TMS051 Matematisk statistik och simuleringsteknik, Z, del A

Tentamen 22 oktober 2005 em V

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.

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 första våningen i Matematisk centrum, Eklandagatan 86. Genomgång av tentan görs torsdagen den 10 november kl 13-15 i HA3. Granskning av tentan kan under terminstid även göras i mottagningsrummet på entréplanet (våning 2) i Matematisk centrum må-fr 12^{30}–13. Obs att Matematiskt centrum flyttar under juluppehållet.

Svar skall motiveras om ej annat sägs i uppgiften.

Uppgifter

1. Visa för två godtyckliga händelser A och B, att

   (a) $P(B \setminus A) = P(B) - P(A \cap B)$ (2 p)

   (b) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (2 p)

2. Du drar två kort ur en välblandad kortlek. Beräkna sannolikheten att båda har samma valör. (För den som aldrig spelat kort eller glömt: En kortlek består av 52 kort, indelade i 13 valörer med 4 kort i varje.) (3 p)

3. I syfte att ta reda på om det verkligen är någon skillnad i smak på en viss känd läskedryck och en billigare kopia gjordes 10 s.k dubbla blindtest, i vilka en försöksperson ombedes tala om vilken av två drycker han/hon tycker bäst om. (Att testet är ett dubbelt blindtest betyder att inte heller försöksledaren vet vilken av dryckerna som är originalet.) Låt X vara antalet gånger försökspersonen väljer originalet. (a) Vilken sannolikhetsfördelning har X om det inte är någon skillnad på dryckerna? I 8 av de 10 försöken valde försökspersonen originalet. (b) Är ett så pass extremt resultat förvånande om det inte är någon skillnad på dryckerna? Tänk dig in i situationen att du är försöksledare och vill dra någon form av slutsats av experimentets resultat. Svara kortfattat, men tydligt. (2+2 p)

4. För en produktionsanläggning vet man att den går ned med intensiteten ≈ 0.0025 per timma. Beräkna under lämpligt antagande (a) förväntat antalet stopp under en period av 400 timmar, och (b) förväntad tid tills nästa stopp. Glöm inte att tydligt ange vilket antagande du gör. (2+2+1 p)

5. Antag att de två stokastiska variablerna X och Y är kontinuerligt fördelade med täthet

   $$f(x, y) = c \quad \text{för} \quad 0 \le x \le y \le 1$$

   Beräkna $\mathrm{Cov}[X, Y]$. (4 p)

6. Besvara med Ja eller Nej följande tre frågor. Motivera inte dina svar.

   (a) Är medelvärdet en väntevärdesriktig skattning av väntevärdet? (1 p)

   (b) Är stickprovsvariansen en väntevärdesriktig skattning av variansen? (1 p)

   (c) Är stickprovets standardavvikelse en väntevärdesriktig skattning av standardavvikelsen? (1 p)

7. Under fem konsekutiva arbetsdagar lät man fem medarbetare utföra ett visst moment i produktionen av en viss vara. Efter introduktion och en timmas träning under överinseende av ansvarig arbetsledare, mätte man tiden det tog att utföra momentet. Härvid erhölls för de fem medarbetarna:

$$9.13 \quad 6.67 \quad 10.25 \quad 10.58 \quad 7.71$$

   (tidsenhet: minuter) Syftet var att skaffa sig en övre gräns för momentets snittid. Den på företaget ansvarige för statistisk bearbetning konstaterade efter att ha gjort en QQ-plot[1] att det inte är orimligt att i den fortsatta analysen utgå ifrån att tiderna är (i alla fall approximativt) normalfördelade. Låt μ beteckna den förväntade tiden det tar att utföra momentet. Beräkna en övre gräns μ_0, sådan att risken att påståendet $\mu \leq \mu_0$ är fel är ca 5%. (4 p)

8. I syfte att få en uppfattning av kvaliteten hos en viss vara, tillverkad av en underleverantör, valde man i 100 konsekutiva skeppade partier slumpmässigt ut en enhet ur varje. Den utvalda enheten kontrollerades noggrant. Här gör vi det förenklande antagandet att kontrollen var m.a.p en viss funktionalitet och att man helt enkelt avgjorde om enheten höll de med underleverantören överenskomna kraven för denna funktionalitet eller ej. Det visade sig att 90 uppfyllde kraven och att 10 ej gjorde det. Punkt- och intervallskatta sannolikheten att en godtycklig enhet levererad av denna underleverantör uppfyller kraven för varan ifråga. (3 p)

Lycka till!

[1] I del B får du lära dig vad en Q(uantile-)Q(uantile)-plot är och vad man kan använda en sådan till

Lösningar till TMS051, del A den 22/10-05

1. (a) $P(B) = P(A \cap B \cup B \setminus A) = P(A \cap B) + P(B \setminus A)$

   (b) $P(A \cup B) = P(A \cup B \setminus A) = P(A) + P(B \setminus A) = P(A) + P(B) - P(A \cap B)$

   Här ska man tydligt visa att argumenteringen baseras på axiomen. Annars blir det inga poäng.

