Tentamen 27 augusti 2003
8^{45}-13^{45} i V

NTM010 Statistik för fysiker

Anvisningar

Hjälpmedel: räknedosa utan lagrad information om kursen, Beta samt kursens formel- och tabellsamling.

Betygsgränser: 12 p för G och 22 p för VG av totalt 30 p.

Lösningar publiceras på webben första vardagen efter tentamensdagen.

Examinator och jour: Tommy Norberg, tel 772 3528, 0730 79 42 09.

Problem

1. I en viss population förekommer en viss sjukdom med frekvensen 1/20. Då man testar om en godtyckligt vald individ har denna sjukdom gäller att testet visar positivt resultat (d.v.s indikerar att individen är sjuk) med sannolikheten 0.9 om individen är sjuk och med sannolikheten 0.2 annars. Antag att testet visat positivt resultat för en godtyckligt vald individ. Vad är sannolikheten att denna individ verkligen är sjuk. (4 p)

2. Låt $F_X(x)$ vara fördelningsfunktion för den stokastiska variabeln X. Visa
$$F_X(b) - F_X(a) = P(a < X \leq b)$$
Här förutsätts att $a \leq b$. (3 p)

3. Låt X vara normalfördelad med väntevärde μ och standardavvikelse σ. Beräkna den 95:e percentilen $x_{0.95}$ (kallas också 95%-kvantilen). (3 p)

4. Den två-dimensionella massfunktionen $p(x, y)$ uppfyller

		y			
		0	1	2	3
	0	.08	.07	.04	.00
	1	.06	.15	.05	.04
x	2	.05		.10	.06
	3	.00	.01	.05	.06
	4	.00	.01	.05	.06

Kalla motsv stokastiska variabler för X och Y.

   (a) Beräkna den saknade sannolikheten, d.v.s $P(X = 2, Y = 1)$. (1 p)

   (b) Beräkna $P(X + Y = 2)$. (1 p)

   (c) Beräkna $P(Y = 2)$. (1 p)

   (d) Beräkna $P(Y - X = 1)$. (1 p)

5. Låt X, Y vara stokastiska variabler med väntevärden $E[X]$ och $E[Y]$. Visa att $E[X + Y] = E[X] + E[Y]$. (4 p)

(Vänd)

6. Pontus påstår att med den splitter nya eneuron han håller i sin hand säkert får minst 60 klave (eller det som motsvarar klave på ett eneuromynt) på 100 kast. Hur stor är hans chans att lyckas. (En approximativ överslagsberäkning duger bra.) (3 p)

7. Visa att stickprovsvariansen S^2 är en väntevärdesriktig skattning av variansen σ^2. Här förutsätts att observationerna är oberoende, n till antalet och att deras fördelning har ändligt andramoment. (4 p)

8. I ett laboratorium gjordes fem väntevärdesriktiga mätningar av en viss storhet μ. Man erhöll

$$11.67 \quad 12.98 \quad 6.99 \quad 9.95 \quad 9.61$$

   Antag normalfördelade fel.

   (a) Punkt- och intervallskatta μ. Låt konfidensgraden vara 95%. (4 p)

   (b) En viktig förutsättning för (a) har ej nämnts. Vilken? (1 p)

— Lycka till! —

Tentamen 20 maj 2003
8^{45}-13^{45} i VV

NTM010 Statistik för fysiker

Anvisningar

Hjälpmedel: räknedosa utan lagrad information om kursen, Beta samt kursens formel- och tabellsamling.

Examinator och jour: Tommy Norberg, tel 772 3528, 0730 79 42 09.

Betygsgränser: 12 p för G och 22 p för VG av totalt 30 p.

Lösningar publiceras på webben den 20 maj. Se kursens hemsida. Tentamensresultat publiceras på anslagstavla i "suckarnas gång" (källaren i Matematiskt centrum) preliminärt onsdag den 28 maj. Nästa tentamensmöjlighet är onsdag 27 augusti, 2003.

