مقاييس التشتت (Measures of dispersion)

أولاً: مقاييس التشتت المطلقة (Absolute measures of dispersion):

مؤشرات إحصائية تستخدم لقياس درجة التباعد أو الاختلاف (**مدى التجانس**) بين مفردات الظاهرة الواحدة. وهي عدة أنواع هي:

أ. المدى (**Range**): أبسط مقاييس التشتت المطلقة وأقلها دقة ، وهو عبارة عن الفرق بين أكبر قيمة وأصغر قيمة للبيانات الغير مبوبة. أما للبيانات المبوبة فهو عبارة عن الفرق بين الحد الاعلى للفئة الأخيرة والحد الادنى للفئة الأولى.

$$\mathbf{R = X_L - X_S}$$

حيث أن $\mathbf{R}$: المدى.

$\mathbf{X_L}$: أكبر قيمة.

$\mathbf{X_S}$: أصغر قيمة.

ب. الانحراف الربيعي (**Quartile Deviation**): وهو عبارة عن متوسط الفرق بين الربيع الثالث والربيع الأول. أي أن

$$\mathbf{Q.D. = \frac{Q_3 - Q_1}{2}}$$

حيث أن $\mathbf{Q.D.}$: الانحراف الربيعي.

$\mathbf{Q_1}$: الربيع الاول.

$\mathbf{Q_3}$: الربيع الثالث.

1. حساب الانحراف الربيعي للبيانات الغير مبوبة

لتكن $X_1, X_2, \dots, X_n$ قيم متغير عشوائي والتي تمثل قياسات عينة قوامها (n) ، فلحساب الانحراف الربيعي لهذه القيم نتبع ما يلي:

أ. حساب قيمة الربيع الأول و الربيع الثالث بعد ترتيب القيم تصاعدياً أو تنازلياً ، حيث أن:

$\mathbf{Q_1}$: القيمة التي ترتيبها $\frac{\mathbf{n}}{\mathbf{4}}$

$\mathbf{Q_3}$: القيمة التي ترتيبها $\frac{\mathbf{3n}}{\mathbf{4}}$

ب. حساب الانحراف الربيعي وفق الصيغة

$$\mathbf{Q.D.} = \frac{\mathbf{Q_3 - Q_1}}{\mathbf{2}}$$

مثال (1): البيانات التالية تمثل اوزان ثمانية اشخاص ، والمطلوب حساب درجة التشتت بين اوزان هؤلاء الاشخاص باستخدام.

1. المدى. 2. الانحراف الربيعي.

$\mathbf{X_i}$	**56**	**68**	**72**	**63**	**68**	**71**	**69**	**62**

1. المدى:

$$\mathbf{R = X_L - X_S}$$

$$\mathbf{R = 72 - 56 = 16}$$

2. الانحراف الربيعي:

ترتب القيم تصاعدياً

X_i	56	62	63	68	68	69	71	72

ترتيب الربيع الاول = $\frac{n}{4} = \frac{8}{4} = 2$

$\therefore Q_1 = 62$

ترتيب الربيع الثالث = $\frac{3n}{4} = \frac{3(8)}{4} = 6$

$\therefore Q_1 = 69$

$$\therefore Q.D. = \frac{Q_3 - Q_1}{2} = \frac{69 - 62}{2} = \frac{7}{2} = 3.5$$

2. حساب الانحراف الربيعي لبيانات مبوبة (التوزيعات التكرارية):

أ. حساب قيمة الربيع الأول و كالآتي:

1. عمل توزيع تكراري متجمع صاعد.
2. حساب ترتيب الربيع الاول والذي يساوي

ترتيب الربيع الاول = $\frac{\sum f_i}{4}$

3. تحديد فئة الربيع الاول وهي الفئة التي تقابل التكرار المتجمع الصاعد اللاحق لترتيب الربيع الاول (F_{k+1}).
4. اذا كان التوزيع التكراري **متقطع** فأن قيمة الربيع الأول تساوي مركز فئة الربيع الأول.

5. اذا كان التوزيع التكراري مستمر فأن الربيع الأول يساوي

$$\mathbf{Q_1} = \mathbf{h_k} + \frac{\frac{\sum \mathbf{f_i}}{4} - \mathbf{F_{k-1}}}{\mathbf{f_k}} * \mathbf{L_k}$$

حيث ان $\mathbf{Q_1}$: الربيع الاول.

$\mathbf{h_k}$: الحد الادنى لفئة الربيع الأول.

$\frac{\sum \mathbf{f_i}}{4}$: ترتيب الربيع الأول.

$\mathbf{F_{k-1}}$: التكرار المتجمع الصاعد السابق لترتيب الربيع الأول.

$\mathbf{f_k}$: تكرار فئة الربيع الأول.

