الانحدار الخطي المتعدد

Multiple Linear Regression

بعد الانتهاء من قراءة هذا الفصل سيكون لدى القارئ القدرة على بناء نموذج الانحدار الخطي المتعدد وذلك من خلال دراسة الموضوعات التالية:

- عرض نموذج الانحدار الخطي المتعدد.
- الاختبارات الإحصائية وتشمل:
  - الاختبارات المعنوية لمعالم الانحدار الخطي المتعدد.
  - فترات الثقة لمعالم الانحدار الخطي المتعدد.
  - اختبار جودة الملاءمة الكلية لنموذج الانحدار الخطي المتعدد.

مقدمة

عرضنا في الفصل السادس نموذج الانحدار الخطي البسيط وذلك بمعرفة متغير مستقل واحد فقط، وهو يقدم تفسيراً لجزء من التغير في المتغير التابع. ولكن معظم البحوث الاقتصادية والاجتماعية تتطلب دراسة العلاقة بين المتغير التابع وعدد من المتغيرات المستقلة (اثنين أو أكثر). ومن أمثلة ذلك ما تفترضه النظرية الاقتصادية من أن الكمية المطلوبة لسلعة ما دالة في كل من سعر السلعة ودخل المستهلك.

في هذه الحالة يخضع المتغير التابع لتأثير عدد من المتغيرات المستقلة التي تشترك في تفسير ما قد يطرأ عليه من تغيرات، وبذلك قد تنجح هذه المتغيرات المستقلة في تفسير نسبة كبيرة من التغيرات الحادثة في المتغير التابع.

لاشك أن التطوير الكبير في استخدام البرامج الإحصائية والاقتصادية هو الذي مَكَّن الباحثين من اجتياز صعوبة العمليات الحسابية الطويلة والمعقدة أحياناً في تحليل الانحدار المتعدد، ومعالجة النماذج التي تتضمن العديد من المعادلات وكذلك وجود أكثر من متغير مستقل في وقت واحد.

ومن الجدير بالذكر بأن المبادئ الأساسية السابق ذكرها في الفصل السادس الخاص بتحليل الانحدار الخطي البسيط تنطبق في حالة الانحدار المتعدد مع وجود اختلاف وحيد وهو أننا نتعامل في حالة الانحدار المتعدد مع أكثر من متغير مستقل.

2.7 نموذج الانحدار الخطي المتعدد

الصورة العامة لنموذج الانحدار الخطي المتعدد الذي يحتوى على K من المتغيرات المستقلة $X_1, X_2, \ldots X_n$ ومتغير تابع واحد هو Y هي كما يلي:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_k X_{ki} + \varepsilon_i \text{ , } for\ i = 1,2,..n \qquad (7.1)$$

حيث β_0: الجزء المقطوع الخاص بنموذج الانحدار الخطي المتعدد.

$\beta_{1,,}\beta_{2,}\ldots\beta_k$: تمثل معاملات الانحدار الجزئية التي تقيس استجابة المتغير التابع للمتغير المستقل X_i $(i = 1,2,\ldots,k)$، مع بقاء أثر بقية المتغيرات المستقلة الأخرى ثابتة. فمثلاً معامل الانحدار الجزئي $\hat{\beta}_1$ في تحليل الانحدار المتعدد يقيس معدل التغير في المتغير التابع نتيجة للتغير في المتغير المستقل X_1 بوحدة قياس واحدة مع بقاء أثر بقية المتغيرات المستقلة ثابتة. أما المتغير ε_i فيمثل تأثير العوامل الأخرى التي أهملتها المعادلة (7.1).

باستخدام بيانات من عينة مسحوبة حجمها (n) من المجتمع المراد دراسته، وذلك خلال فترة زمنية معينة، فإنه يمكن كتابة معادلة نموذج الانحدار الخطي المتعدد المُقدَّرة على النحو التالي:

$$Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \cdots + \hat{\beta}_k X_{ki} + e_i \qquad (7.2)$$

حيث $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \ldots, \beta_k$ هي التقديرات المحسوبة من العينة لمعالم المجتمع، وتسمى معاملات الانحدار الجزئية المقدرة، بينما e_i تمثل تقدير للخطأ ε_i.

