# الفصل الاول :

**تعريفات ماده مبادىء الاحصاء**

**تعريف علم الاحصاء :**

**يختص علم الإحصاء بجمع وتبويب وتنظيم وعرض وتلخيص وتحليل واستقراء النتائج من البيانات الرقمية لعمل استدلالات واتخاذ القرارات في ظل عدم التأكد.**

**وينقسم علم الإحصاء إلي قسمين:**

**أ- الإحصاء الوصفي**

**يشمل جمع وتبويب وتنظيم وعرض وتلخيص البيانات الرقمية.**

**ب- الإحصاء الاستقرائي (التحليلي):**

**تحليل البيانات واستقراء النتائج واتخاذ القرارات**

**طرق عرض البيانات**

| المقارنه | اولا - طريقه الجداول | ثانيا –طريقه العرض البياني للبيانات الاحصائيه |
|---|---|---|
| | | **وله اربعه انواع للطلاق العرض :**<br>**اولا- طريقه المستطيلات او الاعمده**<br>**ثانيا- طريقه الخط المنكسر**<br>**ثالثا- طريقه الخط المنحنى**<br>**رابعا – طريقه الدائره** |

| | | |
|---|---|---|
| التعريفات | وهي وضع البيانات في جداول لعرض تغير ظاهره مع الزمن او مع مسميات كالبلدان والمدارس والجامعات وغيرها اومع الزمن والمسميات معا . | - طريقه المستطيلات او الاعمده :<br>وهي توضع المسميات او الزمن على المحور الافقي او العمودي ورسم مستطيل على كل مسمى ويكون ارتفاع كل مستطيل ممثلا للقيمه المقابله<br>٢- طريقه الخط المنكسر وتستخدم في :<br>تغير ظاهره مع الزمن مثل : تغير درجه حراره مريض مع الزمن بالساعات<br>تغير ظاهره مع المسميات مثل :تغير اعداد الطلاب في جامعه مع السنوات<br>تغير الظاهرتين معا مثل: تغير اعداد الطلاب حسب الكليات على مدى فتره زمنيه معينه<br>٣ – طريقه الخط المنحنى :<br>نفس طريقه الخط المنكسر ولكن بدون زوايا<br>٤- طريقه الدائره:<br>وهي تستعمل بتقسيم المتغير او الظاهره الكليه الى اجزاء حيث ان المجموع الكلي للظاهره يساوي مساحه الدائره ٣٦٠<br>ثم يتم تحويل كل جزء الى قطاع من الدائره |
| عند استعمال طريقه كلا من هذه الطرق وهي :طريقه الجداول والاعمده والمستطيلات | طريقة الجداول:<br>**عند استعمال هذه الطريقة يجب مراعاة :**<br>**(١)عنوان الجدول .**<br>**(٢) الوحدات المستعملة .**<br>**(٣) مذكرات المصادر التي أخذت منها البيانات.**<br>**(٤) مذكرات تفسر القيم الشاذة أن وجدت.** | **طريقة المستطيلات أو الأعمدة:**<br>**ولاستعمال الطريقة يجب مراعاة الأتي:**<br>**(١)عنوان الشكل.**<br>**(٢) الوحدات المستعملة .**<br>**(٣) مذكرات المصادر التي أخذت منها الإشكال.** |

**التوزيع التكراري**:

**التوزيعات التكرارية هي أحدي طرق تنظيم البيانات الكثيرة بحيث لا تخسر هذه البيانات من أهميتها.**

**والطريقة الأساسية لبناء التوزيع التكراري هي تقسيم مدي قيم البيانات إلي فئات وحصر عدد البيانات الواقعة ضمن كل فئة .**

**وعند بناء التوزيع التكراري يجب مراعاة:**

I. **يجب أن تكون الفئات منفصلة عن بعضها البعض.**
II. **أن تكون الفئات متساوية في الطول .**

III. يجب أن تكون الفئات كافئه لاحتواء جميع البيانات.

**التوزيع التكراري النسبي:**

هوالتكرار النسبي لكل فئة هو نسبة تكرار تلك الفئة إلي مجموع التكرارات (n) .

