الارتباط

مفهوم الارتباط :

هو تعيين طبيعة وقوة العلاقة بين متغيرين او عدمها ، والارتباط يتعامل مع الوضع الذي يكون فية متغيرين علي علاقة ببعضهما ومعامل الارتباط هو مؤشر هذه العلاقة

ويوجد نوعان من **الارتباط**

الارتباط

ارتباط موجب

الارتباط الموجب (الطردي):

هو علاقة بين متغيرين (x,y) بحيث اذا تغير احد المتغيرين فان الاخر يتبعه في نفس الاتجاه

ارتباط سالب

الارتباط السالب (العكسي) :

هو علاقة بين متغيرين (x,y) بحيث اذا تغير احد المتغيرين فان الاخر يتبعه في الاتجاه المضاد

قياس الارتباط :

تستخدم معاملات خاصة تسمي معاملات الارتباط لقياس درجة الارتباط بين متغيرين (ظاهرتين)

معامل الارتباط :

يرمز له بالرمز (r) ويعرف بانه عبارة عن مقياس رقمي يقيس قوة ونوع الارتباط بين متغيرين حيث تتراوح قيمته بين (1+_1-)

اي ان ($-1 \leq r \leq +1$) ، وتدل اشارة المعامل الموجبة علي العلاقة الطردية ، بنما تدل اشارة المعامل السالبة علي العلاقة العكسية

ملاحظات هامة :

_ يقال ان معامل الارتباط **طردي** تام اذا كان معامل الارتباط **r=+1**

ويقال ان الارتباط **عكسي** تام اذا كان **r=-1**

ويقال انهما **غير مرتبطين** اذا كانت **r=0**

الارتباط الطردي : (قيمة معامل الارتباط ومعناه)

قيمة معامل الارتباط	المعني
+1	ارتباط طردي تام
من (0.70 الي 0.99	ارتباط طردي قوي
من 0.50 الي 0.69	ارتباط طردي متوسط
من 0.01 الي 0.49	ارتباط طردي ضعيف
0	لا يوجد ارتباط

معامل بيرسون للارتباط الخطي

معامل بيرسون للارتباط الخطي :

معامل بيرسون من اكثر معا ملات الارتباط استخداما وخاصة في العلوم الانسانية والاجتماعية

ومستوي القياس المطلوب عند تطبيقه ان يكون كلا المتغيرين مقياس فترة او نسبة او البيانات تكون كمية

حساب معامل بيرسون

يحسب بدلالة المتغيرات لبيانات المتغيرين (x,y) باستخدام الصيغة التالية :

$$r_p = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n(\sum x^2) - (\sum x)^2]\ [n(\sum y^2) - (\sum y)^2]}}$$

حيث :

$\sum xy$: مجموع حاصل ضرب X في Y

$\sum x$: مجموع قيم المتغير (او الظاهرة) x

$\sum y$: مجموع قيم المتغير (او الظاهرة) x

$\sum x^2$: مجموع مربعات قيم المتغير (او الظاهرة) x

$\sum y^2$: مجموع مربعات المتغير (او الظاهرة) y

n : عدد المفردات

مثال 1

سجلت ست قراءات تقريبية لحجم الانتاج وحجم صادرات النفط الخام بالمملكة العربية السعودية (بالمليار برميل) خلال عدة سنوات كما يلي :

حجم الانتاج(x)	3	4	2	2	2	2
حجم الصادرات(y)	2	2	2	1	1	1

ادرس وجود علاقة ارتباط بين حجم الانتاج وحجم الصادرات للنفط الخام

الحل

X	Y	xy	X^2	Y^2
3	2	6	9	4
4	2	8	16	4
2	2	4	4	4
2	1	2	4	1
2	1	2	4	1
2	1	2	4	1
$\sum 15$	$\sum y = 9$	$\sum xy = 24$	$\sum x^2 = 41$	$\sum y^2 = 15$

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

$$= \frac{6(24)-(15)(9)}{\sqrt{6(41)-15^2][6\times15)-9^2]}} = \frac{9}{\sqrt{189}} = .65$$

