الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
جامعة الشهيد حمه لخضر – الوادي
Université Echahid Hamma Lakhdar-El oued
Faculté des sciences économiques, commerciales et sciences de gestion
كلية العلوم الاقتصادية و التجارية و علوم التسيير
المجلس العلمي للكلية

الرقم: ... /م ع/ك ع ا ق ت و ع ت/2022 الـوادي في: 2022/05/26

شهادة اعتماد مطبوعة

يشهد رئيس المجلس العلمي لكلية العلوم الاقتصادية والتجارية وعلوم التسيير بجامعة الشهيد حمه لخضر الوادي على أن:

د/ أحمد بن أحمد .... أستاذ محاضر أ

قدم للمجلس العلمي للكلية عن طريق اللجنة العلمية لقسم العلوم المالية والمحاسبية مطبوعة دروس بعنوان: **محاضرات في الإحصاء 1**، عدد صفحاتها 128 صفحة مقدمة لطلبة سنة أولى جذع مشترك (مسار العلوم الاقتصادية والتجارية وعلوم التسيير)، للموسم الجامعي 2022/2021.

وبعد التقارير الإيجابية المقدمة من طرف لجنة القراءة والتحكيم:

فإنه يتم اعتماد المطبوعة وتعتبر محكمة علميا.

رئيس المجلس العلمي

رئيس المجلس العلمي لكلية العلوم الاقتصادية والتجارية وعلوم التسيير
أ.د/ إلياس شاهد

الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne démocratique et Populaire
وزارة التعليم العالي والبحث العلمي
Ministère de l'Enseignement Supérieure et de la Recherche Scientifique
جامعة الشهيد حمّه لخضر - الوادي
Université Echahid Hamma Lakhdar - El Oued

Faculté de science économique et commerce et science gestion

Département des Sciences financières et comptables
Comité scientifique du Département

كلية العلوم الاقتصادية والتجارية وعلوم التسيير

قسم العلوم المالية والمحاسبية

اللجنة العلمية للقسم

الرقم: **006**/ ل ع ق ع م م /ك.ع. ق.ت.ع.ت/2022

الوادي في:2022/05/18

شهــــادة إداريــــة

(اعتماد مطبوعة)

يشهد السيد رئيس اللجنة العلمية لقسم العلوم المالية والمحاسبية بأن الدكتور: **بن أحمد أحمد** ، قد قدم أمام اللجنة العلمية لقسم العلوم المالية والمحاسبية مطبوعة بعنوان:

محاضرات في الإحصاء 1

مقدمة لطلبة السنة أولى جذع مشترك (مسار العلوم الاقتصادية والتجارية وعلوم التسيير) ، وبعد تعين لجنة الخبرة وبعد استلام تقارير الخبرة الإيجابية، واستيفاء المطبوعة الجوانب الشكلية والمنهجية والشروط المعتمدة في كلية العلوم الاقتصادية والتجارية وعلوم التسيير بجامعة الشهيد حمه لخضر، تم اعتماد المطبوعة للتدريس في المستوى والتخصص المذكور أعلاه (عدد الصفحات 128).

رئيس اللجنة العلمية للقسم

رئيس اللجنة العلمية لقسم العلوم المالية والمحاسبية
أ.د. محمد الهادي ضيف الله

وزارة التعليم العالي والبحث العلمي
جامعة الشهيد حمة لخضر – الوادي

كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير

مطبوعة موجهة للسنة الأولى جذع مشترك

محاضرات في : الاحصاء 1

من إعداد:

د. بن أحمد أحمد   أستاذ محاضر أ

السنة الجامعية: 2021 / 2022

الإهــداء

أهدي هذا العمل المتواضع إلى :

إلى كل أساتذتي من المرحلة الابتدائية

إلى المرحلة الحالية

وإلى كافة طلبة كلية العلوم الاقتصادية التجارية وعلوم التسيير

بجامعة الشهيد حمه لخضر بالوادي.

د. بن أحمد أحمد

الفهرس

يعتبر الإحصاء من العلوم القديمة التي صاحبت الإنسان في تطوره وإدارة شؤونه، فقد كانت فكرة الإحصاء قديما تقوم على فكرة التعداد فقط، لكن سرعان ما ازداد استعماله لما شعرت بعض الدول بحاجتها إلى معرفة بعض البيانات العددية عن عدد سكانها وتكاثرهم وأحوالهم الشخصية ومقدار ثرواتها الزراعية والمعدنية لمعرفة احتياجاتها في حالتي السلم والحرب.

وقد تطور علم الإحصاء من مجرد فكرة الحصر والعد إلى أن أصبح الآن علما له قواعده ونظرياته، حيث اكتسب مفهوما حديثا اعتبارا من القرنين السابع عشر والثامن عشر، من خلال شموله لجميع الطرق العلمية والوسائل والأدوات الحديثة التي تساعد الباحث على الوصول إلى القرار السليم إزاء مشكلات تتحكم فيها العشوائية، وهذا باستخدام مقاييس النزعة المركزية أو مقاييس التشتت أو دراسة الارتباط والانحدار أو تحليل التباين، ويرجع الفضل في ذلك إلى كثير من علماء الإحصاء أمثال فريديريك جاوس Gauss ، كارل بيرسون Karl Pearson الذي لقب بأبي الإحصاء، ويول Yule، وأرنولد فيشر Fisher، سيبرمان Spearman وغيرهم من علماء الإحصاء والرياضيات.

وأصبح الإحصاء في الوقت الحاضر يعالج بشكل رئيسي النواحي الكمية للظواهر الاقتصادية والاجتماعية والسكانية وغير ذلك باستخدام الطرق الإحصائية المناسبة، حيث تمكن هذه الطرق من تبسيط ظاهرة معينة خلال فترات سابقة لإيجاد حلول مناسبة لمشاكل الحاضر والتنبؤ بقيمها في المستقبل.

سنتطرق من خلال هذا المقياس إلى الإحصاء الوصفي الذي يعالج المرحلة الأولى من الطريقة الإحصائية التي تتضمن ثلاث مراحل، المرحلة الوصفية، المرحلة التنبئية ومرحلة إتخاذ القرار.

أولا: مفاهيم أساسية:

كما يعرف علم الإحصاء بأنه علم اتخاذ القرارات في جميع نواحي الحياة، وذلك من خلال جمع ودراسة وتحليل البيانات المتوفرة واستخلاص النتائج عن الظواهر المدروسة مما سبق يمكن تصنيف الإحصاء كعلم إلى قسمين رئيسيين هما:

1 – **الإحصاء الوصفي** **Statistique descriptive:** وهو ذلك الفرع من الإحصاء الذي يتناول طرق تنظيم وتلخيص وعرض البيانات في صورة مبسطة.

2 – **الإحصاء الاستدلالي** **Statistique inférentielle:** وهو ذلك الفرع من الإحصاء الذي يهتم بطرق الوصول إلى نتائج معينة أو توقعات ما عن المجتمع من خلال دراسة عينة من ذلك المجتمع.

فإذا كانت لدينا كمية كبيرة من البيانات العددية، فإن الإحصائي سيحاول أن يرتبها في صورة تجعل من السهل قراءتها وفهمها، وقد يتضمن هذا:

– تبويب البيانات وتقديمها في شكل جداول تكرارية، أو في شكل منحنيات بيانية ليسهل فهم معناها فورا.

– حساب بعض المقاييس أو المؤشرات الإحصائية مثل النسب أو المتوسطات.

وتدخل العمليات السابقة في نطاق الإحصاء الوصفي، أما الإحصاء الاستدلالي فهو يختص بـــ:

– إجراء التنبؤات والتقديرات والاستنتاجات عن مجموعة من المتغيرات أكبر من تلك التي تمت ملاحظتها فعلا.

ثانيا – التعريف بالمصطلحات الإحصائية الضرورية

1 – **المجتمع** Population: هو جميع العناصر المشتركة في الصفة التي تهم الباحث في دراسته، فقد يكون المجتمع مثلا عدد سكان مدينة، أو طلبة جامعة التكوين المتواصل، أو المساحات الزراعية في الجزائر أو إنتاج آلة معينة ... إلخ.

2 – **العينة** Echantillon: هي جزء من المجتمع تحت الدراسة مثل مجموعة من سكان مدينة، أو مجموعة من طلبة جامعة الوادي، أو بعض المساحات الزراعية في الجزائر ...الخ.

3 – **الظاهرة** Phénomène: هي صفة لعناصر تختلف من عنصر لآخر في شكل أو النوع أو الكمية، ويطلق على الصفة تحت الدراسة متغير variable مثل طول شخص ما، عدد الأخطاء الإملائية في بحث ما، سرعة سيارة بين مدنيين خلال أسبوع ... الخ.

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

ثالثا– أنواع المتغيرات

تنقسم المتغيرات إلى نوعين:

1 – **متغيرات نوعية (كيفية)** Variables qualitatives: وهي عبارة عن صفات أو أنواع معينة ليست عددية، وتنقسم بدورها إلى:

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

ب– بيانات نوعية غير خاضعة للترتيب: مثل الجنسية، أنواع السيارات، أنواع الأمراض ...إلخ.

2 – **متغيرات كمية (عددية) Variables quantitatives**: وهي البيانات التي يعبر عنها في صورة عددية وتنقسم إلى:

أ – متغير متقطع Variable discret: وهو المتغير الذي يأخذ أعداد صحيحة، فمثلا 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.

ب – متغير متصل (مستمر) Variable Continue: وهو المتغير الذي لا يمكن أن يأخذ أي قيمة بين قيمتين معنيتين، وكأمثلة عن المتغيرات المتصلة: الطول، الوزن، الزمن، السرعة ...الخ، فإذا كان x هو متغير الطول فمثلا فإن x يمكن أن تأخذ القيم 15متر، 11.3 متر، 14.75 متر، أي أن المتغير x يمكن أن نأخذ أي قيمة في فترة زمنية معينة.

رابعا – تقريب البيانات

يعتمد الإحصاء في كثير من عملياته على التقريب الذي يهدف من ورائه تبسيط العمليات الحسابية حتى يتيسر للباحث معالجتها وتأكيد معالمها الرئيسية، وتساعد القارئ على فهم نتائجها.

تقوم فكرة التقريب على حذف الرقم الذي يبدأ به العدد من اليمين ثم إضافة واحد صحيح إلى الرقم الذي يتبع إلى يسارة مباشرة إذا كان الرقم المحذوف أكبر من 5 أو يترك كما هو دون إضافة الواحد الصحيح إذا كان الرقم المحذوف أقل من 5.

أما إذا كان الرقم المحذوف يساوي 5 فإن الرقم الذي يقع إلى يساره يقرب إلى أقرب عدد زوجي. فإذا كان الرقم زوجيا ظل كما هو.

مثال 1: قرب الأعداد التالية: 1.2، 16.5 ، 23.4،25.5، 15.6، 15.5، 18.7، 28.5 إلى أعداد صحيحة؟

ومن أهم استخدامات التقريب تقريب النسب المئوية والكسور العشرية إلى أقرب عدد صحيح وأثر هذا التقريب على مجموعها النهائي الذي يجب أن يساوي 100 في حالة النسب المئوية، و واحد صحيح في حالة الكسور العشرية.

الفصل الأول: عرض البيانات

أولا – العرض الجدولي للبيانات

تلخص الجداول التكرارية البيانات الكمية الكثيرة في وضعها على صورة جدول منتظم يوضح كيفية توزيع القيم التي حصلنا عليها من الظاهرة المدروسة حيث يدل العمود (السطر) الأول على قيم الظاهرة، ويدل العمود (السطر) الثاني على التكرار المقابل لهذه القيم.

<u>مثال1</u>: إذا كانت لدينا درجات 30 طالب في أحد الاختبارات كما يلي:

4، 8، 7، 6، 4، 11، 12، 4، 6، 8، 4، 6، 9، 6، 7، 6، 8، 4، 11، 9، 8، 11 ، 9، 4، 8، 9، 11، 12، 4، 5.

وفي حالة ما إذا كنت الدرجات كثيرة ومنتشرة داخل مجال واسع (مثلا التنقيط كان من 100) فإنه توجد طريقة أخرى أكثر اختصارا من السابقة يمكن بواسطتها وضع البيانات في جدول يبين ويوضح الخصائص العامة لهذه البيانات، يسمى هذا الجدول بجدول التوزيع التكراري ولتكوين مثل هذا الجدول نقوم بالتالي:

أولا: نحدد المجال (المدى) الذي تنتشر فيه البيانات، وهو الفرق بين أكبر قيمة للبيانات وأصغر قيمة لها، أي أن:

المدى = أكبر قيمة – أصغر قيمة

ثانيا: نقسم المدى إلى فئات متساوية الطول بحيث يكون عددها مناسبا و هناك عدة طرق لحساب عدد الفئات نذكر منها:

1) معادلة ستيرجس Sturgers التي تنص:

$$\text{عدد الفئات} = 1 + 3.322\log\left(\text{عدد البيانات}\right)$$

2) معادلة يول yule التي تنص على: $\text{عدد الفئات} = 2.5\sqrt[4]{\text{عدد البيانات}}$

ثالثا: نحسب طول الفئة وهو يساوي المدى مقسوما على عدد الفئات $\boxed{= \text{طول الفئة}}$

$$\frac{\text{المدى}}{\text{عدد الفئات}}$$

مما سبق نستخلص الطريقة المرنة في تحديد عدد الفئات وأطوالها والتي لا تعتمد على المعدلات الرياضية بل أن هذه الطريقة مرنة بطبيعتها وهي: $\boxed{\times \text{طول الفئة}}$

$$\boxed{\text{المدى} \geq \text{عدد الفئات}}$$

ملاحظـــات:

I. عند تفريغ البيانات فإنه يجب أن تنتمي كل مفردة إلى فئة واحدة فقط.

II. عند كتابة الفئات فإنه:

- يذكر الحد الأدنى والأعلى لكل فئة(فئات مغلقة)إذا كان المتغير متقطع.
- يذكر الحد الأدنى ويحدد الحد الأعلى ضمنيا أو العكس (فئات مغلقة من جهة و مفتوحة من الجهة الاخرى) إذا كان المتغير متصل.
- يفضل استخدام الفئات المتساوية الطول، إلا أنه في بعض الحالات يمكن أن يستخدم الفئات غير المتساوية، من هذه الحالات ما يلي:

– إذا كان الغرض من الدراسة هو الاهتمام ببعض الفئات والتركيز عليها وإهمال باقي الفئات، فيمكن عندها دمج الفئات التي لا تهم الباحث في فئة واحدة.

– إذا كان التكرار لبعض الفئات صغير جدا مقارنة بباقي الفئات، يمكن دمج هذه الفئات معا.

<u>مثال 2:</u> البيانات التالية تمثل إنتاج 60 ورشة من الكراسي خلال يوم.

87	93	21	46	57	77	62	25	31	72
68	29	89	66	62	73	83	81	72	54
57	81	58	62	73	12	73	83	88	96

67	87	97	52	63	17	29	36	71	63
62	92	73	57	65	71	36	54	21	33
42	58	89	46	49	36	56	62	51	91

المطلوب:1) ما هو نوع المتغير؟ ولماذا؟.

2) أي نوع من الفئات تستخدم في مثل هذه الحالة؟

3) تحديد عدد الفئات باستخدام معادلة Sturgers؟

4) تحديد عدد الفئات باستخدام معادلة Yule؟

5) تكوين جدول تكراري من 10 فئات متساوية الطول؟.

<u>مثال 3</u>: إذا كانت بيانات المثال السابق تمثل إنتاج الحليب باللترات في يوما ما بـــ 60 مزرعة أعد الإجابة على الأسئلة المطروحة؟.

العرض الجدولي في حالة البيانات النوعية

يتكون الجدول من عمودين (سطرين) يحتوي العمود (السطر) الأول على رموز كتابية للخاصية المدروسة (صفات) أما الثاني فيحتوي على تكرارات كل رمز كتابي.

<u>مثال4</u>: أخذت عينة من طلبة جامعة الوادي متكونة من 100 طالب، حيث كانت الخاصية المدروسة هي الانتماء إلى كلية من كليات الجامعة فكانت النتائج كما هي مبينـــة فـــي الجدول أدناه

جدول توزيع مجموعة من طلبة الوادي على كليات الجامعة

الكلية	الحقوق	الاقتصاد	الآداب	العلوم	المجموع
عدد الطلبة	35	30	25	10	100

التكـــرارات المتجمعـــة

في بعض الحالات نرغب في معرفة التكرارات أو البيانات التي تزيد قيمتها عن قيمة معينة أو تقل عن قيمة معينة، فمثلا عندما نرغب في معرفة عدد النـــاجحين (أي الطلبة المتحصلين على درجة تساوي أو تزيد 10) فإن هذه المعلومات غير واضحة في جدول التوزيع التكراري فنكون لهذا الغرض ما يسمى بالجـــدول التكـــراري المتجمـــع الصاعد أو النازل، وكتعريف فإن:

– التكرار المتجمع الصاعد لأي فئة هو تكرار هذه الفئة مضافا إليه مجموع تكرارات الفئات السابقة.

– التكرار المتجمع النازل لأي فئة هو عبارة عن مجموع التكرارات مطروحا منه تكرارات الفئات السابقة.

التكرار النسبي

هو وضع تكرار كل فئة كنسبة من التكرار الكلي، وهذه الطريقة لها عدة استخدامات وفوائد حيث توضح نسبة تكرار كل فئة إلى التكرار الكلي.

مثال 5: الجدول الآتي يبين توزيع دخول عينة من عمال مؤسسة صناعة الكوابل الكهربائية حسب دخولهم بآلاف الدينارات. المطلوب حساب التكرار المتجمع الصاعد والتكرار المتجمع النازل و التكرار النسبي لهذه البيانات.

فئة الدخل	10-15	15-20	20-25	25-30	30-35	35-40	**المجموع**
عدد العمال	4	6	12	8	6	4	**40**

ثانيا – العرض البياني للبيانات

بالإضافة إلى العرض الجدولي للبيانات هناك طريق أخرى تستخدم لتوضيح وتلخيص البيانات وهي طرق العرض البياني. وهذه الطرق أسهل وأبسط من سابقتها، فإذا كان العرض الجدولي لا يفهمه إلا المختصين فإن طرق العرض البياني يمكن أن تجلب غير المتخصصين، كما تمكن من القيام بتحليل سريع للظاهرة المدروسة. وتستخدم أنواع مختلفة للعرض البياني حسب نوعية المتغير المدروس.

1) المتغير الكمي المتقطع: وهنا يمكن استخدام الأعمدة البسيطة، العرض البياني للتكرارات المتجمعة الصاعدة والنازلة.
2) المتغير الكمي المتصل: هنا يمكن استخدام المدرج التكراري، المضلع التكراري، المنحنى البياني للتكرارات المتجمعة الصاعدة والنازلة.
3) المتغير النوعي: وهنا يمكن استخدام الدوائر، الأعمدة المجزأة، الأعمدة المستطيلة.

I – العرض البياني في حالة متغير كمي متقطع:

الأعمدة البسيطة (Diagrammes en bâtons): هي عبارة عن أعمدة بسيطة تتناسب أطوالها مع التكرار المقابل لقيمة العينة للمتغير المدروس

<u>مثال 6:</u> يبين الجدول التالي عدد الأطفال في العائلة لعينة تكون من 125 أسرة، المطلوب عرض هذه البيانات بطريقة الاعمدة البسيطة.

عدد الأطفال في كل أسرة	0	1	2	3	4	5	6	7	8	المجموع
عدد الأسر	6	9	10	14	26	20	25	15	10	125

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

إن العروض البيانية للمتغير الإحصائي الكمي المستمر هي أكثر العروض البيانية استعمالا ومن أهمها:

1 – المدرج التكراري Histogramme:

وهو عبارة عن مستطيلات متلاصقة طول كل منها يتناسب مع التكرار المقابل، وقاعدة كل منها تساوي طول الفئة المقابلة، ويمكن أن نميز بين حالتين عند رسم المدرج التكراري:

أ) الحالة الأولى: عندما تكون الفئات متساوية.

مثال7: يبين الجدول التالي توزيع الدرجات التي حصل عليها 50 طالب في مادة الإحصاء. أرسم المدرج التكراري الذي يمثل توزيع الدرجات؟.

الفئة	0-4	4-8	8-12	12-16	16-20	**المجموع**
التكرار	8	12	20	6	4	**50**

ب) الحالة الثانية: عندما تكون الفئات غير متساوية.

إذا كانت هناك فئات التوزيع غير متساوية نقوم بتعديل التكرارات، حتى يكون هناك تناسب بين طول الفئة والتكرار المقابل لها، وقيم تعديل التكرارات باستخدام المعادلة الآتية.

$$\text{التكرار المعدل} = \frac{\text{تكرار الفئة}}{\text{طول الفئة}} \times \text{طول الفئة المختار}$$

مثال8: يبين الجدول التالي توزيع عينة من 100 عامل حسب الأجر اليومي المطلوب تمثيل هذه البيانات باستخدام المدرج التكراري؟.

فئة الأجر	200-250	250-350	350-400	400-550	550-750	750-800	**المجموع**
عدد العمال	5	15	20	25	30	5	100

2 – **المضلع التكراري** Polygone de fréquence:

هو مجموعة من القطع المستقيمة متصلة ومنكسرة تتحدد بنقاط أحداثياتها مركز الفئة والتكرارات المقابلة.

$$\text{مركز الفئة} = \frac{\text{الحد الأعلى للفئة} + \text{الحد الأدنى للفئة}}{2}$$

مثال9: ليكن التوزيع التكراري الآتي، أرسم المدرج التكراري والمضلع التكراري؟.

الفئة	2-4	4-6	6-8	8-10	10-12	12-14	**المجموع**
التكـــرار	4	8	16	12	6	2	48

ملاحظة: المساحة التي تقع تحت المضلع التكراري تساوي المساحة التي تقع تحت المدرج التكراري، وحتى نحافظ على المساحة التي تقع أسفل المضلع التكراري، نفترض أن لهذا التوزيع فئات إحداهما في بدايته والأخرى في نهايته تكرار كل منهما يساوي صفر، بحيث ننطلق في رسم المضلع من مركز الفئة الافتراضية الأولى (الفئة ما قبل الأولى)، وننتهي عند مركز الفئة الافتراضية الأخيرة.

3 – **منحنى التكرارات المتجمعة الصاعدة والنازلة:**

يرسم منحنى التكرار المتجمع الصاعد عن طريق إيصال مجموعة النقاط ذات الإحداثيات التالية: الحدود العليا للفئات والتكرار المتجمع الصاعد المقابل لها، ويرسم منحنى التكرار المتجمع النازل بإيصال مجموعة النقاط التي إحداثياتها: الحدود الدنيا للفئات والتكرار المتجمع النازل مقابل لها.

مثال10: أرسم على نفس المعلم كل من منحنى التكرار المتجمع الصاعد ومنحنى التكرار المتجمع النازل لبيانات التكراري للمثال السابق؟.

III – العرض البياني في حالة متغير نوعي:

1) **العرض الدائري** Diagramme circulaire:

ويتمثل في دائرة مقسمة إلى عدة أجزاء كل جزء يقابل زاوية مركزية تتناسب مع التكرارت المقابلة لكل خاصية من الخصائص المدروسة، ولتحقيق ذلك نضيف عمودا إلى جدول المعطيات يحتوي على الزوايا المركزية المقابلة لكل تكرار.

مثال11: يبين الجدول التالي عدد طلبة كلية العلوم الاقتصادية و التجارية و علوم التسيير سنة 2011 مقسمين على أقسام الكلية المختلفة.

