جامعـــــة الجزائـــــر -3-
معهـــد التربيـــة البدنيـــة و الرياضيـــة
- دالي إبراهيم -

برنامج السنة الثانية ل.م.د في مقياس :

الإحصاء الوصفي

من إعداد الأستاذة:
د.موساوي أمال

العام الجامعي :2019/2018

الإحصاء

الوصفي

برنامج مقياس الإحصاء الوصفي

لسنة أولى ل م د

المحور الأول: نشأة الإحصاء:

- تمهيد.
- تعريف الإحصاء.
- مبادئ علم الإحصاء.
- المجتمع الإحصائي والعينات.
- خطوات البحث العلمي.

المحور الثاني: الإحصاء وخطوات البحث العلمي:

المحور الثالث: خطوات العمل الإحصائي.

المحور الرابع: تطبيقات حول الدروس.

المحور الخامس: مقاييس النزعة المركزية.

المحور السادس: مقاييس التشتت.

المحور السابع: تطبيقات.

المحور الأول: الإحصاء:

تمهيد:

عرف الإنسان منذ فجر التاريخ الإحصاء، ويقال أن أولى من استعمل الإحصاء هم القدماء المصريين، حيث قاموا بعملية تعداد سكان مصر وثرواتهم قبل إنشاء الأهرام حتى يكون ذلك لهم عونا ومرشدا في عملية البناء، ومع ذلك لم يبرز الإحصاء كعلم إلا في أواخر القرن 18 بعد إنشاء قوانين الاحتمالات والتحليل الإحصائي وبالتالي أصبح الإحصاء أداة فعالة لدراسة مختلف الظواهر الاجتماعية والسياسية والاقتصادية والرياضية.

تعريف الإحصاء:

هو فرع من فروع الرياضيات يهتم بوضع الطرق العلمية القائمة على جمع البيانات المختلفة والتعبير عنها بشكل أرقام، وتعتبر هذه العملية الخطوة الأولى لتحقيق هدف معين. فالعمل الإحصائي يبدأ بجمع البيانات ولا ينتهي إلا بعد مراجعتها وتبويبها ثم تحليلها وتفسيرها وبالتالي اتخاذ القرار المناسب أو وضع التوصيات المناسبة.

مبادئ علم الإحصاء:

أ. **جمع البيانات:**

يقصد بجمع البيانات الحصول على بيانات رقمية أو وصفية تتصف بالصحة والدقة عن ظاهرة معينة، وهي نوعان:

- **بيانات تاريخية:** ويمكن الحصول عليها من البيانات التي تنشرها دائرة الإحصاءات العامة والدوائر الحكومية والخاصة، أو الهيئات المتخصصة في المجالات والكتب.
- **بيانات ميدانية:** ويتم جمع هذه البيانات من قبل الباحث نفسه أو مساعديه من خلال عدة طرق، منها الاستبيان، المقابلات، الملاحظة، تحليل محتوى والاختبارات.

ب. مراجعة البيانات:

بعد الانتهاء من جمع البيانات ترسل إلى الجهة المعينة للقيام بعملية المراجعة التي تهدف إلى: إلغاء الإجابات الناقصة، عد الاستمارات، التأكد من سلامة البيانات.

ت. تبويب البيانات:

بعد مراجعة البيانات يتم عرضها على هيئة جداول ورسومات وتتقسم البيانات إلى قسمين:

- **بينات وصفية (متغيرات كيفية):** وهي تلك المتغيرات التي لا يمكن قياسها أو غير قابلة للقياس مثل: الجنسية، الحالة الاجتماعية، أنواع الرياضات.
- **بينات رقمية (متغيرات كمية):** وهي تلك الخصائص التي يمكن قياسها وهي أكثر المتغيرات انتشارا واستعمالا لأن لغة الإحصاء هي لغة الأرقام مثل: الطول، الوزن، العمر.....

وتنقسم المتغيرات الكمية إلى قسمين:

- **كميات متصلة (متغيرات مستمرة):**

وهي تلك المتغيرات التي تأخذ كل القيم الممكنة لمجال الدراسة، ونظرا للعدد غير المتناهي هذه القيم نقسم مجال الدراسة إلى مجالات جزئية تسمى الفئات.

- **كميات منفصلة (متغيرات منقطعة):**

وهي تلك المتغيرات التي تأخذ قيما صحيحة، لا يمكن تجزئتها مثلا: عدد اللاعبين في فريق ما .... (أي لا تأخذ قيما رقمية كسرية).

ث. تحليل البيانات:

بعد جمع البيانات وتبويبها يتم معالجتها رياضيا باستخدام العلاقات والقوانين الإحصائية لاستخراج قيم عددية لها ولمؤشراتها ومدلولاتها وتفسيراتها، مثل استخدام مقاييس النزعة المركزية، مقاييس التشتت، أو معاملات الارتباط، وبالتالي الوصول إلى النتائج.

ج. تفسير نتائج تحليل البيانات:

ويقصد بها توضيح دلالات النتائج الإحصائية التي تم التوصل إليها، وبالتالي عمل الاستنتاجات وإصدار الأحكام على مجتمع الدراسة، بناء على إطلاع الباحث على الظروف المحيطة بالدراسة، ونظرياتها المختلفة، والتي قد تؤدي إلى تفسيرات تختلف من باحث لآخر.

المجتمع الإحصائي والعينات:

- **تعريف المجتمع الإحصائي: « Population statistique »:**

هو مجموعة المشاهدات والقياسات الخاصة بمجموعة من الوحدات الإحصائية والتي تخص ظاهرة من الظواهر القابلة للقياس.

أو: هو عبارة عن مجموعة من الأفراد أو الأشياء القابلة للقياس مثلا: مجموعة من البشر (طلبة، رياضيين)، او مجموعة من الأشياء المادية (وسائل نقل، ملاعب، كرات)، أو مجموعة من الوسائل الغير مادية (الأعمار، الذكاء، درجة الحرارة).

- **الوحدة الإحصائية: « Unité statistique »:**

هي الوحدة الأساسية لتكوين المجتمع الإحصائي.

مثال: مجتمع إحصائي: رياضيين، الوحدة الإحصائية: رياضي.

- **الظاهرة الإحصائية: « Phénomène statistique »:**

هي الخاصية المدروسة أو المتغير المدروس في المجتمع الإحصائي، مثلا: طول القامة، السن، الوزن، العلامة المحصل عليها في مقياس ما.

- **العينة: « Echantillon »:**

العينة هي جزء من المجتمع الإحصائي أو المجموعة الجزئية من المفردات الداخلة في تركيب المجتمع الذي يجرى عليه البحث شرط أن يكون اختيارها بشكل يجعلها ممثلة

للمجتمع الأصلي تمثيلا صحيح. وعندئذ يستطيع الباحث أن يستخلص من دراسة العينة نتائج تصلح للتعبير عن المجتمع بأكمله (تعميم النتائج).

أنواع العينات وتصنيفاتها:

يمكن تصنيف العينات إلى صنفين رئيسيين هما:

أ. **عينات عشوائية احتمالية:** وتضم كل من:

1. العينة الحصصية.
2. عينة الصدفة.
3. العينة القصدية.
4. عينة الكرة الثلجية.

ب. عينات عشوائية احتمالية:

1. العينة العشوائية البسيطة:

هي مجموعة جزئية بحجم مجدد من المجتمع الإحصائي يتم اختيارها بحيث يملك كل عنصر في المجتمع فرصة اختياره كأحد عناصر العينة بحيث تكون جميع المجموعات بنفس الحجم (وهو مجتمع متجانس لهم نفس الخصائص).

يتم اختيار العينة العشوائية البسيطة كما يأتي:

- يعطى كل عنصر من عناصر مجتمع الدراسة رقما متسلسلا يبدأ من صفر وينتهي عند $N-1$ حيث N هي عدد أفراد المجتمع.

- عمل جداول عشوائية، أو استخدام جداول عشوائية جاهزة.
- اختيار أحد أعمدة الجداول العشوائية مثلا، لأخذ أرقام أفراد أو عناصر العينة n.

مثال:

عدد الطلبة المسجلين في مادة الإحصاء في إحدى الجامعات (N =500) طالب، والمطلوب اختيار عينة عشوائية بسيطة، حجمها (n=50) طالبا.

الحل:

نعطي الطلبة أرقاما متسلسلة من صفر إلى 499 مكون من ثلاث أرقام، لأن 500 مكون من 3 أرقام كما يأتي:

000 001 002 003 004 005 006 007 008 009

010 011 012 013 014 015 016 017 018 019

–

–

490 491 492 493 494 495 496 497 498 499

ثم عن طريق الجداول العشوائية لنا حق اختيار أي عمود مثلا الأول من اليسار أي الذين أرقام كما يأتي:

000, 010 , 020 , 030 , …

ثم قد نحتاج إلى عمود آخر، وهكذا حتى نصل إلى العنصر أو الفرد الخمسين.

2. العينة المنتظمة:

هي عينة تختار من الإطار وفق ترتيب معين، بحيث يتم تحديد مقدار القفز بين كل عنصر والذي يليه حسب القاعدة:

$$\text{مقدار القفز} = \frac{N}{\text{n}}$$

مثال: N = 1000 ، n = 200

الحل: مقدار القفز = 1000 / 200 = 5 فنختار هنا أحد الأرقام من 1 إلى 5 وليكن 2 فيكون هو الفرد الأول، ثم يضاف له 5 وهكذا، أي: 2، 7، 12، 17، ....

3. العينة الطبقية:

وهي عينة مكونة من عدة عينات عشوائية بسيطة نختار كل منها من طبقة من طبقات المجتمع.

يلجأ الباحث إلى العينة الطبقية عندما يكون المجتمع الأصلي مقسم إلى طبقات.

ميزة العينة الطبقية أنها أكثر دقة من البسيطة والمنتظمة لأنها أكثر دقة في تمثيل المجتمع.

