المحور الرابع: الارتباط الجزئي والازدواج الخطي وطرق اختيار المتغيرات التفسيرية

معاملات الارتباط الجزئي:

يقيس معامل الارتباط الجزئي صافي الارتباط بين المتغير التابع ومتغير مستقل بعد حذف التأثير المشترك أي مع تثبيت المتغيرات المستقلة الأخرى في النموذج. فمثلاً $\mathrm{r}_{YX_1.X_2}$ هو الارتباط الجزئي بين Y و X_1 بعد حذف تأثير X_2 من كل من Y و X_1:

المسألة 7–14 (أ):

$$\mathrm{r}_{YX_1.X_2} = \frac{\mathrm{r}_{YX_1} - \mathrm{r}_{YX_2} r_{X_1X_2}}{\sqrt{1 - r_{X_1X_2}^2}\sqrt{1 - r_{YX_2}^2}} \qquad (14-7)$$

$$\mathrm{r}_{YX_2.X_1} = \frac{\mathrm{r}_{YX_2} - \mathrm{r}_{YX_1} r_{X_1X_2}}{\sqrt{1 - r_{X_1X_2}^2}\sqrt{1 - r_{YX_2}^2}} \qquad (15-7)$$

حيث r_{YX_1} معامل الارتباط البسيط بين Y و X_1، ويعرف r_{YX_2} و $r_{X_1X_2}$ على نفس الخط. وتتراوح معاملات الارتباط الجزئية بين 1−،1+ (كما هو الحال بالنسبة لمعاملات الارتباط البسيط)، ويكون لها نفس إشارة معلمة المجتمع المناظرة، وتستخدم لتحديد الأهمية النسبية للمتغيرات المفسرة المختلفة في الانحدار المتعدد (سلفاتور ، 1982 صفحة 169).

مثال5: بالتعويض بقيم جدول 7–1، 7–2 في معادلة (6–18) لمعامل الارتباط البسيط، نحصل على:

$$r_{YX_1} = \frac{\sum x_1 y}{\sqrt{\sum x_1^2}\sqrt{\sum y^2}} = \frac{956}{\sqrt{576}\sqrt{1,634}} \approx 0.9854$$

$$r_{YX_2} = \frac{\sum x_2 y}{\sqrt{\sum x_2^2}\sqrt{\sum y^2}} = \frac{900}{\sqrt{504}\sqrt{1,634}} \approx 0.9917$$

$$r_{X_1X_2} = \frac{\sum x_2 x_1}{\sqrt{\sum x_2^2}\sqrt{\sum x_1^2}} = \frac{524}{\sqrt{504}\sqrt{576}} \approx 0.9725$$

وعليه:

$$r_{YX_1..X_2} = \frac{r_{YX_1} - r_{YX_2} r_{X_1X_2}}{\sqrt{1 - r_{X_1X_2}^2}\sqrt{1 - r_{YX_2}^2}} = \frac{0.9854 - (0.9917)(0.9725)}{\sqrt{1 - 0.9725^2}\sqrt{1 - 0.9917^2}} \cong 0.7023 or 70.23\%$$

و:

$$r_{YX_2.X_1} = \frac{r_{YX_2} - r_{YX_1} r_{X_1X_2}}{\sqrt{1 - r_{X_1X_2}^2}\sqrt{1 - r_{YX_1}^2}} = \frac{0.9917 - (0.9854)(0.9725)}{\sqrt{1 - 0.9725^2}\sqrt{1 - 0.9854^2}} \cong$$

$$\cong 0.8434 \; or \; 84.34\%$$

وعليه، فإن X_2 أكثر أهمية من X_1 في تفسير التغير في Y.

<u>مثال 6:</u> يمكن تلخيص النتائج الكلية لمثال الحنطة – السماد – المبيد كالآتي:

$$\hat{Y} = 31.98 + 0.65X_1 + 1.10X_2$$

قيم t ـــــ (2.70) (4.11)

$$R^2 = 0.992 \qquad \bar{R}^2 = 0.989 \qquad F_{2,7} = 413.17$$

$$r_{YX_1.X_2} = 0.70 \qquad r_{YX_2X_1} = 0.84$$

وبالرغم من الحصــول على النتائج عادة باســتخدام الكمبيوتر، إلا أنه من المهم القيام بحل المســألة (يدويًا)، كما فعلنا لكي نفهم خطوات الحل بوضـــوح. وتعرض المســـألة 7–22 عينة برنامج – كمبيوتر كامل يشـــرح بالكامل كيفية استخدام Statistical Package for the Social Sciences) SPSS وهو أكثر برامج الكمبيوتر شيوعًا في الاستخدام)، لانحدار متعدِّد في ثلاث متغيرات.

<u>مسائل محلولة:</u>

النموذج الخطي ذو المتغيـرات الثلاثة:

7–1: (أ) اكتب معادلة نموذج الانحدار الخطي المتعدِّد لحالة متغيرين مستقلين أو مفسرين وحالة k متغير مستقل أو مفسر. (ب) اذكر فروض النموذج الخطي للانحدار المتعدِّد.

(أ) في حالة متغيرين مستقلين أو مفسرين، المعادلة هي:

$$Y_i = b_0 + b_1X_{1i} + b_2X_{2i} + u_i \qquad (1-7)$$

وفي حالة k متغير مستقل أو مفسر، المعادلة هي:

$$Y_i = b_0 + b_1X_{1i} + b_2X_{2i} + \cdots + b_kX_{ki} + u_i$$

حيث تمثل X_{2i}، على سبيل المثال المشاهدة التي ترتيبها i للمتغير المستقل X_2

(ب) الفروض الخمســـة الأول لنموذج الانحدار الخطي المتعدِّد هي نفس فروض نموذج الانحدار البســـيط OLS (انظر المســـألة 6–1). أي أن الفروض الثلاثة الأول يمكن تلخيصـــها على النحو: $u_i \sim N(O, \sigma)$. الفرض الرابع هو $E_{(u_iu_j)} = 0$ عند $i \neq j$ ؛ والفرض الخامس هو $E(X_iu_j) = 0$. الفرض الإضـــافي الوحيد المطلوب لنموذج الانحدار الخطي المتعدِّد OLS هو أنه لا توجد علاقة خطية تامة بين المتغيرات المستقلة $X_1, X_2, \ldots$ لأنه لو كان بين اثنين أو أكثر من المتغيرات المستقلة ارتباط خطي تام، لاستحالة حساب تقديرات معالم OLS لأن مجموعة المعادلات الطبيعية سوف تشتمل على معادلتين أو أكثر ليست مستقلة. أما إذا كان هناك ارتباط خطي كبير وليس تامًا بين التين أو أكثر من المتغيرات المفســـرة، فإنه يمكن تقدير معالم OLS، ولكن لا يمكن عزل تأثير كل من المتغيرات المستقلة ذات الارتباط الخطي الكبير فيما بينها (أنظر قسم 9–1)

7–2: باستخدام طريقة OLS في حالة متغيرين مستقلين أو مفسرين، اشتق (أ) المعادلة الطبيعية (7–2)، (ب) المعادلة الطبيعية (7–3)، (ج) المعادلة الطبيعية (7–4). (القارئ غير الملم بالتفاضل يمكنه أن يتخطي هذه المسألة).

(أ) تشتق المعادلة الطبيعية (7–2) بإيجاد النهاية الصغرى للمقدار $\sum e_i^2$ بالنسبة إلى $\hat{b}_0$:

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

$$-2\sum\left(Y_i - \hat{b}_0 - \hat{b}_1 X_{1i} - \hat{b}_2 X_{2i}\right) = 0 \qquad (2-7)$$

$$\sum Y_i = n\hat{b}_0 \pm \hat{b}_1 \sum X_{1i} \pm \hat{b}_2 \sum X_{2i}$$

(ب) وتشتق المعادلة (7–3) بإيجاد النهاية الصغرى للمقدار $\sum e_i^2$ بالنسبة إلى $\hat{b}_1$:

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

$$-2X_{1i}\sum\left(Y_i - \hat{b}_0 - \hat{b}_1 X_{1i} - \hat{b}_2 X_{2i}\right) = 0 \qquad (3-7)$$

$$\sum X_{1i} Y_i = \hat{b}_0 \sum X_{1i} + \hat{b}_1 \sum X_{1i}^2 + b_2 \sum X_{1i} X_{2i}$$

(ج) وتشتق المعادلة الطبيعية (7–4) بإيجاد النهاية الصغرى للمقدار $\sum e_i^2$ بالنسبة إلى$\hat{b}_2$:

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

$$-2X_{2i}\sum\left(Y_i - \hat{b}_0 - \hat{b}_1 X_{1i} - \hat{b}_2 X_{2i}\right) = 0 \qquad (4-7)$$

$$\sum X_{2i} Y_i = \hat{b}_0 \sum X_{2i} + \hat{b}_1 \sum X_{1i} X_{2i} + b_2 \sum X_{2i}^2$$

