدة التالية:

$$A_i = \frac{\text{القاعدة الصغرى} + \text{القاعدة الكبرى}}{2} \times \text{الإرتفاع}$$

$$S = \frac{1}{2} - \left(\frac{S_1^{\uparrow} \times f_1}{2} + \frac{(S_2^{\uparrow} + S_1^{\uparrow})f_2}{2} + \frac{(S_3^{\uparrow} + S_2^{\uparrow})f_3}{2} + \cdots + \frac{(S_k^{\uparrow} + S_{k-1}^{\uparrow})f_k}{2}\right)$$

$$S = \frac{1}{2} - \left(\frac{(S_1^{\uparrow} + S_0^{\uparrow})f_1}{2} + \frac{(S_2^{\uparrow} + S_1^{\uparrow})f_2}{2} + \frac{(S_3^{\uparrow} + S_2^{\uparrow})f_3}{2} + \cdots + \frac{(S_k^{\uparrow} + S_{k-1}^{\uparrow})f_k}{2}\right)$$

حيث: $S_0^{\uparrow} = 0$

$$S = \frac{1}{2} - \frac{1}{2}\sum_{i=1}^{k} f_i (S_i^{\uparrow} + S_{i-1}^{\uparrow})$$

بالتعويض في العلاقة (1) نجد:

$$I_{Gini} = 2\left(\frac{1}{2} - \frac{1}{2}\sum_{i=1}^{k} f_i (S_i^{\uparrow} + S_{i-1}^{\uparrow})\right)$$

$$\boxed{I_{Gini} = 1 - \sum_{i=1}^{k} f_i (S_i^{\uparrow} + S_{i-1}^{\uparrow})}$$

وهي علاقة تقريبية لحساب هذا المعامل.

حيث أن:

f_i: تمثل التكرار النسبي، $S_i^{\uparrow}$: تمثل النسبة المتجمعة الصاعدة و $S_0^{\uparrow} = 0$.

إن الحالات التي يأخذها معامل جيني للتمركز هي:

$I_{Gini} = 0$: هذا يعني أن منحنى لورونز ينطبق على خط العدالة وبالتالي فالتوزيع عادل تماما.

$I_{Gini} = 1$: هذا يعني أن منحنى لورونز ينطبق على المثلث (ABC) وبالتالي نقول أن التوزيع جائر تماما.

$I_{Gini} \leq 0{,}3$: هذا يعني أن منحنى لورونز قريب من خط العدالة، وبالتالي فالتوزيع أكثر عدالة.

$I_{Gini} > 0{,}3$: هذا يعني أن منحنى لورونز بعيد عن خط العدالة، وبالتالي فالتوزيع أقل عدالة.

ملاحظة: القيمة 0,3 هي قاعدة تطبيقية للحكم على قوة التمركز من إقتراح بعض الخبراء، كما يعتمد في بعض المراجع على معيار آخر وهو 0,2

بالرجوع إلى المثال(2−12) نقوم بحساب معامل جيني كما يلي:

$f_i(S_i^{\uparrow}+S_{i-1}^{\uparrow})$	$S_i^{\uparrow}+S_{i-1}^{\uparrow}$	$S_i^{\uparrow}$	f_i
0,0055	0,0473	0,0473	0,1163
0,0411	0,2208	0,1735	0,1860
0,1687	0,6043	0,4380	0,2791
0,2721	1,1699	0,7319	0,2326
0,3221	1,7319	1	0,1860
0,8095	/	/	1

$$I_{Gini} = 1 - \sum_{i=1}^{k} f_i\left(S_i^{\uparrow} + S_{i-1}^{\uparrow}\right)$$
$$I_{Gini} = 1 - 0,8095$$
$$I_{Gini} = 0,1905$$

بما أن معامل جيني للتمركز أقل من 0,3 فإن توزيع الدخل الشهري للأسر أكثر عدالة.

– إذا كانت النسب مئوية فإن: $\boxed{I_{Gini} = \frac{S}{ABC} = \frac{S}{5000}}$

$$S = 5000 - (A_1 + A_2 + \cdots + A_K)$$

وبإتباع الطريقة السابقة نحصل على:

$$\boxed{I_{Gini} = 1 - \frac{1}{10000}\sum_{i=1}^{k} f_i\%(S_i^{\uparrow}\% + S_{i-1}^{\uparrow}\%)}$$

إن الحالات التي يأخذها معامل جيني للتمركز في هذه الحالة هي:

$I_{Gini} = 0$: هذا يعني أن منحنى لورونز ينطبق على خط العدالة وبالتالي فالتوزيع عادل تماما.

$I_{Gini} = 1$: هذا يعني أن منحنى لورونز ينطبق على المثلث (ABC) وبالتالي نقول أن التوزيع جائر تماما.

$I_{Gini} \leq 30\%$: هذا يعني أن منحنى لورونز قريب من خط العدالة، وبالتالي فالتوزيع أكثر عدالة.

$I_{Gini} > 0,3\%$: هذا يعني أن منحنى لورونز بعيد عن خط العدالة، وبالتالي فالتوزيع أقل عدالة.

تمارين محلولة للفصل الثاني

التمرين الأول:

لقد تبين من خلال الدراسة الميدانية حول ميزانية العائلات الجزائرية التي أجراها الديوان الوطني للإحصاء ما بين مارس 1979 ومارس 1980 أن مصاريف الاستهلاك للعائلة الجزائرية يتوزع كما يلي: الوحدة: %

الجدول(2−9): مصاريف الاستهلاك للعائلة الجزائرية ما بين مارس 1979 ومارس 1980

الرقم	نوع المصاريف	النسبة المئوية
1	المواد الغذائية والمشروبات الغير الكحولية	%55,7
2	السكن− الإضاءة والتسخين	%5,4
3	الأثاث والأواني المنزلية	%6,4
4	اللباس والأحذية	%9,2
5	الصحة والنظافة الجسمية	%3,1
6	النقل والاتصال	%6,6
7	التعليم والثقافة والترفيه	%3,4
8	مصاريف أخرى	%10,2
/	**المجموع**	**%100**

المصدر: الديوان الوطني للإحصاء − الدليل الإحصائي السنوي 1980.

المطلوب:

1− حدد المجتمع الإحصائي، الوحدة الإحصائية، المتغير الإحصائي ونوعه في هذه الدراسة؟

2− مثل بيانيا هذه المعطيات؟

3− ما هي الخصائص العامة لمصاريف العائلة الجزائرية من خلال هذه البيانات؟

التمرين الثاني:

الجدول التالي يمثل توزيع السكان حسب الدخل السنوي للفرد في المناطق الحضرية بالجزائر.

الجدول(2−10): توزيع السكان حسب الدخل السنوي للفرد في المناطق الحضرية بالجزائر

الوحدة: %

الدخل السنوي للفرد (دج)	نسبة السكان $f_i\%$ (%)
$]1000 - 500]$	3,65
$]3000 - 1000]$	41,10
$]5000 - 3000]$	33,95
$]7000 - 5000]$	8,10
$]10000 - 7000]$	13,20

المصدر: الديوان الوطني للإحصاء بتصرف.

المطلوب:

1- مثل بيانيا معطيات هذا الجدول؟

2- أحسب التكرارات النسبية المئوية التجميعية الصاعدة والنازلة؟ مثلها بيانيا؟ اشرح معنى $f_2\%$ ، $F_2^{\uparrow}\%$ ، $F_2^{\downarrow}\%$؟

3- حدد إحداثيتي نقطة التقاطع بين المنحنى التجميعي الصاعد والمنحنى التجميعي النازل؟ اشرحها إحصائيا؟

4- حدد نسبة السكان الذين دخلهم السنوي يساوي أو يفوق 5000 دج؟ يقل عن 3000 دج؟ يفوق أو يساوي 500 ويقل عن 5000؟

التمرين الثالث:

الجدول التالي يمثل توزيع السكان وإجمالي الدخل السنوي للفرد حسب 5 فئات في المناطق الحضرية بالجزائر. الوحدة: %

الجدول(2-11): توزيع السكان وإجمالي الدخل السنوي للفرد حسب 5 فئات في المناطق الحضرية بالجزائر

الدخل السنوي للفرد (دج)	نسبة السكان $f_i\%$ (%)	نسبة الدخل $s_i\%$
أقل من 1200	3,65	0,80
]3000 − 1200]	41,10	23,80
]5000 − 3000]	33,95	34,20
]6000 − 5000]	8,10	11,50
6000 فما أكثر	13,20	29,70
المجموع	**100,00**	**100,00**

المصدر: الديوان الوطني للإحصاء بتصرف.

المطلوب:

- أدرس قضية التمركز على هذا التوزيع مبينا هل هناك عدالة أم لا في توزيع الدخل على مختلف فئات السكان:

أ- بيانيا ب- حسابيا

الحلــــــــــــــول

حل التمرين الأول:

1- تحديد: الوحدة الاحصائية، المجتمع الاحصائي، الخاصة الاحصائية ونوعها:

المجتمع الإحصائي	الوحدة الإحصائية	المتغير الإحصائي	نوعه
العائلات الجزائرية	العائلة الواحدة	مبلغ المصاريف السنوية للأسر	كمي مستمر

2- التمثيل البياني لهذه المعطيات يكون بواسطة الدائرة كما يلي:

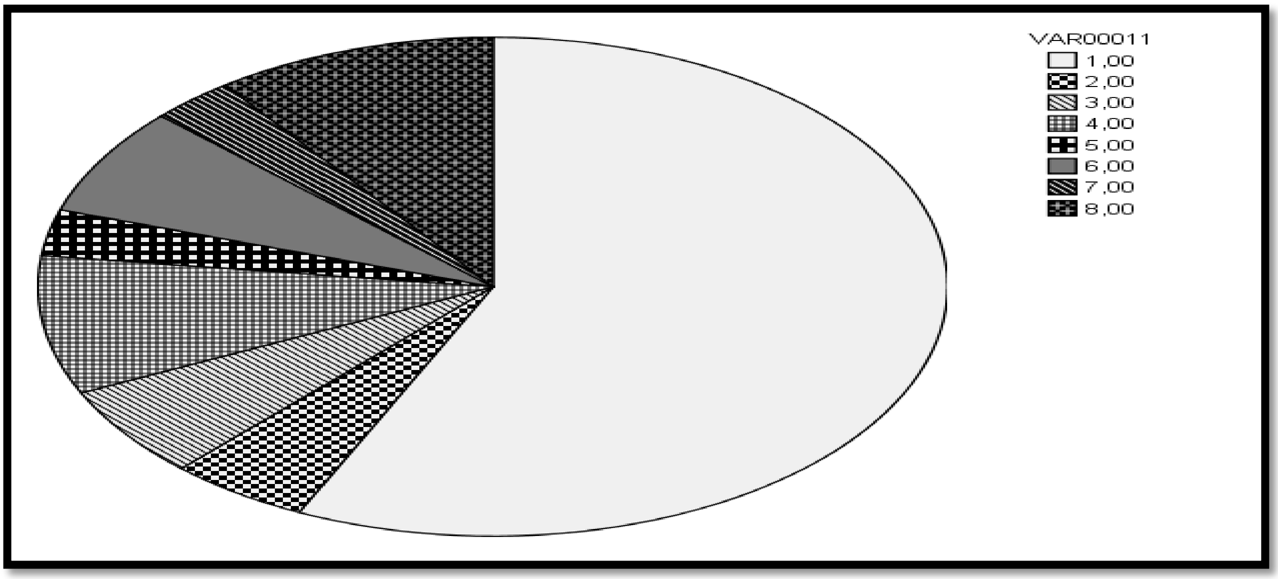
- نحول النسب المئوية إلى زوايا فمثلا المواد الغذائية والمشروبات الغير الكحولية:

$$360^{\circ} \rightarrow 100\%$$
$$\alpha \rightarrow 55,7\% \Rightarrow \alpha = 200,52^{\circ}$$

أما باقي النتائج فهي كما يلي:

الرقم	نوع المصاريف	النسبة المئوية	الدرجات
1	المواد الغذائية والمشروبات الغير الكحولية	%55,7	°200,52
2	السكن- الإضاءة والتسخين	%5,4	°19,44
3	الأثاث والأواني المنزلية	%6,4	°23,04
4	اللباس والأحذية	%9,2	°33,12
5	الصحة والنظافة الجسدية	%3,1	°11,16
6	النقل والاتصال والبريد	%6,6	°23,76
7	التعليم والثقافة والتربية	%3,4	°12,24
8	مصاريف أخرى	%10,2	°36,72
/	**المجموع**	**%100**	**°360**

- التمثيل البياني:



الشكل(2-12): توزيع مصاريف الاستهلاك للعائلة الجزائرية بين مارس1979 ومارس 1980

3- الخصائص العامة لمصاريف العائلة الجزائرية من خلال هذه البيانات:

يبين لنا الجدول أن أكثر من نصف المصاريف السنوية للعائلة الجزائرية تخصص للاستهلاك الغذائي والمشروبات غير الكحولية، بينما بقية أنواع المصاريف يخصص لها أقل من النصف (44,3%) مجتمعة، فمثلا التعليم والثقافة والترفيه لا يخصص لها سوى 3,4 % من مجموع المصاريف.

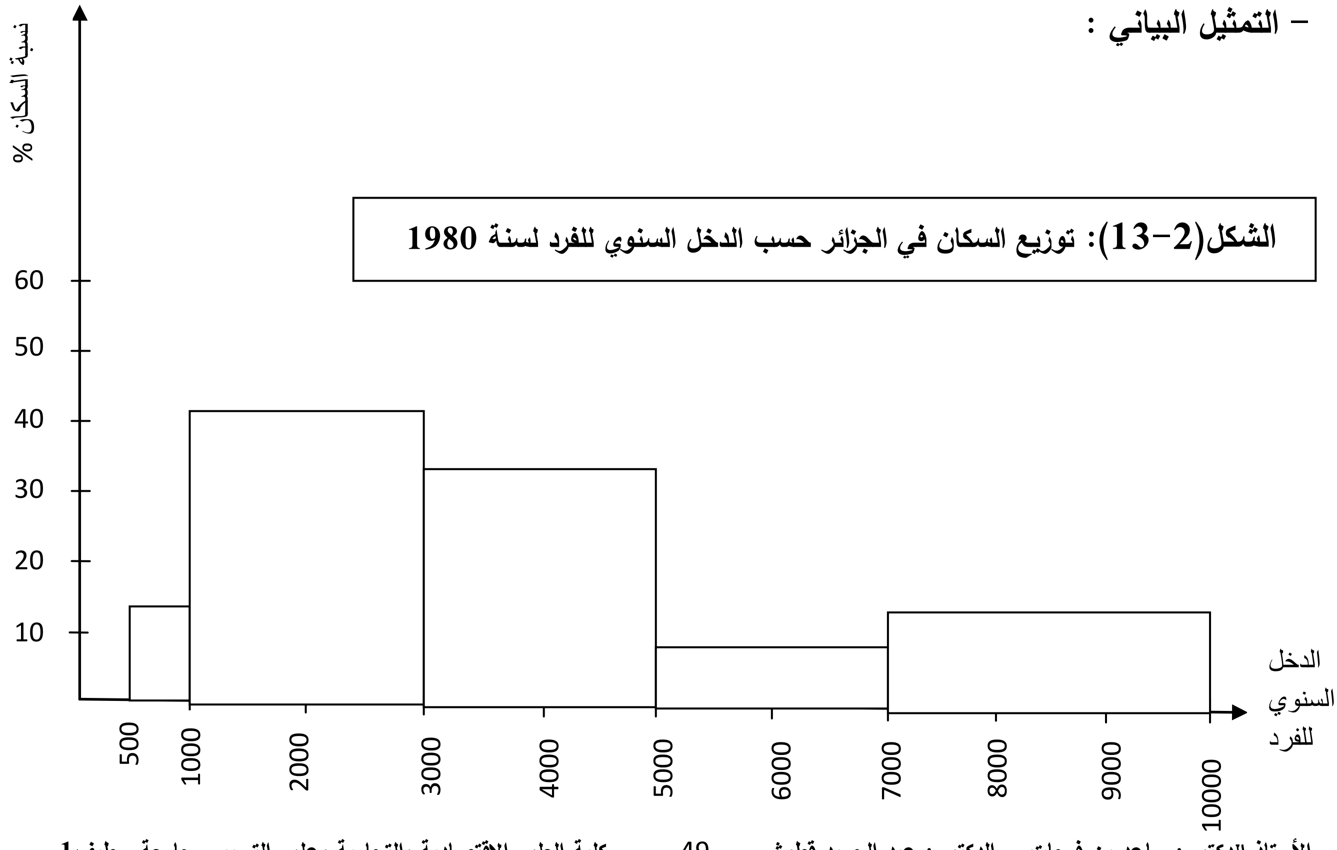
حل التمرين الثاني:

1- التمثيل البياني لمعطيات الجدول:

بما أن التوزيع في شكل فئات (المتغير مستمر والذي يمثل الدخل السنوي للفرد) فإن التمثيل البياني المناسب هو المدرج التكراري، لكننا نلاحظ أن أطوال الفئات غير متساوية وبالتالي يجب تصحيح التكرارات قبل التمثيل، الطول الشائع هو $a = 2000$.

$F_i^{\downarrow}\%$.	$F_i^{\uparrow}\%$.	$f_i\% = \frac{f_i}{a_i} \times a$	طول الفئة a_i	نسبة السكان (%) $f_i\%$	الدخل السنوي للفرد (دج)
100	3,65	14.6	500	3,65	]1000 – 500]
96,35	44,75	41.10	2000	41,10	]3000 – 1000]
55,25	78,70	33.90	2000	33,95	]5000 – 3000]
21,30	86,80	8.10	2000	8,10	]7000 – 5000]
13,20	100	8.83	3000	13,20	]10000 – 7000]
/	/	/	/	**100**	**المجموع**

- التمثيل البياني :



الشكل(2-13): توزيع السكان في الجزائر حسب الدخل السنوي للفرد لسنة 1980

2- حساب التكرارات النسبية المئوية التجميعية الصاعدة والنازلة، تمثيلها بيانيا وشرح معنى $f_2\%$ ، $F_2^{\uparrow}\%$ ،$F_2^{\downarrow}\%$:

أ- حساب التكرارات النسبية المئوية التجميعية الصاعدة والنازلة:

- التكرارات النسبية المئوية التجميعية الصاعدة $F_i^{\uparrow}\%$ تحسب بالعلاقة: $F_i^{\uparrow}\% = F_{i-1}^{\uparrow}\% + f_i\%$

- التكرارات النسبية المئوية التجميعية النازلة $F_i^{\downarrow}\%$ تحسب بالعلاقة: $F_i^{\downarrow}\% = F_{i-1}^{\downarrow}\% - f_{i-1}\%$

- أنظر الجدول السابق.

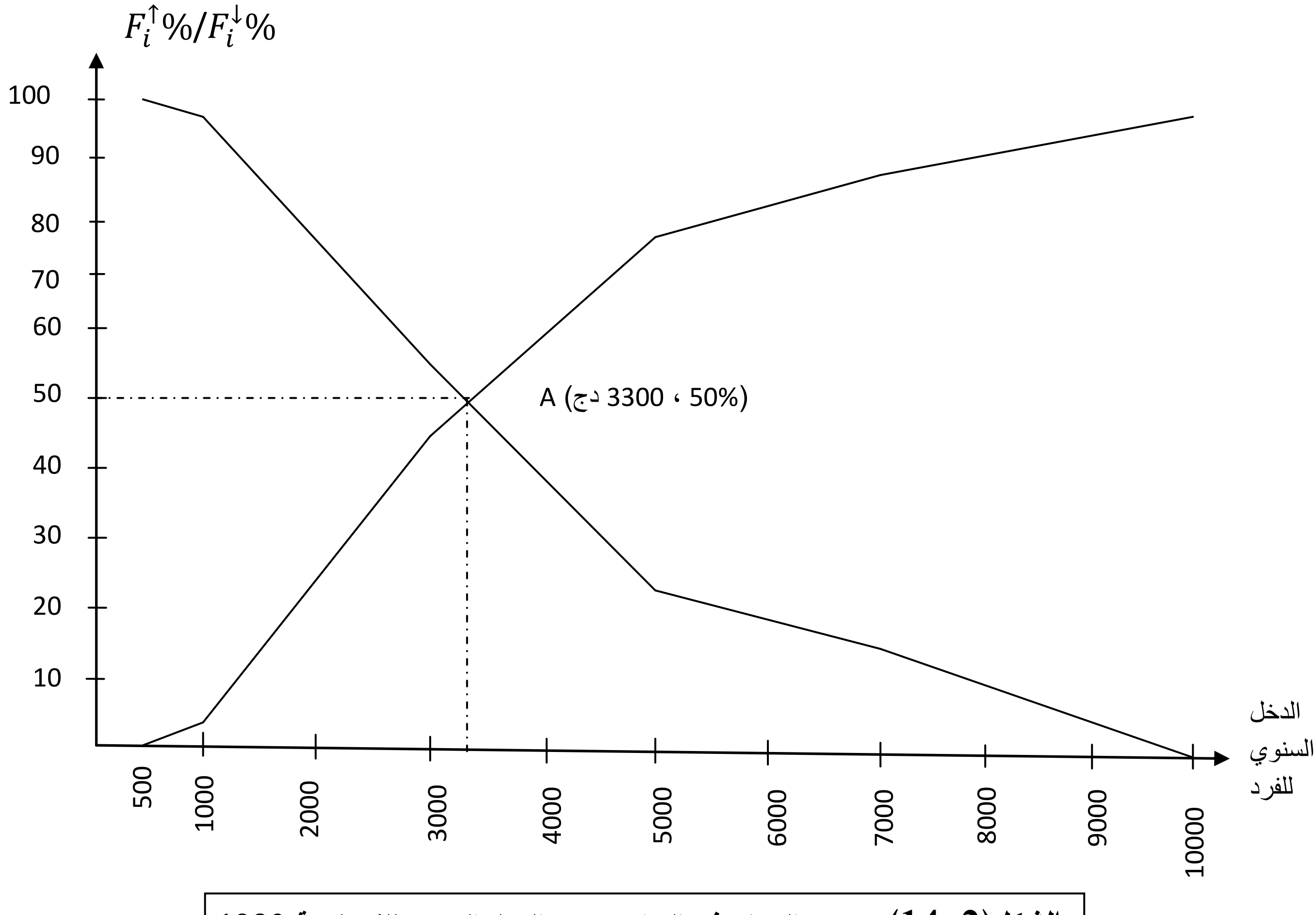
ب- الشرح:

$f_2^{\uparrow}\% = 41,10\%$: هناك 41,10% من السكان دخلهم السنوي يتراوح بين 1000 دج و3000 دج.

$F_2^{\uparrow}\% = 44,75\%$: هناك 44,75% من السكان دخلهم السنوي أقل تماما من 3000 دج.

$F_2^{\downarrow}\% = 96,35\%$: هناك 96,35% من السكان دخلهم السنوي يساوي أو يفوق 1000 دج.

ج- التمثيل البياني: بما أن المتغير الإحصائي مستمر فإنه يمثل عن طريق **المنحنى المتجمع الصاعد والنازل**



الشكل(2-14): توزيع السكان في الجزائر حسب الدخل السنوي للفرد لسنة 1980

3– تحديد إحداثيتي نقطة التقاطع بين المنحنى التجميعي الصاعد والمنحنى التجميعي النازل:

A (3300 , 50%)

– الشرح الإحصائي:

هذه النقطة تعني أن 50% من مجموع السكان يتقاضون دخلا سنويا للفرد يقل عن 3300 دج و 50% المتبقية يتقاضون دخلا أكثر من 3300 دج.

4– نسبة السكان الذين دخلهم السنوي يساوي أو يفوق 5000 دج: %21,30 = 13,2 + 8,10

– نسبة السكان الذين دخلهم السنوي يقل عن 3000 دج: %44,75 = 41,10 + 3,65

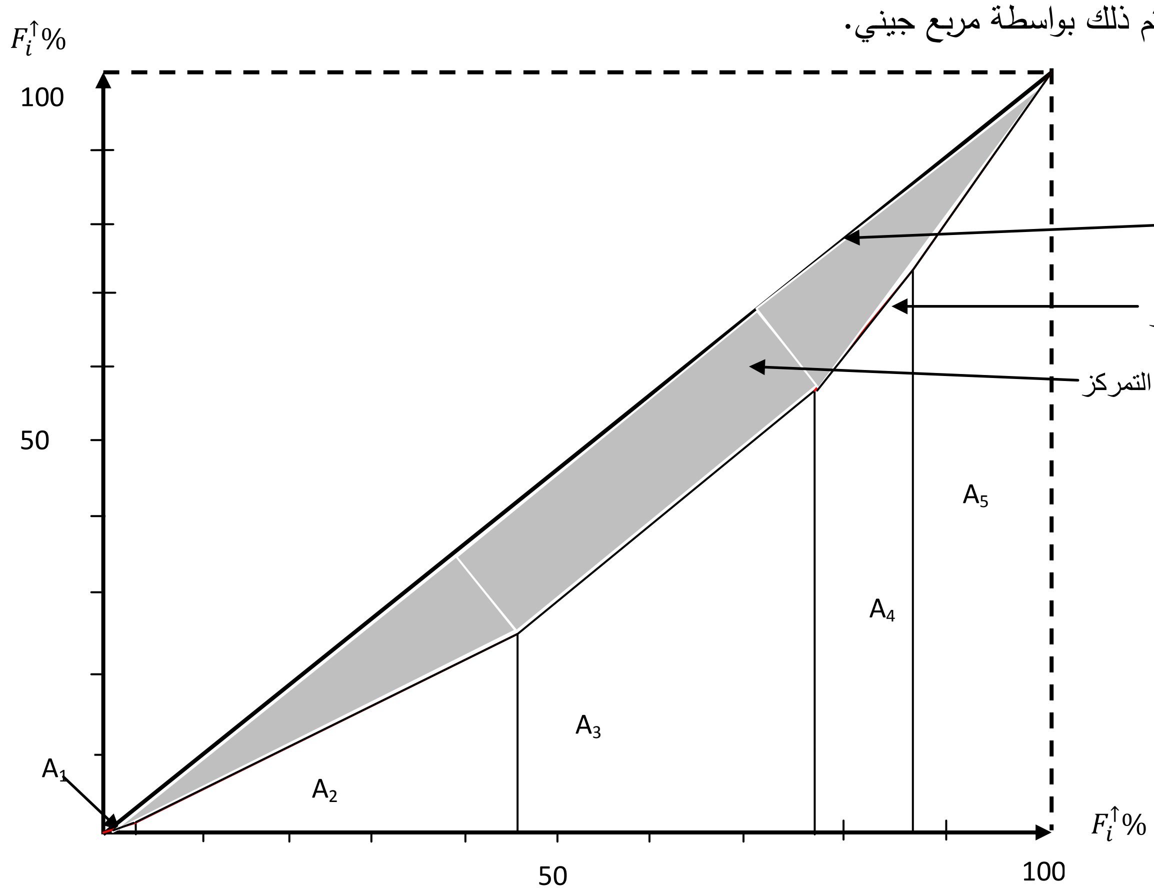
– نسبة السكان الذين دخلهم السنوي يفوق أو يساوي 500 ويقل عن 5000:

%78,70 = 33,95 + 41,10 + 3,65

حل التمرين الثالث:

– دراسة قضية التمركز على هذا التوزيع:

أ– بيانيا: ويتم ذلك بواسطة مربع جيني.



الشكل(2–15): مربع جيني لتوزيع السكان حسب الدخل الشهري للفرد

نلاحظ من خلال الشكل(2−15) أن مساحة التمركز تبدو صغيرة مقارنة بمساحة المثلث ABC، وعليه يمكن القول أن التمركز ضعيف نوعا ما، أي أن توزيع الدخل أكثر عدالة.

ب− حسابيا:

من خلال حساب معامل جيني كما يلي:

$f_i(S_i^{\uparrow} + S_{i-1}^{\uparrow})$	$S_i^{\uparrow} + S_{i-1}^{\uparrow}$	$S_i^{\uparrow}\%$	$f_i\%$
2,92	0,80	0,80	3,65
1043,94	25,40	24,60	41,10
2831,43	83,40	58,80	33,95
1045,71	129,1	70,30	8,10
2247,96	170,3	100	13,20
7171,96	/	/	**100,00**

بما أن النسب مئوية فإن:

$$I_{Gini} = 1 - \frac{1}{10000}\sum_{i=1}^{k} f_i\%(S_i^{\uparrow}\% + S_{i-1}^{\uparrow}\%)$$

$$I_{Gini} = 1 - \frac{1}{10000}(7171,96)$$

$$I_{Gini} = 0,2828 = 28,28\%$$

نلاحظ أن $I_{Gini} = 0,2828$ وهي أقل من 0,30 وعليه يمكن القول أن هذا التوزيع أكثر عدالة، حيث توجد هناك فعلا فوارق في توزيع الدخل الإجمالي على مختلف فئات السكان، ولكن يمكن أن نعتبر هذا التمركز ضعيفا، أي أن هذه الفوارق يمكن تفسيرها بأسباب موضوعية

تمارين مقترحــــــــــــــــة

التمرين الأول:

الجدول التالي يعطينا نتائج دراسة إحصائية حول غيابات العمال في أحد الشركات (الوحدة: يوم)

الجدول(2−12): غيابات العمال في أحد الشركات باليوم

عدد أيام التغيب	$]12-8]$	$]16-12]$	$]20-16]$	$]24-20]$	$]28-24]$	$]32-28]$
نسبة العمال	0,06	0,10	0,30	0,40	0,10	0,04

المطلوب:

1− حدد على هذه المسألة بدقة: المجتمع الإحصائي، الوحدة الإحصائية، المتغير الإحصائي ونوعه وكذا طريقة جمع البيانات؟

2− أحسب التكرارات المطلقة، والتكرارات المطلقة والنسبية الصاعدة والنازلة واشرح معنى:

n_3 ،n_5 ،f_3 ،f_5 ، $N_3^{\uparrow}$ ، $N_5^{\uparrow}$ ،$N_3^{\downarrow}$ ،$N_5^{\downarrow}$ ، $F_3^{\uparrow}$ ، $F_5^{\uparrow}$ ،$F_3^{\downarrow}$ ،$F_5^{\downarrow}$، مع العلم أن عدد عمال الشركة هو 200

3− مثل بيانيا هذا التوزيع بواسطة المدرج التكراري، المضلع التكراري والمنحنى التكراري؟

4− أرسم المنحنى التجميعي الصاعد والنازل باستعمال التكرارات النسبية؟

5− ما هي إحداثيات نقطة التقاطع ما بين المنحنيين، اشرح المعني الإحصائي لهذه الإحصائيات؟

التمرين الثانـي:

الدراسة الإحصائية للأوزان الخاصة بـ 50 شخصا يكونون جمعية رياضية أعطت النتائج التالية (بالكيلوغرام)، معروضة على شكل سلسلة إحصائية كما يلي:

37، 61، 68، 74، 82، 43، 62، 69، 74، 84،

47، 63، 69، 75، 86، 50، 63، 70، 76، 87،

52، 64، 71، 76، 88، 54، 65، 72، 77، 90،

55، 66، 72، 79، 92، 56، 66، 72، 79، 93،

58، 67، 73، 80، 98، 58، 68، 73، 82، 98.

1− حدد على هذه المسألة بدقة: المجتمع الإحصائي، الوحدة الإحصائية، المتغير الإحصائي ونوعه؟

2− أعرض هذه البيانات في جدول تكراري على شكل فئات بحيث يكون:

− عدد الفئات هو تسع (09) فئات.

– الفئة الأولى هي: من 35 إلى أقل من 50، طول الفئة الثانية والأخيرة هو 10.

3– مثل بيانيا هذا التوزيع؟

4– أرسم المنحنى التجميعي الصاعد والنازل باستعمال التكرارات النسبية؟

5– ما هو عدد الأشخاص الذين يتراوح وزنهم ما بين 60 و 70 كلغ؟ يقل وزنهم عن 63 كلغ؟

التمرين الثالث:

قامت شركة سونطراك بمسابقة من أجل التوظيف، فحصلت الشركة إجمالا على 200 ملف طلب من بينها 35 ملف تم رفضه لأنه لم يلبي شروط المسابقة، فكانت نتائج المسابقة موزعة على الجدول التالي:

الجدول(2–13): نتائج مسابقة التوظيف التي قامت بها شركة سونطراك

18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0	x_i
2	2	3	16	11	13	10	β	8	9	7	α	18	6	6	5	4	4	1	n_i

أولا:

1– عرف المجتمع الإحصائي، الوحدة الإحصائية وحجم العينة؟

2– حدد المتغير الإحصائي ونوعه لهذا التوزيع؟ مثله بيانيا؟

3– أحسب التكرارين α و β، إذا علمت أن $\alpha = \frac{3}{2}\beta$؟ مع شرحهما؟

4– أحسب التكرارات النسبية والنسبية المئوية المتجمعة الصاعدة والنازلة؟ مثلها بيانيا؟

5– ما هي نسبة المرشحين الذين حصلوا على معدل أقل من 10 مع الشرح؟ أكثر من 10 مع الشرح؟

ثانيا:

نجمع المعلومات الإحصائية السابقة على شكل فئات ذات المراكز: 2، 6،...، 18.

1– ما الغرض من جمع هذه المعلومات في شكل فئات؟ 2– مثل هذه الفئات في شكل بياني مناسب؟

3– إذا كان شرط النجاح هو الحصول على علامة 16، ما هي نسبة النجاح؟

4– قررت شركة سونطراك تقديم فرصة إضافية للمترشحين الحاصلين على علامة 12 فما أكثر لأن عدد الناجحين غير كاف، فما هو عدد ونسبة المترشحين المعنيين بالفرصة الثانية؟

التمرين الرابع:

يمثل الجدول التالي توزيع عينة من 50 أستاذ حسب درجة الأستاذية في جامعة معينة:

الجدول(2−14): توزيع عينة من 50 أستاذ حسب درجة الأستاذية في جامعة معينة

الدرجة	n_i
أستاذ مساعد قسم أ	20
أستاذ مساعد قسم ب	14
أستاذ محاضر	10
أستاذ التعليم العالي	06
المجموع	50

المطلوب:

1− حدد على هذه المسألة بدقة: المجتمع الإحصائي، الوحدة الإحصائية، المتغير الإحصائي ونوعه؟

2− أحسب كلا من: $N_i^{\uparrow}$ ، $N_i^{\downarrow}$، f_i ، $f_i\%$، $F_i^{\uparrow}\%$، $F_i^{\downarrow}$ % ، ثم إشرح: n_2 ، $N_2^{\uparrow}$ ، $N_4^{\downarrow}$، $F_2^{\uparrow}\%$، $F_4^{\downarrow}\%$ ،

3− مثل بيانيا هذا التوزيع؟

التمرين الخامس:

الجدول التالي يمثل توزيع السكان وإجمالي الدخل السنوي للأسر حسب 6 فئات للدخل السنوي للأسر لسنة 1980، الوحدة: %

الجدول(2−15): توزيع السكان وإجمالي الدخل السنوي للأسر حسب 6 فئات للدخل السنوي للأسر لسنة 1980

الدخل السنوي للأسر (دج)	نسبة الأسر $f_i\%$ (%)	نسبة الدخل $s_i\%$
]5000 − 10000]	13,20	06,28
]10000 − 15000]	19,20	14,50
]15000 − 20000]	25,50	25,02
]20000 − 30000]	22,00	22,08
]30000 − 40000]	15,08	17,02
]40000 − 60000]	05,02	15,10
المجموع	**100,00**	**100,00**

المصدر: الإستقصاء الوطني حول دخل ومصاريف الأسر لسنة 1980 على عينة حجمها 8208 أسرة.

المطلوب:

1− حدد المتغير الإحصائي ونوعه مع التمثيل البياني لهذا التوزيع؟

2− أدرس قضية التمركز على هذا التوزيع مبينا هل هناك عدالة أم لا في توزيع الدخل على مختلف فئات الأسر: أ− بيانيا ب− حسابيا

3− بينت نفس الدراسة أن مساحة خارج منطقة التمركز للمناطق الحضرية 3650 ملم^2، مساحة منطقة التمركز للمناطق الريفية 1550 ملم^2 ، أدرس الفوارق في توزيع الدخل بين المنطقتين (قضية التمركز)؟

التمرين السادس:

الجدول التالي يعطينا توزيع السكان والدخل في الجزائر لخمس مجموعات من السكان حسب الترتيب التصاعدي للدخل، كل مجموعة تحتوي على 20% من إجمالي عدد السكان لسنة 1995.

الجدول(2-16): وزيع السكان والدخل في الجزائر لخمس مجموعات من السكان حسب الترتيب التصاعدي للدخل

	المجموعة 1	المجموعة 2	المجموعة 3	المجموعة 4	المجموعة 5	المجموع
نسبة السكان f_i	0,2	0,2	0,2	0,2	0,2	**1**
نسبة الدخل s_i	0,07	0,116	0,161	0,227	0,426	**1**

المصدر: تقرير البنك الدولي حول التنمية في العالم لسنة 1999

المطلوب:

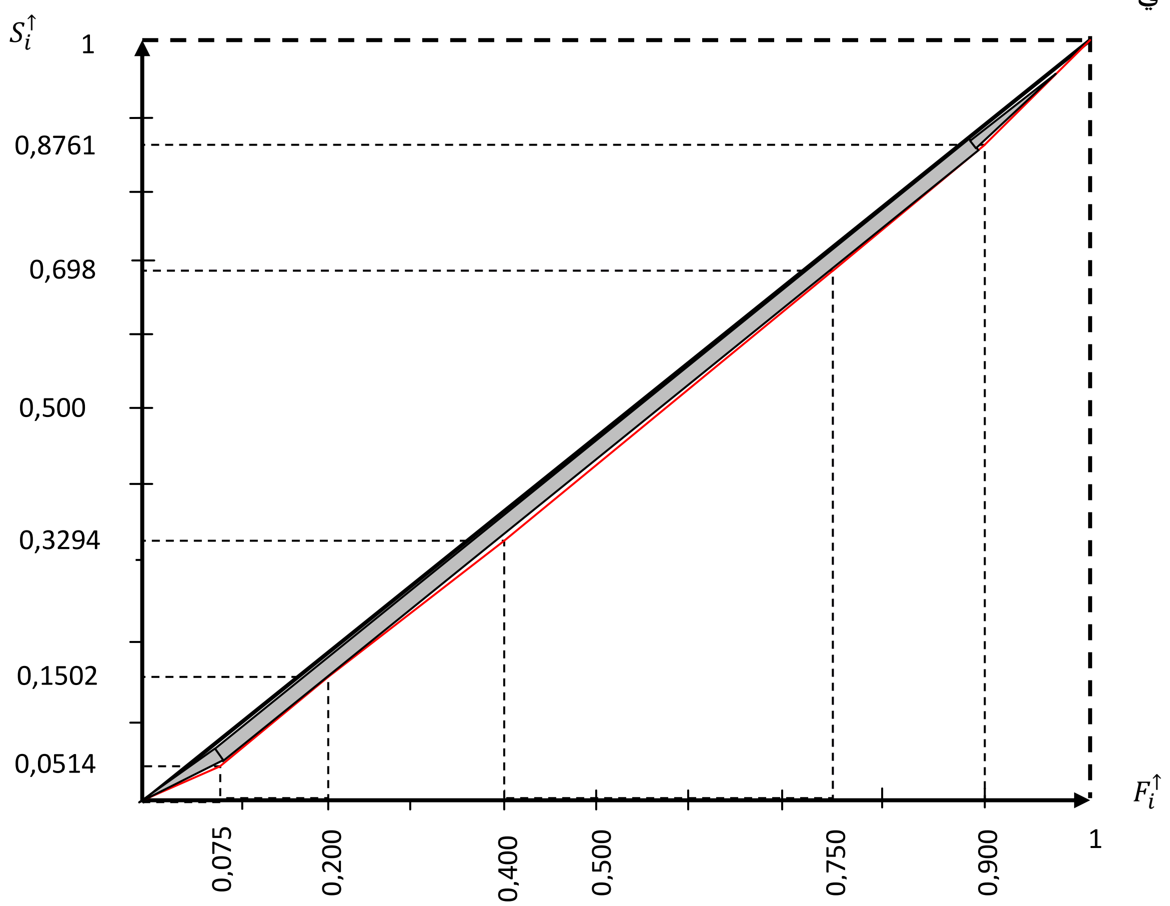
1- أدرس قضية التمركز على توزيع الدخل في الجزائر بيانيا؟

2- إذا علمت أن مساحة التمركز تقدر بـ 0,1646 ملم^2، أدرس التمركز حسابيا مع الشرح الإقتصادي؟

3- في السنة نفسها أجريت دراسة مماثلة حول توزيع المداخيل بالبرازيل، وكانت مساحة خارج منطقة التمركز تقدر بـ 0,2292 ملم^2، قارن توزيع المداخيل بين البلديين؟

التمرين السابع:

إليك الشكل التالي:



حيث $F_i^{\uparrow}$: يمثل التكرار التجميعي الصاعد النسبي للعمال للمؤسسة A.

$S_i^{\uparrow}$: تمثل النسبة المتجمعة الصاعدة للأجور اليومية للعمال للمؤسسة A.

