Norges teknisk–
naturvitenskapelige universitet
Institutt for matematiske fag

Bokmål

Faglig kontakt under eksamen:
Steinar Engen 73 59 17 47/ 90 63 50 53
Arvid Næss 73 59 70 53/ 99 53 83 50

EKSAMEN I EMNE SIF5062/SIF5506 STATISTIKK

4. august 2003
Tid: 09:00–14:00

Hjelpemidler: *Tabeller og formler i statistikk*, Tapir Forlag
K. Rottmann: *Matematisk formelsamling*
Kalkulator HP30S

Sensuren faller: 25. august 2003

Oppgave 1

En laborant skal undersøke måleusikkerheten til et instrument som benyttes til å bestemme konsentrasjonen av et stoff i en oppløsning. Det gjennomføres n målinger med instrumentet på en oppløsning med kjent konsentrasjon $\mu = 1.0$ (mg/l). Observasjonene $X_1, X_2, \ldots, X_n$ kan antas å være uavhengige og normalfordelte med forventning μ og varians σ^2.

a) Anta i dette punktet at $\sigma^2 = 0.04$.

Hva er sannsynligheten for at en (bestemt) måling er mindre enn 1.3?

Finn sannsynligheten for at en måling avviker mer enn σ fra μ.

Gitt at en måling ikke avviker mer enn σ fra μ, hva er sannsynligheten for at målingen er større enn 1.0?

La $\bar{X} = \dfrac{1}{n} \sum_{i=1}^{n} X_i$ være gjennomsnittet av målingene.

b) Hva er fordelingen til $\bar{X}$? Begrunn svaret.

Hva er sannsynligheten for at gjennomsnittet $\bar{X}$ av 5 målinger avviker mer enn σ fra μ?

Vi skal nå finne en estimator for den ukjente variansen σ^2. To alternative estimatorer er foreslått:

$$\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2 \qquad (1)$$

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

c) Hvilke to egenskaper bør en god estimator ha?

Hvilken av de to estimatorene $\widehat{\sigma}^2$ og S^2 vil du foretrekke? Begrunn svaret. (Hint: χ^2-fordelingen).

d) Bruk estimatoren du valgte i c) til å utlede et $(1-\alpha)\cdot 100\%$ konfidensintervall for σ^2.

Hva blir konfidensintervallet når $n = 10$, $\sum_{i=1}^{10}(x_i - \mu)^2 = 0.50$, $\sum_{i=1}^{10}(x_i - \bar{x})^2 = 0.43$ og $\alpha = 0.05$?

Oppgave 2

En fjernsynsstasjon har et debattprogram der seerne kan sende inn SMS-meldinger med kommentarer og spørsmål underveis. Anta at ankomsten av SMS-meldinger kan modelleres som en Poisson-prosess. Hvis X betegner antall meldinger som kommer inn i løpet av ett minutt, vil X være Poissonfordelt med en passende valgt parameter μ, slik at punktsannsynligheten for X er

$$f(x) = \frac{\mu^x}{x!}\exp(-\mu); \quad \text{for} \quad x = 0, 1, 2, \ldots$$

La hendelsene A og B være gitt som $A: X > 10$ og $B: X \leq 14$.

a) Anta i dette punktet at $\mu = 10$.

Finn sannsynlighetene $P(A)$ og $P(B' \mid A)$, der B' er komplementærhendelsen til B.

Er A og B uavhengige? Begrunn svaret.

Anta at programmet varer i n minutter, og la $X_1, X_2, \ldots, X_n$ være antall SMS-meldinger i løpet av hvert av de n minuttene.

b) Finn sannsynlighetsmaksimeringsestimatoren (SME) for μ basert på de n observasjonene.

Er estimatoren forventningsrett?

Det viser seg at det er omtrent dobbelt så mange som ringer inn de siste $n/2$ minuttene som i de første $n/2$ (vi antar at n er et partall), og vi vil justere for dette i modellen. Vi antar fortsatt at $X_i,\ i = 1, \ldots, n$ er uavhengige og Poissonfordelte, men at $X_1, X_2, \ldots, X_{n/2}$ har parameter μ og $X_{n/2+1}, X_{n/2+2}, \ldots, X_n$ har parameter 2μ.

Det kan vises at i dette tilfellet blir SME for μ lik

$$\widetilde{\mu} = \frac{2}{3}\frac{\sum_{i=1}^{n} X_i}{n}.$$

La $Y = \dfrac{3n}{2}\widetilde{\mu}$.

c) Hvilken (eksakt) fordeling har Y?

Bruk denne fordelingen til å finne fordelingen til $\widetilde{\mu}$.

Forklar hvordan du kan bruke sentralgrenseteoremet til å vise at $\widetilde{\mu}$ er tilnærmet normalfordelt for store n.

Oppgave 3

En ny salve, B, for behandling av eksém på hendene skal sammenlignes med en tidligere brukt salve, A. Salvene brukes på $n = 15$ tilfeldig valgte pasienter som har eksém på begge hendene. Salve A brukes på den ene hånden og B på den andre, og det trekkes lodd slik at høyre hånd blir behandlet med A med sannsynlighet 0.5. Effekten av behandlingen vurderes etter en viss tid av en lege, som mener at 11 av de 15 pasientene har bedre effekt av B enn av A.

a) La Z være antall tilfeller der legen vurderer B til å gi bedre effekt enn A.

Anta at Z er binomisk fordelt med parametre n og p, der $p = P(\text{B er bedre enn A})$.

Finn sannsynligheten for at $Z \geq 11$ dersom $p = 0.7$.

En estimator for p er $\widehat{P} = \dfrac{Z}{n}$. Hvor stor måtte n være for at vi skulle være sikre på at standardavviket til $\widehat{P}$ alltid er mindre enn 0.1 (uansett verdien på p)?

Legen har i tillegg har en metode for å tallfeste effekten av salvene på en slik måte at høye verdier svarer til god effekt av salven. La $D_i,\ i = 1, \ldots, n$ være differansen mellom effektene av B og A for de n pasientene, slik at en positiv verdi for D_i betyr at salve B har bedre effekt enn salve A for pasient i.

De observerte verdiene for de 15 pasientene er

-0.3 0.3 1.0 0.5 -0.1 0.7 0.2 1.8 0.4 0.7 0.0 0.5 0.1 -0.1 -0.3

Det oppgis at $\sum_{i=1}^{15} d_i = 5.4$.

Anta at $D_1, D_2, \ldots, D_n$ er uavhengige og normalfordelte med forventning μ_D og varians σ_D^2. Anta videre at variansen σ_D^2 er kjent og lik $\sigma_D^2 = 0.25$. Basert på de observerte differansene vil vi teste om B er bedre enn A.

b) Formulér problemet som et hypotesetestingsproblem. Hvilken testobservator vil du bruke? Begrunn valget ditt. Hva blir resultatet av testen dersom du bruker signifikansnivå $\alpha = 0.05$?

c) Anta at den sanne forventningen er $\mu_D = 0.3$. Dersom du bruker testen i b), hva ville da sannsynligheten være for å oppdage at B er bedre enn A?

Dersom du bruker testen i b), hvor mange observasjoner må du ha for at sannsynligheten skal være minst 0.9 for å oppdage at B er bedre enn A?