من الملاحظ ان علاقة الارتباط الخطي بين حجم الانتاج وحجم الصادرات النفط الخام علاقة طردية متوسطة

مثال 2

لدراسة علاقة الصادرات بالميزان التجاري خلال عدة سنوات ، اخذ نا عشر قراءات تقريبية لقيمة صادرات المملكة العربية السعودية (x) وقيمة الميزان التجاري (y) بعشرات المليارات ريال كما يلي :

X	9	11	17	18	16	16	19	23	83
y	1	3	8	7	5	7	8	12	12

احسب معامل الارتباط الخطي ، ما مدي قوة العلاقة الخطية ؟

X	Y	xy	X^2	Y^2
9	1	9	81	1
11	3	33	121	9
17	8	136	289	64
8	7	126	324	49
19	6	114	361	36
16	5	80	256	25
16	7	112	256	49
19	8	152	361	64
23	12	276	529	144
23	12	276	529	144
$\sum x=171$	$\sum y=69$	$\sum xy=1314$	$\sum x^2=3107$	$\sum y^2=585$

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

$$= \frac{10(1314) - (171)(69)]}{\sqrt{[(10\times 3107) - 171^2][(10\times 585 - 69^2]}} = \frac{1341}{1411.30} = .95$$

نلاحظ ان علاقة الارتباط الخطي بين قيمة صادرات المملكة وقيمة الميزان التجاري موجودة و هي علاقة طردية قوية .

==

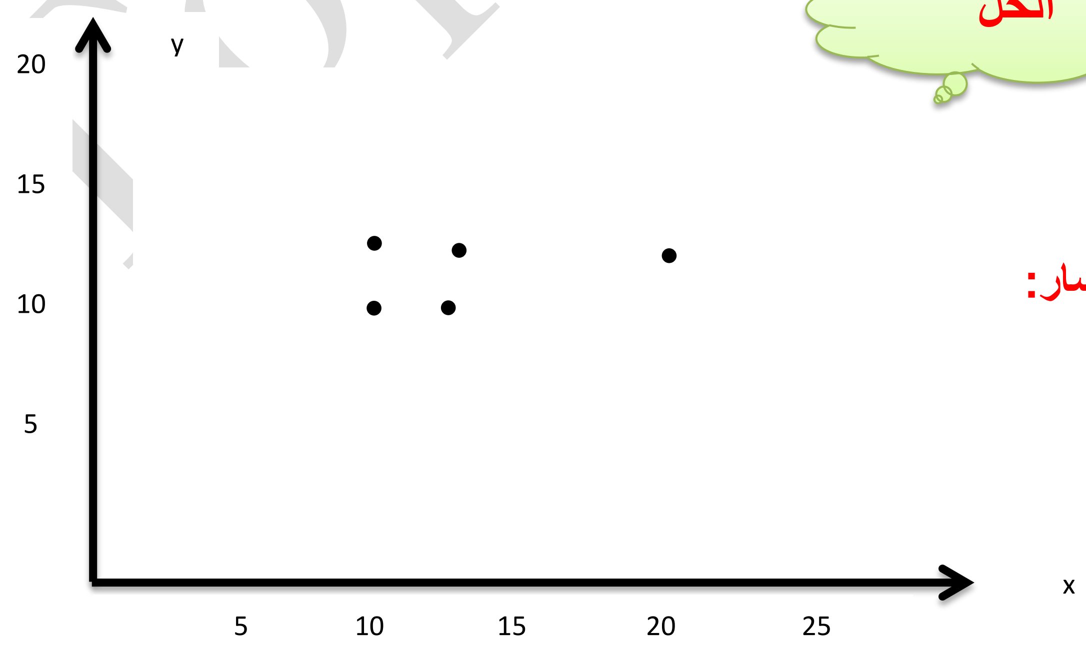
مثال 3

طلب مدرس الدراسات الاجتماعية 20 سؤالا عن الاحداث العالمية والمحلية اذا كان الجدول التالي يوضح درجات 10 من التلاميذ فارسم شكل الانتشار ثم احسب معامل بيرسون للارتباط r

