الباب الرابع

الارتباط والانحدار

(Correlation and Regression)

1.4 مقدمـــة

في الأبواب الســابقة من هذا الكتاب تم عرض بعض المقاييس الإحصائية التي تصف متغير واحد منها المتوسط الحســابي والوسيط والمنوال والوســط الهندسي كمقاييس للنزعة المركزية ، والانحــراف الربــيعي والانحــراف المعياري كمقاييس لتشتت القيم . أما في هذا الباب فسوف نقوم بدراسة العلاقة بين متغيرين أو أكثر بهدف معرفة الارتباط بين هذه المتغيرات .

ولدراسة الارتباط بين متغيرين نحتاج لمقياس يقيس لنا درجــة العلاقــة بينهما واتجاه هذه العلاقة ، فإذا وجدنا أن الزيادة في المتغير الأول تصــاحبها زيادة في المتغير الثاني وأن النقص في المتغير الأول يصــاحبه نقــص فــي المتغير الثاني نقول أنه يــوجد ارتــباط طــــردي موجــب بــين هــذين المتغيرين . أما لو كانت الزيادة في المتغير الأول يصــاحبها نـــقص فــي المنغير الثــاني وأن النقص في المتغير الأول تصاحبه زيادة في المتغير الثاني فأننا نقول أنه يوجد ارتباط عكسي أو سالب بين هذين المتغيرين .

وقد تقابلنا حالات نجد فيها أن الارتباط يكون تاماً ، سواء كان طرديــاً أو عكسياً ، وفي هذه الحالات نســتطيع معــرفة أحد المتغيــرين لــو عــــرفنا المتغير الآخر . والأمثلة على ذلك عديدة منها العلاقة بــين مســاحة الــدائرة ونصف قطرها وطول ضلع المربع ومساحته وغيرها . وقــد تقابلنــا أيضــاً حالات ينعدم فيها الارتباط مثل دراســة العلاقة بين طول الفرد ودخلــه كمــا هــو الحال في الكثير من الحالات الشائعة والتي تواجهنا كثيراً في الدراسات المختلفة والتي تتباين من الحالات التي يكون فيها الارتباط تامــاً ولا يكــون منعدماً ولكن بين هذا وذاك . مثل دراسة العلاقة بين الطول والوزن أو العلاقة بين التقدير الذي حصل عليه بعض الطلبة في مادتين وغيرها وســوف نقــوم بدراسة هذه الأمثلة بالتفصيل .

ويجب ملاحظته أن وجود ارتباط بين متغيرين لا يبين ما إذا كان أحدهما تابع للآخر فإذا كان لدينا متغيرين هما x ، y ووجدنا أن هناك بينهما ارتباطاً قوياً فإن هذا لا يوضح ما إذا كانت x تؤثر في y أو أن y تؤثر في x أم أن هناك عامل مشترك يؤثر في كل منهما وهو الذي أدى إلى زيادة الارتباط بينهما .

2.4 معامل الارتباط (Coefficient of Correlation)

يعني وجود الارتباط بين ظاهرتين أن التغير" بالنقص أو الزيادة " في أحدهما يكون مصحوباً بتغير في الظاهرة الأخرى ، ويكون هذا التغير في نفس الاتجاه في حالة الارتباط الطردي ، وفي الاتجاه المخالف في حالة الارتباط العكسي ، أي أن الارتباط يمكن قياسه بواسطة التغيرات التي تحدث في الظاهرتين .

فإذا كان لدينا المتغيرين x ، y يعبران عن ظاهرتين معينتين فإن أفضل طريقة لمقارنة التغير في هاتين الظاهرتين هي مقارنة القيم المعيارية لهما أي القيم :

$$\left(\frac{x-\bar{x}}{\delta_x}\right),\left(\frac{y-\bar{y}}{\delta_y}\right)$$

حيث أن :

S_x ، S_y – هما الانحرافان المعياريان لقيم x ، y على الترتيب .

وهنا نلاحظ أن حاصل ضرب القيم المعيارية للظاهرتين يكون كبيراً عددياً " بغض النظر عن الإشارة موجبة كانت أم سالبة " في حالة وجود ارتباط قوي بين الظاهرتين وعليه فقد أتفق على اتخاذ متوسط حاصل ضرب

القيم المعيارية كمقياس لدرجة الارتباط بين المتغيرين والذي يسمى بمعامل الارتباط (Coefficient of Correlation) ويرمز له بالرمز R حيث أن :

$$R = \frac{1}{n}\sum\left(\frac{x-\bar{x}}{S_x}\right)\left(\frac{y-\bar{y}}{S_y}\right) \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(1-4)$$

حيث أن :

n - هي عدد أزواج المفردات .

أن معامل الارتباط (R) يتميز بالخصائص التالية :

1- تتراوح قيمته العددية بين الصفر والواحد الصحيح .
2- هذا المقياس يساوي صفراً في حالة انعدام الارتباط ويساوي الواحد في حالة الارتباط التام .
3- تكون قيمة هذا المقياس موجبة حينما يكون الارتباط طردي وتكون سالبة في حالة الارتباط العكسي .
4- قيمة هذا المقياس العددية تزداد كلما ازدادت درجة الارتباط .

3.4 حساب معامل الارتباط (Calculation of Correlation Factor)

تعرف العلاقة السابقة (4-1) بمعامل بيرسون للارتباط (Pearson's Correlation) ويمكن كتابتها على الشكل التالي :

$$\therefore R = \frac{1}{n}\frac{\sum(x-\bar{x})(y-\bar{y})}{S_x . S_y} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(2-4)$$

حيث أن كلاً من S_x ، S_y مقدار ثابت ويمكن أخذه كعامل مشترك في المقام .

وهذه العلاقة وإن كانت أسهل في حسابها من العلاقة السابقة إلاّ أنها تتطلب الكثير من العمليات الحسابية ، وخاصة إذا أحتوى كل من $\bar{x}$ و $\bar{y}$ على كسور وما يترتب على ذلك من صعوبات في العمليات الحسابية ولغرض تبسيط الصيغة الأخيرة نقوم باختزال العلاقة (4 - 2) إلى العلاقة التالية :

$$R = \frac{1}{n}\frac{\sum(x-\bar{x})(y-\bar{y})}{S_x.S_y} = \frac{1}{n}\frac{\sum(x.y - x.\bar{y} - \bar{x}.y + \bar{x}.\bar{y})}{S_x.S_y}$$

$$R = \frac{1}{n}\left(\frac{\sum x.y - \sum x.\bar{y} - \sum \bar{x}.y + \sum \bar{x}.\bar{y}}{S_x.S_y}\right)$$

$$= \frac{\frac{\sum x.y}{n} - \bar{y}\frac{\sum x}{n} - \bar{x}\frac{\sum y}{n} + \frac{n\,\bar{x}.\bar{y}}{n}}{S_x.S_y}$$

$$= \frac{\frac{\sum x.y}{n} - \bar{x}.\bar{y} - \bar{x}.\bar{y} + \bar{x}.\bar{y}}{S_x.S_y} = \frac{\frac{1}{n}\sum x.y - \bar{x}.\bar{y}}{S_x.S_x}$$

$$\therefore R = \frac{\frac{\sum x.y}{n} - \bar{x}.\bar{y}}{S_x.S_y} \qquad \text{..........................}(3-4)$$

حيث أن :

$$\bar{x} = \frac{\sum x}{n}$$

$$\bar{y} = \frac{\sum y}{n}$$

وبما أن :

$$\delta_x = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$$

وكذلك :

$$\delta_y = \sqrt{\frac{\sum y^2}{n} - \left(\frac{\sum y}{n}\right)^2}$$

فأننا نحصل من خلال التعويض في المعادلة (3-4) على العلاقة التي تمكننا من حساب معامل الارتباط " معامل بيرسون " بطريقة سهلة حيث :

$$R = \frac{\frac{\sum x.y}{n} - \bar{x}.\bar{y}}{\sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}\sqrt{\frac{\sum y^2}{n} - \left(\frac{\sum y}{n}\right)^2}} \qquad \ldots\ldots\ldots\ldots(4-4)$$

مثـــال (4-1)

أحســب معامل ارتباط بيرســون بين قيم المتغيرين x و y من البيانـــات المبينة في الجدول (4-1) :

جــدول (4-1)

65	68	62	70	66	67	64	68	71	69	x
28	29	26	28	25	28	25	31	30	28	y

الحــل :

نقوم لحساب معامل الارتباط باستخدام العلاقة (4-4) ، حيث يلزمنا معرفة كل من المجاميع الاتية :

$\sum x$ ، $\sum y$ ، $\sum x^2$ ، $\sum y^2$ و $\sum xy$

وهذه يمكن حسابها من خلال الجــدول (4 -2) .

جــدول (4-2)

إيجاد معامل الارتباط بين المتغيرين x و y

x	y	x^2	y^2	x y
69	28	4761	784	1932
71	30	5041	900	2130
68	31	4624	961	2108
64	25	4096	625	1600
67	28	4489	784	1876
66	25	4356	625	1650
70	28	4900	784	1960
62	26	3844	676	1612
68	29	4624	841	1972
65	28	4225	784	1820
$\sum x = 670$	$\sum y = 278$	$\sum x^2 = 44960$	$\sum y^2 = 7764$	$\sum xy = 18660$

$$\therefore \bar{x} = \frac{670}{10} = 67 \quad \bar{y} = \frac{287}{10} = 27.8$$

$$S_x = \sqrt{\frac{44960}{10} - \left(\frac{670}{10}\right)^2} = \sqrt{4496 - 4489} = \sqrt{7}$$

$$S_y = \sqrt{\frac{7764}{10} - \left(\frac{278}{10}\right)^2} = \sqrt{776.40 - 772.84} = \sqrt{3.56}$$

$$\therefore R = \frac{\frac{18660}{10} - (67)(27.8)}{(\sqrt{7})(\sqrt{3.56})} = \frac{1866 - 1862.6}{\sqrt{24.92}} = \frac{3.4}{4.99} = 0.68$$

ويطلق على معامل ارتباط بيرسون أيضاً معامل الارتباط العزمي ، ويمكن لحسابه استخدام العلاقة التالية :

$$R = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{\left[n\sum x^2 - (\sum x)^2\right]\left[n\sum y^2 - (\sum y)^2\right]}} \ldots\ldots\ldots\ldots\ldots\ldots(5-4)$$

ويشـــترط في اســـتخدام القانون أن تكون العلاقة بين x و y خطيـــة مـــع الأخذ بعين الاعتبار الملاحظات التالية :

1- في حالة كون معامل الارتباط سالب فإن العلاقة بين المتغيرين تكون سالبة تامة . إذا زاد أحد المتغيرين نقص الثاني بمقدار ثابت .

