الجمهورية الجزائرية الديمقراطية الشعبية

وزارة التعليم العالي والبحث العلمي

جامعة ابن خلدون تيارت

مطبوعة دروس في مقياس الإحصاء 1

موجهة لطلبة السنة الأولى جذع مشترك

من إعداد الدكتورة اجري خيرة

السنة الجامعية 2023-2024

فهرس المحتويات

الفصل الأول: مفاهيم عامة حول الإحصاء

1- تمهيد

2- تعريف الإحصاء

3- التطرق إلى مصطلحات إحصائية عامة (المجتمع، العينة..).

4- مصادر وأساليب البيانات الإحصائية.

I) تمهيد

كلمة إحصاء كلمة شائعة الاستخدام في مختلف حقول العلم والمعرفة، خاصة الاقتصاد والإدارة، فهو يهتم بتوفير جميع المقاييس الكمية للظواهر الاقتصادية والاجتماعية، مما يساعد على اتخاذ قرارات وأحكام سليمة ليمهد الطريق لتطبيق أسلوب البحث العلمي في تحليل هذه الظواهر، ومن هنا يمكننا تعريف الإحصاء على أنه:

II) تعريف الإحصاء

1) **الإحصاء بالمفهوم العامي**: من المفاهيم الشائعة بين الناس عن الإحصاء ما هو إلا أرقام أو بيانات رقمية فقط كتعداد السكان، المواليد، الوفيات...

2) **أما الإحصاء بمفهومه المحدود**: فما هو إلا عد وتعداد وحصر الأشياء والتعبير عنها بالأرقام.

3) **أما الإحصاء كعلم**: فهو الذي يهتم بطرق جميع البيانات المتعلقة بظاهرة ما وتلخيصها في جداول إحصائية (تبويب البيانات) وتحليلها باستخدام مجموعة من المقاييس العلمية كمقاييس النزعة المركزية، مقاييس التشتت وغيرها من المقاييس الأخرى، كل ذلك من أجل الوصول إلى قرارات سليمة تستخدم للتنبؤ مستقبلا.

وبدوره، الإحصاء ينقسم إلى فرعين رئيسيين:

الإحصاء الوصفي	الإحصاء الاستدلالي
هو ذلك الجزء من الإحصاء يهتم بتلخيص البيانات ووصفها وتحليلها.	يقوم على فكرة اختيار جزء من المجتمع يسمى العينة، يتم استخدام بيانات هذه العينة بدل المجتمع للتوصل إلى نتائج يمكن تعميمها على مجتمع الدراسة.

III) بعض المصطلحات الإحصائية

1) **المجتمع الإحصائي**: هي مجموع الوحدات الإحصائية والمعرفة بشكل دقيق والتي تشترك فيما بينها في (الصفة) الأساسية محل اهتمام الباحث، مثل: مجتمع الطلبة – مجتمع الأسر...

2) **الوحدات الإحصائية**: هي العناصر أو الأفراد التي تشكل لنا المجموعة المدروسة أو المجتمع الإحصائي، مثل: "طلبة سنة أولى جذع مشترك كلية العلوم الاقتصادية دفعة 2021" ←كل طالب هو عبارة عن وحدة إحصائية في المجتمع الإحصائي هو مجتمع الطلبة.

3) **العينة الإحصائية**: هي جزء من المجتمع الإحصائي يتم استخراجها بطرق إحصائية معينة (عشوائية أو غير عشوائية) حتى تكون ممثلة للمجتمع الإحصائي أحسن تمثيل، ويتم الاعتماد عليها بدل المجتمع للأسباب التالية:

- كبر حجم المجتمع.
- ربح الوقت والجهد والمال.

1.3– أنواع العينات:

1.1.3– العينات الاحتمالية (العشوائية): من أجل أن تكون العينة ممثلة أحسن تمثيل للمجتمع الدراسة، فإن أكثر الطرق استخداما هو أسلوب العشوائية، والتي تعني أن لأفراد أو عناصر المجتمع فرصا متساوية في الاختيار، مما يعني أن لجميع عناصر المجتمع الفرصة نفسها لدخول العينة، ومن أهم أنواعها: العينة العشوائية البسيطة، الطبقية، المنتظمة، العنقودية.

2.1.3– العينات الغير الاحتمالية (الغير العشوائية): هي العينات التي لا يتحقق مبدأ العشوائية عند اختيارها، والتي لا تعطي لأفراد أو عناصر المجتمع فرصا متساوية عند اختيار العينة، ومن أهم أنواعها: العينة بالمصادفة، العينة الحصصية، العينة العمدية أو المقصودة.

4) **الصفة (المتغير الإحصائي)**: عبارة عن الميزة التي تمتاز بها كل وحدة إحصائية مثلا: "نريد أن ندرس أطوال طلبة السنة الأولى جذع مشترك".

المجتمع الإحصائي: مجتمع الطلبة، **الوحدة الإحصائية**: طالب، **المتغير الإحصائي (الصفة)**: طول الطلبة.

وتنقسم الصفة إلى:

أ– **الصفة الكيفية أو النوعية (الإسمية)**: هي تلك الصفات التي لا يمكن قياس كيفياتها، وتنقسم إلى قسمين: قابلة للترتيب كالمستويات التعليمية مثلا، وغير قابلة للترتيب ك:

- الجنس (ذكر – أنثى) ← لدينا كيفيتين.
- الحالة المدنية (أعزب – متزوج – أرمل – مطلق) ← 4 كيفيات.

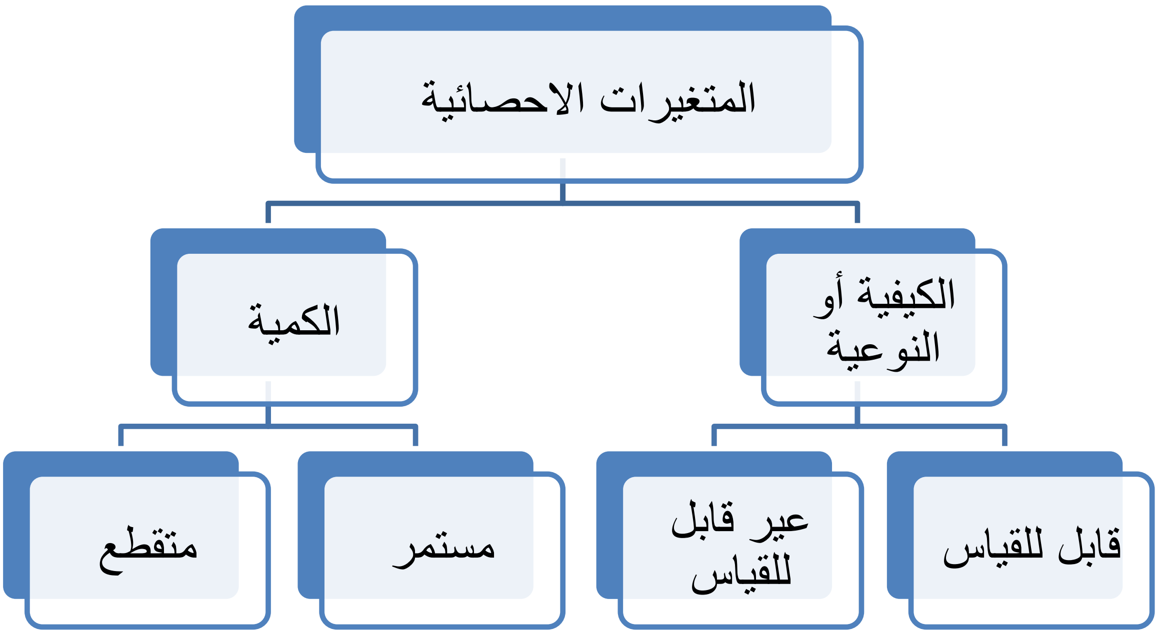
أو بصفة عامة هي تلك الصفات التي لا يمكن التعبير عنها بالأرقام بل يعبر عنها بمفردات إسمية (كلمات).

ب- **الصفة الكمية (متغير احصائي كمي)**: هي الصفة التي يمكن قياس كيفياتها والتي نستطيع التعبير عنها عن طريق قيم وأرقام مثل :الوزن ،الطول، السعر، الأجر، وبدورها تنقسم إلى قسمين:

_ **متغير إحصائي متقطع**: عدد القيم التي نستطيع أن نأخذها منفردة أو منفصلة عن بعضها البعض مثل :عدد الأطفال، عدد السيارات ،عدد الأسر .

_ **متغير إحصائي مستمر**: عندما نكون أمام حالة عدد لانهائي من القيم في مجال محدود مثل : السن، الطول، الأجر، النقاط......

أنواع المتغيرات الاحصائية



مصادر وأساليب جمع البيانات:

إن جمع بيانات تفصيلية حول أي ظاهرة تتطلب توفير معطيات رقمية أو وصفية دقيقة خلال فترة زمنية معينة، وهناك عدة طرق يمكن للباحث اللجوء إليها من أجل الحصول على البيانات الاحصائية والمتمثلة في:

1) مصادر أولية:

هي المصادر التي يتحصل عليها الباحث بشكل مباشر، ويقوم فيها بجمع البيانات بنفسه، فلو كان اهتمام الباحث بجمع بيانات عن الأسرة مثلا، يقوم بإجراء مقابلة مع رب الأسرة لأخذ البيانات مثل: المنطقة التابعة لها، الحي الذي يسكن فيه، الجنسية، المهنة، عدد أفراد الأسرة إلى غير ذلك..

$\Longleftarrow$ يتميز هذا النوع بالدقة و الثقة، ولكن أهم ما يعاب عليها أنها تحتاج إلى وقت وجهد، ومكلفة من الناحية المادية.

2) مصادر ثانوية:

هي المصادر التي يحصل منها الباحث على بيانات بشكل غير مباشر، كالحصول عليها من أشخاص آخرين – أجهزة وهيئات رسمية كالديوان الوطني للإحصائيات ONS، دراسات سابقة على شكل بحوث ومقالات علمية وغيرها.

$\Longleftarrow$ ومن مزايا هذا النوع توفير الوقت والجهد والمال، إلا أن درجة ثقة الباحث فيها ليست بنفس الدرجة في حالة المصادر الأولية.

ومن خلال ما سبق يقوم الباحث بجمع البيانات الاحصائية باستخدام أحد الأسلوبين التاليين:

1) **أسلوب الحصر الشامل**: ويقصد به جمع البيانات حول كل مفردات أو عناصر المجتمع الاحصائي المدروس، ويتميز هذا الأسلوب بالشمول وعدم التحيز ودقة النتائج، ولكن ما يعاب عليه أنه يحتاج إلى وقت وجهد، ومكلف من الناحية المادية.
2) **أسلوب المعاينة**.

IV) تمارين الفصل الأول

- **أسئلة نظرية**:

- قم بإعطاء تعريف مختصر لكل من: علم الاحصاء، الاحصاء الاستدلالي، المجتمع الإحصائي، الوحدة الاحصائية، العينة، المتغير الاحصائي؛
- ما هو الفرق بين مصادر جمع البيانات الأولية والثانوية؛
- ما هو الفرق بين المتغيرات الاحصائية الكمية والنوعية.

- **التمرين الأول**:

حدد كل من المجتمع الاحصائي والوحدة الاحصائية، المتغير الاحصائي ونوعه للعبارات التالية:

- الدخل الشهري لعمال مؤسسة ما؛
- غيابات الطلبة في حصص الأعمال الموجهة؛
- عدد أطفال أسر بلدية ما؛
- جنس المواليد الجدد لدولة ما؛
- المستوى التعليمي لموظفي بنك التنمية المحلية؛
- وظيفة أولياء تلاميذ مدرسة ابتدائية؛
- قامة لاعبين كرة السلة فريق الاتحاد.
- جنسية لاعبين كرة القدم لمنتخب الجزائر.

التمرين الثاني:

على الطالب إعطاء ثلاث أمثلة لكل متغير احصائي: كمي مستمر، كمي متقطع، كيفي قابل للترتيب، كيفي غير قابل للترتيب.

الفصل الثاني: عرض وتبويب البيانات الإحصائية (جداول التوزيعات التكرارية والتمثيلات البيانية).

1- تعريف البيانات الإحصائية.

2- بناء الجداول وأنواعها (إيجاد الفئة – التكرارات – مركز الفئة – تكرار المجتمع ..)

3- التمثيل البياني حول نوع المتغير الإحصائي

1- تمهيد

يعتبر عرض وتبويب البيانات الإحصائية الخطوة الثانية بعد تجميع هذه البيانات في مفهوم التحليل الإحصائي، حيث يلجأ الباحث إلى حصر وتصنيف هذه البيانات وعرضها بطريقة مختصرة تساعد على فهمها وتحليلها إحصائيا، للتعرف عليها ووصفها ومقارنتها بغيرها من الظواهر، والخروج ببعض المدلولات الإحصائية عن مجتمع الدراسة.

2- طريقة عرض البيانات الإحصائية

يتم عرض وتبويب البيانات الإحصائية عن طريق العرض الجدولي والعرض البياني (التمثيلات البيانية) لكل من البيانات النوعية والأخرى الكمية، حيث لا تشكل المعطيات الإحصائية التي يتحصل عليها الباحث إلا كتلة غير قابلة للاستعمال وليس لها أي معنى إلا إذا كانت هذه المعطيات مصنفة ومرتبة بصورة علمية.

أ- العرض الجدولي والبياني للبيانات الاحصائية النوعية

يتم عرض البيانات بعد جمعها ضمن جدول مرتب بشكل متسلسل ومنطقي مما يسهل على الباحث التحليل.

- تصنف الصفة النوعي (الكيفية، الإسمية) في الجدول التالي:

التكرارات	الكيفيات
n_1	m_1
n_2	m_2
.	.
.	.
.	.
n_k.	m_k

n_1: عدد الحالات الممكنة التي تمتاز بالكيفية m_1.

مثال: قمنا بتوزيع عمال جزائريين حسب قطاع النشاط كالتالي:

التكرارات	الصفات (الكيفيات)
537	صناعة
657	بناء وأشغال عمومية
169	النقل والمواصلات
618	التجارة والخدمات
940	إدارة

- التكرارات المطلقة
- التكرار الكلي هو مجموع التكرارات المطلقة

$$\sum n_{\mathrm{i}} = 537 + \cdots + 940 = 2921$$

المطلوب:

1– ما هي المجموعة الإحصائية (المجتمع الإحصائي)؟ **العمال الجزائريون**.

2– ما هي الوحدة الاحصائية؟ **العامل**

3– ما هي الصفة المدروسة؟ **نوع النشاط**.

4– ما نوعها؟ صفة كيفية (نوعية أو إسمية)

5– كم لهذه الصفة من كيفية؟ **05 كيفيات**.

التمثيل البياني: لتمثيل البيانات النوعية يمكن استخدام **الأعمدة** أو باستخدام **شكل الدائرة**.

أ– باستخدام الأعمدة البيانية

ب – باستخدام شكل الدائرة

1– حساب مجموع التكرارات:

$$\sum n_i = 537 + \cdots + 940 = 2921$$

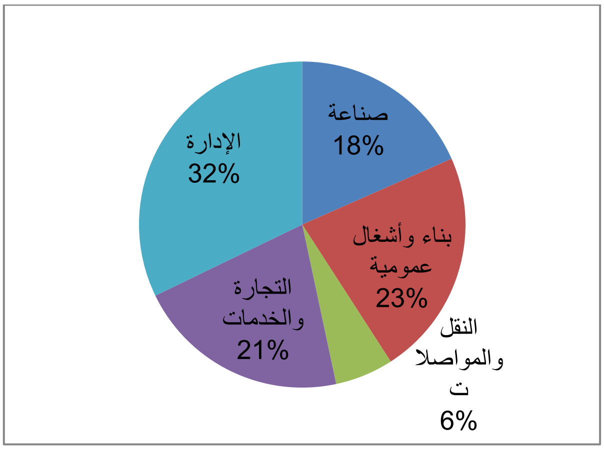
2– تحويل التكرارات إلى زوايا:

$$\left.\begin{array}{l} 2921 \rightarrow 360° \\ \\ 537 \rightarrow x \end{array}\right\} \Rightarrow \frac{360 \times 537}{2921}$$

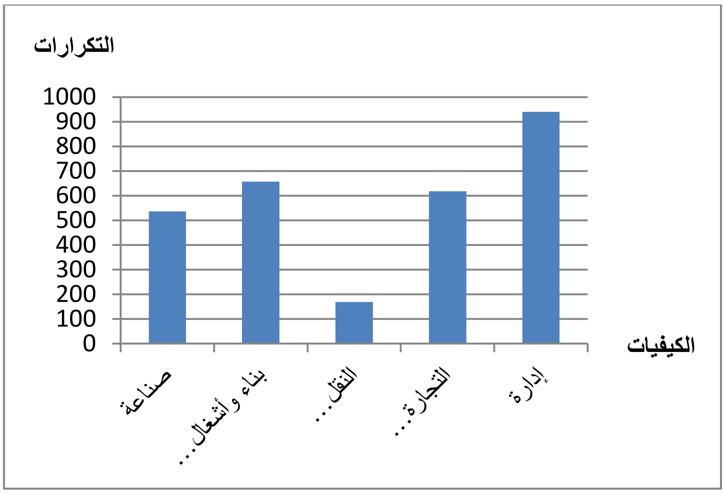
$$= 66,18$$

وهكذا مع باقي التكرارات فنتحصل على:

الزوايا ⇒ 66,18 - 80.97 - 20,82 - 76,16 - 115,85



بالإضافة إلى تمثيلات أخرى كالأعمدة المستطيلة:



ب– العرض الجدولي والبياني للبيانات الاحصائية الكمية (الصفة الكمية):

قبل التطرق لتحليل البيانات الكمية يجب أولا معرفة:

– **السلسلة الإحصائية**: هي قائمة من الأعداد أو القيم الإحصائية، إذا كانت هذه القائمة معطاة بطريقة عشوائية تسمى بـ السلسلة الإحصائية غير المنتظمة، واحتراما للخطوات الإحصائية يجب ترتيب السلسلة بصورة تصاعدية أو تنازلية، وبالتالي نتحصل على **سلسلة إحصائية منتظمة**.

مثال:

70، 60، 50، 40، 30، 20، 10 ← **مدى السلسلة**: هو الفرق بين القيمتين الأعلى والدنيا للسلسلة:

$$R = E = 70 - 10 = 50$$

ـ **الصفة الكمية (متغير احصائي كمي)**: هي الصفة التي يمكن قياس كيفياتها والتي نستطيع التعبير عنها عن طريق قيم وأرقام مثل: الوزن ،الطول، السعر، الأجر، وبدورها تنقسم إلى قسمين:

ـ **متغير إحصائي متقطع**: عدد القيم التي نستطيع أن نأخذها منفردة أو منفصلة عن بعضها البعض مثل :عدد الأطفال، عدد السيارات ،عدد الأسر .

ـ **متغير أحسائي مستمر**: عندما نكون أمام حالة عدد لانهائي من القيم في مجال محدود مثل :السن، الطول، الأجر، النقاط......

1_ حالة متغير إحصائية متقطعة:

مثال: قمنا بدراسة إحصائية لعينة تتكون من 50 أسرة (n=50) وزعناهم في جدول تكراري حسب عدد الأطفال وتحصلها على النتائج التالية:

عدد الأطفال	2	3	4	5	6	7	8
عدد الأسر	2	4	6	9	11	13	5

المطلوب:

- أوجد كل من المجتمع الاحصائي، الوحدة الاحصائية، المتغير الاحصائي، ونوعه؛
- تكوين الجدول الاحصائي مع ايجاد كل من التكرار النسبي، المئوي، التكرار المتجمع الصاعد والنازل،
- ما هو عدد الاسرة الذين لديهم 5 أطفال.
- ما هو عدد الأسر الذي عدد أطفالها أكثر من 6 أطفال
- ما هو عدد الأسر الذي عدد أطفالها أقل من 4 أطفال؟
- ما هو نسبة الأسر الذي عدد أطفالها أقل من 5 ؟

الحل:

مجتمع الدراسة : مجتمع الأسرة.

الوحدة الإحصائية: أسرة واحدة

المتغير الإحصائي : عدد الأطفال

النوع : كمي متقطع.

تكوين الجدول الاحصائي:

عدد الأطفال x_i	**عدد الأسر** n_i	**التكرار النسبي** (f_i)	**التكرار المئوي** $(f_i\%)$	**التكرار المتجمع الصاعد** $(N\uparrow)$	**التكرار المتجمع النازل** $(N\downarrow)$
2	2	0,04	4	2	50
3	4	0,08	8	6	48
4	6	0,12	12	12	44
5	9	0,18	18	21	38
6	11	0,22	22	32	29
7	13	0,26	26	45	18
8	5	0,1	10	50	5

$\sum n_i$	50	1	100	/	/

قبل الشروع في حل المثال السابق، سنتطرق أولا إلى إعطاء المفاهيم التالية،

– **التكرار المطلق**: هو عدد الحالات الممكنة التي تمتاز بكيفية ما مثل:

$n_4 = 9$: تكرار مطلق خاص بالمجموعة 04 التي تمتاز بالكيفية التالية:

هناك 9 أسر من بين 50 أسرة عدد أطفالها يساوي 5.

– **التكرار النسبي (f_i)** : هو حاصل قسمة التكرارات المطلقة على التكرار الكلي :

$$f_i = \frac{\text{التكرار المطلق}}{\text{التكرار الكلي}} = \frac{n_i}{\sum n_i}$$

ملاحظة: مجموع التكرارات النسبة يساوي دائما 1 أي:

$$\sum f_i = 1$$

– **التكرار المئوي $(f_i\%)$: هو حاصل ضرب التكرار النسبي في 100.**

$$f_i\% = \frac{\text{التكرار المطلق}}{\text{التكرار الكلي}} \times 100 = \frac{n_i}{\sum n_i} \times 100$$

مثلا: القيمة $(f_i\% = 22\%)$ معناها أن هناك 22 % من الأسرة عدد أطفالها يساوي 6.

ملاحظة: مجموع التكرارات المئوية يساوي دائما 100 أي:

$$\sum f_i\% = 100$$

– **التكرار المجتمع الصاعد $(N_i \uparrow)$**: هو عبارة عن عدد الأفراد أو الوحدات الإحصائية الذين لديهم الكيفية **أقل أو يساوي** الكيفية الملائمة.

من المثال السابق القيمة $(N_i \uparrow= 45)$ معناها أن هناك 45 أسرة من بين 50 أسرة عدد أطفالها أقل أو يساوي 7.

- **التكرار المجتمع النازل** $(N_i \downarrow)$: هو عبارة عن عدد الأفراد أو الوحدات الإحصائية الذين لديهم الكيفية **أكبر أو يساوي** الكيفية الملائمة.

من المثال السابق القيمة $(N_i \downarrow = 29)$ معناها أن هناك 29 أسرة من بين 50 أسرة عدد أطفالها أكثر أو يساوي 6.

بعض الملاحظات فيما يخص التكرارات المتجمعة الصاعدة والنازلة:

1_القيمة الأولى لتكرار المتجمع الصاعدة هو نفسه القيمة الأولى للتكرار المطلق؛

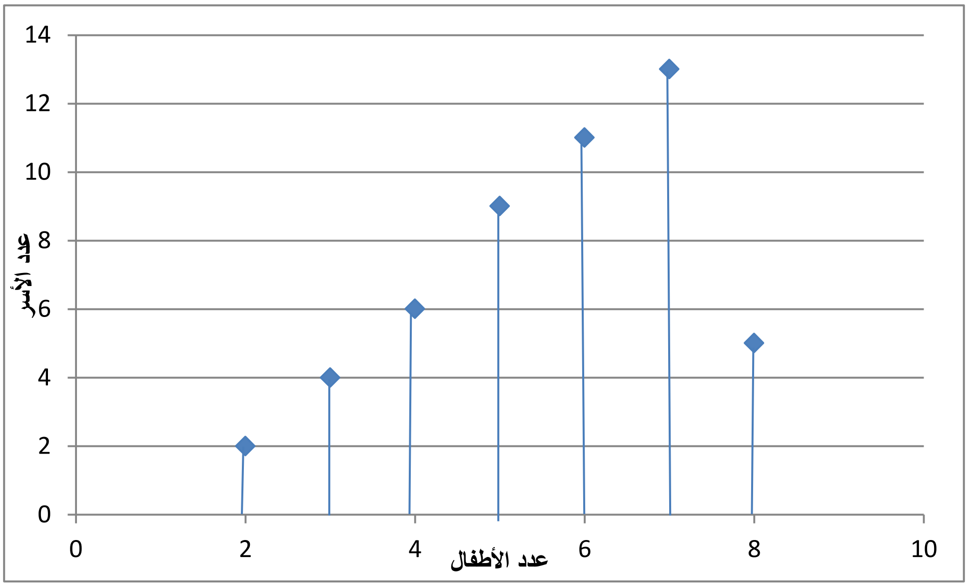
2_ القيمة الأخيرة لتكرار المتجمع الصاعد هو مجموع التكرارات.

