Løsningsforslag, Statistikk 1 eksamen 5.8.1998

Oppgave 1)

$$\mathrm{E}(X) = \sum_{x=-2}^{2} xf(x) = -2 \cdot 0.1 + (-1) \cdot 0.1 + 0 \cdot 0.5 + 1 \cdot 0.2 + 2 \cdot 0.1 = \underline{\underline{0.1}}$$

$$P(X \geq 0) = f(0) + f(1) + f(2) = \underline{\underline{0.8}}$$

$$P(X \geq 0 | X \leq 1) = \frac{P(X \geq 0 \cap X \leq 1)}{P(X \leq 1)} = \frac{f(0) + f(1)}{f(-2) + f(-1) + f(0) + f(1)} = \underline{\underline{0.78}}$$

Oppgave 2)

a)

$$\begin{aligned} P(Z > 10) &= 1 - P(Z \leq 10) = 1 - F(10; 0.05) \\ &= 1 - (1 - e^{-0.05 \cdot 10}) = e^{-0.5} = \underline{\underline{0.607}} \end{aligned}$$

I den neste deloppgaven benytter vi egenskapen at eksponensialfordelingen er "uten hukommelse" og får

$$P(Z > 20 | Z > 10) = P(Z > 10) = \underline{\underline{0.607}}$$

b)

$$\begin{aligned} P(M = m) &= P(m \leq Z < m + 1) \\ &= P(Z \leq m + 1) - P(Z \leq m) \\ &= F(m + 1; \lambda) - F(m; \lambda) \\ &= \big(1 - e^{-\lambda(m+1)}\big) - \big(1 - e^{-\lambda m}\big) \\ &= e^{-\lambda m} - e^{-\lambda(m+1)} \\ &= \underline{\underline{(1 - e^{-\lambda})e^{(-\lambda m)}}} \end{aligned}$$

c)

For å finne SME setter vi opp rimelighetsfunksjonen L, tar ln til denne for å lette regningen, deriverer og setter lik 0 for å finne maksimum.

$$\begin{aligned}
L(\lambda) &= \prod_{i=1}^{n} P(M_i = m_i) = \prod_{i=1}^{n} \left[\left(1 - e^{-\lambda}\right) e^{-\lambda m_i}\right] \\
l(\lambda) &= \ln L = \sum_{i=1}^{n} \left[\ln(1 - e^{-\lambda}) - \lambda m_i\right] \\
l'(\lambda) &= \sum_{i=1}^{n} \left[\frac{e^{-\lambda}}{1 - e^{-\lambda}} - m_i\right] = \frac{ne^{-\lambda}}{1 - e^{-\lambda}} - \sum_{i=1}^{n} m_i \\
l'(\lambda) &= 0 \qquad \Rightarrow \frac{ne^{-\lambda}}{1 - e^{-\lambda}} = \sum_{i=1}^{n} m_i
\end{aligned}$$

Denne siste ligningen løser vi med hensyn på λ og får

$$\begin{aligned}
ne^{-\lambda} &= \sum_{i=1}^{n} m_i - e^{-\lambda} \sum_{i=1}^{n} m_i \\
e^{-\lambda} \left(n + \sum_{i=1}^{n} m_i\right) &= \sum_{i=1}^{n} m_i \\
e^{-\lambda} &= \frac{\sum_{i=1}^{n} m_i}{n + \sum_{i=1}^{n} m_i} \\
\lambda &= -\ln \left[\frac{\sum_{i=1}^{n} m_i}{n + \sum_{i=1}^{n} m_i}\right] \\
&= \ln(n + \sum_{i=1}^{n} m_i) - \ln(\sum_{i=1}^{n} m_i)
\end{aligned}$$

Det betyr at SME blir

$$\underline{\underline{\hat{\lambda} = \ln(n + \sum_{i=1}^{n} M_i) - \ln(\sum_{i=1}^{n} M_i)}}$$

Oppgave 3)

a)

- β angir bilens bensinforbruk (i liter/mil)

- Rimelig med $\alpha = 0$ fordi med $x = 0$ (ingen kjøring) brukes ingen bensin

- en tur av lengde $x_1 = x$ kan tenkes sammensatt av to turer på $x_2 = x/2$ og $x_3 = x/2$. La Y_1, Y_2, Y_3 være tilhørende bensinforbruk. Det er da rimelig å kreve at

$\text{Var}(Y_1) = \text{Var}(Y_2) + \text{Var}(Y_3)$.

