جامعة 8 ماي 1945 ـ قالمة

كلية العلوم الاقتصادية والتجارية وعلوم التسيير

قسم العلوم التجارية

مطبوعة بيداغوجية في مقياس

الإحصاء 1

المستوى : سنة اولى جذع مشترك

السداسي : الأول

من اعداد

د. بيري نورة

السنة الجامعية 2016-2017

تقديم

يعتبر الإحصاء من العلوم التي تحتل مكانة هامة في الحياة الأكاديمية والعملية، لأنه يتعامل مع الحاجات اليومية للإنسان على اختلافها، بأساليب متنوعة تتدرج في سلم متصاعد بدءا من البساطة، والتي لا تتطلب سوى معلومات أولية في المفردات الإحصائية، وانتهاء بالأساليب المتطورة في علم الإحصاء، والتي تحتاج إلى فهم متعمق في النظرية الإحصائية. من منطلق الأهمية هذه، ورغبة منا في إثراء مكتبة الجامعة بمثل هذه المواضيع، وتعبيرا عن المعرفة التي تراكمت لدي خلال سنين طويلة من الدراسة والتدريس، وددت أن أضع بين يدي طلبتنا الأعزاء هذه المطبوعة، لتكون لهم مرجع يساهم في اكتسابهم لهذا المقياس.

لقد اعتمدنا في إعداد محتويات هذه المطبوعة على البرنامج المقرر من طرف اللجنة البيداغوجية الوطنية لميدان التكوين في العلوم الاقتصادية والتجارية وعلوم التسيير سنة **2014**. ولهذا فقد جاءت المطبوعة ضمن ستة فصول. الفصل الأول خصص إلى شرح المفاهيم الإحصائية، والفصل الثاني خصص إلى عرض البيانات جدوليا أما الفصل الثالث فخصص إلى العرض البياني .الفصل الرابع خصص إلى مقاييس النزعة المركزية، والخامس إلى مقاييس التشتت، بينما خصص الفصل السادس إلى مقاييس الشكل.

د. بيري نورة

قالمة في **26** سبتمبر **2016**

قائمة المحتويات

الفصل الأول: المفاهيم الإحصائية

1 – 1 تمهيد

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

1 – 2 تعريف علم الإحصاء وأقسامه

1 – 2 – 1 تعريف علم الإحصاء

يختص علم الإحصاء بجمع وعرض وتحليل واستخدام البيانات الرقمية لعمل استدلالات واتخاذ قرارات في ظل عدم التأكد في مجالات الاقتصاد والأعمال وغيرها من العلوم الاجتماعية و الطبيعية.

1 – 2 – 3 أقسام علم الإحصاء

ينقسم إلى الإحصاء الوصفي والإحصاء الاستدلالي. يختص الإحصاء الوصفي بجمع وتلخيص وتوصيف مجموعة من البيانات. بينما يختص الإحصاء الاستدلالي بالوصول إلى تعميم عن خواص الكل (المجتمع) من واقع فحص جزء من هذا الكل (العينة).

1 – 3 منهجية البحث الإحصائي

لإجراء أي عمل إحصائي نتبع المراحل المنهجية التالية:

1 – 3 – 1 تحديد الظاهرة المدروسة

عند القيام بدراسة ظاهرة ما، أول ما يجب القيام به، هو:

– تحدد مشكلة الدراسة بكل دقة ووضوح وأهداف الدراسة الإحصائية، والإطار المكاني والزماني المناسب لجمع هذه البيانات حول الظاهرة المراد دراستها، وكذا تحديد وحدات القياس،

– تحديد الوحدة الإحصائية، المجتمع الإحصائي والعينة، الصفة الإحصائية وطبيعتها.

1 – 3 – 2 جمع البيانات

تعود عملية جمع البيانات الإحصائية من أساسيات العمل الإحصائي، ولهذه المرحلة أهمية خاصة في أي دراسة إحصائية، إذ أن توفر البيانات الدقيقة والسليمة عن الظاهرة المدروسة، يعطي نتائج سليمة ويساعد على اتخاذ قرار سليم بناء على تلك النتائج، لذا يجب على الباحث تحديد مصادر، أساليب وطرق جمع البيانات المرغوب فيها، وذلك قبل بدء العمل.

1 – 3 – 2 – 1 مصادر جمع البيانات

هناك عدة مصادر للحصول على البيانات، تختلف باختلاف موضوع الدراسة والغرض منها، ويمكن تصنيف مصادر البيانات إلى صنفين هما:

أولا: المصادر غير المباشرة:

وتضم:

– **المصادر التاريخية (الرسمية):** يمكن جمع البيانات عن مفردات المجتمع من السجلات الرسمية المحفوظة نتيجة تدوينها في سجلات رسمية طبقا للقوانين والأنظمة، مثال ذلك: سجل الميلاد، سجل الضرائب، سجل المنازعات،...الخ.

– **المصادر الداخلية:** يقصد بالمصادر الداخلية مجموعة البيانات التي تجمعها الهيئات والمؤسسات بصفة مستمرة لأغراض التنظيم والإدارة داخلها. وذلك بشكل دوري ومستمر، وهي إما وطنية (بيانات ينشرها البنك المركزي، المديرية العامة للضرائب، الديوان الوطني للإحصاء،...الخ) أو دولية (بيانات ينشرها صندوق النقد الدولي، صندوق النقد العربي، مؤتمر الأمم المتحدة للتجارة والتنمية،...الخ).

ثانيا: المصادر المباشرة (الأولية، الميدانية):

يقصد بالمصادر المباشرة استقصاء البيانات من مصادرها الأصلية، وذلك بأن يقوم الباحث بجمعها عن طريق:

– **المقابلة:** تتم عن طريق التقاء الباحث بأفراد العينة والتحدث إليهم بخصوص الظاهرة المراد دراستها، ويقوم الباحث بتدوين المعلومات المستقاة من هذا الحوار لاستخدمها فيما بعد،

- **المقابلة عن طريق الوسائط الحديثة:** وتتم هذه العملية عن طريق اتصال الباحث بأفراد العينة عن طريق الوسائط الحديثة (هاتف، انترنت،...) والتحدث إليهم وجمع المعلومات المراد دراستها.

- **الملاحظة:** هي المشاهدة والمراقبة الدقيقة لسلوك ما أو ظاهرة معينة في ظل ظروف وعوامل بيئية معينة بغرض الحصول على معلومات دقيقة لتشخيص هذا السلوك أو هذه الظاهرة.

- **الاستبيان:** يعتبر من أهم طرق جمع البيانات وأكثرها انتشارا، وهو عبارة عن مجموعة من الأسئلة معدة مسبقا حول الظاهرة المدروسة، حيث يقدم لأفراد العينة بغرض الإجابة عليه، وتعتبر هذه الإجابات كبيانات يستخدمها الباحث لدراسته.

1 – 3 – 2 – 2 أساليب جمع البيانات من المصادر المباشرة

هناك أسلوبان لجمع البيانات الإحصائية هما:

أولا: **الحصر الشامل:**

تقوم هذه الطريقة على جمع البيانات الإحصائية لكل وحدات المجتمع الإحصائي ومن مزايا هذا الأسلوب أنه يعطينا صورة كاملة عن المجتمع الإحصائي، كما يتميز بالدقة المطلوبة إذا ما توفرت شروط البحث الإحصائي، غير أن هذه الطريقة صعبة التنفيذ، كما تحتاج إلى تكاليف كبيرة، ووقت طويل للجمع والفرز، وهذا في أغلب الدراسات الإحصائية،

وخاصة عندما يكون المجتمع كبير جدا. ومن أبرز أمثلة الحصر الشامل نجد: التعداد السكاني، إجمالي الناتج المحلي،...الخ.

ثانيا: العينة

تقوم هذه الطريقة على أخذ جزء من المجتمع الإحصائي يكون ممثل لهذا المجتمع. بحيث يتم دراسة الظاهرة بواسطة العينة وإسقاط النتائج على المجتمع ككل. وتنقسم العينة إلى:

- **العينة العشوائية:** نقوم بسحب عناصر العينة العشوائية بطريقة تخضع للصدفة فقد. ويوجد 4 طرق رئيسية لاختيار العينة العشوائية وهي:

* **العينة العشوائية البسيطة:** هي العينة التي تعطى فيها لجميع أفراد المجتمع فرصة متساوية في الظهور، وتستخدم عادة في المجتمعات التي تكون عناصرها متجانسة.

* **العينة الطبقية:** تستخدم في حالة كون المجتمعات غير متجانسة، بحيث يقسم المجتمع إلى طبقات ويختار بطريقة عشوائية عينة بسيطة من كل طبقة، بحيث تكون كل الطبقات ممثلة بنسبة معينة في العينة الطبقية المسحوبة.

* **العينة العنقودية:** تستخدم في حالة كون المجتمعات كبيرة جدا، حيث يتم تقسيم المجتمع الإحصائي إلى مجموعات جزئية، ثم نقسم كل مجموعة جزئية إلى مجموعات جزئية، وهكذا يكون لدينا شكل هرمي للمجتمع الإحصائي. بعد ذلك، نقوم بسحب عينة عشوائية بسيطة من كل مجموعة جزئية مكونة، بحيث تكون كل هذه المجموعات الجزئية ممثلة في العينة العنقودية بنسب معينة.

* **العينة المنتظمة:** يكون اختيار عناصر العينة عن طريق ترتيب عناصر المجتمع وفق نظام معين، من ثم يتم اختيار ترتيب عنصر ما بطريقة عشوائية، وبعدها نظيف عددا معين لهذا الترتيب ونأخذ العنصر الموالي وهكذا، حتى يتم أخذ جميع عناصر العينة المرغوبة.

– **العينة الغير عشوائية:** يتم اختيار مفرداتها بطريقة غير عشوائية، أي لا تخضع إلى الصدفة، ومن أمثلتها:

* **العينة الحصصية:** وهي عينة طبقية غير عشوائية، حيث يتم تقسيم المجتمع إلى طبقات حسب خصائص معينة، ثم أخذ عينات غير عشوائية من كل طبقة من هذه الطبقات.

* العينة المتاحة: يقوم الباحث بأخذ المفردات المتاحة أمامه بصورة مباشرة في تكوين العينة.

* **العينة العمدية:** حيث يقوم الباحث باختيار أفراد العينة طبقا لأهداف محددة تتطلبها الدراسة.

* **عينة الكرة الثلجية:** يقوم الباحث باختيار فرد في العينة، وعن طريق هذا الفرد يختار الفرد الثاني وهكذا.

1 – 3 – 3 ترتيب، تصنيف وعرض البيانات

بعد جمع البيانات الإحصائية سواء كانت كمية أن كيفية، نجدها في الغالب بشكل خام، أي تفتقر إلى الترتيب والتصنيف، لكون عملية التحصيل تتم بصورة عشوائية في الغالب. وهذا ما يجعل البيانات لا تعطي نظرة وافية عن الظاهرة المدروسة. لذا وجدت طرق إحصائية وصفية يلجأ إليها الباحثون للترتيب وتصنيف البيانات حتى تصبح عملية، سواء بإعطائنا صورة مباشرة عن الظاهرة المراد دراستها أو تصبح قابلة للمعالجة. ويمكن لهذه البيانات المرتبة والمصنفة أن يستفيد منها الباحثون الآخرون، خصوصا إذا

نشرت في دوريات خاصة أو عامة. وترتيب وتصنيف وعرض البيانات يجري حسب طبيعة البيانات ذاتها، وسنتناول في الفصل الثاني والثالث من هذه المطبوعة بالتفصيل لهذه المرحلة.

1 – 3 – 4 تحليل البيانات واستخلاص النتائج

تحليل البيانات هو وسيلة الحصول على الإجابات المطلوبة في إشكالية البحث الإحصائي، إذ تعتبر المراحل المشار إليها آنفا ضرورية، حتى يتمكن الباحث من التحليل الإحصائي لجوانب الظاهرة المدروسة. ويتم ذلك عن طريق أدوات إحصائية كثيرة منها البسيط ومنها المعقد، تسمح كلها باستقراء معلومات واستخلاص نتائج الذي هو هدف البحث الإحصائي.

1 – 4 تعريف بعض المصطلحات الإحصائية

المجتمع: هو مجموعة عناصر تشترك في خاصية أو أكثر قد يكون منته أو غير منته.

العينة: هي جزء من المجتمع.

مثال: نرغب في استخلاص نتائج حول أطوال 10000 طالب من خلال فحص 100 طالب فقط، في هذه الحالة يتكون المجتمع من 10000 طالب، بينما تتكون العينة من 100 طالب مسحوب من هذا المجتمع.

مثال 2: نرغب في استخلاص نتائج تخص عدد الزبائن الوافدين إلى محل. يتشكل المجتمع من كل الزبائن، بينما تتشكل العينة من خلال فحص أول 30 زبون مثلا.

نلاحظ أن المجتمع في المثال 1 منته أما في المثال 2 فهو غير منته.

الوحدة الإحصائية:

تعرف الوحدة الإحصائية بأنها "كل عنصر من المجتمع أو العينة". من هذا التعريف نجد أن الإحصائية هي العنصر الأولي التي تجري عليه الدراسة الإحصائية، وتصنف عادة الوحدات الإحصائية كما يلي:

أ – **وحدات عد:** ويطلق عليها وحدات أو أشياء فردية وهي تتعلق بكائنات أو أشياء مستقلة، وهي على نوعين:

- **وحدات طبيعية:** مثل الإنسان، النباتات،...

- **وحدات منتجة:** مثل الكرسي، الكومبيوتر،...

ب – وحدات قياس: وهي وحدات تتعلق بالنواحي التي لا يمكن عدها ويمكن قياسها، وهي على نوعين:

- **وحدات القياس المادية:** وهي تقيس الكميات المادية وقد تكون:
  - **بسيطة:** عندما تتكون من وحدة واحدة مثل: الطن، الكلغ،...
  - **مركبة:** عندما تكون مكونة من مزيجين لوحدتين بسيطتين مثل، كلم/سا، ن/م،...
- **وحدات القياس النقدية:** وهي وحدات تقيس التبادل السلعي من سلع وخدمات مثل، الدينار، الدولار،...

الصفة الإحصائية:

تخص الصفة الوحدة الإحصائية، وهي الخاصية أو الميزة التي تقوم عليها الدراسة. وتنقسم إلى نوعين:

أ – **صفة كيفية (نوعية):** وهي تلك الصفة التي لا يمكن أن تقاس بأرقام مثل اللون، الجنس، الجنسية،...وتنقسم إلى نوعين:

- صفة كيفية ترتيبية: وهي صفة كيفية قابلة للترتيب. ومن أمثلتها:
  - المستوى التعليمي (ابتدائي، متوسط، ثانوي،...)
  - رتبة أستاذ في الجامعة (معيد، أستاذ مساعد ب، أستاذ مساعد أ،...)
  - رتبة عسكرية (جندي، عريف، عريف أول،...)

ب – **صفة كمية (كمية):** وهي تلك الصفة التي يمكن قياسها بأرقام. مثل الطول، الوزن، الحجم،.... وتنقسم إلى:

- **صفة كمية منفصلة (متقطعة، وثابة):** ويطلق عليها أيضا متغير منفصل، وهو المتغير الذي يمكن أن يأخذ فقط قيما محدودة و متميزة. ومن أمثلتها:
  - عدد العمال بمؤسسة
  - عدد المصانع في دولة
  - عدد الطلقات على هدف

- **صفة كمية متصلة (مستمرة):** ويطلق عليها أيضا متغير متصل، وهو المتغير الذي يمكن أن يأخذ عدد لا نهائيا من القيم داخل أي مجال معلوم. ومن أمثلتها:
  - طول شخص
  - مساحة بلد
  - فترة اشتغال مصباح

1 – 5 تمارين تطبيقية حول الفصل الأول

تمرين رقم (1–1).

في الأمثلة التالية حدد كل من المجتمع، الوحدة الإحصائية، الصفة الإحصائية، طبيعة الصفة الإحصائية.

1– عدد النساء العاملات.

2– شكل التكوين في مركز للتكوين المهني.

3– موظفي المؤسسات التي تشتغل في جهات اقتصادية.

4– طبيعة قطاعات النشاط في اقتصاد ما.

5– ترتيب الفنادق في القطاع السياحي الوطني.

6– فصيلة دم شخص.

7– مساحة الأراضي زراعية مستغلة.

الحل:

نعرض في الجدول الموالي إجابة التمرين رقم (1–1)

الرقم	المجتمع	الوحدة الإحصائية	الصفة الإحصائية	طبيعة الصفة الإحصائية
1	النساء العاملات	امرأة	العمل (الوظيفة)	كيفية
2	التكوينات	تكوين	شكل التكوين	كيفية
3	الجهات الاقتصادية	جهة	عدد المؤسسات	كمية منفصلة
4	قطاعات النشاط	قطاع	طبيعة القطاع	كيفية
5	الفنادق	فندق	ترتيب	كيفية
6	الأشخاص	شخص	فصيلة الدم	كيفية
7	الأراضي الزراعية المستغلة	أرض زراعية مستغلة	المساحة	كمية متصلة

تمرين رقم (1−2).

قام أحد الباحثين بإجراء دراسة لاستقصاء أراء 50 شركة من بين 300 شركة أجنبية تنشط في قطاع الصيدلة بدولة ما. وهذا لمعرفة آراؤهم حول ملائمة الضريبة المفروضة عليهم.

المطلوب:

حدد كل من المجتمع الإحصائي، العينة الإحصائية، الوحدة الإحصائية، الصفة الإحصائية، طبيعة الصفة الإحصائية.

الحل:

المجتمع الإحصائي: 300 شركة أجنبية

العينة الإحصائية: 50 شركة أجنبية

الوحدة الإحصائية: شركة أجنبية

الصفة الإحصائية: درجة الملائمة (ملائم ، غير ملائم)

طبيعة الصفة الإحصائية: كيفية

الفصل الثاني: عرض البيانات جدوليا

2 – 1 تمهيد

تعد عملية عرض البيانات الإحصائية من الأعمال الضرورية حتى تصبح البيانات عملية، ويمكن استخدام أدوات التحليل الإحصائي لتفسيرها واستخلاص النتائج منها، وسوف نتطرق في هذا الفصل إلى عرض البيانات عن طريق الجداول الإحصائية.

2 – 2 تعريف الجدول الإحصائي ومواصفاته

2 – 2 – 1 تعريف الجدول الإحصائي

الجدول الإحصائي (Tableau Statistique) عبارة عن ترتيب منظم للبيانات الإحصائية، في أسطر وأعمدة بقصد إبراز أهمية تلك البيانات وتسهيل المقارنة، يشكل هذا الجدول حلقة وصل بين قام بإعداده ومن سيتطلع عليه في المستقبل لأخذ المعلومات التي يريدها، لذا نجده في الغالب يفسر نفسه بنفسه، إذا ما صمم بطريقة يكون كفيلا فيها بشرح محتواه دون تكرار أو طول.

2 – 2 – 2 مواصفات الجدول الإحصائي

إن كل جدول إحصائي يتطلب مجموعة من المواصفات نذكر منها:

- يجب أن يكون الجدول مصمما بشكل مقبول، ويحمل كل عناصر الظاهرة المدروسة،
- يجب أن يكون لكل جدول عنوان واضح ومختصر قدر الإمكان يوضح محتواه،
- يجب أن يذكر مصدر البيانات الواردة في الجدول،
- يجب أن نعرض مجاميع الأعمدة أو الأسطر إن كان عرضها مفيد،
- يجب إبراز وحدة القياس المعتمدة في الأعلى إن كانت البيانات متجانسة، أو في الأعمدة أو الأسطر إن كانت البيانات غير متجانسة،
- في حالة استخدام أكثر من جدول يجب أن إعطائها رقم خاص ليسهل الإشارة إليه دون الحاجة إلى ذكر عنوانه، وكذا يسهل الرجوع إليه عند الضرورة. ويكون ترقيم الجداول ترقيم متتالي من حيث الترتيب.

وترتب البيانات في داخل الجداول الإحصائية بعدة أساليب أهمها:

- الترتيب التاريخي: حيث يتم وضع البيانات في الجدول حسب تسلسلها التاريخي،

– الترتيب الكمي: عندما يكون عامل الكم هو النقطة الهامة هو النقطة الهامة المرغوبة تأكيدها، فإن ترتيب البيانات بحسب الكم هو الأسلوب الأكثر فاعلية في تكوين الجدول.

2 – 3 أنواع الجداول الإحصائية

يمكن إبراز الجداول بصورة عامة إلى نوعين رئيسيين هما:

2 – 3 – 1 الجداول العادية

يمكن أن تكون الجداول العادية بسيطة أو مركبة.

2 – 3 – 1 – 1 الجداول البسيطة

هي الجداول التي تصف ظاهرة واحدة فقط، وبالتالي تحتوي على صفة واحدة، وكمثال على ذلك نأخذ الجداول التالية:

مثال رقم (2–1).

الجنسية	السياح
أمريكية	12
اسبانية	25
تونسية	32
مصرية	24
المجموع	93

مستوى تعليمي	موظف
متوسط	12
بكالوريا	25
ليسانس	20
ماجستير	1
المجموع	58

الخطوط الهاتفية	المستعملين
0	2
1	25
2	50
3	10
المجموع	87

الخبرة المهنية	العمال
أقل من 5	2
5–10	15
10–20	15
20 فأكثر	11
المجموع	43

2 – 3 – 1 – 2 الجداول المركبة

هي الجداول التي تصف ظاهرتين أو أكثر، وبالتالي تحتوي على صفتين أو أكثر.

مثال رقم (2–2).

المستوى التعليمي	الجنس		المجموع
تخصص	ذكر	أنثى	
آداب	213	287	500
علوم	384	296	680
رياضيات	58	92	150
المجموع	655	675	1330

أما الجدول التالي فيحتوي على 3 صفات إحصائية:

مثال رقم (2-3).

المستوى التعليمي	سنة أولى		سنة ثانية		سنة ثالثة		المجموع
تخصص	ذكر	أنثى	ذكر	أنثى	ذكر	أنثى	
آداب	70	110	85	75	58	102	500
علوم	65	178	127	100	192	18	680
رياضيات	40	27	13	30	5	35	150
المجموع	175	315	225	205	255	155	1330

2 – 3 – 2 الجداول التكرارية

الجدول المنتظم على صورة قيم (أو فئات) بحيث يقابل كل قيمة (أو فئة) تكرارها يسمى بالجدول التكراري. وتقسم الجداول التكرارية إلى ثلاثة أنواع هي:

2 – 3 – 2 – 1 الجداول التكرارية البسيطة

أولا: حالة الصفة الكيفية

تحتوي على صفة نوعية واحدة فقط، ويتم إفراغ البيانات فيها كما يلي:

- يرسم مسودة جدول يناسب حجم البيانات مقسم إلى ثلاثة أعمدة،
- يوضح في العمود الأول الصفات، والعمود الثاني التعداد، وفي العمود الثالث التكرار،
- وفي النهاية نقوم بحذف العمود الثاني من جدول المسودة لنحصل على الجدول المطلوب.

مثال رقم (2-4).

لدينا بيانات لعينة من طلبة كلية العلوم الاقتصادية والتجارية وعلوم التسيير بجامعة ما، تتكون من 30 طالب حسب الشعب التي ينتمون إليها كالتالي:

ع. اقتصادية	ع. تسيير	ع. اقتصادية	ع. تسيير	ع. تجارية	ع. تجارية
ع. تسيير	ع. تسيير	ع. تسيير	ع. اقتصادية	ع. تسيير	ع. اقتصادية
ع. تجارية	ع. اقتصادية	ع. تجارية	ع. تجارية	ع. اقتصادية	ع. تسيير
ع. اقتصادية	ع. تسيير	ع. تسيير	ع. تسيير	ع. تسيير	ع. تسيير
ع. تجارية	ع. تجارية	ع. اقتصادية	ع. تجارية	ع. تجارية	ع. تجارية

المطلوب: وضع هذه البيانات في جدول موضحا تكرار كل فئة.

الحل:

بعد تحديد عدد الشعب المتوفرة في الكلية، نقوم بأخذ شعبة الطلبة شعبة تلو أخرى، ووضع تشطيبه عمودية صغيرة أمام الصفة (الشعبة) التي ينتمي إليها الطالب في عمود التعداد، وكل ما أصبح عدد هذه التشطيبات خمسة نكتبه على هيئة زمرة (حزمة)، أي عندما نصل إلى ألتشطيبه الخامسة نضعها مقاطعة للأربع تشطيبات الأولى، ونستمر هكذا حتى ننتهي من تسجيل كل الإجابات. إن استخدام هذه الرموز الخماسية (الزمر أو الحزم) على هذا المنوال، يسهل لنا عملية عدها، وتسمى هذه العملية بعملية التفريغ، حيث يتم تفريغ البيانات في الجدول كالتالي:

الشعبة	التعداد	التكرار
ع. اقتصادية	𝍸 𝍷𝍷𝍷	8
ع. تجارية	𝍸 𝍸	10
ع. تسيير	𝍸 𝍸 𝍷𝍷	12
المجموع		30

يعتبر الجدول السابق جدولا أوليا، أي جدول مسودة، وعند حذف العمود الثاني نحصل على الجدول النهائي المطلوب كما يلي:

الشعبة	التكرار
ع. اقتصادية	8
ع. تجارية	10
ع. تسيير	12
المجموع	30

ثانيا: حالة الصفة الكمية

وهي الجداول التي نقوم فيها بتصنيف ظاهرة تقاس كميا، أو بمعنى آخر، هي الجداول التي تعرض تكرارات كميات الظواهر كتوزيع عدد من الأشخاص حسب أعمارهم، أو وزنهم، أو أجورهم. وقد تكون هذه الجداول إما متصلة أو منفصلة، وفي كلتا الحالتين، فإنه عند إفراغ البيانات الخام في هذه الجداول، لا بد من مراعاة ترتيب القيم تنازليا، أو تصاعديا، واستخدام قاعدة العد الأساسية وهي التشطيبات لتفادي الأخطاء، خاصة إن كان عدد القيام المراد جدولتها كبيرا.

– الجداول التكرارية للصفة الكمية المنفصلة

وهي الجداول التي تظهر عدد تكرارات كمية واحدة محددة وممثلة في رقم واحد فقط، أي عبارة عن متغيرة إحصائية منفصلة. ويتكون الجدول من عمودين، حيث يمثل العمود الأول الصفة، بينما الثاني يمثل التكرار المطلق.

مثال رقم (2–5):

اختيرت عينة عشوائية من إحدى البلدان تتكون من 100 عائلة، ليتم استقصاؤهم عن عدد أطفالهم فكانت إجاباتهم كالتالي:

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

المطلوب: وضع هذه البيانات في صورة جدول تكراري

الحل:

– نحدد أولا العدد الممكن لأطفال في العائلات وهو: 0، 1، 2، 3، 4، 5، 6، 7، ثم ندرج الجدول التكراري كما يلي

عدد الأطفال	التعداد	التكرار
0	𝍸 𝍸	10
1	𝍸 𝍸 ||	12
2	𝍸 𝍸 𝍸 |	16
3	𝍸 𝍸 𝍸 |||	18
4	𝍸 𝍸 𝍸	17
5	𝍸 𝍸 ||||	14
6	𝍸 |||	8
7	𝍸	5
المجموع		100

يعتبر الجدول السابق جدول مسودة، وعندما نحذف العمود رقم 2 نحصل على الجدول التكراري بالصورة التالية:

عدد الأطفال	التكرار
0	10
1	12
2	16
3	18
4	17
5	14
6	8
7	5
المجموع	100

- الجداول التكرارية للصفة الكمية المتصلة

وهي الجداول التي تكون فيها الظاهرة محصورة في مجال محدد، بحيث يمكن أن تأخذ أي قيمة ضمن هذا المجال. يتم استخدام هذه الجداول في عرض البيانات إذا كان عددها كبيرا، وذلك لتلخيصها، إذ أن هدف التبويب هو عرض البيانات بأقل حيز ممكن، وبأقصى وضوح، حيث يتم تقسيم مجموعة البيانات التي لدينا، إلى مجموعات جزئية حيث تشمل كل مجموعة جزئية عددا من القيم المتقاربة، ويطلق على المجموعة الجزئية مصطلح فئة (Classe). بعد ذلك، نحصر العدد الذي يقع داخل حدود كل فئة ويسمى تكرار (Fréquence) أو تكرار مطلق (Fréquence absolue)، وبذلك تتحول البيانات الأصلية إلى عدد من الفئات وتكراراتها المقابلة. ولا يصح أن يكون عدد الفئات صغير بحيث يقع في كل فئة عدد كبير من القيم قد تتباين أرقامها، ولا يبلغ عدد الفئات قدرا كبيرا كذلك مما يجعل الجدول التكراري قريب الوضع من المجموعة الأصلية، وأخيرا لا يصح أن يترتب على تحديد الفئات أن تتجمع المفردات في عدد منها، بينما تخلو فئات أخرى من التكرارات.

ونعرف فيما يلي بعض المصطلحات المتعلقة بالجداول التكرارية المتصلة.

مدى الفئة (Amplitude de la classe): تعني المدى العددي الواقع بين حدي الفئة الأدنى والأعلى، أي الفرق بينهما، ويكتب بالصورة التالية:

$$a = C_{max} - C_{min}$$

مركز الفئة (Centre de classe): هو عبارة عن القيمة الواقعة في منتصف المسافة بين الحد الأعلى للفئة والحد الأدنى للفئة، ويكتب بالصورة التالية:

$$C_i = \frac{\boldsymbol{C_{max}} + \boldsymbol{C_{min}}}{2}$$

هذه العلاقة توضح لنا أنه بالإمكان إيجاد حدي الفئة الأدنى والأعلى إذا علم مركز الفئة ومداها وهذا باستخدام الصيغ التالية:

$$\boldsymbol{C_{min}} = C_i - \frac{\boldsymbol{a}}{2}$$

$$\boldsymbol{C_{max}} = C_i + \frac{\boldsymbol{a}}{2}$$

المدى العام (Etendue): هو الفرق بين أكبر قيمة في البيانات ناقص أقل قيمة فيها، أي:

$$E = \boldsymbol{x_{max}} - \boldsymbol{x_{min}}$$

تحديد عدد الفئات:

لوضع البيانات في الجدول التكراري يجب علينا أولا تحديد عدد الفئات التي يمكن ان نرتب البيانات على أساسها، وتوجد عدة طرق لتحديد عدد الفئات نذكر منها:

طريقة Sturges: تستخدم معادلة الإحصائي Sturges لتحديد عدد الفئات بالعلاقة التالية:

$$\text{عدد الفئات} = 1 + 1.332(log(n))$$

بحيث n تمثل عدد المفردات في العينة.

