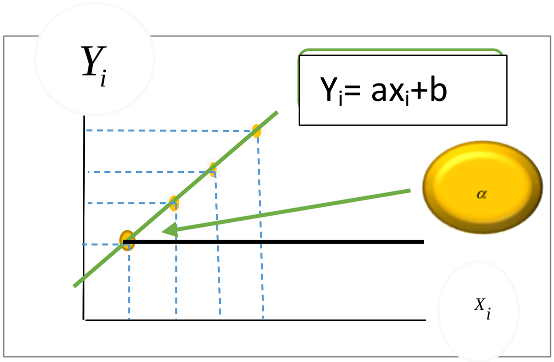
المحور التاسع: الارتباط و الانحدار

1. الانحدار الخطي البسيط:

إن الغرض من استخدام تحليل الانحدار الخطي البسيط هو دراسة وتحليل أثر متغير كمي على متغير كمي آخر.

1.1. الانحدار الخطي البسيط التام:

هنا تكون جميع نقاط شكل الانتشار على استقامة واحدة سواء كانت في الاتجاه الموجب أو في الاتجاه السالب.

تكون معادلة الإنحدار على الشكل التالي: $Y_i= ax_i+b$ حيث a ميل دالة الانحدار، وb ثابت الدالة.	

ملاحظة هامة: لرسم شكل الانتشار لابد من تحديد المتغير التابع والذي نرمز له بالرمز (Y_i) ويمثل على محور التراتيب، طبعا المتغير التابع يتأثر بالمتغير المستقل والذي نرمز له (Xi) والذي يمثل على محور الفواصل .

2.1 الانحدار الخطي البسيط غير التام

في هذه الحالة نقاط شكل الانتشار لا تكون على استقامة تامة، ولكنها تأخذ اتجاها يمكن تقريبه من معادلة خط مستقيم، حيث أنه يستحيل إيجاد المعادلة الحقيقية والتي نفترضها كما يلي:

$$Y_i= ax_i+b+e_i$$

ax_i+b هي معادلة خط مستقيم مضاف إليه المقدار e_i الذي يمثل قيمة البعد بين النقاط الحقيقية ومعادلة خط المستقيم. حيث أنه يستحيل إيجاد هذه المعادلة لذلك يتم تقريبها إلى المعادلة الخطية التالية:

$$\widehat{y_i} = \hat{a}x_i + \hat{b}$$

حيث:

$\widehat{y_i}$ القيمة التقديرية

$\hat{a}$ قيمة تقديرية ل a

$\hat{b}$ قيمة تقديرية ل b

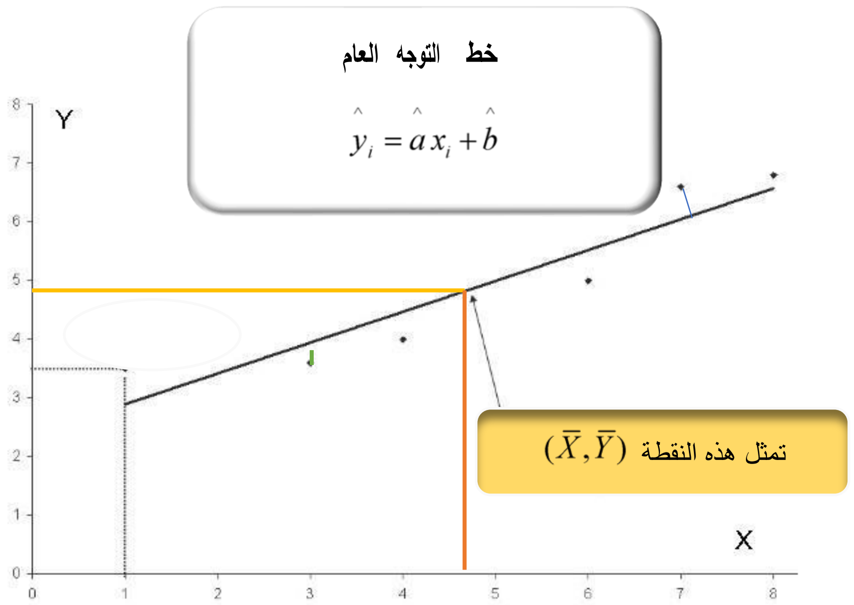
يمر الخط المستقيم من النقطة $(\bar{X},\ \bar{Y})$، حيث يمثل هذا المستقيم أقرب ما يمكن إلى جميع نقاط شكل الانتشار الحقيقية.

