

المحور الثاني: تحليل نموذج الانحدار الخطي الكلاسيكي

1- مفهوم تحليل الانحدار:

1.1. تعريف الانحدار: هو أسلوب يستخدم في وصف وتقييم العلاقة بين متغير ما يسمى بالتابع ومتغير أو مجموعة من المتغيرات تسمى بالمستقلة ، أو هو محاولة لتفسير تحركات متغير ما نتيجة تحركات متغير أو متغيرات أخرى، وهناك إنحدار خطي وغير خطي، وبسيط ومتعدد،

الشكل العام للنموذج الرياضي للإنحدار الخطي الذي هو ترجمة للنموذج الاقتصادي النظري يكتب كما يلي :

$$Y_i = B_0 + B_1 X_i$$

وبما أن النظرية هو تبسيط للواقع ولذلك لا تأخذ بعين الاعتبار التغيرات الأخرى لعدة أسباب ، ويمكن حل هذا المشكل بإضافة للنموذج الرياضي السابق متغير الخطأ العشوائي (U_i) ، وتصبح المعادلة السابقة على النحو التالي:
 $Y_i = B_0 + B_1 X_i + U_i$ ويسمى بالنموذج القياسي .

1.2. تعريف الخطأ العشوائي: وهو رمزا شاملا لكل المتغيرات التي لا يمكن إدراجها في النموذج لعدد من الأسباب، غير أنه يفترض أن متوسط تأثير هذه المتغيرات على المتغير التابع ضعيف جدا.

1.3. أسباب إضافة الحد العشوائي: هناك الكثير من الأسباب تستدعي إضافة الخطأ العشوائي نذكر منها:

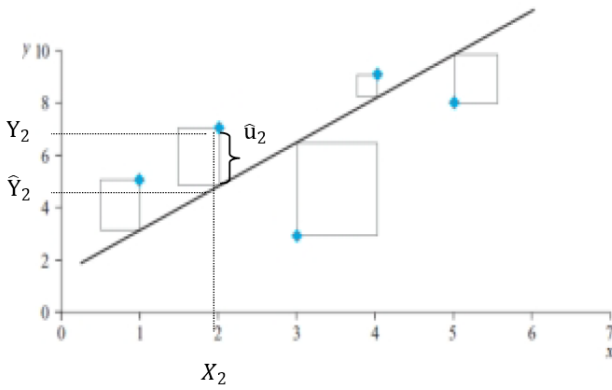
- ① حذف بعض المتغيرات من الدراسة عن غير قصد.
- ② بعض المتغيرات غير قابلة للرصد او للقياس مثل الذوق العادات والتقليد التوقعات ... إلى غير ذلك
- ③ هناك بعض تأثيرات خارجية عشوائية لا يمكن التنبؤ بحدوثها مثل الأوبئة أو الزلازل أو الإضرابات أو هجوم إلكتروني.
- ④ بعض المتغيرات معروفة لكن البيانات عنها غير كافية أو غير دقيقة.
- ⑤ في الكثير من الحالات يكون السلوك البشري عشوائي وغير متوقع اتجاه الظواهر الاقتصادية.
- ⑥ الأخطاء في توصيف النموذج، كالخطأ في تجميع البيانات أو في استخدام النموذج مناسب للعينة.

2- تقدير معالم نموذج الانحدار الخطي البسيط :

تمثل المشكلة الأساسية في الاقتصاد القياسي هو تقدير (إيجاد) قيم معاملات النموذج B_0, B_1 من بيانات واقعية يمكن الحصول عليها إما عن طريق تجربة مسيطر عليها أو مصممة لهذا الغرض أو عن طريق السجلات التاريخية .

نرمز للمقدرات بـ $\hat{B}_0$ ، $\hat{B}_1$ (نضع ^ على المعلمات والمتغيرات والمؤشرات للدلالة على أنها قيم قام النموذج بتقديرها) ،

هناك عدة طرق للتقدير أشهرها طريقة المربعات الصغرى العادية (OLS) The method of ordinary least squares . عند رسم قيم المتغير المستقل X_i مقابل قيم المتغير التابع Y_i تتشكل لنا سحابة نقاط تسمى بسحابة الإنتشار scatter يمكن من خلالها رسم عدد لا نهائي من خطوط الانحدار تمر بنقاط السحابة، لكن الخط الذي يمثل النقاط احسن تمثيل هو الذي يمر على جميع النقاط أو أقرب ما يكون منها ويسمى بالخط المقدّر، والقيمة التي توضح قرب أو بعد القيم المشاهدة للتابع عن القيم المقدرة $\hat{Y}_i$ (التي تقع على الخط المقدّر) تسمى بالباقي ونرمز له بـ $\hat{U}_i = Y_i - \hat{Y}_i$ ويكون مجموع البواقي يساوي الصفر.



ونتلخص طريقة المربعات الصغرى بأن تجعل مجموع مربعات البواقي $\sum \hat{U}_i^2$ أقل ما يمكن.