$\mathbf{L_k}$: طول فئة الربيع الأول.

ب. حساب قيمة الربيع الثالث ، باتباع نفس خطوات حساب الربيع الأول مع ملاحظة أن ترتيب الربيع الثالث يساوي

ترتيب الربيع الثالث = $\frac{3\sum \mathbf{f_i}}{4}$

جـ . وأخيراً فأن الانحراف الربيعي يحسب وفق الصيغة

$$\mathbf{Q.D.} = \frac{\mathbf{Q_3 - Q_1}}{\mathbf{2}}$$

مثال (2): الآتي توزيع تكراري لعينة من الأسر قوامها **(124)** أسرة حسب عدد الافراد ، والمطلوب حساب درجة التشتت بين عدد أفراد الأسرة باستخدام.

1. المدى.　　　　2. الانحراف الربيعي.

	الفئات	f_i	F_i		
	$2-4$	4	4		
	$5-7$	15	19	F_{k-1}	ترتيب الربيع الاول
فئة الربيع الأول	$\boxed{8-10}$	22	41	F_{k+1}	
	$11-13$	36	77	F_{k-1}	ترتيب الربيع الثالث
فئة الربيع الثالث	$\boxed{14-16}$	28	105	F_{k+1}	
	$17-19$	14	119		
	$20-22$	5	124		
		124			

1. المدى:

$$R = X_L - X_S$$

$$R = 22 - 2 = 20$$

2. الانحراف الربيعي:

ترتيب الربيع الاول = $\frac{\sum f_i}{4} = \frac{124}{4} = 31$

نلاحظ بأن التوزيع التكراري **متقطع** فأن قيمة الربيع الأول تساوي مركز فئة الربيع الأول.

$$\therefore Q_1 = \frac{8+10}{2} = 9$$

ترتيب الربيع الثالث = $\frac{3\sum f_i}{4} = 3(31) = 93$

نلاحظ بأن التوزيع التكراري **متقطع** فأن قيمة الربيع الثالث تساوي مركز فئة الربيع الثالث.

$$\therefore Q_3 = \frac{14+16}{2} = 15$$

$$\therefore Q.D. = \frac{Q_3 - Q_1}{2} = \frac{15-9}{2} = \frac{6}{2} = 3$$

مثال (3): الآتي توزيع تكراري لأعمار تلاميذ احدى المدارس الابتدائية ، والمطلوب تقدير درجة التشتت بين اعمار هؤلاء التلاميذ باستخدام الانحراف الربيعي.

	الفئات	f_i	F_i		
	أقل من 6	16	16		
	$6-7$	120	136	F_{k-1}	ترتيب الربيع الاول
فئة الربيع الأول	$\boxed{7-8}$	131	267	F_{k+1}	
	$8-9$	145	412		
	$9-10$	122	534	F_{k-1}	ترتيب الربيع الثالث
فئة الربيع الثالث	$\boxed{10-11}$	115	649	F_{k+1}	
	$11-12$	101	750		
	$12-13$	22	772		
		772			

الانحراف الربيعي:

نلاحظ بأن التوزيع التكراري **مستمر** فأن قيمة الربيع الأول ، والربيع الثالث تحسب وفقاً للآتي:

ترتيب الربيع الاول = $\frac{\sum f_i}{4} = \frac{772}{4} = 193$

$$Q_1 = h_k + \frac{\frac{\sum f_i}{4} - F_{k-1}}{f_k} * L_k$$

$$Q_1 = 7 + \frac{193 - 136}{131} * (1) \rightarrow \therefore Q_1 = 7.435$$

ترتيب الربيع الثالث = $\frac{3\sum f_i}{4} = 3(193) = 579$

$$Q_3 = h_k + \frac{\frac{3\sum f_i}{4} - F_{k-1}}{f_k} * L_k$$

$$Q_3 = 10 + \frac{579 - 534}{115} * (1) \rightarrow \therefore Q_3 = 10.391$$

$$\therefore Q.D. = \frac{Q_3 - Q_1}{2} = \frac{10.391 - 7.435}{2} = \frac{2.956}{2} = 1.478$$

ملاحظات حول الانحراف الربيعي

يعتبر الانحراف الربيعي أفضل من المدى كونه يستخدم (50%) من البيانات المتاحة فقط ويهمل الباقي. أن هذا المقياس قليل الاستخدام في حين أنهُ يمتاز على بقية مقاييس التشتت المطلقة من حيث امكانية حسابه في حالة **التوزيعات التكرارية المفتوحة** في حين يتعذر الحصول على بقية مقاييس التشتت.