الطفل	A	B	C	D	E	F	G	H	I	J
محلي(x)	20	15	12	18	10	13	12	10	18	15
عالمي (y)	15	12	10	14	10	8	6	15	16	13

الحل

شكل الانتشار:



X	Y	xy	X^2	Y^2
20	15	300	400	225
15	12	180	225	144
12	10	120	144	100
18	14	252	324	196
10	10	100	100	100
13	8	104	169	64
12	6	72	144	36
10	15	150	100	825
18	15	288	321	256
15	13	195	225	169
$\sum x = 143$	$\sum y = 119$	$\sum xy = 1761$	$\sum x^2 = 2155$	$\sum y^2 = 1515$

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

$$= \frac{(10)(1761) - (143)(119)}{\sqrt{[(10)(2155) - 143^2][(10)(1515) - 119^2]}}$$

$$= \frac{593}{\sqrt{1101 \times 989}} = .56$$

مثال 4

اختر الاجابة المناسبة للفقرات الاتية

1. العلاقة بين متغيرين (x,y) بحيث اذا تغير احد المتغيرين فان الاخر يتبعه في نفس الاتجاه هو علاقة ....

a) (ارتباط سالب) ارتباط طردي
b) (ارتباط سالب) ارتباط عكسي
c) (ارتباط موجب) ارتباط طردي
d) (ارتباط موجب) ارتباط عكسي

2. تتراوح قيمة معامل الارتباط (r) بين ....

a) $-1 \leq r \leq 1$

b) $-1 \leq r \leq 2$

c) $-2 \leq r \leq 1$

d) $-2 \leq r \leq 2$

معامل بوينت بايسيريال للارتباط :

يستخدم هذا المعامل لقياس علاقة الارتباط بين متغير كمي (x) ومتغير اسمي (y)
كالاجابة(نعم _ لا) او الجنس (ذكر _ انثي)

حساب هذا المعامل

يحسب من العلاقة :

$$r_{pd}=\frac{\overline{x_1}-\overline{x_2}}{s_x}\sqrt{\frac{n_1 n_2}{n(n-1)}}$$

حيث :

$\overline{x_1}$: يمثل المتوسط للمتغير الكمي تحت المستوي الاول

$\overline{x_2}$: المتوسط للمتغير الكمي تحت المستوي الثاني

S_x: يمثل الانحراف المعياري للمتغير الكمي

n : عدد المشاهدات الكلي

n_1 : عدد المشاهدات تحت المستوي الاول

n_2 : عدد المشاهدات تحت المستوي الثاني .

مثال 5

اوجد قيمة معامل الارتباط بين الاجابة علي السؤال الاجباري (x) حيث (1) تعني الاجابة علي السؤال الاجباري و(0) تعني عدم الاجابة علي السؤال الاجباري والدرجة الاجمالية لعدد خمسة من الطلبة حسب البيانات التالية .

زوج القراءات	1	2	3	4	5
Y	1	1	1	0	0
X	18	14	19	11	7
	X_1			X_2	

الحل

n=5 , n_1=3 , n_2=2

$$\overline{x_1}=\frac{18+14+19}{3}=17 \qquad \overline{x_2}=\frac{11+7}{2}=9$$

$$s_x=\sqrt{\frac{\sum x^2-[\frac{(\sum x)^2}{n}]}{n-1}}=\sqrt{\frac{[18^2+14^2+19^2+11^2+7^2]-\frac{[18+14+19+11+7]^2}{5}}{4}}=$$

$$\sqrt{\frac{98.8}{4}}=\sqrt{24.7}=4.97$$

$$r_{pd}=\frac{\overline{x_1}-\overline{x_2}}{s_x}\sqrt{\frac{n_1 n_2}{n(n-1)}} \qquad =\frac{17-9}{4.97}\sqrt{\frac{3\times2}{5\times4}}=.88$$

اي ان قيمة معامل الارتباط تدل علي ان هناك ارتباط قوي بين الاجابة علي السؤال الاجباري والدرجة الاجمالية