وسيتم توضيح خطوات حسابها من خلال المثالين التاليين:

مثال 1:

إذا كان طلبة السنة الأولى موزعين في إحدى الكليات حسب التخصصات الآتية:

العدد N	التخصص
200	**التدريب الرياضي**
150	**الإعلام الرياضي**
150	**إدارة وتسيير رياضي**

المطلوب: اختبر عينة حجمها (100) طالب بطريقة العينة الطبقية:

الحل:

$$N = W_1 + N_2 + N_3$$
$$N = 200 + 150 + 150$$
$$N = 500$$

ولحساب عدد أفراد العينة نحسب: $n.\frac{N_1}{N}$

- عدد أفراد عينة التدريب الرياضي: $1000\left(\frac{200}{500}\right) = 40$
- عدد أفراد عينة الإعلام الرياضي: $1000\left(\frac{150}{500}\right) = 30$
- عدد أفراد عينة إدارة وتسيير رياضي: $1000\left(\frac{150}{500}\right) = 30$

ثم تختار العينة بطريقة العينة العشوائية البسيطة.

4. العينة العنقودية:

هي عينة تختار من المجتمع بحيث تكون وحدات المعاينة عناقيدا بدلا من أن تكون عناصرا. وتستخدم في حالة كون المجتمع كبيرا جدا حيث يقوم الباحث بتقسيم المجتمع إلى أقسام كل قسم يسميه عنقود، وكل عنقود يقسمه إلى أجزاء إلى أن يصل إلى العينة التي يختارها.

ثانيا: العينات غير الاحتمالية (الغير عشوائية):

1-العينة الحصصية:

وهي شبيهة بالعينة الطبقية، غير أنه في الطبقية يتم الاختيار وفقا للنسب، أما في الحصصية فيتم الاختيار وفقا للصدفة أو قصدا.

2-عينة الصدفة:

عينة يختارها الباحث صدفة، ولا يكون على علم مسبق بالأفراد الذين سيجمع منهم البيانات.

- اهتمام الباحث بأسهل الطرق للحصول على مشاهدات من وحدة المعاينة.
- قد يواجه الباحث مشكلة عدم استجابة بعض أفراد العينة التي اختارها عشوائيا.
- أنها من أشهر طرق المعاينة انتشارا في العلوم الطبية ما يسمى بالعينة المتوفرة، ويؤخذ على هذه الطريقة أنها لا يمكن أن تمثل المجتمع.

3-العينة القصدية:

وهي العينة التي يذهب فيها الباحث إلى أشخاص معينين يثق بوجود المعلومات لديهم.

4-عينة الكرة الثلجية:

تستخدم في بعض الدراسات الإدارية، وفيها قد لا يكون واضحا لدى الباحث الأشخاص الذين يجب جمع المعلومات منهم، ومن طرقها:

- **طريقة الشريحة الرأسية:**عن طريق جمع المعلومات بأسلوب ترتيبي من أعلى مسؤول في المؤسسة إلى أقل عامل فيها.
- **عن طريق الشريحة القطرية:** عن طريق جمع المعلومات من شرائح من موظفي المؤسسة دون الأخذ بالتسلسل الوظيفي.
- **طريقة عينة الكرة الثلجية:** تبدأ هذه الطريقة باختيار فرد معين، وبناء على استجابته يقرر الباحث بمفرده أو بأي استعانة بهذا الفرد من سيكون الفرد التالي الذي يتم اختياره. وهكذا تتراكم لدى الباحث البيانات، لذلك سميت بكرة الثلجية، إذ يعتبر الفرد الأول النقطة الأولى التي سيبدأ التكثيف حولها لإكمال الكرة، أي اكتمال العينة.

الإحصاء وخطوات البحث العلمي

إن خطوات البحث العلمي تكاد تكون نفسها في مختلف مناهج البحث العلمي سواء كان وصفي أو تجريبي أو تاريخي حيث أنها لا تختلف إلا في بعض المراحل التي تميز كل منهج عن غيره.

وعموما فإن هذه الخطوات هي:

1-الشعور بمشكل البحث وجمع البيانات والمعلومات التي تساعد على تحديدها.

2-تحديد مشكلة البحث التي يريد الباحث دراستها وصياغتها على شكل سؤال محدد أو أكثر من سؤال.

3-وضع فرضية أو مجموعة من الفرضيات كحلول مؤقتة لمشكلة البحث.

4-اختيار العينة التي ستجرى عليها الدراسة مع توضيح حجمها وأسلوب اختيارها.

5-اختيار أدوات البحث التي سوف يستخدمها الباحث في الحصول على البيانات والمعلومات كاستبيان، المقابلة... وفقا لطبيعة البحث وفروضه، ثم القيام بتعيين هذه وحساب صدقها وثباتها.

6-جمع البيانات والمعلومات المطلوبة بطريقة دقيقة ومنظمة وواضحة.

7-الوصول إلى النتائج وتنظيمها وتصنيفها.

8-تحليل النتائج وتفسيرها واستخلاص التعميمات منها.

9–صياغة توصيات البحث والاقتراحات المختلفة من خلال المرحلة السادسة والمتعلقة بجمع البيانات تبدأ مرحلة هامة وهي المعالجة الإحصائية، حيث أن هذه البيانات تبقى بدون فائدة ولا نستطيع ان نتوصل بها إلى نتائج مفيدة بمجرد فحص بسيط ومباشر فنحن نحتاج إلى فرزها وتلخيصها بالأدوات والطرق الدقيقة.

هنا يظهر الدور الأساسي للإحصاء في تزويد الباحث بالأدوات التي تساعده على ذلك، حيث يفيده في تلخيص النتائج بصورة مفهومة، إضافة إلى تحليل العوامل المسببة لظواهر معقدة، كما تساعد على التنبؤ بالظواهر.

خطوات العمل الإحصائي

العرض الجدولي:

- جدول كيفي (وصفي).
- جدول كمي مستمر.
- جدول كمي منقطع.

العرض البياني للبيانات الإحصائية.

العرض الجدولي:

تستخدم عدة طرق لعرض البيانات الإحصائية لغرض فهم واستيعاب المؤشرات الرئيسية التي تنطوي عليها البيانات الإحصائية ويكون عرض هذه البيانات في جداول مختلفة تكون حسب نوعية البيانات فيقوم الباحث بتفريغ النتائج المحصل عليها في جداول إحصائية بعد عملية فرز الاستبيانات أو الاستمارات التي تحتوي على معلومات وإحصائيات متعددة والجدول الإحصائي عادة يكون ذو عمودين ويشمل كل عمود على عنوان والعنوان يسجل أعلى العمود، ويخصص العمود الأول للظاهرة المدروسة التي ترمز له بـ x_i، أما العمود الثاني فيعنون بالتكرارات ويرمز لها بـ n_i وفي أسفل الجدول الإحصائي مجموع النتائج.

الجدول الإحصائي

الظاهرة المدروسة x_i	التكرار n_i
المجموع $\sum$	$\sum n_i$

وللجدول الإحصائي 3 أنواع:

أ- الجدول الإحصائي الوصفي (الكيفي):

هو ذلك الجدول الذي لا تظهر فيه أرقام في العمود مثلا المخصص للظاهرة المدروسة، نجد إما كلمات أو رموز أما الحالة المخصصة للتكرارات فهي أرقام.

مثال:

البيانات التالية توضح تقدير 30 طالب في امتحان الإحصاء (ممتاز، جيد جدا، جيد، مقبول، ضعيف).

المطلوب:

(إعداد جدول تكراري للبيانات التالية؟)

جيد، ضعيف، جيدا جدا، جيد، مقبول، ضعيف، ممتاز، ضعيف، جيد، جيد جدا، جيد، مقبول، جيدا جدا، جيد، مقبول، جيد، ممتاز، جيد جدا، جيد، جيد، مقبول، جيد، جيد جدا، ممتاز، مقبول، ممتاز، ضعيف، جيد، مقبول، جيد.

المطلوب:

تنظيم البيانات السابقة في جدول إحصائي تكراري ؟

الظاهرة المدروسة x_i	التكرار n_i
ممتاز	4
جيدا جدا	5
جيد	11
مقبول	6
ضعيف	4
المجموع $\sum$	30

ب- جدول كمي منقطع:

يشمل جدول كمي منقطع على الظاهرة المدروسة x_i في شكل قيم (أرقام) وهذه النقاط ترتب ترتيبا تصاعديا (كشرط أساسي)، ويشمل أيضا عمودين الأول للظاهرة المدروسة x_i والثاني للتكرارات n_i.

مثال: البيانات التالية توضح علامات 40 طالبا في استجواب علامته النهائية (5):

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

3، 5، 0، 5، 5، 4، 3، 2، 2، 4، 1، 3، 4، 3.

المطلوب: تنظيم البيانات السابقة في جدول إحصائي (تكراري)؟

الحل:

العلامة x_i	التكرارات n_i عدد الطلبة
0	2
1	6
2	4
3	11
4	10
5	7
المجموع $\sum$	40

ت- جدول إحصائي كمي مستمر:

المتغيرة المدروسة في الجداول الكمية المستمرة تكون في شكل مجالات وكل مجال يسمى فئة، وكل فئة تحتوي على الحد الأدنى للفئة والحد الأعلى للفئة وتكون الفئات مرتبة ترتيبا تصاعديا أما العمود الثاني مخصص للتكرارات n_i وعليه كل تكرار يقابل فئة.

مثال: نأخذ نفس المثال السابق ونعبر عنه بشكل فئات طول الفئة = 1.

المطلوب: إنشاء جدول إحصائي (تكراري)؟

الحل:

تكرارات n_i	الفئات x_i
2	$]1-0]$
6	$]2-1]$
4	$]3-2]$
11	$]4-3]$
17	$]5-4]$
40	المجموع $\sum$

طريقة ستورج « Sturge »

خطوات هذه الطريقة:

1) نقوم بترتيب القيم من الأصغر إلى الأكبر.

