مقاييس النزعة المركزية
Measures of Central Tendency

- إن ما عرضناه في الجزء الأول، هو أسلوب سهل وبسيط ويمتاز بسرعة الإقناع.
- ويستقيم ذلك مع العديد من الباحثين الذين لا يميلون إلى استخدام أساليب أخرى في الدراسة والتحليل.
- يعتمد هذا الأسلوب حسبما أسلفنا على تمثيل الجداول الإحصائية بيانياً من خلال منحنى الظاهرة ، حيث يمكن تحديد العلاقات والخصائص والاتجاهات على أساسه .

- يصعب استخدام هذا الأسلوب كأساس لتحديد الخصائص والاتجاهات.
- وعليه فإنه من الأفضل اللجوء إلى أساليب يستخدم فيها الباحث الطريقة الرياضية في قياس التغيرات وتحديد درجة العلاقات وطبيعة الاتجاهات بطريقة موضوعية غير متحيزة. وذلك للاستدلال على ملخص لحقائق الظاهرة التي تمثلها مجموعة البيانات.
- في هذا الفصل نعرض لأحد المجموعات الرئيسية في طرق القياس الكمي لقيم التمركز في الظواهر وذلك لتحديد الاتجاهات العامة وتلخيص البيانات من خلال هذه القيم .

المتوسطات (مقاييس النزعة المركزية):

لماذا تسمى مقاييس النزعة المركزية؟

- حيث أن هذه القيمة النموذجية (المتوسط) تميل إلى الوقوع في المركز داخل مجموعة بيانات مرتبة حسب قيمتها، فإنها تسمى أيضاً بمقاييس النزعة المركزية.
- هناك صور عديدة للمتوسطات وأكثرها شيوعاً الوسط الحسابي أو ما يطلق عليه عادة الوسط، الوسيط، المنوال ، الوسط الهندسي والوسط التوافقي - ولكل من هذه المقاييس مميزاته وعيوبه وهذا يعتمد على البيانات والهدف من استخدام هذا المقياس أو ذاك .

الوسط الحسابي (The Arithmetic Mean):

- هو القيمة التي لو أعطيت لكل مفردة في مجموعة البيانات فإن مجموعها لا يتغير. وهي عبارة عن مجموع مشاهدات المجتمع مقسومة على عددها. X_1, X_2, X_3, X_n
- الوسط الحسابي أو الوسط للمجموعة n من الأرقام $\overline{X}$ ويرمز له بالرمز (ويقرأ X bar) ويعرف كالآتي:

$$\overline{X} = \frac{X_1 + X_2 + X_3 + \ldots X_n}{n} = \frac{\sum_{j=1}^{n} x_j}{n} = \frac{\sum X_j}{n} \qquad \ldots\ldots(1)$$

- مثال الوسط الحسابي للأرقام 8, 3, 5, 12, 10 هو:

$$\overline{X} = \frac{8+3+5+12+10}{5} = \frac{38}{5} = 7.6$$

- إذا كانت الأرقام $X_1, X_2, \ldots, X_k$ تحدث $f_1, f_2, \ldots, f_k$ مرة على الترتيب بمعنى أنها تحدث بتكرارات $f_1, f_2, \ldots, f_k$ فإن الوسط الحسابي سيكون:

$$\overline{X} = \frac{f_1 x_1 + f_2 x_2 + \ldots + f_k x_k}{f_1 + f_2 + \ldots + f_k} = \frac{\sum_{j=1}^{k} f_j x_j}{\sum_{j=1}^{K} f_j} = \frac{\sum fx}{\sum f} = \frac{\sum fx}{n} \quad \ldots\ldots (2)$$

- حيث $n = \Sigma f$ وهو مجموع التكرارات أي مجموع عدد المجالات

- مثال: إذا كانت 2 ,6 ,8 ,5 تحدث بتكرارات 1 ,4 ,2 ,3 على الترتيب فإن الوسط الحسابي سيكون:

$$\overline{X} = \frac{(3)(5)+(2)(8)+(4)(6)+(1)(2)}{3+2+4+1} = \frac{15+16+24+2}{10} = \frac{57}{10}$$

$$= 5.7$$

الوسط الحسابي المرجح (The Weighted Mean):

