Econometric Modeling

Solutions to Exercises with Even Numbers

David F. Hendry
Bent Nielsen

24 January 2008 - preliminary version

Chapter One

The Bernoulli model

Solution 1.2.

(a) Follow the arguments in the §1.3.2 to derive the likelihood, with parameter space $\Theta_{03} = (0,1)$. First, find the joint density:

$$\mathsf{f}_{\theta_{03}}(x_1, \dots x_{n_{03}}) = \cdots = \left\{ \theta_{03}^{\overline{x}} \, (1-\theta_{03})^{1-\overline{x}} \right\}^{n_{03}},$$

and then the likelihood. The maximum likelihood estimate for the chance of a girl in 2003 is then:

$$\widehat{\theta}_{03} = \overline{X} = \frac{338971}{356578 + 338971} = 0.4873.$$

(b) From (a) we have the joint density for the 2003-data, whereas the joint density for the 2004-data is given in the text. The parameter is now bivariate, $\theta_{03}, \ \theta_{04}$, with parameter space $\Theta_{joint} = (0,1) \times (0,1) = (0,1)^2$. Due to the independence assumption the joint density for all data is the product:

$$\begin{aligned} \mathsf{f}_{\theta_{03},\theta_{04}} & (x_1, \dots, x_{n_{03}}, y_1, \dots, y_{n_{04}}) && \text{(S.1.1)} \\ &= \mathsf{f}_{\theta_{03}}(x_1, \dots, x_{n_{03}}) \mathsf{f}_{\theta_{04}}(y_1, \dots, y_{n_{04}}) \\ &= \left\{ \theta_{03}^{\overline{x}} \, (1-\theta_{03})^{1-\overline{x}} \right\}^{n_{03}} \left\{ \theta_{04}^{\overline{y}} \, (1-\theta_{04})^{1-\overline{y}} \right\}^{n_{04}}. && \text{(S.1.2)} \end{aligned}$$

The likelihood function is found by looking at this as a function of the two parameters $\theta_{03}, \ \theta_{04}$. This likelihood has a product structure with one term involving θ_{03} and another term involving θ_{04}. As there are no cross-restrictions between these parameters each term can be maximized separately giving:

$$\begin{aligned} \widehat{\theta}_{03} &= \overline{X} = \frac{338971}{356578 + 338971} = 0.4873, \\ \widehat{\theta}_{04} &= \overline{X} = \frac{348410}{367586 + 348410} = 0.4866, \end{aligned}$$

since, for instance:

$$\frac{\partial}{\partial \theta_{03}} \ell_{X_1,\dots,X_{n_{03}},Y_1,\dots,Y_{n_{98}}}(\theta_{03}, \theta_{04}) = \frac{\partial}{\partial \theta_{03}} \ell_{X_1,\dots,X_{n_{03}}}(\theta_{03}).$$

Solution 1.4. Make a coordinate system and plot the four points.

(a) Use (1.2.3) so:

$$\begin{aligned}\mathsf{P}(X=-1) &= \mathsf{P}(X=-1, Y=0) = \frac{1}{4},\\ \mathsf{P}(X=0) &= \mathsf{P}(X=0, Y=-1) + \mathsf{P}(X=0, Y=1) = \frac{1}{2},\\ \mathsf{P}(X=1) &= \mathsf{P}(X=1, Y=0) = \frac{1}{4}.\end{aligned}$$

(b) No. Note that Y has the same marginal distribution as X. Then it holds, for instance that:

$$\mathsf{P}(X=1, Y=0) = \frac{1}{4} \neq \mathsf{P}(X=1)\mathsf{P}(Y=0) = \left(\frac{1}{4}\right)\left(\frac{1}{2}\right) = \frac{1}{8}.$$

In order to have independence it is necessary that the two variables vary in a product space. Here the sample space is like a rhomb rather than a square.

Chapter Two

Inference in the Bernoulli model

Solution 2.2. These data can be analyzed using the Bernoulli model. Thus, the chance of a newborn child being female is estimated by:

$$\widehat{\theta} = \frac{1964}{1964 + 2058} = 0.4883$$

As the data in Table 1.6 are measured in thousands, then:

$$n = (1964 + 2058) \times 1000 = 4022000.$$

The fact that the reported data are in thousands simply results in a rounding error of up to 1000, corresponding to a relative error of 0.00025. Thus, in reporting the estimate for θ the rounding error kicks in on the fourth digit after the decimal point.

The standard error of the estimate for θ is:

$$\widehat{\mathsf{se}}(\widehat{\theta}) = \frac{\sqrt{\widehat{\theta}(1-\widehat{\theta})}}{\sqrt{n}} = \frac{\sqrt{0.4883(1-0.4883)}}{\sqrt{4022000}} = 0.000249,$$

Since the 99.5%-quantile of the standard normal distribution is 2.576 (see Table 2.1) this gives a 99%-confidence interval of:

$$0.4883 - 2.576 \times 0.000249 \leq \theta \leq 0.4883 + 2.576 \times 0.000249,$$

or:

$$0.4877 \leq \theta \leq 0.4889.$$

This confidence interval overlaps with the confidence interval for the UK reported in §2.3.1, indicating that sex ratios are not all that different between countries. A formal test could be made as set out in Exercise 2.3.

Solution 2.4. Suppose $X \overset{\mathsf{D}}{=} \mathsf{Bernoulli}[p]$. Then:

$$\begin{aligned}\mathsf{E}(X) &= 0 \cdot (1-p) + 1 \cdot p = p,\\ \mathsf{Var}(X) &= (0-p)^2 \cdot (1-p) + (1-p)^2 \cdot p\\ &= p^2(1-p) + p(1-p) = p(1-p).\end{aligned}$$

The variance can also be found from the general formula:

$$\mathsf{E}\{X - \mathsf{E}(X)\}^2 = \mathsf{E}[X^2 - 2X\mathsf{E}(X) + \{\mathsf{E}(X)\}^2] = \mathsf{E}\left(X^2\right) - \{\mathsf{E}(X)\}^2.$$

using that for the Bernoulli case:

$$\mathsf{E}\left(X^2\right) = 0^2 \cdot (1-p) + 1^2 \cdot p = p.$$

Choosing a success probability p the density and distribution functions are:

$$\mathsf{f}(x) = \begin{cases} 1-p & \text{for } x = 0,\\ p & \text{for } x = 1,\\ 0 & \text{otherwise}, \end{cases}$$
$$\mathsf{F}(x) = \begin{cases} 0 & \text{for } x < 0,\\ 1-p & \text{for } 0 \le x < 1,\\ 1 & \text{for } 1 \le x. \end{cases}$$

Now, the "confidence intervals" are, for $k = 1, 2$:

$$\mathsf{E}(X) \pm k\mathsf{sdv}(X) = p \pm k\sqrt{p(1-p)}.$$

Examples:
Let $p = 1/2$ then $\mathsf{E}(X) = 1/2$, $\mathsf{sdv}(X) = 1/2$ so $\mathsf{E}(X) \pm \mathsf{sdv}(X)$ has probability 0 (if open interval) while $\mathsf{E}(X) \pm 2\mathsf{sdv}(X)$ has probability 1.
let $p = 1/5$ then $\mathsf{E}(X) = 1/5$, $\mathsf{sdv}(X) = 2/5$ so $\mathsf{E}(X) \pm \mathsf{sdv}(X)$ has probability 4/5 while $\mathsf{E}(X) \pm 2\mathsf{sdv}(X)$ has probability 1.

Solution 2.6. Use that expectations are linear operators, and that the expectation, $\mathsf{E}(X)$, is deterministic:

$$\begin{aligned}&\mathsf{E}\{X - \mathsf{E}(X)\}^2\\ &= [\text{ complete square }] = \mathsf{E}[X^2 - 2X\mathsf{E}(X) + \{\mathsf{E}(X)\}^2]\\ &= [\text{ linearity }] = \mathsf{E}(X^2) - 2\{\mathsf{E}(X)\}^2 + \{\mathsf{E}(X)\}^2 = \mathsf{E}(X^2) - \{\mathsf{E}(X)\}^2.\end{aligned}$$

Solution 2.8. The key to the solution is (2.1.8) stating that:

$$X, Y \text{ independent} \qquad \Rightarrow \qquad \mathsf{E}(XY) = \mathsf{E}(X)\mathsf{E}(Y).$$

In particular, for constants μ_x, μ_y:

$$\begin{aligned}
&X, Y \text{ independent} \\
\Leftrightarrow\ &(X - \mu_x), (Y - \mu_y) \text{ independent} \\
\Rightarrow\ &\mathsf{E}\{(X - \mu_x)(Y - \mu_y)\} = \mathsf{E}(X - \mu_x)\mathsf{E}(Y - \mu_y).
\end{aligned}$$

Thus, the problem can be solved as follows:

$$\begin{aligned}
&\mathsf{Cov}(X, Y) \\
&= [\text{ Definition (2.1.12)}] = \mathsf{E}[\{X - \mathsf{E}(X)\}\{Y - \mathsf{E}(Y)\}] \\
&= [\text{ use independence and (2.1.8)}] = \mathsf{E}\{X - \mathsf{E}(X)\}\mathsf{E}\{Y - \mathsf{E}(Y)\} \\
&= [\text{ use } \mathsf{E}\{X - \mathsf{E}(X)\} = \mathsf{E}(X) - \mathsf{E}(X) = 0\} \text{]} = 0.
\end{aligned}$$

Solution 2.10.