إليك الجدول التالي:

الأجر اليومي	$[60-70]$	$[70-80]$	$[80-90]$	$[90-110]$	$[110-115]$	$[115-120]$
$F_i^{\uparrow}$	؟	؟	؟	؟	؟	؟
$S_i^{\uparrow}$	؟	؟	؟	؟	؟	؟
f_i	؟	؟	؟	؟	؟	؟

المطلوب:

1- ماذا يمثل الشكل البياني السابق؟ ماذا نستنتج بالنسبة لمدى عدالة التوزيع؟

2- حدد كلا من المجتمع الإحصائي، الوحدة الإحصائية، المتغير المدروس ونوعه في هذا التوزيع؟

3- أكمل بيانات الجدول أعلاه؟

4- أكد على إستنتاجك الذي أعطيته في السؤال الأول حسابيا؟

5- مثل هذا التوزيع بيانيا بواسطة المدرج التكراري؟

6- بفرض أنه تم دراسة التمركز على توزيع مؤسسة أخرى B، حيث قدرت المساحة خارج منطقة التمركز بـ 3505,8827 ملم^2، حدد أي التوزيعين أكثر عدالة؟

الفصل الثالث

مقاييس النزعة المركزيـة

نتطرق من خلال هذا الفصل إلى العناصر التالية:

أولا- المتوسط الحسابي

ثانيا- المنوال

ثالثا- الوسيط

رابعا- مشتقات الوسيط

خامسا- مشتقات المتوسط الحسابي

تمارين محلولة وتمارين مقترحة

تمهيد:

لقد لاحظ المحللون للسلاسل الإحصائية حول مواضيع مختلفة أن معظم بيانات السلسلة الإحصائية تتجه أو تنزع إلى التمركز أو التجمع حول قيم متميزة تقع في مركز البيانات والقليل منها تتطرف إما بالكبر وإما بالصغر. نسمي هذه الظاهرة بظاهرة النزعة المركزية. أما القيم التي تتمركز حولها البيانات الأخرى فتسمى مقاييس النزعة المركزية أو المتوسطات، فهي إذن تعين لنا موقع الظاهرة ونستعملها في تقدير مستوى الظاهرة المدروسة.

هناك عدة مقاييس للتعبير عن هذه الظاهرة تختلف من ناحية الدقة والمدلول الإحصائي وطريقة الحساب، من أهمها:

- المتوسط الحسابي ومشتقاته(المتوسط الهندسي، المتوسط التوافقي، المتوسط التربيعي).
- الوسيط ومشتقاته(الربيعيات، العشيريات، والمئويات).
- المنوال .

أولا- المتوسط الحسابي $\bar{X}$:

يعتبر المتوسط الحسابي من أهم مقاييس النزعة المركزية وأكثرها استخداما في النواحي التطبيقية، ويعرف عموما على أنه مجموع القيم مقسوما على عددها.

1- الطريقة المباشرة:

1-1- المتوسط الحسابي البسيط (غير المرجح):

لتكن السلسلة الإحصائية X_1، X_2 ... ... ، X_i، ... ... ، X_n ، يحسب $\bar{X}$ كالتالي:

$$\boxed{\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n} \ldots\ldots\ldots\ldots (1)}$$

حيث: X_i تمثل قيم المتغير الإحصائي، و n تمثل عدد القيم.

مثال(3-1):

ليكن لدينا علامات مجموعة من الطلبة في مقياس الإحصاء1:

12، 14، 16، 08، 07، 05، 11، 15، 06، 13.

المطلوب: حساب متوسط علامات الطلبة؟.

الحـــــــل: $$\bar{X} = \frac{\sum_{i=1}^{10} X_i}{n} = \frac{12+14+16+08+07+05+11+15+06+13}{10} = \frac{107}{10} = 10{,}7$$

متوسط علامات الطلبة في مقياس الإحصاء1 هو 10,7.

1-2- المتوسط الحسابي المرجح:

إذا كانت لدينا بيانات مبوبة أي في شكل توزيع تكراري فإن:

$$\bar{X} = \frac{n_1X_1 + n_2X_2 + \cdots + n_kX_k}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{k} n_iX_i}{\sum_{i=1}^{k} n_i} \ldots\ldots\ldots\ldots (2)$$

حيث: X_i تمثل قيم المتغير الإحصائي المتقطع، و n_i تمثل التكرارات المطلقة الموافقة لها و n تمثل مجموع التكرارات.

ملاحظات:

- نسمي $\bar{X}$ في الصيغة (2) بالمتوسط الحسابي المرجح، لأننا نرجح كل قيمة X_i في الجدول بتكرارها المطلق n_i أو تكرارها النسبي f_i.

- نلاحظ من الصيغة (2) أن $f_i = \frac{n_i}{\sum n_i}$ وبالتعويض في الصيغة نفسها نجد:

$$\bar{X} = f_1X_1 + f_2X_2 + \cdots + f_kX_K = \sum_{i=1}^{k} f_iX_i \ldots\ldots\ldots\ldots (3)$$

مثال(3-2): البيانات التالية تمثل عدد الغرف في المسكن الواحد لعينة من 50 مسكن ببلدية سطيف.

الجدول(3-1): توزيع عينة من 50 مسكن ببلدية سطيف حسب عدد الغرف في المسكن الواحد

عدد الغرف (قيم المتغير X_i)	عدد المساكن (التكرار) n_i	n_iX_i	التكرار النسبي f_i	f_iX_i
1	1	1	0,02	0,02
2	8	16	0,16	0,32
3	13	39	0,26	0,78
4	13	52	0,26	1,04
5	6	30	0,12	0,60
6	4	24	0,08	0,48
7	3	21	0,06	0,42
8	2	16	0,04	0,32
المجموع	$\sum n_i = 50$	199	1	3,98

المطلوب: حساب متوسط عدد الغرف في المسكن الواحد؟

الحـــــل:

$$\bar{X} = \frac{\sum_{i=1}^{k} n_iX_i}{\sum_{i=1}^{k} n_i} = \frac{\sum_{i=1}^{8} n_iX_i}{50} = \frac{199}{50} = 3{,}98$$

$$\bar{X} = \sum_{i=1}^{k} f_iX_i = \sum_{i=1}^{8} f_iX_i = 3{,}98$$

متوسط عدد الغرف في المسكن الواحد هو 3,98.

2- الطريقة غير المباشرة:

عندما تكون لدينا بيانات كبيرة القيم فإننا نستخدم طريقة غير مباشرة في حساب المتوسط الحسابي (طريقة

الانحراف عن متوسط فرضي)، الغرض منها هو تصغير قيم البيانات من أجل تسهيل عملية حساب المتوسط الحسابي.

2-1- المتوسط الحسابي البسيط:

إذا كانت لدينا بيانات غير مبوبة أي في شكل سلسلة إحصائية، فإننا نتبع الخطوات التالية:

أ- نفرض قيمة ثابتة a تكون قريبة من القيم الأصلية وتتوسطها (نسميها متوسط فرضي)؛

ب- حساب الانحرافات d_i بين قيم السلسلة والمتوسط الفرضي: $d_i = X_i - a$؛

ج- حساب المتوسط الحسابي للقيم الجديدة: $\bar{d} = \frac{\sum d_i}{n}$؛

د- حساب المتوسط الحسابي الأصلي حيث لدينا: $\bar{d} = \frac{\sum d_i}{n} = \frac{\sum(X_i - a)}{n} = \frac{\sum X_i - \sum a}{n} = \frac{\sum X_i}{n} - \frac{na}{n}$

$$\bar{d} = \bar{X} - a \Rightarrow \boxed{\bar{X} = \bar{d} + a}$$

مثال(3-3):

لتكن لدينا القيم التالية: 12، 10، 15، 8، 5. أحسب الوسط الحسابي عن طريق انحرافات القيم عن وسط فرضي.

الحل: نفرض أن $a = 9$، وعليه نحسب الانحرافات وفق الصيغة التالية: $d_i = X_i - a$، فيتنج لدينا:

X_i	$d_i = X_i - a$
5	−4
8	−1
10	1
12	3
15	6
المجموع	**5**

$$\bar{d} = \frac{\sum d_i}{n} = \frac{5}{5} = 1$$

$$\bar{X} = \bar{d} + a = 1 + 9 = 10$$

كما أننا نحصل على نفس النتيجة بتطبيق العلاقة:

$$\bar{X} = \frac{\sum_{i=1}^{5} X_i}{n} = \frac{50}{5} = 10$$

2-2- المتوسط الحسابي المرجح:

إذا كانت لدينا بيانات مبوبة أي في شكل توزيع تكراري فإننا نتبع نفس الخطوات السابقة ولكن بوضع:

$$\boxed{\bar{d} = \frac{\sum_{i=1}^{k} d_i n_i}{\sum n_i}}$$

كما أن: $\boxed{\bar{X} = \bar{d} + a}$

مثال(3-4): أحسب المتوسط الحسابي بطريقة الانحراف عن متوسط فرضي نفترض أنه $a = 800$؟

X_i	500	700	800	850	900	1000	المجموع
n_i	10	16	02	18	16	10	**72**

الحل:

X_i	n_i	$d_i = x_i - 800$	$d_i n_i$
500	10	−300	−3000
700	16	−100	−1600
800	02	0	0
850	18	50	900
900	16	100	1600
1000	10	200	2000
المجموع	**72**	/	**−100**

$$\bar{d} = \frac{\sum_{i=1}^{k} n_i d_i}{\sum n_i} = \frac{\sum_{i=1}^{6} n_i d_i}{72} = \frac{-100}{72} = -1,39$$

حساب المتوسط الحسابي الأصلي: $\bar{X} = \bar{d} + a = -1,39 + 800 = 798,61$

إذن المتوسط الحسابي لهذه البيانات: 798,61.

ملاحظة: إذا كان المتغير الإحصائي مستمر فإننا نعوض X_i بمراكز الفئات$\mathbf{C_i}$ في كل المعادلات السابقة أي:

$$\boxed{\bar{X} = \frac{n_1c_1 + n_2c_2 + \cdots + n_kc_k}{\sum n_i} = \frac{\sum_{i=1}^{k} n_i c_i}{\sum n_i}}$$

$$\boxed{\bar{X} = f_1c_1 + f_2c_2 + \cdots + f_kc_K = \sum_{i=1}^{k} f_i c_i}$$

$\boxed{\bar{X} = \bar{d} + a}$ حيث $\boxed{\bar{d} = \frac{\sum_{i=1}^{k} d_i n_i}{\sum n_i}}$ و $\boxed{d_i = c_i - a}$

مثال(3−5): البيانات التالية تمثل أوزان 60 طالبا بالكيلوغرام في أحد أقسام الـــ: LMD بكلية العلوم الاقتصادية وعلوم التسيير بجامعة سطيف، والمطلوب حساب متوسط أوزان الطلبة؟

الجدول(3−2): أوزان 60 طالبا بالكيلوغرام في أحد أقسام الـ: LMD بكلية العلوم الاقتصادية وعلوم التسيير بجامعة سطيف

أوزان الطلبةX_i	عدد الطلبة n_i
]55 − 50]	2
]60 − 55]	5
]65 − 60]	12
]70 − 65]	16
]75 − 70]	14
]80 − 75]	8
]85 − 80]	3
المجموع	**60**

الحل:

أوزان الطلبة X_i	عدد الطلبة n_i	C_i	$X_i\,n_i$
]55 – 50]	2	52,5	105
]60 – 55]	5	57,5	287,5
]65 – 60]	12	62,5	750
]70 – 65]	16	67,5	1080
]75 – 70]	14	72,5	1015
]80 – 75]	8	77,5	620
]85 – 80]	3	82,5	247,5
المجموع	60	/	4105

$\bar{X} = \frac{\sum_{i=1}^{7} n_i c_i}{\sum n_i} = \frac{4105}{60} = 68{,}42$ ، ومنه فإن متوسط أوزان الطلبة هو 68,42 كلغ.

3- خصائص المتوسط الحسابي:

أ- يتأثر المتوسط الحسابي بالقيم المتطرفة.

مثلا: تتكون أسرة من 5 أفراد، تبلغ أعمارهم على الترتيب: 55، 49، 13، 11، 10.

المتوسط الحسابي هو: $\bar{X} = \frac{55+49+13+11+10}{5} = \frac{138}{5} = 27{,}6$

نلاحظ أن النتيجة المحصل عليها لا تمثل أي قيمة من القيم المدروسة، وبالتالي نستنتج أنه لا يمكن استعمال المتوسط الحسابي في مثل هذه الحالات، لأنه من المفروض أن تكون قيم المتغير الإحصائي متمركزة حول النتيجة المحصل عليها.

ب- يستعمل المتوسط الحسابي في حالة المتغيرات الكمية أي القابلة للقياس.

ج- لا يمكن حساب المتوسط الحسابي في حالة البيانات المفتوحة (أقل من، أكثر من).

د- مجموع انحرافات القيم عن متوسطها يساوي دائما الصفر. $\boxed{\sum(X_i - \bar{X}) = 0}$

$$\sum(X_i - \bar{X}) = \sum X_i - \sum \bar{X}$$
$$\sum(X_i - \bar{X}) = \sum X_i - n\bar{X} = n\bar{X} - n\bar{X} = 0$$

تدل هذه الخاصية على أن المتوسط الحسابي يقع في مركز البيانات.

هـ- مجموع مربع انحرافات القيم عن متوسطها الحسابي أقل من مجموع مربعات انحراف نفس القيم عن أي قيمة أخرى، أي: $\boxed{\sum(X_i - \bar{X})^2 < \sum(X_i - X_\alpha)^2}$ حيث: $\bar{X} \neq X_\alpha$

تدل هذه الخاصية على أن المتوسط الحسابي أقرب إلى البيانات X_i من أي قيمة أخرى.

مثلا: إذا كانت لدينا القيم التالية: 24، 8، 6، 16، 14، 4، أثبت صحة الخاصيتين د ، هـ؟

نفرض أن النقطة المختارة هي: $X_\alpha = 10$ ولدينا: $\overline{X} = \frac{\sum_{i=1}^{6} X_i}{n} = \frac{72}{6} = 12$

القيم x_i	$(X_i - \overline{X})$	$(X_i - \overline{X})^2$	$(X_i - X_\alpha)$	$(X_i - X_\alpha)^2$
24	12	144	14	196
8	−4	16	−2	4
6	−6	36	−4	16
16	4	16	6	36
14	2	4	4	16
4	−8	64	−6	36
$\sum X_i = 72$	$\sum(X_i - \overline{X}) = 0$	$\sum(X_i - \overline{X})^2 = 280$	$\sum(X_i - X_\alpha) = 12$	$\sum(X_i - X_\alpha)^2 = 304$

بالنسبة للخاصية الأولى نلاحظ من خلال الجدول أن: $\sum(X_i - \overline{X}) = 0$ وهو المطلوب.

أما بالنسبة للخاصية الثانية فمن خلال مقارنة النتائج المتحصل عليها في الجدول أعلاه نجد: $304 > 280$

ينتج عن ذلك: $\sum(X_i - \overline{X})^2 < \sum(X_i - X_\alpha)^2$.

ثانيا- المنوال:

يعرف المنوال على أنه القيمة الأكثر تكرارا من بين مجمل القيم المعطاة، ويرمز لها بـ: $\boldsymbol{M_o}$.

1- حساب المنوال في حالة سلسلة إحصائية:

قيمة المنوال للبيانات: 12، 16، 10، 12، 17، 9 هي: $M_o = 12$ لأنها الأكثر تكرارا من غيرها.

قيمة المنوال للبيانات: 10، 15، 18، 13، 10، 15، 16 هي: $M_o = 10$ و $M_o = 15$.

البيانات التالية: 14، 16، 9، 17، 13 ليس لها منوالا.

2- حساب المنوال في حالة توزيع تكراري لمتغير إحصائي منفصل:

يستنتج مباشرة من جدول التوزيع التكراري، مع الإشارة إلى أنه يمكننا أن نجد أكثر من منوال، كما يمكننا ألا نجد ولا منوال.

مثا(3-6): البيانات التالية تمثل عدد الغرف في المسكن الواحد لعينة من 35 مسكن ببلدية سطيف.

الجدول(3-3): عدد الغرف في المسكن الواحد لعينة من 35 مسكن ببلدية سطيف

عدد الغرف X_i (قيم المتغير)	عدد المساكن n_i (التكرار)
1	3
2	8
3	13
4	5
5	6
المجموع	35

المنوال في هذا التوزيع هو: $M_o = 3$

الشرح:

أغلبية السكنات تحتوي على 3 غرف.

3- حساب المنوال في حالة توزيع تكراري لمتغير إحصائي مستمر:

إذا كان لدينا جدول توزيع تكراري على شكل فئات فإننا نتبع الخطوات التالية لحساب المنوال:

- تحديد الفئة المنوالية: وهي الفئة التي تقابل أكبر تكرار عندما تكون أطوال الفئات متساوية، أو الفئة التي تقابل أكبر تكرار معدل عندما تكون أطوال الفئات غير متساوية.

- حساب المنوال بطريقة المد الداخلي:

$$M_o = Lim_{M_o} + \left[\frac{\Delta_1}{\Delta_1+\Delta_2}\right] \times A_{M_o}$$

حيث:

Lim_{M_o}: الحد الأدنى للفئة المنوالية، Δ_1: الفرق بين تكرار الفئة المنوالية وتكرار الفئة السابقة لها.

Δ_2: الفرق بين تكرار الفئة المنوالية وتكرار الفئة اللاحقة لها، A_{M_o}: طول الفئة المنوالية.

مثال(3-7): أحسب قيمة المنوال لبيانات المثال(3-5)؟

أوزان الطلبة X_i	عدد الطلبة n_i
$]55-50]$	2
$]60-55]$	5
$]65-60]$	12
$]70-65]$	16
$]75-70]$	14
$]80-75]$	8
$]85-80]$	3
المجموع $\sum n_i$	60

بما أن أطوال الفئات متساوية فإن الفئة المنوالية هي: $[65-70]$

وبالتالي فإن: $\Delta_1 = 16-12 = 4$ ، $\Delta_1 = 16-14 = 2$

ومنه :

$$M_o = Lim_{M_o} + \left[\frac{\Delta_1}{\Delta_1+\Delta_2}\right] \times A_{M_o}$$

$$M_o = 65 + \left[\frac{4}{4+2}\right] \times 5 = 68{,}33$$

أغلبية الطلبة وزنهم يقدر بـ: 68,33 كلغ.

4- تحديد المنوال بيانيا:

يحدد المنوال بيانيا بواسطة المدرج التكراري، وهذا بإتباع الخطوات التالية:

أ- نرسم المدرج التكراري للتوزيع.

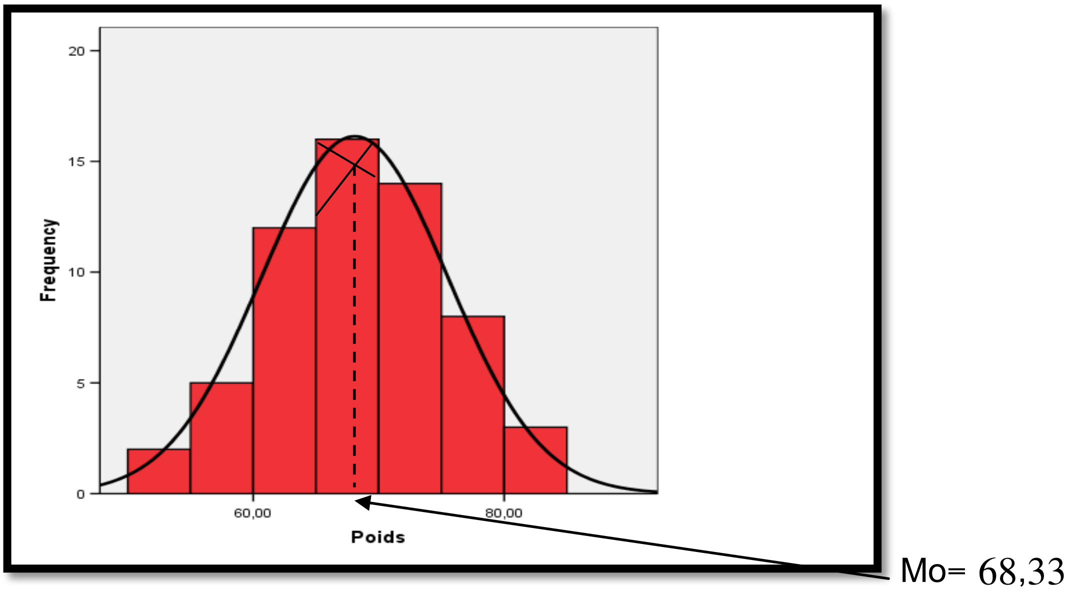
ب- نصل بخط مستقيم رأس الحد الأعلى للفئة المنوالية برأس الحد الأعلى للفئة السابقة لها.

ج- نصل بخط مستقيم رأس الحد الأدنى للفئة المنوالية برأس الحد الأدنى للفئة اللاحقة لها.

د- من تقاطع الخطين السابقين نسقط عمودا على المحور الأفقي ونقطة تقاطعه مع المحور الأفقي تمثل تقديرا لقيمة المنوال بيانيا.

مثال(3-8): حدد قيمة المنوال بيانيا للمثال السابق؟

الحل:



الشكل(3−1): أوزان 60 طالبا بالكيلوغرام في أحد أقسام الـ LMD بكلية العلوم الاقتصادية وعلوم التسيير بجامعة سطيف

ثالثا − الوسيط:

الوسيط هو أحد مقاييس النزعة المركزية الذي يأخذ بعين الاعتبار رتبة القيم، ويعرف الوسيط على أنه القيمة التي تقسم البيانات إلى جزئين متساويين بحيث تكون قيم المتغير الإحصائي مرتبة ترتيبا تصاعديا أو تنازليا، ونرمز له بالرمز M_e

1− حساب الوسيط في حالة سلسلة إحصائية أو توزيع تكراري لمتغير إحصائي متقطع:

لحساب الوسيط في هذه الحالة نتبع الخطوات التالية:

أ− ترتيب البيانات ترتيبا تصاعديا.

ب− إذا كان عدد البيانات n عددا فرديا فإن الوسيط هو القيمة التي رتبتها $\frac{n+1}{2}$ أي: $\boxed{M_e = X_{\left(\frac{n+1}{2}\right)}}$

ج− إذا كان عدد البيانات n عددا زوجيا فإن الوسيط هو متوسط القيمة التي رتبتها $\frac{n}{2}$ والقيمة التي رتبتها $\frac{n}{2}+1$ أي: $\boxed{M_e = \frac{X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}}{2}}$

مثال(3−9): أحسب الوسيط للسلسلتين الإحصائيتين التاليتين:

− السلسلة الأولى: (1، 2، 5، 4، 3، 6، 7، 7، 5، 3، 1، 9)

− السلسلة الثانية: (1، 0، 2,5، 3، 5، 5,5، 1، 6، 7,5)

الحل:

− السلسلة الأولى: (1، 1، 2، 3، 3، 4، 5، 5، 6، 7، 7، 9)

عدد البيانات زوجي أي: 12، ومنه: $M_e = \frac{\left(\frac{n}{2}\right)+X\left(\frac{n}{2}+1\right)}{2} = \frac{X_6+X_7}{2} = \frac{4+5}{2} = 4.5$

50% من البيانات أقل من 4,5، و50% من البيانات أكبر من 4,5.

– السلسلة الثانية: (0، 1، 1، 2,5، 3، 5، 5,5، 6، 7,5)

عدد البيانات فردي أي: 9، ومنه: $M_e = X_{\left(\frac{n+1}{2}\right)} = X_5 = 3$

50% من البيانات على الأكثر أقل من 3، و50% من البيانات على الأكثر أكبر من 3.

2– حساب الوسيط في حالة توزيع تكراري لمتغير إحصائي مستمر:

إذا كان لدينا جدول توزيع تكراري على شكل فئات فإننا نتبع الخطوات التالية لحساب الوسيط:

– تحديد الفئة الوسيطية: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{n}{2}$، أي: $N_{M_e}^{\uparrow} \geq \frac{n}{2}$

– حساب الوسيط بطريقة المد الداخلي:

$$\boxed{M_e = Lim_{M_e} + \left[\frac{\frac{n}{2} - N_{M_e-1}^{\uparrow}}{n_{M_e}}\right] \times A_{M_e}}$$

حيث:

Lim_{M_e}: الحد الأدنى للفئة الوسيطية، $\frac{n}{2}$: رتبة الوسيط

$N_{M_e-1}^{\uparrow}$: التكرار المتجمع الصاعد المطلق للفئة قبل الفئة الوسيطية.

n_{M_e}: تكرار الفئة الوسيطية ، A_{M_e}: طول الفئة الوسيطية.

مثال(3–10): بالعودة إلى المثال(3–5)، أحسب الوسيط واشرح النتيجة؟

أوزان الطلبة X_i	عدد الطلبة n_i	$N_i^{\uparrow}$.
$]55-50]$	2	2
$]60-55]$	5	7
$]65-60]$	12	19
$\mathbf{]70-65]}$	**16**	**35**
$]75-70]$	14	49
$]80-75]$	8	57
$]85-80]$	3	60
المجموع $\sum n_i$	60	/

– تحديد الفئة الوسيطية: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{n}{2}$ ، أي:

$$N_{M_e}^{\uparrow} \geq \left(\frac{n}{2} = \frac{60}{2} = 30\right)$$

ومنه الفئة الوسيطية هي: $[65 - 70[$

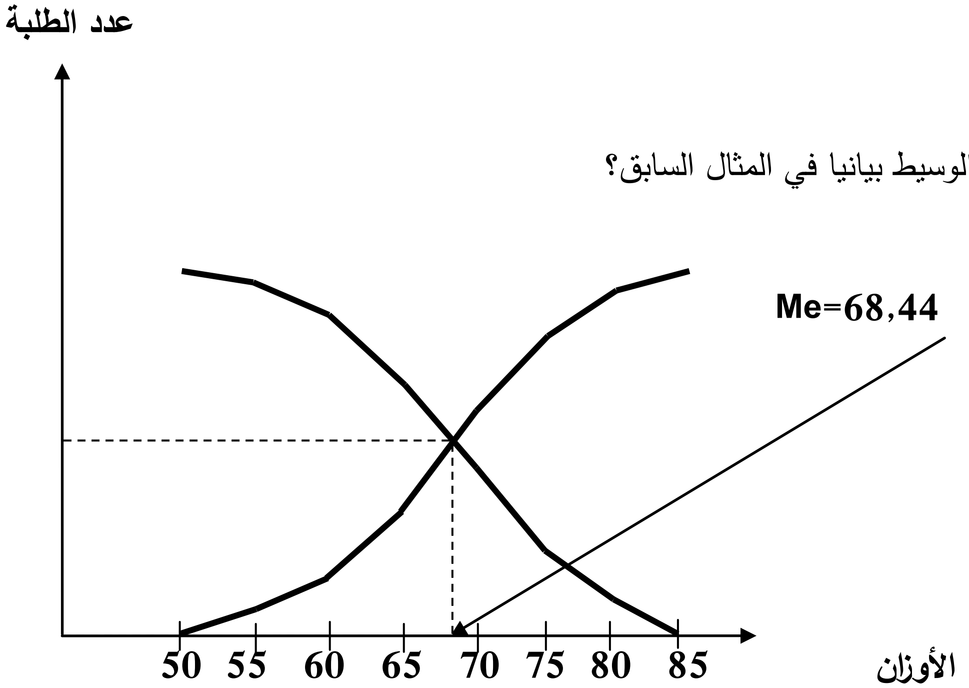
– حساب الوسيط بطريقة المد الداخلي:

$$M_e = Lim_{M_e} + \left[\frac{\frac{n}{2} - N^{\uparrow}_{M_e-1}}{n_{M_e}}\right] \times A_{M_e} = 65 + \left[\frac{30-19}{16}\right] \times 5 = 68,44$$

الشرح: هناك 50% من الطلبة أوزانهم أقل من 68,44 كلغ و 50% من الطلبة أوزانهم أكبر من 68,44 كلغ.

ملاحظة: الوسيط بيانيا هو نقطة التقاطع بين المنحنى المتجمع الصاعد والنازل.

مثال(3−11): حدد قيمة الوسيط بيانيا في المثال السابق؟



الشكل(3−2): أوزان 60 طالبا بالكيلوغرام في أحد أقسام الـ: LMD بكلية العلوم الاقتصادية وعلوم التسيير بجامعة سطيف

رابعا: مشتقات الوسيط:

1– الربيعيات: وهي الناتجة من تقسيم البيانات إلى أربع أقسام متساوية، وبالتالي كل قسم يمثل 25% من البيانات.

1−1− الربيع الأول: تقسم البيانات إلى 25% من القيم أقل من قيمة الربيع الأول و 75% من القيم أكبر من قيمة الربيع الأول، ونرمز له بالرمز $\boldsymbol{Q_1}$.

أ– إذا كانت البيانات في شكل سلسلة إحصائية أو توزيع تكراري لمتغير إحصائي متقطع:

$$\boxed{Q_1 = X_{\left(\frac{n+1}{4}\right)}}$$

إذا كانت $\frac{n+1}{4}$ طبيعية (دون فواصل) نأخذ القيمة مباشرة.

إذا كانت $\frac{n+1}{4}$ غير طبيعية (وجود فواصل) نأخذ متوسط القيمتين، مثلا:1، 3، 5، 7، 9، 11، 13، 15.

$\frac{n+1}{4} = \frac{9}{4} = 2,25$.

الربيع الأول هنا: متوسط القيمة التي رتبتها 2 والقيمة التي رتبتها3 أي $4 = \frac{3+5}{2} = Q_1$

ب- في حالة توزيع تكراري لمتغير إحصائي مستمر:

- تحديد الفئة الربيعية الأولى: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{n}{4}$، أي: $N_{Q_1}^{\uparrow} \geq \frac{n}{4}$

- حساب الربيع الأول بطريقة المد الداخلي: $\boxed{Q_1 = Lim_{Q_1} + \left[\frac{\frac{n}{4} - N_{Q_1-1}^{\uparrow}}{n_{Q_1}}\right] \times A_{Q_1}}$

2-1- الربيع الثالث: تقسم البيانات إلى 75% من القيم أقل من قيمة الربيع الثالث و 25% من القيم أكبر من قيمة الربيع الثالث، ونرمز له بالرمز $\boldsymbol{Q_3}$.

أ- إذا كانت البيانات في شكل سلسلة إحصائية أو توزيع تكراري لمتغير إحصائي متقطع:

$$\boxed{Q_3 = X_{\frac{3(n+1)}{4}}}$$

إذا كانت $\frac{3(n+1)}{4}$ طبيعية (دون فواصل) نأخذ القيمة مباشرة.

إذا كانت $\frac{3(n+1)}{4}$ غير طبيعية (وجود فواصل) نأخذ متوسط القيمتين، مثلا:1، 3، 5، 7، 9، 11، 13، 15.

$\frac{3(n+1)}{4} = \frac{27}{4} = 6,75$.

الربيع الثالث هنا: متوسط القيمة التي رتبتها 6 والقيمة التي رتبتها 7 أي $12 = \frac{11+13}{2} = Q_3$

ب- في حالة توزيع تكراري لمتغير إحصائي مستمر:

- تحديد الفئة الربيعية الثالثة: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{3n}{4}$، أي: $N_{Q_3}^{\uparrow} \geq \frac{3n}{4}$

- حساب الربيع الثالث بطريقة المد الداخلي: $\boxed{Q_3 = Lim_{Q_3} + \left[\frac{\frac{3n}{4} - N_{Q_3-1}^{\uparrow}}{n_{Q_3}}\right] \times A_{Q_3}}$

2- العشيريات: وهي الناتجة من تقسيم البيانات إلى عشر أقسام متساوية، وبالتالي كل قسم يمثل 10% من البيانات.

2-1- العشير الأول: تقسم البيانات إلى 10% من القيم أقل من قيمة العشير الأول و90% من القيم أكبر من قيمة العشير الأول، ونرمز له بالرمز $\boldsymbol{D_1}$.

أ- إذا كانت البيانات في شكل سلسلة إحصائية أو توزيع تكراري لمتغير إحصائي متقطع:

$$\boxed{D_1 = X_{\left(\frac{n+1}{10}\right)}}$$

إذا كانت $\frac{n+1}{10}$ طبيعية (دون فواصل) نأخذ القيمة مباشرة.

إذا كانت $\frac{n+1}{10}$ غير طبيعية (وجود فواصل) نأخذ متوسط القيمتين.

مثلا:1، 3، 5، 7، 9، 11، 13، 15، 17، 19، 21.

$\frac{n+1}{10} = \frac{11}{10} = 1,1$.

العشير الأول هنا: متوسط القيمة التي رتبتها 1 والقيمة التي رتبتها2 أي $D_1 = \frac{1+3}{2} = 2$

ب- في حالة توزيع تكراري لمتغير إحصائي مستمر:

- تحديد الفئة العشيرية الأولى: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{n}{10}$، أي: $N_{D_1}^{\uparrow} \geq \frac{n}{10}$

- حساب العشير الأول بطريقة المد الداخلي: $\boxed{D_1 = Lim_{D_1} + \left[\frac{\frac{n}{10} - N_{D_1-1}^{\uparrow}}{n_{D_1}}\right] \times A_{D_1}}$

2-2- العشير الثامن: تقسم البيانات إلى 80% من القيم أقل من قيمة العشير الثامن و 20% من القيم أكبر من قيمة العشير الثامن، ونرمز له بالرمز $\boldsymbol{D_8}$.

أ- إذا كانت البيانات في شكل سلسلة إحصائية أو توزيع تكراري لمتغير إحصائي متقطع:

$$\boxed{D_8 = X_{\frac{8(n+1)}{10}}}$$

إذا كانت $\frac{8(n+1)}{10}$ طبيعية (دون فواصل) نأخذ القيمة مباشرة.

إذا كانت $\frac{8(n+1)}{10}$ غير طبيعية (وجود فواصل) نأخذ متوسط القيمتين، مثلا:

مثلا:1، 3، 5، 7، 9، 11، 13، 15، 17، 19، 21.

$\frac{8(n+1)}{10} = \frac{96}{10} = 9{,}6$.

العشير الثامن هنا: متوسط القيمة التي رتبتها 9 والقيمة التي رتبتها 10 أي $D_8 = \frac{17+19}{2} = 18$

ب- في حالة توزيع تكراري لمتغير إحصائي مستمر:

- تحديد الفئة العشيرية الثامنة: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{8n}{10}$، أي: $N_{D_8}^{\uparrow} \geq \frac{8n}{10}$

- حساب العشير الثامن بطريقة المد الداخلي: $\boxed{D_8 = Lim_{D_8} + \left[\frac{\frac{8n}{10} - N_{D_8-1}^{\uparrow}}{n_{D_8}}\right] \times A_{D_8}}$

وهكذا بالنسبة لباقي العشيريات، مع مراعاة الرتبة المناسبة.

3- المئويات: وهي الناتجة من تقسيم البيانات إلى مائة قسم متساوي، وبالتالي كل قسم يمثل 1% من البيانات.

3-1- المئوي الأول: تقسم البيانات إلى 1% من القيم أقل من قيمة المئوي الأول و 99% من القيم أكبر من قيمة المئوي الأول، ونرمز له بالرمز $\boldsymbol{C_1}$.

أ- إذا كانت البيانات في شكل سلسلة إحصائية أو توزيع تكراري لمتغير إحصائي متقطع:

$$\boxed{C_1 = X_{\left(\frac{n+1}{100}\right)}}$$

إذا كانت $\frac{n+1}{100}$ طبيعية (دون فواصل) نأخذ القيمة مباشرة.

إذا كانت $\frac{n+1}{100}$ غير طبيعية (وجود فواصل) نأخذ متوسط القيمتين.

مثلا إذا وجدنا: $\frac{n+1}{100} = 6{,}3$، المئوي الأول هنا: متوسط القيمة التي رتبتها 6 والقيمة التي رتبتها 7.

ب- في حالة توزيع تكراري لمتغير إحصائي مستمر:

- تحديد الفئة المئوية الأولى: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{n}{100}$، أي: $N_{C_1}^{\uparrow} \geq \frac{n}{100}$

- حساب المئوي الأول بطريقة المد الداخلي: $$\boxed{C_1 = Lim_{C_1} + \left[\frac{\frac{n}{100} - N_{C_1-1}^{\uparrow}}{n_{C_1}}\right] \times A_{C_1}}$$

3-2- المئوي السابع والستون: تقسم البيانات إلى 67% من القيم أقل من قيمة المئوي السابع والستون و 33% من القيم أكبر من قيمة المئوي السابع والستون ، ونرمز له بالرمز $\boldsymbol{C_{67}}$.

أ- إذا كانت البيانات في شكل سلسلة إحصائية أو توزيع تكراري لمتغير إحصائي متقطع:

$$\boxed{C_{67} = X_{\frac{67(n+1)}{100}}}$$

إذا كانت $\frac{67(n+1)}{100}$ طبيعية (دون فواصل) نأخذ القيمة مباشرة.

إذا كانت $\frac{67(n+1)}{100}$ غير طبيعية (وجود فواصل) نأخذ متوسط القيمتين.

مثلا إذا وجدنا: $\frac{67(n+1)}{100} = 31{,}6$.

المئوي السابع والستون هنا: متوسط القيمة التي رتبتها 31 والقيمة التي رتبتها 32.

ب- في حالة توزيع تكراري لمتغير إحصائي مستمر:

- تحديد الفئة المئوية السابعة والستون: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{67n}{100}$، أي: $N_{C_{67}}^{\uparrow} \geq \frac{67n}{100}$

- حساب المئوي السابع والستون بطريقة المد الداخلي: $$\boxed{C_{67} = Lim_{C_{67}} + \left[\frac{\frac{67n}{100} - N_{C_{67}-1}^{\uparrow}}{n_{C_{67}}}\right] \times A_{C_{67}}}$$

وهكذا بالنسبة لباقي المئويات، مع مراعاة الرتبة المناسبة.

ملخص قوانين مشتقات الوسيط:

الربيعيات إذا كانت البيانات في شكل سلسلة إحصائية أو توزيع تكراري لمتغير إحصائي متقطع:

$$Q_i = X_{\frac{i}{4}(n+1)}$$

الربيعيات إذا كانت البيانات في شكل توزيع تكراري لمتغير إحصائي مستمر:

$$Q_i = Lim_{Q_i} + \left[\frac{\left(\frac{i}{4}\right)n - N^{\uparrow}_{Q_i-1}}{n_{Q_i}}\right] \times A_{Q_i}$$

العشيريات إذا كانت البيانات في شكل سلسلة إحصائية أو توزيع تكراري لمتغير إحصائي متقطع:

$$D_i = X_{\frac{i}{10}(n+1)}$$

العشيريات إذا كانت البيانات في شكل توزيع تكراري لمتغير إحصائي مستمر:

$$D_i = Lim_{D_i} + \left[\frac{\left(\frac{i}{10}\right)n - N^{\uparrow}_{D_i-1}}{n_{D_i}}\right] \times A_{D_i}$$

المئويات إذا كانت البيانات في شكل سلسلة إحصائية أو توزيع تكراري لمتغير إحصائي متقطع:

$$C_i = X_{\frac{i}{100}(n+1)}$$

المئويات إذا كانت البيانات في شكل توزيع تكراري لمتغير إحصائي مستمر:

$$C_i = Lim_{C_i} + \left[\frac{\left(\frac{i}{100}\right)n - N^{\uparrow}_{C_i-1}}{n_{C_i}}\right] \times A_{C_i}$$

خامسا- مشتقات المتوسط الحسابي:

بالإضافة إلى المتوسط الحسابي $\bar{X}$ الذي يحسب بطريقة معينة هناك متوسطات أخرى من نفس درجة الدقة للمتوسط الحسابي، إلا أنها تحسب بطرق متميزة لأن سلسلة البيانات ليس لها نفس الطبيعة في حالة المتوسط الحسابي.

$\bar{X}$ سمي متوسطا حسابيا لأن سلسلة البيانات $x_1, x_2, \dots\dots, x_n$ تحمل من الناحية الرياضية شكل المتتالية الحسابية، وأما إذا كانت السلسلة الإحصائية لا تتزايد بطريقة حسابية فلا يمكن أن نحسب عليها متوسط حسابي بل نستعمل متوسط آخر، فإذا كانت مثلا تتزايد بطريقة هندسية نطبق عليها المتوسط الهندسي.

1- المتوسط الهندسي $\bar{X}_G$:

1-1- في حالة سلسلة إحصائية:

لتكن السلسلة الإحصائية $x_1, x_2, \dots, x_n$ نحسب المتوسط الهندسي $\bar{X}_G$ على هذه السلسلة بالطريقة التالية:

$$\bar{X}_G = \sqrt[n]{x_1 . x_2 . \dots\dots . x_n} \quad \dots\dots\dots (1)$$

$$\boxed{\bar{X}_G = \sqrt[n]{\prod_{i=1}^{n} x_i}}$$

1-2- في حالة توزيع تكراري:

إذا كانت $x_1, x_2, \dots, x_k$ تمثل قيم المتغير الإحصائي، و $n_1, n_2, \dots, n_k$ تمثل تكراراتها على الترتيب، فإن المتوسط الهندسي يحسب وفق العلاقة التالية:

$$\bar{X}_G = \sqrt[\sum n_i]{(x_1 . x_1 . x_1 \dots x_1)(x_2 . x_2 . \dots x_2) \dots\dots\dots (x_k . x_k \dots x_k)}$$

$$\bar{X}_G = \sqrt[\sum n_i]{x_1^{n_1} . x_2^{n_2} \dots\dots\dots x_k^{n_k}} \quad \dots\dots\dots (2)$$

$$\boxed{\bar{X}_G = \sqrt[\sum n_i]{\prod_{i=1}^{k} x_i^{n_i}}}$$

ملاحظة هامة:

نلاحظ من الصيغتين (1) و(2) أنه إذا كانت البيانات في السلسلة أو الجدول تحمل قيم كبيرة تصبح الحسابات ضخمة أو حتى مستحيلة وعليه لا نستعمل الصيغتين (1) و(2)، وإنما نلجأ إلى تبسيط البيانات بإدخال الحساب اللوغاريتمي.