معادلة الانحدار الخطي المتعدد المُقدَّرة من العينة يمكن كتابتها على النحو التالي:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \cdots + \hat{\beta}_k X_{ki} \qquad (7.3)$$

حيث $\hat{Y}_i$ تمثل تقديراً لمتوسط أو (القيمة المتوقعة) للمتغير التابع عند قيم ثابتة للمتغيرات المستقلة $X_1, X_2, \ldots, X_k$ وعلى ذلك فإن المعادلة (7.2) تصبح $Y_i = \hat{Y}_i + e_i$، أي أن:

$$e_i = Y_i - \hat{Y}_i \qquad (7.4)$$

أي أن باقي الانحدار عند المشاهدة i يساوي الفرق بين القيمة الفعلية والقيمة المقدرة عند نفس المشاهدة.

3.7 الاختبارات الإحصائية

1.3.7 الاختبارات المعنوية لمعالم الانحدار الخطي المتعدد

بفرض أنه لدينا نموذج الانحدار الخطي المتعدد في معادلة (7.1) لاختبار الفرضية الصفرية $H_0 : \beta_i = \beta_{H_0}$، $i = 0,1,\ldots,k$، مقابل الفرضية البديلة:

- $H_1 : \beta_i \neq \beta_{H_0}$
- $H_1 : \beta_i > \beta_{H_0}$
- $H_1 : \beta_i < \beta_{H_0}$

فإننا نستعمل إحصاء الاختبار:

$$t_i = \frac{\hat{\beta}_i - \beta_i}{SE(\hat{\beta}_i)} \ , \ i = 0,1,\ldots,k \qquad (7.5)$$

حيث k يمثل عدد المتغيرات المستقلة

إحصاء الاختبار في (7.5) يخضع لتوزيع T بدرجات حرية $(n - k - 1)$.

حالة خاصة:

إذا كانت $\beta_{H_0} = 0$ ، فإن إحصاء الاختبار يصبح على النحو التالي:

$$t_i = \frac{\hat{\beta}_i}{SE(\hat{\beta}_i)} \ , \ i = 0,1,\ldots,k \qquad (7.6)$$

وكذلك إحصاء الاختبار في (7.6) يخضع لتوزيع T بدرجات حرية $(n-k-1)$.

2.3.7 فترات الثقة لمعالم الانحدار الخطي المتعدد

فترة الثقة $100(1-\alpha)\%$ **للمعلمة** β_i **هي:**

$$\hat{\beta}_i \pm t\left(1-\frac{\alpha}{2};n-k-1\right)SE(\hat{\beta}_K) \qquad (7.7)$$

مع ملاحظة أن $t\left(1-\frac{\alpha}{2};n-k-1\right)$ يمكن حسابها من خلال جداول خاصة بتوزيع T.

4.7 اختبار جودة الملاءمة الكلية لنموذج الانحدار الخطي المتعدد

1.4.7 معامل التحديد Coefficient of Determination

سبق لنا تعريف معامل التحديد بأنه نسبة بين مجموع مربعات الانحدار ومجموع المربعات الكلي، بالتالي من الواضح أنه كلما زاد عدد المتغيرات المستقلة في النموذج زادت قيمة معامل التحديد، بصرف النظر أن المتغير المستقل أو المتغيرات المستقلة التي يتم إضافتها مؤثرة في المتغير التابع من عدمه.

بينما قيمة معامل التحديد المعدل $\bar{R}^2$ تزداد بدرجة أقل من زيادة معامل التحديد عند إضافة متغير أو متغيرات مستقلة، وعادة قيمة $\bar{R}^2$ تزداد عند إضافة متغيرات مستقلة مؤثرة في المتغير التابع. ومن الممكن أن قيمة معامل التحديد المعدل $\bar{R}^2$ تأخذ قيمة سالبة

رغم أن معامل التحديد R^2 تكون قيمته موجبة دائماً غير سالبة. وفي الحالة التي تكون فيها قيمة معامل التحديد المعدل سالبة، فإن قيمته تساوي صفراً في هذه الحالة.

عند مقارنة نموذجين لهما نفس المتغير التابع، ولكنهما يختلفان في عدد المتغيرات المستقلة في كل منهما، فإن المقارنة بينهما تتم بناء على قيمة معامل التحديد المعدل وليس معامل التحديد غير المعدل.

