

# Econometrie: tussentijdse test 1

## Oplossingen

- (a) Het gemiddeld gezinsinkomen is 23017.26 euro.  
(b) De standaarddeviatie van het gezinsinkomen is 16412.63 euro. De standaardfout van het gemiddeld gezinsinkomen (als schatter van het populatiegemiddelde) is

$$s_{\bar{X}} = \frac{1}{\sqrt{n}} s_X = \frac{1}{\sqrt{3816}} 16412.63 = 265.69 \text{ euro}$$

- (c) Het 95%-betrouwbaarheidsinterval is

$$[\bar{X} - 1.96 \times s_{\bar{X}}, \bar{X} + 1.96 \times s_{\bar{X}}] = [22496.51, 23538.01].$$

De grenzen van het interval zijn uitgedrukt in euro.

- De correlatie tussen de uitgaven voor voeding en de totale uitgaven voor niet-duurzame goederen is 0.5481.
- Na de beperking van de steekproef is de correlatie 0.2708.
- Het mediaan equivalent inkomen is 11438.05 euro.
- De armoedegrens is  $0.6 \times 11438.05 = 6862.83$  euro. Deze armoedegrens impliceert dat 13.57% van de gezinnen in armoede leeft.
- (a) De schattingsresultaten zijn

$$\widehat{efood} = 1589.2 + 0.05901 etot$$

(79.2)      (0.0031)

$$R^2 = 0.3005 \quad SER = 1658.1$$

$\hat{\beta}_1 = 0.05901$  geeft aan dat, als de totale uitgaven aan niet-duurzame goederen met 1 euro toenemen, de uitgaven aan voeding toenemen met gemiddeld 5.9 eurocent.

- (b) De gemiddelde restterm is  $-1.73 \times 10^{-7}$  (euro), dus praktisch gelijk aan 0. Dit is theoretisch te verklaren: de eerste-orde voorwaarde voor  $\hat{\beta}_0$  impliceert dat de gemiddelde restterm exact 0 is. Stata rapporteert geen exacte 0 maar dat is te wijten aan afrondingsfouten.
- (c) De correlatie is  $-4.519 \times 10^{-10}$ , opnieuw praktisch gelijk aan 0. Ook dit is theoretisch te verklaren: de eerste-orde voorwaarde voor  $\hat{\beta}_1$  impliceert dat de steekproefcovariantie tussen de restterm en de regressor exact 0 is (en dus ook de steekproefcorrelatie).
- (d)  $SER = 1658.1$ . Dit betekent dat de variatie van de gezinsuitgaven aan voeding rond de geschatte regressielijn een standaarddeviatie heeft van 1658.1 euro.
- (e) Het is overduidelijk dat de storingstermen heteroskedastisch zijn: de variantie van de resttermen (en dus ook die van de storingstermen) neemt toe met  $etot$ .
- (f) Defineer  $etot^* = etot/1000$  en  $efood^* = efood/1000$ . Dan kunnen we

$$efood = \beta_0 + \beta_1 etot + u$$

herschrijven als

$$1000efood^* = \beta_0 + \beta_1 1000etot^* + u$$

en dus als

$$efood^* = \underbrace{\beta_0/1000}_{\beta_0^*} + \underbrace{\beta_1}_{\beta_1^*} etot^* + \underbrace{u/1000}_{u^*}$$

Dus:  $\beta_1$  blijft ongewijzigd,  $\beta_0$  wordt duizendmaal kleiner,  $SER$  wordt ook duizendmaal kleiner.

7. (a) (i)  $\beta_2 > 0$ : hoe meer volwassenen in het gezin, hoe groter de uitgaven aan voeding (zelfs voor gegeven  $etot$  en  $nchild$ ). (ii) De covariantie tussen  $etot$  en  $nadult$  is positief (logischerwijze). De combinatie van (i) en (ii) impliceert dat, als  $nadult$  niet als regressor wordt opgenomen,  $\hat{\beta}_1$  naar *boven* toe vertekend is. Dezelfde redenering geldt voor  $nchild$ . Daarom *daalt*  $\hat{\beta}_1$  als  $nadult$  en  $nchild$  worden opgenomen in de regressie.

- (b)
- i. We toetsen  $H_0 : \beta_2 = \beta_3$  tegenover  $H_1 : \beta_2 \neq \beta_3$ .
  - ii. We gebruiken de F-toets.
  - iii. Onder  $H_0$  is de F-statistiek asymptotisch verdeeld volgens de  $F_{1,\infty}$  verdeling.
  - iv. De waarde van de F-statistiek is 78.64.
  - v. De  $p$ -waarde die hoort bij  $F = 78.64$  wordt door Stata gerapporteerd als  $p = 0.0000$ . D.w.z. dat de  $p$ -waarde kleiner is dan 0.00005. Een nauwkeuriger  $p$ -waarde kan je aflezen na *return list*: de  $p$ -waarde is  $1.126 \times 10^{-18}$ , extreem dicht bij 0.
  - vi. We besluiten dat kinderen en volwassenen verschillende effecten hebben op de gezinsuitgaven aan voeding (voor constant gehouden uitgaven aan niet-duurzame goederen). Met name is het effect van een extra kind ongeveer half zo groot als dat van een extra volwassene (en beide effecten zijn positief).