Norges teknisk–
naturvitenskapelige universitet
Institutt for matematiske fag

Nynorsk

Fagleg kontakt under eksamen:
Håkon Tjelmeland 73 59 35 20

SIF5060/SIF5062/SIF5505/SIF5506 Statistikk og 75510/75515 Statistikk 1

Onsdag 9. august 2000
Tid: 09:00–14:00

Hjelpemidel: Godkjent lommekalkulator med tomt minne.
Statistiske tabeller og formler, Tapir forlag.
K. Rottman: Matematisk formelsamling.
Utdelt ordliste.

Oppgave 1

Levetida (målt i år), X, til ein bestemt type mekaniske komponentar har vist seg å følgje ei fordeling med kumulativ fordelingsfunksjon gjeve ved

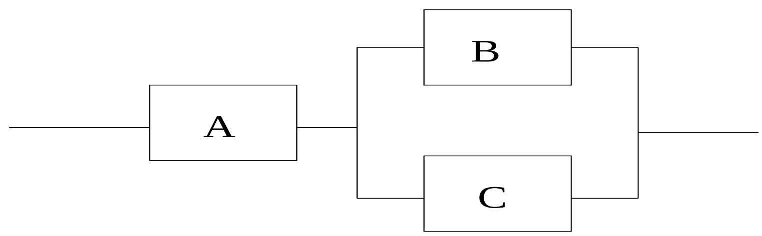
$$F(x) = 1 - \exp\left\{-\frac{x^2}{2\alpha}\right\} \quad ; \quad x > 0$$

der α er ein parameter som skildrar kvaliteten til komponentane.

a) Bestem sannsynstettleiken til X.

Bestem for kva verdi av x sannsynstettleiken $f(x)$ tar sitt maksimum. Skisser $f(x)$.

Eit instrument inneheld tre komponentar av denne typen, alle med samme kvalitetsparameter α. Vi kallar de tre komponentane komponent A, B og C. Det blir dessutan antatt at dei tre komponentane sviktar uavhengig av kvarandre. Komponentane inngår i instrumentet slik at instrumentet kun vil funksjonere så lenge komponent A og minst ein av komponentane B og C funksjonerer. Dette kan illustrerast med følgjande figur.



La følgjande fire hendingar vere definert:

A: Komponent A funksjonerer framleis etter to år.
B: Komponent B funksjonerer framleis etter to år.
C: Komponent C funksjonerer framleis etter to år.
D: Instrumentet funksjonerer framleis etter to år.

b) Teikn inn hendingane A, B og C i eit Venn-diagram.

Skraver hendinga D i Venn-diagrammet.

For $\alpha = 1$, finn sannsynet for at instrumentet framleis funksjonerer etter to år.

Oppgave 2

Miljøkonsulenten i ein kommune ønskjer å undersøke den ukjende pH-verdien i eit vatn. Kall den sanne pH-verdien for μ. Konsulenten har tilgjengeleg to målemetodar. Metode I er rask, men måleresultata er hefta med stor måleusikkerheit. Metode II er mykje meir tidkrevjande, men gjev meir nøyaktige målingar. Begge målemetodane er velbrukte og variansen i målingane er derfor kjende. Miljøkonsulenten velgjer å gjere ein observasjon med kvar metode. La X vere observasjonen ved bruk av metode I og Y observasjonen ved metode II. Vi antar at X og Y er uavhengige og normalfordelt med

$$\mathrm{E}(X) = \mu \ \ , \ \ \mathrm{Var}(X) = \sigma_0^2 \ \ , \ \ \mathrm{E}(Y) = \mu \ \ , \ \ \mathrm{Var}(Y) = \tau_0^2$$

der σ_0^2 og τ_0^2 er kjende storleikar.

Det er oppgjeve at ein forventningsrett estimator (som og er sannsynsmaksimeringsestimatoren) for μ i denne situasjonen er

$$\widehat{\mu} = \frac{\tau_0^2 X + \sigma_0^2 Y}{\tau_0^2 + \sigma_0^2}.$$

Ta utgangspunkt i estimatoren $\widehat{\mu}$ og utlei eit $(1-\alpha)100\%$ konfidensintervall for μ.

