ثانيا: أنواع المقاييس الوصفية

1. مقاييس النزعة المركزية (المتوسطات).
2. مقاييس التشتت.
3. مقاييس الشكل.

1. مقاييس النزعة المركزية (المتوسطات) :

غالبا ما نلاحظ أن البيانات تميل إلى التركيز حول قيمة معينة. وفي هذه الحالة يمكن استخدام هذه "القيمة المركزية" لتمثيل هذه المجموعة من البيانات والمقاييس المستخدمة في التعرف على هذه القيمة المركزية تسمى "مقاييس النزعة المركزية".

ويجب أن يتوفر في هذه المقاييس الصفات الآتية لكي يكون المقياس جيداً:

- أن يعتمد المقياس في حسابه على كل المشاهدات .
- أن يكون المقياس سهل الحساب والفهم .
- أن يتوفر فيه القابلية للتعامل الجبري .
- ألا يتأثر بالقيم الشاذة أو المتطرفة .

ومن بين هذه المقاييس: الوسط الحسابي والوسيط والمنوال .

ولا تتوفر كل الصفات السابق ذكرها في مقياس واحد، ولكن كل مقياس من هذه المقاييس يفضل استخدامه في حالات معينة ولا يفضل استخدامه في حالات /أخرى.

وفيما يلي شرح لكل من هذه المقاييس (المتوسطات) .

أولا- الوسط الحسابي : Arithmetic Mean

تعريفه: يعرف الوسط الحسابي لمجموعة من القيم، بأنه مجموع قيم الظاهرة مقسوما على عدد مفرداتها، ويرمز له بالرمز $(\overline{X})$.

طرق حسابه:

أ) حالة البيانات غير المبوبة: (غير واردة في جدول تكراري) نستخدم القانون الآتي: $\overline{X} = \frac{\sum x}{n}$

حيث :

x: قيم الظاهرة

n: عدد المفردات

$\overline{X}$: الوسط الحسابي للظاهرة X

<u>مثال1</u>: احسب الوسط الحسابي لأوزان خمسة طلاب من البيانات التالية: 50،70،60،80،90

<u>الحل</u>:

x = 50 90 80 60 70

$$\overline{X} = \frac{\sum X}{n} = \frac{50+70+60+80+90}{5} = \frac{350}{5} = 70 \quad Kg$$

ب) حالة البيانات المبوبة : (الجداول التكرارية) نتبع الخطوات الآتية:

- نكون جدول من أربع أعمدة، العمودين الأول والثاني هما المسألة، العمود الثالث هو مراكز الفئات (القيم) ويرمز لها بالرمز (X) ، العمود الرابع هو حاصل ضرب القيم المتناظرة في عمودي 2 ، 3 .

$$\overline{X} = \frac{\sum f\,x}{\sum f}$$

- نطبق القانون الآتي، فنحصل على الوسط الحسابي للظاهرة :

حيث:

x: مراكز الفئات (القيم)

f: تكرار كل فئة

$\sum f$: مجموع التكرارات [n=]

مثال2: احسب الوسط الحسابي للمثال الخاص بتوزيع (50) عداء حسب التوقيت المستقطع في سباق ما بالثواني.

الحل:

فئات الزمن (c)	عدد العدائين (f)	مركز الفئات (x = القيم)	2*3 f x
10-20	3	10+20/2 = 15	45
20-30	6	25	150
30-40	10	35	350
40-50	15	45	675
50-60	8	55	440
60-70	5	65	325
70-80	3	75	225
$\sum$	50	-	2210

$$\overline{X} = \frac{\sum f\, x}{\sum f} = \frac{2210}{50} = 44.2$$ ث

الوسط الحسابي المرجح: la moyenne arithmétique pondéré

في بعض الأحيان يأخذ المتغير قيما تختلف من حيث أهميتها أو وزنها.

فإذا كانت: قيم المتغيرX: X1,X2,……..

الأهمية أو الوزنW: W1,W2,…….

فإن هذه الأوزان يجب أن تؤخذ في الاعتبار عند حساب الوسط الحسابي ، ويكون: $\overline{X}_w = \frac{\sum XW}{\sum W}$

مثل المعدل الفصلي للطالب .

مثال3: حصل طالب على الدرجات الآتية في أحد الفصول الدراسية:

لغة إنجليزية	كرة قدم	منهجية	سباحة	إحصاء	المقــرر	-
70	**87**	**65**	**80**	**75**	**الدرجـــة**	**(X)**
3	**2**	**4**	**3**	**4**	**المعامل**	**(w)**

أحسب المعدل الفصلي لهذا الطالب باستخدام الوسط الحسابي المرجح

الحل:

$$\overline{X}_w = \frac{\sum XW}{\sum W}$$

$$= \frac{(75x4)+(80x3)+(65x4)+(87x2)+(70x3)}{4+3+4+2+3}$$

$$= \frac{1184}{16} = 74 \qquad \text{ةجرد}$$

الوسط الحسابي العام :

يستخدم في حسابه القانون الآتي :
$$\overline{\overline{X}} = \frac{n_1 \overline{X}_1 + n_2 \overline{X}_2 + n_3 \overline{X}_3 + ...}{n_1 + n_2 + n_3 + ...}$$

مثال: إذا كان الوسط الحسابي لأطوال (25) طالب من الفوج الأول هو(175) سم ،الوسط الحسابي لأطوال(35) طالب من الفوج الثاني هو (168) سم .
احسب الوسط الحسابي العام لجميع الطلبة.

