STK1110 - Statistiske metoder og dataanalyse høsten 2019
Løsningsforslag til oppgaver i læreboka for uke 37

Vinnie Ko & Ørnulf Borgan

September 11, 2019

Lærebok: *Modern Mathematical Statistics with Applications (2nd ed.)* - Devore & Berk

Exercise 7.42

a) Punktsannsynligheten til Poissonfordelingen er (for $x = 0, 1, 2, \ldots$)

$$f(x;\lambda) = \frac{\lambda^x e^{-\lambda}}{x!}.$$

Dermed er

$$\ln f(x;\lambda) = x\ln\lambda - \lambda - \ln(x!)$$

$$\frac{\partial}{\partial\lambda}\ln f(x;\lambda) = \frac{x}{\lambda} - 1$$

$$\frac{\partial^2}{\partial\lambda^2}\ln f(x;\lambda) = -\frac{x}{\lambda^2}$$

Vi finner nå Fisher informasjonen i én observasjon på to måter, nemlig ved definisjonen på side 371 i læreboka (metode 1) og ved resultatet i formel (7.9) på side 372 i læreboka (metode 2).

i) Metode 1:

$$I(\lambda) = \mathrm{V}\left(\frac{\partial}{\partial\lambda}\ln f(X,\lambda)\right) = \mathrm{V}\left(\frac{X}{\lambda} - 1\right) = \frac{\mathrm{V}(X)}{\lambda^2} = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}$$

ii) Metode 2:

$$I(\lambda) = -\mathrm{E}\left(\frac{\partial^2}{\partial\theta^2}\ln f(X,\theta)\right) = -\mathrm{E}\left(-\frac{X}{\lambda^2}\right) = \frac{E(X)}{\lambda^2} = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}$$

b) Cramér-Raos nedre grense for variansen til en forventningsrett estimator er gitt på side 374 i læreboka. Fra punkt a) har vi at $I(\lambda) = \frac{1}{\lambda}$. Dermed er Cramér-Raos nedre grense $\frac{1}{nI(\lambda)} = \frac{\lambda}{n}$

c) Score funksjonen er definert på side 373 i læreboka; se også øverst side 376. Her blir score funksjonen

$$S(\lambda) = \sum_{i=1}^{n}\frac{\partial}{\partial\lambda}\ln f(X_i;\lambda) = \sum_{i=1}^{n}\left(\frac{X_i}{\lambda} - 1\right) = \frac{\sum_{i=1}^{n} X_i}{\lambda} - n.$$

Maximum likelihood estimatoren $\widehat{\lambda}$ er gitt ved $S(\widehat{\lambda}) = 0$. Dermed får vi $\widehat{\lambda} = \overline{X}$.

En forventningsrett estimator er *effisient*, dersom dens varians er lik Cramér-Raos nedre grense (jf. side 375 i læreboka). Nå er

$$\mathrm{V}(\widehat{\lambda}) = \mathrm{V}(\overline{X}) = \mathrm{V}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{V}\left(X_i\right) = \frac{1}{n^2}n\lambda = \frac{\lambda}{n}$$

så $\widehat{\lambda}$ en effisient estimator.

d) Av sentralgrensesetningen har vi at $\widehat{\lambda} = \overline{X}$ er tilnærmet normalfordelt med forventning λ og varians λ/n. Det er i samsvar med det vi får av det generelle resultatet om tilnærmet fordeling av maximum likelihood estimatoren gitt nederst på side 375 i læreboka.

Exercise 7.44

Tettheten til eksponentialfordelingen er vanligvis gitt på formen $f(x;\lambda) = \lambda e^{-\lambda x}$ for $x > 0$.

I dennne oppgaven vil vi uttrykke tettheten ved forventningsverdien $\mu = 1/\lambda$.

