Chalmers
Matematiska vetenskaper

TMS051: Matematisk statistik för Z, del B

Tentamen 13 december 2003 em V

Tillåtna hjälpmendel är räknedosa utan information om kursen i minnena, Beta samt kursens formel- och tabellsamling.

För betyget 3 krävs 12 p, för 4:a 18 p och för 5:a 24 p av totalt 30 p.

Jour och examinator är Tommy Norberg, ankn 3528 eller 0730 794209 till kl 16^{45}.

Observera att ofullständiga uträkningar eller motiveringar till resultat eller beräkningar ofelbart leder till poängavdrag för den som har ambitionen att uppnå betyget 4:a eller bättre, t.ex gäller att svar av typen ja, nej (j.f.r uppgift 4a) ej godkänns utan korrekt motivering.

Uppgifter

1. Man har utvecklat ett nytt test som, för individer i åldersgruppen 50+, ska detektera en viss typ av arthritis. Känt är att ca 12% av individerna i åldersgruppen 50+ har besvär av denna typ av arthritis. Man utsatte individer om vilka man redan visste att de besväras av denna typ av arthritis för testet och det visade sig då att testet indikerade arthritis korrekt i ca 87% av fallen. När man sedan testade garanterat friska individer i åldersgruppen, visade det sig att testet indikerade arthritis felaktigt i ca 6% av fallen. Bestäm approximativt sannolikheten att en testad individ besväras av denna typ av arthritis givet att testet indikerat att så är fallet. (3 p)

2. I snitt fungerar 95% av alla bläckstråleskrivare av ett visst fabrikat och typ direkt efter installation i ett hem. Övriga behöver någon form av intrimmning. En försäljare säljer under januari månad 7 st sådana skrivare. (a) Hur stor är sannolikheten att minst 6 av dessa fungerar vid installationen? Antag att försäljaren under ett halvår säljer exakt 7 st sådana skrivare varje månad. (b) Hur stor är sannolikheten att minst 6 av dessa 7 fungerar vid installationen i alla sex månaderna? Sammanlagt säljer alltså försäljaren under halvåret 42 st sådana skrivare. (c) Hur stor är sannolikheten att minst 36 av dessa fungerar vid installationen. (3 p)

3. Visa den s.k "glömskeegenskapen" för exponentialfördelningen, alltså att

$$P(T > s + t \mid T > s) = P(T > t)$$

gäller för exponentialfördelade stokastiska variabler T. (4 p)

4. (a) Gäller centrala gränsvärdessatsen för Cauchydata, alltså för data som följer tätheten

$$f(x) = \frac{1}{\pi(1+x^2)} \quad \text{för} \quad -\infty < x < \infty$$ (2 p)

   (b) Låt X, Y vara bivariat normalfödelade med väntevärden $\mu_x = 3.5$, $\mu_y = 7.5$, standardavvikelser $\sigma_x = 1.5$, $\sigma_y = 2$ och korrelation $\rho = 0.75$. Låt $Z = 2X - Y$. Beräkna Z varians. (2 p)

(Vänd!)

5. Under de senaste 10 åren har antalet allvarliga olycksfall i arbetet varit

$$8 \quad 6 \quad 12 \quad 15 \quad 12 \quad 4 \quad 9 \quad 7 \quad 20 \quad 10$$

Data gäller för en viss region i södra Sverige. Antag att antalet olyckor av denna typ under ett givet år följer en Poissonfördelning med parameter λ och att antalet olyckor under olika år är oberoende. Härled med någon metod en punktskattning $\widehat{\lambda}$ av λ. Är $\widehat{\lambda}$ väntevärdesriktig? Beräkna det observerade utfallet av $\widehat{\lambda}$. Diskutera modellen. Är det t.ex något i den som du tycker är tvivelaktigt och som möjligen gör resultatet av analysen tveksamt? Inga utsvävningar, tack, utan formulera dig kort och koncist. (4 p)

6. En viss typ av tillverkning har felsannolikheten $p = 0.15$. Man tror att man medelst diverse smarta justeringar kan få ned felsannolikheten till ca 0.10 och man vill "designa" ett försök så att det finns en rimlig chans ($\geq$ 90%) att detektera en förbättring samtidigt som man håller risken (sannolikheten) att göra ett s.k typ-I-fel mindre än 5%. Redogör för hur försöket ska genomföras och för hur analysen av försöksresultatet ska göras samt vilka fel som kan göras och hur stora riskerna är för dessa. (4 p)

7. En hyfsad modell för tiden tills en viss typ av halvledarkomponent upphör att fungera tillfredsställande är Exp(λ). Under en tid har man samlat på sig observationer av $n = 25$ sådana komponenters resp funktionstider. Observerat medelvärde och standardavvikelse är

$$\bar{x} = 10719.0 \quad \text{och} \quad s = 10366.9$$

(tidsenheten är timmar). Intervallskatta parametern λ för den aktuella typen av halvledare. (4 p)

8. I den linjära regressionsmodellen $Y|x \sim \mathrm{N}(\beta_0 + \beta_1 x, \sigma)$ gäller att variansen skattas av

$$s^2 = \frac{1}{n-2}\sum_{i=1}^{n}(y_i - b_0 - b_1 x_i)^2$$

där $b_1 = S_{xy}/S_{xx}$ och $b_0 = \bar{y} - b_1\bar{x}$. Visa att

$$(n-2)s^2 = S_{yy} - S_{xy}^2/S_{xx} \qquad \text{(4 p)}$$

— Lycka till —