



Bokmål

Faglig kontakt under eksamen:

Ingelin Steinsland 73 55 02 39/ 92 66 30 96

Arvid Næss 73 59 70 53/ 99 53 83 50

Øyvind Salvesen 73 59 20 25/ 93 64 87 72

EKSAMEN I EMNE TMA4245 STATISTIKK

10. august 2008

Tid: 09:00–13:00

Hjelpemidler: *Tabeller og formler i statistikk*, Tapir Forlag

K. Rottmann: *Matematisk formelsamling*

Kalkulator HP30S

Gult, stemplet A5-ark med egne hndskrevne notater.

Sensuren faller: 30. august 2008

1 Høyde kvinner og menn.

Vi skal i denne oppgava se på høydefordelingene til menn og kvinner. Anta at høyden til menn er normalfordelt med forventning $\mu_M = 179$ og varians $\sigma_M^2 = 6^2$, og at høyden til unge kvinner er normalfordelt med forventning μ_K og varians σ_K^2

- a) Finn sannsynligheten for at en tilfeldig valgt ung mann er over 185 cm.
Finn sannsynligheten for at en mann er over 185 cm gitt at han er over 179 cm.
- b) Vi modellerer nå kvinners gjennomsnittshøyde som $\mu_K = \beta \cdot \mu_M$, der β er en parameter vi ønsker å estimere. Anta at vi har høydedata fra et tilfeldig utvalg på $n = 5$ kvinner;
 $X_i \sim N(\mu_K, \sigma_K^2)$, $i = 1, 2, \dots, n$.

Som estimator for β velges $\hat{\beta} = \frac{\sum_{i=1}^n X_i}{n\mu_M}$

Er $\hat{\beta}$ forventningsrett?

Utled et uttrykk for et 95% konfidensintervall for β , og finn tallsvar når vi fra data får $\bar{x} = \sum_{i=1}^n x_i/n = 167.5$ og $s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = 5.1^2$.

- c) Før har det vært vanlig å bruke $\beta = 1.08$. Vi ønsker å teste denne antakelsen ved hjelp av våre data. Formulr dette som en hypotesetest, og utfør testen med signifikansnivå $\alpha = 0.05$.

Utled et uttrykk for hvor stor n måvære for å få en teststyrke på 0.90 dersom sann $\beta = \beta_1$.

Finn nødvendig n for $\beta_1 = 1.05$.

2

En nrringsmiddelbedrift har et laboratorium for testing av kvaliteten av bedriftens produkter. For testing av mengde pr. volumenhet av et bestemt sporstoff i et av produktene, bruker bedriften to mleapparater, som vi kan kalle A og B. A er det mest nyaktige av de to, og det er ogs det raskeste. Ulempen med A er at det er dyrt i bruk. Bedriften har derfor etablert en testprosedyre som gr ut p ta ut to tilfeldige utvalg fra en gitt produksjonsserie, hvor vi kan anta at mengde pr. volumenhet, betegnet med μ , er konstant. Det ene utvalget, som er av strrelse n , testes i A. Etter denne testingen i A har en da et tilfeldig utvalg $x_1, \dots, x_n$ med mleresultater. Det andre utvalget, som er av strrelse $2n$, testes i B. Det resulterer i et uavhengig tilfeldig utvalg $y_1, \dots, y_{2n}$ av strrelse $2n$ med mleresultater etter testing i B.

Det tilfeldige utvalget $x_1, \dots, x_n$ kan betraktes som et utfall av n stokastiske variabler $X_1, \dots, X_n$, som er uavhengige og normalfordelte med forventningsverdi μ og varians σ^2 . σ^2 reflekterer A's nyaktighet. Tilsvarende, $y_1, \dots, y_{2n}$ kan betraktes som utfall av $2n$ stokastiske variabler $Y_1, \dots, Y_{2n}$, som er uavhengige og normalfordelte med forventningsverdi μ og varians $4\sigma^2$, og de er uavhengige av X_i , $i = 1, \dots, n$.

- a) Mengde pr. volumenhet μ skal estimeres p grunnlag av de tilsammen $3n$ mlingene som er gjort, og en mulig estimator er

$$\hat{M} = \frac{1}{2} (\bar{X} + \bar{Y}) \quad \left(\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, \quad \bar{Y} = \frac{1}{2n} \sum_{j=1}^{2n} Y_j \right)$$

Hva er forventningsverdi og varians til denne estimatoren?

