قسم علم الاجتماع السنة الأولى جذع مشترك	المادة: الاحصاء الوصفي الأفواج: 55 – 56

الدرس الثاني:

مقاييس التشتت (الاختلاف)

Caracteristique de Dispersion

مقدمة: إليك البيانات التالية لمجموعتين من الطلاب تمثل درجات مادة الاحصاء :

مج 1: 76 – 64 – 71 – 70 – 96

مج 2: 60 – 100 – 80 – 40 – 70

$$\bar{X} = \frac{76+64+71+64+76}{5} = 70$$

$$\bar{X} = \frac{60+100+80+40+70}{5} = 70$$

نلاحظ أن الوسطين متساويين وهما (70)، لكن إذا قمنا بمقارنة بين المجموعتين نستنتج أن مستوى الطلاب في المجموعتين متجانس، إلا أن هذه لنتائج تخالف واقع البيانات لأن درجة الطلاب في المجموعة (1) تظهر متقاربة من بعضها وتتمركز حول وسطها الحسابي، في حين أن درجات الطلاب في المجموعة (2) تظهر متباعدة عن بعضها البعض، لذا نحتاج إلى مقاييس أخرى غير مقاييس النزعة المركزية التي لم تعد كافية أحيانا لوصف ظاهرة ما ومقارنتها بأخرى، بحيث أن بعض الظواهر قد تتشابه في أوساطها ولكن تختلف في تباعدها وتقارب بياناتها عن أوساطها الحسابية، ولمعرفة مدى تجانس الظواهر في خصائصها نلجأ إلى استخدام مقاييس أخرى تسمى مقاييس التشتت أو الاختلاف.

التشتت: معناه تقارب أو تباعد البيانات عن وسطها الحسابي، فكلما كانت قريبة تكون البيانات متجانسة، وكلما تباعدت عن وسطها تكون متشتة، وتقاس بعدة مقاييس منها: المدى، المدى الربيعي، الانحراف المتوسط، الانحراف المعياري، والانحراف الربيعي(نصف المدى الربيعي)، معامل الاختلاف.

1 – المدى Range :

1 –1: المدى العام (étendue ou Range): هو الفرق بين أكبر قيمة وأصغر قيمة، ونرمز له بالرمز: R

1 – 1 : البيانات الغير مبوبة: نستخدم القانون التالي:

$$R = X_{MAX} - X_{MIN}$$

مثال: أحسب المدى للبيانات التالية: 95 – 200 – 250 – 300 – 110 – 90 – 150 – 100– 350 – 80.

الحل:

الترتيب تصاعديا : 80–90–95–100–110–150–200– 250–300–350

$$M = 350 - 80 = 270$$

1 –2 : البيانات المبوبة: نستخدم القانون التالي:

المدي = الحد الأعلى للفئة الأخيرة – الحد الأدنى للفئة الأولى

مثال: أحسب المدى للجدول التالي: (–16 تقرأ : أقل من 16 / +16 تقرأ: أكثر من 16)

هذا جدول مستمر وليس متقطع، ونحسب مركز الفئة في مثل هذه الجداول كما يلي: $\frac{16+20}{2}$ ،

$\frac{20+24}{2}$ ، $\frac{24+28}{2}$ ، $\frac{28+32}{2}$ ، $\frac{32+36}{2}$

الفئات	16– الحد الأدنى	20–	24–	28–	32 – 36 الحد الأعلى
عدد المبحوثين	10	15	40	20	15

الحل:

المدي = الحد الأعلى للفئة الأخيرة – الحد الأدنى للفئة الأولى

المدي = 36 – 16 = 20

1 – 2: المدى الربيعي: هو الفرق بين الربيع الثالث والربيع الأول، ويرمز له بالرمز IQ,

ويحسب كما يلي: $IQ = Q_3 - Q_1$

1 – 3: الانحراف الربيعي النسبي: هو الفرق بين الربيع الثالث والربيع الأول على إثنين، ويرمز له بالرمز EQ،

