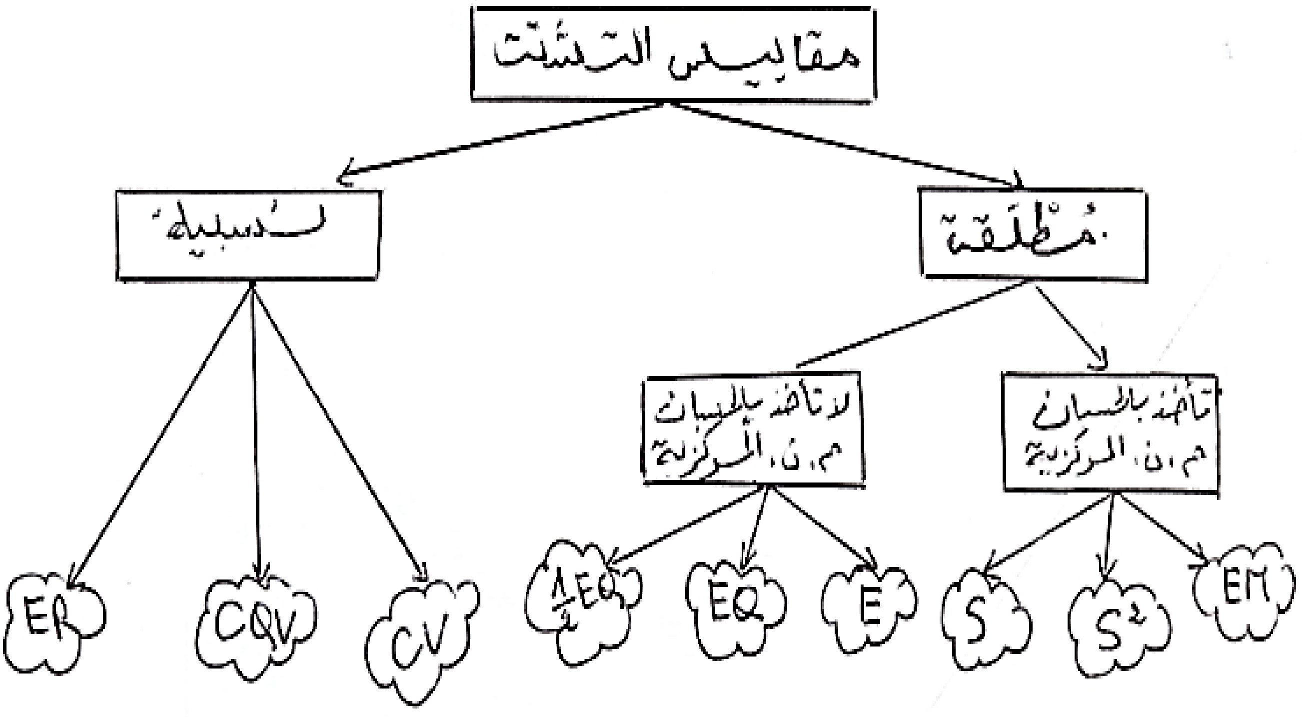
المحور الرابع : مقاييس التشتت

إن الاعتماد على مقاييس النزعة المركزية فحسب لوصف ظاهرة ما غير كاف. لأن هذه المقاييس تغفل عدة جوانب أخرى من الظاهرة، من بينها مدى تبعثر قيم هذه الظاهرة. ولهذا نلجأ إلى مجموعة أخرى من المقاييس تسمى مقاييس التشتت. فتنقسم مقاييس التشتت إلى مجموعتين رئيسيتين تسمى الأولى مقاييس التشتت المطلقة، وتسمى الثانية مقاييس التشتت النسبية.



الأستاذ / هاشمي عبابسة

26

المجموعة الأولى: مقاييس التشتت المطلقة

وتنقسم الى نوعين :-

I- المقاييس التي لا تأخذ بالحسبان مؤشرات النزعة المركزية :-

1- المدى: "E" (L'étendue)

أ- تعريفه :- " المدى عبارة عن مجال تغيّر، يسمح لقياس الفرق بين أعلى قيمة وأدنى قيمة بين المشاهدات"، حيث يسمح بإعطاء فكرة سريعة عن تشتت مفردات المجموعة .

ب- حسابه :-

* بالنسبة لسلسلة عددية :-

$$E = X_k - X_1$$ (21)

حيث X_k : أعلى قيمة

X_1 : أدنى قيمة .

* بالنسبة لتوزيع تكراري :- "حسب المدى بطريقتين :-

الطريقة الأولى :- المدى = مركز الفئة الأخيرة - مركز الفئة الأولى .

الطريقة الثانية :- المدى = الحد الأعلى الفعلي للفئة الأخيرة - الحد الأدنى الفعلي للفئة الأولى .

ج- مزاياه :-

* سهولة حسابه .
* يساعد على اعطاء فكرة سريعة حول تشتت مفردات الظاهرة .

