

مبادئ الإحصاء التربوي

الفصل الدراسي الأول

العام الدراسي ١٤٣١ - ١٤٣٢ هـ

د. سعيد سيف الدين



نظام التعليم المطور للانتساب
كلية التربية

المحاضرة الحادية عشرة

تابع الباب الرابع
[مقاييس التشتت]



عناصر المحاضرة

- حل مسائل "سلي نفسك" الموجودة بالمحاضرة العاشرة
- تابع مقاييس التشتت
 - الانحراف الربيعي [نصف المدى الربيعي]
 - المدى المثني
- علاقات اعتبارية بين مقاييس التشتت
- التشتت النسبي ومقاييسه
- الدرجات المعيارية



حل مسائل "سلي نفسك" الموجودة بالمحاضرة العاشرة

(١) أوجد المدى R ، الوسط الحسابي $\bar{x}$ ، الانحراف المتوسط $M.D$ ، التباين s^2 ، والانحراف المعياري s لمجموعة القيم :

5 3 8 4 7 6 12 4 3 8

الحل : عدد القيم n يساوي 10

• المدى $R = \text{أكبر قيمة} - \text{أقل قيمة} = 12 - 3 = 9$

• الوسط الحسابي $\bar{x} = \frac{\sum x}{n} = \frac{60}{10} = 6$

• الانحراف المتوسط $M.D = \frac{\sum |d|}{n} = \frac{22}{10} = 2.2$

• التباين $s^2 = \frac{\sum d^2}{n} = \frac{76}{10} = 7.6$

• الانحراف المعياري $s = \sqrt{s^2} = \sqrt{7.6} \cong 2.76$

x	d	$ d $	d^2
5	-1	1	1
3	-3	3	9
8	2	2	4
4	-2	2	4
7	1	1	1
6	0	0	0
12	6	6	36
4	-2	2	4
3	-3	3	9
8	2	2	4
60		22	76
$\sum x$		$\sum d $	$\sum d^2$

(٢) أوجدى المدى R ، الوسط الحسابي $\bar{x}$ ، الانحراف المتوسط $M.D$ ، التباين s^2 ، والانحراف المعياري s للبيانات المبينة بالجدول التكراري المقابل .

الحل :

الجدول التكراري	المتغير x	التكرار f	fx	$d = x - \bar{x}$	$ d $	$f d $	d^2	fd^2
	8	20	160	$8 - 4.5 = 3.5$	3.5	$20 \times 3.5 = 70$	12.25	245
	2	30	60	$2 - 4.5 = -2.5$	2.5	$30 \times 2.5 = 75$	6.25	187.5
	4	35	140	$4 - 4.5 = -0.5$	0.5	$35 \times 0.5 = 17.5$	0.25	8.75
	6	15	90	$6 - 4.5 = 1.5$	1.5	$15 \times 1.5 = 22.5$	2.25	33.75
		100	450			185		475
		$\sum f$	$\sum fx$			$\sum f d $		$\sum fd^2$

• المدى $R = \text{أكبر قيمة} - \text{أقل قيمة} = 8 - 2 = 6$

• الانحراف المتوسط $M.D = \frac{\sum f|d|}{\sum f} = \frac{185}{100} = 1.85$

• الانحراف المعياري $s = \sqrt{s^2} = \sqrt{4.75} = 2.18$

• الوسط الحسابي $\bar{x} = \frac{\sum fx}{\sum f} = \frac{450}{100} = 4.5$

• التباين $s^2 = \frac{\sum fd^2}{\sum f} = \frac{475}{100} = 4.75$

المتغير x	$5 \leq x < 15$	$15 \leq x < 25$	$25 \leq x < 45$	$45 \leq x < 55$
التكرار f	20	30	40	10

(٣) أوجد s ، s^2 ، $M.D$ ، $\bar{x}$ ، R للبيانات المبينة بالجدول التكراري المقابل .

الحل :

الجدول التكراري

المتغير x	التكرار f	x_0	fx_0	$d = x_0 - \bar{x}$	$ d $	$f d $	d^2	fd^2
$5 \leq x < 15$	20	10	200	$10 - 27 = -17$	17	$20 \times 17 = 340$	289	5780
$15 \leq x < 25$	30	20	600	$20 - 27 = -7$	7	$30 \times 7 = 210$	49	1470
$25 \leq x < 45$	40	35	1400	$35 - 27 = 8$	8	$40 \times 8 = 320$	64	2560
$45 \leq x < 55$	10	50	500	$50 - 27 = 23$	23	$10 \times 23 = 230$	529	5290
	100		2700			1100		15100
	$\sum f$		$\sum fx_0$			$\sum f d $		$\sum fd^2$

• المدى $R = \text{أكبر قيمة} - \text{أقل قيمة} = \text{الحد الأعلى للفئة الأخيرة} - \text{الحد الأدنى للفئة الأولى} = 55 - 5 = 50$

• الانحراف المتوسط $M.D = \frac{\sum f|d|}{\sum f} = \frac{1100}{100} = 11$

• الانحراف المعياري $s = \sqrt{s^2} = \sqrt{151} \approx 12.29$

• الوسط الحسابي $\bar{x} = \frac{\sum fx_0}{\sum f} = \frac{2700}{100} = 27$

• التباين $s^2 = \frac{\sum fd^2}{\sum f} = \frac{15100}{100} = 151$



رابعاً : الانحراف الربيعي [نصف المدى الربيعي]

لمجموعة من البيانات يُعرف الانحراف الربيعي [أو نصف المدى الربيعي وسنرمز له بالرمز Q] كالآتي :

$$Q = \frac{1}{2}(Q_3 - Q_1)$$

الربيع الثالث

الربيع الأول

ويفضل استخدام هذا المقياس [الانحراف الربيعي] في الكثير من الحالات خاصة تلك الحالات التي يستعصي فيها حساب الانحراف المتوسط أو المعياري [مثل حالة التوزيعات التكرارية المفتوحة أو حالة وجود قيم متطرفة في البيانات]

