الانحدار الخطي المتعدد

يعتبر الانحدار الخطي المتعدد من بين الأدوات المهمة المستعملة في التنبؤ على المستوى القصير والمتوسط حيث يستعمل في اتخاذ القرار والتوقع والرقابة ويهتم كذلك الانحدار المتعدد في الأصل بتحديد ودراسة طبيعة العلاقة بين المتغير التابع (y) وعدة متغيرات مستقلة (x_1, x_2, ...x_k).

نفترض في الانحدار المتعدد الخطي وجود علاقة خطية بين المتغير التابع والمتغيرات المستقلة كما هو مبين في العلاقةI التالية

$$y = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \cdots + \beta_k x_{nk} \ldots\ldots\ldots\ldots.(\mathbf{I})$$

لكن في الواقع الأصلي لا توجد علاقة خطية تامة بين المتغير التابع والمتغيرات المستقلة وذلك لعدة أسباب منها أخطاء القياس و وجود متغيرات أخرى تؤثر على المتغير التابع لكن لا يمكن قياسها أو أخذها بعين الاعتبار .لذلك لتكون العلاقة (I) أكثر تصويراً للواقع نضيف لها مقدار يرمز له ب(ε_i) ويسمى حد الخطأ العشوائي أو معامل الإزعاج، هذا المعامل (ε_i)يفترض أنه يجمع جميع تلك العوامل أو المؤثرات التي يمكنها أن تؤثر على المتغير التابع y والتي لا يمكننا قياسها كذلك يضم(ε_i)أخطاء القياس التي تؤثر على العلاقة السابقة وبالتالي لكي تصبح علاقة خط الانحدار السابقة أكثر تصويرا للواقع أ؟و بالأحرى لتكون أكثر دقة نضيف المتغير العشوائي للعلاقة I لتصبح كما يلي:

$y = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \cdots + \beta_k x_{nk} + \varepsilon_i$

$y_n = \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \beta_3 x_{n3} + \cdots + \beta_k x_{nk} + \cdots + \varepsilon_n$

$i = 1,2,3, \dots n$

K: عدد المعلمات المقدرة

$\beta_0, \beta_1, \beta_2, \beta_3 \dots \beta_k$: معالم أو معاملات أو معلمات النموذج المراد تقديرها

ε_i : المتغير العشوائي أو حد الخطأ العشوائي أو معامل الإزعاج

ملاحظة: يجب أن نعلم أن طريقة الإنحدار الخطي المتعدد لا يمكن تطبيقها إلا على المتغيرات التي توجد بينها علاقة خطية أو تقبل للتحويل إلى علاقة خطية (نفس ما رأيناه سابقا في الانحدار الخطي البسيط).

شكل الانحدار الخطي المتعدد:

$$\begin{pmatrix} y_1 = & \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \beta_3 x_{13} + \cdots + \beta_k x_{1k} \cdots + & \varepsilon_1 \\ \vdots & \ddots & \vdots \\ y_n = & \beta_0 + \beta_1 x_{n1} + \beta_2 x_{n2} + \beta_3 x_{n3} + \cdots + \beta_k x_{nk} + & \varepsilon_n \end{pmatrix}$$

جملة المعادلات السابقة تتضمن $k+1$ من المعلمات المراد تقديرها مع العلم أن الحد الأول منها هو β_0 يمثل الحد الثابت لتقدير هذه المعلمات نحتاج اللجوء إلى علم المصفوفات والأشعة وعليه يمكن صياغة تلك الجملة في صورتها المصفوفية كالتالي:

$$\begin{pmatrix} y_1 = & (1 & x_{11} & x_{12} & x_{13} & \dots x_{1k})\beta_0 + & \varepsilon_1 \\ \vdots & & & & \ddots & & \vdots \\ y_n = & (1 & x_{n1} & x_{n2} & x_{n3} & \dots x_{nk})\beta_k + & \varepsilon_n \end{pmatrix}$$

حيث: $\underline{\underline{y}}$ شعاع عمودي أبعاده ($1 \underline{\underline{x}}$ h) يحتوي مشاهدات المتغير التابع.

* $\underline{\underline{x}}$ مصفوفة المتغيرات المستقلة أبعادها [n x(k+1)] حيث يحتوي العمود الأول فيها على القيمة1 (معاملات المعلمة β_0) وباقي الأعمدة تمثل قيم المتغيرات المستقلة.