لا يمكن المقارنة بين معامل التحديد الذي نحصل عليه من النموذج الخطي بنظيره الذي نحصل عليه من النموذج اللوغارتيمي مثلاً، وذلك لان المتغير التابع في النموذج الأول يعتمد على القيم الأصلية بينما في النموذج الثاني فيعتمد على لوغاريتمات تلك القيم.

2.4.7 اختبار جودة المعنوية الكلية

يستخدم اختبار F (نسبة للعالم Fisher) لاختبار المعنوية الكلية لنموذج الانحدار الخطي البسيط، نستخدم اختبار F لاختبار معنوية العلاقة الخطية للانحدار المتعدد Y على المتغيرات المستقلة $X_1,\ldots,X_k$، وذلك عند مستوى دلالة α ويمكننا صياغة الفرضية الصفرية على النحو التالي:

$$H_0 : \beta_1=\beta_2=\cdots=\beta_K=0$$

أي أن الانحدار المتعدد غير معنوي

مقابل الفرضية البديلة:H_1 على الأقل أحد المعاملات β_i لا تساوي الصفر، أي أن الانحدار الخطي المتعدد معنوياً.

- إذا كانت (Prob. Value) أقل من مستوى الدلالة α، فإننا نرفض الصفرية ونستنتج أن الانحدار الخطي المتعدد معنوياً.
- بينما إذا كانت (Prob. Value) أكبر من مستوى الدلالة α، فإننا لا نرفض الصفرية ونستنتج أن الانحدار المتعدد غير معنوي.

البيانات في الجدول (1.7) يعطي الكميات المطلوبة من سلعة معينة (Y) وســـعرها (X_1) بالدينار ودخل المستهلك (X_2) بالدينار خلال الفترة الزمنية 1990-2012. اسم الملــف (Example7.1).

جدول (1.7): الكمية المطلوبة من سلعة معينة وسعرها ودخل المستهلك

السنة	الكمية	السعر	الدخل	السنة	الكمية	السعر	الدخل
1990	20	120	500	2002	68	65	1170
1991	24	116	530	2003	72	57	1210
1992	28	114	570	2004	76	56	1320
1993	32	109	620	2005	80	53	1290
1994	36	104	670	2006	84	50	1400
1995	40	100	730	2007	88	47	1450
1996	44	92	810	2008	82	40	1380
1997	48	88	840	2009	76	55	1310
1998	52	84	960	2010	70	61	1280
1999	56	79	1020	2011	65	65	1190
2000	60	70	1100	2012	60	72	1090
2001	64	68	1180				

المطلوب: مستخدماً مستوى معنوية 5% أجب عما يلي:

1. التحقق من توافر الفرض الخاص بتحقق العلاقة الخطية بين الكمية المطلوبة من سلعة معينة وسعرها ودخل المستهلك.
2. تقدير معادلة انحدار الكمية المطلوبة من سلعة معينة على كل من سعرها ودخل المستهلك.
3. فسِّر معاملات معادلة الانحدار.
4. اختبار معنوية معالم معادلة الانحدار β_0 , β_1 , β_2
5. تقدير فترة الثقة لمعالم معادلة الانحدار المتعدد.
6. إيجاد قيمة معامل التحديد R^2 ومعامل التحديد المعدل $\overline{R}^2$ مع تفسير النتيجة.
7. اختبار معنوية العلاقة الخطية للانحدار المتعدد باستخدام اختبار F.
8. قدِّر كمية الطلب من السلعة إذا علمت أن دخل الفرد 1100 دينار وسعر السلعة 70 دينار ثم احسب الباقي (الخطأ) مع العلم أن القيمة الفعلية لكمية الطلب تساوي 60 في هذه الحالة.

