اسم المقرر

الإحصاء الاجتماعي

أستاذ المقرر

د. سعيد سيف الدين

عمادة التعلم الإلكتروني والتعليم عن بُعد

جامعة الملك فيصل

عمادة التعلم الإلكتروني والتعليم عن بعد

كلية الآداب

المحاضرة التاسعة
مقاييس التشتت (تابع)

١. التباين والانحراف المعياري
٢. معامل الاختلاف (معامل التشتت)
٣. الانحراف الربيعي [نصف المدى الربيعي]

١. التباين والانحراف المعياري

يُعرف متوسط مربعات الانحرافات عن الوسط الحسابي على أنه تباين مجموعة البيانات [ويُرمز له بالرمز s^2] ، ويُعرف الجذر التربيعي للتباين على أنه الانحراف المعياري للبيانات [ويُرمز له بالرمز s] ، أي أن :

التباين $= \dfrac{\sum d^2}{n} = s^2$ ← ومنه يكون ← الانحراف المعياري $= \sqrt{\dfrac{\sum d^2}{n}} = \sqrt{s^2} = s$

فمثلاً لمجموعة القيم ١٥ ٣ ٧ ٦ ١٢ ١٨ ٥ ٣ ١٣ ١٥ (والتي سبق وحسبنا لها لانحراف المتوسط وكان **4.9**) يمكننا حساب التباين لها والانحراف المعياري كالتالي :

x	$d = x - \bar{x}$	d^2
15	15 – 9.7 = 5.3	28.09
13	13 – 9.7 = 3.3	10.89
3	3 – 9.7 = -6.7	44.89
5	5 – 9.7 = -4.7	22.09
18	18 – 9.7 = 8.3	68.89
12	12 – 9.7 = 2.3	5.29
6	6 – 9.7 = -3.7	13.69
7	7 – 9.7 = -2.7	7.29
3	3 – 9.7 = -6.7	44.89
15	15 – 9.7 = 5.3	28.09
$\sum x$ = 97		$\sum d^2$ = 274.1

$$\bar{x} = \frac{\sum x}{n} = \frac{97}{10} = 9.7$$

التباين $= 27.41 = \dfrac{274.1}{10} = \dfrac{\sum d^2}{n} = s^2$

الانحراف المعياري $= \underline{\underline{5.24}} \cong \sqrt{27.41} = \sqrt{s^2} = s$

المتغير x	التكرار f
4	20
5	40
6	30
7	10

التوزيع التكراري

وفي حالة توزيع تكراري كما هو مبين (بيانات منفصلة) ، يمكن تحديد التباين والانحراف المعياري من العلاقات :

$$s^2 = \frac{\sum fd^2}{\sum f} = \text{التباين}$$

ومنه يكون

$$s = \sqrt{s^2} = \sqrt{\frac{\sum fd^2}{\sum f}} = \text{الانحراف المعياري}$$

فمثلاً ، للتوزيع التكراري المبين يكون :

إتجاه الحل

الجدول التكراري

المتغير x	التكرار f	fx
4	20	80
5	40	200
6	30	180
7	10	70
	100	530

$\sum f = 100$ $\quad \sum fx = 530$

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{530}{100} = 5.3$$

من هنا بداية الحل

$d = x - \bar{x}$	d^2	fd^2
4 – 5.3 = -1.3	1.69	20 × 1.69 = 33.8
5 – 5.3 = -0.3	0.09	40 × 0.09 = 3.6
6 – 5.3 = 0.7	0.49	30 × 0.49 = 14.7
7 – 5.3 = 1.7	2.89	10 × 2.89 = 28.9
		81

$\sum fd^2$

$$s^2 = \frac{\sum fd^2}{\sum f} = \frac{81}{100} = 0.81 = \text{التباين}$$

$$\boxed{s = \sqrt{s^2} = \sqrt{0.81} = \underline{\underline{0.9}} =} \text{ الانحراف المعياري}$$

خاص بحساب الوسط الحسابي

وهذا الجزء يُضاف إذا كان مطلوباً حساب الانحراف المعياري

الفئة	المتغير x	التكرار f
الأولى	$0 \le x < 20$	4
الثانية	$20 \le x < 30$	16
الثالثة	$30 \le x < 35$	12
الرابعة	$35 \le x < 40$	10
الخامسة	$40 \le x < 50$	6
السادسة	$50 \le x < 60$	2