القسم	تجارة	إقتصاد	تسيير	سنة اولي ل. م. د	**المجموع**
عدد الطلبة	120	130	250	520	**1020**

المطلوب: عرض البيانات باستخدام القطع الدائرية؟.

ملاحظة: تحسب الزوايا المركزية المقابلة لكل تكرار بالطريقة التالية:

$$\text{الزاوية المركزية} = \frac{\text{تكرار الخاصية}}{\text{مجموع التكرارات}} \times 360$$

2) العمود المجزأ Diagrammes en Barres: وهو عبارة عن مستطيل مقسم إلى عدة أجزاء كل جزء يقابل تكرار معين للخاصية المدروسة.

مثال12: أعرض بيانات المثال السابق باستخدام العمود المجزأ.

3) الأعمدة المستطيلة: وهو عبارة عن مجموعة من الأعمدة المتجاوزة ذات القواعد المتساوية إلا أن ارتفاعها تتناسب مع تكرار كل خاصية، كما أن هذه الأعمدة تكون متباعدة بمسافات متساوية.

مثال13: أعرض بيانات المثال السابق باستخدام المستطيلة

سلسلة تمارين حول: عرض البيانات

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.
- أوزران طلبة جامعة الوادي.
- الرياضة الممارسة من طرف طلبة كلية ع الاقتصادية والتجارية وع التسيير بجامعة الوادي.
- 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.

التمرين 02: في محاولة لدراسة حالة التعليم كانت نتائج البحث الميداني الذي خص 30 شخص كما يلي:

ماجيستير	بكالوريا	دكتوراة	ماجيستير	دكتوراة	ليسانس	ليسانس	بكالوريا	بكالوريا	ليسانس
دكتوراة	بكالوريا	بكالوريا	بكالوريا	بكالوريا	ماجيستير	ليسانس	تقني	ماجيستير	ليسانس
ليسانس	تقني	تقني	تقني	ماجيستير	دكتوراة	ماجيستير	ليسانس	تقني	تقني

1) حدد المتغير الإحصائي ونوعه؟
2) لخص هذه البيانات في جدول إحصائي ثم أوجد نسب التمثيل لكل صفة في العينة المدروسة؟
3) أنشئ التمثيل البياني المناسب للجدول؟

التمرين 03: البيانات التالية تبين عدد الغيابات التي سجلها عمال مؤسسة ما خلال الثلاثي الأول من السنة:

4	3	5	2	3	2	2	2	5	3
5	5	4	0	0	5	1	2	5	0
1	1	5	3	0	2	3	2	1	4

1) حدد المجتمع الإحصائي والمتغير الإحصائي ونوعه؟

2) لخص هذه البيانات في جدول إحصائي؟ ثم أنشئ التمثيل البياني المناسب للجدول؟

3) أحسب عدد ونسبة العمال الذين لديهم أقل من 3 غيابات؟

التمرين 04: البيانات التالية تمثل مبيعات كشك معين من جريدة الخبر خلال 50 يوم.

36	45	31	28	41	32	29	26	48	52
30	28	33	27	40	31	30	40	35	45
32	37	36	31	35	33	38	39	41	30
38	42	35	20	36	38	39	37	45	23
36	46	30	35	32	51	43	37	22	34

1) شكل جدول توزيع تكراري المناسب؟ ثم أنشئ التمثيل البياني لهذا الجدول؟

التمرين 05: لدينا البيانات التالية والتي تمثل الأجر اليومي لأربعين عاملا بالدينار:

370	250	140	250	490	300	490	170	180	370
450	560	150	600	510	140	220	270	190	450
230	460	640	370	450	260	340	170	320	150
600	160	270	300	700	100	490	190	160	690

1) لخص هذه البيانات في جدول إحصائي؟ ثم أنشئ التمثيـــل البيـــاني المناســـب للجدول؟

2) أحسب عدد و نسبة العمال الذين تساوي أو تزيد أجورهم عن 200 دينار؟

التمرين 06: الجدول التالي يمثل توزيع الأجور الأسبوعية لعمال مؤسسة ما:

الفئات	التكرارات	التكرارات النسبية
0−10	n_1	0.08
10−20	08	−
20−30	12	−
30−50	14	−
50−60	n_5	−
المجموع	50	1

1) أكمل الجدول ثم أحسب نسبة العمال الذين تساوي أو تزيد أجورهم عن 10 ولكنها تقل عن 50 ؟

<u>التمرين 07</u>: الجدول التالي يمثل توزيع عمال مؤسسة ما حسب التخصص:

الشهادة	التكرارات	التكرارات النسبية
تقني	55	0.44
تقني سامي	n_2	0.24
ليسانس	25	0.20
مهندس	n_4	f_4
المجموع	N	1

1) برهن على أن مجموع التكرارات النسبية يساوي 1.

2) استنتج كل من N ، n_4 ، n_2 و f_4 ؟

التمرين 08: **فيما يلي بيانات عن المستوى التعليمي لعينة من 50 فرد.**

متوسط	يقرأ ويكتب	ثانوي	متوسط	ثانوي	أعلى من جامعي	متوسط	ابتدائي
يقرا ويكتب	متوسط	ثانوي	ثانوي	متوسط	ثانوي	ابتدائي	متوسط
ابتدائي	ثانوي	يقرا ويكتب	جامعي	ثانوي	ابتدائي	يقرا ويكتب	ثانوي
متوسط	ابتدائي	متوسط	ثانوي	ابتدائي	متوسط	جامعي	متوسط
ثانوي	متوسط	ابتدائي	ثانوي	يقرا ويكتب	ابتدائي	ثانوي	ابتدائي
جامعي	ثانوي	جامعي	ابتدائي	جامعي	أعلى من جامعي	ثانوي	ثانوي
متوسط	يقرا ويكتب						

والمطلوب: **1**- اعرض البيانات في شكل جدول تكراري.

2-كون التوزيع التكراري النسبي، ثم علق على النتائج.

التمرين 09:

فيما يلي بيانات درجات **70** طالب في الاختبار النهائي لمقــرر مـــادة الإحصـــاء التطبيقي.

56	65	70	65	55	60	66	70	75	56
60	70	61	67	61	71	67	62	71	66
68	72	57	68	72	69	57	71	69	75
72	62	67	73	58	63	66	73	63	65
58	73	74	76	74	80	81	60	74	58
76	82	77	83	77	85	91	78	94	72
79	64	57	79	55	87	64	88	78	62

والمطلوب:

1- كون التوزيع التكراري لدرجات الطلاب.

2- كون التوزيع التكراري النسبي.

3- ما هو نسبة الطلاب الحاصلين على درجة ما بين **70** إلى أقل من **80**؟

4- ما هو نسبة الطلاب الحاصلين على درجة أقل من **70** درجة؟

5- ما هو نسبة الطلاب الحاصلين على درجة **80** أو أكثر ؟

التمرين 10:

فيما يلي التوزيع التكراري لأوزان عينة من الدواجن بالجرام، حجمها **100** اختيرت من أحد المزارع بعد **45** يوم.

الوزن	600-	620-	640-	660-	680-	700-720	**المجموع**
عدد الدجاج	10	15	20	25	20	10	100

والمطلوب:

1- ما هو طول الفئة؟

2- ارسم المدرج التكراري.

3- ارسم المدرج التكراري النسبي، ثم علق على الرسم.

التمرين 11:

الجدول التكراري التالي يبين توزيع **40** بقرة في مزرعة حسب كمية الألبان التي تنتجها البقرة في اليوم باللتر.

كمية الألبان	18-	22-	26-	30-	34-38	**المجموع**
عدد الأبقار	4	9	15	8	4	40

والمطلوب:

1- كون جدول التوزيع التكراري المتجمع الصاعد.

2- كون جدول التوزيع التكراري المتجمع الصاعد النسبي.

3- ارسم المنحنى التكراري المتجمع الصاعد النسبي.

4- من المنحنى المتجمع أوجد الآتي:

- نسبة الأبقار التي يقل إنتاجها عن **28** لتر.
- كمية الإنتاج التي يقل عنها **25%** من الأبقار.
- كمية الإنتاج التي يقل عنها **50%** من الإنتاج.

التمرين 12:

الجدول التكراري التالي يبين توزيع عينة حجمها **500** أسرة حسب المنطقة التي تنتمي إليها.

المنطقة	الرياض	الشرقية	القصيم	الغربية	**المجموع**
عدد الأسر	150	130	50	170	500

المطلوب:

مثل البيانات أعلاه في شكل دائرة بيانية.

الفصل الثاني: مقاييس النزعة المركزية

تمهيد:

عند التمعن في الظواهر التي حولنا والقيم التي تأخذها نلاحظ أن غالبية هذه القيم نقترب من بعضها البعض وتتجمع حول قيمة معينة غير منظورة، فذكاء أو طول أو وزن مجموعة من الأشخاص مثلا وتتجمع حول قيمة معينة متوسطة، والقليل من الأشخاص لهم ذكاء أو طول أو وزن يبتعد كثيرا عن هذه القيمة من ناحية الصغر أو الكبر. سميت هذه الظاهرة بظاهرة النزعة المركزية والتي لها عدد من المتوسطات للتعبير عنها تختلف باختلاف الغرض الذي تستخدم فيه، وطبيعة البيانات المحسوبة منها، أهم هذه المتوسطات: المتوسط الحسابي، المتوسط الهندسي، المتوسط التوافقي، الوسيط، المنوال والتي سنتطرق لها في الصفحات القادمة.

وميزة هذه المتوسطات كقيم عددية وحيدة توفر لنا فكرة عامة عن البيانات، وتصف الظاهرة المدروسة مثل الجداول الإحصائية، إلا أنها أكثر اختصارا وأكثر فائدة، حيث تمكننا من المقارنة بين مجموعة من القيم ومجموعة أخرى، أو بين ظاهرة وأخرى.

أولا – المتوسط الحسابي Moyenne arithmétique

يعتبر المتوسط الحسابي من أسهل وأكثر متوسطات النزعة المركزية اســتخداما في الإحصاء، هو عبارة عن مجموع القيم مقسوما على عددها.

فإذا كانت لدينا القيم $x_1,x_2,x_3,.........x_n$ فإن متوسطها الحسابي يساوي

$$\overline{X} = \frac{x_1 + x_2 + x_3 + ...+ x_n}{n} = \frac{\sum_{i=1}^{n} Xi}{n}$$

حيث: $\overline{X}$= المتوسط الحسابي

X_i = تمثل قيم الظاهرة

n = تمثل عدد البيانات

<u>مثال1:</u>

إذا كانت الدرجات التي تحصل عليها الطالب في خمس مواد هي: 8، 10، 13، 14، 15.

أحسب متوسط درجات هذا الطالب؟.

الحل:

$$\text{المتوسط} = \frac{\text{مجموع الدرجات}}{\text{عدد الدرجات}}$$

$$\overline{X} = \frac{\sum Xi}{n} = \frac{8+10+13+14+15}{5} = \frac{60}{5} = 12$$

المتوسط الحسابي في حالة بيانات مكررة

إذا كانت لدينا القيم

$x_1, x_2, x_3, \ldots\ldots\ldots\ldots\ldots x_4$

ولها تكرارات

$n_1, n_2, n_3, \ldots\ldots\ldots\ldots\ldots n_4$

فإن المتوسط الحساب لها يعطي بالعلاقة

$$\overline{X} = \frac{n_1x_1 + n_2x_2 + n_3x_3 + \ldots\ldots\ldots\ldots n_nx_n}{n_1 + n_2 + n_3 + n_4 + \ldots\ldots\ldots\ldots n_n} = \frac{\sum_{i=1}^{n} n_ix_i}{\sum n_i}$$

أي أن المتوسط الحسابي لبيانات متكررة يساوي مجموع حاصل ضرب كل قيمة في تكرارها على مجموع التكرارات

وحيث: X_i = تمثل قيم الظاهرة.

n_i = تكرار مل قيمة.

Σn_i = مجموع التكرارات.

<u>مثال2:</u>

في امتحان فجائي في مادة الإحصاء الوصفي تحصل طلبة فوج معين على الدرجات المبينة في الجدول التالي:

الدرجة	3	4	5	6	8	9
عدد الطلبة	4	3	6	5	4	2

المطلوب: حساب متوسط الدرجات التي تحصل عليها طلبة هذا الفوج؟.

الحل:

لحساب هذا المتوسط فإننا نقوم أولا بضرب كل قيمة في تكرارها ثم نطبق العلاقة التي تحسب المتوسط الحسابي لبيانات متكررة، وسنقوم بذلك من خلال الجدول التالي:

الدرجة x_i	عدد الطلبة n_i	$n_i\,x_i$
3	4	12
4	3	12
5	6	30
6	5	30
8	4	32
9	2	18
المجموع	**24**	**134**

$$\overline{X} = \frac{\sum n_i x_i}{\sum n_i} = \frac{134}{24} = 5{,}58$$

أي أن متوسط درجات طلبة هذا الفوج الحسابي في الامتحان الفجائي يساوي 5.58.

المتوسط الحسابي في حالة توزيع تكراري:

تعتمد طريقة حساب المتوسط الحسابي لبيانات مبوبة على مراكز الفئات التي يفترق أنها تمثل الفئات التي أخذت منها، ويكون المتوسط الحسابي في هذه الحالة يساوي إلى مجموع حاصل ضرب مراكز الفئات في تكرارها على مجموع التكرارات

$$\overline{X} = \frac{\sum_{i=1}^{n} n_i x_i}{\sum_{i=1}^{n}}$$

حيث: x_i = تمثل مراكز الفئات.

n_i = تكرار الفئات.

Σn_i = مجموع التكرارات.

<u>مثال3:</u>

في دراسة إحصائية حول مادة الحليب بالمزارع الموجودة على مستوى ولاية سوق أهراس توصلنا إلى إعداد الجدول التالي:

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

المطلوب: إيجاد متوسط إنتاج الحليب بهذه المزارع

<u>الحل:</u>

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	240	1360
400-360	2	380	760
المجموع	25	/	7180

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

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

حساب المتوسط الحسابي بالطريقة المختصرة:

عندما تكون البيانات الموجودة في الجداول المعطاة كبيرة القيم فإن استخدام الطريقة السابقة يصبح صعب، كما يزداد احتمال الوقوع في الأخطاء بذلك فإنه في مثل

هذه الحالات يفضل استخدام طرقة مختصرة الهدف منها تبسـيط العمليـات الحسـابية الطويلة حتى يسهل التعامل معها.

فإذا قمنا مثلا بطريقة قيمة ثابتة (a) من جميع القيم (جميع مراكز الفئات) فإن المتوسط الحساب يصبح

$$\overline{X} = a + \frac{\sum_{i=1}^{n}(x_1 - a)}{n}$$

إذا كانت البيانات مفردة أو إذا كانت البيانات مبوبة في جداول توزيع التكراري

$$\overline{X} = a + \frac{\sum_{i=1}^{n} n_1(x_i - a)}{\sum n_i}$$

مثال4:

أحسب متوسط إن تاج مادة الحليب في المثال السابق بإتباع الطريقة المختصرة؟

الحل:

نختار وسط فرضي a=300 وهو مركز الفئة الوسطى ونتبع الخطوات التالية:

1. توجد قيم جديدة y_i = والتي تساوي مراكز الفئات ناقص الوسط الفرض.
2. نضرب هذه القيم الجديدة في تكرار الفئات n_i.
3. نحسب $\overline{Y}$ وهو المتوسط الحسابي لهذه القيم الجديدة.
4. نحسب المتوسط الحسابي للقيم الجديدة والذي يساوي $\overline{X} = \overline{Y} + a$

الإنتـــــاج باللترات	عدد المزارع n_i	x_i	y_i	$n_i\ x_i$
240−200	5	220	−80	−400
280−240	6 8	260 300	−40	−240

0	0	240	4	320−280
160	40	380	2	360−320
160	80			400−360
−320			25	المجموع

$$\overline{Y}=\frac{\sum_{i=1}^{n}n_1y_i}{\sum_{i=1}^{n}n_1}=\frac{-320}{25}=-12{,}8$$

$\overline{X}$ =300 − 12.8= لتر 287.2

المتوسط الحسابي المرجح (الموزون) Moyenne Arithmétique pondérée

في بعض الأحيان فإن القيم المراد حساب المتوسط الحسابي لها لا تكـون لهـا نفس الأهمية بل أهميات نسبية مختلفة تختلف باختلاف عامل الترجيح الخاص بها. في مثل هذه الحالات فإن المتوسط الحسابي البسيط يمكن الاعتماد عليه في إيجاد المتوسط الصحيح والمنطقي، بل يتطلب الأمر استخدام صيغة أخرى تسمى

$$\text{المتوسط الحسابي المربح} = \frac{\text{مجموع حاصل ضرب مل قيمة في معاملها}}{\text{مجموع المعاملات}} = \overline{X}=\frac{\sum_{i=1}^{n}n_ix_i}{\sum_{i=1}^{n}x_i}$$

بصيغة المتوسط الحسابي المرجح:

حيث: x_i = تمثل القيم.

n_i = تمثل المعاملات.

ويستخدم المتوسط الحسابي المرجع كذلك لإيجاد المتوسط الحسابي لمجموع البيانات أو أكثر في حالة دمجهم معا في مجموعة واحدة وبالتالي فإن متوسط الحسـابي المـرجح لمجموعتين من البيانات y,z يساوي:

$$\overline{X}=\frac{\sum_{i=1}^{n}z_i+\sum_{i=1}^{n}y_i}{n+Z}$$

$$\sum^{n}z_i=n\overline{Z}$$

وبما أن

فإن

مثال5:

الجدول التالي يبين عدد العمال ومتوسط الأجر للعامل الواحد في الواحدات 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 مقسوما على ثلاثة فإن الإجابة تكون خاطئة فالإجابة الصحيحة هي تلك التي يمكن الحصول عليها من خلال علاقة المتوسط الحسابي المرجح.

$$\overline{X} = \frac{n_1\overline{x}_1 + n_2\overline{x}_2 + n_3\overline{x}_3 \ldots\ldots\ldots\ldots + n_1\overline{x}_1}{n_1 + n_2 + n_3 + \ldots\ldots\ldots\ldots + n_4}$$

$$\overline{X} = \frac{(130x13000) + (110x14500) + (80x18500)}{130+110+80} = \frac{1690000+1595000+1480000}{320}$$

$$\overline{X} = \frac{4765000}{320} = 14890{,}62 =$$ متوسط أجر عمال الشركة

خواص المتوسط الحسابي:

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

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

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

وللتأكد من ذلك نقوم بتفكيك الطرف الأيمن:

$$\sum_{i=1}^{n}(x_i - \overline{X})$$
$$= \sum X_i - \sum X$$
$$= \sum X_i - n\overline{X}$$

(لأن :

$$\left(n\overline{X} = ...\overline{X} + \overline{X} + \overline{X}\right)$$

نضرب الحد الأول في n ونقسمه على n

$$= \frac{n\sum X_i}{n} - n\overline{X}$$
$$= n\overline{X} - n\overline{X} = 0$$

4 – يتأثر المتوسط الحسابي بالقيم المتطرفة أو الشاذة.

مثال6:

إذا تحصل طالب على الدرجات التالية في خمس مواد: 20.9.8.7.6

فإن المتوسط درجاته = $\frac{20+9+8+7+6}{5} = \frac{50}{5} = 10$

أي أن الطالب ناجح، وفي الواقع فإن الطالب ناجح في مادة واحدة وراسب في أربع مواد، وظهور الطالب ناجح رغم رسوبه في أغلب المواد، راجع إلى أن المتوسط الحسابي قد تأثر بالنتيجة الأخيرة والتي تعتبر متطرفة في الكبر.

5 – لا يمكن إيجاد المتوسط الحسابي من الجداول التكرارية المفتوحة من البداية أو النهاية.

ثانيا: المتوسط الهندسي: Moyenne géométrique

في الحالات التي تكون فيها قيم الظاهرة المدروسة عبارة عن نسب أو معدلات، أي في الحالات التي نرغب فيها بدراسة معدل تغيرات ظاهرة ما فإن المتوسط الحسابي لن يصف هذه الظاهرة الوصف السليم، ولن يعطي أي فكرة صحيحة عن مثل هذه الظاهرة ولهذا دعت الضرورة إلى إيجاد متوسط آخر يصلح لوصف مثل هذه الظواهر سمي هذا المتوسط بالمتوسط الهندسي.

والمتوسط الهندسي واسع الاستخدام في الحياة الاقتصادية (كما سنرى لاحقا) لأن التركيز يكون غالبا منصبا على إيجاد متوسط نسب التغير لبعض الظواهر مثل: تطور الدخول، زيادة الأجور، والنمو السكاني ... إلخ.

تعريف:

إذا كانت لدينا القيم $x_1, x_2, x_3, x_4, \ldots\ldots x_n$ فإن متوسطها الهندسي يساوي الجذر النوني لحاصل ضرب هذه القيم في بعضها البعض.

$$G = \sqrt[n]{x_1.x_2x_3.....x_n}$$

وإذا كانت هذه القيم متكررة فإن:

$$G = \sqrt[N]{x_1^{n_1}.x_2^{n_2}.x_3^{n_3}.....x_3^{n_4}}$$

حيث $\Sigma ni = N$ (مجموع التكرارات)

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

$$G = \sqrt[N]{x_1^{n_1}.x_2^{n_2}.x_3^{n_3}.....x_3^{n_4}}$$

حيث x_i = مراكز الفئات.

$\Sigma ni = N$ (مجموع التكرارات).

مثال7:

أوجد المتوسط الهندسي للأعداد: 2، 4، 5، 6.

الحل: $G = \sqrt[4]{2\times4\times5\times6} = \sqrt[4]{240} = 3,94$

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

في حالة بيانات مفردة: $G = \sqrt[n]{x_1.x_2.x_3.x_4\ldots x_n}$

$$LogG = \frac{1}{N}\left[\log x_1 + \log x_2 + \log x_3 + \ldots\ldots\ldots + \log x_4\right]$$

$$\mathbf{logG} = \frac{\mathbf{1}}{\mathbf{N}}\sum_{\mathbf{i=1}}^{\mathbf{n}}\mathbf{logx_1}$$

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

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

$$LogG = \frac{1}{N}\left[n_1\log x_1 + n_2\log x_2 + n_3\log x_3 + \ldots\ldots\ldots + n_n\log x_4\right]$$

$$\mathbf{logG} = \frac{\mathbf{1}}{\mathbf{N}}\sum_{\mathbf{i=1}}^{\mathbf{n}}\mathbf{n_i logx_1}$$

حيث: x_i تمثل القيمة أو مركز الفئة.

n_i تمثل التكرار.

مثال8:

أوجد المتوسط الهندسي للبيانات المبينة في الجدول أدناه باستخدام طريقة اللوغاريتم

القيم	2	4	5	6
التكرار	2	3	2	4

الحل:

$$LogG = \frac{1}{N}\sum_{i=1}^{n} Ni \log X_i$$

$$= \frac{1}{11}[2\log 2 + 3\log 4 + 2\log 5 + 4\log 6]$$

$$= \frac{1}{11}[(2\times 0,0301) + (3\times 0,601) + (2\times 0,699) + (4\times 0,778)]$$

$$= \frac{1}{11}[0,602 + 1,806 + 1,398 + 3,112]$$

$$= \frac{1}{11}6,918$$

$$\text{LogG} = 0,629$$

ومنه المتوسط الهندسي = 4.24 = G

استخدام المتوسط الهندسي في الحياة الاقتصادي:

لتباين كيفية استخدام المتوسط الهندسي في الحياة الاقتصادية نأخذ المثال التالي:

الجدول التالي يبين تطور إنتاج مؤسسة ما خلال الفترة 2000–2003 .