2- في حالة كون معامل الارتباط موجب فإن العلاقة بين المتغيـــرين تكـــون موجبة تامة .

3- أما في حالة كون معامل الارتباط = صفر فإن العلاقة بين المتغيرين تكون معدومة ويجب أن يكون معامل الارتباط محصوراً بين (− 1 إلى + 1) .

مثـــال (4 -2)

أحســـب معامل الارتباط العزمي لبيرسون للمتغيرين X و Y والمبينة في الجدول (4-3) :

جــدول (4-3)

X	9	8	7	6	5
Y	1	2	3	4	5

الحـــل:

باستخدام العلاقة (4-5) ننظم الجدول (4-3) كما هو مبين في الجدول (4-4) :

جــدول (4-4)

X	Y	X . Y	X^2	Y^2
9	1	9	81	1
8	2	16	64	4
7	3	21	49	9
6	4	24	36	16
5	5	25	25	25
$\sum X = 35$	$\sum Y = 15$	$\sum XY = 95$	$\sum X^2 = 255$	$\sum Y^2 = 55$

$$\therefore R = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{\left[n\sum x^2 - (\sum x)^2\right]\left[n\sum y^2 - (\sum y)^2\right]}}$$

$$\therefore R = \frac{5(95) - (35)(15)}{\sqrt{\left[(5)(255) \quad (35)^2\right]\left[(5)(55) \quad (15)^2\right]}} = -1$$

إذن العلاقة بين المتغيرين سالبة وتامة .

مثـــال (4 -3)

أحســـب معامـــل الارتبـــاط بـــين المتغيـــرين x , y الموضـــحين فـــي الجدول (4-5) :

جـــدول (4-5)

X	1	2	3	4	5
Y	2	1	1	3	2

الحــل:

نضع البيانات الموجودة في الجدول(4-5) بالشكل العمودي الموضح فـــي الجدول (4-6) :

جـــدول (4-6)

X	Y	X Y	X^2	Y^2
1	3	3	1	9
2	1	2	4	1
3	1	3	9	1
4	3	12	16	9
5	2	10	25	4
$\sum X = 15$	$\sum Y = 10$	$\sum XY = 30$	$\sum X^2 = 55$	$\sum Y^2 = 24$

$$\therefore R = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{\left[n\sum x^2 - (\sum x)^2\right]\left[n\sum y^2 - (\sum y)^2\right]}}$$

$$\therefore R = \frac{5(30) - (15)(10)}{\sqrt{\left[(5)(55) - (15)^2\right]\left[(5)(24) - (10)^2\right]}} = 0$$

وعلى هذا الأساس فإن العلاقة تكون معدومة بين المتغيرين .

4.4 إيجـــاد معامـــل الارتبـــاط باستخـــدام الوســـط الفرضـــي
(Calculation of Correlation Factor Using Assumed Mean)

أن العلاقة التي تم استخدامها في حل المثال (4-1) تتطلب الكثيــر مــن العمليات الحسابية ، وأن الحل سيكون أسهل بكثيــر إذا اســتخدمنا وســطين فرضيين للقيم x و y فإذا حسبنا انحرافات (x) عن الوســط الفرضـــي (a) وانحرافات (y) عن الوسط الفرضي(b) سنجد أن :

$$d_X = (x - a)$$
$$d_Y = (y - b)$$

وقد سبق أن بينا أن البسط في صيغة معامل الارتباط هو :

$$\frac{1}{n}\sum(x-\bar{x})(y-\bar{y})$$
$$= \frac{1}{n}\sum\left[(x-a)-(\bar{x}-a)\right]\left[(y-b)-(\bar{y}-b)\right]$$
$$= \frac{1}{n}\sum(d_X - \bar{d}_X)(d_Y - \bar{d}_Y) \qquad \ldots\ldots\ldots\ldots..(6-4)$$

حيث أن :

$$\bar{d}_X = \bar{x} - a$$
$$\bar{d}_Y = \bar{y} - b$$

وعلى الطالب مراجعة علاقة حساب المتوسط الحسابي للبيانـــات المبوبـــة باستخدام الوسط الفرضي التي تم شرحها بالتفصيل في الباب الثاني . والنتيجة الأخيرة لبسط معامل الارتباط يمكن كتابتها على النحو التالي :

$$\frac{\sum d_X . d_Y}{n} - \bar{d}_X . \bar{d}_Y$$

وبالتالي نحصل على العلاقة التي تمكننا من حساب معامل الارتباط باستخدام وسطين فرضيين لقيم x وقيم y وهي بالشكل التالي :

$$\therefore R = \frac{\frac{\sum d_X . d_Y}{n} - \sum \bar{d}_X . \bar{d}_Y}{\delta_X . \delta_Y} \quad \text{.........................} (7-4)$$

حيث أن :

$$\frac{\sum d_X}{n} = \bar{d}_X$$

$$\frac{\sum d_Y}{n} = \bar{d}_Y$$

أما الانحراف المعياري فيساوي :

$$S_X = \sqrt{\frac{\sum d_X^2}{n} - \left(\frac{\sum d_X}{n}\right)^2} \quad \text{............................} (8-4)$$

$$S_Y = \sqrt{\frac{\sum d_Y^2}{n} - \left(\frac{\sum d_Y}{n}\right)^2} \quad \text{............................} (9-4)$$

ويمكن حساب معامل الارتباط بين قيم x ، y في المثال السابق باستخدام هذه العلاقة . وذلك باعتبار القيمة 68 كوسط فرضي لقيم x والقيمة 28 كوسط فرضي لقيم y ، وقد اخترنا هاتين القيمتين نظراً لتكرارهما مما يسهل الحل كما هو مبين في الجدول (4-7) .

جــدول (7-4)

إيجاد معامل الارتباط بين قيم x ، y باستخدام وسطين فرضيين

$d_X d_Y$	d_Y^2	d_X^2	$d_Y (y-28)$	$d_X (x-68)$	y	x
0	0	1	0	1	28	69
6	4	9	2	3	30	71
0	9	0	3	0	31	68
12	9	16	3 -	4 -	25	64
0	0	1	0	1 -	28	67
6	9	4	3 -	2 -	25	66
0	0	4	0	2	28	70
12	4	36	2 -	6 -	26	62
0	1	0	1	0	29	68
0	0	9	0	3 -	28	65
36	36	80	2 -	10 -	المجموع	

$$\therefore \bar{d}_X = \frac{\sum d_X}{n} = \frac{10 -}{10} = -1$$

$$\therefore \bar{d}_Y = \frac{\sum d_Y}{n} = \frac{2 -}{10} = -0.2$$

$$\therefore S_X = \sqrt{\frac{\sum d_X^2}{n} - \left(\frac{\sum d_X}{n}\right)^2} = \sqrt{\frac{8}{10} - \left(\frac{10-}{10}\right)^2}$$

$$= \sqrt{8-1} = \sqrt{7}$$

$$S_Y = \sqrt{\frac{\sum d_Y^2}{n} - \left(\frac{\sum d_Y}{n}\right)^2} = \sqrt{\frac{36}{10} - \left(\frac{2-}{10}\right)^2}$$

$$= \sqrt{3.60 - 0.04} = \sqrt{3.56}$$

$$\therefore R = \frac{\frac{\sum d_X . d_Y}{n} - \sum \bar{d}_X . \bar{d}_Y}{S_X . S_Y} = \frac{\frac{36}{10} - (-1)(-0.2)}{(\sqrt{7})(\sqrt{3.56})}$$

$$= \frac{3.6 - 0.2}{\sqrt{24.92}} = \frac{3.4}{4.99} = 0.68$$

وهي نفس النتيجة التي حصلنا عليها من قبل مع ملاحظة قلة العمليات الحسابية ، ولذلك يجب استخدام هذه العلاقة الأخيرة لتسهيل العمل الحسابي ، فإذا طلب منا حساب معامل بيرسون للارتباط فيجب استخدام هذه الصيغة نظراً لسهولة حسابها .

مثـال (4 -4)

أحسب معامل ارتباط بيرسون بين قيم x , y من البيانات المبينة في الجدول (4-8) :

جـدول (4-8)

164	154	170	169	170	170	165	**x**
62	56	83	65	72	70	61	**y**

الحـــل:

باستخدام القيمة 170 كوسط فرضي لقيم x والقيمة 65 كوسط فرضي لقيم y ومن الجدول (4-8) يمكن تكوين جدول (4-9) الخاص بالمثال على بالشكل التالي :

جـــدول (4-9)

$d_X d_Y$	d_Y^2	d_X^2	$d_Y\,(y-28)$	$d_X\,(x-68)$	y	x
20	16	25	4 -	5 -	61	165
0	25	0	5	0	70	170
0	49	0	7	0	72	170
0	0	1	0	1 -	65	169
0	324	0	18	0	83	170
144	81	256	9 -	16 -	56	154
18	9	36	3 -	6 -	62	164
182	504	318	14	28 -	المجموع	

$$\therefore \bar{d}_X = \frac{\sum d_X}{n} = \frac{-28}{7} = -4$$

$$\therefore \bar{d}_Y = \frac{\sum d_Y}{n} = \frac{14}{7} = 2$$

$$\therefore S_X = \sqrt{\frac{\sum d_X^2}{n} - \left(\frac{\sum d_X}{n}\right)^2} = \sqrt{\frac{318}{7} - \left(\frac{28-}{7}\right)^2}$$

$$= \sqrt{45.43 - 16} = \sqrt{29.43}$$

$$S_Y = \sqrt{\frac{\sum d_Y^2}{n} - \left(\frac{\sum d_Y}{n}\right)^2} = \sqrt{\frac{504}{7} - \left(\frac{14}{7}\right)^2}$$

$$= \sqrt{72 - 4} = \sqrt{68}$$

$$\therefore\ R = \frac{\frac{\sum d_X . d_Y}{n} - \sum \bar{d}_X . \bar{d}_Y}{\delta_X . \delta_Y} = \frac{\frac{182}{7} - (-4)(2)}{(\sqrt{29.43})(\sqrt{68})}$$

$$= \frac{26 + 8}{\sqrt{2001.24}} = \frac{34}{44.55} = 0.76$$

إذن هناك ارتباط طردي قوي .

