

Vraag over studenten die dezelfde cursus in het Engels en in het Nederlands volgden. Dan kreeg je hun punten in het 1ste en 2de semester (dus 4 kolommen) en dan het verschil berekenen. Dus je moest eerst 4 gemiddelden berekenen (een gemiddelde per kolom) en dan de formule toepassen (gelijk een vaccin/controle model)

- 4,9
- $4,5 \times 2 (6x) = (Y_{\text{treat after}} - Y_{\text{treat before}}) - (Y_{\text{controle after}} - Y_{\text{controle before}})$  met Y is telkens het gemiddelde van zo n rij

Vraag met F-test homoskedasticiteit

- 90,5 x3

Vraag kijken of variabelen significant is bij logit: -2xloglikelihood en dan p-waarde berekenen via CHISQ.DIST.RT(..., 2?) ja ging van 6 naar 8 dus 2 (denk ik). Was tussen 0,2 en 0,5 denk ik  $\Rightarrow$  ja denk ik ook!

Moest hier 1 of 2 staan? Hier was het moeilijke of je een 1 of 2 moest zetten, er kwam een variabele bij (dit was 'mentalhealth' met 0= normaal, 1=beetje depressief en 2=depressief)

- Minder dan 0,05 x2
- tussen 0.05 en 0.25
- tussen 0,2 en 0,5 x3
- Groter ...
- 

Vraag over de standaard deviatie van die ANOVA tabel: SSB, SSW en DFB waren gegeven , 17 en 14 ong

- 2.8 x3
- 1.7 x4  $\rightarrow$  was dit denk ik want ge moest nog de vierkantswortel van 2.8 doen.

Vierkantswortel van 2.8 geeft 1.67, afgerond dus 1.7

- 1,6 x2
- 1,3

Hoeveel de prijs van een huis stijgt als er zwembad is:  $\ln(\text{prijs}) = \ln(\text{opp}) + \text{pool} + \text{zeezicht}$

- prijs steeg met 10.4% door de verdubbeling van opp (7x)
- prijs steeg 189% door een zwembad ? nee was toename in procenten dus een Verdubbeling is +100% dus stijging van 97% ofzo (8x) dat hoort toch bij andere vraag? of 10300 euro

dat was toch niet significant? als je beta van pool door de st err van pool zou delen en dan de f test doen, kwam je iets van 0,1 uit? x5 Denk inderdaad dat het niet significant was.

was dat deze vraag van waterfront of pool?

Wat is de teststatistiek of p-waarde (weet ik niet goed meer) als je test dat waterfront gelijk is aan 0,2 ofzo?

Hoeveel stijgt de prijs van woning indien u oppervlakte maal 2 gaat:  $\ln(\text{Price}) = \dots + 0,9 \cdot \ln(\text{Area})$

- 189% x3
- 189.000
- 91,3%  $\rightarrow$  x10 = 100% stijging

Vraag over die grote en kleine ondernemingen, welke nulhypothese voer je uit? je moest testen of het loon significant verschillend was voor mannen en vrouwen in een grote organisatie:  $Y = B_0 + B_1 \times \text{geslacht} + B_2 \times \text{organisatie} + B_3 \times \text{organisatie} \times \text{geslacht}$

beta 2 = beta 3

beta 2 = 0

beta 3 = 0 x11

beta 1 + beta 3 = 0 x10  $\rightarrow$  vrij zeker aangezien het alleen in grote organisatie was (maar dan test je toch niet of ze beiden gelijk aan 0 zijn? **Maar dan test je toch niet of ze significant verschillend zijn van 0 maar B1= -B3??**) hoezo beta 1 + beta 3 juist dat toch super raar? Is het niet zoals vraag 16 op voorbeeldexamen en dus B3=0?