- في بعض الأحيان نقرن بعض الأرقام $X_1, X_2, \ldots, X_k$ بمعاملات ترجيح أو أوزان $W_1, W_2, \ldots, W_k$ وهذه تعتمد على الدلالة أو الأهمية المرتبطة بهذه الأرقام، وفي هذه الحالة فإن:

$$\overline{X} = \frac{W_1 X_1 + W_2 X_2 + \ldots + W_k X_k}{W_1 + W_2 + \ldots W_k} = \frac{\Sigma W_i X_i}{\Sigma W_i} = \frac{\Sigma W X}{\Sigma W} \qquad \ldots\ldots(3)$$

- يسمى بالوسط الحسابي المرجح، ويلاحظ هنا أوجه الشبة بالمعادلة (2) التي يمكن اعتبارها وسطاً حسابياً بأوزان $f_1, f_2, \ldots, f_k$

- مثال: إذا أعطي الامتحان النهائي في مقرر ما وزناً يعادل ثلاثة أمثال الامتحانات الشفوية، وإذا حصل طالب في الامتحان النهائي على 85 وفي الامتحانات الشفهية على 90 و 70 فإن متوسط تقديره هو:

$$\overline{X} = \frac{(1)(70) + (1)(90) + (3)(85)}{1+1+3}$$

$$= \frac{70 + 90 + 255}{5} = \frac{415}{5} = 83$$

خصائص الوسط الحسابي:

(أ) المجموع الجبري لانحرافات القيم عن وسطها الحسابي يساوي صفراً.

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

(ب) مجموع مربعات انحرافات مجموعة من الأرقام X_i عن أي رقم a يكون أصغر ما يمكن في حالة واحدة فقط إذا كانت $a = \overline{X}$.

(جـ) الوسط الحسابي لعدة مجموعات عبارة عن الوسط الحسابي المرجح لكل وسط حسابي لكل مجموعة مرجحاً بحجم هذه المجموعة .

$$\overline{X} = \frac{f_1 m_1 + f_2 m_2 + \dots + f_k m_k}{f_1 + f_2 + \dots + f_k} \qquad ......(4)$$

حيث m_i هي الوسط الحسابي للمجموعة i ، f_i عدد أفراد هذه المجموعة.

(د) إذا كانت A أي وسط فرضى، فإن الوسط الحسابي = الوسط الفرضي + متوسط الانحرافات عن الوسط الفرضي.

الوسيط (The Median):

- الوسيط لمجموعة من الأرقام مرتبة حسب قيمها (في منظومة) هي القيمة التي في المنتصف أو الوسط الحسابي للقيمتين في المنتصف.
- يستخدم عادةً في $\left(\frac{N+1}{2}\right)$ الحالات التي تحتوي فيها المفردات الإحصائية على قيم شاذة أو متطرفة

مثال 1 - مجموعة الأرقام 3,4,4,5,6,8,8,8,10

وسيطها هو 6 .

مثال 2 $\left(\frac{N}{2}\right)$ مجموعة $\left(\frac{N}{2}+1\right)$ الأرقام 5,5,7,9,11,12,15,18

وسيطها هو 10 = ½(9+11)

المنوال (The Mode):

