جامعة غليزان

كلية العلوم الإقتصادية، التجارية وعلوم التسيير

قسم علوم التسيير

مطبوعة بيداغوجية بعنوان :

الإحصـــاء الوصفي (احصاء1)

-محاضرات وتمارين-

موجهة لطلبة السنة الأولى جذع مشترك علوم اقتصادية، تسيير وعلوم تجارية

من إعداد

د. مبرك إبراهيم

السنة الجامعية:2023-2024

جامعة غليزان
كلية العلوم الإقتصادية، التجارية وعلوم التسيير
قسم علوم التسيير

مطبوعة بيداغوجية بعنوان :

الإحصاء الوصفي (احصاء1)
-محاضرات وتمارين-

موجهة لطلبة السنة الأولى جذع مشترك علوم اقتصادية، تسيير وعلوم تجارية

من إعداد

د. مبرك إبراهيم

السنة الجامعية:2023-2024

الفهرس العام

مقدمة:

مما لا شك فيه أن التحليل الاحصائي يحظى بمستوى عال من الاهتمام من قبل الباحثين في شتى العلوم والمجالات المختلفة، وهو وسيلة هامة وأساسية في انجاز الأبحاث والدراسات للإجابة عن تساؤلاتها وفرضياتها، وهو جزء لا يتجزأ منها، فلا تكاد تخلو أي دراسة أو بحث علمي من استخدام أساليب التحليل الاحصائي المختلفة.

ونظرا لأهمية دراسة الإحصاء في مجال العلوم الاقتصادية والعلوم التجارية وعلوم التسيير، ارتأينا تقديم هذه المطبوعة لطلبة السنة الأولى جذع مشترك، والهدف من هذه المادة هو تقديم الأسس العامة للإحصاء الوصفي، والتي تفيد الطالب في اكتساب مهارة اختيار الأساليب المناسبة لوصف البيانات وتحليلها، وهذا بالتعرف على الطرق والأساليب الاحصائية التي تساعدهم في مجال تخصصهم على اتخاذ القرارات المناسبة، من خلال تعريف بأنواع البيانات ووصفها، وعرضها وتلخيصها في شكل جداول ورسومات بيانية وتحليلها وصولا إلى تحليل واستقراء النتائج.

الفصل الأول: مفاهيم عامة

من المفاهيم الشائعة بين الناس عن الإحصاء، ما هي إلا أرقام وبيانات رقمية فقط، كأعداد السكان، وأعداد المواليد، وأعداد الوفيات، وأعداد المزارعين، وأعداد المزارع، وخلافه، ومن ثم ارتبط مفهوم الناس عن الإحصاء بأنه عد أو حصر الأشياء والتعبير عنها بأرقام، وهذا هو المفهوم المحدود لعلم الإحصاء، ولكن الإحصاء كعلم، هو الذي يهتم بطرق جمع البيانات، وتبويبها، وتلخيصها بشكل يمكن الاستفادة منها في وصف البيانات وتحليلها للوصول الى قرارات سليمة في ظل ظروف عدم التأكد.

1) **تعريف علم الإحصاء :**هناك تعريفات عديدة لعلم الاحصاء منها قديم ومنها شامل وحديث وقريب الى الدقة في البحث العلمي، حيث يمكن تعريفه على أنه عبارة عن العد والحساب، أو عبارة عن مجموعة الطرق المستعملة في جمع البيانات والمشاهدات وطرق عرض هذه البيانات وتلخيصها."

كما يعرف ايضا بأنه" عبارة عن مجموعة الطرق المستعملة في تحليل البيانات الاحصائية المتوفرة، واتخاذ القرارات الحكيمة في مواجهة الظواهر العشوائية التي تحيط بنا."

يعتبر الاحصاء علم كبقية العلوم لأنه يمتاز بالمراحل الاربعة التي تمتاز بها بقية العلوم وهي:

أ) **المشاهدة :**العالم يشاهد ما يحدث، وبدون الحقائق المتعلقة بالمشكلة التي يود أن يدرسها .

ب) **الفرضية :**لتفسير الحقائق المتعلقة بالمشكلة التي يود العالم ان يدرسها ويصوغ ما في ذهنه على شكل فرضيات تعبر على تحتويه البيانات التي جمعها.

ج) **التنبؤ :**يستنتج العالم من فرضياته بعض الحقائق .

د) **التحقق :**يقوم العالم بجمع بيانات جديدة ويضع فرضيات جديدة وباستنتاج حقائق جديدة للتأكد من صحة تنبؤه.

2) **التعريف بالمصطلحات الاحصائية:**

أ) **المجتمع الاحصائي Population:**يعرف بأنه مجموعة ذات خصائص مشتركة من الاشياء أو المفردات ذات أهمية خاصة لدراسة علمية .ويقسم المجتمع الاحصائي الى قسمين:

المجتمع المحدود :يعتبر المجتمع محدودا إذا كان بالإمكان حصر جميع وحدات الدراسة فيه.

المجتمع غير المحدود :في المجتمع غير المحدود فان أسلوب دراسة جميع وحدات المجتمع والذي يطلق عليه **بأسلوب الحصر الشامل** يصبح مستحيلا .كذلك الحال في بعض المجتمعات المحدودة والتي لا يقبل المنطق تطبيق أسلوب الحصر الشامل :**مثلا :**فحص دم شخص، حيث لا يمكن سحب جميع دمه مما يؤدي إلى هلاكه(موته) لذا فالأسلوب هنا يكمن في تبني **أسلوب المعاينة.**

ب) **الوحدة الاحصائية :**هي العنصر الأولي محل الدراسة الإحصائية، أو هي القيمة المادية أو المعنوية التي تقع عليها الدراسة الإحصائية، **مثل :**الطالب من مجتمع الطلبة .وبالتالي فإن المجتمع الاحصائي هو مجموعة من الوحدات الاحصائية.

ج) **العينة Echantillon:** هي جزء من المجتمع الاحصائي يتم اختيارها بطرق مختلفة بغرض دراسة هذا المجتمع. نلجأ اليها من اجل استخراج النتائج المطلوبة في وقت قصير كما تسمح لنا العينة بتوفير الجهد والتكاليف.

د) **المتغير Variable :** هي المعيار الذي على أساسه يمكن تقسيم المجتمع الاحصائي وتنظيمه وتوزيع وحداته. وتعرف أيضا بأنها بيانات غير رقمية أو بيانات رقمية في شكل مستويات أو في شكل فئات رقمية وتقاس بمعيارين(نوعين)

1. **متغيرات نوعية (كيفية) Variable qualitative:** لا يمكن التعبير عن حالتها بأرقام حيث لا يمكن قياسها أي هي البيانات التي تصف أفراد المجتمع الاحصائي.**مثل:** لون الشعر أو العيون أو البشرة أو تقديرات النجاح للطلاب في احدى المواد...وتنقسم بدورها إلى نوعين هما:

أ . بيانات نوعية ترتيبية(رتبية) :وهي صفة نوعية يمكن ترتيب حالاتها المختلفة ترتيبا معينا .**مثل:** مستوى الطالب حيث نجد الحالات التالية :ممتاز، جيد جدا، متوسط، ضعيف.

ب . بيانات نوعية غير ترتيبية (غير رتبية) : في هذا النوع لا يوجد أي معيار لترتيب حالاتها .مثل : التخصص وهي صفة حالاتها :علوم اقتصادية، علوم تجارية، علوم التسيير، علوم مالية ومحاسبية.

2. **متغيرات كمية (عددية) Variable quantitatives :** هي صفة يعبر عن حالاتها بأرقام أي يمكن قياس حالاتها المختلفة. اي هي البيانات التي يقاس فيها افراد المجتمع الاحصائي بمقاييس كمية(رقمية.)**مثل** :أطوال الطلاب وتقاس بالسنتيمتر، أوزان الطلاب وتقاس بالكيلوجرام، أعمار الطلاب وتقاس بالسنة، نتيجة الامتحان وتقاس بالدرجات، أجور العمال وتقاس بالدينار ...وتنقسم بدورها الى نوعين هما:

أ .متغير كمي متقطع: هي صفة كمية تأخذ حالاتها قيماً ثابتة ومحددة(رقما واحدا محددا)لا تقبل وحدات قياسها التجزئة **مثل**:عدد الأهداف المسجلة في كل مباراة، عدد الأطفال في كل عائلة...

ب . متغير كمي مستمر(متصل): هي صفة كمية يمكن تجزئة وحدات قياسها ونعبر عنها ليس بقيم ثابتة محددة ولكن بمجالات . **مثل :**قامة(طول)مجموعة من الطلبة أو أجور العمال...

تمارين محلولة:

التمرين الأول:

حدد المجتمع الاحصائي، الوحدة الإحصائية، المتغير الاحصائي ونوعه في كل حالة من الحالات التالية:

- توزيع العمال حسب المنصب الوظيفي.
- ترتيب مجموعة من السيارات حسب النوع.
- أطوال مجموعة من الرياضيين.
- مدة حياة المصابيح الكهربائية المنتجة في المصنع.
- الأجور الشهرية لعمال مؤسسة ما.
- بلديات الإقامة لمجوعة من طلبة جامعة غليزان.
- عدد أسهم شركة مساهمة المخصصة للفرد.
- المستوى التعليمي لمجموعة من العمال.

التمرين الثاني:

حدد نوع المتغيرات في العبارات التالية:

مكان الميلاد، درجات الحرارة، عدد الزبائن، الحالة الاجتماعية، عدد أفراد الأسرة، الجنسية، أوزان مجموعة من الأشخاص، عدد الأسهم، الاستهلاك الأسري، المهن، رضا المستهلكين لمنتوج ما، رقم الأعمال.

حلول التمارين:

حل التمرين الأول:

الحالة	**المجتمع الاحصائي**	**الوحدة الاحصائية**	**المتغير الاحصائي**	**نوعه**
توزيع العمال حسب المنصب الوظيفي	العمال	عامل واحد	المنصب الوظيفي	نوعي ترتيبي
ترتيب مجموعة من السيارات حسب النوع	السيارات	سيارة	نوع السيارة	نوعي غير ترتيبي
أطوال مجموعة من الرياضيين	الرياضيين	رياضي	الطول	كمي مستمر
مدة حياة المصابيح الكهربائية المنتجة في المصنع	المصابيح	مصباح	مدة الحياة	كمي مستمر
الأجور الشهرية لعمال مؤسسة ما	العمال	عامل	الأجور	كمي مستمر
بلديات الإقامة لمجوعة من طلبة جامعة غليزان	الطلبة	طالب	بلديات الإقامة	نوعي غير ترتيبي
عدد أسهم شركة مساهمة المخصصة للفرد	المساهمين	مساهم	عدد الأسهم	كمي متقطع
المستوى التعليمي لمجموعة من العمال	العمال	عامل	المستوى التعليمي	نوعي ترتيبي

حل التمرين الثاني:

متغيرات كمية	**متغيرات نوعية**
درجات الحرارة عدد الزبائن عدد أفراد الأسرة أوزان مجموعة من الأشخاص عدد الأسهم الاستهلاك الأسري رقم الأعمال	مكان الميلاد الحالة الاجتماعية الجنسية المهن رضا المستهلكين لمنتوج ما

الفصل الثاني: عرض البيانات الإحصائية

تمهيد:

الخطوة التالية بعد جمع البيانات في مجال الإحصاء الوصفي، هو تبويب البيانات وعرضها بصورة يمكن الاستفادة منها في وصف الظاهرة محل الدراسة، من حيث تمركز البيانات، ودرجة تجانسها .وهناك طريقتين لعرض البيانات هما:

أولا :عرض البيانات جدوليا.

ثانيا :العرض(التمثيل)البياني.

أولا :عرض البيانات جدوليا

يمكن عرض البيانات في صورة جدول تكراري، ويختلف شكل الجدول طبقا لنوع البيانات، وحسب عدد المتغيرات، ففي الجدول الاحصائي الأولي (البسيط)نضع في العمود الأول جميع الحالات الممكنة للصفة المدروسة ونرمز لها بالرمز X_i ، ونضع في العمود الثاني عدد عناصر المجتمع الاحصائي المقابلة لكل حالة أي التكرار المطلق n_i ويكون الجدول الاحصائي كما يلي:

الجدول رقم 01: الجدول التكراري البسيط

الحالات X_i	التكرار المطلق n_i
X_1	n_1
X_2	n_2
X_3	n_3
- - - X_k	- - - n_k
المجموع	$\sum_{i=1}^{k} n_i = N$

هذا الجدول يبين لنا أو يعطينا توزيع المجتمع الاحصائي حسب الصفة المدروسة يمكن اثراء هذا الجدول بإضافة عمودا ثالثا مخصصا لما يسمى **بالتكرارات النسبية** التي نرمز لها ب f_i حيث:

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

كما يمكننا الحصول على نسب مئوية (**تكرار نسبي مئوي**) بضرب الحاصل في 100 كما يلي

$$f_i(\%) = \frac{n_i}{N} \times 100$$

1) **الجداول التكرارية البسيطة ذات المتغيرات النوعية :**

وهي الجداول التي تتضمن تكرارات صفات نوعية معينة للظاهرة المدروسة، وتحتوي هذه الجداول على صفة نوعية واحدة فقط (جداول تكرارية بسيطة) والمثال التالي يبين لنا كيف يمكن تبويب ويتم افراغ البيانات فيها كما هو مبين في الجدول السابق، والمثال التالي يبين لنا كيف يمكن تبويب البيانات الوصفية الخام في شكل جدول تكراري.

مثال01:

أخذت عينة عشوائية من الطلبة تتكون من 20 طالبا، ليتم استقصاءهم عن شعب البكالوريا التي يحملونها، وتم ذلك من خلال ملأ استمارات خاصة، فكانت الاجابات في الاستمارات كما يلي:

تسيير واقتصاد	أدب	علوم	تسيير واقتصاد	رياضيات
أدب	علوم	تسيير واقتصاد	رياضيات	علوم
تسيير واقتصاد	رياضيات	تسيير واقتصاد	علوم	علوم
تسيير واقتصاد	تسيير واقتصاد	رياضيات	تسيير واقتصاد	أدب

المطلوب:

1. حدد المجتمع الاحصائي، الوحدة الإحصائية، المتغير الاحصائي ونوعه؟
2. مثل هذه البيانات في جدول توزيع تكراري؟
3. كون التوزيع التكراري النسبي؟
4. علق على هذه النتائج؟

الحل:

1. المجتمع الاحصائي: الطلبة،

- الوحدة الإحصائية: طالب،
- المتغير الاحصائي: شعب البكالوريا
- نوعي غير ترتيبي.

2. تمثيل البيانات في جدول:

الفرع	التكرار المطلق n_i	التكرار النسبي $f_i = \frac{n_i}{N}$	التكرار النسبي المئوي $f_i(\%) = \frac{n_i}{N} \times 100$
أداب	3	$f_1 = \frac{3}{20}$ =0.15	$f_1(\%) = \frac{3}{20} \times 100$ =15%
علوم	5	0.25	25%
تسيير واقتصاد	8	0.4	40%
رياضيات	4	0.2	20%
المجموع	**20**	**1**	**100%**

التعليق: نلاحظ أن شعبة البكالوريا الشائعة بين الطلبة في العينة هي شعبة التسيير والاقتصاد وهي أكبر نسبة (40%) مما يدل على أن شعبة التسيير والاقتصاد لا تمثل الأغلبية بين طلبة العينة المستسقاة، أما كل من شعبة العلوم والرياضيات والأداب فهي بنسب متقاربة (25%)و(20%) و(15%) على التوالي.

2) **الجداول التكرارية البسيطة ذات المتغيرات الكمية:**

وهي نوعان هما على التوالي:

أ) **الجداول التكرارية البسيطة ذات المتغيرات الكمية المتقطعة (المنفصلة):** وهي الجداول التي تظهر عدد تكرارات كمية واحدة محددة وممثلة في رقم واحد فقط.

مثال 02:

أراد صاحب مكتبة تبيع الكتب الدراسية لطلاب الجامعة ان يحصر عدد الكتب التي يشتريها الطلاب في السداسي الأول من السنة الدراسية .قام صاحب المحل باختيار عينة عشوائية من 12 طالب وطالبة وسأل كل واحد منهم عن عدد الكتب التي اشتراها في السداسي الأول وكانت الإجابات كما يلي:

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

لكي تكون هذه البيانات أكثر فائدة يجب أن يتم تنظيمها . ونلاحظ أن المتغير الذي ورد في العينة هو عدد الكتب التي يشتريها الطالب في السداسي الاول وهو **متغير كمي متقطع**.

الحل:

جدول توزيع تكراري للكتب المشترات من قبل الطلبة

عدد الكتب Xi	التكرار المطلق ni	التكرار النسبي
0	1	0.833
1	1	0.833
2	2	0.1667
3	4	0.3333
4	3	0.25
5	1	0.833
المجموع	**12**	**1**

ب) **الجداول التكرارية البسيطة ذات المتغيرات الكمية المستمرة (المتصلة):**

وهي الجداول التي تكون فيها الظاهرة محصورة في مجال، بحيث يمكن أن تأخذ أية قيمة ضمنه. ويتم استخدام هذه الطريقة في عرض البيانات إذا كان عددها كبيرا، وذلك لتقليصها، إذ أن هدف التبويب هو عرض البيانات بأقل حيز ممكن وبأقصى وضوح، فيتم حينئذ تحديد فئات طولها أكبر من الصفر، ويتم اهمال القيم التي تقع داخل مجال الفئات.

المدى: هو الفرق بين أكبر قيمة ضمن مجموعة القيم وأصغر قيمة ضمنها، ونرمز له بالرمز W

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

X_{max}: أعظم (أكبر) قيمة ضمن مجموعة القيم.

X_{min}: أدنى (أصغر) قيمة ضمن مجموعة القيم.

- **تحديد طول الفئة:**

تحديد طول الفئة يساعد على تحديد عدد الفئات وبالتالي حجم الجدول، إذ كلما كان طول الفئة كبيرا كلما كان حجم الفئة صغيرا والعكس صحيح، ولتحديد طول الفئة يتم استخدام ستيرجس (H.A. Sturges) والتي تعطى كما يلي:

$$L=\frac{W}{1+3.322LogN}$$

حيث:

L: طول الفئة،

N: عدد القيم،

W: المدى.

➢ **تحديد عدد الفئات:** يحدد عدد الفئات باستخدام القاعدة التالية:

$$N_C = \frac{W}{L}$$

حيث: N_Cتمثل عدد الفئات.

$$W = L(1 + 3.322LogN)$$

$$N_C = 1 + 3.322LogN$$

كذلك يمكن تحديد عدد الفئات من خلال معادلة يول yule كالتالي:

$$N_C = 2.5\sqrt[4]{N}$$

مثال03:

إذا كانت لدينا نقاط 30 طالب في أحد الاختبارات في مادة الإحصاء كما يلي:

2، 8، 7، 6، 4، 18، 12، 4، 6، 8، 15، 6، 9، 6، 14، 6، 8، 4، 11، 9، 8، 11 ، 9، 4، 8، 9، 11، 12، 4، 5.

لخص هذه البيانات في صورة جدول يوضح معالمها الأساسية؟

الحل:

تكوين جدول التوزيع التكراري: نقاط الطلبة في الاختبار متغير كمي مستمر، ولكي يتم تبويب البيانات في شكل جدول تكراري، يتم اتباع الأتي:

حساب المدىW:

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

$W = 19 - 2 = 17$

تحديد طول الفئة:

$$L = \frac{W}{1 + 3.322LogN}$$

$$L = \frac{17}{1 + 3.322Log30}$$

$$L = \frac{17}{1 + 3.322(1.477)}$$

$$L = 2.88 \cong 3$$

تحديد عدد الفئات:

$$N_C = \frac{W}{L} \Rightarrow N_C = \frac{17}{3} = 5.66 \cong 6$$

جدول التوزيع التكراري:

الفئات	التكرارات المطلقة n_i	التكرار النسبي المئوي $f_i(\%)$
[2 - 5[	6	20
[5 - 8[	7	23.33
[8 - 11[	9	30
[11 -14[	5	16.67
[14 -17[	2	6.67
[17 -20[	1	3.33
المجموع	30	100

3) التكرارات المتجمعة:

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 الناجحين (أي الطلبة المتحصلين على درجة تساوي أو تزيد 10) فإن هذه المعلومات غير واضحة في جدول التوزيع التكراري فنكون لهذا الغرض ما يسمى بالجدول التكراري المتجمع 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.

بالرجوع للمثال رقم 03

الفئات	التكرارات المطلقة n_i	التكرار المتجمع الصاعد N^+	التكرار المتجمع النازل N^-
[2 - 5[	6	6	30
[5 - 8[	7	13	24
[8 - 11[	9	22	17
[11 -14[	5	27	8
[14 -17[	2	29	3
[17 -20[	1	30	1
المجموع	30	/	/

ثانيا :العرض(التمثيل)البياني

العرض البياني للبيانات، هو أحد طرق التي يمكن استخدامها في وصف البيانات، من حيث شكل التوزيع ومدى تمركز البيانات، وفي كثير من النواحي التطبيقية يكون العرض البياني أسهل وأسرع في وصف الظاهرة محل الدراسة، وتختلف طرق عرض البيانات بيانيا حسب نوع البيانات المبوبة في شكل جدول تكراري، وفيما يلي عرض للأشكال البيانية المختلفة.

1) **العرض البياني للمتغيرات النوعية:** يمكن عرض البيانات الخاصة بمتغير وصفي في شكل دائرة بيانية أو أعمدة بيانية، يمكن من خلاله وصف ومقارنة مجموعات أو مستويات هذا المتغير.

أ) **الدائرة البيانية:** يمكن ان نرسم دائرة ونقسمها الى قطاعات دائرية تتناسب مساحة كل قطاع مع تكرار الفئة التي تمثلها. فالفئة الاكثر تكرارا تقابل القطاع الاكبر مساحة والفئة الاقل تكرارا تقابل القطاع الاصغر مساحة. طريقة الدائرة هي عبارة عن تقسيم الكل الى أجزاء وكل جزء يُمَثَّلْ بقطاع دائري بحيث ان زاوية راس كل قطاع دائري تعطى حسب القاعدة التالية:

$$\text{قياس الزاوية} = \frac{\text{قيمة الجزء الواحد}}{\text{مجموع قيم الأجزاء}} \times 360^0$$

مثال04:

الجدول التكراري التالي يبين توزيع عينية حجمها 500 أسرة حسب المنطقة التي تنمي إليها.

المنطقة	غليزان	واد رهيو	مازونة	يلل	المجموع
عدد الأسر (التكرار)	170	150	130	50	500

المطلوب: مثل البيانات أعلاه في شكل دائرة بيانية.

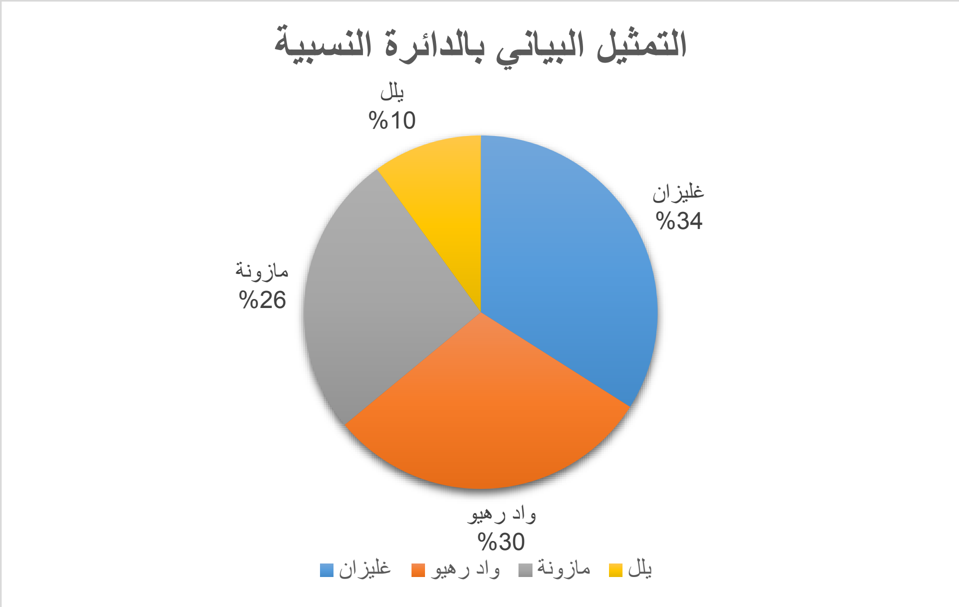
الحل:

جدول التوزيع التكراري

المنطقة	عدد الاسر	مقدار الزاوية
غليزان	170	122.4^0
واد رهيو	150	108^0
مازونة	130	93.6^0
يلل	50	36^0
المجموع	500	360^0

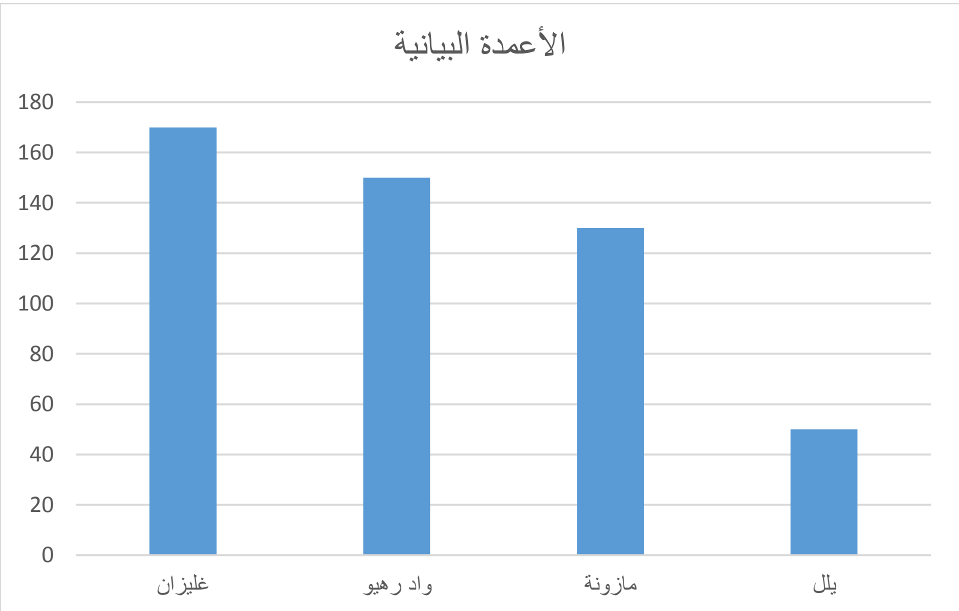
رسم الدائرة :يتم رسم دائرة وتقسيمها إلى أربع أجزاء لكل منطقة جزء يتناسب مع مقدار الزاوية المخصصة له، كما هو مبين في الشكل التالي:

الدائرة البيانية لعينة حجمها 500 أسرة موزعة حسب المنطقة



ب) **الأعمدة البيانية(التكرارية):**
تستخدم لتقديم البيانات الاحصائية التي لا مدى لفئاتها، والبيانات ذات الصفات النوعية، ولتقديم البيانات بهذه الطريقة، يتم اعداد معلم متعامد، بحيث توضع الفئات(سواء كانت صفات كمية، او ذات صفات نوعية)على المحور الافقي، أما على المحور العمودي فيتم وضع التكرارات.

