ثالثا) التباين (variance)

ان مجموع انحرافات أي عينة عن وسطها الحسابي يساوي صفرا. وبدلا من اخذ القيمة المطلقة للانحرافات (أي بدون اشارات) فإننا نستطيع ان نتغلب على ذلك بتربيع قيم الانحرافات وبذلك تصبح جميعها موجبة أي نحصل على مجموع الانحرافات (summation of squares) والتي يرمز لها بالرمز **(S S)** وعلى ذلك فان: -

$$S\,S = \sum (Xi - \bar{X})^2 \quad \text{----------------} \; (1)$$

ولكي نأخذ حجم العينة في الاعتبار حتى نتمكن من مقارنة العينات المختلفة فإننا **نقسم مجموع الانحرافات على درجات الحرية** (degree of freedom) والتي تمثل بـ $(n - 1)$ وبذلك نحصل على ما يسمى التباين (S^2).

مما سبق يمكن تعريف التباين كما يلي: -

١- بيانات غير مبوبة: -

إذا كانت لدينا (n) من المشاهدات $(X_1, X_2, X_3, \ldots, X_n)$ **فان التباين والذي يرمز له بالرمز (S^2) يكون كالاتي:**

$$S^2 = \frac{\sum_{i=1}^{n}(xi - \bar{X})^2}{n-1} = \frac{\sum_{i=1}^{n} Xi^2 - \frac{(\sum_{i=1}^{n} Xi)^2}{n}}{n-1} \quad \text{----------------} (2)$$

ويلاحظ ان هذا القانون هو لحساب تباين العينة اما إذا كانت قيم المشاهدات تمثل المجتمع كله فان التباين للمجتمع ويرمز له $(^2\sigma)$ ويلفظ (سيكما سكوير) ويحسب كالاتي: -

$$^2\sigma = \frac{\sum_{i=1}^{n}(xi - \bar{M})^2}{N} \quad \text{---------------} (3)$$

حيث ان: -

M = الوسط الحسابي للمجتمع

N = عدد المفردات للمجتمع.

ومن الملاحظ في حالة إيجاد تباين العينة نقسم على $(n - 1)$ أي على درجة الحرية وهو ما يعني ان مجموع الانحرافات عن الوسط الحسابي يساوي صفرا. لذلك فعند سحب عينة واحدة فان $(n - 1)$ من المشاهدات هي قيم حرة.

رابعا) الانحراف القياسي (standard deviation)

ويسمى أيضا بالخطأ القياسي او الانحراف المعياري وهو اخذ الجذر التربيعي للتباين وذلك لإرجاع قيمة التباين الى وحداته الاصلية أي ان: -

$$S = \sqrt{S^2} = \sqrt{\frac{\sum_{i=1}^{n}(xi - \bar{X})^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^{n} Xi^2 - \frac{(\sum_{i=1}^{n} Xi)^2}{n}}{n-1}} \quad \text{-------------------- } (4)$$

اما الانحراف المعياري للمجتمع (σ) فهو: -

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{\sum_{i=1}^{n}(xi-\bar{M})^2}{N}} \quad \text{-------------------(5)}$$

مثال/ البيانات الاتية تبين عدد الأساتذة في خمسة اقسام. احسب التباين والانحراف القياسي لهذه القيم.

Xi = 9, 8, 6, 5, 7

الحل: -

يمكن حل هذا السؤال بطريقتين وكما يلي: -

أولا) الطريقة الأولى: -

١- نحسب الوسط الحسابي للقيم. حسب القانون: -

$$\bar{X} = \frac{\sum_{i=1}^{n} Xi}{n} = \frac{9+8+6+5+7}{5} = 7$$

٢- نطرح كل قيمة من قيم المشاهدات من وسطها الحسابي.

٣- نربع القيم المطروحة من الوسط. وكما مبين في الجدول التالي.