أما في حالة البيانات الكمية المتصلة (كما في التوزيع التكراري المبين) ، تُستخدم نفس العلاقات السابقة لتحديد التباين والانحراف المعياري ، أي يكون :

$$\text{التباين} = s^2 = \frac{\sum fd^2}{\sum f}$$

ومنه يكون

$$\text{الانحراف المعياري} = \sqrt{\frac{\sum fd^2}{\sum f}} = \sqrt{s^2} = s$$

حيث $d = x_0 - \bar{x}$ ، x_0 تمثل مراكز الفئات . فمثلاً ، للتوزيع التكراري المبين ، يكون :

إتجاه الحل

من هنا بداية الحل

الفئة	المتغير x	التكرار f	المركز x_0	fx_0	$d = x_0 - \bar{x}$	d^2	$f \times d^2$
الأولى	$0 \le x < 20$	4	10	40	10 – 31.7 = -21.7	470.89	4 × 470.89 = 1883.56
الثانية	$20 \le x < 30$	16	25	400	25 – 31.7 = -6.7	44.89	16 × 44.89 = 718.24
الثالثة	$30 \le x < 35$	12	32.5	390	32.5 – 31.7 = 0.8	0.64	12 × 0.64 = 7.68
الرابعة	$35 \le x < 40$	10	37.5	375	37.5 – 31.7 = 5.8	33.64	10 × 33.64 = 336.4
الخامسة	$40 \le x < 50$	6	45	270	45 – 31.7 = 13.3	176.89	6 × 176.89 = 1061.34
السادسة	$50 \le x < 60$	2	55	110	55 – 31.7 = 23.3	542.89	2 × 542.89 = 1085.78
		50		1585			$\sum fd^2 = 5093$
		$\sum f$		$\sum fx_0$			

$$\therefore \quad \bar{x} = \frac{\sum f\, x_0}{\sum f} = \frac{1585}{50} = \underline{\underline{31.7}}$$

$$\text{التباين} = s^2 = \frac{\sum fd^2}{\sum f} = \frac{5093}{50} = 101.86$$

$$\text{الانحراف المعياري} = s = \sqrt{s^2} = \sqrt{101.86} \cong \underline{\underline{10.09}}$$

من السابق يتضح أن كلاً من الانحراف المتوسط والانحراف المعياري يعتمدا تماماً في حساباتهما على الوسط الحسابي ، وبالتالي فلهما نفس مزايا وعيوب الوسط الحسابي . أي :

المزايا : من السهل حسابهما – يأخذا في الاعتبار جميع البيانات – لا يحتاجا لترتيب معين للبيانات

العيوب : يتأثرا بشدة بالقيم المتطرفة – لا يمكن حسابهما للتوزيعات التكرارية المفتوحة

---

٢. معامل الاختلاف (معامل التشتت)

التغير الفعلي أو التشتت كما نحصل عليه من الانحراف المتوسط أو المعياري 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 . ومن أكثر مقاييس التشتت النسبي استخداماً ما يُسمى بمعامل الاختلاف [أو معامل التشتت] حيث :

$$\text{معامل الاختلاف (كنسبة مئوية)} = \frac{\text{الانحراف المعياري}}{\text{الوسط الحسابي}} \times 100 = \frac{s}{\bar{x}} \times 100$$

وبالتالي يكون معامل الاختلاف لبيانات انحرافها المعياري 10 ووسطها الحسابي 50 يساوي :

$$\frac{10}{50} \times 100 = 20\%$$

أما معامل الاختلاف لبيانات انحرافها المعياري 10 ووسطها الحسابي ٢٠٠ يساوي :

$$\frac{10}{200} \times 100 = 5\%$$

وبالتالي تكون البيانات الأولى أكثر تشتتاً من البيانات الثانية (لأن معامل الاختلاف لها أكبر)

تلخيص لكيفية حساب كلٍ من الانحراف المتوسط والانحراف المعياري

- **بيانات منفصلة (قيم مفردة)**

الوسط الحسابي = $\bar{x} = \dfrac{\sum x}{n}$

الانحراف المتوسط = $M.D = \dfrac{\sum |d|}{n}$

التباين= $s^2 = \dfrac{\sum d^2}{n}$ $\longrightarrow$ الانحراف المعياري = $s = \sqrt{s^2}$

قيم عددها n	الانحرافات عن الوسط	القيم المطلقة للانحرافات	مربع الانحرافات
x	$d = x - \bar{x}$	$\|d\|$	d^2
...	...	...	...
...	...	...	...
$\sum x$		$\sum \|d\|$	$\sum d^2$