(a) This is a straight forward application of the Law of Large Numbers. For instance, the first subsample consists of $n/2$ independently identical distributed variables. Thus, their sample average, here denoted Z_1 converges in probability to their population expectation, $\theta_1 = 1/4$.

(b) Here the definition of convergence in probability as stated in Theorem 2.1 has to be used. Thus, it has to be shown that for any $\delta > 0$ the probability that:

$$\left| \overline{Y} - \frac{1}{2} \right| > \delta$$

converges to zero. Now, decompose:

$$\begin{aligned}
\overline{Y} &= \frac{1}{n}\sum_{i=1}^{n} Y_i = \frac{1}{n}\sum_{i=1}^{n/2} Y_i + \frac{1}{n}\sum_{i=n/2+1}^{n} Y_i \\
&= \left(\frac{1}{2}\right)\frac{2}{n}\sum_{i=1}^{n/2} Y_i + \left(\frac{1}{2}\right)\frac{2}{n}\sum_{i=n/2+1}^{n} Y_i = \frac{Z_1}{2} + \frac{Z_1}{2}.
\end{aligned}$$

Likewise, decompose $1/2 = (1/2)(1/4) + (1/2)(3/4)$, so:

$$\overline{Y} - \frac{1}{2} = \frac{1}{2}(Z_1 - \theta_1) + \frac{1}{2}(Z_2 - \theta_2).$$

In particular, by the triangle inequality:

$$\left| \overline{Y} - \frac{1}{2} \right| \le \frac{1}{2}|Z_1 - \theta_1| + \frac{1}{2}|Z_2 - \theta_2|.$$

Since $Z_1,\ Z_2$ both converge in probability, then for any δ it holds that:

$$\begin{aligned}\mathsf{P}\left(\left|\overline{Y}-\frac{1}{2}\right|>\delta\right) &\leq \mathsf{P}\left(\frac{1}{2}\left|Z_1-\theta_1\right|+\frac{1}{2}\left|Z_2-\theta_2\right|>\delta\right)\\ &\leq [\text{ check }]\\ &\leq \mathsf{P}\left(\frac{1}{2}\left|Z_1-\theta_1\right|>\frac{\delta}{2}\text{ or }\frac{1}{2}\left|Z_2-\theta_2\right|>\frac{\delta}{2}\right)\\ &= \mathsf{P}\left(\left|Z_1-\theta_1\right|>\delta\text{ or }\left|Z_2-\theta_2\right|>\delta\right).\end{aligned}$$

Now, the so-called Boole's inequality is needed. This states that for any events $A,\ B$ then:

$$\mathsf{P}(A\text{ or }B)\leq\mathsf{P}(A)+\mathsf{P}(B). \tag{S.2.1}$$

This comes about as follows. If A and B are disjoint, that is, they cannot both be true, then (S.2.1) holds with equality. This property was used when linking distribution functions and densities in §1.2.2. If A and B are not disjoint, split the sets into three disjoint parts: (A but not B), (B but not A), (A and B). Then the left hand side of (S.2.1) includes these sets once each, whereas the right hand side includes the intersection set (A and B) twice, hence the inequality.

Applying Boole's inequality it then holds:

$$\mathsf{P}\left(\left|\overline{Y}-\frac{1}{2}\right|>\delta\right)\leq\mathsf{P}\left(\left|Z_1-\theta_1\right|>\delta\right)+\mathsf{P}\left(\left|Z_2-\theta_2\right|>\delta\right).$$

Due to (a) both of the probabilities on the right hand side converge to zero.

(c) The Law of Large Numbers requires that the random variables are independent, which is satisfied here, as well as having the same distribution, which is not satisfied here. The theorem therefore does not apply to the full set of observations in this case. As all inference in a given statistical model is based on the the Law of Large Numbers and the Central Limit Theorem, it is therefore important to check model assumptions. The inference may be valid under milder assumptions, but this is not always the case.

Chapter Three

A first regression model

Solution 3.2. Recall $\widehat{\beta}=5.02$ and $\widehat{\sigma}=0.753$.

$$\begin{aligned} n&=100 && 4.86<\beta<5.17,\\ n&=10000 && 5.00<\beta<5.04.\end{aligned}$$

The widths of the confidence intervals vary with $\sqrt{n}$.

Solution 3.4.
(a) The first order derivative is given in (3.3.8). Thus the second derivative is:

$$\frac{\partial^2}{\partial\left(\sigma^2\right)^2}\ell_{Y_1,\ldots,Y_n}\left(\widehat{\beta},\sigma^2\right)=\frac{n}{2\sigma^4}-\frac{1}{\sigma^6}\sum_{i=1}^{n}\widehat{u}_i^2=\frac{n}{2\sigma^6}\left(\sigma^2-2\widehat{\sigma}^2\right),$$

giving that the second order condition is satisfied at $\widehat{\sigma}^2$ since:

$$\frac{n}{2\widehat{\sigma}^6}\left(\widehat{\sigma}^2-2\widehat{\sigma}^2\right)=-\frac{n}{2\widehat{\sigma}^6}\widehat{\sigma}^2=-\frac{n}{2\widehat{\sigma}^4}<0.$$

Therefore the profile likelihood is locally concave. It is not, however, globally concave, since the second derive is positive for $\sigma^2>2\widehat{\sigma}^2$. To show uniqueness of the maximum likelihood estimator, global concavity is not needed, however. Since $\partial\ell/\partial\sigma^2$ is positive for $\sigma^2<\widehat{\sigma}^2$ and negative for $\sigma^2>\widehat{\sigma}^2$ then $\widehat{\sigma}^2$ is a global maximum.

Note: This argument shows that there is a unique *mode* for the likelihood.

(b) The profile log-likelihood for $\theta=\log\sigma$ is found by reparametrizing (3.3.7):

$$\ell_{Y_1,\ldots Y_n}(\widehat{\beta},\log\sigma)=-n\log\sigma-\sum_{i=1}^{n}\widehat{u}_i^2/(2\sigma^2).$$

Thus the profile log-likelihood for θ is:

$$\ell_{Y_1,\ldots Y_n}(\widehat{\beta},\theta)=-n\theta-\exp(-2\theta)\sum_{i=1}^{n}\widehat{u}_i^2/2,$$

with derivatives:

$$\partial\ell/\partial\theta=-n+\exp(-2\theta)\sum_{i=1}^{n}\widehat{u}_i^2,$$
$$\partial^2\ell/\partial\theta^2=-2\exp(-2\theta)\sum_{i=1}^{n}\widehat{u}_i^2.$$

The second derivative is negative for all θ. Thus the profile log-likelihood for θ is concave with a unique maximum.

Note: This argument shows that with the $\theta=\log\sigma$ parametrization the likelihood is concave. In particular, there is a unique mode. Concavity is not invariant to reparametrization, whereas uniqueness of the mode is invariant.

Solution 3.6. We know that if $X\overset{\mathsf{D}}{=}\mathsf{N}[\mu,\sigma^2]$ then $\mathsf{E}\,(X)=\mu$, $\mathsf{Var}\,(X)=\sigma^2$ so by (2.1.6) and (2.1.10) we have $\mathsf{E}\,(aX+b)=a\mu+b$ and $\mathsf{Var}(aX+b)=a^2\sigma^2$.

Solution 3.8. Use Table 2.1.
(a) Due to the symmetry then $\mathsf{P}(X\leq 0)=1/2$. Since $\mathsf{P}(X>2.58)=0.005$ then the lower quantile satisfies $\mathsf{P}(0<X\leq 2.58)=0.995$.

(b) Note that $X = (Y-4)/\sqrt{4} = (Y-4)/2 \overset{\mathsf{D}}{=} \mathsf{N}[0,1]$. Therefore:

$$\mathsf{P}(Y<0) = \mathsf{P}\left(\frac{Y-4}{2} < \frac{0-4}{2}\right) = \mathsf{P}(X<-2)$$

Now, using that the normal distribution is symmetric it follows that:

$$\mathsf{P}(Y<0) = \mathsf{P}(X<-2) = \mathsf{P}(X>2) \approx 0.025$$

Finally, we want to find a quantile q so $\mathsf{P}(Y>q) = 0.025$. The standard normal table gives $\mathsf{P}(X>2) = 1.96$. As $Y = 2X+4$ the 97.5%-quantile for Y is then $2 \times 1.96 + 4 = 7.92$.

Solution 3.10. We have $\widehat{\beta} = \sum_{i=1}^n Y_i x_i / (\sum_{i=1}^n x_i^2)$ where $x_1, \ldots x_n$ are non-random constants. The only source of randomness is therefore the Y_is. Thus $\widehat{\beta}$ is a linear combination of independent normal distributed variables and must be normally distributed, see (3.4.5). Therefore we only need to find the expectation and variance of $\widehat{\beta}$ using the formulas for the expectation and variance of sums, (2.1.7), (2.1.11). First the expectation:

$$\begin{aligned}
\mathsf{E}(\widehat{\beta}_1) &= [\text{ (2.1.7) }] = \frac{\sum_{i=1}^n x_i \mathsf{E}\,(Y_i)}{\sum_{i=1}^n x_i^2} \\
&= [\text{ model : } Y_i \overset{\mathsf{D}}{=} \mathsf{N}[\beta x_i, \sigma^2] \text{ so } \mathsf{E}(Y_i) = \beta x_i \text{ }] \\
&= \frac{\sum_{i=1}^n x_i \beta x_i}{\sum_{i=1}^n x_i^2} = \beta,
\end{aligned}$$

and similarly the variance:

$$\begin{aligned}
\mathsf{Var}(\widehat{\beta}_1) &= [\text{ (2.1.14) using independence }] = \frac{\sum_{i=1}^n x_i^2 \mathsf{Var}\,(Y_i)}{(\sum_{i=1}^n x_i^2)^2} \\
&= [\text{ model : } Y_i \overset{\mathsf{D}}{=} \mathsf{N}[\beta x_i, \sigma^2] \text{ so } \mathsf{Var}(Y_i) = \sigma^2 \text{ }] \\
&= \frac{\sum_{i=1}^n x_i^2 \sigma^2}{(\sum_{i=1}^n x_i^2)^2} = \frac{\sigma^2}{\sum_{i=1}^n x_i^2}.
\end{aligned}$$

Solution 3.12.

