الجمهورية الجزائرية الديمقراطية الشعبية

République Algérienne Démocratique et Populaire

وزارة التعليم العالي والبحث العلمي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Mustapha STAMBOULI de Mascara

Faculté des Sciences Economiques, Sciences Commerciales et Sciences de Gestion

جامعة مصطفى اسطمبولي بمعسكر

كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير

مطبوعة موجهة لطلبة السنة أولى جذع مشترك علوم اقتصادية، علوم تجارية وعلوم التسيير

بعنوان

الإحصاء الوصفي

ملخص وتمارين محلولة

من إعداد: الدكتور زاوي أحمد صادق

السنة الجامعية: **2019/2018**

المـــــحـــــــتويـــــــات

المحور الأول: مدخل إلى علم الإحصاء

1- تعريفات أساسية

1-1- علم الإحصاء:

إن علم الإحصاء هو فرع من الرياضيات الذي يسهل ويساعد في اتخاذ القرارات السليمة المبنية على عدم التأكد، ويستخدم هذا العلم مجموعة من الأدوات والأساليب. ينقسم علم الإحصاء إلى قسمين: الإحصاء الوصفي، وهو مجموعة المبادئ العلمية التي تعالج البيانات، أو بعبارة أخرى هو مجموعة الطرق العلمية التي بواسطتها تجمع، تلخص، تنظم، تمثل وتشرح البيانات أو المعطيات؛ والقسم الآخر هو الإحصاء الاستدلالي أو الرياضي، وهو علم الإحصاء الذي يستخدم الأساليب الإحصائية للحصول على استنتاجات وتقديرات من خلال التحليل السليم للعينات وتعميمها على المجتمع.

إن الإحصاء الوصفي يدرس خصائص من المشاهدات المتعلقة بظاهرة ما أو تسجيل الملاحظات عن طريق القيام بتجربة لمعرفة اتجاه ظاهرة. فالتجربة هي المرحلة الأولية لكل دراسة.

1-1-1- التجربة الإحصائية:

هي التجربة التي يحددها الباحث ولا يعرف مسبقا نتائج وقوعها.

مثلا: أريد معرفة تأثير درجة الحرارة على نبتة الطماطم، فأقوم بغرس مجموعة من البذور في عدة أماكن مختلفة درجة الحرارة، وأقوم بتسجيل الملاحظات وأستخلص النتائج.

1-1-2- المجتمع الإحصائي:

علم الإحصاء يستند على المجتمعات في دراساته، وهذا المفهوم جاء نتيجة اهتمام الديمغرافيا بدراسة السكان خاصة فيما يتعلق بتعداده، وهذا ما جعل هذا الفرع يحتل مكانة عالية في علم الإحصاء. لكن في علم الإحصاء، ينطبق مصطلح المجتمع على أي كائن إحصائي مدروس، سواء ندرس الطلبة في جامعة ما، العمال في مؤسسة ما، الأسر في منطقة ما...إلخ، وأي مجموعة أخرى التي نقوم عليها الدراسة ونجمع حولها معلومات وملاحظات إحصائية.

ملاحظة:

في أغلب الدراسات الإحصائية، يلجأ الباحث في دراساته إلى اختيار جزء من المجتمع الإحصائي المدروس يعرف بالعينة الإحصائية. فالعينة هي جزء من المجتمع الإحصائي يتم اختيارها بعدة طرق، بحيث تمثل جميع أفراد المجتمع تمثيلا جيدا. ويلجأ الإحصائي إلى اختيار العينات لأنها تتميز بمجموعة من المميزات:

- استحالة فحص المجتمع بأكمله؛
- العينة أقل تكلفة وأكثر سرعة؛
- العينة أكثر دقة.

1-1-3- الوحدة الإحصائية:

كل مجتمع إحصائي مكون من أفراد، ويسمى الفرد الواحد بالوحدة الإحصائية. فالوحدة الإحصائية هي العنصر الأساسي المكون لهذا المجتمع الإحصائي.

مثلا: المجتمع الإحصائي هو السيارات، الوحدة الإحصائية هي السيارة؛ الموظفين، الموظف...الخ.

1-1-4- البيانات:

الإحصاء الوصفي يهتم بوصف مجتمع إحصائي معين، فهو يسعى إلى جمع جميع خصائص الوحدات الإحصائية التي يمكن أن تأخذ قيم مختلفة، وهذه القيم تعرف بما يسمى بالبيانات.

البيانات هي مجموعة المشاهدات المأخوذة أثناء دراسة معينة، وقد تكون بيانات وصفية مثل المستوى التعليمي ولون الشعر أو بيانات رقمية مثل أطوال مجموعة من الطلاب.

1-1-5- المتغير:

هو مقدار له خصائص رقمية أو غير رقمية تتغير قيمته من عنصر إلى آخر من عناصر المجتمع أو العينة.

1-1-6- الترتيبات:

إن ترتيبات متغيرة إحصائية هي مختلف القيم التي يمكن أن تأخذها هذه المتغيرة.

مثال: المتغيرة الإحصائية هي نوع الجنس؛ الترتيبات هي: ذكر أو انثى.

1-2- أنواع المتغيرات:

في علم الإحصاء، نميز بين نوعين رئيسيين من المتغيرات: المتغيرات النوعية والمتغيرات الكمية.

1-2-1- المتغيرات النوعية:

المتغيرات النوعية هي الصفات غير القابلة للقياس وترتيباتها تأخذ فئات إسمية معينة.

مثلا: تصنيف الطلبة حسب الاختصاص:

علوم اقتصادية، علوم اجتماعية، آداب، علوم...الخ.

المتغيرات النوعية تأخذ بدورها نوعين، ويكمن الاختلاف في ترتيب الفئات التي تتغير من فرد إلى آخر.

1) **المتغيرات النوعية الإسمية:** هي الصفات غير القابلة للقياس وتكون ترتيباتها لا تتطلب ترتيب معين (عشوائي).

مثلا: نصنف الطلبة حسب نوع الجنس: إما طلبة (ذكور أو إناث) أو طلبة (إناث أو ذكور) نفس الشيء، أي ترتيبات نوع المتغيرة (نوع الجنس) لا تتطلب الترتيب.

2) **المتغيرات النوعية الترتيبية:** هي الصفات غير القابلة للقياس، لكن ترتيباتها تكون مرتبة ومنظمة وفق سلم تدريجي معين.

مثلا: تصنيف الطلبة حسب المستوى الدراسي ما قبل التدرج:

سنة أولى سنة ثانية سنة ثالثة سنة أولى ماستر سنة ثانية ماستر.

فالتصنيف يكون من أدنى مستوى إلى أعلى مستوى أو الحالة العكسية.

1-2-2- المتغيرات الكمية:

المتغيرات الكمية هي المتغيرات القابلة للقياس، وكل قيمة تمثل بعدد أو برقم مثل العمر، عدد الأطفال، الأجر...الخ. والمتغيرات الكمية كذلك تنقسم إلى نوعين:

1) **المتغيرات الكمية المتقطعة:** وهي المتغيرات القابلة للقياس بأعداد ثابتة أو متقطعة (قابلة للعد بشكل منفصل في مجال) مثل: عدد الغرف في المنزل، عدد الأطفال لدى الأسر...الخ.

2) **المتغيرات الكمية المستمرة:** هي المتغيرات القابلة للقياس بأعداد مستمرة، حيث المتغيرة تأخذ أي قيمة في مجال مثل: أوزان الطلبة، أجور العمال...الخ.

إن الهدف من الإحصاء الوصفي هو جمع المعلومات الواردة في مجموعات البيانات وتبسيطها في جداول أو أشكال أو ملخصات رقمية. ويتم تحليل المتغيرات الإحصائية بشكل مختلف حسب طبيعتها سواء كمية أو نوعية.

2- تمارين محلولة:

2-1- التمرين الأول: حدّد من بين هذه البيانات هي نوعية، كمية متقطعة أو كمية مستمرة:

الدخل السنوي	الجنسية	الجنس	نوع العينين	الطول	نقاط الطلبة
مكان الإقامة	السن	مقاس الحذاء	عدد اللغات المتقنة	الحالة الاجتماعية	عدد مقاييس الاختصاص

2-2- التمرين الثاني: ما هي طبيعة المتغيرات التالية:

- عدد الأسهم المباعة يوميا في البورصة؛
- أجور أساتذة التعليم في ثانوية؛
- التفاوت في الأجور بين الرجال والنساء؛
- دول الاتحاد الأوروبي؛
- المستوى التعليمي للأجراء؛
- أشكال عقود العمل؛
- مستوى نمو الناتج الداخلي الخام؛
- معدل التضخم؛
- الميزان التجاري؛
- عدد الأفراد لدى الأسرة.

2-3- التمرين الثالث: من بين الدراسات الآتية، حدّد المجتمع الإحصائي، الوحدة الإحصائية والمتغيرة الإحصائية ونوعها:

- دراسة مدة صلاحية مجموعة من المنتجات؛
- دراسة غياب العمال في اليوم في مصنع ما؛
- توزيع طلبة السنة الأولى جذع مشترك حسب اختصاص شهادة البكالوريا؛
- توزيع العائلات حسب نوع الفاكهة الأكثر استهلاكا؛
- ترتيب السيارات حسب الجودة.

3- الحلول:

3-1- التمرين الأول:

المتغيرات النوعية	المتغيرات الكمية المتقطعة	المتغيرات الكمية المستمرة
مكان الإقامة، الجنسية، نوع العينين، الحالة الاجتماعية	عدد اللغات المتقنة، عدد مقاييس الاختصاص	الدخل السنوي، السن، مقاس الحذاء، الطول، نقاط الطلبة

3-2- التمرين الثاني:

المتغيرات النوعية الإسمية	المتغيرات النوعية الترتيبية	المتغيرات الكمية المتقطعة	المتغيرات الكمية المستمرة
دول الاتحاد الأوروبي أشكال عقود العمل	المستوى التعليمي	عدد الأسهم عدد الأفراد	الأجور، التفاوت في الأجور مستوى النمو معدل التضخم الميزان التجاري

3-3- التمرين الثالث:

المجتمع الإحصائي	الوحدة الإحصائية	المتغيرة الإحصائية	نوعها
المنتجات	المنتج	مدة الصلاحية	كمية مستمرة
العمال في مصنع	العامل	عدد غياب العمال	كمية متقطعة
طلبة السنة أولى جذع مشترك	الطالب	اختصاص البكالوريا	نوعية إسمية
العائلات	العائلة	نوع الفاكهة الأكثر استهلاكا	نوعية اسمية
السيارات	السيارة	الجودة	نوعية ترتيبية

4- تمارين مقترحة:

4-1- التمرين الأول:

حدّد المجتمع الإحصائي، الوحدة الإحصائية، المتغيرات الإحصائية ونوعها:

- أجرى باحث مختص دراسة حول أصناف الضرائب وحجمها ومعرفة نسبة مساهمتها في مجموع الضرائب المحصلة في مجموعة من الهيئات المختصة؛
- قام منتدى رؤساء المؤسسات الجزائرية بدراسة بتصنيف المؤسسات الجزائرية حسب الاختصاص، وحسب جودة المنتوجات من أجل توجيه المنتوجات نحو التصدير؛
- دراسة إحصائية من أجل معرفة تطور النمة السكاني لكل من الذكور والإناث؛
- أجرى المعهد الوطني للفلاحة بحث حول الكميات المنتجة من الحبوب الجافة ولبقوليات الجافة حسب النوع على جميع الولايات الجزائرية؛
- أجرى الديوان الوطني للإحصائيات دراسة حول العزوف الدراسي لكل من الرجال والنساء لجميع الأطوار الدراسية؛
- قامت المديرية الوطنية للوقاية وأمن الطرقات بدراسة حول عدد المخالفات المرورية في اليوم على مدى شهر كامل في الجزائر العاصمة.

4-2- التمرين الثاني:

حدّد مجال كل من البيانات التالي:

- طالبي العمل حسب الاختصاص والعمر؛
- ترتيب البلديات حسب عدد القرى؛
- تصنيف الولايات حسب نسب النجاح في البكالوريا؛
- عدد الحوادث المرورية خلال سنة كاملة؛
- حجم الأمطار المتساقطة في ولايات الغرب؛
- قوة الرياح في السواحل الجزائرية؛
- عدد المعاملات الإلكترونية في مجموعة من البنوك؛
- عدد الحيوانات في الحدائق؛
- نوع الآلات المستخدمة في مشروع ما؛
- تصنيف الحيوانات حسب مصدر الأكل؛
- عدد الذكر والإناث في المنزل؛
- نوعية الخدمات البريدية؛
- الحالة الصحية لمجموعة المهاجرين غير الشرعيين؛
- الأداء الرياضي لمجموعة من المحترفين؛

المحور الثاني: الجداول التكرارية والأشكال البيانية

في المنهج الإحصائي، بعد تحديد الهدف والظاهرة المدروسة والمجتمع الإحصائي المستهدف، تأتي مرحلة جمع البيانات. هذه الأخيرة تكون في شكلها الخام، أي يتمّ جمع البيانات بشكل عشوائي حيث يصعب على القارئ فهم هذه البيانات، فنلجأ إلى ترتيب هذه البيانات في جداول تكرارية تختلف أنواعها حسب المتغيرات المدروسة وعددها.

1- الجدول التكراري:

تختلف الجداول الإحصائية حسب نوع المتغيرات الإحصائية وعددها.

1-1- الجدول التكراري البسيط:

في هذه الحالة، يتمّ دراسة متغير واحد، وهو يتكون من خانتين، الخانة الأولى تسمى خانة الفئات ويرمز لها بالرمز (x_i)، وهي مختلف الترتيبات والقيم التي تأخذها المتغيرة الإحصائية المدروسة، والخانة الثانية هي خانة التكرارات ويرمز لها بالرمز (n_i) وهي عدد مرّات تكرار الفئة (x_i)، كما هو موضح في الجدول التالي:

الجدول رقم 01: نموذج عام عن الجدول التكراري البسيط

الفئات (x_i)	التكرار (n_i)
x_1	n_1
x_2	n_2
x_k	n_k
المجموع	**N**

ملاحظة: مجموع التكرارات (N) هو عدد عناصر المجتمع الإحصائي المدروس، ورياضيا:

$$\sum_{i=1}^{k} n_i = N$$

مثال رقم (01): يهدف التعرف على عدد المحاضرات التي يحضرها الطلاب في إحدى الجامعات، قام باحث بإجراء دراسة على 30 طالب فكانت النتائج التالية:

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

الحل: هذه البيانات تعتبر بيانات خام أو بيانات أولية (غير مرتبة)، وسنقوم بتنظيمها وترتيبها في جدول تكراري:

الجدول رقم (02): توزيع الطلبة حسب عدد المحاضرات التي يحضرونها

n_1	x_i
1	0
6	1
7	2
7	3
6	4
3	5
30	**المجموع**

المصدر: انطلاقا من البيانات أعلاه

ملاحظة: من حيث أنواع المتغيرات، يوجد جدول إحصائي بسيط نوعي (المتغيرة الإحصائية نوعية)، جدول إحصائي بسيط كمي متقطع (المتغيرة الإحصائية كمية متقطعة) وجدول إحصائي بسيط كمي مستمر (المتغيرة الإحصائية كمية مستمرة).

1-2- الجدول التكراري المزدوج:

في حالة دراسة متغيرتين إحصائيتين لنفس المجتمع الإحصائي، الجدول الإحصائي المناسب يسمى جدول إحصائي مزدوج، وينقسم إلى أنواع وفقا لطبيعة المتغيرات الإحصائية المدروسة: جدول إحصائي مزدوج نوعي نوعي (المتغيرتين الإحصائيتين نوعيتين)، جدول إحصائي نوعي كمي

(المتغيرتين الإحصائيتين نوعية كمية) وجدول إحصائي مزدوج كمي كمي (المتغيرتين الإحصائيتين كميتين). وفي المثال التالي نعرض شكل الجدول الإحصائي المزدوج.

مثال رقم (02): الجدول التالي يمثل تصنيف 280 طالب حسب الجنس وحسب شهادة البكالوريا:

الجدول رقم (03): تصنيف الطلبة حسب الجنس وشهادة البكالوريا

نوع الجنس / شهادة البكالوريا	ذكر	أنثى	المجموع
رياضيات	12	09	21
علوم تجريبية	23	20	43
تقني رياضي	14	11	25
تسيير واقتصاد	33	35	68
آداب وفلسفة	28	30	58
آداب ولغات	20	45	65
المجموع	**130**	**150**	**280**

المصدر: بيانات عشوائية

هذا الجدول هو جدول إحصائي مزدوج نوعي نوعي، حيث يعرض متغيرتين احصائيتين نوعيتين لنفس المجتمع الإحصائي (280 طالب).

2- دراسة متغيرة إحصائية نوعية:

2-1- الجدول الإحصائي:

كما ذكرنا سابقا، المتغيرات الإحصائية النوعية تنقسم إلى نوعين: نوعية إسمية ونوعية ترتيبية. ويتمّ تكوين الجدول الإحصائي بالنسبة للنوع الأول بطريقة عشوائية في ترتيب الفئات، أم بالنسبة للمتغيرات النوعية الترتيبية، يتم ترتيب الفئات وفق السلم التدريجي الخاص بهذه المتغيرة.

مثال رقم (03): البيانات التالية تمثل نوع العين لمجموعة من السكان:

بني	أسود	بني	أخضر	بني	أخضر	أسود	بني	أزرق	أخضر
أزرق	بني	أخضر	بني	أزرق	بني	أسود	أزرق	بني	أسود
بني	أسود	أخضر	بني	أزرق	أزرق	أخضر	أسود	بني	بني
أسود	أزرق	بني	أخضر	أخضر	بني	أسود	بني	أزرق	أسود

الجدول رقم (04): توزيع السكان حسب لون العين

x_i	n_i
بني	15
أسود	09
أخضر	08
أزرق	08
المجموع	**40**

المصدر: بيانات المثال رقم (03).

مثال رقم (04): أجرى مستثمر في الصناعات الغذائية استبيانا على مجموعة من المستهلكين لمعرفة رأيهم حول منتوج جديد عرض في السوق فكانت النتائج كالآتي:

جيد	ضعيف	جيد	ضعيف	متوسط	متوسط
ممتاز	جيد جدا	ضعيف	متوسط	جيد	متوسط
ضعيف	متوسط	ممتاز	جيد	جيد	جيد جدّا
متوسط	جيد	جيد جدا	ممتاز	ضعيف	متوسط

الجدول رقم (05): توزيع المستهلكين حسب رأيهم في المنتوج

x_i	n_i
ضعيف	05
متوسط	07
جيد	06
جيد جدّا	03
ممتاز	03
المجموع	**24**

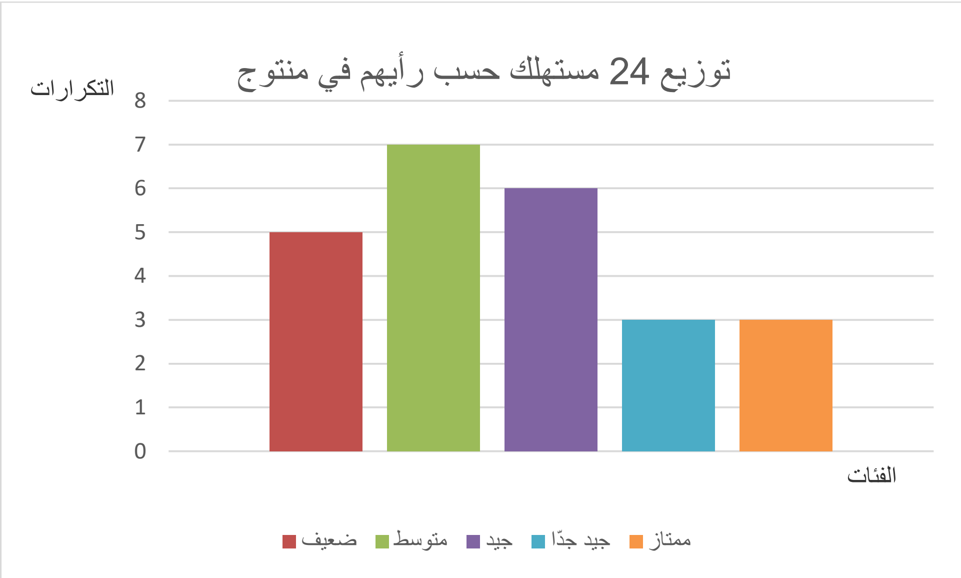
المصدر: بيانات المثال رقم (04).