التوزيع التكراري المئوي:

وهو ضرب كل تكرار نسبي في ١٠٠%

**التوزيع التكراري المتجمع:**

وهويتم وضع الحدود الفعلية للفئات مع التكرارات المتجمعة المقابلة لها .ونبدأ دائما بالحد الادني الفعلي للفئة الأولي ونعتبر تكراره المتجمع صفراً.

ويمكن أن نحصل علي التكرار المتجمع النسبي اذا استعملنا التكرارات النسبية بدلاً عن التكرارات. وأيضاً بالنسبة للتكرار المئوي.

تمثيل التوزيعات التكرارية بيانيا:

**(١)المدرج التكراري :**

هو عبارة عن تمثيل تكرار كل فئة من فئات التوزيع التكراري بمستطيل حدود قاعدته الحدود الفعلية لتلك الفئة.

**(٢) المضلع التكراري:**

هو مضلع مغلق نحصل عليه بتنصيف الإضلاع العلوية للمستطيلات في المدرج التكراري ثم نوصل النقاط بعضها ببعض بخطوط مستقيمة، ولكي نغلق الخط المنكسر الذي حصلنا عليه نعتبر أن هناك فئتين متطرفتين واحدة إلي أقصي اليسار والثانية إلي أقصي اليمين تكرار كل منهما صفراً.وذلك بأخذ مركز كل من هاتين الفئتين.

**المضلع التكراري:**

وهناك طريقة أخري لرسم المضلع التكراري نعتبر مركز كل فئة احداثياً أفقيا لنقطة ونعتبر تكرار هذه الفئة ألأحداثي العمودي لتلك النقطة نبين جميع النقاط ونوصل فيما بينها بخطوط مستقيمة.نعين مركز فئة قبل الأولي مباشرة ونعتبر تكرارها صفراً ونعين مركز فئة بعد الفئة الأخيرة مباشرة ونعتبر تكرارها صفراً.نعين هاتين النقطتين ونغلق بواسطتهما المضلع التكراري

## (٣) المنحني التكراري :

هو تمهيد للمضلع التكراري نوصل المراكز بمنحني بدلاً من الخطوط المنكسرة

١-المضلع التكراري المتجمع:

نحصل علي المضلع التكراري المتجمع برصد التكرار المتجمع لأي فئة مقابل الحد الاعلي الفعلي لها ثم نوصل هذه النقاط فيما بينها بخطوط مستقيمة.

٢-المنحني التكرار المتجمع:

ونحصل علية بتمهيد المضلع التكراري

## أشكال التوزيعات التكرارية:

عند وصف البيانات توجد ثلاث خواص لابد من معرفتها وهي(الشكل والنزعة المركزية والتغير).

يتم التعرف علي شكل التوزيع التكراري من التوزيع نفسه أومن مضلعة أو مدرجه التكراري أو العرض بالساق والورقة الخاص به.

/ التمييز بين التوزيعات المتماثلة والتوزيعات غير المتماثلة .

٢/ التمييز بين التوزيعات ذات المنوال الواحد والتوزيعات ذات العدة منوالات.

٣/ التمييز بين التوزيعات كبيرة التفرطح ومتوسطة التفرطح وقليلة التفرطح (مدبب).

## الفصل الثاني :

مقدمة:

- ن الهدف الأساسي من استخدام مقاييس النزعة المركزية ومقاييس التشتت هو تلخيص البيانات في محاولة أخرى لوصفها عن طريق التعرف على مركزها ومقدار تشتت البيانات حول هذا المركز (درجة تجانس البيانات) ومن خلال هذين المؤشرين يتمكن الباحث من فهم أبعاد الظاهرة قيد الدراسة.
- كل ظاهرة في الحياة العامة لها ميل للتجمع حول نقطة معينة ، ومن ثم إذا استطعنا تحديد هذه النقطة فإننا سنصل إلى قيمة متوسطة تتجمع حولها القيم.

