جامـــــــعة وهـــران 2
كلية العلــوم الاقتصادية، التجارية وعلوم التسيير

مطبوعة

الإحصاء 1

لطلبة السنة الأولى (ل.م.د) جذع مشترك
ميدان العلوم الاقتصادية، التسيير والعلوم التجارية

مقدمة من طرف :

السيد: شنوف صادق
الرتبة: أستاذ محاضر " أ "

مارس 2022

مقدمــة

يعتبر الإحصاء من العلوم القديمة التي رافقت الإنسان عبر التاريخ في تطوره وتسيير شؤونه. وكانت فكرة التعداد Recensement فقط، وقد ازداد استعماله لما شعرت بعض الحضارات والدول بحاجاتها إلى معرفة بعض البيانات العددية والمعطيات الرقمية عن عدد سكانها لعدة أهداف كتحصيل الضرائب ومعرفة مقدار ثرواتهم الزراعية والمعدنية وإحصاء إمكانياتها البشرية والمادية من أجل الدفاع عن نفسها ومعرفة احتياجاتها في حالتي السلم والحرب.

تطور الإحصاء من مجرد مفهوم الحصر والعد إلى أن أصبح علما له قواعده ونظرياته، فقد اكتسب الإحصاء ابتداء من القرنين السابع عشر والثامن عشر ميلادي طابعه العلمي، حيث يشمل جميع الطرق العلمية والوسائل والتقنيات والأدوات الحديثة التي تساعد الباحث على اتخاذ القرارات السليمة إزاء مشكلات تتحكم فيها العشوائية، من خلال استخدام مقاييس كمية، و يرجع الفضل في ذلك إلى الكثير من علماء الإحصاء من أمثال فريديريك جاوس F.Gauss و كارل بيرسون Karl Pearson الذي يلقب بأب الإحصاء و يول Yule و أرنولد فيشرR.A.Fisher و سبيرمان Spearman و غيرهم من علماء الإحصاء و الرياضيات.

وقد أصبح الإحصاء في الوقت الحاضر يعالج بشكل رئيسي المجالات الكمية للظواهر الاقتصادية، الاجتماعية...وغير ذلك باستخدام الطرق والأدوات الإحصائية المناسبة، حيث أن التقدم التقني وتكنولوجيات الإعلام الآلي مكن من توفير الوقت والجهد في استخلاص النتائج الحسابية.

هذه المطبوعة عبارة عن مجموعة من المحاضرات الخاصة بمادة الإحصاء 1، الذي يتوجه أساسا لطلاب السنة الأولى (ل.م.د) التعليم القاعدي (Socle commun) لميدان العلوم الاقتصادية، التسيير و العلوم التجارية (Domaine SEGC) .

نقدم هذه المطبوعة في الإحصاء 1 آملين أن تكون عونا لطلاب السنة الأولى. إذ حاولنا عرض مختلف الطرق المتعلقة بالإحصاء، اعتمادا على مجموعة من المراجع التي تهتم بهذا الموضوع، وقد قسمناها إلى ستة فصول، كل فصل منها مدعم بكثير من الأمثلة التوضيحية والتطبيقية التي تمكن من زيادة المكتسبات لدى الطالب.

فقد تناولنا في الفصل الأول تطور مفهوم علم الإحصاء (لمحة تاريخية) وبعض المفاهيم الأساسية في الإحصاء، وفي الفصل الثاني كيفية إعداد جداول التوزيع التكراري. وفي الفصل الثالث تناولنا كيفية العرض البياني لها.

أما في الفصل الرابع فقد تناولنا فيه النزعة المركزية وطرق حساب مقاييسها بالنسبة للبيانات الغير المبوبة والمبوبة، وكذلك العلاقة بين المتوسطات.

وفي الفصل الخامس تعرضنا إلى مقاييس التشتت المطلقة والنسبية وطرق المقارنة بين تشتت الظواهر.

إضافة الى فصل سادس، حيث تعرضنا الى ملحق الرياضيات حيث ذكرنا ببعض الخواص الجبرية لدليل الجمع Σ ودليل الضرب (الجداء) π .

ونحن نقدم هذه المطبوعة وكل رجاؤنا أن يجد فيها الطلبة ما يساعدهم على الإلمام بمبادئ الإحصاء 1 وتعميق معارفهم ومكتسباتهم والتحضير الجيد لمختلف الامتحانات التي تخص هذه المادة.

ولا يفوتني في الأخير أن أشكر لكل ناقد نصيحته سلفا تلك النصيحة التي تساعد في التطوير والتحسين.

والله ولي التوفيق

الفصل الأول

مفاهيم عامة حول الإحصاء

I. **تطور مفهوم علم الإحصاء**:

الإحصاء هو علم ينتمي إلى فرع الرياضيات، يُعتبر علم الإحصاء في الوقت الحالي واحد من أهم العلوم الحديثة التي تلعب دورا حيويا في كثير من العلوم والدراسات المختلفة حيث يُستخدم كمنهج للبحث في مختلف المجالات العلمية (علوم الفيزياء، العلوم الاجتماعية، الاقتصادية، السياسية، الإنسانية، الطبية...)، وبالتالي يساعد على الفهم الأفضل للمشكلة المدروسة والوصول إلى الحلول الموضوعية وأخذ القرار السليم.

1) **لمحة تاريخية**:

لقد ظهرت أهمية علم الإحصاء منذ زمن بعيد) مع ظهور الإنسانية) ويمكن إظهارها عبر ثلاث فترات.

أ -قديما: اقتصر مفهوم الإحصاء على التعداد (Le dénombrement).

عندما نشأت الدول العظمى (الفرس، الروم، الإغريق، مصر...) أصبح ضروريا (تعداد) جمع معلومات تخص عدد أفراد المجتمع، ممتلكاتهم، وتوزيعهم من أجل تقييم القوة العسكرية والدفاعية لخوض الحروب وجمع الضرائب.

في هذه الفترة من الزمن كان الإحصاء مقتصرا على الحقائق الرقمية الخاصة بالدولة.

ب - مع انتشار الإسلام: تطور مفهوم الإحصاء بحيث أصبح يخص كل فرد من المجتمع. يقوم كل شخص بتعداد الممتلكات الخاصة به (نقدية، زراعية، حيوانية) وهذا من أجل إخراج الزكاة، وفي هذه الفترة توسع مفهوم الإحصاء بفضل علماء عرب ومسلمين، استعملوا الإحصاء في بحوثهم العلمية ومن أشهرهم:

- محمد بن موسى الخوارزمي (354 ه /781 م – 430 ه /1040 م)
- ابن الهيثم (أبو علي الحسن بن الحسن بن الهيثم) (354 ه /965 م – 430 ه /1040 م)
- عبد الرحمن محمد ابن خلدون (732 ه /1332 م – 808 ه /1406 م)

ج- نهاية القرن الثامن عشر: انطلاقا من البحوث العلمية العربية في الرياضيات تمكّن باحثون أوروبيون من تطوير علم الإحصاء بإدخال في دراستهم قوانين الاحتمالات. و يرجع الفضل في ذلك الى كثير من علماء الإحصاء أمثال فريدريك قوس Gauss و كارل بيرسون Karl Pearson الذي يلقب بأب الإحصاء و يول Yule و أرنولد فيشر R.A. Fisher و لابلاس Laplace و سبيرمان Spearman وغيرهم من علماء الإحصاء و الرياضيات.

منذ هذا الوقت أصبحت لعلم الإحصاء أهمية كبيرة ويُستعمل في كل ميادين الحياة.

1) **تعريف علم الإحصاء:**

هناك مفهومان لكلمة الإحصاء. يجب أن نفرق بين كلمة إحصاء في المفرد وكلمة إحصائيات في الجمع.

<u>المفهوم الأول:</u> وهو مفهوم عام يتصل بلغة الأرقام، وهذا المدلول يمثل وصفا عدديا للوجهات الكمية للظواهر المدروسة. كإحصاءات الإنتاج، إحصاءات العمال أو إحصاءات الطلبة في جامعة وهران 2 محمد بن أحمد... ونطلق على هذا المفهوم كلمة **"إحصاءات"** أو **" إحصائيات "**بالجمع فنقول إحصاءات الإنتاج، إحصاءات العمال، ...

اذن كلمة **"إحصاءات"** بالجمع هي مجموعة معلومات رقمية أو بيانات عددية حول موضوع ما.

<u>المفهوم الثاني</u>: وهو ذو مدلول علمي يعبر على الطرق العلمية التي تتبع في الحصول على المعلومات الكمية ودراسة خصائصها من أجل الوصول الى نتائج صحيحة في المستقبل. والإحصاء بهذا المعنى هو طريقة علمية للحصول على البيانات أو المعلومات. ونطلق على هذا المفهوم كلمة إحصاء بالمفرد.

اذن كلمة إحصاء بالمفرد هي مجموعة وسائل تقنية تُستعمل لجمع وعرض وتحليل البيانات الرقمية من أجل أخذ القرارات.

ومنه يُعرف علم الإحصاء بصفة عامة على أنه ذلك العلم الذي يستعمل وسائل علمية وتقنيات رياضية لجمع، تبويب وتنظيم وتلخيص وعرض البيانات العددية، ومن ثم تحليلها وفق طرق وأساليب علمية للحصول على الاستدلالات وأخذ القرارات السليمة.

ينقسم حاليا علم الإحصاء إلى قسمين رئيسيين.

<u>الإحصاء</u>

الإحصاء الاستدلالي الرياضي

- استعمال النتائج
- وضع الاحتمالات
- أخذ القرار

الإحصاء الوصفي

- جمع البيانات
- عرضها في جداول وفي رسومات
- تلخيص البيانات
- دراسة العلاقة الموجودة بين الظواهر
- مقارنة الظواهر

I. **التعبير الإحصائي:**

الإحصاء له تعبير خاص به.

1) **المجتمع الإحصائي: Population statistique:**

هو المجموعة التي تجري عليها الدراسة.

مثال1: – عمال مؤسسة – طلاب الجامعة – مخزون قطع الغيار في وقت معيّن – إنتاج السيارات في سنة معيّنة.

المجتمع المدروس يجب تحديده بصفة دقيقة ويُرمز له عادة ب N.

مثال2: عند دراسة حول سكان مدينة ما يجب تحديد إذا نأخذ بعين الاعتبار الطلاب في النظام الداخلي، المرضى في المستشفيات، المحبوسين إلى غير ذلك.

2) **الوحدة الإحصائية أو الفرد: L’unité ou individu statistique**

كل عنصر من المجتمع الإحصائي يُسمّى وحدة إحصائية أو فرد إحصائي: تُطلق كلمة وحدة إحصائية للدلالة على الكائن الواحد الملموس أو الغير الملموس (إنسان، حيوان، مؤسسة، مدينة...).

مثال1: إذا قمنا بدراسة حول الأجور في مؤسسة ما:

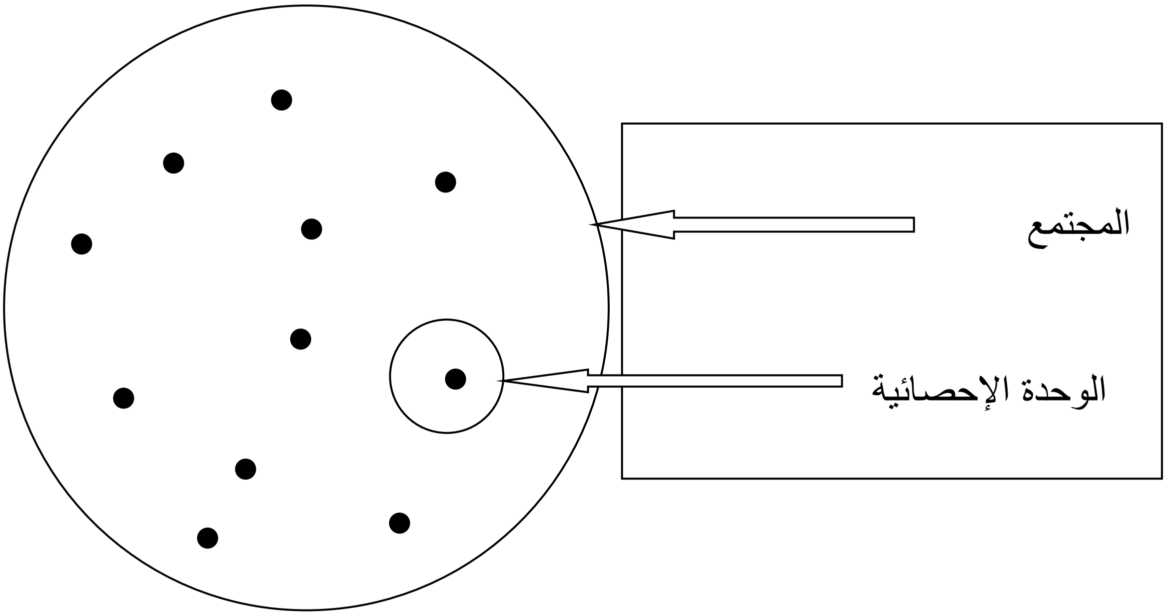
المجتمع هو مجموعة عمال المؤسسة.

الوحدة الإحصائية هي عامل واحد في المؤسسة.

مثال2: نقوم بدراسة حول بلدان القارة الإفريقية.

المجتمع المدروس :54 بلدا إفريقيا.

الوحدة الإحصائية: بلدا أفريقيا معينا واحدا.



المجتمع الإحصائي مكون من الوحدات الإحصائية. مجموع هذه الوحدات هو Nوهي المجتمع.

3) العيّنة: L'échantillon

في كثير من الأحيان تصعب دراسة المجتمع الإحصائي ككل بأكمله. ولهذا نقوم بدراسة جزء من المجتمع الذي يضم عددا ما من الوحدات الإحصائية المتشابهة في إحدى الصفات على الأقل، بحيث يمثل المجتمع تمثيلا جيدا. دراسة العيّنة توفر الوقت، الجهد والمادة.

4) الصفة المدروسة أو الخاصية: Le caractère (variable)

كل وحدة إحصائية يمكن وصفها بعدة خصائص هو الشيء الذي نريد دراسته.

مثال1: عند دراسة عمال مؤسسة يمكن أن تكون هذه الدراسة حول: – الجنس – السن – الحالة المدنية – الأجور – عدد الأطفال.

مثال2: عند دراسة مؤسسة ما يمكن أن تكون هذه الدراسة حول: – رقم الأعمال – الاستثمار– الربح ....

5) الكيفية: Les modalités

هي كل الحالات التي يمكن أن تأخذها الصفة (غير عددية أو عددية).

<u>مثال 1:</u> الجنس: هذه الصفة يمكن أن نجدها في حالتين إمّا ذكر أو أنثى.

الجنس
ذ M
أ F

هناك حالتين أي كيفيتين 2 modalités

<u>مثال 2:</u> الحالة العائلية
متزوج
مطلق
أرمل
أعزب

04 modalités

<u>مثال 3:</u> توزيع عمال المؤسسة حسب الكفاءة المهنية (C.S.P).

الصفة هي الكفاءة المهنية: الكيفيات
عامل بسيط
عامل متخصص
تقني
إطار

6) <u>طبيعة الصفة</u> : Nature du caractère

يمكن أن نميّز بين نوعين من الصفات أو الخاصيات.

- **<u>الصفة النوعية</u>: Caractère qualitatif**

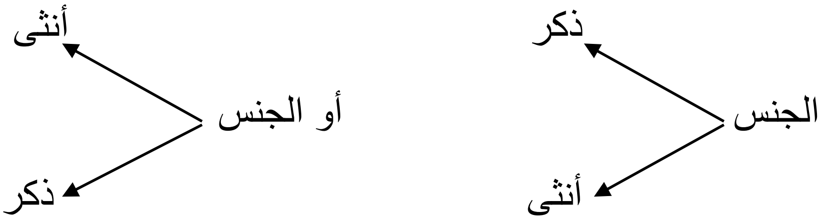
تكون الصفة نوعية لمّا تكون كيفياتها غير قابلة للقياس أي لا تكون هذه الكيفيات على شكل أرقام بل هي كلمات.

مثل: لون الأعين، الجنس، شعبة البكالوريا، الكفاءة المهنية، الحالة العائلية، الديانة، الجنسية...الخ.

ملاحظة : كيفيات صفة نوعية تكون إمّا إسمية (nominale)أو ترتيبية (ordinales).

- تكون الصفة نوعية إسمية لمّا لا يمكننا أن نرتب كيفياتها.

مثال 1:

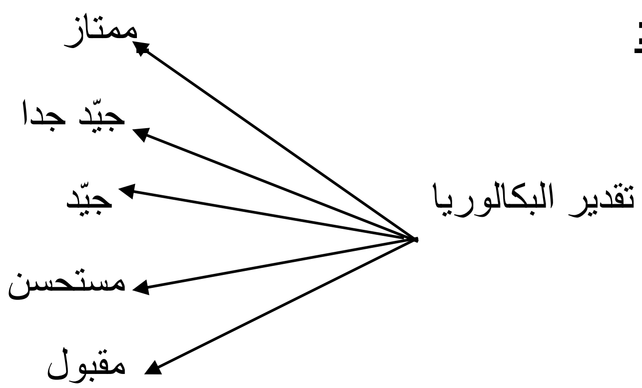


مثال 2: ترتيب مجموعة من الطلّاب حسب مكان الازدياد: – وهران – الجزائر – سطيف – مستغانم – بشار...

هذه الكيفيات لصفة مكان الازدياد لا تتبع أي ترتيب ⇐ الصفة نوعية إسمية.

- تكون الصفة نوعية ترتيبية لمّا يمكن ترتيب مختلف كيفياتها (يوجد ترتيب).

مثال 1:



مثال 2: دراسة حول نوعية العلاقة الموجودة بين الجيران في حي ما:

– جيّدة – حسنة – مقبولة – رديئة – عدوانية.

هذه الصفات نوعية ترتيبية لأن يمكن ترتيب كيفياتها.

- **الصفة الكمية أو المتغيّر:** **Caractère quantitatif ou variable**

تكون الصفة كمية لمّا تكون كيفيتها قابلة للقياس أي تكون على شكل أرقام أو قيم، تُكتب هذه الكيفيات بواسطة أرقام ولهذا نسميها أيضا متغير إحصائي.

مثال: القامة، الوزن، عدد الأطفال في الأسرة، أجور العمال، درجة الحرارة، المساحة، الإنتاج.

ملاحظة: تنقسم الصفات الكمية بدورها إلى قسمين:

- **متغير إحصائي منفصل أو متقطع:**

هي الصفات أو المتغيرات التي تأخذ قيم صحيحة ومحدودة في مجال معين.

مثال1: عدد الأطفال في الأسرة (0، 1،2، 3،...).

مثال2: عدد الغرف في المنزل (1، 2، 3، 4، 5).

مثال3: عدد التلاميذ في القسم (1، 2، 3، ....30).

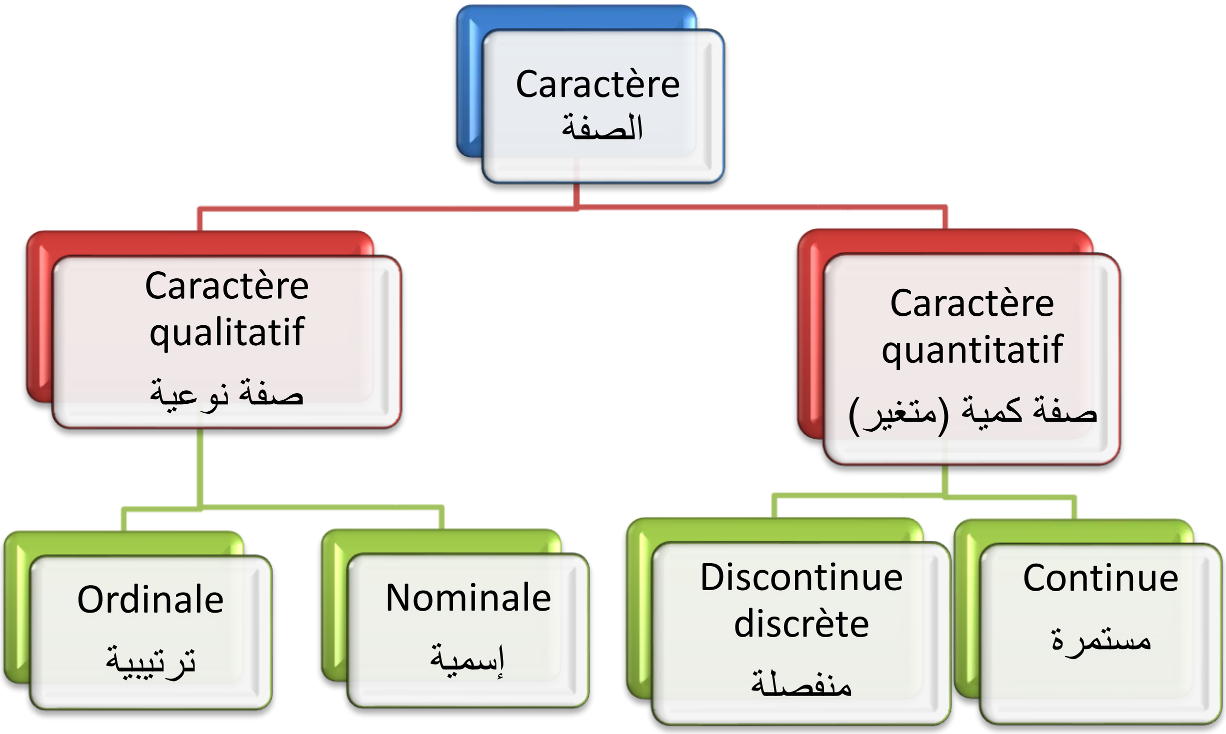
كل كيفية هي قيمة صحيحة موجبة ويمكننا معرفة مقدما عدد الكيفيات وهي قيم متباعدة عن بعضها البعض.

- **متغير إحصائي متصل أو مستمر:**

هي الصفات التي يمكنها أخذ كل القيم الممكنة في مجال معين. لا توجد أي قيمة في المجال لا يمكن أن يأخذها المتغير، القيم متصلة ببعضها البعض ولا يوجد بينها فواصل.

مثال: – الزمن – العمر – السرعة – الوزن – الأطوال – المساحات.

في هذه الحالة كيفيات القيم يجب قياسها ولا يمكننا تحديدها مقدما. يكون عدد القيم في غالب الأحيان كبير جدا، ولهذا نقسم مجال الدراسة إلى مجالات جزئية تسمى الفئات (Classes).



7) مصادر المعلومات و البيانات الإحصائية: Les sources statistiques

يعتمد الباحث للحصول على المعلومات الإحصائية الخاصة بدراسة معيّنة على مصدرين أساسيين.

- **المصدر المباشر:**

يتحصل الباحث على المعلومات بطريقة مباشرة أي يقوم الباحث بوسائله الخاصة بدراسة ميدانية على كل وحدات المجتمع أو عيّنة من المجتمع وباستعمال عدة طرق (كالاستبيانquestionnaire، البريد، الهاتف، استطلاع الرأيsondage، الأنترنت Internet...).

- **المصدر الغير مباشر:**

يتحصل الباحث على البيانات من الدراسات والتحقيقات المنشورة في الصحف، الكتب والوثائق الرسمية (الديوان الوطني للإحصائيات ONS).

الفصل الثاني

عرض وتبويب البيانات الإحصائية

مقدمة:

بعد تحديد موضوع البحث والمنهج المتبع في الدراسة، و بعد جمع المعلومات والبيانات الخاصة بهذه الدراسة، يأتي دور تصنيف وترتيب هذه المعطيات في جداول تسمى بالجداول الإحصائية.

الجدول الإحصائي عبارة عن صورة تنقل المعلومات، دون الإنقاص منها، من حالتها الأولى الى حالة جديدة تتسم بالتنظيم والترتيب والسهولة والوضوح.

وتختلف طرق ترتيب المعلومات في جداول إحصائية، باختلاف الأسلوب المستخدم و المنهج المتبع في الدراسة. كما تختلف الجداول الإحصائية باختلاف وتنوع المعطيات، كأن تكون كمية أو نوعية.

تشكل المعطيات الإحصائية التي يتحصل عليها الباحث عند القيام بدراسة ظاهرة ما، كتلة ضخمة خامة، غير منتظمة ويصعب دراستها أو استنتاج أي شيء منها. لذلك يلجأ الباحث إلى حصر وتصنيف هذه البيانات وعرضها بطريقة تساعد على فهمها وتحليلها في جداول أو على شكل رسوم بيانية.

I. - **العرض الجدولي** :

تعتبر الجداول الإحصائية وسيلة لتصنيف وترتيب وعرض الوحدات الإحصائية.

1. **جدول التوزيع التكراري:**

إذا كان لدينا مجتمع إحصائي N موصوف حسب صفة متكونة من k كيفية، يمكن عرض المعطيات داخل جدول وهذا بحساب عدد الوحدات الإحصائية التي تناسب كل كيفية (هذا العدد يسمى التكرار المطلق).

- يمكننا وضع جدول تكراري مهما كانت طبيعة الصفة (كمية أو نوعية).
- البيانات المصنفة والمرتبة داخل الجدول تشكل توزيع.
- يتكون جدول التوزيع التكراري من عمودين (أو سطرين).

التكرار n_i	الكيفية	
	صفة نوعية m_i	صفة كمية x_i
n_1	m_1	x_1
n_2	m_2	x_2
.	.	.
.	.	.
n_k	m_k	x_k
$N= \sum_{i=1}^{k} ni$	المجموع	

- التكرار n_i هو عدد المشاهدات لكل كيفية.
- كيفيات صفة نوعية يرمز لها ب m_i .
- كيفيات صفة كمية يرمز لها ب x_i .
- المجتمع الإحصائي N = مجموع التكرارات.
-

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

1-2. **العرض الجدولي لصفة نوعية :**

مثال 1 : تمثل السلسلة الإحصائية التالية التقدير المتحصل عليه في الامتحان ل 20 طالب.