جـ. الانحراف المتوسط (**Mean Deviation**):

يعـرف الانحــراف المتوســط بأنــهُ مجمــوع الانحرافــات المطلقــة لقــيم المتغيــر العشــوائي عــن نقطــة اختياريــة مثــل (**A**) مقســوماً علــى عــدد هــذه القــيم وغالبــاً نختــار النقطــة (القيمــة) (**A**) لان تكــون أحــد مقــاييس النزعــة المركزيــة الــثلاث (الوســط الحســابي ، الوســيط ، المنــوال) ، وســنتناول الانحــراف المتوســط بالاعتمــاد علــى الوســط الحسابي.

1. حساب الانحراف المتوسط لبيانات غير مبوبة:

لتكن $X_1 , X_2 , ... , X_n$ قــيم متغيــر عشــوائي والتــي تمثــل قياســات عينــة قوامهــا (n) ، ولــيكن $(\overline{X})$ الوســط الحســابي لتلــك القياســات. فــأن الانحــراف المتوســط (M.D.) لتلــك القياســات يحســب وفق الصيغة الآتية:

$$\mathrm{M.D.} = \frac{\sum |X_i - A|}{n}$$

$$\mathrm{M.D.} = \frac{\sum |X_i - \overline{X}|}{n} = \frac{|X_1 - \overline{X}| + |X_2 - \overline{X}| + \cdots + |X_n - \overline{X}|}{n}$$

مثال (4): للبيانات التالية جد الانحراف المتوسط

X_i	2	3	4	5	5	6	7	10	13	14	19

X_i	$\lvert X_i - \overline{X} \rvert$
2	6
3	5
4	4
5	3
5	3
6	2
7	1
10	2
13	5
14	6
19	11
88	48

1. نجد الوسط الحسابي

$$\overline{X} = \frac{\sum X_i}{n} = X_1 + X_2 + \cdots + X_n$$

$$\overline{X} = \frac{2 + 3 + \cdots + 19}{11} = \frac{88}{11} = 8$$

2. نجد الانحراف المتوسط

$$M.D. = \frac{\sum \lvert X_i - \overline{X} \rvert}{n} = \frac{\lvert X_1 - \overline{X} \rvert + \lvert X_2 - \overline{X} \rvert + \cdots + \lvert X_n - \overline{X} \rvert}{n}$$

$$M.D. = \frac{\lvert 2 - 8 \rvert + \lvert 3 - 8 \rvert + \cdots + \lvert 19 - 8 \rvert}{11} = \frac{6 + 5 + \cdots + 11}{11}$$

$$M.D. = \frac{48}{11} = 4.367$$

2. حساب الانحراف المتوسط لبيانات مبوبة (التوزيعات التكرارية):

لتكن $X_1, X_2, \dots, X_m$ مراكز الفئات لتوزيع تكراري عدد فئاته (m) وأن $f_1, f_2, \dots, f_m$ التكرارات المقابلة لهذه الفئات. فأن الانحراف المتوسط (M.D.) لتلك البيانات يحسب وفق الصيغة الآتية:

$$M.D. = \frac{\sum |X_i - \bar{X}| * f_i}{\sum f_i}$$

$$M.D. = \frac{|X_1 - \bar{X}| * f_1 + |X_2 - \bar{X}| * f_2 + \cdots + |X_m - \bar{X}| * f_m}{f_1 + f_2 + \cdots + f_m}$$

حيث أن:

$$\bar{X} = \frac{\sum X_i * f_i}{\sum f_i} = \frac{X_1 * f_1 + X_2 * f_2 + \cdots + X_m * f_m}{f_1 + f_2 + \cdots + f_m}$$

مثال (5): الآتي توزيع تكراري لدرجات مجموعة من طلبة في امتحان معين. يطلب حساب درجة التشتت بين درجات هؤلاء الطلبة باستخدام الانحراف المتوسط

الفئات	f_i	x_i	$f_i * x_i$	$\lvert X_i - \bar{X}\rvert * f_i$
$0-10$	2	5	10	109
$10-20$	4	15	60	178
$20-30$	8	25	200	276
$30-40$	16	35	560	392
$40-50$	25	45	1125	362.5
$50-60$	60	55	3300	270
$60-70$	42	65	2730	231
$70-80$	35	75	2625	542.5
$80-90$	18	85	1530	459
$90-100$	10	95	950	355
	220		13090	3175

1. نجد الوسط الحسابي

$$\overline{X} = \frac{\sum X_i * f_i}{\sum f_i} = \frac{X_1 * f_1 + X_2 * f_2 + \cdots + X_m * f_m}{f_1 + f_2 + \cdots + f_m}$$

$$\overline{X} = \frac{5 * 2 + 4 * 15 + \cdots + 10 * 95}{2 + 4 + \cdots + 10} = \frac{13090}{220} = 59.5$$