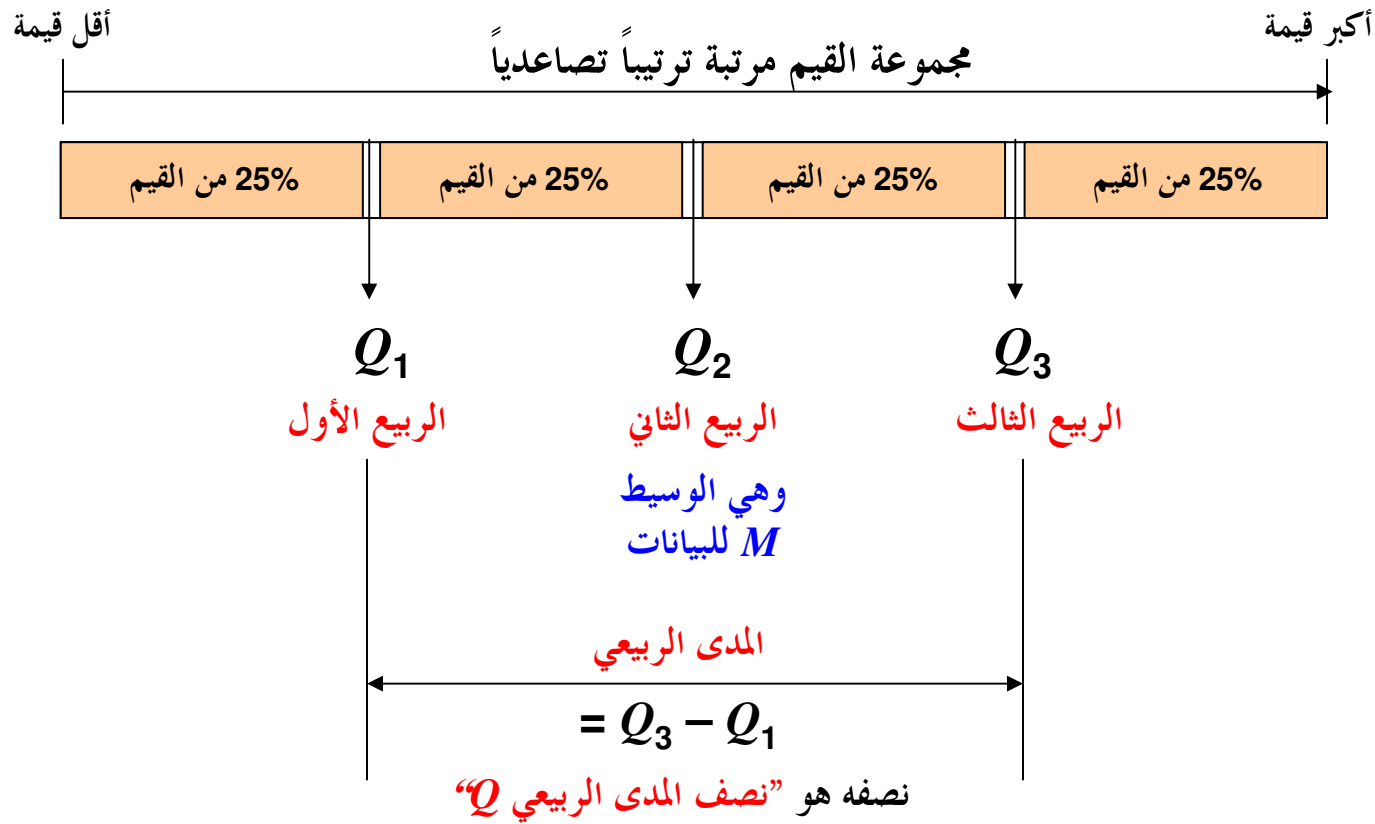
وفي بعض الأحيان يُستخدم المدى الربيعي $Q_3 - Q_1$ كمقياس للتشتت بدلاً من نصف المدى الربيعي

س : ما هي الربيعات ؟

ج : إذا رتبنا مجموعة من القيم ترتيباً تصاعدياً فإن القيمة التي تقسم المجموعة إلى مجموعتين متساويتين في العدد تُسمى بالوسيط M . بتعميم هذه الفكرة ، يمكن أن نقسم مجموعة القيم إلى أربعة أجزاء متساوية في العدد وذلك بثلاثة قيم [سنرمز لها بالرموز Q_1 ، Q_2 ، Q_3] . هذه القيم تُسمى بالربيعات حيث :

Q_1 تُسمى بالربيع الأول ، Q_2 تُسمى بالربيع الثاني ، Q_3 تُسمى بالربيع الثالث





أي أن :

Q_1 [الربع الأول] هي تلك القيمة التي يقع تحتها 25% من القيم [وبالطبع وفوقها 75% من القيم]

Q_2 [الربع الثاني] هي تلك القيمة التي يقع تحتها 50% من القيم [وبالطبع وفوقها 50% من القيم] [أي الوسيط M]

Q_3 [الربع الثالث] هي تلك القيمة التي يقع تحتها 75% من القيم [وبالطبع وفوقها 25% من القيم]