1-3- الصيغة اللوغاريتمية للمتوسط الهندسي:

أ- في حالة سلسلة إحصائية:

لتكن الصيغة (1) السابقة:

$$\bar{X}_{\mathrm{G}} = \sqrt[n]{x_1.x_2.\dots.x_n} \Rightarrow \bar{X}_{\mathrm{G}} = (x_1.x_2.\dots.x_n)^{\frac{1}{n}}$$

$$\Rightarrow log(\bar{X}_{\mathrm{G}}) = \log(x_1.x_2.\dots.x_n)^{\frac{1}{n}}$$

$$\Rightarrow log(\bar{X}_{\mathrm{G}}) = \frac{1}{n}\log(x_1.x_2.\dots.x_n)$$

$$\Rightarrow log(\bar{X}_{\mathrm{G}}) = \frac{1}{n}\left(\log(x_1) + \log(x_2) + \cdots + log(x_n)\right)$$

$$\Rightarrow log(\bar{X}_{\mathrm{G}}) = \frac{1}{n}\sum_{i=1}^{n}\log(x_i)$$

بعد إجراء الحسابات في الصيغة اللوغاريتمية نحسب $\bar{X}_{\mathrm{G}}$ كما يلي: $log(\bar{X}_{\mathrm{G}}) = \frac{1}{n}\sum_{i=1}^{n}\log(x_i)$

من هذه الصيغة نستنتج: $\boxed{\bar{X}_{\mathrm{G}} = 10^{\left(\frac{1}{n}\sum_{i=1}^{n}\log(x_i)\right)}}$

ب- في حالة توزيع تكراري:

لتكن الصيغة (2) السابقة:

$$\bar{X}_{\mathrm{G}} = \sqrt[\sum n_i]{x_1^{n_1}.x_2^{n_2}\dots\dots\dots x_k^{n_k}} \Rightarrow \bar{X}_{\mathrm{G}} = \left(x_1^{n_1}.x_2^{n_2}\dots\dots\dots x_k^{n_k}\right)^{\frac{1}{\sum n_i}}$$

$$\Rightarrow log(\bar{X}_{\mathrm{G}}) = \log\left(x_1^{n_1}.x_2^{n_2}\dots\dots\dots x_k^{n_k}\right)^{\frac{1}{\sum n_i}}$$

$$\Rightarrow log(\bar{X}_{\mathrm{G}}) = \frac{1}{\sum n_i}\log\left(x_1^{n_1}.x_2^{n_2}\dots\dots\dots x_k^{n_k}\right)$$

$$\Rightarrow log(\bar{X}_{\mathrm{G}}) = \frac{1}{\sum n_i}\left(\log\left(x_1^{n_1}\right) + \log\left(x_2^{n_2}\right) + \cdots + log\left(x_k^{n_k}\right)\right)$$

$$\Rightarrow log(\bar{X}_{\mathrm{G}}) = \frac{1}{\sum n_i}\left(n_1\log(x_1) + n_2\log(x_2) + \cdots + n_k log(x_k)\right)$$

$$\Rightarrow log(\bar{X}_{\mathrm{G}}) = \frac{1}{\sum n_i}\sum_{i=1}^{k} n_i\log(x_i)$$

بعد إجراء الحسابات في الصيغة اللوغاريتمية نحسب $\bar{X}_{\mathrm{G}}$ كما يلي: $log(\bar{X}_{\mathrm{G}}) = \frac{1}{n}\sum_{i=1}^{k} n_i\log(x_i)$

من هذه الصيغة نستنتج: $\boxed{\bar{X}_{\mathrm{G}} = 10^{\left(\frac{1}{n}\sum_{i=1}^{n} n_i\log(x_i)\right)}}$

ملاحظة: نستعمل المتوسط الهندسي في المجال الاقتصادي لحساب المعدلات (معدل الفائدة، معدل نمو السكان...إلخ)، متوسط نسبة النمو الاقتصادي لعدة سنوات، الأرقام القياسية.

2- المتوسط التوافقي $\bar{X}_H$:

المتوسط التوافقي هو متوسط نادر الاستعمال نوظفه في بعض الحالات لحساب القدرة الشرائية المتوسطة لسلعة ما بدلالة سلعة أخرى، وعادة ما تكون هذه الوحدة هي الوحدة النقدية لبلد معين، وكذلك لحساب السرعة المتوسطة لمركبة على مسالك مختلفة.

2-1- في حالة سلسلة إحصائية:

لتكن السلسلة الإحصائية $x_1, x_2, \dots, x_n$ نحسب المتوسط التوافقي $\bar{X}_{\mathrm{H}}$ على هذه السلسلة بالطريقة التالية:

- نحسب مقلوبات القيم X_i: $\frac{1}{x_1}, \frac{1}{x_2}, \dots, \frac{1}{x_n}$
- نحسب متوسط المقلوبات (مجموع المقلوبات تقسيم n): $\frac{\sum_{i=1}^{n}\frac{1}{x_i}}{n}$
- نقلب هذا المتوسط فنحصل على $\bar{X}_H$ أي: $\boxed{\bar{X}_H = \frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}}}$

2-2- في حالة توزيع تكراري:

في حالة التوزيع التكراري فإن المتوسط التوافقي يحسب بالعلاقة التالية: $\boxed{\bar{X}_H = \frac{\sum_{i=1}^{k} n_i}{\sum_{i=1}^{k}\frac{n_i}{x_i}}}$

3- المتوسط التربيعي $\bar{X}_Q$:

3-1- في حالة سلسلة إحصائية:

لتكن السلسلة الإحصائية $x_1, x_2, \dots, x_n$ نحسب المتوسط التربيعي $\bar{X}_Q$ على هذه السلسلة بالطريقة التالية:

$$\bar{X}_Q = \sqrt{\frac{x_1^2 + x_2^2 + \dots + x_n^2}{n}}$$

$$\boxed{\bar{X}_Q = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}}}$$

3-2- في حالة توزيع تكراري:

في حالة التوزيع التكراري فإن المتوسط التربيعي يحسب بالعلاقة التالية: $\boxed{\bar{X}_Q = \sqrt{\frac{\sum_{i=1}^{k} n_i x_i^2}{\sum_{i=1}^{k} n_i}}}$

نتيجة: مهما تكن البيانات فإنه في كل الأحوال: $\boxed{\bar{X}_H < \bar{X}_G < \bar{X} < \bar{X}_Q}$

ملاحظة: إذا كان المتغير الإحصائي مستمر فإننا نعوض $\boldsymbol{X_i}$ بمراكز الفئات $\boldsymbol{C_i}$ في كل المعادلات السابقة.

مثال(3−12): لتكن البيانات التالية: 3، 2، 4، 6

المطلوب: 1− أحسب: المتوسط الحسابي، المتوسط الهندسي، المتوسط التوافقي، المتوسط التربيعي؟

− الحل:

1− حساب المتوسطات:

أ− المتوسط الحسابي:

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{\sum_{i=1}^{4} X_i}{4} = \frac{2+3+4+6}{4} = \frac{15}{4} = 3{,}75$$

ب− المتوسط الهندسي:

$$\bar{X}_{\mathrm{G}} = \sqrt[n]{x_1 . x_2 . \ldots \ldots x_k \ldots \ldots . x_n} = \sqrt[4]{2 \times 3 \times 4 \times 6} = 3{,}46$$

ج− المتوسط التوافقي:

$$\bar{X}_{\mathrm{H}} = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}} = \frac{4}{\sum_{i=1}^{4} \frac{1}{x_i}} = \frac{4}{\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{6}} = \frac{4}{1{,}247} = 3{,}21$$

د− المتوسط التربيعي:

$$\bar{X}_Q = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}} = \sqrt{\frac{\sum_{i=1}^{4} x_i^2}{4}} = \sqrt{\frac{(2)^2+(3)^2+(4)^2+(6)^2}{4}} = 4{,}03$$

تمارين محلولة للفصل الثالث

التمرين الأول:

إليك التوزيع التكراري لعدد الرضع حسب الوزن عند الولادة في أحد العيادات خلال سنة 2006:

الجدول(3-4): التوزيع التكراري لعدد الرضع حسب الوزن عند الولادة في أحد العيادات خلال سنة 2006

الوزن بالكيلوغرام	$]2-1]$	$]2,5-2]$	$]3-2,5]$	$]3,5-3]$	$]4-3,5]$	$]6-4]$
نسبة المواليد%	10	15	26	31	10	8

المطلوب:

1- أحسب الوزن المتوسط، الوسيط والمنوال واشرح النتائج؟

2- حدد قيمتي المنوال والوسيط بيانيا؟

3- حدد قيمة الربيع الثالث، العشير السادس، المئوي 43، حسابيا ثم بيانيا مع شرح النتائج؟

التمرين الثاني:

البيانات التالية تبين عدد الغرف في المسكن الواحد لعينة من 50 مسكنا في بلدية سطيف.

4	2	4	6	3	3	4	2	5	4
4	7	2	8	4	5	7	3	4	2
6	3	4	8	2	5	5	7	4	3
3	4	3	6	3	6	4	2	5	4
2	3	3	5	4	1	3	3	4	2

المطلوب:

1- أعرض هذه البيانات في جدول تكراري ثم أحسب التكرارات النسبية؟

2- أحسب العدد المتوسط للغرف في المسكن الواحد، على السلسلة وعلى الجدول؟

3- في دراسة مماثلة في بلدية أخرى، تبين أن العدد الإجمالي للغرف هو 261 غرفة وأن العدد المتوسط للغرف هو 3 غرف.

أ- ما هو حجم العينة التي أجريت عليها الدراسة؟

ب- قارن نوع المسكن بين البلديتين؟

ج- أي الدراستين أكثر دقة؟ لماذا؟

4- من خلال التأمل في بيانات السلسلة الإحصائية نعتقد أن العدد المتوسط للغرف في المسكن الواحد هو 5، ماهو العدد المتوسط الحقيقي للغرف (فقط على الجدول)؟

5- أحسب المنوال مع شرح النتيجة؟

6- أحسب الوسيط مع شرح النتيجة؟

7- أحسب الربيع الأول والعشير الرابع والمئوي67 مع شرح النتائج؟

التمرين الثالث:

قامت مصلحة الأسعار والنوعية بمراقبة أسعار الكيلوغرام من البرتقال في 10 أسواق مختلفة لمدينة سطيف فكانت النتائج التالية:

الجدول(3-5): أسعار الكيلوغرام من البرتقال في 10 أسواق مختلفة لمدينة سطيف

السعر (دج)	[20 - 30[	[30 - 50[	[50 - 80[	[80 - 120[	المجموع
عدد الأسواق	1	3	4	2	10

المطلوب:

أحسب السعر المتوسط للكيلوغرام من البرتقال باستعمال:

أ- المتوسط الحسابي، ب- المتوسط الهندسي، ج- المتوسط التوافقي، د- المتوسط التربيعي؟

التمرين الرابع:

تشير تقارير وزارة التخطيط أن المعدلات السنوية للنمو الاقتصادي لمختلف المخططات التنموية الوطنية للفترة 1967-1992 كانت كالتالي:

الجدول(3-6): المعدلات السنوية للنمو الاقتصادي لمختلف المخططات التنموية الوطنية للفترة 1967-1992

المخطط الثلاثي	المخطط الرباعي 1	المخطط الرباعي 2	المخطط الخماسي 1	المخطط الخماسي 2	المخطط الخماسي 3
%4	%6,07	%7,20	%7,80	%5,50	%3,20

- أحسب النسبة السنوية المتوسطة للنمو الاقتصادي للفترة المذكورة؟

التمرين الخامس:

قامت أحد الشركات الصناعية بتجربة على أحد شاحناتها الجديدة بقطع مسافة 50 كلم على 4 مسالك مختلفة بغرض اختبار سرعتها فقطعت هذه الشاحنة:

- المسلك الأول بسرعة 50 كلم/سا
- المسلك الثاني بسرعة 150 كلم/سا
- المسلك الثالث بسرعة 75 كلم/سا
- المسلك الرابع بسرعة 100 كلم/سا

المطلوب: ما هو متوسط السرعة لهذه الشاحنة على العموم؟

الحلـــــــــــــــــــول

حل التمرين الأول:

الوزن X_i(كلغ)	نسبة الولادة $f_i\%$	نسبة الولادة f_i	C_i	$C_i f_i$	a_i	$n'_i = \frac{f_i\%}{a_i} \times a$	$F_i^{\uparrow}\%$	$F_i^{\uparrow}\%$
$]2 - 1]$	10	0,1	1,5	0,15	1	5	10	100
$]2,5 - 2]$	15	0,15	2,25	0,3375	0,5	15	25	90
$]3 - 2,5]$	26	0,26	2,75	0,715	0,5	26	51	75
$]3,5 - 3]$	31	0,31	3,25	1,0075	0,5	31	82	49
$]4 - 3,5]$	10	0,1	3,75	0,375	0,5	10	92	18
$]6 - 4]$	8	0,08	5	0,4	2	2	100	8
المجموع $\sum n_i$	100	1	/	2,985	/	/	/	/

1− حساب الوزن المتوسط، الوسيط والمنوال مع شرح النتائج:

أ− المتوسط الحسابي: $\bar{X} = \sum_{i=1}^{k} f_i c_i = 2,985\ kg$

الشرح: يقدر وزن الرضيع عند الولادة في المتوسط بـــ: 2,985 كلغ.

ب− المنوال: بما أن أطوال الفئات غير متساوية فإننا نقوم بتعديل التكرارات أولا ثم نحسب المنوال.

الطول الشائع هو 0,5، أما التعديل فهو وفق الجدول أعلاه.

بعد التعديل تصبح الفئة المنوالية هي: $[3 - 3,5[$

وبالتالي فإن: $\Delta_1 = 31 - 26 = 5$ ، $\Delta_1 = 31 - 10 = 21$

ومنه :

$$M_o = Lim_{M_o} + \left[\frac{\Delta_1}{\Delta_1 + \Delta_2}\right] \times A_{M_o}$$

$$M_o = 3 + \left[\frac{5}{5+21}\right] \times 0,5 = 3,1\ kg$$

الشرح: أغلبية المواليد الرضع وزنهم عند الولادة يقدر بـــ: 3,1 كلغ.

ج− الوسيط:

− تحديد الفئة الوسيطية: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{n}{2}$ ، وبما أن التكرارات نسبية فإن الفئة الوسيطية هي أول فئة تكرارها المتجمع الصاعد النسبي أكبر أو يساوي 50% أي:

$$F_{M_e}^{\uparrow} \geq 50\%$$

ومنه الفئة الوسيطية هي: $[2,5 - 3[$

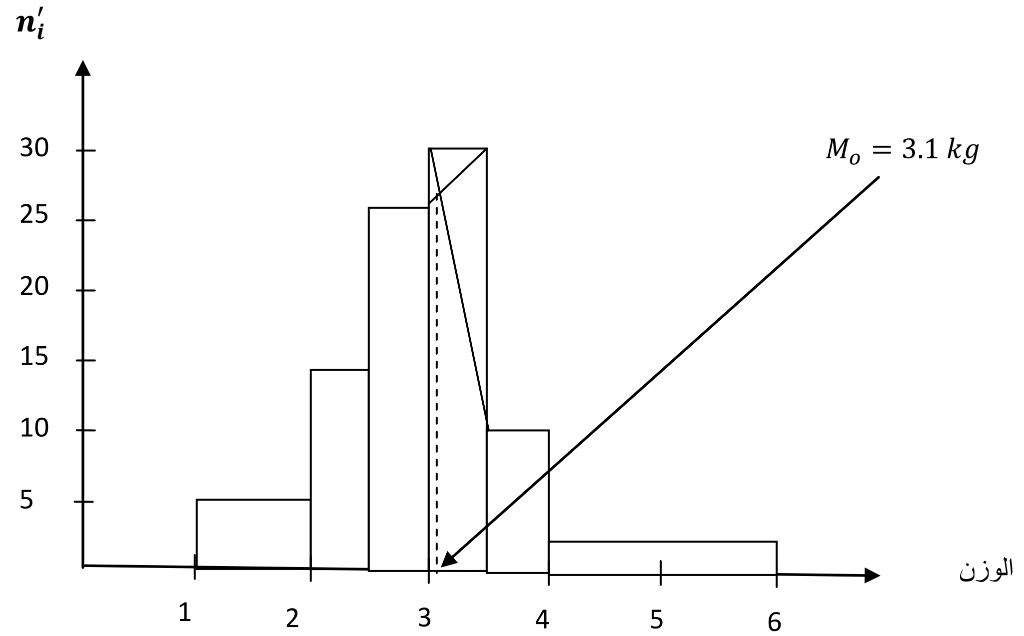
− حساب الوسيط بطريقة المد الداخلي:

$$M_e = Lim_{M_e} + \left[\frac{50 - F^{\uparrow}_{M_{e-1}}\%}{f_{M_e}\%}\right] \times A_{M_e} = 2.5 + \left[\frac{50-25}{26}\right] \times 0.5 = 2.981\ kg$$

الشرح: 50% من المواليد الرضع أوزانهم أقل من 2,981 كلغ و50% المتبقية أوزانهم أكبر من2,981 كلغ.

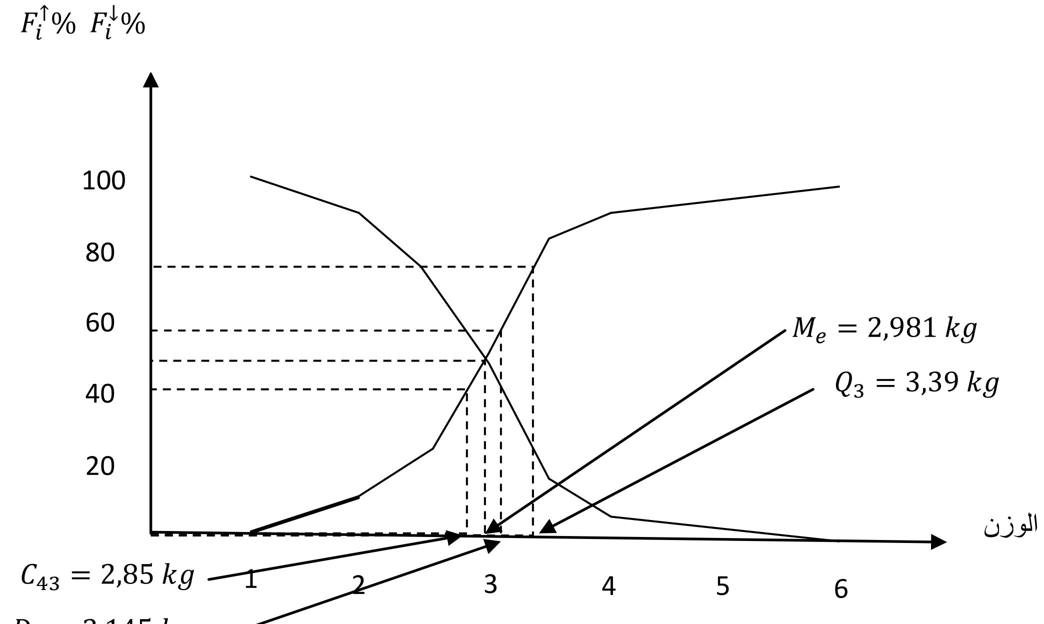
2- تحديد قيمتي المنوال والوسيط بيانيا:

أ- المنوال: عن طريق المدرج التكراري بعد تعديل التكرارات.



الشكل(3-3): التوزيع التكراري لعدد الرضع حسب الوزن عند الولادة في أحد العيادات خلال سنة 2006

ب- الوسيط: عن طريق المنحنى التجميعي الصاعد والنازل، وبما أن التكرارت نسبية مئوية نرسم المنحنى التجميعي الصاعد والنازل بدلالة $F_i^{\uparrow}\%$ و $F_i^{\downarrow}\%$



الشكل(3-4): المنحنى المتجمع الصاعد والنازل النسبي للرضع حسب الوزن عند الولادة في أحد العيادات لسنة 2006

3- حساب الربيع الثالث، العشير السادس، المئوي 43:

3-1- حسابيا:

أ- الربيع الثالث:

- تحديد الفئة الربيعية الثالثة: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{3}{4}$ ، وبما أن التكرارات نسبية فإن الفئة الربيعية الثالثة هي أول فئة تكرارها المتجمع الصاعد النسبي أكبر أو يساوي 75% أي:

$$F_{Q_3}^{\uparrow} \geq 75\%$$

ومنه الفئة الربيعية الثالثة هي: $[3 - 3{,}5[$

- حساب الربيع الثالث بطريقة المد الداخلي:

$$Q_3 = Lim_{Q_3} + \left[\frac{75 - F_{Q_3-1}^{\uparrow}\%}{f_{Q_3}\%}\right] \times A_{Q_3} = 3 + \left[\frac{75-51}{31}\right] \times 0.5 = 3{,}39\ kg$$

الشرح:

هناك 75% من المواليد الرضع أوزانهم أقل من 3,39 كلغ و 25% المتبقية أوزانهم أكبر من 3,39 كلغ.

ب- العشير السادس:

- تحديد الفئة العشير السادس: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{6n}{10}$ ، وبما أن التكرارات نسبية فإن الفئة العشير السادس هي أول فئة تكرارها المتجمع الصاعد النسبي أكبر أو يساوي 60% أي:

$$F_{D_6}^{\uparrow} \geq 60\%$$

ومنه الفئة العشير السادس هي: $[3 - 3{,}5[$

- حساب العشير السادس بطريقة المد الداخلي:

$$D_6 = Lim_{D_6} + \left[\frac{60 - F_{D_6-1}^{\uparrow}\%}{f_{D_6}\%}\right] \times A_{D_6} = 3 + \left[\frac{60-51}{31}\right] \times 0.5 = 3{,}145\ kg$$

الشرح:

هناك 60% من المواليد الرضع أوزانهم أقل من 3,145 كلغ و40% المتبقية أوزانهم أكبر من 3,145 كلغ.

ج- المئوي 43:

- تحديد فئة المئوي 43: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{43}{100}$ ، وبما أن التكرارات نسبية فإن فئة المئوي 43 هي أول فئة تكرارها المتجمع الصاعد النسبي أكبر أو يساوي 43% أي:

$$F_{C_{43}}^{\uparrow} \geq 43\%$$

ومنه فئة المئوي 43 هي: $[2{,}5 - 3[$

- حساب المئوي 43 بطريقة المد الداخلي:

$$C_{43} = Lim_{C_{43}} + \left[\frac{43 - F^{\uparrow}_{C_{43-1}}\%}{f_{C_{43}}\%}\right] \times A_{C_{43}} = 2,5 + \left[\frac{43-25}{26}\right] \times 0.5 = 2,85\ kg$$

الشرح:

هناك 43% من المواليد الرضع أوزانهم أقل من 2,85 كلغ و57% المتبقية أوزانهم أكبر من 2,85 كلغ.

3-2- بيانيا: عن طريق المنحنى التجميعي الصاعد والنازل كما يلي:

- **الربيع الثالث:** نقوم برسم المستقيم الأفقي ذو القيمة 75% الواقعة على المحور العمودي حتى يتقاطع مع المنحنى التجميعي الصاعد ثم نسقط ذلك التقاطع على المحور الأفقي فنحصل على قيمة الربيع الثالث كما هو مبين في البيان السابق.

- **العشير السادس:** نقوم برسم المستقيم الأفقي ذو القيمة 60% الواقعة على المحور العمودي حتى يتقاطع مع المنحنى التجميعي الصاعد ثم نسقط ذلك التقاطع على المحور الأفقي فنحصل على قيمة العشير السادس كما هو مبين في البيان السابق.

- **المئوي 43:** نقوم برسم المستقيم الأفقي ذو القيمة 43% الواقعة على المحور العمودي حتى يتقاطع مع المنحنى التجميعي الصاعد ثم نسقط ذلك التقاطع على المحور الأفقي فنحصل على قيمة المئوي 43 كما هو مبين في البيان السابق.

حل التمرين الثاني:

1- عرض هذه البيانات في جدول تكراري ثم حساب التكرارات النسبية:

نقوم بترتيب البيانات كما يلي:

3	2	2	2	2	2	2	2	2	1
3	3	3	3	3	3	3	3	3	3
4	4	4	4	4	4	4	4	4	3
5	5	5	5	5	4	4	4	4	4
8	8	7	7	7	6	6	6	6	5

ثم نعرض البيانات في الجدول التكراري كما يلي:

عدد الغرف X_i	عدد السكنات n_i	التكرار النسبي f_i	$n_i \times X_i$	$n_i(X_i - 5)$
1	1	0,02	1	−4
2	8	0,16	16	−24
3	12	0,24	36	−24
4	14	0,28	56	−14
5	6	0,12	30	0
6	4	0,08	24	4
7	3	0,06	21	6
8	2	0,04	16	6
المجموع	50	1	200	−50

2− حساب العدد المتوسط للغرف في المسكن الواحد:

أ− على السلسلة: $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{\sum_{i=1}^{50} X_i}{50} = \frac{4+2+4+6+\cdots+3+3+4+2}{50} = \frac{200}{50} = 4$

ب− على الجدول: $\bar{X} = \frac{\sum_{i=1}^{k} n_i X_i}{\sum n_i} = \frac{\sum_{i=1}^{8} n_i X_i}{50} = \frac{200}{50} = 4$

3− في دراسة مماثلة في بلدية أخرى، تبين أن:

− العدد الإجمالي للغرف هو 261 غرفة: أي أن $\sum(n_i \times y_i) = 261$

− العدد المتوسط للغرف هو 3 غرف: أي أن $\bar{y} = 3$

أ− إيجاد حجم العينة التي أجريت عليها الدراسة:

$\bar{y} = \frac{\sum(n_i \times y_i)}{\sum n_i} \Rightarrow \sum n_i = \frac{\sum(n_i \times y_i)}{\bar{y}} = \frac{261}{3} = 87$ سكن

ب− المقارنة بين نوع المسكن بين البلديتين:

من خلال العدد المتوسط للغرف في البلديتين نلاحظ أن السكنات في البلدية الأولى أكبر من السكنات في البلدية الثانية من حيث عدد الغرف.

ج− الدراسة الأكثر دقة:

− تكون الدراسة أكثر دقة إذا كان حجم العينة كبيرا، وبالتالي فإن الدراسة الأكثر دقة هنا هي الدراسة الثانية لأن: حجم العينة في الدراسة الثانية (87) يفوق حجم العينة في الدراسة الأولى (50).

4− إيجاد العدد المتوسط الحقيقي للغرف إذا علمنا أن المتوسط الفرضي هو $a = 5$

$$\bar{d} = \frac{\sum_{i=1}^{k} n_i d_i}{\sum n_i} = \frac{\sum_{i=1}^{8} n_i d_i}{72} = \frac{-50}{50} = -1$$

حساب المتوسط الحسابي الأصلي: $\bar{X} = \bar{d} + a = -1 + 5 = 4$

إذن المتوسط الحقيقي لهذه البيانات هو 4 غرف.

5- حساب المنوال مع شرح النتيجة:

يعرف المنوال على أنه القيمة الأكثر تكرارا في التوزيع وبالتالي فإن المنوال يساوي 4 أي: $M_o = 4$

الشرح: أغلبية السكنات في هذه البلدية تحتوي على 4 غرف.

6- حساب الوسيط مع شرح النتيجة:

بما أن عدد التكرارات هو 50 أي أنه عدد زوجي، فإن الوسيط هنا هو متوسط القيمة التي رتبتها $\frac{n}{2}$ والقيمة التي رتبتها $\frac{n}{2}+1$ أي:

$$M_e = \frac{X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}}{2} = \frac{X_{\left(\frac{50}{2}\right)} + X_{\left(\frac{50}{2}+1\right)}}{2} = \frac{X_{(25)} + X_{(26)}}{2} = \frac{4+4}{2} = 4$$

الشرح: 50% من السكنات عدد الغرف فيها أقل أو يساوي 4 غرف و50% من السكنات المتبقية عدد الغرف فيها أكبر أو يساوي 4 غرف.

7- حساب الربيع الأول والعشير الرابع والمئوي67 مع شرح النتائج:

أ- حساب الربيع الأول:

رتبة الربيع الأول هي: $\frac{n+1}{4} = \frac{50+1}{4} = 12{,}75$

ومنه قيمة الربيع الأول هي المتوسط الحسابي للقيمة التي ترتيبها 12 وهي 3 والقيمة التي ترتيبها 13 وهي 3

أي: $$Q_1 = X_{\left(\frac{n+1}{4}\right)} = X_{\left(\frac{50+1}{4}\right)} = X_{(12{,}75)} = \frac{X_{(12)} + X_{(13)}}{2} = \frac{3+3}{2} = 3$$

الشرح: 25% من السكنات عدد الغرف فيها أقل أو يساوي 3 غرف و50% من السكنات المتبقية عدد الغرف فيها أكبر أو يساوي 3 غرف.

ب- حساب العشير الرابع:

رتبة العشير الرابع هي: $\frac{4(n+1)}{10} = \frac{4(50+1)}{10} = 20{,}4$

ومنه قيمة العشير الرابع هي المتوسط الحسابي للقيمة التي ترتيبها 20 وهي 3 والقيمة التي ترتيبها 21 وهي 3

أي: $$D_4 = X_{\left(\frac{4(n+1)}{10}\right)} = X_{\left(\frac{4(50+1)}{10}\right)} = X_{(20{,}4)} = \frac{X_{(20)} + X_{(21)}}{2} = \frac{3+3}{2} = 3$$

الشرح: 40% من السكنات عدد الغرف فيها أقل أو يساوي 3 غرف و60% من السكنات المتبقية عدد الغرف فيها أكبر أو يساوي 3 غرف.

ج- حساب المئوي 67:

رتبة المئوي 67 هي: $\frac{67(n+1)}{100} = \frac{67(50+1)}{100} = 34{,}17$

ومنه قيمة المئوي 67 هي المتوسط الحسابي للقيمة التي ترتيبها 34 وهي 4 والقيمة التي ترتيبها 35 وهي 4

أي: $C_{67} = X_{\left(\frac{67(n+1)}{100}\right)} = X_{\left(\frac{67(50+1)}{100}\right)} = X_{(34,17)} = \frac{X_{(34)} + X_{(35)}}{2} = \frac{4+4}{2} = 4$

الشرح: 67% من السكنات عدد الغرف فيها أقل أو يساوي 4 غرف و33% من السكنات المتبقية عدد الغرف فيها أكبر أو يساوي 4 غرف.

حل التمرين الثالث:

السعر (دج)	عدد الأسواق	C_i	n_iC_i	$n_i\log(c_i)$	$\frac{n_i}{c_i}$	$n_ic_i^2$
]30 – 20]	1	25	25	1,398	0,04	625
]50 – 30]	3	40	120	4,806	0,075	4800
]80 – 50]	4	65	260	7,252	0,061	16900
]120 – 80]	2	100	200	4	0,02	20000
المجموع	**10**	/	**605**	**17,456**	**0,196**	**42325**

حساب السعر المتوسط للكيلوغرام من البرتقال باستعمال:

أ- المتوسط الحسابي:

$$\bar{X} = \frac{\sum_{i=1}^{k} n_iC_i}{n} = \frac{\sum_{i=1}^{4} n_iC_i}{10} == \frac{605}{10} = 60,5\ kg$$

ب- المتوسط الهندسي:

$$log(\bar{X}_{\mathrm{G}}) = \frac{1}{n}\sum_{i=1}^{k} n_i\log(c_i) = \frac{1}{10}\sum_{i=1}^{4} n_i\log(c_i) = \frac{1}{10}(17,456) = 1,7456$$

$$\Rightarrow \bar{X}_{\mathrm{G}} = 10^{\left(\frac{1}{n}\sum_{i=1}^{n} n_i\log(x_i)\right)} = 10^{1,7456} = 55,667\ kg$$

ج- المتوسط التوافقي:

$$\bar{X}_{\mathrm{H}} = \frac{\sum_{i=1}^{k} n_i}{\sum_{i=1}^{k}\frac{n_i}{c_i}} = \frac{\sum_{i=1}^{4} n_i}{\sum_{i=1}^{4}\frac{n_i}{c_i}} = \frac{10}{0,196} = 51,02$$

د- المتوسط التربيعي:

$$\bar{X}_Q = \sqrt{\frac{\sum_{i=1}^{k} n_ic_i^2}{\sum_{i=1}^{k} n_i}} = \sqrt{\frac{\sum_{i=1}^{4} n_ic_i^2}{\sum_{i=1}^{4} n_i}} = \sqrt{\frac{42325}{10}} = 65,058\ kg$$

نلاحظ أن: $\bar{X}_{\mathrm{H}} < \bar{X}_G < \bar{\mathrm{X}} < \bar{\mathrm{X}}_{\mathrm{Q}}$ محققة.

حل التمرين الرابع:

$n_i \log(x_i)$	التكرار	المخطط	معدل النمو الإقتصادي %
1,806	3	الثلاثي	4
3,133	4	الرباعي الأول	6,07
3,429	4	الرباعي الثاني	7,20
4,460	5	الخماسي الأول	7,80
3,702	5	الخماسي الثاني	5,50
2,526	5	الخماسي الثالث	3,20
19,056	**26**	/	**المجموع**

– **حساب النسبة السنوية المتوسطة للنمو الاقتصادي للفترة المذكورة:**

نستخدم المتوسط الهندسي لأنه الأنسب في حساب معدلات النمو:

أ– الطريقة البسيطة:

$$\bar{X}_{\mathrm{G}} = \sqrt[n]{x_1 . x_2 . \dots\dots\dots . x_n} = \sqrt[6]{4 \times 6{,}07 \times 7{,}20 \times 7{,}80 \times 5{,}50 \times 3{,}20} = 5{,}37$$

ومنه فإن: النسبة السنوية المتوسطة للنمو الاقتصادي هي 5,37%.

ب– الطريقة المرجحة:

$$log(\bar{X}_{\mathrm{G}}) = \frac{1}{n}\sum_{i=1}^{k} n_i \log(x_i) = \frac{1}{26}\sum_{i=1}^{6} n_i \log(x_i) = \frac{1}{26}(19{,}056) = 0{,}733$$

$$\Rightarrow \ \bar{X}_{\mathrm{G}} = 10^{\left(\frac{1}{n}\sum_{i=1}^{n} n_i \log(x_i)\right)} = 10^{0{,}733} = 5{,}41\ \%$$

ومنه فإن: النسبة السنوية المتوسطة للنمو الاقتصادي هي 5,41%.

نلاحظ أن الطريقة الثانية هي الأدق.

حل التمرين الخامس:

– **حساب متوسط السرعة للشاحنة:**

نستخدم المتوسط التوافقي، ومنه:

$$\bar{X}_{\mathrm{H}} = \frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}} = \frac{4}{\sum_{i=1}^{4}\frac{1}{x_i}} = \frac{4}{\frac{1}{50}+\frac{1}{150}+\frac{1}{75}+\frac{1}{100}} = \frac{4}{0{,}05} = 80\ km/h$$

إذن: متوسط السرعة للشاحنة هو 80 كلم/سا.

تمارين مقترحــــــــــــــة

التمرين الأول:

1- ماذا نعني بظاهرة النزعة المركزية في الدراسات الإحصائية؟

2- ما هي وظيفة مقاييس النزعة المركزية في تحليل نتائج الدراسة الإحصائية حول ظاهرة ما؟ أذكر أهم هذه المقاييس؟

3- لتكن السلسلة الإحصائية $x_n, \dots\dots, x_2, x_1$ لمتغير إحصائي، ولتكن العبارتين التاليتين:

$$\sum(X_{\mathrm{i}} - \overline{\mathrm{X}}) = 0 \dots\dots\dots\dots (1)$$
$$\sum(X_{\mathrm{i}} - \overline{\mathrm{X}})^2 < \sum(X_{\mathrm{i}} - \mathrm{X}_\alpha)^2 \dots\dots\dots\dots.. (2)$$

حيث: $\overline{\mathrm{X}}$ هو المتوسط الحسابي، و X_α هي أحد القيم بحيث: $\bar{X} \neq \mathrm{X}_\alpha$

المطلوب:

1- ماذا تمثل هاتين العبارتين؟

2- ما هي القراءة الإحصائية للعبارتين ومدلولهما؟ ماذا تستنتج؟

3- لتكن السلسلة الإحصائية التالية: 2، 5، 8، 11، 14، تأكد من أن العبارتين السابقتين محققتين على هذه السلسلة، حيث نأخذ مثلا $\mathrm{X}_\alpha = 7$؟

التمرين الثاني:

الدراسة الإحصائية للأوزان الخاصة بـ 50 شخصا يكونون جمعية رياضية أعطت النتائج التالية (بالكيلوغرام)، معروضة على شكل جدول كما يلي:

الجدول(3-7): أوزان الخاصة بـ 50 شخصا يكونون جمعية رياضية أعطت النتائج التالية (بالكيلوغرام)

الأوزان $_i$	$]50-35]$	$]60-50]$	$]70-60]$	$]80-70]$	$]90-80]$	$]100-90]$
نسبة الأشخاص f_i	0,06	0,14	0,26	0,30	0,14	0,10

المطلوب:

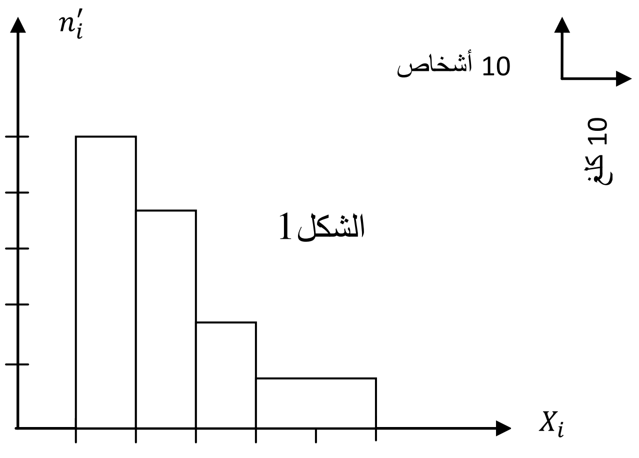
1- حدد المجتمع الإحصائي، الوحدة الإحصائية، المتغير المدروس ونوعه لهذا التوزيع؟ مثله بيانيا؟

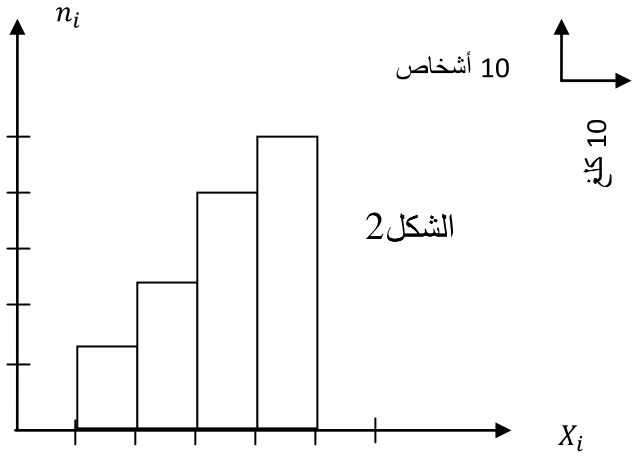
2- أحسب الوزن المتوسط، المتوسط الهندسي، المتوسط التوافقي والمتوسط التربيعي؟ مع الشرح؟

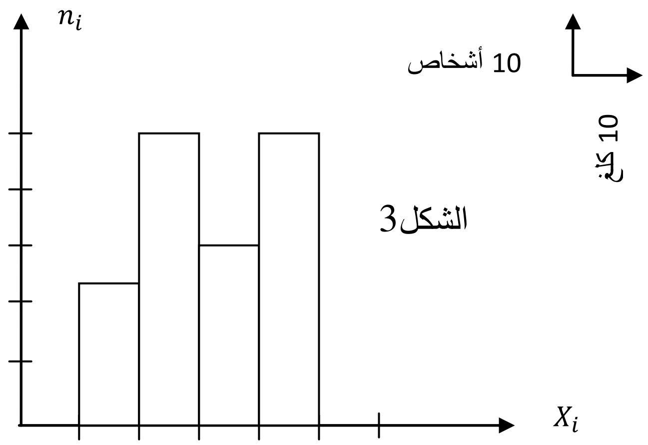
3- حدد قيمة كلا من: الوزن الوسيط، المنوال، الربيع الأول، العشير الثالث والمئوي 20 حسابيا وبيانيا؟ مع الشرح؟

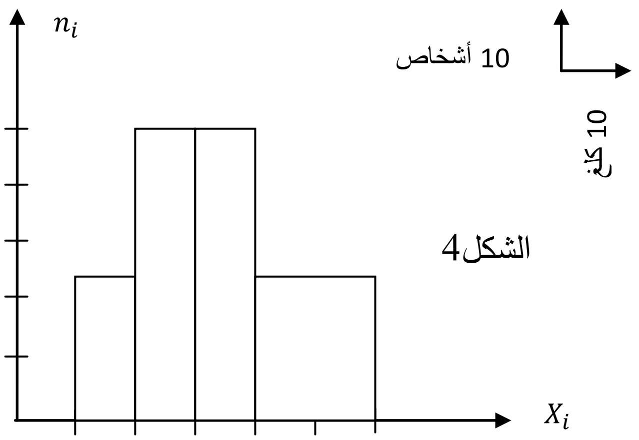
4- ما هي نسبة وعدد الأشخاص الذين يقل وزنهم عن: 67,69 كلغ، يتراوح وزنهم بين 67,69 و71,33 كلغ؟

5- حدد المنوال بيانيا على الأشكال التالية:









التمرين الثالث:

أجريت دراسة حول عدد الصغار التي تضعها الأرانب في الدفعة الواحدة، وذلك على عينة من 63 أرنبا في أحد المزارع فكانت النتائج كالتالي:

7	2	1	5	4	6	3	5	2	6	7	4	6	3	6
5	6	7	5	7	3	6	3	6	1	8	4	8	4	2
6	3	5	3	4	2	6	5	6	5	1	7	5	4	6
1	5	6	2	5	2	6	5	6	3	4	5	6	5	4
						4	6	3						

المطلوب:

1- أعرض هذه البيانات في جدول تكراري ثم مثله بيانيا؟

2- أحسب ما يلي:

أ- عدد ونسبة الأرانب التي تضع 6 صغار في الدفعة الواحدة؟

ب- عدد ونسبة الأرانب التي تضع 6 صغار فما أقل في الدفعة الواحدة؟

ج- عدد ونسبة الأرانب التي تضع أكثر من 6 صغار في الدفعة الواحدة؟

د- عدد ونسبة الأرانب التي تضع ما بين 2 و 7 صغار فما في الدفعة الواحدة؟

3- أحسب المتوسط الحسابي، الوسيط، المنوال، الربيع الثالث، العشير الثاني والمئوي 60 مع شرح النتائج؟

4- إذا كانت كل أنثى تلد تقريبا كل شهرين، وعلما أن المزرعة تقوم بتربية 1000 أنثى، ما هو عدد الصغار التي تولد في السنة؟

التمرين الرابع:

يقوم أحد الأشخاص خلال خمسة أعوام بصرف نفس المبلغ في شهر رمضان والذي قدره 20000 دج لشراء اللحم، فإذا علمت أن سعر الكيلوغرام الواحد كان كما يلي:

- السنة الأولى: 550 دج
- السنة الثانية: 580 دج
- السنة الثالثة: 660 دج
- السنة الرابعة: 810 دج
- السنة الخامسة: 900 دج

والمطلوب: حساب متوسط سعر الكيلوغرام الواحد خلال هذه الفترة؟

التمرين الخامس:

يرتفع إنتاج مادة معينة خلال أربع سنوات بالشكل التالي على التوالي: 100، 200، 300، 400، أما إنتاج مادة أخرى فيتناقص خلال نفس الفترة بالشكل التالي على التوالي: 20، 15، 10، 5

المطلوب:

1- ما هي معدلات النمو من سنة لأخرى؟

2- ما هو متوسط معدل النمو للفترة المدروسة؟ ماذا تلاحظ؟

التمرين السادس:

1- أثبت أن مجموع انحرافات القيم عن متوسطها الحسابي معدوما؟

2- بين صحة العلاقة التالية مع العلم أن $\bar{X} = M_e = M_o$:

$$2\sum n_{\mathrm{i}}\, x_i = 2Lim_{M_o}\sum n_i + A_{M_o}\sum n_i$$

الفصل الرابع

مقاييس التشتت

نتطرق من خلال هذا الفصل إلى العناصر التالية:

أولا- مقاييس التشتت المطلقة

ثانيا- مقاييس التشتت النسبية

تمارين محلولة وتمارين مقترحة

تمهيد:

رأينا أن الهدف من مقاييس النزعة المركزية هو قياس مستوى الظاهرة المدروسة، فمثلا نقيس مستوى الأجور في مؤسسة اقتصادية بواسطة الأجر المتوسط، ونقيس مستوى الدخل في بلد ما بواسطة الدخل المتوسط، ونقيس مستوى الطالب في آخر السنة بواسطة المعدل وهو متوسط حسابي...إلخ، إلا أن هذه المقاييس تشوبها مساوئ عدة، أهمها إخفاء الفروق الموجودة بين القيم، فعند إجراء مقارنة بين ظاهرتين يمكن أن يتساوى متوسطهما الحسابي، ورغم ذلك نجد أن انتشار البيانات في الظاهرتين مختلف كثيرا لأن البيانات غير متجانسة، لهذا وجدت مقاييس أخرى تعطينا فكرة عن مدى تباعد البيانات عن بعضها البعض، تسمى هذه المقاييس بمقاييس التشتت.