7–3: بالنسبة لنموذج الانحدار الخطى المتعدِّد ذي المتغيرين المستقلين، (أ) اشتق المعادلات الطبيعية باستخدام الانحرافات (إرشاد: ابدأ باشتقاق تعبير $\hat{y}_i$: يمكن للقارئ غير الملم بالتفاضل أن يتخطى هذا الجزء من المسألة). (ب) كيف يمكن اشتقاق المعادلات 7–5، 7–6، 7–7، لإيجاد $\hat{b}_1$، $\hat{b}_2$، $\hat{b}_0$:

(أ)

$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_{1i} + \hat{b}_2 X_{2i}$$

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

بالطرح، نحصل على:

$$\hat{y}_i = \hat{Y}_i - \bar{Y} = \hat{b}_1 x_{i1} + \hat{b}_2 x_{2i}$$

وعليه:

$$e_i = y_i - \hat{y}_i = \hat{b}_1 x_{1i} - \hat{b}_2 x_{2i}$$

$$\sum e_i^2 = \sum (y_i - \hat{y}_i)^2 = \sum\left(y_i - \hat{b}_1 x_{1i} - \hat{b}_2 x_{2i}\right)^2$$

$$\frac{\partial \sum e_i^2}{\partial \hat{b}_1} = \frac{\partial \sum \left(y_i - \hat{b}_1 x_{1i} - \hat{b}_2 x_{2i}\right)^2}{\partial \hat{b}_1} = 0$$

$$-2x_{1i} \sum \left(y_i - \hat{b}_1 x_{1i} - \hat{b}_2 x_{2i}\right) = 0 \qquad (16-7)$$

$$\sum x_{1i} y_i = \hat{b}_1 \sum x_{1i}^2 + \hat{b}_2 \sum x_{1i} x_{2i} \qquad (16-7)$$

$$\frac{\partial \sum e_i^2}{\partial \hat{b}_2} = \frac{\partial \sum \left(y_i - \hat{b}_1 x_{1i} - \hat{b}_2 x_{2i}\right)^2}{\partial \hat{b}_2} = 0$$

$$-2x_{2i} \sum \left(y_i - \hat{b}_1 x_{1i} - \hat{b}_2 x_{2i}\right) = 0$$

$$\sum x_{2i} y_i = \hat{b}_1 \sum x_{1i} x_{2i} + \hat{b}_2 \sum x_{2i}^2 \qquad (17-7)$$

(ب) المعادلات (7−5)، (7−6) لحساب $\hat{b}_1$ و $\hat{b}_2$ على الترتيب، يتم الحصول عليها بحل معادلات (7−16)، (7−17) آنيًا، ويمكن دائمًا حساب $\hat{b}_1$ و $\hat{b}_2$ ، إلاَّ إذا كانت هناك علاقة خطية تامة بين X_1 و X_2 أو كان عدد المشاهدات عن كل متغير في النموذج 3 أو أقل. ويمكن حساب المعلمة $\hat{b}_0$ بالتعويض في معادلة (7−7) بقيم $\hat{b}_1$ و $\hat{b}_2$ (المحسوبة باستخدام معادلات (7−5)، (7−6))، وقيم $\bar{Y}$ و $\bar{X}_1$ و $\bar{X}_2$ (المحسوبة من معطيات المسألة).

7−4: بالنسبة لتحليل الانحدار المتعدِّد ذي المتغيرين المستقلين بين معنى (أ) b_0 (ب) b_1، (ج) b_2 (د) هل $\hat{b}_0$ و $\hat{b}_1$ و $\hat{b}_2$ BLUE؟

(أ) المعلمة b_0 هي الحد الثابت أو مقطع الانحدار وتعطى قيمة المتغير Y_i، عندما $X_{i1} = X_{2i} = 0$

(ب) تقيس المعلمة b_1 التغير في Y لكل وحدة تغير في X_1 مع إبقاء X_2 ثابتة. ومعلمة الميل b_1 هي معامل انحدار جزئي لأنها تناظر المشتقة الجزئية للمتغير Y بالنسبة إلى X_1، أي $\partial Y / \partial X_1$.

(ج) تقيس المعلمة b_2 التغير في Y لكل وحدة تغير في X_2 مع إبقاء X_1 ثابتة. ومعلمة الميل b_2 هي المعامل الجزئي الثاني للانحدار لأنها تناظر المشتقة الجزئية للمتغير Y بالنسبة إلى X_2، أي: $\partial Y / \partial X_2$

(د) حيث أنه يتم الحصول على $\hat{b}_0$ و $\hat{b}_1$ و $\hat{b}_2$ بطريقة OLS، فإنها أيضًا أفضل مقدرات خطية غير متحيزة (BLUE، انظر قسم 6−5). أي أن $E(\hat{b}_0) = b_0$، $E(\hat{b}_1) = b_1$ ، $E(\hat{b}_2) = b_2$، و $\hat{S}_{b_0}$،$\hat{S}_{b_1}$، $\hat{S}_{b_2}$ أصغر منها لأي مقدرات خطية غير متحيزة أخرى، ولما كان إثبات هذه الخصائص يمثل عبئًا ثقيلاً بدون استخدام جبر المصفوفات لذا لا تتناولها هنا (سلفاتور، 1982 صفحة 175).

7−5: جدول 7−3 هو امتداد للجدول 6−11 ويعطي دخل الفرد الحقيقي بآلاف الدولارات Y، مع نسبة القوة العاملة في الزراعة، X_1 ومتوسط سنوات التعليم للسكان فوق سن 25 سنة، X_2 لعدد 15 دولة متقدمة في

1981. (أ) أوجد معادلة انحدار المربعات الصغرى للمتغير Y على X_1 وX_2 (ب) فسر النتائج في (أ) وقارنها بنتائج المسألة 6–30.

جدول 3–7 دخل الفرد، القوة العاملة في الزراعة، وسنوات التعليم

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Y	6	8	8	7	7	12	9	8	9	10	10	11	9	10	11
X_1	9	10	8	7	10	4	5	5	6	8	7	4	9	5	8
X_2	8	13	11	10	12	16	10	10	12	14	12	16	14	10	12

جدول 7–4 مسودة لتقدير المعالم لبيانات جدول 7–3

n	y	X_1	X_2	y	x_1	x_2	x_1y	x_2y	x_1x_2	x_1^2	x_2^2
1	6	9	8	3-	2	4-	6-	12	8-	4	16
2	8	10	13	-1	3	1	-3	-1	3	9	1
3	8	8	11	-1	1	-1	-1	1	-1	1	1
4	7	1	10	-2	0	-2	0	4	0	0	4
5	7	10	12	-2	3	0	-6	0	0	9	0
6	12	4	16	3	-3	4	-9	12	-12	9	16
7	9	5	10	0	-2	-2	0	0	4	4	4
8	8	5	10	-1	-2	-2	2	2	4	4	4
9	9	6	12	0	-1	0	0	0	0	1	0
10	10	8	14	1	1	2	1	2	2	1	4
11	10	7	12	1	0	0	0	0	0	0	0
12	11	4	16	2	-3	4	-6	8	-12	9	16
13	9	9	14	0	2	2	-0	0	4	4	4
14	10	5	10	1	-2	-2	-2	-2	4	4	4
15	11	8	12	2	1	0	2	0	0	1	0
$n = 15$	$\sum Y = 139$ $\bar{Y} = 9$	$\sum X_1 = 105$ $\bar{X}_1 = 7$	$\sum X_2 = 180$ $\bar{X}_2 = 12$	$\sum y = 0$	$\sum x_1 = 0$	$\sum x_2 = 0$	$\sum x_1y = -28$	$\sum x_2y = 38$	$\sum x_1x_2 = -12$	$\sum x_1^2 = 60$	$\sum x_2^2 = 74$

(أ) يبين جدول 7–4 الحســــابات اللازمة لتقدير معالم معادلة انحدار OLS للمتغير Y على المتغيرين X_1 وX_2:

$$\hat{b}_1 = \frac{(\sum x_1y)(\sum x_2^2) - (\sum x_2y)(\sum x_1x_2)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1x_2)^2} = \frac{(-28)(74) - (38)(-12)}{(60)(74) - (-12)^2}$$

$$= \frac{-2{,}075 + 456}{4{,}440 - 144} \cong -0.38$$

$$\hat{b}_2 = \frac{(\sum x_2 y)(\sum x_1^2) - (\sum x_1 y)(\sum x_1 x_2)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2} = \frac{(38)(60) - (-28)(-12)}{(60)(74) - (-12)^2}$$
$$= \frac{2{,}280 - 336}{4{,}440 - 144} \cong 0.45$$
$$\hat{b}_0 = \bar{Y} - \hat{b}_1 \bar{X}_1 - \hat{b}_2 \bar{X}_2 = 9 - (-0.38)(7) - (0.45)(12) = 9 + 2.66 - 5.40 \cong 6.26$$

وعليه فمعادلة OLS لانحدار Y على X_1 و X_2 هي:

$$\hat{Y}_i = 6.26 - 0.38X_{1i} + 0.45X_{2i}$$

(ب) تشير معادلة انحدار OLS المقدرة على أن مستوى دخل الفرد الحقيقي Y ، يرتبط عكسيًا مع نسبة القوة العاملة في الزراعة X_1، وطرديًا مع عدد سنوات التعليم للسكان فوق من 25 (كما قد يكون متوقعًا). بالتحديد تشير $\hat{b}_1$، إلى أن نقص نسبة القوة العاملة في الزراعة بمقدار 1% من إجمالي القوة العاملة سوف يصاحبه زيادة قدرها 380 دولارًا أمريكيًا في دخل الفرد مع تثبيت X_2 ولكن زيادة سنة واحدة في سنوات التعليم للسكان فوق سن 25 سنة يصاحبها زيادة في دخل الفرد قدرها 450 دولارًا أمريكيًا، مع تثبيت X_1. وعند: $X_{1i} = X_{2i}| = 0,\ \hat{Y}_i = \hat{b}_0 = 6.26$. وبقدر ما اتضح أن X_2 معنوية إحصائيًا (انظر المسألة 8-7 (ب))، وبالتالي يجب أن تدخل في علاقة الانحدار، فقد اتضح أيضًا أن $\hat{b}_1 \cong -0.47$ السابق إيجادها في تمرين6-30، لا تكون تقديرًا موثوقًا للمعلمة b_1.