3_القيمة الأولى لتكرار المتجمع النازل هو مجموع التكرارات .

4_ القيمة الأخيرة لتكرار المتجمع النازل هو القيمة الأخيرة للتكرار المطلق.

التمثيل البياني: لتمثيل البيانات الاحصائية الخاصة بالمتغير الاحصائي الكمي المتقطع نلجأ إلى الأعمدة البيانية، كما هو الحال من خلال الشكل التالي:

توزيع 50 أسرة حسب عدد الأطفال



الإجابة على أسئلة المثال السابق:

1_ ما هو عدد الاسرة الذين لديهم 5 أطفال؟

ج1 - هناك 9 أسر لديهم 5 أطفال.

2_ ما هو عدد الأسر الذي عدد أطفالها أكثر من 6 أطفال؟

ج2 - (أكثر من كذا) ← مباشرة من العمود الخاص بالتكرار المتجمع النازل ← هناك 29 أسرة من بين 50 أسرة عدد أطفالها أكثر من 6 أطفال.

3_ ما هو عدد الأسر الذي عدد أطفالها أقل من 4 أطفال؟

ج2- (أقل من كذا) ← مباشرة من العمود الخاص بالتكرار المتجمع الصاعد ← هناك 12 أسرة من بين 50 أسرة عدد أطفالها أقل من 4 أطفال.

4 – ما هو نسبة الأسر الذي عدد أطفالها أقل من 5؟ (مهمة)

لإيجاد النسبة يجب علينا أولا :

- ✓ **ايجاد عدد الأسرة الذي عدد أطفالها أقل من 5**:

(أقل من كذا) ← مباشرة من العمود الخاص بالتكرار المتجمع الصاعد ← هناك 21 أسرة من بين 50 أسرة عدد أطفالها أقل من 5 أطفال.

- ✓ **ثم نقوم بحساب النسبة:**

$$\left.\begin{array}{l} 50 \text{ أسرة} \rightarrow 100\ \% \\ 21 \text{ أسرة} \rightarrow x\ \% \end{array}\right\} \Rightarrow \frac{21 \times 100}{50} = 42\%$$

هناك 42% أسرة من بين 50 أسرة عدد أطفالها أقل من 5 أطفال.

2_ حالة متغير إحصائية مستمرة:

نلجأ إلى فكرة الفئات لتمثيل هذه المعطيات ولتكوين جدول التوزيع التكراري لمتغير كمي مستمر نتبع الخطوات التالية:

- **تحديد المدى** : (RANGE) : هو المجال الذي تنتشر فيه البيانات وهو الفرق بين أكبر قيمة وأصغر قيمة.

$$R = X_{MAX} - X_{MIN}$$

- **تحديد عدد الفئات** :هناك معادلتين شهريتين:

أ-**معادلة STURGERS**:

$$k = 1 + 3{,}32\ Log(N)$$

أو:

$$k = 1 + 1{,}32\ Ln(N)$$

حيث: $N = \sum n_i$ وهي مجموع التكرارات.

ب_ المعادلة **YULE**:

$$k = 2{,}5\sqrt[4]{N} = 2{,}5(N)^{\frac{1}{4}} = 2{,}5(N)^{0{,}25}$$

حيث: $\boldsymbol{N = \sum n_i}$ وهي مجموع التكرارات.

- **تحديد طول الفئة**: يتم تحديد طول الفئة بالعلاقة التالية:

$$l = \frac{R}{k} = \frac{\text{المدى}}{\text{عدد الفئات}}$$

بحيث عند تحديد عدد الفئات يجب مراعاة المتباينة التالية:

$$\text{المدى} \geq \text{الفئا عدد} \times \text{الفئة طول}$$

أي:

$$\mathrm{l} \times \mathrm{k} \geq \mathrm{R}$$

- التأكيد دائما من أن مجموع التكرارات يساوي عدد قيم السلسلة.
- **تحديد مراكز الفئات**:

$$c_i = \frac{l_i + l_{i+1}}{2} = \frac{\text{الحد الأعلى للفئة} + \text{الحد الأدنى للفئة}}{2}$$

وأيضا

الفرق بين مراكز فئتين متتاليتين يساوي طول الفئة.

- **التكرار المتجمع الصاعد**: هو عدد الأفراد أو الوحدات الإحصائية الذين لديهم الكيفية **أقل تماما من الحد الأقصى للفئة**.
- **التكرار المتجمع النازل** : هو العدد الأفراد أو الوحدات الإحصائية الذين لديهم الكيفية **أكبر أو يساوي من الحد الأدنى للفئة**.

مثال : البيانات التالية تمثل الدخل الشهري لـ 50 عامل في إحدى المؤسسات (**الوحدة**: 10^3)

36	45	31	28	41	32	29	23	26	48
32	30	28	33	27	40	31	30	40	35
45	32	37	36	31	35	33	38	39	41
30	38	42	35	33	36	38	39	37	45
.34	24	36	46	30	35	32	50	43	37

المطلوب:- كون جدول التوزيع التكراري، مع تفسير النتائج.

- ما هو عدد العمال الذين يتقاضون أجر أقل من 43 ألف دينار؟
- ما هو نسبة العمال الذين يتقاضون أجر أكثر من 43 ألف دينار؟

الحل:

مجتمع الدراسة : عمال المؤسسة.

الوحدة الإحصائية: عامل واحد

المتغير الإحصائي : الدخل الشهري

النوع : كمي مستمر.

- **تكوين جذول التوزيع التكراري:** لتكوين الجدول الاحصائي نتبع الخطوات التالية:

1- حساب المدى:

$$R = X_{MAX} - X_{MIN} = 50 - 23 = 27$$

2- حساب عدد الفئات : باستخدام المعادلتين الشهريتين:

أ- معادلة STURGERS:

$$k = 1 + 3{,}32\ Log(N) = 1 + 3{,}32(50) = 6{,}6405 \cong 7$$

ب- معادلة YULE:

$$k = 2{,}5\sqrt[4]{N} = 2{,}5(N)^{\frac{1}{4}} = 2{,}5(N)^{0{,}25} = 2{,}5(50^{0{,}25}) = 6{,}647 \cong 7$$

3- تحديد طول الفئة:

$$l = \frac{R}{k} = \frac{\text{المدى}}{\text{عدد الفئات}} = \frac{27}{7} = 3{,}86 \cong 4$$

4- تحديد مركز الفئات:

$$\boldsymbol{c_i = \frac{l_i + l_{i+1}}{2}} = \frac{\text{الحد الأعلى للفئة} + \text{الحد الأدنى للفئة}}{\mathbf{2}}$$

تكوين الجدول الإحصائي

الأجر x_i	عدد العمال n_i	مركز الفئات (c_i)	التكرار النسبي (f_i)	التكرار المئوي $(f_i\%)$	التكرار المتجمع الصاعد$(N\uparrow)$	التكرار المتجمع النازل$(N\downarrow)$
$[\mathbf{23-27}[$	3	25	0,06	6	3	50
$[\mathbf{27-31}[$	8	29	0,16	16	11	47
$[\mathbf{31-35}[$	11	33	0,22	22	22	39
$[\mathbf{35-39}[$	14	37	0,28	28	36	28
$[\mathbf{39-43}[$	7	41	0,14	14	43	14
$[\mathbf{43-47}[$	5	45	0,1	10	48	7
$[\mathbf{47-51}[$	2	49	0,04	4	50	2
$\sum n_i$	50	/	1	100	/	/

- **تفسير بعض قيم الجدول:**

$(\boldsymbol{n_3 = 11})$ ←← لدينا 11 عامل من بين 50 عامل يتقاضون أجرا ما بين 31 إلى أقل من 35 ألف دينار.

$(\boldsymbol{f_i\% = 16\%})$ ←← معناها أن هناك 16 % من العمال يتقاضون أجر يتراوح ما بين 27 إلى أقل من 31 ألف دينار.

$(\boldsymbol{N_i \uparrow= 36})$ ←← لدينا 36 عامل من بين 50 عامل يتقاضون أجر أقل تماما من 39 ألف دينار.

$(\boldsymbol{N_i \downarrow= 28})$ ←← لدينا 28 عامل من بين 50 عامل يتقاضون أجر أكثر أو يساوي 35 ألف دينار.

- **الإجابة على باقي أسئلة المثال:**

1– ما هو عدد العمال الذين يتقاضون أجر أقل من 43 ألف دينار؟

ج1 – (أقل من كذا) ← مباشرة من العمود الخاص بالتكرار المتجمع الصاعد ← هناك 43 عامل من بين 50 عامل يتقاضون أجر أقل من 43 ألف دينار.

2– ما هو نسبة العمال الذين يتقاضون أجر أكثر من 43 ألف دينار؟

ج2– لإيجاد النسبة يجب علينا أولا :

✓ **إيجاد عدد العمال الذين يتقاضون أجر أكثر من 43 ألف دينار:**

(أكثر من كذا) ← مباشرة من العمود الخاص بالتكرار المتجمع النازل ← هناك 7 عمال من بين 50 عامل يتقاضون أجر أكثر من 43 ألف دينار.

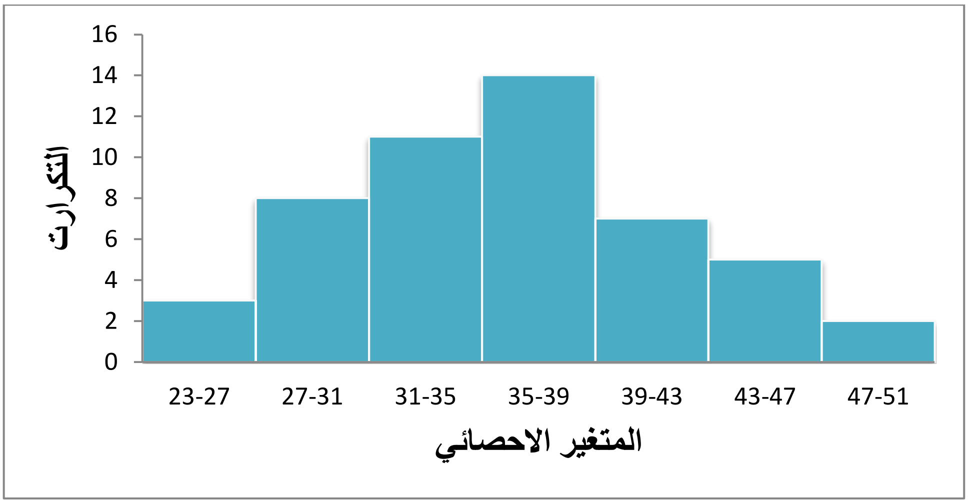
✓ **ثم نقوم بحساب النسبة:**

$$\left.\begin{array}{l} 50 \text{ أسرة} \rightarrow 100\ \% \\ 7 \text{ عمال} \rightarrow x\ \% \end{array}\right\} \Rightarrow \frac{7\ \times 100}{50} = 14\%$$

هناك 14 % من العمال يتقاضون آجر آكثر من 43 ألف دينار.

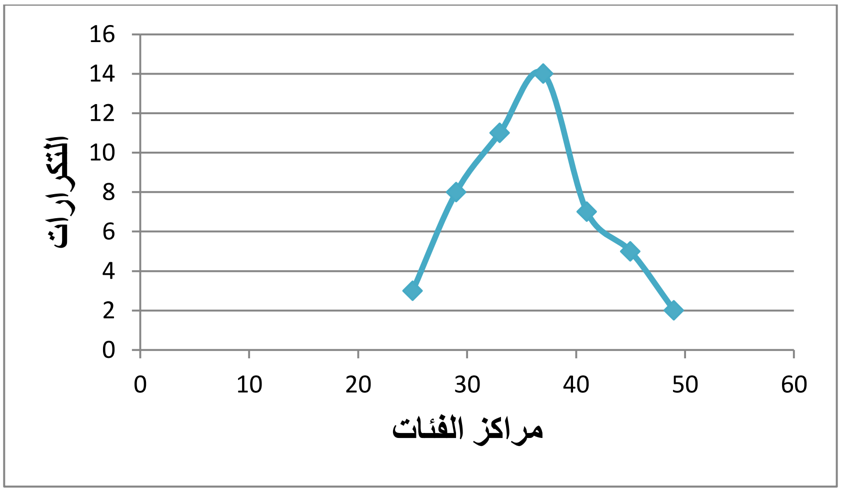
التمثيل البياني: عندما نكون بصدد متغيرة احصائية مستمرة ونريد تمثيلها بيانيا نلجأ إلى المدرج التكراري، الذي نستطيع من خلاله رسم كل من المنحنى والمضلع التكرارين.

المدرج التكراري



المضلع التكراري

المنحنى التكراري



تمارين الفصل:

التمرين الأول:

البيانات التالية تبين المصاريف السنوية لإحدى العائلات:

الميدان	**النقل**	**المواد الغذائية**	**العطل**	**الصحة**	**الكهرباء والغاز**	**اللباس**
المصاريف ب 10^3 دج	**10**	**82**	**8**	**19**	**8**	**28**

- مثل هذه البيانات باستخدام الأعمدة البيانية، وباستخدام شكل الدائرة.

التمرين الثاني:

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

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

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

التمرين الثالث:

الجدول الآتي يبين أوزان 133 طالب:

الوزن (كغ)	[56,58[	[58,60[	[60,62[	[62,64[	[64,66[	[66,68[	[68,70[
عدد الطلبة (n_i)	5	12	18	39	36	15	8

- أرسم المدرج التكراري، والمضلع التكراري.
- مثل بيانيا كل من التكرار المتجمع الصاعد والنازل.

التمرين الرابع:

البيانات التالية تمثل مجموعة من الإحصائيات تخص نقاط عينة من طلبة السنة أولى في مقياس الاقتصاد الجزئي:

19، 17، 15، 03، 07، 02، 12، 13، 19، 14، 19.5، 18.5، 16، 15، 11.5، 05، 08، 11.5، 12، 13، 03.5، 07.5، 13، 16، 17، 00، 13، 11، 02، 04، 13، 02، 08، 19.5، 08، 18.5، 05.5، 09، 10، 04، 08، 03.5، 10، 08، 01، 17.5، 04، 15، 07.

__المطلوب__:

- ما هو عدد عناصر المجموعة الإحصائية (N)، ثم حدد عدد الفئات باستخدام المعادلتين الشهيرتين.

- أنشئ الجدول الإحصائي مبينا كل من أطوال الفئة، مركز الفئة، التكرارات المطلقة، التكرارات النسبية والمئوية، التكرار المتجمع الصاعد والنازل.

- ما هو عدد الطلبة الذين تحصلوا على علامة أقل من 12.

- ما هو عدد الطلبة الذين تحصلوا على العلامة 15 وأكثر.

التمرين الخامس:

البيانات التالية متعلقة بالمداخيل اليومية لـ 50 شخص في إحدى المؤسسات بالدينار الجزائري.

375	370	360	200	250	230	180	180	180	170	120	120	120
350	280	520	520	520	460	110	110	100	390	380	380	375
440	420	420	400	400	400	390	660	640	360	360	360	350
		630	620	620	620	620	640	600	600	540	540	460

- حدد المجتمع الإحصائي والمتغير الإحصائي ونوعه.

- شكل جدول توزيع تكراري من 7 فئات متساوية الطول.

- تحديد التكرارات النسبية والمئوية والتكرارات المتجمعة الصاعدة والنازلة.

- مثل هذه البيانات بواسطة المدرج والمضلع التكراري.

– مثل بيانيا التكرارات المتجمعة الصاعدة والنازلة.

– ما هي نسبة العمال الذين يتقاضون أجرا ما بين 340 إلى أقل من 420 دينار.

– ما هي نسبة العمال الذين يتقاضون أجرا أقل من 340 دينار.

– ما هي نسبة العمال الذين يتقاضون 420 و أكثر.

التمرين السادس:

بفرض أن البيانات التالية تمثل إجمالي ما أنفقه 75 شخص خلال أسبوع (الوحدة :102 دج)

62	72	68	53	73	82	68	78	66	62	65	74	73	67	73
69	74	81	63	63	83	60	79	75	71	79	62	69	97	78
83	75	61	76	65	82	78	75	73	66	75	82	73	84	77
93	73	57	90	60	96	78	79	71	85	75	60	90	71	79
62	88	68	76	83	65	75	87	74	85	91	80	79	89	76

– حدد المجتمع الإحصائي والمتغير الإحصائي ونوعه.

– شكل جدول توزيع تكراري باستخدام طريقة يول.

– تحديد التكرارات النسبية والمئوية والتكرارات المتجمعة الصاعدة والنازلة.

– تحديد عدد الأشخاص الذين تتراوح نفقاتهم بين 65 و 80 دينار.

– تحديد نسبة الأشخاص الذين تزيد نفقاتهم الأسبوعية عن 77 دينار.

الفصل الثالث: مقاييس النزعة المركزية

1- تمهيد؛

2- تعريف مقاييس النزعة المركزية؛

3- مقاييس النزعة المركزية (المنوال- الوسيط- المتوسط الحسابي)

1- تمهيد

إن تلخيص البيانات العددية في جداول احصائية وعرضها في الصور أشكال بيانية يعطي للباحث وصف عام وسريع حول الظاهرة المدروسة، غير ان لهذه الطريقة حدود نذكر منها :

ـ لا يمكن استخدامها في تحليل المعطيات.

ـ لا يمكن استفادة منها في التنبؤ واتخاذ القرارات.

لهذه الاسباب وضعت مقاييس عديدة وصفية تستخدم في التحليل و التنبؤ واتخاذ القرارات تسمى: مقاييس النزعة المركزية.

2- تعريف مقاييس النزعة المركزية

هي تلك المقاييس التي تبحث في تقدير قيمة تتمركز حولها أغلبية القيم، هذه القيمة المتوسطة أو المتمركزة هي رقم واحد نعبر وتمثل جميع بيانات تلك المجموعة.

إذن :إن مقاييس النزعة المركزية لا تحل محل البيانات التفصيلية ولكن تعطي فكرة واضحة عن الظاهرة قيد الدراسة، ولعل من أهم مقاييس النزعة المركزية:

أ_ المنوال

يعتبر أداة من أدوات التحليل يلجأ إليها المحلل اعتمادا على الفئة المهيمنة (يعبر عن المشاهدة أكثر تكرار) ويعتبر المنوال أفضل مقياس لقياس البيانات النوعية ويرمز له بالرمز M_O فكيف يتم حسابه؟

1_ حالة سلسلة إحصائية بسيطة:

مثال: لتكن لدينا السلسلة الاحصائية التالية: 1، 2، 4، 6، 5، 7، 8، 6.

بما أن: المنوال يعبر عن القيمة الأكثر تكرارا، والملاحظ من خلال المثال أن القيمة 6 مكررة مرتين، إذن مباشرة قيمة المنوال هي: $M_O = 6$.

2_ حالة متغير احصائية متقطعة (بيانات مبوبة):

في حالة البيانات المبوبة ونوع المتغير الاحصائي متقطع، فمباشرة قيمة المنوال هو القيمة المقابلة لأكبر تكرار.

مثال: الجدول التالي يبين عدد أيام الغيابات لعمال مؤسسة ما:

عدد الغيابات (المتغير الاحصائي)	0	1	2	3	4	5
عدد العمال (التكرارات)	22	18	15	25	12	8

مباشرة من الجدول أكبر تكرار هو 25 ، **إذن** قيمة المنوال هي: $M_O = 3$.

تفسير قيمة المنوال : أغلبية العمال عدد غياباتها 3 غيابات.

3_ حالة المتغير احصائية مستمرة (بيانات مبوبة):نميز حالتين:

✓ **الحالة الأولى**: **طول الفئة المتساوي**

نقوم في البداية تحديد الفئة المنوالية، ثم حساب قيمة المنوال وذلك باستخدام القانون التالي:

$$M_O = L_1 + a\frac{d_1}{d_1 + d_2}$$

حيث:

(L_1): الحد الأدنى للفئة المنوالية؛

(a): طول الفئة المنوالية؛

(d_1): الفرق بين تكرار الفئة المنوالية والتكرار السابق لها.

(d_2): الفرق بين تكرار الفئة المنوالية وتكرار الفئة اللاحقة لها.

تحديد قيمة المنوال بيانيا: يمكن تحديد المنوال بيانيا وذلك من خلال رسم المدرج التكراري للفئة المنوالية والفئتين السابقة و اللاحقة لها ، حيث نقوم بإيصال بداية المستطيل للفئة المنوالية ببداية المستطيل للفئة اللاحقة لها، ونهاية المستطيل للفئة المنوالية بنهاية المستطيل للفئة السابقة لها، ومن نقطة القاطع نسقط عمودا على المحور الأفقي (محور الفواصل) ونقطة التقاطع مع هذا المحور تعطي قيمة المنوال.

مثال: لإيجاد قيمة المنوال حسابيا وبيانيا نأخذ المثال التالي:

x_i	$[0-3[$	$[3-6[$	$[6-9[$	$[9-12[$	$[12-15[$	$[15-18[$	$[18-21[$
n_i	4	5	7	8	3	1	5

الحل:

1_ إيجاد قيمة المنوال حسابيا:

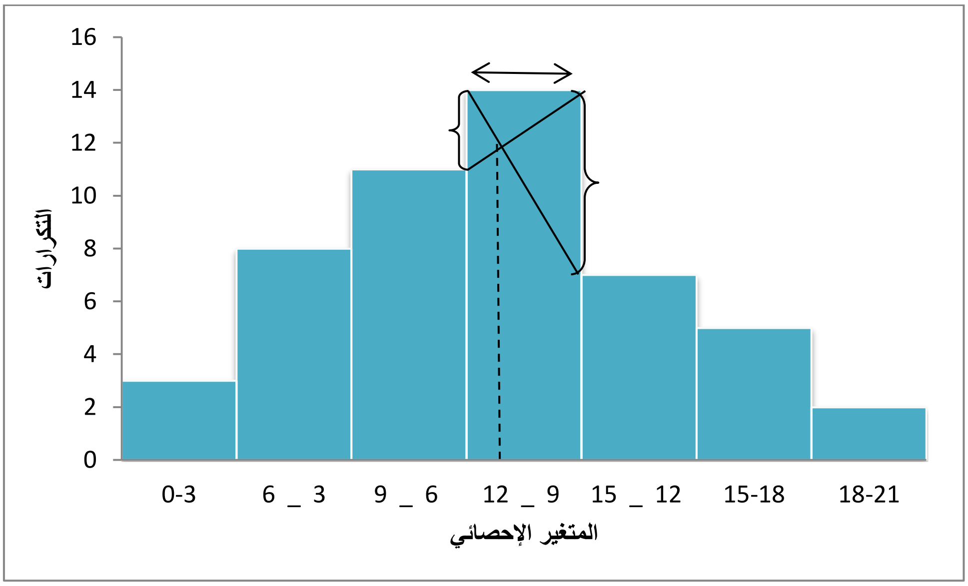
- **بما أن**: طول الفئة متساوي $(a = 3)$ **إذن** مباشرة يتم ايجاد الفئة المنوالية بدون أي تعديلات.
- **الفئة المنوالية** :هي التي تقابل أكبر تكرار:

$$[9 - 12[$$

- **تطبيق القانون**:

$$M_O = L_1 + a\frac{d_1}{d_1 + d_2} = 9 + 3\frac{(8-7)}{(8-7)+(8-3)} = 9{,}5$$

2_ **إيجاد قيمة المنوال بيانيا**:



✓ **الحالة الثانية** :**طول الفئة غير متساوية**

إذا كان طول الفئة غير متساوي يجب علينا أولا تعديل التكرارات باستخدام القانون التالي:

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

$$n_i^* = \frac{n_i}{a} \times a^*$$

ثم يتم استخدام نفس القانون السابق، **إذن** في حالة الفئات الغير المتساوية تستعمل التكرارات المعدلة بدلا من التكرارات العادية.

مثال: أوجد قيمة المنوال حسابيا للمعطيات التالية:

بما أن: طول الفئة غير متساوي، يجب علينا تعديل الفئات كالتالي:

الأجر x_i	عدد العمال n_i	طول الفئة (a)	$\frac{n_i}{a}$	التكرار المعدل (n_i^*)
$[1-2[$	15	1	15	15
$[2-3[$	22	1	22	22
$[3-5[$	30	2	15	15
$[5-8[$	40	3	13,33	13,33
$[8-9[$	13	1	13	13
$\sum n_i$	120	/	/	/

طول الفئة الشائع هو 1.

- **الفئة المنوالية** : هي التي تقابل أكبر تكرار معدل.

$$[2-3[$$

- **تطبيق القانون**:

$$M_O = L_1 + a\frac{d_1}{d_1+d_2} = 2 + 1\frac{(22-15)}{(22-15)+(22-15)} = 2{,}5$$

3- خصائص المنوال :

_أسهل مقاييس النزعة المركزية.

