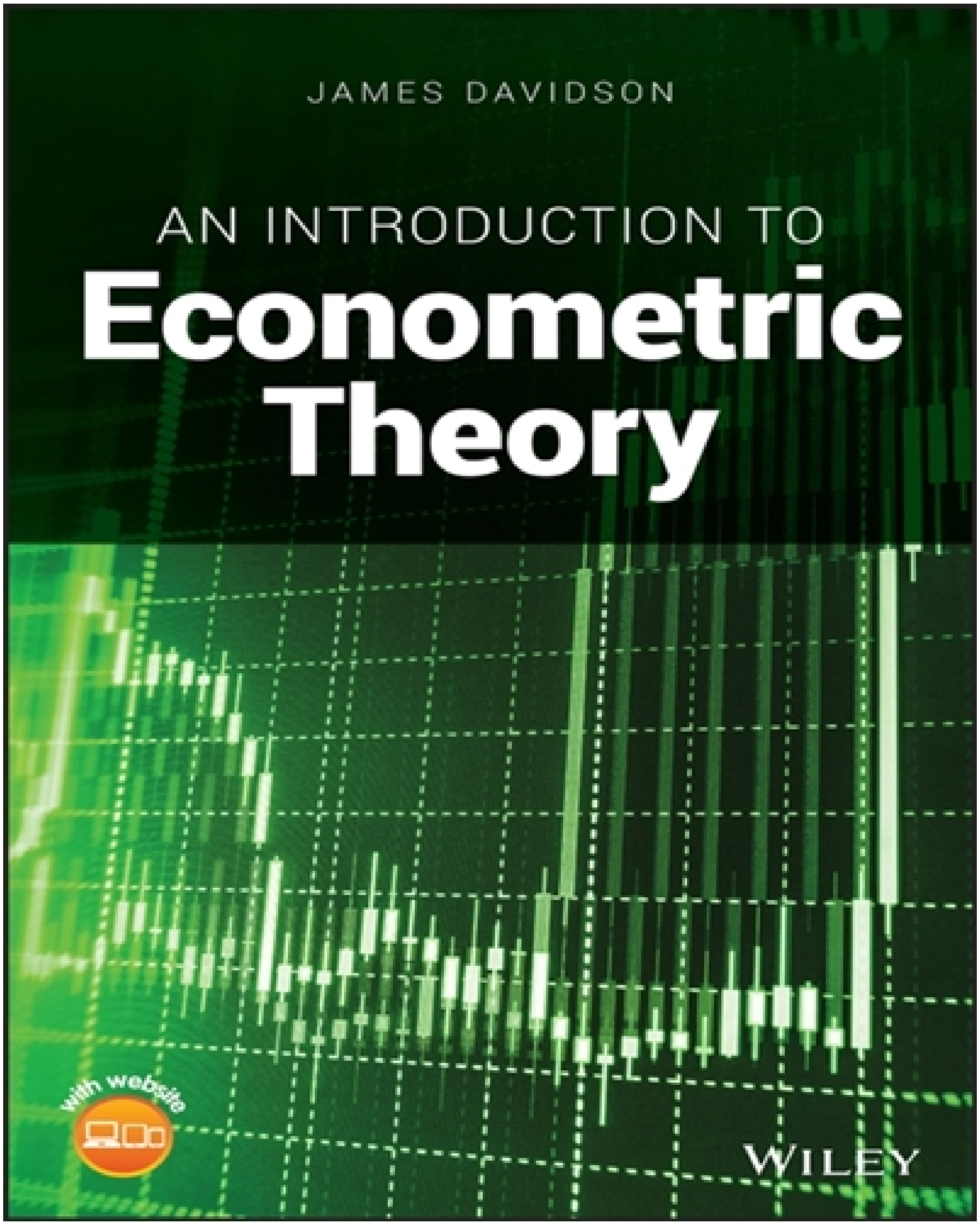
Solutions for Introduction to Econometric Theory 1st Edition by Davidson



Solutions

An Introduction to Econometric Theory: Solutions Manual for the Exercises

James Davidson

John Wiley and Sons, 2018

[copyright page here]

Chapter 1

Elementary Data Analysis: Exercises

1. *Are the following statements True or False?*

   (a) *The correlation of a variable with itself is always 1.*
   TRUE. Put $y_t = x_t$ in (1.11)

   (b) *Chebyshev's rule says that at least a proportion* $1/m^2$ *of any sample lies beyond* m *standard deviations from the mean.*
   FALSE. "At most", not "At least".

