٦ ـ ٨ : معاملات التشتت coefficients of dispersion

قد يتطلب الامر في بعض الاحيان اجراء المقارنة بين تشتت مجموعتين او اكثر من القيم المختلفة عن بعضها من حيث الوسط الحسابي ، او ان قيم مفردات كل مجموعة مقاسة بوحدات قياس تختلف عن الاخرى . وعندئذٍ فان مقياس التشتت المطلق (اياً كان) سوف لن يكون مقياس نافع لوحده في اجراء مقارنات من هذا النوع انما يستوجب الامر ايجاد مقياس تشتت آخر اكثر ملائمة لهذه الحالات . هذا النوع من المقاييس تسمى بمقاييس التشتت النسبية وكما ذكرنا سابقاً فان هذه المقاييس خالية من وحدات القياس ، هذه المقاييس هي :

٦ ـ ٨ ـ ١ : معامل التشتت المستند الى المدى :

افرض ان $X_1, X_2, \ldots, X_n$ تمثل قياسات مفردات عينة قوامها n وان X_L تمثل اكبر قيمة في هذه المجموعة وان X_S تمثل اصغر قيمة فيها . عندئذٍ يعرف معامل التشتت المدوي C.D على النحو الاتي :

$$C.D(R) = \frac{X_L - X_S}{X_L + X_S} = \frac{R}{X_L + X_S} \cdot 100$$

حيث ان R تمثل المدى لهذه المجموعة :

وحيث ان R لايعول عليه كثيراً بسبب حساسيته كونه يعتمد على قيمتين متطرفتين فقط لذا فان معامل التشتت المدوي هو الاخر لايعول عليه كثيراً .

مثال : ـ قارن بين درجة تجانس قيم المجموعتين التاليتين .

المجموعة الاولى : 2 , 5 , 7 , 11 , 9 , 12 , 8 , 6 , 20 , 13 , 4 , 5

المجموعة الثانية : 6 , 8 , 12 , 4 , 7 , 3 , 7 , 9 , 11 , 12 , 8 , 5

الحل : ان الوسط الحسابي لكل مجموعة وعلى التوالي هو 8.5 و 7.7 وذلك امر غير نافع في اجراء المقارنة . عليه نستخدم معامل التشتت المدوي في ذلك . ان

$$R_1 = 20 - 2 = 18 \qquad R_2 = 12 - 3 = 9$$

$$\therefore C.D(R_1) = \frac{18}{22} \cdot 100 = 81.82\%$$

$$C.D(R_2) = \frac{9}{15} \cdot 100 = 60\%$$

٦ ـ ٨ ـ ٤ : معامل التشتت المستند الى الانحراف المعياري (معامل الاختلاف) Coefficient of variation

افرض ان $\bar{X}$ يمثل الوسط الحسابي لمجموعة قيم وان S يمثل الانحراف المعياري لها . عندئذٍ يعرف معامل الاختلاف على النحو التالي :

$$C.V = \frac{S}{\bar{X}} \cdot 100$$

ان معامل الاختلاف يعتبر بحق افضل انواع معاملات التشتت الانفة الذكر كونه يعتمد على افضل مقياس نزعة مركزية وافضل مقياس تشتت ان هذا المعامل يوضح نسبة حصة كل وحدة من وحدات الوسط الحسابي من الانحراف المعياري . وعليه وعند اجراء مقارنة بين قيم مجموعتين تتم مقارنة معامل اختلاف الاولى مع معامل اختلاف الثانية وعندئذ يقال عن المجموعة بانها اكثر تجانساً اذا كان معامل اختلافها اقل من الاخرى . .

وبرغم كون هذا المعامل افضل الانواع الاخرى الا انه يصعب الحصول عليه من توزيعات تكرارية مفتوحة . كذلك فهو مقياس حساس في حالة وجود قيم شاذة او متطرفة . وبشكل عام يمكن القول بان العيوب التي تخص الوسط الحسابي والتي سبق ذكرها تؤثر وبشكل مباشر على قيمة هذا المعامل .

مثال : كان متوسط درجات طلبة الصف الاول احصاء في امتحان الرياضيات 69 درجة بانحراف معياري قدره 19.3 في حين كان متوسط درجاتهم في امتحان الاحصاء 75 درجة بانحراف معياري قدره 25.5 . في اي من الامتحانين كان مستوى اداء الطلبة اكثر تقارباً .

الحل : نجد معامل الاختلاف لكل امتحان :

$$C.V(\text{stat.}) = \frac{25.5}{75} \cdot 100 = 34\% \,,\quad C.V(\text{math}) = \frac{19.3}{69} \cdot 100 = 28\%$$

وحيث ان معامل الاختلاف في امتحان الرياضيات اقل من معامل اختلاف امتحان الاحصاء عليه فان مستوى اداء الطلبة في امتحان الرياضيات كان اكثر تقارباً .

٦ ـ ٨ ـ ٥ : اختيار معامل التشتت الملائم

كما لاحظنا مما سبق فانه لايوجد معامل تشتت يلائم كافة الحالات التي تواجه الباحث . حيث لكل منها ميزاته وعيوبه . الا ان معامل الاختلاف يعتبره معظم علماء الاحصاء بانه افضل هذه المعاملات كونه يعتمد على اهم مقياس نزعة مركزية (الوسط الحسابي) واهم مقياس تشتت (الانحراف المعياري) . ان لهذا المعامل تطبيقات كثيرة في مجال الرقابة الاحصائية على جودة الانتاج وكذلك في المجالات الطبية والاقتصادية والزراعية وغيرها من المجالات التي يتطلب الامر فيها اجراء المقارنات بين مجموعتين أو اكثر . وكما ذكرنا فانه يصعب في بعض الاحيان ايجاد قيمة هذا المعامل لاسباب سبق ذكرها . وفي مثل هذه الاحوال لايوجد امام الباحث من سبيل سوى استخدام معامل التشتت المستند الى الانحراف الربيعي كافضل معامل يلائم هذه الحالات .