D-B-E-C-D-B-D-C-E-A-B-E-C-D-B-D-D-A-E-C

A = ممتاز ، B = حسن جدا ، C = حسن ، D = مستحسن ، E = متوسط

المطلوب : وضع البيانات في جدول توزيع تكراري.

الحل :

ملاحظة : لما يكون عدد البيانات كبير جدا، نضيف إلى الجدول عمود وراء العمود الأول أين نقوم بعملية التفريغ حتى تسهل عملية حساب التكرارات.

في المثال 1 :

- المجتمع الإحصائي متكون من 20 طالب :N = 20.
- الصفة المدروسة هي التقدير.
- طبيعة الصفة هي نوعية ترتيبية.
- عدد الكيفيات (modalités) : 5 (A-B-C-D-E)

جدول التوزيع التكراري لصفة نوعية

Modalités الكيفيات m_i	التفريغ	التكرار n_i
A	II	$n_1 = 2$
B	IIII	$n_2 = 3$
C	~~IIII~~I	$n_3 = 5$
D	~~IIII~~ I	$n_4 = 6$
E	III	$n_5 = 3$
المجموع		$N= \sum_{i=1}^{5} ni = 20$

1-2. العرض الجدولي لصفة كمية :

1-1-2. المتغيرات الكمية المتقطعة (المنفصلة) :

عندما تكون لدينا بيانات على شكل صفة كمية متقطعة (متغير كمي متقطع) ولغرض تبويبها وعرضها في جدول توزيع تكراري يتم إعداد جدول مكون من ثلاثة أعمدة كما فعلنا سابقا. يخصص العمود الأول للمتغيرات التي نرمز لها ب x_i مع مراعاة ترتيبها وتصنيفها ترتيبا تصاعديا، والعمود الثاني للتفريغ البيانات والعمود الثالث للعد وكتابة تكرارات تلك المتغيرات التي نرمز لها بالرمز n_i.

مثال 2 : تمثل البيانات التالية عدد التلاميذ في كل قسم من أقسام مدرسة ابتدائية :

30-25-32-26-30-26-26-28-30-30-27-26-27-25-30

المطلوب: وضع البيانات في جدول توزيع تكراري.

الحل :

في المثال 2 :

- المجتمع الإحصائي متكون من 15قسم: N = 15.
- الصفة المدروسة هي التلاميذ.
- طبيعة الصفة هي كمية متقطعة (منفصلة).
- عدد قيم هذه الصفة هو 6.

جدول التوزيع التكراري لصفة كمية متقطعة

n_i التكرار	x_i
$n_1 = 2$	$x_1 = 25$
$n_2 = 4$	$x_2 = 26$
$n_3 = 2$	$x_3 = 27$
$n_4 = 1$	$x_4 = 28$
$n_5 = 5$	$x_5 = 30$
$n_6 = 1$	$x_6 = 32$
$N=\sum_{i=1}^{6} ni = 15$	**المجموع**

هذا الجدول، جدول تكراري كمّي متقطع أو منفصل وذلك لانفصال القيم عن بعضها البعض وعدم ارتباطها، وكلقيمة منفصلة عن التي تليها والتي تسبقها.

2-1-2. المتغيرات الكمية المستمرة (المتصلة):

في حالة وجود عدد كبير جدا من المشاهدات التي يمكن أن يأخذها المتغير، وحتى لا يكون الجدول مطّول وممّل يجب تشكيل فئات، تشمل كل فئة عدد ما من القيم. بمعنى إن إعداد جدول التوزيع التكراري يتم بتجميع أو توزيع مجموعة من القيم (المفردات) على عدد من المجموعات الجزئية تسمى فئات." Les classes" غالبا ما تكون متساوية الطول (طول الفئة ثابت) وتكون على شكل مجال. ويجب أن يتلاءم عدد الفئات مع إجمالي عدد المفردات المشاهدة، فلا يكون عددا قليلا ولا كثيرا للمحافظة على الهدف من عملية التوزيع والتجميع في فئات.

لتكوين جدول توزيع تكراري ذو فئات نتبّع الخطوات التالية:

– تحديد المدى العام :(Etendu) الذي يُرمز له ب (E)

$$E = Max_{(xi)} - Min_{(xi)}$$

– **تحديد عدد فئات مناسب :(NC : Nombre de classes)**

رغم أنه لا توجد هناك قاعدة ثابتة لتحديد عدد الفئات المناسب، وذلك لأن هذه العملية تتوقف على طبيعة البيانات وعدد المفردات بالإضافة إلى اجتهاد الإحصائي (الباحث) وخبرته زيادة على أخذ بعين الاعتبار الأهداف المرجوة من الدراسة.
يتم تحديد عدد الفئات حسب حجم العيّنة أو المجتمع المدروس أو عدد المشاهدات التي نرمز لها ب N.
يجب أن يكون عدد الفئات مناسبا لعدد المشاهدات، ولتفادي التقليل من دقة القياس ان عدد الفئات يفضل أن يكون بين 5 و15 فئة (وذلك حسب الإحصائيين).

عموما هناك طريقتين لتحديد عدد الفئات المتساوية:

- **إذا $N \leq 100$:** نحسب عدد الفئات باستعمال قانون YULE :

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

- **إذا $N > 100$:** نحسب عدد الفئات باستعمال قانون **STURGE:**

$$NCs = 1 + 3{,}322\ .(\log N)$$

N يمثل عدد المشاهدات.
NC عدد الفئات.

- **تحديد طول الفئة** : (a: amplitude de classes)

لدينا مدى عام نريد أن نقسمه إلى عدد معين من الفئات (مجالات) المتساوية الطول. بمعنى نبحث في مجال كل فئة (كل مجال جزئي) والذي يساوي المدى العام مقسوما على عدد هذه الفئات ويكون كالتالي:

$$\text{طول الفئة} = \frac{\text{المدى العام}}{\text{عدد الفئات}}$$

أو باستخدام الصيغة:

$$\mathbf{a = \frac{E+1}{NC}}$$

يجب التأكد من تحقيق الصيغة: $\mathbf{E \leq NC.a}$ لضمان توزيع كافة المشاهدات على الفئات.

- **إيجاد حدود الفئات** : (Limites de classes)

لكل فئة حدين، حد أدنى نرمز له بالرمز $(\mathbf{L_{i-1}})$ و**حد** أعلى نرمز له $(\mathbf{L_i})$. فمثلا كتابة حدود الفئة الأولى:$[\mathbf{L_0 - L_1}[$

$\text{الحد الأدنىللفئةالأولى} = \mathbf{L_0 = Min_{x_i}}$

هذا الحد يساوي أصغر قيمة من بين المشاهدات.

$\text{الحد الأعلى للفئةالأولى} = \mathbf{L_1 = L_0 + a}$

ثم نضيف للحد الأدنى طول الفئة للحصول على الحد الأعلى للفئة الأولى $(\mathbf{L_1})$ الذي يساوي الحد الأدنى للفئة الثانية وهكذا بالنسبة لبقية الفئات كما هو مبين من خلال المثال.

- **كتابة حدود كل الفئات الموالية** :

علما أن الحد الأعلى يساوي دائما الحد الأدنى للفئة السابقة زائد طول الفئة.

ملاحظة:

- إذا كانت الصفة المدروسة كمية منفصلة (متقطعة) نشكل فئات متباعدة أي منفصلة عن بعضها البعض، حيث لا توجد حدود مشتركة بينها: $[\mathbf{24-20}]$ ، $[\mathbf{19-15}]$، $[\mathbf{14-10}]$.
- أما إذا كانت الظاهرة كمية مستمرة نشكل فئات متداخلة فيما بينها، أي أن الحد الأعلى للفئة الأولى هو الحد الأدنى للفئة الموالية: ....$[\mathbf{40-30}]$، $]\mathbf{30-20}]$، $]\mathbf{20-10}]$

مثال 3 : تمثل البيانات التالية الدخل بالساعة (DA) ل 50 عاملا في المؤسسة "س".

-71-70-68-77-68-62-73-76-70-62-63-82-63-64-65-66-68-65-81-62-67-70-68-79
-73-62-68-68-80-65-62-65-64-85-65-64-67-65-70-67-66-63-72-62-75-69-63-63
.88-70

المطلوب :

1- ما هو المتغير المدروس وما طبيعته؟

2-قم بإعداد جدول توزيع تكراري ذو فئات متساوية (بطول ثابت) لهذا المتغير.

الحل :

في المثال 3 :

- المجتمع الإحصائي متكون من 50عاملا: N = 50
- الصفة المدروسة هي الدخل في الساعة (DA).
- طبيعة الصفة هي كمية مستمرة.
- بما أن عدد القيم معتبر، يستحسن تجميع هذه القيم داخل فئات

1- المتغير المدروس هو الدخل في الساعة، وهذا المتغير كمي قابل للقياس وبالتالي مستمر (متصل).

2- من أجل إعداد جدول توزيع تكراري نتبع الخطوات المذكورة سابقا كما يلي:

- **تحديد المدى العام (E) :**

$$\mathbf{E = Max_{(xi)} - Min_{(xi)} \Rightarrow E = 88 - 62}$$

$$\Rightarrow \mathbf{E = 26}$$

وهذا يعني أن المدى العام للدراسة هو 26 دينار جزائري، حيث أن أقل أجر بالساعة 62 دج وأقصى أجر بالساعة 88 دج وهو مجال الدراسة.

- **تحديد عدد الفئات (NC):**

لتحديد عدد الفئات نستخدم علاقة يول Yuleلأن $N = 50 \leq 100$

$$NC = 2{,}5\sqrt[4]{N} = 2{,}5\sqrt[4]{50} \Rightarrow \quad NC = 6{,}64 \quad \Rightarrow \boxed{NC \approx 7}$$

نقرب العدد إلى أقرب عدد صحيح لضمان تغطية كافة المشاهدات في فئات.

- **تحديد طول الفئات(a) :**

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

$$a = \frac{26+1}{7} \Rightarrow \mathbf{a} = 3.857 \approx 4$$

نقرب النتيجة إلى الوحدة القريبة.

- **إيجاد حدود الفئات : (Limites de classes)**

نبدأ بالفئة الأولى فمثلا كتابة حدود الفئة الأولى :$[\mathbf{L_0} - \mathbf{L_1}[$ثم التي تلي ...و هكذا.

الحد الأولللفئةالأدنى $= \mathbf{L_0} = \mathbf{Min_{x_i}} = 62$

هذا الحد يساوي أصغر قيمة من بين المشاهدات.

الحد الأعلى الأولللفئة $= \mathbf{L_1} = \mathbf{L_0} + \mathbf{a}$

$$\mathbf{L_1} = \mathbf{L_0} + \mathbf{a} = 62 + 4 = 66$$

حدود الفئة الأولى $[62 - 66[$

وعند عملية التفريغ وعد التكرار المقابل لهذه الفئة الأولى نقوم بعد عدد مرات تكرار كافة القيم من 62 إلى أقل من 66، ثم يمكن تحديد حدود الفئة الثانية بوضع الحد الأدنى للفئة يساوي الحد الأعلى للفئة السابقة وللحصول على الحد الأعلى يجب إضافة في كل مرة طول الفئة a إلى الحد الأدنى وهكذا...

جدول التوزيع التكراري لصفة كمية مستمرة

الفئات (الدخلDA)	التفريغ أو عمود العلامات التكرارية	التكرار n_i
[62-66[	𝍸𝍸𝍸 𝍸	20
[66-70[	𝍸 𝍸 II	12
[70-74[	𝍸 IIII	9
[74-78[	III	3
[78-82[	III	3
[82-86[	II	2
[86-90[	I	1
المجموع	/	N= $\sum_{i=1}^{7} \mathbf{ni} = 50$

– **مراكز الفئات :** بعد كتابة الفئات نحسب مركز كل فئة xi الذي هو عبارة عن الحد الأدنى + الحد الأعلى قسمة 2.

$$x_1 = \frac{L0+L1}{2} ، \quad x_2 = \frac{L1+L2}{2} ، \quad x_3 = \frac{L2+L3}{2}$$

– **جدول التوزيع التكراري النسبي و المئوي:**

يمكن حساب أنواع أخرى من التكرارات في جدول التوزيع التكراري وهذه التكرارات هي التكرار النسبي والتكرار المئوي.

❖ **التكرار النسبي:** الذي يرمز له ب (f_i) هو عبارة عن التكرار المطلق مقسوم على مجموع التكرارات.

يكون مجموع التكرارات النسبية يساوي دائما: 1.

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

$$\Sigma f_i = 1 \qquad \Rightarrow$$

التكرار المئوي: الذي يرمز له ب ($f_i\%$) هو عبارة عن التكرار النسبي مضروب في 100.

$$f_i\% = \frac{n_i}{\sum n_i} \times 100 = f_i \times 100$$

يكون مجموع التكرارات المئوية يساوي دائما:100% .

$$\Sigma f_{i\%} = 100$$

مثال: نأخذ المثال رقم 3 ونحسب التكرار النسبي والتكرار المئوي.

التكرار المئوي $f_i\%$	التكرار النسبي f_i	مراكز الفئات x_i	التكرار المطلق n_i	الفئات (الدخلDA)
0.4×100= 40	$\frac{20}{50}$= **0.4**	$\frac{62+66}{2}$= 64	20	**[62-66[**
24	0.24	68	12	**[66-70[**
18	0.18	72	9	**[70-74[**
6	0.06	76	3	**[74-78[**
6	0.06	80	3	**[78-82[**
4	0.04	84	2	**[82-86[**
2	0.02	88	1	**[86-90[**
100	$\sum f_i = 1$	/	$N= \sum_{i=1}^{7} ni = 50$	**المجموع**

II. التوزيعات التكرارية المتجمعة (الجداول المتجمعة):

من التوزيع التكراري المطلق يمكن معرفة عدد الوحدات التي تقابل قيمة معينة أو فئة معينة. أما لمعرفة عدد الوحدات التي تقع قبل أو بعد قيمة معينة أو فئة معينة، يجب حساب التكرارات التجميعية الصاعدة ($nc_i\uparrow$) والتكرارات التجميعية النازلة ($nc_{i\downarrow}$).

إن الفكرة الأساسية في التوزيعات التكرارية المتجمعة هي تجميع التكرارات (مطلقة أو نسبية أو مئوية):

- ❖ أمام الحد الأعلى لكل فئة وفي هذه الحالة نحصل على التكرار المتجمع الصاعد. تكون التكرارات في تصاعد مستمر.
- ❖ أمام الحد الأدنى لكل فئة وفي هذه الحالة نحصل على التكرار المتجمع النازل وتكون هذه التكرارات في نزول مستمر.

<u>مثال 4</u> : يمثل الجدول التالي عدد الغرف في كل منزل.

عدد الغرف x_i	عدد المنازل n_i	≤ $nc_i\uparrow$	≥ $nc_i\downarrow$	f_i	≤ $fc_i\uparrow$	≥ $fc_i\downarrow$
1	5	5	10	0,05	0,05	1
2	10	15	95	0,1	0,15	0,95
3	20	35	85	0,2	0,35	0,85
4	30	65	65	0,3	0,65	0,65
5	100	90	35	0,25	0,90	0,35
6	10	100	10	0,1	1	0,1
المجموع				1		

المطلوب:

1- أوجد عدد المنازل الذين لهم عدد غرف يقل أو يساوي 3.

2- أوجد عدد المنازل الذين لهم عدد غرف يزيد أو يساوي 4.

3- أوجد نسبة المنازل الذين لهم عدد غرف يقل أو يساوي 2.

4- أوجد نسبة المنازل الذين لهم عدد غرف يزيد أو يساوي 5.

الحل:

1- أوجد عدد المنازل الذين لهم عدد غرف يقل أو يساوي 3 هو 35.

2- أوجد عدد المنازل الذين لهم عدد غرف يزيد أو يساوي 4 هو 65.

3- أوجد نسبة المنازل الذين لهم عدد غرف يقل أو يساوي 2 هي : 0.15.100= 15%

4- أوجد نسبة المنازل الذين لهم عدد غرف يزيد أو يساوي 5 هي: 0.35.100 = 35%

مثال 5 : يمثل الجدول التالي تأخر العمال (m_n) عن الالتحاق بعملهم في بداية الفترة الزمنية في مؤسسة ما.

التأخر (دقائق) فئات	عدد العمال ni	الحد الأعلى -أقل من-	تكرار متجمع صاعد nci↑	الحد الأدنى -أكثر من-	تكرار نازل nci↓	fci%	fci%↑	fci%
[0-5[	4	أقل من 5	4	أكثر من 0	72	5	5	100
[5-10[	7	أقل من 10	11	أكثر من 5	68	10	15	95
[10-15[	10	أقل من 15	21	أكثر من 10	61	14	29	85
[15-20[	15	أقل من 20	36	أكثر من15	51	21	50	71
[20-25[	18	أقل من 25	54	أكثر من 20	36	25	75	50
[25-30[	12	أقل من 30	66	أكثر من 25	18	17	92	25
[30-35[	6	أقل من 35	72	أكثر من 30	6	8	100	8
Total	72					100		

المطلوب : أوجد

1- عدد العمال الذين تأخرهم يفوق 10 دقائق.

2- عدد العمال الذين تأخرهم يقل عن 20 دقيقة.

3- النسبة المئوية للعمال الذين تأخرهم يقل عن 25 دقيقة.

4- نسبة العمال الذين تأخرهم يقل عن 20 دقيقة.

5- عدد العمال الذين لهم تأخر ما بين 10 و25 دقيقة.

الحل :

1- عدد العمال الذين تأخرهم يفوق 10 دقائق هو: 61 عامل.
2- عدد العمال الذين تأخرهم يقل عن 20 دقيقة هو: 36 عامل.
3- النسبة المئوية للعمال الذين تأخرهم يقل عن 25 دقيقة هي: 75%.
4- نسبة العمال الذين تأخرهم يقل عن 20 دقيقة هي: 05 ($\frac{50}{100}$. أو $\frac{36}{72}$).
5- عدد العمال الذين لهم تأخر ما بين 10 و25دقيقة هو: 43 عامل.

من $nc_{i\uparrow}$: $\underline{54}$–$\underline{11}$ = 43.

من $nc_{i\downarrow}$: $\underline{61}$–$\underline{18}$ = 43.

الفصل الثالث

التمثيل البياني للتوزيعات التكرارية

إن تبويب البيانات في صورة منظمة تساعد على فهم الموضوع الذي نقوم بدراسته، لكن هذه الصورة لا تكفي أحيانا للتوضيح ولهذا يحاول الباحث استعمال البيانات بطريقة أخرى، وهي الرسوم البيانية حيث تساعد في تكوين فكرة سريعة ودقيقة عن الظاهرة المدروسة.

تختلف وسائل العرض البياني باختلاف طبيعة الصفة.

I. العرض البياني في حالة صفة نوعية:

أهم الرسوم البيانية المستعملة في هذه الحالة هي:

1- القضبان البيانية (Tuyau d'orgue) :

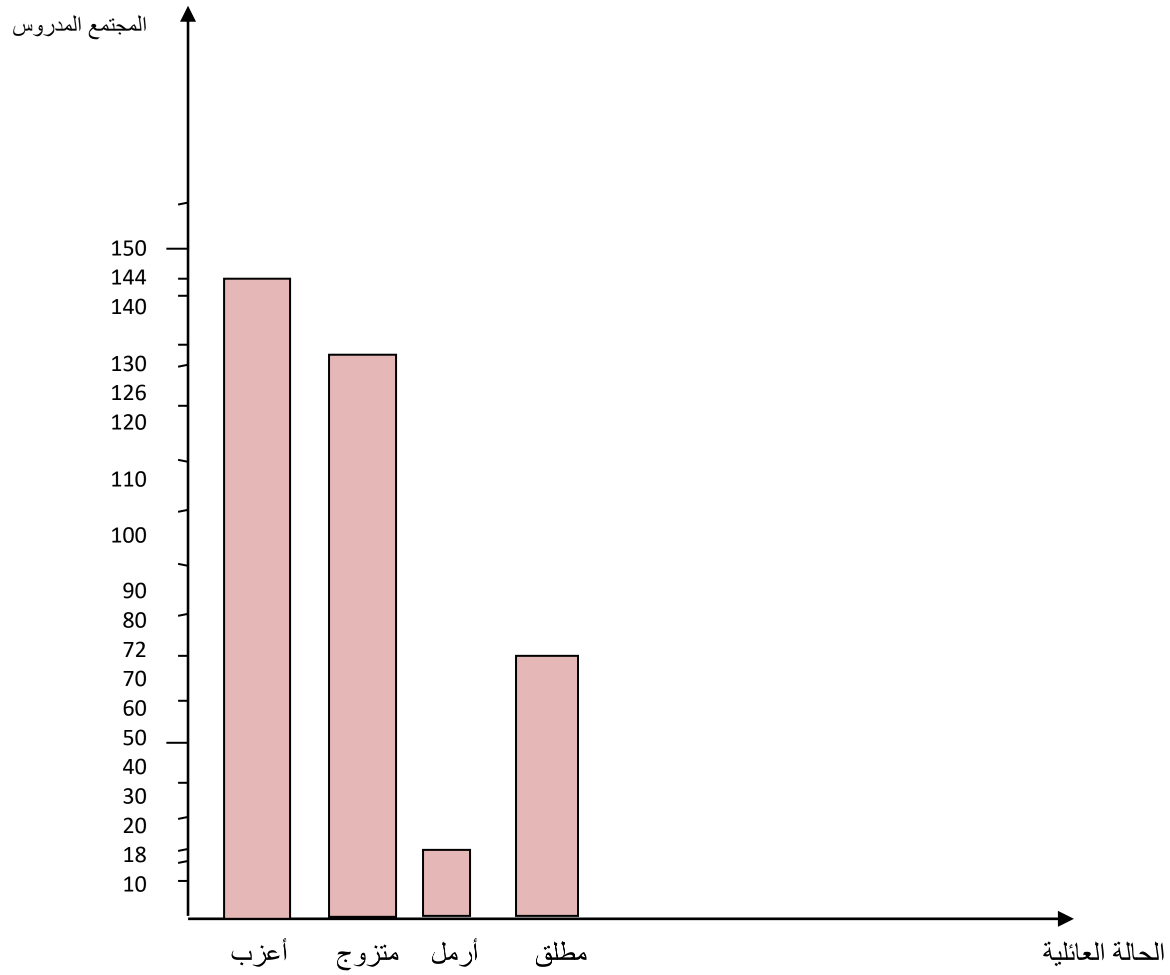
هي عبارة عن مستطيلات متساوية العرض ومتساوية البعد عن بعضها ويكون طول هذه المستطيلات متناسب مع التكرارات (مطلق، نسبي، مئوي). تقوم هذه الطريقة بوضع التكرارات على المحور العمودي والكيفيات على المحور الأفقي.

مثال 1 : يمثل الجدول التكراري التالي عينة من أشخاص موزعين حسب حالتهم العائلية.

الكيفيات m_i	n_i	f_i	f_i%
أعزب	144	0.40	40
متزوج	126	0.35	35
أرمل	18	0.05	5
مطلق	72	0.20	20
المجموع	**360**	**1**	**100**

الحل :

- المجتمع الإحصائي متكون من 360شخص: N = 360.
- الصفة المدروسة هي الحالة العائلية.
- طبيعة الصفة : نوعية اسمية.
- عدد الكيفيات 4 و هي : أعزب، مطلق، أرمل، متزوج.

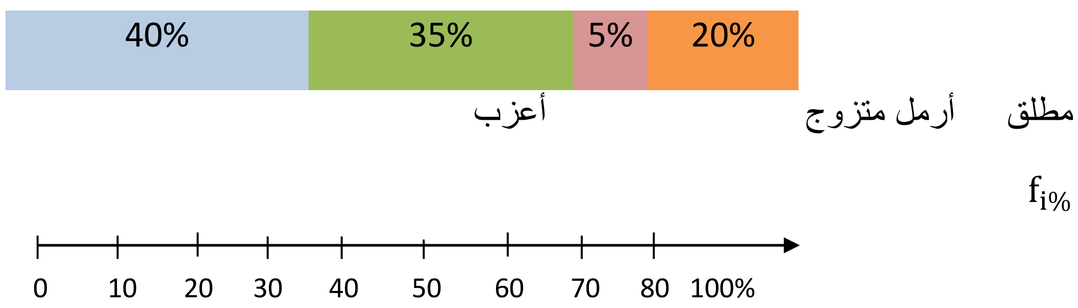


القضبان البيانية

2- العمود المجزأ(Le diagramme rectiligne) :

وهو عبارة عن مستطيل مقسم إلى أجزاء، كل جزء يمثل كيفية تتناسب مع التكرار المئوي: طول العمود يساوي 100%.

مثال 2 : نستعمل المثال الأول و نرسم العمود المجزأ.



العمود المجزأ

3- العرض الدائري (القطاعات) Diagramme circulaire ou par secteurs

وهو عبارة عن دائرة مقسمة إلى أجزاء، كل جزء يمثل كيفية للصفة المدروسة.

الزاوية التي تمثل كل جزء تعطى كالتالي:

$$\alpha_i° = f_i \times 360° = \frac{n_i}{\sum n_i} \times 360°$$

مثال 3 : الجدول التالي يمثل توزيع الأجانب في فرنسا حسب جنسياتهم.