2-2- التمثيل البياني:

2-2-1. الأعمدة البيانية:

سنقوم بتمثيل الجدول رقم (05) بالأعمدة البيانية

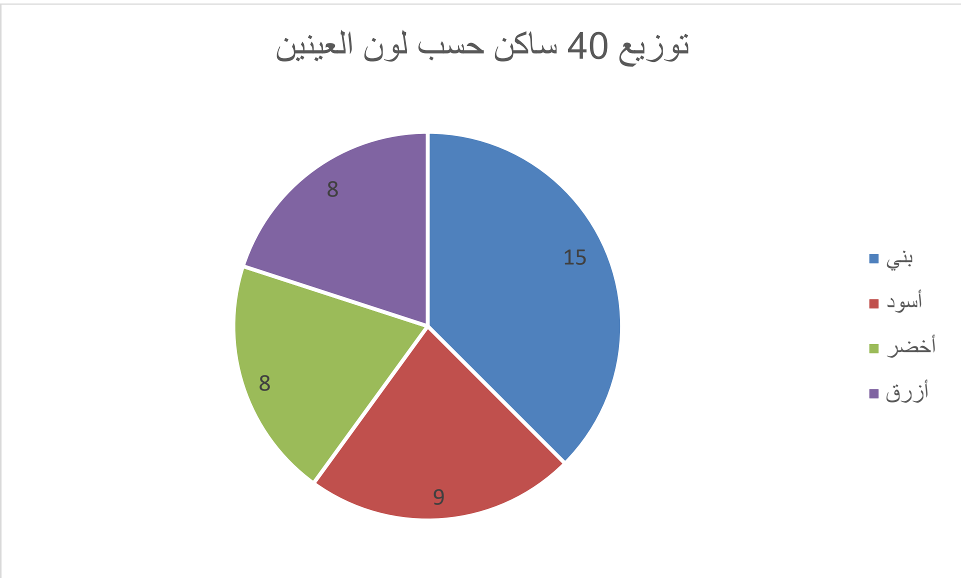
الشكل رقم (01): توزيع 24 مستهلك حسب رأيهم في منتوج



المصدر: بيانات الجدول رقم (05)

2-2-2. الدائرة النسبية:

الشكل رقم (02): توزيع 40 ساكن حسب لون العينين



المصدر: بيانات الجدول رقم (04).

3- دراسة المتغيرة الكمية المتقطعة:

كما ذكرنا سابقا، المتغيرة الكمية المتقطعة هي التي تأخذ قيما محدودة في مجال، أي مثلا تأخذ إما القيمة 4 أو القيمة 5 في المجال (4-5) كعدد الغرف في البيت، عدد الغيابات، عدد الأسهم...الخ.

3-1- الجدول الإحصائي الكمي المتقطع:

جدول التوزيع التكراري الذي يدرس المتغيرة الكمية المتقطعة يسمى جدول توزيع تكراري متقطع، ويتم ترتيب الفئات (x_i) من أصغر قيمة إلى أكبر قيمة أو العكس، وسنوضح ذلك في المثال التالي:

المثال رقم (05): قامت دراسة إحصائية لمعرفة عدد السيارات لدى كل عائلة في قرية ما فكانت النتائج كالتالي:

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

الجدول رقم (06): توزيع العائلات حسب عدد السيارات في قرية ما

x_i	n_I
0	13
1	22
2	14
3	06
4	03
5	02
المجموع	**60**

المصدر: بيانات المثال رقم (05)

حيث: x_i: عدد السيارات n_I: عدد العائلات

$$\sum_{i=1}^{n} n_i = N$$

المثال رقم (06): أجرت هيئة مختصة مسح على مستوى إحدى بلديات ولاية تلمسان لمعرفة عدد الفلاحين العاملين حسب عدد الهكتارات المستغلة فكانت النتائج في هذه السلسلة الإحصائية التالية:

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

- كوّن جدول إحصائي

الجدول رقم (07): توزيع الفلاحين حسب عدد الهكتارات المستغلة في إحدى بلديات ولاية تلمسان

x_i	n_I
1	02
2	05
3	04
4	05
5	05
6	05
7	04
المجموع	**30**

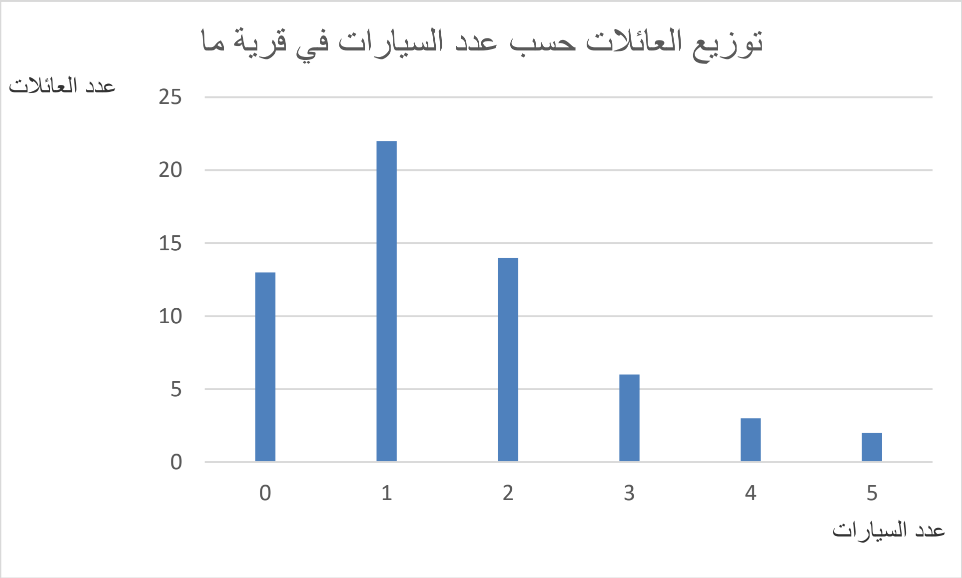
المصدر: بيانات المثال رقم (06)

حيث: x_i: عدد الهكتارات؛ n_I: عدد الفلاحين؛ N=30

3-2- التمثيل البياني:

3-2-1. الأعمدة البيانية:

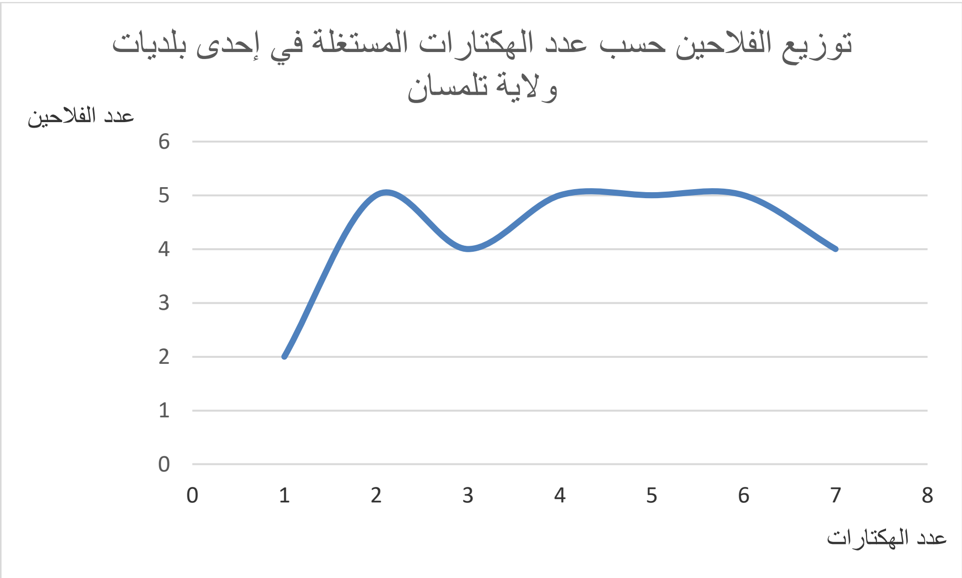
الشكل رقم (03): توزيع العائلات حسب عدد السيارات في قرية ما



المصدر: بيانات الجدول رقم (06)

3-2-2. المنحنى التكراري:

الشكل رقم (04): توزيع الفلاحين حسب عدد الهكتارات المستغلة في إحدى بلديات ولاية تلمسان



المصدر: بيانات الجدول رقم (07).

4- دراسة المتغيرة الكمية المستمرة:

كما ذكرنا سابق، المتغيرة الكمية المستمرة أو المتصلة هي المتغيرة التي تأخذ أي قيمة في مجال معين، مثلا تأخذ أي قيمة في المجال [4-5]، سواء 4؛ 4.1؛ 4.2....الخ، وتكون الوحدات التي تقاس عليها المتغيرة قابلة للتجزئة. والجدول الإحصائي الذي يكون السلسلة الإحصائية المتعلقة بهذه المتغيرات يسمى جدول توزيع تكراري كمي مستمر.

4-1- الجدول الإحصائي الكمي المستمر:

الجدول الإحصائي الكمي المستمر يتكون من فئات ذات مجال مستمر، حيث يبدأ من أصغر فئة إلى أكبر فئة، ويتم تحديد فئاته وفق قاعدة إحصائية أو نظرا لخبرة الإحصائي.

4-1-1. تحديد عدد الفئات:

يتطلب تكوين الجدول الإحصائي الكمي المستمرة تحديد عدد الفئات، حيث يتحدد هذا الأخير وفق القانون التالي الذي يعتمد على علاقة ستورجز (H. A STURGERS) كالآتي:

$$L = \frac{w}{n_c} = \frac{w}{1 + 3.322 * log_{10}(N)}$$

حيث:

L: طول الفئة

W: المدى وهو الفرق بين أكبر قيمة وأصغر قيمة

$$w = MAX_{x_i} - MIN_{x_i}$$

n_c: عدد الفئات

log_{10}: اللوغاريتم العشري

N: عدد أفراد المجتمع الإحصائي (العينة)

ومنه نستنتج أن:

$$n_c = 1 + 3.322log_{10}(N) = \text{عدد الفئات}$$

ملاحظة: يجب التنويه إلى ضرورة تدوير نتائج طول الفئة إذا كان كسرا إلى أقرب عدد بصورة عامة.

مثال رقم (07): البيانات التالية تمثل أعمار ل 30 مسافر يستقلون القطار بمحطة العاصمة:

29	30	31	27	23	18	68	55	41	22
21	32	33	34	38	53	57	52	49	44
33	29	46	37	24	23	19	15	17	20

جدول التوزيع التكراري:

- **تحديد عدد الفئات:** يجب ألا يكون عدد الفئات قليل لنفقد بعض البيانات أو كبيرة لنفقد صفة الاختصار وخلق فئات خالية من التكرار ولذا نهتم بمدى الاختلاف بين البيانات للوصول لعدد الفئات والذي يتراوح بين 5 ، 20 فئة ويمكن تطبيق القاعدة $n_c = 1 + 3.322log_{10}(N)$

$$n_c = 1 + 3.322log_{10}(30) = 5.9 \approx 6$$

- **تحديد طول الفئة:** طول الفئة هو قسمة المدى (الفرق بين أكبر قيمة وأصغر قيمة) على عدد الفئات، أي $L = \frac{w}{n_c}$

وبالتالي: W= 68-15=53

$$\boldsymbol{L = \frac{53}{6} = 8.83 \approx 9}$$

طول الفئة هو 9

الجدول رقم (08): توزيع 30 مسافر حسب العمر في محطة القطار بالعاصمة

n_i	x_i
09	]24-15]
07	]33-24]
06	]42-33]
03	]51-42]
04	]60-51]
01	]69-60]
30	**المجموع**

المصدر: بيانات المثال رقم (08).

2-1-4. أنواع الجداول الإحصائية الكمية المستمرة:

- جداول منتظمة: عندما تكون أطوال فئاتها متساوية كما في المثال السابق حيث جميع الفئات طولها هو 9؛
- جداول غير منتظمة: عندما تكون أطوال فئاتها غير متساوية كما يوضح الجدول رقم
- جداول إحصائية مغلقة: هي الجداول التكرارية التي تكون محددة المبدأ البداية والنهاية؛
- جداول إحصائية مفتوحة من الأسفل: هي الجداول التكرارية التي تكون مفتوحة المبدأ في البداية على سبيل المثال [أقل من5-10]؛
- جداول إحصائية مفتوحة من الأعلى: هي الجداول التكرارية التي تكون مفتوحة المبدأ في النهاية على سبيل المثال [40-أكثر من 60]؛
- جداول إحصائية مفتوحة من الجهتين: هي الجداول التكرارية التي تكون مفتوحة المبدأ البداية والنهاية على سبيل المثال [أقل من5-10] و [40-أكثر من 60]

جدول توزيع تكراري مغلق على الجهتين

x_i	n_i
[2-0]	5
[4-2]	4
[6-4]	9
[8-6]	2
المجموع	20

جدول توزيع تكراري مفتوح من الأسفل

x_i	n_i
[أقل من 5-10]	5
[15-10]	4
[20-15]	9
[25-20]	2
المجموع	20

جدول توزيع تكراري مفتوح من الأعلى

x_i	n_i
[20-10]	5
[30-20]	4
[40-30]	9
[40-أكثر من 50]	2
المجموع	20

جدول توزيع تكراري غير منتظم

x_i	n_i
[2-0]	5
[6-2]	4
[12-6]	9
[20-12]	2
المجموع	20

جدول توزيع تكراري مفتوح من الجهتين

x_i	n_i
[أقل من 5-10]	5
[15-10]	4
[20-15]	9
[20-أكثر من 25]	2
المجموع	20

4-1-3. مركز الفئة:

كل فئة لها حدين [a, b[حيث a هو الحد الأدنى وb هو الحد الأعلى، ومركز هذين الحدين يسمى مركز الفئة ويرمز له بالرمز c_i ونكتب:

$$c_i = \frac{a+b}{2}$$

4-1-4. الحدود الفعلية:

في حالة أن يكون الجدول التكراري مكون من فئات غير متصلة ببعضها البعض، نقوم بتحديد الحدود الفعلية للفئات، أي الحد الأدنى الفعلي و الحد الأعلى الفعلي كما يلي:

- الحد الأدنى الفعلي للفئة المطلوبة= (الحد الأعلى للفئة السابقة + الحد الأدنى للفئة المطلوبة)÷2؛
- الحد الأعلى الفعلي للفئة المطلوبة= (الحد الأدنى للفئة اللاحقة + الحد الأعلى للفئة المطلوبة)÷2.

مثال رقم (08): الجدول التالي يمثل توزيع مجموعة من الطلبة حسب الوزن

الجدول رقم (09): توزيع مجموعة من الطلبة حسب الوزن

c_i	الحدود الفعلية	الحد الأعلى الفعلي	الحدود الأدنى الفعلي	n_i	x_i
45.5	]50.5-40.5]	50.5 =2/(51+50)	40.5 =2/(40+41)	12	]50-41]
55.5	]60.5-50.5]	60.5 =2/(61+60)	50.5 =2/(50+51)	27	]60-51]
65.5	]70.5-60.5]	70.5 =2/(71+70)	60.5 =2/(60+61)	43	]70-61]
75.5	]80.5-70.5]	80.5 =2/(81+80)	70.5 =2/(70+71)	50	]80-71]
85.5	]90.5-80.5]	90.5 =2/(91+90)	80.5 =2/(80+81)	18	]90-81]
				150	المجموع

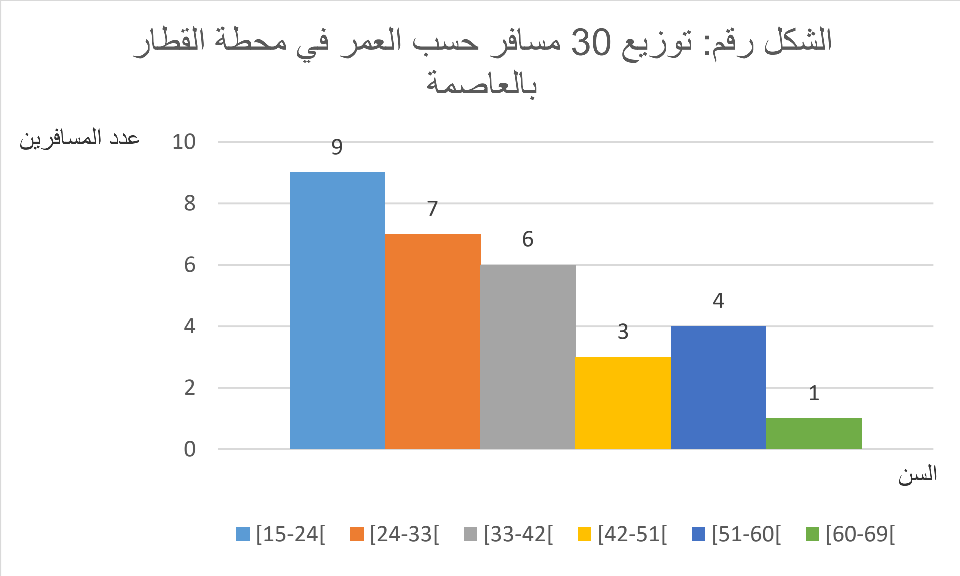
المصدر: بيانات فرضية

4-2- التمثيل البياني:

4-2-1. المدرج التكراري:

سنقوم بتمثيل الجدول رقم (08) في شكل المدرج التكراري

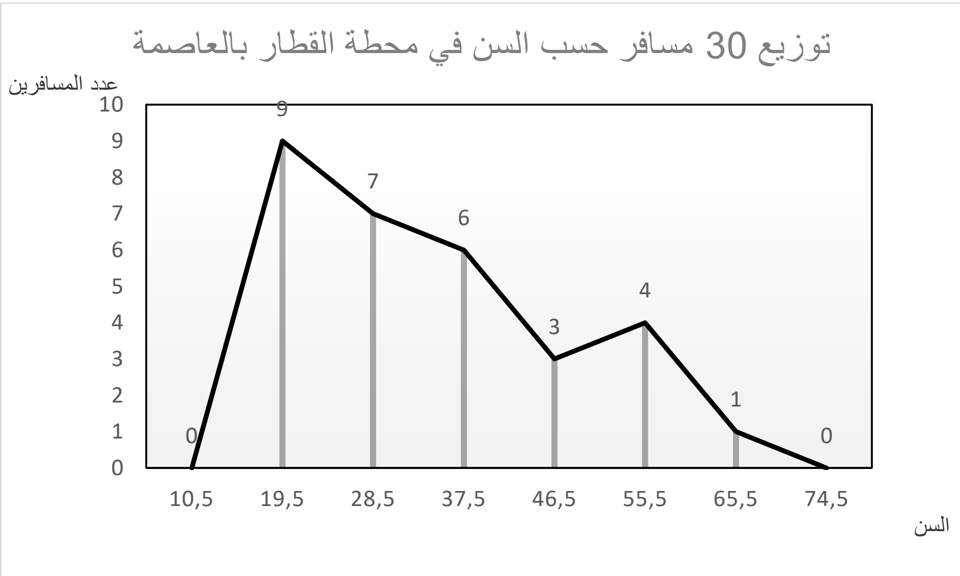
الشكل رقم (05): توزيع 30 مسافر حسب العمر في محطة القطار بالعاصمة



المصدر: بيانات الجدول رقم (08)

4-2-2. المضلع التكراري:

الشكل رقم (06): توزيع 30 مسافر حسب العمر في محطة القطار بالعاصمة



المصدر: بيانات الجدول رقم (08).

5- أنواع التكرارات:

التكرار المطلق: هو عدد مرات ظهور الفئة (x_i)، ويرمز له بالرمز n_i؛

التكرار النسبي: هو عبارة عن حاصل قسمة تكرار كل فئة إلى مجموع التكرارات، أي هو يعبر عن نسبة المفردات التي قيمتها تنتمي إلى كل فئة ويرمز له بالرمز f_iحيث: $f_i = \frac{n_i}{N}$؛

التكرار المتجمع (التجميعي) الصاعد: هو عبارة عن عدد المفردات في جدول تكراري تقل عن الحد الأعلى لكل فئة، ويرمز له بالرمز N_i^+؛

التكرار المتجمع (التجميعي) النازل: هو عبارة عن عدد المفردات في جدول تكراري تزيد عن الحد الأدنى لكل فئة، ويرمز له ب: N_i^+؛

التكرار المتجمع (التجميعي) النسبي الصاعد: يعبر عن نسبة المفردات التي تقل فيها عن الحد الأعلى لكل فئة ، وذلك من خلال تقسيم التكرار التجميعي الصاعد على مجموع التكرارات، ويرمز له ب: F_i^+؛

التكرار المتجمع (التجميعي) النسبي النازل: فهو يعطي نسبة المفردات التي تزيد قيمتها عن الحد الأدنى لكل فئة ، وذلك من خلال تقسيم التكرار التجميعي الهابط على مجموع التكرارات، ويرمز له ب: F_i^-.

والجدول التالي يوضح مختلف التكرارات:

الجدول رقم(10): أنواع التكرارات

x_i	n_i	f_i	N_i^+		N_i^-		F_i^+	F_i^-
]10-5]	4	**0.09**	4	**4**	45	**45**	0.09	1
]15-10]	7	**0.15**	7+4	**11**	41-45	**41**	0.24	0.91
]20-15]	12	**0.27**	12+11	**23**	7-41	**34**	0.51	0.76
]25-20]	8	**0.18**	8+23	**31**	12-34	**22**	0.69	0.49
]30-25]	14	**0.31**	14+31	**45**	8-22	**14**	1	0.31
المجموع	45	**1**						

المصدر: بيانات مفترضة

5-1- التكرار المعدّل:

عندما تكون الفئات المستمرة في جدول التوزيع التكراري غير منتظمة، أي طول الفئة غير متساوي في جميع الفئات، نلجأ إلى تعديل التكرارات وهذا في حالتين فقط:

- في حالة رسم المدرج التكراري؛
- في حالة حساب المنوال (أحد مقاييس النزعة المركزية).

