

Econometrie: tussentijdse test 1

Voor deze test werken we met de data van een Belgische budgetenquête uit het jaar 2000. De dataset bevat 3816 waarnemingen (huishoudens/gezinnen) en is beschikbaar in Stata formaat in de folder

\\stud-shares.icts.kuleuven.be\data\GDhaene\Econometrie\Test1.

Maak een folder Test1 aan op jouw bureaublad, kopieer budgetenquete.dta en test1.do daarin, en open de dataset in Stata (dubbelklik op het .dta bestand). Inspecteer de dataset (gebruik *browse*) en stel vast dat de dataset onder andere de volgende variabelen bevat:

nadult: aantal volwassenen

nchi1: aantal kinderen ≤ 3 jaar

nchi2: aantal kinderen > 3 en ≤ 6 jaar

nchi3: aantal kinderen > 6 en ≤ 13 jaar

nchi4: aantal kinderen > 13 jaar

efood: uitgaven voor voeding

etot: totale uitgaven voor niet-duurzame goederen

income: gezinsinkomen

Inkomens en uitgaven hebben betrekking op het jaar 2000 en zijn uitgedrukt in euro.

Instructies:

- Voor het met werken met Stata mag je gebruik maken van test1.do en dat bestand verder aanvullen.
- Gebruik steeds heteroskedasticiteits-robuste standaardfouten.
- Voor hypothesetoetsen, wees zorgvuldig: geef steeds alle stappen aan zoals opgelijst op p. 17 van Lab2.pdf.
- Als je een getal rapporteert, vermeld steeds de eenheid.
- Beantwoord de volgende vragen en schrijf je antwoorden neer op het antwoordblad.

1. Maak een 0/1-variabele *chi* aan die 0 is voor gezinnen zonder kinderen en 1 voor gezinnen met kinderen (d.w.z. gezinnen met minstens één kind).
 - (a) (1 punt) Hoeveel procent van de gezinnen heeft kinderen?
 - (b) (1 punt) Bereken het steekproefgemiddelde van *efood* voor de gezinnen zonder kinderen ($\hat{\mu}_0$) en ook voor de gezinnen met kinderen ($\hat{\mu}_1$). Rapporteer $\hat{\mu}_0$ en $\hat{\mu}_1$.
 - (c) (3 punten) Definieer μ_0 en μ_1 als de populatiegemiddelden waarvan $\hat{\mu}_0$ en $\hat{\mu}_1$ de steekproefgemiddelden zijn. Toets de nulhypothese dat $\mu_0 = \mu_1$ tegenover de alternatieve hypothese dat $\mu_0 < \mu_1$. Gebruik de *t*-statistiek zoals gedefinieerd op p. 18 van de cursusslides.
2. Schat het regressiemodel

$$efood_i = \beta_0 + \beta_1 income_i + u_i$$

met kleinste kwadraten.

- (a) (1 punt) Rapporteer de schattingsresultaten in standaardvorm.
- (b) (2 punten) Toets de nulhypothese dat $\beta_1 = 0.05$ tegenover de alternatieve hypothese dat $\beta_1 > 0.05$.
- (c) (3 punten) Welke van de volgende uitspraken zijn juist/fout? Motiveer kort.
 - i. Volgens de schattingsresultaten stijgen de uitgaven aan voeding met (gemiddeld) 5.65 eurocent als het gezinsinkomen verdubbelt (onder de veronderstellingen van het model).
 - ii. De standaarddeviatie van *efood* is 1752 euro.
 - iii. De standaardfout van $\hat{\beta}_1$ is 0.00578, zeer dicht bij 0. Dit wijst er op dat β_1 waarschijnlijk 0 is.

3. Maak een variabele $nchi$ aan die gelijk is aan het aantal kinderen in het gezin. Schat het regressiemodel

$$efood_i = \beta_0 + \beta_1 income_i + \beta_2 nchi_i + \beta_3 nadult_i + u_i$$

met kleinste kwadraten (rapporteren hoeft niet).

- (a) (1 punt) Stel een 99.99% betrouwbaarheidsinterval op voor β_2 .
- (b) (2 punten) Leg uit waarom $\hat{\beta}_1$ gedaald is t.o.v. de regressie zonder $nchi$ en $nadult$.
- (c) (2 punten) Toets de nulhypothese dat $2\beta_2 = \beta_3$ tegenover de alternatieve hypothese dat $2\beta_2 \neq \beta_3$. Gebruik de F-toets.
- (d) (1 punt) In welke eenheid is β_3 uitgedrukt?
- (e) (1 punt) In welke eenheid is de standaardfout van de regressie uitgedrukt?
- (f) (2 punten) Beschouw een gezin van 2 volwassenen en 1 kind, met een inkomen van 30000 euro. Wat is, volgens de schattingsresultaten, de voorspelde waarde van $efood$ voor dit gezin?
- (g) (4 bonuspunten) Stel dat we $efood$ en $income$ zouden uitdrukken in 1000 euro (i.p.v. euro). Hoe zouden de coëfficiënten β_0, \dots, β_3 (en hun schattingen) dan veranderen? En R^2 ? En SER ? Leg uit.