Norges teknisk-naturvitenskapelige universitet
Instituut for matematiske fag

Løsningsforslag - Eksamen august 2003

Oppgave 1

a)

Vi har normalfordelte variabler med $\mu = 1.0$ og $\sigma = 0.2$.

En måling mindre enn 1.3:

$$P(X \leq 1.3) = P\left(\frac{X - 1.0}{0.2} \leq \frac{1.3 - 1.0}{0.2}\right) = \Phi(1.5) = 0.9332.$$

Avvik på mer enn σ:

$$\begin{aligned} &P(|X - \mu| > \sigma) = P(|Z| > 1) = P(Z < -1 \cup Z > 1) \\ &= 2P(Z < -1) = 2 \cdot 0.1587 = 0.3174. \end{aligned}$$

Sannsynlighet for at måling er større enn 1.0 gitt avvik fra μ mindre enn eller lik σ:

$$\begin{aligned} &P(X > 1.0| \quad | \quad |X - \mu| \leq \sigma) = P(X - \mu > 0 \quad | \quad |X - \mu| \leq \sigma) \\ &= P(Z > 0 \quad | \quad |Z| \leq 1) = 0.5 \end{aligned}$$

pga symmetri (eller ved utregning).

b)

$\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$ siden $\overline{X}$ er en lineærkombinasjon av uavhengige $N(\mu, \sigma^2)$-variabler.

Sannsynligheten for at gjennomsnittet av 5 målinger avviker mer enn σ fra μ:

$$\begin{aligned} &P(|\overline{X} - \mu| > \sigma) = P\left(\left|\frac{\overline{X} - \mu}{\sigma/\sqrt{5}}\right| > \sqrt{5}\right) \\ &= P(|Z| > 2.23) = 2P(Z \leq -2.23) = 2 \cdot 0.0129 = 0.0258. \end{aligned}$$

c)

En god estimator bør være forventningsrett med liten varians.
$n\frac{\widehat{\sigma}^2}{\sigma^2}$ er χ^2-fordelt med n frihetsgrader.
$(n-1)\frac{S^2}{\sigma^2}$ er χ^2-fordelt med $n-1$ frihetsgrader.

$$\begin{aligned}
\mathrm{E}(\widehat{\sigma}^2) &= \frac{\sigma^2}{n}\mathrm{E}\left(\frac{n\widehat{\sigma}^2}{\sigma^2}\right) = \sigma^2. \\
\mathrm{E}(S^2) &= \frac{\sigma^2}{n-1}E\left(\frac{(n-1)S^2}{\sigma^2}\right) = \sigma^2.
\end{aligned}$$

Begge estimatorene er forventningsrette.

$$\begin{aligned}
\mathrm{Var}(\widehat{\sigma}^2) &= \left(\frac{\sigma^2}{n}\right)^2 \mathrm{Var}\left(\frac{n\widehat{\sigma}^2}{\sigma^2}\right) = \frac{\sigma^4}{n^2}\cdot 2n = \frac{2\sigma^4}{n}. \\
\mathrm{Var}(S^2) &= \left(\frac{\sigma^2}{n-1}\right)^2 \mathrm{Var}\left(\frac{(n-1)S^2}{\sigma^2}\right) = \frac{\sigma^4}{(n-1)^2}\cdot 2(n-1) = \frac{2\sigma^4}{n-1}.
\end{aligned}$$

$\mathrm{Var}(\widehat{\sigma}^2) < \mathrm{Var}(S^2)$, dermed foretrekkes $\widehat{\sigma}^2$ da den er mer effektiv enn S^2. (Minst varians.)

d)

$$\begin{aligned}
\mathrm{P}(\chi^2_{1-\alpha/2,n} &< \frac{n\widehat{\sigma}^2}{\sigma^2} < \chi^2_{\alpha/2,n}) = 1-\alpha \\
\mathrm{P}(\frac{n\widehat{\sigma}^2}{\chi^2_{\alpha/2,n}} &< \sigma^2 < \frac{n\widehat{\sigma}^2}{\chi^2_{1-\alpha/2,n}}) = 1-\alpha.
\end{aligned}$$