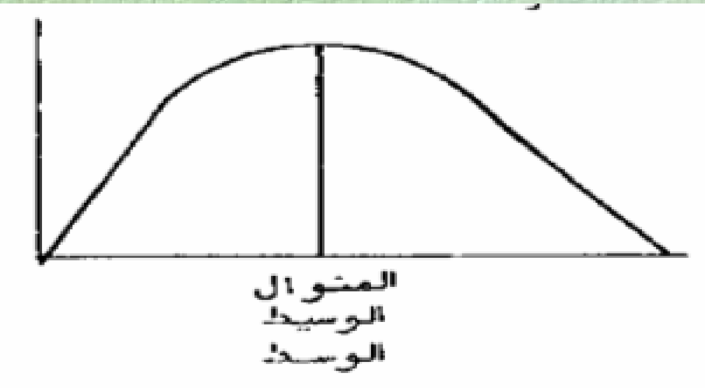
- المنوال لمجموعة من القيم هي القيمة التي تتكرر أكثر من غيرها ، أو القيمة الأكثر شيوعاً . وقد لا يكون للقيم منوال وقد يوجد منوال ولكنه غير وحيد.
  - مثال 1 - المجموعة 22,5,7,9,9,9,10,10,11,12,18 لها منوال واحد وهو 9 .
  - مثال 2 - المجموعة 3,5,8,10,15,16 ليس لها منوال .
  - مثال 3 - المجموعة 2,3,4,4,4,5,5,7,7,7,9 لها منوالان وهما 7,4 وتسمى مجموعة ذات منوالين Bimodal .

التوزيع الذي له منوال واحد يسمى وحيد المنوال

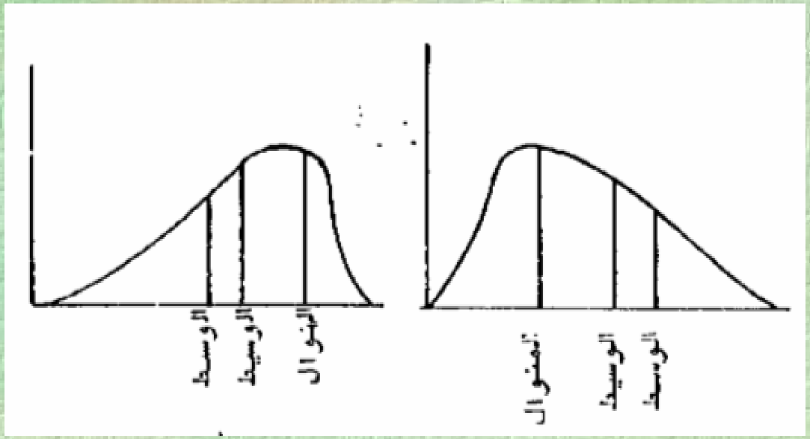
مقارنة المقاييس المركزية:

الوسط الحسابي	الوسيط	المنوال
▪ يوجد له نظرية رياضية. ▪ كل البيانات تستخدم في إيجاده. ▪ لبعض مجموعات البيانات قيمة ليست قريبة من وسط الجسم الأساسي لهذه البيانات. وبالتالي لا تعتبر مقياساً مقبولاً للنزعة المركزية.	▪ لا يوجد له نظرية رياضية صريحة تتعلق به. ▪ سهل الاحتساب، ولا يحتاج إلى أكثر من عددين. ▪ الوسيط يعتبر مقياساً مقنعاً، لأنه لا يتأثر بتفاوت القيم، حيث أن قيمته تتحدد بمنتصف المفردات فقط.	▪ لا يوجد له خواص رياضية. ▪ لا يحتاج إلى حسابات كثيرة وخصوصاً إذا نظر إليه بالنسبة لمركز الفئة المنوالية. ▪ يمكن استخدامه في حالة البيانات غير العددية مثل الأسماء، لون العيون، لون الشعر، وغيرها. ▪ يمكن أن يكون هناك أكثر من منوال.

العلاقة بين المنوال والوسط والوسيط:



1. تتطابق رقم مقاييس النزعة المركزية (المنوال والوسط والوسيط) في التوزيعات التكرارية المتماثلة.