الحل:
$$\overline{\overline{X}} = \frac{n_1 \overline{X}_1 + n_2 \overline{X}_2}{n_1 + n_2} = \frac{(25x175) + (35x168)}{25 + 35}$$

$$= \frac{10255}{60} = 170,92 \ \ c.m$$

حساب المتوسط الحسابي بالطريقة المختصرة:

عندما تكون البيانات الموجودة في الجداول المعطاة كبيرة القيم فإن استخدام الطريقة السابقة يصبح صعبا، كما يزداد احتمال الوقوع في الأخطاء، لذلك فإنه في مثل هذه الحالات يفضل استخدام طريقة مختصرة الهدف منها تبسيط العمليات الحسابية الطويلة حتى يسهل التعامل معها.
فإذا قمنا مثلا بطرح قيمة ثابتة (a) من جميع القيم (جميع مراكز الفئات) فإن المتوسط الحساب يصبح

إذا كانت البيانات مفردة:
$$\boxed{\overline{X} = a + \frac{\sum_{i=1}^{n}(x_1 - a)}{n}}$$

أو إذا كانت البيانات مبوبة في جداول توزيع تكراري:
$$\boxed{\overline{X} = a + \frac{\sum_{i=1}^{n} f_1(x_i - a)}{\sum f_i}}$$

خواص المتوسط الحسابي:

1 – يعتبر المتوسط الحسابي أبسط مقاييس النزعة المركزية حسابا وأكثرها استخداما.
2 – يأخذ المتوسط الحسابي بعين الاعتبار جميع قيم الظاهرة المدروسة.
3 – مجموع انحرافات القيم عن متوسطها الحسابي يساوي صفر= 0

$$\sum_{i=1}^{n}(x_i - \overline{X}) = 0$$

4 – يتأثر المتوسط الحسابي بالقيم المتطرفة أو الشاذة

ثانيا- الوسيط : Median

أ- تعريفه: هو القيمة التي تقسم مجموعة القيم بعد ترتيبها تصاعديا (أو تنازليا) إلى قسمين متساويين، بحيث يكون عدد القيم الأصغر منها مساويا لعدد القيم الأكبر منها، ويرمز له بالرمز(m) .

ب- طرق حسابه:

ب-1) حالة البيانات غير المبوبة: يتم إيجاد قيمة الوسيط من التعريف مباشرة .

مثال4: إذا كان لدينا القيم الآتية: 50 60 80 70 90

- أحسب قيمة الوسيط

<u>الحل:</u> لحساب قيمة الوسيط، نرتب القيم تصاعديا فتصبح: 50 60 70 80 90

من التعريف، نجد أن قيمة الوسيط $m = 70$

وإذا كان لدينا القيم الآتية: 50 60 80 70 90 100

لحساب قيمة الوسيط، نرتب القيم تصاعديا فتصبح: 50 60 70 80 90 100

من التعريف، نجد أن قيمة الوسيط

$$m = \frac{70+80}{2} = 75$$

ب-2 حالة البيانات المبوبة: يتم إيجاد قيمة الوسيط بالحساب وبالرسم.

أ) بالحساب: نتبع الخطوات الآتية:

- نكون جدول متجمع صاعد (نازل) كالمعتاد.
- نوجد ترتيب الوسيط من العلاقة الآتية: $c_1 = \frac{\sum f}{2}$
- نوجد قيمة الوسيط من العلاقة الآتية: $m = L + \frac{c_1 - c_2}{c_3}.h$

حيث:

L: بداية (الحد الأدنى) لفئة الوسيط .

C_1: ترتيب الوسيط $c_1 = \frac{\sum f}{2}$

C_2: ك.م.ص السابق لفئة الوسيط

C_3: التكرار الأصلي لفئة الوسيط

h: طول فئة الوسيط

وهذا القانون يستخدم في حالة الجدول الصاعد والذي سنكتفي به في حالة الحساب.

ويلاحظ أن قيمة الوسيط لا بد وأن تقع داخل حدود فئة الوسيط، أي لا تقل عن بداية فئة الوسيط ولا تزيد عن نهايتها.

<u>مثال5</u>: أوجد قيمة الوسيط للمثال الخاص بتوزيع (50) عداء حسب التوقيت المستقطع في سباق ما بالثواني.

<u>الحل:</u>

ك . م . ص	اقل من الحد الأعلى للفئة	عدد العدائين (f)	فئات الزمن (c)
3	أقل من 20	3	10-20
9	أقل من 30	6	20-30
19	أقل من 40	10	30-40
34	أقل من 50	15	40-50 (فئة الوسيط)
42	أقل من 60	8	50-60
47	أقل من 70	5	60-70
50	أقل من 80	3	70-80
-	-	50	Σ

- **ترتيب الوسيط**

$$C_1 = \frac{\sum f}{2} = \frac{50}{2} = 25$$

$$m = L + \frac{c_1 - c_2}{c_3}.h \quad = 40 + \frac{25-19}{15} x10$$

$$m = 40 + 4 = 44$$ ث

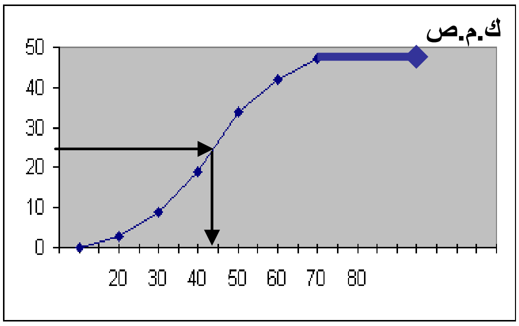
- **قيمة الوسيط**

ب) بالرسم: نتبع الخطوات الآتية:

- نرسم منحنى متجمع صاعد (أو نازل) كالمعتاد.
- نحدد ترتيب الوسيط $\left(c_1 = \frac{\sum f}{2} \right)$ على المحور الرأسي.
- نوجد قيمة الوسيط على المحور الأفقي.

مثال: أوجد قيمة الوسيط بالرسم من المنحنى المتجمع الصاعد للمثال السابق

فئة الوسيط



ك.م.ص

قيمة الوسيط : m=44 s

خواص الوسيط:

يتصف الوسيط بعدة خصائص أهمها:

1 – لا يتأثر بالقيم المتطرفة وبالتالي فإنه يعتبر أصلح المقاييس عن وجود مثل هذه القيم.
2 – يمكن إيجاد الوسيط من الرسم.
3 – يمكن استخدامه في حالة البيانات النوعية.
4 - يمكن حسابه من الجداول التكرارية المفتوحة.

ثالثا- المنوال : Mode

أ- تعريفه : هو القيمة التي تتكرر أكثر من غيرها، أو هو القيمة الأكثر شيوعا أو تكراراً، ويرمز له بالرمز D .

