



جامعة وهران 2

كلية العلوم الاقتصادية التجارية و علوم التسيير

مطبوعة

الإحصاء 1 Statistique1

ملخص دروس مع تمارين محلولة

السنة الأولى ليسانس

السداسي الأول

مقدمة من طرف :

السيد(ة): شامي يمينة

الرتبة: أستاذة محاضرة ب

السنة: 2024/2023

وحدة التعليم : منهجية

المادة : إحصاء 1

الرصيد: 5

المعامل: 3

نمط التعليم: حضوري

أهداف التعليم: يهدف هذا المقياس إلى وصف مجموعة من البيانات وتنظيمها وتصنيفها وتلخيصها وعرضها بطريقة واضحة في

صورة جداول أو أشكال بيانية وحساب المقاييس الإحصائية المختلفة لوصف متغير ما (أو أكثر) في مجتمع ما.

المعارف المسبقة المطلوبة: يحتاج الطالب فقط إلى معرفة أهم العمليات و القواعد الرياضية التي تم التطرق إليها في مرحلة التعليم

المتوسط و الثانوي..

المعارف والمهارات المستهدفة للمقرر : اكساب الطالب مهارات استخدام الاساليب الاحصائية في مجالات عديدة.

- استخدام الاساليب الاحصائية لجمع البيانات.
- استخدام الاساليب الاحصائية لعرض البيانات.
- استخدام الاساليب الاحصائية للتنبؤ.
- معرفة بعلم الإحصاء، وخطوات البحث العلمي، وأنواع العينات.
- يستطيع أن يبني جدول تكراري بناء على ما تعلمه بشكل صحيح.
- يميز بين مقاييس المركزية ومقاييس التشتت.
- يعرف الارتباط والانحدار.
- يحلل نتائج معامل الارتباط بين متغيرين.

طريقة التقييم: تقييم مستمر + إمتحان نهائي ويقاس معدل المادة بالوزن الترجيحي للدروس (60%) والأعمال الموجهة (40%)

فهرس المحتويات

1	مقدمة
3	الفصل الاول : مفاهيم اساسية في الاحصاء
3	تعريف علم الإحصاء:
4	الخطوات المنهجية لتحليل الإحصائي في البحث العلمي :
5	أنواع المتغيرات : تنقسم المتغيرات إلى نوعين:
5	متغيرات نوعية (كيفية)
6	متغيرات كمية (عددية)
7	المفاهيم الاحصائية :
8	سلسلة تمارين الفصل الأول محلولة :
17	الفصل الثاني : عرض البيانات الاحصائية
17	العرض الجدولي و البياني للمتغير الإحصائي الكيفي
19	العرض الجدولي والبياني للمتغير الإحصائي المنفصل:
20	العرض الجدولي والبياني للمتغير الإحصائي المستمر
24	تمارين الفصل الثاني محلولة
49	الفصل الثالث مقاييس النزعة المركزية
49	المتوسط الحسابي
50	المتوسط الهندسي
52	المتوسط التوافقي
53	المتوسط التربيعي
53	المنوال
54	الوسيط
56	تمارين الفصل الثالث محلولة
70	تمارين مقترحة :

76.....	الفصل الرابع مقاييس التشنت والشكل
76.....	تعريف مقاييس التشنت
76.....	مقاييس تشنت مطلقة
77.....	مقاييس تشنت نسبية
78.....	مقاييس الشكلن
78.....	مقاييس الالتواء
80.....	التفرطح
82.....	تمارين الفصل الرابع محلولة
92.....	الفصل الخامس الارتباط والانحدار الخطي
93.....	الارتباط الخطي
94.....	الانحدار الخطي البسيط
95.....	تمارين الفصل الخامس محلولة

مقدمة

مقدمة :

يعتبر الإحصاء من الوسائل العامة التي تستخدمها الباحثين في شتى المجالات، حيث يزودهم الإحصاء بالأدوات التي تساعد في تحليل المعطيات بشكل علمي دقيق، ومن ثم استخراج النتائج والتي بناء عليها يتم اتخاذ القرارات بناء على ما تقدم وانطلاقاً من أهمية الإحصاء جاءت هذه المطبوعة ملخص دروس و تمارين محلولة في مادة الإحصاء¹ بما ينسجم مع البرنامج المعتمد من طرف وزارة التعليم العالي و البحث العلمي والموجه لطلبة السنة الأولى ليسانس " السداسي الأول "الكلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير والعلوم المالية، ووتعتبر كمبادئ للإحصاء وفي نفس الوقت كتمهيد لباقي أنواع الإحصاء الأخرى مثل الاحتمالات و الإحصاء القياسي و الإحصاء الاستدلالي. تركز هذه المطبوعة على ملخص شرح المفاهيم الأساسية لمبادئ الإحصاء الوصفي ومتبوعة بسلسلة من تمارين محلولة.

تحتوي هذه المطبوعة على 5 فصول متسلسلة تتناسب مع القدرات المختلفة للطالب مع تبسيط واختصار لبعض المفاهيم، والمنهجية في العرض و على الفهم السريع، مطبوعة الإحصاء¹تناول العرض الجدولي والعرض البياني ومقاييس النزعة المركزية و مقاييس التشتت و الشكل و الارتباط و الانحدار و الأرقام القياسية، يجد الطالب في بداية كل فصل ملخص لكل عنصر من عناصر الدرس مع أمثلة تطبيقية بسيطة ثم يليه سلسلة من التمارين مع الحلول و أخرى مقترحة لتكون قاعدة عمل في حصص الأعمال الموجهة، ليتمكن الطلبة من حلها والاستعانة بها في مراجعتهم واستعدادهم للامتحانات النهائية .

الفصل الأول: مفاهيم أساسية في الإحصاء

سنتطرق من خلال هذا الفصل إلى العناصر التالية:

أولاً- تعريف الاحصاء

ثانياً- الخطوات المنهجية للتحليل الإحصائي في البحث العلمي

ثالثاً- أنواع المتغيرات .

رابعاً- سلسلة تمارين محلولة للفصل الاول

الفصل الاول: مفاهيم اساسية في الاحصاء

تعريف علم الإحصاء:

تعريف مصطفى الخواجة¹: "يقصد بعلم الإحصاء، الطريقة الإحصائية، وهي تلك الطريقة التي تمكن من:

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

و هو العلم الذي يبحث في تصميم أساليب جمع البيانات و التقنيات المختلفة لتنظيم و تصنيف و عرض البيانات و تلخيصها في صورة مؤشرات رقمية لوصف و قياس خصائصها الأساسية و تحليلها بغرض اتخاذ القرارات المناسبة

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؛ ويقسم الإحصاء إلى قسمين:

1 إحصاء وصفي Descriptive statistics: الذي يشمل الطرق المستخدمة في وصف وتنظيم وعرض البيانات، وتتمثل هذه الطرق في التوزيعات ومقاييس النزعة المركزية والتشتت والرسومات البيانية.

2 إحصاء استدلاي Inferential statistics: والذي يشمل استخلاص النتائج من بيانات محدودة وتعميمها على كل البيانات حيز الدراسة و من تطبيقاته التقدير و الاختبارات والتنبؤ.

تعريف جلاطو جيلاي²: "الإحصاء هو علم جمع وترتيب معلومات خاصة بظاهرة معينة وقياس الوقائع كأساس للاستقراء".

¹ مصطفى الخواجة: مقدمة في الإحصاء، الدار الجامعية، الإسكندرية، 2002، ص 2

² جلاطو جيلاي، الإحصاء مع تمارين ومسائل محلولة، 2002، ديوان المطبوعات الجامعية، ص 3.

تعريف جون جاك دراوزنيك³ : يعرف هذا العالم الإحصاء بأنه ذلك العلم الذي يشمل " مجموعة الطرق التي تهدف إلى معالجة المعطيات ... " ويقول عن موضوع علم الإحصاء: "يتعلق الأمر بمعرفة كيفية الحصول على تلك البيانات، معرفة من أين يتم تجميعها وبأي شكل يكون ذلك التجميع."

تعريف دومنيك سلفاتور⁴ : " الإحصاء هو مجموع الطرق الرياضية المتعلقة بجمع، وعرض، وتحليل، واستخدام المعطيات الرقمية. هذه العمليات تمكن من استخلاص استنتاجات واتخاذ قرارات إزاء حالة عدم التأكد التي نواجهها في مجال الاقتصاد ومجال الأعمال أو في علوم اجتماعية وفيزيائية أخرى."

يعتبر علم الإحصاء احد الوسائل المهمة في البحث العلمي من خلال استخدام قواعده وقوانينه وطرقه في جمع البيانات و المعلومات اللازمة في البحث العلمي وتحليل هذه المعلومات لغرض الوصول إلى النتائج التي يهدف لها البحث كما وان للإحصاء دورا بارزا في وضع الخطط المستقبلية عن طريق التنبؤ بالنتائج لكافة القطاعات حيث يمكن تطبيق علم الإحصاء في مجالات العلوم الصرفة أو العلوم الإنسانية وفيما يلي سنورد أمثلة لبعض المجالات التي يستعمل فيها الإحصاء والتي كان له دور بارز في حل كثير من مشاكلها:

- يستخدم الإحصاء في تطوير التعليم و خططه.
- يستخدم الإحصاء في دراسة مختلف العلوم.
- يستخدم الإحصاء في مجال الدعاية والإعلانات التجارية.
- يستخدم بشكل كبير في مجال التأمين.
- يستخدم في مجال الاقتصاد والصناعة

الخطوات المنهجية للتحليل الإحصائي في البحث العلمي :

³ جون جاك دراوزنيك، أساسيات في الإحصاء، سلسلة (SMA)، دار (Ellips)، 1996، ص 2.

⁴ دومينيك سالفاتور، الاقتصاد القياسي والإحصاء التطبيقي، سلسلة شوم، دار ماك قراو هيل، 1985، ص 1.

جمع البيانات : عملية الحصول على المعلومات أو قيم المشاهدات أو القياسات للتجارب التي يجريها الإحصائي

تنظيم و عرض المعلومات : عملية وضع المعلومات في جداول منسقة و عرضها بطرق مناسبة كالأشكال الهندسية و الرسوم البيانية و غيرها

تحليل البيانات : عملية إيجاد قيم لمقاييس و اقترانات معينة تحدد قيمها من البيانات موضع الدراسة.

استقراء النتائج و اتخاذ القرارات: الاستنتاجات التي يصل إليها الباحث و تكون على شكل تقديرات أو تنبؤات أو تعميمات أو قرار برفض أو قبول الفرضيات الإحصائية.

مصادر البيانات :

البيانات (Data) : مجموعة القيم التي يتم جمعها من مفردات المجتمع أو العينة لخاصية معينة (متغير).

يعتمد الباحثون على مصدرين أساسيين للحصول على المعلومات الخاصة بظاهرة معينة:

1المصادر المباشرة : يجمع الإحصائي معلوماته إما بالاتصال المباشر بوحدة المجتمع الإحصائي أو بالاعتماد

على الوثائق التي تكون فيها المعلومات لا زلت خامة و في هذه الحالة يمثل الاستبيان أو المقابلة.

2المصادر غير المباشرة : يتحصل الباحث على المعلومات الإحصائية من الدراسات والتحقيقات السابقة مثلا

البيانات من طرف باحثين سابقين أو هيئات رسمية أو غير رسمية وتم نشرها في نشرات خاصة أو دوريات.

أنواع المتغيرات : تنقسم المتغيرات إلى نوعين:

1 – متغيرات نوعية (كيفية) Variable qualitative : وهي عبارة عن صفات أو أنواع معينة ليست

عددية، وتنقسم بدورها إلى:

– متغيرات نوعية خاضعة للترتيب: مثل المستوى التعليمي، الرتب العسكرية تقديرات النجاح المستوى الاقتصادي... الخ.

—متغيرات نوعية غير خاضعة للترتيب: مثل الجنسية، أنواع السيارات، أنواع الأمراض... إلخ.

2 - متغيرات كمية (عددية) **Variable quantitatives**: وهي البيانات التي يعبر عنها في صورة

عددية وتنقسم إلى:

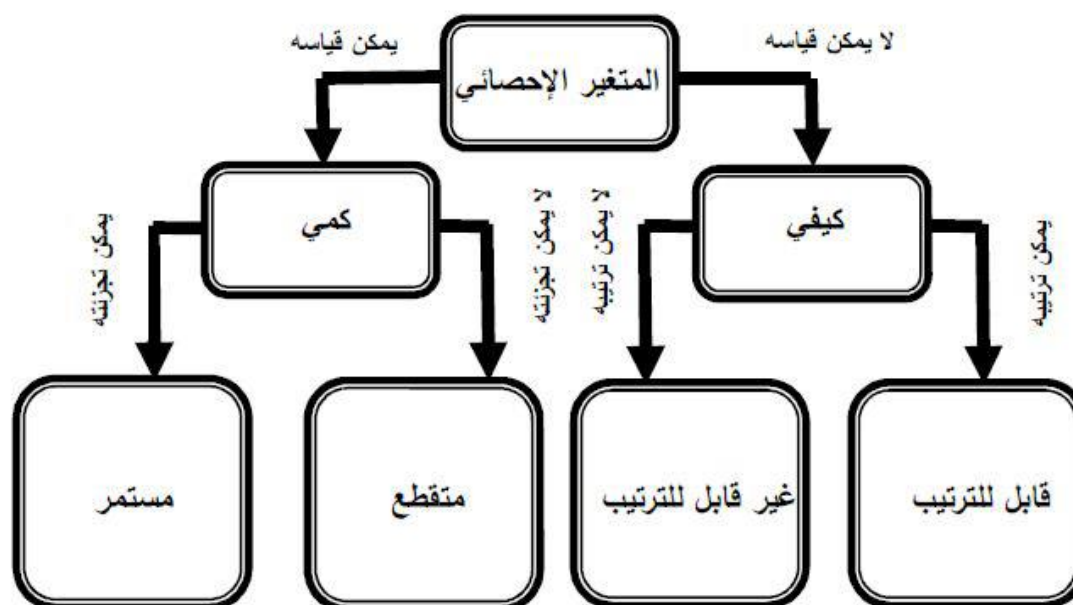
— متغير متقطع **Variable discret**: وهو المتغير الذي يأخذ أعداد صحيحة، فمثلاً إذا كان x متغير يمثل عدد أفراد

الأسرة، فإنه يمكن أن يأخذ القيم 2، 3، 4، 5... ولا يمكن أن نأخذ x القيم 1.5، 3.25، 5.17.

— متغير متصل (مستمر) **Variable Continue**: وهو المتغير الذي لا يمكن أن يأخذ أي قيمة بين قيمتين معينتين،

وكأمثلة عن المتغيرات المتصلة: الطول، الوزن، الزمن، السرعة.

مخطط أنواع المتغيرات الإحصائية



المصدر: مساعد بن فرحات، عبد الحميد قطوش، (الإحصاء 1: مدعمة بتمارين وامتحانات محلولة) جامعة فرحات عباس سطيف

2 (2013/2014) ص 15

المفاهيم الاحصائية **Statistical concepts** : علم الإحصاء مثل العلوم الأخرى بأنه مليء بالتعاريف والمصطلحات الجوهرية التي يستلزم فهم معانيها منها:

- **المجتمع الإحصائي (Statistical Population)**: يمثل جميع الوحدات الإحصائية التي تشترك في صفة أو أكثر ونرغب في دراستها، فقد يكون المجتمع يمثل مجموعة من مندوبي المبيعات أو المؤسسات، وقد يكون المجتمع محدوداً مثل عدد المندوبين في مدينة أو غير محدود مثل عدد النجوم في السماء..

- **الوحدة الإحصائية**: هي الوحدة الأساسية و الفرد الأساسي لتكوين المجتمع الإحصائي.

- **المتغير variable**: هو الصفة تحت الدراسة كما أشرنا أعلاه أو هو الشيء الذي يمكن أن يأخذ قيماً مختلفة في الظروف المختلفة (زمنية، مكانية، سياسية، اقتصادية ... الخ) فمثلاً سعر التمر يختلف من يوم لآخر ويختلف في نفس السوق من سنة لأخرى.

- **العينة الإحصائية (Statistical Sample)** : هي جزء من المجتمع الإحصائي يتم اختيار وحداتها لتكون ممثلة للمجتمع كله أحسن تمثيل. والعينة الإحصائية أنواع نذكر منها:

1. **العينة العشوائية البسيطة**: هي العينة التي تعطي فيها لجميع وحدات المجتمع المراد بحثه نفس الفرصة في الاختيار.
2. **العينة الطبقية**: هي العينة التي تخص المجتمعات غير المتجانسة حيث تكون في حالة كان المجتمع يتكون من الوحدات التي تتصف بالتجانس داخل كل مجموعة وبالتباين بين المجموعات المختلفة.
3. **العينة متعددة المراحل**: تكون في حالة كان المجتمع يتكون من أقسام متجانسة فيتم البدء باختيار هذه الأقسام عشوائياً كمرحلة أولى ثم يتم إختيار عينة عشوائية بسيطة من كل قسم من الأقسام التي تم إختيارها في المرحلة الأولى كمرحلة ثانية.

سلسلة تمارين الفصل الأول محلولة :

التمرين الأول:

حدد المجتمع الإحصائي، الوحدة الإحصائية، المتغير الإحصائي ونوعه من واقع العبارات التالية:

- مدة حياة المصابيح الكهربائية المنتجة في مصنع.
- الأجور الشهرية لعمال مؤسسة ما.
- أوزان طلبة جامعة التكوين المتواصل.
- الرياضة الممارسة من طرف طلبة جامعة بسكرة.
- عدد الغرف المملوكة من طرف كل عائلة في حي معين.
- تصنيف عمال مصنع حسب المؤهل.
- عدد السيارات في دولة حسب الصنف.
- تصنيف الأحزاب السياسية حسب عدد الأصوات المكتسبة في الانتخابات

حل التمرين الأول:

الرقم	المجتمع الإحصائي	الوحدة الإحصائية	المتغير	نوع المتغير
01	المصابيح الكهربائية	المصباح	مدة الحياة	كمي متميز
02	عمال المؤسسة	العامل	الأجر	كمي متميز
03	طلبة جامعة التكوين المتواصل	الطالب	الوزن	كمي متميز
04	طلبة جامعة بسكرة	الطالب	الرياضة	نوعي
05	عائلات الحي	العائلة	عدد الغرف	كمي متقطع
06	العمال	العامل	المؤهل	نوعي
07	السيارات	السيارة	الصنف	نوعي
08	الأحزاب السياسية	الحزب	عدد الأصوات	كمي متقطع

التمرين الثاني : أعط طبيعة الصفة للخصائص التالية:

فصيلة الدم- لون البشرة -دخل رب الأسرة- عدد أطفال الأسرة -درجة الذكاء -نوع الكلية- رقم الهاتف.

حل التمرين الثاني :

المتغير	وصفي	كمي منفصل	كمي متصل
فصيلة الدم	نعم		
دخل رب الأسرة			نعم
درجة الذكاء	نعم		
نوع الكلية	نعم		
رقم الهاتف		نعم	
عدد أطفال الأسرة		نعم	
لون البشرة	نعم		

التمرين الثالث : حدد نوع العينة للحالات التالية

- عند إجراء دراسة على مجتمع ما تم تقسيم المجتمع إلى متزوج و أعزب ثم اختيار عينة عشوائية

بسيطة من كل منهما تتناسب و نسبة كل منهما.

- عن طريق استخدام جهاز الحاسب في توليد 50 رقماً أكاديمياً تم اختيار أصحاب هذه الأرقام

كعينة.

- لإجراء دراسة على الطلبة مبادئ الإحصاء تم اختيار 4 شعب عشوائياً و دراسة جميع طلبة هذه

الشعب.

حل التمرين الثالث :

عشوائية بسيطة	عند إجراء دراسة على مجتمع ما تم تقسيم المجتمع إلى متزوج و أعزب ثم اختيار عينة من كل منهما تتناسب و نسبة كل منهما.
عينة بسيطة	عن طريق استخدام جهاز الحاسب في توليد 50 رقماً أكاديمياً تم اختيار أصحاب هذه الأرقام كعينة.
عينة عنقودية	لإجراء دراسة على طالبات مبادئ الإحصاء تم اختيار 4 شعب عشوائياً و دراسة جميع طالبات هذه الشعب

التمرين الرابع : حدد الأسلوب الأنسب لجمع البيانات في كل مما يلي :

- أراد باحث اقتصادي دراسة حجم الصادرات السنوية للجزائر من النفط أفضل أسلوب لجمع البيانات
- للبحث عن أفضل جهاز Tablet للاستخدام المكتبي نستخدم الانترنت لجمع معلومات عن جميع الأجهزة المتوفرة في السوق.
- تطبيق عدة أساليب غذائية لاكتشاف أفضلها في انقاص الوزن.