(a) The joint density is:

$$\begin{aligned}
\mathsf{f}_\lambda(y_1, \ldots, y_n) &= [\text{ independence, use (3.2.1) }] = \prod_{i=1}^n \mathsf{f}_\lambda(y_i) \\
&= [\text{ Poisson distribution }] = \prod_{i=1}^n \frac{\lambda^{y_i}}{y_i!} \exp(-\lambda).
\end{aligned}$$

Using the functional equation for powers, see (1.3.1), this can be rewritten as:

$$\mathsf{f}_\lambda(y_1, \ldots, y_n) = \lambda^{\sum_{i=1}^n y_i} \exp(-n\lambda) \prod_{i=1}^n \frac{1}{y_i!} = \left\{\lambda^{\overline{y}} \exp(-\lambda)\right\}^n \prod_{i=1}^n \frac{1}{y_i!},$$

where $\overline{y} = n^{-1}\sum_{i=1}^{n} y_i$ is the average of $y_1, \ldots, y_n$. The likelihood function is then:

$$\mathsf{L}_{Y_1,\ldots,Y_n}(\lambda) = \left\{\lambda^{\overline{Y}} \exp(-\lambda)\right\}^n \prod_{i=1}^{n} \frac{1}{Y_i!},$$

with the corresponding log-likelihood function:

$$\begin{aligned}\ell_{Y_1,\ldots,Y_n}(\lambda) &= n \log\left\{\lambda^{\overline{Y}} \exp(-\lambda)\right\} + \log\left(\prod_{i=1}^{n} \frac{1}{Y_i!}\right) \\ &= [\text{ functional equation for logarithm, (1.3.5) }] \\ &= n\left\{\overline{Y}\log(\lambda) - \lambda\right\} - \sum_{i=1}^{n} \log(Y_i!).\end{aligned}$$

(b) The maximum likelihood equation is found by differentiating the log-likelihood function ℓ:

$$\frac{\partial}{\partial\lambda}\ell_{Y_1,\ldots,Y_n}(\lambda) = n\left(\frac{\overline{Y}}{\lambda} - 1\right),$$

giving the likelihood equation:

$$n\left(\frac{\overline{Y}}{\widehat{\lambda}} - 1\right) = 0,$$

which is solved by $\widehat{\lambda} = \overline{Y}$. To show that the solution is unique note *either* that the first derivative of ℓ is positive for $\lambda < \widehat{\lambda}$ and negative for $\lambda > \widehat{\lambda}$ *or* find the second derivative:

$$\frac{\partial^2}{\partial\lambda^2}\ell_{Y_1,\ldots,Y_n}(\lambda) = -n\frac{\overline{Y}}{\lambda^2},$$

which is negative for all values of λ so the log-likelihood function is concave. Note, that these arguments only apply for $\overline{Y} > 0$. If, however, $\overline{Y} = 0$ there is no variation to model.

Chapter Four

The logit model

Solution 4.2. Use (4.2.4) and (4.2.5) to see that:

$$\begin{aligned}\mathsf{P}(Y_i = 0|X_i) &= [\text{ (4.2.5) }] = \mathsf{P}(Y_i^* < 0|X_i) \\ &= [\text{ (4.2.4) }] = \mathsf{P}(\beta_1 + \beta_2 X_i + u_i < 0|X_i) \\ &= \mathsf{P}(u_i < -\beta_1 - \beta_2 X_i|X_i).\end{aligned}$$

Using the definition of the logistic distribution, which is continuous, we get:

$$\mathsf{P}(Y_i = 0|X_i) = \frac{\exp(-\beta_1 - \beta_2 X_i)}{1 + \exp(-\beta_1 - \beta_2 X_i)}.$$

Extend the fraction by $\exp(\beta_1 + \beta_2 X_i)$ and use the functional equation for powers, see (1.3.1), to get:

$$\mathsf{P}(Y_i = 0|X_i) = \frac{1}{\exp(\beta_1 + \beta_2 X_i) + 1}.$$

In particular:

$$\mathsf{P}(Y_i = 1|X_i) = 1 - \mathsf{P}(Y_i = 0|X_i) = \frac{\exp(\beta_1 + \beta_2 X_i)}{1 + \exp(\beta_1 + \beta_2 X_i)}$$

as desired.

Solution 4.4. Without loss of generality we can assume that the observations are ordered so the first m regressors X_i are all 1. Then (4.2.8), (4.2.9) reduce as:

$$m\frac{\exp\left(\widehat{\beta}_1 + \widehat{\beta}_2\right)}{1 + \exp\left(\widehat{\beta}_1 + \widehat{\beta}_2\right)} + (n-m)\frac{\exp\left(\widehat{\beta}_1\right)}{1 + \exp\left(\widehat{\beta}_1\right)} = \sum_{i=1}^{n} Y_i, \tag{S.4.1}$$

$$m\frac{\exp\left(\widehat{\beta}_1 + \widehat{\beta}_2\right)}{1 + \exp\left(\widehat{\beta}_1 + \widehat{\beta}_2\right)} = \sum_{i=1}^{m} Y_i. \tag{S.4.2}$$

Subtracting equation (S.4.2) from (S.4.1) shows :

$$(n-m)\frac{\exp\left(\widehat{\beta}_1\right)}{1 + \exp\left(\widehat{\beta}_1\right)} = \sum_{i=m+1}^{n} Y_i.$$

Since the function $y = g(x) = \exp(x)/\{1 + \exp(x)\}$ has the inverse $x = g^{-1}(y) = \log\{y/(1-y)\}$ this implies:

$$\widehat{\beta}_1 = \log \frac{(n-m)^{-1} \sum_{i=m+1}^{n} Y_i}{1 - (n-m)^{-1} \sum_{i=m+1}^{n} Y_i}.$$

Correspondingly, from (S.4.2):

$$\widehat{\beta_1 + \beta_2} = \log \frac{m^{-1} \sum_{i=1}^{m} Y_i}{1 - m^{-1} \sum_{i=1}^{m} Y_i},$$

and thus:

$$\widehat{\beta}_2 = \log \frac{m^{-1} \sum_{i=1}^{m} Y_i}{1 - m^{-1} \sum_{i=1}^{m} Y_i} - \log \frac{(n-m)^{-1} \sum_{i=m+1}^{n} Y_i}{1 - (n-m)^{-1} \sum_{i=m+1}^{n} Y_i}.$$

Solution 4.6.

(a) Inspect that inserting the expression for $p_i = p(X_i)$ in (4.2.6), (4.2.7) gives the desired answer.

(b) This follows by noting that:

$$\mathsf{E}(Y_i \mid \mathcal{I}) = p_i, \qquad \mathsf{E}(Y_i X_i \mid \mathcal{I}) = p_i X_i,$$

(c) The key to the result is that:

$$\frac{\partial}{\partial \beta_1} p_i = p_i(1-p_i), \qquad \frac{\partial}{\partial \beta_1} p_i = p_i(1-p_i)X_i,$$

Apply this to the first derivatives given in (a).

(d) Start with the element $\frac{\partial^2}{\partial \beta_1^2}\ell(\beta_1, \beta_2)$. Since $0 < p_i < 1$ the diagonal terms this equals minus one times a sums of positive elements. When it comes to $\frac{\partial^2}{\partial \beta_2^2}\ell(\beta_1, \beta_2)$ the summands are zero if $X_i = 0$. Provided the X_is are not all zero the overall sum is positive.

(e) By the Cauchy-Schwarz inequality:

$$\begin{aligned}\left\{\sum_{i=1}^n p_i(1-p_i)X_i\right\}^2 &= \left\{\sum_{i=1}^n \sqrt{p_i(1-p_i)}\sqrt{p_i(1-p_i)X_i}\right\}^2 \\ &\leq \left\{\sum_{i=1}^n p_i(1-p_i)\right\}\left\{\sum_{i=1}^n p_i(1-p_i)X_i\right\},\end{aligned}$$

with identity only if $p_i(1-p_i) = p_i(1-p_i)X_i$ for all i, that is if $X_i = 1$ for all i.

Chapter Five

The two-variable regression model

Solution 5.2. On the one hand, if $(Y_i|X_i) \overset{\mathsf{D}}{=} \mathsf{N}[\beta_1 + \beta_2 X_i, \sigma^2]$ is the correct model with $\beta_1 \neq 0$, it must be a bad idea to use the simpler model! We will return to that issue in connection with omitted variable bias. On the other hand if the simple model is correct this must be better to use. We will return to that issue in connection with efficiency. In practice it is extremely rare to have models with regressors, but no intercept. When it comes to time series models without intercepts are often used in expositions, but rarely in practice

If it is deemed appropriate to omit the intercept a complication one has to be careful with using the standard output of econometric and statistical software. While the estimators and t-statistics are fine, a few statistics are not applicable in the usual way for models without an intercept. This applies for instance to the R^2-statistic and to Q-Q plots (see Engler and Nielsen, 2007 for the latter).