يسمى ذلك الميل إلى التجمع حول هذه القيمة بالنزعة المركزية
تسمى المقاييس المستخدمة

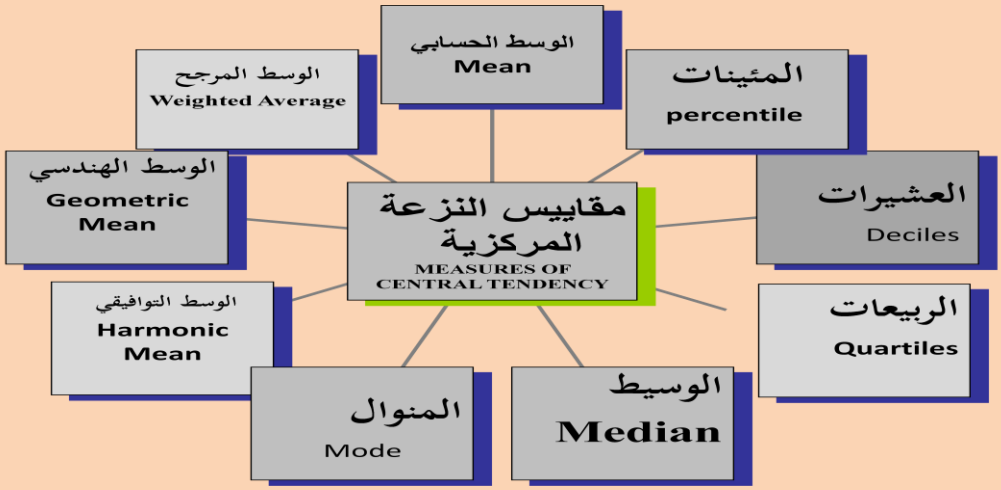
## مقاييس النزعة المركزية:

وهي القيم التي تتوزع حولها القراءات

## شروط المقياس الجيد:

- يحسب بطريقة سهلة لا تؤثر على دقة البيانات.
- يأخذ في الاعتبار جميع المفردات المطلوب حساب المقياس لها.
- يكون له معنى طبيعي مفهوم يستخدم في الحياة العامة .
- يعكس التغير في الظاهرة ، ولا يتغير بتغير طرق حسابه.
- يخضع للعمليات الجبرية خضوعا تاما .
- لا يتأثر بالقيم الشاذة أو المتطرفة.
- لا يتأثر باختلاف العينات ذات الحجم الواحد.

## مقاييس النزعة المركزية بأشكالها المتعارف عليها في علم الإحصاء




## مقاييس النزعة المركزية (الأساسية):

١-الوسط الحسابي ٢- الوسيط

٣- الوسط الهندسي ٤-المنوال

٥-الوسط التوافقي

## (أولا) الوسط الحسابي *Arithmetic Mean*:

- الوسط الحسابي يعتبر من أهم مقاييس النزعة المركزية والأكثر استخداما في الإحصاء والحياة العملية إذ يستخدم عاده في الكثير من المقارنات بين الظواهر المختلفة.
- يعتبر الوسط أكثر المقاييس الإحصائية انتشارا وشيوعا بين الباحثين لسهولته وفائدته التي تضفي عليه أهمية كبرى في حياتنا اليومية فكثيرا ما يتحدث الأفراد عن متوسطات الأسعار في الشهر الأول أو العام الأول ومتوسطات الأعمار واختلافاتها من جيل إلى جيل ، ومن بلد إلى بلد آخر ومتوسطات الدخل الشهري والسنوي ، وغير ذلك من الأمور العملية التي تتصل من قريب بحياتنا اليومية .

### الوسط الحسابي :

اي ان الوسط الحسابي يساوي مجموع القراءات مقسوما علي عددها

لاحظ أن المتوسط الحسابي يستخدم كل البيانات المتوفرة في طريقة حسابه.

### ملاحظة :

الوسط الحسابي يعتبر من أهم مقاييس البيانات الكمية ولا يستخدم مع البيانات النوعية

الوسط الحسابي المرجح:

- إذا كان لدينا مجموعات ذات إعداد مختلفة من البيانات وعلم الوسط الحسابي لكل مجموعة. كيف نحصل علي الوسط الحسابي للمجموعات إذا دمجت معا؟

قاااااااعده :