2) نعين المدى ونرمز لها بـ E:

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

المدى = القيمة الكبرى − القيمة الصغرى

3) نقوم بتعيين طول الفئة بطريقة التالية حيث:

$$K = \frac{E}{1 + 1,32\ lnx(N)}$$

E: المدى

Ln: اللوغاريتم النيبيري

K: طول الفئة

N: هو عدد عناصر الإحصائيات

كما يمكن كتابة هذا القانون باستعمال اللوغاريتم العشري بحيث:

Log: اللوغاريتم العشري (لغ$_{10}$)

$$K = \frac{E}{1 + 3,32\, log(N)}$$

4) مركز الفئات:

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

مثال:

ينظم معهد قسم ت ب ر كل سنة مسابقة الدخول إلى القسم فيما يلي نتائج سباق 100م لعينة تتكون من 60 متسابق:

12.0 – 3 مرات	12.7 – 5 مرات	13.2 – 5 مرات
12.2 – 3 مرات	12.8 – 7 مرات	13.3 – 3 مرات
12.4 – 2 مرتين	12.9 – مرى	13.4 – 3 مرات
12.5 – 3 مرات	13.0 – 10 مرات	13.5 – مرة
12.6 – مرتين	13.1 – 5 مرات	13.6 – مرة
13.7 – مرة		
13.8 – مرتين		
13.9 – مرتين		
14.0 – مرة		

المطلوب:

ضع هذه البيانات ضمن جدول التوزيع التكراري مع العلم أن: log60 = 1.7

الحل:

1/ E = 14.0 – 12 = 2

$$2/\ K = \frac{E}{1+3.32\log(N)} = \frac{2}{1+3.32\ (1.7)} = \frac{2}{6.644}$$

K = 0.30

مراكز الفئات	n_i	x_i
12.15	6	$]\mathbf{12.3-12.0}]$
12.45	5	$]\mathbf{12.6-12.3}]$
12.75	14	$]\mathbf{12.9-12.6}]$
	16	$]\mathbf{13.2-12.9}]$
	11	$]\mathbf{13.5-13.2}]$
	3	$]\mathbf{13..8-13.5}]$
	5	$]\mathbf{14.1-13.8}]$
	60	**المجموع** $\sum$

يمكن التمييز بين 3 مجموعات من التكرارات:

أ. **التكرار المطلق:** وهو التكرار العادي n_i

ب. **التكرار النسبي:** ويتم حسابه بقسمة تكرار كل فئة على مجموع التكرارات، هذا فيما يخص التكرار النسبي المطلق.

أما فيما يخص التكرار النسبي المئوي فيتم حسابه بضرب التكرار النسبي المطلق × 100.

ت. **التكرارات التجميعية:** وينقسم هذا النوع إلى قسمين:

- **تكرارات تجميعية صاعدة:** وهي عبارة عن تكرار الفئة مضاف إليه مجموع التكرارات اللاحقة.
- **تكرارات التجميعية النازلة:** وزهي عبارة عن مجموع التكرارات منقوص منها تكرار الفئة السابقة.

مثال2: إليك الجدول التكراري التالي:

المطلوب:

1/ كم عدد الطلبة الذين يفوق طولهم عن 167 سم؟

2/ ما هو عدد الطلبة الذين يقل طولهم عن 183 سم؟

3/ ضف خانتين تبين فيهما التكرارات التجميعية الصاعدة والنازلة؟

4/ احسب مراكز الفئات؟

الحل:

$F \swarrow$	$F \nearrow$	مراكز الفئات	n_i	x_i
100	13	139	13	$]143 - 135]$
(13-100)87	(15+13)28	147	15	$]151 - 143]$
(15-87)72	(15+28)43	155	15	$]159 - 151]$
(15-72)57	(20+43)63	163	20	$]167 - 159]$
(20-57)37	(15+63)78	171	15	$]175 - 167]$
(15-37)22	(12+78)90	179	12	$]183 - 175]$
(12-22)10	(8+90)98	187	8	$]191 - 183]$
(8-10)2	(2+98)100	195	2	$]199 - 191]$
00			100	**المجموع** $\sum$

1/ عدد الأشخاص الذين يفوق طولهم 167 سم هو:

إما بجمع التكرارات أو نأخذ القيمة المقابلة لـ 167 سم في ت ق ف: **37 شخص (2+8+12+15)**

2/ عدد الأشخاص الذين يقل طولهم عن 183 سم: **90 شخص.**

استنتاجه من التكرار المتجمع الصاعد أو جمع التكرارات:

90 شخص (13+15+15+20+15+12)

العرض البياني للبيانات الإحصائية

I. العرض البياني للمتغير الكيفي:

يمكن التمييز بين عدة أشكال بيانية خاصة بالمتغيرات الكيفية، أهم هذه العروض البيانية طريقة الدائرة، طريقة الأعمدة المستطيلة، طريقة العود المجزأ.

1. **طريق الدائرة:** يتمثل في دائرة مقسمة إلى عدة أجزاء كل جزء يقال زاوية مركزية تتناسب مع التكرارات المقابلة لكل خاصية من الخصائص المدروسة، ولتحقيق ذلك نضيف عمودا إلى جدول المعطيات يحتوي على الزوايا المركزية المقابلة لكل تكرار:

$$x_i = \frac{n_i \times 360°}{\sum n_i}$$

مثال:

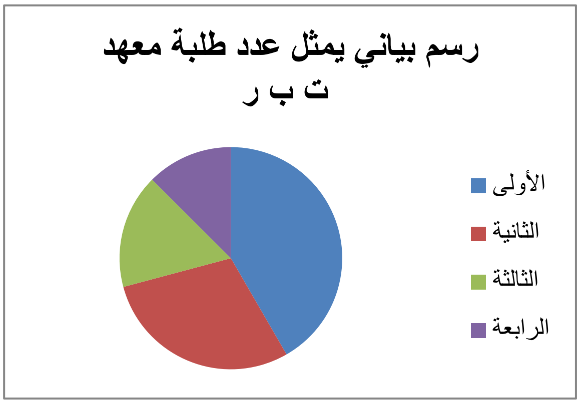
عدد الطلبة المسجلين بمعهد ت ب ر 12000 طالب، عدد الطلبة المسجلين في السنة الأولى 5000 طالب، السنة الثانية 3500، السنة الثالث 2000 طالب، السنة الرابعة 1500 طالب.

المطلوب:

تنظيم البيانات السابقة في جدول إحصائي وتمثيله بيانيا؟

الحل:

السنوات	n_i	x_i
الأولى	5000	150°
الثانية	3500	105°
الثالثة	2000	60°
الرابعة	1500	45°
المجموع	12000	360°



$$\alpha_1 = \frac{5000 \times 360°}{12000} = 150° \qquad \alpha_3 = \frac{2000 \times 360°}{12000} = 60°$$

$$\alpha_1 = \frac{3500 \times 360°}{12000} = 105° \qquad \alpha_4 = \frac{1500 \times 360°}{12000} = 45°$$

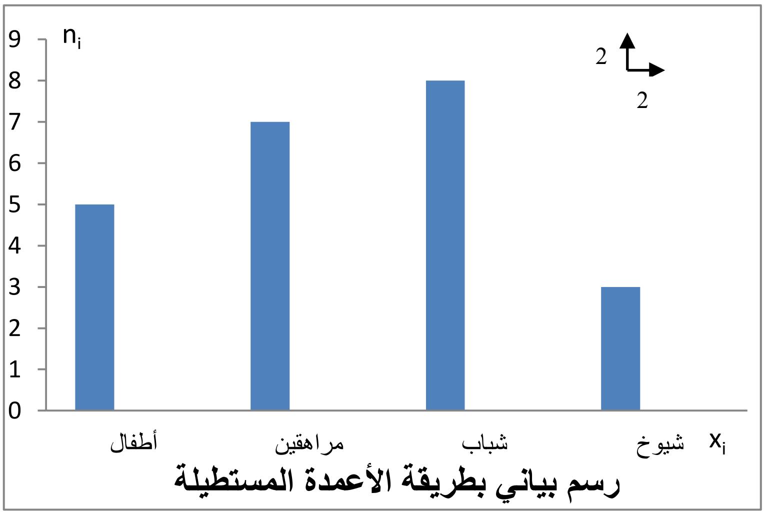
2. الأعمدة المستطيلة:

هي عبارة عن مستطيلات متباعدة بمسافات ثابتة، ولها قواعد متساوية، وتتناسب أطوالها مع التكرارات المقابلة لمكونات الخاصية المدروسة.

مثال: قدم لكم الإحصاء العام لسنة 2012 لعدة عائلات فكان كما يلي:

فئة الأعمار x_i	n_i
أطفال	5
مراهقين	7
شباب	8
شيوخ	3
المجموع	23

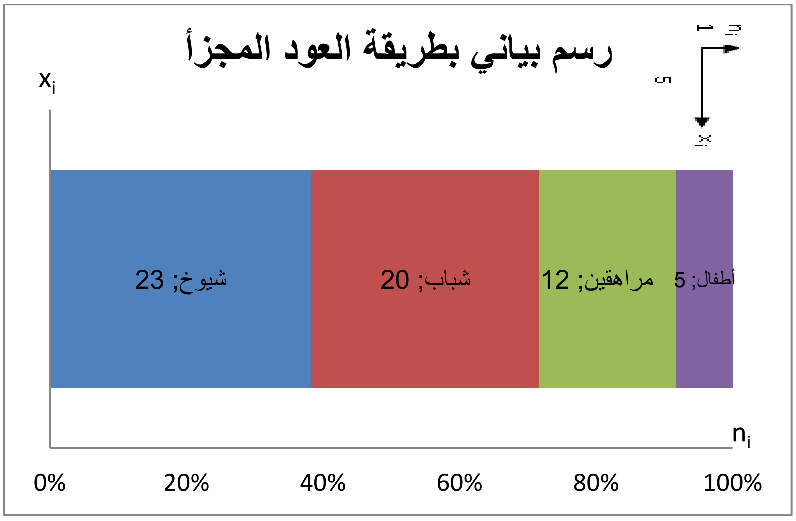
المطلوب: مثل هذه المعطيات في شكل أعمدة مستطيلة؟



3. العمود المجزأ:

هو عبارة عن مستطيل مقسم إلى عدة أجزاء، كل جزء يقال تكرار معين للخاصية المدروسة.