2. نجد الانحراف المتوسط

$$M.D. = \frac{\sum |X_i - \overline{X}| * f_i}{\sum f_i}$$

$$M.D. = \frac{|X_1 - \overline{X}| * f_1 + |X_2 - \overline{X}| * f_2 + \cdots + |X_m - \overline{X}| * f_m}{f_1 + f_2 + \cdots + f_m}$$

$$M.D. = \frac{|5 - 59.5| * 2 + |15 - 59.5| * 4 + \cdots + |95 - 59.5| * 10}{2 + 4 + \cdots + 10}$$

$M.D. = \frac{3175}{220} = 14.432$ درجة

مزايا الانحراف المتوسط:

1. أن حسابه يعتمد على كافة البيانات المتاحة.
2. أنهُ مقياس سهل الفهم والحساب.
3. خضوعهِ للعمليات الجبرية.
4. أن قيمتهِ تكون أقل ما يمكن عندما يحسب باستخدام الوسيط.

عيوب الانحراف المتوسط:

I. اهمال الاشارات السالبة للفروق عند عملية حسابهِ.
II. لا يمكن حسابهِ في حالة التوزيعات التكرارية المفتوحة.
III. لا يمكن حسابهِ في حالة البيانات الوصفية.
IV. تتأثر قيمتهِ في حالة وجود قيم شاذة أو متطرفة.
V. يتأثر وعلى نحو كبير بأخطاء المعاينة.

د. الانحراف المعياري (**Standard Deviation**):

يعتبر هذا المقياس أفضل مقاييس التشتت المطلقة ، ويعرف بأنهُ الجذر التربيعي الموجب للتباين ، و**التباين** (**Variance**) عبارة عن متوسط مجموع مربعات انحرافات قيم المتغير العشوائي عن وسطها الحسابي.

1. حساب الانحراف المعياري لبيانات غير مبوبة:

لتكن $X_1, X_2, \dots, X_n$ قيم متغير عشوائي والتي تمثل قياسات عينة قوامها (n) ، وليكن $(\overline{X})$ الوسط الحسابي لتلك القياسات. فأن الانحراف المعياري (**S**) لتلك القياسات يحسب وفق الصيغة الآتية:

$$S = \sqrt{\frac{\sum (X_i - \overline{X})^2}{n}} = \sqrt{\frac{(X_1 - \overline{X})^2 + (X_2 - \overline{X})^2 + \cdots + (X_n - \overline{X})^2}{n}}$$

أما التباين لهذه القياسات فيساوي

$$S^2 = \frac{\sum (X_i - \overline{X})^2}{n} = \frac{(X_1 - \overline{X})^2 + (X_2 - \overline{X})^2 + \cdots + (X_n - \overline{X})^2}{n}$$

هذا يعني أن:

$$S = \sqrt{S^2}$$

أن الصيغة السابقة لحساب الانحراف المعياري أو التباين قد تكون مطولة ، ولغرض تسهيل العمليات الحسابية يمكن اتباع احدى الصيغتين الآتيتين بالاعتماد على الاشتقاق الآتي:

من خلال النظر للصيغة $\sum(X_i - \bar{X})^2$: مجموع مربعات انحرافات القيم عن وسطها الحسابي في حل الأمثلة تكون مطولة لذا يتم اختصارها وكما يلي:

$$\sum(X_i - \bar{X})^2 = \sum(X_i{}^2 - 2X_i\bar{X} + \bar{X}^2) = \sum X_i{}^2 - 2\bar{X}\sum X_i + n\bar{X}^2$$

$$\bar{X} = \frac{\sum X_i}{n} \rightarrow \sum X_i = n\bar{X}$$

$$\sum(X_i - \bar{X})^2 = \sum x_i{}^2 - n\bar{X}^2$$

$$S = \sqrt{\frac{\sum x_i{}^2 - n\bar{X}^2}{n}} \qquad \dots(3)$$

$$S = \sqrt{\frac{\sum x_i{}^2}{n} - \left(\frac{\sum x_i}{n}\right)^2} \qquad \dots(4)$$

أي أن التباين

$$S^2 = \frac{\sum x_i{}^2 - n\bar{X}^2}{n} \qquad \dots(5)$$

$$S^2 = \frac{\sum x_i{}^2}{n} - \left(\frac{\sum x_i}{n}\right)^2 \qquad \dots(6)$$

ملاحظات:

1. تستخدم الصيغة رقم (3) اذا توفرت قيمة الوسط الحسابي.
2. تستخدم الصيغة رقم (4) اذا لم تتوفر قيمة الوسط الحسابي.
3. يمكن حساب التباين أولاً ثم اخذ الجذر التربيعي لهُ لنحصل على قيمة الانحراف المعياري.

مثال (6): البيانات التالية تمثل أوزان عينة من الطلبة قوامها عشرة طلاب. قدر درجة التشتت بين أوزانهم باستخدام الانحراف المعياري.