2. في التوزيعات التكرارية وحيدة المنوال وغير المتماثلة تتحقق العلاقة الاعتيادية التالية: الوسط الحسابي – المنوال = 3 (الوسط الحسابي – الوسيط)

مقاييس التشتت
Measures of Dispersion

- لماذا هذه المقاييس؟
- لأن مقاييس النزعة المركزية لا تعطي وصفا كاملا للبيانات، حيث أنها تبين القيمة التي تتركز حولها بقية قيم العينة، ولكنها لا تعطي صورة عن مدى تباعد أو تقارب باقي القيم. بمعنى أنها لا توضح مدى التجانس في البيانات أو مدى تشتتها، لأن الاعتماد على مقياس واحد قد يكون مضللا. لذلك فالحاجة ملحة إلى مقياس آخر يوضح مدى تقارب أو تباعد هذه البيانات. فقد تتساوى مقاييس النزعة المركزية لمجموعتين أو أكثر من البيانات، إلا أنها قد تختلف في طبيعة توزيع

مثال: الوسط الحسابي

(أكثر تجانسا) مجموعة (أ): 16,15,14,13,12,11,10 13

(أكثر تشتتا) مجموعة (ب): 24,20,19,13,7,6,2 13

- بالرغم من تساوي الوسط (13) والوسيط (13) للمجموعتين (تشير إلى احتمال التكافؤ) إلا أن التوزيع مختلف.
- هناك عدة مقاييس للتشتت أهمها: الانحراف المعياري، التباين، الانحراف المتوسط، المدى، ونصف المدى.

المدى ونصف المدى

- يعرف المدى بأنه الفرق بين أكبر وأصغر قيمة في مجموعة البيانات.

في المجموعة (أ): 16,15,14,13,12,11,10

فإن المدى هو 6-10=6، ونصف المدى = $\frac{6}{2}=3$

أما المجموعة (ب): 24,20,19,13,7,6,2

فإن المدى هو 24-2=22 ، ونصف المدى = $\frac{22}{2}=11$

- ويعكس ذلك مدى التشتت في المجموعتين حيث أن مدى (أ) أقل من مدى (ب). وبالتالي فإن التشتت في (أ) أقل منه في (ب).

- ويتم اللجوء لنصف المدى الربيعي الذي يعرف بأنه نصف الفرق بين أكبر وأصغر قيمة من القيم المتبقية بعد استبعاد الربع الأعلى والأدنى من البيانات. (ليحل مشكلة المدى من حيث عيوبه ونواقصه).

- يعتبر المدى أقل مقاييس التشتت دقة وكفاءة ، وبالتالي أقلها صدقا:

(1) لا يأخذ بالاعتبار جميع قيم المجموعة (يعتمد على قيمتين فقط الكبرى والصغرى).

(2) تقل كفاءته مع زيادة أفراد العينة، حيث يسمح ذلك بوجود قيم متطرفة وشاذة.

(3) لا يعطي فكرة جيدة عن توزيع البيانات وتشتتها، فقد تتساوى مجموعتان أو عينتان في مدى توزيعها في حين تختلف متوسطاتهما، وقد يكون لأحدهما توزيع طبيعي

- للحصول على نصف المدى الربيعي تتبع الخطوات التالية:

(1) ترتيب البيانات تصاعديا أو تنازليا.

(أ) عندما تكون البيانات غير مبوبة يكون ترتيب الربيع الأول = $\frac{N+1}{4}$

(ب) عندما تكون البيانات مبوبة يكون ترتيب الربيع الأول = $\frac{\sum f}{4}$

(2) يحدد ترتيب موضع القيمة التي تقع في آخر الربيع الأول وتسمى Q_1.

(3) إيجاد قيمة Q_1 مع ملاحظة أن Q_2 = الوسيط.

(4) تحديد مرتبة Q_3 وهي التي تقع في آخر الربيع الثالث وتسمى بقيمة الربيع الثالث.

وتتحدد مرتبة الربيع الثالث = 3 X مرتبة الربيع الأول

(5) نوجد قيمة الربيع الثالث.