طريقة Yule: تستخدم معادلة الإحصائي Yule لتحديد عدد الفئات بالعلاقة التالية:

$$\text{عدد الفئات} = \sqrt[4]{n}$$

ملاحظة:

- لا توجد قاعدة تلزم الباحث لإتباع طريقة معينة في تحديد عدد الفئات، والأمر متروك للتقدير الشخصي لكل باحث في تحديد هذه الفئات بشرط احترام الخصائص المطلوبة في هذا الصنف من الجداول التكرارية.

- يتم دائما تقريب عدد الفئات للعد الأعلى.

تحديد طول الفئة:

بعد تحديد عدد الفئات، يمكن لنا أن نقوم بتحديد طول كل فئة، وذلك من خلال استخدام العلاقة التالية:

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

مثال رقم (2-6):

أوزان عينة تتكون من 40 شخص بالكلغ كما يلي:

60	47	70	65	67	62	67	60	45	50
49	63	58	54	53	48	47	51	50	70
50	50	57	49	47	67	57	64	61	60
53	51	60	49	47	52	46	56	58	50

من خلال إطلاعنا على البيانات المعروضة نلاحظ أنه لا يمكن استيعاب الأرقام بسهولة، لذلك نلجأ للتبسيط عن طريق إنشاء جدول توزيع تكراري بإتباع الخطوات سالفة الذكر على النحو التالي:

حساب المدى العام:

نلاحظ أن أقل وزن للأشخاص هو 45 كلغ، وأكبر وزن هو 70 كلغ، بالتالي مدى الفئة يساوي:

$$E = x_{max} - x_{min} = 70 - 45 = 25kg$$

حساب عدد الفئات:

يمكن تحديد عدد الفئات بإحدى المعادلتين التاليتين:

معادلة Sturges

$$\text{عدد الفئات} = 1 + 1.332\big(log(40)\big) = 6.32 \cong 7$$

معادلة Yule

عدد الفئات: هو $\sqrt[4]{40} = 6.29 \cong 7$

حساب طول الفئة:

$$\text{طول الفئة} = \frac{\text{المدى العام}}{\text{عدد الفئات}} = \frac{25}{7} = 3.57 \cong 4$$

نلاحظ من خلال الحسابات أن أوزان الأشخاص تتوزع على 7 فئات طول كل منها 4 كلغ كالتالي:

الفئة الأولى: تشمل الأوزان من 45 كلغ إلى أقل من 49 كلغ

الفئة: تشمل الأوزان من 49 كلغ إلى أقل من 53 كلغ

الفئة الثالثة: تشمل الأوزان من 53 كلغ إلى أقل من 57 كلغ

الفئة الرابعة: تشمل الأوزان من 57 كلغ إلى أقل من 61 كلغ

الفئة الخامسة: تشمل الأوزان من 61 كلغ إلى أقل من 65 كلغ

الفئة السادسة: تشمل الأوزان من 65 كلغ إلى أقل من 69 كلغ

الفئة السابعة: تشمل الأوزان من 69 كلغ إلى أقل من 73 كلغ

ويمكن تفريغ البيانات في الجدول التكراري التالي.

التكرار (Fi)	الفئات
7	$[45-49[$
11	$[49-53[$
4	$[53-57[$
8	$[57-61[$
4	$[61-65[$
4	$[65-69[$
2	$[69-73[$
40	$\sum$

ملاحظة: يفضل أن يكون عدد الفئات محصورا بين 5−15.

2 – 3 – 2 – 2 جداول التوزيع التكراري النسبي (Fr) و النسبي المئوي (Fr%)

وهو جدول يوضح الأهمية النسبية لكل قيمة من قيم الظاهرة محل الدراسة، مقارنة بإجمالي عدد القيم، ونستخدم القانونين التاليين لحسابهما:

$$Fr = \frac{Fi}{\sum Fi} \qquad Fr_{(\%)} = \frac{Fi}{\sum Fi} \times 100 = Fr \times 100$$

مثال رقم (2–7).

بأخذ المثال السابق، نقوم بحساب التوزيع التكراري النسبي والنسبي المئوي في الجدول التالي.

الفئات	التكرار (Fi)	التكرار النسبي (Fr)	التكرار النسبي المئوي (Fr%)
[45 – 49[	7	0.175	17.5
[49 – 53[	11	0.275	27.5
[53 – 57[	4	0.1	10
[57 – 61[	8	0.2	20
[61 – 65[	4	0.1	10
[65 – 69[	4	0.1	10
[69 – 73[	2	0.05	5
$\sum$	40	1	100

ملاحظة: يجب أن يكون دائما مجموع التكرار النسبي مساويا 1 ومجموع التكرار النسبي المئوي مساويا 100%.

2 – 3 – 2 – 3 جدول التوزيع التكراري المتجمع الصاعد (Fcc) والمتجمع النازل (Fcd)

يستخدم التوزيع التكراري المتجمع الصاعد (Fréquence cumulée croissante)، والتوزيع التكرار المتجمع النازل (Fréquence cumulée décroissante)، لمعرفة عدد العناصر التي تقل عن مستوى معين (أقل من) أو تزيد عن قيمة معينة (أكبر من) على الترتيب، وفي هذه الحالات نستخدم التوزيعات التكرارية المتجمعة وفق الصيغتين التاليتين:

$$Fcc = \begin{cases} 0 \\ F_1 \\ F_1 + F_2 \\ F_1 + F_2 + F_3 \\ \dots\dots\dots\dots \\ \sum Fi \end{cases} \qquad Fcd = \begin{cases} \sum Fi \\ \sum Fi - F_1 \\ \sum Fi - F_1 - F_2 \\ \dots\dots\dots\dots \\ 0 \end{cases}$$

مثال رقم (2−8):

بأخذ المثال السابق، نقوم بحساب التوزيع التكراري المتجمع الصاعد والمتجمع النازل في الجدول التالي.

الفئات	Fi	الحدود	Fcc	الحدود	Fcd
[45 − 49[	7	أقل من 49	7	45 فأكثر	40
[49 − 53[	11	أقل من 53	18	49 فأكثر	33
[53 − 57[	4	أقل من 57	22	53 فأكثر	22
[57 − 61[	8	أقل من 61	30	57 فأكثر	14
[61 − 65[	4	أقل من 65	34	61 فأكثر	10
[65 − 69[	4	أقل من 69	38	65 فأكثر	6
[69 − 73[	2	أقل من 73	40	69 فأكثر	2
$\sum$	40				

يتضح لنا من الجدول السابق أن عدد الأشخاص الذين وزنهم أقل من 65 كلغ هو 34 شخص، بينما عدد الأشخاص الذي وزنهم مثلا 61 كلغ فأكبر هو 10 أشخاص.

2 – 4 أصناف الجداول التكرارية

يمكن أن يصادف الباحث عدد لا نهائي من البيانات، كل منها يصلح لدارسة ظاهرة ما أو عدة ظواهر. لهذا، توجد عدة أصناف للجدول التكرارية بحيث يستجيب كل منها لطبيعة بيانات محددة.

2 – 4 – 1 الجداول المقفلة والمفتوحة

يمكن أن تكون الجداول التكرارية مغلقة أو مفتوحة من طرف أو من الطرفين كما يلي:

الفئات
أقل من C_{min1}
$C_{min1} - C_{max1}$
$C_{min2} - C_{max2}$
$C_{min3} - C_{max3}$
مفتوح من طرفه الأدنى

الفئات
$C_{min1} - C_{max1}$
$C_{min2} - C_{max2}$
$C_{min3} - C_{max3}$
C_{max3} فأكثر
مفتوح من طرفه الأعلى

الفئات
أقل من C_{min1}
$C_{min2} - C_{max2}$
$C_{min3} - C_{max3}$
C_{max3} فأكثر
مفتوح من الطرفين

الفئات
$C_{min1} - C_{max1}$
$C_{min2} - C_{max2}$
$C_{min3} - C_{max3}$
$C_{min4} - C_{max4}$
مقفل

مثال رقم (2-9).

يظهر الجدول التكراري التالي بيانات أجور شهرية لـ 435 بالأورو لمؤسسة X.

رقم الفئة	الفئات	طول الفئة a	مركز الفئة Ci	Fr
I	$[800 - 1000[$	200	900	10
II	$[1000 - 1100[$	100	1050	90
III	$[1100 - 1400[$	300	1250	220
IV	$[1400 - 1650[$	150	1525	60
V	$[1650 - 1900[$	250	1775	40
VI	$[1900 - 2000[$	100	1950	10
VII	2000 فأكثر	غير محدد	-	5
	$\sum$			435

نلاحظ من خلال الجدول السابق أن الفئة السابعة غير محدودة من الأعلى، ولحساب مركز الفئة الأخيرة يعتمد الباحث على المنطق ومعارفه السابقة حول الظاهرة المدروسة. مثلا يمكننا أن نقدر أن الحد الأعلى للفئة الأخيرة يقدر بـ 2500 أورو، بالتالي فإن مركز الفئة الأخيرة يكون 2250 أورو. كما يمكن للباحث في حال البيانات أن يختار مثلا أكبر قيمة في المشاهدات ويعتبرها هي حد الفئة الأعلى.

كذلك نلاحظ في نفس الجدول أن الفئة الثالثة تم اختيارها بشكل سيء، حيث تحتوي على عدد كبير من المشاهدات مقارنة بالفئات الأخرى، وهذا ما يجعل بعض خصائص أن تكون غامضة أثناء معالجتها، ولأجل الحصول على أكثر معلومات يفضل أن تقسم الفئة الرابعة إلى 3 فئات كل منها بطول 100 أورو.

2 – 4 – 2 الجداول المنتظمة وغير المنتظمة

يعرف الجدول المنتظم بأنه الجدول الذي تكون فيه أطوال الفئات متساوية، أما الجدول غير المنتظم فهو الجدول الذي تكون فيه أطوال الفئات غير متساوية.

مثال رقم (2–10).

لديك بيانات خام حول وزن 47 شخص كالتالي:

وزن	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80
تكرار	1	0	0	0	1	0	1	0	0	3	1	3	3	3	4	2	3	4	0	4
وزن	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
تكرار	3	0	2	3	0	1	1	1	0	0	1	0	0	0	0	0	0	1	1	0

المطلوب: إعادة تجميع البيانات في فئات غير متساوية مع حساب التكرار النسبي والتكرار النسبي المئوي والتجميعي الصاعد والتجميعي النازل.

الحل:

يمكن وضع المشاهدات السابقة في الجدول التكراري الموالي.

الفئات	Fi	Fr	Fr(%)	Fcc	Fcd	Frcc(%)	Frcd(%)
$[60-70[$	3	0.0638	6.38	3	47	6.38	100%
$[70-75[$	13	0.2766	27.66	16	44	34.04	93.62
$[75-80[$	13	0.2766	27.66	29	31	61.70	65.96
$[80-85[$	12	0.2553	25.53	41	18	87.23	38.30
$[85-100[$	6	0.1277	12.77	47	6	100%	12.77
$\sum$	47	1	100%	↑	↑	↑	↑
				أقل من $C_{max(i)}$	$C_{min(i)}$ فأكثر	أقل من $C_{max(i)}$	$C_{min(i)}$ فأكثر

من الجدول السابق نجد مثلا أن:

- 61.7% من المشاهدات (ما يمثل 29 مشاهدة) لهم وزن أقل من 80 كلغ.

2 – 4 – 3 الحدود الفعليـة (الحقيقيـة) للفئات

في بعض الحيان قد يصادف الباحث جداول تكرارية غير محددة فئاتها بواسطة مجالات تظهر بالضبط قيمة فئاتها العليا والدنيا. لهذا، يلجأ الباحث إلى تحديد ما يعرف بالحدود الفعلية للفئات قبل أن يقوم بعملية تفريغ البيانات.

مثال رقم (2–11).

لديك جدول تكراري يمثل أوزان 100 طالب كالتالي.

الفئات	Fi
60 – 62	5
63 – 65	18
66 – 68	42
69 – 71	27
72 – 74	8
$\sum$	100

الحدود الفعلية للفئات
59.5–62.5
62.5–65.5
65.5–68.5
68.5–71.5
71.5–74.5
$\sum$

تتضمن الفئة 60 – 62 من الناحية الحقيقية كل القياسات من 59.5000 إلى 62.5000 هذه الأرقم إذا عبرنا عنها بـ 59.5 إلى 62.5 تسمى بالحدود الفعلية للفئات. وبالتالي يمكن تحديد الحدود الفعلية للفئات في الجدول السابق.

2 – 5 تمارين تطبيقية حول الفصل الثاني

تمرين رقم (2–1)

تبرع لمصلحة الدم 20 شخص، وبعد الاختبارات تبين أن فصيلة دمهم كما يلي.

O	B	AB	B	A
B	AB	O	AB	A
O	O	AB	AB	B
B	O	O	B	A

المطلوب: ضع هذه البيانات في جدول تكراري

الحل:

فئة الدم	Fi
A	3
B	6
AB	5
O	6
المجموع	20

تمرين رقم (2–2)

يحتوي قسم دراسي على 30 طالب تحصلوا على النقاط التالية في مقياس الإحصاء 1.

10	12.5	5	3.5	14	6	3	5	12	13
9	4	5	7.5	8	12	0	10.5	9	8.5
11.5	5	7.5	6	12	10	13	16.5	2	9.5

المطلوب

1 – حدد المجتمع الإحصائي، العينة الإحصائية، الوحدة الإحصائية، الصفة الإحصائية، طبيعة الصفة الإحصائية؟

2– ضع هذه البيانات في جدول تكراري؟ مع حساب التكرار النسبي والتكرار النسبي المئوي؟

3 – أوجد التكرار المتجمع الصاعد، المتجمع النازل، المتجمع الصاعد النسبي، المتجمع النازل النسبي؟

الحل:

1 – هذه البيانات تمثل علامات 30 طالب لقسم في مقياس الإحصاء 1، وهي تتراوح بين 0 كأدنى نقطة و16.5 كأعلى نقطة.

المجتمع الإحصائي: الطلبة

العينة الإحصائية: 30 طالب

الوحدة الإحصائية: طالب

الصفة الإحصائية: علامة مقياس الإحصاء

طبيعة الصفة الإحصائية: كمي متصل

2 – في ظل غياب أي شرط محدد مسبقا لتكوين الجدول التكراري، فإنه يمكننا استخدام طريقة Sturges أو Yule لتحديد عدد الفئات. كما يمكننا تحديد عدد الفئات مباشرة واختيار 4 فئات حيث أن علامات الطلبة تتراوح بين 0–20 عادة. ويكون الجدول التكراري كما يلي.

الفئات	Fi	Fr	Fr(%)
$[0-5[$	5	0.1667	16.67
$[5-10[$	13	0.4333	43.33
$[10-15[$	11	0.3667	36.67
$[15-20[$	1	0.0333	3.33
$\sum$	30	1	100

3 –

الفئات	Fi	Fcc	Fcd	Frcc(%)	Frcd(%)
$[0-5[$	5	5	30	16.67	100
$[5-10[$	13	18	25	60	83.33
$[10-15[$	11	29	12	96.66	40
$[15-20[$	1	30	1	100%	12.77
$\sum$	30				
		أقل من $C_{max(i)}$	$C_{min(i)}$ فأكثر	أقل من $C_{max(i)}$	$C_{min(i)}$ فأكثر

تمرين رقم (2−3)

يوضح الجدول الموالي تكرارات أجور شهرية لـ 65 عامل بالأورو في شركة ما.

الفئات	Fi
$[50-60[$	8
$[60-70[$	10
$[70-80[$	16
$[80-90[$	14
$[90-100[$	10
$[100-110[$	5
$[110-120[$	2
$\sum$	65

المطلوب:

1 – حدد الحد الأعلى للفئة الرابعة؟ والحد الأدنى للفئة السادسة؟

2 – أحسب مركز الفئة الثالثة؟

3 – أحسب طول الفئة؟

4 – ما هو تكرار الفئة الثانية؟

5 – أوجد Fr، (%)Fr، Fcc، (%)Fcc، Fcd، (%)Fcd؟

6 – أوجد عدد ونسبة العاملين الذين يتقاضون دخل شهري أقل من 80 أورو؟

7 – أوجد عدد ونسبة العاملين الذين يتقاضون دخل شهري 100 أورو فأكثر؟

8 – عين 5 عمال جدد في الشركة وكانت مرتباتهم كالتالي:

85.5، 113، 132، 159، 179.9.

أوجد توزيع تكراري لأجور 70 عامل في هذه الشركة؟ ماذا تلاحظ؟ وما هي البدائل المقترحة؟

الحل:

1 – الأعلى للفئة الرابعة = 90 أورو.

الحد الأدنى للفئة السادسة = 100 أورو.

2 – مركز الفئة الثالثة = $\frac{70+80}{2}$ = 75 أورو.

3 – حساب طول الفئة

نلاحظ أن جميع الفئات طولها متساوي ويساوي 10 أورو.

4 – تكرار الفئة الثانية = 16 عامل.

5 – تكوين جدول تكراري يضم Fr، Fr(%)، Fcc، Fcc(%)، Fcd، Fcd(%).

الفئات	Fi	Fr	Fr(%)	Fcc	Fcd	Frcc(%)	Frcd(%)
[50 – 60[	8	0,1231	12,31	8	65	12.31	100
[60 – 70[	10	0,1538	15,38	18	57	27.69	87.69
[70 – 80[	16	0,2462	24,62	34	47	52.31	72.31
[80 – 90[	14	0,2154	21,54	48	31	73.85	47.69
[90 – 100[	10	0,1538	15,38	58	17	89.23	26.15
[100 – 110[	5	0,0769	7,692	63	7	96.92	10.77
[110 – 120[	2	0,0308	3,077	65	2	100%	3.077
Σ	65	1	100%				
				أقل من $C_{max(i)}$	$C_{min(i)}$ فأكثر	أقل من $C_{max(i)}$	$C_{min(i)}$ فأكثر

6 – عدد العاملين الذين يتقاضون دخل شهري أقل من 80 أورو هو 34 عامل أما نسبتهم فهي 52.31%.

7 – عدد العاملين الذين يتقاضون دخل شهري 100 أورو فأكثر هو 7 عمال ونسبتهم هي 10.77%.

8 – إيجاد التوزيع التكراري بعد تعيين 5 عمال جدد.

نقوم أولا بوضع فئات جديدة التي يجب أن تضم كل أجور العمل المضافين. ثم نضع تكراراتهم المقابلة.

الفئات	Fi
$[50-60[$	8
$[60-70[$	10
$[70-80[$	16
$[80-90[$	15
$[90-100[$	10
$[100-110[$	5
$[110-120[$	3
$[120-130[$	0
$[130-140[$	1
$[140-150[$	0
$[150-160[$	1
$[160-170[$	0
$[170-180[$	1
$\sum$	70

نلاحظ أننا احتفظنا بطول الفئة الذي قدره 10 أورو حيث تم إضافة 6 فئات جديدة لكي تشمل المشاهدات المضافة للعمال الخمسة الجدد. وكنتيجة لذلك ظهرت لنا فئات خالية من التكرارات تماما، وفئات تكراراتها ضعيفة جدا مقارنة بالفئات الأخرى.

يمكن لنا أن نضع جدول توزيع تكراري يكون مفتوح من أعلى كما يلي.

الفئات	Fi
$[50-60[$	8
$[60-70[$	10
$[70-80[$	16
$[80-90[$	15
$[90-100[$	10
$[100-110[$	5
$[110-120[$	3
120 فأكثر	3
$\sum$	70

هذا التمثيل يحمل عدة عيوب منها أنه لا يمكن إجراء بعض العمليات الحسابية، كذلك فإن عدم تحديد الفئة الأخيرة من الأعلى يمكن أن يفسر بأن هناك أجور كبيرة تدفعها المؤسسة قد تصل إلى أي قيمة، وهذا ما يتنافى والواقع.

يمكن أن نمثل الجدول التكراري بتغيير طول الفئة لتصبح 20 أورو بدلا من 10 أورو كما يلي.

الفئات	Fi
$[50-70[$	18
$[70-90[$	31
$[90-110[$	15
$[110-130[$	3
$[130-150[$	1
$[150-170[$	1
$[170-190[$	1
$\sum$	70

نلاحظ من هذا الجدول أن معظم البيانات تتركز في الفئات الثلاث الأولى، وبالتالي فقدان الكثير من المعلومات ضمن هذه الفئات، بينما الفئات العليا (من الفئة الرابعة حتى 7) مازالت دقيقة في وصف هيكل الأجور. ويمكن كذلك تكوين جدول تكراري غير متساوي كما يلي.

الفئات	Fi
$[50-60[$	8
$[60-70[$	10
$[70-80[$	16
$[80-90[$	15
$[90-100[$	10
$[100-120[$	8
$[120-180[$	3
$\sum$	70

نلاحظ أن طول الفئة السادسة يساوي 20 أورو، وطول الفئة السابعة يساوي 60 أورو، بينما الفئات الخمسة الأولى طولها 10 أورو. وأحد عيوب هذه الطريقة هو صعوبة بعض العمليات الحسابية التي ستجرى لاحقا، كذلك فإنه كلما زاد طول الفئة تزداد معه أخطاء التجميع وتصبح المعلومات أكثر غموضا.

الفصل الثالث: عرض البيانات بيانيا

3 – 1 تمهيد

العرض البياني نوع آخر لتمثيل البيانات الإحصائية، إذ يمكن من تحليلها ومراقبتها والتعليق عليها، ويمكن القارئ من أخذ صورة مرئية مباشرة عن المعلومات المعروضة. وبالتالي يعتبر العرض البياني أكثر وضوحا وأكثر تعبيرا من الجدول البياني.

وتختلف التماثيل البيانية باختلاف البيانات المحصل عليها، واختيار عرض بياني يرجع إلى طبيعة الصفة المدروسة وإلى مقتضيات الدراسة في حد ذاتها. وسوف نقسم عرض البيانات إلى قسمين:

3 – 2 عرض البيانات الكيفية بيانيا

من أهم التقنيات المستخدمة لعرض البيانات الكيفية نجد:

3 – 2 – 1 الصور

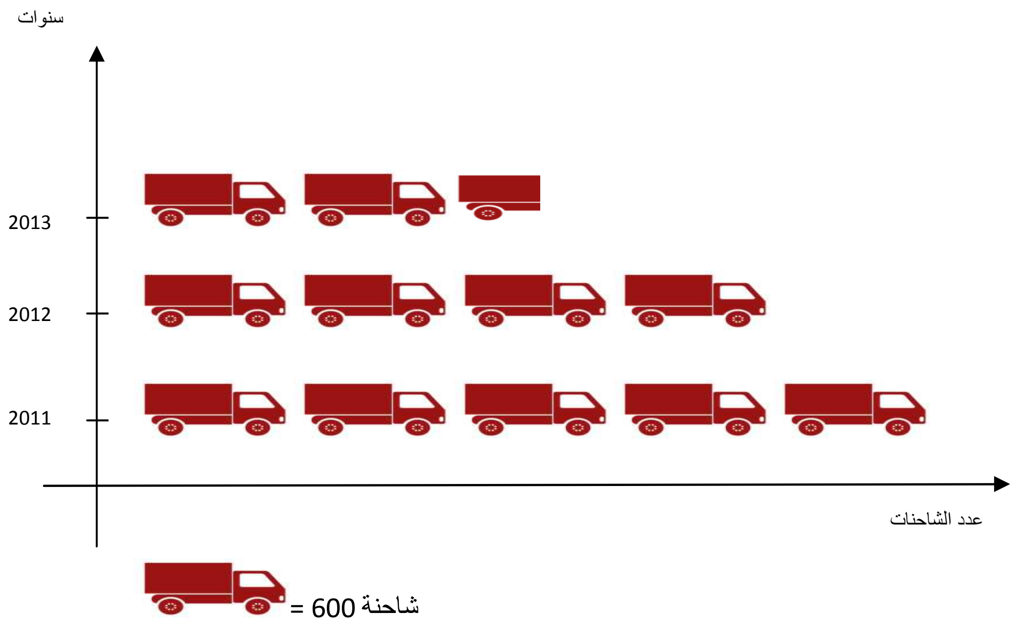
تستخدم الصورة المعبرة عن البيانات المراد عرضها كوسيلة إيضاحية جاذبة لانتباه القارئ، أو المشاهد لأنها تعطي انطباعا بصريا جيدا عن مجموعة البيانات المبحوثة. ويراعى في هذه الحالة عدة اعتبارات أساسية لاستخدام الصور منها:

- أن تستخدم صور مصغرة ومعبرة تلقائيا عن الظاهرة المراد عرضها،
- أن تكون صورة تقريبية وليست صورة دقيقة للمتغير،
- أن تستخدم الصور في عمل مقارنات فقط،
- أن تستخدم الصور لعرض تكرار معين ويعبر عن التغير في الكمية بتغيير عدد الصور المستخدمة وليس بتغيير حجمها.

ملاحظة: هذه الطريقة ليست دقيقة إذا أنجزت يدويا، لذلك يفضل عدم استخدامها إلا باستخدام الحاسوب.

مثال رقم (3-1):

لنفرض أننا نود إبراز إنتاج مصنع شاحنات نقل البضائع بمصنع سوناكوم بالرويبة الجزائر، بين عامي 2011 و2013 باعتماد الصور البيانية فإن ذلك يكون مثلا (افتراضيا) كما يعكسه الشكل الموالي.



نلاحظ من خلال هذا الشكل إن إنتاج المصنع في تناقص مستمر خلال الفترة 2011-2013.

3 - 2 - 2 الأعمدة (المستطيلات)

تتلخص هذه الطريقة في رسم أعمدة تتناسب أطوالها مع أرقام الظاهرة المعروضة، ووضع هذه الأعمدة بطريقة متقاربة يسهل عملية المقارنة بمجرد إلقاء النظر عليها. ويراعى وضع الأعمدة أن تكون متناسبة مع الظواهر المعروضة، بمعنى أنه إذا كان الرسم يمثل أرقام ظاهرة واحدة في عدد من السنين فإننا نرسم عمود واحد لكل سنة. أما إذا كان الرسم يمثل ظاهرتين أو ثلاثا من عند السنين، وجب أن يشمل الرسم عمودين أو ثلاثة لكل سنة مع التمييز بين الأعمدة عن طريق التظليل أو التلوين بنسق واحد لكل ظاهرة منعا من الالتباس، وإن كانت الظاهرة تتكون من جزأين أو أكثر أمكن تقسيم العمود مع ما يتناسب وأرقام الأجزاء مع التفرقة بين الأجزاء عن طريق التظليل أو التلوين.

وليس هناك ما يحول دون كتابة الأرقام على رأس العمود أو بداخله، هذا لتسهيل قراءة الشكل.

من أهم طرق عرض البيانات بواسطة الأعمدة نجد: الأعمدة البسيطة، المجزأة والمتلاصقة

3 – 2 – 2 – 1 الأعمدة البسيطة

تتلخص هذه الطريقة في وضع محورين رأسي وعمودي، حيث يمثل المحور الرأسي الظاهرة محل الدراسة، بينما يمثل المحور العمودي التكرارات المناظرة لكل قيمة من الظاهرة المدروسة، ونترك بين الأعمدة مسافات تكون متساوية.

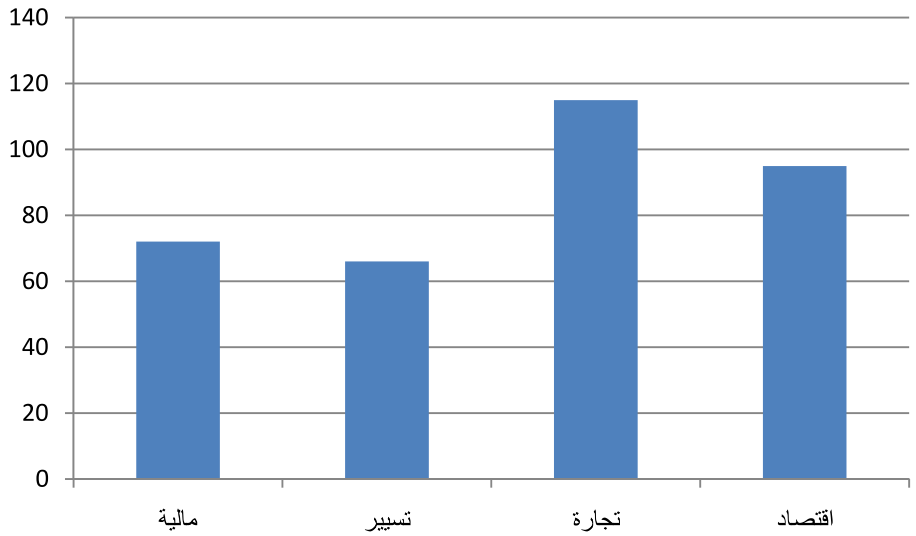
مثال رقم (3–2):

يوضح الجدول التالي توزيع الطلبة الناجحين إلى السنة الثانية بكلية العلوم الاقتصادية وعلوم التسيير وجنسهم لسنة 2010:

الجنس / الشعبة	مالية	تسيير	تجارة	اقتصاد
ذكور	20	32	51	38
إناث	52	34	64	57
المجموع	72	66	115	95

المطلوب:

عرض عدد الطلبة المنتقلين إلى السنة الثانية حسب الشعبة (أعمدة بسيطة).