_ لا يتأثر بالقيم المتطرفة وغير قابل للعمليات الجبرية.

_ يعتبر أفضل المتوسطات لوصف الظواهر النوعية.

_ يمكن أن يوجد أكثر من منوال للتوزيع الواحد.

ب_ الوسيط:

هو أحد مقايس النزعة المركزية يؤخذ في عين الاعتبار رتبة القيم، ويعرف بأنه القيمة التي تقع وسط القيم بعد ترتيبها إما تصاعديا أو تنازليا، ويرمز له بالرمز M_e، إذن نصف القيم (50 %) تكون أقل من الوسيط ونصفها الاخر (50 %) أكبر منه، ولتحديد قيمة الوسيط تصادفنا عدة حالات :

1- **إيجاد الوسيط حسابيا:**

✓ <u>**الحالة الأولى**</u>: **حالة سلسلة إحصائية بسيطة:**

الوسيط هو القيمة أو المتغيرة الإحصائية التي رتبته:

$\frac{N+1}{2}$ في حالة سلسلة إحصائية عدد قيمها عدد فردي، و$\frac{N}{2}$ ، $\frac{N}{2}+1$ في حالة سلسلة إحصائية عدد قيمها عدد زوجي.

مثال: أحسب الوسيط للقيم التالية: 112، 6، 4، 3، 5

<u>**لحل التمرين يجب أولا:**</u>

- **ترتيب السلسلة الإحصائية تصاعديا:**

112، 6، 5، 4، 3

- **إيجاد رتبة الوسيط:**

بما أن $N=5$ (فردي) $\leftarrow\leftarrow$ $\frac{N+1}{2}=\frac{5+1}{2}=3$.

إذن: الوسيط هو القيمة الذي رتبته 3 أي $M_e=5$.

مثال آخر: أحسب الوسيط للقيم التالية: - 3، 7، 8، 6، 3، - 1,

- **ترتيب السلسلة الإحصائية تصاعديا:**

-3 ،-1 ،3 ،6 ،8،7

- **إيجاد رتبة الوسيط:**

بما أن $\begin{cases} \frac{N}{2} = \frac{6}{2} = 3 \\ \text{و} \\ \frac{N}{2} + 1 = \frac{6}{2} + 1 = 4 \end{cases}$ (زوجي) $N = 6 \to\to$.

إذن: الوسيط هو القيمة الذي رتبته 3 و4 (متوسط)، أي هو $M_e = \frac{3+6}{2} = 4,5$.

✓ **الحالة الثانية**: **حالة متغيرة إحصائية متقطعة (بيانات مبوبة)**:

مثال: أوجد الوسيط للجدول الإحصائي التالي:

x_i	n_i	التكرار المتجمع الصاعد ($N_i \uparrow$)
1	**5**	**5**
2	**10**	**15**
3	**20**	**35**
4	**30**	**65**
$\sum n_i$	**65**	/

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

الرتبة في حالة البيانات مبوبة دائما هي:

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

حل المثال: لإيجاد قيمة الوسيط نتبع الخطوات التالية:

1_ **الرتبة**:

$$\frac{N}{2} = \frac{\sum n_i}{2} = \frac{65}{2} = 32,5$$

2_ نقوم بإضافة عمود خاص بالتكرار المتجمع الصاعد لإيجاد **الوسيط**، وهو أول قيمة تكرار متجمعها الصاعد **أكبر أو يساوي**:

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

بمعنى: إذا كانت رتبة الوسيط هي إحدى قيم التكرار المتجمع الصاعد فيؤخذ مباشرة القيمة التي تقابل الرتبة.

_ أما في حالة عدم ظهور الرتبة مباشرة في العمود الخاص بالتكرار المتجمع الصاعد، نقوم بحصرها ما بين قيمة تكرار متجمع صاعد لاحق وقيمة تكرار متجمع نازل، وتأخذ القيمة التي تقابل التكرار المتجمع الصاعد اللاحق.

إذن: الوسيط في حالة المتغير الاحصائي المتقطع هو القيمة التي تقابل التكرار المتجمع الصاعد اللاحق:

$$M_e = 3.$$

✓ **الحالة الثالثة: حالة متغيرة إحصائية مستمرة (بيانات مبوبة):**

مثال: يمثل الجدول توزيع عمال مؤسسة ما حسب الأجور:

الأجر x_i	عدد العمال n_i	مركز الفئات (c_i)	التكرار المتجمع الصاعد $(N\uparrow)$	التكرار المتجمع النازل $(N\downarrow)$
$[40-50[$	10	45	10	100
$[50-60[$	15	55	25	90
$[60-70[$	40	65	65	75
$[70-80[$	20	75	85	35
$[80-90[$	15	85	100	15
$\sum n_i$	100	/	/	/

المطلوب: حدد قيمة الوسيط حسابيا وبيانيا، مع تفسير النتيجة.

حل المثال: لإيجاد قيمة الوسيط نتبع الخطوات التالية:

1_ **الرتبة**:

$$\frac{N}{2} = \frac{\sum n_i}{2} = \frac{100}{2} = 50$$

2_ **الفئة الوسيطية**:

✓ إذا كانت رتبة الوسيط هي إحدى قيم التكرار المتجمع الصاعد فيؤخذ مباشرة الفئة التي تقابل الرتبة.

✓ أما في حالة عدم ظهور الرتبة مباشرة، نقوم بحصرها ما بين قيمة التكرار المتجمع الصاعد اللاحق وقيمة التكرار المتجمع النازل، وتأخذ الفئة التي تقابل التكرار المتجمع الصاعد اللاحق، إذن : الفئة الوسيطية هي التي تقابل التكرار المتجمع الصاعد اللاحق.

←← الفئة الوسيطية هي: $[60-70[$

3_ لحساب M_e نقوم أخيرا بتطبيق القانون التالي:

$$M_e = L_1 + a\left(\frac{\frac{\sum n_i}{2} - \text{تكرار متجمع صاعد سابق}}{\text{تكرار الفئة الوسيطية}}\right)$$

حيث:

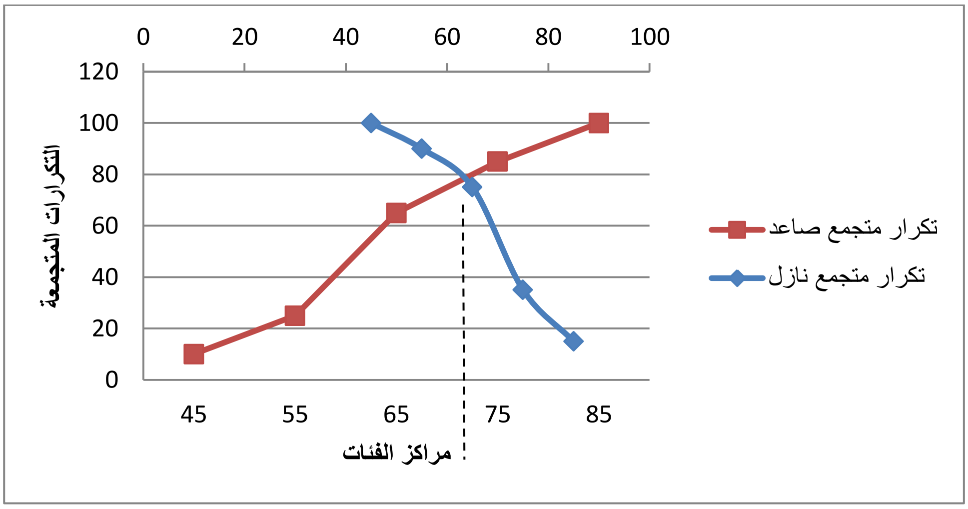
(L_1): الحد الأدنى للفئة الوسيطية؛

(a): طول الفئة الوسيطية.

$$M_e = 60 + 10\left(\frac{50-25}{40}\right) = 66{,}25$$

التفسير: 50% من العمال يتقاضون أجر أقل من **66,25 دينار**، و50% الأخرى أكثر من 66,25 دينار.

2− تحديد قيمة الوسيط بيانيا: يمكننا إيجاد الوسيط بيانيا M_e وهو نقطة تقاطع التكرار المتجمع الصاعد والتكرار المتجمع النازل.



3– خصائص الوسيط:

ـــ يمكن حسابه بيانيا؛

ـــ لا يخضع للعمليات الجبرية؛

ـــ يستعمل في التوزيعات الملتوية؛

ـــ لا يتأثر بالقيم المتطرفة؛

ـــ لا يدخل في حسابه جمع القيم ويتحدد بعدد البيانات وليس بقيمها؛

ـــ يمكن حسابه في حالة الجداول المفتوحة.

ـــلا يتأثر بطول الفئة الغير المتساوية.

3– أشباه الوسيط:

هناك مقاييس أخرى تعمل بنفس مبدأ الوسيط، فإن كان الوسيط هو القيمة التي تقسم مجموع البيانات إلى قسمين متساويين فإننا نستطيع تقسيمها إلى عدة أقسام متساوية :

1ـــ إذا تم تقسيم البيانات إلى 4 أقسام متساوية ←← المقياس يدعى الربيعي "Q"؛

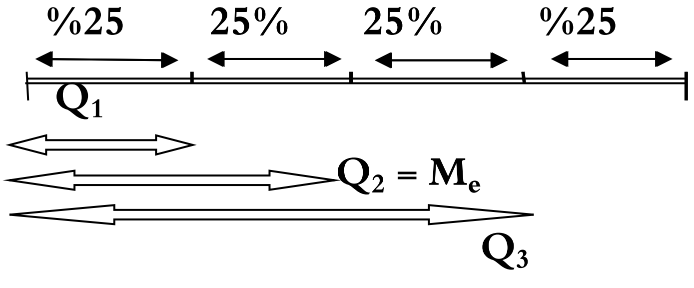
2ـــ إذا تم تقسيم البيانات إلى 10 أقسام متساوية ←← المقياس يدعى العشري "D"

3ـــ إذا تم تقسيم البيانات إلى 100 أقسام متساوية←← المقياس يدعى المئوي أو المئين "P"

أـــ الربعيات: "Q" les Quartiles:

هي القيم التي تقسم البيانات إلى 4 أجزاء متساوية ←←كل جزء يمثل 25% (لدينا 3 ربيعيات)، رتبته من الشكل:

$$\frac{iN}{4} = \frac{i\sum n_i}{4}$$



حيث:

$\mathbf{Q_1}$: هو الربيع الأول ذو الرتبة $\frac{1N}{4} = \frac{\sum n_i}{4}$، وهو القيمة التي تكون فيها 25 % من القيم أقل منها و75 % أكبر منها.

$\mathbf{Q_2}$: هو الربيع الثاني ذو الرتبة $\frac{2N}{4} = \frac{2\sum n_i}{4} = \frac{\sum n_i}{2}$، وهو القيمة التي تكون فيها 50 % من القيم أقل منها و50 % أكبر منها، ومنه نستنتج أن قيمة الربيعي الثاني هو نفسه الوسيط $Q_2 = M_e$

$\mathbf{Q_3}$: هو الربيع الثالث ذو الرتبة $\frac{3N}{4} = \frac{3\sum n_i}{4}$، وهو القيمة التي تكون فيها 75 % من القيم أقل منها و25 % أكبر منها.

ب_ العشريات "D" les Déciles:

هي القيم التي تقسم البيانات إلى 10 أجزاء متساوية ←← كل جزء يمثل 10% (لدينا 9 عشريات)، رتبته من الشكل:

$$\frac{iN}{10} = \frac{i\sum n_i}{10}$$

حيث: $\mathbf{D_3}$ هو العشيري الثالث ذو الرتبة $\frac{\mathbf{3N}}{\mathbf{10}} = \frac{\mathbf{3\sum n_i}}{\mathbf{10}}$، وهو القيمة التي تكون فيها 30 % من القيم أقل منها و70 % أكبر منها، وهكذا مع بقية العشيريات.

ملاحظة: العشيري الخامس هو نفسه الوسيط أي: $D_5 = M_e$

ج _ المئويات: "P" les Percentiles:

هي القيم التي نقسم البيانات إلى 100 قيمة متساوية ←← كل جزء يمثل 1% (لدينا 99مئوي) ورتبته من الشكل:

$$\frac{iN}{100} = \frac{i\sum n_i}{100}$$

حيث: $\mathbf{P_{40}}$ هو المئين أو المئوي الأربعون ذو الرتبة $\frac{40N}{100} = \frac{40\sum n_i}{100} = \frac{\sum n_i}{10}$، وهو القيمة التي تكون فيها 40 % من القيم أقل منها و60 % أكبر منها، وهكذا مع بقية المئويات.

ملاحظة: المئوي أو المئين الخمسين هو نفسه الوسيط والعشيري الخامس أي: $P_{50} = D_5 = M_e$

مثال: الجدول التالي يمثل توزيع 110 مؤسسة حسب أرباحها الشهرية (وحدة الاف دنانير)

الأرباح الشهرية x_i	$[10-20[$	$[20-30[$	$[30-40[$	$[40-50[$	$[50-60[$	$[60-70[$
عدد n_i المؤسسات	18	30	25	17	12	8

المطلوب: أحسب كل من:

ـ الربيع الأول مع التفسير؛

ـ العشير السادس مع التفسير؛

ـ المئين (المئوي) 85 مع التفسير؟

الحل:

تكوين الجدول الاحصائي المساعد للحساب

الأجر x_i	عدد العمال n_i	التكرار المتجمع الصاعد $(N\uparrow)$
$[10-20[$	18	18
$[20-30[$	30	48
$[30-40[$	25	73
$[40-50[$	17	90
$[50-60[$	12	102
$[60-70[$	8	110
$\sum n_i$	110	/

1) **حساب الربيعي الأول Q_1** :

– **الرتبة:**

$$\frac{1N}{4} = \frac{1\sum n_i}{4} = \frac{110}{4} = 27,5$$

- **الفئة الربيعية:** التي تقابل التكرار المتجمع الصاعد اللاحق.

←← الفئة الربيعية هي: $]20 - 30[$

- لحساب Q_1 نقوم أخيرا بتطبيق القانون التالي:

$$Q_1 = L_1 + a\left(\frac{\frac{i\sum n_i}{4} - \text{تكرار متجمع صاعد سابق}}{\text{تكرار الفئة الربيعية}}\right)$$

حيث:

(L_1): الحد الأدنى للفئة الربيعية؛

(a): طول الفئة الربيعية.

$$Q_1 = 20 + 10\left(\frac{27,5 - 18}{30}\right) = 23,166$$

التفسير: 25% من المؤسسات أرباحها أقل من 23,166 ألف دينار، و75% الأخرى أكثر من 23,166 ألف دينار.

2) **حساب العشيري السادس** D_6:

- **الرتبة:**

$$\frac{iN}{10} = \frac{6N}{10} = \frac{6\sum n_i}{10} = \frac{6 \times (110)}{10} = 66$$

- **الفئة العشيرية:** التي تقابل التكرار المتجمع الصاعد اللاحق.

←← الفئة العشيرية هي: $]30 - 40[$

- لحساب D_6 نقوم أخيرا بتطبيق القانون التالي:

$$D_6 = L_1 + a\left(\frac{\frac{i\sum n_i}{10} - \text{تكرار متجمع صاعد سابق}}{\text{تكرار الفئة العشيرية}}\right)$$

حيث:

(L_1): الحد الأدنى للفئة العشيرية؛

(a): طول الفئة العشيرية.

$$D_6 = 30 + 10\left(\frac{66-48}{25}\right) = 37,2$$

<u>التفسير</u>: % 60 من المؤسسات أرباحها أقل من 37,2 ألف دينار، و% 40 الأخرى أكثر من 37,2 ألف دينار.

3) **حساب المئوي 85:** P_{85}

- **الرتبة:**

$$\frac{iN}{100} = \frac{85N}{100} = \frac{85\sum n_i}{100} = \frac{85 \times (110)}{100} = 93,5$$

- **الفئة المئوية:** التي تقابل التكرار المتجمع الصاعد اللاحق.

←← الفئة العشيرية هي: $]50-60[$

- لحساب P_{85} نقوم أخيرا بتطبيق القانون التالي:

$$P_{85} = L_1 + a\left(\frac{\frac{i\sum n_i}{100} - \text{تكرار متجمع صاعد سابق}}{\text{تكرار الفئة المئوية}}\right)$$

حيث:

(L_1): الحد الأدنى للفئة المئوية؛

(a): طول الفئة المئوية.

$$P_{85} = 50 + 10\left(\frac{93,5-90}{12}\right) = 52,916$$

<u>التفسير</u>: % 85 من المؤسسات أرباحها أقل من 52,916 ألف دينار، و% 15 الأخرى أكثر من 52,916 ألف دينار.

ج_ المتوسط الحسابي: la Moyenne Arithmétique:

يعتبر من أشهر وأكثر المتوسطات النزعة المركزية استخداما وشيوعا في الإحصاء، وهو مركز التوازن لأي ظاهرة ونرمز له بالرمز $\bar{X}$، ويحسب كالتالي:

أ_ حالة البيانات الغير المبوبة (حالة سلسلة إحصائية بسيطة):

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \cdots X_n}{N} \rightarrow\rightarrow \bar{X} = \frac{\sum X_i}{N}$$

حيث: $(\boldsymbol{N})$ عدد قيم السلسلة.

مثال: أوجد المتوسط الحسابي للسلسلة الاحصائية التالية:

10، 12، 5، 3، 8

الحل:

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \cdots X_n}{N} = \frac{8 + 3 + 5 + 12 + 10}{5} = 7{,}6$$

ب_ حالة بيانات مبوبة: نستخدم القانون التالي

_ حالة متغيرة احصائية متقطعة :

$$\bar{X} = \frac{\sum n_i X_i}{\sum n_i}$$

_ حالة متغيرة احصائية مستمرة: نقوم بتعويض $\boldsymbol{x_i}$ بمركز الفئة $\boldsymbol{c_i}$

$$\bar{X} = \frac{\sum n_i c_i}{\sum n_i}$$

مثال:

ليكن لدينا الجدول الاحصائي التالي:

c_i $\bar{X}_0$

$\boldsymbol{x_i}$	$\boldsymbol{n_i}$	$\boldsymbol{c_i}$	$\boldsymbol{n_i c_i}$	$\boldsymbol{x_i - \bar{X}_0}$	$\boldsymbol{n_i(X_i - \bar{X}_0)}$
$[200 - 240[$	5	220	1100	– 80	– 400
$[240 - 280[$	6	260	1560	– 40	– 240
$[280 - 320[$	8	300	2400	0	0
$[320 - 360[$	4	400	1360	40	160
$[360 - 400[$	2	380	760	80	160
$\sum n_i$	25	/	7180	/	–320

هناك طريقة أخرى لحساب $\overline{(X)}$ تدعى بالطريقة المختصرة وتستخدم خاصة في حالة القيم الكبيرة للبيانات.

المطلوب: من خلال الجدول السابق أحسب المتوسط الحسابي باستخدام طريقتين مختلفتين؟

- ✓ **الطريقة الأولى**: الطريقة المباشرة أي باستخدام القانون المباشر التالي:

$$\bar{X} = \frac{\sum n_i c_i}{\sum n_i} = \frac{7180}{25} = 287{,}2$$

- ✓ **الطريقة الثانية**: الطريقة الغير المباشرة وتسمى أيضا بالمختصرة ونستخدم القانون التالي:

$$\bar{X} = \bar{X}_0 + \frac{\sum n_i (X_I - \bar{X}_0)}{\sum n_i}$$

يتم فرض **متوسط الفرضي** $(\bar{X}_0)$ ويكون قيمة في الوسط، بالرغم من أن كل القيم تحقق ولكن من الأفضل قيمة في الوسط لتفادي حسابات قيم كبيرة.

إذن: نأخذ $(\mathbf{300 = \bar{X}_0})$ $\leftarrow\leftarrow\leftarrow$ $\bar{X} = 300 - \frac{320}{25} = 300 - 12{,}8 = 287{,}2$

ملاحظة 1: هناك متوسط آخر يدعى بالمتوسط المرجح يستعمل في حالة البيانات التي تكون لها أوزان أو معاملات (تكرارات)، فإذا كان لدينا القيم $X_1, X_2, X_3, \dots \dots X_n$، ولها تكرارات (معاملات) n_1، n_2،n_3، ... ... n_k، فإن المتوسط الحسابي يعطى بالعلاقة التالية:

$$\bar{X} = \frac{n_1 X_1 + n_2 X_2 + n_3 X_3 + \cdots n_k X_n}{n_1 + n_2 + n_3 + \cdots n_k}$$

مثال: الجدول التالي يمثل علامات الامتحان لمقياس الاحصاء للفوج الأول سنة أولى علوم اقتصادية:

5	3	4	6	7	6	**عدد الطلبة (المتغير الاحصائي)**
10	9,25	9	8,5	7,25	5,75	**العلامات (التكرارات)**

المطلوب: أحسب المتوسط الحسابي مع تفسير النتيجة.

الحل:

$$\bar{X} = \frac{6 \times 5{,}75 + 7 \times 7{,}25 + \cdots 5 \times 10}{6 + 7 + \cdots + 5} = \frac{250}{31} = 8{,}06$$

التفسير: متوسط علامات امتحان الطلبة للفوج الأول في مقياس الإحصاء هو 8,06.

ملاحظة 2: إذا كان لدينا مجموعة من العينات داخل مجتمع إحصائي N ولكل عينة $\bar{X}$ خاص بها، فإننا نستطيع حساب المتوسط الحسابي الكلي (الخاص بالمجتمع بأكمله)، فإن المتوسط الحسابي يعطى بالعلاقة التالية:

$$\begin{cases} \bar{X}_1 ، \bar{X}_2 ، \bar{X}_3 ، \dots \dots \bar{X}_n \\ n_1 ، n_2 ، n_3 ، \dots \dots n_k \end{cases} \rightarrow\rightarrow\rightarrow \bar{X}_{\text{الكلي}} = \frac{n_1\bar{X}_1 + n_2\bar{X}_2 + n_3\bar{X}_3 + \cdots n_k\bar{X}_n}{n_1 + n_2 + n_3 \dots + n_k}$$

خواص المتوسط الحسابي:

1_ من أبسط المقاييس وأكثرها استخداما؛

2_ يخضع للعمليات الجبرية؛

3_ يؤخذ بعين الاعتبار جميع قيم الظاهرة؛

4_ مجموع انحرافات القيم عن متوسطها الحسابي يساوي دائما 0 وللتأكيد نبرهن ذلك:

$$\Sigma\left(X_i - \bar{X}\right) = 0 \Rightarrow \sum X_i - \sum \bar{X} = n\bar{X} - n\bar{X} = 0$$

5_ عند إضافة أو طرح عدد ثابت إلى كل قيم التوزيع فإن المتوسط الحسابي يرتفع أو (ينخفض) بمقدار العدد الثابت، بمعنى إذا كان لدينا متغير احصائي آخر Y_i مثلا وكان له علاقة مع المتغير الإحصائي X_i كما يلي: $Y_i = X_i + b$ فيمكننا حساب المتوسط الحسابي للمتغير Y_i مباشرة باستخدام المتوسط الحسابي للمتغير X_i ، إذن: $\bar{Y} = \bar{X} \pm b$ أي دون الحاجة إلى إضافة (b) إلى كل قيم التوزيع؛

6_ يتأثر بالقيم المتطرفة والشاذة، لمعرفة معنى ذلك نعطي المثال التالي: إذا تحصل الطالب على النقاط التالية في 5 مواد: 20، 9، 8، 7، 6، فإن متوسطه:

$$\bar{X} = \frac{X_1 + X_2 + X_3 + \cdots X_n}{N} = \frac{6 + 7 + 8 + 9 + 20}{5} = 10$$

$\leftarrow\leftarrow\leftarrow$ النتيجة أن الطالب ناجح، ولكن في الواقع أن هذا الطالب نجح في مادة واحدة فقط ورسب في أربع مواد.

التفسير: ظهور الطالب ناجح رغم رسوبه في أغلب المواد راجع إلى أن المتوسط الحسابي فد تأثر بالنتيجة الأخيرة والتي تعتبر متطرفة في الكبر.

7_ لا يحسب المتوسط الحسابي في حالة الجداول المفتوحة وحالة البيانات الوصفية (النوعية).