اختبارات معنوية تقديرات المعالم:

7-6: عرف (أ) σ_u^2 و s^2، (ب) تباين $\hat{b}_1$ وتباين $\hat{b}_2$، (ج) و، $s_{\hat{b}_1}^2$ و $s_{\hat{b}_2}^2$ ، (د) $s_{\hat{b}_1}$ و $s_{\hat{b}_2}$. (هـ) لماذا لا تكون b_0 عادة موضع اهتمام أساسي ؟

(أ)σ_u^2 هو تباين حدا لخطأ في العلاقة الحقيقية بين X_{1i} و X_{2i} و Y_i ولكن $s^2 = \hat{\sigma}_u^2 = \sum e_i^2/(n-k)$ هي تباين البواقي وهي تقدير غير متحيز للتباين غير المعلوم σ_u^2. k هي عدد المعالم المقدرة. في حالة الانحدار المتعدد ذي المتغيرين، $k = 3$. وعليه $n - k = n - 3 = df$

(ب): $$Var\hat{b}_1 = \sigma_u^2 \frac{\sum x_2^2}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2}$$

بينما: $$Var\hat{b}_2 = \sigma_u^2 \frac{\sum x_1^2}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2}$$

إن تبايني $\hat{b}_1$ و $\hat{b}_2$ (أو تقديراتها) مطلوبة لاختبار الفروض وتكوين فترات الثقة لكل من b_1 و b_2 .

$$s_{\hat{b}_1}^2 = s^2 \frac{\sum x_2^2}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2} = \frac{\sum e_i^2}{n-k} \frac{\sum x_2^2}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2}$$

(ج): $$s_{\hat{b}_2}^2 = s^2 \frac{\sum x_1^2}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2} = \frac{\sum e_i^2}{n-k} \frac{\sum x_1^2}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2}$$

و $s^2_{\hat{b}_1}$ و $s^2_{\hat{b}_2}$ هما على الترتيب، تقديران غير متحيزين لتباين b_1 وتباين b_2 غير المعلومين حيث أن σ^2_u غير معلومة.

(د): $S_{\hat{b}_1} = \sqrt{s^2_{\hat{b}_1}}$ و $S_{\hat{b}_2} = \sqrt{s^2_{\hat{b}_2}}$ ، $S_{\hat{b}_1}$ و $S_{\hat{b}_2}$ هما على الترتيب، الانحراف المعياري لكل من $\hat{b}_2$ و $\hat{b}_2$ ويسميان بالأخطاء المعيارية.

(هـ) ما لم تتوفر مشـاهدات كافية بالقرب من $X_{1i} = X_{2i} = 0$ فإن معلمة المقطع b_0 لا تكون عادة ذات أهمية أسـاسـية ويمكن حذف اختبار معنوية الإحصـائية الخاص بها ومعادلة (18−7) لتباين $\hat{\mathrm{b}}_0$ معادلة معقدة في الحساب ولهذا السبب أيضًا فمن النادر أن تذكر أو تستخدم:

$$Var\hat{\mathrm{b}}_0 = \sigma^2_u = \frac{\sum X_1^2 \sum X_2^2 - (\sum X_1X_2)^2}{n[\sum X_1^2X_2^2 - (\sum X_1X_2)^2] - \sum X_1 (\sum X_1 \sum X_2^2 - \sum X_2 \sum X_1X_2) + \sum X_1 (\sum X_1X_2 - \sum X_2 \sum X_1^2)} \dots\dots\dots\dots (18-7)$$

ومع ذلك، ترد $S_{\hat{b}_0}$ أحيانًا في نتائج الكمبيوتر، ويمكن إجراء الاختبارات الإحصائية لمعنوية b_0 بسهولة.

__7−7:__ من بيانات جدول 3−7، أوجد (أ) S^2، (ب) $s^2_{\hat{b}_1}$ و $S_{\hat{b}_1}$ (ج) $s^2_{\hat{b}_2}$ و $S_{\hat{b}_2}$

(أ) الحسابات اللازمة لإيجاد S^2 موضـحة في جدول 5−7، وهو امتداد لجدول 4−7. وقد تم الحصـول على قيم $\hat{Y}_i$ بالتعويض بقيم X_{i1} و X_{2i} في معادلة انحدار OLS المقدرة السابق إيجادها في المسألة 5−7(أ).

$$S^2 = \hat{\sigma}^2_u = \frac{\sum e_i^2}{n-k} = \frac{12.2730}{15-3} \cong 1.02$$

جدول 5−7 انحدار دخل الفرد: حسابات اختبار معنوية المعالم

الدولة	Y	X_1	X_2	$\hat{Y}$	e	e^2
1	6	9	8	6.44	−0.44	0.1936
2	8	10	13	8.31	−0.31	0.0961
3	8	8	11	8.17	−0.17	0.0289
4	7	7	10	8.10	−1.10	1.2100
5	7	10	12	7.86	−0.86	0.7396
5	7	10	12	7.86	−0.86	0.0036
7	9	5	10	8.86	0.14	0.0196
8	8	5	10	8.86	−0.86	0.7396
9	9	6	12	9.38	−0.38	0.1444
10	10	8	14	9.52	0.48	0.2304
11	10	7	12	9.00	1.00	1.0000
12	11	4	16	11.94	−0.94	0.8836
13	9	9	14	9.14	−0.14	0.0196
14	10	5	10	8.86	1.14	1.2996
15	11	8	12	8.62	2.38	5.6644

n = 15					$\sum e = 0$	$\sum e^2 = 12.2730$

(ب) باستخدام قيمة S^2 السابق إيجادها في (أ) وقيم جدول 7–4، نحصل على:

$$s_{\hat{b}_1}^2 = S^2 \frac{\sum x_2^2}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2} \cong 1.02 \frac{74}{(60)(74) - (-12)^2} \cong 0.02$$

$$s_{\hat{b}_1} \cong \sqrt{0.02} \cong 0.14$$

(ج)
$$s_{\hat{b}_1}^2 = s^2 \frac{\sum x_1^2}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2} \cong 1.02 \frac{60}{(60)(74) - (-12)^2} \simeq 0.01$$

$$s_{\hat{b}_2} \cong \sqrt{0.01} \cong 0.10$$

7–8: اختبر عند مستوى معنوية 5% كل من (أ) b_1 و (ب)b_2 في المسألة 7–5 (أ).

(أ)
$$t_1 = \frac{\hat{b}_1 - b_1}{s_{\hat{b}_1}} = \frac{-0.38-0}{0.14} \cong 2.71$$

وحيث أن قيمة t_1 المطلقة تتجاوز القيمة الجدولية $t = 2.179$ (من ملحق5) بمستوى معنوية 5% اختبار ذو ذيلين)، $n - k = 15 - 3 = 12df$ فإننا نستنتج أن b_1 معنوية إحصائيًا عند مستوى معنوية 5% (أي أننا لا نستطيع أن نرفض H_1، بأن $b_1 \neq 0$).

(ب)
$$t_2 = \frac{\hat{b}_2 - b_2}{S_{\hat{b}_2}} \cong \frac{0.45-0}{0.10} \cong 4.50$$

أي أن b_2 معنوية إحصائيًا عند مستوى معنوية 5% (وأيضًا 1%) (أي أنه لا يمكن رفض H_1 بأن $b_2 \neq 0$ كون فترة الثقة 95% لكل من (أ)b_1 ، (ب) b_2 في المسألة 7–5 (أ).

(أ) فترة الثقة 95% للمعلمة b_1:

$$b_1 = \hat{b}_1 \pm 2.179 s_{\hat{b}_1} = -0.38 \pm 2.179(0.14) = -0.38 \pm 0.31$$

أي أن b_1 بين -0.69 ، -0.07 (أي $-0.69 \leq b_1 \leq -0.07$) بدرجة ثقة 95%.