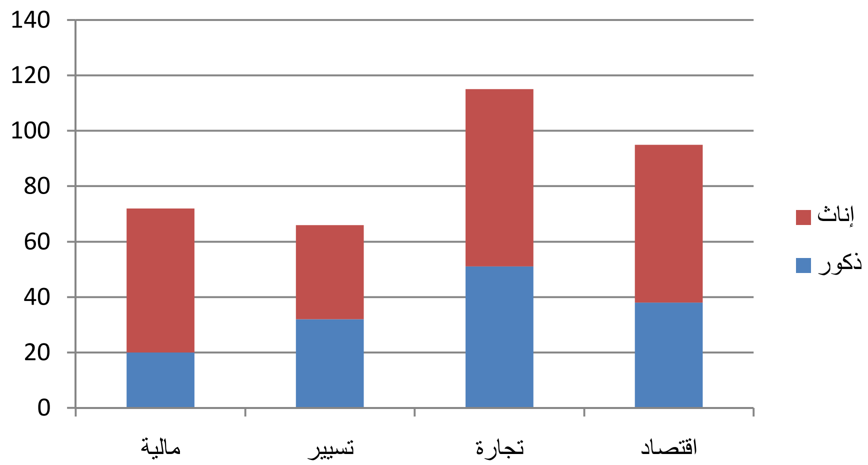


3 – 2 – 2 – 2 أعمدة مجزأة

هي عبارة عن أعمدة بيانية بسيطة، ولكن كل عمود منها يوضح قراءات أخرى لظاهرتين أو أكثر، حيث يكون كل جزأ في العمود متناسب مع تكراره. ويكون طول العمود كاملا ممثلا للمجموع الكلي للتكرارات التي تخص الصفة الممثلة.

مثال رقم (3-3):

بأخذ بيانات المثال رقم (3-2) نعرض الطلبة المنتقلين إلى السنة الثانية حسب الشعبة والجنس (أعمدة مجزأة بالقيم المطلقة)



ويمكن كذلك أن يتم عرض البيانات باستخدام النسب المئوية، بحيث يمثل طول كل الأعمدة الممثلة للصفات المعروضة 100%، ثم يقسم كل عمود إلى أجزاء، بحيث كل جزء يتناسب مع تكرار الجزء الممثل.

مثال رقم (3-4):

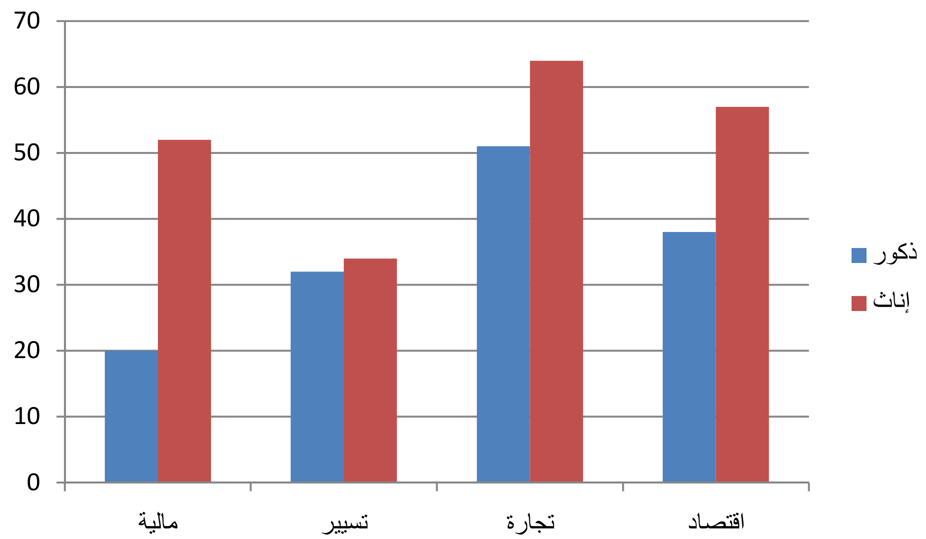
بأخذ بيانات المثال رقم (3-2) نعرض الطلبة المنتقلين إلى السنة الثانية حسب الشعبة والجنس (أعمدة مجزاة بالنسب المئوية).

3 – 2 – 2 – 3 أعمدة متلاصقة

تستخدم الأعمدة المتلاصقة عندما يكون لدينا ظواهر مركبة نريد تمثيلها، بحيث تسمح لنا بإجراء مقارنة بينها، ومقارنة التطور بينها، وهي عبارة عن مجموعة من الأعمدة تكون متلاصقة نميز بينها بالألوان أو التظليل.

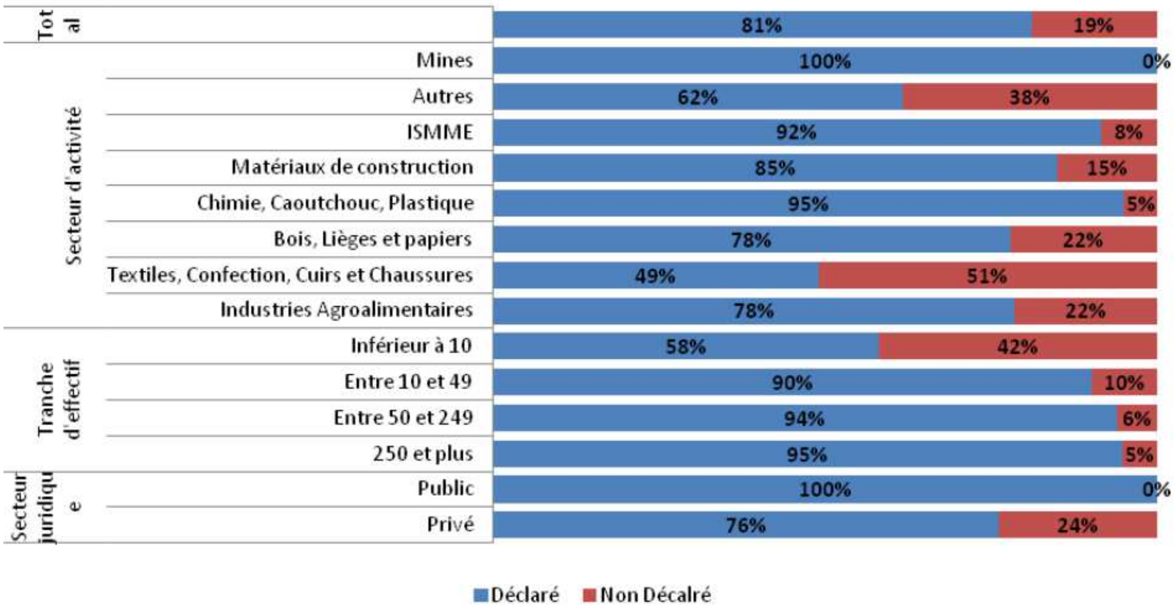
مثال رقم (3-5):

بأخذ بيانات المثال رقم (3-2) نعرض الطلبة المنتقلين إلى السنة الثانية حسب الشعبة والجنس (أعمدة متلاصقة)



ويمكن استخدام العديد من أشكال الأعمدة لتمثيل البيانات، وتكون مركبة أحيانا وتوفر البرمجيات (SPSS،Excel ، ...الخ) العديد منها نعرض إحداها في الشكل الموالي.

مثال رقم (3-6): شكل يوضح عدد العمال في سنة 2014 حسب اشتراكاتهم في الصندوق الوطني للتأمين على الأجراء (CNAS) في الجزائر خلال سنة 2014.



يظهر لنا الشكل السابق أن 81% من العمال مصرح بهم لدى CNAS سنة 2014 في الجزائر، كما يوضح الشكل أن 100% من عمال قطاع المناجم (Mines) مصرح بهم. يوضح الشكل كذلك توزيع العمل شرائح عدد الموظفين، حيث أن المؤسسات التي بها أكثر من 250 عامل (250 et plus) هي أكبر المؤسسات التي عمالها مشتركين في الصندوق، ونلاحظ كذلك أنه كلما زاد عدد العمال في المؤسسة (كلما كانت المؤسسة كبيرة) كلما زادت نسبة إشتراك عمالها في الصندوق. وأخيرا يظهر الشكل اشتراك العمال حسب الشكل القانوني (Secteur juridique)، حيث أن 100% من عمال المؤسسات العمومية (public) مشتركين في الصندوق.

3 – 2 – 2 الدوائر

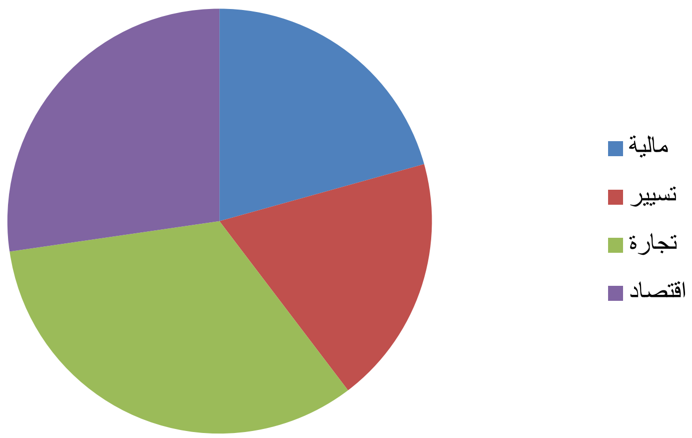
تعتبر الدوائر أحسن الأشكال وأبسطها خصوصا إذا كانت الظاهرة المعروضة مقسمة إلى أجزاء، ففي هذه الحالة يمكن رسم دائرة منقسمة إلى أجزاء تتجمع في مركز الدائرة، وتكون زوايا هذه الأجزاء متناسبة مع قيم أجزاء الظاهرة.

ويمكن الحصول على هذه الزوايا باستخدام العلاقة الثلاثية التالية:

$$\text{زاوية القطاع} = \frac{\text{تكرار الصفة}}{\text{مجموع التكرارات}} \times 360$$

مثال رقم (3−7):

بأخذ بيانات المثال رقم (3−2) نعرض شكل يمثل عرض الطلبة المنتقلين إلى السنة الثانية حسب الشعبة (دائرة).



ويمكن استخدام العديد من أشكال الدوائر لتمثيل البيانات، وتوفر البرمجيات (Excel، SPSS،...الخ) العديد منها نعرض إحداها في الشكل الموالي.

مثال رقم (3−8): شكل يوضح عدد المشاريع الاستثمارية المصرح بها في ولاية قالمة للفترة 2002−2015

يظهر لنا الشكل السابق أن 59% من عدد المشاريع المصرح بها في ولاية قالمة للفترة 2002-2015 يخص قطاع النقل، يليه قطاع البناء، الأشغال العمومية والسكن بنسبة 18%، وقطاع الصناعة بـ 15%، بينما القطاعات المتبقية وهي: الخدمات، الفلاحة، السياحة والصحة فمجموع عددها هو 8%.

3 – 3 عرض البيانات الكمية بيانيا

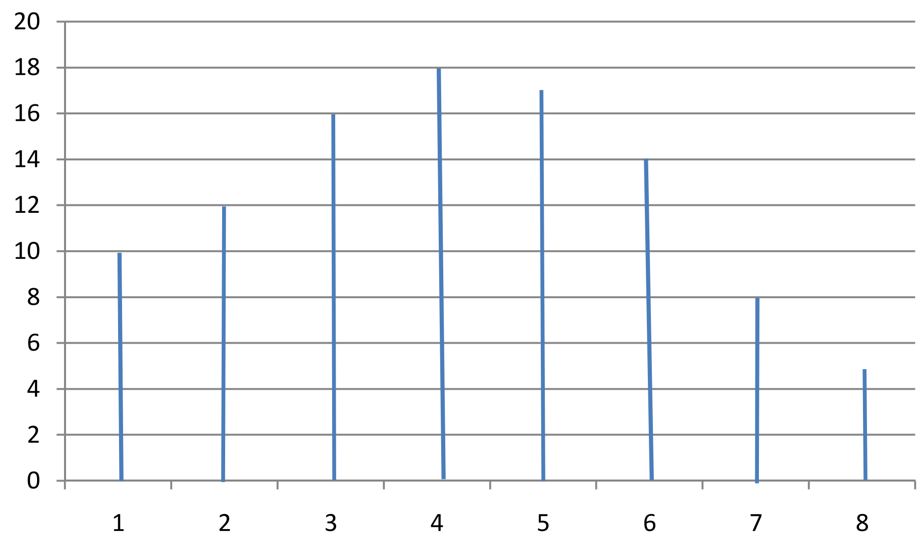
3 – 3 – 1 عرض بيانات المتغير الكمي المنفصل

3 – 3 – 1 – 1 مخطط الأعمدة البيانية

يعتمد هذا النوع على رسم معلم، حيث نضع على محور الفواصل المتغير، وعلى محور التراتيب نضع التكرارات المناسبة لقيم المتغير الإحصائي المنفصل، ثم نرسم خطوطا عمودية عند كل نقطة متناسبة مع تكرارها.

مثال رقم (3-9).

نعرض في الشكل الموالي بيانات المثال رقم (2-5) بواسطة أعمدة بيانية.



3 – 3 – 1 – 2 التكرار المتجمع الصاعد والنازل (والتكرارات المتجمع الصاعد والنازل النسبي)

يمكن تمثيل التكرار المتجمع الصاعد والنازل وكذا التكرار المتجمع الصاعد والنازل النسبي بواسطة شكل بياني، بحيث يكون هذا الشكل درجي وتكون درجاته أفقية.

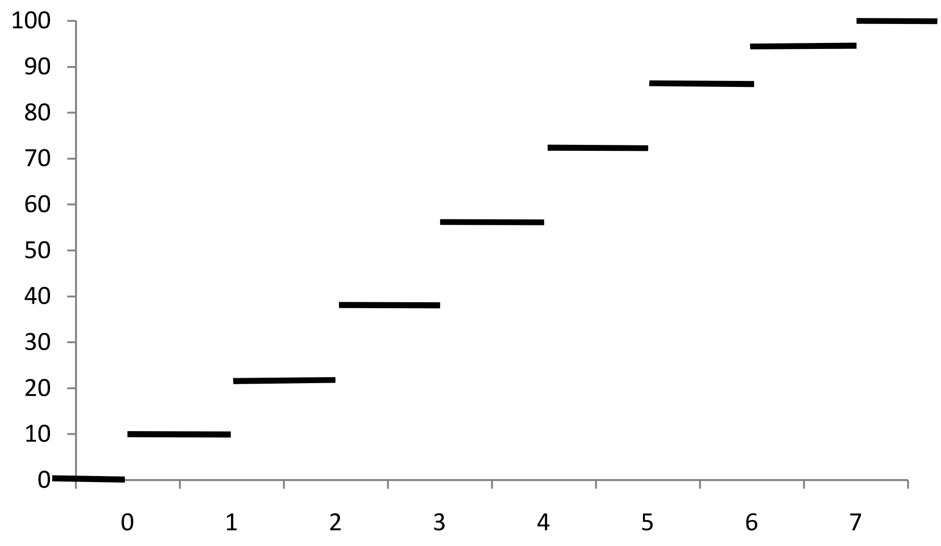
مثال رقم (3−10).

بأخذ بيانات المثال رقم (2−5) أوجد التكرار المتجمع الصاعد ومثله بيانيا.

الحل: نعرض في الجدول الموالي التكرار المتجمع الصاعد.

عدد الأطفال	Fi	Fcc
0	10	10
1	12	22
2	16	38
3	18	56
4	17	73
5	14	87
6	8	95
7	5	100
$\sum$	100	

وتمثيلها البياني نعرضه في الشكل الموالي.



3 – 3 – 2 عرض بيانات المتغير الكمي المتصل

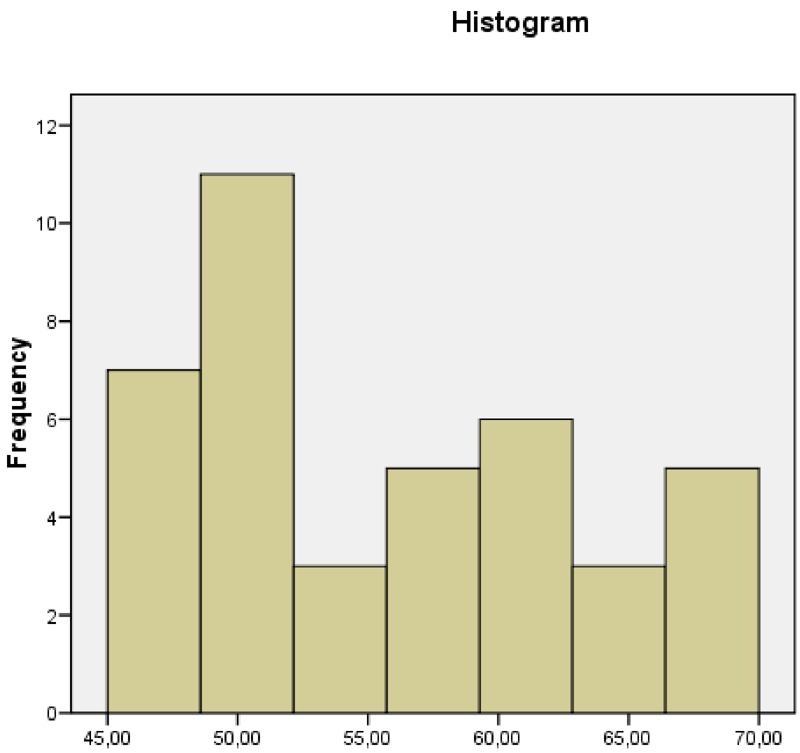
3 – 3 – 2 – 1 المدرج التكراري

المدرج التكراري (Histogramme) عبارة عن مجموعة من المستطيلات العمودية والمتلاصقة، كل مستطيل قاعدته على محور الفواصل تمثل طول الفئة، وارتفاعه يتناسب مع تكرار[1] الفئة على محور التراتيب. ونعرض البيانات بواسطة مدرج تكراري في حالتين: إما يكون طول الفئات متساوي، أو غير متساوي.

أولا: المدرج التكراري في حالة طول الفئات متساوي

مثال رقم (3–11).

كون المدرج التكراري لبيانات المثال رقم (2–6)



[1] - يمكن أن يكون المدرج التكراري ممثلا للتكرارات النسبية

ثانيا: المدرج التكراري في حالة طول الفئات غير متساوي

في حالة كون الفئات غير متساوية، فإنه لا يجوز أن نستخدم التكرارات الأصلية لتمثيل البيانات في مدرج تكراري. لذلك، نقوم باستبدال التكرارات الأصلية لتحل محلها التكرارات المعدلة والتي يتم الحصول عليها بالعلاقة التالية:

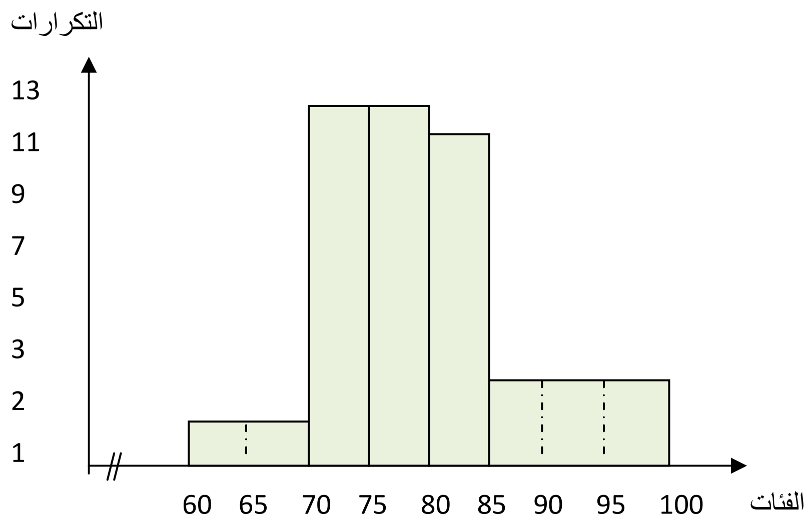
$$\text{التكرار المعدل} = \frac{\text{طول الفئة المنتظمة}}{\text{طول الفئة غير المنتظمة}} \times \text{التكرار الفعلي للفئة غير المنتظمة}$$

مثال رقم (3−12)

كون مضلع تكراري من المدرج التكراري في المثال رقم (3−11)

نقوم بحساب التكرار المعدل في الجدول الموالي

الفئات	Fi	ai	hi
[60 – 70[	3	10	1.5
[70 – 75[	13	5	13
[75 – 80[	13	5	13
[80 – 85[	12	5	12
[85 – 100[	6	15	2
$\sum$	47		



3 – 3 – 2 – 2 المضلع التكراري

المضلع التكراري (Polygone) هو عبارة عن خط بياني منكسر يصل بين النقط التي إحداثياتها مراكز الفئات والتكرارات، فعند توضيح الأرقام التي ترد بالجدول التكراري بمضلع. ويمتاز المضلع التكراري عن المدرج التكراري بأنه يسمح لنا بمقارنة أكثر من ظاهرتين عند رسمهما في نفس المعلم. ولتشكيل المضلع التكراري نبدأ برسم المحورين المتعامدين ويخصص المحور الرأسي للتكرارات، أما المحور الأفقي فترصد عليه مراكز الفئات، ونوضح نقاط تتناسب مع مراكز الفئات وتكراراتها المقابلة، لتحديد النقد النقط المكونة لمضلع التكرار بعد توصيلها ببعض، ونقوم بوصل حدي المضلع التكراري مع محور الفواصل.

كما يمكن أن نرسم المضلع التكراري من واقع المدرج التكراري، وذلك بأن نصل بين منصفات قسم المستطيلات، لأن منتصف القمة يقع عند مركز الفئة. وعلى غرار المدرج التكراري، فإن المضلع التكراري يمكن أن يمثل بواسطة جدول تكراري فئاته متساوية أو غير متساوية.

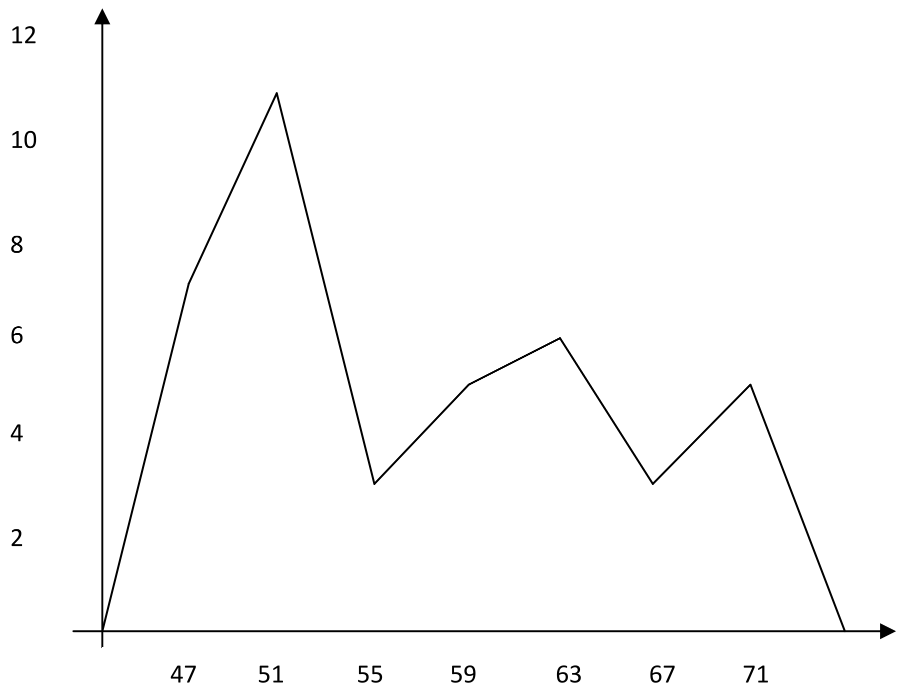
إن المساحة الواقعة تحت المضلع التكراري تساوي بالضبط المساحة الواقعة تحت المدرج التكراري (يمكن استنتاج هذه الخاصية باستخدام حسابات المثلثات)

أولا: المضلع التكراري في حالة طول الفئات متساوي

نقوم مباشرة بتمثيل المدرج التكراري.

مثال رقم (3–13):

كون مضلع تكراري من بيانات المثال رقم (2–6).

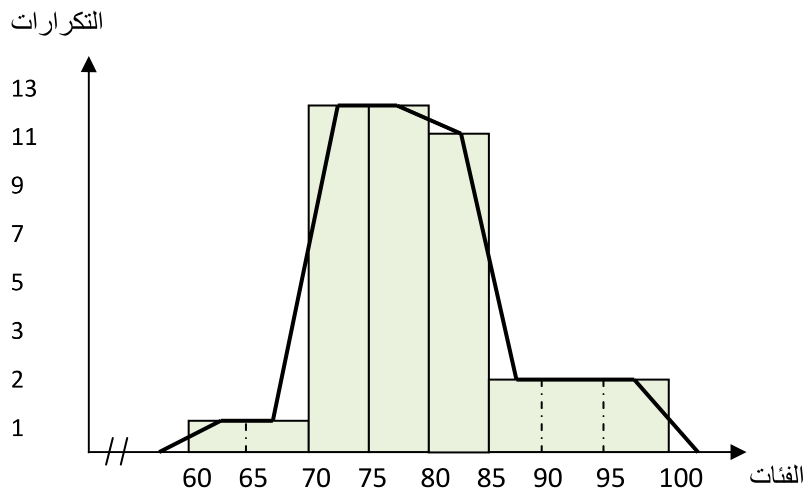


ثانيا: المضلع التكراري في حالة طول الفئات غير متساوي

في حالة كون الفئات غير متساوية الطول، فإننا نرسم المضلع التكراري بواسطة التكرار المعدل.

مثال رقم (3–14).

كون مضلع تكراري من المدرج التكراري الوارد في المثال رقم (3–11)

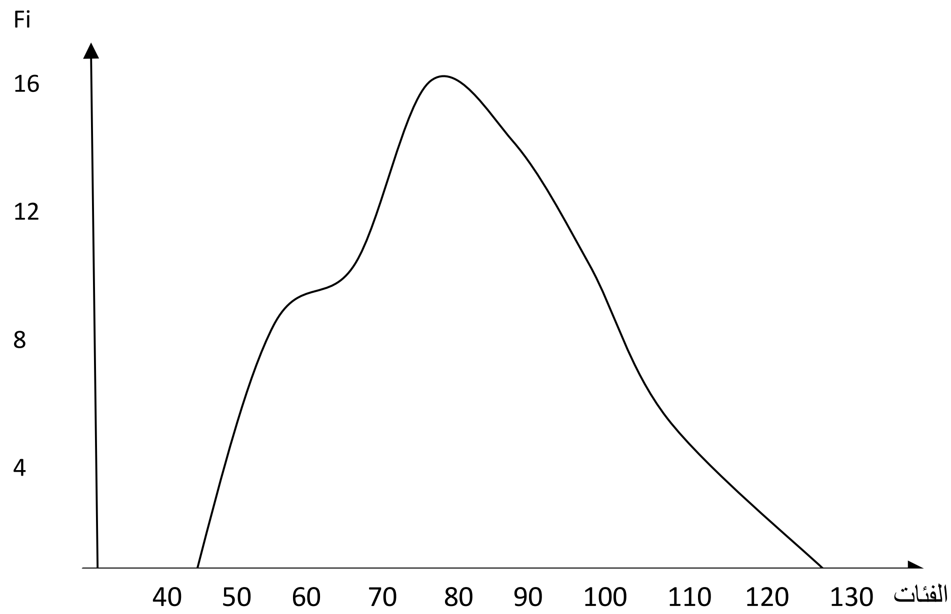


نلاحظ في الشكل السابق أننا قسمنا الفئة الأولى إلى قسمين، والفئة الخامسة إلى ثلاثة أقسام، وهذا ليتساوى طول كل الفئة. بعد ذلك نقوم برسم المضلع التكراري باستخدام الفئات الجديدة المستحدثة. ويمكن إزالة المدرج التكراري بعد رسم المضلع التكراري.

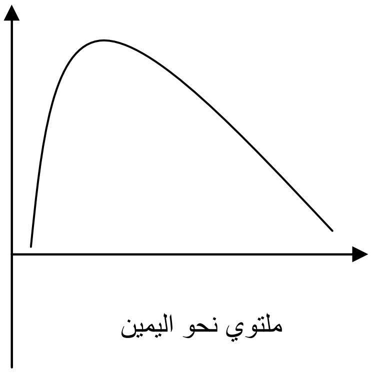
3 – 3 – 2 – 3 المنحنى التكراري

المنحنى التكراري (courbe de fréquence) عبارة عن تمهيد للمضلع التكراري، ويكون رسمه عن طريق تحديد النقط أمام مراكز الفئات، كما في حالة المضلع التكراري، ثم نوصل بينها بواسطة خط ممهد غير متكسر. إن الخط الممهد يعتمد على المهارة الشخصية، حيث قد يقتضي التمهيد عدم وقوع جميع النقاط على الخط المنحني، بل يعبر بينها. وعليه فإن المساحة التي يحصرها المنحنى التكراري قد لا تساوي مساحة المدرج التكراري أو المضلع التكراري.