ملاحظات هامة : من بين مقاييس النزعة المركزية يعتبر:

1_ المنوال M_O أحسن مقياس للبيانات النوعية (الوصفية)؛

2_ المتوسط الحسابي $\bar{X}$ أحسن مقياس في حالة الجداول المغلقة؛

3_ M_e الوسيط (الربيعي، العشري ،المئوي) أحسن المقاييس في حالة الجداول المفتوحة؛

4_ هناك علاقة تسمح لنا بالتأكد من صحة حسابتنا تسمى علاقة يول وكندال تعطى بالعلاقة التالية:

العلاقة بين مقاييس النزعة المركزية الثلاثة:

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

المتوسطات الخاصة:

إن المتوسط الحسابي وبالرغم من كثرة استخدامه في مجالات كثيرة، إلا أنه في بعض الظواهر لا يعطينا الوصف الدقيق والصحيح وتكون نتائجه خاطئة، لهذا دعت الضرورة في البحث عن متوسطات أخرى أكثر دقة نذكرها كالتالي:

1_ المتوسط الهندسي: la Moyenne Géométrique:

إذا كانت الظاهرة عبارة عن نسب أو معدلات فإن المتوسط الحسابي لا يوصف هذه الظاهرة الوصف السليم، ولهذا دعت الضرورة إلى إيجاد متوسط آخر يصلح لوصف مثل هذه الظواهر سمي بالمتوسط الهندسي ونرمز له بالرمز "G" وهو واسع الاستخدام في الحياة الاقتصادية ،لأن التركيز يكون غالبا منصب على إيجاد متوسط نسب التغير لبعض الظواهر مثل معدل النمو الناتج ،معدل زيادة الأجور ، معدل النمو السكاني وهكذا، ويعطى بالقانون التالي:

1- **حالة سلسلة احصائية بسيطة (بيانات غير مبوبة):**

$$G = \sqrt[N]{X_1 \times X_2 \times X_3 \dots \dots X_n}$$

حيث: N عدد قيم السلسلة الاحصائية.

2- **حالة البيانات المبوبة (متقطعة +مستمرة):**

$$G = \sqrt[N]{X_1^{n_1} \times X_2^{n_2} \times X_3^{n_3} \dots \dots X_n^{n_k}}$$

حيث: $N = \sum n_i$.

ملاحظة: في حالة متغيرة احصائية مستمرة تعوض X_i بمركز الفئة C_i.

2- **في حالة البيانات الكبيرة فإنه يفضل استخدام طريقة اللوغاريتم:**

- ✓ **حالة سلسلة احصائية بسيطة** (**بيانات غير مبوبة**):

$$logG = \frac{\sum logX_i}{N} \rightarrow\rightarrow G = 10^{logG}$$

حيث: N عدد قيم السلسلة الاحصائية.

- ✓ **حالة البيانات المبوبة** (**متقطعة +مستمرة**):

$$logG = \frac{\sum n_i logX_i}{\sum n_i} \rightarrow\rightarrow G = 10^{logG}$$

حيث: $N = \sum n_i$.

من أهم خواص المتوسط الهندسي:

_ يدخل في حسابه جميع القيم ولكن أقل تأثر بالقيم المتطرفة من المتوسط الحسابي؛

_ لا يمكن حسابه في حالة وجود قيم سالبة أو معدومة؛

_ يستخدم بأكثر واقعية عند وصف الظواهر النسبية.

_ يكون دائما المتوسط الهندسي أقل من المتوسط الحسابي أي: $G < \bar{X}$

2_ المتوسط التوافقي: la Moyenne Harmonique

هو من المقاييس الخاصة التي تستخدم لتحديد معدلات السرعة ومتوسط الأسعار ومتوسط كثافة الأمطار، الكثافة السكانية، ونرمز له بالرمز "H"، ويحسب كما يلي:

1- **حالة سلسلة احصائية بسيطة** (**بيانات غير مبوبة**):

$$H = \frac{N}{\sum \left(\frac{1}{X_i}\right)}$$

حيث: N عدد قيم السلسلة الاحصائية.

2- حالة البيانات المبوبة (متقطعة +مستمرة):

$$H = \frac{\sum n_i}{\sum \left(\frac{n_i}{X_i}\right)}$$

حيث: $N = \sum n_i$.

من أهم خواص المتوسط التوافقي:

_ لا يمكن حسابه في حالة البيانات المعدومة؛

_ يأخذ بعين الاعتبار جميع القيم، وتأثره بالقيم المتطرفة أقل من تأثير المتوسط الحسابي.

_ يكون دائما المتوسط التوافقي أقل من المتوسط الهندسي وأقل من المتوسط الحسابي أي: $H < G < \bar{X}$

3_ المتوسط التربيعي: la Moyenne Quadratique

يرمز له بالرمز Q يستعمل لتفادي القيم السالبة ويحسب كما يلي:

1- **حالة سلسلة احصائية بسيطة (بيانات غير مبوبة):**

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

حيث: N عدد قيم السلسلة الاحصائية.

2- **حالة البيانات المبوبة (متقطعة +مستمرة):**

$$Q^2 = \frac{\sum n_i X_i^2}{\sum n_i} \rightarrow Q = \sqrt{\frac{\sum n_i X_i^2}{\sum n_i}}$$

حيث: $N = \sum n_i$.

خلاصة عامة فيما يتعلق بالمتوسطات: دائما تكون العلاقة بين المتوسطات من الشكل التالي:

$$\boxed{H < G < \bar{X} < Q}$$

أمثلة عن كيفية تطبيق القوانين السابقة:

مثال 1: لتكن لدينا السلسلة الإحصائية التالية: 6، 5، 4، 2، فأوجد عند إذن كل من المتوسط الحسابي والهندسي والتوافقي والتربيعي.

1- **حساب المتوسط الحسابي**:

$$\bar{X} = \frac{X_1 + X_2 + X_3 + X_4}{N} = \frac{2+4+5+6}{4} = 4{,}25$$

2- **حساب المتوسط الهندسي**:

$$G = \sqrt[N]{X_1 \times X_2 \times X_3 \times X_3} = \sqrt[4]{2 \times 4 \times 5 \times 6} = 3{,}94$$

3- **المتوسط التوافقي**:

$$H = \frac{N}{\sum\left(\frac{1}{X_i}\right)} = \frac{4}{\frac{1}{2}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}} = 3{,}58$$

4- **المتوسط التربيعي**:

$$Q^2 = \frac{\sum X_i^2}{N} \rightarrow Q = \sqrt{\frac{\sum X_i^2}{N}} = \sqrt{\frac{2^2+4^2+5^2+6^2}{4}} = 4{,}5$$

نتحقق من النتائج التالية حسب العلاقة:

$$H < G < \bar{X} < Q$$
$$3{,}58 < 3{,}94 < 4{,}25 < 4{,}5$$

مثال 2: أحسب كافة المتوسطات للمعطيات التالية:

x_i	n_i	$n_i x_i$	$log x_i$	$n_i log x_i$	$\frac{n_i}{x_i}$	x_i^2	$n_i x_i^2$
30	5	150	1,477	7,385	0,166	900	4500
25	3	75	1,398	4,194	0,12	625	1875

35	4	140	1,544	6,176	0,114	1225	4900
40	3	120	1,602	4,806	0,075	1600	4800
$\sum n_i$	15	485	/	22,561	0,475	/	16075

1- **حساب المتوسط الحسابي**:

$$\bar{X} = \frac{\sum n_i X_i}{\sum n_i} = \frac{485}{15} = 32,33$$

2- **حساب المتوسط الهندسي**:

$$logG = \frac{\sum n_i logX_i}{\sum n_i} = \frac{22,561}{15} = 1,504$$

$$G = 10^{1,504} = 31,91$$

3- **المتوسط التوافقي**:

$$H = \frac{\sum n_i}{\sum \left(\frac{n_i}{X_i}\right)} = \frac{15}{0,475} = 31,58$$

4- **المتوسط التربيعي**:

$$Q^2 = \frac{\sum n_i X_i^2}{\sum n_i} \rightarrow Q = \sqrt{\frac{\sum n_i X_i^2}{\sum n_i} = \frac{16075}{15}} = 32,74$$

نتحقق من النتائج التالية حسب العلاقة:

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

$$31,58 < 31,91 < 32,33 < 32,74$$

تمارين الفصل:

التمرين الأول:

ليكن التكرار المتجمع الصاعد للظاهرة (X_i) على الشكل التالي 10 30 70 90 100، فإذا علمت أن طول الفئة عبارة عن جداء التكرار الأول في التكرار الأخير وأن الحد الأعلى للفئة الثالثة عبارة عن جداء تكرار الفئة الثانية في تكرار الفئة الرابعة.

المطلوب: - لخص المعطيات في جدول؛

- حساب المتوسط الحسابي.

التمرين الثاني:

ليكن لدينا السلسلة الإحصائية التي تمثل الأعداد التالية: 2، 7، 8، 3، 4، 12

المطلوب:

-أحسب كل من المتوسط الحسابي، المتوسط الهندسي، المتوسط التربيعي والتوافقي؛

- قارن بين مختلف المتوسطات المتحصل عليها.

التمرين الثالث:

إذا كان أجر الساعة لعينة من 25 عامل بأحد المصانع هو كالتالي:

أجر الساعة	35-36	36-37	37-38	38-39	39-40	40-41	41-42	42-43
التكرار	**1**	**2**	**2**	**4**	**5**	**6**	**3**	**2**

- أحسب الوسيط بيانيا وحسابيا؛

- أحسب المتوسط الحسابي بالطريقة المباشرة والمختصرة؛

- ثم حدد المنوال باستخدام طريقتين.

التمرين الرابع:

المعلومات التالية تخص نقط عينة من الطلبة في مادة الاحصاء الوصفي:

16-20	12-16	10-12	8-10	4-8	0-4	الفئات
5	6	10	7	3	1	التكرار المتجمع الصاعد

– أحسب كل من المتوسط الحسابي و المتوسط الهندسي و المتوسط التوافقي؛

– أوجد العلامة التي تحصل عليها أغلب الطلبة.

التمرين الخامس:

– **برهن أن:**

$$\sum \left(X_i - \overline{X}\right) = 0$$

– ليكن لدينا السلسلتين (X_i) و (Z_i) حيث: $\forall i = 1, n$ ، $X_i = aZ_i + b$ ، a, b عبارة عن ثوابت:

برهن أن: $\overline{X} = a\overline{Z} + b$

التمرين السادس:

ليكن لدينا الجدول الإحصائي التالي:

2175	2125	2075	2025	1975	1925	1875	1825	مركز الفئات
6	7	16	25	18	14	9	5	التكرارات المئوية %

– إذا علمت أن مجموع التكرارات $\sum n_i = 200$ ، أوجد حسابيا قيم المتوسط الحسابي، المنوال، الوسيط؛

- أرسم المدرج التكراري للتوزيع واستنتج منه المنوال؛

- أرسم منحنى التكرار المتجمع الصاعد والنازل واستنتج منه الوسيط؛

التمرين السابع:

يمثل الجدول التالي توزيع N على مجموع تكرارات مجهول حسب الأعمار في مصنع ما:

الفئات	$[40, L_2[$	$[L_2, L_3[$	$[L_3, L_4[$	$[L_4, L_5[$	$[L_5, L_6[$	$[L_6, 280[$	$\sum n_i$
(n_i)	n_1	n_2	**34**	**28**	**22**	**6**	**200**

- أوجد قيمة التكرارات إذا علمت أن قيمة العشير الرابع $D_4 = 95$.

التمرين الثامن:

تمت دراسة الأجور لـ 100 عامل حسب الجدول الموالي:

الأجور	$[0,2[$	$[2,4[$	$[4,6[$	$[6,8[$	$[8,10[$	$\sum n_i$
عدد العمال (n_i)	n_1	**5**	**6**	**3**	n_5	**20**

- إذا علمت أن قيمة المتوسط الحسابي $\overline{X} = 5{,}2$ ، أحسب الوسيط والمنوال.

التمرين التاسع:

- ليكن لدينا السلسلتين (X_i) و (Z_i) **حيث:** $\forall i = 1, n$ ، $X_i = aZ_i$ ، $(a \succ 0)$:
  - ما هي العلاقة بين: $\overline{Z}$ و$\overline{X}$ ؛
  - ما هي العلاقة التي تربط بين: G_Z.,G_X .

التمرين العاشر:

ليكن لدينا الجدول الإحصائي التالي:

الفئات	[62,64[	[64,66[	[66,68[	[68,70[	[70,72[	[72,74[	[74,76[
التكرار (n_i)	**3**	**8**	**28**	**34**	**18**	**5**	**4**

المطلوب:

- أحسب كل من المتوسط الحسابي والتوافقي.

التمرين الحادي عشر:

حدد الربيع الأول والثاني والثالث للسلسلة التالية:

7، 2، 25، 15، 7، 13، 12، 5، 8

التمرين الثاني عشر:

لدينا التوزيع التكراري في الجدول الموالي، والذي يمثل أوزان أكياس من الحبوب:

أوزان الأكياس	50 أقل من 60	60 أقل من 70	70 أقل من 80	80 أقل من 90
عدد الأكياس (n_i)	**20**	**40**	**30**	**10**

- أحسب كافة الربيعيات.

التمرين الثالث عشر:

أجور 100 عامل موزعة على الشكل التالي:

الفئات	[100,110[	[110,120[	[120,130[	[130,140[	[140,150[	[150,160[	[160,170[	[170,180[
التكرارات	**8**	**14**	**20**	**27**	**15**	**9**	**5**	**2**

المطلوب:

- حدد مختلف الربيعيات؛
- حدد العشير الثاني والثالث والمئين الثلاثين؛
- ما هو عدد العمال الذي يزيد أجورهم عن **130**؛

الفصل الرابع: مقاييس التشتت

I) تمهيد

عرفنا في الفصل السابق أن مقاييس النزعة المركزية تسمح لنا بالحصول على القيم المتوسطة للبيانات، غير أن هذه المقاييس لا نكفي لوحدها لمعرفة الصفات الإحصائية اللازمة لوصف الظواهر، فقد نجد أن لسلسلتين مختلفتين تفس الوسيط، فيما مدى بيانات السلسلتين مختلف، هذا الفرق في البيانات يجعل من الصروري استخدام مقاييس أخرى مكملة بمقاييس النزعة المركزية، وهي **مقاييس التشتت**.

مثال:

أرباح الشركة x $\rightarrow 10-50-45-65-10 \rightarrow \bar{x} = \frac{\sum X_i}{n} = \frac{180}{5} = 36 \rightarrow$ R = E = 65–10=55

أرباح الشركة y $\rightarrow 40-30-35-40-35$ $\bar{y} = \frac{\sum y_i}{n} = \frac{180}{5} = 36 \rightarrow$ R = E = 40–30 = 10

نلاحظ أن المتوسط الحسابي لكلا السلسلتين متساوي ($\bar{x} = \bar{y} = 36$) ولكن مدى البيانات للسلسلتين مختلف، إضافة إلى أن الفرث بين قيم مختلف وحدات السلسلتين المرتبة مختلف، إذ يتراوح الفرق بين 0 و 35 في السلسلة الأولى، أما السلسلة الثانية فيتراوح بين 0 و 5 $\Leftarrow$ بذلك فإن الاختلاق بين بيانات السلسلة الثانية أقل منه في بيانات السلسلة الأولى و يقال اصطلاحا أن السلسلة الثانية أقل تشتتا من السلسلة الأولى أي (أرباح التركة الثانية أقل تشتت من أرباح التركة الأولى)

إن هذه النتيجة تبين إن مقياس النزعة المركزية لا تكفي لوحدها معرفة الصفات الإحصائية للظواهر واختلاف قيمتها لذلك فإن مقاييس النزعة المركزية لابدا أن تكون مصحوبة بمقاييس أخرى تدرس مدى قرب وتباعد قيم البيانات عن بعض البعض أو عن متوسطها سند بمقاييس التشتت.

II) معنى التشتت:

تشتت بيانات ظاهرة، يقصد بها التفاوت أو اختلاف بين قيم هذه الظاهرة.

أ) تعتبر هذه الظاهرة متجانسة عندما تكون قيمها قريبة (متقاربة) حول بعضها البعض ومن القيمة المتوسطة ونقول في هذه الحالة انها غير مشتتة .

ب) تعتبر بيانات هذه الظاهرة غير متجانسة إذا كانت بياناتها متباعدة عن بعضها البعض وعن القيمة المتوسطة و نقول في هذه الحالة أنها مشتتة.

III) قياس تشتت البيانات:

هناك بعض المقاييس تقيس لنا تقارب وتباعد القيم عن بعضها البعض وهي المدى والانحراف الربيعي وهناك مقاييس أخرى تقيس لنا قرب وبعد القيم عن القيمة المتوسطة وهي الانحراف المتوسط، الانحراف المعياري والتباين.

أ) مقاييس التشتت المطلقة:

1- المدى R أو E:هو الفرق بين أكبر قيمة في السلسلة و أصغر قيمة لها، ويعطى بالعلاقة التالية:

$$\text{المدى} = R = E = X_{max} - X_{min}$$

ما يعيب على هذا المقياس هو تأثره الشديد بالقيم المتطرفة ويعتمد في حسابه على قيمتين فقط وبسبب ذلك لا يستعمل هذا المقياس بشكل واسع كمقياس للتشتت.

2- المدى الربيعي و الانحراف الربيعي (نصف المدى الربيعي):

يحسب المدى الربيعي بالفرق بين الربيعي الثالث والربيعي الأول، ويعطينا فكرة عن المجال الذي تنتشر فيه نصف عدد البيانات ويحسب بالعلاقة التالية :

$$\text{المدى الربيعي} = R_q = Q_3 - Q_1$$

(يعتبر المدى الربيعي كحل لاستبعاد أثر القيم المتطرفة).

أما الانحراف الربيعي فهو **المدى الربيعي مقسوم على 2**: $\text{الإنحراف الربيعي} = E_q = \frac{Q_3 - Q_1}{2}$

ويسمى أيضا بنصف المدى الربيعي.

3- الانحراف المتوسط: ويحسب كالتالي:

ـ سلسلة بسيطة: $E_{\overline{X}} = \frac{\sum |x_i - \overline{x}|}{N}$

ـ بيانات مبوبة: $E_{\overline{X}} = \frac{\sum n_i |x_i - \overline{x}|}{\sum n_i}$

ملاحظة: في حالة متغيرة احصائية مستمرة تعوض X_i بمركز الفئة C_i.

مثال: لدينا السلسلة الإحصائية التالية: 21، 7، 2، 10، 3، 4، 9

- حساب الانحراف المتوسط:

$$E_{\overline{x}} = \frac{\sum|x_i - \overline{x}|}{N}$$

$$\overline{x} = \frac{\sum x_i}{N} = \frac{56}{7} = 8$$

$$E_{\overline{x}} = \frac{|9-8|+|4-8|+|3-8|+|10-8|+|2-8|+|7-8|+|21-8|}{7}$$

$$= \frac{32}{7} = 4{,}57$$

4- التباين والانحراف المعياري: التباين والانحراف المعياري من أكثر مقاييس التشتت استخداما في النواحي التطبيقية، يحسب الانحراف المعياري بالعلاقة التالية:

أ- حالة البيانات المبوبة:

$$\sigma_x = \sqrt{\frac{\sum n_i\left(x_i - \overline{X}\right)^2}{\sum n_i}}$$

$$\sigma_x = \sqrt{\frac{\sum n_i x_i^2}{\sum n_i} - \overline{X}^2} \quad \text{(طريقة مختصرة)}$$

ب- حالة سلسلة بسيطة:

$$\sigma_x = \sqrt{\frac{\sum x_i^2}{N} - \overline{X}^2} \qquad \text{أو} \quad \sigma_x = \sqrt{\frac{\sum\left(x_i - \overline{X}\right)^2}{N}}$$

$$\textbf{(II)} \longleftarrow \textbf{(I)}$$

يمكن لنا تفكيك وتبسيط القانون(**I**) للوصول إلى القانون (**II**)

$$\sigma_x = \sqrt{\frac{\sum (x_i - \overline{X})^2}{N}}$$

$$\sigma_x = \sqrt{\frac{\sum \left(x_i^2 - 2x_i\overline{X} + \overline{X}^2\right)}{N}}$$

$$= \sqrt{\frac{\sum x_i^2}{N} - 2\overline{X}\frac{\sum x_i}{N} + \frac{\sum \overline{X}^2}{N}}$$

$$= \sqrt{\frac{\sum x_i^2}{N} - 2\overline{X}\overline{X} + \frac{N\overline{X}^2}{N}}$$

(مجموع عدد ثابت هو $\sum a = aN$)

$$= \sqrt{\frac{\sum x_i^2}{N} - 2\overline{X}^2 + \overline{X}^2}$$

ومنه:

$$\boxed{= \sqrt{\frac{\sum x_i^2}{N} - \overline{X}^2}} \qquad ...(\mathrm{II})$$

أما التباين فهو **الانحراف المعياري المربع**

$$V(x) = \sigma_x^2 = \frac{\sum n(x_i - \overline{X})^2}{N}$$ (سلسلة بسيطة)

$$V(x) = \sigma_x^2 = \frac{\sum n_i(x_i - \overline{X})^2}{\sum n_i}$$ (بيانات مبوبة)

ملاحظات:

– يعتبر الانحراف المعياري أحسن مقياس لقياس تشتت الظواهر في حالة **تساوي** المتوسطات الحسابية.

- يأخذ الانحراف المعياري نفس وحدة قياس المتغير الأصلي (كلغ، متر، لتر ..الخ) لذلك لا يمكن استخدامه كأساس للمقارنة بين توزيعين لهما وحدات قياس مختلفة.
- إذا كان لدينا مجموعة من العينات داخل مجتمع إحصائي N ولكل عينة $\bar{X}$ خاص بها، فإننا نستطيع حساب التباين الكلي (أو الانحراف الكلي) (الخاص بالمجتمع بأكمله)، فإن التباين الكلي يعطى بالعلاقة التالية:

$$\frac{[n_1V(x_1)+n_2V(x_2)+\cdots n_kV(x_k)]+n_1(\overline{X_1}-\bar{X})^2+n_2(\overline{X_2}-\bar{X})^2+\cdots n_k\left(\overline{X_k}-\bar{X}\right)^2}{n_1+n_2+n_3\ldots+n_k}$$

← نعود للمثال الأول الخاص بأرباح الشركات:

إذا كانت لدينا المعطيات التالية:

شركة A → 10-50-45-65-10

شركة B → 40-30-35-40-35

السؤال: أي من الشركتين أفضل بالنسبة للمستثمرين؟

للإجابة على السؤال يجب علينا معرفة أي الشركتين أكثر استقرارا (أي أقل تشتتا):

أحسن مقياس هو الانحراف المعياري لأن المتوسطات الحسابية متساوية.

$$\overline{X_A}=36$$

$$\overline{X_B}=36$$

$$\sigma_{X_A}=\sqrt{(10-36)^2+(50-36)^2+(45-36)^2+(65-36)^2+(10-36)^2}$$
$$=22{,}23$$

$$\sigma_{X_A}=\sqrt{(40-36)^2+(30-36)^2+(35-36)^2+(40-36)^2+(35-36)^2}$$
$$=3{,}74$$

بما أن:

$$\boxed{\begin{array}{c}\sigma_{X_A}=22,23\\ \sigma_{X_B}=3,74\end{array}} \rightarrow \quad \sigma_{X_B} < \sigma_{X_A}$$

⟵إذن أرباح الشركة B أقل تشتت، ومنه الشركة B أفضل بالنسبة للمستثمرين.

ب_ مقاييس التشتت النسبية:

1_معامل الاختلاف الربيعي:

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

2_ معامل الاختلاف النسبي:

$$CV = \frac{\sigma_X}{\overline{\overline{X}}} \times 100$$

ملاحظة: يستخدم هذا المقياس CV في حالة ما إذا كانت المتوسطات الظواهر غير متساوية وما يميز هذا المقياس أنه غير مرتبط بوحدة القياس، وهذا ما يسمح بقياس ومقارنة تشتت ظاهرتين مختلفتين من حيث وحدات القياس المستعملة.

مثال: إذا كانت لدينا البيانات التالية:

الظاهرة A $\quad \sigma_{X_A} = 3\ ,\ \overline{X_A} = 14$

الظاهرة B $\quad \sigma_{X_B} = 2\ ,\ \overline{X_B} = 15$

بما أن المتوسط احسابي مختلف فإن أحسن مقياس هو معامل الاختلاف CV (وليس الانحراف المعياري)

$$CV_A = \frac{\sigma_{X_A}}{\overline{\overline{X_A}}} \times 100 = \frac{3}{14} \times 100 = 21,43$$

$$CV_B = \frac{\sigma_{X_B}}{\overline{\overline{X_B}}} \times 100 = \frac{2}{15} \times 100 = 13,33$$

بما أن: $\quad CV_B < CV_A$

إذن: الظاهرة B أقل تشتتا من الظاهرة A.

ملاحظات هامة :

من بين مقاييس النزعة المركزية يعتبر: أحسن مقياس التشتت إذا كانت:

الجداول المفتوحة: الانحراف الربيعي (نصف المدى الربيعي) ومعامل الاختلاف الربيعي.

الجداول المغلقة: الانحراف المعياري، التباين، معامل الاختلاف النسبي.

