LØSNING, EKSAMEN I STATISTIKK, SIF 5062, MAI 2002

OPPGAVE 1

a)

I dette punktet er $\beta = 10$.

$P(X \leq 4) = 1 - e^{-4/10} = 0.33$

$P(X > 7) = e^{-7/10} = 0.50$

$P(X > 7|X > 4) = \frac{P(X>7)}{P(X>4)} = \frac{0.5}{1-0.33} = 0.74$

b)

Utleder SME for β basert på $X_1, X_2, \ldots, X_k$.

Vi setter opp rimelighetsfunksjonen, tar logaritmen, og finner nullpunkt:

$L(\beta) = \frac{1}{\beta}e^{-x_1\beta} \ldots \frac{1}{\beta}e^{-x_k/\beta}$

$L(\beta) = \frac{1}{\beta^k}e^{-\sum_{i=1}^k x_i/\beta}$

$\ln L(\beta) = -k\ln(\beta) - \sum_{i=1}^k x_i/\beta$

$\frac{dlnL(\beta)}{d\beta} = -\frac{k}{\beta} + \frac{\sum_{i=1}^k x_i}{\beta^2}$

Setter man $\frac{dlnL(\beta)}{d\beta} = 0$ får man $\hat{\beta} = \frac{\sum_{i=1}^k x_i}{k}$.

c)

$P(X > c) = 1 - F(c) = e^{-c/\beta}$

$P(Y > y) = P(X - c > y|X > c) = \frac{P(X>y+c)}{P(X>c)} = e^{-y/\beta}$

Fordelingen til Y gjenkjennes som en eksponensialfordeling med forventningsverdi β.

d)

Hvert av de n stråene i oppsamleren blir enten klipt eller ikke klipt. Sannsynligheten for å bli klipt er lik for alle strå. Om et strå blir klipt eller ikke antas uavhengig av hva som skjer med andre strå. Under disse antakelsene er Z binomisk fordelt.

Suksesssansynligheten: $P(X > c) = e^{-c/\beta}$. Et estimat for suksessannsynligheten gis ved $\frac{Z}{n}$. Et estimat for β finnes ved å løse $e^{-c/\beta} = \frac{Z}{n}$, som gir $\hat{\beta} = \frac{c}{ln(n/Z)}$.

Alternativt kan man maksimere rimelighetsfunksjonen for binomisk fordeling med innsatt suksessansynlighet som funksjon av β:

$\binom{n}{z}(e^{-c/\beta})^z(1 - e^{-c/\beta})^{n-z}$

Dette gir samme svar: $\hat{\beta} = \frac{c}{ln(n/Z)}$.

OPPGAVE 2

a)

$P(B) = 2P(X < 0.1 - 3\sigma) = 2P(\frac{X-0.1}{\sigma} < -3) = 2 \cdot 0.0013 = 0.0026$

$P(B|A) = \frac{P(A\cap B)}{P(A)} = \frac{P(B)}{P(A)} = \frac{0.0026}{2P(X<0.1-2\sigma)} = \frac{0.0026}{2\cdot 0.0228} = 0.057$

Med $\mu = 0.11$ blir $P(B)$ lik:

$P(B) = P(X > 0.1 + 3\sigma) + P(X < 0.1 - 3\sigma) = P(\frac{X-0.11}{\sigma} > \frac{-0.01}{\sigma} + 3) + P(\frac{X-0.11}{\sigma} < \frac{-0.01}{\sigma} - 3) = P(\frac{X-0.11}{\sigma} > 2) + P(\frac{X-0.11}{\sigma} < -4) = 0.0228 + 0 = 0.0228$

b)

En god estimator er forventningsrett og har liten varians.

I beregningene benytter vi at $\frac{\sum(X_i-\mu)^2}{\sigma^2}$ er χ^2 fordelt med n frihetsgrader. Dvs at $E(\frac{\sum(X_i-\mu)^2}{\sigma^2}) = n$ og at $Var(\frac{\sum(X_i-\mu)^2}{\sigma^2}) = 2n$.

$E(\hat{\sigma}^2) = E(\frac{\sigma^2}{n}\frac{\sum(X_i-\mu)^2}{\sigma^2}) = \frac{\sigma^2}{n}E(\frac{\sum(X_i-\mu)^2}{\sigma^2}) = \sigma^2$

$Var(\hat{\sigma}^2) = \frac{\sigma^4}{n^2}Var(\frac{\sum(X_i-\mu)^2}{\sigma^2}) = 2\sigma^4/n$

I beregningene benytter vi at $\frac{\sum(X_i-\bar{X})^2}{\sigma^2}$ er χ^2 fordelt med $n-1$ frihetsgrader. Dvs at $E(\frac{\sum(X_i-\bar{X})^2}{\sigma^2}) = n-1$ og at $Var(\frac{\sum(X_i-\bar{X})^2}{\sigma^2}) = 2(n-1)$.