ونرمز له بـ (ESS) Error sum of squares :

$$\text{ونكتب } \min \sum \hat{U}_i^2 = \min \sum (Y_i - \hat{Y}_i)^2$$

ويمكن كتابة مجموع مربعات البواقي كما يلي $\sum \hat{U}_i^2 = \sum (Y_i - \hat{B}_0 - \hat{B}_1 X_i)^2$.

ورياضيا المقدرات $\hat{B}_0$ ، $\hat{B}_1$ التي تصغر مجموع مربعات البواقي هي تلك التي تعدل المشتقة الجزئية للعلاقة السابقة

$$\min \sum \hat{U}_i^2 = \min \sum (Y_i - \hat{B}_0 - \hat{B}_1 X_i)^2 \Rightarrow \begin{cases} \frac{\partial \sum (Y_i - \hat{B}_0 - \hat{B}_1 X_i)^2}{\partial \hat{B}_0} = 0 \\ \frac{\partial \sum (Y_i - \hat{B}_0 - \hat{B}_1 X_i)^2}{\partial \hat{B}_1} = 0 \end{cases} \text{ ونكتب}$$

$$\begin{cases} -2 \sum (Y_i - \hat{B}_0 - \hat{B}_1 X_i) = 0 \\ -2 \sum X_i (Y_i - \hat{B}_0 - \hat{B}_1 X_i) = 0 \end{cases} \text{ بالإشتقاق نجد}$$

$$\begin{cases} \sum (Y_i - \hat{B}_0 - \hat{B}_1 X_i) = 0 \dots\dots\dots ① \\ \sum X_i (Y_i - \hat{B}_0 - \hat{B}_1 X_i) = 0 \dots\dots\dots ② \end{cases} \text{ بقسمة طرفي المعادلتين على -2 نجد}$$

$$\sum Y_i - n\hat{B}_0 - \hat{B}_1 \sum X_i = 0 \text{ المعادلة 1 نجد}$$

$$n\hat{B}_0 = \sum Y_i - \hat{B}_1 \sum X_i \text{ منه}$$

$$\hat{B}_0 = \frac{\sum Y_i}{n} - \hat{B}_1 \frac{\sum X_i}{n} \text{ وعليه}$$

$$\boxed{\hat{B}_0 = \bar{Y} - \hat{B}_1 \bar{X}} \text{ في الاخير نجد}$$

$$\sum Y_i X_i - \hat{B}_0 \sum X_i - \hat{B}_1 \sum X_i^2 = 0 \text{ المعادلة 2 نجد}$$

$$\sum Y_i X_i - (\bar{Y} - \hat{B}_1 \bar{X}) \sum X_i - \hat{B}_1 \sum X_i^2 = 0 \text{ بتعويض قيمة } \hat{B}_0 \text{ بما يساويها نجد}$$

$$\sum Y_i X_i - \bar{Y} \sum X_i - \hat{B}_1 \bar{X} \sum X_i - \hat{B}_1 \sum X_i^2 = 0 \text{ أي}$$

$$\sum Y_i X_i - \bar{Y} \sum X_i - \hat{B}_1 (\sum X_i^2 - \bar{X} \sum X_i) = 0 \text{ وعليه}$$

$$\hat{B}_1 = \frac{\sum Y_i X_i - n \bar{Y} \bar{X}}{\sum X_i^2 - n \bar{X}^2} \text{ أو حسب القانون العام } \hat{B}_1 = \frac{\sum Y_i X_i - \bar{Y} \sum X_i}{\sum X_i^2 - \bar{X} \sum X_i}$$

فائدة: يتضح من خلال المقدّر $\hat{B}_0$ أن الخط المقدّر يمر حتماً بالنقطة $(\bar{X}, \bar{Y})$

ملاحظة: يمكن تفسير قيمة $\hat{B}_1$ على أنها إذا زادت قيمة x بمقدار وحدة واحدة فيتوقع أن تزيد y بقيمة $\hat{B}_1$ وحدة مع افتراض بقاء العوامل الأخرى على حالها ، أما $\hat{B}_0$ فتفسر على أنها القيمة التي من المتوقع أن يأخذها المتغير التابع y عندما تكون قيمة x معدومة

$$\hat{Y}_i = \hat{B}_0 + \hat{B}_1 \hat{X}_i \text{ تكتب معادلة خط الانحدار بعد التقدير كما يلي}$$

$$Y_i = \hat{B}_0 + \hat{B}_1 \hat{X}_i + \hat{U}_i \text{ النموذج القياسي بعد التقدير}$$

تقسم المعادلة السابقة القيمة المشاهدة لـ Y إلى عنصرين القيمة المتحصل عليها من النموذج بالإضافة للبواقي يعرف النموذج السابق بنموذج الانحدار الخطي الكلاسيكي

3- شروط تطبيق طريقة المربعات الصغرى:

حتى تعطي لنا طريقة المربعات الصغرى مقدرات لها الخصائص المرغوبة يجب تحقق مجموعة من الافتراضات الخاصة بالمتغير المستقل والخطأ العشوائي

A- الافتراضات الخاصة بالخطأ العشوائي: وحتى يتم التعامل مع المتغير العشوائي إحصائياً نضع بعض الافتراضات وهي
① الخطأ العشوائي U_i هو متغير حقيقي عشوائي، بمعنى يأخذ قيم قد تكون موجبة أو سالبة أو صفرية ، لكن هذه القيم غير مؤكدة الحدوث أي غير معلومة للدارس.

