الفصل الثالث: الانحدار الخطي المتعدد

يتضح مما سبق أن الانحدار الخطي البسيط يركز على دراسة العلاقة بين متغيرين أحدهما المتغير المستقل (X) والآخر المتغير التابع (Y)، غير أن واقع الحياة الاقتصادية والاجتماعية مبني بشكل عام على تأثر أية ظاهرة بأكثر من متغير مستقل، فدالة الطلب مثلا، تحدد العلاقة بين الكميات المطلوبة من سلعة معينة، والاسعار الخاصة بهذه السلعة، وغيرها من السلع (أسعار السلع البديلة) ويدخل فيها دخل المستهلك كأحد المتغيرات فضلاً عن المتغيرات الأخرى، لذلك لابد من توسيع نموذج الانحدار السابق ليشتمل على انحدار للمتغير التابع (Y) على العديد من المتغيرات المستقلة $(X_1, X_2, \dots, X_k)$، ويسمى هذا بنموذج الانحدار الخطي المتعدد (Multiple Linear regression)

ويهدف هذا المحور إلى توضيح كيفية تقدير نموذج الانحدار الخطي المتعدد، الذي يتكون من متغير تابع ومتغيرين أو أكثر من المتغيرات المستقلة، لذلك سيتم مناقشة طبيعة نموذج الانحدار المتعدد، ثم تحديد أهم افتراضات النموذج التي سبق وأن تعرفنا عليها في المور السابق، يضاف إلى ذلك بيان عدم وجود علاقة خطية تامة بين المتغيرات المستقلة وكيف ان المصفوفة $(X'X)$، تكون مصفوفة غير شاذة $(Non-Singular)$ إذا كان محددها لا يساوي صفرا، ثم يتم بعد ذلك تقدير معلمات النموذج، تقدير التباين والتباين المشترك والانحراف المعياري لها للوصول إلى اختبار معاملات النموذج.

- **النموذج الخطي المتعدد:**

يستند النموذج الخطي المتعدد على افتراض وجود علاقة بين متغير مستقل تابع

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik} + \mu_i \dots\dots\dots (I)$$

وفي واقع الأمر، فإن هذه المعادلة هي واحدة من جملة معادلات يبلغ عددها (n) تكون نظام المعادلات الآتي:

$$Y_1 = \beta_0 + \beta_1 X_{11} + \beta_2 X_{12} + \cdots + \beta_k X_{1k} + \mu_i$$

$$Y_i = \beta_0 + \beta_1 X_{21} + \beta_2 X_{22} + \cdots + \beta_k X_{2k} + \mu_2$$

$$\vdots \quad \vdots \quad \vdots \qquad \vdots \qquad\quad \vdots \qquad \vdots$$

$$Y_i = \beta_0 + \beta_1 X_{n1} + \beta_2 X_{n2} + \cdots + \beta_k X_{nk} + \mu_n$$

وهذه المعادلة تتضمن $(k+1)$ من المعلمات المطلوب تقديرها، علما أن الحد الأول منها (β_0) يمثل الحد الثابت، الأمر الذي يتطلب اللجوء إلى المصفوفات والمتجهات لتقدير تلك المعلمات (بخيت و فتح الله، 2006، الصفحات 145-144)، وعليه؛ يمكن صياغة هذه المعادلات في صورة مصفوفات كالتالي:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_3 \end{bmatrix} = \begin{bmatrix} 1 & X_{11} & X_{12} & X_{13} & X_{1k} \\ 1 & X_{21} & X_{22} & X_{23} & X_{2k} \\ 1 & X_{31} & X_{32} & X_{33} & X_{3k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & X_{N1} & X_{n2} & X_{n3} & X_{nk} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} + \begin{bmatrix} \mu_0 \\ \mu_1 \\ \mu_2 \\ \vdots \\ \mu_k \end{bmatrix}$$

وباختصار نقول:

$$Y = X\beta + \mu \dots \dots \dots (II)$$

Y: متجه عمودي أبعاده $(n \times 1)$ يحتوي مشاهدات المتغير التابع.

X: مصفوفة أبعادها $(n \times k + 1)$ تحتوي مشاهدات المتغيرات المستقلة، ويحتوي عمودها الأول على قيم الواحد الصحيح ليمثل الثابت.

β: متجه عمودي أبعاده $(k + 1 \times 1)$ ، ويحتوي على المعالم المطلوب تقديرها.

μ: متجه عمودي أبعاده $(n \times 1)$ يحتوي على الأخطاء العشوائية.

وبما أن المعادلة (II) هي العلاقة الحقيقية المجهولة والمراد تقديرها بإستخدام الإحصاءات المتوفرة عن المتغير التابع Y، والمتغيرات المستقلة $X_1, X_2, \dots, X_k$، فإنه يستوجب تحقق الفروض الأساسية الخاصة بـ μ_i التالية:

$$\mu_i \sim N(0, \sigma^2 In)$$

والذي يعني أن μ_i يتوزع توزيعا طبيعيا (N) متعدد المتغيرات لمتجه وسطه صفري (0) ومصفوفة تباين وتبيان مشترك عددية هي: $(\sigma^2 In)$.

- فرضيات النموذج الخطي المتعدد:

عند استخدام طريقة (OLS) في تقدير نموذج الانحدار الخطي المتعدد/ فإنخ يجب توفر الافتراضات الآتية:

1. القيمة المتوقعة لمتجه حد الخطأ تساوي صفرا أي أن $E(\mu_i) = 0$

$$E(\mu_i) = E\begin{bmatrix}\mu_1\\ \mu_2\\ \mu_3\\ \vdots\\ \mu_n\end{bmatrix} = \begin{bmatrix}E(\mu_1)\\ E(\mu_2)\\ E(\mu_3)\\ \vdots\\ E(\mu_n)\end{bmatrix} = \begin{bmatrix}0\\ 0\\ 0\\ \vdots\\ 0\end{bmatrix} = 0$$

- تباين العناصر العشوائية ثابت، والتباين المشترك بينها يساوي صفرا؛ أي أن:

$$Cov(\mu) = E(\mu\mu') = \sigma^2 In$$

$$E(\mu\mu') = E\begin{bmatrix}\mu_1\\ \mu_2\\ \vdots\\ \mu_n\end{bmatrix}\begin{bmatrix}\mu_1 & \mu_2 & \cdots & \mu_n\end{bmatrix}$$

$$E(\mu\mu') = \begin{bmatrix}\mu_1^2 & \mu_1\mu_2 & \cdots & \mu_1\mu_n\\ \mu_2\mu_1 & \mu_2^2 & \cdots & \mu_2\mu_n\\ \vdots & \vdots & & \vdots\\ \mu_n\mu_1 & \mu_n\mu_2 & \cdots & \mu_n^2\end{bmatrix}$$

$$E(\mu\mu') = \begin{bmatrix}E(\mu_1^2) & E(\mu_1\mu_2) & \dots & E(\mu_1\mu_n)\\ E(\mu_2\mu_1) & E(\mu_2^2) & \dots & E(\mu_2\mu_n)\\ \vdots & \vdots & & \vdots\\ E(\mu_n\mu_1) & E(\mu_n\mu_2) & \cdots & E(\mu_n^2)\end{bmatrix}$$

وبما أن: $Var(\mu_i) = E\left(\mu_i^2\right) = \sigma^2$، فإن:

$$E(\mu\mu') = \begin{bmatrix}Var(\mu_1) & Cov(\mu_1\mu_2) & \dots & Cov(\mu_1\mu_n)\\ Cov(\mu_2\mu_1) & Var(\mu_2) & \dots & Cov(\mu_2\mu_n)\\ \vdots & \vdots & & \vdots\\ Cov(\mu_n\mu_1) & Cov(\mu_n\mu_2) & \cdots & Var(\mu_n)\end{bmatrix}$$

و بما أن $Cov\left(\mu_i\mu_j\right) = E\left(\mu_i\mu_j\right) = 0$ حيث أن $i \neq j$؛ يصبح لدينا:

$$E(\mu\mu') = \begin{bmatrix}\sigma_1^2 & 0 & \dots & 0\\ 0 & \sigma_2^2 & \dots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & \sigma_n^2\end{bmatrix}$$

حيث أن: $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_n^2$، نجد ان:

$$E(\mu\mu') = \sigma^2 \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$E(\mu\mu') = \sigma^2 In$$

وتسمى المصفوفة العددية أعلاه بمصفوفة التباين والتباين المشترك، $(Variance -$ $Covariance\ Matrix)$ لحد الخطأ (μ)، حيث تشكل العناصر القطرية في المصفوفة، تبيان قيم (μ) بينما تبقى العناصر غير القطرية (أعلى وأسفل القطر) (بخيت و فتح الله، 2006، صفحة 150)، مساوية للصفر لإنعدام التباين المشترك والترابط بين قين قيم (μ_i)

- ليس هناك علاقة خطية تامة بين المتغيرات المستقلة كما وأن عدد المشاهدات يجب أن يزيد على عدد المعلمات المطلوب تقديرها، أي أن:

$$r(x) = k + 1 < n$$

حيث أن (r) رتبة مصفوفة البيانات، (X) تساوي عدد المتغيرات المستقلة، (k) زائد (1) الحد الثابت، وهي أصغر من عدد المشاهدات (n)، وهذه الفرضية ضرورية جدا لضمان إيجاد معكوس المصفوفة $(X'X)$، إذ ان وجود هذا الفرض يجعل رتبة المصفوفة (X) أقل من $(k + 1)$، وبالتالي فإن رتبة $(X'X)$ التي تستخدم في الحصول على مقدرات (OLS) بدورها أقل من $(k + 1)$، ولا يمكن إيجاد معكوسها بسبب ما يسمى بمشكلة الارتباط الخطي المتعدد، وبالتالي لا يمكن الحصول على مقدرات المربعات الصغرى العادية (OLS).