ويحسب كما يلي $EQ = \frac{Q3 - Q1}{2}$

1 – 4: الانحراف الربيعي (نصف المدى الربيعي)Le Semi–Interquartile : للتغلب على بعض عيوب المدى والتي من أهمها التأثر بالقيم الشاذة، وعدم امكانية حسابه في حال وجود الجداول التكرارية المفتوحة، نلجأ إلى مقياس آخر وهو الانحراف الربيعي، ويعرف على أنه النصف المدى الربيعي، والمدى الربيعي هو الفرق بين الربيع الثالث والربيع الأول،كما ذكرت سابقا

ويحسب كما يلي:

$$EQp = \frac{Q_3 - Q1}{Q2}$$

من مزاياه:

- **هو المقياس الوحيد الذي يستخدم في حساب التشتت، إذا كانت بعض فئات الجدول التكراري مفتوحة.**
- **هو أقل مقاييس التشتت تأثيرا بالقيم الشاذة والمتطرفة في جانبي التوزيع,**

عيوبه:

- **يأخذ في الحسبان 50% من البيانات ويهمل الـ 50% الاخر.**
- **يعطي فكرة عن التشتت، لكن لا يعطي القيمة الفعلية بدقة، أنه يظهر على نحو واضح تباعد المفردات الواقعة بين الربيع الثالث والربيع الأول أو تقاربها,**
- **لا يتصف بخواص رياضية وجبرية، وهو أمر لا يساعد استخدامه في الدراسات الإحصائية المتقدمة.**

1- 1 :البيانات الغير مبوبة:

مثال: لدينا القيم التالية: 71 -70- 69- 73- 68- 74- 77

1, حساب المدى الربيعي:

- الترتيب التصاعدي: 68- 69- 70- 71- 73- 74- 77

ترتيب الربيع الأول: Q_1	ترتيب الربيع الثالث: Q_3
$C_1 = \frac{1\,(N)}{4} = \frac{1(7+1)}{4} = \frac{8}{4} = 2$	$C_3 = \frac{3\,(N)}{4} = \frac{3(7+1)}{4} = \frac{24}{4} = 6$

- موقع الربيع الأول هو: $Q_1 = 69$ / موقع الربيع الثالث هو :

$Q_3 = 74$

- $$IQ = Q_3 - Q_1 = 74 - 69 = 5$$

2. حساب الانحراف الربيعي النسبي:

- $$EQ = \frac{Q3 - Q1}{2} = \frac{5}{2} = 2,5$$

3. حساب الانحراف الربيعي:

$$Q_2 = \frac{Q_3 + Q_1}{2} = \frac{74 + 69}{2} = \frac{143}{2} = 71,5$$

$$EQp = \frac{Q_3 - Q1}{Q2} = \frac{5}{71,5} = -0,06$$

1- 2 : البيانات المبوبة:

$$Q = L_1 + \left(\frac{\frac{\Sigma N}{4} - f_1}{f_2} \right) \times L_c$$

مثال:

1 ، ايجاد موقع Q_1 و Q3، ثم تحديد الفئة الربيعية :

C	f	$f\uparrow$
50-55	30	30
55- 60	34	64
60-65	40	104
65- 70	28	132
70- 75	12	144
75-80	6	150
Σ	150	/

37,5

ترتيب الربيع الأول: Q_1	ترتيب الربيع الثالث: Q_3
$C_1 = \frac{1\,(N)}{4} = \frac{1(150)}{4} = 37,5$	$C_3 = \frac{3\,(N)}{4} = \frac{3(150)}{4} = 112,5$

65- 70	55- 60

122,5

2, حساب الربيع الأول والثالث:

L_1 : الحد الأدنى للفئة الربيعية وهو (55) و(65)

$\frac{\Sigma N}{4}$: رتبة الربيع الأول هو 37,5 والربيع الثالث هو 112,5

F_1 : تكرار المتجمع الصاعد السابق للفئة الربيعية

F_2 :تكرار الفئة الربيعية

L_C: طول الفئة الربيعية وهي: 60-55= 5 / 70- 65=5

$$Q_1 = L_1 + \left(\frac{\frac{\Sigma 1(N)}{4} - f_1}{f_2}\right) \times L_c = 55 + \left(\frac{37,5 - 30}{34}\right) \times 5 = 56,108$$

$$Q_3 = L_1 + \left(\frac{\frac{\Sigma 3(N)}{4} - f_1}{f_2}\right) \times L_c = 65 + \left(\frac{112,5 - 104}{28}\right) \times 5$$
$$= 66,517$$