Solution 5.4. From (5.2.1) we have:

$$\widehat{\beta}_1 = \overline{Y} - \widehat{\beta}_2 \overline{X}, \qquad \widehat{\beta}_2 = \sum_{i=1}^{n} \frac{Y_i \left(X_i - \overline{X} \right)}{\sum_{j=1}^{n} \left(X_j - \overline{X} \right)^2}.$$

Therefore:

$$\widehat{\beta}_1 = \overline{Y} - \overline{X} \frac{\sum_{i=1}^{n} Y_i \left(X_i - \overline{X} \right)}{\sum_{j=1}^{n} \left(X_j - \overline{X} \right)^2} = \sum_{i=1}^{n} w_i Y_i,$$

where

$$w_i = \frac{1}{n} - \frac{\overline{X}}{\sum_{j=1}^{n} (X_j - \overline{X})^2} (X_i - \overline{X}).$$

It has to be argued that $w_i = X_{1\cdot 2,i} / \sum_{j=1}^{n} X_{1\cdot 2,j}^2$. Thus, from (5.2.12) find:

$$\sum_{j=1}^{n} X_{1\cdot 2,j}^2 = n - 2 \frac{\left(\sum_{j=1}^{n} X_j \right)^2}{\sum_{j=1}^{n} X_j^2} + \frac{\left(\sum_{j=1}^{n} X_j \right)^2}{\sum_{j=1}^{n} X_j^2} = n - \frac{\left(\sum_{j=1}^{n} X_j \right)^2}{\sum_{j=1}^{n} X_j^2},$$

and find a common denominator for w_i so $w_i = N_i / D$ where

$$N_i = \sum_{j=1}^{n} (X_j - \overline{X})^2 - n\overline{X}(X_i - \overline{X}), \qquad D_i = n \sum_{j=1}^{n} (X_j - \overline{X})^2.$$

Finish, by noting that:

$$\begin{aligned} D_i &= n \left(\sum_{j=1}^{n} X_j^2 - n\overline{X}^2 \right) = \sum_{j=1}^{n} X_j^2 \sum_{j=1}^{n} X_{1\cdot 2,j}^2 \\ N_i &= \sum_{j=1}^{n} X_j^2 - n\overline{X}^2 - n\overline{X} X_i + n\overline{X}^2 = \sum_{j=1}^{n} X_j^2 - n\overline{X} X_i \\ &= \sum_{j=1}^{n} X_j^2 \left(1 - \frac{\sum_{j=1}^{n} X_j}{\sum_{j=1}^{n} X_j^2} X_i \right) = \left(\sum_{j=1}^{n} X_j^2 \right) X_{1\cdot 2,i}. \end{aligned}$$

Solution 5.6.
(a) See Figure S.5.1.
(b) The following statistics have sufficient information about the data:

$$\sum_{t=1}^{T} Y_t = 71.7424, \qquad \sum_{t=1}^{T} Y_t^2 = 756.3848, \qquad T = 7,$$

$$\sum_{t=1}^{T} t = 28, \qquad \sum_{t=1}^{T} t^2 = 140, \qquad \sum_{t=1}^{T} t Y_t = 311.2105.$$

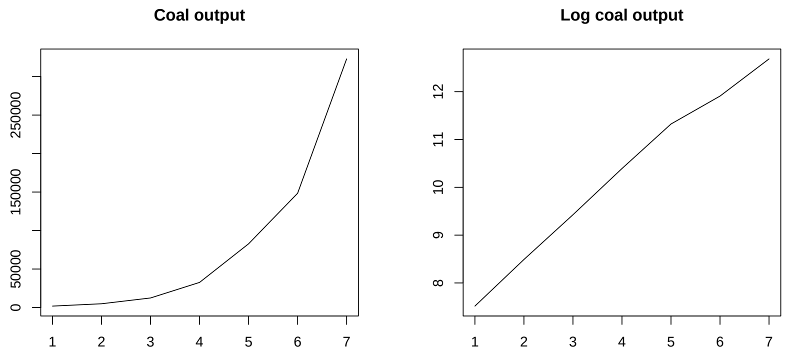

Figure S.5.1 Industrial output by period

It follows that:

$$\widehat{\beta}_1 = 6.79, \qquad \widehat{\beta}_2 = 0.866,$$
$$\widehat{\sigma}^2 = (0.129)^2, \qquad s^2 = (0.153)^2$$
$$\widehat{\delta}_1 = \widehat{\beta}_1 + \widehat{\beta}_2\overline{X} = 10.25, \qquad \widehat{\delta}_2 = \widehat{\beta}_2.$$

The prediction for the log-output in the period $1975 - 1999$ is:

$$\widehat{Y}_i = \widehat{\beta}_1 + \widehat{\beta}_2 8 = \widehat{\delta}_1 + \widehat{\delta}_2(8-4) = 13.72.$$

Solution 5.8. Model equation $Y_i = \beta_1 + \beta_2 X_i + u_i$ with $\beta_2 = 0$ in data generating process. The estimators in question are

$$\widehat{\beta} = \overline{Y}, \qquad \widehat{\beta}_1 = \frac{\sum_{i=1}^n X_{1\cdot 2,i} Y_i}{\sum_{j=1}^n X_{1\cdot 2,j}^2}$$

(a) For $\widehat{\beta}$ the result can be taken from §3.4.1, that is $\mathsf{E}(\widehat{\beta}) = \beta_1$.
For $\widehat{\beta}_1$ the result can be taken from §5.5.1, that is $\mathsf{E}(\widehat{\beta} \mid \mathcal{I}) = \beta_1$. The Law of Iterated Expectations, see (4.1.8), $\mathsf{E}(\widehat{\beta}) = \beta_1$.

(b) The calculation of the variance reported in §3.4.1 also holds conditionally on $\mathcal{I}$. From this, and from §5.5.1, we have

$$\mathsf{Var}\left(\widehat{\beta}\middle|\mathcal{I}\right) = \frac{\sigma^2}{n}, \qquad \mathsf{Var}\left(\widehat{\beta}_1\middle|\mathcal{I}\right) = \frac{\sigma^2}{\sum_{i=1}^n X_{1\cdot 2,i}^2}.$$

In general it holds $\mathsf{Var}(\widehat{\beta}|\mathcal{I}) \leq \mathsf{Var}(\widehat{\beta}_1|\mathcal{I})$ since:

$$\sum_{i=1}^n X_{1\cdot 2,i}^2 = n - \frac{\left(\sum_{i=1}^n X_i\right)^2}{\sum_{i=1}^n X_i^2} \leq n$$

since the subtracted ratio is non-negative.

(c) We have equality, $\mathsf{Var}(\widehat{\beta}|\mathcal{I}) = \mathsf{Var}(\widehat{\beta}_1|\mathcal{I})$, when the numerator of the subtracted ratio is zero, that is $\sum_{i=1}^n X_i = 0$.

(d) The ratio of the variances is:

$$\frac{\mathsf{Var}\left(\widehat{\beta}_1 \middle| \mathcal{I}\right)}{\mathsf{Var}\left(\widehat{\beta} \middle| \mathcal{I}\right)} = \frac{\sigma^2}{\sum_{i=1}^n X_{1\cdot 2,i}^2} \Bigg/ \frac{\sigma^2}{n} = \left\{1 - \frac{(\sum_{i=1}^n X_i)^2}{n \sum_{i=1}^n X_i^2}\right\}^{-1},$$

that is the inverse of one minus the square of a correlation type coefficient based on the intercept $X_{1,i}$ and the regressor $X_{2,i} = X_i$. It this correlation type quantity is close to one the variance ratio is large. For instance, if it holds that $X_1 = \cdots = X_m = M$ while $X_{m+1} = \cdots X_n = 0$ the variance ratio is

$$\frac{\mathsf{Var}\left(\widehat{\beta}_1 \middle| \mathcal{I}\right)}{\mathsf{Var}\left(\widehat{\beta} \middle| \mathcal{I}\right)} = \left\{1 - \frac{(mM)^2}{nn(M^2)}\right\}^{-1} = \left(1 - \frac{m}{n}\right)^{-1},$$

which is large if m is close to n. Note, that the actual value M of the regressors does not matter in this particular calculation.

Solution 5.10. The results in (5.2.10) are:

$$\widehat{Y}_i = \widehat{\delta}_1 + \widehat{\delta}_2 X_{2\cdot 1,i},$$

where the estimators are:

$$\widehat{\delta}_1 = \overline{Y}, \qquad \widehat{\delta}_2 = \frac{\sum_{i=1}^n X_{2\cdot 1,i}(Y_i - \overline{Y})}{\sum_{j=1}^n X_{2\cdot 1,j}^2}.$$

It follows from this that:

$$\begin{aligned} Y_i - \widehat{Y}_i &= Y_i - \widehat{\delta}_1 - \widehat{\delta}_2 X_{2\cdot 1,i} = Y_i - \overline{Y} - \widehat{\delta}_2 X_{2\cdot 1,i}, \\ \widehat{Y}_i - \overline{Y} &= \widehat{\delta}_1 + \widehat{\delta}_2 X_{2\cdot 1,i} - \overline{Y} = \widehat{\delta}_2 X_{2\cdot 1,i}. \end{aligned}$$

The sum of cross products of these terms are:

$$\begin{aligned} \sum_{i=1}^n \left(Y_i - \widehat{Y}_i\right)\left(\widehat{Y}_i - \overline{Y}\right) &= \sum_{i=1}^n \left(Y_i - \overline{Y} - \widehat{\delta}_2 X_{2\cdot 1,i}\right)\left(\widehat{\delta}_2 X_{2\cdot 1,i}\right) \\ &= \widehat{\delta}_2 \left\{\sum_{i=1}^n \left(Y_i - \overline{Y}\right) X_{2\cdot 1,i} - \widehat{\delta}_2 \sum_{i=1}^n X_{2\cdot 1,i}^2\right\}. \end{aligned}$$

It follows from the above definition of $\widehat{\delta}_2$ that the expression in curly brackets is zero.