ب- طرق حسابه :

ب-1)حالة البيانات غير المبوبة: يتم إيجاد قيمة المنوال من التعريف مباشرة

فمثلاً : في القيم 7 8 7 6 5
نجد أن قيمة المنوال D = 7
وفي القيم 3 2 5 3 4 3 2
نجد أن قيمة المنوال D = 3

ب-2) حالة البيانات المبوبة: يتم إيجاد قيمة المنوال بالحساب وبالرسم .

أ) بالحساب: نستخدم القانون الآتي :

$$D = L + \frac{\Delta_1}{\Delta_1 + \Delta_2}.h$$

حيث :
L: الحد الأدنى (بداية) لفئة المنوال.
$\Delta 1$: الفرق بين أكبر تكرار والسابق له .
$\Delta 2$: الفرق بين أكبر تكرار واللاحق له .
h: طول فئة المنوال .
وتعرف فئة المنوال، بأنها الفئة التي تقابل أكبر تكرار.

مثال6: أوجد قيمة المنوال للمثال الخاص بتوزيع (50) عداء حسب التوقيت المستقطع في سباق ما بالثواني

الحل:

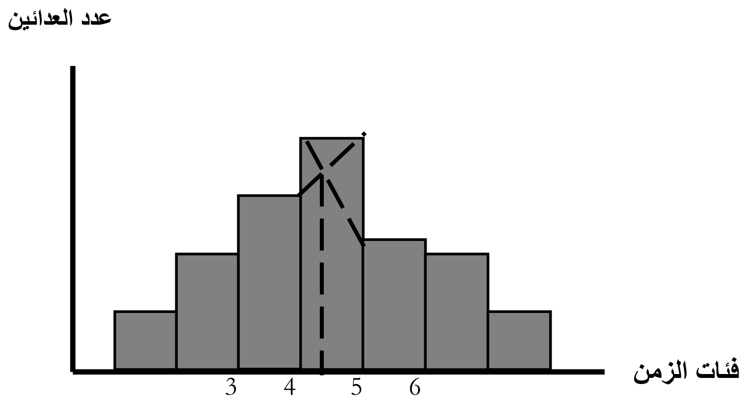
$$D = L + \frac{\Delta_1}{\Delta_1 + \Delta_2}.h \;= 40 + \frac{5}{5+7}x10$$

$$= 40 + 4.17 = 44.17 \qquad c.m$$

ب) بالرسم : يتم إيجاد قيمة المنوال بالرسم من المدرج التكراري، ويمكن الاكتفاء برسم ثلاث مستطيلات، مستطيل يمثل فئة المنوال (الفئة التي تقابل أكبر تكرار) ومستطيل يمثل الفئة السابقة وآخر يمثل الفئة اللاحقة لها، ثم نصل رؤوس المستطيلات ببعضها فتتقابل في نقطة نسقط منها عمود على المحور الأفقي، فتكون هي قيمـة المنوال.

مثال: أجد قيمة المنوال بالرسم للمثال السابق.

الحل:



D= 44 cm

خواص المنوال:

1 – لا يأخذ بعين الاعتبار جميع البيانات المعطاة وبالتالي فهو لا يتأثر بالقيم المتطرفة.

2 – يمكن حسابه بيانيا.

3 – يمكن أن يوجد أكثر من منوال لتوزيع واحد.

4 – يمكن حسابه من الجداول الإحصائية المفتوحة.

5 – يعتبر أفضل المتوسطات لوصف الظواهر النوعية.

مقاييس التشتت:

عرفنا فيما سبق أن مقاييس النزعة المركزية (من متوسط ووسيط ومنوال) تسمح لنا بالحصول على القيم المتوسطة للبيانات أو على تجمعها، غير أن هذه المقاييس لا تكفي لوحدها لمعرفة الصفات الإحصائية اللازمة لوصف الظواهر، لأن الفروق بين قيم الظواهر قد تزداد أو تنقص رغم تساوي المتوسطات لهذه الظواهر، ولتوضيح ما سبق نفترض أن طالبين تحصلا على النتائج التالية في خمس مواد دراسية:

الطالب (X): 10،11،13،14،15.

الطالب (Y):8،9،13،15،18.

فمتوسط درجات الطالب (X) يساوي 12,6 وكذلك متوسط درجات الطالب (Y) يساوي 12,6 ووسيط درجات الطالب (X) يساوي 13 وكذلك وسيط درجات الطالب (Y) يساوي 13.

قد يفهم مما سبق أن الطالبين (X) و (Y) لهما نفس المستوى غير أن التمعن الجيد في الدرجات التي تحصل عليها الطالبين تبين أن الطالب (X) ناجح في كل المواد المدروسة في حين أن الطالب (Y) ناجح في ثلاث مواد فقط.

إن هذه الحقيقية تبين أن مقاييس النزعة المركزية لا تعطي فكرة وافية عن اختلاف قيم الظواهر، ولا تحقق كل الأغراض التي نرغب الوصول إليها من دراستنا لذلك فإن مقاييس النزعة المركزية لا بد أن تكون مصحوبة بمقاييس أخرى لقياس مدى تباعد أو تقارب البيانات من بعضها البعض أو من متوسطها، تسمى هذه المقاييس بمقاييس التشتت.

<u>ما معنى التشتت؟</u>

تشتت بيانات ظاهرة ما يقصد به درجة أو مقدار التفاوت أو الاختلاف بين مفردات هذه الظاهرة، وتعتبر بيانات الظاهرة متجانسة عندما تكون قيمتها قريبة من بعضها البعض ونقول في هذه الحالة أنها غير مشتتة، أما إذا كانت بيانات الظاهرة متباعدة وغير متجانسة فنقول أن مفردات الظاهرة مشتتة وغير مركزة، ويقاس تشتت البيانات بعدة مقاييس منها:

أولا – المدى (المطلق)

المدى لمجموعة من البيانات هو الفرق بين أكبر قيمة وأصغر قيمة لها ويرمز له بالرمز R

المدى= أكبر قيمة – أصغر قيمة.

أما المدى للتوزيعات التكرارية فيحسب بعدة طرق منها:

المدى = مركز الفئة الأخيرة – مركز الفئة الأولى

المدى = الحد الأعلى للفئة الأخيرة – الحد الأدنى للفئة الأولى

<u>مثال1</u>: أوجد المدى للبيانات التالية: 12،18،22،28،30.