ويتم تعديد التكرارات وفق طريقتين:

الطريقة الأولى: هي الطريقة المشهورة وفق العلاقة ($n_i^* = \frac{n_i}{L}$) حيث،

n_i^*: التكرار المعدل؛ n_i: التكرار المطلق؛ L: طول الفئة x_i

الطريقة الثانية: وفق العلاقة $n_i^* = \frac{(n_i \times L_i)}{L_i^*}$ حيث:

n_i^*: التكرار المعدل؛ n_i: التكرار الفعلي للفئة غير المعدلة؛ L_i: طول الفئة المنتظمة؛ L_i^*: طول الفئة غير المنتظمة

ملاحظة: نلجأ إلى تعديل التكرارات في حالة رسم المدرج التكراري، لأن الفئات المتساوية في أطوالها، فإن المدرج التكراري سيكون عبارة عن مستطيلات متلاصقة ومتساوية القاعدة وارتفاعاتها تتناسب مع التكرار. أما إذا كانت الفئات غير متساوية الطول تكون مساحة هذه المستطيلات المتلاصقة غير متناسبة مع التكرار وكذلك ارتفاعاتها، لذلك يجب تعديل التكرار قبل رسم المدرج التكراري للفئات غير المتساوية حتى يصبح التكرار المعدل متناسبا مع ارتفاع المستطيل الخاص بالفئة غير منتظمة الطول.

الجدول التالي يوضح ذلك:

الجدول رقم (11): تعديل التكرارات

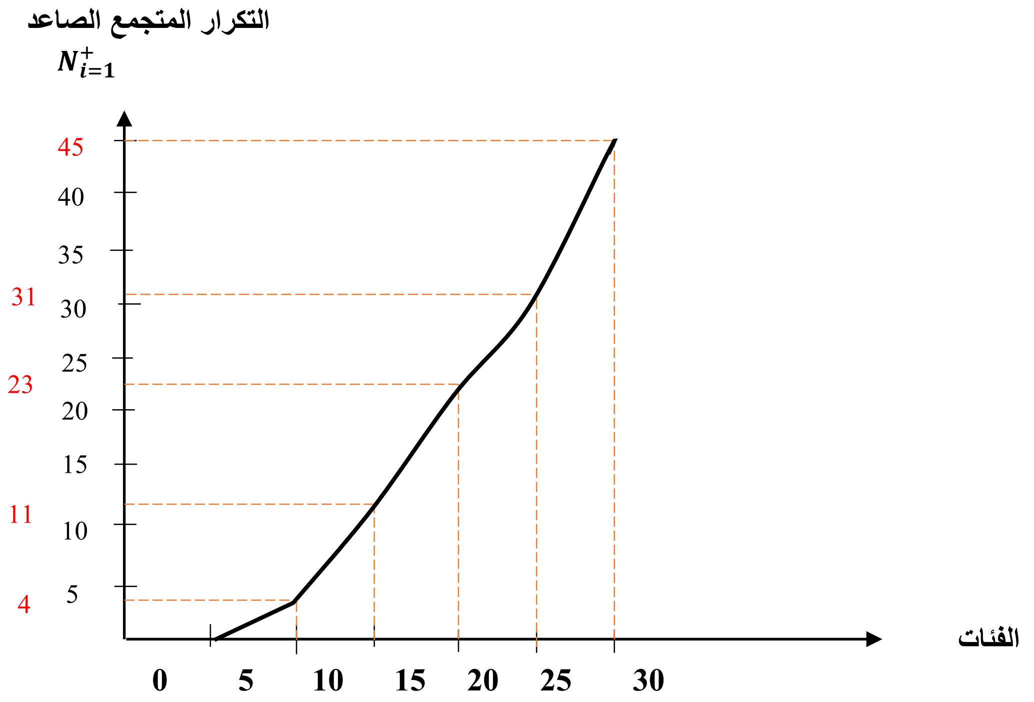
x_i	n_i	L_i	الطريقة الأولى	n_i^*	الطريقة الثانية	n_i^*
]10-0]	120	10	=10/120	12	=10/(10*120)	120
]15-10]	40	5	=5/40	8	=5/(10*40)	80
]30-15]	150	15	=15/150	10	=15/(10*150)	100
]40-30]	20	10	=10/20	2	=10/(10*20)	20
]50-40]	60	10	=10/60	6	=10/(10*60)	60

المصدر: بيانات فرضية.

5-2- التمثيل البياني:

5-2-1. المنحنى التجميعي الصاعد:

الشكل رقم (07): المنحنى المتجمع الصاعد

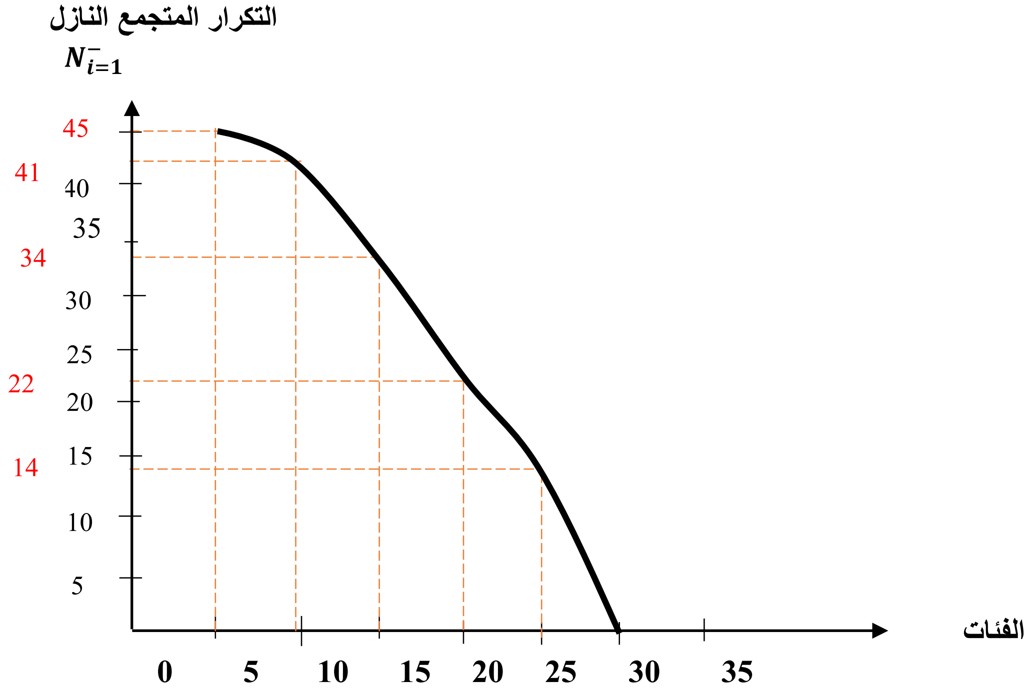


المصدر: بيانات الجدول رقم (10).

ملاحظة: المنحنى المتجمع الصاعد في حالة جدول توزيع تكراري كمي مستمر، يُرسم من خلال تعيين الإحداثيات بين الحد الأعلى للفئة والتكرار المتجمع الصاعد للفئة المقابلة، (مع تعيين الحد الأعلى للفئة السابقة الوهمية التي تسبق الفئة الأولى الرئيسية مع القيمة الصفر (0) للتكرار المتجمع الصاعد).

5-2-2. المنحنى المتجمع النازل:

الشكل رقم (08): المنحنى المتجمع النازل



المصدر: بيانات الجدول رقم (10)

ملاحظة: المنحنى المتجمع النازل في حالة جدول توزيع تكراري كمي مستمر، يُرسم من خلال تعيين الإحداثيات بين الحد الأدنى للفئة والتكرار المتجمع النازل للفئة المقابلة، (مع تعيين الحد الأدنى للفئة اللاحقة الوهمية التي تأتي بعد الفئة الأخيرة الرئيسية مع القيمة الصفر (0) للتكرار المتجمع النازل).

6- تمارين محلولة:

6-1- التمرين الأول: ليكن لدينا التوزيع الآتي الذي يمثل توزيع الكتب حسب الأسعار بالدينار الجزائري في إحدى المكتبات:

x_i	n_i
]200-100]	15
]300-200]	25
]400-300]	70
]500-400]	90
]600-500]	80
]700-600]	80

المطلوب:

- حدّد المجتمع الإحصائي، الوحدة الإحصائية والمتغيرة الإحصائية ونوعها
- كون جدول توزيع تكراري نسبي
- كون جدول توزيع تكراري تجميعي صاعد ونازل
- كون جدول توزيع تكراري نسبي صاعد ونازل
- ما هي نسبة الكتب التي سعرها على الأكثر 400 دج؟
- ما هي نسبة الكتب التي سعرها على الأقل 300 دج؟

6-2- التمرين الثاني:

أخذت أطوال لعينة من 30 طالب من جامعة معسكر فكانت كالآتي:

155	173	184	164	166	158	180	177	174	170
182	174	175	169	171	176	159	166	167	181
176	179	182	174	178	157	176	180	167	168

المطلوب:

- حدد المجتمع الإحصائي، الوحدة الإحصائية والمتغيرة الإحصائية ونوعها؛
- كون جدول توزيع تكراري مقسم على خمسة فئات
- كون جدول توزيع تكراري نسبي ثم مثله بيانيا بالمدرج التكراري
- كون جدول توزيع تكراري تجميعي صاعد ومثله بيانيا.

7- الحل:

7-1- التمرين الأول:

المجتمع الإحصائي	الوحدة الإحصائية	المتغيرة الإحصائية	نوعها
360 كتاب في إحدى المكتبات	الكتاب	السعر	كمية مستمرة

الجدول رقم (12): توزيع 360 كتاب حسب السعر في إحدى المكتبات

الفئات	التكرار المطلق	التكرار النسبي	التكرار المتجمع الصاعد	التكرار المتجمع النازل	التكرار النسبي المتجمع الصاعد	التكرار النسبي المتجمع النازل
x_i	n_i	f_i	N_i^+	N_i^-	F_i^+	F_i^-
]200-100]	15	0.042	15	360	0.042	1
]300-200]	25	0.07	40	345	0.112	0.958
]400-300]	70	0.194	110	320	0.306	0.888
]500-400]	90	0.25	200	250	0.556	0.694
]600-500]	80	0.222	280	160	0.778	0.444
]700-600]	80	0.222	360	80	1	0.222
المجموع	360	1				

المصدر: بيانات التمرين رقم 6-1.

- نسبة الكتب التي سعرها على الأكثر 400 دج هي 30.6% (0.306 *100%).

ملاحظة: عبارة على الأكثر معناه أننا لا نتجاوز القيمة المعطاة، أي في هذه الحالة لا نتجاوز 400 دج، وإحصائيا نختار التكرار المتجمع الصاعد المقابل لهذه الفئة، أي التكرار المتجمع النسبي الصاعد للفئة]300 - 400].

- نسبة الكتب التي سعرها على الأقل 300 دج هي 88.8% (0.888*100%).

ملاحظة: عبارة على الأقل معناه أن الحد الأدنى هي القيمة المعطاة، أي في هذه الحالة السعر الأدنى للكتب هو 300 دج، وإحصائيا نختار التكرار المتجمع النازل لهذه القيمة، أي التكرار المتجمع النسبي النازل للفئة]300-400].

2-7- التمرين الثاني:

المجتمع الإحصائي	الوحدة الإحصائية	المتغيرة الإحصائية	نوعها
30 طالب في إحدى الجامعات	الطالب	الطول	كمية مستمرة

- جدول التوزيع التكراري :

تحديد طول الفئة:

طول الفئة = $\frac{\text{المدى}}{\text{عدد الفئات}}$؛ $L_i = \frac{w}{n_c}$

المدى = الفرق بين أقصى قيمة وأدنى قيمة = 184-155= 29

عدد الفئات = 5

طول الفئة = 29÷5= 5.8 ≈6

L_i=6

تكوين التوزيعات التكرارية

الجدول رقم (13): توزيع 30 طالب حسب الطول في إحدى الجامعات

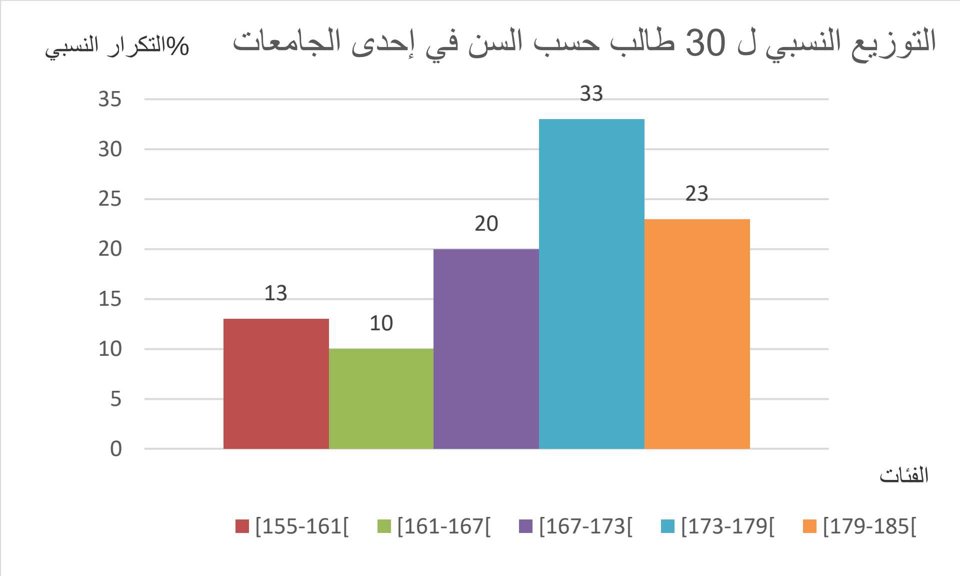
الفئات	التكرار المطلق	التكرار النسبي	التكرار المتجمع الصاعد	التكرار المتجمع النازل	التكرار النسبي المتجمع الصاعد	التكرار النسبي المتجمع النازل
x_i	n_i	f_i	N_i^+	N_i^-	F_i^+	F_i^-
]161-155]	4	0.13	4	30	0.13	1
]167-161]	3	0.1	7	26	0.23	0.87
]173-167]	6	0.2	13	23	0.43	0.77
]179-173]	10	0.33	23	17	0.76	0.57
]185-179]	7	0.23	30	7	1	0.23
المجموع	**30**	**1**				

المصدر: بيانات التمرين رقم 6-2.

- **التمثيل البياني:**

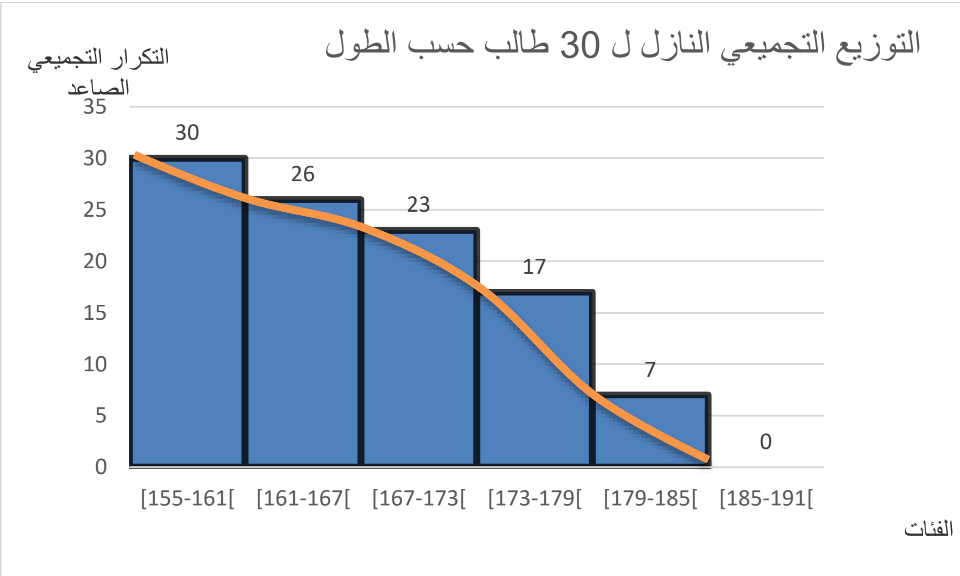
المدرج التكراري النسبي

الشكل رقم (09): التوزيع النسبي ل 30 طالب حسب السن في إحدى الجامعات



المصدر: بيانات الجدول رقم (13).

الشكل رقم (10): التوزيع التجميعي النازل ل 30 طالب حسب الطول



المصدر: بيانات الجدول رقم (13).

8- تمارين غير محلولة:

8-1- التمرين الأول:

البيانات التالية تمثل درجات مجموعة من الطلبة في إحدى المقاييس الدراسية:

51	59	70	74	73	90
91	72	83	89	50	80
62	82	87	76	91	76
71	96	81	88	64	82
80	81	75	85	74	90

المطلوب:

- حدّد المجتمع الإحصائي، الوحدة الإحصائية والمتغيرة الإحصائية ونوعها؛
- كون جدول توزيع تكراري ثم مثله بيانيا؛
- كون جدول توزيع تكراري تجميعي نازل ثم مثله بيانيا؛
- ما هي نسبة الطلبة الذين تحصلوا على الأقل 45 درجة؟

8-2- التمرين الثاني:

البيانات التالية تمثل أوزان مجموعة من الرياضيين بالكيلوغرام في أحد المركبات الرياضية بولاية تلمسان:

67	64	68	73	73	54	61	74	60	78
80	74	65	63	60	69	72	66	77	65
74	50	76	69	68	66	78	63	70	55
67	67	64	76	61	72	72	57	65	77
59	71	79	78	58	63	74	66	73	67
61	71	69	68	73	81	64	61	84	55

المطلوب:

- حدّد المجتمع الإحصائي والمتغيرة الإحصائية ونوعها؟
- كون جدول إحصائي مقسم إلى سبعة فئات
- مثل بيانيا هذا الجدول بواسطة المظلع التكراري
- أحسب التكرارات النسبية التجميعية الصاعدة والنازلة ومثلها بيانيا بالمنحنيات التجميعية على نفس المعلم
- ما هي نسبة الرياضيين الذين يزنون على الأكثر 70 كلغ؟
- ما هي نسبة الرياضيين الذين يزنون على الأقل 60 كلغ؟

8-3- التمرين الثالث:

الجدول التالي يمثل التوزيع التكراري ل 100 طفل حسب مقاسات الحذاء:

المقاسات	16-	20-	26-	30-	34-32	المجموع
عدد الأطفال	10	24	30	22	14	100

المطلوب:

- مثل بيانيا هذا التوزيع بالمدرج التكراري؛
- أحسب التكرارات النسبية؛
- ما هي نسبة الأطفال التي مقاساتهم على الأقل 26؟
- كون جدول توزيع تكراري تجميعي صاعد ومثله بيانيا؟

المحور الثالث: مقاييس النزعة المركزية

تكلمنا في الفصول السابقة عن جمع المعلومات الإحصائية عن الظاهرة المدروسة وتبويبها في جداول إحصائية وتمثيلها في أشكال بيانية، إلاّ أن المعلومات التي نحصل عليها غير قابلة للتحليل العلمي الأولي لأنها تتشكل من قيم مطلقة غير قابلة للمقارنة بين عناصر الظاهرة المدروسة في المقارنات الآنية أو الزمنية لعدم وجود أساس مشترك، لذلك تمثل الخطوات السابقة من الوجهة الإحصائية المرحلة التمهيدية للتحليل الإحصائي الذي يهدف إلى الحصول على نتائج دقيقة وواضحة، ويلجأ الإحصائي إلى تطبيق أساليب إحصائية تساعد بدقة في فهم الظاهرة المدروسة واتخاذ القرارات السليمة فيما يتعلق باتجاهها وذلك من خلال حساب مقاييس النزعة المركزية، مقاييس التشتت وحساب مقاييس الالتواء والتفلطح.

1- مفهوم مقاييس النزعة المركزية:

مقاييس النزعة المركزية هي مقاييس عددية تستخدم لقياس موضع تركز أو تجمع البيانات؛ إذ أن بيانات أي ظاهرة تنزع في الغالب إلى التركز والتجمع حول قيم معينة. هذه القيم هي ما يسمى بمقاييس النزعة المركزية. فعند تحليل التوزيعات التكرارية سواء كانت منفصلة أو متقطعة أو مقارنة توزيعين تكراريين أو أكثر، يجب بالضرورة تلخيص كل توزيع من هذه التوزيعات التكرارية بمقاييس إحصائية وصفية لأجل الوصف الدقيق لهذه التوزيعات. فهذه المقاييس تستخدم لتلخيص البيانات عدديا إذ أنها تعتبر قيم نموذجية أو مثالية للبيانات. وإحدى خصائص التوزيعات التكرارية هي ميل مفردات إلى التمركز حول قيمة معينة والتناقص التدريجي عن كل من جانبي قيمة أحد مقاييس النزعة المركزية. وكل مقياس من المقاييس المذكورة يقيس نقطة النزعة المركزية بطريقة مختلفة عن الأخرى، لذا يجب على الباحث الإحصائي اختيار المتوسط الملائم الذي يصف شكل النزعة المركزية وذلك تبعا لطبيعة البيانات الإحصائية ونوعها. وهذا ما يتطلب منا التعرف على ماهية كل مقياس من المقاييس المشار إليها، وسنقتصر في منهجنا هذا على المقاييس المذكورة أعلاه (الوسط الحسابي، الوسيط، المنوال).