2. Totala antalet utfall är $\binom{52}{2} = \frac{52 \cdot 51}{2}$. Antalet utfall där båda korten har samma valör är enligt multiplikationsprincipen $13 \cdot \binom{4}{2} \cdot \binom{48}{0} = 13 \cdot \frac{4 \cdot 3}{2} \cdot 1$. Sannolikheten för lika valör (par) blir således enl den klassiska sannolikhetsdefinitionen

$$\frac{13 \cdot \binom{4}{2} \cdot \binom{48}{0}}{\binom{52}{2}} = 13 \cdot \frac{4}{52}\frac{3}{51} = \frac{3}{51} = \frac{1}{17} \approx 0.0589$$

   Man kan också tänka så här: Första kortet spelar ingen roll. Det bestämmer dock i vilken valör man kan få sitt par. Sedan gäller det att andra kortet får samma valör som det första. Denna sannolikhet är naturligtvis 3/51.

3. (a) Man gör 10 oberoende och likafördelade försök. Om det inte är någon skillnad mellan dryckerna blir valet av dryck fullständigt slumpmässig. Sannolikheten att försökspersonen väljer originalet måste då vara 1/2. Härur drar vi slutsatsen att X är binomialfördelad med parametrar $n = 10$ och $p = 1/2$. (b) Vi har att

$$P(X \geq 8) = \sum_{k=8}^{10} \binom{10}{k} (1/2)^k (1/2)^{10-k} = \frac{45 + 10 + 1}{1024} = \frac{56}{1024} \approx 0.055$$

   Att en försöksperson väljer någon av dryckerna minst 8 gånger av 10, sker alltså oftare än 1 gång på 10 (om det inte är någon skillnad). Jag skulle därför inte bli sådär väldigt förvånad. Resultatet 8-av-10 kan dock tyda på att försökspersonen tycker att originalet smakar bättre. Men att statistiskt säkerställa detta kräver fler försök.

4. Här är det rimligt att antaga att stoppen följer en Poissonprocess, vilket innebär (1) att antalet stopp under tiden $[0, t]$ är $\text{Poi}(\lambda t)$ och (2) att tiden till nästa stopp är $\exp(\lambda)$. Här är $\lambda \approx 0.0025 = 1/400$. Ur formelsamling eller Beta får vi nu reda på att svaret på (a) är $\lambda t = 1$ och att svaret på (b) är $1/\lambda = 400$ timmar.

5. Arean av utfallsrummet för paret (X, Y) är 1/2, så $c = 2$. De två marginaltätheterna för X och Y är

$$f(x) = \int_x^1 c\, dy = 2(1 - x) \quad \text{för} \quad 0 \leq x \leq 1$$
$$f(y) = \int_0^y c\, dx = 2y \qquad \text{för} \quad 0 \leq y \leq 1$$

   Vi får nu att

$$E[X] = \int_0^1 2x(1 - x)\, dx = \int_0^1 (2x - 2x^2)\, dx = 1 - \frac{2}{3} = \frac{1}{3}$$
$$E[Y] = \int_0^1 2y^2\, dy = \frac{2}{3}$$
$$E[XY] = \int_0^1 \int_0^y xyc\, dxdy = 2\int_0^1 y \int_0^y x\, dxdy = \int_0^1 y^3\, dy = \frac{1}{4}$$

(Att $E[X] = 1 - E[Y]$ är väl ganska självklart, eller... Titta på marginaltätheterna.) Det följer nu att

$$\mathrm{Cov}[X, Y] = E[XY] - E[X]E[Y] = \frac{1}{4} - \frac{2}{9} = \frac{1}{36}$$

6. (a) Ja, ty $E[\bar{X}] = \mu$
   (b) Ja, ty $E[S^2] = \sigma^2$
   (c) Nej, ty $E[S] \neq \sigma$ utom i ointressanta undantagsfall

7. Vi noterar att $n = 5$, $\sum x = 44.34$, $\sum x^2 = 404.2888$. Härur fås att

$$\bar{x} = \frac{44.34}{5} = 8.868 \quad \text{och} \quad s = \sqrt{\frac{404.2888 - \frac{44.34^2}{5}}{4}} = 1.664$$

Ur t-tabell (4 frihetsgrader) fås kritiska värdet $t_{0.05} = 2.132$ och vi ser till sist att

$$\mu \leq \bar{x} + t_{0.05} s/\sqrt{n} = 8.868 + 1.587 \approx 10.45 \text{ minuter}$$

gäller med konfidensen ca 95%.

8. Punktskattningen av den sökta sannolikheten är $\widehat{p} = 90/100 = 0.90$. Härur fås att $n \min(p, 1-p) \approx 10 > 5$. Vi kan därför använda normalapproximationen av $\mathrm{bin}(n, p)$ då vi beräknar ett konfidensintervall för p. Då inget sagts om konfidensgraden väljer vi att sikta på ≈ 0.95. Motsv kvantil (d.v.s 0.975-kvantilen) i normalfördelningen är ≈ 2, så

$$p = \widehat{p} \pm 2\sqrt{\widehat{p}(1-\widehat{p})/n} = 0.90 \pm 0.06$$

eller

$$p \in [0.84,\ 0.96]$$

gäller med ca 95% konfidens. Med ca 90% konfidens gäller $p = 0.90 \pm 0.05$ och med ca 99% konfidens gäller $p = 0.90 \pm 0.08$.