Norges teknisk–
naturvitenskapelige universitet
Institutt for matematiske fag

Faglege kontaktar under eksamen:
Mette Langaas 98847649
Eirik Mo 73593541/41106633

EKSAMEN I FAG TMA4245 STATISTIKK

Laurdag 5. juni 2004

Tid: 09:00–14:00

Tillatne hjelpemiddel:
Gult A5 ark med eigne handskrevne notatar.
Tabeller og formler i statistikk (Tapir Forlag).
K. Rottmann: *Matematisk formelsamling.*
Kalkulator: HP30S.

NYNORSK

Sensur: 28. juni 2004.

Oppgave 1 Atle, du lyg!

Utspørjing av deltakarar i humor- og realityprogram på TV har den siste tida blitt svært populært. Spørsmåla som blir stilte kan vi dele inn i tre typer, og vi definerer følgjande disjunkte hendingar:

A_1="det blir stilt eit spørsmål av ikkje-sensitiv natur, f.eks. kva heiter du?",
A_2="det blir stilt eit spørsmål av delvis sensitiv natur, f.eks. kor gammal er du?",
A_3="det blir stilt eit spørsmål av sensitiv natur, f.eks. har du vore utru?".

I tillegg definerer vi hendinga:
L="deltakaren lyg"

Føljande sannsyn er gjevne:
$P(A_1) = 0.1$, $P(A_2) = 0.4$, $P(A_3) = 0.5$, $P(L|A_1) = 0.05$, $P(L|A_2) = 0.2$, $P(L|A_3) = 0.6$.

a) Vis dei fire hendingane i eit venndiagram.

Gitt at ein deltakar blir spurt eit spørsmål av type A_2, kva er sannsynet for at deltakaren ikkje lyg, $P(L'|A_2)$?

Kva er sannsynet for at ein tilfeldig vald deltakar lyg, $P(L)$?

Eit av spørsmåla som blir rekna å vere av delvis sensitiv natur er "kor gammal er du?". Ei gruppe på n personar vart stilt dette spørsmålet, deretter vart svara registrerte og samanlikna med informasjon i offentlege register. Lat X vere ein stokastisk variabel som angir talet på personar som lyg blant n personar, og lat p vere sannsynet for at ein person lyg.

b) Under kva for føresetnader vil X vere binomisk fordelt?

Vi føreset at $p = 0.2$ og at vi spør $n = 20$ personar. Kva er $P(X = 4)$?

Kva er $P[(X \leq 2) \cup (X > 5)]$?

Vi føreset no at p er ukjent. For å estimere p er det foreslått to estimatorar,

$$\hat{p} = \frac{X}{n} \quad \text{og} \quad p^* = \frac{X}{n-1}.$$

c) Finn forventningsverdi og varians til estimatorane $\hat{p}$ og p^*.

Kva for to eigenskapar kjenneteiknar ein god estimator?

Kva for ein av estimatorane $\hat{p}$ og p^* vil du føretrekke? Grunngi svaret.

Vi føreset at n er stor og vel å bruke estimatoren $\hat{p}$ vidare. Då er

$$Z = \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

tilnærma standard normalfordelt.

Basert på tidlegare undersøkingar føreset vi at $p = 0.2$. Vi ønskjer å undersøkje om dei innsamla dataene gir oss grunn til å tru at p er større enn 0.2.

d) Formulér dette som ein hypotesetest ved å definere nullhypotese og alternativ hypotese.

Bruk at Z er tilnærma standard normalfordelt til å bestemme eit forkastingsområde for nullhypotesen når vi vel signifikansnivå $\alpha = 0.01$.

Kva blir konklusjonen på testen når $n = 200$ personar vart spurde og $x = 55$ personar laug?

Vil p-verdien til testen vere mindre eller større enn 0.01? Grunngi svaret. (Det blir ikkje kravd at du reknar ut p-verdien.)

Oppgave 2 Aksjekursar

Selskapet Agderfrukt er notert på børsen. Vi føreset at endringa X i verdien på ein Agderfrukt-aksje i løpet av ein dag er normalfordelt med forventningsverdi $\mu_X = 0.15$ kroner og standardavvik $\sigma_X = 0.60$ kroner. Har du ein aksje i Agderfrukt vil $X > 0$ bety forteneste, medan $X < 0$ er tap.

a) Kva er sannsynet for å tape pengar i løpet av ein dag, dvs. $P(X < 0)$?

Kva er $P(0 \leq X \leq 0.15)$?

Dersom du kjøper 10 aksjar i Agderfrukt idag og sel imorgon, kva er forventa forteneste? Finn også variansen til fortenesta.

Selskapet Trønderfrukt er også notert på børsen. Vi kallar endringa på ein Trønderfrukt-aksje i løpet av ein dag for Y, der Y er normalfordelt med forventningsverdi $\mu_Y = 0.15$ kroner og standardavvik $\sigma_Y = 0.80$ kroner.

b) Vi ser på aksjekursendringa på den same dagen for Agderfrukt og Trønderfrukt og føreset i dette punktet at aksjekursendringane X og Y er uavhengige.