② الوسط أو التوقع الرياضي (الأمل الرياضي) للأخطاء معدوم ، بمعنى أن مجموع قيم الحد العشوائي يساوي الصفر، ونكتب:

$$E(U_i) = 0 \Rightarrow \frac{\sum X_i}{n} = 0 \Rightarrow \sum X_i = 0 \quad i = 1, 2, \dots, n.$$

③ ثبات تباين الخطأ العشوائي، وتعني أن تشتت الأخطاء حول المتوسط ثابت ، ونكتب:

$$\sigma_{U_1}^2 = \sigma_{U_2}^2 \dots \dots = \sigma_{U_i}^2.$$

④ عدم وجود ارتباط ذاتي بين الأخطاء العشوائية ، أي الخطأ في الفترة ما مستقل عن الخطأ في الفترات التالية ونكتب:

$$\text{cov}(U_i, U_j) = 0; \quad i \neq j \quad i, j = 1, 2, \dots, n.$$

⑤ الخطأ العشوائي يتبع التوزيع الطبيعي أي أنه قيم المتغير العشوائي قريبة للتوزيع الواقعي، ويمكن أن نمثل الافتراضات

$$U_i \sim N(0, \sigma_u^2) \text{ السابقة بالكتابة التالية:}$$

B- الافتراضات الخاصة بالمتغيرات المستقلة:

① قيم المتغير X_i حقيقية غير عشوائية ، وكذلك غير متساوية جميعها فيما بينها.

- ② عدم وجود ارتباط ذاتي بين الخطأ العشوائي وقيم المتغير المستقل X_i ، أي أن هذا الأخير لا يؤثر ولا يتأثر بالخطأ العشوائي لأن قيمته يتم تحديدها خارج النموذج (عكس المتغير التابع Y_i)
ونكتب: $cov(X_i, U_i) = 0 ; \quad i = 1, 2, \dots, n$
- ③ عدم وجود ارتباط بين المتغيرات المستقلة في حالة الانحدار المتعدد ، ونكتب:
 $cov(X_i, X_j) = 0 ; \quad i \neq j \quad i, j = 1, 2, \dots, n$.

4- خواص مقدر المربعات الصغرى العادية:

- إذا تحققت الافتراضات السابقة سيكون للمقدرات $\hat{B}_0$ و $\hat{B}_1$ المتحصل عليها بهذه الطريقة تمتاز بالخصائص التالية:
1. المقدّر $\hat{B}_0$ و $\hat{B}_1$ هي مقدرات القيم الحقيقية لـ B_0 و B_1
 2. الخطئية تعني أن المقدرات $\hat{B}_0$ و $\hat{B}_1$ لها علاقة خطية مع المتغير التابع.
 3. غير متحيز وتعني في المتوسط تكون القيم الفعلية لـ $\hat{B}_0$ و $\hat{B}_1$ مساوية لقيمتها الحقيقية
 4. الكفاءة ويعني أن المقدرات $\hat{B}_0$ و $\hat{B}_1$ تتميز بأصغر تباين من بين الطرق الأخرى
 5. الاتساق ونعني به أن احتمال اختلاف المقدّر عن القيمة الحقيقية صفراً ، وكلما نتقارب التقديرات من قيمها الحقيقية كلما زاد حجم العينة إلى ما لانهاية

5- الأخطاء المعيارية واختبار المعنوية :

- 3- 1 الدقة والأخطاء المعيارية: تعتبر القيم المقدرة للانحدار هي خاصة بالعينة التي استعملت في التقدير، فإذا قمنا باختيار عينة مختلفة لنفس المجتمع سيؤدي إلى قيم مختلفة للتقديرات، وعليه يجب أخذ فكرة عن مدى جودة ودقة هذه القيم المقدرة $\hat{B}_0$ و $\hat{B}_1$ ومدى إختلافها من عينة إلى أخرى ، ومن أجل ذلك تستعمل مؤشر الخطأ المعياري الذي هو الجذر التربيعي للتباين حيث يعطى بالعلاقات التالية :

$$\hat{\sigma}_{\hat{U}}^2 = \frac{\sum \hat{U}^2}{n-K} \quad / \quad \text{حيث } n \text{ عدد المشاهدات } k \text{ عدد المعلمات المقدرة}$$

$$\hat{\sigma}_{\hat{B}_1}^2 = \hat{\sigma}_{\hat{U}}^2 \left[\frac{1}{\sum (X_i - \bar{X})^2} \right] \quad \text{تباين } \hat{B}_1$$

$$\hat{\sigma}_{\hat{B}_0}^2 = \hat{\sigma}_{\hat{U}}^2 \left[\frac{\sum X_i^2}{n \sum (X_i - \bar{X})^2} \right] \quad \text{تباين } \hat{B}_0$$