السنة	2000	2001	2002	2003
الإنتاج	1000	1250	1875	2625

المطلوب :

5. إيجاد نسبة زيادة الإنتاج من سنة إلى أخرى؟
6. إيجاد متوسط نسبة الزيادة خلال الفترة؟

الحل:

السنة	200	2001	2002	2003
الإنتاج	1000	1250	1875	2625
نسبة الزيادة	–	0.25	0.50	0.40

إذا رمزنا لمتوسط نسبة الزيادة خلال الفترة بالرمز (t) و لنسبة الزيادة خــلال السنوات الأولى والثانية والثالثة بالرموز (t_1)، (t_2) و (t_3) فإنه حتى يكون (t) ممــثلا فعلا لمتوسط نسبة الزيادة خلال الفترة لابد من تحقق الشــرط التــالي: (1+t) (1+t) (1+t) = ($1+t_1$) ($1+t_2$) ($1+t_3$).

ومنه $(1+t)^3 = (1+t_1)(1+t_2)(1+t_3)$.

ونأخذ الجذر التكعيبي لطرفي المعادلة نحصل على

$$1+t=\sqrt[3]{(1+t_1)(1+t_2)(1+t_3)}$$

ومنه

$$t=\sqrt[3]{(1+t_1)(1+t_2)(1+t_3)}-1$$

وبصفة عامة إذا لدينا نسبة الزيادة لــ (n) فترة متوسط نســبة هــذه الزيــادة (متوسط النسب) يعطي بالعلاقة:

$$\mathbf{t=\sqrt[N]{(1+t_1)(1+t_2)(1+t_3)........(1+t_3)}-1}$$

وبتطبيق هذه العلاقة على بيانات المثال السابق فإن:

$$t=\sqrt[3]{(1,25)(1,50)(1,40)}-1$$

t = 0,3795 = سنويا = 37,95%

<u>مثال9:</u>

إذا كان عدد سكان مدينة ما في سنة 1995 يساوي 500000 نسمة وبلغ في سنة 2005، 609497 نسمة أوجد متوسط نسبة زيادة السكان خلال هذه الفترة؟.

<u>الحل:</u>

- نسبة زيادة السكان خلال الفترة = [عدد السكان في سنة 2005− عدد السكان في سنة 1995]/عدد السكان في 1995 = (609497 − 500000)/ 500000 = 21.19%

ومنه يمكن إيجاد متوسط نسبة الزيادة خلال الفترة بتطبيق العلاقة السابقة

$t = \sqrt[10]{(1,2119)} - 1$

$t = 0,02 = 2\%$

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

ثالثا: المتوسط التوافقي: Moyenne Harmonique

المتوسط التوافقي هو من المقاييس الخاصة التي تستخدم لتحديد معدلات السرعة ومتوسط الأسعار، ومتوسط الكثافة السكانية. وكمثال على ذلك نقول أن سائق قطع المسافة الفاصلة بين مدينتين على أربع مراحل متساوية، المسافة المقطوعة في كل منها 100 كم.

فإذا قطع المرحلة الأولى بسرعة 100كم/ساعة والمرحلة الثانية بسرعة 120كم/ساعة والمرحلة الثالثة بسرعة 150كم/ساعة والمرحلة الرابعة بسرعة 80كم/ساعة، أوجد متوسط سرعة هذا السائق على طول المرحلة؟.

الحل:

إذا استخدمنا المتوسط الحسابي لتحديد متوسط السرعة لهذا السائق على طول المسافة فإننا سنجده يساوي: 112.5 كم/ساعة = $\overline{X} = \frac{100 + 120 + 150 + 80}{4} =$

وهذه النتيجة غير صحيحة ومضللة بدليل أن:

الزمن الذي إستغرقه السائق في المرحلة الأولى = 1 ساعة.

الزمن الذي إستغرقه السائق في المرحلة الثانية = 6/5 ساعة.

الزمن الذي إستغرقه السائق في المرحلة الثالثة = 3/2 ساعة.

الزمن الذي إستغرقه السائق في المرحلة الرابعة = 4/5 ساعة.

ومنه الزمن الذي استغرقه السائق على طـــول المســـافة = $1+\frac{5}{6}+\frac{2}{3}+\frac{5}{4}=\frac{45}{12}$ ساعة.

فإذا كان متوسط السرعة المتوصل إليه صحيحا فإن هذا السائق سيقطع خلال المدة $\frac{45}{12}$ ساعة المسافة 422 كم (112.5 × $\frac{45}{12}$ ساعة) وهذا غير صحيح لأن المســـافة الكليـــة هي 400 كم فقط.

الآن إذا قسمنا المسافة الكلية على المدة الزمنية الفعلية فإننـــا سنحصـــل علـــى متوسط السرعة الصحيح وهو 106.67 كم/ساعة = $\frac{4800}{45}=\frac{400}{\frac{45}{12}}$

وهذا المتوسط كما سنرى لاحقا يمكن إيجاده بالمتوسط التوافقي.

تعريف: المتوسط التوافقي لمجموعة من القيم هو مقلوب المتوسط الحسابي لمقاليب هذه القيم.

فإذا كانت لدينا القيم: $x_1,x_2,x_3,x_4..........x_n$.

فإن مقاليب هذه القيم هو $\frac{1}{x_1},\frac{1}{x_2},\frac{1}{x_3},\frac{1}{x_4},..........\frac{1}{x_n}$

والمتوسط الحسابي لمقاليب هذه القيم هو $\frac{\frac{1}{x_1}+\frac{1}{x_2}+\frac{1}{x_3}+.......\frac{1}{x_n}}{n}$

ومقلـــوب المتوســـط الحســـابي لمقاليـــب هـــذه القـــيم هـــو المتوســـط التـــوافقي.

$$H=\frac{n}{\frac{1}{x_1}+\frac{1}{x_2}+\frac{1}{x_3}+.......\frac{1}{x_n}}$$

وباختصار $H=\frac{n}{\sum_{i=1}^{n}\frac{1}{x_i}}$

أما إذا كانت البيانات متكررة أو مبوبة في جداول توزيع تكراري فإن: $H = \frac{\sum_{i=1}^{n} N_i}{\sum_{i=1}^{n} \frac{n_i}{x_i}}$

حيث: n_i تمثل التكرار.

x_i تمثل القيم أو مراكز الفئات.

فإذا طبقنا علاقة المتوسط التوافقي على بيانات المثال السابق فإننا سنجد

متوسط السرعة

$$H = \frac{400}{\frac{100}{100} + \frac{100}{120} + \frac{100}{150} + \frac{100}{80}} = \frac{400}{\frac{1200 + 1000 + 800 + 1500}{1200}}$$

$H = \frac{48000}{4500} =$ 106.67 كم/ساعة

فإذا ضربنا هذه السرعة في زمن المرحلة $\frac{45}{12}$ ساعة فإننا نحصل على 400 كم هي المسافة المقطوعة فعلا.

<u>مثال10:</u>

أحسب المتوسطات الثلاث (الحسابي والهندسي والتوافقي) للبيانات التالية: 2، 4، 6، 8.

<u>الحل:</u>

7. المتوسط الحسابي $\overline{X} = \frac{2+4+6+8}{4} = \frac{20}{4} = 5$
8. المتوسط الهندسي $G = \sqrt[4]{2 \times 4 \times 6 \times 8} = \sqrt[4]{384} = 4,42$
9. المتوسط التوافقي $H = \frac{4}{\frac{1}{2} + \frac{1}{4} + \frac{1}{6} + \frac{1}{8}} = \frac{4}{\frac{24+12+8+6}{48}} = \frac{192}{50} = 3,84$

أي أن $\overline{X} > G > H$ وهذه العلاقة صحيحة في كل الحالات.

<u>مثال11:</u>

اشترى شخص من نفس السوق الكميات التالية من مادة البطاطس:

4كغ بقيمة 100 دينار ثم 5 كغ بقيمة 100 دينار أيضا ثم 8 كغ بقيمة 120 دينار.

المطلوب: إيجاد متوسط سعر البطاطس المشتراة؟.

الحل:

سعر الكيلو غرام في الحالة الأولى = $\frac{100}{4}$ = 25 دينار.

سعر الكيلو غرام في الحالة الثانية = $\frac{100}{5}$ = 20 دينار.

سعر الكيلو غرام في الحالة الثالثة = $\frac{120}{8}$ = 15 دينار.

و متوسط السعر للبطاطس المشتراة هو:

$$H = \frac{100+100+120}{\frac{100}{25}+\frac{100}{20}+\frac{120}{15}} = \frac{320}{\frac{1200+1500+2400}{300}} = \frac{96000}{1200+1500+2400}$$

$H = \frac{96000}{5100} =$ 18,82 DA ≈ 18,8235 دينار/كغ

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.

مما سبق فإن $\mathbf{\bar{X} > G > H}$

مثال12:

أوجد المتوسط الحسابي والمتوسط الهندسي والمتوسط التوافقي للبيانات المبوبة في الجدول الإحصائي التالي ثم قارن بينها؟.

الفئة	3−1	5−3	7−5	9−7	11−9	المجموع
التكرار	2	4	6	4	2	18

الحل:

الفئة	التكرار	X_i	$N_i\,X_i$
3−1	2	2	4
5−3	4	4	16
7−5	6	6	36
9−7	4	8	32
11−9	2	10	20
المجموع	18		108

المتوسط الحسابي= $\overline{X} = \frac{\sum n_i x_i}{\sum n_i} = \frac{108}{18} = 6$

المتوسط الهندسي: $LogG = \frac{1}{N}\left[\sum N_i LogX_i\right]$

$$= \frac{1}{18}\left[2Log2 + 4Log4 + 6Log6 + 4Log8 + 2Log10\right]$$

$$= \frac{1}{18}\left[0,602 + 2,408 + 4,669 + 3,612 + 2\right]$$

$$= \frac{1}{18} \times 13,291$$

$$LogG = 0,7384$$

ومنه G= 5,48

المتوسط التوافقي:

$$H=\frac{\sum N_i}{\sum \frac{N_i}{X_i}}=\frac{18}{\frac{2}{2}+\frac{4}{4}+\frac{6}{6}+\frac{4}{8}+\frac{2}{10}}$$

$$H=\frac{18}{1+1+1+\frac{1}{2}+\frac{1}{5}}=\frac{180}{10+10+10+5+2}=\frac{180}{37}=4{,}86$$

ويلاحظ بوضوح: $\bar{\mathbf{X}}>\mathbf{G}>\mathbf{H}$

رابعا: الوسيط Mediane

تبين لنا عند دراستنا للمتوسط الحسابي أن هذا المتوسط يعطي نتيجة صحيحة ومنطقية عندما تكون البيانات التي حسب منها متجانسة ومتقاربة، أما إذا كانت تحتوي على قيم متطرفة في الصغر أو الكبر فإن النتيجة التي يعطيها تكون غير واقعية، في مثل هذه الحالات فقد وجد متوسط آخر سمي بالوسيط الذي هو أكثر واقعية ودلالة وصحة للحصول على فكرة عامة عن حالة البيانات التي بها قيم متطرفة.

تعريف: وكتعريف فإن الوسيط هو القيمة التي تقسم مجموع البيانات إلى قسمين بحيث يكون نصف عدد البيانات أكبر منه ونصف عدد البيانات أصغر منه ويرمز له بالرمز: Me.

<u>مثال13:</u>

أوجد الوسيط للبيانات التالية: 8، 6، 8، 10، 12، 15، 9؟

<u>الحل:</u>

لإيجاد الوسيط نقوم بالتالي:

أ) نرتب البيانات تصاعديا أو تنازليا:

6، 8، 8، 9، 10، 12، 15.

ب) نبحث عن الوسيط لهذه البيانات، 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 $\frac{n}{2}+1$.

وفي مثالنا فإن عدد البيانات المعطاة هو 7 أي فردي وبالتالي فإن الوسيط هو القيمة التي ترتيبها $\frac{7+1}{8}=4$ وهو **9**.

ويلاحظ جليا أن عدد البيانات أقل من 9 يساوي عدد البيانات أكبر من 9.

<u>مثال14:</u>

البيانات التالية تمثل الدرجات التي تحصل عليها 10 طلبة في امتحان معين:

16، 17، 17، 15، 14، 16، 15، 13، 4، 3.

المطلوب:أ) إيجاد متوسط درجات هؤلاء الطلبة؟.

ب) إيجاد وسيط الدرجات؟

ج) أيهما أكثر تعبيرا على نتائج الطلبة ولماذا؟.

الحل: أ) متوسط الدرجات $\overline{X}=\frac{16+17+17+15+14+16+15+13+4+3}{10}$ **=**

$$\overline{X}=\frac{130}{10}=13$$

ب) الوسيط: نرتب أولا هذه البيانات ترتيبا تصاعديا: 3، 4، 13، 14، 15، 15، 16، 16، 17، 17.

بما أن عدد البيانات زوجي فإن الوسيط هو المتوسط الحسابي للقيمتين اللتين ترتيبهما $\frac{n}{2}$ و $\frac{n}{2}+1$ وهما 15 و15 وبالتالي فإن الوسيط هو =15= Me.

ج) نلاحظ من النتائج المتحصل عليها أن المتوسط الحسابي $(\overline{X}=13)$ لم ينصف أغلب (8) الطلبة الذين تحصلوا على درجات أكبر بكثير من هذا المتوسط وأنحاز ناحية نتيجة الطالبين اللذين تحصلا على نتائج سيئة، في حين أن وسيط هذه الدرجات (Me = 15)

قسم نتائج الطلبة إلى قسمين بحيث نصف عدد الطلبة تحصلوا على درجات أعلى منـــه ونصف عدد الطلبة تحصلوا على درجة أقل منه وهو هنا أصدق من المتوسط الحسابي:

الوسيط في حالة بيانات متكررة:

الوسيط لبيانات متكررة:

إذا كانت البيانات المراد حساب الوسيط لها متكررة (لها تكرارات) فإن الوسيط يوجد بإتباع الخطوات التالية:

10. نحسب التكرار المتجمع الصاعد لقيم الظاهرة.
11. نحدد ترتيب الوسيط $\frac{N}{2}$ (حيث N = مجموع التكرارات).
12. نبحث عن القيمة التي تكرارها المتجمع الصاعد أكبر مـــن $\frac{N}{2}$ مباشـــرة وهي القيمة التي تمثل الوسيط.

مثال15:

الجدول التالي يمثل توزيع الطلبة فوج معين حسب الدرجات التي تحصلوا عليها في الفرض الفجائي في مادة الإحصاء. المطلوب إيجاد الوسيط لهذه البيانات؟.

الدرجة	4	5	6	7	8	9	المجموع
عدد الطلبة	2	4	8	6	5	4	**29**

1) نحسب التكرار المتجمع الصاعد للبيانات المعطاة:

الدرجة X_i	4	5	6	7	8	9	المجموع
عدد الطلبة N_i	2	4	8	6	5	4	**29**

التكرار المتجمع الصاعد	2	6	14	20	25	29	/

2) نحسب ترتيب الوسيط = $\frac{N}{2}$ = 14.5.

3) القيمة التي تكرارها المتجمع الصاعد أكبر مباشرة من $\frac{N}{2}$ هي 7 وبالتالي فإن الوسيط = 7.

الوسيط لبيانات 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 الوسيطية 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.

$$Me = L_1 + \frac{\frac{N}{2} - N_0}{n_e}.K$$

استنتاج العلاقة التي تحسب الوسيط:

لتبيان كيف تحصلنا على العلاقة التي تحسب الوسيط نأخذ المثال التالي:

<u>مثال16:</u>

يبين التوزيع التكراري التالي توزيع 50 طالب حسب الدرجة المتحصل عليها في امتحان مادة ما.

المطلوب:1) استنتج المعادلة التي تحسب الوسيط؟.

2) أحسب قيمة الوسيط لهذه البيانات؟.

فئة الدرجات	4−2	6−4	8−6	10−8	12−10	14−12	16−14	المجموع
عدد الطلبة	2	6	8	10	14	6	4	**50**

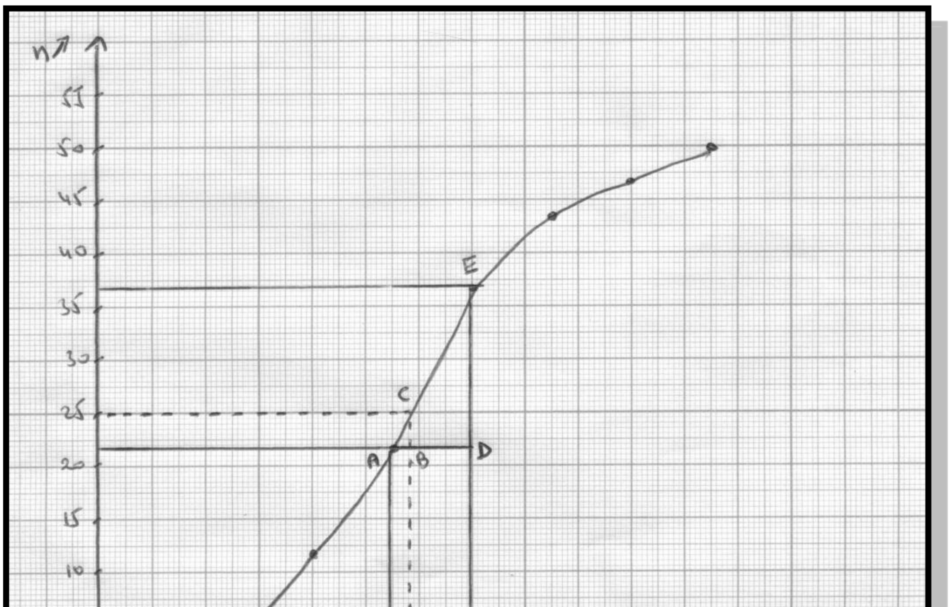
<u>الحل:</u>

أ) نحسب التكرار المتجمع الصاعد لبيانات التوزيع كما هو مبين في الجدول الآتي:.

ب) نحدد ترتيب الوسيط = $\frac{50}{2} = 25$

ج) نرسم منحنى التكرار المتجمع الصاعد لبيانات التوزيع.

الفئة	التكرار	التكرار المتجمع الصاعد
4−2 6−4 8−6 10−8 12−10 14−12 16−14	4 8 10 15 6 4 3	4 12 22 37 43 47< 50
المجموع	**50**	/



د) نحدد الفئة الوسيطية أي الفئة التي تكرارها المتجمع يساوي ترتيب الوسيط أو الأكبر منه مباشرة، (أي الفئة التي تكرارها المتجمع الصاعد ≤ 25 وهي هنا الفئة [8–10[).

هــ) نحدد التكرار المتجمع الصاعد للفئة الوسيطية والتكرار المتجمع الصاعد للفئة السابقة ونربط بينهما وبين الحد الأدنى والحد الأعلى للفئة الوسيطية على الرسم.

و) من النقطة التي تمثل ترتيب الوسيط نرسم خط مستقيم أفقي يقطع المنحنى في نقطة فاصلتها تساوي الوسيط.

وإذا افترضنا أن منحنى التكرار المتجمع الصاعد هو عبارة عن خط مستقيم عند الفئة [8–10[فإن الميل يكون ثابتا.

ي) نستخرج علاقة الوسيط كما يلي:

المثلثان ABC و ADE متشابهان، ومنه وحسب نظرية طاليس فإنه

$$\frac{\overline{AB}}{\overline{\overline{BC}}} = \frac{\overline{AD}}{\overline{\overline{DE}}}$$

وبالتعويض:

$$\frac{M_e - L_1}{\frac{N}{2} - N_0} = \frac{L_2 - L_1}{N_1 - N_0}$$

حيث: M_e = الوسيط.

L_1 = الحد الأدنى للفئة الوسيطية.

N = مجموع التكرارات.

N_0 = التكرار المتجمع الصاعد للفئة قبل الفئة الوسيطية.

L_2 = الحد الأعلى للفئة الوسيطية.

N_1 = التكرار المتجمع الصاعد للفئة الوسطية.

$$M_e - L_1 = \frac{(\frac{N}{2} - N_0)(L_2 - L_1)}{N_1 - N_0}$$

$$\mathbf{M_e = L_1 + \frac{\frac{N}{2} - N_0}{n_e}.K}$$

وذلك لأن (L_2-L_1) = K = طول الفئة و $N_1 - N_0$ = ne = تكرار الفئة الوسيطية

2 – نحسب الآن قيمة الوسيط باستخدام العلاقة السابقة

$$M_e = 8 + \frac{25-22}{15} 2 = 8 + 0,4$$

M_e = 8,4

خواص الوسيط:

يتصف الوسيط بعدة خصائص أهمها:

1 – لا يتأثر بالقيم المتطرفة وبالتالي فإنه يعتبر أصلح 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.

خامسا – أشباه الوسيط

الوسيط هو القيمة التي تقسم مجموع البيانات إلى قسمين متساويين، بحيث نصف عدد البيانات أقل منه ونصف البيانات أكبر منه، وما دام يمكن تقسيم بيانات أي ظاهرة إلى عدة أقسام متساوية وليس إلى قسمين فقط فإنه يمكن التعامل معه القيم التي تقسم هذه البيانات بنفس طريقة التعامل مع الوسيط.

فإذا تم تقسيم البيانات إلى أربعة أقسام فإن المقياس يسمى بالربيع.

وإذا تم تقسيم البيانات إلى عشرة أقسام فإن المقياس يسمى بالعشير.

أما إذا تم تقسيم البيانات إلى 100 قسم فإن المقياس يسمى بالمئين.

الربيعات Les quartils:

هي القيم التي تقسم مجموع البيانات إلى أربع أجزاء متساوية فمثلا:

الربيع الأول Q1: ويسمى كذلك بالربيع الأدنى وهو القيمة التي تقسم مجموع البيانات إلى قسمين بحيث ربع عدد البيانات أقل منه وثلاثة أرباع البيانات أكبر منه وترتيب الربيع الأول هو $\frac{N}{4}$ ويحسب كالتالي:

$$Q_1 = L_1 + \frac{\frac{N}{4} - N_0}{n_{Q1}}.K$$

الربيع الثالث Q3: ويسمى كذلك الربيع الأعلى وهو القيمة التي تقسم مجموع البيانات إلى قسمين بحيث ثلاثة أرباع عدد البيانات أقل منه وربع عدد البيانات أكبر منه وترتيب الربيع الثالث هو $\frac{3N}{4}$.