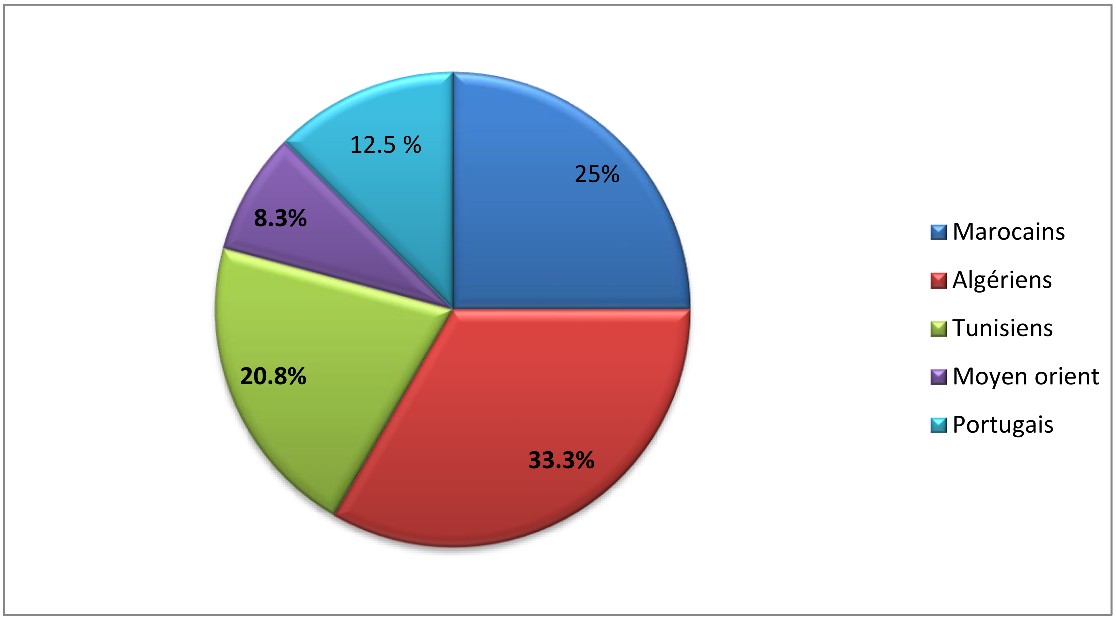
الجنسية m_i	n_i
جزائرية	200
مغربية	150
تونسية	125
من الشرق الأوسط	50
برتغالية	75
المجموع	**600**

المطلوب : مثّل بيانيا هذه البيانات باستعمال العرض الدائري.

الحل :

نحسب الزوايا (الدرجات) العائدة لكل كيفية:

الجنسية m_i	n_i	f_i	$\alpha_i° = f_i \times 360°$
جزائرية	200	0.333	0.333×360° =120°
مغربية	150	0.25	90°
تونسية	125	0.208	75°
من الشرق الأوسط	50	0.083	30°
برتغالية	75	0.125	45°
المجموع	**600**	**1**	**360°**



العرض الدائري لتوزيع الأجانب في فرنسا حسب جنسياتهم

II. العرض البياني في حالة متغير كمي :

يجب التمييز بين العرض البياني لمتغير كمي منفصل (متقطع) ومتغير كمي مستمر (متصل).

1- حالة متغير كمي متقطع (منفصل):

أهم العروض البيانية هي:

1-1. الأعمدة البيانية: Le diagramme en bâtons

هي عبارة عن أعمدة بسيطة تتناسب أطوالها مع التكرار (مطلق، نسبي، أو مئوي) المقابل لقيمة المتغير.

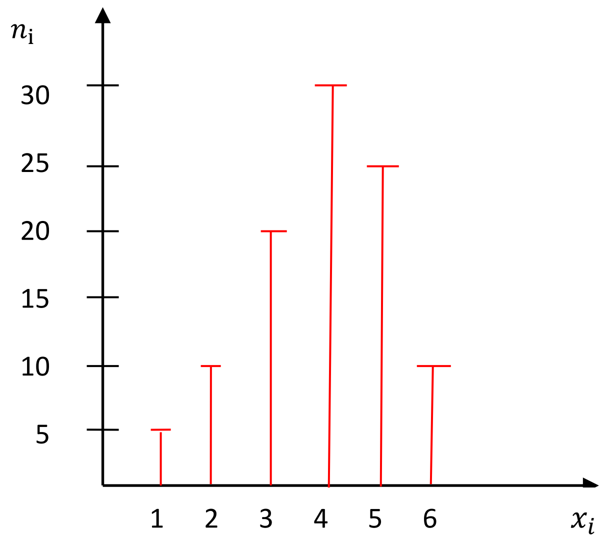
مثال 4 : الجدول التالي يمثل توزيع عيّنة من المنازل حسب عدد الغرف.

x_i	n_i	≤ $nc_i\uparrow$	≥ $nc_i\downarrow$
1	5	5	100
2	10	15	95
3	20	35	85
4	30	65	65
5	25	90	35
6	10	100	10
المجموع	100		

المطلوب :

- مثّل بيانيا البيانات بواسطة الأعمدة البيانية.
- مثل بانيا المعطيات بواسطة التكرارات التجميعية الصاعدة والنازلة.

تعريف : الأعمدة البيانية هي أعمدة رأسية قواعدها متساوية و ارتفاعها يناسب أو يساوي عدد تكراراتها.

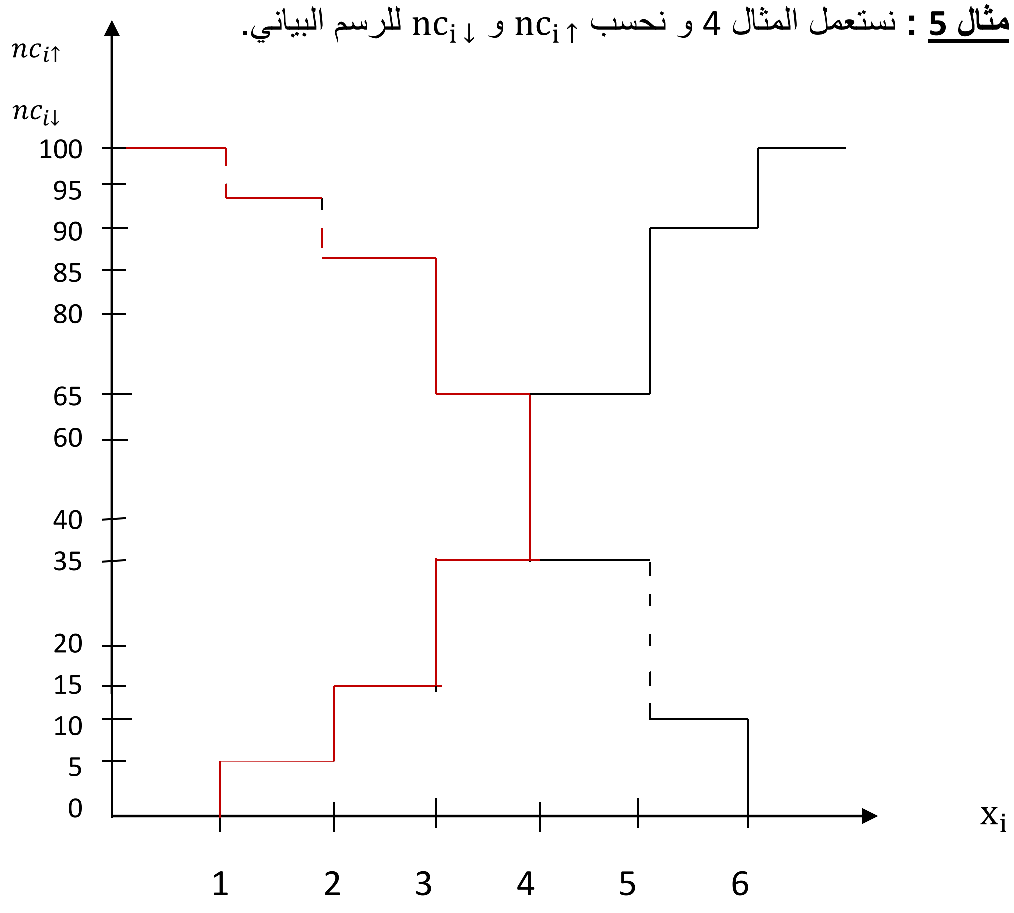


<u>الأعمدة البيانية</u>

2-1. <u>العرض البياني للتكرارات التجميعية الصاعدة و النازلة</u> <u>:Diagramme cumulatif</u>

وهو عبارة عن قطع مستقيمة (متصاعدة أو نازلة) تتناسب مع التكرار التجميعي (الصاعد أو النازل) المقابل لكل قيمة من قيم المتغير.

<u>مثال 5</u> : نستعمل المثال 4 و نحسب $nc_{i\uparrow}$ و $nc_{i\downarrow}$ للرسم البياني.



<u>العرض البياني للتكرارات التجميعية</u>

<u>البيان التجميعي التكاملي</u>

<u>2-</u> <u>حالة متغير كمي مستمر (متصل):</u>

أهم العروض البيانية لمتغير كمي مستمر نجد:

2-2- <u>المدرج التكراري</u>: L'histogramme

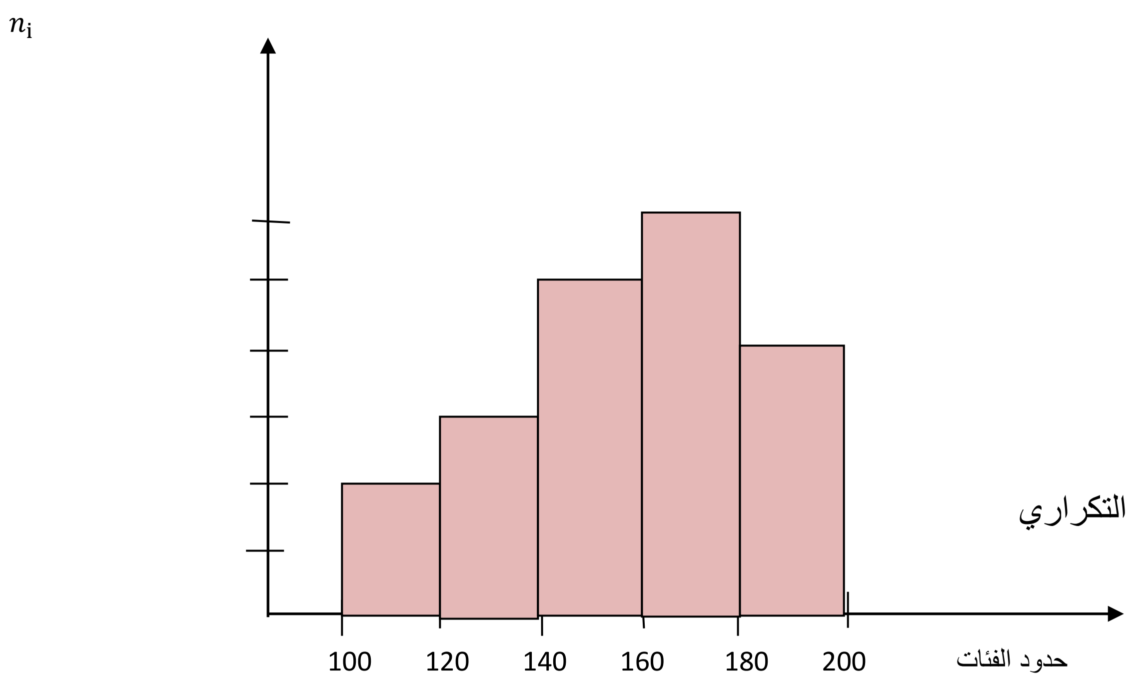
لما تكون البيانات مصنفة داخل فئات أي لما يكون المتغير كمي مستمر نرسم مدرج تكراري وهو عبارة عن مستطيلات متلاصقة حيث يتناسب طول المستطيلات مع تكرار الفئة ويتناسب عرض المستطيلات مع طول الفئة أي حدود الفئات.

ملاحظة : قبل رسم المدرج التكراري يجب التمييز بين حالتين:

<u>الحالة الأولى</u> : طول الفئات متساوي : نقوم برسم المدرج التكراري مباشرة.

<u>مثال 6</u> : يمثل الجدول التالي الأجر في الساعة لمجموعة من العمال.

الفئات	n_i
[100-120[	10
[120-140[	15
[140-160[	25
[160-180[	30
[180-200[	20
المجموع	**100**



المدرج التكراري

<u>الحالة الثانية</u> : طول الفئات غير متساوي : في هذه الحالة قبل رسم المدرج التكراري يجب تعديل التكرارات حتى تصبح قاعدة المقارنة ثابتة و لتعديل التكرارات نتبع الخطوات التالية:

- حساب طول كل فئة (a).
- استنتاج طول الفئة الوحدوي (au) الذي يكون :
-
  - ❖ أكبر قاسم مشترك لأطوال الفئات (PGDC) إن وجد.
  - ❖ أو طول الفئة الأكثر شيوعا.
  - ❖ أو أصغر طول فئة.
- حساب المعاملات (c_i): $\mathbf{c_i = \frac{a}{au}}$

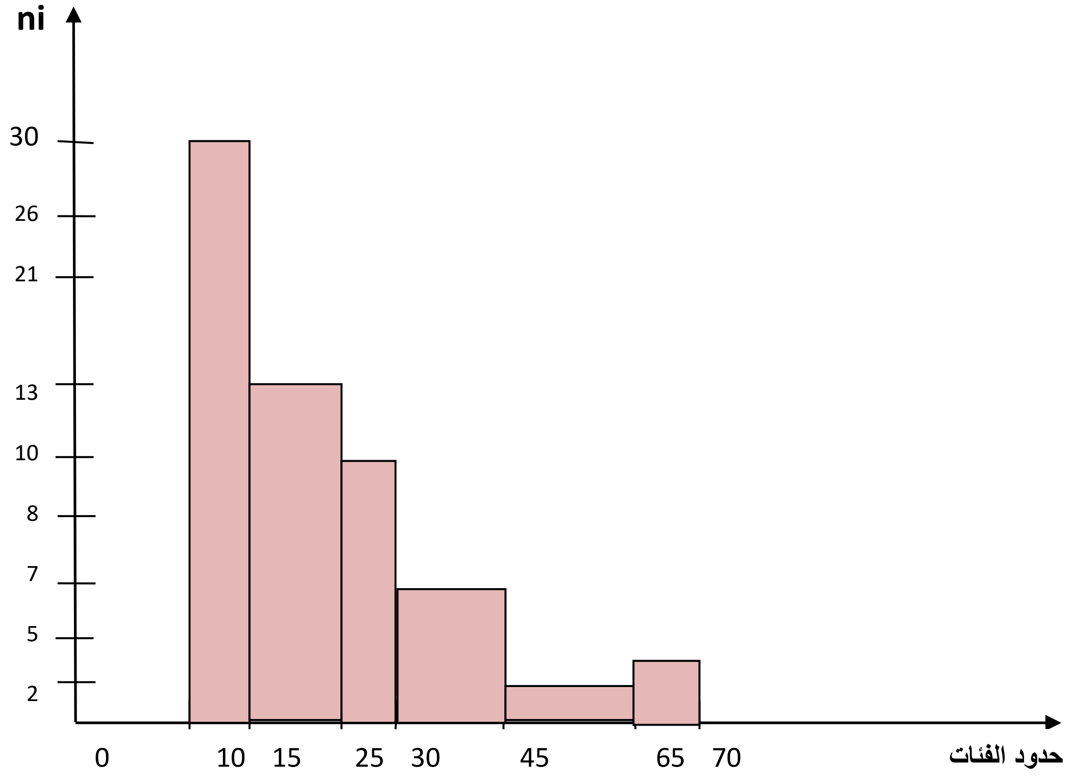
– حساب التكرار المعدل (h_i) $:h_i = \frac{n_i}{c_i}$

مثال 7 : يمثل الجدول التالي توزيع الدخل الشهري بالآلاف لمجموعة من العمال.

n_i	**الفئات**
30	**[10-15[**
26	**[15-25[**
10	**[25-30[**
21	**[30-45[**
8	**[45-65[**
5	**[65-70[**
100	المجموع

الحل : قبل الرسم يجب تعديل التكرارات و ذلك لأن أطوال الفئات غير متساوية.
أكبر قاسم مشترك لأطوال الفئات هو: au = 5

$h_i = \frac{n_i}{c_i}$	$c_i = \frac{a}{au}$	a	n_i	**الفئات**
30	1	5	30	**[10-15[**
13	2	10	26	**[15-25[**
10	1	5	10	**[25-30[**
7	3	15	21	**[30-45[**
2	4	20	8	**[45-65[**
5	1	5	5	**[65-70[**
			100	المجموع



<u>المدرج التكراري المعدّل</u>

3-2- المضلع التكراري: Le polygone des fréquences

يتكوّن من مجموعة من القطع المستقيمة، متصلة والمنكسرة تتحدّد بنقاط إحداثياتها مراكز الفئات والتكرارات المقابلة لها. ولكي نغلق الخط المنكسر الذي نحصل عليه، نعتبر أن هناك فئتين متطرفتين واحدة أقصى اليمين والأخرى أقصى اليسار، تكرار كل واحدة منهما يساوي صفر (0).

مثال 8: الجدول التالي الذي يمثل توزيع 40 عاملا في مؤسسة ما حسب الأجر الساعي.

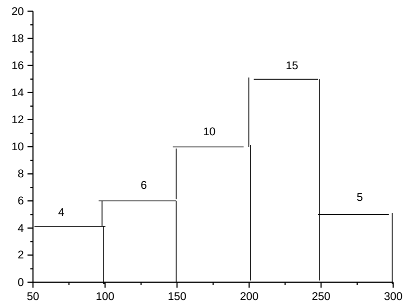
الفئات	n_i
[50-100[	4
[100-150[	6
[150-200[	10
[200-250[	15
[250-300[	5
المجموع	**40**

المطلوب :

1- مثل المدرج التكراري الخاص بهذه البيانات
2- حدد كل من المضلع والمنحنى التكراري.

الحل :

1- نرسم المدرج التكراري المتعلق بهذا التوزيع:

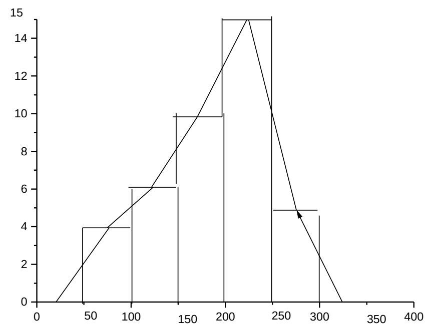


المدرج التكراري

تمثل مساحة كل مستطيل عدد الأفراد (التكرارات) الذين ينتمون إلى الفئة.
بالنسبة للفئة الأولى من**]50-100]** نقول 4 أشخاص.
مساحة هذا المستطيل هي **العرض × الطول**وتساوي**1×4=4** وحدات؛ العدد 1 يعبر عن فئة وهي وحدة القياس على محور (x) والعدد4 يعبر عن الأفراد الموجودين داخل هذه الفئة.
بالنسبة للفئة الثانية من**]100-150]** ونقرأ 6 أشخاص وتكون مساحة المستطيل الخاص بها تساوي **1×6=6** وحدات، إلى غير ذلك.

2- نرسم المضلع التكراري **Polygone des fréquences** المتعلق بهذا التوزيع:

لرسم المضلع التكراري نحدد منتصف كل الفئات- ونقول مركز الفئة- في أعلى المستطيلات المكونة للمدرج التكراري، ثم نمرر خطا منكسرا بكل هذه المراكز. المضلع التكراري يبدأ من 25 أي من منتصف الفئة **]0-50]** وينتهي في منتصف الفئة **]300-350]** بمعنى 325 وهما فئتان افتراضيان رسمناهما حتى تكون المساحة المحصورة بالمضلع التكراري مطابقة لمجموع التكرارات.



مضلع تكراري

Polygone de Fréquences

ملاحظة :

نلاحظ أننا كلما أقصينا مساحة على شكل مثلث من أحد مستطيلات المدرج التكراري عوضناها بمساحة مثلث تساويها، لهذا تمت إضافة الفئتين الافتراضيتين في البداية والنهاية.

و تكون المساحة الإجمالية للمدرج التكراري **aire de l'histogramme** تساوي مجموع مساحات المستطيلات و هي 40 وحدة، و تمثل 40 عاملا أي مجموع التكرارات.

4-2 - المنحنى التكراري التجميعي الصاعد و النازل: Courbes cumulatives

يُرسم المنحنى التكراري التجميعي (الصاعد أو النازل) في معلم متعامد ومتجانس وهذا بربط النقاط التي تكون إحداثيتها حدود الفئات (فواصل) والتكرارات التجميعية (التراتيب).

مثال 9 :

يعطى الجدول التالي توزيع مجموعة من العمال حسب أجرهم الشهري بالدينار الجزائري (DA):

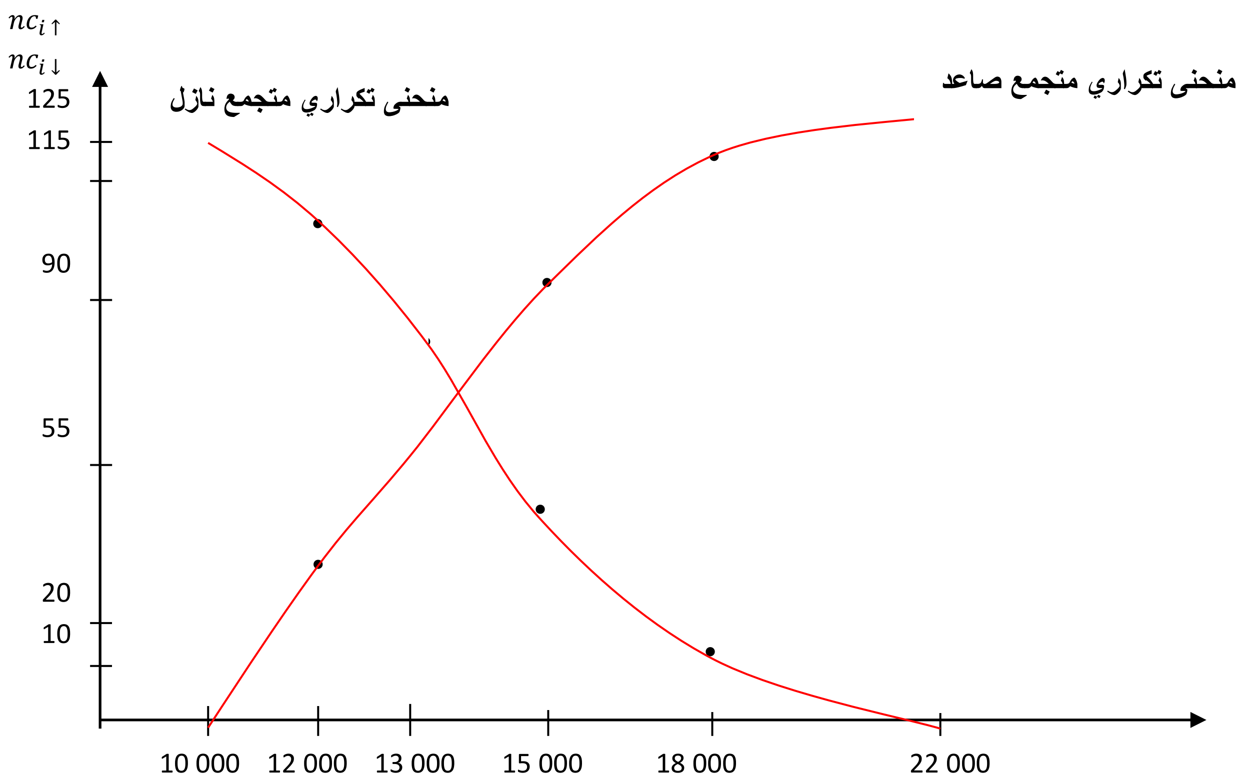
الفئات	n_i
[10000-12000[	25
[12000-13000[	30
[13000-15000[	35
[15000-18000[	25
[18000-22000[	10
المجموع	**125**

الحل:

إن التكرار المتجمع الصاعد هو ترتيب المجتمع الإحصائي ترتيبا تصاعديا.
لحساب التكرار المتجمع النازل فإننا نبدأ من مجموع التكرار ونطرح تكرار الفئات الواحدة بعد الأخرى مع استعمال صيغة " أكبر من الحدود الدنيا للفئات " وهو ما نحصل عليه في الجدول الآتي:

الفئات	n_i	الحد الأعلى (أقل من)	$nc_i\uparrow$	الحد الأدنى (أكثر من)	$nc_i\downarrow$
[10000-12000[	25	أقل من 12 000	25	10 000 فأكثر	125
[12000-13000[	30	أقل من 13 000	55	12 000 فأكثر	100
[13000-15000[	35	أقل من 15 000	90	13 000 فأكثر	70
[15000-18000[	25	أقل من 18 000	115	14 000 فأكثر	35
[18000-22000[	10	أقل من 22 000	125	15 000 فأكثر	10
المجموع	**125**	/	/	/	/

لنرسم التكرار المتجمع الصاعد والنازل:



<u>المنحنى التكراري التجميعي الصاعد و النازل</u>

Courbes cumulatives ascendantes et descendantes

<u>ملاحظة:</u> نحصل على المنحنيات التكرارية المتجّمعة بوصل النقاط المتحصّل عليها والتي تمثل إحداثياتها التالية: حدود الفئات والتكرارات المتجمّعة.

