STK1110 – Oblig 1

Jonas Gahr Sturtzel Lunde (`jonassl`)

October 2, 2018

Oppgave 2

a)

Ettersom ligning (1) fra oppgaven er en t-fordeling med $n-1$ frihetsgrader, vet vi at den følger

$$P\left(t_{\alpha/2,\,n-1} < \frac{\bar{X}-\mu}{S/\sqrt{n}} < t_{1-\alpha/2,\,n-1}\right) = 1-\alpha \tag{1}$$

der $t_{\alpha/2,\,n-1}$ og $t_{1-\alpha/2,\,n-1}$ er $\alpha/2$ og $1-\alpha/2$ persentilene til en t-fordeling med $n-1$ frihetsgrader. Løser vi ulikheten inni parantesen for μ får vi at

$$\bar{X} - t_{\alpha/2,\,n-1}\cdot\frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{1-\alpha/2,n-1}\cdot\frac{S}{\sqrt{n}}$$

som er $100(1-\alpha)\%$ konfidensintervallet til μ.

b

Ettersom ligning (1) fra oppgaven er kjikvadrat-fordelt med $n-1$ frihetsgrader, vet vi at den tilfredsstiller

$$P\left(\chi_{\alpha/2,\,n-1} < \frac{(n-1)}{\sigma^2}S^2 < \chi_{\alpha/2,\,n-1}\right) = 1-\alpha \tag{2}$$

der $\chi_{\alpha/2,\,n-1}$ og $\chi_{1-\alpha/2,\,n-1}$ er $\alpha/2$ og $1-\alpha/2$ persentilene til en kjikvadrat-fordeling med $n-1$ frihetsgrader. Løser vi ulikheten inni parantesen for σ får vi at

$$\sqrt{\frac{(n-1)}{\chi_{\alpha/2,\,n-1}}}S < \sigma < \sqrt{\frac{(n-1)}{\chi_{1-\alpha/2,\,n-1}}}S \tag{3}$$

Oppgave 3

a)

$$F(x) = \int_{-\infty}^{x} f(x)\,\mathrm{d}x = \int_{\kappa}^{x} \theta\kappa^\theta x^{-\theta-1}\,\mathrm{d}x = \left[\theta\kappa^\theta\frac{x^{-\theta}}{-\theta}\right]_\kappa^\theta = 1-\left(\frac{\kappa}{x}\right)^\theta$$

$$F(x) = 1-\left(\frac{\kappa}{x}\right)^\theta = \frac{1}{2} \;\Rightarrow\; \frac{\kappa}{x} = \frac{1}{2}$$

b)

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\,\mathrm{d}x = \int_{\kappa}^{\infty} \theta\kappa^{\theta} x^{-\theta}\,\mathrm{d}x = \left[\theta\kappa^{\theta}\frac{x^{-\theta+1}}{-\theta+1}\right]_{\kappa}^{\infty} = 0 - \theta\kappa^{\theta}\frac{\kappa^{-\theta+1}}{-\theta+1} = \frac{\theta\kappa}{\theta-1}$$

c)

Vi omskriver den gitte definisjonen til å gi X som definisjon av Y:

$$\begin{aligned} Y = 2\theta[\ln(X) - \ln(\kappa)] &= 2\theta\ln(X/\kappa) \\ e^{Y/2\theta} &= X/\kappa \\ X &= \kappa e^{Y/2\theta} \end{aligned}$$

Vi setter dette inn i uttrykket vårt for den kummulative fordelingsfunksjonen:

$$F(Y) = 1 - \left(\frac{\kappa}{\kappa e^{y/2\theta}}\right)^{\theta} = 1 - e^{-y/2}$$

$$f(y) = F'(y) = \frac{1}{2}e^{-y/2} = \frac{1}{2^{2/2}\Gamma(2/2)}x^{2/2-1}e^{-x/2}$$

som vi ser er en kjikvadratfordeling med 2 frihetsgrader.

d)

$$E(X) = \frac{\theta\kappa}{\theta-1} = \bar{X} \;\Rightarrow\; \theta\kappa = \theta\bar{X} - \bar{X} \;\Rightarrow\; \theta = \frac{\bar{X}}{\bar{X}-\kappa}$$

Momentestimatoren til θ er altså $\hat{\theta} = \frac{\bar{X}}{\bar{X}-\kappa}$.

e)

