4. إذا كانت لديك البيانات التالية:

(3, 13, 12, 7, 5) أثبت بأن مجموع انحرافاتها عن وسطها الحسابي يساوي صفراً.

5. إذا كان الوسط الحسابي لمجموعة من البيانات هو 15، وتم ضرب جميع القيم بالعدد 2 فما هو الوسط الحسابي الجديد بعد عملية الضرب.

6. أذكر خصائص المنوال.

7. إذا كان الوسط الحسابي لمجموعة من البيانات هو 60 والوسيط 56 فما حجم المنوال.

8. البيانات التالية تمثل علامات سبعة طلاب في مادة مبادئ الإحصاء والمطلوب حساب المئين 40 لهذه العلامات وهي كما يلي:

(95, 67, 72, 55, 90, 60, 80)

9. الجدول التالي يمثل اجور عمال لمصنع أدوية:

فئات الأجور	عدد العمال
199 – 150	15
249 – 200	10
299 – 250	8
349 – 300	5
399 – 350	16
449 – 400	2

المطلوب:

أ. حساب الوسيط. ب. حساب المئين 60.

ج. حساب Q_3 (P_{75}). د. حساب D_5 (P_{50}).

هـ. تمثيل هذه البيانات بيانياً بالمضلع التكراري.

الفصل الرابع

مقاييس التشتت

Measures of Dispersion

الفصل الرابع

مقاييس التشتت

Measures of Dispersion

4-1 مقدمة:

إن المقاييس السابقة ومنها مقاييس النزعة المركزية غير كافية في بعض الأوقات لتحديد صفات التوزيعات التكرارية والبيانات الإحصائية. عندما يكون هناك ظاهرتان متساويتان في مقاييس الموقع كالوسط الحسابي والوسيط إلا أن الظاهرتين مختلفتان، في هذه الحالات لا بد من استعمال مقاييس أخرى تبين مدى اختلاف البيانات فيما بينها ومدى التفاوت والتغير بين مفرداتها وهل هي متقاربة من بعضها البعض أم هي متباعدة وعليه تستطيع مقاييس التشتت الإجابة عن مثل هذه الأسئلة بمعنى آخر تستخدم مقاييس التشتت للحكم على دقة مقاييس النزعة المركزية ومدى الاعتماد عليه في وصف البيانات. فإذا كان لدينا مجموعتان من القيم، الأولى (5, 4, 7) والثانية (8, 3, 1) فإن الوسط الحسابي لكل مجموعة هو (4). فنلاحظ أن الوسط الحسابي يعتبر قيمة نموذجية للمجموعة الأولى ويمثل بدقة كل قيمة، بينما نجد في المجموعة الثانية أن وسطها الحسابي قيمة غير نموذجية ولا يمثل مفردات هذه المجموعة بدقة. لذلك فإن مقاييس التشتت هي التي تعطي وصفاً أفضل من المقاييس السابقة.

التشتت يعني تباعد القيم عن بعضها البعض.

ويمكن تعريف التشتت (الاختلاف) على أنه التباعد بين المشاهدات داخل المجموعة الواحدة. ومقاييس التشتت هي تلك المقاييس التي تقيس مدى التشتت للمشاهدات عن وسطها الحسابي، مما يعني كلما كان التشتت كبير كلما دل ذلك على أن الفروقات كبيرة بين المشاهدات وعدم تجانسها (غالباً عدم التجانس يكون عند جمع المشاهدات من مجتمعات مختلفة)، وعندما تكون مقاييس التشتت قليلة فهذا يعني أن الفروقات بين المشاهدات داخل المجموعة الواحدة قليلة وأنها متجانسة مع بعضها البعض.

ومن مقاييس التشتت المدى، الانحراف المعياري والتباين والانحراف المتوسط ومعامل الاختلاف.

4–2 **المدى** (Range):

المدى في البيانات هو الفرق بين أعلى قيمة وأصغر قيمة. ونلاحظ في هذا التعريف بأن المدى لا يعتمد على جميع البيانات بل على أكبرها وأصغرها فقط، مما يقلل من أهميته. فإذا كانت القيمتان متطرفتان (قيم شاذة) عند ذلك يكون المدى كبيراً بينما المفردات الأخرى للبيانات ليست متباعدة عن بعضها البعض.