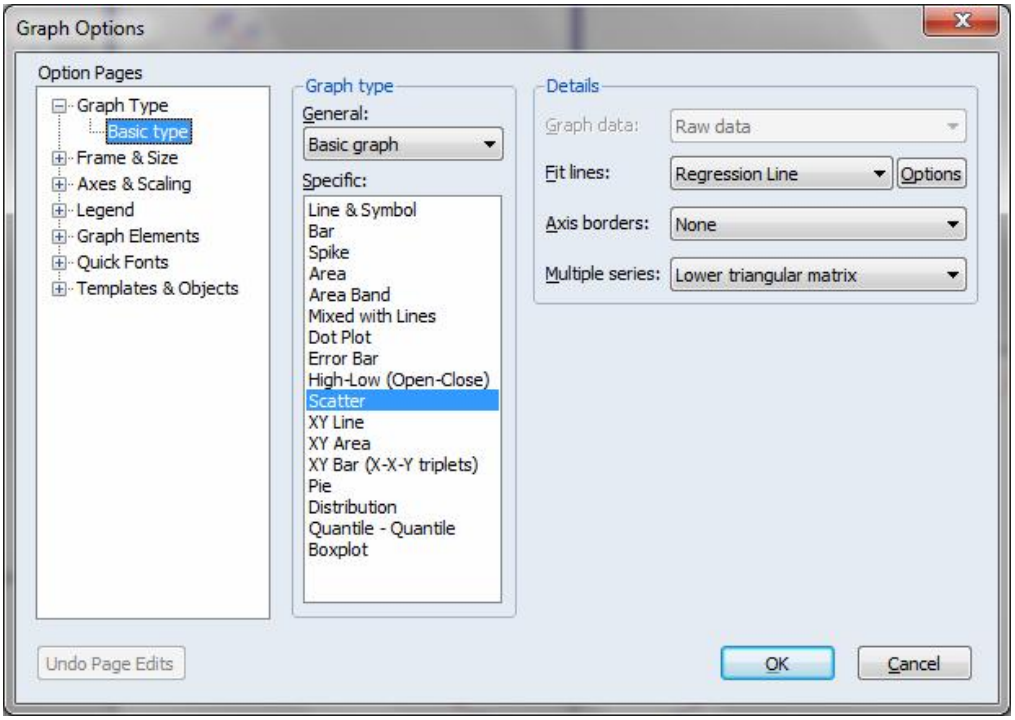
الحل:

1. **التحقق من وجود العلاقة الخطية.**

يتم ذلك من خلال رسم شكل الانتشار بين المتغير التابع (الكمية المطلوبة) والمتغيرين المستقلين (سعر السلعة، دخل المستهلك) كما يلي.

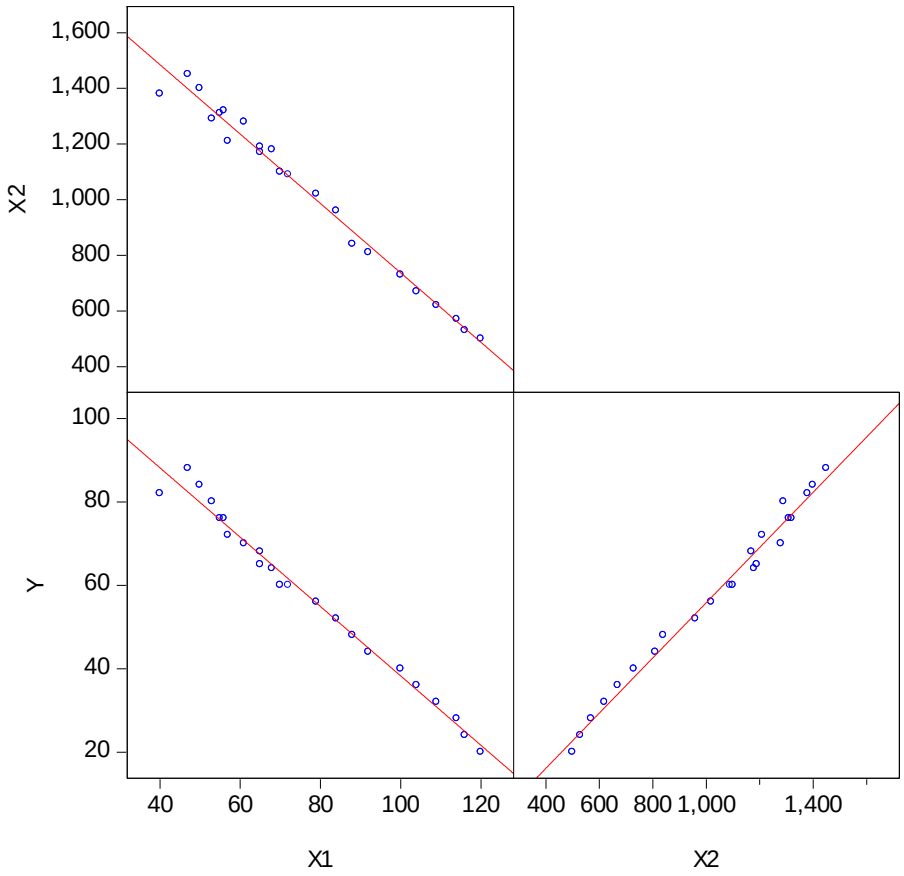
- اختر المتغيرات Y، X1، X2 ثم اضغط على مفتاح الإدخال.
- اختر View ثم Graph.
- اختر Scatter أسفل قائمة Graph Type.