ويحسب كالتالي: $\mathbf{Q_3 = L_1 + \frac{\frac{3N}{4} - N_0}{nQ_3}.K}$

وواضح أن الربيع الثاني هو نفسه الوسيط $Q_2 = L_1 + \frac{\frac{2N}{4} - N_0}{n_e}.K = M_e$

العشريات Les Diciles:

العشير الأول D_1: هو القيمة التي تقسم مجموع عدد البيانات إلى قسمين بحيث عشر عدد البيانات اقل منه وتسعة أعشار عدد البيانات أكبر منه وترتيبه هو $\frac{N}{10}$ ونحسب بالتالي: $\mathbf{D_1 = L_1 + \frac{\frac{N}{10} - N_0}{n_{d3}}.K}$

المئينات: Les centiles :

إذا قسمت البيانات إلى مائة قسم متساوي فإن نقاط التقسيم هذه تسمى المئينات. **فالمئين الأول C_1** هو القيمة التي يسبقها 1% من البيانات ويليها 99% من البيانات ويحسب كالتالي:

$$C_1 = L_1 + \frac{\frac{N}{100} - N_0}{n_{C1}}.K$$

والمئين العشرين C_{20} يحسب كالتالي:

$$C_{20} = L_1 + \frac{\frac{20N}{100} - N_0}{n_{C20}}.K$$

<u>مثال17</u>:

أحسب الربيع الأول والعشير الثالث والمئين الستين للبيانات التالية التي توضح درجات 50 طالب في مادة ما؟

الدرجة	4−2	6−4	8−6	10−8	−10 12	−12 14	−14 16	**المجموع**
عدد الطلبة	4	8	10	15	6	4	3	50

<u>الحل</u>:

الدرجة	**عدد الطلبة**	**التكرار المتجمع الصاعد**
4−2	4	4
6−4	8	12
8−6	10	22
10−8	15	37
	6	43
	4	47

12−10 14−12 16−14	3	50
المجموع	**50**	/

$$Q_1 = L_1 + \frac{\frac{N}{4} - N_0}{n_{Q1}}.K = 6 + \frac{12.5 - 12}{10} \times 2 = 6,1$$

$$D_3 = L_1 + \frac{\frac{3N}{10} - N_0}{n_{D3}}.K = 6 + \frac{15 - 12}{10} \times 2 = 6,6$$

$$C_{60} = L_1 + \frac{\frac{60N}{100} - N_0}{n_{C60}}.K = 8 + \frac{30 - 22}{15} \times 2 = 9,067$$

سادسا: 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 **Mo**.

مثال18:

الجدول التالي يبين إنتاج مصنع للأحذية من المقاسات المختلفة؟ أوجد المنوال؟

المقاس	40	41	42	43	44	45	46
عدد الأزواج	300	800	500	300	200	100	10

الحل:

المنوال هنا هو المقاس 41.

مثال19:

البيانات التالية تمثل التقديرات التي تحصل عليها 10 طلاب في مادة الرياضيات.

المطلوب: إيجاد المنوال

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

الحل:

المنوال هنا هو جيد.

المنوال لبيانات مبوبة في جداول توزيع تكراري:

لإيجاد المنوال من الجداول التوزيع التكراري نبحث عن الفئة المنوالية وهي الفئة التي يقابلها أكبر تكرار. وهناك أكثر من طريقة لحساب المنوال:

أ) يمكن اعتبار مركز الفئة المنوالية كمنوال على وجه التقريب.

ب) يمكن حساب المنوال بالاعتماد على الفرق بين تكرار الفئة المنوالية والفئتين التي قبلها والتي بعدها طريقة بيرسون ويكون ذلك على النحو التالي:

إذا رمزنا للفرق بين تكرار الفئة المنوالية وتكرار الفئة التي قبلها بالرمز (d_1) وللفرق بين تكرار الفئة المنوالية التي بعدها بالرمز (d_2) ولطول الفئة بالرمز (K) ولحدها الأدنى بالرمز (L_1) فنحصل على العلاقة التالية:

$$M_0 = L_1 + \frac{d_1}{d_1 + d_2}.K$$ *

يمكن كذلك إيجاد المنوال من الرسم، وذلك من خلال رسم المدرج التكراري للفئة المنوالية وللفئتين التي قبلها والتي بعدها. نقوم بعد ذلك بإيصال نهاية المستطيل للفئة المنوالية من الناحية اليسرى بنهاية المستطيل للفئة التي بعدها من الناحية اليسرى. كذلك نهاية المستطيل للفئة المنوالية من الجهة اليمنى بنهاية المستطيل للفئة التي قبلها من الجهة اليمنى.

ومن نقطة التقاطع المستقيمين ننزل عمودا على المحور الأفقي فتكون نقطة تقاطع هذا العمود المحور مع المحور الأفقي هي قيمة المنوال.

مثال20:

الجدول التالي يبين توزيع عمال مؤسسة ما حسب الأجور الموزعة المطلوب:

1) تحديد قيمة المنوال؟

* ملاحظة: توجد طريقة حسابية أخرى لإيجاد المنوال تسمى بطريقة الرافعة ولكنها أقل دقة من طريقة بيرسون.

2) قارن بين قيمة المنوال والمتوسط الحسابي والوسيط؟.

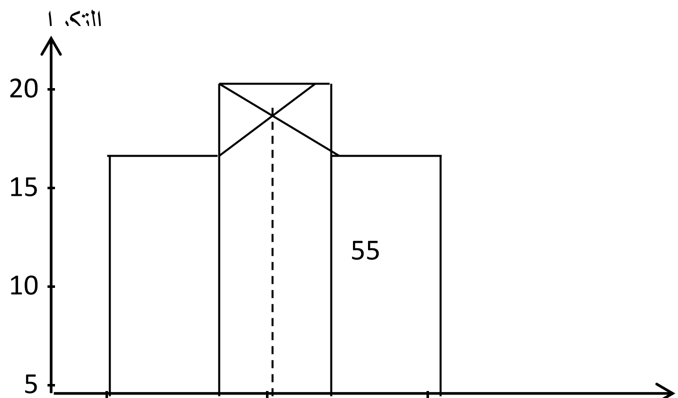
الأجــور بــآلاف الدينارات	10- 15	15- 20	20- 25	25- 30	30- 35	35- 40	المجموع
عدد العمال	10	15	20	15	10	5	75

الحل:

الأجر بالآلاف	عدد العمال	X_i	n_ix_i	$n_i \nearrow$
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.

$$M_0 = L_1 + \frac{d_1}{d_1 + d_2}.K = 20 + \frac{5}{5+5}.5$$
$$M_0 = 22{,}5$$

2) لمقارنة بين $\overline{X}$,M$_1$, M$_e$ نحس نحسب كل من المتوسط الحسابي والوسيط

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

$$M_e = L_1 + \frac{\frac{N}{2} - N_0}{n_2}.K = 20 + \frac{37{,}5 - 25}{20}.5 = 32{,}125$$

حيث أن الفئة الوسيطية هي نفسها الفئة المنوالية

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

<u>خواص المنوال:</u>

1 – لا يأخذ بعين الاعتبار جميع البيانات المعطاة وبالتالي فهو لا يتأثر بالقيم المتطرفة.

2 – يمكن حسابه بيانيا.

3 – يمكن أن يوجد أكثر من منوال لتوزيع واحد.

4 – يمكن حساب من الجداول الإحصائية المفتوحة.

5 – يعتبر أفضل المتوسطات لوصف الظواهر النوعية

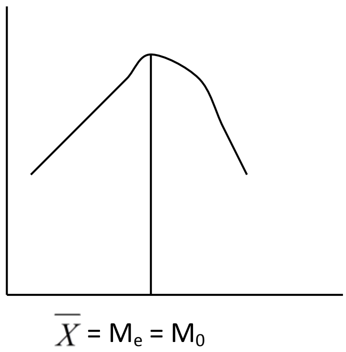
العلاقة بين المتوسط الحسابي والوسيط والمنوال:

1 – تنطبق هذه المقاييس على بعضها وتتساوى في حالة التوزيع التكـراري المعتـدل $\overline{X} = M_e = M_0$ وعندها يكون شكل منحنى التوزيع على شكل جرس.

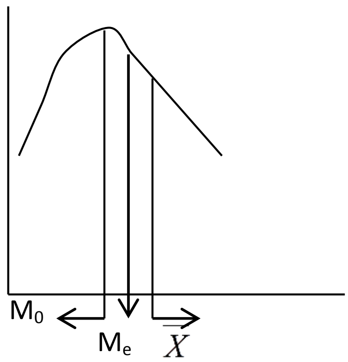
2 – عندما يكون التوزيع التكراري المدروس غير متناظر من اليمن، أي عندما تكـــون البيانات الصغيرة كثيرة (أكثر تكرارا) تكون المقاييس بالشكل التـــالي: $\overline{X} > M_e > M_0$ ويكون منحنى التوزيع ملتوي ناحية اليمين.

3 – عندما يكون التوزيع التكراري المدروس غير منتظرا من اليسار، أي عندما تكون البيانات الكبيرة كثيرة فإن المقاييس تكون بالشكل التالي: $\overline{X} < M_e < M_0$

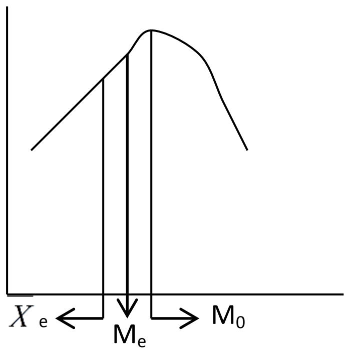
الحالة الأولى



الحالة الثانية



الحالة الثالثة



وفي كل الحالات فقد حدد كارل بيرسون علاقة تجريبية بين المقاييس الثلاث إذا كان التوزيع قريبا جدا من التناظر وهي

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

وبمقتضى هذه العلاقة يمكن استنتاج أي من المتوسطات بمعلوميـــة المقياســـين الآخرين.

سلسلة تمارين حول : مقاييس النزعة المركزية

التمرين 01:

- **الحالة الأولى:** اشترى شخص بقيمة 210000 دينار مجموعة من الأسهم بسعر 10 دينار للسهم واشترى مرة أخرى بنفس القيمة مجموعة أخرى من الأسهم بسعر 7 دينار للسهم. ما هو متوسط السعر للسهم.
- **الحالة الثانية:** الجدول التالي يبين معدل زيادة الأرباح التي حققتها شركة معينة خلال فترة رئاسة ثلاث مدراء:

الفترة	المدة بالسنوات	معدل زيادة الربح
المدير الأول	3	%5.8
المدير الثاني	1	%4.6
المدير الثالث	2	%11.2

- إيجاد معدل زيادة الأرباح خلال الفترة كاملة؟.

- **الحالة الثالثة:** قطع مسافر المسافة الكلية بين مدينتين على ثلاث مر احل كما هو مبين في الجدول أدناه:

المرحلــة	المسافة	السرعة
الأولى	40	120
الثانية	35	100
الثالثة	25	80

- أوجد متوسط سرعة هذا المسافر على طول المسافة؟

- **الحالة الرابعة:** متوسط الدخل السنوي لموظفي أحد البنوك الجزائرية لسنة 2011 كان 65110 دينار، حيث أن متوسط الدخل السنوي للرجال هو 65380 دينار أما بالنسبة للنساء فهو 65000 دينار. ماهي نسبة العمال الرجال و النساء لهذا البنك في سنة 2011؟

<u>التمرين 02</u>: الجدول التالي يمثل توزيع عمال مؤسسة ما حسب السن والجنس:

الجنس / السن	الرجال	النساء
20-25	29	38
25-30	48	57
30-35	36	42
35-40	45	39
40-45	49	41
45-50	32	30
50-55	37	18
55-60	28	20

1) أحسب متوسط السن لكل من الرجال و النساء؟
2) أحسب متوسط السن لعمال المؤسسة؟

<u>التمرين 03</u>: أقيمت دراسة حول عدد أيام العطل المرضية على 100 عامل في شركة لسكك الحديدية، فكانت نتائج الدراسة كما يلي:

عدد الغيابات	0	1	2	3	4	5
عدد العمال	9	15	11	8	4	3

1) مثل بيانيا الجدول التالي باستخدام طريقة الأعمدة البسيطة؟ ثم استنتج قيمة المنوال بيانيا؟
2) أرسم منحنى التكرارات المتجمعة الصاعدة والنازلة؟ ثم استنتج قيمة الوسيط بيانيا؟
3) أحسب كل من المقاييس التالية: المتوسط الحسابي، المنوال و الوسيط، الربيع الأول والربيع الثالث، العشير الثالث و العشير السابع، بواسطة:

أ- التكرارات المتجمعة المطلقة. ب- التكرارات المتجمعة النسبية العادية والمئوية.

ب- أحسب عدد و نسبة العمال الذين لديهم: أ- غيابين أو أقل. ب- أقل من غيابين. ج- على الأقل غيابين.

التمرين 04: يبين التوزيع التالي بعد مساكن 60 عاملا عن مكان عملهم بالكلم:

الفئات –المسافة–	02–04	04 – 06	06 –08	08 –10	10 –12	14 – 12
التكرار –عدد العمال–	06	07	15	12	؟	05

4) ماهو تقدير عدد العمال الذين تبعد مساكنهم عن مكان عملهم بأكثر أو تساوي 03 كلم ولكنها تقل أو تساوي 9.5 كلم؟

5) أحسب كل من المتوسط الحسابي، الوسط الهندسي، التوافقي ؟ ماذا تلاحظ؟

6) أحسب الوسيط والمنوال مع تفسير معنى كل منهم ؟

التمرين 05: يمثل الجدول التالي تعداد الكريات البيضاء بالآلاف عند 50 شخص:

الفئات	1-3	3-5	7-9	9-11	11-13	13-15	15-17	17-19
التكرارات	03	05	17	14	07	02	01	01

1) أحسب المتوسط الحسابي، وأحسب قيمة المنوال وفسر معناه؟

2) حدد قيمة الوسيط و الربيع الثالث؟

3) ما هو عدد ونسبة الأشخاص الذين لهم أقل من 9000 كرية بيضاء؟ ما هو عدد ونسبة الأشخاص الذين لهم 11000 كرية بيضاء على الأقل؟

التمرين 06: الجدول التالي يبين توزيع العمال حسب عدد ساعات العمل الأسبوعية.

الفئات	38 –40	40 – 42	42 – 46	– 52	52 –56	56 –58

			46			
التكرارات	30	110	240	90	20	10

1) أرسم المدرج التكراري لهذا التوزيع واستنتج قيمة المنوال بيانيا ؟

2) أرسم منحنى المتجمع الصاعد والنازل ثم استنتج قيمة الوسيط بيانيا ؟

3) أحسب المتوسط الحسابي، الوسيط والمنوال ؟ ماذا تلاحظ ؟

4) أحسب عدد و نسبة العمال الذين يفوق أو يساوي عدد ساعات عملهم عن 40 ساعة ؟

<u>التمرين 07</u>: الجدول التالي يبين التوزيع التكراري لنفقات الاستهلاكية الشهرية بآلاف الدينارات لعينة من الأسر سحبت بطريقة عشوائية من مجتمع معين:

الفئات	0 −4	4 − 8	− 12 8	12 −e_4	e_4 −22	22 −30	30 −42	المجموع
التكرارات	6	n_2	n_3	17	14	11	3	100

1) أوجد قيمة كل من n_2 و n_3 علما ان العشير الرابع يساوي 9.5 ؟

2) أوجد قيمة e_4 علما أن المتوسط الحسابي يساوي 13؟

3) حدد الانفاق الاستهلاكي للأسرة ذات المرتبة 75%؟

<u>التمرين 08</u>

قام أستاذ بإجراء امتحان لطلبته المقسمين إلى مجموعتين فإذا توفرت لك المعلومات التالية.

عدد طلبة المجموعة الأولى = 248 المتوسط الحسابي لدرجاتهم = 09.5.

عدد طلبة المجموعة الثانية = 230 المتوسط الحسابي لدرجاتهم = 10.4.

أوجد المتوسط الحسابي لمجموع الطلبة؟

<u>التمرين 09</u>

إذا علمت بأن متوسط عدد الطلبة بالحافلات الصغيرة في إحدى رحلات الجامعة كان 23 طالبا، والمتوسط بالحافلات الكبيرة كان 65 طالبا. أوجد عدد الطلبة المشاركين في الرحلة إذا كان عدد الحافلات الصغير = 3 وعدد الحافلات الكبيرة =4؟

التمرين 10

ليكن التكرار المتجمع الصاعد للظاهرة (×) على الشكل التالي 10 30 70 90 100

فإذا علمت أن طول الفئة عبارة عن جداء التكرار الأول في التكرار الأخير وأن الحد الأعلى للفئة الثالثة عبارة عن جداء تكرار الفئة الثانية في تكرار الفئة الرابعة.

المطلوب:

- إعادة تكوين حساب الجدول
- حساب المتوسط الحسابي.

التمرين 11

الجدول التالي يمثل مبيعات 500 حذاء خلال أسبوع الدخول المدرسي

قياس الحذاء	35	36	37	38	39	40	41	42	43	44
العدد	50	60	55	35	60	65	**α**	**β**	45	10

المطلوب :

- تحديد طبيعة المتغير الإحصائي.

حساب α β إذا علمت أن α = β ماذا يمثل كل من α، β.

- إيجاد قيمة المنوال من الرسم أن أمكن؟
- حساب قيمة المنوال؟

التمرين 12

إذا كان مستهلكي سلعة ما موزعين حسب فئات العمر كما يظهر في الجدول التالي:

السن	15- 20	20- 25	25- 30	30- 35	35- 40	40- 50	أكبر من 50	المجموع
النسبة	13	26	28	15	10	5	3	%100

المطلوب:

- أرسم المدرج التكراري للتوزيع واستنتج منه المنوال
- أرسم منحنى التكرار النسبي المتجمع الصاعد والنازل واستنتج منه الوسيط
- أوجد حسابيا قيم المنوال، الوسيط، المتوسط الحسابي

التمرين 13

الجدول التالي يمثل توزيع إنتاج 300 عامل بإحدى المؤسسات:

الإنتاج	60-50	70-60	80-70	90-80	-90 100	-100 110	-110 120	-120 130	-130 140	-140 150
عدد العمال	4	6	10	15	20	40	80	50	45	30

- أوجد الربيع الأول والعشير الثاني والمئين الستين وما هي دلالات كل منهما؟.
- أوجد أقل إنتاج سجل بواسطة 25 بالمائة الأحسن إنتاجا؟.
- أعلى إنتاج سجل بواسطة 20 بالمائة الأقل إنتاجا في المؤسسة؟.

تمهيد:

عرفنا في الفصل السابق أن مقاييس النزعة المركزية (من متوسط و وسيط ومنوال) تسمح لنا بالحصول على القيم المتوسطة للبيانات أو على تجمعها، غير أن هذه المقاييس لا تكفي لوحدها لمعرفة الصفات الإحصائية اللازمة لوصف الظواهر، لأن الفروق بين قيم الظواهر قد تزداد أو تنقص رغم تساوي المتوسطات لهذه الظواهر، ولتوضيح ما سبق نفترض أن طالبين تحصلا على النتائج التالية في خمس مواد دراسية:

الطالب x	10	11	13	14	15
الطالب y	8	9	13	15	18

فمتوسط درجات الطالب (x) يساوي 12.6 وكذلك متوسط درجات الطالب (y) يساوي 12.6 و وسيط درجات الطالب (x) يساوي 13 وكذلك وسيط درجات الطالب (y) يساوي 13.

قد يفهم مما سبق أن الطالبين (x)و (y) لهما نفس المستوى غير أن التمعن الجيد في الدرجات التي تحصل عليها الطالبين تبين أن الطالب (x) ناجح في كل المواد المدروسة في حين أن الطالب (y) ناجح في ثلاث مواد فقط. إن هذه الحقيقة تبين أن مقاييس النزعة المركزية لا تعطي فكرة وافية عن اختلاف قيم الظواهر ولا تحقق كل الأغراض التي نرغب الوصول إليها من دراستنا لذلك فإن مقاييس النزعة المركزية لا بد أن تكون مصحوبة بمقاييس أخرى لقياس مدى تباعد أو تقارب البيانات من بعضها البعض أو من متوسطها، تسمى هذه المقاييس بمقاييس التشتت.

ما معنى التشتت؟.

تشتت بيانات ظاهرة ما يقصد به درجة أو مقدار التفاوت أو الاختلاف بين مفردات هذه الظاهرة، وتعتبر بيانات الظاهرة متجانسة عندما تكون قيمتها قريبة من بعضها البعض ونقول في هذه الحالة أنها غير مشتتة. أما إذا كانت بيانات الظاهرة

متباعدة وغير متجانسة فنقول أن مفردات الظاهرة مشتتة وغير مركزة. ويقاس تشتت البيانات بعدة مقاييس منها:

أولا – المدى (المطلق)

المدى لمجموعة من البيانات هو الفرق بين أكبر قيمة وأصغر قيمة لها ويرمز له بالرمز R. المدى= أكبر قيمة – أصغر قيمة.

أما المدى للتوزيعات التكرارية فيحسب كما يلي: المدى = الحد الأعلى لفئة الأخيرة – الحد الأدنى للفئة الأولى

مثال1: أوجد المدى للبيانات التالية: 12،18،22،28،30.

مثال2: أوجد المدى للبيانات التالية 65،20،17،‒4، 18، 19، 14

نلاحظ المدى في هذا المثال قد تأثر بشكل كبير جدا بالقيم المتطرفة، إذ نلاحظ أن معظم البيانات متقاربة باستثناء القيمة 65 والقيمة (‒4)، فإذا استبعدنا هذه القيم المتطرفة فإن المدى يصبح: $R = x_{max} - x_{min} = 20 - 14 = 6$

وبسبب هذا العيب فإن المدى كمقياس للتشتت لا يستخدم إلا عندما نرغب في مقياس تقريبي وسريع لتشتت البيانات دون الاهتمام بالدقة في المقياس، أو عندما يكون للبيانات المتطرفة أهمية خاصة كتوزيعات درجات الحرارة على سبيل المثال، حيث تعلن درجات الحرارة اليومية بحدها الأقصى وحدها الأدنى خلال اليوم، كما يشيع استخدام هذا المقياس في حالات مراقبة جودة الإنتاج أو متابعة المبيعات التي يحققها رجال البيع لمؤسسة ما.

أما إذا أردنا أن نقلل من أثر القيم المتطرفة فإننا نقوم باستبعادها ويمكن أن يتم ذلك باستخدام الطرق التالية:

- المدى الربيعي = الربيع الثالث – الربيع الأول.
- المدى العشيري = العشير التاسع – العشير الأول.
- المدى المئيني = المئين 99 – المئين الأول.

خواص المدى:

1. يتصف المدى بسهولة حسابه.
2. يعتمد في حساب على قيمتين فقط هما القيمة الكبرى والقيمة الصغرى.
3. بسبب الخاصية الثانية فإن المدى شديد التأثر بالقيم المتطرفة.

ثانيا: الانحراف المتوسط L'écart moyen:

لأن مقاييس التشتت هي مقاييس لقوة تجمع البيانات حول بعضها، وحيث أن التجمع يكون حول القيم المتوسطة، فإنه إذا كان مقدار الاختلاف بين القيم ومتوسطها كبيرا دل ذلك على أن التشتت كبير والعكس صحيح. وحيث أن مجموع الاختلافات (الانحرافات) عن المتوسط يساوي صفرا فإنه لو حسبنا القيم المطلقة لمقدار الاختلاف عن المتوسط يكون متوسط هذه الاختلافات مقياسا مناسبا لمقدار التشتت، يسمى هذا المقياس بالانحراف المتوسط. ويعرف الانحراف المتوسط بأنه المتوسط الحسابي للقيم المطلقة لانحرافات القيم عن متوسطها الحسابي، وسوف نرمز للانحراف المتوسط في دراستنا بالرمز E_x وعليه:

إذا كانت لدينا القيم التالية: $x_1, x_2, x_3, \dots, x_n$ فإن الانحراف المتوسط لها هو:

$$E_x = \frac{|x_1-\bar{x}|+|x_2-\bar{x}|+|x_3-\bar{x}|+\cdots+|x_n-\bar{x}|}{n} = \frac{\sum_{i=1}^{n}|x_i-\bar{x}|}{n}$$

أما إذا كانت البيانات مكررة أو مبوبة في جدول توزيع تكراري فإن الانحراف المتوسط لها يعطي العلاقة:

$$E_x = \frac{\sum_{i=1}^{n} n_i|x_i-\bar{x}|}{\sum_{i=1}^{n} n_i}$$

مثال3: أوجد الانحراف المتوسط للبيانات التالية: 2،4،5،6،8؟

مثال4: أوجد تشتت البيانات المبوبة في جدول التوزيع التكراري الآتي باستخدام الانحراف المتوسط؟

الفئة	0−2	2−4	4−6	6−8	المجموع
التكرار	2	3	4	3	12

خواص الانحراف المتوسط:

1. يعتمد في حسابه على جميع القيم وليس على القيمة الكبرى والصغرى فقط.
2. يتأثر بالقيم المتطرفة، لأن انحرافها عن المتوسط الحسابي يكون كبيرا.
3. يعتبر الانحراف المتوسط أفضل من سابقه (المدى) لأنه أقل تأثر بالقيم المتطرفة غير أنه لا يستعمل بشكل واسع بسبب إعتماده على القيمة المطلقة لانحرافات القيم عن متوسطها الحسابي.