إن مقاييس النزعة المركزية تعتبر مهمة نظرا لقدرتها على إعطاء فكرة عامة وسهلة وواضحة عن قيم الظاهرة المدروسة، وسوف نقسم هذه المقاييس إلى الآتي:

- مقاييس كمية وتشمل: الوسط الحسابي، الوسط الربيعي، الوسط الهندسي والوسط التوافقي؛
- مقاييس موضعية أو مكانية وتشمل: الوسيط، الربيعات، العشيرات والمؤينات؛
- مقاييس تكرارية وتشمل: المنوال.

2- الوسط الحسابي:

يعتبر الوسط الحسابي من أهم وأفضل مقاييس النزعة المركزية ومن أكثرها شيوعا واستخداما في التحليل الإحصائي وذلك لما يتمتع به من خصائص وصفات إحصائية جيدة. ويعرف المتوسط الحسابي على أنه حاصل قسمة مجموع مفردات السلسلة على عددها. ولإيجاد المتوسط الحسابي للبيانات فإننا لابد أن نفرق بين البيانات غير المبوبة (غير ملخصة في جدول تكراري) والبيانات المبوبة (جدول التوزيع التكراري).

2-1- البيانات غير المبوبة:

إذا كانت لدينا السلسلة الإحصائية التالية: $X = \{x_1, x_2, \dots x_n\}$، فإن وسطها الحسابي $\bar{X}$ سيكون:

$$\mathbf{\overline{X_s} = \frac{x_1 + x_2 + x_3 + \cdots x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i}$$

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

مثال رقم (9): البيانات التالية تمثل نقاط 10 من الطلبة في مقياس الإحصاء الرياضي:

12	05	08	13	14	11	07	10	12	15

ما هي النقطة المتوسطة؟

$$\bar{X} = \frac{x_1 + x_2 + \cdots x_{10}}{10} = \frac{1}{10}\sum_{i=1}^{10} x_i$$

$$\bar{X} = \frac{12 + 5 + 8 + 13 + 14 + 11 + 7 + 10 + 12 + 15}{10} = \frac{107}{10} = \frac{1}{10}\sum_{i=1}^{10} 107 = 10.7$$

النقطة المتوسطة هي 10.7.

ملاحظة:

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

$$\sum_{i=1}^{n} (x_i + y_i) = \sum_{i=1}^{n} x_1 + \sum_{i=1}^{n} y_i$$

$$\sum_{i=1}^{n} cx_i = c\sum_{i=1}^{n} x_i$$

$$\sum_{i=1}^{n} c = nc$$

2-2- البيانات المبوبة:

إذا كانت لدينا بيانات عددها N ملخصة في جداول توزيعات تكرارية حيث:

- عدد الفئات x_i هو k $(x_1, x_2, \dots x_k)$؛
- تكرارات الفئات هي: $n_1, n_2, \dots n_k$؛ فإن:

$\boldsymbol{x_i \times n_i}$	$\boldsymbol{n_i}$	$\boldsymbol{x_i}$
$x_1 \times n_1$	n_1	x_1
$x_2 \times n_2$	n_2	x_2
:	:	:
:	:	:
$x_k \times n_k$	n_k	x_k
$\boldsymbol{\sum_{i=1}^{k} x_i \times n_i}$	**N**	**المجموع**

فإن المتوسط الحسابي $\bar{X}$ سيحسب بالعلاقة التالية:

$$\bar{X} = \frac{x_1 \times n_1 + x_2 \times n_2 + \cdots + x_k \times n_k}{n_1 + n_2 + \cdots + n_k} = \frac{\sum_{i=1}^{k} x_i \times n_i}{N}$$

مثال رقم (10): يمثل الجدول التالي توزيع 100 بيت حسب عدد الغرف في إحدى البلديات:

عدد الغرف	1	2	3	4	5	6	المجموع
عدد البيوت	8	14	32	23	16	7	100

المطلوب: حساب متوسط عدد الغرف لكل بيت؟

لحساب المتوسط الحسابي نلخص البيانات التي نحتاجها في هذا الجدول الإحصائي:

الجدول رقم (14): توزيع 100 بيت حسب عدد الغرف في إحدى البلديات

عدد الغرف x_i	عدد البيوت n_i	$x_i \times n_i$
1	8	8
2	14	28
3	32	96
4	23	92
5	16	80
6	7	42
المجموع	100	346
	N	$\sum_{i=1}^{6} x_i \times n_i$

وبالتطبيق العددي للنتائج المحسوبة في الجدول فإن:

$$\bar{X} = \frac{\sum_{i=1}^{k} x_i \times n_i}{N} = \frac{346}{100} = 3.46 \cong 3$$

وبالتالي، متوسط عدد الغرف في البيوت هو 3.

هذا المثال المدروس هو عبارة عن متغيرة كمية متقطعة، أمّا اذا كانت المتغيرة المدروسة كمية مستمرة، أي الفئات عبارة عن مجال فإن الوسط الحسابي هو:

$$\bar{X} = \frac{c_1 \times n_1 + c_2 \times n_2 + \cdots + c_k \times n_k}{n_1 + n_2 + \cdots + n_k} = \frac{\sum_{i=1}^{k} c_i \times n_i}{N}$$

حيث c_i هو مركز كل فئة x_i.

مثال رقم (11): الجدول التالي يمثل توزيع أطوال 130 طالب في إحدى المعاهد بالعاصمة بالسنتيمتر:

الطول x_i	]158-150]	]166-158]	]174-166]	]182-174]	]190-182]	المجموع
عدد الطلبة n_i	22	30	45	20	13	**130**

المطلوب: حساب متوسط الطول للطلبة؟

سنلخص البيانات التي نحتاجها في حساب المتوسط الحسابي في هذا الجدول:

الجدول رقم (15): توزيع 130 طالب حسب الطول (سم)

الطول x_i	عدد الطلبة n_i	مركز الفئة c_i	$c_i \times n_i$
]158-150]	22	154	3388
]166-158]	30	162	4860
]174-166]	45	170	7650
]182-174]	20	178	3560
]190-182]	13	186	2418
المجموع	**130**		**21876**
	N		$\sum_{i=1}^{k} c_i \times n_i$

وبالتطبيق العددي على نتائج الجدول فإن:

$$\bar{X} = \frac{\sum_{i=1}^{5} c_i \times n_i}{N} = \frac{21876}{130} = 168.3$$

ومنه فإن متوسط الطول لدى الطلبة هو 168.3 سم.

2-3- الوسط الحسابي الفرضي:

أحيانا يلجأ الباحث الإحصائي إلى حساب الوسط الحسابي بطريقة غير مباشرة يسمى الوسط الحسابي الفرضي، ويتم ذلك عند معالجة عدد كبير من البيانات حول ظاهرة ما، حيث يحسب وفق العلاقة التالية:

- **في حالة البيانات غير المبوبة:**

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

حيث: $\bar{X}$ هو الوسط الحسابي؛ Z هو الوسط الحسابي المفترض مسبقا.

- **في حالة البيانات المبوبة:**

$$\bar{X} = Z + \frac{\sum_{i=1}^{k}[(x_i - Z) \times n_i]}{N}$$

وفي حالة دراسة متغيرة كمية مستمرة، نعوض مراكز الفئات في القانون.

مثال رقم (12): نفس المثال السابق رقم (09)، نطبق قانون الوسط الفرضي.

12 05 08 13 14 11 07 10 12 15

$$\bar{X} = 10.$$

Z= 12

x_i	12	5	8	13	14	11	7	10	12	15	المجموع
$x_i - Z$	0	-7	-4	1	2	-1	-5	-2	0	3	-13

$$\bar{X} = Z + \frac{\sum_{i=1}^{10}(x_i - Z)}{N} = 12 + \left(\frac{-13}{10}\right) = 12 - 1.3 = 10.7$$

ومنه متوسط النقاط لذى الطلبة هو 10.7.

مثال رقم (13): نفس المثال رقم (10) المذكور أعلاه، نقوم بحساب الوسط الحسابي غير المباشر

$\bar{X} \cong 3$ Z= 4

عدد الغرف x_i	عدد البيوت n_i	$x_i - Z$	$(x_i - Z)n_i$
1	8	-3	-24
2	14	-2	-28
3	32	-1	-32
4	23	0	0
5	16	1	16
6	7	2	14
المجموع	**100**		**-54**
	N		$\sum_{i=1}^{6}((x_i - Z)n_i)$

$$\bar{X} = Z + \frac{\sum_{i=1}^{6}((x_i - Z)n_i)}{N} = 4 + \frac{(-54)}{100} = 4 - 0.54 = 3.46 \cong 3$$

ومنه متوسط الغرف في البيوت هو 3.

2-4- خصائص الوسط الحسابي:

- المجموع الجبري لانحرافات القيم عن وسطها الحسابي $\bar{X}$ يساوي الصفر، أي أنّ:

- **بيانات غير مبوبة:**

$$\sum_{i=1}^{n}(x_i-\bar{X})=(x_1-\bar{X})+(x_2-\bar{X})+\cdots+(x_n-\bar{X})=0$$

- **بيانات غير مبوبة:**

$$\sum_{i=1}^{k}(x_i-\bar{X})n_i=(x_1-\bar{X})n_1+(x_2-\bar{X})n_2+\cdots+(x_k-\bar{X})n_k=0$$

حيث أن القيمة $(x_i-\bar{X})$ تمثل انحراف القيمة x_i عن وسطها الحسابي $\bar{X}$.

- لا تتغير قيمة الوسط الحسابي إذا (ضربنا، جمعنا، طرحنا أو قسمنا) جميع قيم المشاهدات x_i بقيم ثابتة أو كسرية، أي أن:

بيانات مبوبة	بيانات غير مبوبة
$\bar{X}\mp c=\frac{\sum_{i=1}^{n}(x_i\mp c)n_i}{N}$	$\bar{X}\mp c=\frac{\sum_{i=1}^{n}(x_i\mp c)}{n}$
$\bar{X}\times c=\frac{\sum_{i=1}^{n}(x_i\times c)\times n_i}{N}$	$\bar{X}\times c=\frac{\sum_{i=1}^{n}(x_i\times c)}{n}$
$\bar{X}\div c=\frac{\sum_{i=1}^{n}(x_i\div c)\times n_i}{N}$	$\bar{X}\div c=\frac{\sum_{i=1}^{n}(x_i\div c)}{n}$

مثال رقم (14): ليكن لدينا القيم التالية: 3، 7، 4، 8، 9، 11

$$\bar{X}=\frac{3+7+4+8+9+11}{6}=\frac{42}{6}=7$$

إذا أضفنا العدد 2 لجميع القيم فإن:

$$\bar{X}=7+2=\frac{(3+2)+(7+2)+(4+2)+(8+2)+(9+2)+(11+2)}{6}=\frac{54}{6}$$
$$=9$$

إذا ضربنا جميع القيم في العدد 3 فإن:

$$\bar{X} = 7 \times 3 = \frac{(3 \times 3) + (7 \times 3) + (4 \times 3) + (8 \times 3) + (9 \times 3) + (11 \times 3)}{6} = \frac{126}{6}$$
$$= 21$$

مثال رقم (15): ليكن لدينا التوزيع التكراري التالي:

$(x_i \times 4) \times n_i$	$x_i \times 4$	$x_i \times n_i$	n_i	x_i
24	8	6	3	2
24	12	6	2	3
96	16	24	6	4
144		**36**	**12**	**المجموع**

$$\bar{X} = \frac{36}{12} = 3$$

عندما نضرب جميع الفئات في العدد 4، سيصبح الوسط الحسابي $\bar{X}$ يساوي:

$$\bar{X} \times 4 = \frac{\sum_{i=1}^{3}((x_i \times 4) \times n_i)}{N} = \frac{144}{12} = 12 = 3 \times 4$$

- إذا كان لدينا مجموعتين من البيانات بحيث أن عدد عناصر المجموعة الأولى n_1 ووسطها الحسابي $\bar{X}_1$، وعدد عناصر المجموعة الثانية n_2 ووسطها الحسابي $\bar{x}_2$، فإن متوسط المجموعة الكلية المكونة من دمج هذين المجموعتين مع بعض يمكن حسابه بالعلاقة التالية:

$$\bar{X} = \frac{n_1\bar{X}_1 + n_2\bar{X}_2}{n_1 + n_2}$$

مثال رقم (16): يوجد فوجين من الطلبة، الفوج الأول يتكون من 25 طالب، وعلامته المتوسطة هي 12.5، وفوج ثاني يتكون من 35 طالب وعلامته المتوسطة هي 11.5، فما هي العلامة المتوسطة لمجموع الطلبة؟

$$\bar{X} = \frac{n_1\bar{X}_1 + n_2\bar{X}_2}{n_1 + n_2} = \frac{25 \times 12.5 + 35 \times 11.5}{25 + 35} = \frac{715}{60} = 11.9$$

العلامة المتوسطة لمجموع الطلبة هي 11.9.

- مجموع مربعات انحرافات القيم عن وسطها الحسابي هو أصغر من مجموع انحرافات القيم عن أي نقطة أخرى، أي:

$$\sum_{i=1}^{n}(x_i-\bar{X})^2<\sum_{i=1}^{n}(x_i-c)^2 \in c\neq \bar{X}$$

2-5- مميزات وعيوب الوسط الحسابي

مميزات الوسط الحسابي	عيوب الوسط الحسابي
- سهل التعريف وعملية الحساب بسيطة في إجراء العمليات الجبرية؛ - يأخذ جميع البيانات بعين الإعتبار؛ - متوسط واحد لمجموع البيانات الواحدة؛ -	- يتأثر بالقيم الشاذة؛ - لا يمكن حسابه في حالة البيانات النوعية؛

3- المتوسط الهندسي (G):

إنه نوع من المتوسطات ويشاع استخدامه في حساب معدلات النمو، حساب الأرقام القياسية للأسعار، حساب متوسطات النسب...الخ، وحسابه يكون وفق العلاقة التالية:

3-1- في حالة البيانات غير مبوبة:

$$G=\sqrt[n]{x_1\times x_2\times\ldots\times x_n}=\left[\prod_{i=1}^{n}x_1\right]^{\frac{1}{n}}$$

ولتبسيط العمليات الحسابية، يرفع الجذر ويؤخذ اللوغاريتم العشري للطرفين، أي:

$$LogG=\frac{1}{n}\log(x_1\times x_2\times\ldots x_n)=\frac{1}{n}(logx_1+logx_2+\cdots+logx_n)=\frac{1}{n}\sum_{i=1}^{n}logx_i$$

مثال رقم (17): حساب المتوسط الهندسي للقيم التالية: 10، 08، 06، 05، 12

$$G = \sqrt[5]{10 \times 8 \times 6 \times 5 \times 12} = (10 \times 8 \times 6 \times 5 \times 12)^{5} = 7.8$$

أو باستخدام اللوغاريتم العشري:

$$LogG = \frac{1}{5}(log10 + log8 + log6 + log5 + log12) = 0.89 \Rightarrow G = 10^{0.89} = 7.8$$

2-3 حالة البيانات المبوبة:

- **البيانات الكمية المتقطعة:** إذا كان البيانات على شكل جدول توزيع تكراري كمي متقطع، فالمتوسط الهندسي يعبر يحسب بالعلاقة التالية:

$$G = \sqrt[N]{{x_1}^{n_1} \times {x_2}^{n_2} \times \ldots \times {x_k}^{n_k}} = \left[\prod_{i=1}^{k}\left({x_i}^{n_i}\right)\right]^{\frac{1}{N}}$$

وباستخدام اللوغاريتم العشري في كلا الطرفين، نحصل على:

$$LogG = \frac{1}{N}(n_i log x_i)$$

مثال رقم (18): حساب المتوسط الهندسي للتوزيع التالي:

المجموع	6	5	4	3	x_i
14	2	3	5	4	n_i

$$G = \sqrt[14]{3^4 \times 4^5 \times 5^3 \times 6^2} = (373248000)^{\frac{1}{14}} = 4$$

وباستخدام اللوغاريتم:

x_i	3	4	5	6	المجموع
n_i	4	5	3	2	14
$logx_i$	0.47	0.6	0.7	0.77	
$n_i logx_i$	1.88	3	2.1	1.54	8.52

$$LogG = \frac{1}{N}\sum_{i=1}^{k} n_i logx_i = \frac{1}{14}(8.52) = 0.6 \Rightarrow G = 10^{0.6} = 4$$

- حالة البيانات الكمية المستمرة: في حالة وجود توزيع تكراري كمي مستمر، نفس المعادلة السابقة أعله مع تعويض x_i ب C_i و هو مركز الفئة، حيث نكتب:

$$G = \sqrt[N]{c_1^{n_1} \times c_2^{n_2} \times \ldots \times c_k^{n_k}} = \left[\prod_{i=1}^{k} c_i^{n_i}\right]^{\frac{1}{N}}$$

وباستخدام اللوغاريتم نكتب:

$$LogG = \frac{1}{N}\sum_{i=1}^{k}(n_i logc_i)$$

4- المتوسط التوافقي (H):

هذا المتوسط عادة يستخدم عند حساب معدل التغير، السرعة ومتوسط الأسعار متى أعطيت على أساس الوحدات، ويحسب بالصيغ التالية:

4-1- بيانات غير مبوبة:

يحسب المتوسط التوافقي وفق الصيغة التالية:

$$H = \frac{N}{\sum_{i=1}^{n}(\frac{1}{x_i})}$$

ويعرف المتوسط التوافقي على أنه مقلوب المتوسط الحسابي لمقلوبات قيم السلسلة.

مثال رقم (19): حساب الوسط التوافقي للقيم التالية: 3، 7، 6، 4، 5، 8

$$H = \frac{6}{\frac{1}{3}+\frac{1}{7}+\frac{1}{6}+\frac{1}{4}+\frac{1}{5}+\frac{1}{8}} = \frac{6}{1.205} \cong 5$$

4-2- بيانات مبوبة:

- **حالة بيانات كمية متقطعة:** يحسب المتوسط التوافقي في حالة التوزيعات التكرارية الكمية المتقطعة وفق الصيغة التالية:

$$H = \frac{N}{\sum_{i=1}^{k}(\frac{n_i}{x_i})}$$

- **حالة بيانات كمية مستمرة:** يحسب المتوسط التوافقي في حالة التوزيعات التكرارية الكمية المستمرة وفق الصيغة التالية:

$$H = \frac{N}{\sum_{i=1}^{k}(\frac{n_i}{c_i})}$$

مثال رقم (20): حساب المتوسط التوافقي للتوزيع التالي:

x_i	]8-4]	]12-8]	]16-12]	]20-16]	]24-20]	المجموع
n_i	3	2	6	4	1	16
c_i	6	10	14	18	22	
$\frac{n_i}{c_i}$	0.5	0.2	0.43	0.22	0.045	1.395

$$H = \frac{N}{\sum_{i=1}^{k}(\frac{n_i}{x_i})} = \frac{16}{1.395} = 11.5$$

5- المتوسط التربيعي (Q):

يستخدم هذا المتوسط في الحالات التي لا نستطيع استخدام المتوسط الهندسي (عندما تتضمن قيم سالبة) ولا نستطيع استخدام المتوسط الحسابي (عندما تلغى القيم السالبة والموجبة بعضها البعض)، ويحسب وفق الصيغ التالية:

5-1- بيانات غير مبوبة:

يعرف بأنه الجذر التربيعي لمتوسط مربعات قيم سلسلة إحصائية ما، ويحسب بالعبارة التالية:

$$Q = \sqrt[2]{\frac{\sum_{i=1}^{n} x_i^2}{n}}$$

مثال رقم (20): حساب المتوسط التربيعي للبيانات التالية: 4، 6، 2، 5، 8

$$H = \sqrt{\frac{8^2 + 5^2 + 2^2 + 6^2 + 4^2}{5}} = \sqrt{\frac{145}{5}} \cong 5.4$$

5-2- بيانات مبوبة:

- **حالة البيانات الكمية المتقطعة:** في حالة هذا التوزيع التكراري يحسب المتوسط التربيعي بالعلاقة التالية:

$$Q = \sqrt{\frac{\sum_{i=1}^{k} n_i x_i^2}{N}}$$

- **حالة البيانات الكمية المستمرة:** نفس الصيغة الرياضية بالنسبة للمعادلة السابقة مع تعويض x_i ب c_i:

$$Q = \sqrt{\frac{\sum_{i=1}^{k} n_i c_i^2}{N}}$$

مثال رقم (21): حساب المتوسط التربيعي للجدول التالي:

x_i	1	2	3	4	5	المجموع
n_i	4	3	2	6	5	20
x_i^2	1	4	9	16	25	
$n_i x_i^2$	4	12	18	96	125	255

$$\mathbf{Q} = \sqrt{\frac{\sum_{i=1}^{k} n_i c_i^2}{N}} = \sqrt{\frac{255}{20}} = 3.6$$

ملاحظة: هناك علاقة بين $\bar{X}, G, H, Q$ حيث:

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

$$\mathrm{G} = \sqrt{\bar{X} \times H}$$

6- الوسيط *(Me)*:

إن الوسيط يمثل قيمة موضعية حيث يصنف البيانات إلى النصف، وهي القيمة المفردة التي يتساوى عندها عدد المشاهدات السابقة مع عدد المشاهدات اللاحقة. ولحسابه، يتطلب ترتيب المشاهدات ترتيبا تصاعديا أو تنازليا؛ كما يعتبر هذا المقياس من أكثر المقاييس تفضلا، نظرا لعدم تأثر قيمته بالقيم الشاذة والمتطرفة في التوزيع. كما يتطلب حساب الوسيط الخطوتين التاليتين:

أولا: تحديد رتبة الوسيط؛

ثانيا: تعيين قيمة الوسيط

ويتم حساب الوسيط بطرق مختلفة في الحالات التالية كما يلي:

6-1- حالة البيانات غير المبوبة:

- إذا كان عدد البيانات فردي: في حالة دراسة مجتمع إحصائي يكون عدد مفرداته (N) فردي فإن:

أولا: رتبة الوسيط تساوي $C_{me} = \frac{N+1}{2}$

حيث C_{me} هي رتبة الوسيط، وN هي عدد قيم المجتمع الإحصائي.