لدينا مجموعه ذات $N_1$ من القيم ووسطها الحسابي $X_1$

ومجموعه ثانيه ذات $N_2$ من القيم ووسطها الحسابي X2

فأن الوسط الحسابي للمجموعات ذات $N_1 + N_2$ عن القيم الناتجه

من دمج المجموعتين

**الوسط الهندسي**

إذا كان لدينا N من الإعداد الموجبة $x_1, x_2, \ldots, x_n$

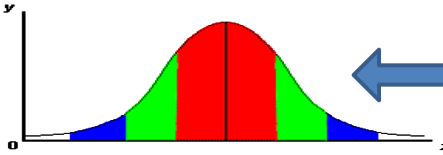
فان وسطها الهندسي يعرف بالمعادلة:

$$G = \sqrt[N]{X_1 X_2 \cdots\cdots\cdots X_N}$$

## (ثانيا) الوسيط Median

- **الوسيط :** هو مقياس أخر من مقاييس النزعة المركزية ويمثل ألمشاهده التي تكون التكرارات التي تسبقها تساوي التكرارات التي تليها. أوهو النقطة التي تقع تماما في منتصف توزيع الدرجات بحيث يسبقها نصف عدد الدرجات ويتلوها النصف الآخر أو هو القيمة التي تقسم البيانات إلى مجموعتين متساويتين وذلك بعد ترتيبها تصاعدياً أو تنازلياً. ومن خلال التعريف للوسيط نجد أنه يتأثر بالدرجات الوسطى أكثر مما يتأثر بالدرجات المتطرفة، وهو يصبح بهذه الصفة على.

نقيض المتوسط الذي يتأثر بالدرجات المتطرفة أكثر من تأثره بالدرجات الوسطى . ولذا يصلح الوسيط كمقياس للنزعة المركزية أكثر من المتوسط عندما تكون أطراف التوزيع متراكمة متجمعة غير مستوية كأي يلتوي التوزيع التكراري ناحية اليمين أو يلتوي ناحية



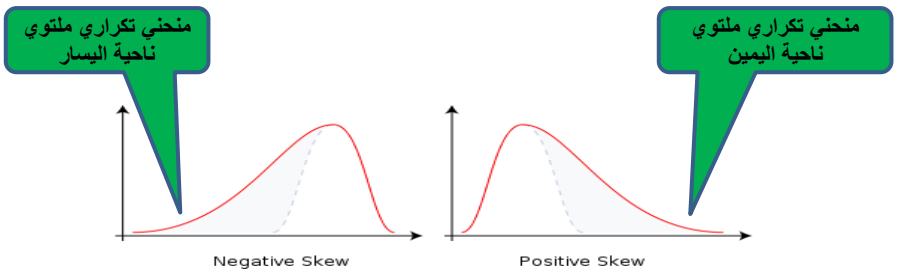
لمنحني التكراري المعتدل

يصلح الوسيط لنفس الميادين التي صلح فيها المتوسط ، أي في المعايير والمقارنة وخاصة عندما يكون التوزيع التكراري للدرجات ملتويا أي مرتفعا من أحد طرفيه ، والالتواء قد يكون موجبا أو سالبا : فإذا زاد تجمع تكرار الدرجات نحو الطرف الأول للتوزيع سمي الالتواء موجبا ، وإذا زاد تجمع تكرار الدرجات نحو الطرف الثاني للتوزيع سمي الالتواء سالبا ، وإذا اعتدل التوزيع التكراري سمي التوزيع معتدل وهذا يعني أن الوسيط يصلح كمقياس للنزعة المركزية في الالتواء الموجب والسالب ، فيما يصلح المتوسط كمقياس للنزعة المركزية إذا كان التوزيع معتدلا .

## (ثالثا) المنوال Mode

هو عبارة عن القيمة الأكثر تكراراً أو شيوعاً في العينة أو يدل المنوال على أكثر الدرجات شيوعا ،أي هي النقطة التي تدل على أكثر درجات التوزيع تكرارا

**المنوال**
**أقل مقاييس النزعة المركزية تأثر بالقيم الشاذة**




### طريقة حساب الوسيط من البيانات الغير مبوبة

إذا كانت قيم المتغير (x) هـــى $x_1 , x_2 , \dots , x_n$

حيث (n) يمثل حجم المجموعة ؛ فإن الوسيط يكون هو المفردة التي رتبتها الأتي (بعد الترتيب *إما تصاعديا أو تنازليا*.)

