Norges teknisk-naturvitenskapelige universitet
Institutt for matematiske fag

Øving nummer 9, blokk II
Løsningsskisse

Oppgave 1

a) The probability is $\int_{0.5}^{0.9} 6x(1-x)\,dx = \int_{0.5}^{0.9}(6x-6x^2)\,dx = [3x^2-2x^3]_{0.5}^{0.9} = 0.472.$

b) The likelihood function is given by

$$L(\beta) = \prod_{i=1}^{n} \beta(\beta+1)x_i(1-x_i)^{\beta-1} = \beta^n(\beta+1)^n\left(\prod_{i=1}^{n} x_i\right)\prod_{i=1}^{n}(1-x_i)^{\beta-1},$$

and the log likelihood

$$\ln L(\beta) = n\ln\beta + n\ln(\beta+1) + \sum_{i=1}^{n}\ln x_i + (\beta-1)\sum_{i=1}^{n}\ln(1-x_i),$$

which has derivative

$$(\ln L)'(\beta) = \frac{n}{\beta} + \frac{n}{\beta+1} + \sum_{i=1}^{n}\ln(1-x_i).$$

$(\ln L)'$ is decreasing on $(0,\infty)$ and the sum of two first terms tends to ∞ when $\beta\to 0^+$ and to 0 when $\beta\to\infty$, so that $(\ln L)'$ will have a single zero (the third term is negative) for $\beta > 0$ and be positive left of the zero and negative right of the zero. This means that L has its maximum at this zero. Solving for the zero,

$$\beta^2\sum_{i=1}^{n}\ln(1-x_i) + \Big(2n + \sum_{i=1}^{n}\ln(1-x_i)\Big)\beta + n = 0,$$

we get

$$\beta = \frac{-2n - \sum_{i=1}^{n}\ln(1-x_i) \pm \sqrt{4n^2 + (\sum_{i=1}^{n}\ln(1-x_i))^2}}{2\sum_{i=1}^{n}\ln(1-x_i)}$$
$$= -\frac{n}{\sum_{i=1}^{n}\ln(1-x_i)} - \frac{1}{2} \pm \sqrt{\Big(\frac{n}{\sum_{i=1}^{n}\ln(1-x_i)}\Big)^2 + \frac{1}{4}}.$$

We choose the larger zero since $(\ln L)'$ has only one zero for positive arguments (the other we found must be negative), and get the maximum likelihood estimator

$$\sqrt{\Big(\frac{n}{\sum_{i=1}^{n}\ln(1-X_i)}\Big)^2 + \frac{1}{4}} - \frac{n}{\sum_{i=1}^{n}\ln(1-X_i)} - \frac{1}{2} = \sqrt{\frac{1}{\big(\overline{\ln(1-X)}\big)^2} + \frac{1}{4}} - \frac{1}{\overline{\ln(1-X)}} - \frac{1}{2}.$$

For $n = 100$ and $\sum_{i=1}^{n} \ln(1 - x_i) = -104.0$ the estimate is $\sqrt{1/1.04^2 + 1/4} + 1/1.04 - 1/2 = 1.545$.

(The discussion of actual attainment of maximum at the zero and of which zero to be chosen, is not required.)

Oppgave 2

Antar antall grove fartsoverskridelser, X, over et tidsrom t er Poissonfordelt med parameter λt.

a) Sannsynlighetsfordelingen til X er gitt ved:

$$P(X = x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}$$

Med $\lambda = 0.5$ og $t = 5$ får vi :

$$P(X = x) = \frac{(2.5)^x e^{-2.5}}{x!}$$

Sannsynligheten for at det skjer ingen grove overskridelser i perioden:

$$P(X = 0) = e^{-2.5} = \underline{\underline{0.082}}$$

Sannsynligheten for at det skjer mer enn 2 grove overskridelser i perioden:

$$\begin{aligned} P(X > 2) &= 1 - P(X \leq 2) \\ &= 1 - (P(X = 0) + P(X = 1) + P(X = 2)) \\ &= 1 - (0.082 + 0.205 + 0.257) = \underline{\underline{0.456}} \end{aligned}$$

$$\mathrm{E}[X] = 2.5, \ \mathrm{Var}[X] = 2.5$$

b) Sann fart på bilene er μ, lasermålingene er $N(\mu, 1.5^2)$. Skal finne sannsynligheten for at laseren viser mer enn 130 km/h for en bil som kjører i 129 km/h.