Vraag over de nulhypothese berekenen bij  $H_0 : \beta_2 = 0,2$

coeff was 0.235 met een sd van 0.04  $\rightarrow t = 0.875$

wat was de p-waarde

- $p > 0.25$  x14 -
- 
- $p < 0.1$

$\beta_2 = \beta_3$  hoe schatten  $\rightarrow$  invoegen van B6 term zoals in de les gezien :  $-B_6X + B_6X$

- $C: \alpha_2 = 0$  x11
- $\alpha_3 = 0$

Nog een gelijkaardige vraag als die hierboven maar met interactieeffect, interactie term toevoegen

- $C: H_0 : \beta_5 = 0$  x4

Vraag met die grafiek waar ge 5 observaties had met 4 entiteiten

- beta 1 = -2 x9  $\rightarrow$  gewoon helling berekenen tussen die punten easy
- beta 1 = -1
- beta 1 = 1
- beta 1 = 2

Vraag over de rho waar je variantie tussen de entiteiten vergelijkt met totale variantie

- 1 x1
- 0
- 0,5
- -1 1x

Hoe kan je het gemakkelijkste zien of u residu homoscedastisch was

- residuplot bekijken x15
- variantie van de residus berekenen
- histogram van de residus
- residus van groot naar klein

Een vraag met deze opties

- paneldata
- regressie
- fixed model x2
- random model x11 → want ze wouden kijken naar ras, betere score

jongen/meisje en dat kan niet geschat worden in FE wel in RE

Probit model waarbij de opa 80 jaar was, 15 keer naar de dokter geweest was etc. wat was de kans dat de opa opgenomen wordt in het ziekenhuis

had een z waarde van 1.611 → fout denk ik jij hebt 2 ingevuld bij da van de dokter bezoeken maar je moest uiteindelijk toch 1 invullen omdat dat een kwalitatieve dummyvariabele was: dan kwam je op 0,15 → dan dat in excel en kreeg je 0,55  
probit volgt de normale verdeling en geen t verdeling?  $\text{normalcdf}(0.15; 10^{99}) = 0.44$ ? ja dat klopt maar dan kwam ik 0,56 uit

| <div> <div>▼</div> <div>:</div> <div>✕</div> <div>✓</div> <div><math>f_x</math></div> <div>=NORM.S.VERD(0,15;WAAR)</div> </div> |   |    |   |   |   |
|---------------------------------------------------------------------------------------------------------------------------------|---|----|---|---|---|
| A                                                                                                                               | B | C  | D | E | F |
| 0,559618                                                                                                                        |   |    |   |   |   |
| -81%                                                                                                                            |   | x2 |   |   |   |
| -44% ong                                                                                                                        |   | x5 |   |   |   |
| -56% ong                                                                                                                        |   | x9 |   |   |   |
| -16% ong                                                                                                                        |   |    |   |   |   |

Vraag met 2 stellingen

1. Als de steekproef groter wordt neemt de  $\text{var}(\beta_1)$  toe
  2. Als de varianties toenemen neemt ook de  $\text{var}(\beta_1)$  toe
- alleen eerste stelling is juist

- alleen 2de stelling is juist x1
- allebei juist
- allebei fout x11 → i. dan neemt varB1 af ii. dan neemt varB1 af

Vraag of de coëfficiënt van vrouw verandert als je de extra variabele toevoegt van niet en wel winstgevend ondernemingen

- coëff neemt toe (dus minder neg) x11
- coëff neemt af (dus meer neg) x2
- coëff veranderd niet SE wordt groter
- coëff veranderd niet SE wordt kleiner x1

redenering: (antwoord b is hieronder juist btw)

5. Wat is het gevolg van imperfecte multicollineariteit?
- A. De schatters blijven zuiver en hun standaardfout wordt kleiner.
  - B. De schatters blijven zuiver en hun standaardfout wordt groter.
  - C. De schatters zijn niet meer zuiver en hun standaardfout wordt kleiner.
  - D. De schatters zijn niet meer zuiver en hun standaardfout wordt groter.