ثالثا – التباين والانحراف المعياري La variance et l'écart type:

أ – التباين La variance: وهو عبارة عن المتوسط الحسابي لمربعات الفروق بين قيم المتغير الإحصائي ومتوسطها الحسابي، ونستخدم مربعات الفروق هنا تفاديا لاستخدام القيم المطلقة كما هو الشأن في الانحراف المتوسط.

فإذا كانت لدينا البيانات التالية: $x_1, x_2, x_3, \dots, x_n$ فإن التباين لهذه البيانات يعطي بالعلاقة:

$$V_x = \frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+(x_3-\bar{x})^2+\cdots+(x_n-\bar{x})^2}{n} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sum_{i=1}^{n} n_i} = \frac{\sum_{i=1}^{n} x_i^2}{n} - \bar{x}^2$$

أما اذا كانت البيانات مبوبة او في جدول توزيع تكراري فيمكن حساب التباين بالعلاقة التالية:

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

مثال5:أوجد التباين للبيانات التالية: 3،7،6،1،11،5،6،9.

ب – الانحراف المعياري: ويعتبر الانحراف المعياري من أهم المقاييس الإحصائية للتشتت، وهو أكثر استخداما في النظريات والقوانين الإحصائية، لأنه يعطي فكرة سليمة ومنطقية عن ظاهرة التشتت، ويعرف الانحراف المعياري بأنه الجذر التربيعي لمتوسط مجموع مربع انحراف القيم عن متوسطها، أي أنه الجذر التربيعي للتباين. و نرمز له عادة بالرمز : S_x. و منه:$S_x = \sqrt{V_x}$

مثال6: أوجد التباين والانحراف المعياري للبيانات المبوبة في الجدول الإحصائي التالي:

الفئة	4-8	9-13	14-18	19-23	24-28	**المجموع**
التكرار	3	4	6	2	4	**19**

خصائص الانحراف المعياري:

1. إذا كان الانحراف المعيار للقيم ${}_1, x_2, x_3, \dots, x_n$ هو S_x فإنه إذا أضيفت أو طرحت قيمة ثابتة (a) إلى أو من جميع القيم فإن الانحراف المعياري لا يتغير
2. إذا كان الانحراف المعيار للقيم ${}_1, x_2, x_3, \dots, x_n$ هو S_x فإنه إذا ضربت كل قيمة بالمقدار a (قسمت كل قيمة على المقدار a) فإن الانحراف المعياري يتأثر بالمقدار نفسه

3. يأخذ الانحراف المعياري نفس وحدة القياس للمتغير الأصلي (كلغ، متر، لتر ....) لذلك لا يمكن استخدامها كأساس للمقارنة بين تشتت توزيعين لهما وحدات قياس مختلفة.

4. بما أن الانحراف المعياري يتأثر بالمتوسط الحسابي لبيانات الظاهرة فإنه لا يمكن استخدامه للمقارنة بين تشتت بيانات توزيعين لهما متوسط حسابي مختلف ولو كان هذين التوزيعين من نفس النوعية

5. إذا كان لدينا مجموعة كلية متكونة من مجموعتين جزئتين أو أكثر فإنه يمكن حساب الانحراف المعياري لها من خلال العلاقة التالي:

$$s_x = \sqrt{\frac{1}{n}[n_1 S_{x1}^2 + n_2 S_{x2}^2] + \frac{1}{n}[n_1(\bar{x}_1 - \bar{x})^2 + n_2(\bar{x}_2 - \bar{x})^2]}$$

رابعا – معامل الاختلاف Coefficient de variation:

رأينا سابقا أن الانحراف المعياري هو مقياس واقعي ومؤشر صحيح عن مقدار التشتت غير أن الخاصيتين 3و 4 السابقتين تبينان أنه إذا استخدمنا هذا المقياس للمقارنة بين تشتت ظاهرتين أو أكثر فإن المقارنة تكون واقعية فقط إذا كانت الظواهر من نوعية واحدة ولها متوسطات متساوية. أي يمكن مقارنة تشتت درجات مادة ما بدرجات مادة أخرى أو مقارنة تشتت دخل مجموعة من العمال بدخل مجموعة أخرى، وتكون المقارنة أكثر واقعية إذا كانت المتوسطات متساوية أو قريبة من بعضها. أما إذا كانت الظواهر من صفات مختلفة أو إذا كانت متوسطاتها متباعدة، فإن المقارنة اعتمادا على الانحراف المعياري ستكون غير منطقية وغير واقعية، ولهذا السبب وجدت مقاييس أخرى سميت مقاييس التشتت النسبي تعتمد على تمييز البيانات وتقيس التشتت كنسبة مئوية للمتوسط، أهم هذه المقاييس هو معامل الاختلاف. نرمز له عادة بالرمز: CV

$$CV = \frac{S_x}{\bar{X}} \times 100$$

مثال7: إذا كان متوسط درجات مجموعة من الطلبة في مادة ما هو 15 بانحراف معياري 3 ومتوسط درجاتهم في مادة أخرى هو 8 بانحراف معياري 2، فأي الدرجات في نظرك أكثر تشتتا؟

مثال8: ينتج مصنع نوعين من المصابيح الكهربائية، فإذا علمت أن المتوسط الحسابي والانحراف المعياري لعمر المصباح في كل نوع هما:

	المتوسط الحسابي	الانحراف المعياري
المصنع الأول	1500	300
المصنع الثاني	1800	325

أي المصابيح لها مدة حياة أكثر تشتت؟

سلسلة تمارين حول : مقاييس التشتت و الشكل

التمرين 01: برهن أن:

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

2) إذا كان الانحراف المعياري للقيم $x_1, x_2, x_3, \ldots, x_n$ هو S_x فإنه إذا أضيفت أو طرحت قيمة ثابتة (a) إلى أو من جميع القيم فإن الانحراف المعياري لا يتغير.

3) إذا كان الانحراف المعياري للقيم $x_1, x_2, x_3, \ldots, x_n$ هو S_x فإنه إذا ضربت كل قيمة بالمقدار a (قسمت كل قيمة على المقدار a) فإن الانحراف المعياري يتأثر بالمقدار نفسه

<u>التمرين 02</u>: كان توزيع 50 طالب حسب عدد الغيابات كالتالي:

عدد الغيابات	0	1	2	3	4	5
عدد الطلبة	9	15	11	8	4	3

1) أحسب كل من المقاييس التالية: المدى العام، المدى الربيعي، الانحراف المتوسط والانحراف المعياري ؟

2) أدرس شكل منحنى التوزيع ؟

<u>التمرين 03</u>: البيانات التالية تمثل توزيع المستخدمين بالآلاف حسب فئات الأجور (الساعة/ دينار) وذلك في إحدى المزارع الموسمية:

الفئات	**5-10**	**10-15**	**−20 15**	**−25 20**	**−30 25**	**−35 30**	**−40 35**	**−45 40**
n_i	**39**	**82**	**95**	**44**	**33**	**22**	**5**	**3**

1) أحسب كل من المقاييس التالية: المدى العام، المدى الربيعي، الانحراف المتوسط والانحراف المعياري ؟

2) أدرس شكل منحنى التوزيع التكراري ؟

<u>التمرين 04</u>: الجدول الآتي يبين أرباح الشركتين X، Y لفترة ما بملايين الدينارات: أي الشركتين أفضل بالنسبة للمستثمرين ولماذا؟

الشركة X	**10**	**65**	**45**	**50**	**10**

الشركة Y	35	40	35	30	40

التمرين 05: أخذت عينتان من مجتمعين مستقلين عن بعضهما البعض فأعطت النتائج التالية:

العينة الأولى	العينة الثانية
$\sum_{i=1}^{30} x_i = 450$	$\sum_{i=1}^{20} y_i = 100$
$\sum_{i=1}^{30} x_i^2 = 6900$	$\sum_{i=1}^{20} y_i^2 = 3500$

3) أحسب المتوسط الحسابي و الانحراف المعياري لكل عينة ؟

4) دمجت العينتان، أحسب كل من المتوسط الحسابي و الانحراف المعياري الناتج عن دمج العينتين؟

5) أحسب معامل الاختلاف لكل عينة، أي العينتان أكثر تشتتا ؟

الفصل الرابع :

الارتباط والانحدار الخطي البسيطين

عرضنا المقاييس التي تصف شكل المنحنى مثل النزعة المركزية والتشتت والتشكل من خلال وصف وعرض البيانات التي نجمعها من متغير واحد، وننتقل الآن للتعامل مع متغيرين (أو أكثر) باستخدام بعض طرق التحليل الإحصائي: مثل تحليل الارتباط الخطي البسيط لدراسة العلاقة بين المتغيرين ونوعها وقوتها، أو دراسة وتحليل تأثير أحدهما في الآخر بتحليل الانحدار الخطي البسيط، ويعتمد الانحدار والارتباط على دراسة العلاقة بين متغيرين أحدهما يتغير تبعا لتغير الأخر أي أحدهما يؤثر في الآخر(فيسمى الذي يؤثر متغير **مستقل X**، والمتأثر يسمى متغير **تابع Y**)، والغرض من كل ذلك صياغة نموذج (معادلة) للعلاقة بينهما من أجل إمكانية التنبؤ مستقبلا بقيم المتغير التابع وفق قيم المتغير المستقل (إذا بقيت نفس الظروف على حالها)، ومن أمثلة ذلك:

- تأثير تغير الدخل في تغير الإنفاق العائلي.
- سعر السلعة والكمية المطلوبة منها.
- علامات الطلاب في مقياس الإحصاء وعلاقتها بعلاماتهم في مقياس آخر.

العلاقة الخطية تعني أن نسبة الزيادة في المتغير X تساوي نسبة الزيادة في المتغير Y

وكلمة البسيطة تعني العلاقة بين متغيرين فقط.

والأمثلة في مجال العلوم الاجتماعية كثيرة لا يمكن حصرها، فإذا جمعنا بيانات عن متغيرين (X Y) وتم تمثيل قيمهما بيانيا (فيما يسمى شكل الانتشار)، فإن العلاقة بينهما غالبا ما تكون على أحد الأشكال التالية:

- علاقة خطية طردية.
- علاقة خطية عكسية.
- لا توجد علاقة.
- علاقة غير خطية.

1) الانحدار الخطي البسيط:

نستعمل تحليل الانحدار الخطي البسيط لدراسة وتحليل أثر متغير مستقل X على متغير آخر تابع Y، والتنبؤ X من خلال Y فإننا نضع نموذج للانحدار الخطي البسيط في شكل معادلة خطية من الدرجة الأولى تبين المتغير التابع كدالة في المتغير المستقل $\boxed{y=\beta_0+\beta_1 x+e}$ وتقدير هذا النموذج من خلال العينة هو $\boxed{y=\beta_0+\beta_1 x}$.

حيث:

y هو المتغير التابع (الذي يتأثر).

x هو المتغير المستقل (الذي يؤثر).

β_0 هي قيمة المتغير التابع عند انعدام المتغير المستقل.

β_1 هي مقدار التغير في y إذا تغير x بقيمة واحدة.

e هو الخطأ العشوائي، فالقيمة y التي نجدها في النموذج لا تعبر بالضرورة عن القيمة الفعلية لها لذا قد ينتج عن ذلك خطأ نسميه الخطأ العشوائي وهو الفرق بين y الفعلية وy المقدرة.

<u>النموذج الثانية هو تقدير للنموذج الأول وفيما يلي سنستعمل فقط النموذج الثاني.</u>

ونقصد بتقدير النموذج بإيجاد القيمتين β_1 و β_0 (أي تقديرهما من خلال العينة) بطريقة المربعات الصغرى OLS فتكون قيمتاهما كما يلي: $\boxed{\beta_1 = \frac{\sum\left((x_i - \overline{X})(y_i - \overline{Y})\right)}{\sum\left(x_i - \overline{X}\right)^2}}$

أو تحسب بالطريقة التالية: $\boxed{\beta_1 = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2}}$.

وتحسب B_0 كما يلي: $\boxed{\beta_0 = \overline{Y} - \beta_1 \overline{X}}$ أو تحسب بالطريقة التالية: $\boxed{\beta_0 = \frac{\sum y_i - \beta_1 \sum x_i}{n}}$

.

ومن أمثلة ذلك:

- أثر كمية السماد على إنتاجية هكتار من الأرض.
- أثر الدخل على الإنفاق (أو الاستهلاك أو الادخار).
- عدد ساعات مشاهدة التلفاز وتأثير ذلك في درجة الاستيعاب لدى الطفل.
- تأثير عدد ساعات المراجعة في نقطة الطالب.

2) الارتباط الخطي البسيط R_p:

يسعى الانحدار الخطي البسيط إلى تحديد **نوع العلاقة وقوتها** بين متغيرين:

- **نوع العلاقة:** وتأخذ ثلاثة أشكال حسب قيمة معامل الارتباط R حيث أن قيمته تتراوح بين (+1 و−1)
  - $R < 0$ ارتباط سالب أي العلاقة بين المتغيرين عكسية، وتغير أحدهما بالزيادة يؤدي لتغير الآخر بالنقصان والعكس.
  - $R = 0$ لا يوجد ارتباط أي لا توجد علاقة وتغير أحد المتغيرين ليس له علاقة بتغير الآخر ولا يؤثر فيه.
  - $R > 0$ ارتباط موجب أي العلاقة بين المتغيرين طردية، وتغير أحدهما بالزيادة يؤدي لتغير الآخر بالزيادة والعكس.
- **قوة العلاقة:** وذلك من حيث قربها أو بعدها عن (+1 أو −1):

درجات قوة معامل الارتباط (قوة العلاقة أو درجة الارتباط)										
ارتباط عكسي						ارتباط طردي				
قوي جدا	قوي	قوي	متوسط	ضعيف	ضعيف جدا	ضعيف جدا	ضعيف	متوسط	قوي	قوي جدا
1+	0.9−	0.7−	0.5−	0.3−	0	0.3+	0.5+	0.7+	0.9+	1+
تامة					منعدمة					تامة

ولقياس الارتباط توجد عدة معاملات منها:

2–1) معامل الارتباط الخطي لبيرسون "Pearson":

عند جمع بيانات عن متغيرين كميين (X, Y) (الطول والوزن، الإنتاج والتكلفة، الدخل والاستهلاك، علامة الطالب وساعات المراجعة) يمكن قياس الارتبـــاط بينهمـــا بطريقة بيرسون، ويحسب معامل الارتباط لبيرسون كما يلي:

$$\boxed{R_p = \frac{\sum\left((x_i-\overline{X})(y_i-\overline{Y})\right)}{\sqrt{\sum\left(x_i-\overline{X}\right)^2}\sqrt{\sum\left(y_i-\overline{Y}\right)^2}}}$$

أي نحتاج لحسابه أن نضيف 5 أعمدة نضع فيها القـــيم: $(x_i-\overline{X})$، و $(y_i-\overline{Y})$، $(x_i-\overline{X})(y_i-\overline{Y})$، ثم $(x_i-\overline{X})^2$ و $(y_i-\overline{Y})^2$. وكل ذلك بعد حساب الوسط الحسابي.

أو نحسبه بالطريقة التالية:

$$\boxed{R_p = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2}\sqrt{n\sum y_i^2 - (\sum y_i)^2}}}$$

أي نحتاج لحسابه أن نضيف 3 أعمدة نضع فيها القـــيم $x_i y_i$ و x_i^2 و y_i^2، وذلـــك دون حساب الوسط الحسابي.

2–2) معامل الارتباط الرتبي (ارتباط الرتب) لسبيرمان "Spearman":

عند جمع بيانات عن متغيرين أحـــدهما أو كلاهمـــا وصـــفي ترتيبـــي (X, Y) (علامات الطلبة في مادتين X Y، مستوى الدخل ودرجة التفضيل لسلعة معين) فـــيمكن استخدام طريقة بيرسون لحساب معامل ارتباط يعتمد على رتب ومستويات المتغيـــرين

كبديل للقيم الأصلية لهما ويسمى هذا المعامل **بمعامل ارتباط الرتب لسبيرمان** ويحسب بالمعادلة التالية: $\boxed{R_S = 1 - \frac{6\sum d^2}{n(n^2-1)}}$

حيث: d هو الفرق بين رتب مستويات المتغير الأول x ورتب مستويات المتغير الثاني y. أي: $\boxed{d = R_x - R_y}$

<u>مثال</u>: الملاحظات غير مكررة حيث عدد الملاحظات يساوي عدد الطلبة وهي غير مكررة

أوجد R_S معامل ارتباط الرتب لسبيرمان بين ملاحظات الطلبة في مقياسي الإحصاء وعلم الاجتماع لعينة عشوائية من 6 طلاب:

d^2	d	رتب الإحصاء	رتب الاقتصاد	ملاحظات الإحصاء Y	ملاحظات الاقتصاد X	رقم الطالب
1	1	4	5	مقبول	ضعيف	1
1	1-	2	1	جيد جدا	ممتاز	2
0	0	3	3	جيد	جيد	3
1	1	5	6	ضعيف	ضعيف جدا	4
4	2	6	4	ضعيف جدا	مقبول	5
1	1	1	2	ممتاز	جيد جدا	6
8						المجموع

- نرتب الملاحظات تصاعديا (أو تنازليا): بداية من: ممتاز1، جيد جدا2، جيد3، مقبول4، ضعيف5 ثم ضعيف جدا6.
- نضع في خانتي الرتب الأرقام التي تقابل كل ملاحظة.

وبتطبيق المعادلة السابقة نجد: $R_S = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6\times 8}{6(36-1)} = 1 - 0.2286 = 0.7714$

ويمكن أن نقول أن العلاقة بين ملاحظات الإحصاء وملاحظات الاقتصاد **<u>طردية وقوية</u>**.

<u>مثال</u> : (الملاحظات مكررة) فلو كانت ملاحظات الطالب في الإحصاء والاقتصاد بالصيغة التالية:

n	X	Y	رتب X	رتب Y	d	d^2
1	A	A+	2	1	1	1
2	C+	D	7.5	10	2.5	6.25
3	D	C	10	8	2	4
4	D+	C	9	8	1	1
5	B+	A	4	2	2	4
6	C+	B	7.5	5	2.5	6.25
7	A+	B+	1	3	2	4
8	B	B	6	5	1	1
9	B+	C	4	8	4	16
10	B+	B	4	5	1	1
						44.5

- نرتب الملاحظات من A+ إلى C كما يلي:

الرتب	1	2	3	4	5	6	7	8	9	10
X	A+	A	B+	B+	B+	B	C+	C+	D+	D
رتب X	1	2	4=3/(5+4+3)			6	7.5		9	10
Y	A+	A	B+	B	B	B	C	C	C	D
رتب Y	1	2	3	5			8			10

- في عمودي رتب X وY في الجدول نضع الرتبة المقابلة لكل ملاحظة في عمودي X وY.
- ثم نحسب الفروق بين رتب الملاحظات ومربعات الفروق.
- ونحسب المعامل كما يلي:

$$R_S = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6 \times 44.5}{10(100-1)} = 0.7303$$

فنقول أن العلاقة بين ملاحظات الإحصاء والاقتصاد قوية.

الفصل الخامس :

المتغيرات العشوائية ذات التوزيع الاحتمالي

- **التجربة العشوائية:** أي عملية أو تجربة لا يمكن تحديد نتائجها.
- **فراغ العينة:** مجموع النتائج أو الحودث الممكنة للتجربة.(قطعة نقد مرتين ((n(s)=4))، قطعة نرد مرتين (n(s)=36).
- **الحادث:** هو كل مفردة في فراغ العينة (ويكون بسيطا أو مركبا) (وله حالات الاتحاد، التقاطع، التكامل، التنافي).
- **الإمكانية:** تعبر عن فرصة أو نسبة وقوع حادث معين(**التكرار**).
- **الاحتمال التجريبي:** هو **التكرار النسبي** أي تكرار الحادث مقسوما على فراغ العينة.

يُدْرَسُ المتغير المتقطع بعدة توزيعات احتمالية منها (توزيع ذي الحدين، وتوزيع Poisson). و يُدْرَسُ المتغير المستمر بعدة توزيعات منها (التوزيع المنتظم، التوزيع الأسي السالب، توزيع Student والتوزيع الطبيعي).

1) التوزيع الاحتمالي للمتغير الكمي المستمر (باستخدام التوزيع الطبيعي):

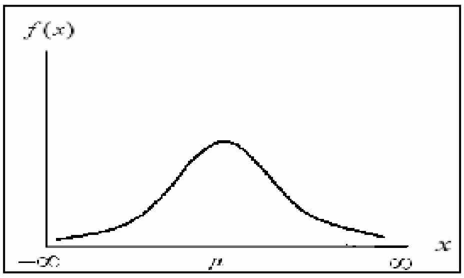
يعتبر هذا التوزيع من أكثر التوزيعات الاحتمالية استخداما في النـــواحي التطبيقيـــة كالاستدلال الإحصائي (التقدير، التنبؤ واختبار الفروض)، كما أن معظـــم التوزيعـــات يمكن تقريبها إلى هذا التوزيع.

1-1) شكل دالة الكثافة الاحتمالية للتوزيع الطبيعي:

إذا كان x_i متغير عشوائي يتوزع توزيعا طبيعيا مداه هو $-\infty < x_i < +\infty$ فـــإن دالة كثافة احتماله هي:

$$\boxed{f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x_i - \overline{X}}{\sigma}\right)^2}}$$

حيث $\boxed{-\infty < x_i < +\infty}$ ، $\boxed{\pi = \frac{22}{7}}$



الشكل (5- 1): تماثل على جانبي الوسط الحسابي

ولهذا التوزيع منحنى متماثل على جانبي الوسط الحسابي كما في الشكل المقابل:

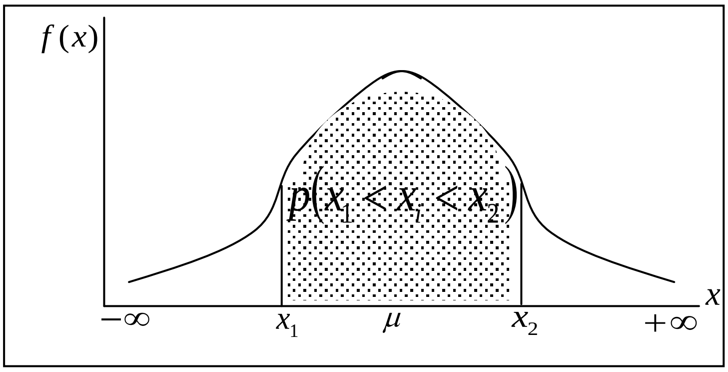
1-2) معالم وخصائص التوزيع الطبيعي:

لهذا التوزيع معلمتين هما: الوسط الحسابي $E(x)=\mu$ والتباين $\text{var}(x)=\sigma^2$ ويعبر عن هذا التوزيع كما يلي: $\boxed{x \sim N(\mu, \sigma^2)}$

ويعني أن المتغير العشوائي x يتبع قانون التوزيع الطبيعي (loi de **N**ormalité) بمتوسط قدره μ وتباين قدره σ^2. ومن خصائص هذا التوزيع أنه أكثر التوزيعات الاحتمالية استعمالا، وتشتق منه كل التوزيعات الاحتمالية المستعملة في الاستدلال الإحصائي: ووسطه μ وتباينه σ^2 وهو متماثل على جانبي الوسط الحسابي.