Med denne reparameterisering er tettheten gitt ved $f(x;\mu) = \dfrac{1}{\mu}e^{-\frac{x}{\mu}}$ for $x > 0$.

a) Vi har her at

$$\begin{aligned}
\ln f(x;\mu) &= -\ln\mu - \frac{x}{\mu} \\
\frac{\partial}{\partial\mu}\ln f(x;\mu) &= -\frac{1}{\mu} + \frac{x}{\mu^2} \\
\frac{\partial^2}{\partial\mu^2}\ln f(x;\mu) &= \frac{1}{\mu^2} - 2\frac{x}{\mu^3}
\end{aligned}$$

Fisher informasjon i én observasjon er

$$I(\mu) = -\mathrm{E}\left(\frac{\partial^2}{\partial\mu^2}\ln f(X,\mu)\right) = -\mathrm{E}\left(\frac{1}{\mu^2} - 2\frac{X}{\mu^3}\right) = -\frac{1}{\mu^2} + 2\frac{E(X)}{\mu^3} = -\frac{1}{\mu^2} + 2\frac{\mu}{\mu^3} = \frac{1}{\mu^2}.$$

Dermed er Cramér-Raos nedre grense $\dfrac{1}{nI(\mu)} = \dfrac{\mu^2}{n}$.

b) Score funksjon er her

$$S(\mu) = \sum_{i=1}^{n}\frac{\partial}{\partial\mu}\ln f(X_i;\mu) = \sum_{i=1}^{n}\left(-\frac{1}{\mu} + \frac{X_i}{\mu^2}\right) = -\frac{n}{\mu} + \frac{\sum_{i=1}^{n} X_i}{\mu^2}.$$

Maximum likelihood estimatoren $\widehat{\mu}$ er gitt ved $S(\widehat{\mu}) = 0$. Dermed får vi $\widehat{\mu} = \overline{X}$.

c) For eksponentialfordelingen har vi at $E(X_i) = 1/\lambda = \mu$ og $\mathrm{V}(X_i) = 1/\lambda^2 = \mu^2$.

Vi finner dermed at

$$\begin{aligned}
\mathrm{E}(\widehat{\mu}) &= \mathrm{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n}\mathrm{E}(X_i) = \frac{1}{n}n\mu = \mu, \\
\mathrm{V}(\widehat{\mu}) &= \mathrm{V}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{V}\left(X_i\right) = \frac{1}{n^2}n\mu^2 = \frac{\mu^2}{n}
\end{aligned}$$

d) Vi ser at variansen til maximum likelihood estimatoren $\widehat{\mu}$ er lik Cramér-Raos nedre grense $\frac{\mu^2}{n}$.

I tillegg har vi vist at $\widehat{\mu}$ er en forventningsrett estimator. Dermed er $\widehat{\mu}$ en effisient estimator.

Exercise 7.45

Tetthetsfunksjonen til normalfordelingen er $f(x;\mu) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

(Siden vi i denne oppgaven antar at σ er kjent, er det bare μ som er en (ukjent) parameter i tettheten.)

a) Vi har likelihooden

$$\mathcal{L}(\mu) = \prod_{i=1}^{n} f(X_i;\mu) = \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(X_i-\mu)^2}{2\sigma^2}} \right)$$

og log-likelihooden

$$\ell(\mu) = \sum_{i=1}^{n} \ln f(X_i;\mu) = \sum_{i=1}^{n} \left(-\frac{1}{2}\ln 2\pi - \ln\sigma - \frac{(X_i-\mu)^2}{2\sigma^2} \right)$$

Score funksjon er dermed

$$S(\mu) = \frac{\partial}{\partial\mu}\ell(\mu) = \sum_{i=1}^{n} \frac{\partial}{\partial\mu} \left(-\frac{1}{2}\ln 2\pi - \ln\sigma - \frac{(X_i-\mu)^2}{2\sigma^2} \right) = \sum_{i=1}^{n} \frac{X_i-\mu}{\sigma^2} = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma^2}$$

Maximum likelihood estimatoren $\widehat{\mu}$ er gitt ved $S(\widehat{\mu}) = 0$. Dermed får vi $\widehat{\mu} = \overline{X}$.

b) Vi har at $X_i \overset{u.i.f.}{\sim} N(\mu, \sigma^2)$. Fra STK1100 vet vi at $\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$; jf. sidene 297 og 311 i læreboka. Altså er

$$\widehat{\mu} = \overline{X} \sim N(\mu, \frac{\sigma^2}{n}).$$

c) Fisher informasjonen i én observasjon (med hensyn til μ) er

$$I(\mu) = -\mathrm{E}\left[\frac{\partial^2}{\partial\mu^2} \ln f(X,\mu) \right] = -\mathrm{E}\left[\frac{\partial^2}{\partial\mu^2} \left(-\frac{1}{2}\ln 2\pi - \ln\sigma - \frac{(X-\mu)^2}{2\sigma^2} \right) \right] = -\mathrm{E}\left[-\frac{1}{\sigma^2} \right] \quad = \frac{1}{\sigma^2}$$

Dermed er Cramér-Raos nedre grense $\frac{1}{nI(\mu)} = \frac{\sigma^2}{n}$.