## مقاييس النزعة المركزية للتوزيع التكراري:

### الوسط الحسابي من التوزيع التكراري:

تعريف: اذا كان لدينا توزيع تكراري عدد فئاته h وكانت مراكز الفئات (اوالقيم) $X_1,X_2,X_3,...,X_n$ وكانت التكرارات المقابلة لها $f_1,f_2,f_3,...,f_n$

**خطوات حساب الوسط الحسابي**

١/ إيجاد مجموع التكرارات $n = \sum_{i=1}^{h} f_i$

**٢/** حساب مراكز الفئات $X_i$

**٣/** ضرب مركز الفئة في التكرار المناظر له $x_i * f_i$

وحساب مجموع $x*f$

٤/ حساب الوسط الحسابي بتطبيق القاعدة

**الوسط الهندسي للتوزيع التكراري:**

إذا كان لدينا توزيع تكراري عدد فئاته h وكانت مراكز الفئات أو(القيم) $x_1, x_2, \ldots, x_n$

وكانت التكرارات المقابلة لها $f_1, f_2, \ldots\ldots\ldots, f_h$ *فان الوسط الهندسي:*

$$G = \sqrt[N]{X_1^{f_1} X_2^{f_2} \ldots\ldots X_h^{f_h}}$$

$$N = \sum_{i=1}^{h} f_i$$

- الوسيط للتوزيع التكراري

حساب الوسيط من الجدول التكراري (البيانات المبوبة)

من جدول التكرار المتجمع الصاعد أو النازل أو الرسم البياني لكلاهم أو أحدهم وسنورد هنا تفصيلا كاملا لذلك .

- تعريف :

الفئة الوسيطية هي أول فئة يزيد تكرارها المتجمع عن $\frac{n}{2}$ أو يساويه حيث n هي مجموع التكرارات .

ولإيجاد الوسيط للتوزيع التكراري ذي الفئات نفرض أن:

h:عدد الفئات .

n :مجموع التكرارات أي : $n=\sum_{i=1}^{h} f_i$

C :طول الفئة الوسيطية.

a:الحد الادني الفعلي للفئة الوسيطية

$n_1$:التكرار المتجمع للفئة التي تسبق الفئة الوسيطية مباشرة.

b : الحد الادني الفعلي للفئة الوسيطية .

$f_m$ :تكرار الفئة الوسيطية .

**المنوال للتوزيع التكراري:**

❖الفئة المنوالية (أو الفئات)التي يقابلها اكبر تكرار تسمي الفئة المنوالية (الفئات المنوالية ).

❖مركز الفئة المنوالية يسمي المنوال التقريبي وإذا كان هناك عدة فئات منوالية فيكون لدينا عدة منوالات تقريبية

## مقارنة بين صفات الوسط الحسابي والوسيط والمنوال

| | الوسط الحسابي | الوسيط | المنوال |
|---|---|---|---|
| ١ | الوسط الحسابي أكثر مقاييس النزعة المركزية استعمالا لسهولة حسابه وسهولة تعريفه ويخضع للعمليات الحسابية. | أما الوسيط فهو سهل الحساب وسهل التعريف ولا يتأثر بالقيم المتطرفة | المنوال أقل مقاييس النزعة المركزية استعمالا وفي البيانات قليلة العدد عديم الفائدة ، أما في البيانات الكبيرة العدد فله معني معقول . |
| ٢ | نستطيع معرفة مجموع القيم إذا علمنا الوسط الحسابي ومجموع التكرارات .(من التعريف هو مجموع البيانات مقسوماً علي عددها). | ولا يعتمد علي جميع القيم دائماً ،فتغيير قيمه من القيم ربما يؤثر في قيمته أو لا يؤاثر. | وهو يتأثر بطريقة اختيار الفئات. وهو لا يتأثر بالقيم الشاذة. |