1-3) كيفية حساب الاحتمالات (التكرارات النسبية $f(x)$):

نفرض أن الاحتمال المراد حسابه هو: ما هو احتمال أن تقع القيمة x_i مثلا بين قيمتين هما x_1 وx_2. أي $p(x_1 < x_i < x_2)$ وهذا الاحتمال يحدد بالمساحة الموجودة تحت المنحنى حيث أن المساحة الإجمالية تحت المنحنى تساوي الواحد الصحيح (أي مجموع التكرارات النسبية).



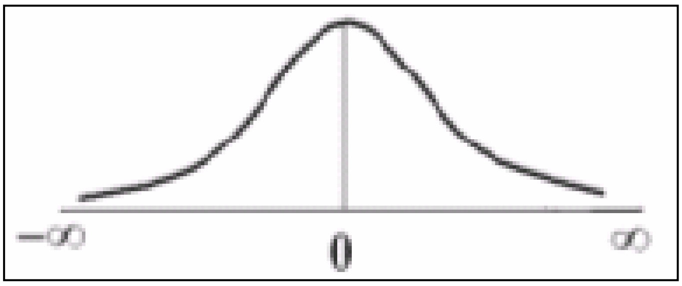
وحيث أن هذا التوزيع من التوزيعات المستمرة فإن هذه المساحة تحت المنحنى تحسب بتكامل معادلة دالة الكثافة الاحتمالية السابقة (الطلبة غير مطالبين بذلك). ونظرا لصعوبته لجأ الإحصائيون إلى عملية تحويل رياضية وبتعويض متغير جديد (Z) بدل (X) حيث: $\boxed{Z=\frac{x_i-\overline{X}}{\sigma}}$.

وحيث أن المتغير x يعرف بالمتغير الطبيعي.فالمتغير الجديد z يعرف بالمتغير الطبيعي القياسي($Standard\ Normal\ Variable$)، وهذا المتغير دالة كثافته الاحتمالية:

$$\boxed{f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{z^2}{2}}}$$

حيث $\boxed{-\infty < z_i < +\infty}$، $\boxed{\pi=\frac{22}{7}}$

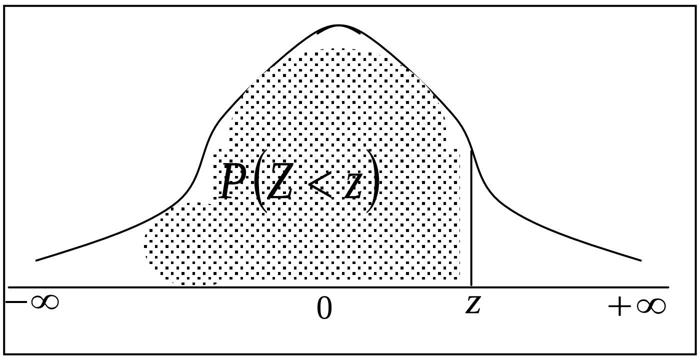
لهذا التوزيع معلمتين هما: الوسط الحسابي $E(z)=0$ والتباين $\mathrm{var}(z)=1$



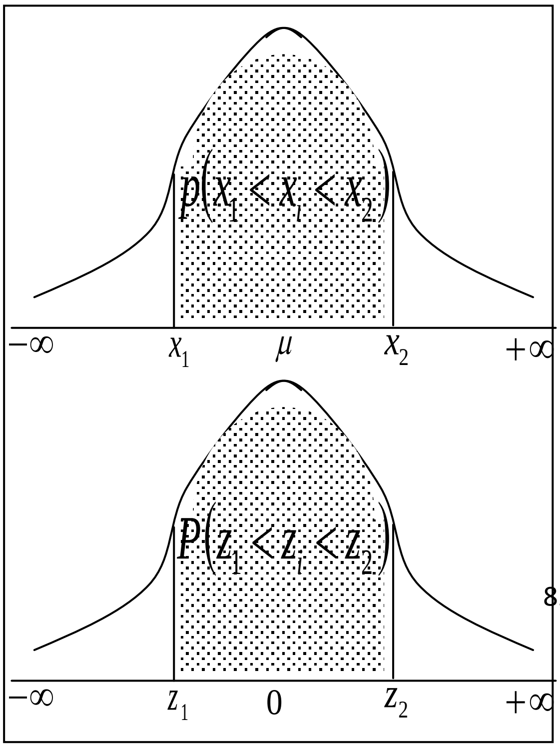
ويعبر عن هذا التوزيع كما يلي: $z \sim N(0,1)$ ويعني ذلك أن المتغير العشوائي z يتبع قانون التوزيع الطبيعي (loi de **N**ormalité) بمتوسط حسابي قدره 0 وتباين قدره 1.

ويأخذ هذا المنحنى الشكل الناقوس أو الجرسي المتماثل على جانبي الوسط الحسابي 0.

Z



وصمم الإحصائيون جداول إحصائية لحساب دالة التوزيع التجميعي F حيث أن $F(z)=P(Z<z)$ كما يبين الشكل التالي: وحيث أن المساحة تحت المنحنى تساوي 1 والمنحنى متماثل فإن $P(Z<0)=0.5$ أي أن نصف مساحة المنحنى الأقل من الوسط الحسابي 0 تساوي 0.5.



- أوجد الاحتمالات التالية باستخدام جدول التوزيع الطبيعي:

$P(z<0.57)$ $\quad$ $P(z<1.57)$

$P(z<2.57)$

$P(z > 1.96) \quad P(z > 1.21) \quad P(z > 0.96)$

$P(z > -2.68) \quad P(z > -1.68) \quad P(z > -0.68)$

$P(z < -2.33) \quad P(z < -1.33) \quad P(z < -0.33)$

ونفـــس الشـــيء بالنســـبة لاحتمـــال وقـــوع قيمـــة x_i بـــين قيمتـــين x_1 وx_2. أي

$$p(x_1 < x_i < x_2) = p(z_1 < z_i < z_2)$$

- نحول القيم الطبيعة(x_1,x_2) إلى قيم طبيعية قياسية(z_1, z_2):

$$Z_2 = \frac{x_2 - \overline{X}}{\sigma} \quad و \quad Z_1 = \frac{x_1 - \overline{X}}{\sigma}$$

- ومن ثم يكون الاحتمال: $p(x_1 < x_i < x_2) = p(z_1 < z_i < z_2)$
- نستخدم جداول التوزيع الطبيعي القياسي لحساب الاحتمالات.
- ثم أوجد الاحتمالات التالية: $P(2.01 < z < 1.28)$ $\quad P(-2.01 < z < 1.28)$

$P(-2.01 < z < -1.28)$

ونفس الطريقة عند استعمال جدول توزيع ستيودنت (وذلك عند كون حجم العينة أقل من 30)

السلسلة الأولى (عرض البيانات):

التمرين الأول:

سجلت مصلحة طب العيون بمستشفى، حضور 80 مريضا يوم: 2008/11/16، حيث تم قياس ضغط دمهم (وهو المتغير المدروس ونعبر عنه بـــ X_i).

المطلوب:

1- يتكون المجتمع المدروس من:

أ- كل المرضى في المستشفى.

ب- كل المرضى المصابين بمرض العيون.

د- كل المرضى المسجلين بهذه المصلحة.

ج- 80 مريضا الذين حضروا يوم 2008/11/16.

2- المتغير X_i المدروس هو متغير:

أ- كمي مستمر. ب- كمي متقطع.

ج- كيفي اسمي. د- كيفي ترتيبي.

التمرين الثاني:

أجري استطلاع للرأي على 200 مواطن جزائري لمعرفة رأيهم حول تعديل الدستور بصيغته المقترحة، وكانت الإجابة المطلوبة من المواطنين هي إما: "موافق أو غير موافق" على تعديل الدستور، أما الإجابة المطلوبة من طلبة العلوم الاقتصادية سنة أولى "ل م د" هي:

1- يتكون المجتمع الإحصائي من:

أ- كل المواطنين.

ب- المواطنين الجزائريين.

ج- من المواطنين المستجوبين.

د- من المواطنين الذين اختير منهم 200 مستجوب.

2- المتغير المدروس هو:

أ- كمي مستمر. ب- كمي متقطع.

ج- وصفي اسمي. د- وصفي ترتيبي.

التمرين الثالث:

يتكون مجتمع ما من 4 فئات اجتماعية (حسب المهنة) N=5000، وحجم كل فئة هو كالتالي:N1 = 1000 ; N2 = 1800 ; N3 = 1600 ; N4 = 600، ونريد سحب عينة حجمها n = 180.

المطلوب:

1−ما هي طبيعة المجتمع المدروس؟.

2−حدد عدد الوحدات الإحصائية التي يجب سحبها من كل فئة.

السلسلة الثانية (عرض البيانات):

التمرين الأول:

فيما يلي بيانات عن المستوى التعليمي لعينة من 50 فرد.

ابتدائي	أساسي	ثانوي	جامعي	ابتدائي	أساسي	ثانوي	أساسي	جامعي	أميٌّ
ثانوي	أميٌّ	جامعي	ثانوي	أساسي	ثانوي	أميٌّ	جامعي	ثانوي	أساسي
أساسي	ثانوي	أساسي	ثانوي	جامعي	ثانوي	أساسي	ثانوي	ابتدائي	ثانوي
ثانوي	أساسي	ثانوي	ثانوي	أساسي	جامعي	ثانوي	ابتدائي	أميٌّ	ابتدائي
جامعي	ابتدائي	ثانوي	ابتدائي	جامعي	أميٌّ	أساسي	ثانوي	أساسي	جامعي

المطلوب:

1- اعرض البيانات في شكل جدول تكراري وكون التوزيع التكراري النسبي.

2- علق على النتائج.

التمرين الثاني:

فيما يلي بيانات درجات 70 طالب في اختبار مقياس الإحصاء الوصفي.

56	65	70	65	55	60	66	70	75	56
60	70	61	67	61	71	67	62	71	66
68	72	57	68	72	69	57	71	69	75
72	62	67	73	58	63	66	73	63	65
58	73	74	76	74	80	81	60	74	58
76	82	77	83	77	85	91	78	94	72
79	64	57	79	55	87	64	88	78	62

المطلوب:

1- كون جميع أنواع التوزيعات التكرارية، بما فيها النسبية.

2- ما هى نسبة الطلاب الحاصلين على درجة ما بين 70 إلى أقل من 80 ؟، نسبة الطلاب الحاصلين على درجة أقل من 70 درجة ؟، وما هى نسبة الطلاب الحاصلين على درجة 80 أو أكثر ؟.

3- أرسم المدرج البياني والمضلع البياني، والمنحنى البياني لجميع التوزيعات السابقة.

حل السلسلة الثانية (عرض البيانات):

حل التمرين الأول:

1- عرض البيانات في جدول تكراري وتكوين التوزيع التكراري النسبي:

نلاحظ من البيانات وجود 5 مستويات أو حالات لمتغير المستوى التعليمي:

المتغير X_i	التكرار	التكرار النسبي
أميّ	5	0.10=
ابتدائي	7	0.14
أساسي	12	0.24
ثانوي	16	0.32
جامعي	10	0.20
المجموع	50	1.00

2-التعليق على النتائج: (تعميم النتائج على المجتمع)

- درجة الأمية بلغت 10 % من أفراد المجتمع.
- 14% من أفراد المجتمع لا يتعدى مستواهم التعليمي الابتدائي.
- 24% من أفراد المجتمع لا يتعدى مستواهم التعليمي الأساسي.
- 32% من أفراد المجتمع يملكون المستوى الثانوي.
- 20% فقط من أفراد المجتمع يملكون المستوى الجامعي.

<u>حل التمرين الثاني:</u>

1-تكوين جميع أنواع التوزيعات التكرارية، بما فيها النسبية:

فئات في مثال الدرس		فئات في ت السلسلة		تكرار بسيط	تكرار صاعد	تكرار نازل	نسبي بسيط	نسبي صاعد أقل من الحد الأعلى للفئة	نسبي نازل أكبر من الحد الأدنى للفئة
18	22	55	60	10	10	70	0.14	0.14	1.00
22	26	60	65	12	22	60	0.17	0.31	0.86
26	30	65	70	13	35	48	0.19	0.50	0.69
30	34	70	75	16	51	35	0.23	0.73	0.50
34	38	75	80	10	61	19	0.14	0.87	0.27
38	42	80	85	4	65	9	0.06	0.93	0.13
42	46	85	90	3	68	5	0.04	0.97	0.07
46	50	90	95	2	70	2	0.03	1.00	0.03
				70			1.00		

2-نسبة الطلاب الحاصلين على درجة ما بين 70 إلى أقل من 80: هي مجموع نسبة الفئة 70-75 و75-80 وهي: (0.23+0.14 = 0.37 أي 37%)، نسبة الطلاب الحاصلين على درجة أقل من 70 درجة هي مجموع نسب الفئات الأقل من 70 وهي (في التكرار الصاعد النسبي 0.50 أي 50%)، ونسبة الطلاب الحاصلين على درجة 80 أو أكثر هي مجموع نسب الفئات الأكثر من 80 وهي (في التكرار النازل النسبي 0.13 أي 13%).

3-أرسم المدرج البياني والمضلع البياني، والمنحنى البياني لجميع التوزيعات السابقة.

السلسلة الثالثة (النزعة المركزية):

التمرين الأول:

قرر مدير إحدى المؤسسات بمكافأة أكثر العمال كفاءة بسحب عينة عشوائية من عمال مؤسسته فوجد كمية إنتاجهم الشهرية كالتالي: 115 123 119 123 124 119 123 121 123 121.

المطلوب:

1- ما هو متوسط الإنتاج الشهري لعمال العينة المختارة؟، وما هي القيمة التي تزيد عنها نصف كمية الإنتاج؟، وما هي قيمة الإنتاج الأكثر تكرارا؟، ثم حدد شكل الالتواء لهذه البيانات.

2- أحسب الوسط الحسابي المعدل إذا أضفنا قيمة = 5 لكل القيم العشرة السابقة.

3- أحسب الوسط الحسابي المعدل إذا ضربنا قيمة = 3 في كل القيم العشرة السابقة.

4- أحسب مجموع مربعات انحرافات القيم عن وسطها الحسابي.

5- أحسب الربيع الأول والثاني والثالث، ماذا تلاحظ؟.

التمرين الثاني:

الجدول التالي يبين توزيع 1000 ساكن حسب المساحة بالكلم^2:

المساحة بالكلم^2	70-50	90-70	110-90	130-110	150-130	170-150	190-170
عدد السكان	80	150	280	200	150	80	60

المطلوب:

1- أحسب المتوسط، والوسيط والمنوال؟، وماذا تمثل هذه القيم؟.

2- بين شكل التواء المنحنى المتعلق بتوزيع البيانات.

3- أحسب العشير الخامس والسابع، ماذا تلاحظ؟.

4- أحسب الميئي الخمسين والسبعين، والثمانين، ماذا تلاحظ؟

التمرين الثالث:

يبين الجدول التالي مستوى مبالغ كراء 2000 منزل في مدينة ما لسنة 2007:

حجم مبالغ الكراء بـ: دج	4000- 5000	5000- 6000	6000- 7000	7000- 8000	8000- 9000	9000- 10000
عدد البنايات	184	306	436	410	376	288

المطلوب:

1- بين شكل توزيع مبالغ كراء المنازل.

2- أحسب العشير السادس والثامن، ثم الميئي الثمانين والتسعين، ماذا تلاحظ.

حل السلسلة الثالثة (النزعة المركزية):

حل التمرين الأول:

1- **متوسط** الإنتاج الشهري لعمال العينة المختارة هو :

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{115+123+119+123+124+119+123+121+123+121}{10} = \frac{1211}{10} = 121.1$$

القيمة التي تزيد عنها **نصف** كمية الإنتاج: بما أن البيانات غير مبوبة أي في شكل سلسلة فإن القيمة هي:

- نرتب القيم تصاعديا: 115 119 119 121 121 123 123 123 123 124.
- نحدد رتبة الوسيط وهي: بما أن عدد المشاهدات 10 عدد زوجي فإن قيمة الوسيط هي الوسط الحسابي للقيمتين ذات الرتبتين 2/10 و (10+1)/2 أي الرتبتين 5 و6، قيمتاهما هما: 121 و123 ووسطهما الحسابي هو $\boxed{122}$، وهو يساوي وسيط البيانات أي القيمة التي تزيد عنها نصف كمية الإنتاج.

المنوال وبما أنها بيانات في شكل سلسلة فإن المنوال هو القيمة الأكثر تكـرارا وهـي بالملاحظة $\boxed{123}$.

نلاحظ أن 123>122>121.1 أي: المنوال > الوسيط > الوسط، ومنه فإن المنحنى ملتوي جهة اليسار.

2- إذا أضفنا قيمة = 5 لكل القيم السابقة تصبح: <u>120</u> <u>128</u> <u>124</u> <u>128</u> <u>129</u> <u>124</u> <u>128</u> <u>126</u> <u>128</u> <u>126</u>، والوسـط الحسـابي المعـدل يصـبح

$$\overline{Y} = \frac{120+128+124+128+129+124+128+126+128+126}{10} = \frac{1261}{10} = 126.1$$

نلاحظ أن: $\overline{Y} = \overline{X} + 5 \Leftarrow 126.1 = 121.1 + 5$

3- إذا ضربنا قيمة = 3 في كل القيم السابقة تصـبح <u>345</u> <u>369</u> <u>357</u> <u>369</u> <u>372</u> <u>357</u> <u>369</u> <u>363</u> <u>369</u> <u>363</u>، والوسـط الحسـابي يصـبح:

$$\overline{Y} = \frac{345+369+357+369+372+357+369+363+369+363}{10} = \frac{3633}{10} = 363.3$$

نلاحظ أن: $\overline{Y} = \overline{X} \times 3 \Leftarrow 363.3 = 121.1 \times 3$

4- حساب مجموع مربعات انحراف القيم عن وسطها الحسابي:

$$\sum(x_i - \overline{X})^2 = (115-121.1)^2 + ... + (121-121.1)^2 = 68.9$$ ونلاحظ أن:

$$\sum(x_i - \overline{X}) \cong 0$$

5- حساب الربيع الأول والثاني والثالث:

- <u>تحديد رتبة الربيع الأول</u>: $R = (n+1)\times(\frac{i}{4}) = 11 \times \frac{1}{4} = 2.75$
- الرتبة K التي تسبق رتبة الربيع الأول R=2.75 هي: K=2. والقيمـة التي تسبق قيمته هي $X_K = X_2 = 119$.
- الرتبة U التي تلي رتبة الربيع الأول R=2.75 هـي: U=3. والقيمـة التي تلي قيمته $X_U = X_3 = 119$.
- قيمة الربيع **الأول** هي: $Q_3 = x_K + (R-K)\times(x_U - x_K) = 119 + (2.75-2)\times(119-119) = 119$

- تحديد رتبة الربيع الثاني: $R=(n+1)\times(\frac{i}{4})=11\times\frac{2}{4}=5.5$
- الرتبة K التي تسبق رتبة الربيع الثاني R=5.25 هـــي: K=5. والقيمـــة التي تسبق قيمته هي $X_K=X_5=121$.
- الرتبة U التي تلي رتبة الربيع الثاني R=5.25 هـــي: U=6. والقيمـــة التي تلي قيمته $X_U=X_6=123$.
- قيمة الربيع **الثاني** هي: $Q_3=x_K+(R-K)\times(x_U-x_K)=$ $121+(5.5-5)\times(123-121)=122$
- تحديد رتبة الربيع الثالث: $R=(n+1)\times(\frac{i}{4})=11\times\frac{3}{4}=8.25$
- الرتبة K التي تسبق رتبة الربيع الثالث R=8.25 هـــي: K=8. والقيمـــة التي تسبق قيمته هي $X_K=X_8=123$.
- الرتبة U التي تلي رتبة الربيع الثالث R=8.25 هـــي: U=9. والقيمـــة التي تلي قيمته $X_U=X_9=123$.
- قيمة الربيع **الثالث** هي: $Q_3=x_K+(R-K)\times(x_U-x_K)=$ $123+(8.25-8)\times(123-123)=123$

نلاحظ أن الربيع الثاني يساوي الوسيط.

حل التمرين الثاني:

1- حساب مقاييس النزعة المركزية الثلاثة.

نازل	صاعد	f_ix_i	مراكز الفئات X_i	عدد السكان (التكرارات f)	فئات المساحة بالكلم²
1000	80	4800	60	80	70-50
920	230	12000	80	150	90-70
770	**510**	28000	100	280	110-90
490	710	24000	120	200	130-110
290	860	21000	140	150	150-130
140	940	12800	160	80	170-150
60	1000	10800	180	60	190-170
		113400		1000	المجموع

$\overline{X}=\frac{\sum_{i=1}^{7}x_i f_i}{n}=\frac{113400}{1000}=113.4$ إذن متوسط المساحة هو 113.4 كلم^2 لكل 1000 ساكن.

- المساحة التي يسكن فيها نصف عدد السكان تمثل وسيط المساحة ويحسب كما يلي:

بعد تكوين التكرار الصاعد، نحدد رتبة الوسيط وهي (n/2 = 2/1000 =500) والتكرار الصاعد الأكبر من هذه القيمة مباشرة هو تكـــرار الفئـــة الثالثـــة (90–110) وطولها L=20، ويتم حساب الوسيط بالمعادلة التالية:

$$Med = A+\frac{\frac{n}{2}-f_1}{f_2-f_1}\times L = 90+\frac{500-230}{510-230}\times 20 = 109.3$$

نصف عدد السكان (500) يسكنون في مساحة قدرها 109.3 كلم^2،

- حساب المنوال بالمعادلة التالية: (طريقة الفروقات):

$$\text{Mod}= A+\frac{d_1}{d_1+d_2}\times L = 90+\frac{280-150}{(280-150)+(280-200)}\times 20 = 102.4$$

معظم السكان يسكنون في مساحة قدرها: 102.4 كلم^2.

2– نلاحظ أن 113.4>109.3>102.4 أي: الوسط > الوسيط > المنوال ، ومنه فـــإن المنحنى ملتوي جهة اليمين.

3– حساب العشير الخامس والسابع:

- العشير الخامس:

$$D_5 = A+\frac{\frac{n}{10}\times 5-f_1}{f_2-f_1}\times L = 90+\frac{500-230}{510-230}\times 20 = 109.29$$

- العشير السابع:

- $D_7 = \mathrm{A} + \dfrac{\frac{n}{10}\times 7 - f_1}{f_2 - f_1}\times L = 110 + \dfrac{700-510}{710-510}\times 20 = 129$

- الميئي الخمسين: $P_{50} = \mathrm{A} + \dfrac{\frac{n}{100}\times 50 - f_1}{f_2 - f_1}\times L = 90 + \dfrac{500-230}{510-230}\times 20 = 109.29$

- الميئي السبعين: $P_{70} = \mathrm{A} + \dfrac{\frac{n}{100}\times 70 - f_1}{f_2 - f_1}\times L = 110 + \dfrac{700-510}{710-510}\times 20 = 129$

- الميئي الثمانين: $P_{80} = \mathrm{A} + \dfrac{\frac{n}{100}\times 80 - f_1}{f_2 - f_1}\times L = 130 + \dfrac{800-710}{860-710}\times 20 = 142$

- نلاحظ أن الميئي الخمسين يساوي العشير الخامس ويساوي الوسيط؛ وأن الميئي السبعين يساوي العشير السابع.

