

Empiriske Metoder - Anden halvdel

Sommer 2003

OPGAVE 2

Figuren viser andelen af høj-teknologiske produkter i eksporten i $n = 105$ lande sammenholdt med disse landes elektricitetsforbrug pr. indbygger. På figuren ses en tendens til, at de to variable samvarierer, men billedet er sløret af en kraftig heteroskedasticitet. Man kan derfor betragte en model af formen

$$y_i = \beta x_i \times e_i$$

med $\beta > 0$, hvor fejleddet i regressionsmodellen indgår multiplikativt og er uafhængige normalfordelte med middelværdi 1 og varians σ^2 . Der gælder ifølge modellen, at sandsynlighedsfordelingen af andelen høj teknologisk eksport, y , givet elektricitetsforbruget pr. indbygger, x , opfylder

$$E[y_i] = \beta x_i \quad \text{og} \quad \text{var}(y_i) = \beta^2 x_i^2 \sigma^2.$$

Antag nu, at der foreligger n observationer af y_i og x_i . I denne model er den simultane tæthedsfunktion givet ved

$$\begin{aligned} & \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\beta^2 x_i^2 \sigma^2}} \exp\left(-\frac{1}{2} \frac{(y_i - \beta x_i)^2}{\beta^2 x_i^2 \sigma^2}\right) \\ &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\beta^2 x_i^2 \sigma^2}} \exp\left(-\frac{y_i^2}{2\beta^2 x_i^2 \sigma^2} - \frac{1}{2\sigma^2} + \frac{y_i}{\beta x_i \sigma^2}\right). \end{aligned}$$

Scorevektoren i denne model er rækkevektoren $\mathbf{q} = (q_\beta, q_\sigma)$, hvor

$$q_\beta = -\frac{n}{\beta} + \sum \frac{y_i(y_i - \beta x_i)}{\beta^3 x_i^2 \sigma^2}$$

$$q_\sigma = -\frac{n}{2\sigma^2} + \frac{1}{2} \sum \frac{(y_i - \beta x_i)^2}{\beta^2 x_i^2 \sigma^4}$$

Dette skal IKKE vises.

For det konkrete datamateriale estimeres parametrene β og σ^2 ved PROC NLP i SAS til

$$\hat{\beta} = 11.7575 \quad \hat{\sigma}^2 = 4.7191 ,$$

og ud fra matricen \mathbf{Q} med parametrene i rækkefølgen β, σ^2 beregnes

$$-\mathbf{Q}^{-1} = \begin{bmatrix} 6.1946 & -4.9721 \\ -4.9721 & 4.4148 \end{bmatrix} .$$

mens maksimum af log likelihoodfunktionen er -531.1798.

a)

Test hypotesen $\beta = 15$ ved et Wald test og et Likelihood ratio test, idet maksimum for log likelihoodfunktionen beregnet under denne hypotese er -532.0165.

b)

For Danmark er $x = 6.03$. Estimer forudsigelsen for andelen af høj teknologisk eksport fra Danmark, $\beta \times 6.03$, og angiv den asymptotiske fordeling af dette estimat, herunder en talværdi for den asymptotiske varians.

c)

Beregn talværdien af estimatet for $\beta^2 \times 6.03^2 \times \sigma^2$, der er variansen på andelen høj teknologisk eksport fra Danmark. Angiv den asymptotiske fordeling af denne størrelse og beskriv hvorledes talværdien af den asymptotiske varians kan beregnes. Selve talværdien af kræves ikke beregnet.

d)

Beregn $Q_{\sigma\sigma}$, der er andenordens partielle afledede af log likelihoodfunktionen efter parameteren σ^2 og vis at

$$H_{\sigma\sigma} = \frac{n}{2\sigma^4}$$

e)

Den samlede 2×2 matrix \mathbf{H} med parametrene i rækkefølgen β, σ^2 er givet ved

$$\mathbf{H} = \begin{bmatrix} \frac{2n\sigma^2 + n}{\beta^2\sigma^2} & \frac{n}{\beta\sigma^2} \\ \frac{n}{\beta\sigma^2} & \frac{n}{2\sigma^4} \end{bmatrix} .$$

Beskriv hvorledes en formel for LM testet af hypotesen $\beta = 15$ kan stilles op, idet estimatet for σ^2 under hypotesen betegnes værdien $\tilde{\sigma}^2$.