الفصل الرابع

مقاييس النزعة المركزية

في الفصل السابق رأينا أن عرض البيانات لإحصائية لظاهرة ما يتم بواسطة جداول تكرارية وتمثيلات بيانية. للقيام بعملية الاستقراء الإحصائي (التنبؤ وأخذ القرار)، وضعت مقاييس عديدة يمكن استخدامها في عملية التحليل، من بين هذه المقاييس نجد مقاييس النزعة المركزية. هذه المقاييس تُلخص العدد الكبير من البيانات بقيمة واحدة وتُستعمل أساسا من أجل المقارنة بين عدة سلاسل إحصائية. أُطلق عليها مصطلح مقاييس النزعة المركزية لأنها تحتل مركز المعطيات أي تشكّل القيمة التي تتجمّع حولها بقيمة القيم.

I. **المنوال :**

1- تعريفه :

المنوال هو الكيفية أو القيمة الأكثر شيوعا أو الأكثر تكرارا (انتشارا) من غيرها من قيم المجموعة المعطاة ويرمز له ب $\mathbf{M_o}$.

2- **حالة البيانات الغير المبوبة:**

1-2. **الطريقة الحسابية :**

مثال 1:أوجد قيمة المنوال من مجموعة البيانات التالية:

3 – 5 – 7 – 15 – 16 – 16 – 16 – 17 – 17 – 30

القيمة التي تتكرر أكثر من غيرها هي 16:

$$\Rightarrow M_O = 16$$

نقول **سلسلة أحادية المنوال.**

مثال 2:أوجد قيمة المنوال من مجموعة البيانات التالية:

16 – 15 -10 – 13 – 18 – 15 -10

هناك منوالان(**Deux modes : Série bimodal**) وهما القيمتين 10 و15:

$$M_{O_1} = 10 \ , \ M_{O_2} = 15$$

نقول **سلسلة ثنائية المنوال.**

مثال 3:أوجد قيمة المنوال من مجموعة البيانات التالية:

13 – 17 – 9 – 16 – 14

سلسلة عديمة منوال (سلسلة ليس لها منوال) **Série uniforme**

مثال 4:أوجد قيمة المنوال من مجموعة البيانات التالية:

8 – 10 – 8 – 13 – 13 – 15 – 14 – 15

سلسلة ثلاثية المنوال نقول سلسلة متعددة المنوال (**Série plurimodale ou multimodale**).

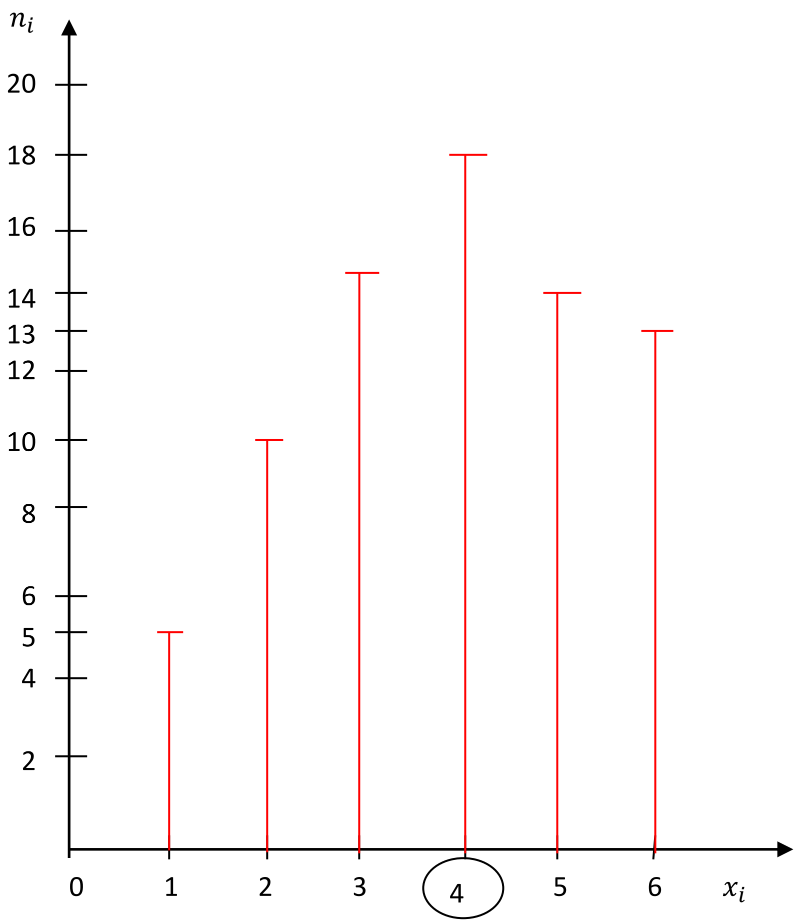
مثال 5:أوجد قيمة المنوال من مجموعة البيانات التالية:

7 – 9 – 7 – 12 – 9 – 12 – 9 – 12 – 7

ليس هناك منوال حيث لا تغلب أي قيمة على بقية القيم **Série uniforme**

2-2. الطريقة البيانية :

نقوم برسم الأعمدة البيانية للسلسلة، تكون فاصلة أطول الأعمدة في قيمة المنوال.



الأعمدة البيانية التي تمثل عدد الأطفال في الأسر

ملاحظة: عند تحديد المنوال توجد عدة حالات ممكنة:

- **توزيع تكراري عديم المنوال:** لما لا توجد أي قيمة متكررة أكثر من غيرها. في هذه الحالة كل قيم المتغير الإحصائي لها نفس التكرار. نقول إن السلسلة الإحصائية عديمة المنوال، لا يوجد منوال.
- **توزيع تكراري وحيد المنوال:** وهو التوزيع الذي يحتوي على قيمة واحدة تتكرر أكثر من غيرها. نقول سلسلة الإحصائية أحادية المنوال.(**Série uni-modale**).
- **توزيع تكراري ثنائي المنوال:** وهو التوزيع التكراري الذي يحتوي على قيمتين تتكرر أكثر من غيرها. نقول سلسلة ثنائية المنوال: يوجد منوالين(**Série bimodale**).
- **توزيع تكراري متعدد المنوال:** وهو التوزيع التكراري الذي يضم عدة قيم تتكرر أكثر من غيرها أي أكثر من منوالين. نقول سلسلة متعددة المنوال (**Série multi ou plurimodale**). في هذه الحالة لا يكون للمنوال أي مدلول إحصائي، ولا يستعمل كمقياس من مقاييس النزعة المركزية إلا في حالات نادرة. في هذه الحالة يفقد المنوال من أهميته كمقياس ولا يصبح له أي مدلول إحصائي.

2- حالة البيانات المُبوبّة :

1.2-حالة بيانات نوعية (صفة نوعية):

منوال الصفة النوعية هو الكيفية الأكثر تكرارا.

مثال : الجدول التالي يمثل التقدير في امتحان لمجموعة من الطلبة :

C_i	n_i
ممتاز	30
جيد	50
متوسط	70
ضعيف	20
ضعيف جدا	5
المجموع	**175**

n_i = عدد الطلبة

الكيفية التي تتكرر أكثر من غيرها هي: متوسط.

الكيفية متوسط تتكرر 70 مرة.

2.2-حالة بيانات كمية (متغيركمي):

1.2.2 -حالة متغير كمي منفصل (متقطع) :

تحديد المنوال تكون سهلة، يتّم تحديدها مباشرة من الجدول وهي القيمة التي يقابلها أكبر تكرار.

مثال : يمثل الجدول التالي توزيع 68 أسرة حسب عدد الأطفال.

عدد الأطفال x_i	عدد الأسر n_i
0	3
1	4
2	8
3	14
4	16
5	12
6	11
المجموع	**68**

القيمة التي يقابلها أكبر تكرار هي : 4.

إذن المنوال هو : $\boxed{M_{o=4}}$

التفسير : معظم الأسر لديهم 4 أطفال.

2.2.2 - حالة متغير كمي مستمر (حساب المنوال من بيانات مبوبة)في فئات متساوية الطول :

هناك عدة طرق لحساب المنوال من بيانات مبوبة، سنستعرضها فيما يلي:

أ- طريقة مركز الفئة المنوالية :

هذه الطريقة بسيطة وتقريبية في حساب المنوال، والفئة المنوالية هي الفئة التي يقابلها أكبر تكرار في التوزيع، حيث يمكن اعتبار مركز الفئة المنوالية منوالا للتوزيع، غير أن حساب المنوال بهذه الطريقة يعتبر غير دقيق ويكون صحيح في حلة التوزيع المتماثل (المتناظر) فقط.

مثال 1 : يمثل الجدول التالي توزيع 110 موظف حسب أجورهم الشهرية (ب 10^2 دج).

الفئات متساوية الطول نطبق القانون مباشرة.

الفئات (الأجور) دج10^2	التكرار n_i
[100– 110 [	9
[110– 120 [	14
[120– 130 [	23
[130– 140 [	32
[140– 150 [	18
[150– 160 [	14
المجموع	**110**

الفئة المنوالية هي الفئة التي تقابل أكبر تكرار(32)⇐ [130– 140[

$$M_O = \frac{130+140}{2} = 135$$

$$M_O = 135.10^2 = 13500 \text{ دج}$$

ب- طريقة بيرسون أو الفروق (طريقة الاستكمال الخطي) :

اذا أردنا أكثر دقة لحساب المنوال نستعمل طريقة الفروق لبيرسون (Pearson)، و التي تُسمى أيضا طريقة الاستكمال الخطيMéthode d’interpolation linéaire و التي تعطى كالتالي:

$$M_O = L_{MIN} + \left[\frac{\Delta 1}{\Delta 1+\Delta 2}\right] . a_{M_O}$$

حيث أن :

- L_{MIN} هو الحد الأدنى للفئة المنوالية أي الفئة التي يقابلها أكبر تكرار.
- Δ_1 هو الفرق بين تكرار الفئة المنوالية وتكرار الفئة السابقة لها.
- Δ_2 . هو الفرق بين تكرار الفئة المنوالية وتكرار الفئة اللاحقة لها.
- a_{M_O} طول الفئة المنوالية.

لدينا فقئات متساوية الطول ⇐الفئة المنوالية هي الفئة التي تقابل أكبر تكرار(32)
الفئة المنوالية هي[130– 140[

$$M_O = 130 + \left(\frac{(32-23)}{(32-23)+(32-18)}\right).(140\text{-}130) = 133{,}913.10^2 = 13391{,}30 \text{ DA}$$

$$\Rightarrow \boxed{M_o = 13391{,}30 \text{ DA}}$$

التفسير : أغلبية أو معظم الموظفين لهم أجر شهري يساوي 13391.30 دج.

3.2.2 - <u>المنوال في حال فئات غير متساوية الطول :</u>

ملاحظة :إذا كانت الفئات متساوية الطول نطبّق القانون مباشرة. أمّا إذا كانت الفئات غير متساوية يجب تعديل التكرارات (بالطريقة التي رأيناها عند رسم المدرّج التكراري)، وتصبح الفئة المنوالية هي الفئة التي تقابل أكبر تكرار معدل.

إذا كانت البيانات مبوبة في فئات غير متساوية الطول نقوم بدراسة جداول تكرارية غير منتظمة والتي تتطلب منا تعديل التكرارات قبل حساب المنوال أو عند رسم المدرج التكراري.

<u>مثال 2</u> : يمثل الجدول التالي توزيع 100 موظف حسب أجورهم الشهرية (ب 10^3 دج).

$h_i = \frac{n_i}{c_i}$	$c_i = \frac{a}{a_u}$	a	عدد الموظفين n_i	الأجر الشهري 10^3
$h_1 = \frac{10}{2} = 5$	$c_1 = \frac{10}{5} = 2$	10	10	**[50– 60 [**
12.5	4	20	50	**[60– 80 [**
15	2	10	30	**[80– 90 [**
10	1	5	10	**[90– 95 [**
/	/	/	**100**	**المجموع**

- نلاحظ أن الفئات غير متساوية⇐يجب تعديل التكرارات.
- أكبر القاسم المشترك الأكبر لأطوال الفئات هو a_u =5
- $c_i = \frac{a}{a_u}$ المعاملات.
- $h_i = \frac{n_i}{c_i}$ التكرارات المعدلة.

الفئة المنوالية هي الفئة التي تقابل أكبر تكرار معدل.

أكبر تكرار معدل هو 15 ⇐ الفئة المنوالية هي [90 –80[

ومنه:

$$M_O = + \left[\frac{\Delta 1}{\Delta 1+\Delta 2}\right] . a_{M_O} L_{MIN}$$

$$M_O = 80 + \left(\frac{(15-12.5)}{(15-12.5)+(15-10)}\right) .(90 - 80) = 83,333.\ 10^3 = 83333 \text{ DA}$$

$$\Rightarrow \boxed{M_o = 83333 \text{ DA}}$$

التفسير : أغلبية أو معظم الموظفين لهم أجر شهري يساوي 83333 دج.

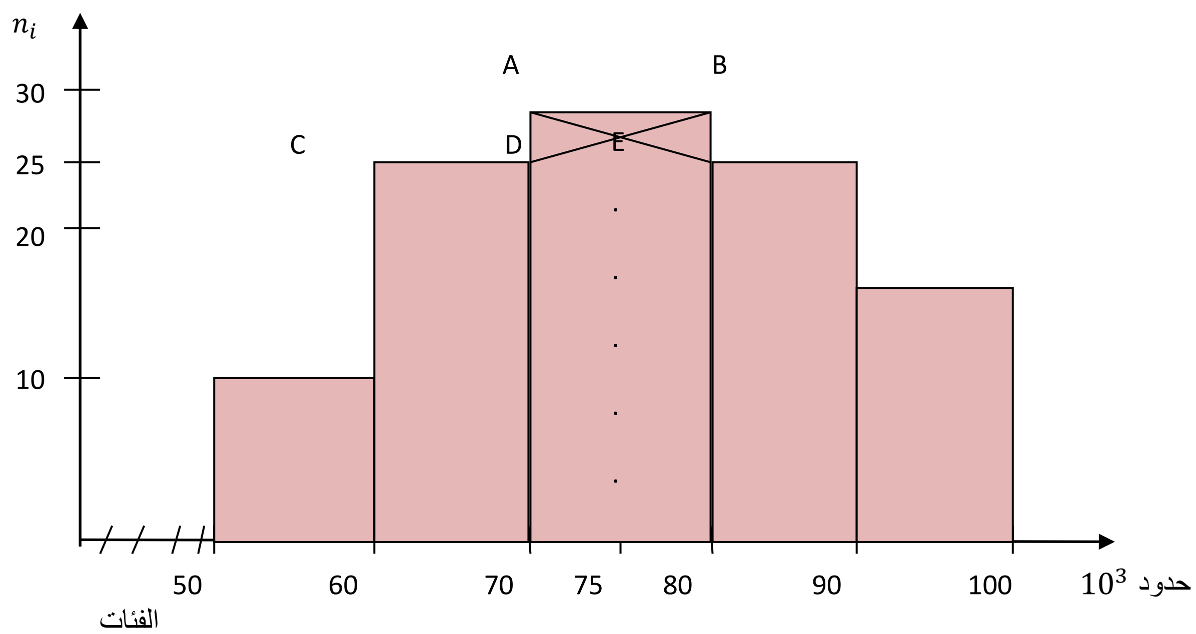
3-2. الطريقة البيانية :

نقوم برسم المدرج التكراري لثلاثة فئات فقط: الفئة المنوالية، الفئتين السابقة واللاحقة لها.

كما هو مبيّن في الشكل. نتحصّل على القطعة [AD] والقطعة [CD] ولتكن النقطة (E) نقطة تقاطع هاتين القطعتين. بإسقاط النقطة (E) على محور الفواصل نتحصّل على قيمة المنوال بيانيا.

قيمة المنوال بيانيا هي: $M_o = 75 \text{ x}10^3 = 75000$ DA

وهي نفس القيمة التي وجدناها حسابيا.



المدرج التكراري الذي يمثل توزيع الموظفين حسب الأجر الشهري ب 10^3دج

II. الوسيط (La médiane)

1- تعريفه :

الوسيط هو قيمة المتغير التي تقسم المجتمع الإحصائي إلى قسمين متساويين ويرمز له ب M_e. لتحديد قيمة الوسيط، يجب أن تكون قيم المتغير الإحصائي مرتبة ترتيبا تصاعديا أو تنازليا. قيمة الوسيط تقسم سلسلة إحصائية مرتبة إلى مجموعتين متكونتين من 50% من البيانات، 50% من القيم. تكون أكبر من الوسيط و50% من القيم تكون أصغر من الوسيط.

2- حالة البيانات الغير المبوبة:

2-1. الطريقة الحسابية :

بعد ترتيب السلسلة وتحديد المجتمع N، نميز بين حالتين:

- **إذا كان N عدد فردي :**

تكون قيمة الوسيط في المنتصف ولتحديدها، يجب معرفة رتبة الوسيط $\mathbf{RM_e}$:

$$\mathbf{RM_e} = \textbf{Rang de la médiane}$$

$$\mathbf{RM_e = \frac{N+1}{2}}$$

مثال1 : تمثل السلسلة التالية الوزن (kg) لمجموعة من الأشخاص.

45 – 68 – 89 – 74 – 62 – 56 – 49 – 52 – 63

- نرتب السلسلة ترتيبا تصاعديا:

$$\underbrace{45 – 49 – 52 – 56}_{\%50} – \underset{M_e}{62} – \underbrace{63 – 68 – 74 – 89}_{\%50}$$

- نحدد N : $N = 9 \Leftarrow$ عدد فردي
- نحدد رتبة الوسيط $\mathbf{RM_e = \frac{N+1}{2}} = \frac{9+1}{2} = 5$
- نحدد قيمة الوسيط M_e والتي هي القيمة M_e التي تحتل المرتبة الخامسة في السلسلة المرتبة:

$$\mathbf{M_e = 62\ Kg}$$

- **اذا كان N عدد زوجي :**

إذا كان المجتمع N عدد زوجي قيمة الوسيط تكون تساوي الوسط الحسابي للقيمتان اللتان تقعان في المنتصف، ولهذا بعد الترتيب يجب تحديد مرتبتين للوسيط:

$$\mathbf{RM_{e1}} = \frac{\mathbf{N}}{\mathbf{2}} \quad \text{و} \quad \mathbf{RM_{e2}} = \frac{\mathbf{N}}{\mathbf{2}} + \mathbf{1}$$

مثال2 : تمثل السلسلة التالية الوزن (kg) لمجموعة من الأشخاص:

45- 89 – 68 – 55 – 74 – 56 – 62 – 52- 63 – 49

- نرتب السلسلة ترتيبا تصاعديا :

$$\underbrace{45 - 49 - 52 - 55 - 56}_{} \downarrow \underbrace{62 - 63 - 68 - 74 - 89}_{M_e}$$

- نحدد N : ⇐N = 10 عدد زوجي
- نجد رتبتين للوسيط

$$\mathbf{RM_{e1}} = \frac{\mathbf{N}}{\mathbf{2}} = \frac{10}{2} = 5 \quad \text{و} \quad \mathbf{RM_{e2}} = \frac{\mathbf{N}}{\mathbf{2}} + \mathbf{1} = \frac{10}{2} + 1 = 6$$

القيمة التي تتوسط هذه المجموعة تقطع بين 56و62وهي لا تظهر. لدينا موقعها في المجموعة وهو بين المرتبة الخامسة والمرتبة السادسة.

$$\Rightarrow \mathbf{M_e} = \frac{\mathbf{56+62}}{\mathbf{2}} \Rightarrow \mathbf{M_e} = \mathbf{59\ kg}$$

-3 حالة البيانات المبوبة:

هنا يجب مراعاة الصفة التي ندرسها، هل المتغير متقطع أم مستمر.

1-3. حالة متغير كمي متقطع (منفصل) :

1.1.3- الطريقة الحسابية :

في هذه الحالة لتحديد قيمة الوسيط، يجب حساب التكرارات التجميعية الصاعدة. الوسيط هو أول قيمة للمتغير الإحصائي التي تتناسب مع تكرار تجميعي صاعد أكبر على $\frac{\mathbf{N}}{2}$ =RM_e، يكون تحديد قيمة الوسيط مباشرة من الجدول.

في هذه الحالة لحساب الوسيط نستعمل يجب إتباع المراحل التالية:
أ- حساب التكرار المتجمع الصاعد.
ب- تحديد ترتيب الوسيط $\mathbf{RM_e} = \frac{\mathbf{N}}{\mathbf{2}}$
ج- حساب الوسيط مباشرة من الجدول.

مثال3 : يمثل الجدول التالي توزيع عدد من الأسر حسب عدد الأطفال في الأسرة.

عدد الأطفال x_i	عدد الأسر n_i	التكرار المتجمع الصاعد $N_i\nearrow$
0	4	4
1	5	9
2	10	19
3	16	35
4	18	53
5	14	67
6	7	4
7	6	80
المجموع	**80**	

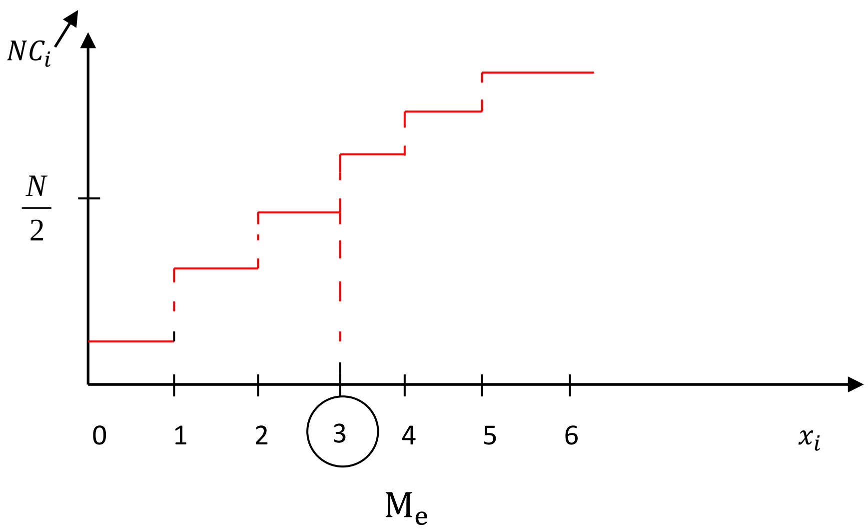
$$RM_e = \frac{N}{2} = \frac{80}{2} = 40$$

الوسيط هو القيمة التي تحتل الرتبة ال40 و هي : $\boxed{M_e = 4}$

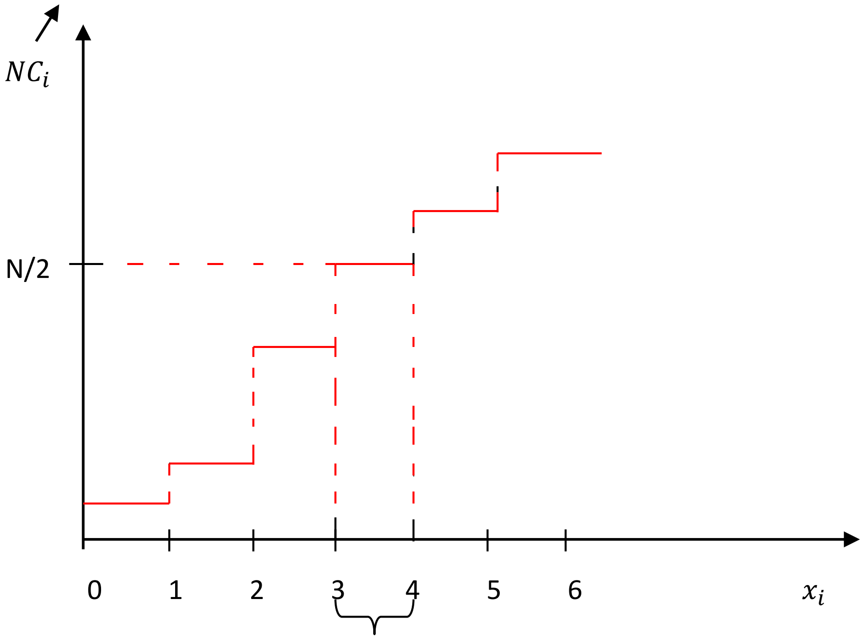
التفسير: 50% من الأسر (40 عائلة) لديهم أكثر من 4 أطفال و 50%الآخرون لديهم أقل من 4 أطفال.

2.1.3- الطريقة البيانية :

في حلة المتغير الإحصائي المتقطع نرسم المدرج للتكرارات التجميعية الصاعدة وحسب رتبة الوسيط $\frac{N}{2}$ يتبين لنا إن كانت قيمة الوسيط قيمة معينة أو مجال وسطي.



$\Rightarrow M_e = 3$



مجال وسطي

قيمة الوسيط تساوي المتوسط الحسابي لقيمتي المجال الوسيطي :

$$M_e = \frac{3+4}{2} = 3,5$$

ملاحظة :

نلاحظ أن هذه القيمة لا توجد ضمن البيانات المعطاة، ولكنها قيمة توجد داخل المجال الوسيطي وهي التي تقسم البيانات إلى قسمين متساويين.

2-3. حالة متغير كمي مستمر:

1.2.3- الطريقة الحسابية :

لحساب الوسيط نستعمل طريقة الاستكمال الخطي ويجب إتباع المراحل التالية:

أ- حساب التكرار المتجمع الصاعد.

ب- تحديد ترتيب الوسيط

ج- تحديد الفئة الوسيطية.

د- تحديد ترتيب الوسيط داخل الفئة الوسيطية: نقوم بحصر رتبة الوسيط RMe في التكرارات التجميعية الصاعدة ومنها يمكننا تحديد الفئة الوسيطية.

ه- حساب حصة الوسيط داخلالفئة الوسيطية.