- طرق تقدير معلمات النموذج:

في ضوء الفرضيات المذكورة أعلاه، يمكن استخدام (OLS)، في تقدير نعلمات النموذج الخطي المتعدد، ولهذا الغرض يمكن إعادة كتابة المعادلة (I) بصيغتاه التقديرية كالتالي:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2}$$

وبما أن هدفنا هو الحصول على قيم كل من $\left(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2\right)$ التي تجعل مجموع مربعات الانحرافات أقل ما يمكن، أي تصغير القيمة $\left(\sum e_i^2\right)$ (مبدأ المربعات الصغرى) إلى أقل قيمة ممكنة، أي:

$$\min \rightarrow \sum_{i=1}^{n} e_i^2$$

$$e_i = Y_i - \hat{Y}_i$$

$$\sum e_i^2 = \sum \left(Y_i - \hat{Y}_i\right)^2$$

ومن خلال التعويض بـ $\hat{Y}_i$ بما يساويها وأخذ المشتقات الجزئية بالنسبة إلى $\left(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2\right)$ ومساوتها بالصفر نحصل على:

$$\sum e_i^2 = \sum \left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{i1} - \hat{\beta}_2 X_{i2}\right)^2$$

$$\frac{\partial \sum e_i^2}{\partial \hat{\beta}_0} = 2 \sum \left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{i1} - \hat{\beta}_2 X_{i2}\right)(-1) = 0$$

$$2 \sum \left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{i1} - \hat{\beta}_2 X_{i2}\right)(-1) = 0$$

بالقسمة على (-2) وفك القوس؛ نجد:

$$\sum Y_i - n\hat{\beta}_0 - \hat{\beta}_1 \sum X_{i1} - \hat{\beta}_2 \sum X_{i2} = 0$$

$$\sum Y_i = n\hat{\beta}_0 + \hat{\beta}_1 \sum X_{i1} + \hat{\beta}_2 \sum X_{i2} \dots\dots\dots (III)$$

$$\frac{\partial \sum e_i^2}{\partial \hat{\beta}_1} = 2 \sum \left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{i1} - \hat{\beta}_2 X_{i2}\right)(-X_{i1}) = 0$$

$$-2 \sum X_{i1}\left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{i1} - \hat{\beta}_2 X_{i2}\right) = 0$$

بالقسمة على (-2) وفك القوس؛ نجد:

$$\sum X_{i1} Y_i - \hat{\beta}_0 \sum X_{i1} - \hat{\beta}_1 \sum X_{i1}^2 - \hat{\beta}_2 \sum X_{i1} X_{i2} = 0$$

$$\sum X_{i1} Y_i = \hat{\beta}_0 \sum X_{i1} + \hat{\beta}_1 \sum X_{i1}^2 + \hat{\beta}_2 \sum X_{i1} X_{i2} \dots\dots\dots (IV)$$

$$\frac{\partial \sum e_i^2}{\partial \hat{\beta}_2} = 2 \sum \left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{i1} - \hat{\beta}_2 X_{i2}\right)(-X_{i2}) = 0$$

$$-2 \sum X_{i2}\left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{i1} - \hat{\beta}_2 X_{i2}\right) = 0$$

بالقسمة على (-2) وفك القوس؛ نجد:

$$\sum X_{i2} Y_i - \hat{\beta}_0 \sum X_{i2} - \hat{\beta}_1 \sum X_{i1} X_{i2} - \hat{\beta}_2 \sum X_{i2}^2 = 0$$

$$\sum X_{i2}\,Y_i = \hat{\beta}_0 \sum X_{i2} + \hat{\beta}_1 \sum X_{i1}\,X_{i2} + \hat{\beta}_2 \sum X_{i2}^2 \dots\dots\dots (V)$$

تمثل المعادلات (III) و (IV) و (V) المعادلات الطبيعية الثلاثة التي تستخدم في تقدير المعالم الثلاثة المجهولة $(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2)$، ويمكن حل هذه المعادلات بالطرق التالية:

- طريقة المحددات:

ويمكن أن تحل هذه المعادلات بواسطة قاعدة كرايمر للحصول على قيم $\hat{\beta}_k$ من المعلمات وعلى النحو الآتي:

$$\sum Y_i = n\hat{\beta}_0 + \hat{\beta}_1 \sum X_{i1} + \hat{\beta}_2 \sum X_{i2}$$

$$\sum X_{i1} Y_i = \hat{\beta}_0 \sum X_{i1} + \hat{\beta}_1 \sum X_{i1}^2 + \hat{\beta}_2 \sum X_{i1} X_{i2}$$

$$\begin{bmatrix} \sum Y_i \\ \sum X_{i1} Y_i \\ \sum X_{i2} Y_i \end{bmatrix} = \begin{bmatrix} n & \sum X_{i1} & \sum X_{i2} \\ \sum X_{i1} & \sum X_{i1}^2 & \sum X_{i1} X_{i2} \\ \sum X_{i2} & \sum X_{i1} X_{i2} & \sum X_{i2}^2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix}$$

ومن خلال ذلك، نستطيع إيجاد المحددات التالية:

$$|D| = \begin{vmatrix} \sum Y_i & \sum X_{i1} & \sum X_{i2} \\ \sum X_{i1} Y_i & \sum X_{i1}^2 & \sum X_{i1} X_{i2} \\ \sum X_{i2}\,Y_i & \sum X_{i1} X_{i2} & \sum X_{i2}^2 \end{vmatrix}$$

$$|N_1| = \begin{vmatrix} \sum Y_i & \sum Y_i & \sum X_{i2} \\ \sum X_{i1} & \sum X_{i1} Y_i & \sum X_{i1} X_{i2} \\ \sum X_{i2} & \sum X_{i2}\,Y_i & \sum X_{i2}^2 \end{vmatrix}$$

$$|N_2| = \begin{vmatrix} n & \sum X_{i1} & \sum Y_i \\ \sum X_{i1} & \sum X_{i1}^2 & \sum X_{i1} Y_i \\ \sum X_{i2} & \sum X_{i1} X_{i2} & \sum X_{i2}\,Y_i \end{vmatrix}$$

$$\hat{\beta}_1 = \frac{|N_1|}{|D|} = \frac{\begin{vmatrix} \sum Y_i & \sum Y_i & \sum X_{i2} \\ \sum X_{i1} & \sum X_{i1}Y_i & \sum X_{i1}X_{i2} \\ \sum X_{i2} & \sum X_{i2}\,Y_i & \sum X_{i2}^2 \end{vmatrix}}{\begin{vmatrix} \sum Y_i & \sum X_{i1} & \sum X_{i2} \\ \sum X_{i1}Y_i & \sum X_{i1}^2 & \sum X_{i1}X_{i2} \\ \sum X_{i2}\,Y_i & \sum X_{i1}X_{i2} & \sum X_{i2}^2 \end{vmatrix}}$$

$$\hat{\beta}_2 = \frac{|N_2|}{|D|} = \frac{\begin{vmatrix} n & \sum X_{i1} & \sum Y_i \\ \sum X_{i1} & \sum X_{i1}^2 & \sum X_{i1}Y_i \\ \sum X_{i2} & \sum X_{i1}X_{i2} & \sum X_{i2}\,Y_i \end{vmatrix}}{\begin{vmatrix} \sum Y_i & \sum X_{i1} & \sum X_{i2} \\ \sum X_{i1}Y_i & \sum X_{i1}^2 & \sum X_{i1}X_{i2} \\ \sum X_{i2}\,Y_i & \sum X_{i1}X_{i2} & \sum X_{i2}^2 \end{vmatrix}}$$

أما بالنسبة لـ $\hat{\beta}_0$ فيتم الحصول عليه عن طريق:

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1\bar{X}_1 - \hat{\beta}_2\bar{X}_2$$

- **طريقة الانحرافات:**

ويمكن تقدير معاملات الانحدار المتعدد باستخدام الانحرافات أو ما يسمى بالمتوسطات ، أي إنحرافات القيم الاصلية عن وسطها الحسابي كالآتي:

ولهذا نأخذ نموذج يحتوي متغيرين مستقلين X_1 و X_2:

$$\hat{Y}_i = \hat{\beta}_1 X_{i1} + \hat{\beta}_{i2} X_{i2} + e_i$$

وبأخذ المتوسط لهه المعادلة نجد:

$$\bar{Y} = \hat{\beta}_0 + \hat{\beta}_1\bar{X}_1 + \hat{\beta}_2\bar{X} + \bar{e}_i \qquad , \qquad \bar{e}_i = 0$$

$$\hat{Y}_i - \bar{Y} = \hat{\beta}_1(X_{i1} - \bar{X}_1) + \hat{\beta}_2(X_{i2} - \bar{X}) + e_i$$

$$\hat{y}_i = \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + e_i \ldots\ldots\ldots (V)$$

في واقع الأمر فإن المعادلة أعلاه هي واحدة من جملة معادلات يبلغ عددها n معادلة تكون نظام المعادلات التالي:

$$y_1 = \hat{\beta}_1 x_{11} + \hat{\beta}_2 x_{12} + \cdots + \hat{\beta}_k x_{1k} + e_1$$

$$y_2 = \beta_1 x_{21} + \beta_2 x_{22} + \cdots + \hat{\beta}_k x_{2k} + e_2$$

$$\vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots$$

$$y_n = \hat{\beta}_1 x_{n1} + \hat{\beta}_2 x_{n2} + \cdots + \hat{\beta}_k x_{nk} + e_n$$

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \cdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

حيث يمكن التعبير عن ذلك بصيغة المصفوفات كما يلي:

$$y = x\hat{\beta} + e$$

y: متجه عمودي أبعاده $(n \times 1)$ يحتوي على انحرافات قيم المتغير التابع.

x: مصفوفة أبعادها $(n \times k - 1)$ تحتوي على انحرافات قيم المتغيرات المستقلة، حيث أنها لا تتضمن العمود الأول الذي يمثل الحد لاثابت، حيث يمكن بذلك استخراج الحد الثابت $\hat{\beta}_0$ من خارج المصفوفة باستخدام القانون التالي:

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}_1 - \hat{\beta}_2 \bar{X}_2$$

$\hat{\beta}$: متجه عمودي أبعاده $(k - 1 \times 1)$ تحتوي على المعالم المجهولة.

e: متجه عمودي ابعاده $(n \times 1)$ يحتوي على البواقي.

بإعادة كتابة المعادلة (V) على النحو التالي:

$$e_i = y_i - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2}$$

ولما كانت أفضل طريقة للحصول على أصغر قيمة ممكنة للإنحرافات تتم بواسطة تربيعها وبجعل مربعاتها أصغر ما يمكن، وبأخذ المشتقة الجزئية لها بالنسبة لكل من $\hat{\beta}_1$ و $\hat{\beta}_2$ ومساواتها بالصفر؛ نحصل على ما يلي:

$$\sum e_i^2 = \sum \left(y_i - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2}\right)^2$$

$$\frac{\partial \sum e_i^2}{\partial \hat{\beta}_1} = 2 \sum \left(y_i - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2}\right)(-x_{i1}) = 0$$

$$-2 \sum x_{i1}\left(y_i - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2}\right) = 0$$

وبالقسمة على (-2) وفك القوس، نحصل على:

$$\sum x_{i1}\, y_i - \hat{\beta}_1 \sum (x_{i1})^2 - \hat{\beta}_2 \sum x_{i1}\, x_{i2} = 0$$

$$\sum x_{i1}\, y_i = \hat{\beta}_1 \sum (x_{i1})^2 + \hat{\beta}_2 \sum x_{i1}\, x_{i2} \ldots\ldots\ldots (VI)$$

$$\frac{\partial \sum e_i^2}{\partial \hat{\beta}_2} = 2 \sum \left(y_i - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2}\right)(-x_{i2}) = 0$$

$$-2 \sum x_{i2}\left(y_i - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2}\right) = 0$$

وبالقسمة على (-2) وفك القوس، نحصل على:

$$\sum x_{i2}\, y_i - \hat{\beta}_1 \sum x_{i1} x_{i2} - \hat{\beta}_2 \sum (x_{i2})^2 = 0$$

$$\sum x_{i2}\, y_i = \hat{\beta}_1 \sum x_{i1} x_{i2} + \hat{\beta}_2 \sum (x_{i2})^2 \ \ldots\ldots\ldots (VII)$$

ويمكن صياغة المعادلتين السابقين على شكل مصفوفة كما يلي:

$$\begin{bmatrix} \sum x_{i1}\, y_i \\ \sum x_{i2}\, y_i \end{bmatrix} = \begin{bmatrix} \sum (x_{i1})^2 & \sum x_{i1} x_{i2} \\ \sum x_{i1} x_{i2} & \sum (x_{i2})^2 \end{bmatrix} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix}$$