Solution 5.12.
(a) First, show the identity for the variance estimators. From §5.4.3 it follows that

$$n\left(\widehat{\sigma}_R^2-\widehat{\sigma}^2\right)=\sum_{i=1}^n\left(Y_i-\overline{Y}\right)^2\left\{1-\left(1-r^2\right)\right\}=r^2\sum_{i=1}^n\left(Y_i-\overline{Y}\right)^2$$
$$=\frac{\left\{\sum_{i=1}^n X_{2\cdot1,i}\left(Y_i-\overline{Y}\right)\right\}^2}{\sum_{i=1}^n X_{2\cdot1,i}^2}=\widehat{\beta}_2^2\sum_{i=1}^n X_{2\cdot1,i}^2.$$

Secondly, using that $F=Z^2$ and the expression for $\widehat{\mathsf{se}}(\widehat{\beta}_2)$ and for s^2 stated in §5.5.3 it holds:

$$F=Z^2=\frac{\widehat{\beta}_2^2}{\left\{\widehat{\mathsf{se}}(\widehat{\beta}_2)\right\}^2}$$
$$=\frac{\widehat{\beta}_2^2\sum_{i=1}^n X_{2\cdot1,i}^2}{s^2}=\frac{\widehat{\beta}_2^2\sum_{i=1}^n X_{2\cdot1,i}^2}{\widehat{\sigma}^2 n/(n-2)}$$

Insert the above expression for the variance terms to get:

$$F=\frac{n\left(\widehat{\sigma}_R^2-\widehat{\sigma}^2\right)}{\widehat{\sigma}^2 n/(n-2)}=\frac{\left(\widehat{\sigma}_R^2-\widehat{\sigma}^2\right)}{\widehat{\sigma}^2/(n-2)}.$$

(b) For the second expression extend the fraction by $1/\widehat{\sigma}_R^2$ and use (5.4.6) to get:

$$F=\frac{\left(\widehat{\sigma}_R^2-\widehat{\sigma}^2\right)}{\widehat{\sigma}^2/(n-2)}=\frac{\left(1-\widehat{\sigma}^2/\widehat{\sigma}_R^2\right)}{\widehat{\sigma}^2/\widehat{\sigma}_R^2/(n-2)}=\frac{r^2}{(1-r^2)/(n-2)}.$$

For the third expression use (5.4.10):

$$F=\frac{r^2}{(1-r^2)/(n-2)}=\frac{\mathsf{ESS}/\mathsf{TSS}}{(\mathsf{RSS}/\mathsf{TSS})/(n-2)}=\frac{\mathsf{ESS}}{\mathsf{RSS}/(n-2)}.$$

Solution 5.14. To get r^2 we can use the formula $r^2=(\widehat{\sigma}_R^2-\widehat{\sigma}^2)/\widehat{\sigma}_R^2$. Here:

$$\widehat{\sigma}_R^2=n^{-1}\sum_{i=1}^n(Y_i-\overline{Y})^2=n^{-1}\sum_{i=1}^n Y_i^2-\overline{Y}^2$$
$$=(756.3848/7)-(71.7424/7)^2=3.014727.$$

Therefore:

$$r^2=\frac{3.014727-(0.129)^2}{3.014727}=0.99448.$$

The test statistics are:

$$F=\frac{r^2/1}{(1-r^2)/(7-2)}=900.7971,\qquad Z=\sqrt{F}=30.01.$$

The tests are:
One-sided t-test: reject since $30.01 > 2.02$. (95% quantile of $\mathsf{t}[5]$ is 2.02)
Two sided t-test: reject since $|30.01| > 2.57$. (97.5% quantile of $\mathsf{t}[5]$ is 2.57)
F-test: reject since $901 > 6.61$. (95% quantile of $\mathsf{f}[1,5]$ is 6.61)

Chapter Six

The matrix algebra of two-variable regression

Solution 6.2.

(a) First multiply matrices:

$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}$$

$$\mathbf{X}'\mathbf{Y} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 10 \\ -2 \end{pmatrix}.$$

Then find the inverse of $\mathbf{X}'\mathbf{X}$. This is a diagonal matrix so the inverse is simply the inverse of the diagonal elements:

$$(\mathbf{X}'\mathbf{X})^{-1} = \begin{pmatrix} 1/4 & 0 \\ 0 & 1/4 \end{pmatrix}.$$

Then multiply matrices to get the least squares estimator:

$$(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \begin{pmatrix} 1/4 & 0 \\ 0 & 1/4 \end{pmatrix} \begin{pmatrix} 10 \\ -2 \end{pmatrix} = \begin{pmatrix} 10/4 \\ -2/4 \end{pmatrix} = \begin{pmatrix} 5/2 \\ -1/2 \end{pmatrix}.$$

(b) The matrix multiplication "$\mathbf{X}'\mathbf{Y}(\mathbf{X}'\mathbf{X})^{-1}$" makes no sense, since $\mathbf{X}'\mathbf{Y}$ is a (2×1)-matrix, while $\mathbf{X}'\mathbf{X}$ is a (2×2)-matrix, so the column-dimension of $\mathbf{X}'\mathbf{Y}$ is one, while the row-dimension of $\mathbf{X}'\mathbf{Y}$ is two.

The matrix multiplication "$\dfrac{\mathbf{X}'\mathbf{Y}}{\mathbf{X}'\mathbf{X}}$" is not well-defined. Even if $\mathbf{X}'\mathbf{Y}$ and $\mathbf{X}'\mathbf{X}$ had the same dimension it is not clear whether it would mean multiplying $\mathbf{X}'\mathbf{Y}$ from the left by the inverse of $\mathbf{X}'\mathbf{X}$, or from the right by the inverse of $\mathbf{X}'\mathbf{X}$, or dividing each of the elements of $\mathbf{X}'\mathbf{Y}$ by the respective elements of $\mathbf{X}'\mathbf{X}$. In general, these operations would give different results.

Solution 6.4.

(a) The inverse of a (2×2)-matrix can be found using the formula (6.2.1). Thus, the inverse of the matrix $\mathbf{X}'\mathbf{X}$ in (6.3.3) it then:

$$\left(\mathbf{X}'\mathbf{X}\right)^{-1} = \begin{pmatrix} \sum_{i=1}^n 1^2 & \sum_{i=1}^n X_i \\ \sum_{i=1}^n X_i & \sum_{i=1}^n X_i^2 \end{pmatrix}^{-1} = \frac{1}{\det\left(\mathbf{X}'\mathbf{X}\right)} \begin{pmatrix} \sum_{i=1}^n X_i^2 & -\sum_{i=1}^n X_i \\ -\sum_{i=1}^n X_i & \sum_{i=1}^n 1^2 \end{pmatrix},$$

where $\det(\mathbf{X}'\mathbf{X})$ is the determinant of $\mathbf{X}'\mathbf{X}$ given by:

$$\det\left(\mathbf{X}'\mathbf{X}\right) = \left(\sum_{i=1}^n 1^2\right)\left(\sum_{i=1}^n X_i^2\right) - \left(\sum_{i=1}^n X_i\right)^2.$$

The least squares estimator is then

$$\widehat{\beta} = \frac{1}{\det\left(\mathbf{X}'\mathbf{X}\right)} \begin{pmatrix} \sum_{i=1}^n X_i^2 & -\sum_{i=1}^n X_i \\ -\sum_{i=1}^n X_i & \sum_{i=1}^n 1^2 \end{pmatrix} \begin{pmatrix} \sum_{i=1}^n Y_i \\ \sum_{i=1}^n X_i Y_i \end{pmatrix} = \frac{1}{\det\left(\mathbf{X}'\mathbf{X}\right)} \begin{pmatrix} \sum_{i=1}^n X_i^2 \sum_{i=1}^n Y_i - \sum_{i=1}^n X_i \sum_{i=1}^n X_i Y_i \\ \sum_{i=1}^n 1^2 \sum_{i=1}^n X_i Y_i - \sum_{i=1}^n X_i \sum_{i=1}^n Y_i \end{pmatrix}.$$

(b) The expressions in (5.2.1) arise by simplification of the above expression. Consider initially the second element and divide numerator and denominator by $\sum_{i=1}^n 1^2 = n$:

$$\widehat{\beta}_2 = \frac{\left(\sum_{i=1}^n 1^2 \sum_{i=1}^n X_i Y_i - \sum_{i=1}^n X_i \sum_{i=1}^n Y_i\right)/n}{\left\{\left(\sum_{i=1}^n 1^2\right)\left(\sum_{i=1}^n X_i^2\right) - \left(\sum_{i=1}^n X_i\right)^2\right\}/n} = \frac{\sum_{i=1}^n X_i Y_i - \left(\sum_{i=1}^n X_i/n\right)\sum_{i=1}^n Y_i}{\sum_{i=1}^n X_i^2 - n\left(\sum_{i=1}^n X_i/n\right)^2}. = \frac{\sum_{i=1}^n X_i Y_i - \overline{X}\sum_{i=1}^n Y_i}{\sum_{i=1}^n X_i^2 - n\left(\overline{X}\right)^2}.$$