حل التمرين الرابع :

السلاسل الزمنية	أراد باحث اقتصادي دراسة حجم الصادرات السنوية للجزائر من النفط أفضل
مسح شامل	للبحث عن أفضل جهاز Tablet للاستخدام المكتبي نستخدم الانترنت لجمع معلومات عن جميع الأجهزة المتوفرة في السوق.
الاسلوب الشامل	تطبيق عدة أساليب غذائية لاكتشاف أفضلها في انقاص الوزن

التمرين الخامس: تحديد كلا من المجتمع الإحصائي، الوحدة الإحصائية، المتغير الإحصائي، نوع المتغير الإحصائي من الأمثلة المعطاة:

الدراسة	المجتمع الإحصائي	الوحدة الإحصائية	المتغير الإحصائي	نوع المتغير الإحصائي
الاولى	عمال شركة متعددة الجنسيات للمحروقات و الغاز	عامل	الوظيفة	كيفي قابل للترتيب
			السن	كمي متصل
			الحالة العائلية	كيفي غير قابل للترتيب
			الدخل	كمي متصل
			المستوى الدراسي	كيفي قابل للترتيب
			الجنس	كيفي غير قابل للترتيب
			مكان الميلاد	كيفي غير قابل للترتيب
			الجنسية	كيفي غير قابل للترتيب
الثانية	مبيعات مؤسسة		تاريخ البيع	
			الوزن	
			مكان البيع	كيفي قابل للترتيب
			السعر	
الثالثة	طلبة جامعيين		المستوى الدراسي	كيفي قابل للترتيب
			طبيعة التخصص	كيفي غير قابل للترتيب
			السن	كمي متصل

الطول	كي متصل			
مكان الميلاد	كيفي غير قابل للترتيب			
سعر المتر مربع	كمي متصل		مجموعة سكنات	الرابعة
عدد الغرف	كمي منفصل			
موقع السكنات	كيفي غير قابل للترتيب			
طابق السكن	كيفي قابل للترتيب			

تمارين مقترحة :

تمرين 1

- عرف كل من الإحصاء الوصفي والإحصاء الاستدلالي؟
2. المجتمع الإحصائي هو جزء من المجتمع الكلي في منطقة ما. صح أو خطأ أو غير دقيق؟
3. أيهما أدق القول ان العينة الإحصائية جزء من المجتمع الإحصائي أم جزء من المجتمع الكلي في منطقة ما؟
4. حدد المجتمع المحدود وغير المحدود في التالي:
- أ. سمك الشعور في المحيط الهندي
- ب. شباب المملكة بين الأعمار 20 و 25 سنة.
- ج. المكالمات الهاتفية الواردة لمدينة الرياض خلال يوم.
- د. علب العصير في مصنع الإنتاج لدورة إنتاجية.
- هـ. الأعداد الحقيقية بين الصفر والواحد.
- و. ذرات الهيدروجين في زجاجة مياه سعة لتر.
5. هل يمكن للعينة الإحصائية أن تتكون من جميع مفردات المجتمع الإحصائي؟ علل إجابتك.
6. أي من الجمل التالية صحيح أو خطأ أو غير دقيق:
- أ. كل مفردات المجتمع الإحصائي هي متغيرات.
- ب. كل مفردات العينة الإحصائية هي متغيرات.
- ج. طول الفرد في مجتمع إحصائي هو متغير.
- د. وزن الفرد في عينة إحصائية هو متغير.

هـ. نوع الجنس في عينة من 50 شابا هو متغير.

و. عدد ركعات النوافل والسنن هو متغير.

7. صنف المتغيرات التالية لكمية ووصفية:

أ. أطوال عينة من الأطفال.

ب. أوزان حمولة من البطيخ.

ج. أنواع الطماطم في سوق عتيقة.

د. أصناف الكتب في معرض دولي.

هـ. عدد كل صنف من الكتب في معرض دولي.

8. ما الفرق بين البيان (مفرد بيانات) والمتغير؟

9. صنف التالي لمعلمة أو إحصاءة:

أ. متوسط أطوال عينة من الأطفال.

ب. متوسط أوزان حمولة من البطيخ.

ج. أغلب نوع من أنواع الطماطم في سوق عتيقة.

د. أكثر صنف من أصناف الكتب في معرض دولي.

هـ. متوسط عدد كل صنف من الكتب في معرض دولي.

تمرين 2:

(1) اعط طبيعة الصفة للخصائص التالية:

درجة الحرارة ، الجنسية ، التقدير في شهادة البكالوريا ، عدد الطلاب في الفوج، التوزيع حسب الفئات المهنية الاجتماعية، نوعية المنتج، السن ، الحالة العائلية، تزايد مجموعة من الطلاب حسب مكان الازدياد ،درجة رضا المستهلك من منتج جديد

(2) حدد كصفات الخصائص التالية :

الجنس ، سن العامل ، الحالة العائلية، عدد الأطفال في العائلة ، الفئات المهنية الاجتماعية

(3) ماهي الصفات التي يمكن أن نجدها لما نقوم بدراسة حول مجموعة من المؤسسات في ولاية ما؟

الفصل الثاني: عرض البيانات الإحصائية

سننتقل من خلال هذا الفصل إلى العناصر التالية:

أولاً- التوزيع التكراري للمتغير الإحصائي الكيفي

ثانياً- العرض الجدولي والبياني الإحصائي المنفصل (المتقطع).

ثالثاً- العرض الجدولي للمتغير الإحصائي المتصل (المستمر

رابعاً- سلسلة تمارين الفصل الثاني محلولة.

الفصل الثاني: عرض البيانات الإحصائية

العرض الجدولي و البياني للمتغير الإحصائي الكيفي (الصفة النوعية):

يعتبر العرض الجدولي أحد الأدوات الإحصائية المستعملة قصد تصنيف ، ترتيب ، تجميع و عرض الوحدات الإحصائية⁵.

يستخدم جدول التوزيع التكراري لتصنيف و تلخيص البيانات الإحصائية بشكل بسيط و يسهل عملية قراءة مفردات موضوع الدراسة و يتكون هذا الجدول من عمودين (أو سطرين). أولهما يخص للظاهرة أو الصفة المدروسة (نوعية أو كمية) ، بينما يخص العمود الثاني لعملية تفرغ البيانات موضوع الدراسة و ذلك بعدد مشاهداتها و تسجيل كل منها مقابل الصفة أو الظاهرة المدروسة .

- حيث أن n_i هو عدد المشاهدات لكل كيفية

- C_i هو كيفيات الصفة النوعية

- X_i هو كيفيات الصفة الكمية

- N أو $\sum n_i$ هو المجتمع الإحصائي أو مجموع التكرارات المطلقة

$$N = n_1 + n_2 + n_3 + \dots + n_k = \sum_{i=1}^k n_i$$

f_i التكرارات النسبية و هي عبارة عن ناتج قسمة التكرارات المطلقة على مجموع التكرارات

$$f_i = \frac{n_i}{\sum n_i} = \frac{n_i}{N} \quad \text{و} \quad \sum f_i = 1$$

⁵ هشام بورمة, احصاء 1 دروس و امثلة ,تطبيقية 2017/2018, جامعة جيجل. ص.16.

fi% التكرارات المئوية : هي عبارة عن حاصل ضرب التكرارات النسبية في 100

$$fi\% = f_i \times 100 = \frac{n_i}{\sum n_i} \times 100 \quad \text{و} \quad \sum f_i\% = 100\%$$

- التوزيع التكراري المتجمع (الصاعد) $\sum N_{Ci}$ و النازل (N_{Ci}) نحصل على التكرار المتجمع

الصاعد بتجميع التكرارات بطريقة متتالية من بداية الجدول و نحصل على التكرار المتجمع

النازل بوضع مجموع التكرارات في أول العمود ثم نطرح بطريقة متتالية .

العرض الجدولي صفة نوعية :

التكرارات		الكيفية (طبيعة الصفة المدروسة)
n_i		صفة نوعية
		c_i
n_1		C1
n_2		C2
n_3		C3
:		:
n_k		Ck
$N = \sum n_i$		total

العرض البياني للصفة النوعية⁶ :

التمثيل الدائري: تعتمد هذه الطريقة على المساحات ، و هي عبارة عن دائرة مقسمة إلى قطاعات، كل جزء

يمثل كيفية للصفة المدروسة. الزاوية المركزية التي تمثل كل جزء تحسب كما يلي: $\alpha^\circ = f_i \times$.

⁶ Dominick SALVATORE & Derrick REAGLE, Theory and Problems of Statistics and Econometrics, Second Edition, Schaum's Outline Series, McGRAW-HILL, companies, New York, 2002

القضبان البيانية (المستطيلات أو الأشرطة البيانية) : عبارة عن مستطيلات أو أشرطة عمودية الشكل، قواعدها

متساوية بينما ارتفاعها يتناسب مع التكرارات التي تتخذها الظاهرة (تكرارات مطلقة، نسبية أو مئوية)، مع وضع

كل مستطيل على بعد متساوي مع المستطيل الآخر.

القضبان البيانية المزدوجة: يستخدم هذا الشكل من التمثيلات في حالة الصفات النوعية ذات عدة أوجه.

العمود المجزأ: يستعمل هذا التمثيل هو الآخر في حالة الصفة النوعية و هو عبارة عن مستطيل مقسم إلى أجزاء

حيث يمثل كل جزء كيفية من كفيات الظاهرة المدروسة و التي تناسب تكرارها المئوي. طول العمود يساوي %

100.

العرض الجدولي والبياني للمتغير الإحصائي المنفصل (المتقطع) :

العرض الجدولي للمتغير الإحصائي المنفصل (المتقطع)

التكرارات	الكيفية (طبيعة الصفة المدروسة)
ni	صفة كمية xi
n1	X1
n2	X2
:	:
:	:
:	:
nk	Xk
$N = \sum ni$	total

البياني الإحصائي المنفصل (المتقطع)⁷ :

⁷ فريد بن ختو, محاضرات في الإحصاء الوصفي, 2018/2017 جامعة بسكرة, ص 23.

الأعمدة البيانية : هي عبارة أعمدة بسيطة عمودية الشكل تنطلق من المحور الأفقي الخاص بالمتغير و حيث تتناسب أطوالها مع التكرار (المطلق، النسبي أو المئوي) المقابل لقيمة المتغير.

العرض البياني للتكرارات التجميعية الصاعدة و النازلة: هو عبارة عن قطع مستقيمة أفقية متصاعدة أو متنازلة لقيم المتغير و التي تتناسب مع التكرار الصاعد أو النازل المقابل له

العرض الجدولي والبياني للمتغير الإحصائي المستمر (متصل):
العرض الجدولي للمتغير الإحصائي المستمر (متصل) :

هو المتغير الإحصائي الذي يمثل قيم الظاهرة المدروسة التي تخضع لمنطق التجزئة كأوزان مجموعة من الطلبة أو أجور مجموعة من عمال مؤسسة ما، وفي أغلب الأحيان يتم ترتيب هذه القيم داخل فئات حيث تحتوي هذه الأخيرة (الفئات) على مجموعة من قيم الظاهرة، و لأجل ذلك و قصد إعداد جدول التوزيع التكراري يجب إتباع الخطوات التالية⁸:

(1) تحديد المدى العام:

يمثل المدى العام الفرق بين أكبر قيمة و أصغر قيمة للسلسلة الإحصائية و يرمز له بـ **E**.

$$E = X_{i \max} - X_{i \min}$$

(2) تحديد عدد الفئات:

لحساب عدد الفئات يتطلب منا أخذ بعين الاعتبار حجم المجتمع المدروس أو العينة المختارة و نرمز لها بـ **NC**

♦ إذا كان: **N < 100** نقوم باستعمال قاعدة يول (YULE)

⁸ حيدوشي عاشور، محاضرات في الإحصاء الوصفي 2015، / 2016 جامعة البويرة. ص. 20.

$$NC = 2,5\sqrt[4]{N}$$

♦ إذا كان: $N \geq 100$ نقوم باستعمال قاعدة ستروج (STURGE) و التي تعتمد في حسابها على:

$$NC = 1 + [3,322(\log N)]$$

نقوم باستعمال اللوغاريتم العشري

(3) تحديد طول الفئات:

و هو محاولة تجزئة المدى العام إلى مجموعات مصغرة، أطوالها متساوية تسمى بالفئات و تحسب كالتالي:

$$a_i = \frac{E + 1}{NC}$$

(4) تحديد حدود الفئات:

لكل فئة حد أدنى و حد أعلى حسب الشكل المقابل:

- L_0 : هو الحد الأدنى للفئة الأولى وهو عبارة عن

أصغر قيمة في السلسلة الإحصائية . $L_0 = X \min$

- L_1 : هو الحد الأعلى للفئة الأولى و هو عبارة عن الحد

الأدنى للفئة الأولى مضاف إليه طول الفئة a_i

$$L_1 = L_0 + a_i$$

نحدد حدود الفئات المتبقية و ذلك علما أن الحد الأعلى لكل فئة يصبح الحد الأدنى للفئة الموالية، وبعد إضافة طول الفئة

نحصل على الحد الأعلى للفئة، و هكذا إلى أن نصل إلى الحد الأعلى لأخر فئة

(5) تحديد مراكز الفئات:

تعتبر مراكز الفئات X_i عبارة عن منتصفات الفئات التي تم تحديدها، و يتم حسابها بالشكل التالي:

$$X_1 = \frac{L_0 + L_1}{2}, X_2 = \frac{L_1 + L_2}{2}, X_3 = \frac{L_2 + L_3}{2} \dots \dots \dots$$

نموذج جدول توزيع تكراري لصفة كمية متصلة او مستمرة

الفئات	عدد الموظفين n_i	الحدود العليا "أقل من الحد الأعلى"	N_{Ci}	الحدود الدنيا "الحد الأدنى فأكثر"	N_{Ci}
$[L_0 - L_1[$	n_1	أقل من L_1	n_1	L_0 فأكثر	$\sum n_i$
$[L_1 - L_2[$	n_2	أقل من L_2	$n_1 + n_2$	L_1 فأكثر	$\sum n_i - n_1$
$[L_{k-1} - L_k[$	n_k	أقل من L_k	$\sum n_i$	L_{k-1} فأكثر	N_k
$\sum$ المجموع	$\sum n_i$	/	/	/	/

العرض البياني للمتغير الإحصائي المستمر (متصل):

المدرج التكرارى : هو عبارة عن أعمدة مستطيلة متلاصقة يعبر ارتفاع العمود فيها على التكرار المناظر للفئة.

ويستخدم هذا النوع من الرسوم البيانية لتمثيل البيانات التي تم عرضها في جدول توزيع تكرارى، وفيه يمثل كل

مستطيل فئة من فئات التوزيع التكرارى

المضلع التكرارى

هو ذلك النوع من الرسوم البيانية الذي يمكن الحصول عليه من خلال حساب مراكز الفئات أو بتتصيف

الأضلاع العلوية للمستطيلات في المدرج التكرارى، ثم نوصل هذه النقاط بعضها مع بعض.

المنحنى التكرارى : ونحصل عليه إذا مهدنا المضلع التكرارى وجعلناه منحنى بدلا من خطوط منكسرة فإننا نحصل

على المنحنى التكرارى

المنحنى المتجمع الصاعد : لتمثيل التوزيع التكرارى المتجمع الصاعد، سواء أكان بالقيم المطلقة للتكرارات، أو

بالتكرار النسبي. ويراعى وضع النقاط الخاصه بالتكرارات في حالة المنحنى المتجمع الصاعد عند الحد الأعلى لكل

فئة، لأنه يعبر عن العدد الاجمالى لأوجه الظاهرة الواقع أسفل الحد الأعلى للفئة

المنحنى المتجمع الهابط (النازل) : لتمثيل التوزيع التكرارى المتجمع الهابط (النازل) أيضاً بالقيم المطلقة للتكرارات

أو بالتكرار النسبي، ويراعى وضع النقاط الخاصه بالتكرارات المتجمعه الهابطه (النازلة) عند الحد الأدنى لكل فئة،

لأنه يعبر عن العدد الاجمالى لأوجه الظاهرة الواقع أعلي الحد الأدنى للفئة.

ملاحظة: المدرج التكرارى في حالة عدم تساوي أطوال الفئات: في هذه الحالة نلاحظ بان قواعد المستطيلات غير

متساوية و بالتالى مساحة المستطيلات لا تتناسب مع التكرارات الأصلية، وحتى تصبح مساحة المستطيلات متناسبة مع

ارتفاعها نقوم بتعديل التكرارات قبل رسم المدرج التكرارى و ذلك بإتباع الخطوات التالية:

— حساب طول كل فئة (ai)

— تحديد طول الفئة الواحدوي (au) و ذلك باختياره بأحد المعايير التالية:

أكبر قاسم مشترك (PGDC) لأطوال الفئات إن وج دأو طول الفئة الأكثر شيوعاً أو أصغر

طول فئة

$$ci = \frac{ai}{au} : (ci) \text{ حساب المعاملات}$$

$$hi = \frac{ni}{ci} : (hi) \text{ حساب التكرار المعدل}$$

تمارين الفصل الثاني محلولة

التمرين الاول : الجدول التالي يبين مؤهلات منسوبي إحدى الشركات

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

والمطلوب :

- 1 - عرض البيانات. في جدول تكراري
- 2- كون التوزيع التكراري النسبي، المئوي ثم علق على النتائج.

حل التمرين الاول :

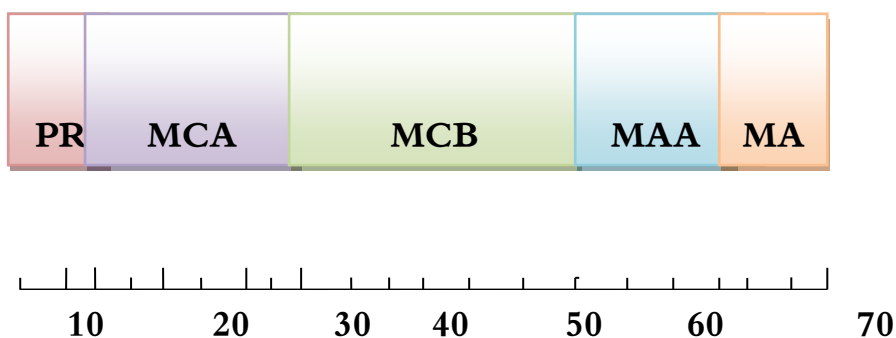
1- نكون الجدول التكراري

المؤهل العلمي (الفئات)	العلامات	Fi	Fi%
دكتوراه	2	$2/35=0.057$	$0.057*100=5.7\%$
جامعي	13	$13/35=0.371$	$=37.1\%$
ثانوي	16	$16/35=0.457$	45.7%
متوسط	3	$3/35=0.086$	8.6%
ابتدائي	1	$1/35=0.029$	2.9%
	$\sum n_i = 35$	$\sum f_i = 1$	$\sum f_i \% = 100\%$

التمرين الثاني يمثل الجدول التالي توزيع الأساتذة الجامعيين حسب درجتهم العلمية لكلية العلوم الاقتصادية، علوم التسيير و العلوم التجارية، جامعة وهران 2 .

1. حدد المجتمع المدروس ، الصفة ، طبيعتها و الكيفيات؟
2. مثل بيانيا هذه المعطيات بواسطة العمود المجرأ .
3. مثل بيانيا هذه المعطيات بواسطة القطاعات الدائرية.