تمرين شامل حول مقاييس التشتت (حالة البيانات المبوبة)

المطلوب: أحسب مختلف مقاييس التشتت المطلقة والنسبية

X_i	n_i	c_i	ت.م صاعد ↗	$\lvert X_i-\bar{X}\rvert$	n_ix_i	$n_i\lvert X_i-\bar{X}\rvert$	$(X_i-\bar{X})$	$n_i(X_i-\bar{X})$	X_i^2	$n_iX_i^2$
$[0-20[$	2	10	2 $\frac{n}{4}$	45	20	90	2025	4050	100	200
$[20-40[$	5	30	7	25	150	125	625	3125	900	4500
$[40-60[$	3	50	10 $\frac{3n}{4}$	5	150	150	25	75	2500	7500
$[60-80[$	6	70	16	15	420	90	225	1350	4900	29400
$[80-100[$	4	90	20	35	360	140	1225	4900	8100	32400
$\sum n_i$	20	/	/	/	1100	460	/	13500	/	74000

1- حساب الانحراف الربيعي:

$$E_Q=\frac{\text{المدى الربيعي}}{2}=\frac{Q_3-Q_1}{2}$$

حساب Q_1:

الرتبة: $\frac{n}{4}=\frac{20}{4}=5$

الفئة الربيعية: $[20-40[$

$$Q_1 = 20 + 20\left(\frac{5-2}{54}\right) = 32$$

حساب Q_3:

الرتبة: $\frac{3n}{4} = \frac{3(20)}{4} = 15$

الفئة الربيعية: $[60 - 80[$

$$Q_3 = 60 + 20\left(\frac{15-10}{6}\right) = 60 + 20{,}66$$

$$Q_3 = 76{,}66$$

← إذن: **الانحراف** الربيعي:

$$E_Q = \frac{Q_3 - Q_1}{2} = \frac{76{,}66 - 32}{2} = \boxed{\mathbf{22,33}}$$

2- حساب الانحراف المتوسط:

$$E_{\overline{X}} = \frac{\sum n_i \left|X_i - \overline{X}\right|}{\sum n_i}$$

$$\overline{X} = \frac{\sum n_i x_i}{\sum n_i} = \frac{1100}{20} = 55$$

$$E_{\overline{X}} = \frac{\sum n_i \left|X_i - \overline{X}\right|}{\sum n_i} = \frac{460}{20} = \boxed{\mathbf{23}}$$

3- حساب الانحراف المعياري:

ط1) ← $\sigma_X = \sqrt{\frac{\sum n_i (X_i - \overline{X})^2}{\sum n_i}} = \sqrt{\frac{13500}{20}} = \sqrt{675}$

$$\boxed{\sigma_X = \mathbf{25,98}}$$

ط2) ←

$$\sigma_x = \sqrt{\frac{\sum n_i x_i^2}{\sum n_i} - \overline{X}^2}$$

$$\sigma_x = \sqrt{\frac{74000}{20} - (55)^2} = \sqrt{3700 - 3025} = \mathbf{25,98}$$

4- حساب التباين:

$$V(x) = \sigma_x^2 = \frac{\sum n_i \left(x_i - \overline{X}\right)^2}{\sum n_i} = \left(\sqrt{675}\right)^2 = \mathbf{675}$$

5- حساب معامل الاختلاف النسبي:

$$CV = \frac{\sigma_X}{\overline{X}} \times 100 = \frac{25,98}{55} \times 100 = \mathbf{47,24}$$

6- حساب معامل الاختلاف الربيعي:

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

حساب Q_2:

الرتبة: $\frac{n}{2} = \frac{20}{2} = 10$

الفئة الربيعية: $[40 - 60[$

$$Q_2 = 40 + 20\left(\frac{10 - 7}{3}\right) = 60$$

إذن:

$$CVQ = \frac{Q_3 - Q_1}{Q_2} \times 100 = \frac{76,66 - 32}{60} \times 100 = \mathbf{74,43\%}$$

تمارين الفصل:

التمرين الأول:

الجدول الآتي يبين أرباح الشركتين X، Y لفترة ما بملايين الدينارات، أي الشركتين أفضل في نظرك ولماذا؟

10	**65**	**45**	**50**	**10**	**الشركة** X
35	**40**	**35**	**30**	**40**	**الشركة** Y

التمرين الثاني:

لدينا التوزيع التكراري في الجدول الموالي:

الفئات	[52,5 – 57,5[	[57,5 – 62,5[	[62,5 – 67,5[	[67,5 – 72,5[	[72,5 – 77,5[	[77,5 – 82,5[	[82,5 – 87,5[
التكرار (n_i)	**10**	**24**	**50**	**58**	**36**	**14**	**8**

المطلوب:

- أحسب كل من المتوسط الحسابي، الانحراف المتوسط ، الانحراف المعياري، التباين؛
- حدد المدى الربيعي، الانحراف الربيعي.
- حدد العشير الخامس تم السادس.

التمرين الثالث:

لدينا التوزيع التكراري في الجدول الموالي:

الفئات	**1900-1950**	**1950-2000**	**2000-2050**	**2050-2100**	**2100-2150**	**2150-2200**	**2200-2250**
التكرار (n_i)	**152**	**244**	**363**	**502**	**363**	**244**	**152**

المطلوب:

- أحسب المتوسط الحسابي، التباين؛
- حدد مقاييس التشتت النسبي: (معامل الاختلاف النسبي، معامل الاختلاف الربيعي).

التمرين الرابع:

- برهن أن: $\sigma_x = \sqrt{\left(\frac{\Sigma ni.X_i^{\ 2}}{\Sigma ni} - \overline{X}^2\right)}$
- برهن على أن العلاقة التالية صحيحة:

$$\frac{\sum n_i\left(x_i - b^2\right)}{\sum n_i} = \sigma_x^{\ 2} + \left(\overline{X} - b^2\right)$$

التمرين الخامس:

ليكن لدينا السلسلتين $\left(X_i\right)$ و $\left(Z_i\right)$ **حيث:** $\forall i = 1, n$ ، $X_i = aZ_i + b$ ، $(a \succ 0)$ برهن أن:

$$\sigma_x^2 = a^2\sigma_z^2$$

التمرين السادس

ليكن لدينا مجموعتين من نفس المجتمع، مجموعة حجمها 10 ومتوسطها يقدر بـ 16 وانحرافها المعياري 4، ومجموعة ثانية حجمها 9 ومتوسطها 15 وانحرافها المعياري 3.

المطلوب:

- أحسب المتوسط الحسابي الكلي؛
- أحسب التباين الكلي؛
- أي المجموعتين أقل تشتت؟

التمرين السابع

- تمت دراسة الأجور لـ 100 عامل حسب الجدول الموالي:

الأجور	[10,15[	[15,20[	[20,25[	[25,30[	[30,35[	[35,40[
عدد العمال $\left(n_i\right)$	n_1	**15**	**18**	**25**	**20**	n_6

- أوجد قيم التكرارات n_1 و n_6 إذا علمت أن قيمة المتوسط الحسابي $\overline{X} = 25{,}8$ **.**

- حدد كل من معامل الاختلاف النسبي ومعامل الاختلاف الربيعي.

التمرين الثامن:

ليكن لدينا الجدول التالي:

الفئات	$[0,2[$	$[4,2[$	$[4,6[$	$[6,8[$	$[8,10[$	$\sum n_i$
(n_i)	n_1	**5**	**6**	n_4	n_5	**20**

- حدد قيمة الوسيط إذا علمت أن قيمة المتوسط الحسابي $\overline{X} = 4,7$ والتباين $\sigma^2 = 5,71$.

الفصل الخامس: مقاييس الشكل (تحديد شكل التوزيع)

I) تمهيد:

إذا كانت مقاييس النزعة المركزية ومقاييس التشتت تسمح لنا بتلخيص بيانات أي ظاهرة في صورة أرقام، تعطي فكرة عن خصائص توزيع هذه البيانات ودرجة تجانسها أو اختلافها، فإن هذا الوصف تبقى تنقصه الدقة المطلوبة للتعرف على خواص التوزيع ، خاصة فيما يخص انتشار البيانات على المنحنى البياني، الممثل لها من حيث التوائه وتفلطحه عن الوضع الطبيعي ،لذلك دعت الحاجة إلى استخدام مقاييس أخرى لتحقيق هذا الغرض سميت بمقاييس الشكل (التفلطح والالتواء). ولكن هذه المقاييس تعتمد على العزوم البسيطة والمركزية في حسابها، وعليه لابد لنا أولا من إعطاء علاقات العزوم:

1- العزوم: قد تكون العزوم حول نقطة الأصل أو حول المتوسط او حول أي نقطة معينة، أما رتبة العزم فتتحدد بدرجة القوة (الأس) التي ترفع إليها القيم، أو انحرافاتها عن المتوسط الحسابي ونميز نوعان:

1.1_ العزوم البسيطة: إذا كانت الظاهرة (X) تأخذ القيم $X_1, X_2, X_3, ...X_n$ ، فإن العزم البسيط من الدرجة (r) لهذا المتغير يعطى بالعلاقة التالية:

- **حالة السلسلة إحصائية البسيطة:**

$$m_r = \frac{\Sigma X_i^{\ r}}{N}$$

حيث: (N) عدد قيم السلسلة الاحصائية، و(r) تمثل درجة الأس.

- **حالة بيانات احصائية مبوبة (متغيرة احصائية متقطعة ومستمرة):**

$$m_r = \frac{\Sigma n_i x_i^r}{\Sigma n_i}$$

حيث: $N = \Sigma n_i$ ، و(r) تمثل درجة الأس.

مثال: إذا كانت لدينا القيم التالية: 8، 5، 2، 1، فأوجد العزم الأول والثاني حول نقطة الأصل.

- **العزم الأول هو**: $m_1 = \frac{\Sigma X_i^{\ r}}{N} = \frac{\Sigma X_i^1}{4} = \frac{1+2+5+8}{4} = 4$
- **العزم الثاني هو**: $m_2 = \frac{\Sigma X_i^{\ r}}{N} = \frac{\Sigma X_i^2}{4} = \frac{1^2+2^2+5^2+8^2}{4} = 23{,}5$

2.1- العزوم المركزية : إذا كانت الظاهرة (X) تأخذ القيم $X_1, X_2, X_3, ...X_n$ ، وكان $(\overline{X})$ متوسطها الحسابي، فإن العزم المركزي من الدرجة (r) حول المتوسط الحسابي لهذا المتغير يعطى بالعلاقة التالية:

- **حالة السلسلة إحصائية البسيطة:**

$$M_r = \frac{\Sigma\left(X_i - \overline{X}\right)^r}{N}$$

حيث: (N) عدد قيم السلسلة الاحصائية، و(r) تمثل درجة الأس.

- **حالة بيانات احصائية مبوبة (متغيرة احصائية متقطعة ومستمرة):**

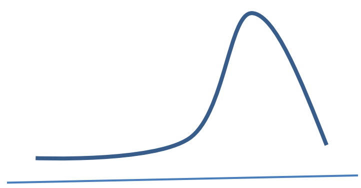
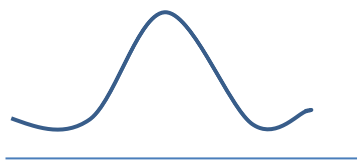
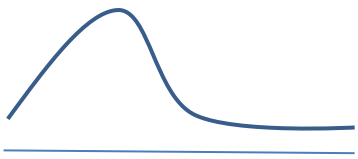
$$M_r = \frac{\Sigma n_i\left(X_i - \overline{X}\right)^r}{\Sigma n_i}$$

حيث: $N = \Sigma n_i$ ، و(r) تمثل درجة الأس.

ملاحظة: يمكن تعريف العزم المركزي حول أي نقطة أخرى خلاف المتوسط الحسابي وذلك من خلال استبدال المتوسط الحسابي $(\overline{X})$ بقيمة ذلك المقياس.

II) الالتواء:

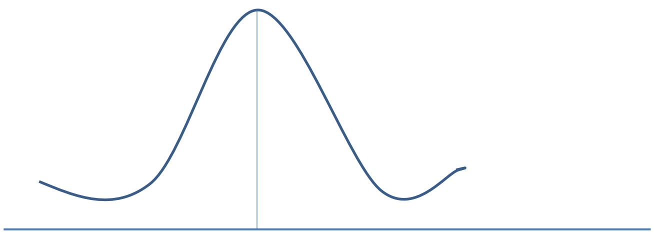
يعبّر الالتواء عن درجة توزيع البيانات حول نقطة التمركز فيها، فوجود الالتواء دليل على انعدام الانتظام في التوزيع، ويمكن معرفة طبيعة أي توزيع بمجرد النظر إلى منحنى التوزيع الذي يأخذ الأشكال التالية:

		
منحنى ملتوي جهة اليسار	**منحنى متماثل (متناظر)**	**منحنى ملتوي جهة اليمين**

ويمكن استخدام المتوسط الحسابي والوسيط والمنوال في وصف شكل منحنى التوزيع التكراري، والذي يعبر عن شكل توزيع البيانات، فالعلاقة بين مقاييس النزعة المركزية (متوسط الحسابي، الوسيط، المنوال) لبيانات واحدة تربطهم إحدى العلاقات التالية:

أ- **منحنى التوزيع التكراري المتماثل (المتناظر)**: المنحنى يكون متماثل وله قيمة واحدة وشكله يشبه الجرس، وفي هذه الحالة يكون:

$$\boxed{\overline{X} = M_e = M_O}$$

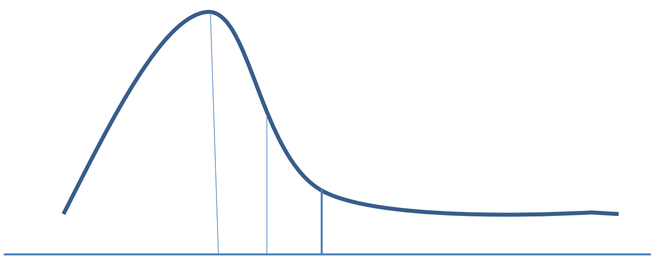


$$\overline{X} = M_e = M_O$$

منحنى توزيع تكراري متماثل (المتناظر)

ب- **منحنى التوزيع التكراري ملتوي (مائل) جهة اليمين (موجب الالتواء)**: عندما يكون :

$$\boxed{M_O \prec M_e \prec \overline{X}}$$

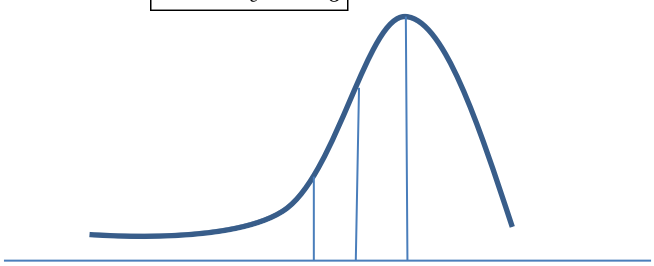


M_O M_e $\overline{X}$

منحنى توزيع تكراري ملتوي جهة اليمين (موجب الالتواء)

ج- **التوزيع التكراري ملتوي (مائل) جهة اليسار (سالب الالتواء)**: عندما يكون :

$$\boxed{\overline{X} \prec M_e \prec M_O}$$



$\overline{X}$ M_e M_O

منحنى توزيع تكراري ملتوي جهة اليسار (سالب الالتواء)

ويمكن استخدام بعض المقاييس لقياس الالتواء:

1- معامل بيرسون للالتواء (P):

أ- **معامل بيرسون الاول**: يعطى معامل بيرسون للالتواء بالعلاقة التالية:

$$P_1 = \frac{\overline{X} - M_O}{\sigma_X}$$

حيث: σ_X : الانحراف المعياري.

ويمكن تحديد شكل التوزيع كما يلي:

- ✓ إذا كان $P_1 = 0 \longrightarrow\longrightarrow$ يكون منحنى التوزيع متماثل (متناظر).
- ✓ إذا كان $P_1 \succ 0 \longrightarrow\longrightarrow$ يكون منحنى التوزيع ملتوي جهة اليمين (موجب الالتواء).
- ✓ إذا كان $P_1 \prec 0 \longrightarrow\longrightarrow$ يكون منحنى التوزيع ملتوي جهة اليسار (سالب الالتواء).

ب- **معامل بيرسون الثاني**: يعطى معامل بيرسون للالتواء بالعلاقة التالية:

$$P_1 = \frac{3.\left(\overline{X} - M_e\right)}{\sigma_X}$$

ويمكن تحديد شكل التوزيع كما يلي:

- ✓ إذا كان $P_2 = 0 \longrightarrow\longrightarrow$ يكون منحنى التوزيع متماثل (متناظر).
- ✓ إذا كان $P_2 \succ 0 \longrightarrow\longrightarrow$ يكون منحنى التوزيع ملتوي جهة اليمين (موجب الالتواء).
- ✓ إذا كان $P_2 \prec 0 \longrightarrow\longrightarrow$ يكون منحنى التوزيع ملتوي جهة اليسار (سالب الالتواء).

ج- **معامل بيرسون للالتواء بدلالة العزوم** (β_1): يعطى هذا المعامل بالعلاقة التالية:

$$\beta_1 = \frac{(M_3)^2}{(M_2)^3}$$

حيث: M_2 : العزم المركزي من الدرجة الثانية $(r = 2)$؛

M_3 : العزم المركزي من الدرجة الثالثة $(r = 3)$.

ملاحظة:

نلاحظ أن (β_1) دائما موجب (البسط والمقام دائما موجبين) لهذا لا يعطينا هذا المعامل فكرة عن الالتواء، أما إذا كان $(\beta_1 = 0)$ فالتوزيع هنا يكون متماثل (متناظر).

2- معامل فيشر للالتواء (F_1): يعطى معامل فيشر للالتواء بالعلاقة التالية:

$$F_1 = \frac{M_3}{\sigma^3}$$

حيث: M_3 : العزم المركزي من الدرجة الثالثة $(r=3)$؛

σ_X : الانحراف المعياري.

ملاحظة:

يعتمد معامل فيشر على العزم المركزي من الدرجة الثالثة بشكل خاص في حساب الالتواء، لأن قيمته في حالة التوزيع المتناظر يساوي دائما صفر $(M_3=0)$، وبناءا على هذه القيمة يتم تحديد شكل التوزيع:

- ✓ إذا كان $F_1=0 \longleftarrow\longleftarrow$ يكون منحنى التوزيع متماثل (متناظر).
- ✓ إذا كان $F_1 \succ 0 \longleftarrow\longleftarrow$ يكون منحنى التوزيع ملتوي جهة اليمين (موجب الالتواء).
- ✓ إذا كان $F_1 \prec 0 \longleftarrow\longleftarrow$ يكون منحنى التوزيع ملتوي جهة اليسار (سالب الالتواء).

3- معامل يول للالتواء (C_Y):

يستخدم هذا المعامل في حالة الجداول المفتوحة ويسمى أيضا بمعامل الالتواء الربيعي وهو معطى بالعلاقة التالية:

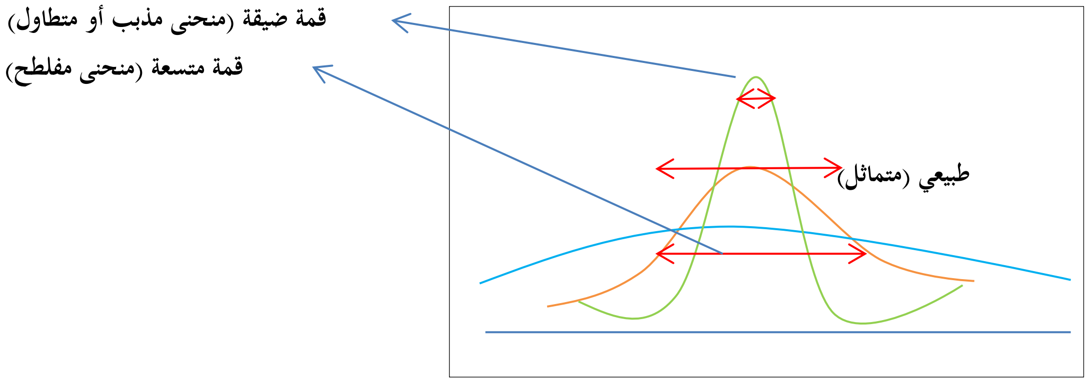
$$C_Y=\frac{(Q_3-Q_2)-(Q_2-Q_1)}{Q_3-Q_1}=\frac{Q_3-2Q_2+Q_1}{Q_3-Q_1}$$

ويتم تحديد شكل التوزيع كما يلي:

- ✓ إذا كان $C_Y=0 \longleftarrow\longleftarrow$ يكون منحنى التوزيع متماثل (متناظر).
- ✓ إذا كان $C_Y \succ 0 \longleftarrow\longleftarrow$ يكون منحنى التوزيع ملتوي جهة اليمين (موجب الالتواء).
- ✓ إذا كان $C_Y \prec 0 \longleftarrow\longleftarrow$ يكون منحنى التوزيع ملتوي جهة اليسار (سالب الالتواء).

III) التفلطح (التفرطح):

التفلطح هو قياس درجة علو قمة التوزيع بالنسبة للتوزيع الطبيعي، أي يقصد به مدى اتساع أو ضيق قمة منحنى التوزيع، فكلما كان الشكل أكثر ارتفاعا من الشكل الطبيعي، نقول أن الشكل **مذبب (متطاول)** معنى أن قمة التوزيع تكون ضيقة، أما إذا كان أقل ارتفاع من الشكل الطبيعي، نقول أن الشكل **مفلطح** (قمة التوزيع متسعة) والتمثيل البياني يبين ذلك:



ويقاس التفلطح بأحد المعاملات التالية:

1- معامل بيرسون للتفلطح (β_2): يعطى بالعلاقة التالية:

$$\beta_2 = \frac{(M_4)}{(M_2)^2} = \frac{M_4}{\sigma^4}$$

حيث: M_2 : العزم المركزي من الدرجة الثانية $(r=2)$؛

M_4 : العزم المركزي من الدرجة الرابعة $(r=4)$؛

σ_X : الانحراف المعياري.

ويتم تحديد شكل التوزيع كما يلي:

- ✓ إذا كان $\beta_2 = 3$ ⟵⟵ يكون منحنى التوزيع طبيعي.
- ✓ إذا كان $\beta_2 \succ 3$ ⟵⟵ يكون منحنى التوزيع مذبب (متطاول).
- ✓ إذا كان $\beta_2 \prec 3$ ⟵⟵ يكون منحنى التوزيع مفلطح.

2- معامل فيشر للتفلطح (F_1): وهو عبارة عن معامل بيرسون مطروح منه القيمة 3، ويعطى معامل فيشر للتفلطح بالعلاقة التالية :

$$F_2 = \beta_2 - 3 = \frac{(M_4)}{(M_2)^2} - 3 = \frac{M_4}{\sigma^4} - 3$$

ويتم تحديد شكل التوزيع كما يلي:

- ✓ إذا كان $F_2 = 0$ ⟵⟵ يكون منحنى التوزيع طبيعي.
- ✓ إذا كان $F_2 \succ 0$ ⟵⟵ يكون منحنى التوزيع مذبب (متطاول).
- ✓ إذا كان $F_2 \prec 0$ ⟵⟵ يكون منحنى التوزيع مفلطح.

ملاحظة:

يعتمد كل من فيشر وبيرسون على العزوم المركزية من الدرجة الرابعة في حساب التفلطح بشكل خاص، لأن قيمته في حالة التوزيع المتناظر يساوي دائما 3، أي $(M_4 = 3)$.

3- معامل التفلطح المئوي: يستخدم هذا المعامل في حالة جداول التوزيع المفتوحة وهو معطى بالعلاقة التالية:

$$\beta = \frac{Q_3 - Q_1}{2(C_{90} - C_{10})} = \frac{Q_3 - Q_1}{2(D_9 - D_1)}$$

حيث: D_1 و D_9 : العشيري الأول والعشيري التاسع على التوالي؛

C_{10} و C_{90} : المئوي العاشر والمئوي التسعون على التوالي.

ويتم تحديد شكل التوزيع كما يلي:

- ✓ إذا كان $\beta = 0,263$ ⟵⟵ يكون منحنى التوزيع طبيعي.

- ✓ إذا كان $\beta \succ 0,263$ $\longleftarrow\longleftarrow$ يكون منحنى التوزيع مذبب (متطاول).
- ✓ إذا كان $\beta \prec 0,263$ $\longleftarrow\longleftarrow$ يكون منحنى التوزيع مفلطح.

تمارين الفصل:

التمرين الأول:

المعلومات تخص أجور عينة من العمال تتكون من 42 عامل بأحد المصانع هو كالتالي:

(وحدة القياس 10دج)

الفئات	35-50	50-65	65-80	80-95	95-110
العمال	7	8	12	10	5

- ما هو عدد العمال الذين يقل أجرهم عن 650 دج.
- باستخدام المقارنة بين مقاييس النزعة المركزية استنتج طبيعة أو شكل التوزيع الإحصائي.
- أدرس معامل التفلطح (فيشر وبيرسون)، ماذا تستنتج؟

التمرين الثاني:

المعلومات تخص أجور عينة من العمال تتكون من 96 عامل بأحد المصانع هو كالتالي:

(وحدة القياس 10دج)

الفئات	500-550	550-600	600 -650	650 -700	700 -750	750 -800	800 -850
العمال	6	10	18	28	18	10	6

- ما هو عدد العمال الذين يزيد أجرهم عن 650 دج.
- أوجد معامل يول للالتواء، ماذا تستنتج؟
- أوجد معامل التفلطح المئوي، ماذا تستنتج.