مثال(4−1):

إذا كان لدينا مجموعتان من الموظفين، وكل مجموعة تتكون من خمسة أشخاص حيث الأجور الشهرية (دج) لكل فرد في المجموعتين هي كالآتي:

- المجموعة الأولى: 30000 ، 31000 ، 32000 ، 33000 ، 34000.
- المجموعة الثانية: 16000 ، 20000 ، 24000 ، 44000 ، 56000.

لو قمنا بحساب المتوسط الحسابي لكل مجموعة، نجد أنه متساوي في كل منهما ويساوي 32000، ومع ذلك أجور المجموعة الأولى أكثر تجانسا من أجور المجموعة الثانية، من أجل ذلك لجأ الإحصائيون إلى استخدام مقاييس أخرى لقياس مدى تجانس البيانات أو مدى انتشار البيانات حول مقياس النزعة المركزية، ويمكن استخدامها في المقارنة بين مجموعتين أو أكثر من البيانات، ومن هذه المقاييس، مقاييس التشتت و مقاييس الشكل.

سنتطرق في هذا الفصل إلى مقاييس التشتت على أن نتطرق إلى مقاييس الشكل في الفصل اللاحق.

أولا- مقاييس التشتت المطلقة:

هذا النوع من المقاييس يقيس التشتت بقيمة مطلقة أي بمقادير لها وحدة قياس وهي نفس وحدة قياس المتغير الإحصائي موضوع الدراسة، وهناك العديد من مقاييس التشتت المطلقة تتفاوت أو تختلف في طريقة الحساب، المعنى الاقتصادي والإحصائي، الدقة في قياس التشتت، نذكر منها:

- المدى E، المدى الربيعي IQ، الانحراف الربيعي E_Q، الانحراف المتوسط بالنسبة للمتوسط الحسابي $EM_{\bar{X}}$، الانحراف المتوسط بالنسبة للوسيط EM_{M_e}، التباين $V(X)$، والانحراف المعياري $\delta(X)$.

1- المدى E:

هو الفرق بين أكبر قيمة وأقل قيمة في البيانات (السلسلة أو الجدول) أي:

- القانون في حالة سلسلة إحصائية أو حالة توزيع تكراري: $\boxed{E = Max(X_i) - Min(X_i)}$

ملاحظة: نلاحظ أن المدى يقيس التشتت بدلالة قيمتين فقط، وهما قيمتان متطرفتان لا تعكسان حقيقة الظاهرة المدروسة فهو لا يعطي القياس الحقيقي للتشتت.

2- المدى الربيعي IQ :

هو الفرق بين الربيع الثالث والربيع الأول أي: $\boxed{IQ = Q_3 - Q_1}$

ملاحظة: نلاحظ أن المدى الربيعي قد ابتعد عن القيم المتطرفة فهو أحسن من المدى، ولكن بقي يستعمل قيمتين فقط ويهمل باقي البيانات فهو لا يعكس حقيقة التشتت.

3- الانحراف الربيعي E_Q:

وهو نصف المدى الربيعي أي: $\boxed{E_Q = \frac{IQ}{2} = \frac{Q_3 - Q_1}{2}}$

ملاحظة: حتى الانحراف الربيعي رغم أنه يقترب من مركز البيانات إلا أنه يبقي يستعمل قيمتين فقط ويهمل باقي البيانات فهو لا يعكس حقيقة التشتت.

4- الانحراف المتوسط بالنسبة للوسيط EM_{M_e}:

هو متوسط البعد بالقيمة المطلقة لقيم المتغير الإحصائي عن الوسيط:

- القانون في حالة سلسلة إحصائية: $\boxed{EM_{M_e} = \frac{\sum_{i=1}^{n}|X_i - M_e|}{n} \ldots\ldots\ldots\ldots (1)}$

- القانون في حالة توزيع تكراري: $\boxed{EM_{M_e} = \frac{\sum_{i=1}^{k} n_i|X_i - M_e|}{\sum n_i} \ldots\ldots\ldots\ldots (2)}$

ملاحظة:

- يمكن أن نعوض $f_i = \frac{n_i}{\sum n_i}$ فتصبح العلاقة (2) كما يلي: $\boxed{EM_{M_e} = \sum_{i=1}^{k} f_i|X_i - M_e|}$

- نلاحظ أن هذا المقياس يقيس التشتت بدلالة كل البيانات إلا أنه يقيس التشتت بالنسبة للوسيط الذي ليس هو

أحسن مقاييس النزعة المركزية.

5- الانحراف المتوسط بالنسبة للمتوسط الحسابي $EM_{\bar{X}}$:

هو متوسط البعد بالقيمة المطلقة لقيم المتغير الإحصائي عن المتوسط الحسابي:

- القانون في حالة سلسلة إحصائية: $$EM_{\bar{X}} = \frac{\sum_{i=1}^{n}|X_i - \bar{X}|}{n} \dots\dots\dots\dots (1)$$
- القانون في حالة توزيع تكراري: $$EM_{\bar{X}} = \frac{\sum_{i=1}^{k} n_i|X_i - \bar{X}|}{\sum n_i} \dots\dots\dots\dots (2)$$

ملاحظة:

- يمكن أن نعوض $f_i = \frac{n_i}{\sum n_i}$ فتصبح العلاقة (2) كما يلي: $$EM_{\bar{X}} = \sum_{i=1}^{k} f_i|X_i - \bar{X}|$$
- نلاحظ أن $EM_{\bar{X}}$ يحسب التشتت بدلالة كل البيانات ويقيس التشتت حول أدق وأحسن مقاييس النزعة المركزية وهو المتوسط الحسابي، فهو إذن مقياس جيد للتشتت، إلا أن هاتين الصيغتين ستصبحان في الإحصاء الرياضي دوال ذات مدلول إحصائي وهي الدوال بالقيمة المطلقة، ونحن نعرف أن الدراسة الرياضية للدوال ذات القيمة المطلقة ثقيلة نوعا ما، وعليه نسعى إلى تسهيل المهمة بإلغاء القيمة المطلقة وتعويضها بما يكافؤها لقياس التشتت، والذي يؤدي هذا الغرض، أي مقياس للتشتت صالح وسهل الاستعمال في الإحصاء الوصفي وفي الإحصاء الرياضي، وهو الانحراف المعياري والذي رمزه $\delta(X)$.

6- التباين $V(X)$ والانحراف المعياري $\delta(X)$:

أ- التباين $V(X)$:

- القانون في حالة سلسلة إحصائية: $$V(X) = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n} \dots\dots\dots\dots (1)$$
- القانون في حالة توزيع تكراري: $$V(X) = \frac{\sum_{i=1}^{k} n_i(X_i - \bar{X})^2}{\sum n_i} \dots\dots\dots\dots (2)$$
- **ملاحظة:** يمكن أن نعوض $f_i = \frac{n_i}{\sum n_i}$ فتصبح العلاقة (2) كما يلي: $$V(X) = \sum_{i=1}^{k} f_i(X_i - \bar{X})^2$$

ب- الانحراف المعياري $\delta(X)$:

وهو الجذر التربيعي للتباين، أي:

- القانون في حالة سلسلة إحصائية: $$\delta(X) = \sqrt{V(X)} = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n}} \dots\dots\dots\dots (1)$$
- القانون في حالة توزيع تكراري: $$\delta(X) = \sqrt{V(X)} = \sqrt{\frac{\sum_{i=1}^{k} n_i(X_i - \bar{X})^2}{\sum n_i}} \dots\dots\dots\dots (2)$$

ملاحظة:

- المقياس الحقيقي للتشتت هو الانحراف المعياري وليس التباين لأن هذا الأخير هو عبارة عن قيمة إحصائية

ليس لها وحدة قياس، بينما الانحراف المعياري فهو قيمة إحصائية تعبر عن التشتت ولها وحدة قياس (نفس وحدة قياس X).

- يمكن أن نعوض $f_i = \frac{n_i}{\sum n_i}$ فتصبح العلاقة (2) كما يلي:

$$\delta(X) = \sqrt{\sum_{i=1}^{k} f_i (X_i - \bar{X})^2}$$

ثانيا- مقاييس التشتت النسبية:

إذا كنا بصدد إجراء مقارنة بين توزيعات تكرارية ليست لها نفس وحدة القياس أو ليس لهما نفس المتوسط، فمن الضروري هنا استخدام مقاييس التشتت النسبية والتي من أهمها:

1- المدى النسبي E%:

$$E\% = \frac{E}{\bar{X}} \times 100$$

2- المدى الربيعي النسبي $IQ\%$

$$IQ\% = \frac{IQ}{M_e} \times 100$$

3- الانحراف الربيعي النسبي $E_Q\%$:

$$E_Q\% = \frac{E_Q}{M_e} \times 100$$

4- الانحراف المتوسط بالنسبة للوسيط النسبي EM_{M_e}:

$$EM_{M_e}\% = \frac{EM_{M_e}}{M_e} \times 100$$

5- الانحراف المتوسط بالنسبة للمتوسط الحسابي النسبي $EM_{\bar{X}}$:

$$EM_{\bar{X}}\% = \frac{EM_{\bar{X}}}{\bar{X}} \times 100$$

6- الانحراف المعياري النسبي (معامل الاختلاف) CV :

$$CV = \frac{\delta(X)}{\bar{X}} \times 100$$

ملاحظة: إذا كان المتغير الإحصائي مستمر فإننا نعوض X_i بمراكز الفئات C_i في كل المعادلات السابقة.

مثال(4−2):

البيانات التالية تمثل أوزان 60 طالبا بالكيلوغرام في أحد أقسام الـــــ: LMD بكلية العلوم الاقتصادية وعلوم التسيير بجامعة سطيف:

الجدول(4−1): أوزان 60 طالبا بالكيلوغرام في أحد أقسام الـ: LMD بكلية العلوم الاقتصادية وعلوم التسيير بجامعة سطيف

أوزان الطلبة X_i	عدد الطلبة n_i
$]55-50]$	2
$]60-55]$	5
$]65-60]$	12
$]70-65]$	16
$]75-70]$	14
$]80-75]$	8
$]85-80]$	3
المجموع $\sum n_i$	60

المطلوب:

1− حساب المدى المطلق والنسبي؟

2− حساب المدى الربيعي المطلق والنسبي؟

3− حساب الإنحراف الربيعي المطلق والنسبي؟

4− الانحراف المتوسط عن المتوسط الحسابي المطلق والنسبي؟

5− الانحراف المتوسط عن الوسيط المطلق والنسبي؟

6− التباين، الانحراف المعياري ؟

7− في دراسة مماثلة عن أوزان الطلبة بجامعة أخرى تحصلنا على النتائج التالية: $\bar{X}=75$ ، $\delta(X)=12$

− قارن بين تشتت الأوزان في الدراستين؟

الحل:

أوزان الطلبة X_i	n_i	C_i	$N_i^{\uparrow}$	$n_i \times C_i$	$n_i\lvert C_i-\bar{X}\rvert$	$n_i\lvert C_i-M_e\rvert$	$n_i(C_i-\bar{X})^2$
]55 – 50]	2	52,5	2	105	31.84	31.88	506,8928
]60 – 55]	5	57,5	7	287.5	54.6	54.7	596,232
]65 – 60]	12	62,5	19	750	71,04	71.28	420,5568
]70 – 65]	16	67,5	35	1080	14,72	15.04	13,5424
]75 – 70]	14	72,5	49	1015	57,12	56.84	233,0496
]80 – 75]	8	77,5	57	620	72,64	72.48	659,5712
]85 – 80]	3	82,5	60	247,5	42,24	42.18	594,7392
المجموع $\sum n_i$	60	/	/	4105	344,2	344.4	3024,584

1– حساب المدى:

أ– المطلق:

$$\bar{X} = \frac{\sum_{i=1}^{7} n_i c_i}{n} = \frac{4105}{60} = 68,42$$

$$E = Max(X_i) - Min(X_i) = 85 - 50 = 35$$

الفرق ما بين أكبر وأصغر وزن يقدر بـ: 35 كلغ.

ب– النسبي:

$$E\% = \frac{E}{\bar{X}} \times 100 = \frac{35}{68,42} \times 100 = 51,15\%$$

2– حساب المدى الربيعي:

أ– المطلق:

$$IQ = Q_3 - Q_1$$

نقوم بحساب Q_1 كمايلي:

تحديد الفئة الربيعية الأولى: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{n}{4}$ ، أي:

$$N_{Q_1}^{\uparrow} \geq \left(\frac{n}{4} = \frac{60}{4} = 15\right)$$

ومنه الفئة الربيعية الأولى هي: $[60 - 65[$

– حساب الربيع الأول بطريقة المد الداخلي:

$$Q_1 = Lim_{Q_1} + \left[\frac{\frac{n}{4} - N_{Q_1-1}^{\uparrow}}{n_{Q_1}}\right] \times A_{Q_1} = 60 + \left[\frac{15-7}{12}\right] \times 5 = 63,33$$

نقوم بحساب Q_3 كمايلي:

تحديد الفئة الربيعية الثالثة: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{3n}{4}$ ، أي:

$$N_{Q_3}^{\uparrow} \geq \left(\frac{3n}{4} = \frac{3(60)}{4} = 45\right)$$

ومنه الفئة الربيعية الثالثة هي: $[70 - 75[$

- حساب الربيع الثالث بطريقة المد الداخلي:

$$Q_3 = Lim_{Q_3} + \left[\frac{\frac{3n}{4} - N^{\uparrow}_{Q_3-1}}{n_{Q_3}}\right] \times A_{Q_3} = 70 + \left[\frac{45-35}{14}\right] \times 5 = 73{,}57$$

ومنه: $$IQ = 73{,}57 - 63{,}33 = 10{,}24$$

الشرح: طول المجال الذي تنتشر فيه الأوزان المتوسطة لنصف عدد الطلبة هو: 10,24 كلغ.

ب- النسبي:

نقوم بحساب الوسيط في حالة البيانات المبوبة فنجده يساوي: 68,44 كلغ، أي: $M_e = 68.44$

$$IQ\% = \frac{IQ}{M_e} \times 100 = \frac{10.24}{68{,}44} \times 100 = 14{,}96\%$$

3- حساب الانحراف الربيعي E_Q:

أ- المطلق:

وهو نصف المدى الربيعي أي: $$E_Q = \frac{IQ}{2} = \frac{Q_3 - Q_1}{2} = \frac{73{,}57 - 63{,}33}{2} = 5{,}12$$

ب- النسبي:

$$E_Q\% = \frac{E_Q}{M_e} \times 100 = \frac{5{,}12}{68{,}44} \times 100 = 7{,}48\%$$

4- الانحراف المتوسط عن المتوسط الحسابي:

$$\bar{X} = \frac{\sum_{i=1}^{7} n_i c_i}{n} = \frac{4105}{60} = 68{,}42$$

أ- المطلق: $$EM_{\bar{X}} = \frac{\sum_{i=1}^{k} n_i |X_i - \bar{X}|}{n} = \frac{344{,}2}{60} = 5{,}74$$

يقدر الانحراف المتوسط عن المتوسط الحسابي لأوزان الطلبة بـــ: 5,74 كلغ.

ب- النسبي: $$EM_{\bar{X}}\% = \frac{EM_{\bar{X}}}{\bar{X}} \times 100 = \frac{5{,}74}{68{,}42} \times 100 = 8{,}39\%$$

5- الانحراف المتوسط عن الوسيط:

$$M_e = 68{,}44$$

أ- المطلق: $$EM_{M_e} = \frac{\sum_{i=1}^{k} n_i |X_i - M_e|}{n} = \frac{344{,}4}{60} = 5{,}74$$

يقدر الانحراف المتوسط عن الوسيط لأوزان الطلبة بـــ: 5,74 كلغ.

ب- النسبي: $$EM_{M_e}\% = \frac{EM_{M_e}}{M_e} \times 100 = \frac{5{,}74}{68{,}42} \times 100 = 8{,}39\%$$

6-حساب التباين والانحراف المعياري:

أ- التباين: $$V(X) = \frac{\sum_{i=1}^{k} n_i (X_i - \bar{X})^2}{n} = \frac{3024{,}584}{60} = 50{,}41$$

ب- الانحراف المعياري: $\delta(X) = \sqrt{\frac{\sum_{i=1}^{k} n_i(C_i-\bar{X})^2}{n}} = \sqrt{V(X)} = \sqrt{50.41} = 7.1$

الفرق المعياري بين الأوزان الحقيقية للطلبة والوزن المتوسط يقدر بــ: 7,1 كلغ.

7- المقارنة بين تشتت الأوزان في الدراستين:

بما أن: $\bar{X}_1 \neq \bar{X}_2$ فإننا نستخدم معامل الاختلاف CV.

أ- الدراسة الأولى:

الانحراف المعياري النسبي (معامل الاختلاف): $CV_1 = \frac{\delta(X)}{\bar{X}} \times 100 = \frac{7,1}{68,42} \times 100 = 10,38\%$

ب- الدراسة الثانية:

الانحراف المعياري النسبي (معامل الاختلاف): $CV_2 = \frac{\delta(X)}{\bar{X}} \times 100 = \frac{12}{75} \times 100 = 16\%$

نلاحظ أن تشتت الأوزان في الدراسة الثانية أكبر منه في الدراسة الأولى.

تمارين محلولة للفصل الرابع

التمرين الأول:

لتكن العلامات التالية المحصل عليها من طرف طالبين في 5 إمتحانات:

الطالب الأول: 8، 10، 13، 11، 15.

الطالب الثاني: 10، 12، 12، 10، 13.

المطلوب:

1- قارن بين مستوى الطالبين؟

2- أدرس التشتت في علامات كل طالب؟

التمرين الثاني:

أجريت دراسة إحصائية حول الأجور في مؤسستبن فأعطت النتائج التالية:

المؤسسة الأولى: $\overline{X}_1 = 36\ \mathrm{DA}$ و $\delta(X_1) = 9\ \mathrm{DA}$

المؤسسة الثانية: $\overline{X}_2 = 50\ \mathrm{DA}$ و $\delta(X_2) = 10\ \mathrm{DA}$

المطلوب: قارن مستوى وتشتت الأجور بين المؤسستين؟ أيهما أحسن؟ لماذا؟

التمرين الثالث:

1- أثبت أن: $V(X) = \frac{\sum X_i^2}{n} - \overline{X}^2$ واستنتج أن: $\delta(X) = \sqrt{\overline{X}_Q^2 - \overline{X}^2}$

2- بهدف دراسة معدلات التلاميذ للسنة الخامسة إبتدائي أخذت عينتان من التلاميذ من مدرستين مختلفتين، فأعطتا النتائج التالية:

العينة الأولى: $\sum_{i=1}^{40} Y_i = 280$ و $\sum_{i=1}^{40} Y_i^2 = 2100$

العينة الثانية: $\sum_{i=1}^{50} X_i = 300$ و $\sum_{i=1}^{50} X_i^2 = 1950$

أ- أوجد الوسط الحسابي والانحراف المعياري والمتوسط التربيعي لكل عينة؟ مع الشرح؟

ب- قارن بين مستوى وتشتت معدلات التلاميذ في المدرستين؟ أيهما الأحسن؟ لماذا؟

ج- أحسب متوسط معدلات جميع التلاميذ؟

التمرين الرابع:

عند مراقبة الوصول إلى مقر العمل لمجموعة من العمال (100 عامل) في إحدى المؤسسات تم الحصول على المعلومات التالية: (الوحدة: الدقيقة)

الجدول(4−2): توزيع مجموعة من 100 عامل حسب زمن تأخرهم عن العمل

زمن التأخر X_i	$[5 - 10[$	$[10 - 15[$	$[15 - 20[$	$[20 - 30[$	$[30 - 40[$	$[40 - 45[$
عدد العمال n_i	10	18	40	20	8	04

المطلوب:

1− حدد المجتمع الإحصائي، الوحدة الإحصائية، المتغير المدروس ونوعه لهذا التوزيع؟

2− أحسب كلا من: المتوسط الحسابي، الوسيط؟ مع الشرح؟

3− أحسب كلا من: المدى المطلق والنسبي، المدى الربيعي المطلق والنسبي، الانحراف المطلق بالنسبة للوسيط المطلق والنسبي؟ مع الشرح؟

4− إذا علمت أن: المتوسط التربيعي لزمن التأخر يساوي 21,25 دقيقة، أحسب التباين والانحراف المعياري؟

5− نفرض أن زمن التأخر الوسيط في مؤسسة ثانية يقدر بـ: 17,75 دقيقة وأن الانحراف المتوسط بالنسبة للوسيط يساوي 4 دقائق، قارن مستوى التأخر والتشتت في المؤسستين، ما هي القراءة الإحصائية لذلك؟

التمرين الخامس:

أخذت عينات من عمال أربعة شركات، فكانت النتائج التالية:

الجدول(4−3): توزيع عينات من عمال أربع شركات حسب أجورهم.

الشركات	A	B	C	D
عدد العمال	160	130	200	400
مجموع الأجور (دج)	80000	78000	148000	160000
التباين	900	2500	625	400

المطلوب:

1− أحسب الأجر المتوسط والانحراف المعياري لكل مؤسسة؟

2− ما هي الشركة الأكثر تجانسا والشركة الأقل تجانسا في هذه الدراسة؟

3− أحسب متوسط الأجر الإجمالي لكل الشركات؟ المتوسط التربيعي لكل شركة؟

الحلــــــــــــول

حل التمرين الأول:

الطالب الأول: 8، 10، 13، 11، 15.

الطالب الثاني: 10، 12، 12، 10، 13.

1- المقارنة بين مستوى الطالبين:

نقارن مستوى الطالبين بواسطة المتوسط الحسابي باعتباره أدق مقاييس النزعة المركزية.

الطالب الأول: $\bar{X}_1 = \frac{\sum_{i=1}^{5} X_i}{n} = \frac{8+10+13+11+15}{5} = \frac{57}{5} = 11{,}4$

الطالب الثاني: $\bar{X}_2 = \frac{\sum_{i=1}^{5} X_i}{n} = \frac{10+12+12+10+13}{5} = \frac{57}{5} = 11{,}4$

بما أن: $\bar{X}_1 = \bar{X}_2$ فإن الطالبين لديهما نفس المستوى على العموم.

2- دراسة التشتت في علامات كل طالب:

بما أن: $\bar{X}_1 = \bar{X}_2$ فإننا نستخدم مقاييس التشتت المطلقة، وأحسن مقياس يمكن استخدامه هو الانحراف المعياري $\delta(X)$.

الطالب الأول: $\delta(X_1) = \sqrt{\frac{\sum_{i=1}^{5}(X_i-\bar{X})^2}{n}} = \sqrt{\frac{(8-11{,}4)^2+\cdots\ldots\ldots+(15-11{,}4)^2}{5}} = 2{,}417$

الطالب الثاني: $\delta(X_2) = \sqrt{\frac{\sum_{i=1}^{5}(X_i-\bar{X})^2}{n}} = \sqrt{\frac{(10-11{,}4)^2+\cdots\ldots\ldots+(13-11{,}4)^2}{5}} = 1{,}342$

بما أن $\delta(X_1) > \delta(X_2)$: فإنه يمكننا القول على العموم أن تشتت العلامات للطالب الثاني أقل من تشتت علامات الطالب الأول، أي أن الفوارق في توزيع العلامات هي أقل عند الطالب الثاني وبالتالي علامات الطالب الثاني أكثر تجانس من علامات الطالب الأول.

حل التمرين الثاني:

المؤسسة الأولى: $\bar{X}_1 = 39\ DA$ و $\delta(X_1) = 9\ DA$

المؤسسة الثانية: $\bar{X}_2 = 50\ DA$ و $\delta(X_2) = 10\ DA$

- المقارنة بين مستوى وتشتت الأجور بين المؤسستين:

أ- مقارنة مستوى الأجور: نلاحظ أن $\bar{X}_1 < \bar{X}_2$ وعليه فإنه على العموم مستوى الأجور في المؤسسة الثانية أكبر مما هي عليه في المؤسسة الأولى.

ب- مقارنة تشتت الأجور:

بما أن $\bar{X}_1 \neq \bar{X}_2$ فإننا نستخدم مقاييس التشتت النسبية لمقارنة تشتت الأجور في المؤسستين، حيث نقوم بحساب كلا من CV_1 و CV_2 كما يلي:

$$CV_1 = \delta(X_1)\% = \frac{\delta(X_1)}{\bar{X}_1} \times 100 = \frac{9}{36} \times 100 = 25\%$$

$$CV_2 = \delta(X_2)\% = \frac{\delta(X_2)}{\bar{X}_2} \times 100 = \frac{10}{50} \times 100 = 20\%$$

بما أن $CV_1 > CV_2$ فإننا نقول أن تشتت الأجور في المؤسسة الأولى أكبر مما هي عليه في المؤسسة الثانية، أي أن الفوارق في توزيع الأجور أكبر في المؤسسة الأولى، وعليه فإن أجور المؤسسة الثانية أكثر تجانس من أجور المؤسسة الاولى.

- **المؤسسة الثانية هي الأحسن،** لأن المؤسسة الأولى توزع أجور منخفضة والفوارق بين الأجور كبيرة، والمؤسسة الثانية توزع أجور مرتفعة والفوارق بين الأجور قليلة.

حل التمرين الثالث:

1- إثبات أن: $\boldsymbol{V(X) = \frac{\sum X_i^2}{n} - \bar{X}^2}$ **واستنتاج أن:** $\boldsymbol{\delta(X) = \sqrt{\bar{X}_Q^2 - \bar{X}^2}}$

$$V(X) = \frac{\sum(X_i-\bar{X})^2}{n} = \frac{\sum\left(X_i^2-2X_i\bar{X}+\bar{X}^2\right)}{n} = \frac{\sum X_i^2-2\bar{X}\sum X_i+\sum \bar{X}^2}{n}$$

ولدينا: $\sum \bar{X}^2 = n\bar{X}^2$ و $\bar{X} = \frac{\sum X_i}{n} \Rightarrow \sum X_i = n\bar{X}$

ومنه: $V(X) = \frac{\sum X_i^2-2n\bar{X}^2+n\bar{X}^2}{n} = \frac{\sum X_i^2-n\bar{X}^2}{n} = \frac{\sum X_i^2}{n} - \frac{n\bar{X}^2}{n}$

أي: $\boldsymbol{V(X) = \frac{\sum X_i^2}{n} - \bar{X}^2}$

ولدينا: $\delta(X) = \sqrt{V(X)}$ و المتوسط التربيعي: $\bar{X}_Q = \sqrt{\frac{\sum X_i^2}{n}}$ أي: $\frac{\sum X_i^2}{n} = \bar{X}_Q^2$ ومنه:

$$\delta(X) = \sqrt{V(X)} = \sqrt{\frac{\sum X_i^2}{n} - \bar{X}^2} = \sqrt{\bar{X}_Q^2 - \bar{X}^2}$$

أي: $\boldsymbol{\delta(X) = \sqrt{\bar{X}_Q^2 - \bar{X}^2}}$

2- أ- إيجاد الوسط الحسابي والانحراف المعياري والمتوسط التربيعي لكل عينة مع الشرح:

العينة الأولى: $\sum_{i=1}^{40} Y_i = 280$ و $\sum_{i=1}^{40} Y_i^2 = 2100$

- **المتوسط الحسابي:** $\bar{Y} = \frac{\sum_{i=1}^{40} Y_i}{n} = \frac{280}{40} = 7$

- **الشرح:** متوسط معدلات تلاميذ العينة الأولى هو 7.

- **الانحراف المعياري:** $\delta(Y) = \sqrt{\frac{\sum Y_i^2}{n} - \bar{Y}^2} = \sqrt{\frac{2100}{40} - (7)^2} = 1{,}87$

الشرح: تشتت علامات تلاميذ العينة الأولى يقدر بـ: 1,87

- **المتوسط التربيعي:**

$$\delta(Y) = \sqrt{\bar{Y}_Q^2 - \bar{Y}^2} \Rightarrow \bar{Y}_Q^2 = \left(\delta(Y)\right)^2 + \bar{Y}^2$$

$$\Rightarrow \bar{Y}_Q = \sqrt{\left(\delta(Y)\right)^2 + \bar{Y}^2} = \sqrt{(1{,}87)^2 + (7)^2} = \sqrt{52{,}5} = 7{,}24$$

- **الشرح:** المتوسط التربيعي لمعدلات تلاميذ العينة الأولى هو7,24.

العينة الثانية: $\sum_{i=1}^{50} X_i = 300$ و $\sum_{i=1}^{50} X_i^2 = 1950$

- **المتوسط الحسابي:** $\bar{X} = \frac{\sum_{i=1}^{50} X_i}{n} = \frac{300}{50} = 6$

- **الشرح:** متوسط معدلات تلاميذ العينة الثانية هو 6.

- **الانحراف المعياري:** $\delta(X) = \sqrt{\frac{\sum_{i=1}^{50} X_i^2}{n} - \bar{X}^2} = \sqrt{\frac{1950}{50} - (6)^2} = 1{,}73$

الشرح: تشتت علامات تلاميذ العينة الثانية يقدر بـ: 1,73

- **المتوسط التربيعي:**

$$\delta(X) = \sqrt{\bar{X}_Q^2 - \bar{X}^2} \Rightarrow \bar{X}_Q^2 = \left(\delta(X)\right)^2 + \bar{X}^2$$

$$\Rightarrow \bar{X}_Q = \sqrt{\left(\delta(X)\right)^2 + \bar{X}^2} = \sqrt{(1{,}73)^2 + (6)^2} = \sqrt{39} = 6{,}24$$

- **الشرح:** المتوسط التربيعي لمعدلات تلاميذ العينة الثانية هو6,24.

ب- المقارنة بين مستوى وتشتت معدلات التلاميذ في المدرستين (العينتين):

- **مقارنة مستوى المعدلات:** نلاحظ أن $\bar{X} < \bar{Y}$ وعليه فإنه على العموم المستوى الدراسي لتلاميذ العينة الأولى أعلى مما هو عليه في تلاميذ العينة الثانية.

ب- مقارنة تشتت المعدلات:

بما أن $\bar{X} \neq \bar{Y}$ فإننا نستخدم مقاييس التشتت النسبية لمقارنة تشتت التعدلات في العينتين، حيث نقوم بحساب كلا من CV_1 و CV_2 كما يلي:

$$CV_1 = \delta(Y)\% = \frac{\delta(Y)}{\bar{Y}} \times 100 = \frac{1,87}{7} \times 100 = 26,71\%$$

$$CV_2 = \delta(X)\% = \frac{\delta(X)}{\bar{X}} \times 100 = \frac{1,73}{6} \times 100 = 28,83\%$$

بما أن $CV_1 < CV_2$ فإننا نقول أن تشتت المعدلات في العينة الأولى أقل مما هي عليه في العينة الثانية، أي أن الفوارق في توزيع المعدلات أكبر في العينة الثانية، وعليه فإن معدلات العينة الأولى أكثر تجانس من معدلات العينة الثانية.

- **العينة (المدرسة) الأولى هي الأحسن:** لأن العينة الأولى مستوى المعدلات فيها مرتفع والفوارق بين المعدلات قليلة، بينما العينة الثانية مستوى المعدلات فيها منخفض والفوارق بين المعدلات كبيرة.

ج- حساب متوسط معدلات جميع التلاميذ:

نرمز لمتوسط معدلات جميع التلاميذ بـــ: $\bar{Z}$ وبالتالي:

$$\bar{Z} = \frac{n_1\bar{Y}+n_2\bar{X}}{n_1+n_2} = \frac{40(7)+50(6)}{40+50} = 6,44$$

حل التمرين الرابع:

X_i زمن التأخر	n_i عدد العمال	C_i	$n_i \times C_i$	$N_i^{\uparrow}$	$\sum_{i=1}^{k} n_i\lvert c_i - M_e\rvert$
$]10 - 5]$	10	7,5	75	10	102,5
$]15 - 10]$	18	12,5	225	28	94,5
$]20 - 15]$	40	17,5	700	68	10
$]30 - 20]$	20	25	500	88	145
$]40 - 30]$	8	35	280	96	138
$]45 - 40]$	4	42,5	170	100	99
المجموع	100	/	1950	/	589

1- تحديد المجتمع الإحصائي، الوحدة الإحصائية، المتغير المدروس ونوعه لهذا التوزيع:

المجتمع الإحصائي	الوحدة الإحصائية	المتغير الإحصائي	نوعه
جميع عمال المؤسسة	العامل الواحد	زمن التأخر	كمي مستمر

2- حساب كلا من: المتوسط الحسابي، الوسيط مع الشرح:

أ- المتوسط الحسابي:

$$\bar{X}_1 = \frac{\sum_{i=1}^{6} n_i \times c_i}{\sum n_i} = \frac{1950}{100} = 19,50\, mn$$

الشرح: متوسط زمن التأخر للعامل الواحد هو 19,50 دقيقة.

ب- الوسيط:

- تحديد الفئة الوسيطية: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{n}{2}$ ، أي:

$$N_{M_e}^{\uparrow} \geq \left(\frac{n}{2} = \frac{100}{2} = 50\right)$$

ومنه الفئة الوسيطية هي: $[15 - 20[$

– حساب الوسيط بطريقة المد الداخلي:

$$M_e = Lim_{M_e} + \left[\frac{\frac{n}{2} - N_{M_e-1}^{\uparrow}}{n_{M_e}}\right] \times A_{M_e} = 15 + \left[\frac{50-28}{40}\right] \times 5 = 17{,}75\ mn$$

الشرح: هناك 50% من العمال زمن تأخرهم أقل من 17,75 دقيقة و 50% من العمال المتبقيين زمن تأخرهم أكبر من 17,75 دقيقة

3– حساب كلا من: المدى المطلق والنسبي، المدى الربيعي المطلق والنسبي، الانحراف المطلق بالنسبة للوسيط المطلق والنسبي مع الشرح:

أ– **المدى المطلق والنسبي:**

– **المطلق:**

$$E = Max(X_i) - Min(X_i) = 45 - 5 = 40$$

يقدر التشتت في زمن التأخر للعمال باستعمال المدى المطلق بـ 40 دقيقة.

– **النسبي:** $E\% = \frac{E}{\bar{X}} \times 100 = \frac{40}{19{,}50} \times 100 = 205{,}13\%$

يقدر التشتت في زمن التأخر للعمال باستعمال المدى النسبي بـ 205,13%.

ب– حساب المدى الربيعي المطلق والنسبي:

– **المطلق:**

$$IQ \ = Q_3 - Q_1$$

نقوم بحساب Q_1 كمايلي:

تحديد الفئة الربيعية الأولى: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{n}{4}$ ، أي:

$$N_{Q_1}^{\uparrow} \geq \left(\frac{n}{4} = \frac{100}{4} = 25\right)$$

ومنه الفئة الربيعية الأولى هي: $[10 - 15[$

– حساب الربيع الأول بطريقة المد الداخلي:

$$Q_1 = Lim_{Q_1} + \left[\frac{\frac{n}{4} - N_{Q_1-1}^{\uparrow}}{n_{Q_1}}\right] \times A_{Q_1} = 10 + \left[\frac{25-10}{18}\right] \times 5 = 14{,}16$$

نقوم بحساب Q_3 كمايلي:

تحديد الفئة الربيعية الثالثة: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{3}{4}$ ، أي:

$$N_{Q_3}^{\uparrow} \geq \left(\frac{3n}{4} = \frac{3(100)}{4} = 75\right)$$

ومنه الفئة الربيعية الثالثة هي: $[20 - 30]$

- حساب الربيع الثالث بطريق المد الداخلي:

$$Q_3 = Lim_{Q_3} + \left[\frac{\frac{3n}{4} - N_{Q_3-1}^{\uparrow}}{n_{Q_3}}\right] \times A_{Q_3} = 20 + \left[\frac{75-68}{20}\right] \times 10 = 23,50\ mn$$

ومنه:

$$IQ = 23,50 - 14,16 = 9,34\ mn$$

الشرح: يقدر التشتت في زمن التأخر للعمال باستعمال المدى الربيعي المطلق بـ 9,34 دقيقة.

- **النسبي:**

لدينا: $\mathrm{M_e} = 17,75\ mn$

$$IQ\% = \frac{IQ}{M_e} \times 100 = \frac{9,34}{17,75} \times 100 = 52,62\%$$

الشرح: يقدر التشتت في زمن التأخر للعمال باستعمال المدى الربيعي النسبي بـ 52,62%.

ج- حساب الانحراف الربيعي المطلق والنسبي:

- **المطلق:**

وهو نصف المدى الربيعي أي: $E_{\mathrm{Q}} = \frac{IQ}{2} = \frac{Q_3 - Q_1}{2} = \frac{23,50-14,16}{2} = 4,67\ mn$

الشرح: يقدر التشتت في زمن التأخر للعمال باستعمال الانحراف الربيعي المطلق بـ 4,67 دقيقة.

- **النسبي:**

$$E_{\mathrm{Q}}\% = \frac{E_Q}{M_e} \times 100 = \frac{4,67}{17,75} \times 100 = 26,31\%$$

الشرح: يقدر التشتت في زمن التأخر للعمال باستعمال الانحراف الربيعي النسبي بـ 26,31%.

د- الانحراف المتوسط عن الوسيط:

$$M_{\mathrm{e}} = 17,75$$

- **المطلق:** $EM_{M_e} = \frac{\sum_{i=1}^{k} n_i |X_i - M_e|}{n} = \frac{589}{100} = 5,89$

الشرح: يقدر التشتت في زمن التأخر للعمال باستعمال الانحراف المتوسط عن الوسيط المطلق بـ 5,89 دقيقة

- **النسبي:** $EM_{M_e}\% = \frac{EM_{M_e}}{M_e} \times 100 = \frac{5,89}{17,75} \times 100 = 33,18\%$:

الشرح: يقدر التشتت في زمن التأخر للعمال باستعمال الانحراف المتوسط عن الوسيط النسبي بـ 33,18%.