Oppgave 3 Kolibakterier i drikkevatn

Eit viktig kriterium for god drikkevasskvalitet er at talet på kolibakteriar ikke er for høgt. For å overvåke dette blir det jamnleg teke prøvar av vatnet og talet på kolibakteriar i desse blir bestemt. Erfaring tilseier at talet på kolibakteriar (X) i v liter drikkevatn er poissonfordelt med parameter λv, dvs.

$$\mathrm{P}(X = x) = \frac{(\lambda v)^x}{x!} \exp(-\lambda v) \quad ; \quad x = 0, 1, 2, \ldots$$

a) Skriv ned $\mathrm{E}(X)$. Bruk dette til å gje ei tolking av parameteren λ.

Dersom $\lambda = 3$ og $v = 0.5$, bestem sannsyna $\mathrm{P}(X = 0)$ og $\mathrm{P}(X > 3)$.

b) Dersom $\lambda = 3$ og $v = 0.5$, bestem $\mathrm{P}(X > 3 | X > 0)$.

La X_1 og X_2 vere talet på kolibakteriar i to uavhengige prøver på henhaldsvis $v_1 = 1$ og $v_2 = 2$ liter vatn. For $\lambda = 3$, bestem sannsyna $\mathrm{P}(X_1 + X_2 > 3)$ og $\mathrm{P}(X_1 + X_2 > 3 | X_1 > 0 \cap X_2 > 0)$.

Parameteren λ vil normalt vere ukjend og det er av interesse å estimere denne frå data. Anta at det for dette formålet blir tatt n prøvar av drikkevatnet og talet på kolibakteriar i kvar prøve blir bestemt. La v_i vere mengda med vatn (i liter) i prøve nummer i og la X_i vere talet kolibakteriar i denne prøven. Vi antar at $X_1, X_2, \ldots, X_n$ er uavhengige.

c) Vis at sannsynsmaksimeringsestimatoren for λ blir

$$\widehat{\lambda} = \frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n v_i}.$$

Vis vidare at $\widehat{\lambda}$ er forventningsrett og at

$$\mathrm{Var}(\widehat{\lambda}) = \frac{\lambda}{\sum_{i=1}^n v_i}.$$

I resten av oppgåva kan du utan prov nytte at $\widehat{\lambda}$ er tilnærma normalfordelt med forventning og varians som gjeve i punkt **c**).

Helsemyndigheitene har fastsett at ei øvre grense for kva verdi λ kan ha for at vatnet skal seiast å vere av god drikkevasskvalitet. Grenseverdien er $\lambda_0 = 3$ og ut frå målingane $X_1, X_2, \ldots, X_n$ er det ønskjeleg å avgjere om det er grunn til å påstå om denne verdien er overskriden.

d) Formuler dette som ein hypotesetest og lag ein test for dette formålet basert på $\widehat{\lambda}$ med (tilnærma) signifikansnivå α.

Kva blir konklusjonen på testen når $\alpha = 0.025$, $n = 10$, $v_1 = v_2 = \ldots = v_{10} = 2.0$ og talet på observerte kolibakteriar i dei 10 prøvane er

$$6 \quad 12 \quad 14 \quad 11 \quad 8 \quad 3 \quad 8 \quad 4 \quad 4 \quad 8$$

Det blir oppgjeve at $\sum_{i=1}^{n} x_i = 78$.

Vi antar fortsatt at vi tar $n = 10$ prøver, kvar på $v = 2.0$ liter. I staden for å lage et forkastningskriterium basert på $\widehat{\lambda}$ slik du har gjort i punkt **d**), kan ein då alternativt definere

$$Z = \text{talet på prøvar der } \{X_i > \lambda_0 v = 6\}$$

og forkaste H_0 dersom $Z \geq k$ der k er ein konstant som må bestemmast.

e) Bestem største verdi k slik at denne testen får signifikansnivå $\alpha \leq 0.025$.

Kva blir konklusjonen på testen når dataene er som gjeve i punkt **d**)?

f) Anta så at $\lambda = 3.5$, dvs talet på kolibakteriar i drikkevatnet overskrid grensa sete av helsemyndigheitene. For kvar av dei to testane diskutert over, bestem sannsynet for at det for høge innhaldet av kolibakteriar vil bli oppdaga dersom ein tar $n = 10$ prøvar, kvar på $v = 2.0$ liter.

Oppgave 4

La X og Y vere uavhengige og eksponensialfordelte med $\mathrm{E}(X) = \mathrm{E}(Y) = 1/\lambda$. La vidare Z vere lik absoluttverdien av differansen mellom X og Y, dvs $Z = |X - Y|$.

Finn sannsynstettleiken til Z. Kva blir $\mathrm{E}(Z)$?