X_i	X_i^2
56	3136
62	3844
69	4761
71	5041
68	4624
65	4225
63	3969
72	5184
68	4624
56	3136
650	42544

$$S = \sqrt{\frac{\sum x_i^2}{n} - \left(\frac{\sum x_i}{n}\right)^2} = \sqrt{\frac{42544}{10} - \left(\frac{650}{10}\right)^2} = \sqrt{4254.4 - 4225}$$

$$S = \sqrt{29.4} = 5.422$$

ملاحظة : إذا أريد حساب التباين فأن التباين هو:

$$S^2 = \frac{\sum x_i^2}{n} - \left(\frac{\sum x_i}{n}\right)^2 = \frac{42544}{10} - \left(\frac{650}{10}\right)^2 = 4254.4 - 4225 = 29.4$$

ملاحظة : في المثال رقم (6) ، استخدمت الصيغة رقم (4) ، والصيغة رقم (5) لحساب الانحراف المعياري والتباين لعدم وجود الوسط الحسابي.

مثـال (7): أحسـب الانحـراف المعيـاري والتبـاين للبيانـات الـواردة فـي المثال رقم (4). علماً أن الوسط الحسابي هو $\bar{X} = 8$

X_i	X_i^2
2	4
3	9
4	16
5	25
5	25
6	36
7	49
10	100
13	169
14	196
19	361
88	990

$$S = \sqrt{\frac{\sum x_i^2 - n\bar{X}^2}{n}} = \sqrt{\frac{990 - (11)(8)^2}{11}} = \sqrt{\frac{990 - 704}{11}}$$

$$S = \sqrt{\frac{286}{11}} = \sqrt{26} = 5.099$$

$$S^2 = \frac{\sum x_i^2 - n\bar{X}^2}{n} = \frac{990 - (11)(8)^2}{11} = \frac{990 - 704}{11} = \frac{286}{11}$$

$$S^2 = 26$$

2. حساب الانحراف المعياري للبيانات المبوبة (التوزيعات التكرارية):

لتكن $X_1, X_2, \dots, X_m$ مراكـــز الفئـــات لتوزيـــع تكـــراري عـــدد فئاتـــه (m) وأن $f_1, f_2, \dots, f_m$ التكـــرارات المقابلـــة لهـــذه الفئـــات. فـــأن الانحـــراف المعيـــاري (S) لهـــذه البيانـــات يحســـب وفـــق الصيغة الآتية:

$$S = \sqrt{\frac{\sum (X_i - \overline{X})^2 * f_i}{\sum f_i}} \quad \dots (1)$$

$$S = \sqrt{\frac{(X_1 - \overline{X})^2 * f_1 + (X_2 - \overline{X})^2 * f_2 + \cdots + (X_n - \overline{X})^2 * f_n}{f_1 + f_2 + \cdots + f_n}}$$

أما التباين لهذه القياسات فيساوي

$$S^2 = \frac{\sum (X_i - \overline{X})^2 * f_i}{\sum f_i} \quad \dots (2)$$

$$S^2 = \frac{(X_1 - \overline{X})^2 * f_1 + (X_2 - \overline{X})^2 * f_2 + \cdots + (X_n - \overline{X})^2 * f_n}{f_1 + f_2 + \cdots + f_n}$$

هذا يعني أن:

$$S = \sqrt{S^2}$$

وكمـــا أوردنـــا فـــي البيانـــات الغيـــر مبوبـــة أن الصـــيغتين (1) ، و (2) الســـابقة لحســـاب الانحـــراف المعيـــاري أو التبـــاين تكـــون مطولـــة ، ولغـــرض تســـهيل العمليـــات الحســـابية يمكـــن اتبـــاع احـــدى الصـــيغتين الآتيتين بالاعتماد على الاشتقاق الآتي:

من خلال النظر للصيغة $\sum (X_i - \overline{X})^2 * f_i$ في حل الأمثلة تكون مطولـة لذا يتم اختصارها وكما يلي:

$$\sum(X_i-\bar{X})^2*f_i=\sum\left(X_i{}^2-2X_i\bar{X}+\bar{X}^2\right)*f_i=\sum X_i{}^2*f_i-2\bar{X}\sum X_i*f_i+\bar{X}^2\sum f_i$$

$$\bar{X}=\frac{\sum X_i*f_i}{\sum f_i}\rightarrow\sum X_i*f_i=\left(\sum f_i\right)*\bar{X}$$

$$\sum(X_i-\bar{X})^2*f_i=\sum x_i{}^2*f_i-\left(\sum f_i\right)\bar{X}^2$$

$$S=\sqrt{\frac{\sum x_i{}^2*f_i-(\sum f_i)\bar{X}^2}{\sum f_i}}\qquad\ldots(3)$$

$$S=\sqrt{\frac{\sum x_i{}^2*f_i}{\sum f_i}-\left(\frac{\sum x_i*f_i}{\sum f_i}\right)^2}\qquad\ldots(4)$$