(ب) فترة الثقة 95% المعلمة b_2:

$$b_2 = \hat{b}_2 \pm 2.179\, s_{\hat{b}_2} = 0.45 \pm 1.179(0.10) = 0.45 \pm 0.22$$

أي أن b_2 بين 0.23، 0.67 (أي $0.23 \leq b_2 \leq 0.67$) بدرجة ثقة 95%

معامل التحديد المتعدد:

7–10: بدءًا باستخدام $R^2 = 1 - \sum e_i^2 / \sum y_i^2$ اشتق $R^2 = \left(\hat{b}_1 \sum y x_1 + \hat{b}_2 \sum y x_2\right) / \sum y_i^2 |$ (إرشاد: ابدأ بتبيان أن $\sum e_i^2 = \sum y_i^2 - \hat{b}_1 \sum y x_1 - \hat{b}_2 \sum y x_2$. يمكن القارئ غير الملم بالتفاضل أن يتخطي هذه المسألة).

$$\sum e_i^2 = \sum e_i\,(y_i - \hat{y}_i) = \sum e_i\left(y_1 - \hat{b}_1 x_{1i} - \hat{b}_2 x_{2i}\right)$$
$$= \sum e_i y_i - \hat{b}_1 \sum e_i x_{1i} - \hat{b}_2 \sum e_i x_{2i}$$

ولكن من طريقة OLS وجدنا:

$$\frac{\partial \sum e_i^2}{\partial \sum \hat{b}_1} = -\sum e_i x_{1i} = 0 \quad and \quad \sum e_i x_{1i} = 0$$
$$\frac{\partial \sum e_i^2}{\partial \sum \hat{b}_2} = -\sum e_i x_{2i} = 0 \quad and \quad \sum e_i x_{2i} = 0$$

والتالي:

$$\sum e_i^2 = \sum e_i y_i = \sum (y_i - \hat{y}_i)\, y_i = \sum y_i \left(y_i - \hat{b}_1 \sum y_1 x_{1i} - \hat{b}_2 \sum y_i x_{2i}\right)$$
$$= \sum y_i^2 - \hat{b}_1 \sum y_1 x_{1i} - \hat{b}_2 \sum y_i x_{2i}$$

بالتعويض في معادلة R^2 :

$$\mathrm{R}^2 = 1 - \frac{\sum e_i^2}{\sum y_i^2} = 1 - \frac{\sum y_i^2 - \hat{b}_1 \sum y_1 x_{1i} - \hat{b}_2 \sum y_i x_{2i}}{\sum y_i^2} = \frac{\hat{b}_1 \sum y_1 x_{1i} - \hat{b}_2 \sum y_i x_{2i}}{\sum y_i^2}$$

أو بحذف i للتبسيط، نحصل عل (كما في قسم 7–3):

$$\mathrm{R}^2 = \frac{\hat{b}_1 \sum x y_1 + \hat{b}_2 \sum x y_2}{\sum y^2}$$

7–11: أوجد R^2 من معادلة انحدار OLS المقدرة في المسألة 7–5 (أ)، باستخدام: (أ) $R^2 = \sum \hat{y}_i^2 / \sum y_i^2$. (ب) $R^2 = 1 - \sum e_i^2 / \sum y_i^2$ (ج) $R^2 = \left(\hat{b}_1 \sum y x_1 + \hat{b}_2 \sum y x_2\right) / \sum y_i^2$

(أ) نعرف من المسألة 6–20 أن:

$$\sum y_i^2 = \sum \hat{y}_i^2 + \sum e_i^2 \quad so\ that \quad \sum \hat{y}_i^2 = \sum y_i^2 - \sum e_i^2$$

حيث $\sum y_i^2 = 40$ (بتربيع وجمع قيم y_i من جدول 7–4) $\sum e_i^2 = 12.2730$ (من جدول 7–5)،

$$\sum \hat{y}_i^2 = 40 - 12.2730 = 27.7270$$

وبالتالي $R^2 = \sum \hat{y}_i^2 / \sum y_i^2 = 27.7270/40 \cong 0.6930 \ or\ 69.32\%$

(ب) باستخدام $\sum e_i^2 = 12.2730$ و $\sum y_i^2 = 40$ نحصل على $R^2 = 1 - \sum e_i^2 / \sum y_i^2 = 1 - 12.2730/40 \cong 0.6932$ أو 69.32% كما في (أ).

(ج) باستخدام $b_1 = 0.38$ و $b_1 = 0.45$ (السابق إيجادها في مسالة 7–5 (أ)، $\sum y x_1 = -28$ و $\sum y x_2 = 38$ (من جدول 7–4)، $\sum y_i^2 = 40$ نحصل على:

$$R^2 = \frac{\hat{b}_1 \sum yx_1 + \hat{b}_2 \sum yx_2}{\sum y_i^2} = \frac{(-0.38)(-28) + (0.45)(38)}{40} \cong \frac{27.74}{40}$$
$$= 0.6935 \; or \; 69.35\%$$

وتختلف قيمة R^2 هذه قليلا عن تلك السابق إيجادها في (أ)، (ب) كنتيجة لأخطاء التقريب.

7–12: من $R^2 = 1 - \left(\sum e_i^2 / \sum y_i^2\right)$ اشتق $\bar{R}^2$ (ب) ما هو المدى لقيم $\bar{R}^2$ (إرشاد بالنسبة لجزء (أ): ابدأ بالتشابه بين $\sum e_i^2$ و تباين $\sum y_i^2$ وتباين y).

(أ) صعوبة R^2 (غير المعدلة) أنها لا تأخذ في الاعتبار درجات الحرية ولكن:

$\mathrm{e} = s^2 = \sum e_i^2/(n-k)$ حيث: $n - k = df$ وبالتالي: $\sum e_i^2 = s^2(\mathrm{n}-\mathrm{k})$

و $\sum(Y_i - \bar{Y})^2 = \sum y_i^2 = varY(n-1)$

$$R^2 = 1 - \frac{\sum e_i^2}{\sum y_i^2} = 1 - \frac{s^2(n-k)}{var\,Y(n-k)}$$

وعليه: $1 - R^2 = (s^2/var\,Y)(n-k)/(n-1)$ ولكن: $1 - \bar{R}^2 = s^2/var\,Y$ فتكون:

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

وبالحل لإيجاد R^2، نحصل على:

$$\bar{R}^2 = 1 - (1 - \bar{R}^2)\frac{(n-1)}{(n-k)} \qquad (12-7)$$

(ب) عندما $k = 1$ ، $\frac{(n-1)}{(n-k)} = 1$ تكون $R^2 = \bar{R}^2$

عندما $k > 1$ ، $\frac{(n-1)}{(n-k)} > 1$ تكون $R^2 > \bar{R}^2$

عندما تكون n كبيرة، لقيمة معينة k ، تكون $\frac{(n-1)}{(n-k)}$ قريبة من الوحدة، ولن تختلف $\bar{R}^2$ عن R^2 كثيراً.

عندما تكون n صـغيرة وتكون k كبيرة بالنسـبة إلىn ، فإن $\bar{R}^2$ سـوف تكون أصـغر كثيرًا من R^2 وقد تكون R^2 سالبة بالرغم من أن $0 \le R^2 \le 1$: (انظر المسائل من 7–29 إلى 7–32.

7–13: (أ) أوجد $\bar{R}^2$ بالنسبة لمعادلة انحدار OLS المقدرة في مسألة 7–5 (أ).

(ب) كيف تقارن $\bar{R}^2$ المحسوبة في (أ) مع $\bar{R}^2$ في مسألة 7–11 (أ)، في مسألة 6–31 (ج)؟.

(أ) باستخدام $R^2 = 0.6932$ السابق إيجادها في المسألة 7–11 (ب)، نحصل على:

$$\bar{R}^2 = 1 - (1 - \bar{R}^2)\frac{(n-1)}{(n-k)} = 1 - (1 - 0.6932)\frac{15-1}{15-3} \cong 0.6410$$

(ب) $R^2 = 0.33$ في حالة الانحدار البسيط، باستخدام نسبة القوة العاملة في الزراعة X_1، كمتغير مستقل وحيد (انظر المسألة 6–31 (ج)). $R^2 = 0.69$ بعد إضافة سنوات التعليم للسكان فوق سن 25 سنة، X_2 كمتغير مستقل ثان. ولكن عندما نأخذ في الاعتبار حقيقة أن إضافة X_2 يقلل درجات الحرية بمقدار 1 في $n - k = 15 - 2 = 13$ في حالة الانحدار البسيط للمتغير Y على X_1، إلى $n - k = 15 - 3 = 12$ في الانحدار المتعدد للمتغير Y على X_1 و X_2)، فإن $\bar{R}^2$ تنخفض إلى 0.64

وحقيقة أن b_2 وجدت معنوية إحصائيًا (في المسألة 7–8 (ب)) $R^2 = \bar{R}^2 = 0.33$ في حالة الانحدار البسيط للمتغير Y على X_1 ترتفع إلى $\bar{R}^2 = 0.64$ في حالة الانحدار المتعدد للمتغير Y على X_1 و X_2 يبرر الإبقاء على X_2 كمتغير مستقل إضافي في معادلة الانحدار (سلفاتور، 1982 صفحة 180).

7–14: (أ) كيف يمكن إيجاد $\sum e_i^2$ (المطلوبة لإجراء اختبارات المعنوية) بدون إيجاد $\hat{Y}_i$ أولا ؟ (ب) أوجد $\sum e_i^2$ لبيانات جدول 7–3 بدون إيجاد $\hat{Y}_i$ (جدول 7–5).