$E(S^2) = E(\frac{\sigma^2}{n-1}\frac{\sum(X_i-\bar{X})^2}{\sigma^2}) = \frac{\sigma^2}{n-1}E(\frac{\sum(X_i-\bar{X})^2}{\sigma^2}) = \sigma^2$

$Var(S^2) = \frac{\sigma^4}{(n-1)^2}Var(\frac{\sum(X_i-\bar{X})^2}{\sigma^2}) = 2\sigma^4/(n-1)$

Både $\hat{\sigma}^2$ og S^2 er forventningsrette, men $\hat{\sigma}^2$ har mindre varians, og er derfor å foretrekke.

c)

$P(\chi^2_{0.95,20} < \frac{\sum(X_i-\mu)^2}{\sigma^2} < \chi^2_{0.05,20}) = 0.9$

Vi flytter om innenfor P tegnet til vi får σ^2 for seg selv.

$P(\frac{\sum(X_i-\mu)^2}{\chi^2_{0.05,20}} < \sigma^2 < \frac{\sum(X_i-\mu)^2}{\chi^2_{0.95,20}}) = 0.9$

Ved å innsette tall: $\sum(X_i-\mu)^2 = 0.0018$, samt $\chi^2_{0.95,20} = 10,85$ og $\chi^2_{0.05,20} = 31,41$, blir et 0.9 konfidensintervall for σ^2 gitt ved $(5.7\cdot 10^{-5}, 17\cdot 10^{-5})$.

d)

Vårt ønske på konfidensintervallet er at: $\frac{\sum(X_i-\mu)^2}{\chi^2_{0.95,n}} - \frac{\sum(X_i-\mu)^2}{\chi^2_{0.05,n}} = \frac{n\hat{\sigma}^2}{\chi^2_{0.95,n}} - \frac{n\hat{\sigma}^2}{\chi^2_{0.05,n}} \leq \frac{\hat{\sigma}^2}{2}$

$\hat{\sigma}^2$ forkortes slik at: $\frac{n}{\chi^2_{0.95,n}} - \frac{n}{\chi^2_{0.05,n}} \leq \frac{1}{2}$

Ved å prøve seg fram i tabellen får man at dette er oppfyllt for $n = 100$, og dermed også for høyere verdier av n.

OPPGAVE 3

a)

La μ være fosforinnholdet. Hypotesene blir som følger:

H_0: $\mu = 0.20$, H_1: $\mu > 0.20$

Testobservator $Z = \frac{\bar{X}-0.20}{0.02/\sqrt{10}}$ er standard normalfordelt under H_0. Vi forkaster H_0 når Z er stor. Med signifikansnivå 0.05 forkaster vi H_0 for $Z > 1.645$.

$z = \frac{\bar{x}-0.20}{0.02/\sqrt{10}} = 1.51$. Siden observert $z < 1.645$ beholder vi H_0.

b)

Hypotese H_0 forkastes kun hvis $Z = \frac{\bar{X}-0.20}{0.02/\sqrt{10}} > 1.645$. Under alternativ hypotese at $\mu = 0.21$ vil størrelsen $Y = \frac{\bar{X}-0.21}{0.02/\sqrt{10}}$ være standard normalfordelt.

$P(Z > 1.645) = P(\frac{\bar{X}-0.20}{0.02/\sqrt{10}} > 1.645) = P(\frac{\bar{X}-0.21}{0.02/\sqrt{10}} > \frac{-0.01}{0.02/\sqrt{10}} + 1.645)$
$= P(Y > 0.0639) = 0.48$

Vi ønsker en n slik at sannsynligheten for å forkaste H_0 er større enn 0.9:

$P(Y > \frac{-0.01}{0.02/\sqrt{n}} + 1.645) > 0.9$. Dette vil si at $\frac{-0.01}{0.02/\sqrt{n}} + 1.645 < z_{0.9}$, der $z_{0.9}$ er 0.9 percentilen i standard normalfordelingen. Siden $z_{0.9} = -1.28$, finner vi at:

$n > (\frac{0.02(1.645+1.28)}{0.01})^2 = 34.$

Dvs at $n > 35$ oppfyller kravet om 0.9 sannsynlighet for å gjenkjenne at $\mu = 0.21$.