* $\underline{\underline{\beta}}$ شعاع عمودي أبعاده [(k+1)x1] يحتوي على المعالم المراد تغييرها.

* ε_i شعاع عمودي أبعاده) (n x 1) يحتوي على حدود الخطأ العشوائي.

1.3. فرضيات النموذج الخطي المتعدد:

عند استعمال طريقة المربعات الصغرى التي رأينهاها سابقا في فإنه من الضروري تحقق الفرضيات التالية:

1/ القيمة المتوقعة لمتجه (شعاع) حد الخطأ العشوائي تساوي $E(\varepsilon_i)=0$

2/ تباين العناصر العشوائية ثابتة، والتباين المشترك بينها =0.

3/ ليست هناك علاقة خطية تامة بين المتغيرات المستقلة وأن عدد المشاهدات يجب أن يزيد على عدد المعالم المراد تقديرها.

4/ مصفوفة المتغيرات المستقلة x محددة ومعطاة فهي مقاسة بدون أخطاء.

5/ يوجد استقلال إحصائي بين المشاهدات المستقلة ($x_1, x_2 \dots\dots x_k$) والخطأ العشوائي ε_i

6/ الأخطاء العشوائية ε_i لها توزيع طبيعي متوسطه صفروتباينه ثابت

2.3. تقدير معلمات النموذج (β_i) باستخدام طريقة المربعات الصغرى:

بصفة عامة يتم تقدير معالم النموذج β_j باستخدام حساب المصفوفات والأشعة لطريقة المربعات الصغرى بالعلاقات التالية:

$$Y = XB + \varepsilon$$

$$\widehat{\beta_J} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_k \end{bmatrix}$$

$$\widehat{\beta_J} = (x'x)^{-1}x'y$$

$$(x'x)^{-1} = \frac{adj(x'x)}{det(x'x)}$$

مثال:

تعطي الكمية المباعة من السلعة y بدلالة كل من: سعرها (X_1) ، دخل المستهلك (X_2)، سعر السلعة البديلة (X_3) في الجدول التالي:

Y	**40**	**45**	**50**	**55**	**60**	**65**	**70**
سعر السلعة x_1	**9**	**8**	**9**	**8**	**7**	**6**	**6**
دخل المستهلك x_2	**400**	**400**	**400**	**500**	**500**	**500**	**500**
سعر السلعة البديلة x_3	**10**	**14**	**12**	**13**	**11**	**15**	**16**

المصدر: محاضرات الأستاذ شويخي إسماعيل بتصرف

المطلوب: بافتراض أن هناك علاقة انحدار خطية بين المتغيرات المستقلة المذكورة والمتغير التابع y مستخدماً طريقة المربعات الصغرى قدر معلمات النموذج:

$$\boxed{\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_3 x_3}$$

الحـــل:

$$(\hat{\beta}_j) = (x'x)^{-1} x'y$$

مصفوفة المتغيرات المستقلة

$$\mathbf{X} = \begin{array}{cccc} B_j & x_1 & x_2 & x_3 \\ \end{array}$$
$$\left|\begin{array}{cccc} 1 & 9 & 400 & 10 \\ 1 & 8 & 400 & 14 \\ 1 & 9 & 400 & 12 \\ 1 & 8 & 500 & 13 \end{array}\right|$$

$$\begin{matrix} 1 & 7 & 500 & 11 \\ 1 & 6 & 500 & 15 \\ 1 & 6 & 500 & 16 \end{matrix}_{7x4}$$

* منقول مصفوفة المتغيرات المستقلة

$$(x') = \begin{vmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 9 & 8 & 9 & 8 & 7 & 6 & 6 \\ 400 & 400 & 400 & 500 & 500 & 500 & 500 \\ 10 & 14 & 12 & 13 & 11 & 15 & 16 \end{vmatrix}_{4x7}$$

* مصفوفة المتغير التابع:

$$y = \begin{vmatrix} 40 \\ 45 \\ 50 \\ 55 \\ 60 \\ 65 \\ \end{vmatrix}$$

70

7x1

$$\mathbf{x'x} = \begin{vmatrix} 7 & 53 & 3200 & 91 \\ 53 & 411 & 23900 & 677 \\ 3200 & 23900 & 1480000 & 41900 \\ 91 & 677 & 41900 & 1211 \end{vmatrix}_{4x4}$$