ومن النظام أعلاه، يمكن إعادة كتابته بالشكل التالي:

$$x' y = (x' x)\hat{\beta}$$

وعليه فإن تقدير المعالم باستخدام المصفوفة بأسلوب الانحرافات يأخذ الصيغة التالية:

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

وبعد احتساب المتجه $x' y$ ومحدد المصفوفة $|x' x|$ الذي ينبغي أن لا يساوي صفرا، نوجد مقلوب المصفوفة الذي هو عبارة عن:

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

ومن ثم تطبيق القانون أعلاه، أما $\hat{\beta}_0$ فيمكن حسابه بموجب القانون التالي:

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}_1 - \hat{\beta}_2 \bar{X}_2$$

هذا ويمكن استخراج القيم بالانحرافات دون الرجوع الى البيانات الاصلية كما يلي:

$$\sum x_1 y = \sum X_1 Y - \frac{(\sum X_1)(\sum Y)}{n}$$

$$\sum x_2 y = \sum X_2 Y - \frac{(\sum X_2)(\sum Y)}{n}$$

$$\sum x_1 x_2 = \sum X_1 X_2 - \frac{(\sum X_1)(\sum X_2)}{n}$$

$$\sum x_1^2 = \sum X_1^2 - \frac{(\sum X_1)^2}{n}$$

$$\sum x_2^2 = \sum X_2^2 - \frac{(\sum X_2)^2}{n}$$

$$\sum y_i^2 = \sum Y_i^2 - \frac{(\sum Y_i)^2}{n}$$

طرق تقدير نموذج الانحدار الخطي المتعدد:

لغرض التبسيط نفترض أن نموذج الانحدار الخطي المتعدد يتكون من متغير تابع ومتغيرين مستقلين (كامل علاوي، 2014، الصفحات 103-102). أي:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + U_i \ldots\ldots\ldots\ldots. (4.9)$$

والعلاقة التقديرية للمعادلة (4.9) هي:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + e_i \ldots\ldots\ldots. (4.10)$$

وبتطبيق نظرية كاوس – ماركوف لتقدير معلمات النموذج:

$$e_i = Y_i - \hat{\beta}_0 X_{i1} - \hat{\beta}_2 X_{i2} \ldots\ldots\ldots\ldots (4.11)$$

إذ إن: $e_i = Y - \hat{Y}$

وبتربيع طرفي المعادلة (4.11) وإدخال (Σ) عليها:

$$\sum e_i^2 = \sum \left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{i1} - \hat{\beta}_2 X_{i2}\right)^2 \ldots\ldots (4.12)$$

ولتصغير مجموع مربعات الخطأ إلى أدنى ما يمكن، نأخذ المشتقات الجزئية للمعلمات في المعادلة (4.12) مع ($\sum e_i^2$) ومساواتها مع الصفر.

$$\frac{\partial \sum e_i^2}{\partial \hat{\beta}_0} = 2 \sum \left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{i1} - \hat{\beta}_2 X_{i2}\right)(-1) = 0$$

$$\sum Y_i = n\,\hat{\beta}_0 - \hat{\beta}_1 \sum X_{i1} - \hat{\beta}_2 \sum X_{i2} = 0$$

$$\sum Y_i = n\hat{\beta}_0 + \hat{\beta}_1 \sum X_{i1} + \hat{\beta}_2 \sum X_{i2} \dots \dots \dots (4.13)$$

وبأخذ المشتقة الجزئية للعلاقة بين $\hat{\beta}_1$ و $\sum e_i^2$:

$$\frac{\partial \sum e_i^2}{\partial \hat{\beta}_1} = 2 \sum \left(Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{i1} - \hat{\beta}_2 X_{i2}\right)(-X_{i1}) = 0$$

$$\sum X_{i1} Y_i = \hat{\beta}_0 \sum X_{i1} + \hat{\beta}_1 \sum X_{i1}^2 + \hat{\beta}_2 \sum X_{i1} X_{i2} \dots \dots \dots (4.14)$$

وبنفس الطريقة نشتق العلاقة بين $\hat{\beta}_{i2}$ مع $\sum e_i^2$:

$$\sum X_{i1} Y_i = \hat{\beta}_0 \sum X_{i1} + \hat{\beta}_1 \sum X_{i1} X_{i2} + \hat{\beta}_2 \sum X_{i2}^2 \dots \dots \dots (4.15)$$

وبإعادة كتابة المعادلات (4.13) و (4.14) و (4.15):

$$\left.\begin{aligned} \sum Y_i &= n\hat{\beta}_0 + \hat{\beta}_1 \sum X_{i1} + \hat{\beta}_2 \sum X_{i2} \\ \sum X_{i1} Y_i &= \hat{\beta}_0 \sum X_{i1} + \hat{\beta}_1 \sum X_{i1}^2 + \hat{\beta}_2 \sum X_{i1} X_{i2} \\ \sum X_{i1} Y_i &= \hat{\beta}_0 \sum X_{i2} + \hat{\beta}_1 \sum X_{i1} X_{i2} + \hat{\beta}_2 \sum X_{i2}^2 \end{aligned}\right\} \dots \dots (4.16)$$

وتسمى منظومة المعادلات في (4.16) بالمعادلات الطبيعية، وتحل بعدة طرق مثل الحذف والتعويض وكرايمر والمصفوفات.

ويمكن كتابة المنظومة (4.16) على شكل مصفوفات وكالآتي: $Y = XB$

وباستخدام قانون المصفوفات الذي مر بنا في الانحدار البسيط المعادلة(3.50)[1] :

$$\hat{B} = (X'X)^{-1} X'Y \dots \dots \dots \dots \dots (4.17)$$

فإنّ:

$$(X'X) = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_{i1} & X_{i2} & \cdots & X_{in} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nn} \end{bmatrix} \begin{bmatrix} 1 & X_{11} & \cdots & X_{n1} \\ 1 & X_{12} & \cdots & X_{n2} \\ 1 & X_{11} & \cdots & X_{n3} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n} & \cdots & X_{nn} \end{bmatrix} \dots \dots (4.18)$$

$$= \begin{bmatrix} n & \sum X_{i1} & \sum X_{i2} & \cdots & \sum X_{in} \\ \sum X_{i1} & \sum X_{i1}^2 & \sum X_{i1} X_{i2} & \cdots & \sum X \\ \vdots & \vdots & \vdots & & \vdots \\ \sum X_{in} & \sum X_{i1} X_{in} & \cdots & & \sum X_{in}^2 \end{bmatrix}$$

[1] إن المعادلة (4.17) صحيحة في حالة الانحدار البسيط أو المتعدد لذا تم الاقتصار على ذكرها فقط.-

وأن:

$$X'Y = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ X_{11} & X_{12} & \cdots & X_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{2n} & \cdots & X_{nn} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_n \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} \sum Y \\ \sum X_{i1}Y \\ \sum X_{i2}Y \\ \vdots \\ \sum X_{in}Y \end{bmatrix} \ldots\ldots (4.19)$$

مثال11: استخدم البيانات الآتية لتقدير منظومة المعادلات الآنية الواردة في النموذج (4.16):

Y	X_1	X_2	X_1^2	X_2^2	X_1X_2	X_1Y	X_2Y
7	1	2	1	4	2	7	14
8	2	1	4	1	2	16	8
5	1	3	1	9	3	5	15
6	3	1	9	1	3	18	6
4	1	2	1	4	2	4	8
30	8	9	16	19	12	50	51

وبالتعويض في منظومة المعادلات (4.16):

$$30 = 5\hat{\beta}_0 + 8\hat{\beta}_1 + 9\hat{\beta}_2 \ldots\ldots..(1)$$

$$30 = 5\hat{\beta}_0 + 8\hat{\beta}_1 + 9\hat{\beta}_2 \ldots\ldots\ldots..(1)$$

$$50 = 8\hat{\beta}_0 + 16\hat{\beta}_1 + 12\hat{\beta}_2 \ldots\ldots..(2)$$

$$51 = 9\hat{\beta}_0 + 12\hat{\beta}_1 + 19\hat{\beta}_2 \ldots\ldots..(3)$$

نضرب المعادلة (1) بالعدد الصحيح (2):

$$60 = 10\hat{\beta}_0 + 19\hat{\beta}_1 + 18\hat{\beta}_2 \ldots\ldots..(4)$$

وبطرح المعادلة (2) من المعادلة (4):

$$10 = 2\hat{\beta}_0 + 6\hat{\beta}_2 \ldots\ldots..(5)$$

$$\hat{\beta}_0 = 5 - 3\hat{\beta}_2 \ldots\ldots\ldots..(6)$$

وبضرب المعادلة (2) بالعدد الصحيح (2)، والمعادلة (3) بالعدد الصحيح (4):

$$150 = 24\hat{\beta}_0 + 48\hat{\beta}_1 + 36\hat{\beta}_2 \ldots\ldots..(7)$$

$$204 = 36\beta_0 + 48\hat{\beta}_1 + 76\hat{\beta}_2 \ldots\ldots..(8)$$

وبالطرح:

$$54 = 12\beta_0 + 40\beta_2 \ldots\ldots..(9)$$

$$54 = 12(5 - 3\beta_2) + 40\beta_2 \ldots\ldots..(10)$$

$$-6 = 4\beta_2$$

$$\beta_2 = -1.5$$

وبالتعويض عن قيمة (β_2) في المعادلة (6) نحصل على:

$$\hat{\beta}_0 = 5 - 3(-1.5) = 9.5$$

أما قيمة (β_1):

$$30 = 5(9.5) + 8\hat{\beta}_1 + 9(-1.5) \ldots\ldots\ldots\ldots.(1)$$

$$30 = 47.5 + 8\hat{\beta}_1 - 13.5$$

$$-4 = 8\hat{\beta}_1$$

$$\hat{\beta}_1 = \frac{-4}{8} = -0.5$$

وبذلك تكون المعادلة التقديرية كالآتي:

$$\hat{Y} = 9.5 - 0.5X_1 - 1.5X_2$$

وبالاختصار على متغيرين مستقلين نستخدم المنظومة (4.16) في الحل.