- اختر Regression Line مقابل :Options.
- اختر Lower triangular matrix مقابل :Multiple series كمـــا فـــي شـــكل (1.7).



شكل (1.7): رسم شكل الانتشار لأكثر من متغيرين

- اضغط OK، فنحصل على شكل الانتشار الموضحة في شكل في شكل (2.7).



شكل (2.7): شكل الانتشار لنموذج انحدار الكمية المطلوبة من سلعة معينة

من خلال شكل الانتشار بين Y ، X_1، X_2 الموضحة في شكل (2.7)، نجد أن معظم النقاط تقع على الخط المستقيم المشار إليه في الشكل، وهذا يدل على وجود علاقة خطية بين Y وكلٍ من X_1 ، X_2 على حده. ويتبين من الرسم البياني كذلك أن العلاقة بين الكمية

المطلوبة (Y) وسعر السلعة (X_1) علاقة عكسية، بينما العلاقة بين الكمية المطلوبة (Y) ودخل المستهلك هي علاقة طردية (X_2).

2. تقدير معادلة انحدار الكمية المطلوبة من سلعة معينة على كل من سعرها ودخل المستهلك.

لإيجاد معادلة الانحدار الخطي نتبع الخطوات التالية في برنامج E Views

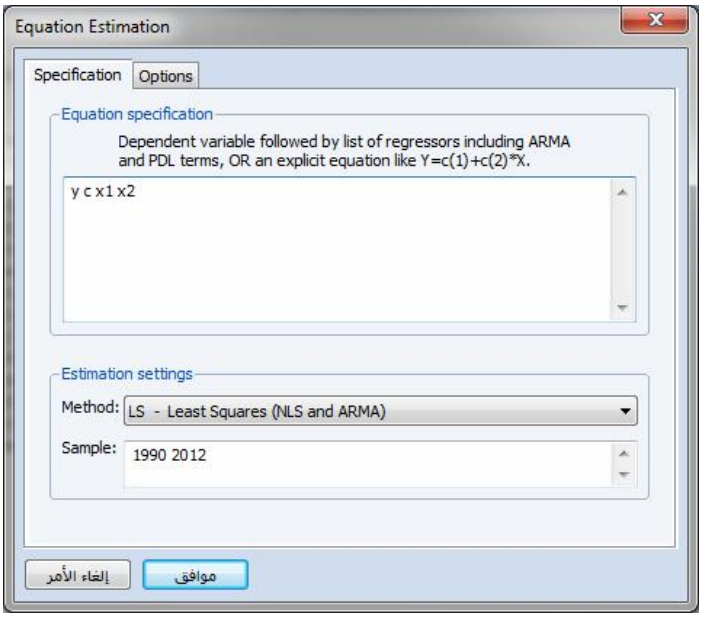
- من شريط القوائم اختر

Quick ► Estimate Equation

- ندخل معادلة خط الانحدار الخطي المتعدد كما يلي:

Y C X1 X2

كما هو موضح في شكل (3.7):



شكل (3.7): المربع الحواري الخاص بنموذج انحدار الكمية المطلوبة من سلعة معينة

حيث: Y هو المتغير التابع ثم يتبعه C والذي يمثل الجزء الثابت (المقطوع من محور Y) ثم المتغيرين المستقل X1، X2.

- يجب ملاحظة أن الترتيب ضروري في هذه الحالة حيث يجب أن نبدأ بكتابة المتغير التابع ثم ***c*** للدلالة على الجزء الثابت ثم المتغيرين المستقلين.
- اضغط موافق سنحصل على النتائج الموضحة في جدول (2.7).