يتم إيجاد ثوابت الانحدار بما يسمى طريقة المربعات الصغرى والتي تعطي المعادلتين التقديريتين كما يلي :

$$\hat{a} = \frac{\sum_{i=1}^{n}(x_i - \bar{X})(y_i - \bar{Y})}{\sum_{i=1}^{n}(x_i - \bar{X})^2}$$

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

حيث أن: $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$ و $\bar{Y} = \frac{\sum_{i=1}^{n} y_i}{n}$



2. معاملات الارتباط

الهدف من تحليل الارتباط الخطي البسيط هو تحديد نوع وقوة العلاقة بين متغيرين. يرمز له بالرمز r، حيث $-1 \leq r \leq +1$. قد تكون علاقة عكسية اذا كان $r < 0$ أو طردية اذا كان $r > 0$ أو لا توجد علاقة اذا كان $r = 0$

ارتباط طردي					ارتباط عكسي				
قوي جدا	قو ي	متوسط	ضعيف	ضعيف جدا	ضعيف جدا	ضعيف	متوسط	قو ي	قوي جدا
[0.9,1[	[0.7,0.9[	[0.5,0.7[	[0.3,0.5[	[0,0.3[	]−0.3,0[	[−0.5, −0.3[	[−0.7, −0.5[	[−0.9, −0.7[	[−1, −0.9[

ملاحظات هامة:

- ✓ لما $r= \pm 1$ نقول أنه يوجد ارتباط تام بين المتغيرين؛
- ✓ لما $r = 0$ نقول لا يوجد ارتباط بين المتغيرين ؛
- ✓ $r < 0$ هناك علاقة عكسية بين المتغيرين؛
- ✓ $r > 0$هناك علاقة طردية بين المتغيرين.

1.2 معامل الارتباط الخطي البسيط:

يعرف بمعامل الارتباط لبيرسون وهو خاص بالمتغيرات الكمية ويحسب بالعلاقة التالية :

$r = \frac{cov(x,y)}{\sigma_X . \sigma_y}$ ، حيث cov (x,y) هو التباين المشترك (التغاير) بين المتغيرين (x,y) حيث:

- $cov(x,y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{X})(y_i - \bar{Y})$
- σ_X , σ_y الانحراف المعياري لكل من x و y.

يمكن اختصار علاقة حساب r كما يلي:

$$r = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{X})(y_i - \bar{Y})}{\sqrt{\frac{\sum(x_i - \bar{X})^2}{n}} \times \sqrt{\frac{\sum(y_i - \bar{Y})^2}{n}}}$$

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{X})(y_i - \bar{Y})}{\sqrt{\sum(x_i - \bar{X})^2} \times \sqrt{\sum(y_i - \bar{Y})^2}}$$

2.2. معامل ارتباط الرتب:

إذا كانت متغيرات الدراسة متغيرات وصفية ترتيبية، فإنه يمكن حساب معامل ارتباط لسبيرمان يعتمد أساسا على رتب المتغيرين، يعبر عنه بالعلاقة التالية:

$$r_{spearman} = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}$$

حيث:

d هي الفرق بين ترتيب x و y؛

$$d_i = R_X - R_Y$$

N عدد البيانات (x,y)

ملاحظة: يمكن كذلك استخدام صيغة معامل ارتباط الرتب في حساب الارتباط بين متغيرين كميين، حيث يتم استخدام رتب القيم التي يأخذها المتغير محل الدراسة .