Vi ser at maximum likelihood estimatoren $\widehat{\mu} = \overline{X}$ er en forventningsrett estimator for μ og at variansen til denne estimatoren er lik Cramér-Raos nedre grense $\frac{\sigma^2}{n}$. Dermed er $\widehat{\mu}$ en effisient estimator.

d) Det generelle resultatet om tilnærmet fordeling av maximum likelihood estimatoren gitt nederst på side 375 i læreboka, vil her gi at $\widehat{\mu} = \overline{X}$ er *tilnærmet* normalfordelt med forventning μ og varians σ^2/n. Ved resultatet i punkt b har vi nå at $\widehat{\mu}$ er *eksakt* normalfordelt med forventning μ og varians σ^2/n for alle vedier av n (og ikke bare som en tilnærming når n er stor).

Exercise 7.46

a) I denne oppgaven antar vi at μ er kjent, så det er bare σ^2 som er en (ukjent) parameter i tettheten. For enkelhets skyld innfører vi $\theta = \sigma^2$ for denne parameteren. (Det gjør det mere oversiktlig når vi skal partiellderivere med hensyn på parameteren.)

Tetthetsfunksjonen til normalfordelingen er da gitt ved $f(x,\theta) = \dfrac{1}{\sqrt{2\pi\theta}} e^{-\frac{(x-\mu)^2}{2\theta}}$

Dermed er

$$\ln f(x;\theta) = -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln\theta - \frac{(x-\mu)^2}{2\theta}$$

$$\frac{\partial}{\partial\theta}\ln f(x;\theta) = -\frac{1}{2\theta} + \frac{(x-\mu)^2}{2\theta^2}$$

$$\frac{\partial^2}{\partial\theta^2}\ln f(x;\theta) = \frac{1}{2\theta^2} - \frac{(x-\mu)^2}{\theta^3}$$

Fisher informasjonen i én observasjon (med hensyn til $\theta = \sigma^2$) er

$$I(\theta) = -\mathrm{E}\left[\frac{\partial^2}{\partial\theta^2}\ln f(X,\theta)\right] = -\mathrm{E}\left[\frac{1}{2\theta^2} - \frac{(X-\mu)^2}{\theta^3}\right] = -\frac{1}{2\theta^2} + \frac{\mathrm{E}\left[(X-\mu)^2\right]}{\theta^3}$$

$$= -\frac{1}{2\theta^2} + \frac{\mathrm{V}(X)}{\theta^3} = -\frac{1}{2\theta^2} + \frac{\theta}{\theta^3} = \frac{1}{2\theta^2}.$$

Dermed er Cramér-Raos nedre grense $\dfrac{1}{nI(\theta)} = \dfrac{2\theta^2}{n} = \dfrac{2\sigma^4}{n}$.

b) Score funksjonen er

$$S(\theta) = \sum_{i=1}^{n}\frac{\partial}{\partial\theta}\ln f(X_i,\theta) = \sum_{i=1}^{n}\left(-\frac{1}{2\theta} + \frac{(X_i-\mu)^2}{2\theta^2}\right) = -\frac{n}{2\theta} + \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{2\theta^2}$$

Maximum likelihood estimatoren $\widehat{\theta}$ er gitt ved $S(\widehat{\theta}) = 0$. Dermed får vi

$$\widehat{\theta} = \widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2.$$

(Merk at siden vi har en estimator for σ^2, ville det formelt sett vært riktig å betegne estimatoren $\widehat{\sigma^2}$. Men vi velger likevel å bruke den "penere" skrivemåten $\widehat{\sigma}^2$, som jo strengt tatt betyr kvadratet av en estimator for σ. Grunnen til at det er i orden å gjøre det slik, er at på grunn av invariansprinsippet for maximum likelihood estimatorene (jf. side 357 i læreboka), vil kvadratet av maximum likelihood estimatoren for σ være maximum likelihood estimatoren for σ^2.)

c) Vi vet at $Z_i = (X_i-\mu)/\sigma$ er standardnormalfordelt. Nå finner vi at

$$\widehat{\theta} = \widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2 = \frac{\sigma^2}{n}\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 = \frac{\sigma^2}{n}\sum_{i=1}^{n}Z_i^2$$

Ved resultatet øverst på side 317 i læreboka har vi at $\displaystyle\sum_{i=1}^{n}Z_i^2 \sim \chi_n^2$.