(أ) باستخدام القيم المقدرة لكمل من b_1 و b_2 وكذلك $\sum yx_1$، $\sum yx_2$ نحصل أولاً على:

$$R^2 = \frac{\hat{b}_1 \sum yx_1 + \hat{b}_2 \sum yx_2}{\sum y^2}$$

وبالتالي $R^2 = 1 - \left(\sum e_i^2 / \sum y_i^2\right)$ فتكون $\sum e_i^2 = (1 - R^2) \sum y_i^2$. وهذه الطريقة لإيجاد $\sum e_i^2$ تتضمن حسابات أقل عن استخدام Y_i (فالحسابات الوحيدة الإضافية بجانب تلك المطلوبة لتقدير b_1 و b_2 هي $\sum y_i^2$).

(ب) من قيمة $R^2 = 0.6935$ السابق إيجادها في المسألة 7–11 (ج) (التي تستخدم فقط القيم المقدرة لكل من b_1 و b_2 السابق إيجادها في مسألة 7–5 (أ) والقيم المحسوبة في جدول 7–4). ومن قيمة $\sum y_i^2 = 40$ من المسألة 7 – 11 (أ)، نحصل على:

$$\sum e_i^2 = (1 - R^2) \sum y_i^2 = (1 - 0.6935)(40) = 12.26$$

قارن هذه بقيمة $\sum e_i^2 = 12.2730$ السابق إيجادها في جدول 7 –5. الفرق الصغير في قيمتي $\sum e_i^2$ اللتين حصلنا عليهما باستخدام الطريقتين راجع إلى أخطاء التقريب). لاحظ، أن إيجاد $\sum e_i^2$ بالطريقة السابقة يلغى تمامًا الحاجة إلى جدول 7–5.

اختبار المعنوية الإجمالية للانحدار:

7 –15: أذكر الفرض العدمي والفرض البديل لاختبار معنوية الانحدار ككل. (ب) كيف تختبر المعنوية الكلية للانحدار؟ ما هو منطق هذا الاختبار؟ (ج) أعط صيغة التباين المفسر، التباين غير المفسر أو تباين البواق.

(أ) يشير اختبار المعنوية الإجمالية للانحدار إلى اختبار الفرض أن كل المتغيرات المستقلة لا تساعد على تفسير التغير في المتغير التابع حول وسطه. وبشكل محدّد.

الفرض العدمي هو: $H_0: b_1 = b_2 = \cdots = b_k = 0$ مقابل الفرض البديل.

ليست كل قيم b_i تساوي الصفر: H_i.

(ب) تختبر المعنوية الكلية للانحدار بحساب النسبة F بين التباين المفسر والتباين غير المفسر أو تباين البواقي وتوحى القيمة «المرتفعة» الإحصائية F بعلاقة معنوية بين المتغير التابع والمتغيرات المستقلة، مؤدية إلى رفض الفرض العدمي بأن معاملات كل المتغيرات المفسرة كلها أصفار.

(ج) التباين المفسّر: $\left(\hat{Y}_i - \bar{Y}\right)^2/(k-1) = RSS/(k-1) = \sum \hat{Y}_i^2/(k-1)$ حيث k عدد المعالم المقدرة (انظر قسم 6–4).

والتباين غير المفسر: $\sum\left(Y_i - \hat{Y}_i\right)^2/(n-k) = ESS/(n-k) = \sum e_i^2/(n-k)$

7–16: (أ) أعط صيغة إحصائية أو نسبة F المحسوبة لحالة الانحدار البسيط وللانحدار عند $n = 15$، $k = 3$. (ب) هل يمكن أن تكون F المحسوبة «كبيرة» ومع ذلك فكل المعالم المقدرة ليست معنوية إحصائيًا؟

(أ)
$$F_{1.n-2} = \frac{\sum \hat{y}_i^2/1}{\sum e_i^2/(n-2)}$$

حيث تشير رموز دليل F إلى عدد درجات الحرية في البسط والمقام على الترتيب في حالة الانحدار البسيط، $F_{1.n-2} = t_{n-2}^2$ لنفس مستوى الثقة. بالنسبة للانحدار المتعدّد عند $n = 15$ |, $k = 3$ و $F_{2.12} = \left(\sum \hat{y}_i^2/2\right)/\left(\sum e_i^2/12\right)$.

(ب) من الممكن أن تكون F المحسوبة كبيرة، وليس بين المعالم المحسوبة ما هو معنوي إحصائيًا. وقد يحدث هذا عندما يكون هناك ارتباط مرتفع بين المتغيرات المستقلة بعضها البعض (انظر قسم 9–2). وغالبًا ما يكون اختبار F ذا فائدة محدودة لأنه من الممكن أن يرفض الفرض العدمي، بصرف النظر عما إذا كان النموذج يشرح «جزءًا كبيرًا» من التغير في Y.

7–17: (أ) أثبت أن:

$$\left[\sum \hat{y}_i^2/(k-1)\right]/\left[\sum e_i^2/(n-k)\right] = [R^2/(k-1)]/[(1-R^2)/(n-k)]$$

(ب) على ضوء نتائج (أ)، ما هي الطريقة البديلة للتعبير عن الفرض لاختبار المعنوية الكلية للانحدار؟

(أ)
$$\frac{\sum \hat{y}_i^2/(k-1)}{\sum e_i^2/(n-k)} = \frac{\sum \hat{y}_i^2}{\sum e_i^2}\frac{n-k}{k-1} = \frac{\sum \hat{y}_i^2/\sum y_i^2}{\sum e_i^2/\sum y_i^2}\frac{n-k}{k-1} = \frac{R^2}{(1-R^2)}\frac{n-k}{k-1} = \frac{R^2/(k-1)}{(1-R^2)/(n-k)}$$

(ب) نسـبة F ، كاختبار لمعنوية القدرة التفسـيرية لكل المتغيرات المسـتقلة معًا، تعادل تقريبًا اختبار معنوية الإحصائية R^2 فإذا قبل الفرض البديل فإننا نتوقع أن تكون R^2، وبالتالي F «عالية».

7 – 18 اختبر عند مستوى معنوية 5% المعنوية الإجمالية لانحدار OLS المقدر في المسألة 7–5 (أ) $[\sum \hat{y}_i^2/(k-1)]/\left(\sum e_i^2/(n-k)\right)$.

(ب)– $[R^2/(k-1)]/[(1-R^2)/(n-k)]$

(أ) باسـتخدام: $\sum \hat{y}_i^2 = 27.727$ من المسـألة 7–11(أ) ، $\sum e_i^2 = 12.2730$ من جدول 7–5،

نحصل على: $$F_{2.12} = \frac{27.727/2}{12.273/12} \cong 13.59$$

وحيث أن القيمة المحسـوبة للنسـبة F تفوق القيمة الجدولية $\mathrm{F} = 3.88$ عند مسـتوى معنوية 5% ودرجات حرية 2، 12 (انظر ملحق 7)، فإننا تقبل الفرض البديل بأنه ليسـت كل قيم b_i تسـاوى الصـفر عند مسـتوى معنوية 5%.

(ب) باستخدام $R^2 = 0.6932$ من تمرين 7–11 (ب)، نحصل على:

$$F_{2.12} = \frac{R^2/(k-1)}{(1-R^2)/n-k} = \frac{0.6932/2}{(1-0.6932)/12} \cong 13.54$$

ونقبل الفرض أن R^2 تخطف معنويًا عن الصفر عند مستوى معنوية 5%

معاملات الارتباط الجزئـي:

7 – 19 (أ) كيف يمكن إبعاد تأثير X_2 من كل من Y وX_1 عند إيجاد $r_{YX_i.X_{2i}}$؟ (ب) ما هو المدى لقيم معادلات الارتباط الجزئي ؟ (ج) ما هي إشـــارة معاملات الارتباط الجزئي ؟ (د) ما فائدة معاملات الارتباط الجزئي ؟

(أ) لإبعاد تأثير X_2 على Y ، فإننا نوجد انحدار Y على X_2، ونوجد البواقي $e_1 = Y^*$. ولإبعاد تأثير X_2 على X_1 فإننا نوجد انحدار X_1 و X_2 ونوجد البواقي $e_2 = X_1^*$. وفي هذه الحالة فإن Y^*. و X_1^* تمثلان التغير في Y وX_1 على الترتيب، الباقي بدون تفسـير بعد إزاحة تفسـير X_2 على كل من Y وX_1 وبالتالي، فعامل الارتباط الجزئي ليس إلا معامل الارتباط البسيط بين البواقي Y^* و X_1^* (أى أن، $r_{YX_1.X_2} = r_{Y^*X_1^*}$)

(ب) مدى معاملات الارتباط الجزئي هو من 1− إلى 1+ (تمامًا كما في حالة معاملات الارتباط البسـيط). على سبيل المثال، تشير $r_{YX_1.X_2} = -1$ إلى الحالة عندما توجد علاقة خطية تامة عكسية بين Y وX_1 بعد إزاحة التأثير المشـــترك للمتغير X_2 على كل من Y وX_1. ولكن، تثير $r_{YX_1.X_2} = 1$ إلى علاقة خطية تامة طردية صـافية بين Y وY_1.. تشير $r_{YX_1.X_2} = 0$ إلى عدم وجود علاقة بين Y وX_1 بعد إزاحة تأثير X_2 على كل من Y وX_1 وكنتيجة، فإنه يمكن حذف X_1 من الانحدار (سلفاتور، 1982 صفحة 184).

