TMS051 Matematisk statistik och simuleringsteknik, Z, del A

Tentamen 31 augusti 2006 fm

Även omtentamen på TMS050 Matematisk statistik och simuleringsteknik Z, del A.

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

Examinator är Tommy Norberg, ankn 3528 eller 0730 79 42 09. Tommy går att nås per mobiltelefon under tentamen.

Maximalt antal tentamenspoäng är 30, av dessa krävs 12 för godkänt betyg på tentan och 18 resp 24 för 4:a och 5:a. Lösningar går att ladda ner från kurshemsidan. Rättningsprotokoll anslås i MV:F, plan 2.

Svar skall motiveras om ej annat sägs i uppgiften.

Uppgifter

1. Varje dag skippas 20 enheter till en sammansättningsindustri. Fyra av dessa sorteras ut slumpmässigt och funktionskontrolleras. Antag att sex enheter ej fungerar tillfredsställande. Hur stor är sannolikheten att man hittar

   (a) exakt två enheter som ej fungerar tillfredsställande? (2 p)

   (b) minst två enheter som ej fungerar tillfredsställande? (2 p)

2. När Kvast-Hilda, stadens mest kända häxa, kokar ihop en soppa som kan förvandla en groda till en prinsessa händer det ibland att hon istället får en soppa som istället förvandlar grodan till en häst. Vis av mångårig erfarenhet vet Kvast-Hilda att detta händer ca 3 gånger av 17. Under tre på varandra följande veckor kokar Kvast-Hilda denna soppa. Ungefär hur stor är sannolikheten att det blir fel

   (a) en gång? (2 p)

   (b) minst en gång? (2 p)

3. Beräkna väntevärde och 75%-kvantil för exponentialfördelningen. (3 p)

4. Antag att tiden tills en viss utrustning "går ner" (alltså slutar att fungera) är exponentialfördelad med väntevärde μ. Parametern $\lambda = 1/\mu$ brukar då kallas för intensiteten. I ett test av 7 sådana utrustningar uppmättes funktionstiderna

   76.6   4.3   3.8   1.7   30.1   4.2   22.5

   ML-skatta intensiteten λ. (4 p)

5. Man mätte tiden t för utförandet av en viss uppgift i sammansättningen av en viss produkt. I $n = 100$ på varandra följande utföranden erhölls $\sum_i t_i = 7436$ s, $\sum_i t_i^2 = 553948$. Under lämpliga villkor kan man punkt- och intervallskatta väntevärdet μ. (a) Vilka är dessa villkor? (b) Gör nu punkt- och intervallskattningen. Den senare ska ha konfidensen 99%. (c) Granska kritiskt villkoren. Finns det något — och i så fall, vilket — som kanske inte är uppfyllt? (1+3+1 p)

6. Under samma villkor som i ovanstående problem kan man intervallskatta standardavvikelsen σ. Intervallet ska vara uppåt begränsat (men ej nedåt) och konfidensen ska här vara 99%. (3 p)

7. Nu när vi börjar närma oss valtider görs många politiska partundersökningar. I en fick alliansen 50.1% medan det andra blocket tillsammans erhöll 45.2%. Antag att 2 000 personer besvarade enkäten och att den är gjord enligt vedertagen statistisk metodik. Intervallskatta med konfidensen 95% skattningen av sannolikheten att en godtycklig väljare röstar på något av allianspartierna. (3 p)

8. Antar man att inga nya partier kommer in i riksdagen, så är det bara $(0.501 + 0.452) \cdot 2\,000 = 1\,906$ av de 2 000 avlagda svaren som bör räknas, eftersom det är rösterna på partierna som kommer in i riksdagen som avgör mandatfördelningen. Är alliansens ledning med detta sätt att räkna statistiskt säkerställd? (4 p)

Lycka till!

Lösningar till TMS051, del A den 31/8-06

1. (a) $p = \binom{6}{2}\binom{14}{2}/\binom{20}{4} = 0.282$

   (b) $p = 1 - \left(\binom{6}{0}\binom{14}{4}/\binom{20}{4} + \binom{6}{1}\binom{14}{3}/\binom{20}{4}\right) = 0.343$

2. (a) $p = \binom{3}{1}\left(\frac{3}{17}\right)^1\left(\frac{14}{17}\right)^2 = 0.359$

   (b) $p = 1 - \binom{3}{0}\left(\frac{3}{17}\right)^0\left(\frac{14}{17}\right)^3 = 0.441$

3. Exponentialfördelningens täthet är $f(x) = \lambda e^{-\lambda x}$ eller $f(x) = \frac{1}{\beta}e^{-x/\beta}$, beroende på om man väljer intensiteten λ eller $\beta = 1/\lambda$ som parameter. Utfallsrummet är positiva talaxeln, så $x > 0$, Med den förstnämnda parametriseringen fås

$$\mu = \int_0^\infty x\,\lambda e^{-\lambda x}\,dx = \{\text{part int}\} = 1/\lambda$$

   och

$$0.75 = \int_0^{x_{0.75}} \lambda e^{-\lambda x}\,dx = 1 - e^{-\lambda x_{0.75}}$$

$$\Rightarrow \quad 0.25 = e^{-\lambda x_{0.75}} \quad \Rightarrow \quad x_{0.75} = \frac{-\ln 0.25}{\lambda} = \frac{\ln 4}{\lambda}$$