التمرين الثالث:

المعلومات تخص أجور عينة من العمال تتكون من 40 عامل بأحد المصانع هو كالتالي:

(وحدة القياس 10دج)

الفئات	100-150	150-200	200-250	250-300	300-350
العمال	7	8	10	10	5

- ما هو عدد العمال الذين يقل أجرهم عن 2500 دج.
- أدرس التواء وتفلطح هذا التوزيع باستخدام قوانين بيرسون.

التمرين الرابع:

المعلومات تخص أجور عينة من العمال تتكون من 60 عامل بأحد المصانع هو كالتالي:

(وحدة القياس 10^3 دج)

الأجر	7-17	17-27	27-37	37-47	47-57	57-67	67-77
التكرار	7	8	9	12	9	8	7

- ما هو عدد العمال الذين يقل أجرهم عن 57000 دج.
- أوجد معامل فيشر للالتواء، ماذا تستنتج.

الفصل السابع: مقاييس التمركز

المفهوم

مفهوم التمركز[1] يشبه على العموم مفهوم التشتت، فالتشتت يدرس انحراف القيم عن القيمة المتوسطة للبيانات الاحصائية، بينما التمركز فهو يدرس مدى عدالة توزيع البيانات وانتشارها، مثل دراسة مدى عدالة توزيع الدخول على الأفراد (حصول الأفراد على قيم متساوية من الدخل)، أو توزيع السكان على وحدة المساحة مثلا أي مدى ميل السكان إلى التمركز في منطقة واحدة داخل دولة ما، ويكون توزيع السكان منظم من الناحية الاحصائية هنا كلما كانت نسبة السكان **مساوية** لنسبة المساحة، وبعبارة أخرى عندما يتوزعون السكان بالتساوي على مساحة الدولة.

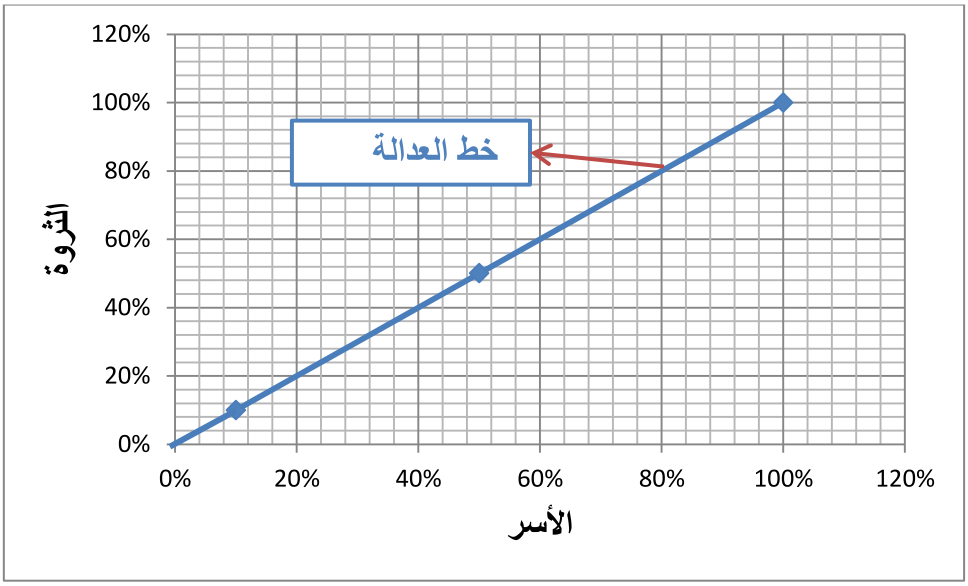
ولقياس التفاوت أو عدم التساوي في توزيع ظاهرة ما، هناك العديد من الطرق أهمها الطريقة البيانية والطريقة الحسابية:

1- منحنى لورنز لقياس التمركز (LORENZ) (الطريقة البيانية)

ويسمى أيضا بمنحنى تركز الثروات، ويعرف بأنه التوزيع التراكمي لمتغير ما، ويعد هذا المنحنى من أكثر الأشكال البيانية استخداما للتعبير عن حجم التفاوت في توزيع الدخل أو الثروة أو متغيرات أخرى، ويقوم هذا المنحنى على المقارنة بين التوزيع الفعلي للظاهرة المدروسة والتوزيع المثالي، وقد بدأ استخدام هذا المنحنى أساسا لقياس توزيع الثروة أو الدخل (المنحنى يعبر عن العلاقة بين النسب التراكمية للدخل والنسب التراكمية للأفراد)، ولكن استخداماته توسعت فيما بعد وأصبح يستخدم في دراسات السكان والمراكز العمرانية والخدمات وغيرها، ويتشكل منحنى لورنز من:

- ✓ المحور الأفقي (محور الفواصل): يسجل فيه النسب التراكمية للأسر أو الأفراد أو العمال أو السكان (حسب المتغير المدروس) تبعا لمداخيلها (الثروة بصورة عامة).
- ✓ المحور العمودي (محور التراتيب): يسجل فيه النسب التراكمية للمداخيل المتحصل عليها من طرف كل نسبة من الأسر أو الأفراد أو العمال.
- ✓ الخط القطري 45°: هو التوزيع المتساوي للمداخيل (التوزيع المثالي)، اذ نجد أنه عند النسبة 10 % من الأفراد أو الأسر يكون الدخل المتحصل عليه هو 10 % ، وعند النسبة 50 % من الأفراد أو الأسر يكون الدخل المتحصل عليه هو 50 % وهكذا، وبعبارة أخرى يمكن القول أن10 % من الأفراد حصلوا على 10 % من الثروة وهكذا، والشكل البياني التالي يبين ذلك:

[1]- ظاهرة التمركز لا تطبق إلا على المتغيرات الاحصائية المستمرة ذات القيم الموجبة، والقابلة للجمع وفقط، كالدخل، الخدمات الصحية، التعليمية، توزيع السكان وغيرها.



ولتوضيح كيفية رسم منحنى لورنز نتبع الخطوات التالية:

أ- **تكوين الجدول الإحصائي الذي يتكون من:**

- التكرار المئوي للأفراد: $f_{i\%} = \frac{n_i}{\sum n_i} \times 100$؛
- النسبة المئوية التراكمية للأفراد $(F_{i\%} \uparrow)$؛
- النسبة المئوية لتوزيع الدخل لكل فئة: $q_{i\%} = \frac{n_i \times C_i}{\sum n_i \times C_i} \times 100$؛
- النسبة المئوية التراكمية لتوزيع الدخل لكل فئة $(Q_{i\%} \uparrow)$؛

ب- **رسم منحنى لورنز:** (النسبة المئوية التراكمية للأفراد $(F_{i\%} \uparrow)$، النسبة المئوية التراكمية للدخل $(Q_{i\%} \uparrow)$).

ج- تفسير المنحنى:

- كلما زادت الفجوة بين منحنى لورنز (التوزيع الفعلي) والتوزيع المثالي (خط المساواة أو خط العدالة)، كلما زاد التفاوت في التوزيع (بمعنى ازدادت اللامساواة في التوزيع أو بمعنى آخر يكون التوزيع التكراري أقل عدالة) والعكس صحيح، أي كلما اقترب منحنى لورنز من خط المساواة كلما كان التوزيع التكراري أكثر عدالة.

مثال: الجدول التالي يبين الدخل الشهري لـ 50 أسرة جزائرية:

(وحدة القياس 10^3 دج)

الأجر	10-15	15-20	20-25	25-30	30-35	35-40
الأسر	7	8	10	12	5	8

المطلوب:

– أدرس تمركز أجور هذه الأسر؟

الحل: لدراسة تمركز الأجور نتبع الخطوات التالية:

– **تكوين الجدول الإحصائي:**

الأجر x_i	عدد الأسر n_i	مركز الفئات (c_i)	$(f_{i\%})$	$(F_{i\%}\uparrow)$	$(n_i \times C_i)$	$(q_{i\%})$	$(Q_{i\%}\uparrow)$
$[10-15[$	7	12,5	14	14	87,5	7,03	7,03
$[15-20[$	8	17,5	16	30	140	11,24	18,27
$[20-25[$	10	22,5	20	50	225	18,07	36,34
$[25-30[$	12	27,5	24	74	330	26,51	62,85
$[30-35[$	5	32,5	10	84	162,5	13,05	75,9
$[35-40[$	8	37,5	16	100	300	24,10	100
$\sum n_i$	50	/	100	/	1245	100	/

– **التعليق على قيم الجدول:**

– هناك 14% من الأسر يحصلون على 7,03% من الدخل الشهري الإجمالي.

– هناك 30% من الأسر يحصلون على 18,27% من الدخل الشهري الإجمالي.

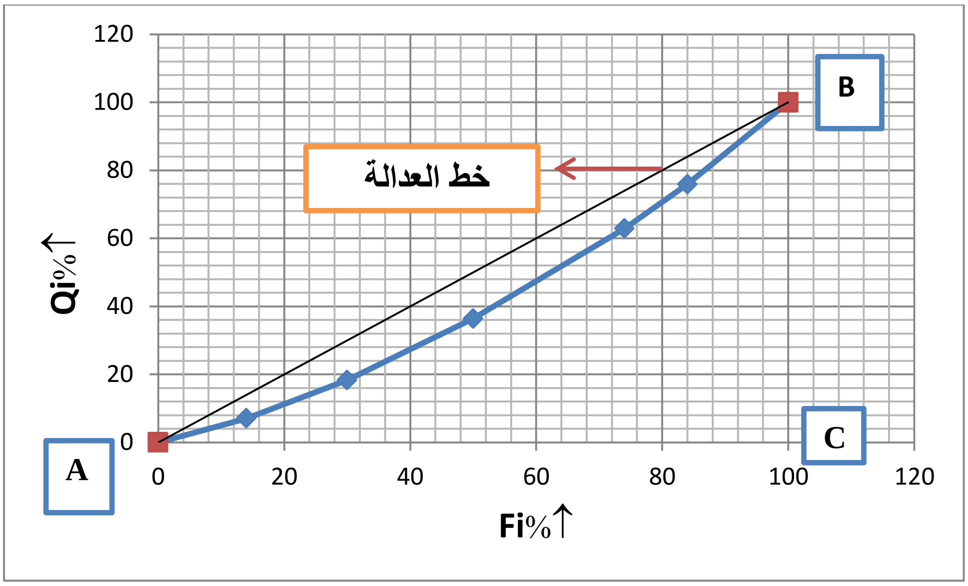
– هناك 50% من الأسر يحصلون على 36,34% من الدخل الشهري الإجمالي.

– هناك 74% من الأسر يحصلون على 62,85% من الدخل الشهري الإجمالي.

– هناك 84% من الأسر يحصلون على 75,9% من الدخل الشهري الإجمالي.

* نلاحظ أن توزيع الدخل الشهري للأسر أكثر عدالة.

– **رسم منحنى لورنز:** (النسبة المئوية التراكمية للأسر $(F_{i\%}\uparrow)$، النسبة المئوية التراكمية للدخل $(Q_{i\%}\uparrow)$).



التفسير:

- كلما زادت الفجوة بين منحنى لورنز (التوزيع الفعلي) والتوزيع المثالي (خط المساواة أو خط العدالة)، كلما زاد التفاوت في التوزيع (بمعنى ازدياد اللامساواة في التوزيع أو بمعنى آخر يكون التوزيع التكراري أقل عدالة) والعكس صحيح، أي كلما اقترب منحنى لورنز من خط المساواة كلما كان التوزيع التكراري أكثر عدالة.

- إذا انطبق منحنى لورنز على خط العدالة فإننا نقول أن التوزيع التكراري عادل تماما (مساواة تامة)، أما إذا انطبق المنحنى على المثلث (ABC) فإننا نقول أن التوزيع التكراري غير عادل (هناك اللامساواة في توزيع الدخل).

ومن خلال الشكل أعلاه نلاحظ أن توزيع الدخل الشهري للأسر أكثر عدالة لاقتراب منحنى لورنز من خط العدالة.

2- معامل جيني لقياس التمركز (GINI) (الطريقة الحسابية)

بالاعتماد على منحنى لورنز قدم عالم الرياضيات الإيطالي "جيني" مقياسا جبريا لقياس درجة التفاوت في توزيع الدخل، الذي يمثل قيمة المساحة المحصورة بين منحنى لورنز وخط العدالة (المساواة) مقسومة على قيمة مساحة المثلث (ABC) الموجود تحت خط العدالة، ويكون معامل جيني محصورا دائما بين 0 و 1، وبين 0 و100 إذا أخذت القيم بالنسب المئوية، أي:

$$0 \leq G \leq 1$$

أو:

$$0 \leq G \leq 100$$

أ- حساب معامل جيني: لحساب قيمة معامل جيني بشكل نسب مئوية نستخدم القانون التالي:

$$I_{Gini} = 100 - \frac{1}{100}\sum_{i=1}^{k} f_i\%(Q_i^{\uparrow}\% + \mathrm{Q}_{i-1}^{\uparrow}\%)$$

أو باستخدام التكرارت النسبية:

$$I_{Gini} = 1 - \sum_{i=1}^{k} f_i(Q_i^{\uparrow} + \mathrm{Q}_{i-1}^{\uparrow})$$

حيث:

$f_{i\%}$: تمثل التكرار المئوي؛

$Q_{i\%}^{\uparrow}$: النسبة المئوية التراكمية لتوزيع الدخل لكل فئة؛

$\mathrm{Q}_{(i-1)\%}^{\uparrow}$: النسبة المئوية التراكمية لتوزيع الدخل للفئة السابقة $(i-1)$.

ب– تفسير نتائج معامل جيني:

- عندما ينطبق منحنى لورنز على خط العدالة أو المساواة فتصبح المساحة بينهما معدومة أي مساوية للصفر، فيصبح معامل جيني في هذه الحالة **مساوي للصفر**: $G = 0$ ←← هناك مساواة تامة لتوزيع الدخل، أو بمعنى آخر التوزيع عادل تماما؛
- عندما ينطبق منحنى لورنز على المثلث (ABC) فتصبح المساحة بينهما متساوية، بمعنى أن معامل جيني في هذه الحالة يكون **مساوي للواحد الصحيح**: $G = 1$ ←← فإن التوزيع التكراري هنا غير عادل (هناك اللامساواة في توزيع الدخل).

وبالرجوع إلى المثال السابق نقوم بحساب معامل جيني كما يلي:

$(f_{i\%})$	$(\mathrm{Q}_{i\%}\uparrow)$	$(\mathrm{Q}_{(i-1)\%}^{\uparrow})$	$(Q_i^{\uparrow} + \mathrm{Q}_{i-1}^{\uparrow})$	$f_i\%(Q_i^{\uparrow}\% + \mathrm{Q}_{i-1}^{\uparrow}\%)$
14	7,03	-	7,03	98,42
16	18,27	7,03	25,3	404,8
20	36,34	18,27	54,61	1092,2
24	62,85	36,34	99,19	2380,56
10	75,9	62,85	138,75	1387,5
16	100	75,9	175,9	2814,4
100	/	/	/	8177,88

$$I_{Gini} = 100 - \frac{1}{100}\sum_{i=1}^{k} f_i\%\left(Q_i^{\uparrow}\% + Q_{i-1}^{\uparrow}\%\right)$$

$$I_{Gini} = 100 - \frac{8177,88}{100} = 100 - 81,7788 = 18,2212$$

بما أن معامل جيني $I_G = 18,22$ وهو يقترب من الصفر، فإن هذا التوزيع أكثر عدالة (أقرب إلى المساواة)، ما يدعم النتيجة المتحصل عليها باستخدام منحنى لورنز.

تمارين الفصل

التمرين الأول:

المعلومات التالية تخص أجور عينة من العمال تتكون من 36 عامل بمؤسسة خاصة مقاسة بـ 100 دينار:

الفئات	[75,90[	[90,105[	[105,120[	[120,135[	[135,150[	[150,165[	[165,180[
العمال	2	6	4	10	5	4	5

- حدد مقاييس النزعة المركزية (الوسيط، المتوسط الحسابي، المنوال) مع التفسير؛
- أحسب الربيعي الأول والمئين 50، مع التفسير؛
- ما هو عدد العمال الذي يقل أجرهم عن 130 دينار؛
- ما هو طبيعة هذا التوزيع الاحصائي؛
- أدرس تمركز الأجور باستخدام منحنى لورنز ومعامل جيني.

التمرين الثاني:

المعلومات تخص أجور عينة من العمال تتكون من 60 عامل بأحد المصانع هو كالتالي:

(وحدة القياس 10^3 دج)

الأجر	7-17	17-27	27-37	37-47	47-57	57-67
التكرار	7	8	9	12	9	8

- **المطلوب:**
- أحسب المتوسط الحسابي؛
- أحسب الوسيط حسابيا وبيانيا؛
- ثم حدد المنوال؛
- ما هو طبيعة أو شكل التوزيع الإحصائي.
- ما هو عدد العمال الذين يقل أجرهم عن 30 دج.
- أوجد قيمة العشير السابع، مع تفسير النتيجة؛
- أدرس التمركز بيانيا.

التمرين الثالث:

الجدول التالي يبين الدخل الشهري لـ 110 أسرة جزائرية:

(وحدة القياس 10^3 دج)

الأجر	15-20	20-25	25-30	30-35	35-40
الأسر	25	30	20	25	10

المطلوب:

– أذكر خطوات تمثيل منحنى لورنز؛

– هل هذا التوزيع أكثر عدالة، برر إجابتك باستخدام الطريقة الحسابية؟

المحور السابع: الأرقام القياسية

1- تمهيد

2- مفهوم الرقم القياسي

3- أنواعها: - الأرقام القياسية البسيطة؛ التجميعية؛ والمرجحة.

تمهيد:

كما هو معلوم فإن علم الاقتصاد يعرف بعدم الاستقرار وحالة عدم التأكد، لهذا نجد دائما هناك تغير مستمر من فترة لأخرى للعديد من الظواهر الاقتصادية، لهذا نلجأ إلى تقنيات الأرقام القياسية لتسهيل على الباحث دراسة وتحليل الظواهر الاقتصادية، خاصة في حالة المقارنة بين فترتين زمنيتين مختلفتين أو مكانين مختلفين.

1- المفهوم:

أداة تستخدم لقياس التغير النسبي أو المئوي في قيم الظواهر من زمن لآخر ومن مكان إلى آخر كأسعار السلع مثلا، بمعنى أنه اذا نسبنا ظاهرة ما في زمن معين إلى قيمتها في زمن آخر فإننا سوف نحسب الرقم القياسي ونرمز له بـ (I).

مثال: كان سعر منتوج ما في سنة 2020 ، 100 دج وفي سنة 2021، ارتفع سعره إلى 140دج:

هناك زمنان اثنان أحدهما يمثل الأساس والآخر يمثل المقارن، ونحسب الرقم القياسي كالتالي:

$$I_{2021/2020} = \frac{140}{100} \times 100 = 140\%$$

سنة 2020 تسمى سنة الأساس، أما سنة 2021 فهي سنة المقارنة "الراهنة".

ونستنتج مما سبق أنه يوجد صيغتين للأرقام القياسية

✓ **الرقم القياسي الزمني:**

$$\text{الرقم القياسي للظاهرة} = \frac{\text{قيمة الظاهرة في سنة المقارنة}}{\text{قيمة الظاهرة في سنة الأساس}} \times 100$$

✓ **الرقم القياسي المكاني:**

$$\text{الرقم القياسى للظاهرة} = \frac{\text{قيمة الظاهرة في مكان المقارنة}}{\text{قيمة الظاهرة في مكان الأساس}} \times 100$$

2- أنواع الأرقام القياسية: هناك عدة أنواع من الأرقام القياسية نذكرها كالتالي:

1.2- الأرقام القياسية البسيطة: يقيس الرقم القياسي البسيط تطور سعر أو كمية أو قيمة مادة واحدة فقط بين فترتين زمنيتين مختلفتين أو مكانين مختلفين، وهو عبارة عن النسبة بين قيمة المتغير في فترة المقارنة (t_1) إلى قيمة نفس المتغير في سنة الأساس (t_0).

➢ **الرقم القياسي البسيط للأسعار:**

ويقصد به إظهار سعر سلعة واحدة معينة في فترة المقارنة منسوبا إلى سعر نفس السلعة في فترة الأساس، ويعبر عنه بالعلاقة التالية

$$\text{الرقم القياسى البسيط للأسعار} = \frac{\text{السعر في سنة المقارنة}}{\text{السعر في سنة الأساس}} \times 100$$

أي:

$$I_{P\,t_1/t_0} = \frac{P_1}{P_0} \times 100$$

➢ **الرقم القياسي البسيط للكميات:**

ويستخدم هذا الرقم القياسي في حالة المقارنة بين كميات السلع بدلا من أسعارها، ويعبر عنه بالعلاقة التالية:

$$I_{Q\,t_1/t_0} = \frac{Q_1}{Q_0} \times 100$$

➢ **الرقم القياسي البسيط للقيمة:**

تعبر القيمة الاجمالية للسلعة بكمية هذه السلعة مضروبة في سعرها $(V = P \times Q)$، فإذا كانت (P_0, Q_0) تعبر عن سعر السلعة والكمية المنتجة منها في سنة الأساس و (P_1, Q_1) تعبر عن سعر السلعة والكمية المنتجة منها في سنة المقارنة، فإن القيمة الإجمالية خلال سنة الأساس هي (V_0) وخلال سنة المقارنة هي (V_1) وعليه فإن:

$$I_{V\,t_1/t_0} = \frac{V_1}{V_0} \times 100 = \frac{P_1 \times Q_1}{P_0 \times Q_0} \times 100$$

مثال: الجدول التالي يبين الأسعار والكميات المنتجة من سلعة معينة للفترة 2021-2022، والمطلوب إيجاد الأرقام القياسية البسيطة للأسعار، الكميات والقيمة:

السنة	البيانات الأساسية			حسابات الأرقام القياسية البسيطة		
	السعر	**الكمية**	**القيمة**	**للأسعار %**	**الكميات%**	**للقيم %**
2021 **2022**	**20** **35**	**30** **44**	**600** **1540**	$I_{P\,t_1/t_0} = \frac{P_1}{P_0} \times 100 = \frac{20}{20} \times 100 = 100\%$ **%175**	**%100** $I_{Q\,t_1/t_0} = \frac{Q_1}{Q_0} \times 100$ $= \frac{44}{30} \times 100 = 146{,}67\%$	**100%** $I_{V\,t_1/t_0} = \frac{V_1}{V_0} \times 100$ $= \frac{1540}{600} \times 100 = 256{,}67\%$

تفسير قيم الجدول:

- الارقام القياسية تساوي دائما 100 إذا تطابقت السنة الراهنة (المقارنة) مع سنة الأساس؛

- $\left(I_{P\,t_1/t_0} = 175 \right)$: أي أن الأسعار في سنة 2022 ارتفعت بنسبة 75% مقارنة بسنة 2021: نسبة الزيادة هي الفرق بين (175-100=75%)؛

- $\left(I_{Q\,t_1/t_0} = 146{,}67 \right)$: هناك ارتفاع في الكمية المنتجة لهذه السلعة سنة 2022 بنسبة 46,67% مقارنة بسنة 2021: نسبة الزيادة هي الفرق بين (146,67-100=46,67%)؛

- $\left(I_{V\,t_1/t_0} = 256{,}67 \right)$: يعني أن قيمة السلعة زادت بنسبة 156,67% في سنة 2022 عما كانت عليه في سنة 2021: نسبة الزيادة هي الفرق بين (256,67-100=156,67%).

ملاحظة: إذا كانت النتيجة المتحصل عليها بالنسبة للأرقام القياسية أيا كان نوعها أقل من 100، فهذا يعني أن هناك انخفاض في الكمية أو السعر أو القيمة وهكذا.

مثال: كان سعر سلعة ما في سنة 2021 (سنة أساس) 100 دج وفي سنة 2022 انخفض سعرها إلى 75دج، فإن الرقم القياسي للسعر يقدر بـ: $I_{P\,t_1/t_0} = \frac{P_1}{P_0} \times 100 = \frac{75}{100} \times 100 = 75$ ، أي أن هناك انخفاض في سعر السلعة بنسبة (100-75=25%).