(6) نصف المدى الربيع (الانحراف الربيعي) = $\frac{Q_3 - Q_1}{2}$

- في المثال السابق

مجموعة (أ): 10, 11, 12, 13, 14, 15, 16

مجموعة (ب): 2 , 6 , 7, 13, 19, 20, 24

ترتيب $Q_1 = \frac{7+1}{4}$ ، ترتيب $Q_3 = 3 \times 2 = 6$

في المجموعة (أ) قيمة Q_1 هي المفردة رقم 2 = 11

Q_3 هي المفردة رقم 6 = 15

نصف المدى الربيعي = $2 = \frac{11-15}{2}$

الانحراف المتوسط

- يستخدم لتجنب النواقص أو العيوب الجوهرية في المقاييس السابقة (المدى ونصف المدى) فالمدى تعتمد قيمته على أكبر قيمة وأصغر قيمة، أما نصف المدى فيشمل نصف المفردات الإحصائية.
- ويمكن التعبير عنه بما يلي:

هو معدل الانحرافات المطلقة عن الوسط الحسابي.

لو كان لدينا مجموعة n من الأرقام $x_1, x_2 \ldots\ldots x_n$

فإن الانحراف المتوسط $= \dfrac{\sum \left| X_i - \overline{X} \right|}{n}$

$\overline{X}$ = الوسط الحسابي

$\left| X_i - \overline{X} \right|$ = القيمة المطلقة لانحرافات x_i عن الوسط الحسابي

$$= \frac{\text{مجموع الانحرافات المطلقة عن الوسط}}{\text{عدد المفردات}}$$

- مثال: أوجد الانحراف المتوسط لمجموع الأرقام:

$$2,3,6,8,11$$

$$\overline{X} = \frac{2+3+6+8+11}{5} = 6$$

$$\left[|6-11| + |6-8| + |6-6| + |6-3| + |6-2|\right]\frac{1}{5} =$$

الانحراف المتوسط $= \left[5+2+0+3+4\right]\frac{1}{5}$

$$= \frac{14}{5} = 2.8$$

وفي حالة البيانات المبوبة يكون الانحراف المتوسط = $\dfrac{\sum f\,|x_i - \bar{x}|}{n}$

- بالرغم من أن الانحراف المتوسط كانت له قيمة كبيرة لمدة طويلة، إلا أنه فقد أهميته في حالات معينة أهمها:

(1) يصعب إيجاده في التوزيعات التكرارية المفتوحة.

(2) يصعب الاستفادة منه خاصة في الإحصاءات الاستدلالية كما هو في الانحراف والتباين.

(3) يهمل الإشارات الجبرية بطريقة غير مبررة رياضيا، بالإضافة إلى صعوبة احتسابه عندما يكون الوسط الحسابي كسرا.

الانحراف المعياري

- مقياس للتشتت، مرتبط إلى حد كبير بالوسط الحسابي، حيث يعتمد على انحراف كل قيم البيانات عن الوسط الحسابي لها.
- في المثال السابق: بخصوص الأرقام في المجموعتين (أ) و (ب) نجد أن انحرافات قيم المجموعة (أ) عن وسطها الحسابي هي:

  [-3, -2, -1, 0, 1, 2, 3]
- بينما انحرافات قيم المجموعة (ب) عن وسطها الحسابي هي:

  [11, 7, 6, 0, -6, -7, -11]
- وهذا يدل أن القيم (أ) أكثر اقترابا من وسطها الحسابي من قيم (ب) عن وسطها الحسابي.

- ومن أجل تسهيل عملية المقارنة، فإنه يتوجب إيجاد قيمة واحدة تدل على كل مجموعة. وحيث أن مجموع انحرافات القيم عن وسطها الحسابي = صفر. لذلك فإنها لا يمكن أن تستخدم كمقياس. لذلك يتم اللجوء إلى استخدام مجموع مربعات الانحرافات:

مجموعة (أ) : 9, 4, 1, 0, 1, 4, 9 المجموع = 28

مجموعة (ب): 121,49,36,0,36,49,121 المجموع = 412

$$\text{التباين} = \frac{\text{مجموع مربعات الانحرافات}}{\text{عدد البيانات}}$$

$= \frac{28}{7} = \boxed{4}$ تباين مجموعة (أ)

$= \frac{412}{7} = 58.9$ تباين مجموعة (ب)