<u>الحل</u>: المدى = أكبر قيمة – أصغر قيمة

= 30-12=22.

<u>مثال2</u>: أوجد المدى للبيانات التالية 65،20،17، -04،18،19.4

<u>الحل</u>: المدى = 65 –(-4) = 69.

نلاحظ المدى في هذا المثال قد تأثر بشكل كبير جدا بالقيم المتطرفة، إذ نلاحظ أن معظم البيانات متقاربة باستثناء القيمة 65 والقيمة (-4)، فإذا استبعدنا هذه القيم المتطرفة فإن المدى يصبح

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

R = 20-14=6

وبسبب هذا العيب فإن المدى كمقياس للتشتت لا يستخدم إلا عندما نرغب في مقياس تقريبي وسريع لتشتت البيانات دون الاهتمام بالدقة في المقياس، أو عندما يكون للبيانات المتطرفة أهمية خاصة كتوزيعات درجات الحرارة على سبيل المثال، حيث تعلن درجات الحرارة اليومية بحدها الأقصى وحدها الأدنى خلال اليوم.

خواص المدى:

1 – يتصف المدى بسهولة حسابه.

2 – يعتمد في حسابه على قيمتين فقط هما القيمة الكبرى والقيمة الصغرى.

3- بسبب الخاصية الثانية فإن المدى شديد التأثر بالقيم المتطرفة.

ثانيا: الانحراف المتوسط L'écart moyen:

لأن مقاييس التشتت هي مقاييس لقوة تجمع البيانات حول بعضها، وحيث أن التجمع يكون حول القيم المتوسطة، فإنه إذا كان مقدار الاختلاف بين القيم ومتوسطها كبيرا دل ذلك على أن التشتت كبير والعكس صحيح.

وحيث أن مجموع الاختلافات (الانحرافات) عن المتوسط يساوي صفرا، فإنه لو حسبنا القيم المطلقة لمقدار الاختلاف عن المتوسط يكون متوسط هذه الاختلافات مقياسا مناسبا لمقدار التشتت، يسمى هذا المقياس بالانحراف المتوسط.

ويعرف الانحراف المتوسط بأنه المتوسط الحسابي للقيم المطلقة لانحرافات القيم عن متوسطها الحسابي، وسوف نرمز للانحراف المتوسط في دراستنا بالرمز E_x وعليه:

إذا كانت لدينا القيم التالية: $X_1 , X_2 , X_3 , X_4 \dots X_n$ فإن الانحراف المتوسط لها هو:

$$E_X = \frac{|X_1 - \overline{X}| + |X_2 - \overline{X}| + |X_3 - \overline{X}| + \dots + |X_n - \overline{X}|}{n}$$

أو

$$E_X = \frac{\sum_{i=1}^{n} |X_i - \overline{X}|}{n}$$

أما إذا كانت البيانات مكررة أو مبوبة في جداول توزيع تكراري فإن الانحراف المتوسط لها يعطي بالعلاقة:

$$E_X = \frac{\sum_{i=1}^{n} fi|X_i - \overline{X}|}{\sum_{i=1}^{n} fi}$$

<u>مثال3</u>: أوجد الانحراف المتوسط للبيانات التالية: 2،4،5،6،8

<u>الحل</u>:

$\left\|X_I - \overline{X}\right\|$	X_i
3	2
1	4
0	5
1	6
3	8
8	المجموع

$$\overline{X} = \frac{\sum x_i}{n} = \frac{25}{5} = 5$$

$$E_x = \frac{\sum\left|X_i - \overline{X}\right|}{n} = \frac{8}{5} = 1.6$$

مثال4: أوجد تشتت البيانات المبوبة في جدول التوزيع التكراري الآتي باستخدام الانحراف المتوسط

الفئة	2-0	4-2	6-4	8-6	المجموع
التكرار	2	3	4	3	12

الحل:

الفئة	التكرار f_i	X_i	$X_i.f_i$	$\left\|X_I - \overline{X}\right\|$	$f_i\left\|X_i - \overline{X}\right\|$
0-2	2	1	2	3.33	6.66
4-2	3	3	9	1.33	4
6-4	4	5	20	0.67	2.68
8-6	3	7	21	2.67	8.01
المجموع	12		52		21.35

$$\overline{X} = \frac{\sum f_i X_i}{\sum f_i} = \frac{52}{12} = 4.33$$

$$E_x = \frac{\sum f_i\left|X_i - \overline{X}\right|}{\sum f_i} = \frac{21.35}{12} = 1.78$$

ويعتبر الانحراف المتوسط أفضل من سابقه (المدى) لأنه أقل تأثر ا بالقيم المتطرفة غير أنه لا يستعمل بشكل واسع بسبب اعتماده على القيمة المطلقة لانحرافات القيم عن متوسطها الحسابي.

خواص الانحراف المتوسط:

1 - يعتمد في حسابه على جميع القيم وليس على القيمة الكبرى والصغرى فقط.

2 - لا يمكن حسابه في حالة التوزيعات التكرارية المفتوحة.

3 - يتأثر بالقيم المتطرفة، لأن انحرافها عن المتوسط الحسابي يكون كبيرا.

ثالثا - التباين والانحراف المعياري La variance et l’écart type:

أ – التباين La variance:

وهو عبارة عن المتوسط الحسابي لمربعات الفروق بين قيم المتغير الإحصائي ومتوسطها الحسابي، ونستخدم مربعات الفروق هنا تفاديا لاستخدام القيم المطلقة كما هو الشأن في الانحراف المتوسط، فإذا كانت لدينا البيانات التالية: $X_1,X_2,X_3,X_4 \ldots\ldots X_n$

فإن التباين لهذه البيانات يعطي بالعلاقة:

$$V_X = \frac{(X_1-\overline{X})^2+(X_2-\overline{X})^2+(X_3-\overline{X})^2+\ldots\ldots+(X_n-\overline{X})^2}{n}$$

$$V_X = \frac{\sum_{i=1}^{n}(X_1-\overline{X})^2}{n}$$

في حالة البيانات المبوبة:

$$V_X = \frac{\sum_{i=1}^{n} fi(X_1-\overline{X})^2}{\sum fi}$$

مثال5: أوجد التباين للبيانات التالية: 3،7،6،1،11،5،6،9.