Sannsynligheten for at kombinasjonen av tilfeldige variable $X_1, X_2, ..., X_n$ i n uavhengige forsøk blir $x_1, x_2, ..., x_n$ vil være produktet av de individuelle sannsynlighetene

$$f(x_1, x_2, ..., x_n; \theta) = \prod_{i=1}^{n} \theta\kappa^{\theta}\left(\frac{1}{x_i}\right)^{\theta+1} = \theta^n\kappa^n\theta\left(\prod_{i=1}^{n}(x_i)^{-1}\right)^{\theta+1}$$

Vi skal finne maks-verdien til denne fordelingen. Vi tar først logaritmen av fordelingen, ettersom den deler toppunkt med sin logaritme, og dette er enklere å regne med.

$$\begin{aligned} \ln[f(x_1, x_2, ..., x_n; \theta)] &= \ln(\theta^n) + \ln(\kappa^n\theta) + \ln\left[\left(\prod_{i=1}^{n} x_i^{-1}\right)^{\theta+1}\right] \\ &= n\ln(\theta) + n\theta\ln(\kappa) + (\theta+1)\left[\sum_{i=1}^{n} -\ln(x_i)\right] \end{aligned}$$

Deriverer, setter lik 0, og løser for θ:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\theta}\ln[f(x_1, x_2, ..., x_n; \theta)] &= \frac{n}{\theta} + n\ln(\kappa) - \sum_{i=1}^{n}\ln(x_i) = 0 \\ n &= \theta\left[\sum_{i=1}^{n}\ln(x_i) - n\ln(\kappa)\right] \\ \hat{\theta} &= \frac{n}{\sum_{i=1}^{n}\ln(x_i) - n\ln(\kappa)} \end{aligned}$$

som da er maximum likelihood estimatoren for θ.

f)

Vi har en ny stokastisk variabel

$$Y = 2n\frac{\theta}{\hat{\theta}} = 2\theta\left[\sum_{i=1}^{n}\ln(x_i) - n\ln(\kappa)\right] = \sum_{i=1}^{n} 2\theta[\ln(x_i) - \ln(\kappa)] = \sum_{i=1}^{n} Z$$

hvor $Z \sim \chi^2_2$, altså Z er kjikvadratfordelt med 2 frihetsgrader.

Fra s.316 i læreboka har vi at summen av kjikvadratfordelinger selv er en kjikvadratfordeling, med ny frihetsgrad lik summen av frihetsgradene, som betyr at Z er en kjikvadratfordeling med $2n$ frihetsgrader:

$$\sum_{i=1}^{n} Z \sim \chi^2_{2n}$$

g)

$$E[\hat{\theta}] = 2n\theta \cdot E[Y^{-1}]$$

Ettersom Y er en kjikvadratfordelt tilfeldig variabel med $\nu = 2n$ frihetsgrader, setter vi inn for dens forventningsverdi, definert i ligning (5) i oppgaven

$$E[\hat{\theta}] = 2n\theta \cdot \frac{2^{-1}\Gamma\left(\frac{2n}{2} - 1\right)}{\Gamma\left(\frac{2n}{2}\right)} = n\theta\frac{\Gamma(n-1)}{\Gamma(n)} = n\theta\frac{(n-2)!}{(n-1)!} = \theta\frac{n}{n-1}$$

For å finne variansen finner vi først $E[\hat{\theta}^2]$.

$$E[\hat{\theta}^2] = 2^2n^2\theta^2 \cdot \frac{2^{-2}\Gamma\left(\frac{2n}{2} - 2\right)}{\Gamma\left(\frac{2n}{2}\right)} = n^2\theta^2\frac{\Gamma(n-2)}{\Gamma(n)} = n\theta\frac{(n-3)!}{(n-1)!} = \theta^2\frac{n^2}{(n-1)(n-2)}$$

Vi bruker da definisjonen

$$\begin{aligned} V[\hat{\theta}] = E[\hat{\theta}^2] - E[\hat{\theta}]^2 &= \theta^2\frac{n^2}{(n-1)(n-2)} - \theta^2\frac{n^2}{(n-1)^2} \\ &= \theta^2\frac{n^2(n-1) - n^2(n-2)}{(n-1)^2(n-2)} = \theta^2\frac{n^2}{(n-1)^2(n-2)} \end{aligned}$$

h)

Estimatoren er ikke unbiased, fordi forventningsverdien ikke er θ. Biasen til estimatoren er definert som

$$Bias(\hat{\theta}) = E[\hat{\theta}] - \theta = \frac{1}{n-1}\theta$$

Vi ser at vi kan gjøre estimatoren unbiased ved å gange den med $n - 1/n$:

$$\frac{n-1}{n}\hat{\theta} = \frac{n-1}{\sum_{i=1}^{n}\ln(X_i) - n\ln(\kappa)}$$