و- حساب الوسيط الذي يكون يساوي: بداية الفئة الوسيطية + حصة الوسيط داخل الفئة الوسيطية.

$$M_e = L_{min} + \left(\frac{RMe - S}{S' - S}\right).a_{Me}$$

حيث L_{min} هو الأحد الأدنى للفئة الوسيطية.

- RMe : تمثل رتبة الوسيط : $\mathbf{RM_e = \frac{N}{2}}$
- S : التكرار التجميعي الذي يقع قبل RMe.
- S' : التكرار التجميعي الذي يقع بعد RMe.
- a_{Me} : طول الفئة الوسيطية.

<u>مثال4</u> : يمثل الجدول التالي توزيع 50 طفل حسب الوزن (kg).

الوزن (الفئات) Kg	عدد الأطفال n_i	التكرار المتجمع الصاعد $N_i\uparrow$
[20-25[	8	8
[25-30[	4	12
[30-35[	30	42
[35-40[	6	48
[40-45[	2	50
المجموع	**50**	

الفئة الوسيطية هي: **[30-35[** $RM_e = \frac{N}{2} = \frac{50}{2} = 25 \Rightarrow$

- L_{min} هو الأحد الأدنى للفئة الوسيطية = 30
- 12 =S
- 42 =S'
- 5 =a_{Me}

$$\mathbf{M_e = L_{MIN} + \left[\frac{RMe-S}{S'+S}\right] . a_{M_e}}$$

$$\mathbf{M_e = 30+ \left(\frac{(25-12)}{(42-12))}\right) .(35 -30) = 32.166Kg}$$

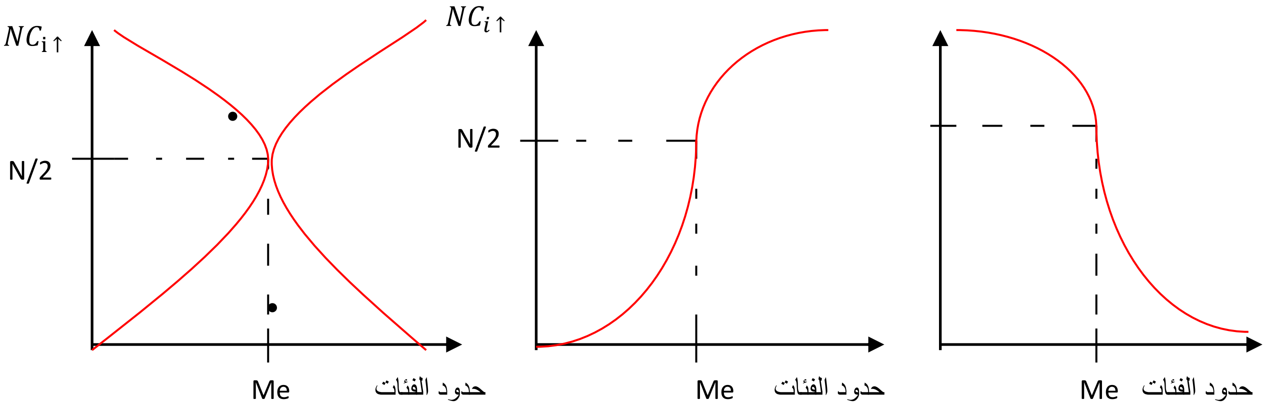
$$\Rightarrow \boxed{\mathbf{M_e = 32.166\ Kg}}$$

التفسير: هناك 50% (25 طفل) من الأطفال الذين وزنهم يقل على 32,2 kg و 50 % (25 طفل) الذين وزنهم يفوق 32,2 kg.

2.2.3- الطريقة البيانية :

يمكن تحديد قيمة الوسيط بيانيا، في هذه الحالة بعد رسم المنحنى التكراري الصاعد والمنحنى التجميعي النازل (أو أحدهما)، نحدد قيمة N/2 على محور التراتيب ثم تتبين لنا قيمة الوسيط (كما في الرسم 1، 2 و3).

لتحديد الوسيط بيانيا نرسم على محور (س) الفئات وعلى محور (ع) التكرار المتجمع الصاعد النسبي ثم نحدد ما هي (س) التي تناسب 50% من (ع).



ملاحظة :

نقرأ قيمة الوسيط على محور (س). نجد قيمة مساوية للتي وجدناها بالطريقة الحسابية.

يمكن أن نلجأ في تحديد الوسيط بيانيا إلى رسم منحنى التكرار المتجمع النسبي الصاعد ومنحنى التكرار المتجمع النازل، ويكون الوسيط هو نقطة تقاطع المنحنيين.

4- خواص الوسيط:

- إن مجموع انحرافات عناصر السلسلة الإحصائية بالنسبة للوسيط أقل من مجموع انحرافات عناصر السلسلة بالنسبة لأي قيمة أخرى. هذا ما يدل على أن الوسيط أقرب للمجموعة من أي عنصر[1] آخر وبالتالي فإنه يمثلها أحسن من أي عنصر آخر.

نبرهن على هذا:

$$\begin{array}{cccccc} & & M_e & & M & \\ & & \downarrow & & \downarrow & \\ x_1 & x_2 & & x_3 & & x_4 \end{array}$$

نكتب مجموع الانحرافات بالنسبة للوسيط[2]:

$$S_1 = (M_e - x_1) + (M_e - x_2) + (x_3 - M_e) + (x_4 - M_e)$$

نكتب مجموع الانحرافات بالنسبة للعنصر M :

$$S_2 = (M - x_1) + (M - x_2) + (M - x_3) + (x_4 - M)$$

---

[1]-نستثني من هذا الوسط الحسابي.

[2]-نكتب الانحرافات بالشكل الذي يمكننا من التخلص من الإشارة السالبة.

نطرح المجموع S_1 من المجموع S_2:

$$S_2 - S_1 = M - x_1 + M - x_2 + M - x_3 + x_4 - M$$
$$- M_e + x_1 - M_e + x_2 - x_3 + M_e - x_4 + M_e$$
$$S_2 - S_1 = 2M - 2x_3$$

ويكون:

$$s_2 - s_1 \succ 0$$

ومنه:

$$S_2 \succ S_1$$

للتنبيه مهما كان عدد العناصر ومهما كان ترتيب العنصر M نحصل دائما على هذا الاستنتاج.

- من مميزات الوسيط أنه سهل الحساب.
- يقسم المدرج التكراري الى مساحتين متساويتين.
- يمكن تحديد الوسيط إذا كانت السلسلة الإحصائية مرتبة حتى لو كنا نجهل بعض عناصرها لأن المهم في حساب الوسيط هو معرفة الترتيب وليس معرفة القيم.
- -العيب في الوسيط أنه غير ملائم في المعالجة الرياضية بمعنى انه لا يخضع للحسابات الجبرية.

III. <u>المقاييس الشبيهة بالوسيط أو التقسيمات:</u>

هناك مقاييس شبيهة بالوسيط من حيث طريقة الحساب وكذا تقسيم البيانات، الا أن نسبة التقسيم تختلف من مقياس إلى آخر. نسميها أيضا مقاييس الوضع وهي قيم تقسم السلسلة الإحصائية المرتبة إلى عدة أجزاء متساوية. ومن هذه المقاييس نذكر ما يلي:

1. <u>الربيعيات:</u>

وهي ناتجة من تقسيم البيانات إلى أربع أقسام متساوية، وبالتالي كل قسم يمثل 25% من البيانات. الربيعيات عددها3 $(\mathbf{Q_i})$ (i = 1,2,3).

- **الربيع الأول (أو الأدنى):** نرمز له بالرمز $\mathbf{Q_1}$ والذي ترتيبه في البيانات المبوبة $\frac{N}{4}$ أي $\frac{\sum n_i}{4}$ يقسم البيانات إلى قسمين، 25% منها أقل منه و75% منها أكبر منه.
- **الربيع الثاني (أو الأوسط):** نرمز له بالرمز $\mathbf{Q_2}$ والذي ترتيبه في البيانات المبوبة $\frac{2N}{4}$ أي $\frac{2\sum n_i}{4}$ يقسم البيانات إلى قسمين، 50% منها أقل منه و50% منها أكبر منه. وهو الوسيط M_e.
- **الربيع الثالث (أو الأعلى):** نرمز له بالرمز $\mathbf{Q_3}$ والذي ترتيبه في البيانات المبوبة $\frac{3N}{4}$ أي $\frac{3\sum n_i}{4}$ يقسم البيانات إلى قسمين، 75% منها أقل منه و25% منها أكبر منه.

نجد قيمة الربيعيات بالاستناد إلى الطريقة المستعملة لحساب الوسيط. ولحساب الربيعيات Q_i من بيانات مبوبة نطبق الصيغة التالية:

$$Q_i = \mathrm{L_{MIN}} + a_{Q_i}\left(\frac{\frac{i.N}{4} - S}{S' - S}\right)$$

حيث Q_i الربيعيات (i= 1,2,3)وأن $RQ_i = \frac{i.N}{4}$.

2. الخميسات:

وهي ناتجة من تقسيم البيانات إلى خمس أقسام متساوية، وبالتالي كل خمس أو قسم يمثل 20% من البيانات.

الخميسات عددها4 (q_i), (i = 1,2,3,4).

- **الخميس الأول:** نرمز له بالرمز q_1 و الذي ترتيبه في البيانات المبوبة $\frac{N}{5}$ أي $\frac{\sum n_i}{5}$ يقسم البيانات إلى قسمين، 20% منها أقل منه و 80% منها أكبر منه.
- **الخميس الثالث:** نرمز له بالرمز q_3و الذي ترتيبه في البيانات المبوبة $\frac{3N}{5}$أي $\frac{3\sum n_i}{5}$ يقسم البيانات إلى قسمين، 60% منها أقل منه و 40% منها أكبر منه.
- **الخميس الرابع:** نرمز له بالرمز q_4 و الذي ترتيبه في البيانات المبوبة $\frac{4N}{5}$أي $\frac{4\sum n_i}{5}$ يقسم البيانات إلى قسمين، 80% منها أقل منه و 20% منها أكبر منه.

نجد قيمة الخميسات بالاستناد إلى الطريقة المستعملة لحساب الوسيط. ولحساب الخميسات q_i من بيانات مبوبة نطبق الصيغة التالية:

$$q_i = \mathrm{L_{MIN}} + a_{q_i}\left(\frac{\frac{i.N}{5} - S}{S' - S}\right)$$

3. العشيرات:

وهي ناتجة من تقسيم البيانات إلى عشرة أقسام متساوية، وبالتالي كل قسم يمثل 10% من البيانات.
العشيرات عددها9 $(\mathbf{D_i})$ (i = 1,2,3,...,8,9).

- **العشير الأول:** نرمز له بالرمز D_1 و الذي ترتيبه في البيانات المبوبة $\frac{N}{10}$ أي $\frac{\sum n_i}{10}$ يقسم البيانات إلى قسمين، 10% منها أقل منه و 90% منها أكبر منه.
- **العشير السادس:** نرمز له بالرمز D_6 و الذي ترتيبه في البيانات المبوبة $\frac{6N}{10}$ أي $\frac{6\sum n_i}{10}$ يقسم البيانات إلى قسمين، 60% منها أقل منه و 40% منها أكبر منه.
- **العشير التاسع و الأخير:** نرمز له بالرمز D_9 و الذي ترتيبه في البيانات المبوبة $\frac{9N}{10}$ أي $\frac{9\sum n_i}{10}$ يقسم البيانات إلى قسمين، 90% منها أقل منه و 10% منها أكبر منه.

ملاحظة: العشير الخامس هو الوسيط: $D_5 = M_e$

نجد قيمة العشيرات بالاستناد إلى الطريقة المستعملة لحساب الوسيط. ولحساب الربيعيات D_i من بيانات مبوبة نطبق الصيغة التالية:

$$D_i = \mathbf{L_{MIN}} + a_{D_i}\left(\frac{\frac{i.N}{10} - S}{S' - S}\right)$$

حيث D_i الشعيرات (i= 1,2,3,...,8,9) و أن $RD_i = \frac{i.N}{10}$.

4. المئينات أو المئينيات:

وهي ناتجة من تقسيم البيانات إلى مئة قسم متساوي، وبالتالي كل قسم يمثل 1% من البيانات.
المئينات عددها 99 $(\mathbf{P_i})$ (i = 1,2,3,...,98,99).

- **المئين الأول:** نرمز له بالرمز P_1 والذي ترتيبه في البيانات المبوبة $\frac{N}{100}$ أي $\frac{\sum n_i}{100}$ يقسم البيانات إلى قسمين، 1% منها أقل منه و99% منها أكبر منه.

– **المئين الخامس والعشرون:** نرمز له بالرمز$\boldsymbol{P_{25}}$والذي ترتيبه في البيانات المبوبة $\frac{25N}{100}$أي $\frac{25\sum n_i}{100}$ يقسم البيانات إلى قسمين، 25% منها أقل منه و75% منها أكبر منه. وهو الربيع الأول.

ملاحظة: المئين الخامس والعشرون: $\boldsymbol{P_{25} = Q_1}$

– **المئين التاسع والتسعون والأخير:** نرمز له بالرمز$\boldsymbol{P_{99}}$والذي ترتيبه في البيانات المبوبة $\frac{99N}{100}$أي $\frac{99\sum n_i}{100}$ يقسم البيانات إلى قسمين، 99% منها أقل منه و1% منها أكبر منه.

ملاحظة: المئين الخمسون:

$$\boldsymbol{M_e = Q_2 = D_5 = P_{50}}$$

نفس الرتبة نفس القيمة.

نجد قيمة المئينات بالاستناد إلى الطريقة المستعملة لحساب الوسيط. ولحساب المئينات$\boldsymbol{P_i}$ من بيانات مبوبة نطبق الصيغة التالية:

$$\boldsymbol{P_i = \mathrm{L_{MIN}} + a_{P_i}\left(\frac{\frac{i.N}{100} - S}{S' - S}\right)}$$

حيث $\boldsymbol{P_i}$المئينات (i= 1,2,3,...,98,99) وأن$\boldsymbol{RP_i = \frac{i.N}{100}}$.

مثال : الجدول التالي يبين مبيعات محل تجاري خلال أسبوع :

الجدول رقم (1)

فئات المبيعات (10 آلاف د.ج)	10-20	20-30	30-40	40-50	50-60	60-70	70-80
عدد المحلات n_i	15	27	30	58	26	14	10

المطلوب :

أحسب كل من: الربيع الثالث، الخميس الرابع، العشير السادس والمئين الخامس والأربعون.

الحل :

فئات المبيعات (10 آلاف د.ج)	عدد المحلاتn_i	$nc_i \nearrow$ تكرار متجمع صاعد
[10-20[	15	15
[20-30[	27	42
[30-40[	30	72
[40-50[	58	130
[50-60[	26	156
[60-70[	14	170
[70-80[	10	180
المجموع	**180**	

حساب Q_3 :

ترتيب Q_3 = $\frac{3N}{4}$ $\frac{3(180)}{4}$ =135، $\Leftarrow$ **فئة الربيع الثالث** $]60-50[$

و منه نجد :

$$Q_3 = \mathrm{L_{MIN}} + a_{Q_3}\left(\frac{\frac{3.N}{4} - S}{S' - S}\right)$$

$$\Rightarrow Q_3 = 50 + 10.\left(\frac{135-130}{156-130}\right) = 51.923$$

حساب q_4 :

ترتيب q_4 = $\frac{4N}{5}$ $\frac{4.(180)}{5}$ =144، $\Leftarrow$ **فئة الربيع الثالث** $]60-50[$

و منه نجد :

$$q_4 = \mathrm{L_{MIN}} + a_{q_4}\left(\frac{\frac{4.N}{5} - S}{S' - S}\right)$$

$$\Rightarrow q_4 = 50 + 10.\left(\frac{144-130}{156-130}\right) = 55.384$$

حساب D_6 :

ترتيب D_6 = $\frac{6N}{10}$ $\frac{6(180)}{10}$ = 108، ⇐**فئة العشير السادس** $[40 - 50]$

و منه نجد :

$$D_6 = \mathrm{L_{MIN}} + a_{D_6}\left(\frac{\frac{6.N}{10} - S}{S' - S}\right)$$

$$\Rightarrow Q_3 = 40 + 10.\left(\frac{108-72}{130-72}\right) = 46.206$$

حساب P_{45} :

ترتيب P_{45} = $\frac{45N}{100}$ $\frac{45(180)}{100}$ **=**81، ⇐**فئة المئين الخامس والأربعون** $[40 - 50]$

و منه نجد :

$$P_{45} = \mathrm{L_{MIN}} + a_{P_{45}}\left(\frac{\frac{45.N}{100} - S}{S' - S}\right)$$

$$\Rightarrow P_{45} = 40 + 10.\left(\frac{81-72}{130-72}\right) = 41.551$$

ملاحظة: يجب الإشارة الى أن ترتيب كل من الوسيط، الربيع الثاني، العشير الخامس و المئين الخمسين كلها متساوية و منه نستنتج أن هذه المقاييس متساوية أي :

$$M_e = Q_2 = D_5 = P_{50}$$

IV. المتوسط (أو الوسط) الحسابي $\overline{x}$: (La moyenne arithmétique)

1- تعريفه :

يعرف المتوسط الحسابي بأنه مجموع القيم مقسوما على عددها، أو بعبارة أخرى هو القيمة التي إذا ضربت في عدد القيم لأعطت مجموعها. ونرمز له بالرمز $\bar{x}$.

2- حالة البيانات الغير المبوبة:

2-1. الطريقة المباشرة (المتوسط الحسابي البسيط):

في حالة عدم وجود التكرار فإن الوسط الحسابي الذي نحصل عليه يسمى الوسط الحسابي البسيط[3] ويساوي:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_N}{N} = \frac{\sum x_i}{N} \quad (1)$$

حيث x_i تمثل قيم المتغير الإحصائي، وN تمثل عدد القيم.

مثال1: أحسب المتوسط الحسابي للقيم التالية: 6، 8، 5، 11، 4، 2.

$$\bar{x} = \frac{6+8+5+11+4+2}{6} = \frac{33}{6} = 5.5$$

ملاحظة: نلاحظ أننا لو ضربنا قيمة المتوسط الحسابي 5.5 بعدد القيم 6 لأعطى مجموع تلك القيم وهو 33.

و بتطبيق الصيغة رقم (1) نجد :

$$\bar{x} = \frac{\sum x_i}{N} = \frac{x_1 + x_2 + \dots + x_N}{N} = \frac{6+8+5+9+3+2}{6} = 5.5$$

ملاحظة: نلاحظ أن قيمة المتوسط الحسابي تنتمي الى مجال القيم المعطاة، هذا المجال الذي حده الأدنى 2 وحده الأقصى 9، ولا يمكن أن يكون بأي حال من الأحوال خارج هذا المجال.

مثال2: إذا كانت نقاط 5 طلبة في أحد المواد هي : 8 -20 -10 -15 -12

المطلوب : أوجد معدل هذه المادة؟

الحل : معدل المادة هو المتوسط الحسابي :

$$\overline{x} = \frac{1}{N}\sum x_i = \frac{1}{5}[12 + 15 + 10 + 20 + 8] = \frac{65}{5} = 13$$

[3] في الواقع هذا الوسط الحسابي هو وسط حسابي مرجح لكن تكرار كل العناصر يساوي1.

2-2. المتوسط الحسابي المرجح (متغير إحصائي متقطع):

فإذا رمزنا للقيم بالرموز: $x_1,x_2,x_3,\dots,x_{p-1},x_p$ و للتكرارات المقابلة لها بالرموز $n_1,n_2,n_3,\dots,n_{p-1},n_p$ على الترتيب، يكون لدينا :

$$\bar{x} = \frac{n_1.x_1 + n_2.x_2 + n_3.x_3 + \cdots + n_{p-1}.x_{p-1} + n_p.x_p}{n_1 + n_2 + n_3 + \cdots + n_{p-1} + n_p}$$

أو بشكل مختصر:

$$\bar{x} = \frac{\sum_{i=1}^{i=p} n_i x_i}{\sum_{i=1}^{i=p} n_i} \quad (2)$$

حيث xi تمثل قيم المتغير الإحصائي المتقطع، و ni تمثل التكرارات المطلقة لها.

مثال3:

لدينا الجدول التالي :

x_i	التكرار n_i	$n_i x_i$
2	1	2
6	2	12
10	3	30
12	2	24
المجموع	**8**	**68**

المطلوب :

أحسب المتوسط الحسابي.

الحل :

$$\bar{x} = \frac{\sum_{i=1}^{i=4} n_i x_i}{\sum_{i=1}^{i=4} n_i} = \frac{68}{8} = 8.5$$

مثال 4: امتحن طالب في 3 مواد و تحصل على النقاط التالية : 8 – 12 -10 علما أن معامل كل مادة هي على التوالي : 3 -2 -2 .

المطلوب: أوجد معدل هذا الطالب؟

الحل:

x_i	n_i	$n_i x_i$
8	3	24
10	2	20
12	2	24
Σ	7	68

$$\bar{x} = \frac{\sum_{i=1}^{i=3} n_i x_i}{\sum_{i=1}^{i=3} n_i} = \frac{68}{7} = 9.71$$

3-2. المتوسط الحسابي بالطريقة المباشرة لبيانات مبوبة :

في حالة وجود فئات نستخدم نفس العلاقة (2) مع xi تمثل مراكز الفئات.

المتوسط الحسابي باستخدام التكرار النسبي:

$$\bar{x} = \sum_{i=1}^{k}[fi.xi]$$

مثال 5:

من الجدول رقم (01) أحسب المتوسط الحسابي لمبيعات المحلات.

فئات المبيعات (10 آلاف د.ج)	عدد المحلات n_i	مراكز الفئات x_i	$n_i.x_i$	f_i	$f_i.x_i$
[10-20[	15	15	225	0.083	1.2454
[20-30[	27	25	675	0.15	3.75
[30-40[	30	35	1050	0.167	5.845
[40-50[	58	45	2610	0.322	14.49
[50-60[	26	55	1430	0.144	7.92
[60-70[	14	65	910	0.078	4.125
[70-80[	10	75	750	0.056	750
المجموع	180		7650	1	42.52

المتوسط الحسابي لمبيعات المحلات :

$$\bar{x} = \frac{\sum_{i=1}^{i=7} n_i x_i}{\sum_{i=1}^{i=7} n_i} = \frac{7650}{180} = 42.5$$

المتوسط الحسابي عن طريق التكرار النسبي :

$$\bar{x} = \sum_{i=1}^{k}[fi.xi] = 42.5$$

3-2. المتوسط الحسابي بطريقة تبديل المتغير:

في بعض الأحيان تصادفنا بيانات كبيرة من حيث الحجم وكثيرة من حيث العدد، ولحساب متوسطها الحسابي تنتج لدينا أرقام كبيرة جدا، ولتبسيط الحساب نستعمل طريقة تبديل المتغير (أو الانحرافات المختزلة عن وسط فرضي).

في هذه الطريقة نستبدل المتغير x_i بمتغير جديد z_i، بحيث تكون العلاقة بين المتغيرينكالتالي:

$$x_i = x_o + a.z_i$$

ونعوّض المتغير x_i بالمتغيرفي z_i فيعلاقة $\bar{x}$:

$$\bar{x} = \frac{\sum n_i.x_i}{\Sigma ni} = \frac{\Sigma(ni.[x_0 + a.z_i])}{\Sigma ni} = \frac{\Sigma[ni.x_0 + a.ni.z_i)}{\Sigma ni}$$

$$\Rightarrow \quad \boxed{\bar{x} = x_0 + a.\bar{z}}$$

الشيء الذي يجب أن ننبه إليه هو استحسان اختيار القيمة الثابتة x_0 التي نطلق عليها اسم الوسط الحسابي الفرضي.

a: طول الفئة

x_0 : قيمة ثابتة اختيارية من بين قيم x_i (مراكز الفئات) يفضل تتوسط السلسلة الإحصائية حتى نحصل على قيم متناظرة/أو تقابل أكبر تكرار. هذه القيمة تمكننا من تسهيل عملية الحساب كما سنراه في المثال اللاحق.

: متوسط مؤقت

$$\bar{z} = \frac{\sum[n_i.z_i]}{\sum n_i}$$

مثال 6:

من الجدول رقم (01) أحسب المتوسط الحسابي لمبيعات المحلات بطريقة تبديل المتغير.

نأخذ: $\begin{cases} x_0 = 45 \\ a = 10 \end{cases}$

فئات المبيعات (10 آلاف د.ج)	عدد المحلات n_i	مراكز الفئات x_i	$z_i = \frac{x_i - x_0}{a}$	f_i	$n_i.z_i$
[10-20[	15	15	-3	0.083	- 45
[20-30[	27	25	-2	0.15	- 54
[30-40[	30	35	-1	0.167	- 30
[40-50[	58	45	0	0.322	0
[50-60[	26	55	+1	0.144	+26
[60-70[	14	65	+2	0.078	+28
[70-80[	10	75	+3	0.056	+30
المجموع	**180**			**1**	**-45**

وسط مؤقت Moyenne provisoire

$$\bar{z} = \frac{\sum[n_i.z_i]}{\sum n_i} = \frac{-45}{180} = -0.25$$

المتوسط الحسابي

$$\Rightarrow \quad \boxed{\bar{x} = x_0 + a.\bar{z}}$$

$$\bar{x} = 45 + [(10).(-0.25)] = 42.5$$

نفس النتيجة بالطريقة المباشرة.