مثال (1): أوجد المدى للبيانات التالية:

60, 30, 24, 16, 18, 12, 40

الحل:

نجد أكبر قيمة وأصغر قيمة فيكون المدى (R):

$$R = 60 - 12 = 48$$

مثال (2): أوجد المدى للبيانات التالية:

10, 100, 120, 130, 150, 1000

الحل:

نجد أكبر قيمة وأصغر قيمة:

$$R = 1000 - 10 = 990$$

ولكن هنا نلاحظ القيم المتطرقة (الشاذة) وهي أكبر قيمة وأصغر قيمة وبما أن المدى يعتمد عليها ينصح بحذف القيم المتطرفة الصغرى والكبرى. فيصبح المدى الجديد هو:

$$R = 150 - 100 = 50$$

4–3 **المدى الربيعي** (Interquartile Range):

الفرق بين الربيع الثالث (الأعلى) والربيع الأول (الأدنى) $Q_3 - Q_1$.

مثال (3): أوجد المدى الربيعي للبيانات التالية:

260, 190, 150, 160, 200, 170,180,220,250,210,280

الحل:

نرتب البيانات ترتيباً تصاعدياً ثم نوجد الربع الثالث والربيع الأول كما يلي:

150	160	170	180	190	200	210	220	250	260	280
1	2	3	4	5	6	7	8	9	10	11

ونجد أن رتبة الربع الأول:

$$Q_1 = P_{25} = \frac{25}{100} \times 12 = 3$$

$$\boxed{\text{Or } Q_1 = \frac{1}{4} \times n + 1 = 3}$$ القانون

إذاً الرتبة 3 يقابلها القيمة $Q_1 = 170$.

رتبة الربع الثالث (الأعلى) هي:

$$Q_3 = P75 = \frac{75}{100} \times 12 = 9$$

$$\text{Or } Q_3 = \frac{3}{4} \times 12 = 9$$

القيمة (value) 250.

إذاً المدى الربيعي:

$$Q3 - Q1$$

$$250 - 170 = 80$$

وعليه نصف المدى الربيعي يكون المدى الربيع/ 2

إذاً نصف المدى الربيعي:

$$\frac{250 - 170}{2} = 40$$

مثال (4): إذا كان الربع الثالث لمجموعة من البيانات يساوي 89 والربع الأول يساوي 39 وأعلى قيمة هي 100 وأدناها 40.

المطلوب: جد المدى ونصف المدى الربيعي.

الحل:

المدى = أعلى قيمة – أدنى قيمة

$$100 - 40 = 60$$

نصف المدى الربيعي:

$$= \frac{Q_3 - Q_1}{2}$$

$$= \frac{89 - 39}{2} = 25$$

4–4 وبنفس الإجراء نستطيع إيجاد التالي:

- المدى المئيني = المئين الأعلى – المئين الأدنى

$$Pa_9 - P_1$$

- نصف المدى المئيني = $\frac{\text{المدى المئيني}}{2}$

- المدى العشيري = العشر التاسع (الأعلى) – العشير الأول (الأدنى)

$$D_9 - D_1$$

- نصف المدى العشيري = $\frac{\text{المدى العشيري}}{2} = \frac{D_9 - D_1}{2}$

4–5 المدى في البيانات المبوبة:

نجد المدى من العلاقة التالية:

الحد الفعلي للفئة العليا – الحد الفعلي الأدنى للفئة الدنيا

أو:

المدى = مركز الفئة العليا – مركز الفئة الدنيا

6–4 المدى الربيعي ونصف المدى الربيعي:

نستخدم نفس الصيغ للقيم المبوبة مع الأخذ بعين الاعتبار طرق استخراجها.

× المدى الربيعي = $Q_3 - Q_1$

× نصف المدى الربيعي = $\frac{Q_3 - Q_1}{2}$

ولتوضيح ذلك نأخذ المثال التالي:

مثال (5): البيانات التالية تمثل أعمار (60) شخصاً:

الفئات الأعمار	22–26	27–31	32–36	37–41	42–46	47–51
تكرار الأشخاص	10	8	12	9	13	8

المطلوب: جد كل من: 1. المدى المطلق، 2. نصف المدى الربيعي.