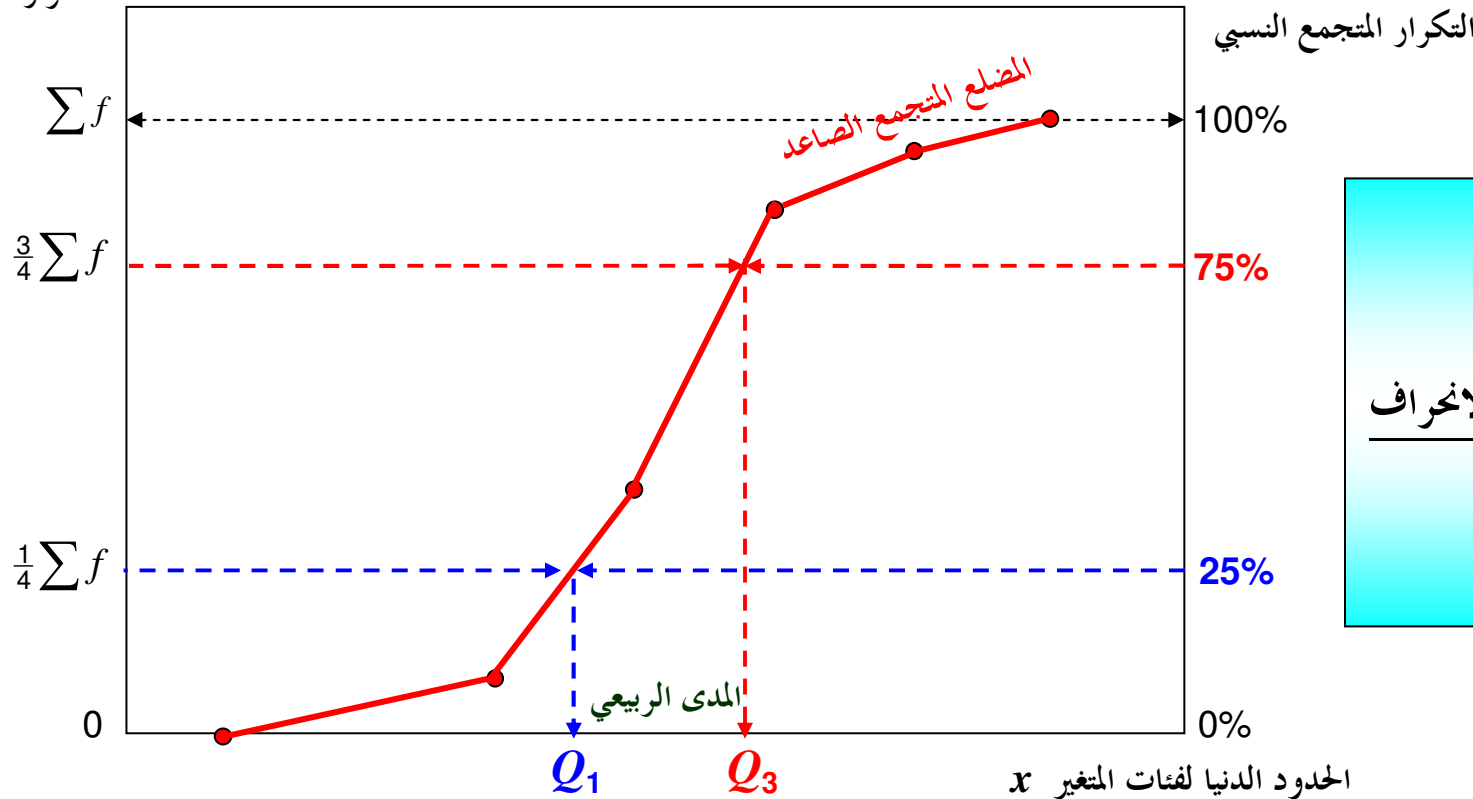


ويكن تحديد الربعين Q_1 (الأول) ، Q_3 (الثالث) بنفس الطريقة التي حددنا بها الوسيط M [الربع الثاني Q_2] ، إلا أننا سنكتفي هنا بتحديد Q_1 و Q_3 ثم تحديد نصف المدى الربيعي Q للبيانات الكمية المتصلة كالتالي :

على المضلع التكراري المتجمع الصاعد :

- حدد قيمة المتغير المناظرة لتكرار متجمع قدره $\frac{1}{4} \sum f$ أو تكرار متجمع نسبي قدره **25%** فتكون تلك القيمة هي Q_1 [الربع الأول] .
 حدد قيمة المتغير المناظرة لتكرار متجمع قدره $\frac{3}{4} \sum f$ أو تكرار متجمع نسبي قدره **75%** فتكون تلك القيمة هي Q_3 [الربع الثالث] .

التكرار المتجمع



ويكون المدى الربيعي هو :

$$Q_3 - Q_1$$

ونصف المدى الربيعي [أو الانحراف الربيعي] هو :

$$Q = \frac{1}{2} (Q_3 - Q_1)$$

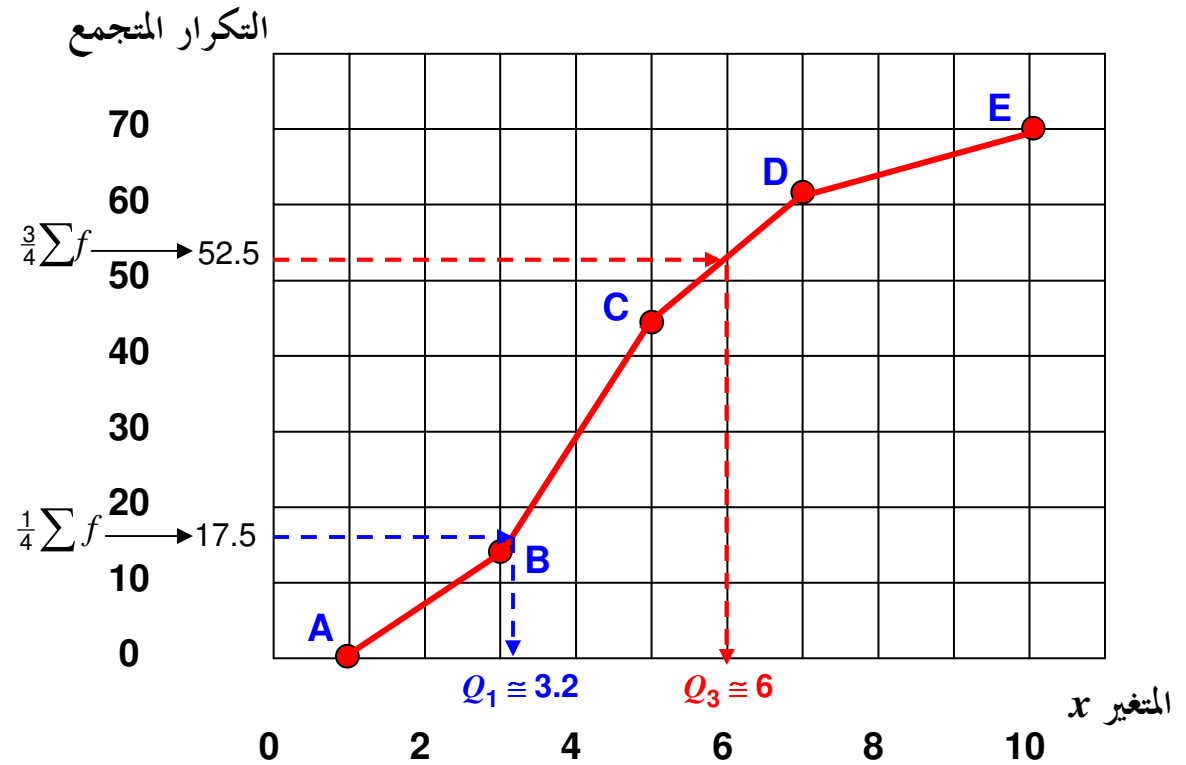
المتغير x	$1 \leq x < 3$	$3 \leq x < 5$	$5 \leq x < 7$	$7 \leq x < 10$
التكرار f	14	29	18	9