حل التمرين الثالث:

1-لتبيان شكل التوزيع نحسب الوسط، الوسيط والمنوال، ونحتاج لحساب مراكز الفئات والتكرار المتجمع الصاعد:

فئات المبالغ دج	عدد السكان (التكرارات f)	مراكز الفئات X_i	$f_i x_i$	صاعد	نازل
5000-4000	184	4500	828000	184	**2000**
6000-5000	306	5500	1683000	490	1816
7000-6000	436	6500	2834000	926	1510
8000-7000	410	7500	3075000	1336	1074
9000-8000	376	8500	3196000	1712	664
10000-9000	288	9500	2736000	**2000**	288
المجموع	2000		14352000		

- الوسط الحسابي: $\overline{X} = \dfrac{\sum_{i=1}^{6} x_i f_i}{n} = \dfrac{14352000}{2000} = 7176$

- الوسيط: $Med = \mathrm{A} + \dfrac{\frac{n}{2} - f_1}{f_2 - f_1}\times L = 7000 + \dfrac{1000-926}{1336-926}\times 1000 = 7180.5$

● المنوال:

$$\text{Mod} = A + \frac{d_1}{d_1 + d_2} \times L = 6000 + \frac{436-306}{(436-306)+(436-410)} \times 1000 = 6833.3$$

نلاحظ أن 7180.5 < 7176 < 6833.3 أي: الوسيط > الوسط > المنوال ، ومنه فإن المنحنى ملتوي جهة اليمين.

2- حساب العشير السادس والثامن، والميئي الثمانين والتسعين: ونلاحظ أن العشير الثامن يساوي الميئي الثمانين

- العشير السادس:

$$D_6 = \text{A} + \frac{\frac{n}{10} \times 6 - f_1}{f_2 - f_1} \times L = 7000 + \frac{1200-926}{1336-926} \times 1000 = 7668.3$$

- العشير الثامن:

$$D_8 = \text{A} + \frac{\frac{n}{10} \times 8 - f_1}{f_2 - f_1} \times L = 8000 + \frac{1600-1336}{1712-1336} \times 1000 = 8702.13$$

- الميئي الثمانين:

$$P_{80} = \text{A} + \frac{\frac{n}{100} \times 80 - f_1}{f_2 - f_1} \times L = 8000 + \frac{1600-1336}{1712-1336} \times 1000 = 8702.13$$

- الميئي السبعين: $P_{90} = \text{A} + \frac{\frac{n}{100} \times 90 - f_1}{f_2 - f_1} \times L = 9000 + \frac{1800-1712}{2000-1712} \times 1000 = 9305.6$

السلسلة الرابعة (التشتت والتشكل):

التمرين الأول:

إذا كانت الطاقة التصديرية لعشر محطات لتحلية المياه بالمليون متر مكعب كما يلي:

15 11 13 14 18 16 12 19 20 10 17

المطلوب:

1- أوجد المدى العام الذي تتراوح فيه قيم الطاقة التصديرية، ثـم أوجـد المـدى الربيعي ثم الانحراف الربيعي.

2- أحسب الانحراف المتوسط ، والتباين ثم الانحراف المعياري.

3- حدد شكل التواء وتفرطح هذه البيانات (بدون حساب المنوال).

التمرين الثاني:

فيما يلي الأجور الشهري بالألف دينار لـ:40 عامل في مؤسسة ما.

فئات الأجر	3-5	5-7	7-9	9-11	11-13	13-15	15-17	17-19
عدد العمال	3	4	6	7	8	5	4	3

المطلوب:

1- أوجد المدى والمدى الربيعي والانحراف الربيعي، ثم الانحراف المتوسط.

2- أحسب التباين والانحراف المعياري.

3- حدد شكل التواء البيانات بطريقة بيرسون ثم بطريقة الميئين.

4- هل شكل البيانات مدبب أم مفرطح.

التمرين الثالث:

فيما يلي 3 أنواع من المصابيح وسحبنا من كل نوع 50 مصباحا وقمنا بقيـاس مدة اشتعالها قبل أن تحترق:

العينة							
العينة الأولى	فئات الساعات	3-1	5-3	7-5	9-7	11-9	13-11
	عدد المصابيح	2	4	6	18	15	5
العينة الثانية	فئات الساعات	3-1	5-3	7-5	9-7	11-9	13-11
	عدد المصابيح	7	9	10	9	8	7
العينة الثالثة	فئات الساعات	3-1	5-3	7-5	9-7	11-9	13-11
	عدد المصابيح	2	7	9	14	13	5

المطلوب:

1- أوجد المدى العام لكل نوع، ثم المدى الربيعي والانحراف الربيعي، ثم الانحراف المتوسط.

2-ما هو أحسن نوع يجب استعماله من المصابيح؟(النوع الذي تشتته صغير وأحسن مقياس للتشتت هو الانحراف المعياري).

3-بالنسبة للنوع الأحسن:

- حدد شكل التواء توزيع البيانات بطريقتين ؟.
- هل شكل توزيع البيانات مدبب أم مفرطح؟.

حل السلسلة الرابعة (التشتت والتشكل):

حل التمرين الأول:

1- المدى العام الذي تتراوح فيه قيم الطاقة التصديرية يساوي الفرق بين أكبر قيمة وأصغر قيمة

المدى = 20−10 =10

ولحساب المدى الربيعي والانحراف الربيعي نحتاج لحساب الربيع الأول والربيع الثالث:

أولا نرتب القيم تصاعديا: 10 11 12 13 14 15 16 17 18 19 20

- تحديد رتبة الربيع الأول: $R=(n+1)\times(\frac{i}{4})=12\times\frac{1}{4}=3$

وبما أن R عدد صحيح فإن قيمة الربيع الأول هي: $Q_1 = X_R = X_3 = 12$.

- تحديد رتبة الربيع الثالث: $R=(n+1)\times(\frac{i}{4})=12\times\frac{3}{4}=9$

وبما أن R عدد صحيح فإن قيمة الربيع الأول هي: $Q_3 = X_R = X_9 = 18$.

إذن المدى الربيعي هو: $IQ=Q_3-Q_1=17-11=6$ أي أن 50% من قيم الطاقة التصديرية تتراوح في مجال قدره 6 مليون متر مكعب. والانحراف الربيعي هو: $Q=\frac{IQ}{2}=\frac{Q_3-Q_1}{2}=\frac{6}{2}=3$ أي أن نصف عدد قيم الطاقة التصديرية لا تبعد عن القيمة الوسيطية إلا بمقدار 3 مليون متر مكعب.

2− حساب الانحراف المتوسط والتباين والانحراف المعياري:

- تكوين الجدول التكراري: حيث أن الوسط الحسابي = 15

القيم	17	10	20	19	12	16	18	14	13	11	15	المجموع
$x_i-\overline{X}$	2	5-	5	4	3-	1	3	1-	2-	4-	0	0
$\lvert x_i-\overline{X}\rvert$	2	5	5	4	3	1	3	1	2	4	0	30
$(x_i-\overline{X})^2$	4	25	25	16	9	1	9	1	4	16	0	110
$(x_i-\overline{X})^4$	16	625	625	256	81	1	81	1	16	256	0	1958

- حساب الانحراف المتوسط $MD = \frac{\sum |x_i - \overline{X}|}{n} = \frac{30}{11} = 2.73$.
- حساب التباين: $s^2 = \frac{\sum (x_i - \overline{X})^2}{n} = \frac{110}{11} = 10$ والانحـــــــراف المعيـــــــاري:

$$s = \sqrt{s^2} = \sqrt{10} = 3.16$$

3-يتحدد شكل التواء البيانات بمعامل بيرسون أو معامل الميئين، أما شكل تفرطح البيانات فيتحدد بمعامل التفرطح:

الوسط=15، والوسيط رتبته $\frac{n+1}{2} = 6$ (لأن n فردي) وبالتالي قيمته هي قيمة الرتبة 6 بعد ترتيب القيم: Med =15.

- معامل بيرسون α معـــدوم: $a = \frac{3(\overline{X} - Med)}{s} = \frac{3(15-15)}{3.16} = 0$ وبالتـــالي فـــإن منحنى التوزيع التكراري متماثل.
- معامل التفرطح k: $k = \frac{\frac{1}{n}\sum (x_i - \overline{X})^4}{S^4} = \frac{\frac{1}{11} \times 1958}{100} = 1.78$
- نلاحظ أن k<3 وبالتالي فإن منحنى التوزيع التكراري منبسط أي مفرطح.

حل التمرين الثاني:

1-تحديد المدى، المدى الربيعي، الانحراف الربيعي والانحراف المتوسط، وقبـــل ذلك نضع جدول لتوزيع التكراري:

فئات الأجر	f	x_i	$f\uparrow$	$x_i f_i$	$\lvert x_i - \overline{X}\rvert$	$\lvert x_i - \overline{X}\rvert \times f$	$(x_i - \overline{X})^2$	$f \times (x_i - \overline{X})^2$	$f \times (x_i - \overline{X})^4$
5-3	3	4	2	12	7	21	49	147	7203
7-5	4	6	7	24	5	20	25	100	2500
9-7	6	8	13	48	3	18	9	54	486
11-9	7	10	20	70	1	7	1	7	7
13-11	8	12	28	96	1	8	1	8	8
15-13	5	14	33	70	3	15	9	45	405
17-15	4	16	37	64	5	20	25	100	2500
19-17	3	18	40	54	7	21	49	147	7203
المجموع	40			438	32	130	168	608	20312

- المدى = الحد الأعلى للفئة الأخيرة – الحد الأدنى للفئة الأولـــى = 115−35 = 80.
- الانحـــــراف الربيعـــــي يحســــــب بحســــــاب الربيـــــع الأول والثالـــــث:

$$Q_1 = A + \frac{\frac{n}{4}\times 1 - f_1}{f_2 - f_1}\times L = 7 + \frac{10-7}{13-7}\times 2 = 8$$

$$Q_3 = A + \frac{\frac{n}{4}\times 3 - f_1}{f_2 - f_1}\times L = 13 + \frac{30-28}{33-28}\times 2 = 13.8$$

المدى الربيعي: $IQ = Q_3 - Q_1 = 5.8$ والانحراف الربيعي: $Q = \frac{IQ}{2} = 2.9$

الانحـــــراف المتوســـــط للبيانـــــات المبوبـــــة هـــــو $\boxed{MD = \frac{\sum\left|x_i - \overline{X}\right|\times f}{n}}$ حيـــــث

$$\overline{X} = \frac{\sum_{i=1}^{k} x_i f_i}{\sum_{i=1}^{k} f_i = n} = \frac{438}{40} \cong 11$$

وبالتالي الانحراف المتوسط: $MD = \frac{\sum\left|x_i - \overline{X}\right|\times f}{n} = \frac{130}{40} = 3.25$

2−حساب التباين: $s^2 = \frac{\sum f \times \left(x_i - \overline{X}\right)^2}{n} = \frac{608}{40} = 15.2$ والانحراف المعيـــاري:

$$s = \sqrt{s^2} = \sqrt{15.2} = 3.9$$

3−تحديد شكل التواء البيانات بطريقة الميئين: (طريقة بيرسون رأيناها في التمرين الأول):

$\alpha_{i,\,100-i} = \frac{(P_{100-i} - P_{50}) - (P_{50} - Pi)}{P_{100-i} - Pi}$ فنحسب الميئين P_{20} والميئين P_{80} مثلا:

- الميئي العشرين: $P_{20} = A + \frac{\frac{n}{100}\times 20 - f_1}{f_2 - f_1}\times L = 7 + \frac{8-7}{13-7}\times 2 = 7.33$

- الميئي الثمانين: $P_{80} = A + \dfrac{\dfrac{n}{100} \times 80 - f_1}{f_2 - f_1} \times L = 13 + \dfrac{32-28}{33-28} \times 2 = 14.6$

4-نحدد شكل تفرطح البيانـــات بحســـاب معامـــل التفـــرطح بالمعادلـــة التاليـــة:

$$k = \frac{\frac{1}{n}\sum (x_i - \overline{X})^4}{S^4} = \frac{\frac{1}{40} \times 20312}{3.899^4} = 2.2$$

نلاحظ أن k>3 وبالتالي فإن منحنى التوزيع التكراري مدبب.

حل التمرين الثالث:

1-تحديد المدى العام لكل نوع:

- نلاحظ أن المدى لكل نوع نفسه = الحد الأعلى للفئة الأخيرة – الحد الأعلى للفئة الأولى = 1300−100=1200
- لحساب المدى الربيعي والانحراف الربيعي نكون جدول التوزيع التكراري:

فئات الساعات	مركز	$f1$	$f2$	ب $f3$	ت ص 1	ت ص 2	ت ص 3	النوع الأول $(x_i - \overline{X})^2$	النوع الثاني $(x_i - \overline{X})^2$	النوع الثالث $(x_i - \overline{X})^2$
3-1	2	2	7	2	2	7	2	38.44	24.21	0.58
5-3	4	4	9	7	6	16	9	17.64	8.53	1.54
7-5	6	6	10	9	12	26	18	4.84	0.85	5.02
9-7	8	18	9	14	30	35	32	0.04	122.77	1.54
11-9	10	15	8	13	45	43	45	3.24	65.29	0.06
13-11	12	5	7	5	50	50	50	14.44	3.69	0.58
المجموع		50	50	50						
الوسط الحسابي لكل نوع		8.2	6.9	7.8						

الفئات	النوع الأول $f(x_i - \overline{X})^2$	النوع الثاني $f(x_i - \overline{X})^2$	النوع الثالث $f(x_i - \overline{X})^2$	النوع الأول $f(x_i - \overline{X})^4$	النوع الثاني $f(x_i - \overline{X})^4$	النوع الثالث $f(x_i - \overline{X})^4$
1	76.88	169.44	1.16	1477.63	585.95	0.33
2	70.56	76.74	10.76	311.17	72.70	2.36
3	29.04	8.46	45.16	23.43	0.72	25.18
4	0.72	1104.90	21.53	0.00	15071.59	2.36
5	48.60	522.29	0.75	10.50	4262.31	0.00
6	72.20	25.80	2.89	208.51	13.59	0.33
المجموع	298.00	1907.64	82.24	2031.24	20006.86	30.58

المدى الربيعي والانحراف الربيعي للنوع الأول:

- نحسب الربيع الأول: $Q_1 = A + \frac{\frac{n}{4}\times 1 - f_1}{f_2 - f_1}\times L = 7 + \frac{12.5-12}{30-12}\times 2 = 7.06$
- نحسب الربيع الثالث: $Q_3 = A + \frac{\frac{n}{4}\times 3 - f_1}{f_2 - f_1}\times L = 9 + \frac{37.5-30}{45-30}\times 2 = 10$

بالنسبة للنوع الأول: المدى الربيعي: $IQ = Q_3 - Q_1 = 2.94$ والانحراف الربيعي:

$$Q = \frac{IQ}{2} = 1.47$$

المدى الربيعي والانحراف الربيعي للنوع الثاني :

- نحسب الربيع الأول: $Q_1 = A + \frac{\frac{n}{4}\times 1 - f_1}{f_2 - f_1}\times L = 3 + \frac{12.5-7}{16-7}\times 2 = 4.22$
- نحسب الربيع الثالث: $Q_3 = A + \frac{\frac{n}{4}\times 3 - f_1}{f_2 - f_1}\times L = 9 + \frac{37.5-35}{43-35}\times 2 = 9.63$

بالنسبة للنوع الثاني: المدى الربيعي: $IQ = Q_3 - Q_1 = 5.41$ والانحراف الربيعي:

$$Q = \frac{IQ}{2} = 2.75$$

المدى الربيعي والانحراف الربيعي للنوع الثالث:

- نحسب الربيع الأول: $Q_1 = A + \frac{\frac{n}{4}\times 1 - f_1}{f_2 - f_1}\times L = 5 + \frac{12.5-9}{18-9}\times 2 = 5.8$
- نحسب الربيع الثالث: $Q_3 = A + \frac{\frac{n}{4}\times 3 - f_1}{f_2 - f_1}\times L = 9 + \frac{37.5-32}{45-32}\times 2 = 9.8$

بالنسبة للنوع الثالث: المدى الربيعي: $IQ = Q_3 - Q_1 = 4$ والانحراف الربيعي:

$$Q = \frac{IQ}{2} = 2$$

2- حساب الانحراف المعياري:

بالنسبة للنوع الأول: $s=\sqrt{s^2}=\sqrt{\dfrac{\sum f\left(x_i-\overline{X}\right)^2}{n}}=\sqrt{\dfrac{298}{50}}=2.44$

بالنسبة للنوع الثاني: $s=\sqrt{s^2}=\sqrt{\dfrac{\sum f\left(x_i-\overline{X}\right)^2}{n}}=\sqrt{\dfrac{1907.64}{50}}=6.18$

بالنسبة للنوع الثالث: $s=\sqrt{s^2}=\sqrt{\dfrac{\sum f\left(x_i-\overline{X}\right)^2}{n}}=\sqrt{\dfrac{82.24}{50}}=1.28$

ونعرف أن الانحراف المعياري من عيوبه أنه يتأثر بالقيم الشاذة ومن مزاياه أنه يعتمد على جميع القيم، بينما الانحراف الربيعي يمتاز بعدم تأثره بالقيم الشاذة ويعاب عليه اعتماده على قيمتين فقط هما الربيع الأول والثالث واللتان تعتمدان بدورهما على قيمتين فقط

رغم أن مقياس التشتت: الانحراف الربيعي كان في النوع الأول أصغر ويبين أن أحسن نوع هو النوع الأول وبما أنه لا توجد قيم شاذة واضحة فمقياس الانحراف المعياري يبين أن أحسن نوع هو النوع الثالث.

3- شكل الالتواء بطريقة بيرسون:

بالنسبة للنوع الأول فإن: $Med=\text{A}+\dfrac{\frac{n}{2}-f_1}{f_2-f_1}\times L=7+\dfrac{25-12}{30-12}\times 2=8.44$

$$a=\frac{3(\overline{X}-Med)}{s}=\frac{3(8.2-8.44)}{2.44}=-0.3$$

$\alpha<0$ إذن فإن منحنى التوزيع التكراري ملتوي جهة اليسار (سالب الالتواء).

بالنسبة للنوع الثاني فإن: $Med=\text{A}+\dfrac{\frac{n}{2}-f_1}{f_2-f_1}\times L=5+\dfrac{25-16}{26-16}\times 2=6.8$

$$a = \frac{3(\overline{X} - Med)}{s} = \frac{3(6.9 - 6.8)}{6.18} = 0.05$$

$\alpha > 0$ إذن فإن منحنى التوزيع التكراري ملتوي جهة اليمين (موجب الالتواء).

بالنسبة للنوع الثالث فإن: $Med = \mathrm{A} + \frac{\frac{n}{2} - f_1}{f_2 - f_1} \times L = 7 + \frac{25 - 18}{32 - 18} \times 2 = 8$

$$a = \frac{3(\overline{X} - Med)}{s} = \frac{3(8 - 7.8)}{1.28} = 0.47$$

$\alpha > 0$ إذن فإن منحنى التوزيع التكراري ملتوي جهة اليمين (موجب الالتواء).

شكل تفرطح البيانات:

- معامل التفرطح k بالنسبة للنوع الأول:

$$k = \frac{\frac{1}{n}\sum f(x_i - \overline{X})^4}{S^4} = \frac{\frac{1}{50} \times 203124}{2.44^4} = 1.14$$

- معامل التفرطح k بالنسبة للنوع الثاني:

$$k = \frac{\frac{1}{n}\sum f(x_i - \overline{X})^4}{S^4} = \frac{\frac{1}{50} \times 2000686}{6.18^4} = 0.27$$

- معامل التفرطح k بالنسبة للنوع الثالث:

$$k = \frac{\frac{1}{n}\sum f(x_i - \overline{X})^4}{S^4} = \frac{\frac{1}{50} \times 30.58}{1.28^4} = 0.23$$

نلاحظ أن k>3 بالنسبة للتوزيعات التكرارية للأنواع الثلاثة، وبالتالي فإن المنحنيات كلها منبسطة أي مفرطحة.

السلسلة الخامسة (الانحدار):

التمرين الأول:

إذا كان انحدار الرضا الوظيفي لدى العامل على أجره الشهري ممثلا بالمعادلة التقديرية التالية:

$$Y = B_0 + B_1 X + e$$

المطلوب:

1- فسر كل متغير ومعلمة في المعادلة السابقة.

2- مالفائدة من استخدام المعادلات التقديرية لانحدار العلاقات بين المتغيرات.

3- إذا كان: $B_1 = 0.8 \quad B_0 = 2$ أرسم خط الانحدار.

التمرين الثاني:

فيما يلي عدد ساعات المذاكرة يوميا لـ 8 طلاب وعلاماتهم في مقياس معين:

الطلبة	1	2	3	4	5	6	7	8
عدد ساعات المذاكرة	2	3	4	5	6	8	9	7
علامة الاختبار	12	13	14	15	16	17	19	20

المطلوب:

1- حدد المتغير المستقل والمتغير التابع.

2- أوجد المعادلة التقديرية لانحدار x على y.

3- ماذا تمثل كل معلمة(B_1B_2).

4- إذا كان إذا كان عدد ساعات المذاكرة للطالب هو 10، فقدر علامة اختبار الطالب.

التمرين الثالث:

فيما يلي المساحة المزروعة بالأعلاف بالألف هكتار وكمية إنتاج اللحوم بالألف طن في الفترة (1995–2002):

السنوات	1995	1996	1997	1998	1999	2000	2001	2002
المساحة المزروعة	305	313	297	289	233	214	240	217
كمية اللحوم المنتجة	592	603	662	607	635	699	719	747

المطلوب:

1. حدد المتغير المستقل والمتغير التابع.
2. أوجد المعادلة التقديرية لانحدار x على y.
3. ماذا تمثل كل معلمة(B_1B_2).
4. إذا كانت المساحة المزروعة في سنة 2010 تساوي 400 ألف هكتار، فقدر كمية اللحوم المنتجة سنة 2010.

السلسلة السادسة (الارتباط):

التمرين الأول:

فيما يلي عدد ساعات المذاكرة يوميا لـ 8 طلاب وعلاماتهم في مقياس معين:

الطلبة	1	2	3	4	5	6	7	8
عدد ساعات المذاكرة	2	3	4	5	6	8	9	7
علامة الاختبار	12	13	14	15	16	17	19	20

المطلوب:

1. أوجد قيمة معامل الارتباط البسيط بين عدد ساعات المذاكرة وعلامة الاختبار؟.
2. ما هي درجة قوة العلاقة بينهما وما نوعها؟.

التمرين الثاني:

فيما يلي المساحة المزروعة بالأعلاف بالألف هكتار وكمية إنتاج اللحوم بالألف طن في الفترة (1995–2002):

السنوات	1995	1996	1997	1998	1999	2000	2001	2002
المساحة المزروعة	305	313	297	289	233	214	240	217
كمية اللحوم المنتجة	592	603	662	607	635	699	719	747

المطلوب:

1. أوجد قيمة معامل الارتباط البسيط بين المساحة المزروعة وكمية اللحوم المنتجة؟.

