Norges teknisk–
naturvitenskapelige universitet
Institutt for matematiske fag

Faglig kontakt under eksamen:
Håkon Tjelmeland   73 59 35 20

EKSAMEN I FAG 75510/75515 STATISTIKK 1

Fredag 15. august 1997
Tid: 09:00–14:00

Hjelpemidler:
Godkjent lommekalkulator
Statistiske tabeller og formler, TAPIR.

Oppgave 1

a) La A og B betegne to hendeler der $\mathrm{P}(A \cap B) > 0$. Skraver følgende hendelser i et Venn-diagram,

$$A \cap B, \qquad A \cap B', \quad \text{og} \quad A' \cup B'$$

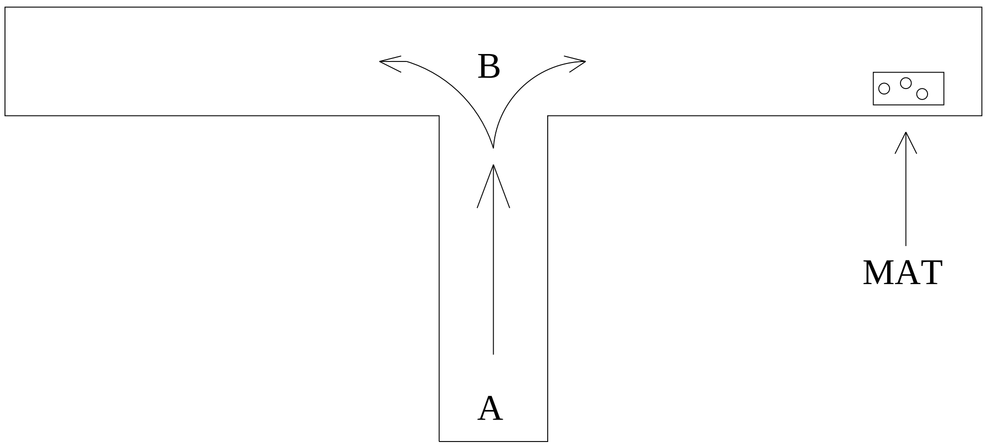
der A' er komplementærhendelsen til A og tilsvarende med B'.

b) La X og Y være uavhengige normalfordelte stokastiske variable der begge har forventningsverdi 0 og varians 4. Beregn følgende sannsynligheter,

$$\mathrm{P}(X > 1), \qquad \mathrm{P}(X + Y > 1), \quad \text{og} \quad \mathrm{P}(X - 2Y > 1).$$

Oppgave 2 Rotters læringsevne

Kari er nylig ferdig med sine studier i Trondheim og på sin første arbeidsdag får hun en interessant oppgave på laboratoriet. Hun blir bedt om å undersøke rotters læringsevne. Dette skal gjøres ved hjelp av et T-formet bur, se figuren under. En rotte slippes inn ved punkt A.



Rotta går så til punkt B og må der gå enten til høyre eller til venstre. Hvis den går til høyre, finner den mat. Hvis den går til venstre finner den ingenting. Kari antar at en rotte som slippes inn i buret for første gang går til høyre med sannsynlighet 1/2 eller venstre med sannsynlighet 1/2. For en rotte som slippes inn i buret for andre gang, gjelder derimot følgende regel: Hvis den første gang gikk til venstre, vil den igjen velge høyre med sannsynlighet 1/2 og venstre med sannsynlighet 1/2. Hvis den første gang gikk til høyre vil den neste gang gå til høyre med sannsynlighet θ og til venstre med sannsynlighet $1-\theta$, der $0<\theta<1$. Dersom $\theta > 1/2$ innebærer det altså at rotta har lært av å finne mat første gangen og derfor har større sannsynlighet for å gjøre samme valget neste gang.

Kari bestemmer seg for å undersøke rotters lærningsevne med følgende eksperiment. Hun slipper en rotte inn i buret to ganger og definerer

$$X = \begin{cases} 1, & \text{dersom rotta begge ganger går til høyre,} \\ 0, & \text{ellers.} \end{cases}$$

a) Vis at sannsynlighetsfordelingen til X, $f(x)$, kan skrives som

$$f(x) = \left(\frac{\theta}{2}\right)^x \cdot \left(1-\frac{\theta}{2}\right)^{1-x}, \qquad x = 0, 1.$$

Vis at

$$\mathrm{E}(X) = \frac{\theta}{2} \qquad \text{og} \qquad \mathrm{Var}(X) = \frac{\theta}{2}\left(1-\frac{\theta}{2}\right).$$

b) Finn sannsynligheten for at rotta går til høyre andre gangen den blir slippet inn i buret.

Dersom rotta gikk til høyre andre gang den ble sluppet inn i buret, hva er sannsynligheten for at den gikk til høyre også første gang ?

Dersom $x = 0$, hva er sannsynligheten for at rotta gikk til høyre andre gang den ble slippet inn i buret ?