الحل:

نكون جدول جديد كما يلي:

الفئات	التكرار	الحدود الفعلية	التكرار المتجمع الصاعد
22 – 26	10	21.5 – 26.5	10
27 – 31	8	26.5 – 31.5	18
32 – 36	12	31.5 – 36.5	30
37 – 41	9	36.5 – 41.5	39
42 – 46	13	41.5 – 46.5	52
47 – 51	8	46.5 – 51.5	60
المجموع	60		

1. المدى= الحد الفعلي الأعلى للفئة العليا – الحد الفعلي الأدنى للفئة الأخيرة

$51.5 - 21.5 = 30$

أو عن طريق مراكز الفئات:

$49 - 24 = 25$

2. نصف المدى الربيعي = $\frac{Q_3 - Q_1}{2}$

لهذا نجد المدى ومن ثم نصف المدى ونتبع الخطوات التالية:

نجد رتبة الربع الأول (الأدنى):

$Q_1 = \frac{25}{100} \times 60 = 15$

نحدد موقع الربع الأول في عمود التكرار الصاعد حيث يقع بين 10 و 18، نحدد الفئة الربيعية وهي: 26.5 – 31.5.

نحدد الحد الأدنى للفئة الربيعية (a): = 26.5

نجد الربع الأول من اللعلاقة التالية:

$$Q_1 = a + \frac{\frac{k}{100} \times n - n_1}{fm} \times C$$

$$Q_1 = 26.5 + \frac{15-10}{8} \times 5$$

$$= 26.5 + \frac{25}{8}$$

$$= 26.5 + 3.1 = 29.6$$

والآن نجد الربع الأعلى بنفس الأسلوب:

رتبة الربع الأعلى:

$$Q_3 = P_{75} = \frac{75}{100} \times 60 = 45$$

نحدد موقع الربع الأعلى في عمود التكرار الصاعد حيث يقع بين 39 و 52. نحدد الفئة الربيعية وهي: 46.5 – 41.5، ثم نحدد الحد الأدنى للفئة الربيعية (a) وهي 41.5 ونجد الربع الثالث (الأعلى) في نفس العلاقة أعلاه:

$$Q_3 = 41.5 + \frac{45-39}{52-39} \times 5$$

$$= 41.5 + \frac{6}{13} \times 5$$

$$= 41.5 + \frac{30}{13}$$

$$= 41.5 + 2.3 = 43.8$$

الآن نجد المدى الربيعي:

$$Q_3 - Q_1$$

$$43.8 - 29.6 = 14.2$$

ونجد نصف المدى الربيعي:

$$\frac{Q_3 - Q_1}{2}$$

$$\frac{14.2}{2} = 7.1$$

4–7 التباين والانحراف المعياري:

Variance and standard Deviation

التباين: مجموع تربيع انحرافات البيانات عن وسطها الحسابي مقسوم على عدد البيانات – 1.

ونرمز للتباين بالرمز (S^2) وتكون صيغة معادلته كما يلي:

$$S^2 = \frac{\sum (xi - \bar{x})^2}{n-1}$$

ونرمز للانحراف المعياري بالرمز (S):

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

أو:

$$S = \sqrt{\frac{\sum (xi - \bar{x})^2}{n-1}}$$

مثال (6): أوجد الانحراف المعياري من البيانات غير المبوبة التالية:

19, 10, 14, 6, 11

الحل:

1. نجد الوسط الحسابي:

$$\bar{x} = \frac{19+10+14+6+11}{5}$$

$$= \frac{60}{5} = 12$$

2. نجد أولاً التباين من خلال إيجاد فروق البيانات عن وسطها الحسابي ثـم نربعها ونجمعها ونقسمها على عدد البيانات مطروح منه (n − 1).

$$S^2 = \frac{(19-12)^2+(10-12)^2+(14-12)^2+(6-12)^2+(11-12)^2}{5-1}$$

$$= \frac{(7)^2+(-2)^2+(2)^2+(-6)^2+(-1)^2}{4}$$

$$= \frac{49+4+4+36+1}{4}$$

$S^2 = \frac{94}{4} = 23.5 =$ التباين

ونلاحظ أنه تم إزالة الإشارة السالبة بتربيع الانحرافات.