فمثلاً للتوزيع التكراري المبين :

- قم بتكوين الجدول التكراري المتجمع الصاعد
- قم برسم المضلع التكراري المتجمع الصاعد ومنه حدد قيم الربع الأول Q_1 والربع الثالث Q_3 بالطريقة المذكورة سابقاً

الجدول التكراري المتجمع الصاعد		
المتغير x	التكرار المتجمع	النقطة على المضلع
< 1	0	A (1 , 0)
< 3	14	B (3 , 14)
< 5	43	C (5 , 43)
< 7	61	D (7 , 61)
< 10	$\sum f = 70$	E (10 , 70)

ملحوظة : $\frac{1}{4} \sum f = 17.5$ ، $\frac{3}{4} \sum f = 52.5$



إذن المدى الربيعي هو : $Q_3 - Q_1 = 6 - 3.2 = 2.8$

ونصف المدى الربيعي [أو الانحراف الربيعي] هو : $Q = \frac{1}{2}(Q_3 - Q_1) = \frac{1}{2} \times 2.8 = 1.4$

خامساً : المدى المئيني : لمجموعة من البيانات يُعرف المدى المئيني [وسنرمز له بالرمز P] كالاتي :

$$P = P_{90} - P_{10}$$

المدى المئيني :

المئين العاشر المئين التسعون

ويفضل أيضاً استخدامه في الحالات التي يستعصي فيها حساب الانحراف المتوسط أو المعياري [مثل حالة التوزيعات التكرارية المفتوحة أو حالة وجود قيم متطرفة في البيانات]

س : ما هي المئينات ؟

ج : بنفس الطريقة التي تم بها تقسيم مجموعة من القيم إلى مجموعتين متساويتين في العدد [عن طريق الوسيط M] أو تقسيمها إلى أربعة مجموعات متساوية في العدد [عن طريق الربيعات Q_1, Q_2, Q_3] ، يمكن تقسيم مجموعة القيم إلى 100 مجموعة متساوية في العدد عن طريق قيم عددها 99 سنرمز لها بالرموز :

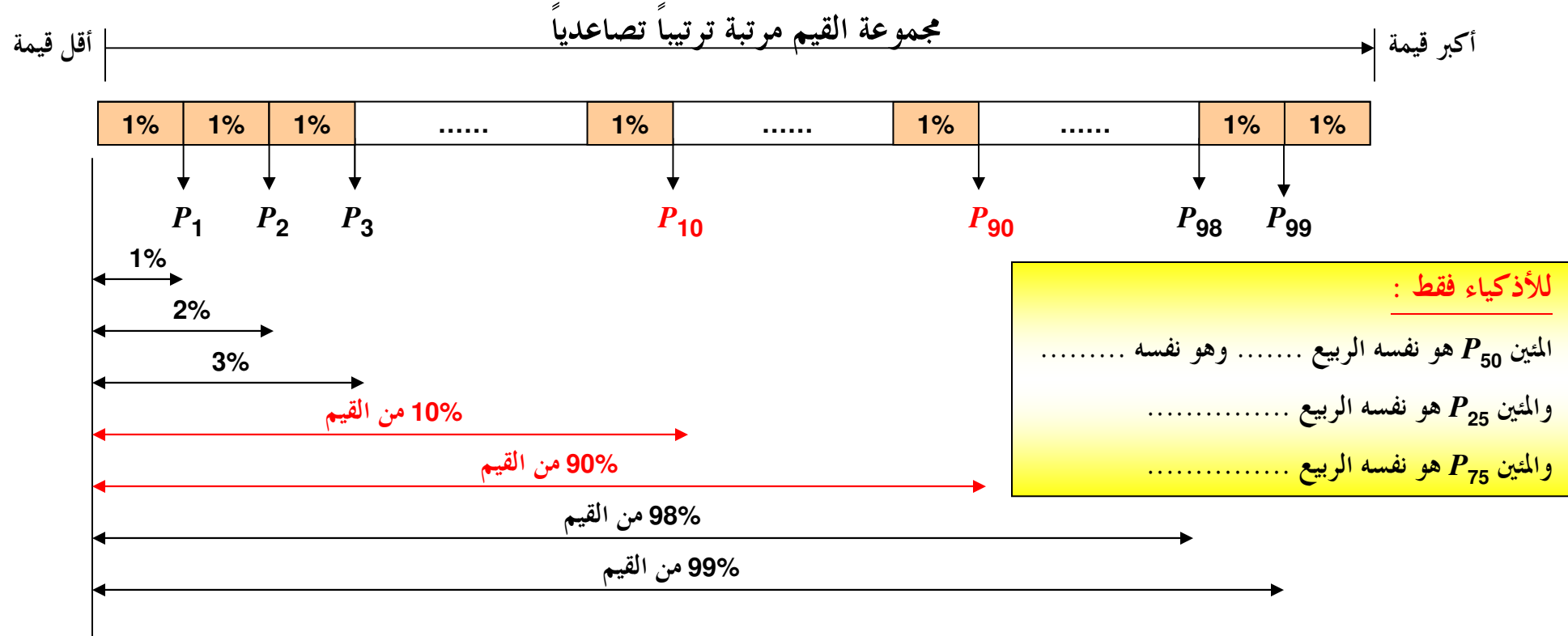
$$P_1, P_2, \dots, P_{10}, \dots, P_{90}, \dots, P_{98}, P_{99}$$