$(1-\alpha)\cdot 100\%$ konfidensintervall for σ^2 basert på estimatoren $\widehat{\sigma}^2$ blir da

$$\left(\frac{n\widehat{\sigma}^2}{\chi^2_{\alpha/2,n}}, \frac{n\widehat{\sigma}^2}{\chi^2_{1-\alpha/2,n}}\right)$$

hvor $n\widehat{\sigma}^2 = \sum_{i=1}^n (x_i-\mu)^2$.
Innsatt $\alpha = 0.05$, $n = 10$, $\sum_{i=1}^{10}(x_i-\mu)^2 = 0.50$ fås et 95% konfidensintervall

$$\left(\frac{0.50}{20.483}, \frac{0.50}{3.247}\right) = (0.024, 0.154).$$

Oppgave 2

Vi har hendelsene

$$A = \{X > 10\}$$
$$B = \{X \leq 14\}$$

a)

Vi setter $\mu = 10$ og slår opp kvantiler i tabell for Poissonfordelingen.

$$\mathrm{P}(A) = 1 - \mathrm{P}(X \leq 10) = 1 - 0.583 = 0.417.$$
$$\mathrm{P}(B'|A) = \mathrm{P}(X > 14|X > 10) = \frac{\mathrm{P}(X > 14)}{\mathrm{P}(X > 10)} = \frac{1 - 0.9165}{0.417} = \frac{0.0835}{0.417} = 0.2.$$

Dermed har vi $\mathrm{P}(B|A) = 0.8$.

Hvis A og B er uavhengige, er $\mathrm{P}(B|A) = P(B)$. Her er

$$\mathrm{P}(B) = \mathrm{P}(X \leq 14) = 0.9165 \neq \mathrm{P}(B|A) = 0.8.$$

Altså er A og B avhengige hendelser.

b)

Anta $x_1, x_2, \ldots, x_n$ er de observerte verdiene på $X_1, X_2, \ldots, X_n$. Rimelighetsfunksjonen blir

$$L(x_1, \ldots, x_n, \mu) = \prod_{i=1}^{n} \frac{\mu^{x_i}}{x_i!} e^{-\mu}$$
$$l(x_1, \ldots, x_n, \mu) = \ln(L) = (\sum x_i) \ln \mu - \ln(\prod x_i!) - n\mu$$
$$\frac{\partial l}{\partial \mu} = \frac{\sum x_i}{\mu} - n.$$

Setter det siste uttrykket lik null med $\widehat{\mu}$ for μ og får

$$\widehat{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i.$$

Sannsynlighetsmaksimeringsestimatoren (SME) blir dermed

$$\widehat{\mu} = \frac{1}{n} \sum_{i=1}^{n} X_i, \qquad \mathrm{E}(\widehat{\mu}) = \frac{1}{n} \sum \mathrm{E}(X_i) = \frac{1}{n} \cdot n\mu = \mu.$$

$\widehat{\mu}$ er forventningsrett.

c)

$$Y = \frac{3n}{2}\widetilde{\mu} = \sum_{i=1}^{n} X_i = \sum_{i=1}^{n/2} X_i + \sum_{i=n/2+1}^{n} X_i = Y_1 + Y_2.$$

Y_1 er Poisson-fordelt med parameter $\sum_{i=1}^{n/2} \mu = \frac{n}{2}\mu$,
Y_2 er Poisson-fordelt med parameter $\sum_{i=n/2+1}^{n} 2\mu = n\mu$,
$Y = Y_1 + Y_2$ er Poisson-fordelt med parameter $\frac{n}{2}\mu + n\mu = \frac{3n}{2}\mu$.

Y er er en diskret tilfeldig variabel med verdier $x = 0, 1, 2, \ldots$. Dermed er $\widetilde{\mu}$ diskret tilfeldig variabel med mulige verdier $\frac{2}{3n}x$, for $x = 0, 1, 2, \ldots$.

$$\mathrm{P}(\widetilde{\mu} = \frac{2}{3n}x) = \mathrm{P}(Y = x) = \frac{\left(\frac{3n}{2}\mu\right)^x}{x!} e^{-\frac{3n\mu}{2}}$$

for $x = 0, 1, 2, \ldots$.

Sentralgrensetroremet: $n \to \infty$ innebærer også $\frac{n}{2} \to \infty$, og dermed blir Y_1 og Y_2 tilnærmet normalfordelte for store n. Siden Y_1 og Y_2 er uavhengige, blir $Y = Y_1 + Y_2$ også normalfordelt. Dermed blir $\widetilde{\mu} = \frac{2}{3n}Y$ normalfordelt, siden vi bare multipliserer med en konstant.