- **بيانات منفصلة (جدول تكراري)**

$$\bar{x} = \frac{\sum fx}{\sum f} \quad , \quad M.D = \frac{\sum f|d|}{\sum f}$$

$$s^2 = \frac{\sum fd^2}{\sum f} \rightarrow s = \sqrt{s^2}$$

القيم	التكرار		الانحرافات عن الوسط	القيم المطلقة للانحرافات	مربع الانحرافات		
x	f	fx	$d = x - \bar{x}$	$\|d\|$	d^2	$f\|d\|$	fd^2
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
	$\sum f$	$\sum fx$				$\sum f\|d\|$	$\sum fd^2$

- **بيانات متصلة**

$$\bar{x} = \frac{\sum fx}{\sum f}$$

$$M.D = \frac{\sum f|d|}{\sum f}$$

$$s^2 = \frac{\sum fd^2}{\sum f} \rightarrow s = \sqrt{s^2}$$

الفئات	التكرار			الانحرافات عن الوسط	القيم المطلقة للانحرافات	مربع الانحرافات		
x	f	مراكز الفئات	fx_0	$d = x_0 - \bar{x}$	$\|d\|$	d^2	$f\|d\|$	fd^2
...	...	x_0	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...
	$\sum f$	...	$\sum fx$				$\sum f\|d\|$	$\sum fd^2$

٣. الانحراف الربيعي [نصف المدى الربيعي] :

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

Q [نصف المدى الربيعي] الانحراف الربيعي $\frac{1}{2}(Q_3 - Q_1)$

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]

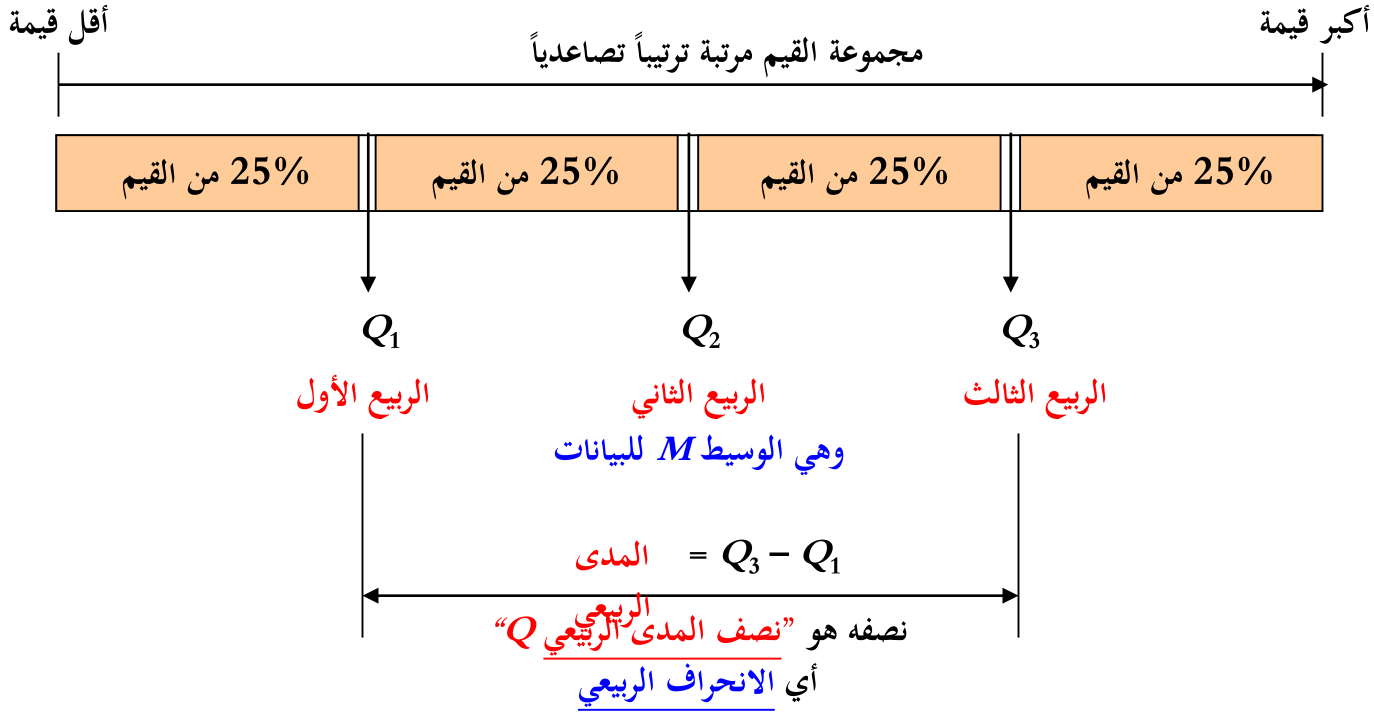
وفي بعض الأحيان يُستخدم المدى الربيعي $Q_3 - Q_1$ كمقياس للتشتت بدلاً من نصف المدى الربيعي