| | | | |
|---|---|---|---|
| ٤ | بمأن الوسط الحسابي يعتمد علي جميع القيم فإن قيمته تتغير إذا خذ فنا أو غيرنا في مفردات البيانات . | الوسط الحسابي أكثر ثبوتا من الوسيط. | |
| ٥ | / مجموع انحرافات قيم البيانات عن وسطها الحسابي يساوي صفر: أذن الوسط الحسابي هو نقطة اتزان للمدرج التكراري وإذا أضفنا أي عدد من القيم المساوية للوسط الحسابي إلي البيانات فان هذه الإضافة لا تؤثر عليه أما إذا أضفنا مفردات تختلف قيمتها عن قيمة الوسط الحسابي فإن قيمته تتأثر. | | |
| ٦ | ومن أهم نواقصه تأثره الشديد بالقيم المتطرفة. | | |
| ٧ | يفضل استعمال الوسط الحسابي إذا كان التوزيع متماثلاً. | / يفضل استعمال الوسيط إذا كان التوزيع ملتوياً. | يفضل استعمال المنوال إذا كان التوزيع متعدد المنوالات. |

- مقاييس التشتت هي مقاييس عددية تستخدم لقياس اختلاف أو تشتت البيانات . والاختلاف أو التشتت لمجموعة من البيانات هو مقدار تفرق أو تباعد أو انتشار البيانات فيما بينها . فتشتت البيانات يكون صغيرا إذا كانت البيانات متقاربة فيما بينها والعكس بالعكس وأما البيانات المتساوية فلا اختلاف ولا تشتت فيها .ومقاييس التشتت تستخدم لوصف مجموعة البيانات وكذلك لمقارنة مجموعات البيانات المختلفة إذ أن مقاييس النزعة المركزية لا تكفي وحدها لوصف مجموعة البيانات أو مقارنة مجموعات البيانات المختلفة .ومن أشهر مقاييس التشتت نذكر :

- **المدى : Range** **التباين : Variance**

- **الانحراف المعياري : Standard Deviation . الانحراف المتوسط: The Mean Deviation**

- **معامل الاختلاف (أو التغير): Coefficient o Variation**

**١ / المدى Range** :

- يعتبر المدى من أسهل مقاييس التشتت تعريًفا وحسابا ويعطينا فكرة سريعة عن مدى تفرق البيانات .ويعرف المدى لمجموعة من البيانات بالصيغة التالية :

$$\text{Range} = X_{\max} - X_{nin}$$

حيث أن :

$X_{\max}$ = أكبر قيمة (للبيانات المفردة ) = الحد الاعلي للفئة العليا (للبيانات المبوبة)

$X_{\min}$ = أصغر قيمة ( للبيانات المفردة = الحد الادني للفئة الدنيا ( للبيانات المبوبة)

ألمدي = Upper – Lower

## التباين : Variance

التباين للبيانات المفردة ( غير المبوبة) إذا كانت عينه حجمها n وكان متوسطها هو $\overline{X}$ $X_1, X_2, ..., X_n$ فإن تباين العينة يعرف كما يلي :

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{(n-1)}$$

## الانحراف المعياري : Standard Deviation

الانحراف المعياري هو الجذر الربيعي الموجب للتباين

## الانحراف المتوسط : Deviation The Mean

القيم المطلقة للانحراف هي مجموع الانحرافات المطلقة عن المتوسط بينما يعرف الانحراف المتوسط علي أنه انحراف قيم المفردات عن متوسطها الحسابي بغض النظر عن إشارات الانحرافات ويرمز له بالرمز M.D . ولما كان مجموع الانحرافات عن المتوسط يساوي صفرا إذا راعينا الإشارة التي تسبق الانحراف ولكن إذا أهملنا الإشارة نحصل علي مقياس آخر للتشتت هو مجموع القيم المطلقة للانحراف.

$$\sum_{i=1}^{n} \left| X_i - \overline{X} \right|$$

## التباين للبيانات المبوبة: (للتوزيع التكراري) :

**تعريف** :اذا كانت مراكز فئات توزيع تكراري هي $X_1, X_2, \ldots\ldots, X_n$ وكانت التكرارات المقابلة لها $f_1, f_2, \ldots\ldots\ldots\ldots, f_n$ فالتباين يكون

**المدى للبيانات المبوبة:**

المدى = الحد الأعلي للفئه العليا – الحد الأدني للفئه الدنيا

**الانحراف المتوسط للبيانات المبوبة (توزيع التكراري):**

إذا كانت مراكز التوزيع التكراري $X_1, X_2, X_3, \ldots\ldots\ldots, X_n$

وكانت التكرارات المقابلة لها $f_1, f_2, f_3, \ldots\ldots\ldots\ldots, f_n$

فإن الانحراف المتوسط يكون

## الخطوات :

1. إيجاد الوسط الحسابي $\overline{X}$ للتوزيع التكراري.
2. طرح كل مركز من الوسط الحسابي $X_i - \overline{X}$ .
3. اخذ القيمة المطلقة $|X_i - \overline{X}|$ .
4. ضرب $|X_i - \overline{X}|$ في التكرار $f_i$ .

