Norges teknisk–
naturvitenskapelige universitet
Institutt for matematiske fag

Faglige kontakter under eksamen:
Turid Follestad 98066880
Jo Eidsvik 90127472
Mette Langaas 98847649

EKSAMEN I FAG TMA4245 STATISTIKK

Fredag 19.mai 2006

Tid: 09:00–13:00

Tillatte hjelpemidler:
Gult A5-ark med egne håndskrevne notater.
Tabeller og formler i statistikk (Tapir Forlag).
K. Rottmann: *Matematisk formelsamling.*
Kalkulator: HP30S.

BOKMÅL

Sensur: 14.juni 2006.

Oppgave 1 Feil på mobilnett

La tiden X (målt i uker) mellom to påfølgende feil i et mobilnett være en kontinuerlig stokastisk variabel med sannsynlighetstetthet

$$f(x) = \beta x^{-\beta-1}, \quad x > 1,\ \beta > 1.$$

a) Vis at den kumulative fordelingsfunksjonen, $F(x)$, for X er $F(x) = 1 - x^{-\beta}$, for $x > 1$.

Anta i resten av dette punktet at $\beta = 3$.

Hva er sannsynligheten for at det tar mer enn 2 uker mellom to påfølgende feil?

Dersom det er gått 2 uker siden sist det var en feil på nettet, hva er sannsynligheten for at det svikter innen det er gått 3.5 uker fra siste feil?

Vi vil estimere parameteren β basert på data for tidligere tilfeller av feil på nettet. La $X_i,\ i = 1, \ldots, n$ være lengden på n tidsintervaller (målt i uker) mellom to påfølgende feil. Vi antar at $X_1, X_2, \ldots, X_n$ er uavhengige og identisk fordelte stokastiske variabler, med sannsynlighetstetthet $f(x)$ som gitt i starten av oppgaven.

Tre alternative estimatorer for β er

$$\hat{\beta}_1 = \frac{n}{\sum_{i=1}^{n} \ln(X_i)}, \quad \hat{\beta}_2 = \frac{n-1}{\sum_{i=1}^{n} \ln(X_i)} \quad \text{og} \quad \hat{\beta}_3 = \frac{\sum_{i=1}^{n} \ln(X_i)}{n},$$

der ln er den naturlige logaritmen.

b) Hvilken av estimatorene over er sannsynlighetsmaksimeringsestimatoren (SME)? Begrunn svaret ved å utlede SME. Beregn estimatet når $n = 10$ og de observerte verdiene er som følger:

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
1.23	2.04	1.27	1.79	1.10	1.29	2.74	1.15	1.10	1.06

Det oppgis at $\sum_{i=1}^{10} \ln(x_i) = 3.39$.

c) Vis at $2\beta \ln(X_i)$ er kjikvadratfordelt med 2 frihetsgrader, og videre at $2\beta \sum_{i=1}^{n} \ln(X_i)$ er kjikvadratfordelt med $2n$ frihetsgrader.

Utled et 95% konfidensintervall for β. Hva blir intervallet når dataene er som i punkt b)?

Oppgave 2 Transport av masse

En entreprenør har leid inn et transportfirma til å transportere masse bort fra en byggeplass. Det er to mulige veivalg fra byggeplassen til stedet der massen skal deponeres, gjennom eller utenom bykjernen. Av helse, miljø og sikkerhetshensyn velger entreprenøren at massen skal transporteres utenom bykjernen.

Når oppdraget er ferdig har transportfirmaet kjørt 1000 turer, og transportfirmaet informerer entreprenøren om at av de 1000 turene har 5 blitt kjørt gjennom bykjernen og 995 er blitt kjørt utenom bykjernen.

I løpet av transportperioden har entreprenøren ved 5 tilfeldig valgte transporter sjekket om transporten har skjedd gjennom bykjernen. La X være antall ganger, av de 5 transportene som ble sjekket, entreprenøren finner at massen er blitt transportert gjennom bykjernen.

Hvilken fordeling har X? Begrunn svaret.

Hvilken verdi av X har høyest punktsannsynlighet?

Nå viser det seg at entreprenøren fant at i 5 av de 5 tilfellene han sjekket så ble massen transportert gjennom bykjernen. Hva er sannsynligheten for dette, dvs. $P(X = 5)$?

