Løsningsforslag øving 6

Oppgave 1:

a) $Y \sim Poisson(5t)$

$$\begin{aligned} t=1: &\quad P(Y>12) = 1-P(Y\le 12) \stackrel{tabell}{=} 1-0.9980 = \underline{\underline{0.002}} \\ t=0.5: &\quad P(Y>6) = 1-P(Y\le 6) \stackrel{tabell}{=} 1-0.9858 = \underline{\underline{0.014}} \end{aligned}$$

$P(\text{minst en av 10 prøver} > 12) = 1 - P(\text{ingen av 10 prøver} > 12)$
$= 1 - P(\text{alle prøver} \le 12) = 1 - P(\text{en prøve} \le 12)^{10} = 1-(0.9980)^{10} = \underline{\underline{0.020}}$

b) $Y_i \sim Poisson(\lambda t_i)$

$$\begin{aligned} L(\lambda) &= \prod_{i=1}^n f(y_i;t_i,\lambda) = \prod_{i=1}^n \frac{(\lambda t_i)^{y_i}}{y_i!}e^{-\lambda t_i} = \frac{\prod_{i=1}^n(\lambda t_i)^{y_i}}{\prod_{i=1}^n y_i!}e^{-\lambda\sum_{i=1}^n t_i} \\ \ln(L(\lambda)) &= \sum_{i=1}^n y_i\ln(\lambda t_i) - \ln(\prod_{i=1}^n y_i!) - \lambda\sum_{i=1}^n t_i \\ \frac{\partial \ln L(\lambda)}{\partial\lambda} &= \sum_{i=1}^n \frac{y_i}{\lambda} - \sum_{i=1}^n t_i = \frac{1}{\lambda}\sum_{i=1}^n y_i - \sum_{i=1}^n t_i = 0 \\ \Rightarrow &\quad \hat\lambda = \underline{\underline{\frac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n t_i}}} \end{aligned}$$

$$\begin{aligned} \mathrm{E}(\hat\lambda) &= \frac{1}{\sum_{i=1}^n t_i}\sum_{i=1}^n \mathrm{E}(Y_i) = \frac{1}{\sum_{i=1}^n t_i}\sum_{i=1}^n \lambda t_i = \lambda \\ \mathrm{Var}(\hat\lambda) &= (\frac{1}{\sum_{i=1}^n t_i})^2\mathrm{Var}(\sum_{i=1}^n Y_i) \stackrel{\text{uavh.}}{=} (\frac{1}{\sum_{i=1}^n t_i})^2\sum_{i=1}^n \mathrm{Var}(Y_i) \\ &= (\frac{1}{\sum_{i=1}^n t_i})^2\cdot\sum_{i=1}^n \lambda t_i = \frac{\lambda}{\sum_{i=1}^n t_i} \end{aligned}$$

c)

$$H_0: \ \lambda = 6.0 \qquad \text{mot} \qquad H_1: \ \lambda > 6.0$$

Estimator: $\hat\lambda = \frac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n t_i}$

Siden Y_i er Poissonfordelt med parameter λt_i, vil $\sum_{i=1}^n Y_i$ også være Poissonfordelt med parameter $\lambda\sum_{i=1}^n t_i$. Under H_0 er $\lambda = 6$, dvs dersom $\sum_{i=1}^n t_i > 2.5$ har vi at $\lambda\sum_{i=1}^n t_i > 6\cdot 2.5 = 15$ og fra oppgitt resultat får vi da at $\sum_{i=1}^n Y_i$ vil være tilnærmet normalfordelt. Dermed vil også $\hat\lambda$ være tilnærmet normalfordelt med (under H_0):

$$\begin{aligned} \mathrm{E}(\hat\lambda) &= \lambda = 6 \\ \mathrm{Var}(\hat\lambda) &= \frac{\lambda}{\sum_{i=1}^n t_i} = \frac{6}{\sum_{i=1}^n t_i} \\ \Rightarrow &\quad Z = \frac{\hat\lambda - 6}{\sqrt{\frac{6}{\sum_{i=1}^n t_i}}} \approx N(0,1) \text{ under } H_0 \end{aligned}$$

Vi forkaster H_o dersom $Z \ge z_{0.05} = 1.645$.

Observert : $z_{obs} = \frac{65/10-6}{\sqrt{6/10}} = 0.645 < z_{0.05} = 1.645 \Longrightarrow$ $\underline{\underline{\text{Forkaster ikke } H_0}}$.