بالرجوع إلى نفس المثال السابق، الأعمدة البيانية لتوزيع الأسر على المناطق:



2) **العرض البياني للمتغيرات الكمية:**

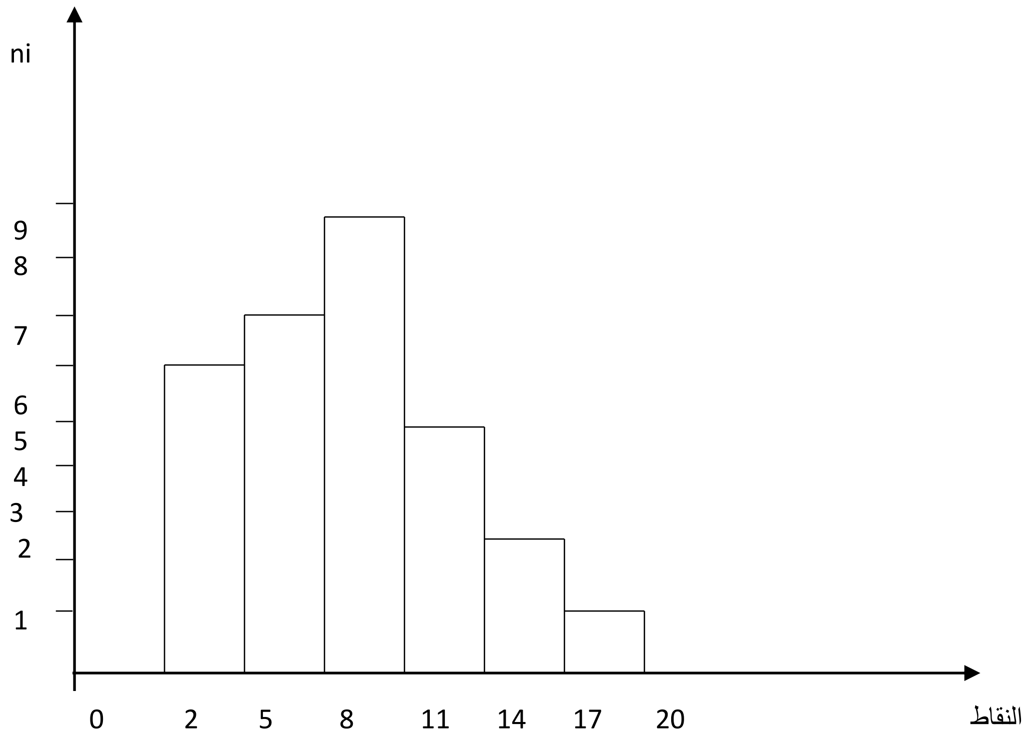
أ) **المدرج التكراري Histogram:** المدرج التكراري هو التمثيل البياني للجدول التكراري البسيط الخاص بالبيانات الكمية المتصلة، وهو عبارة عن أعمدة بيانية متلاصقة، حيث تمثل التكرارات على المحور العمودي، بينما تمثل قيم المتغير (حدود الفئات) على المحور الأفقي، ويتم تمثيل كل فئة بعمود، ارتفاعه هو تكرار الفئة، طول قاعدته هو طول الفئة.

بالرجوع إلى المثال رقم 03:

الفئات	التكرارات المطلقة n_i	مركز الفئة
[2 - 5[	6	3.5
[5 - 8[	7	6.5
[8 - 11[	9	9.5
[11 -14[	5	12.5
[14 -17[	2	15.5
[17 -20[	1	18.5
المجموع	30	/

المطلوب: أنشئ المدرج التكراري؟

الحل:



ب) **المضلع التكراري:** هو تمثيل بياني أيضا للجدول التكراري البسيط، حيث تمثل التكرارات على المحور العمودي، ومراكز الفئات على المحور الأفقي، ثم التوصيل بين الإحداثيات بخطوط منكسرة، وبعد ذلك بتوصيل طرفي المضلع بالمحور الأفقي. ومركز الفئة هي القيمة التي تقع في منتصف الفئة، وتحسب بتطبيق المعادلة التالية:

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

ج) **المنحنى التكراري**: إتباع نفس الخطوات السابقة في رسم المضلع يمكن رسم المنحنى التكراري، ولكن يتم تمهيد الخطوط المنكسرة في شكل منحنى بحيث يمر بأكثر عدد من النقاط، وفي المثال السابق يمكن رسم المنحنى التكراري.

المنحنى التكراري هو خط منحنى ممهد للمضلع التكراري، يعطي لنا فكرة عن شكل التوزيع هل هو قريب إلى التوزيع الطبيعي (متناظر أو غير متناظر) .

تمارين محلولة:

التمرين الأول:

الجدول التكراري التالي يمثل توزيع مجموعة من الطلبة حسب التخصصات المفتوحة في احدى الجامعات.

التخصصات	العلوم الاقتصادية	الحقوق	العلوم الاجتماعية	الأداب واللغات	العلوم والتكنولوجيا
عدد الطلبة	350	220	430	180	170

المطلوب:

1. حدد المجتمع الاحصائي، الوحدة الإحصائية، المتغير الاحصائي ونوعه؟
2. أحسب قيمة التكرار النسبي؟
3. مثل هذه البيانات باستخدام الأعمدة البيانية وباستخدام شكل الدائرة؟

التمرين الثاني:

تمثل البيانات التالية المستوى التعليمي لعينة من 50 شخص.

متوسط	ابتدائي	ثانوي	ابتدائي	جامعي	متوسط	ثانوي
ابتدائي	جامعي	ثانوي	متوسط	جامعي	ثانوي	دراسات عليا
ابتدائي	متوسط	ثانوي	متوسط	ابتدائي	دراسات عليا	ثانوي
متوسط	دراسات عليا	جامعي	جامعي	متوسط	جامعي	جامعي
متوسط	دراسات عليا	جامعي	متوسط	ثانوي	ثانوي	جامعي
ثانوي	ثانوي	دراسات عليا	جامعي	متوسط	جامعي	ثانوي
جامعي	ابتدائي	ثانوي	جامعي	ثانوي	ثانوي	جامعي
ثانوي						

المطلوب:

1. حدد المجتمع الاحصائي، الوحدة الإحصائية، المتغير الاحصائي ونوعه؟
2. أعرض البيانات في شكل جدول تكراري.
3. كون التوزيع التكراري النسبي، والمئوي، ثم علق على النتائج.

التمرين الثالث:

الجدول التالي يمثل التوزيع التكراري لمجموعة من المؤسسات الناشئة (الصغيرة) حسب عدد المشغلين

عدد العمال	1	2	3	4	5	6	7	المجموع
عدد المؤسسات	14	13	12	15	16	17	13	100

المطلوب:

1. أنشئ جدول التوزيع التكراري؟
2. أحسب قيمة كل من التكرارات النسبية والمئوية، والتكرارات المتجمعة الصاعدة والنازلة؟
3. ماهو عدد المؤسسات التي تشغل على الأقل 5 عمال؟
4. ماهو عدد المؤسسات التي تشغل على الأكثر 4عمال؟

التمرين الرابع:

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

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

المطلوب:

1. حدد المجتمع الإحصائي والمتغير الإحصائي (الصفة المدروسة) ونوعه.
2. لخص هذه البيانات في جدول إحصائي.

التمرين الخامس:

لتكن لدينا سلسلة إحصائية تخص أوزان 60طالب

80	55	60	62	58	75	95	90	89	84	75	76	58	69	77
76	78	67	62	97	90	87	89	71	72	75	84	75	80	76
64	79	81	83	87	92	60	59	88	70	70	76	81	82	76
66	65	63	64	94	86	68	70	72	75	78	80	74	75	80

المطلوب:

1. حدد عدد الفئات باستخدام معادلة يول؟
2. حدد طول الفئة، ومركز كل فئة؟
3. أنشئ الجدول التوزيع التكراري؟ وأحسب مختلف التكرارات (النسبية، المئوية، المتجمعة الصاعدة والنازلة)؟

التمرين السادس:

المعلومات الواردة أدناه تمثل الأجر الشهري لـــــ 45 عامل في احدى المؤسسات خلال شهر جوان 2022.

(الوحدة: بالألف دج)

23	35	46	29	47	56	39	43	28	25	45	52
23	28	43	57	42	51	32	24	49	34	31	42
34	29	54	57	49	19	45	53	29	38	44	24
			37	19	21	41	43	35	47	37	41

المطلوب:

1. حدد المجتمع الاحصائي، المتغيرة الاحصائية، بين طبيعتها.
2. أنشئ جدول التوزيع التكراري بعد تحديد طول الفئة حسب طريقة sturges، عدد الفئات.
3. أحسب التكرارات النسبية والمئوية، ومراكز الفئات والتكرارات المتجمعة الصاعدة والنازلة.

حلول التمارين:

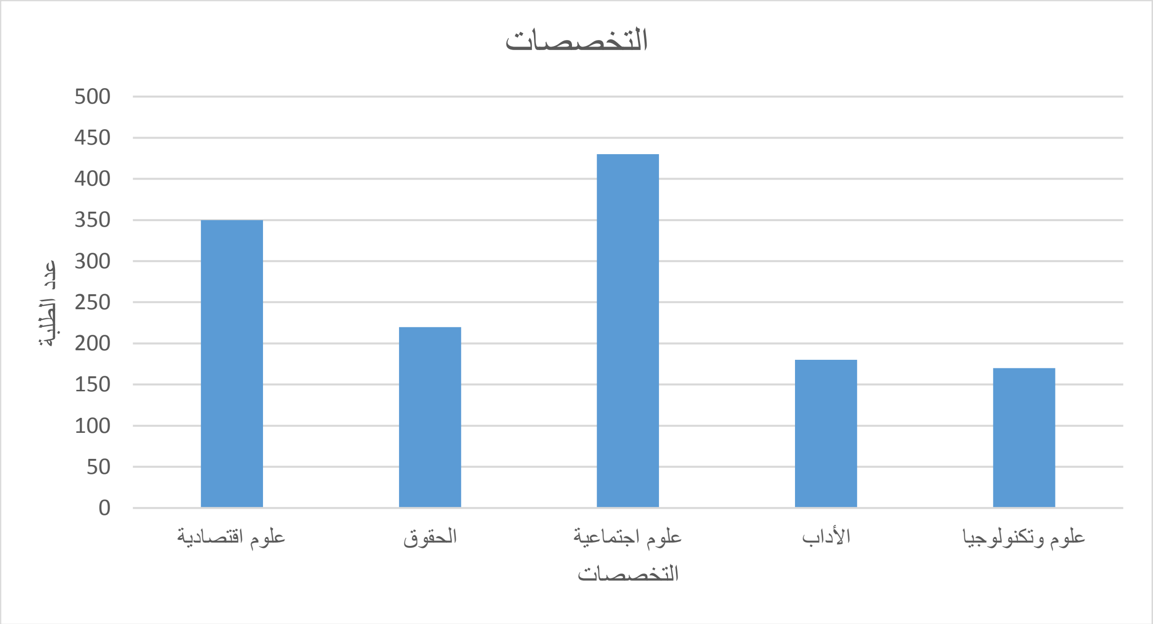
حل التمرين الأول:

- المجتمع الاحصائي: الطلبة،
- الوحدة الإحصائية: طالب،
- المتغير الاحصائي: التخصص
- نوعي غير ترتيبي.

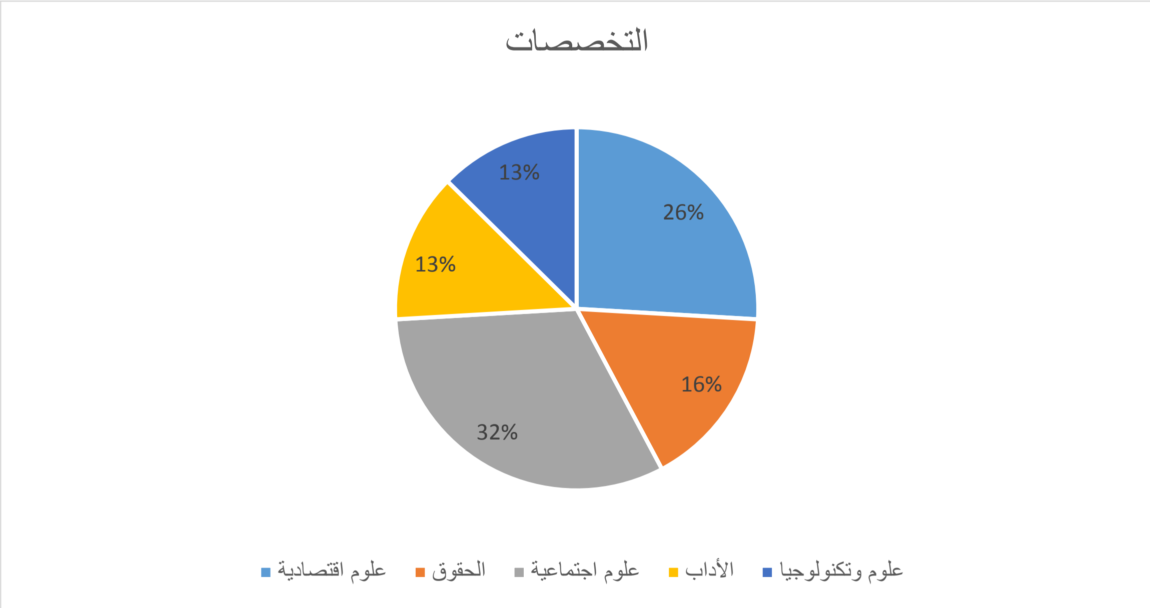
تمثيل البيانات في جدول:

التخصصات	التكرار المطلق n_i	التكرار النسبي $f_i = \frac{n_i}{N}$	التحويل إلى الزوايا $f_i \times 360$
علوم اقتصادية	350	0.26	$0.15 \times 360 = 93.6^0$
الحقوق	220	0.16	57.6^0
علوم اجتماعية	430	0.32	115.2^0
الأداب	180	0.13	46.8^0
علوم وتكنولوجيا	170	0.12	43.2^0
المجموع	**1350**	**1**	360^0

تمثيل هذه البيانات باستخدام الأعمدة البيانية:



تمثيل هذه البيانات باستخدام شكل الدائرة:



حل التمرين الثاني:

- المجتمع الاحصائي: أشخاص،
- الوحدة الإحصائية: شخص،
- المتغير الاحصائي: المستوى التعليمي،
- نوعي ترتيبي.

جدول التوزيع التكراري:

المستوى الدراسي	التكرار المطلق n_i	التكرار النسبي $f_i = \frac{n_i}{N}$	التكرار النسبي المئوي $f_i(\%) = \frac{n_i}{N} \times 100$
ابتدائي	6	0.12	$f_1(\%) = \frac{6}{20} \times 100 = 12\%$
متوسط	10	0.2	20%
ثانوي	15	0.3	30%
جامعي	14	0.28	28%
دراسات عليا	5	0.1	10%
المجموع	**20**	**1**	**100%**

حل التمرين الثالث:

جدول التوزيع التكراري:

عدد العمال	التكرارات n_i	التكرار النسبي f_i	التكرار النسبي المئوي $f_i(\%)$	التكرار المتجمع الصاعد N^+	التكرار المتجمع النازل N^-
1	14	0.14	14%	14	100
2	13	0.13	13%	27	86
3	12	0.12	12%	39	73
4	15	0.15	15%	54	61
5	16	0.16	16%	70	46
6	17	0.17	17%	87	30
7	13	0.13	13%	100	13
المجموع	**100**	**1**	**100%**	/	/

- ✓ **عدد المؤسسات التي تشغل على الأقل 5 عمال:**

من التكرار المتجمع النازل نجد أن عدد المؤسسات التي تشعل على الأقل 5 عمال تقدر ب 46 مؤسسة.

- ✓ **عدد المؤسسات التي تشغل على الأكثر 4 عمال:**

من التكرار المتجمع الصاعد نجد أن عدد المؤسسات التي تشغل على الأكثر 4 عمال تقدر ب 54 مؤسسة.

حل التمرين الرابع:

المجتمع الاحصائي: عمال المؤسسة،

المتغير الاحصائي: عدد الغيابات،

نوعه: كمي متقطع.

جدول التوزيع التكراري:

Xi عدد الغيابات	ni عدد العمال
0	4
1	5
2	7
3	6
4	5
5	8
6	2
7	1
9	2
المجموع	**40**

حل التمرين الخامس:

تحديد عدد الفئات باستخدام معادلة يول:

$$N_C = 2.5\sqrt[4]{N} = 2.5\sqrt[4]{60} = 6.95 \cong 7$$

تحديد طول الفئة:

$$N_C = \frac{W}{L} \Rightarrow N_C = \frac{X_{max} - X_{min}}{L} = \frac{97-55}{7} \cong 6$$

جدول التوزيع التكراري:

الفئات	n_i	C_i	f_i	$f_i(\%)$	$N\uparrow$	$N\downarrow$
]61-55]	6	58	0.10	10%	6	60
]67-61]	7	64	0.1167	11.67%	13	54
]73-67]	9	70	0.15	15%	22	47
]79-73]	15	76	0.25	25%	37	38
]85-79]	11	82	0.1833	18.33%	48	23
]91-85]	8	88	0.1333	13.33%	56	12
]97-91]	4	94	0.0667	6.67%	60	4
المجموع	**60**	/	**1**	**100%**	/	/

حل التمرين السادس:

المجتمع الاحصائي : العمال، المتغير الاحصائي: الأجر، نوعه: كمي مستمر.

تحديد طول الفئة حسب طريقة sturges :

$$L = \frac{W}{1 + 3.322LogN} = \frac{57-19}{1 + 3.322Log45} = \frac{38}{6.48} = 5.86 \cong 6$$

تحديد عدد الفئات:

$$N_C = \frac{W}{L} = \frac{38}{6} = 6.33 \cong 7$$

جدول التوزيع التكراري:

الفئات	**n_i**	C_i	f_i	$N\uparrow$	$N\downarrow$
]25-19]	7	22	0.16	7	45
]31-25]	6	28	0.13	13	38
]37-31]	6	34	0.13	19	32
]43-37]	9	40	0.2	28	26
]49-43]	8	46	0.18	36	17
]55-49]	6	52	0.13	42	9
]61-55]	3	58	0.07	45	3
المجموع	**45**	/	**1**	/	/

الفصل الثالث: مقاييس النزعة المركزية

في كثير من النواحي التطبيقية يكون الباحث في حاجة إلى حساب بعض المؤشرات التي يمكن الاعتماد عليها في وصف الظاهرة من حيث القيمة التي تتوسط القيم أو تترع إليها القيم، ومن حيث التعرف على مدى تجانس القيم التي يأخذها المتغير، وأيضا ما إذا كان هناك قيم شاذة أم لا. والاعتماد على العرض البياني وحدة لا يكفي، ولذا يتناول هذا الفصل، والذي يليه عرض بعض المقاييس الإحصائية التي يمكن من خلالها التعرف على خصائص الظاهرة محل البحث، وكذلك إمكانية مقارنة ظاهرتين أو أكثر، ومن أهم هذه المقاييس، مقاييس الترعة المركزية والتشتت.

مقاييس الترعة المركزية

تسمى مقاييس الترعة المركزية بمقاييس الموضع أو المتوسطات، وهي القيم التي تتركز القيم حولها، ومن هذه المقاييس، الوسط الحسابي، والمنوال، والوسيط، والوسط الهندسي والوسط التوافقي والرباعيات والمئينات، وفيما يلي عرض لأهم هذه المقاييس:

أولا :الوسط الحسابي:

من أهم مقاييس الترعة المركزية، وأكثرها استخداما في النواحي التطبيقية، ويمكن حسابه للبيانات المبوبة وغير المبوبة، كما يلي:

1) الوسط الحسابي للبيانات غير المبوبة

يعرف الوسط الحسابي بشكل عام على أنه مجموع القيم مقسوما على عددها، فإذا كان لدينا n من القيم، ويرمز لها بالرمز: $x_1, x_2, \ldots, x_n$ يحسب بالمعادلة التالية:

$$\text{الوسط الحسابي} = \frac{\text{مجموع القيم}}{\text{عدد القيم}}$$

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

مثال01:

أوجد الوسط الحسابي للقيم التالية: 12،15،11،17،14،13،16،14

الحل:

$$\bar{x} = \frac{14+16+13+14+17+11+15+12}{8} = 14$$

2) حالة البيانات المبوبة:

أ) حالة المتغير المتقطع:

إذا كانت $x_1, x_2, \ldots, x_k$ قيم ميزة إحصائية، وكانت $n_1, n_2, \ldots, n_k$ تكراراتها على الترتيب.

فإن الوسط الحسابي لهذه السلسلة الإحصائية يحسب بالمعادلة التالية:

$$\bar{x} = \frac{n_1x_1 + n_2x_2 + \cdots + n_kx_k}{n_1 + n_2 + \cdots + n_k} = \frac{\sum_{i=1}^{k} n_i x_i}{N}$$

حيث:

K: عدد القيم المختلفة.

N: حجم المجتمع.

n_i: التكرار المطلق.

x_i: القيم.

مثال 02:

البيانات الأتية تمثل التوزيع التكراري لأعداد العمال في 100مؤسسة صغيرة.

عدد العمال x_i	عدد المؤسسات n_i	$n_i * x_i$
1	15	15
2	10	20
3	5	15
4	20	80
5	30	150
6	20	120
المجموع	**100**	**400**

$$\bar{x} = \frac{\sum_{i=1}^{k} n_i x_i}{N} = \frac{400}{100} = 4$$

متوسط عدد العمال في المؤسسات هو 04

ب) حالة المتغير المستمر:

من المعلوم أن القيم الأصلية لا يمكن معرفتها من جدول التوزيع التكراري، حيث أن هذه القيم موضوعة في فئات، لذا يتم التعبير عن كل قيمة من القيم التي تقع داخل حدود الفئة بمركز الفئة، ومن ثم يؤخذ في الاعتبار أن مركز الفئة هو القيمة التقديرية لكل مفردة تقع في هذه الفئة.

فإذا كانت K هي عدد الفئات وكانت c_1 , c_2 , ..., c_k هي مراكز هذه الفئات، وكانت n_1, n_2, ..., n_k تكراراتها على الترتيب.

فإن المتوسط الحسابي يحسب بالمعادلة التالية:

$$\bar{x} = \frac{\sum_{i=1}^{k} n_i c_i}{\sum_{i=1}^{k} n_i}$$

مثال03:

الجدول التاي يعطينا توزيع 30 طالب حسب نقاط مادة الإحصاء:

الفئات	التكرارات المطلقة n_i	مركز الفئة c_i	n_i*c_i
]5 - 2]	7	3.5	24.5
]8 - 5]	7	6.5	45.5
]11 - 8]	9	9.5	85.5
]14- 11]	4	12.5	50
]17- 14]	2	15.5	31
]20- 17]	1	18.5	18.5
المجموع	**30**	/	**255**

$$\bar{x} = \frac{\sum_{i=1}^{k} n_i c_i}{\sum_{i=1}^{k} n_i} = \frac{255}{30} = 8.5$$

معدل الطلبة هو 8.5 في مادة الإحصاء.

ثانيا: الوسيط

الوسيط هو ثاني مقاييس النزعة المركزية هو أقل دقة من الوسط الحسابي، ولكنه يعتبر جيد في تمثيل الفئات المفتوحة والتي لا يستطيع الوسط الحسابي أن يعبر عنها بدقة .الوسيط هو الدرجة التي تقسم توزيع الدرجات إلى قسمين أو نصفين متساويين. أي هو القيمة التي تقع في منتصف المجموعة بعد ترتيب القيم تصاعديا أو تنازليا، أي هو القيمة التي يكون نصف عدد القيم أصغر منها أو يساويها والنصف الآخر أكبر منها أو يساويها، من هذا التعريف للوسيط نجد أنه يعالج العيوب الثلاثة التي يعاني منها الوسط الحسابي، فالوسيط لا يتأثر بالقيم الشاذة أو المتطرفة .كما أنه يمكن حسابه في حالة الفئات المفتوحة، ويمكن إيجاده بيانيا.

1) **البيانات غير المبوبة :يتم حساب الوسيط لهذه البيانات باتباع الخطوات التالية:**

- نقوم بترتيب البيانات تصاعديا او تنازليا .
- نقوم بحساب ترتيب الوسيط C حسب عدد القيم إذا كان زوجيا او فرديا.

أ) إذا كان عدد القيم N فرديا يكون :

$$c = \frac{n+1}{2}$$

حيث أن: c: رتبة الوسيط. n: عدد البيانات.

ب) إذاكان عدد القيم N زوجيا يكون :

$$M_e = \frac{M_{e1} + M_{e2}}{2}$$

$$c = \frac{n}{2}, \qquad c = \frac{n}{2} + 1$$

حيث أن:

M_{e1}: الوسيط الأول.

M_{e2}: الوسيط الثاني.

c: رتبة الوسيط.

مثال4:

أحسب الوسيط للقيم التالية: 9، 7، 3، 5، 3، 6، 9.

الحل:

-نرتب الأرقام تصاعديا (مهما تكررت الأرقام).

-حساب الرتبة:

$$c = \frac{n+1}{2} = \frac{7+1}{2} = 4$$

رقم الرتبة	1	2	3	4	5	6	7
القيم	3	3	5	6	7	9	9

إذن قيمة الوسيط هو Me=6

مثال5:

أحسب الوسيط للقيم التالية: 9،7،2،4،7،6.

الحل:

-نرتب الأرقام تصاعديا (مهما تكررت الأرقام).