ملاحظة :

1. تستعمل طريقة تبديل المتغير في حالة المتغيرة الإحصائية المستمرة (وجود فئات).
2. في حالة البيانات الغير المبوبة نأخذ $a = 1$ وذلك لعدم وجود فئات والعلاقة السابقة تصبح:

$$\boxed{\bar{x} = x_0 + \bar{z}}$$

مثال 7:

الجدول التالي يمثل الأجور الشهرية ب 10^2 دج لمجموعة من العمّال:

الأجور (10^2 د.ج)	n_i
[130-150[	21
[150-170[	49
[170-190[	100
[190-210[	24
[210-230[	6
Σ	**200**

المطلوب:

حساب المتوسط الحسابي ب:

أ- علاقة التعريف.

ب- طريقة تبديل المتغير.

الحل:

الأجور (10^2 د.ج)	n_i	x_i	$n_i.x_i$	$z_i = \frac{x_i - x_0}{a}$	$n_i.z_i$
[130-150[	21	140	2940	-2	- 42
[150-170[	49	160	7840	-1	- 49
[170-190[	100	180	18000	0	0
[190-210[	24	200	4800	+1	+24
[210-230[	6	220	13200	+2	+12
Σ	**200**		**34900**		**-55**

أ- **علاقة التعريف:**

$$\overline{x} = \frac{\sum[n_i.x_i]}{\sum n_i}$$

$$\bar{x} = \frac{\sum[n_i.x_i]}{\sum n_i} = \frac{3900}{200} = 174.5.10^2 = 17450 \text{ دج}$$

ب- طريقة تبديل المتغير:

نأخذ: $\begin{cases} x_0 = 180 \\ a = 20 \end{cases}$

وسط مؤقت Moyenne provisoire

$$\bar{z} = \frac{\sum[n_i.z_i]}{\sum n_i} = \frac{-55}{200} = -0.275$$

المتوسط الحسابي

$$\Rightarrow \quad \boxed{\bar{x} = x_0 + a.\bar{z}}$$

$$\bar{x} = 180 + [(20).(-0.275)] = 174.5.10^2 = 17450 \text{ دج}$$

نفس النتيجة بالطريقة المباشرة.

4-2.خصائص المتوسط الحسابي:

– مجموع انحرافات قيم السلسلة عن المتوسط الحسابي يساوي دائما الصفر ويمكن أن نبرهن على ذلك:

$$\Sigma(x_i - \bar{x}) = 0$$

$$\Sigma(x_i - \bar{x}) = \sum x_i - \sum \bar{x}$$

$$\Sigma(x_i - \bar{x}) = \sum x_i - N.\bar{x}$$

لدينا:

$$\bar{x} = \frac{\sum x_i}{N} \Rightarrow \sum x_i = N.\bar{x}$$

إذا عوضنا $N\bar{x}$ بقيمته نجد أن:

$$\Rightarrow \Sigma(x_i - \bar{x}) = \sum x_i - \sum x_i = 0$$

ملاحظة: هذه الخاصية تستنتج من تعريف المتوسط الحسابي الذي قلنا عنه أنه يمكن أن يحل محل أي قيمة من السلسلة الإحصائية وبالتالي فإن مجموع الفروق بينه وبين القيم الأخرى لا بد أن يساوي الصفر. وبعبارة أخرى فإن المتوسط الحسابي يمثل مركز الثقل أو مركز التوازن بالنسبة للسلسلة الإحصائية لأن الانحرافات بالنسبة للمتوسط الحسابي تمثل القوى التي تتجاذب هذا المقياس وعندما نقول

أن مجموع الانحرافات يساوي الصفر هذا معناه أن مجموع القوى يساوي الصفر و يكون المتوسط الحسابي هو مركز التوازن.

- مجموع مربع الانحرافات بالنسبة للمتوسط الحسابي أقل من مجموع مربع الانحرافات بالنسبة لأي قيمة أخرى، أي:

$$\Sigma(x_i-\bar{x})^2 \prec \Sigma\left(x_i-c\right)^2$$

حيث يمثل c أي قيمة ثابتة. لنبرهن على ذلك:

ننطلق من الحد الثاني للعبارة السابقةونطرح $\bar{x}$ ثم نضيفه، فنحصل على:

$$\Sigma\left(x_i-c\right)^2=\Sigma\left[(x_i-\bar{x})+(\bar{x}-c)\right]^2$$

$$=\Sigma\left[(x_i-\bar{x})^2+2(\bar{x}-c)(x_i-\bar{x})+(\bar{x}-c)^2\right]$$

$$=\Sigma(x_i-\bar{x})^2+2(\bar{x}-c)\Sigma(x_i-\bar{x})+\Sigma(\bar{x}-c)^2$$

وبما أن:

$$\Sigma(x_i-\bar{x})=0$$

يكون لدينا:

$$\Sigma(x_i-c)^2=\Sigma(x_i-\bar{x})^2+N(\bar{x}-c)^2$$

وهذا معناه أن:

$$\Sigma(x_i-\bar{x})^2 \prec \Sigma\left(x_i-c\right)^2$$

وهذه من أهم خواص الوسط الحسابي وسنستعملها في مقاييس التشتت وتبرهن هذه الخاصية على أن الوسط الحسابي أقرب للمجموع القيم المدروسة من أي قيمة أخرى وهو ما يجعل الوسط الحسابي أحسن ممثل لباقي القيم.

ملاحظة: إذا لدينا تكرار فإن البرهان لا يختلف ويجب أن يتم ضرب كل x_i في التكرار n_i. يكون لدينا بالنسبة للخاصية الأولى:

$$\Sigma(n_i x_i-\bar{x})=\Sigma n_i x_i-\Sigma\,\bar{x}$$

نقسم هذه العبارة على N:

$$\frac{\Sigma n_i x_i}{N}-\frac{N\,\bar{x}}{N}=\bar{x}-\bar{x}=0$$

أما بالنسبة للخاصية الثانية نكتب:

$$\Sigma(n_i x_i-\bar{x})^2 \prec \Sigma\left(n_i x_i-c\right)^2$$

ننطلق من الحد الثاني لهذه المعادلة ونكتب:

$$\Sigma\left(n_i x_i - c\right)^2 = \Sigma\left|(n_i x_i - \bar{x}) + (\bar{x} - c)\right|^2$$

$$= \Sigma\left|(n_i x_i - \bar{x})^2 + 2(n_i x_i - \bar{x})(\bar{x} - c) + (\bar{x} - c)^2\right|$$

$$= \Sigma(n_i x_i - \bar{x})^2 - 2(\bar{x} - c)\Sigma(n_i x_i - \bar{x}) + \Sigma(\bar{x} - c)^2$$

مع العلم أن $\Sigma n_i x_i = N\bar{x}$ فيصبح الحد الأوسط من العبارة السابقة يساوي 0 ويكون لدينا:

$$\Sigma\left(n_i x_i - c\right)^2 = \Sigma(n_i x_i - \bar{x})^2 + N(\bar{x} - c)^2$$

ويتضح أن:

$$\Sigma(n_i x_i - \bar{x})^2 \prec \Sigma\left(n_i x_i - c\right)^2$$

- من خواص المتوسط الحسابي أنه قابل للمعالجة الرياضية على عكس الوسيط والمنوال.

- من عيوب المتوسط الحسابي أنن لا نتمكن من حسابه إلا بمعرفة قيم كل عناصر السلسلة الإحصائية ولم نر هذا في الوسيط.
- إن المتوسط الحسابي يتأثر بالقيم المتطرفة أو الشاذة فإذا كان لدينا مثلا السلسلة التالية:

1-2-1-4-3-1-4-19-19-19 فإن المتوسط الحسابي لهذه المجموعة هو:

$$\bar{x} = \frac{1+2+1+4+3+1+4+19+19+19}{10} = \frac{73}{10} = 7.3$$

نلاحظ أن متوسط هذه المجموعة ضعيف نسبيا ولو أننا حسبنا المتوسط الحسابي لهذه النقاط دون القيم الشاذة وهي19-19-19لأصبح الوسط الحسابي يساوي:

$$\bar{x} = \frac{1+2+1+4+3+1+4}{7} = \frac{16}{7} = 2,285$$

فالقيم الشاذة في المجموعة هي التي رفعت من مستوى المتوسط الحاسبي وجعلت المجموعة تبدو ضعيفة نسبيا.

المتوسط الحسابي التجميعي: إذا كان لدينا مجموعة إحصائية متوسطها الحسابي هو $\bar{x}_1$ وتتكون من n_1عنصرا ومجموعة ثانية متوسطها الحسابي $\bar{x}_2$وتحتوي على n_2عنصرا، يمكن دمج المجموعتين في مجموعة واحدة يصبح عدد عناصرها $N=n_1+n_2$ومتوسطها الحسابي:

$$\bar{x}_{1+2} = \frac{n_1.\bar{x}_1 + n_2.\bar{x}_2}{n_1 + n_2}$$

تعميم : المتوسط الحسابي التجميعي $\bar{x}_c$ **Moyenne arithmétique combinée** لـ k مجموعة جزئية هو :

$$\bar{x}_c = \bar{x}_{1+2+\cdots+K} = \frac{n_1.\bar{x}_1 + n_2.\bar{x}_2 + \cdots + n_{k-1}.\bar{x}_{k-1} + n_k.\bar{x}_k}{n_1 + n_2 + \cdots + n_{k-1} + n_k}$$

$$\Rightarrow \bar{\mathbf{x}}_{\mathbf{c}} = \frac{\sum_{i=1}^{k}[\mathbf{ni}.\bar{\mathbf{x}}_{\mathbf{i}}]}{\sum_{i=1}^{k}\mathbf{n}_{\mathbf{i}}}$$

3-لا يمكن تحديد المتوسط الحسابي بيانيا.

مثال 8: تنقسم احدى الشركات الى ثلاثة فروع في ثلاثة مناطق مختلفة. متوسط الأجر الشهري لعمال الفرع الأول هو 30000 دج، ومتوسطي الأجور الشهرية لعمال الفرع الثاني والثالث هما 36000 دج و42000 دج على الترتيب.

إذا علمت أن عدد عمال هذه الفروع الثلاثة هو على التالي: 200، 150 و130 عامل.

المطلوب: فما هو متوسط الأجر الشهري لعمال هذه الشركة؟

الحل:

نلاحظ أن حجم الفروع الثلاثة هو على الترتيب: 200، 150 و130. ة بالتالي فانه بعد دمجها مع بعضها البعض سوف تشكل مجموعة واحدة عدد مفرداتها:480 = 130 + 150 + 200. اذن المتوسط المطلوب بعد الدمج هو المتوسط التجميعي:

$$\bar{x}_c = \bar{x}_{1+2+3} = \frac{n_1.\bar{x}_1 + n_2.\bar{x}_2 + n_3.\bar{x}_3}{n_1 + n_2 + n_3}$$

$$\bar{x}_c = \bar{x}_{1+2+3} = \frac{[200\times30000]+[150\times36000]+[130\times42000]}{200+150+130}$$

$$\bar{x}_c = \bar{x}_{1+2+3} = 35125 \text{ دج}$$

متوسط الأجر الشهري لعمال الشركة هو: 35125 دج.

ملاحظة:

في هذا المثال، وفي التمارين المشابهة له، المتوسط العام الناتج عن دمج المجموعات (العينات) الجزئية لا يساوي متوسط المجموعات الجزئية إلا إذا كانت أحجام المجموعات متساوية، بمعنى:

$$\boldsymbol{n_1 = n_2 = n_3 = \cdots = n_k}$$

في المثال رقم 8 مثلا لدينا:

$$\bar{x}_c \neq \bar{x}_{1+2+3} \neq \frac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3}{3}$$

العلاقة بين $\bar{x}$ و M_e و M_o:

لاحظ كارل بيرسون أنه إذا كان التوزيع التكراري المدروس قريب من التماثل (التناظر) أي قليل الالتواء، هناك شبه علاقة دائمة Relation Empirique بين كل من مقاييس النزعة المركزية الثلاثة و هي:

$$\frac{\bar{x} - M_o}{\bar{x} - M_e} \approx 3$$

$$\bar{x} - M_o \approx 3\,(\bar{x} - M_e)$$

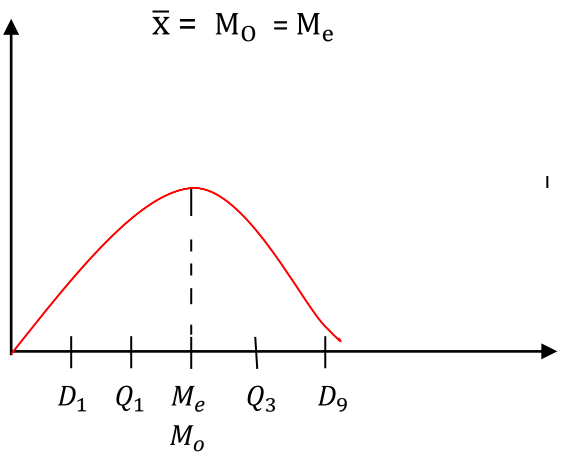
– عندما يكون التوزيع التكراري متماثلا (متناظرا) symétrique يتساوى مقاييس النزعة المركزية الثلاثة: $\bar{X} = M_o = M_e$ وتنطبق الأوساط الثلاثة على بعضها البعض **(الشكل 1)**.

– وعندما يختل التماثل يبتعد $\bar{x}$ عن M_o ويقع M_e بينهما ويكون M_e أقرب إلى $\bar{x}$:

أ. فإذا كان المنحنى التكراري ملتويا نحو اليسار oblique à gauche كان M_o أكبر من M_e و M_e أكبر من $\bar{x}$ **(الشكل 2)**.

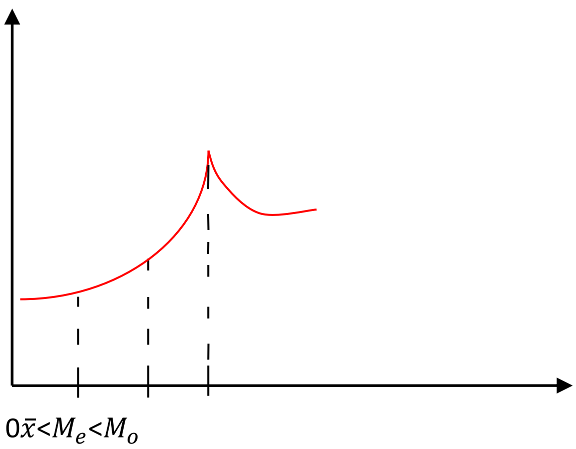
ب. إذا كان المنحنى التكراري ملتويا نحو اليمين oblique à droite كان M_o أصغر من M_e، و M_e أصغر من $\bar{x}$ **(الشكل 3)**.

حسب ما تظهره الأشكال البيانية التالية:



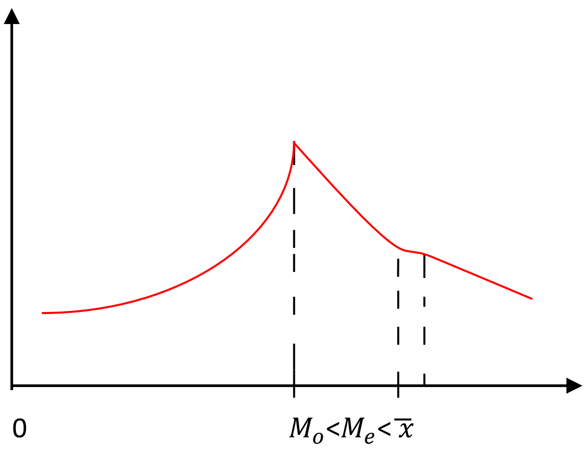
توزيع متماثل

الشكل (1)



منحنى ملتو نحو اليسار

الشكل (2)



منحنى ملتو نحو اليمين

الشكل (3)

V. <u>متوسطات أخرى:</u>(Autres moyenne)

1. <u>المتوسط التوافقي</u> : (H) <u>Moyenne harmonique</u>

تعريف : يعرّف المتوسط التوافقي لمجموعة من القيم على أنه مقلوب المتوسط الحسابي لمقلوبات تلك القيم أي $\frac{1}{x_i}$.

يرمز له ب H :

في حالة البيانات الغير المبوبة:

$$H = \frac{n}{\sum\left(\frac{1}{x_i}\right)}$$

حيث هو **n** عدد هذه المشاهدات.

في حالة البيانات المبوبة (التوزيعات التكرارية):

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

وبدلالة التكرارات النسبية، الوسط التوافقي يساوي:

$$H = \frac{1}{\sum \left(\frac{f_i}{x_i} \right)}$$

ملاحظة : في حالة التوزيع التكراري المعبر عنه بفئات، نعوض x_iبمراكز الفئات.

مجالات استخدام المتوسط التوافقي :

- لا يتم حساب المتوسط التوافقي إلا إذا كانت القيم لا تساوي الصفر (غير معدومة).
- لا يكون للمتوسط التوافقي مدلول منة حسابه إلا بوجود العلاقة العكسية ولمقلوب قيم x_i معنى حقيقي. إذن يستعمل المتوسط التوافقي عند حساب معدل السرعة، الأثمان (الأسعار) والنسب. مثلا :

أ. **عند حساب معدل السرعة:** نلاحظ أن المسافة مقسومة على السرعة لها معنى وتعطينا الزمن كذلك كلما زادت السرعة نقص الزمن والعكس صحيح.

ب. **عند حساب القدرة الشرائية:** بمعنى يكون المتوسط التوافقي مرتبط بالأسعار والأثمان. في هذه الحالة نلاحظ كلما ارتفع السعر انخفضت القدرة الشرائية والعكس صحيح. وهكذا...

مثال :

لنفرض أن شخص قطع مسافة من وهران إلى الجزائر العاصمة مستعملا الطريق السيار شرق-غرب، فكانت المسافة المقطوعة 424 كم ذهابا بسرعة متوسطة 100 كم/ساعة، ونفس المسافة إيابا بسرعة متوسطة مقدارها 90 كم/ساعة.

المطلوب : كم كانت سرعة الشخص المتوسطة ذهابا و إيابا ؟

الحل :

نحن نبحث عن متوسط السرعة الإجمالية ذهابا و إيابا، و انطلاقا من قانون السرعة المعروف يمكن أن نكتب :

$$\text{السرعة} = \frac{\text{المسافة}}{\text{الزمن}} = \frac{\mathbf{424 + 424}}{\frac{\mathbf{424}}{\mathbf{90}} + \frac{\mathbf{424}}{\mathbf{100}}} = \frac{\mathbf{848}}{\mathbf{4.711 + 4.24}} = \frac{\mathbf{848}}{\mathbf{8.951}}$$

$\Leftarrow$ **السرعة = 94.738 كم/ساعة.**

يمكن القول أن السرعة المتوسطة ذهابا و إيابا كانت 94.738 كم/ساعة.

ملاحظة:

- تعرف هذه القيمة المتوسطة (94.738) بالمتوسط التوافقي، الذي يرمز له بالرمز H.
- نلاحظ أن المتوسط الحسابي يعطي سرعة متوسطة تساوي $\frac{90+100}{2} = 95$ كم/سا ذهابا وإيابا، بخلاف الواقع الذي يتطابق مع النتيجة التي نحصل عليها بتطبيق المتوسط التوافقي وهي كالتالي:

$$H = \frac{n}{\sum\left(\frac{1}{x_i}\right)} = \frac{2}{\frac{1}{100}+\frac{1}{90}} = \frac{2}{\frac{2}{190}} = \frac{2}{0.021111} = 94.737 km/heure$$

إذن المتوسط المناسب لحساب السرعة المتوسطة هو المتوسط التوافقي.

2. <u>المتوسط التربيعي</u> :(متوسط المربع) (Q) <u>Moyenne quadratique</u>

<u>تعريف:</u> المتوسط التربيعي (Q) لمجموعة من القيم هو الجذر التربيعي لمتوسط مربعات هذه القيم.

في حالة البيانات الغير المبوبة:

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

حيث هو **n** عدد هذه المشاهدات.

في حالة البيانات المبوبة (التوزيعات التكرارية):

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

وباستخدام التكرارات النسبية، الوسط التربيعي يساوي:

$$Q = \sqrt{\Sigma[fi . xi^2]}$$

<u>ملاحظة</u> :في حالة التوزيع التكراري المعبر عنه بفئات، نعوض $\boldsymbol{x_i}$بمراكز الفئات.

مثال1 : أحسب المتوسط التربيعي للقيم : 12 -16 – 10 – 8 – 15.

$$Q = \sqrt{\frac{\sum x_i^2}{N}} = \frac{12^2 + 16^2 + 10^2 + 8^2 + 15^2}{5} = \sqrt{\frac{789}{5}} = \sqrt{157.8}$$

Q = 25.561

مثال2: الجدول التالي يمثل أجور 86 عاملا في مؤسسة صغيرة ومتوسطة (PME). أحسب المتوسط التربيعي.

فئات المبيعات (10 آلاف د.ج)	عدد المحلات n_i	مراكز الفئات x_i	x_i^2	$n_i.x_i^2$
[10-20[	12	15	225	2700
[20-30[	14	25	625	8750
[30-40[	16	35	1225	19600
[40-50[	20	45	2025	40500
[50-60[	12	55	3025	36300
[60-70[	8	65	4225	33800
[70-80[	4	75	5625	22500
المجموع	**86**			**164150**

$$\mathbf{Q = \sqrt{\frac{\sum n_i.x_i^2}{\sum n_i}} = \sqrt{\frac{164150}{86}} = \sqrt{1908.72} = 43.688}$$

مجالات استخدام المتوسط التربيعي :

- يُستعمل **Q** عندما تكون بعض القيم سالبة الإشارة والأخرى موجبة. ومن الأمثلة على ذلك في المصارف عندما تكون الحاجة إلى التعرف على حركة النقود في البنك من مسحوبات وودائع. فإذا كانت المسحوبات تساوي الودائع في يوم 28 فبراير2022 فإن حساب المتوسط الحسابي يعطينا الصفر ($\bar{x} = 0$). معنى ذلك أنه كأنما في يوم 02/28 لا توجد أي حركة نقدية في المصرف. ففي مثل هذه الحالات من الأفضل استخدام **Q** كمتوسط.
- يستعمل **Q** خاصة في المجالات الفيزيائية كحساب متوسط الأضلاع ومتوسط المساحات المربعة وغيرها.
- كما أننا نجد صيغة **Q** عند حساب التباين والانحراف المعياري (أنظر الفصل الموالي دراسة مقاييس التشتت).

3. المتوسط الهندسي: (متوسط الضرب) (G) Moyenne géométrique

تعريف : المتوسط الهندسي هو الجذر النوني لحاصل ضرب (N) قيمة معطاة. أو بعبارة أخرى هو حاصل ضرب القيم جذر القوة تساوي عدد القيم.

في حالة البيانات الغير المبوبة (المتوسط الهندسي للمشاهدات المفردة غير المكررة) :

اذا كان لدينا مجموعة من القيم **الموجبة تماما** $(x_1, x_2, x_3, ..., x_n)$المفردة و بدون تكرار فان متوسطها الهندسي هو الجذر النوني لحاصل ضرب هذه القيم، و يحسب بواسطة الصيغة التالية:

$$G = \sqrt[N]{\prod_{i=1}^{N} x_i} = \sqrt[N]{x_1 . x_2 . x_3 \dots x_{N-1} . x_N}$$

$$\Rightarrow G = (x_1 . x_2 . x_3 \dots x_{N-1} . x_N)^{\frac{1}{N}}$$

ولاستخراج قيمة Gنستعمل اللوغاريتمات، فإذا أخذنا لوغاريتم طرفي العلاقة السابقة نجد:

$$\log G = \log\left(x_1 \cdot x_2 \cdot x_3 \cdots x_N\right)^{\frac{1}{N}} = \frac{1}{N}\left(x_1 \cdot x_2 \cdot x_3 \cdots x_{N-1} \cdot x_N\right)$$

$$\log G = \frac{1}{N}\left(\log x_1 + \log x_2 + \log x_3 + \dots + \log x_{N-1} + \log x_N\right)$$

$$\log G = \frac{1}{N}\left(\sum_{i=1}^{N} \log x_i\right) = \frac{\sum \log x_i}{N}$$

نبحث عن العدد المقابل لـlog G في الجداول اللوغاريتمية أو بواسطة الآلة الحاسبة العلمية لنحصل على قيمة التي تمثل قيمة المتوسط الهندسي.

$$\Rightarrow G = 10^k$$

ملاحظة :و نذكر هنا بأنه لا يمكن حساب متوسط هندسي لمجموعة من القيم تحتوي على قيم معدومة أو سالبة.

مثال1 : أحسب المتوسط الهندسي للقيم : 12، 15، 9، 18.

الحل :

$$\log G = \frac{\sum log x_i}{N} = \frac{1}{4}(log x_1 + log x_2 + log x_3 + log x_4)$$

$$\log G = \frac{1}{4}(log 12 + log 15 + log 9 + log 18)$$

$$\log G = \frac{1}{4}(1.0791 + 1.1760 + 0.9542 + 1.2552)$$
$$\log G = \frac{1}{4}(4.4645) = 1.1161$$

ومنه : $G = 10^{1.1161}$

$$G = 13.065$$

حساب المتوسط الهندسي من بيانات المبوبة (متوسط هندسي مرجح) :

إذا كانت لدينا القيم $x_1, x_2, x_3, \ldots, x_k$ مع التكرارات المقابلة لها $n_1, n_2, n_3, \ldots, n_k$

$$\mathbf{G} = \sqrt[n_1+n_2+\cdots+n_k]{x_1^{n^1} . x_2^{n^2} . \ldots . x_k^{n^k}}$$

$$\mathbf{G} = \sqrt[\sum n_i]{x_1^{n^1} . x_2^{n^2} . \ldots . x_k^{n^k}}$$

$$\mathbf{G} = \sqrt[\sum n_i]{\prod_{i=1}^{k} x_i^{n_i}}$$

$$\mathbf{G} = \left[\prod_{i=1}^{k} x_i^{n_i}\right]^{\frac{1}{\sum n_i}}$$

و بأخذ لوغاريتم الطرفين يكون لدينا :

$$\log G = \frac{1}{\sum n_i} = \left(\log x_1^{n_1} + \log x_2^{n_2} + \log x_3^{n_3} + \cdots + \log x_K^{n_K}\right)$$

$$\log G = \frac{1}{\sum n_i} = (n_1 . \log x_1 + n_2 . \log x_2 + n_3 . \log x_3 + \cdots + n_K . \log x_K)$$

$$\log G = \frac{\sum [n_i . \log x_i]}{\sum n_i} \quad (1)$$

ومنه نستخرج قيمة G كما فعلنا في السابق.