5.4 حساب معامل الارتباط للبيانات المبوبة

(Correlation Factor for Tabulated Data)

عندما يكبر حجم العينة يكون من الصعب حساب معامل الارتباط بالطريقة السابقة ، لذلك يمكن تبويب البيانات في شكل جدول توزيع تكراري مزدوج والذي تم عرضه في الباب الأول من هذا الكتاب . ومن ثم يتم إيجاد معامل الارتباط باستخدام العلاقة التالية :

$$R = \frac{\frac{\sum (d_X . d_Y).f}{n} - \bar{d}_X . \bar{d}_Y}{S_X . S_Y} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (10-4)$$

حيث أن :

$$n = \sum f_i$$

$$\bar{d}_X = \frac{\sum d_X . f}{\sum f}$$

$$\bar{d}_Y = \frac{\sum d_Y . f}{\sum f}$$

أما :

$$S_X = \sqrt{\frac{(\sum d_X^2) f_i}{\sum f} - \left(\frac{\sum d_X . f_i}{\sum f}\right)^2} \qquad ....(11-4)$$

$$S_Y = \sqrt{\frac{(\sum d_Y^2) f_i}{\sum f} - \left(\frac{\sum d_Y . f_i}{\sum f}\right)^2} \qquad ....(12-4)$$

حيث أن :

d_X ، d_Y – هي انحرافات مراكز الفئات للمتغيرين x و y عـن وسـطيهما الفرضيين .

وفي حالة تساوي أطوال فئات كل من المتغيرين فإن استخدام الانحرافات المختصرة بدلاً من الانحرافات لن يؤثر على النتيجة وبذلك يمكن استعمال العلاقة التالية :

$$R = \frac{\frac{\sum (\bar{d}_X . \bar{d}_Y) . f}{\sum f} - \left(\frac{\bar{d}_X . f_X}{\sum f_X}\right)\left(\frac{\bar{d}_Y . f_Y}{\sum f_Y}\right)}{\sqrt{\frac{\sum (\bar{d}_X^2 . f_X)}{\sum f_X} - \left(\frac{\sum \bar{d}_X . f_X}{\sum f_X}\right)^2} \sqrt{\frac{(\bar{d}_Y^2 . f_Y)}{\sum f_Y} - \left(\frac{\sum \bar{d}_Y . f_Y}{\sum f_Y}\right)^2}} \quad ......(13-4)$$

مثـال (4-5)

جدول (4-10) التكراري المزدوج يمثل العلاقة بين الطول والوزن لعينة مكونة من 100 طالب من إحدى المدارس . أحسب معامل ارتباط بيرسون لتلك البيانات .

جـــدول (4-10)

الوزن \ الطول	150 -	155 -	160 -	165 -	170 - 175	المجموع
40 -	5	3				8
50 -	1	12	14			27
60 -		2	28	22		52
70 - 80			3	8	2	13
المجموع	6	17	45	30	2	100

الحـــل:

نلاحظ هنا تساوي أطوال الفئات بالنسبة للطول (5cm) والوزن (10kg) وبذلك يمكننا اســتخدام الصيغة الســابقة . ولحسابها نضيف إلى الجدول سبعة صفوف وسبعة أعمدة كما في الجدول (4–11) . ففي العمــود الأول نحســب مراكز فئات المتغير(y) ، ثم نختار مــن بينهــا وسـطاً فرضـياً ونحســب الانحرافات عنه d_Y ، وفي العمود التالي وحيث أن أطوال الفئــات متســاوية نقسم كل من الانحرافات d_Y على طول الفئة فنحصل على $\bar{d}_Y$ في العمــود الثالث .

ونلاحظ أنها تساوي صفر أمام الفئــة التــي اختيــر مركزهــا كوســط فرضي −1 ، − 2..... في الفئات السابقة لها ،1 ، 2 ، 3 .... فــي الفئــات اللاحقة لها ، وفي العمود الرابع نضرب كل انحراف مختصر فــي التكــرار المناظر له فنحصل على $\bar{d}_Y \cdot f_i$ وفي العمود الخامس نضرب كل انحــراف مختصر $\bar{d}_Y$ (العمود الثــالث) في $\bar{d}_Y \cdot f_Y$ (العمود الرابع) فنحصــل علــى

$\overline{d}_Y^2 . f_Y$ وبتكرار نفس الخطوات بالنسبة للصفوف نحصل على مراكز فئات المتغير x وهي :

$$\overline{d}_X^2 . f_X , \overline{d}_X . f_X , \overline{d}_X , d_X$$

في الخمسة صفوف التي أضفناها ، والبيانات التي حسبناها حتى الآن تكفي لحساب المقام في صيغة معامل الارتباط .

وقبل حساب بيانات العموديين السادس والسابع ننقل بيانات الانحرافات المختصرة $\overline{d}_Y$ خارج الجدول الأصلي (إلى يمين الفئات y) والانحرافات المختصرة $\overline{d}_X$ خارج الجدول (فوق الفئات x) ونحسب قيم العمود السادس بأن نضرب التكرارات في كل صف (من الجدول الأصلي) في الانحرافات المناظرة والتي كتبناها أعلى الجدول ثم نجمع النتائج . فأول قيمة في العمود السادس تكون :

$$13 - = (1-)(3) + (2 -)(5)$$

وثاني قيمة تكون :

$$14 - = (\text{صفر})(14) + (1-)(12) + (2 -)(1)$$

وبذلك نحصل على بيانات العمود السادس التي تعطي لنا $\sum \overline{d}_Y . f_Y$ أما بيانات العمود السابع فنحصل عليها بضرب كل قيمة من قيم العمود السادس في الانحراف المناظر، وهو $\overline{d}_X$ فنحصل على $\sum \overline{d}_X . \overline{d}_Y . f$ وبذلك يمكننا حساب معامل الارتباط الآن . أما الصفين السادس والسابع فيمكن تكملة الحل بدونهما ولكن الهدف منهما هو التأكد من صحة العمليات الحسابية السابقة . فالصف السادس نحسبه بنفس الطريقة

التي حسبت بها بيانات العمود السادس ، ونحصل على قيمته بأن نضرب التكرارات في كل عمود من الجدول الأصلي في الانحرافات المناظرة والتي كتبناها إلى يمين الجدول ثم نجمع النتائج . فأول قيمة في الصف السادس تكون :

11−=(1 −)(1)+(2 −)(5)

وثاني قيمة فيه تكون:

18 −=(صفر)(2)+(1−)(12)+(1−)(12) + (2−)(3)

وبذلك نحصل على بيانات الصف السادس وتعطي لنا :

$$\sum \bar{d}_X . f_X$$

أما بيانات الصف السابع فنحصل عليها بضرب كل قيمة من قيم الصف السادس في الانحراف المناظر $\bar{d}_Y$ ونكون بذلك قد حصلنا على :

$$\sum \bar{d}_X . \bar{d}_Y . f$$

نلاحظ هنا أن مجموع الصف الرابع لا بد وأن يساوي مجموع العمود السادس وأن مجموع الصف السادس لا بد وأن يساوي مجموع العمود الرابع وأن مجموع الصف السابع سوف يساوي مجموع العمود السابع كما يتضح من الأسهم في الجدول (4–11) ، وبعد تكملة الجدول بالشكل الذي شرحناه يمكن حساب معامل الارتباط :

جدول (4 -11) إيجاد معامل الارتباط من الجدول التكراري المزدوج

	2 -	1 -	0	1	2	
X / Y	- 150	- 155	- 160	- 165	175 - 170	f_Y
- 40	5	3				8
- 50	1	12	14			27
- 60		2	28	22		52
80 - 70			3	8	2	13
f_X	6	17	45	30	2	100

مركز الفئة x	152.5	157.5	$\boxed{162.5}$	167.5	172.5	
d_X	10 -	5 -	0	5	10	----
$\bar{d}_X$	2 -	1 -	0	1	2	-----
$\bar{d}_X \cdot f_X$	12 -	17 -	0	30	4	5
$\bar{d}_X^2 \cdot f_X$	24	17	0	30	8	79
$\sum \bar{d}_Y \cdot f$	11 -	18 -	11 -	8	2	30 -
$\sum \bar{d}_X \cdot \bar{d}_Y \cdot f$	22	18	0	8		52

مركز الفئة y	d_Y	$\bar{d}_Y$	$\bar{d}_Y \cdot f_Y$	$\bar{d}_Y^2 \cdot f_Y$	$\sum \bar{d}_X \cdot f$	$\sum \bar{d}_X \cdot \bar{d}_Y \cdot f$
55	20 -	2 -	16 -	32	13 -	26
68	10 -	1 -	27 -	27	14 -	14
$\boxed{78}$	0	0	0	0	20	0
85	10	1	13	13	12	12
----			30 -	72	5	52

$$\therefore R=\frac{\frac{52}{100}-\left(\frac{5}{100}\right)\left(\frac{30-}{100}\right)}{\sqrt{\frac{79}{100}-\left(\frac{5}{100}\right)^2}\sqrt{\frac{72}{100}-\left(\frac{30-}{100}\right)^2}}$$

$$=\frac{0.52+(0.05)(0.3)}{(\sqrt{0.7875})(\sqrt{0.63})}$$

$$=\frac{0.52+0.15}{\sqrt{0.496125}}$$

$$=\frac{0.535}{0.7043}= \ 0.76$$

إذن يوجد ارتباط طردي قــوي بــين الطــول والــوزن لعينــة الطلبــة المدروسة . وتجد الإشارة إلى أنه يمكن اختصار الحل في حالة تساوي أطوال الفئات وذلك بعدم إضافة كل من العموديين الأول والثــاني والصــفين الأول والثاني في الجدول (4-11) ، وذلك بأن نكتب الانحرافات المختصرة مباشرة بوضع صفر أمام الفئة التي كنا سنختار مركزها كوسط فرضي (ويفضــل أن تكون في منتصف الجدول وأمــام أكبــر تكــرار) ثــم نكتــب − 1,− 2 ... الانحرافات المختصــرة للفئــات الســابقة لهــا, 1 ، 2 ، 3 ...للانحرافــات المختصرة للفئات اللاحقة لها ثم نكمل الحل .