أما تباين المعلمات فنحصل عليه باستخدام القانون الآتي:

$$var(\hat{\beta}_1) = \sigma_u^2 (X'X)^{-1} \ldots\ldots\ldots.(4.20)$$

التقدير باستخدام طريقة المصفوفات (كامل علاوي، 2014، الصفحات 107-106):

$$\sum Y = 30 \qquad \sum X_1 = 8 \qquad \sum X_2 = 9 \qquad \sum X_1X_2 = 12$$

$$\sum X_1^2 = 16 \qquad \sum X_2^2 = 19 \qquad \sum X_1Y = 50 \qquad \sum X_2Y = 51$$

وبتطبيق القانون الآتي:

$$\hat{B} = (X'X)^{-1}X'Y \ldots\ldots\ldots\ldots\ldots\ldots(4.17)$$

$$X'X = \begin{bmatrix} n & \sum X_1 & \sum X_2 \\ \sum X_1 & \sum X_1^2 & \sum X_1X_2 \\ \sum X_2 & \sum X_2X_1 & \sum X_2^2 \end{bmatrix} = \begin{bmatrix} 5 & 8 & 9 \\ 8 & 16 & 12 \\ 9 & 12 & 19 \end{bmatrix}$$

$$X'Y = \begin{bmatrix} \sum Y \\ \sum X_1Y \\ \sum X_2Y \end{bmatrix} = \begin{bmatrix} 30 \\ 50 \\ 51 \end{bmatrix}$$

$$(X'X)^{-1} = \frac{1}{|X'X|} adj X'X$$

$$|X'X| = 5\begin{vmatrix} 16 & 12 \\ 12 & 19 \end{vmatrix} - 8\begin{vmatrix} 8 & 12 \\ 9 & 19 \end{vmatrix} + 9\begin{vmatrix} 8 & 16 \\ 9 & 12 \end{vmatrix} = 16$$

$$(X'X)^{-1} = \frac{1}{16}\begin{bmatrix} 160 & -44 & -48 \\ -44 & 14 & 12 \\ -48 & 12 & 16 \end{bmatrix}$$

$$\hat{\beta}_1 = (X'X)^{-1}X'Y = \frac{1}{16}\begin{bmatrix}160 & -44 & -48\\ -44 & 14 & 12\\ -48 & 12 & 16\end{bmatrix} = \begin{bmatrix}30\\50\\51\end{bmatrix}$$

$$\hat{\beta}_0 = \frac{1}{16}[160 \cdot 30 - 44 \cdot 50 - 48 \cdot 51] = 9.5$$

$$\hat{\beta}_1 = \frac{1}{16}[-44 \cdot 30 + 14 \cdot 50 + 128 \cdot 51] = -1.5$$

$$\hat{\beta}_2 = \frac{1}{16}[-48 \cdot 30 + 12 \cdot 50 + 16 \cdot 51] = -0.5$$

وبذلك تكون المعادلة التقديرية كالآتي:

$$\hat{Y} = 9.5 - 0.5X_1 - 1.5X_2$$

التقدير حول نقطة المتوسط (استخدام الانحرافات):

يمكن استخدام انحرافات القيم عن أوساطها الحسابية في تقدير معالم الانحدار الخطي المتعدد. وبافتراض متغيرين مستقلين (X_1 و X_2) فإن المعادلة التقديرية هي:

$$Y_1 = \hat{\beta}_0 + \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + e_i \dots \dots .(4.21)$$

$$\bar{Y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{X}_1 + \hat{\beta}_2 \bar{X}_2 \dots \dots .(4.22)$$

$$\mathrm{y} = \hat{\beta}_1(X_{i1} + \bar{X}_1) + \hat{\beta}_2(X_{i2} + \bar{X}_2) + e_i$$

$$y = \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2} + e_i \dots \dots \dots \dots .(4.23)$$

$$e_i = y - \hat{\beta}_1 X_{i1} - \hat{\beta}_2 X_{i2} \dots \dots \dots .(4.24)$$

وباستخدام طريقة كاوس – ماركوف وكما مر سابقا:

$$\sum e_i^2 = \sum \left(-\hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2}\right)^2 \dots \dots \dots (4.25)$$

$$\frac{\partial \sum e_i^2}{\partial \hat{\beta}_1} = 2\sum \left(y - \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2}\right)(-X_{i1}) = 0$$

$$\sum x_{i1}y_i = \hat{\beta}_1 \sum x_{i1}^2 - \hat{\beta}_2 \sum x_{i1}x_{i2} \dots \dots \dots \dots (4.26)$$

$$\frac{\partial \sum e_i^2}{\partial \hat{\beta}_2} = 2\sum \left(y - \hat{\beta}_1 X_{i1} + \hat{\beta}_2 X_{i2}\right)(-X_{i2}) = 0$$

$$\sum x_{i2}y_i = \hat{\beta}_1 \sum x_{i1}x_{i2} - \hat{\beta}_2 \sum x_{i2}^2 \dots \dots \dots \dots (4.27)$$

أما الحد الثابت فقيد وفق الصيغة الرياضية الآتية:

$$\hat{\beta}_0 = \bar{Y} + \hat{\beta}_1 \bar{X}_1 + \hat{\beta}_2 \bar{X}_2 \dots \dots \dots \dots \dots .(4.28)$$

والمعادلتان (4.26) و (4.27) هي المعادلات الطبيعية باستخدام البيانات بانحرافاتها عن أوساطها الحسابية وتحل بالطرق مارة الذكر.

$$\left.\begin{aligned}\sum x_{i1}y_i &= \hat{\beta}_1\sum x_{i1}^2 - \hat{\beta}_2\sum x_{i1}x_{i2}\\ \sum x_{i2}y_i &= \hat{\beta}_1\sum x_{i1}x_{i2} - \hat{\beta}_2\sum x_{i2}^2\end{aligned}\right\} \dots\dots\dots(4.29)$$

وبكتابة المنظومة (4.29) على شكل مصفوفات:

$$\begin{bmatrix}\sum x_{i1}y_i\\ \sum x_{i2}y_i\end{bmatrix} = \begin{bmatrix}\sum x_{i1}^2 & \sum x_{i1}x_{i2}\\ \sum x_{i1}x_{i2} & \sum x_{i2}^2\end{bmatrix} = \begin{bmatrix}\hat{\beta}_1\\ \hat{\beta}_2\end{bmatrix} \dots\dots\dots(4.30)$$

وباستخدام قاعدة كرايمر نجد:

$$\hat{\beta}_1 = \frac{|A_1|}{|D|} = \frac{\begin{bmatrix}\sum x_{i1}y_i & \sum x_{i1}x_{i2}\\ \sum x_{i2}y_i & \sum x_{i2}^2\end{bmatrix}}{\begin{bmatrix}\sum x_{i2}^2 & \sum x_{i1}x_{i2}\\ \sum x_{i1}x_{i2} & \sum x_{i2}^2\end{bmatrix}}$$

$$\hat{\beta}_1 = \frac{\sum x_{i1}y_i\sum x_{i2}^2 - \sum x_{i2}y_i\sum x_{i1}x_{i2}}{\sum x_{i1}^2\sum x_{i2}^2 - (\sum x_{i1}x_{i2})^2} \dots\dots\dots(4.31)$$

$$\hat{\beta}_2 = \frac{|A_2|}{|D|} = \frac{\begin{bmatrix}\sum x_{i2}^2 & \sum x_{1i}y_i\\ \sum x_{i1}x_{i2} & \sum x_{i2}y_i\end{bmatrix}}{\begin{bmatrix}\sum x_{i1}^2 & \sum x_{i1}x_{i2}\\ \sum x_{i1}x_{i2} & \sum x_{i2}^2\end{bmatrix}}$$

$$\hat{\beta}_2 = \frac{\sum x_{i1}^2\sum x_{i1}y_i - \sum x_{i2}y_i\sum x_{i1}x_{i2}}{\sum x_{i1}^2\sum x_{i2}^2 - (\sum x_{i1}x_{i2})^2} \dots\dots\dots(4.32)$$

وباستخدام المصفوفات تقدر معالم النموذج وفق القانون الآتي:

$$\begin{bmatrix}\hat{\beta}_1\\ \hat{\beta}_2\end{bmatrix} = (x'x)^{-1}x'y \dots\dots\dots(4.33)$$

إذ أن:

$$(x'x) = \overbrace{\begin{bmatrix}x_{11} & x_{12} & \cdots & x_{1n}\\ x_{21} & x_{22} & \cdots & x_{2n}\end{bmatrix}}^{x'}\overbrace{\begin{bmatrix}x_{11} & x_{21}\\ x_{12} & x_{22}\\ \vdots & \vdots\\ x_{1n} & x_{2n}\end{bmatrix}}^{x} = \begin{bmatrix}\sum x_{i1}^2 & \sum x_{11}x_{12}\\ \sum x_{i1}x_{i2} & \sum x_{i2}^2\end{bmatrix} \dots\dots(4.34)$$

$$(x'y) = \begin{bmatrix}x_{11} & x_{12} & \cdots & x_{1n}\\ x_{21} & x_{22} & \cdots & x_{2n}\end{bmatrix} = \begin{bmatrix}y_1\\ y_2\\ \vdots\\ y_n\end{bmatrix} = \begin{bmatrix}\sum x_{i1}y\\ \sum x_{i2}y\end{bmatrix} \dots\dots(4.35)$$

أما التباين والتباين المشترك فنحصل عليه من القانون الآتي:

$$var\begin{bmatrix}\hat{B}_1\\ \hat{B}_2\end{bmatrix} = \sigma_u^2(x'x)^{-1} \dots\dots\dots(4.36)$$

وتباين الحد الثابت:

$$var\hat{B}_0 = \sigma_u^2\left[\bar{X}_i'(x'x)^{-1}\bar{X}_i + \frac{1}{n}\right] \ldots\ldots\ldots\ldots(4.37)$$

إذ إن:

$$\sigma_u^2 = \frac{\sum_{i=1}^{n} e^2}{N-K}$$

مثال (12): البيانات الآتية تمثل العلاقة بين المتغير المعتمد (Y) والمتغيرات المستقلة (X_1) و (X_2) لعينة من خمسة مشاهدات:

Y	X_1	X_2	y	x_1	x_1^2	x_2	x_2^2	x_1y	x_2y	x_1x_2	y^2
40	4	8	−22	−4	16	−0.4	0.16	88	8.8	1.6	484
60	6	12	−2	−2	4	3.6	12.96	4	−7.2	−7.2	4
50	7	10	−12	−1	1	1.6	2.56	12	−19.2	−1.6	144
70	10	5	8	2	4	−3.4	11.56	16	−27.8	−6.8	64
90	13	7	28	5	25	−1.4	1.96	140	−39.2	−7	784
310	40	42	−−	−−	50	−−	29.7	260	−84	−21	1480