-حساب الرتبة:

$$c = \frac{n}{2} = \frac{6}{2} = 3$$

رقم الرتبة	1	2	3	4	5	6
القيم	2	4	6	7	7	7

إذن قيمة الوسيط هو $M_{e1}=6$, $M_{e2}=7$

$$M_e = \frac{M_{e1} + M_{e2}}{2} = \frac{6+7}{2} = 6.5$$

2) **البيانات المبوبة:**

البيانات المبوبة قد تكون غير مستمرة اي مدى فئاتها معدوم، وقد تكون مستمرة اي مدى فئاتها أكبر من الصفر، ويتم ايجاد الوسيط حسب كل حالة كما يلي:

أ) **حالة البيانات المنفصلة(المتقطعة)(طول الفئات معدوم) :** في هذه الحالة يتم ايجاد الوسيط كما يلي:

-ترتيب الوسيط باستخدام احدى المعادلتين التاليتين :

- **اذا كان مجموع التكرارات فرديا:**

$$c = \frac{\sum n_i + 1}{2}$$

- **اذا كان مجموع التكرارات زوجيا:**

$$c = \frac{\sum n_i}{2}$$

-**نحسب التكرار المتجمع الصاعد** ونبحث عن مكان ترتيب الوسيط بين التكرارات المتجمعة، فتجده بين تكرارين من التكرارات المتجمعة، وتكون قيمة الوسيط هي القيمة المقابلة للتكرار المتجمع اللاحق لترتيب الوسيط. اذا ما كانت C تساوي احدى قيم التكرارات المتجمعة فانM_e تساوي الفئة المقابلة لها، سواء كان مجموع التكرارات زوجيا او فرديا:

مثال 6:

نفس المثال(02)البيانات الآتية تمثل التوزيع التكراري لأعداد العمال في 100مؤسسة صغيرة

x_i	n_i	N^+
1	15	15
2	10	25
3	5	30
Me=4	20	50
5	30	80
6	20	100
المجموع	**100**	/

$$c = \frac{N}{2} = \frac{100}{2} = 50$$

ب) **حالة البيانات المستمرة (المتصلة) (طول الفئات اكبر من الصفر):** يتم ايجاد الوسيط بعدة طرق تعطي نتائج متقاربة في الغالب، وهي:

- ✓ **الطريقة الأولى (وهي الاكثر استخداما):** يتم استخدام المنهجية التالية لايجاد الوسيط :

-نحسب التكرار المتجمع الصاعد او التكرار المتجمع النازل.

-نجد ترتيب (رتبة)الوسيط باستخدام المعادلة (سواء كان مجموع التكرارات **فرديا او زوجيا** تستخدم نفس المعادلة):

$$c = \frac{\sum_{i=1}^{k} n_i}{2}$$

-نبحث عن مكان ترتيب الوسيط بين التكرارات المتجمعة، فنجده بين تكرارين من التكرارات المتجمعة احدهما سابق له والاخر لاحق له.

-نبحث عن الفئة الوسيطية في حدود الفئات التي تحدد التكرار المتجمع، بحيث يكون الحد الادنى للفئة الوسيطية هو الحد المقابل للتكرار المتجمع السابق لترتيب الوسيط ، وحدها الاعلى هو الحد المقابل للمتجمع اللاحق لترتيب الوسيط.

$$M_e = d + \frac{C - N_{i-1}^{+}}{N_{i+1}^{+} - N_{i-1}^{+}} L$$

حيث أن:

M_e: قيمة الوسيط.

d: الحد الأدنى للفئة الوسيطية.

C: رتبة الوسيط.

L: طول الفئة الوسيطية.

N_{i-1}^{+}: التكرار المتجمع السابق لرتبة الوسيط.

N_{i+1}^{+}: التكرار المتجمع اللاحق لرتبة الوسيط.

مثال 07:

بالرجوع إلى نفس المثال رقم 04

الفئات	التكرارات المطلقة n_i	N^+
[2 - 5[	7	7
[5 - 8[	7	$14=N_{i-1}^{+}$
[8 - 11[	9	$23=N_{i+1}^{+}$
[11 -14[	4	27
[14 -17[	2	29
[17 -20[	1	30
المجموع	**30**	/

حساب الرتبة:

$$c = \frac{\sum_{i=1}^{k} n_i}{2} = \frac{30}{2} = 15$$

$$M_e = d + \frac{C - N_{i-1}^{+}}{N_{i+1}^{+} - N_{i-1}^{+}} L = 8 + \frac{15 - 14}{23 - 14} * 3 = \frac{25}{3} = 8.33$$

الطريقة الثانية :نفس المنهجية المستخدمة في الطريقة الاولى لإيجاد الوسيط:

$$M_e = d + \frac{\frac{\sum n_i}{2} - N_{i-1}^{+}}{n_i} L$$

حيث أن:

M_e: قيمة الوسيط.

d: الحد الأدنى للفئة الوسيطية.

$C = \frac{\sum n_i}{2}$: رتبة الوسيط.

L: طول الفئة الوسيطية.

N_{i-1}^{+}: التكرار المتجمع السابق لرتبة الوسيط.

n_i: التكرار المطلق للفئة الوسيطية.

مثال 08:

بالرجوع إلى نفس المثال السابق

$$M_e = d + \frac{\frac{\sum n_i}{2} - N_{i-1}^{+}}{n_i} L = 8 + \frac{\frac{30}{2} - 14}{9} * 3 = 8.33$$

الطريقة الثالثة: الطريقة البيانية :

يمكن ايجاد الوسيط بيانيا، وذلك برسم، المنحنى التكراري المتجمع الصاعد والنازل وتحديد نقطة تقاطع كل من المنحنى التكراري المتجمع الصاعد والنازل.

بالرجوع إلى المثال السابق: نجد أن قيمة الوسيط يمكن تحديدها بيانيا من خلال المنحنى التالي

المقاييس الشبيهة بالوسيط

بما أنه يمكن تقسيم بيانات أي ظاهرة إلى عدة أقسام متساوية وليس إلى قسمين فقط كما يفعل الوسيط، فإنه يمكن التعامل مع هذه القيم بنفس طريقة التعامل مع الوسيط، حيث نميز:

الربيعات Q: هي القيم التي تقسم مجموع البيانات إلى أربعة أقسام متساوية، كل قسم يأخذ 25%.

العشيرات D: هي القيم التي تقسم مجموع البيانات إلى عشرة أقسام متساوية، كل قسم يأخذ 10%.

المئينات P: هي القيم التي تقسم مجموع البيانات إلى مئة قسم متساوي، كل قسم يأخذ 1%.

المقاييس الشبيهة بالوسيط في البيانات غير المبوبة

-يجب أولا أن نرتب البيانات ترتيبا تصاعديا وليس تنازليا؛
-نحسب التكرار التجميعي الصاعد؛
-نبحث عن المقياس الشبيه بالوسيط المراد حسابه لهذه البيانات وذلك بحساب رتبته، والرتب هي كمايلي:

الربيعات Q: $m = 1,2,3,4.$ $CQ_m = N\frac{m}{4}$

العشيرات D: $m = 1,2, \dots\dots, 10.$ $CD_m = N\frac{m}{10}$

المئينات P: $m = 1,2, \dots, 100.$ $CP_m = N\frac{m}{100}$

حالة البيانات المبوبة :في هذه الحالة يتم استبدال N بمجموع التكرارات

في البيانات المبوبة نعتمد فقط على التكرار التجميعي الصاعد
-نحسب التكرار التجميعي الصاعد؛
-نحدد ترتيب المقياس الشبيه بالوسيط المراد حسابه كمايلي:

الربيعات Q: $m = 1,2,3,4.$ $CQ_m = \frac{m}{4}\sum_{i=1}^{k} n_i$

العشيرات D: $m = 1,2, \dots\dots, 10.$ $CD_m = \frac{m}{10}\sum_{i=1}^{k} n_i$

المئينات P: $m = 1,2, \dots, 100$ $CP_m = \frac{m}{100}\sum_{i=1}^{k} n_i$

- نحدد فئة المقياس الشبيه بالوسيط المراد حسابه، وهي التي تقابل التكرار المتجمع الصاعد الذي يساوي ترتيب المقياس الشبيه بالوسيط المراد حسابه أو أكبر منه مباشرة؛
- نحدد ونحسب المقياس الشبيه بالوسيط المراد حسابه بتطبيق العلاقة السابقة الخاصة بالوسيط كمايلي**:**

$$B = d + \frac{C - N_{i-1}^{+}}{N_{i+1}^{+} - N_{i-1}^{+}} L$$

حيث أن: B: قيمة الوسيط، d: الحد الأدنى للفئة الوسيطية، C: الرتبة المراد ايجادها، L: طول الفئة.

N_{i-1}^{+}: التكرار المتجمع السابق لرتبة الوسيط، N_{i+1}^{+}: التكرار المتجمع اللاحق لرتبة الوسيط.

ثالثا: المنوال

يعرف المنوال لمجموعة من البيانات بانه الفئة الاكثر تكراراً بين مجموعة القيم، ويتم حسابه كمايلي:

1) **البيانات غير المبوبة:** يمثل المنوال في هذه الحالة القيمة الاكثر تكرارا

مثال09:
إليك السلسلة التالية 13،15،10، 13، 12، 12، 13.
إذن قيمة المنوال هي : Mo=13

ملاحظة : يمكن ان نجد أكثر من منوال واحد في نفس السلسلة للقيم، كما يمكن أن لا نجد منوالا لسلسلة القيم.

أ) **البيانات المبوبة:**

ـ**حالة البيانات المنفصلة(المتقطعة) (طول الفئات معدوم):** في هذه الحالة يكون المنوال هو قيمة المتغير ذات التكرار المطلق الاكبر.

مثال 10:
بالرجوع إلى المثال رقم 02.

x_i	n_i
1	15
2	10
3	5
4	20
5	30
6	20
المجموع	**100**

قيمة المنوال هي Mo=5

حالة البيانات المستمرة(المتصلة) (طول الفئات أكبر من الصفر): تسمى هذه الطريقة بطريقة الرافعة، وفيها يتم ايجاد المنوال باستخدام المعادلة التالية:

$$M_O = d + \frac{n_{i+1}}{n_{i+1} + n_{i-1}} L$$

حيث أن:
M_O: قيمة المنوال.
d: الحد الأدنى للفئة المنوالية.
L: طول الفئة المنوالية.
n_{i+1}: التكرار اللاحق لتكرار الفئة المنوالية.
n_{i-1}: التكرار السابق لتكرار الفئة المنوالية.

مثال 11:

بالرجوع إلى المثال رقم 03:

الفئات	التكرارات المطلقة n_i
[2 - 5[	7
[5 - 8[	7
[8 - 11[	9
[11 -14[	4
[14 -17[	2
[17 -20[	1
المجموع	30

نلاحظ أن الفئة المنوالية هي الفئة [8 -11[وهي الفئة التي تقابل اكبر تكرار.

إذن قيمة المنوال هي

$$M_O = d + \frac{n_{i+1}}{n_{i+1} + n_{i-1}} L = 8 + \frac{4}{4+7} * 3 = 9.09$$

أو من خلال:

$$M_O = d + \frac{\Delta_{i-1}}{\Delta_{i+1} + \Delta_{i-1}} L$$

حيث أن:

M_O: قيمة المنوال.

d: الحد الأدنى للفئة المنوالية.

L: طول الفئة المنوالية.

Δ_{i+1}: الفرق بين تكرار الفئة المنوالية وتكرار الفئة اللاحقة لها.

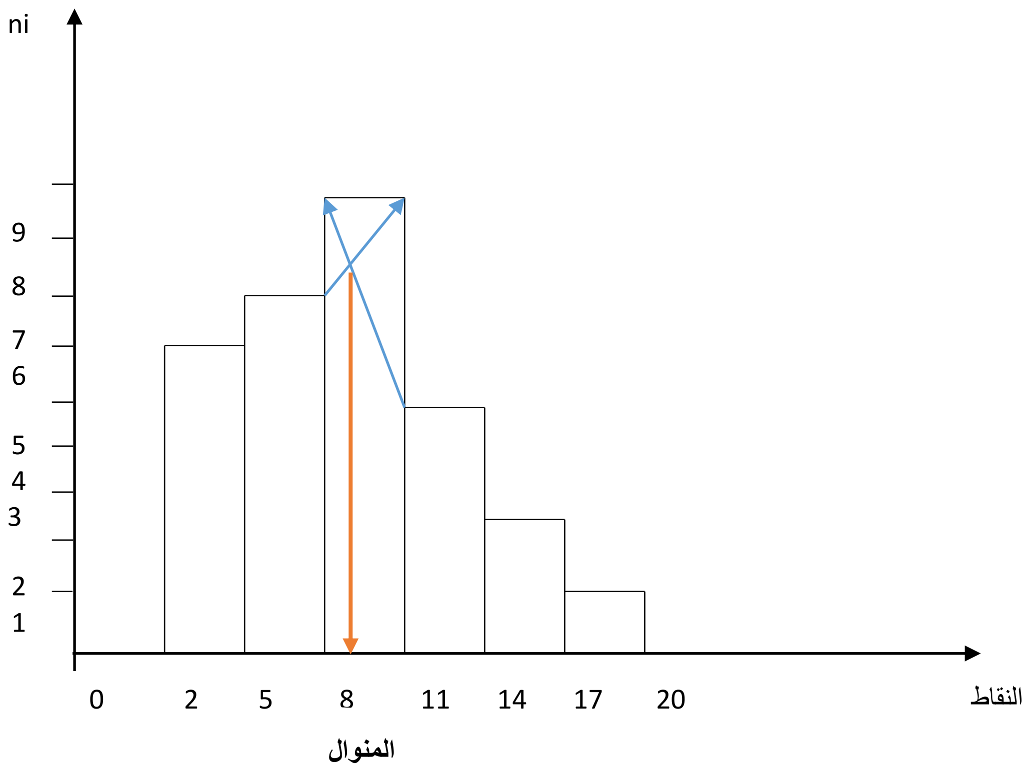
Δ_{i-1}: الفرق بين تكرار الفئة المنوالية وتكرار الفئة السابقة لها.

مثال 12:

وبطريقة أخرى يمكن حساب المنوال

$$M_O = d + \frac{\Delta_{i-1}}{\Delta_{i+1} + \Delta_{i-1}} L = 8 + \frac{2}{2+5} * 3 = 8.86$$

كيفية إيجاد منوال هذه الفئة من العرض البياني:



ملاحظة: تربط بين الوسط الحسابي $\bar{X}$ والوسيط M_e والمنوال M_O العلاقة التالية:

$$\bar{X} - M_O = 3(\bar{X} - M_e)$$

رابعا: المتوسط الهندسي، المتوسط التوافقي، المتوسط الربيعي:

	البيانات غير المبوبة	البيانات المبوبة
المتوسط الهندسي	$G = \sqrt[n]{X_1X_2X_3 \ldots .X_K}$ $G = 10^{\frac{1}{N}\sum_{i=1}^{k}\mathrm{Log}X_i}$	$G = 10^{\frac{1}{\sum_{i=1}^{k} n_i}\sum_{i=1}^{k} n_i \mathrm{Log}X_i}$
المتوسط التربيعي	$Q = \sqrt{\frac{\sum_{i=1}^{k} x_i^2}{N}}$	$Q = \sqrt{\frac{\sum_{i=1}^{k} n_i x_i^2}{\sum_{i=1}^{k} n_i}}$
المتوسط التوافقي	$H = \frac{N}{\sum_{i=1}^{k}\left[\frac{1}{X_i}\right]}$	$H = \frac{\sum_{i=1}^{k} n_i}{\sum_{i=1}^{k}\left[\frac{1}{X_i}\right] n_i}$

مثال 13:

لتكن لدينا السلسلة الإحصائية التالية: 12،15،14،13،16،14

المتوسط الهندسي:

$$G = \sqrt[n]{X_1X_2X_3 \ldots .X_K} = \sqrt[6]{14 * 16 * 13 * 14 * 15 * 12} = 13.94$$

المتوسط التربيعي:

$$Q=\sqrt{\frac{\sum_{i=1}^{k}x_i^2}{N}}=\sqrt{\frac{14^2+16^2+13^2+14^2+15^2+12^2}{6}}=14.05$$

المتوسط التوافقي:

$$H=\frac{N}{\sum_{i=1}^{k}\left[\frac{1}{X_i}\right]}=\frac{6}{\frac{1}{14}+\frac{1}{16}+\frac{1}{13}+\frac{1}{14}+\frac{1}{15}+\frac{1}{12}}=13.88$$

المتوسط الحسابي:

$$\bar{x}=\frac{\sum_{i=1}^{k}x_i}{n}=\frac{14+16+13+14+15+12}{6}=14$$

مثال 14:

إليك جدول التوزيع التكراري المطلوب حساب المتوسط التربيعي والهندسي والتوافقي.

x_i	n_i
1	15
2	10
3	5
4	20
المجموع	**50**

المتوسط الهندسي:

$$G=\sqrt[\sum n_i]{x_1^{n_1}*x_2^{n_2}*x_3^{n_3}\dots .x_k^{n_k}}=\sqrt[50]{1^{15}*2^{10}*3^{5}*4^{20}}=2.23$$

المتوسط التربيعي:

$$Q=\sqrt{\frac{\sum_{i=1}^{k}x_i^2 n_i}{\sum n_i}}=\sqrt{\frac{1^2*15+2^2*10+3^2*5+4^2*20}{50}}=2.89$$

المتوسط التوافقي:

$$H=\frac{\sum n_i}{\sum_{i=1}^{k}\left[\frac{1}{X_i}\right]n_i}=\frac{50}{\frac{15}{1}+\frac{10}{2}+\frac{5}{3}+\frac{20}{4}}=2.11$$

المتوسط الحسابي:

$$\bar{x}=\frac{\sum_{i=1}^{k}x_i n_i}{\sum n_i}=\frac{1*15+2*10+3*5+4*20}{50}=2.6$$

تمارين محلولة:

التمرين الأول:

ليكن لدينا السلسلة الإحصائية التي تمثل الأعداد التالية: 6، 9، 8، 3، 2، 13

المطلوب:

1. أحسب كل من المتوسط الحسابي، المتوسط الهندسي، المتوسط التربيعي والتوافقي؛
2. قارن بين مختلف المتوسطات المتحصل عليها.

التمرين الثاني:

ليكن لدينا التوزيع التكراري التالي

X_i	4	6	8	10	12	14	16
n_i	8	10	12	16	10	8	6

المطلوب:

1. أحسب المتوسط الحسابي، التوافقي، والتربيعي، الهندسي.
2. حدد قيمة الوسيط والمنوال.
3. حدد قيمة الربيع الثالث والعشير التاسع.

التمرين الثالث:

ليكن لدينا التوزيع التكراري التالي

الفئات	8-2	14-8	20-14	26-20	32-26
التكرارات	15	35	25	15	30

المطلوب:

حدد قيمة كل من

1. المتوسط الحسابي.
2. القيمة الأكثر شيوعا في السلسلة.
3. الوسيط.
4. الربيع الأول، والعشير الثالث، والمئين الثالث والأربعون.
5. المتوسط التوافقي والتربيعي والهندسي.

حلول التمارين:

حل التمرين الأول:

المتوسط الحسابي:

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i}{N} = \frac{13+2+3+8+9+6}{6} = \frac{41}{6} = 6.83$$

المتوسط الهندسي:

$$\mathrm{G} = \sqrt[\mathrm{N}]{\mathrm{X}_1\mathrm{X}_2\mathrm{X}_3 \ldots . \mathrm{X}_\mathrm{K}} = \sqrt[6]{13*2*3*9*8*6} = \sqrt[6]{33696} = 5.68$$

المتوسط التربيعي:

$$\mathrm{Q} = \sqrt{\frac{\sum_{\mathrm{i}=1}^{\mathrm{k}} \mathrm{x}_\mathrm{i}^2}{\mathrm{N}}} = \sqrt{\frac{13^2+2^2+3^2+9^2+8^2+6^2}{6}} = \sqrt{\frac{363}{6}} = 7.78$$

المتوسط التوافقي:

$$H = \frac{N}{\sum_{i=1}^{k}\left[\frac{1}{X_i}\right]} = \frac{6}{\frac{1}{13}+\frac{1}{2}+\frac{1}{3}+\frac{1}{9}+\frac{1}{8}+\frac{1}{6}} = 4.58$$

المقارنة: نلاحظ أن:

$$H < G < \bar{x} < Q$$

حل التمرين الثاني:

x_i	n_i	$n_i * x_i$	$n_i * x_i^2$	n_i/x_i	$n_i * \log x_i$	N^+
4	8	32	128	2	4.81	8
6	10	60	360	1.66	7.78	18
8	12	96	768	1.5	10.83	30
10	16	160	1600	1.6	16	46
12	10	120	1440	0.83	10.79	56
14	8	112	1568	0.57	9.12	64
16	6	96	1536	0.37	7.22	70
المجموع	**70**	676	7394	8.53	66.67	/

المتوسط الحسابي:

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i n_i}{\sum n_i} = \frac{676}{70} = 9.65$$

المتوسط الهندسي:

$$\mathrm{G} = \sqrt[\sum n_i]{x_1^{n_1} * x_2^{n_2} * x_3^{n_3} \ldots . x_k^{n_k}} = 10^{\frac{1}{\sum n_i}(\sum n_i \log x_i)} = 10^{\frac{66.67}{70}} = 8.91$$

المتوسط التربيعي:

$$Q=\sqrt{\frac{\sum_{i=1}^{k}x_i^2 n_i}{\sum n_i}}=\sqrt{\frac{7394}{70}}=\sqrt{105.54}=10.27$$

المتوسط التوافقي:

$$H=\frac{\sum n_i}{\sum_{i=1}^{k}\left[\frac{1}{X_i}\right]n_i}=\frac{70}{8.53}=8.20$$

1. **حساب الوسيط:**

-تحديد رتبة الوسيط

$$c=\frac{\sum n_i}{2}=\frac{70}{2}=35$$

-من التكرار المتجمع الصاعد نجد قيمة الوسيط هي: Me=10

2. **تحديد قيمة المنوال:**

المنوال هو القيمة التي تقابل أكبر تكرار، إذن قيمة المنوال تقدر ب Mo=10

تحديد قيمة الربيع الثالث:

تحديد رتبته:

$$CQ_3=\frac{3*\sum n_i}{4}=\frac{210}{4}=52.5$$

-من التكرار المتجمع الصاعد نجد قيمة الربيع الثالث هي:

Q_3=12

تحديد قيمة العشير التاسع:

تحديد رتبته:

$$CD_9=\frac{9*\sum n_i}{10}=\frac{630}{10}=63$$

-من التكرار المتجمع الصاعد نجد قيمة العشير التاسع هي:

D_9=14

حل التمرين الثالث:

الفئات	(n_i)	C_i	N^+	$n_i * c_i$	$n_i * c_i^2$	$n_i * log\ c_i$	$\frac{ni}{C_i}$
]8-2]	15	5	15	75	375	10.35	3
]14-8]	35	11	50	385	4235	36.4	3.18
]20-14]	25	17	75	425	7225	30.75	1.47
]26-20]	15	23	90	345	7935	20.4	0.65
]32-26]	30	29	120	870	25230	43.08	1.03
المجموع	**120**	/	/	**2100**	**45000**	**141.47**	**9.33**

1. المتوسط الحسابي:

$$\bar{x} = \frac{\sum_{i=1}^{k} n_i c_i}{\sum_{i=1}^{k} n_i} = \frac{2100}{120} = 17.5$$

2. القيمة الأكثر شيوعا في السلسلة (المنوال):

نلاحظ أن الفئة المنوالية هي الفئة [8 -14[وهي الفئة التي تقابل اكبر تكرار.

$$M_O = d + \frac{\Delta_{i-1}}{\Delta_{i+1} + \Delta_{i-1}} L$$

حيث أن:

M_O: قيمة المنوال.
d: الحد الأدنى للفئة المنوالية.
L: طول الفئة المنوالية.
Δ_{i+1}: الفرق بين تكرار الفئة المنوالية وتكرار الفئة اللاحقة لها.
Δ_{i-1}: الفرق بين تكرار الفئة المنوالية وتكرار الفئة السابقة لها.

$$M_O = d + \frac{\Delta_{i-1}}{\Delta_{i+1} + \Delta_{i-1}} L = 8 + \frac{(35-15)}{(35-15)+(35-25)} 6 = 11.75$$

3. الوسيط:

حساب الرتبة:

$$\mathrm{c} = \frac{\sum_{\mathrm{i}=1}^{\mathrm{k}} \mathrm{n_i}}{2} = \frac{120}{2} = 60$$

نلاحظ أن الفئة الوسيطية هي الفئة [14-20[.

$$M_e = d + \frac{C - N_{i-1}^{+}}{N_{i+1}^{+} - N_{i-1}^{+}} L = 14 + \frac{60-50}{25} * 6 = 16.4$$

4. الربيع الأول، والعشير الثالث، والمئين الثالث والأربعون:

الربيع الأول:

حساب الرتبة:

$$\mathrm{c}Q_1 = \frac{\sum_{\mathrm{i}=1}^{\mathrm{k}} \mathrm{n_i}}{4} = \frac{120}{4} = 30$$

نلاحظ أن الفئة الربيعية الأولى هي الفئة [8-14[.

$$Q_1 = d + \frac{CQ_1 - N_{i-1}^{+}}{N_{i+1}^{+} - N_{i-1}^{+}} L = 8 + \frac{30-15}{35} * 6 = 10.57$$

العشير الثالث:

حساب الرتبة:

$$cD_3 = \frac{\sum_{i=1}^{k} n_i}{4} = \frac{3 * 120}{10} = 36$$

نلاحظ أن الفئة العشيرية الثالثة هي الفئة [8-14[.

$$D_3 = d + \frac{CD_3 - N_{i-1}^{+}}{N_{i+1}^{+} - N_{i-1}^{+}} L = 8 + \frac{36 - 15}{35} * 6 = 11.6$$

المئين الثالث والأربعون:

حساب الرتبة:

$$cP_{43} = \frac{\sum_{i=1}^{k} n_i}{4} = \frac{43 * 120}{100} = 51.6$$

نلاحظ أن الفئة المئينية الثالثة والأربعون هي الفئة [*14*-20[.

$$P_{43} = d + \frac{CP_{43} - N_{i-1}^{+}}{N_{i+1}^{+} - N_{i-1}^{+}} L = 14 + \frac{51.6 - 50}{25} * 6 = 14.38$$

5. المتوسط التوافقي والتربيعي والهندسي:

المتوسط التوافقي:

$$H = \frac{\sum n_i}{\sum_{i=1}^{k} \left[\frac{1}{C_i}\right] n_i} = \frac{120}{9.33} = 12.86$$

المتوسط الهندسي:

$$\mathrm{G} = 10^{\frac{1}{\sum n_i}(\sum n_i \log c_i)} = 10^{\frac{141.47}{120}} = 15.13$$

المتوسط التربيعي:

$$\mathrm{Q} = \sqrt{\frac{\sum_{i=1}^{k} c_i^2 n_i}{\sum n_i}} = \sqrt{\frac{45000}{120}} = 19.36$$

الفصل الرابع: مقاييس التشتت

رأينا في الفصل السابق أن مقاييس الترعة المركزية تسمح لنا بتحديد القيم المتوسطة للبيانات أو أماكن تجمعها، غير أن هذه المقاييس لا تعطي فكرة واضحة ووافية عن اختلاف قيم مفردات مجموعة البيانات، ولا تحقق كل الأغراض التي يريد الباحث الوصول إليها من دراسته وتحليلاته للبيانات. فالمتوسطات لا تبين طبيعة المجموعة ولا كيفية توزيع مفرداتها حول القيمة المركزية، ولتوضيح ما سبق نورد المثال التالي الخاص ببيانات تساقط الأمطار خلال الأسبوع الأول من شهر ديسمبر(المجموعة الأولى)، ثم خلال الأسبوع الأول من شهر جانفي (المجموعة الثانية) كما يلي:

المجموعة الأولى: 10،20،30،40،50،60،70.