جدول (2.7): نتائج نموذج انحدار الكمية المطلوبة من سلعة معينة

```
Dependent Variable: Y
Method: Least Squares
Date: 04/07/13   Time: 00:00
Sample: 1990 2012
Included observations: 23
```

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	52.02873	24.03827	2.164412	0.0427
X1	-0.395365	0.151986	-2.601328	0.0171
X2	0.034977	0.012092	2.892660	0.0090

R-squared	0.989807	Mean dependent var	57.60870
Adjusted R-squared	0.988787	S.D. dependent var	20.37543
S.E. of regression	2.157538	Akaike info criterion	4.496921
Sum squared resid	93.09945	Schwarz criterion	4.645029
Log likelihood	-48.71459	Hannan-Quinn criter	4.534169
F-statistic	971.0454	Durbin-Watson stat	1.089594
Prob(F-statistic)	0.000000		

من خلال النتائج الموضحة في جدول (2.7)، فإن معادلة انحدار الكمية المطلوبة من سلعة معينة على كل من سعرها ودخل المستهلك هي:

$$\hat{Y}_i = 52.029 - 0.395X_{1i} + 0.035X_{2i}$$

$$\hat{SE}(\beta) = (24.038) \quad (0.152) \quad (0.012)$$

$$T = (2.164) \quad (-2.601) \quad (2.893)$$

3. تفسير معاملات معادلة الانحدار.

- معامل الانحدار $\hat{\beta}_0$ في المعادلة السابقة يساوي 52.029، وهو يمثل الكمية المطلوبة من السلعة عندما يكون كلاً من سعر السلعة والدخل مساوياً صفراً. وهذا غير مقبول من الناحية العملية؛ لذا تفسير $\hat{\beta}_0$ ليس منطقياً في هذه الحالة.
- معامل الانحدار $\hat{\beta}_1$ في المعادلة السابقة يساوي 0.395- وهو يعبر عن التغير في الكمية المطلوبة من السلعة الناتج عن تغير سعرها بدينار واحد مع ثبات الدخل، وحيث أن إشارة معامل الانحدار $\hat{\beta}_1$ سالبة فهذا يعني أنه إذا زاد سعر السلعة بمقدار دينار واحد فإن الكمية المطلوبة تقل بمقدار 0.395 وحدة. أي أن زيادة في سعر السلعة بمقدار 100 دينار فإن الكمية المطلوبة سوف تقل بمقدار 39.5 وحدة (أي تقريباً 40 وحدة).
- أما معامل الانحدار $\hat{\beta}_2$ في المعادلة السابقة يساوي 0.035 وهو يعبر عن التغير في الكمية المطلوبة من السلعة نتيجة التغير في الدخل بدينار واحد مع ثبات السعر، وحيث أن إشارة معامل الانحدار $\hat{\beta}_2$ موجبة فهذا يعني أنه إذا زاد الدخل بمقدار دينار واحد فإن الكمية المطلوبة تزداد بمقدار 0.035، أي أن زيادة في الدخل بمقدار 100 دينار فإن الكمية المطلوبة سوف تزداد بمقدار 3.5 وحدة.

4. اختبار معنوية معالم الانحدار المتعدد.

باستخدام القيمة الاحتمالية (.Prob) نجد أن جميعها أقل من $\alpha = 0.5$، وبالتالي نستنتج أن $\beta_0 \neq 0$, $\beta_1 \neq 0$, $\beta_2 \neq 0$، أي أن المعالم جميعها معنوية من الناحية الإحصائية.

5. تقدير فترة الثقة لمعالم معادلة الانحدار المتعدد.

- من نافذة نتائج الانحدار اختر التالي:

View ► Coefficient Diagnostic ►Confidence Intervals

كما تم شرحه في الفصل السادس – انظر شكل (8.6).

- أكتب 0.95 في أسفل ***Confidence Intervals*** ثم اضغط ***OK*** فنحصل على النتائج الموضحة في جدول (3.7).

