Norges teknisk–
naturvitenskapelige universitet
Institutt for matematiske fag

Bokmål

Faglig kontakt under eksamen:

Jan Terje Kvaløy	73 59 35 20
Håkon Tjelmeland	73 59 35 20
Hilde Grude Borgos	73 59 35 20

EKSAMEN I FAG SIF5060/SIF5505 STATISTIKK

Torsdag 2. desember 1999
Tid: 09:00–14:00

Hjelpemidler:
Godkjent lommekalkulator, Statistiske tabeller og formler (Tapir forlag).

Sensuren faller: 6. januar 2000.

Oppgave 1
La X og Y være to uavhengige stokastiske variable og anta at X er normalfordelt med forventning 0 og varians 2^2, mens Y er normalfordelt med forventning 4 og varians 1.5^2.

Bestem sannsynlighetene

$$\mathrm{P}(X \leq 0.5), \;\; \mathrm{P}(X \geq 0.5 | Y \leq 0.5) \;\; \text{og} \;\; \mathrm{P}(X + Y \leq 0.5).$$

Oppgave 2 Bremselengder

Bremselengde for bil med to ulike dekktyper skal undersøkes. En bremseprøve utføres ved at man begynner å bremse når bilen kjører i $80 km/t$ og bremselengde måles. For dekktype 1 utføres n slike prøver. La $X_1, X_2, \ldots, X_n$ betegne bremselengdene målt i disse prøvene. Helt tilsvarende utføres m bremseprøver for dekktype 2. La $Y_1, Y_2, \ldots, Y_m$ betegne bremselengdene målt her.

Anta at $X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots, Y_m$ alle er uavhengige og normalfordelt. Anta videre at $X_1, X_2, \ldots, X_n$ har ukjent forventningsverdi μ_1, at $Y_1, Y_2, \ldots, Y_m$ har ukjent forventningsverdi μ_2 og at $X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots, Y_m$ alle har samme kjente varians σ_0^2.

Utled et $(1-\alpha) \cdot 100\%$ konfidensintervall for differansen $\mu_1 - \mu_2$.

Regn også ut intervallet numerisk når $\alpha = 0.05$, $n = m = 10$, $\sigma_0^2 = 2^2$ og observerte bremselengder er som gitt i følgende tabell

i	1	2	3	4	5	6	7	8	9	10
x_i	33.0	30.8	28.0	28.7	28.9	26.6	27.9	28.9	27.8	27.4
y_i	23.4	25.3	25.0	28.9	26.7	25.9	24.4	26.8	28.8	25.5

Det oppgis at $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = 28.80$ og $\bar{y} = \frac{1}{m}\sum_{i=1}^{m} y_i = 26.07$.

Oppgave 3 Farskap

I situasjoner der det er uklart hvem som er den biologiske faren til et barn kan farskapet avklares ved å sammenligne DNA-prøver fra barnet med mulige fedre. For en mulig far gjøres dette ved å sammenligne n ulike deler av DNA-strukturen til mannen med de samme n deler av DNA-strukturen hos barnet. De n undersøkte delene av DNA-strukturen antas uavhengige.

Hos et barn og en tilfeldig valgt mann (som ikke er biologisk far) er det for hver enkel del av DNA-strukturen som undersøkes en sannsynlighet $p = 0.15$ for at delen er sammenfallende hos barnet og mannen. Anta videre at en biologisk far alltid har alle de undersøkte delene av DNA-strukturen sammenfallende med barnets (dvs. vi ser bort fra mutasjoner o.l.), slik at hver undersøkte del av DNA-strukturen hos biologisk far og barn er sammenfallende med sannsynlighet $p = 1$.

La X være antall sammenfallende deler i DNA-strukturen hos et barn og en tilfeldig valgt mann (som ikke er biologisk far).

a) Begrunn at X er binomisk fordelt med parametre n og $p = 0.15$.

Dersom $n = 5$, beregn sannsynlighetene $P(X = 2)$, $P(X \geq 2)$ og $P(X = 2 | X \geq 2)$.

I en farsskapssak blir en mann erklært å være biologisk far dersom alle undersøkte deler av DNA-strukturen er sammenfallende hos mannen og barnet. Dette kan vi se på som en hypotesetest der vi tester

$$H_0: \quad p = 0.15 \text{ (ikke far)} \qquad \text{mot} \qquad H_1: \quad p = 1.0 \text{ (far)}$$

der H_0 forkastes (dvs. mannen erklæres som far til barnet) dersom $X = n$.

b) For $n = 5$, finn sannsynligheten for å begå type 1 feil i testen over.

For $n = 5$, finn sannsynligheten for å begå type 2 feil i testen over.

Hvor mange ulike deler, n, av DNA-strukturen må man minst sammenligne dersom man ønsker at sannsynligheten for feilaktig å erklære en mann som far skal være mindre enn 0.000001?