S^2 = تباين العينة

σ^2 = تباين المجتمع

- الانحراف المعياري هو الجذر التربيعي للتباين.
- وبالتالي فإن الانحراف المعياري وهو σ لمجموعة أ=

$$\sqrt{4} = 2$$

- ولمجموعة ب = $\sqrt{58.9} = 7.7$

الوسط الحسابي = $\bar{x} = \frac{\sum x_i}{N}$

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

انحرافات القيم عن وسطها الحسابي = $(x - \bar{x})$

$$\sigma = \sqrt{\frac{\sum x_i^2}{N} - \bar{x}_i^2}$$

- إذا كانت كل من $x_1, x_2, x_3, \ldots, x_n$ تحدث بتكرارات $f_1, f_2, f_3, \ldots, f_n$

- فإن الانحراف المعياري يكتب كالتالي:

$$\sigma = \sqrt{\frac{\sum f_i \left(x_i - \bar{x}\right)^2}{N}}$$

$$= \sqrt{\frac{\sum x_i^2 f_i}{N} - \left(\bar{x}_i\right)^2}$$

الخطوات اللازمة لاحتساب الانحراف المعياري (التباين)

أولا: احتساب الوسط الحسابي لمفردات المجموعة $\left(\frac{\sum x}{N}\right)$

ثانيا: إيجاد انحرافات القيم (المفردات) عن الوسط الحسابي $\sum(X_i - \bar{X})$

ثالثا: من أجل التخلص من الإشارات السالبة نربع الانحرافات، ثم نضربها بالتكرارات المناظرة (في حالة القيم المكررة أو التوزيعات التكرارية) إن وجدت.

رابعا: جمع مربعات انحرافات القيم عن الوسط الحسابي.

خامسا: قسمة مجموع مربعات انحرافات القيم (رابعا) على عدد المفردات، نحصل على معدل مربع الانحرافات أو ما يسمى بالتباين (Variance) ويساوي $\frac{\sum (x_i - \bar{x})^2}{N}$ وفي حالة القيم المبوبة يساوي $\frac{\sum f (x_i - \bar{x})^2}{\sum f}$

سادسا: نأخذ الجذر التربيعي (للناتج من خطوة 5)، نحصل على الانحراف المعياري = $\sqrt{\text{التباين}}$

$$\sqrt{\frac{\sum (x_i - \bar{x})^2}{N}} \quad \text{OR} \quad \sqrt{\frac{\sum f (x_i - \bar{x})^2}{\sum f}}$$

مثال: أوجد الانحراف المعياري لأجور العمال في مثال سابق:

فئات الأجور	عدد العمال f_i	مركز الفئات x_i	f_ix_i	x_i^2f	الانحرافات البسيطة $d=x_i-A$	df_i	d^2f_i
-20	9	30	270	8100	-40	-360	14400
-40	22	50	1100	55000	-20	-440	8800
-60	68	70	4760	333200	0	0	0
-80	64	90	5760	518400	20	1280	25600
-100	27	110	2970	326700	40	1680	43200
140-120	10	130	1300	169000	60	60	36000
المجموع	200		16160	1410400		2160	128000

$$\bar{x} = \frac{\sum f_i x_i}{N} = \frac{16160}{200} = 80.8$$

الانحراف المعياري

$$\sigma = \sqrt{\frac{1410400}{200} - (80.8)^2}$$

$$= \sqrt{7052 - 6528.64}$$

$$= \sqrt{523.36} = \boxed{22.87}$$

أما التباين فيساوي 523.36

- ويمكن تبسيط الحل باستخدام الانحرافات البسيطة (بأخذ وسط فرضي)

$$\sigma^2 = \frac{\sum d^2 f_i}{\sum f_i} - \left(\frac{\sum df_i}{\sum f}\right)\left[d = x_i - A, 70 = A\right]$$

$$= \frac{128000}{200} - \left(\frac{2160}{200}\right)^2 = 640 - 116.64$$

$$= 523.36$$

التباين

$$= \sqrt{523.36} = 22.87$$

الانحراف المعياري

خصائص الانحراف المعياري

(1) $$S = \sqrt{\frac{(x_i - a)^2}{n}}$$

a= وسط فرضي. أصغر الانحرافات المعيارية التي يمكن الحصول عليها هو عندما تكون a= $\bar{X}$