جدول (3.7): نتائج فترة الثقة لنموذج انحدار الكمية المطلوبة من سلعة معينة

```
Coefficient Confidence Intervals
Date: 04/07/13   Time: 00:03
Sample: 1990 2012
Included observations: 23
```

Variable	Coefficient	95% CI Low	High
C	52.02873	1.885773	102.1717
X1	-0.395365	-0.712401	-0.078328
X2	0.034977	0.009754	0.060200

من خلال النتائج الموضحة في جدول (3.7) لتقدير فترات الثقة للمعالم نجد أن:

- تقدير فترة الثقة للمعلمة β_0 هي $1.886 \le \beta_0 \le 102.172$.
- وهي تعني أننا واثقون بنسبة 95% أن فترة الثقة تحتوى على قيمة المعلمة β_0 المجهولة، حيث أن فترة الثقة السابقة لا تشمل على الصفر (البداية والنهاية موجبتان) وبالتالي يمكن القول بأن $\beta_0 \neq 0$، أي أن β_0 معنوية إحصائياً وهي نفس النتيجة التي حصلنا عليها في المطلوب (4).
- تقدير فترة الثقة للمعلمة β_1 هي $-0.712 \le \beta_1 \le -0.078$.

- وهي تعني أننا واثقون بنسبة 95% أن فترة الثقة تحتوى على قيمة المعلمة β_1 المجهولة، حيث أن فترة الثقة السابقة لا تشمل على الصفر (البداية والنهاية سالبتان) وبالتالي يمكن القول بأن $\beta_1 \neq 0$ ، أي أن β_1 معنوية إحصائياً وهي نفس النتيجة التي حصلنا عليها في المطلوب (4).
- تقدير فترة الثقة للمعلمة β_2 هي $0.010 \leq \beta_2 \leq 0.060$.
- وهي تعني أننا واثقون بنسبة 95% أن فترة الثقة تحتوى على قيمة المعلمة β_2 المجهولة، حيث أن فترة الثقة السابقة لا تشمل على الصفر (البداية والنهاية موجبتان) وبالتالي يمكن القول بأن $\beta_2 \neq 0$ ، أي أن β_2 معنوية إحصائياً وهي نفس النتيجة التي حصلنا عليها في المطلوب (4).

6. معامل التحديد R^2 ومعامل التحديد المعدل $\overline{R}^2$.

من خلال النتائج الموضحة في جدول (2.7) نجد أن معامل التحديد $R^2 = 0.9898$ ومعامل التحديد المعدل $\overline{R}^2 = 0.9888$، وهذا يعني أن 98.88% من التغيرات الكلية في الكمية المطلوبة من السلعة يمكن تفسيرها من خلال المتغيرين المستقلين سعر السلعة والدخل، أما النسبة المتبقية 1.12% فقد ترجع إلى عوامل مستقلة أخرى قد تؤثر في الكمية المطلوبة من تلك السلعة.

7. اختبار معنوية العلاقة الخطية للانحدار المتعدد باستخدام اختبار F.

في هذه الحالة نقوم باختبار الفرضية الصفرية $H_0: \beta_1 = \beta_2 = 0$ مقابل الفرضية البديلة والتي تنص على أنه توجد واحدة على الأقل من معاملات الانحدار لا تساوي الصفر، أي أن $H_1: \beta_i \neq 0$.

من خلال النتائج الموضحة في جدول (2.7) نجـد أن $F = 971.045$ ، $Prob. = 0.0000$ ، لذلك نرفض الفرضية الصفرية ونستنتج أنه توجد واحدة على الأقل من معاملات الانحدار لا تساوي الصفر. بالتالي يمكن القول بأن العلاقة الخطيـة للانحـدار المتعـدد معنويـة إحصائيا وهذا الاختبار يكافئ $H_0 : R^2 = 0$ مقابل $H_1 : R^2 \neq 0$ وهذا يدل علـى أن R^2 تختلف معنوياً عن الصفر.

8. التنبؤ والباقي (الخطأ)

$$\hat{Y}(X_1=70, X_2=1100) = 52.029 + 0.395\times 70 + 0.035\times 1100$$
$$= 62.879$$

الباقي (الخطأ) = القيمة الفعلية – القيمة المقدرة

-2.879 = 62.879 -60 =

ملاحظة حسب برنامج ***EViews***:

القيمة المقدرة تساوي 62.828 وبالتالي فإن قيمة الباقي تساوي 2.828-، مع العلم بأن تقريب الإجابة هو سبب وجود الاختلاف البسيط.