المحور الخامس: مقاييس التشتت

إن حساب المتوسطات تعطينا فكره عن التوزيع التكراري، إلا أن هذه الفكرة ليست دقيقه لان المتوسط وحده لا يكفي لإعطائنا فكره دقيقه عن المعلومة من حيث طبيعتها وكيفيه توزيع مفرداتها. كما أن المتوسط لا يكفي عند قيامنا مثلا بالمقارنة بين عده مجموعات فقد يتساوى المتوسط في حين تختلف مفردات المجموعات اختلافا كليا عن بعضها البعض .

تعتبر مقاييس التشتت مقاييس عددية تعبر عن مدى تباعد البيانات عن وسطها الحسابي، أي تبين كيفية توزيع وانتشار قيم المتغير الإحصائي حول القيمة المركزية. فكلما ارتفعت قيم البيانات دل ذلك على التباعد والاختلاف بين قيم البيانات، وكلما كانت القيم صغيره فهذا دليل على إن الاختلاف بين قيم البيانات اقل وبذلك فهذه المقاييس تعطينا فكره عن مدى تجانس أو اختلاف البيانات.

1- المدى العام:

هو الفرق بين اكبر قيمة واصغر قيمة للتوزيع الإحصائي. يأخذ المقياس بعين الاعتبار القيمتين الأولى و الأخيرة فقط. وبالتالي فانه يتأثر بالقيم المتطرفة ويستخدم للمقارنة بين توزيعين إحصائيين أو أكثر، ويرمز له بالرمز « ET ».

$$ET = X_{max} - X_{min}$$

$$ET = X_n - X_1$$

مثال: تبين السلسلتان الإحصائيتان التاليتان توزيع مجموعتين من العمال في مؤسستين مختلفتين حسب الأجور اليومية :

التوزيع الأول:

05	04	03	02	01	العامل
165	150	140	135	120	الأجر اليومي

التوزيع الثاني:

05	04	03	02	01	العامل
200	180	140	100	80	الاجر اليومي

حساب المدى العام للسلسلتين:

$ET_1=165-120=45$ ، $ET_2=200-80=120$ لدينا $ET_2>ET_1$ إذا التوزيع الثاني أكثر تشتتا من التوزيع الأول ونستنتج أن أجور المؤسسة الأولى أكثر تجانسا وعدالة في توزيع الأجور مقارنة بالمؤسسة الثانية.

أما المدى في حالة البيانات المبوبة فيحسب كالتالي: $ET= A_{1n}-A_{01}$

حيث A_{1n} هو الحد الأعلى للفئة الأخيرة و A_{01} هو الحد الأدنى للفئة الأولى.

2- الانحراف المتوسط:

يعرف الانحراف المتوسط لمجموعة من العلامات مثلا بأنه متوسط الانحرافات المطلقة لتلك العلامات عن وسطها الحسابي، ويقصد بالانحراف المطلق بأنه الفرق بين العلامة والوسط الحسابي بغض النظر عن اتجاه هذا الفرق سواء كان سالبا أو موجبا. يرمز له بالرمز MD ويحسب بالعلاقة: $\mathbf{MD}=\frac{\sum|\bar{X}-\mathbf{Xi}|}{n}$

لإيجاد الانحراف المتوسط نقوم بإتباع الخطوات التالية:

- حساب المتوسط الحسابي؛
- ثم إيجاد الانحرافات المطلقة وهي الفرق بين المتوسط الحسابي وقيم المتغير X؛
- تجميع كل الانحرافات المطلقة وقسمتها على عددها وبالتالي نجد الانحراف المتوسط.

ويعتبر الانحراف المتوسط أفضل من المقياس السابق (المدى) من حيث الدقة وإعطاء صورة أوضح للبيانات، حيث يمكن بسهولة التعرف على تشتت التوزيعات والمقارنة بينها.

مثال: لدينا العلامات التالية وانحرافها المطلق موضحة في الجدول التالي:

العلامة X	$\bar{X}$-X	الانحراف المطلق
3	10-3	7
4	10-4	6
6	10-6	4
9	10-9	1
12	10-12	2
16	10-16	6
20	10-20	10

المتوسط الحسابي:

$$\bar{X}=\frac{3+4+6+9+12+16+20}{7} \text{ ، } \bar{X}=10$$

الانحراف المتوسط:

$$\mathbf{MD}=\frac{\sum|\mathbf{Xi}-\bar{X}|}{n}=\frac{7+6+4+1+2+6+10}{7}=5.14$$

3- التباين:

المحور الخامس: مقاييس التشتت

يعتبر التباين احد أشهر مقاييس التشتت وأكثرها استخداما في المجالات التطبيقية ويعبر عن متوسط مربعات انحرافات القيم عن متوسطها الحسابي، يرمز له ب σ^2 . ويحسب في حالة البيانات غير المبوبة بالعلاقة:

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

مثال: يعمل 15 موظف بقسم الإحصاء وكانت عدد سنوات الخبرة لديهم كالتالي:

10-12-11-6-14-13-10-8-6-9-12-14-7-13-5. المطلوب: اوجد التباين لسنوات الخبرة.