Due to the formula found in Exercise 5.5(b) this reduces to the desired expression:

$$\widehat{\beta}_2 = \frac{\sum_{i=1}^n Y_i\left(X_i - \overline{X}\right)}{\sum_{i=1}^n \left(X_i - \overline{X}\right)^2}.$$

Consider now the first element and divide both numerator and denominator by

$\sum_{i=1}^{n} 1^2 = n$:

$$\widehat{\beta}_1 = \frac{\left(\sum_{i=1}^{n} X_i^2 \sum_{i=1}^{n} Y_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} X_i Y_i\right)/n}{\left\{\left(\sum_{i=1}^{n} 1^2\right)\left(\sum_{i=1}^{n} X_i^2\right) - \left(\sum_{i=1}^{n} X_i\right)^2\right\}/n}$$
$$= \frac{\sum_{i=1}^{n} X_i^2 \left(\sum_{i=1}^{n} Y_i/n\right) - \left(\sum_{i=1}^{n} X_i/n\right)\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2 - n\left(\sum_{i=1}^{n} X_i/n\right)^2}.$$
$$= \frac{\sum_{i=1}^{n} X_i^2 \overline{Y} - \overline{X}\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2 - n\left(\overline{X}\right)^2}.$$

Now, add and subtract $n\overline{X}^2\overline{Y}$ to get:

$$\widehat{\beta}_1 = \frac{\left(\sum_{i=1}^{n} X_i^2 - n\overline{X}^2\right)\overline{Y} - \overline{X}\left(\sum_{i=1}^{n} X_i Y_i n\overline{XY}\right)}{\sum_{i=1}^{n} X_i^2 - n\left(\overline{X}\right)^2}.$$

Due to the formula found in Exercise 5.5(b) this reduces to the desired expression:

$$\widehat{\beta}_1 = \overline{Y} - \widehat{\beta}_2\overline{X}.$$

Solution 6.6.

(a) When $Y_i \stackrel{\mathsf{D}}{=} \mathsf{N}[\beta, \sigma^2]$ then:

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}, \qquad \mathbf{X} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

By matrix multiplication:

$$\mathbf{X}'\mathbf{Y} = \sum_{i=1}^{n} Y_i = n\overline{Y}, \qquad \mathbf{X}'\mathbf{X} = \sum_{i=1}^{n} 1 = n, \qquad \left(\mathbf{X}'\mathbf{X}\right)^{-1} = \frac{1}{n},$$

so that:

$$\widehat{\beta} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{Y} = \overline{Y}.$$

(b) When $Y_i \stackrel{\mathsf{D}}{=} \mathsf{N}[\beta X_i, \sigma^2]$ then:

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}, \qquad \mathbf{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}.$$

By matrix multiplication:

$$\mathbf{X}'\mathbf{Y} = \sum_{i=1}^{n} X_i Y_i, \qquad \mathbf{X}'\mathbf{X} = \sum_{i=1}^{n} X_i^2, \qquad \left(\mathbf{X}'\mathbf{X}\right)^{-1} = \left(\sum_{i=1}^{n} X_i^2\right)^{-1},$$

so that:

$$\widehat{\beta} = \frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2}.$$

(c) When $Y_i \stackrel{\mathsf{D}}{=} \mathsf{N}[\beta_1 + \beta_2 X_i, \sigma^2]$ then:

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}, \qquad \mathbf{X} = \begin{pmatrix} 1 & X_1 \\ \vdots & \vdots \\ 1 & X_n \end{pmatrix}.$$

By matrix multiplication:

$$\mathbf{X}'\mathbf{Y} = \begin{pmatrix} \sum_{i=1}^{n} Y_i \\ \sum_{i=1}^{n} X_i Y_i \end{pmatrix},$$
$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} n & \sum_{i=1}^{n} X_i \\ \sum_{i=1}^{n} X_i & \sum_{i=1}^{n} X_i^2 \end{pmatrix},$$
$$(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{\det(\mathbf{X}'\mathbf{X})} \begin{pmatrix} \sum_{i=1}^{n} X_i^2 & -\sum_{i=1}^{n} X_i \\ -\sum_{i=1}^{n} X_i & n \end{pmatrix},$$

where $\det(\mathbf{X}'\mathbf{X}) = n \sum_{i=1}^{n} X_i^2 - (\sum_{i=1}^{n} X_i)^2$, so that:

$$\widehat{\beta} = \frac{1}{\det(\mathbf{X}'\mathbf{X})} \begin{pmatrix} \sum_{i=1}^{n} X_i^2 \sum_{i=1}^{n} Y_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i X_i \\ n \sum_{i=1}^{n} Y_i X_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i \end{pmatrix}.$$

By Exercise 6.4(b) this reduces further to the expressions in (5.2.1).

(d) When $Y_i \stackrel{\mathsf{D}}{=} \mathsf{N}[\beta_1 + \beta_2 X_i, \sigma^2]$ with $\sum_{i=1}^{n} X_i = 0$ the expressions in (c) reduce to:

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}, \qquad \mathbf{X} = \begin{pmatrix} 1 & X_1 \\ \vdots & \vdots \\ 1 & X_n \end{pmatrix}.$$

By matrix multiplication:

$$\mathbf{X}'\mathbf{Y} = \begin{pmatrix} \sum_{i=1}^{n} Y_i \\ \sum_{i=1}^{n} X_i Y_i \end{pmatrix},$$
$$\mathbf{X}'\mathbf{X} = \begin{pmatrix} n & 0 \\ 0 & \sum_{i=1}^{n} X_i^2 \end{pmatrix},$$
$$(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{\det(\mathbf{X}'\mathbf{X})} \begin{pmatrix} \sum_{i=1}^{n} X_i^2 & 0 \\ 0 & n \end{pmatrix} = \begin{pmatrix} \frac{1}{n} & 0 \\ 0 & \frac{1}{\sum_{i=1}^{n} X_i^2} \end{pmatrix},$$

where $\det(\mathbf{X}'\mathbf{X}) = n \sum_{i=1}^{n} X_i^2$, so that:

$$\widehat{\beta} = \begin{pmatrix} \frac{1}{n} \sum_{i=1}^{n} Y_i \\ \frac{\sum_{i=1}^{n} Y_i X_i}{\sum_{i=1}^{n} X_i^2} \end{pmatrix}.$$

Note, that unlike in (c) this result is a combination of the results in (a) and (b).

Chapter Seven

The multiple regression model

Solution 7.2. The calculations in this exercise are somewhat long. They could be reduced by using matrix algebra as set out in Chapter 8.

(a) The formula (7.2.8) gives:

$$\widehat{\delta}_2 = \widehat{\beta}_2 + a\widehat{\beta}_3, \qquad \widehat{\delta}_3 = \widehat{\beta}_3, \qquad \text{so} \qquad \widehat{\beta}_2 = \widehat{\delta}_2 - a\widehat{\delta}_3.$$

Inserting the expression for a in (7.2.6) and for $\widehat{\delta}_2$ and $\widehat{\delta}_3$ in (7.2.12) then results in:

$$\widehat{\beta}_2 = \frac{\sum_{i=1}^n Y_i X_{2\cdot 1,i}}{\sum_{j=1}^n X_{2\cdot 1,j}^2} - \left(\frac{\sum_{j=1}^n X_{3\cdot 1,j} X_{2\cdot 1,j}}{\sum_{k=1}^n X_{2\cdot 1,k}^2}\right) \frac{\sum_{i=1}^n Y_i X_{3\cdot 1,2,i}}{\sum_{j=1}^n X_{3\cdot 1,2,j}^2}.$$

Put these expressions into a common fraction:

$$\widehat{\beta}_2 = \frac{\sum_{i=1}^n Y_i X_{2\cdot 1,i} \left\{\sum_{j=1}^n X_{3\cdot 1,j}^2 - \frac{\left(\sum_{j=1}^n X_{2\cdot 1,j} X_{3\cdot 1,j}\right)^2}{\sum_{j=1}^n X_{2\cdot 1,j}^2}\right\}}{\sum_{j=1}^n X_{2\cdot 1,j}^2 \sum_{j=1}^n X_{3\cdot 1,j}^2 - \left(\sum_{j=1}^n X_{2\cdot 1,j} X_{3\cdot 1,j}\right)^2}$$
$$- \frac{\sum_{j=1}^n X_{2\cdot 1,j} X_{3\cdot 1,j} \sum_{i=1}^n Y_i X_{3\cdot 1,i}}{\sum_{j=1}^n X_{2\cdot 1,j}^2 \sum_{j=1}^n X_{3\cdot 1,j}^2 - \left(\sum_{j=1}^n X_{2\cdot 1,j} X_{3\cdot 1,j}\right)^2}$$
$$+ \frac{\sum_{i=1}^n Y_i X_{2\cdot 1,i} \frac{\left(\sum_{j=1}^n X_{2\cdot 1,j} X_{3\cdot 1,j}\right)^2}{\sum_{j=1}^n X_{2\cdot 1,j}^2}}{\sum_{j=1}^n X_{2\cdot 1,j}^2 \sum_{j=1}^n X_{3\cdot 1,j}^2 - \left(\sum_{j=1}^n X_{2\cdot 1,j} X_{3\cdot 1,j}\right)^2},$$

which reduces to:

$$\widehat{\beta}_2 = \frac{\sum_{i=1}^n Y_i X_{2\cdot 1,i} \sum_{j=1}^n X_{3\cdot 1,j}^2 - \sum_{j=1}^n X_{2\cdot 1,j} X_{3\cdot 1,j} \sum_{i=1}^n Y_i X_{3\cdot 1,i}}{\sum_{j=1}^n X_{2\cdot 1,j}^2 \sum_{j=1}^n X_{3\cdot 1,j}^2 - \left(\sum_{j=1}^n X_{2\cdot 1,j} X_{3\cdot 1,j}\right)^2}.$$