$$\begin{aligned} P(Y > 130 \mid \mu = 129) &= P\left(\frac{Y - 129}{1.5} > \frac{130 - 129}{1.5}\right) \\ &= P(Z > 0.67) \\ &= 1 - P(Z \leq 0.67) = 1 - 0.749 = \underline{\underline{0.251}} \end{aligned}$$

Konstanten må oppfylle:

$$P(Y \geq k \mid \mu = 130) = 0.01$$

som kan omformes til

$$P\left(\frac{Y - 130}{1.5} \geq \frac{k - 130}{1.5} \mid \mu = 130\right) = 0.01$$

Konstanten er dermed gitt ved:

$$\Leftrightarrow \frac{k - 130}{1.5} = 2.325 \Leftrightarrow k = 130 + 1.5 \cdot 2.325 \approx \underline{\underline{133.5}}$$

c) Sannsynlighetstetthetsfunksjonen er gitt ved:

$$f(x_1, x_2, x_3, x_4 \mid \lambda, t_1, t_2, t_3, t_4) = \frac{(\lambda t_1)^{x_1} e^{-\lambda t_1}}{x_1!} \cdot \frac{(\lambda t_2)^{x_2} e^{-\lambda t_2}}{x_2!} \cdot \frac{(\lambda t_3)^{x_3} e^{-\lambda t_3}}{x_3!} \cdot \frac{(\lambda t_4)^{x_4} e^{-\lambda t_4}}{x_4!}$$

Som gir følgende likelihood:

$$L(\lambda \mid x_1, x_2, x_3, x_4, t_1, t_2, t_3, t_4) = \frac{\lambda^{\sum_{i=1}^{4} x_i} \cdot \prod_{i=1}^{4} t_i^{x_i} \cdot e^{-\lambda \sum_{i=1}^{4} t_i}}{\prod_{i=1}^{4} x_i!}$$

$$\Rightarrow \ln L(\lambda \mid x_1, x_2, x_3, x_4, t_1, t_2, t_3, t_4) = \ln \lambda \sum_{i=1}^{4} x_i + \ln(\prod_{t=1}^{4} t_i^{x_i}) - \lambda \sum_{i=1}^{4} t_i - \ln(\prod_{i=1}^{4} x_i!)$$

Deriverer og setter lik null:

$$\frac{\partial \ln L(\lambda \mid \ldots)}{\partial \lambda} = \frac{1}{\lambda} \sum_{i=1}^{4} x_i - \sum_{i=1}^{4} t_i = 0 \ \Rightarrow \ \lambda = \frac{\sum_{i=1}^{4} x_i}{\sum_{i=1}^{4} t_i}$$

dvs. sannsynlighetsmaksimeringsestimatoren er:

$$\widehat{\lambda} = \frac{\sum_{i=1}^{4} X_i}{\sum_{i=1}^{4} t_i} = \frac{\sum_{i=1}^{4} X_i}{30}$$

Forventningsverdien til estimatoren er:

$$\mathrm{E}[\widehat{\lambda}] = \frac{1}{30}\{\mathrm{E}[X_1] + \mathrm{E}[X_2] + \mathrm{E}[X_3] + \mathrm{E}[X_4]\} = \frac{1}{30}\{5\lambda + 5\lambda + 10\lambda + 10\lambda\} = \underline{\underline{\lambda}}$$

og variansen er:

$$\mathrm{Var}[\widehat{\lambda}] = \frac{1}{30^2}\{\mathrm{Var}[X_1] + \mathrm{Var}[X_2] + \mathrm{Var}[X_3] + \mathrm{Var}[X_4]\} = \frac{1}{30^2}\{5\lambda + 5\lambda + 10\lambda + 10\lambda\} = \underline{\underline{\frac{\lambda}{30}}}$$

d) $Y = \sum_{i=1}^{4} X_i$ er Poissonfordelt med parameter $\lambda \sum_{i=1}^{4} t_i = 30\lambda$: Med $\lambda \approx 0.5$ er $\mathrm{Var}[Y] \approx 15$ $\Rightarrow$ det er rimelig grunn til å tro at fordelingen til Y kan tilnærmes med en normalfordeling.