Iemand van 40 jaar wordt een half jaar ouder, wat is het effect van leeftijd

- 8 x8 → ½ invullen ipv 1 en ook rekening houden met leeftijd<sup>2</sup>
- 16 x1
- 24? x2
- 32? x1

vraag over de VIF bereken van een beta

- 1.029 x1
- 48,3 x8 → correlatie was 0,98 ofzo dus  $1-R^2 = 0,02...$  en  $VIF = 1/0,02... = 48,...$

- iets anders x1 → correlatie moest nog <sup>2</sup> en dan zoals hierboven
- 

iets met kleinste kwadratenmethode

t waarde van leeftijd?

- 275
- 90 x4
- 5 x1
- 231,... x1

Iets van aantal vrijheidsgraden als er 1080 observaties zijn en het is een student T verdeling. gwn  $n-k-1$ ? was een t test dus was het niet  $n-2$  dus 1078? ja heb ik ook. t test is toch  $n-2$ ?

-1075 x7 → waarom zou dit kloppen?  $n - k - 1$   
-1074  
-1078 x 3 → ttest gewoon  $n-2$  toch?? correct me if i am wrong  
-1079

## PC DEEL

### vraag 1:

Ging over een mexicaanse prostituee die allemaal karakteristieken had (attractive, school gedaan, condoom of niet) en dan waar het gebeurde (bar, straat of other) en dan de karakteristieken van de klant (rijk, dronken en regular visit).

WELKE VRAAG STOND OP HOE VEEL PUNTEN? ik denk 1.75 2.75 2.5 3 ofzo iets

Ze vroegen aantal en gemiddelde transacties met condoom. → ik had ttest price, by (nocondom) unequal en dan uit het resultaat van dat commando kon je alles aflezen

Geen condoom: 80      iets in 300

Wel condoom: 260      iets in 400

→ wow ik had 66 geen connie en 688 wel, ik ook, ik ook (we bedoelen hetzelfde denk ik) → ik vond t wel vreemd da gemiddeld zonder minder kost dan met ma bon

Bl opstellen voor ui -u0 (als je ttest had gedaan zag je combined Bl staan)

Moest je dan hier gewoon de difference Bl die stata gaf opschrijven, of moest je zelf nog aanpassen want hij vroeg verschil (1-0), terwijl stata gaf verschil in andere vorm namelijk (0-1), ofzoiets ⇒ ja je moest dit zelf in de omgekeerde richting interpreteren!!

Coëfficiënt interpreteren en in odds

vraag wrm ge beter paneldata gebruikte ipv cross-sectionele data

vraag waarom omitted bias bij fe optreden

Je kreeg karakteristieken van de persoon en met behulp van margins, at(...) dan interpreteren wat de verschillen waren (als er één variabele anders was, denk dat dit heeft school gedaan of niet). Daarna moest je dan zeggen of het verschil significant was ja of nee → welke test? ik deed lrttest → ik deed dit met marginale effect van school dus dydx en dan kijken of dit verschil significant was (ah da kan idd, lrttest was een gokje bij mij)

hausman test

Iemand die hier ook vond dat opgave niet goed gesteld was en onduidelijk vond over welke variabele je moest opnemen voor bij u FE en RE voor hausman test??? **x9**!!!!!! → jep ik heb dit wel gevraagd en ge moest WEL alle var opnemen anders werkte uw hausmantest niet, dan kwam ge  $p=0,00$  uit en dus fe verkiezen MAAR er stond toch duidelijk in de opgave om enkel te

werken met de variabele die gemeenschappelijk waren, dus enkel diegene die zowel in FE als in RE model stonden? - stond idd onduidelijk in de opgave ja da ontken k ni, ma kebt em echt gevraagd. Mss de prof is over sturen? Heb ik gedaan, prof zei dat je de gemeenschappelijke variabelen moest opnemen. Als dit niet werkte in stata krijg je nogsteeds punten als je de juiste commando's hebt opgeschreven tijdens het examen!

enige verschil bij een RE is dat je de karakteristieken van sekswerkers meetelt dus die moesten erbij → nee, want er stond in e opgave gebruik de variabelen die in BEIDE modellen gebruikt worden en die karakteristieken van de sekswerkers stonden niet in in het FE model