(ج) إشـارة معامل الانحدار الجزئي في نفس إشـارة المعلمة المقدرة المناظرة. فمثلا، بالنسـبة لمعادلة الانحدار المقدرة: $\hat{Y} = \hat{b}_0 + \hat{b}_1 X_1 + \hat{b}_2 X_2 + r_{YX_1.X_2}$ فإن $r_{YX_1.X_2}$ له نفس إشـــارة $\hat{b}_1$ وكذلك $r_{YX_2.X_1}$ له نفس إشارة في $\hat{b}_2$.

(د) تستخدم معاملات الارتباط الجزئية في تحليل الانحدار المتعدّد لتحديد الأهمية النسبية لكل متغير مفسر في النموذج. والمتغير المستقل صـاحب أعلى معامل ارتباط جزئي مع المتغير التابع يسـاهم أكثر من المتغيرات المســـتقلة الأخرى في القدرة التفســـيرية للنموذج ويدخل أولاً في تحليل الانحدار المتعدّد خطوة – بخطوة. ولكن يجب ملاحظة أن معامل الارتباط الجزئي يعطي مقياسًـا لترتيب صـافى الارتباط وليس مقياسًـا لقيمته، فمجموع معاملات الارتباط الجزئي بين المتغير التابع وكل المتغيرات المستقلة لا يساوى 1 بالضرورة.

7–20: بالفنية للانحدار المقدر في المسـألة 7–5 (أ)، أوجد (أ) $r_{YX_1.X_2}$، (ب) $r_{YX_2.X_1}$ (ج) هل تسـاهم X_1 أو X_2 أكثر في القدرة التفسيرية للنموذج ؟

(أ) لإيجاد $r_{YX_1.X_2}$ فإننا نحتاج أولا إلى إيجاد r_{YX_1}، r_{YX_2} و$r_{X_1X_2}$. باســـتخدام القيم من جدول 7–4، نحصل على:

$$r_{YX_1} = \frac{\sum x_1 y}{\sqrt{\sum x_1^2}\sqrt{\sum y^2}} = \frac{-28}{\sqrt{60}\sqrt{40}} \cong -0.5715$$

(أ)

$$r_{YX_2} = \frac{\sum x_2 y}{\sqrt{\sum x_2^2}\sqrt{\sum y^2}} = \frac{38}{\sqrt{74}\sqrt{40}} \cong 0.6984$$

$$r_{YX_2} = \frac{\sum x_2 x_1}{\sqrt{\sum x_2^2}\sqrt{\sum x_1^2}} = \frac{-12}{\sqrt{74}\sqrt{60}} \cong -0.1801$$

وبالتالي:

$$r_{YX_1.X_2} = \frac{r_{YX_1} - r_{YX_2} r_{X_1X_2}}{\sqrt{1-r^2_{X_1X_2}}\sqrt{1-r^2_{YX_2}}} = \frac{(-0.5715)-(0.6984)(-0.1801)}{\sqrt{1-(-0.1801)^2}\sqrt{1-0.6984^2}} \cong -0.6331$$

(ب) باستخدام قيم r_{YX_1}، r_{YX_2}، $r_{X_1.X_2}$ السابق حسابها في (أ)، نحصل على:

$$r_{YX_2.X_1} = \frac{r_{YX_2} - r_{YX_1} r_{X_1X_2}}{\sqrt{1 - r^2_{X_1X_2}}\sqrt{1 - r^2_{YX_1}}} = \frac{(0.6984) - (-0.5715)(-0.1801)}{\sqrt{1 - (-0.1801)^2}\sqrt{1 - (-0.5715)^2}} \cong 0.8072$$

(ج) حيث أن: $r_{YX_1.X_2}$ تتجاوز القيمة المطلقة للمعامل ..... فإننا نسـتنتج أن X_2 تسـاهم أكثر من Y_1 في القدرة التفسيرية النموذج.

مسائل شاملة:

7–21: يعطى جدول 7–6 الكمية المطلوبة (فرضًـا) من سـلعة ما، Y وسـعرها X_1، ودخل المسـتهلك، X_2 من عام 1971 إلى 1985. (أ) هيء انحدار OLS لهذه المشاهدات.

جدول 7–6 الكمية المطلوبة من سلعة ما، سعرها، ودخل المستهلك، 1971 – 1980:

السنة	Y	X_1	X_2
1971	40	9	400
1972	45	8	500
1973	50	9	600
1974	55	8	700
1975	60	7	800
1976	70	6	900
1977	65	6	1000
1978	65	8	1100
1979	75	5	1200
1980	75	5	1300
1981	80	5	1400
1982	100	3	1500
1983	90	4	1600
1984	95	3	1700
1985	85	4	1800

(ب) اختبر عند مستوى معنوية 5% المعنوية الإحصائية لمعالم الميل. (ج) أوجد معامل الارتباط المتعدد غير المعدل والمعدل. (د) اختبر المعنوية الكلية للانحدار (هـ) أوجد معاملات الارتباط الجزئي وحدّد أي متغير مستقل يساهم أكثر في قدرة النموذج التفسيرية (و) أوجد معامل المرونة السعرية للطلب η_P، والمرونة الداخلية الطلب، η_M، عند المتوسطات. (ز) ضع جميع النتائج في شكل ملخص مع تقريب كل الحسابات إلى 4 علامات عشرية.

(أ) يعطي جدول 7–7 الحسابات اللازمة لتوفيق الانحدار الخطي.

$$\hat{b}_1 = \frac{(\sum x_1 y)(\sum x_2^2) - (\sum x_2 y)(\sum x_1 x_2)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2}$$

$$= \frac{(-505)(2,800,000) - (107,500)(-11,900)}{(60)(2,800,000) - (-11,900)^2} \cong 5.1061$$

$$\hat{b}_2 = \frac{(\sum x_2 y)(\sum x_1^2) - (\sum x_1 y)(\sum x_1 x_2)}{(\sum x_1^2)(\sum x_2^2) - \left(\sum x_1\, x_2\right)^2} = \frac{(107{,}500)(60) - (-505)(-11{,}900)}{(60)(2{,}800{,}000) - (-11{,}900)^2}$$
$$\cong 0.1607$$
$$\hat{b}_0 = \bar{Y} - \hat{b}_1 \bar{X}_1 - \hat{b}_2 \bar{X}_2 = 70 - (-5.1061)(6) - (0.0167)(1{,}100) \cong 82.2666$$
$$\hat{Y} = 82.2666 - 5.1061 X_1 + 0.0167 X_2$$

(ب) ويمكننا إيجاد $\sum e_i^2$ بحساب R^2 أولًا من جدول 7–7:

$$R^2 = \frac{\hat{b}_1 \sum y x_1 + \hat{b}_2 \sum y x_2}{\sum y^2} = \frac{(-5.1061)(-505) + (0.0167)(107{,}500)}{4{,}600} \cong 0.9508$$

ولكن:
$$R^2 = 1 - \frac{\sum e^2}{\sum y^2}$$

وبالتالي:
$$\sum e^2 = (1 - R^2) \sum y^2 = (1 - 0.9508)600 \cong 226.32$$

$$s_{\hat{b}_1}^2 = \frac{\sum e_i^2}{n-k} \frac{\sum x_2^2}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2} = \frac{226.32}{15-3} \frac{2{,}800{,}000}{(60)(2{,}800{,}000) - (-11{,}900)^2}$$
$$\cong 2.0011 \quad and \qquad s_{\hat{b}_1^2} \cong 1.4146$$

$$s_{\hat{b}_2}^2 = \frac{\sum e_i^2}{n-k} \frac{\sum x_i}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2} = \frac{226.32}{15.3} \frac{60}{(60)(2{,}800{,}000) - (-11{,}900)^2}$$
$$\cong 0.00004 \quad and \qquad s_{\hat{b}_2} \cong 0.0065$$

$$t_1 = \frac{\hat{b}_1}{s_{\hat{b}_1}} = \frac{-5.1061}{1.4146} \cong -3.6096 \quad and \quad t_2 = \frac{\hat{b}_2}{s_{\hat{b}_2}} = \frac{0.0167}{0.0065} \cong 2.5692$$

وبالتالي، فإن كلا من $\hat{b}_1$ و $\hat{b}_2$ معنوية إحصائيًا عند مستوى معنوية 5%

(ج) $R^2 = 0.9508$ (من (ب))، وعليه:

$$\bar{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-k} = 1 - (1 - 0.9508)\frac{15-1}{15-3} \cong 0.9426$$

(د)–
$$F_{k-1,n-k} = \frac{R^2/(k-1)}{(1-R^2)/(n-k)} = \frac{0.9508/(3-1)}{(1-0.9508)/(15-3)} \cong 115.9512$$