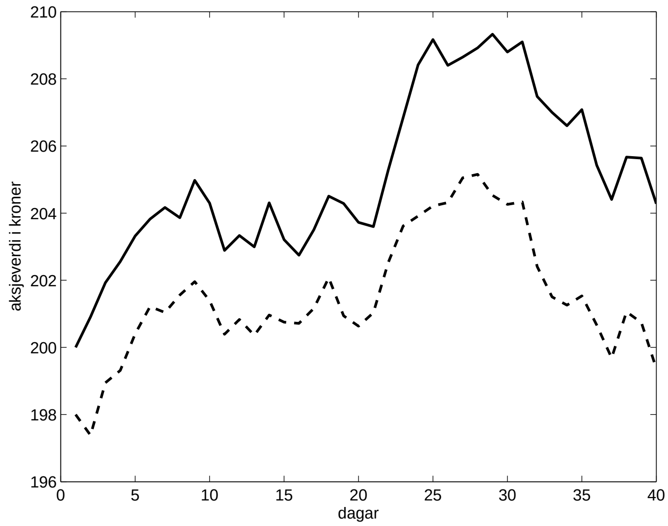
Idag er verdien til ein Agderfrukt-aksje den same som verdien til ein Trønderfrukt-aksje. Vi ønskjer å undersøkje tre moglege strategiar for aksjekjøp, der vi kjøper aksjar idag og sel imorgon.

i) Kjøp to aksjar i Agderfrukt.

ii) Kjøp ein aksje i Agderfrukt og ein aksje i Trønderfrukt.

iii) Kjøp to aksjar i Trønderfrukt.

Dersom du vil ha minst mogeleg risiko for investeringa di, kva for ein av dei tre investeringsstrategiane over vil du velje? Grunngi svaret.



Figuren viser utviklinga av aksjekursen til Agderfrukt (stipla) saman med aksjekursen til Trønderfrukt (heiltrukken).

Kursendringa dag i for Agderfrukt kallar vi X_i, og vi føreset at X_i er normalfordelt med forventning $\mu_X = 0.15$ kroner og standardavvik $\sigma_X = 0.60$ kroner.

Kursendringa dag i for Trønderfrukt kallar vi Y_i, og vi føreset at Y_i er normalfordelt med forventning $\mu_Y = 0.15$ kroner og standardavvik $\sigma_Y = 0.80$ kroner.

Kursendringar ulike dagar føreset vi er uavhengige.

Vi samanliknar dei to selskapa ved å måle differansen mellom dei daglege kursendringane, $D_i = X_i - Y_i$, og ta gjennomsnitt. Vi ser på 10 dagar og får $\overline{D} = \frac{1}{10}\sum_{i=1}^{10} D_i = \frac{1}{10}\sum_{i=1}^{10}(X_i - Y_i)$.

c) Gir figuren grunn til å tru at endringane i dei to aksjekursane same dag, X_i og Y_i, er uavhengige?

Korrelasjonen mellom X_i og Y_i for desse to selskapa, $\rho(X_i, Y_i)$, er anten -0.5, 0.0 eller 0.5. Kva for ein av desse verdiane verkar mest rimeleg frå figuren? Grunngi kort.

Kva blir forventningsverdi og varians for $\overline{D}$? Bruk verdien for korrelasjonen, $\rho(X_i, Y_i)$, som du valde over.

Oppgave 3 Bølgjehøgde

For å kunne dimensjonere ei oljeplattform er det viktig å vite kor store bølgjene kan bli i området der plattforma skal plasserast. Det blir difor sett ut ein bølgjehøgdemålar. Lat X vere største bølgjehøgde ein tilfeldig valt dag. Vi føreset at sannsynstettleiken til X er gitt ved

$$f(x;\theta) = \frac{2x}{\theta} e^{-\frac{x^2}{\theta}}, \qquad x \geq 0, \qquad \theta > 0.$$

Det blir gjeve at $\mathrm{E}[X^2] = \theta$ og $E[X^4] = 2\theta^2$.

a) Vis at den kumulative fordelingsfunksjonen, $F(x) = P(X \leq x)$, er $F(x) = 1 - e^{-\frac{x^2}{\theta}}$. (Hint: Bruk substitusjon med $u = x^2$).

Gitt at største bølgjehøgde er større enn 10 meter, finn sannsynet for at ho er større enn 15 meter dersom $\theta = 25$, dvs. $P(X > 15 | X > 10)$?

I resten av oppgåva blir θ rekna som ukjent.

Vi har observert største bølgjehøgde n dagar. Lat X_i vere største bølgjehøgde på dag i. Vi føreset at $X_1, ..., X_n$ er uavhengige og identisk fordelte med sannsynstettleik $f(x;\theta)$.

b) Finn sannsynsmaksimeringsestimatoren (SME) $\hat{\theta}$ for θ.

Er estimatoren $\hat{\theta}$ forventningsrett?

Finn variansen til $\hat{\theta}$.

c) Bruk sentralgrenseteoremet til å argumentere for at

$$Z = \frac{\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \theta}{\sqrt{\frac{\theta^2}{n}}}$$

er tilnærma standard normalfordelt.

Bruk Z til å finne eit tilnærma 95% konfidensintervall for θ.

Sannsynet for at største bølgjehøgde ein tilfeldig valt dag overskrider 10 meter er $P(X > 10) = e^{-\frac{100}{\theta}}$. Bruk det tilnærma konfidensintervallet for θ til å finne eit tilnærma 95% konfidensintervall for $e^{-\frac{100}{\theta}}$.