تحليل الانحدار الخطي المتعدد

Multiple Linear Regression Analysis

المحاضـــــــرة الثالثة

II-4. تفسير معنى معاملات الانحدار

كما رأينا في الانحدار الخطي البسيط فان **المعامل الثابت** (β_0) يمثل القيمة المتوسطة للمتغير التابع عندما تكون قيم المتغيرات المستقلة مساوية للصفر. ولكن هناك ملاحظتان يجب أخذهما في الحسبان عند تفسير المعامل الثابت وهما:

- يجب أن يكون التنبؤ بقيم المتغير التابع بالتعويض في نطاق قيم مشاهدات المتغيرات المستقلة التي أُستُخدمت في تقدير النموذج.
- يصعب تفسير المعامل الثابت اذا كانت قيمته سالبة وقيم المتغير التابع الفعلية موجبة.

في حين **المعامل** (β_1) يمثل التغير في القيمة المتوسطة للمتغير التابع الناتج عن تغير المتغير المستقل (x_1) بوحدة واحدة بافتراض ثبات قيم المتغيرات المستقلة الأخرى $(X_2, X_3, \dots, X_k)$. بعبارة أخرى **المعامل** (β_1) يقيس الأثر المباشر لتغير (x_1) بوحدة واحدة على القيمة المتوسطة للمتغير التابع.

بينما **المعامل** (β_2) يمثل التغير في القيمة المتوسطة للمتغير التابع الناتج عن تغير المتغير المستقل (x_2) بوحدة واحدة بافتراض ثبات قيم المتغيرات المستقلة الأخرى $(X_1, X_3, \dots, X_k)$. وهكذا يستمر التفسير لبقية معاملات الانحدار.

والشيء الذي يجب الإشارة إليه هو أن "حجم وإشارة" معاملات نموذج الانحدار لهما دلالات معينة؛

-فالمعامل الموجب يُشير الى وجود علاقة طردية بين المتغير المستقل والمتغير التابع، والمعامل السالب يُشير الى وجود علاقة عكسية بين المتغير المستقل والمتغير التابع.

-أما حجم المعامل فهو يشير الى مقدار التغير الذي يحدث في المتغير التابع الناتج عن زيادة المتغير المستقل المرتبط بهذا المعامل بوحدة واحدة مع افتراض ثبات قيم جميع المتغيرات المستقلة الأخرى.

II-5. خصائص مقدرات المربعات الصغرى

تتميز مقدرات المربعات الصغرى بعدة خصائص وهي:

1-خاصية عدم التحيز:

ان عدم التحيز يعني أن القيمة المتوقعة لكل عنصر من عناصر المتجه (b) تساوي العنصر المقابل في متجه المعالم الحقيقية (β) أي:

$$E(b) = \beta \Leftrightarrow E\begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ . \\ . \\ b_K \end{pmatrix} = \begin{pmatrix} E(b_0) \\ E(b_1) \\ E(b_2) \\ . \\ . \\ E(b_K) \end{pmatrix} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ . \\ . \\ \beta_K \end{pmatrix} = \beta$$

الإثبات:

نعلم أن معادلة نموذج الانحدار بصيغة المصفوفات تأخذ الشكل التالي:

$$Y = X\beta + \varepsilon$$

ولدينا المعادلة الأساسية الخاصة بمقدرات المربعات الصغرى هي: $b = (X^TX)^{-1}X^TY$

وبتعويض قيمة (Y) في المعادلة الأخيرة نجد:

$$b = (X^TX)^{-1}X^T.(X\beta + \varepsilon)$$

$$b = (X^TX)^{-1}X^TX\beta + (X^TX)^{-1}X^T.\varepsilon$$

$$b = \beta + (X^TX)^{-1}X^T.\varepsilon$$

وبأخذ التوقع الرياضي لطرفي هذه المعادلة نجد:

$$E(b) = E(\beta + (X^TX)^{-1}X^T.\varepsilon)$$

$$E(b) = E(\beta) + (X^T X)^{-1} X^T . E(\varepsilon)$$

وبما أن: $E(\varepsilon) = 0$ ، فيكون:

$$\mathbf{E(b) = \beta}$$

2-خاصية الخطية:

لدينا: $b = (X^T X)^{-1} X^T . \mathbf{Y}$

وبما أن: $[(X^T X)^{-1} X^T]$ هي مصفوفة أرقام ثابتة، فان **(b) دالة خطية لـــ** (Y)، **ومن هنا نستنتج أن مقدرات المربعات الصغرى هي مقدرات خطية.**

3- خاصية الكفاءة أو خاصية أقل تباين:

ان تباين مقدرات المربعات الصغرى أقل من أو يساوي تباين أية مقدرات أخرى خطية وغير متحيزة، ويُشار إليها بالتعبير (BLUE) أي (Best Linear Unbiased Estimators).

ويعرف تباين مقدرات المربعات الصغرى بالصيغة التالية:

$$\mathbf{Var(b) = \sigma^2 . (X^T X)^{-1}} \qquad \mathbf{(11)}$$

يمكن إثبات هذه العلاقة كما يلي:

$$Var(b) = E[b - E(b)]^2$$

$$Var(b) = E[(b - \beta).(b - \beta)^{\mathbf{T}}]$$

$$Var(b) = E[\{(X^T X)^{-1} X^T . Y - \beta\}.\{(X^T X)^{-1} X^T . Y - \beta\}^{\mathbf{T}}]$$

$$Var(b) = E[\{(X^T X)^{-1} X^T . (X\beta + \varepsilon) - \beta\}.\{(X^T X)^{-1} X^T . (X\beta + \varepsilon) - \beta\}^{\mathbf{T}}]$$

$$Var(b) = E[\{(X^T X)^{-1} X^T . \varepsilon\}.\{(X^T X)^{-1} X^T . \varepsilon\}^{\mathbf{T}}]$$

$$Var(b) = E[(X^T X)^{-1} X^T . \varepsilon . (X^T X)^{-1} X . \varepsilon^{\mathbf{T}}]$$

$$Var(b) = E[(X^T X)^{-1} X^T . \varepsilon . \varepsilon^{\mathbf{T}} . X . (X^T X)^{-1}]$$

$$\mathrm{Var}(b) = [(X^TX)^{-1}(X^T.X).E(\varepsilon.\varepsilon^T).(X^TX)^{-1}]$$

$$\mathrm{Var}(b) = (X^TX)^{-1}(X^T.X).\sigma^2.(X^TX)^{-1}$$

$$\mathbf{Var(b) = \sigma^2.(X^TX)^{-1}}$$

وهو المطلوب

وهي مصفوفة مربعة من الترتيب (k+1).(k+1) والعناصر القطرية هي **تباينات مقدرات المربعات الصغرى**، **حيث:**

-العنصر الأول هو: $\mathbf{Var(b_0)}$

-العنصر الثاني هو: $\mathbf{Var(b_1)}$

-العنصر الثالث هو: $\mathbf{Var(b_2)}$

.

.

.

-العنصر الأخير هو: $\mathbf{Var(\ \ _k)}$

تسمى هذه المصفوفة مصفوفة التباين / التغاير : "Variance / Covariance Matrix"

ملاحظة:

بما أن التباين (σ^2) مجهول فإننا نقدره بتباين البواقي (S_e^2) وبذلك تصبح العلاقة رقم (11) على الشكل التالي:

$$\mathbf{S_b^2 = \widehat{Var(b)} = S_e^2.(X^TX)^{-1}} \qquad \mathbf{(12)}$$

ان التباين (S_e^2) هو مقدر غير متحيز للتباين (σ_ε^2) ويتم حسابه وفق العلاقة التالية:

$$\mathbf{S_e^2 = \frac{SSE}{n-(k+1)} = \frac{Y^T.Y - b^T.X^T.Y}{n-(k+1)}} \qquad \mathbf{(13)}$$