**معامل التغير أو الاختلاف:**

أحد المقاييس المستخدمة لقياس درجة التشتت، وفيه يحسب قيمة التشتت كنسبة مئوية من قيمة مقياس النزعة المركزية ، ومن ثم يفضل استخدام معامل الاختلاف عند مقارنة درجة تشتت بيانات مجموعتين أو أكثر مختلفة لها وحدات قياس مختلفة، بدلا من الانحراف المعياري ، أو الانحراف الربيعي ، لأن معامل الاختلاف يعتمد على التغيرات النسبية في القيم عن مقياس النزعة المركزية، بينما يعتمد الانحراف الانحراف المعياري أو الانحراف الربيعي على التغيرات المطلقة للقيم ،

فعند مقارنة درجة تشتت بيانات الأطوال بالسنتمتر، وببيانات الأوزان بالكيلوجرام، لا يمكن الاعتماد على الانحراف المعياري في هذه المقارنة، وإنما يستخدم معامل الاختلاف

$$C.V = \frac{S}{\overline{X}} \times 100\%$$

## مقاييس الالتواء Measures of SKewness

- **تعريف** :

نعرف مقياس الالتواء لتوزيع تكراري أو مجموعة من البيانات بالمعادلة:

$$\gamma_1 = \frac{3(\overline{X} - M)}{S}$$

- ويستفاد من مقياس الالتواء في معرفة نوعية التواء التوزيع.إذا كان الالتواء موجبا يعني إن الوسط الحسابي اكبر من الوسيط وان الطرف الأيمن ممتد أكثر .وأيضا للمقارنة بين التواء توزيعين تكراريين أو مجموعتين من البيانات.

**الفصل الثالث :**

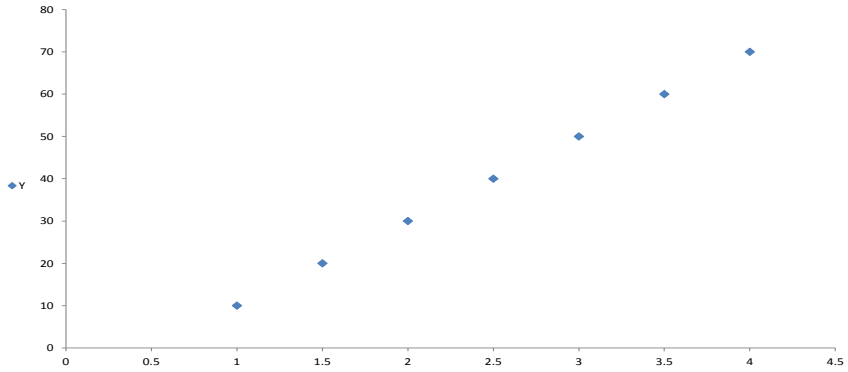
## معامل ارتباط بيرسون:

- لتكن (xi,yi) ،i=1,2,...,n عينة من الازواج المرتبة التي تعطي قيمتي متغيرين عشوائيين على n من العناصر. يعالج هذا البند العلاقة بين المتغيرين (x,y)
- ان هذا المعامل يصلح لقياس الارتباط الخطي أي عندما يبدوا ان النقاط في شكل الانتشار قريبة من خط مستقيم.
- الاشكال الانتشارية:

g

## هنالك اتجاه عام نحو ازدياد yكلما زادت x

# عدم وجود ارتباط




- تعريف معامل الارتباط لبيرسون على النحو التالي:

r = r(x , y) $\frac{SS_{xy}}{\sqrt{SS_x SS_y}}$

$$SS_{xy} = \sum_{i=1}^{n}(X_i - \overline{X})(y_i - \overline{y}) = \sum_{i=1}^{n}(x_i y_i) - n\overline{X}\,\overline{y}$$

$$SS_x = \sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 = \sum_{i=1}^{n} X_i^{\,2} - n\,\overline{X}^2$$

$$SS_y = \sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2 = \sum_{i=1}^{n} y_i^{\,2} - n\,\overline{y}^2$$