(2) في التوزيع الطبيعي نجد أن:

(أ) 68.27% من الحالات تقع بين μ +σ و μ-σ

(ب) 95.45% من الحالات تقع بين μ +2σ و μ -2σ

(ج) 99.73% من الحالات تقع بين μ +3σ و μ -3σ

(3) لمجموعتين N_1 و N_2 (تكراراتهما N_1 و N_2) وتبايناتهما σ_1^2 و σ_2^2 لهما نفس الوسط فإن

التباين المشترك ($\bar{\sigma}^2$) $$\frac{N_1\sigma_1^2 + N_2\sigma_2^2}{N_1 N_2}$$

والانحراف المعياري = جذر المعادلة أعلاه.

(4) لا يتأثر الانحراف المعياري بعمليتي الجمع والطرح، لذلك فإنه إذا زيد أو طرح عدد ثابت من كل مفردة أصلية فإن قيمته تبقى ثابتة. ولكن قيمته تتأثر بعمليتي الضرب والقسمة.

(5) يستخدم في حساب إحصاءات أخرى كالدرجة (القيمة) المعيارية والخطأ المعياري.

التشتت المطلق والتشتت النسبي

- التشتت المطلق: هو التشتت الذي نحصل عليه من الانحراف المعياري وغيره من مقاييس التشتت.
- ولكن تشتت متر واحد من مسافة 1000 متر، يختلف في تأثيره عن تغير 1 متر في مسافة 20 متر.
- ويقاس هذا التأثير بما يعرف بالتشتت النسبي:

$$\text{التشتت النسبي} = \frac{\text{التشتت المطلق}}{\text{المتوسط}} \times 100$$

- ويطلق على التشتت النسبي معامل التشتت $= \frac{a}{\bar{X}} \times 100$

استعمالات الانحراف المعياري

مثال: أجرت شركة تأمين امتحانين للتدقيق في اختبار المتقدمين لها، وكانت نتائج أحد المتقدمين كما يلي:

أحرز 135 نقطة في الامتحان الأول و 265 في الامتحان الثاني.

يبدو منذ اللحظة الأولى أن المتقدم أحرز نتيجة أفضل في المادة الثانية، ولكن عند معرفة الوسط الحسابي والانحراف المعياري للدرجات في كلا الامتحانين يمكن إجراء مقارنة مختلفة.

$\mu_1=100 \qquad \sigma_1=15$

- تحول النقاط إلى قيم معيارية، فيكون قد حصل في الامتحان الأول على

$$= \frac{x - \bar{x}}{\sigma} = \frac{135 - 100}{15} = \boxed{2.5}$$

القيمة المعيارية

- أما القيمة المعيارية للامتحان الثاني فهي

$$= \frac{265 - 250}{30} = \boxed{0.5}$$

- من ذلك نستنتج أن الممتحن قد حصل في الامتحان الأول على نتيجة أفضل من الامتحان الثاني.
- نستدل من التحليل أعلاه أنه بدون الانحراف المعياري والوسط الحسابي لا يمكن أن يكون التحليل دقيقا، وبوجودها تتحول القيم إلى قيم معيارية (Z).

