المحاضرة الثالثة في الاحصاء الاستدلالي: معامل الارتباط بيرسون

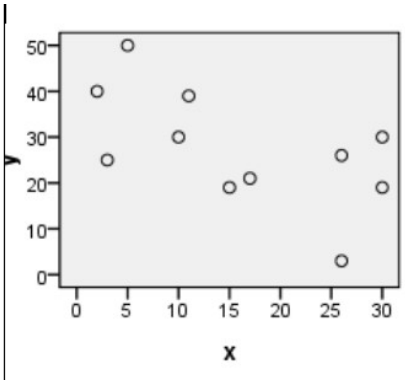
معامل الارتباط (Correlation Coefficient): يقيس معامل الارتباط اتجاه وحجم العلاقة بين المتغيرات، يأخذ قيمة بين (+1 و –1) وتشير الإشارة الى اتجاه العلاقة (موجبة أي العلاقة طردية وسالبة تعني علاقة عكسية)، والقيمة المطلقة للمعامل تشير الى حجم العلاقة. اما الصفر فيظهر عدم وجود علاقة او ارتباط بين المتغيرات المدروسة. ويتم تحديد حجم العلاقة بين المتغيرين من خلال قيمة معامل الارتباط المحسوبة باستخدام قانون الارتباط ثم مقارنتها بالقيمة الجدولية المحددة بمستوى الدلالة ودرجة الحرية (عدد العينة –2). فعندما تكون درجة الارتباط المحسوبة أكبر من القيمة الجدولية فهذا يعني وجود ارتباط قوي، وفي حالة العكس هذا يعني عدم وجود ارتباط معنوي بين المتغيرين.

أشكال الانتشار:

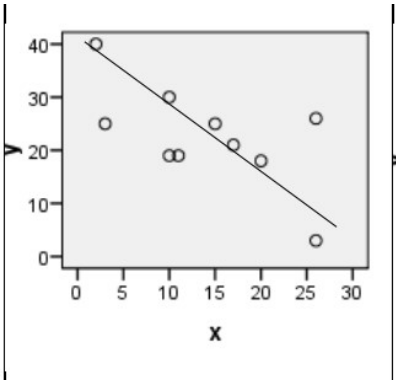
يمكن تمثيل العلاقة بين المتغيرين بيانيا وذلك بتعيين النقاط التي تمثل المتغيرين معا على الرسم البياني، وتعرف الطريقة التي تتوزع بها نقاط البيانات على الرسم البياني بأشكال لوحة الانتشار. ومن هذه الاشكال يمكن التعرف على نوع العلاقة بينهما كالتالي:

- اذا كانت النقاط موزعة على شكل خط مستقيم تكون العلاقة بين المتغيرين خطية؛
- اذا كان الخط المستقيم متزايدا تكون العلاقة طردية موجبة؛
- اذا كان الخط المستقيم متناقصا فالعلاقة عكسية سالبة؛
- كلما كانت النقاط قريبة من الخط يكون الارتباط أقوى.
- اذا كانت العلاقة موزعة على شكل منحنى فالعلاقة ليست خطية؛
- اذا كانت النقاط مبعثرة ولا تأخذ الشكلين الخطي والمنحنى تعني عدم وجود ارتباط (مبعثرة).

ويمكن توضيح بعض هذه الحالات في الاشكال التالية:

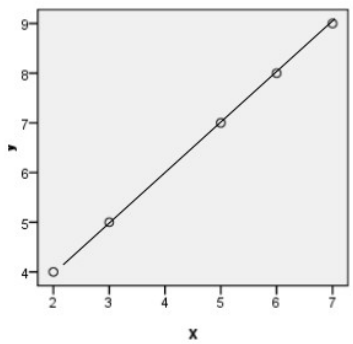


عدم وجود ارتباط



ارتباط خطي سالب

علاقة خطية عكسية



ارتباط خطي موجب

علاقة خطية طردية تامة

معامل الارتباط البسيط (بيرسون):

يستعمل معامل ارتباط بيرسون لقياس قوة واتجاه العلاقة الخطية بين متغيرين رقميين اثنين، يرمز لهما عادة ب (Xو Y)، تتراوح قيمة معامل بيرسون بين -1 و+1 . حيث يشير -1 الى علاقة سالبة وارتباط تام (أي عندما يزداد أحد المتغيرين ينقص الاخر)، ويشير 1+ الى علاقة موجبة وارتباط تام (يزداد أحد المتغيرين بزيادة المتغير الاخر) بينما يشير الصفر الى عدم وجود علاقة خطية بين المتغيرين.

ولحساب معامل ارتباط بيرسون نستخدم القانون التالي:

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

حيث: x هو المتغير الاول

Y هو المتغير الثاني

n حجم العينة

بعد حساب معامل الارتباط نقوم بتوضيح نوع العلاقة (طردية ام عكسية، او لا توجد علاقة) ثم نقوم بتحديد قوة العلاقة حيث:

- إذا كانت قيمة (r) أكبر من (0) وأقل من (0.3) فالارتباط ضعيف جدا
- إذا كانت قيمة (r) محصورة بين (0.3-0.5) فالارتباط ضعيف
- 0.5-0.7 فالارتباط متوسط
- 0.7-0.9 فالارتباط قوي
- 0.9-1 فالارتباط قوي جدا

مثال: تم اجراء اختبارين رياضيين مختلفين على 08 لاعبين وكانت النتائج مسجلة في الجدول التالي:

والمطلوب: تحديد نوع وقوة العلاقة بين الاختبارين وشكل الانتشار

X	Y
10	15
2	5
5	11
6	10
7	14
1	3

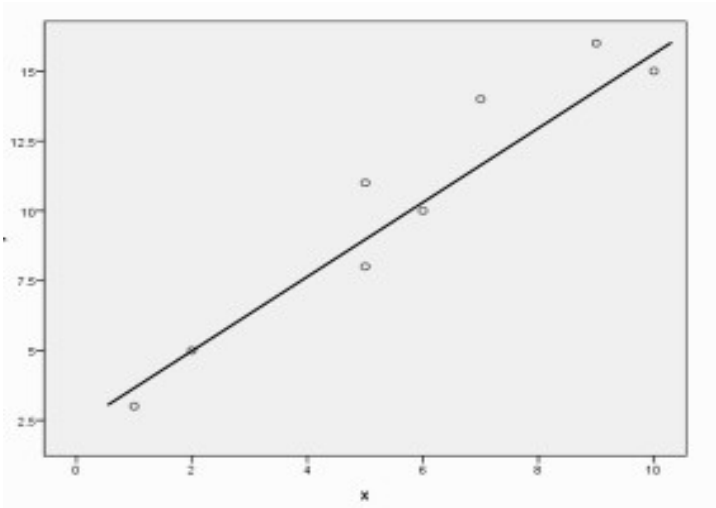
9	16
5	8
∑ 45	82

الحل:

x	y	x.y	x^2	y^2
10	15	150	100	225
2	5	10	4	25
5	11	55	25	121
6	10	60	36	100
7	14	98	49	196
1	3	3	1	9
9	16	144	81	256
5	8	40	25	64
∑ 45	82	560	321	996

حساب r:

$$R_p = \frac{n\sum(x.y)-(\sum x)(\sum y)}{\sqrt{[n(\sum x^2)-(\sum x)^2\ ][n(\sum y^2)-(\sum y)^2]}} = \frac{8.560-45.82}{\sqrt{[8.321-2025][8.996-6724]}} 0.96$$



R = 0.96

وبالتالي العلاقة طردية موجبة قوية جدا