تُسمى هذه القيم بـ المئينات ، حيث :

P_1 [المئين الأول] : هو قيمة يقع تحتها 1% من مجموع القيم [وبالطبع يقع فوقها 99% من القيم]

P_2 [المئين الثاني] : هو قيمة يقع تحتها 2% من مجموع القيم [وبالطبع يقع فوقها 98% من القيم]





P_{10} [المئين العاشر] : هو قيمة يقع تحتها 10% من مجموع القيم [وبالطبع يقع فوقها 90% من القيم]

P_{90} [المئين التسعون] : هو قيمة يقع تحتها 90% من مجموع القيم [وبالطبع يقع فوقها 10% من القيم]

وهكذا

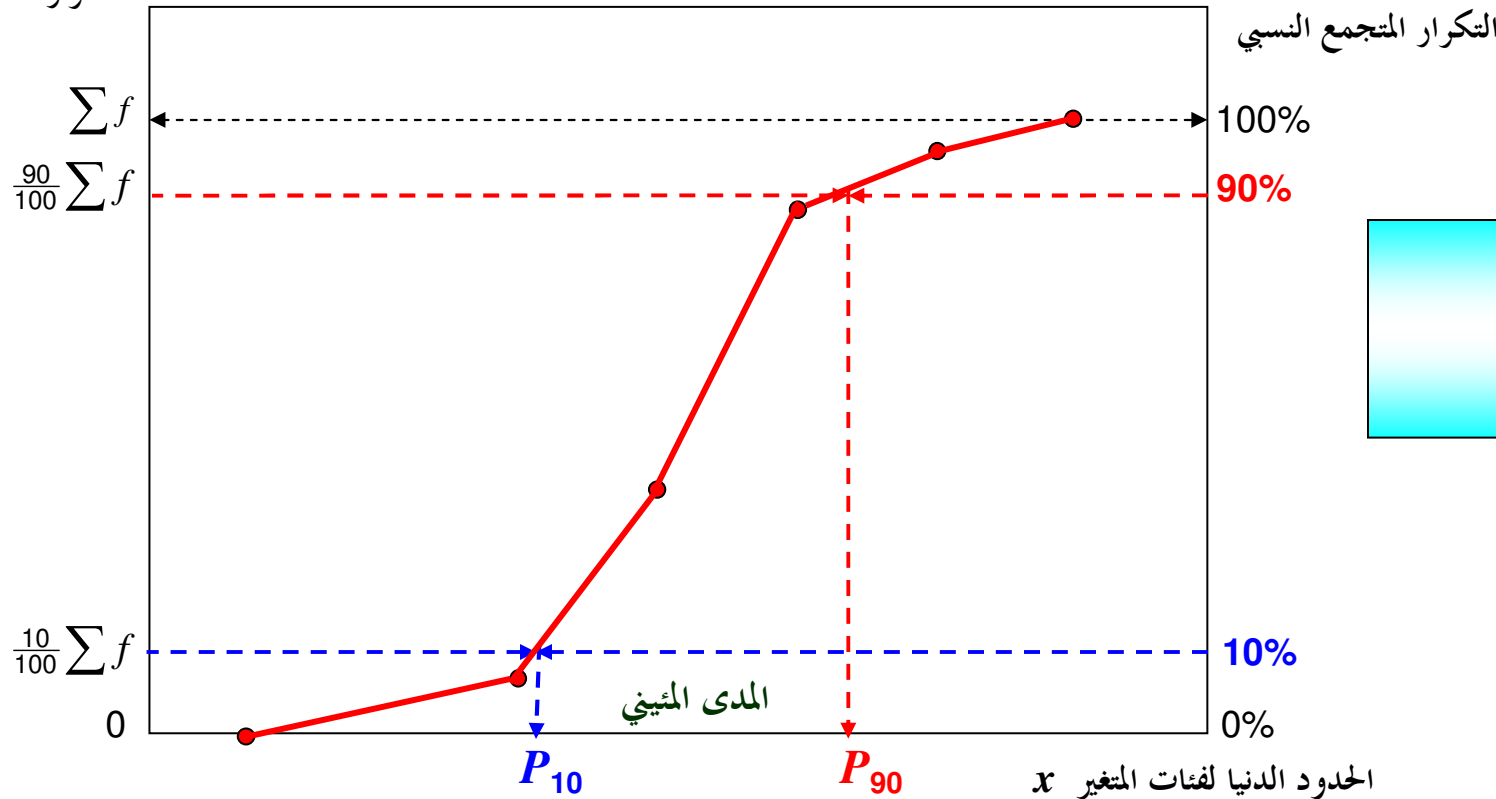
ويكن تحديد المئينات P_{10} (العاشر) ، P_{90} (التسعون) بنفس الطريقة التي حددنا بها الوسيط M والربيعات ، إلا أننا سنكتفي هنا بتحديد P_{10} و P_{90} [ومن ثم تحديد المدى المئيني P] للبيانات الكمية المتصلة كالتالي :

على المصطلح التكراري المتجمع الصاعد :

حدد قيمة المتغير المناظرة لتكرار متجمع قدره $\sum f$ أو $\frac{10}{100} \sum f$ [تكرار متجمع نسبي قدره 10%] فتكون تلك القيمة هي P_{10} [المئين العاشر] .