ثانيا: نرتب القيم ترتيبا تصاعديا أو تنازليا ونحدد قيمة الوسيط، حيث $\mathrm{Me} = X_{\frac{N+1}{2}}$.

مثال رقم (21): لدينا سلسلة إحصائية تمثل معدلات بعض الطلبة في إحدى الأقسام:

8 9 11 13 9 10 10 12 11 9 10

أولا: رتبة الوسيط هي $C_{me} = \frac{N+1}{2} = \frac{11+1}{2} = 6$

ثانيا: نرتب القيم تصاعديا ثم نحدد قيمة الوسيط

					$C_{me} = 6$					
13	12	11	11	10	**10**	10	9	9	9	8

ومنه قيمة الوسيط تساوي 10، أي $Me = 10$.

➢إذا كان عدد البيانات زوجي: في حالة دراسة مجتمع إحصائي يكون عدد مفرداته (N) عدد زوجي فإن:

أولا: رتبة الوسيط تساوي $C_{me} = \frac{N}{2}$

ثانيا: نرتب القيم ترتيبا تصاعديا أو تنازليا، ونحدد قيمة الوسيط، حيث $Me = \frac{Me_1+Me_2}{2}$

حيث: $Me_2 = X_{\frac{N}{2}+1}$; $Me_1 = X_{\frac{N}{2}}$

مثال رقم (22): نفس المثال السابق مع إضافة معدل طالب آخر:

11 8 9 11 13 9 11 10 12 11 9 10

أولا: رتبة الوسيط هي: $C_{me} = \frac{N}{2} = \frac{12}{2} = 6$

ثانيا: ترتيب القيم تصاعديا أو تنازليا وتحديد قيمة الوسيط:

					$Me_2 = X_{\frac{N}{2}+1}$	$Me_1 = X_{\frac{N}{2}}$					
13	12	11	11	11	11	10	10	9	9	9	8

ومنه قيمة الوسيط تساوي $\boldsymbol{Me = \frac{Me_1+Me_2}{2} = \frac{10+11}{2} = 10.5}$

ملاحظة: إذا كان عدد المشاهدات كبير، يقدر الوسيط على أنه القيمة ذات الترتيب $\frac{N}{2}$ أيا كانت N سواء فردية أو زوجية.

2-6- حالة البيانات المبوبة:

- **حالة المتغيرة الكمية المتقطعة (المنفصلة):** إذا كان لدينا توزيع تكراري كمي متقطع، تُحدد قيمة الوسيط وفق الخطوات التالية:

أولا: حساب التكرار المتجمع الصاعد؛

ثانيا: تحديد رتبة الوسيط C_{me} حيث:

إذا كان عدد القيم فردي $C_{me}=\frac{N}{2}$ إذا كان عدد القيم زوجي $C_{me}=\frac{N+1}{2}$

ثالثا: تحديد قيمة الوسيط **Me** في الحالتين التاليتين:

قيمة الوسيط = قيمة الفئة التي تقابل التكرار المتجمع الصاعد	إذا كانت رتبة الوسيط تساوي أحد قيم التكرار المتجمع الصاعد
قيمة الوسيط = قيمة الفئة التي تقابل التكرار المتجمع الصاعد الأكبر	إذا كانت رتبة الوسيط محصورة بين التكرار المتجمع الأصغر و التكرار المتجمع الأكبر

مثال رقم (23): ليكن لدينا التوزيع التكراري التالي:

x_i	2	3	4	5	6
n_i	4	2	3	4	1

حساب الوسيط:

$N_{i=1}^{+}$	n_i	x_i
4	4	2
6	2	3
9	3	4
12	4	5
14	1	6
	14	المجموع

رتبة الوسيط = $C_{me} = \frac{14}{2} = 7$

التكرار المتجمع الصاعد الأصغر ←

التكرار المتجمع الصاعد الأكبر ←

قيمة الوسيط = **Me = 4**

6-3- حالة متغيرة كمية مستمرة:

إذا كان لدينا توزيع كمي مستمر، يُحسب الوسيط باتباع الخطوات التالية:

أولا: حساب التكرار المتجمع الصاعد $N_{i=1}^{+}$؛

ثانيا: تعيين الفئة الوسيطية من خلال رتبة الوسيط C_{me} حيث $C_{me} = \frac{N}{2}$؛

(الفئة الوسيطية هي الفئة التي تقابل أول تكرار متجمع صاعد أكبر من رتبة الوسيط)

ثالثا: حساب قيمة الوسيط بالصيغة التالية:

$$Me = d + \left(\frac{C_{me} - N_{i-1}^{+}}{n_{me}}\right) \times L$$

حيث:

d	الحد الأدنى للفئة الوسيطية
C_{med}	رتبة الوسيط
N_{i-1}^{+}	التكرار المتجمع الصاعد السابق للفئة الوسيطية (مجموع التكرارات التي تسبق الفئة الوسيطية)
n_{me}	التكرار المطلق للفئة الوسيطية
L	طول الفئة الوسيطية

مثال رقم (24): ليكن التوزيع التالي:

x_i	]4-2]	]8-4]	]14-8]	]20-14]
n_i	5	7	12	6

حساب الوسيط:

	x_i	n_i	$N_{i=1}^{+}$	رتبة الوسيط	$C_{me}=\frac{N}{2}=\frac{30}{2}=15$ ←
	]4-2]	5	5		
	]8-4]	7	12	15	← التكرار المتجمع الصاعد السابق للفئة الوسيطية N_{i-1}^{+}
					→ رتبة الوسيط
الفئة الوسيطية	**]14-8]**	**12**	**24**		← أول تكرار متجمع صاعد أكبر من رتبة الوسيط N_{i+1}^{+}
	]20-14]	6	30		$n_{me}=12\ ;\ d=8\ ;\ L=6$
	المجموع	30			

$$Me=d+\left(\frac{C_{me}-N_{i-1}^{+}}{n_{me}}\right)\times L=8+\left(\frac{15-12}{12}\right)\times 6=9.5$$

قيمة الوسيط هي 9.5، أي أن 50% من المشاهدات أقل من 9.5.

4-6- تعيين قيمة الوسيط بيانيا:

يمكن تحديد قيمة الوسيط بيانيا وهذا من خلال المراحل التالية:

- حالة أولى:

أولا: رسم المنحنى المتجمع الصاعد أو النازل؛

ثانيا: تحديد رتبة الوسيط على المحور العمودي (محور التكرارات) ثم إسقاطها على المنحنى؛

ثالثا: إسقاط النقطة المعينة الأخيرة في المنحنى على محور الفواصل (محور الفئات) وتحديد قيمة الوسيط مباشرة.

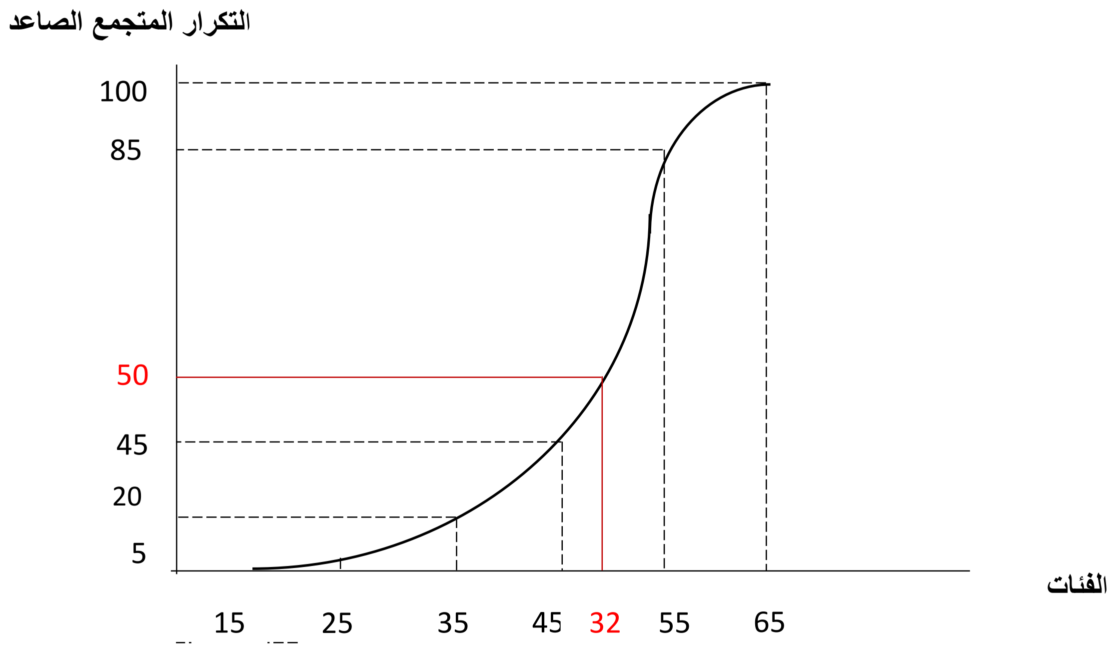
- حالة ثانية:

أولا: رسم المنحنى المتجمع الصاعد والنازل في نفس الشكل؛

ثانيا: إسقاط نقطة التقاطع بين المنحنيين على محور الفواصل وتحديد قيمة الوسيط مباشرة.

مثال رقم (25): ليكن التمثيل التالي يمثل توزيع 100 عامل حسب العمر:

الشكل رقم (11): التوزيع المتجمع الصاعد ل 100 عامل حسب العمر



من خلال المعطيات، نستنتج أن رتبة الوسيط تساوي 50، ومن تتحدد قيمة الوسيط من خلال الشكل والتي تساوي 32 سنة، أي أن 50 عامل أعمارهم تقل عن 32 سنة.

7- الربيعات، العشيرات والمؤينات:

على غرار الوسيط، تعتبر هذه المقاييس بمقاييس موضعية حيث تقسم موضع البيانات حسب كل مؤشر. ويتم حسابها بنفس طريقة حساب الوسيط، إلاّ أن الاختلاف يكمن فقط في تحديد الرتبة (سواء رتبة الربيع، العشير أو المؤين)، ويتم تحديد هذه الرتب كما يلي:

7-1- الربيعات Q:

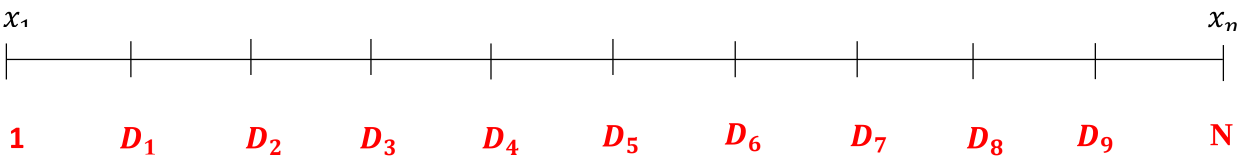
يوجد ثلاثة ربيعات، حيث كل ربيع يقسم البيانات إلى الربع ورتبهم محددة كما يلي:

رتبة الربيع الأول	رتبة الربيع الثاني	رتبة الربيع الثالث
$N \times \frac{1}{4}$	$N \times \frac{2}{4}$	$N \times \frac{3}{4}$
Q_1	Q_2	Q_3

7-2- العشيرات (D):

يوجد تسعة عشيرات مرقمة من 1 إلى 9، حيث كل عشير يقسم البيانات إلى العشر ورتبهم محددة كما يلي:

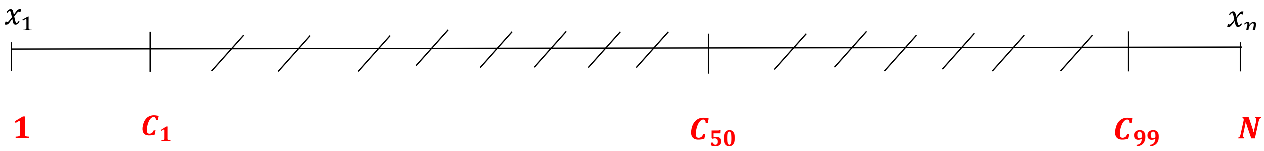
D_1	D_2	D_3	D_4	D_5	D_6	D_7	D_8	D_9	العشير
$N \times \frac{1}{10}$	$N \times \frac{2}{10}$	$N \times \frac{3}{10}$	$N \times \frac{4}{10}$	$N \times \frac{5}{10}$	$N \times \frac{6}{10}$	$N \times \frac{7}{10}$	$N \times \frac{8}{10}$	$N \times \frac{9}{10}$	الرتبة



7-3- المؤينات (C):

يوجد 99 مؤين وكل واحد يقسم البيانات إلى 1 من مائة ورتبهم محددة كما يلي:

المؤين	C_{99}	C_{98}	....	C_{50}	....	C_2	C_1
الرتبة	$N\times\frac{99}{100}$	$N\times\frac{98}{100}$	....	$N\times\frac{50}{100}$	....	$N\times\frac{2}{100}$	$N\times\frac{1}{100}$



8- المنوال (MO):

تعرف قيمة المنوال على أنها أكثر المفردات (المشاهدات) شيوعا أو أكبرها تكرارا، وتعرف بالقيمة السائدة أو القيمة المسيطرة. ويتطلب وجود قيمة المنوال على توفر تكرارات مختلفة لعدد المفردات المكونة للمجتمع الإحصائي المدروس. كما يعتبر المنوال أقل مقاييس النزعة المركزية استخداما نظرا لعدم وجود فائدة في حسابه عندما يكون عدد قليل من المفردات، كما يوجد في بعض الحالات أكثر من منوالين، ويتم تحديد المنوال في الحالات التالية كما يلي:

8-1- بيانات غير مبوبة:

إذا كانت البيانات غير مبوبة، قيمة المنوال هي المفردة التي تتكرر أكثر.

مثال رقم (26): لتكن لدينا السلسلة الإحصائية التالية:

5	9	5	10	12	7	7	5	11	10

قيمة المنوال تساوي 5، لأنها المفردة التي تتكرر أكثر (3 مرات).

مثال (27): لتكن لدينا البيانات التالية:

3	4	9	4	5	4	5	3	3	4	8	3	5

عندنا منوالين في هذا المثال، الأول يساوي 3 والثاني يساوي 4.

8-2- بيانات مبوبة:

- **حالة متغيرة كمية متقطعة:** اذا كان هناك توزيع تكراري كمي متقطع، فإن قيمة المنوال هي قيمة الفئة التي تقابل أكبر تكرار مباشرة.

مثال رقم (28): ليكن لدينا التوزيع التكراري التالي:

x_i	9	10	11	12	13	14	**المجموع**
n_i	4	22	14	33	17	20	**110**

قيمة المنوال (Mo) تساوي 12 (تقابل أكبر تكرار).

- **حالة متغيرة كمية مستمرة:** إذا كان هناك توزيع تكراري كمي مستمر، فإن قيمة المنوال تحسب بالصيغة التالية:

$$Mo = d + \left(\frac{n_0 - n_1}{(n_0 - n_1) + (n_0 - n_2)}\right) \times L$$

حيث:

d	الحد الأدنى للفئة المنوالية
n_0	تكرار الفئة المنوالية
n_1	التكرار السابق للفئة المنوالية
n_2	التكرار اللاحق للفئة المنوالية
L	طول الفئة المنوالية
الفئة المنوالية هي الفئة التي تقابل أكبر تكرار	

مثال رقم (29): ليكن لدينا التوزيع التكراري التالي:

x_i	]5-0]	]10-5]	]15-10]	]20-15]	]25-20]	]30-25]
n_i	4	7	15	33	20	14

حساب المنوال:

الفئة المنوالية هي الفئة التي تقابل أكبر تكرار وهي [15-20[		
$$Mo = d + \left(\frac{n_0 - n_1}{(n_0 - n_{1)} + (n_0 - n_2)}\right) \times L$$	15	d
	33	n_0
$$Mo = 15 + \left(\frac{33 - 15}{(33 - 15) + (33 - 20)}\right) \times 5 \cong 18$$	15	n_1
	20	n_2
	5	L

3-8- تحديد قيمة المنوال بيانيا:

على غرار الوسيط، يمكن تحديد قيمة المنوال بيانيا ويتم تحديد ذلك من خلال المراحل التالية:

أولا: رسم المدرج التكراري للتوزيع التكراري المدروس؛

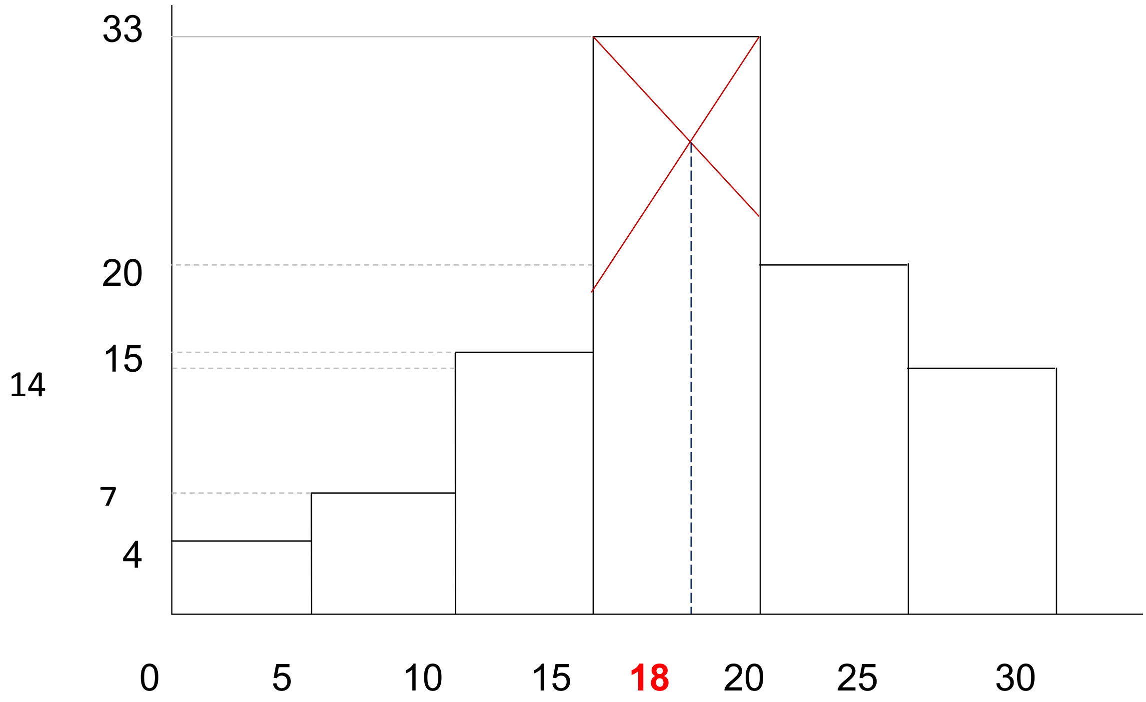
ثانيا: تحديد الفئة المنوالية التي تقابل أكبر تكرار (أطول مستطيل)؛

ثالثا: رسم خطين داخل مستطيل الفئة المنوالية، حيث الأول يبدأ من الحد الأدنى للفئة المنوالية وينتهي بالحد الأدنى للفئة الموالية، والثاني يبدأ بالحد الأعلى للفئة المنوالية وينتهي بالحد الأعلى للفئة السابقة؛

رابعا: نقطة تقاطع الخطين نسقطها على محور الفواصل ونحدد قيمة المنوال.

مثال رقم (30): نحسب قيمة المنوال للمثال السابق بالطريقة البيانية:

الشكل رقم (12): تمثيل التوزيع التكراري السابق بالمدرج التكراري



من المنحنى تتحدد قيمة المنوال التي تساوي 18.

ملاحظة: إذا كان التوزيع التكراري المستمر غير منتظم (أي طول الفئة غير متساوي بين جميع الفئات)، يتطلب تعديل التكرارات في حساب المنوال سواء بالطريقة الرياضية أو البيانية (أنظر طريقة تعديل التكرارات في المحاور السابقة).

