المحور الثاني: الانحدار الخطي البسيط وطريقة المربعات الصغرى العادية

المحاضرة الثالثة: نموذج الانحدار الخطي البسيط

1- مفهوم نموذج الانحدار الخطي البسيط:

يعتبر نموذج الانحدار الخطي البسيط أبسط النماذج القياسية، التي يعتمد عليها في دراسة العلاقات والتأثيرات بين المتغيرات، وهو يعبر عن وجود علاقة خطية (سواء طردية أو عكسية) بين متغيرين أحدهما متغير تابع والثاني متغير مستقل.

2- استخدامات الانحدار الخطي البسيط:

أهم استخدامات معادلة الانحدار الخطي البسيط هو التنبؤ بقيم المتغير التابع والتي تسمى بالمتغيرات التقديرية (Estimate Value) وذلك من خلال تحديد قيم المتغير المستقل.

3- معادلة الانحدار الخطي البسيط:

تأخذ معادلة الانحدار الخطي البسيط الشكل التالي:

$$Y_i = \alpha + \beta X_i + U_i$$

حيث: i رقم المشاهدة $i = (1,2, \dots, n)$

X_i : قيمة المشاهدة i الخاصة بالمتغير المستقل

Y_i: قيمة المشاهدة i الخاصة بالمتغير التابع

U_i : قيمة الخطأ i الخاصة بالمتغير العشوائي، وتعتبر المشاهدات الخاصة بالمتغير العشوائي مقدرة وغير حقيقية.

α, β : معلمات النموذج التي تحدد طبيعة العلاقة بين المتغيرات.

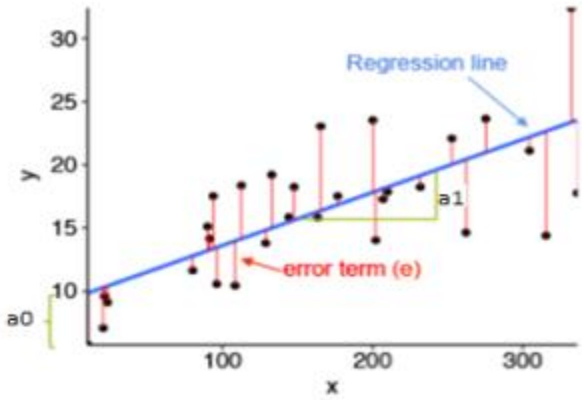
ويسمى النموذج بالانحدار البسيط لوجود العلاقات التالية:

- المتغير Y_i ينحدر على المتغير X_i باعتبار الأول تابع والثاني مستقل.
- بسيط لأنه يمثل العلاقة بين متغيرين فقط.
- خطي لأن العلاقة بين متغيراته ومعلماته تأخذ شكل خط مستقيم.
- طبيعة العلاقة بين المتغيرين تنحدر من خلال المعلمة β، فإذا كانت سالبة فالعلاقة ستكون عكسية وبالتالي $Y = \alpha - \beta X$، وإذا كانت موجبة فالعلاقة تكون طردية $Y = \alpha + \beta X$

4- فرضيات النموذج:

1- النموذج خطي.
2- المتغير المستقل X_i محدد وثابت وعدم وجود أخطاء في قياس المتغير المستقل.
3- الأخطاء العشوائية مستقلة ولها توزيع طبيعي وسطه الحسابي يساوي الصفر $E(\blacksquare_i) = 0$، أي أن القيم تتمركز حول 0 وتباينها يساوي قيمة ثابتة وموجبة $V(U_i) = \sigma^2$ (ثبات تباين حد الخطأ).
4- المتغير المستقل X_i مستقل عن الخطأ العشوائي U_i أي: $Cov(X_i, U_i) = 0$
5- قيم الخطأ العشوائي غير مرتبطة فيما بينها $Cov(U_i, U_{i'}) = 0$ أي لا يوجد مشكل الارتباط الذاتي للأخطاء