2.2– الأرقام القياسية التجميعية: يقيس الرقم القياسي التجميعي النسبة بين مجموعة من الأسعار أو الكميات أو القيم في سنة المقارنة (t_1) إلى جميع أسعارها وكمياتها وقيمها في سنة الأساس (t_0)، ويحسب الرقم القياسي التجميعي حسب الصيغ أذناه:

➢ **الرقم القياسي التجميعي للأسعار:**

$$I_{P\, t_1/t_0} = \frac{\sum P_1}{\sum P_0} \times 100$$

➢ **الرقم القياسي التجميعي لكميات:**

$$I_{Q\, t_1/t_0} = \frac{\sum Q_1}{\sum Q_0} \times 100$$

➢ **الرقم القياسي التجميعي للقيم:**

$$I_{V\, t_1/t_0} = \frac{\sum V_1}{\sum V_0} \times 100 = \frac{\sum(P_1 \times Q_1)}{\sum(P_0 \times Q_0)} \times 100$$

مثال: الجدول التالي يبين أسعار والكميات المنتجة لثلاث سلع في سنتي 2020 و2023، والمطلوب إيجاد الأرقام القياسية التجميعية للأسعار والكميات:

نوع السلعة	2020 (سنة الأساس)		سنة 2023	
	السعر	الكمية	السعر	الكمية
السلعة A (كغ)	50	30	53	40
السلعة B (لتر)	25	40	45	36
السلعة C (كغ)	1200	16	1000	20
المجموع	1275	86	1098	96

الحل:

– إيجاد الأرقام القياسية التجميعية:

1– للأسعار:

$$I_{P\, t_1/t_0} = \frac{\sum P_1}{\sum P_0} \times 100 = \frac{1098}{1275} \times 100 = 86{,}12\%$$

التفسير: نلاحظ أن أسعار السلع انخفضت سنة 2023 بنسبة (100-86,12=13,88%) عما كانت عليه في سنة 2020.

2- للكميات:

$$I_{Q\,t_1/t_0}=\frac{\sum Q_1}{\sum Q_0}\times 100=\frac{96}{86}\times 100=111{,}63\%$$

التفسير: هناك ارتفاع في الكميات المنتجة للسلع سنة 2023 بنسبة 11,63% مقارنة بسنة 2020.

ملاحظة: من بين عيوب الارقام القياسية التجميعية بالرغم من سهولة حسابها وهو عدم الاهتمام بوحدات القياس، اضافة أنها لا تأخذ الاهمية النسبية للسلع، فنلاحظ من خلال المثال أعلاه أن الرقم القياسي التجميعي يتأثر بالأسعار المرتفعة.

3.2- الأرقام القياسية المرجحة: للتغلب على عيوب الطريقة التجميعية البسيطة فإن الصيغ المرجحة للأرقام القياسية تعتمد على ترجيح أسعار أو كميات كل سلعة باستخدام وزن معين، ويستخدم عادة كمية السلعة المباعة أو سعرها خلال سنة الأساس أو خلال سنة المقارنة أو خلال سنة نموذجية (قد تكون متوسط عدد من السنوات). وهذه الأوزان تشير إلى الأهمية النسبية للسلعة، وهناك ثلاث صيغ للأرقام القياسية المرجحة تعتمد على ما إذا كنا نستخدم كميات أو أسعار سنة الأساس أو المقارنة.

- **الرقم القياسي المرجح لـ "لاسبير": Indices de Laspeyres**: وهو الرقم القياسي التجميعي المرجح باستخدام سنة الأساس. وقد اقترح هذا المؤشر سنة 1844 من طرف لاسبير. وهناك صيغتين لهذا الرقم:

- **صيغة الرقم القياسي المرجح للأسعار لـ "لاسبير"**: $$I_PL_{t_1/t_0}=\frac{\sum(P_1\times Q_0)}{\sum(P_0\times Q_0)}\times 100$$
- **صيغة الرقم القياسي المرجح للكميات لـ "لاسبير"**: $$I_QL_{t_1/t_0}=\frac{\sum(Q_1\times P_0)}{\sum(Q_0\times P_0)}\times 100$$

- **الرقم القياسي المرجح لـ "باش": Indices de Pache**: : وهو الرقم القياسي التجميعي المرجح باستخدام سنة المقارنة، وهناك صيغتين لهذا الرقم:

- **صيغة الرقم القياسي المرجح للأسعار لـ "باش"**: $I_P P_{t_1/t_0} = \frac{\sum(P_1 \times Q_1)}{\sum(P_0 \times Q_1)} \times 100$

- **صيغة الرقم القياسي المرجح للكميات لـ "باش"**: $I_Q P_{t_1/t_0} = \frac{\sum(Q_1 \times P_1)}{\sum(Q_0 \times P_1)} \times 100$

➢ **الرقم القياسي المرجح لـ "فيشر"**: **Indices Fisher**: تبين مما سبق أن رقم لاسبير يجعل صيغة الرقم القياسي متحيز إلى أعلى بالنظر إلى أنه مبني على الترجيح بأوزان نسبة الأساس، أما الرقم باش فيجعل الرقم القياسي متحيز إلى أسفل لأنه يستند على الترجيح بأوزان سنة المقارنة، وعليه فقد اقترحت عدة صيغ لمعالجة الفرق بين الترجحين، وقد كانت صيغة فيشر اشهرها والذي جاء كحل وسط بين رقمي لاسبير وباش لتكوين رقما قياسيا أمثلا الذي يساوي المتوسط الهندسي للرقمين.

الرقم القياسي المرجح لـ "فيشر" = $\sqrt{\text{رقم لاسبير} \times \text{رقم باش}}$

$$IF_{t_1/t_0} = \sqrt{IL_{t_1/t_0} \times IP_{t_1/t_0}}$$

ومنه يكون لدينا صيغتين لهذا الرقم:

- **صيغة الرقم القياسي المرجح للأسعار لـ "فيشر"**:

$$IF_{t_1/t_0} = \sqrt{IL_{t_1/t_0} \times IP_{t_1/t_0}} = \sqrt{\frac{\sum(P_1 \times Q_0)}{\sum(P_0 \times Q_0)} \times \frac{\sum(P_1 \times Q_1)}{\sum(P_0 \times Q_1)}} \times 100$$

- **صيغة الرقم القياسي المرجح للكميات لـ "فيشر"**:

$$IF_{t_1/t_0} = \sqrt{IL_{t_1/t_0} \times IP_{t_1/t_0}} = \sqrt{\frac{\sum(Q_1 \times P_0)}{\sum(Q_0 \times P_0)} \times \frac{\sum(Q_1 \times P_1)}{\sum(Q_0 \times P_1)}} \times 100$$

المثال السابق: أحسب جميع صيغ الرقم القياس لكل من "لاسبير، باش، فيشر" ؟.

نوع السلعة	2020 (سنة الأساس)		سنة 2023		حسابات مساعدة			
	السعر P_0	الكمية Q_0	السعر P_1	الكمية Q_1	$P_0 \times Q_0$	$P_1 \times Q_0$	$Q_1 \times P_0$	$P_1 \times Q_1$
السلعة **A** (كغ)	50	30	53	40	1500	1590	2000	2120
السلعة **B** (لتر)	25	40	45	36	1000	1800	900	1620
السلعة **C** (كغ)	1200	16	1000	20	19200	16000	24000	20000

23740	26900	19390	21700	96	1098	86	1275	**المجموع**

– الرقم القياسي المرجح لـ "لاسبير":

✓ **للأسعار:**

$I_P L_{t_1/t_0} = \frac{\sum(P_1 \times Q_0)}{\sum(P_0 \times Q_0)} \times 100 = \frac{19390}{21700} \times 100 = 89{,}36\%$ $\Leftarrow$ هناك انخفاض في الأسعار بنسبة 10,64%.

✓ **للكميات:**

$I_Q L_{t_1/t_0} = \frac{\sum(Q_1 \times P_0)}{\sum(Q_0 \times P_0)} \times 100 = \frac{26900}{21700} \times 100 = 123{,}96\%$ $\Leftarrow$ هناك ارتفاع في الكميات المنتجة بنسبة 23,96%.

– الرقم القياسي المرجح لـ "باش":

✓ **للأسعار:**

$I_P P_{t_1/t_0} = \frac{\sum(P_1 \times Q_1)}{\sum(P_0 \times Q_1)} \times 100 = \frac{23740}{26900} \times 100 = 88{,}25\%$ $\Leftarrow$ هناك انخفاض في الأسعار بنسبة 11,75%.

✓ **للكميات:**

$I_Q P_{t_1/t_0} = \frac{\sum(Q_1 \times P_1)}{\sum(Q_0 \times P_1)} \times 100 = \frac{23740}{19390} \times 100 = 122{,}43$ $\Leftarrow$ هناك ارتفاع في الكميات المنتجة بنسبة 22,43%.

– الرقم القياسي المرجح لـ "فيشر":

✓ **للأسعار:**

$IF_{t_1/t_0} = \sqrt{IL_{t_1/t_0} \times IP_{t_1/t_0}} = \sqrt{89{,}36 \times 88{,}25} = 88{,}80$ $\Leftarrow$ هناك انخفاض في الأسعار بنسبة 11,2%.

✓ **للكميات**

$IF_{t_1/t_0} = \sqrt{IL_{t_1/t_0} \times IP_{t_1/t_0}}\sqrt{123{,}96 \times 122{,}43} = 123{,}19$ $\Leftarrow$ هناك ارتفاع في الكميات المنتجة بنسبة 23,19%.

خصائص الأرقام القياسية المرجحة:

تتوفر الأرقام القياسية على الشروط التالية:

✓ المطابقة: بمعنى تطابق قيمة ظاهرة ما (سعر، كمية، أو قيمة) لسنة الأساس مع السنة الراهنة (المقارنة)، أي:

$$I_{P\,t_0/t_0} = \frac{P_0}{P_0} \times 100 = 100\%$$

أو

$$I_{P\,t_1/t_1} = \frac{P_1}{P_1} \times 100 = 100\%$$

- قابلية الانعكاس: إذا تحققا العلاقة التالية، نستطيع القول أن هذا الرقم القياسي حقق خاصية الانعكاس، والعكس صحيح.

$$I_{t_1/t_0} \times I_{t_0/t_1} \times 100 = 100\% = 1$$

- قابلية تغيير الأساس:
- قابلية الدوران (التحول): إذا كان لدينا الأرقام القياسية $\left(I_{t_1/t_0}\right)$، $\left(I_{t_2/t_1}\right)$، $\left(I_{t_3/t_2}\right)$، ونريد حساب الرقم القياسي $\left(I_{t_3/t_0}\right)$، يكون ذلك محقق إذا كان:

$$I_{P\, t_3/t_0} = \frac{I_{t_3/t_2} \times I_{t_2/t_1} \times I_{t_1/t_0}}{100^{n-1}}$$

تمارين الفصل

التمرين الأول:

ارتفعت مبيعات سلعة ما من سنة 2020 إلى 2022 بنسبة 35%، في حين ارتفع سعرها في هذه الفترة بنسبة 25%، أحسب معدل نمو كميات السلعة المستهلكة؟

التمرين الثاني:

الجدول التالي يبين تغير سعر سلعة معينة خلال الفترة 2019-2022:

السنوات	2019	2020	2021	2022
P_x	20	25	36	30

- أحسب الرقم القياسي لسعر السلعة لجميع السنوات باعتبار 2019 سنة الأساس؛
- إعادة حساب الرقم القياسي باعتبار سنة الأساس هي 2020؛
- قم باختبار تحقق خاصية الدوران للرقم القياسي $\left(I_{2022/2019}\right)$.

التمرين الثالث:

الجدول التالي يبين أسعار وكميات السلع خلال سنتي 2020 و2022:

نوع السلعة	2020 (سنة الأساس)		سنة 2022	
	السعر P_0	الكمية Q_0	السعر P_1	الكمية Q_1
السلعة **A**	20	20	25	35
السلعة **B**	15	25	20	30
السلعة **C**	20	30	22	40

- أحسب الرقم القياسي البسيط لسعر السلعة **A**، مع تفسير النتيجة؛
- أحسب الرقم القياسي البسيط لكمية السلعة **B**، مع تفسير النتيجة؛
- أحسب الرقم القياسي البسيط لقيمة السلعة **C**، مع تفسير النتيجة؛
- أحسب الرقم القياسي التجميعي لأسعار وكميات السلع؛
- أحسب الأرقام القياسية للأسعار والكميات لكل من لاسبير، باش، وفيشر.

التمرين الرابع:

من خلال البيانات التالية:

نوع السلعة	2019 (سنة الأساس)		سنة 2020	
	السعر P_0	الكمية Q_0	نسبة الزيادة في الأسعار %	الكمية Q_1
السلعة **A**	15	10	50	؟
السلعة **B**	30	15	16,67	20
السلعة **C**	55	30	18,18	35
السلعة **D**	120	15	16,67	20

- أحسب أسعار السلع في سنة 2020؛
- أحسب كمية السلعة **A** علما أن الكميات انخفضت سنة 2020 بنسبة 30%؛
- أحسب معامل فيشر للأسعار.

المحور الثامن: الانحدار والارتباط

تمهيد

يعرّف الإحصاء بأنه مجموعة الطرق العلمية التي تستخدم في جمع وعرض وتحليل البيانات، وذلك باستخدام مجموعة من المقاييس، فالهدف بطبيعة الحال هو دراسة ظاهرة واحدة (متغير احصائي واحد)، أما إذا كانت البيانات تتعلق بسلوك ظاهرتين (متغيرين احصائيين) كالكمية المطلوبة من سلعة ما وسعرها، الدخل المتاح والاستهلاك...، فالهدف هنا هو محاولة ترجمة العلاقة إلى علاقة رياضية من أجل التحليل والتنبؤ.

- ✓ فإذا كان اهتمام الباحث هو دراسة العلاقة بين متغيرين نستخدم أسلوب الارتباط (نوع وقوة العلاقة بين المتغيرين)؛
- ✓ إذا كان اهتمام الباحث هو دراسة أثر أحد المتغيرين على الآخر نستخدم هنا أسلوب الانحدار.

1- توزيعات المتغيرات الثنائية التغير (الجداول الاحصائية الثنائية)

يقصد بالجداول الثنائية أو التوزيع التكراري المزدوج (ثنائي البعد)، بالجداول التي تتوزع فيها بيانات لمتغيرين احصائيين، أحدهما يكون في الأسطر والمتغير الآخر يكون في الأعمدة، مما ينتج عنهما مربعات تحتوى على تكرارات مشتركة بين المتغيرين، والجدول التالي يبين ذلك:

1.1- الجداول الاحصائية الثنائية

(Y_j) / (X_i)	Y_1	Y_2	$\dots$	Y_m	التوزيع الهامشي (الحدي) لـ (X_i)
X_1	n_{11}	n_{12}	$\dots$	n_{1m}	$= n_1$
X_2	n_{21}	n_{22}	$\dots$	n_{2m}	$= n_2$
$\vdots$	$\iddots$	$\iddots$	$\iddots$	$\iddots$	$\vdots$
X_k	n_{k1}	n_{k2}	$\iddots$	n_{km}	$= n_k$
التوزيع الهامشي (الحدي) لـ (Y_j)	$= n_1$	$= n_2$	$\dots$	$= n_m$	المجموع $\sum n_i$ / المجموع $\sum n_j$

حيث: $N_{\text{الكلي}} = \sum n_i = \sum n_j$

ملاحظات:

- **بالنسبة للأسطر** n_i:

➢ $n_1 = n_{11} + n_{12} + \dots\dots + n_{1m}$.

➢ $n_2 = n_{21} + n_{22} + \dots\dots + n_{2m}$

➢ $n_k = n_{k1} + k_{k2} + \dots\dots n_{km}$

- **بالنسبة للأعمدة** n_j:

➢ $n_1 = n_{11} + n_{21} + \dots\dots + n_{k1}$

➢ $n_2 = n_{12} + n_{22} + \dots\dots + n_{k2}$

➢ $n_m = n_{1m} + n_{2m} + \dots\dots n_{km}$

- دائما التكرار الكلي يساوي مجموع الأعمدة يساوي مجموع الأسطر **أي**:

$$N = \sum n_i = \sum n_j$$

- بإمكاننا بدل من أخذ التكرارات المطلقة نأخذ التكرارات النسبية، ويصبح لدينا جدول التوزيع الثنائي النسبي كالتالي:

$$f_{ij} = \frac{n_{ij}}{N}$$

(Y_j) / (X_i)	Y_1	Y_2	$\dots$	Y_m	**التوزيع الهامشي (الحدي) ل** (X_i)
X_1	f_{11}	f_{12}	$\dots$	f_{1m}	$= f_1$
X_2	f_{21}	f_{22}	$\dots$	f_{2m}	$= f_2$
$\vdots$	$\ddots$	$\ddots$	$\ddots$	$\ddots$	$\vdots$
X_k	f_{k1}	f_{k2}	$\ddots$	f_{km}	$= f_k$
التوزيع الهامشي (الحدي) ل (Y_j)	$= f_1$	$= f_2$	$\dots$	$= f_m$	**المجموع** $\sum f_i = 1$ / **المجموع** $\sum f_j = 1$

2.1- جداول التكرارات الحدية (التوزيعات الهامشية):

التوزيع الهامشي أو الحدي للمتغير (X) و(Y) هو التوزيع الأحادي للمتغير (X) و(Y) مع عدم اعتبار ما حدث للمتغير (X) أو(Y)، وهو نتيجة تجميع التكرارات المشتركة، والجدول التالي يبين ذلك:

- **بالنسبة للمتغير** (X):

(X_i)	X_1	X_2	$\dots$	X_k
n_i	n_1	n_2	$\dots$	n_k

- **بالنسبة للمتغير** (Y):

(Y_j)	Y_1	Y_2	$\dots$	Y_m
n_J	n_1	n_2	$\dots$	n_m

3.1- جداول التكرارات الشرطية (والنسبية الشرطية):

يمثل التوزيع التكراري (أو النسبي) للمتغير (X) عند القيمة الثابتة معطاة للمتغير (Y) بالتوزيع الشرطي (أو شرطي نسبي) للمتغير (X) بشرط (Y_j)، والتوزيع التكراري (أو النسبي) للمتغير (Y) عند القيمة الثابتة معطاة للمتغير (X) بالتوزيع الشرطي (أو شرطي نسبي) للمتغير (Y) بشرط (X_i)، ونعبر عنها بالجداول التالية:

- **بالنسبة للمتغير** (X) **شرط** (Y_m) مثلا: **(وهكذا مع باقي الأعمدة):**

(X_i)	X_1	X_2	$\dots$	X_k
n_{ij}	n_{1m}	n_{2m}	$\dots$	n_{km}

- **بالنسبة للمتغير** (Y) **شرط** (X_1) **مثلا**: **(وهكذا مع باقي الأسطر)**

(Y_j)	Y_1	Y_2	$\dots$	Y_m
n_{ij}	n_{11}	n_{12}	$\dots$	n_{1m}

ملاحظة:

- التوزيع التكراري النسبي الشرطي للمتغير (X) بشرط (Y_j):

$$\frac{n_{ij}}{n_j}$$

- التوزيع التكراري النسبي الشرطي للمتغير (Y) بشرط (X_i):

$$\frac{n_{ij}}{n_i}$$

2- أسلوب الارتباط

1.2- تعريف

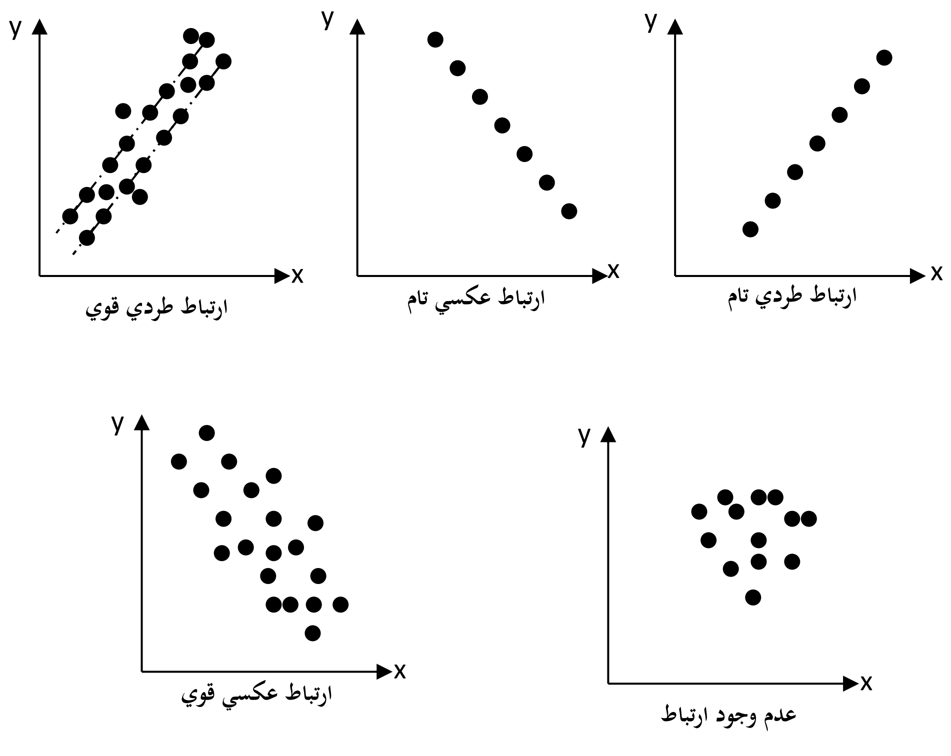
يعني الارتباط وجود علاقة بين ظاهرتين أو متغيرين، بمعنى التغير في إحدى المتغيرين يؤدي إلى التغير في المتغير الآخر بالزيادة أو النقصان، فالاعتماد على معامل الارتباط يبين طبيعة وقوة العلاقة بين المتغيرات الاقتصادية.

2.2- قياس الارتباط الخطي البسيط[2]

يمكن قياس الارتباط بين متغيرين باستخدام طريقتين (بيانيا وحسابيا):

أ- **بيانيا: برسم شكل الانتشار:**

إن شكل الانتشار يبين لنا العلاقة التي تربط بين المتغيرين بشكل مباشر وواضح، أما بالنسبة لقوة العلاقة فيمكن التعرف عليها بشكل تقريبي، وذلك من خلال ملاحظة مدى قرب وبعد نقط شكل الانتشار عن بعضها البعض، والأشكال التالية تبين ذلك:



ارتباط طردي تام — ارتباط عكسي تام — ارتباط طردي قوي

عدم وجود ارتباط — ارتباط عكسي قوي

ب- **حسابيا: معامل الارتباط الخطي البسيط لـ "بيرسون"** (r_P) :

يستخدم لقياس الارتباط بين متغيرين احصائيين فقط في حالة البيانات الكمية، وقيمته تتراوح دائما بين (1) و(-1)، ويتم حسابه باستخدام الصيغ التالية:

---

[2] - **الخطي** يعني أن العلاقة التي تربط بين المتغيرين علاقة خطية، أما **البسيط** فمعناها أن العلاقة تكون بين متغيرين احصائيين فقط.

- **طريقة التباين المشترك:**

$$r_P = \frac{COV(x,y)}{\sigma_x \sigma_y}$$

- **طريقة الانحرافات:**

$$r_P = \frac{\sum\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{\sqrt{\sum\left(X_i - \overline{X}\right)^2}\sqrt{\sum\left(Y_i - \overline{Y}\right)^2}}$$

- **الطريقة المختصرة:**

$$r_P = \frac{\sum X_i Y_i - n\overline{X}\,\overline{Y}}{\sqrt{\sum X_i^2 - n\overline{X}^2}\sqrt{\sum Y_i^2 - n\overline{Y}^2}}$$

حيث:

$COV(x,y)$: التباين المشترك بين (x) و (y)؛

$(\overline{X})$ و $(\overline{Y})$: المتوسط الحسابي لكل من (x) و (y) على التوالي؛

(σ_X) و (σ_Y): الانحراف المعياري لكل من (x) و (y) على التوالي؛

(n): عدد قيم المتغيرين.

ج- **خصائص معامل الارتباط الخطي البسيط لـ "بيرسون"** (r_P) :

- ✓ معامل الارتباط يأخذ قيمة عددية واحدة تلخص لنا اتجاه وقوة العلاقة بين المتغيرين؛
- ✓ تتراوح قيمته بين 1 و-1، أي: $-1 \leq r_P \leq 1$؛
- ✓ إذا كان معامل الارتباط سالب $(r_P \prec 0)$ دلّ ذلك على وجود علاقة عكسية، وتزداد قوة العلاقة كلما اقترب المعامل من (-1)؛
- ✓ إذا كان معامل الارتباط موجب $(r_P \succ 0)$ دلّ ذلك على وجود علاقة طردية، وتزداد قوة العلاقة كلما اقترب المعامل من $(+1)$؛
- ✓ إذا كان معامل الارتباط معدوم $(r_P = 0)$ دلّ ذلك على عدم وجود علاقة، وكلما اقترب المعامل من (0) ذل ذلك على أن العلاقة بين المتغيرين ضعيفة؛
- ✓ إذا كان معامل الارتباط $(r_P = \pm 1)$ دل ذلك على وجود ارتباط تام؛

✓ الارتباط بين (x) و (y) هو نفسه بين (y) و (x)؛

✓ الارتباط بين أي متغير ونفسه (الارتباط بين (x) و (x)) يساوي دائما 1.