$$\widehat{\lambda} = \frac{Y}{30} \approx n(v; \lambda, \sqrt{\frac{\lambda}{30}}) \ \Rightarrow \ \frac{\widehat{\lambda} - \lambda}{\sqrt{\frac{\lambda}{30}}} \approx n(z; 0, 1)$$

$$\frac{\widehat{\lambda} - \lambda}{\sqrt{\frac{\lambda}{30}}} \approx \frac{\widehat{\lambda} - \lambda}{\sqrt{\frac{\widehat{\lambda}}{30}}}$$

Vi får:

$$P\left(-z_{\frac{\alpha}{2}} < \frac{\widehat{\lambda}-\lambda}{\sqrt{\frac{\widehat{\lambda}}{30}}} < z_{\frac{\alpha}{2}}\right) \approx 1-\alpha$$

$$\Leftrightarrow P\left(\widehat{\lambda} - z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{\lambda}}{30}} < \lambda < \widehat{\lambda} + z_{\frac{\alpha}{2}}\sqrt{\frac{\widehat{\lambda}}{30}}\right) \approx 1-\alpha$$

$$\alpha = 0.05,\ \ \widehat{\lambda} = \frac{20}{30} \quad \Rightarrow \quad \text{et 95 \% konfidensintervall blir:}$$

$$(0.67 - 1.97\sqrt{\frac{2}{90}}, 0.67 + 1.97\sqrt{\frac{2}{90}}) = \underline{\underline{(0.38, 0.96)}}$$

e)

$$\begin{aligned} P(T \leq t) &= 1 - P(T > t) = 1 - P(X = 0 \text{ i tidsrommet } [0,t]) \\ &= \begin{cases} 1 - e^{-\lambda t} & t > 0 \\ 0 & \text{ellers.} \end{cases} \quad \Rightarrow \text{T er eksponensialfordelt.} \end{aligned}$$

La $U = \min\{T_1, T_2, \ldots, T_8\}$

$$\begin{aligned} P(U \leq u) &= 1 - P(T_1 > u, T_2 > u, \ldots, T_8 > u) = 1 - \prod_{i=1}^{8} P(T_i > u) \\ &= \begin{cases} 1 - \prod_{i=1}^{8} e^{-\lambda u} & u > 0 \\ 0 & \text{ellers.} \end{cases} = \begin{cases} 1 - e^{-8\lambda u} & u > 0 \\ 0 & \text{ellers.} \end{cases} \\ &= \begin{cases} 1 - e^{-4u} & u > 0 \\ 0 & \text{ellers.} \end{cases} \end{aligned}$$

$$P(U \leq \frac{1}{4}) = 1 - e^{-1} = 1 - 0.368 = \underline{\underline{0.632}}$$

Oppgave 3

(Merk: I følgje oppgåveteksten skal konfidensintervallet *utleias*, ikkje berre setjas opp!)

Vi har at $X_1, \ldots, X_n$ er u.i.f. $\mathrm{N}(\mu_1, \sigma_0^2)$ og at $Y_1, \ldots, Y_m$ er u.i.f. $\mathrm{N}(\mu_2, \sigma_0^2)$, og også at alle X_i-ane er uavhengige av alle Y_j-ane, $i = 1, \ldots, n,\ j = 1, \ldots, m$. Forventningsverdiane μ_1 og μ_2 er ukjende, medan variansen σ_0^2 er felles og kjend.

Naturleg estimator: $\hat{\mu_1} - \hat{\mu_2} = \bar{X} - \bar{Y}$

Då estimatoren er ein lineærkombinasjon av uavhengige normalfordelte variable er han sjølv normalfordelt med:

$$\begin{aligned} \mathrm{E}(\hat{\mu_1} - \hat{\mu_2}) &= \mathrm{E}(\bar{X}) - \mathrm{E}(\bar{Y}) = \mu_1 - \mu_2 \\ \mathrm{Var}(\hat{\mu_1} - \hat{\mu_2}) &= \mathrm{Var}(\bar{X}) + \mathrm{Var}(\bar{Y}) = \frac{\sigma_0^2}{n} + \frac{\sigma_0^2}{m}. \end{aligned}$$

