Noregs teknisk–
naturvitskaplege universitet
Institutt for matematiske fag

Nynorsk

Fagleg kontakt under eksamen:

Steinar Engen	73 59 17 47/ 90 63 50 53
Arvid Næss	73 59 70 53/ 99 53 83 50
Turid Follestad	73 59 35 37

EKSAMEN I EMNE SIF5062/SIF5506 STATISTIKK

Laurdag 24. mai 2003

Tid: 09:00–14:00

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

Sensuren fell: 28. juni 2003

Oppgåve 1

Lat X vere høgda til ei tilfeldig valt 6-årig jente. Vi går ut ifrå at høgda er normalfordelt med forventning $\mathrm{E}(X) = 115$ cm og standardavvik $\mathrm{SD}(X) = 5$ cm.

a) Rekn ut sannsyna

$$P(X \leq 120) \text{ og } P(120 < X \leq 125).$$

Vi let videre følgjande hendingar vere definerte:

$A: \quad X \leq 120$
$B: \quad X > 125$

Er A og B disjunkte? Er A og B uavhengige? Grunngi svara.

Vi tenkjer oss at vi går til år 2010, og er interesserte i jentenamn blant førsteklassingar, dvs. blant dei som er fødde i 2004. Namnestatistikken viser at 2% av jentene som vart fødde dette året fekk namnet Maud.

b) Lat Z vere talet på jenter som heiter Maud i ei tilfeldig valt førsteklasse der det er n jenter. Forklar kvifor det er rimeleg å gå ut ifrå at at Z er binomisk fordelt med parametrar n og p, der $p = 0.02$.

Vi vil i resten av oppgåva gå ut ifrå at Z er binomisk fordelt med $n = 15$ og $p = 0.02$. Lat hendingane C og D vere definerte ved

C : minst ei av jentene i klassa heiter Maud
D : akkurat to jenter i klassa heiter Maud

Rekn ut sannsyna $P(C)$ og $P(D \mid C)$.

Gå ut ifrå at det er totalt 25 elevar i klassa. Kva er då sannsynet for at ein tilfeldig valt elev blant alle elevane i klassa heiter Maud?

Oppgåve 2

Produsenten av ein bestemt bilmodell hevdar at denne modellen kan ventast å køyre minst 16 km pr. liter bensin på motorveg. Forbrukarorganisasjonen FO testar denne påstanden ved å køyre eit tilfeldig utval bilar av denne modellen ein passende distanse på ein representativ motorveg og måle bensinforbruket.

På bakgrunn av erfaringar frå tidlegare forsøk av same type, går FO ut ifrå at bensinforbruket til ein tilfeldig valt bil av den modellen som blir testa, kan modellerast med god tilnærming som ein normalfordelt tilfeldig variabel X med forventningsverdi μ og varians σ^2, dvs. $X \sim N(\mu, \sigma^2)$. Både forventningsverdien μ og standardavviket σ er i utgangspunktet ukjente storleikar.

Av praktiske grunnar avgrensar FO storleiken på det tilfeldige utvalet til $n = 20$ bilar. Etter forsøket vart alle målingane analyserte, og resulterte i ein gjennomsnittsverdi $\bar{x} = 15.56$ og eit sample (empirisk) standardavvik $s = 0.94$.

a) Set opp ein hypotesetest for dette forsøket. Lat produsenten sin påstand representere nullhypotesen. Kva for ein testobservator vil du bruke for å kontrollere hypotesen? Grunngi kort valet ditt. I forhold til eit valt signifikansnivå $\alpha = 0.05$, vil du akseptere produsenten sin påstand?

b) Finn P-verdien (signifikanssannsynet) for testen i punkt a) som svarer til dei observerte verdiane.

Kva for ei tilnærming kan du gjere for at testobservatoren skal bli normalfordelt? Kva for ein P-verdi får du dersom du bruker denne tilnærminga?

c) Bestem teststyrken for den alternative hypotesen $\mathrm{H}_1' : \mu = 15.5$ for signifikansnivå $\alpha = 0.05$ ved å bruke den same normaltilnærminga som i punkt b). Gi eit forslag til korleis teststyrken kan aukast.

Oppgåve 3

Eit apparat inneheld k like komponentar og fungerer berre dersom alle desse er i orden. Komponentane sine levetider $T_1, T_2, \ldots T_k$ er uavhengige og eksponensielt fordelte med parameter β (> 0), dvs. sannsynstettleiken er

$$f(t) = \begin{cases} \frac{1}{\beta} e^{-t/\beta} & \text{for} \quad t \geq 0, \\ 0 & \text{for} \quad t < 0. \end{cases}$$

a) Finn den kumulative fordelingsfunksjonen for levetida til ein komponent. Kva blir $P(T_1 < 3)$ og $P(2 < T_1 < 4)$ når $\beta = 5$?

b) Lat X vere apparatet si levetid. Vis at X er eksponensielt fordelt med parameter β/k. Kva blir den forventa levetida til apparatet når $k = 4$ og $\beta = 5$?

Verksemda har laga fleire utgåver av apparatet med ulikt tal på komponentar. Apparatet fungerer betre med mange komponentar, men har samtidig kortare forventa levetid. Lat $X_1, X_2, \ldots,$ X_n vere levetidene for n apparat med $k_1, k_2, \ldots, k_n$ komponentar i gitt rekkefølgje. To estimatorar for β basert på levetidene til apparata er føreslegne,

$$\hat{\beta} = \frac{1}{n} \sum_{i=1}^{n} X_i k_i$$

og

$$\tilde{\beta} = \frac{\sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} k_i^{-1}}.$$

c) Finn forventningsverdi og varians til begge estimatorane.

d) Vis at ein av estimatorane i pkt. (b) er sannsynsmaksimeringsestimatoren (SME) og vis at denne har varians som alltid er mindre enn eller lik variansen til den andre estimatoren.

Hint: Set $r_i = 1/k_i$ og bruk resultatet $\frac{1}{n} \sum r_i^2 - (\frac{1}{n} \sum r_i)^2 \geq 0$.

e) Vis at $2k_i X_i/\beta$ er χ^2-fordelt med 2 fridomsgrader og at $2n\hat{\beta}/\beta$ er χ^2-fordelt. Bruk dette til å utleie eit $(1-\alpha) \cdot 100\%$ konfidensintervall for β. Kva blir intervallet dersom $\alpha = 0.05$, $n = 8$ og $\hat{\beta} = 8.3$?