الدرجة العلمية للأساتذة	التكرارات n_i	$F_i\%$
C_i		
PRF	20	8,696%
MCA	65	28,261%
MCB	75	32,609%
MAA	45	19,565%
MAB	25	10,87%
$\sum$ المجموع	230	100%



العمود المجزأ لتوزيع الأساتذة حسب الدرجة العلمية

3. التمثيل الدائري:

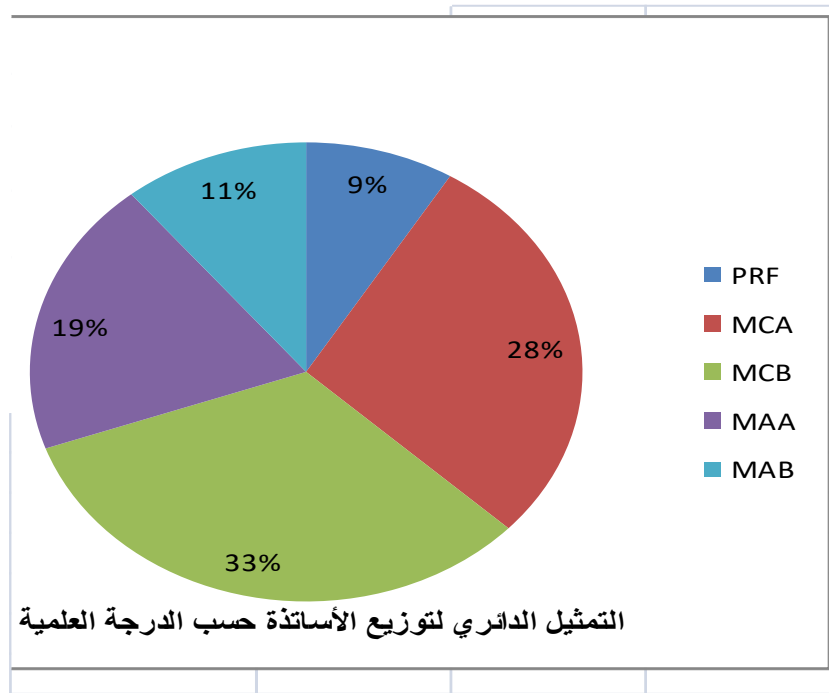
الدرجة العلمية للأساتذة Ci	التكرارات ni	fi=ni/N	$\alpha^\circ = f_i \times 360^\circ$
PRF	20	0,08696	31,30435°
MCA	65	0,28261	101,73913°
MCB	75	0,32609	117,3913°
MAA	45	0,19565	70,43478°
MAB	25	0,1087	39,13023°
المجموع Σ	230	1	360°

المجتمع المدروس: الأساتذة الجامعيين

الصفة : الدرجة العلمية

طبيعتها : نوعية ترتيبية

الكيفيات : MAB MAA MCB MCA PRF

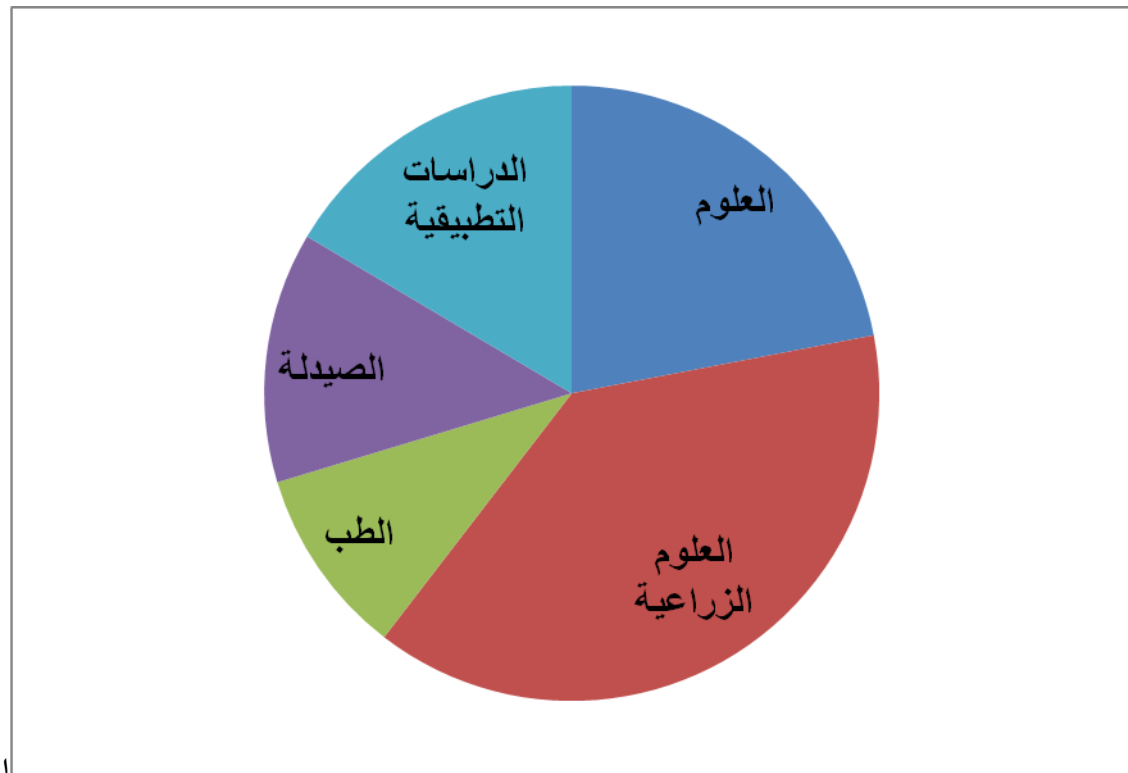


التمرين الثالث : الجدول الموالي يعطي بيانات حول عدد الطلبة لكل اختصاص جامعي :

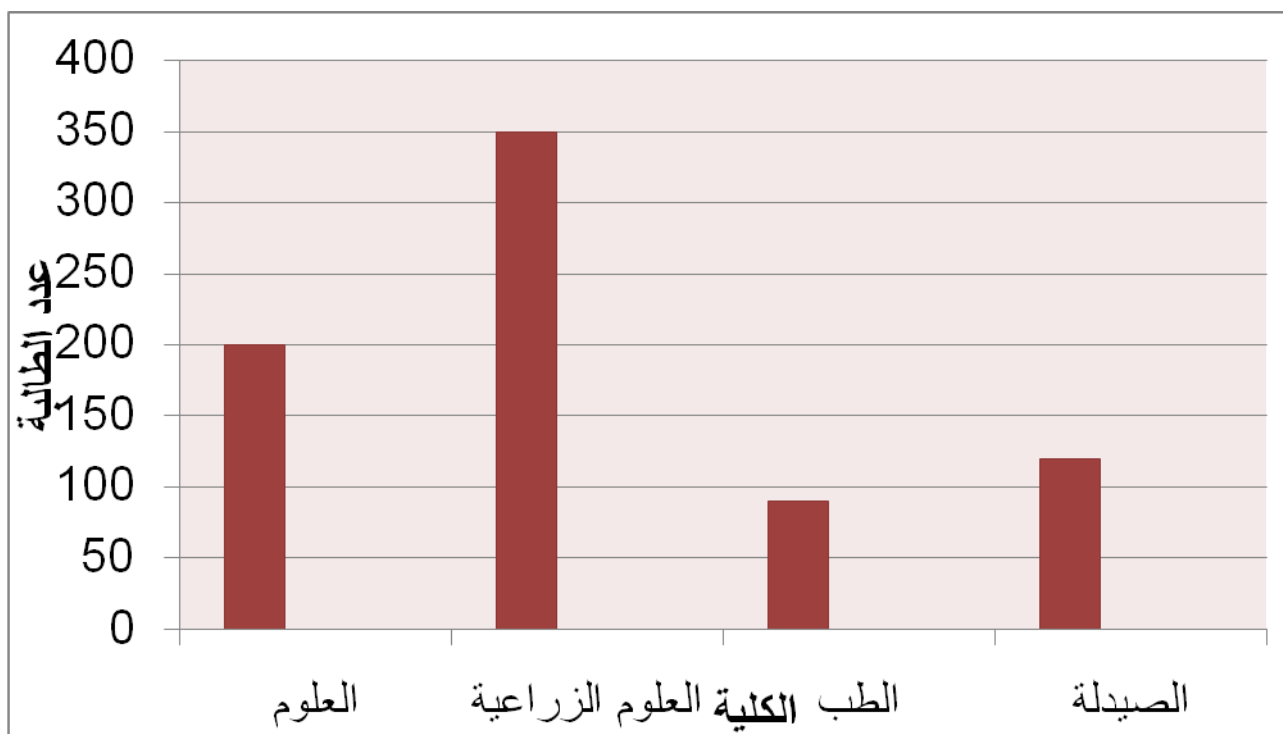
الكلية	عدد الطلبة	fi	fi%	Nci	Nci	$\alpha^\circ = \frac{fi}{Nci} \times 360^\circ$
العلوم	200	0.219	21.97	200	910	79.120
العلوم الزراعية	350	0.384	38.46	550	710	138.46
الطب	90	0.098	9.8	640	360	35.60
الصيدلة	120	0.131	13.18	720	270	47.47
الدراسات التطبيقية	150	0.164	16.4	840	150	29.34
المجموع	910	1	100%	910	0	360

المطلوب :

مثل بيانيا هذه المعطيات باستعمال القضبان البيانية و الدائرة النسبية



الدائرة النسبية للطلبة



القضبان البيانية للطلبة حسب الاختصاص

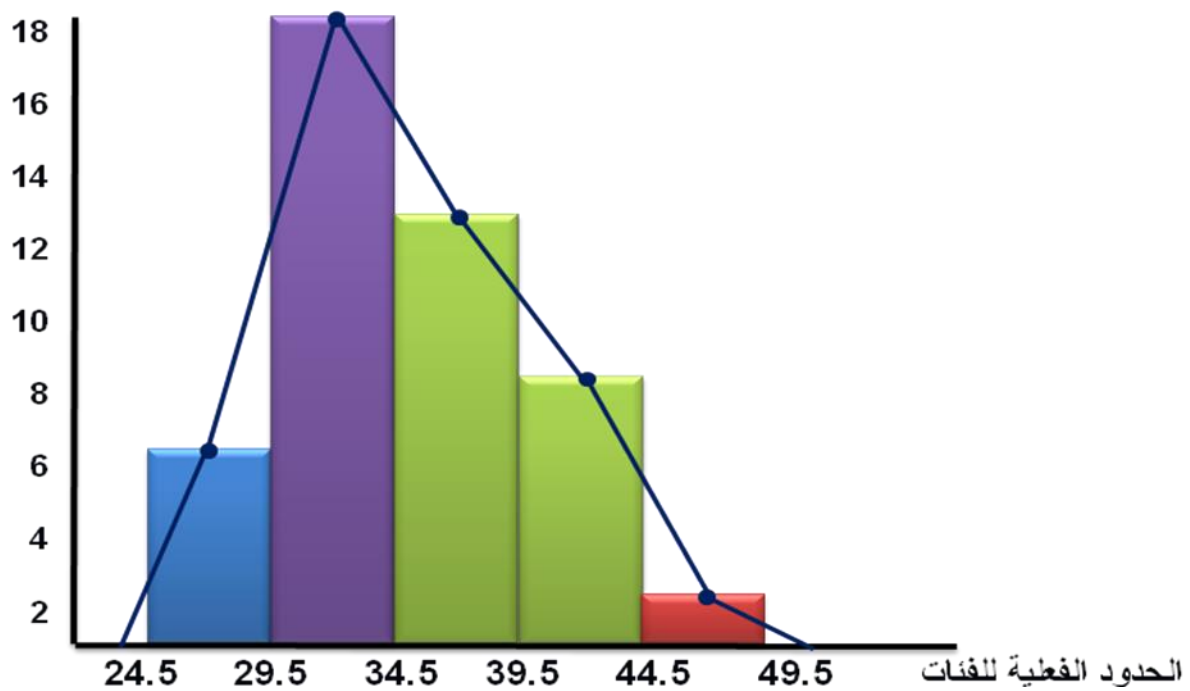
التمرين الرابع : البيانات الملخصة الجدول التالي تمثل درجات الحرارة في 50 بلد افريقي :

الفئات	مراكز الفئات X_i	التكرارات n_i	التكرار النسبي f_i	التكرار المئوي
[24.5-29.5]	$X_1=24.5+29.5/2=27$	7	$=7/50$	14
[29.5-34.5]	32	18	0.14	36
[34.5-39.5]	32	13	0.36	26
[39.5-44.5]	42	9	0.26	18
[44.5-49.5]	47	3	0.18	6
			0.06	
$\sum$ المجموع		$\sum n_i=50$	$\sum f_i=1$	$\%=100\%\sum f_i$

المطلوب :

1. احسب مراكز الفئات ، التكرارات النسبية و التكرارات المئوية
2. ارسم المدرج التكراري ، المضلع التكراري والمنحنى التكراري
3. ماهي النسبة المئوية للبلدان الذين لهم متوسط درجة حرارة تفوق 39.5؟
4. ماهو عدد البلدان الذين لهم متوسط حرارة يقل عن 39.5

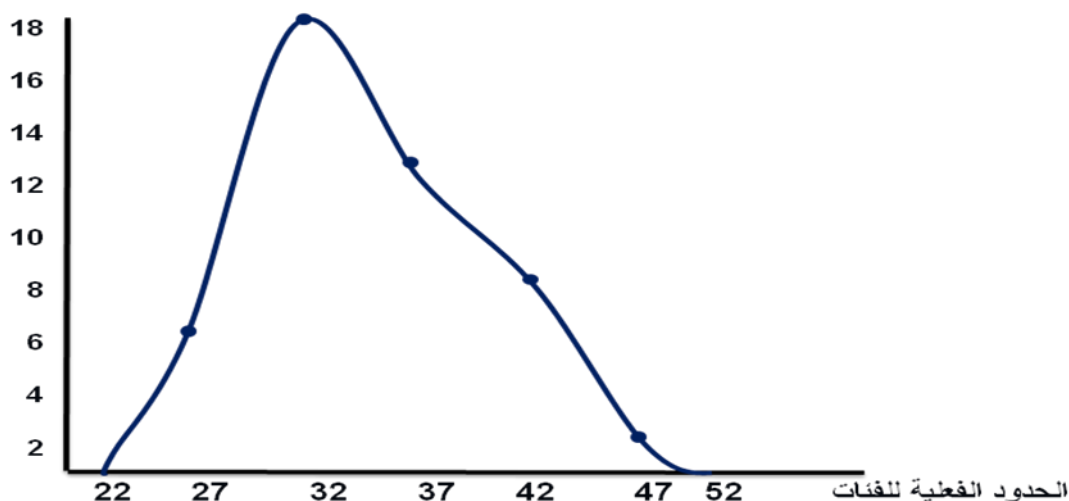
الحل



المدرج التكراري و المضلع التكراري لمتوسط درجة الحرارة

3. النسبة المئوية للبلدان الذين لهم متوسط درجة حرارة تفوق 39.5 تقرا من التكرار النسبي: $0.18 + 0.26$

4 عدد البلدان الذين لهم متوسط حرارة يقل عن 39.5 تقرا من التكرار المطلق: $7 + 18$



المنحنى التكراري لمتوسط درجة الحرارة

التمرين الخامس:

البيانات التالية تمثل 60 طالبا بالكيلوغرام في أحد أقسام الـ LMD بكلية العلوم الاقتصادية وعلوم التسيير

بجامعة وهران 2:

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

1. كون جدول التوزيع التكراري واحسب كلا من: $Nc_i^{\downarrow}$, f_i , $f_i\%$, $Fc_i^{\uparrow}\%$, $Fc_i^{\downarrow}\%$, ثم فسر

كلا من: n_2 , $Nc_2^{\uparrow}$, $Nc_5^{\downarrow}$, $Fc_2^{\uparrow}\%$, $Fc_5^{\downarrow}\%$

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

- حساب المدى: $E = \text{Max}(X_i) - \text{Min}(X_i) = 84 - 50 = 34$

- حساب عدد الفئات: $NC = 2,5\sqrt[4]{N} = 6,9$

- حساب طول الفئة:

$$ai = a_i = \frac{E+1}{NC} = \frac{35}{6,9} = 5.07 \approx 5$$

أوزان الطلبة X_i	عدد الطلبة n_i	X_i	f_i	$f_i\%$	$Nci_i^{\uparrow}$	$Nci_i^{\downarrow}$	$\%Fci_i^{\uparrow}$	$\%Fci_i^{\downarrow}$
]55 – 50]	2	52,5	0,033	3,3	2	60	3,3	100
]60 – 55]	5	57,5	0,083	8,3	7	58	11,6	96,7
]65 – 60]	12	62,5	0,2	20	19	53	31,6	88,4
]70 – 65]	16	67,5	0,268	26,8	35	41	58,4	68,4
]75 – 70]	14	72,5	0,233	23,3	49	25	81,7	41,6
]80 – 75]	8	77,5	0,133	13,3	57	11	95	18,3

5	100	3	60	5	0,05	82,5	3	[85 – 80]
/	/	/	/	100	1	/	60	المجموع $\sum n_i$

الشرح:

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

$Nc_2^{\uparrow} = 7$: هناك 7 طلبة من بين 60 طالبا أوزانهم أقل تماما من 60 كلغ.

$Nc_5^{\downarrow} = 25$: هناك 25 طالبا من بين 60 طالبا أوزانهم أكبر أو يساوي 70 كلغ.

$Fc_2^{\uparrow} \% = 11,6\%$: هناك 11,6% من الطلبة أوزانهم أقل تماما من 60 كلغ.

$Fci_5^{\downarrow} \% = 41,6\%$: هناك 41,6% من الطلبة أوزانهم أكبر أو تساوي 70 كلغ.

التمرين السادس يمثل الجدول التالي توزيع الدخل الشهري بالآلاف لمجموعة من العمال في مؤسسة إنتاج العصير المطلوب ارسم المدرج التكراري.

المدرج التكراري في حالة عدم تساوي أطوال الفئات: في هذه الحالة نلاحظ بان قواعد المستطيلات غير

متساوية و بالتالي مساحة المستطيلات لا تتناسب مع التكرارات الأصلية، وحتى تصبح مساحة المستطيلات

متناسبة مع ارتفاعها نقوم بتعديل التكرارات قبل رسم المدرج التكراري و ذلك بإتباع الخطوات التالية:

– حساب طول كل فئة (ai)

– تحديد طول الفئة الوجدوي (au) و ذلك باختياره بأحد المعايير التالية:

أكبر قاسم مشترك (PGDC) لأطوال الفئات إن وجد أو طول الفئة الأكثر شيوعا أو أصغر

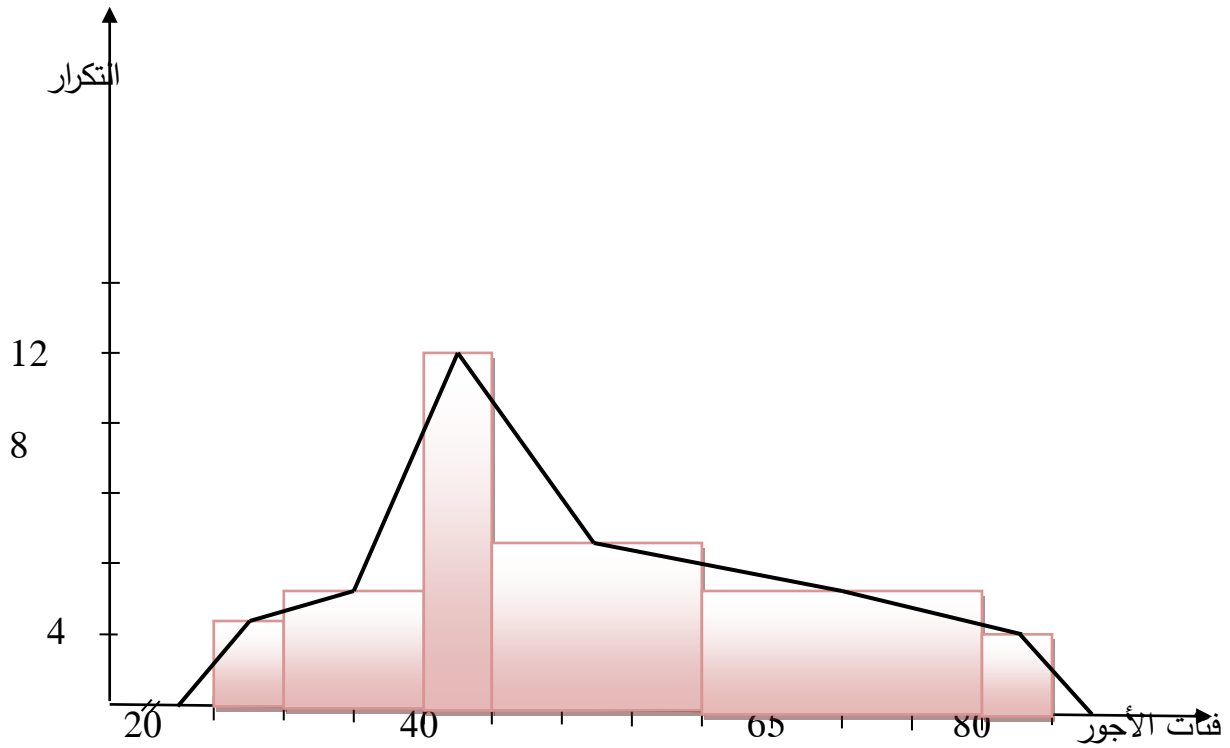
طول فئة

– حساب المعاملات (ci): $ci = \frac{ai}{au}$

– حساب التكرار المعدل (h_i) : $h_i = \frac{n_i}{c_i}$

أجور العمال ب 10^3 الفئات	عدد العمال n_i	a_i	$c_i = a_i/au$	$h_i = n_i/c_i$
[20 – 25[	30	5	$5/5 = 1$	$30/1 = 30$
[25 – 35[	26	10	$10/5 = 2$	$26/2 = 13$
[35 – 40[	10	5	$5/5 = 1$	$10/1 = 10$
[40 – 55[	21	15	$15/5 = 3$	$21/3 = 7$
[55 – 75[	8	20	$20/5 = 4$	$8/4 = 2$
[75 – 80[	5	5	$5/5 = 1$	$5/1 = 5$
المجموع Σ	100			

التمثيل البياني للمدرج التكراري و المضلع التكراري:



التمرين السابع يمثل الجدول التالي توزيع عينة من 50 فرد حسب التأهيل العلمي، والمطلوب حساب كلا من:

$F_4^{\downarrow}\%$ ، $F_2^{\uparrow}\%$ ، $N_4^{\downarrow}$ ، $N_2^{\uparrow}$ ، n_2 ، ثم فسر كلا من: $F_1^{\downarrow}\%$ ، $F_1^{\uparrow}\%$ ، $f_i\%$ ، f_i ، $N_1^{\downarrow}$ ، $N_1^{\uparrow}$ مع التمثيل

البياني لهذا التوزيع.

التأهيل العلمي	n_i	f_i	$f_i\%$	$N_{ci} \nearrow$	$N_{ci} \searrow$	$F_i^{\uparrow}$	$F_i^{\downarrow}\%$
المستوى الابتدائي	4	0,08	8	4	50	8	100
المستوى المتوسط	14	0,28	28	18	46	36	92
المستوى الثانوي	20	0,40	40	38	32	76	64
المستوى الجامعي	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.

التمرين الثامن يمثل الجدول التالي توزيع عينة من 40 فردا من الجالية المغاربية في فرنسا حسب البلد الأصلي والمطلوب :

حساب كلا من f_i ، $f_i \%$ مع شرح n_2 و n_4 و $f_2 \%$ ؟

مثل بيانيا باستعمال الدائرة النسبية.

