

Naam: .....  
 Voornaam: .....

## Examen: Kwantitatieve Beleidsmethoden Deel Econometrie, 3e Bach T.E.W., Juni 2013

*Het examen is gesloten boek. Je mag een rekentoestel gebruiken. Schrijf je naam op elk antwoordenblad. Geef korte antwoorden. Gebruik indien nodig de achterkant van een blad. Je kan het laatste blad gebruiken als kladpapier. Het nietje mag niet worden losgemaakt. De tabel van de cumulatieve normale verdelingsfunctie (Tabel 1) vind je in bijlage.*

Een van de vragen behelst de interpretatie van een STATA-output (gelijkaardig als in de oefenzittingen), en een andere vraag de interpretatie van een regressieanalyse weergegeven in tabelvorm zoals in het boek (bvb pag 326, pag 408). Voorbeelden van andere vragen.

1. Geef een voorbeeld van een randomized controlled experiment.
2. Maak een schets van een t-verdeling met 3 vrijheidsgraden. Wat kan je zeggen over zijn Skewness en zijn Kurtosis?
3. Gegeven twee kansvariabelen  $X$  en  $u$ . Toon aan dat  $u$  en  $X$  ongecorreleerd zijn als  $E[u|X] = 0$ .
4. Onderstel dat het dagelijks rendement van een aandeel een normale verdeling volgt met gemiddelde 0,10 % en standaarddeviatie 2 %. Wat is dan de kans om een rendement te observeren van minstens 10% in absolute waarde?
5. Bereken de variantie van het steekproefgemiddelde van  $Y_1, \dots, Y_n$ , waar de kansvariabelen onderling gecorreleerd zijn met correlatie  $\rho > 0$ . Bespreek je uitkomst voor  $n$  groot.
6. Wanneer is een schatter efficiënter dan een andere schatter?
7. Noteer  $Y$  een variabele die gelijk is aan 1 wanneer je werkloos bent, en gelijk aan nul anders. Wat is de interpretatie van  $E(Y)$ ?
8. Software gebruikt standaard de  $t$ -verdeling bij het toetsen van hypothesen voor een verwachte waarde. Is die aanpak geldig indien de steekproefgrootte groot is en de verdeling niet normaal? Motiveer kort je antwoord.
9. De steekproefcorrelatie is een consistente schatter voor de correlatie tussen twee kansvariabelen. Wat betekent dit? Noteer deze eigenschap ook in wiskundige notatie.
10. Gegeven het regressiemodel  $Y_i = \beta_0 + \beta_1 X_i + u_i$ .
  - (10.1) Hoe interpreteer je de term  $u_i$  in dit model
  - (10.2) Geef de definitie van de kleinste kwadraten schatter  $\hat{\beta}_0$  en  $\hat{\beta}_1$  voor de parameters  $\beta_0$  en  $\beta_1$ .

(10.3) Wat betekent het als men zegt dat deze schatter asymptotisch normaal is.

(10.4) Hoe ziet een “confidence set” rond  $(\hat{\beta}_0, \hat{\beta}_1)$  eruit. Maak een tekening.

11. De formule voor de variantie van  $\beta_1$  in enkelvoudige regressie is, gebruik makende van de notaties van het boek,

$$\frac{1}{n} \frac{\text{Var}[(X_i - \mu_X)u_i]}{\text{Var}[X_i]^2}.$$

Hoe vereenvoudigt bovenstaande formule zich onder de extra assumptie van homoskedasticiteit.

12. De Gauss-Markov stelling zegt dat de kleinste kwadraten schatter (onder bepaalde voorwaarden) optimaal is in een bepaalde klasse van schatters. (i) In welke klasse van schatters? (ii) Wat bedoelt men met “optimaal”?
13. We hebben een steekproef van 500 klassen, en noteren  $Y$  het klasgemiddelde op een gestandaardizeerde test. Noteer  $X$  de klasgrootte. We beschouwen het regressiemodel  $Y_i = \beta_0 + \beta_1 X_i + u_i$ . (i) Geef twee voorbeelden van variabelen die niet in het model opgenomen zijn en die “omitted variable bias” kunnen veroorzaken. Motiveer goed waarom we de “omitted variable bias” kunnen bekomen. (ii) Geef een voorbeeld van een variabele die niet in het model opgenomen is, maar die geen “omitted variable bias” veroorzaakt.
14. Waarom gebruikt men soms de “adjusted R-kwadraat,  $\bar{R}^2$ ”, in plaats van de gebruikelijke R-kwadraat  $R^2$  in regressieanalyse?
15. Noteer  $Y$  het maandelijks inkomen in Euro van een werknemer,  $X$  het aantal jaren werker-  
varing, en  $Age$  zijn leeftijd. Op basis van een steekproef van grootte 100 bekomen we

$$\hat{Y} = \underset{(79.4)}{905} + \underset{(7.6)}{30} X + \underset{(7.6)}{12} Age$$

