

Økonomisk Kandidateksamen
Likelihoodteori
Eksamenstermin: vinter 2005/2006
2/1 2006
3 Timer

Alle spørgsmål ønskes besvaret
Alle sædvanlige hjælpemidler er tilladt

Betragt en sædvanlig simpel lineær regressionsmodel uden konstantled

$$y_i = \beta x_i + e_i$$

hvor restleddene e_i er uafhængige normalfordelte med middelværdi nul og samme varians σ^2 .

Denne model suppleres med en model for heteroskedasticitet, hvori variationen afhænger af niveauet på følgende måde

$$\text{var}(e_i) = \text{var}(y_i) = \sigma^2(\beta x_i)^\gamma.$$

Bemærk at restriktionen $\gamma = 0$, svarer til modellen med den almindelige homoskedasticitetsantagelse. For n uafhængige observationer af y_i og x_i fra denne model er den simultane tæthedsfunktion givet ved

$$\prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\sigma^2(\beta x_i)^\gamma}} \exp\left(-\frac{(y_i - \beta x_i)^2}{2\sigma^2(\beta x_i)^\gamma}\right)$$

a)

Scorevektoren er den tre-dimensionale rækkevektor $\mathbf{q} = (q_\beta, q_\sigma, q_\gamma)$. Vis at komponenten q_γ er givet ved

$$q_\gamma = \frac{\partial}{\partial \gamma} \log L = \sum_{i=1}^n \left[-\frac{1}{2} \log(\beta) - \frac{1}{2} \log(x_i) + \frac{(y_i - \beta x_i)^2 (\beta x_i)^\gamma \log(\beta x_i)}{2\sigma^2(\beta x_i)^{2\gamma}} \right]$$

Husk at

$$\frac{d}{d\gamma} x^\gamma = x^\gamma \log(x)$$

I en konkret analyse benyttes data, hvor y betegner antallet af personlige computere pr. 1000 indbyggere og x betegner urbaniseringsgraden i procent for $n = 134$ lande, hvorfra data er vist på figuren. Først estimeres en lineær model med restriktionen $\gamma = 0$ til

$$\beta = 2.041 \quad \sigma^2 = 15081 \quad (\gamma = 0)$$

Den dertil svarende rette linie er indlagt på figuren. Loglikelihoodfunktionens værdi for disse parameterværdier er - 834.3.

Maksimum likelihood estimererne, når alle parametre estimeres frit, er

$$\beta = 0.59 \quad \sigma^2 = 0.0027 \quad \gamma = 4.30$$

idet maksimum af loglikelihoodfunktionen er - 774.1. Den dertil svarende rette linie er også indlagt på figuren.

Desuden estimeres variansen (med parametrene i rækkefølgen β, σ^2, γ) ved $-Q^{-1}$ til

	beta	sigma2	gamma
beta	0.007377	-0.00007335	-0.007651
sigma2	-0.00007335	0.000004638	-0.0003419
gamma	-0.007651	-0.0003419	0.05389

b)

Test hypotesen $H_0: \gamma = 0$ ved et Likelihoodratio test.

c)

Test hypotesen $H_0: \gamma = 0$ ved et Wald test.

d)

Et Lagrange Multiplikator test for hypotesen $H_0: \gamma = 0$ bygger på værdien af scorevektoren. Den er i dette tilfælde

$$q_\gamma = 19.87$$

idet værdierne $\beta = 2.041$, $\sigma^2 = 15081$ og $\gamma = 0$ er indsat.

Som approksimation til variansen anvendes $\Sigma \mathbf{q}^T \mathbf{q}$, hvis inverse for disse parameterverdier beregnes til

	beta	sigma2	gamma
beta	0.002569	46.50	-0.0009063
sigma2	46.50	23250952.0	-316.3
gamma	-0.0009063	-316.3	0.004359

Gennemfør scoretestet (Lagrange Multiplikator testet) for hypotesen $H_0: \gamma = 0$. Redegør for, hvorfor der kan være så store forskelle mellem konklusionen på Lagrange Multiplikatortestet på den ene side og Likelihood ratio- og Waldtestene på den anden side.

e)

En alternativ tilgang til Lagrange Multiplikator testet bygger på, at scorevektoren q_γ for $\gamma = 0$ er givet ved formelen

$$q_\gamma = \sum_{i=1}^n \left[-\frac{1}{2} \log(\beta) - \frac{1}{2} \log(x_i) + \frac{(y_i - \beta x_i)^2 \log(\beta x_i)}{2\sigma^2} \right]$$

$$= \sum_{i=1}^n \frac{(y_i - \beta x_i)^2 \left(\log(x_i) - \frac{1}{n} \sum_{i=1}^n \log(x_i) \right)}{2\sigma^2}$$

Denne udregning kan tages for givet og skal IKKE eftervises. Den viser, at Lagrange Multiplikator teststørrelsen er proportional med den estimerede hældningskoefficient i en regressionsanalyse af de kvadrerede residualer $(y_i - \beta x_i)^2$ som responsvariabel og $\log(x)$ som forklarende variabel. Er et almindeligt test i regressionsanalysen for hypotesen om at denne hældningskoefficient er nul teoretisk tilfredsstillende? Testet giver en t-teststørrelse på 4.45 og dermed en χ^2 størrelse på 19.81.

f)

Der gælder, at da residualerne $(y_i - \beta x_i)$ er normalfordelte med middelværdi nul og varians σ^2 , er varianserne på de kvadrerede residualer $\text{Var}((y_i - \beta x_i)^2) = 2\sigma^4$. Vis ud fra denne oplysning, at variansen af q_γ er givet ved formelen

$$\text{var}(q_\gamma) = \frac{1}{2} \sum_{i=1}^n \left(\log(x_i) - \frac{1}{n} \sum_{i=1}^n \log(x_i) \right)^2 = 17.90$$

og test herudfra hypotesen $H_0: q_\gamma = 0$ og dermed $H_0: \gamma = 0$.

Figur

Antallet af personlige computere pr. 1000 indbygger for 134 lande afsat mod landenes urbaniseringsgrad. De indtegnede kurver er i rækkefølge fra neden: Nedre 95% sikkerhedsgrænse, Regressionslinien estimeret med hensyntagen til heteroskedasticiteten, Regressionslinien estimeret uden hensyntagen til heteroskedasticiteten og Øvre 95% sikkerhedsgrænse. Sikkerhedsgrænserne er beregnet i modellen med hensyntagen til heteroskedasticiteten.