د- عيوبه :-

* التأثر بالقيم المتطرفة .
* يصعب حسابه في التوزيعات التكرارية المفتوحة .
* يعتمد فقط على القيمتين الكبرى والصغرى ، ويهمل بقية القيم .
* غير دقيق لما يكفي لقياس التشتت .

الأستاذ / هاشمي عبابسة

27

2- المدى الربيعي (EQ)

أ- تعريفه :- هو عبارة عن مجال التغير الأوسط، الخالي من القيم المتطرفة الناتج عن استبعاد الربيعين الأول والأخير لمجموعة المعطيات، حيث يسمح بقياس الفرق بين أعلى قيمة في هذا المجال (Q_3) وأدنى قيمة فيه (Q_1).

ب- حسابه :- سواء بالنسبة لسلسلة عددية أو لتوزيع تكراري نحسب EQ كما يلي :-

$$EQ = Q_3 - Q_1 \quad (22)$$

ج- مزاياه :-

* سهولة حسابه.
* لا يتأثر كثيراً بالقيم المتطرفة.
* عادة ما يمكن حسابه في التوزيعات التكرارية المفتوحة.
* يساعد على إعطاء فكرة سريعة عن تشتت المجموعة.

د- عيوبه :-

* لا يعتبر من مقاييس التشتت الدقيقة، فهو مؤشر تقريبي، لا يستخدم إلا لأخذ فكرة سريعة عن تشتت المعطيات.
* يعتمد فقط على القيمتين الكبرى والصغرى في المجال الأوسط للقيم، ويهمل البقية.

3- نصف المدى الربيعي ($\frac{1}{2}EQ$)

هو مقياس تشتت شائع الاستخدام، لا يتأثر بالقيم المتطرفة وينتج عن قسمة المدى الربيعي على 2.

أي أن :

$$\frac{1}{2}EQ = \frac{1}{2}(Q_3 - Q_1) \quad (23)$$

II- المقاييس التي تأخذ بالحسبان مؤشرات النزعة المركزية:

1- الانحراف المتوسط: EM (L'écart Moyen)

أ- تعريف: هو الوسط الحسابي لانحرافات القيم عن مؤشر من مؤشرات النزعة المركزية، عادة ما تكون $\bar{X}$ و Me.
تؤخذ هذه الانحرافات بقيمها المطلقة.

ب- حسابه:
✓ بالنسبة لسلسلة من القيم:

$$EM = \frac{\sum |x_i - Me|}{n} \quad \text{أو} \quad EM = \frac{\sum |x_i - \bar{X}|}{n} \qquad (24)$$

✓ بالنسبة لتوزيع تكراري:

$$EM = \frac{\sum n_i |x_i - Me|}{n} \quad \text{أو} \quad EM = \frac{\sum n_i |x_i - \bar{X}|}{n} \qquad (25)$$

ج- مميزاته:
* يعتمد في حسابه على جميع القيم.
* سهل الحساب.

د- عيوبه:
* يتأثر بالقيم المتطرفة.
* يصعب حسابه في التوزيعات التكرارية المفتوحة.
* لا يمكن إيجاده لبيانات نوعية.

2- التباين S^2 La variance

أ- تعريفه: التباين مؤشر يقيس تشتت القيم حول وسطها الحسابي من خلال حساب متوسط مجموع مربعات (قيم) انحرافات القيم عن وسطها الحسابي.

وقد لجأ الإحصائيون الى ترييع هذه الانحرافات للتخلص من الاشارة السالبة لبعض الانحرافات تفاديا لانعدام مجموعها، وذلك عوض استخدام القيمة المطلقة كما هو الحال في EM.

ب- حسابه:

$$S^2 = \frac{\sum n_i (x_i - \bar{x})^2}{n} \quad \text{——} \quad (26)$$

* ملاحظات:

1- اذا كانت لدينا سلسلة قيم دون تكرارات، يبقى القانون نفسه، فقط نحذف "n_i".

2- أحيانا نحسب S^2 بالقسمة على $(n-1)$ بدل القسمة على n، وذلك للحصول على تقدير أفضل لتباين المجتمع الذي تسحب منه العينة.

لكن اذا كانت العينة كبيرة، اي قيمة n كبيرة كما يكفي $(n \geqslant 30)$ فإنه لا حاجة للتعديل والقسمة على $(n-1)$، لأن الفرق لن يكون معتبرا بين الحالتين.

3- هناك طرق أخرى لحساب S^2 وتعطي النتيجة ذاتها دائما، وهي:

* طريقة البيانات الأصلية:

حيث نعتمد على البيانات الأصلية x_i دون الحاجة الى حساب $\bar{x}$.

$$S^2 = \frac{\sum n_i x_i^2}{n} - \left[\frac{\sum n_i x_i}{n}\right]^2 = \overline{x^2} - (\bar{x})^2 \quad \text{——} \quad (27)$$

ونحذف n_i اذا كانت لدينا سلسلة قيم x_i دون تكرارات.