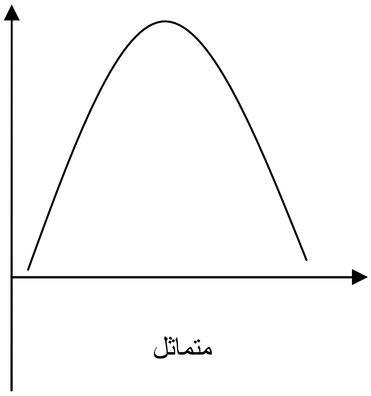
مثال رقم (3–15): أرسم منحنى تكراري لبيانات التمرين رقم (2–1)



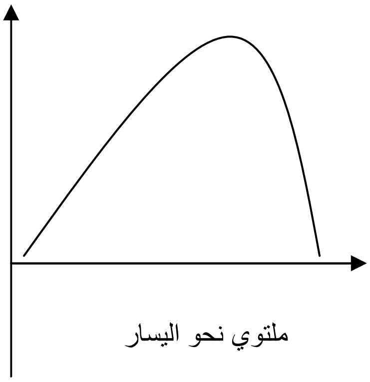
تأخذ المنحنيات التكرارية أشكالا مختلفة، وذلك تبعا لتغير الظاهرة وقيمها، ويمكن لنا أن نعرض بعض هذه الأشكال فيما يلي:



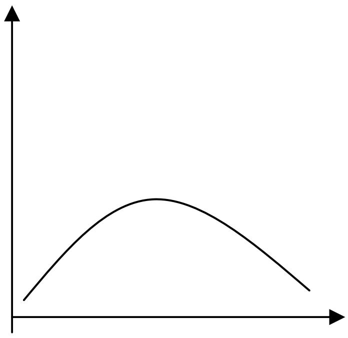
ملتوي نحو اليمين



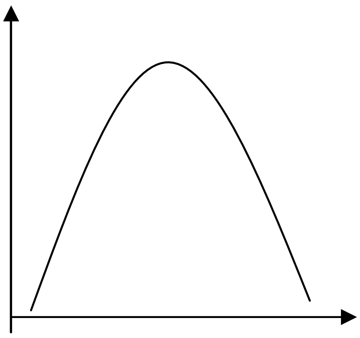
متماثل



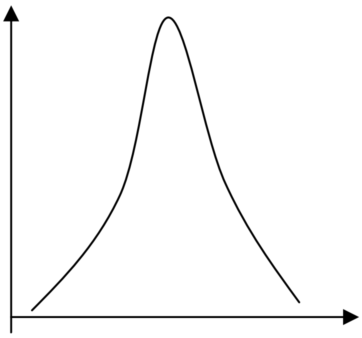
ملتوي نحو اليسار



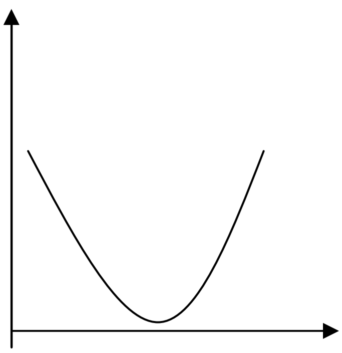
مقعر



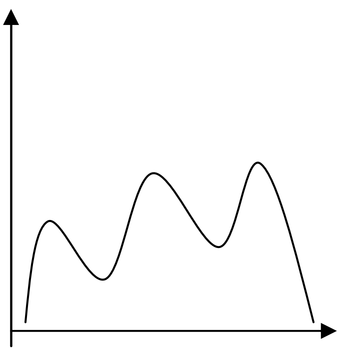
متماثل



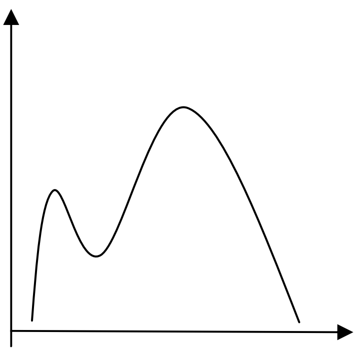
مدبب



نوني



متعدد القيم



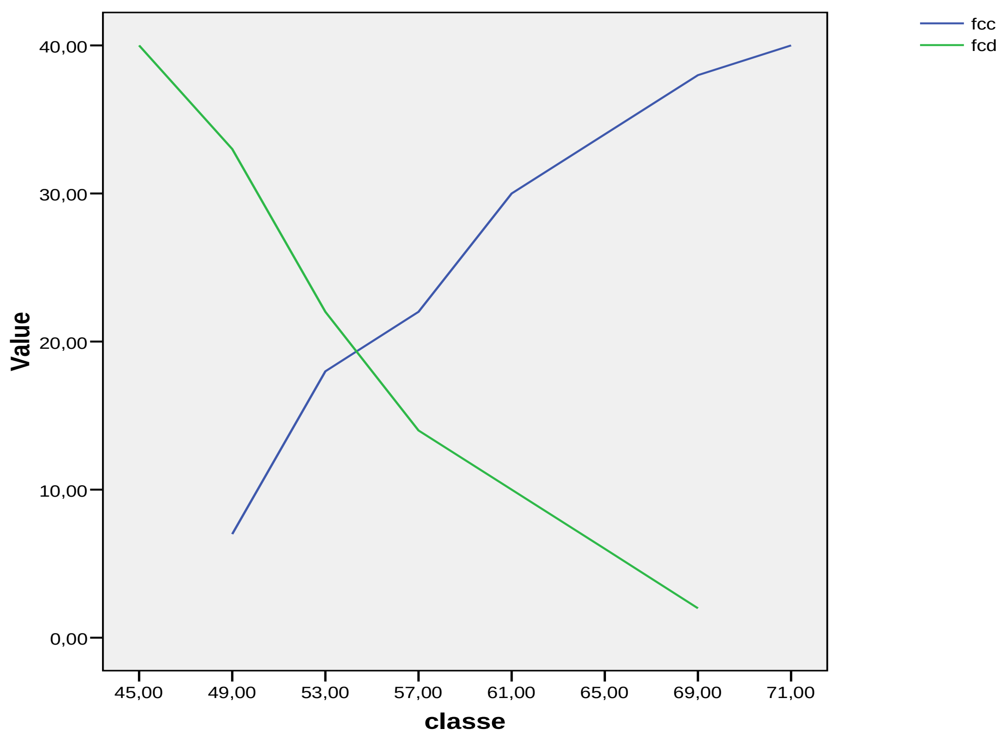
ذو قمتين

3 – 3 – 2 – 4 المنحنى التكراري المتجمع الصاعد والنازل

المنحنى التكراري المتجمع الصاعد والنازل[1] هو الطريقة الرابعة من طرق تمثيل البيانات التكرارية الكمية المتصلة. فإذا رسمنا المنحنى الذي يمثل التكرارات المتجمعة الصاعدة، فإننا نستخدم الحدود العليا للفئات في تحديد إحداثيات النقط التي سنوصل بينها بواسطة المنحنى. أما إذا أردنا رسم منحنى يمثل تكرار تجميعي نازل، فإننا نستخدم الحدود الدنيا للفئات في تحديد هذه الإحداثيات.

مثال رقم (3–16).

أرسم المنحنى التكراري المتجمع الصاعد والمنحنى التكراري المتجمع النازل لبيانات المثال رقم (2–8).



[1] - يمكن رسم منحنى تكراري تجميعي صاعد ونازل نسبي بنفس الطريقة.

3 – 4 تمارين حول الفصل الثالث

تمرين رقم (3–1)

لدينا مجتمع يتكون من 150 عامل، نفترض أن أجورهم بالأورو موزعة كما يلي.

عدد العمال	فئة الأجر
50	أقل من 700
45	700–1200
30	1200–1700
25	1700–2200
15	2200–2700
5	2700 فأكثر
170	المجموع

المطلوب:

1 – ما هي الصفة الإحصائية المدروسة، وما هي طبيعتها؟

2 – كم هو عدد العمال الذين يحصلوا على أجر 1700 أورو فأكثر؟

3 – ما هي نسبة العمل الذين أجورهم تصل إلى 2200 أورو؟

4 – مثل بيانيا هذا الجدول بواسطة مدرج تكراري منحنى تكراري؟ ما هو تعليقك على الشكل المتحصل عليه؟

الحل:

1 – الصفة الإحصائية المدروسة هي أجر العامل وطبيعتها كمية متصلة.

2 – حساب التكرار التجميعي الصاعد والنازل والتكرار التجميعي

الفئات	Fi	Fcc	Frcc(%)	Fcd	Frcd(%)
أقل من 700	50	50	29,41	170	100
700−1200	45	95	55,88	120	70,59
1200−1700	30	125	73,53	75	44,12
1700−2200	25	150	88,24	45	26,47
2200−2700	15	165	97,06	20	11,76
2700 فأكثر	5	170	100	5	2,941
Σ	170				

2 – عدد العمال الذين يحصلوا على أجر 1700 أورو فأكثر هو 45 عامل.

3 – نسبة العمل الذين أجورهم تصل إلى 2200 أورو هو 88.24%.

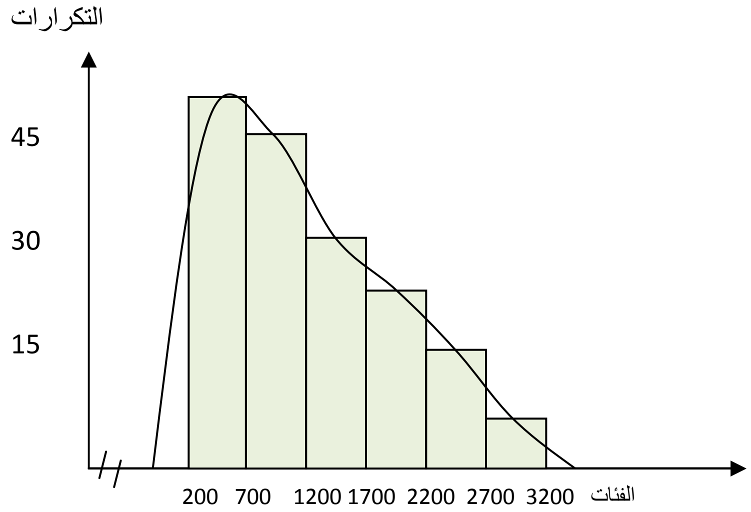
4 – التمثيل البياني

نلاحظ أن الجدول التكراري مفتوح من الأسفل ومن الأعلى. وبالتالي لا يمكن تمثيله بيانيا بدون إجراء بعض الخطوات والتي هي:

– تحديد حد الفئة الأدنى والأعلى. نلاحظ أن طول الفئات متساوي ويساوي 500 أورو. بالتالي فإننا نفرض أن طول الفئتين الأولى والأخيرة متساوي كذلك ويقدر بـ 500أورو. لهذا، فإن الحد الأدنى للفئة الأولى يساوي 200 أورو، والحد الأعلى للفئة الأخيرة يساوي 3200 أورو. وبالتالي فإن الجدول التكراري يصبح بالشكل الموالي.

الفئات	Fi
200−700	50
700−1200	45
1200−1700	30
1700−2200	25
2200−2700	15
2700−**3200**	5
Σ	170

– نقوم برسم المدرج التكراري المرافق للبيانات الجديدة



نلاحظ أن المنحنى التكراري ملتوي نحو اليمين.

الفصل الرابع: مقاييس النزعة المركزية Caractéristiques de tendances centrales

4 – 1 تمهيد

تستخدم مقاييس النزعة المركزية لوصف مجموعة من البيانات عن طريق تقدير قيمة من قيم البيانات بوضع معين. حيث أن قيم البيانات تميل إلى الالتفاف حول قيمة متوسطة. هذا الالتفاف نطلق عليه "النزعة المركزية"، أما القيمة التي تتركز حولها البيانات أكثر من غيرها، وهي النقطة الأكثر تمثيلا لهذه البيانات فنطلق عليها "مقاييس النزعة المركزية". هذا التركز يختلف من مقياس إلى آخر، ولعل أهم هذه المقاييس هي: المتوسط الحسابي، الوسيط والمنوال.

ولقد وضع Yule سنة 1945، ستة معايير (خصائص) أساسية يجب أن تتوفر في معايير النزعة المركزية لتتوفر على سمة النزعة المركزية، وهي:

– يجب أن تكون معرفة بطريقة موضوعية: يجب أن يجد شخصان مختلفان نفس النتيجة العددية من نفس البيانات الإحصائية،

– تعتمد على جميع المشاهدات وليس على جزء منها،

– يكون لها معنى ملموس،

– تكون سهلة الحساب،

– يمكن تمثيلها بسهولة عند نمذجتها رياضيا،

– لا تكون حساسة كثيرا للتقلبات في أخذ العينات.

ونتطرق في هذا الفصل إلى أهم مقاييس النزعة المركزية، من خلال ذكر مزايا كل مقياس وكيفية حسابه سواء كانت البيانات متقطعة أم مستمرة. حيث أنه من الجدير بالذكر، أن مقاييس النزعة المركزية لا تحسب في هذا المستوى إلا في حالة البيانات الكمية، أما طرق حساب هذه المقاييس في البيانات الكيفية فتدرس في مستويات أعلى.

4 – 2 المتوسط الحسابي ($\bar{X}$) Moyen

4 – 2 – 1 تعريف المتوسط الحسابي

المتوسط الحسابي لمجموعة من القيم هو مجموع هذه القيم مقسوما على عددها(معدل القيم). ويرمز للمتوسط الحسابي في حالة حسابه من مشاهدات العينة بالرمز $\bar{X}$، أما في حالة حسابه من مشاهدات المجتمع فيرمز له بـ μ.

والمتوسط الحسابي تتوفر فيه جميع خصائص Yule، ما عدا الخاصية السادسة، أي أنه حساس للتقلبات في أخذ العينات.

4 – 2 – 2 طرق حسابه

4 – 2 – 2 – 1 إيجاد المتوسط الحسابي للبيانات غير المبوبة

هنا يمكن التمييز بين:

– حالة البيانات المفردة (الوسط الحسابي البسيط)

إذا كانت لدينا القيم $X_1, X_2, \dots X_n$ فإن المتوسط الحسابي لهذه القيم هو:

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

مثال رقم (4–1).

لديك بيانات تمثل علامات 5 طلبة في مقياس مدخل للاقتصاد 1 وهي: 15، 7، 10، 8، 12.

المتوسط الحسابي هو:

$$\bar{X} = \frac{15 + 7 + 10 + 8 + 12}{5} = 10.4$$

– حالة البيانات المتكررة (الوسط الحسابي المرجح)

إذا كانت لدينا القيم $x_1, x_2, \dots x_n$، وتكراراتها المقابلة هي $F_1, F_2, \dots F_n$، فإن المتوسط الحسابي لهذه القيم هو:

$$\overline{X}=\frac{f_1\times x_1+f_2\times x_2+......f_n\times x_n}{f_1+f_2+......f_n}=\frac{\sum_{i=1}^{n}f_ix_i}{\sum_{i=1}^{n}f_i}=\sum_{i=1}^{n}fr_i\times x_i$$

مثال رقم (4-2).

في مستوى دراسي معين، تحصل الطالب على النقاط المدرجة في الجدول الموالي.

المقياس	مقياس 1	مقياس 2	مقياس 3	مقياس 4	مقياس 5
المعامل	1	2	4	2	2
النقطة	15	10	7	5.5	10

حساب المتوسط الحسابي

$$\bar{X}=\frac{\sum_{i=1}^{5}f_ix_i}{\sum_{i=1}^{5}f_i}=\frac{(15\times 1)+(10\times 2)+\cdots+(10\times 2)}{11}=8.54$$

4 – 2 – 2 – 2 إيجاد المتوسط الحسابي للبيانات المبوبة

هناك عدة طرق لحسابه ونستعرض فيما يلي أهمها:

– الطريقة المباشرة (طريقة القانون العام)

تكون صيغة الوسط الحسابي بالشكل التالي:

$$\overline{X}=\frac{\sum_{i=1}^{n}f_ic_i}{\sum_{i=1}^{n}f_i}=\sum_{i=1}^{n}fr_ic_i$$

– الطريقة غير المباشرة

يمكن الحصول على نفس نتائج المتوسط الحسابي المحسوب بالطريقة المباشرة، وذلك باستخدام إحدى الطريقتين غير المباشرتين التاليتين:

*** طريقة استخدام الوسط الفرضي (طريقة الانحرافات)**

تكون صيغة المتوسط الحسابي كما يلي:

$$\overline{X} = \alpha + \frac{\sum_{i=1}^{n} f_i w_i}{\sum_{i=1}^{n} f_i}$$

حيث αهو الوسط الفرضي، وغالبا ما يمثل مركز الفئة المقابلة للأكبر تكرار، و w_i يمثل انحراف مراكز الفئات عن الوسط الفرضي $w_i = c_i - \alpha$

*** طريقة استخدام الانحرافات المختصرة**

تكون صيغة المتوسط الحسابي بهذه الطريقة كما يلي:

$$\overline{X} = \alpha + \frac{\sum_{i=1}^{n} f_i w_i'}{\sum_{i=1}^{n} f_i} \times k$$

حيث α هو الوسط الفرضي، w_i يمثل انحراف مراكز الفئات عن الوسط الفرضي، k هو طول الفئة، $w_i' = \frac{w_i}{k}$

ملاحظة:

تستخدم كل من الطريقة المباشرة وطريقة الوسط الفرضي في الجداول ذات الفئات المتساوية أو غير المتساوية. بينما تستخدم طريقة الانحرافات المختصرة في الجداول التكرارية ذات المتساوية فقط.

مثال رقم (4-3).

استخدم التوزيع التكراري التالي لحساب المتوسط الحسابي.

الفئات	Fi
59.5-62.5	5
62.5-65.5	18
65.5-68.5	42
68.5-71.5	27
71.5-74.5	8
$\sum$	100

الحل:

نقوم أولا بإعداد جدول يتضمن كل المجاميع اللازمة لحساب المتوسط الحسابي.

الفئات	Fi	ci	$Fi\times ci$	$w_i=c_i-\alpha$	$Fi\times wi$	$w'=w/k$	$Fi\times w'$
59.5-62.5	5	61	305	-6	-30	-2	-10
62.5-65.5	18	64	1152	-3	-54	-1	-18
65.5-68.5	42	67	2814	0	0	0	0
68.5-71.5	27	70	1890	3	81	1	27
71.5-74.5	8	73	584	6	48	2	16
Σ	100	-	6745	0	45	0	15

– حساب الوسط الحسابي باستخدام الطريقة المباشرة:

$$\overline{X}=\frac{\sum_{i=1}^{n} f_i c_i}{\sum_{i=1}^{n} f_i}=\frac{6745}{100}=67.45$$

– حساب الوسط الحسابي باستخدام الطريقة غير المباشرة:

* طريقة استخدام الوسط الفرضي: يمكن أن نختار مركز الفئة المقابلة لأكبر تكرار كوسط فرضي والتي تساوي 67 ومنه:

$$\overline{X}=\alpha+\frac{\sum_{i=1}^{n} f_i w_i}{\sum_{i=1}^{n} f_i}=67+\frac{45}{100}=67.45$$

* طريقة استخدام الانحرافات المختصرة: طول الفئة هو 3 ومنه:

$$\overline{X}=\alpha+\frac{\sum_{i=1}^{n} f_i w'_i}{\sum_{i=1}^{n} f_i}\times k=67+\frac{15}{100}\times 3=67.45$$

4 – 2 – 3 خصائص المتوسط الحسابي

يتميز المتوسط الحسابي بعدة خصائص نذكر منها:

– يعتبر أبسط مقاييس النزعة المركزية حسابا وأكثرها استخداما،

– يأخذ بعين الاعتبار جميع قيم الظاهرة المدروسة،

– مجموع انحرافات القيم عن وسطها الحسابي يساوي الصفر،

$$\sum_{i=1}^{n} f_i(x_i - \bar{X}) = \sum_{i=1}^{n} f_i x_i - n\bar{X} = \sum_{i=1}^{n} f_i x_i - n\left(\frac{\sum_{i=1}^{n} f_i x_i}{n}\right)$$
$$= \sum_{i=1}^{n} f_i x_i - \sum_{i=1}^{n} f_i x_i = 0$$

ملاحظة: يكون إدخال تعبير المجموع $\sum$ على القوس في ثلاث حالات نشرحها بالمثال التالي:

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

– عند إضافة (طرح أو ضرب) مقدار ثابت a إلى كل قيمة من القيم في العينة أو المجتمع، فإن الوسط الحسابي يساوي الوسط الحسابي الأصلي مضافا (مطروحا أو مضروبا) إليه المقدار الثابت.

$$\bar{X} = \frac{\sum_{i=1}^{n} f_i(x_i + a)}{\sum_{i=1}^{n} f_i} = \bar{X} + \mathrm{a}$$

– يتأثر بالقيم المتطرفة،

– لا يمكن إيجاده من الجداول التكرارية المفتوحة.

4 – 3 الوسيط (Me) Médiane

4 – 3 – 1 تعريف الوسيط

هو قيمة المتغير الإحصائي الذي يفصل السلسلة الإحصائية إلى قسمين متساويين بعد ترتيب القيم تصاعديا أو تنازليا.

4 – 3 – 2 طرق حسابه

تختلف طرق حسابه باختلاف البيانات المراد حساب الوسيط لها، حيث نميز بين:

4 – 3 – 2 – 1 الوسيط للبيانات غير المبوبة

إذا كان لدينا عدد معين من قيم المشاهدات (n)، فلحساب الوسيط نتبع الخطوات التالية:

أ – نقوم بترتيب البيانات تصاعديا أو تنازليا.

ب – نقوم بتحديد رتبة الوسيط (Rme) حسب عدد القيم إذا كان زوجيا أو فرديا:

الحالة الأولى:

إذا كان عدد القيم (n) فرديا فيكون $Rme = \frac{n+1}{2}$ و تكون قيمة الوسيط Me هي القيمة التي ترتيبها Rme.

مثال رقم (4–4).

لتكن القيم التالية 16.11.13.13.10.14.12.11.15،

نرتب البيانات تصاعديا 10 ، 11 ، 11 ، 12 ، 13 ، 13 ، 14 ، 15 ، 16.

نحدد رتبة الوسيط بالعلاقة: $Rme = \frac{9+1}{2}$

الوسيط يشغل الرتبة الخامسة ويساوي Me=13.

الحالة الثانية:

إذا كان عدد القيم زوجيا:

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

في هذه الحالة نحدد قيمتين للوسيط الأولى ترتيبها $\frac{n}{2}$ وهي Me_1 والثانية ترتيبها $\frac{n}{2}+1$ وهي Me_2، وبالتالي قيمة الوسيط هي:

$$Me = \frac{Me_1 + Me_2}{2}$$

مثال رقم (4–5):

لتكن القيم التالية 3.15.9.7.5.10.18.16.11.20

لإيجاد الوسيط نتبع الخطوات التالية:

نرتب البيانات تصاعديا 3 . 5 . 7 . 9 . 10 . 11 . 15 . 16 . 18 . 20

تحديد رتبة الوسيط:

$$Rme = \begin{cases} \frac{n}{2} = \frac{10}{2} = 5 \Rightarrow Me_1 = 10 \\ \frac{n}{2} + 1 = \frac{10}{2} + 1 = 6 \Rightarrow Me_2 = 11 \end{cases}$$

ومنه: $Me = \frac{Me_1 + Me_2}{2} = \frac{10+11}{2} = 10.5$

4 – 3 – 2 – 2 الوسيط في حالة البيانات المبوبة

لإيجاد الوسيط في هذه الحالة نتبع الخطوات التالية في كلتا الحالتين:

– حالة المتغير المتصل:

أ – جعل التوزيع التكراري توزيعا متجمعا صاعدا أو نازلا،

ب – تحديد رتبة الوسيط و ذلك باستخدام العلاقة $Rme = \frac{\sum fi}{2}$

ج – يستخدم ترتيب الوسيط لتحديد الفئة التي يقع فيها الوسيط (من جدول التوزيع التكراري المتجمع) وتدعى بالفئة الوسيطية،

د – نحسب الوسيط باستخدام العلاقة في حالة التكرار المتجمع الصاعد:

$$Me = L_0 + \frac{\frac{\sum fi}{2} - F_1}{F_2 - F_1} \times K$$

K: هو طول الفئة الوسيطية،

L_0: هو الحد الأدنى للفئة الوسيطية،

F_1: هو التكرار المتجمع الصاعد السابق لرتبة الوسيط ،

F_2: هو التكرار المتجمع الصاعد اللاحق لرتبة الوسيط.

مثال رقم (4–6).

أوجد قيمة الوسيط لأوزان الـ 100 طالب.

الحل:

الفئات	Fi	حدود الفئات	Fcc	
		أقل من 59.5	0	
59.5-62.5	5	أقل من 62.5	5	
62.5-65.5	18	أقل من 65.5	**23**	
65.5-68.5	42	أقل من 68.5	**65**	الفئة الوسيطية
68.5-71.5	27	أقل من 71.5	92	
71.5-74.5	8	أقل من 74.5	100	
$\sum$	100	-	-	

- تحديد رتبة الوسيط

$$Rme = \frac{\sum fi}{2} = \frac{100}{2} = 50$$

- نحسب الوسيط بالعلاقة

$$Me = L_0 + \frac{\frac{\sum fi}{2} - F_1}{F_2 - F_1} \times K = 65.5 + \frac{50-23}{65-23} \times 3 = 67.42$$

ملاحظة: إذا كانت رتبة الوسيط هي إحدى قيم التكرار المتجمع الصاعد فأن قيمة الوسيط تحدد مباشرة وهي الحد الأعلى للفئة الوسيطية.

- حالة المتغير المنفصل

أ - جعل التوزيع التكراري متجمعا صاعدا،

ب - تحديد رتبة الوسيط و ذلك باستخدام العلاقة $Rme = \frac{\sum fi}{2}$ وهنا نجد حالتين:

*** الحالة الأولى:**

إذا كانت رتبة الوسيط مساوية لإحدى قيم التكرار المتجمع الصاعد فإن الوسيط يكون مساويا لمتوسط القيمتين،

*** الحالة الثانية:**

إذا كانت رتبة الوسيط غير مساوية لإحدى قيم التكرار المتجمع الصاعد فإن الوسيط يكون مساويا للقيمة المقابلة لرتبته.

مثال رقم (4–7).

ليكن الجدولين الإحصائيين التاليين:

– الحالة الأولى.

Xi	Fi	Fcc
		0
0	3	
		3
1	5	
		8
2	6	
		14
3	5	
		19
4	4	
		23
5	2	
		25
6	3	
		28
المجموع	28	-

$$\frac{n}{2}=\frac{28}{2}=14$$

القيمة 14 موجودة ضمن التكرارات المتجمعة الصاعدة إذن الوسيط محصور بين القيمة 2 و 3 و يساوي:

$$me=\frac{2+3}{2}=2.5$$

– الحالة الثانية

Xi	Fi	Fcc
		0
0	2	
		2
1	3	
		5
2	5	
		10
3	4	
		14
4	2	
		16
5	2	
		18
المجموع	18	-

$$\frac{n}{2}=\frac{18}{2}=9$$

القيمة 9 غير موجودة ضمن التكرارات المتجمعة الصاعدة إذن الوسيط يقابل القيمة 2 في الجدول

$$me=2$$

4 – 3 – 2 – 3 تعيين قيمة الوسيط بيانيا

يمكن تحديد قيمة الوسيط بيانيا وفقا لإحدى الحالتين التاليتين:

– حالة المتغير المنفصل

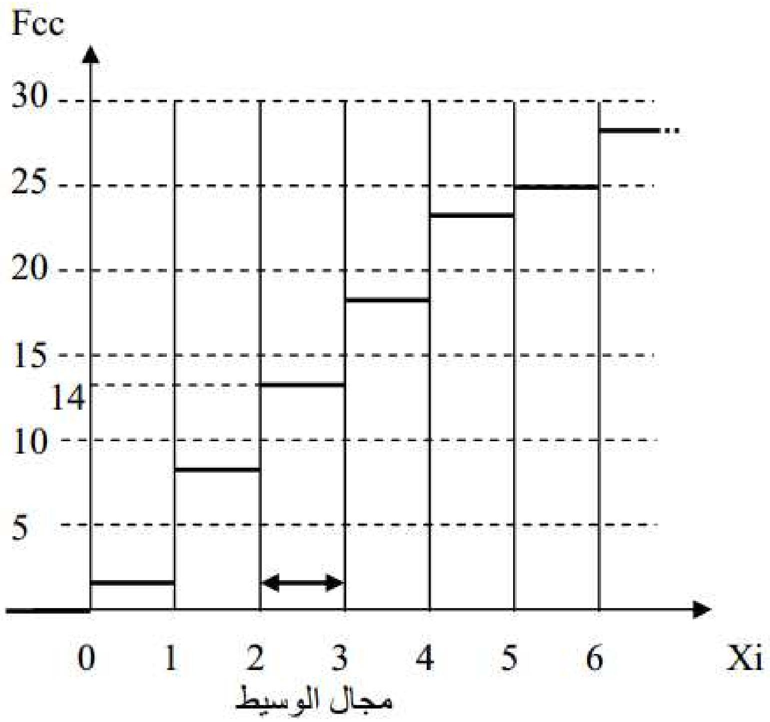
نقوم برسم التكرار التجميعي الصاعد، ثم نقوم بتحديد رتبة الوسيط على محور التراتيب، حيث يمكن أن نجد حالتين:

*** الحالة الأولى**

رتبة الوسيط تقع ضمن مجال محدد بقيمتين على محور الفواصل، ومنه تأخذ قيمة الوسيط المتوسط الحسابي للقيمتين الناتجتين.

مثال رقم (4–8).

أوجد قيمة الوسيط في التكرار التجميعي الصاعد التالي.

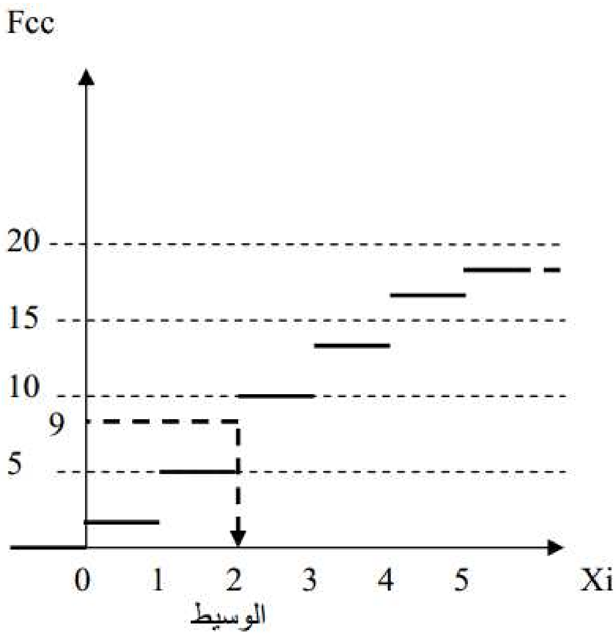


*** الحالة الثانية**

نقوم بإسقاط رتبة الوسيط على التمثيل البياني للتكرار التجميعي الصاعد، ثم نقوم بإسقاط تلك النقطة على محور الفواصل التي تكون قيمة الوسيط.