Xi	$\bar{X}$	$(Xi - \bar{X})$	$(Xi - \bar{X})^2$
9	7	2	4
8	7	1	1
6	7	-1	1
5	7	-2	4
7	7	0	0
$\sum_{i=1}^{5} Xi = 35$			$\sum_{i=1}^{5}(Xi - \bar{X})^2 = 10$

٤- نطبق القانون: -

$$S^2 = \frac{\sum_{i=1}^{n}(xi-\bar{X})^2}{n-1} = \frac{10}{5-1} = \frac{10}{4} = 2.5$$ (التباين)

$$S = \sqrt{S^2} = \sqrt{2.5} = 1.58$$ (الانحراف القياسي)

ثانيا) الطريقة الثانية

١- نربع قيم المشاهدات مباشرة ونحسب مجموعهما. وكما مبين في الجدول التالي: -

---

Xi^2	Xi
81	9
64	8
36	6
25	5
35	7
$\sum_{i=1}^{5} Xi^2 = 225$	$\sum_{i=1}^{5} Xi = 35$

٢- نطبق القانون الاتي: -

(التباين)

$$S^2 = \frac{\sum_{i=1}^{n} Xi^2 - \frac{(\sum_{i=1}^{n} Xi)^2}{n}}{n-1} = \frac{225 - \frac{35}{5}}{4} = 2.5$$

(الانحراف القياسي)

$$S = \sqrt{S^2} = \sqrt{2.5} = 1.58$$

مقاييس الارتباط (Measures of Correlation)

سبق ان درسنا بعض المقاييس الإحصائية مثل مقاييس النزعة المركزية ومقاييس التشتت والتي كانت تستخدم في وصف توزيع واحد بصورة منفصلة في علاقته بالتوزيعات الأخرى. بينما مقاييس الارتباط تتطلب طرق إحصائية لدراسة العلاقة بين متغيرين او توزيعين بدلا من دراسة خصائص توزيع واحد ومن هذه الدراسات: -

١- علاقة مستوى الذكاء بالميول المهنية.
٢- علاقة تحصيل الطالب بذكائه.
٣- علاقة المستوى الاقتصادي بالتطور الاجتماعي لمجتمع معين.

وهذه الدراسات تأخذ الابعاد الآتية: -

- إذا كانت قيمة المتغير الأول عالية وقيمة المتغير الثاني عالية تكون العلاقة موجبة او طردية.
- إذا كانت قيمة المتغير الأول واطئة وقيمة المتغير الثاني واطئة تكون العلاقة ايضا موجبة او طردية.
- إذا كانت قيمة المتغير الأول عالية وقيمة المتغير الثاني واطئة او العكس تكون العلاقة سالبة او عكسية.
- إذا كانت قيم المتغيرين غير واضحة الاتجاه لا توجد علاقة بين المتغيرين.

وهناك طريقة لتفسير تلك العلاقات بين أي متغيرين تقاس بمعامل هو معامل الارتباط ويرمز له بالرمز (r) ويتخذ قيم عددية محصورة ضمن المدى (١،-١) او كالاتي: -

$$-1 \le r \le 1$$

فاذا كانت قيمة (r) أكبر او أصغر من هذه الحدود فهذا يدل على وجود خطأ حسابي.

ملاحظات: -

١- إذا كانت قيمة $(r = 1)$ فان العلاقة موجبة تامة او طردية تامة.
٢- إذا كانت قيمة $(r = -1)$ فان العلاقة سالبة تامة او عكسية تامة.
٣- إذا كانت قيمة $(r = 0)$ فانه لا توجد علاقة بين المتغيرين.
٤- إذا كانت قيمة $(-1 \le r \le 0)$ فان العلاقة سالبه او عكسية.
٥- إذا كانت قيمة $(0 \le r \le -1)$ فان العلاقة موجبه او طردية تزداد قوتها كلما اقتربنا من واحد صحيح.

أنواع معاملات الارتباط: -

١- معامل ارتباط بيرسون (معامل الارتباط الخطي البسيط): -

يستخدم إذا كان (x, y) متصلين او مستمرين على شكل ارقام او قيم عددية والعلاقة بينهما علاقة خطية. ويحسب من العلاقة التالية: -

$$r = \frac{n\sum_{i=1}^{n} XiYi - \sum_{i=1}^{n} Xi * \sum_{i=1}^{n} Yi}{\sqrt{[n\sum (Xi)^2 - (\sum Xi)^2] * [n\sum (Yi)^2 - (\sum Yi)^2]}}$$