* تذكرة لمقلوب المصفوفة:

مثال:

$$A = \begin{pmatrix} 2 & 1 & -1 \\ 2 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

- عين مقلوب المصفوفة **A**
- يرمز لها بـ $\mathbf{A^{-1}}$

$$\boxed{A^{-1} = \frac{1}{\det(A)}(CoA)^t}$$

$$\det(A) = \begin{matrix} + \\ - \\ + \end{matrix}\begin{vmatrix} 2 & 1 & -1 \\ 2 & 2 & 1 \\ 0 & 1 & 1 \end{vmatrix}$$

$$= +2\begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} \quad -2\begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix} \quad +0\begin{vmatrix} 1 & -1 \\ 2 & 1 \end{vmatrix}$$

$$= 2(1) - 2(2) + 0 = -2$$

$$\boxed{\det(A) = 2}$$

$$\textbf{Co A=} \begin{vmatrix} +\begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} & -\begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix} & +\begin{vmatrix} 2 & 2 \\ 0 & 1 \end{vmatrix} \\ -\begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix} & +\begin{vmatrix} 2 & -1 \\ 0 & 1 \end{vmatrix} & -\begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix} \\ +\begin{vmatrix} 1 & -1 \\ 2 & 1 \end{vmatrix} & -\begin{vmatrix} 2 & -1 \\ 2 & 1 \end{vmatrix} & +\begin{vmatrix} 2 & 1 \\ 2 & 2 \end{vmatrix} \end{vmatrix} = \begin{vmatrix} +1 & -2 & +2 \\ -2 & +2 & -2 \\ 3 & -4 & +2 \end{vmatrix}$$

$$(\textbf{Co A})^t = \begin{vmatrix} 1 & -2 & 3 \\ -2 & 2 & -4 \\ 2 & -2 & 2 \end{vmatrix}$$

نقسم على **det** (**A**)

ومنه يصبح لدينا:

$$A^{-1} = \begin{pmatrix} -0,5 & 1 & -1,5 \\ 1 & -1 & 2 \\ -1 & 1 & -1 \end{pmatrix}$$

$$x'y = \begin{vmatrix} 35 \\ 240 \\ 179000 \\ 5100 \end{vmatrix}_{4x1}$$

حساب $(\mathbf{x'x})^{-1}$

بنفس الطريقة التي رأيناها في المثال السابق:

$$\boxed{(x'x)^{-1} = \frac{1}{\det(x'x)}(Cox'x)^{t}}$$

نجد:

$$(x'x)^{-1} = \begin{vmatrix} 197,91 & -10,42 & -0,17 & -3,08 \\ -10,42 & 0,59 & 0,01 & -0,16 \\ -0,17 & 0,01 & 0,0002 & 0,00016 \\ -3,08 & 0,16 & 0,0016 & 0,09 \end{vmatrix}_{4x4}$$

إذا: $(\hat{\beta}_j) = (x'x)^{-1}x'y$

$$(\hat{\beta}_j) = \begin{vmatrix} 26,34 \\ -3,43 \\ 0,09 \\ 0,93 \end{vmatrix}_{4x4} \longrightarrow \begin{vmatrix} \hat{\beta}_0 = 26,34 \\ \hat{\beta}_1 = -3,43 \\ \hat{\beta}_2 = 0,09 \\ \hat{\beta}_3 = 0,93 \end{vmatrix}_{4x1}$$

ومنه:

$$\boxed{\hat{y} = 26,34 - 3,43x_1 + 0,09x_2 + 0,93x_3}$$

من خلال معادلة خط الانحدار نلاحظ أن المتغير (x_1) الذي يمثل سعر السلعة له علاقة عكسية مع المبيعات y لأنه إشارة x_1 سالبة أي كلما زاد سعر السلعة كلما انخفض الطلب عليها و هذا يدل أن النموذج موافق للنظرية الاقتصادية لكن ذلك وحده لا يكفي لاستعمال النموذج في التبأ بل يجب عرضه على مجموعة من الاختارات كما تم تبيانه بحمد الله في الانحدار البسيط.وقبل ذلك يجب حساب معامل التحديد كما يلي.