Tekniska högskolan i Linköping
Matematiska institutionen
Matematisk statistik,Jan Olheim

MATEMATIK:Statistik

92MA31 STN2 , 92MA37 STN2

TENTAMEN MÅNDAGEN DEN 22 OKTOBER 2012 KL 14.00-18.00.

Hjälpmedel:Formler och tabeller för 92MA31/92MA37 samt räknedosa med tömda minnen.
Varje korrekt löst uppgift ger 3 poäng. Tredelade uppgifter ger 1 poäng per del.
För betyget godkänd krävs minst 8 poäng.För väl godkänd minst 14.
Jourhavande lärare: Jan Olheim,tel 28 14 53.

1. A , B och C är tre oberoende händelser.
P (A) = 0.6 , P (B) = 0.2 , P (C) = 0.3

a) Bestäm P(ingen av händelserna inträffar).

b) Bestäm P(exakt två av händelserna inträffar).

c) Givet : Minst en av A och B inträffar.
Sök: Den betingade sannolikheten att endast en av A och B inträffar.

2. Ett undersökningsföretag jobbar med snabba tittarfrågor,
dvs efter ett TV-program ringer man upp ett slumpmässigt urval
av Sveriges befolkning och ställer frågor.
Man är i första hand ute efter att få en skattning av andelen
personer (p) som sett programmet.
Antag att 28 personer av 500 såg programmet "Ruskaby skola"
Bilda ett approximativt 95 % konfidensintervall för p. (2.5p)

Följdfråga: Om det finns 5.000.000 tänkbara tittare,hur ser motsvarande
intervall för antalet tittare ut? (0.5 p)

3. Antalet kunder som under en timme anländer till en urmakare kan ses som en Poissonfördelad slumpvariabel med väntevärde 3.

a) Beräkna sannolikheten att minst 5 kunder anländer mellan kl 9 och 10.

b) Beräkna approximativt sannolikheten att högst 20 kunder kommer under ett 8-timmars arbetspass.

c) Endast 80% av kundbesöken leder till ett uppdrag för urmakaren. Hur stor är sannolikheten att 10 kundbesök leder till högst 7 uppdrag?

4. Barn drabbas ibland av blyförgiftning genom att de av någon anledning äter blyhaltiga ämnen. En tänkbar förklaring är att de lider av kalkbrist.
I en studie användes 20 råttor som slumpmässigt delades in i två grupper. Kontrollgruppen fick normal diet medan experimentgruppen fick kalkfattig diet. Råttorna hade tillgång till en lösning med blyacetat. Mängden blyacetat som konsumerades mättes för varje råtta.

Kont.	5.4	6.2	3.1	3.8	6.5	5.8	6.4	4.5	4.9	4.0
Expr.	8.8	9.5	10.6	9.6	7.5	6.9	7.4	6.5	10.5	8.3

$\bar{x} = 5.06 \quad s_1 = 1.189 \quad \bar{y} = 8.51 \quad s_2 = 1.471$

Vi har observationer $x_1, x_2.....x_{10}$ och $y_1, y_2.....y_{10}$
av oberoende slumpvariabler X som är$N(\mu_1, \sigma)$ resp.Y som är $N(\mu_2, \sigma)$

Bilda ett 95% konfidensintervall för $\mu_2 - \mu_1$
Kan man ur intervallet dra slutsatsen att $\mu_2 > \mu_1 + 2$?
(Dvs konsumerar råttor som får kalkfattig kost avsevärt mer bly än de som får normal kost?)

5. Låt den kontinuerliga tvådimensionella slumpvariabeln (X , Y) ha täthetsfunktionen

$$f(x,y) = x + y \qquad 0 \leq x \leq 1 \text{ , } 0 \leq y \leq 1$$

a) Bestäm de marginella täthetsfunktionerna $f_X(x)$ och $f_Y(y)$

b) Bestäm väntevärdena E (X) och E (Y)

c) Beräkna kovariansen mellan X och Y , C(X,Y)

6. En rak stav av längden 1 delas slumpmässigt(likformig fördelning) i två delar, med längder X resp. 1-X.
Låt längderna utgöra kateter i en rätvinklig triangel.
Bestäm väntevärde och varians för triangelns area.

1 92MA31/37-Lösningsförslag 121022

1.