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

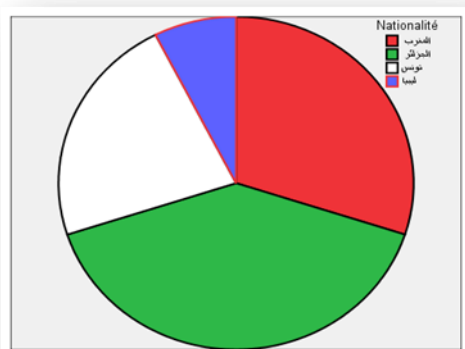
الشرح:

$n_2 = 16$: 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:



التمرين التاسع:

البيانات التالية تبين عدد الغيابات التي سجلها عمال مؤسسة ما خلال الثلاثي الأول من السنة.

9	5	4	1	6	4	3	5	7	3	2	6	2	5	3
2	3	3	4	9	5	5	4	0	0	5	1	2	5	0
1	2	2	2	1	1	1	5	3	0	2	3	2	1	4

المطلوب:

- حدد المجتمع الاحصائي والمتغير الاحصائي ونوعه؟

- لخص هذه البيانات في جدول إحصائي؟.

الحل

المجتمع الاحصائي هو 45 عامل

المتغير الإحصائي: عدد الغيابات نوعه هو كمي متقطع لأن يأخذ قيم صحيحة (معزولة)

عدد الغيابات	1	2	3	4	5	6	المجموع
عدد العمال	5	12	18	4	4	2	45

التمرين العاشر

في أحد الأحياء السكنية قمنا بجمع بيانات حول عدد الغرف المملوكة من طرف العائلات وتوصلنا إلى النتائج التالية:

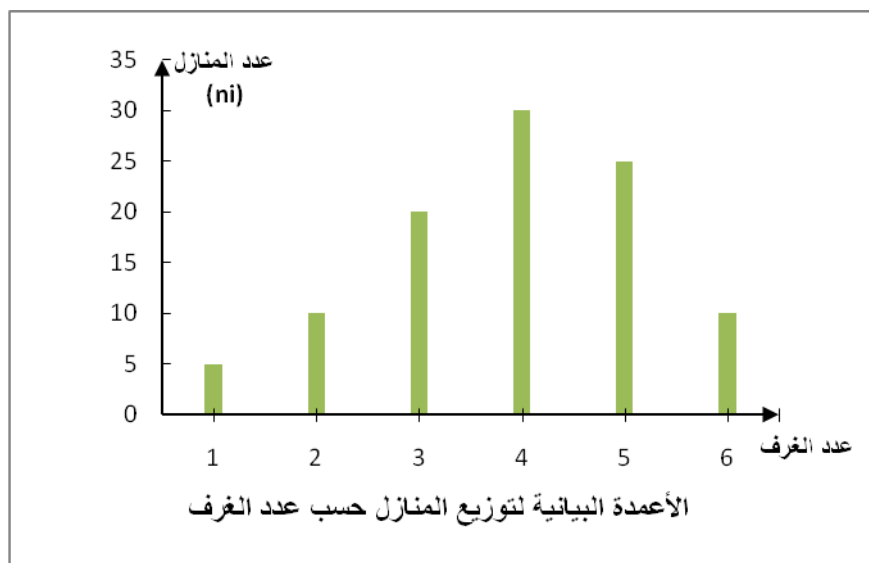
5	2	4	3	3	6	3	2	4	4
2	2	4	3	7	5	4	8	7	4
3	4	7	3	5	2	8	4	3	6
4	5	2	4	6	3	6	3	4	3
2	4	3	5	1	4	5	3	3	2

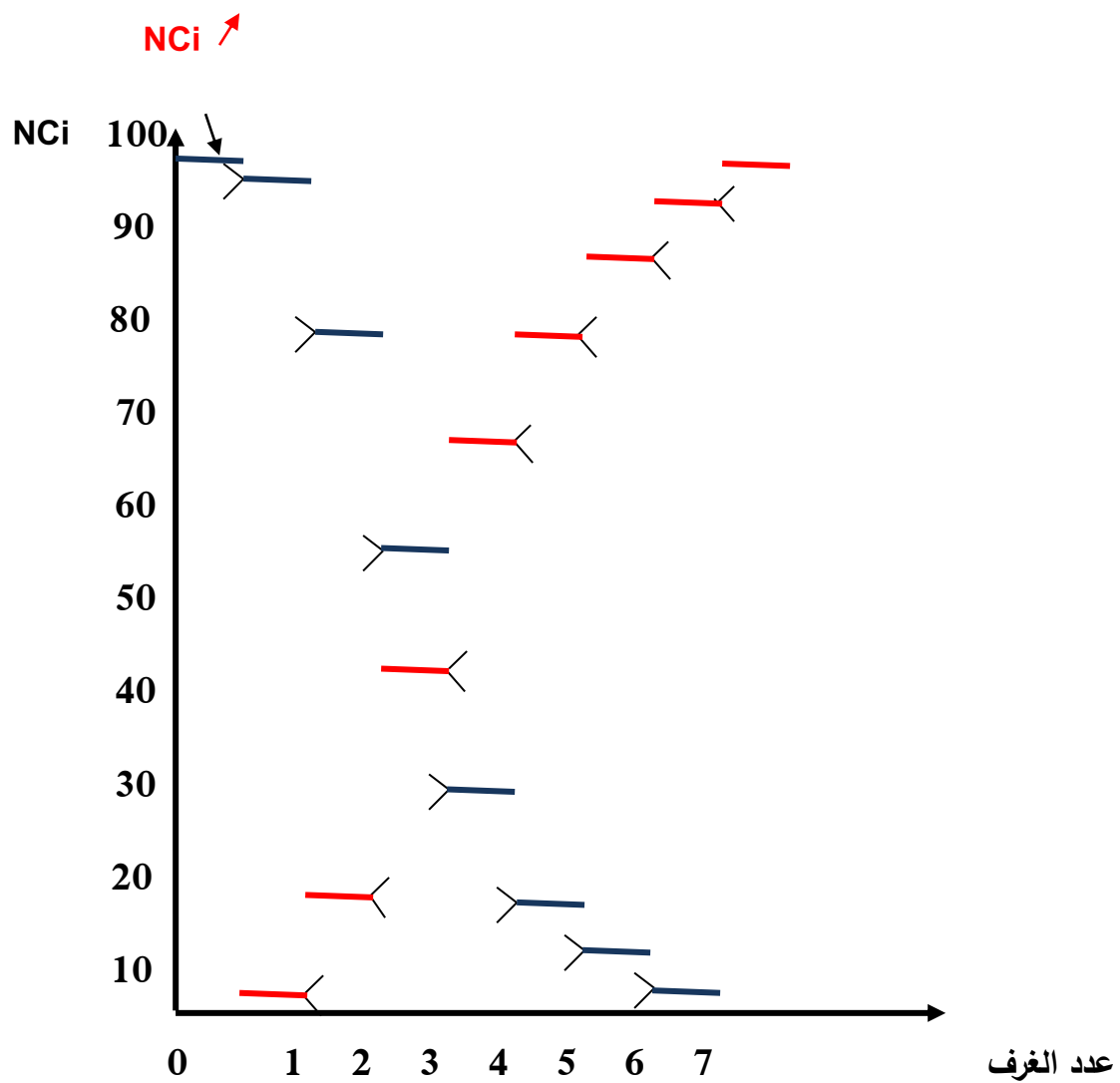
المطلوب:

- حدد طبيعة المتغير الإحصائي؟ لماذا؟.
- تبويب هذه البيانات في جدول إحصائي؟
- مثلها بيانيا
- الحل
- طبيعة المتغير الإحصائي هو كمي منفصل (متقطع) لانه يحمل قيم صحيحة.
- الجدول الاحصائي

عدد الغرف	عدد المنازل	$\sum N_i$ ↗	$\sum N_i$ ↘
1	5	5	100
2	10	15	95
3	20	35	85
4	30	65	65
5	25	90	35
6	10	100	10
$\sum$ المجموع	100	/	/

التمثيل البياني





التمثيل البياني للتكرارات التجميعية الصاعدة و النازلة للمنازل حسب عدد الغرف

التمرين 11 : البيانات التالية تمثل إنتاج 60 ورشة من الكراسي خلال يوم.

66	48	93	84	39	57	39	12	21	12
54	84	66	30	66	57	66	66	84	66
66	57	66	57	84	57	21	75	57	57
93	57	84	39	21	21	84	66	57	75
48	21	75	48	30	57	57	57	57	75
21	48	66	75	84	48	48	57	30	30

المطلوب:

- (1) ما هو المتغير المدروس؟ ما هو المجتمع الإحصائي المدروس؟ حدد الكيفيات.
- (2) قم بتكوين الجدول التكراري
- (3) احسب التكرارات النسبية
- (4) احسب التكرارات التجميعية الصاعدة ($NCi\uparrow$) و احسب التكرارات التجميعية ($FCi\downarrow$)

الحل:

1- المتغير المدروس هو عدد الطاولات طبيعته هو متغير كمي منفصل

المجتمع الإحصائي المدروس هو 60 ورشة

الكيفيات عبارة عن اعداد صحيحة من 12.....93

2- اعداد الجدول التكراري

جدول التفرغ

Xi	التفرغ	التكرار fi
12		2
21		5
30		5
39		3
48		6
57		13
66		11
75		5
84		7
93		3

جدول التوزيع التكراري

Xi	التكرار fi
12	2
21	5
30	5
39	3
48	6
57	13
66	11
75	5
84	7
93	3
المجموع	60

3- التكرارات النسبية

4- التكرارات التجميعية الصاعدة ($N_{Ci} \uparrow$) و احسب التكرارات التجميعية ($F_{Ci} \downarrow$)

Xi	التكرار fi	fi	fi	Nci	Nci
12	2	2/60			
21	5				
30	5				
39	3				
48	6				

57	13				
66	11				
75	5				
98	7				
93	3				
المجموع	60	1	100		

تمارين مقترحة

التمرين 1:

لخص كل من السلسلتين من البيانات في جدول توزيع تكراري من 6 فئات متساوية الطول؟

-1

12	19	13	07	01	15	12	30	24	29	08	07	03	23	18
						25	22	17	18	13	12	07	18	09

-2

51.5	44.7	31.6	37.5	34.8	36.8	40.1	53.4	51.8	43.8
48.0	46.2	41.5	44.0	47.8	41.5	42.3	39.9	46.0	45.3
					45.3	44.3	43.9	50.5	55.9

التمرين 2:

البيانات التالية تمثل مبيعات كشك معين من جريدة الخبر خلال 50 يوم. شكل جدول توزيع تكراري من 10 فئات مغلقة متساوية الطول؟.

36	45	31	28	41	32	29	26	48	32
30	28	33	27	40	31	30	40	35	45
32	37	36	31	35	33	38	39	41	30
38	42	35	33	36	38	39	37	45	23
36	46	30	35	32	51	43	37	22	34

التمرين 3:

البيانات التالية متعلقة بمداخيل 50 شخص في إحدى المؤسسات يوميا بالدنانير.

375	370	360	200	250	230	180	180	180	170	120	120	120
350	280	520	520	520	460	110	110	100	390	380	380	375
440	420	420	400	400	400	390	650	640	360	360	360	350
630	620	620	620	620	620	640	600	600	540	540	460	

المطلوب: شكل جدول توزيع تكراري من 7 فئات متساوية الطول؟

التمرين 4:

إذا كان القطر الداخلي لسلك كهربائي يمكن قياسه إلى أقرب وحدة من مائة من الملمتر، وكانت مراكز الفئات

للتوزيع التكراري لهذه القياسات معطاة بالملمتر كالتالي:

3.36	3.33	3.30	3.27	3.24	3.21
------	------	------	------	------	------

أوجد: طول الفئة، الحدود الحقيقية للفئات، الحدود الظاهرية للفئات؟ علما بأن فئات التوزيع متساوية الطول.

التمرين 5:

أجريت دراسة على أطوال جنود الاحتياط بثكنة ما تم التوصل من خلالها إلى النتائج التالية:

1.79	1.78	1.77	1.76	1.75	1.74	1.73	1.72	1.71	1.70	الطول
3	2	1	2	4	0	2	3	2	1	العدد
1.89	1.88	1.87	1.86	1.85	1.84	1.83	1.82	1.81	1.80	الطول
1	2	3	1	2	4	0	4	6	4	العدد

انطلاقا من هذه البيانات شكل:

- جدول توزيع تكراري من فئات متساوية طول كل منها 5 سم؟.

- جدول التكرار المتجمع الصاعد والتكرار المتجمع النازل؟.

التمرين 6:

إذا أعطيت البيانات الخاصة بسن 341 سيدة:

السن	أقل من 20 سنة	-20	-25	-30	-35	-40	-50	60	المجموع
		25	30	35	40	50	60		
العدد	55	161	66	17	16	17	7	2	341

المطلوب:

- بين في جدول ولكل فئة عمر التكرارات المعدلة والتكرارات المتجمعة الصاعدة والنازلة؟.
- أرسم المدرج التكراري؟.
- أرسم على نفس المنحنى المضلع التكراري واشتق منه منحنى التوزيع التكراري؟

التمرين 7:

الجدول التالي يبين التوزيع التكراري لمدة حياة 500 مصباح كهربائي:

مدة الحياة بالساعات	500-300	700-500	1100-700	1100 فأكثر	المجموع
عدد المصابيح	50	150	275	25	500

المطلوب:

- رسم المدرج التكراري والمضلع التكراري للتكرارات النسبية؟
- رسم المنحنى التكراري المتجمع النسبي؟
- إيجاد عدد المصابيح ذات مدة حياة أقل من 900 ساعة؟

التمرين 8:

الآتي توزيع عمال مؤسسة ما حسب التخصص:

التخصص	العدد
مهندس	50
تقني سامي	50
تقني	100
عامل متخصص	200
عامل بسيط	400
المجموع	800

المطلوب:

- حدد طبيعة المتغير الاحصائي؟.

- تمثيل هذه البيانات عن طريق القطاعات الدائرية؟.

التمرين 9:

عند إلقاء زهرة نرد 200 مرة سجل عدد المرات التي شوهد فيها كل وجه من الأوجه الستة وكانت نتائج التجربة كالتالي:

الوجه	1	2	3	4	5	6
عدد المرات	8	29	68	57	33	5

المطلوب:

عرض هذه البيانات بيانياً؟.

تمرين 10:

التمرين 9: - بهدف تقييم مدى قراءة الشباب للكتب على مدار العام حصل باحث على البيانات التالية

المطلوب

عدد الكتب (قيم المتغير) X_i	عدد الشباب (التكرار)
1	35
2	20
3	15
4	12
5	8
6	5
7	3
8	2
المجموع $\sum n_i$	100

1 - ما هو المجتمع المدروس؟ ما هي الصفة المدروسة و طبيعتها؟

2- باستعمال التكرارات التجميعية :

حدد العدد و النسبة المئوية للشباب الذي يقرأ أقل من 4 كتب

حدد العدد و النسبة المئوية للشباب الذي يقرأ أكثر من 4 كتب

حدد العدد و النسبة المئوية للشباب الذي يقرأ على الأكثر 4 كتب

حدد العدد و النسبة المئوية للشباب الذي يقرأ أكثر من 3 وأقل من

6 كتب

تمرين 9:

سجلت شركة في الأشهر الأخيرة عددا من الشكاوي اليومية ، و التي وصلت الى مصلحة العملاء . كانت النتائج كالتالي :

المجموع	5	4	3	2	1	عدد الشكاوي
100	2	6	18	35	39	عدد الأيام

1. احسب التكرارات النسبية و التكرارات التجميعية المناسبة (الصاعدة و النازلة)
2. ارسم الأعمدة البيانية و مثل بيانيا التكرارات التجميعية الصاعدة و النازلة لعدد الشكاوي
3. ماهي نسبة الأيام التي سجل فيها على الأقل شكاوتين ؟

تمرين 10:

اليك الجدول التالي الذي يمثل متوسط العمر المتوقع في 50 بلد إفريقي

	[64-68[	[68-70[	[70-72[	[72-74[	[74-76[	[76-78[	[78-80[
ni		8		8		15	
fi	0.1		0.12		0.08		

1. انه الجدول
2. ارسم المدرج التكراري ، المضلع التكراري ، منحني التكرارات التجميعية النازلة و منحني التكرارات التجميعية الصاعدة
3. ماهي النسبة المئوية للبلدان الذين لهم متوسط عمر متوقع يفوق 69 سنة ؟
4. ماهو عدد البلدان الذين لهم متوسط عمر متوقع أكبر من الذي هو متوقع في الجزائر (75 سنة)؟

Measures of Central Tendency

الفصل الثالث: مقاييس النزعة المركزية

سنتطرق من خلال هذا الفصل إلى العناصر التالية:

أولاً- التوزيع التكراري للمتغير الإحصائي المنفصل (المتقطع).

ثانياً- التوزيع التكراري للمتغير الإحصائي المتصل (المستمر).

ثالثاً- التوزيع التكراري للمتغير الإحصائي الكيفي.

الفصل الثالث مقاييس النزعة المركزية

بعد جمع البيانات الإحصائية حول الظاهرة المدروسة - في شكل توزيعات تكرارية ومنحنيات تكرارية ينتقل الإحصائي إلى دارستها وتحليلها و إستخلاص النتائج، ولأجل ذلك يكون من بين أولى إهتماماته بحث مدى تركز القيم التي جمعها حول قيمة ما ضمن مجموعة البيانات الإحصائية، وبمعنى آخر بحث مدى نزوع مختلف قيم الظاهرة حول قيمة مركزية منها، ويتم ذلك عن طريق أدوات تحليلية سميت لأجل ذلك بمقاييس النزعة المركزية وهي⁹:

1- المتوسط الحسابي Arithmetic Mean ومشتقاته) المتوسط الهندسي، المتوسط التوافقي، المتوسط التربيعي الذي يعد من أهم المقاييس الإحصائية لكونه يستخدم جميع البيانات الإحصائية.

المتوسط الحسابي **Arithmetic Mean** بشكل عام الوسط الحسابي هو مجموع القيم مقسوما على عددها

و يحسب بالطريقة التالية

$$\bar{x} = \frac{\sum x_i}{N} \quad \blacklozenge \quad \text{في حالة البيانات غير المبوبة}$$

$$\bar{x} = \frac{\sum n_i x_i}{\sum n_i} \quad \blacklozenge \quad \text{في حالة البيانات المبوبة باستخدام التكرارات المطلقة}$$

$$\bar{x} = \sum f_i x_i \quad \blacklozenge \quad \text{في حالة البيانات المبوبة باستخدام التكرارات النسبية}$$

$$\bar{x} = x_0 + a_i \bar{z} \quad \blacklozenge \quad \text{في حالة بطريقة تبديل المتغير (الانحرافات المختزلة):}$$

$$\bar{X} = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2 + n_3 \bar{x}_3 + \dots + n_1 \bar{x}_1}{n_1 + n_2 + n_3 + \dots + n_4} \quad \blacklozenge \quad \text{في حالة مجموعتين من البيانات}$$

خواص المتوسط الحسابي:

1 - يعتبر المتوسط الحسابي أبسط مقاييس المركزية حسابا وأكثرها استخداما.

⁹ بنية صابرينة محاضرات في الاحصاء الوصفي جامعة تيارت 2017/2018 ص 63

2 - يأخذ المتوسط الحسابي بعين الاعتبار جميع قيم الظاهرة المدروسة.

3 - مجموع انحرافات القيم عن متوسطها الحسابي يساوي صفر $= \sum_{i=1}^n (x_i - \bar{X}) = 0$

وللتأكد من ذلك نقوم بتفكيك الطرف الأيمن:

$$\begin{aligned} & \sum_{i=1}^n (x_i - \bar{X}) \\ &= \sum_{i=1}^n x_i - \sum_{i=1}^n \bar{X} \\ &= \sum_{i=1}^n x_i - n\bar{X} \end{aligned}$$

$$(لأن : \bar{X} + \bar{X} + \bar{X} + \dots + \bar{X} = n\bar{X})$$

نضرب الحد الأول في n ونقسمه على n

$$\begin{aligned} &= \frac{n \sum_{i=1}^n x_i}{n} - n\bar{X} \\ &= n\bar{X} - n\bar{X} = 0 \end{aligned}$$

4 - يتأثر المتوسط الحسابي بالقيم المتطرفة أو الشاذة.

مزايا وعيوب الوسط الحسابي

يتميز الوسط الحسابي بالمزايا التالية :

أنه سهل الحساب .

يأخذ في الاعتبار كل القيم .

أنه أكثر المقاييس استخداما وفهما . ومن عيوبه .

أنه يتأثر بالقيم الشاذة والمتطرفة .

يصعب حسابه في حالة البيانات الوصف

قد يكون خادعا مالم يصاحبه مقياس للتشتت

المتوسط الهندسي Geometric Mean : إذا كانت لدينا القيم $x_1, x_2, x_3, x_4, \dots, x_n$ فإن

متوسطها الهندسي يساوي الجذر النوني لحاصل ضرب هذه القيم في بعضها البعض.