4- حساب التباين والانحراف المعياري:

$$V(X) = \overline{\mathrm{X}}_{\mathrm{Q}}^2 - \overline{\mathrm{X}}^2 = (21,25)^2 - (19,5)^2 = 71,3125$$

$$\delta(X) = \sqrt{V(X)} = \sqrt{71,3125} = 8,44\ mn$$

5- المقارنة بين مستوى التأخر والتشتت في المؤسستين، مع القراءة الإحصائية:

المؤسسة الأولى: $M_{e1} = 17,75$ و $EM_{M_{e1}} = 5,89$

المؤسسة الثانية: $M_{e2} = 17,75$ و $EM_{M_{e2}} = 4$

أ- **مقارنة مستوى التأخر:** نلاحظ أن $M_{e1} = M_{e2}$ وعليه فإنه على العموم مستوى تأخر العمال في المؤسستين متساوي.

ب- مقارنة تشتت التأخر:

بما أن $M_{e1} = M_{e2}$ فإننا نستخدم مقاييس التشتت المطلقة لمقارنة تشتت التأخر في المؤسستين.

بما أن $M_{M_{e1}} > EM_{M_{e2}}$ فإننا نقول أن تشتت التأخر في المؤسسة الأولى أكبر مما هي عليه في المؤسسة الثانية، أي أن الفوارق في زمن التأخر أكبر في المؤسسة الأولى، وعليه فإن زمن التأخر في المؤسسة الثانية أكثر تجانس من زمن التأخر المؤسسة الأولى.

القراءة الإحصائية: عمال المؤسسة الأولى والثانية لهما على العموم نفس زمن التأخر، غير أن زمن تأخر عمال المؤسسة الثانية أكثر تجانس أو تقارب من المؤسسة الأولى، أي أن المؤسسة الثانية هي الأحسن من حيث زمن التأخر.

حل التمرين الخامس:

أخذت عينات من عمال أربعة شركات، فكانت النتائج التالية:

الشركات	A	B	C	D
عدد العمال	160	130	200	400
مجموع الأجور (دج)	80000	78000	148000	160000
التباين	900	2500	625	400
الأجر المتوسط (دج)	500	600	740	400
الانحراف المعياري (دج)	30	50	25	20
المتوسط التربيعي (دج)	500,90	602,08	740,42	400,50

1- حساب الأجر المتوسط والانحراف المعياري لكل مؤسسة:

- **الأجر المتوسط:** هو حاصل قسمة مجموع الأجور على عدد العمال، مثلا بالنسبة للشركة A:

$$\bar{X}_A = \frac{\sum_{i=1}^{n} X_{i(A)}}{n_{(A)}} = \frac{\sum_{i=1}^{160} X_{i(A)}}{160} = \frac{80000}{160} = 500$$

وهكذا بالنسبة لباقي الشركات (أنظر النتائج في الجدول أعلاه).

- **الانحراف المعياري:** يساوي الجذر التربيعي للتباين، مثلا بالنسبة للشركة A:

$$\delta(X_A) = \sqrt{V(X_A)} = \sqrt{900} = 30$$

وهكذا بالنسبة لباقي الشركات (أنظر النتائج في الجدول أعلاه).

2- الشركة الأكثر تجانسا والشركة الأقل تجانسا في هذه الدراسة:

بما أن: $\bar{X}_A \neq \bar{X}_B \neq \bar{X}_C \neq \bar{X}_D$ فإننا نستخدم مقاييس التشتت النسبية لمقارنة تشتت الأجور في الشركات.

نستخدم هنا الانحراف المعياري النسبي (معامل الاختلاف) كمايلي:

$$CV_A = \delta(X_A)\% = \frac{\delta(X_A)}{\bar{X}_A} \times 100 = \frac{30}{500} \times 100 = 6\%$$

$$CV_B = \delta(X_B)\% = \frac{\delta(X_B)}{\bar{X}_B} \times 100 = \frac{50}{600} \times 100 = 8{,}33\%$$

$$CV_C = \delta(X_C)\% = \frac{\delta(X_C)}{\bar{X}_C} \times 100 = \frac{25}{740} \times 100 = 3{,}38\%$$

$$CV_D = \delta(X_D)\% = \frac{\delta(X_D)}{\bar{X}_D} \times 100 = \frac{20}{400} \times 100 = 5\%$$

ترتيب الشركات حسب التجانس:

أ- الشركة C...........3,38%

ب- الشركة D 5%

ج- الشركة A 6%

د- الشركة B 8,33%

إذن: الشركة الأكثر تجانسا هي الشركة C: أي أجور عمالها أكثر تقاربا.

الشركة الأقل تجانسا هي الشركة B: أي أجور عمالها أقل تقاربا.

3- حساب متوسط الأجر الإجمالي لكل الشركات والمتوسط التربيعي لكل شركة:

أ- حساب متوسط الأجر الإجمالي لكل الشركات:

نرمز لمتوسط الأجر الإجمالي لكل الشركات بـــ: $\bar{X}$ وبالتالي:

$$\bar{X} = \frac{n_{(A)}\bar{X}_A + n_{(B)}\bar{X}_B + n_{(C)}\bar{X}_C + n_{(D)}\bar{X}_D}{n_{(A)} + n_{(B)} + n_{(C)} + n_{(D)}} = \frac{160(500)+130(600)+200(740)+400(400)}{160+130+200+400} = \frac{466000}{890} = 523{,}6$$

ب- المتوسط التربيعي: مثلا المؤسسة A :

$$\bar{X}_{Q(A)} = \sqrt{\left(\delta\left(X_{(A)}\right)\right)^2 + \bar{X}_{(A)}^2} = \sqrt{V\left(X_{(A)}\right) + X_{(A)}^2} = \sqrt{900 + (500)^2} = 500{,}9$$

وهكذا بالنسبة لباقي الشركات (أنظر النتائج في الجدول أعلاه)

تمارين مقترحـــــة

التمرين الأول:

1- ما هي وظيفة مقاييس التشتت في تحليل نتائج الدراسة الإحصائية حول ظاهرة ما؟

2- ما هي أنواع مقاييس التشتت؟ ومتى نستخدم كل نوع؟ أذكر أهم مقاييس كل نوع؟

3- ما هو الفرق بين المدى العام والمدى الربيعي؟ لماذا لا يعتبران من أحسن مقاييس التشتت؟

4- وما هو الفرق بين المدى الربيعي والانحراف الربيعي؟

التمرين الثاني:

في آخر السنة الدراسية الماضية تم تحليل المعدلات السنوية لتلاميذ القسم النهائي لثانويتين بسطيف فكان ما يلي:

- **الثانوية الأولى:** المعدل السنوي يساوي 12، التباين يساوي 9
- **الثانوية الثانية:** المعدل السنوي يساوي 10,5، الانحراف المعياري يساوي 4,41

المطلوب:

1- قارن المستوى التعليمي في الثانويتين؟ وكذلك مدى التقارب أو التباعد بين التلاميذ داخل الثانوية الواحدة؟

2- ما هي الثانوية الأنجح من ناحية المستوى التعليمي؟ برر ذلك؟

التمرين الثالث:

في دراسة حول دخل العائلات في الجزائر لسنة 1979، وبعد تحليل النتائج تبين ما يلي:

- **منطقة الجزائر العاصمة:**
  - الدخل السنوي المتوسط للفرد يساوي 2334 دج، الانحراف المعياري يساوي 420 دج
- **المناطق الحضرية (دون العاصمة):**
  - الدخل السنوي المتوسط للفرد يساوي 2149 دج، المتوسط التربيعي يساوي 2191 دج
- **المناطق الريفية:**
  - المتوسط التربيعي يساوي 1481 دج، الانحراف المعياري يساوي 359 دج

المطلوب:

1- أحسب الانحراف المعياري في المناطق الحضرية؟ والدخل المتوسط السنوي للفرد في المناطق الريفية؟

2- قارن مستوى الدخل والتشتت بين المناطق الثلاث؟

3- أي منطقة تمتاز بأكثر عدالة في توزيع الدخل بين الأفراد؟ علل ذلك؟

التمرين الرابع:

يبين التوزيع التكراري التالي نتائج الدراسة التي قامت بها مصلحة متابعة الجودة لمؤسسة صناعة المصابيح الكهربائية على عينة من 50 مصباحا حسب مدة الحياة (مدة الاشتغال) (الوحدة الزمنية: اليوم)

الجدول(4-4): توزيع عينة من 50 مصباحا حسب مدة الحياة (مدة الاشتغال)

مدة الحياة X_i	$]50-10]$	$]80-50]$	$]90-80]$	$]100-90]$	$]110-100]$	$]150-110]$
عدد المصابيح n_i	2	4	6	18	15	5

المطلوب

1- حدد المجتمع الإحصائي، الوحدة الإحصائية، المتغير المدروس ونوعه لهذا التوزيع؟

2- أحسب كلا من: متوسط مدة الحياة اليومية للمصباح الواحد، الوسيط والمنوال، ماذا تستنتج؟

3- أحسب كلا من: المدى المطلق والنسبي، المدى الربيعي المطلق والنسبي، الانحراف المطلق بالنسبة للوسيط المطلق والنسبي؟ مع الشرح؟

4- أحسب التباين والانحراف المعياري؟

5- في دراسة ثانية على عينة أخرى حجمها يساوي ضعف حجم العينة الأولى كانت النتائج كالتالي:

متوسط مدة الحياة اليومية للمصباح الواحد يقدر بـ 94 يوما، التباين يساوي 202,705.

- قارن بين مدة الحياة المصباح والتشتت في الدراستين؟

التمرين الخامس:

أخذت عينة من الطلاب تتكون من 40 طالبا، وسجلت أوزانهم في الجدول التالي: (الوحدة: كلغ)

الجدول(4-5): توزيع عينة من 40 طالبا حسب أوزانهم

فئات الأوزان X_i	$]45-40]$	$]50-45]$	$]55-50]$	$]60-55]$	$]65-60]$
عدد الطلبة n_i	3	12	15	6	4

المطلوب:

1- حدد المجتمع الإحصائي، الوحدة الإحصائية، المتغير المدروس ونوعه لهذا التوزيع؟

2- أحسب كلا من: متوسط أوزان الطلبة، والانحراف المعياري؟

3- في دراسة ثانية تخص أطوال الطلبة (بالسنتيمتر) على نفس العينة، كانت النتائج كالتالي:

متوسط أطوال الطلبة هو 162 سم ، الانحراف المعياري 4,72 سم

أ- قارن بين التشتت في الدراستين؟ ب- أحسب على الدراسة الثانية:

- نسبة وعدد الطلبة الذين تتراوح وزنهم بين الربيع الأول والوسيط؟
- نسبة وعدد الطلبة الذين أوزنهم تقل عن العشير الثامن؟
- نسبة وعدد الطلبة الذين أوزنهم تفوق المئوي 44؟

التمرين السادس:

تمتلك أحد المؤسسات الصناعية وحدتين إنتاجيتين، وبغرض دراسة ظاهرة التغيب في الوحدتين أجريت دراسة خاصة حيث حصلنا على النتائج التالية:

الوحدة الأولى: المدة المتوسطة لتغيب العامل الواحد سنويا يساوي 35 يوم، التباين يساوي 36.

الوحدة الثانية: عدد العمال 400 عامل، المدة الكلية للتغيب خلال السنة هي 8000 يوم، معامل الاختلاف (الانحراف المعياري النسبي) يساوي 15%.

المطلوب:

1- حساب المتوسط الحسابي والانحراف المعياري في الوحدة الثانية؟ مع الشرح؟

2- قارن مستوى التغيب والتشتت بين الوحدتين؟ ما هي أحسن وحدة في رأيك؟

3- أحسب على الوحدة الثانية:

أ- نسبة وعدد العمال الذين تتراوح مدة تغيبهم بين الربيع الأول والربيع الثالث؟

ب- نسبة وعدد العمال الذين تتراوح مدة تغيبهم بين العشير السادس والعشير التاسع؟

ج- نسبة وعدد العمال الذين تتراوح مدة تغيبهم بين المئوي 47 والمئوي 95؟

الفصل الخامس

مقاييس الشكل

نتطرق من خلال هذا الفصل إلى العناصر التالية:

أولا– مفاهيم حول العزوم

ثانيا– أشكال المنحنيات التكرارية

ثالثا– مقاييس الالتواء

رابعا– مقاييس التفرطح

خامسا– حساب المساحات في حالة التوزيع المعتدل (الطبيعي).

تمارين محلولة وتمارين مقترحة

تمهيد:

إن مقاييس النزعة المركزية ومقاييس التشتت لا يكفيان لوحدهما لوصف البيانات وإجراء المقارنات بين التوزيعات التكرارية، فقد يتساوى توزيعان في متوسطهما وفي الانحراف المعياري ومع ذلك نجدهما مختلفين تماما من حيث الشكل، إن الاختلاف من حيث الشكل يتجلى في إحدى الأمرين: الإلتواء أو التفرطح.

لدراسة الإلتواء أو التفرطح نحتاج إلى معلومات رياضية حول ما يعرف بالعزوم.

أولا- مفاهيم حول العزوم:

هي مفاهيم رياضية فيزيائية نحتاج إليها لدراسة بعض المفاهيم الإحصائية وبعض العزوم، بالإضافة إلى مدلولها الرياضي والفيزيائي فلها مدلول إحصائي.

ليكن $x_1, x_2, \dots, x_n$ سلسلة إحصائية، ولتكن a قيمة حقيقية، نقول أن العزم من الدرجة k لهذه السلسلة بالنسبة لـ a هو:

- في حالة سلسلة إحصائية: $\boxed{k\text{العزم من الدرجة} = \frac{\sum_{i=1}^{n}(X_i-a)^k}{n} \dots\dots\dots (1)}$

- في حالـــة توزيع تكراري: $\boxed{k\text{العزم من الدرجة} = \frac{\sum_{i=1}^{n}n_i(X_i-a)^k}{\sum n_i} \dots\dots\dots (2)}$

1- العزوم البسيطة m_k (العزم بالنسبة لنقطة الأصل $a = 0$):

العزم البسيط من الدرجة k هو عبارة عن الوسط الحسابي لقيم المتغير الإحصائي مرفوعة إلى القوة k.

إنطلاقا من تعريف العزم البسيط نحصل على الصيغتين التاليتين:

- في حالة سلسلة إحصائية: $\boxed{m_k = \frac{\sum_{i=1}^{n}X_i^k}{n} \dots\dots\dots (1)}$

- في حالـــة توزيع تكراري: $\boxed{m_k = \frac{\sum_{i=1}^{n}n_iX_i^k}{\sum n_i} \dots\dots\dots (2)}$

1-1- حالات خاصة من العزوم البسيطة:

أ- إذا كان $k = 0$ فإن:

$$m_0 = \frac{\sum_{i=1}^{n}X_i^0}{n} = 1 \dots\dots\dots (1)$$

$$m_0 = \frac{\sum_{i=1}^{n}n_iX_i^0}{\sum n_i} = \frac{\sum_{i=1}^{n}n_i}{\sum n_i} = 1 \dots\dots (2)$$

ب- إذا كان $k = 1$ فإن:

$$m_1 = \frac{\sum_{i=1}^{n}X_i^1}{n} = \bar{X} \dots\dots\dots (1)$$

$$m_1 = \frac{\sum_{i=1}^{n}n_iX_i^1}{\sum n_i} = \bar{X} \dots\dots (2)$$

ومنه إذا كان $k = 1$ فإن: $m_1 = \bar{X}$ أي أن m_1 هو المتوسط الحسابي.

ج- إذا كان $k = 2$ فإن:

$$m_2 = \frac{\sum_{i=1}^{n} X_i^2}{n} = \bar{X}_Q^2 \ldots\ldots\ldots (1)$$

$$m_2 = \frac{\sum_{i=1}^{n} n_i X_i^2}{\sum n_i} = \bar{X}_Q^2 \ldots\ldots (2)$$

ومنه إذا كان $k = 2$ فإن: $m_2 = \bar{X}_Q^2$ أي أن m_2 هو مربع المتوسط التربيعي.

1-2- التعبير على التباين والانحراف المعياري بدلالة العزوم:

رأينا أن: $V(X) = \bar{X}_Q^2 - \bar{X}^2$

ومنه فإن: $\boxed{V(X) = m_2 - m_1^2}$

وبالتالي: $\boxed{\delta(X) = \sqrt{(X)} = \sqrt{m_2 - m_1^2}}$

2- العزوم المركزية m_k (العزم بالنسبة للمتوسط الحسابي $a = \bar{X}$):

في هذه الحالة نعتبر $a = \bar{X}$، وعليه نحصل على العلاقتين التاليتين:

- في حالة سلسلة إحصائية $\boxed{\mu_k = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^k}{n} \ldots\ldots\ldots (1)}$
- في حالـــة توزيع تكراري: $\boxed{\mu_k = \frac{\sum_{i=1}^{n} n_i(X_i-\bar{X})^k}{\sum n_i} \ldots\ldots\ldots (2)}$

- حالات خاصة من العزوم المركزية:

أ- إذا كان $k = 0$ فإن:

$$\mu_0 = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^0}{n} = 1 \ldots\ldots\ldots (1)$$

$$\mu_0 = \frac{\sum_{i=1}^{n} n_i(X_i-\bar{X})^0}{\sum n_i} = 1 \ldots\ldots\ldots (2)$$

ب- إذا كان $k = 1$ فإن:

$$\mu_1 = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^1}{n} = 0 \ldots\ldots\ldots (1)$$

$$\mu_1 = \frac{\sum_{i=1}^{n} n_i(X_i-\bar{X})^1}{\sum n_i} = 0 \ldots\ldots\ldots (2)$$

ج- إذا كان $k = 2$ فإن:

$$\mu_2 = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n} = V(X) \ldots\ldots\ldots (1)$$

$$\mu_2 = \frac{\sum_{i=1}^{n} n_i(X_i-\bar{X})^2}{\sum n_i} = V(X) \ldots\ldots\ldots (2)$$

3- العلاقة بين العزوم المركزية والبسيطة:

لإيجاد العلاقة بين العزوم المركزية والعزوم البسيطة نقوم بنشر ثنائي نيوتن للمقدار $\boxed{\mu_k = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^k}{n}}$

العزوم المركزية الخمسة الأولى بدلالة العزوم البسيطة هي:

$k = 0 \Rightarrow \mu_0 = 1$

$$k = 1 \Rightarrow \mu_0 = 0$$

$$k = 2 \Rightarrow \mu_2 = \frac{\sum X_i^2}{n} - \bar{X}^2 = m_2 - m_1^2$$

$$k = 3 \Rightarrow \mu_3 = \frac{\sum (X_i - \bar{X})^3}{n} = m_3 - 3m_1m_2 + 2m_1^3$$

$$k = 4 \Rightarrow \mu_4 = \frac{\sum (X_i - \bar{X})^4}{n} = m_4 - 4m_1m_3 + 6m_1^2m_2 - 3m_1^4$$

ثانيا– أشكال المنحنيات التكرارية:

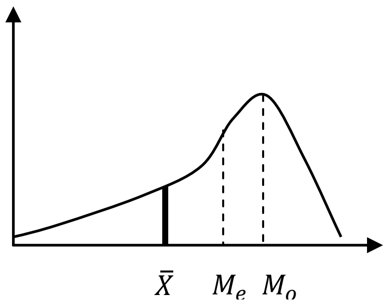
يمكن أن يأخذ المنحنى التكراري أشكالا مختلفة منها ما يعبر عن حالة التناظر والالتواء ومنها ما يعبر عن حالة التفرطح والتطاول والتوزيع الطبيعي.

1– أشكال الإلتواء والتناظر: يمكن أن نميز بين ثلاث أشكال وهي:

أ– توزيع ملتوي نحو اليسار:

في هذه الحالة نجد أن المساحة على يسار $\bar{X}$ أقل من المساحة على يمين $\bar{X}$ أي: $M_o > M_e > \bar{X}$

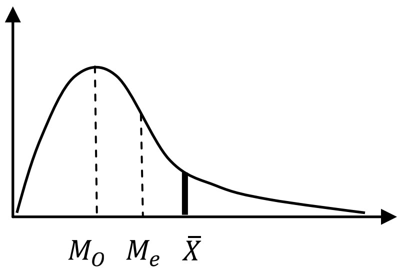
كما هو موضح من خلال الشكل التالي:



ب– توزيع ملتوي نحو اليمين:

في هذه الحالة نجد أن المساحة على يمين $\bar{X}$ أقل من المساحة على يسار $\bar{X}$ أي: $M_o < M_e < \bar{X}$

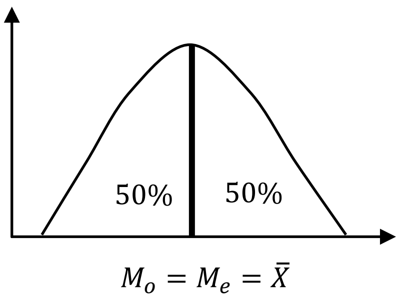
كما هو موضح من خلال الشكل التالي:



ب– توزيع متناظر:

في هذه الحالة نجد أن المساحة على يمين $\bar{X}$ تساوي المساحة على يسار $\bar{X}$ أي: $M_o = M_e = \bar{X}$

كما هو موضح من خلال الشكل التالي:

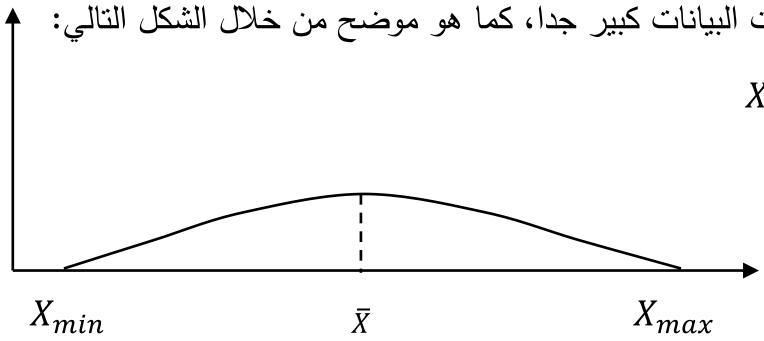


2- أشكال التفرطح والتطاول والتوزيع الطبيعي:

إذا تبين أن التوزيع متناظر، فإنه يمكن أن يأخذ المنحنى التكراري الأشكال التالية:

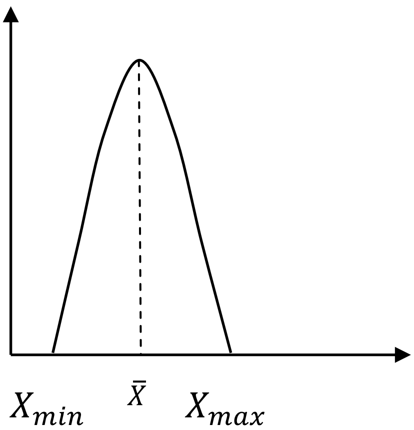
أ- **توزيع متفرطح:** وفيه يكون تشتت البيانات كبير جدا، كما هو موضح من خلال الشكل التالي:

أي: $X_{max} - X_{min} =$ كبير



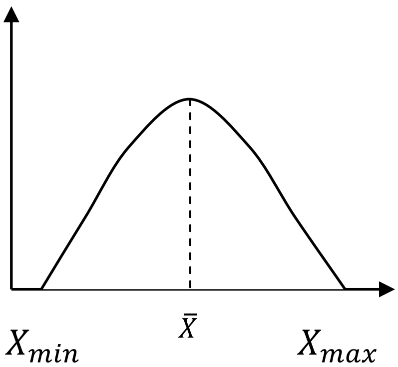
ب- **توزيع متطاول:** وفيه يكون تشتت البيانات ضعيف جدا، كما هو موضح من خلال الشكل التالي:

أي: $X_{max} - X_{min} =$ صغير



ج- **توزيع طبيعي:** وفيه يكون تشتت البيانات لا هو بالكبير ولا هو بالصغير، كما هو موضح من خلال الشكل التالي:

أي: $X_{max} - X_{min} =$ معتدل



ملاحظة:

في أغلب الأحيان لا يكون شكل المنحنى التكراري واضحا مما يصعب علينا أن نحكم بالعين المجردة على نوع التوزيع (متناظر، ملتوي، مفرطح، مدبب أو طبيعي)، وعليه لا بد من مقاييس علمية حسابية للحكم على شكل المنحنى التكراري.

ثالثا- مقاييس الالتواء:

هناك عدة مقاييس للالتواء، أهمها:

1- معامل فيشر α_F:

يقيس هذا المعامل درجة التواء شكل التوزيع الإحصائي، ونعتمد في ذلك على قيمة العزم المركزي من الدرجة الثالثة (لأن قيمته تساوي الصفر في حالة توزيع متناظر)، ولاستبعاد وحدة القياس نقسمه على الانحراف المعياري من الدرجة نفسها.

- حالة سلسلة إحصائية:

$$\boxed{\alpha_F = \frac{\mu_3}{[\delta(X)]^3}}$$

حيث: $\mu_3 = \frac{\sum_{i=1}^{n}(X_i-\bar{X})^3}{n}$ و $\delta(X)$: يمثل الانحراف المعياري

- حالة توزيع تكراري:

$$\boxed{\alpha_F = \frac{\mu_3}{[\delta(X)]^3}}$$

حيث: $\mu_3 = \frac{\sum_{i=1}^{k}n_i(X_i-\bar{X})^3}{n}$ و $\delta(X)$: يمثل الانحراف المعياري

يحتمل معامل فيشر للالتواء ثلاث حالات:

- إذا كان $\alpha_F = 0$: منحنى التوزيع التكراري متناظر، وهذا يعني أن: $M_O = \mathrm{M_e} = \bar{\mathrm{X}}$
- إذا كان $\alpha_F > 0$: منحنى التوزيع التكراري ملتوي نحو اليمين، وهذا يعني أن: $M_O < \mathrm{M_e} < \bar{\mathrm{X}}$
- إذا كان $\alpha_F < 0$: منحنى التوزيع التكراري ملتوي نحو اليسار، وهذا يعني أن: $M_O > \mathrm{M_e} > \bar{\mathrm{X}}$

ملاحظة: يعتبر معامل فيشر أدق مقاييس الالتواء والتفرطح لأنه يوظف كل البيانات بدون استثناء.

2- معامل بيرسون P :

نعلم أنه في حالة التوزيع المتناظر يتساوى كلا من المتوسط الحسابي والوسيط والمنوال، وكلما قل تناظر التوزيع نحو اليمين أو اليسار اختلفت قيم هذه المتوسطات، وفي غالب الأحيان يقع الوسيط في ثلث المسافة بين المتوسط الحسابي والمنوال، ومعامل بيرسون للالتواء هو الفرق بين المتوسط الحسابي والمنوال مقسوما على الانحراف المعياري، أي: $\boxed{P = \frac{(\bar{X}-M_O)}{\delta(X)}}$ أو $\boxed{P = \frac{3(\bar{X}-M_e)}{\delta(X)}}$

يحتمل معامل بيرسون ثلاث حالات:

- إذا كان $P = 0$: منحنى التوزيع التكراري متناظر، وهذا يعني أن: $M_O = \mathrm{M_e} = \bar{\mathrm{X}}$
- إذا كان $P > 0$: منحنى التوزيع التكراري ملتوي نحو اليمين، وهذا يعني أن: $M_O < \mathrm{M_e} < \bar{\mathrm{X}}$
- إذا كان $P < 0$: منحنى التوزيع التكراري ملتوي نحو اليسار، وهذا يعني أن: $M_O > \mathrm{M_e} > \bar{\mathrm{X}}$

ملاحظة: يعتبر معامل بيرسون أقل دقة من معامل فيشر لأنه يعتمد على ثلاث مقاييس ولا يوظف كل البيانات فهو يعطينا فكرة أولية حول نوع التوزيع.

3- معامل يول وكاندال C_{YK}:

يستعمل هذا المعامل في حالة الجداول المفتوحة، أما صيغته فهي:

$$\boxed{C_{YK}=\frac{(Q_3-M_e)-(M_e-Q_1)}{(Q_3-M_e)+(M_e-Q_1)}}$$

يحتمل معامل يول وكندال ثلاث حالات:

- إذا كان $C_{YK}=0$: منحنى التوزيع التكراري متناظر، وهذا يعني أن: $M_O=\mathrm{M_e}=\bar{\mathrm{X}}$
- إذا كان $C_{YK}>0$: منحنى التوزيع التكراري ملتوي نحو اليمين، وهذا يعني أن: $M_O<\mathrm{M_e}<\bar{\mathrm{X}}$
- إذا كان $C_{YK}<0$: منحنى التوزيع التكراري ملتوي نحو اليسار، وهذا يعني أن: $M_O>\mathrm{M_e}>\bar{\mathrm{X}}$

ملاحظة: يعتبر هذا المعامل أقل دقة من معامل فيشر، وهو يعطينا فكرة أولية حول التوزيع إذا كان الالتواء واضح.

رابعا- مقاييس التفرطح

هناك عدة مقاييس للتفرطح، أهمها:

- معامل فيشر β_F:

تستعمل هذه المقاييس في حالة التوزيعات المتناظرة، حيث يقيس هذا المعامل درجة التشتت من خلال شكل المنحني، ونعتمد في ذلك على قيمة العزم المركزي من الدرجة الرابعة (لأن قيمته تساوي 3 في حالة التوزيع الطبيعي أي التوزيع الذي يكون على شكل جرسي)، ولاستبعاد وحدة القياس نقسمه على الانحراف المعياري من الدرجة نفسها.

- حالة سلسلة إحصائية:

$$\boxed{\beta_F=\frac{\mu_4}{[\delta(X)]^4}-3}$$

حيث: $\mu_4=\frac{\sum_{i=1}^{n}(X_i-\bar{X})^4}{n}$ و $\delta(X)$: يمثل الانحراف المعياري

- حالة توزيع تكراري:

$$\boxed{\beta_F=\frac{\mu_4}{[\delta(X)]^4}-3}$$

حيث: $\mu_4=\frac{\sum_{i=1}^{k}n_i(X_i-\bar{X})^4}{n}$ و $\delta(X)$: يمثل الانحراف المعياري

يحتمل معامل فيشر للتفرطح ثلاث حالات:

– إذا كان $\beta_F = 0$: منحنى التوزيع التكراري طبيعي، أي غير متشتت كثيرا ولا متمركز كثيرا.

– إذا كان $\beta_F < 0$: منحنى التوزيع التكراري متفرطح، أي تشتت كبير (قمة منبسطة).

– إذا كان $\beta_F > 0$: منحنى التوزيع التكراري متطاول (مدبب – قمة حادة)، أي تشتت ضعيف.

وهناك مقياس آخر سهل ولكنه أقل دقة يستخدم في حالة البيانات المفتوحة وهو:

$$\boxed{A = \frac{E_Q}{D_9 - D_1}}$$

حيث: E_Q: الانحراف الربيعي، D_1 و D_9 : العشيرين الأول والتاسع على الترتيب.

وقد تبين بالتجربة والحساب أن:

$A = 0,263$: في حالة التوزيع الطبيعي.

$A < 0,263$: في حالة التوزيع المتفرطح.

$A > 0,263$: في حالة التوزيع المتطاول.

مثال(5−1):

في دراسة طبية أجريت على عينة من 40 شخص لمعرفة نسبة السكر في دمهم، تحصلنا على النتائج التالية:

الجدول(5−1): توزيع عينة من 40 شخص حسب نسبة السكر في دمهم

نسبة السكر(غ/ل) X_i	عدد الأشخاص n_i
$]0,3 - 00]$	3
$]0.6 - 0.3]$	15
$]0,9 - 0,6]$	12
$]1,2 - 0,9]$	6
$]1,5 - 1,2]$	2
$]1,8 - 1,5]$	2
المجموع $\sum n_i$	40

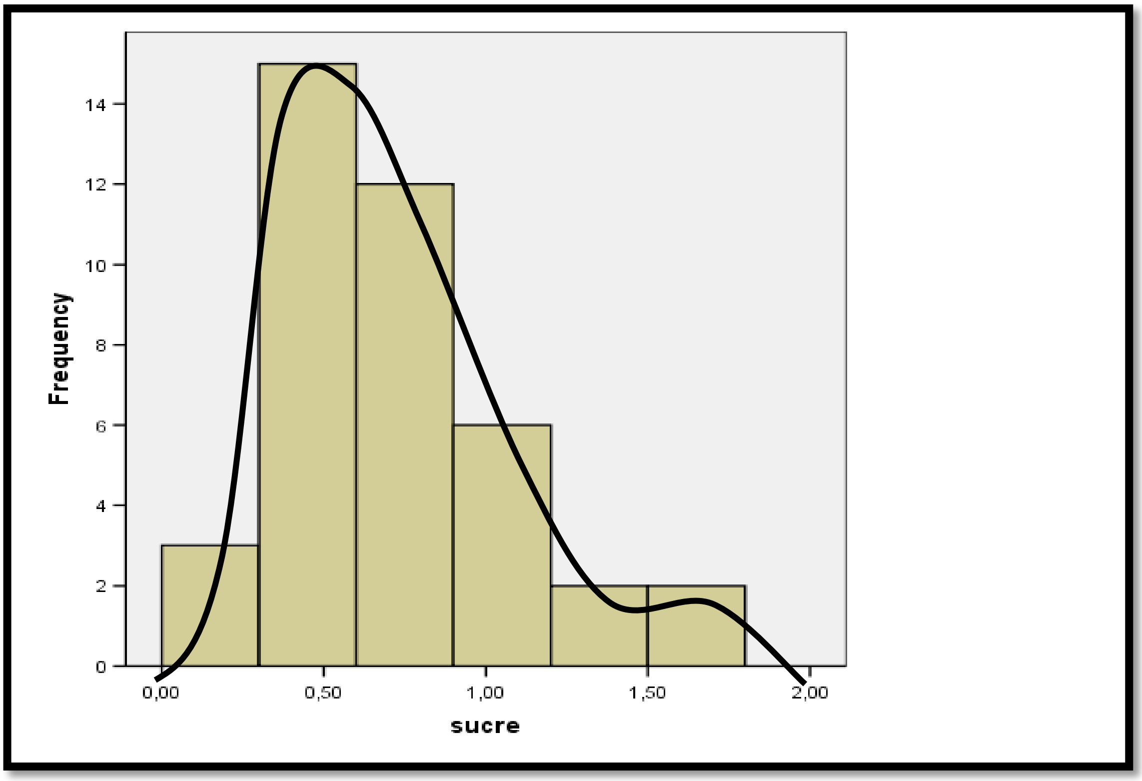
المطلوب:

1− أرسم المنحنى التكراري، ماذا تلاحظ بالنسبة لشكل المنحني،

2− أدرس شكل التوزيع من حيث الالتواء التفرطح؟.

الحل:

1− رسم المنحنى التكراري:



الشكل(5−1): توزيع عينة من 40 شخص حسب نسبة السكر في دمهم

نلاحظ أن المنحني ملتوي نحو اليمين، ومدبب نوعا ما.

$N_i^{\uparrow}$	$n_i(C_i-\bar{X})^4$	$n_i(C_i-\bar{X})^3$	$n_i(C_i-\bar{X})^2$	$n_i\times C_i$	C_i	n_i	نسبة السكر X_i
3	0,300338745	−0,53393555	0,94921875	0,45	0,15	3	$]0,3-00]$
18	0,07122107	−0,27131836	1,03359375	6,75	0,45	15	$]0.6-0.3]$
30	0,00002373	0,00063281	0,016875	9	0,75	12	$]0,9-0,6]$
36	0,0778478	0,23066016	0,6834375	6,3	1,05	6	$]1,2-0,9]$
38	0,33033208	0,51816797	0,8128125	2,7	1,35	2	$]1,5-1,2]$
40	1,54495239	1,64794922	1,7578125	3,3	1,65	2	$]1,8-1,5]$
/	2,32471582	1,59215625	5,25375	28,5	/	40	المجموع $\sum n_i$

2- دراسة الالتواء:

أ- حساب معامل فيشر :

$$\bar{X} = \frac{\sum_{i=1}^{6} n_i C_i}{n} = \frac{28.5}{40} = 0{,}7125\ g/l$$

$$\delta(X) = \sqrt{\frac{\sum_{i=1}^{6} n_i (C_i - \bar{X})^2}{n}} = \sqrt{\frac{5.25375}{40}} = \sqrt{0{,}131} = 0.362 g/l$$

$$\mu_3 = \frac{\sum_{i=1}^{6} n_i (X_i - \bar{X})^3}{n} = \frac{1{,}5921}{40} = 0{,}04$$

$$\alpha_F = \frac{\mu_3}{[\delta(X)]^3} == \frac{0{,}04}{(0{,}362)^3} = 0{,}851$$

نلاحظ أن $\alpha_F > 0$ ومنه فإن منحنى التوزيع التكراري ملتوي نحو اليمين، لكن بدرجة طفيفة.

كما نلاحظ: $\bar{X} = 0{,}7125\ g/l$ ، $M_e = 0{,}66\ g/l$ ، $M_O = 0{,}54\ g/l$

أي: $M_O < M_e < \bar{X}$

ب- معامل بيرسون:

$$P = \frac{(\bar{X} - M_O)}{\delta(X)} = \frac{0{,}7125 - 0{,}54}{0{,}362} = 0{,}476$$

بما أن $P > 0$: فإن منحنى التوزيع التكراري ملتوي نحو اليمين، لكن بدرجة طفيفة.

ج- معامل يول وكندال:

$$C_{YK} = \frac{(Q_3 - M_e) - (M_e - Q_1)}{(Q_3 - M_e) + (M_e - Q_1)}$$

- الربيع الثالث:

- تحديد الفئة الربيعية الثالثة: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{3n}{4}$ ، أي:

$$N_{Q_3}^{\uparrow} \geq \frac{3n}{4} = \frac{3(40)}{4} = 30$$

ومنه الفئة الربيعية الثالثة هي: $[0{,}6 - 0{,}9[$

- حساب الربيع الثالث بطريقة المد الداخلي:

$$Q_3 = Lim_{Q_3} + \left[\frac{\frac{3n}{4} - N_{Q_3-1}^{\uparrow}}{N_{Q_3}}\right] \times A_{Q_3} = 0{,}6 + \left[\frac{30-18}{12}\right] \times 0.3 = 0{,}9\ g/l$$

- الربيع الاول:

- تحديد الفئة الربيعية الأولى: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{n}{4}$ ، أي:

$$N_{Q_1}^{\uparrow} \geq \frac{n}{4} = \frac{(40)}{4} = 10$$

ومنه الفئة الربيعية الأولى هي: $[0{,}3 - 0{,}6[$

- حساب الربيع الثالث بطريقة المد الداخلي:

$$Q_1 = Lim_{Q_1} + \left[\frac{\frac{n}{4} - N^{\uparrow}_{Q_1-1}}{N_{Q_1}}\right] \times A_{Q_1} = 0{,}3 + \left[\frac{10-3}{15}\right] \times 0.3 = 0{,}44\ g/l$$

ومنه:

$$C_{YK} = \frac{(Q_3-M_e)-(M_e-Q_1)}{(Q_3-M_e)+(M_e-Q_1)} = \frac{(0{,}9-0{,}66)-(0{,}66-0{,}44)}{(0{,}9-0{,}66)+(0{,}66-0{,}44)} = \frac{0{,}02}{0{,}46} = 0{,}043$$

بما أن $C_{YK} > 0$: فإن منحنى التوزيع التكراري ملتوي نحو اليمين، لكن بدرجة طفيفة.

3- حساب معامل فيشر للتفرطح:

$$\beta_F = \frac{\mu_4}{[\delta(X)]^4} - 3$$

$$\mu_4 = \frac{\sum_{i=1}^{6} n_i(X_i-\bar{X})^4}{n} = \frac{3{,}3247}{40} = 0{,}058$$

$$\beta_F = \frac{0{,}058}{(0{,}362)^4} - 3 = 0{,}412$$

نلاحظ أن $\alpha_F > 0$ ومنه فإن منحنى التوزيع التكراري متطاول (مدبب)، لكن بدرجة طفيفة.

وهناك مقياس آخر سهل ولكنه أقل دقة وهو:

$$A = \frac{E_Q}{D_9-D_1}$$

$$E_{\mathrm{Q}} = \frac{IQ}{2} = \frac{0{,}9-0{,}44}{2} = 0{,}23\ g/l$$

- حساب العشير التاسع D_9:

- تحديد الفئة العشيرية التاسعة: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{9n}{10}$، أي:

$$N^{\uparrow}_{D_9} \geq \frac{9n}{10} = \frac{9(40)}{10} = \frac{360}{10} = 36$$

ومنه الفئة العشيرية التاسعة هي: $[0{,}9 - 1{,}2]$

- حساب العشير التاسع بطريقة المد الداخلي:

$$D_9 = Lim_{D_9} + \left[\frac{\frac{9n}{10} - N^{\uparrow}_{D_9-1}}{n_{D_9}}\right] \times A_{D_9} = 0{,}9 + \left[\frac{36-30}{6}\right] \times 0{,}3 = 1{,}2\ g/l$$

- حساب العشير الأول D_1:

- تحديد الفئة العشيرية الأولى: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{n}{10}$، أي:

$$N^{\uparrow}_{D_1} \geq \frac{n}{10} = \frac{(40)}{10} = 4$$

ومنه الفئة العشيرية الأولى هي: $[0{,}3 - 0{,}6]$

- حساب العشير الأول بطريقة المد الداخلي:

$$D_1 = Lim_{D_1} + \left[\frac{\frac{n}{10} - N^{\uparrow}_{D_1-1}}{n_{D_1}}\right] \times A_{D_1} = 0{,}3 + \left[\frac{4-3}{15}\right] \times 0{,}3 = 0{,}32\ g/l$$

ومنه:

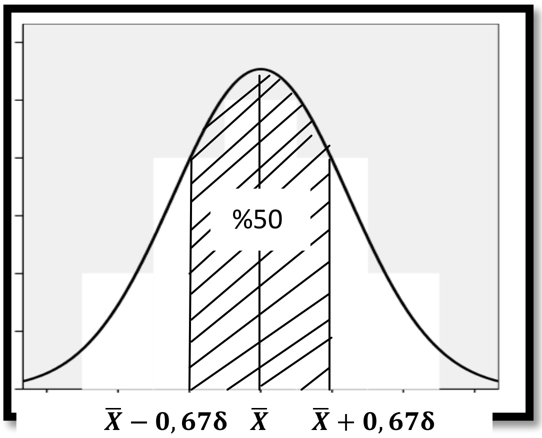
$$A = \frac{E_Q}{D_9-D_1} = \frac{0{,}46}{1{,}2-0{,}32} = 0{,}523$$

بما أن: $A > 0,263$ فإن منحنى التوزيع التكراري متطاول (مدبب)، لكن بدرجة طفيفة.