⇐ المتوسط الهندسي لمجموعة من القيم $x_1, x_2, x_3, \dots, x_k$ تتكرر حسب $n_1, n_2, n_3, \dots, n_k$ هو المتوسط الحسابي للوغاريتمات هذه القيم.

ملاحظة: في حالة التوزيع التكراري المعبر عنه بفئات، نستخدم الصيغة (1) على أن تدل $\boldsymbol{x_i}$ على مراكز الفئات بدلا من القيم. بمعنى:

$$\boldsymbol{logG = \frac{\sum[n_i . logx_i]}{\sum n_i}}$$

حيث $\boldsymbol{x_i}$: مركز الفئة. ومن ثم نستخرج قيمة المتوسط الهندسي كما فعلنا سابقا.

مجالات استخدام المتوسط الهندسي:

من أهم مجالات استخدام المتوسط الهندسي نجده يستعمل في حساب المتوسطات الخاصة بالأرقام القياسية وخاصة السلاسل التي تتغير حسب متتالية هندسية، حيث يعد من أحسن المتوسطات في هذا المجال: كما أنه يستعمل في حساب معدلات التغير والنسب، كمعدل النمو السكاني، معدل نمو الإنتاج، أو نسبة الزيادة أو(النقصان) في سعر سلعة ما، وغيرها من المجالات المشابهة، ولا يستخدم مع القيم السالبة أو المعدومة.

يعتبر المتوسط الهندسي مقياسا مهما في بعض الحالات التي لا يمكن معها استخدام المتوسط الحسابي كمؤشر له مدلول إحصائي، وبصفة خاصة تلك الحالات التي يكون فيها لضرب القيم معنى وتفسير إحصائي.

4. العلاقة بين المتوسطات:

إذا كان لدينا مجموعة من القيم، بدون أو مع وجود تكرارات، أو بيانات مبوبة في فئات، عند حساب كل

$$\mathbf{H} < G < \overline{\boldsymbol{X}} < Q$$

مثال 2 : أحسب المتوسط الهندسي من المعطيات التالية : 12، 16، 10، 8، 15.

المطلوب: أحسب كل من المتوسط الحسابي ($\overline{x}$)، المتوسط التوافقي (H)، المتوسط التربيعي (Q) والمتوسط الهندسي (G). ما ذا تلاحظ؟

الحل :

x_i	$\frac{1}{x_i}$	x_i^2	$\log x_i$
8	$= 0.125\frac{1}{8}$	64	log 8 = 0.9030
10	0.1	100	1
12	0.083	144	1.0791
15	0.066	225	1.1760
16	0.062	256	1.2041
61	**0.436**	**789**	**5.3622**

المتوسط الحسابى:

$$\bar{x} = \frac{\sum_{i=1}^{i=7} n_i x_i}{\sum_{i=1}^{i=7} n_i} = \frac{61}{5} = 12.2$$

المتوسط التوافقى:

$$H = \frac{N}{\sum\left(\frac{1}{x_i}\right)} = \frac{5}{0.436} = 11.647$$

المتوسط التربيعى:

$$Q = \sqrt{\frac{\sum x_i^2}{N}} = \sqrt{\frac{789}{5}} = \sqrt{157.8} = 12.561$$

المتوسط الهندسى:

$$\log G = \frac{\sum \log x_i}{N} = \frac{5.3622}{5} = 1.07244$$

$$G = 10^{1.07244} = 11.815$$

نلاحظ أن العلاقة التي تربط بين المتوسطات المختلفة محققة:

$$\begin{array}{ccccccccccc} \boldsymbol{X}_{\mathbf{Min}} & < & H & < & G & < & \overline{\boldsymbol{X}} & < & Q & < & \boldsymbol{X}_{\boldsymbol{Max}} \\ \downarrow && \downarrow && \downarrow && \downarrow && \downarrow && \downarrow \\ \mathbf{8} && \mathbf{11.467} && \mathbf{11.815} && \mathbf{12.2} && \mathbf{12.561} && \mathbf{16} \end{array}$$

مثال3 : يمثل الجدول التالي توزيع 60 موظف حسب أجورهم الشهرية (ب 10^2 دج).

التكرار n_i	الفئات (الأجور) دج10^3
10	**[10– 20 [**
30	**[20– 30 [**
20	**[30 – 40 [**
60	**المجموع**

المطلوب: أحسب كل من المتوسط الحسابي ($\bar{x}$)، المتوسط التوافقي (H)، المتوسط التربيعي (Q) والمتوسط الهندسي (G). ما ذا تلاحظ؟

الحل :

الفئات (الأجور) دج10^3	n_i	x_i	$n_i.x_i$	$\frac{n_i}{x_i}$	x_i^2	$n_i.x_i^2$	$\log x_i$	$n_i.\log x_i$
[10– 20 [	10	15	150	0.666	225	2250	1.1760	11.76
[20– 30 [	30	25	750	1.2	625	18750	1.3979	41.938
[30 –40 [	20	35	700	0.571	1225	24500	1.5440	30.881
Σ	**60**	**/**	**1600**	**2.437**	**/**	**45500**	**/**	**84.579**

المتوسط الحسابي:

$$\bar{x} = \frac{\sum_{i=1}^{i=7} n_i x_i}{\sum_{i=1}^{i=7} n_i} = \frac{1600}{60} = 26.666$$

المتوسط التوافقي:

$$H = \frac{\sum n_i}{\sum \left(\frac{n_i}{x_i}\right)} = \frac{60}{2.437} = 24.62$$

المتوسط التربيعي:

$$Q = \sqrt{\frac{\sum n_i . x_i^2}{\sum n_i}} = \sqrt{\frac{45500}{60}} = \sqrt{758.333} = 27.537$$

المتوسط الهندسي:

$$logG = \frac{\sum [n_i . log x_i]}{\sum n_i} = \frac{84.579}{60} = 1.4096$$

$$G = 10^{1.4096} = 25.683$$

نلاحظ أن العلاقة التي تربط بين المتوسطات المختلفة محققة :

$$X_{Min} < H < G < \bar{X} < Q < X_{Max}$$

15 25.683 27.537

26.62 26.666 35

مثال 4 :عرف حجم التبادل التجاري بين بلدين ارتفاعا سنويا قدر ب:

- 5 % خلال الثلاثة سنوات الأولى،
- 7 % خلال الخمس سنوات الموالية،
- 9 % خلال السنتين الأخيرتين.

المطلوب :

ما هو حجم الارتفاع السنوي المتوسط لحجم التبادل التجاري بين البلدين خلال العشر سنوات؟ فسر النتيجة.

الحل: طالما نحن نتعامل مع تغيرات معبر عنها بنسب مئوية فإننا نستخدم المتوسط الهندسي من أجل حساب مدى التغير (الارتفاع) في حجم التبادل بين البلدين، و بالتالي فإن المتوسط المناسب الذي يعبر عن معدل الارتفاع السنوي خلال هذه الفترة المقدرة بعشر سنوات هو المتوسط الهندسي كما يلي :

$$G = \sqrt[\sum n_i]{\prod_{i=1}^{k} x_i^{n_i}} - 1 = \sqrt[\sum n_i]{{x_1}^{n^1} . {x_2}^{n^2} . \ldots . {x_k}^{n^k}} - 1$$

$$G = \sqrt[10]{(1+0.05)^3 x (1+0.07)^5 x (1+0.09)^2} - 1 \Rightarrow$$

$$G = \sqrt[10]{1.929033} - 1 \Rightarrow$$

$$G = 1.0679 - 1 \Rightarrow$$

$$G = 0.0679 = 6.79\ \% \Rightarrow$$

التفسير :

وهو معدل الارتفاع السنوي المتوسط في حجم التبادل التجاري بين هذين البلدين خلال العشر سنوات المعنية بالدراسة.

الفصل الخامس

مقاييس التشتت

رأينا أن الهدف من مقاييس النزعة المركزية $(M_o, M_e, \bar{x})$ هو قياس مستوى الظاهرة المدروسة، فمثلا نقيس مستوى الأجور في مؤسسة اقتصادية بواسطة الأجر المتوسط، ونقيس مستوى الدخل في بلد ما بواسطة الدخل المتوسط، ونقيس مستوى الطالب في آخر السنة بواسطة المعدل وهو متوسط حسابي...إلخ ، إلا أن هذه المقاييس تشوبها مساوئ عدة، أهمها إخفاء الفروق الموجودة بين القيم، فعند إجراء مقارنة بين ظاهرتين يمكن أن يتساوى متوسطهما الحسابي، ورغم ذلك نجد أن انتشار البيانات في الظاهرتين مختلف كثيرا لأن البيانات غير متجانسة، لهذا وجدت مقاييس أخرى تعطينا فكرة عن مدى تباعد البيانات عن بعضها البعض، تسمى هذه المؤشرات بمقاييس التشتت.

<u>**مثال 1 :**</u> إذا كان لدينا توزيعين A و B يمثلان مجموعة من المصابيح الكهربائية مع فترات الإضاءة لكل مصباح بالساعة.

السلسلة A: 780- 790- 790- 800- 800- 800- 810- 810- 820

السلسلة B: 400- 600- 600- 800- 800- 800- 1000- 1000- 1200

نلاحظ أن التوزيعين لهما نفس مقاييس النزعة المركزية.

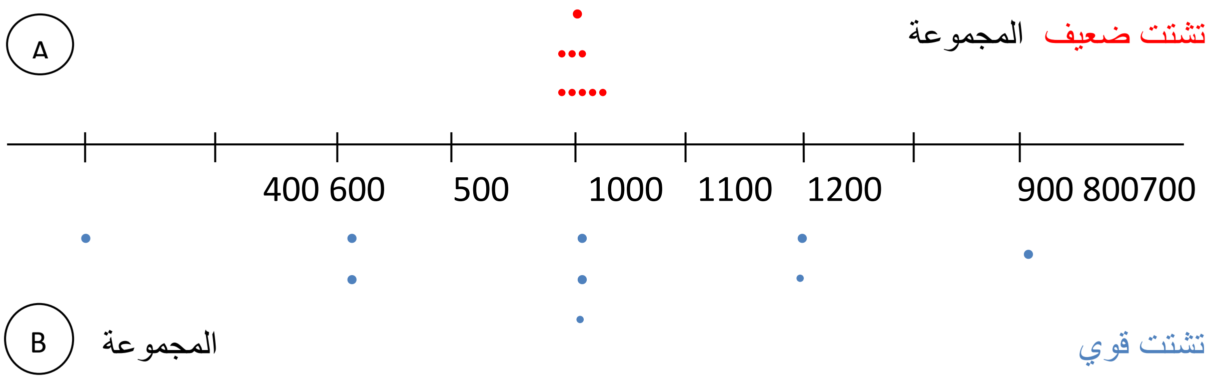
$$\bar{x}_A = \bar{x}_B = 800$$

$$Me_A = Me_B = 800$$

$$Mo_A = Mo_B = 800$$

إذن إن الخصائص الثلاثة للنزعة المركزية متساوية في التوزيعين $\mathbf{M_e = M_o = \bar{x}}$، غير أن طول مجال الدراسة في التوزيعين (المدى العام E) يختلف بين التوزيعين. معنى ذلك أن التوزيعين يختلفان اختلافا كبيرا من حيث انتشار وتوزيع قيمهما على مجال الدراسة.

نضع معطيات السلسلتين على خط :



إن قيم المجموعة B منتشرة ومتباعدة حول متوسطها، بينما قيم المجوعة**A**متجمعة حول متوسطها عندئذ يقال ⇐ إن المجموعة **A** أقل تشتت من المجموعة **B**، بعبارة أخرى إن قيم المجموعة **A** أكثر تجانس من قيم المجموعة **B**.

خلاصة : لمقارنة مجموعتين إحصائيتين أو أكثر، لا نكتفي بمقاييس النزعة المركزية ($\bar{x}$، M_e وM_o) و حتى نستكمل دراسة التوزيع الإحصائي، نتطرق إلى مقاييس التشتت.

سنكتفي بالمقاييس الأكثر استعمالا.

سنتطرق إلى مقاييس التشتت حسب درجة أهميتها، فمن الأقل أهمية إلى الأكثر أهمية.

تقسم مقاييس التشتت إلى قسمين:

I - **مقاييس تشتت مطلقة.**

II - **مقاييس نشتت نسبية.**

I)<u>مقاييس التشتت المطلقة</u> :

<u>تنقسم بدورها إلى مجموعتين</u> :

1–<u>مقاييس التشتت التي تقيس تقارب أو تباعد القيم عن بعضها البعض</u>:

أ–<u>المدى</u> : Etendu

هو الفرق بين أكبر وأصغر قيمة من مفردات السلسلة.

$$E = Max_{(xi)} - Min_{(xi)}$$

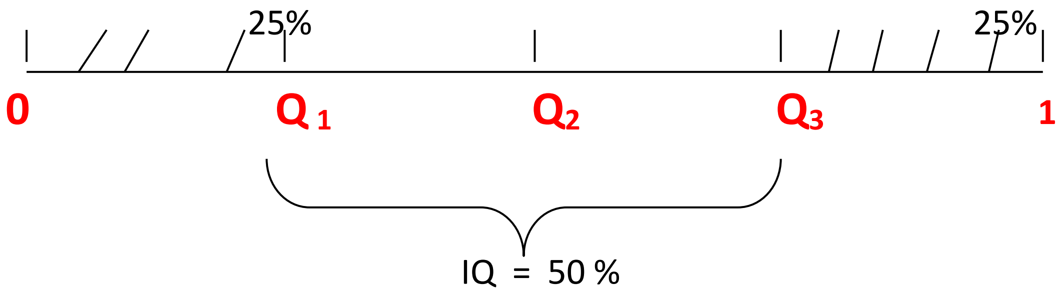
$$\begin{cases} E_A = 820 - 780 = 40 \\ E_B = 1200 - 400 = 800 \end{cases}$$

نستنتج أن السلسلة A أكثر تجانس وأقل تشتت من السلسلة B.

من مميزات المدى E أنه سهل الحساب، ومن عيوبه أنه يعتمد على القيمتين المتطرفتين فقط والتي من الممكن أن تكون قيم شاذة **Valeurs aberrantes.**

ب–<u>المدى (أو المجال الربيعي)</u> : Intervalle interquartile

$$I_Q = Q_3 - Q_1$$



يضم المدى الربيعي 50% من المجتمع مهما كان التوزيع الإحصائي. للتخلص من القيم المتطرفة نحسب I_Q .

كلما كان I_Q صغيرا كلما كان التشتت ضعيفا و التمركز قويا.

I_Q استعمالاته محدودة نظرا لبساطته، غير أنه أحسن من المدى العام.

يستعمل I_Q في المقارنة بين توزيعين أو أكثر. من مزايا هذا المقياس أنه لا يعتمد على القيم المتطرفة، و لكن من عيوبه أنه يأخذ بعين الاعتبار فقط50% من قيم السلسلة.

ملاحظة : في بعض الأحيان نحسب المدى **(المجال) العشري Intervalle interdécile** و الذي يضم 80% من المعطيات :

$$\boxed{I_D = D_9 - D_1}$$

أو **المدى المئيني Intervalle interpercentile** و الذي يضم 98% من المعطيات:

$$\boxed{I_P = P_{99} - P_1}$$

2 -مقاييس التشتت التي تقيس قرب أو بعد القيم من قيمة مركزية غالبا الوسط الحسابي :

مثال : لدينا نقاط (علامات) لستة طلبة في الإحصاء :

2 -17 -7 -18 - 3 -13

متوسط النقاط هو : $\bar{x} = \frac{\sum X_i}{N} = \frac{60}{6} = 10$

هل يمكن استنتاج أن هذه المجموعة متجانسة؟

بمعنى هل الطلبة لديهم نفس المستوى؟ الجواب لا، لأن 50% من الطلبة فقط لديهم المعدل.

لقياس التشتت حول $\bar{x}$ نحسب مختلف الانحرافات بين القيم و $\bar{x}$ نحصل على :

x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
2	− 8	64
17	7	49
7	− 3	9
18	8	64
3	− 7	49
13	3	9
Σ	**0**	**244**

مجموع انحرافات القيم عن $\bar{x}$ = 0 : $\sum(x_i - \bar{x}) = 0$

—— هناك **حل أول** يتمثل في أخذ **القيم المطلقة valeurs absolues** لهذه الانحرافات $|x_i - \bar{x}|$ لكن بما أن القيم المطلقة لا تخضع للحسابات الجبرية، لذلك نفضل **حل ثاني**و هو أن نقوم بتربيع الانحرافات. $(x_i - \bar{x})^2$.

عند حساب متوسط الانحرافات المربعة نحصل على التباين (Variance).

أ–<u>التباين variance</u> :

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

a–<u>علاقة التعريف</u> : Formule de définition

<u>بيانات غير مبوبة:</u>

$$V(x) = \frac{\sum(x_i - \bar{x})^2}{N}$$

<u>بيانات مبوبة:</u>

$$V(x) = \frac{\sum[n_i.(x_i - \bar{x})^2]}{\sum n_i}$$

التباين باستخدام التكرار النسبي

$$V(x)=\sum f_i(x_i-\overline{x})^2$$

ملاحظة : إن وحدة التباين هي مربع وحدة البيانات. مثال تباين المتغير الأجور وحدته DA^2 إذا أردنا نفس وحدات البيانات نأخذ الجذر التربيعي للتباين فنحصل على الانحراف المعياري.

ب– الانحراف المعياري : Ecart type

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

ملاحظة : وحدة الانحراف المعياري هي نفس وحدة البيانات. يستعمل الانحراف المعياري لمقارنة توزيعين تكراريين أو أكثر.
كلما كانت قيمة الانحراف المعياري كبيرة، كلما كان تشتت القيم حول وسطها الحسابي كبيرا. (قيم متباعدة) والعكس صحيح.

مثال : أحسب الانحراف المعياري للمثال السابق :

$$\delta_x=\sqrt{\frac{\sum(x_i-\overline{x})^2}{N}}=\sqrt{\frac{(-8)^2+7^2+(-3)^2+8^2+(-7)^2+3^2}{6}}$$

$$\Rightarrow \delta_x=\sqrt{40,66}=\boxed{6,37}$$

التفسير : بعض الطلبة (الممتازين) يكون لديهم العلامة (بالتقريب) المتوسطة 10 زائد 6,37،
الآخرين (les mauvais) لديهم العلامة 10 ناقص 6,37.

$$\bar{x}-\delta_x \text{ et } \bar{x}+\delta_x$$

(10 – 6,37) et (10 + 6,37)

العلامات محصورة بين 3,63 و 16,37

مثال 2 : لدينا النقاط التالية في الإحصاء لمجموعة ثانية من الطلبة : 11 – 9 – 12 – 8
المطلوب : أحسب الانحراف المعياري ثم قارنه مع الانحراف المعياري للمجموعة الأولى .

$$\bar{x}_2 = \frac{\Sigma xi}{N} = \frac{40}{4} = \boxed{10}$$

تباين المجموعة الثانية :

$$V(x)_2 = \frac{\Sigma(xi - \bar{x})^2}{4} = \frac{1}{4}[(8-10)^2 + (12-10)^2 + (9-10)^2 + (11-10)^2]$$

$$\Rightarrow \quad V(x)_2 = \frac{10}{4} = 2,5.$$

الانحراف معياري للمجموعة 2 :

$$\delta_{x2} = \sqrt{2,5} = \boxed{1,581}$$

المقارنة :

$\delta_{x1} = 6,37$ et $\delta_{x2} = 1,581$

$$\delta_{x1} < \delta_{x2}$$

الاستنتاج: المجموعة الثانية أكثر تجانس من المجموعة الأولى، لأن تشتت النقاط في المجموعة 1 أكثر من تشتت النقاط في المجموعة 2 بأربع مرات.

القانون المنشور : علاقة كونينغ Formule développée : Relation de Köning

بيانات غير مبّوبة : $V(x) = \delta_x^2 = \frac{\Sigma xi^2}{N} - \bar{x}^2$

بيانات مبّوبة : $V(x) = \delta_x^2 = \frac{\Sigma xi.xi^2}{\Sigma xi} - \bar{x}^2$

باستخدام التكرارات النسبية: $V(x) = \delta_x^2 = \Sigma Fi.x^2 - \bar{x}^2$

ملاحظة: إذا كان $\delta_x = \sqrt{V(x)} \Rightarrow V(x) = \delta_x^2$

<u>مثال 3</u> : لتكن المعطيات التالية : 15 -11 -10 -9 -5

المطلوب: أحسب الانحراف المعياري بواسطة :

1) علاقة التعريف　　　　　　　　2) القانون المنشور

x_i	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	x_i^2
5	-5	25	25
9	-1	1	81
10	0	0	100
11	1	1	121
15	5	25	225
50	**/**	**52**	**552**

<u>الحل:</u>

1)<u>علاقة التعريف :</u>

$$\delta_x = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{N}}$$

$$= \frac{\Sigma xi}{N}\bar{x}$$

$$= \frac{50}{5}\bar{x} = 10$$

N=5

$$\delta_x = \sqrt{\frac{52}{5}} \Rightarrow$$

$$\delta_x = \sqrt{10,4} \Rightarrow \Rightarrow \boxed{\delta_x = 3,224}$$

2)<u>القانون المنشور</u> :

$$\delta_x = \sqrt{\frac{\Sigma . x_i^2}{N} - \bar{x}^2} = \sqrt{\frac{552}{5} - 10^2}\sqrt{110{,}4 - 100}$$

$$\Rightarrow \quad \delta_x = \sqrt{10{,}4} \quad \Rightarrow \quad \boxed{\delta_x = 3{,}224}$$ **نفس النتيجة السابقة**

<u>مثال 4</u> : لدينا الجدول التالي :

الفئات	n_i
[10-20[	2
[20-30[	5
[30-40[	3
Σ	**10**

المطلوب : أحسب الانحراف المعياري بواسطة :

1) علاقة التعريف　　　　2) القانون المنشور

الفئات	n_i	x_i	$n_i . x_i$	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$	$n_i . (x_i - \bar{x})^2$	$n_i . x_i^2$
[10-20[	2	15	30	-11	121	242	450
[20-30[	5	25	125	-1	1	5	3125
[30-40[	3	35	105	9	81	243	3675
Σ	**10**	**/**	**/**	**/**	**/**	**490**	**7250**

<u>الحل:</u>

1)<u>علاقة التعريف</u> :

$$\bar{x} = \frac{\Sigma ni . xi}{\Sigma ni} = \frac{260}{10} = \boxed{26}$$

$$\delta_X=\sqrt{\frac{\Sigma[ni.(xi-\bar{x})^2]}{\Sigma ni}} = \sqrt{\frac{490}{10}} = \sqrt{49} = \boxed{7}$$

2)القانون المنشور : Koening

$$\delta_X= \sqrt{\frac{\Sigma n_i.x_i^2}{\Sigma ni}-\bar{x}^2} = \sqrt{\frac{7250}{10}-26^2}= \boxed{\mathbf{7}}$$

نفس النتيجة السابقة.

ملاحظة: يعترض تطبيق علاقة التعريف أحيانا حسابات طويلة و شاقة خاصة عندما تكون $\bar{x}$ قيمة عشرية valeur décimale لذلك من الأفضل استعمال القانون المنشور لحساب التباين و الانحراف المعياري و ذلك لسهولة العمليات الحسابية.

II–مقاييس التشتت النسبية :

إن المقاييس التي قمنا بدراستها لحد الآن هي مقاييس تشتت مطلقة ونستعملها عندما نريد مقارنة مجموعتين لهما نفس وحدات القياس، كأن نقارن أجور مجموعة A أجورهم بالدج مع مجموعة ثانية B أجورهم معّبر عنها كذلك بالدينار الجزائري (دج).

أما إذا تعلق الأمر بمجموعة نعّبر عن أجورها بالدج مع مجموعة ثانية نعّبر عن أجورها بالأورو (عملة أخرى) ونطلب المقارنة.

في هذه الحالة مقياس التشتت المطلق (الانحراف المعياري δ_X) لا يصلح للمقارنة ⇐ يجب أن نستعمل مقياس تشتت نسبي حتى نحصل على رقم خال من وحدات القياس.
يوجد عدة معاملات للتشتت النسبي، نذكر:

1) **معامل المدى الربيعي النسبي Coefficient interquartile relatif**

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

2) معامل الإختلاف (المعياري)   Coefficient de variation

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

ملاحظة: إن معامل الاختلاف CV يستعمل بشكل واسع في التحليل الإحصائي لمقارنة توزيعين أو أكثر.

مثال5: نفترض أننا نريد مقارنة أجور عمال قطاع معّين في الجزائر و في بلد أوروبي، نفترض أننا حسبنا كل من $\bar{x}$ و δ_x لهذه الأجور في كل بلد و وجدناهم حسب الجدول التالي :

الجزائر: $\bar{x}$ = 25000 DA ، δ_x = 350 DA

بلد أوروبي : $\bar{x}$ = 25**euros** ، δ_x = 7000 **euros**

قارن بين الأجور في البلدين.