Oppgave 2:

a)

$$\begin{aligned} P(X>4) &= \int_4^\infty \frac{2x}{3}e^{-x^2/3} = [-e^{-x^2/3}]_4^\infty = e^{-4^2/3} = \underline{\underline{0.0048}} \\ p &= P(X_1>4\cap X_2>4) \stackrel{\text{uavh.}}{=} P(X_1>4)P(X_2>4) = e^{-4^2/3}e^{-4^2/3} = \underline{\underline{0.000023}} \end{aligned}$$

Generelt er

$$p = P(X_1>4\cap X_2>4) = P(X_1>4|X_2>4)P(X_2>4)$$

der $P(X_1>4|X_2>4) > P(X_1>4)$ dersom X_1 og X_2 er positivt korrelerte, dvs $\underline{\underline{p \text{ vil øke}}}$ dersom vi har positiv korrelasjon mellom bølgehøydene.

b) $Z = 2X^2/\theta$ gir at $X = \sqrt{\frac{\theta Z}{2}} = w(Z)$ og dermed $w'(Z) = \frac{1}{2}(\frac{\theta Z}{2})^{-1/2}\frac{\theta}{2}$. Transformasjonsformelen gir nå:

$$f_Z(z) = f_X(w(z))|w'(z)| = \frac{2}{\theta}\sqrt{\frac{\theta z}{2}}e^{-(\frac{\theta z}{2})/\theta}\frac{\theta}{4}(\frac{\theta z}{2})^{-1/2} = \underline{\underline{\frac{1}{2}e^{-z/2}}}$$

Dersom vi setter $\nu = 2$ i sannsynlighetstettheten til χ^2-fordelingen får vi:

$$f(x) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)}x^{\nu/2-1}e^{-x/2} \stackrel{\nu=2}{=} \frac{1}{2^{2/2}\Gamma(2/2)}x^{2/2-1}e^{-x/2} = \frac{1}{2}e^{-x/2}$$

dvs vi ser at fordelingen til Z er en $\underline{\underline{\chi^2_2\text{-fordeling}}}$.

c)

$$\begin{aligned} L(\theta;x_1,..x_n) &= \prod_{i=1}^n(\frac{2x_i}{\theta}e^{-x_i^2/\theta}) = \frac{2^n}{\theta^n}(\prod_{i=1}^n x_i)e^{-\sum_{i=1}^n x_i^2/\theta} \\ l(\theta) &= \ln L(\theta;x_1,..x_n) = n\ln(2) - n\ln(\theta) + \ln(\prod_{i=1}^n x_i) - \frac{\sum_{i=1}^n x_i^2}{\theta} \\ \frac{dl(\theta)}{d\theta} &= -\frac{n}{\theta} + \frac{\sum_{i=1}^n x_i^2}{\theta^2} = 0 \qquad \Rightarrow \hat\theta = \underline{\underline{\frac{1}{n}\sum_{i=1}^n X_i^2}} \end{aligned}$$

Merk at $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 = \frac{\theta}{2n}\sum_{i=1}^{n} \frac{2X_i^2}{\theta} = \frac{\theta}{2n}\sum_{i=1}^{n} Z_i$ der vi har fra b) at Z_i-ene er uavh. og χ^2_2-fordelte. Vi har videre fra resultatene om forventning og varians i χ^2-fordelingen at $\mathrm{E}(Z_i) = 2$ og $\mathrm{Var}(Z_i) = 2 \cdot 2 = 4$. Får da:

$$\begin{aligned}
\mathrm{E}(\hat{\theta}) &= \mathrm{E}(\frac{\theta}{2n}\sum_{i=1}^{n} Z_i) = \frac{\theta}{2n}\sum_{i=1}^{n}\mathrm{E}(Z_i) = \frac{\theta}{2n}\sum_{i=1}^{n} 2 = \frac{\theta}{2n}2n = \underline{\underline{\theta}} \\
\mathrm{Var}(\hat{\theta}) &= \mathrm{Var}(\frac{\theta}{2n}\sum_{i=1}^{n} Z_i) = \frac{\theta^2}{(2n)^2}\sum_{i=1}^{n}\mathrm{Var}(Z_i) = \frac{\theta^2}{(2n)^2}\sum_{i=1}^{n} 4 = \frac{\theta^2}{(2n)^2}4n = \underline{\underline{\frac{\theta^2}{n}}}
\end{aligned}$$

d)

Vi ser at $\frac{2n\hat{\theta}}{\theta} = \frac{2n}{\theta}\frac{1}{n}\sum_{i=1}^{n} X_i^2 = \sum_{i=1}^{n}\frac{2}{\theta}X_i^2 = \sum_{i=1}^{n} Z_i$ der $Z_i \sim \chi^2_2$. Siden en sum av n χ^2_2-fordelte variable er χ^2_{2n}-fordelt (f.saml. s. 35) har vi at $\frac{2n\hat{\theta}}{\theta} = \sum_{i=1}^{n} Z_i \sim \underline{\underline{\chi^2_{2n}}}$.