مثال: نأخذ نفس المثال السابق ونمثله بطريقة العمود المجزأ.



II. العرض البياني في حالة متغير كمي منقطع:

نكتفي في هذه الحالة بنوعين من العروض البيانية:

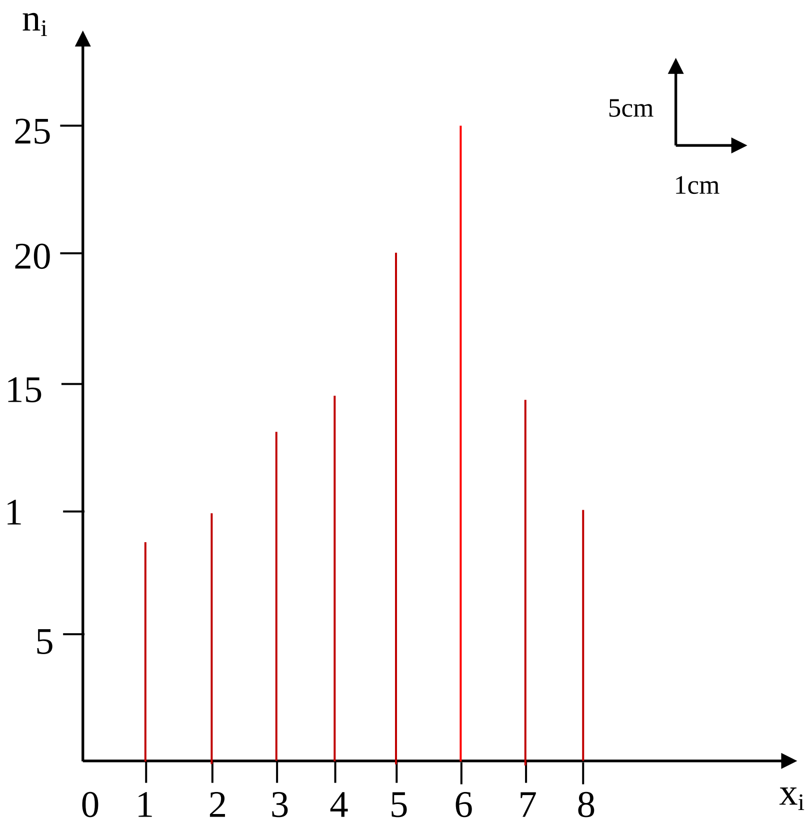
أ. العرض البياني للتكرارات البسيطة:

هي عبارة عن أعمدة بسيطة تتناسب أطوالها مع التكرار المقابل لقيمة معينة للمتغير المدروس وتسمى **بالأعمدة البسيطة**.

مثال: يبين الجدول التالي عدد الأطفال في العائلة لـ 125 أسرة.

المطلوب: العرض البياني للتكرارات البسيطة.

$F\swarrow$	$F\nearrow$	n_i	x_i
125	6	6	**0**
119	15	9	**1**
110	25	10	**2**
100	39	14	**3**
86	55	16	**4**
70	75	20	**5**
50	100	25	**6**
25	115	15	**7**
10	125	10	**8**
00	–	125	$\sum$



عرض بياني للتكرارات البسيطة يبين عدد الأطفال في 125 أسرة

ب. العرض البياني للتكرارات التجميعية الصاعدة والنازلة:

ب.1. العرض البياني للتكرارات التجميعية الصاعدة:

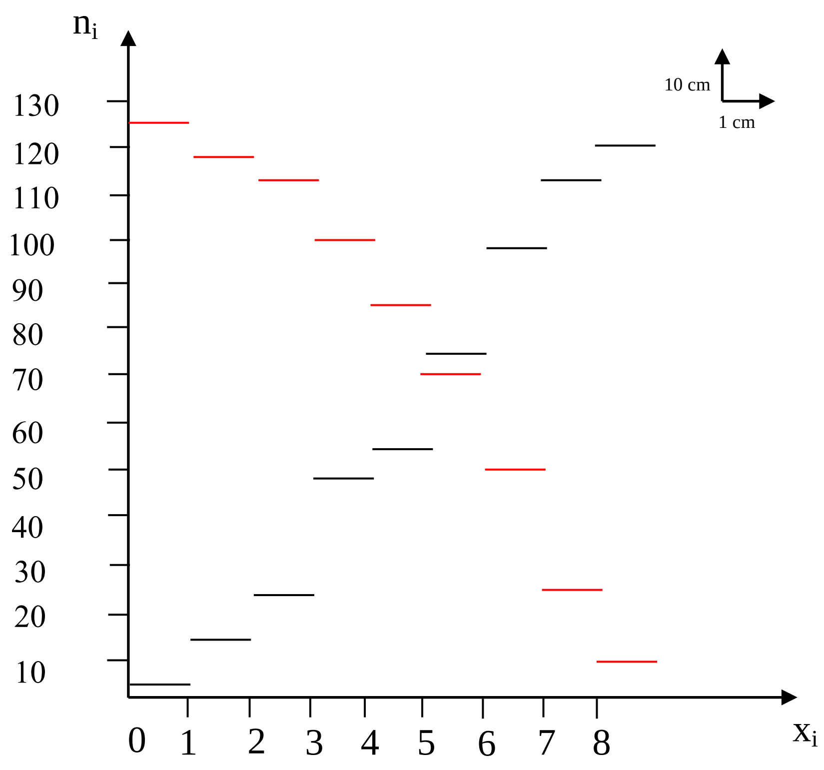
تسمى بطريقة السلم، هو عبارة عن قطع مستقيمة متصاعدة حسب تصاعد التكرارات التجميعية الصاعدة المقابلة لكل قيمة من قيم المتغير الإحصائي المدروس.

ب.1. العرض البياني للتكرارات التجميعية النازلة:

هو عبارة عن قطع مستقيمة متنازلة حسب تنازل التكرارات التجميعية النازلة، حيث أن القطعة الأولى تقابل مجموع التكرارات....

مثال: نفس المثال السابق مع إضافة عمودين للتكرارات التجميعية الصاعدة والنازلة.

المطلوب: العرض البياني للتكرارات التجميعية الصاعدة والنازلة.



عرض بياني للتكرارات التجميعية الصاعدة والنازلة بطريقة السلم

III. العرض البياني في حالة متغير كمي مستمر:

إن العروض البيانية للمتغير الإحصائي المستمر هي أكثر العروض البيانية استعمالا، ومن أهمها ما يلي:

1. المدرج التكراري:

يكون على شكل مستطيلات متلاصقة طول كل مستطيل منها يتناسب مع التكرار المقابل، وقاعدة كل منها تساوي طول الفئة المقابلة.

يمكن أن نميز بين حالتين عند وضع المدرج التكراري.

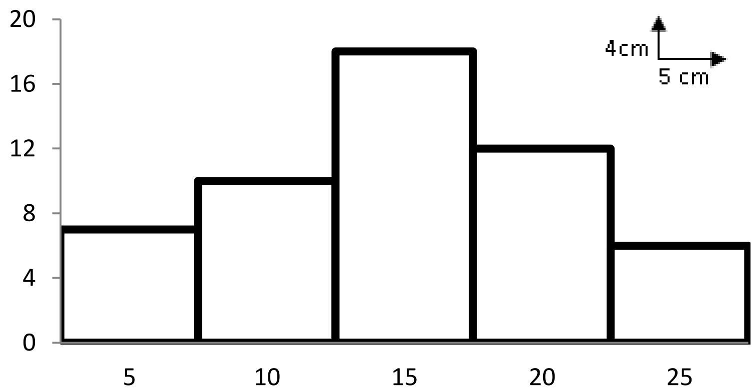
مثال:

- **حالة عدم انتظام طول الفئة:** يقصد بهذا الانتظام بتساوي طول الفئة لمختلف الفئات المشكل للمتغيرة المدروسة، ويعطى التمثيل البياني للمدرج التكراري كما في المثال التالي:

مثال: ليكن لديك جدول التوزيع التكراري التالي:

المطلوب: التمثيل البياني؟

x_i	n_i
$]5-0]$	7
$]10-5]$	10
$]15-10]$	18
$]20-15]$	12
$]25-20]$	6
$\sum$	53



عرض بياني يمثل المدرج التكراري

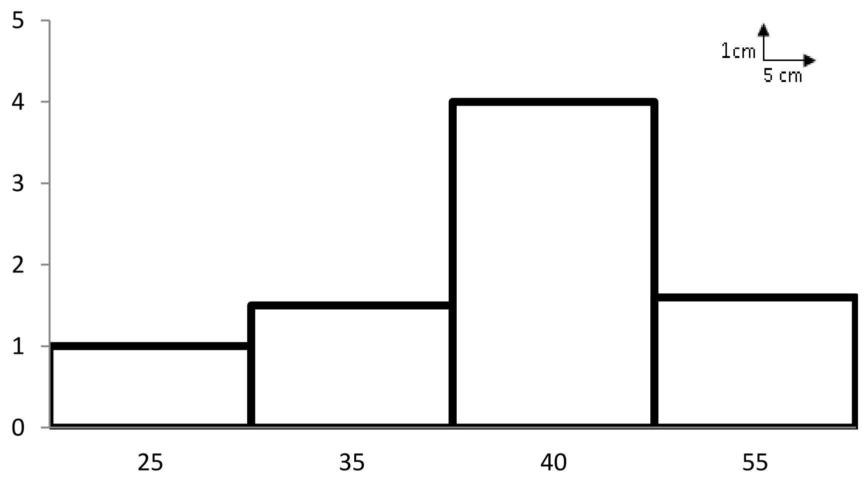
- **حالة عدم انتظام طول الفئة:**

في هذه الحالة يتم تعديل التكرارات وذلك وفق طول كل فئة حتى نتمكن من إجراء المقارنة بين الفئات ويجري هذا التعديل فنتحصل بذلك على تكرارات معدلة تقوم على إثرها برسم المدرج التكراري.