مثــال (4 -6)

أحسب معامل الارتباط بين قيم x , y من البيانات المبينــة فــي الجــدول (12-4) .

الجــدول (4-12)

Y \ X	- 40	- 50	- 60	- 70	90 - 80	المجموع
- 40	6	—	—	—	—	6
- 50	6	12	4	—	—	22
- 60	—	10	20	2	—	42
80 - 70	—	—	12	16	2	30
المجموع	12	22	46	18	2	100

الحــل :

نضيف إلى الجدول خمسة صفوف وخمسة أعمدة ونتبع نفــس خطــوات المثال السابق فنجد أن :

$$\sum \bar{d}_X . \bar{d}_Y . f = 68 \;,\; \sum \bar{d}_X . f_X = -24 \;,\; \sum \bar{d}_Y . f_Y = -4$$

وأن :

$$\sum \bar{d}_Y^2 . f_Y = 76 \;,\; \sum \bar{d}_X^2 . f_X = 96$$

كما مبين في الجدول(4 – 13) .

$$\therefore\ R = \frac{\frac{68}{100} - \left(\frac{24\ -}{100}\right)\left(\frac{4\ -}{100}\right)}{\sqrt{\frac{96}{100} - \left(\frac{24\ -}{100}\right)^2}\sqrt{\frac{76}{100} - \left(\frac{4\ -}{100}\right)^2}}$$

$$= \frac{0.68 - (0.24\ -)(0.04\ -)}{(\sqrt{0.9024})(\sqrt{0.7584})}$$

$$= \frac{0.68 - 0.0096}{\sqrt{0.6838}}$$

$$= \frac{0.6704}{0.8272} = 0.81$$

جدول (4 - 13) إيجاد معامل الارتباط من الجدول التكراري المزدوج

	2 -	1 -	0	1	2	
X / Y	- 40	- 50	- 60	- 70	90 - 80	f_Y
- 40	6					6
- 50	6	12	4			22
- 60		10	30	2		42
80 - 70			12	16	2	30
f_X	12	22	46	18	2	100

$\bar{d}_Y$	$\bar{d}_Y \cdot f_Y$	$\bar{d}_Y^2 \cdot f_Y$	$\sum \bar{d}_X \cdot f$	$\sum \bar{d}_X . \bar{d}_Y . f$
2 -	12 -	24	12 -	24
1 -	22 -	22	24 -	24
0	0	0	8 -	0
1	30	30	20	20
	4 -	76	24 -	68

$\bar{d}_X$	2 -	1 -	0	1	2	----
$\bar{d}_X \cdot f_X$	12 -	22 -	0	18	4	24 -
$\bar{d}_X^2 \cdot f_X$	24 -	22	0	18	8	96
$\sum \bar{d}_Y \cdot f$	18 -	12 -	8	16	2	4 -
$\sum \bar{d}_X . \bar{d}_Y . f$	36	12	0	16	4	68

إذن هنالك ارتباط طردي قوي بين قيم x و y .

6.4 معامل ارتباط الرتب (سبيرمان)

(Spearman's Order Correlation Factor)

يستخدم هذا المعامل لدراسة الارتباط بين البيانات النوعية ، أي تلك التي لا يمكن قياسها كمياً . وتعتمد هذه الطريقة على إعطاء المتغيرات رتباً لتحل محل القياس العددي . فإذا رتبنا مفردات المتغير x ترتيباً تصاعدياً ووجدنا أن مفردات المتغير y المناظرة لها مرتبة ترتيباً تصاعدياً أيضاً نستنتج وجود ارتباط طردي تام بين المتغيرين x ، y . أما إذا رتبنا مفردات المتغير x ترتيباً تصاعدياً ووجدنا أن مفردات المتغير y المناظرة لها مرتبة ترتيباً تنازلياً نستنتج وجود ارتباط عكسي تام بين المتغيرين x ، y غير أن هذا الارتباط التام نادراً ما يصادفنا في الدراسات الاجتماعية والاقتصادية .

ولقياس معامل الارتباط بين مفردات المتغيرين x و y نرتب كلاً منهما حسب أفضليته ثم نحسب الفرق بين كل رتبتين متتاليتين فنجد أن :

$$\sum F = 0$$

وبحساب مربعات هذه الفروقات يمكن إيجاد معامل الارتباط باستخدام العلاقة :

$$R = 1 - \frac{6\sum F^2}{n(n^2-1)} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots..(14-4)$$

حيث أن :

n = عدد الرتب .

$\sum F$ = مجموع مربعات الفروقات بين الرتب .

مثـــال (4-7)

الجدول (4-14) يبين تقديرات ســتة من الطــلبة فــي امتحــان مــادتي الفيزياء والكيمياء أحسب معــامل ارتباط الرتب لســبيرمان بــين تقــديرات المادتين .

جــدول (4-14)

رقم الطالب	1	2	3	4	5	6
تقدير الفيزياء	ضعيف	ممتاز	جيد	ضعيف جداً	مقبول	جيد جداً
تقدير الكيمياء	مقبول	جيد جداً	جيد	ضعيف	ضعيف جداً	ممتاز

الحــل:

لحساب معامل الارتباط من هذه البيانات نرتب تقديرات كل من المادتين ترتيباً تصاعدياً أو تنازلياً وذلك بإعطاء التقدير ممتاز الرتبة(1) والتقدير الذي يليه الرتبة(2) و.....هكذا ثم نحسب الفروقات بين كل رتبتين متناظرتين كمــا في جدول(4 -15) فنجد أن :

$$\therefore R = 1 - \frac{6\sum F^2}{n(n^2-1)} = 1 - \frac{(6)(8)}{(6)(36-1)} = 1 - \frac{8}{35}$$

$$= \frac{35-8}{35} = \frac{27}{35} = 0.77$$

إذن يوجد ارتباط طردي قوي بين تقديرات الطلبة الستة في هاتين المادتين .

جـــدول (4-15)

حساب معامل ارتباط الرتب - سبيرمان - بين تقديرات مادتي الفيزياء والكيمياء

تقدير الفيزياء	تقدير الكيمياء	رتب تقدير الفيزياء	رتب تقدير الكيمياء	الفروق F	مربع الفروق F^2
ضعيف	مقبول	5	4	1	1
ممتاز	جيد جداً	1	2	1−	1
جيد	جيد	3	3	صفر	صفر
ضعيف جداً	ضعيف	6	5	1	1
مقبول	ضعيف جداً	4	6	2−	4
جيد جداً	ممتاز	2	1	1	1
المجمـــــــــوع					8

7.4 معامل ارتباط الرتب (سبيرمان) في حالة الرتب التكرارية
(Spearman's Correlation Coefficient for Frequent Orders)

في المثـــال الســـابق نلاحظ أنه لم تتكرر أي من التقديرات التي حصـــل عليها الطلبة . فإذا صادفنا مثالاً آخراً تتكرر فيه بعض التقديرات فإننا نعطـــي القيم المتكررة رتباً تساوي متوسط الرتب التي كانت ستعطى لو لـــم تتكـــرر التقديرات .

مثـــال (4-8)

الجدول (4-16) يبين تقديرات عشرة من الطلبـــة فـــي امتحـــان مـــادتي الإحصاء والاقتصاد والمطلوب حساب معامل الارتباط بين تقديرات المادتين .

جـدول (4-16)

رقم الطالب	1	2	3	4	5	6	7	8	9	10
تقدير الإحصاء	ضعيف جداً	مقبول	ممتاز	مقبول	ضعيف	جيد جداً	جيد	جيد	مقبول	مقبول
تقدير الاقتصاد	مقبول	جيد	جيد جداً	مقبول	جيد	مقبول	ممتاز	ضعيف جداً	ضعيف	جيد جداً

الحـــل :

عند إعطاء رتب التقدير لمادة الإحصاء سـنجد أن الطالب رقم 3 يأخـذ الرتبة (1) والطالب رقم 6 يأخذ الرتبة (2) والطالب رقم 7 يأخذ الرتبـة (3) بينما الطلبة رقم 2 , 4 , 9 , 10 لهم نفس التقدير ويستحقون الرتب (4) ، (5) ، (6) ، (7) ونظراً لتقاربهم في التقدير يعطى لكل منهم متوسط هذه الرتـب وهو حاصل قسـمة مجموع هذه الرتب على عددها أي:

$$5.5 = \frac{7+6+5+4}{4}$$

ويلي ذلك الطالبان رقـم (5) ، (8) ولما كان لكل منهما نفس التقدير لذلك يعطى لكل منهما متوسط الرتبتين أي :

$$8.5 = \frac{9+8}{2}$$

ويلي ذلك الطالب رقم 1 حيث يأخذ الرتـبة (10) ، وبإتباع نفس الطريقة عـند إعطاء رتب التقدير لمادة الاقتصاد يمكن أن نحسـب الفروقـات كمـا موضح في الجدول (4-17) ثم نكمل حل المسألة بالطريقة المعتادة فنجد أن :

جـــدول (4 -17)

حساب معامل الارتباط لسبيرمان في حالة الرتب التكرارية

رقم الطالب	تقدير الإحصاء	تقدير الاقتصاد	رتب تقدير الإحصاء	رتب تقدير الاقتصاد	الفروقات F	مربع الفروقات F^2
1	ضعيف جداً	مقبول	10	7	3	9
2	مقبول	جيد	5.5	4.5	1	1
3	ممتاز	جيد جداً	1	2.5	-1.5	2.25
4	مقبول	مقبول	5.5	7	-1.5	2.25
5	ضعيف	جيد	8.5	4.5	4	16
6	جيد جداً	مقبول	2	7	-5	25
7	جيد	ممتاز	3	1	2	4
8	ضعيف	ضعيف جداً	8.5	10	-1.5	2.25
9	مقبول	ضعيف	5.5	9	-3.5	12.25
10	مقبول	جيد جداً	5.5	2.5	3	9
المجموع						83.00

$$\therefore R = 1 - \frac{6\sum F^2}{n(n^2-1)} = 1 - \frac{(6)(83)}{(10)(100-1)} = 1 - \frac{(6)(83)}{(10)(99)}$$

$$= 1 - \frac{498}{990} = 1 - 0.503 = 0.497$$

وهذه القيمة لمعامل الارتباط تبين أن هناك ارتباطاً طردياً ليس بالقوي ولـــيس بالضعيف .