المجموعة الثانية: 10،30،30،40،50،50،70.

نلاحظ أن المتوسط الحسابي لكل واحدة من هاتين المجموعتين هو 40 ، وكذلك قيمة وسيطها تساوي 40 ، لكن الفرق بين قيم مختلف وحدات المجموعتين المرتبة يختلف من المجموعة الأولى عنه في المجموعة الثانية، إذ يساوي 10 بين مختلف قيم عناصر المجموعة الأولى، ويتراوح بين الصفر و 20 بين قيم المجموعة الثانية .وبذلك فإن الاختلافات بين مفردات المجموعة الأولى أقل منه بين مفردات المجموعة الثانية، ويقال اصطلاحا أن المجموعة الأولى أقل تشتتا من المجموعة الثانية .إن هذه الحقيقية تبين أن مقاييس الترعة المركزية فعلا لا تعطي فكرة وافية عن اختلاف قيم الظواهر، لذلك فإن مقاييس الترعة 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 .

إن مقاييس التشتت تقيس مدى تبعثر وتشتت مفردات السلسلة الإحصائية حول القيمة الوسطى (عادة ماتكون المتوسط الحسابي) وهي تدرس مدى اختلاف قيم السلسلة الإحصائية أو العينة المدروسة فيما بينها أو مدى قربها أو بعدها من أحد مقاييس النزعة المركزية. وكلما كانت هذه المقاييس أقرب إلى الصفر كلما كان التشتت ضعيفا أو العكس صحيح.

1) المدى:

أ) البيانات غير مبوبة:

المدى هو الفرق بين أكبر قيمة في البيانات وأصغر قيمة فيها، ويعطى بالعلاقة التالية:

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

كلما زادت قيمة المدى كلما كانت القيم غير متجانسة والعكس صحيح.

مثال01:

أوجد المدى للقيم التالية: 3، 4، 5، 7، 8، 9، 11 القيم مرتبة. أعلى قيمة 11 أقل قيمة 3 فيكون المدى W=11-3=8

ب) البيانات المبوبة:

يحسب المدى في البيانات المبوبة بعدة طرق؛ فمنها من يعتبره الفرق بين مركز الفئة الأخيرة ومركز الفئة الأولى، ومنها من يعتبره الفرق بين الحد الأعلى للفئة الأخيرة والحد الأدنى للفئة الأولى.

2) المدى الربيعي:

يحسب بالفرق بين الربيع الثالث والربيع الأول، ويعطينا فكرة عن المجال الذي تنتشر فيه نصف عدد البيانات متوسطة القيمة، ويحسب بالعلاقة:

$$\mathrm{IQ} = Q_3 - Q_1$$

3) **الانحراف المتوسط :**

هو الوسط الحسابي لفروقات القيم عن وسطها الحسابي بالقيمة المطلقة، وتختلف طريقة حسابه باختلاف طريقة تقديم البيانات:

أ) **بيانات غير مبوبة :**اذا كانت لدينا القيم التالية: $x_1, x_2, \ldots, x_k$ فإن انحرافها المتوسط يعطى بالمعادلة التالية:

$$e = \frac{\sum_{i=1}^{k} |(x_i - \bar{x})|}{N}$$

مثال02:

إليك القيم التالية لإيجاد الانحراف المتوسط: 4،6،18،9،14

$$\bar{x} = \frac{4 + 6 + 18 + 9 + 14}{5} = 10.2$$

ومنه:

$$e = \frac{|4 - 10.2| + |6 - 10.2| + |18 - 10.2| + |9 - 10.2| + |14 - 10.2|}{5} = 4.64$$

ب) البيانات المبوبة:

إذا كانت لدينا البيانات التالية: $x_1, x_2, \ldots, x_k$ ، وكانت $n_1, n_2, \ldots, n_k$ تكراراتها على الترتيب.

فإن الانحراف المتوسط يحسب بالمعادلة التالية:

- **حالة البيانات المتقطعة:**

$$e = \frac{\sum_{i=1}^{k} n_i |(x_i - \bar{x})|}{\sum_{i=1}^{k} n_i}$$

مثال 03:

عدد العمال x_i	عدد المؤسسات n_i	$n_i * x_i$	$\|(x_i - \bar{x})\|$	$n_i\|(x_i - \bar{x})\|$
1	15	*15*	1.6	24
2	10	*20*	0.6	12
3	5	*15*	0.4	6
4	20	*80*	1.4	28
المجموع	**50**	**130**	/	**70**

$$\bar{x} = \frac{130}{50} = 2.6$$

$$e = \frac{70}{50} = 1.4$$

- **حالة البيانات المستمرة:**

إذا كانت K هي عدد الفئات وكانت c_1 , c_2, …, c_k هي مراكز هذه الفئات، وكانت n_1, n_2, …, n_k تكراراتها على الترتيب. فإن الانحراف المتوسط يحسب بالمعادلة التالية:

$$e = \frac{\sum_{i=1}^{k} n_i |(c_i - \bar{x})|}{\sum_{i=1}^{k} n_i}$$

ملاحظة: الخاصية الايجابية للانحراف المتوسط انه يأخذ جميع القيم، لذلك درجة تأثره بالقيم الشاذة ضعيفة. على عكس المدى.

4) التباين والانحراف المعياري

أ) التباين:

التباين V_X أو δ^2 هو أحد مقاييس التشتت وأكثرها استخداما في النواحي التطبيقية ، ويعبر عن التباين
متوسط مربعات انحرافات القيم عن وسطها الحسابي.

-التباين في البيانات غير المبوبة:

إذا كانت القيم التالية: $X_1, X_2, X_3, \dots, X_k$ فإن التباين يعطى بالعلاقة العادية:

$$V_x = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + (X_3 - \bar{X})^2 + \cdots + (X_k - \bar{X})^2}{N} = \frac{\sum_{i=1}^{k} (X_i - \bar{X})^2}{n}$$

كما يمكن حسابه بطريقة مختصرة كما يلي:

$$V_x = \sigma^2 = \frac{\sum_{i=1}^{k} {X_i}^2}{n} - \bar{X}^2$$

مثال04:

إليك القيم التالية: 4،6،18،9،14

$$\bar{x} = \frac{4+6+18+9+14}{5} = 10.2$$

$$V_x = \frac{(4-10.2)^2 + (6-10.2)^2 + +(18-10.2)^2 + (9-10.2)^2 + (14-10.2)^2}{5}$$

$$V_x = 23.96$$

-التباين في البيانات المبوبة:

إذا كانت لدينا البيانات التالية: $x_1, x_2, \ldots, x_k$ ، وكانت $n_1, n_2, \ldots, n_k$ تكراراتها على الترتيب. فإن الانحراف المتوسط يحسب بالمعادلة التالية:

حالة البيانات المتقطعة: يحسب التباين بالعلاقة التالية:

$$V_x = \frac{\sum_{i=1}^{k} n_i (X_i - \bar{X})^2}{\sum_{i=1}^{k} n_i}$$

كما يمكن حسابه بطريقة مختصرة كما يلي:

$$V_x = \frac{\sum_{i=1}^{n} n_i {X_i}^2}{\sum_{i=1}^{k} n_i} - \bar{X}^2$$

مثال05: بالرجوع إلى المثال 03:

عدد العمال x_i	عدد المؤسسات n_i	$n_i * x_i$	$(x_i-\bar{x})^2$	$n_i(x_i-\bar{x})^2$
1	15	*15*	2.56	38.4
2	10	*20*	0.36	3.6
3	5	*15*	0.16	0.8
4	20	*80*	1.96	39.2
المجموع	50	130	/	114.4

$$V_x = \frac{114.4}{50} = 2.28$$

– **حالة البيانات المستمرة:**

إذا كانت K هي عدد الفئات وكانت $c_1, c_2, \ldots, c_k$ هي مراكز هذه الفئات، وكانت $n_1, n_2, \ldots, n_k$ تكراراتها على الترتيب. فإن الانحراف المتوسط يحسب بالمعادلة التالية:

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

ب) الانحراف المعياري

يعتمد التباين علي مجموع مربعات الانحرافات، وهو ما لا يتماشى مع وحدات قياس المتغير محل الدراسة، من أجل ذلك لجأ الإحصائيين إلى مقياس منطقي يأخذ في الاعتبار الجذر التربيعي للتباين يناسب وحدات قياس المتغير، هذا المقياس هو الانحراف المعياري δ_X، إذا الانحراف المعياري هو الجذر التربيعي الموجب للتباين، أي أن:

$$\delta_X = \sqrt{V_X}$$

-الانحراف المعياري في البيانات غير المبوبة:

$$\delta_X = \sqrt{V_X} = \sqrt{\frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + (X_3 - \bar{X})^2 + \cdots + (X_n - \bar{X})^2}{N}}$$

$$\delta_X = \sqrt{V_X} = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{N}}$$

كما يمكن حسابه بطريقة مختصرة كما يلي:

$$\delta_X = \sqrt{V_X} = \sqrt{\frac{\sum_{i=1}^{n} {X_i}^2}{n} - \bar{X}^2}$$

مثال06 :

إليك القيم التالية: 4،6،18،9،14

$$\bar{x} = \frac{4 + 6 + 18 + 9 + 14}{5} = 10.2$$

$$\delta_X = \sqrt{V_X}$$

$$= \sqrt{\frac{(4 - 10.2)^2 + (6 - 10.2)^2 + +(18 - 10.2)^2 + (9 - 10.2)^2 + (14 - 10.2)^2}{5}}$$

$$\delta_X = \sqrt{V_X} = 4.89$$

-الانحراف المعياري في البيانات المبوبة

إذا كانت لدينا البيانات التالية: $x_1, x_2, \ldots, x_k$ ، وكانت $n_1, n_2, \ldots, n_k$ تكراراتها على الترتيب. فإن الانحراف المتوسط يحسب بالمعادلة التالية:

حالة البيانات المتقطعة: يحسب التباين بالعلاقة التالية:

$$\delta_X = \sqrt{V_X} = \sqrt{\frac{\sum_{i=1}^{k} n_i (X_i - \bar{X})^2}{\sum_{i=1}^{k} n_i}}$$

كما يمكن حسابه بطريقة مختصرة كما يلي:

$$\delta_X = \sqrt{V_X} = \sqrt{\frac{\sum_{i=1}^{n} n_i {X_i}^2}{\sum_{i=1}^{k} n_i} - \bar{X}^2}$$

مثال07: بالرجوع إلى المثال 03:

عدد العمال x_i	عدد المؤسسات n_i	$n_i * x_i$	$(x_i - \bar{x})^2$	$n_i(x_i - \bar{x})^2$
1	15	*15*	2.56	38.4
2	10	*20*	0.36	3.6
3	5	*15*	0.16	0.8
4	20	*80*	1.96	39.2
المجموع	**50**	**130**	/	**114.4**

$$\delta_X = \sqrt{V}_x = \sqrt{\frac{114.4}{50}} = \sqrt{2.28} = 1.50$$

حالة البيانات المستمرة:

إذا كانت K هي عدد الفئات وكانت c_1 , c_2 , ..., c_k هي مراكز هذه الفئات، وكانت n_1, n_2, ..., n_k تكراراتها على الترتيب. فإن الانحراف المتوسط يحسب بالمعادلة التالية:

$$\delta_X = \sqrt{V}_x = \sqrt{\frac{\sum_{i=1}^{k} n_i (c_i - \bar{x})^2}{\sum_{i=1}^{k} n_i}}$$

ملاحظة:

يأخذ الانحراف المعياري في الحسبان جميع القيم، كما ان قيمته صغيرة وبالتالي يمكن ان تعطي خلاصة واضحة عن مدى تباعد القيم، اذ كلما كانت هذه القيمة صغيرة دل ذلك على ان القيم ليست متباعدة عن الوسط الحسابي وبالتالي فهي اقل تشتتا ووسطها الحسابي يمثلها تمثيلا جيدا.

5) الانحراف الربيعي ومعامل الاختلاف

أ) **الانحراف الربيعي:**

إذا كان لدينا بيانات إحصائية ربيعها الأول هو Q1 وربيعها الثالث هو Q3، فإن انحرافها الربيعي يعطى بالعلاقة التالية كما يلي:

$$EQ = \frac{Q_3 - Q_1}{2}$$

أما الانحراف الربيعي النسبي:

$$EQP = \frac{Q_3 - Q_1}{Q_2}$$

ملاحظة :الانحراف المتوسط=اربعة اخماس الانحراف المعياري، اي:

$$e = \frac{4}{5}\delta$$

والانحراف الربيعي =ثلثي الانحراف المعياري، اي:

$$EQ = \frac{2}{3}\delta$$

ب) معامل الاختلاف:

رأينا في خصائص الانحراف المعياري أنه إذا كانت الظواهر من صفات مختلفة أو إذا كانت متوسطاتها مختلفة، فإن المقارنة اعتمادا عليه ستكون غير منطقية وغير صحيحة، لهذا وجدت مقاييس أخرى تعرف بمقاييس التشتت النسبي تعتمد على فكرة مقارنة البيانات والظواهر في شكل نسبة مئوية اعتمادا على متوسطاتها، من أهم المقاييس معامل الاختلاف CV.

يحسب معامل الاختلاف بقسمة الانحراف المعياري للبيانات على المتوسط الذي حسبت حوله، ويعطى بالعلاقة التالية:

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

مثال08:

إذا كانت لديك البيانات التالية حول ظاهرتين من نفس النوع:

الظاهرة A: $\delta_x = 1.7 \quad \bar{X} = 15$

الظاهرة B: $\delta_x = 1.2 \quad \bar{X} = 11$

أي الظاهرتين أكثر تشتت؟

الحل: إذا ما استخدمنا الانحراف المعياري في المقارنة، فإننا نلاحظ أن الظاهرة B أكثر تجانسا من الظاهرة A ، أو أن الظاهرةA أكثر تشتتا من الظاهرةB ، لكن بما أن المتوسط الحسابي الذي حسب من حوله الانحراف المعياري مختلف، فإن المقارنة هنا لا تصح، بل يجب الاحتكام الى مقياس يوحد نمط المقارنة وهو معامل الاختلاف كما يلي:

$$CV_A = \frac{\delta_x}{\bar{X}} \times 100 = \frac{1.7}{15} \times 100 = 11.33\%$$

$$CV_B = \frac{\delta_x}{\bar{X}} \times 100 = \frac{1.2}{11} \times 100 = 10.90\%$$

إذن بعد استخدام معامل الاختلاف تبين أن الظاهرة B أقل تشتتا من الظاهرة A لأن

$$CV_B < CV_A$$

-خصائص معامل الاختلاف:

ما يميز معامل الاختلاف عن بقية مقاييس التشتت أنه غير مرتبط بوحدة القياس المستعملة، وهذا ما يسمح بقياس ومقارنة تشتت ظاهرتين مختلفتين من حيث الوحدات المستعملة.

تمارين محلولة:

التمرين الأول:

ليكن لدينا السلسلة الإحصائية التي تمثل الأعداد التالية: 6، 9، 8، 3،11، 2، 7،13.

المطلوب:

-أحسب كل من المدى الربيعي، الانحراف المتوسط، التباين، الانحراف المعياري.

التمرين الثاني: ليكن لدينا التوزيع التكراري التالي

X_i	4	6	8	10	12	14	16
n_i	8	10	12	16	10	8	6

المطلوب:

1. أحسب كل من المدى الربيعي، الانحراف المتوسط، التباين، الانحراف المعياري.
2. أحسب كل من الانحراف الربيعي، معامل الاختلاف.

التمرين الثالث:

يكن لدينا التوزيع التكراري التالي

الفئات	8-2	14-8	20-14	26-20	32-26
التكرارات	15	35	25	15	30

المطلوب:

1. أحسب كل من الانحراف المتوسط، التباين، الانحراف المعياري.
2. أحسب معامل الاختلاف.

التمرين الرابع:

ليكن السلسلتان الاحصائيتان التالتان A، B:

A : 5 ;18 ;10 ;15 ;3 ;7 ;6 ;12

B : 18 ;9 ;8 ;9 ;8 ;8 ;3 ;9

المطلوب:

1. هل يمكن استعمال المدى العام للمقارنة بين السلسلتين؟
2. ماهو مقياس التشتت المناسب في المقارنة بين السلسلتين؟

حلول التمارين:

حل التمرين الأول:

حساب المدى الربيعي:

$$\mathrm{IQ} = Q_3 - Q_1$$

الربيع الأول: تحديد رتبته

$$CQ_1 = \frac{N}{4} = \frac{8}{4} = 2$$

ترتيب القيم:

2/3/6/7/8/9/11/13

$$Q_1 = 3$$

الربيع الثالث: تحديد رتبته

$$CQ_3 = \frac{3 * N}{4} = \frac{24}{4} = 6$$

$$Q_3 = 9$$

إذن المدى الربيعي يساوي:

$$\mathrm{IQ} = Q_3 - Q_1 = 9 - 3 = 6$$

الانحراف المتوسط:

$$\bar{x} = \frac{\sum x_i}{N} = \frac{59}{8} = 7.375$$

$$e = \frac{\sum_{i=1}^{k} |(x_i - \bar{x})|}{N} = \frac{23}{8} = 2.875$$

التباين:

$$V_x = \frac{\sum_{i=1}^{k} (x_i - \bar{x})^2}{N} = \frac{97.875}{8} = 12.23$$

الانحراف المعياري:

$$\delta_X = \sqrt{V_X} = \sqrt{12.23} = 3.49$$

حل التمرين الثاني:

x_i	n_i	$n_i * x_i$	$n_i\|x_i - \bar{x}\|$	$n_i(x_i - \bar{x})^2$	N^+
4	8	32	45.2	255.38	8
6	10	60	36.5	133.225	18
8	12	96	19.8	32.67	30
10	16	160	5.6	1.96	46
12	10	120	23.5	55.225	56
14	8	112	34.8	151.38	64
16	6	96	38.1	241.935	70
المجموع	**70**	676	203.5	871.775	/

المتوسط الحسابي:

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i n_i}{\sum n_i} = \frac{676}{70} = 9.65$$

حساب المدى الربيعي:

$$\text{IQ} = Q_3 - Q_1$$

الربيع الأول:

$$CQ_1 = \frac{70}{4} = 17.5 \Rightarrow Q_1 = 6$$

الربيع الثالث:

$$CQ_3 = \frac{3 * 70}{4} = 52.5 \Rightarrow Q_3 = 12$$

إذن المدى الربيعي يساوي:

$$\text{IQ} = Q_3 - Q_1 = 12 - 6 = 6$$

الانحراف المتوسط:

$$e = \frac{\sum_{i=1}^{k} n_i |(x_i - \bar{x})|}{\sum_{i=1}^{k} n_i} = \frac{203.5}{70} = 2.90$$

التباين:

$$V_x = \frac{\sum_{i=1}^{k} n_i (x_i - \bar{x})^2}{\sum_{i=1}^{k} n_i} = \frac{871.775}{70} = 12.45$$

الانحراف المعياري:

$$\delta_X = \sqrt{V_X} = \sqrt{12.45} = 3.52$$

الانحراف الربيعي:

$$EQ = \frac{Q_3 - Q_1}{2} = \frac{12 - 6}{2} = 3$$

معامل الاختلاف:

$$CV = \frac{\delta_x}{\overline{X}} \times 100 = \frac{3.52}{9.65} * 100 = 36.47\%$$

حل التمرين الثالث:

الفئات	n_i	C_i	$n_i\lvert c_i-\bar{x}\rvert$	$n_i(c_i-\bar{x})^2$
]8-2]	15	5	187.5	2343.75
]14-8]	35	11	227.5	1478.75
]20-14]	25	17	12.5	6.25
]26-20]	15	23	82.5	453.75
]32-26]	30	29	34.5	3967.5
المجموع	**120**	/	**855**	**8250**

المتوسط الحسابي:

$$\bar{x}=\frac{\sum_{i=1}^{k}n_ic_i}{\sum_{i=1}^{k}n_i}=\frac{2100}{120}=17.5$$

الانحراف المتوسط:

$$e=\frac{\sum_{i=1}^{k}n_i|(c_i-\bar{x})|}{\sum_{i=1}^{k}n_i}=\frac{855}{120}=7.125$$

التباين:

$$V_x=\frac{\sum_{i=1}^{k}n_i(c_i-\bar{x})^2}{\sum_{i=1}^{k}n_i}=\frac{8250}{120}=68.75$$

الانحراف المعياري:

$$\delta_X=\sqrt{V_X}=\sqrt{68.75}=8.29$$

معامل الاختلاف:

$$CV=\frac{\delta_x}{\bar{X}}\times 100=\frac{8.29}{17.5}*100=47.37\%$$

حل التمرين الرابع:

لا يمكن استعمال المدى العام للمقارنة بين السلسلتين A و B لأن:

$$W_A=18-3=15$$

$$W_B=18-3=15$$

مقياس التشتت المناسب للمقارنة هو معامل الاختلاف

$$\bar{X}_A=\frac{\sum x_{iA}}{N}=\frac{76}{8}=9.5$$

$$\bar{X}_B = \frac{\sum x_{iB}}{N} = \frac{72}{8} = 9$$

$$\delta_A = \sqrt{\frac{\sum {x_{iA}}^2}{N} - {\bar{X}_A}^2} = \sqrt{\frac{912}{8} - 9.5^2} = 4.87$$

$$\delta_B = \sqrt{\frac{\sum {x_{iB}}^2}{N} - {\bar{X}_B}^2} = \sqrt{\frac{768}{8} - 9^2} = 3.87$$

$$CV_A = \frac{\delta_A}{\bar{X}_A} \times 100 = \frac{4.87}{9.5} * 100 = 51.26\%$$

$$CV_B = \frac{\delta_B}{\bar{X}_B} \times 100 = \frac{3.87}{9} * 100 = 43\%$$

معامل الاختلاف للسلسلة A أكبر من معامل الاختلاف للسلسلة B إذن السلسلة A أكثر تشتتا من السلسلة B.

الفصل الخامس: مقاييس الشكل

بعدما تطرقنا إلى مقاييس النزعة المركزية التي تقيس مدى تمركز القيم، ومقاييس التشتت التي تقيس مدى تباعد أو تقارب القيم، نتطرق في هذا المحور إلى قياس تماثل التوزيع من خلال مقاييس الالتواء والتفرطح. فمقاييس النزعة المركزية ومقاييس التشتت لا تكفي وحدها في وصف البيانات، فقد نجد التمثيلات متساوية في وسطها الحسابي وانحرافها المعياري، لكنها تختلف من حيث الشكل، أي إما تكون إحداها ملتوية أو متفرطحة عن الاخرى، وتسمى هذه المقاييس بمقاييس الشكل، وتعتمد هذه الأخيرة على العزوم البسيطة والعزوم المركزية.

أولا: العزوم:

العزوم تحسب في ثلاثة حالات: إما حول نقطة الأصل، أو حول المتوسط الحسابي أو أي نقطة معينة. مثلا العزم المركزي هو العزم حول الوسط الحسابي وهو انحرافات القيم عن الوسط الحسابي، أما رتبة العزم فهي درجة الأس هذه الانحرافات عن القيمة (سواء قيمة أصلية، وسط حسابي أو أي قيمة أخرى)

1) العزم البسيط:

وهو العزم البسيط من الدرجة r (حول الصفر)، عبارة عن المتوسط الحسابي لقيم المتغير الاحصائي مرفوعة إلى القوة (الاس) r، ولتكن $x_1^r, x_2^r, x_3^r, \dots, x_n^r$ ويحسب رياضيا بالعلاقة التالية:

أ) **بيانات غير مبوبة:** يعطى بالعلاقة التالية:

$$m_r = \frac{\sum_{i=1}^{n} x_i^r}{N}$$

ب) **بيانات مبوبة** يعطى بالعلاقة التالية:

$$m_r = \frac{\sum_{i=1}^{n} n_i x_i^r}{N}$$

حيث r تمثل درجة العزم (r=1.2…)

إن مراتب العزوم البسيطة تتراوح بين 0 وr، وهذه بعض الحالات الخاصة:

$$m_0 = 1;\ m_1 = \bar{X};\ \ m_2 = Q^2.$$

2) العزم المركزي:

وهو العزم المركزي من الدرجة r حول المتوسط الحسابي ويحسب رياضيا بالعلاقة التالية:

أ) **بيانات غير مبوبة:** يعطى بالعلاقة التالية:

$$\mu_r = \frac{\sum_{i=1}^{n}(x_i - \bar{X})^r}{N}$$

ب) **بيانات مبوبة** يعطى بالعلاقة التالية:

$$\mu_r = \frac{\sum_{i=1}^{n} n_i(x_i - \bar{X})^r}{N}$$

حيث r تمثل درجة العزم (r=1.2…)

ملاحظة: العزم المركزي من الدرجة الأولى يساوي الصفر أي $\mu_1 = 0$، والعزم المركزي من الدرجة الثانية هو التباين أي $\mu_1 = \delta_x^2$

مثال 1:

البيانات الآتية تمثل التوزيع التكراري لأعداد العمال في 100مؤسسة صغيرة

x_i	1	2	3	4	5	6	المجموع
n_i	15	10	5	20	30	20	100

المطلوب:

1. أحسب العزم البسيط من الدرجة الأولى ومن الدرجة الثانية ومن الدرجة الثالثة؟
2. أحسب العزم المركزي من الدرجة الأولى والثانية والرابعة؟

الحل:

x_i	n_i	$n_i x_i^1$	$n_i x_i^2$	$n_i x_i^3$	$n_i(x_i - \bar{X})^1$	$n_i(x_i - \bar{X})^2$	$n_i(x_i - \bar{X})^4$
1	15	15	15	15	-45	135	1215
2	10	20	40	80	-20	40	160
3	5	15	45	135	-5	5	5
4	20	80	320	1280	0	0	0
5	30	150	750	3750	30	30	30
6	20	120	720	4320	40	80	320
المجموع	**100**	400	1890	9580	0	290	1730

1. **حساب العزم البسيط:**

$$m_r = \frac{\sum_{i=1}^{n} n_i x_i^r}{N}$$

أ) حساب العزم البسيط من الدرجة الأولى:

$$m_1 = \frac{\sum_{i=1}^{n} n_i x_i^1}{N} = \frac{400}{100} = 4$$

ب) حساب العزم البسيط من الدرجة الثانية:

$$m_2 = \frac{\sum_{i=1}^{n} n_i x_i^2}{N} = \frac{1890}{100} = 18.9$$

ج) حساب العزم البسيط من الدرجة الثالثة:

$$m_3 = \frac{\sum_{i=1}^{n} n_i x_i^3}{N} = \frac{9580}{100} = 95.8$$

2. حساب العزم المركزي:

$$\mu_r = \frac{\sum_{i=1}^{n} n_i (x_i - \bar{X})^r}{N}$$

أ) حساب العزم المركزي من الدرجة الأولى:

$$\mu_1 = \frac{\sum_{i=1}^{n} n_i (x_i - \bar{X})^1}{N} = \frac{\sum_{i=1}^{n} n_i (x_i - 4)^1}{100} = \frac{0}{100} = 0$$

ب) حساب العزم المركزي من الدرجة الثانية:

$$\mu_2 = \frac{\sum_{i=1}^{n} n_i (x_i - \bar{X})^2}{N} = \frac{\sum_{i=1}^{n} n_i (x_i - 4)^2}{100} = \frac{290}{4} = 2.9$$

ج) حساب العزم المركزي من الدرجة الرابعة:

$$\mu_4 = \frac{\sum_{i=1}^{n} n_i (x_i - \bar{X})^4}{N} = \frac{1730}{100} = 17.30$$

ثانيا: الالتواء

1) أشكال التوزيعات التكرارية:

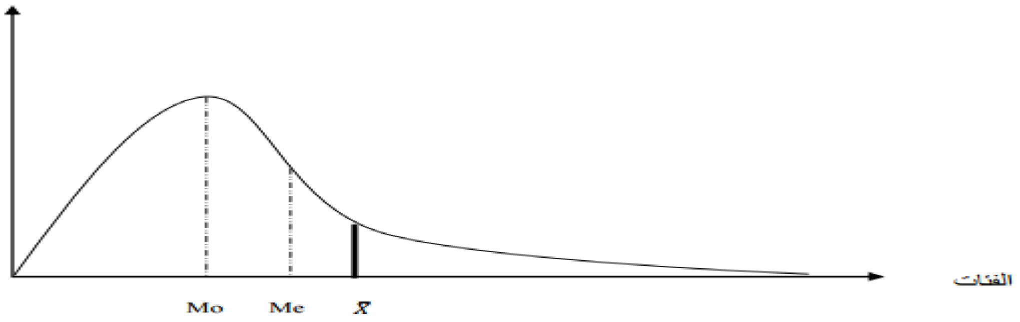
إن التوزيع التكراري يمكن أن يأخذ عدة أشكال، فقد يكون التوزيع مائلا نحو الموجب، نقول عنه أنه توزيع تكراري ملتوي نحو اليمين (أو موجب الالتواء)، قد يكون مائلا نحو السالب، نقول عنه أنه توزيع تكراري ملتوي نحو اليسار (سالب الالتواء)، وقد يكون توزيع تكراري متماثل، أي متناظر، وكل نصف ينطبق على الآخر.

