Norges teknisk-naturvitenskapelige universitet
Institutt for matematiske fag

Bokmål

Faglig kontakt under eksamen:

Arvid Næss	73 59 70 53/ 99 53 83 50
Jarle Tufto	99 70 55 19
Ola Diserud	93 21 88 23

EKSAMEN I EMNE TMA4245 STATISTIKK

3. juni 2010
Tid: 09:00–13:00

Hjelpemidler: *Tabeller og formler i statistikk*, Tapir Forlag
K. Rottmann: *Matematisk formelsamling*
Kalkulator HP30S
Gult, stemplet A5-ark med egne håndskrevne notater.

Oppgave 1 En sjelden genvariant a disponerer for en bestemt sykdom. I en populasjon er den relative frekvensen av individ som bærer to kopier av genvarianten (individ av genotype aa) 0.0001, frekvensen av individ som bærer én kopi (genotypen Aa) er 0.0198 og frekvensen av individ som ikke har den sjeldne genvarianten (genotypen AA) er 0.9801. Anta at sannsynligheten for at sykdommen kommer til uttrykk blant personer med genotype aa, aA og AA er henoholdvis 0.6, 0.02 og 0.01.

a) Hva er sannsynligheten for at sykdommen kommer til uttrykk i et tilfeldig valgt individ fra populasjonen?

b) Hva er de respektive sannsynlighetene for at et individ er av genotypene aa, Aa og AA gitt at sykdommen har kommet til uttrykk i individet?

Oppgave 2 En fabrikk produserer en spesiell type maskinkomponenter. Tiden fra en komponent blir tatt i bruk til den bryter sammen for første gang, kaller vi levetiden for komponenten. Erfaring har vist at levetiden T, målt i uker, kan modelleres som en kontinuerlig stokastisk variabel med sannsynlighetstetthet (sannsynlighetsfordeling):

$$\begin{aligned} f_T(t) &= 2\lambda t e^{-\lambda t^2}\,, \quad t \geq 0\,, \\ &= 0\,, \qquad \text{ellers}\,, \end{aligned} \tag{1}$$

der $\lambda > 0$ er en ukjent parameter.

a) Bestem kumulativ fordelingsfunksjon (KF) $F_T(t)$ til T, og beregn $P(20 < T \leq 30)$ når $\lambda = 1.5 \cdot 10^{-3}$.

b) Parameteren λ skal estimeres på basis av levetidene $T_1, \ldots, T_n$ for $n > 2$ tilfeldig valgte komponenter. $T_1, \ldots, T_n$ antas uavhengige og identisk fordelte med sannsynlighetstetthet $f_T(t)$. Vis at sannsynlighetsmaksimerings-estimatoren (SME) for λ blir

$$\Lambda^* = \frac{n}{\sum_{i=1}^n T_i^2}\,. \tag{2}$$

c) La X være χ^2-fordelt med $2n$ frihetsgrader ($n > 2$). Vis at

$$\mathrm{E}(X^{-1}) = \frac{1}{2(n-1)} \quad \text{og} \quad \mathrm{E}(X^{-2}) = \frac{1}{4(n-1)(n-2)}\,. \tag{3}$$

Vis at $Y = 2\lambda T^2$ er χ^2-fordelt med 2 frihetsgrader (husk at $T \geq 0$). Bruk så dette resultatet sammen med ligning (3) til å undersøke om Λ^* er forventningsrett. Om nødvendig, korrigér for å få en estimator som er forventningsrett. Bestem også denne estimatorens varians. (Hint: De resultatene som kan hjelpe deg her, står i 'Tabeller og formler i statistikk'.)

d) Utled et $100(1-\alpha)\%$ konfidensintervall for λ. Bestem intervallet numerisk når $\alpha = 0.05$, $n = 5$ og de observerte verdiene er

23.63 35.97 18.65 18.18 11.59

e) En dag oppdages feil ved produksjonsprosessen. Komponentene testes ved at 5 komponenter velges tilfeldig fra dagsproduksjonen, og levetiden bestemmes for disse ved akselerert levetidstesting. Observasjonene brukes til å teste

$$H_0 \,:\, \lambda \leq \lambda_0 = 1.5 \cdot 10^{-3}$$

mot

$$H_1 \,:\, \lambda > \lambda_0 = 1.5 \cdot 10^{-3}$$

Påvis at $2\lambda_0 \sum_{i=1}^{n} T_i^2$ kan brukes som testobservator, og bestem det kritiske området for testen for signifikansnivå α. Konkludér for $\alpha = 0.05$ når observasjonene er

$$12.06 \quad 18.02 \quad 19.86 \quad 16.60 \quad 9.36$$

Komponentene blir pakket i kasser med 5 komponenter i hver kasse. Fabrikken garanterer at alle komponentene i en kasse har levetid på minst a uker. Fabrikken får reklamasjon på en kasse hvis én eller flere av de 5 komponentene i kassen har levetid kortere enn a uker.

f) Hvis $\lambda = 1.5 \cdot 10^{-3}$, hvor stor kan a velges for at sannsynligheten for reklamasjon skal være høyst 0.05? Du kan anta uavhengige levetider.

g) La a og λ være som i punkt f), og anta at i et visst tidsrom selges 1000 kasser. La U være antall kasser som det blir reklamert på. Hvilke forutsetninger må en gjøre for at U skal være binomisk fordelt? Anta at disse forutsetningene er oppfylt og bestem $P(U \leq 60)$ ved å bruke tilnærming til normalfordeling.