مثال/ احسب معامل ارتباط بيرسون للبيانات التالية: -

$Xi = 1, 2, 3, 4, 5$　　　　$Yi = 3, 6, 9, 12, 15$

الحل: -

Xi	Yi	Xi^2	Yi^2	Xi * Yi
1	3	1	9	3
2	6	4	36	12
3	9	9	91	27
4	12	16	144	48
5	15	25	225	75
15	45	55	495	165

نطبق القانون: -

$$r = \frac{n\sum_{i=1}^{n} XiYi - \sum_{i=1}^{n} Xi * \sum_{i=1}^{n} Yi}{\sqrt{[n\sum (Xi)^2 - (\sum Xi)^2] * [n\sum (Yi)^2 - (\sum Yi)^2]}}$$

$$r = \frac{5*165 - 15*45}{\sqrt{[5*55 - (15)^2][5*495 - (465)^2]}}$$

$$r = \frac{825 - 675}{\sqrt{[270 - 225][2475 - 2025]}}$$

$$r = \frac{150}{\sqrt{50*450}}$$

$$\mathbf{r} = \frac{\mathbf{150}}{\sqrt{\mathbf{22500}}} = \frac{\mathbf{150}}{\mathbf{150}} = \mathbf{1}$$

(أي ان الارتباط إيجابي تام او طردي تام)

٢- معامل ارتباط سبيرمان (معامل ارتباط الرتب): -

اشتق هذا القانون لمعالجة حالات خاصة تعتمد على رتب القيم بدلا من استخدام القيم العددية الاصلية التي تجري معالجتها باستخدام معامل ارتباط بيرسون. والسبب في استخدام معامل ارتباط الرتب هو سهولة حسابه ولتعذر التعامل مع القيم الاصلية بدقة كافية وخاصة عندما تكون حساباتها طويلة ومعقدة ويشيع استخدام هذا المعامل عندما يكون عدد أزواج البيانات للمتغيرين قليلة نسبيا بحيث لا تزيد عن ثلاثين زوجا لهذ يهتم الباحث بالرتب للبيانات أكثر من اهتمامه بقيمها الحقيقية. ويرمز له بالرمز $(\mathbf{r_s})$ ويحسب من المعادلة التالية: -

$$\mathbf{r_s = 1 - \frac{6*\sum(di)^2}{n(n^2-1)}}$$

حيث: -

d_i = الفرق بين رتب المتغير الأول ورتب المتغير الثاني

مثال/ اوجد معامل ارتباط سبيرمان لقيم المتغيرين **(x, y)**

Xi = 5, 3, 1, 4, 2 **Yi = 2, 1, 4, 5, 3**

الحل: -

-- ----------------------

Xi	Yi	di	$(di)^2$
5	2	3	9
3	1	2	4
1	4	-3	9
4	5	-1	1
2	3	-1	1
			$\sum(di)^2 = 24$

نطبق القانون: -

$$r_s = 1 - \frac{6*\sum(di)^2}{n(n^2-1)} = 1 - \frac{6*24}{5(5^2-1)} = 1 - \frac{144}{5*24} = 1 - 1.20 = -0.20$$

اذن العلاقة عكسية او سالبة

مقارنة بين معامل ارتباط بيرسون ومعامل ارتباط سبيرمان: -

١- لا يشترط تساوي قيمة معامل ارتباط بيرسون مع معامل ارتباط سبيرمان وذلك لان هذين المعاملين مختلفان تماما من حيث الأهداف في استخدامهما والفروقات بينهما تكون طفيفة على الاغلب.
٢- معامل ارتباط بيرسون ادق من معامل ارتباط سبيرمان بسبب ان معامل ارتباط بيرسون يستخدم القيم الاصلية بينما سبيرمان يستخدم الرتب المشتقة من القيم الاصلية.
٣- سهولة حساب معامل ارتباط سبيرمان يجعله مفضلا على استخدام معامل ارتباط بيرسون الذي يتضمن عمليات حسابية معقدة.
٤- يستخدم سبيرمان البيانات سواء كانت كمية ام نوعية ترتيبية حتى وان كان أحد المتغيرين كميا والأخر نوعيا بينما لا يمكن حساب معامل الارتباط لبيرسون الا إذا كان المتغيرين كميين.

واجب / احسب معاملي ارتباط بيرسون وسبيرمان من الجدول الاتي: -

Xi	Yi
8	2
3	2
6	5
6	5
1	2
3	7
6	5
3	8