   (c) *The least absolute values fitting method is more influenced by outlying observations than least squares.*
   FALSE. Less not more.

   (d) *The sample mean is the least squares measure of location.*
   TRUE, see (1.29).

   (e) *The slope coefficient in the simple regression is the tangent of the angle of the regression line with the horizontal axis.*
   TRUE, see Figure 1.3

   (f) *The least squares estimator of the slope coefficient* (y on x) *is the sum of the products of* y *with the mean deviations of* x*, divided by the sum of squares of the mean deviations of* x*.*
   TRUE, see (1.27).

   (g) *Run a regression in both directions* (y on x *and* x on y)*, and the product of the two slope coefficients is equal to the squared correlation coefficient.*
   TRUE, compare (1.11) with (1.27).

2. *Here are 12 observations on a variable* x:

$$86, 109, 81, 53, 86, -14, 65, -39, 65, 34, 79, -27$$

(a) *Compute the mean.*

$$\bar{x} = 48.1667$$

(b) *Compute the sequence of mean deviations.*

$$60.83, 32.83, 4.83, 37.83, -62.17, 16.83, -87.17, 16.83, -14.17, 30.83, -75.17$$

(c) *Compute the standard deviation.*

$$s = 49.0785$$

(d) *How many of these data points lie more than (i) one, (ii) two, (iii) three standard deviations from the mean?*

(i) 1 SD bounds: $[-0.91, 97.24]$ and 4/12 lie outside.

(ii) 2SD bounds: $[-49.99, 146.32]$ and none lie outside.

(e) *Include the following observations in the set, and obtain the mean and standard deviation for this case.*

209, 475, −114, 46

$$\bar{x} = 74.625, \quad s = 129.009$$

(f) *Repeat exercise (d) for the enlarged data set.*

(i) 1 SD bounds: $[-54.38, 203.63]$ and 3/12 lie outside.

(ii) 2SD bounds: $[-183.99, 332.64]$ and none lie outside.

3. *Here are 12 observations on a variable y.*

43,62,26,46,48,3,52,-10,37,5,42,4

(a) *Compute the mean.*

$$\bar{x} = 29.833$$

(b) *Compute the standard deviation.*

$$s = 23.54$$

(c) *Compute the correlation coefficient of y with x in Question 1.*

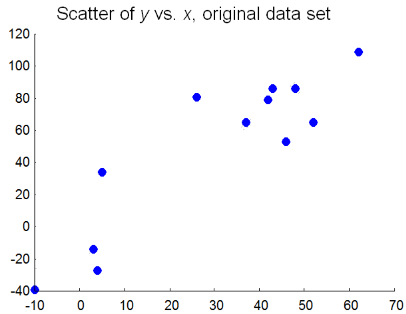
$$r = 0.896$$

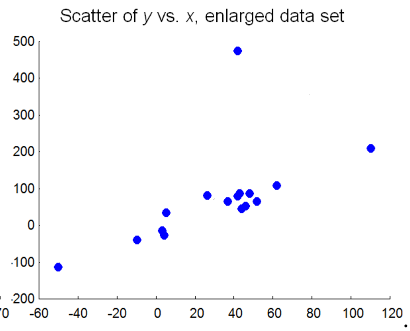
(d) *Include the following data points in the set and compute the correlation coefficient with the enlarged data set of Question 1.*