مثال رقم (4-9).

أوجد قيمة الوسيط في التكرار التجميعي الصاعد التالي.



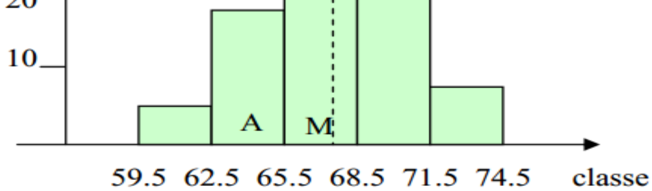
- حالة المتغير المتصل

***باستخدام المدرج التكراري**

قيمة الوسيط هي النقطة على محور الفواصل للخط LM، الذي يقسم المدرج التكراري إلى مساحتين متساويتين، أي أن التكرارات على يمينه والتكرارات على يساره تكون متساوية.

مثال رقم (4-10).

باستخدام بيانات المثال رقم (4-3)، أوجد قيمة الوسيط باستخدام المدرج التكراري.



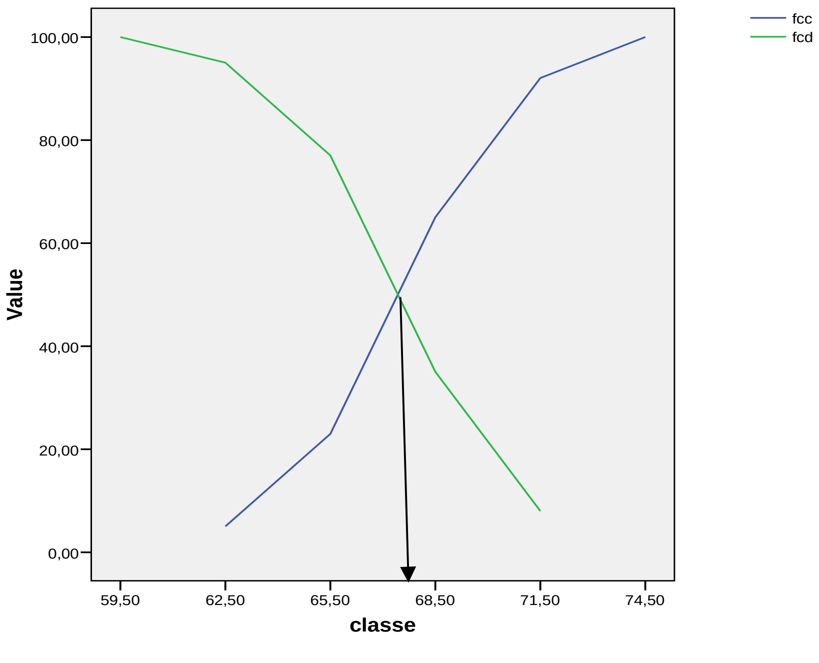
يوضح الشكل المدرج التكراري المقابل لأوزان الـ 100 طالب، وقيمة الوسيط هي النقطة على محور الفواصل للخط LM، الذي يقسم المدرج التكراري إلى مساحتين متساويتين، أي أن التكرارات على يمينه و التكرارات على يساره تكون متساوية وتساوي 50، وحيث أن مجموع تكرار الفئة الأولى و الثانية هو 23 = 18 + 5 بالتالي فإننا نريد 27 رقم من الـ 42 تكرار الموجود في الفئة الثالثة، بالتالي الوسيط يقع في 27/42 المسافة بين 65.5 و 68.5، وبهذا فإن $AM = \frac{27}{42} \times 3 = 1.92$ وقيمة الوسيط هي $65.5+1.92=67.42$

***باستخدام التكرار التجميعي الصاعد والنازل**

عند تمثيل التكرار التجميعي الصاعد والنزل بيانيا، فإن نقطة تقاطعهما تمثل منتصف البيانات الإحصائية، والمسقط العمودي لهذه النقطة على محور الفواصل يمثل قيمة الوسيط.

مثال رقم (4−11).

باستخدام بيانات المثال رقم (4−3)، أوجد قيمة الوسيط باستخدام التكرار التجميعي.



الشكل السابق يوضح المنحنى التكراري المتجمع الصاعد والنازل المقابل لأوزان الـ 100 طالب، وقيمة الوسيط هي النقطة على محور الفواصل التي تمثل المسقط العمودي لنقطة تقاطع المنحنيين.

4 – 3 – 3 خصائص الوسيط

للوسيط عدة خصائص نذكر منها:

- لا يتأثر بالقيم المتطرفة،
- يحسب من الجداول المفتوحة،
- يبين الوسيط القيمة الوسطى للتوزيع.

بهذا، فإن الوسيط يستجيب لقواعد Yule رقم 1، 3، 4 و6. حيث أنه لا يتبع قيم المشاهدات بل يتبع ترتيبها، وهو الأمر الذي يؤدي إلى وجود تحيز عندما تكون هناك مشاهدات متطرفة في العينة (القاعدة 2) كذلك توجد عدة صعوبات في تطبيق القواعد الرياضية في حسابه عندما يكون المتغير منفصل (القاعدة 5).

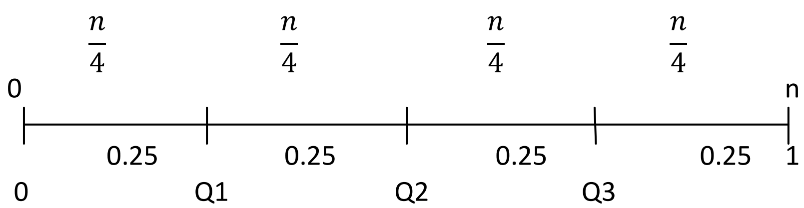
4 – 3 – 4 مشتقات الوسيط

للوسيط عدة مشتقات تسمى بالمجزآت (Quantiles) تؤدي نفس المبدأ وهو تقسيم التكرار الكلي إلى أجزاء بنسب معينة[1]. ومن هذه المشتقات نجد:

4 – 3 – 4 – 1 الربيعيات Quartiles

تعرف الربيعيات بأنها القيم xi التي تجزئ البيانات إلى أربعة أجزاء متساوية. من هذا التعريف نستنتج أن عدد الربيعيات هو 3 ونرمز لهم بـ: Q1، Q2 وQ3.

إن المجالات التي تحددها الربيعيات يكون طولها 25% من المشاهدات أو يكون ربع المشاهدات n/4 كما يوضحه الشكل الموالي. ويسمى المجال Q3–Q1 بالمجال داخل الربيعي Interval interquartile، ويحتوي على 50% من المشاهدات. وبهذا، فإن قيمة الربيع الثاني تكون نفسها قيمة الوسيط.



1- هذه المشتقات لا تعتبر مقاييس نزعة مركزية باستثناء تلك التي تتساوى قيمها مع قيمة الوسيط.

يتم حساب الربيعيات بنفس منهجية حساب الوسيط.

– حالة البيانات المفردة

يتم ترتيب البيانات تصاعديا، ثم نقوم بتحديد رتبة الربيع Qi بالعلاقة $Rme_i = \frac{in}{4}$ التي تقع عندها قيمة الربيع.

مثال رقم (4−12).

لتكن لديك البيانات التالية: 1، 3، 5، 6، 7، 9، 10، 12، 14، 15، 16، 20.

نحدد رتبة الربيع بالعلاقة:

$$Rme_1 = \frac{1(12)}{4} = 3$$

$$Rme_2 = \frac{2(12)}{4} = 6 = Me$$

$$Rme_3 = \frac{3(12)}{4} = 9$$

1	3	5	Q1	6	7	9	Q2	10	12	14	Q3	15	16	20
			5.5				10.5				14.5			

– حالة التوزيع التكراري

*** حالة المتغير الإحصائي المنفصل**

عندما لا تكون لدينا بيانات مفردة، فإنه يكون لدينا جدول تكراري يضم قيم كل مشاهدة منفصلة مع تكرارها. ولحساب الربيع Q_i حيث (i=1.2.3)، نقوم بتحديد موقع القيمة $\frac{i\sum fi}{4}$ بالنسبة للتكرارات المتجمعة الصاعدة. ونستخرج قيمة الربيع بنفس القاعدة المستخدمة لاستخراج الوسيط كالتالي:

– إذا كانت رتبة الربيع مساوية لتكرار متجمع صاعد، فإن قيمة الربيع تكون قيمة المتغير الواقعة في السطر التالي لقيمة رتبة الربيع،

– إذا كانت رتبة الربيع غير مساوية لتكرار متجمع صاعد، فإن قيمة الربيع تكون مساوية للمتوسط الحسابي بين قيمة المتغيرين التي وقعت بينهما.

مثال رقم (4–13).

باستخدام بيانات المثال رقم (4–7) أوجد الربيعيات.

xi	ni	Fcc
0	3	3
1	5	8
2	6	14
3	5	19
4	4	23
5	2	25
6	3	28
$\sum$	28	

$n/4 = 28/4 = 7$ ← $Q_1 = 1$

$n/2 = 28/2 = 14$ ← $Q_2 = Me = 2.5$

$3n/4 = 3 \times 28/4 = 21$ ← $Q_3 = 4$

*** حالة المتغير الإحصائي المتصل**

نحسب الربيع Q_i حيث (i=1.2.3) من العلاقة التالية:

$$Q_i = L_0 + \frac{\frac{i\sum fi}{4} - F_1}{F_2 - F_1} \times K$$

K: هو طول الفئة الربيعية،

L_0: هو الحد الأدنى للفئة الربيعية ،

F_1: هو التكرار المتجمع الصاعد السابق لتكرار فئة الربيع ،

F_2: هو التكرار المتجمع الصاعد اللاحق لتكرار فئة الربيع.

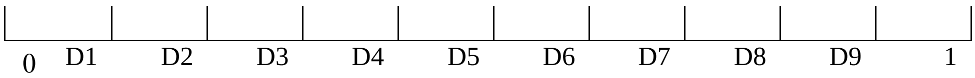
4 – 3 – 4 – 2 العشيريات Deciles

العشيريات عبارة عن قيم خاصة بـ xi التي تقسم المشاهدات إلى 10 مجموعات جزئية متساوية. وبالتالي عدد العشيريات هو 9، والتي نرمز لها بالرمز Di بحيث (i=1,2…9).

يتم حساب العشيريات بنفس منهجية حساب الوسيط.

4 – 3 – 4 – 3 المئينات centiles

المئينات عبارة عن قيم خاصة بـ xi التي تقسم المشاهدات إلى 100 مجموعة جزئية متساوية. وبالتالي عدد المئينات هو 99، والتي نرمز لها بالرمز Pi بحيث (i=1,2,..99).



يتم حساب العشيريات بنفس منهجية حساب الوسيط.

4 – 4 المنوال (Mo) Mode

4 – 4 – 1 تعريف المنوال

المنوال هو القيمة الأكثر تكرارا أو شيوعا بين قيم المشاهدات، وبالتالي فهي القيمة السائدة في المشاهدات.

4 – 4 – 2 طرق حسابه

يتم حساب المنوال حسب طبيعة البيانات الكمية كما يلي:

4 – 4 – 2 – 1 إيجاد المنوال للبيانات غير المبوبة

وهنا يمكن التمييز بين:

– إذا لم تتكرر أي من القيم فلا وجود للمنوال.

مثال رقم (4–14).

أوجد المنوال لمجموعة القيم التالية: 7.5.3.2.810

لا يوجد منوال لأن أي من القيم لم يتكرر.

– إذا تكرر أحد القيم فيكون هناك منوال واحد

مثال رقم (4–15).

أوجد المنوال لمجموعة القيم التالية:7.5.3.10.2.7.8.10.10

يوجد منوال واحد هو 10.

- إذا كان لقيمتين نفس العدد من التكرار،

يكون لمجموعة القيم منوالان، وهكذا تزداد المنوالات بزيادة الأعداد المتساوية للتكرارات.

مثال رقم (4−16).

أوجد المنوال لمجموعة القيم التالية:7.5.3.10.2.7.8.10

الحل: يوجد منوالان هما 7.10.

4 – 4 – 2 – 2 إيجاد المنوال للبيانات المبوبة

أولا: إيجاد قيمة المنوال حسابيا

- طريقة مركز الفئة المنوالية

تعتبر هذه الطريقة من أبسط طرق حساب المنوال وأقلها دقة، والمقصود بالفئة المنوالية هي تلك الفئة التي تقابل أكبر تكرار، ومن ثم فإن المنوال يساوي مركز الفئة المنوالية. وتعتبر قيمة المنوال المحسوب بهذه الطريقة قيمة تقريبية لأنه ليس من الضروري أن تكون قيمة المنحنى تناظر مركز الفئة إلا في حالة كون التوزيع متماثل.

مثال رقم (4−17).

يمثل الجدول الموالي أجور 65 عامل بالأورو.

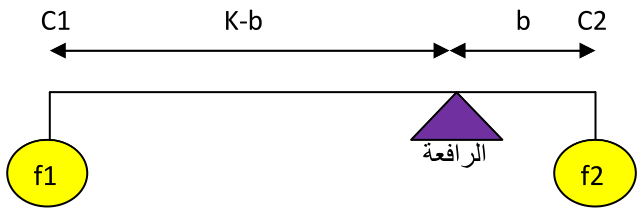
عدد العمال	فئات الأجور
10	20−30
12	30−40
27	40−50
11	50−60
5	60−70
65	$\sum$

الفئة المنوالية هي الفئة المقابلة لأكبر تكرار (27) وهي 40−50، وبالتالي فإن قيمة المنوال تساوي:

$$Mo = \frac{40+50}{2} = 45$$

– طريقة الرافعة

كما أشرنا سابقا، فإنه ليس من الضروري أن يكون قمة التوزيع عند مركز الفئة المنوالية، إلا في حالة كون التوزيع متماثل. وفي حالة كون المنحنى غير متماثل، فإن المنوال تبتعد قيمته عن مركز الفئة المنوالية قليلا سواء بالزيادة أو النقصان. ويرجع السبب إلى تسمية هذه الطريقة بطريقة الرافعة في أنها تعتمد على نفس مبدئها، بحيث يكون تكرار الفئة السابقة للفئة المنوالية القوة الدافعة، وتكرار الفئة اللاحقة للفئة المنوالية القوة المقاوَمة. وعلى هذا الأساس يتحدد موضع المنوال عند محور ارتكاز هذه الرافعة والتي تمثل نقطة توازنها. وحسب القانون الفيزيائي للرافعة، فإن عزم قوة الدفع في ذراعها يكون مساوي لعزم قوة المقاوَمة في ذراعها، كما يظهره الشكل التالي.



بتطبيق هذه القاعدة نجد:

$$f_1 b = f_2(k - b)$$
$$f_1 b = f_2(k) - f_2(b)$$

ومنه فإن قيمة المنوال تساوي الحد الأدنى للفئة المنوالية زائد القيمة b، أي:

$$Mo = C_{min} + b$$

مثال رقم (4–18).

أوجد المنوال بطريقة الرافعة لبيانات المثال رقم (4–17).

بتطبيق القانون نجد:

$$b = \frac{f_2}{f_1 + f_2} \times K = \frac{11}{11 + 12} \times 10 = 4.78$$

وبالتالي فإن قيمة المنوال تساوي:

$$Mo=C_{min}+b=40+4.78= 44.78$$

- طريقة الفروق

لحساب المنوال نتبع الخطوات التالية:

- تحديد الفئة المنوالية: وهي الفئة التي تقابل أكبر تكرار.

- تقدير قيمة المنوال باستخدام طريقة الفروق:

$$Mo = Lo + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times K$$

أو

$$Mo = L_1 - \frac{\Delta_2}{\Delta_1 + \Delta_2} \times K$$

$$\Delta_1 = f_0 - f_1$$

$$\Delta_2 = f_0 - f_2$$

L_0: الحد الأدنى للفئة المنوالية، f_0: تكرار الفئة المنوالية.

L_1: الحد الأعلى للفئة المنوالية، f_1: تكرار الفئة السابقة للفئة المنوالية.

K: طول الفئة المنوالية، f_2: تكرار الفئة اللاحقة للفئة المنوالية.

مثال (4-19).

أوجد المنوال بطريقة الرافعة لبيانات المثال رقم (4-17).

لحساب المنوال نتبع الخطوات التالية:

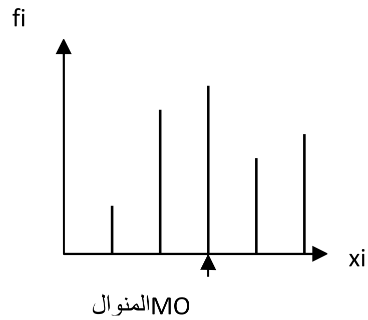
الفئة المنوالية هي الفئة المقابلة لأكبر تكرار (27) وهي 40-50،

$$Mo = L_0 + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times K = 40 + \frac{27 - 12}{(27 - 12) + (27 - 11)} \times 10 = 44.84$$

ثانيا: استخراج قيمة المنوال بيانيا

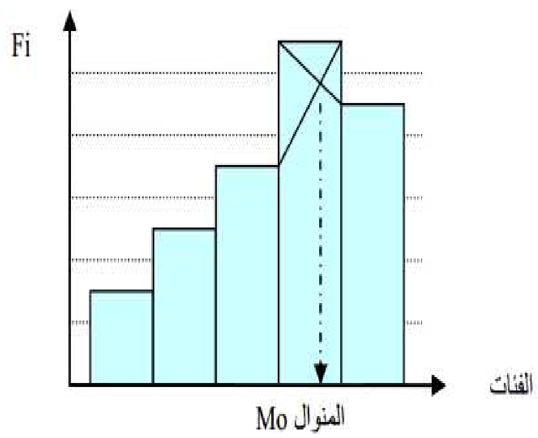
أ – حالة المتغير الإحصائي المنفصل:

نستخرج المنوال من مخطط الأعمدة التكرارية، و هو قيمة التي تناسب الخط العمودي الأكثر Xالمتغير الإحصائي ارتفاعا (أكثر تكرارا)



ب – حالة المتغير الإحصائي المتصل

لتحديد قيمة المنوال بيانيا يجب أولا تحديد الفئة المنوالية، الفئة السابقة لها و الفئة اللاحقة لها ثم نقوم برسم خط مستقيم يصل بين الحد الأدنى للفئة المنوالية والحد الأدنى للفئة اللاحقة للفئة المنوالية، ثم نرسم مستقيم يصل بين الحد الأعلى للفئة المنوالية والحد الأعلى للفئة السابقة للفئة المنوالية، ثم نقوط بإسقاط نقطة تقاطع المستقيمين على محور الفواصل التي تمثل لنا قيمة المنوال.



ملاحظة: إذا كانت أطوال الفئات غير متساوية فإنه يجب استخدام التكرارات المعدلة hi بدلا من التكرارات Fi لحساب قيمة المنوال سواء حسابيا أو بيانيا.

4 – 4 – 3 خصائص المنوال

يتميز المنوال بعدة خصائص هي:

– لا يتأثر بالقيم المتطرفة،

– من أفضل المتوسطات لوصف الظواهر الكيفية،

– المقياس الوحيد القابل للاستعمال عندما تكون الظواهر غير قابلة للقياس،

– يتبع طريقة اختيار الفئات،

– يتأثر بالفئات غير المتساوية،

عموما، يتوافق المنوال مع القواعد 1، 3 و4 لـ Yule، ولا يتوافق مع باقي القواعد.

4 – 5 متوسطات أخرى

توجد العديد من المتوسطات التي لا تقل أهميتها عن المتوسطات السابقة، ولها العديد من الاستعمالات في المجالات البحثية، ومن بين هذه المتوسط نجد:

4 – 5 – 1 المتوسط الهندسي: Moyen Géométrique (G)

يستخدم هذا المتوسط لوصف ظاهرة حسب نسبة تغيرها، وخصوصا عندما يكون سلوك الظاهرة يتبع نمط المتتالية الهندسية.

4 – 5 – 1 – 1 المتوسط الهندسي البسيط

وهو عبارة عن الجذر النوني لجداء القيم X_i ويعبر عنه بالشكل التالي:

$$G_s = \sqrt[n]{x_1 \times x_2 \times \cdots \times x_n}$$

وبشكل مبسط

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

مثال رقم (4–20).

لتكن لدينا القيم التالية: 3، 5، 6، 6، 7، 10، 12.

الوسط الهندسي البسيط لهذه القيم هو:

$$G_s = \sqrt[7]{3 \times 5 \times 6 \times 6 \times 7 \times 10 \times 12} = 6.34$$

4 – 5 – 1 – 2 المتوسط الهندسي المرجح

– حالة المتغير المنفصل

يعطى بالعلاقة التالية:

$$G_P = \sqrt[\sum f_i]{x_1^{f_1} \times x_2^{f_2} \times \cdots \times x_n^{f_n}}$$

وبشكل مبسط

$$\log G_P = \frac{\sum_{i=1}^{n} f_i \log x_i}{\sum_{i=1}^{n} f_i} = \sum_{i=1}^{n} fr \log x_i$$

- **حالة المتغير المتصل**

يعطى بالعلاقة التالية:

$$G_P = \sqrt[\sum f_i]{c_1^{f_1} \times c_2^{f_2} \times \cdots \times c_n^{f_n}}$$

وبشكل مبسط

$$logG_p = \frac{\sum_{i=1}^{n} f_i logc_i}{\sum_{i=1}^{n} f_i} = \sum_{i=1}^{n} f_r logc_i$$

4 – 5 – 2 المتوسط التوافقي: Moyen Harmonique (H)

هناك حالات أخرى وجد أن المتوسطات السابقة لا تصلح لوصفها جيدا، خصوصا في الظواهر التي تكون قيمها ناتج قسمة قيمة على قيمة أخرى مثل السرعة التي تساوي المسافة على الزمن. وبالتالي فأحسن طريقة لوصف هذا النمط من المتغيرات الكمية هو المتوسط التوافقي

4 – 5 – 2 – 1 المتوسط التوافقي البسيط

يعرف على أنه مقلوب المتوسط الحسابي لمقلوب القيم Xi ويحسب بالعلاقة:

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

4 – 5 – 2 – 2 المتوسط التوافقي المرجح

- **حالة المتغير المنفصل:**

يعبر عنه بالعلاقة: $H_p = \frac{\sum_{i=1}^{n} f_i}{\sum_{i=1}^{n} \frac{f_i}{x_i}}$

- **حالة المتغير المتصل:** يعبر عنه بالعلاقة:

$$H_p = \frac{\sum_{i=1}^{n} f_i}{\sum_{i=1}^{n} \frac{f_i}{c_i}}$$

4 – 5 – 3 المتوسط التربيعي: Moyen quadratique (MQ)

بالرغم من أن هذا المتوسط هو قليل الاستخدام في مجال العلوم الإجتماعية، إلا أن له الكثير من الاستخدامات في مجال العلوم الطبيعية.

4 – 5 – 3 – 1 المتوسط التربيعي البسيط

يعرف على أنه الجذر التربيعي للمتوسط الحسابي لمربعات القيم ويعطى بالعلاقة:

$$MQ_s = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}}$$

4 – 5 – 3 – 2 المتوسط التربيعي المرجح:

- **حالة المتغير المنفصل**

يعبر عنه بالعلاقة التالية:

$$MQ_p = \sqrt{\frac{\sum_{i=1}^{n} f_i x_i^2}{\sum_{i=1}^{n} f_i}} = \sqrt{\sum_{i=1}^{n} fr x_i^2}$$

- **حالة المتغير المتصل**

يعبر عنه بالعلاقة التالية:

$$MQ_p = \sqrt{\frac{\sum_{i=1}^{n} f_i c_i^2}{\sum_{i=1}^{n} f_i}} = \sqrt{\sum_{i=1}^{n} fr c_i^2}$$

ملاحظة:

* العلاقة بين $\overline{X}, G, H, MQ$ هي $H \le G \le \overline{X} \le MQ$

* العلاقة بين $\overline{X}, G, H$ هي $G = \sqrt{\overline{X} \times H}$

* العلاقة بين $\overline{X}, Me, Mo$ هي $\overline{X} - Mo = 3(\overline{X} - Me)$ وذلك بشرط أن يكون التوزيع قريب من التماثل.

4 – 6 تمارين تطبيقية حول الفصل الرابع

تمرين رقم (4–1).

تم تسجيل المعدلات التالية لـ 600 طالب في مدرسة عليا ما:

- 35% من الطلبة تحصلوا على نقاط تصل إلى 04،
- 40% من الطلبة تحصلوا على نقاط تصل إلى 08،
- 15% من الطلبة تحصلوا على نقاط تصل إلى 12،
- 10% من الطلبة تحصلوا على نقاط تصل إلى 16.

المطلوب:

1 – تكوين جدول تكراري لهذه المشاهدات تكون فئاته متساوية، مع العلم أن اكبر نقطة هي 16.

2 – حساب المتوسط الحسابي مع التعليق؟

3 – حساب الوسيط مع التعليق؟

4 – حساب المنوال بيانيا وجبريا مع التعليق؟

الحل.

1 – تكوين الجدول التكراري

يجب أن يكون الجدول التكراري الذي يشمل بيانات الطلبة كما يلي:

- تكون فئاته متساوية،
- الحد الأدنى للفئة الأولى هو 0 والحد الأعلى للفئة الأخيرة هو 16.

الفئات	Fr(%)	Fi
0–4	35	210
4–8	40	240
8–12	15	90
12–16	10	60
$\sum$	100	600

2 – حساب المتوسط الحسابي

الفئات	Fi	Ci	FiCi	Fcc
0−4	210	2	420	210
4−8	240	6	1440	450
8−12	90	10	900	540
12−16	60	14	840	600
$\sum$	600		3600	

المتوسط الحسابي يساوي:

$$\bar{X} = \frac{\sum_{i=1}^{n} f_i c_i}{\sum_{i=1}^{n} f_i} = \frac{3600}{600} = 6$$

متوسط معدل نقاط 600 طالب في هذه المدرسة العليا هو 20/6.

3 – حساب الوسيط

– رتبة الوسيط هي:

$$Rme = \frac{\sum_{i=1}^{n} f_i}{2} = \frac{600}{2} = 300$$

وبالتالي فإن الفئة الوسيطة هي الفئة الثانية 4−8.

– حساب الوسيط:

$$Me = L_o + \frac{\frac{\sum_{i=1}^{n} f_i}{2} - f_1}{f_2 - f_1} \times K = 4 + \frac{300 - 210}{540 - 210} \times 4 = 5.5$$

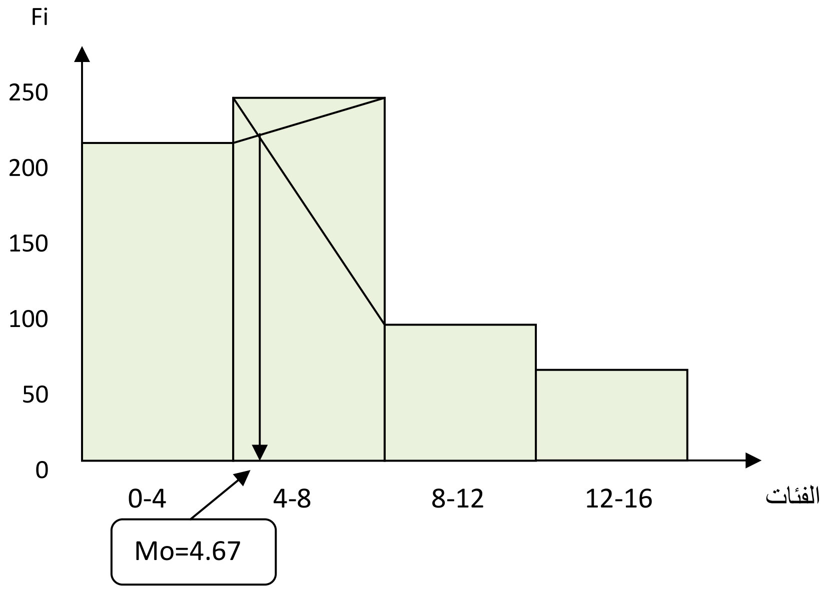
تدل النتيجة على أن 50% من الطلبة في هذه المدرسة العليا لهم معدل نقاط أقل من 20/5.5، والنصف الثاني لهم معدل نقاط أكبر من 20/5.5.

4 – حساب المنوال

– حساب المنوال جبريا

الفئة المنوالية هي الفئة التي تقابل أكبر تكرار، وحيث أن أكبر تكرار في المشاهدات هو 240 فإن الفئة المنوالية هي الفئة الثانية 4−8.

$$Mo = L_0 + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times K = 4 + \frac{240-210}{(240-210)+(240-90)} \times 4 = 4.67$$



تظهر النتيجة أن أغلبية (النسبة الكبيرة) طلبة المدرسة العليا قد تحصلوا على معدل نقاط قدره 20/4.67.

تمرين رقم (4–2).

ليكن لدينا التوزيع التكراري التالي:

الفئات	25–20	30–25	35–30	40–35	45–40	المجموع
التكرارات	7	13	21	6	3	50

أحسب:

المتوسط الحسابي بالطريقة المباشرة، طريقة الوسط الفرضي وطريقة الانحرافات المختصرة.

الوسيط، المنوال، المتوسط الهندسي، المتوسط التوافقي والمتوسط التربيعي.