أي أن التباين

$$S^2=\frac{\sum x_i{}^2*f_i-(\sum f_i)\bar{X}^2}{\sum f_i}\qquad\ldots(5)$$

$$S^2=\frac{\sum x_i{}^2*f_i}{\sum f_i}-\left(\frac{\sum x_i*f_i}{\sum f_i}\right)^2\qquad\ldots(6)$$

ملاحظات:

1. تستخدم الصيغة رقم (3) اذا توفرت قيمة الوسط الحسابي.
2. تستخدم الصيغة رقم (4) اذا لم تتوفر قيمة الوسط الحسابي.
3. يمكن حساب التباين أولاً ثم اخذ الجذر التربيعي لهُ لنحصل على قيمة الانحراف المعياري.

مثـــال (8): الآتـــي توزيــع تكــراري لــدرجات مجموعـــة مــن طلبـــة فـــي امتحـــان معــين. يطلـــب حســـاب درجـــة التشـــتت بـــين درجـــات هـــؤلاء الطلبة باستخدام الانحراف المعياري ثم حساب التباين.

الفئات	f_i	x_i	$f_i * x_i$	$X_i^2 * f_i$
$0-10$	2	5	10	50
$10-20$	4	15	60	900
$20-30$	8	25	200	5000
$30-40$	16	35	560	19600
$40-50$	25	45	1125	50625
$50-60$	60	55	3300	181500
$60-70$	42	65	2730	177450
$70-80$	35	75	2625	196875
$80-90$	18	85	1530	130050
$90-100$	10	95	950	90250
	220		13090	852300

$$S = \sqrt{\frac{\sum x_i^2 * f_i}{\sum f_i} - \left(\frac{\sum x_i * f_i}{\sum f_i}\right)^2} = \sqrt{\frac{852300}{220} - \left(\frac{13090}{220}\right)^2}$$

$$\mathbf{S = \sqrt{3874.091 - 3540.25} = \sqrt{333.841} = 18.271}$$

$$S^2 = \frac{\sum x_i^2 * f_i}{\sum f_i} - \left(\frac{\sum x_i * f_i}{\sum f_i}\right)^2$$

$$S^2 = \frac{852300}{220} - \left(\frac{13090}{220}\right)^2 = \mathbf{3874.091 - 3540.25 = 333.841}$$

مثـــال (9): الآتـــي توزيــع تكــراري لرواتـــب مجموعـــة مـــن المـــوظفين علـــى فئـــات الراتـــب الأســـمي للمـــوظفين. أحســـب الانحـــراف المعيـــاري لقيـــاس درجـــة التشـــتت بـــين رواتـــب هـــؤلاء المـــوظفين إذا علمـــت $\bar{X} = 90.917$

الفئات	f_i	x_i	$X_i^2 * f_i$
$69.5 - 72.5$	6	71	30246
$72.5 - 76.5$	10	74.5	55502.5
$76.5 - 88.5$	18	82.5	122512.5
$88.5 - 97.5$	25	93	216225
$97.5 - 112.5$	12	105	132300
$112.5 - 130.5$	7	121.5	103335.75
	78		660121.75

$$S = \sqrt{\frac{\sum x_i^2 * f_i - (\sum f_i)\bar{X}^2}{\sum f_i}} = \sqrt{\frac{660121.75 - (78) * (90.917)^2}{78}}$$

$$S = \sqrt{\frac{660121.75 - 644740.27}{78}} = \sqrt{\frac{15381.48}{78}} = \sqrt{197.199} = 14.043$$

أي أن التباين يساوي

$$S^2 = \frac{\sum x_i^2 * f_i - (\sum f_i)\bar{X}^2}{\sum f_i} = \frac{660121.75 - (78) * (90.917)^2}{78}$$

$$S^2 = \frac{660121.75 - 644740.27}{78} = \frac{15381.48}{78} = 197.199$$

مزايا و عيوب الانحراف المعياري

أ. مميزات الانحراف المعياري:

1) حسابه يعتمد على كافة البيانات.

2) أنهُ مقياس سهل الفهم والحساب.

3) خضوعهِ للعمليات الجبرية.

4) قابليتهِ للتجزئة والاندماج.

ب.عيوب الانحراف المعياري:

1) لا يمكن حسابه للتوزيعات التكرارية المفتوحة.

2) لا يمكن حسابه للمتغيرات الوصفية.

3) تتأثر قيمته في حالة وجود قيم شاذة أو متطرفة.