D.v.s:

$$\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sigma_0\sqrt{\frac{1}{n}+\frac{1}{m}}} \sim \mathrm{N}(0,1)$$

som gjev:

$$P\left(-z_{\frac{\alpha}{2}} \leq \frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sigma_0\sqrt{\frac{1}{n}+\frac{1}{m}}} \leq z_{\frac{\alpha}{2}}\right) = 1-\alpha$$

$$P\left(\bar{X}-\bar{Y}-z_{\frac{\alpha}{2}}\sigma_0\sqrt{\frac{1}{n}+\frac{1}{m}} \leq \mu_1-\mu_2 \leq \bar{X}-\bar{Y}+z_{\frac{\alpha}{2}}\sigma_0\sqrt{\frac{1}{n}+\frac{1}{m}}\right) = 1-\alpha$$

D.v.s. at vi får $(1-\alpha)100\%$ konfidensintervall ved:

$$\left[\bar{X}-\bar{Y}-z_{\frac{\alpha}{2}}\sigma_0\sqrt{\frac{1}{n}+\frac{1}{m}}, \bar{X}-\bar{Y}+z_{\frac{\alpha}{2}}\sigma_0\sqrt{\frac{1}{n}+\frac{1}{m}}\right]$$

For å få numerisk svar set vi inn talverdiane: $\bar{x} = 28.80$, $\bar{y} = 26.07$, $\sigma_0 = 2$, $m = n = 10$ og $z_{0.025} = 1.96$. Får då eit 95%-konfidensintervall på:

$\underline{\underline{[0.977, 4.483]}}$

Oppgave 4

a) $T \sim \text{eksp}(\frac{z}{\mu}) \quad \mathrm{E}(T) = \frac{\mu}{z}$

$\mu = 1000, \;\; z = 2.0$

$P(T \leq 1000) = \int_0^{1000} \frac{z}{\mu} e^{-\frac{z}{\mu}x} dx = \int_0^{1000} \frac{1}{500} e^{-\frac{x}{500}} dx = [-e^{-\frac{x}{500}}]_0^{1000} = 1 - e^{-2} = \underline{\underline{0.86}}$

$P(T \leq 1000) = 0.5 \quad \Leftrightarrow \quad 1 - e^{-\frac{1000z}{1000}} = 0.5$

$e^{-z} = 0.5 \quad \Leftrightarrow \quad z = -\ln 0.5 = \underline{\underline{0.69}}$

$z_1 = 1.0, \quad z_2 = 2.0$

$P(T_2 \geq T_1) =?$

Finner simultanfordelingen til T_1 og T_2:

$f(t_1, t_2) = \frac{z_1}{\mu} e^{-\frac{z_1}{\mu}t_1} \frac{z_2}{\mu} e^{-\frac{z_2}{\mu}t_2}$ siden T_1 og T_2 er uavhengige.

$$\begin{aligned}
P(T_2 \geq T_1) &= \int_0^\infty \int_{t_1}^\infty f(t_1,t_2) dt_2 dt_1 = \tfrac{z_1 z_2}{\mu^2} \int_0^\infty \int_{t_1}^\infty e^{-\frac{z_1}{\mu}t_1} e^{-\frac{z_2}{\mu}t_2} dt_2 dt_1 \\
&= \tfrac{z_1 z_2}{\mu^2} \int_0^\infty [-\tfrac{\mu}{z_2} e^{-\frac{z_1}{\mu}t_1 - \frac{z_2}{\mu}t_2}]_{t_1}^\infty dt_1 = \tfrac{z_1 z_2}{\mu^2} \tfrac{\mu}{z_2} \int_0^\infty e^{-\frac{z_1}{\mu}t_1 - \frac{z_2}{\mu}t_1} dt_1 \\
&= \tfrac{z_1}{\mu} [-\tfrac{\mu}{z_1+z_2} e^{-(\frac{z_1+z_2}{\mu})t_1}]_0^\infty = \tfrac{z_1}{z_1+z_2} = \tfrac{1.0}{1.0+2.0} = \underline{\underline{\tfrac{1}{3}}}
\end{aligned}$$

b) SME for μ:

$f(t_1,\ldots,t_n;\mu,z_1,\ldots,z_n) = \prod_{i=1}^n \frac{z_i}{\mu} e^{-\frac{z_i}{\mu}t_i}$

$L(\mu;t_1,\ldots,t_n,z_1,\ldots,z_n) = \prod_{i=1}^n \frac{z_i}{\mu} e^{-\frac{z_i}{\mu}t_i}$

$l(\mu) = \ln L(\mu) = \sum_{i=1}^n \ln z_i - n\ln\mu - \sum_{i=1}^n \frac{z_i}{\mu}t_i$