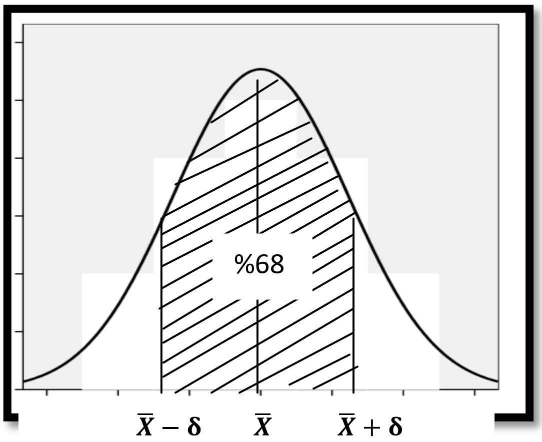
خامسا– حساب المساحات في حالة التوزيع المعتدل (الطبيعي):

يعتبر التوزيع الطبيعي من أكثر التوزيعات التكرارية التي تنطبق على مختلف الظواهر الاقتصادية والاجتماعية، والتوزيع الطبيعي هو كل توزيع تكراري يكون متناظر ومعتدل، أي يحقق الخصائص التالية:

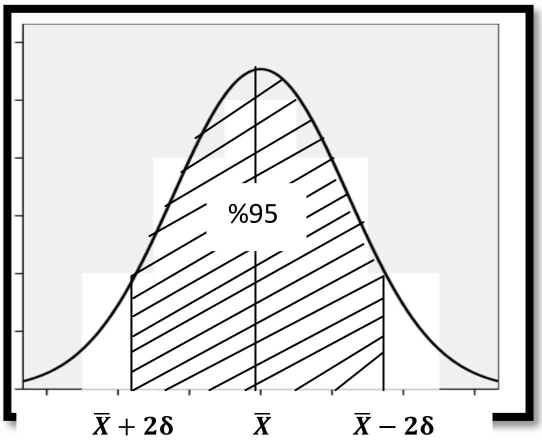
- $M_O = M_e = \bar{X}$
- $\alpha_F = \beta_F = 0$
- المجال: $[\bar{X} - 0,67\delta(X) , \bar{X} + 0,67\delta(X)]$ يحتوي على 50% من البيانات.



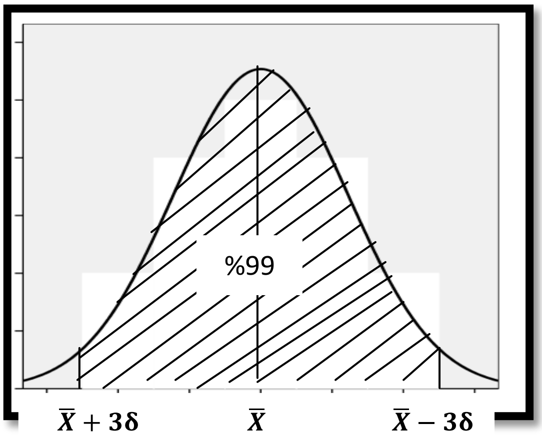
- المجال: $[\bar{X} - \delta(X) , \bar{X} + \delta(X)]$ يحتوي على 68% من البيانات.



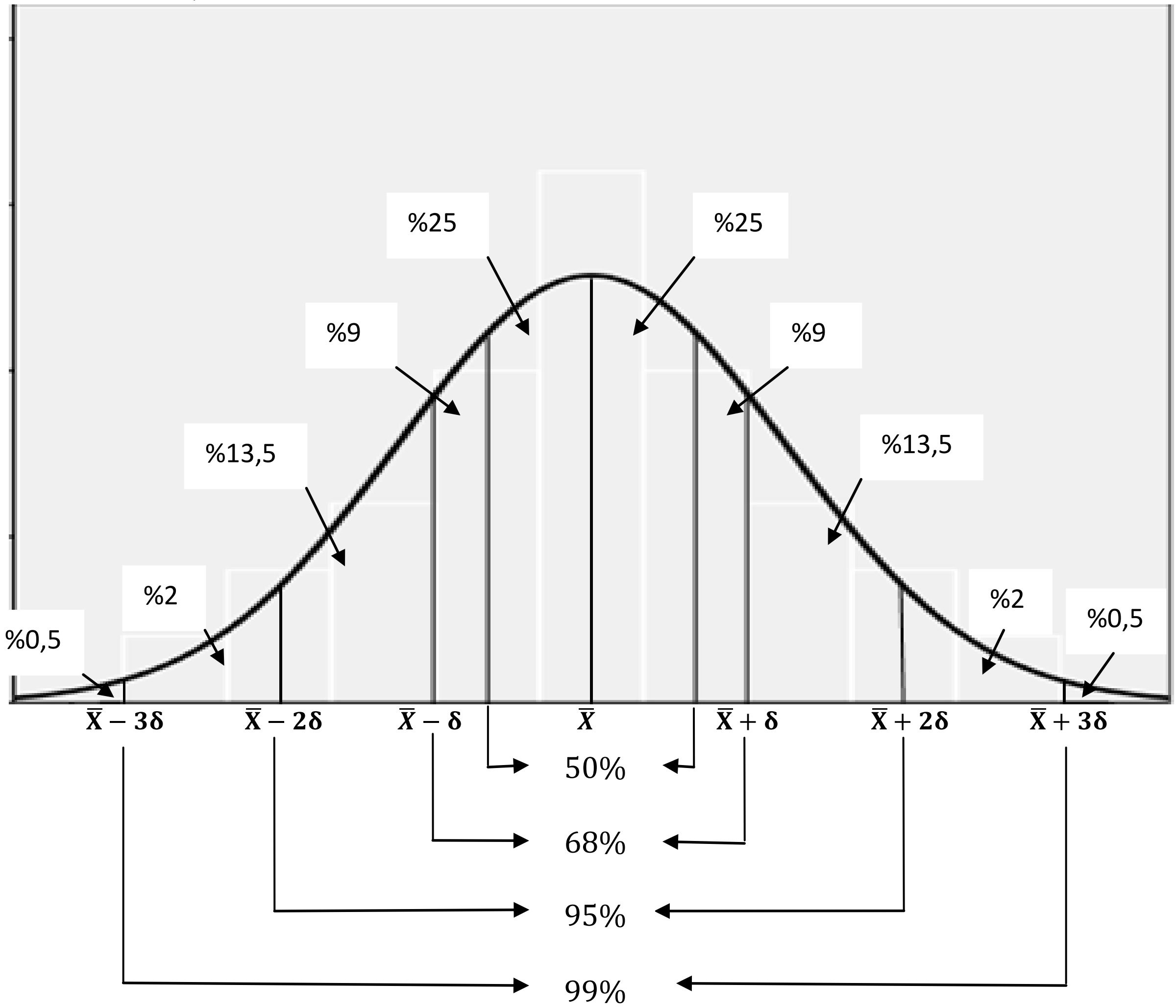
- المجال: $[\bar{X} - 2\delta(X) , \bar{X} + 2\delta(X)]$ يحتوي على 95% من البيانات.



- المجال: $[\bar{X}-3\delta(X)\,,\bar{X}+3\delta(X)]$ يحتوي على 99% من البيانات.



والشكل (5−2) يوضح كل المساحات السابقة في تمثيل واحد:



الشكل(5−2): توزيع المساحات في حالة التوزيع المعتدل (الطبيعي)

مثال(5−2):

بالعودة للمثال السابق، وبافتراض أن التوزيع طبيعي، أحسب مايلي:

1− نسبة وعدد الأشخاص الذين يحملون في دمهم نسبة السكر أقل من 0,7125 غ/ل.

2− نسبة وعدد الأشخاص الذين يحملون في دمهم نسبة السكر أقل من 0,95504 غ/ل.

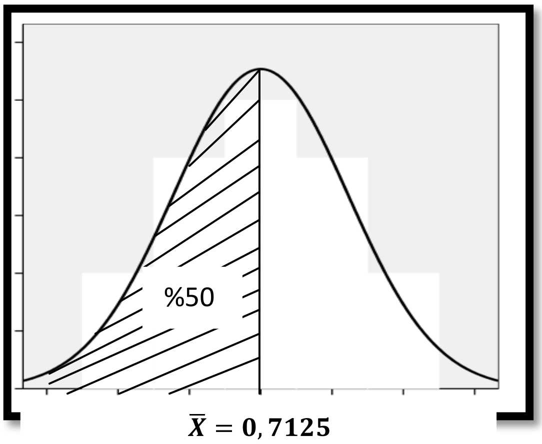
3− نسبة وعدد الأشخاص الذين يحملون في دمهم نسبة السكر بين 0,95504 غ/ل و 1,4365 غ/ل.

4− نسبة وعدد الأشخاص الذين يحملون في دمهم نسبة السكر أكبر من 0,3505 غ/ل.

الحل:

1− نسبة وعدد الأشخاص الذين يحملون في دمهم نسبة السكر أقل من 0,7125 غ/ل:

نلاحظ أن 0,7125 غ/ل تمثل قيمة المتوسط الحسابي، الذي يقسم البيانات إلى قسمين متساويين وبالتالي فإن النسبة كما يتضح من خلال الشكل (5−2) هي 50% ، أما عددهم فهو: $\frac{50}{100}\times 40 = 20$ شخصا.

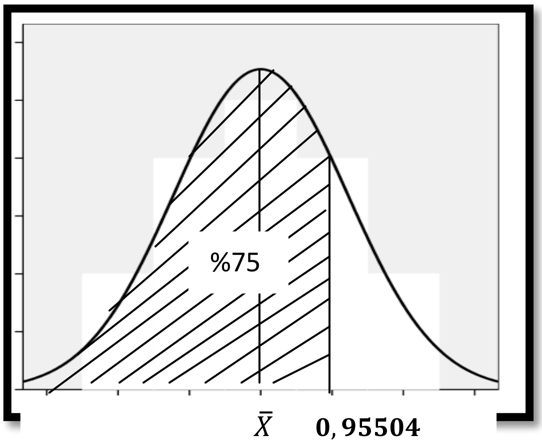


2- نسبة وعدد الأشخاص الذين يحملون في دمهم نسبة السكر أقل من 0,95504 غ/ل.

نلاحظ أن: $0,95504 = 0,7125 + 0,67(0,362) \Rightarrow 0,95504 = \bar{X} + 0,67\delta(X)$

أي أن النسبة كما يتضح من خلال الشكل (5-2) هي: $50 + \frac{50}{2} = 75\ \%$.

أما عددهم فهو: $\frac{75}{100} \times 40 = 30$ شخصا



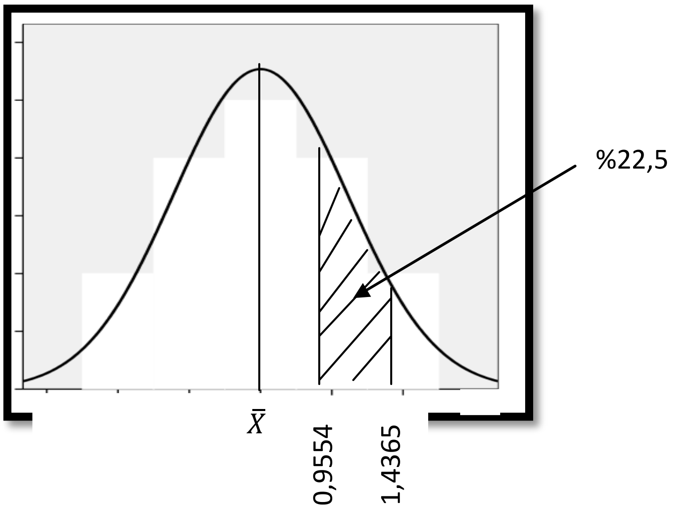
3- نسبة وعدد الأشخاص الذين يحملون في دمهم نسبة السكر بين 0,95504 غ/ل و 1,4365 غ/ل:

نلاحظ أن: $0,95504 = 0,7125 + 0,67(0,362) \Rightarrow 0,95504 = \bar{X} + 0,67\delta(X)$

وأن: $1,4365 = 0,7125 + 2(0,362) \Rightarrow 1,4365 = \bar{X} + 2\delta(X)$

أي أن النسبة كما يتضح من خلال الشكل (5-2) هي: $\frac{95}{2} - \frac{50}{2} = 22,5\ \%$ ،

أما عددهم فهو: $\frac{22,5}{100} \times 40 = 9$ أشخاص.

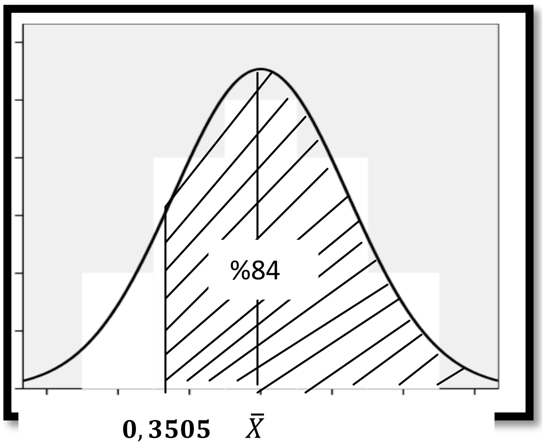


4- نسبة وعدد الأشخاص الذين يحملون في دمهم نسبة السكر أكبر من $0,3505$ غ/ل:

نلاحظ أن: $0,3505 = 0,7125 - (0,362) \Rightarrow 0,3505 = \bar{X} - \delta(X)$

أي أن النسبة كما يتضح من خلال الشكل (5-2) هي: 84 % ، أما عددهم فهو:

$40 \times \frac{84}{100} = 34$ شخص.



تمارين محلولة للفصل الخامس

التمرين الأول:

1- أكتب العزمين المركزيين من الدرجتين الثالثة والرابعة بدلالة العزوم الابتدائية؟

2- أحسب قيمة العزمين المركزيين من الدرجتين الثالثة والرابعة للقيم التالية: 4، 8، 5، 10، 6؟ ثم استنتج شكل هذا التوزيع؟

التمرين الثاني:

في دراسة إحصائية حول أجور العمال في أحد المؤسسات، وبعد جمع البيانات وتحليلها تبين أن 68% من مجموع العمال تتراوح أجورهم بين 60 و84 دج، وقد تبين من خلال دراسة شكل المنحني أن توزيع الأجور طبيعيا.

المطلوب:

1- أحسب الأجر المتوسط والانحراف المعياري؟

2- ما هي نسبة العمال الذين:

أ- تفوق أجورهم 60 دج؟ ب- تقل أجورهم عن 72 دج؟

ج- تتراوح أجورهم ما بين 48 و60 دج؟ د- ما هو بالتقريب أدنى وأعلى أجر في هذه المؤسسة؟

التمرين الثالث:

أولا- في دراسة إحصائية حول حجم المبيعات الشهرية للتجار الصغار من الملابس والأحذية في مدينة سطيف والذي يبلغ عددهم 1500 تاجر تبين أن نصف عدد هؤلاء التجار يحقق الواحد فيهم حجما من المبيعات يتراوح ما بين 23611 دج و 31389 دج.

1- أحسب الحجم المتوسط للمبيعات والانحراف المعياري؟

2- ما هي نسبة وعدد التجار الذين تقل مبيعاتهم عن 31389 دج؟

3- ما هي نسبة وعدد التجار الذين تتراوح مبيعاتهم ما بين 10000 دج و21667 دج؟

4- ما هو الحجم الإجمالي لمبيعات تجار الملابس والأحذية في مدينة سطيف؟

ثانيا- في دراسة مماثلة على تجار الخضر والفواكه تبين أن 99% منهم يحققون حجما من المبيعات يتراوح ما بين 40000 دج و12000 دج.

1- قارن مستوى وتشتت النشاط التجاري بين توزيع تجار الملابس والأحذية وتوزيع تجار الخضر والفواكه مع الشرح الاقتصادي؟

2- ما هي نسبة وعدد التجار الذين تتراوح مبيعاتهم ما بين 16667 دج و 29111 دج إذا افترضنا أن العدد الإجمالي لتجار الخضر والفواكه هو 1200 تاجر؟

ملاحظة: نفرض أن كلا التوزيعين طبيعيين.

التمرين الرابع:

أجريت دراسة إحصائية حول مردودية القمح بولايتي الجلفة والمسيلة للموسم الزراعي 2000-2001 (وحدة القياس: قنطار/هكتار)، وأعطت لنا المعلومات التالية:

ولاية الجلفة: أجريت الدراسة على عينة من 10 مزارع، تم حساب ما يلي:

$$\sum_{i=1}^{10} X_i = 120$$
$$\sum_{i=1}^{10} X_i^2 = 1518{,}4$$
$$\sum_{i=1}^{10} (X_i - \bar{X})^3 = 1{,}1756$$
$$\sum_{i=1}^{10} (X_i - \bar{X})^4 = 1721$$

ولاية المسيلة: أجريت الدراسة على عينة من 15 مزرعة، تم حساب ما يلي:

$$\sum_{i=1}^{15} Y_i = 225$$
$$\sum_{i=1}^{15} (Y_i - \bar{Y})^2 = 183{,}75$$
$$\mu_3 = 0.08575$$
$$\mu_4 = 468$$

المطلوب:

1- أحسب المردودية المتوسطة في الولايتين؟

2- أحسب الانحراف المعياري في الولايتين؟

3- قارن مستوى المردودية وكذلك التشتت في الولايتين؟ ما المدلول الاقتصادي لذلك؟

4- أدرس الالتواء والتفرطح على كلا التوزيعين؟

5- إذا اعتبرنا توزيع المزارع حسب مردوديتها طبيعيا في كلا الولايتين:

أ- أحسب نسبة المزارع في ولاية المسيلة التي تفوق مردوديتها 22 ق/هـ؟

ب- أحسب نسبة المزارع في ولاية الجلفة التي تتراوح مردوديتها ما بين 14,8 ق/هـ و 20,4 ق/هـ؟

6- ما هو بالتقريب أدنى وأقصى مستوى للمردودية في الولايتين؟

الحلـــــــــــول

حل التمرين الأول:

1- كتابة العزمين المركزيين من الدرجتين الثالثة والرابعة بدلالة العزوم الابتدائية:

أ- العزم المركزي من الدرجة الثالثة:

لإيجاد العلاقة بين العزم المركزي من الدرجة الثالثة والعزوم البسيطة نقوم بنشر ثنائي نيوتن للمقدار:

$$\mu_3 = \frac{\sum(X_i-\bar{X})^3}{n} = \frac{\sum\left(X_i^3-3X_i^2\bar{X}+3X_i\bar{X}^2-\bar{X}^3\right)}{n} = \frac{\sum X_i^3-3\bar{X}\sum X_i^2+3\bar{X}^2\sum X_i-n\bar{X}^3}{n}$$

$$\Rightarrow \mu_3 = \frac{\sum X_i^3}{n} - 3\bar{X}\frac{\sum X_i^2}{n} + 3\bar{X}^2\frac{\sum X_i}{n} - \frac{n\bar{X}^3}{n} = m_3 - 3m_1m_2 + 3m_1^3 - m_1^3$$

$$\Rightarrow \mu_3 = m_3 - 3m_1m_2 + 2m_1^3$$

أ- العزم المركزي من الدرجة الرابعة:

لإيجاد العلاقة بين العزم المركزي من الدرجة الثالثة والعزوم البسيطة نقوم بنشر ثنائي نيوتن للمقدار:

$$\mu_4 = \frac{\sum(X_i-\bar{X})^4}{n} = \frac{\sum\left(X_i^4-4X_i^3\bar{X}+6{X_i}^2\bar{X}^2-4X_i\bar{X}^3+\bar{X}^4\right)}{n} = \frac{\sum X_i^4-4\bar{X}\sum X_i^3+6\bar{X}^2\sum X_i^2-4\bar{X}^3\sum X_i+n\bar{X}^4}{n}$$

$$\Rightarrow \mu_4 = \frac{\sum X_i^4}{n} - 4\bar{X}\frac{\sum X_i^3}{n} + 6\bar{X}^2\frac{\sum {X_i}^2}{n} - 4\bar{X}^3\frac{\sum X_i}{n} + \frac{n\bar{X}^4}{n}$$

$$\Rightarrow \mu_4 = m_4 - 4m_1m_3 + 6m_1^2m_2 - 4m_1^3m_1 + m_1^4$$

$$\Rightarrow \mu_4 = m_4 - 4m_1m_3 + 6m_1^2m_2 - 3m_1^4$$

2- حساب قيمة العزمين المركزيين من الدرجتين الثالثة والرابعة للقيم التالية: 4، 8، 5، 10، 6، ثم استنتاج شكل هذا التوزيع:

$$m_1 = \frac{\sum_{i=1}^{5} X_i}{n} = \frac{4+8+5+10+6}{5} = \frac{33}{5} = 6{,}6$$

$$m_2 = \frac{\sum_{i=1}^{5} X_i^2}{n} = \frac{4^2+8^2+5^2+10^2+6^2}{5} = 48{,}2$$

$$m_3 = \frac{\sum_{i=1}^{5} X_i^3}{n} = \frac{4^3+8^3+5^3+10^3+6^3}{5} = 383{,}4$$

$$m_4 = \frac{\sum_{i=1}^{5} X_i^4}{n} = \frac{4^4+8^4+5^4+10^4+6^4}{5} = 3254{,}6$$

بتطبيق العلاقة بين العزم المركزي من الدرجة الثالثة والعزوم الابتدائية نجد:

$$\mu_3 = m_3 - 3m_1m_2 + 2m_1^3 = 383{,}4 - 3(6{,}6)(48{,}2) + 2(6{,}6)^3 = 4{,}032$$

بتطبيق العلاقة بين العزم المركزي من الدرجة الثالثة والعزوم الابتدائية نجد:

$$\mu_4 = m_4 - 4m_1m_3 + 6m_1^2m_2 - 3m_1^4$$

$$\Rightarrow \mu_4 = 3254{,}6 - 4(6{,}6)(383{,}4) + 6(6{,}6)^2(48{,}2) - 3(6{,}6)^4 = 37{,}9712$$

- **استنتاج شكل هذا التوزيع:**

- بالنسبة للإلتواء نستخدم معامل فيشر $\alpha_F = \frac{\mu_3}{[\delta(X)]^3} = \frac{4,032}{(2,154)^3} = 0,403$

بما أن $\alpha_F > 0$: فإن هذا التوزيع ملتوي نحو اليمين ولكن بدرجة طفيفة.

- بالنسبة للتفرطح نستخدم معامل فيشر $\beta_F = \frac{\mu_4}{[\delta(X)]^4} - 3 = \frac{37,9712}{(2,154)^4} - 3 = -1,23$

بما أن $\beta_F < 0$: فإن هذا التوزيع مفرطح ولكن بدرجة طفيفة.

حل التمرين الثاني:

1- حساب الأجر المتوسط والانحراف المعياري:

لدينا 68% من العمال أجورهم تتراوح بين $[\,60\ ,\ 84\,]$

وبما أن توزيع الأجور طبيعي فإن: 68% من العمال أجورهم تتراوح بين $[\,\bar{X} - \delta(X)\ ,\ \bar{X} + \delta(X)\,]$

وبالتالي نخلص إلى حل جملة المعادلتين: $\begin{cases} \bar{X} - \delta(X) = 60 \ldots\ldots\ldots\ldots (1) \\ \bar{X} + \delta(X) = 84 \ldots\ldots\ldots\ldots (2) \end{cases}$

بالجمع بين (1) و (2) نجد: $2\,\bar{X} = 144 \Rightarrow \bar{X} = 72\ DA$

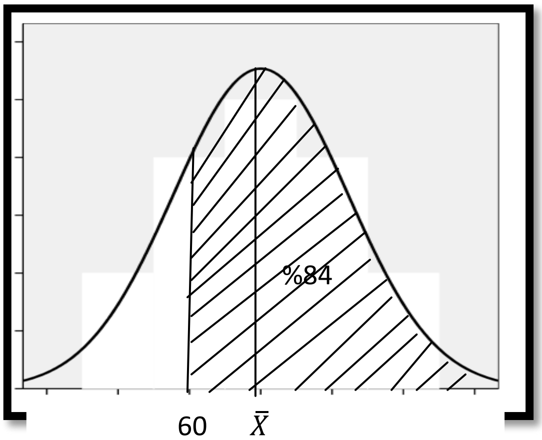
بالتعويض في المعادلة (1) نجد: $72 - \delta(X) = 60 \Rightarrow \delta(X) = 12\ DA$

2- نسبة العمال الذين:

أ- تفوق أجورهم 60 دج

لدينا: $60 = 72 - (1)12 = \bar{X} - \delta(X)$

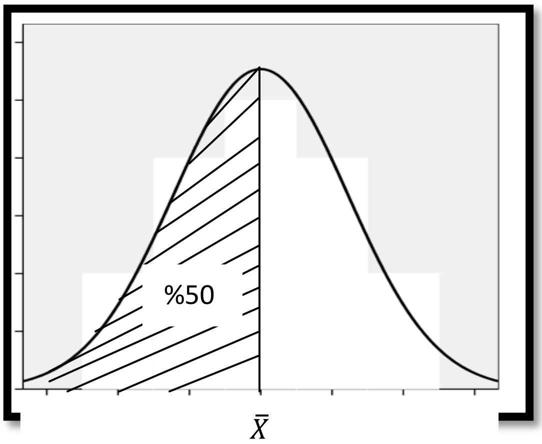
وبالتالي المساحة المقصودة من خلال الشكل (5-2) هي التي تفوق $\bar{X} - \delta(X)$ أي: 50+25+9 = 84%



ب- تقل أجورهم عن 72 دج:

لدينا: $72 = \bar{X}$

وبالتالي المساحة المقصودة من خلال الشكل (5-2) هي التي تقل عن $\bar{X}$ أي: 50%

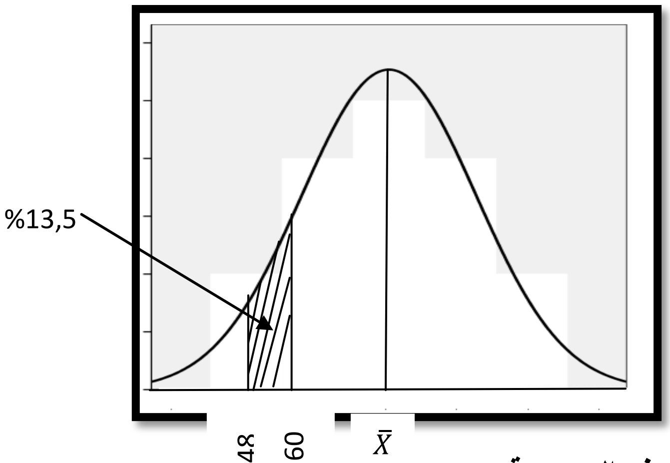


ج- تتراوح أجورهم ما بين 48 و60 دج:

لدينا: $60 = 72 - (1)12 = \bar{X} - \delta(X)$

$48 = 72 - (2)12 = \bar{X} - 2\delta(X)$

وبالتالي المساحة المقصودة من خلال الشكل (5-2)هي المحصورة بين $\bar{X} - \delta(X)$ و $\bar{X} - 2\delta(X)$ أي: %13,5



د- حساب أدنى وأعلى أجر في هذه المؤسسة:

أدنى وأعلى أجر في هذه المؤسسة هو المحصور بين $[\bar{X} - 3\delta(X) \ , \ \bar{X} + 3\delta(X)]$

- **أدنى أجر** هو: $\bar{X} - 3\delta(X) = 72 - 3(12) = 36\ DA$
- **أعلى أجر** هو: $\bar{X} - 3\delta(X) = 72 + 3(12) = 108\ DA$

حل التمرين الثالث:

أولا-1- حساب الحجم المتوسط للمبيعات والانحراف المعياري:

لدينا 50% من تجار الملابس والأحذية يحقق الواحد فيهم حجما من المبيعات يتراوح ما بين 23611 دج و 31389 دج

وبما أن هذا التوزيع طبيعي فإن: 50% من حجم المبيعات تتراوح بين $\left[\bar{X}+\frac{2}{3}\delta(X)\ ,\ \bar{X}-\frac{2}{3}\delta(X)\ \right]$

وبالتالي نخلص إلى حل جملة المعادلتين:

$$\begin{cases}\bar{X}-\frac{2}{3}\delta(X)\ =23611\ \dots\dots\dots(1)\\ \bar{X}+\frac{2}{3}\delta(X)=31389\dots\dots\dots(2)\end{cases}$$

بالجمع بين (1) و (2) نجد: $2\,\bar{X}\ =55000\Rightarrow\bar{X}\ =27500\,DA$

بالتعويض في المعادلة (1) نجد: $27500-\frac{2}{3}\delta(X)\ =23611\Rightarrow\delta(X)\ =5833{,}5\,DA$

2- نسبة وعدد التجار الذين تقل مبيعاتهم عن 31389 دج:

لدينا: $31389=27500+\left(\frac{2}{3}\right)5833{,}5=\bar{X}+\frac{2}{3}\delta(X)$

وبالتالي المساحة المقصودة من خلال الشكل (5-2)هي التي تقل عن $\bar{X}+\frac{2}{3}\delta(X)$ أي: 50 + 9 + 13,5 + 2 + 0,5 = 75%، أما عدد العمال فهو: 0,75(1500) = 1125 تاجر.

3- نسبة وعدد التجار الذين تتراوح مبيعاتهم ما بين 10000 دج و21667 دج:

لدينا: $21667\approx27500-5833{,}5\approx\bar{X}\ -\delta(X)$

$$10000\approx27500-3(5833{,}5)\approx\bar{X}\ -3\delta(X)$$

وبالتالي المساحة المقصودة من خلال الشكل (5-2)هي المحصورة بين $\bar{X}-\delta(X)$ و $\bar{X}\ -3\delta(X)$ أي: 13,5 + 2 = 15,5%، أما عدد العمال فهو: 0,155(1500) ≈ 233 تاجر.

4- الحجم الإجمالي لمبيعات تجار الملابس والأحذية في مدينة سطيف:

$$\bar{X}=\frac{\sum X_i}{n}\Rightarrow\sum X_i=n\bar{X}$$
$$\Rightarrow\sum X_i=1500(27500)$$
$$\Rightarrow\sum X_i=41250000\,DA$$

ثانيا- في دراسة مماثلة على تجار الخضر والفواكه تبين أن 99% منهم يحققون حجما من المبيعات يتراوح ما بين 40000 دج و12000 دج

وبما أن هذا التوزيع طبيعي فإن: 99% من حجم المبيعات تتراوح بين $[\bar{Y}+3\delta(Y)\ ,\ \bar{Y}-3\delta(Y)\ ]$

وبالتالي نخلص إلى حل جملة المعادلتين:

$$\begin{cases}\bar{Y}-3\delta(Y)\ =12000\ \dots\dots\dots(1)\\ \bar{Y}+3\delta(Y)=40000\dots\dots\dots(2)\end{cases}$$

بالجمع بين (1) و (2) نجد: $2\,\bar{Y}\ =52000\Rightarrow\bar{Y}\ =26000\,DA$

بالتعويض في المعادلة (1) نجد: $26000-3\delta(Y)\ =12000\Rightarrow\delta(Y)\approx4667\,DA$

1- مقارنة مستوى وتشتت النشاط التجاري بين توزيع تجار الملابس والأحذية وتوزيع تجار الخضر والفواكه:

أ- مقارنة مستوى النشاط التجاري: بما أن: $\bar{Y} < \bar{X}$ فإن المبيعات في تجارة الملابس والأحذية أحسن من المبيعات في الخضر والفواكه على العموم.

ب- مقارنة التشتت:

بما أن $\bar{Y} \neq \bar{X}$ فإننا نستخدم مقاييس التشتت النسبية لمقارنة تشتت المبيعات الفردية في المجموعتين، حيث نقوم بحساب كلا من CV_1 و CV_2 كما يلي:

$$CV_1 = \delta(\bar{X})\% = \frac{\delta(X)}{\bar{X}} \times 100 = \frac{5833,5}{27500} \times 100 = 21,21\%$$

$$CV_2 = \delta(Y)\% = \frac{\delta(Y)}{\bar{Y}} \times 100 = \frac{4667}{26000} \times 100 = 17,95\%$$

بما أن $CV_1 > CV_2$ فإننا نقول أن تشتت المبيعات الفردية حول المبيعات المتوسطة في قطاع تجارة الملابس والأحذية أكبر مما هي عليه في قطاع الخضر والفواكه، أي أن القطاع الأول غير متجانس حيث هناك فوارق بين حجم المبيعات بين تجار هذا القطاع فمنهم التجار الكبار جدا ومنهم التجار الصغار جدا.

2- حساب نسبة وعدد التجار الذين تتراوح مبيعاتهم ما بين 16667 دج و 29111 دج إذا افترضنا أن العدد الإجمالي لتجار الخضر والفواكه هو 1200 دج:

لدينا:
$$16667 \approx 26000 - 2(4667) \approx \bar{Y} - 2\delta(Y)$$
$$29111 \approx 26000 + \frac{2}{3}(4667) \approx \bar{Y} + \frac{2}{3}\delta(Y)$$

وبالتالي المساحة المقصودة من خلال الشكل (5-2) هي المحصورة بين $\bar{Y} - 2\delta(Y)$ و $\bar{Y} + \frac{2}{3}\delta(Y)$

أي: 50 + 13,5 + 9 = 72,5%، أما عدد التجار فهو: 0,725(1200) = 870 تاجر.

حل التمرين الرابع:

1- حساب المردودية المتوسطة في الولايتين:

أ- ولاية الجلفة: $\bar{X} = \frac{\sum X_i}{n} \Rightarrow \bar{X} = \frac{120}{10} = 12$ ق/هـ

ب- ولاية المسيلة: $\bar{Y} = \frac{\sum Y_i}{n} \Rightarrow \bar{Y} = \frac{225}{15} = 15$ ق/هـ

2- حساب الانحراف المعياري في الولايتين:

أ- ولاية الجلفة: $\delta(X) = \sqrt{\frac{\sum X_i^2}{n} - \bar{X}^2} \Rightarrow \delta(X) = \sqrt{\frac{1518,4}{10} - (12)^2} = 2,8$ ق/هـ

ب- ولاية المسيلة: $\delta(Y) = \sqrt{\frac{(Y_i - \bar{Y})^2}{n}} \Rightarrow \delta(Y) = \sqrt{\frac{183,75}{15}} = 3,5$ ق/هـ

3- مقارنة مستوى المردودية وكذلك التشتت في الولايتين:

أ- مقارنة مستوى المردودية: بما أن: $\bar{Y} > \bar{X}$ فإننا على العموم نقول أن مردودية الهكتار الواحد من القمح في ولاية المسيلة أكبر مما هي عليه في ولاية الجلفة.

ب- مقارنة التشتت:

بما أن $\bar{Y} \neq \bar{X}$ فإننا نستخدم مقاييس التشتت النسبية لمقارنة تشتت المردودية من القمح في الولايتين، حيث نقوم بحساب كلا من CV_1 و CV_2 كما يلي:

$$CV_1 = \delta(\bar{X})\% = \frac{\delta(X)}{\bar{X}} \times 100 = \frac{2,8}{12} \times 100 = 23,33\%$$

$$CV_2 = \delta(Y)\% = \frac{\delta(Y)}{\bar{Y}} \times 100 = \frac{3,5}{15} \times 100 = 23,33\%$$

بما أن $CV_1 = CV_2$ فإننا نقول أن تشتت المردوديات الفعلية حول المردودية المتوسطة في الولايتين متساوي، أي أن الفوارق في المردودية بين المزارع هي نفسها في الولايتين.

4- دراسة الالتواء والتفرطح على كلا التوزيعين:

ولاية الجلفة:

أ- الالتواء:

$$\alpha_F = \frac{\mu_3}{[\delta(X)]^3}$$

$$\mu_3 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^3}{n} = \frac{1,1756}{10} = 0,11756$$

$$\alpha_F = \frac{\mu_3}{[\delta(X)]^3} = \frac{0,11756}{(2,8)^3} = 0,005$$

نلاحظ أن $\alpha_F \approx 0$ ومنه فإن منحنى التوزيع التكراري متناظر نوعا ما، أي أن المردوديات الفعلية في ولاية الجلفة موزعة بالتناظر حول المردودية المتوسطة.

ب- التفرطح:

$$\beta_F = \frac{\mu_4}{[\delta(X)]^4} - 3$$

$$\mu_4 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^4}{n} = \frac{1721}{10} = 172,1$$

$$\beta_F = \frac{\mu_4}{[\delta(X)]^4} - 3 == \frac{172,1}{(2,8)^4} - 3 = -0,2$$

نلاحظ أن $\beta_F \approx 0$ ومنه فإن توزيع المردوديات في ولاية الجلفة طبيعي.

ولاية المسيلة:

أ- الالتواء: $$\alpha_F = \frac{\mu_3}{[\delta(Y)]^3} = \frac{0,08575}{(3,5)^3} = 0,002$$

نلاحظ أن $\alpha_F \approx 0$ ومنه فإن منحنى التوزيع التكراري متناظر نوعا ما، أي أن المردوديات الفعلية في ولاية المسيلة موزعة بالتناظر حول المردودية المتوسطة.

ب- التفرطح: $\beta_F = \frac{\mu_4}{[\delta(X)]^4} - 3 = \frac{468}{(3,5)^4} - 3 = 0,19$

نلاحظ أن $\beta_F \approx 0$ ومنه فإن توزيع المردوديات في ولاية المسيلة طبيعي.