2. ما هي درجة قوة العلاقة بينهما وما نوعها؟.

<u>التمرين الثالث</u>:

فيما يلي إجابات 12 طالب حول الأداء التعليمي لأستاذين (نوعين من الإجابات: ملاحظات وعلامات):

لما كان السؤال عن **<u>إعطاء علامة</u>** لكل أستاذ من طرف الطلبة كانت **<u>العلامات</u>** كما فـــي الجدول الأول.

ولما كان السؤال عن **<u>رأي الطلبة</u>** حول أداء الأستاذين كانت **<u>الملاحظات</u>** كما في الجدول الثاني.

الجدول الأول

الطلبة	أستاذ 1	أستاذ 2
1	F	B
2	A	H
3	L	J
4	H	L
5	C	G
6	J	A
7	B	K
8	I	F
9	G	I
10	D	C
11	E	E
12	K	D

الجدول الثاني

الطلبة	أستاذ 1	أستاذ 2
1	ضعيف	حسن
2	جيد	ضعيف
3	ممتاز	ممتاز
4	متوسط	جيد
5	حسن	ضعيف
6	جيد	ممتاز
7	حسن	ممتاز
8	ممتاز	جيد
9	جيد	ضعيف
10	ضعيف	جيد
11	ضعيف	حسن
12	جيد	حسن

أحسب معامل الارتباط وحدد نوع العلاقة وقوتها؟. (في كل من الجدولين)

حل السلسلتين الخامسة والسادسة (الارتباط والانحدار):

حل التمرين الأول:

1. المتغير المستقل هو المساحة المزروعة والتي تؤثر في كمية إنتاج اللحوم وهو المتغير التابع:

$\sum y^2$	$\sum x^2$	$\sum x_i y_i$	الكمية Y	المساحة X	السنوات
350464.00	93025.00	180560.00	592.00	305.00	1995
363609.00	97969.00	188739.00	603.00	313.00	1996
438244.00	88209.00	196614.00	662.00	297.00	1997
368449.00	83521.00	175423.00	607.00	289.00	1998
403225.00	54289.00	147955.00	635.00	233.00	1999
488601.00	45796.00	149586.00	699.00	214.00	2000
516961.00	57600.00	172560.00	719.00	240.00	2001
558009.00	47089.00	162099.00	747.00	217.00	2002
3487562.00	567498.00	1373536.00	5264.00	2108.00	المجموع

لدينا الوسط الحسابي لـ x: $\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{2108}{8} = 263.5$

لدينا الوسط الحسابي لـ y: $\overline{y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{5264}{8} = 658$

- نستطيع حساب معامل الارتباط بطريقة بيرسون بدون حساب الوسط الحسابي:

$$R_p = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - \sum (x_i)^2}\sqrt{n\sum y_i^2 - \sum (y_i)^2}} = \frac{8(1373536) - (11096512)}{13556495} = -0.789$$

- نلاحظ أن معامل الارتباط سالب وبالتالي فالعلاقة عكسية بين المساحة المزروعة كمية إنتاج اللحوم.
- نلاحظ من قيمة معامل الارتباط أن العلاقة قوية بين المساحة المزروعة وكمية إنتاج اللحوم.

2. إيجاد المعادلة التقديرية لانحدار y على x: وذلك بإيجاد المعلمات.

$$\beta_1 = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2} = \frac{-108224}{190800} = -0.57$$

$$\beta_0 = \frac{\sum y_i - \beta_1 \sum x_i}{n} = \frac{5264 + 0.57(2108)}{8} = 807.46$$

إذن المعادلة التقديرية لانحدار x على y هي: $y = \beta_0 + \beta_1 x = 807.46 - 0.57x$

3. إذا كانت المساحة المزروعة سنة 2003 تساوي 400 ألف هكتار، فإن:

$$y = 807.46 - 0.57(400) = 580.58$$

كمية اللحوم المنتجة سنة 2010 هي: 580.58 ألف طن.

حل التمرين الثاني:

1. المتغير المستقل هو عدد ساعات المذاكرة والتي تؤثر في علامة الطالب وهو المتغير التابع:

الطالب	الساعات x	العلامة y	$\sum x_i y_i$	$\sum x^2$	$\sum y^2$
1	2	12	24.00	4.00	144.00
2	3	13	39.00	9.00	169.00
3	4	14	56.00	16.00	196.00
4	5	15	75.00	25.00	225.00
5	6	16	96.00	36.00	256.00
6	8	17	112.00	64.00	196.00
7	9	19	171.00	81.00	361.00
8	7	20	140.00	49.00	400.00
المجموع	44.00	123.00	713.00	284.00	1947.00

لدينا الوسط الحسابي لـx: $\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{44}{8} = 5.5$

لدينا الوسط الحسابي لـy: $\overline{y} = \frac{\sum_{i=1}^{n} y_i}{n} = \frac{123}{8} = 15.38$

- نستطيع حساب معامل الارتباط بطريقة بيرسون بدون حساب الوسط الحسابي:

$$R_p = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - \sum (x_i)^2}\sqrt{n\sum y_i^2 - \sum (y_i)^2}} =$$

$$\frac{8(713) - (5412)}{387.55} = 0.7535$$

- نلاحظ أن معامل الارتباط موجب وبالتالي فالعلاقة طردية بين ساعات الاســـتذكار وعلامة الاختبار.
- نلاحظ من قيمة معامل الارتباط أن العلاقة قوية بين المساحة المزروعة وكمية إنتاج اللحوم.

2. إيجاد المعادلة التقديرية لانحدار x على y: وذلك بإيجاد المعالم.

$$\beta_1 = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - \sum (x_i)^2} = \frac{292}{447} = 0.65$$

$$\beta_0 = \frac{\sum y_i - \beta_1 \sum x_i}{n} = \frac{123 - 0.65(44)}{8} = 11.78$$

إذن المعادلة التقديرية لانحدار x على y هي: $y = \beta_0 + \beta_1 x = 11.78 - 0.65x$

3. إذا كانت المساحة المزروعة سنة 2003 تساوي 400 ألف هكتار، فإن

$$y = 11.78 - 0.65(10) = 18.31$$

علامة الطالب المقدرة لو ذاكر 10 ساعات هي: 18.31.

<u>حل التمرين الثالث</u>:

1. حساب معامل الارتباط وتحدد نوع العلاقة وقوتها في الجدول الأول:

- أولا ترتيب العلامات كما يلي:

ترتيب العلامات للأستاذ 1	رتب العلامات	ترتيب العلامات لأستاذ 2	رتب العلامات
A	1	A	1
B	2	B	2
C	3	C	3
D	4	D	4
E	5	E	5
F	6	F	6
G	7	G	7
H	8	H	8
I	9	I	9
J	10	J	10
K	11	K	11
L	12	L	12

– تكوين جدول الرتب لحساب مجموع مربعات الفروق بين رتب الملاحظات: $\sum d^2$

الطلبة	أستاذ 1	أستاذ 2	رتب 1	رتب2	d	d^2
1	F	B	6	2	4	16
2	A	H	1	8	7	49
3	L	J	12	10	2	4
4	H	L	8	12	4	16
5	C	G	3	7	4	16
6	J	A	10	1	9	81
7	B	K	2	11	9	81
8	I	F	9	6	3	9
9	G	I	7	9	2	4
10	D	C	4	3	1	1
11	E	E	5	5	0	0
12	K	D	11	4	7	49
مجموع						326

–ونحسب المعامل كما يلي: $R_S = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6\times 326}{12(144-1)} = -0.1399$

فنقول أن العلاقة بين الملاحظات للأستاذين عكسية وضعيفة جدا.

2. حساب معامل الارتباط وتحدد نوع العلاقة وقوتها في الجدول الثاني:

– أولا ترتيب الملاحظات كما يلي:

الطلبة	ترتيب الملاحظات للأستاذ 1	رتب الملاحظات	ترتيب الملاحظات لأستاذ 2	رتب الملاحظات
1	ضعيف		ضعيف	1.5
2	ضعيف	2	ضعيف	
3	ضعيف		حسن	3.5
4	حسن	4.5	حسن	
5	حسن		متوسط	
6	متوسط	6	متوسط	6
7	جيد		متوسط	
8	جيد	8.5	جيد	8.5
9	جيد		جيد	
10	جيد		ممتاز	
11	ممتاز	11.5	ممتاز	11
12	ممتاز		ممتاز	

- تكوين جدول الرتب لحساب مجموع مربعات الفروق بين رتب الملاحظات: $\sum d^2$

الطلبة	أستاذ 1	أستاذ 2	رتب 1	رتب 2	d	d^2
1	ضعيف	حسن	2	3.5	1.5	2.25
2	جيد	متوسط	8.5	6	2.5	6.25
3	ممتاز	ممتاز	11.5	11	0.5	0.25
4	متوسط	متوسط	6	6	0	0
5	حسن	ضعيف	4.5	1.5	3	9
6	جيد	ممتاز	8.5	11	2.5	6.25
7	حسن	ممتاز	4.5	11	6.5	42.25
8	ممتاز	جيد	11.5	8.5	3	9
9	جيد	ضعيف	8.5	1.5	7	49
10	ضعيف	جيد	2	8.5	6.5	42.25
11	ضعيف	متوسط	2	6	4	16
12	جيد	حسن	8.5	3.5	5	25
مجموع						207.50

-ونحسب المعامل كما يلي: $R_S = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6\times 207.5}{12(144-1)} = 0.2745$

فنقول أن العلاقة بين الملاحظات للأستاذين طردية وضعيفة جدا.

السلسلة السابعة (الانحدار والارتباط):

التمرين الأول:

تقدير علامة الاقتصاد	تقدير علامة الإحصاء
A+	A
D	C+
C	D
C	D+
A	B+
B	C+
B+	A+
B	B
C	B+
B	B+

المطلوب:

أوجد نوع وقوة العلاقة بين تقدير العلامات في المقياسين؟.

(0.703)

<u>التمرين الثاني:</u>

أراد طالب أن يدرس العلاقة الارتباطية بين عدد المتخرجين من قسم العلوم الاقتصادية في جامعة ما والذين تمكنوا من العمل مباشرة فرصد الطالب البيانات التالية خلال 7 سنوات:

عدد العاملين	عدد المتخرجين	السنوات
10	74	2000
9	65	2001
12	79	2002
15	77	2003
12	69	2004
16	82	2005
17	80	2006

المطلوب:

أوجد نوع وقوة العلاقة بين عدد المتخرجين ومن تمكنوا من العمل مباشرة ؟ (0.79).

التمرين الثالث:

1- قدر علامة أحمد في اختبار ما إذا كانت نسبة توتره هي 94، حيث عند دراسة العلاقة بين (X) نسبة التوتر و(Y) علامات 8 طلاب حصلنا على:

$\sum x^2 = 750$ ، $\sum xy = 510$ ، $\sum x = 240$ ،

$\sum y = 195$ ، $\sum y^2 = 300$.

2- أوجد نوع العلاقة ودرجة قوتها بين التوتر وعلامة الاختبار، وأوجد نسبة تأثير X في Y.

(0.99-ـــــ77.4-)

التمرين الرابع:

1- هل يوجد ارتباط بين علامات تحصيل 8 طلبة في مقياسي الإحصاء والمحاسبة . إذا كانت نتائجهم كما يلي:

إحصاء	13	9	19	15	11	8	16	11
المحاسد	15	7	17	15	10	9	14	10

بة								

2- هل يوجد ارتباط بين علامات تحصـــيل 8 طلبـــة فـــي مقياســـي الإحصـــاء والمحاسبة. إذا كانت نتائجهم كما يلي:

إحصاء	D	F	A	C	E	G	B	E
المحاسبة	B	F	A	B	D	E	C	D

الحلول (0.94 0.89)

<u>تابع للسلسلة السابعة (الانحدار والارتباط):</u>

<u>التمرين الأول:</u>

يمثل الجدول التالي توزيع عدد سنوات الخدمة لعينة من 7 عمال، وتقابل عـــدد سنوات الخدمة لكل عامل نسبة رضاه عن منصبه:

عدد سنوات الخدمة	نسبة الرضا الوظيفي

93	11
80	4
88	9
85	5
83	8
80	6
91	7

المطلوب:

1- أوجد المعادلة التقديرية لانحدار الرضا على سنوات الخدمة؟.

2- أرسم خط الانحدار من خلال معادلة الانحدار؟.

3- اشرح كلا من B_1 و B_0؟.

4- حدد نوع العلاقة بين المتغيرين؟.

5- أوجد نوع العلاقة ودرجة قوتها بين المتغيرين باستخدام معامل الارتباط المناسب؟.

6- أوجد نوع العلاقة ودرجة قوتها بين المتغيرين باستخدام معامل ارتباط سبيرمان للرتب؟. (ماذا تلاحظ علما أن معامل سبيرمان حالة خاصة من معامل بيرسون).

7- حدد نسبة تأثير عدد سنوات الخدمة في الرضا الوظيفي ونسبة تأثير باقي المتغيرات غير المدروسة؟.

<u>التمرين الثاني:</u>

أراد أحد طلبة العلوم الاقتصادية سنة أولى LMD بالجامعة أن يدرس العلاقة بين عدد أو حجم الطلبة في الفوج ونسبة الذين يواصلون دراساتهم العليا فحصل على البيانات في الجدول المقابل:

عدد الطلبة	نسبة المواصلين للدراسات العليا
30	3
25	4
20	1
15	7
10	5
34	2
29	4
26	8

المطلوب:

1- أوجد المعادلة التقديرية للانحدار.

2- أرسم خط الانحدار من خلال معادلة الانحدار.

3- اشرح كلا من B_1 و B_0.

4- حدد نوع العلاقة بين المتغيرين.

5- أوجد نوع العلاقة ودرجة قوتها بين المتغيرين باستخدام معامل الارتباط المناسب.

6- أوجد نوع العلاقة ودرجة قوتها بين المتغيرين باستخدام معامل ارتباط سبيرمان للرتب (ماذا تلاحظ علما أن معامل سبيرمان حالة خاصة من معامل بيرسون).

7- حدد نسبة تأثير حجم أو عدد طلبة الفوج في نسبة المواصلين لدراساتهم العليا ونسبة تأثير باقي المتغيرات غير المدروسة.

السلسلة الثامنة

(الارتباط بين المتغيرات الوصفية ذات النوع الإسمي 2*2):

التمرين الأول:

لتحديد مدى علاقة العمل بالتدخين أي هل يرتبط تدخين الأفراد بالعمل أم لا وإلى أي درجة تم رصد النتائج التالية.

التدخين / العمل	يدخن	لا يدخن	$\sum$
يعمل	20	20	40
لا يعمل	50	10	60
$\sum$	70	30	100

(الحل 0.34)

التمرين الثاني:

أراد طبيب أن يحدد مدى وجود علاقة بين التطعيم ومقاومة مرض معين في عينة من 60 مريضا، حيث وجد:

-من كل المرضى في العينة تم تطعيم 40 منهم فقط.
-من كل المرضى 20 فقط قاوم الأمراض.
مع العلم أن من بين الذين تم تطعيمهم يوجد 15 منهم قاوم الأمراض.

التدخين / العمل	تم تطعيمهم	لم يتم تطعيم	$\sum$
قاوم	15	5	20
لم يقاوم	25	15	40
$\sum$	40	20	60

هل توجد علاقة بين التطعيم ومقاومة الأمراض في هذه العينة؟.

(الحل 0.12)

التمرين الثالث:

في عينة من 100 فرد وعند تصنيفهم وجدنا:

- حسب متغير الجنس: 40 رجلا و60 امرأة.
- حسب التوجه السياسي: 70 وطنيا.

علما أن عدد الرجال الذين توجههم وطني هو 30.

المطلوب:

2- أوجد قوة العلاقة بين المتغيرين؟.

(الحل 0.09)

التمرين الرابع:

اخترنا 50 طالبا من الجامعة وعند تصنيفهم وجدنا:

- حسب متغير الجنس: 15 ذكور.
- حسب متغير فرع التخصص: 20 علميا.
- حسب متغير الحضور والغياب: 10 يتغيبون.

علما أن عدد الإناث في الفروع العلمية هو 10.

علما أن معدل الإناث اللواتي يتغيبن هو 6.

علما معدل الغيابات في الفروع العلمية هو 8.

المطلوب:

- أوجد قوة العلاقة بين متغيري الجنس وفرع التخصص؟.
- أوجد قوة العلاقة بين متغيري الجنس والغياب؟.

- أوجد قوة العلاقة بين متغيري الفرع والغياب؟.

(الحلول 0.16 0.00 0.20)

السلسلة التاسعة

(الارتباط بين المتغيرات ذات النوع الوصفي

الإسمي أكثر من 2*2)

التمرين الأول:

في الجدول التالي بيانات لعينة من 80 أسرة حسب متغيري عمل الأم والتحصيل العلمي للأبناء.

التحصيل / عمل الأم	جيد	متوسط	ضعيف	$\sum$
تعمل	6	5	4	15
لا تعمل	30	15	20	65
$\sum$	31	25	24	80

(ملاحظة هذه البيانات وهمية ولا تمد للواقع بصلة)

المطلوب:

أوجد العلاقة بين عمل الأم والتحصيل العلمي للأبناء؟

التمرين الثاني:

أوجد العلاقة بين التدخين والمستوى التعليمي؟. في البيانات التالية لـــــ 200 فرد:

التدخين / المستوى	يدخن	لا يدخن	$\sum$
جامعي	30	40	70
ثانوي	40	20	60
أساسي	10	40	50
ابتدائي	15	5	20
$\sum$	95	105	200

التمرين الثالث:

أوجد العلاقة بين لون الزهور وقوة الرائحة؟. لـــــــ 30 زهــرة حســب بيانــات الجدول التالي:

لون الزهرة / الرائحة	أصفر	أبيض	أحمر	$\sum$
قوية	3	5	3	11
متوسطة	4	2	4	10
ضعيفة	4	1	4	9
	11	8	11	30

التمرين الرابع:

اخترنا 100 فرد من عنابة ووهران وعند تصنيفهم وجدنا:

-حسب متغير (الهجرة غير الشرعية): 20 ينوون الهجرة غير الشرعية.

-حسب المستوى المعيشي: 60 منخفض، 30 متوسط، 10 مرتفع.

- عدد من مستواهم المعيشي متوسط وينوون الهجرة غير الشرعية هو 3.
- عدد من مستواهم المعيشي منخفض وينوون الهجرة غير الشرعية هو16.

المطلوب:

1- أوجد قوة العلاقة بين متغيري المستوى المعيشي و الهجرة غير الشرعية؟.

(الحل 0.20)

<u>التمرين الخامس</u>:

درسنا 80 أسرة في مدينة ورقلة من خـــلال متغيـــرين همـــا المســـتوى المعيشـــي والتحصيل العلمي للأبناء (قيم افتراضية) فوجدنا:

- حسب التحصيل العلمي: 30 جيد، 20 متوسط، 30 ضعيف.
- حسب المستوى المعيشي: 50 منخفض، 20 متوسط، 10 مرتفع.
  - عدد ذوي التحصيل الجيد والمستوى المعيشي المنخفض هو17.
  - عدد ذوي التحصيل المتوسط والمستوى المعيشي المتوسط هو04.
  - عدد ذوي التحصيل الجيد والمستوى المعيشي المرتفع هو 01.
  - عدد ذوي التحصيل المتوسط والمستوى المعيشي المرتفع هو03

1- أوجد قوة العلاقة بين متغيري المستوى المعيشي والتحصيل العلمي؟. (الحل 0.31).

السلسلة العاشرة

(الارتباط بين المتغيرات ذات النوع الوصفي

الإسمي أكثر من 2*2)

المسألة:

أخذنا عينة من 50 عاملا في الجامعة وعند تصنيفهم وجدنا:

- حسب متغير الجنس: 20 ذكور.
- حسب متغير المهنة: 5 رؤساء أقسام و15 أستاذا و30 إداريا.
- حسب المستوى التعليمي: 4 دكاترة 16 ماجستير و30 ليسانس.
  - مع العلم أن عدد الأساتذة الذكور هو 5.
  - عدد رؤساء الأقسام الذكور هو 4.
  - عدد الذكور ذوي المستوى ماجستير هو 6.
  - عدد الإناث ذوات المستوى ليسانس هو 19.
  - عدد رؤساء الأقسام ذوي المستوى دكتور هو 4.
  - عدد رؤساء الأقسام ذوي المستوى ماجستير هو 1.
  - عدد الأساتذة ذوي المستوى ماجستير هو 13.
  - عدد الإداريين ذوي المستوى ليسانس هو 16.

المطلوب:

1- أوجد قوة العلاقة بين متغيري الجنس والمهنة؟.

2- أوجد قوة العلاقة بين متغيري الجنس والمستوى التعليمي؟.

3- أوجد قوة العلاقة بين متغيري المهنة والمستوى التعليمي؟.

(الحلول 0.26 0.21 0.76)

<u>سلسلة التوزيع الطبيعي (المعياري):</u>

<u>تمرين رقم 01</u>:

أوجد المساحة تحت التوزيع الطبيعي المعياري (القياسي)؟. في القيم التالية:

($-1<Z<+1$) ($-2<Z<+2$) ($-3<Z<+3$)

($0<Z<0.88$) ($1.6<Z<2.55$) ($Z<1.6$)

($2.55<Z$) **خارج المجال** ($-1.6<Z<2.55$).

<u>تمرين رقم 02</u>: (استعمل قاعدة التحويل)

إذا كان المتوسط الحسابي لأوزان الطلبة هو 70 بانحراف معياري 07:

- ما هو احتمال أن يكون وزن الطالب أقل من 80.

- ما هو احتمال أن يكون وزن الطالب أقل من 84.

- ما هو احتمال أن يكون وزن الطالب أقل من 63.

- ما هي نسبة الطلبة الذين تتراوح أوزانهم بين 55 و85.

<u>تمرين رقم 03</u>: (استعمل قاعدة التحويل)

يتبع دخل الأسر توزيعا طبيعيا بوسط حسابي 30.000دج وانحراف معيـــاري 9.000دج.

حددت قيمة 12.000دج كعتبة للفقر، وأن دخل الفئة المتوسطة يتـــراوح بـــين 15.000دج و60.000. وأكثر من 60.0000 هم فئة الأغنياء.

- فما هي نسبة كل فئة في المجتمع؟.

<u>تمرين رقم 04</u>: (استعمل قاعدة التحويل)

كان لدينا متوسط أطوال الطلبة في كلية الرياضة 165 سم وانحراف معيـــاري 5سم.

- الطول المطلوب في رياضة كرة القدم بين 157سم و172سم.
- الطول المطلوب في كرة السلة هو أكثر من 172سم.
- باقي الطلبة ينضمون لرياضة كرة اليد.

إذا علمت أن عدد الطلبة في الكلية هو 4000 طالب فما هو عدد الطلبة الـــذين سيشاركون في الرياضات التالية (كرة القدم، كرة السلة، كرة اليد، الشطرنج).

قائمة المراجع

- أحمد فاروق عبد العظيم , الإحصاء , المكتب الجامعي الحديث , الإسكندرية , 1984.
- دومنيك سالفتور , الإحصاء والاقتصاد القياسي , سلسلة ملخصات شوم , 1982.
- عز الدين جوني , الإحصاء الاقتصادي , ديوان المطبوعات الجامعية , الجزائر ,1983.
- محمد صبحي أبو صالح , مقدمة في الإحصاء , ديوان المطبوعات الجامعية , الجزائر , 1984.
- محمد كلاس , محاضرات في الإحصاء التطبيقي , ديوان المطبوعات الجامعية , الجزائر , 1984.