الحل: تقدير النموذج بطريقة المعادلات الطبيعية:

$$\left.\begin{aligned}\sum x_{i1}y_i &= \hat{\beta}_1\sum x_{i1}^2 - \hat{\beta}_2\sum x_{i1}x_{i2}\\ \sum x_{i2}y_i &= \hat{\beta}_1\sum x_{i1}x_{i2} - \hat{\beta}_2\sum x_{i2}^2\end{aligned}\right\} \ldots\ldots\ldots(4.29)$$

$$260 = 50\hat{\beta}_1 - 21\hat{\beta}_2 \ldots\ldots\ldots\ldots\ldots(1)$$

$$-84 = -21\hat{\beta}_1 - 29.7\hat{\beta}_2 \ldots\ldots\ldots(2)$$

نضرب المعادلة (1) بـ (21) والمعادلة (2) بـ (50):

$$5460 = \cancel{1050}\hat{\beta}_1 - 441\hat{\beta}_2 \ldots\ldots\ldots(3)$$

$$-4200 = -\cancel{1050}\hat{\beta}_1 - 1485\hat{\beta}_2 \ldots\ldots(4)$$

بالجمع نحصل على:

$$1260 = 1044\hat{\beta}_2$$

$$\hat{\beta}_2 = 1.2069$$

$$260 = 50\hat{\beta}_1 - 21(1.2069)$$

$$285.3449 = 50\hat{\beta}_1$$

$$\hat{\beta}_1 = 5.707$$

$$\hat{\beta}_0 = \bar{Y} + \hat{\beta}_1\bar{X}_1 + \hat{\beta}_2\bar{X}_2$$

$$\hat{\beta}_0 = 62 - 5.707(8) - 1.2069(8.4) = 6.20604$$

المعادلة التقديرية هي:

$$\hat{Y} = 6.20604 + 5.707X_1 + 1.2064X_2$$

1) التقدير بطريقة كرايمر: نعيد كتابة المعادلات بصيغة المصفوفات:

$$\begin{bmatrix}\sum x_{i1}y_i\\ \sum x_{i2}y_i\end{bmatrix}=\begin{bmatrix}\sum x_{i1}^2 & \sum x_{i1}x_{i2}\\ \sum x_{i1}x_{i2} & \sum x_{i2}^2\end{bmatrix}=\begin{bmatrix}\hat{\beta}_1\\ \hat{\beta}_2\end{bmatrix}\dots\dots\dots..(4.30)$$

$$\begin{bmatrix}260\\ -84\end{bmatrix}=\begin{bmatrix}50 & -21\\ -21 & 29.7\end{bmatrix}\begin{bmatrix}\hat{\beta}_1\\ \hat{\beta}_2\end{bmatrix}\dots\dots\dots\dots.(5)$$

$$\hat{\beta}_1=\frac{|A_1|}{|D|}=\frac{\begin{bmatrix}260 & -21\\ -84 & 29.7\end{bmatrix}}{\begin{bmatrix}50 & -21\\ -21 & 29.7\end{bmatrix}}=\frac{5958}{1044}=5.707$$

$$\hat{\beta}_2=\frac{|A_2|}{|D|}=\frac{\begin{bmatrix}50 & 260\\ -21 & -84\end{bmatrix}}{\begin{bmatrix}50 & -21\\ -21 & 29.7\end{bmatrix}}=\frac{1260}{1044}=1.2069$$

أما الحد الثابت فيتم تقديره كما مر سابقا.

3) التقدير بطريقة معكوس المصفوفة:

$$\hat{B}_1=(x'x)^{-1}x'y$$
$$\hat{B}_1=\frac{1}{|x'x|}adj\ x'x\cdot x'y$$

من المصفوفة (5):

$$\hat{\beta}_1=\frac{1}{1044}\begin{bmatrix}29.7 & 21\\ 21 & 50\end{bmatrix}\begin{bmatrix}260\\ -84\end{bmatrix}$$
$$\hat{\beta}_1=\frac{1}{1044}[29.7\ \cdot 260-84\cdot 21]=5.707$$
$$\hat{\beta}_2=\frac{1}{1044}[21\ \cdot 260-50\cdot 84]=1.2069$$

أما الحد الثابت فيقدر بالطريقة السابقة.

تحليل الانحرافات في حالة الانحدار الخطي المتعدد:

لقد أوضحنا بأن الانحرافات الكلية $\sum_{i=1}^{n}y^2=\sum_{i=1}^{n}(Y-\bar{Y})^2$ تنقسم إلى جزأين، الأول يمثل الانحرافات الموضحة $\sum_{i=1}^{n}\bar{y}^2=\sum_{i=1}^{n}\left(\hat{Y}-\bar{Y}\right)^2$ والثاني الانحرافات غير الموضحة $\sum_{i=1}^{n}\hat{e}_i^2=\sum_{i=1}^{n}(Y-\bar{Y})^2$ أي إن:

$$\sum_{i=1}^{n}y^2=\sum_{i=1}^{n}\hat{y}^2+\sum_{i=1}^{n}\sigma_1^2\dots\dots\dots\dots(4.38)$$
$$SST=SSR+SSE$$

إذ إن:

الانحرافات الكلية: SST

الانحرافات الموضحة: SSR

الانحرافات غير الموضحة: SSE

وبما إن معامل التحديد المتعدد (R^2) يعد مؤشرًا على جودة توفيق خط الانحدار، فهو يحسب بقسمة الانحرافات الموضحة على الانحرافات الكلية.

$$R^2 = \frac{\sum_{i=1}^{n} \hat{y}^2}{\sum_{i=1}^{n} y^2} = \frac{SSR}{SST}$$

وإذا كان لدينا (K) من المتغيرات فإن:

$$R^2 = \frac{\hat{\beta}_1 \sum_{i=1}^{n} x_{i1} y + \hat{\beta}_2 \sum_{i=1}^{n} x_{i2} y + \cdots .. + \hat{\beta}_k \sum_{i=1}^{n} x_{ik} y}{\sum_{i=1}^{n} y_i^2} \ldots \ldots . (4.39)$$

وباستخدام المصفوفات يمكن أن نشتق صيغة معامل التحديد المتعدد (R^2):

$$e'e = \left(y - x\hat{\beta}\right)'\left(y - x\hat{\beta}\right)$$
$$e'e = y'y - y'x\hat{\beta} - \hat{\beta}'yx + \hat{\beta}'x'x\hat{\beta}$$
$$e'e = y'y - \hat{\beta}'x'y + \hat{\beta}'x'x\hat{\beta}$$
$$e'e = y'y - \hat{\beta}'x'y \ldots \ldots \ldots . (4.40)$$

إذ إن:

الانحرافات غير الموضحة: $e'e$

الانحرافات الكلية $y'y$

الانحرافات الموضحة $\hat{\beta}'x'y$

وبإعادة كتابة المعادلة: $y'y = \hat{\beta}'x'y + e'e$

ومنها: $R^2 = \frac{\hat{\beta}' x' y}{y' y} \ldots \ldots \ldots . (4.41)$

ويمكن الحصول على الانحرافات الموضحة وغير الموضحة بدلالة معامل التحديد المتعدد (R^2) وكالآتي:

$$\hat{\beta}'x'y = R^2 \;\; y'y \ldots \ldots \ldots \ldots . (4.42)$$
$$e'e = (1 - R^2)y'y \ldots \ldots \ldots (4.43)$$

أما باستخدام البيانات الأصلية تحول إلى انحرافات للحصول على (R^2) وكالاتي:

$$R^2 = \frac{\hat{\beta}'X'Y - n\bar{Y}^2}{Y'Y - n\bar{Y}^2} \ldots \ldots . (4.44)$$

إن إضافة متغيرات مستقلة إلى نموذج الانحدار الخطي المتعدد يجعل من معامل التحديد المتعدد (R^2) متحيز علويا Upward Based، إذ إن إضافة المتغيرات تؤدي إلى تخفيض درجات الحرية $(N - K)$ مما يتطلب تصحيح المعامل وفق الصيغة الآتية:

$$\bar{R}^2 = 1 - \left[(1-R^2)\frac{n-1}{n-k}\right] \dots\dots\dots\dots (4.45)$$

إذ إن معامل التحديد المصحح $(\bar{R}^2)$.

ومما تجدر الإشارة إليه إن قيمة معامل التحديد تتراوح بين الصفر والواحد: $0 \leq R^2 \leq 1$

في حين يمكن أن يكون معامل التحديد المصحح $(\bar{R}^2)$ سالبًا، إذا كان عدد المتغيرات كبيرًا وحجم العينة (n) صغيراً.

وبالرجوع إلى المثال (12) يمكن الحصول على (R^2):

$$\mathrm{R}^2 = \frac{\hat{\beta}_1 \sum_{i=1}^{n} x_{i1}y + \hat{\beta}_2 \sum_{i=1}^{n} x_{i2}y}{\sum_{i=1}^{n} y_i^2} = \frac{5.707(260) + 1.2069(-84)}{1480}$$

$$= \frac{1382.4404}{1480} = 0.93$$

$$\mathrm{R}^2 = \frac{\hat{\beta}' X' Y}{Y'Y} = \frac{[5.707 \quad 1.2069]\begin{bmatrix}260\\-84\end{bmatrix}}{1480} = \frac{1382.4404}{1480} = 0.93$$

وهذه تفسر أن 93% من التغيرات التي تحدث في المتغير التابع سببها المتغيرات المستقلة الداخلة في النموذج.

$$\bar{R}^2 = 1 - \left[(1-R^2)\frac{n-1}{n-k}\right] = 1 - \left[(1-0.93)\frac{5-1}{5-3}\right] = 0.86$$

اختبار الفروض Hypothesis Test:

يكتسب تحليل عملية اختبار الفروض أهمية في تحليل الانحدار الخطي المتعدد (عبد القادر عطية، 2004، صفحة 313)، وذلك كونها تعطي الباحث مؤشرًا على استبعاد المتغيرات المستقلة التي لا تمارس تأثيرها على المتغير التابع. لذا تنطوي فرضية العدم على عدم وجود العلاقة بين المتغيرات المستقلة والمتغير التابع، أي إن مصدر الانحرافات هو المتغير العشوائي فقط، وعادة ما يستخدم اختبار (F) في ذلك. وتصاغ فرضية العدم كالاتي:

$$H_0 = \hat{\beta}_0 = \hat{\beta}_1 = \cdots \dots \dots . = \hat{\beta}_k = 0 \dots \dots . . (4.46)$$

أما إذا تم رفض الفرضية أعلاه أي هناك علاقة خطية بين المتغيرات المستقلة والمتغير التابع فإن الفرض البديل يصاغ بالشكل الآتي:

$$H_1 = \hat{\beta}_0 \neq \hat{\beta}_1 \neq \cdots \dots \dots \neq \hat{\beta}_k \neq 0 \dots \dots . . (4.47)$$

ويحسب اختبار (F) حسب الصيغة الآتية:

$$F = \frac{\hat{\beta}' x' y / k-1}{e'e / n-k} = \frac{y' y R^2 / k-1}{y' y\,(1-R^2) / n-k}$$

$$F = \frac{\sum \hat{y}2/k-1}{\sum e^2/n-k}$$

$$F = \frac{R^2/k-1}{(1-R^2)/n-k} \dots\dots(4.48)$$

وتكوين قاعدة القرار تحدد كالآتي:

إذا كانت القيمة المحتسبة لاختبار (F) أكبر من القيمة الجدولية بمستوى معنوي معين، وبدرجات حرية (V_1) للبسط و (V_2) للمقام، نرفض فرضية العدم ونقبل الفرض البديل والعكس في حالة كون (F) المحتسبة أقل من الجدولية إذ نقبل فرضية العدم ونرفض الفرض البديل (عبد القادر عطية، 2004، صفحة 318).