وبالتالي، فإن R^2 تختلف معنويًا من الصفر عند مستوى معنوية 5%

السنة	Y	X_1	X_2	y	x_1	x_2	yx_1	yx_2	x_1x_2	x_1^2	x_2^2	y^2
1971	40	9	400	−30	3	−700	−90	21,000	−2,100	9	490,000	900
1972	45	8	500	−25	2	−600	−50	15,000	−1,200	4	260,000	625
1973	50	9	600	−20	3	−500	−60	10,000	−1,500	9	250,000	400
1974	55	8	700	−15	2	−400	−30	6,000	−800	4	160,000	225
1975	60	7	800	−10	1	−300	−10	3,000	−300	1	90,000	100
1976	70	6	900	0	0	−200	0	0	0	0	40,000	0
1977	65	6	1,000	−5	0	−100	0	500	0	9	10,000	25
1978	65	8	1,100	−5	2	0	−10	0	0	4	0	25
1979	57	5	1,200	5	−1	100	−5	500	−100	1	10,000	25
1980	57	5	1,300	5	−1	200	−5	1,000	−200	1	40,000	25
1981	080	5	1,400	10	−1	300	−10	3,000	−300	1	90,000	100
1982	100	3	1,500	30	−3	400	−90	12,000	−1,200	9	160,000	900
1983	90	4	1,600	20	−2	500	−40	10,000	−1000	4	250,000	400
1984	95	3	1,700	25	−3	600	−75	15,000	−1800	9	360,000	625
1985	85	4	1,800	15	−2	700	−30	10,500	−1400	4	490,000	225
0	$\sum Y = 1,05$ $\bar{Y} = 75$	$\sum X_1 = 90$ $\bar{X}_1 = 6$	$\sum X_2 = 16,50$ $\bar{X}_2 = 1,100$	$\sum y = 0$	$\sum x_1 = 0$	$\sum x_2 = 0$	$\sum yx_1 = -50$	$\sum yx_2 = 107,50$	$\sum x_1x_2 = 11,90$	$\sum x_1^2 = 60$	$\sum x_2^2 = 2,800,0$	$\sum y^2 = 4,60$

(هـ) لإيجاد $r_{YX_1.X_2}$ و $r_{YX_2.X_1}$، فيجب أولاً إيجاد (من جدول 7−7).

$$r_{YX_1} = \frac{\sum x_1 y}{\sqrt{\sum x_1^2}\sqrt{\sum y^2}} = \frac{-505}{\sqrt{60}\sqrt{4,600}} \cong 0.9613$$

$$r_{YX_2} = \frac{\sum x_2 y}{\sqrt{\sum x_2^2}\sqrt{\sum y^2}} = \frac{107,500}{\sqrt{2,800,000}\sqrt{4,600}} \cong 0.9472$$

$$r_{X_1X_2} = \frac{\sum x_1 x_2}{\sqrt{\sum x_2^2}\sqrt{\sum x_1^2}} = \frac{-11,900}{\sqrt{2,800,000}\sqrt{60}} \cong 0.9181$$

$$r_{YX_1.X_2} = \frac{r_{YX_1} - r_{YX_2} r_{X_1X_2}}{\sqrt{1 - r_{X_1X_2}^2}\sqrt{1 - r_{YX_2}^2}} = \frac{(-0.9613) - (0.9472)(-0.9181)}{\sqrt{1 - (-0.9181)^2}\sqrt{1 - (-0.9472)^2}} \cong -0.7213$$

$$r_{YX_2.X_1} = \frac{r_{YX_2} - r_{YX_1} r_{X_1X_2}}{\sqrt{1 - r_{X_1X_2}^2}\sqrt{1 - r_{YX_1}^2}} = \frac{(0.9472) - (0.9613)(-0.9181)}{\sqrt{1 - (-0.9181)^2}\sqrt{1 - (-0.9613)^2}} \cong 0.5919$$

وبالتالي فإن X_1 تساهم أكثر من X_2 في القدرة التفسيرية للنموذج.

$$\eta_P = \hat{b}_1 \frac{\bar{X}_1}{\bar{Y}} = -5.1061 \frac{6}{70} \cong 0.4377$$

(و)
$$\eta_M = \hat{b}_2 \frac{\bar{X}_2}{\bar{Y}} = 0.0167 \frac{1,100}{70} \cong 0.2624$$

(ز)
$$\hat{Y}_1 = 82.2666 - 5.1061X_1 + 0.0167X_2$$
$$value \quad (-3.096) \quad (2.5692)$$
$$R^2 = 0.9508 \quad \bar{R}^2 = 0.9426 \quad , F_{2.12} = 155.9512$$
$$r_{YX_1.X_2} = 0.7023 \quad , \quad r_{YX_2..X_1} = 0.8434$$
$$n_P = -0.4377 \qquad n_M = 0.2624$$

7-22: جدول 7-8 هو نفس جدول 7-6 فيما عدا أنه يشـــمل ســـعر ســـلعة بديلة بالدولار، X_3 كمتغير مســتقل ثالث. جدول 7-9 يعطي صــورة من مخرجات الكمبيوتر للانحدار الخطي للمتغير Y على X_1، X_2 ، X_3 باستخدام SPSS (Statistical Package for the Social Sciences أكثر برامج الكمبيوتر شيوعًا.

جدول 7-8 الكمية المطلوبة، السعر، دخل المستهلك، وسلعة بديلة

السنة	Y	X_1	X_2	X_3
1971	40	9	400	10
1972	45	8	500	14
1973	50	9	600	12
1974	55	8	700	13
1974	60	7	800	11
1976	70	6	900	15
1977	65	6	1,000	16
1978	65	8	1,100	17
1979	75	5	1,200	22
1980	75	5	1,300	19
1981	80	5	1,400	20
1982	100	3	1,500	23
1983	90	4	1,600	18
1984	95	3	1,700	24
1985	85	4	1,800	21

أجب عن الأســـئلة التالية من خلال فحص مخرجات الكمبيوتر في جدول 7-9. (أ) اكتب معادلة انحدار OLS مع قيم t, R^2, $\bar{R}^2$ نســـبة F مع درجات الحرية، الخطأ المعياري للانحدار، ومجموع مربعات البواقي، وفسر النتائج. (ب) كيف أجري برنامج SPSS ؟

(أ) من صفحة 4 من مخرجات الكمبيوتر في جدول 7-9، نحصل على:

$$t_1 = \frac{\hat{b}_1}{s_{\hat{b}_1}} = \frac{-4.9281}{1.6111} = -3.05, \quad t_2 = \frac{\hat{b}_2}{s_{\hat{b}_2}} = \frac{0.0159}{0.0074} = 2.149, t_1 = \frac{\hat{b}_3}{s_{\hat{b}_3}} = \frac{0.1748}{0.6367} = 0.275$$

وعليه:

$$Y = 79.1063 - 4.9281X_1 + 0.0159X_2 + 0.1748X_3$$
$$(-3.059) \quad (2.149) \quad (0.275)$$
$$R^2 = 0.95 \;, \quad \bar{R}^2 = 0.94 \;, \quad F_{3.11} = 71.13$$
$$s = 4.53 \;, \quad SEE = 225.49$$

إشـارات $\hat{b}_1$، $\hat{b}_2$، $\hat{b}_3$ تتفق مع ما هو متوقع طبقًا للنظرية الاقتصـادية. ولكن $\hat{b}_1$ فقط معنوية إحصـائيًا عند مستوى معنوية 5%، بينما كادت $\hat{b}_2$ أن تكون معنوية. من نسبة F، تنبّه إلى أن R^2 (والانحدار ككل) معنوى إحصـائيًا عند مستوى 5%. يجب ملاحظة أنه في صفحة 4 من جدول 9–7، R المتعددة هي الحذر التربيعي للمقدار R تربيع وبالتالي تشير إلى معامل الارتباط المتعدد r.

جدول 9–7 مخرجات الكمبيوتر SPSS لانحدار الطالب

Capture d'ecran

وتشـير بيتا إلى المعاملات المعيارية، أو المعاملات المقدرة مضـروبة في نسـبة الانحراف المعياري لمتغير مستقل معين إلى الانحراف المعياري المتغير التابع. قيمة F لكل معامل ليست إلا مربع قيمة r، لكل معامل.

(ب) الخطوة الأولى لإجراء برنامج كمبيوتر هي «الدخول» من إحدى محطات الكمبيوتر ويحتاج هذا إلى رقم حساب و «وكلمة مرور» (والتي تحصل عليها من مركز الكمبيوتر في جامعتك، أو شركتك). لاستخدام SPSS فإنك تستخدم أمرًا، يمدك به أيضًا مركز الكمبيوتر. الخطوات التالية موجودة في صفحة 1 (السطور من 5 إلى 15) من مخرجات الكمبيوتر في جدول 9–7. وهذه في معظمها لا تحتاج إلى شـــرح. في الســـطر الرابع (INPUT FORMAT) تشـــير FREEFIELD إلى الطريقة التي يتغذى بها الكمبيوتر بالبيانات، حيث تدخل قيمة كل متغير في التسلسل المشار إليه في سطر 3 (أي $75ect$ 24 1,700 3 95 12 600 9 50) في ســطر 7، تشــير الأقواس إلى أن المطلوب هو انحدار متعدّد عادي، بينما تطلب RESID = 0 حســاب البواقي المعيارية وقيم التنبؤ للمتغير Y. ويطلب OPITON 11,12, على سطر 8 كتابة البواق المعيارية وقيم التنبؤ للمتغير Y (المعطاة في صـــفحة 7 من مخرجات الكمبيوتر). وتطلب STATISTICS 1 to 8 في سطر 9 كتابة المتوسط، الانحراف المعياري وعدد الحالات الصحيحة (السنوات) لكل متغير (المعطاة في صفحة 2 من مخرجات الكمبيوتر؛ مصفوفة الارتباط البسيط (المعطاة في صفحة 3) ؛ كل الإحصـائيات في صفحة4 (السابق مناقشتها في (أ))، وإحصائية دير بين واتسون. ارجع إلى دليل SPSS في مركز الكمبيوتر، وأخيرًا، فإن الجدول الموجز في صفحة 5 (ويتضمن كل خرجات الكمبيوتر بشكل أوتوماتيكي (يعالج المتغيرات كما لو كانت أدخلت الكمبيوتر واحدًا وراء الآخر وليس معًا.

