Chalmers tekniska högskola
Matematisk statistik

TMS051 Matematisk statistik och simuleringsteknik Z, del A

Tentamen 13 januari 2005 f 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 794209.

Jour är

Maximalt antal tentamenspoäng är 30, av dessa krävs 12 för godkänt betyg 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. Granskning av tentan kan under terminstid göras i mottagningsrummet på entréplanet (våning 2) i Matematisk centrum må-fr 12^{30}–13.

Svar skall motiveras om ej annat sägs i uppgiften.

Uppgifter

1. Om tre händelser A, B och C vet man följande:

$$P(A) = 0.43, \quad P(B) = 0.43, \quad P(C) = 0.28, \quad P(A \cap B) = 0.12$$
$$P(A \cap C) = 0.10, \quad P(B \cap C) = 0.15, \quad P(A \cap B \cap C) = 0.07$$

   Hur stor är (a) sannolikheten att ingen av händelserna inträffar, (b) sannolikheten att exakt en av dem inträffar, (c) sannolikheten att exakt två av dem inträffar och (d) sannolikheten att alla inträffar. (4 p)

2. Låt C och D vara händelser i ett försök. Visa att

$$P(C|D) = \frac{P(C)P(D|C)}{P(C)P(D|C) + P(C')P(D|C')} \qquad (4\text{ p})$$

3. I ett visst försök förekommer en intressant händelse A med sannolikheten $P(A) = p$. Man gjorde oberoende upprepningar av detta försök tills dess att A inträffade för första gången. Låt X vara antalet försök som gjordes. Härled X:s sannolikhetsfördelning. (4 p)

4. Låt de stokastiska variablerna X, Y vara sådana att $E[X] = 4$, $E[Y] = 3$, $\mathrm{Var}[X] = 2$, $\mathrm{Var}[Y] = 1$ och $\mathrm{Cov}[X, Y] = 1$. Låt $U = 3Y - 2X$. Beräkna U:s väntevärde μ och varians σ^2. (4 p)

5. Antag att du i ett försök observerar en stokastisk variabel X med väntevärde $\mu = 2$ och standardavvikelse $\sigma = 3$. Antag vidare att du gör 5 oberoende upprepningar av försöket. Låt de oberoende stokastiska variablerna $X_1, \ldots, X_5$ representera de erhållna resultaten och bilda medelvärdet $\bar{X} = (X_1 + \ldots + X_5)/5$. Beräkna $E[\bar{X}]$ och $\mathrm{Var}[\bar{X}]$. (3 p)

6. I ett försök att evaluera ett mätinstruments noggranhet gjordes 5 oberoende mätningar på samma objekt - en s.k standard. Därvid erhölls följande avvikelser från det sanna mätvärdet:

$$-0.019, \quad -0.120, \quad -0.324, \quad 0.097, \quad 0.021$$

Anta att dessa avvikelser är $N(\mu, \sigma)$-fördelade.

   (a) Punkt- och intervallskatta mätinstrumentets standardavvikelse σ. Önskad konfidensgrad: 0.99. (3 p)

   (b) Är din punktskattning väntevärdesriktig? Svara bara Ja eller Nej. (1 p)

7. Använd data från ovanstående uppgift till att

   (a) punkt- och intervallskatta väntevärdet μ som är ett mått på hur mycket mätresultatet avviker från det sanna värdet.. Önskad konfidensgrad: 0.90. (3 p)

   (b) Går det att utifrån dessa 5 mätningar hävda att $\mu \neq 0$? Vilken felrisk har man i så fall? (1 p)

8. Inför ett allmänt val har man frågat 1000 slumpmässigt utvalda ”röstare”. Det visade sig att 375 skulle rösta med partiet X. Punkt- och intervallskatta proportionen väljare som tänker rösta på detta parti. Önskad konfidensgrad: 0.95. (3 p)

Lösningar till TMS051, del A den 13/1-05

1. Beräkna successivt (rita ett Venn-diagram innehållande tre händelser, så förstår du beräkningarna)

$$\begin{aligned}
P(A \cap B \cap C') &= 0.12 - 0.07 = 0.05 \\
P(A \cap B' \cap C) &= 0.10 - 0.07 = 0.03 \\
P(A' \cap B \cap C) &= 0.15 - 0.07 = 0.08 \\
P(A \cap B' \cap C') &= 0.43 - (0.07 + 0.05 + 0.03) = 0.28 \\
P(A' \cap B \cap C') &= 0.43 - (0.07 + 0.05 - 0.08) = 0.23 \\
P(A' \cap B' \cap C) &= 0.28 - (0.07 + 0.03 + 0.08) = 0.10 \\
P(A' \cap B' \cap C') &= 1 - (0.07 + 0.05 + 0.03 + 0.08 + 0.28 + 0.23 + 0.10) = 0.16
\end{aligned}$$

Vi kan nu beräkna eller ange de sökta sannolikheterna. Låt $p(k)$ beteckna sannolikheten att exakt k av händelserna inträffar. (a) $p(0) = 0.16$, (b) $p(1) = 0.28+0.23+0.10 = 0.61$, (c) $p(2) = 0.05 + 0.03 + 0.08 = 0.16$ och (d) $p(3) = 0.07$.