الحل:

الفئات	Fi	Ci	FiCi	Wi	FiWi	Wi'	FiWi'	FCC	LogCi	FiLogCi	Fi/Ci	Ci^2	$FiCi^2$
20–25	7	22.5	157.5	−10	−70	−2	−14	7	1.352	9.465	0.31	506.3	3543
25–30	13	27.5	357.5	−5	−65	−1	−13	20	1.439	18.71	0.47	756.3	9831
30–35	21	32.5	682.5	0	0	0	0	41	1.512	31.75	0.65	1056	22181
35–40	6	37.5	225	5	30	1	6	47	1.574	9.444	0.16	1406	8437
40–45	3	42.5	127.5	10	30	2	6	50	1.628	4.885	0.07	1806	5418
Σ	50		1550		−75		−15				1.66		49412

1 – حساب الوسط الحسابي بالطريقة المباشرة

$$\overline{X}=\frac{\sum_{i=1}^{n}f_ic_i}{\sum_{i=1}^{n}f_i}=\frac{1550}{50}=31$$

2– حساب الوسط الحسابي بطريقة الوسط الفرضي

$$\overline{X}=\alpha+\frac{\sum_{i=1}^{n}f_iw_i}{\sum_{i=1}^{n}f_i}=32.5+\frac{-75}{50}=31$$

3– حساب الوسط الحسابي بطريقة الانحرافات المختصرة

$$\overline{X}=\alpha+\frac{\sum_{i=1}^{n}f_iw'_i}{\sum_{i=1}^{n}f_i}\times k=32.5+\frac{-15}{50}\times 5=31$$

4 – حساب قيمة الوسيط

$$Me=L_0+\frac{\frac{\sum fi}{2}-F_1}{F_2-F_1}\times K=30+\frac{25-20}{41-20}\times 5=31.2$$

5 – حساب قيمة المنوال

$$Mo=Lo+\frac{\Delta_1}{\Delta_1+\Delta_2}\times K=30+\frac{21-13}{(21-13)+(21-6)}\times 5=31.7$$

6 – حساب المتوسطات

$$MQ_p=\sqrt{\frac{49412.5}{50}}=31.4,H_p=\frac{50}{1.66}=30.1,G_p=10^{1.4851}=30.6$$

نلاحظ أن

$$H\leq G\leq \overline{X}\leq MQ \Leftrightarrow 30.1\leq 30.6\leq 31\leq 31.4$$

تمرين رقم (4–3)

يمثل الجدول الموالي توزيع الأجور الشهرية بالأورو لعمال في مؤسسة.

الأجور	900-800	1000-900	1050-1000	1150-1050	1300-1150
عدد العمال	42	49	75	19	16

المطلوب:

1 – أحسب متوسط الأجر في هذه المؤسسة؟ ماذا تستنتج؟

2 – أرسم التكرار المتجمع الصاعد، ثم استنتج قيم الوسيط والربيعيات؟

3 – أوجد قيمة الربيع الأول والثالث جبريا؟

4 – تم إضافة فئة سادسة جديدة في هذه السلسلة الإحصائية وهي 1300–1500. ما هو تكرار هذه الفئة إذا علمت أن المتوسط الحسابي للمؤسسة يبلغ 1200 أورو.

الحل

1 – حساب المتوسط الحسابي

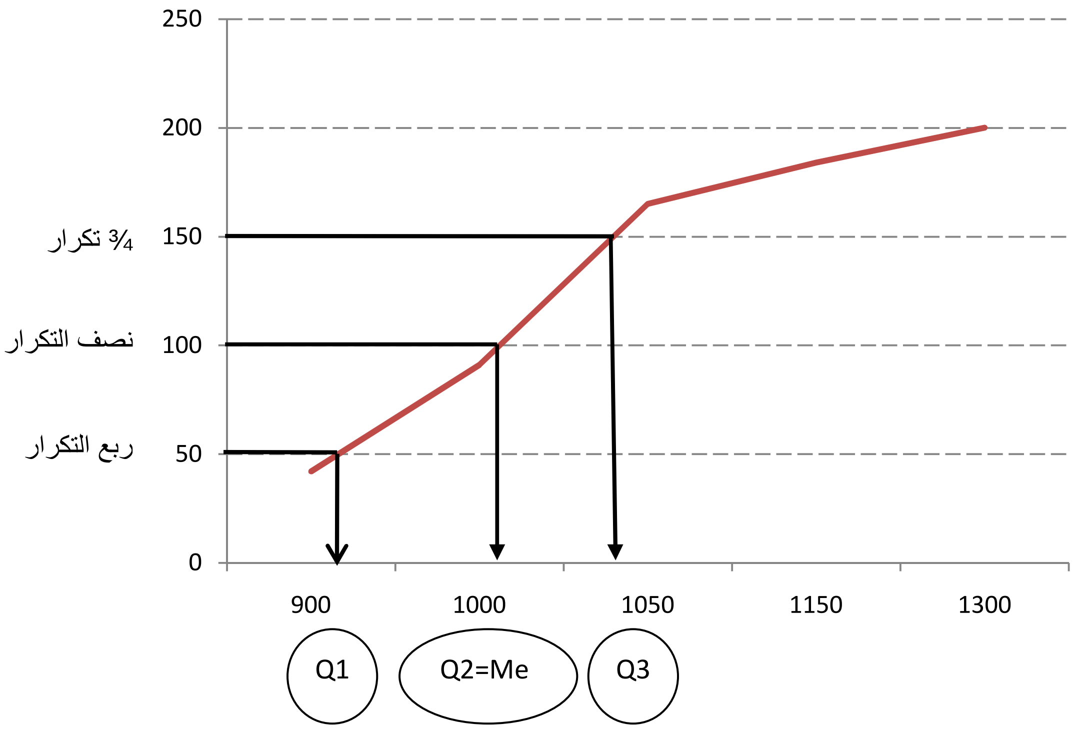
الفئات	Fi	Ci	FiCi	Fcc
800-900	42	850	35700	42
900-1000	49	950	46550	91
1000-1050	74	1025	75850	165
1050-1150	19	1100	20900	184
1150-1300	16	1225	19600	200
$\sum$	200		198600	

المتوسط الحسابي يساوي:

$$\bar{X} = \frac{\sum_{i=1}^{n} f_i c_i}{\sum_{i=1}^{n} f_i} = \frac{198600}{200} = 993$$

المتوسط الحسابي لأجور عمال هذه المؤسسة هو 993 أورو. نلاحظ أن المتوسط الحسابي له تمثيل ضعيف للنزعة المركزية لأجور المؤسسة، وذلك لأن أكثر من نصف العمال أجورهم أكبر من 1000 أورو.

2 – رسم التكرار المتجمع الصاعد



نقوم بتحديد رتبة الربيعيات وهي:

$$RQ_1 = \frac{\sum_{i=1}^{n} f_i}{4} = \frac{200}{4} = 50$$

$$RQ_2 = \frac{2\sum_{i=1}^{n} f_i}{4} = \frac{400}{4} = 100$$

$$RQ_3 = \frac{3\sum_{i=1}^{n} f_i}{4} = \frac{600}{4} = 150$$

نلاحظ من خلال الشكل أنه يصعب تحديد القيمة الدقيقة للربيعيات بدون استخدام أوراق ميليمترية للرسم، لهذا نلجأ إلى الطرق الجبرية في حسابها.

3 – حساب قيمة الربيع الأول والثالث

$$Q_1 = \mathrm{L}_0 + \frac{\frac{\sum_{i=1}^{n} f_i}{4} - f_1}{f_2 - f_1} \times \mathrm{k} = 900 + \frac{50-42}{91-42} \times 100 = 916.32$$

$$Q_3 = \mathrm{L}_0 + \frac{\frac{3\sum_{i=1}^{n} f_i}{4} - f_1}{f_2 - f_1} \times \mathrm{k} = 1000 + \frac{150-91}{165-91} \times 50 = 1039.86$$

4 – حساب تكرار الفئة الجديدة

$$\bar{X} = \frac{\sum_{i=1}^{n} f_i c_i}{\sum_{i=1}^{n} f_i} = \frac{\sum_{i=1}^{5} f_i c_i + f_6 C_6}{\sum_{i=1}^{5} f_i + f_6} = \frac{198600 + f_6 1400}{200 + f_6} = 1200$$

$$198600 + f_6 1400 = 1200(200 + f_6) \Rightarrow f_6 = \frac{41400}{200} = 207$$

تكرار الفئة السادسة هو 207 عامل.

الفصل الخامس: مقاييس التشتت Caractéristique de dispersion

5 – 1 تمهيد

قد تم التعرض في الفصل السابق إلى مقاييس النزعة المركزية، والتي تسمح لنا بالحصول على القيم المتوسطة للبيانات. إلا أن القيم المتوسطة وحدها لا تكفي لإعطاء صورة كاملة عن توزيع ظاهرة ما، والوسط وحده لا يعطي فكرة دقيقة عن مجموعة من القيم، فلا يبين طبيعتها ولا كيفية توزيعها. كما أن استخدام الوسط فقط لمقارنة عدة مجموعات لا يكفي لإظهار حقيقة المقارنة، فقد يتساوى متوسط مجموعتين بين تختلف المجموعتان عن بعضهما كل الاختلاف، فقد تكون مفردات إحدى المجموعتين متقاربة بعضها من بعض، ومفردات المجموعة الثانية متباعدة.

مثال رقم (5–1).

لنأخذ السلسلتين التاليتين:

السلسلة الأولى: 15، 16، 17، 18، 18، 18، 19، 19، 20، 20.

السلسلة الثانية: 3، 4، 6، 9، 18، 18، 19، 29، 30، 44.

المتوسط الحسابي للسلسة الأولى من البيانات متوسطها الحسابي يساوي 18 وهو يساوي المتوسط الحسابي لبيانات السلسلة الثانية. إذا اكتفينا بمقارنة الوسطين الحسابيين للمجموعتين فقد نستنتج أن السلستين متساويتين، غير أن نظرة على بياناتها تظهر أنهما مختلفتان في الواقع. حيث أن البيانات في السلسلة الأولى متقاربة إلى بعضها أو قريبة إلى مقياس النزعة المركزية المحسوب وهو المتوسط الحسابي، بينما بيانات السلسلة الثانية هي مشتتة (متباعدة) عن نفس قيمة الوسط الحسابي. ولهذا، فإننا نلجأ لما يعرف بمقاييس التشتت التي تسمح لنا بقياس تشتت البيانات عن مركزها.

تعريف:

تعرف مقاييس التشتت بأنها "الدرجة التي تنتشر بها البيانات حول قيمتها المتوسطة". أو بعبارة أخرى تظهر مدى اختلاف القيم المدروسة فيما بينها، أو مدى قربها أو بعدها من احد مقاييس النزعة المركزية.

5 – 2 مقاييس التشتت المطلق

5 – 2 – 1 المدى العام Etendue E

هو أبسط مقياس لتشتت البيانات وأسهلها حسابا، وهو عبارة عن الفرق بين أكبر قيمة وأقل قيمة في السلسلة الإحصائية. أي:

$$E = \boldsymbol{x_{max}} - \boldsymbol{x_{min}}$$

وكمثال على المدى نأخذ نقاط مقياس طلبة التي يمكن أن تكون محصورة بين 20/2 و20/18.

5 – 2 – 2 الانحراف المتوسط: $e_{\bar{x}}$ Ecart absolu moyen

يعرف على أنه "متوسط القيمة المطلقة لانحرافات القيم عن وسطها الحسابي". ونستخدم القيمة المطلقة لأن مجموع انحرافات القيم عن وسطها الحسابي يساوي 0.

5 – 2 – 2 – 1 حالة البيانات المفردة

$$e_{\bar{x}} = \frac{\sum_{i=1}^{n} \left| X_i - \bar{X} \right|}{n}$$

مثال رقم (5–2).

لديك البيانات التالية: 4، 3، 5، 9، 1، 10، 8، 12

أحسب قيمة الانحراف المتوسط؟

الحل:

حساب المتوسط الحسابي:

$$\bar{X} = \frac{\sum_{i=1}^{8} X_i}{n} = \frac{4 + \cdots + 12}{8} = 6.5$$

حساب الانحراف المتوسط:

$$e_{\bar{X}} = \frac{\sum_{i=1}^{n} |X_i - \bar{X}|}{n} = \frac{|4 - 6.5| + \cdots |12 - 6.5|}{8} = 3.25$$

5 – 2 – 2 – 2 حالة التوزيع التكراري

*** حالة المتغير الإحصائي المنفصل**

$$e_{\bar{x}} = \frac{\sum_{i=1}^{n} f_i \left| X_i - \overline{X} \right|}{\sum_{i=1}^{n} f_i} = \sum_{i=1}^{n} fr_i \times \left| X_i - \overline{X} \right|$$

مثال رقم (5–3).

باستخدام الجدول الأول في المثال رقم (4–2) أوجد قيمة الانحراف المتوسط.

$$e_{\bar{X}} = \frac{\sum_{i=1}^{n} f_i |x_i - \bar{X}|}{\sum_{i=1}^{n} f_i} = \frac{1|15 - 8.54| + \cdots + 2|10 - 8.54|}{11} = 2.23$$

*** حالة المتغير الإحصائي المتصل**

$$e_{\bar{x}} = \frac{\sum_{i=1}^{n} f_i \left| c_i - \overline{X} \right|}{\sum_{i=1}^{n} f_i} = \sum_{i=1}^{n} fr_i \times \left| c_i - \overline{X} \right|$$

5 – 2 – 3 الانحراف الوسيط (e_{Me}) Ecart absolu médiane

يعرف على أنه "متوسط القيمة المطلقة لانحرافات القيم عن وسيطها".

5 – 2 – 3 – 1 حالة البيانات المفردة:

نقوم بحساب قيمة الوسيط باستخدام العلاقة التالية:

$$e_{Me} = \frac{\sum_{i=1}^{n} |x_i - Me|}{n}$$

مثال رقم (5–4).

باستخدام بيانات المثال رقم (4–4) أوجد قيمة الانحراف الوسيط.

$$e_{Me} = \frac{\sum_{i=1}^{n} |x_i - Me|}{n} = \frac{|10 - 13| + \cdots + |16 - 13|}{9} = 1.56$$

5 – 2 – 3 – 2 حالة التوزيع التكراري

*** حالة المتغير الإحصائي المنفصل**

نقوم بحساب قيمة الوسيط باستخدام العلاقة التالية:

$$e_{Me} = \frac{\sum_{i=1}^{n} f_i |x_i - Me|}{\sum_{i=1}^{n} f_i} = \sum_{i=1}^{n} f_r |x_i - Me|$$

*** حالة المتغير الإحصائي المتصل**

نقوم بحساب قيمة الوسيط باستخدام العلاقة التالية:

$$e_{Me} = \frac{\sum_{i=1}^{n} f_i |c_i - Me|}{\sum_{i=1}^{n} f_i} = \sum_{i=1}^{n} f_r |c_i - Me|$$

مثال رقم (5–6).

باستخدام بيانات المثال رقم (4–6) أوجد قيمة الانحراف الوسيط.

قيمة الوسيط هي 67.42.

الفئات	Fi	Ci	$\lvert c_i - Me\rvert$	$f_i\lvert c_i - Me\rvert$
59.5–62.5	5	61	6,42	32,1
62.5–65.5	18	64	3,42	61,56
65.5–68.5	42	67	0,42	17,64
68.5–71.5	27	70	2,58	69,66
71.5–74.5	8	73	5,58	44,64
$\sum$	100			225,6

قيمة الانحراف الوسيط هي:

$$e_{Me} = \frac{\sum_{i=1}^{n} f_i |c_i - Me|}{\sum_{i=1}^{n} f_i} = \frac{225.6}{100} = 2.256$$

5 – 2 – 4 الانحراف الربيعي: Ecart qunatile (e_Q)

يعرف على أنه نصف المدى بين الربيع الأول Q_1 والربيع الثالث Q_3.

إذن يحسب الانحراف الربيعي بالعلاقة:

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

مثال رقم (5−7).

باستخدام نتائج المثال رقم (4−12) اوجد قيمة الانحراف الربيعي.

$$e_Q = \frac{Q_3 - Q_1}{2} = \frac{14.5 - 5.5}{2} = 10$$

ملاحظة: تتجلى أهمية الانحراف الربيعي في كونه مقياس للتشتت قابل للحساب من الجداول التكرارية المفتوحة.

5 − 2 − 5 الانحراف المعياري Ecart−Type S

يعرف الانحراف المعياري بأنه الجذر التربيعي لمتوسط مجموع مربعات انحرافات القيم عن وسطها الحسابي. ويعتبر الانحراف المعياري واحد من أهم المقاييس الإحصائية لاستخدامه في العديد من النظريات الإحصائية. ويرتبط الانحراف المعياري بمفهوم أخر وهو التباين Variance الذي نرمز له بالرمز S^2 أو σ أو الرمز $V(x)$ ويساوي مربع الانحراف المعياري.

5 − 2 − 5 − 1 حالة البيانات المفردة

لإيجاد قيمة الانحراف المعياري في حالة البيانات المفردة نستخدم القانون التالي:

$$S_x = \sqrt{S_x^2} = \sqrt{V_x} = \sqrt{\frac{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}{n-1}}$$

يمكن حسابه كذلك بالعلاقة التالية

$$S_x^2 = \frac{\sum X_i^2 - n\overline{X}^2}{n-1} = \frac{1}{n-1}\left[\sum X_i^2 - \frac{\left(\sum X_i\right)^2}{n}\right]$$

مثال رقم (5−8).

باستخدام بيانات المثال رقم (5−2) أوجد قيمة الانحراف المعياري والتباين.

$$S_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{X})^2}{n-1}} = \sqrt{\frac{(4-6.5)^2 + \cdots + (3-6.5)^2}{8-1}} \cong 3.82$$

أما التباين فهو يساوي مربع الانحراف المعياري وبالتالي:

$$V(x) = S^2 = (3.82)^2 = 14.59$$

5 – 2 – 5 – 2 حالة التوزيع التكراري

*** حالة المتغير الإحصائي المنفصل**

لإيجاد قيمة الانحراف المعياري في حالة المتغير الإحصائي المنفصل نستخدم القانون التالي:

$$S_x = \sqrt{\frac{\sum_{i=1}^{n} f_i (x_i - \bar{X})^2}{\sum_{i=1}^{n} f_i - 1}} =$$

ويمكن كذلك حساب الانحراف المعياري باستخدام الصيغة المختصرة التالية:

$$S_x = \sqrt{\frac{\sum_{i=1}^{n} f_i (x_i)^2 - n\bar{X}^2}{\sum_{i=1}^{n} f_i - 1}}$$

*** حالة المتغير الإحصائي المتصل**

لإيجاد قيمة الانحراف المعياري في حالة المتغير الإحصائي المتصل نستخدم القانون التالي:

$$S_x = \sqrt{\frac{\sum_{i=1}^{n} f_i (C_i - \bar{X})^2}{\sum_{i=1}^{n} f_i - 1}}$$

أما الصيغة المختصرة فهي:

$$S_x = \sqrt{\frac{\sum_{i=1}^{n} f_i (C_i)^2 - n\bar{X}^2}{\sum_{i=1}^{n} f_i - 1}}$$

مثال رقم (5–9).

أوجد قيمة الانحراف المعياري بطريقتين، والتباين باستخدام بيانات المثال رقم (4–3).

الفئات	Fi	ci	$c_i-\bar{X}$	$(c_i-\bar{X})^2$	$f_i(c_i-\bar{X})^2$	Ci^2	$FiCi^2$
59.5–62.5	5	61	−6,45	41,6025	208,0125	3721	18605
62.5–65.5	18	64	−3,45	11,9025	214,245	4096	73728
65.5–68.5	42	67	−0,45	0,2025	8,505	4489	188538
68.5–71.5	27	70	2,55	6,5025	175,5675	4900	132300
71.5–74.5	8	73	5,55	30,8025	246,42	5329	42632
$\sum$	100	–			852.75		455803

– حساب قيمة الانحراف المعياري (المتوسط الحسابي يساوي 67.45)

$$S_x=\sqrt{\frac{\sum_{i=1}^{n} f_i(C_i-\bar{X})^2}{\sum_{i=1}^{n} f_i-1}}=\sqrt{\frac{852.75}{100-1}}=2.935$$

ويمكن كذلك حساب الانحراف المعياري باستخدام الصيغة المختصرة التالية:

$$S_x=\sqrt{\frac{\sum_{i=1}^{n} f_i(C_i)^2-n\bar{X}^2}{\sum_{i=1}^{n} f_i-1}}=\sqrt{\frac{455803-100(67.45)^2}{100-1}}=2.935$$

– حساب التباين

$$V(x) = S^2 = (2.935)^2 = 8.614$$

ملاحظات:

– يعبر S_x عن الانحراف المعياري للعينة، ونعبر عن الانحراف المعياري للمجتمع بالرمز δx والذي يساوي في حالة البيانات المفردة:

$$\sigma_x=\sqrt{\frac{\sum_{i=0}^{n}(x_i-\bar{X})^2}{n}}$$

أما في حالة التوزيع التكراري فإننا نقسم مربع انحرافات القيم على مجموع التكرارات بدون أن ننقص 1.

مثال رقم (5–10).

أوجد المتوسط الحسابي لبيانات المثال رقم (5–9) باعتبارها مجتمعا.

$$S_x = \sqrt{\frac{\sum_{i=1}^{n} f_i (C_i - \bar{X})^2}{\sum_{i=1}^{n} f_i - 1}} = \sqrt{\frac{852.75}{100}} = 8.92$$

– هناك علاقة تجريبية تربط بين الانحراف الربيعي والانحراف المعياري والانحراف المتوسط في حالة التوزيع القريب من التماثل وهي:

$$e_Q = \frac{2}{3} S_x \qquad , e_{\bar{X}} = \frac{4}{5} S_x$$

مثال رقم (5–11).

باستخدام بيانات المثال رقم (4–3) تحقق من العلاقة التجريبية السابقة.

– المتوسط الحسابي يساوي: 67.45

– الانحراف المعياري يساوي: 8.614

– حساب الانحراف المتوسط

الفئات	Fi	ci	$c_i - \bar{X}$	$\lvert c_i - \bar{X} \rvert$	Fcc
59.5–62.5	5	61	−6,45	6,45	5
62.5–65.5	18	64	−3,45	3,45	23
65.5–68.5	42	67	−0,45	0,45	65
68.5–71.5	27	70	2,55	2,55	82
71.5–74.5	8	73	5,55	5,55	100
$\sum$	100	–		226.5	

$$e_{\bar{X}} = \frac{\sum_{i=1}^{n} f_i \lvert x_i - \bar{X} \rvert}{\sum_{i=1}^{n} f_i} = \frac{226.5}{100} = 2.256$$

– حساب الانحراف الربيعي

رتبة الربيع الأول والثالث هي:

$$Rme_1 = \frac{1(100)}{4} = 25$$

$$Rme_3 = \frac{3(100)}{4} = 75$$

الربيع الأول والثالث يساويان:

$$Q_1 = \mathrm{L}_0 + \frac{\frac{\sum_{i=1}^{n} f_i}{4} - f_1}{f_2 - f_1} \times \mathrm{k} = 65.5 + \frac{25-23}{65-23} \times 3 = 65.643$$

$$Q_3 = \mathrm{L}_0 + \frac{\frac{3\sum_{i=1}^{n} f_i}{4} - f_1}{f_2 - f_1} \times \mathrm{k} = 68.5 + \frac{75-65}{82-65} \times 3 = 70.264$$

الانحراف الربيعي يساوي:

$$e_{\mathrm{Q}} = \frac{\mathrm{Q}_3 - \mathrm{Q}_1}{2} = \frac{70.264 - 65.643}{2} = 2.31$$

- **العلاقة التجريبية الأولى**

$$e_{\mathrm{Q}} \cong \frac{2}{3}\mathrm{S}_{\mathrm{x}} \Rightarrow 2.31 \cong \frac{2}{3}2.935 \Rightarrow 2.31 \cong 1.95$$

$$e_{\overline{\mathrm{X}}} \cong \frac{4}{5}\mathrm{S}_{\mathrm{x}} \Rightarrow 2.256 \cong \frac{4}{5}2.935 \Rightarrow 2.31 \cong 2.348$$

ملاحظة: كلما كان التوزيع أقرب إلى التماثل كلما زاد تحقق العلاقة التجريبية السابقة.

5 – 3 مقاييس التشتت النسبي

يقيس كل مقياس من مقاييس التشتت التي استعرضناها بنفس وحدات قياس الظاهرة موضع الدراسة. مثلا نقيس التشتت بين أوزان طلبة في قسمين مختلفين، أو نقيس التشتت بين إنتاج مصنع في سنتين مختلفتين. وكلا الظاهرتين لها نفس المقياس، الأولى بالكلغ والثانية بالطن مثلا. أما عند مقارنة التشتت لمفردات ظاهرتين مختلفتين أو لمفردات ظاهرة واحدة بمستويات مختلفة للقياس، يجب أن نحول مقياس التشتت المطلق للمفردات إلى مقياس نسبي حتى يمكن التخلص من تأثير اختلاف وحدات القياس وكذلك التخلص من تأثير اختلاف المتوسط لكل منها.

ولإجراء ذلك نستخدم مقياس يسمى معامل الاختلاف Cœfficient de Variation ويرمز له بالرمز CV وله صورتان:

الصورة الأولى:

تعتمد في حسابها على الانحراف المعياري والمتوسط الحسابي ويرمز له بالرمز CV_1 حيث

لاحظ أن معامل الاختلاف لا يعتمد على الوحدات المستخدمة. لهذا السبب يكون مناسبا لمقارنة توزيعات مختلفة للوحدات.

من عيوبه هو صعوبة حسابه عندما يكون المتوسط الحسابي قريب من الصفر ولا يستخدم إلا في حالة التوزيعات المقفلة.

مثال رقم (5−12).

أحسب معامل الاختلاف الأول لبيانات المثال رقم (5−9).

$$CV_1 = \frac{S_x}{\bar{X}} \times 100 = \frac{2.935}{67.45} \times 100 = 4.35\%$$

الصورة الثانية:

تعتمد في حسابها على الانحراف الربيعي والوسيط ويرمز لها بـ CV_2 حيث

$$CV_2 = \frac{Q_3 - Q_1}{M_e} \times 100$$

ويستخدم هذا المقياس في حالة التوزيعات المفتوحة.

مثال رقم (5−31).

أحسب معامل الاختلاف الأول لبيانات المثال رقم (4−13).

$$CV_1 = \frac{Q_3 - Q_1}{Me} \times 100 = \frac{4 - 1}{2.5} \times 100 = 120\%$$

5 – 4 تمارين حول الفصل الخامس

تمرين رقم (5−1)

باستخدام بيانات المثال رقم (2−11)، أوجد مقاييس التشتت لأوزان الـ 100 طالب.

الحل: $\overline{X} = 67.45kg, Me = 67.42kg$

Classes	Fi	ci	$\lvert ci - \bar{x}\rvert$	$Fi\lvert ci - \bar{x}\rvert$	$\lvert ci - Me\rvert$	$Fi\lvert ci - Me\rvert$	Fcc	$(ci - \bar{x})^2$	$Fi(ci - \bar{x})^2$
59.5−62.5	5	61	6.45	32.25	6.42	32.1	5	41.6025	208.0125
62.5−65.5	18	64	3.45	62.10	3.42	61.56	23	11.9025	214.2450
65.5−68.5	42	67	0.45	18.90	0.42	18.9	**65**	0.2025	8.5050
68.5−71.5	27	70	2.55	68.85	2.58	69.66	**92**	6.5025	175.5657
71.5−74.5	8	73	5.55	44.40	5.58	44.64	100	30.8025	246.4200
المجموع	100	−		226.5		226.86	−		852.7500

حساب المتوسط الحسابي:

$$\overline{X} = 67.45kg, Me = 67.42kg$$

1 – حساب الانحراف المتوسط:

$$e_{\bar{x}} = \frac{\sum_{i=1}^{n} f_i \lvert c_i - \overline{X}\rvert}{\sum_{i=1}^{n} f_i} = \frac{226.5}{100} = 2.265kg$$

2 – حساب الانحراف الوسيط:

$$e_{Me} = \frac{\sum_{i=1}^{n} f_i \lvert c_i - Me\rvert}{\sum_{i=1}^{n} f_i} = \frac{226.86}{100} = 2.2686kg$$

3 – حساب الانحراف الربيعي:

$$Q_1 = L_0 + \frac{\frac{\sum fi}{4} - F_1}{F_2 - F_1} \times K = 65.5 + \frac{25-23}{65-23} \times 3 = 65.64kg$$

$$e_Q = \frac{Q_3 - Q_1}{2} = \frac{69.61 - 65.64}{2} = 1.985kg$$

$$Q_3 = L_0 + \frac{\frac{3\sum fi}{4} - F_1}{F_2 - F_1} \times K = 68.5 + \frac{75-65}{92-65} \times 3 = 69.61kg$$

4 – حساب الانحراف المعياري:

$$S_x = \sqrt{\frac{\sum fi(c_i - \overline{X})^2}{(\sum fi) - 1}} = \sqrt{\frac{852.7500}{99}} = 2.934kg$$

5 – حساب التباين:

$$S_x^2 = (2.934)^2 = 8.613kg$$

6 – معامل الاختلاف الأول:

$$CV_1 = \frac{s_x}{\overline{X}} \times 100 = \frac{2.934}{67.45} \times 100 = 4.34\%$$

7 – معامل الاختلاف الثاني:

$$CV_2 = \frac{Q_3 - Q_1}{M_e} \times 100 = \frac{69.61 - 65.64}{67.42} \times 100 = 5.88\%$$

تمرين رقم (5–2).

ناقش العلاقة التجريبية التي تربط بين الانحراف الربيعي والانحراف المعياري والانحراف المتوسط في حالة التوزيع السابق:

$\frac{e_{\bar{x}}}{S_x} = \frac{2.265}{2.934} = 0.77 \cong \frac{4}{5}$ ، $\frac{e_Q}{S_x} = \frac{1.985}{2.985} = 0.66 \cong \frac{2}{3}$ وبهذا فالعلاقة التجريبية صحيحة.

تمرين رقم (5−3).

أثبت أن التباين في حالة المجتمع يساوي $\sigma_x^2=\frac{\sum x_i^2}{n}-\bar{x}^2$.