أما الانحراف المعياري $S = \sqrt{S^2}$ = Standard Deviation

$$S = \sqrt{23.5} = 4.84$$

ويمكن إيجاد التباين للبيانات غير المبوبة من الصيغة التالية:

$$S^2 = \frac{\sum xi^2 - n\bar{x}^2}{n-1}$$

$$= \frac{814 - 5\times(12)^2}{5-1}$$

$$= \frac{814 - 5\times 144}{4}$$

$$= \frac{814-720}{4}$$

$$= \frac{94}{4} = 23.5$$

$$S = \sqrt{S^2} = \sqrt{23.5} = 4.84$$

مثال (7): أوجد الانحـراف المعيـاري في القيـم المبوبـة مـن الجـدول الـذي يمثـل التوزيع التكراري لأوزان مائة مسافر.

حدود فئات الأوزان	10–15	16–21	22–27	28–33	34–39	40–45
تكرار المسافرين	20	18	16	14	12	20

الحل:

نستخدم صيغة المعادلة التالية:

$$S^2 = \frac{\sum\left(xi-\bar{x})^2.f\right)}{n-1} \text{S}$$

أو:

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

$$S = \sqrt{\frac{\sum(xi-\bar{x})^2.f}{n-1}}$$

وعليه نكون جدول جديد كما يلي:

الفئات	التكرار f	مراكز الفئة xi	xi.f	$xi-\bar{x}$	$(xi-\bar{x})^2$	$(xi-\bar{x})^2.f$
10 – 15	20	12.5	250	−14.4	207.36	4147.2
16 – 21	18	18.5	333	−8.4	70.56	1270.08
22 – 27	16	24.5	392	−2.4	5.76	92.16
28 – 33	14	30.5	427	3.6	12.96	181.44
34 – 39	12	36.5	438	9.6	92.16	1105.92
40 – 45	20	42.5	850	15.6	243.36	4867.2
المجموع	100		2690			11664

حيث أوجدنا مراكز الفئات(xi) وضربنها بالتكرار المقابل لها لكي نحسب الوسط الحسابي ثم أوجدنا انحرافات مراكز الفئات عن وسطها الحسابي وبعد ذلك نربع هذه الانحرافات ونضرب كل منها بالتكرار المقابل لها كما يظهر في الجدول السابق ونقسمها على مجموع التكرار مطروح منه راحد.

نجد الوسط الحسابي:

$$\bar{x}=\frac{2690}{100}=26.9$$

نطبق قانون التباين:

$$S^2=\frac{\sum(xi-\bar{x})^2.f}{n-1}$$

$$S^2=\frac{11664}{100-1}$$

$$=117.818$$

ثم نجد الانحراف المعياري حسب الصيغة التالية:

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

$$S=\sqrt{117.818}$$

$$=10.85$$

وهناك صيغة أخرى لحساب التباين والانحراف المعياري.

التباين للبيانات غير المبوبة:

$$S^2=\frac{\sum\left(xi^2-n\bar{x}^2\right)}{n-1}$$

	التكرار f	مراكز الفئة xi	xi.f	xi^2	xi^2f
	20	12.5	250	156.25	3125
	18	18.5	333	342.25	6160.5
	16	24.5	392	600.25	9604
	14	30.5	427	930.25	13023.5
	12	36.5	438	1332.25	15987
	20	42.5	850	1806.25	36125
المجموع	100		2690		84025

الوسط الحسابي هو:

$$\bar{x} = \frac{2690}{100} = 26.9$$

نعوض القانون ونجد التباين:

$$S^2 = \frac{84025 - 100 \times 26.9^2}{100 - 1}$$

$$S^2 = \frac{84025 - 100 \times 723.61}{99}$$

$$S^2 = \frac{84025 - 72361}{99}$$

$$S^2 = \frac{11664}{99} = 117.818$$

والانحراف المعياري:

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

$$S = \sqrt{117.818}$$

$$S = 10.85$$

4-8 الانحراف المتوسط The Mean Deviation:

هو متوسط الانحرافات المطلقة عن وسطها الحسابي ونلاحظ بأن الإشـارة السالبة تلغى بسبب القيمة المطلقة ويرمز لها بالرمز (M.D).

الانحراف المتوسط للبيانات الأولية (غير المبوبة):

نستخدم القانون التالي:

$$M.D = \frac{\sum |xi - \bar{x}|}{n}$$

حيث أن $|xi - \bar{x}|$ هي القيمة المطلقة. بمعنى قيمة موجبة للفروقات.