5- إذا اعتبرنا توزيع المزارع حسب مردوديتها طبيعيا في كلا الولايتين:

أ- حساب نسبة المزارع في ولاية المسيلة التي تفوق مردوديتها 22 ق/هـ:

لدينا: $22 = 15 + (2)3,5 = \bar{Y} + 2\delta(Y)$

وبالتالي المساحة المقصودة من خلال الشكل (5-2) هي التي تفوق $\bar{Y} + 2\delta(Y)$ أي: 2 + 0,5 = 2,5%

ب- حساب نسبة المزارع في ولاية الجلفة التي تتراوح مردوديتها ما بين 14,8 ق/هـ و 20,4 ق/هـ:

لدينا: $14,8 = 12 + 2,8 = \bar{X} + \delta(X)$

$$20,4 = 12 + 3(2,8) = \bar{X} + 3\delta(X)$$

وبالتالي المساحة المقصودة من خلال الشكل (5-2) هي المحصورة بين $\bar{X} + \delta(X)$ و $\bar{X} + 3\delta(X)$ أي:

13,5 + 2 = 15,5%

6- حساب أدنى وأقصى مستوى للمردودية في الولايتين:

أ- ولاية الجلفة:

أدنى وأقصى مردودية في الولاية محصورة بين $[\bar{X} - 3\delta(X) \ , \ \bar{X} + 3\delta(X)]$

- **أدنى مردودية** هي: $\bar{X} - 3\delta(X) = 12 - 3(2,8) = 3,6$ ق/هـ
- **أقصى مردودية** هي: $\bar{X} + 3\delta(X) = 12 + 3(2,8) = 20,4$ ق/هـ

ب- ولاية المسيلة:

أدنى وأقصى مردودية في الولاية محصورة بين $[\bar{Y} - 3\delta(Y) \ , \ \bar{Y} + 3\delta(Y)]$

- **أدنى مردودية** هي: $\bar{Y} - 3\delta(Y) = 15 - 3(3,5) = 4,5$ ق/هـ
- **أقصى مردودية** هي: $\bar{Y} + 3\delta(Y) = 15 + 3(3,5) = 25,5$ ق/هـ

تمارين مقترحة

التمرين الأول:

1- ماهي الأشكال التي تأخذها المنحنيات التكرارية؟ مثل هذه الأشكال مع الشرح؟

2- ما هي وظيفة كلا من مقاييس الالتواء والتفرطح؟ متى نستخدمها؟ أذكر أهم وأحسن هذه المقاييس؟ علل؟

3- أكتب العزمين المركزيين من الدرجتين الخامسة والسادسة بدلالة العزوم الابتدائية؟

4- أحسب قيمة العزمين المركزيين من الدرجتين الخامسة والسادسة للقيم التالية: 4، 8، 5، 10، 6؟

التمرين الثاني:

لتقييم مدى نجاح أحد المزارع في تربية الأغنام الموجهة لإستهلاك اللحوم، تم إجراء دراسة على وزن الخرفان التي بلغت 6 أشهر من العمر، وذلك على عينة من 80 خروفا، فكانت النتائج التالية:

الوزن – كلغ- X_i	[15 − 20[	[20 − 25[	[25 − 30[	[30 − 40[	[40 − 50[	المجموع
عدد الخرفان n_i	10	18	30	15	07	80

المطلوب:

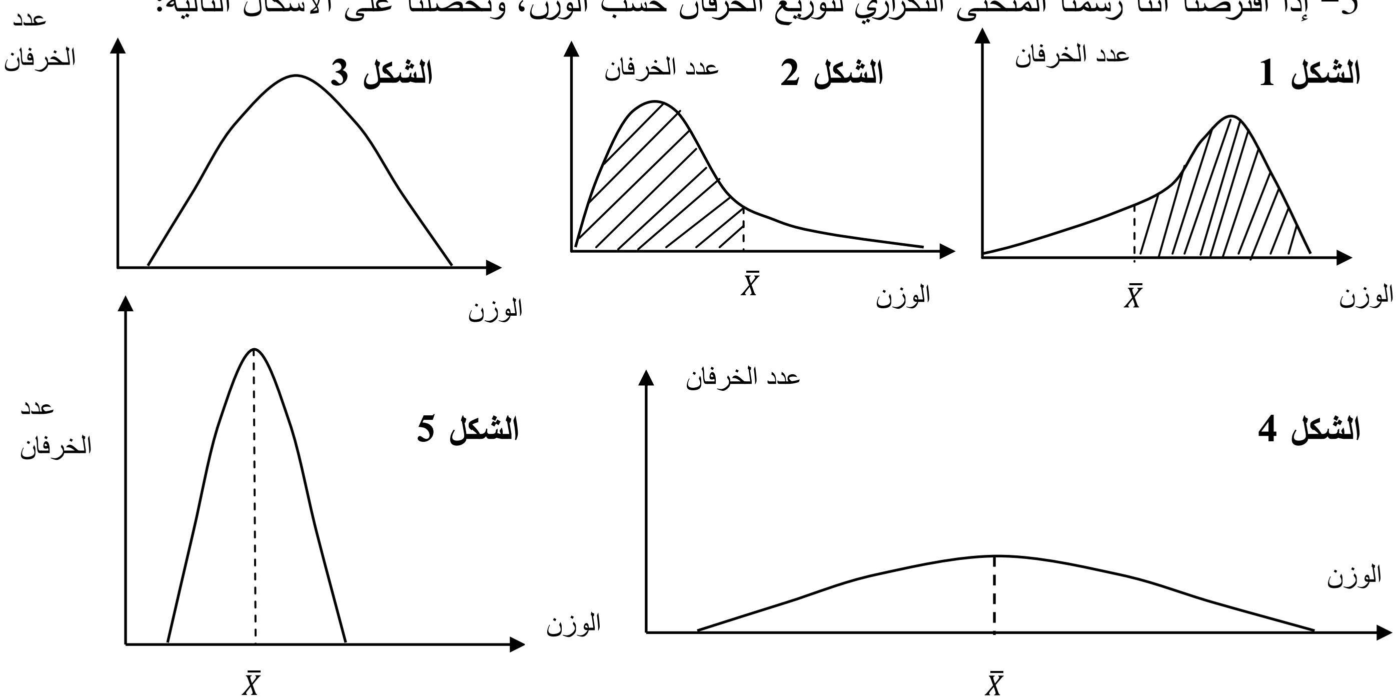
1- مثل بيانيا هذا التوزيع، وأرسم المنحنى التكراري على نفس الرسم البياني؟

2- أحسب الوزن المتوسط، الوسيط والمنوال مع شرح النتائج؟ ماذا تستنتج؟

3- إذا أفترضنا أن الوزن المتوسط للخروف هو 25 كلغ، ما هو الوزن الحقيقي المتوسط بطريقة الانحرافات؟

4- أدرس قضيتي الالتواء والتفرطح لهذا التوزيع؟

5- إذا افترضنا أننا رسمنا المنحنى التكراري لتوزيع الخرفان حسب الوزن، وتحصلنا على الأشكال التالية:



أ- أدرس إحصائيا هذه الأشكال مبينا مدلولها الاقتصادي؟

ب- أي شكل من الأشكال السابقة مطابق لهذه الدراسة؟

ج- ما هي الأشكال التي تعبر على نجاح المزرعة موضحا كيف ذلك؟

التمرين الثالث:

بطلب من إدارة إحدى المركبات الصناعية أجريت دراسة حول الأجور اليومية الفئات المختلفة للعمال على عينة من 55 عاملا، فكانت النتائج التالية:

X_i	$]60-50]$	$]70-60]$	$]80-70]$	$]90-80]$	$]100-90]$	$]110-100]$	$]120-110]$	$]130-120]$
n_i	3	5	8	14	19	9	4	2

ولقد تم حساب ما يلي: الأجر اليومي المتوسط = 88,80 دج، الوسيط = 88,21 دج

المنوال = 86,00 دج، الانحراف المعياري = 17,1 دج.

المطلوب:

1- أرسم المنحنى التكراري لهذا التوزيع؟ حدد مواقع المتوسطات الثلاثة للنزعة المركزية؟ ماذا تستنتج؟

2- أدرس قضية الالتواء باستعمال معامل فيشر، ماذا تعني هذه النتيجة بالنسبة لتوزيع الأجور في هذا المركب؟

3- ماذا يعني الالتواء إلى اليسار والالتواء إلى اليمين من الناحية الاقتصادية بالنسبة لتوزيع الأجور في هذا المركب؟

4- أدرس قضية التفرطح باستعمال معامل فيشر، ماذا تعني هذه النتيجة بالنسبة لتوزيع الأجور في هذا المركب؟

5- إذا تقرر زيادة 10 دج في اليوم للعمال الذين يتقاضون أجرا أقل من المئوي 25، و8 دج للذين يتقاضون أجرا مابين الربيع الأول والعشير الخامس، و 6 دج للذين يتقاضون أجرا ما بين الوسيط والربيع الثالث، و4 دج للذين يتقاضون أجرا أكبر من المئوي 75، إذا علمت أن العدد الإجمالي للعمال في المؤسسة هو 1200 عاملا، فما هي المبالغ المالية الإجمالية الاضافية التي يجب على المؤسسة أن ترصدها لهذه الزيادة في الأجور في اليوم الواحد؟

التمرين الرابع:

في آخر السنة الدراسية الماضية تم تحليل المعدلات السنوية لتلاميذ القسم النهائي لثلاث ثانويات بسطيف فكان ما يلي:

الثانوية الأولى: المعدل السنوي العام 12، الانحراف المعياري 3

الثانوية الثانية: المعدل السنوي العام 10,5، التبـــــاين 4,41

الثانوية الثالثة: بينت الدراسة أن توزيع التلاميذ حسب معدلاتهم السنوية طبيعي، وأن تقريبا 68% تتراوح بين 10,2 و 13,8

المطلوب:

1- قارن المستوى التعليمي في الثانويات الثلاث وكذلك مدى التقارب أو التباعد بين التلاميذ داخل الثانوية الواحدة؟

2- ما هي الثانوية الأنجح من ناحية المستوى التعليمي؟ برر ذلك؟

3- إذا علمت أن العدد الاجمالي للتلاميذ في الثانوية الثالثة هو 200 تلميذ

أ- أحسب نسبة وعدد التلاميذ الذين معدلاتهم تفوق 13,2؟

ب- أحسب نسبة وعدد التلاميذ الذين معدلاتهم تتراوح بين 10,8 و 11,2؟

ج- أحسب نسبة وعدد التلاميذ الذين معدلاتهم تقل عن 6,6؟

4- ما هو بالتقريب أدنى وأقصى مستوى للمعدلات في الثانوية الثالثة؟

التمرين الخامس:

في دراسة إحصائية بأحد ولايات الشرق الجزائري حول المصاريف اليومية المخصصة لشراء السجائر لعينة من 4000 مدخن في سنة 2000م، تبين أن المصاريف الإجمالية اليومية تقدر بـ: 200000 دج، وأن العزم الابتدائي من الدرجة الثانية يقدر بـ: 2564 دج، إذا علمنا أن هذا التوزيع طبيعي:

المطلوب:

1- أحسب كلا المتوسط الحسابي، المتوسط التربيعي، الوسيط، المنوال والانحراف المعياري؟

2- أحسب نسبة وعدد الأفراد التي تفوق مصاريفهم 44,70 دج؟

3- أحسب نسبة وعدد الأفراد التي تتراوح مصاريفهم بين 44,70 دج و 66 دج؟

4- ما هو بالتقريب أدنى وأقصى مصاريف يومية مخصصة لشراء السجائر؟

5- في دراسة مماثلة في 2005م للولاية نفسها، تبين أن 95% من المدخنين تتراوح مصاريفهم اليومية على شراء المخصصة لشراء السجائر ما بين 25 دج و65 دج، قارن مستوى المصاريف والتشتت ما بين الفترتين؟

إمتحانات سابقة

محلولة

جامعة سطيف1

كلية العلوم الاقتصادية والتجارية وعلوم التسيير

السنة أولى LMD   امتحان في مقياس الاحصاء1 (جانفي 2013)   المدة: ساعة ونصف

(الاقسام 1-2-3)

التمرين الأول: (4 ن)

لتكن السلسلة الإحصائية $x_1, x_2, \dots\dots, x_n$ لمتغير إحصائي، ولتكن العبارتين التاليتين:

$$\sum(X_i - \overline{X}) \quad = 0 \dots\dots\dots\dots\dots (1)$$

$$\sum(X_i - \overline{X})^2 < \sum(X_i - X_\alpha)^2 \dots\dots\dots\dots\dots (2)$$

حيث: $\overline{X}$ هو المتوسط الحسابي، و X_α هي أحد القيم بحيث: $\overline{X} \neq X_\alpha$

1-ماذا تمثل هاتين العبارتين؟   2- ماهي القراءة الإحصائية لكل عبارة ؟   3- ماذا تستنتج من ذلك؟

التمرين الثاني: (10 ن)

لتقييم مدى نجاح أحد المزارع في تربية الأغنام الموجهة لاستهلاك اللحوم، تم إجراء دراسة على وزن الخرفان التي بلغت ستة (6) أشهر من العمر، وذلك على عينة من 80 خروفا، فكانت النتائج التالية:

الأوزان (كلغ) X_i	$]15 - 20]$	$]20 - 25]$	$]25 - 30]$	$]30 - 35]$	$]35 - 40]$	**المجموع**
عدد الخرفان	**10**	**18**	**30**	**15**	**07**	**80**

لقد تم حساب مايلي: $\sum n_i(X_i - \overline{X})^3 = 838{,}71$ ، $\mu_4 = 2347{,}4$ ، $\sum n_i(X_i - \overline{X})^2 = 2499{,}688$

المطلوب:

1-مثل بيانيا هذا التوزيع.

2-اشرح n_3 و $N_3^{\uparrow}$

3- أحسب: الوزن المتوسط، الوزن الوسيط، والوزن الأكثر شيوعا

4-أحسب الانحراف المعياري

5-ادرس قضية الالتواء والتفرطح.

6-في مزرعة أخرى تبين أن: الوزن المتوسط هو 30 والانحراف المعياري هو 4.5 ، قارن بين مستوى الوزن والتشتت في المزرعتين.

التمرين الثالث: (6 ن)

إليك التوزيع التالي لنسبة الأسر ونسبة دخلها الشهري حسب خمس فئات للدخل:

الدخل الشهري	**نسبة الأسر%**	**نسبة الدخل%**
10000-20000	12	7
20000-30000	22	17
30000-40000	38	41
40000-50000	28	35
المجموع	**100**	**100**

المطلوب :

1- أدرس قضية التمركز بيانيا.

2- أدرس قضية التمركز حسابيا.

جامعة سطيف1

كلية العلوم الاقتصادية والتجارية وعلوم التسيير

السنة أولى LMD حل امتحان مقياس الاحصاء1 (جانفي 2013) المدة: ساعة ونصف

(الاقسام 1-2-3)

حل التمرين الأول: (4 ن)

1- تمثل هاتين العبارتين: خاصيتين أساسيتين من خصائص المتوسط الحسابي.

2- القراءة الإحصائية لكل عبارة:

$\sum(X_i - \overline{X}) = 0$: مجموع انحرافات القيم عن متوسطها الحسابي يساوي دائما الصفر.

$\sum(X_i - \overline{X})^2 < \sum(X_i - X_\alpha)^2$: مجموع مربع انحرافات القيم عن متوسطها الحسابي أقل من مجموع مربع انحرافات نفس القيم عن أي قيمة أخرى.

3- الاستنتاج:

العبارة (1) تدل على أن المتوسط الحسابي يقع في مركز البيانات، أما العبارة (2) فتدل على أن المتوسط الحسابي أقرب من جميع القيم من أي قيمة أخرى، وبالتالي نستنتج أن المتوسط الحسابي أدق مقاييس النزعة المركزية.

حل التمرين الثاني: (10 ن)

الوزن (كلغ)X_i	عدد الخرفان n_i	C_i	$C_i n_i$.	$N_i^{\uparrow}$.
$[15 - 20[$	10	17,5	175	10
$[20 - 25[$	18	22,5	405	28
$[25 - 30[$	30	27,5	825	58
$[30 - 35[$	15	32,5	487,5	73
$[35 - 30[$	07	37,5	262,5	80
المجموع $\sum n_i$	80	/	2155	/

لقد تم حساب مايلي: $\sum n_i(X_i - \overline{X})^3 = 838{,}71$ ، $\mu_4 = 2347{,}4$ ، $\sum n_i(X_i - \overline{X})^2 = 2499{,}688$

1- التمثيل البياني لهذا التوزيع: بواسطة المدرج التكراري، وبما أن أطوال الفئات متساوية فإننا نمثل مباشرة بدون تعديل التكرارات

عدد الخرفان

وزن الخرفان

2– الشرح:

$\boldsymbol{n_3 = 30}$: هناك 30 خروفا من بين 80 خروفا أوزانهم تتراوح بين 25 كلغ وأقل تماما من 30 كلغ.

$\boldsymbol{N_3^{\uparrow} = 58}$: هناك 58 خروفا من بين 80 خروفا أوزانهم أقل تماما من 30 كلغ

3– حساب الوزن المتوسط، الوزن الوسيط، والوزن الأكثر شيوعا:

أ– المتوسط الحسابي: $\overline{\mathrm{X}} = \frac{\sum_{i=1}^{5} \mathrm{n}_i C_i}{n} = \frac{2155}{80} = 26.94 \text{ kg}$

ب– الوسيط:

– تحديد الفئة الوسيطية: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{\mathrm{n}}{2}$ ، أي: $N_{\mathrm{M_e}}^{\uparrow} \geq 40$

ومنه الفئة الوسيطية هي: $[25 - 30[$

– حساب الوسيط:

$$\mathrm{M_e} = \mathrm{Lim}_{\mathrm{M_e}} + \left[\frac{\frac{n}{2} - N_{\mathrm{M_e}-1}^{\uparrow}}{n_{\mathrm{M_e}}}\right] \times \mathrm{A}_{\mathrm{M_e}} = 25 + \left[\frac{40-28}{30}\right] \times 5 = 27 \text{ kg}$$

ج– الوزن الأكثر شيوعا (المنوال):

– الفئة المنوالية هي: $[25 - 30[$

وبالتالي فإن: $\Delta_2 = 30 - 15 = 15$ ، $\Delta_1 = 30 - 18 = 12$

ومنه :

$$\mathrm{M_o} = \mathrm{Lim}_{\mathrm{M_o}} + \left[\frac{\Delta_1}{\Delta_1 + \Delta_2}\right] \times \mathrm{A}_{\mathrm{M_o}}$$

$$\mathrm{M_o} = 25 + \left[\frac{12}{12+15}\right] \times 5 = 27{,}22 \text{ kg}$$

4– حساب الانحراف المعياري:

$$\delta(X) = \sqrt{\frac{\sum_{i=1}^{5} n_i (X_i - \bar{X})^2}{n}} = \sqrt{\frac{2499{,}688}{80}} = 5.59 \text{ kg}$$

5– دراسة قضيتي الالتواء والتفرطح:

دراسة الالتواء: $\alpha_F = \frac{\alpha_3}{[\delta(X)]^3} = \frac{\frac{838{,}71}{\mathbf{80}}}{[\mathbf{5{,}59}]^3} = 0{,}06 \approx 0$

بما أن $\alpha_F > 0$: فإن التوزيع ملتوي نحو اليمين لكن بدرجة طفيفة جدا أي يميل إلى التناظر.

– **دراسة التفرطح:** $B_F = \frac{\mu_4}{[\delta(X)]^4} - 3 = \frac{2347{,}4}{[\mathbf{5{,}59}]^4} - 3 = -\,0{,}6 \approx 0$

بما أن $B_F < 0$: فإن التوزيع مفرطح لكن بدرجة طفيفة جدا أي يميل إلى الإعتدال أو إلى الشكل الطبيعي.

6- في مزرعة أخرى تبين أن الوزن المتوسط هو 30 والانحراف المعياري هو 4.5 ، المقارنة بين مستوى الوزن والتشتت في المزرعتين:

المزرعة الأولى: $\overline{X}_1 = 26{,}94\ kg$ و $\delta(X_1) = 5{,}59\ \text{kg}$

المزرعة الثانية: $\overline{X}_2 = 30\ \text{kg}$ و $\delta(X_2) = 4{,}5\ \text{kg}$

أ- مقارنة مستوى الوزن:

نلاحظ أن $\overline{X}_1 < \overline{X}_2$ وعليه فإنه على العموم مستوى وزن الخرفان في المزرعة الثانية أكبر مما هو عليه في المزرعة الأولى.

ب- مقارنة تشتت الوزن: بما أن $\overline{X}_1 \neq \overline{X}_2$ فإننا نستخدم مقاييس التشتت النسبية لمقارنة تشتت الوزن في المزرعتين.

$$CV_1 = \delta(X_1)\% = \frac{\delta(X_1)}{\overline{X}_1} \times 100 == \frac{5{,}59}{26{,}94} \times 100 = 20{,}75\ \%$$

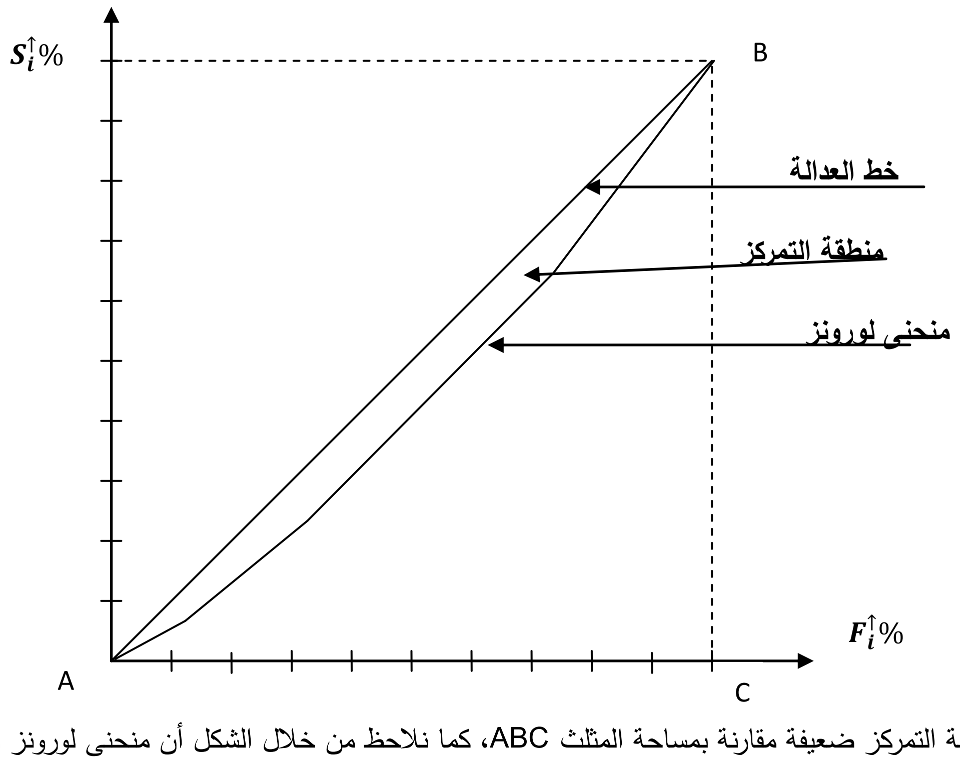
$$CV_2 = \delta(X_2)\% = \frac{\delta(X_2)}{\overline{X}_2} \times 100 == \frac{4{,}5}{30} \times 100 = 15\ \%$$

بما أن $CV_1 > CV_2$ فإننا نقول أن تشتت الوزن في المزرعة الأولى أكبر مما هو عليه في المزرعة الثانية، أي أن الفوارق في وزن الخرفان أكبر في المزرعة الأولى، وعليه فإن وزن الخرفان في المزرعة الثانية أكثر تجانس من وزن الخرفان المزرعة الأولى.

حل التمرين الثالث: (6 ن)

الدخل الشهري	نسبة الأسر $f_i\%$	نسبة الدخل $s_i\%$	$F_i^{\uparrow}\%$	$S_i^{\uparrow}\%$
20000−10000	12	7	12	7
30000−20000	22	17	34	24
40000−30000	38	41	72	65
50000−40000	28	35	100	100
المجموع	**100**	**100**	**/**	**/**

1- دراسة قضية التمركز بيانيا: بواسطة منحنى لورونز



- نلاحظ أن مساحة التمركز ضعيفة مقارنة بمساحة المثلث ABC، كما نلاحظ من خلال الشكل أن منحنى لورونز يقترب من خط العدالة ومنه فإن توزيع الأسر حسب الدخل الشهري أكثر عدالة (التمركز ضعيف).

2– دراسة قضية التمركز حسابيا: بواسطة معامل جيني للتمركز

$f_i\%(S_i^{\uparrow}\% + S_{i-1}^{\uparrow}\%)$	$S_i^{\uparrow}\% + S_{i-1}^{\uparrow}\%$	$S_i^{\uparrow}\%$	$f_i\%$
84	7	7	12
682	31	24	22
3382	89	65	38
4620	165	100	28
8768	/	/	**100**

بما أن النسب مئوية فإن:

$$I_{Gini} = 1 - \frac{1}{10000}\sum_{i=1}^{k} f_i\%(S_i^{\uparrow}\% + S_{i-1}^{\uparrow}\%)$$

$$I_{Gini} = 1 - \frac{1}{10000}(8768)$$

$$I_{Gini} = 0{,}1232 = 12{,}32\%$$

نلاحظ أن $I_{Gini} = 0{,}1232 = 12{,}32\%$ وهي أقل من 30% وعليه يمكن القول أن هذا التوزيع أكثر عدالة، حيث توجد هناك فعلا فوارق في توزيع الدخل الشهري الإجمالي على مختلف فئات السكان، ، لكن هذه الفوارق يمكن تفسيرها بأسباب موضوعية.

جامعة سطيف1

كلية العلوم الاقتصادية والتجارية وعلوم التسيير

السنة أولى LMD امتحان في مقياس الاحصاء1 (جانفي 2013) المدة: ساعة ونصف

(الأقسام 4-5-6)

التمرين الأول: (4 ن)

لتكن لدينا السلسلة الإحصائية التالية: 4، 12، 16، 11، 7،

1. أثبت حسابيا أن مجموع انحرافات هذه القيم عن متوسطها الحسابي معدوما؟
2. أثبت حسابيا أن المتوسط الحسابي لهذه البيانات هو أقرب لها من أي قيمة أخرى (نفرض $X_\alpha = 9$)
3. ماذا تستنتج؟

التمرين الثاني: (10 ن)

الجدول التالي يمثل توزيع 50 مؤسسة حسب مبيعاتها الشهرية (الوحدة: مليون دج).

المبيعات	0-4	4-8	8-12	12-16	16-20	المجموع
عدد المؤسسات	3	15	20	10	2	50

<u>المطلوب</u> : $\sum n_i (X_i - \overline{X})^2 = 704.32$

1- مثل بيانيا هذا التوزيع.

2- احسب التكرارات المطلقة الصاعدة النازلة، ثم اشرح $N_3^{\uparrow}$ ، $N_3^{\downarrow}$

3- أحسب المتوسط الحسابي الوسيط و المنوال.

4- أحسب الربيع الأول والربيع الثالث.

5- أحسب الانحراف المعياري المطلق والنسبي.

6- احسب نسبة المؤسسات التي تقل مبيعاتها عن 9.4 مليون دج.

التمرين الثالث: (6 ن)

لتقييم مدى نجاح أحد المزارع في تربية الأغنام الموجهة لاستهلاك اللحوم، تم إجراء دراسة على وزن الخرفان التي بلغت ستة (6) أشهر من العمر، وذلك على عينة من 80 خروفا، فكانت النتائج التالية:

الأوزان (كلغ) X_i	$[15 - 20[$	$[20 - 25[$	$[25 - 30[$	$[30 - 35[$	$[35 - 40[$	**المجموع**
عدد الخرفان	**10**	**18**	**30**	**15**	**07**	**80**

لقد تم حساب مايلي: $\sum n_i (X_i - \overline{X})^3 = 838{,}71$ ، $\mu_4 = 2347{,}4$ ، $\sum n_i (X_i - \overline{X})^2 = 2499{,}688$

المطلوب:

1-ادرس قضية الالتواء والتفرطح.

2-بفرض أن توزيع الأوزان طبيعي، بمتوسط قدره 27kg وانحراف معياري قدره 5.5kg ، أحسب مايلي:

أ- نسبة الخرفان التي يقل وزنها عن 10.5kg

ب- نسبة الخرفان التي يزيد وزنها عن 16kg

ج- نسبة الخرفان التي يتراوح وزنها بين 21.5 و 32.5kg

د- نسبة الخرفان التي يتراوح وزنها بين 16 و 32.5kg

جامعة سطيف1

كلية العلوم الاقتصادية والتجارية وعلوم التسيير

السنة أولى LMD حل امتحان مقياس الاحصاء1 (جانفي 2013) المدة: ساعة ونصف

(الأقسام 4-5-6)

حل التمرين الأول: (4 ن)

لتكن لدينا السلسلة الإحصائية التالية: 4، 12، 16، 11، 7،

1- إثبات حسابيا أن مجموع انحرافات هذه القيم عن متوسطها الحسابي معدوما: أي $\sum(X_i - \overline{X}) = 0$

لدينا: $\overline{X} = \frac{\sum_{i=1}^{5} X_i}{n} = \frac{50}{5} = 10$ ، $X_\alpha = 9$

القيم xi	$(X_i - \overline{X})$	$(X_i - X_\alpha)$.
4	-6	-5
12	2	3
16	6	7
11	1	2
7	-3	-2
$\sum \boldsymbol{X_i} = 50$	$\sum(X_i - \overline{X}) = 0$	$\sum(\boldsymbol{X_i} - \boldsymbol{X_\alpha}) = \mathbf{5}$

نلاحظ من خلال الجدول أن: $\sum(X_i - \overline{X}) = 0$ وهو المطلوب

2- إثبات حسابيا أن المتوسط الحسابي لهذه البيانات هو أقرب لها من أي قيمة أخرى (نفرض $\boldsymbol{X_\alpha = 9}$): أي
$\sum(X_i - \overline{X})^2 < \sum(X_i - X_\alpha)^2$

القيم xi	$(X_i - \overline{X})^2$.	$(X_i - X_\alpha)^2$.
4	36	25
12	4	9
16	36	49
11	1	4
7	9	4
$\sum \boldsymbol{X_i} = 50$	$\sum(\boldsymbol{X_i} - \overline{\boldsymbol{X}})^2 = 86$	$\sum(\boldsymbol{X_i} - \boldsymbol{X_\alpha})^2 = \mathbf{91}$

من خلال مقارنة النتائج المتحصل عليها في الجدول أعلاه نجد: $91 > 86$ ينتج عن ذلك: $\sum(X_i - \overline{X})^2 < \sum(X_i - X_\alpha)^2$.

3- الاستنتاج:

العبارة $\sum(X_i - \overline{X}) = 0$ تدل على أن المتوسط الحسابي يقع في مركز البيانات،

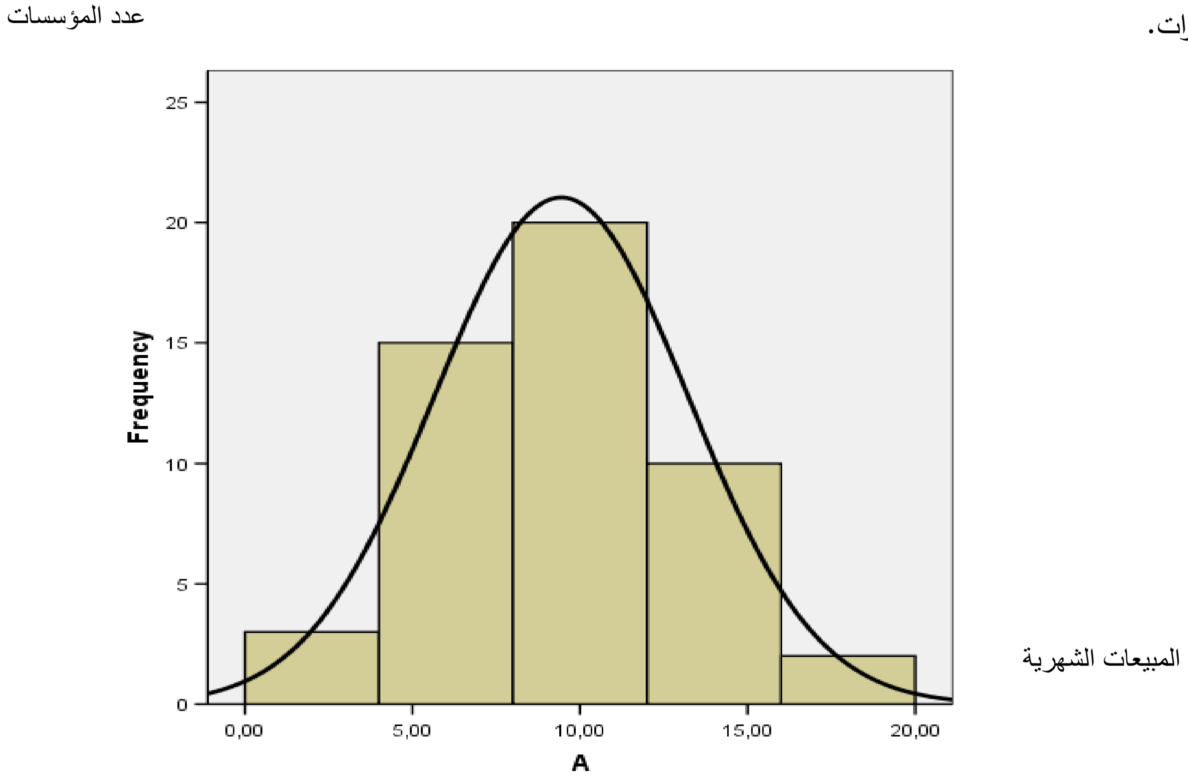
أما العبارة $\sum(X_i - \overline{X})^2 < \sum(X_i - X_\alpha)^2$ فتدل على أن المتوسط الحسابي أقرب من جميع القيم من أي قيمة أخرى، وبالتالي نستنتج أن المتوسط الحسابي أدق مقاييس النزعة المركزية.

حل التمرين الثاني: (10 ن)

المبيعات الشهرية	عدد المؤسسات n_i	C_i	$C_i n_i$.	$N_i^{\uparrow}$.	$N_i^{\downarrow}$.
$]4-0]$	3	2	6	3	50
$]8-4]$	15	6	90	18	47
$]12-8]$	20	10	200	38	32
$]16-12]$	10	14	140	48	12
$]20-16]$	2	18	36	50	2
المجموع $\sum n_i$	50	/	472	/	/

لدينا: $\sum n_i (X_i - \overline{X})^2 = 704.32$

1- التمثيل البياني لهذا التوزيع: بواسطة المدرج التكراري، وبما أن أطوال الفئات متساوية فإننا نمثل مباشرة بدون تعديل التكرارات.



2- حساب التكرارات المطلقة الصاعدة النازلة، ثم شرح $N_3^{\uparrow}$ و $N_3^{\downarrow}$: (أنظر الجدول أعلاه)

$N_3^{\uparrow} = 38$: هناك 38 مؤسسة من بين 50 مؤسسة مبيعاتهم الشهرية أقل تماما من 12 مليون دينار.

$N_3^{\downarrow} = 32$: هناك 32 مؤسسة من بين 50 مؤسسة مبيعاتهم الشهرية أكبر أو تساوي 8 مليون دينار.

3- حساب المتوسط الحسابي، الوسيط و المنوال:

أ- المتوسط الحسابي: $\overline{X} = \frac{\sum_{i=1}^{5} n_i c_i}{n} = \frac{2155}{50} = 9{,}44 \times 10^6 DA$

ب- الوسيط: - تحديد الفئة الوسيطية: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{n}{2}$ ، أي: $N^{\uparrow}_{M_e} \geq 25$

ومنه الفئة الوسيطية هي: $[8 - 12]$

- حساب الوسيط: $M_e = Lim_{M_e} + \left[\frac{\frac{n}{2} - N^{\uparrow}_{M_e-1}}{n_{M_e}}\right] \times A_{M_e} = 8 + \left[\frac{25-18}{20}\right] \times 4 = 9,4 \times 10^6 DA$

ج- المنوال:

- الفئة المنوالية هي: $[8 - 12[$، وبالتالي فإن: $\Delta_1 = 20 - 15 = 5$ ، $\Delta_2 = 20 - 10 = 10$

ومنه :

$$M_o = Lim_{M_o} + \left[\frac{\Delta_1}{\Delta_1 + \Delta_2}\right] \times A_{M_o}$$

$$M_o = 8 + \left[\frac{5}{5+10}\right] \times 4 = 9,33 \times 10^6 DA$$

4- حساب الربيع الأول والربيع الثالث:

أ- الربيع الأول: تحديد الفئة الربيعية الأولى: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{n}{4}$ ، أي: $N^{\uparrow}_{Q_1} \geq 12,5$

ومنه الفئة الربيعية الأولى هي: $[4 - 8]$

- حساب الربيع الأول:

$$Q_1 = Lim_{Q_1} + \left[\frac{\frac{n}{4} - N^{\uparrow}_{Q_1-1}}{n_{Q_1}}\right] \times A_{Q_1} = 4 + \left[\frac{12.5-3}{15}\right] \times 4 = 6,53 \times 10^6 DA$$

أ- الربيع الثالث:

- تحديد الفئة الربيعية الثالثة: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{3n}{4}$ ، أي: $N^{\uparrow}_{Q_3} \geq 37,5$

ومنه الفئة الربيعية الثالثة هي: $[8 - 12]$

- حساب الربيع الثالث:

$$Q_3 = Lim_{Q_3} + \left[\frac{\frac{3n}{4} - N^{\uparrow}_{Q_3-1}}{n_{Q_3}}\right] \times A_{Q_3} = 8 + \left[\frac{37.5-18}{20}\right] \times 4 = 11,9 \times 10^6 DA$$

5- حساب الانحراف المعياري المطلق والنسبي:

أ- الانحراف المعياري المطلق: $\delta(X) = \sqrt{\frac{\sum_{i=1}^{5} n_i (X_i - \bar{X})^2}{n}} = \sqrt{\frac{704.32}{50}} = 3,75 \times 10^6 DA$

أ- الانحراف المعياري النسبي: $CV = \delta(X)\% = \frac{\delta(X)}{\bar{X}} \times 100 == \frac{3,75}{9,44} \times 100 = 39,72\ \%$

6- حساب نسبة المؤسسات التي تقل مبيعاتها عن 9.4 مليون دج:

نلاحظ أن 9.4 مليون دج تمثل قيمة الوسيط، ونعلم أن الوسيط يقسم البيانات إلى 50% إلى اليمين و50% إلى اليسار، وبالتالي فإن نسبة المؤسسات التي تقل مبيعاتها عن 9.4 مليون دج هو 50%، أي 25 مؤسسة.

حل التمرين الثالث: (6 ن)

الوزن (كلغ) X_i	عدد الخرفان n_i	C_i	$C_i n_i$.	$N^{\uparrow}_i$.
$]15 - 20]$	10	17,5	175	10
$]20 - 25]$	18	22,5	405	28
$]25 - 30]$	30	27,5	825	58
$]30 - 35]$	15	32,5	487,5	73
$]35 - 30]$	07	37,5	262,5	80
المجموع $\sum n_i$	80	/	2155	/

لقد تم حساب مايلي: $\sum n_i(X_i-\overline{X})^3 = 838{,}71$ ، $\mu_4 = 2347{,}4$ ، $\sum n_i(X_i-\overline{X})^2 = 2499{,}688$

5- دراسة قضيتي الالتواء والتفرطح:

- دراسة الالتواء:
$$\alpha_F = \frac{\alpha_3}{[\delta(X)]^3} = \frac{\frac{838{,}71}{80}}{[5{,}59]^3} = 0{,}06 \approx 0$$

بما أن $\alpha_F > 0$: فإن التوزيع ملتوي نحو اليمين لكن بدرجة طفيفة جدا أي يميل إلى التناظر.

- دراسة التفرطح:

$$B_F = \frac{\mu_4}{[\delta(X)]^4} - 3 = \frac{2347{,}4}{[5{,}59]^4} - 3 = -\,0{,}6 \approx 0$$

بما أن $B_F < 0$: فإن التوزيع مفرطح لكن بدرجة طفيفة جدا أي يميل إلى الإعتدال أو إلى الشكل الطبيعي.

2- بفرض أن توزيع الأوزان طبيعي، بمتوسط قدره 27kg وانحراف معياري قدره 5.5kg ، حساب مايلي:

أ- نسبة الخرفان التي يقل وزنها عن 10.5kg:

$$10{,}5 = 27 - 16{,}5 = 15 - 3(5{,}5) = \bar{X} - 3\sigma(X)$$

من خلال الشكل البياني لمنحنى التوزيع الطبيعي يتبين أن نسبة الخرفان التي يقل وزنها عن 10.5kg تقدر بـــــ: 0,5 %.

ب- نسبة الخرفان التي يزيد وزنها عن 16kg:

$$16 = 27 - 11 = 15 - 2(5{,}5) = \bar{X} - 2\sigma(X)$$

من خلال الشكل البياني لمنحنى التوزيع الطبيعي يتبين أن نسبة الخرفان التي يزيد وزنها عن 16kg تقدر بـــــ:97,5 %.

ج- نسبة الخرفان التي يتراوح وزنها بين 21.5 و 32.5kg:

$$21{,}5 = 27 - 5{,}5 = 27 - 1(5{,}5) = \bar{X} - \sigma(X)$$
$$32{,}5 = 27 + 5{,}5 = 27 + 1(5{,}5) = \bar{X} + \sigma(X)$$

من خلال الشكل البياني لمنحنى التوزيع الطبيعي يتبين أن نسبة الخرفان التي يتراوح وزنها بين 21.5 و 32.5kg هي:

% 68

د- نسبة الخرفان التي يتراوح وزنها بين 16 و 32.5kg:

$$16 = 27 - 11 = 15 - 2(5{,}5) = \bar{X} - 2\sigma(X)$$
$$32{,}5 = 27 + 5{,}5 = 27 + 1(5{,}5) = \bar{X} + \sigma(X)$$

من خلال الشكل البياني لمنحنى التوزيع الطبيعي يتبين أن نسبة الخرفان التي يتراوح وزنها بين 16 و 32.5kg هي:

% 81,5

جامعة سطيف1

كلية العلوم الاقتصادية والتجارية وعلوم التسيير

السنة أولى LMD　　　　امتحان في مقياس الاحصاء1 (جانفي 2013)　　　　المدة: ساعة ونصف

(الأقسام 7-8-9-10)

التمرين الأول: (4 ن)

لتكن العبارتين التاليتين:

- مجموع انحرافات القيم عن متوسطها الحسابي دوما معدوم.
- مجموع مربع انحرافات القيم عن متوسطها الحسابي أقل من مجموع مربع انحرافاتها عن أي قيمة أخرى.

1. أكتب كل عبارة من العبارتين السابقتين في صيغة رياضية مناسبة.
2. ما هي الدلالة الإحصائية لذلك.

التمرين الثاني: (10 ن)

لتقييم مدى نجاح أحد المزارع في تربية الأغنام الموجهة لاستهلاك اللحوم، تم إجراء دراسة على وزن الخرفان التي بلغت ستة (6) أشهر من العمر، وذلك على عينة من 80 خروفا، فكانت النتائج التالية:

الأوزان (كلغ) X_i	[15 – 20[	[20 – 25[	[25 – 30[	[30 – 35[	[35 – 40[	**المجموع**
عدد الخرفان	**10**	**18**	**30**	**15**	**07**	**80**

لقد تم حساب مايلي: $\sum n_i(X_i - \overline{X})^2 = 2499{,}688$

1- أحسب التكرارات النسبية الصاعدة والنازلة، ومثلها بيانيا على نفس الشكل.

2- أحسب: الوزن المتوسط، الوزن الوسيط، والوزن الأكثر شيوعا

3- احسب العشير 4 والمئوي85.

4- في مزرعة أخرى تبين أن: الوزن المتوسط هو 30 والانحراف المعياري هو 4.5 ، قارن بين مستوى الوزن والتشتت في المزرعتين.

التمرين الثالث: (6 ن)

الجدول التالي يمثل توزيع 50 مؤسسة حسب مبيعاتها الشهرية (الوحدة: مليون دج).

