



Faglege kontaktar under eksamen:

Mette Langaas 98847649

Ole Petter Løvdøen 99484158

EKSAMEN I FAG TMA4240 STATISTIKK

Torsdag 9. desember 2004

Tid: 09:00–13:00

Tillatne hjelpemiddel:

Gult A5-ark med egne handskrivne notatar.

Tabeller og formler i statistikk (Tapir Forlag).

K. Rottmann: *Matematisk formelsamling*.

Kalkulator: HP30S.

NYNORSK

Sensur: 6. januar 2005.

Oppgåve 1 Bronseboltar

Bronse er ei legering der kopar og tinn er dei viktigaste komponentane. Vi studerer koparinnhaldet i bronseboltar av ein gitt dimensjon som er laga av ei spesiell type bronselegering.

Ved bedrifta *Bronsespesialisten* vert det produsert bronseboltar, og det er teke stikkprøver av $n = 10$ bronseboltar frå produksjonen. Koparinnhaldet, X_i , $i = 1, \dots, n$, er målt. Vi føreset at X_1, X_2, \dots, X_n er uavhengjege og normalfordelte stokastiske variablar med $E(X_i) = \mu_x$ og $\text{Var}(X_i) = \sigma^2$.

- a) Vi føreset at forventninga er $\mu_x = 85$ gram og variansen er $\sigma^2 = 1$ gram² (kun i dette punktet).

Kva er sannsynet for at koparinnhaldet i ein tilfeldig valt bronsebolt er mindre enn 84 gram?

Finn eit tal, k , slik at sannsynet er 0.01 for at koparinnhaldet i ein tilfeldig valt bronsebolt er større enn k .

Vi ser på koparinnhaldet i to tilfeldig valte og uavhengjege bronseboltar. Kva er sannsynet for at koparinnhaldet i dei to bronseboltane avvik med meir enn 1.5 gram?

Vi føreset i resten av oppgåva at både μ_x og σ^2 er ukjende parametrar. Vi ser først på estimatorar for σ^2 .

- b) Kva eigenskapar kjenneteiknar ein god estimator?

To aktuelle estimatorar for σ^2 er $\hat{\sigma}^2$ og S^2

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

Finn forventningsverdien og variansen til dei to estimatorane, og kommenter. (Hint: bruk relasjon til kji-kvadrat-fordeling).

Ved bedrifta *Metalleksperten* vert det òg produsert bronseboltar av same legering og dimensjon som hjå Bronsespesialisten. Hjå Metalleksperten er det teke stikkprøver av $m = 10$ bronseboltar, og koparinnhaldet Y_j , $j = 1, \dots, m$, er målt. La Y_1, Y_2, \dots, Y_m vere uavhengjege og normalfordelte stokastiske variablar med $E(Y_j) = \mu_y$ og $\text{Var}(Y_j) = \sigma^2$, der både μ_y og σ^2 er ukjende parametrar. Vi føreset at målingar av koparinnhald i bronseboltar frå dei to bedriftene er uavhengjege. Vidare føreset vi at det er same (ukjende) varians for målt koparinnhald i bronseboltar frå dei to bedriftene, dvs. $\text{Var}(X_i) = \text{Var}(Y_j) = \sigma^2$.

Data for stikkprøvene frå Bronsespesialisten og Metalleksperten er presentert i tabell 1.

Bronsespesialisten	84.44	84.77	86.56	85.07	85.13	86.72	85.46	83.73	84.31	84.55
Metalleksperten	85.22	84.36	84.40	84.11	83.44	85.79	84.50	82.03	84.70	83.53

Tabell 1: Koparinnhald i bronselegering i stikkprøver frå Bronsespesialisten og Metalleksperten. Det vert gjeve opp at for Bronsespesialisten er $\sum_{i=1}^{10} x_i = 850.75$ og $\sum_{i=1}^{10} (x_i - \bar{x})^2 = 8.19$, og for Metalleksperten er $\sum_{j=1}^{10} y_j = 842.10$ og $\sum_{j=1}^{10} (y_j - \bar{y})^2 = 9.70$.

Leiinga i Bronsespesialisten har i den siste tida gjentekne gongar skulda Metalleksperten for at koparinnhaldet i bronseboltane fra Metalleksperten er lavare enn koparinnhaldet i bronseboltane frå Bronsespesialisten. Vi ønskjer å undersøkje om dette er tilfellet.

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

Set opp ein testobservator og finn forkastningsområdet. Kva blir konklusjonen på testen, med data gjeve i tabell 1, når signifikansnivået er $\alpha = 0.05$?

Rekn ut p -verdien ved å bruke tabell 2.

t	1.82	1.86	1.90	1.94	1.98	2.02	2.06
$\nu = 18$	0.957	0.960	0.963	0.966	0.968	0.971	0.973
$\nu = 19$	0.958	0.961	0.964	0.966	0.969	0.971	0.973
$\nu = 20$	0.958	0.961	0.964	0.967	0.969	0.972	0.974

Tabell 2: Kumulativt sannsyn i t -fordelinga. For T t -fordelt med ν friheitsgrader, så viser tabellen $P(T \leq t)$ for ulike verdier av t .

Metalleksperten har blitt utsett for hærverk og vindauget i hovudkontoret til bedrifta vart knust ved at ein bronsebolt vart kasta igjennom det. Politiet etterforskar saka, og er interessert i å finne ut om bronsebolten som knuste vindauget er produsert av Metalleksperten eller av Bronsespesialisten.

