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
Turid Follestad	73 59 35 37

EKSAMEN I EMNE SIF5062/SIF5506 STATISTIKK

Lørdag 24. mai 2003
Tid: 09:00–14:00

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

Sensuren faller: 28. juni 2003

Oppgave 1

La X være høyden til en tilfeldig valgt 6-årig jente. Vi antar at høyden er normalfordelt med forventning $\mathrm{E}(X) = 115$ cm og standardavvik $\mathrm{SD}(X) = 5$ cm.

a) Beregn sannsynlighetene

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

Vi lar videre følgende hendelser være definert:

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

Er A og B disjunkte? Er A og B uavhengige? Begrunn svarene.

Vi tenker oss at vi går til år 2010, og er interessert i jentenavn blant førsteklassinger, dvs. blant de som er født i 2004. Navnestatistikken viser at 2% av jentene som ble født dette året fikk navnet Maud.

b) La Z være antall jenter som heter Maud i en tilfeldig valgt førsteklasse der det er n jenter. Forklar hvorfor det er rimelig å anta at Z er binomisk fordelt med parametre n og p, der $p = 0.02$.

Vi vil i resten av oppgaven anta at Z er binomisk fordelt med $n = 15$ og $p = 0.02$. La hendelsene C og D være definert ved

C : minst en av jentene i klassen heter Maud
D : akkurat to jenter i klassen heter Maud

Beregn sannsynlighetene $P(C)$ og $P(D \mid C)$.

Anta at det er totalt 25 elever i klassen. Hva er da sannsynligheten for at en tilfeldig valgt elev blant alle elevene i klassen heter Maud?

Oppgave 2

Produsenten av en bestemt bilmodell hevder at denne modellen kan forventes å kjøre minst 16 km pr. liter bensin på motorvei. Forbrukerorganisasjonen FO tester denne påstanden ved å kjøre et tilfeldig utvalg biler av denne modellen en passende distanse på en representativ motorvei og måle bensinforbruket.

På bakgrunn av erfaringer fra tidligere forsøk av samme type, antar FO at bensinforbruket til en tilfeldig valgt bil av den modellen som testes, kan modelleres med god tilnærmelse som en 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 størrelser.

Av praktiske grunner begrenser FO størrelsen på det tilfeldige utvalget til $n = 20$ biler. Etter forsøket ble alle målingene analysert, og resulterte i en gjennomsnittsverdi $\bar{x} = 15.56$ og et sample (empirisk) standardavvik $s = 0.94$.

a) Sett opp en hypotesetest for dette forsøket. La produsentens påstand representere nullhypotesen. Hvilken testobservator vil du bruke for å kontrollere hypotesen? Gi en kort begrunnelse for valget ditt. I forhold til et valgt signifikansnivå $\alpha = 0.05$, vil du akseptere produsentens påstand?

b) Finn P-verdien (signifikanssannsynligheten) for testen i punkt a) som svarer til de observerte verdiene.

Hvilken tilnærmelse kan du gjøre for at testobservatoren skal bli normalfordelt? Hvilken P-verdi får du hvis du bruker denne tilnærmelsen?

c) Bestem teststyrken for den alternative hypotesen $\mathrm{H}_1' : \mu = 15.5$ for signifikansnivå $\alpha = 0.05$ ved å bruke den samme normaltilnærmelsen som i punkt b). Gi et forslag til hvordan teststyrken kan økes.

Oppgave 3

Et apparat inneholder k like komponenter og fungerer bare dersom alle disse er i orden. Komponentenes levetider $T_1, T_2, \ldots T_k$ er uavhengige og eksponensielt fordelte med parameter β (> 0), dvs. sannsynlighetstettheten er

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

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

b) La X betegne apparatets levetid. Vis at X er eksponensielt fordelt med parameter β/k. Hva blir apparatets forventede levetid når $k = 4$ og $\beta = 5$?

Bedriften har laget flere utgaver av apparatet med forskjellig antall komponenter. Apparatet fungerer bedre med mange komponenter, men har samtidig kortere forventet levetid. La $X_1, X_2, \ldots, X_n$ være levetidene for n apparater med hhv. $k_1, k_2, \ldots, k_n$ komponenter. To estimatorer for β basert på apparatenes levetider er foreslått,

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

og

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

c) Finn forventningsverdi og varians til begge estimatorene.

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

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

e) Vis at $2k_iX_i/\beta$ er χ^2-fordelt med 2 frihetsgrader og at $2n\hat{\beta}/\beta$ er χ^2-fordelt. Bruk dette til å utlede et $(1-\alpha)\cdot 100\%$ konfidensintervall for β. Hva blir intervallet hvis $\alpha = 0.05$, $n = 8$ og $\hat{\beta} = 8.3$?