أ) توزيع تكراري ملتوي نحو اليمين (موجب الالتواء):

هذا التوزيع يكون ممتد نحو اليمين (نحو ∞+)، ورياضيا يكون:

$$\bar{X} > \mathrm{Me} > Mo$$

كما نلاحظ في الشكل التالي:

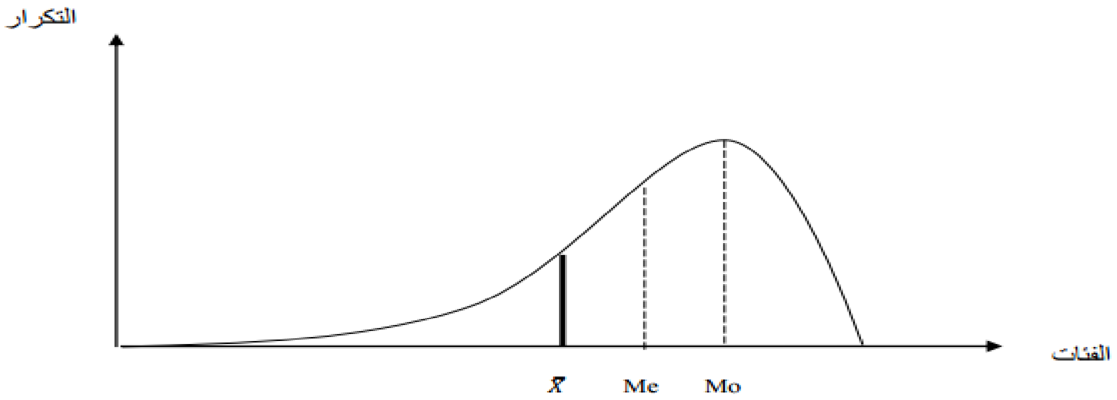


ب) توزيع تكراري ملتوي نحو اليسار (سالب الالتواء):

هذا التوزيع يكون ممتد نحو اليسار (نحو المبدأ)، ورياضيا يكون:

$$\bar{X} < \mathrm{Me} < Mo$$

كما نلاحظ في الشكل التالي:

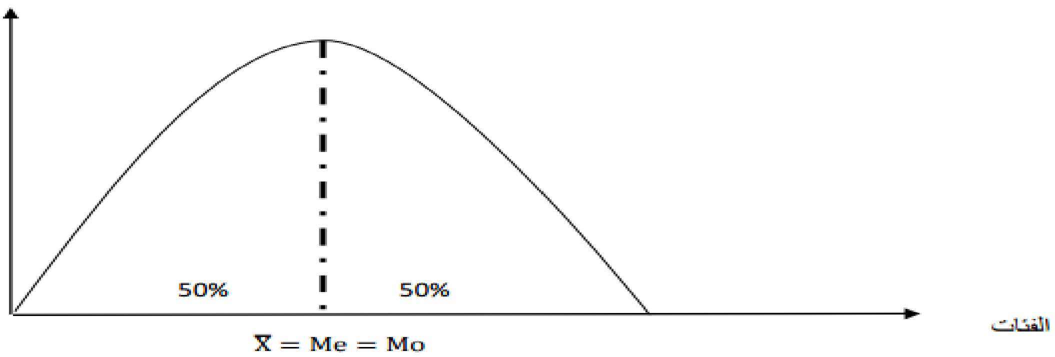


ج) توزيع تكراري متماثل:

هو التوزيع الذي يكون متناظر والوسط الحسابي يقسمه إلى صنفين متساويين، ورياضيا يتساوى فيه الوسط الحسابي، الوسيط والمنوال، أي:

$$\bar{X} = \mathrm{Me} = Mo$$

كما نلاحظ في الشكل التالي:



2) مقاييس الالتواء:

أ) مقاييس بيرسون (Person) للالتواء:

1. معامل بيرسون الأول P_1:

يتحدد هذا المقياس باستخدام العلاقة بين الوسط الحسابي والمنوال، حيث يحسب بالصيغة التالية:

$$P_1 = \frac{\bar{X} - Mo}{\delta_x}$$

من هذه الصيغة يتحدد شكل التوزيع حيث:

- ✓ إذا كان $P_1 = 0$ فإن التوزيع يكون متماثل.
- ✓ إذا كان $P_1 > 0$ فإن التوزيع يكون مائل (ملتوي) نحو اليمين (موجب الالتواء).
- ✓ إذا كان $P_1 < 0$ فإن التوزيع يكون مائل (ملتوي) نحو اليسار (سالب الالتواء).

2. معامل بيرسون الثاني P_2:

يتحدد هذا المقياس باستخدام العلاقة بين الوسط الحسابي والوسيط، حيث يحسب بالصيغة التالية:

$$P_2 = \frac{3(\bar{X} - Me)}{\delta_x}$$

من هذه الصيغة يتحدد شكل التوزيع حيث:

- ✓ إذا كان $P_2 = 0$ فإن التوزيع يكون متماثل.
- ✓ إذا كان $P_2 > 0$ فإن التوزيع يكون مائل (ملتوي) نحو اليمين (موجب الالتواء).
- ✓ إذا كان $P_2 < 0$ فإن التوزيع يكون مائل (ملتوي) نحو اليسار (سالب الالتواء).

3. معامل فيشر للالتواء:

يحسب هذا المقياس باستخدام العلاقة بين العزم من الدرجة الثالثة والانحراف المعياري، حيث يحسب رياضيا:

$$F_1 = \frac{\mu_3}{\delta^3}$$

حيث: μ_3هو العزم المركزي من الدرجة الثالثة، δ هو الانحراف المعياري.

ويتم تحديد شكل التوزيع بناء على قيمة العزم المركزي من الدرجة الثالثة، حيث إذا كان يساوي الصفر (μ_3=0)فإن التوزيع يكون متناظر، وفي جميع الحلات يكون كما يلي:

- ✓ إذا كان $F_1 = 0$ فإن التوزيع يكون متناظر(متماثل).
- ✓ إذا كان $F_1 > 0$ فإن التوزيع يكون مائل (ملتوي) نحو اليمين (موجب الالتواء).
- ✓ إذا كان $F_1 < 0$ فإن التوزيع يكون مائل (ملتوي) نحو اليسار (سالب الالتواء).

ويستخدم هذا المعامل في حالة إذا كان التوزيع وحيد المنوال، وتنحصر قيمته بين +1 و-1.

4. معامل يول للالتواء Y_1:

يحسب هذا المقياس باستخدام الربيعات، ويستخدم في حالة يكون التوزيع التكراري مفتوح من أحد الجهتين أو من الجهتين مع بعض، ويسمى كذلك معامل الالتواء الربيعي، ويحسب رياضيا بالعلاقة التالية:

$$Y_1 = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{Q_3 - Q_1} = \frac{Q_3 - 2Q_2 + Q_1}{Q_3 - Q_1}$$

أما شكل التوزيع حسب هذا المقياس يتحدد كما يلي:

- ✓ إذا كان $Y_1 = 0$ فإن التوزيع يكون متناظر(متماثل).
- ✓ إذا كان $Y_1 > 0$ فإن التوزيع يكون مائل (ملتوي) نحو اليمين (موجب الالتواء).
- ✓ إذا كان $Y_1 < 0$ فإن التوزيع يكون مائل (ملتوي) نحو اليسار (سالب الالتواء).

مثال02:

الجدولين التاليين يمثل مجموعة من العمال حسب الأجور التي يتقاضونها في المؤسستين AوB.

المؤسسة A	
الأجر	**عدد العمال**
15000	3
18000	17
21000	10
25000	9
30000	11
المجموع	50

المؤسسة B	
الأجر	**عدد العمال**
26000	4
30000	9
35000	12
40000	16
45000	9
المجموع	50

المطلوب:

1. أوجد معامل بيرسون وفيشر للالتواء؟
2. قارن بين التواء التوزيعين، أثبت النتيجة أيضا بالرسم؟

الحل:

حساب معامل بيرسون وفيشر للالتواء:

قبل حساب معامل بيرسون وفيشر للالتواء يجب أولا حساب المتوسط الحسابي والوسيط والمنوال والانحراف المعياري لكل مؤسسة:

المؤسسة B				المؤسسة A			
$n_{iB}x_{iB}^2$	$n_{iB}x_{iB}$	n_{iB}	x_{iB}	$n_{iA}x_{iA}^2$	$n_{iA}x_{iA}$	n_{iA}	x_{iA}
2704000000	104000	4	26000	675000000	45000	3	15000
7200000000	240000	9	30000	5508000000	306000	17	18000
14700000000	420000	12	35000	4410000000	210000	10	21000
27200000000	680000	16	40000	5625000000	225000	9	25000
18225000000	405000	9	45000	9900000000	330000	11	30000
70029000000	**1849000**	**50**	**المجموع**	**675000000**	**1116000**	**50**	**المجموع**

حساب المتوسط الحسابي للمؤسستين A وB:

$$\overline{X_A} = \frac{\sum_{i=1}^{k} n_{iA}x_{iA}}{N} = \frac{1116000}{50} = 22320$$

$$\overline{X_B} = \frac{\sum_{i=1}^{k} n_{iB}x_{iB}}{N} = \frac{1849000}{50} = 36980$$

حساب الوسيط للمؤسستين A وB:

$$M_{eA} = 21000$$

$$M_{eB} = 40000$$

حساب المنوال للمؤسستين A وB:

$$M_{oA} = 18000$$

$$M_{oB} = 40000$$

حساب الانحراف المعياري للمؤسستين A وB:

$$\delta_A = \sqrt{V_A} = \sqrt{\frac{\sum_{i=1}^{n} n_{iA}{X_{iA}}^2}{\sum_{i=1}^{k} n_{iA}} - \overline{X_A}^2} = \sqrt{24177600} = 4917.07$$

$$\delta_B = \sqrt{V_B} = \sqrt{\frac{\sum_{i=1}^{n} n_{iB}{X_{iB}}^2}{\sum_{i=1}^{k} n_{iB}} - \bar{X}^2} = \sqrt{33059600} = 5947.75$$

معامل بيرسون الأول P_1:

$$P_{1A} = \frac{\overline{X_A} - M_{oA}}{\delta_A} = \frac{22320 - 18000}{4917.07} = \frac{4320}{4917.07} = 0.87$$

$$P_{1B} = \frac{\overline{X_B} - M_{oB}}{\delta_B} = \frac{36980 - 40000}{5947.75} = \frac{-3020}{5947.75} = -0.52$$

معامل بيرسون الأول P_2:

$$P_{2A} = \frac{3(\overline{X_A} - M_{eA})}{\delta_A} = \frac{3(22320 - 21000)}{4917.07} = \frac{9360}{4917.07} = 0.80$$

$$P_{2B} = \frac{3(\overline{X_B} - M_{eB})}{\delta_B} = \frac{3(36980 - 40000)}{5947.75} = \frac{-9060}{5947.75} = -1.57$$

نلاحظ حسب معامل بيرسون الأول والثاني، أن التوزيع الخاص بالمؤسسة B ملتوي نحو اليسار، بينما التوزيع الخاص بالمؤسسة A ملتوي نحو اليمين.

معامل فيشر للالتواء:

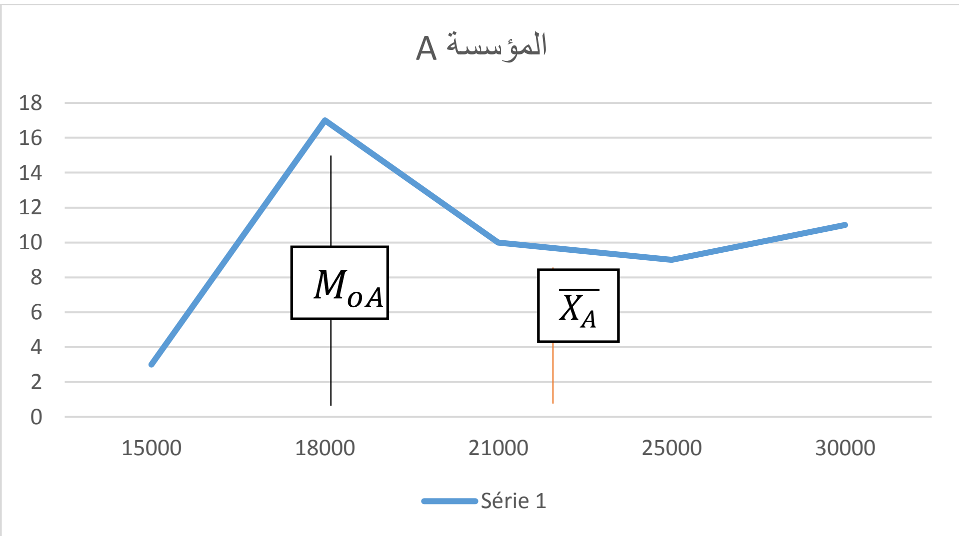
$$F_{1A} = \frac{\mu_{3A}}{{\delta_A}^3} = \frac{\frac{\sum_{i=1}^{n} n_{iA}(x_{iA}-\overline{X_A})^3}{N}}{{\delta_A}^3} = 0.43$$

$$F_{1B} = \frac{\mu_{3B}}{{\delta_B}^3} = \frac{\frac{\sum_{i=1}^{n} n_{iB}(x_{iB}-\overline{X_B})^3}{N}}{{\delta_B}^3} = -0.31$$

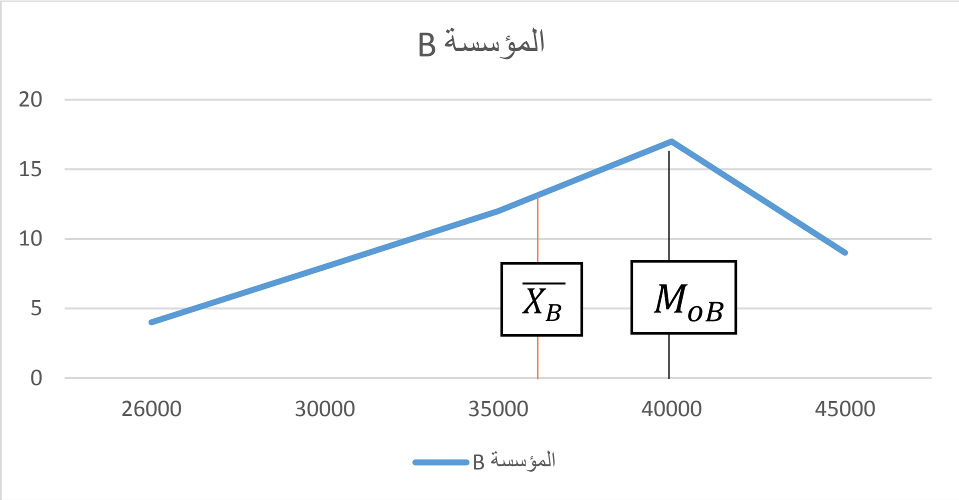
نفس الشيء حيث نلاحظ حسب معامل فيشر، أن التوزيع الخاص بالمؤسسة B ملتوي نحو اليسار (سالب)، بينما التوزيع الخاص بالمؤسسة A ملتوي نحو اليمين (موجب).

و هذا ما يبينه الشكل البياني:

المنحنى التكراري للتوزيع الخاص بالمؤسسة A



المنحنى التكراري للتوزيع الخاص بالمؤسسة B



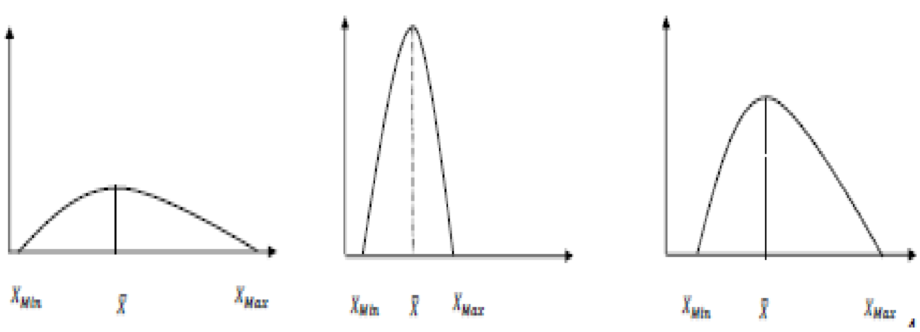
ثالثا: التفرطح

كما سبق وأن رأينا فقد يكون المنحنى التكراري متماثلا أو ملتويا نحو اليمين أو إلى اليسار، وفي كلا الحالتين قد يكون المنحنى عالي القمة (غير مفرطح) أو منبسط القمة (متوسط التفرطح أو مفرطح)، وقياس التفرطح أو الانبساط يعني قياس مدى تشتت القيم عن بعضها البعض، أو يقيس قمة التوزيع بالنسبة للتوزيع الطبيعي، والتفرطح يأخذ ثلاث أشكال رئيسية :

- ✓ **توزيع متطاول أو مدبب:** حيث يكون تشتت البيانات ضعيف جدا وتكون قمة التوزيع عالية؛
- ✓ **توزيع طبيعي**: حيث يكون تشتت البيانات ليس كبير وليس صغير، أي معتدل؛
- ✓ **توزيع متفرطح (منبسط):** حيث يكون تشتت البيانات كبير جدا.

والشكل البياني التالي يوضح مختلف أشكال التفرطح:

توزيع معتدل **توزيع مدبب** **توزيع متفرطح**



1) قياس التفرطح:

أ) معامل بيرسون للتفرطح P_4:

يحسب هذا المقياس باستخدام العلاقة بين العزوم المركزية حيث يكتب رياضيا بالعلاقة التالية:

$$P_4 = \frac{\mu_4}{\mu_2^2} = \frac{\mu_4}{\delta^4}$$

حيث، μ_4 : العزم المركزي من الدرجة الرابعة، μ_2 : العزم المركزي من الدرجة الثانية.

ويتم تحديد شكل التوزيع من جانب التفرطح كما يلي:

- ✓ إذا كان $P_4 = 3$ فإن منحنى التوزيع يكون طبيعي أو على شكل جرس.
- ✓ إذا كان $P_4 > 3$ فإن منحنى التوزيع يكون مدبدب (متطاول).
- ✓ إذا كان $P_4 < 3$ فإن منحنى التوزيع يكون متفرطح (تشتت قوي).

ب) معامل فيشر للتفرطح F_2:

هو عبارة عن مقياس بيرسون للتفرطح مطروحا منه قيمة 3، ويكتب رياضيا كما يلي:

$$F_2 = P_4 - 3$$

$$F_2 = \frac{\mu_4}{\delta^4} - 3$$

ويتم تحديد شكل التوزيع وفق هذا المقياس (من جانب التفرطح)كما يلي:

- ✓ إذا كان $F_2 = 0$ فإن منحنى التوزيع يكون طبيعي.
- ✓ إذا كان $F_2 > 0$ فإن منحنى التوزيع يكون مدبدب.
- ✓ إذا كان $F_2 < 0$ فإن منحنى التوزيع يكون متفرطح.

ملاحظة:

يرجع اعتماد معامل فيشر على القيمة 3 في المقياس، لأن العزم المركزي من الدرجة الرابعة يساوي 3 عندما يكون التوزيع طبيعي.

ج) معامل التفرطح باستخدام الانحراف الربيعي (معامل كيللي):

يحسب هذا المقياس باستخدام العلاقة بين الانحراف الربيعي والعشيرات، حيث يحسب رياضيا بالعلاقة التالية:

$$C_K = \frac{Q_3 - Q_1}{2(D_9 - D_1)}$$

حيث D_1: العشير الأول، D_9: العشير التاسع.

ويتم تحديد شكل التوزيع وفق هذا المقياس (من جانب التفرطح)كما يلي:

- ✓ إذا كان $C_K = 0.263$ فإن منحنى التوزيع يكون طبيعي.
- ✓ إذا كان $C_K > 0.263$ فإن منحنى التوزيع يكون مدبدب.
- ✓ إذا كان $C_K < 0.263$ فإن منحنى التوزيع يكون متفرطح.

مثال 3:

البيانات الآتية تمثل التوزيع التكراري لأعداد العمال في 100مؤسسة صغيرة

x_i	n_i	$n_i(x_i - \bar{X})^2$	$n_i(x_i - \bar{X})^4$	N^+
1	15	135	1215	15
2	10	40	160	25
3	5	5	5	30
4	20	0	0	50
5	30	30	30	80
6	20	80	320	100
المجموع	**100**	290	1730	/

المطلوب:

- ✓ حدد معامل التفرطح؟

الحل:

معامل بيرسون للتفرطح P_4:

$$P_4 = \frac{\mu_4}{\mu_2^2} = \frac{\mu_4}{\delta^4} = \frac{17.3}{2.9^2} = 2.05$$

وبما أن $P_4 < 3$ فإن منحنى التوزيع يكون متفرطح (تشتت قوي).

معامل فيشر للتفرطح F_2:

$$F_2 = P_4 - 3 = \frac{\mu_4}{\delta^4} - 3 = 2.05 - 3 = -0.95$$

بما أن $F_2 < 0$ فإن منحنى التوزيع يكون متفرطح.

معامل كيللي:

$$C_K = \frac{Q_3 - Q_1}{2(D_9 - D_1)}$$

حيث D_1: العشير الأول، D_9: العشير التاسع.

حساب Q_1 و Q_3 و D_1 و D_9:

$$CQ_1 = \frac{\sum n_i}{4} = \frac{100}{4} = 25 \quad \Rightarrow Q_1 = 2$$

$$CQ_3 = \frac{3\sum n_i}{4} = \frac{300}{4} = 75 \quad \Rightarrow Q_3 = 5$$

$$CD_1 = \frac{\sum n_i}{10} = \frac{100}{10} = 10 \quad \Rightarrow D_1 = 1$$

$$CD_9 = \frac{9\sum n_i}{10} = \frac{900}{10} = 90 \quad \Rightarrow D_9 = 6$$

$$C_K = \frac{Q_3 - Q_1}{2(D_9 - D_1)} = \frac{5-2}{2(6-1)} = 0.3$$

تمارين محلولة:

التمرين الأول:

عدد زبائن أحد مراكز البريد خلال اليوم كما يلي: 94، 85، 80، 52،75، 80، 68، 60.

المطلوب:

1. أحسب مقاييس النزعة المركزية ثم قارن بينهم؟
2. أحسب العزم البسيط للعزوم الثلاثة الأولى والعزم المركزي للعزوم الأربعة الأولى؟
3. حدد شكل التوزيع من ناحية الالتواء من ناحية التفرطح؟

التمرين الثاني:

ليكن لدينا التوزيع التكراري التالي:

X_i	4	6	8	10	12	14	16
n_i	8	16	12	10	10	8	6

المطلوب:

1. أحسب المتوسط الحسابي، الوسيط، المنوال ثم قارن بينهم؟
2. حدد شكل التوزيع من ناحية الالتواء ؟
3. حدد شكل التوزيع من ناحية التفرطح؟

التمرين الثالث:

البيانات التالية تمثل نفقات استهلاك الطاقة الكهربائية خلال الثلاثي الرابع لسنة 2022 لمجموعة من الأسر بأحد الأحياء السكنية.