الحل:

X_i	$(X_i-\overline{X})$	$(X_i-\overline{X})^2$
3	**-3**	**9**
7	**1**	**1**
6	**0**	**0**
1	**-5**	**25**
11	**5**	**25**
5	**-1**	**1**
6	**0**	**0**
9	**3**	**9**
48		**70**

المتوسط الحسابي:

$$\overline{X} = \frac{\sum Xi}{n} = \frac{48}{8} = 6$$

التباين:

$$V_x = \frac{\sum(X_i-\overline{X})^2}{n} = \frac{70}{8} = 8.75$$

في بعض الأحيان عندما يكون المتوسط الحسابي للبيانات عبارة عن كسر، فإن عملية حساب التباين تكون مضنية وعرضة للأخطاء الحسابية لذلك فإنه تم تطوير طريقة مختصرة لحساب التباين.

الصيغة المختصرة لحساب التباين:

$$\boxed{V_x = \frac{\sum X_i^2}{n} - \overline{X}^2}$$

أما في حالة البيانات المبوبة فإن العلاقة تصبح:

$$V_x = \frac{\sum f_i X_i^2}{\sum fi} - \overline{X}^2$$

ب – الانحراف المعياري:

ويعتبر الانحراف المعياري من أهم المقاييس الإحصائية للتشتت، وهو أكثرها استخداما في النظريات والقوانين الإحصائية، لأنه يعطي فكرة سليمة ومنطقية عن ظاهرة التشتت، ويعرف الانحراف المعياري بأنه الجذر التربيعي لمتوسط مجموع مربع انحراف القيم عن متوسطها، أي أنه الجذر التربيعي للتباين.

سوف نرمز للانحراف المعياري في دراستنا بالرمز (S_x).

الانحراف المعياري لبيانات مفردة:

$$Sx = \sqrt{\frac{\sum (X_i - \overline{X})^2}{n}}$$

الانحراف المعياري لبيانات متكررة أو مبوبة:

$$Sx = \sqrt{\frac{\sum fi(X_i - \overline{X})^2}{\sum fi}}$$

أما الصيغة المختصرة للانحراف المعياري فتعطى بالعلاقات التالية:

الانحراف المعياري لبيانات مفردة:

$$Sx = \sqrt{\frac{\sum X_i^2}{n} - \overline{X}^2}$$

الانحراف المعياري لبيانات متكررة أو مبوبة:

$$Sx = \sqrt{\frac{\sum fiX_i^2}{\sum fi} - \overline{X}^2}$$

مثال6:

أوجد التباين والانحراف المعياري بالصيغة الأصلية ثم بالصيغة المختصرة للبيانات المبوبة في الجدول الإحصائي التالي:

الفئة	8-4	13-9	18-14	23-19	28-24	المجموع
التكرار	3	4	6	2	4	19

الحل:

الفئة	التكرارf_i	X_i	$f_i x_i$
8-4	3	6	18
13-9	4	11	44
18-14	6	16	96
23-19	2	21	42
28-24	4	26	104
المجموع	19		304

$$\overline{X} = \frac{\sum f_i X_i}{\sum f_i} = \frac{304}{19} = 16$$

$f_i(X_i - \overline{X})^2$	$f_i(X_i - \overline{X})$	$(X_i - \overline{X})$
300	-30	-10
100	-20	-5
0	0	0
50	10	5
400	40	10
850		

$fi(Xi)^2$	X_i^2
108	36
484	121
1536	256
882	441
2704	676
5714	

التباين بالصيغة الأصلية:

$$V_X = \frac{\sum f_i (X_i - \overline{X})^2}{\sum f_i} = \frac{850}{19} = 44.74$$

الانحراف المعياري بالصيغة الأصلية:

$$S_x = \sqrt{\frac{\sum f_i (X_i - \overline{X})^2}{\sum f_i}} = \sqrt{44.74} = 6.69$$

التباين بالصيغة المختصرة:

$$V_x = \frac{\sum f_i X_i^2}{\sum f_i} - \overline{X}^2$$

$$= \frac{5714}{19} - (16)^2 = 300.74 - 256 = 44.74$$

الانحراف المعياري بالصيغة المختصرة

$$S_x = \sqrt{\frac{\sum f_i X_i^2}{\sum f_i} - \overline{X}^2}$$

$$S_x = \sqrt{\frac{5714}{19} - 256} = \sqrt{44.73} = 6.69$$

خصائص الانحراف المعياري:

1 -إذا كان الانحراف المعياري للقيم $X_1, X_2, X_3, X_4 \ldots\ldots X_n$ هو S_x فإنه إذا أضيفت أو طرحت قيمة ثابتة (a) إلى أو من جميع القيم فإن الانحراف المعياري لا يتغير

2- إذا كان الانحراف المعياري للقيم $X_1, X_2, X_3, X_4 \ldots\ldots X_n$ هو S_x فإنه إذا ضربت كل قيمة بالمقدار a (قسمت كل قيمة على المقدار a) فإن الانحراف المعياري يتأثر بالمقدار نفسه.

3- بالنسبة للتوزيع الطبيعي فإن:

a. 68.27% من البيانات تقع في المجال $\overline{X} \pm Sx$

b. 95.45% من البيانات تقع في المجال $\overline{X} \pm 2Sx$

c. 99.73% من البيانات تقع في المجال $\overline{X} \pm 3Sx$

4 -يأخذ الانحراف المعياري نفس وحدة القياس للمتغير الأصلي (كلغ، متر، لتر ....) لذلك لا يمكن استخدامها كأساس للمقارنة بين تشتت توزيعين لهما وحدات قياس مختلفة.

5 -بما أن الانحراف المعياري يتأثر بالمتوسط الحسابي لبيانات الظاهرة فإنه لا يمكن استخدامه للمقارنة بين تشتت بيانات توزيعين لهما متوسط حسابي مختلف ولو كان هذين التوزيعين من نفس النوعية

6 -لا يمكن إيجاده بالنسبة للتوزيعات التكرارية المفتوحة من البداية أو النهاية.