حدد قيمة المتغير المناظرة لتكرار متجمع قدره $\sum f$ أو $\frac{90}{100} \sum f$ [تكرار متجمع نسبي قدره 90%] فتكون تلك القيمة هي P_{90} [المئين التسعون] .

التكرار المتجمع



ويكون المدى المئيني هو :

$$P = P_{90} - P_{10}$$

المتغير x	$1 \leq x < 3$	$3 \leq x < 5$	$5 \leq x < 7$	$7 \leq x < 10$
التكرار f	14	29	18	9

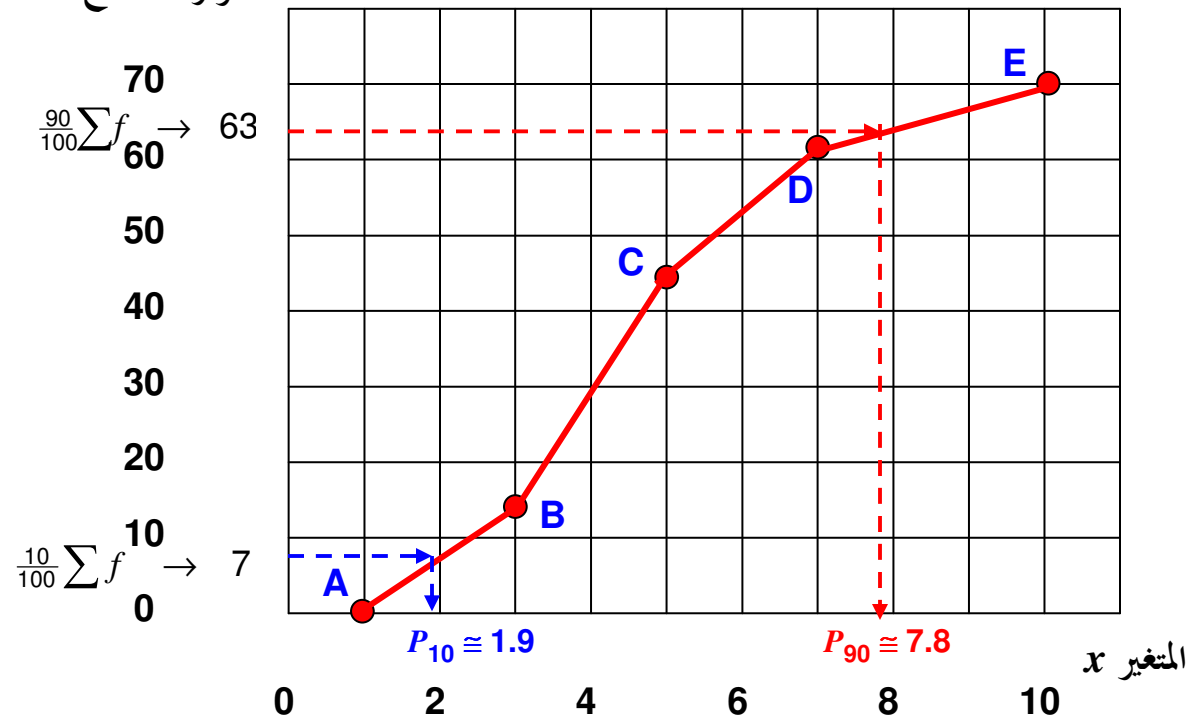
فمثلاً للتوزيع التكراري المبين :

- قم بتكوين الجدول التكراري المتجمع الصاعد
- قم برسم المضلع التكراري المتجمع الصاعد ومنه حدد قيم المئين العاشر P_{10} والمئين التسعين P_{90} بالطريقة المذكورة سابقاً

الجدول التكراري المتجمع الصاعد		
المتغير x	التكرار المتجمع	النقطة على المضلع
< 1	0	A (1 , 0)
< 3	14	B (3 , 14)
< 5	43	C (5 , 43)
< 7	61	D (7 , 61)
< 10	$\sum f = 70$	E (10 , 70)

ملحوظة : $\frac{10}{100} \sum f = 7$ ، $\frac{90}{100} \sum f = 63$

التكرار المتجمع



إذن المدى المئيني هو : $P = P_{90} - P_{10} \cong 7.8 - 1.9 = \underline{\underline{6.9}}$

علاقات اعتبارية بين مقاييس التشتت

في التوزيعات متوسطة الالتواء هناك علاقات اعتبارية (تقريبية) بين مقاييس التشتت السابقة كالآتي :

$$s = \frac{5}{4} M.D \quad \text{الانحراف المعياري} = \frac{5}{4} \times \text{الانحراف المتوسط}$$

$$\text{أو} \quad M.D = \frac{4}{5} s$$

$$\text{الانحراف المتوسط} = \frac{4}{5} \times \text{الانحراف المعياري}$$

$$s = \frac{3}{2} Q \quad \text{الانحراف المعياري} = \frac{3}{2} \times \text{الانحراف الربيعي}$$

$$\text{أو} \quad Q = \frac{2}{3} s$$

$$\text{الانحراف الربيعي} = \frac{2}{3} \times \text{الانحراف المعياري}$$

$$Q = \frac{5}{6} M.D \quad \text{الانحراف الربيعي} = \frac{5}{6} \times \text{الانحراف المتوسط}$$

$$\text{أو} \quad M.D = \frac{6}{5} Q$$

$$\text{الانحراف المتوسط} = \frac{6}{5} \times \text{الانحراف الربيعي}$$

هذه العلاقات الاعتبارية تمكننا من حساب قيم تقريبية لبقية مقاييس التشتت متى عُلم أحدها [وذلك في حالة صلاحيتها .. أي في حالة التوزيعات التكرارية متوسطة الالتواء]

$$M.D = \frac{4}{5} s = \frac{4}{5} \times 30 = \underline{\underline{24}} \quad , \quad Q = \frac{2}{3} s = \frac{2}{3} \times 30 = \underline{\underline{20}}$$

فمثلاً : • إذا كان $s = 30$ فإن :

$$M.D = \frac{6}{5} Q = \frac{6}{5} \times 20 = \underline{\underline{24}} \quad , \quad s = \frac{3}{2} Q = \frac{3}{2} \times 20 = \underline{\underline{30}}$$