110,42,-50,44

$$r = 0.602$$

(e) *Draw scatter plots of the two cases*





4. *Compute the regression of y on x (original sample).*

(a) *Report the fitted slope and intercept coefficients and the residuals, $\hat{u}$.*

$$\hat{y} = 9.11 + 0.43x$$

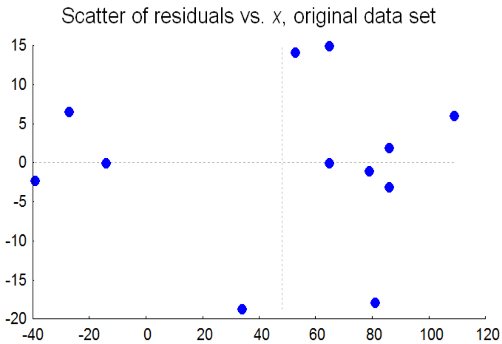
Residuals:

$-3.1, 5.99, -17.96, 14.09, 1.89, -0.091, 14.93, -2.34, -0.074, -18.7, -1.09, 6.5$

Fitted values:

$46.11, 56, 43.96, 31.91, 46.11, 3.09, 37.07, -7.66, 37.07, 23.74, 43.1, -2.5$

(b) *Verify that $\hat{u}$ has zero correlation with x. Draw the scatter plot of the two variables.*



Correlation coefficient is zero.

(c) *Consider the prediction equation*

$$\breve{y} = 5 + 0.6x$$

Show that the prediction errors in this case are correlated with x, and also have a larger mean squared error than the regression predictions.

The predictions are

$56.6, 70.4, 53.6, 36.8, 56.6, -3.4, 44, -18.4, 44, 25.4, 52.4, -11.2$

Correlation with x is -0.62. MSE is 194.47. MSE of the least squares residuals is 108.45

(d) *Compute the regression of y on x (extended sample). Report the fitted coefficients and comment*

$$\hat{y} = 19.037 + 0.167x.$$

Comment: large shift in the coefficients shows that extreme observations are influential in the least squares fit. See scatter plot - a single observation makes the difference.

5.

(a) *Solve the following equation system for $\hat{\beta}_1$.*

$$\hat{\beta}_1 \sum_{t=1}^{T} x_{1t}^2 + \hat{\beta}_2 \sum_{t=1}^{T} x_{1t}x_{2t} = \sum_{t=1}^{T} x_{1t}y_t \qquad (1)$$

$$\hat{\beta}_1 \sum_{t=1}^{T} x_{2t}x_{1t} + \hat{\beta}_2 \sum_{t=1}^{T} x_{2t}^2 = \sum_{t=1}^{T} x_{2t}y_t \qquad (2)$$

First step: get $\hat{\beta}_2$ a function of $\hat{\beta}_1$ in equation (2)

$$\hat{\beta}_2 = \frac{\sum_{t=1}^{T} x_{2t}y_t - \hat{\beta}_1 \sum_{t=1}^{T} x_{2t}x_{1t}}{\sum_{t=1}^{T} x_{2t}^2}$$

Second step. Substitute into equation (1)

$$\hat{\beta}_1 \sum_{t=1}^{T} x_{1t}^2 \sum_{t=1}^{T} x_{2t}^2 + \sum_{t=1}^{T} x_{2t}y_t \sum_{t=1}^{T} x_{1t}x_{2t} - \hat{\beta}_1 \sum_{t=1}^{T} x_{2t}x_{1t} \sum_{t=1}^{T} x_{1t}x_{2t} = \sum_{t=1}^{T} x_{1t}y_t \sum_{t=1}^{T} x_{2t}^2$$

Third step. Solve for $\hat{\beta}_1$

$$\hat{\beta}_1 = \frac{\sum_{t=1}^{T} x_{1t}y_t \sum_{t=1}^{T} x_{2t}^2 - \sum_{t=1}^{T} x_{2t}y_t \sum_{t=1}^{T} x_{1t}x_{2t}}{\sum_{t=1}^{T} x_{1t}^2 \sum_{t=1}^{T} x_{2t}^2 - \left(\sum_{t=1}^{T} x_{2t}x_{1t}\right)^2}.$$