ويجب ملاحظة أنه لا يقتصر استخدام معامل سبيرمان للارتباط على المتغيرات غير القابلة للقياس الكمي(كما تم توضيحه من خلال حل المثالين السابقين ولكن قد يستخدم أيضاً لحساب الارتباط بين المتغيرات القابلة للقياس الكمي وذلك رغبةً في تقليل واختصار العمليات الحسابية كما يتضح من حل المثال (4-9) .

مثال (4 -9)

أحسب معامل الارتباط لسبيرمان بين قيم x , y من البيانات المبينة في الجدول (4-18) :

جدول (4-18)

x	11	14	12	14	15
y	12	13	14	13	18

الحل :

من الجدول (4-18) نعطي المتغيرين x ، y رتباً ثم نحسب الفروقات بين الرتب المتقابلة ونوجد مربعاتها كما هو مبين في الجدول (4-19) .

جدول (4-19)

x	y	رتب x	رتب y	F	F^2
11	12	5	5	0	0
14	13	2.5	3.5	1 -	1
13	14	4	2	2	4
14	13	2.5	3.5	1 -	1
15	18	1	1	0	0
المجموع					6

$$\therefore R = 1 - \frac{6\sum F^2}{n(n^2-1)} = 1 - \frac{(6)(6)}{(5)(25-1)} = 1 - \frac{(6)(6)}{(5)(24)}$$

$$= 1 - \frac{3}{10} = 1 - 0.3 = 0.7$$

إذن يوجد ارتباط طردي قوي بين x ، y

مثـــال (4 -10)

إذا كانت تقديرات 5 طلبة في مادتين دراسيتين مبينة فـــي الجـــدول (4-20) المطلوب حساب معامل الارتباط بين هاتين المادتين .

جدول (4-20)

الطالب	1	2	3	4	5
تقدير المادة الأولى	جيد	ممتاز	جيد جداً	مقبول	ضعيف
تقدير المادة الثانية	مقبول	جيد جداً	جيد	ممتاز	ضعيف

الحـــل:

لاحظ أن البيانات وصفية خاضعة للترتيب وبالتالي فإن معامل الارتبـــاط المناسب لها هو معامل ارتباط الرتب لسبيرمان ويحسب كما هو مبـــين فـــي الجدول (4-21) :

جـــدول (4-21)

الطالب	تقدير المادة الأولى	تقدير المادة الثانية	رتب تقدير المادة الأولى	رتب تقدير المادة الثانية	الفروق F	مربعات الفروق F^2
1	جيد	مقبول	3	4	-1	1
2	ممتاز	جيد جداً	1	2	-1	1
3	جيد جداً	جيد	2	3	-1	1
4	مقبول	ممتاز	4	1	3	9
5	ضعيف	ضعيف	5	5	0	0
المجموع			15	15	0	12

$$\therefore R = 1 - \frac{6\sum F^2}{n(n^2-1)} = 1 - \frac{(6)(12)}{(5)(25-1)} = 1 - \frac{(6)(12)}{(5)(24)}$$

$$= 1 - \frac{3}{5} = 1 - 0.6 = 0.4$$

إذن يوجد ارتباط طردي ولكنه ليس قوياً .

مثـــال (4-11)

ذكر أحد المختصين بأن الارتباط بين عدد الأطفال والمســـتوى التعليمـــي لرب الأســـرة ارتباط طردي قوي ، فهل تؤيد رأيه بناءً على البيانات التاليـــة والمبينة في الجدول (4-22).

جـــدول (4-22)

الطالب	تقدير المادة الأولى	تقدير المادة الثانية	رتب تقدير المادة الأولى	رتب تقدير المادة الثانية	الفروق F	مربعات الفروق F^2
1	جيد	مقبول	3	4	-1	1
2	ممتاز	جيد جداً	1	2	-1	1
3	جيد جداً	جيد	2	3	-1	1
4	مقبول	ممتاز	4	1	3	9
5	ضعيف	ضعيف	5	5	0	0
المجموع			15	15	0	12

الأسرة	1	2	3	4	5	6	7	8	9
عدد الأطفال	4	3	2	7	3	4	7	3	5
المستوى التعليمي	يقرأ ويكتب	شهادة عليا	شهادة متوسطة	يقرأ ويكتب	شهادة متوسطة	شهادة متوسطة	أمي	شهادة متوسطة	أمي

الحـــل:

نلاحظ أنه لا يمكن حساب معامل ارتباط الرتب لسبيرمان لوجود بيانـــات إلاّ بعد تعديل الرتب للصفات المكررة أي الحصول على الرتب المعدّلة وذلك بإتباع الآتي كما هو مبين في الجدول (4-23) .

جـــدول (4-23)

عدد الأطفال	المستوى التعليمي	رتب عدد الأطفال	رتب المستوى التعليمي	الرتب المعدلة لعدد الأطفال	الرتب المعدلة للمستوى التعليمي	الفروق F	مربعات الفروق
4	يقرأ ويكتب	5	3	5.5	3.5	2	4
3	شهادة عليا	2	9	3	9	6 -	36
2	شهادة متوسطة	1	5	1	6.5	5.5 -	30.25
7	يقرأ ويكتب	8	4	8.5	3.5	5	25
3	شهادة متوسطة	3	6	3	6.5	3.5 -	12.25
4	شهادة متوسطة	6	7	5.5	6.5	1 -	1
7	أمي	9	1	8.5	1.5	7	49
3	شهادة متوسطة	4	8	3	6.5	3.5 -	12.25
5	أمي	7	2	7	1.5	5.5	30.25
المجموع		45	45	45	45	0	200

$$\therefore R = 1 - \frac{6\sum F^2}{n(n^2-1)} = 1 - \frac{(6)(200)}{(9)(81-1)} = 1 - \frac{(6)(200)}{(9)(80)}$$

$$= 1 - \frac{5}{3} = 1 - 1.67 = 0.67 -$$

وسنلاحظ أن الارتباط قوي إلى حدٍ ما وعكســـي وهذا لا يؤيد رأي الإحصائي الاجتماعي .

8.4 معامــل الاقتــران (Association Coefficient)

يســتخدم معامل الاقتران لدراســة قــوة واتجاه العلاقة بــين ظــاهرتين (لا يمكن التعبير عنهما رقمياً وغير خاضعتين لترتيب) ولكل منهما صــفتان متعاكستان مثل موجود وغير موجود ، وعادةً ما تظهر البيانات في هذه الحالة كما هو مبين في الجدول (4-24) .

جــدول (4-24)

الظاهرة الثانية / الظاهرة الأولى	الصفة الأولى	الصفة الثانية	المجموع
الصفة الأولى	a_{11}	a_{12}	$k_1 = a_{11} + a_{12}$
الصفة الثانية	a_{21}	a_{22}	$k_2 = a_{21} + a_{22}$
المجموع	$L_1 = a_{11} + a_{21}$	$L_2 = a_{12} + a_{22}$	$k_1 + k_2$

يحسب معامل الاقتران (A) من القانون التالي:

$$A = \frac{a_{11}\,a_{22} - a_{12}\,a_{21}}{a_{11}\,a_{22} + a_{12}\,a_{21}} \quad \text{..............................} (15-4)$$

مثـــال (4-12)

لتحديد قوة واتجاه العلاقـــة بـــين حالـــة التـــدخين والمســـتوى التعليمـــي للأشخاص ، جمعت بيانات من 100 شخص والمبينة في الجدول (4-25) على النحو التالي :

جــدول (4-25)

المستوى التعليمي ← / حالة التدخين ↓	متعلم	غير متعلم	المجموع
مدخن	30	20	50
غير مدخن	40	10	50
المجموع	70	30	100

أحسب معامل الاقتران بين المستوى التعليمي وظاهرة التدخين .

الحـــل:

حيث أن لدينا ظاهرتين فـــقط هما حالتي التدخين والمستوى التعليمي ولكل منهما صفتان متعاكستين لذا يستخدم معامل الاقتران على النحو التالي :

$$A=\frac{a_{11}a_{22}-a_{12}a_{21}}{a_{11}a_{22}+a_{12}a_{21}}=\frac{(30)(10)-(40)(20)}{(30)(10)+(40)(20)}=\frac{300-800}{300+800}$$

$$=\frac{-500}{1100}=-0.46$$

وبالتالي فإن معامل الاقتران بين المســـتوى التعليمي وظاهـــرة التـــدخين اقتران عكسي ليس بالقوي . أي الاقتران بين المتعلم وغير المدخن أو الاقتران بين غير المتعلم والمدخن .

مثـــال (4-13)

سئل 60 رجلاً عن رأيهم في حق المرأة في العمل فأجاب 47 رجلاً منهم بالرفض . وسئلت 40 أمرأة فأجابت 5 منهن بالرفض . أدرس قـــوة واتجـــاه العلاقة بين نوع الجنس وحق المرأة في العمل .

الحـــل:

حيث إنه لدينا ظاهرتين وصفيتين لكل منهما صفتان متعاكســـتان وبالتـــالي نستخدم معامل الاقتران وكما هو مبين في الجدول (4-26).