Konfidensintervall:

$$\begin{aligned}
P(\chi^2_{1-\alpha/2,2n} \le \frac{2n\hat{\theta}}{\theta} \le \chi^2_{\alpha/2,2n}) &= 1-\alpha \\
P(\frac{\chi^2_{1-\alpha/2,2n}}{2n\hat{\theta}} \le \frac{1}{\theta} \le \frac{\chi^2_{\alpha/2,2n}}{2n\hat{\theta}}) &= 1-\alpha \\
P(\frac{2n\hat{\theta}}{\chi^2_{1-\alpha/2,2n}} \ge \theta \ge \frac{2n\hat{\theta}}{\chi^2_{\alpha/2,2n}}) &= 1-\alpha \\
P(\frac{2n\hat{\theta}}{\chi^2_{\alpha/2,2n}} \le \theta \le \frac{2n\hat{\theta}}{\chi^2_{1-\alpha/2,2n}}) &= 1-\alpha
\end{aligned}$$

Dvs et $(1-\alpha)100\%$ konfidensintervall for θ er gitt ved:

$$\underline{\underline{\left[\frac{2n\hat{\theta}}{\chi^2_{\alpha/2,2n}},\ \frac{2n\hat{\theta}}{\chi^2_{1-\alpha/2,2n}}\right]}}$$

Med $n = 50$, $\hat{\theta} = 1.047$ og $\alpha = 0.05$ som gir $\chi^2_{1-\alpha/2,2n} = \chi^2_{0.975,100} = 74.222$ og $\chi^2_{\alpha/2,2n} = \chi^2_{0.025,100} = 129.561$ får vi:

$$\left[\frac{2\cdot 50\cdot 1.047}{129.561},\ \frac{2\cdot 50\cdot 1.047}{74.222}\right] = \underline{\underline{[0.81,\ \ 1.41]}}$$

e)

$$\begin{aligned}
F_Y(y) &= P(Y \le y) = P(\max(X_1, \dots X_m) \le y) \\
&= P(X_1 \le y \cap \dots \cap X_m \le y) \overset{uavh}{=} P(X_1 \le y)\cdots P(X_m \le y) = F_X(y)^m
\end{aligned}$$

Siden

$$F_X(y) = \int_{-\infty}^{y} f_X(x)dx = \int_0^y \frac{2x}{\theta}e^{-x^2/\theta}dx = 1 - e^{-y^2/\theta}$$

blir

$$F_Y(y) = \underline{\underline{[1 - e^{-y^2/\theta}]^m}}$$

Siden $P(Y > y_k) = 1/100$ er $F_Y(y_k) = P(Y \le y_k) = 1 - 1/100$:

$$\begin{aligned}
F_Y(y_k) &= 1 - \frac{1}{100} \\
\left[1 - e^{-y_k^2/\theta}\right]^m &= 1 - \frac{1}{100} \\
1 - e^{-y_k^2/\theta} &= \left[1 - \frac{1}{100}\right]^{1/m} \\
1 - \left[1 - \frac{1}{100}\right]^{1/m} &= e^{-y_k^2/\theta} \\
\ln\left(1 - \left[1 - \frac{1}{100}\right]^{1/m}\right) &= -y_k^2/\theta \\
y_k &= \underline{\underline{\sqrt{-\theta\ln\left(1 - \left[1 - \frac{1}{100}\right]^{1/m}\right)}}}
\end{aligned}$$

f) Dersom $y_k < 4$ er sannsynligheten for å få en bølge større enn 4 meter mindre enn 0.01. Arbeidet startes dersom målingene gir grunnlag for å hevde at $y_k < 4$. I forrige punkt fant vi sammenhengen mellom y_k og parameteren θ. Fra dette får vi:

$$\begin{aligned}
y_k &= \sqrt{-\theta\ln\left(1 - [1 - \frac{1}{100}]^{1/m}\right)} \\
y_k^2 &= -\theta\ln\left(1 - [1 - \frac{1}{100}]^{1/m}\right) \\
\theta &= -y_k^2/\ln\left(1 - [1 - \frac{1}{100}]^{1/m}\right)
\end{aligned}$$

Dvs $y_k = 4$ og $m = 900$ gir:

$$\theta = -4^2/\ln\left(1 - [1 - \frac{1}{100}]^{1/900}\right) = 1.40$$

Dvs $y_k < 4 \Leftrightarrow \theta < 1.40$ slik at vi kan avgjøre om dataene gir grunnlag for å hevde at $y_k < 4$ ved å teste om $\theta < 1.40$.

$$H_0:\ \theta = 1.40 \qquad \text{mot} \qquad H_1:\ \theta < 1.40$$

Under H_0 har vi at: $\frac{2n\hat{\theta}}{\theta_0} = \frac{2n\hat{\theta}}{1.40} \sim \chi^2_{2n}$.

Vi forkaster H_0 dersom $\frac{2n\hat{\theta}}{1.40} \le \chi^2_{1-\alpha,2n} = \chi^2_{0.95,100} = 77.929$. Observert

$$\frac{2\cdot 50\cdot 1.047}{1.40} = 74.8 < 77.929$$

dvs vi $\underline{\underline{\text{forkaster } H_0}}$. Observasjonene gir grunnlag for å hevde at $y_k < 4$. Arbeidet kan starte!