Problem

1. Betrakta följande försök: Drag utan återläggning ett kort ur en välblandad kortlek; drag sedan ytterligare ett kort. Är händelserna det första kortet är hjärter och det andra kortet är hjärter oberoende? Antag att ett av de två dragna korten har färgen hjärter. Beräkna den betingade sannolikheten att det återstående kortet är hjärter. (4 p)

2. Härled följande variant av Bayes formel:
$$P(A|B) = \frac{P(A)P(B|A)}{P(A)P(B|A) + P(A^c)P(B|A^c)} \qquad (4\text{ p})$$

3. Låt $f(x) = k(\alpha)x^\alpha$ för $0 < x < 1$. För vilka α är f en täthetsfunktion för en stokastisk variabel och bestäm $k(\alpha)$ för varje sådant α. Beräkna även väntevärde och varians för en stokastisk variabel med denna täthetsfunktion. (4 p)

4. Vid utvinning av olja i Nordsjön är det inte ovanligt att pump-plattformen hålls på plats av ett satellitbaserat positioneringssystem. Då vinden är stark riskerar plattformen att driva för långt ifrån riktpunkten. Man uppskattar att detta sker i genomsnitt ca 8 gånger per år och när det sker måste pumputrustningen kopplas ifrån, ty ifall plattformen driver tillräckligt fel slits pumputrustningen sönder och en s.k. "blow out" inträffar. Man räknar med att frånkopplingen lyckas i ca 19 fall av 20 och att plattformen driver iväg till blow out-läge i ca 1 fall av 40. Beräkna sannolikheten för ingen blow out under en tidsperiod av 1 år. (4 p)

5. Låt X vara exponentialfördelad med väntevärde $1/\lambda$. Beräkna kvantilen x_α. (3 p)

6. (a) Visa följande generella formel för beräkning av variansen för en summa av två stokastiska variabler X och Y:
$$\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] + 2\mathrm{Cov}[X,Y] \qquad (2\text{ p})$$

   (b) Antag att paret X, Y har en två-dimensionell normalfördelning och
$$\mu_X = 5, \quad \mu_Y = 3, \quad \sigma_X = 1, \quad \sigma_Y = 2$$

   Om korrelationen är 0.5, vad blir då $\mathrm{Var}[X+Y]$? (1 p)

(Vänd)

7. Låt $X_1, X_2, \ldots, X_n$ vara oberoende och ha väntevärdet μ och standardavvikelsen σ.
   (a) Beräkna väntevärde och standardavvikelse för medelvärdet. (2 p)
   (b) Beräkna väntevärde av variansen. (2 p)

8. En metallurg gjorde fyra bestämningar av mangans smältpunkt, varvid han erhöll följande smältpunkter i grader Celcius:

   1269 1271 1263 1265

   (a) Punktskatta mangans smältpunkt. (1 p)
   (b) Bestäm ett 99% konfidensintervall för smältpunkten. (2 p)
   Ange vilka förutsättningar du behöver göra för att lösa deluppgifterna ovan. (1 p)

Lycka till!

Svar eller lösningar till tentamen i NTM010 Statistik för fysiker, 27/8-2003

1. Låt S betyda att individen har sjukdomen och D att den detekteras. Följande är givet: $P(S) = 1/20 = 0.05$, $P(D|S) = 0.9$ och $P(D|S^c) = 0.2$. Bayes formel ger

$$P(S|D) = \frac{P(S)P(D|S)}{P(S)P(D|S) + P(S^c)P(D|S^c)} = \frac{0.05 \cdot 0.9}{0.05 \cdot 0.9 + 0.95 \cdot 0.2} \approx 0.191$$

2. Se Blom sida 52.

3. Ur definitionen $P(X \leq x_{0.95}) = 0.95$ fås

$$P\left(\frac{X-\mu}{\sigma} \leq \frac{x_{0.95}-\mu}{\sigma}\right) = 0.95 \quad \Rightarrow \quad \Phi\left(\frac{x_{0.95}-\mu}{\sigma}\right) = 0.95$$

Ur lämplig tabell (t.ex sida x i kursens formel- och tabellsamling) fås nu

$$\frac{x_{0.95}-\mu}{\sigma} = 1.645 \quad \Rightarrow \quad x_{0.95} = \mu + 1.645\sigma$$

4. (a) $P(X = 2, Y = 1) = 0.06$, ty annars är ej $\sum_{x,y} p(x,y) = 1.00$.
   (b) $P(X + Y = 2) = 0.05 + 0.15 + 0.04 = 0.24$
   (c) $P(Y = 2) = 0.04 + 0.05 + 0.10 + 0.05 + 0.05 = 0.29$.
   (d) $P(Y - X = 1) = 0.07 + 0.05 + 0.06 = 0.18$

5. Se Blom sida 120.

6. Vi utgår ifrån att myntet är symmetriskt så att sannolikheten för klave i ett godtyckligt kast blir 0.5. Vi utgår vidare ifrån att de 100 kasten är oberoende. Då blir frekvensen klave, f säg, enl cgs approximativt normalfördelad med väntevärde $100 \cdot 0.5 = 50$ och standardavvikelse $\sqrt{100 \cdot 0.5 \cdot 0.5} = 5$. Således gäller att

$$P(f \geq 60) = P\left(Z \geq \frac{60-50}{5}\right) \approx 1 - \Phi(2) = 0.025$$

där $Z = (f - 50)/5$ alltså är approximativt $\mathrm{N}(0,1)$-fördelad.

