**Norges teknisk–
naturvitenskapelige universitet
Institutt for matematiske fag**

Bokmål

Faglig 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

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

Oppgave 1

Levetiden (målt i år), X, til en bestemt type mekaniske komponenter har vist seg å følge en fordeling med kumulativ fordelingsfunksjon gitt ved

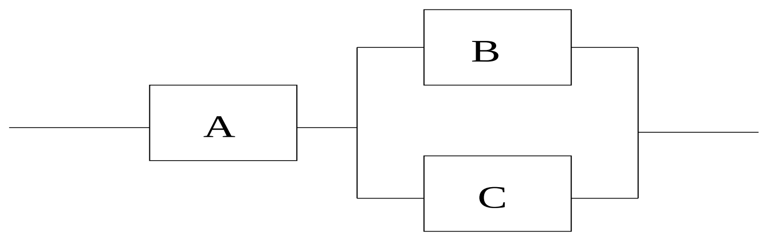
$$F(x) = 1 - \exp\left\{-\frac{x^2}{2\alpha}\right\} \quad ; \quad x > 0$$

der α er en parameter som beskriver kvaliteten til komponentene.

a) Bestem sannsynlighetstettheten til X.

Bestem for hvilken verdi av x sannsynlighetstettheten $f(x)$ tar sitt maksimum. Skisser $f(x)$.

Et instrument inneholder tre komponenter av denne typen, alle med samme kvalitetsparameter α. Vi benevner de tre komponentene henholdsvis komponent A, B og C. Det antas dessuten at de tre komponentene svikter uavhengig av hverandre. Komponentene inngår i instrumentet slik at instrumentet kun vil funksjonere så lenge komponent A og minst en av komponentene B og C funksjonerer. Dette kan illustreres med følgende figur.



La følgende fire hendelser være definert:

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

b) Tegn inn hendelsene A, B og C i et Venn-diagram.

Skraver hendelsen D i Venn-diagrammet.

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

Oppgave 2

Miljøkonsulenten i en kommune ønsker å undersøke den ukjente pH-verdien i et vann. Betegn den sanne pH-verdien for μ. Konsulenten har tilgjengelig to målemetoder. Metode I er rask, men måleresultatene er beheftet med betydelig måleusikkerhet. Metode II er mye mer tidkrevende, men gir mer nøyaktige målinger. Begge målemetodene er velbrukte og variansen i målingene er derfor kjent. Miljøkonsulenten velger å gjøre en observasjon med hver metode. La X betegne observasjonen ved bruk av metode I og Y observasjonen ved metode II. Vi antar at X og Y uavhengige og normalfordelt med

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

der σ_0^2 og τ_0^2 er kjente størrelser.

Det oppgis at en forventningsrett estimator (som forøvrig også er sannsynlighetsmaksimerings-estimator) 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 utled et $(1-\alpha)100\%$ konfidensintervall for μ.

Oppgave 3 Kolibakterier i drikkevann

Et viktig kriterium for god drikkevannskvalitet er at antall kolibakterier ikke er for høyt. For å overvåke dette tas det jevnlige prøver av vannet og antall kolibakterier i disse bestemmes. Erfaring tilsier at antall (X) kolibakterier i v liter drikkevann 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 gi en tolkning av parameteren λ.

Dersom $\lambda = 3$ og $v = 0.5$, bestem sannsynlighetene $\mathrm{P}(X = 0)$ 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 være antall kolibakterier i to uavhengige prøver på henholdsvis $v_1 = 1$ og $v_2 = 2$ liter vann. For $\lambda = 3$, bestem sannsynlighetene $\mathrm{P}(X_1 + X_2 > 3)$ og $\mathrm{P}(X_1 + X_2 > 3 | X_1 > 0 \cap X_2 > 0)$.

Parameteren λ vil normalt være ukjent og det er av interesse å estimere denne fra data. Anta at det for dette formål tas n prøver av drikkevannet og antall kolibakterier i hver prøve bestemmes. La v_i betegne antall liter vann i prøve nummer i og la X_i være antall kolibakterier i denne prøven. Det antas at $X_1, X_2, \ldots, X_n$ er uavhengige.

c) Vis at sannsynlighetsmaksimeringsestimatoren for λ blir

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

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

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

I resten av oppgaven kan du uten bevis benytte at $\widehat{\lambda}$ er tilnærmet normalfordelt med forventning og varians som gitt i punkt **c**).

Helsemyndighetene har fastsatt at en øvre grense for hvilken verdi λ kan ha for at vannet skal sies å være av god drikkevannskvalitet. Grenseverdien er $\lambda_0 = 3$ og ut fra målingene $X_1, X_2, \ldots, X_n$ er det ønskelig å avgjøre om det er grunn til å påstå om denne verdien er overskredet.

d) Formuler dette som en hypotesetest og lag en test for dette formål basert på $\widehat{\lambda}$ med (tilnærmet) signifikansnivå α.

Hva blir konklusjonen på testen når $\alpha = 0.025$, $n = 10$, $v_1 = v_2 = \ldots = v_{10} = 2.0$ og observert antall kolibakterier i de 10 prøvene er

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

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

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

$$Z = \text{antall prøver der } \{X_i > \lambda_0 v = 6\}$$

og forkaste H_0 dersom $Z \geq k$ der k er en konstant som må bestemmes.

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

Hva blir konklusjonen på testen når dataene er som gitt i punkt **d**)?

f) Anta så at $\lambda = 3.5$, dvs antall kolibakterier i drikkevannet overskrider grensen satt av helsemyndighetene. For hver av de to testene diskutert over, bestem sannsynligheten for at det for høye innholdet av kolibakterier vil bli oppdaget dersom man tar $n = 10$ prøver, hver på $v = 2.0$ liter.

Oppgave 4

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

Finn sannsynlighetstettheten til Z. Hva blir $\mathrm{E}(Z)$?