رابعا - معامل الاختلاف Coefficient de variation:

رأينا في الصفحات السابقة أن الانحراف المعياري هو مقياس واقعي ومؤشر صحيح عن مقدار التشتت غير أن الخاصيتين 4و5 السابقتين تبينان أنه إذا استخدمنا هذا المقياس للمقارنة بين تشتت ظاهرتين أو أكثر فإن المقارنة تكون واقعية وواقعية فقط إذا كانت الظواهر من نوعية واحدة ولها متوسطات متساوية ، أي يمكن مقارنة تشتت درجات مادة ما بدرجات مادة أخرى أو مقارنة تشتت دخل مجموعة من العمال بدخل مجموعة أخرى، وتكون المقارنة أكثر واقعية إذا كانت المتوسطات متساوية أو قريبة من بعضها. أما إذا كانت الظواهر من صفات مختلفة أو إذا كانت متوسطاتها متباعدة، فإن المقارنة اعتمادا على الانحراف المعياري ستكون غير منطقية وغير واقعية، ولهذا السبب وجدت مقاييس أخرى سميت مقاييس

التشتت النسبي، وتعتمد هذه المقاييس على تمييز البيانات وتقيس التشتت كنسبة مئوية للمتوسط، ومن أهم هذه المقاييس نجد معامل الاختلاف.

معامل الاختلاف = $\frac{\text{الانحراف المعياري}}{\text{المتوسط الحسابي}} \times 100$

$$CV = \frac{Sx}{\overline{X}} \times 100$$

مثال7:

إذا كان متوسط درجات مجموعة من الطلبة في مادة ما هو 15 بانحراف معياري 3 ومتوسط درجاتهم في مادة أخرى هو 8 بانحراف معياري 2، فأي الدرجات في نظرك أكثر تشتتا؟

الحل:

إذا اعتمدنا على الانحراف المعياري فإننا نحكم على أن درجات المادة الأولى أكثر تشتتا (Sx =3) من درجات المادة الثانية (Sx =2)، وهذا غير صحيح لأننا إذا أدخلنا المتوسط الحسابي لدرجات الطلبة في المادتين في الحسبان سنحصل على النتائج التالية:

$$CV_1 = \frac{3}{15} \times 100 = 20\%$$

$$CV_2 = \frac{2}{8} \times 100 = 25\%$$

أي أن درجات المادة الثانية أكثر تشتتا

مقاييس الشكل:

إذا كانت مقاييس النزعة المركزية ومقاييس التشتت تسمح بتلخيص بيانات أي ظاهرة في صورة أرقام بسيطة تعطي فكرة عن خصائص توزيع هذه البيانات ودرجة تجانسها أو اختلافها فإن هذا الوصف الذي تعطيه يبقى تنقصه الدقة الكافية المطلوبة للتعرف على خواص التوزيع خاصة من حيث انتشار البيانات على المنحنى البياني الممثل لها من حيث التوائه أو تفلطحه عن الوضع الطبيعي، كذلك دعت الحاجة لاستخدام مقاييس أخرى لتحقيق هذا الغرض سميت هذه المقاييس بمقاييس الالتواء والتفلطح التي سنتعرف عليها بعد التطرق لموضوع العزوم.

أولا: العزوم: les Moments

العزوم قد تكون حول نقطة الأصل أو حول المتوسط الحسابي أو حول أي نقطة معينة، فالعزم الأول حول نقطة الأصل مثلا هو متوسط قيم الظاهرة، والعزم حول المتوسط الحسابي هو متوسط انحرافات قيم التوزيع عن المتوسط الحسابي لها. أما رتبة العزم فتتحدد بدرجة القوة (الأس) التي ترفع إليها القيم أو انحرافاتها عن المتوسط الحسابي، وتستخدم العزوم في إيجاد المعامل العزمي للإلتواء وكذلك معامل التفلطح.

I- العزوم حول نقطة الأصل:

إذا كانت لدينا القيم x1, x2, x3…………xn فإن

العزم الأول = $\frac{\sum X_i}{n}$

العزم الثاني = $\frac{\sum X_i^2}{n}$

العزم الثالث = $\frac{\sum X_i^3}{n}$

العزم النوني = $\frac{\sum X_i^n}{n}$

مثال1:

إذا كانت لدينا القيم 1،2،5،8 أوجد العزم الأول والثاني والثالث والرابع حول نقطة الأصل

الحل:

العزم الأول: $\frac{\sum X_i}{n} = \frac{8+5+2+1}{4} = \frac{16}{4} = 4$

العزم الثاني: $\frac{\sum X_i^2}{n} = \frac{64+25+4+1}{4} = \frac{94}{4} = 23.5$

العزم الثالث: $\frac{\sum X_i^3}{n} = \frac{512+125+8+1}{4} = 161.5$

العزم الرابع : $\frac{\sum X_i^4}{n} = \frac{4096+625+16+1}{4} = 1184.5$

نلاحظ أن: العزم الأول = المتوسط الحسابي ، والتباين = العزم الثاني – مربع العزم الأول ومنه فإن التباين في المثال السابق = 23.5 – 16 = 7.5.

أما إذا كانت البيانات متكررة أو مبوبة في فئات فإن:

العزم الأول = $\frac{\sum f_i X_i}{\sum f_i}$

العزم الثاني = $\frac{\sum f_i X_i^2}{\sum f_i}$

العزم الثالث = $\frac{\sum f_i X_i^3}{\sum f_i}$

العزم النوني = $\frac{\sum f_i X_i^n}{\sum f_i}$

حيث: fi = التكرار.

$\sum f_i$ = مجموع التكرارت.

Xi = القيم أو مراكز الفئات.

مثال2: أوجد العزم الأول والثاني والثالث للبيانات المبوبة في جدول التوزيع التالي، ثم أوجد التباين والانحراف المعياري

الفئة	**1 – 3**	**3 – 5**	**5 – 7**	**7 – 9**	**9 – 11**	**المجموع**
التكرار	**1**	**2**	**3**	**4**	**6**	**16**

الحل:

$f_ix_i^3$	$f_ix_i^2$	f_ix_i	مركز الفئة (x_i)	التكرار (f_i)	الفئة
8	**4**	**2**	**2**	**1**	**1-3**
128	**32**	**8**	**4**	**2**	**3-5**
648	**108**	**18**	**6**	**3**	**5-7**
2048	**256**	**32**	**8**	**4**	**7-9**
6000	**600**	**60**	**10**	**5**	**9-11**
8832	**1000**	**120**		**16**	**المجموع**

العزم الأول = $$M_1 = \frac{\sum f_i X_i}{\sum f_i} = \frac{120}{16} = 7.5 = \overline{X}$$

العزم الثاني = $$M_2 = \frac{\sum f_i X_i^2}{\sum f_i} = \frac{1000}{16} = 62.5$$

العزم الثالث = $$M_3 = \frac{\sum f_i X_i^3}{\sum f_i} = \frac{8832}{16} = 552$$

التباين = العزم الثاني – مربع العزم الأول.