معامل سبيرمان لارتباط الرتب

يستخدم معامل سيبر مان اذا كان قياس المتغيرين (كليهما) مقياس ترتيبي او اسمي

حساب معامل سيبرمان لارتباط الرتب

نفرض ان المتغير x له الرتب R_x وان المتغير Y له الرتب R_y ونفرض ان d ترمز لفرق الرتبتين ، بمعني $d=R_x-R_y$

فان معامل سيبرمان يعطي بالصيغة :

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

حيث (n) هي عدد الازواج المرتبة

مثال 6

لدراسة علاقة ارتباط تقديرات الطلاب في مادة الاحصاء وتقديراتهم في مادة الرياضيات ، اخترنا خمس طلاب وكانت تقديراتهم كما يلي :

تقديرات الاحصاء(x)	F	A	C	D	B
تقديرات(y) الرياضيات	D	C	B	F	A

هل توجد علاقة ارتباط ؟ ما نوعها ومدي قوتها ؟

X	Y	رتب x	رتبY	d	D^2
F	D	1	2	-1	1
A	C	5	3	2	4
C	B	3	4	-1	1
D	F	2	1	1	1
B	A	4	5	-1	1
			$\sum$	0	8
				$\sum d$	$\sum d^2$

$r_p = 1 - \frac{6\sum d^2}{n(n^2-1)}$

$= 1 - \frac{6\times 8}{5(25-1)} = 1 - 04 = .6$

نلاحظ وجود علاقة ارتباط طردية متوسطة بين تقديرات الطلاب في مادة الاحصاء وتقديراتهم في مادة الرياضيات

مثال 1

في دراسة لمعرفة العلاقة بين عدد الحقول المكتشفة وطول الانابيب (بالكيلو متر) الناقلة للنفط الخام بالمملكة العربية السعودية خلال عدة سنوات ، سجلت سبع قراءات علي النحو التالي :

عدد الحقول(x)	55	54	56	61	62	63	67
طول الانابيب(y)	21960	22027	23006	23008	23020	23125	23120

هل توجد علاقة ارتباط بين عدد الحقول وطول الانابيب

X	y	رتبX	رتب Y	d	d^2
55	21690	2	1	1	1
54	22027	1	2	-1	1
56	23006	3	3	0	0
61	23008	4	4	0	0
62	23020	5	5	0	0
63	23125	6	7	-1	1
67	23120	7	6	1	1
				$\sum d = 0$	$\sum d^2 = 4$

$$r_p=1-\frac{6\sum d^2}{n(n^2-1)} = 1-\frac{6\times4}{7\times(49-1)} =1-\frac{24}{336} = 1-0.07 = .93$$

← نلاحظ وجود علاقة ارتباط طردية قوية بين عدد الحقول وطول الانابيب

مثال 8

عند تقييم مجموعة من الناقدين الرياضين لعدد 10 من الاعبيين تبعا للتحمل التدريبي قبل المسابقة وترتيب هؤلاء اللا عبين بعد المسابقة كان الترتيب التالي

اللا عب	A	B	C	D	E	F	G	H	I	J
رتبة التحمل التدريبي	5	9	10	2	8	7	4	1	6	3
رتبة اللاعب النهائية	4	8	10	2	9	6	3	1	7	5

احسب معامل الارتباط لدراسة العلاقة بين الحمل التدريبي والترتيب النهائي

مثال 1

اللاعب	رتبة التحمل التدريبي	رتبة الترتيب	$D=r_x-r_y$	D^2
A	5	4	+1	1
B	9	8	+1	1
C	10	10	0	0
D	2	2	0	0
E	8	9	-1	1
F	7	6	1	1
G	4	3	1	1
H	1	1	0	0
I	6	7	-1	1
J	3	5	-2	4
				$\sum d^2 = 10$

$$r_s=1-\frac{6\sum d^2}{n(n-1)}=1-\frac{(6)(10)}{(10)(99)}=1-0.06=.94$$

←نلاحظ انه توجد علاقة ارتباط طردية قوية بين المتغيرين ،بمعني انه كلما زاد الحمل التدريبي كلما تم الحصول علي ترتيب متقدم

