

الانحدار والأرتباط البسيط Simple Regression & Correlation

عندما تكون هناك مشاهدات مزدوجة فإن احدى الافتراضات التالية توضح لتفسير ماهي العلاقة بين المشاهدات :

1. عدم وجود علاقة بين المشاهدات وتحلل بصورة منفردة او منفصلة .
2. وجود علاقة بينهما وتحدد هذه العلاقة بأستخدام اختبار الارتباط .
3. لتقدير هذه العلاقة الحقيقية بين x و y ووصفها في شكل معادلة بحيث يمكن التنبأ فيها عن قيمة y بدلالة x فإن هذا ما يسمى بالانحدار Regression اما لقياس درجة العلاقة بين المتغيرين او قياس مدى التلازم او الترابط بين متغيرين مستقلين فإن هذا ما يسمى بالارتباط Correlation .

معامل الارتباط Correlation Coefficient:

يمثل درجة الترابط او التلازم بين المتغيرين المستقلين والمعرفة فيما اذا كان هناك علاقة بين هذين المتغيرين ام لا يحسب معامل الارتباط (R) وان قيمة معامل الارتباط تقع بين (-1، 1) حيث عندما تكون قيمة R سالبة اي زيادة احد المتغيرين ينقص الاخر (الارتباط عكسي) وعندما تكون R موجبة اي زيادة احد المتغيرين يزيد المتغير الاخر (الارتباط طردي) ، ويكون الارتباط قوي كلما قربت القيم من $+1$ و -1 ونادرا ما تكون القيمة (1) وكلما اقترب الارتباط من الصفر اصبح الارتباط ضعيف واذا كانت قيمة الارتباط صفر هذا يعني عدم وجود علاقة او اي ارتباط بين المتغيرين .

* لتقدير واختبار معنوية معامل الارتباط البسيط بين المتغيرين x و y تجرى العمليات التالية :

1. حساب قيمة متوسطات المتغيرين \bar{x} , \bar{y}
2. ايجاد مجاميع مربعات انحرافات قيم x_i و y_i عن وسطها الحسابي SSX و SSY وكذلك مجموع مربعات الانحرافات المشتركة بين x و y وتسمى $SSXY$.
3. حساب قيمة معامل الارتباط البسيط (r) كالآتي :

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$$

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

مثال :

احسب معامل الارتباط للبيانات التالية والتي تمثل طول وعرض الورقة لنبات ما

عرض الورقة : $x = 13, 15, 17, 14, 17, 14, 18, 13, 19, 13$

طول الورقة : $y = 18, 15, 19, 15, 20, 13, 20, 15, 22, 15$

الجواب :

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
13	15	195	169	225
19	22	418	361	484
13	13	169	169	169
18	20	360	324	400
14	13	182	196	169
17	20	340	289	400
14	15	210	196	225
17	19	323	289	361
15	15	225	225	225
16	18	288	256	324
156	170	2710	2474	2982

$$r = \frac{\sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n}}{\sqrt{\left(\sum x_i^2 - \frac{(\sum x_i)^2}{n}\right) \left(\sum y_i^2 - \frac{(\sum y_i)^2}{n}\right)}}$$

$$= \frac{2710 - \frac{(156)(170)}{10}}{\sqrt{\left(2474 - \frac{(156)^2}{10}\right) \left(2982 - \frac{(170)^2}{10}\right)}} = 0.95$$

الانحدار البسيط : Simple regression

يطلق مصطلح الانحدار على العلاقة الكمية بين المتغيرات والتي تشمل متغير تابع ومتغير مستقل .

المتغير التابع : هو المتغير الذي يتأثر او يتغير بمتغير اخر .

المتغير المستقل : هو المتغير الذي يؤثر على متغير او على متغيرات ولا يتأثر بها وعادة يكون المتغير المستقل سابقا من الناحية الزمنية للمتغير التابع . كما هو الحال بدراسة العلاقة بين المتغير المستقل (x) وهو درجة الحرارة واعداد البكتريا (y) وهو المتغير التابع .

تطلق كلمة الانحدار الخطي البسيط على كمية بين المتغير الواحد المستقل (x) والمتغير التابع (y) اما اذا زادت عدد المتغيرات المستقلة فيطلق عليه بالانحدار المتعدد .

تعود الفائدة في قياس الانحدار البسيط الى :

- 1 . وصف العلاقة الكمية بين المتغير التابع والمتغير المستقل .
- 2 . السيطرة على مستوى المتغير التابع عن طريق التحكم بمستوى المتغير المستقل .
- 3 . التنبؤ بقيم المتغير التابع عند مستويات محددة من المتغير المستقل .