2. Notera först att

$$P(C|D) = \frac{P(C \cap D)}{P(D)}$$

Medelst partitionering av D i delarna $D \cap C$ och $D \cap C'$ inses att

$$P(D) = P(D \cap C) + P(D \cap C') = P(C)P(D|C) + P(C')P(D|C')$$

Dessutom gäller ju att $P(C \cap D) = P(C)P(D|C)$. Nu är det sökta resultatet, som f.ö kallas Bayes formel, omedelbart.

3. $P(X = k) = (1-p)^{k-1}p$ för $k = 1, 2, \ldots$, ty händelsen $X = k$ inträffar precis då man gör k försök varav A' inträffar i den $k-1$ första medan A inträffar i det k:te försöket. Detta räcker som härledning. Vi ser att fördelningen för X är av den geometriska typen.

4. $\mu = E[3Y - 2X] = 3E[Y] - 2E[X] = 1$ och $\sigma^2 = \mathrm{Var}[3Y - 2X] = \mathrm{Var}[3Y] + \mathrm{Var}[2X] + 2\mathrm{Cov}[3X, -2Y] = 9\,\mathrm{Var}[Y] + 4\,\mathrm{Var}[X] - 12\,\mathrm{Cov}[X, Y] = 9 + 8 - 12 = 5$

5. $E[\bar{X}] = \mu = 2$ och $\mathrm{Var}[\bar{X}] = \sigma^2/n = 3^2/5 = 1.8$

6. Det är praktiskt att först räkna ut $\sum x = -0.345$ och $\sum x^2 = 0.129587$ samt notera att $n = 5$, och därefter $\bar{x} = \sum x/n = -0.069$ och $s^2 = \left(\sum x^2 - n\bar{x}^2\right)/(n-1) = 0.02645$ samt $s = 0.1626$. Ur tabell fås kvantilerna $\chi^2_{0.995,4} = 0.206989$ och $\chi^2_{0.005,4} = 14.8603$. Ur faktumet $(n-1)s^2/\sigma^2 \sim \chi^2(n-1)$ fås nu konfidensintervallet $\sqrt{4 \cdot 0.02645/14.8603} = \sqrt{0.007120} = 0.0844 \leq \sigma \leq 0.7149 = \sqrt{0.5111} = \sqrt{4 \cdot 0.02645/0.206989}$. Svaret på delfråga (a) blir följaktligen $\hat{\sigma} = 0.163$ och $\sigma \in (0.0844, 0.7149)$ med konfidensen 99%. På delfråga (b) är svaret nej (s^2 skattar ju σ^2 väntevärdesriktigt och det vore orimligt om då s vore en väntevärdesriktig skattning av σ).

7. Ur t-tabell fås $t_{0.05,4} = 2.13185$. ”Felet” i punktskattningen $\hat{\mu} = \bar{x} = -0.069$ (j.f.r ovanstående uppgift) av μ blir således $\pm t_{0.05,4} s/\sqrt{n} = 2.13185 \cdot 0.1626/\sqrt{5} = \pm 0.155$. Således gäller (a) med konfidensen 90% att $\mu = -0.069 \pm 0.155$. (b) Då $\mu = 0$ tillhör konfidensintervallet kan vi inte på basis av det hävda att $\mu \neq 0$. M.a.o, vill vi att den maximala risken att ha fel ska vara högst 10% (eller lägre) kan vi inte hävda att $\mu \neq 0$.

8. Uttrycket $p = \hat{p} \pm 1.96\sqrt{\hat{p}(1-\hat{p})/n}$, där $\hat{p} = f/n$ (f är antalet som säger sig rösta med partiet och n är totala antalet tillfrågade), härleds ur faktumet $(\hat{p} - p)/\sqrt{p(1-p)/n} \sim \mathrm{N}(0, 1)$ (se formelsamlingen). Vi får $p = 0.375 \pm 0.030$.