الحل:

- أولا نقوم بحساب المتوسط الحسابي $\bar{X} = \frac{\sum Xi}{n} = \frac{150}{15} = 10$
- ثانيا نقوم بحساب الانحرافات المطلقة ثم مربعها أي حساب $(Xi-\bar{X})$ ثم $(Xi-\bar{X})^2$

Xi	$Xi-\bar{X}$	$(Xi-\bar{X})^2$
5	5-	25
13	3	9
7	3-	9
14	4	16
12	2	4
9	1-	1
6	4-	16
8	2-	4
10	0	0
13	3	9
14	4	16
6	4-	16
11	1	1
12	2	4
10	0	0
المجموع/		130

$$\sigma^2 = \frac{\sum(X_i - \bar{X})^2}{n} = \frac{130}{15} = 8.67$$

أما في حالة البيانات المبوبة فهناك حالتين:

- في حالة متغير كمي منفصل فيحسب التباين بالعلاقة التالية: $\sigma^2 = \frac{\sum F(Xi-\bar{X})^2}{\sum F}$
- في حالة متغير كمي متصل (فئات) فيحسب بالعلاقة: $\sigma^2 = \frac{\sum F(Ci-\bar{X})^2}{\sum F}$

مثال: لدينا جدول التوزيع التكراري التالي، والمطلوب حساب التباين.

الفئات	Fi	Ci			(
]5-2]	1	3.5	-7.2	51.84	51.84
]8-5]	8	6.5	-4.2	17.64	141.12
]11-8]	13	9.5	-1.2	1.44	18.44
]14-11]	10	12.5	1.8	3.24	32.4
]17-14]	8	15.5	4.8	23.04	184.32
المجموع	40	/			428.4

$$\bar{X} = \frac{\Sigma(Ci\times Fi)}{\Sigma F} = \frac{428}{40} = 10.7$$

بالتالي فالتباين يساوي: $\sigma^2 = \frac{\sum F(Ci-\bar{X})^2}{\sum F} = \frac{428.4}{40} = \mathbf{10.71}$

4- الانحراف المعياري:

يعرف الانحراف المعياري على انه أدق مقاييس التشتت وأكثرها استخداما، وهو الجذر التربيعي لمتوسط مربعات انحرافات القيم عن وسطها الحسابي ويرمز له بالرمز (σ) يعرف غالبا عن طريق التباين (σ^2) ولحساب الانحراف المعياري في حالة بيانات غير مبوبة نقوم أولا بحساب التباين ثم نأخذ الجذر التربيعي للنتيجة النهائية لنحصل على المعادلة التالية:

$$\sigma = \sqrt{\frac{\sum(X_i - \bar{X})^2}{n}}$$

مثال: لدينا السلسلة الإحصائية من مثال سابق كالتالي:

5-13-7-14-12-9-6-8-10-13-14-6-11-12-10 .احسب الانحراف المعياري.

الحل: نقوم أولا بحساب التباين، (حسب نتائج المثال السابق وجدنا أن $\sigma^2 = 8.67$

بالتالي يكون الانحراف المعياري: $\boldsymbol{\sigma} = \sqrt{8.67} = \mathbf{2.94}$

أما في حالة بيانات مبوبة فيتم حساب الانحراف المعياري في حالة متغير كمي منفصل حسب العلاقة:

$$\sigma = \sqrt{\frac{\sum Fi(Xi-\bar{X})^2}{\sum Fi}}$$

وفي حالة متغير كمي متصل وفق العلاقة:

$$\sigma = \sqrt{\frac{\sum Fi(Ci-\bar{X})^2}{\sum Fi}}$$

مثال: لدينا التوزيع التكراري التالي، احسب الانحراف المعياري

			Fi	Xi
17.71	1.61	-1.27	11	2
1.26	0.07	0.27	18	3
4.77	0.53	0.73	09	4
20.93	2.99	1.73	07	5
44.67	/	/	**45**	**المجموع**

$$\bar{X} = \frac{\sum FiXi}{\sum Fi} = \frac{147}{45} = 3.27$$

$$\sigma = \sqrt{\frac{\sum Fi(Xi-\bar{X})^2}{\sum Fi}} = \sqrt{\frac{44.67}{45}} = 0.996$$

5- الانحراف نصف الربيعي:

لقد سبق التعرف على القيم التي تقسم البيانات إلى أربعة أقسام متساوية والتي تسمى بالربيعيات، يرمز لها ب Q_1،Q_2،Q_3. يتم استخدام المقدار الذي يمثل النصف الأوسط للقيم $|Q3 - Q1|$ المسمى بالانحراف الربيعي. كمقياس تشتت اصح من المدى. و بشكل أكثر دقة يتم اخذ الانحراف نصف الربيعي ويرمز له بالرمز Q وهو متوسط الفرق بين الربيعيين الثالث والأول كالتالي:

$$Q = \frac{Q3 - Q1}{2}$$