الحل:

$$\sigma_x^2=\frac{\sum(x_i-\bar{x})^2}{n}=\frac{\sum(x_i^2-2x_i\bar{x}+\bar{x}^2)}{n}=\frac{\sum x_i^2-2\bar{x}\sum x_i+n\bar{x}^2}{n}=\frac{\sum x_i^2}{n}-2\bar{x}\frac{\sum x_i}{n}+\frac{n\bar{x}^2}{n}$$

$$=\frac{\sum x_i^2}{n}-2\bar{x}\bar{x}+\bar{x}^2=\frac{\sum x_i^2}{n}-2\bar{x}^2+\bar{x}^2=\frac{\sum x_i^2}{n}-\bar{x}^2$$

تمرين رقم (5−4).

ما هي مقاييس النزعة المركزية ومقاييس التشتت التي يمكن حسابها من التوزيع التكراري التالي:

الفئات	10 − 15	15 − 25	25 − 30	30 − 40	أكبر من 40
التكرارات	2	12	25	15	6

الحل:

1 − نحسب الوسيط بالعلاقة

$$Me=L_0+\frac{\frac{\sum fi}{2}-F_1}{F_2-F_1}\times K=25+\frac{30-14}{39-14}\times 5=28.2$$

2 − حساب الربيع الأول والثالث:

$$Q_1=25+\frac{15-14}{39-14}\times 5=25.2, Q_3=30+\frac{45-39}{54-39}\times 10=34$$

3 − حساب الانحراف الربيعي:

$$e_Q=\frac{Q_3-Q_1}{2}=\frac{34-25.2}{2}=4.4$$

4 − حساب معامل الاختلاف الثاني:

$$CV_2=31.20\%$$

الفصل السادس: مقاييس الشكل Caractéristiques de forme

6 – 1 تمهيد

على غرار مقاييس النزعة المركزية ومقاييس التشتت، التي وكما رأينا سابقا تساعدنا على معرفة توزيع وانتشار المعطيات حول قيمة مركزية. سنتطرق إلى مقاييس تساعدنا على معرفة تمركز وشكل التوزيع الإحصائي (الالتواء والتفرطح) بدون اللجوء إلى تمثيل البيانات. تدعى هذه المقاييس بمقاييس الشكل، والتي تعتمد في حسابها على العزوم.

6 – 2 العزوم Les moments

يعرف عزم أي قوة، بأنه مقدار العمل الذي تحدثه. ويتوقف هذا العمل على القوة نفسها والمسافة بين هذه القوة والنقطة التي عندها يحدث أثر هذه القوة. وبالنسبة للتوزيع التكراري تكون تكرارات تكرارات التوزيع هي القوة المؤثرة عليه. أي تكرار يقاس بحاصل ضرب التكرار في انحرافه عن نقطة تمركز التوزيع، التي يعبر عنها عادة بالمتوسط الحسابي. والعزوم نوعان: عزوم مركزية ولا مركزية.

6 – 2 – 1 العزوم اللامركزية: Moments simples (m)

يعرف العزم اللامركزي من الرتبة r $(r \in N)$، لمتغير إحصائي بالعلاقة التالية:

6 – 2 – 1 – 1 حالة البيانات غير المبوبة

$$m_r = \frac{\sum x_i^r}{n}$$

والعزوم اللامركزية الثلاثة الأولى هي:

$$m_1 = \frac{\sum x_i}{n} = \bar{x}; m_2 = \frac{\sum x_i^2}{n}; m_3 = \frac{\sum x_i^3}{n}$$

مثال رقم (6–1).

لديك البيانات التالية 12، 17، 20، 7، 2، 16، 10. أحسب العزوم اللامركزية الأربعة الأولى.

Xi	Xi^2	Xi^3	Xi^4
12	144	1728	20736
17	289	4913	83521
20	400	8000	160000
7	49	343	2401
2	4	8	16
16	256	4096	65536
10	100	1000	10000
84	1242	20088	342210

حساب العزوم اللامركزية الأربعة الأولى:

$$m_1 = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{84}{7} = 12$$

$$m_2 = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{1242}{7} = 177.43$$

$$m_3 = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{20088}{7} = 2869.71$$

$$m_4 = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{342210}{7} = 48887.14$$

6 – 2 – 1 – 2 حالة البيانات المبوبة

– متغير منفصل

$$m_r = \frac{\sum f_i x_i^r}{\sum f_i}$$

العزوم اللامركزية الأولى والثانية هي:

$$m_1 = \frac{\sum f_i x_i}{\sum f_i} = \bar{x}; m_2 = \frac{\sum f_i x_i^2}{\sum f_i} = \sum fr_i x_i^2$$

– متغير متصل

$$m_r = \frac{\sum f_i c_i^r}{\sum f_i}$$

مثال رقم (6−2).

لديك التوزيع التالي.

الفئات	4−2	8−4	12−8	16−12
التكرارات	11	23	6	10

أوجد العزوم اللامركزية الأربعة الأولى.

الفئات	Fi	Ci	FiCi	Ci^2	$FiCi^2$	Ci^3	$FiCi^3$	Ci^4	$FiCi^4$
2−4	11	3	33	9	99	27	297	81	891
4−8	23	6	138	36	828	216	4968	1296	29808
8−12	6	10	60	100	600	1000	6000	10000	60000
12−16	10	14	140	196	1960	2744	27440	38416	384160
Σ	50		371		3487		38705		474859

حساب العزوم اللامركزية الأربعة الأولى:

$$m_1 = \frac{\sum_{\mathrm{i}=1}^{\mathrm{n}} f_i c_{\mathrm{i}}}{\sum_{\mathrm{i}=1}^{\mathrm{n}} f_i c_{\mathrm{i}}} = \frac{371}{50} = 7.42$$

$$m_2 = \frac{\sum_{\mathrm{i}=1}^{\mathrm{n}} f_i {c_{\mathrm{i}}}^2}{\sum_{\mathrm{i}=1}^{\mathrm{n}} f_i c_{\mathrm{i}}} = \frac{3487}{50} = 69.74$$

$$m_3 = \frac{\sum_{\mathrm{i}=1}^{\mathrm{n}} f_i {c_{\mathrm{i}}}^3}{\sum_{\mathrm{i}=1}^{\mathrm{n}} f_i c_{\mathrm{i}}} = \frac{38705}{50} = 774.1$$

$$m_4 = \frac{\sum_{\mathrm{i}=1}^{\mathrm{n}} f_i {c_{\mathrm{i}}}^4}{\sum_{\mathrm{i}=1}^{\mathrm{n}} f_i c_{\mathrm{i}}} = \frac{342210}{7} = 9497.2$$

العزوم اللامركزية الأولى والثانية هي:

$$m_1 = \frac{\sum f_i c_i}{\sum f_i} = \bar{x}; m_2 = \frac{\sum f_i c_i^2}{\sum f_i} = \sum fr_i c_i^2$$

6 − 2 − 2 العزوم المركزية: Moments centrés (μ)

يعرف العزم المركزي من الرتبة r $(r \in N)$، لمتغير إحصائي بالعلاقة التالية:

6 – 2 – 2 – 1 حالة البيانات غير المبوبة

$$\mu_r = \frac{\sum (x_i - \bar{x})^r}{n}$$

العزوم المركزية الثلاثة الأولى هي:

$$\mu_1 = \frac{\sum (x_i - \bar{x})}{n} = 0; \mu_2 = \frac{\sum (x_i - \bar{x})^2}{n} = \sigma_x^2; \mu_3 = \frac{\sum (x_i - \bar{x})^3}{n}$$

مثال رقم (6–3).

أوجد قيمة العزوم المركزية الأربعة للقيم التالية: 2، 9، 6، 3.

Xi	$(x_i - \bar{X})$	$(x_i - \bar{X})^2$	$(x_i - \bar{X})^3$	$(x_i - \bar{X})^4$
2	−3	9	−27	81
3	−2	4	−8	16
6	1	1	01	1
9	4	16	64	256
20	0	30	30	354

العزوم المركزية الأربعة الأولى هي: (المتوسط الحسابي يساوي 5)

$$\mu_1 = \frac{\sum(x_i - \bar{X})}{n} = \frac{0}{4} = 0$$

$$\mu_2 = \frac{\sum(x_i - \bar{X})^2}{n} = \frac{30}{4} = 7.5$$

$$\mu_3 = \frac{\sum(x_i - \bar{X})^3}{n} = \frac{30}{4} = 7.5$$

$$\mu_4 = \frac{\sum(x_i - \bar{X})^4}{n} = \frac{354}{4} = 88.5$$

6 – 2 – 2 – 2 حالة البيانات المبوبة

– متغير منفصل

$$\mu_r = \frac{\sum f_i (x_i - \bar{x})^r}{\sum f_i}$$

العزوم المركزية الأولى والثانية هي:

$$\mu_1 = \frac{\sum f_i (x_i - \bar{x})}{\sum f_i} = 0; \mu_2 = \frac{\sum f_i (x_i - \bar{x})^2}{\sum f_i} = \sigma_x^2$$

– متغير متصل:

$$\mu_r = \frac{\sum f_i (c_i - \bar{x})^r}{\sum f_i}$$

العزوم المركزية الأولى والثانية هي:

$$\mu_1 = \frac{\sum f_i (c_i - \bar{x})}{\sum f_i} = 0; \mu_2 = \frac{\sum f_i (c_i - \bar{x})^2}{\sum f_i} = \sigma_x^2$$

6 – 2 – 3 العلاقة بين العزوم اللامركزية والعزوم المركزية

توجد علاقة بين العزوم اللامركزية والعزوم المركزية يمكن لنا أن نستخدم التوزيع الثنائي لنيوتن ومثلث باسكال لإيضاحها على النحو التالي:

لدينا:

$$(x+a)^r = x^r + C_r^1 a x^{r-1} + C_r^2 a^2 x^{r-2} + \cdots + C_r^p a^p x^{r-p} + a^r$$
$$(x-a)^r = x^r + C_r^1 x^{r-1}(-a) + \cdots + C_r^p x^{r-p}(-a)^p + (-a)^r$$

مثلث باسكال

r \ P	0	1	2	3	4	5	…
0	1						
1	1	1					
2	1	2	1				
3	1	3	3	1			
4	1	4	6	4	1		
5	1	5	10	10	5	1	
…							

من أجل حساب العزم المركزي الثاني مثلا:

$$(x_i - \bar{X})^2 = (x_i^2 - 2x_i\bar{X} + \bar{X}^2)$$

$$\Rightarrow \frac{\sum_{i=1}^{n}(x_i - \bar{X})^2}{n} = \frac{\sum_{i=1}^{n}(x_i^2 - 2x_i\bar{X} + \bar{X}^2)}{n}$$

$$\Rightarrow \frac{\sum_{i=1}^{n}(x_i - \bar{X})^2}{n} = \frac{\sum_{i=1}^{n} x_i^2 - 2\bar{X}\sum_{i=1}^{n} x_i + n\bar{X}^2}{n}$$

$$\Rightarrow \mu_2 = \frac{\sum_{i=1}^{n} x_i^2}{n} - 2\bar{X}^2 + \bar{X}^2$$

$$\Rightarrow \mu_2 = m_2 - m_1^2$$

بالمثل، تتحقق العلاقات التالية بين العزوم:

$$\mu_2 = m_2 - m_1^2$$

$$\mu_3 = m_3 - 3m_2m_1 + 2m_1^3$$

$$\mu_4 = m_4 - 4m_3m_1 + 6m_2m_1^2 - 3m_1^4$$

مثال رقم (6−4).

برهن أن:

$$\mu_3 = m_3 - 3m_2m_1 + 2m_1^3$$

البرهان.

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

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

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

$$= \frac{\sum_{i=1}^{n}(x_i^3 - 2x_i^2\bar{X} + x_i\bar{X}^2 - x_i^2\bar{X} + 2x_i\bar{X}^2 - \bar{X}^3)}{n}$$

$$= \frac{\sum_{i=1}^{n}(x_i^3 - 3x_i^2\bar{X} + 3x_i\bar{X}^2 - \bar{X}^3)}{n}$$

$$= \frac{\sum_{i=1}^{n}\left(x_i^3\right) - 3\bar{X}\sum_{i=1}^{n}\left(x_i^2\right) + 3\bar{X}^2\sum_{i=1}^{n}(x_i) - n\bar{X}^3}{n}$$

$$= \frac{\sum_{i=1}^{n}\left(x_i^3\right)}{n} - \frac{3\bar{X}\sum_{i=1}^{n}\left(x_i^2\right)}{n} + \frac{3\bar{X}^2\sum_{i=1}^{n}\left(x_i\right)}{n} - \frac{n\bar{X}^3}{n}$$

$$= \frac{\sum_{i=1}^{n}\left(x_i^3\right)}{n} - 3\bar{X}\frac{\sum_{i=1}^{n}\left(x_i^2\right)}{n} + 3\bar{X}^2\frac{\sum_{i=1}^{n}\left(x_i\right)}{n} - \bar{X}^3$$

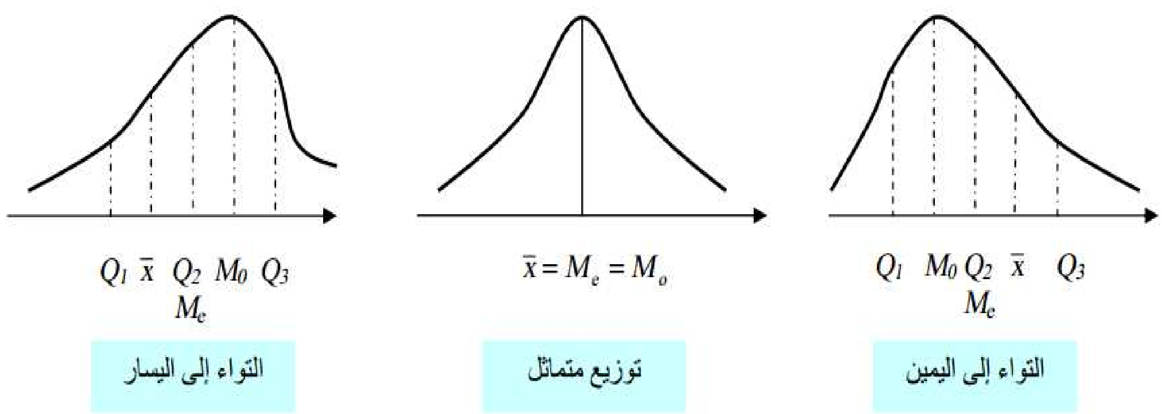
$$= \frac{\sum_{i=1}^{n}\left(x_i^3\right)}{n} - 3\bar{X}\frac{\sum_{i=1}^{n}\left(x_i^2\right)}{n} + 3\bar{X}^2\bar{X} - \bar{X}^3$$

$$= \frac{\sum_{i=1}^{n}\left(x_i^3\right)}{n} - 3\bar{X}\frac{\sum_{i=1}^{n}\left(x_i^2\right)}{n} + 2\bar{X}^3$$

$$= \mathrm{m}_3 - 3\mathrm{m}_2\mathrm{m}_1 + 2\mathrm{m}_1^3$$

6 – 3 الالتواء Oblique

الالتواء هو درجة تماثل (Symétrie) أو البعد عن التماثل لتوزيع. إذا كان المنحنى التكراري لتوزيع له ذيل أكبر إلى يمين مركز النهاية العظمى عنه إلى يسارها يسمى التوزيع بأنه ملتو إلى اليمين أو موجب الالتواء (Oblique à droite). أما إذا كان العكس صحيح فيقال أنه ملتو إلى اليسار أو سالب الالتواء (Oblique à gauche).



- من خصائص التوزيع المتماثل (المتناظر) أن الوسط الحسابي يكون مساويا لقيمة الوسيط والمنوال، ويكون الربيع الأول والثالث على نفس البعد بالنسبة للوسيط.
- أما في التوزيعات الملتوية فإن الوسط الحسابي يقع على نفس جانب الوسيط وذلك على نفس جانب الطرف الأطول (إلى اليمين أو اليسار).

يمكن معرفة وجود التواء و درجة حدته بإحدى المقاييس الإحصائية التالية:

6 – 3 – 1 معامل فيشر للالتواء Coefficient de Fisher

يمكن حساب معامل فيشر للالتواء بالعلاقة التالية:

$$\gamma_F = \frac{\mu_3}{\sigma^3}$$

مثال رقم (6–5).

أوجد معامل فيشر للالتواء من البيانات الواردة في المثال رقم (6–3).

نحسب قيمة الانحراف المعياري والتي تساوي:

$$\sigma_x^2 = \mu_2 = 7.5 \Rightarrow \sigma_x = \sqrt{7.5} = 2.739$$

$$\gamma_F = \frac{\mu_3}{\sigma^3} = \frac{7.5}{2.379^3} = 0.557$$

6 – 3 – 2 معامل بيرسون للالتواء Coefficient de Pearson

يمكن أخذ الفرق (الوسط – المنوال) كمقياس للالتواء. وهذا المقياس يمكن تخليصه من الوحدات بقسمته على الانحراف المعياري، وهذا ما يؤدي إلى التعريف التالي:

6 – 3 – 2 – 1 معامل بيرسون للالتواء الأول

$$\gamma_1 = \frac{\bar{X} - M0}{\sigma}$$

يمكن استخدام الوسيط بدل المنوال في العلاقة السابقة وذلك باستخدام العلاقة التجريبية التي تربط بين المتوسطات $\bar{X} - Mo = 3(\bar{X} - Me)$.

6 – 3 – 2 – 2 معامل بيرسون للالتواء الثاني

يمكن كذلك قياس الالتواء باستخدام معامل بيرسون التالي:

$$\gamma_2 = \frac{\mu_3^2}{\mu_2^3}$$

6 – 3 – 2 – 3 معامل يول للالتواء Coefficient de Yule

يعتمد حساب معامل يول للالتواء على الربيعيات، وبالتالي **يعتبر أداة للحساب في الجداول المفتوحة** التي لا يمكن فيها حساب المعاملات السابقة لأنها تعتمد على المتوسط الحسابي، و منه:

معامل يول للالتواء:

$$\gamma_3 = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_1)}$$

باستخدام إحدى المعاملات السابقة نجد:

* إذا كان معامل الالتواء موجبا، يكون التوزيع ملتويا إلى اليمين،
* إذا كان معامل الالتواء سالبا، يكون التوزيع ملتويا إلى اليسار،
* إذا كان معامل الالتواء معدوما، يكون التوزيع متماثل، وكلما ابتعدت عن الصفر زادت حدة الالتواء يمينا أو يسارا.

مثال رقم (6–6).

أوجد مقاييس الشكل للتوزيع التكراري التالي.

الفئات	Fi	Ci	Fcc	FiCi	$FiCi^2$	$FiCi^3$
[50 – 60[	8	55	8	440	24200	1331000
[60 – 70[	10	65	18	650	42250	2746250
[70 – 80[	16	75	34	1200	90000	6750000
[80 – 90[	14	85	48	1190	101150	8597750
[90 – 100[	10	95	58	950	90250	8573750
[100 – 110[	5	105	63	525	55125	5788125
[110 – 120[	2	115	65	230	26450	3041750
$\sum$	65			5185	429425	36828625

بعد إجراء العديد من العمليات نحصل على النتائج التالية:

$$\begin{pmatrix} Mo \cong 75 & m_1 = \bar{X} = 79.8 \\ Me \cong 79.1 & m_2 = 6606.5 \\ Q_1 \cong 68.2 & m_3 = 566594.23 \\ Q_3 \cong 90.7 & \end{pmatrix}$$

بالتالي:

$$\begin{cases} \mu_2 = m_2 - m_1^2 = 238.46 \Rightarrow \sigma_x = 15.44 \\ \mu_3 = m_3 - 3m_2m_1 + 2m_1^3 = 1337.31 \end{cases}$$

– حساب معامل فيشر للالتواء

$$\gamma_F = \frac{\mu_3}{\sigma^3} = \frac{1337.31}{3680.8} = 0.363$$

– حساب معامل بيرسون الأول للالتواء

$$\gamma_1 = \frac{\bar{X} - M0}{\sigma} = \frac{4.8}{15.44} = 0.3$$

– حساب معامل بيرسون الثاني للالتواء

$$\gamma_2 = \frac{\mu_3^2}{\mu_2^3} = \frac{1788398}{13559592} = 0.131$$

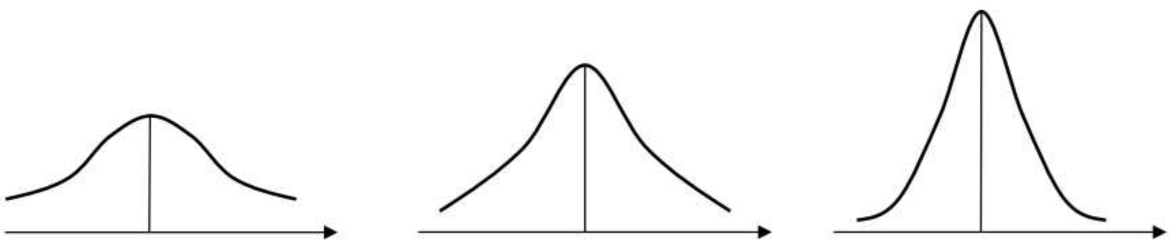
– حساب معامل يول للالتواء

$$\gamma_3 = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_1)} = \frac{(90.7 - 79.1) - (79.1 - 68.2)}{(90.7 - 68.2)} = 0.03$$

كل المقاييس المستخدمة لحساب الالتواء موجبة وبالتالي فالتوزيع ملتوي نحو اليمين.

6 – 4 التفرطح Aplantissement

التفرطح هو درجة تدبب قمة التوزيع، ويؤخذ عادة بالقياس إلى التوزيع الطبيعي Loi normal. وبالتالي يكون التفرطح مرتبط بتشتت البيانات حول قيمة مركزية في التوزيع الإحصائي. فكلما كان التشتت كبيرا كلما كان الشكل أكثر تفرطح.



توزيع مدبب   توزيع طبيعي   توزيع مفرطح

إذا كان شكل التوزيع أكثر استطالة (ارتفاع) من الشكل الطبيعي فهو مدبب Leptocurtique،

إذا كان شكل التوزيع أقل استطالة (ارتفاعا) من الشكل الطبيعي فهو مفرطح Platicurtique.

من أهم المقاييس المستخدمة لقياس التفرطح نجد:

6 – 4 – 1 معامل بيرسون للتفرطح

يعتمد بيرسون على العزوم لقياس تفرطح توزيع الذي يعطى بالعلاقة:

$$K=\frac{\mu_4}{\mu_2^2}=\frac{\mu_4}{\sigma^4}$$

- إذا كان $K\langle 3$ فالتوزيع مفرطح،
- إذا كان $K=3$ فالتوزيع طبيعي،
- إذا كان $K\rangle 3$ فالتوزيع مدبب.

مثال رقم (6–7).

باستخدام بيانات المثال رقم (6–3) أوجد تفرطح التوزيع.

$$K=\frac{\mu_4}{\mu_2^2}=\frac{88.5}{7.5^2}=1.573$$

بما أن قيمة المعامل أقل من 3 فإن التوزيع مفرطح.

6 – 4 – 2 معامل فيشر للتفرطح

يعتمد على معامل بيرسون للتفرطح ويعطى بالعلاقة:

$$K_F=K-3$$

- إذا كان K_F أقل من 0 فالتوزيع مفرطح،
- إذا كان K_F يساوي 0 فالتوزيع طبيعي،
- إذا كان K_F أكبر من 0 فالتوزيع مدبب.

6– 5 تمارين حول الفصل السادس

تمرين رقم (6–1).

باستخدام بيانات المثال رقم (2–11)، أوجد معامل فيشر للالتواء ومعامل التفرطح لأوزان الـ 100 طالب؟ مع التعليق؟

الحل:

الفئات	Fi	ci	$ci-\overline{x}$	$(ci-\overline{x})^3$	$Fi(ci-\overline{x})^3$	$(ci-\overline{x})^4$	$Fi(ci-\overline{x})^4$
59.5–62.5	5	61	−6.45	−268,34	−1341,7	1730,77	8653,84
62.5–65.5	18	64	−3.45	−41,064	−739,15	141,67	2550,05
65.5–68.5	42	67	0.45	−0,0911	−3,8273	0,04101	1,72226
68.5–71.5	27	70	2.55	16,5814	447,697	42,2825	1141,63
71.5–74.5	8	73	5.55	170,954	1367,63	948,794	7590,35
$\sum$	100	–	0	–	−269,33	–	19937,6

1 – حساب الانحراف المعياري

$$\sigma_x = \sqrt{\frac{\sum fi(c_i-\overline{X})^2}{\sum fi}} = \sqrt{\frac{852.7500}{100}} = 2.92kg$$

2 – حساب معامل فيشر للالتواء

$$\mu_3 = \frac{\sum f_i(c_i-\overline{x})^3}{\sum f_i} = \frac{-269.325}{100} = -2.69325kg$$

$$\gamma_F = \frac{\mu_3}{\sigma^3} = \frac{-2.69325}{(2.92)^3} = -0.108$$

التعليق:

التوزيع ذو التواء ضعيف إلى اليسار .

3 – حساب معامل التفرطح

$$\mu_4 = \frac{\sum f_i (c_i - \bar{x})^4}{\sum f_i} = \frac{19937.6}{100} = 199.376kg$$

$$K = \frac{\mu_4}{\sigma^4} = \frac{199.376}{(2.92)^4} = 2.74$$

التعليق:

التوزيع مفرطح قليلا.

تمرين رقم (6−2).

ليكن لديك التوزيع المعطى في الجدول التالي.

Xi	0	1	2	3
Fi	0.216	0.432	0.288	0.064

أوجد مقياسي فيشر وبيرسون للتماثل والتفرطح؟ علق على النتائج؟

الحل

xi	Fi	Fixi	$Fixi^2$	$(x_i - \bar{X})$	$f_i(x_i - \bar{X})^2$	$(x_i - \bar{X})^2$	$f_i(x_i - \bar{X})^3$
0	0.216	0	0	−1.2	0.311	−0.373	0.448
1	0.432	0.432	0.432	−0.2	0.017	−0.0035	0.00069
2	0.288	0.576	1.152	0.8	0.184	0.147	0.11796
3	0.064	0.192	0.576	0.8	0.207	0.373	0.6718
Σ	1	1.2	2.16		0.72	0.144	1.238
		m_1	m_2		μ_2	μ_3	μ_4

− معامل فيشر للالتواء

$$\gamma_F = \frac{\mu_3}{\sigma^3} = 0.24$$

− حساب معامل بيرسون الثاني للالتواء

$$\gamma_2 = \frac{\mu_3^2}{\mu_2^3} = 0.05$$

التوزيع ملتو إلى اليمين

- معامل فيشر للتفرطح

- معامل بيرسون للتفرطح

$$K_F = K-3= 2.39 -3 =-0.61$$

التوزيع التكراري مفرطح.

مسائل محلولة

مسألة رقم (1)

مؤسسة P تتكون من وحدتين P1 وP2. يريد المدير العام للمؤسسة الذي قام بقبولك لإجراء تربص ميداني، أن يستفيد من معارفك في الإحصاء الوصفي، حيث قدم لك الجدول الموالي الذي يمثل توزيع أجور الموظفين (10^2 أورو) وهذا حسب فئات محددة في الوحدتين.

	P1		P2	
	x_{1i}	n_{1i}	x_{2i}	n_{1i}
عمال	$[11-13[$	60	$[9-11[$	5
رؤساء ورشات	$[15-17[$	95	$[13-15[$	15
إطارات	$[19-23[$	5	$[17-19[$	30

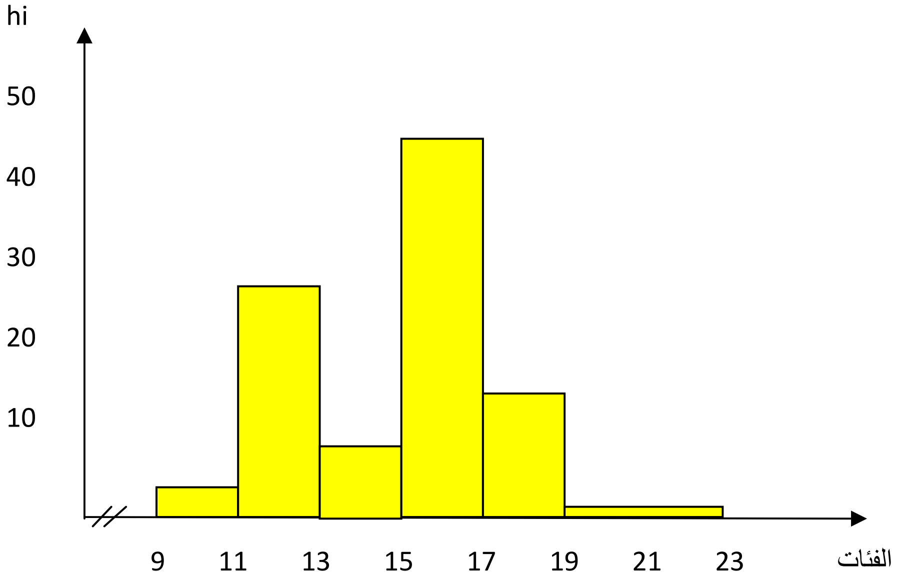
طلب منك المدير العام ما يلي:

1. تكوين مدرج تكراري لأجور موظفي المؤسسة؟
2. حساب الأجر الوسيط لـ P، وتفسير مدلوله؟
3. معرفة تماثل التوزيع لـ P، مع التعليل؟

الحل:

1.من أجل تكوين المدرج التكراري نقوم أولا بإعداد الجدول التكراري العام. هذا الجدول يحتوي على فئات غير متساوية. بالتالي نقوم بحساب التكرار المعدل h_i.

الفئات	F_i	Fr(%)	a_i	h_i
$[9-11[$	5	2.38	2	2.38
$[11-13[$	60	28.57	2	28.57
$[13-15[$	15	7.14	2	7.14
$[15-17[$	95	45.24	2	45.24
$[17-19[$	30	14.29	2	14.29
$[19-23[$	5	2.38	4	1.19
Σ	210			



2. حساب الأجر الوسيط للمؤسسة.