مثال: أجد الانحراف المتوسط للبيانات التاليةت:

12, 10, 8, 14, 16

الحل:

نوجد أولاً الوسط الحسابي ($\bar{x}$).

$$\bar{x} = \frac{12+10+8+14+16}{5}$$

$$\bar{x} = \frac{60}{5} = 12$$

ثم نجد القيمة المطلقة للفروق كما يلي:

$$M.D = \frac{|12-12| + |10-12| + |8-12| + |14-12| + |16-12|}{5}$$

$$= \frac{0+2+4+2+4}{5} = \frac{12}{5}$$

الانحراف المتوسط M.D = 2.4

الانحراف المتوسط في البيانات المجمعة (المبوبة):

ويأخذ الصيغة التالية:

$$M.D = \frac{\sum |xi - \bar{x}|.fi}{\sum fi}$$

مثال: أوجد الانحراف المتوسط للجدول التكراري التالي:

الفئات	التكرار f	مراكز الفئة xi	$xifi$	$\lvert xi-\bar{x}\rvert$	$\lvert xi-\bar{x}\rvert .f$
4 – 8	2	6	12	12.62	25.24
9 – 13	3	11	33	7.62	22.86
14 – 18	5	16	80	2.62	13.1
19 – 23	6	21	126	2.38	14.28
24 – 28	3	26	78	7.38	22.14
29 – 33	2	31	62	12.38	24.76
المجموع	21		391		122.38

$$\bar{x}=\frac{\sum xifi}{\sum fi}$$

$$\bar{x}.=\frac{391}{21}=18.62$$ تقريباً

ونطبق القانون:

$$M.D=\frac{122.38}{21}=5.83$$ تقريباً

4–9 معامل الاختلاف (التغير) :Coefficient of Variation

في بعض الأحيان يكون الانحراف المعياري وحده لا يكفي لإعطاء صورة واضحة عن التشتت للمجموعة من البيانات. وعلى هذا في أحياناً أخرى يكون

معامل الاختلاف ملائم لمنح الدارس صورة أفضل عن التشتت داخل مجموعة القيم أو البيانات. وبذلك فإن معامل الاختلاف أو التغير يعطي فكرة أفضل من الانحراف المعياري وغيره من مقاييس التشتت لمقارنة تشتت البيانات أو القيم بين عدة مجاميع منها.

ويستخدم معامل الاختلاف أو التغير لمقارنة التشتت داخل عدة مجاميع من البيانات وإن كانت وحدات القياس للمجاميع مختلفة.

وكلما كانت النسبة أقل كان التشتت أقل ويرمز له بالرمز (C.V) ويعطى بالصيغة التالية:

$$C.V=\frac{S}{\bar{x}}\times 100\%$$

مثال: إذا كانت أجور عمال في مصنعين متشابهين كما يلي المصنع الأول الوسط الحسابي للأجور هو (250) والمصنع الثاني الوسط الحسابي للأجور (400) دينار والانحراف المعياري (5) والمطلوب حساب معامل الاختلاف لمعرفة في أي من المصنعين تتوزع الأجور بشكل أكثر عدالة.

الحل:

نجد معامل الاختلاف للمصنع الأول:

$$C.V=\frac{2.5}{250}\times 100$$

$$=\frac{250}{250}=1\%$$

ثم نجد معامل الاختلاف للمصنع الثاني:

$$C.V = \frac{5}{400} \times 100 = 1.25\%$$

بالمقارنة نجد أن التشتت للأجور في المصنع الأول أقل منه في المصنع الثـاني وعلى هذا الأساس الأجور أكثر عدالة في المصنع الأول.

تمارين الفصل الرابع

1. وضح المقصود بمقاييس التشتت والمدى.
2. إذا كانت أصغر قيمة هي 23 وأكبر قيمة هي 83. جد المدى.
3. البيانات التالية متوفرة لدينا وهي كما يلي:

(50, 40, 35, 28, 14, 20, 8, 15)

المطلوب:

أ. حساب المدى.

ب. حساب المئين 25.

ج. حساب المد الربيعي.

4. البيانات التالية تمثل أوزان مجموعة من الطلاب:

الفئات (الأوزان)	التكرار (عدد الطلاب)
59 – 50	4
69 – 60	16
79 – 70	10
89 – 80	5
99 – 90	2

المطلوب:

أ. حساب نصف المدى الربيعي.

ب. الوسط الحسابي.

ج. العشير السابع.