- d) Utlei eit intervall som med 95 % sannsyn inneheld ein ny observasjon av koparinnhaldet i ein bronsebolt frå Bronsespesialisten. Finn numeriske verdier for intervallet ved å bruke dataa frå tabell 1. Lag eit tilsvarande intervall for Metalleksperten.

Bronsebolten som vart brukt til å knuse vindauget hjå Metalleksperten vart målt til å ha eit koparinnhald på 86.30 gram. Kan du ut fra intervalla du har laga over seie noko om kva produsent som kan ha laga bronsebolten?

Oppgave 2 Pyramidespelet

Ole Petter har vorte spurt om å bli med i eit pyramidespel. Alt han treng å gjere, er å betale inn ein viss sum pengar, og deretter rekruttere fem nye personar, som vil bli plassert i nivået under han i pyramida. Etter at det er gjort, vert Ole Petter forsikra om at pengane vil byrje å strøyme inn. Ifølgje personen som spurde Ole Petter, vil ein person som blir beden om å ta del i pyramidespelet ha eit sannsyn $p = 1/3$ for å bli med, så det å få fem personar til å bli med, skal ikkje vere noko problem.

Før Ole Petter bestemmer seg, vil han rekne litt på sannsyn og forventningsverdiar, for å finne ut kor mykje arbeid rekrutteringsprosessen vil føre med seg. Han bestemmer seg for å sjå på den stokastiske variabelen X , som angjev antal personar Ole Petter må spørje, inntil den *første personen* blir med i pyramidespelet.

- a) Under kva føresetnadar vil X være geometrisk fordelt? I resten av oppgåva kan du føresette at X er geometrisk fordelt med punktsannsyn

$$f(x) = p(1-p)^{(x-1)} \quad \text{for } x = 1, 2, \dots$$

Om Ole Petter bestemmer seg for å delta i pyramidespelet, kva er forventa antal personar han må spørje for å få med ein ny person, når $p = 1/3$?

Kva er sannsynet for at han må spørje fleire enn fem personar for å få ein person til å delta i pyramidespelet, når $p = 1/3$?

Ole Petter er skeptisk til informasjonen om at kvar tredje person som blir spurt vil melde seg inn, og bestemmer seg for å estimere sannsynet p for at ein person som blir spurt vil delta i pyramidespelet. Han bestemmer seg for å samle inn data, ved at han tel antal personar han må spørje inntil den første svarer ja. Dette vil han prøve å få nokre kameratar til å hjelpe seg med. Den *ite* kameraten observerer at han må spørje X_i personar. Frå n undersøkingar set Ole Petter då igjen med n uavhengjege observasjonar X_1, X_2, \dots, X_n av den stokastiske variabelen X .

- b) Finn sannsynsmaksimeringsestimatoren \hat{p} for p basert på observasjonane X_1, X_2, \dots, X_n . Ingen av kameratane til Ole Petter er interessert i å hjelpe han med undersøkinga. Han set derfor igjen med kun den eine observasjonen som han har gjort sjølv, X_1 . Er sannsynsmaksimeringsestimatoren basert på kun denne eine observasjonen forventningsrett?

$$\text{Hint : } \sum_{n=1}^{\infty} \frac{1}{n} a^n = -\ln(1-a) ; 0 < a < 1$$

Oppg ve 3 Test nasjonen

Laurdag 27. november 2004 vart TV-programmet “Test nasjonen” sendt p  NRK1. I programmet vart 270 deltakarar i studio stilt sp rsm l innan ulike tema. Basert p  alder og antal rette svar fekk kvar deltakar tildelt ein IQ-score.

I programmet vart det opplyst at testen var laga slik at ein forventa at IQ-score til ein tilfeldig valt person skulle vere normalfordelt med forventningsverdi 100 og standardavvik 15.

Den maksimale IQ-score som vart registrert i studio var 122, og det var to personar som hadde denne IQ-scoren.

- a) Kva er sannsynet for at ein tilfeldig valt person skal f  ein IQ-score p  minst 122?

Om vi testar eit representativt utval p  270 personar, kva er d  forventa antal personar som f r ein IQ-score p  minst 122?

Kva er sannsynet for at maksimal IQ-score i eit tilfeldig utval av storleik 270 vil vere st rre enn 122?

Deltakarane i studio var delt inn i seks grupper, der ei av gruppene var sett saman av 42 tidlegare “Reality-deltakarar”. N r det gjeld IQ, bestemte vi oss f r programmet starta for at vi skulle sj  p  dei 42 tidlegare Reality-deltakarane som eit representativt utval fr  befolkninga. Reality-deltakarane fekk ein gjennomsnittleg IQ-score p  94.

- b) Vi f reset at IQ-score til ein tilfeldig valt person er normalfordelt med ukjent forventningsverdi μ , men kjend standardavvik 15. Utlei eit 95% konfidensintervall for forventa IQ-score basert p  data for dei tidlegare Reality-deltakarane.

Etter programmet har det blitt stilt sp rsm l om oppg vene i IQ-testen var for vanskelege, slik at IQ-scorane som vart oppn dde var lavare enn ein kunne forvente. Om vi f reset at IQ-score til ein tilfeldig valt person er normalfordelt med forventningsverdi 100 og standardavvik 15, kva er d  sannsynet for at ein i eit tilfeldig utval p  42 personar oppn r ein gjennomsnittleg IQ-score p  mindre eller lik 94?