مثال:

إذا كانت لدينا المعطيات التالية: حيث (x_i) هو متغير سنوات الخبرة و (y_i) يمثل كمية الإنتاج لمؤسسة ما.

$$n=12,\quad \sum X_i=72,\quad \sum Y_i=96,\quad \sum XY=624,\quad \sum X_i^2=478,\quad \sum Y_i^2=828.$$

المطلوب:

- هل هناك علاقة بين سنوات خبرة العمال وكمية الإنتاج لهذه المؤسسة؟ وما نوعها؟

الحل: حسب معطيات التمرين فإن أنسب قانون هو القانون المختصر:

ولكن علينا أولا إيجاد المتوسطات الحسابية:

$$\overline{X}=\frac{\sum X_i}{n}=\frac{72}{12}=6$$
$$\overline{Y}=\frac{\sum Y_i}{n}=\frac{96}{12}=8$$

ثم نطبق القانون التالي:

$$r_P=\frac{\sum X_iY_i-n\overline{X}\,\overline{Y}}{\sqrt{\sum X_i^2-n\overline{X}^2}\sqrt{\sum Y_i^2-n\overline{Y}^2}}=\frac{621-12(6)(8)}{\sqrt{478-12(36)}\sqrt{828-12(64)}}=\frac{45}{\sqrt{46}\sqrt{60}}=0{,}856$$

التفسير: يدل معامل الارتباط على وجود علاقة طردية قوية بين سنوات خبرة العمال وكمية الإنتاج لهذه المؤسسة.

3.2- قياس معامل الارتباط الرتبي (سبيرمان):

يستخدم لقياس الارتباط بين متغيرين احصائيين فقط في حالة البيانات الكمية، أو النوعية أو بين بيانات أحدها كمي والآخر نوعي، ويعطي الأهمية لرتبة القيم (ترتيبها) بدلا من قيمها، ويحسب كما يلي:

$$r_S=1-\frac{6\sum di^2}{n(n^2-1)}$$

حيث: (d_i): هي الفرق بين رتبة المتغير (x_i) ورتبة المتغير (y_i)؛

(n): عدد قيم المتغيرين.

مثال: ليكن لدينا المتغيرين الاحصائيين التاليين:

$(x_i)\rightarrow$ 5، 2، 7، 2

$(y_i)\rightarrow$ 3، 2، 9، 0

المطلوب:

- أوجد معامل الارتباط سبيرمان، مع تفسير النتيجة؟

الحل: لإيجاد معامل سبيرمان يجب علينا أولا ترتيب السلاسل الاحصائية تصاعديا من أجل ايجاد رتب القيم وفي حالة وجود أكثر من رتبة نحسب المتوسط بينها، ثم في الأخير نقوم بتطبيق القانون.

- **ترتيب قيم المتغير** (x_i):

$$\rightarrow \; 2،2 ،5 ،7$$

- **ترتيب قيم المتغير** (y_i):

$$\rightarrow \; 0 ،2 ،3 ،9$$

(x_i)	(y_i)	**ترتيب** (x_i)	**ترتيب** (y_i)	d_i	$(d_i)^2$
5	3	3	3	0	0
2	2	1,5	2	0,5-	0,25
7	9	4	4	0	0
2	0	1,5	1	0,5	0,25
/	/	/	/	0	0,5

ثم نطبق القانون التالي:

$$r_S = 1 - \frac{6\sum di^2}{n(n^2-1)} = 1 - \frac{6\times(0,5)}{4(16-1)} = 1 - \frac{3}{60} = 1 - 0,05 = 0,95$$

3– أسلوب الانحدار

بعد تقرير وجود علاقة تربط بين متغيرين أو أكثر، والتعرف على طبيعة وقوة العلاقة بينهما، تأتي الخطوة الثانية وهي صياغة نموذج رياضي بين المتغيرات يدعى بنموذج الانحدار.

1.3– تعريف

تشرح نماذج الانحدار المتغير التابع بواسطة متغير أو مجموعة من المتغيرات المستقلة (المفسرة) في أبسط أشكالها، ويمكن صياغتها على النحو التالي:

$$y = f(x)$$

إذن فالانحدار هو عبارة عن أسلوب إحصائي يستخدم لصياغة معادلة رياضية يمكن بها قياس أثر أحد المتغيرين على الآخر، ويعتبر من أسهل وأبسط نماذج التقدير والتحليل والتنبؤ.

2.3- الصياغة العامة للنموذج

تتطلب صياغة النموذج تحديد المتغير التابع ويرمز له بالرمز (y) والمتغيرات المستقلة ويرمز لها بالرمز $(x_1, x_2, x_3........x_n)$، ويكون النموذج **خطيا** أو **غير خطي** فنماذج الإنحدار تأخذ أشكالا مختلفة، حيث عجزت النظرية الاقتصادية من تحديد شكل الدالة (خطي أو لا)، ويتم التخلص من هذه المشكلة عن طريق **الأشكال البيانية** بواسطة رسم بيان مزدوج بين المتغير التابع والمستقل.

ملاحظة:

بافتراض أن العلاقة التي تربط بين المتغير التابع والمستقل علاقة خطية، فإذا كان هناك متغير تابع واحد وآخر مستقل فيطلق على هذا النموذج **الانحدار الخطي البسيط**، أما إذا كان هناك متغير تابع ومجموعة من المتغيرات المستقلة فيطلق عليه **بالانحدار الخطي المتعدد**.

3.3- الصياغة الرياضية للنموذج

$$y = f(x) \Rightarrow\Rightarrow y_i = \alpha + \beta.x_i$$

حيث: (y_i): المتغير التابع؛

(x_i): المتغير المستقل (المفسر)؛

(α): ثابت الانحدار، وهو قيمة المتغير التابع (y_i) لما يكون المتغير المستقل معدوم $(x_i = 0)$؛

(β): معامل الانحدر، وهو مقدار التغير في المتغير التابع (y_i) لما يتغير المتغير المستقل (x_i) بوحدة واحدة ← أثر المتغير المستقل (x_i) غلى المتغير التابع (y_i).

ومما سبق نلاحظ أن معاملات الانحدار مجهولة (α) و(β)، والإشكال هنا كيف يتم استخراج قيمها حسابيا؟

في الواقع لا يمكننا تحديد قيمة فعلية لكل من (α) و(β) لإيجاد قيمة المتغير التابع الحقيقية أو لتحديد العلاقة بين المتغير التابع والمستقل، ولكن يمكننا الحصول على قيم تقديرية تقريبية نرمز لها بـ $(\hat{y})$ و $(\hat{\alpha})$ و$(\hat{\beta})$، حيث تمثل القيم الآخيرة كل من القيم المقدّرة المناظرة للمشاهدة الفعلية (y_i)، والمقدّرة للمعلمة المجهولة الحقيقية (α) والمقدّرة للمعلمة الحقيقية المجهولة $(\beta.)$ على التوالي.

ولايجاد أحسن المقدّرات نستخدم طريقة المربعات الصغرى، التي تسعى إلى تقليل ابتعاد القيم الحقيقية عن القيم المقدرة لها، بمعنى تجعل مجموع مربعات الأخطاء أقل ما يمكن.

بافتراض أن النموذج الفعلي الحقيقي هو:

$$y_i = \alpha + \beta.x_i$$

والنموذج المقدّر:

$$^{1}\ \hat{y}_i = \hat{\alpha} + \hat{\beta}.x_i + u_i$$

فإن: $(\hat{\alpha})$ و$(\hat{\beta})$ هما اللذان يجعلان مجموع مربعات الأخطاء العشوائية أقل ما يمكن:

$$MIN \quad \sum u_i^2 = MIN \quad \sum (y_i - \hat{y}_i)^2$$

ولايجاد $(\hat{\alpha})$ و$(\hat{\beta})$ نشتق العلاقة أعلاه:

$$MIN \quad \sum u_i^2 = MIN \quad \sum (y_i - \hat{y}_i)^2 \Rightarrow \begin{cases} \dfrac{\delta \sum (y_i - \hat{y}_i)^2}{\delta \hat{\alpha}} = 0 \\ \text{و} \\ \dfrac{\delta \sum (y_i - \hat{y}_i)^2}{\delta \hat{\beta}} = 0 \end{cases}$$

لتسهيل عملية الاشتقاق يجب علينا تبسيط العبارة التالية:

$$\sum (y_i - \hat{y}_i)^2 = \sum \left[y_i - (\hat{\alpha} + \hat{\beta}.x_i) \right]^2 = \sum \left[y_i^2 - 2y_i(\hat{\alpha} + \hat{\beta}.x_i) + (\hat{\alpha} + \hat{\beta}.x_i)^2 \right] =$$

$$\sum \left[y_i^2 - 2\hat{\alpha}.y_i - 2\hat{\beta}.x_i y_i + \hat{\alpha}^2 + 2\hat{\alpha}.\hat{\beta}.x_i + \hat{\beta}^2 x_i^2 \right] ....(I)$$

الآن نقوم باشتقاق العبارة السابقة (I). بالنسبة لـ $(\hat{\alpha})$ و$(\hat{\beta})$

$$\frac{\delta \sum (y_i - \hat{y}_i)^2}{\delta \hat{\alpha}} = 0 \Rightarrow \sum \left(-2y + 2\hat{\alpha} + 2\hat{\beta}.x_i \right) = 0 \Rightarrow -2\sum y_i + 2\hat{\alpha}.n + 2\hat{\beta} \sum x_i$$
$$\Rightarrow -2\left(\sum y_i - \hat{\alpha}.n - \hat{\beta} \sum x_i \right) \Rightarrow \sum y_i - \hat{\alpha}.n - \hat{\beta} \sum x_i = 0$$

لدينا:

$$\overline{X} = \frac{\sum X_i}{n} \Rightarrow \sum X_i = \overline{X}n$$
$$\overline{Y} = \frac{\sum Y_i}{n} \Rightarrow \sum Y_i = \overline{Y}n$$

إذن:

$$\boxed{\overline{Y}n - \hat{\alpha}.n - \hat{\beta}\overline{X}n = 0 \Rightarrow \overline{Y} - \hat{\alpha}. - \hat{\beta}\overline{X} = 0 \Rightarrow \hat{\alpha} = \overline{Y} - \hat{\beta}.\overline{X} ....(II)}$$

[1] (u_i) هي عبارة عن قيمة عشوائية أو الخطأ العشوائي (البواقي) تنتج انطلاقا من الفارق بين القيم الفعلية الحقيقية للتابع والقيم المقدرة له.

$$\frac{\delta\sum\left(y_i-\hat{y}_i\right)^2}{\delta\hat{\beta}}=0\Rightarrow\sum\left(-2x_iy_i+2\hat{\alpha}.x_i+2\hat{\beta}.x_i^2\right)=0\Rightarrow-2\sum x_iy_i+2\hat{\alpha}\sum x_i+2\hat{\beta}\sum x_i^2=0$$
$$\Rightarrow-2\left(\sum x_iy_i-\hat{\alpha}\sum x_i-\hat{\beta}\sum x_i^2\right)=0\Rightarrow\sum x_iy_i-\hat{\alpha}\sum x_i-\hat{\beta}\sum x_i^2=0....(III)$$

نقوم بتعويض (II). في (III).:

$$\sum x_iy_i-\hat{\alpha}\sum x_i-\hat{\beta}\sum x_i^2=0\Rightarrow\sum x_iy_i-\left(\overline{Y}-\hat{\beta}.\overline{X}\right)\sum x_i-\hat{\beta}\sum x_i^2=0\Rightarrow$$
$$\sum x_iy_i-\overline{Y}\sum x_i+\hat{\beta}.\overline{X}\sum x_i-\hat{\beta}\sum x_i^2=0\Rightarrow\sum x_iy_i-\overline{Y}\sum x_i=\hat{\beta}\sum x_i^2-\hat{\beta}.\overline{X}\sum x_i$$
$$\Rightarrow\hat{\beta}=\frac{\sum x_iy_i-\overline{Y}\sum x_i}{\sum x_i^2-.\overline{X}\sum x_i}$$

لدينا:

$$\overline{X}=\frac{\sum X_i}{n}\Rightarrow\sum X_i=\overline{X}n$$

إذن:

$$\boxed{\hat{\beta}=\frac{\sum x_iy_i-\overline{Y}\,\overline{X}n}{\sum x_i^2-.\overline{X}\,\overline{X}n}=\frac{\sum x_iy_i-n\overline{X}\,\overline{Y}}{\sum x_i^2-.n\overline{X}^2}}$$

وهناك علاقة أخرى لإيجاد معامل الانحدار: (طريقة الانحرافات):

$$r_P=\frac{\sum\left(X_i-\overline{X}\right)\left(Y_i-\overline{Y}\right)}{\sum\left(X_i-\overline{X}\right)^2}$$

4.3- القدرة التفسيرية للنموذج

يتم الحكم على القدرة التفسيرية لنموذج الانحدار من خلال معامل التحديد (R^2)، وهو مقياس يوضح نسبة التغير في المتغير التابع (y_i) الذي سببها التغير في المتغير المستقل (x_i)، بمعنى قياس مدى قدرة المتغير المستقل على شرح التغيرات التي تحدث في المتغير التابع، ويمكن حسابه كالتالي:

$$R^2=r^2$$

حيث: (r^2): معامل الارتباط.

مثال: توفرت لدى باحث اقتصادي بيانات احصائية حول الدخل النقدي (x_i) والاستهلاك (y_i) لعينة من الأسر:

(x_i)	(y_i)	$(X_i - \overline{X})$	$(Y_i - \overline{Y})$	$(X_i - \overline{X})(Y_i - \overline{Y})$	$(X_i - \overline{X})^2$
5	31	0	1	0	0
11	40	6	10	60	36
4	30	1-	0	0	1
5	34	0	4	0	0
3	25	2-	5-	10	4
2	20	3-	10-	30	9
30	180	/	/	100	50

المطلوب:

- أوجد معادلة الانحدار الخطي البسيط؛ مع اعطاء تفسير معاملات الانحدار؛
- أحسب معامل التحديد، وفسره؟
- أوجد قيمة الاستهلاك عندما يكون الدخل يساوي 20 دينار.

الحل:

- **إيجاد معادلة الانحدار الخطي البسيط:**

طريقة الانحرافات:

$$\hat{\beta} = \frac{\sum (X_i - \overline{X})(Y_i - \overline{Y})}{\sum (X_i - \overline{X})^2} = \frac{100}{50} = 2$$

حساب المتوسطات:

$$\overline{X} = \frac{\sum X_i}{n} = \frac{30}{6} = 5$$

$$\overline{Y} = \frac{\sum Y_i}{n} = \frac{180}{6} = 30$$

الطريقة المختصرة:

(x_i)	(y_i)	$(x_i y_i)$	(x_i^2)	(y_i^2)
5	31	155	25	961
11	40	440	121	1600
4	30	120	16	900
5	34	170	25	1156
3	25	75	9	625
2	20	40	4	400
30	180	1000	200	5642

$$\hat{\beta}=\frac{\sum x_i y_i - n\overline{X}\,\overline{Y}}{\sum x_i^2 - .n\overline{X}^2}=\frac{1000-6(5)(30)}{200-6(25)}=\frac{100}{50}=2$$

أما ثابت الانحدار:

$$\hat{\alpha}=\overline{Y}-\hat{\beta}.\overline{X}=30-2(5)=20$$

إذن: معادلة الانحدار الخطي البسيط هي:

$$\hat{y}=20+2x_i$$

التفسير:

$(\alpha=20)$: وهو قيمة الاستهلاك (y_i) لما يكون الدخل المتاح معدوم $(x_i=0)$؛

$(\beta=2)$: هو مقدار التغير في الاستهلاك (y_i) لما يتغير الدخل المتاح (x_i) بوحدة واحدة ⟵ إذا تغير الدخل المتاح بوحدة واحدة يتغير الاستهلاك بنفس الاتجاه بـ 2.

- **ايجاد معامل التحديد، مع التفسير:**

$$R^2=r^2=(0{,}91)^2=0{,}83$$

أولا يجب علينا ايجاد معامل الارتباط:

$$r_P=\frac{\sum X_i Y_i - n\overline{X}\,\overline{Y}}{\sqrt{\sum X_i^2-n\overline{X}^2}\sqrt{\sum Y_i^2-n\overline{Y}^2}}=\frac{1000-6(5)(30)}{\sqrt{200-6(25)}\sqrt{5642-6(900)}}=\frac{100}{\sqrt{50}\sqrt{242}}=0{,}91$$

التفسير:

$(R^2=0{,}83)$ معناه أن المتغير المستقل (الدخل المتاح) **يفسّر** نسبة (83%) من التغيرات التي تحدث في المتغير التابع (الاستهلاك) ، والباقي حوالي (17%) ترجع إلى عوامل أخرى.

- **إيجاد قيمة الاستهلاك عندما يكون الدخل يساوي 20 دينار:**

يمكننا التنبؤ بقيمة المتغير التابع عند مستويات مختلفة للمتغير المستقل مباشرة في معادلة الانحدار، ومنه:

$$\hat{y}=20+2(20)=60$$

تمارين الفصل

التمرين الأول:

من أجل معرفة العلاقة بين الرضا الوظيفي ومستوى الدخل في مؤسسة معينة، قمنا باختيار عينة مكونة من 5 عمال، والجدول التالي يبين ذلك: (تم قياس مستوى الرضا الوظيفي على مقياس 100%).

العامل	1	2	3	4	5
الدخل	350	200	250	200	550
علامة الرضا	65	70	70	60	70

المطلوب:

- أحسب معامل الارتباط الرتبي سبيرمان، ثم علق على النتيجة.

التمرين الثاني:

تقدم عدد كبير من المترشحين للتوظيف في احدى الشركات الكبرى، وكانت هذه الشركة توظف الشخص بعد اجتيازه كل من الامتحان الكتابي والمقابلة الشخصية، ونظرا للتوافد الكبير للمترشحين قرر مسؤول الشركة الاكتفاء بالامتحان الكتابي، بافتراض انه يوجد علاقة قوية بين الامتحان الكتابي والمقابلة الشخصية، ولاختبار هذا الافتراض تم أخذ عينة مكونة من 8 مرشحين، ثم سجلت علاماتهم الوصفية في كل من الامتحان والمقابلة، والجدول التالي يبين ذلك:

المترشح	1	2	3	4	5	6	7	8
الامتحان	ممتاز	ضعيف	جيد	مقبول	جيد جدا	ممتاز	مقبول	جيد جدا
المقابلة الشخصية	جيد	جيد	جيد جدا	ضعيف	ضعيف	ممتاز	جيد	مقبول

المطلوب:

- ما هو مؤشر الارتباط المناسب في هذه الحالة، ولماذا؟ وعلى ماذا يعتمد في حسابه؟
- أحسب قيمة هذا المؤشر، ماذا تستنتج؟
- بناءا على النتيجة السابقة هل توقع مسؤول الشركة صحيح؟ ماذا تقترح.

التمرين الثالث:

الجدول التالي يبين لنا الكميات المعروضة من سلعة ما وسعرها في السوق:

السعر	5	5	10	15	20	20
الكمية المعروضة	16	22	26	27	29	30

المطلوب:

- ما هو مؤشر الارتباط المناسب في هذه الحالة، ولماذا؟ وعلى ماذا يعتمد في حسابه؟
- هل هناك علاقة بين الكمية المعروضة وسعرها؟ وما نوعها؟
- أوجد معادلة الانحدار الخطي البسيط وفسرها.

التمرين الرابع:

إذا كانت لدينا المعطيات التالية، حيث X_i هو متغير سنوات خبرة العمال و Y_i يمثل كمية الإنتاج لمؤسسة ما:

$$n = 12 ،\ \sum X_i = 72 ، \sum Y_i = 96 ، \sum X_i Y_i = 621 ، \sum X_i^{\,2} = 478 ، \sum Y_i^{\,2} = 828$$

المطلوب:

- هل هناك علاقة بين سنوات الخبرة للعمال وكمية الإنتاج لهذه المؤسسة، وما نوعها؟
- حدد أي المتغيرين المتغير المستقل والمتغير التابع، برر اجابتك؛
- قم بتقدير هذه العلاقة بافتراض وجود علاقة خطية بينهما.

التمرين الخامس:

إذا كانت لدينا المعطيات التالية وبافتراض أن (y_i) هو المتغير التابع:

إذا كانت لدينا المعطيات التالية: $n = 10 ،\ \sum X_i = 261 ، \sum Y_i = 304 ، \sum X_i Y_i = 8286 ، \sum X_i^{\,2} = 7127 ،$

$$\sum Y_i^{\,2} = 9734$$

المطلوب:

- قدّر نموذج الانحدار الخطي البسيط؛
- أوجد معامل التحديد وفسره.

التمرين السادس:

باحث اقتصادي يهتم بدراسة العلاقة بين الناتج المحلي الاجمالي والاستثمار المحلي خلال الفترة 2018-2022:

السنوات	2018	2019	2020	2021	2022
الناتج المحلي	240	249	304	345	355
الاستثمار	58	69	81	94	125

المطلوب:

- علق على شكل الانتشار؛
- أحسب معامل الارتباط الخطي البسيط (بيرسون) باستخدام طريقة الانحرافات والطريقة المختصرة لها؛
- أوجد معادلة الانحدار الخطي البسيط، ثم قم بتفسير معاملات الانحدار (α) و(β)؛
- أوجد القدرة التفسيرية للنموذج، ماذا تستنتج.

التمرين السابع:

- برهن أن:

$$\frac{\sum\left(X_i-\overline{X}\right)\left(Y_i-\overline{Y}\right)}{\sum\left(X_i-\overline{X}\right)^2}=\frac{\sum x_i y_i-n\overline{X}\,\overline{Y}}{\sum x_i^2-.n\overline{X}^2}$$

قائمة المراجع:

مراجع باللغة العربية

الكتب:

- أحمد عبد السميع طيبة، مبادئ الإحصاء ، دار البداية للنشر والتوزيع، الطبعة الألى، عمان، الأردن، 2008.
- بوحفص عبد الكريم، الإحصاء المطبق في العلوم الاجتماعية والإنسانية – مدعم بتطبيقات وتمارين محلولة–، ديوان المطبوعات الجامعية، الجزائر، 2006.
- جلاطو جيلالي، الإحصاء مع مسائل وتمارين محلولة، ديوان المطبوعات الجامعية، الجزائر، 2005.
- جلال صياد، عادل سمرة، مبادئ الإحصاء ، دار حافظ للنشر والتوزيع، الطبعة الثانية، المملكة العربية السعودية، 2003.
- سعدي شاكر حمودي، مبادئ علم الإحصاء وتطبيقاته، دار الثقافة للنشر والتوزيع، عمان، الأردن، 2009.
- عبد الحميد عبد المجيد البلداوي، أساليب الإحصاء، دار وائل للنشر والتوزيع، الطبعة الأولى، عمان، الأردن، 2009.
- عزوز عبد الرزاق، الكامل في الإحصاء – دروس مفصلة، تمارين ومسائل مع الحلول–، جزء 1، ديوان المطبوعات الجامعية، الجزائر،2010.
- علي أحمد السقاف، الإحصاء الوصفي والاستدلالي، إصدارات المركز الديمقراطي العربي، الطبعة الأولى، برلين، ألمانيا، 2020.
- كمال فليفل، فتحي حمدان، الإحصاء، دار المناهج للنشر والتوزيع، الطبعة الأولى، عمان، الأردن، 2013.
- محمد حسين محمد رشيد، الإحصاء الوصفي والتطبيقي والحيوي، دار صفاء للنشر والتوزيع ، الطبعة الأولى، عمان، الأردن، 2008.
- محمد صبحي أبو صالح، مبادئ الإحصاء، دار اليازوري العلمية للطباعة والنشر، عمان، الأردن، 2007.
- مراد كمال عوض، أساسيات الإحصاء، دار البداية للنشر والتوزيع، الطبعة الأولى، عمان، الأردن، 2010.
- محمد راتول، الاحصاء الوصفي، ديوان المطبوعات الجامعية، الجزائر ، 2005.

المطبوعات الجامعية:

- بولعباس مختار، محاضرات في الاحصاء الوصفي، جامعة ابن خلدون، تيارت، 2018.
- حيدوشي عاشور، محاضرات في الإحصاء الوصفي، جامعة أكلي محند أولحاج، البويرة، 2016.
- موسى عبد الناصر، دروس في الاحصاء الوصفي، جامعة محمد خيضر، بسكرة، 2007.
- الحاج بوروبة، دروس في الاحصاء الوصفي، جامعة ابن باديس، مستغانم، 2008.