جـــدول (4-26)

الجنس ← / حق المرأة في العمل ↓	ذكر	أنثى	المجموع
مؤيد	13	35	48
رافض	47	5	52
المجموع	60	40	100

يحسب معامل الاقتران من المعادلة (4 -27) حيث :

$$A = \frac{a_{11}\, a_{22} - a_{12}\, a_{21}}{a_{11}\, a_{22} + a_{12}\, a_{21}}$$

$$= \frac{(13)(5) - (35)(47)}{(13)(5) + (35)(47)} = \frac{65 - 1645}{65 + 1645}$$

$$= -\frac{1580}{1710} = -0.92$$

وبالتالي يتضح أن الاقتران بين نوع الجنس وحق المرأة في العمل اقتران عكسي قوي . أي بين الذكر والرافض أو بين الأنثى والمؤيد .

ويجب ملاحظة أن معامل الاقتران يحسب في اقتران الصفة الأولى للظاهرة الأولى مع الصفة الأولى للظاهرة الثانية بالجدول ، وعليه إذا تغير وضع الجدول فإنه يغير إشارة معامل الاقتران .

9.4 معامـــل التوافـــق $\overline{H}$ (Harmonic Coefficient)

يســـتخدم معامل التوافق لدراســـة قوة العلاقة بين ظاهرتين وصفيتين غير خاضعتين للترتيب ولكل منهما أكـــثر من صفتين . وعادة ما تكون البيانـــات في هذه الحالة وكما هو مبين في الجدول (4-27) .

جـــدول (4-27)

الظاهرة الأولى / الظاهرة الثانية	الصفة الأولى	الصفة الثانية		الصفة الثالثة	المجموع
الصفة الأولى	a_{11}	a_{12}	....	a_{1n}	$\sum a_{1j}$
الصفة الثانية	a_{21}	a_{22}	....	a_{2n}	$\sum a_{2j}$
⋮	⋮	⋮		⋮	⋮
الصفة الأخيرة	a_{n1}	a_{n2}	......	a_{nn}	$\sum a_{nj}$
المجموع	$\sum a_{i1}$	$\sum a_{i2}$		$\sum a_{ij}$	$\sum a_{ij}$

ويحسب معامل التوافق من المعادلة التالية :

$$\overline{H} = \sqrt{\frac{K'-1}{K'}} \qquad \text{.......................................}(16-4)$$

حيث أن K' يحسب من المعادلة التالية :

$$K' = \frac{a_{11}^2}{\left(\sum_{i=1}^{n} a_{i1}\right)\left(\sum_{j=1}^{n} a_{1j}\right)} + \frac{a_{12}^2}{\left(\sum_{i=2}^{n} a_{i2}\right)\left(\sum_{j=2}^{n} a_{2j}\right)} + \cdots + \frac{a_{ij}^2}{\left(\sum_{i=1}^{n} a_{ij}\right)\left(\sum_{j=1}^{n} a_{ij}\right)} \qquad (17\text{-}4)$$

مثــال (4-14)

البيانات التالية تمثــل حالــة التــدخين والمســتوى التعليمــي لثلاثمائــة شخص . هل يوجد توافق بين التدخين والمستوى التعليمــي لهــذه البيانــات والمبينة في الجدول (4-28) .

جــدول (4-28)

حالة التدخين ← / المستوى التعليمي ↓	مدخن	غير مدخن	المجموع
متعلم	75	15	90
يقرأ ويكتب	90	60	150
أمي	15	45	60
المجموع	180	120	300

الحــل:

نحسب معامل التوافق بالشكل التالي :

$$K' = \frac{(75)^2}{(180)(90)} + \frac{(15)^2}{(120)(90)} + \frac{(90)^2}{(180)(150)} + \frac{(60)^2}{(120)(150)}$$

$$+ \frac{(15)^2}{(180)(60)} + \frac{(45)^2}{(120)(60)} = 1.17$$

$$\therefore \overline{H} = \sqrt{\frac{K'-1}{K'}} = \sqrt{\frac{1.17-1}{1.17}} = \sqrt{\frac{0.17}{1.17}} = 0.38$$

وهذا يدل على أن هناك توافقاً بين المستوى التعليمي والتدخين ولكنه ضعيف .

مثـــال (4-15)

الجدول (4-29) يبين لون الزهور ورائحتها فهل تعتقد أن هناك توافقاً بين لون الزهور ورائحتها ؟

جــدول (4-29)

اللون ← / الرائحة ↓	أحمر	أبيض	أخضر	المجموع
قوية	30	40	10	80
متوسطة	60	20	15	95
ضعيفة	10	10	5	25
المجموع	100	70	30	200

الحــل :

نحسب معامل التوافق لهذه البيانات حيث نجد أن :

$$K' = \frac{(30)^2}{(100)(80)} + \frac{(40)^2}{(70)(80)} + \frac{(10)^2}{(30)(80)} + \frac{(60)^2}{(100)(95)}$$
$$+ \frac{(20)^2}{(70)(95)} + \frac{(15)^2}{(30)(95)} + \frac{(10)^2}{(100)(25)} + \frac{(10)^2}{(70)(25)}$$
$$+ \frac{(5)^2}{(30)(25)} = 1.088$$

$$\therefore \overline{H} = \sqrt{\frac{K'-1}{K'}} = \sqrt{\frac{1.088-1}{1.088}} = \sqrt{\frac{0.088}{1.088}} = 0.3$$

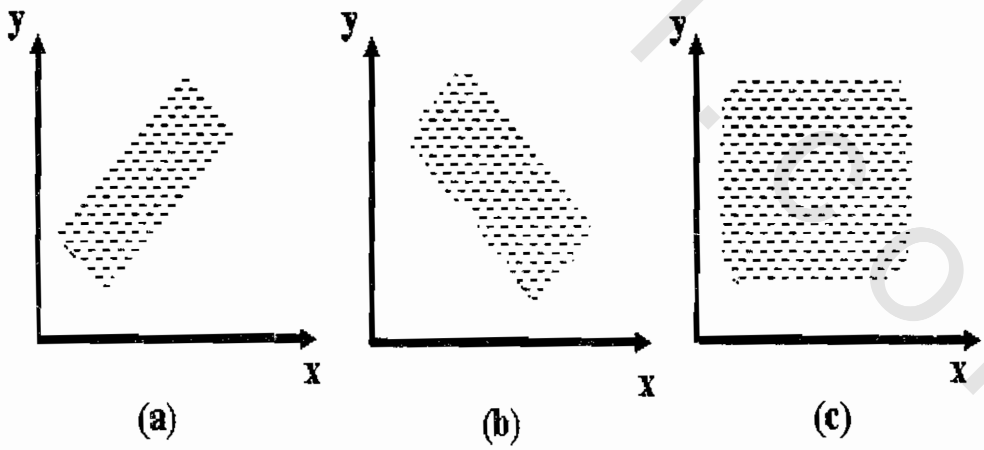
وهذا يدل على وجود توافق بين اللون والرائحة ولكنه ضعيف .

10.4 الانحـــدار (Regression)

لدراسة العلاقة بين ظاهرتين يمكن تكوين فكرة مبدئية عن نـــوع العلاقـــة وقوتها باســـتخدام ما يعرف بشـــكل الانتشـــار(Scatter Diagram) ، فـــإذا مثلنا أزواج المشـــاهدات الخاصة بالظاهرتين بيانياً نحصل على عدد من النقط في مستوى محورين كما هو مبين في الشـــكل (4-1) ، حيـــث يتضـــح مـــن الشـــكل (a) أنه توجد علاقة طردية بين المتغيرين ، بينما العلاقة في الشـــكل (b) علاقة عكسية ، ويظهر شكل الانتشار (c) أنـــه لا توجـــد علاقـــة بـــين المتغيرين حيث نجد أن النقط مبعثرة بطريقة غير منتظمة .

وواضح من الشـــكلين(a) و(b) أن النقط تـــقع على خط مســـتقيم بمعنى أنه توجد عـــلاقة خطية بين المتغيرين يمكن وضعها في شكل معادلـــة مـــن الدرجة الأولى على الصورة التالية :

$$y = m\,x + b \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (18-4)$$



الشكـــل (4 – 1)

أشكـــال الانتشـــار

حيث أن y هو المتغير التابع (Dependent Variable) ، والذي نريد تقديره و x هو المتغير المستقل (Independent Variable) ، أما m و b فهي مقادير ثابتة يمكن حسابها من واقع البيانات المشاهدة . وبمعرفة قيمة كل من m ، b يمكن استنتاج قيم y عندما تأخذ x قيماً معينة لذلك تعرف هذه المعادلة بمعادلة خط انحدار y على x حيث m تعطي ميل الخط و b هو الجزء المقطوع من المحور الرأسي .

ولتوصيل خط مستقيم يتوسط النقاط في شكل الانتشار خير توسط ليمثل العلاقة بين المتغيرين x ، y يمكننا أن نمهد هذا الخط باليد . ولكن هذا التمهيد سوف يكون تقريبياً ويختلف من شخص لآخر لذلك سنلجأ لاستخدام طريقة جبرية تعرف بطريقة المربعات الصغرى وهي طريقة دقيقة تمكننا من تحديد أفضل موضع لهذا الخط .

11.4 طريقة المربعات الصغرى (The Least Square Method)

من المعلوم أن الخط الذي نريد تمهيده سوف لا يمر بجميع النقاط في شكل الانتشار ، ولكن بعض هذه النقاط سيقع فوقه وبعضها سيقع تحته وبالتالي إذا اخترنا أي قيمة للمتغير x وقدرنا قيمة y المناظرة لها من واقع معادلة هذا الخط فإن قيمة y المقدّرة سوف تختلف عن قيمة y الفعلية أي المشاهدة في حالة عدم انطباق النقطة على الخط تماماً وهذا الاختلاف يعطي لنا انحراف النقطة أي البعد الرأسي لها عن خط الانحدار .

وتهدف طريقة المربعات الصغرى إلى إيجاد معادلة لهذا الخط بحيث يكون مجموع مربعات الانحرافات أي الأبعاد الرأسية للنقط عنه أصغر ما يمكن اي ذو نهاية صغرى .