4. Tätheten, uttryckt m.h.a parametern λ, är $f(x) = \lambda e^{-\lambda x}$, $x > 0$. Utfallet $x_1, \ldots, x_n$ har därför troligheten

$$L(\lambda) = f(x_1)\cdot \ldots \cdot f(x_n) = \lambda e^{-\lambda x_1}\cdot \ldots \cdot \lambda e^{-\lambda x_n} = \lambda^n e^{-\lambda \sum_i x_i}$$

   ML-skattningen är det λ som maximerar trolighetsfunktionen. Logaritmera, sätt derivatan till noll och lös ut λ:

$$\mathcal{L}(\lambda) = \log L(\lambda) = n\log\lambda - \lambda\sum_i x_i$$

$$\mathcal{L}'(\lambda) = \frac{n}{\lambda} - \sum_i x_i$$

$$\mathcal{L}'(\lambda) = 0 \quad \Leftrightarrow \quad \frac{n}{\lambda} = \sum_i x_i \quad \Leftrightarrow \quad \lambda = \frac{n}{\sum_i x_i} = \frac{1}{\bar{x}}$$

   ML-skattningen är således $\hat{\lambda} = 1/\bar{x}$. Till sist, $\sum_i x_i = 76.6 + 4.3 + 3.8 + 1.7 + 30.1 + 4.2 + 22.5 = 143.2 \Rightarrow \hat{\lambda} = 7/143.2 = 1/20.457 = 0.0489$.

5. (a) Observationerna måste vara oberoende och likafördelade

   (b) Vi får $\bar{t} = 7436/100 = 74.36$ och $s^2 = \frac{1}{99}\left(553948 - \frac{1}{100}7436^2\right) = 10.1721 \Rightarrow s = 3.189$ etc. Vi får nu att $\mu = \bar{t} \pm 2.6s/\sqrt{n} = 74.4 \pm 0.83$ etc. När man har så många observationer som här är det ingen mening med att hämta kvantilen ur t-tabellen. Det är tillräckligt noggrant att ange det statistiska 99%-felet till 2.6 gånger standardfelet. Hade önskad konfidensgrad varit 95%, hade jag istället multiplicerat standardfelet med 2.

   (c) Oberoendeantagandet behöver verifieras på något sätt, eftersom det typiskt är samma operatör som gör de 100 monteringarna och det kan ju finnas beroenden mellan på varandra följande monteringar. Dessutom skulle jag vilja veta säkert att man inte bytt operatör under försöket, för då kanske inte observationerna är likafördelade.

6. Vi utgår ifrån att $(n-1)s^2/\sigma^2 \sim \chi^2(n-1)$. I $\chi^2(100)$-tabellen hittar vi 0.99-kvantilen 70.0649, vilket värde är aningen för stort, eftersom antalet frihetsgrader är 99 inte 100. Händelsen $70.0649 \leq 99s^2/\sigma^2$ inträffar alltså med den ungefärliga sannolikheten 0.99, och genom att stoppa in $s^2 = 10.172$ och omvandla olikheten får vi att det sökta 99%-intervallet är $\sigma \leq \sqrt{99 \cdot 10.172/70.0649} = 3.79$. Se fotnot 1.

7. $p = 0.501 \pm 2\sqrt{0.501 \cdot 0.499/2000} = 0.501 \pm 0.022$ ("felet" är alltså drygt 2 procentenheter, trots att man tillfrågar så många som 2000 väljare)

8. Av 1906 tillfrågade väljare är det $0.501 \cdot 2000 = 1002$ som väljer något allliansparti. Ett nedåt begr konfidensintervall för proportionen alliansväljare är därför $p \geq 1002/1906 - 1.65\sqrt{\frac{1002}{1906}\left(1 - \frac{1002}{1906}\right)/1906} = 0.526 - 0.019 = 0.507$. Det förefaller alltså som att alliansen hade ett övertag då mätningen gjordes. (Obs 1.65 är 0.95-kvantilen i normalfördelningen.)

Fotnot 1 I ett tidigare förslag till lösning av uppgift 6 användes felaktigt 0.01-kvantilen 135.807. Tack till er som påpekade detta misstag. Till alla: att $\sigma \leq 2.7$ gäller med 99% konfidens, vilket ju var svaret i den felaktiga lösningen, är orimligt eftersom $s = 3.189$. Däremot gäller $\sigma \geq 2.7$ med 99% konfidens, men det är en annan historia.

Resultat:

	U	3	4	5	S:a
antal	0	3	0	0	3
%					