مثال: ليكن لدينا الجدول التالي:

x_i	n_i	K_i طول الفئة	تكرار معدل n_i/k_i
$]25-20]$	5	5	1
$]35-25]$	15	10	1.5
$]40-35]$	20	5	4
$]55-40]$	25	15	1.66
$\sum$	65	-	-

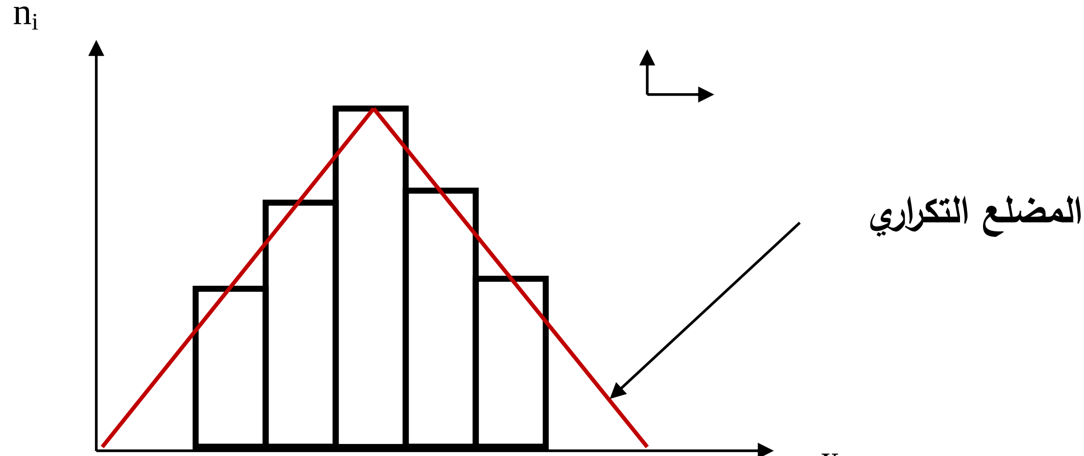
المطلوب: التمثيل البياني؟



رسم بياني يمثل المدرج التكراري بالتكرارات المعدلة

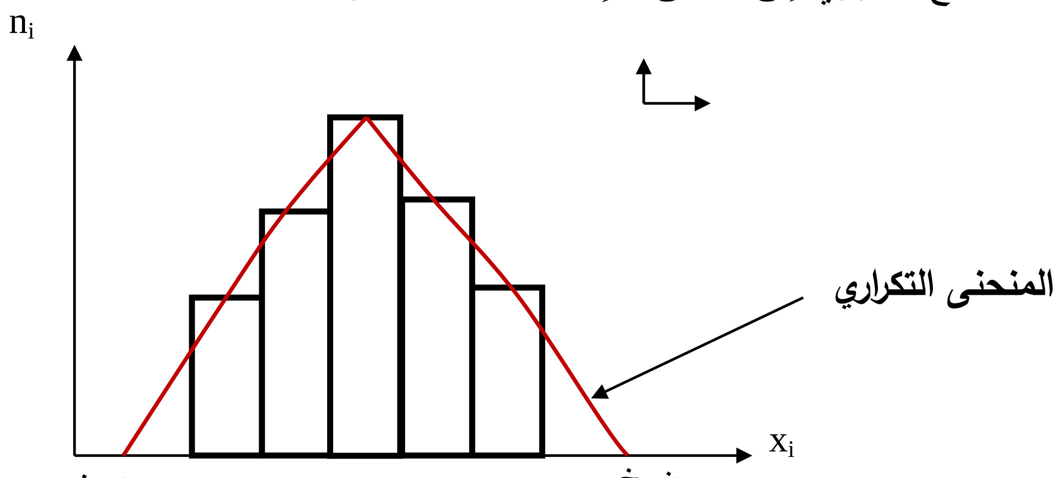
2. المضلع التكراري:

يتم رسم هذا العرض البياني بتحديد مراكز الفئات بحيث يتم تحديد مراكز الفئات في قمة المدرجات التكرارية وإيصال هذه المراكز (النقاط) بخطوط مستقيمة حيث نبدأ رسم هذه الخطوط من وسط فئة خيالية طولها يكون بطول الفئة الأولى ونفس الأمر بالنسبة للخط الأخير ينتهي بوسط فئة خيالية يكون منطبق على محور الفئات على النحو التالي:



3. المنحنى التكراري:

هو عبارة عن جعل المضلع التكراري يتخلى عن الانكسارات والذبذبات المميزة له حيث يتم تحويل المضلع التكراري إلى منحنى بالإيصال النقط المميزة لوسط الفئات عند كل تكرار.



ب.العرض البياني للتكراري التجميعية في حالة متغير مستمر:

ب.1. العرض البياني للتكرارات التجميعية الصاعدة:

يسمى بالمنحنى التجميعي الصاعد، يرسم هذا المنحنى عن طريق إيصال مجموعة النقاط ذات الإحداثيات التالية:

الحدود العليا للفئات والتكرارات التجميعية الصاعدة المقابلة لها.

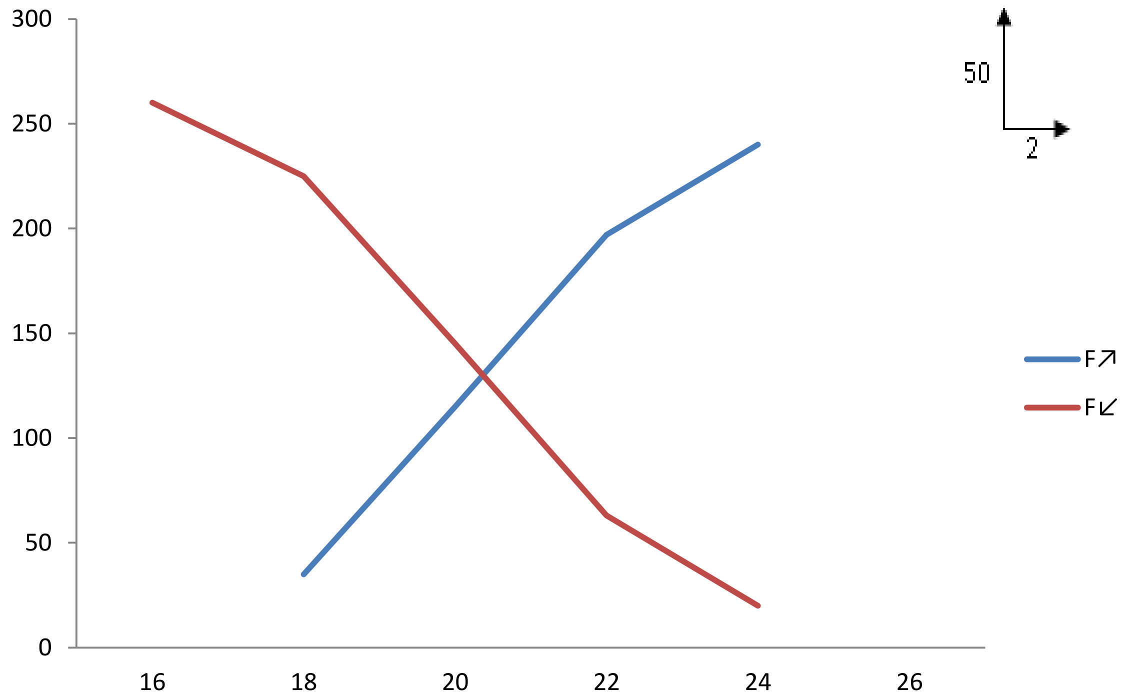
ب.2. العرض البياني للتكرارات التجميعية النازلة:

وهو عبارة عن المنحنى التجميعي النازل، يتم رسمه بإيصال مجموعة النقاط ذات الإحداثيات التالية:

الحدود الدنيا للفئات والتكرارات النازلة المقابلة لها.

مثال: إليك جدول التوزيع التكراري التالي:

$F \swarrow$	$F \nearrow$	n_i	x_i
260	35	35	$]18 - 16]$
225	115	80	$]20 - 18]$
145	197	82	$]22 - 20]$
63	240	43	$]24 - 22]$
20	260	20	$]26 - 24]$
-	-	260	$\sum$



رسم بياني للتكرارات التجميعية الصاعدة والنازلة

مقاييس التركة المركزية

تمهيد:

رأينا في الفصل السابق كيف يتم عرض البيانات الإحصائية جدوليا وبيانيا من أجل نقل وصف عام وسريع للظاهرة المدروسة ومن أجل وضع ترتيب معين وضروري لهذه المعلومات الإحصائية. غير أن لهذه الطريقة حدود من بينها:

1- لا يمكن استخدامها في الأسلوب الشفهي.

2- لا يمكن استخدامها لتحليل المعطيات.

3- لا يمكن الاستفادة منها في مجال الاستقراء الإحصائي (التنبؤ واتخاذ القرارات).

ولهذه الأسباب وضعت مقاييس عددية وصفية يمكن استخدامها في مجالات عديد منها التحليل والتنبؤ واتخاذ القرار، ومن بين هذه المقاييس، مقاييس النزعة المركزية أو مقاييس الوضع.

مفهوم النزعة المركزية:

تميل البيانات الإحصائية إلى التمركز حول قيمة معينة وكلما ابتعدنا عن هذه القيمة فإن عدد المعلومات يبدأ في التناقص، نسمي هذه الظواهر بالتركة المركزية.