(b) *Define $\hat{\alpha} = \bar{y} - \hat{\beta}_1\bar{x}_1 - \hat{\beta}_2\bar{x}_2$. Show that if $T\hat{\alpha}\bar{x}_1$ is subtracted from the right-hand side of equation (1), and $T\hat{\alpha}\bar{x}_2$ is subtracted from the right-hand side of equation (2), the resulting equations are modified by having the variables expressed in mean deviation form.*

$$\hat{\beta}_1 \sum_{t=1}^{T} x_{1t}^2 + \hat{\beta}_2 \sum_{t=1}^{T} x_{1t}x_{2t} = \sum_{t=1}^{T} x_{1t}y_t - T(\bar{y} - \hat{\beta}_1\bar{x}_1 - \hat{\beta}_2\bar{x}_2)\bar{x}_1$$

$$\hat{\beta}_1 \sum_{t=1}^{T} x_{2t}x_{1t} + \hat{\beta}_2 \sum_{t=1}^{T} x_{2t}^2 = \sum_{t=1}^{T} x_{2t}y_t - T(\bar{y} - \hat{\beta}_1\bar{x}_1 - \hat{\beta}_2\bar{x}_2)\bar{x}_2$$

are the same as

$$\hat{\beta}_1\left(\sum_{t=1}^{T} x_{1t}^2 - T\bar{x}_1^2\right) + \hat{\beta}_2\left(\sum_{t=1}^{T} x_{1t}x_{2t} - T\bar{x}_1\bar{x}_2\right) = \left(\sum_{t=1}^{T} x_{1t}y_t - T\bar{y}\bar{x}_1\right)$$
$$\hat{\beta}_1\left(\sum_{t=1}^{T} x_{2t}x_{1t} - T\bar{x}_1\bar{x}_2\right) + \hat{\beta}_2\left(\sum_{t=1}^{T} x_{2t}^2 - T\bar{x}_2^2\right) = \left(\sum_{t=1}^{T} x_{2t}y_t - T\bar{y}\bar{x}_2\right)$$

(c) *What is* $\hat{\alpha}$*?*

Write the equations as

$$\hat{\beta}_1\sum_{t=1}^{T} x_{1t}^2 + \hat{\beta}_2\sum_{t=1}^{T} x_{1t}x_{2t} + \hat{\alpha}\sum_{t=1}^{T} x_{1t} = \sum_{t=1}^{T} x_{1t}y_t$$
$$\hat{\beta}_1\sum_{t=1}^{T} x_{2t}x_{1t} + \hat{\beta}_2\sum_{t=1}^{T} x_{2t}^2 + \hat{\alpha}\sum_{t=1}^{T} x_{2t} = \sum_{t=1}^{T} x_{2t}y_t$$
$$\hat{\beta}_1\sum_{t=1}^{T} x_{1t} + \hat{\beta}_2\sum_{t=1}^{T} x_{2t} + T\hat{\alpha} = \sum_{t=1}^{T} y_t$$

and these are the normal equations with intercept, with the solution for $\hat{\alpha}$ indicated.

6. *Show that the* α *that minimizes* $\sum_{t=1}^{T}(y_t-\alpha)^2$ *is the sample mean of* $y_1, \ldots, y_T$.

Write

$$\begin{aligned}\sum_{t=1}^{T}(y_t-\alpha)^2 &= \sum_{t=1}^{T}(y_t-\bar{y}+\bar{y}-\alpha)^2\\ &= \sum_{t=1}^{T}(y_t-\bar{y})^2 + T(\bar{y}-\alpha)^2 + 2(\bar{y}-\alpha)\sum_{t=1}^{T}(y_t-\bar{y})\\ &\geq \sum_{t=1}^{T}(y_t-\bar{y})^2\end{aligned}$$

because

$$\sum_{t=1}^{T}(y_t-\bar{y}) = 0.$$