Divide through by $\sum_{j=1}^n X_{3\cdot 1,j}^2$ to see that:

$$\widehat{\beta}_2 = \frac{\sum_{i=1}^n Y_i X_{2\cdot 1,3,i}}{\sum_{j=1}^n X_{2\cdot 1,3,j}^2}.$$

(b) The formula (7.2.8) gives:

$$\widehat{\beta}_1 = \widehat{\delta}_1 - \widehat{\delta}_2 \overline{X}_2 - \widehat{\delta}_3 \left(\overline{X}_3 - a \overline{X}_2 \right).$$

Start by rewriting the first two components:

$$\widehat{\delta}_1 - \widehat{\delta}_2 \overline{X}_2 = \frac{\sum_{j=1}^n X_{1,j} Y_j}{\sum_{j=1}^n X_{1,j}^2} - \frac{\sum_{j=1}^n X_{2\cdot 1,j} Y_j}{\sum_{j=1}^n X_{2\cdot 1,j}^2} \left(\frac{\sum_{j=1}^n X_{1,j} X_{2,j}}{\sum_{j=1}^n X_{1,j}^2} \right).$$

Bring them on a common fraction as:

$$\widehat{\delta}_1 - \widehat{\delta}_2 \overline{X}_2 = \frac{\sum_{j=1}^n X_{1,j} Y_j \sum_{j=1}^n X_{2\cdot 1,j}^2 - \sum_{j=1}^n X_{2\cdot 1,j} Y_j \sum_{j=1}^n X_{1,j} X_{2,j}}{\sum_{j=1}^n X_{1,j}^2 \sum_{j=1}^n X_{2\cdot 1,j}^2}.$$

Divide numerator and denominator by $\sum_{j=1}^n X_{2,j}^2$ to get:

$$\begin{aligned}
\widehat{\delta}_1 - \widehat{\delta}_2 \overline{X}_2 &= \frac{\sum_{j=1}^n X_{1,j} Y_j \left\{ 1 - \frac{\left(\sum_{j=1}^n X_{1,j} X_{2,j} \right)^2}{\sum_{j=1}^n X_{1,j}^2 \sum_{j=1}^n X_{2,j}^2} \right\}}{\sum_{j=1}^n X_{1\cdot 2,j}^2} \\
&\quad - \frac{\left(\sum_{j=1}^n X_{2,j} Y_j - \frac{\sum_{j=1}^n X_{2,j} X_{1,j} \sum_{j=1}^n X_{1,j} Y_j}{\sum_{j=1}^n X_{1,j}^2} \right) \frac{\sum_{j=1}^n X_{1,j} X_{2,j}}{\sum_{j=1}^n X_{2,j}^2}}{\sum_{j=1}^n X_{1\cdot 2,j}^2} \\
&= \frac{\sum_{j=1}^n X_{1\cdot 2,j} Y_j}{\sum_{j=1}^n X_{1\cdot 2,j}^2}.
\end{aligned}$$

Continue by analysing the third component the third component in the same way. This is:

$$\begin{aligned}
&\widehat{\delta}_3 \left(\overline{X}_3 - a \overline{X}_2 \right) \\
&= \widehat{\delta}_3 \left\{ \frac{\sum_{j=1}^n X_{1,j} X_{3,j}}{\sum_{j=1}^n X_{1,j}^2} - \left(\frac{\sum_{j=1}^n X_{3\cdot 1,j} X_{2\cdot 1,j}}{\sum_{j=1}^n X_{2\cdot 1,j}^2} \right) \frac{\sum_{j=1}^n X_{1,j} X_{2,j}}{\sum_{j=1}^n X_{1,j}^2} \right\}
\end{aligned}$$

Bring this on a common fraction:

$$\begin{aligned}
&\widehat{\delta}_3 \left(\overline{X}_3 - a \overline{X}_2 \right) \\
&= \frac{\widehat{\delta}_3}{\sum_{j=1}^n X_{1,j}^2 \sum_{j=1}^n X_{2\cdot 1,j}^2} \left[\sum_{j=1}^n X_{1,j} X_{3,j} \left\{ \sum_{j=1}^n X_{2,j}^2 - \frac{\left(\sum_{j=1}^n X_{1,j} X_{2,j} \right)^2}{\sum_{j=1}^n X_{1,j}^2} \right\} \right. \\
&\qquad \left. - \left(\sum_{j=1}^n X_{3,j} X_{2,j} - \frac{\sum_{j=1}^n X_{3,j} X_{1,j} \sum_{j=1}^n X_{1,j} X_{2,j}}{\sum_{j=1}^n X_{1,j}^2} \right) \frac{\sum_{j=1}^n X_{1,j} X_{2,j}}{\sum_{j=1}^n X_{2,j}^2} \right].
\end{aligned}$$

Divide numerator and denominator by $\sum_{j=1}^{n} X_{2,j}^2$ to get:

$$\widehat{\delta}_3 \left(\overline{X}_3 - a\overline{X}_2\right) = \widehat{\delta}_3 \frac{\sum_{j=1}^{n} X_{3\cdot 2,j} X_{1\cdot 2,j}}{\sum_{j=1}^{n} X_{1\cdot 2,j}^2}.$$

Now bring all the components together:

$$\widehat{\beta}_1 = \frac{\sum_{j=1}^{n} X_{1\cdot 2,j} Y_j}{\sum_{j=1}^{n} X_{1\cdot 2,j}^2} - \frac{\sum_{j=1}^{n} X_{3\cdot 1,2,j} Y_j}{\sum_{j=1}^{n} X_{3\cdot 1,2,j}^2} \left(\frac{\sum_{j=1}^{n} X_{3\cdot 2,j} X_{1\cdot 2,j}}{\sum_{j=1}^{n} X_{1\cdot 2,j}^2}. \right).$$

As before, bring these on a common fraction. Then divide numerator and denominator by $\sum_{j=1}^{n} X_{3\cdot 2,j}^2$ to get the desired expression:

$$\widehat{\beta}_1 = \frac{\sum_{j=1}^{n} X_{1\cdot 2,3,j} Y_j}{\sum_{j=1}^{n} X_{1\cdot 2,3,j}^2}.$$

Solution 7.4. The residuals from the two-variable model are given in (5.2.11) as:

$$\widehat{Y}_i = (Y_i - \overline{Y}) - \widehat{\gamma}_Y (X_{2,i} - \overline{X}_2),$$
$$\widehat{X}_{3,i} = (X_{3,i} - \overline{X}_3) - \widehat{\gamma}_X (X_{2,i} - \overline{X}_2),$$

where $\widehat{\gamma}_Y$ and $\widehat{\gamma}_X$ are the least squares estimators:

$$\widehat{\gamma}_Y = \frac{\sum_{i=1}^{n} (Y_i - \overline{Y})(X_{2,i} - \overline{X}_2)}{\sum_{i=1}^{n} (X_{2,i} - \overline{X}_2)^2},$$
$$\widehat{\gamma}_X = \frac{\sum_{i=1}^{n} (X_{3,i} - \overline{X}_3)(X_{2,i} - \overline{X}_2)}{\sum_{i=1}^{n} (X_{2,i} - \overline{X}_2)^2}.$$

The least-squares estimator from the regression of $\widehat{Y}_i$ on $\widehat{X}_{3,i}$ is

$$\widehat{\gamma} = [\text{ one variable model }] = \frac{\sum_{i=1}^{n} \widehat{Y}_i \widehat{X}_{3,i}}{\sum_{i=1}^{n} \widehat{X}_{3,i}^2}$$
$$= [\text{ see below }] = \frac{\sum_{i=1}^{n} Y_i X_{3\cdot 1,2,i}}{\sum_{i=1}^{n} X_{3\cdot 1,2,i}^2}.$$

The denominator is:

$$\begin{aligned}
\sum_{i=1}^{n} \widehat{X}_{3,i}^2 &= \sum_{i=1}^{n} \{(X_{3,i} - \overline{X}_3) - \widehat{\gamma}_X (X_{2,i} - \overline{X}_2)\}^2 \\
&= \sum_{i=1}^{n} (X_{3,i} - \overline{X}_3)^2 - \frac{\{\sum_{i=1}^{n} (X_{3,i} - \overline{X}_3)(X_{2,i} - \overline{X}_2)\}^2}{\sum_{i=1}^{n} (X_{2,i} - \overline{X}_2)^2} \\
&= \sum_{i=1}^{n} X_{3\cdot 1,2,i}^2.
\end{aligned}$$

The numerator is:

$$
\begin{aligned}
&\sum_{i=1}^{n} \widehat{Y}_i \widehat{X}_{3,i} \\
&= \sum_{i=1}^{n} \{(Y_i - \overline{Y}) - \widehat{\gamma}_Y (X_{2,i} - \overline{X}_2)\}\{(X_{3,i} - \overline{X}_3) - \widehat{\gamma}_X (X_{2,i} - \overline{X}_2)\} \\
&= \sum_{i=1}^{n} (Y_i - \overline{Y})(X_{3,i} - \overline{X}_3) \\
&\quad - \frac{\sum_{i=1}^{n} (Y_i - \overline{Y})(X_{2,i} - \overline{X}_2) \sum_{i=1}^{n} (X_{3,i} - \overline{X}_3)(X_{2,i} - \overline{X}_2)}{\sum_{i=1}^{n} (X_{2,i} - \overline{X}_2)^2} \\
&= \sum_{i=1}^{n} Y_i X_{3\cdot 1,2,i}.
\end{aligned}
$$