===

ملاحظات هامة :

_ معامل سيبرمان لارتباط الرتب يمكن حسابه سواء اكانت البيانات كمية او وصفة ترتيبة بينما معامل بيرسون للارتباط الخطي لايمكن حسابه الا علي المتغيرات ذات قياس الفترة او قياس النسبة

_ يتميز معامل سيبرمان لارتباط الرتب بسهولة حسابه حتي لو كانت البيانات غير مرتبة

_ يعاب عليه اهماله لفروق الاعداد عند حساب الرتب وبالتالي فهو اقل دقة

_ يصعب حسابه للمتغيرات الكمية اذا كانت كبيرة العدد ، ولذلك يفضل استخدامه لتحديد درجة ارتباط بيانات كمية عددها اقل من 30

مثال 9

لدراسة الارتباط بين درجات الطلاب في مادتي الاحصاء والرياضيات ، اخذت عينة من ست طلاب وكانت نتائجهم كالتالي :

درجات الاحصاء (x)	80	90	70	65	60	50
درجات الرياضيات (y)	95	70	85	65	60	45

اجب عن الفقرتين الاتيين :

1. معامل الارتباط الخطي (بيرسون) بين درجات الطلاب في مادة الاحصاء ودرجاتهم في الرياضيات يساوي ....

a. 0.69

b. 0.96

c. 0.75

d. 0.37

اذا علمت ان :

$\sum xy = 29925$ ، y^2=31000 ، $\sum x^2 = 29725$ ، $\sum y = 420$ ، $\sum x = 415$

$$r_p = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n(\sum x^2) - (\sum x)^2]\,[n(\sum y^2) - (\sum y)^2]}}$$

$$= \frac{6(29925) - (415)(420)}{\sqrt{[(6)(29925) - (415)^2][(6(131000) - (420)^2]}} = \frac{5250}{\sqrt{58800000}} = 0.685 \cong 0.69$$

2. معامل ارتباط الرتب (سيبرمان) بين درجات الطلاب في مادة الاحصاء ودرجاتهم يساوي ....

a. 0.73
b. 0.63
c. 0.93
d. 0.83

$$R_s = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{36}{210} = 1 - 0.17 = 0.83$$

==

احسب معامل بيرسون ومعامل سيبرمان للنتائج التالية :

X	18	13	9	17	12	10	16	11
Y	22	14	18	25	19	20	24	17

لايجاد معامل بيرسون :

X	Y	Xy	X^2	Y^2
18	22	396	324	484
13	14	182	169	196
9	18	162	81	324
17	25	425	289	625
12	19	228	144	361
10	20	200	100	400
16	24	384	256	576
11	17	187	121	289
$\sum x = 106$	$\sum y = 159$	$\sum xy == 2164$	$\sum x^2 = 1484$	$\sum y^2 = 3255$

$$r_p = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n(\sum x^2) - (\sum x)^2][n(\sum y^2) - (\sum y)^2]}}$$

$$= \frac{8(2164)-(106)(159)}{\sqrt{[8(1484)-(106)^2][8(3255)-(159)^2}} = \frac{458}{\sqrt{636\times759}} = \frac{458}{\sqrt{482724}} = 0.659$$

لايجاد معامل بيرسون :

X	Y	R_x رتب(x)	R_y (y) رتب	d	d^2
18	22	8	6	2	4
13	14	5	1	4	16
9	18	1	3	-2	4
17	25	7	8	-1	1
12	19	4	4	0	0
10	20	2	5	-3	9
16	24	6	7	-1	1
11	17	3	2	1	1
					$\sum d^2 = 36$

$$r_s = 1 - \frac{6\sum d^2}{n(n-1)} = 1 - \frac{6\times36}{8(64-1)} = 1 - \frac{216}{504} = 1 - \frac{3}{7} = 0.57$$

نلاحظ ان معامل سيبرمان اقل من معامل بيرسون

مثال 11

لدراسة العلاقة بين الدخل (x) والاستهلاك (y) بمئات الريالات في مدينة ما ، اخذت عينة من الاسر فاعطت النتائج الاتية :