- ملاحظة: الأخطاء المعيارية لا تعطي سوى مؤشر عام عن الدقة المحتملة لمعاملات الانحدار ، ولهذا إذا كانت أصغر من نصف المعلمة $(\frac{\hat{B}_1}{2} > \hat{\sigma}_{\hat{B}_1}^2)$ فمن المرجح أن تكون المعلمة دقيقة

6- اختبار المعنوية (اختبار الدلالة): Test of Significance

تناولنا سابقا كيفية تقدير معاملات النموذج $\widehat{B}_1, \widehat{B}_0$ ، لكن المنطق يحتم علينا عدم قبول هذا النموذج إلا بعد إجراء اختبار جودتهما ودرجة الثقة فيهما إحصائيا كأساس جيد للوصول لمعلمة المجتمع الحقيقية B_0, B_1 ، أي أنهما لهما معنوية إحصائية يختلفان عن الصفر

ومن أجل ذلك نستخدم إختبار إحصائية t (t-statistic) لـ studante في حالة العينة أقل من 30 مشاهدة ، ويقوم هذا الاختبار على إتباع الخطوات التالية (نفس الخطوات في حالة اختبار معنوية B_1):

① وضع فرضيتين

ويتم وضع فرضيتين وهما: الفرضية الصفرية (العدم) Null Hypothesis والفرضية البديلة Alternative Hypothesis وتعتبر فرضية العدم الفرضية الإحصائية التي يتم اختبارها في حين الفرضية البديلة النتائج المتبقية

$$\begin{cases} H_0 : B_0 = 0 & \text{وتعني عدم وجود ثابت بالنموذج} \\ H_1 : B_0 \neq 0 & \text{وتعني وجود ثابت بالنموذج} \end{cases}$$

$$\begin{cases} H_0 : B_1 = 0 & \text{وتعني عدم وجود علاقة بين المتغير المستقل والتابع} \\ H_1 : B_1 \neq 0 & \text{وتعني وجود علاقة بين المتغير المستقل والتابع} \end{cases}$$

② حساب قيمة الإختبار ، نستخدم هنا إحصائية t لتسودت (t-statistic) حسب الصيغة التالية: $t_{cal} = \frac{\widehat{B}_0 - B_0}{\widehat{\sigma}_{\widehat{B}_0}}$

③ إستخراج قيمة الاختبار الجدولية t_{tab} (القيم الحرجة - Critical Values) لتحديد منطقة الرفض، من الجدول الخاص بتوزيع ستودنت studante عند درجة الحرية n-k ومستوى المعنوية $\alpha/2$

④ المقارنة بين القيمة المحسوبة والمجدولة ، حيث تتشكل لنا حالتين وكل واحدة لها قرارا ونتيجة.

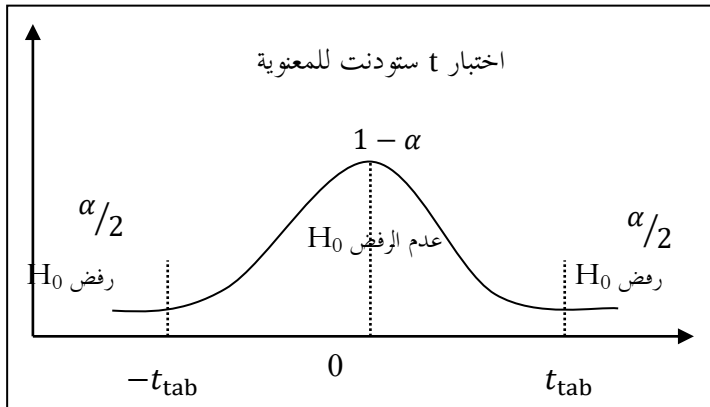
- إذا كان $|t_{cal}| \leq |t_{tab}|$ يكون القرار بعدم رفض

H_0 ورفض H_1 ، وتستخلص أن المعلمة المقدرة ليس لها معنوية إحصائية وإختلافها عن الصفر يرجع لمجرد الدقة ونثق في هذا القرار بـ $(1-\alpha) \%$ ويكون خاطئ بنسبة $\alpha \%$

- إذا كان $|t_{cal}| > |t_{tab}|$ يكون القرار برفض H_0 وعدم رفض H_1 ، وتستخلص أن المعلمة المقدرة لها معنوية إحصائية وإختلافها عن الصفر يعتبر إختلافا

