الانحدار Regression

الانحدار الخطي البسيط: Simple Linear Regression

كان اهتمامنا في في المحاضرات السابقه منصب حول قضايا الاحصاء الاستنتاجي التي تعود الى توزيع ذو متغير واحد ورمزنا لهذا المتغير بالرمز (y)

اما الان فتحول اهتمامنا الى قضايا تخص توزيع ذو متغيرين وسنرمز لهذين المتغيرين بالرمز (x) و(y) فمثلاً قد يكون المتغير (x) هو عدد نباتات القطن في وحده المساحه، بينما المتغير (y) هو كميه المحصول الناتج او قد يكون (x) هو درجات الحرارة، بينما المتغير (y) هو الكميه المذابة في 100 غرام من الماء من ماده كيميائيه معينه او قد يكون (x) هو المعدل الفصلي للطلبة، بينما (y) هو الدرجات النهائيه لهم لمادة الاحصاء ومن ذلك يتضح بأن كل فرد من افراد العينه له قياسان ان احدهما للمتغير (x) والاخر للمتغير (y) فمثلاً لكل طالب درجتان هما المعدل الفصلي (x) ودرجته النهائية (x)

ان الغاية الرئيسيه من دراسه توزيع ذو متغيرين هي:

أ. لتحديد العلاقه الحقيقيه بين (x) و(y) ووضعهما بشكل معادله بحيث يمكن التنبؤ منها عن (y) بدلاله (x) قوس (وهذا ما يسمى بالانحدار Regression)

ب. لقياس درجة العلاقة بين المتغيرين (وهذا مايسمى correlation او بمعنى اخر قياس مدى التلازم والقرابةبين متغيرين مستقلين.

إن الغرض من استخدام أسلوب تحليل الانحدار الخطي البسيط، هو دراسة وتحليل اثر متغير كمي على متغير كمي أخر، ومن الأمثلة على ذلك ما يلي:

- دراسة أثر كمية السماد على إنتاجية الدونم
- دراسة أثر الإنتاج على التكلفة
- دراسة أثر كمية البروتين التي يتناولها الأبقار على الزيادة في الوزن
- أثر الدخل على الإنفاق الاستهلاكي

وهكذا هناك أمثلة في كثير من النواحي الاقتصادية، والزراعية، والتجارية.

في تحليل الانحدار البسيط، نجد أن الباحث يهتم بدراسة أثر أحد المتغيرين ويسمى بالمتغير المستقل او المتنبا منه، على المتغير الثاني ويسمى بالمتغير التابع أو المتنبأ به، ومن ثم يمكن عرض نموذج الانحدار الخطي في شكل معادلة خطية من الدرجة الأولى، تعكس المتغير التابع كدالة في المتغير المستقل كما يلي:

$$\bar{y}_x = \alpha + \beta x$$

حيث أن:

y هو المتغير التابع الذي يتأثر

x هو المتغير المستقل الذي يؤثر

Dr. Ahmed Shehab

α هو الجزء المقطوع من المحور الصادي y ، وهو يعكس قيمة المتغير التابع في حالة انعدام قيمة المتغير المستقل x ، أي في حالة (x = 0).

β ميل خط المستقيم، ويعكس مقدار التغير في y اذا تغيرت x بوحدة واحدة

ان المعادلة اعلاه تسمى معادلة خط الانحدار للمجتمع وان α و β هي ثوابت

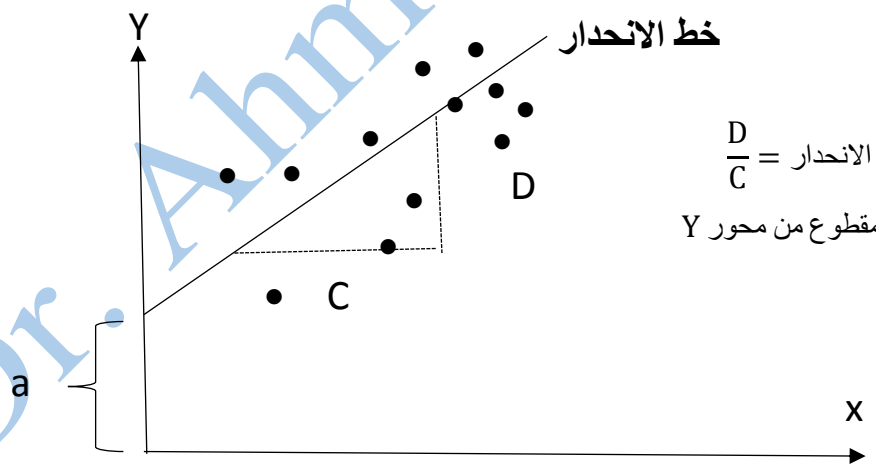
يعرف الثابت α بانه معدل قيمة y عندما x تساوي صفر وتسمى نقطة تقاطع خط الانحدار مع المحور الصادي (y- Intercept)

يعرف الثابت β بميل خط الانحدار للمجتمع (ويسمى معامل انحدار y على x بأنه معدل التغيير في y عندما تتغير قيمة x قيمة واحدة وعند تقدير الميل من العينة يرمز له b

وبذلك نمثل معادلة خط الانحدار للعينة بالمعادلة التالية

$$y = a + bx$$

ومن المعادلة اعلاه يمكن التنبؤ بقيمة y من قيم x عند معرفة الثوابت a و b



$\frac{D}{C}$ = ميل خط الانحدار = b

a = الجزء المقطوع من محور Y

كيفية حساب معامل الانحدار b

يحسب معامل الانحدار او ما يسمى ايضا ميل خط الانحدار b بالمعادلة التالية:

$$\mathbf{b} = \frac{\sum(\mathbf{X_i} - \overline{\mathbf{X}})(\mathbf{Y_i} - \overline{\mathbf{Y}})}{\sum(\mathbf{X_i} - \overline{\mathbf{X}})^2}$$