4) يتأثر وعلى نحو كبير بأخطاء المعاينة.

ملاحظة (مهمة): أن مزايا و عيوب **التباين** هي نفسها مزايا و عيوب الانحراف المعياري.

خصائص الانحراف المعياري

أ. أن قيمة (S) الانحراف المعياري دائماً موجبة أي ان

$$S \geq 0$$

وتكون قيمة الانحراف المعياري تساوي صفراً إذا كانت قيمة مفردات العينة متساوية. أي أن العينة ذات قيم ثابتة.

ب. إذا كانت $X_1, X_2, \dots, X_n$ تمثل قيم المتغير العشوائي X_i والذي يمثل قياسات العينة وأن (a) ثابت **(Constant)**، وأن Y_i هو متغير عشوائي حيث أن:

$$Y_i = a * X_i$$

فأن الانحراف المعياري للمتغير العشوائي Y_i يعطى وفق الصيغة الآتية:

$$S_Y = |a| * S_X$$

ج. إذا كانت (a) كمية ثابتة وأن $Y_i = X_i \pm a$ فأن :

$$S_Y = S_X$$

بمعنى إذا اضيفت أو طرحت قيمة ثابتة لقيم العينة فأن الانحراف المعياري قيمته لا تتغير.

مثال (10): إذا كانت لدينا عينة بحجم (5) مفردات لها القياسات الآتية (6,1,4,2,5) ، وبافتراض أن (a=-3) فأثبت حسابياً أن:

1) $S_Y = |a| * S_X$

إذا علمت أن

$Y_i = a * X_i$

2) $S_Y = S_X$

إذا علمت أن

$Y_i = X_i \pm a$

الحل: أولاً

X_i	X_i^2	$Y_i = a * X_i$	Y_i^2	$Y_i = X_i + a$	Y_i^2
6	36	-3*6=-18	324	6+(-3)=3	9
1	1	-3*1=-3	9	1+(-3)=-2	4
4	16	-3*4=-12	144	4+(-3)=1	1
2	4	-3*2=-6	36	2+(-3)=-1	1
5	25	-3*5=-15	225	5+(-3)=2	4
18	82	-54	738	3	19

$$S = \sqrt{\frac{\sum x_i^2}{n} - \left(\frac{\sum x_i}{n}\right)^2}$$

$$S_x = \sqrt{\frac{82}{5} - \left(\frac{18}{5}\right)^2} = \sqrt{16.4 - 12.96} = \sqrt{3.44} = 1.855$$

$$S_y = \sqrt{\frac{\sum y_i^2}{n} - \left(\frac{\sum y_i}{n}\right)^2}$$

في الخاصية الاولى :

$$S_y = \sqrt{\frac{738}{5} - \left(\frac{-54}{5}\right)^2} = \sqrt{147.6 - 116.64} = \sqrt{30.96} = 5.564$$

$$|a| * S_X = |-3| * 1.855 = 5.565 = S_Y = 5.564$$

في الخاصية الثانية:

$$S_y = \sqrt{\frac{19}{5} - \left(\frac{3}{5}\right)^2} = \sqrt{3.8 - 0.36} = \sqrt{3.44} = 1.855$$

$$S_Y = S_X$$

تمرين: أستخدم بيانات المثال رقم (10) لإثبات أن

$$\mathbf{S_Y = |a| * S_X}$$

إذا علمت أن

$$Y_i = X_i \pm a$$

خصائص التباين

أ. أن قيمة (S^2) التباين دائماً موجبة أي ان

$$\mathbf{S^2 \geq 0}$$

وتكون قيمة التباين تساوي صفراً إذا كانت قيمة مفردات العينة متساوية. أي أن العينة ذات قيم ثابتة.

ب. إذا كانت $X_1, X_2, \dots, X_n$ تمثل قيم المتغير العشوائي X_i والذي يمثل قياسات العينة وأن (a) ثابت (**Constant**) ، وأن Y_i هو متغير عشوائي حيث أن:

$$Y_i = a * X_i$$

فأن التباين للمتغير العشوائي Y_i يعطى وفق الصيغة الآتية:

$$\mathbf{S^2{}_Y = a^2 * S^2{}_X}$$

ج. إذا كانت (a) كمية ثابتة وأن $Y_i = X_i \pm a$ فأن :

$$\mathbf{S^2{}_Y = S^2{}_X}$$

د. إذا كانت (a) كمية ثابتة وأن $Y_i = aX_i \pm b$ فأن :

$$\mathbf{S^2{}_Y = a^2 * S^2{}_X}$$

مثال (11):تمرين: أستخدم بيانات المثال رقم (10) لإثبات الخواص ((ب) ، و (ج) ، و (د)).