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

وإذا كانت هذه القيم متكررة فإن:

$$G = \sqrt[N]{x_1^{n_1} \cdot x_2^{n_2} \cdot x_3^{n_3} \cdot \dots \cdot x_n^{n_n}}$$

حيث $\sum n_i = N$ (مجموع التكرارات)

نفس الشيء إذا كانت هذه القيم مبوبة في جداول توزيع تكراري فإن:

$$G = \sqrt[N]{x_1^{n_1} \cdot x_2^{n_2} \cdot x_3^{n_3} \cdot \dots \cdot x_n^{n_n}}$$

حيث x_i = مراكز الفئات.

$\sum n_i = N$ (مجموع التكرارات).

في حالة كون البيانات كبيرة فإنه يفضل استخدام طريقة اللوغاريتم ويكون ذلك كالتالي:

$$G = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdot x_4 \cdot \dots \cdot x_n} \quad \text{في حالة بيانات مفردة:}$$

$$\text{Log} G = \frac{1}{N} [\log x_1 + \log x_2 + \log x_3 + \dots + \log x_n]$$

$$\log G = \frac{1}{N} \sum_{i=1}^n \log x_i$$

وفي حالة البيانات المتكررة أو المبوبة في جداول توزيع تكرار فإن:

$$G = \sqrt[N]{x_1^{n_1} \cdot x_2^{n_2} \cdot x_3^{n_3} \cdot \dots \cdot x_n^{n_n}}$$

$$\text{Log} G = \frac{1}{N} [n_1 \log x_1 + n_2 \log x_2 + n_3 \log x_3 + \dots + n_n \log x_n]$$

$$\log G = \frac{1}{N} \sum_{i=1}^n n_i \log x_i$$

حيث: x_i تمثل القيمة أو مركز الفئة.

n_i تمثل التكرار.

خواص المتوسط الهندسي:

من أهم خواص المتوسط الهندسي ما يلي:

1- يدخل في حساب جميع القيم ولكنه أقل تأثر بالقيم المتطرفة من المتوسط الحسابي.

2- لا يمكن حساب من الجداول التكرارية المفتوحة من البداية أو النهاية.

3- لا يمكن حسابه في حالة وجود قيمة سالبة أو معدومة.

4- يستخدم بأكثر واقعية عند وصف الظواهر النسبية.

قيمة المتوسط الهندسي لأي ظاهرة أصغر دائما من قيمة المتوسط الحسابي $G < \bar{X}$

المتوسط التوافقي: *Harmonic Mean*

المتوسط التوافقي هو من المقاييس الخاصة التي تستخدم لتحديد معدلات السرعة ومتوسط الأسعار،

ومتوسط الكثافة السكانية. ويحسب بالطريقة التالية ;

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

في حالة بيانات غير مبوبة

$$H = \frac{\sum n_i}{\sum \frac{n_i}{x_i}}$$

في حالة بيانات مبوبة

-خواص المتوسط التوافقي:

1 - يأخذ بعين الاعتبار جميع القيم وتأثره بالقيم الشاذة أقل من تأثر المتوسط الحسابي.

2 - لا يمكن حسابه في حالة وجود بيانات معدومة.

3 - يعطي نتائج أكثر واقعية في حالة حساب متوسطات الأسعار والسرعة.

4 - قيمته دائما أقل من قيمة المتوسط الهندسي.

المتوسط التريبيعي : المتوسط التريبيعي لمجموعة من البيانات على انه الجذر التريبيعي لمتوسط مربعات البيانات.

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

◆ في حالة بيانات غير مبوبة

$$= \sqrt{\frac{\sum n_i x_i^2}{\sum n_i}}$$

◆ في حالة بيانات مبوبة

◆ العلاقة بين المتوسطات الاربعة

$$X \text{ Min} < H < G < \bar{x} < Q < X \text{ Max} \quad \blacklozenge$$

ملاحظة : لا يوجد تمثيل بائي للمتوسطات الاربعة

المنوال Mode : المنوال هو القيمة الأكثر شيوعا أو تكرار في مجموعة القيم؛ والمنوال قد يكون وحيد القيمة كما

قد يكون هناك أكثر من منوال لنفس التوزيع، وسنرمز له بالرمز M_o . ويحسب بالطريقة التالية

القيمة الأكثر تكرارا في السلسلة في حالة بيانات غير مبوبة

$$M_o = L_{Min} + \left[\frac{\Delta_1}{\Delta_1 + \Delta_2} \right] a_{M_o}$$

في حالة البيانات المبوبة

بيانيا نحدد المنوال الآن بالرسم وذلك من خلال رسم المستطيلات التي تمثل كل من الفئة المنوالية والفئتين التي قبلها

والتي بعدها في المدرج التكرار.

مزايا وعيوب المنوال : يتميز المنوال بالمزايا التالية :

سهولة حسابه و إيجاد

لا يتأثر بالقيم الشاذة.

يمكن حسابه في الجداول التكرارية المفتوحة.

يمكن إيجاد جميع أنواع البيانات الوصفية و الكمية.

يمكن إيجاد من خلال الرسم

يعتبر المقياس الوحيد الذي يمكن استخدامه للبيانات الاسمية. ومن بين عيوبه

لا تدخل في حسابه جميع البيانات.

قد لا يقع في مركز البيانات بل في طرفها.

تتغير قيمته باختلاف طريقة اختيار الفئات.

يصعب التعامل معه في الإحصاء الاستدلالي لأنه قد تكون له أكثر من قيمة

- الوسيط¹⁰ Median ومشتقاته الربيعات، العشيريات، والمئويات: الذي يعتبر من المقاييس المهمة

والمستخدم بشكل واسع جدا وخاصة عندما يكون قسم من البيانات كبيرة جدا أو صغيرة جدا، أو ما تسمى في

الإحصاء بالقيم الشاذة.

الوسيط Median: القيمة التي تقسم مجموع البيانات إلى قسمين بحيث يكون نصف عدد البيانات أكبر منه

ونصف عدد البيانات أصغر منه ويرمز له بالرمز: Me.

إذا كانت البيانات عبارة عن سلسلة، وهناك حالتين:

1- إذا كان عدد البيانات فردي فإن الوسيط هو القيمة التي ترتيبها $\frac{n+1}{2}$

2- إذا كان عدد البيانات زوجي فإن الوسيط هو متوسط القيمتين اللتين ترتيبهما $\frac{n}{2}$ و $\frac{n}{2} + 1$.

و إذا كانت البيانات لها تكرارات فإن الوسيط يوجد بإتباع الخطوات التالية:

- نحسب التكرار المتجمع الصاعد لقيم الظاهرة.

- نحدد ترتيب الوسيط $\frac{N}{2}$ (حيث $N = \text{مجموع التكرارات}$).

¹⁰ ساعد بن فرحات، عبد الحميد قطوش، الإحصاء: 1 مدعمة بتمارين وامتحانات محلولة جامعة سطيف، 2013/2014. ص 53

- نبحث عن القيمة التي تكرارها المتجمع الصاعد أكبر من $\frac{N}{2}$ مباشرة وهي القيمة التي تمثل الوسيط.

ليانات مبوبة في جداول توزيع تكراري:

لتحديد الوسيط لبيانات مبوبة في جداول توزيع تكراري فإننا نقوم بالتالي:

- (1) نحسب التكرار المتجمع الصاعد أو النازل.
- (2) نحدد ترتيب الوسيط وهو عبارة عن نصف مجموع التكرارات $\frac{N}{2}$.
- (3) نحدد الفئة الوسيطة أي الفئة التي يقع فيها الوسيط، وهي الفئة التي تقابل التكرار المتجمع الصاعد الذي يساوي ترتيب الوسيط أو أكبر منه مباشرة.
- (4) نحدد ونحسب الوسيط بتطبيق العلاقة الإحصائية للوسيط

$$M_s = L_{Min} + \left[\frac{RM_s - s}{s' - s} \right] a_{M_s}$$

بيانيا : نقطة تقاطع منحنى التكرار المتجمع الصاعد والنازل لبيانات التوزيع.

خواص الوسيط:

يتصف الوسيط بعدة خصائص أهمها:

- 1 - لا يتأثر بالقيم المتطرفة وبالتالي فإنه يعتبر أفضل المقاييس عن وجود مثل هذه القيم.
- 2 - يمكن إيجاد الوسيط من الرسم.
- 3 - يمكن استخدامه في حالة البيانات النوعية.
- 4 - يمكن حسابه من الجداول التكرارية المفتوحة

تمارين الفصل الثالث محلولة

التمرين الاول : في دراسة إحصائية حول مادة الحليب بالمزارع الموجودة على مستوى ولاية وهران توصلنا

إلى إعداد الجدول التالي:

360-320	320-280	280-240	240-200	الإنتاج باللترات
4	8	6	5	عدد المزارع

المطلوب: إيجاد متوسط إنتاج الحليب بهذه المزارع

1. احسب الوسط الحسابي بعلاقة التعريف باستخدام التكرارات المطلقة.
2. احسب الوسط الحسابي بعلاقة التعريف باستخدام التكرارات النسبية .
3. احسب الوسط الحسابي باستعمال طريقة تبديل المتغير (الانحرافات المختزلة) .

الحل:

1. الوسط الحسابي بعلاقة التعريف باستخدام التكرارات المطلقة.

الإنتاج باللترات	عدد المزارع n_i	مراكز الفئات x_i	$n_i x_i$
240-200	5	220	1100
280-240	6	260	1560
320-280	8	300	2400
360-320	4	340	1360
400-360	2	380	760
المجموع	25		7180

$$\bar{X} = \frac{\sum n_i x_i}{\sum n_i} = \frac{7180}{25} = 287.2$$

أي أن متوسط إنتاج الحليب بهذه المزارع هو أكثر بقليل من 287 لتر للمزرعة.

2. الوسط الحسابي بعلاقة التعريف باستخدام التكرارات النسبية .

. حساب الوسط الحسابي باستخدام التكرارات النسبية ($f_i = \frac{n_i}{\sum n_i}$) :

$$= \frac{n_i}{\sum n_i} \rightarrow n_i = f_i \sum n_i \dots\dots\dots (1) f_i$$

$$= \frac{\sum n_i x_i}{\sum n_i} \dots\dots\dots (2) \bar{x}$$

نعوض (1) في (2):

$$\bar{x} = \frac{\sum (f_i \sum n_i x_i)}{\sum n_i} \rightarrow \bar{x} = \frac{\sum n_i \sum f_i x_i}{\sum n_i} \rightarrow \bar{x} = \sum f_i x_i \rightarrow \bar{x} = 287.2$$

نختار وسط فرضي $X_0 = 300$ وهو مركز الفئة الوسطى (القيمة التي تقابل أكبر تكرار). ونتبع الخطوات

التالية:

- توجد قيم جديدة $z_i = \frac{x_i - x_0}{a_i}$ والتي تساوي مراكز الفئات ناقص الوسط الفرض.

- نضرب هذه القيم الجديدة في تكرار الفئات n_i .

- نحسب $\bar{Y}$ وهو المتوسط الحسابي لهذه القيم الجديدة.

- نحسب المتوسط الحسابي للقيم الجديدة والذي يساوي $\bar{x} = x_0 + a_i \bar{z}$

الفئات	n_i	x_i	$n_i x_i$	f_i	$f_i x_i$	a_i	$z_i = \frac{x_i - x_0}{a_i}$	$n_i z_i$
240-200	5	220	1100	0.2	44	40	-80	-400
280-240	6	260	1560	0.24	62.4	40	-40	-240
320-280	8	300	2400	0.32	96	40	0	0
360-320	4	340	1360	0.16	54.4	40	40	160
400-360	2	380	760	0.08	30.4	40	80	160
	25		7180	1	287.2			-320

$$\bar{Z} = \frac{\sum_{i=1}^n n_i Z_i}{\sum_{i=1}^n n_i} = \frac{-320}{25} = -12.8$$

$$\bar{x} = 300 - 12.8 = 287.2 \text{ لتر}$$

التمرين الثاني: الجدول التالي يقدم بيانات حول دراسة تمت لسعر منتج البطاطا عبر تجزئة 500 سوق ناحية الغرب.

المطلوب :

1. احسب متوسط سعر البطاطا

2. قارن بين المتوسطات الاربعة

سعر البطاطا	]10-0]	]20-10]	]30-20]	]40-30]	]50-40]
عدد الاسواق	56	125	175	114	30

الحل:

1. حساب المتوسطات :

الفئات	n_i	x_i	$n_i x_i$	$\log x_i$	$n_i \cdot \log x_i$	$\frac{n_i}{x_i}$	$n_i \cdot x_i^2$
[40-44[	56	42	2352	1,6232	90,8992	1,333	98784
[44-48[	125	46	5750	1,6627	207,8375	2,717	264500
[48-52[	175	50	8750	1,6989	297,3075	3,5	437500
[52-56[	114	54	6156	1,7323	197,4822	2,111	332424
[56-60[	30	58	1740	1,7634	52,902	0,517	100920
المجموع	500	/	24748	/	846,4284	10,178	1234128

الحل:

أ. المتوسط الحسابي:

$$\bar{x} = \frac{\sum x_i n_i}{\sum n_i} = \frac{24748}{500} = 49,496 \Rightarrow \boxed{\bar{x} = 49,496}$$

ب. المتوسط الهندسي:

$$\log G = \frac{\sum (n_i \log x_i)}{\sum n_i} = \frac{846,4284}{500} = 1,6928$$

$$G = 10^{1,6928} = 49,294 \Rightarrow \boxed{G = 49,294}$$

ت. المتوسط التريبيعي:

$$Q = \sqrt{\frac{1}{\sum n_i} \sum n_i x_i^2} = \sqrt{\frac{1234128}{500}} = \sqrt{2468,258} \Rightarrow \boxed{Q = 49,68}$$

ث. المتوسط التوافقي:

$$H = \frac{\sum n_i}{\sum \frac{n_i}{x_i}} = \frac{500}{10,178} = 49,125 \Rightarrow \boxed{H = 49,125}$$

العلاقة الموجودة بين المتوسطات:

$$X_{\min} < H < G < \bar{X} < Q < x_{\max}$$

$$40 < 49,125 < 49,294 < 49,496 < 49,68 < 60 \Rightarrow \text{محققة العلاقة}$$

2. تحديد المتوسط الحسابي بياتيا :

- لا يمكن تحديد المتوسط الحسابي بياتيا (من خواص $\bar{X}$).

التمرين الثالث :

الجدول التالي يبين عدد العمال ومتوسط الأجر للعامل الواحد في الوحدات المختلفة التي تشكل الشركة

الوطنية لإنتاج الأنابيب البلاستيكية. المطلوب حساب متوسط الأجور التي توزعها هذه الشركة؟.

الفرع	وحدة الشمال	وحدة الشرق	وحدة الجنوب
عدد العمال	130	110	80
متوسط الأجور	13000	14500	18500

الحل:

إذا اعتبرنا أن متوسط الأجر في شركة هو عبارة عن مجموع متوسط الأجر في الوحدات الثلاث مقسوماً على ثلاثة فإن الإجابة تكون خاطئة فالإجابة الصحيحة هي تلك التي يمكن الحصول عليها من خلال علاقة المتوسط الحسابي المرجح.

$$\bar{X} = \frac{n_1 \bar{x}_1 + n_2 \bar{x}_2 + n_3 \bar{x}_3 + \dots + n_k \bar{x}_k}{n_1 + n_2 + n_3 + \dots + n_k}$$

$$\bar{X} = \frac{(130 \times 13000) + (110 \times 14500) + (80 \times 18500)}{130 + 110 + 80} = \frac{1690000 + 1595000 + 1480000}{320}$$

$$\bar{X} = \frac{4765000}{320} = 14890,62 =$$

متوسط أجر عمال الشركة

التمرين الرابع : الجدول التالي يبين عدد الصالونات ومتوسط المتر الواحد للجلد في الورشات المختلفة التي تستعملها الشركة الوطنية لإنتاج الصالونات والديكور. المطلوب حساب متوسط سعر الجلد الذي تستعمله هذه الشركة؟.

(1) تحديد قيمة المنوال؟

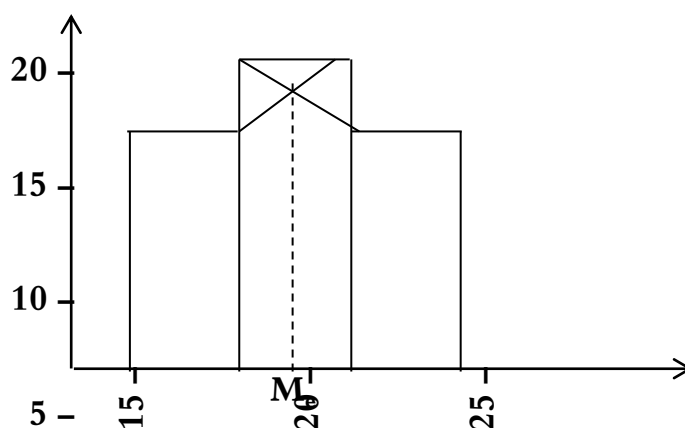
(2) قارن بين قيمة المنوال والمتوسط الحسابي والوسيط؟.

سعر الجلد	15-10	20-15	25-20	30-25	35-30	40-35	المجموع
عدد الصالونات	10	15	20	15	10	5	75

الحل:

سعر بالآلاف	عدد الصالونات	X_i	$n_i X_i$	NCI
15-10	10	12.5	125	10
20-15	15	17.5	262.5	25
25-20	20	22.5	450	45
30-25	15	27.5	412.5	60
35-30	10	32.5	325	70
40-35	5	37.5	187.5	75
المجموع	75		1762.5	

1) الفئة المنوالية هي الفئة [20-25] وبالتالي فإنه يمكن اعتبار مركز هذه الفئة (22.5) كمنوال تقريبي.
نحدد المنوال الآن بالرسم وذلك من خلال رسم المستطيلات التي تمثل كل من الفئة المنوالية والفئتين التي قبلها والتي بعدها.



- نحسب المنوال من خلال العلاقة المتوصل إليها.

أولاً يجب تحديد الفئة المنوالية و هي الفئة التي تقابل أكبر تكرار. أكبر تكرار في الجدول هو 20 و بالتالي الفئة المنوالية [20-25] :

$$\begin{aligned}
 M_O &= L_{Min} + \left[\frac{\Delta_1}{\Delta_1 + \Delta_2} \right] a_{M_O} \\
 &= 20 + \left[\frac{(20 - 15)}{(20 - 15) + (20 - 15)} \right] 5 \\
 &= 20 + \left[\frac{(5)}{(5) + (5)} \right] 5
 \end{aligned}$$

$$= 22.5 \times 10^3 \text{ DA}$$

التفسير: المنوال يساوي 22500 د.ج، مما يعني أن سعر الجلد المستعمل بكثرة هو

22500 د.ج

(2) لمقارنة بين $\bar{X}$, M_e , M_1 نحسب كل من المتوسط الحسابي والوسيط

$$\bar{X} = \frac{\sum n_i x_i}{\sum n_i} = \frac{1762,5}{75} = 23,5$$

n_i	الحدود العليا	Nci ↑
10	اقل من 10	10
15	اقل من 15	25
20	اقل من 20	45
15	اقل من 25	55
10	اقل من 30	65
5	اقل من 35	75
75	/	/

L_{min} : الحد الأدنى للفئة الوسيطة

S : التكرار المتجمع السابق لرتبة الوسيط

S' : التكرار المتجمع اللاحق لرتبة الوسيط

A_{me} : طول الفئة الوسيطة

S
 $R_{me} =$
 S'

نحصر رتبة الوسيط في عمود $Nci \uparrow \leftrightarrow$ الرتبة 50 محصورة بين 45 و 55 إذن $[20-25]$ هي الفئة الوسيطة.

$$RM_s = \frac{\sum n_i}{2} = \frac{75}{2} = 37.5 \quad \text{رتبة الوسيط}$$

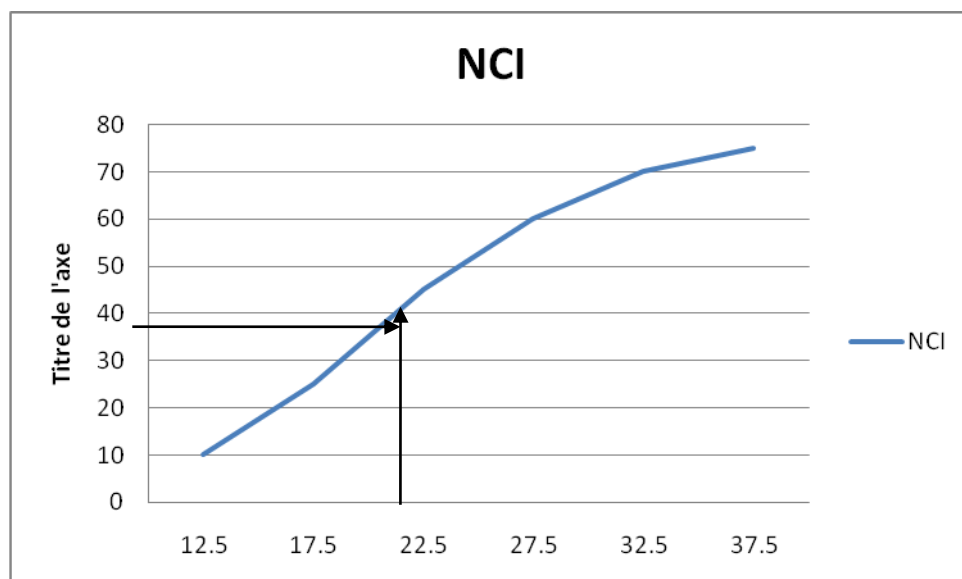
$$M_s = L_{Min} + \left[\frac{RM_s - S}{S' - S} \right] a_{M_s}$$

$$= 20 + \left[\frac{10-45}{55-45} \right] 5$$

$$= 32.125 \text{ da}$$

التفسير: 50 % من الصاونات تستعمل فيها سعر الجلد أقل من 32125 دج و 50% من الصاونات

المتبقين يستعمل فيهم جلد سعره يفوق 32125 دج



التمثيل البياني للوسيط

نلاحظ من النتائج السابق أن $\bar{X} > M_e > M_0$

التمرين الخامس :

أوجد المنوال في الحالات التالية

8 9 9 14 8 8 10 7 6 9 7 8 10 14 11 8 14 11
75 110 731 1031 1162 1977 103

الحل :

1. لتحديد قيمة الوسيط يجب علينا اتباع الخطوات التالية :

- ترتيب قيم السلسلة ترتيباً تصاعدياً أو تنازلياً.