5- تقدير نموذج الانحدار الخطي البسيط باستعمال طريقة المربعات الصغرى MCO:



الشكل أعلاه يعرف بسحابة النقط التي تتمثل في المساحة التي تحددها النقط السوداء في الشكل، وتهدف عملية تقدير معلمات النموذج αوβ إلى تحديد أمثل خط انحدار يمر عبر السحابة مقلصا الأخطاء إلى أدنى حد ممكن، حيث من غير الممكن وجود خط مستقيم يعبر كل النقاط، لهذا الهدف من التقدير هو إيجاد أمثل خط (الخط الأزرق في الشكل) بحيث يتم تقليل الأخطاء، (الخطأ هو تلك الفجوة بين النقطة وإسقاطها على خط الانحدار)، وبالتالي يمكننا تسمية النقط السوداء Y بالقيم الحقيقية في حين تسمى إسقاطاتها على خط الانحدار بالقيم المقدرة $\widehat{Y}$، فيكون الخطأ $U_i = Y_i - \widehat{Y}_i$

ولغرض تقدير معلمات النموذج نستعمل طريقة المربعات الصغرى MCO (Moindres Carrés (Ordinary Least Squares Ordinares)/ OLS، حيث تعتمد هذه الطريقة على تقليص وتدنية مجموع مربع الأخطاء إلى أدنى حد ممكن $Min\sum U_i^2$.

$$y_i = \alpha + \beta x_i + u_i$$

$$\widehat{y}_i = \hat{\alpha} + \hat{\beta} x_i$$

$$Min\sum U_i^2 = Min\sum (y_i - \widehat{y}_i)^2 = Min\sum \left(y_i - \hat{\alpha} - \hat{\beta}x_i\right)^2 = Min\, S$$

$$\begin{cases} \dfrac{\partial s}{\partial \alpha} = 2\sum \left(y_i - \hat{\alpha} - \hat{\beta}x_i\right)(-1) = 0 \\ \dfrac{\partial s}{\partial \beta} = 2\sum \left(y_i - \hat{\alpha} - \hat{\beta}x_i\right)(-x_i) = 0 \end{cases}$$

$$\begin{cases} \sum y_i - n\hat{\alpha} - \hat{\beta}\sum x_i = 0 \\ \sum x_i y_i - \hat{\alpha}\sum x_i - \hat{\beta}\sum x_i^2 = 0 \end{cases}$$

$$\begin{cases} \sum y_i = n\hat{\alpha} + \hat{\beta}\sum x_i \\ \sum x_i y_i = \hat{\alpha}\sum x_i + \hat{\beta}\sum x_i^2 \end{cases}$$

$$\begin{cases} \dfrac{\sum y_i}{n} = \hat{\alpha} + \hat{\beta}\dfrac{\sum x_i}{n} \dots\dots (1) \\ \dfrac{\sum x_i y_i}{n} = \hat{\alpha}\dfrac{\sum x_i}{n} + \hat{\beta}\dfrac{\sum x_i^2}{n} \dots\dots (2) \end{cases}$$

من العلاقة (1) نجد $\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$

$$\frac{\sum x_i y_i}{n} = (\bar{y} - \hat{\beta}\bar{x})\bar{x} + \hat{\beta}\frac{\sum x_i^2}{n}$$

$$\frac{\sum x_i y_i}{n} = \bar{x}\bar{y} - \hat{\beta}\bar{x}^2 + \hat{\beta}\frac{\sum x_i^2}{n}$$

$$\frac{\sum x_i y_i}{n} = \bar{x}\bar{y} + \hat{\beta}(\frac{\sum x_i^2}{n} - \bar{x}^2)$$

$$\hat{\beta} = \frac{\frac{\sum x_i y_i}{n} - \bar{x}\bar{y}}{\frac{\sum x_i^2}{n} - \bar{x}^2} = \frac{Cov\,(x,y)}{\sigma^2(x)}$$