Løsningsforslag Eksamen i Statistikk 1 18/5 1998

Oppgave 1)

$$f(x) = \begin{cases} k(1-x^2) & \text{for } -1 \leq x \leq 1 \\ 0 & \text{ellers} \end{cases}$$

For at $f(x)$ skal være en sannsynlighetstetthet, må $\int_{-1}^{1} f(x)dx = 1$.

$$\int_{-1}^{1} f(x)dx = k[x - \tfrac{1}{3}x^3]_{-1}^{1} = k(1 - \tfrac{1}{3} - (-1 + \tfrac{1}{3})) = k\tfrac{4}{3} = 1$$

Det gir $\underline{\underline{k = \tfrac{3}{4}}}$.

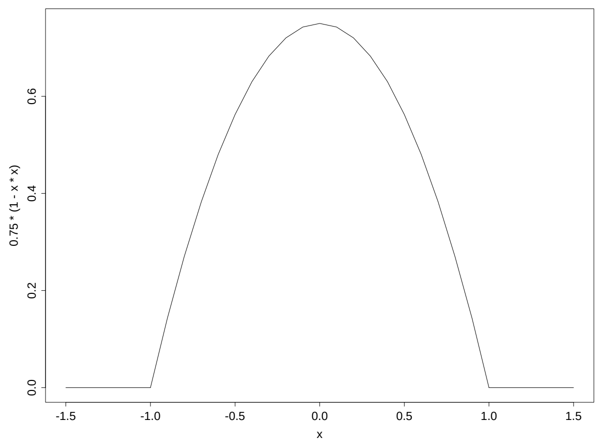

Figure 1: Skisse av $f(x)$.

$$P(X \leq 0.5) = \int_{-1}^{0.5} \tfrac{3}{4}(1-x^2)dx = \tfrac{3}{4}[x - \tfrac{1}{3}x^3]_{-1}^{0.5} = \tfrac{3}{4}(-0.5 - 0.04167 - (-1 + 0.3333)) = \underline{\underline{0.8438}}$$

$$P(X \leq 0.8 | X > 0.5) = \frac{P(X \leq 0.8 \cap X > 0.5)}{P(X > 0.5)} = \frac{P(0.5 < X \leq 0.8)}{P(X > 0.5)}$$

$$P(0.5 < X \leq 0.8) = \int_{0.5}^{0.8} \tfrac{3}{4}(1-x^2)dx = \tfrac{3}{4}[x - \tfrac{1}{3}x^3]_{0.5}^{0.8} = \tfrac{3}{4}(0.629 - 0.458) = 0.128$$

Det gir $P(X \leq 0.8 | X > 0.5) = \frac{0.128}{1 - 0.8438} = \underline{\underline{0.821}}$

Oppgave 2

a)

R: En bolle inneholder minst 10 rosiner.

S: En bolle inneholder minst 4 sukater.

A: En bolle inneholder minst 10 rosiner og minst 4 sukater. $A = R \cap S$.

B: En bolle inneholder færre enn 10 rosiner, men minst 4 sukater. $B = R^c \cap S$.

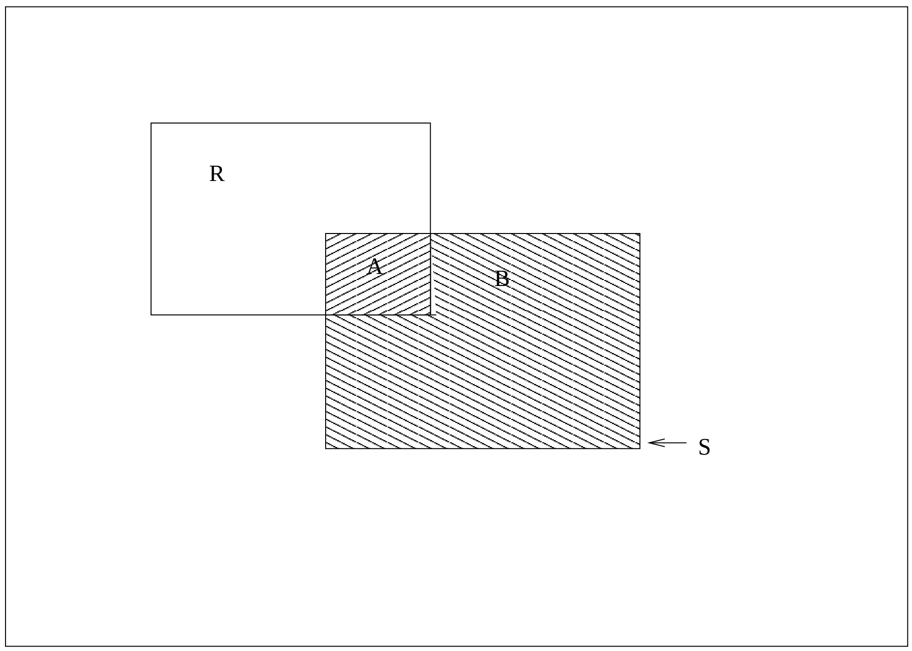

Figure 2: Venndiagram

A og B er disjunkte, fordi $A \cap B = \emptyset$. En bolle kan ikke samtidig inneholde både minst 10 rosiner og færre enn 10 rosiner.

b)

$\lambda_R = 10$, $\lambda_S = 4$

$X \sim \text{Po}(\lambda_R)$, $Y \sim \text{Po}(\lambda_S)$.