## معادلة خط الانحدار البسيط:

- يتناول هذا البند كيفية ايجاد معادلة خط الانحدار بين المتغيرين (y,x)
- ليكن نموذج خط الانحدار هو:

$$y = \beta_0 + \beta_1 x + \varepsilon$$

حيث:

- x: المتغير المستغل
- y : المتغير التابع
- E : متغير الخطأ

## نموذج خط الانحدار المقدر:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

يمكن ايجاد قيم $\hat{\beta}_0, \hat{\beta}_1$ بالعلاقات التالية:

$$\hat{\beta}_1 = \frac{SS_{xy}}{SS_x}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

## ملاحظة:

- يسمى المقدار (SSE) مجموع مربعات الخطأ

$$SSE = SS_y - \hat{\beta}_1 SS_{xy}$$

- معادلة خط انحدار Y علي X هي :

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{90}{5} = 18$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{215}{5} = 43$$

معامل التحديد $r^2$ : ( هو مربع الارتباط)

إذا كانت قيمة معامل الارتباط r=.9 فأن $r^2$ =0. 81

أسلوب خط الانحدار يقلل الخطأ في التقدير بمقدار 81%

وهذا يبرر استخدام خط الانحدار $\bar{y}$

إذا $r^2$=.01 فأن استخدام خط الانحدار للتنبؤ بقيم y عوضا عن $\bar{y}$ يقلل الخطأ بنسبة 1% وفي هذه الحالة لا جدوى من استخدام خط الانحدار ولهذا فأن قيمة $r^2$=0.1 تحديد فعلية خط الانحدار ،عليه يسمي **معامل التحديد**.

## الفصل الرابع :

### الاختبارات الإحصائية

- تنقسم اختبارات الإحصائية إلي :
  الاختبارات معلميه .
  الاختبارات اللامعلميه .
- **في معنى الإحصاء ألمعلمي والإحصاء اللامعلمي :**
  المعلم (parametric) مفردة تعني صفة أو خاصية لمجتمع معين في مقابل تقدير (estimate) التي تكون صفة أو خاصية لعينة ما ، واهم مايميز الاحصاء المعلمي عن اللامعلمي هو الوسط الحسابي والانحراف المعياري ولذلك فان علم الاحصاء

يميز بين شروط اختبار (ت) للعينات المستقلة و (ت) للعينات المترابطة لان الاخير لا يتعامل مع اوساط وانحرافات بينما نلاحظ ان اختبار (ت) للعينات المستقلة يتعامل معها اسوة بقوانين تحليل التباين (F) وقوانين (Z) ، من هنا ممكن فهم الاحصاء المعلمي بانه مجموعة من الطرق التي تتطلب تحقق افتراضات محددة حول المجتمع الذي تسحب منه العينة -وهنا مقتضى الدقة الانتباه للتعبير (حول المجتمع) كونه يختلف عن (حول العينة)

وبالتالي فان الإحصاء اللامعلمي هو مجموعة من الطرق البديلة التي تستخدم في حالات عدم تحقق الافتراضات حول المجتمع الذي تسحب منه العينة أو في حالة البيانات الاسمية والرتيبة ، وكلا الاحصائين (المعلمي،اللامعلمي) من طرق الاحصاء الاستدلالي التي يمكن تعميم نتائجها على المجتمع الا ان لكل منها مستوى ثقة معين يتحدد على ضوء البيانات المتوفرة وكذلك شروط تحقق الافتراضات .

**اختبار مربع كأي تربيع من الاختبارات اللامعلميه ويضم ثلاثة اختبارات هي:**

**١/ اختبارات الاستقلالية Independence Tests**

**٢/ اختبارات التجانس Homogeneity Tests**

**٣/ اختبارات حسن المطابقة Goodness-of-fit tests**

☺تحقيق الاماني ☺