جوهريا ومعنويا لا مجرد الصدفة وإنما يرجع لعوامل واقعية حقيقية ، ونثق في هذا القرار بـ $(1-\alpha) \%$ ويكون خاطئ بنسبة $\alpha \%$

مفهوم مستوى المعنوية (الدلالة) : تأخذ ألفا في العادة القيم التالية (1%, 5%, 10%) ، فعندما نأخذ مستوى المعنوية مثلا ($\alpha = 5\%$) ويكون القرار برفض H_0 فهذا يعني ان هناك احتمال 5% قرار الرفض خاطئا (أو 5% من المساحة



تحت المنحني تمثل منطقة الرفض) ، وعليه فمستوى المعنوية يعبر عن احتمال الخطأ عند اتخاذ القرار الرفض لفرض الصفري H_0 أي أننا لا نثق في قرارنا بـ 5%
ملاحظة: يتم إختيار مستوى المعنوية حسب أهمية الدراسة ، مثل عند دراسة المخاطر الاستثمارية في الاصول فيجب أن لا يكون الخطأ لا يتعدى من 1 % ، كذلك في حالة حجم العينة الكبيرة خاصة في البيانات المالية مما يؤدي إلى إنخفاض الأخطاء المعيارية فيؤدي إلى ارتفاع إحصائية t ، ويكون القرار برفض أية فرضية عدم ، ولحل هذه المسألة يجب استخدام مستوى المعنوية 1 %

7- النموذج الانحدار الخطي المتعدد:

إن وجود النموذج متغير مستقل واحد سيكون غير دقيق في النمذجة المالية فأغلب المتغيرات المالية والاقتصادية تتأثر بمجموعة من العوامل، فمثلا نجد عوائد الأسهم تتأثر بالتضخم الإنتاج الصناعي والمخاطر وغير ذلك من المتغيرات، ومن السهل تعميم النموذج الخطي البسيط إلى نموذج يحتوي على أكثر من متغير انحداري. حيث يكون الشكل العام للنموذج القياسي المتعدد على النحو التالي:

$$Y_i = B_0 + B_1 X_{1i} + B_2 X_{2i} + \dots + B_k X_{ki} + U_i.$$

ويكون الانحدار الخطي المتعدد على النحو التالي:

$$Y_i = \hat{B}_0 + \hat{B}_1 X_{1i} + \hat{B}_2 X_{2i} + \dots + \hat{B}_k X_{ki} + \hat{U}_i$$

حيث تمثل $X_{1i}, X_{2i}, \dots, X_{ki}$ المتغيرات التي تؤثر في المتغير التابع Y_i

وقيم المقدرة $\hat{B}_0, \hat{B}_1, \hat{B}_2, \dots, \hat{B}_k$ تحدد مقدار تأثير المتغيرات المستقلة على Y_i ، وتفسر على أنها القيمة المتوسطة لمتغير التابع نتيجة تغير المتغير المستقل بمقدار وحدة واحدة مع بقاء المتغيرات المستقلة الأخرى ثابتة، أما الحد الثابت $\hat{B}_0$ فيفسر على أنه القيمة المتوقعة لـ Y_i لما جميع قيم المتغيرات المفسرة تكون معدومة

8- تقدير نموذج الانحدار الخطي المتعدد:

ليكن لدينا النموذج القياسي المقدر ذو متغيرين مستقلين فقط للتبسيط: $Y_i = \hat{B}_0 + \hat{B}_1 X_{1i} + \hat{B}_2 X_{2i} + \hat{U}_i$. لتقدير المعلمات سوف نستخدم طريقة المربعات الصغرى OLS التي تعتمد على تدنئه (تصغير) مجموع مربعات البواقي أي: $\sum \hat{U}_i^2$.

ورياضيا المقدرات $\hat{B}_0, \hat{B}_1, \hat{B}_2$ التي تصغر مجموع مربعات البواقي هي تلك التي تعدم المشتقة الجزئية للعلاقة الأخيرة $\Rightarrow \text{Min} \sum \hat{U}_i^2 = \text{Min} \sum (Y_i - \hat{B}_0 - \hat{B}_1 X_{1i} - \hat{B}_2 X_{2i})^2$

$$\begin{cases} \frac{\partial \sum (Y_i - \hat{B}_0 - \hat{B}_1 X_{1i} - \hat{B}_2 X_{2i})^2}{\partial \hat{B}_0} = 0 \\ \frac{\partial \sum (Y_i - \hat{B}_0 - \hat{B}_1 X_{1i} - \hat{B}_2 X_{2i})^2}{\partial \hat{B}_1} = 0 \\ \frac{\partial \sum (Y_i - \hat{B}_0 - \hat{B}_1 X_{1i} - \hat{B}_2 X_{2i})^2}{\partial \hat{B}_2} = 0 \end{cases} \Rightarrow \begin{cases} -2 \sum (Y_i - \hat{B}_0 - \hat{B}_1 X_{1i} - \hat{B}_2 X_{2i}) = 0 \\ -2 X_{1i} \sum (Y_i - \hat{B}_0 - \hat{B}_1 X_{1i} - \hat{B}_2 X_{2i}) = 0 \\ -2 X_{2i} \sum (Y_i - \hat{B}_0 - \hat{B}_1 X_{1i} - \hat{B}_2 X_{2i}) = 0 \end{cases} \Rightarrow$$