ولإيجاد معادلة هذا الخط على صورة المعادلة(4-16) حيث إن b هو الجزء المقطوع من المحور الرأسي ، و m هو ميل خط الانحدار ويسمى أيضاً بمعامل انحدار y على x نجد أن قيم m ، b والتي تحقق هذا الشرط يمكن الحصول عليها من المعادلتين :

$$\sum y = m\sum x + N b \quad \dots\dots\dots\dots\dots\dots\dots\dots(19-4)$$

$$\sum x y = m\sum x^{2} + c\sum x \quad \dots\dots\dots\dots\dots\dots\dots(20-4)$$

وبقسمة المعادلة (4 – 19) على N عدد المفردات نجد أن :

$$\frac{\sum y}{N} = m\frac{\sum x}{N} + b$$

$$\therefore b = \frac{\sum y}{N} - m\frac{\sum x}{N} \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots(21-4)$$

أي أن:

$$b = \bar{y} - m\bar{x} \quad \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots(22-4)$$

حيث أن :

$\bar{y}$ - هي الوسط الحسابي لقيم y .

$\bar{x}$ - هي الوسط الحسابي لقيم x .

وبالتعويض عن قيمة b الموجودة في المعادلة (4-19) بما يساويها في المعادلة (4-20) ينتج أن :

$$\sum x\,y = m\sum x^2 + \sum x\left(\frac{\sum y}{N} - m\frac{\sum x}{N}\right)$$

$$\sum x\,y = m\sum x^2 + \frac{(\sum x)(\sum y)}{N} - m\frac{(\sum x)^2}{N}$$

$$\therefore \sum x\,y - \frac{(\sum x)(\sum y)}{N} = m\left[\sum x^2 - \frac{(\sum x)}{N}\right]$$

وبقسمة المعادلة على N ينتج أن :

$$\frac{\sum x\,y}{N} - \left(\frac{\sum x}{N}\right)\left(\frac{\sum y}{N}\right) = m\left[\frac{\sum x^2}{N} - \left(\frac{\sum x}{N}\right)^2\right]$$

$$\therefore m = \frac{\frac{\sum x\,y}{N} - \bar{x}\,\bar{y}}{S_x^2} \qquad \text{...............}(23-4)$$

أي أنه يمكن معرفة كل من m ، b من المعـــادلتين (4-22) و(4-23) لكـــي نحصل على معادلة انحدار y على x .

مثـــال (4-16)

أوجد معادلة خط انحدار y على x من البيانات المبينة في الجدول (4-31) :

جـــدول (4-30)

10	2	5	4	5	6	3	x
7	1	5	6	2	4	3	y

الحـــل :

لإيجاد هذه المعادلة يلزم معرفة مفردات المعادلة(4-23) وهي مبينة فـــي الجدول (4-31) :

جـــدول (4-31)

x y	x^2	y	x
9	9	3	3
24	36	4	6
4	25	2	5
36	16	6	4
25	25	5	5
2	4	1	2
70	100	7	10
$\sum xy=164$	$\sum x^2=215$	$\sum y=28$	$\sum x=35$

$$\bar{x} = \frac{35}{7} = 5 \qquad \bar{y} = \frac{28}{7} = 4$$

$$\therefore \delta_x^2 = \frac{215}{7} - \left(\frac{35}{7}\right)^2 = 30.7 - 25 = 5.7$$

$$\therefore m = \frac{\frac{164}{7} - (5)(4)}{5.7} = \frac{23.43 - 20}{5.7} = \frac{3.43}{5.7} = 0.6$$

$$\therefore b = \bar{y} - m\,\bar{x} = 4 - (0.6)(5) = 4 - 3 = 1$$

إذن معادلة انحدار y على x هي:

y = 0.6 x + 1

ويمكن كتابة المعادلة(4 – 22) الخاصة بطريقة المربعـــات الصــــغرى علـــى الصورة التالية :

$$m = \frac{N(\sum x_n y_n) - (\sum x_n)(\sum y_n)}{N(\sum x_n^2) - (\sum x_n)^2} \qquad ....(24-4)$$

$$b = \frac{(\sum x_n^2)(\sum y_n) - (\sum x_n y_n)(\sum x_n)}{N(\sum x_n^2) - (\sum x_n)^2} \qquad ....(25-4)$$

حيث أن :

n تمثل رتبة القيمة في الجدول .

N = عدد القيم .

وفي هذه الحالة تستخدم المعادلة (4 – 19) لإيجاد انحدار y على x .

مثـــال (4-17)

تسافر سيارة على طريق مســتقيم بســـرعة ثابتـــة $V = b_1 (m/\sec)$ ، ويعطى موقع تلك السيارة y_n عند أي زمن t_n وفقاً للمعادلة $y_n = b_0 + b_1 t_n$. أفرض أن هذه القياسات المبينة في الجدول (4-32) كانت كالتالي :

جــدول (4-32)

t_n , sec	0	3	5	8	10
y_n , m	200	230	240	270	290

استخدم طريقة المربعات الصغرى لإيجاد سرعة السيارة الثابتة(b_1) .

الحـــل :

باستخدام المعادلات(4 – 24) و(4 – 25) ننظم الجدول (4-32) على النحـــو المبين في الجدول (4-33) :

جـــدول (4-33)

t_n	y_n	$t_n y_n$	t_n^2
0	200	0	0
3	230	690	9
5	240	1200	25
8	270	2160	64
10	290	2900	100
$\sum t_n = 26$	$\sum y_n = 1230$	$\sum t_n y_n = 6950$	$\sum t_n^2 = 198$

بما أن عدد النقاط هو 5 وباستخدام المعادلات المذكورة أعـــلاه نحصـــل على :

$$m = \frac{N(\sum x_n y_n) - (\sum x_n)(\sum y_n)}{N(\sum x_n^2) - (\sum x_n)^2} = \frac{(5)(6950) - (26)(1230)}{(5)(198) - (26)^2}$$

$$\therefore \ m = \frac{34750 - 31980}{990 - 676} = \frac{2770}{314} = 8.821 \frac{m}{\text{sec}}$$

$$b = \frac{(\sum x_n^2)(\sum y_n) - (\sum x_n y_n)(\sum x_n)}{N(\sum x_n^2) - (\sum x_n)^2} = \frac{(198)(1230) - (6950)(26)}{(5)(198) - (26)^2}$$

$$\therefore \ b = \frac{243540 - 180700}{990 - 676} = \frac{62840}{314} = 200.13$$

وبذلك تكون معادلة الانحدار هي :

$$y_n = 200.13 + 8.821\ t_n$$

إذن سرعة السيارة الثابتة هي :

8.821 m / sec

مثـــال (4 -18)

أوجد معادلة انحدار y على x من البيانات المبينة في الجدول (4-34) :

جـــدول (4-34)

x	5	6	7	8	9	10	11	12	13	14	15
y	45	50	53	56	60	64	70	75	78	85	91

الحــل :

لاختصار العمليات الحسابية يمكن نقل نقطة الأصل وذلك بأخذ وســطين فرضيين لقيم x و y فإذا اختيرت القيم 10 , 60 لهذا الغرض وكما هو مبين في الجدول (4-35) ينتج أن :

$$\therefore\ m = \frac{\frac{493}{11} - (0)(\frac{67}{11})}{\frac{110}{11} - 0} = \frac{493}{110} = 4.482$$

$$\therefore\ b = \frac{67}{11} - (4.482)(0) = \frac{67}{11} = 6.091$$

$$\therefore Y = 4.482\ X + 6.091$$

جــدول (35-4)

x	y	X (x – 10)	Y (y – 60)	X^2	X Y
5	45	5 -	15 -	25	75
6	50	4 -	10 -	16	40
7	53	3 -	7 -	9	21
8	56	2 -	4 -	4	8
9	60	1 -	0	1	0
10	64	0	4	0	0
11	70	1	10	1	10
12	75	2	15	4	30
13	78	3	18	9	54
14	85	4	25	16	100
15	91	5	31	25	155
		$\sum X = 0$	$\sum Y = 67$	$\sum X^2 = 110$	$\sum XY = 493$

ولحساب معادلة الانحدار بدلالة القيم الأصلية نضع Y = y – 60 ونضع كذلك X = x –10 ونعوض في المعادلة الأخيرة أعلاه فنحصل على :

$$(y-60) = 4.482(x-10) + 6.091$$

$$\therefore y = 4.482x + 31.27$$

12.4 معادلـــة خـــط انحـــدار (x) علـــى (y)

(Regression Line Equation of x on y)

إذا استخدمنا y كمتغير مســـتقل و x كمتغير تابع فإنه يمكن إيجاد معادلـــة تمكننا من تقدير قيمة x عندما تكون قـــيمة y معلومة وتسمى بمعادلـــة خـــط انحدار x على y ويمكن كتابتها على الصورة التالية :

$$x = \overline{m}\, y + d \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(26-4)$$

حيث أن :

d- يمثل الجزء المقطوع من المحور الأفقي .

$\overline{m}$ - هي ميـــل خـــط الانحـــدار وتســـمى أيضـــاً بمعامـــل انحـــدار x على y .

ويمكن إيجاد هذه باســـتخدام طريـــقة المربعات الصغـــرى وذلك بجعـــل مجموع مربعات الأبعاد الأفقية للنقاط عن خط الانحدار أصغر ما يمكن وفـــي هذه الحالة يمكن حساب قيم $\overline{m}$ و d من المعادلتين التاليتين :

$$\sum x = \overline{m}\sum y + N d \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(27-4)$$

$$\sum y\,x = \overline{m}\sum y^2 + d\sum y \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(28-4)$$

ومن هاتين المعادلتين نجد أن :

$$\overline{m} = \frac{\dfrac{\sum y\,x}{N} - \overline{y}\,\overline{x}}{\delta_Y^2} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(29-4)$$

$$d = \overline{x} - \overline{m}\,\overline{y} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(30-4)$$

مثـــال (4-19)

أوجد معادلـــة خــط انحــدار x علـــى y مــن البيانـــات المبينـــة فـــي الجدول (4-36) .