7. Notera först att

$$(n-1)S^2 = \sum_i (X_i - \bar{X})^2 = \sum_i X_i^2 - n\bar{X}^2$$

(Den som inte kommer ihåg detta får naturligtvis härleda uttrycket.) Vi får nu att

$$E[(n-1)S^2] = \sum_i E[X_i^2] - nE[\bar{X}^2] = nE[X_1^2] - nE[\bar{X}^2]$$

Notera att

$$E[X_1^2] - E[X_1]^2 = \sigma^2 \quad \Rightarrow \quad E[X_1^2] = \sigma^2 + \mu^2$$

där $\mu = E[X_1]$. Analogt

$$E[\bar{X}^2] = \mathrm{Var}[\bar{X}] + E[\bar{X}]^2 = \sigma^2/n + \mu^2$$

Vi får följaktligen att

$$E[(n-1)S^2] = n(\sigma^2 + \mu^2) - n(\sigma^2/n + \mu^2) = (n-1)\sigma^2$$

vilket visar att S^2 skattar σ^2 väntevärdesriktigt.

8. (a) Börja med att räkna ut att $\bar{x} = 10.24$ och att $s = 2.269$. Ur tabell fås $t_{0.025}(4) = 2.776$. Felet med konfidens 95% är

$$2.776 \cdot 2.269/\sqrt{5} = 2.817 \approx 2.82$$

och $\mu = 10.24 \pm 2.82$ är den sökta punkt- och intervallskattningen. Ett alternativ sätt att ge intervallskattningen är $\mu \in (7.42,\ 13.06)$. Punktskattningen av μ är $\widehat{\mu} = 10.24$.

(b) Det är oerhört viktigt att observationerna är oberoende.

Svar eller lösningar till tentamen i NTM010 Statistik för fysiker, 20 maj 2003

1. Låt A_1 vara händelsen att det först dragna kortet är hjärter och A_2 händelsen att det sist dragna kortet är hjärter. Vi får att $P(A_1) = 13/52 = 1/4$, $P(A_1 \cap A_2) = P(A_1)P(A_2|A_1) = (13/52)(12/51) = 3/51 = 1/17$ samt $P(A_1^c \cap A_2) = (39/52)(13/51) = (3/4)(13/51) = 13/68$. Således är $P(A_2) = (3/51) + (3/4)(13/51) = 1/4$. Vi ser att $P(A_1)P(A_2) = (1/4)(1/4) = 1/16$, medan $P(A_1 \cap A_2) = 3/51 \neq 1/16$. Händelserna A_1 och A_2 är alltså ej oberoende.

   Notera även att $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2) = (1/4) + (1/4) - (3/51) = 45/102 = 15/34$ och $P(A_1 \cap A_2 | A_1 \cup A_2) = (1/17)/(15/34) = 2/15$ följer.

2. $P(A|B) = P(A \cap B)/P(B)$ Nu är $P(A \cap B) = P(A)P(B|A)$ och $P(B) = P(A \cap B) + P(A^c \cap B) = P(A)P(B|A) + P(A^c)P(B|A^c)$. Vi får $P(A|B) = P(A)P(B|A)/\left(P(A)P(B|A) + P(A^c)P(B|A^c)\right)$, vilket skulle visas.

3. Ur $1 = \int_0^1 f(x)\,dx = \int_0^1 k(\alpha)x^\alpha\,dx = k(\alpha)\int_0^1 x^\alpha\,dx = k(\alpha)/(\alpha+1)$ för $\alpha > -1$ fås $k(\alpha) = \alpha + 1$. Integralen är divergent för $\alpha \leq -1$. Väntevärdet $\mu = \int_0^1 x\,(\alpha+1)x^\alpha\,dx = (\alpha+1)/(\alpha+2)$ och för variansen behöver vi $\int_0^1 x^2\,(\alpha+1)x^\alpha\,dx = (\alpha+1)/(\alpha+3)$. Vi får nu att variansen är $\sigma^2 = (\alpha+1)/(\alpha+3) - \left((\alpha+1)/(\alpha+2)\right)^2 = (\alpha+1)/\left((\alpha+3)(\alpha+2)^2\right)$.