مسائـل إضافيـة: النموذج الخطي ذو المتغيرات الثلاثة:

7-23: جدول 7−10 امتداد لجدول 6−12 ويعطي مشـــاهدات عن Y، X_1، X_2. أوجد معادلة انحدار OLS للمتغير Y علي X_1 و X_2.

الإجابة: $\hat{Y}_1 = 4.76 + 5.29X_{1i} + 2.13X_{2i}$

جدول 7−10: مشاهدات عن Y ، X_1 و X_2.

n	1	2	3	4	5	6	7	8	9	10
Y	20	28	40	45	37	52	54	43	65	56
X_1	2	3	5	4	3	5	7	6	7	8
X_2	5	6	6	5	5	7	6	6	7	7

7−24 بالرجوع إلى معادلة انحدار OLS المقدرة في المسألة 7−23 فسر (أ)$\hat{b}_0$ و، (ب) $\hat{b}_1$، (ج) $\hat{b}_2$.

الإجابة:

(أ) $\hat{b}_0 = 4,76$ هي الثابت أو مقطع Y، $\hat{Y}_i = \hat{b}_0 = 4.76$ ، عند $X_{i1} = X_{2i} = 0$

(ب) $b_1 = 5.29$ تشير إلى أن زيادة قدرها الوحدة في X_1 (مع تثبيت X_2) ينتج عنها زيادة في $\hat{Y}_i$ قدرها 5.29 وحدة.

(ج) $b_2 = 2.13$، تشـــير إلى أن زيادة قدرها الوحدة في X_2 (مع تثبيت X_1) ينتج عنها زيادة في $\hat{Y}_i$ قدرها 2.13 وحدة.

اختبارات معنوية تقديرات المعالم:

7−25: بالرجوع إلى بيانات جدول 7−10، أوجد (أ) S^2، (ب) $s^2_{\hat{b}_1}$ و $s_{\hat{b}_1}$ (ج) $s^2_{\hat{b}_2}$ و $s_{\hat{b}_2}$.

الإجابة: (أ)$S^2 = 50$ (ب) $s^2_{\hat{b}_1} \cong 3.16\ and\ s_{\hat{b}_1} \cong 1.78$ (ج) $s^2_{\hat{b}_2} \cong 18.95$ و $s_{\hat{b}_2} \cong 4.35$

7−26 اختبر عند مستوى معنوية 5% كلا من (أ) b_1، (ب) b_2 في المسألة 7−23.

الإجابة: (أ) b_1 معنوية إحصائيًا عند مستوى 5%.

(ب)b_2 ليست معنوية إحصائيًا عند مستوى 5%

7−27 كون فترة الثقة 95% لكل من (أ) b_1، (ب) b_2 في المسألة 7−23.

الإجابة: (أ) $1.08 \le b_1 \le 9.50$ (ب) $-8.16 \le b_2 \le 12.42$

معامل التحديد المتعدد:

7−28 بالنسبة لانحدار OLS المقدرة في المسألة 7−23، أوجد (أ) R^2، (ب) $\bar{R}^2$ (ج) هل يجب أن تدخل X_2 في الانحدار؟

الإجابة: (أ) $R^2 \cong 0.79$ (باستخدام $R^2 = 1 - \left(\sum e_i^2/\sum y_i^2\right)$ (ب) $\bar{R}^2 \cong 0.73$ (ج) حيث أن b_2 وجدت غير معنوية إحصـائيًا (في المسـألة 7−26 (ب))، R^2 نقصـت من $R^2 = \bar{R}^2 = 0.77$ عندما كانت

X_1 المتغير المســتقل الوحيد (انظر المســألة 6–41 (أ)) إلى $\bar{R}^2 = 0.73$ (أعلاه)، فإن X_2 يجب ألا تدخل الانحدار.

7–29 بالنسبة إلى $R^2 = 0.60$، $n = 10$ ، $K = 1$ أوجد $\bar{R}^2$

الإجابة: $\bar{R}^2 = 0.60$

7–30 بالنسبة إلى $R^2 = 0.60$، $n = 10$ ، $K = 2$ ، أوجد $\bar{R}^2$

الإجابة: $\bar{R}^2 = 0.55$

7–31 بالنسبة إلى $\bar{R}^2 = 0.60$، $K = 2$ (كما في المسألة 7–30) ولكن $n = 100$، أوجد $\bar{R}^2$

الإجابة: أوجد $\bar{R}^2 = 0.569$

7–32 بالنسبة إلى $R^2 = 0.40$، $n = 10$ ، $K = 5$ أوجد $\bar{R}^2$.

الإجابة: $\bar{R}^2 = -0.08$ (ولكنها تفسر على أنها تساوى الصفر).

اختبار المعنوية الكلية للانحدار:

7–33 من انحدار OLS المقدر في المسألة 7–23، أوجد (أ) التباين المفسر، (ب) التباين غير المفسر أو تباين البواقي (ج) نسبة أو إحصائية F

الإجابة: (أ) $\sum \hat{y}^2/(k-1) \cong 649$ (ب) $\sum e^2/(n-k) = 50$ (ج) $F_{2,7} = 12.98$

7–34 اختبر المعنوية الإجمالية لانحدار OLS المقدر في المسألة 7–33 عند (أ) مستوى 5% (ب) مستوى 1%.

الإجابة: (أ) حيث أن نسبة F المحسوبة (12.98) تتجاوز قيمة F النظرية أو الجدولية (4,74) عند$\alpha = 0$ ودرجات حرية 2، 7، فإننا تقبل الفرض بأن معالم انحدار OLS المقدرة هي معًا معنوية عند مستوى 5%.

(ب) حيث أن قيمة F الجدولية عند مســـتوى $\alpha = 0.01$ هي F = 9.55 ، يقبل الفرض البديل عند مستوى معنوية 1% أيضًا.

معاملات الارتباط الجزئية:

7–35 بالنســـبة لانحدار OLS المقدر في المســـألة 7–23، أوجد (1) r_{YX_1,X_2}، (ب) r_{YX_2,X_1} (ج) أي ستغير مستقل يساهم أكثر في القدرة التسييرية للنموذج ؟

الإجابة: (1) 0.74 (ب) $r_{YX_2,X_1} = 0.18$ (ج) X_1

مسألة شاملة:

7–36 جدول 7–11 امتداد للجدول 6–13 ويعطى بيانات عينة عشـــوائية من 12 من الأســـر عن عدد الأطفال في الأسـرة Y، وعدد الأطفال الذين قالوا إنهم كانوا يرغبون في إنجابهم وقت الزواج X_1، وعدد سـنوات

تعليم الزوجة، X_2. (أ) أوجد معادلة انحدار OLS للمتغير Y على X_1 و X_2. (ب) احسب قيم t، واختبر عند مستوى 5% المعنوية الإحصائية لمعالم الميل. (ج) أوجد معامل الارتباط المتعدد غير المعدل والمعدل (د) اختبر المعنوية الإجمالية للانحدار (هـ) أوجد معاملات الارتباط الجزئية وحدد أي المتغيرات المستقلة يساهم أكثر في القدرة التفسيرية للنموذج. قم بجميع الحسابات إلى رقمين عشريين.

جدول 11–7 عدد الأطفال في الأسرة وعدد الأطفال المرغوب فيهم وتعليم الزوجة

الأسرة	1	2	3	4	5	6	7	8	9	10	11	12
Y	4	3	0	4	4	3	0	4	3	1	3	1
X_1	3	3	0	2	2	3	0	3	2	1	3	2
X_2	12	14	18	10	10	14	18	12	15	16	14	15

الإجابة: (أ) $\hat{Y} = 6.90 + 0.53X_1 - 0.39X_2$ (ب) حيث أن $t_1 = 3.12$ و $t_2 = -5.57$ فإن كلا من $\hat{b}_1$ $\hat{b}_2$ معنوية إحصائيًا عند مستوى 5% (ج) $R^2 = 0.95$ و $\bar{R}^2 = 0.90$ (د) حيث أن $F_{2,9} = 51.31$ فإن R^2 معنوية إحصائيًا عند مستوى 5% (هـ) $r_{YX_2.X_1} = 0.87$ و $r_{YX_1.X_2} = 0.71$ وبالتالي فإن X_2 تساهم أكثر من X_1 في القدرة التفسيرية النموذج.