قياس التركة المركزية:

يعتبر قياس النزعة المركزية عن مركز التوزيع الإحصائي، ولقياس هذه النزعة نستعمل عدة مقاييس من بينها:

1-المتوسط الحسابي:

وهو عبارة عن مجموعة القياسات الخاصة بظاهرة معينة على عدد هذه القياسات.

أ- حالة المتوسط الحسابي للبيانات الغير مبوبة:

هو حاصل جمع القيم مقسوم على عددها.

فإذا كان لدينا القيم: $x_1, x_2, x_3, \ldots\ldots x_n$

فإن ‾ يكون: $\bar{X} = \frac{\sum x_i}{n}$

ط2: طريقة الانحرافات: الوسط الفرضي:

$$\bar{X} = x_0 + \frac{\sum(x_i - x_0)}{n}$$

مثال: لتكن لدينا القيم التالية:

2، 8، 4، 3، 9، 6، 7، 10

احسب المتوسط الحسابي لهذه القيم؟

الحل:

$$\bar{X} = \frac{\sum x_i}{n} = \frac{49}{8} = 6.125$$

ت. حالة المتوسط الحسابي للبيانات المبوبة:

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

مثال:

الجدول التالي يبين لنا العلامة المحصل عليها من طرف مجموعة من الطلبة:

x_i	n_i	$x_i n_i$
5	1	5
6	2	12
7	4	28
8	1	8
9	1	9
10	1	10
$\sum$	10	72

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

$$\bar{X} = \frac{72}{10}$$

$$\bar{X} = 7.2$$

- **حالة جدول كمي مستمر:**

نستعمل نفس علاقة الوسط الحسابي في حالة التوزيع التكراري غير أن الذي ينقصنا هي القيم النقطية للمتغير الإحصائي x_i، لأن هذه القيم معطاة على شكل مجالات جزئية (فئات)، وكل نستبدل هذه الفئات بمراكزها، ليصبح x_i عبارة عن مراكز الفئة.

$x_i n_i$	x_i	n_i	الفئات
3	1	3	$[0-2[$
18	3	6	$[2-4[$
45	5	9	$[4-6[$
70	7	10	$[6-8[$
72	9	8	$[8-10[$
77	11	7	$[10-12[$
26	13	2	$[12-14[$
331	-	45	Σ

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

$$\bar{X} = \frac{311}{45}$$

$$\bar{X} = 6.91$$

2. الوسيط « Md »:

هي تلك القيمة التي تقسم المجتمع الإحصائي إلى قسمين متساويين بحيث تكون قيم المتغير الإحصائي مرتبة.

أ. الوسيط في حالة البيانات الغير مبوبة:

يتم تحديده حسب الخطوات التالية:

1-نرتب قيم السلسلة الإحصائية ترتيبا تصاعديا أو تنازليا.

2-نحدد ترتيب الوسيط: وهنا لابد أن نميز بين حالتين:

- عندما يكون عدد القيم N فردي، في هذه الحالة، ترتيب الوسيط هو عبارة عن:

$$R = \frac{N+1}{2}$$

- أما عندما يكون عدد القيم N زوجي، فلا توجد قيمة وسيطية، وفي هذه الحالة فإن قيمة الوسيط هي عبارة عن الوسط الحسابي للقيمتين ذات الترتيب:

$$R_1 = \frac{N}{2}\ ;\ R_2 = \frac{N}{2} + 1$$

3-نضع قيمة الوسيط، ونرمز له بالرمز Md.

مثال: تبين السلسلة الإحصائية التالية علامات 9 طلبة في مقياس الإحصاء.

المطلوب: تحديد الوسيط؟

13، 14، 7، 17، 15، 9، 8، 11، 10

الحل:

- **الترتيب التصاعدي:** 7، 8، 9، 10، 11، 13، 14، 15، 17.
- **تحديد ترتيب الوسيط:**

$$R = \frac{N+1}{2} = \frac{9+1}{2} = 5$$

- **تحديد الوسيط:** Md $\pm$11

مثال2: تبين السلسلة الإحصائية التالية علامات 10 طلبة في مقياس علم الاجتماع الرياضي:

6، 7، 9، 10، 11، 12، 12، 13، 15، 16.

المطلوب: حساب الوسيط؟

الحل:

- **نحدد ترتيب الوسيط:**

$$R_1 = \frac{N}{2} = \frac{10}{2} = 5\ ;\ R_2 = \frac{N}{2} + 1 = \frac{10}{2} + 1 = 6$$

$$Md_1 = 11 \qquad ; \qquad Md_2 = 12$$

- **تحديد الوسيط:**

$$Md = \frac{Md_1 + Md_2}{2} = \frac{11+12}{2}$$

$$Md = 11.5$$

ب. الوسيط في حالة البيانات المبوبية أو في حالة التوزيع التكراري:

تختلف طريقة حساب الوسيط في هذه الحالة عن الحالة الأولى وهذا حسب الخطوات التالية:

1−تحديد التكرار التجميعي الصاعد.

2−تحديد ترتيب الوسيط وهو عبارة عن نصف مجموع التكرارات:

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

3−تحديد الفئة الوسيطية أي الفئة التي يقع فيها الوسيط، وهي الفئة التي تقابل التكرار التجميعي الصاعد الذي يساوي ترتيب الوسيط أو الأكبر منه مباشرة.

4−تحديد وحساب الوسيط بتطبيق العلاقة الإحصائية التالية:

$$Md = A + \frac{\frac{\sum n_i}{2} - M_{n-1}}{ESMd} \times M_{Md}$$

حيث:

A: الحد الأدنى للفئة الوسيطية.

K_{Md}: طول الفئة الوسيطية.

M_{n-1}: التكرار التجميعي الصاعد السابق للفئة الوسيطية.

ESMd: التكرار الأصلي للفئة الوسيطية.

مثال: ليكن لديك جدول التوزيع التكراري التالي:

$F\nearrow$	n_i	x_i
4	4	$]130-120]$
11	7	$]140-130]$
21	10	$]150-140]$
31	10	$\boxed{]160-150]}$
38	7	$]170-160]$
42	4	$]180-170]$
45	3	$]190-180]$
-	45	$\sum$

$$R=\frac{\sum n_i}{2}=\frac{45}{2}$$

$$R=22.5$$

الفئة الوسيطية:

$$A=150\,;\;M_{n-1}=21$$

$$R=22.5\,;ESMd=10$$

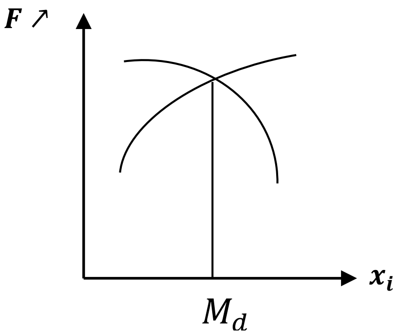
$$K=10$$

$$Md=A+\frac{R-M_{n-1}}{ESMd}\times K_{Md}$$

$$Md=150+\frac{22.5-21}{10}\times 10$$

$$Md=151.5$$

ملاحظة: يتم تحديد الوسيط بيانيا من خلال رسم المنحى البياني للتكرارات التجميعية الصاعدة والنازلة، التقاطع بينهما يحدد قيمة الوسيط.



3.الربيعيات:

نستعمل نفس الطريقة المتبعة لإيجاد الوسيط، غير أن الذي يتغير هو الترتيب وما يترتب عنه.

تتقسم الربيعيات إلى ثلاثة أقسام:

الربيعي الأول، الربيعي الثاني (الوسيط)، الربيعي الثالث.

أ. **الربيعي الأول:** هي قيمة المتغير الإحصائي التي تقسم المجتمع الإحصائي إلى قسمين، يحتوي القسم الأول على 25% من الوحدات الإحصائية، أما القسم الثاني فيحتوي على 75% من هذه الوحدات، تقع قيمة الربيعي الأول في نهاية الربع الأول من التوزيع الإحصائي، وفي هذا الموضوع تكون مرتبته = 25% أو: $R=\frac{\sum n_i}{4}$ حسب الترتيب التصاعدي لقيم المتغير الإحصائي، ويرمز له بالرمز: Q_1

$$Q_1 = A + \frac{R - M_{n-1}}{ESQ_1} \times K_{Q_1}$$

مثال: تأخذ نفس المثال السابق:

$$R = \frac{\sum n_i}{4} = \frac{45}{4} = 11.25$$

الفئة الربيعية الأولى: [140−150]

$$Q_1 = A + \frac{R - M_{n-1}}{ESQ_1} \times K_{Q_1}$$

$$Q_1 = 140 + \frac{11.25 - 11}{11} \times 10$$

$$Q_1 = 140.25$$

ب. الربيع الثالث: Q_3: هو عبارة عن قيمة المتغير الإحصائي التي تقسم المجتمع إلى قسمين، يحتوي القسم الأول على 75% من المجتمع الإحصائي المدروس، أما القسم الثاني فيحتوي على 25% المتبقين منه حسب الترتيب التصاعدي للمتغير المدروس، ويرمز له بالرمزي Q_3 ومرتبته 75% أو $R = \frac{\sum n_i}{4}$

$$Q_3 = A + \frac{R - M_{n-1}}{ESQ_3} \times K_{Q_3}$$

مثال: نفس المثال السابق:

$$R = \frac{\sum n_i}{4} = \frac{3(45)}{4} = 33.75$$

الفئة الوسطية الثالث: [160−170[

$$Q_3 = A + \frac{R - M_{n-1}}{ESQ_3} \times K_{Q_3}$$

$$Q_3 = 160 + \frac{33.75 - 31}{7} \times 10$$

$$Q_3 = 163.92$$

4. العشريات:

العلاقة العامة للعشريات:

$$D_i = A + \frac{(\frac{\sum n_i}{100} - M_{n-1})}{ESD_i} \times K$$

5. المئينات:

العلاقة الإحصائية العامة للمئينات:

$$P_i = A + \frac{(\frac{\sum n_i}{100} - M_{n-1})}{ESP_i} \times K$$

حيث أن i هو مرتبة المتغير الإحصائي

مثال: ليكن التوزيع التكراري التالي:

x_i	n_i	$F \nearrow$
]4 – 0]	3	3
]8 – 4]	6	9
]10 – 8]	7	16
]12 – 10]	9	25
$\boxed{]16 – 12]}$	16	41
]20 – 16]	13	54
]26 – 20]	8	62
]30 – 26]	5	67
]40 – 30]	3	70

المطلوب:

1- حدد الوسيط؟

2- حدد الربيع الأول والثالث؟

3- حدد العشيري الرابع والسادس؟

4- احسب P_{75} ، P_{45}؟

$D_3 \quad , D_5$

$P_{45} \quad , P_{30}$

الحل:

1- حساب الوسيط:

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

$$Md = A + \frac{R - M_{n-1}}{ESMd} \times K_{Md}$$

$$Md = 12 + \frac{35 - 25}{16} \times 4$$

$$Md = 14.5$$

2. حساب Q_1:

$$R = \frac{\sum n_i}{4} = \frac{70}{4} = 17.5$$

$$Q_1 = A + \frac{R - M_{n-1}}{ESQ_1} \times K_{Q_1}$$

$$Q_1 = 10 + \frac{17.5 - 16}{9} \times 2$$

$$Q_1 = 10.33$$

حساب Q_3:

$$R = \frac{3\sum n_i}{4} = \frac{3(70)}{4} = 52.5$$

$$Q_3 = A + \frac{R - M_{n-1}}{ESQ_3} \times K_{Q_3}$$

$$Q_3 = 16 + \frac{52.5 - 41}{13} \times 4$$

$$Q_3 = 19.53$$

3.حساب العشيري الرابع:

$$R = \frac{3\sum n_i}{10} = \frac{4(70)}{10} = 28$$

فئته [12−16[

$$D_4 = A + \frac{R - M_{n-1}}{ESD_4} \times K_{D_4}$$

$$D_4 = 12 + \frac{28 - 25}{16} \times 4$$

$$D_4 = 12.75$$

حساب D_6:

$$R = \frac{6\sum n_i}{10} = \frac{4(70)}{10} = 42$$

$$D_6 = A + \frac{R - M_{n-1}}{ESD_6} \times K_{D_6}$$

$$D_6 = 16 + \frac{42 - 41}{13} \times 4$$

$$D_6 = 16.3$$

5.حساب P_{75}:

$$R = \frac{75\sum n_i}{100} = \frac{75(70)}{100} = 52.5$$

[16-20[

$$P_{75} = A + \frac{R - M_{n-1}}{ESP_{75}} \times K_{P_{75}}$$

$$P_{75} = 16 + \frac{52.5 - 41}{13} \times 4$$

$$P_{75} = 19.53$$

حساب P_{45}:

$$R = \frac{45\sum n_i}{100} = \frac{45(70)}{100} = 31.5$$

]12-16[

$$P_{45} = A + \frac{R - M_{n-1}}{ESP_{45}} \times K_{P_{45}}$$

$$P_{45} = 12 + \frac{31.5 - 25}{16} \times 4$$

$$P_{45} = A + \frac{R - M_{n-1}}{ESP_{45}} \times K_{P_{45}}$$

$$P_{45} = 13.62$$

6. المنوال « Mo »:

هي قيمة المتغير الإحصائي الأكثر انتشارا أو تكرارا، ويرمز له بالرمز « Mo ».

أ. في حالة المنوال للبيانات الغير مبوبة:

مثال: لدينا القيم التالية:

2، 7، 8، 3، 5، 7، 6، 7

المطلوب: أوجد المنوال:

الحل: Mo = 7

ملاحظة:

قد تتساوى قيمتين في تكراراتها فيصبح المنوال في هذه الحالة مزدوج (سلسلة ذات منوالين).

مثال2: إليك سلسلة القيم التالية:

25، 40، 40، 60، 35، 60، 15

أوجد المنوال؟

الحل: $Mo_1 = 40$, $Mo_2 = 60$ لأن كلا الرقمين تكرار مرتين.

ب. حالة المنوال للبيانات المبوبة:

إن طريقة حساب المنوال في هذه الحالة تختلف عن الحالة الأولى، ولحسابه يتبع الخطوات التالية:

- تحديد الفئة المنوالية: وهي الفئة التي تقابل أكبر تكرارا عندما يكون طول الفئة ثابت أو الفئة التي تقابل أكبر تكرار معدل عندما تكون أطول الفئات غير متساوية.
- تحديد المنوال:

$$Mo = A + \frac{D_1}{D_1 + D_2} \times K_{M_o}$$

A: الحد الأدنى للفئة المنوالية.

D1: تكرار الفئة المنوالية –تكرار الفئة التي سبقتها.

D2: تكرار الفئة المنوالية – تكرار الفئة التي تليها.

KMo: طول الفئة المنوالية.

مثال: إليك جدول التوزيع التكراري التالي:

n_i	x_i
23	]14 − 10]
50	]18 − 14]
80	]22 − 18]
60	]26 − 22]
25	]30 − 26]
15	]34 − 30]
253	$\sum$

الفئة المنوالية:]22−18]

لأنها تقابل أكبر تكرار

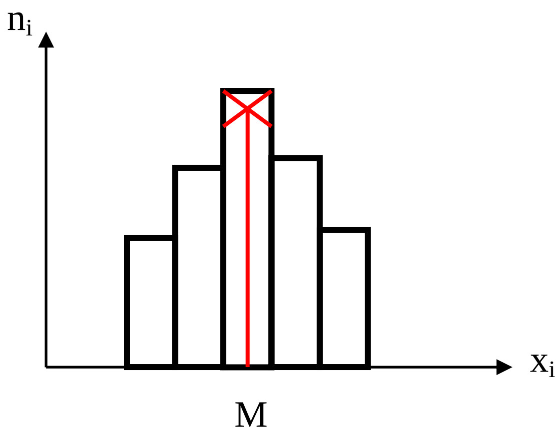
$$Mo = A + \frac{D_1}{D_1 + D_2}$$

$$Mo = 18 + \frac{30}{30 + 20}$$

$$Mo = 20.4$$

ملاحظة:

يمكن استنتاج المنوال بيانيا من خلال رسم المجرد التكراري، نأخذ أعلى درج (عمود) ونقوم بإيصاله بجانبتيه.



مقاييس التشتت

تمهيد:

إن مقاييس القيمة المركزية التي درسناها سابقا لا تسمح لنا بالوصف الكاملة لسلسلة ما، نأخذ مثلا المجموعتين التاليتين:

- **مج1:** 78 – 79 – 80 – 80 – 81 – 82.
- **مج2:** 40 – 60 – 60 – 80 – 80 – 80 – 100 – 100 – 120.

إن لهاتين المجموعتين نفس الوسيط والمنوال والمتوسط الحسابي والقيمة هي 80، رغم ذلك فهما غير متماثلتين في مج1 نجد أن القيم تتمحور حول القيمة النموذجية 80 بينما مج2 متشتتة، فنقول عندئذ أن المج1 أقل تشتتا من مج2..

1-المدى العام « E »:

يعتبر المدى أبسط مقاييس التشتت وأسهلها حسابا، وهو عبارة عن الفرق بين الحد الأعلى للفئة الأخيرة والأدنى للفئة الأولى أو الفرق بين أكبر قيمة وأصغر قيمة.

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

مثال: مج1: $E_1 = 82 – 78 = 04$

مج2: $E_2 = 120 – 40 = 80$

خصائص هذا المقياس:

- حسب تعريفه فإنه يتأثر بالقيم المتطرفة.
- لا يمكن تحديده في التوزيعات المبوبة المفتوحة.

- لا يعطينا نظرة دقيقة على باقي القيم للظاهرة المدروسة.

- يستعمل غالبا للمقارنة بين مجموعتين فأكثر.

2-الانحراف الربيعي (المدى الربيعي):

حسب تسميته فهو يعتمد في حسابه على الربيعيات أي الربيع الأول والربيع الثالث:

$$E_Q = Q_3 - Q_1$$

خصائصه:

- وهو يمثل المجال الذي تتوزع فيه 5% من القيمة الوسطى للظاهرة المدروسة.

- يعد أدق من المقياس الأول.

- يستعمل في المقارنة بين ظاهرتين فأكثر.

مثال: لتكن لدينا المجموعتين التاليتين:

- مج1: 19، 17، 16، 9،12، 7/ 3، 1.

- مج2: 15، 14، 13، 11، 9، 8، 7، 5.

الحل:

مج1: $R = \frac{N}{4} = \frac{8}{4} = 2$

$$Q_1 = 3$$
$$R = \frac{3N}{4} = \frac{24}{4} = 6$$
$$Q_3 = 16$$
$$E_Q = Q_3 - Q_1 = 16 - 3$$
$$E_Q = 6$$

مج1 أكثر تشتتا من مج2.

