

الانحدار Regression

هو محاولة لنمذجة العلاقة الخطية بين متغيرين وذلك بايجاد معادلة العلاقة الخطية بين البيانات الاحصائية ، وفيها احد المتغيرات يعد متغيرا استكشافيا والثاني متغير معتمدا ، حيث يمثل الخط المستقيم للعلاقة بين المتغيرين هو خط الانحدار ، والمعادلة الممثلة له هي معادلة الانحدار.

قبل البدا بايجاد علاقة الانحدار لابد من فحص البيانات والتأكد من وجود علاقة ارتباط خطية بينهما ، وذلك برسم العلاقة البيانية لتوزيع النقاط بين المتغيرين على شكل بياني .

Slope: $b_1 = \frac{n(\sum xy) - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2}$

y-intercept: $b_0 = \bar{y} - b_1\bar{x}$

$$b_0 = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2}$$

$$Y = b_0 + b_1x$$

حيث ان \bar{x} و \bar{y} يمثلان الوسط الحسابي او المعدل لكل من المتغيرين x و y ، لذا فان خط الانحدار يمر بهما لذلك فهما يحققان معادلة خط الانحدار ، b_0 يمثل القيمة التي يتقاطع بها خط الانحدار مع المحور الصادر لشكل البياني. b_1 يمثل قيمة انحدار خط الانحدار. Y يمثل المتغير المعتمد و x يمثل المتغير المستقل .

مثال : جد معادلة الانحدار الخطي بين مقدار المسافة التي قطعها السيارة خلال

Mileage (x)	Price (y)	X ²	Y ²	xy
21.8	16	475.24	256	348.8
34	11.5	1156	132.25	391
41.7	13.4	1738.89	179.56	558.78
53.5	14.8	2862.25	219.04	791.8
65.8	10.5	4329.64	110.25	690.9
72.1	12.3	5198.41	151.29	886.83
76.5	8.2	5852.25	67.24	627.3
84.7	9.5	7174.09	90.25	804.65
450.1	96.2	28786.77	1205.88	5100.06
sum(x)	sum(y)	sum(x ²)	sum(y ²)	sum(xy)
56.2625	12.025			
avg	avg			

استخدامها وسعرها

$$b_1 = -0.09021$$

$$b_0 = 17.10034$$

$$y = 17.10034 - 0.09021x$$

ملاحظة : خلال تطبيق معادلة الانحدار يجب ازالة القيم الشاذة من البيانات

شروط استخدام الانحدار الخطي:

1. في حالة عدم وجود علاقة ارتباط خطية لا تستخدم معادلة الانحدار الخطي للتنبؤ بقيم المتغير المعتمد.
2. في حالة استخدام معادلة الانحدار الخطي يجب ان تبقى في مجال تطبيقها ولايجوز تعميمها على حالات اخرى.
3. معادلة الانحدار الخطي التي استنتجت من بيانات قديمة ، ليس بالضرورة ان تنطبق على البيانات الحالية.
4. لايجوز استخدام معادلة الانحدار الخطي لبيانات تختلف في موقعها وظروفها الجغرافية .