$P(\#\text{ rosiner} \geq 10) = P(X \geq 10) = 1 - P(X \leq 9) = 1 - 0.45793 = \underline{\underline{0.54207}}$

$P(\#\text{ rosiner og sukater} \geq 14) = P(X + Y \geq 14) = 1 - P(X + Y \leq 13)$

Siden X og Y er uavhengige og Poissonfordelt, er $X + Y \sim \text{Po}(\lambda_R + \lambda_S) = \text{Po}(10 + 4) = \text{Po}(14)$.

$1 - P(X + Y \leq 13) = 1 - 0.4655 = \underline{\underline{0.53555}}$

$P(3$ av 6 boller har mindre enn 10 rosiner$)$: La Z være antall boller med mindre enn 10 rosiner. $Z \sim \text{bin}(6, 0.45793)$.

$P(Z = 3) = \binom{6}{3} (0.45793)^3 (0.54207)^3 = \underline{\underline{0.306}}$

c)

$f(x_1, \dots, x_n; \lambda_R) = \prod_{i=1}^n \frac{\lambda_R{}^{x_i}}{x_i!} e^{-\lambda_R}$

$L(\lambda_R; x_1, \dots, x_n) = \prod_{i=1}^n \frac{\lambda_R{}^{x_i}}{x_i!} e^{-\lambda_R}$

$l(\lambda_R; x_1, \dots, x_n) = \ln L(\lambda_R; x_1, \dots, x_n) = \sum_{i=1}^n (x_i \ln \lambda_R - \ln x_i!) - n\lambda_R$

$\frac{\partial l}{\partial \lambda_R} = \sum_{i=1}^n \frac{x_i}{\lambda_R} - n = 0 \Leftrightarrow \lambda_R = \frac{1}{n} \sum_{i=1}^n x_i$

Dette gir oss SME: $\underline{\underline{\widehat{\lambda_R} = \frac{1}{n} \sum_{i=1}^n X_i}}$

$\text{E}(\widehat{\lambda_R}) = \text{E}(\frac{1}{n} \sum_{i=1}^n X_i) = \frac{1}{n} \sum_{i=1}^n \text{E}(X_i) = \frac{1}{n} \sum_{i=1}^n \lambda_R = \underline{\underline{\lambda_R}}$

$\text{Var}(\widehat{\lambda_R}) = \text{Var}(\frac{1}{n} \sum_{i=1}^n X_i) = \frac{1}{n^2} \sum_{i=1}^n \text{Var}(X_i) = \frac{1}{n^2} \sum_{i=1}^n \lambda_R = \underline{\underline{\frac{\lambda_R}{n}}}$

$\hat{\lambda}$ er tilnærmet normalfordelt som følge av Sentralgrenseteoremet, som sier at hvis n er stor, er et gjennomsnitt av uavhengige variable $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ tilnærmet normalfordelt.

d)

Hypotesetest: $H_0: \quad \lambda_R = 10 \quad$ mot $H_1: \quad \lambda_R < 10$

Testobservator: $Z = \frac{\widehat{\lambda_R} - 10}{\sqrt{10/50}} \sim N(0, 1)$ under H_0.

Forkast H_0 når $Z < k$ der

$P(Z < k | \lambda_R = 10) = \alpha$

$P(\frac{\widehat{\lambda_R} - 10}{\sqrt{10/50}} < k) = \alpha$

$\alpha = 0.05 \quad \Rightarrow \quad k = -1.645$

$\widehat{\lambda_R} = \frac{470}{50} = 9.40 \qquad \frac{9.40 - 10}{\sqrt{10/50}} = -1.34$

$\underline{\underline{\text{Ikke forkast } H_0 \text{ på } 5\% \text{ nivå.}}}$

e)

$P(\text{Forkast } H_0 | \lambda_R = 9) = 0.9$

$P(\frac{\widehat{\lambda_R}-10}{\sqrt{10/n}} < -1.645 | \lambda_R = 9) = 0.9$

$P(\widehat{\lambda_R} < -1.645\sqrt{10/n} + 10 | \lambda_R = 9) = 0.9$

Når $\lambda_R = 9$, har vi at $\frac{\widehat{\lambda_R}-9}{\sqrt{9/n}} \sim N(0,1)$.

$P(\frac{\widehat{\lambda_R}-9}{\sqrt{9/n}} < \frac{-1.645\sqrt{10/n}+10-9}{\sqrt{9/n}}) = 0.9$

$\Phi(\frac{-1.645\sqrt{10/n}+1}{\sqrt{9/n}}) = 0.9$

$\frac{-1.645\sqrt{10/n}+1}{\sqrt{9/n}} = 1.28$

$-1.645\sqrt{10/n} + 1 = 1.28\sqrt{9/n}$

$\sqrt{n} = 1.28\sqrt{9} + 1.645\sqrt{10} = 9.04 \quad \Rightarrow \quad n = 81.76$

$\underline{\underline{\text{Olav må sjekke 82 boller.}}}$

Oppgave 3

a)

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

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

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

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

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

$z_1 = 1.0, \quad z_2 = 2.0$

$P(T_2 \geq T_1) = ?$

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.

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

b)

SME for μ:

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

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

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

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

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

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

$$\text{E}(\widehat{\mu}) = \text{E}(\frac{1}{n} \sum_{i=1}^n z_i T_i) = \frac{1}{n} \sum_{i=1}^n z_i \, \text{E}(T_i) = \frac{1}{n} \sum_{i=1}^n z_i \frac{\mu}{z_i} = \frac{1}{n} \sum_{i=1}^n \mu = \underline{\underline{\mu}}$$

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 $\underline{\underline{[808.90, 2341.71]}}$