- تحديد رتبة الوسيط و ذلك حسب عدد قيم السلسلة :

* إذا كان عدد قيم السلسلة فردي فإن للسلسلة قيمة وسيطية واحدة تقابل الرتبة الوحيدة :

$$\rightarrow RM_e = \frac{n+1}{2} \text{ تقابل الرتبة الوحيدة}$$

* إذا كان عدد قيم السلسلة زوجي فإن القيمة الوسيطة هي متوسط القيمتان اللتان تقابلان رتبنا الوسيط :

$$\text{الرتبة الاولى} \leftarrow RM_{e_1} = \frac{n}{2} \rightarrow \text{قيمة الوسيط الاولى تقابل}$$

$$\text{الرتبة الثانية} \leftarrow RM_{e_2} = \frac{n}{2} + 1 \rightarrow \text{قيمة الوسيط الاولى تقابل}$$

قيمة الوسيط هي متوسط القيمتان : $M_e = \frac{e_1 + M_{e2}}{2}$

2. المنوال هو القيمة الأكثر تكرارا أو شيوعا.

المنوال = 8

لا يوجد منوال

التمرين السادس الجدول التالي يبين عدد الأيام التي اشتغل فيها العمال في الشهر

Xi	10	15	18	20	24
ni	10	12	8	6	4

1. ما هي المتغيرة المدروسة ماهي طبيعتها

2. احسب المتوسط الحسابي

3. احسب المنوال والوسيط

4. مثل بيانيا باستعمال الدائرة النسبية هذه المعطيات

الحل

1. المتغيرة المدروسة: عدد الأيام التي اشتغل فيها العمال في الشهر.

طبيعتها: كمية متقطعة.

xi	ni	nixi	Nci ↗
10	10	100	10
15	12	180	22
18	8	144	30
20	6	120	36
24	4	96	40
المجموع	40	640	

2. حساب متوسط أيام اشتغال العمال :

$$\bar{x} = \frac{\sum x_i n_i}{\sum n_i} = \frac{640}{40} = 16 \Rightarrow \boxed{\bar{x} = 16}$$

اشتغل عمال ورشة البناء شهر فيفري 2018 في المتوسط 16 يوما.

3. حساب الوسيط و المنوال :

حساب الوسيط : بما أن المتغيرة المدروسة كمية متقطعة تتبع الخطوات التالية :

$$- \text{ نحدد رتبة الوسيط : } \frac{\sum ni}{2} = RMe = \frac{40}{2} = 20$$

- نلاحظ أن القيمة 20 رتبة الوسيط ($RMe = 20$) في التكرارات المتجمعة الصاعدة تقابل

القيمة $xi = 15$ و منه فإن قيمة الوسيط هي : $Me = 15$.

حساب المنوال : بما أن أكبر تكرار هو 12 فإن قيمة المنوال هي القيمة المقابلة له، وعليه نجد أن :

$$Mo = 15$$

4. غير ممكن استخدام العرض الدائري لتمثيل متغيرة كمية متقطعة.

التمرين السابع

يبين الجدول التالي توزيع 40 عامل حسب أطوالهم بالسنتيمتر لإحدى المؤسسات الخدمانية :

أطوال الطلبة	[150 – 140]	[160 – 150]	[170–160]	[180–170]	[190 – 180]
عدد الطلاب	4	8	16	8	4

المطلوب:

1- أحسب الربيع الأول، الربيع الثاني و الربيع الثالث ثم إستنتج القيمة الوسيطة لأطوال الطلبة مع تفسير النتائج.

2- أحسب قيمة العشير الأول، الخامس ثم التاسع مع تفسير النتائج.

3- أحسب قيمة المئين الأول، الخمسون ثم الأخير مع تفسير النتائج.

الحل :

ملاحظة: تحسب الربيعات, العشيرات والمئينيات بنفس خطوات الوسيط لهذا تسمى بأشباه الوسيط..

الفئات	n_i	الحدود العليا	$N_{ci} \uparrow$
[140-150[	4	اقل من 150	4
[150-160[	8	اقل من 160	12
[160-170[	16	اقل من 170	28
[170-180[	8	اقل من 180	36
[180-190[	4	اقل من 190	40
المجموع	40	/	/

1. الربيعات :

الربيع الأول Q_1 :

فئة [160-1500

$$[1 = \frac{\sum n_i}{4} = \frac{40}{4} = RQ_1$$

الربيع الاول

$$Q_1 = L_{Min} + \left[\frac{RQ_1 - s}{s' - s} \right] a_{Q_1}$$

$$= 20 + \left[\frac{10 - 45}{55 - 45} \right] 5$$

157.5 سم

=

التفسير: 25% من الطلبة (أي 10 طلاب) تقل اطولهم عن 157.5 سم اما 75% من الطلبة المتبقين (أي 30 طالب)

فتفوق اطولهم 157.5 سم .

الربيع الثاني Q2:

فئة [170-1600]

$$RQ_2 = \frac{2 \sum n_i}{4} = \frac{\sum n_i}{2} = \frac{40}{2} = 20$$

الربيع الثاني

$$Q_2 = L_{Min} + \left[\frac{RQ_2 - s}{s' - s} \right] a_{Q_2}$$

$$= 160 + \left[\frac{20 - 12}{28 - 12} \right] 10$$

$$= 165 \text{ سم}$$

$$Me = Q_2 = 165 \text{ سم}$$

التفسير: 50 % من الطلبة (أي 20 طالب) تقل أطوالهم عن 165 سم أما 50 % من الطلبة المتبقين (أي 20 طالب) فتتفوق أطوالهم 165 سم.

الربيع الثالث :

-1700

$$RQ_3 = \frac{3 \sum n_i}{4} = \frac{3 \times 40}{4} = 30$$

180 فئة الربيع الثالث

$$Q_3 = L_{Min} + \left[\frac{RQ_3 - s}{s' - s} \right] a_{Q_3}$$

$$= 170 + \left[\frac{30 - 28}{36 - 28} \right] 10$$

$$= 172.5 \text{ سم}$$

التفسير: 75 % من الطلبة (أي 30 طالب) تقل أطوالهم عن 172.5 سم أما 25 % من الطلبة المتبقين (أي 10 طلاب) فتتفوق أطوالهم 172.5 سم.

2. العشيريات :

العشير الاول :

$$RD_1 = \frac{\sum n_i}{10} = \frac{40}{10} = 4$$

$RD_1 = 4$ ليس محصورا بل يقابل مباشرة الحد 150. اذن قيمة العشير الاول تساوي 150 سم.

$$D_1 = 015 \text{ سم}$$

التفسير: 10% من الطلبة تقل اطوالهم عن 150 سم اما 90% من الطلبة المتبقين فتفوق اطوالهم 150 سم.
العشير الخامس :

$$RD_5 = \frac{5 \sum n_i}{10} = \frac{\sum n_i}{2} = RM_s = RQ_2$$

و بالتالي رتبة العشير الخامس هي نفس رتبة الوسيط و الربيع الثاني لذلك نستنتج ان :

$$165 D_5 = M_e = Q_2 = \text{سم}$$

التفسير: 50% من الطلبة (أي 20 طالب) تقل اطوالهم عن 165 سم اما 50% من الطلبة المتبقين (أي 20 طالب) فتفوق اطوالهم 165 سم.

العشير التاسع :

$$36 = \frac{9 \sum n_i}{10} = \frac{9 \times 40}{10} = RD_9$$

$RD_9 = 36$ ليس محصورا بل يقابل مباشرة الحد 180. اذن قيمة العشير التاسع تساوي 180 سم.

$$D_9 = 018 \text{ سم}$$

التفسير: 90% من الطلبة تقل اطوالهم عن 180 سم اما 10% من الطلبة المتبقين فتفوق اطوالهم 180

3. المئينيات :

المئين الاول :

$$RP_1 = \left[\frac{\sum n_i}{100} = \frac{40}{100} = 0.4 \right] 051 - 041 \text{ فئة المئين}$$

الاول

$$P_1 = L_{Min} + \left[\frac{RQ_{P_1} - s}{s' - s} \right] a_{P_1}$$

$$= 140 + \left[\frac{0.4 - 0}{4 - 0} \right] 10$$

$$= 411 \text{ سم}$$

التفسير: 1% من الطلبة تقل اطوالهم عن 141 سم اما 99% من الطلبة المتبقين فتفوق اطوالهم 141 سم.

المئين الخمسون :

$$RP_{50} = \frac{50 \sum n_i}{100} = \frac{\sum n_i}{2} = RM_s = RQ_2 = RD_5$$

لذلك نستنتج ان :

$$P_{50} = D_5 = M_e = Q_2 = 165 \text{ سم}$$

التفسير: 50% من الطلبة (أي 20 طالب) تقل أطوالهم عن 165 سم أما 50% من الطلبة المتبقين (أي 20 طالب) فتفوق أطوالهم 165 سم.

المئين التاسع و التسعون :

فئة المئين التاسع [180-190]

$$RP_{99} = \frac{99 \sum n_i}{100} = \frac{99 \times 40}{100} = 39.6$$

و التسعون

$$P_{99} = L_{Min} + \left[\frac{RQ_{P_{99}} - s}{s' - s} \right] a_{P_{99}}$$

$$= 180 + \left[\frac{39.6 - 36}{40 - 36} \right] 10$$

189 سم

=

التفسير: 99% من الطلبة تقل أطوالهم عن 189 سم أما 1% من الطلبة المتبقين فتفوق أطوالهم 189 سم

تمارين مقترحة :

التمرين الاول

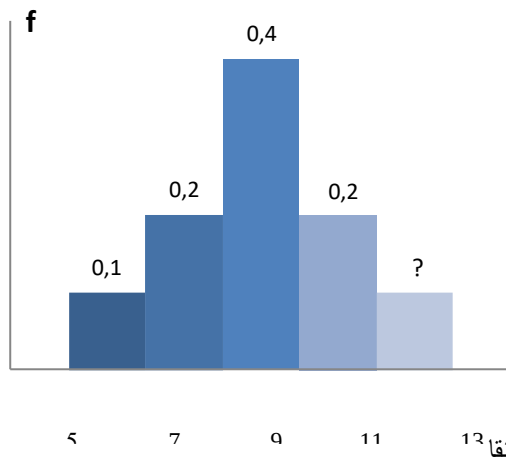
8 9 9 14 8 8 10 7 6 9 7 8 10 14 11 8 14 11 -
110 731 1031 84 20 118 1162 1977 103 752-
104 104 104 104 104 107 109 109 10 109-

1. احسب متوسط السلسلتين باستعمال الوسط الحسابي، الهندسي، التوافقي ثم التربيعي.

2. ماذا تستنتج؟

تمرين الثاني:

يبيّن الجدول التالي توزيع 100 عامل لمؤسسة تصفية المياه حسب الأجور الشهرية بآلاف الدينارات:



الأجور 10 ³ دج	[44 - 42]	[42 - 40]	[40 - 38]	[38 - 36]	[36 - 34]	[34 - 32]	[32 - 30]
العمال	12	13	n ₅	20	15	13	12

المطلوب:

- 1- ما هو المتغير الإحصائي المدروس و ما هي طبيعته؟
- 2- أحسب قيمة الأجر الوسيط ثم الربع الثاني مع تفسير النتيجة.
- 3- أحسب قيمة المنوال مع تفسير النتيجة.
- 4- مثل بيانيا الوسيط ثم المنوال.
- 5- ماذا تستنتج؟

التمرين الثالث: الجدول التالي يعرض مبيعات 40 مصنع من منتجاته خلال شهر بالطن

فئات قيمة المبيعات	32-34	34-36	36-38	38-40	40-42	42-44
عدد المصانع	4	7	13	10	5	1

المطلوب

1. أحسب الوسط الحسابي بعلاقة التعريف (الطريقة المباشرة) .
2. احسب الوسط الحسابي بطريقة تبديل المتغير .
3. احسب متوسط المردود باستعمال الوسط التوافقي، التربيعي. ثم الهندسي. و ماذا تستنتج

التمرين الرابع

يمثل الجدول التالي توزيع مجموعة من الطلاب سنة أولى LMD حسب نفقاتهم اليومية بـ دج :

النفقات اليومية	[10 - 50[	[50 - 70 [	[70 - 150 [	[150 - 250 [	المجموع
عدد الطلبة	4	6	7	3	20

المطلوب

1. ما هو المجتمع المدروس؟ ما هي الصفة المدروسة و طبيعتها؟
2. ارسم المدرج التكراري لهذا التوزيع.
3. حدد قيمة المنوال بيانيا و حسابيا مع تفسير النتيجة.
4. حدد قيمة الوسيط بيانيا و حسابيا مع تفسير النتيجة .

التمرين الخامس :

البيانات التالية تمثل عدد أفراد الأسرة لثلاث عينات مختلفة:

العينة (1): 2, 3, 5, 6, 7, 8, 10, 13, 2, 6, 4, 7

العينة (2): 3, 2, 5, 2, 3, 4, 6

العينة (3): 1, 5, 6, 7, 8, 9, 10

أحسب الوسيط ثم المنوال ثم الوسط الحسابي

التمرين السادس: الجدول التالي يوضح توزيع عينة من العقارات حسب أسعارها

فئات الاسعار	عدد العقارات				
15 - 17	3				
17 - 19	7				
19 - 21	21				
21 - 23	5				
23 - 25	4				

المطلوب :

1. قارن بين مقاييس النزعة المركزية

2. مثلها بيانيا

التمرين السابع: الجدول التالي يوضح توزيع عينة من المرضى حسب ساعات التماثل للشفاء

الفئات	عدد المرضى					
15 - 19	6					
19 - 23	14					
23 - 27	42					
27 - 31	10					
31 - 35	8					
	عدد المرضى					

1. أحسب الوسط الحسابي بعلاقة التعريف (الطريقة المباشرة) .
2. احسب الوسط الحسابي بطريقة تبديل المتغير .
3. احسب متوسط المردود باستعمال الوسط التوافقي، التربيعي. ثم الهندسي. و ماذا تستنتج
4. مثل بيانا الوسط الحسابي

التمرين الثامن

البيانات التالية تمثل قيمة المنتجات المباعة بآلاف الريالات لإحدى الشركات التجارية في الـ 70 يوم الأولى من العام الحالي.

56 75 70 66 60 55 65 70 65 56 66 71 62 67 71 61 67 61 70

60 75 69 71 57 69 72 68 57 72 68 65 63 73 66 63 58 73 67

62 72 58 74 60 81 80 74 76 74 73 58 72 94 78 91 85 77 83

77 82 76 62 78 88 64 87 55 79 57 64 79

والمطلوب:

1. كون التوزيع التكراري لقيمة المبيعات.

2. كون التوزيع التكراري النسبي.

3. أحسب الربيع الأول، الربيع الثاني و الربيع الثالث ثم إستنتج القيمة الوسيطة لأطوال الطلبة مع تفسير النتائج.

4. أحسب قيمة العشير الأول، الخامس ثم التاسع مع تفسير النتائج.

5. أحسب قيمة المئين الأول، الخمسون ثم الأخير مع تفسير النتائج

الفصل الرابع مقاييس التشتت والشكل

Measures of Dispersion

سنتطرق من خلال هذا الفصل إلى العناصر التالية

أولاً- مقاييس تشتت مطلقة

ثانياً- مقاييس تشتت نسبية

ثالثاً- مقاييس الالتواء

رابعاً- مقاييس التفرطح

الفصل الرابع مقاييس التشتت والشكل

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

I - مقاييس تشتت مطلقة.

1- المدى : Rang

هو الفرق بين أكبر و أصغر قيمة في السلسلة. $E = \text{Max}_{(x_i)} - \text{Min}_{(x_i)}$

2. التباين variance :

التباين هو عبارة عن الوسط الحسابي لمربعات الفروق بين قيم المتغير الإحصائي و الوسط الحسابي و يرمز له بالرمز $V(x)$.

$$V(x) = \frac{\sum (x_i - \bar{x})^2}{N} \quad \text{بيانات غير مَبَوَّبة :}$$

$$V(x) = \frac{\sum [n_i (x_i - \bar{x})^2]}{\sum n_i} \quad \text{بيانات مَبَوَّبة :}$$

$$V(x) = \sum f_i (\bar{x}_i - \bar{x})^2 \quad \text{التباين باستخدام التكرارات النسبية}$$

القانون المنشور : علاقة كونيغ : Relation of Köning :

$$V(x) = \delta_x^2 = \frac{\sum xi^2}{N} - \bar{x}^2 \quad \text{بيانات غير مَبَوَّبة :}$$

$$V(x) = \delta_x^2 = \frac{\sum xi \cdot xi^2}{\sum xi} - \bar{x}^2 \quad \text{بيانات مَبَوَّبة :}$$

$$V(x) = \delta_x^2 = \sum Fi \cdot x^2 - \bar{x}^2 \quad \text{باستخدام التكرارات النسبية :}$$

3. الانحراف المعياري : Standard Deviation :

$$\delta_x = \sqrt{\text{التباين}} = \sqrt{v(x)}$$

II - مقاييس نشتت نسبية : Relative measures of dispersion :

نستعمل مقياس نشتت نسبي حتى نحصل على رقم خال من وحدات القياس (في حالة المقارنة بين مجموعات

بوحداث قياس مختلفة)

1. معامل الاختلاف (C.V) Coefficient of Variation :

$$CV = \frac{\delta_x}{\bar{x}} \times 100$$

2. معامل المدى الربيعي النسبي : Relative interquartile Coefficient :

$$C_Q = \frac{Q_3 - Q_1}{Q_2} \times 100$$

مقاييس الشكل : كما أن 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 نجد أنهما مختلفين تماما من حيث الشكل، إن الاختلاف من حيث الشكل يتجلى في إحدى الأمورين: الالتواء أو التفرطح. ولإتمام الصورة يوفق الوصف السابق بمقاييس آخرين (مقاييس الشكل) يقيس الأول التواء منحنى التوزيع *Asymétrie* والثاني يقيس قيمة المنحنى كونها حادة أو مفرطة *Aplatissement*.

سندرس في هذا الفصل مقاييس الالتواء (مقاييس عدم التماثل أي عدم التناظر) ونخص بالذكر هنا معامل الالتواء لبيرسون.