التباين (Variance):

- تباين مجموعة من البيانات يعرف بأنه مربع الانحراف المعياري وبهذا يعرف بـ S^2 .
- يجب التمييز بين الانحراف المعياري للمجتمع والانحراف σ^2, S^2 المعياري لعينة مسحوبة من هذا المجتمع، لذا فإننا نستخدم دائماً الرمز S للأخير والرمز σ للأول. وبهذا فإن يمثلان تباين العينة وتباين المجتمع على الترتيب.
- ملحوظة : يعرف التباين للمجتمع والذي يرمز له بالحرف σ^2 كما يلي:

$$\sigma^2 = \frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}$$

$$f_1, f_2, \ldots, f_k$$

- أما إذا كانت مشاهدات المجتمع تكرر بتكرار فإن

$$\sigma^2 = \frac{\sum_{i=1}^{k} f_i (X_i - \mu)^2}{\Sigma f_i} = \frac{\Sigma f_i (X_i - \mu)^2}{N}$$

- وتكون S^2 هي قيمة تقديرية للتباين σ^2 ومن ثم فإن الانحراف المعياري للمجتمع هو

- طريقة مختصرة لحساب الانحراف المعياري:

المعادلات (3) و (4) يمكن كتابتهما على الترتيب في الصيغ المتكافئة التالية:

$$S = \sqrt{\frac{\sum_{i=1}^{n} X^{2} - \frac{(\sum_{i=1}^{n} X_i)^2}{n}}{n-1}} \quad \text{.........} \ (5)$$

$$S = \sqrt{\frac{\sum_{i=1}^{k} f_i X^{2} - \frac{(\sum_{i=1}^{k} f_i X_i)^2}{n}}{n-1}} \quad \text{.........} \ (6)$$

وإذا كانت $d_i = X_i - A$ هي انحرافات X_i عن ثابت اختياري أو وسط فرضي A فالنتائج (5) و (6) تصبح على الترتيب.

$$S = \sqrt{\frac{\sum_{i=1}^{n} d_i^2 - \frac{(\Sigma d_i)^2}{n}}{n-1}} \qquad \text{.........} (7)$$

$$S = \sqrt{\frac{\sum_{i=1}^{k} f_i d_i^2 - \frac{(\sum_{i=1}^{k} f_i d_i)^2}{n}}{n-1}} \qquad \text{.........} (8)$$

وعندما تجمع البيانات في توزيع تكراري تتساوى فيه أطوال الفئات حيث يساوي طول الفئة C، فإن $X_i = A + Cu_i$ أو $d_i = Cu_i$

وعليه تصبح المعادلة (8) كما يأتي:

$$S = C\sqrt{\frac{\sum_{i=1}^{k} f_i u_i^2 - \frac{(\sum_{i=1}^{k} f_i u)^2}{n}}{n = 1}} = \frac{C\sqrt{\Sigma fu^2 - \frac{(\Sigma fu)^2}{n}}}{n-1} \quad \text{.........} \ (9)$$

وتعطي الصيغة الأخيرة طريقة مختصرة جداً لحساب الانحراف المعياري وتستخدم للبيانات المجمعة إذا كانت أطوال الفئات متساوية ، وهذه تسمى بطريقة الترميز وهي مماثلة بالضبط للطريقة المستخدمة في حساب الوسط الحسابي من البيانات المجمعة في الفصل السابق .

خصائص الانحراف المعياري:

1. يمكن تعريف الانحراف المعياري كالآتي :

$$S=\sqrt{\frac{\sum_{i=1}^{n}(x_i-a)^2}{n-1}} \qquad \text{.........} \ (10)$$

حيث a أي وسط بالإضافة للوسط الحسابي. ومن كل هذه الانحرافات المعيارية، نجد أن أصغرها يمكن الحصول عليه عندما تأخذ $a=\bar{X}$ وذلك حسب الخاصية (ب) من خصائص الوسط الحسابي في الفصل السابق.

2. في التوزيع الطبيعي (وهو أحد التوزيعات الإحصائية المهمة) نجد أن:

أ. 68.27% من الحالات تقع بين $\mu - \sigma$ و $\mu + \sigma$ (أي على بعد انحراف معياري واحد على كل جانب من الوسط الحسابي)

ب. 95.45% من الحالات تقع بين $\mu - 2\sigma$ و $\mu + 2\sigma$ (أي على بعد انحرافين معياريين على كل جانب من الوسط الحسابي) .

جـ. 99.73% من الحالات تقع بين $\mu - 3\sigma$ و $\mu + 3\sigma$