المبالغ (10^3 دج)	[5-10[	[10-15[	[15-20[	[20-25[	[25-30[	[30-35[	المجموع
عدد الأسر	8	14	16	10	9	5	62

حيث: $\sum_{i=1}^{k} c_i * n_i = 1150;\ \sum_{i=1}^{k} n_i (c_i - \bar{X})^2 = 3356{,}86;\ \sum_{i=1}^{k} n_i (c_i - \bar{X})^4 = 387580{,}75$

المطلوب:

1. حدد المجتمع الاحصائي، الوحدة الإحصائية، المتغير الاحصائي ونوعه؟
2. أحسب كل من المتوسط الحسابي، الوسيط، المنوال؟
3. أحسب كل من التباين والانحراف المعياري؟
4. حدد شكل التوزيع من ناحية الالتواء والتفرطح؟

حلول التمارين:

حل التمرين الأول:

Xi	x_i^2	x_i^3	$(x_i-\bar{X})^1$	$(x_i-\bar{X})^2$	$(x_i-\bar{X})^3$	$(x_i-\bar{X})^4$
94	8836	830584	19,75	390,06	7703,73	152148,75
85	7225	614125	10,75	115,56	1242,29	13354,69
80	6400	512000	5,75	33,06	190,11	1093,12
75	5625	421875	0,75	0,56	0,42	0,31
52	2704	140608	-22,25	495,06	-11015,14	245086,87
80	6400	512000	5,75	33,06	190,10	1093,12
68	4624	314432	-6,25	39,06	-244,14	1525,87
60	3600	216000	-14,25	203,06	-2893,64	41234,37
المجموع	45414	3561624	0	1309,5	-4826,25	455537,16

حساب مقاييس النزعة المركزية:

المتوسط الحسابي:

$$\bar{x}=\frac{\sum x_i}{N}=\frac{594}{8}=74.25$$

الوسيط:

$$Me=\frac{80+75}{2}=77.5$$

المنوال: Mo=80

المقارنة بينهم:

$\bar{X}<\mathrm{Me}<Mo$

التوزيع ملتوي نحو اليسار

3. **حساب العزم البسيط:**

$$m_r=\frac{\sum_{i=1}^{n} n_i x_i^r}{N}$$

د) **حساب العزم البسيط من الدرجة الأولى:**

$$m_1=\frac{\sum_{i=1}^{n} x_i^1}{N}=\bar{X}=74.25$$

ه) **حساب العزم البسيط من الدرجة الثانية:**

$$m_2=\frac{\sum_{i=1}^{n} x_i^2}{N}=\frac{45414}{8}=5676.75$$

و) **حساب العزم البسيط من الدرجة الثالثة:**

$$m_3 = \frac{\sum_{i=1}^{n} x_i^3}{N} = \frac{3561624}{8} = 445203$$

4. حساب العزم المركزي:

$$\mu_r = \frac{\sum_{i=1}^{n}(x_i - \bar{X})^r}{N}$$

د) حساب العزم المركزي من الدرجة الأولى:

$$\mu_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{X})^1}{N} = \frac{0}{8} = 0$$

ه) حساب العزم المركزي من الدرجة الثانية:

$$\mu_2 = \frac{\sum_{i=1}^{n}(x_i - \bar{X})^2}{N} = \frac{1309.5}{8} = 163.68$$

و) حساب العزم المركزي من الدرجة الثالثة:

$$\mu_3 = \frac{\sum_{i=1}^{n}(x_i - \bar{X})^3}{N} = \frac{-4826.25}{8} = -603.28$$

ز) حساب العزم المركزي من الدرجة الرابعة:

$$\mu_4 = \frac{\sum_{i=1}^{n}(x_i - \bar{X})^4}{N} = \frac{455537.15}{8} = 56942.14$$

3. تحديد شكل التوزيع من ناحية الالتواء:

باستخدام معامل بيرسون الأول P_1 للالتواء نجد:

$$P_1 = \frac{\bar{X} - Mo}{\delta_x}$$

$$\delta_x = \sqrt{{\delta_x}^2} = \sqrt{163.68} = 12.79$$

$$P_1 = \frac{\bar{X} - Mo}{\delta_x} = \frac{74.25 - 80}{12.79} = -0.44$$

بما أن $P_1 < 0$ فإن التوزيع يكون ملتوي نحو اليسار (سالب الالتواء).

- **تحديد شكل التوزيع التفرطح:**

باستخدام معامل بيرسون للتفرطح P_4:

$$P_4 = \frac{\mu_4}{\mu_2^2} = \frac{56942.14}{26759.75} = 2{,}12$$

بما أن $P_4 < 3$ فإن منحنى التوزيع يكون متفرطح (تشتت قوي).

حل التمرين الثاني:

x_i	n_i	$x_i * n_i$	$n_i(x_i - \bar{X})^2$	$n_i(x_i - \bar{X})^3$	$n_i(x_i - \bar{X})^4$
4	8	32	225,56	-1197,77	6360,16
6	16	96	175,29	-580,23	1920,57
8	12	96	20,59	-26,97	35,33
10	10	100	4,76	3,28	2,26
12	10	120	72,36	194,65	523,61
14	8	112	175,96	825,29	3870,62
16	6	96	268,53	1796,50	12018,65
المجموع	70	652	943,08	1014,75	24731,23

المتوسط الحسابي:

$$\bar{x} = \frac{\sum n_i x_i}{\sum n_i} = \frac{652}{70} = 9.31$$

الوسيط:

$$Me = 8$$

المنوال: Mo=6

الانحراف المعياري:

$$\delta_x = \sqrt{{\delta_x}^2} = \sqrt{\frac{943.08}{70}} = 3.67$$

4. تحديد شكل التوزيع من ناحية الالتواء:

باستخدام معامل فيشر للالتواء:

$$F_1 = \frac{\mu_3}{{\delta_x}^3} = \frac{\frac{\sum_{i=1}^{n} n_i(x_i - \bar{x})^3}{\sum_{i=1}^{k} n_i}}{{\delta_x}^3} = 0.29$$

بما أن $F_1 > 0$ فإن التوزيع يكون مائل (ملتوي) نحو اليمين (موجب الالتواء).

- **تحديد شكل التوزيع التفرطح:**

باستخدام معامل بيرسون للتفرطح P_4:

$$P_4 = \frac{\mu_4}{\mu_2^2} = \frac{\frac{\sum_{i=1}^{k} n_i(x_i - \bar{X})^4}{\sum_{i=1}^{k} n_i}}{\delta^4} = \frac{\frac{24731.23}{70}}{3{,}67^4} = 1.94$$

بما أن $P_4 < 3$ فإن منحنى التوزيع يكون متفرطح (تشتت قوي).

حل التمرين الثالث:

1. **المجتمع الاحصائي:** الأسر
الوحدة الإحصائية: أسرة واحدة
المتغير الاحصائي: نفقات استهلاك الطاقة الكهربائية
نوعه: كمي مستمر.

2. **حساب المتوسط الحسابي:**

$$\bar{x} = \frac{\sum_{i=1}^{k} n_i c_i}{\sum_{i=1}^{k} n_i} = \frac{1150}{62} = 18,55$$

➢ **حساب الوسيط:**

-تحديد رتبة الوسيط

$$c = \frac{\sum n_i}{2} = \frac{62}{2} = 31$$

-من التكرار المتجمع الصاعد نجد الفئة الوسيطية هي: [15-20[
لأن التكرار المتجمع الصاعد لها يساوي (8+14+16=38)

$$M_e = d + \frac{C - N_{i-1}^{+}}{n_i} L = 15 + \frac{31 - 22}{16} 5 = 17,81$$

➢ **تحديد قيمة المنوال:**

الفئة المنوال هي الفئة التي تقابل أكبر تكرار: [15-20[

$$M_O = d + \frac{\Delta_{i-1}}{\Delta_{i+1} + \Delta_{i-1}} L = 15 + \frac{(16-14)}{(16-14) + (16-10)} * 5 = 16,25$$

3. **حساب التباين:**

$$V_x = \frac{\sum_{i=1}^{k} n_i (c_i - \bar{x})^2}{\sum_{i=1}^{k} n_i} = \frac{3356.86}{62} = 54,14$$

➢ **حساب الانحراف المعياري:**

$$\delta_X = \sqrt{V_X} = \sqrt{54,14} = 7,36$$

4. **تحديد شكل التوزيع من ناحية الالتواء:**

باستخدام معامل بيرسون الأول P_1 للالتواء نجد:

$$P_1 = \frac{\bar{X} - Mo}{\delta_x} = \frac{18,55 - 16,25}{7,36} = 0.31$$

بما أن $P_1 > 0$ فإن التوزيع يكون مائل (ملتوي) نحو اليمين (موجب الالتواء).

➢ **تحديد شكل التوزيع التفرطح:**

باستخدام معامل بيرسون للتفرطح P_4:

$$P_4 = \frac{\mu_4}{\mu_2^2} = \frac{\frac{\sum_{i=1}^{k} n_i (c_i - \bar{X})^4}{\sum_{i=1}^{k} n_i}}{\delta^4} = \frac{\frac{387580{,}75}{62}}{7{,}36^4} = \frac{6251{,}30}{2934{,}35} = 2{,}13$$

بما أن $P_4 < 3$ فإن منحنى التوزيع يكون متفرطح (تشتت قوي).

الفصل السادس: منحنى لورنز ومعامل جيني (Lorenz Curve and Gini Coefficient)

1) منحنى لورنز:

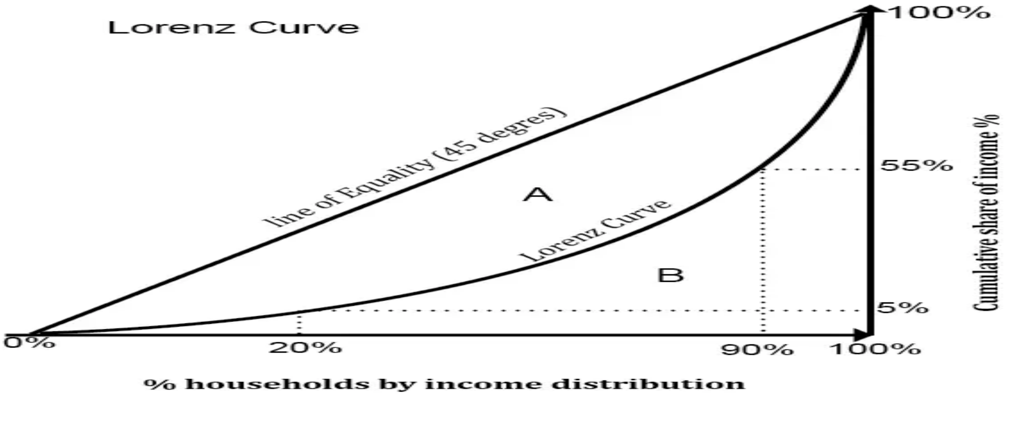
منحنى لورنز (Lorenz Curve): قدمه الخبير الاقتصادي ماكيس لورنز في ورقة بحثية نشرها عام 1905، وهو تمثيل بياني لتوزيع الثروة أو الدخل، يبين النسبة المئوية لإجمالي الدخل الذي تجنيه شرائح مختلفة من السكان بترتيب السكان حسب حجم دخلهم.

يقيس منحنى لورنز مايطلق عليه التوزيع التراكمي (cumulative distribution) لمتغير ما، ومدى ابتعاد هذا المتغير عن معيار مطلق وهو التوزيع المتساوي (Equal Distribution) للمتغير تحت الدراسة، كالدخل أو الثروة، أو أي متغير عن ملكية الأفراد، أو الناتج المحلي الإجمالي للدول. وفي معظم الدراسات يقيس هذا المنحنى مدى التفاوت في توزيع الدخل أو الثروة في المجتمع أو لمجموعة من الناس.

يرتكز منحنى لورنز على محورين، تمثل بموجبهما النسب المئوية للسكان بناء على الثروة أو الدخل على المحور الأفقي، وتمثل الدخل التراكمي على المحور العمودي، ويفيد استخدامه إلى جانب توضيح اللامساواة في الدخل، تمثيل التوزيع المتفاوت في أي نظام، وعليه كلما ابتعد الخط القطري في المنحنى عن نقطة البداية زاد حجم التفاوت.

بصفة عامة فإن منحنى لورنز يستخدم لتوضيح العلاقة بين متغيرين الأول منتغير مستقل ويمثل في المحور الأفقي، والثاني متغير تابع ويمثل في المحور العمودي. أي أن منحنى لورنز يستخدم لتوضيح سوء عدالة /سوء توزيع المتغير المستقل على المتغير التابع.

رسم تخطيطي لمنحنى لورنز



في منحنى لورنز هذا والذي يمثل توزيع السكان حسب الدخل المكتسب، حيث نلاحظ أن أفقر 20٪ من الأسر تمتلك 5٪ من إجمالي الدخل، بينما يمتلك 90٪ من السكان يمتلكون 55٪ من إجمالي الدخل. وهذا يعني أن أغنى 10٪ من أصحاب الدخل يكسبون 45٪ من إجمالي الدخل.

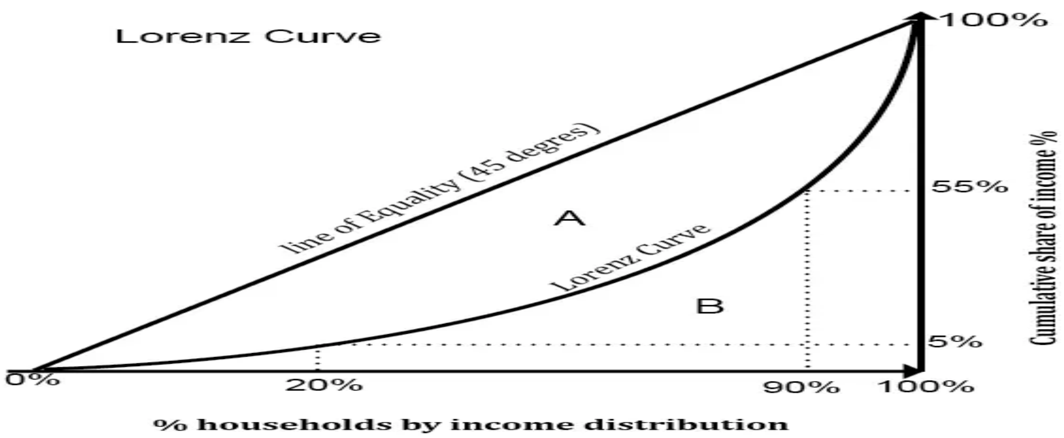
2) معامل جيني:

معامل جيني في علم الاقتصاد، يعد المعامل الجيني، الذي يطلق عليه أحيانًا مؤشر جيني مقياسا للتشتت الإحصائي الذي يهدف إلى تمثيل عدم المساواة في الدخل أو عدم المساواة في الثروة داخل دولة أو أي مجموعة أخرى من الناس.

معامل جيني (G) هو مقياس لعمق التفاوت في التوزيع، تنحصر قيمته بين الصفر والواحد الصحيح، فإذا كانت قيمته صفرا فإن ذلك يدل على أن كل الفئات (أشخاص، أسر، مجموعات، دول) متساوية بحصتها من الشيء المراد معرفة توزيعه، كالدخل أو الثروة. وإذا كانت قيمته 1 فإن ذلك يدل على أن شخصا واحدا (أو أسرة واحدة، أو دولة واحدة) استأثر بكل الشيء ولم يترك أي جزء منه للبقية. يأخذ المعامل (G) أية قيمة بين 0 و1.

3) منحنى لورنز ومعامل جيني:

يمكن استخدام منحنى لورنز لحساب معامل جيني - وهو مقياس آخر لعدم المساواة.



- كلما اقترب منحنى لورنز من خط المساواة ، كانت المنطقة A أصغر. وسيكون معامل جيني منخفضًا.
- إذا كانت هناك درجة عالية من عدم المساواة ، فإن المنطقة A ستكون نسبة مئوية أكبر من إجمالي المساحة.
- يُظهر ارتفاع معامل جيني ارتفاعًا في عدم المساواة - فهو يظهر أن منحنى لورنز بعيدا عن خط المساواة.

معامل جيني هو المنطقة A+ B / A

مثال1:

لنفترض وجود البيانات الآتية عن 10 أشخاص والتي تمثل ثروة لكل شخص منهم.

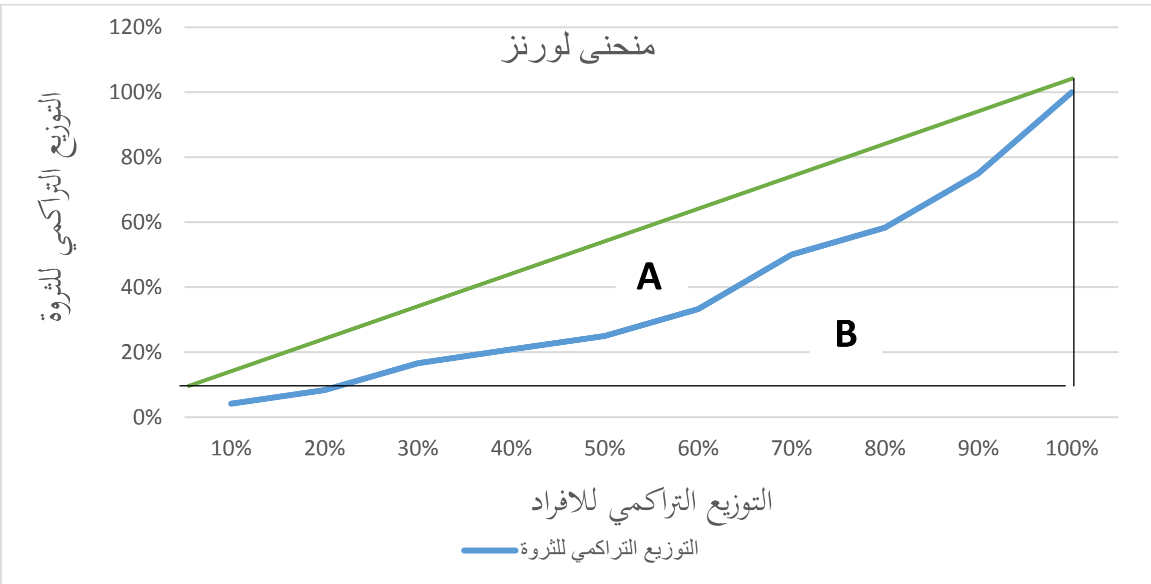
الشخص	1	2	3	4	5	6	7	8	9	10	المجموع
الثروة	25	25	50	25	25	50	100	50	100	150	600

لرسم منحنى لورنز يجب حساب أولا التوزيع التراكمي للأفراد والتوزيع التراكمي للثروة أي التكرار المتجمع الصاعد المئوي للأفراد والثروة، ونتحصل على الجدول التالي:

الشخص	1	2	3	4	5	6	7	8	9	10	المجموع
الثروة	25	25	50	25	25	50	100	50	100	150	600
التوزيع التراكمي للأفراد	10%	%20	%30	%40	%50	%60	%70	%80	%90	%100	
التوزيع التراكمي للثروة	%4.17	%8.33	%16.67	%20.83	%25	%33.33	%50	%58.33	%75	%100	

الشكل البياني لمنحنى لورنز:

بعد حساب التوزيع التراكمي للمتغيرين، نمثلهما على معلم متعامد متجانس فنتحصل على الشكل البياني التالي والذي يمثل منحنى لورنز:



الخط الأخضر هو الذي يمثل التوزيع التراكمي المتساوي للثروة ويسمى بخط التماثل أو التوزيع المتماثل.

منحنى لورنز هو المنحنى الأزرق يسمى بمنحنى التوزيع الفعلي، **A** تمثل منطقة عدم التماثل بُعد منحنى لورنز عن خط التساوي أي كلما اقترب منحنى لورنز من الخط الأخضر كلما كان التوزيع متساويا وكلما ابتعد عن هذا الخط كلما ازداد عدم اللامساواة.

4) قيمة معامل جيني (G) :

$$0 \leq G \leq 1$$

$$G = \frac{A}{A+B}$$

بما أن خط التوازي يقسم الشكل البياني إلى نصفين متساويين فإن:

$$A + B = 0.5$$

$$G = 2A = 1 - 2B$$

إذا تم تمثيل منحنى لورنز بالمعادلة Y = F (X) ، يمكن العثور على قيمة **B** بالتكامل

$$G = 1 - 2\int_0^1 F(x)dx$$

في بعض الأحيان يكون منحنى لورنز غير معروف بالكامل، ويتم إعطاء القيم فقط على فترات زمنية معينة. في هذه الحالة ، يمكن تقريب معامل جيني باستخدام تقنيات مختلفة لاستيفاء القيم المفقودة لمنحنى لورنز. إذا كانت **(**X_k ،Y_k**)** هي النقاط المعروفة على منحنى Lorenz ، مع العلم أن **(**X_{k-1} > X_k**)**، بحيث:

X_k:هي النسبة التراكمية لمتغير السكان

Y_k:هي النسبة المتراكمة لمتغير الدخل

إذا تم تقريب منحنى لورنز في كل فترة زمنية كخط بين نقاط متتالية، فيمكن تقريب المنطقة **B** باستخدام المعادلة التالية:

$$G = 1 - \sum_{k=1}^{n} (X_k - X_{k-1})\,(Y_k - Y_{k-1})$$

التقريب الناتج لـ G. يمكن الحصول على نتائج أكثر دقة باستخدام طرق أخرى لتقريب المنطقة B ، مثل تقريب منحنى لورنز بوظيفة تربيعية عبر أزواج من الفواصل الزمنية، أو بناء تقريب سلس مناسب لوظيفة التوزيع الأساسية التي تطابق البيانات المعروفة. إذا كان متوسط عدد السكان وقيم الحدود لكل فاصل زمني معروفين أيضًا، فيمكن أيضًا استخدامها في كثير من الأحيان لتحسين دقة التقريب.

تمارين محلولة:

التمرين الأول:

الجدول التالي يمثل أجور عمال المؤسسة

الأجر الوحدوي	1124	1561	1969	2149	2257	2365	2473	2977	3559	4060
عدد العمال	101	83	54	49	33	29	38	16	10	5

المطلوب:

1. أرسم منحنى لورنز، واشرحه؟
2. إذا افترضنا أن F(x) دالة منحنى لورنز معرفة على المجال [0-1] من الشكل:

$$F(x) = 1.5x^4 - 2x^3 + 1.4x^2 + 0.1x$$

-حدد قيمة معامل جيني؟

التمرين الثاني:

الجدول التالي يمثل أجور عمال احدى المؤسسات

الأجور (10^3 دج)	[25-40[	[40-55[	[55-75]
عدد العمال	50	130	20

المطلوب:

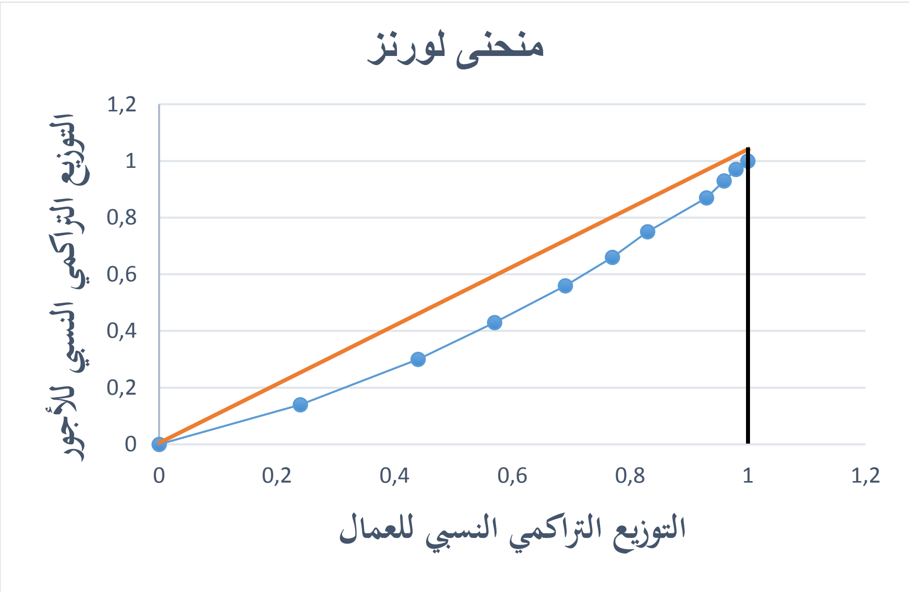
1. أرسم منحنى لورنز، واشرحه؟
2. حدد قيمة معامل جيني انطلاقا من منحنى لورنز وفسره؟

حلول التمارين:

حل التمرين الأول:

الأجر	عدد العمال	التوزيع التراكمي للعمال	التوزيع التراكمي النسبي للعمال	كتلة الأجور	الأجر النسبي	التوزيع التراكمي النسبي للأجور
1124	101	101	0.24	113524	0.14	0.14
1561	83	184	0.44	129563	0.16	0.30
1969	54	238	0.57	106326	0.13	0.43
2149	49	287	0.69	105301	0.13	0.56
2257	33	320	0.77	74481	0.10	0.66
2365	29	349	0.83	68585	0.09	0.75
2473	38	387	0.93	93974	0.12	0.87
2977	16	403	0.96	47632	0.06	0.93
3559	10	413	0.98	35590	0.04	0.97
4060	5	418	1	20300	0.03	1
المجموع	**418**	/	/	**795276**	**1**	/

1. رسم منحنى لورنز:



بما أن منحنى لورنز قريب من خط التماثل إذن هناك نوع من العدالة في توزيع الأجور على العمال في هذه المؤسسة.

2. **تحديد قيمة معامل جيني:**

إذا افترضنا أن F(x) دالة منحنى لورنز معرفة على المجال [0-1] من الشكل:

$$F(x) = 1.5x^4 - 2x^3 + 1.4x^2 + 0.1x$$

قيمة معامل جيني:

$$G = 1 - 2\int_0^1 F(x)dx$$

$$= 1 - 2\int_0^1 (1.5x^4 - 2x^3 + 1.4x^2 + 0.1x)dx$$

$$= 1 - 2\left[\frac{1.5}{5}x^5 - \frac{2}{4}x^4 + \frac{1.4}{3}x^3 + \frac{0.1}{2}x^2\right]_0^1$$

$$= 1 - 2[(0.3 - 0.5 + 0.46 + 0.05) - 0]$$

$$= 1 - 2(0.31)$$

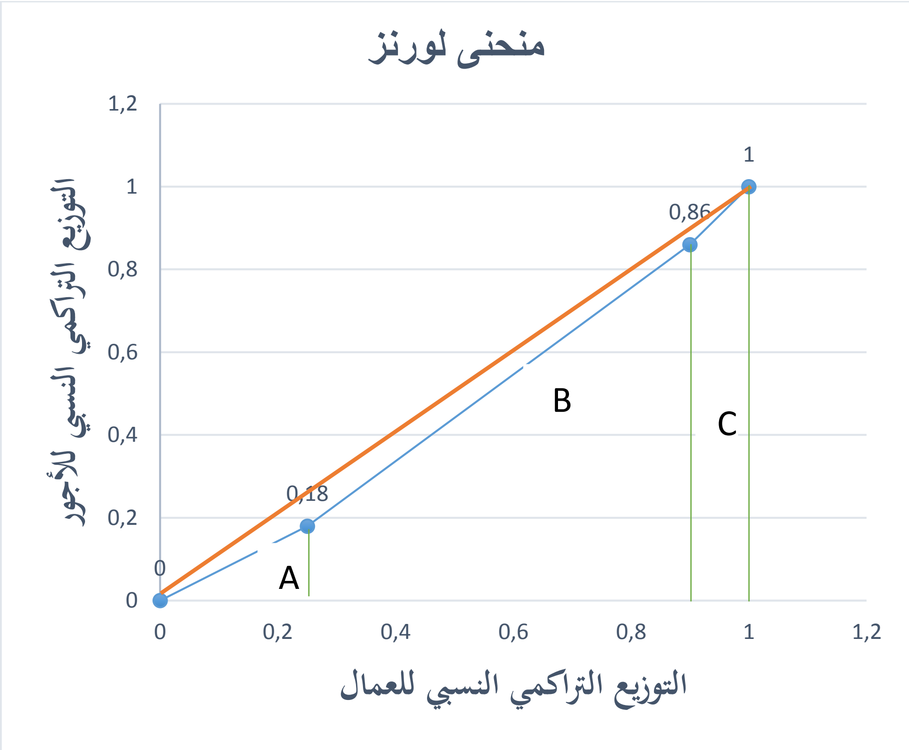
$$G = 0.38$$

حسب قيمة معامل جيني فأن هناك نوع من العدالة في توزيع الأجور داخل هذه المؤسسة.