مقاييس الالتواء

هناك طرق كثيرة لقياس الالتواء ومنها ما يلي:

I. معامل عدم التماثل (معامل الالتواء)¹¹ Cp_1

$$-1 < Cp_1 < +1 \quad Cp_1 = \frac{\bar{x} - M_0}{\delta_x}$$

avec

نتيجة هذا المعامل لديها ثلاث حالات :

الحالة الأولى : حاله التماثل أو التناظر وهذا عندما يكون المعامل معدوم $\bar{x} = M_0$

$$Cp_1 = 0$$

الحالة الثانية : التواء (عدم التماثل) نحو اليمين وهذا عندما يكون المعامل موجب $\bar{x} > M_0$

$$0 < Cp_1$$

¹¹ Fabrice MAZEROLLE, Statistique Descriptive, Notes de cours, 2009

الحالة الثالثة : التواء (عدم التماثل) نحو اليسار وهذا عندما يكون المعامل سالب $\bar{x} < M_0$

$$Cp_1 < 0$$

II. معامل ثاني للتواء لبيرسون: Coefficient d'Asymétrie de Pearson

تستعمل في حالة التوزيعات القليلة التوزيعات الالتواء. تأخذ هذه الطريقة في الاعتبار العلاقة بين الوسط والوسيط والمنوال، في حالة ما إذا كان التوزيع قريب من التماثل وليس شديد الالتواء ، وهذه العلاقة هي :

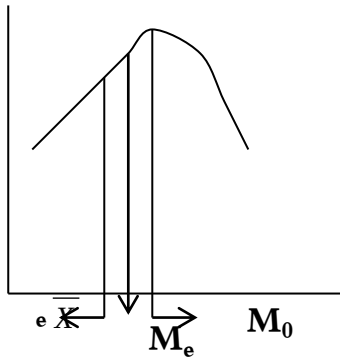
المتوسط الحسابي - المنوال = 3 (المتوسط الحسابي - الوسيط)

$$Cp_2 = \frac{3(\bar{x} - M_0)}{\delta_x} \quad -1 < Cp_2 < +1$$

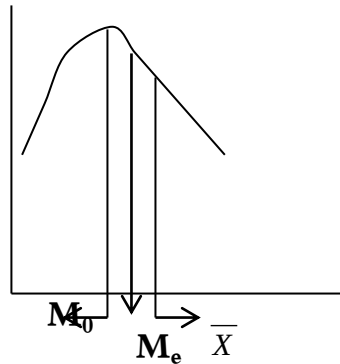
Cp_2 وهو صالح فقط بالتوزيعات قليلة الالتواء

أشكال الالتواء والتناظر: يمكن أن نميز بين ثلاث أشكال وهي:

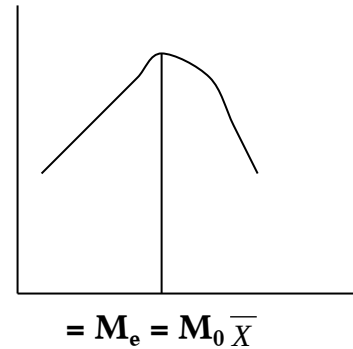
الحالة الثالثة



الحالة الثانية



الحالة الأولى



الحالة الأولى : توزيع متناظر:

الحالة الثانية : توزيع ملتوي نحو اليمين $\bar{x} > Me > M_0$

الحالة الثالثة :توزيع ملتوي نحو اليسار: أي $\bar{x} < Me < Mo$

التفرطح¹²:

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:

$$Max_{(xi)} - Min_{(xi)} = \text{كبير}$$

ب -توزيع متطاوول :وفيه يكون تشتت البيانات ضعيف جدا، كما هو موضح من خلال الشكل التالي:

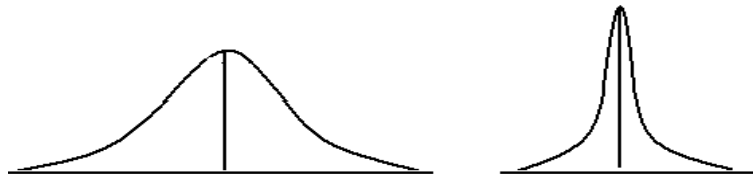
$$Max_{(xi)} - Min_{(xi)} = \text{صغير}$$



عند تمثيل التوزيع التكراري في شكل منحنى تكراري ، قد يكون هذا المنحنى منبسط ، أو مدبب ، فعندما يتركز عدد

أكبر من القيم بالقرب من منتصف المنحنى، ويقل في طرفيه، يكون المنحنى مدببا ، وعندما يتركز عدد أكبر على طرفي المنحنى ،

ويقل بالقرب من المنتصف يكون المنحنى مفرطحا ، أو منبسطا، ويظهر ذلك من الشكل التالي :



¹² فاروق عبد العظيم وآخرون مبادئ الإحصاء دار المعرفة الجامعية 2003 ص 81-83

منحنى مفرطح

منحنى مدبب

ويمكن قياس التفرطح باستخدام عدد من الطرق، ومنها طريقة العزوم، حيث يحسب معامل التفرطح (K) بتطبيق

المعادلة التالية :

$$k = \sum (x - \bar{x})^4 / n$$

حيث أن المقدار $\sum (x - \bar{x})^4 / n$ هو العزم الرابع حول الوسط، S هو الانحراف المعياري. ومعامل التفرطح

في التوزيع الطبيعي يساوي 3، ومن ثم يمكن وصف منحنى التوزيع من حيث التفرطح، والتدبب كما يلي :

- إذا كان $k=3$ كان منحنى التوزيع معتدلاً .
- إذا كان $k>3$ كان منحنى التوزيع مدبباً .
- إذا كان $k<3$ كان منحنى التوزيع منبسطاً (مفرطحاً) .

تمارين الفصل الخامس محلولة

التمرين رقم 1:

يمثل الجدول التالي الأجور الشهرية بـ 10^2 DA دج لـ 50 عامل في مؤسسة صناعية :

الأجور	[120-130[	[130-140[	[140-150[	[150-160[	[160-170[	[170-180[	[180-190[
عدد العمال	4	7	10	15	7	4	3

المطلوب :

1 - أحسب الانحراف المعياري : باستخدام القانون المنشور (علاقة كونيغ) .

2- أحسب معامل الالتواء لبيرسون Pearson حسابيا ثم بيانيا مع تفسير النتيجة.

الحل

الفئات	n_i	x_i	$n_i x_i$	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$n_i(x_i - \bar{x})^2$	$n_i x_i^2$
[120-130[	4	125	500	-27,6	761,76	3047,04	62500
[130-140[	7	135	945	-17,6	309,76	2168,32	127575
[140-150[	10	145	1450	-7,6	57,76	577,6	210250
[150-160[	15	155	2325	2,4	5,76	86,4	360375
[160-170[	7	165	1155	12,4	153,76	1076,32	190575
[170-180[	4	175	700	22,4	501,76	2007,04	122500
[180-190[	3	185	555	32,4	1049,76	3149,28	102675
المجموع	50	/	7630	/	/	12112	1176450

$$\bar{x} = \frac{\sum n_i x_i}{\sum n_i} = \frac{7630}{50} \Rightarrow \boxed{\bar{x} = 152,6}$$

1. حساب الانحراف المعياري:

حساب الانحراف المعياري بعلاقة التعريف:

$$\delta_x = \sqrt{\frac{\sum n_i (x_i - \bar{x})^2}{\sum n_i}} = \sqrt{\frac{12112}{50}} = \sqrt{242,24} = 15,564$$

حساب الانحراف المعياري بقانون المنشور: Koenig

$$\delta_x = \sqrt{\frac{\sum n_i x_i^2}{\sum n_i} - \bar{x}^2} = \sqrt{\frac{1176450}{50} - 152,6^2} = \sqrt{242,24} = 15,564$$

نفس النتيجة.

2. حساب معامل عدم التماثل (معامل الالتواء) حسابيا ثم بيانيا مع تفسير النتيجة:

حساب معامل عدم التماثل (معامل الالتواء): Pearson:

$$Cp_1 = \frac{\bar{x} - M_0}{\delta_x}$$

لدينا: $\bar{x} = 152,6$, $\delta_x = 15,564$ المنوال بعد الحساب:

$$846,3 M_0 = 15$$

$$Cp_1 = \frac{152,6 - 153,846}{15,564} = -0,08$$

$$Cp_1 < 0$$

سالب $\Leftarrow$ التواء قليل نحو اليسار (توزيع قريب من التماثل).

الالتواء بيانيا : من خلال الربط بين نقاط محور الفواصل (مراكز الفئات) x_i والمحور العمودي (التكرارات المطلقة

n_i) نلاحظ وجود التواء ضعيف جدا نحو اليسار وهذا لان معامل Pearson حسابيا سالب اي ان المتوسط

الحسابي $\bar{x} = 152,6$ أصغر من المنوال $M_0 = 153,846$

 $\bar{x}$ Moالتواء قليل نحو اليسار $Cp_1 < 0$ التمرين رقم 2:

إذا كان متوسط الأجور ل 100 عامل في مشروع معين 24 دج بانحراف معياري 1,56 دج. أخذت عينة من

العاملات في نفس المشروع و تبين أن أجورهن موزعة كالتالي:

الفئات	-20 [	-22 [	-24 [	-26 [	-28 [	-30 [
	[22	[24	[26	[28	[30	[32
التكرارات	10	20	30	25	8	7

المطلوب :

قارن بين التوزيعين باستخدام معامل الاختلاف المعياري. ماذا يمكن استنتاجه؟ (هل تتفق مع الرأي القائل بأن

توزيع أجور العمال أكثر عدالة من توزيع أجور العاملات؟).

الحل

الفئات	n_i	x_i	$n_i x_i$	$n_i x_i^2$
[20-22[	10	21	210	4410
[22-24[	20	23	460	10580
[24-26[	30	25	750	18750
[26-28[	25	27	675	18225
[28-30[	8	29	232	6728
[30-32[	7	31	217	6727
المجموع	100	/	2544	65420

معامل الاختلاف المعياري لأجور العمال (الرجال):

$$\bar{x}_H = 24 \text{ DA}$$

$$\delta_{x_H} = 1,56 \text{ DA}$$

$$CV_H = \frac{\delta_{x_H}}{\bar{x}_H} * 100 = \frac{1,56}{24} * 100 = 6,5\%$$

معامل الاختلاف المعياري لأجور العاملات (النساء):

المتوسط الحسابي للعاملات (النساء):

$$\bar{x}_F = \frac{\sum n_{iF} x_i}{\sum n_{iF}} = \frac{2544}{100} \Rightarrow \boxed{\bar{x} = 25,44}$$

انحراف معياري للعاملات (النساء):

$$\delta_{x_F} = \sqrt{\frac{\sum n_{iF} x_i^2}{\sum n_{iF}} - \bar{x}_F^2} = \sqrt{\frac{65420}{100} - (25,44)^2} = 2,64$$

معامل الاختلاف المعياري لأجور العاملات (النساء):

$$CV_F = \frac{\delta_{x_F}}{\bar{x}_F} * 100 = \frac{2,64}{25,44} * 100 = 10,37\%$$

المقارنة:

$$CV_H < CV_F$$

$$6,5\% < 10,37\%$$

نستنتج ان توزيع الاجور اكثر تجانسا عند العمال الرجال من توزيع اجور العاملات النساء.

اذن اجور الرجال اقل تشتتا (اقل تفاوت) و انها متقاربة $\Leftrightarrow$ ان توزيع اجور العمال الرجال اكثر عدالة من توزيع اجور العاملات.

التمرين 6: الجدول التالي يخلص معلومات عن رقم اعمال مؤسسة صناعية بمليون دينار

الوضع و النزعة المركزية	التشتت
3500 ادنى قيمة	650 الانحراف المعياري
4900 متوسط حسابي	1100 المدى الربيعي
4550 المنوال	2800 المدى العشيري
4600 الوسيط	5000 المدى العام
4100 الربيع الاول	
3700 العشير الاول	

المطلوب :

1. احسب اقصى رقم اعمال:

2. احسب الربيع الثالث

3. احسب العشير التاسع

4. رتب هذه المقاييس على محور

الحل

1. حساب اقصى رقم اعمال:

علما بقيمة أدنى رقم اعمال و المدى العام, نستطيع حساب اقصى رقم اعمال:

$$MAX = \text{المدى العام} + 3500 = 5000 + 3500 = 8500$$

اقصى رقم اعمال هو: $CA_{MAX} = 8500000 \text{ DA}$ لان المعطيات بآلاف الدينانير الجزائرية.

2. حساب الربيع الثالث و العشير التاسع:

حساب الربيع الثالث:

لحساب Q_3 نستعمل Q_1 و المدى الربيعي. لدينا:

$$I_Q = Q_3 - Q_1 \Rightarrow Q_3 = Q_1 + I_Q \Rightarrow Q_3 = 4100 + 1100 = 5200$$

الربيع الثالث يساوي: $Q_3 = 5200000 \text{ DA}$

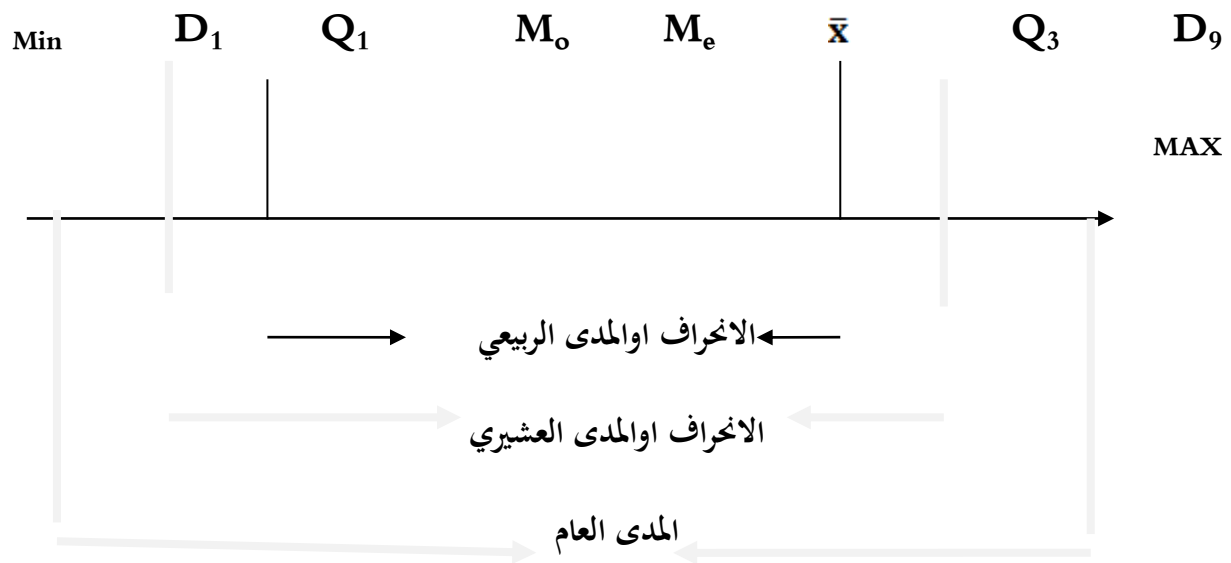
3. حساب العشير التاسع:

بنفس الطريقة نتحصل على D_9

$$I_D = D_9 - D_1 \Rightarrow D_9 = D_1 + I_D \Rightarrow D_9 = 3700 + 2800 = 6500$$

العشير التاسع يساوي: $D_9 = 6500000 \text{ DA}$

4. وضع في محور مقاييس هذه السلسلة:



التمرين 4 :

كانت درجات 8 مستهلكين على منتج جديد طرح بالسوق ، كالتالي.

66 85 52 78 80 91 74 58

والمطلوب : 1- حساب معامل الالتواء بطريقة " بيرسون " .

الحل

1- حساب معامل الالتواء بطريقة " بيرسون " .

• حساب الوسط الحسابي ، والانحراف المعياري :

الدرجة x	x^2
66	4356
85	7225

$$\sum x = 584 , \sum x^2 = 43890$$

ويكون :

52	2704
78	6084
80	6400
91	8281
74	5476
58	3364
584	43890

$$\bar{x} = \frac{\sum x}{n} = \frac{584}{8} = 73$$

$$\begin{aligned}\sigma_X &= \sqrt{\frac{\sum x^2 - (\sum x)^2/n}{n-1}} = \sqrt{\frac{43890 - (584)^2/8}{8-1}} \\ &= \sqrt{\frac{1258}{7}} = \sqrt{179.71428} = 13.406\end{aligned}$$

• حساب الوسيط :

$$Me = 74 + 0.5(78 - 74) = 76$$

معامل الالتواء "بيرسون"

$$cP = \frac{(\bar{x} - Me)}{\sigma} = \frac{(73 - 76)}{13.406} = -0.22$$

إذا منحني توزيع درجات الطلاب ملتوي جهة اليسار

التمرين 5: بنفس معطيات التمرين 4

احسب معامل التفرطح

$$\bar{x} = 73$$

x	66	85	52	78	80	91	74	58	584
$(x - \bar{x})$	-7	12	-21	5	7	18	1	-15	0
$(x - \bar{x})^2$	49	144	441	25	49	324	1	225	1258
$(x - \bar{x})^4$	2401	20736	194481	625	2401	104976	1	50625	376246

ومن البيانات أعلاه نجد أن:

$$\sigma_x = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}} = \sqrt{\frac{1258}{7}} = 13.406$$

$$\frac{1}{n} \sum (x - \bar{x})^2 = \frac{1}{8} (376246) = 47030.75$$

إذا معامل التفرطح هو:

$$K = \frac{47030.75}{(13.406)^4} = \frac{47030.75}{32299.58} = 1.456$$

إذا شكل توزيع بيانات الدرجات مفطح

الفصل الخامس الانحدار و الارتباط

سنتطرق من خلال هذا الفصل إلى العناصر التالية:

اولا الارتباط بين متغيرتين

ثانيا الانحدار الخطي البسيط

الفصل الخامس الارتباط والانحدار الخطي

Simple Linear regression¹³ الانحدار الخطي البسيط

يعرف الانحدار الخطي البسيط بأنه عملية تقدير العلاقة الخطية بين متغيرين أحدهما مستقل و الآخر تابع إن مفهوم الانحدار الخطي البسيط يقتزن بمفهوم الارتباط الخطي البسيط . و يهدف الانحدار الخطي البسيط إلى تقدير قيم عددية لمعالم النموذج ، أي تقدير قيم عددية لكل من $\hat{a}$ و $\hat{b}$ و معادلة خط

$$\hat{b} = \frac{\sum X_i Y_i - n \bar{X} \bar{Y}}{\sum X_i^2 - n \bar{X}^2}$$

الانحدار البسيط هي كالأتي:

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

ومن أعلاه نلاحظ بان $\hat{a}, \hat{b}$ هي دالة بدلالة قياسات مفردات العينة (X_i, Y_i) إن $\hat{a}$ تمثل في الحقيقة مقطع خط الانحدار مع المحور الصادي Y أي ما نعنيه قيمة Y التقديرية عندما $X=0$ و إذا كانت $\hat{a}=0$ عندئذ يقال إن خط الانحدار يمر من نقطة الأصل $(0,0)$ وان $\hat{b}$ تسمى معامل الانحدار أي مقياس يوضح مقدار تغير Y إذا ما تغيرت X بوحدة واحدة و يلاحظ إن $\hat{b}$ قد تكون موجبة أو سالبة تبعا لقيمة الانحراف المعياري لـ X, Y (S_{xy}) إذا كانت موجبة أم سالبة فقيمة $\hat{b}$ الموجبة تعني إن العلاقة بين X, Y موجبة و القيمة السالبة لها تعني إن العلاقة سالبة (عكسية) و غالبا ما يقال لما سبق أنه تم تقدير معادلة الانحدار (Y / X) أي إيجاد معادلة الانحدار فيها المتغير التابع هو Y و المستقل X و العكس ممكن.

¹³ خليل شرف الدين، الإحصاء الوصفي، شبكة الأبحاث والدراسات الاقتصادية، ص 96

1- الارتباط الخطي¹⁴: Coefficient de corrélation

لدراسة الارتباط نستعين بالانحدار. نبحث عن انحدار الظاهرة Y بالنسبة للظاهرة X و عن انحدار الظاهرة X بالنسبة للظاهرة Y .

بما أن كل ظاهرة تتغير بدلالة الأخرى فإن هذا يؤدي ألا وجود ارتباط بين الظاهرتين.
رأينا أن:

$$a = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

و أن:

$$a' = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (y_i - \bar{y})^2}$$

نحصل على معامل الارتباط بالعلاقة التالية:

$$r = \sqrt{a \cdot a'} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

بمعنى أن:

$$r^2 = a \cdot a'$$

من معامل الارتباط، نجد أنه يركز على نقطتين هما:

- نوع العلاقة: وتأخذ ثلاث أنواع حسب إشارة معامل الارتباط كما يلي:
 - 1- إذا كانت إشارة معامل الارتباط سالبة ($r < 0$) توجد علاقة عكسية بين المتغيرين، بمعنى أن زيادة أحد المتغيرين يصاحبه انخفاض في المتغير الثاني، والعكس.
 - 2- إذا كانت إشارة معامل الارتباط موجبة ($r > 0$) توجد علاقة طردية بين المتغيرين، بمعنى أن زيادة أحد المتغيرين يصاحبه زيادة في المتغير الثاني، والعكس.
 - 3- إذا كان معامل الارتباط قيمته صفراً ($r = 0$) دل ذلك على انعدام العلاقة بين المتغيرين.

¹⁴ عوض منصور مبادئ الإحصاء الوصفي دار صفاء للنشر. الأردن 2000. ص 213-215

قوة العلاقة: ويمكن الحكم على قوة العلاقة من حيث درجة قربها أو بعدها عن (± 1) ، حيث أن قيمة معامل الارتباط تقع في المدى $(-1 < r < 1)$ ، وقد صنف بعض الإحصائيين درجات لقوة العلاقة يمكن تمثيلها على الشكل التالي

1- معامل التحديد : Coefficient de détermination

هو مربع معامل الارتباط :

$$r_{x,y}^2 = a.a'$$

معامل التحديد مقياس يبين أو يقيس مدى تأثير المتغير المستقل (X_i) على المتغير التابع (Y_i) .

$$0 \leq r_{x,y}^2 \leq 1 \quad \text{ملاحظة :}$$

تمارين الفصل الخامس محلولة

التمرين 1 :

فيما يلي المساحة المنزرعة بالأعلاف الخضراء بالألف هكتار، وإجمالي إنتاج اللحوم بالألف طن.