Vis at sannsynlighetsmaksimerings-estimatoren (SME) for μ er gitt som

$$M^* = \frac{1}{3} (2\bar{X} + \bar{Y}) ,$$

og beregn dens forventningsverdi og varians.

Hvilken av de to estimatorene ville du anbefale? Begrunn svaret.

b) Anta i dette punktet at σ^2 er kjent, og lik 1.0. En skal n teste hypotesen

$$H_0 : \mu = \mu_0 = 100 ,$$

mot

$$H_1 : \mu < \mu_0 .$$

p signifikansniv $\alpha = 0.05$. Foresl en testobservator, og angi testens kritiske omrde.

Hva blir konklusjonen nr $n = 4$ og mleresultatene er:

X_i : 100.3, 100.8, 99.5, 98.8

Y_j : 100.1, 98.7, 99.2, 102.3, 103.0, 97.5, 96.6, 103.1

c) Anta i dette punktet at σ^2 er ukjent. Flgende to estimatorer for variansen σ^2 innfres,

$$S_1^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \quad \left(E[S_1^2] = \sigma^2 \right)$$

og

$$S_2^2 = \frac{1}{2n-1} \sum_{j=1}^{2n} (Y_j - \bar{Y})^2 \quad \left(E[S_2^2] = 4\sigma^2 \right)$$

Disse to kombineres i en ny stokastisk variabel,

$$V = \frac{(n-1)S_1^2}{\sigma^2} + \frac{(2n-1)S_2^2}{4\sigma^2} .$$

Begrunn hvorfor V blir χ^2 -fordelt. Hvor mange frihetsgrader har V ?

Foresl n en testobservator for μ basert p M^* og V som for $n = 4$ kan skrives som

$$T = \frac{M^* - \mu}{\sqrt{\frac{1}{20} (S_1^2 + \frac{7}{12} S_2^2)}} .$$

Begrunn svaret ved vise at ditt forslag til testobservator fr den angitte form for $n = 4$.

Hva slags variabel blir T ? Hvor mange frihetsgrader har T ?

3

I semifinalen mellom Frankrike og Tyskland i et verdensmesterskap i fotball er resultatet uavgjort etter ekstraomganger.

Vi definerer en runde i straffesparkkonkuransen til å være at hvert lag skter en straffe hver.

I første del av straffesparkkonkuransen er det 5 runder, 5 straffespark fra hvert lag, og laget som skårer flest ganger vinner.

Dersom lagene skårer like mange ganger i første del, går man over til andre del av straffesparkkonkuransen. Nå spilles det en og en runde: Hvert lag har en straffe hver. Dersom det ene laget skårer og det andre bommer, har vi en vinner. Ellers spilles en ny runde (hvert lag får en straffe hver) inntil vi har en vinner.

Anta at de tyske spillerene har sannsynlighet $p_T = 0.80$ for å skåre på straffe, at de franske spillerene har sannsynlighet $p_F = 0.70$ for å skåre, og at utfallene av straffesparkene er uavhengige av hverandre.

- a) Hva er sannsynligheten for at stillingen blir 5-5 etter første del?
Hva er sannsynligheten for at stillingen blir 3-3 etter første del?

Sett opp uttrykk for / algoritme for hvordan man kan finne sannsynligheten p_{D1-lik} for at de skårer like mangen mål i første del?

Du trenger ikke å finne tallsvar, og kan senere i oppgavan benytte at $p_{D1-lik} = 0.27$

- b) Anta i dette punktet at vi er kommet til del 2 av straffesparkkonkuransen.

Hva er sannsynligheten $p_{V1|D2}$ for at vi har en vinner etter første runde i del to?

Gitt at vi har en vinner etter første runde i del 2, hva er sannsynligheten $p_{T|V1,D2}$ for at dette er Tyskland?

Hva er sannsynligheten $p_{T|D2}$ for at Tyskland vinner konkurransen (gitt at vi er kommet til del 2)?

- c) La X være antall runder i del 2 til og med runden det blir kåret en vinner i.
Hvilken fordeling har X ? Begrunn svaret og oppgi parameter/re.
Hva er forventet og varians i antall runder (i del 2) til det blir kåret en vinner?

- d) Vi antar at det alltid blir spilt 5 runder i del 1. Hva er forventet antall runder (i del 1 og evt. del 2) til det blir kåret en vinner?

Finn også variansen i antall runder (i del 1 og evt. del 2) inntil det blir kåret en vinner.