Dr. Ahmed Sh

او يمكن استعمال المعادلة المختصرة التالية

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

كيفية حساب معامل تقاطع خط الانحدار مع المحور الصادي a:

يمكن حساب معامل تقاطع خط الانحدار مع المحور الصادي a بالمعادلة التالية:

$$a = \overline{Y} - b\overline{X}$$

استخدامات تحليل الانحدار:

يستخدم تحليل الانحدار في المجالات التالية:

1. وصف العلاقة الكمية بين المتغير التابع والمتغير المستقل.
2. التنبؤ بقيم المتغير التابع عند مستويات محددة للمتغير المستقل.
3. السيطرة على المتغير التابع عن طريق التحكم بمستوى المتغير المستقل الذي يؤثر فيه.

الدرجة الفصلية x	الدرجة النهائية y
65	85
50	74
55	76
65	90
55	85
70	87
65	94
70	98
55	81
70	91
50	76
55	74

Dr. Ahmed Shehab

x_i	y_i	x_i^2	y_i^2	x_iy_i
65	85	4225	7225	5525
50	74	2500	5476	3700
55	76	3025	5776	4180
65	90	4225	8100	5850
55	85	3025	7225	4675
70	87	4900	7569	6090
65	94	4225	8836	6110
70	98	4900	9604	6860
55	81	3025	6561	4455
70	91	4900	8281	6370
50	76	2500	5776	3800
55	74	3025	5476	4070
$\sum x_i = 725$	$\sum y_i = 1011$	$\sum x_i^2 = 44475$	$\sum y_i^2 = 85905$	$\sum x_iy_i = 61685$

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

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

$a = \bar{Y} - b\bar{X}$

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

$\bar{y}_x = \alpha + bx$

وبالتعويض بمعادلة خط الانحدار نحصل على:

$\bar{y}_x = 30.05 + 0.897x$

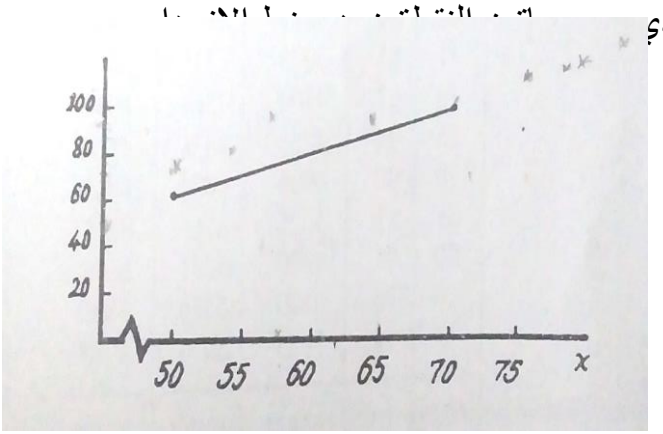
والان لو عوضنا عن اي قيمتين من قيم x المعطاة بالسؤال ولتكونا 50 و 70 فأن

$\bar{y}_{50} = 30.05 + 0.897(50) = 74.9$

$\bar{y}_{70} = 30.056 + 0.897(70) = 92.8$

Dr. Ahmed Shehab

والخط المستقيم الذي



هذا بالتعويض في هذه المعادلة عن كل قيمة من قيم x فأنة يمكن حساب القيم التقديرية لـ y اي $\bar{y}_x$ او $(\hat{y})$ والتي تقع جميعها على خط الانحدار البسيط وهذه القيم موضحة في الجدول التالي:

x_i	y_i	$(\hat{y})$ او $\bar{y}_x$
65	85	88.4
50	74	74.9
55	76	79.4
65	90	88.4
55	85	79.4
70	87	92.8
65	94	88.4
70	98	92.8
55	81	79.4
70	91	92.8
50	76	74.4
55	74	79.4

هذا ومن خواص معادلة الانحدار الخطي البسيط هو:

1. ان قيمة النقطة $(\bar{y}_1, \bar{x})$ تقع على خط الانحدار.
2. ان خط الانحدار يمر من جميع قيم $(x_i, y_{i.....})$
3. ان مجموع الانحرافات عن خط الانحدار يساوي صفراً اي $\sum(X_i - \bar{X}_x)^2 = 0$
4. وان مجموع مربعات الانحرافات عن خط الانحدار هي اقل ما يمكن اي

$$\sum(X_i - \bar{X}_x)^2 = \text{minimum}$$

هذا ويمكن استخدام معادلة خط الانحدار للتنبؤ عن قيمة **y** لقيمة معينة من **x** غير موجودة في العينة

فمثلاً الدرجة النهائية المتوقعة لطالب معدلة الفصلي 73 هي

Dr. Ahmed Shehab

$$\bar{y}_{50} = 30.05 + 0.897(73) = 95.5$$

المصادر:

1. الراوي، خاشع محمود. 1989. المدخل الى الإحصاء. وزارة التعليم العالي والبحث العلمي. جامعة الموصل.
2. طبيه، احمد عبدالسميع. 2007. مبادئ الإحصاء، عمان. دار البداية. ر.أ: www.daralbedayah.com (2007/6/1709)
3. David, M. Lane. Introduction to Statistics. Online Edition.

Dr. Ahmed Shehab