المساحة x	305	313	297	289	233	214	240	217
الكمية y	592	603	662	607	635	699	719	747

والمطلوب: حساب معامل الارتباط بين المساحة والكمية، وما هو مدلوله

1. معادلة خط انحدار المساحة على الكمية

2. حساب معامل الارتباط بطريقتين مختلفتين مع تفسير النتائج

3. معامل التحديد مع تفسير النتائج

الحل

بفرض أن (x) هي المساحة المنزرعة، (y) هي الكمية، ولحساب معامل الارتباط بين (y, x)

يتم تطبيق المعادلة ، وذلك على النحو التالي:

3. حساب الوسط الحسابي لكل من المساحة، والكمية $(\bar{y}, \bar{x})$.

$$\bar{x} = \frac{\sum x}{n} = \frac{2108}{8} = 263.5, \quad \bar{y} = \frac{\sum y}{n} = \frac{5264}{8} = 658$$

x	y	xy	x^2	y^2	المجاميع المطلوبة
305	592	180560	93025	350464	$\sum x = 2108, \sum y = 5264$

313	603	188739	97969	363609	$\sum xy = 1373536$
297	662	196614	88209	438244	$\sum x^2 = 567498$
289	607	175423	83521	368449	$\sum y^2 = 3487562$
233	635	147955	54289	403225	
214	699	149586	45796	488601	
240	719	172560	57600	516961	
217	747	162099	47089	558009	
2108	5264	1373536	567498	3487562	

4. معادلة خط انحدار الاستهلاك على الدخل

5. معادلة خط انحدار الدخل على الاستهلاك

6. حساب معامل الارتباط بطريقتين مختلفتين مع تفسير النتائج

معامل التحديد مع تفسير النتائج

باستخدام المجاميع السابقة، وبالتطبيق على المعادلة ، نجد أن معامل الارتباط قيمته هي:

$$\begin{aligned}
 r &= \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{\left(\sum x^2 - \frac{(\sum x)^2}{n}\right)\left(\sum y^2 - \frac{(\sum y)^2}{n}\right)}} \\
 &= \frac{1373536 - \frac{(2108)(5264)}{8}}{\sqrt{\left(567498 - \frac{(2108)^2}{8}\right)\left(3487562 - \frac{(5264)^2}{8}\right)}} \\
 &= \frac{-13528}{\sqrt{(12040)(23850)}} = \frac{-13528}{16945.619} = -0.798
 \end{aligned}$$

يوجد ارتباط عكسي قوي بين المساحة المنزرعة، وكمية إنتاج اللحوم

التمرين 2 :

لدينا المتغيرين x و y : حيث أن x يمثل الدخل الشهري و y يمثل النفقات الشهرية بمليون سنتيم ل 6 أسر

حسب الجدول التالي:

54	15	12	09	07	06	05	X_i
42	12	08	07	06	05	04	y_i

المطلوب :

أوجد معادلة انحدار النفقات على الدخل Y/X (باستخدام علاقة التعريف).

قدر النفقات إذا كان الدخل يساوي 20 مليون سنتيم.

أوجد معادلة انحدار الدخل على النفقات X/Y (باستخدام علاقة التعريف).

قدر النفقات إذا كان الدخل يساوي 35 مليون سنتيم.

حساب التباين المشترك بين X و y. فسر النتيجة.

الحل :

حساب المتوسطان الحسابيان ل X و y :

$$\bar{x} = \frac{54}{6} = 9$$

$$\bar{y} = \frac{42}{6} = 7$$

$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$	$(y_i - \bar{y})$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})$	y_i	x_i
9	12	-3	16	-4	04	05
4	6	-2	9	-3	05	06
1	2	-1	4	-2	06	07
0	0	0	0	0	07	09
1	3	1	9	3	08	12
25	30	5	36	6	12	15
40	+ 53		74		42	54

1- معادلة انحدار النفقات على الدخل Y/X : (باستخدام علاقة التعريف).

$$D : y/x \quad y_i = ax_i + b$$

$$a = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} = \frac{+53}{74} = +0,716$$

$$b = \bar{y} - a\bar{x} = 7 - (0,716 \times 9) = 0,556$$

و تكون المعادلة التي نبحث عنها هي :

$$y_i = 0,716x_i + 0,556$$

و تعبر عن العلاقة النظرية الموجودة بين الظاهرتين X و Y .

العلاقة التي وجدناها هي دالة $y = f(x)$ أي أن الظاهرة Y تتغير بدلالة الظاهرة X و تعبر عن العلاقة عن

الانحدار Y/X .

2- تقدير النفقات إذا كان الدخل يساوي 20 مليون سنتيم :

$$y = 0,716(20) + 0,556 = 14,876 \approx 15 \Rightarrow x = 20 \text{ إذا كان}$$

3- معادلة انحدار الدخل على النفقات X/Y : (باستخدام علاقة التعريف).

$$D : x/y \quad x_i = a'y_i + b'$$

يمكن أن نريد معرفة الانحدار X/Y الذي يعبر عن تغيير الظاهرة Y بدلالة الظاهرة X ، فيجب أن نبحث عن

$$\text{دالة } x = f(y).$$

نستعمل نفس الطريقة و نجد في هذه الحالة أن الميل هو :

$$a' = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (y_i - \bar{y})^2} = \frac{+53}{40} = +1,325$$

$$b' = \bar{x} - a'\bar{y} = 9 - (1,325 \times 7) = -0,275$$

و تكون المعادلة التي نبحث عنها هي :

$$x_i = 1,325y_i - 0,275$$

4- تقدير الدخل إذا كانت النفقات تساوي 35 مليون سنتيم :

$$x = 1,325(35) - 0,275 = 46,1 \approx 46 \Rightarrow y_i = 46 \text{ إذا كان}$$

5- حساب التباين المشترك بين X و Y. مع تفسير النتيجة :

$$Cov(x, y) = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n} = \frac{+53}{6} = +8,833$$

التفسير : $Cov(x, y) > 0$ موجب $\Leftrightarrow$ علاقة طردية معنى ذلك أن المتغيران X و Y يتغيران في نفس الاتجاه.

نرجع إلى المثال السابق و نحسب معامل الارتباط :

$$r = \sqrt{a.a'} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} = \frac{+53}{\sqrt{74 \times 40}} = +0,974$$

و بعبارة أخرى:

$$r = \sqrt{aa'} = \sqrt{0,716 \times 1,325} = \pm 0,974$$

نرجع إلى المثال السابق و نحسب نحسب معامل التحديد :

$$r_{x,y}^2 = (0,974)^2 = 0,948 \times 100 = 94,8\%$$

التفسير :

يمكن القول إن تغيرات الإنفاق (y_i) ترجع بنسبة 94,8% إلى تغيرات الدخل (x_i) أما النسبة الباقية و هي

(100 - 94.8 = 5.2%) ترجع إلى عوامل أخرى غير مدرجة في المعادلة (نجهلها).

لتكن السلسلة الإحصائية التالية التي تمثل نفقات البحث (X) و الربح (Y):

60	45	26	42	40	نفقات البحث X
80	50	40	60	50	Y

المطلوب:

1- أوجد معادلة مستقيم الانحدار Y على X ثم معادلة مستقيم الانحدار X على Y.

2- أحسب معامل الارتباط الخطي بين X و Y. فسر النتيجة.

3 - إذا افترضنا أن الظاهرة تتبع نفس التطور في المستقبل، حدد قيمة الربح لما تكون قيمة نفقات البحث تساوي 80.

الحل:

Yi ²	x _i ²	XiYi	الربح Yi	نفقات البحث Xi
2500	1600	2000	50	40
3600	1764	2520	60	42
1600	676	1040	40	26
2500	2025	2250	50	45
6400	3600	4800	80	60
16600	9665	12610	280	213

1: معادلة مستقيم الانحدار Y على X: نحدد قيمة الثابتين a و b :

$$a = \frac{\text{cov}_{xy}}{\sigma_x^2}.$$

$$\text{cov}_{xy} = \frac{\sum x_i y_i}{N} - \bar{x} \bar{y}$$

$$\bar{x} = \frac{\sum X_i}{N} = \frac{213}{5} \Rightarrow \boxed{\bar{x} = 42,6}$$

$$\bar{y} = \frac{\sum y_i}{N} = \frac{280}{5} \Rightarrow \boxed{\bar{y} = 56}$$

$$\text{cov}_{xy} = \frac{12610}{5} - 42,6 \times 56 \Rightarrow \boxed{\text{cov}_{xy} = 136,4}$$

$$\sigma_x^2 = \frac{\sum x_i^2}{N} - \bar{x}^2 = \frac{9665}{5} - (42,6)^2 \Rightarrow \boxed{\sigma_x^2 = 118,24}$$

$$a = \frac{136,4}{118,24} = 1,153 \Rightarrow \boxed{a = 1,153}$$

$$b = \bar{y} - a \bar{x} = 56 - (1,153 \times 42,6) \Rightarrow \boxed{b = 6,88}$$

$$\boxed{y_i = 1,153x_i + 6,88} \quad \text{معادلة مستقيم الانحدار Y على X:}$$

$$x_i = a'y_i + b'$$

معادلة مستقيم الانحدار X على Y:

نحدد قيمة الثابتين a' و b' :

$$a' = \frac{\text{COV}_{xy}}{\delta_y^2}.$$

$$\delta_y^2 = \frac{\sum y_i^2}{N} - \bar{y}^2 = \frac{16600}{5} - (56)^2 \Rightarrow \boxed{\delta_y^2 = 184}$$

$$a' = \frac{136,4}{184} = 1,153 \Rightarrow \boxed{a = 0,741}$$

$$b' = \bar{x} - a'\bar{y} = 42,6 - (0,741 \times 56) \Rightarrow \boxed{b' = 1,104}$$

$$\boxed{x_i = 0,741y_i + 1,104}$$
 معادلة مستقيم الانحدار X على Y:

2- معامل الارتباط الخطي بين X و Y : (r_{xy})

$$r_{xy} = \frac{\text{COV}_{xy}}{\delta_x \delta_y} = \frac{136,4}{\sqrt{118,24} \sqrt{184}} \Rightarrow \boxed{r_{xy} = 0,924}$$

العلاقة بين نفقات البحث والربح طردية وقوية جدا قريبة من التطابق

3- تحديد قيمة الربح:

$$\boxed{y_i = 1,153x_i + 6,88}$$

$$x = 80$$

$$y = (1,153 \times 80) + 6,88 \Rightarrow \boxed{y = 99,12}$$

تمارين مقترحة

تمرين 1: البيانات التالية تمثل قيمة الانتاج (Y) و العمل (L) و رأس المال (K) لعينة بحجم (20)

معمل.

K	L	Y	K	L	Y
80	30	115	90	15	82
30	60	64	40	39	73
60	100	140	20	99	58
40	95	85	60	12	68
20	75	56	60	42	98
90	90	150	30	95	83
30	25	65	60	45	100
10	80	36	80	36	110
40	12	57	80	40	120
20	65	50	40	65	95

المطلوب :

- قدر النموذج انحدار الانتاج على العمل ورأس المال.

- احسب معامل الارتباط بين الانتاج و العمل

- احسب معامل الارتباط بين الانتاج ورأس المال

تمرين 2 : البيانات التالية تمثل قيمة الدخل (Y) و الاستهلاك (C) لعينة بحجم (16) عائلة.

Y	C	Y	C
430.6	336.6	306.6	230.7
460.6	356.6	304.9	245.1
485.3	376.6	333.0	255.2
521.7	402.9	352.8	265.5
568.4	434.7	368.2	280.3
625.1	468.3	370.0	291.1
659.0	494.3	402.4	308.2
719.8	538.9	417.1	326.5

المطلوب :

- قدر نموذج انحدار الاستهلاك على الدخل.

- قدر نموذج انحدار الدخل على الاستهلاك

- احسب معامل الارتباط بين الاستهلاك والدخل

تمرين 3 فيما يلي المساحة المنزرعة (x) بالأعلاف الخضراء ، وإجمالي إنتاج اللحوم (y)

$$\sum x y = 516 \quad \sum y = 100 \quad \sum x = 120$$

$$n = 40 \quad \sum y^2 = 410 \quad \sum x^2 = 720$$

1. معادلة خط انحدار إنتاج اللحوم على المساحة المنزرعة بالأعلاف الخضراء
2. حساب معامل الارتباط مع تفسير النتائج
3. معامل التحديد مع تفسير النتائج

تمرين 4 لدينا العلاقة بين الدخل X و الإنفاق Y لمجموعة من الأسر ممثلة من خلال المعادلتين التاليتين:

$$Y_i = 0,5 x_i + 30$$

$$X_i = Y_i + 20$$

1. أحسب معامل الارتباط الخطي مع تفسير النتيجة
2. احسب معامل التحديد مع تفسير النتيجة
3. احسب تباين y علما أن التباين المشترك بين x و y يساوي 20

تمارين لمراجعة عامة السداسي الاول

Test Durée : 30 min

Sujet : A

التمرين الأول:

يمثل الجدول التالي الأجر اليومي لـ 500 عامل في المصنع A:

الأجر اليومي (^{210})	[1-5[	[5-9[	[9-13[	[13-17[	[17-21[	[21-25[
عدد العمال	30	50	60	230	100	30

$$\sum n_i x_{iA} = 7140 \quad \sum n_i x_{iA}^2 = 113700$$

إذا علمت أن في المصنع B الذي يؤجر نفس عدد العمال ، الأجر المتوسط يساوي 16,88 و الانحراف المعياري 4,255.

1. أحسب معامل الاختلاف المعياري للمصنع A

2. أحسب معامل الاختلاف المعياري للمصنع B

3. ما هو المصنع الذي يحتوي على الأجور الأكثر تجانس؟

التمرين الثاني:

لدينا توزيع إحصائي (Y, X) ممثل في الجدول التالي:

x_i	y_i					
1	4					
2	2					
2	0					
3	-2					

1. أحسب التباين المشترك بين X و Y مع تفسير النتيجة

2. أوجد معادلة الانحدار Y على X

3. إذا كان $X = 5$ كم سيكون قيمة Y

التمرين الأول:

يمثل الجدول التالي الاجر ايومي ل 500 عامل في المصنع A:

الأجر اليومي (^{210})	[2-6[	[6-10[	[10-14[	[14-18[	[18-22[	[22-26[
عدد العمال	10	20	70	200	150	50

$$\sum n_i x_i = 8440 \quad \sum n_i x_i^2 = 151520$$

إذا علمت أن في المصنع B الذي يؤجر نفس عدد العمال ، الأجر المتوسط يساوي 28,14 و الانحراف المعياري 4,845.

1. أحسب معامل الاختلاف المعياري للمصنع A

2. أحسب معامل الاختلاف المعياري للمصنع B

3. ما هو المصنع الذي يحتوي على الأجور الأكثر تجانس؟

التمرين الثاني: لدينا توزيع إحصائي (Y, X) ممثل في الجدول التالي:

x_i	y_i					
4	1					
2	2					
0	2					
-2	3					

1. أحسب التباين المشترك بين X و Y مع تفسير النتيجة

2. احسب معامل الارتباط الخطي ثم معامل التحديد

3. ادرس جودة معادلة الانحدار

Test

Durée : 30 min

Sujet : F

التمرين الأول:

أخذت عينات متساوية من الأسر لأربع ولايات و تم حساب متوسطات الإنفاق الشهري لها و كذلك انحرافاتها المعيارية فكانت النتائج كما يلي:

الولاية	A	B	C	D
متوسط الإنفاق (دج)	400	740	600	500
الانحراف المعياري (دج)	20	25	50	30

1. أحسب معامل الاختلاف المعياري

2. ما هي الولاية التي إنفاق أسرها أقل تشتت من غيرها؟

التمرين الثاني:

لدينا توزيع إحصائي (X, Y) ممثل في الجدول التالي:

x_i	y_i					
2	2					
1	4					
-2	3					
0	2					

1. أحسب التباين المشترك بين x و y مع تفسير النتيجة

2. أوجد معادلة الانحدار y على x

3. إذا كان معامل الارتباط الخطي يساوي -1 ، ما هي نوعية العلاقة الموجودة بين x و y ؟

Test

Durée : 30 min

Sujet : E

التمرين الأول:

يمثل الجدول التالي الإنفاق الشهري لـ 500 أسرة في البلدية A:

الانفاق الشهري (10^2) (دج)	[1-5[	[5-9[	[9-13[	[13-17[	[17-21[	[21-25[
عدد الاسر	30	50	60	230	100	30

$$\sum n_i x_{iA} = 7140$$

$$\sum n_i x_{iA}^2 = 113700$$

إذا علمت أن في البلدية B التي تحتوي على نفس عدد الأسر ، الانفاق المتوسط يساوي 16,88 و الانحراف المعياري 4,255.

1. أحسب معامل الاختلاف المعياري للبلدية A

2. أحسب معامل الاختلاف المعياري للبلدية B

3. ما هي البلدية التي إنفاق أسرها الأكثر تجانس؟

التمرين الثاني:

لدينا توزيع إحصائي (X, Y) ممثل في الجدول التالي:

x_i	y_i					
3	-2					
2	0					
2	2					
1	4					

1. أحسب التباين المشترك بين X و Y مع تفسير النتيجة

2. أحسب معامل الارتباط الخطي مع التفسير

3. إذا كان معامل الارتباط الخطي يساوي +1، ما هي نوعية العلاقة الموجودة بين X و Y؟

Test Statistique 2

Durée : 30 min

Sujet : C

التمرين الأول:

تم دراسة رقم أعمال عينة من المؤسسات الصغيرة و المتوسطة لأربع دول و تم حساب متوسطات رقم الأعمال و كذلك انحرافات المعيارية فكانت النتائج كما يلي:

الوحدة: 10^4 دج

الدول	A	B	C	D
متوسط رقم الأعمال (دج)	400	600	740	500
الانحراف المعياري (دج)	20	50	25	30

1. أحسب معامل الاختلاف المعياري
2. ما هي الدولة التي رقم أعمال مؤسساتها أكثر تجانس من غيرها؟

التمرين الثاني: لدينا توزيع إحصائي (Y, X) ممثل في الجدول التالي:

x_i	y_i					
1	4					
2	2					
2	0					
3	-2					

1. أحسب التباين المشترك بين X و Y مع تفسير النتيجة
2. احسب معامل الارتباط الخطي مع التفسير
3. إذا كان معامل الارتباط الخطي يساوي 0 ، ما هي نوعية العلاقة الموجودة بين x و y ؟

Test

Durée : 30min

Sujet : D

التمرين الأول:

تم دراسة قيمة الصادرات لأربع دول خلال فترة معينة و تم حساب متوسطات الصادرات لها وكذلك انحرافاتها المعيارية فكانت النتائج كما يلي:

الوحدة: 10^5 دج

الدول	A	B	C	D
متوسط الصادرات (دج)	600	500	740	400
الانحراف المعياري (دج)	50	30	25	20

1. أحسب معامل الاختلاف المعياري

2. ما هي الدولة التي قيمة صادراتها أقل تشتت من غيرها؟

التمرين الثاني:

لدينا توزيع إحصائي (Y, X) ممثل في الجدول التالي:

x_i	y_i					
4	1					
2	2					
0	2					
-2	3					

--	--	--	--	--	--	--

1. أحسب التباين المشترك بين X و Y مع تفسير النتيجة

2. أوجد معادلة الانحدار Y على X

3. إذا كان $X = 4$ كم سيكون قيمة Y

المراجع

- جلاطو جيلالي، الإحصاء مع تمارين ومسائل محلولة، ديوان المطبوعات الجامعية، الطبعة الثامنة، الجزائر، 2010.

- حللمي عبد القادر، مدخل إلى الإحصاء، ديوان المطبوعات الجامعية، الطبعة الخامسة، الجزائر، 2004 .
- حيدوشي عاشور، محاضرات في الإحصاء الوصفي، مطبوعة موجهة لطلبة السنة الأولى جذع مشترك، جامعة أكلي محمد أولحاج البويرة، الجزائر، 2016.

- خليل شرف الدين، الإحصاء الوصفي، شبكة الأبحاث والدراسات الاقتصادية.
- راتول محمد، الإحصاء الوصفي، ديوان المطبوعات الجامعية، الطبعة الثالثة، الجزائر، 2009 .
- عبيدات محمد وآخرون، منهجية البحث العلمي، القواعد والمراحل والتطبيقات. دار وائل للطباعة والنشر، الطبعة الثانية، عمان، 2006 .

- عوض مراد كمال، أساسيات الإحصاء، دار البداية ناشرون وموزعون، الطبعة الأولى، عمان، 2008
- موساوي عبد النور، بركان يوسف، الإحصاء 1 ، دروس-تمارين محلولة- تطبيقات، دار العلوم للنشر والتوزيع، عنابة، الجزائر.

- جون جاك دروزيك، أساسيات في الإحصاء، سلسلة (SMA)، دار (Ellips)، 1996، ص 2.
- دومينيك سالفاتور، الاقتصاد القياسي والإحصاء التطبيقي، سلسلة شوم، دار ماك فراو هيل، 1985.
- مصطفى الخواجة: مقدمة في الإحصاء، الدار الجامعية، الإسكندرية، 2002،
- فريد بن ختو، محاضرات في الإحصاء الوصفي، 2018/2017 جامعة بسكرة اطو جيلالي، الإحصاء مع تمارين ومسائل محلولة، 2002، ديوان المطبوعات الجامعية.

- فاروق عبد العظيم وآخرون مبادئ الإحصاء دار المعرفة الجامعية 2003.

- عوض منصور مبادئ الاحصاء الوصفي دار صفاء للنشر. الاردن 2000.

- Dominick SALVATORE & Derrick REAGLE, Theory and Problems of Statistics and Econometrics, Second Edition, Schaum's Outline Series, McGRAW-HILL, companies, NewYork, 2002.

- Fabrice MAZEROLLE, Statistique Descriptive, Notes de cours, 2009