Dette oppnås ved å velge

$\text{Var}(Y) = x\sigma^2$

b)

$\beta = 0.75 \quad , x = 5.0 \quad , \sigma^2 = 0.1^2$

Dette betyr at

$Y \sim n(y; \beta x, \sqrt{x\sigma^2}) \sim n(y; 3.75, \sqrt{0.05})$

$$\begin{aligned} P(Y > 4) &= 1 - P(Y \leq 4) = 1 - P\left(\frac{Y - 3.75}{\sqrt{0.05}} \leq \frac{4 - 3.75}{\sqrt{0.05}}\right) \\ &= 1 - \Phi(1.12) = 1 - 0.869 = \underline{\underline{0.131}} \end{aligned}$$

Ser så på to kjøreturer

$Y_1 \sim n(y; 3.75, \sqrt{0.05})$ og

$Y_2 \sim n(y; 7.5, \sqrt{0.1})$

P.g.a. uavhengighet har vi at $Z = Y_1 + Y_2 \sim n(z; 3.75 + 7.5, \sqrt{0.05 + 0.10})$.

$$\begin{aligned} P(Z < 12) &= P\left(\frac{z - 11.25}{\sqrt{0.15}} \leq \frac{12 - 11.25}{\sqrt{0.15}}\right) = \Phi(1.94) \\ &= \underline{\underline{0.974}} \end{aligned}$$

$$\begin{aligned} U = Y_2 - 2Y_1 &\sim n(z; 0, \sqrt{0.1 + 4 \cdot 0.05}) \\ P(Y_2 - 2Y_1 > 0) &= P(U > 0) = \underline{\underline{0.5}} \end{aligned}$$

Siden fordelingen til U er symmetrisk om $u = 0$.

c)

Studerer to estimatorer $\hat{\beta}$ og $\tilde{\beta}$

$$
\begin{aligned}
\mathrm{E}(\hat\beta) &= \mathrm{E}\left(\frac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n x_i}\right) = \frac{\mathrm{E}(\sum_{i=1}^n Y_i)}{\sum_{i=1}^n x_i} = \frac{\sum_{i=1}^n \mathrm{E}(Y_i)}{\sum_{i=1}^n x_i} \\
&= \frac{\sum_{i=1}^n \beta x_i}{\sum_{i=1}^n x_i} = \beta\frac{\sum_{i=1}^n x_i}{\sum_{i=1}^n x_i} = \underline{\underline{\beta}} \\
\mathrm{Var}(\hat\beta) &= \mathrm{Var}\left(\frac{\sum_{i=1}^n Y_i}{\sum_{i=1}^n x_i}\right) = \frac{\mathrm{Var}(\sum_{i=1}^n Y_i)}{(\sum_{i=1}^n x_i)^2} = \frac{\sum_{i=1}^n \mathrm{Var}(Y_i)}{(\sum_{i=1}^n x_i)^2} = \frac{\sum_{i=1}^n x_i\sigma^2}{\sum_{i=1}^n x_i} \\
&= \sigma^2\frac{\sum_{i=1}^n x_i}{(\sum_{i=1}^n x_i)^2} = \underline{\underline{\frac{\sigma^2}{\sum_{i=1}^n x_i}}}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}(\tilde\beta) &= \mathrm{E}\left(\frac{1}{n}\sum_{i=1}^n \frac{Y_i}{x_i}\right) = \frac{1}{n}\sum_{i=1}^n \mathrm{E}\left(\frac{Y_i}{x_i}\right) = \frac{1}{n}\sum_{i=1}^n \frac{\mathrm{E}(Y_i)}{x_i} \\
&= \frac{1}{n}\sum_{i=1}^n \frac{\beta x_i}{x_i} = \frac{\beta}{n}\sum_{i=1}^n \frac{x_i}{x_i} = \frac{\beta}{n}n = \underline{\underline{\beta}} \\
\mathrm{Var}(\tilde\beta) &= \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^n \frac{Y_i}{x_i}\right) = \frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}\left(\frac{Y_i}{x_i}\right) = \frac{1}{n^2}\sum_{i=1}^n \frac{\mathrm{Var}(Y_i)}{x_i^2} \\
&= \frac{1}{n^2}\sum_{i=1}^n \frac{x_i\sigma^2}{x_i^2} = \underline{\underline{\frac{\sigma^2}{n^2}\sum_{i=1}^n \frac{1}{x_i}}}
\end{aligned}
$$

Vi ser at begge estimatorene er forventingsrette. Vi foretrekker den med minst varians. Med oppitte tall for x_i'ene får vi

$\mathrm{Var}(\hat\beta) = \sigma^2 \cdot 0.00299$ og $\mathrm{Var}(\tilde\beta) = \sigma^2 \cdot 0.0107$

Det vil si at vi $\underline{\underline{\text{foretrekker } \hat\beta}}$

d)

$$H_0 : \beta = 0.56 \quad \text{mot } H_1 : \beta > 0.56$$

$\hat\beta$ blir normalfordelt siden den er en lineærkombinasjon av uavhengige, normalfordelte variabler.