- (15.1) Welke variabelen zijn significant?
- (15.2) Constueer benaderend een 95% betrouwbaarheidsinterval voor de parameter die hoort bij  $Age$ .
- (15.3) Interpreteer de waarde van de geschatte coëfficiënt die hoort bij  $X$ .
- (15.4) Hoe verandert je antwoord op bovenstaande vraag indien  $Y$  de logaritme van het maandelijks inkomen zou voorstellen.
- (15.5) Je wil in dit model een categorische verklarende variabele “Opleidingsniveau” opnemen, die aanduidt of iemand een secundair, hogeschool, of universitair diploma heeft. (i) Hoe doe je dit? (ii) Hoe toets je of “Opleidingsniveau” significante verklarende kracht heeft.
16. Gegeven het regressiemodel  $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i$ . Onderstel dat  $E[u_i | X_{1i}, X_{2i}] = 7X_{2i}$ . Bereken dan het effect van een verhoging van een eenheid van de eerste verklarende variabele (de tweede verklarende variabele constant gehouden) op de verwachte waarde van de responsvariabele.

17. Men kan aantonen dat de variantie van de kleinste kwadraten schatter in een model met twee verklarende variabelen  $X_1$  en  $X_2$  proportioneel is aan

$$1/(1 - \rho_{X_1, X_2}^2),$$

Gebruik bovenstaande formule om het begrip “multicollineariteit” uit te leggen.

18. Noteer  $Y$  het maandelijks inkomen in Euro van een werknemer, en  $X$  het aantal jaren werkervaring. Op basis van een steekproef van grootte 100 bekomen we

$$\hat{Y} = \underset{(79.4)}{905} + \underset{(7.6)}{+12}X + \underset{(0.04)}{-0.6}X^2$$

- (18.1) Verkies je het lineaire of het kwadratische model?
- (18.2) Wat is het effect van een stijging van 1 jaar van het aantal jaren werkervaring op het verwachte maandelijks inkomen?
- (18.3) Je wil in dit model een verklarende variabele “Geslacht” opnemen, en je denkt dat er interactie kan zijn tussen “Geslacht” en “aantal jaren werkervaring”. (i) Hoe doe je dit? (ii) Hoe toets je of er een significante interactie is.
19. Beschouw een fixed effects regression met “entity” en “period” fixed effects. We hebben slechts twee tijdsperiodes ( $T = 2$ ), maar wel  $n = 100$  entiteiten. Het model is, met de notaties uit de cursus:

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \alpha_i + \lambda_t + u_{it}.$$

- (19.1) Toon aan hoe je  $\beta_1$  kan schatten door een “before and after” of “differences” vergelijking.
- (19.2) Een andere manier om het model te schatten is om dummy variabelen in te voeren. Hoeveel dummy variabelen moet je dan aanmaken?
- (19.3) De storingsterm  $u_{it}$  bevat de “omitted variables” die .... (vul aan)
- (19.4) Om de standard errors te berekenen in regressie met panel data gebruik je best de optie “clustered standard errors”. Waarom is dit nodig? Wat zijn de verschillende clusters?
20. Noteer  $Y = 1$  als een student slaagt voor al zijn vakken in eerste Bach, en  $Y = 0$  anders. Als regressoren nemen we  $X_1 = \text{geslacht}$  (1 voor jongens, 0 voor meisjes),  $X_2 = \text{behaald resultaat (op 100)}$  in middelbaar onderwijs. Op basis van een steekproef schatten we het Probit model

$$P(Y = 1 | X_1, X_2) = \Phi\left(\underset{(7.0)}{-3.2} + \underset{(0.01)}{-0.04} X_1 + \underset{(0.022)}{0.07} X_2\right)$$

- (20.1) Welke regressoren zijn significant? Interpreteer het teken van de coëfficiënten.
- (20.2) Wat is het geschatte verschil in slaagkans tussen jongens en meisjes, voor  $X_2 = 70$ ?
- (20.3) Wat is het geschatte effect van een verhoging van  $X_2$  met 1 punt op de slagingskans voor meisjes. En voor jongens?
- (20.4) Hoe zou je het geschatte verschil in slaagkans tussen jongens en meisjes kunnen samenvatten in één enkel getal?
- (20.5) Wat is het voordeel van het Probit model ten op zichte van een lineair model.
- (20.6) Wat is het verschil tussen een Probit en een Logit model?

**Naam:** .....

**Voornaam:** .....

KLAD