بعد حذف 2- من جملة معادلات ثم حذف الاقواس وإدخال المجموع نجد

$$\begin{cases} \sum Y_i - n\hat{B}_0 - \hat{B}_1 \sum X_{1i} - \hat{B}_2 \sum X_{2i} = 0 \\ X_{1i} \sum Y_i - \hat{B}_0 \sum X_{1i} - \hat{B}_1 \sum X_{1i}^2 - \hat{B}_2 \sum X_{1i} X_{2i} = 0 \\ X_{2i} \sum Y_i - \hat{B}_0 \sum X_{2i} - \hat{B}_1 \sum X_{2i} X_{1i} - \hat{B}_2 \sum X_{2i}^2 = 0 \end{cases}$$

$$\begin{cases} \sum Y_i = n\hat{B}_0 + \hat{B}_1 \sum X_{1i} + \hat{B}_2 \sum X_{2i} \\ X_{1i} \sum Y_i = \hat{B}_0 \sum X_{1i} + \hat{B}_1 \sum X_{1i}^2 + \hat{B}_2 \sum X_{1i} X_{2i} \\ X_{2i} \sum Y_i = \hat{B}_0 \sum X_{2i} + \hat{B}_1 \sum X_{2i} X_{1i} + \hat{B}_2 \sum X_{2i}^2 \end{cases} \quad \text{أي}$$

تسمى هذه المعادلات بالمعادلات الطبيعية للمربعات الصغرى، يمكن حلها باستخدام أسلوب التعويض أو المحددات أو أسلوب كرامر

وباستخدام الطريقة الأخيرة يمكن كتابة جملة المعادلات على الشكل المصفوفي كما يلي:

$$\begin{bmatrix} \sum Y_i \\ \sum X_{1i} Y_i \\ \sum X_{2i} Y_i \end{bmatrix} = \begin{bmatrix} n & \sum X_{1i} & \sum X_{2i} \\ \sum X_{1i} & \sum X_{1i}^2 & \sum X_{1i} X_{2i} \\ \sum X_{2i} & \sum X_{2i} X_{1i} & \sum X_{2i}^2 \end{bmatrix} * \begin{bmatrix} \hat{B}_0 \\ \hat{B}_1 \\ \hat{B}_2 \end{bmatrix}$$

أو الشكل المصفوفي المختصر ① $X'Y = X'X\hat{B}$

حيث: $X = \begin{bmatrix} 1 & x_{11} & x_{21} \\ 1 & x_{12} & x_{22} \\ \vdots & \vdots & \vdots \\ 1 & x_{1i} & x_{2i} \end{bmatrix}$; $Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \end{bmatrix}$; $\hat{B} = \begin{bmatrix} \hat{B}_0 \\ \hat{B}_1 \\ \hat{B}_2 \end{bmatrix}$

X' هو منقول المصفوفة حيث يتم تبديل الأسطر بالأعمدة ،
العمود الأول بالمصفوفة X يلزم الثابت فيفترض أن يساوي 1 ، والعمود الثاني يمثل قيم X_{1i} والعمود الثالث يمثل قيم X_{2i} ، اما الشعاع Y يمثل قيم Y_i .

من ① نجد $(X'X)^{-1}X'Y = I\hat{B}$

حيث I تمثل مصفوفة الوحدة ، و $(X'X)^{-1}$ تمثل معكوس (مقلوب) المصفوفة ويحسب بالعلاقة التالية :

$$(X'X)^{-1} = \frac{[con(X'X)]^t}{|X'X|} \quad \text{حيث} \quad con(X'X) = (-1)^{i+j} det(X'X)_{ij}$$

وفي الأخير نجد شعاع (متجه) المقدرات كالتالي $\hat{B} = (X'X)^{-1}X'Y$

مصفوفة التباين والتباين المشترك:

تعطى مصفوفة التباين والتباين المشترك بالعلاقة التالية :

$$\hat{\Omega}_{\hat{B}} = \sigma_{\hat{U}}^2 * (X'X)^{-1} = \begin{pmatrix} \sigma_{\hat{B}_0}^2 & COV(\hat{B}_0, \hat{B}_1) & COV(\hat{B}_0, \hat{B}_2) \\ COV(\hat{B}_1, \hat{B}_0) & \sigma_{\hat{B}_1}^2 & COV(\hat{B}_1, \hat{B}_2) \\ COV(\hat{B}_2, \hat{B}_0) & COV(\hat{B}_2, \hat{B}_1) & \sigma_{\hat{B}_2}^2 \end{pmatrix}$$