X	5	4	5	6	9	10	9	12	11	9
y	5	4	5	5	8	6	8	11	10	8

1. احسب معامل ارتباط بيرسون بين الظاهرتين
2. احسب معامل ارتباط سيبرمان بين الظاهرتين

الحل

معامل بيرسون

X	Y	Xy	X^2	Y^2
5	5	25	25	25
4	4	16	16	16
5	5	25	25	25
6	5	30	36	25
9	8	72	81	64
10	6	60	100	36
9	8	72	81	64
12	11	132	144	121
11	10	110	121	100
9	8	72	81	64
$\sum x = 80$	$\sum y = 70$	$\sum xy = 614$	$\sum x^2 = 710$	$\sum y^2 = 540$

$$r_p=\frac{10(614)-(80)(70)}{\sqrt{[10(710)-(80)^2][10(540)-(70)^2}}$$

$$=\frac{540}{\sqrt{700\times500}}=\frac{540}{\sqrt{350000}}=0.91$$

معامل سيبرمان :

n =10

X	Y	$R_x(x)$ رتب	$R_y(y)$ رتب	d	d^2
5	5	2	2	0	0
4	4	1	1	0	0
5	5	3	3	0	0
6	5	4	4	0	0
9	8	5	6	-1	1
10	6	8	5	13	9
9	8	6	8	-2	4
12	11	10	10	0	0
11	10	9	9	0	0
9	8	7	7	0	0
				$\sum d=0$	$\sum d^2=14$

$$r_s=1-\frac{6\sum d^2}{n(n^2-1)}=1-\frac{6\times14}{10\times99}=0.91$$

← نلاحظ ان معامل سيبرمان (r_s)= معامل بيرسون (r_p)

معامل الاقتران (فاي)

يستخدم العلاقة بين متغيرين اسمين كل منهما ثنائي التقسيم كالنوع(ذكر / انثي) والاصابة بالمرض (مصاب/ غير مصاب) والتدخين (مدخن/ غير مدخن)

نفترض ان المتغرين معرضين علي صورة جدول ثنائي مزدوج كما يلي :

	X_1	X_2	المجموع
Y_1	a	b	a +b
Y_2	C	d	C+ d
المجموع	a+c	b+d	

فان معامل فاي للاقتران يعطي من العلاقة :

$$r_{\varnothing}=\frac{ad-bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}$$

مثال 12

اوجد قيمة معامل الاقتران بين النوع x (ذكر/ انثي) والاصابة بمرض الاكتئاب (y) (مصاب/ غير مصاب) للبيانات التالية .

مرض الاكتئاب / النوع	مصاب	غير مصاب
ذكر	12	8
انثي	4	6

الحل

مرض الاكتئاب / النوع	مصاب	غير مصاب	المجموع
ذكر	12	8	20
انثي	4	6	10
المجموع	16	14	30

a =12 , b= 8 , c=4 , d=6

$$r_{\phi}=\frac{ad-bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}=\frac{12\times6-8\times4}{\sqrt{20\times10\times16\times14}}=0.19$$

اي انه توجد علاقة ضعيفة بين النوع والاصابة بمرض الاكتئاب .

مثال 13

اوجد قيمة معامل الارتباط بين النوع x (ذكر /انثي) والتدخين y (مدخن /غير مدخن) للبيانات التالية موضحا اسم المعامل الذي اخترته .

النوع \ مرض الاكتئاب	مدخن	غير مدخن
ذكر	7	3
انثي	4	6

الحل

النوع \ الحالة	مدخن	غير مدخن	المجموع
ذكر	7	3	10
انثي	4	6	10
المجموع	11	9	20

a=7 , b=3 , c=4 , d=6

$$r_{\varnothing} = \frac{ad-bc}{\sqrt{(a+b)(c+d)(a+c)(b+d)}}$$

$$= \frac{7(6)-3(4)}{\sqrt{10\times10\times11\times9}} = \frac{30}{\sqrt{9900}} = 0.30$$

استخدمنا معامل الاقتران (فاي) لان العلاقة بين متغيرين اسميين