Norges teknisk–
naturvitenskapelige universitet
Institutt for matematiske fag

Bokmål

Faglig kontakt under eksamen:
Øyvind Bakke 73 59 81 26
Håkon Tjelmeland 73 59 35 38

SIF5060/SIF5062/SIF5505/SIF5506 Statistikk

Onsdag 7. august 2002
Tid: 09:00–14:00

Hjelpemidler: Bestemt enkel kalkulator
Statistiske tabeller og formler, Tapir forlag.
K. Rottman: Matematisk formelsamling.

Sensuren faller: 28. august 2002.

Oppgave 1 Surhetsgrad i ferskvann

Man skal undersøke pH-verdien, μ, i et ferskvann. Til dette formål benyttes to ulike målemetoder, metode A og metode B. La X betegne målt verdi ved metode A og la Y tilsvarende betegne målt verdi med metode B. Anta at X og Y er uavhengige og at begge er normalfordelt med forventning μ. Målenøyaktigheten for de to målemetodene er forskjellige. Anta at variansen til X er $\sigma_A^2 = 0.2^2$, mens variansen til Y er $\sigma_B^2 = 0.1^2$. For å estimere μ er det foreslått tre estimatorer

$$\widehat{\mu} = Y \quad , \quad \widetilde{\mu} = \frac{1}{2}X + \frac{1}{2}Y \quad \text{og} \quad \mu^* = \frac{1}{5}X + \frac{4}{5}Y.$$

a) Hvilke to egenskaper kjennetegner en god estimator?

Hvilken av estimatorene $\widehat{\mu}$, $\widetilde{\mu}$ og μ^* vil du foretrekke? Begrunn svaret.

b) Hvilken sannsynlighetsfordeling har μ^*? Begrunn svaret.

Utled et 90%-konfidensintervall for μ ved å ta utgangspunkt i estimatoren μ^*.

Oppgave 2 Helseundersøkelse

Man ønsker å estimere hvor stor andel av befolkningen i Trøndelag som lider av en bestemt sykdom. La $\alpha \in [0,1]$ betegne den ukjente andelen som skal estimeres. For å avgjøre om en person har sykdommen kan man ta en prøve. Dessverre viser prøven ikke alltid riktig konklusjon. Det er kjent at for en person som har sykdommen vil prøven slå ut (dvs. indikere at personen har sykdommen) med sannsynlighet 9/10, mens den med sannsynlighet 1/10 ikke vil slå ut. For en person som ikke har sykdommen vil prøven slå ut med sannsynlighet 1/20 og ikke slå ut med sannsynlighet 19/20. For en tilfeldig valgt person, la hendelse A være at han/hun har sykdommen og la hendelse B være at hans/hennes prøve slår ut.

a) Tegn hendelsene A og B inn i et Venn-diagram.

Bestem sannsynligheten for at en tilfeldig valgt person både har sykdommen og at prøven hans slår ut. (Gi svaret som funksjon av α.)

Vis at sannsynligheten for at prøven til en tilfeldig valgt person slår ut er

$$\frac{17\alpha + 1}{20}.$$

For å estimere α trekker man tilfeldig n personer bosatt i Trøndelag og utfører prøven på disse. La X betegne antall personer hvor prøven slår ut. Vi antar at X er binomisk fordelt med n forsøk og suksessannsynlighet $p = (17\alpha + 1)/20$.

b) Vis at sannsynlighetsmaksimeringsestimatoren (SME) for α kan skrives

$$\widehat{\alpha} = \begin{cases} 0 & \text{hvis } X < \frac{n}{20}, \\ \frac{1}{17}\left(\frac{20X}{n} - 1\right) & \text{hvis } \frac{n}{20} \leq X \leq \frac{18n}{20}, \\ 1 & \text{hvis } X > \frac{18n}{20}. \end{cases}$$

Når n er stor vil sannsynligheten for å få $\widehat{\alpha} = 0$ eller $\widehat{\alpha} = 1$ være svært liten og man kan med god approksimasjon se bort fra disse tilfellene. Dette betyr altså at man i stedet for $\widehat{\alpha}$ benytter estimatoren

$$\widetilde{\alpha} = \frac{1}{17}\left(\frac{20X}{n} - 1\right).$$

c) Vis at

$$\mathrm{E}(\widetilde{\alpha}) = \alpha \qquad \text{og} \qquad \mathrm{Var}(\widetilde{\alpha}) = \frac{1}{n}\left(\alpha + \frac{1}{17}\right)\left(\frac{19}{17} - \alpha\right).$$

d) Redegjør kort for det generelle resultatet i sentralgrenseteoremet.

Vis hvordan man fra sentralgrenseteoremet får at $\widetilde{\alpha}$ er tilnærmet normalfordelt når n er stor.

Fra tidligere studier vet man at i hele Norges befolkning er det 5.3% av befolkningen som lider av den aktuelle sykdommen. Det ønskes nå å avgjøre om det er grunnlag for å påstå at hyppigheten av sykdommen er større i Trøndelag enn i Norge generelt.

e) Formuler dette som en hypotesetestingsproblem og lag en test basert på $\widetilde{\alpha}$ for dette formål med (tilnærmet) signifikansnivå lik 5%.

Hva blir konklusjonen på testen når 1000 personer i Trøndelag ble undersøkt og prøven slo ut for 101 av disse.

Oppgave 3 Levetid for lyspærer

To typer lyspærer, A og B, har levetider henholdsvis X og Y, der X og Y antas uavhengige. Videre antas at X har sannsynlighetstetthet $f_1(x)$ og Y har sannsynlighetstetthet $f_2(y)$, der

$$f_1(x) = \begin{cases} \frac{1}{\beta_1} e^{-x/\beta_1} & \text{for } x \geq 0 \\ 0 & \text{ellers} \end{cases} \quad , \quad f_2(y) = \begin{cases} \frac{1}{\beta_2} e^{-y/\beta_2} & \text{for } y \geq 0 \\ 0 & \text{ellers} \end{cases}$$

og $\beta_1, \beta_2 > 0$.

a) Vis at forholdet mellom forventet levetid for pære A og forventet levetid for pære B er β_1/β_2.

La $U = X/Y$ være forholdet mellom levetidene og la $V = Y$.

b) Finn simultan sannsynlighetstetthet for U og V.

Vis at marginal sannsynlighetstetthet for U er

$$g(u) = \begin{cases} \frac{\beta_1 \beta_2}{(\beta_1 + \beta_2 u)^2} & \text{for } u \geq 0, \\ 0 & \text{ellers.} \end{cases}$$

c) Vis at $\mathrm{E}(U)$ er uendelig.