الفصل الثاني

إعداد : د . قصي العمر

مقاييس النزعة المركزية

Measures of central tendency

مقدمة :

تمثل مقاييس النزعة المركزية أو المتوسطات مقاييس موقع ، حيث تعطينا قيمة تمثل البيانات أكثر من باقي القيم ، وسميت كذلك لأن جميع البيانات تنزع أن تتمركز حول هذه القيمة ،وهذه المقاييس شائعة الاستخدام في كثير من مجالات الحياة ، ويهتم بدراستها المشتغلين في هذه المجالات ، فهي تهم واضعي البرامج الاقتصادية والإنتاجية وذلك للتعرف على معدل الزيادة في السكان مثلاً ، أو متوسط إنتاج الدونم من المحاصيل المختلفة أو متوسط دخل الفرد .....الخ . وهي مقاييس متعددة ومتنوعة ، ومن أشهر هذه المقاييس : المتوسط أو الوسط الحسابي ، الوسيط ، والمنوال .

المتوسط الحسابي : The arithmetic mean

ويعتبر من أهم مقاييس النزعة المركزية وأكثرها استخداماً ، ويستخدم لإعطاء فكرة سريعة عن العينة والأرقام الموجودة فيها.

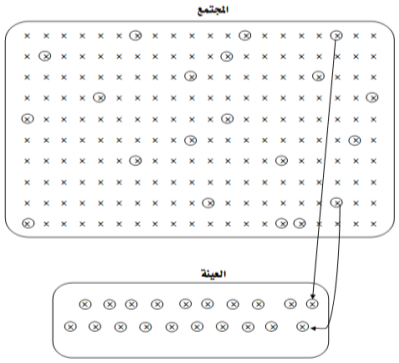
ويعرف المتوسط بأنه : **مجموع القيم مقسوماً على عددها .**

ويمكن التعبير عن هذا التعريف رياضياً بالعلاقة التالية :

$$\overline{x}=\frac{x_1+x2.....+.x_n}{n}$$

أو

$$\overline{x}=\sum_{i=1}^{n} x_i \ /n$$



المتوسط الحسابي

المتوسط الحسابي من البيانات الخام

مثال : احسب متوسط البيانات التالية:

3 , 5 , 1 , 4 , 2 , 3 , 5 , 1 , 4 , 2 , 3 , 5 , 1 , 4 , 2, 3 , 5 , 1 , 4 , 2

$$\overline{x}=\frac{3+5+1+4+2+3+5+1+4+2+3+5+1+4+2+3+5+1+4+2}{20}$$

$$\overline{x}=\frac{60}{20}=3$$

ومن مزايا المتوسط ما يلي :

١ -سهل الحساب وسهل التعريف .

٢ -وحيد لمجموعة من البيانات .

٣ -يدخل في الاعتبار عند حسابه جميع البيانات .

٤ -مجموع انحرافات القيم عن متوسطها الحسابي يساوي الصفر .

ومن عيوب المتوسط الحسابي ما يلي :

١ -تأثره الشديد بالقيم الشاذة أو المتطرفة .

٢ -لا يمكن حسابه لبيانات نوعية مثل نوع التربة ، لون العيون ، زمر الدم.....الخ .

الوسيط : The Median

يعتبر الوسيط أحد مقاييس النزعة المركزية وهو يعتبر متوسط مكاني حيث يكون عدد الأفراد التي قيمتها أقل منه مساوي لتلك التي قيمتها أكبر منه ، ويعرف الوسيط لمجموعة من الأعداد المرتبة ترتيبا تصاعديا أو تنازليا بأنه العدد الأوسط منها إذا كان عددها فرديا، وهو الوسط الحسابي للعددين الأوسطين إذا كان عددها زوجيا.

الوسيط

مثال (1): أوجد قيمة الوسيط للبيانات التالية:

27 , 12 , 15 , 300 , 22 , 18 , 14 , 24 , 17

الحل: 1- ترتيب البيانات: 12 , 14 , 15 , 17 , 18 , 22 , 24 , 27 , 300

الوسيط

2- نحدد موقع الوسيط (رتبته) كالتالي:

بما أن عدد البيانات : $n = 9$ وهو عدد فردي، إذن رتبة الوسيط هي:

$$\frac{9+1}{2} = 5 \quad \text{أي} \quad \frac{n+1}{2}$$

3- إذن قيمة الوسيط هي القيمة الخامسة في البيانات المرتبة وهي:

$$Med = 18$$

الوسيط

مثال (2): أوجد قيمة الوسيط للبيانات التالية:

3 , 9 , 12 , 16 , 14 , 33 , 10 , 15 , 20 , 7

الحل: 1- ترتيب البيانات: 3 , 7 , 9 , 10 , 12 , 14 , 15 , 16 , 20 , 33

القيمتين الوسيطيتين

2- تحديد موقع الوسيط: عدد البيانات $n = 10$ ، بما أن عدد البيانات عدد زوجي، فإن

الوسيط له رتبتين: $\frac{10}{2} = 5$ و $\frac{10}{2} + 1 = 6$

3- نحدد قيمة الوسيط، وهي متوسط القيمتين الوسيطيتين: الخامسة والسادسة، وهي:

$$\frac{12+14}{2} = 13$$

المنوال : The Mode

وهو القيمة التي تتكرر أكثر من غيرها في المفردات الإحصائية ، أو القيمة الأكثر تكرارا أو شيوعا ولهذا يطلق عليه أحيانا الشائع أو القيمة الشائعة . ويعتبر المنوال من أسهل مقاييس النزعة المركزية لأن الحصول عليه يتم بالنظر أكثر مما يتم بالحساب .

وقد يكون للمجموعة منوال واحد أو أكثر من منوال أو قد لا يوجد لها منوال.

مثال (١) :

أوجد المنوال للأعداد التالية: ٢٥، ٢٠، ٢٨، ٢٠، ١٥، ٣٠.

الحل:

بما أن القيمة (٢٠) تتكرر أكثر من غيرها و بناء على تعريف المنوال إذاً يكون المنوال لهذه المجموعة يساوي (٢٠).

مثال (٢) :

ما المنوال لمجموعة الأعداد التالية: ٥٦، ٤٠، ٥٦، ٥٦، ٧٢، ٣٤، ٤٩، ٧٢، ٧٢؟

الحل:

يوجد لهذه المجموعة منوالان هما: ٥٦، ٧٢ لأن كلا منهما يتكرر بنفس عدد المرات التي يتكرر فيها الآخر.

مثال (٣) :

ما المنوال لمجموعة الأعداد: ٢١، ١٦، ٢٤، ٩، ٨٤، ٣٨؟

الحل: لا يوجد منوال لهذه المجموعة لأن كلا منها يتكرر مرة واحدة

الفصل الثالث

إعداد : د . قصي العمر

مقاييس التشتت أو الاختلاف

Measures of Dispersion

مقدمة :

إن مقاييس النزعة المركزية غير كافية لتحديد صفات البيانات الإحصائية، فربما يكون لدينا ظاهرتان متساويتان في الوسط الحسابي والوسيط إلا أنهما مختلفتان.

فمثلاً إذا كانت درجات الحرارة في بلد ما خلال الليل و النهار هي: ٢٢، ٢٨, ٢١ ، ١٩، ١٠ فيكون معدل أو متوسط درجات الحرارة ٢٠ .

وإذا كانت درجات الحرارة في بلد آخر هي: ٢٤، ٢٩، ٢١، ١٦، ١٠ فيكون معدل أو متوسط

درجات الحرارة فيها ٢٠ أيضاً .

وهذا يعني أن معدل درجات الحرارة في البلدين متساوي ، إلا أنه بالنظر إلى مفردات البيانات في كل من البلدين نجد اختلافاً بينهما ، وهذا يعني أن الوسط الحسابي لا يكفي لوصف البيانات أو للحكم على تشابهها.

ويقاس التباعد بدرجة انتشار البيانات حول معدل أو وسط، وكلما كان التغير صغيراً اعتبرت البيانات متجانسة ، وكلما كبر التغير اعتبرت البيانات متباعدة أو متغايرة. وعلى ما تقدم فإننا بحاجة إلى مقاييس توضح مقدار تشتت أو تبعثر أو اختلاف البيانات عن بعضها البعض لاكتمال صورة وصف البيانات ، وتقوم مقاييس التشتت بهذا الدور .

ومن أهم هذه المقاييس ما يلي : ١- المدى المطلق ٢- التباين ٣- الانحراف المعياري ٤- الخطأ القياسي ٥- معامل الاختلاف

وفيما يلي شرح مبسط عن كل واحد من هذه المقاييس .

المدى المطلق : The range

يعرف المدى للبيانات بأنه الفرق بين أكبر قيمة وأصغر قيمة.

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

ويستخدم المدى عادةً لإعطاء فكرة سريعة أولية عن طبيعة توزيع المفردات الإحصائية لأنه يمتاز ببساطته وسهولة حسابه، لكنه لا يعكس أثر جميع المشاهدات لأنه يعتمد على أكبر وأصغر قيمتين فقط