٦ ـ ٩ : الدرجة المعيارية Standard Score

في كثير من الاحيان نحتاج الى تحويل قيم المتغير العشوائي X الى شكل آخر يدعى بـ « الشكل المعياري » او « الشكل القياسي » لهذه القيم ، وخصوصاً في موضوعي الاحتمالات والاختبارات الاحصائية المتعلقة ببعض التوزيعات ، كالتوزيع الطبيعي مثلاً ، التي سيأتي ذكرها لاحقا . كذلك نحتاج الى الشكل المعياري للقيم في حالات المقارنة بين مجموعتين او اكثر من القيم ذات الاوساط الحسابية والانحرافات المعيارية المختلفة اقيامها .

فاذا كانت $X_1, X_2, \ldots, X_n$ تمثل قياسات عينة من المفردات قوامها n وان $\bar{X}$, S يمثلان على التوالي الوسط الحسابي والانحراف المعياري لقيم هذه العينة عندئذ تعرف الدرجة المعيارية Z_i لاية قيمة من قيم X على النحو التالي :

$$Z_i = \frac{X_i - \bar{X}}{S} , i = 1 , 2 , 3 , \ldots , n$$

ويلاحظ من هذا التعريف ان الدرجة المعيارية خالية من وحدات قياس المتغير الاصلي X .

اما في حالة التوزيعات التكرارية فان X_i تمثل مراكز الفئات وان $\bar{X}$ ، S يمثلان على التوالي الوسط الحسابي والانحراف المعياري للتوزيع .

وفيما يلي خصائص الدرجات المعيارية :

١ ـ ان الوسط الحسابي للدرجات المعيارية مساوٍ للصفر .

البرهان :

لتكن $Z_1, Z_2, \ldots, Z_n$ تمثل الدرجات المعيارية المقابلة لقيم X_i في المجموعة . وليكن $\bar{Z}$ يمثل الوسط الحسابي لهذه الدرجات عندئذ :

$$Z_i = \frac{X_i - \bar{X}}{S} \rightarrow \frac{\sum_{i=1}^{n} Z_i}{n} = \frac{\sum_{i=1}^{n} (X_i - \bar{X})}{n}$$

لكن $\sum_{i=1}^{n} (X_i - \bar{X}) = 0$ (لماذا) ، وان $\sum_{i=1}^{n} Z_i / n = \bar{Z}$

عليه نستنتج ان $\bar{Z} = 0$.

ونترك برهنة هذه الخاصية في التوزيعات التكرارية كتمرين للقارىء .

٢ ـ ان تباين الدرجات المعيارية مساوٍ للواحد .

البرهان :

لتكن $Z_1, Z_2, \ldots, Z_n$ تمثل الدرجات المعيارية المقابلة لقيم X_i في المجموعة . عندئذ :

$$S_z^2 = \frac{\sum_{i=1}^{n} (Z_i - \bar{Z})^2}{n} = \frac{\sum_{i=1}^{n} Z_i^2}{n} \cdot (\bar{Z} = 0)$$

وحيث ان

$$Z_i = \frac{X_i - \bar{X}}{S} \rightarrow Z_i^2 = \frac{(X_i - \bar{X})^2}{S^2}$$

$$\therefore S_z^2 = \frac{1}{n} \cdot \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{S^2} = \frac{nS^2}{nS^2} = 1$$

طالما ان

$$\sum_{i=1}^{n} (X_i - \bar{X})^2 = nS^2$$

ونترك برهنة هذه الخاصية في التوزيعات التكرارية كتمرين للقاريء .

مثال : كانت درجات خمسة طلاب في مادتي الرياضيات والاحصاء هي :

تسلسل الطالب :	1	2	3	4	5
درجة الرياضيات X :	50	52	69	72	81
درجة الاحصاء Y :	62	77	82	85	86

يطلب حساب الدرجات المعيارية لكل موضوع .

الحل :

ان الوسط الحسابي والانحراف المعياري الى X هي $\bar{X} = 64.8 , S_x = 11.96$

وان الوسط الحسابي والانحراف المعياري الى Y هي $\bar{Y} = 78.4 , S_y = 8.78$

$\therefore$ الدرجات المعيارية لكل موضوع هي :

Z_x : -1.24 , -1.07 , 0.35 , 0.60 , 1.36

Z_y : -1.87 , -0.16 , 0.41 , 0.75 , 0.87

لاحظ هنا مثلاً ولاغراض المقارنة ان الطالب الخامس كانت درجته في الاحصاء (86) افضل منها في الرياضيات (81) كمقارنة مطلقة . الا انه من حيث الاهمية فان درجته في الرياضيات هي في الحقيقة افضل منها في الاحصاء حيث ان الدرجة المعيارية في الرياضيات كانت أعلى من الدرجة المعيارية في الاحصاء .

كذلك ومن خلال هذا المثال يمكن التحقق بان $\bar{Z} = 0$ و $S_z^2 = 1$ لكل موضوع . ونترك هذه العملية كتمرين للقارىء .