3-نصف المدى الربيعي:

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

4-الانحراف المتوسط:

يعرف الانحراف المتوسط بمجموعات المفردات على أنه متوسط حسابي للانحرافات، هذه المفردات عن متوسطها الحسابي بغض النظر عن إشارتها:

أ. **في حالة ب غ م:** $E_{\overline{X}} = \frac{\sum|x_i - \overline{X}|}{N}$

ب. **في حالة ب م:** $E_{\overline{X}} = \frac{\sum n_i |x_i - \overline{X}|}{\sum n_i}$

مثال1: لتكن السلسلة الإحصائية التالية:

13، 11، 10، 9، 8، 7، 5، 4، 1،2

المطلوب: تحديد الانحراف المتوسط؟

الحل:

- **حساب المتوسط الحسابي:** $\bar{X} = \frac{\sum n_i}{n} = \frac{20}{10} = 7$

$$\sum|x_i - \overline{X}| = 6 + 5 + 3 + 2 + 0 + 1 + 2 + 3 + 4 + 6 = 32$$

$$E_{\overline{X}} = \frac{\sum|x_i - \overline{X}|}{N} = \frac{32}{10} = 3.2$$

مثال2: ليكن التوزيع التكراري التالي:

المطلوب: تحديد الانحراف المتوسط؟

x_i	n_i	x_i	$x_i n_i$	$\lvert x_i - \bar{X}\rvert$	$n_i \lvert x_i - \bar{X}\rvert$
$[89.5 - 99.5[$	5	94.5	472.5	29.37	146.85
$[99.5 - 109.5[$	9	10.5	940.5	19.37	174.33
$[109.5 - 119.5[$	16	114.5	1832	9.37	149.92
$[119.5 - 129.5[$	25	124.5	3112.5	0.63	15.75
$[129.5 - 139.5[$	13	134.5	1748.5	10.63	138.19
$[139.5 - 149.5[$	7	144.5	1011.5	20.63	144.41
$[149.5 - 159.5[$	3	154.5	463.5	30.63	91.89
$[159.5 - 169.5[$	2	164.5	329	40.63	81.26
$\sum$	**80**	-	**5910**	-	**942.26**

الحل:

$$\bar{X} = \frac{\sum n_i x_i}{\sum n_i} = \frac{9910}{80}$$

$$\bar{X} = 123.87$$

$$E_{\bar{X}} = \frac{\sum \mathrm{n_i} |\mathrm{x_i} - \bar{\mathrm{X}}|}{\sum n_i} = \frac{942.6}{80}$$

$$E_{\bar{X}} = 11.76$$

ملاحظة:

يعتبر الانحراف المتوسط أحسن من سابقيه (المدى العام، المدى الربيعي) غير أنه لا يستعمل بشكل واسع نظرا لوجود القيمة المطلقة.

5-التباين والانحراف المعياري:

أ. التباين:

هو عبارة عن الوسط الحسابي لمربعات الفروق بين قيم المتغير الإحصائي والوسط الحسابي، ويرمز له بالرمز *V(x)*.

- **في حالة البيانات غ م:** $V(x) = \frac{\sum(x_i-\bar{X})^2}{N}$

- **في حالة البيانات م:** $V(x) = \frac{\sum n_i(x_i-\bar{X})^2}{\sum n_i}$

ب. الانحراف المعياري: وهو التجذر التربيعي للتباين، ويرمز له بالرمز SD.

- **في حالة البيانات غ م:** $SD = \sqrt{\frac{\sum(x_i-\bar{X})^2}{N}} = \sqrt{V(x)}$

- **في حالة البيانات م:** $SD = \sqrt{\frac{\sum n_i(x_i-\bar{X})^2}{\sum n_i}} = \sqrt{V(x)}$

مثال1: في حالة البيانات غ م:

إليك القيم التالية:

2، 4، 7، 9، 10، 12، 13، 15، 16

مطلوب: أوجد التباين والانحراف المعياري؟

x_i	$x_i - \bar{X}$	$(x_i - \bar{X})^2$
2	9.4 – 2= -7.4	54.76
4	-5.4	29.16
6	-3.4	11.56
7	-2.4	5.76
9	-0.4	0.16
10	-0.6	0.36
12	2.6	6.76
13	3.6	12.96
15	5.6	31.36
16	6.6	43.56

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

$$V(x) = \frac{196.4}{10}$$

$$V(x) = 19.64$$

$$SD = \sqrt{V(x)}$$

$$SD = \sqrt{19.64}$$

$$SD = 4.43$$

6- معامل الاختلاف (معامل التغاير):

ولمعرفة دلالة معنى الانحراف المعياري، وللحكم بوجود أو بعدم وجود تجانس بين القيم فإننا نلجأ إلى استخدام معامل التغاير CV.

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

$$CV = \frac{4.43}{9.4} \times 100$$

$$CV = 47.12\%$$

حيث إذا كانت قيمة معامل التغاير أقل من 30% يعني أنه يوجد تجانس بين القيم، أما إذا كان معامل التغاير أكبر أو يساوي 30% فهذا يعني أنه يوجد تباين واختلاف بين القيم.

مثال: إليك جدول التوزيع التكراري التالي:

$\mathbf{n_i(x_i - \bar{X})^2}$	$\mathbf{(x_i - \bar{X})^2}$	$\mathbf{x_i n_i}$	$\mathbf{n_i}$	$\mathbf{x_i}$
32.8	16.40	4	2	2
27.9	9.30	9	3	3
0.01	0.025	24	4	6
5.4	0.90	42	6	7
46.8	15.60	30	3	10
112.91	**-**	**109**	**18**	**3**

$$\bar{X} = \frac{\sum n_i x_i}{\sum n_i} = \frac{109}{18} = 6.05$$

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

$$V(x) = \frac{112.91}{18}$$
$$V(x) = 6.27$$
$$SD = \sqrt{V(x)}$$
$$SD = \sqrt{6.27}$$
$$SD = 2.5$$
$$CV = \frac{SD}{\bar{X}} \times 100$$
$$CV = \frac{2.5}{6.05} \times 100$$
$$CV = 41.32\%$$

$CV > 30\%$ إذن: يوجد تباين واختلاف بين القيم.

معامل الارتباط

يستخدم معامل الارتباط في المنهج الوصفي للتعرف على طبيعة وقوة العلاقة بين متغيرين أو أكثر، فعندما تغير في المتغير (x) يتبعه تغير في المتغير (y) فإن الباحث يهتم بدراسة التي تربط هذين المتغيرين، والتعرف على نوعية وقوة هذه العلاقة ويرمز له بالرمز « v ».

تقدير قوة العلاقة بين متغيرين:

-1   -0.80   -0.50   0   +0.50   +0.80   1

- إذا كان معامل *الارتباط* = +1: فالعلاقة موجبة تامة.
- إذا كان معامل الارتباط = −1 : فالعلاقة سالبة تامة.
- إذا كان معامل الارتباط = 0: فالعلاقة منعدمة.
- إذا كان معامل الارتباط أقل من 0.50 : فالعلاقة موجبة أو سالبة ضعيفة.

- إذا كان معامل الارتباط بين 0.80 و 0.50: فالعلاقة موجبة أو سالبة متوسطة.

- إذا كان معامل الارتباط يفوق 0.80: فالعلاقة موجبة أو سالبة قوية.

1-معامل الارتباط لبيرسون:

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

حيث: r: معامل الارتباط.

n: حجم العينة.

$x.y$: متغيرات.

مثال:

أوجد العلاقة بين متغير التحصيل الدراسي ومتغير التغيب من خلال الجدول التالي:

xy	y^2	x^2	**التحصيل** y	**التغيب** x	n
30	9	100	3	10	1
12	144	1	12	1	2
15	1	225	1	15	3
32	64	16	8	4	4
21	49	9	7	3	5
20	100	4	10	2	6
15	225	1	15	1	7
36	36	36	6	6	8
30	4	225	2	15	9
38	361	4	19	2	10
249	**993**	**621**	**83**	**59**	$\sum$**10**

لقياس العلاقة بين متغير التحصيل ومتغير التغيب نحتاج أولا إلى حساب الحدود التالية:

$\sum x$, $\sum y$, $\sum x^2$, $\sum y^2$, $\sum xy$ ثم نستخدم المعادلة التالية:

$$r = \frac{10(249) - (59)(83)}{\sqrt{[10(621) - (59)^2]} . \sqrt{[10(993) - (83)^2]}}$$

$$r = \frac{2490 - 4897}{\sqrt{(6210 - 3481).(9930 - 6889)}}$$

$$r = \frac{-2407}{\sqrt{(2729)(3041)}} = \frac{-2407}{2880.77}$$

$$r = -0.83$$

إن العلاقة بين التحصيل الدراسي والغياب علاقة عكسية سالبة ضعيفة قوية، ذلك أنه كلما زاد الغياب، قل التحصيل الدراسي، وكلما قل الغياب زاد التحصيل الدراسي.

2-معامل الارتباط سبيرمان:

يتعامل الباحث في كثير من الأحيان مع مستوى الرتب حيث تكون بيانات على شكل رتب، كونها لا تتوفر على شروط المسافات المتساوية أو الكمية، أو لأن أحد المتغيرين معبر عنه برتب، فيلجأ الباحث إلى تحول المتغير الكمي إلى رتب لقياس العلاقة بين المتغيرين والمعامل الأكثر شيوعا في هذه الحالة هو معامل سبيرمان يعطي هذا المعامل بالمعادلة: $r_s = 1 - \frac{\sum 6D^2}{n(n-1)}$

حيث:

r: معامل الارتباط سبيرمان.

1 و6: ثوابت (لا تتغير).

D: الفرق بين رتب نفس الفرد على المتغير x و y يعد ترتيب القيم تصاعديا أو تنازليا.

D^2: مربع الفرق بين رتب نفس الفرد على المتغيرين.

n: حجم العينة.

مثال: احسب معامل ارتباط سبيرمان للبيانات التالية:

D²	***D= Rx - Ry***	***Ry***	***Rx***	***y***	***x***	***n***
1	1	11	12	12	30	1
1	1-	12	11	13	28	2
0.25	0.5	9.5	10	11	26	3
0.25	0.5-	9.5	9	11	24	4
0	0	8	8	10	22	5
0	0	6.5	6.5	9	20	6
2.25	1.5	5	6.5	8	20	7
2.25	1.5-	6.5	5	9	18	8
0.25	0.5	3.5	4	7	16	9
0.25	0.5-	3.5	3	7	14	10
0	0	2	2	6	12	11
0	0	1	1	5	10	12
7.50	-	-	-	-	-	$\sum n$

$$r_s = 1 - \frac{6(7.50)}{12(144-1)}$$

$$r_s = 1 - \frac{45}{1716}$$

$$r_s = 0.97$$

العلاقة بين المتغيرين قوية وموجبة. (طردية).