

الفصل السادس

Simple Regression and Correlation الانحدار والارتباط البسيط

مقدمة :

في الفصول السابقة كان اهتمامنا ينصب على قضايا الاحصاء الاستنتاجي التي نعود الى توزيع ذو متغير واحد هو X .

اما الان فسنحول اهتمامنا الى قضايا تخص توزيع ذو متغيرين (Birariate Dist) وسنرمز لهذين المتغيرين بالرمز X, Y فمثلا قد يكون المتغير X هو عدد نبات القطن في وحدة المساحة بينما المتغير Y هو كمية المحصول الناتج او قد يكون X هو درجات الحرارة بينما المتغير Y هو الكمية المذابة غم من الماء في مادة كيميائية معينة او قد تكون X هو المعدل الفصلي للطلبة بينما Y هو الدرجات النهائية لهم في مادة الاحصاء ومن ذلك يتضح بان كل فرد من الافراد العينة له قياسان احدهما للمتغير X والآخر للمتغير Y فمثلا لكل طالب درجتان هما معدله الفصلي (X) ودرجته النهائية (Y) .

وهناك نوعان من الانحدار :

- انحدار بسيط (Simple Regression)

- انحدار متعدد (Multiple Regression)

الانحدار البسيط (Simple Regression)

يشمل على متغيرين فقط هما مستقل (X) ومعتمد (Y)

اما الانحدار المتعدد (Multiple Regression)

يشمل على اكثر من متغيرين احدهما معتمد (Y) ومستقل ($x_1, x_2, \dots, x_n$)

مثال:

البيانات التالية تمثل الدرجة الفصلية والدرجة النهائية في درس الاحصاء لاثني عشر طالبا . اوجد معالجة الانحدار ؟

Xi	65	50	55	65	55	70	65	70	55	70	50	55
Yi	85	74	76	90	85	87	94	98	81	91	76	74

الحل :

Xi	65	50	55	65	55	70	65	70	55	70	50	55
Yi	85	74	76	90	85	87	94	98	81	91	76	74
Xi ²	4225	2500	3025	4225	3025	7569	4225	7569	3025	7569	2500	3025
Xi\Yi	5525	3700	4180	5850	4675	6090	6110	6860	4455	6370	3800	4070

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

$$= \frac{61685 - \frac{(725)(1011)}{12}}{4475 - \frac{(725)^2}{12}} = 0.897$$

$$a = y - bx$$

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

$$Y^* = Y = a + bx = 30.056 + (0.897)x$$

اي قيمة نأخذ لـ X ولتكونان 50 , 70 سوف نجد منها قيمة الـ Y* او Y فتكون :

$$Y(50) = 30.056 + 0.897(50) = 74.9$$

$$Y(70) = 30.056 + 0.897(70) = 92.8$$

والخط المستقيم الذي يمر بهاتين النقطتين هو الانحدار

وبالتعويض في هذه المعادلة عن كل قيمة من قيم فانه يمكن حساب قيم تقديرية الـ (Y, Y*) والتي

تقع جميعها على خط الانحدار البسيط وهذه القيم موضحة في الجدول التالي :

X_i	65	50	55	65	55	70	65	70	55	70	50	55
Y_i	85	74	76	90	85	87	94	98	81	91	76	74
$Y(x)$	88.4	74.9	79.4	88.4	79.4	92.8	88.4	92.8	79.4	92.8	74.9	79.4

اذن خطوات الحل لإيجاد معادلة الانحدار الخطي هي :

- نجد X_i^2 و $X_i Y_i$ في الجدول

- ثم نجد b, a

- ثم نجد معادلة خط الانحدار والتي هي :

$$Y^* = Y(x) = a + bx$$

ومن خواص معادلة الانحدار الخطي البسيط هو

- ان قيمة النقطة (x_i, y_i) تقع على خط الانحدار

- ان خط الانحدار يمر من جميع قيم (x_i, y_i)

- ان مجموع الانحرافات عن خط الانحدار يساوي صفر اي ان $\sum(y_i - y_x) = 0$

- ان مجموع مربعات الانحرافات عن خط الانحدار هي اقل ما يمكن اي ان

$$\sum(y_i - y_x)^2 = \min$$

هذا يمكن ان نستخدم معادلة خط الانحدار للتنبؤ عن قيمة Y لقيمة معينة من Y غير موجودة في

العينة مثل الدرجة النهائية المتوقعة لطالب معدل الفصلي (73) هي .

$$Y(73) = 30.056 + 0.897(73) = 95.5$$