4. Antalet blow-out under en tidsperiod av ett år är Poissonfördelat med väntevärde ca $8 \cdot (1/20) \cdot (1/40) = 1/100$. Således är den sökta sannolikheten $\approx e^{-1/100} \approx 1 - 1/100 = 0.99$.

5. Kvantilen x_α definieras av att $P(X > x_\alpha) = \alpha$. För exponentialfördelningen är $P(X > x) = \int_x^\infty \lambda e^{-\lambda y}\,dy = e^{-\lambda x}$. Således är $-\lambda x_\alpha = \ln\alpha$ ur vilket $x_\alpha = (-\ln\alpha)/\lambda$ följer. Rätt får också de som ur definitionen $P(X \leq x_\alpha) = \alpha$ härlett $x_\alpha = (-\ln(1-\alpha))/\lambda$.

6. (a) $\mathrm{Var}[X+Y] = \mathrm{Cov}[X+Y, X+Y] = \mathrm{Cov}[X, X+Y] + \mathrm{Cov}[Y, X+Y] = \mathrm{Cov}[X,X] + \mathrm{Cov}[X,Y] + \mathrm{Cov}[Y,X] + \mathrm{Cov}[Y,Y] = \mathrm{Var}[X] + \mathrm{Var}[Y] + 2\mathrm{Cov}[X,Y]$

   (b) Ur $\rho = \mathrm{Cov}[X,Y]/\sqrt{\mathrm{Var}[X]\,\mathrm{Var}[Y]}$ fås $\mathrm{Cov}[X,Y] = \sqrt{1^2 \cdot 2^2} \cdot (1/2) = 1$ och $\mathrm{Var}[X+Y] = 1^2 + 2^2 + 2 \cdot 1 = 7$ följer.

7. Medelvärdet är $\bar{X} = (1/n)\sum_{i=1}^n X_i$ och variansen är $S^2 = \sum_{i=1}^n (X_i - \bar{X})^2/(n-1)$. Vi får

   (a) $E[\bar{X}] = (1/n)\sum_{i=1}^n E[X_i] = (1/n)n\mu = \mu$ och $\mathrm{Var}[\bar{X}] = (1/n)^2\sum_{i=1}^n \mathrm{Var}[X_i] = (1/n)^2 n\sigma^2 = \sigma^2/n$. Medelvärdets standardavvikelse är således $\sigma/\sqrt{n}$.

   (b) Ur $(n-1)S^2 = \sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n (X_i - \mu + \mu - \bar{X})^2 = \sum_{i=1}^n (X_i - \mu)^2 + 2\sum_{i=1}^n (X_i - \mu)(\mu - \bar{X}) + \sum_{i=1}^n (\mu - \bar{X})^2 = \sum_{i=1}^n (X_i - \mu)^2 + 2n(\bar{X} - \mu)(\mu - \bar{X}) + n(\mu - \bar{X})^2 = \sum_{i=1}^n (X_i - \mu)^2 - n(\bar{X} - \mu)^2$ fås att $(n-1)E[S^2] = n\sigma^2 - n\sigma^2/n = (n-1)\sigma^2$ och $E[S^2] = \sigma^2$ följer.

8. Vi har i $n = 4$ mätningar erhållit medelvärdet $\bar{x} = 1267$ och standardavvikelsen $s = 3.651$.

   (a) Jag punktskattar väntevärdet μ med medelvärdet $\bar{x} = 1267$ och om mätmetoden är väntevärdesriktig är detta även en punktskattning av mangans smältpunkt i °C..

   (b) Vi förutsätter även att mätningarna är oberoende och har normalfördelade fel. Om den önskade konfidensgraden är 99%, så är $\alpha = 0.01$. Ur tabell fås att $t_{n-1,\alpha/2} = t_{3,0.005} = 5.841$, vilket ger "felet" $t_{n-1,\alpha/2}s/\sqrt{n} = 5.841 \cdot 3.651/\sqrt{4} = 10.66$. Således är [1256.3, 1277.7] ett 99% konfidensintervall för smältpunkten.

   Resultatet från (a) och (b) sammanfattas i $\mu = 1267 \pm 10.7$ med konfidensen 99%. Förutsättningarna för beräkningarna är som redan nämnts att mätmetoden är väntevärdesriktig, att mätningarna är oberoende och att felen är normalfördelade.