2, المدى الربيعي: $IQ = Q_3 - Q_2 = 66,517 - 56,108 = 10,415$

3, حساب الانحراف النسبي: $EQ = \frac{3 - Q_1}{2} = \frac{66,517 - 56,108}{2} = \frac{10,409}{2} = 5,20$

4. حساب الانحراف الربيعي:

$$Q_2 = \frac{Q_3 + Q_1}{2} = \frac{66,517 + 56,108}{2} = \frac{122,625}{2} = 61,3125$$

$$EQp = \frac{Q_3 - Q1}{Q2} = \frac{66,517 - 56,108}{61,3125} = \frac{10,409}{61,3125} = 0,16$$

2 – الانحراف المتوسط (Ecart-Moyen): المدى المطلق يعتمد في حسابه على القيم المتطرفة في توزيع وأنه لا يهتم بباقي القيم، كما أن المدى الربيعي يهمل نصف عدد القيم التي تقع في طرفي التوزيع، وأنه لا يهتم إلا بالنصف المتوسط من القيم، لذا لا بد من التفكير في مقياس آخر يأخذ في الحسبان تشتت جميع القيم عن القيمة الموسطة، من أجل ذلك تم الاعتماد على الانحراف عن المتوسط ويعرف على أنع هو الوسط الحسابي لفروقات القيم عن وسطها الحسابي بالقيمة المطلقة، ويرمز له بالرمز: e ، ومن مزايا المتوسط الحسابي أنه يأخذ بجميع القيم، ودرجة تأثره بالقيم الشاذة ضعيفة. و تكون القيمة المطلقة موجة مهما وجدتم القيمة سالبة

2 – 1 :البيانات الغير مبوبة: ويحسب بالقانون التالي:

$$e = \frac{\sum |x - \bar{X}|}{N}$$

مثال: أوجد الانحراف المتوسط للبيانات التالية: 32-31-30-26-25-25 -20

الحل: n= 7

1. نحسب الوسط الحسابي: $\bar{X} = \frac{20+25+25+26+30+31+32}{7} = 27$

2. نشكل جدولا تدعيميا لتسهيل حساب الفروق المطلقة:(هذا ليس جدولا متقطعا ولا جدولا مستمرا):

x	20	25	25	25	26	30	31	32	Σ

24	5	4	3	3	-1	-2	-2	$20 - 27 = -7$	$\lvert x - \bar{X} \rvert$

$$e = \frac{\Sigma \lvert x - \bar{X} \rvert}{n}$$

$$e = \frac{24}{7} = 3{,}43$$

هذا يعني أن قيم الظاهرة المدروسة تنحرف عن وسطها في المتوسط عن وسطها الحسابي بمقدار هو 3,43 ,

2 – 2 : البيانات المبوبة: نستخدم القانون التالي:

$$e = \frac{\Sigma f \lvert x - \bar{X} \rvert}{\Sigma F}$$

2 – 2 – 1: في حالة البيانات المتقطعة:

مثال : إليك البيانات التالية: والمطلوب : أوجد الانحراف المتوسط للبيانات التالية:

x	f	$(x \times f)$	$\lvert x - \bar{X} \rvert$	$f \lvert x - \bar{X} \rvert$
1	7	7	1,85	12,95
2	5	10	0,85	4,25
3	8	40	2,15	17 ,2
Σ	20	57	/	34,4

الحل:

$$\bar{X} = \frac{\Sigma (x \times f)}{\Sigma F} \Rightarrow \frac{57}{20} = 2{,}85$$

$$e = \frac{\Sigma f \lvert x - \bar{X} \rvert}{\Sigma F} = \frac{34{,}4}{20} = 1{,}72$$

2 – 2 – 2: في حالة البيانات المستمرة:

هنا نستخدم مركز الفئة C_I، ونستخدم القانون التالي:

$$e = \frac{\Sigma f\,(c\,i - \bar{X}\,)}{\Sigma F}$$

مثال: إليك البيانات التالية:

C	f	C_i	$(f \times x)$	$\lvert x - \bar{X} \rvert$	$f \lvert x - \bar{X} \rvert$
10 −14	1	12	12	−9	9
15− 19	7	17	119	−4	28
20− 24	8	22	176	1	8
25−29	3	27	81	6	18
30−34	1	32	32	11	11
Σ	20	/	420	/	74

$$\bar{X} = \frac{\Sigma(\,x \times f)}{\Sigma f} \Rightarrow \frac{420}{20} = 21$$

$$e = \frac{\Sigma f\,(\,c\,i - \bar{X}\,)}{\Sigma f} = \frac{74}{20} = 3,7$$

مزاياه:

- يأخذ في الحسبان جميع القيم.
- يمكن حسابه استنادا إلى الوسط الحسابي أو الوسيط، يفضل في التوزيعات الغير المتماثلة والقريبة من التماثل حسابه الاعتماد على الوسط الحسابي، أما التوزيعات الغير المتماثلة فيفضل حسابه إستنادا إلى الوسيط.

عيوبه:

- نادر الاستعمال، فاستعماله يعتمد على سرعته في قياس التشتت.
- لا يمكن حسابه من الجداول التكرارية ذات الفئات المفتوحة.

3 – التباين La Variance: هو عبارة عن متوسط مربعات انحرافات القيم عن وسطها الحسابي، ويسمى هذا المتوسط بالتباين، وهو المقياس الأكثر استخداما ويرمز له بالرمز: V= $S^2 = \sigma^2$

3 – 1: البيانات الغير مبوبة: ويحسب بالقانون التالي:

$$V = \sigma^2 = S^2 = \frac{\Sigma (x - \bar{X})^2}{N}$$

مثال: أحسب التباين البيانات التالية: 12-10-11-7-5 / 5 = N

الحل: 1 – نحسب الوسط الحسابي:

$$\bar{X} = \frac{12 + 10 + 11 + 7 + 5}{5} = \frac{45}{5} = 9$$

2- نحصل على الانحرافات ومربعاتها بالاعتماد على جدول تدعيمي:

x	X²	$(x - \bar{X})$	$(x - \bar{X})^2$
12	144	3	9
10	100	1	1
11	121	2	4
7	49	−2	4
5	25	−4	16
Σ	439	/	34

$$V = \frac{\Sigma (x_i - \bar{X})^2}{N} = \frac{34}{5} = 6{,}8$$

3 – 2: البيانات المبوبة:

3 – 2 – 1: في حالة البيانات المتقطعة: نستخدم القانون التالي:

$$V = \frac{\Sigma f (x_i - \bar{X})^2}{\Sigma F}$$

مثال: إليك الجدول التالي:

X	f	$(x \times f)$	$(x - \bar{X})$	$(x - \bar{X})^2$
1	7	7	3,42	23,94
2	5	10	0,493	2,465
5	8	40	4,622	36,98
Σ	20	57		63,385

أحسب التباين؟

$$\bar{X} = \frac{\Sigma(x \times f)}{\Sigma F} = \frac{57}{20} = 2,85$$

$$V = \frac{\Sigma f\,(x_i - \bar{X})^2}{\Sigma F} = \frac{63,385}{20} = 3,169$$

3 – 2 – 2: في حالة البيانات المستمرة: هناك طريقتين للحساب، لكننا سنعتمد طريقة واحدة فقط.

$$V = \frac{\Sigma f\,(C_i - \bar{X})^2}{\Sigma F}$$

مثال:

إليك الجدول التالي:

والمطلوب حساب التباين؟

X	f	C_i	$(C_i \times f)$	$(C_i - \bar{X})$	$(C_i - \bar{X})^2$	$f(C_i - \bar{X})^2$
10−20	7	15	105	−10,5	110,25	771,75
20−30	5	25	125	−0,5	0.25	1,25
30− 40	8	35	280	10	100	800
Σ	20	/	510	/	/	1573

$$\bar{X} = \frac{\Sigma(C_i \times f)}{\Sigma F} = \frac{510}{20} = 25,5$$

$$V = \frac{\Sigma f\,(C_i - \bar{X})^2}{\Sigma F} = \frac{1573}{20} = 78,65$$