Oppgave 3

a)

Vi har Z binomisk fordelt $b(z; n, p) = b(z; 15, 0.7)$.

$$\mathrm{P}(Z \geq 11) = 1 - \mathrm{P}(Z \leq 10) = 1 - 0.485 = 0.515.$$

Variansen til estimatoren $\widehat{P} = \frac{Z}{n}$:

$$\mathrm{Var}(\widehat{P}) = \frac{1}{n^2}\mathrm{Var}(Z) = \frac{1}{n^2} \cdot np(1-p) = \frac{p(1-p)}{n}.$$

Uttrykket over er størst for $p = \frac{1}{2}$. (Da er også standardavviket størst.) Maksimal varians er $\frac{1}{4n}$.

$\mathrm{Var}(\widehat{P}) \leq (0.1)^2$ krever $\frac{1}{4n} \leq (0.1)^2$, dvs. $4n \geq 100$, $n \geq 25$.

b)

Vi har testobservator $\overline{D} = \frac{1}{15}\sum_{i=1}^{15} D_i$. En lineærkombinasjon av uavhengige normalfordelte variabler er normalfordelt, altså

$$\overline{D} \sim N\left(\mu_D, \frac{\sigma_D^2}{15}\right).$$

Hypotesetesten blir

$H_0 \quad : \quad \mu_D \leq 0 \quad$ (B ikke bedre enn A i snitt)

$H_1 \quad : \quad \mu_D > 0 \quad$ (B bedre enn A i gjennomsnitt)

$$\mathrm{P}(\text{Type I feil}) = \mathrm{P}(\text{Forkaste } H_0 | H_0 \text{ sann}) = \mathrm{P}\left(\frac{\overline{D} - \mu_D}{\sigma_D/\sqrt{15}} > z_{0.05}\right) = 0.05.$$

Fra tabell: $z_{0.05} = 1.645$. Utfallet av testobservatoren er $\overline{d} = \frac{5.4}{15} = 0.36$. Dermed blir beregnet z-verdi $\frac{0.36}{0.5} \cdot \sqrt{15} = 2.78 > 1.645$.

Dermed forkastes H_0, dvs. at H_1 aksepteres.

c)

Vi vet nå at den korrekte forventningsverdien $\mu_D = 0.3$. Sannsynligheten for å oppdage at B

er bedre enn A er

$$\begin{aligned}
&\mathrm{P}\left(\frac{\overline{D}}{\sigma_D/\sqrt{15}} > z_{0.05}|\mu_D = 0.3\right) = \mathrm{P}\left(\overline{D} > \frac{\sigma_D}{\sqrt{15}} z_{0.05}|\mu_D = 0.3\right) \\
&= \mathrm{P}\left(Z > z_{0.05} - \frac{0.3}{\sigma_D/\sqrt{15}}\right) \qquad \text{(Dvs: trekker fra korrekt } \mu_D \text{ og deler på standardavviket)} \\
&= \mathrm{P}(Z > 1.645 - 2.32) = \mathrm{P}(Z > -0.68 = \mathrm{P}(Z \leq 0.68) = 0.75.
\end{aligned}$$

Skal til slutt finne ut hvor mange observasjoner som må gjøres for å oppnå sannsynlighet på 0.9 for å avsløre at B er bedre enn A. Setter opp samme uttrykk, men lar være å sette inn for n.

$$\begin{aligned}
\mathrm{P}\left(Z > 1.645 - \frac{0.3}{0.5}\sqrt{n}\right) &\geq 0.9 \\
\mathrm{P}(Z \leq 0.6\sqrt{n} - 1.645) &\geq 0.9 \qquad \text{av symmetri.}
\end{aligned}$$

Vi vet at $\mathrm{P}(Z \leq z_{0.1}) = 0.9$, dermed vil vi ha

$$0.6\sqrt{n} - 1.645 \geq z_{0.1}.$$

Fra tabellen er $z_{0.1} = 1.282$, som gir

$$n \geq \left(\frac{1.645 + 1.282}{0.6}\right)^2 = 23.8,$$

dvs. at $n = 24$ er minste antall observasjoner du kan gjøre.