Solution 7.6. The idea is to show that:

$$
\begin{aligned}
\sum_{i=1}^{n} \widehat{u}_{y\cdot 1,2,i}^2 &= (1 - r_{y,2\cdot 1}^2) \sum_{i=1}^{n} (Y_i - \overline{Y})^2, \\
\sum_{i=1}^{n} \widehat{u}_{3\cdot 1,2,i}^2 &= (1 - r_{3,2\cdot 1}^2) \sum_{i=1}^{n} (X_{3,i} - \overline{X}_3)^2, \\
\sum_{i=1}^{n} \widehat{u}_{y\cdot 1,2,i} \widehat{u}_{3\cdot 1,2,i} &= (r_{y,3\cdot 1} - r_{y,2\cdot 1} r_{2,3\cdot 1}) \\
&\qquad \times \left\{ \sum_{i=1}^{n} (Y_i - \overline{Y})^2 \sum_{i=1}^{n} (X_{3,i} - \overline{X}_3)^2 \right\}^{1/2}.
\end{aligned}
$$

The desired expression then arise by dividing the latter expression with the square root of the two first. The first two expressions are proved as following. From the two-variable regression analysis of Y_i on $X_{1,i} = 1, X_{2,i}$ it is found that:

$$\begin{aligned}
&\sum_{i=1}^{n} \widehat{u}_{y\cdot 1,2,i}^{2} \\
&= \sum_{i=1}^{n} \left\{ (Y_i - \overline{Y}) - \frac{\sum_{j=1}^{n} (Y_i - \overline{Y})(X_{2,j} - \overline{X}_2)}{\sum_{j=1}^{n} (X_{2,j} - \overline{X}_2)^2} (X_{2,j} - \overline{X}_2) \right\}^2 \\
&= \sum_{i=1}^{n} (Y_i - \overline{Y})^2 - \frac{\{\sum_{j=1}^{n} (Y_i - \overline{Y})(X_{2,j} - \overline{X})\}^2}{\sum_{j=1}^{n} (X_{2,j} - \overline{X}_2)^2} \\
&= \sum_{i=1}^{n} (Y_i - \overline{Y})^2 \left[1 - \frac{\{\sum_{j=1}^{n} (Y_i - \overline{Y})(X_{2,j} - \overline{X}_2)\}^2}{\sum_{i=1}^{n} (Y_i - \overline{Y})^2 \sum_{j=1}^{n} (X_{2,j} - \overline{X}_2)^2} \right] \\
&= \left(1 - r_{y,2\cdot 1}^2\right) \sum_{i=1}^{n} (Y_i - \overline{Y})^2.
\end{aligned}$$

as desired. The last expression is proved in the same way:

$$\begin{aligned}
&\sum_{i=1}^{n} \widehat{u}_{y\cdot 1,2,i} \widehat{u}_{3\cdot 1,2,i} \\
&= \sum_{i=1}^{n} \left\{ (Y_i - \overline{Y}) - \frac{\sum_{j=1}^{n} (Y_i - \overline{Y})(X_{2,j} - \overline{X}_2)}{\sum_{j=1}^{n} (X_{2,j} - \overline{X}_2)^2} (X_{2,i} - \overline{X}_2) \right\} \\
&\qquad \times \left\{ (X_{3,i} - \overline{X}_3) - \frac{\sum_{j=1}^{n} (X_{3,j} - \overline{X}_3)(X_{2,j} - \overline{X}_2)}{\sum_{j=1}^{n} (X_{2,j} - \overline{X}_2)^2} (X_{2,i} - \overline{X}_2) \right\} \\
&= \sum_{i=1}^{n} (Y_i - \overline{Y})(X_{3,i} - \overline{X}_3) \\
&\quad - \frac{\{\sum_{j=1}^{n} (Y_j - \overline{Y})(X_{2,j} - \overline{X}_2)\}\{\sum_{j=1}^{n} (X_{3,j} - \overline{X}_3)(X_{2,j} - \overline{X}_2)\}}{\sum_{j=1}^{n} (X_{2,j} - \overline{X}_2)^2} \\
&= (r_{y,3\cdot 1} - r_{y,2\cdot 1} r_{2,3\cdot 1}) \left\{ \sum_{i=1}^{n} (Y_i - \overline{Y})^2 \sum_{i=1}^{n} (X_{3,i} - \overline{X}_3)^2 \right\}^{1/2}.
\end{aligned}$$

Solution 7.8.
(a) The formula (7.4.1) gives:

$$R^2 = \frac{\mathsf{E}SS}{\mathsf{T}SS} = \frac{\sum_{i=1}^{n} (\widehat{Y}_i - \overline{Y})^2}{\sum_{i=1}^{n} (Y_i - \overline{Y})^2}.$$

Noting that $\widehat{\delta}_1 = \overline{Y}$ then apply (7.2.15) to the numerator, and note that the denominator is the sum of squared residuals in a regression of Y_i on a constant.

Thus:

$$R^2 = \frac{\widehat{\delta}_2^2 \sum_{i=1}^n X_{2\cdot 1,i}^2 + \widehat{\delta}_3^2 \sum_{i=1}^n X_{3\cdot 1,2,i}^2}{\sum_{i=1}^n \widehat{u}_{y\cdot 1}^2}.$$

(b) The same argument as in (a), but for a two-variable regression, where R^2 equals $r_{y\cdot 2,1}^2$ So for the numerator apply (5.2.10) instead of (7.2.15).

(c) From the definition in §7.3.1 we have:

$$r_{y,2\cdot 1,2}^2 \sum_{i=1}^n \widehat{u}_{y\cdot 1,2,i}^2 = \frac{\left(\sum_{i=1}^n \widehat{u}_{y\cdot 1,2,i}\widehat{u}_{3\cdot 2,1,i}\right)^2}{\sum_{i=1}^n \widehat{u}_{3\cdot 1,2,i}^2}.$$

But; $\sum_{i=1}^n \widehat{u}_{y\cdot 1,2,i}\widehat{u}_{3\cdot 2,1,i} = \sum_{i=1}^n y_i \widehat{u}_{3\cdot 2,1,i}$ and $\widehat{u}_{3\cdot 1,2,i} = X_{3\cdot 1,2,i}$ so:

$$r_{y,2\cdot 1,2}^2 \sum_{i=1}^n \widehat{u}_{y\cdot 1,2,i}^2 = \frac{\left(\sum_{i=1}^n y_i X_{3\cdot 2,1,i}\right)^2}{\sum_{i=1}^n X_{3\cdot 1,2,i}^2} = \widehat{\delta}_3^2 \sum_{i=1}^n X_{3\cdot 1,2,i}^2,$$

which gives the desired formula.

(d) This is the formula (5.4.10).

(e) Use the results in the sequence (a), (b), (d), (c) to get:

$$\begin{aligned}
& 1 - R^2 \\
&= [\text{ use } (a) \text{ }] = 1 - \frac{\widehat{\delta}_2^2 \sum_{i=1}^n X_{2\cdot 1,i}^2}{\sum_{i=1}^n \widehat{u}_{y\cdot 1}^2} - \frac{\widehat{\delta}_3^2 \sum_{i=1}^n X_{3\cdot 1,2,i}^2}{\sum_{i=1}^n \widehat{u}_{y\cdot 1}^2} \\
&= [\text{ use } (b) \text{ }] = 1 - r_{y,2\cdot 1}^2 - \frac{\widehat{\delta}_3^2 \sum_{i=1}^n X_{3\cdot 1,2,i}^2}{\sum_{i=1}^n \widehat{u}_{y\cdot 1}^2} \\
&= [\text{ extend fraction }] = 1 - r_{y,2\cdot 1}^2 - \left(\frac{\widehat{\delta}_3^2 \sum_{i=1}^n X_{3\cdot 1,2,i}^2}{\sum_{i=1}^n \widehat{u}_{y\cdot 1,2}^2}\right)\left(\frac{\sum_{i=1}^n \widehat{u}_{y\cdot 1,2}^2}{\sum_{i=1}^n \widehat{u}_{y\cdot 1}^2}\right) \\
&= [\text{ use } (d) \text{ }] = \left(1 - r_{y,2\cdot 1}^2\right)\left(1 - \frac{\widehat{\delta}_3^2 \sum_{i=1}^n X_{3\cdot 1,2,i}^2}{\sum_{i=1}^n \widehat{u}_{y\cdot 1,2}^2}\right) \\
&= [\text{ use } (c) \text{ }] = \left(1 - r_{y,2\cdot 1}^2\right)\left(1 - r_{y,3\cdot 2,1\cdot}^2\right)
\end{aligned}$$

Solution 7.10. From (7.6.2) with $\mathsf{RSS}_R = \mathsf{TSS}$, $m = 2$, $k = 3$ it follows:

$$F = \frac{(\mathsf{TSS} - \mathsf{RSS})/2}{\mathsf{RSS}/(n-3)}.$$

Exploiting the relation (7.4.4) that $\mathsf{TSS} = \mathsf{ESS} + \mathsf{RSS}$ then gives:

$$F = \frac{\mathsf{ESS}/2}{(\mathsf{TSS} - \mathsf{ESS})/(n-3)} = \frac{(\mathsf{ESS}/\mathsf{TSS})/2}{\{1 - (\mathsf{ESS}/\mathsf{TSS})\}/(n-3)}$$

The desired expression then follows from (7.4.1).