المبيعات	0-4	4-8	8-12	12-16	16-20	المجموع
عدد المؤسسات	3	15	20	10	2	50

المطلوب: $\sum n_i(X_i - \overline{X})^3 = 360.04$ ، $\mu_4 = 527{,}13$ ، $\sum n_i(X_i - \overline{X})^2 = 704.32$

1. ادرس قضية الالتواء والتفرطح
2. بافتراض أن توزيع المبيعات طبيعي، بمتوسط قدره 10 مليون دج وانحراف معياري قدره 2.5 مليون دج، احسب مايلي:
   - أ. احسب نسبة المؤسسات التي تقل مبيعاتها عن 7.5 مليون دج
   - ب. احسب نسبة المؤسسات التي تزيد مبيعاتها عن 15 مليون دج
   - ج. احسب نسبة المؤسسات التي تتراوح مبيعاتها بين 5 و 15 مليون دج
   - د. احسب نسبة المؤسسات التي تتراوح مبيعاتها بين 2.5 و 12.5 مليون دج

جامعة سطيف1

كلية العلوم الاقتصادية والتجارية وعلوم التسيير

السنة أولى LMD   حل امتحان في مقياس الاحصاء1 (جانفي 2013)   المدة: ساعة ونصف

(الأقسام 7-8-9-10)

حل التمرين الأول: (4 ن)

1- كتابة كل عبارة في صيغة رياضية مناسبة:

- مجموع انحرافات القيم عن متوسطها الحسابي دوما معدوم أي: $\sum(\mathrm{X_i}-\overline{\mathrm{X}})=0$
- مجموع مربع انحرافات القيم عن متوسطها الحسابي أقل من مجموع مربع انحرافاتها عن أي قيمة أخرى:

$$\sum(\mathrm{X_i}-\overline{\mathrm{X}})^2<\sum(\mathrm{X_i}-\mathrm{X}_\alpha)^2$$

2-الدلالة الإحصائية لذلك:

- العبارة $\sum(\mathrm{X_i}-\overline{\mathrm{X}})=0$ تدل على أن المتوسط الحسابي يقع في مركز البيانات.
- العبارة $\sum(\mathrm{X_i}-\overline{\mathrm{X}})^2<\sum(\mathrm{X_i}-\mathrm{X}_\alpha)^2$ تدل على أن المتوسط الحسابي أقرب من جميع القيم من أي قيمة أخرى.

وبالتالي نستنتج أن المتوسط الحسابي أدق مقاييس النزعة المركزية.

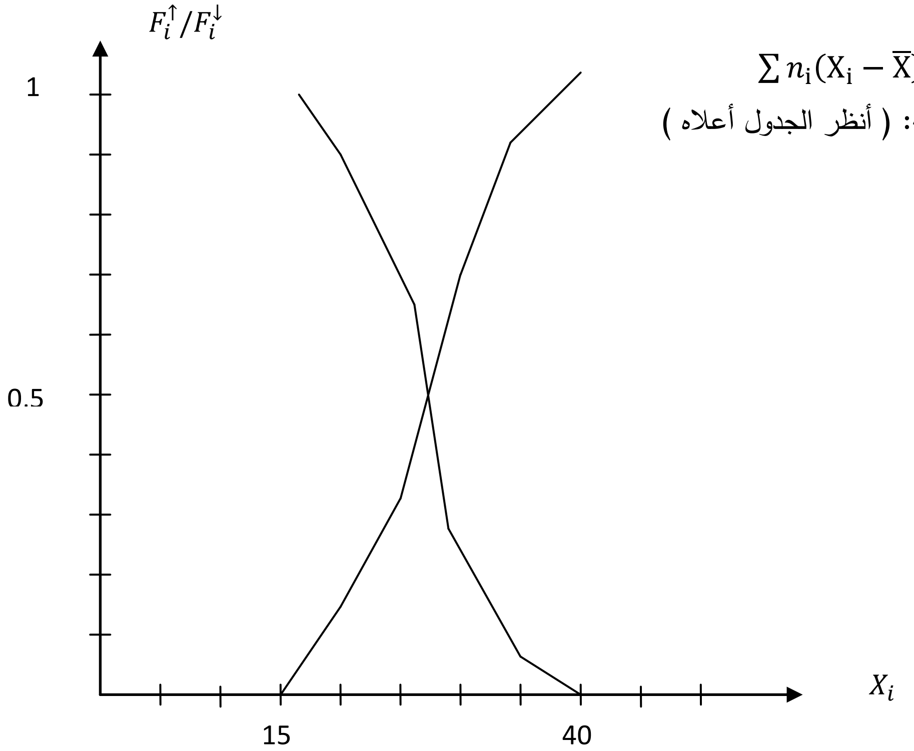
حل التمرين الثاني: (10 ن)

الوزن (كلغ)$\mathrm{X_i}$	عدد الخرفان n_i	f_i	$F_i^{\uparrow}$.	$F_i^{\downarrow}$.	C_i	$\mathrm{C}_i n_i$.
$]20-15]$	10	0.125	0,125	1	17,5	175
$]25-20]$	18	0.225	0,35	0,875	22,5	405
$]30-25]$	30	0.375	0,725	0,65	27,5	825
$]35-30]$	15	0.1875	0,9125	0,275	32,5	487,5
$]30-35]$	07	0.0875	1	0,0875	37,5	262,5
المجموع $\sum \mathrm{n_i}$	80	1	/		/	2155

لقد تم حساب مايلي: $\sum n_\mathrm{i}(\mathrm{X_i}-\overline{\mathrm{X}})^2=2499{,}688$

1- حساب التكرارات النسبية الصاعدة النازلة: (أنظر الجدول أعلاه)

- **التمثيل البياني على نفس الشكل:**



2- حساب الوزن المتوسط، الوزن الوسيط، والوزن الأكثر شيوعا:

أ- المتوسط الحسابي: $\overline{X} = \frac{\sum_{i=1}^{5} n_i c_i}{n} = \frac{2155}{80} = 26.94\ \text{kg}$

ب- الوسيط:

- تحديد الفئة الوسيطية: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{n}{2}$ ، أي: $N_{M_e}^{\uparrow} \geq 40$

ومنه الفئة الوسيطية هي: $[25 - 30[$

- حساب الوسيط: $M_e = Lim_{M_e} + \left[\frac{\frac{n}{2} - N_{M_e-1}^{\uparrow}}{n_{M_e}}\right] \times A_{M_e} = 25 + \left[\frac{40-28}{30}\right] \times 5 = 27\ \text{kg}$

ج- الوزن الأكثر انتشارا (المنوال):

- الفئة المنوالية هي: $[25 - 30[$، وبالتالي فإن: $\Delta_2 = 30 - 15 = 15$ ، $\Delta_1 = 30 - 18 = 12$

ومنه :
$$M_o = Lim_{M_o} + \left[\frac{\Delta_1}{\Delta_1 + \Delta_2}\right] \times A_{M_o}$$
$$M_o = 25 + \left[\frac{12}{12+15}\right] \times 5 = 27,22\ \text{kg}$$

3- حساب العشير 4 والمئوي85:

أ- العشير 4:

- تحديد فئة العشير الرابع: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{4n}{10}$ ، وبما أن التكرارات نسبية فإن فئة العشير الرابع هي أول فئة تكرارها المتجمع الصاعد النسبي أكبر أو يساوي 0,4 أي:

$$F_{D_4}^{\uparrow} \geq 0,4$$

ومنه فئة العشير الرابع هي: $[25 - 30[$

- حساب العشير الرابع:

$$D_4 = Lim_{D_4} + \left[\frac{0,4 - F_{D_4-1}^{\uparrow}}{f_{D_4}}\right] \times A_{D_4} = 25 + \left[\frac{0,4-0,35}{0,375}\right] \times 5 = 25,67\ \text{kg}$$

ب- المئوي85:

- تحديد فئة المئوي 85: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{85n}{100}$ ، وبما أن التكرارات نسبية فإن فئة المئوي85 هي أول فئة تكرارها المتجمع الصاعد النسبي أكبر أو يساوي 0,85 أي:

$$F_{C_{85}}^{\uparrow} \geq 0,85$$

ومنه فئة المئوي 85 هي: $[30 - 35[$

- حساب المئوي 85:

$$C_{85} = Lim_{C_{85}} + \left[\frac{0,85 - F_{C_{85}-1}^{\uparrow}}{f_{C_{85}}}\right] \times A_{C_{85}} = 30 + \left[\frac{0,85-0,725}{0.1875}\right] \times 5 = 33,33\ kg$$

4- في مزرعة أخرى تبين أن الوزن المتوسط هو 30 والانحراف المعياري هو 4.5 ، المقارنة بين مستوى الوزن والتشتت في المزرعتين:

المزرعة الأولى: $\overline{X}_1 = 26,94\ kg$ و $\delta(X_1) = 5,59\ \text{kg}$

المزرعة الثانية: $\overline{X}_2 = 30\ \text{kg}$ و $\delta(X_2) = 4,5\ \text{kg}$

أ- مقارنة مستوى الوزن:

نلاحظ أن $\overline{X}_1 < \overline{X}_2$ وعليه فإنه على العموم مستوى وزن الخرفان في المزرعة الثانية أكبر مما هي عليه في المزرعة الأولى.

ب- مقارنة تشتت الوزن:

بما أن $\bar{X}_1 \neq \bar{X}_2$ فإننا نستخدم مقاييس التشتت النسبية لمقارنة تشتت الوزن في المزرعتين.

$$CV_1 = \delta(X_1)\% = \frac{\delta(X_1)}{\bar{X}_1} \times 100 == \frac{5,59}{26,94} \times 100 = 20,75\ \%$$

$$CV_2 = \delta(X_2)\% = \frac{\delta(X_2)}{\bar{X}_2} \times 100 == \frac{4,5}{30} \times 100 = 15\ \%$$

بما أن $CV_1 > CV_2$ فإننا نقول أن تشتت الوزن في المزرعة الأولى أكبر مما هو عليه في المزرعة الثانية، أي أن الفوارق في وزن الخرفان أكبر في المزرعة الأولى، وعليه فإن وزن الخرفان في المزرعة الثانية أكثر تجانس من وزن الخرفان المزرعة الأولى.

حل التمرين الثالث: (6 ن)

المبيعات الشهرية X_i	عدد المؤسسات n_i	C_i	$C_i n_i$	$N_i^{\uparrow}$	$N_i^{\downarrow}$
$]4-0]$	3	2	6	3	50
$]8-4]$	15	6	90	18	47
$]12-8]$	20	10	200	38	32
$]16-12]$	10	14	140	48	12
$]20-16]$	2	18	36	50	2
المجموع $\sum n_i$	50	/	472	/	/

$\sum n_i(X_i - \bar{X})^2 = 704.32$ ، $\mu_4 = 527,13$ ، $\sum n_i(X_i - \bar{X})^3 = 360.04$

1- دراسة قضيتي الالتواء والتفرطح:

- دراسة الالتواء:

$$\alpha_F = \frac{\alpha_3}{[\delta(X)]^3} = \frac{\frac{360.04}{50}}{\left[\sqrt{\frac{704.32}{50}}\right]^3} = 0,14 \approx 0$$

بما أن $\alpha_F > 0$: فإن التوزيع ملتوي نحو اليمين لكن بدرجة طفيفة جدا أي يميل إلى التناظر.

- دراسة التفرطح:

$$B_F = \frac{\mu_4}{[\delta(X)]^4} - 3 = \frac{527,13}{\left[\sqrt{\frac{704.32}{50}}\right]^4} - 3 = -0,34 \approx 0$$

بما أن $B_F < 0$: فإن التوزيع مفرطح لكن بدرجة طفيفة جدا أي يميل إلى الإعتدال أو إلى الشكل الطبيعي.

2- بافتراض أن توزيع المبيعات طبيعي، بمتوسط قدره 10 مليون دج وانحراف معياري قدره 2.5 مليون دج، حساب مايلي:

أ- حساب نسبة المؤسسات التي تقل مبيعاتها عن 7.5 مليون دج:

$$7,5 = 10 - 2,5 = 10 - 1(2,5) = \bar{X} - \sigma(X)$$

من الشكل البياني لمنحنى التوزيع الطبيعي يتبين أن نسبة الخرفان التي تقل مبيعاتها عن 7.5 مليون دج تقدر بـ: 16 %.

ب- حساب نسبة المؤسسات التي تزيد مبيعاتها عن 15 مليون دج:

$$15 = 10 + 5 = 10 + 2(2,5) = \bar{X} + 2\sigma(X)$$

من خلال الشكل البياني لمنحنى التوزيع الطبيعي يتبين أن نسبة الخرفان التي يزيد وزنها عن 16kg تقدر بــــ: 2,5 %.

ج- حساب نسبة المؤسسات التي تتراوح مبيعاتها بين 5 و 15 مليون دج:

$$5 = 10 - 5 = 10 - 2(2,5) = \bar{X} - 2\sigma(X)$$
$$15 = 10 + 5 = 10 + 2(2,5) = \bar{X} + 2\sigma(X)$$

من خلال الشكل البياني لمنحنى التوزيع الطبيعي يتبين أن نسبة الخرفان التي تتراوح مبيعاتها بين 5 و15 مليون دج هي:

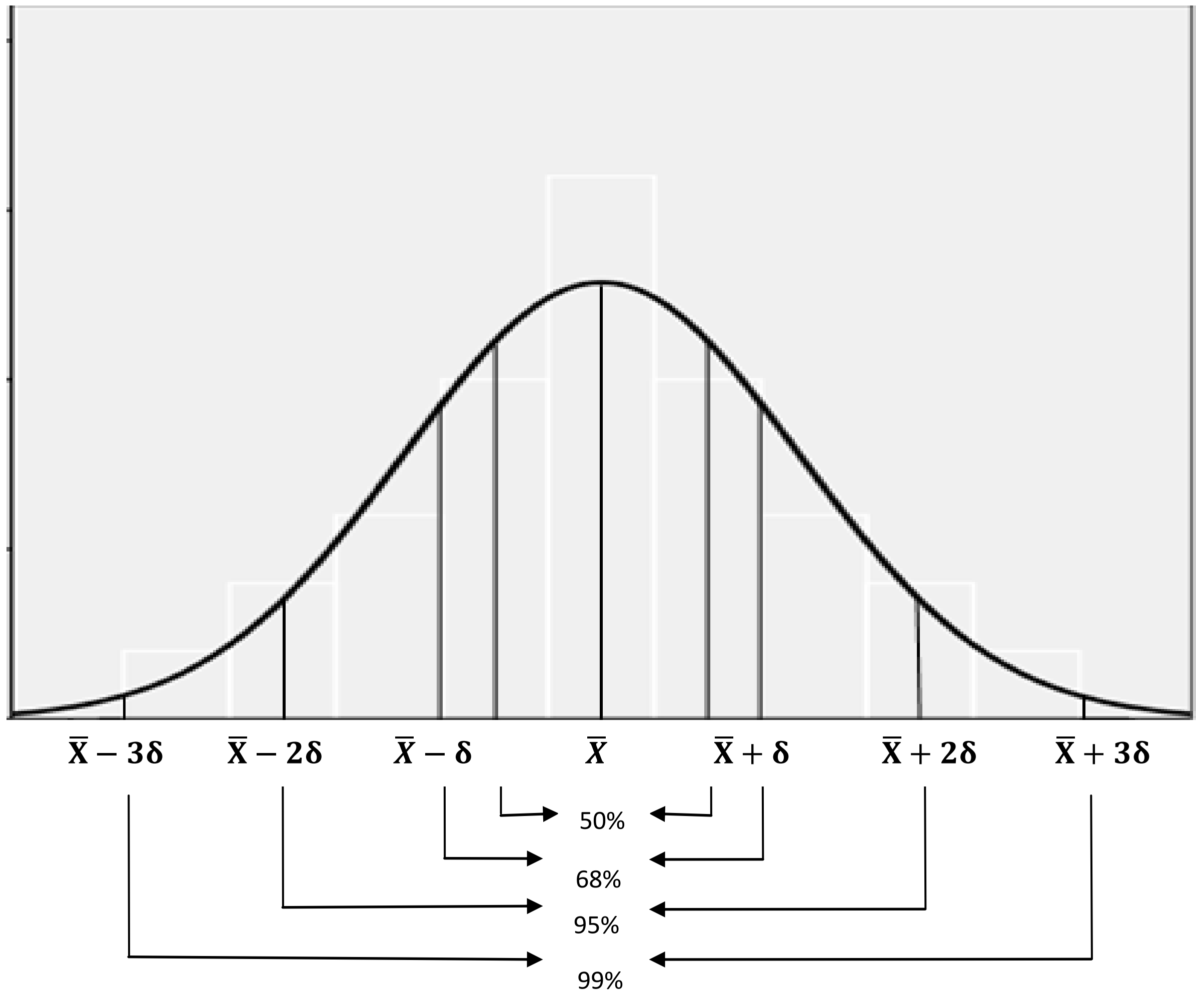
% 95

د- حساب نسبة المؤسسات التي تتراوح مبيعاتها بين 2.5 و 12.5 مليون دج:

$$2,5 = 10 - 7,5 = 10 - 3(2,5) = \bar{X} - 3\sigma(X)$$
$$12,5 = 10 + 2,5 = 10 + 1(2,5) = \bar{X} + \sigma(X)$$

من خلال الشكل البياني لمنحنى التوزيع الطبيعي يتبين أن نسبة الخرفان التي يتراوح وزنها بين 16 و 32.5kg هي:

% 83,5



الامتحان الاستدراكي في مقياس الإحصاء1

الأقسام (من 1 إلى 5)

التمرين الأول:(6 نقاط)

1- في دراسة إحصائية وجدنا أن $\overline{X}=5$ و $\sum x_i = 120$ ، ما هو حجم العينة في هذه الدراسة؟

2- إذا كان $\beta_F = 4 \quad , \quad \alpha_F = -1$ ، ماذا يمكن أن نقول عن الالتواء و التفرطح في هذا التوزيع؟

3- في دراسة حول الأجور وجدنا أن مساحة التمركز تساوي 1700، ماذا تستنتج ؟

4- إذا وجدنا أن الرقم القياسي لسعر سلعة يساوي 90%، ما نوع ومقدار التغير في سعر هذه السلعة؟

التمرين الثاني: (14 نقاط)

الجدول التالي يعطينا توزيع عينة من 40 سيارة من نوع Renault حسب كمية البنزين المستهلكة في 100 كلم (الوحدة: لتر)

الفئات	5-6	6-7	7-8	8-9	9-10	المجموع
عدد السيارات	3	10	20	5	2	40

المطلوب:

1- حدد المجتمع الإحصائي، الوحدة الإحصائية، المتغير الإحصائي المدروس ونوعه في هذا التوزيع،

2- مثل بيانيا هذا التوزيع.

3- أحسب المتوسط الحسابي واشرح النتيجة،

4- احسب الوسيط واشرح النتيجة،

5- احسب المنوال واشرح النتيجة،

6- إذا علمت أن: $\sum n_i (x_i - \bar{x})^2 = 33{,}775$ ، احسب الانحراف المعياري النسبي،

7- في دراسة مماثلة على سيارات Peugeot وجدنا أن: $S(x)=1{,}2 \quad , \quad \overline{X}=7{,}5$ ، قارن بين مستوى استهلاك البنزين و التشتت لكلا النوعين من السيارات، أي النوعين أفضل؟

بالتوفيق للجميع

جامعة سطيف-1-
كلية العلوم الاقتصادية والتجارية وعلوم التسيير
السنة الأولى ل.م.د –

يوم: السبت 2013/03/16

المدة: ساعة و ربع

حل الامتحان الاستدراكي في مقياس الإحصاء1

الأقسام (من 1 إلى 5)

حل التمرين الأول:(6 نقاط)

1- في دراسة إحصائية وجدنا أن $\bar{X} = 5$ و $\sum x_i = 120$ ، حجم العينة في هذه الدراسة:

$$\bar{X} = \frac{\sum x_i}{n} \Rightarrow n = \frac{\sum x_i}{\bar{X}} = \frac{120}{5} = 24$$

2- إذا كان $\beta_F = 4 \quad , \quad \alpha_F = -1$ ، يمكن أن نقول عن الالتواء و التفرطح في هذا التوزيع:

$\alpha_F = -1$: المنحني الخاص بهذا التوزيع ملتوي نحو اليسار، أي أن عدد القيم التي تكون على يمين المتوسط الحسابي أكبر من عدد القيم التي تكون على اليسار، وهذا يعني أن: $Mo \rangle Me \rangle \bar{X}$.

$\beta_F = 4$: المنحنى الخاص بهذا التوزيع مدبب أي متطاول، وهذا يدل على أن البيانات غير متباعدة عن بعضها(تشتت ضعيف).

3- في دراسة حول الأجور وجدنا أن مساحة التمركز تساوي 1700، نستنتج أن:

$$S = 1700 \Rightarrow I_{Gini} = \frac{S}{ABC} = \frac{1700}{5000} = 0,34 = 34\,\%$$

بما أن $I_{Gini} \rangle 0,3$: فإنه يمكن القول أن هذا التوزيع أقل عدالة، حيث توجد هناك فعلا فوارق في توزيع الأجر الإجمالي على مختلف الفئات ، وبالتالي يمكن أن نعتبر هذا التمركز قوي، أي أن هذه الفوارق لا يمكن تفسيرها بأسباب موضوعية.

4- إذا وجدنا أن الرقم القياسي لسعر سلعة يساوي 90%، نوع ومقدار التغير في سعر هذه السلعة:

$$I_{(P)1/0} = 90\,\% \Rightarrow \Delta = 90 - 100 = -10\,\%$$

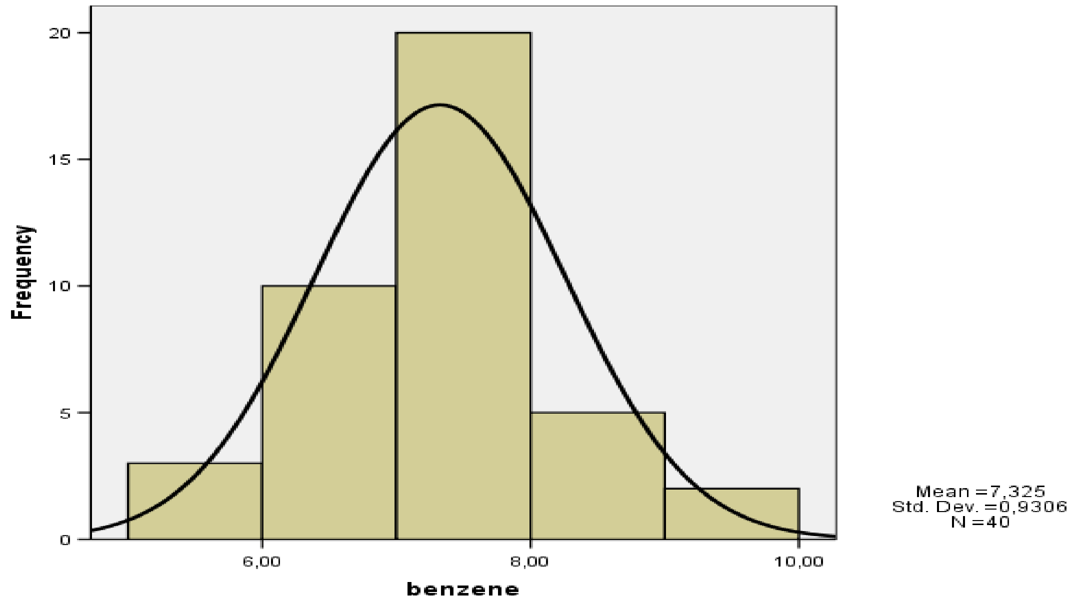
أي أنه حدث **انخفاض** بنسبة **10%** في سعر هذه السلعة.

حل التمرين الثاني: (14 نقاط)

1- تحديد المجتمع الإحصائي، الوحدة الإحصائية، المتغير الإحصائي المدروس ونوعه في هذا التوزيع:

المجتمع الإحصائي	الوحدة الإحصائية	المتغير الإحصائي	نوعه
جميع السيارات من نوع رونو	السيارة الواحدة	كمية البنزين المستهلكة	كمي مستمر

2- التمثيل البياني لهذا التوزيع: بواسطة المدرج التكراري.



تمثيل بياني بواسطة المدرج التكراري لتوزيع 40 سيارة رونو حسب كمية البنزين المستهلكة في 100 كلم

3- حساب المتوسط الحسابي وشرح النتيجة:

$N_i^{\uparrow}$.	.$C_i n_i$	C_i	n_i عدد السيارات	X_i(لتر) كمية البنزين
03	16,5	5,5	03	$]6-5]$
13	65	6,5	10	$]7-6]$
33	150	7,5	20	$]8-7]$
38	42,5	8,5	05	$]9-8]$
40	19	9,5	02	$]10-9]$
/	293	/	40	$\sum n_i$ المجموع

$$\bar{X} = \frac{\sum_{i=1}^{5} n_i C_i}{n} = \frac{293}{40} = 7{,}325\ l$$

الشرح: متوسط كمية البنزين المستهلكة في 100 كلم بالنسبة لسيارات نوع رونو تقدر بـ: 7,325 لتر.

4- حساب الوسيط وشرح النتيجة:

- تحديد الفئة الوسيطية: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{n}{2}$، أي: $N_{M_e}^{\uparrow} \geq 20$

ومنه الفئة الوسيطية هي: $]8-7]$

- حساب الوسيط:

$$M_e = Lim_{M_e} + \left[\frac{\frac{n}{2} - N_{M_e-1}^{\uparrow}}{n_{M_e}}\right] \times A_{M_e} = 7 + \left[\frac{20-13}{20}\right] \times 1 = 7{,}35\ l$$

الشرح: 50% من السيارات نوع رونو يستهلكون كمية من البنزين تقل عن 7,35 لتر في 100 كلم و50% من السيارات المتبقية يستهلكون كمية من البنزين تفوق 7,35 لتر في 100 كلم.

5- حساب المنوال وشرح النتيجة:

- الفئة المنوالية هي: $[7-8[$

وبالتالي فإن: $\Delta_2 = 20 - 5 = 15$ ، $\Delta_1 = 20 - 10 = 10$

ومنه : $M_o = Lim_{M_o} + \left[\frac{\Delta_1}{\Delta_1+\Delta_2}\right] \times A_{M_o}$

$$M_o = 7 + \left[\frac{10}{10+15}\right] \times 1 = 7,4\ l$$

الشرح: أغلبية سيارات نوع رونو يستهلكون كمية من البنزين تقدر بـ: 7,4 لتر في 100 كلم.

6- إذا علمت أن: $\sum n_i(x_i - \bar{x})^2 = 33,775$ ، **حساب الانحراف المعياري النسبي:**

$$CV = \frac{S(x)}{\bar{X}} \times 100 = \frac{\sqrt{\frac{\sum n_i(x_i-\bar{x})^2}{n}}}{\bar{X}} \times 100 = \frac{\sqrt{\frac{33,775}{40}}}{7,325} \times 100 = 12,54\ \%$$

7- في دراسة مماثلة على سيارات Peugeot وجدنا أن: $S(x) = 1,2$, $\bar{X} = 7,5$ ، **المقارنة بين مستوى استهلاك البنزين و التشتت لكلا النوعين من السيارات، وأي النوعين أفضل:**

:Renault $\bar{X}_1 = 7,325\ l$ و $S(X_1) = 0,92\ l$

:Peugeot $\bar{X}_2 = 7,5\ l$ و $S(X_2) = 1,2\ l$

أ- مقارنة مستوى الوزن:

نلاحظ أن $\bar{X}_1 < \bar{X}_2$ وعليه فإنه على العموم كمية البنزين التي تستهلكها سيارات رونو أقل مما هي عليه في سيارات بيجو.

ب- مقارنة تشتت الوزن: بما أن $\bar{X}_1 \neq \bar{X}_2$ فإننا نستخدم مقاييس التشتت النسبية لمقارنة تشتت كمية البنزين في الشركتين.

$$CV_1 = S(X_1)\% = \frac{S(X_1)}{\bar{X}_1} \times 100 == \frac{0,92}{7,325} \times 100 = 12,54\ \%$$

$$CV_2 = S(X_2)\% = \frac{S(X_2)}{\bar{X}_2} \times 100 == \frac{1,2}{7,5} \times 100 = 16\ \%$$

بما أن $CV_1 < CV_2$ فإننا نقول أن تشتت كميات البنزين المستهلكة من طرف سيارات شركة رونو أقل مما هي عليه في سيارات شركة بيجو، وعليه فإن كمية البنزين المستهلكة من قبل سيارات شركة رونو أكثر تجانس أو تقارب من كمية البنزين المستهلكة من قبل سيارات شركة بيجو.

الاستنتاج: سيارات رونو تستهلك كمية أقل من البنزين والفوارق بين الكميات قليلة (أكثر تجانس)، بينما سيارات بيجو تستهلك كمية أكبر من البنزين والفوارق بين الكميات كبيرة (أقل تجانس) ومنه السيارات من نوع رونو أفضل من السيارات من نوع بيجو.

بالتوفيق للجميع

يوم: السبت **2013/03/16**

المدة: ساعة و ربع

الامتحان الاستدراكي في مقياس الإحصاء1

الأقسام (من 6 إلى 10)

التمرين الأول:(12 نقاط)

البيانات التالية تمثل مدة اشتغال عينة من 30 مصباحا كهربائيا(الوحدة شهر) من إنتاج أحد المؤسسات:

0.5 – 2 – 2.5 – 3– 5 – 6 – 7.5– 8– 9– 9.5 – 10.5 – 12 – 12.5 – 13 – 13.5 –

14– 14 – 14.5 –15– 16 – 18 – 18– 19– 19.5– 20– 22– 25 – 26– 28 – 29.5

1- حدد المجتمع الإحصائي، الوحدة الإحصائية، المتغير الإحصائي المدروس ونوعه في هذا التوزيع.

2- أعرض هذه البيانات في توزيع تكراري على شكل فئات هي:[0-6[، [6-12[، [12-18[، [18-24[، [24-30]

3- مثل هذا التوزيع بيانيا.

4- أحسب المتوسط الحسابي على الجدول، واشرح النتيجة.

5- أحسب الوسيط على الجدول، واشرح النتيجة.

6- عين المنوال بيانيا (حدد قيمة تقريبية له من الشكل).

7- ما هي نسبة المصابيح التي تقل مدة حياتها عن 12 شهر.

8- ما هي نسبة المصابيح التي تفوق مدة حياتها عن 14.67 شهر.

التمرين الثاني: (8 نقاط)

أجريت دراسة إحصائية حول مردودية القمح على عينتين من المزارع بولايتي سطيف و تيارت للموسم الزراعي 2009-2010 (الوحدة ق/هـ) و أعطيت لنا المعلومات التالية :

ولاية تيارت: حجم العينة :10 مزارع ، $\sigma(x)=2.8$ ، $\sum x_i=120$ ، $\sum(x_i-\bar{x})^3=1.1756$ ، $\sum(x_i-\bar{x})^4=1721$

ولاية سطيف: حجم العينة :15 مزرعة ، $\sum x_i=225$ ، $\sum(x_i-\bar{x})^2=183.75$ ،

المطلوب:

1- قارن مستوى المردوية و التشتت في الولايتين.

2- أي الولايتين أفضل.

3- أدرس الإلتواء و التفرطح في ولاية تيارت.

بالتوفيق للجميع

جامعة سطيف-1-
كلية العلوم الاقتصادية والتجارية وعلوم التسيير
السنة الأولى ل.م.د –

يوم: السبت 2013/03/16

المدة: ساعة و ربع

حل الامتحان الاستدراكي في مقياس الإحصاء1

الأقسام (من 6 إلى 10)

حل التمرين الأول:(12 نقاط)

البيانات التالية تمثل مدة اشتغال عينة من 30 مصباحا كهربائيا(الوحدة شهر) من إنتاج أحد المؤسسات:

– 13.5 – 13 – 12.5 – 12 – 10.5 – 9.5 –9 –8 –7.5 – 6 – 5 –3 – 2.5 – 2 – 0.5

29.5 – 28 –26 – 25 –22 –20 –19.5 –19 –18 – 18 – 16 –15– 14.5 – 14 –14

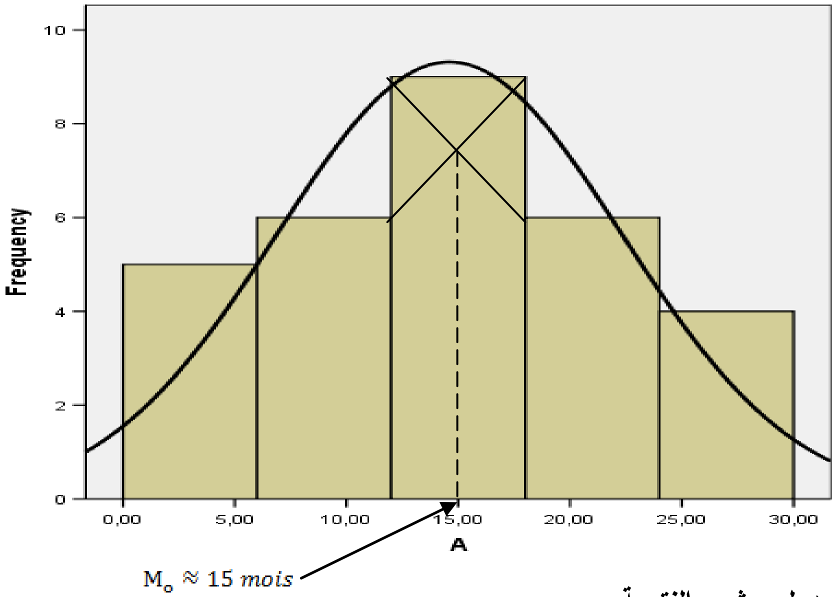
1– تحدد المجتمع الإحصائي، الوحدة الإحصائية، المتغير الإحصائي المدروس ونوعه في هذا التوزيع:

المجتمع الإحصائي	الوحدة الإحصائية	المتغير الإحصائي	نوعه
جميع المصابيح الكهربائية المنتجة بالمؤسسة	المصباح الواحد	مدة الاشتغال	كمي مستمر

2– عرض هذه البيانات في توزيع تكراري على شكل فئات هي:[0–6[، [6–12[، [12–18[، [18–24[، [24–30[

مدة الاشتغال (شهر)X_i	عدد المصابيح n_i	C_i	$C_i n_i$.	$N_i^{\uparrow}$.
$[0-6[$	05	3	15	05
$[6-12[$	06	9	54	11
$[12-18[$	09	15	135	20
$[18-24[$	06	21	126	26
$[24-30[$	04	27	108	30
المجموع $\sum n_i$	30	/	438	/

3– تمثيل هذا التوزيع بيانيا: بواسطة المدرج التكراري



4- حساب المتوسط الحسابي على الجدول، وشرح النتيجة:

$$\overline{X} = \frac{\sum_{i=1}^{5} n_i c_i}{n} = \frac{438}{30} = 14{,}6 \ mois$$

الشرح: متوسط مدة الاشتغال للمصابيح الكهربائية تقدر بـ: 14,6 شهرا

5- حساب الوسيط على الجدول، وشرح النتيجة:

– تحديد الفئة الوسيطية: وهي أول فئة تكرارها المتجمع الصاعد أكبر أو يساوي $\frac{n}{2}$ ، أي: $N^{\uparrow}_{M_e} \geq 15$

ومنه الفئة الوسيطية هي: $]18 - 12]$

– حساب الوسيط:

$$M_e = \mathrm{Lim}_{M_e} + \left[\frac{\frac{n}{2} - N^{\uparrow}_{M_e - 1}}{n_{M_e}}\right] \times A_{M_e} = 12 + \left[\frac{15 - 11}{9}\right] \times 6 = 14{,}67 \ \ mois$$

الشرح: 50% من المصابيح الكهربائية مدة اشتغالها تقل عن 14,67 شهرا و50% من المصابيح المتبقية مدة اشتغالها تفوق 14,67 شهرا.

6- تعيين المنوال بيانيا (تحديد قيمة تقريبية له من الشكل):

من خلال المدرج التكراري: $M_o \approx 15 \ mois$

7- نسبة المصابيح التي تقل مدة حياتها عن 12 شهر: $\frac{11}{30} \times 100 = 36{,}67\%$

8- نسبة المصابيح التي تفوق مدة حياتها عن 14.67 شهر: 50% لأنها تمثل قيمة الوسيط.

حل التمرين الثاني: **(8 نقاط)**

أجريت دراسة إحصائية حول مردودية القمح على عينتين من المزارع بولايتي سطيف و تيارت للموسم الزراعي 2009-2010 (الوحدة ق/ه) و أعطيت لنا المعلومات التالية :

<u>ولاية تيارت:</u> حجم العينة :10 مزارع ، $\sigma(x) = 2.8$ ، $\sum x_i = 120$ ، $\sum (x_i - \bar{x})^3 = 1.1756$ ، $\sum (x_i - \bar{x})^4 = 1721$

<u>ولاية سطيف:</u> حجم العينة: 15 مزرعة ، ، $\sum x_i = 225$ ، $\sum (x_i - \bar{x})^2 = 183.75$ ،

1- المقارنة بين مستوى المردوية و التشتت في الولايتين:

أ- حساب المردودية المتوسطة والانحراف المعياري في الولايتين:

- ولاية تيارت:

- المردودية المتوسطة: $\bar{X} = \frac{\sum_{i=1}^{10} X_i}{n} = \frac{120}{10} = 12$ ق/هـ

- الانحراف المعياري: $\sigma(X) = 2,8$ ق/هـ

- ولاية سطيف:

- المردودية المتوسطة: $\bar{Y} = \frac{\sum_{i=1}^{15} Y_i}{n} = \frac{225}{15} = 15$ ق/هـ

- الانحراف المعياري: $\delta(Y) = \sqrt{\frac{\sum_{i=1}^{15}(Y_i - \bar{Y})^2}{n}} = \sqrt{\frac{183,75}{15}} = 3,5$ ق/هـ

ب- المقارنة بين مستوى المردودية وكذلك التشتت في الولايتين:

- مقارنة مستوى المردودية:

بما أن $\bar{X} < \bar{Y}$ فإننا نقول على العموم مردودية القمح في ولاية سطيف أكبر مما هي عليه في ولاية تيارت.

- مقارنة التشتت:

بما أن $\bar{X} \neq \bar{Y}$ فإننا نستخدم مقاييس التشتت النسبية لمقارنة تشتت المردودية بين الولايتين، حيث نقوم بحساب كلا من CV_1 و CV_2 كما يلي:

$$CV_1 = \delta(X)\% = \frac{\delta(X)}{\bar{X}} \times 100 = \frac{2,8}{12} \times 100 = 23,33\%$$

$$CV_2 = \delta(Y)\% = \frac{\delta(Y)}{\bar{Y}} \times 100 = \frac{3,5}{15} \times 100 = 23,33\%$$

بما أن $CV_1 = CV_2$ أي هناك تساوي في التشتت بين الولايتين، أي أن الفوارق بين المزارع في مستوى المردودية المحققة هي على العموم نفسها.

2- أفضل ولاية:

مستوى مردودية مزارع ولاية سطيف من القمح أكبر مما هي عليه في ولاية تيارت، بينما الفوارق بين المزارع لكلا الولايتين متساوي ومنه ولاية سطيف أفضل من ولاية تيارت من حيث مرددية القمح.

3- دراسة الإلتواء والتفرطح في ولاية تيارت:

- دراسة الالتواء:

$$\alpha_F = \frac{\mu_3}{[\delta(X)]^3} = \frac{\frac{1,1756}{10}}{[2,8]^3} = 0,0053 \approx 0$$

بما أن $\alpha_F \approx 0$: فإن التوزيع متماثل أي 50% من المزارع تحقق مردودية أقل من 12 ق/هـ، 50% من المزارع تحقق مردودية أكبر من 12 ق/هـ

- دراسة التفرطح:

$$B_F = \frac{\mu_4}{[\delta(X)]^4} - 3 = \frac{\frac{1721}{10}}{[2,8]^4} - 3 = -0,2 \approx 0$$

بما أن $B_F \approx 0$: فإن التوزيع طبيعي.

بالتوفيق للجميع

المراجـــــــــــــــــــــــــع

1 – **موراي ر. شبيجل**، الإحصاء – سلسلة ملخصات شوم – ، ترجمة د. شعبان عبد الحميد شعبان و مراجعة د. أحمد حسن الموازيني، ط8؛ القاهرة: الدار الدولية للاستثمارات الثقافية، 2006.

2– **محمد عبد العال النعيمي وحسن ياسين طعمة**، الاحصاء التطبيقي، ط1، عمان، دار وائل للنشر والتوزيع، 2008.

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

4– **فتحي حمدان وكمال فليفل**، الإحصاء، ط1، عمان، دار المناهج للنشر والتوزيع، 2006.

5– **أحمد شكري الرماوي وسامي مسعود**، مقدمة في علم الإحصاء الوصفي والتحليلي، ط1، عمان، دار حنين، 1998.

6– **عبد العزيز فهمي هيكل**، مبادئ في الإحصاء التطبيقي، بيروت، دار الجامعة، 1985.

7– **أحمد فاروق عبد العظيم**، الإحصاء، الإسكندرية، المكتب الجامعي الحديث، 1984.

8– **أنيس كانجو**، الإحصاء وأسس تطبيقه في ميادين البحث العلمي، بيروت، مؤسسة الرسالة، 1980.

9– **ناظم حيدر**، مبادئ الإحصاء، ط8، دمشق، منشورات جامعة دمشق، 1996.

10– **محمد صبحي أبو صالح وعدنان محمد عوض**، مقدمة في الإحصاء، الجزائر، ديوان المطبوعات الجامعية، 1984.

11– **دلال صادق الجواد وحميد ناصر الفتال**، الأساليب الإحصائية في الإدارة، عمان: دار زهران للنشر والتوزيع، 2006.

12– **محمد عبد الرحمن إسماعيل محمد**، الرقابة الإحصائية على العمليات، الرياض: الإدارة العامة للطباعة والنشر بمعهد الإدارة العامة، 2006.

13– **ساعد بن فرحات**، محاضرات في الإحصاء1 (غير منشورة)، كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير، 2012 – 2013.

14– **شريف شطايبي**، محاضرات في الإحصاء الوصفي، الجزائر، مطبعة جامعة منتوري، 2004.

15- Bernard Verlant, *Statistique & Probabilités , Manuel de cours exercices corrigés – Sujets d'examens*

16- *GRAIS .B*, *Statistique Descriptive, France, dunod, 1977*

17- TORTRAT.A, Principe de Statistique, paris, dunod, 1967.

18- Lancry. P.J : *Statistique , étude de cas , Economica , paris , 1983*

فهرس المحتويات