9- تمارين محلولة:

9-1- التمرين الأول:

لتكن البيانات التالية تمثل مبيعات السيارات الشهرية في عدد من الوكالات:

12 9 8 10 7 8 7 10 7 10 11 8 11 10

المطلوب:

- حدد المجتمع الإحصائي والمتغيرة الإحصائية ونوعها؛

- أحسب متوسط عدد السيارات المباعة شهريا (الوسط الحسابي)؛
- أحسب الوسط التوافقي؛
- عين قيمة الربع الثالث؛
- حدد قيمة الوسيط والمنوال

9-2- التمرين الثاني:

ليكن الجدول التالي الذي يمثل توزيع عدد من المساكن حسب عدد الغرف:

عدد الغرف	1	2	3	4	5	المجموع
عدد المساكن	20	26	35	37	22	140

المطلوب:

- أحسب الوسط الحسابي والوسط الهندسي؛
- أحسب قيمة الوسيط والعشير الثالث؛
- عين قيمة المنوال.

9-3- التمرين الثالث:

الجدول التالي يمثل توزيع السكان (بالمليون) حسب العمر:

الفئات(سنة)	]20 – 10]	]30 – 20]	]40 – 30]	]50 – 40]	]60 – 50]
التكرار(عدد السكان)	4	11	13	8	6

المطلوب:

- حدد قيمة الوسط الحسابي؛
- أحسب الوسط التوافقي والتربيعي؛
- أحسب الوسيط، الربيع الأول والمؤين رقم 60؛
- أحسب قيمة المنوال.

10- الحل:

10-1- التمرين الأول:

12	9	8	10	7	8	7	10	7	10	11	8	11	10

المجتمع الإحصائي: 14 وكالة بيع السيارات؛

المتغيرة الإحصائية ونوعها: عدد السيارات المباعة (متغيرة كمية متقطعة)؛

حساب الوسط الحسابي $\bar{X}$:

$$\bar{X} = \frac{12+9+8+10+7+8+7+10+7+10+11+8+11+10}{14} = 9.14$$

متوسط مبيعات السيارات يساوي 9 سيارات.

حساب قيمة الوسط التوافقي H:

$$H = \frac{14}{\frac{1}{12}+\frac{1}{9}+\frac{1}{8}+\frac{1}{10}+\frac{1}{7}+\frac{1}{8}+\frac{1}{7}+\frac{1}{10}+\frac{1}{7}+\frac{1}{10}+\frac{1}{11}+\frac{1}{8}+\frac{1}{11}+\frac{1}{10}} \cong 9$$

الوسط التوافقي يساوي 9.

حساب الربيع الثالث Q_3:

رتبة الربيع Q_3 هي: $N \times \frac{3}{4} = 14 \times \frac{3}{4} \cong 11 \leftarrow$ $Q_3 = 10$

7	7	7	8	8	8	9	10	10	10	10	11	11	12

قيمة الربيع الثالث تساوي 10، ومنه 75% من الوكالات لا تتجاوز مبيعاتها 10 سيارات.

حساب قيمة الوسيط **Me**:

رتبة الوسيط $= \frac{N}{2} = \frac{14}{2} = 7$ ومنه:

$$Me_1 = 9\ ;\ Me_2 = 10 \Rightarrow Me = \frac{9+10}{2} = 9.5$$

قيمة الوسيط تساوي 9.5 ومنه 50% من الوكالات لا تتجاوز مبيعاتها 9 سيارات.

حساب قيمة المنوال **Mo**:

المنوال = 10 = **Mo**

10-2- التمرين الثاني:

x_i	n_i	$x_i \times n_i$	$logx_i$	$n_i \times logx_i$	$N_{i=1}^{+}$
1	20	20	0	0	20
2	26	52	0.3	7.8	46
3	35	105	0.48	16.8	81
4	37	148	0.6	22.8	118
5	22	110	0.7	15.4	140
المجموع	**140**	**435**		**62.8**	
	N	$\sum_{i=1}^{5} x_i \times n_i$		$\sum_{i=1}^{5} n_{i \times logx_i}$	

حساب الوسط الحسابي:

$$\bar{X} = \frac{\sum_{i=1}^{5} x_i \times n_i}{N} = \frac{435}{140} = 3.10$$

متوسط عدد الغرف في المساكن هو 3 غرف.

حساب الوسط الهندسي:

$$logG = \frac{1}{N}\sum_{i=1}^{5}(n_i logx_i) = \frac{1}{140} \times 62.8 = 0.44 \Rightarrow G = 2.75$$

متوسط الرف في المساكن هو 3 غرف.

حساب الوسيط

رتبة الوسيط هي $\frac{N}{2} = \frac{140}{2} = 70$

ومنه قيمة الوسيط تساوي 3 غرف، أي أن 50% من المساكن لا تتجاوز عدد غرفها 3.

حساب العشير الثالث D_3:

رتبة العشير الثالث هي $N \times \frac{3}{10} = 140 \times \frac{3}{10} = 42$

ومنه قيمة العشير الثالث تساوي 2 $(D_3 = 2)$، أي أن 30% من المساكن لا يتجاوز عدد غرفها 2.

تحديد قيمة المنوال:

قيمة المنوال تساوي 4.

3-10- التمرين الثالث:

الفئات(سنة)	$]20-10]$	$]30-20]$	$]40-30]$	$]50-40]$	$]60-50]$	المجموع
التكرار (عدد السكان)	4	11	13	8	6	42

x_i	n_i	C_i	$C_i \times n_i$	$\frac{C_i}{n_i}$	${c_i}^2$	$n_i \times {c_i}^2$	$N_{i=1}^{+}$
$]20-10]$	4	15	60	0.26	225	900	4
$]30-20]$	11	25	275	0.44	625	6875	15
$]40-30]$	13	35	455	0.37	1225	15925	28
$]50-40]$	8	45	360	0.17	2025	16200	36
$]60-50]$	6	55	330	0.11	3025	18150	42
المجموع	42		1480	1.352		58050	
	N		$\sum_{i=1}^{5} C_i \times n_i$	$\sum_{i=1}^{5} \frac{n_i}{c_i}$		$\sum_{i=1}^{5} n_i \times {c_i}^2$	

حساب الوسط الحسابي:

$$\bar{X} = \frac{\sum_{i=1}^{5} c_i \times n_i}{N} = \frac{1480}{42} = 35.23$$

متوسط العمر هو 35 سنة.

حساب الوسط التوافقي:

$$H = \frac{N}{\sum_{i=1}^{5} \frac{n_i}{c_i}} = \frac{42}{1.35} = 31.11$$

متوسط العمر هو 31 سنة.

حساب الوسط الحسابي التربيعي:

$$H = \sqrt{\frac{\sum_{i=1}^{5} n_i \times {c_i}^2}{N}} = \sqrt{\frac{58050}{42}} = 37.2$$

متوسط العمر هو 37 سنة.

حساب قيمة الوسيط:

رتبة الوسيط تساوي $\frac{N}{2} = \frac{42}{2} = 21$

الفئة الوسيطية هي [30-40[

$$Me = d + \left(\frac{\frac{N}{2} - N_{i-1}^{+}}{n_{me}}\right) \times L = 30 + \left(\frac{21 - 15}{13}\right) \times 10 = 34.61$$

قيمة الوسيط تساوي 34.61، أي أن 50% (21 مليون نسمة) لا يتجاوز سنهم 35 سنة.

حساب الربيع الأول:

رتبة الربيع الأول Q_1 هي $N \times \frac{1}{4} = \frac{42}{4} = 10.5$

فئة الربيع الأول هي [20-30[

$$Q_1 = d + \left(\frac{\frac{N}{4} - N_{i-1}^{+}}{n_{Q_1}}\right) \times L = 20 + \left(\frac{10.5 - 4}{11}\right) \times 10 = 25.9$$

قيمة الربيع الأول تساوي 25.9، أي أن 25% من السكان لا يتجاوز سنهم 30 سنة.

حساب المؤين رقم 60 (C_{60}):

رتبة المؤين 60 هي $N \times \frac{60}{100} = 42 \times \frac{60}{100} = 25.2$

فئة المؤين 60 هي [30-40[

$$C_{60} = d + \left(\frac{\frac{60N}{100} - N_{i-1}^{+}}{n_{C_{60}}}\right) \times L = 30 + \left(\frac{25.2 - 15}{13}\right) \times 10 = 37.84$$

قيمة المؤين 60 تساوي 37.84، أي أن 60% من النسمة لا يتجاوز سنهم 38 سنة.

حساب قيمة المنوال:

أكبر تكرار مطلق هو 13، إذن الفئة المنوالية هي [30-40[

$$Mo = d_0 + \left(\frac{(N_0 - N_1)}{(N_0 - N_1) + (N_0 - N_2)}\right) = 30 + \left(\frac{13 - 11}{(13 - 11) + (13 - 8)}\right) \times 10 = 32.85$$

قيمة المنوال تساوي 32.85، ومنه العمر السائد في المجتمع هو 33 سنة.

11- تمارين غير محلولة:

11-1- التمرين الأول:

لتكن السلسلة الإحصائية التالية تمثل عدد الولادات الحديثة خلال أسبوع في أحد مستشفيات الأمومة والطفولة في إحدى ولايات الجزائر:

8 5 7 11 6 4 3

المطلوب:

- أحسب متوسط عدد الولادات في الأسبوع؟
- أحسب الوسط التوافقي، الهندسي والتربيعي؟
- أحسب قيمة الوسيط والربيع الأول؟
- عين قيمة المنوال.

11-2- التمرين الثاني:

التوزيع التكراري التالي يمثل توزيع الهكتارات حسب عدد الفلاحين في منطقة فلاحية:

عدد الهكتارات	1	2	3	4	5	6
عدد الفلاحين	12	11	22	14	11	7

المطلوب:

- حدد المجتمع الإحصائي، الوحدة الإحصائية والمتغيرة الإحصائية ونوعها؟
- أحسب الوسط الحسابي؟
- أحسب الوسط التوافقي والوسط التربيعي ثم قارن بين النتائج؟
- أحسب قيمة الوسيط والربيع الثالث؟
- عين قيمة المنوال.

11-3- التمرين الثالث:

الجدول التالي بمثل توزيع الأبقار في مزرعة حسب الوزن (الكيلوغرام):

الوزن	]50-70]	]70-100]	]100-150]	]150-200]	]200-250]
عدد الأبقار	10	14	30	35	11

المطلوب:

- حدد المجتمع الإحصائي والمتغيرة الإحصائية ونوعها؟
- أحسب الوسط الحسابي، الوسط التربيعي؟
- أحسب قيمة الوسيط وقيمة العشير السابع؟
- أحسب قيمة المنوال؟

المحور الرابع: مقاييس التشتت

في الفصل السابق، تطرقنا إلى مقاييس النزعة المركزية، وهي مقاييس تقيس مدى تمركز القيم حول قيمة معينة أو مؤشر معين، إلا أن هذه المقاييس غير كافية لإعطاء نظرة كاملة عن حقيقة المجموعة أو المجتمع الإحصائي المدروس الذي يمثله هذا التوزيع، فقد نجد القيم تتباعد وتتباعد عن القيمة المركزية بأبعاد مختلفة، وهو ما يعرف عن تشتت القيم عن هذه المقاييس المركزية، ولمعرفة ذلك، نلجأ إلى حساب بعض المقاييس تسمى مقاييس التشتت وهي قيم مكملة لمقاييس النزعة المركزية.

1- مفهوم مقاييس التشتت:

مقاييس التشتت تقيس مدى تشتت أو انتشار قيم الظاهرة المدروسة حول أحد قيم المتوسطات (أحد مقاييس النزعة المركزية). ويهدف قياس التشتت إلى قياس درجة التباعد بين مفردات القيم، وبالتالي يحدد درجة التجانس والمصداقية بين مفرداتها.

وبشكل عام، كلما قلّ التشتت زادت درجة التجانس بين المفردات، أي يكون المتوسط يمثل أحسن تمثيل لهذا التوزيع والعكس بالعكس. وتصنف مقاييس التشتت إلى قسمين:

- مجموع المقاييس التي توضح مدى وكيفية اختلاف القيم فيما بينها، أي تقيس تباعد (تقارب) أو انتشار المشاهدات من خلال قياس الاختلاف بين قيمتين في مجموع القيم كمقياس المدى أو مقياس الانحراف الربيعي؛
- مجموع المقاييس التي تقيس انحرافات أو اختلافات القيم عن أحد المتوسطات كالانحراف المتوسط أو الانحراف المعياري؛

يتم حساب هذه المقاييس بشكل مطلق أو نسبي؛ فمقاييس التشتت المطلقة تُحسب في حالة شمولية وحدة القياس بالنسبة لجميع الوحدات حتى في حالة وجود مجتمعات إحصائية كثيرة، ومقاييس التشتت النسبية تُحسب في حالة عدم شمولية وحدة القياس بالنسبة للوحدات الإحصائية. وأهم مقاييس التشتت نعرضها كما يلي:

2- المدى:

يعتبر هذا المقياس من أبسط مقاييس التشتت، ويعرف بأنه: " الفرق بين أكبر قيمة وأصغر قيمة في مجموعة واحدة من القيم"؛ ويكتب:

$$E = X_{Max} - X_{Min}$$

مثال رقم (31): البيانات التالية تمثل نقاط بعض الطلبة في مقياس الإحصاء:

10	14	9	5	16	8	10	11	10	9

المدى هو الفرق بين أكبر قيمة (16) وأصغر قيمة (5)، أي

$$E = X_{max} - X_{min} = 16 - 5 = 11$$

ملاحظة: المدى هو سهل الحساب إلا أنه لا يستخدم لأنه يتأثر بقوة بالقيم الشاذة ولا يأخذ بعين الاعتبار جميع قيم المجموعة.

في حالة وجود توزيع تكراري مستمر، يُحسب المدى بحساب الفرق بين الحد الأعلى للفئة الأخيرة والحد الأدنى للفئة الأولى (أو مركز الفئة الأخيرة ومركز الفئة الأولى).

3- المدى النسبي:

يستخدم هذا المقياس في المقارنة بين مجموعتين من البيانات المختلفة في وحدات قياسها أو في مقاييسها الإحصائية، ويحسب بالعلاقة التالية:

إذا كان الوسط الحسابي معلوم	$E\% = \frac{X_{max} - X_{min}}{\bar{x}} \times 100$
إذا كان الوسط الحسابي غير معلوم	$E\% = \frac{X_{max} - X_{min}}{X_{max} - X_{min}} \times 100$

4- الانحراف المتوسط (E_X):

يعرف الانحراف المتوسط ب: "القيمة المطلقة لمتوسط مجموع انحرافات القيم عن وسطها المركزي (أو أحد مقاييس النزعة المركزية)، ويحسب هذا المقياس حسب الحالات التالية:

4-1- بيانات غير مبوبة:

يحسب وفق العلاقة التالية:

$$E_X = \frac{\sum_{i=1}^{n} |x_i - \bar{X}|}{n}$$

مثال رقم (32): لنفرض المثال السابق رقم (31) الذي يمثل نقاط الطلبة في مقياس الإحصاء:

9 10 11 10 8 16 5 9 14 10

أولا: نحسب الوسط الحسابي:

$$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{10+14+9+5+16+8+10+11+10+9}{10} = 10.2$$

ثانيا: حساب الانحراف المتوسط:

المجموع	9	10	11	10	8	16	5	9	14	10	x_i
21.2	1.2	0.2	1.2	0.2	2.2	5.8	5.2	1.2	3.8	0.2	$\lvert x_i - \bar{X}\rvert$

$$E_X = \frac{\sum_{i=1}^{n} |x_i - \bar{X}|}{n} = \frac{21.2}{10} = 2.12$$

ومنه القيم متجانسة مع بعضها البعض.

4-2- بيانات غير مبوبة:

4-2-1- حالة بيانات كمية متقطعة:

في حالة توزيع تكراري متقطع، يحسب الإنحراف المتوسط وفق الصيغة التالية:

$$E_X = \frac{\sum_{i=1}^{k} (|x_i - \bar{X}| \times n_i)}{N}$$

مثال رقم (33): نستخدم المثال رقم (2-10): مع العلم أن: $\bar{X} = 3.10$

x_i	n_i	$\lvert x_i - \bar{X} \rvert$	$n_i \times \lvert x_i - \bar{X} \rvert$
1	20	2.1	42
2	26	1.1	28.6
3	35	0.1	3.5
4	37	0.9	33.3
5	22	1.9	41.8
المجموع	140		149.2
	N		$\sum_{i=1}^{k} n_i \times (\lvert x_i - \bar{X} \rvert)$

حساب الانحراف المتوسط:

$$E_X = \frac{\sum_{i=1}^{k}(\lvert x_i - \bar{X} \rvert \times n_i)}{N} = \frac{149.2}{140} = 1.06$$

ومنه القيم متجانسة مع بعضها البعض.

4-2-2- حالة بيانات كمية مستمرة:

في حالة توزيع تكراري مستمر، يحسب الانحراف المتوسط بالصيغة التالية:

$$E_X = \frac{\sum_{i=1}^{k}(\lvert c_i - \bar{X} \rvert \times n_i)}{N}$$

مثال رقم (34): نطبق المثال على التمرين رقم (10-3):

مع العلم أن الوسط الحسابي لهذا المثال هو $\bar{X} = 35.23$

x_i	n_i	C_i	$\lvert C_i - \bar{X}\rvert$	$n_i \times (\lvert c_i - \bar{X}\rvert)$
$]20 - 10]$	4	15	20.23	80.92
$]30 - 20]$	11	25	10.23	112.53
$]40 - 30]$	13	35	0.23	3
$]50 - 40]$	8	45	9.77	78.16
$]60 - 50]$	6	55	19.77	118.62
المجموع	42			393.23
	N			$\sum_{i=1}^{k}(n_i(\lvert c_i - \bar{X}\rvert)$

حساب الانحراف المتوسط:

$$E_X = \frac{\sum_{i=1}^{k}(|c_i - \bar{X}| \times n_i)}{N} = \frac{393.23}{42} = 9.36$$

5- التباين $(\sigma_x{}^2)$:

يعتبر هذا المقياس من أشهر مقاييس التشتت، ويعرف بأنه " متوسط مربع انحرافات القيم عن وسطها الحسابي و هو يقيس مدى انتشار أو ابتعاد كل قيمة من القيم التي تأخذها الظاهرة المدروسة عن وسطها الحسابي، ويحسب كما يلي في الحالات التالية:

5-1- بيانات غير مبوبة:

في حالة وجود سلسلة إحصائية غير مبوبة، يحسب التباين بالعلاقة:

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

مثال رقم (35): نطبق القانون على المثال السابق رقم (32) مع العلم أن $\bar{X} = 10.2$

المجموع	9	10	11	10	8	16	5	9	14	10	x_i
	-1.2	-0.2	-1.2	-0.2	-2.2	5.8	-5.2	-1.2	3.8	-0.2	$(x_i - \bar{X})$
84.4	1.44	0.04	1.44	0.04	4.84	33.64	27.04	1.44	14.44	0.04	$(x_i - \bar{X})^2$

$$\sigma_x{}^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{X})^2}{N} = \frac{84.4}{10} = 8.44$$

5-2- بيانات مبوبة:

5-2-1- حالة متغيرة كمية متقطعة:

في حالة هذا التوزيع، يحسب التباين بالصيغة التالية:

$$\sigma_x{}^2 = \frac{\sum_{i=1}^{k}[(x_i - \bar{X})^2 \times n_i]}{N}$$

مثال رقم (36): نطبق القانون على التمرين السابق رقم (10-2) مع العلم أن $\bar{X} = 3.10$

$[n_i \times (x_i - \bar{X})^2]$	$(x_i - \bar{X})^2$	$x_i - \bar{X}$	n_i	$.x_i$
88.2	4.41	-2.1	20	1
31.46	1.21	-1.1	26	2
0.35	0.01	-0.1	35	3
29.97	0.81	0.9	37	4
79.42	3.61	1.9	22	5
229.4			**140**	المجموع
$\sum_{i=1}^{k}[n_i \times (x_i - \bar{X})^2]$			N	

$$\sigma_x{}^2 = \frac{\sum_{i=1}^{k}[(x_i - \bar{X})^2 \times n_i]}{N} = \frac{229.4}{140} = 1.63$$

5-2-2- حالة متغيرة كمية مستمرة:

في هذا التوزيع، يحسب التباين بالعلاقة التالية:

$$\sigma_x{}^2=\frac{\sum_{i=1}^{k}[n_i\times(x_i-\bar{X})^2]}{N}$$

مثال رقم (37): نطبق العلاقة على التمرين السابق رقم (3-10) مع العلم أن $\bar{X}=35.23$

x_i	n_i	C_i	$C_i-\bar{X}$	$(c_i-\bar{X})^2$	$[n_i(c_i-\bar{X})^2]$
$]20-10]$	4	15	-20.23	409.25	1637
$]30-20]$	11	25	-10.23	104.65	1151.15
$]40-30]$	13	35	-0.23	0.05	0.65
$]50-40]$	8	45	9.77	95.45	763.6
$]60-50]$	6	55	19.77	390.85	2345.1
المجموع	42				5897.5
	N				$\sum_{i=1}^{k}[n_i(c_i-\bar{X})^2]$

$$\sigma_x{}^2=\frac{\sum_{i=1}^{k}[n_i\times(x_i-\bar{X})^2]}{N}=\frac{5897.5}{42}=140.41$$