المدى

الفرق بين أكبر قيمة وأصغر قيمة في مجموعة ما

A : 5 , 6 , 8 , 10 , 12 , 14 , 15

Range = 15-5 =10

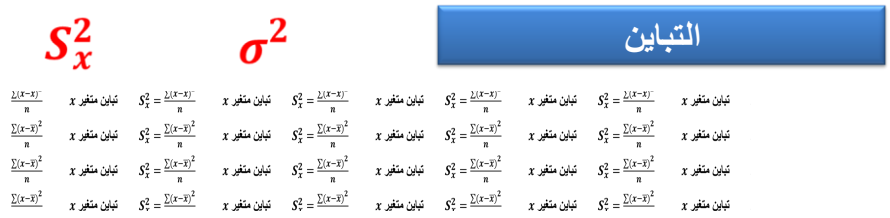
B : 1 , 2 , 5 , 10 , 15 , 18 , 19

Range = 19 -1 =18

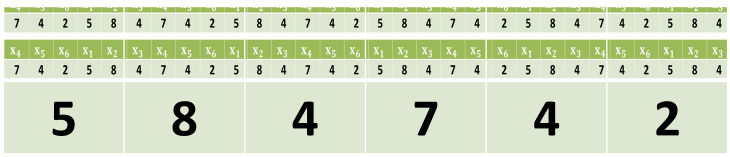
التباين والانحراف المعياري Variance and Standard Deviation

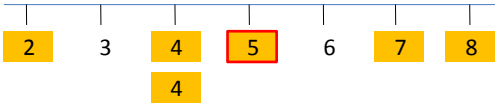
ويوضح هذان المقياسان مدى تقارب أو تباعد الدرجات عن المتوسط الحسابي ، و هما من أدق مقاييس التشتت وأكثرها استعمالا ، ويعرف التباين بأنه : متوسط مربعات انحرافات القيم عن متوسطها و يرمز له بالرمز (S_2) .بينما يعرف الانحراف المعياري بأنه : الجذر التربيعي الموجب لمتوسط مربعات انحرافات القيم عن متوسطها و يرمز له بالرمز (S) . ويمكن التعبير عن التعريفين رياضياً بالصيغتين التاليتين :

التباين داخل متغير x $S_x^2 = \frac{\sum(x-\bar{x})^2}{n}$ الانحراف المعياري $S_x = \sqrt{\frac{\sum(x-\bar{x})^2}{n}}$



مثال : أوجد تباين العينة الممثلة بالبيانات التالية:





التباين

(الحل)

$$S_x^2 = \frac{\sum(x-\bar{x})^2}{n}$$

$$= \frac{(5-5)^2+(5-8)^2+(5-4)^2+(5-7)^2+(5-4)^2+(5-2)^2}{6}$$

$$= \frac{(0)^2+(-3)^2+(1)^2+(-2)^2+(1)^2+(3)^2}{6}$$

$$= \frac{0+9+1+4+1+9}{6} = \frac{24}{6} = 4$$

الخطأ القياسي : Standard Error

يعبر الخطأ القياسي عن مقياس يمكن التعرف بواسطته عن مدى تشتت المتوسطات لعدد من العينات المأخوذة من نفس المجتمع ، فإذا كانت قيمته كبيرة دل ذلك على تبعثر المتوسطات ، وإذا كانت قيمته صغيرة اعتبر ذلك دليلاً على تركيز المتوسطات حول متوسط المجتم ع . وقد اعتاد الاحصائيون لاستخراج قيمة الخطأ القياسي بحسابه من العينة التي يعمل في إطارها معتمداً في ذلك على قيمة الانحراف القياسي للعينة وفق المعادلة التالية :

$$SX = S / \sqrt{n}$$

SX : الخطأ القياسي

S : الانحراف القياسي للعينة

n : عدد أفراد العينة

معامل الاختلاف : Coefficient of Variability

وهو أحد مقاييس التشتت النسبية ، ويستخدم لمقارنة المجتمعات أو العينات المختلفة ، وذلك عن طريق التعبير عن الانحراف القياسي للعينة أو المجتمع في صورة نسبة مئوية من المتوسط الحسابي ، وهو مقياس مجرد من وحدة قياس البيانات ، إذ يمكن مقارنة تشتت أطوال النباتات في سلالة الشعير مقيسة بوحدات من السنتمترات بتشتت أوزان محصول من نفس السلالة مقدراً بالكيلوغرام .

ويمكن التعبير رياضياً عن هذا المقياس وفق المعادلة التالية :

$$C.V = S / X * 100$$

C.v : معامل الاختلاف

S : الانحراف المعياري

X : المتوسط الحسابي

* : ضرب

/ : مقسوماً على

مثال :

بلغ متوسط أطوال عينة من نباتات الشعير 90cm وانحراف معياري 9 ، كما بلغ متوسط أوزان عينة من نباتات الشعير 30 kilogram وانحراف معياري 2 ، والمطلوب : أي العينتين أكثر اختلافاً أو تغيراً ؟

الحل :

نحسب معامل الاختلاف للعينة الأولى (الأطوال) ومعامل الاختلاف للعينة الثانية (الأوزان) ، بحيث يكون المعامل الأكبر هو الأكثر اختلافاً .

C.V 1 = 9 / 90 * 100 = 10 %

C.V 2 = 2 / 30 * 100 = 6.67 %

أي أن العينة الأولى أكثر اختلافا من العينة الثانية .