حل التمرين الثاني:

الأجر الفردي	**عدد العمال**	**التوزيع النسبي للعمال**	**التوزيع النسبي التراكمي للعمال**	***مركز الفئات للأجور***	***الأجور***	**التوزيع النسبي للأجور**	**التوزيع النسبي التراكمي للأجور**
]40-25]	50	0,25	0,25	32,5	1625	0,18	0,18
]55-40]	130	0,65	0,9	47,5	6175	0,68	0,86
]75-55]	20	0,1	1	65	1300	0,14	1
المجموع	**200**	**1**	/	/	9100	1	/

1. رسم منحنى لورنز:



منحنى لورنز يدرس مدى عدالة توزيع الأجور في المؤسسة ونلاحظ أن هذا المنحنى (الخط الأزرق) يقترب من خط التماثل (الخط الأحمر) إذن هناك عدالة في توزيع الأجور داخل هذه المؤسسة. و هذا ما سيبينه معامل جيني.

2. تحديد قيمة معامل جيني انطلاقا من منحنى لورنز:

لحساب المساحة تحت منحنى لورنز تم تقسيمها إلى ثلاثة أشكال هندسية A.B.C حيث:

$$G = 1 - 2(A + B + C)$$

- **مساحة المثلث A:**

$$A = \frac{\text{القاعدة} \times \text{الارتفاع}}{2} = \frac{0.18 \times 0.25}{2} = 0.0225$$

- **مساحة الرباعي B:**

$$B = \frac{\text{الارتفاع} \times (\text{القاعدة الصغرى} + \text{القاعدة الكبرى})}{2} = \frac{(0.18 + 0.86) \times 0.65}{2} = 0.338$$

➢ **مساحة الرباعي** C:

$$C=\frac{(\text{القاعدة الصغرى}+\text{القاعدة الكبرى})\times\text{الارتفاع}}{2}=\frac{(1+0.86)\times 0.1}{2}=0.093$$

➢ **معامل جيني:**

$$G=1-2(A+B+C)=1-2(0.0225+0.338+0.093)=0.093$$

بما أن معامل جيني قريب من الصفر إذن توجد عدالة في توزيع الأجور داخل هذه المؤسسة.

الفصل السابع: الأرقام القياسية

الأرقام القياسية هي عبارة عن مؤشرات إحصائية، تستعمل لقياس تطور سعر أو كمية مادة أو عدة مواد بين فترتين زمنيتين مختلفتين أو بين مجموعتين أو بين مكانين.

تشكل الأرقام القياسية (للأسعار، للأجور، للعملة، ...) مؤشرات هامة في سيرورة الاقتصاديات المتطورة، لما لها من آثار بليغة على النشاطات الاقتصادية.

1) تعريف:

الرقم القياسي هو أداة لقياس التغير النسبي الحاصل في قيم أية ظاهرة أو مجموعة من الظواهر، من ظرف أول يسمى بظرف الأساس إلى ظرف آخر يسمى بظرف المقارنة، سواء كان الظرف زمانيا أو مكانيا، وتكون قيمة الرقم القياسي في ظرف الأساس مساوية دائما للمقدار 100.

والأرقام القياسية كثيرة الاستخدام في دراسة تطور الظواهر الاقتصادية والاجتماعية، كالاستهلاك، الإنتاج، الصادرات، الواردات، ...إلخ، غير أنها أكثر استخداما في دراسة تطور الأسعار والنفقات الاستهلاكية.

2) أنواع الأرقام القياسية:

أ) الرقم القياسي البسيط:

يقيس تطور سعر أو كمية مادة واحدة فقط بين فترتين زمنيتين مختلفتين أو بين مكانيين مختلفين، وهو عبارة عن النسبة بين سعر أو كمية الفترة أو السنة الحالية أو المدروسة، وسعر أو كمية فترة أو سنة الأساس، يرمز للسنة الحالية أو المدروسة بالرمز t ويرمز لسنة أو فترة الأساس بالرمز 0.

تكتب العلاقة الإحصائية للرقم القياسي البسيط بالشكل التالي:

العلاقة الإحصائية للرقم القياسي البسيط للأسعار:

$$I_{(t|0)}(P) = \frac{P_t}{P_0} \times 100$$

العلاقة الإحصائية للرقم القياسي البسيط للكميات:

$$I_{(t|0)}(Q) = \frac{Q_t}{Q_0} \times 100$$

حيث:

$I_{t/0}(P)$: الرقم القياسي البسيط للأسعار، P_t السعر للفترة t وP_0 السعر للفترة 0والتي تدعى بسنة الأساس.

$I_{t/0}(Q)$:الرقم القياسي البسيط للكميات، Q_t الكمية للفترة t وQ_0 الكمية للفترة 0وهي سنة الأساس.

ويمكن أن نميز بين ثلاثة حالات لقيم الرقم القياسي البسيط:

الحالة الأولى: ثبات تطور السعر أو الكمية، ففي هذه الحالة الرقم القياسي يساوي 100.

الحالة الثانية: انخفاض في السعر أو الكمية، ففي هذه الحالة قيمة الرقم القياسي تكون أقل من 100.

الحالة الثالثة: زيادة في السعر أو الكمية، ففي هذه الحالة قيمة الرقم القياسي تكون أكبر من 100.

ملاحظة: الزيادة أو الارتفاع هو عبارة عن الفرق لبن القيمة المتحصل عليها و100، أما الانخفاض فهو عبارة عن 100 ناقص النتيجة المتحصل عليها.

مثال01 :

في 2005/11/11 كان سعر صرف الأورو يساوي 83.06دج، وفي2022/11/19 أصبح سعر الأورو يساوي 143.86دج. أوجد الرقم القياسي لتطور سعر صرف الأورو بالنسبة للدينار الجزائري، (2005 هي سنة الأساس)؟

الحل:

$$I_{(t|0)}(P) = \frac{P_t}{P_0} \times 100 = \frac{143.86}{83.06} \times 100 = 173.2\%$$

يعني هذا أنه إذا كان سعر صرف الأورو مقابل الدينار الجزائري يساوي 100دج بأسعار سنة 2005 فإنه ارتفع ليصبح 173.2دينار بأسعار سنة 2022.

تستخدم هذه الطريقة في دراسة تطور قيمة ظاهرة واحدة فقط، غير أنه عمليا، وفي كثير من الأحيان يتطلب الأمر، دراسة تطور أسعار عدة مواد، في مثل هذه الحالة غالبا ما يتم استخدام الوسط الحسابي أو الوسط الهندسي كمتوسط الأرقام القياسية.

1. طريقة الوسط الحسابي:

إذا كانت لدينا $P_{0,1}$,$P_{0,2}$,….,$P_{0,n}$ أسعار فترة الأساس للمواد N,….,1,2 على التوالي، و $P_{1,1}$,$P_{1,2}$,….,$P_{1,n}$ أسعار فترة المقارنة لنفس المواد، حيث N: عدد المواد.

فالوسط الحسابي للأرقام القياسية يعرف بالمعادلة التالية:

$$I = \frac{\sum_{i=1}^{n}\left[\frac{P_{1,i}}{P_{0,i}}\right]}{N} \times 100$$

2. طريقة الوسط الهندسي:

إذا كانت لدينا $P_{0,1}$, $P_{0,2}$,….,$P_{0,n}$ أسعار فترة الأساس للمواد 1,2,….,N على التوالي، و $P_{1,1}$,$P_{1,2}$,….,$P_{1,n}$ أسعار فترة المقارنة لنفس المواد، حيث N: عدد المواد.

فالوسط الهندسي للأرقام القياسية يعرف بالمعادلة التالية:

$$I = \sqrt[n]{\prod_{i=1}^{n}\left[\frac{P_{1,i}}{P_{0,i}}\right]} \times 100$$

مثال 02:

أوجد الرقم القياسي للأسعار بطريقتي الوسط الحسابي والوسط الهندسي.

المادة / السعر	A	B	C	D
P_0	80	70	50	30
P_1	88	90	63	37

الحل:

لإيجاد الوسط الحسابي والوسط الهندسي للأرقام القياسية، يتطلب أولا إيجاد الأرقام القياسية البسيطة

$$I_{(1|0)}(A) = \frac{P_{1A}}{P_{0A}} \times 100 = \frac{88}{80} \times 100 = 110\%$$

$$I_{(1|0)}(B) = \frac{P_{1B}}{P_{0B}} \times 100 = \frac{90}{70} \times 100 = 128.57\%$$

$$I_{(1|0)}(C) = \frac{P_{1C}}{P_{0C}} \times 100 = \frac{63}{50} \times 100 = 126\%$$

$$I_{(1|0)}(D) = \frac{P_{1D}}{P_{0D}} \times 100 = \frac{37}{30} \times 100 = 123.33\%$$

الوسط الحسابي للأرقام القياسية:

$$I = \frac{\sum_{i=1}^{n}\left[\frac{P_{1,i}}{P_{0,i}}\right]}{N} \times 100$$

$$I = \frac{110 + 128.57 + 126 + 123.33}{4}$$

$$I = 121.97\%$$

الوسط الهندسي للأرقام القياسية:

$$I = \sqrt[n]{\prod_{i=1}^{n}\left[\frac{P_{1,i}}{P_{0,i}}\right]} \times 100$$

$$I = \sqrt[4]{(1.1 \times 1.2857 \times 1.26 \times 1.2333} \times 100$$

$$I = \sqrt[4]{2.1977} \times 100 = 121.75\%$$

ب) الرقم القياسي التجميعي:

وهو عبارة عن النسبة بين أسعار أو كميات مجموعة من المواد في السنة المدروسة ومجموع أسعار أو كميات هذه المواد في سنة الأساس للفترة (0).

تعطى العلاقة الإحصائية للرقم القياسي التجميعي بالشكل التالي:

$$I_{(t|0)}(P_j) = \frac{\sum_j P_{tj}}{\sum_j P_{0j}} \times 100$$

$$I_{(t|0)}(Q_j) = \frac{\sum_j Q_{tj}}{\sum_j Q_{0j}} \times 100$$

حيث:

$\sum_j P_{tj}$: يمثل مجموع أسعار سلع مختلفة للفترة t، $\sum_j Q_{tj}$مجموع الكميات لسلع مختلفة للفترة t.

$\sum_j P_{0j}$: يمثل مجموع أسعار سلع مختلفة في سنة الأساس، $\sum_j Q_{0j}$مجموع الكميات لسلع مختلفة في سنة الأساس.

مثال 03:

يمثل الجدول التالي أسعار 4 مواد من الوقود في الجزائر خلال 5 سنوات.

مازوت ديزل (ل/دج)	بنزين عادي(ل/دج)	بنزين بدون رصاص(ل/دج)	بنزين ممتاز(ل/دج)	
13.77	21.2	22.6	23	2010
18.76	28.45	31.02	31.42	2016
20.42	32.69	35.33	35.72	2017
23.06	38.95	41.62	41.97	2018
29.01	43.71	45.62	45.97	2020

المطلوب: تحديد الرقم القياسي التجميعي علما ان سنة الأساس هي سنة 2010؟

الحل:

$$I_{(16|10)} = \frac{\sum_j P_{tj}}{\sum_j P_{0j}} \times 100 = \frac{31.42 + 31.02 + 28.45 + 18.76}{23 + 22.6 + 21.2 + 13.77} = \times 100 = 136.09\%$$

$$I_{(17|10)} = \frac{35.72 + 35.33 + 32.69 + 20.42}{23 + 22.6 + 21.2 + 13.77} = \times 100 = 154.10\%$$

$$I_{(18|10)} = \frac{41.97 + 41.62 + 38.95 + 23.06}{23 + 22.6 + 21.2 + 13.77} = \times 100 = 180.71\%$$

$$I_{(20|10)} = \frac{45.97 + 45.62 + 43.71 + 29.01}{23 + 22.6 + 21.2 + 13.77} = \times 100 = 203.93\%$$

انطلاقا من هذه النتائج نلاحظ أن أسعار الوقود ارتفعت بمقدار 36.09% من سنة 2010 إلى سنة 2016، ثم تواصل هذا الارتفاع حتى وصل إلى 103.93% في سنة 2020.

ج) الأرقام القياسية المرجحة:

في هذه الطريقة يتم أخذ الكميات المستهلكة من كل مادة كأوزان، بحيث ننتقل من الرقم القياسي التجميعي البسيط للأسعار إلى الرقم القياسي التجميعي للنفقات على المواد، ومن أهم الأرقام القياسية المرجحة:

1. **الرقم القياسي لـــــ لاسبير (Laspeyres):**

ويعتبر من الأرقام القياسية المرجحة، حيث يتم حسابه بقسمة مجموع مرجح للفترة t على مجموع مرجح لسنة الأساس، والترجيح يكون بالنسبة لسنة الأساس. وهما نوعان:

الرقم القياسي لاسبير للأسعار: وهو عبارة عن النسبة بين الكتلة النقدية المدفوعة في الفترة t، والكتلة النقدية المدفوعة في سنة الأساس لاقتناء نفس كمية سنة الأساس. وتعطى بالعلاقة التالية:

$$I_{LAS}(P) = \frac{\sum_j P_{tj}\, Q_{0j}}{\sum_j P_{0j} Q_{0j}} \times 100$$

الرقم القياسي لاسبير للكميات: وهو عبارة عن النسبة بين الكمية الكلية للفترة t، والكمية الكلية لسنة الأساس حسب سعر الأساس. وتعطى بالعلاقة التالية:

$$I_{LAS}(Q) = \frac{\sum_j Q_{tj} P_{0j}}{\sum_j Q_{0j} P_{0j}} \times 100$$

2. **الرقم القياسي ل باش (Passche):**

وهو نفسه لاسبير فقط الاختلاف يكمن في أن باش يرجح بالنسبة للفترة t، وتعطى بالعلاقة التالية:

$$I_{PAS}(P) = \frac{\sum_j P_{tj}\, Q_{tj}}{\sum_j P_{0j} Q_{tj}} \times 100$$

$$I_{PAS}(Q) = \frac{\sum_j Q_{tj} P_{tj}}{\sum_j Q_{0j} P_{tj}} \times 100$$

3. **الرقم القياسي ل فيشر (Fisher):**

يعتبر من أفضل الأرقام القياسية، حيث يعتمد على الرقمين القياسيين السابقين للاسبير وباش، وتعطى علاقته كالتالي:

$$I_{FP}(P) = \sqrt{I_{LAS}(P) \times I_{PAS}(P)}$$

$$I_{FP}(Q) = \sqrt{I_{LAS}(Q) \times I_{PAS}(Q)}$$

مثال 04:

يبين الجدول التالي أسعار وكميات 4 مواد خلال فترتين، المطلوب حساب الأرقام القياسية للأسعار لكل من لاسبير وباش وفيشر.

المواد	P_0	Q_0	P_1	Q_1
A	80	5	88	9
B	70	6	90	8
C	50	9	63	14
D	30	3	37	7

الحل:

الرقم القياسي للاسبير:

$$I_{LAS}(P) = \frac{\sum_j P_{tj}Q_{0j}}{\sum_j P_{0j}Q_{0j}} \times 100 = \frac{88\times5+90\times6+63\times9+37\times3}{80\times5+70\times6+50\times9+30\times3} \times 100$$

$$I_{LAS}(P) = 121.91\%$$

الرقم القياسي لباش (Passche):

$$I_{PAS}(P) = \frac{\sum_j P_{tj}Q_{tj}}{\sum_j P_{0j}Q_{tj}} \times 100 = \frac{88\times9+90\times8+63\times14+37\times7}{80\times9+70\times8+50\times14+30\times7} \times 100$$

$$I_{PAS}(P) = 121.14\%$$

الرقم القياسي ل فيشر (Fisher):

$$I_{FP} = \sqrt{I_{LAS}(P) \times I_{PAS}(P)} = \sqrt{121.91 \times 121.14}$$

$$I_{FP} = 121.52\%$$

ملاحظة: يتميز الرقم القياسي ل لاسبير بثبات قاعدة المقارنة مما يسهل في العمليات الحسابية، غير أن بعد سنة الأساس عن السنة المدروسة يعطي صورة غير حقيقية عن الظاهرة المدروسة.

يصبح ترجيح المواد عائقا ومشكلة لقاعدة لاسبير عند تطور أهمية المواد بسرعة في الزمن، وفي هذه الحالة يجب تغيير سنة الأساس حتى تتماشى مع المعطيات الجديدة.

أما الرقم القياسي لباش يتغير نظام ترجيحه باستمرار، حيث يبين في نفس الوقت تطور ظاهرتين: أثر تغير الأسعار وأثر تغير بنية الاستهلاك. ومن هنا يصبح اجراء دراسات ميدانية مستمرة ضرورة ملزمة لكل باحث يريد الحصول على نتائج تعبر عن الواقع فعلا.

3) خصائص الأرقام القياسية:

أ) **خاصية المطابقة:** يمكن التعبير عن ذلك بواسطة العلاقات التالية:

$$I_{\%} = \frac{P_0}{P_0} \times 100 = 100$$

أو

$$I_{\mathrm{t/t}} = \frac{P_t}{P_t} \times 100 = 100$$

ب) **خاصية الانعكاس:** يمكن التعبير عن ذلك بواسطة العلاقة التالية:

$$I_{\mathrm{t/0}} \times I_{0/t} = 100^2$$

ج) **خاصية الدوران:** يمكن التعبير عن ذلك بواسطة العلاقة التالية:

$$I_{\mathrm{a/b}} \times I_{\mathrm{b/c}} \times I_{\mathrm{c/a}} = 100^3$$

د) **خاصية التحويل:** يمكن التعبير عن ذلك بواسطة العلاقة التالية:

$$I_{\mathrm{c/b}} = \frac{I_{\mathrm{c/a}}}{I_{\mathrm{b/a}}} \times 100$$

تمارين محلولة:

التمرين الأول: الجدول التالي يمثل انتاج القمح للجزائر للفترة (2010-2016)

السنوات	2010	2011	2012	2013	2014	2015	2016
انتاج القمح(مليون طن)	2.9	2.8	3.4	3.6	1.9	2.6	1.9

المطلوب: تحديد الرقم القياسي البسيط لإنتاج القمح في الجزائر، علما أن سنة الأساس هي سنة 2010.

التمرين الثاني:

يمثل الجدول التالي أسعار وكميات 4 مواد خلال سنتي 2015 و2020

السلع	سنة2015		سنة 2020	
	السعر	الكمية	السعر	الكمية
A	5	150	10	100
B	8	200	8	220
C	6	80	15	100
D	7	60	21	90

المطلوب:

باعتبار 2015 سنة الأساس أحسب كل من:

1. الوسط الحسابي والهندسي للأرقام القياسية.
2. الرقم القياسي التجميعي البسيط للأسعار.
3. الرقم القياسي المرجح للأسعار والكميات لاسبير.
4. الرقم القياسي المرجح للأسعار والكميات لباش .
5. الرقم القياسي الأمثل للأسعار والكميات لفيشر.

حلول التمارين:

حل التمرين الأول:

الأرقام القياسية لانتاج القمح:

$$I_{(2011|2010)}(Q) = \frac{Q_{2011}}{Q_{2010}} \times 100 = \frac{2.8}{2.9} \times 100 = 96.55\%$$

$$I_{(2012|2010)}(Q) = \frac{Q_{2012}}{Q_{2010}} \times 100 = \frac{3.4}{2.9} \times 100 = 117.24\%$$

$$I_{(2013|2010)}(Q) = \frac{Q_{2013}}{Q_{2010}} \times 100 = \frac{3.6}{2.9} \times 100 = 124.13\%$$

$$I_{(2014|2010)}(Q) = \frac{Q_{2014}}{Q_{2010}} \times 100 = \frac{1.9}{2.9} \times 100 = 65.51\%$$

$$I_{(2015|2010)}(Q) = \frac{Q_{2015}}{Q_{2010}} \times 100 = \frac{2.6}{2.9} \times 100 = 89.65\%$$

$$I_{(2016|2010)}(Q) = \frac{Q_{2016}}{Q_{2010}} \times 100 = \frac{1.9}{2.9} \times 100 = 65.51\%$$

حل التمرين الثاني:

1. **الوسط الحسابي للأرقام القياسية:**

$$I = \frac{\sum_{i=1}^{n}\left[\frac{P_{20,i}}{P_{15,i}}\right]}{N} \times 100 = \frac{\frac{10}{5} + \frac{8}{8} + \frac{15}{6} + \frac{21}{7}}{4} \times 100 = \frac{2 + 1 + 2.5 + 3}{4} \times 100$$
$$= 212.5\%$$

2. **الوسط الهندسي للأرقام القياسية:**

$$I = \sqrt[n]{\prod_{i=1}^{n}\left[\frac{P_{20,i}}{P_{15,i}}\right]} \times 100 = \sqrt[4]{2 \times 1 \times 2.5 \times 3} \times 100 = \sqrt[4]{15} \times 100$$
$$= 196.79\%$$

3. **الرقم القياسي التجميعي البسيط للأسعار:**

$$I_{(20|15)} = \frac{\sum_j P_{20i}}{\sum_j P_{15i}} \times 100 = \frac{10 + 8 + 15 + 21}{5 + 8 + 6 + 7} \times 100 = \frac{54}{26} \times 100$$
$$= 207.69\%$$

4. الرقم القياسي المرجح للأسعار والكميات لاسبير:

$$I_{LAS}(P) = \frac{\sum_j P_{20,i}\, Q_{15,i}}{\sum_j P_{15,i} Q_{15,i}} \times 100$$
$$= \frac{10 \times 150 + 8 \times 200 + 15 \times 80 + 21 \times 60}{5 \times 150 + 8 \times 200 + 6 \times 80 + 7 \times 60} \times 100 = 171.07\%$$

$$I_{LAS}(Q) = \frac{\sum_j Q_{20,i} P_{15,i}}{\sum_j Q_{15,i} P_{15,i}} \times 100$$
$$= \frac{100 \times 5 + 220 \times 8 + 100 \times 6 + 90 \times 7}{150 \times 5 + 200 \times 8 + 80 \times 6 + 60 \times 7} \times 100 = 107.38\%$$

5. الرقم القياسي المرجح للأسعار والكميات لباش:

$$I_{PAS}(P) = \frac{\sum_j P_{20,i}\, Q_{20,i}}{\sum_j P_{15,i} Q_{20,i}} \times 100$$
$$= \frac{10 \times 100 + 8 \times 220 + 15 \times 100 + 21 \times 90}{5 \times 100 + 8 \times 220 + 6 \times 100 + 7 \times 90} \times 100 = 176.21\%$$

$$I_{PAS}(Q) = \frac{\sum_j Q_{20,i} P_{20,i}}{\sum_j Q_{15,i} P_{20,i}} \times 100$$
$$= \frac{100 \times 10 + 220 \times 8 + 100 \times 15 + 90 \times 21}{150 \times 10 + 200 \times 8 + 80 \times 15 + 60 \times 21} \times 100 = 110.61\%$$

6. الرقم القياسي المرجح للأسعار والكميات لفيشر:

$$I_{FP} = \sqrt{I_{LAS}(P) \times I_{PAS}(P)} = \sqrt{171.07 \times 176.21} = 173.62\%$$
$$I_{FP} = \sqrt{I_{LAS}(Q) \times I_{PAS}(Q)} = \sqrt{107.38 \times 110.61} = 108.98\%$$

الفصل الثامن: الارتباط والانحدار

تقابلنا كثيرا في الحياة العملية مواقف تتضمن متغيرين (ظاهرتين) وأكثر ويكون المطلوب معرفة ما إذا كان هناك علاقة بين هذه المتغيرات وما هو شكل هذه العلاقة وأيضا كيفية التنبؤ بأحد هذين المتغيرين في حالة معرفتنا بالمتغير الآخر فمثلا: دراسة العلاقة بين دخل الفرد وعدد ساعات العمل، دراسة العلاقة بين أوزان وأطوال مجموعة من الأطفال في عمر معين.

أولا: الارتباط Correlation

هو تعيين طبيعة وقوة العلاقة بين متغيرين أو عدمها. أما معامل الارتباط Correlation Coefficient هو مؤشر هذه العلاقة. أول خطوة في تحديد طبيعة العلاقة هي رسم شكل الانتشار.

إذا كان لدينا متغيران فقط نرمز لهما ب X ,Y، حيث المتغير X هو متغير يتم تحديده من قبل الباحث الذي يقوم بالدراسة ويسمى بالمتغير المستقل Independent variable.

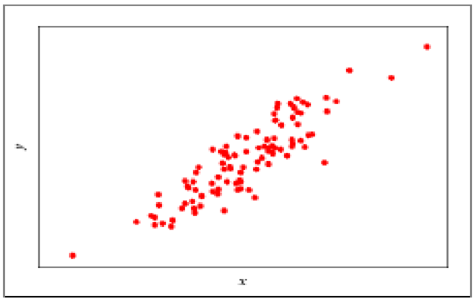
أما المتغير Y فيسمى بالمتغير التابع dependent variable لأن نتيجته غير محددة وتعتمد على قيم المتغير المستقل.

1. أنواع الارتباط:

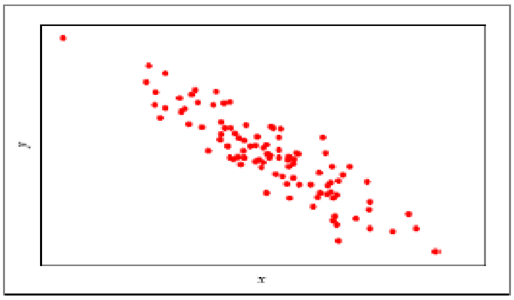
الارتباط الموجب (الطردي) (Positive Correlation)

بأنه علاقة بين متغيرين (X,Y) بحيث إذا تغير أحد المتغرين فإن الآخر يتبعه في نفس الاتجاه.