6- الانحراف المعياري (σ_x):

يعتبر هذا المقياس من أكثر مقاييس التشتت أهمية، لأنه يستند على عمليات جبرية محددة بدقة ويعتمد على خصائص رياضية مناسبة، ويعرف الانحراف المعياري بأنه "الجذر التربيعي لمتوسط انحرافات القيم عن وسطها الحسابي"، وهو يقيس مدى انتشار أو ابتعاد كل قيمة من القيم التي تأخذها الظاهرة المدروسة عن وسطها الحسابي، ويحسب كما يلي في الحالات التالية:

6-1- بيانات غير مبوبة:

في حالة وجود سلسلة إحصائية غير مبوبة، يحسب الانحراف المعياري ب:

$$\sigma_s = \sqrt{{\sigma_x}^2} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{X})^2}{N}} = \frac{\sum_{i=1}^{n} x^2}{N} - \bar{X}^2$$

مثال رقم (38): نحسب الانحراف المعياري للمثال السابق رقم (32):

$$\bar{X} = 10.2$$

x_i	10	14	9	5	16	8	10	11	10	9	المجموع
$(x_i - \bar{X})$	-0.2	3.8	-1.2	-5.2	5.8	-2.2	-0.2	-1.2	-0.2	-1.2	
$(x_i - \bar{X})^2$	0.04	14.44	1.44	27.04	33.64	4.84	0.04	1.44	0.04	1.44	84.4

$$\sigma_s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{X})^2}{N}} = \sqrt{\frac{84.4}{10}} = 2$$

6-2- بيانات مبوبة:

6-2-1- حالة متغيرة كمية متقطعة:

يحسب الانحراف المعياري في حالة التوزيع الكمي المتقطع ب:

$$\sigma_x = \sqrt{{\sigma_x}^2} = \sqrt{\frac{n_i \times (x_i - \bar{X})^2}{N}} = \frac{\sum_{i=1}^{k}(n_i \times {x_i}^2)}{N} - \bar{X}^2$$

مثال رقم (39): نحسب الانحراف المعياري للتمرين السابق رقم (2-10):

$$\bar{X} = 3.10$$

$n_i \times [(x_i - \bar{X})^2]$	$(x_i - \bar{X})^2$	$x_i - \bar{X}$	n_i	X_i
88.2	4.41	-2.1	20	1
31.46	1.21	-1.1	26	2
0.35	0.01	-0.1	35	3
29.97	0.81	0.9	37	4
79.42	3.61	1.9	22	5
229.4			**140**	المجموع
$\sum_{i=1}^{k} n_i \times [(x_i - \bar{X})^2]$			N	

$$\sigma_x = \sqrt{\frac{n_i \times (x_i - \bar{X})^2}{N}} = \sqrt{\frac{229.4}{140}} = 1.28$$

2-2-6- حالة متغيرة كمية مستمرة:

في حالة التوزيع التكراري المستمر، يحسب الانحراف المعياري ب:

$$\sigma_x = \sqrt{{\sigma_x}^2} = \sqrt{\frac{n_i \times (c_i - \bar{X})^2}{N}} = \frac{\sum_{i=1}^{k} (n_i \times {c_i}^2)}{N} - \bar{X}^2$$

مثال رقم (40): نحسب الانحراف المعياري للتمرين السابق رقم (3-10) مع العلم أن

$$\bar{X} = 35.23$$

x_i	n_i	C_i	$C_i - \bar{X}$	$(c_i - \bar{X})^2$	$[n_i(c_i - \bar{X})^2]$
$]20 - 10]$	4	15	-20.23	409.25	1637
$]30 - 20]$	11	25	-10.23	104.65	1151.15
$]40 - 30]$	13	35	-0.23	0.05	0.65
$]50 - 40]$	8	45	9.77	95.45	763.6
$]60 - 50]$	6	55	19.77	390.85	2345.1
المجموع	42				5897.5
	N				$\sum_{i=1}^{k}[n_i(c_i - \bar{X})^2]$

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^{k}[n_i \times (c_i - \bar{X})^2]}{N}} = \sqrt{\frac{5897.5}{42}} = 11.84$$

7- الانحراف الربيعي (Q):

يعتبر هذا المقياس مؤشر موضعي للتشتت حيث يقوم على الترتيب، ويعرف على أنه " متوسط الفرق بين الربيع الثالث والربيع الأول، ويعرف أيضا بنصف المدى الربيعي، ورياضيا يحسب ب:

$$\mathbf{Q} = \frac{Q_3 - Q_1}{2}$$

حيث: Q_1: الربيع الأول، Q_3: الربيع الثالث.

مثال رقم (41): نطبق المقياس على التمرين السابق رقم (10-3) :

$N^{+}_{i=1}$	n_i	x_i
4	4	$]20-10]$
15	11	$]30-20]$
28	13	$]40-30]$
36	8	$]50-40]$
42	6	$]60-50]$
	42	**المجموع**
	N	

حساب Q_1:

رتبة الربيع الأول هي $\frac{N}{4}=\frac{42}{4}=10.5$، ومنه فئة الربيع الأول هي: [20-30[، وبالتالي:

$$Q_1=d+\left(\frac{\frac{N}{4}-N^{+}_{-1}}{n_{Q_1}}\right)\times L=20+\left(\frac{10.5-4}{11}\right)\times 10=25.9$$

حساب Q_3:

رتبة الربيع الثالث هي $N\times\frac{3}{4}=42\times\frac{3}{4}=31.5$ ومنه فئة الربيع الثالث هي: [40-50[، وبالتالي:

$$Q_3=d+\left(\frac{\frac{3N}{4}-N^{+}_{-1}}{n_{Q_3}}\right)\times L=40+\left(\frac{31.5-28}{8}\right)\times 10=44.375$$

حساب الانحراف الربيعي:

$$\mathrm{Q}=\frac{Q_3-Q_1}{2}=\frac{44.375-25.9}{2}=9.23$$

8- معامل الاختلاف:

هذا المقياس يقارن بين المقاييس النسبية بين مجموعتين أو أكثر، ويعرف أيضا بمقياس الانحراف النسبي، فالمجموعة ذات الانحرافات النسبية الأقل هي المجموعات ذات التشتت الأقل وفي نفس الوقت هي المجموعات ذات التجانس الأكبر في مقياس نزعتها المركزية، ويحسب هذا المقياس بالصيغة التالية:

$$V = \frac{\sigma_x}{\overline{X}} \times 100$$

مثال رقم (42): نحسب معامل الاختلاف للمثالين التاليين:

التوزيع الأول: تمرين سابق رقم (10-2)

X_i	n_i	$x_i - \overline{X}$	$(x_i - \overline{X})^2$	$n_i \times [(x_i - \overline{X})^2]$
1	20	-2.1	4.41	88.2
2	26	-1.1	1.21	31.46
3	35	-0.1	0.01	0.35
4	37	0.9	0.81	29.97
5	22	1.9	3.61	79.42
المجموع	**140**			**229.4**
	N			$\sum_{i=1}^{k} n_i \times [(x_i - \overline{X})^2]$

الوسط الحسابي $\overline{X} = 3.1$

الانحراف المعياري $\sigma_x = 1.28$

معامل الاختلاف $V = \frac{\sigma_x}{\overline{X}} \times 100 = \frac{1.28}{3.1} \times 100 = 41.29$

التوزيع الثاني:

$n_i \times {x_i}^2$	${x_i}^2$	$n_i \times x_i$	n_i	x_i
8	4	4	2	2
90	9	30	10	3
336	16	84	21	4
175	25	35	7	5
609		153	40	المجموع
$\sum_{i=1}^{k}(n_i \times {x_i}^2)$		$\sum_{i=1}^{k}(n_i \times x_i)$	N	

الوسط الحسابي $\bar{X} = \frac{\sum_{i=1}^{k} n_i \times x_i}{N} = \frac{153}{40} = 3.825$

الانحراف المعياري $\sigma_x = \sqrt{\frac{\sum_{i=1}^{k}(n_i \times (x_i^2))}{N} - (\bar{X})^2} = \sqrt{\frac{609}{40} - 3.825^2} = 0.77$

معامل الاختلاف $V = \frac{\sigma_x}{\bar{X}} \times 100 = \frac{0.77}{3.825} \times 100 = 20.13$

النتائج أظهرت ان التوزيع الثاني أقل تشتتا من التوزيع الأول، ومن التوزيع الثاني أكثر تجانس وأفضل من الأول.

9- تمارين محلولة:

9-1- التمرين الأول:

السلسلة الإحصائية التالية تمثل عدد مراكز البريد لعدد من الدوائر بالشرق الجزائري:

3 4 5 3 5 5 3

المطلوب:

- أحسب المتوسط الحسابي؛
- أحسب الانحراف الربيعي والانحراف المعياري،

- استنتج معامل الاختلاف.

9-2- التمرين الثاني:

الجدول التالي يمثل توزيع الدوائر حسب عدد البلديات:

عدد البلديات	2	3	4	5	6	المجموع
عدد الدوائر	8	24	11	10	7	60

المطلوب:

- أحسب متوسط عدد البلديات؛
- أحسب التباين واستنتج الانحراف المعياري؛
- أحسب الانحراف الربيعي.

9-3- التمرين الثالث:

الجدول التالي يمثل توزيع النسمة (بالألف) حسب عدد البلديات:

النسمة	]10-5]	]15-10]	]20-15]	]25-20]	المجموع
عدد البلديات	7	12	8	13	40

المطلوب:

- أحسب الوسط الحسابي؛
- أحسب عدد النسمة الشائع؛
- أحسب الانحراف المعياري؛
- استنتج معامل الاختلاف.

10- الحل:

10-1- التمرين الأول:

3	3	3	4	5	5	5

- الوسط الحسابي $\bar{X}$:

$$\bar{X} = \frac{\sum_{i=1}^{7} x_i}{n} = \frac{3+4+5+3+5+5+3}{7} = 4$$

متوسط عدد مراكز البريد في الدوائر هو 4.

- الانحراف الربيعي Q:

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

رتبة Q_1: $C_{Q_1} = \frac{N+1}{4} = \frac{7+1}{4} = 2$

$$Q_1 = 3$$

رتبة Q_3: $C_{Q_3} = \frac{3N}{4} = \frac{3\times 7}{4} = 5$

$$Q_3 = 5$$

$$Q = \frac{5-3}{2} = 1$$

- الانحراف المعياري σ_x:

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^{7} x^2}{n} - \bar{X}^2} = \sqrt{\frac{3^2+4^2+5^2+3^2+5^2+5^2+3^2}{7} - 4^2} \Rightarrow \sigma_x = 0.92$$

التوزيع متجانس وأقل تشتت.

- معامل الاختلاف V:

$$V = \frac{\sigma_x}{\bar{X}} \times 100 = \frac{0.92}{4} \times 100\% = 23\%$$

10-2- التمرين الثاني:

$N_{i=1}^{+}$	$n_i \times x_i^2$	x_i^2	$n_i \times x_i$	n_i	x_i
8	32	4	16	8	**2**
32	216	9	72	24	**3**
34	176	16	44	11	**4**
53	250	25	50	10	**5**
60	252	36	42	7	**6**
	926		***224***	**60**	المجموع
	$\sum_{i=1}^{k} n_i \times x_i^2$		$\sum_{i=1}^{k} n_i \times x_i$	N	

- متوسط عدد البلديات (الوسط الحسابي $\bar{X}$):

$$\bar{X} = \frac{\sum_{i=1}^{k} n_i \times x_i}{N} = \frac{224}{60} = 3.73$$

متوسط عدد البلديات هو 4.

- التباين σ_x^2:

$$\sigma_x^2 = \frac{\sum_{i=1}^{k} n_i \times (x_i)^2}{N} - \bar{X}^2 = \frac{926}{60} - (3.73)^2 = 1.51$$

- الانحراف المعياري σ_x:

$$\sigma_x = \sqrt{\sigma_x^2} = \sqrt{1.51} = 1.22$$

- الانحراف الربيعي Q:

رتبة الربيع الأول $C_{Q_1} = \frac{N}{4} = \frac{60}{4} = 15$ $\qquad Q_1 = 3$

رتبة الربيع الثالث $C_{Q_3} = \frac{3N}{4} = \frac{3\times 60}{4} = 45$ $\qquad Q_3 = 5$

$$Q = \frac{Q_3 - Q_1}{2} = \frac{5-3}{2} = 1$$

التوزيع أقل تشتت وأكثر تجانس.

10-3- التمرين الثالث:

x_i	n_i	c_i	$n_i \times c_i$	${c_i}^2$	$n_i \times {c_i}^2$
]10-5]	7	7.5	52.5	56.25	393.75
]15-10]	12	12.5	150	156.25	1875
]20-15]	8	17.5	140	306.25	2450
]25-20]	13	22.5	292.5	506.25	6581.25
المجموع	**40**		**635**		**11300**
	N		$\sum_{i=1}^{k} n_i \times c_i$		$\sum_{i=1}^{k} n_i \times {c_i}^2$

- متوسط عدد النسمة (الوسط الحسابي $\bar{X}$):

$$\bar{X} = \frac{\sum_{i=1}^{k} n_i \times c_i}{N} = \frac{635}{40} = 15.875$$

متوسط عدد النسمة في البلديات هو 15875 نسمة.

- العدد الشائع للنسمة في البلديات (المنوال *Mo*):

الفئة المنوالية هي :]25-20] ومنه:

$$Mo = d + \left(\frac{(N_0 - N_1)}{(N_0 - N_1) + (N_0 - N_2)}\right) \times L = 20 + \left(\frac{(13-8)}{(13-8)+(13-0)}\right) \times 5 = 21.388$$

عدد النسمة الشائع في البلديات هو21388 نسمة.

- الانحراف المعياري σ_x:

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^{k} n_i \times c_i^2}{N} - \bar{X}^2} = \sqrt{\frac{11300}{40} - (15.875)^2} = 5.52$$

- معامل الاختلاف V:

$$V = \frac{\sigma_x}{\bar{X}} \times 100 = \frac{5.52}{15.875} \times 100 = 34.78\%$$

11- تمارين غير محلولة:

11-1- التمرين الأول:

لتكن السلسلة الإحصائية التالية تمثل عدد المدرجات المتاحة لكل كلية في جامعة معسكر:

4 2 5 4 2 3 2

المطلوب:

- أحسب الوسط الحسابي؛
- أحسب التباين ثم استنتج الانحراف المعياري؛
- استنتج معامل الاختلاف.

11-2- التمرين الثاني:

ليكن الجدول التالي الذي يمثل توزيع عدد مراكز الشرطة حسب عدد الشكاوي المقدمة يوميا في ولايات الغرب:

عدد الشكاوي	1	2	3	4	5	6
عدد مراكز الشرطة	5	4	6	4	3	2

المطلوب:

- أحسب الوسط الحسابي؛
- أحسب الانحراف المتوسط والانحراف المعياري؛
- أحسب الانحراف الربيعي؛
- استنتج معامل الاختلاف.

3-11- التمرين الثالث:

الجدول التالي يمثل توزيع الأبقار حسب حجم الحليب المعطى يوميا في إحدى المزارع:

حجم الحليب	]10-5]	]15-10]	]20-15]	]25-20]	]30-25]
عدد الأبقار	4	7	22	14	15

المطلوب:

- أحسب حجم الحليب المتوسط؛
- أحسب الربيع الأول والثالث؛
- استنتج الانحراف الربيعي؛
- أحسب التباين؛
- استنتج معامل الاختلاف.

المحور الخامس: مقاييس الالتواء والتفرطح (مقاييس الشكل)

بعدما تطرقنا إلى مقاييس النزعة المركزية التي تقيس مدى تمركز القيم حول نقطة أو مقياس، ومقاييس التشتت التي تقيس مدى تباعد أو تقارب القيم عن بعضها البعض، نتطرق في هذا المحور إلى قياس تماثل التوزيع من خلال مقاييس الالتواء والتفرطح. فمقاييس النزعة المركزية ومقاييس التشتت لا تكفي وحدها في وصف البيانات، فقد نجد التمثيلات متساوية في وسطها الحسابي وانحرافها المعياري، لكنها تختلف من حيث الشكل، أي إما تكون إحداها ملتوية أو متفرطحة عن الأخرى، ونقدم أهم مقاييس الالتواء والتفرطح في النقاط الموالية.

1- الالتواء:

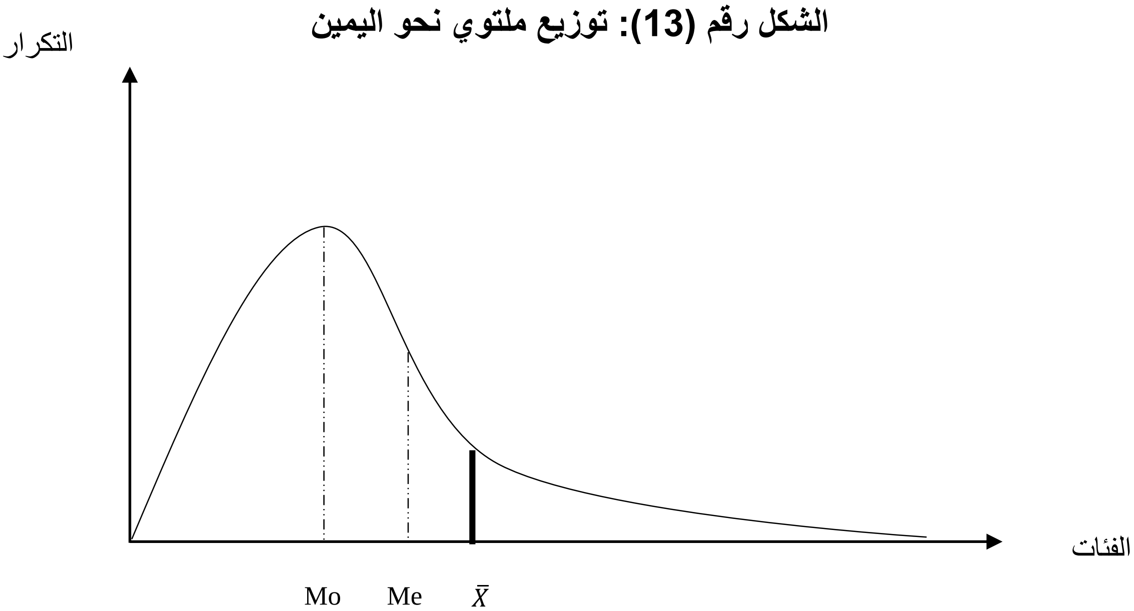
1-1- أشكال التوزيعات التكرارية:

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

1-1-1- توزيع تكراري ملتوي نحو اليمين (موجب الالتواء):

هذا التوزيع يكون ممتد نحو اليمين (نحو + ∞)، ورياضيا يكون: $\bar{X} > Me > Mo$، كما نلاحظ في الشكل التالي:

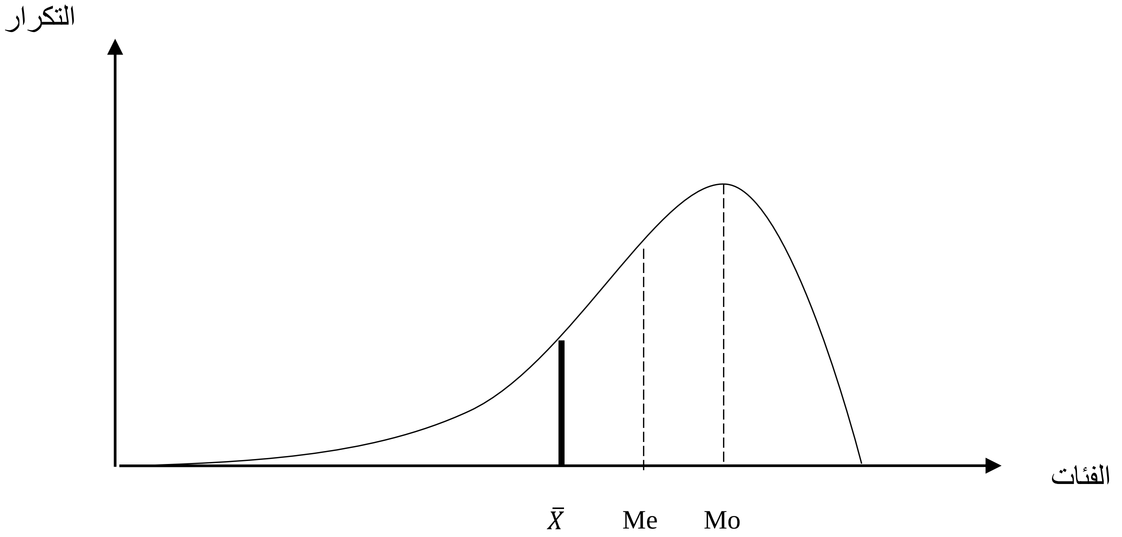
الشكل رقم (13): توزيع ملتوي نحو اليمين



1-1-2- توزيع تكراري ملتوي نحو اليسار (سالب الالتواء):