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

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

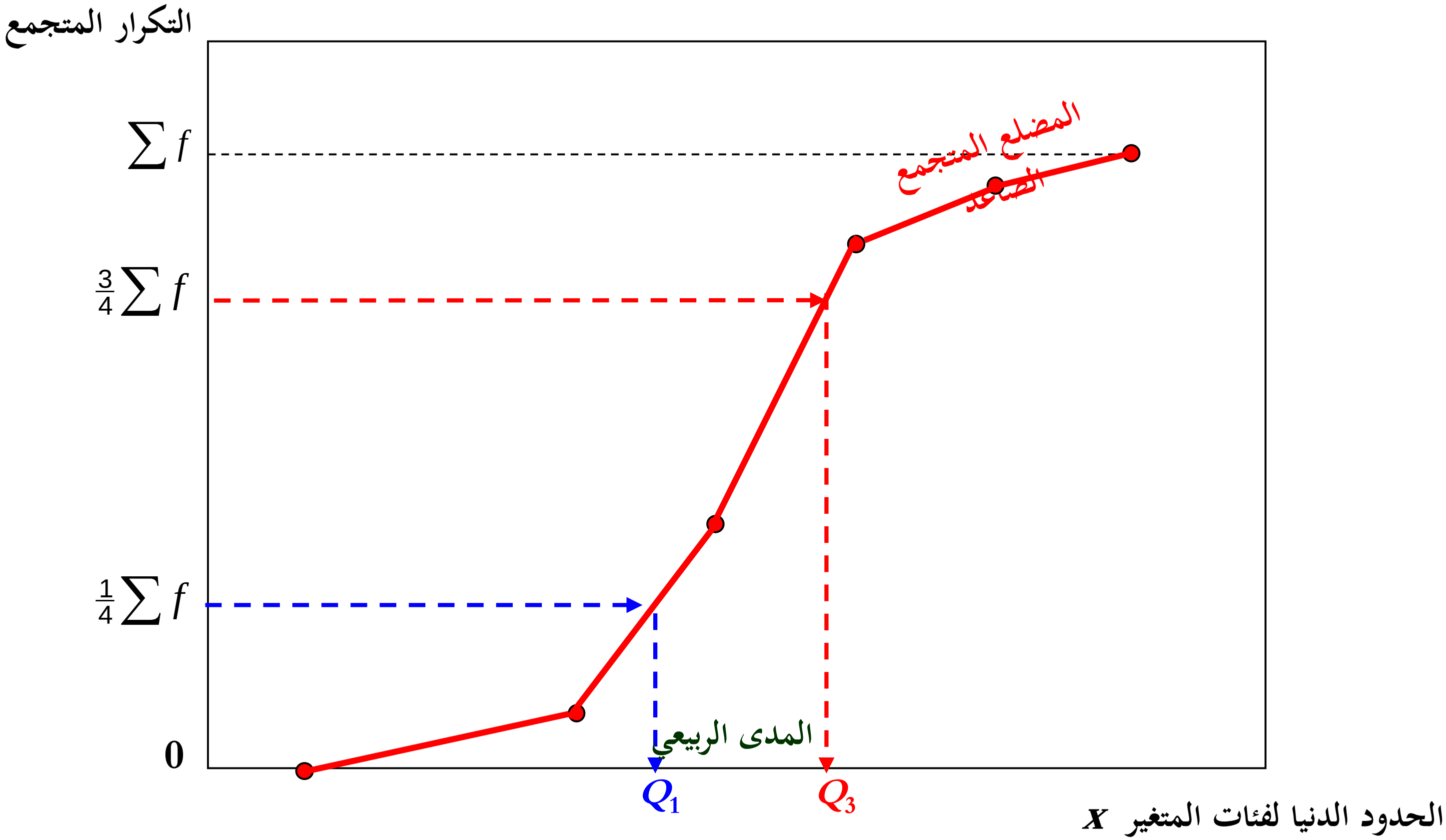
Q_3 [الربيع الثالث] هي تلك القيمة التي يقع تحتها 75% من القيم [وبالطبع فوقها 25% من القيم]