تحليل الانحدار الخطي البسيط:

معنى الانحدار الخطي اي انه النقاط المحسوبة للمتغيرين تكون على خط واحد مستقيم واحد او قريب من الخط المستقيم .
ان تحليل الانحدار الخطي البسيط يهتم بتقدير واختبار المعنوية لقيمتين A و B اذ يرمز B الى مقدار الانحدار في الخط المستقيم و يرمز A الى نقطة تقاطع الانحدار مع المحور y والتي توضع في المعادلة :

$$\hat{y} = a + bx$$

خطوات التحليل :

- 1 . حساب قيمة b من المعادلة التالية
- 2 . حساب قيمة a من المعادلة
- 3 . رسم خط الانحدار البسيط بتحديد نقطتين على الاقل من قيمة \hat{y} المقدره في المعادلة اعلاه .

مثال :

البيانات التالية تمثل الدرجة الفصلية والدرجة النهائية في درس الاحصاء لاثني عشر طالب
اوجد معادلة خط الانحدار .

الدرجة النهائية y	الدرجة الفصلية x
٨٥	٦٥
٧٤	٥٠
٧٦	٥٥

٩٠	٦٥
٨٥	٥٥
٨٧	٧٠
٩٤	٦٥
٩٨	٧٠
٨١	٥٥
٩١	٧٠
٧٦	٥٠
٧٤	٥٥

الجواب :

الجدول التالي يوضح خطوات الحل لحساب معامل الانحدار

x_i	y_i	x_i^2	y_i^2	$x_i y_i$
٦٥	٨٥	4225	7225	5525
٥٠	٧٤	2500	5476	3700
٥٥	٧٦	3025	5776	4180
٦٥	٩٠	4225	8100	5850
٥٥	٨٥	3025	7225	4675
٧٠	٨٧	4900	7569	6090
٦٥	٩٤	4225	8836	6110
٧٠	٩٨	4900	9604	6860
٥٥	٨١	3025	6561	4455
٧٠	٩١	4900	8281	6370
٥٠	٧٦	2500	5776	3800
٥٥	٧٤	3025	5476	4070
725	1011	44475	85905	61685

$$b = \frac{\sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}$$

$$b = \frac{61685 - \frac{725 \cdot 1011}{12}}{44475 - \frac{(725)^2}{12}} = 0.897$$

$$a = \bar{y} - b\bar{x}$$

$$\bar{y} = \frac{\sum y_i}{n} = \frac{1011}{12} = 84.250$$

$$\bar{x} = \frac{\sum x_i}{n} = \frac{725}{12} = 60.417$$

$$a = 84.250 - (0.897)(60.417) = 30.056$$

$$\hat{y} = a + bx = 30.05 + 0.897x$$

لو عوضنا مثلا اي قيمتين من x المعطاة بالسؤال ولتكن ٥٠ و ٧٠ فإن

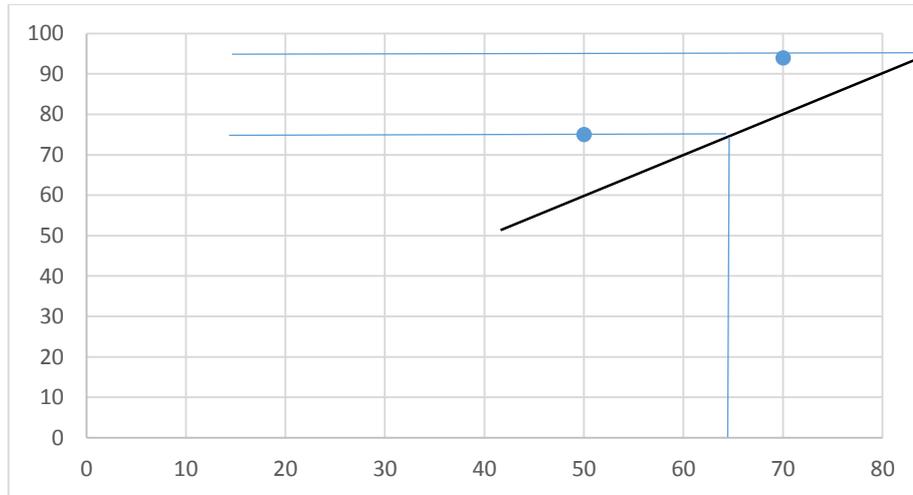
$$Y_{50} = 30.05 - 0.897(50) = 74.9$$

$$Y_{70} = 30.05 - 0.897(70) = 92.8$$

الخط الذي يمر بالنقطتين هو خط الانحدار

النقطة الاولى (٥٠ ، ٧٤,٩)

النقطة الثانية (٧٠ ، ٩٢,٨)



* من صفات خط الانحدار البسيط ان جميع القيم (النقط) (x,y) تقع على خط المستقيم فلو فرضنا وقمنا بالتعويض عن جميع القيم في معادلة الخط المستقيم ليجاد قيم y بدلالة x ستكون جميعها على خط مستقيم واحد وهو خط الانحدار .

** يلاحظ من خلال الفرق بين y و \hat{y} المقدرة هناك فرق وهذا الفرق يسمى الخطأ التجريبي .

*** على الطالب المحاولة ليجاد ورسم القيم الاخرى .