Oppgave 4 Levetid til elektroniske komponenter

Levetiden (målt i måneder), X, til en del typer elektroniske komponenter har vist seg å følge en fordeling med sannsynlighetstetthet

$$f(x) = \begin{cases} \frac{1}{2\theta}\frac{1}{\sqrt{x}}e^{-\frac{\sqrt{x}}{\theta}} & \text{for } x > 0 \\ 0 & \text{ellers,} \end{cases} \tag{1}$$

der θ er en parameter som beskriver kvaliteten til komponentene. Tilhørende kumulative fordelingsfunksjon er gitt ved

$$F(x) = \begin{cases} 1 - e^{-\frac{\sqrt{x}}{\theta}} & \text{for } x > 0 \\ 0 & \text{ellers.} \end{cases}$$

I hele denne oppgaven antar vi at levetider til ulike komponenter er uavhengige.

[Hint: Du kan i denne oppgaven uten bevis benytte at

$$\int_0^\infty x^\alpha e^{-\frac{\sqrt{x}}{\theta}} \mathrm{d}x = 2\theta^{2\alpha+2}\Gamma(2\alpha+2), \quad \text{for } \alpha > -1, \theta > 0,$$

der $\Gamma(\alpha)$ er gamma-funksjonen (se også tabell).]

a) For $\theta = 2.0$, finn sannsynligheten for at en elektronisk komponent av denne typen fremdeles skal funksjonere etter 10 måneder.

Gitt at en komponent fremdeles funksjonerer etter 10 måneder, hva er sannsynligheten for at den også vil funksjonere etter 20 måneder dersom $\theta = 2.0$?

Vis at

$$\mathrm{E}(X) = 2\theta^2.$$

b) Et bestemt instrument inneholder to elektroniske komponenter, komponent A og komponent B. For at instrumentet skal funksjonere må begge de to elektroniske komponentene funksjonere. La X_A betegne levetiden til komponent A og la X_B være levetiden til komponent B. Levetiden til instrumentet blir dermed

$$U = \min(X_A, X_B).$$

Levetidene for begge komponentene er på formen gitt i ligning (1), men med ulik kvalitetsparameter θ, dvs. komponent A har kvalitetsparameter θ_A, mens komponent B har kvalitetsparameter θ_B.

Finn sannsynlighetstettheten for instrumentets levetid (uttrykt ved θ_A og θ_B).

Hva blir forventet levetid for instrumentet ?

For å undersøke kvaliteten, θ, til en ny type elektronisk komponent, har man undersøkt $n = 20$ komponenter. Vi antar fremdeles at sannsynlighetsfordelingen til levetidene følger ligning (1). La $X_1, X_2, \ldots, X_n$ betegne de tilhørende levetider. De observerte levetider er

1.27	0.40	11.37	1.91	0.083	2.15	4.68	5.54	0.88	23.45
7.12	0.18	7.75	1.41	2.33	0.020	3.00	0.37	0.0088	0.029

Det oppgis at $\sum_{i=1}^{n} \sqrt{x_i} = 29.902$.

c) Vis at sannsynlighetsmaksimeringsestimatoren (SME) for θ er

$$\widehat{\theta} = \frac{1}{n} \sum_{i=1}^{n} \sqrt{X_i}.$$

Er $\widehat{\theta}$ forventningsrett ?

La $Z_i = 2\sqrt{X_i}/\theta$ for $i = 1, 2, \ldots, n$.

d) Vis ved hjelp av transformasjonsformelen at Z_i er χ^2-fordelt med 2 frihetsgrader.

Bruk så dette til å begrunne at

$$\frac{2n\widehat{\theta}}{\theta}$$

er χ^2-fordelt med $2n$ frihetsgrader.

e) Benytt resultatet i punkt **d**) til å utlede et $(1 - \alpha) \cdot 100\%$ konfidensintervall for θ.

Regn også ut intervallet numerisk når $\alpha = 0.05$ og dataene er som gitt over.

Oppgave 5

I en Poisson-prosess med intensitet λ, la X_1 betegne tiden frem til første hendelse og la X_2 betegne tiden mellom første og andre hendelse. Som kjent er da X_1 og X_2 uavhengige og eksponensialfordelte med forventning $1/\lambda$ (dette trenger ikke du vise). La Y betegne tiden frem til andre hendelse, dvs.

$$Y = X_1 + X_2.$$

Vis at da er Y gamma-fordelt med parametre $\alpha = 2$ og $\beta = 1/\lambda$, dvs. vis at sannsynlighetstettheten for Y er

$$f(y) = \begin{cases} \lambda^2 y e^{-\lambda y} & \text{for } y > 0, \\ 0 & \text{ellers.} \end{cases}$$