الفصل الرابع

مقاييس التشتت

مفهوم التشتت والهدف من حسابه:

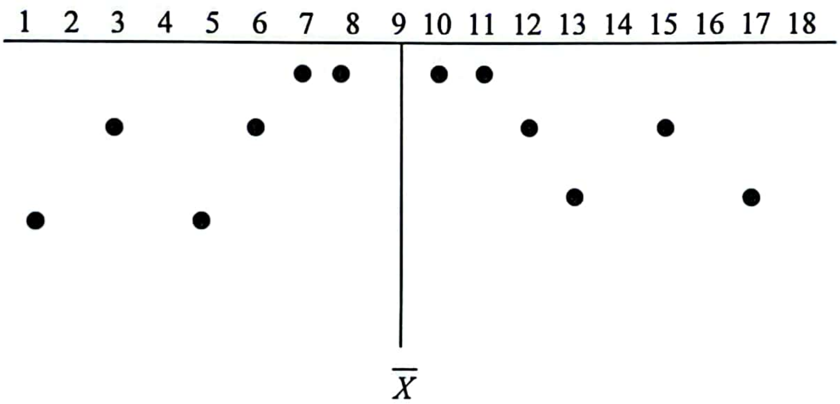
يعرف التشتت بأنه تباعد أو انتشار قيم مجموعة من المفردات عن بعضها البعض أو عن قيمة معينة ثابتة كالوسط الحسابي مثلاً.

إن الهدف من دراسة التشتت هو تكوين فكرة عن مدى تجانس قيم مجموعة من المفردات. اذ كلما كان مقياس التشتت كبيراً دل ذلك على عدم التجانس بين القيم وكلما كان مقياس التشتت صغيراً دل ذلك على ان الاختلافات بين قيم المشاهدات قليلة. وهذا يعني ان دراسة التشتت أمر مفيد في إجراء المقارنة بين قيم مجموعتين أو أكثر من البيانات عن ظاهرة معينة. فمثلاً نلاحظ أن الوسط الحسابي لكل مجموعة من المجموعات التالية هو (9).

المجموعة الأولى: 7, 8, 9, 10, 11

المجموعة الثانية: 3, 6, 9, 12, 15

المجموعة الثالثة: 1, 5, 9, 13, 17



نلاحظ أن المجموعة الأولى أكثر تجانساً (اقل انتشاراً) من المجموعتين الثانية والثالثة.
كذلك فان المجموعة الثانية أكثر تجانساً من المجموعة الثالثة.

وتقسم مقاييس التشتت إلى قسمين رئيسيين هما:

أ- مقاييس التشتت المطلقة:

وهي المقاييس التي تبين درجة تجانس قيم مجموعة من البيانات بشكل مطلق وتكون مقاسه بنفس وحدات قياس المتغير العشوائي (وحدات طول, وزن, زمن, كثافة, عدد, ... الخ) .

واهم هذه المقاييس هي:

1. المدى
2. الانحراف المتوسط
3. الانحراف المعياري (القياسي)
4. التباين

ب- مقاييس التشتت النسبية:

وهي المقاييس التي تبين درجة تجانس قيم مجموعة من البيانات بشكل نسبي وتكون خالية من وحدات قياس المتغير العشوائي وهذه المقاييس تسمى معاملات التشتت وأهمها معامل التشتت النسبي المستند إلى الانحراف المعياري (معامل الاختلاف) والدرجة المعيارية.

أ- مقاييس التشتت المطلقة:

1. المدى Range :

يعتبر المدى ابسط أنواع مقاييس التشتت. ويعرف المدى بأنه الفرق ما بين اكبر قيمة في مجموعة من البيانات واصغر قيمة فيها, ويرمز له بالرمز R .

$$R = X_L - X_S$$

إذ أن:

X_L : اكبر قيمة في مجموعة البيانات.

X_S : اصغر قيمة في مجموعة البيانات.

مثال:

جد المدى للبيانات التالية :

12, 6, 7, 3, 15, 10, 18, 5

الحل:

$$R = X_L - X_S$$
$$= 18 - 3 = 15$$

ملاحظة:

أما في حالة البيانات المبوبة في توزيع تكراري فان المدى في هذه الحالة يمثل الفرق ما بين الحد الأعلى للفئة الأخيرة والحد الأدنى للفئة الأولى.

2. الانحراف المتوسط Mean Deviation:

يعرف الانحراف المتوسط بأنه مجموع الانحرافات المطلقة لقيم المتغير العشوائي عن وسطها الحسابي مقسوماً على عدد هذه القيم ويرمز له بالرمز M.D.