* طريقة الانحرافات:

حيث نحسب من القيم x_i أي $Z_i = x_i - A$ (A عدد حقيقي).

$$S^2 = \frac{\sum n_i (Z_i - \bar{Z})^2}{n} = \frac{\sum n_i Z_i^2}{n} - \left[\frac{\sum n_i Z_i}{n}\right]^2 = \overline{Z^2} - (\bar{Z})^2 \quad \text{——} \quad (28)$$

ونحذف n_i اذا كانت لدينا سلسلة قيم دون تكرارات، وهذا ينطبق على جميع الحالات الآتية.

* طريقة الترميز:

حيث نحسبه من القيم u_i حيث $u_i = \frac{x_i - A}{L}$ اين A عدد حقيقي و L هو طول الفئات أو القاسم المشترك لأطوال الفئات (انظر الصفحة 17)

$$S^2 = \frac{\sum n_i (u_i - \bar{u})^2}{n} = \frac{\sum n_i u_i^2}{n} - \left[\frac{\sum n_i u_i}{n}\right]^2 \quad (29)$$

$$= \overline{u^2} - (\bar{u})^2$$

د- الانحراف المعياري: S (L'écart Type)

أ- تعريفه: يعد من أهم مقاييس التشتت وأكثرها دقة واستخداما لأنه يعطي فكرة ملموسة ومنطقية عن مدى اختلاف قيم ظاهرة معينة عن وسطها الحسابي. ويعرف بأنه الجذر التربيعي لمتوسط مجموع مربعات انحرافات القيم عن وسطها الحسابي.

ب- حسابه: وفقا للتعريف السابق فإن الانحراف المعياري يحسب كالآتي:

$$S = \sqrt{\frac{\sum n_i (x_i - \bar{x})^2}{n}} \quad (30)$$

ملاحظات:

1- اذا كانت لدينا سلسلة من القيم دون تكرارات، فإننا نطبق القانون السابق ذاته، لكن نحذف n_i.

2- كل ما قيل عن التباين يمكن أن نطبقه على S، بشأن القسمة على $(n-1)$ أو حساب S بطرق أخرى (انظر ملاحظات الصفحة 29).

ه- خصائص الانحراف المعياري: [هي نفسها خصائص التباين S^2]

1- اذا كانت لدينا القيم x_i والقيم y_i، حيث $y_i = x_i \pm A$ (A قيمة افتراضية) فإن $S_x = S_y$.

ويعني هذا أن الانحراف المعياري (والتشتت عموما) لا يتأثر ولا يتغير بإضافة قيمة ثابتة الى سلسلة القيم أو طرحها منها.

2- اذا كانت لدينا القيم x_i والقيم y_i، حيث $y_i = Ax_i$ (A عدد حقيقي) فإن $S_y = AS_x$.

أما إذا كان $y_i = \frac{x_i}{A}$ فإن $S_y = \frac{S_x}{A}$ اي أن الانحراف المعياري يتأثر ويتغير بالضرب في قيمة او القسمة عليها.

3- الانحراف المعياري S للقيم X حول أي عدد حقيقي A يصل الى أدنى قيمة ممكنة له فقط عندما يكون A مساويا للوسط الحسابي $\bar{X}$ لهذه القيم.

4- اذا كانت لدينا مجموعتان من القيم، وسطاهما الحسابيين هما $\bar{X}_1$، $\bar{X}_2$ وعدد عناصرهما n_1 و n_2، وتباينهما S_1^2، S_2^2 على الترتيب. فإن تباين المجموعة الناتجة عن دمج هاتين المجموعتين هو S^2 حيث:

$$S^2 = \frac{n_1 S_1^2 + n_2 S_2^2 + n_1(\bar{X}_1 - \bar{X})^2 + n_2(\bar{X}_2 - \bar{X})^2}{n_1 + n_2}$$

أين $\bar{X}$ هو الوسط الحسابي المشترك لكل من $\bar{X}_1$ و $\bar{X}_2$. (انظر قانون حسابه في خصائص $\bar{X}$ الخاصية الخامسة في الصفحة 17) يمكن تعميم هذا القانون على أكثر من مجموعتين.

5- كلما كانت قيمة S صغيرة، دل ذلك على قلة تشتت المعطيات حول $\bar{X}$، ودل على جودة تمثيل هذا الأخير للظاهرة.

6- وحدة قياس S هي نفسها وحدة قياس المتغير المدروس. ولذلك لا يمكن استخدامه لمقارنة تشتت مجموعتين غير متجانستين.

7- قيمة دوما موجبة (كبقية مقاييس التشتت).

هـ- عيوب الانحراف المعياري:

* يتأثر بالقيم المتطرفة. * يتعذر حسابه للبيانات النوعية.
* يصعب حسابه في التوزيعات التكرارية المفتوحة.
* لا يمكن استخدامه لمقارنة تشتت مجموعتين غير متجانستين.

و- العلاقة الاعتبارية بين مقاييس التشتت المطلقة:

في التوزيعات ذات الالتواءات البسيطة يتحقق مايلي:

$EM = \frac{4}{5} S$, $\frac{1}{2} EQ = \frac{2}{3} S$