هذا التوزيع يكون ممتد نحو اليسار (نحو المبدأ)، ورياضيا يكون: $\bar{X} < Me < Mo$، كما نلاحظ في الشكل التالي:

الشكل رقم (14): توزيع ملتوي نحو اليسار

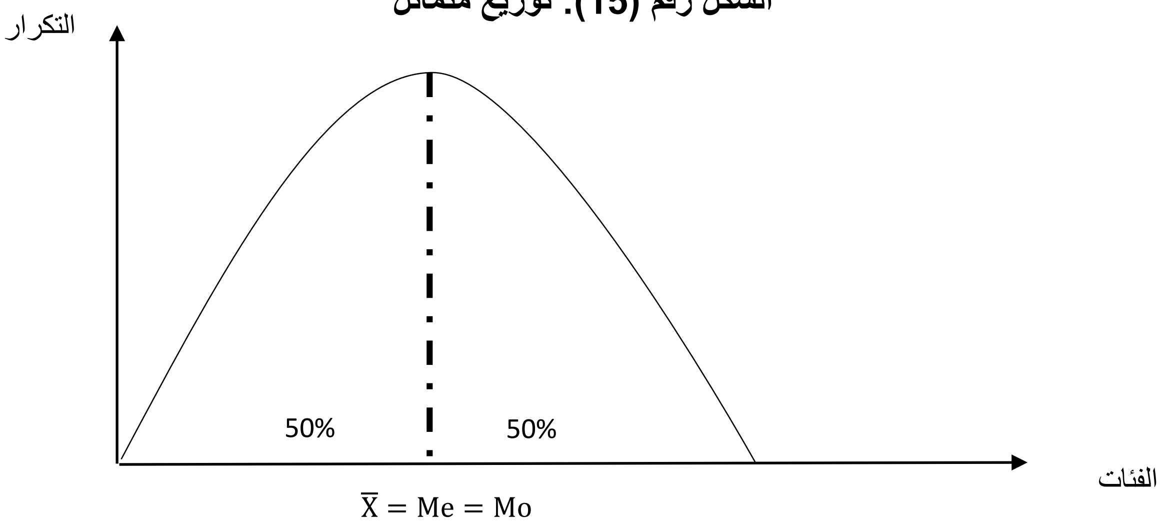


1-1-3- توزيع تكراري متماثل:

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

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

الشكل رقم (15): توزيع متماثل



1-2- مقاييس الالتواء:

1-2-1- مقاييس بيرسون (Person) للالتواء:

- معامل بيرسون الأول P_1:

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

$$P_1 = \frac{\bar{X} - Me}{\sigma_x}$$

من هذه الصيغة يتحدد شكل التوزيع حيث:

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

- معامل بيرسون الثاني P_2:

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

$$P_1 = \frac{3(\bar{X} - Mo)}{\sigma_x}$$

من هذه الصيغة يتحدد شكل التوزيع حيث:

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

- **معامل بيرسون الثالث للإلتواء P_3:**

يتحدد هذا المقياس باستخدام العزوم من الدرجة الثانية والدرجة الثالثة، حيث يحسب بالصيغة التالية:

$$P_3 = \frac{M_3^2}{M_2^3}$$

حيث: M_2 هو العزم المركزي من الدرجة الثانيّة؛ M_3 هو العزم المركزي من الدرجة الثالثة.

ملاحظة: معامل بيرسون للالتواء باستخدام العزوم P_3 دائما موجبا، لهذا لا يعطي هذا المقياس أي توضيح لشكل التوزيع إلا في حالة يكون يساوي الصفر فإن التوزيع يكون متناظر.

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

ورياضيا يكتب في هذه الحالات:

1) **العزم البسيط:** وهو العزم البسيط من الدرجة r حول الصفر ويحسب رياضيا بالعلاقة التالية:

- **بيانات غير مبوبة:** يعطى بالعلاقة التالية:

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

- **بيانات مبوبة:** يعطى بالعلاقة التالية:

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

حيث r تمثل درجة العزم $(r=1.2....)$.

2) **العزم المركزي:** هو العزم المركزي من الدرجة r حول المتوسط الحسابي ويحسب رياضيا ب:

- **بيانات غير مبوبة:** يعطى بالعلاقة التالية:

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

- **بيانات مبوبة:** يعطى بالعلاقة التالية:

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

حيث r تمثل درجة العزم المركزي $(r=1,2.....)$.

ملاحظة: العزم المركزي من الدرجة الأولى يساوي الصفر، أي $M_1 = 0$

والعزم المركزي من الدرجة الثانية هو التباين، أي $M_2 = {\sigma_x}^2$

1-2-2- معامل فيشر للإلتواء:

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

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

حيث M_3 هو العزم المركزي من الدرجة الثالثة، σ هو الانحراف المعياري.

ويتم تحديد شكل التوزيع بناء على قيمة العزم المركزي من الدرجة الثالثة، حيث إذا كان يساوي الصفر $(M_3 = 0)$ فإن التوزيع يكون متناظر، وفي جميع الحالات يكون كما يلي:

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

1-2-3- معامل يول للالتواء Y_1:

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

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

أما شكل التوزيع حسب هذا المقياس يتحدد كما يلي:

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

مثال رقم (43): يمثل الجدول التالي توزيع أبناء أحد الأحياء في ولاية تلمسان حسب الوزن (بالكيلوغرام):

الوزن	]20-10]	]30-20]	]40-30]	]50-40]	]60-50]	]70-60]
عدد الصغار	18	30	25	17	12	8

المطلوب: أحسب مقاييس الالتواء؟

الحل:

الوزن	عدد الصغار	c_i	$n_i \times c_i$	$N_{i=1}^{+}$	$(c_i - \bar{X})$	$(c_i - \bar{X})^2$	$(c_i - \bar{X})^2 \times n_i$	$(c_i - \bar{X})^3$	$(c_i - \bar{X})^3 \times n_i$
]20-10]	18	15	270	18	-19.9	396.01	7128.18	-7880.6	-141850.8
]30-20]	30	25	750	48	-9.9	98.01	2940.3	-970.3	-29109
]40-30]	25	35	875	73	0.1	0.01	0.25	0.001	0.025
]50-40]	17	45	765	90	10.1	102.01	1734.17	1030.3	17515.1
]60-50]	12	55	660	102	20.1	404.01	4848.12	8120.6	97447.2
]70-60]	8	65	520	110	30.1	906.01	7248.08	27270.9	218167.2
المجموع	**110**		**3840**				**23899.1**		**162169.72**

- مقياس بيرسون الأول للإلتواء P_1:

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

الوسط الحسابي $\bar{X}$:

$$\bar{X} = \frac{\sum_{i=1}^{k} n_i \times c_i}{N} = \frac{3840}{110} = 34.9$$

المنوال **Mo**:

الفئة المنوالية هي :]30-20]، ومنه:

$$Mo = d + \left(\frac{(N_0 - N_1)}{(N_0 - N_1) + (N_0 - N_2)}\right) \times L$$

$$= 20 + (\frac{30 - 18}{(30 - 18) + (30 - 25)} \times 10 = 27.05$$

الإنحراف المعياري σ:

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^{k}[(c_i - \bar{X})^2 \times n_i]}{N}} = \sqrt{\frac{23899.1}{110}} = 14.74$$

$$\boldsymbol{P_1 = \frac{\bar{X} - Mo}{\sigma_x} = \frac{34.9 - 27.05}{14.74} = 0.53 > 0}$$

معامل بيرسون الأول أكبر من الصفر، ومنه التوزيع ملتوي نحو اليمين (موجب الالتواء).

معامل بيرسون الثاني للإلتواء P_2:

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

الوسيط **Me** :

رتبة الوسيط = 55 ومنه الفئة الوسيطية هي: [30-40[، وبالتالي:

$$Me = d + \left(\frac{\frac{N}{2} - N_{i-1}^{+}}{n_{me}}\right) \times L = 30 + \left(\frac{55 - 48}{25}\right) \times 10 = 32.8$$

$$P_2 = \frac{3(\bar{X} - Me)}{\sigma_x} = \frac{3(34.9 - 32.8)}{14.74} = 0.42 > 0$$

معامل بيرسون الثاني للإلتواء أكبر من الصفر، ومنه التوزيع ملتوي نحو اليمين (موجب الإلتواء).

معامل بيرسون الثالث للإلتواء P_3:

$$P_3 = \frac{M_3^2}{M_2^3}$$

العزم المركزي من الدرجة الثانية M_2:

إن العزم المركزي من الدرجة الثانية M_2 هو التباين، أي $M_2 = \sigma_x^2$

$$M_2 = (14.74)^2 = 217.26$$

العزم المركزي من الدرجة الثالثة M_3:

$$M_3=\frac{\sum_{i=1}^{k}[(c_i-\bar{X})]^3\times n_i}{N}=\frac{162169.72}{110}=1474.27$$

$$P_3=\frac{M_3^2}{M_2^3}=\frac{(1474.27)^2}{(217.26)^3}=\frac{2173472.7}{10255086.44}=0.21$$

معامل بيرسون الثالث للالتواء أكبر من الصفر، ومنه التوزيع ملتوي نحو اليمين (موجب الالتواء).

معامل فيشر للالتواء F_1:

$$\boldsymbol{F_1=\frac{M_3}{\sigma_x^3}=\frac{1474.27}{(14.74)^3}=0.46>0}$$

معامل فيشر للالتواء أكبر من الصفر، ومنه التوزيع ملتوي نحو اليمين (موجب الالتواء).

معامل يول للالتواء Y_1:

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

الربيع الأول Q_1:

رتبة الربيع الأول هي 27.5 ومنه فئة الربيع الأول هي: [20-30[، وبالتالي:

$$\boldsymbol{Q_1=d+\left(\frac{\frac{N}{4}-N_{i-1}^{+}}{n_{Q_1}}\right)\times L=20+\left(\frac{27.5-18}{30}\right)\times 10=23.16}$$

الربيع الثاني Q_2 هي قيمة الوسيط *Me* وبالتالي:

$$\boldsymbol{Q_2=Me=32.8}$$

الربيع الثالث Q_3:

رتبة الربيع الثالث هي 82.5 ومه فئة الربيع الثالث هي : [40-50[، وبالتالي:

$$\boldsymbol{Q_3=d+\left(\frac{\frac{3N}{4}-N_{i-1}^{+}}{n_{Q_3}}\right)\times L=40+\left(\frac{82.5-73}{17}\right)\times 10=45.6}$$

$$Y_1 = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{(Q_3 - Q_1)} = \frac{(45.6 - 32.8) - (32.8 - 23.16)}{(45.6 - 23.16)}$$

$$= \frac{3.16}{22.44} = 0.14 > 0$$

معامل يول للالتواء أكبر من الصفر ومنه التوزيع ملتوي نحو اليمين (موجب الالتواء).

2- التفرطح:

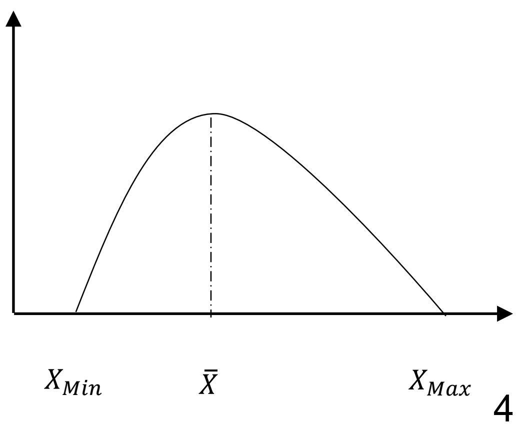
إن قياس التفرطح أو الانبساط يعني قياس مدى تشتت القيم عن بعضها البعض، أو يقيس قمة التوزيع بالنسبة للتوزيع الطبيعي، والتفرطح يأخذ ثلاث أشكال رئيسية:

توزيع متطاول أو مدبب: حيث يكون تشتت البيانات ضعيف جدا وتكون قمة التوزيع عالية؛

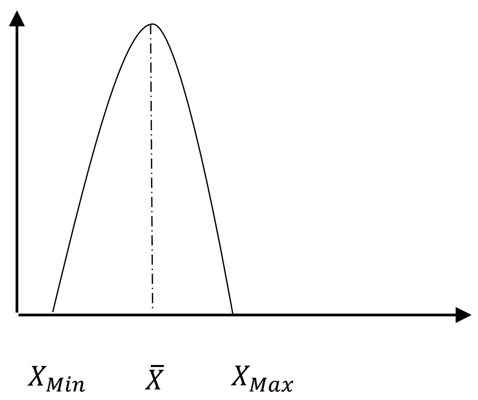
توزيع طبيعي: حيث يكون تشتت البيانات ليس كبير وليس صغير، أي معتدل؛

توزيع متفرطح (منبسط): حيث يكون تشتت البيانات كبير جدا.

الشكل رقم (16):توزيع معتدل

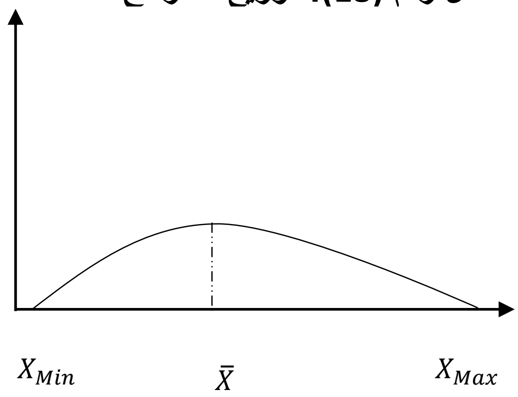


الشكل رقم (17): توزيع مدبب



4

الشكل رقم (18): توزيع متفرطح



2-1- قياس التفرطح.

2-1-1- مقياس بيرسون للتفرطح P_4:

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

$$P_4 = \frac{M_4}{M_2^2}$$

حيث، M_4 : العزم المركزي من الدرجة الرابعة، M_2: العزم المركزي من الدرجة الثانية.

ويتم تحديد شكل التوزيع من جانب التفرطح كما يلي:

- إذا كان $P_4 = 3$ فإن منحنى التوزيع يكون طبيعي؛
- إذا كان $P_4 > 3$ فإن منحنى التوزيع يكون مدبب؛
- إذا كان $P_4 < 3$ فإن منحنى التوزيع يكون متفرطح.

2-1-2- *معامل فيشر للتفرطح* F_2:

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

$$F_2 = \frac{M_4}{M_2^2} - 3$$

ويتم تحديد شكل التوزيع وفق هذا المقياس كما يلي:

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

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

2-1-3- معامل التفرطح باستخدام الانحراف الربيعي:

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

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

حيث: D_1: العشير الأول، D_9: العشير التاسع.

أمّا تحديد شكل التوزيع من حيث التفرطح يكون كالتالي:

- إذا كان $C_K = 0.263$ فإن منحنى التوزيع يكون طبيعي؛
- إذا كان $C_k > 0.263$ فإن منحنى التوزيع يكون متدبب؛
- إذا كان $C_k < 0.263$ فإن منحنى التوزيع يكون متفرطح.

مثال رقم (43): يمثل الجدول التالي توزيع عدد المدارس الابتدائية في مجموعة من البلديات:

عدد المدارس	1	2	3	4	5
عدد البلديات	30	45	22	11	2

أحسب مقاييس التفرطح وحدد شكله؟

الحل:

x_i	n_i	$n_i \times x_i$	$N_{i=1}^{+}$	$(x_i - \bar{X})$	$(x_i - \bar{X})^2$	$(x_i - \bar{X})^2 \times n_i$	$(x_i - \bar{X})^4$	$(x_i - \bar{X})^4 \times n_i$
1	30	30	*30*	*-1.18*	*1.4*	*42*	1.94	*58.2*
2	45	90	*75*	*-0.18*	*0.03*	*1.35*	0.001	*0.045*
3	22	66	*97*	*0.82*	*0.67*	*14.74*	0.45	*9.9*
4	11	44	*108*	*1.82*	*3.31*	*36.41*	10.95	*120.45*
5	2	10	*110*	*2.82*	*7.95*	*15.9*	63.20	*126.4*
المجموع	**110**	**240**				***110.4***		***315***

حساب مقاييس التفرطح:

- **معامل بيرسون للتفرطح P_4:**

$$\boldsymbol{P_4 = \frac{M_4}{M_2^2} = \frac{\frac{\sum_{i=1}^{k}[(x_i - \bar{X})^4 \times n_i]}{N}}{\left[\frac{\sum_{i=1}^{k}[(x_i - \bar{X})^2 \times n_i]}{N}\right]^2}}$$

الوسط الحسابي $\bar{X}$:

$$\boldsymbol{\bar{X} = \frac{\sum_{i=1}^{k}(x_i \times n_i)}{N} = \frac{240}{110} = 2.18}$$

العزم المركزي من الدرجة الثانية M_2:

$$M_2 = \frac{\sum_{i=1}^{k}[(x_i - \bar{X})^2 \times n_i]}{N} = \frac{110.4}{110} = 1$$

العزم المركزي من الدرجة الرابعة M_4:

$$M_4 = \frac{\sum_{i=1}^{k}[(x_i - \overline{X})^4 \times n_i]}{N} = \frac{315}{110} = 2.86 < 3$$

معامل بيرسون للتفرطح أقل من قيمة 3، ومنه التوزيع متفرطح.

- **معامل فيشر للتفرطح F_2:**

$$P_4 = \frac{M_4}{M_2^2} - 3 = 2.86 - 3 = -0.1 < 0$$

معامل فيشر للتفرطح أقل من الصفر، وبالتالي التوزيع هو متفرطح.

- **معامل الانحراف الربيعي للتفرطح C_k:**

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

الربيع الأول Q_1:

رتبة الربيع الأول هي $\frac{110}{4}$ وتساوي 27.5، أي $Q_1 = 2$.

الربيع الثالث Q_3:

رتبة الربيع الثالث هي $\frac{3\times110}{4}$ وتساوي 82.5، أي $Q_3 = 3$.

العشير الأول D_1:

رتبة العشير الأول هي $\frac{110}{10}$ وتساوي 11، أي $D_1 = 1$.

العشير التاسع D_9:

رتبة العشير التاسع هي $\frac{9\times110}{10}$ وتساوي 99، أي $D_9 = 4$.

$$C_k = \frac{Q_3 - Q_1}{2(D_9 - D_1)} = \frac{3 - 2}{2(4 - 1)} = 0.16$$

قيمة هذا المقياس أصغر من 0.263، وبالتالي يعتبر هذا التوزيع متفرطح.

3- تمارين غير محلولة:

3-1- التمرين الأول:

تمثل السلسلتين الإحصائيتين التاليتين نقاط طالبين في مختلف مقاييس الطور:

الطالب الأول	12	10	08	11	10	14	12
الطالب الثاني	6	16	07	20	05	06	17

المطلوب:

قارن بين التوزيعين من خلال معامل الاختلاف؟

أدرس شكل التوزيعين من حيث الالتواء والتفرطح؟

3-2- التمرين الثاني:

يمثل الجدول التالي توزيع عدد المصانع حسب عدد الفروع الإنتاجية في ولاية الجزائر:

عدد الفروع	1	2	3	4	5	6	7
عدد المصانع	7	12	20	16	14	8	3

المطلوب:

أحسب الانحراف المعياري؟

أحسب معاملات بيرسون للالتواء؟

أحسب معامل فيشر للتفرطح؟

3-3- التمرين الثالث:

قام مستثمر فلاحي دراسة إحصائية لمعرفة حجم الإنتاج (بالكلغ)التي تنتجه كل شجرة زيتون فكانت النتائج التالية:

كمية الإنتاج	]80-50]	]110-80]	]140-110]	]170-140]	]200-170]	]230-200]
عدد الأشجار	150	130	400	320	300	200

المطلوب:

أحسب التباين واستنتج الانحراف الربيعي؟

أحسب معاملات الالتواء؟

أحسب معاملين للتفرطح؟

ما نوع التوزيع؟

المراجع:

1- جيلالي جلاطو، الإحصاء مع تمارين ومسائل محلولة، ديوان المطبوعات الجامعية، الجزائر، 2002؛
2- عبد الرزاق عزوز، الكامل في الإحصاء: الجزء الأول، ديوان المطبوعات الجامعية، الجزائر، 2010؛
3- محمد راتول، الإحصاء الوصفي، ديوان المطبوعات الجامعية، الجزائر، الطبعة الثانية، 2006؛
4- محمد بوهزة، محاضرات في الإحصاء الوصفي، دار المحمدية العامة، الجزائر، 2011؛
5- محمد رشيد، مبادئ الإحصاء والاحتمالات، دار صفاء، عمان، الأردن، 2012؛
6- مختار الهاشمب، مبادئ الإحصاء، الدار الجامعية، بيروت، 1993؛
7- محمد أبو صالح، مقدمة في الإحصاء، مركز الكتب الأردني، عمان، 1990؛
8- أحمد عبادة سرحان، الإحصاء، مؤسسة شباب الجامعة، الإسكندرية، 1977؛
9- عبد الحسين زيني، الإحصاء السكاني، وزارة التعليم العالي، بغداد، 1980.