= 62.5 – 56.25 = 6.25.

ومنه الانحراف المعياري = $\sqrt{\text{التباين}}$ = $\sqrt{6.25}$ = 2.5

II – العزوم حول المتوسط الحسابي:

ونرمز للعزم حول المتوسط الحسابي بالرمز µ.

العزم الأول حول المتوسط الحسابي: $$\mu_1 = \frac{\sum \left(X_i - \overline{x}\right)}{n}$$

العزم الثاني حول المتوسط الحسابي: $$\mu_2 = \frac{\sum \left(X_i - \overline{x}\right)^2}{n}$$

العزم الثالث حول المتوسط الحسابي: $$\mu_3 = \frac{\sum \left(Xi - \overline{x}\right)^3}{n}$$

العزم النوني حول المتوسط الحسابي: $\mu_n = \frac{\sum (X_i - \bar{x})^n}{n}$

مثال3:

أوجد العزم الأول والثاني والثالث حول المتوسط الحسابي للقيم التالية 2، 4، 6

الحل: $\bar{X} = \frac{\sum x_i}{n} = \frac{2+4+6}{3} = \frac{12}{3} = 4$

X1	$(X_1 - \bar{X})$	$(X_1 - \bar{X})^2$	$(X_1 - \bar{X})^3$
2 **4** **6**	**2-** **0** **2**	**4** **0** **4**	**-8** **0** **8**
المجموع	**0**	**8**	**0**

$$\mu_1 = \frac{\sum (X_i - \bar{x})}{n} = \frac{0}{3} = 0$$

$$\mu_2 = \frac{\sum (X_i - \bar{x})^2}{n} = \frac{8}{3} = 2.66$$

$$\mu_2 = \frac{\sum (X_i - \bar{x})^3}{n} = \frac{0}{3} = 0$$

ويلاحظ من خلال هذه النتائج أن:

- العزم الأول حول المتوسط الحسابي يساوي صفر دائما.
- العزم الثاني حول المتوسط الحسابي يساوي التباين

أما إذا كانت البيانات متكررة أو مبوبة في جداول توزيع تكراري فإن العزوم حول المتوسط الحسابي يمكن حسابها على النحو التالي:

العزم الأول حول المتوسط الحسابي = $\mu_1 = \frac{\sum f_i (X_i - \bar{x})}{\sum f_i}$

العزم الثاني حول المتوسط الحسابي: $\mu_2 = \frac{\sum f_i (X_i - \bar{x})^2}{\sum f_i}$

العزم الثالث حول المتوسط الحسابي: $\mu_3 = \frac{\sum f_i (X_i - \bar{x})^3}{\sum f_i}$

العزم النوني حول المتوسط الحسابي: $\mu_n = \frac{\sum f_i (X_i - \bar{x})^n}{\sum f_i}$

مثال4: أوجد العزم الأول والثاني حول المتوسط الحسابي للبيانات المبوبة في الجدول التالي:

الفئة	0 – 2	2 – 4	4 – 6	8 – 6	المجموع
التكرار	2	4	6	4	16

الحل:

الفئة	f_i	X_i	fixi	$(x_i - \bar{x})$	$f_i(x_i - \bar{x})$	$f_i(x_i - \bar{x})^2$
0 – 2	2	1	2	-3.5	-7	24.5
2 – 4	4	3	12	-1.5	-6	9
4 – 6	6	5	30	0.5	3	1.5
6 – 8	4	7	28	2.5	10	25
المجموع	16		72		0	60

العزم الأول حول المتوسط الحسابي : $\mu_1 = \frac{\sum f_i (X_i - \bar{x})}{\sum f_i} = \frac{0}{4}$

العزم الثاني حول المتوسط الحسابي: $\mu_2 = \frac{\sum f_i (X_i - \bar{x})^2}{\sum f_i} = \frac{60}{4}$

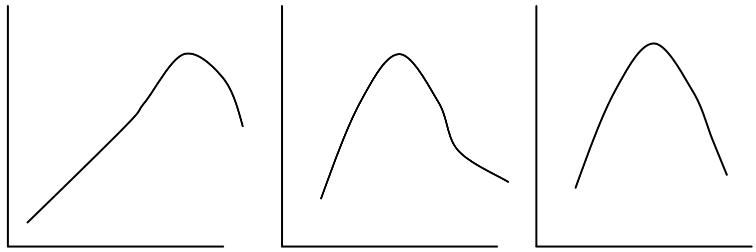
ثانيا: تحديد شكل التوزيع:

يمكن تحديد شكل التوزيع باستخدام مقاييس الالتواء والتفلطح.

1 – الالتواء Asymétrie: (عدم التناظر من اليمين أو من اليسار مقارنة بتوزيع متناظر بالنسبة للقيمة المركزية) يعتبر منحنى التوزيع التكراري المعتدل هاما جدا في الدراسات والتحليلات الإحصائية.

إن هذا المنحنى الذي تتساوى عنده مقاييس النزعة المركزية الثلاثة $(\bar{X} = Me = M0)$ نظري ونادر الوقوع، فالمنحيات التي نحصل عليها عادة تكون ملتوية ناحية اليمين أو ناحية اليسار أو قريبة من الاعتدال .

فقد عرفنا عند دراستنا لمقاييس النزعة المركزية أن التوزيعات الإحصائية يمكن أن تأخذ أحد الأشكال التالية حسب العلاقة بين المقاييس الثلاثة:



التواء ناحية اليسار   التواء ناحية اليمين   توزيع متناظر

$Mo>Me>\overline{X}$ $Mo<Me<\overline{X}$ $\overline{X}=Me=M0$

أما الآن فسنحاول معرفة درجة تناظر (تماثل) توزيع البيانات حول المتوسط الحسابي باستخدام العزوم والانحراف المعياري.

