Econometrie: tussentijdse test 1

Oplossingen

1. (a) 38.65% van de gezinnen heeft kinderen.

   (b) $\widehat{\mu}_0 = 2395$ euro; $\widehat{\mu}_1 = 4564$ euro.

   (c) We toetsen $H_0 : \mu_0 = \mu_1$ tegenover $H_1 : \mu_0 < \mu_1$. We gebruiken de t-statistiek. Onder H_0 is de t-statistiek asymptotisch standaardnormaal verdeeld. De waarde van de t-statistiek is -35.83. De p-waarde is $\Phi(-35.83) = 2.09 \times 10^{-281}$ (extreem dicht bij 0). We besluiten dat, zo goed als zeker, gezinnen met kinderen gemiddeld meer uitgeven aan voeding dan gezinnen zonder kinderen.

2. (a) De schattingsresultaten zijn als volgt:

$$\begin{array}{rcccl} \widehat{efood} & = & 1933 & + & 0.0565 income \\ & & (127) & & (0.0058) \end{array}$$

$$R^2 = 0.219 \qquad SER = 1752$$

   (b) We toetsen $H_0 : \beta_1 = 0.05$ tegenover $H_1 : \beta_1 > 0.05$. We gebruiken de t-statistiek. Onder H_0 is de t-statistiek asymptotisch standaardnormaal verdeeld. De waarde van de t-statistiek is 1.121. De p-waarde is $1 - \Phi(1.121) = 0.131$. We besluiten dat de data redelijk goed in lijn zijn met de nulhypothese. In elk geval kunnen we *niet* met grote zekerheid besluiten dat $\beta_1 > 0.05$.

   (c) (i) is fout: de uitgaven aan voeding stijgen (gemiddeld) met 5.65 eurocent als het gezinsinkomen met 1 euro toeneemt. (ii) is fout: de standaardfout van de regressie is 1752 euro; de standaarddeviatie van $efood$ is 1982 euro. (iii) is fout: de kleine standaardfout van $\widehat{\beta}_1$ wijst er op dat $\widehat{\beta}_1$ een nauwkeurige schatter is van β_1.

3. (a) Het 99.99% betrouwbaarheidsinterval voor β_2 is $[614.2, 854.3]$ (euro/kind).

   (b) Zowel $nchi$ en $nadult$ hebben een positief effect op $efood$. Bovendien is $\widehat{cov}(income, nchi) > 0$ en $\widehat{cov}(income, nadult) > 0$. Daarom neemt $\widehat{\beta}_1$ in de regressie *zonder nchi* en *nadult* een deel van de positieve effecten β_2 en β_3 over. Daarom wordt $\widehat{\beta}_1$ kleiner in de regressie *met nchi* en *nadult*.

(c) We toetsen $H_0 : 2\beta_2 = \beta_3$ tegenover $H_1 : 2\beta_2 \neq \beta_3$. (Merk op: volgens H_0 tellen 2 kinderen hetzelfde als 1 volwassene voor de uitgaven aan voeding). We gebruiken de F-statistiek. Onder H_0 is de F-statistiek asymptotisch $F(1, \infty)$ verdeeld. De waarde van de F-statistiek is 4.93. De p-waarde is 0.0265. We besluiten met redelijk grote zekerheid dat de nulhypothese fout is; we zijn daar met name 97% zeker van.

(d) β_3 is uitgedrukt in euro per volwassene.

(e) De standaardfout van de regressie is uitgedrukt in euro.

(f) De voorspelde waarde van $efood$ voor dit gezin is

$$\widehat{\beta}_0 + \widehat{\beta}_1 \times 30000 + \widehat{\beta}_2 \times 1 + \widehat{\beta}_3 \times 2 = 4105.6 \text{ euro}.$$

(g) Definieer $efood^* = efood/1000$ en $income^* = income/1000$. De regressievergelijking is nu

$$efood_i^* = \beta_0^* + \beta_1^* income_i^* + \beta_2^* nchi_i + \beta_3^* nadult_i + u_i^* \qquad (1)$$

met $\beta_0^*, \ldots, \beta_4^*$ de nieuwe coefficiënten en u_i^* de nieuwe storingsterm. Herschrijf (1) als

$$\frac{efood_i}{1000} = \beta_0^* + \beta_1^* \frac{income_i}{1000} + \beta_2^* nchi_i + \beta_3^* nadult_i + u_i^*.$$

Beide zijden vermenigvuldigen met 1000 geeft

$$efood_i = 1000\beta_0^* + \beta_1^* income_i + 1000\beta_2^* nchi_i + 1000\beta_3^* nadult_i + 1000u_i^*. \qquad (2)$$

Merk nu op dat de regressand in (2), $efood_i$, identiek is aan de regressand in de oorspronkelijke vergelijking,

$$efood_i = \beta_0 + \beta_1 income_i + \beta_2 nchi_i + \beta_3 nadult_i + u_i. \qquad (3)$$

Uit (2)–(3) volgt dat

$$\beta_0 = 1000\beta_0^*, \quad \beta_1 = \beta_1^*, \quad \beta_2 = 1000\beta_2^*, \quad \beta_3 = 1000\beta_3^*, \quad u_i = 1000u_i^*,$$

en dus

$$\beta_0^* = \beta_0/1000, \quad \beta_1^* = \beta_1, \quad \beta_2^* = \beta_2/1000, \quad \beta_3^* = \beta_3/1000, \quad u_i^* = u_i/1000.$$

De nieuwe coëfficienten $\beta_0^*, \beta_2^*, \beta_3^*$ zijn duizend keer kleiner dan de oorspronkelijke, terwijl $\beta_1^* = \beta_1$ onveranderd blijft. De nieuwe SER is

$$\begin{aligned} SER^* &= \sqrt{\frac{1}{n-3-1}\sum_{i=1}^{3816}(\widehat{u}_i^*)^2} \\ &= \sqrt{\frac{1}{n-3-1}\sum_{i=1}^{3816}\left(\frac{\widehat{u}_i}{1000}\right)^2} \\ &= \frac{1}{1000}\sqrt{\frac{1}{n-3-1}\sum_{i=1}^{3816}\widehat{u}_i^2} \\ &= \frac{SER}{1000}, \end{aligned}$$

dus duizend keer kleiner dan de oorspronkelijke SER. De nieuwe R^2 is

$$(R^2)^* = \frac{ESS^*}{TSS^*} = \frac{ESS/1000^2}{TSS/1000^2} = \frac{ESS}{TSS} = R^2$$

(in voor de hand liggende notatie) en is dus onveranderd.