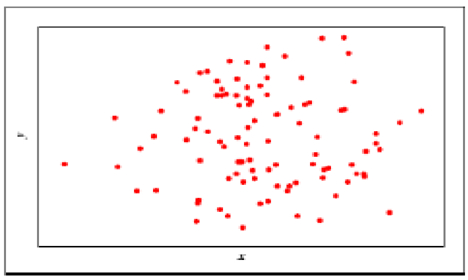
الارتباط السالب (العكسي) (Negative Correlation) بأنه علاقة بين متغيرين (X,Y) بحيث إذا تغير أحد المتغيرين فإن الآخر يتبعه في الاتجاه العكسي.



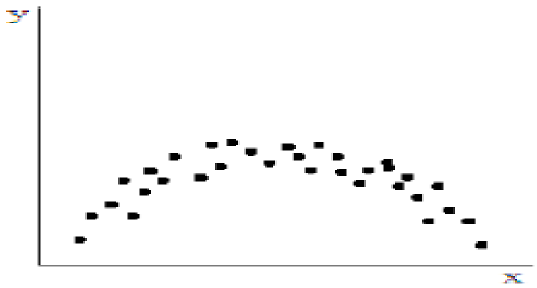
شكل الانتشار الخاص بالارتباط الموجب (الطردي)



شكل الانتشار الخاص بالارتباط السالب (العكسي)

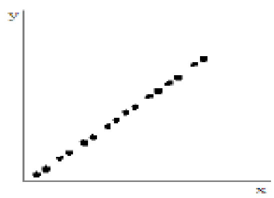


شكل الانتشار الخاص باستقلال متغيرين (ظاهرتين)

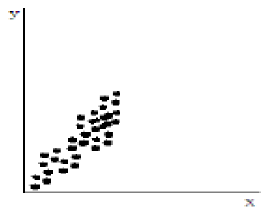


شكل الانتشار الخاص بالعلاقه الغير خطيه بين متغيرين (ظاهرتين)

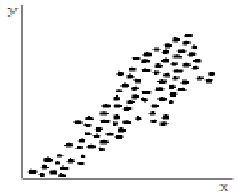
2. شكل الانتشار:



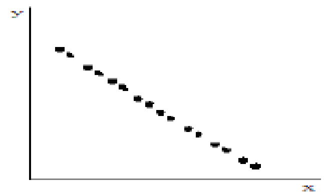
ارتباط طردي تام



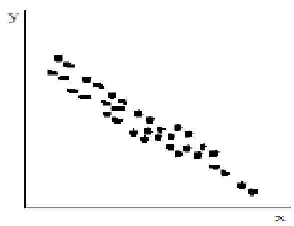
ارتباط طردي قوي



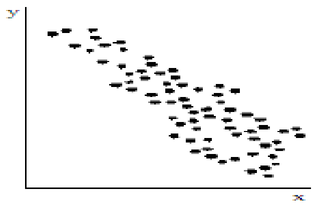
ارتباط طردي



ارتباط عكسي تام



ارتباط عكسي قوي



ارتباط عكسي

3. قياس الارتباط:

يقاس الارتباط بين متغيرين بمقياس إحصائي يسمى " معامل الارتباط " ويعكس هذا المقياس درجة أو قوة العلاقة بين المتغيرين واتجاه هذه العلاقة. وتنحصر قيمة معامل الارتباط بين + 1، - 1.
ويرمز لمعامل الارتباط بالرمز r حيث $-1 \leq r \leq +1$

- فإذا كانت قيمة معامل الارتباط تساوي + 1 فمعنى ذلك أن الارتباط بين المتغيرين طردي تام، وهو أقوى أنواع الارتباط الطردي بين متغيرين.
- وإذا كانت قيمة معامل الارتباط تساوي – 1 فمعنى ذلك أن الارتباط بين المتغيرين عكسي تام وإذا كانت قيمته معامل تساوي صفر، فمعنى ذلك أنه لا يوجد ارتباط بين المتغيرين.
- وكلما اقتربت قيمة معامل الارتباط من + 1 أو – 1 كلما كان الارتباط قوياً، وكلما اقترب من الصفر كلما كان الارتباط ضعيفاً.

كما تدل إشارة معامل الارتباط الموجبة على العلاقة الطردية بيمنا تدل الإشارة السالبة على العلاقة العكسية.

4. معامل بيرسون للارتباط الخطي (Pearson Linear Correlation Coefficient):

معامل بيرسون للارتباط الخطي من أكثر معاملات الارتباط استخداما خاصة في العلوم الاجتماعية. ويفترض Pearson أن المتغيرين كميان، وأن العلاقة بينهما خطية (أي تأخذ شكل خط مستقيم، ويرى بيرسون أن أفضل مقياس للارتباط بين متغيرين قد يختلفان في وحدات القياس و/ أو في مستواهما العام (مثل الارتباط بين العمر والدخل) حيث يقاس العمر بالسنوات ويقاس الدخل بالعملة. ويعطى معامل بيرسون بالعلاقة التالية:

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

أو بطريقة أخرى:

$$r = \frac{n\sum(xy) - (\sum x)(\sum y)}{\sqrt{(n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}}$$

حيث:

$\sum(xy)$: مجموع حاصل ضرب x في y.

$\sum x$: مجموع قيم المتغير x، $\sum y$: مجموع قيم المتغير y

$\sum x^2$: مجموع مربعات قيم المتغير x، $\sum y^2$: مجموع مربعات قيم المتغير y

مثال01:

البيانات التالية تظهر تطور أسعار مادة ما بالدينار والكميات المطلوبة منها (طن) في احدى المدن:

السعر	80	85	87	100	110	123	140	135
الكمية	130	130	125	120	90	80	72	70

المطلوب: أحسب معامل الارتباط الخطي وما مدى قوة العلاقة الخطية؟

الحل:

السعر (x)	**الكمية (y)**	**xy**	**X^2**	**Y^2**
80	130	10400	6400	16900
85	130	11050	7225	16900
87	125	10875	7569	15625
100	120	12000	10000	14400
110	90	9900	12100	8100
123	80	9840	15129	6400
140	72	10080	19600	5184
135	70	9450	18225	4900
860	**817**	**83595**	**96248**	**88409**

معامل الارتباط يمكن حسابه كما يلي:

$$r = \frac{n\sum(x*y) - (\sum x)(\sum y)}{\sqrt{(n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}}$$

$$r = \frac{8*83595 - 860*817}{\sqrt{(8*96248 - 860^2)(8*88409 - 817^2)}}$$

$$r = \frac{668760 - 702620}{\sqrt{(30384)(39783)}} = \frac{-33860}{34767.32} = -0.97$$

من الملاحظ أن علاقة الارتباط الخطي بين الكمية المطلوبة وسعرها هي علاقة ارتباط عكسية قوية.

5. معامل سبيرمان لارتباط الرتب (Spearman Rank Correlation Coefficient) :

يستخدم معامل سبيرمان لارتباط الرتب إذا كان المتغيرين كيفيين، حيث تستعمل رتب تصاعدية أو تنازلية عوضا عن القيم العددية للمتغيرات المدروسة، فإذا كان الترتيب التصاعدي للمتغير المستقل يقابله ترتيب تصاعدي للمتغير التابع نقول أن الارتباط موجب (علاقة طردية) والعكس صحيح.

فإذا فرضنا أن المتغير X له الرتب R_X وأن المتغير Yله الرتب R_Y، وبفرض أن d ترمز لفرق الرتبتين، بمعنى $d= R_X- R_Y$.

فإن معامل سبيرمان لارتباط الرتب يعطى بالصيغة التالية:

$$r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$$

حيث n: هي عدد الأزواج المرتبة.

مثال02:

لدراسة علاقة ارتباط تقديرات الطلاب في مادة الإحصاء وتقديراتهم في مادة الرياضيات، اخترنا خمسة طلبة وكانت تقديراتهم كما يلي:

الطالب	1	2	3	4	5
تقديرات الإحصاء (x)	ممتاز	ضعيف	جيد	جيد جدا	مقبول
تقديرات الرياضيات (y)	جيد جدا	جيد	مقبول	ممتاز	ضعيف

المطلوب: هل توجد علاقة ارتباط؟ مانوعها ومدى قوتها؟

الحل:

1-نرتب الدرجات التقديرية ترتيبا تصاعديا أو تنازليا.

2-نحسب مربع الفروق بين الدرجات التقديرية للمقياسين حسب الجدول التالي:

رقم الطالب	x	y	رتب x	رتب y	d_i	d_i^2
1	ممتاز	جيد جدا	1	2	-1	1
2	ضعيف	جيد	5	3	2	4
3	جيد	مقبول	3	4	-1	1
4	جيد جدا	ممتاز	2	1	1	1
5	مقبول	ضعيف	4	5	1-	1
المجموع					0	8

$$r_s = 1 - \frac{6\sum d_i^2}{n(n^2-1)} = 1 - \frac{6\times 8}{5(25-1)} = 1 - \frac{48}{120} = 1 - 0.4 = 0.6$$

نلاحظ وجود علاقة ارتباط طردية متوسطة بين تقديرات الطلبة في مادة الإحصاء وتقديراتهم في مادة الرياضيات.

ثانيا: الانحدار

الانحدار هو أسلوب يمكن بواسطته تقدير قيمة أحد المتغيرين بمعلومية قيمة المتغير الآخر عن طريق معادلة الانحدار، وله أنواع:

الانحدار الخطي البسيط: يعني أن المتغير التابع Y يعتمد على متغير مستقل واحد وهو X بينما خطي فتعني أن العلاقة بين المتغيرين (X,Y) علاقة خطية.

-الانحدار المتعدد: إذا كان المتغير التابع y يعتمد على أكثر من متغير مستقل.

الانحدار غير الخطي: إذا كانت العلاقة بين المتغير التابع Y والمتغيرات المستقلة غير خطية كأن تكون من الدرجة الثانية أو أسية.

1. الانحدار الخطي البسيط التام:

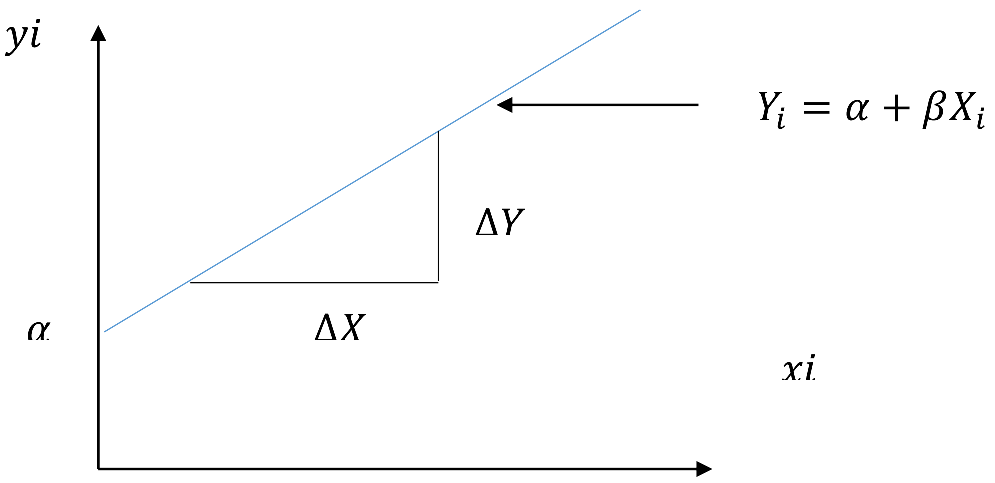
الانحدار الخطي البسيط التام، هو الذي تكون كل نقاط انتشاره من تمثيل متغيراته على معلم متعامد، على استقامة واحدة، سواء كانت في الاتجاه الموجب أو في الاتجاه السالب، وتكون معادلة الانحدار كما يلي:

$$Y_i = \alpha + \beta X_i$$

حيث:α : ميل دالة الانحدار.

β: ثابت الدالة، وتدل على قيمة الدالة عندما ينعدم x_i

وهي دالة خط مستقيم يمكن إيجاد معالمها α و β بسهولة وذلك باستخدام الطرق الهندسية، حيث يتم إيصال بين هذه النقاط لشكل الانتشار لنتحصل على خط مستقيم ثم نبحث عن ميله الذي يساوي ظل الزاوية المحصورة بين المنحنى والمستقيم الافقي الموازي لمحور الفواصل، ونجد بعد ذلك قيمة الثابت الدالة α ، وهو عبارة عن نقطة تقاطع المنحنى مع محور التراتيب، كما يوضحه الشكل التالي:



بما أن B هو ميل الدالة فهو عبارة عن ظل الزاوية المحصورة بين المنحنى والمستقيم الافقي الموازي لمحور الفواصل، أي:

$$\beta = \frac{\Delta Y}{\Delta X}$$

بينما α هي عبارة عن نقطة تقاطع خط الانحدار مع المحور العمودي، ويمكن ايجادها هندسيا من الرسم.

2. الانحدار الخطي البسيط غير التام:

بعد تمثيل الأزواج المرتبة بالمستوى نحصل على شكل الانتشار لا تكون على استقامة واحدة، ولكنها تأخذ اتجاها يمكن تقريبه من معادلة خط مستقيم، فإذا أظهر الشكل الانتشاري للبيانات أن هناك علاقة خطية بين المتغيرين نقوم بتقدير خط الانحدار Y على X حيث يستحيل إيجاد المعادلة الحقيقية، التي نفترضها كما يلي:

$$Y = \alpha + \beta X_i + \varepsilon_i$$

حيث الشق الأول من المعادلة وهو$\alpha + \beta X_i$ ، هو معادلة خط مستقيم وقد أضيف لها المقدار ε_iالذي هو قيمة البعد بين النقاط الحقيقية ومعادلة الخط المستقيم، حيث أنه يستحيل إيجاد هذه المعادلة لذلك يتم تقريبها تقديريا إلى المعادلة الخطية التالية:

$$\hat{Y} = \alpha + \beta X_i$$

حيث:

α:ثابت الانحدار

β: ميل الخط المستقيم أو معامل انحدار Yعلى X

وتحسب معالم النموذج α،β بطريقة المربعات الصغرى والتي تعطى من العلاقتين التاليتين:

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

$$\alpha = \bar{y} - \beta\bar{x} = \frac{\sum y_i - \beta\sum x_i}{n}$$

مثال 03:

بالرجوع للمثال 01 والذي يظهر تطور أسعار مادة ما بالدينار والكميات المطلوبة منها (طن) في احدى المدن:

السعر	80	85	87	100	110	123	140	135
الكمية	130	130	125	120	90	80	72	70

المطلوب:

نفرض أن العلاقة خطية بين السعر والكمية، قدر معلمات النموذج الخطي البسيط؟

الحل:

تقدير معلمات النموذج الخطي البسيط:

السعر (x)	الكمية (y)	xy	X^2	Y^2
80	130	10400	6400	16900
85	130	11050	7225	16900
87	125	10875	7569	15625
100	120	12000	10000	14400
110	90	9900	12100	8100
123	80	9840	15129	6400
140	72	10080	19600	5184
135	70	9450	18225	4900
860	**817**	**83595**	**96248**	**88409**

معادلة الانحدار من الشكل:

$$\hat{Y} = \alpha + \beta X_i$$

وتحسب معالم النموذج α،β بطريقة المربعات الصغرى والتي تعطى من العلاقتين التاليتين:

$$\beta = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sum(x_i-\bar{x})^2} = \frac{n\sum(x_iy_i)-(\sum x_i)(\sum y_i)}{n\sum x_i^2-(\sum x_i)^2} = \frac{8*83595-860*817}{8*96248-860^2}$$

$$= \frac{-33860}{30384} = -0,11$$

$$\alpha = \bar{y} - \beta\bar{x}$$

$$= \frac{\sum y_i - \beta\sum x_i}{n}$$

$$= \frac{817-(-0.11)*860}{8}$$

$$= \frac{87,32}{8} = 113,95$$

$$\hat{Y} = 113,95 - 0,11X_i$$

تمارين محلولة:

التمرين الأول:

الجدول التالي يمثل الوقت المستغرق اليومي في مواقع التواصل الاجتماعي ل 10 أشخاص:

السن	18	24	28	32	40	45	57	60	64	68
الوقت المخصص (ساعة/يوم)	6	7	6	4	3	3	2	2	1	1

المطلوب:

1. حدد المتغير التابع والمتغير المستقل؟
2. أرسم شكل الانتشار وحدد طبيعة ونوع العلاقة بين المتغيرين؟
3. أحسب معامل الارتباط وفسره؟
4. نفرض العلاقة خطية بين المتغيرين، قدر معادلة خط الانحدار بطريقة المربعات الصغرى؟
5. ماهو الوقت المستغرق في مواقع التواصل الاجتماعي لشخص عمره 21 سنة؟

التمرين الثاني:

لدراسة علاقة الاستهلاك المحلي بالإنتاج لمادة معينة (الوحدة : مليون)خلال عدة سنوات، أخذت 10 سنوات كما يلي:

الاستهلاك المحلي	6	8	9	8	7	6	5	6	5	5
الانتاج	10	13	15	14	9	7	6	6	5	5

المطلوب:

1. حدد المتغير التابع والمتغير المستقل؟
2. أرسم شكل الانتشار ماذا تستنتج؟
3. أحسب معامل الارتباط وفسره؟
4. قدر معادلة خط الانحدار بطريقة المربعات الصغرى؟
5. توقع قيمة الاستهلاك عندما يبلغ الإنتاج 16 مليون وحدة؟

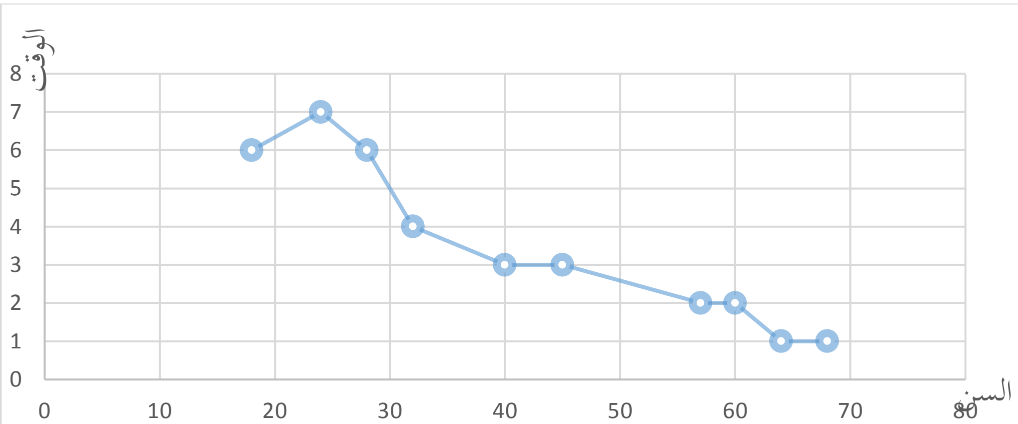
حلول التمارين:

حل التمرين الأول:

رقم	السن x_i	الوقت y_i	$X*x$	Y*y	X*y
1	18	6	324	36	108
2	24	7	576	49	168
3	28	6	784	36	168
4	32	4	1024	16	128
5	40	3	1600	9	120
6	45	3	2025	9	135
7	57	2	3249	4	114
8	60	2	3600	4	120
9	64	1	4096	1	64
10	68	1	4624	1	68
المجموع	**436**	**35**	**21902**	**165**	**1193**

1. **تحديد المتغير التابع والمتغير المستقل:**
المتغير التابع هو الوقت المستغرق على مواقع التواصل الاجتماعي بينما المتغير المستقل هو السن.

2. **رسم شكل الانتشار وتحديد طبيعة ونوع العلاقة بين المتغيرين:**



من خلال شكل الانتشار بين المتغيرين نلاحظ وجود علاقة خطية عكسية بين السن والوقت المستغرق على مواقع التواصل.

3. حساب معامل الارتباط:

$$r = \frac{n\sum(x*y) - (\sum x)(\sum y)}{\sqrt{(n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}} = \frac{10*1193 - 436*35}{\sqrt{(10*21902 - 436^2)(10*165 - 35^2)}}$$

$$r = \frac{-3330}{\sqrt{(28924)(425)}} = \frac{-3330}{3506,09} = -0,95$$

من خلال معامل الارتباط نلاحظ وجود علاقة عكسية قوية بين المتغيرين.

4. تقدير معادلة خط الانحدار بطريقة المربعات الصغرى:

معادلة الانحدار من الشكل:

$$\hat{Y} = \alpha + \beta X_i$$

وتحسب معالم النموذج α،β بطريقة المربعات الصغرى والتي تعطى من العلاقتين التاليتين:

$$\beta = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} = \frac{n\sum(x_i y_i) - (\sum x_i)(\sum y_i)}{n\sum x_i^2 - (\sum x_i)^2} = \frac{10*1193 - 436*35}{10*21902 - 436^2} = \frac{-3330}{28924} = -0,12$$

$$\alpha = \bar{y} - \beta\bar{x} = \frac{\sum y_i - \beta\sum x_i}{n} = \frac{35 - (-0.12)*436}{10} = \frac{87,32}{10} = 8,73$$

$$\hat{Y} = 8,73 - 0,12X_i$$

5. الوقت المستغرق في مواقع التواصل الاجتماعي لشخص عمره 21 سنة:

بتعويض السن (xi) في معادلة الانحدار المقدرة نجد الوقت المستغرق كما يلي:

$$\hat{Y} = 8,73 - 0,12*(21) = 6.21$$

إذن الوقت المستغرق في مواقع التواصل الاجتماعي لشخص عمره 21 سنة هو 6.21 ساعة في اليوم.

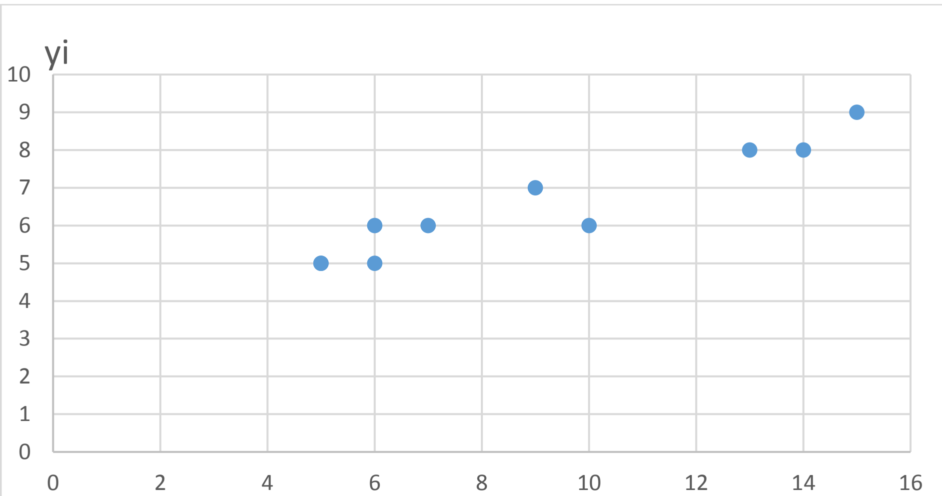
حل التمرين الثاني:

y_i	x_i	$x_i * y_i$	x_i^2	y_i^2
6	10	60	100	36
8	13	104	169	64
9	15	135	225	81
8	14	112	196	64
7	9	63	81	49
6	7	42	49	36
5	6	30	36	25
6	6	36	36	36
5	5	25	25	25
5	5	25	25	25
المجموع	**90**	**632**	**942**	**441**

1. تحديد المتغير التابع والمتغير المستقل:

المتغير التابع هو الاستهلاك المحلي بينما المتغير المستقل هو الانتاج.

2. رسم شكل الانتشار وتحديد طبيعة ونوع العلاقة بين المتغيرين:



من خلال شكل الانتشار بين المتغيرين نلاحظ وجود علاقة خطية بين الاستهلاك المحلي والإنتاج لهذه المادة.

3. حساب معامل الارتباط:

$$r = \frac{n\sum(x*y) - (\sum x)(\sum y)}{\sqrt{(n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}}$$
$$= \frac{10*632 - 90*65}{\sqrt{(10*441 - 65^2)(10*942 - 90^2)}}$$

$$r = 0.95$$

من خلال معامل الارتباط نلاحظ وجود علاقة طردية قوية بين الاستهلاك المحلي والإنتاج لهذه المادة.

4. تقدير معادلة خط الانحدار بطريقة المربعات الصغرى:

معادلة الانحدار من الشكل:

$$\hat{Y} = \alpha + \beta X_i$$

وتحسب معالم النموذج α،β بطريقة المربعات الصغرى والتي تعطى من العلاقتين التاليتين:

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

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

$$= \frac{10*632 - 90*65}{10*942 - 90^2}$$

$$\beta = 0.36$$

$$\alpha = \bar{y} - \beta\bar{x}$$

$$= \frac{65}{10} - 0.36*\frac{90}{10}$$

$$\alpha = 3.26$$

وبالتالي معادلة خط الانحدار هي:

$$\hat{Y} = 3,26 + 0,36X_i$$

5. توقع الاستهلاك عندما يبلغ الإنتاج 16 مليون وحدة:

$$\hat{Y} = 3.26 + 0,36 * (16) = 9.02$$

أي أن الاستهلاك المحلي يصل إلى 9.02 مليون وحدة عندما يبلغ الإنتاج 16 مليون وحدة لهذه المادة.

المراجع:

- محمد راتول، (2006)، **الإحصاء الوصفي**، ديوان المطبوعات الجامعية، الطبعة الثانية، الجزائر.
- جلاطو جيلالي، (2001)، **الإحصاء مع تمارين ومسائل محلولة**، ديوان المطبوعات الجامعية، الجزائر.
- طيار أحسن، (2019)، **الإحصاء الوصفي: دروس مفصلة وتمارين محلولة**، دار هومة.
- طبيه أحمد عبد السميع، (2008)، **مبادئ الإحصاء**، دار البداية، عمان، الأردن.
- محمد صبحي أبو صالح و عدنان محمد عوض، (2007)، **مقدمة في الإحصاء**، دار المسيرة، عمان.

- David P. Doane, Lori E. Seward, 2016, **Applied Statistics In Business And Economics**, Fifth Edition, McGraw-Hill Education.
- Hurlin Christophe, 2015, **Statistique et Probabilités en économie-gestion**, Dunod.
- *Goldfarb, B., & Pardoux, C. (2011).* ***Introduction à la méthode statistique****, 6° édition. Dunod.*
- *Anderson, D. R., Sweeney, D. J., Camm, J. D., Williams, T. A., & Cochran, J. J. (2015).* ***Statistiques pour l'économie et la gestion****. De Boeck Supérieur.*