أ) معامل فيشر للالتواء Coefficient de ficher:

يقيس هذا المعامل درجة التواء شكل التوزيع الإحصائي ويعتمد في ذلك على قيمة العزم الثالث حول المتوسط الحسابي، ولاستبعاد وحدة القياس نقسمه على الانحراف المعياري من نفس المرتبة.

$$F_1 = \frac{u_3}{S_x^3}$$

ويكون لدينا ثلاث حالات هي:

توزيع إحصائي متناظر ⇨ $F1=0$

منحنى التوزيع غير متناظر ملتوي ناحية اليمين ⇨ $F1>0$

منحنى التوزيع غير متناظر ملتوي ناحية اليسار ⇨ $F1<0$

ب) معامل بيرسون للالتواء Coefficient de Pearson:

$$P_1 = \frac{(u3)^2}{(u_2)^3}$$

وتكون لدينا ثلاث حالات كذلك:

توزيع إحصائي متناظر ⇨ $P1=0$

منحنى التوزيع غير متناظر ناحية اليمين ⇨ $P1>0$

منحنى التوزيع غير متناظر ملتوي ناحية اليسار ⇨ $P1<0$

ج) معامل يول وكندال للالتواء Coefficient de yule et kendall:

ويستعمل هذا المعامل بالنسبة للجداول الإحصائية المفتوحة.

$$\text{Cyk} = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{Q_3 - Q_1}$$

أما الحالات الممكنة فهي:

توزيع إحصائي متناظر ⇨ $Cyk = 0$

منحنى التوزيع غير متناظر ملتوي ناحية اليمين ⇨ $Cyk > 0$

منحنى التوزيع غير متناظر ملتوي ناحية اليسار ⇨ $Cyk < 0$

II – التفلطح Applatissment (تطاول أو تفلطح المنحنى مقارنة بالتوزيع الطبيعي):

ويقصد بالتفلطح مدى اتساع وضعف قمة منحنى التوزيع ولقد أصطلح على اعتبار منحنى التوزيع الطبيعي متوسط التفلطح.

وتوجد كذلك عدة معاملات لقياس التفلطح أهمها:

أ) معامل بيرسون للتفلطح Coefficient de Pearson:

$$P_2 = \frac{u_4}{(S_x)^4} = \frac{\mu_4}{(\mu_2)^2}$$

والحالات الممكنة هي:

توزيع معتدل التفلطح (توزيع طبيعي) ⇨ P2 = 3

منحنى التوزيع متطاول (مدبب) ⇨ P2>3

منحنى التوزيع متفلطح ⇨ P2<3

ب) معامل fisher للتفلطح Coefficient de fisher

وهو عبارة عن معامل بيرسون مطروحا منه 3.

$$F_2 = P_2 - 3 = \frac{\mu_4}{(\mu_2)^2} - 3$$

والحالات الممكنة هي:

منحنى التوزيع معتدل التفلطح F2 = 0

منحنى التوزيع متطاول F2 > 0

منحنى التوزيع متفلطح F2 < 0

مثال4:

أدرس شكل منحنى التوزيع التكراري الآتي باستخدام معامل فيشر للالتواء ومعامل بيرسون للتفلطح.

المجموع	4	3	2	1	XI
20	1	4	9	6	fi

الحل:

fi(Xi- $\bar{x}$)	fi(Xi- $\bar{x}$)	fi(Xi- $\bar{x}$)^3	fi(Xi- $\bar{x}$)²	(Xi- $\bar{x}$)	fixi	fi	Xi
6	**-6**	**6**	**-6**	**-1**	**6**	**6**	**1**
0	**0**	**0**	**0**	**0**	**18**	**9**	**2**
4	**4**	**4**	**4**	**1**	**12**	**4**	**3**
16	**8**	**4**	**2**	**2**	**4**	**1**	**4**
26	**6**	**14**	**0**		**40**	**20**	**المجموع**

$$\overline{X} = \frac{\sum fiXi}{\sum f_i} = \frac{40}{20} = 2$$

$$\mu_1 = \frac{\sum f_i (X_i - \overline{X})}{\sum f_i} = \frac{0}{20} = 0$$

التباين:

$$\mu_2 = \frac{\sum f_i (X_i - \overline{X})^2}{\sum f_i} = \frac{14}{20} = 0.7 =$$

والانحراف المعياري:

$$S_x = \sqrt{0.7} = 0.83$$

العزم الثالث:

$$\mu_3 = \frac{\sum f_i (X_i - \overline{X})^3}{\sum f_i} = \frac{6}{20} = 0.3$$

العزم الرابع:

$$\mu_4 = \frac{\sum f_i (X_i - \overline{X})^4}{\sum f_i} = \frac{26}{20} = 1.3$$

ومنه معامل فيشر للالتواء:

$$F_1 = \frac{\mu 3}{(S_x)^3} = \frac{0.3}{(0.83)^3} = 0.52$$

معامل فيشر للالتواء موجب هذا يعني أن منحنى التوزيع التكراري ملتوي ناحية اليمين.

$$P_2 = \frac{\mu_4}{(\mu_2)^2} = \frac{1.3}{(0.7)^2} = 2.653$$

معامل بيرسون للتفلطح.

معامل بيرسون للتفلطح أقل من 3 هذا يعني أم منحنى التوزيع يميل للتفلطح.

3 - معامل كيلي للتفلطح Coefficient de kelly

ويستخدم عندما يكون جدول التوزيع التكراري مفتوح من البداية أو من النهاية، ويعطي هذا المعامل بالعلاقة التالية:

$$C_k = \frac{1}{2} \cdot \frac{Q_3 - Q_1}{D_9 - D_1}$$

تمرين:

أدرس شكل التوزيع التكراري باستخدام معاملات بيرسون للالتواء والتفلطح:

الفئة	10 – 8	12 - 10	14 - 12	16 – 14	16 فأكثر	المجموع
التكرار	12	16	20	25	17	90

المطلوب: حساب معاملات الالتواء والتفلطح