طرق حساب الانحراف المتوسط:

أ. لبيانات غير مبوبة:

لتكن $X_1, X_2, ..., X_n$ تمثل قياسات عينة من المفردات حجمها n وليكن $\overline{X}$ الوسط الحسابي لهذه القيم عندئذٍ يمكن حساب الانحراف المتوسط وفق الصيغة التالية:

$$M.D = \frac{\sum_{i=1}^{n} \left| X_i - \overline{X} \right|}{n}$$

مثال:-

جد الانحراف المتوسط للقيم التالية:

$X_i = 9, 8, 6, 5, 7$

الحل:-

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} = \frac{35}{5} = 7$$

$$X_i - \overline{X} = 2, \quad 1, \quad -1, \quad -2, \quad 0$$

$$\left| X_i - \overline{X} \right| = 2, \quad 1, \quad 1, \quad 2, \quad 0$$

لذا فان الانحراف المتوسط هو:

$$M.D = \frac{\sum_{i=1}^{5} \left| X_i - \overline{X} \right|}{n} = \frac{6}{5} = 1.2$$

ب- لبيانات مبوبة:

لتكن $X_1, X_2, ... X_m$ تمثل مراكز فئات توزيع تكراري عدد فئاته m وإن $f_1, f_2, ... f_m$ تمثل التكرارات المقابلة لهذه الفئات. عندئذٍ يمكن حساب الانحراف المتوسط وفق الصيغة التالية:

$$M.D = \frac{\sum_{i=1}^{m} f_i \left| X_i - \overline{X} \right|}{\sum_{i=1}^{m} f_i}$$

إذ أن:

X_i : يمثل مركز الفئة i.

$\overline{X}$: يمثل الوسط الحسابي للتوزيع.

f_i : يمثل تكرار الفئة i.

مثال:- من جدول التوزيع التكراري التالي جد الانحراف المتوسط.

الفئات	التكرار f_i
60-62	5
63-65	18
66-68	42
69-71	27
72-74	8
المجموع	100

الحل:-

الفئات	التكرار f_i	مركز الفئة X_i	$f_i X_i$	$\left\| X_i - \overline{X} \right\|$	$f_i \left\| X_i - \overline{X} \right\|$
60-62	5	61	305	6.45	32.25
63-65	18	64	1152	3.45	62.10
66-68	42	67	2814	0.45	18.90
69-71	27	70	1890	2.55	68.85
72-74	8	73	584	5.55	44.40
المجموع	100		6745		226.5

$$\overline{X} = \frac{\sum_{i=1}^{m} f_i X_i}{\sum_{i=1}^{m} f_i} = \frac{6745}{100} = 67.45$$

$$M.D = \frac{\sum_{i=1}^{m} f_i \left| X_i - \overline{X} \right|}{\sum_{i=1}^{m} f_i} = \frac{226.5}{100} = 2.265$$

3. الانحراف المعياري (القياسي) Standard Deviation:

يعرف الانحراف القياسي بأنه الجذر التربيعي الموجب لمتوسط مجموع مربعات انحرافات قيم المتغير العشوائي عن وسطها الحسابي. ويعتبر هذا المقياس أفضل مقاييس التشتت لما يمتاز به من مميزات مثلى ويرمز له بالرمز S.

طرق حساب الانحراف القياسي:

أ. في حالة البيانات الغير مبوبة:

لتكن $X_1, X_2, ..., X_n$ تمثل قياسات عينة من المفردات حجمها n وليكن $\overline{X}$ الوسط الحسابي لهذه القيم عندئذٍ يمكن حساب الانحراف القياسي وفق الصيغة التالية:

$$S = \sqrt{\frac{\sum_{i=1}^{n} \left(X_i - \overline{X} \right)^2}{n}}$$

ملاحظة: يتم القسمة على { n } عندما يكون حجم العينة $n \geq 30$ أي من العينات الكبيرة وعندما تكون $n < 30$ فانه يقسم على { n-1 } أي انه يفضل استخدام التالي:

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

ب. في حالة البيانات المبوبة:

لتكن $X_1, X_2, ...X_m$ تمثل مراكز فئات توزيع تكراري عدد فئاته m وإن $f_1, f_2, ...f_m$ تمثل التكرارات المقابلة لهذه الفئات. عندئذٍ يمكن حساب الانحراف القياسي وفق الصيغة التالية:

$$S = \sqrt{\frac{\sum_{i=1}^{m} f_i \left(X_i - \overline{X} \right)^2}{\sum_{i=1}^{m} f_i}}$$

إذ أن:

X_i : يمثل مركز الفئة i.

$\overline{X}$: يمثل الوسط الحسابي للتوزيع.