ويمكن توضيح الانحرافات في جدول تحليل التباين وكالآتي:

Anova Table

Source of Varriation	Sun of Square	d.f	Mean Error Sun	F
Explained Variation	$\hat{\beta}'x'y = \sum \hat{y}^2 = y'yR^2$	$k-1$	$\frac{\sum \hat{y}^2}{k-1}$	$F = \frac{\sum \hat{y}^2}{k-1}$
Unexplained Variation	$\sum e^2 = y'y(1-R^2) = y'y - \hat{\beta}'x'y$	$N-K$	$\sum e^2/N-K$	
Total Variation	$y'y = \sum y^2$	$n-1$		

وباستخدام البيانات السابقة يمكن بناء جدول تحليل التباين وكالاتي:

Anova Table

Source of Varriation	Sun of Square	d.f	Mean Error Sun	F
Explained Variation	$\sum \hat{y}^2 = 13824404$	$k-1 = 3-1 = 2$	$\frac{\sum \hat{y}^2}{k-1} = \frac{1382.4404}{2} = 691.2202$	$F = \frac{\sum \hat{y}^2/k-1}{\sum e^2/k-1} = \frac{691.2202}{48.7798} = 14.17$
Unexplained Variation	$\sum e^2 = 97.5596$	$N-K = 5-3 = 2$	$\sum e^2/N-K$	

Total Variation	$\sum y^2 = 1480$	$n-1$ $=5-1$ $=4$	$=\frac{97.5596}{2}$ $=48.7798$	

وبمقارنة F المحتسبة مع F الجدولية فإن النموذج المقدر معنوي.

اختبار معنوية المعلمات المقدرة:

يستخدم اختبار (t–Statistic) لمعرفة معنوية المعلمات المقدرة سواءً في نموذج الانحدار البسيط أم المتعدد إذا كان عدد المشاهدات يقل عن (30) مشاهدة. واختبار (Z) إذا كانت أكثر من ذلك. ويستخدم اختبار (t) وفق الآتي بإعادة كتابة المعادلة (4.20).

$$var-con\hat{\beta}_i = \sigma_u^2 (X'X)^{-1} \ldots \ldots (4.46)$$

فإذا كانت البيانات المستخدمة هي البيانات الحقيقية فتمثل عناصر القطر الرئيسي تباين المعلمات المقدرة بما فيها الحد الثابت وعلى التوالي والعناصر خارج القطر الرئيسي تمثل التباين المشترك. وبأخذ الجذر التربيعي لها نحصل على الخطأ المعياري للمعلمات ومنها نحسب اختبار (t) وكالاتي:

$$t_{\hat{\beta}_i} = \frac{\hat{\beta}_i - \beta_i}{S.e_{\hat{\beta}_i}} \ldots \ldots \ldots (4.47)$$

وتكوين قاعدة القرار يتحدد وفق ما مر ذكره في الانحدار البسيط.

أما إذا كانت البيانات المستخدمة هي بالانحرافات فإن أخذ الجذر التربيعي لعناصر المصفوفة في المعادلة (4.36) تعطي الخطأ المعياري للمعلمات عدا الحد الثابت، وعلى معنويتها يحسب اختبار (t) ابتداء من المعلمة $(\hat{\beta}_i)$. أما الاختبار للحد الثابت $(\hat{\beta}_0)$ فالخطأ المعياري له يحسب بأخذ الجذر التربيعي للمعادلة (4.37).

$$Se_{\hat{\beta}_0} = \sqrt{\sigma_u^2 \left[\bar{X}_i' (x'x)^{-1} \bar{X}_i + \frac{1}{n}\right]} \ldots \ldots \ldots \ldots (4.48)$$

$$t_{\hat{\beta}_i} = \frac{\hat{\beta}_0 - \beta_0}{S.e_{\hat{\beta}_0}} \ldots \ldots \ldots (4.49)$$

وبالرجوع إلى المثال السابق وإعادة كتابة $(x'x)^{-1}$:

$$(x'x) = \begin{bmatrix} \frac{29.7}{1044} & \frac{21}{1044} \\ \frac{21}{1044} & \frac{50}{1044} \end{bmatrix}$$

$$\sigma_u^2 = \frac{\sum e^2}{N-K} = 48.7798$$

$$\sigma_u^2 x'x = 48.7798 \begin{bmatrix} \frac{29.7}{1044} & \frac{21}{1044} \\ & \frac{50}{1044} \end{bmatrix} = \begin{bmatrix} 1.3877 & 0.9812 \\ & 2.3362 \end{bmatrix}$$

$$t_{\hat{\beta}_1} = \frac{5.707}{\sqrt{1.3877}} = \frac{5.707}{1.178} = 4.845$$

$$t_{\hat{\beta}_2} = \frac{1.2069}{\sqrt{2.3362}} = \frac{1.2069}{1.5285} = 0.7896$$

$$Se_{\hat{\beta}_0} = \sqrt{48.7798 \left[(8 \quad 8.4) \begin{bmatrix} 50 & -21 \\ -21 & 29.7 \end{bmatrix} \begin{bmatrix} 8 \\ 8.4 \end{bmatrix} + \frac{1}{5} \right]}$$

$$Se_{\hat{\beta}_0} = \sqrt{48.7798 \left[[223.6 \quad 81.48] \begin{bmatrix} 8 \\ 8.4 \end{bmatrix} + \frac{1}{5} \right]}$$

$$Se_{\hat{\beta}_0} = \sqrt{120643.7623 + \frac{1}{5}} = \sqrt{120643.9623}$$

$$Se_{\hat{\beta}_0} = 347.338$$

$$t_{\hat{\beta}_0} = \frac{6.20604}{347.338} = 0.0179$$

وبمقارنة القيم المحتسبة مع القيمة الجدولية البالغة 2.35 بمستوى معنوية 5% ودرجات حرية 3 نجد أن المعلمتين $\hat{\beta}_2 , \hat{\beta}_1$ غير معنويتين.

5.4 التباين والخطأ المعياري لمقدرات OLS:

مصفوفة التباين والتباين المشترك لمقدرات نموذج الانحدار الخطي المتعدد:

في حالة K من المتغيرات المستقلة فإن التباين إلى B_k يمكن الحصول عليه كما يلي:

$$\hat{B} = (X'X)^{-1}X'Y$$

$$\because Y = XB + U$$

وبالتعويض في المعادلة أعلاه نحصل:

$$\hat{B} = (X'X)^{-1}X'(XB + U)$$

ويمكن إعادة كتابة هذه المعادلة كما يلي:

$$\hat{B} = (X'X)^{-1}X'XB + (X'X)^{-1}X'U$$

ولما كانت: $(X'X)^{-1}X'X = 1$

فإن: $\hat{B} = B + (X'X)^{-1}X'U$

وبأخذ التوقع لطرفي المعادلة:

$$\mathrm{E}(\hat{B}) = B + E[(X'X)^{-1}X'U]$$

$$= B + (X'X)^{-1}X'E(U)$$
$$\because E(U) = 0$$
$$\because E(\hat{B}) = 0$$

إذًا $\hat{B}$ هي عبارة عن مقدار غير متحيز إلى B الحقيقية. ويمكن الحصول على تباين قيمة المعلمة المقدرة $\hat{B}$ حيث يؤدي ذلك إلى مصفوفة التباين والتباين المشترك الخاصة بالمتجه $\hat{B}$:

$$\hat{B} - E(\hat{B}) = \hat{B} - B = (X'X)^{-1}X'U$$
$$var(\hat{B}) = E\left\{[\hat{B} - E(\hat{B})][\hat{B} - E(\hat{B})]'\right\}$$
$$\because E(\hat{B}) = B$$
$$var\ (\hat{B}) = E\left\{(\hat{B} - B)(\hat{B} - B)'\right\} = E\{[(X'X)^{-1}X'U][(X'X)^{-1}X'U]'\ \}$$
$$= E\{(X'X)^{-1}X'UU'(X'X)^{-1}\}$$
$$= (X'X)^{-1}X'E(U'U)XU'(X'X)^{-1}$$

ولما كانت:
$$E(UU') = \sigma^2 In$$
$$var(\hat{B}) = (X'X)^{-1}X'.\sigma^2 In.X(X'X)^{-1}$$

وبإعادة الترتيب نحصل:

$$var(\hat{B}) = \sigma^2 In(X'X)^{-1}X'X(X'X)^{-1}$$

مصفوفة الوحدة:
$$\therefore\ (X'X)^{-1}X'X = 1$$
$$var(\hat{B}) = \sigma^2(X'X)^{-1}$$

هذا يعني أن قيمة تباين أي عنصر من عناصر المتجه $(\hat{B})$ هو عبارة عن حاصل ضرب قيمة σ^2 بما يقابلها من العناصر الواقعة على القطر الرئيسي لمقلوب المصفوفة $(X'X)^{-1}$. أما التباين المشترك بين أي اثنين من عناصر المتجه $(\hat{B})$ فهو عبارة عن حاصل ضرب σ^2 بالعنصر المقابل لها والواقع خارج القطر الرئيسي لمقلوب المصفوفة $(X'X)^{-1}$. ويمكن توضيح ذلك لنموذج الانحدار الخطي المتعدِّد الذي يحتوي على متغيرين مستقلين، وكما يلي:

$$var(\hat{B}) = \sigma^2(X'X)^{-1}$$
$$= \frac{\sigma^2}{n\sum X^2 - (\sum X)^2}\begin{bmatrix} \sum X^2 & \sum X \\ -\sum X & n \end{bmatrix}$$
$$= \begin{bmatrix} \dfrac{\sigma^2 \sum X^2}{n\sum X^2 - (\sum X)^2} & \dfrac{-\sigma^2 \sum X}{n\sum X^2 - (\sum X)^2} \\ \dfrac{-\sigma^2 \sum X}{n\sum X^2 - (\sum X)^2} & \dfrac{n\sigma^2}{n\sum X^2 - (\sum X)^2} \end{bmatrix}$$
$$var(\hat{B}_0) = \frac{\sigma^2 \sum X^2}{n\sum X^2 - (\sum X)^2} = \sigma^2\left[\frac{1}{n} + \frac{\bar{X}^2}{\sum X^2}\right]$$

$$var(\hat{B}_1) = \frac{n\sigma^2}{n\sum X^2 - (\sum X)^2} = \frac{\sigma^2}{\sum X^2}$$

$$Cov(\hat{B}_0, \hat{B}_1) = \frac{-\sigma^2 \sum X}{n\sum X^2 - (\sum X)^2} = \frac{-\sigma^2 \bar{X}}{\sum X^2}$$

أما الصيغة التقديرية لتباين الخطأ العشوائي σ^2، ($\mathrm{S}^2\mathrm{e}$) فيمكن الوصول إليها كالآتي:

$$\mathrm{S}^2\mathrm{e} = \frac{e'e}{n-k-1} \ldots\ldots\ldots (10.4)$$

$$e'e = (Y - \hat{Y})'(Y - \hat{Y})$$

$$\because \hat{Y} = X\hat{B}$$

$$e'e = (Y - X\hat{B})'(Y - X\hat{B})$$

$$= Y'Y - Y'X\hat{B} - \hat{B}'X'Y + \hat{B}'X'X\hat{B} \ldots\ldots (11.4)$$

وبما أن الحدين الثاني والثالث يمثلان قيمة واحدة كما وأن كلاً منهما يمثل مبدلة للآخر فإن:

$$Y'X\hat{B} = (Y'X\hat{B})' = Y'X\hat{B}'$$

عليه يمكن إعادة كتابة المعادلة (11.4) بالشكل التالي:

$$e'e = Y'Y - 2\hat{B}'X'Y + \hat{B}'X'X\hat{B} \ldots\ldots (12.4)$$

$$\because \hat{B} = (X'X)^{-1}X'Y$$

$$\therefore (X'X)\hat{B} = X'Y$$

$$e'e = Y'Y - 2\hat{B}'X'Y + \hat{B}'X'Y \ldots\ldots\ldots (13.4)$$

$$e'e = Y'Y - \hat{B}'X'Y \ldots\ldots\ldots\ldots\ldots\ldots\ldots (14.4)$$

وبتعويض هذه القيمة في البسط من المعادلة (10.4) نحصل:

$$\mathrm{S}^2\mathrm{e} = \frac{Y'Y - \hat{B}'X'Y}{n-k-1}$$

6.4 اختبار فرضيات النموذج الخطي المتعدد:

يهدف هذا المبحث إلى توسيع معارفنا الأساسية لنموذج الانحدار وذلك بإجراء اختبار معنوية الانحدار المتعدد والمقدر باستخدام توزيع اختبار إحصاؤه F ومقارنته باختبار t ومن ثم تقييم كفاءة الأداء العام لنموذج الانحدار المتعدد R^2 ومقارنته بمعامل التحديد المقدر المعدل $\overline{\mathrm{R}}^2$ ، وكذلك اختبار العلاقة بين F و R^2 من خلال جدول تحليل التباين، ANOVA ، ثم علاقة R^2 بقيمة المتغير العشوائي، $\sum e_i^2$. (عبد القادر عطية، 2004، صفحة 340)

1.6.4 اختبار معنوية المعالم (t):

يستخدم اختبار (t) لتقييم معنوية تأثير المتغيرات المستقلة $x_1, x_2, \dots x_k$ في المتغير التابع y في نموذج الانحدار المتعدد، وكما ذكرنا سابقاً عند تناول اختبار (t) في نموذج الانحدار الخطي البسيط، أنه يعتمد على نوعين من الفروض:

فرضية العدم: $H_0 \;:\; B_1 = B_2 = B_3 = \cdots .. = B_k = 0$

الفرضية البديلة: $H_1 \;:\; B_1 \neq B_2 \neq B_3 \neq \cdots .. \neq B_k = 0$

وبعد احتساب قيمة (t) تقارن مع قيمتها الجدولية لتحديد قبول أو رفض فرضية العدم ومن ثم تقييم معنوية معلمات النموذج المقدر والصيغة الرياضية لهذا الاختبار يمكن بيانها كما يلي:

أ – بالنسبة الى $\hat{B}_1$:

$$t_{\hat{B}_1} = \frac{\hat{B}_1}{S_{\hat{B}_1}}$$

$$S_{\hat{B}_1} = \sqrt{S^2_{\hat{B}_1}}$$

$$S^2_{\hat{B}_1} = var(\hat{B}_1) = S^2 e\, a_{11}$$

$$var(\hat{B}_1) = S^2 e (x'x)^{-1}$$

$$S^2 e = \frac{e'e}{n-k-1} = \frac{Y'Y - \hat{B}'X'Y}{n-k-1} = \frac{\sum y^2 - \left(\hat{B}_1 \sum x_1 y + \hat{B}_2 \sum x_2 y\right)}{n-k-1}$$

ب– بالنسبة الى $\hat{B}_2$:

$$t_{\hat{B}_2} = \frac{\hat{B}_2}{S_{\hat{B}_2}}$$

$$S_{\hat{B}_2} = \sqrt{S^2_{\hat{B}_2}}$$

$$S^2_{\hat{B}_2} = var(\hat{B}_2) = S^2 e\, a_{22}$$

$$S^2 e = \frac{e'e}{n-k-1}$$

2.6.4 معامل التحديد المتعدد Multiple Coefficient of Determination (R):

يعد مؤشر أساس في تقييم مدى معنوية العلاقة بين المتغير التابع (Y)، والمتغيرات المستقلة (X_k)، إذ $(k = 1,2, \dots, k)$ بعبارة أخرى هو مقياس يوضح نسبة مساهمة المتغيرات المستقلة في تفسير التغير الحاصل في المتغير التابع. ويمكن اشتقاقه باستخدام المصفوفات بالانحرافات كالآتي:

$$\because\; y = x\hat{B} + e$$

$$e = y - x\hat{B}$$

$$e'e = \left(y - x\hat{B}\right)'\left(y - x\hat{B}\right)$$
$$e'e = y'y - y'x\hat{B} - x'\hat{B}'y + \hat{B}'x'x\,\hat{B}$$

وبما أن الحدين الثاني والثالث يمثلان قيمة واحدة كما وأن كلاً منهما يمثل مبدلاً للآخر فإن:

$$\therefore\ e'e = y'y - 2\hat{B}'x'y + \hat{B}'x'x\,\hat{B}$$
$$\because\ \hat{B}(x'x)^{-1}x'y$$
$$(x'x)\hat{B} = x'y$$
$$e'e = y'y - 2\hat{B}'x'y + \hat{B}'x'y$$
$$e'e = y'y - \hat{B}'x'y$$

بذلك يمكن كتابة معادلة الانحرافات الكلية كالآتي:

$$y'y = \hat{B}'x'y - e'e$$

إذ أن:

$y'y$: تمثل الانحرافات الكلية.

$\hat{B}'x'y$: تمثل الانحرافات الموضحة من قبل خط الانحدار.

$e'e$: تمثل الانحرافات غير الموضحة.

وبما أن معامل التحديد R^2 عبارة عن نسبة الانحرافات الموضحة من قبل خط الانحدار إلى الانحرافات الكلية Total variation ، فإنه يمثل نسبة مجموع مربعات التغير في المتغيرات المستقلة إلى مجموع المربعات الكلية:

$$\therefore\ R^2 = \frac{\hat{B}'x'y}{y'y} = \frac{\hat{B}'x'y}{\sum y^2}$$
$$or\ :\quad R^2 = 1 - \frac{e'e}{y'y - n\bar{Y}^2}$$
$$or:\quad R^2 = \frac{\hat{B}_1\sum x_1 y + \hat{B}_2\sum x_2 y}{\sum y^2}$$

إن إضافة متغيرات مستقلة جديدة إلى المعادلة يؤدي الى رفع قيمة R^2 ، وذلك لثبات قيمة المقام وتغير قيمة البسط بمقدار $(\hat{B}xy)$ غير أن الاستمرار بإضافة المتغيرات المستقلة سيؤدي إلى انخفاض درجات الحرية $(n-k-1)$، مما يتطلب استخراج معامل التحديد المعدل أو المصحح R^2 وعلى النحو الآتي:

$$\bar{R}^2 = 1 - \left[(1 - R^2)\frac{n-1}{n-k-1}\right]$$

اختبار إحصائية F F–Statistics:

.... هذا الاختبار معرفة مدى معنوية العلاقة الخطية بين المتغيرات المستقلة $X_1, X_2, \ldots., X_i$ على المتغير التابع Y ، وكما هو الحال في الانحدار البسيط فإنه يعتمد على نوعين من الفروض:

فرضية العدم H_0: وتنص على انعدام العلاقة بين كل متغير من المتغيرات المستقلة $(X_1, X_2, \dots., X_i)$وبين المتغير التابع Y، أي:

$$H_0:\ \hat{B}_1 = \hat{B}_2 = \cdots . = \hat{B}_k = 0$$

الفرضية البديلة H_1: وتنص على وجود علاقة معنوية بين المتغيرات المستقلة والمتغير التابع، أي:

$$H_0:\ \hat{B}_1 \neq \hat{B}_2 \neq \cdots . \neq \hat{B}_k \neq 0$$

والصيغة الرياضية لهذا الاختبار هي:

$$F = \frac{\hat{B}' x' y / k}{e' e / n - k - 1}$$

$$or: \quad F = \frac{R^2 / k}{1 - R^2 / n - k - 1}$$

وبعد احتساب قيمة F نقارن مع قيمتها بالجدولية بدرجة حرية (k) و $(n - k - 1)$ للبسط والمقام ولمستوى معنوية معين. فإذا كانت القيمة المحتسبة أكبر من القيمة الجدولية نرفض H_0 ونقبل H_1، أي أن العلاقة المدروسة معنوية، وهناك على الأقل متغير مستقل واحد من المتغيرات X_k ذو تأثير في Y. أما إذا كنت القيمة المحتسبة أصغر من الجدولية فإن ذلك يعني قبول H_0 أي أن العلاقة الخطية المدروسة غير معنوية، أي أنه ليس ثمة تأثير من أي متغير من المتغيرات المستقلة على المتغير التابع.

4-7. تحليل جدول التباين، ANOVA

لغرض الوقوف على تأثير كل من X_1 ، X_2 في المتغير التابع Y، لابد من عمل جدول تحليل التباين لبيان أثر المتغيرين المستقلين X_1 ، X_2 في النموذج.

جدول تحليل التباين

<table>
<tr><th>مصدر التباين</th><th>مجموع مربعات الخطأ</th><th>درجات الحرية</th><th>متوسط مربعات الخطأ</th><th>اختبار F</th></tr>
<tr><td>الانحراف الموضح من قبل X_1وX_2</td><td>$\hat{B}' x' y$
1201.622859</td><td>K_2</td><td>$\hat{B}' x' y / k$
600.8114295</td><td rowspan="3">$F_1 = \frac{\hat{B}' x'' y / k}{e' e / n - k - 1} = \frac{600.8114295}{12.06285679} = 49.8067$</td></tr>
<tr><td>الانحرافات غير الموضحة</td><td>$e' e$
72.37714076</td><td>$n - k - 1$
6</td><td rowspan="2">$e' e / n - k - 1$
12.06285679</td></tr>
<tr><td>الانحراف الكلي</td><td>$y' y$
1274</td><td>$n - k$
8</td></tr>
</table>

ولمعرفة تأثير كل متغير مستقل في المتغير التابع بصورة منفردة فإننا نختبر خلال المرحلة الأولى تأثير المتغير X_2 بصورة مستقلة في Y. وفي المرحلة الثانية نختبر تأثير المتغير X_2 بصورة مستقلة في Y وبالرجوع إلى مثالنا نختبر ما يلي:

أولاً: تأثير عنصر الدخل X_i في الاستيرادات Y:

لغرض اختبار التأثير المستقل لعنصر الدخل (X_i) في دالة الاستيرادات (Y) يجب معرفة مقدار الزيادة المتحققة في قيمة مجموع مربعات الانحرافات الموضحة من قبل خط انحدار المتغير التابع Y على المتغير المستقل X_2 نتيجة إضافة المتغير X_i الى الدالة، ويتم ذلك بافتراض نموذج يتضمن المتغير X_2 أي أن:

$$Y_i = B_0 + B_2X_2 + U_i$$

$$\hat{Y} = \hat{B}_0 + \hat{B}_2X_2$$

$$\hat{B}_2 = \frac{\sum x_2y}{\sum x_2^2} = \frac{-83}{648} = -0.128086419$$

أما مجموع مربعات الانحرافات الناجمة عن المتغير X_2 فهي:

$$\hat{B}_2 \sum x_2y = (= -0.128086419)(-83) = 10.63117284$$

أما التأثير الذي يضيفه المتغير المستقل X_1 إلى الدالة فهو:

$$\hat{B}_1 \sum x_1y = \hat{B}'X'y - \hat{B}_2 \sum x_2y$$
$$= 1201.622859 - 10.63117284 = 1190.991686$$

ولاختبار مدى معنوية هذا التأثير المضاف من قبل المتغير المستقل X_1 ننظم جدول تحليل التباين الآتي:

تحليل التأثير المستقل للمتغير المستقل X_1 في النموذج

مصدر التباين	مجموع مربعات الخطأ	درجات الحرية	متوسط مربعات الخطأ	اختبـار F
الانحراف الموضح من قبل X_2	10.63117284	1	1190.9911686	$F_1 = \frac{1190.9914686}{12.06285679} = 9873214171$
الانحراف الموضح من قبل X_1	1190.991686	1		
الانحراف الموضح من	1201.622859	2		

قبل X_1 و X_2				
الانحرافات غير الموضحة	72.37714076	6	12.06285679	
الانحرافات الكلية	1274	8		

وبمقارنة قيمة (F_1) المحتسبة والبالغة (98.73) مع مثيلتها الجدولية عند مستوى معنوية (5%) ودرجة حرية (2، 6) للبسط والمقام والبالغة (5.14) والتي يتضح أنها أكبر من الجدولية عليه ترفض فرضية العدم وتقبل الفرضية البديلة، مما يدل على وجود تأثير معنوي للمتغير المستقل X_1 الذي يمثل الدخل القومي على المتغير التابع Y الذي يمثل الاستيرادات.