Av dette følger det at $n\widehat{\theta}/\theta = n\widehat{\sigma}^2/\sigma^2 \sim \chi_n^2$

d) Hvis en stokastisk variabel er kji-kvadratfordelt med n frihetsgrader, har den forventning n og varians $2n$ (jf. side 315 i læreboka). Derfor har vi at

$$\mathrm{E}\left(\frac{n\widehat{\sigma}^2}{\sigma^2}\right) = n \qquad \text{og} \qquad \mathrm{V}\left(\frac{n\widehat{\sigma}^2}{\sigma^2}\right) = 2n$$

Det gir

$$\mathrm{E}\left(\widehat{\sigma}^2\right) = \frac{\sigma^2}{n}\mathrm{E}\left(\frac{n\widehat{\sigma}^2}{\sigma^2}\right) = \frac{\sigma^2}{n}\, n = \sigma^2$$
$$\mathrm{V}\left(\widehat{\sigma}^2\right) = \left(\frac{\sigma^2}{n}\right)^2 \mathrm{V}\left(\frac{n\widehat{\sigma^2}}{\sigma^2}\right) = \frac{\sigma^4}{n^2}\, 2n = \frac{2\sigma^4}{n}$$

Vi ser at maximum likelihood estimatoren $\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$ er en forventningsrett estimator for σ^2 og at variansen til denne estimatoren er lik Cramér-Raos nedre grense $\frac{2\sigma^4}{n}$. Dermed er $\widehat{\sigma}^2$ en effisient estimator for σ^2.

e) Det generelle resultatet om tilnærmet fordeling av maximum likelihood estimatoren gitt nederst på side 375 i læreboka, vil her gi at $\widehat{\sigma}^2$ er tilnærmet normalfordelt med forventning σ^2 og varians $2\sigma^4/n$. Så forventningen og variansen for den tilnærmede normalfordelingen svarer til svarer til forventningen og variansen vi fant i punkt d). I punkt c) fant vi at $n\widehat{\sigma}^2/\sigma^2 \sim \chi^2_n$. Nå vil kji-kvadratfordelingen være tilnærmet like en normalfordeling når n er stor (jf. side 317 i læreboka), så resultatet i punkt c) er ikke i konflikt med det generelle resultatet om tilnærmet fordeling av maximum likelihood estimatoren.

Exercise 7.47

a) I denne oppgaven antar vi at μ er kjent, så det er bare σ som er en (ukjent) parameter i tettheten.

Tetthetsfunksjonen til normalfordelingen er gitt ved $f(x,\sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\theta}}$

Dermed er

$$\ln f(x;\sigma) = -\frac{1}{2}\ln 2\pi - \ln\sigma - \frac{(x-\mu)^2}{2\sigma^2}$$
$$\frac{\partial}{\partial\sigma}\ln f(x;\sigma) = -\frac{1}{\sigma} + \frac{(x-\mu)^2}{\sigma^3}$$
$$\frac{\partial^2}{\partial\sigma^2}\ln f(x;\sigma) = \frac{1}{\sigma^2} - 3\frac{(x-\mu)^2}{\sigma^4}$$

Fisher informasjonen i én observasjon (med hensyn til σ) er

$$I(\sigma) = -\mathrm{E}\left[\frac{\partial^2}{\partial\sigma^2}\ln f(X,\sigma)\right] = -\mathrm{E}\left[\frac{1}{\sigma^2} - 3\,\frac{(X-\mu)^2}{\sigma^4}\right] = -\frac{1}{\sigma^2} + 3\,\frac{\mathrm{E}[(X-\mu)^2]}{\sigma^4}$$
$$= -\frac{1}{\sigma^2} + 3\,\frac{\mathrm{V}(X)}{\sigma^4} = -\frac{1}{\sigma^2} + 3\frac{\sigma^2}{\sigma^4} = \frac{2}{\sigma^2}.$$

b) Informasjon vi fikk fra Exercise 7.46 a) og den vi fikk her er forskjellige. Altså er informasjonen avhengig av parametrisering.