• وإذا كان $Q = 20$ فإن :

$$s = \frac{5}{4} M.D = \frac{5}{4} \times 24 = \underline{\underline{30}} \quad , \quad Q = \frac{5}{6} M.D = \frac{5}{6} \times 24 = \underline{\underline{20}}$$

• وإذا كان $M.D = 24$ فإن :

التشتت النسبي ومقاييسه

التغير الفعلي أو التشتت كما نحصل عليه من الانحراف المتوسط أو المعياري أو الربيعي أو غيره من مقاييس التشتت يُسمى **بالتشتت المطلق** ، ولكن تشتت قدره 10 درجات عن قيمة متوسطة 50 درجة (مثلاً) يختلف عن تشتت قدره 10 درجات عن قيمة متوسطة 200 ، لذا من المناسب تعريف ما يُسمى بـ **التشتت النسبي** وهو :

$$\text{التشتت النسبي (كنسبة مئوية)} = \frac{\text{التشتت المطلق}}{\text{المتوسط}} \times 100$$

وبالتالي يكون التشتت النسبي لتشتت مطلق قدره 10 لبيانات متوسطة 50 : $\frac{10}{50} \times 100 = 20\%$

أما التشتت النسبي لتشتت مطلق قدره 10 لبيانات متوسطة 200 فهو : $\frac{10}{200} \times 100 = 5\%$

ومن أكثر مقاييس التشتت النسبي استخداماً ما يُسمى **بمعامل الاختلاف** [أو **معامل التشتت**] و **معامل الاختلاف الربيعي** ، حيث :

$$\text{معامل الاختلاف الربيعي} = 100 \times \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

$$\frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100$$

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

$$\frac{s}{\bar{x}} \times 100$$

المتغير x	التكرار f
$6 \leq x < 10$	8
$10 \leq x < 12$	20
$12 \leq x < 14$	12
$14 \leq x < 18$	10

فعلي سبيل المثال ، إذا كانت لدينا البيانات الموضحة بالجدول المقابل عن توزيع الوحدات السكنية حسب الإيجار السنوي بأحد الأحياء $[x$ يمثل الإيجار بالآلاف ريال ، f يمثل عدد الوحدات السكنية] ، وكان مطلوباً تحديد كلٍّ من معامل الاختلاف للإيجار و معامل الاختلاف الربيعي له .

أولاً : بالنسبة لمعامل الاختلاف : لابد أولاً من تحديد كلٍّ من الوسط الحسابي والانحراف المعياري [فرصة للمراجعة]

المتغير x	التكرار f	x_0	fx_0	$d = x_0 - \bar{x}$	d^2	fd^2
$6 \leq x < 10$	8	8	64	$8 - 12 = -4$	16	128
$10 \leq x < 12$	20	11	220	$11 - 12 = -1$	1	20
$12 \leq x < 14$	12	13	156	$13 - 12 = 1$	1	12
$14 \leq x < 18$	10	16	160	$16 - 12 = 4$	16	160
	50		600			320
	$\sum f$		$\sum fx_0$			$\sum fd^2$

$$\bar{x} = \frac{\sum fx_0}{\sum f} = \frac{600}{50} = 12 \quad s^2 = \frac{\sum fd^2}{\sum f} = \frac{320}{50} = 6.4 \rightarrow s = \sqrt{s^2} = \sqrt{6.4} \approx 2.53$$

وبالتالي يكون معامل الاختلاف $21.1\% \approx 100 \times \frac{2.53}{12} = 100 \times \frac{s}{\bar{x}}$ ، أي أن الإيجار يتغير بنسبة ٢١,١ % تقريباً

ثانياً : بالنسبة لمعامل الاختلاف الربيعي : لابد أولاً من تحديد الربيعين الأول والثالث [فرصة للمراجعة]



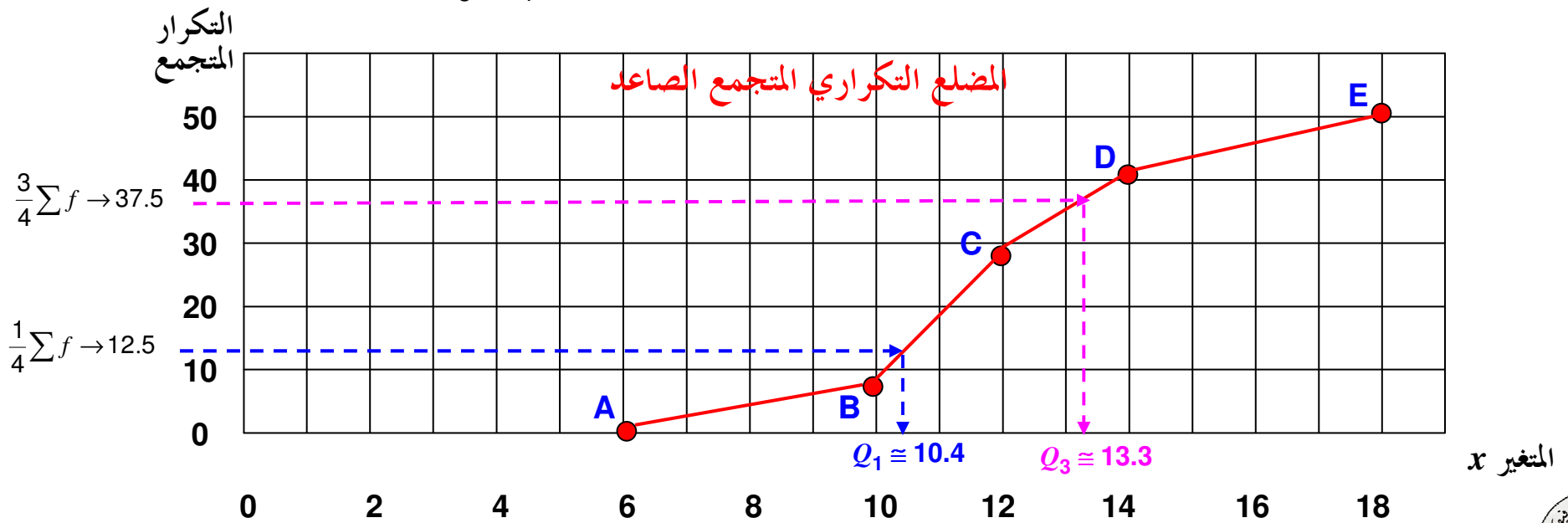
المتغير x	التكرار f
$6 \leq x < 10$	8
$10 \leq x < 12$	20
$12 \leq x < 14$	12
$14 \leq x < 18$	10