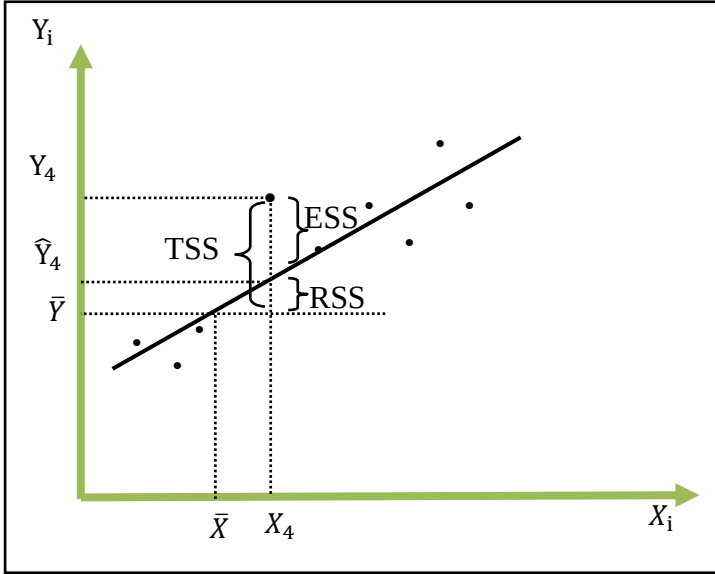
حيث تباين البواقي هو $\sigma_{\hat{U}}^2 = \frac{\sum \hat{U}^2}{n-K}$ ، k عدد المعلمات المقدرة و n حجم العينة ، و $\sigma_{\hat{B}_0}^2, \sigma_{\hat{B}_1}^2, \sigma_{\hat{B}_2}^2$ تمثل تباين المعلمات المقدرة ، وباقي العناصر تمثل التباين المشترك بين المعلمات المقدرة

9- القدرة التفسيرية واختبار المعنوية الكلية للنموذج :

عند رسم خط الانحدار المقدّر $\hat{Y}_i = \hat{B}_0 + \hat{B}_1 X_i$ حتما سوف يمر بالنقطة $(\bar{X}, \bar{Y})$ اما بقية الشائيات (X_i, Y_i) قد يكون لنا أحد الحالات :

① خط الانحدار المقدّر يمر بجميع الثنائيات ، أي أن الانحرافات قيم Y_i عن قيم $\hat{Y}_i$ يساوي الصفر ، بمعنى لا يوجد أخطاء عشوائية في النموذج ، ونقول أننا موفقين في إختيار النموذج أو جودة التوفيق في حدها الأقصى ، لكن هذه الحالة نادرا ما تحدث.

② خط الانحدار المقدّر لا يمر بجميع الثنائيات، أي يوجد انحرافات بين القيم المشاهدات Y_i والقيم المقدرة $\hat{Y}_i$ بسبب وجود أخطاء عشوائية $Y_i - \hat{Y}_i = \hat{u}_i > 0$ في هذه الحالة تنخفض جودة التوفيق، فكلما زادت الانحرافات أي زيادة الأخطاء العشوائية تنخفض القدرة التفسيرية للنموذج .



لنأخذ النقطة (X_4, Y_4) مثلاً نجد انحراف Y_4 عن وسطها هو $Y_4 - \bar{Y}$ وهو الجزء المراد تفسيره ، هذا الجزء مقسم إلى قسمين، جزء مفسر بمعادلة الانحدار المقدرة وهو $\hat{Y}_4 - \bar{Y}$ ، وجزء غير مفسر وهو $Y_4 - \hat{Y}_4$ بسبب وجود الخطأ العشوائي ، إذا عند هذه النقطة لدينا : الانحرافات المراد تفسيرها = الانحرافات المفسرة + الانحرافات غير المفسرة
ونكتب $Y_4 - \bar{Y} = (\hat{Y}_4 - \bar{Y}) + (Y_4 - \hat{Y}_4)$

$$\sum(Y_i - \bar{Y}) = \sum(\hat{Y}_i - \bar{Y}) + \sum(Y_i - \hat{Y}_i) \text{ نجد جميع النقاط نجد}$$

$$\sum(Y_i - \bar{Y})^2 = \sum(\hat{Y}_i - \bar{Y})^2 + \sum(Y_i - \hat{Y}_i)^2 \text{ وبترتيب الجاميع حتى لا تساوي مجموع الانحرافات الصفر نجد}$$

$\sum(\hat{Y}_i - \bar{Y})^2$: هو مجموع مربعات الانحدار (RSS) Regression sum of squares وهو الجزء الذي يفسر تغيرات Y_i بواسطة النموذج المقدّر

$\sum(Y_i - \hat{Y}_i)^2$: هو مجموع مربعات الأخطاء (ESS) Error sum of squares وهو الجزء الذي لم يتم تفسيره من تغيرات Y_i في