هـ . لتكن $S_1{}^2, S_2{}^2, \dots, S_k{}^2$ تمثل التباينات المحسوبة لسلسلة من العينات المتعاقبة عن المتغير X التي قوام كل منها وعلى التتابع $\mathbf{n_1, n_2, \dots, n_k}$ عندئذ التباين الكلي لحاصل دمج قياسات هذه العينات معاً والذي غالباً ما يصطلح عليه بمحصلة التباين (**Pooled Variance**) هو

$$\mathbf{S^2{}_X} = \frac{\sum n_j * S^2{}_j}{\sum n_j} \qquad , \text{Where } j = 1,2, \dots, k$$

مع ملاحظة أنهُ إذا كان عدد مفردات كافة هذه العينات متساوي أي أن $\mathbf{n_1 = n_2 = \cdots = n_k = n}$ ، عندئذ فأن محصلة التباين $\mathbf{S^2{}_X}$ تساوي الوسط الحسابي لتباينات هذه العينات أي أن:

$$\mathbf{S^2{}_X} = \frac{\sum S^2{}_j}{k} \qquad , \text{Where } j = 1,2, \dots, k$$

مثال (**12**): إذا علمت أن تباين أعمار طلبة الصف الأول إحصاء الشعبة (أ) عددهم (**52**) يساوي (**2.5**) ، وأن تباين أعمار طلبة الشعبة (**ب**) البالغ عددهم (**63**) يساوي (**2.1**). جد تباين أعمار الطلبة في هذه المرحلة.

$$\mathbf{S^2{}_X} = \frac{\sum n_j * S^2{}_j}{\sum n_j}$$

$$\mathbf{S^2} = \frac{\mathbf{n_1 * S^2{}_1 + n_2 * S^2{}_2}}{\mathbf{n_1 + n_2}} = \frac{\mathbf{(52)(2.5) + (63)(2.1)}}{\mathbf{52 + 63}} = \frac{\mathbf{262.3}}{\mathbf{115}} = \mathbf{2.281}$$

مثـــال (**13**): إذا علمــت أن تبـــاين أعمـــار طلبـــة الصـــف الأول إحصـــاء الشـــعبة (أ) عـــددهم (**50**) يســـاوي (**2.5**) ، وأن تبـــاين أعمـــار طلبـــة الشـــعبة (**ب**) البـــالغ عـــددهم (**50**) أيضـــاً يســـاوي (**2.1**). جـــد تبـــاين أعمار الطلبة في هذه المرحلة.

$$\mathbf{S^2}_{\mathrm{X}} = \frac{\sum \mathrm{S^2}_{\mathrm{j}}}{\mathrm{k}}$$

$$\mathbf{S^2} = \frac{\mathbf{S^2}_1 + \mathbf{S^2}_2}{\mathbf{2}} = \frac{(\mathbf{2.5}) + (\mathbf{2.1})}{\mathbf{2}} = \frac{\mathbf{4.6}}{\mathbf{2}} = \mathbf{2.3}$$

تمارين عن مقاييس التشتت المطلقة:

1. لمجاميع البيانات التالية جد كافة مقاييس التشتت المطلقة.

I.

X_i	2	1	3	4	6	7	6	5	9	10	2	3	1	4	6	2	3	8	11

II.

X_i	-1	-3	-2.5	-4	0	1.5	-2	-4.5	-6	-5

III.

$$X_i \quad , \text{Where } i = 1, 2, \ldots, 20$$

2. الآتي توزيع تكراري لأعمار عينة من الأفراد قوامها 542 فرد. جد:

1. الانحراف المتوسط. 2. الانحراف المعياري. 3. التباين.

الفئات	f_i
20 –	3
30 –	61
40 –	132
50 –	153
60 –	140
70 –	51
80 – 90	2

3. الآتي توزيع تكراري لأوزان مجموعة من الأطفال. يطلب إيجاد مقياس التشتت الملائم.

الفئات	f_i
أقل من 4	10
4 –	25
6 –	36
8 –	32
10 فأكثر	17

4. شركة للغزل والنسيج مؤلفة من ثلاثة مصانع. فإذا علمت أن متوسط الأجر الشهري والانحراف المعياري للأجر في كل مصنع كان على النحو الآتي :

متوسط الأجر $\bar{X}_j$	الانحراف المعياري $(S^2{}_j)$	عدد العاملين (n_j)	تسلسل المصنع
107	4.62	122	1
103	5.53	157	2
111	6.32	198	3

جد متوسط الأجر والتباين للأجر في هذه الشركة.

5. جد قيمة (Z) التي تجعل التباين لقيم المجموعة التالية يساوي (4).

X_i	2	5	4	3	8	7	Z

6. للبيانات الواردة في التمرين رقم (4) ، جد تباين الأجر للشركة بافتراض أن عدد العاملين متساوي في كل مصنع.