الجدول التكراري المتجمع الصاعد		
المتغير x	التكرار المتجمع	النقطة على المضلع
< 6	0	A (6 , 0)
< 10	8	B (10 , 8)
< 12	28	C (12 , 28)
< 14	40	D (14 , 40)
< 18	$\sum f = 50$	E (18 , 50)

- قم بتكوين الجدول التكراري المتجمع الصاعد
 - قم برسم المضلع التكراري المتجمع الصاعد
- ومنه حدد قيم الربع الأول Q_1 والربع الثالث Q_3 .

$$\frac{Q_3 - Q_1}{Q_3 + Q_1} \times 100 = \frac{13.3 - 10.4}{13.3 + 10.4} \times 100 = \frac{2.9}{23.7} \times 100 \cong \underline{\underline{12.2\%}}$$



الدرجات المعيارية

لمجموعة من القيم $x_1, x_2, \dots, x_n$ وسطها الحسابي $\bar{x}$ وانحرافها المعياري s تسمى :

$$z = \frac{x - \bar{x}}{s}$$

بـ الدرجة المعيارية للقيمة x .

فمثلاً لمجموعة القيم 8 3 4 12 6 7 4 8 3 5 [سؤال سلمي نفسك/المحاضرة ١٠/شريحة ١٦ والذي قمنا بحله في بداية هذه المحاضرة/شريحة ٤] ، كان الوسط الحسابي 6 والانحراف المعياري 2.76 . إذن الدرجات المعيارية لهذه القيم هي :

5	3	8	4	7	6	12	4	3	8	القيم
---	---	---	---	---	---	----	---	---	---	-------

-0.36	-1.09	0.72	-0.72	0.36	0	2.17	-0.72	-1.09	0.72	الدرجات المعيارية للقيم
$\frac{5-6}{2.76}$	$\frac{3-6}{2.76}$	$\frac{8-6}{2.76}$	$\frac{4-6}{2.76}$	$\frac{7-6}{2.76}$	$\frac{6-6}{2.76}$	$\frac{12-6}{2.76}$	$\frac{4-6}{2.76}$	$\frac{3-6}{2.76}$	$\frac{8-6}{2.76}$	

وللدرجات المعيارية للقيم أهمية كبيرة في مقارنة نتائج بيانات مختلفة ببعضها حيث قد يؤدي الاعتماد على القيم الحقيقية إلى استنتاجات غير سليمة أو مضللة . لتوضيح ذلك دعنا نعتبر المثال التالي :

في الاختبار النهائي لمقرر الإحصاء حصل طالب على **82** درجة [حيث كان الوسط الحسابي للدرجات **76** بانحراف معياري **10**] وحصل في مقرر الصحة واللياقة على **90** درجة [حيث كان الوسط الحسابي للدرجات **82** بانحراف معياري **16** . هل يمكن القول أن الطالب درجة استيعابه لمقرر الصحة واللياقة كانت أعلى من درجة استيعابه لمقرر الإحصاء ؟

الاعتماد على درجات الطالب في المقررين [82 في الإحصاء ، 90 في الصحة واللياقة] تجعل الإجابة : نعم درجة استيعاب الطالب لمقرر الصحة واللياقة أعلى من درجة استيعابه لمقرر الإحصاء .

ولكن الإجابة الصحيحة تعتمد على الدرجة المعيارية للطالب في كل من المقررين :

في مقرر الإحصاء	في مقرر الصحة واللياقة
$x=84$, $\bar{x}=76$, $s=10$	$x=90$, $\bar{x}=82$, $s=16$
$\therefore z = \frac{x - \bar{x}}{s} = \frac{84 - 76}{10} = \frac{8}{10} = \underline{0.8}$	$\therefore z = \frac{x - \bar{x}}{s} = \frac{90 - 82}{16} = \frac{8}{16} = \underline{0.5}$

أي أن الدرجة المعيارية للطالب في مقرر الإحصاء أعلى من نظيرتها في مقرر الصحة واللياقة ، مما يعني أن درجة استيعاب الطالب لمقرر الإحصاء أعلى من درجة استيعابه لمقرر الصحة واللياقة .

سلي نفسك لغاية ما نتقابل بإذن الله

الوزن x	$x < 50$	$50 \leq x < 60$	$60 \leq x < 70$	$70 \leq x < 80$	$x \geq 80$
العدد f	14	29	18	9	9

(١) البيانات الموضحة بالجدول المقابل تعبر عن أوزان مجموعة من

الطلبة (بالكيلوجرام) في المرحلة الجامعية . المطلوب حساب مقياس مناسب للترعة المركزية وآخر للتشتت ، ثم أوجد مقياساً لمعامل الاختلاف .

(٢) حصل أحد الطلاب في مقرر المحاسبة على درجة ٨٠ في الاختبار النهائي وعلى درجة ٧٠ في مقرر الرياضيات . هل يمكن القول بأن درجة

استيعاب الطالب لمادة المحاسبة أفضل من درجة استيعابه لمادة الرياضيات علماً بأن الوسط الحسابي لدرجات الطلاب في المادتين هو ٨٣ [في المحاسبة] ، ٦٥ [في الرياضيات] بانحراف معياري قدره ٥ في المادتين .



مَشْرِقٌ
بِحَمْدِ اللَّهِ