– نقوم بحساب التكرار التجميعي الصاعد Fcc

الفئات	F_i	Fcc
$[9-11[$	5	5
$[11-13[$	60	65
$[13-15[$	15	80
$[15-17[$	95	175
$[17-19[$	30	205
$[19-23[$	5	210
$\sum$	210	

– نقوم بتحديد رتبة الوسيط بالعلاقة

$$Rme = \frac{\sum_{i=o}^{n} F_i}{2} = \frac{210}{2} = 105$$

– نحسب الوسيط:

$$Me = L_0 + \frac{\frac{\sum_{i=o}^{n} F_i}{2} - F_1}{F_2 - F_1} \times K$$

$$Me = 15 + \frac{105 - 80}{175 - 80} \times 2 = 15.5263$$

بالتالي فإن وسيط الأجور للمؤسسة هو 1552.63 أورو.

3. تحديد تماثل التوزيع.

– حساب المتوسط الحسابي

الفئات	F_i	C_i	$Fi \times Ci$
$[9 - 11[$	5	10	50
$[11 - 13[$	60	12	720
$[13 - 15[$	15	14	210
$[15 - 17[$	95	16	1520
$[17 - 19[$	30	18	540
$[19 - 23[$	5	21	105
$\sum$	210		3145

المتوسط الحسابي يساوي:

$$\bar{X} = \frac{\sum_{i=o}^{n} F_i C_i}{\sum_{i=o}^{n} F_i} = \frac{3145}{210} = 14.9762$$

بالتالي فإن متوسط أجور المؤسسة هو 1497.62 أورو.

– حساب المنوال

بما أن طول الفئات غير متساوي فإننا نستخدم التكرار المعدل h_i لحساب المنوال. وبالتالي فإن الفئة المنوالية التي تقابل أكبر تكرار معدل هي الفئة الرابعة.

$$M_0 = L_o + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times K$$

$$M_0 = 15 + \frac{45.25 - 7.14}{(45.25 - 7.14) + (45.25 - 14.29)} \times 2 = 16.1035$$

بالتالي فإن منوال أجور المؤسسة هو 1610.35 أورو.

– نقارن بين المتوسط الحسابي، الوسيط والمنوال فنجد أن:

$$\bar{X} < Me < M_0$$

وبالتالي فإن المنحنى ملتو إلى اليسار.

مسألة رقم (2).

ليكن لدينا التوزيع التكراري التالي:

12	10	8	6	4	Xi
6	9	16	10	8	Fi

المطلوب: أحسب قيمة كل من

$$K, \gamma_F, CV_{1,} CV_2, \sigma_X, \delta_X, e_Q, e_{MQ}, e_{\bar{X}}, Me, Mo, MQ, H, G, \bar{X}$$

الحل:

Xi	Fi	XiFi	log(Xi)	Fi.log(Xi)	Fi/Xi	$(Xi)^2$	$Fi.(Xi)^2$
1	6	6	0	0	6	1	6
2	12	24	0,301	3,61236	6	4	48
3	16	48	0,477	7,63394	5,33	9	144
4	14	56	0,602	8,42884	3,5	16	224
5	5	25	0,699	3,49485	1	25	125
∑	53	159		23,17	21,8	55	547

- حساب قيمة المتوسطات

$$\bar{X} = \frac{\sum fixi}{\sum fi} = \frac{159}{53} = 3$$

$$\log G = \frac{\sum_{i=1}^{n} f_i \log x_i}{\sum_{i=1}^{n} f_i} = \frac{23.17}{53} = 0.4372 \Rightarrow G = 10^{0.4372} = 2.73$$

$$H = \frac{\sum_{i=1}^{n} f_i}{\sum_{i=1}^{n} \frac{f_i}{x_i}} = \frac{53}{21.8} = 2.34$$

$$MQ=\sqrt{\frac{\sum_{i=1}^{n} f_i x_i^2}{\sum_{i=1}^{n} f_i}}=\sqrt{\frac{547}{53}}=3.21$$

لاحظ أن: $MQ > \bar{X} > G > H \leftrightarrow 3.21 > 3 > 2.73 > 2.34$

حساب قيمة المنوال والوسيط:

Xi	Fi	Fcc
1	6	6
2	12	18
3	16	34
4	14	48
5	5	53
$\sum$	53	

$\frac{\sum fi}{4}=13.25$

$\frac{3\sum fi}{4}=26.5$

$\frac{\sum fi}{4}=39.75$

بالتالي:

قيمة الربيع الأول تساوي 2، قيمة الربيع الثاني (الوسيط) تساوي 3، قيمة الربيع الثالث تساوي 4.

قيمة المنوال تساوي 3.

من خلال الحسابات السابقة نلاحظ أن قيمة المتوسط الحسابي=الوسيط=المنوال=3 بالتالي فهذا التوزيع متماثل.

- حساب مقاييس التشتت

Xi	Fi	$\lvert Xi-\bar{X}\rvert$	$Fi\lvert Xi-\bar{X}\rvert$	$\lvert Xi-Me\rvert$	$Fi\lvert Xi-Me\rvert$	$(Xi-\bar{X})^2$	$Fi(Xi-\bar{X})^2$
1	6	2	12	2	12	4	24
2	12	1	12	1	12	1	12
3	16	0	0	0	0	0	0
4	14	1	14	1	14	1	14
5	5	2	10	2	10	4	20
$\sum$	53		48		48		70

$$e_{\bar{X}} = \frac{\sum fi|Xi-\bar{X}|}{\sum fi} = \frac{48}{53} = 1.09,$$

$$e_{Me} = \frac{\sum fi|Xi-Me|}{\sum fi} = \frac{48}{53} = 1.09,$$

$$e_Q = \frac{Q_3-Q_1}{2} = \frac{39.75-13.25}{2} = 13.25,$$

$$S_{\bar{X}} = \sqrt{\frac{\sum fi(Xi-\bar{X})^2}{\sum fi-1}} = \sqrt{\frac{70}{53-1}} = 1.16,$$

$$\sigma_{\bar{X}} = \sqrt{\frac{\sum fi(Xi-\bar{X})^2}{\sum fi}} = \sqrt{\frac{70}{53}} = 1.15\,,$$

$$CV_1 = \frac{S_{\bar{X}}}{\bar{X}}100 = 38.66,$$

$$CV_2 = \frac{Q_3-Q_1}{Q_2}100 = 100,$$

حساب مقاييس الشكل:

Xi	Fi	$(Xi-\bar{X})$	$(Xi-\bar{X})^3$	$Fi(Xi-\bar{X})^3$	$(Xi-\bar{X})^4$	$Fi(Xi-\bar{X})^4$
1	6	−2	−8	−48	16	96
2	12	−1	−1	−12	1	12
3	16	0	0	0	0	0
4	14	1	1	14	1	14
5	5	2	8	40	16	80
∑	53	0	0	−6		202

$$\gamma_F = \frac{\mu_3}{\sigma_{\bar{X}}^3} = \frac{-0.12}{1.52} = -0.078$$

من خلال معامل فيشر للالتواء نلاحظ أنه يوجد التواء سالب بسيط، هذه النتيجة لم نلاحظها عند مقارنة كل من المتوسط الحسابي، الوسيط والمنوال، لهذا تدعى هذه الطريقة بالطريقة الدقيقة.

من خلال حساب معامل فيشر نلاحظ أن التوزيع يقترب من التوزيع الطبيعي وهو مفرطح قليلا.

مسألة رقم (3).

تمرين رقم (1)

يمثل الجدول التالي توزيع أوزان 100 شخص (كلغ):

الفئات	45−40	50−45	55−50	60−55
التكرار	22	n1	24	n2

– إذا علمت أن الوسط الحسابي $\bar{X} = 49.7$ فأحسب n1 ، n2 ؟

تمرين رقم (2)

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

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

المطلوب:

– إعادة تكوين الجدول التكراري؟

تمرين رقم (3)

الجدول التالي يمثل مبيعات 500 حذاء خلال أسبوع الدخول المدرسي

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

المطلوب:

– حساب α و β إذا علمت أن $\alpha = 2\beta$ ماذا يمثل كل من α، β؟ ماذا تلاحظ؟

تمرين رقم (4)

إذا عملت أن معامل الاختلاف لإنتاج أحد المصانع في فترة ما هو 20%، أوجد عدد أيام هذه الفترة إذا كان الانحراف المعياري للإنتاج هو 10 ومجموع إنتاج الفترة يساوي 500 وحدة؟

الحل

حل التمرين رقم (1)

$$\begin{cases} \bar{X} = \dfrac{X_1n_1 + X_2n_2 + X_3n_3 + X_4n_4}{n_1 + n_2 + n_3 + n_4} \\ n_1 + n_2 + n_2 + n_2 = 1000 \end{cases}$$

$$\begin{cases} 49.7 = \dfrac{935 + 47.5n_2 + 1260 + 57.5n_4}{100} \\ 22 + n_2 + 24 + n_4 = 100 \end{cases}$$

$$\begin{cases} 47.5n_2 + 57.5n_4 = 2775 \\ n_2 + n_4 = 54 \end{cases}$$

$$\begin{cases} n_2 = 33 \\ n_4 = 21 \end{cases}$$

حل التمرين رقم (2)

الفئات	Fi	Fcc
100−200	10	10
200−300	20	30
300−400	40	70
400−500	20	90
500−600	10	100
	100	

$a = F_1 * F_5 = 10*10 = 100$

$C_{max3} = F_2 * F_4 = 20*20 = 400$

حل التمرين رقم (3)

$$\sum_{i=1}^{10} F_i = 500 \Rightarrow 380 + \alpha + \beta = 500 \Rightarrow 380 + 2\beta + \beta = 500 \Rightarrow \beta = 40, \alpha = 80$$

تمثل α عدد الأحذية المباعة من قياس 41 وβ عدد الأحذية المباعة من قياس 42. ونلاحظ أن الأحذية قياس 41 هي القيمة السائدة الأكثر بيعا.

حل التمرين رقم (4)

$$CV_1 = \frac{S_x}{\bar{X}}100 \Rightarrow 20 = \frac{10}{\bar{X}}100 \Rightarrow \bar{X} = 50$$

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n} \Rightarrow 50 = \frac{500}{n} \Rightarrow n = 10$$

مسألة رقم (3)

التمرين الأول: (10 ن)

مكنت دراسة أجريت حول أجور 100 عامل من عرضها جدوليا في صورة التوزيع التكراري التالي (الوحدة ألف دج)

الفئات	30–20	40–30	50–40	60–50	70–60	80–70	90–80	المجموع
التكرارات	8	12	15	18	24	15	8	100

المطلوب:

1 – مثل بيانيا الظاهرة الإحصائية المدروسة في صورة مدرج تكراري ومضلع تكراري؟

2 – حدد نسبة العمال الذين تصل أجور كل منهم 70 دج أو تزيد؟ ونسبة العمال الذين تصل أجور كل منهم أقل من 50 دج؟

3 – أحسب كل من: أ – المتوسط الحسابي؟ الوسيط Me؟ المنوال Mo (جبريا، وبيانيا)؟

4 – أحسب الربيع الأول و الثالث؟ وفسر معناهما؟

5 – أحسب كل من: أ – الانحراف المتوسط $e_{\bar{X}}$؟ الانحراف الربيعي e_Q ؟ الانحراف المعياري S_x ؟

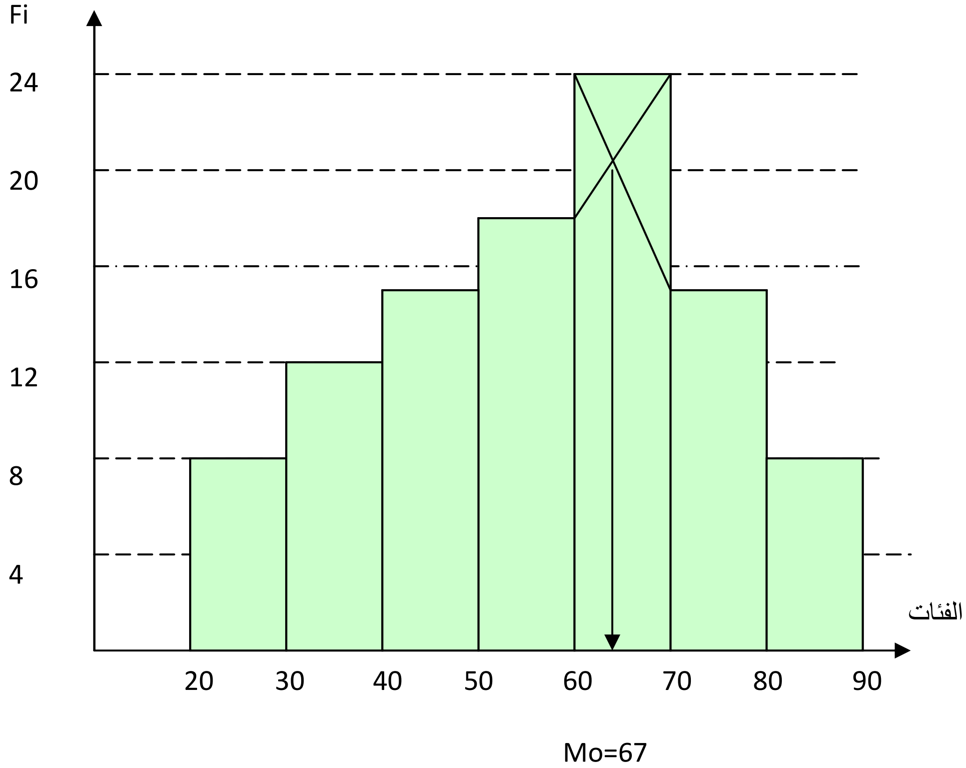
6 – أحسب معامل الاختلاف الأول CV_1؟ وضح الحالات التي نستخدم فيها معاملات التشتت النسبي؟

7 – مثل بيانيا المنحنى التكراري؟ وضح فيه كل من $\bar{X}, Me, Mo$ ؟ ما هي ملاحظاتك حول تماثل التوزيع؟

8 – أحسب معامل فيشر للالتواء؟ ماذا تستنتج؟ وهل تؤكد هذه النتيجة نتيجة السؤال 7؟

حل التمرين الأول:

1 – رسم المدرج التكراري



2 –

– نسبة العمال الذين تصل أجور كل منهم 70 دج أو تزيد: $\frac{23}{100}\times 100=23\%$

– نسبة العمال الذين تصل أجور كل منهم أقل من 50 دج: $\frac{35}{100}\times 100=35\%$

3 –

الفئات	Fi	Ci	Fcc	Fcd	Fici
20–30	8	25	8	100	200
30–40	12	35	20	92	420
40–50	15	45	35	80	675
50–60	18	55	53	65	990
60–70	24	65	77	57	1560
70–80	15	75	92	23	1125
80–90	8	85	100	8	680
المجموع	100				5650

– حساب المتوسط الحسابي

$$\overline{X} = \frac{\sum_{i=1}^{n} f_i c_i}{\sum_{i=1}^{n} f_i} = \frac{5650}{100} = 56.5$$

– الوسيط:

$$Me = L_0 + \frac{\frac{\sum fi}{2} - F_1}{F_2 - F_1} \times K = 50 + \frac{50-35}{53-35} \times 10 = 58.33$$

– المنوال

أ – جبريا

$$Mo = Lo + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times K = 60 + \frac{6}{6+9} \times 10 = 64$$

– بيانيا: في الشكل السابق.

4 – حساب الربيع الأول

$$Q_1 = L_0 + \frac{\frac{\sum fi}{4} - F_1}{F_2 - F_1} \times K = 40 + \frac{25-20}{35-20} \times 10 = 43.33$$

– حساب الربيع الثاني

$$Q_2 = L_0 + \frac{\frac{3\sum fi}{4} - F_1}{F_2 - F_1} \times K = 60 + \frac{75-53}{77-53} \times 10 = 69.16$$

التفسير:

– 25% من العمال يتقاضون أجور تساوي أو تقل عن 43.33 دج،

– 75% من العمال يتقاضون أجور تساوي أو تقل عن 69.16 دج.

5 –

الفئات	Fi	ci	$(ci-\bar{x})$	$\lvert ci-\bar{x}\rvert$	$Fi\lvert ci-\bar{x}\rvert$	$(ci-\bar{x})^2$	$Fi(ci-\bar{x})^2$	$(ci-\bar{x})^3$	$Fi(ci-\bar{x})^3$
20−30	8	25	−32	32	252	992	7938	−31255,875	−3E+05
30−40	12	35	−22	22	258	462	5547	−9938,375	−1E+05
40−50	15	45	−12	12	173	132	1983,75	−1520,875	−22813
50−60	18	55	−2	1,5	27	2,25	40,5	−3,375	−60,75
60−70	24	65	8,5	8,5	204	72,3	1734	614,125	14739
70−80	15	75	19	19	278	342	5133,75	6331,625	94974
80−90	8	85	29	29	228	812	6498	23149,125	185193
المجموع	100				1419		28875		−97275

- حساب الانحراف المتوسط: $e_{\bar{x}}=\dfrac{\sum_{i=1}^{n} f_i\lvert c_i-\overline{X}\rvert}{\sum_{i=1}^{n} f_i}=\dfrac{1419}{100}=14.19$

- حساب الانحراف الربيعي: $e_Q=\dfrac{Q_3-Q_1}{2}=\dfrac{69.16-43.33}{2}=12.92$

- حساب الانحراف المعياري: $S_x=\sqrt{\dfrac{\sum fi(c_i-\overline{X})^2}{(\sum fi)-1}}=\sqrt{\dfrac{28875}{99}}=17.07$

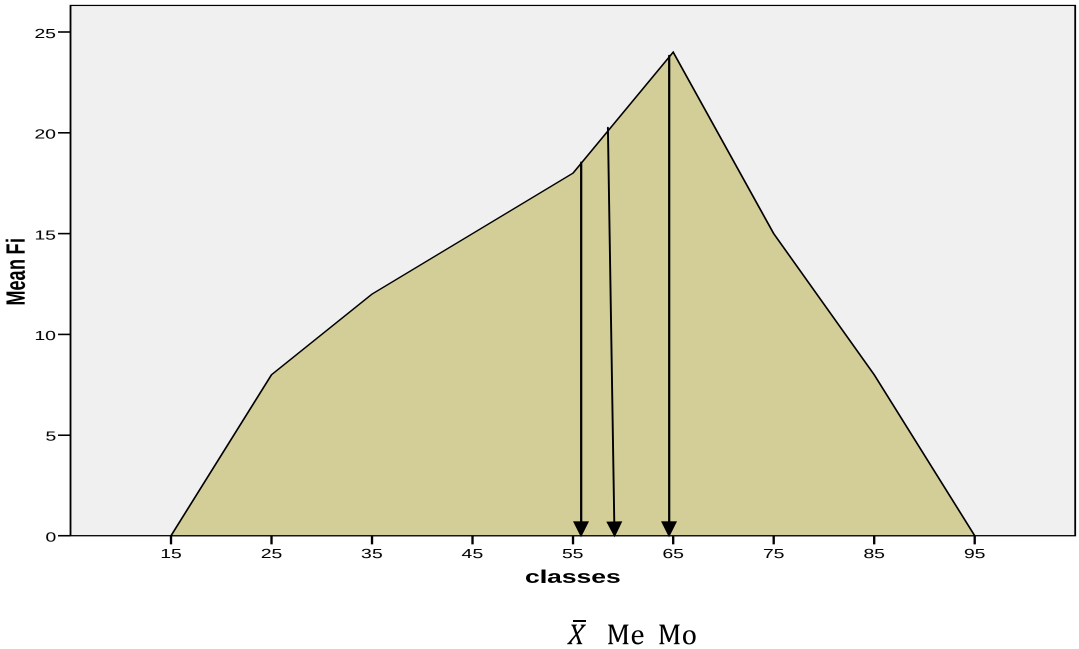
6 – حساب معامل الاختلاف الأول: $CV_1=\dfrac{s_x}{\overline{X}}\times 100=\dfrac{17.07}{56.5}\times 100=30.21\%$

– تستخدم معاملات التشتت النسبي عند:

أ– مقارنة التشتت لمفردات ظاهرتين مختلفتين،

ب– أو لمفردات ظاهرة واحدة على مستويات مختلفة

7 – تمثيل المنحنى التكراري:



الملاحظة:

نلاحظ أن $\bar{x}\langle M_e \langle M_o$ بالتالي المنحنى ملتوي إلى اليسار.

8 – حساب معامل فيشر للالتواء:

$$\mu_3 = \frac{\sum f_i (x_i - \bar{x})^3}{\sum f_i} = \frac{-97275}{100} = -972.75$$

$$\sigma_x = \sqrt{\frac{\sum fi(c_i - \bar{X})^2}{\sum fi}} = \sqrt{\frac{28875}{100}} = 16.99$$

$$\gamma_F = \frac{\mu_3}{\sigma^3} = \frac{-972.75}{16.99} = -0.19$$

- بما أن قيمة معامل فيشر سالبة نستنتج أن التوزيع سالب الالتواء.
- نعم تؤكد هذه النتائج نتائج السؤال 7؟

اختبارات مقترحة

اختبار رقم 01

وزارة التعليم العالي والبحث العلمي
جامعة 8 ماي 45 - قالمة
كلية العلوم الاقتصادية وعلوم التسيير

المقياس : إحصاء وصفي
المدة : ساعتان
الأستاذ : شريف غياط

الامتحان المتوسط المدى رقم (02)

التمرين الأول :

- إذا كان $\bar{X} = M_o = M_e$ لتوزيعين تكراريين، فهل هذا يعني أنهما متشابهان؟ بين ذلك؟
- انطلاقا من المساواة السابقة بين صحة العلاقة التالية :

$$2 \sum n_i x_i = 2A \sum n_i + C \sum n_i$$

حيث (C : طول الفئة)

التمرين الثاني :

يبين الجدول التكراري التالي توزيع 65 عاملا حسب أجورهم اليومية :

الفئات	ن%
$X_i < 60$	12,3
60 - 70	15,38
70 - 80	24,61
80 - 90	21,53
90 - 100	15,38
100 - 110	7,69
$X_i \geq 110$	3,07

المطلوب :

1. ما هو الأجر المتوسط ؟
2. قس تشتت هذا التوزيع باستعمال مقياس مناسب؟
3. قس التواء هذا التوزيع باستعمال مقياس بيرسون؟
4. وضح لماذا لا يعد الانحراف المعياري مقياسا مثاليا للتشتت ؟
5. إذا فرضنا أن طول كل من الفئة الأولى والأخيرة يساوي 10. قارن بين النسبة النظرية والنسبة التجريبية للمجال التالي $[\bar{X} \pm S]$

التمرين الثالث :

يبين الجدول التكراري التالي توزيع 100 طالب حسب علاماتهم في مقياسي الإحصاء والرياضيات :

- 1/2 -

x_i	2	4	5	7	9	10	11	13	14	15	16	$\sum n_i$
n_i1	1	3	6	8	14	20	25	18	2	2	1	100
n_i2	2	3	5	7	10	15	20	22	10	5	1	100

المطلوب:

1- احسب معامل الاختلاف في الحالتين؟ ماذا تلاحظ؟

2- احسب معامل الالتواء لبيرسون في الحالتين؟

3- قارن تشتت العلامات في الحالتين بناءا على السؤالين 1 و 2؟

4- متى نستخدم مقاييس التشتت المطلقة والنسبية؟

مع تمنيات أستاذ المقياس
لكم بالتوفيق والنجاح

2002/03/09

2/2

اختبار رقم 02.

جامعة 8 ماي 1945 قالمة

كلية العلوم الاقتصادية ، التجارية و علوم التسيير

قسم العلوم التجارية

السنة الأولى جذع مشترك

الموسم الجامعي 2015 - 2016

الامتحان النهائي لمادة الإحصاء 1

نص التمرين الأول (05 نقاط): ليكن لدينا التوزيع التكراري الآتي

الفئات	20 - 10	30 - 20	40 - 30	50 - 40	60 - 50	70 - 60
F_i	20	15	20	10	15	20

المطلوب: حدد قيمة المنوال ، الوسيط و الوسط الحسابي بالطريقة المباشرة ، طريقة الوسط الفرضي $(\alpha = 35)$ و طريقة الانحرافات المختصرة

نص التمرين الثاني(07 نقاط): عند دراسة عينة حجمها 80 ، تحصلنا على المعلومات المبينة في الجدول التالي

X_i	3	8	14	16	20	24
Fr	0.2	0.05	0.3	0.25	0.1	0.1

المطلوب: حساب كافة مقاييس النزعة المركزية و كافة مقاييس التشتت (المطلقة و النسبية)

ملاحظة: تقرب كل نتائج التمرين الثاني إلى ثلاثة (03) أرقام بعد الفاصلة

نص التمرين الثالث (08 نقاط): ليكن لدينا التوزيع التكراري التالي

الفئات	10 - 0	30 - 10	70 - 30	80 - 70	90 - 80	100 - 90
F_i	10	40	40	20	30	20

المطلوب: حساب قيمة معامل الالتواء بالطريقة الدقيقة و الطرق التقريبية ثم قيمة معامل التفرطح. ما تعليقك على النتائج المتحصل عليها ؟

بالتوفيق للجميع

اختبار رقم 03

Année 2009/2010
L1 AES – 8 Avril 2010

CONTROLE CONTINU – STATISTIQUES DESCRIPTIVES

Documents interdits – Seules les calculatrices « 4 opérations » sont autorisées – Tout résultat non justifié ne sera pas comptabilisé.
Le barème signalé n'est qu'indicatif et est susceptible de se modifier à la marge

N'oubliez pas d'indiquer sur votre copie : Le TD auquel vous assistez (N° de groupe + jour et heure de TD+nom de votre chargé de TD)

Les exercices 1 et 2 sont indépendants et peuvent être traités séparément.

EXERCICE 1 (8,5 pts)

Les deux parties sont indépendantes et peuvent être traitées séparément.

Une étude sur le budget consacré aux vacances d'été auprès de ménages Bretons a donné les résultats suivants.

Budget	Fréquence cumulée
	0
[800-1000[	
	0,08
[1000-1400[	
	0,18
[1400-1600[	
	0,34
[1600-**y**[	
	0,64
[**y**-2400[	
	0,73
[2400-**x**[	
	1

PARTIE 1

1/ Certaines données sont manquantes. Calculer la borne manquante **x** sachant que l'étendue de la série est égale à 3200 *(0,5pt).*

2/ Calculer la borne manquante **y** dans les deux cas suivants :

a- Le budget moyen est égal à 1995 euros *(1,5 pts).*

b- Le budget médian est égal à 1920 euros *(1,5 pts).*

PARTIE 2

Considérons maintenant que la borne manquante y ±ségale à 2000 euros.

3/ Donner une représentation graphique de la distribution des budgets « vacances » *(1,5 pts).*

4/Calculer et interpréter le budget moyen et médian. Que pouvez-vous conclure de la comparaison entre ces deux valeurs *(1,5 pts)*?

5/ Sachant que $\sum_i n_i x_i^2 = 4741200000$ et V(x)=604044, retrouver les effectifs correspondant à chacune des tranches de budgets ainsi que l'effectif total des ménages enquêtés *(2 pts).*

EXERCICE 2 (11,5 pts)

La répartition des salaries mensuels d'une entreprise est donnée par le tableau suivant :

Salaire	**nb. de salariés**
[1000-1400[	100
[1400-1800[	150
[1800-2200[	40
[2200-3000[	10

1 / Calculer le salaire moyen ainsi que l'écart-type de cette distribution *(1,5pts).* Commenter.

2/ Tracer la boîte à moustache de cette série. Commenter. *(2 pt)*

3/ Calculer les fréquences et les fréquences cumulées de la distribution de la masse salariale (masse salariale détenue par chaque catégorie d'individus) *(1,5 pts).*

4/ A partir du tableau précédent, calculer la médiale. Que vous indique la comparaison entre la médiane et la médiale *(1,5 pts)*?

5/ Tracez la courbe de Lorenz. Que remarquez-vous ? *(1,5 pt)*

6/ Après avoir rappelé ce que mesure l'indice de Gini et comment il se calcule, calculer et interpréter l'indice de Gini. *(2,5 pts)*

7/ Supposons que, sur une période de 5 ans, le salaire moyen augmente de 20% :

a – Quelle est l'évolution annuelle moyenne du salaire moyen ? *(0,5pt)*

b - Au bout de combien d'années le salaire moyen aura-t-il doublé ? *(0,5pt)*

المراجع

– إسماعيل سعيد السيد علي، مبادئ الإحصاء الوصفي والتطبيقي، مؤسسة حورس الدولية، الأردن، 2011.

– المغربي محمد محمد جبر، الإحصاء التحليلي، المكتبة العصرية، مصر، 2011.

– بوهزة محمد، محاضرات في الإحصاء الوصفي، دار المحمدية العامة، الجزائر، 2011.

– جيلاطو جيلالي، الإحصاء، ديوان المطبوعات الجامعية، الجزائر، 2007.

– راتول محمد، الإحصاء الوصفي، ديوان المطبوعات الجامعية، الجزائر، 2009.

– رجال السعدي، الإحصاء الوصفي، مؤسسة الرجاء للطباعة والنشر، الجزائر، 2013.

– شبيجل، مواري ر.، الإحصاء، الدار الدولية للاستثمارات الثقافية، مصر، 2008.

– شطيبي شريف، محاضرات في الإحصاء الوصفي، مطبوعات جامعة منتوري، الجزائر، 2006.

– عزوز عبد القادر، الكامل في الإحصاء، ديوان المطبوعات الجامعية، الجزائر، 2010.

– علاء الدين محمد، فرحان نور الدين حسن، مبادئ الأسلوب الإحصائي، هيئة المعاهد الفنية، العراق، 1988.

– موساوي عبد النور، بركان يوسف، الإحصاء، دار العلوم للنشر والتوزيع، الجزائر، 2009.

– نوروسيس ماريجا، تحليل البيانات باستخدام SPSS، شعاع للنشر والعلوم، سورية، 2010.

- Boutti Rachid, Statistique descriptive appliquée, sans Maison d’édition, Maroc, sans année d’édition.

- Elabjani Abdelaziz, Exercice de statistique descriptive, Imprimerie Papeterie El-watanya, Maroc, 2013.

- Py Bernard, Statistique descriptive, Economica, France, 2007.