Oppgave 3 Trykkfasthet av murblokker

Ved en bedrift produseres en spesiell type murblokker, og vi skal i denne oppgaven se på trykkfastheten, Y, til murblokkene. Med trykkfasthet menes det maksimale trykket en murblokk kan utsettes for uten at den skades. Anta at Y er normalfordelt med forventningsverdi $\mu = \mathrm{E}(Y)$, gitt i MPa (10^6 Pascal), og standardavvik $\sigma = \mathrm{SD}(Y) = 0.21$ MPa.

a) Anta i dette punktet at $\mu = 2.10$ MPa.

Hva er sannsynligheten for at en tilfeldig valgt murblokk har en trykkfasthet som er høyere enn 1.83 MPa, dvs. $P(Y > 1.83)$?

Hva er sannsynligheten for at en tilfeldig valgt murblokk har en trykkfasthet som avviker mindre enn 0.3 MPa fra forventningsverdien $\mu = 2.10$ MPa?

Vi ser på måling av trykkfasthet for $n = 24$ tilfeldig valgte murblokker fra produksjonen. Hva er sannsynligheten for at den minste målingen vil være lavere enn 1.83 MPa?

Bedriften har startet produksjon av en ny type murblokker, som de mener har forventet trykkfasthet $\mu = 2.40$ MPa. Anta at det er kjent at standardavviket til trykkfastheten til de nye murblokkene er $\sigma = 0.21$ MPa. Bedriften ønsker å undersøke om det er grunn til å tro at forventet trykkfasthet for den nye typen murblokker er lavere enn 2.40 MPa.

b) Formulér dette som en hypotesetest ved å definere nullhypotese og alternativ hypotese.

Sett opp en testobservator og finn forkastningsområdet til testen, når vi ønsker å benytte signifikansnivå $\alpha = 0.05$.

Hva blir konklusjonen på testen, når vi har observert trykkfasthet til $n = 24$ murblokker, og gjennomsnittlig trykkfasthet ble 2.30 MPa?

Bestem teststyrken til den alternative hypotesten $H_1 : \mu = 2.30$ MPa for signifikansnivå 0.05 og $n = 24$ observasjoner.

Oppgave 4 Hubble

En viktig vitenskapelig oppdagelse fant sted i 1929 da Edwin Hubble oppdaget at universet er ekspanderende. Hubble's tallmateriale bestod blant annet av; x_i = avstanden til galakse i (målt i millioner lysår), og y_i = hastigheten til galakse i (målt i 1000 km/s). Verdiene Hubble benyttet i en av sine analyser er som følger:

Navn	Avstand, x_i	Hastighet, y_i
Virgo	22	1.2
Pegasus	68	3.8
Perseus	108	5.1
Coma Berenices	137	7.5
Ursa Major 1	255	14.9
Leo	315	19.2
Corona Borealis	390	21.4
Gemini	405	23.0
Bootes	685	39.2
Ursa Major 2	700	41.6
Hydra	1100	60.8

Det oppgis her at $\sum_{i=1}^{11} x_i = 4185$, $\sum_{i=1}^{11} y_i = 237.7$, $\sum_{i=1}^{11} x_i^2 = 2685141$ og $\sum_{i=1}^{11} x_i y_i = 152220$.

Hubble foreslo en modell for hastighet som funksjon av avstand på formen $y = \beta x$, der β senere har blitt kalt Hubble's konstant. En statistisk versjon av ligningen kan gis ved:

$$Y_i = \beta x_i + \varepsilon_i, \qquad i = 1, \dots, 11, \tag{1}$$

der ε_i, $i = 1, \dots, 11$, er uavhengige og normalfordelte stokastiske variabler med forventning 0 og varians σ^2.

a) Vi vil i første omgang finne en estimator for β.

Bruk minste kvadraters metode (method of least squares) til å estimere β med utgangspunkt i ligning (1), og vis at estimatoren for β da blir gitt ved $\hat{\beta} = \frac{\sum_{i=1}^{11} x_i Y_i}{\sum_{i=1}^{11} x_i^2}$. Regn ut estimatet for β basert på dataene over.

Finn også forventning og varians til $\hat{\beta}$.

b) Anta at en annen galakse befinner seg en avstand $x_0 = 900$ millioner lysår borte.

Finn predikert hastighet, $\hat{y}_0$, til denne galaksen.

Utled et 95% prediksjonsintervall for en måling av hastigheten til denne galaksen. Det oppgis at $\sum_{i=1}^{11} (y_i - \hat{y}_i)^2 = 9.87$, der $\hat{y}_i = \hat{\beta} x_i$.