Under H_0 vil en ha at $\mathrm{E}(\hat\beta) = 0.56$ og $\mathrm{Var}(\hat\beta) = \frac{\sigma^2}{\sum_{i=1}^n x_i}$

Vi benytter testobservatoren

$$U = \frac{\hat\beta - 0.56}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^n x_i}}} \sim n(u;0,1) \quad \text{under } H_0$$

Vi forkaster H_0 dersom $U > k$, der k bestemmes fra kravet

$$P(\text{Forkast } H_0 \text{ når } H_0 \text{ er riktig }) = 0.05$$

det vil si at $k = u_{0.05} = 1.645$
Innsatt observasjonene:

$$\hat{\beta} = 0.584 \quad \sigma^2 = 0.1^2 \quad \sum_{i=1}^{n} x_i = 335 \quad \Rightarrow U = \frac{0.584 - 0.56}{\sqrt{\frac{0.1^2}{335}}} = 4.38 > k$$

Det vil si $\underline{\underline{\text{Forkast } H_0}}$. Vi vil da påstå at bilen bruker mer bensin enn forhandleren sier.
e)

Vet at

$$V = \frac{\hat{\beta} - \beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n} x_i}}} \sim n(v; 0, 1)$$

$$\begin{aligned} P\left(-u_{0.025} \leq V \leq u_{0.025}\right) &= 0.95 \\ P\left(-u_{0.025} \leq \frac{\hat{\beta} - \beta}{\sqrt{\frac{\sigma^2}{\sum_{i=1}^{n} x_i}}} \leq u_{0.025}\right) &= 0.95 \\ P\left(\hat{\beta} - u_{0.025}\sigma\sqrt{\frac{1}{\sum_{i=1}^{n} x_1}} \leq \beta \leq \hat{\beta} + u_{0.025}\sigma\sqrt{\frac{1}{\sum_{i=1}^{n} x_1}}\right) &= 0.95 \end{aligned}$$

Vi finner da et 95% konfidensintervall for β

$$\left[\hat{\beta} - u_{0.025}\sigma\sqrt{\frac{1}{\sum_{i=1}^{n} x_1}} \;,\; \hat{\beta} + u_{0.025}\sigma\sqrt{\frac{1}{\sum_{i=1}^{n} x_1}}\;\right]$$

Innsatt for tallverdiene $\hat{\beta} = 0.584$, $\sigma = 0.1$, $\sum_{i=1}^{n} x_i = 335$ og $u_{0.025} = 1.96$ får vi da

$$\underline{\underline{[0.573, 0.595]}}$$

Oppgave 4)

$$\begin{aligned}
f(t|\lambda) &= \lambda e^{-\lambda t} \quad ,t \geq 0 \\
f(\lambda) &= \theta e^{-\lambda\theta} \quad ,\lambda \geq 0 \\
& \quad \text{Simultanfordelingen er da gitt ved} \\
f(t,\lambda) &= \lambda\theta e^{-\lambda(t+\theta)} \\
& \text{Marginalfordelingen kan da finnes ved å integrere vekk } \lambda \\
f(t) &= \int_0^\infty f(t,\lambda)d\lambda = \int_0^\infty \lambda\theta e^{-\lambda(t+\theta)}d\lambda \\
&= \theta\left(\left[\lambda\left(-\frac{1}{t+\theta}\right)e^{-\lambda(t+\theta)}\right]_0^\infty + \int_0^\infty \frac{1}{t+\theta}e^{-\lambda(t+\theta)}d\lambda\right) \\
&= \theta\left(0 - 0 + \left[\frac{1}{(t+\theta)^2}e^{-\lambda(t+\theta)}\right]_0^\infty\right) \\
&= \theta\left(0 + \frac{1}{(t+\theta)^2}\right) = \frac{\theta}{(t+\theta)^2}
\end{aligned}$$

Det vil altså si at fordelingen er gitt ved

$$\underline{\underline{f(t) = \frac{\theta}{(t+\theta)^2} \qquad ; t \geq 0}}$$