ويكن تحديد الربيعين Q_1 (الأول) ، Q_3 (الثالث) بنفس الطريقة التي حددنا بها الوسيط M [الربيع الثاني Q_2] ، إلا أننا سنكتفي هنا بتحديدها [ومن ثم تحديد نصف المدى الربيعي Q] للبيانات الكمية المتصلة وبالطريقة التخطيطية (الرسم) كالتالي :

على المضلع التكراري المتجمع الصاعد

- حدد قيمة المتغير المناظرة لتكرار متجمع قدره $\frac{1}{4}\sum f$ فتكون تلك القيمة هي Q_1 [الربيع الأول] .
- حدد قيمة المتغير المناظرة لتكرار متجمع قدره $\frac{3}{4}\sum f$ فتكون تلك القيمة هي Q_3 [الربيع الثالث] .

فيكون المدى الربيعي هو $\boxed{Q_3 - Q_1}$ ونصف المدى الربيعي [أو الانحراف الربيعي] هو $\boxed{Q = \frac{1}{2}(Q_3 - Q_1)}$

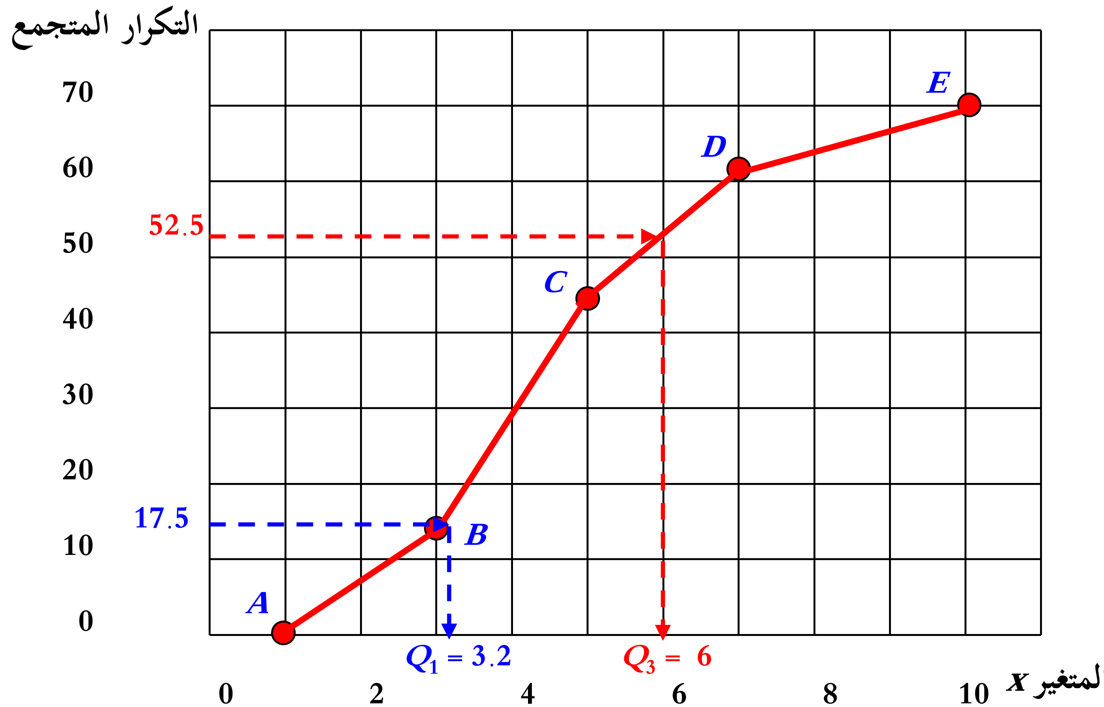


فمثلاً للتوزيع التكراري المبين :

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

- قم بتكوين الجدول التكراري المتجمع الصاعد
- قم برسم المضلع التكراري المتجمع الصاعد
- حدد قيم الربيع الأول Q_1 [وهي قيمة المتغير x المناظرة لتكرار متجمع صاعد قدره: $\frac{1}{4}\sum f = \frac{1}{4}\times 70 = 17.5$
- حدد قيم الربيع الأول Q_3 [وهي قيمة المتغير x المناظرة لتكرار متجمع صاعد قدره: $\frac{3}{4}\sum f = \frac{3}{4}\times 70 = 52.5$

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



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

ومنه يكون

الانحراف الربيعي هو : $Q = \frac{1}{2}(Q_3 - Q_1) = 1.4$

تمت
بحمد الله