ثانياً: تأثير عنصر السعر X_2 في الاستيرادات Y:

لبيان أثر المتغير المستقل X_2 في الدالة نفترض نموذجًا يتضمن المتغير المستقل X_1، أي:

$$Y_i = B_0 + B_1 X_1 + U_i$$

$$\hat{Y} = \hat{B}_0 + \hat{B}_1 X_1$$

$$\hat{B}_1 = \frac{\sum x_1 y}{\sum x_1^1} = \frac{881}{650} = 1.355384615$$

أما مجموع مربعات الانحرافات الناجمة عن المتغير X_1 فهي:

$$\hat{B}_1 \sum x_1 y = (1.355384612)(881) = 1194.093846$$

أما التأثير الذي يضيفه المتغير المستقل X_2 إلى الدالة فهو :

$$\hat{B}_2 \sum x_2 y = \hat{B}' x' y - \hat{B}_1 \sum x_1 y$$
$$= 1201.622859 - 1194.093846 = 7.529013$$

ولاختبار مدى معنوية هذا التأثير المضاف من قبل المتغير المستقل X_2 ننظم جدول تحليل التباين الآتي:

تحليل التأثير المستقل للمتغير المستقل X_2 في النموذج

مصدر التباين	مجموع مربعات الخطأ	درجات الحرية	متوسط مربعات الخطأ	اختبار F
الانحراف الموضح من قبل X_1	1194.093846	1	7.529013	$F_1 = \frac{7.529013}{12.06285679} = 0.624148419$

الانحراف الموضح من قبل X_2	7.529013	1		
الانحراف الموضح من قبل X_1 و X_2	1201.622859	2		
الانحرافات غير الموضحة	72.37714076	6	12.06285679	
الانحرافات الكلية	1274	8		

وبمقارنة قيمة (F_2) المحتسبة والبالغة (0.62) مع مثيلتها الجدولية عند مستوى معنوية (5%) ودرجة حرية (2، 6) للبسط والمقام والبالغة (5.14) والتي يتضح بأنها أقل من الجدولية عليه نقبل فرضية العدم، مما يدل بأن المتغير المستقل X_2 لا يمارس تأثيرًا معنويًا على المتغير التابع Y.

وعليه نستنتج بأن المتغير المستقل X_1 الذي يمثل الدخل القومي يمارس تأثيرًا معنويًا على المتغير التابع Y الذي يمثل الاستيرادات. في حين أن المتغير المستقل X_2 الذي يمثل السعر لا يمارس تأثيرًا معنويًا على Y ومن ثم يجب استبعاده من النموذج المدروس واعتماد النموذج الذي يحتوي على المتغير X_1 ذو التأثير المعنوي في المتغير Y والذي ندرجه أدناه:

$$Y_i = B_0 + B_1 X_1 + U_i$$

حيث يتم تقديره وتقييمه على غرار النموذج الوارد في الفصل الخاص بالانحدار الخطي البسيط وكما يلي:

$$\hat{B}_1 = \frac{\sum x_1 y}{\sum x_1^2} = \frac{881}{650} = 1.355384615$$

$$\hat{B}_0 = \bar{Y} - \hat{B}_1 \bar{X}$$

$$= 117 - (1.355384615)(113)$$

$$= 117 - 153.1584615$$

$$= -36.1584645$$

$$\hat{Y} = \hat{B}_0 + \hat{B}_1 X_i$$

$$\hat{Y} = -36.1584615 + 1.355384615\, X_i$$

وتعني بأن زيادة المتغير المستقل X والذي يمثل الدخل القومي بمقدار وحدة واحدة تسبب زيادة في المتغير التابع Y الذي يمثل الاستيرادات بمقدار (1.355) وحدة.

$$R^2 = \frac{\hat{B}_1 \sum x_1 y}{\sum y^2} = \frac{(1.355384615)(881)}{1274} = \frac{1194.093846}{1274} = 0.937279313$$
$$= 93.72\%$$
$$F = \frac{R^2/k}{1 - R^2/n - k - 1} = \frac{0.93727931\,3/1}{1 - 0.93727931\,3/9 - 1 - 1} = \frac{0.937279313}{0.008960098}$$
$$= 104.6059221$$

ولاختبار معنوية المعلمات $\hat{B}_1$ و $\hat{B}_0$ نحتاج البيانات الآتية:

Y	$\hat{Y}$	e_i	e_i^2
100	99.38	0.62	0.3844
106	104.8015385	1.1984615	1.436309967
107	107.5123077	−0.5123077	0.262459179
120	114.2892308	5.7107692	32.61288486
110	114.2892308	−4.2892308	18.39750086
116	119.7107692	−3.7107692	13.7698086
124	126.4876923	−2.4876923	6.188612979
133	131.9092308	1.0907692	1.18977448
137	134.62	2.38	5.6644
$\sum \mathrm{Y} = 1053$	$\sum \hat{Y} = 1053$	$\sum e_i = 0$	79.90615335

بالنسبة لـ $\hat{B}_1$:

$$S^2 e_i = \frac{\sum e_i^2}{n-2} = \frac{79.90615335}{9-2} = 11.41516476$$
$$S_{\hat{B}_1} = \frac{S^2 e_i}{\sum x_i^2} = \frac{11.41516476}{650} = 0.017561791$$
$$S_{\hat{B}_1}^2 = \sqrt{S_{\hat{B}_1}^2} = \sqrt{0.017561791} = 0.132520911$$
$$t_{\hat{B}_1} = \frac{\hat{B}_1}{S_{\hat{B}_1}} = \frac{1.355384615}{0.132520911} = 10.22770372$$

بالنسبة لـ $\hat{B}_0$:

$$S^2 e = 11.41516476$$
$$S_{\hat{B}_0}^2 = S^2 e\left[\frac{1}{n} + \frac{\bar{X}^2}{\sum x_i^2}\right] = 11.41516476\ \left[\frac{1}{9} + \frac{(113)^2}{650}\right]$$
$$= 11.41516476[0.111111111 + 19.64461538] = 225.5148729$$
$$S_{\hat{B}_0} = \sqrt{S_{\hat{B}_0}^2} = \sqrt{225.5148729} = 15.01715262$$

$$t_{\hat{B}_0} = \frac{\hat{B}_0}{S_{\hat{B}_0}} = \frac{-36.15846415}{15.01715262} = -2.407810749$$

وعليه فإن الصيغة التقديرية للنموذج المدروس تكون كما يلي:

$$\therefore \hat{Y} = \hat{B}_0 + \hat{B}_1 X_i$$

$$\hat{Y} = -36.1584615 + 1.355384615 X_i$$

$$(2.407) \qquad (10.227)$$

$$R^2 = \%93.72 \quad F = 104.60$$

4-8. قياس حدود الثقة:

لاحتساب حدود الثقة لأية مشاهدة (نقطة) من مشاهدات خط الانحدار للمجتمع أو بعبارة أخرى لحساب القيمة المتوسطة الحقيقية إلى Y عند مستوى معنوية معين للمتغير المستقل في النموذج. نفترض بأن النقطة المراد تقدير حدود ثقتها هي $F(Y_0)$. ولتقدير المجال الذي يمكن أن تقع فيه قيمة $F(Y_0)$ المقابلة لتشكيلة معينة من قيم المتغيرات المستقلة (k) يجب اشتقاق متباينة القيمة $F(Y_0)$.

$$Y_0 = [1 \quad X_{01} \quad X_{02} \quad \dots X_{0k}]$$

$$\hat{Y}_0 = [1 \quad X_{01} \quad X_{02} \quad \dots X_{0k}] \begin{bmatrix} \hat{B}_0 \\ \hat{B}_1 \\ . \\ . \\ \hat{B}_k \end{bmatrix}$$

$$\hat{Y}_0 = \hat{B}_0 + \hat{B}_1 X_{01} + \cdots + \hat{B}_k X_{0k}$$

وباختصار:

$$\hat{Y}_0 = X_0 \hat{B}$$

ولغرض اشتقاق المتباينة الخاصة بتقدير فترات حدود الثقة للقيمة $E(\text{Yo})$ يجب اشتقاق وسط وتباين القيمة $(\hat{Y}_0)$ وكالآتي:

لإيجاد الوسط فإننا نأخذ القيمة المتوقعة لـ $(\hat{Y}_0)$:

$$E(\hat{Y}_0) = E(X_0 \hat{B})$$

$$E(\hat{Y}_0) = X_0 E(\hat{B})$$

$$\because E(\hat{B}) = B$$

$$\because E(\hat{Y}_0) = X_0 B$$

ولإيجاد التباين:

$$var(\hat{Y}_0) = E\{\hat{Y}_0 - E(\hat{Y}_0)(\hat{Y}_0 - E(\hat{Y}_0)'\}$$

$$= E\left\{(\hat{Y}_0 - X_0 B)(\hat{Y}_0 - X_0 B)'\right\}$$

$$\therefore \hat{Y}_0 = X_0 B$$

$$\therefore var(\hat{Y}_0) = E\left\{(X_0 \hat{B} - X_0 B)(\hat{Y}_0 - X_0 B)'\right\}$$

$$\therefore var(\hat{Y}_0) = E\{(X_0\hat{B} - X_0B)(X_0B' - X_0B)'\}$$
$$= X_0\left\{E(\hat{B} - B)(\hat{B} - B)'\right\}X_0'$$
$$= \sigma^2 X_0 (X'X)^{-1} X_0'$$

وإذا رمزنا للقيمة التقديرية لتباين قيمة حدود الثقة $var(\hat{Y}_0)$ بالرمز $S^2(\hat{Y}_0)$ فإن:

$$S^2(\hat{Y}_0) = S^2 X_0 (X'X)^{-1} X_0'$$

وعليه فإن حدود الثقة للقيمة $E(\text{Y}_0)$ تكون:

$$E(\hat{Y}_0) = \hat{Y}_0 \pm t_{1/2}.S(\hat{Y}_0)$$
$$E(\hat{Y}_0) = X_0\hat{B} \mp t_{1/2}.S(\hat{Y}_0)$$