For å undersøke rotters læringsevne utfører Kari eksperimentet beskrevet over for $n = 100$ rotter og observerer $X_1, X_2, \ldots, X_n$. Hun antar at $X_1, X_2, \ldots, X_n$ er et tilfeldig utvalg fra $f(x)$. (For å være på den sikre siden, registrerer hun også begge valg til alle rottene, i fall hun skulle ønske å analysere dataene på en annen måte senere.)

c) Vis at sannsynlighetsmaksimeringsestimatoren for θ, $\widehat{\theta}$, basert på $X_1, X_2, \ldots, X_n$ er

$$\widehat{\theta} = \frac{2}{n} \sum_{i=1}^{n} X_i.$$

Finn forventningsverdi og varians til $\widehat{\theta}$.

d) Finn et (tilnærmet) 95% konfindensintervall for θ ved å benytte at

$$\frac{\widehat{\theta} - \theta}{\sqrt{\frac{2\widehat{\theta}}{n}(1 - \frac{\widehat{\theta}}{2})}}$$

er tilnærmet standard normalfordelt når n er stor.

Hva blir konfidensintervallet når Kari observerer $\sum_{i=1}^{n} x_i = 32$?

Kari går nå til lunsj, godt fornøyd etter å ha utført eksperimentet beskrevet over. På lunsjrommet ser Kari tilfeldigvis en rapport som omhandler en annen lignende studie. Hun finner raskt ut at forfatterne har utført et tilsvarende eksperiment, men de har estimert θ basert på andre observasjoner enn Karis X'er. Forfatterne har definert

$$Y = \begin{cases} 1, & \text{dersom rotta går til høyre første gang } \textit{og} \text{ til venstre andre gang,} \\ 0, & \text{ellers.} \end{cases}$$

Tilsvarende til Kari, observerer de $Y_1, Y_2, \ldots, Y_n$ for rotte nummer $1, 2, \ldots, n$. De antar videre at $Y_1, Y_2, \ldots, Y_n$ er et tilfeldig utvalg fra en sannsynlighetsfordeling $g(y)$ (som du ikke trenger å regne ut). Sannsynlighetsmaksimeringsestimatoren for θ, $\widetilde{\theta}$, basert på $Y_1, \ldots, Y_n$, er funnet av forfatterne som

$$\widetilde{\theta} = 1 - \frac{2}{n} \sum_{i=1}^{n} Y_i.$$

Videre er det oppgitt at

$$\mathrm{E}(\widetilde{\theta}) = \theta \qquad \text{og} \qquad \mathrm{Var}(\widetilde{\theta}) = (1-\theta^2)/n.$$

Tilbake etter lunsj, vil Kari undersøke hvilken av estimatorene $\widehat{\theta}$ og $\widetilde{\theta}$ som hun bør benytte til å estimere θ.

e) Hvilken estimator, $\widehat{\theta}$ eller $\widetilde{\theta}$, bør Kari benytte? Begrunn svaret.

Kari vil nå undersøke om det er grunn til å påstå at rotter lærer av å finne mat første gangen de slippes inn i buret.

f) Formuler dette som et hypotesetestingsproblem og lag en test for dette formål med (tilnærmet) signifikansnivå α basert på at $\widehat{\theta}$ er tilnærmet normalfordelt når n er stor.

Hva blir konklusjonen på hypotesetesten når $n = 100$ og $\sum_{i=1}^{n} x_i = 32$ (som spesifisert over) og $\alpha = 5\%$?

g) For samme hypotesetestingsproblem som i punkt **f**), lag en ny hypotesetest med (tilnærmet) signifikansnivå α basert på at $\widetilde{\theta}$ er tilnærmet normalfordelt når n er stor.

Hva blir konklusjonen på hypotesetesten når $\sum_{i=1}^{n} y_i = 17$ ($n = 100$) og $\alpha = 5\%$?

Sammenlign med konklusjonen i punkt **f**) og kommenter.

h) For hypotesetestene i punkt **f**) og **g**), finn sannsynligheten for å forkaste H_0 dersom $\theta = 0.65$?

Kommenter konklusjonene i **f**) og **g**) basert på disse sannsynlighetene.

Noe forvirret over alle de forskjellige muligheter som analysen av målingene tilsynelatende gir, bestemmer Kari seg for å bruke all tilgjengelig informasjon. La derfor

$$Z = \begin{cases} 1, & \text{dersom rotta går til høyre begge ganger,} \\ 2, & \text{dersom rotta går til høyre første gang } \textit{og} \text{ til venstre andre gang,} \\ 3, & \text{dersom rotta går til venstre første gang } \textit{og} \text{ til høyre andre gang,} \\ 4, & \text{dersom rotta går til venstre begge ganger.} \end{cases}$$

La videre $Z_1, Z_2, \ldots, Z_n$ være observasjoner for rotte nummer $1, 2, \ldots, n$.

i) Finn et uttrykk for sannsynlighetsmaksimeringsestimatoren for θ basert på $Z_1, Z_2, \ldots, Z_n$.

Forklar hvorfor det er intuitivt rimelig at denne estimatoren kun avhenger av antall ganger $Z_i = 1$, samt antall ganger $Z_i = 2$.

Kommentar: Det kan vises at for n stor, er estimatoren du fant i punkt **i**) forventningsrett og har varians lik $2\theta(1-\theta)/n$, dvs. estimatoren er bedre enn både $\hat{\theta}$ og $\tilde{\theta}$ for alle $\theta \in [0, 1]$.