$\frac{\partial l}{\partial\mu} = -\frac{n}{\mu} + \sum_{i=1}^n \frac{z_i t_i}{\mu^2} = 0$

$n = \sum_{i=1}^n \frac{z_i t_i}{\mu}$

$\mu = \frac{1}{n}\sum_{i=1}^n z_i t_i$ Dermed er SME $\widehat{\mu} = \frac{1}{n}\sum_{i=1}^n z_i T_i$.

$$\mathrm{E}(\widehat{\mu}) = \mathrm{E}(\tfrac{1}{n}\sum_{i=1}^n z_i T_i) = \tfrac{1}{n}\sum_{i=1}^n z_i \mathrm{E}(T_i) = \tfrac{1}{n}\sum_{i=1}^n z_i \tfrac{\mu}{z_i} = \tfrac{1}{n}\sum_{i=1}^n \mu = \underline{\underline{\mu}}$$

Dvs. estimatoren er forventningsrett.

$$\begin{aligned}\mathrm{Var}(\hat{\mu}) &= \mathrm{Var}(\tfrac{1}{n}\sum_{i=1}^n z_i T_i) = \tfrac{1}{n^2}\sum_{i=1}^n \mathrm{Var}(z_i T_i) = \tfrac{1}{n^2}\sum_{i=1}^n z_i^2 \mathrm{Var}(T_i)\\ &= \tfrac{1}{n^2}\sum_{i=1}^n z_i^2 \tfrac{\mu^2}{z_i^2} = \tfrac{1}{n^2}\sum_{i=1}^n \mu^2 = \underline{\underline{\tfrac{\mu^2}{n}}}\end{aligned}$$

c) MGF for T_i: $M_{T_i}(t) = \frac{\frac{z_i}{\mu}}{\frac{z_i}{\mu}-t}$ (Funnet i tabell.)

$V = \frac{2n\widehat{\mu}}{\mu} = \frac{2\sum_{i=1}^n z_i T_i}{\mu} = \sum_{i=1}^n \frac{2z_i}{\mu}T_i$

$M_{\frac{2z_i}{\mu}T_i}(t) = \frac{\frac{z_i}{\mu}}{\frac{z_i}{\mu}-\frac{2z_i}{\mu}t} = (1-2t)^{-1}$ (Bruker at $M_{aX}(t) = M_X(at)$)

$M_V(t) = \prod_{i=1}^n (1-2t)^{-1} = (1-2t)^{-n}$

(Bruker at $M_{\sum_{i=1}^n X_i}(t) = \prod_{i=1}^n M_{X_i}(t)$)

$(1-2t)^{-n}$ er MGF for kji-kvadratfordelingen med $2n$ frihetsgrader. V har samme MGF som kji-kvadratfordelingen med $2n$ frihetsgrader, derfor er $V \sim \chi^2_{2n}$.

d) $(1-\alpha)100\%$ konfidensintervall for μ:
Bruker at $V = \frac{2n\widehat{\mu}}{\mu} \sim \chi^2_{2n}$.

$$\begin{aligned}&P(z_{1-\alpha/2,2n} \le V \le z_{\alpha/2,2n}) = 1-\alpha\\ &P(z_{1-\alpha/2,2n} \le \tfrac{2n\widehat{\mu}}{\mu} \le z_{\alpha/2,2n}) = 1-\alpha\\ &P(\frac{z_{1-\alpha/2,2n}}{2n\widehat{\mu}} \le \tfrac{1}{\mu} \le \frac{z_{\alpha/2,2n}}{2n\widehat{\mu}} \le \tfrac{1}{\mu}) = 1-\alpha\\ &P(\frac{2n\widehat{\mu}}{z_{\alpha/2,2n}} \le \mu \le \frac{2n\widehat{\mu}}{z_{1-\alpha/2,2n}}) = 1-\alpha\end{aligned}$$

Det gir konfidensintervallet $\underline{\underline{[\frac{2n\widehat{\mu}}{z_{\alpha/2,2n}}, \frac{2n\widehat{\mu}}{z_{1-\alpha/2,2n}}]}}$

$\alpha = 0.10$, $n = 10$, $\widehat{\mu} = 1270.38$

$z_{1-\alpha/2,2n} = z_{0.95,20} = 10.85$, $z_{\alpha/2,2n} = z_{0.05,20} = 31.41$

Innsatt disse tallverdiene blir konfidensintervallet $\underline{\underline{[808.90, 2341.71]}}$