$\sum(Y_i - \bar{Y})^2$: هو مجموع المربعات الكلية (TSS) Total sum of squares وهو مجموع التغيرات في Y_i التي نبحث عن تفسيرها لها ، ويمكن أن نكتب $TSS = RSS + ESS$

- معامل التحديد :

وهو النسبة المئوية من التغير الكلي في المتغير التابع Y_i التي يمكن تفسيرها بدلالة المتغير المستقل X_i ،

ويمكن إيجادها بالعلاقة التالية : حيث لدينا $TSS = RSS + ESS$ ومنه $R^2 = \frac{RSS}{TSS}$ أو $R^2 = 1 - \frac{ESS}{TSS}$

$$R^2 = 1 - \frac{\sum(Y_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} \quad \text{أو} \quad R^2 = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2}$$

- في حالة النموذج المقدر يفسر كل التغيرات في Y_i أي لا يوجد أخطاء فإن $ESS=0$ ويكون $R^2 = 1$

- أما إذا كان النموذج لا يفسر شيء من تغيرات Y_i فإن ESS تصل لحدها الأقصى و $TSS=ESS$ ومن ثم $R^2 = 0$

ومنه نستنتج أن $0 \leq R^2 \leq 1$

مثال: $R^2 = 94\%$ فهذا يعني أن 94% من التغيرات في Y_i يفسرها النموذج والباقي أي 8% نفسرها الأخطاء العشوائية وهو ما يدل على ارتفاع الجودة الإحصائية للنموذج المقدر .

- معامل التحديد المعدل :

معامل التحديد هو يشير إلى النسبة التي يمكن تفسيرها من التغيرات في المتغير التابع بدلالة المتغيرات المستقلة المدرجة في النموذج ، وعليه كلما زادنا متغير مستقل حتى لو كان لا يفسر تغيرات التابع سوف يزيد R^2 ، ومن أجل معالجة هذا المشكل نقوم بتعديل قيمة معامل التحديد عن طريق أخذ عدد درجات الحرية $n-k$ في الحسبان فهي تقل عند زيادة المتغيرات

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

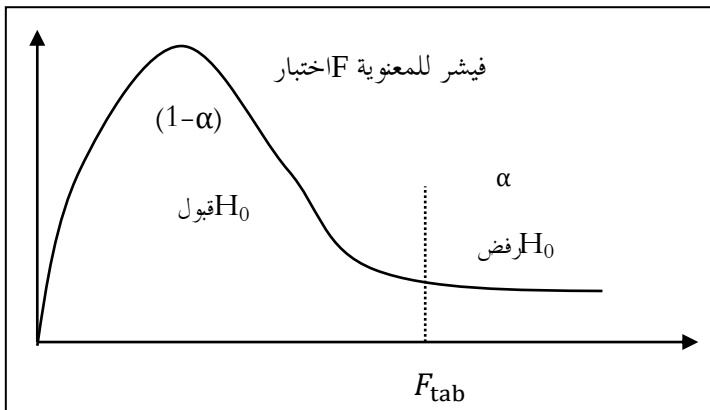
- اختبار المعنوية الكلية للنموذج المقدر:

لاختبار معنوية معادلة الانحدار المقدرة ككل نستخدم اختبار فيشر F ، ويتم إجراء هذا الاختبار إتباع الخطوات التالية:

$$\begin{cases} H_0 : B_0 = 0 \text{ و } B_1 = 0 \text{ النموذج غير معنوي} \\ H_1 : B_0 \neq 0 \text{ أو } B_1 \neq 0 \text{ النموذج معنوي} \end{cases}$$

$$\textcircled{2} \text{ حساب قيمة اختبار فيشر } F_{cal} \text{ المحسوبة بإستخدام الصيغة التالية : } F_{cal} = \frac{R^2/k-1}{(1-R^2)/(n-k)}$$

$\textcircled{3}$ إستخراج قيمة الجدولية F_{tab} من الجدول الخاص بتوزيع فيشر عند مستوى المعنوية $\alpha\%$ ودرجة الحرية $k-1$ و $n-k$



$\textcircled{4}$ المقارنة بين القيمة المحسوبة والجدولية لاختبار فيشر

حيث تكون لنا حالتين وكل واحدة لها قرارا ونتيجة.

- إذا كان $F_{cal} \leq F_{tab}$ يكون القرار بقبول H_0

ورفض H_1 ، وتستنتج أن النموذج المقدر ليس له

معنوية إحصائية ولا تصلح لتفسير سلوك المتغير التابع

وقيمة معامل التحديد R^2 مزيفة غير حقيقية.

- إذا كان $F_{tab} < F_{cal}$ يكون القرار برفض H_0

وقبول H_1 ، وتستنتج أن النموذج المقدر له معنوية إحصائية، وأن قيمة معامل التحديد R^2 التي تحصلنا عليها هي قيمة

موضوعية تصلح لإستخدامها كمقياس لنسبة القيم المفسرة بالنموذج المقدر