الحل :فضلا على أنه لا يمكننا أن نقارن DA مع **euros** نحسب معامل التشتت النسبي ⇐ و هو معامل الاختلاف المعياري في كلتا الحالتين و ذلك حتى نحصل على رقم خال من وحدات القياس :

$$CV_{Algérie} = \frac{\delta x}{\bar{x}} \times 100 = \frac{350}{25000} \times 100 = \underline{1,4\ \%}$$

$$CV_{Pays\ européen} = \frac{25}{7000} \times 100 = \underline{0,357\ \%}$$

$$\Rightarrow \quad \mathbf{CV_{Pays\ européen} < CV_{Algérie}}$$

نلاحظ أن الاختلاف والتشتت في الأجور عند العمال في البلد الأوروبي أقل منه تشتت الأجور في الجزائر ⇐ الأجور في البلد الأوروبي أكثر تجانس.

مثال5: في سنة معينة كان هناك أستاذا جامعيا يتقاضى راتبا شهريا وسطا قدره 46256 و.ن، و كان الانحراف المعياري 4334، و كان عمال صناعة النسيج في نفس الفترة يتقاضون 16.38و.ن/سا، و كان الانحراف المعياري 2.89 و.ن، فأي هما تشتته أكبر، أجر عمال قطاع النسيج أم راتب الأستاذ الجامعي ؟ ما ذا تستنتج؟

الحل :

بما أن أجور الأساتذة أعدت على أساس شهري وأجور عمال قطاع النسيج أعد على أساس الساعة، فإن السلسلتين غير قابلتين للمقارنة مباشرة، ولذلك نستعمل معامل الاختلاف (التغير):

$$\mathbf{CV_{Ens.\ Univ}} = \frac{\delta x}{\bar{x}} \times 100 = \frac{4334}{46256} \times 100 = \underline{9.36\ \%}$$

$$\mathbf{CV_{Trav.\ Textiles}} = \frac{2.89}{16.38} \times 100 = \underline{17.64\ \%}$$

$$\mathbf{CV_{Ens.\ Univ}} < \mathbf{CV_{Trav.\ Textiles}}$$

الاستنتاج: و منه فان رواتب الأساتذة الجامعيين أكثر تجانسا من أجور عمال قطاع النسيج، و هذا غير مفاجئ، لأن عمال النسيج يتميزون عن بعضهم البعض بالعمر و التأهيل المهني و نوع العمل، بينما الأساتذة الجامعيين يشكلون مجموعة متجانسة نسبيا.

تمرين1:

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

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

المطلوب:

أحسب الانحراف المعياري: باستخدام علاقة التعريف ثم باستخدام القانون المنشور (علاقة كونينغ).

تمرين2:

إذا كان متوسط الأجور ل 100 عامل في مشروع معين 24 دج بانحراف معياري 1,56 دج. أخذت عينة من العاملات في نفس المشروع وتبين أن أجورهن موزعة كالتالي:

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

المطلوب: قارن بين التوزيعين باستخدام معامل الاختلاف المعياري. ماذا يمكن استنتاجه؟ (هل تتفق مع الرأي القائل بأن توزيع أجور العمال أكثر عدالة من توزيع أجور العاملات؟)

تمرين3:

لدينا المعلومات التالية عن توزيع تكراري متماثل لأجور 50 عامل، حيث أن:

V(x) = Variance = 18,6 = التباين و أن $\sum$ **(ni. x_i^2) = 1820**

أوجد الوسيط ثم المنوال.

حل التمرين1:

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

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

حساب الانحراف المعياري:

حساب الانحراف المعياري بعلاقة التعريف:

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

حساب الانحراف المعياري بقانون المنشور: **Koening**

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

حل التمرين2:

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

معامل الاختلاف المعياري لأجور العمال (الرجال):

متوسط اجور الرجال: $\bar{\mathrm{x}}_{H} = 24\ DA$

انحراف معياري الرجال: $\delta_{x_H} = 1{,}56\ DA$

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

معامل الاختلاف المعياري لأجور العاملات (النساء):

المتوسط الحسابي العاملات (النساء):

$$\bar{\mathrm{x}}_{\mathrm{F}} = \frac{\sum \mathrm{n_{i_F} x_i}}{\sum \mathrm{n_{i_F}}} = \frac{2544}{100} \Longrightarrow \boxed{\bar{\mathrm{x}} = 25{,}44}$$

انحراف معياري للعاملات (النساء):

$$\delta_{x_F} = \sqrt{\frac{\sum \mathrm{n_{i_F}} x_i^2}{\sum \mathrm{n_{i_F}}} - \bar{\mathrm{x}}_{\mathrm{F}}{}^2} = \sqrt{\frac{65420}{100} - (25{,}44)^2} = 2.64$$

معامل الاختلاف المعياري لأجور العاملات (النساء):

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

المقارنة:

$$CV_H < CV_F$$

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

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

اذن اجور الرجال اقل تشتتا (اقل تفاوت) وأنها متقاربة . ان توزيع اجور العمال الرجال أكثر عدالة من توزيع اجور العاملات.

حل التمرين3:

توزيع متماثل معنى ذلك:

$$\bar{\mathrm{x}} = \mathrm{M_o} = \mathrm{M_e}$$

$$\sum \mathrm{n_i} = 50, \qquad \sum \mathrm{n_i} x_i^2 = 1820, \qquad V(x) = 18{,}6$$

قانون المنشور للتباين: $V(x) = \frac{\sum \mathrm{n_i} x_i^2}{\sum \mathrm{n_i}} - \bar{x}^2$

بالتعويض:

$$V(x) = \frac{\sum \mathrm{n_i} x_i^2}{\sum \mathrm{n_i}} - \bar{x}^2 \Longrightarrow 18{,}6 = \frac{1820}{50} - \bar{x}^2$$

$$\Longmapsto 18{,}6 = 36{,}4 - \bar{x}^2$$

$$\Longmapsto \bar{x}^2 = 36{,}4 - 18{,}6 = 17{,}4$$

$$\Longmapsto \bar{\mathrm{x}} = \sqrt{17{,}8} = 4{,}219$$

بما ان التوزيع متماثل:

$\bar{\mathrm{x}} = \mathrm{M_o} = \mathrm{M_e} = 4{,}219$ وهو المطلوب

الفصل السادس

ملحق الرياضيات

تذكير ببعض الخواص الجبرية لدليل الجمع Σ

الخاصية 1: si a est une constante $\sum_{i=1}^{n} a = n.a$

لأن

$$\sum_{i=1}^{n} a = a + a + a + ... + a = n.a$$

الخاصية2: $\sum_{i=1}^{n} a.x_i = a\sum_{i=1}^{n} x_i$

لأن

$$\sum_{i=1}^{n} a.x_i = a.x_1 + a.x_2 + ... + a.x_n = a.(x_1 + x_2 + ... + x_n) = a.\sum_{i=1}^{n} x_i$$

الخاصية3: $\sum_{i=1}^{n} (x_i + y_i) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i$

لأن

$$\sum_{i=1}^{n} (x_i + y_i) = (x_1 + y_1) + (x_2 + y_2) + ... + (x_n + y_n)$$

$$= (x_1 + x_2 ..... + x_n) + (y_1 + y_2 + ... + y_n) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i$$

ملاحظات:

1) $\sum_{i=1}^{n} x_i . \sum_{i=1}^{n} y_i \neq \sum_{i=1}^{n} x_i . y_i$

لأن

$$\sum_{i=1}^{n} x_i . \sum_{i=1}^{n} y_i = (x_1 + x_2 + x_3 ..... + x_n).(y_1 + y_2 + y_3 ...... + y_n)$$

$$\Rightarrow \sum_{i=1}^{n} x_i . y_i = (x_1.y_1)+(x_2.y_2)+(x_3.y_3).....+(x_n.y_n)$$

2) $\left(\sum_{i=1}^{n} x_i\right)^2 \neq \sum_{i=1}^{n} x_i^{\,2}$

لأن

$$\left(\sum_{i=1}^{n} x_i\right)^2 = (x_1+x_2+x_3+.......+x_n)^2$$

$$\Rightarrow \sum_{i=1}^{n} x_i^{\,2} = x_1^2+x_2^2+x_3^2+.......+x_n^2$$

<u>مثال1</u>:<u>لتكن</u> السلسلة الإحصائية التالية :

$$x_1=1;\ \ x_2=4;\ \ x_3=2;\ \ x_4=3;\ \ x_5=5$$

أنشر ما يلي:

1) $\sum_{i=1}^{5} x_i$ 2) $\frac{1}{5}\sum_{i=1}^{5} x_i$ 3) $\sum_{i=1}^{5}(x_i-3)$ 4) $\sum_{i=1}^{5} x_i^{\,2}$

5) $\sum_{i=1}^{5} 4x_i$ 6) $\frac{1}{5}\sum_{i=1}^{5}(x_i-3)^2$ 7) $\frac{1}{5}\sum_{i=1}^{5}(x_i-3)^3$ 8) $\frac{1}{5}\sum_{i=1}^{5}|x_i-3|$

<u>الحل</u>:

$x_1=1;\ \ x_2=4;\ \ x_3=2;\ \ x_4=3;\ \ x_5=5$

1) $\sum_{i=1}^{5} x_i = x_1+x_2+x_3+x_4+x_5 = 1+4+2+3+5=15$

2) $\frac{1}{5}\sum_{i=1}^{5} x_i = \frac{1}{5}(x_1+x_2+x_3+x_4+x_5) = \frac{1}{5}\times 15 = 3$

$$3)\sum_{i=1}^{5}(x_i-3)=\sum_{i=1}^{5}x_i-\sum_{i=1}^{5}3$$

$$=\sum_{i=1}^{5}x_i-(5\times3)=15-15=0$$

$$4)\sum_{i=1}^{5}x_i^2=x_1^2+x_2^2+x_3^2+x_4^2+x_5^2$$

$$\sum_{i=1}^{5}x_i^2=1^2+4^2+2^2+3^2+5^2=55$$

ملاحظة:

$$\sum_{i=1}^{5}x_i^2\neq\left(\sum_{i=1}^{5}x_i\right)^2\rightarrow 55\neq(15)^2$$

$$5)\sum_{i=1}^{5}4x_i=4\sum_{i=1}^{5}x_i=4\times15=60$$

$$6)\frac{1}{5}\sum_{i=1}^{5}(x_i-3)^2=\frac{1}{5}\sum_{i=1}^{5}(x_i^2-2\times3x_i+9)$$

$$=\frac{1}{5}\left[\sum_{i=1}^{5}x_i^2-6\sum_{i=1}^{5}x_i+\sum_{i=1}^{5}9\right]$$

$$=\frac{1}{5}\left[\sum_{i=1}^{5}x_i^2-6\sum_{i=1}^{5}x_i+\sum_{i=1}^{5}9\right]$$

$$=\frac{1}{5}\left[55-6\times15+5\times9\right]$$

$$=\frac{1}{5}\left[55-6\times15+5\times9\right]$$

$$=\frac{1}{5}\left[55-90+45\right]=11-18+9=2$$

$$7)\frac{1}{5}\sum_{i=1}^{5}(x_i-3)^3=\frac{1}{5}\sum_{i=1}^{5}(x_i^3-9x_i^2+27x_i-27)$$

$$=\frac{1}{5}\left[\sum_{i=1}^{5}x_i^3-9\sum_{i=1}^{5}x_i^2+27\sum_{i=1}^{5}x_i-\sum_{i=1}^{5}27\right]$$

$$=\frac{1}{5}\left[(1^3+4^3+2^3+3^3+5^3)-9\times 55+(27\times 15)-(5\times 27)\right]$$

$$=\frac{1}{5}\left[255-495+405-135\right]=\frac{0}{5}=0$$

$$8)\frac{1}{5}\sum_{i=1}^{5}|x_i-3|=\frac{1}{5}\left[|x_1-3|+|x_2-3|+|x_3-3|+|x_4-3|+|x_5-3|\right]$$

$$=\frac{1}{5}\left[|1-3|+|4-3|+|2-3|+|3-3|+|5-3|\right]$$

$$=\frac{1}{5}(2+1+1+0+2)=\frac{6}{5}$$

<u>مثال 2</u>:أنشر العبارات التالية:

$$1)\sum_{i=1}^{5}(x_i-a) \qquad 2)\sum_{i=1}^{5}(x_i-a)^2 \qquad 3)\sum_{i=1}^{5}(a-2x_i)^2 \qquad 4)\sum_{i=1}^{5}a(x_i-b)^2$$

$$5)\sum_{i=1}^{4}(2x_i-3)^2 \qquad 6)\sum_{i=1}^{4}(x_i-a)^3 \qquad 7)\frac{1}{n}\sum_{i=1}^{n}(x_i-a) \qquad 8)\frac{1}{n}\sum_{i=1}^{n}(x_i-a)^2 \qquad 9)\frac{1}{n}\sum_{i=1}^{n}(\frac{x_i-a}{b})$$

<u>الحل:</u>

$$1)\sum_{i=1}^{5}(x_i-a)=\sum_{i=1}^{5}x_i-\sum_{i=1}^{5}a=\sum_{i=1}^{5}x_i-5.a$$

$$2)\sum_{i=1}^{5}(x_i-a)^2=\sum_{i=1}^{5}(x_i^{\,2}-2ax_i+a^2)$$

$$=\sum_{i=1}^{5}x_i^{\,2}-2a\sum_{i=1}^{5}x_i+5a^2$$

$$3)\sum_{i=1}^{4}(a-2x_i)^2=\sum_{i=1}^{4}(a^2-4ax_i+4x_i^2)$$

$$=\sum_{i=1}^{4}a^2-4a\sum_{i=1}^{4}x_i+4\sum_{i=1}^{4}x_i^2$$

$$=4a^2-4a\sum_{i=1}^{4}x_i+4\sum_{i=1}^{4}x_i^2$$

$$4)\sum_{i=1}^{5}a(x_i-b)\ \ =\sum_{i=1}^{5}(ax_i-ab)$$

$$=\sum_{i=1}^{5}ax_i-\sum_{i=1}^{5}a.b=a.\sum_{i=1}^{5}x_i-5a.b$$

$$5)\sum_{i=1}^{4}(2x_i-3)^2=\sum_{i=1}^{4}(4x_i^{\,2}-12x_i+9)$$

$$=\sum_{i=1}^{4}4x_i^{\,2}-\sum_{i=1}^{4}12x_i+\sum_{i=1}^{4}9$$

$$=4\sum_{i=1}^{4}x_i^{\,2}-12\sum_{i=1}^{4}x_i+36$$

$$6)\sum_{i=1}^{4}(x_i-a)^3=\sum_{i=1}^{4}(x_i^{\,3}-3ax_i^{\,2}+3x_ia^2-a^3)$$

$$=\sum_{i=1}^{4}x_i^{\,3}-3a\sum_{i=1}^{4}x_i^{\,2}+3a^2\sum_{i=1}^{4}x_i\ \ -4a^3$$

$$7)\frac{1}{n}\sum_{i=1}^{n}(x_i-a)=\frac{1}{n}\left[\sum_{i=1}^{n}x_i-\sum_{i=1}^{n}a\right]$$

$$=\frac{1}{n}\left[\sum_{i=1}^{n}x_i-n.a\right]=\frac{1}{n}\sum_{i=1}^{n}x_i-a$$

$$8)\frac{1}{n}\sum_{i=1}^{n}(x_i-a)^2=\frac{1}{n}\sum_{i=1}^{n}(x_i^{\,2}-2a.x_i+a^2)$$

$$=\frac{1}{n}(\sum_{i=1}^{n}x_i^2-2a.\sum_{i=1}^{n}x_i+\sum_{i=1}^{n}a^2)$$

$$=\frac{1}{n}\sum_{i=1}^{n}x_i^2-\frac{2a}{n}\sum_{i=1}^{n}x_i+\frac{\cancel{n}.a^2}{\cancel{n}}$$

$$=\frac{1}{n}\sum_{i=1}^{n}x_i^2-\frac{2a}{n}\sum_{i=1}^{n}x_i+a^2$$

$$9)\ \frac{1}{n}\sum_{i=1}^{n}(\frac{x_i-a}{b})=\frac{1}{n}.\frac{1}{b}.\sum_{i=1}^{n}(x_i-a)$$

$$=\frac{1}{n.b}\left[\sum_{i=1}^{n}x_i-\sum_{i=1}^{n}a\right]=\frac{1}{n.b}\left[\sum_{i=1}^{n}x_i-n.a\right]$$

$$=\frac{1}{n.b}\sum_{i=1}^{n}x_i-\frac{1}{\cancel{n}.b}\cancel{n}.a=\frac{1}{n.b}\sum_{i=1}^{n}x_i-\frac{a}{b}$$

__مثال3:__ استعمل الرمز $\sum$ لكتابة ما يلي:

$1)k_1+k_2+k_3+.......+k_9$

$2)F_1X_1+F_2X_2+F_3X_3+.......+F_LX_L$

$3)B_1X^1+B_2X^2+B_3X^3+.......+B_nX^n$

$4)F_1(X_1-a)^2+F_2(X_2-a)^2+.......+F_j(X_j-a)^2$

$5)\frac{CF_1(X_1+b)^2}{L}+\frac{CF_2(X_2+b)^2}{L}+..........+\frac{CF_n(X_n+b)^2}{L}$

$6)1+4+9+16+25$

$7)-1+4-27+256-3125$

__الحل:__

$$1)\ k_1+k_2+k_3+.......+k_9=\sum_{i=1}^{9}k_i$$

$$2)\ F_1X_1+F_2X_2+F_3X_3+.......+F_LX_L=\sum_{i=1}^{L}F_iX_i$$

$$3)\ B_1X^1+B_2X^2+B_3X^3+.......+B_nX^n=\sum_{i=1}^{n}B_iX^i$$

$$4)\ F_1(X_1-a)^2+F_2(X_2-a)^2+.......+F_j(X_j-a)^2=\sum_{i=1}^{j}F_i(X_i-a)^2$$

$$5)\ \frac{CF_1(X_1+b)^2}{L}+\frac{CF_2(X_2+b)^2}{L}+.........+\frac{CF_n(X_n+b)^2}{L}=\frac{C}{L}\sum_{i=1}^{n}F_i.(x_i+b)^2$$

$$6)\ 1+4+9+16+25=(1)^2+(2)^2+(3)^2+(4)^2+(5)^2=\sum_{i=1}^{5}i^2$$

$$7)\ -1+4-27+256-3125=(-1)^1+(-2)^2+(-3)^3+(-4)^4+(-5)^5$$

$$=\sum_{i=1}^{5}(-i)^i$$

دليل عملية الضرب (π)

ضرب **x** متغيرة يكتب كالتالي:

$$\boldsymbol{X_1}\times \mathbf{X_2}\times \mathbf{X_3}\times \ldots \times \mathbf{X_N}=\prod_{\mathbf{i=1}}^{\mathbf{N}}\mathbf{X_i}$$

الخواص :

$$\prod_{i=1}^{n}a=a^{n} \qquad\qquad \prod_{i=1}^{n}axi=a^{n}\prod_{i=1}^{n}xi$$

قائمة المراجع

-**جاك لوكايون و كريستيان لابروس :**"الإحصاء الوصفي"، ديوان المطبوعات الجامعية الجزائر 1985

-**محمد صبحي أبو صالح و عدنان محمد عوض :**"مقدمة في الإحصاء"، دار المسيرة عمان 2007.

-**عز الدين جوني :** "نظرية الإحصاء"، ديوان المطبوعات الجامعية الجزائر 1984.

-**عبد العزيز فهمي هيكل :**"طرق التحليل الإحصائي"، دار النهضة العربية بيروت.

- **عبد الجبار توفيق البياتي :**"الإحصاء و تطبيقاته"، إثراء للنشر و التوزيع، عمان 2008.

- **عبد الناصر رويسات :**"الإحصاء الوصفي و مدخل الاحتمالات دروس و تمارين"، ديوان المطبوعات الجامعية المطبعة الجهوية بوهران 2006.

- **جلاطو جيلالي:** "الإحصاء مع تمارين ومسائل محلولة "، ديوان المطبوعات الجامعية - الجزائر- 1998.

- **حامد نور الدين:"** دروس في مقياس الإحصاء الوصفي"، دار اليازوري العلمية للنشر والتوزيع – عمان - 2016.

- مواقع الأنترنت: قوقل، يوتيوب: كتابة دروس، محاضرات، تمارين وامتحانات في الإحصاء الوصفي.

- AIT HABOUCHE-MIHOUB O., « **Statistique et mathématique en économie et en gestion** » , Dar El Adib, 1989.
- BAILLARGEON G., « **Probabilité, statistique descriptive et technique de régression** » , SMG, 1989.
- BARTHE René., « **La Statistique Descriptive en 10 leçons : Méthode progressive "ABCD"** » , Edition Economica, 1989.
- CALOT Gérard, « **Cours de statistique descriptive »**, édition Dunod, Paris, 1973.

- CHAUVAT Gérard, REAU Jean-Philippe, « **Statistiques descriptives** », ARMAND COLIN, 2002.
- DAGNELIE P., « **Théorie et méthodes statistiques 2 tomes** », Presses agronomiques Gembloux 1985.
- DROESBEKE Jean-Jacques, « **Eléments de Statistique** », OPU, 1988 by Editions de l'Université de Bruxelles (Belgique).
- GIARD V., « **Statistique appliquée à la gestion** » , Economica, 1987.
- GOLDFARB Bernard, PARDOUX Catherine, **« Introduction à la Méthode Statistique, Exercices Corrigés** », DUNOD, 2003.
- GOLDFARB Bernard, PARDOUX Catherine, « **Introduction à la Méthode Statistique** », DUNOD, 2003.
- GRAIS B., « **Statistique descriptive** », Edition Dunod, 1995.
- GRAIS B., « **Exercices corrigés de statistique descriptive avec rappel de cours** », Dunod, 1991.
- GRAIS B., « **Statistique pour économistes** », Economica, 1987.
- GUITTON Henri, « **Statistique** «, Edition Dalloz,1976.
- Léonard G Kazmier, « **Statistiques de la gestion** », McGRAW-HILL Québec 1982.
- Khaldi Khaled, « **Méthodes statistiques** », OPU Constantine 2005.
- LABROUSSE Christian, Statistique, « **Exercices corrigés avec rappel de cours** », Dunod, 292 p., Paris, 1980.
- LETHIELLEUX Maurice, « **Statistique descriptive** », Dunod, 2003.
- MASIERI Walder, « Méthodes **Quantitatives** », 2ème Dunod, Paris 1990.

- MASIERI Walder, « **Statistique et calcul des probabilités** », Dalloz, Paris 2001.

- MONINO Jean-Louis, KOSIANSKI Jean-Michel, LE CORNU François, « **Statistiques descriptives - Travaux dirigés** », DUNOD, 2000.

- MURRAY R. Spiegel, « **Statistique** », Série Schaum, 2ème édition, traduction de la 2ème américaine, France, 1996.

- Murray R. Spiégel, « **Théorie et application de la statistique** », McGRAW-HILL Québec 1985.

- P. Pacé et R Cluzel, « **Statistique et probabilité** », Ed Delagrave Paris 1982.

- PY Bernard, « **Exercices corrigés de Statistique Descriptive**», ECONOMICA, 2ème édition, Paris, 1994.

- PY Bernard, « **Statistique Descriptive** », ECONOMICA, 4ème édition, Paris, 2001.

- TENENHAUS Michel, **« Statistique : Méthodes pour décrire, expliquer et prévoir** », DUNOD, 2006.

- SALVAORE Dominick, « **Econométrie et statistique appliquée, Cours et problème** », Série Schaum, 297 p., Paris, 1985.

- *TUFFERY Stéphane*, ***« Data Mining et statistique décisionnelle TECHNIP 2012 ».*** Un panorama très complet du *data mining* avec quelques rappels de statistiques. Plus axé sur l'opérationnel que sur les démonstrations.

- *LEBART L., PIRON M et MORINEAU A.*, « ***Statistique exploratoire multidimensionnelle*** », ***DUNOD 2006.***Un incontournable de niveau master.

- *MELARD Guy*, « ***Méthodes de prévision à court terme*** », ***ELLIPSES 2007.*** Un ouvrage particulièrement clair et opérationnel qui explore à fond certaines techniques. Avec CD-ROM. Mais il n'est plus édité...

- *SAPORTA Gilbert*, ***« Probabilités, analyse des données et statistiques*** » ***TECHNIP 2011.*** La référence, tant pour les statistiques que pour l'analyse de données. Bon niveau en maths exigé !

- *TRIBOUT Brigitte*, ***« Statistiques pour économistes et gestionnaires*** », ***PEARSON 2013.*** Initiation très claire qui ne s'adresse pas qu'aux économistes...

- *DEHON Catherine, DROESBEKE Jean-Jacques et VERMANDELE Catherine*, ***« Eléments de statistiques*** », ***SMA 2008.*** Une copieuse introduction aux statistiques qui aborde aussi l'analyse de données. Lecture facile.

- GEORGIN Jean-Pierre, GOUET Michel***, « Statistiques avec Excel, descriptives, tests paramétriques et non paramétriques à partir de la version Excel 2000*** », ***PUR 2005.*** Je dois beaucoup à cet ouvrage qui détaille des tests pour certains peu connus. Avec CD-ROM.

- Webographie et Sites internet en arabe et en langues étrangères : Google, YouTube, … : Ecrire cours, exercices et examens de statistique descriptive.

الفهرس

الفهرس