جــدول (4-36)

x	y	y^2	x y
3	3	9	9
6	4	16	24
5	2	4	10
4	6	36	24
5	5	25	25
2	1	1	2
10	7	49	70
$\sum x = 35$	$\sum y = 28$	$\sum y^2 = 140$	$\sum xy = 164$

الحــل :

$$\bar{x} = \frac{35}{7} = 5 \qquad \bar{y} = \frac{28}{7} = 4$$

$$\therefore\ S_Y^2 = \frac{140}{7} - \left(\frac{28}{7}\right)^2 = 20 - (4)^2 = 20 - 16 = 4$$

$$\therefore\ \bar{m} = \frac{\frac{164}{7} - (5)(4)}{4} = \frac{23.43 - 20}{4} = \frac{3.43}{4} = 0.857$$

$$\therefore\ d = \bar{x} - \bar{m}\,\bar{y} = 5 - (0.857)(4) = 5 - 3.428 = 1.572$$

إذن معادلة خط انحدار x على y هي:

$$x = 0.857\,y + 1.572$$

13.4 العلاقـــة بين معامـــل الارتبـــاط ومعـــاملات الانحـــدار
(Relation between Correlation and Regression Coefficients)

1- إن حاصل ضرب معامل انحدار y على x (m) في معامل انحدار x على y ($\overline{m}$) يساوي مربع معامل الارتباط أي أن :

$$m\ x\ \overline{m} = \frac{\left(\frac{\sum x\,y}{N} - \bar{x}\,\bar{y}\right)}{S_X^2}\ x\ \frac{\left(\frac{\sum x\,y}{N} - \bar{x}\,\bar{y}\right)}{S_Y^2}$$

$$= \left(\frac{\frac{\sum x\,y}{N} - \bar{x}\,\bar{y}}{S_X^2\,S_Y^2}\right)^2 = R^2$$

$$\therefore R = \sqrt{m\ x\ \overline{m}} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \qquad (31-4)$$

مثـــال (4-20)

أحسب معامل الارتباط للبيانات المبينة في الجدول (4-36) للمثال السابق .

الحـــل :

$$R = \sqrt{(0.6)(0.857)} \approx 0.72$$

2- إن حاصل ضرب (m) معامل انحدار y على x في $\frac{S_X}{S_Y}$ يساوي معامل الارتباط .

$$m\,x\,\frac{S_X}{S_Y} = \frac{\left(\frac{\sum xy}{N} - \bar{x}\,\bar{y}\right)}{S_X^2}\,x\,\frac{S_X}{S_Y}$$

$$= \frac{\frac{\sum xy}{N} - \bar{x}\,\bar{y}}{S_X S_Y} = R$$

$$\therefore R = m\,x\,\frac{S_X}{SY} \qquad \text{.......................}(32-4)$$

أو:

$$m = R\,x\,\frac{S_X}{S_Y} \qquad \text{.......... ....}(33-4)$$

مثـــال (4-20)

إذا علمت أن الوســط الحسابي والانحراف المعياري للمتغيــر x همــا 2 و 0.89 على الترتيب وأن الوسط الحسابي والانحراف المعياري للمتغيــر y هما 8 و1.67 على الترتيب . أوجد معادلة خط انحــدارy علــى x علمــاً أن معامل الارتباط بين قيم x ، y يساوي 0.8 .

الحــل :

$$m = R\,x\,\frac{S_X}{S_Y} = \frac{(0.8)(1.67)}{(0.89)} = 1.5$$

$$\therefore b = 8 - (1.5)(2) = 8 - 3 = 5$$

$$\therefore y = 1.5\,x + 5$$

مثـــال (4-21)

إذا علمت أن معادلتي انحدار y على x وانحدار x على y هما :

$$y = 0.9\,x + 4.1$$
$$x = 2.1\,y - 1.3$$

ومن هاتين المعادلتين يتضح لنا أنه يوجد خطأ في أحداها .

الحــل :

نحسب معامل الارتباط فنجد أن :

$$R=\sqrt{m\,x\,\overline{m}}=\sqrt{(0.9)(2.1)}=\sqrt{1.89}=1.375$$

إذن يوجد خطــأ في إحدى المعادلتين لأن معامل الارتــباط لا يمكن أن يكون أكبر من الواحد الصحيح .

14.4 تماريــن

س1: باستخدام طريقة المربعات الصغرى أوجد كلاً من معادلتي خط انحدار y على x ، وخط انحدار x على y للبيانات المبينة في الجدول (4-37) .

جــدول (4-37)

x	1	3	4	6	8	9	11	14
y	1	3	4	4	5	7	8	9

س2: باســتخدام معادلتي الانحدار في السؤال السابق أحسب معامل الارتباط بين قيم x و y .

س3:أحسب معامل ارتباط الرتب " لسبيرمان " بين قيم x و y مــن البيانـــات المبينة في الجدول (4-38) .

جــدول (4-38)

x	13	14	15	11	16	12	13
y	13	16	15	12	14	15	17

س4: إذا علمت أن الانحراف المعياري لقيم x هو 0.61 والانحراف المعياري لقيم y هو 1.32 . أحســب معامل الارتباط بين قــيم x و y إذا عــــلمت أن معادلة انحدار y على x هي :

$$y = 0.84\,x + 21.38$$

س5: إذا علمت أن معادلتي انحدار y على x وانحدار x على y هما :

$$y = 0.72\,x + 3.12$$
$$x = 1.43 - 0.81y$$

بين أنه يوجد خطأ في إحدى هاتين المعادلتين .

س6: الجدول (39-6) يبين التقديرات التي حصل عليها ثمانية من الطلبة في كل من مادتي الرياضيات والإحصاء والمطلوب حساب معامل الارتباط بــين تقديرات هاتين المادتين .

جــدول (39-4)

تقدير الرياضيات	ضعيف	مقبول	ممتاز	جيد جداً	جيد	مقبول	مقبول	ضعيف جداً
تقدير الإحصاء	ضعيف	مقبول	جيد جداً	ممتاز	مقبول	جيد	ضعيف	ضعيف جداً

س7: أحسب معامل ارتباط بيرسون بين قيم x و y من البيانات المبينــة فــي الجدول (40-4) .

جــدول (4-40)

X	67	56	65	60	69	58
Y	177	171	170	169	174	171

أستخدم القيمتين 60 و 171 كوسطين فرضيين لقيم x و y على التوالي .

س8: أمكن التوصل إلى البيانات التالية عن المتغيرين x و y :

$$\sum x = 48 \quad , \quad \sum y = 76 \quad , \quad \sum x^2 = 536$$
$$\sum y^2 = 1108 \quad , \quad \sum xy = 720 \quad , \quad N = 6$$

والمطلوب :

a) إيجاد معادلة انحدار y على x .

b) إيجاد معامل الارتباط (بيرسون) بين قيم x و y .

س9: أحسب معامل الارتباط بين قيم x و y مــن البيانـــات الموضـــحة فـــي الجدول (4-41) :

جــدول (4-41)

المجموع	100 - 90	- 80	- 70	- 60	- 50	X / Y
2					2	- 50
20			4	11	5	- 60
35		6	25	4		- 70
30	10	14	6			- 80
13	8	5				100 - 90
100	18	25	35	15	7	المجموع

س10: أحسب معامل الارتباط العزمي لبيرسون بين قيم x , y من البيانـــات الموضحة في الجدول (4-42) .

جـــدول (4-42)

X	**170**	**172**	**169**	**165**	**168**	**167**	**171**	**163**	**169**	**166**
Y	**69**	**71**	**72**	**66**	**69**	**66**	**69**	**67**	**70**	**69**

أستخدم القيمتين 169 و69 كوسطين فرضيين لقيم x وy على الترتيب .

س11: البيانات المبينة في الجدول (4-43) تمثل الدخل والإنفـــاق الشـــهري بالدينار لعشر أسر والمطلوب :

1- تحديد العلاقة المناسبة من خلال شكل الانتشار .

2- تقدير تلك العلاقة مع رسمها على الشكل ألانتشاري .

3- تحديد (تقدير) الإنفاق الشهري عندما يكون الدخل 110 دينار .

4- تحديد معامل الارتباط بين الدخل والإنفاق .

جـــدول (4-43)

الدخل (x)	73	50	128	170	87	108	135	69	148	132
الإنفاق (y)	30	20	60	80	40	50	60	30	70	60

س12: البيانات المبينة في الجدول (4-44) تمثل قيماً للمتغيرين x ، y .

جـدول (4-44)

X	2	4	5	6	8	11
Y	18	12	10	8	7	5

a) أوجد معامل ارتباط بيرسون للبيانات .

b) أوجد معامل ارتباط الرتب لسبيرمان . أيهما أفضل مع التعليق على النتائج .

س13: من خلال البيانات أدرس العلاقة بين الإصابة بمرض الكوليرا والتطعيم بمصل مضاد من بين 200 شخص تم تطعيمهم أصيب 50 شخصاً ، ومن بين 200 شخص لم يطعّموا أصيب 150 شخصاً .

س14: أحسب معامل الارتباط بين قيم x و y من البيانات المبينة (4-45) .

جـدول (4-45)

Y \ X	140 -	150 -	160 -	170 -	180 -	190- 200	المجموع
40 -	3	5	4				12
50 -	3	6	6	2			17
60 -	1	4	9	5	2		21
70 -			5	10	8	1	24
80 -			1	4	6	5	16
90 - 100				2	4	4	10
المجموع	7	15	25	23	20	10	100

س15: أدرس العلاقة بين نوع النبات ودرجة الإصابة بمرض ما من خلال البيانات المبينة في الجدول (4-46) والتي تمثل عدد النباتات المصابة من كل نوع .

جدول (4-46)

نوع النبات / درجة الإصابة	A	B	C
بسيطة	300	30	20
متوسطة	50	250	50
عالية	30	100	300

س16: إذا علمت أن معامل الارتباط بين درجات الكيمياء ودرجات الفيزياء مصححة من 50 هو 0.75.فإذا تم تحويل درجات المادتين من 100 فما هو معامل الارتباط الجديد .