P (A) = 0.6 P (B) = 0.2 P (C) = 0.3

a) P (ingen) = P ($A^* \cap B^* \cap C^*$) = (oberoende) =
P ($A^*) \cdot P(B^*) \cdot P(C^*)$ = 0.4 · 0.8 ·0.7 = 0.224

b) P (exakt två) = P (A) · P (B) · P (C^*) +
P (A) · P (B^*) · P (C) + P (A^*) · P (B) · P (C) =
= 0.6 · 0.2 · 0.7 + 0.6 · 0.8 · 0.3 + 0.4 · 0.2 · 0.3 = 0.252

c) Minst en av A och B = A ∪ B
Exakt en av A och B = (A $\cap B^*) \cup (A^* \cap B)$

$$P(((A \cap B^*) \cup (A^* \cap B)) \mid (A \cup B)) = \frac{P((A \cap B^*) \cup (A^* \cap B))}{P(A \cup B)} =$$

$$\frac{P(A) \cdot P(B^*) + P(A^*) \cdot P(B)}{P(A) + P(B) - P(A) \cdot P(B)} = \frac{0.6 \cdot 0.8 + 0.4 \cdot 0.2}{0.6 + 0.2 - 0.12} = \frac{0.56}{0.68} \approx 0.82$$

2.

a) X = antalet personer som sett programmet
X är Bin (500 , p) ≈ N (500p , $\sqrt{500p(1-p)}$
p skattas med $\hat{p} = \frac{X}{500}$, observerat värde 0.056

$\frac{\hat{p}-p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{500}}}$ är approximativt N (0 , 1)

Ett approximativt 95 % konfidensintervall för p fås ur:

P (-1.96 ≤ $\frac{\hat{p}-p}{\sqrt{\frac{\hat{p}(1-\hat{p})}{500}}}$ ≤ 1.96) ≈ 0.95 1.96 fås ur N(0,1)-tabell

Lös ut p ,vilket ger: p = $\hat{p} \pm 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{500}}$

Ett observerat intervall blir: 0.056 ± 0.020 eller (0.036,0.076)

b) Multiplicera gränserna med 5.000.000.
Dvs av 5 miljoner tänkbara tittare såg mellan 179000 och 380000 programmet.
(Namnet Ruskaby skola är hämtad från Realskolans Läsebok,1933).

3.

a) X = antalet kunder mellan 9.00 och 10.00
X är Po (3)
P (X $\geq$ 5) = 1 - P (X$\leq$ 4) = (tabell) = 1- 0.8153 = 0.1847

b) Y = antalet kunder på 8 timmar
Y är Po (8 $\cdot$3) = Po (24) $\approx$ N (24 , $\sqrt{24}$)

P (X $\leq$ 20) = P ($\frac{X-24}{\sqrt{24}} \leq \frac{20-24}{\sqrt{24}}) \approx \Phi(\frac{-4}{\sqrt{24}}) = \Phi(-0.8165) =$

$1 - \Phi(0.8165) \approx 1 - 0.79 = 0.21$

c) Z = antalet uppdrag Z är Bin (10 , 0.8)
P (Z $\leq$ 7) = 0.3222 (tabell)

4.

$\frac{\bar{Y}-\bar{X}-(\mu_2-\mu_1)}{s\cdot\sqrt{\frac{1}{10}+\frac{1}{10}}}$ är t(10+10-2) dvs t(18)

där $s^2 = \frac{(10-1)\cdot s_1^2+(10-1)\cdot s_2^2}{10+10-2} = \frac{9\cdot(1.189)^2+9\cdot(1.471)^2}{18} = 1.7888$; s= 1.3375

95 % konfidensgrad ger värdet 2.101 ur t(18)-tabell
Ett konfidensintervall för $\mu_2 - \mu_1$ blir då:

$8.51 - 5.05 \pm 2.101 \cdot 1.3375 \cdot \sqrt{\frac{1}{10} + \frac{1}{10}}$

3.45 ± 1.26 d.v.s. (2.19,4.71)
Intervallet ligger över 2.0 , dvs $\mu_2 > \mu_1 + 2$

5.

a) $f_X(x) = \int f(x,y)\,dy = \int_{y=0}^{1} (x+y)\,dy = [xy+\frac{y^2}{2}]_{y=0}^{1} = x+\frac{1}{2}$; $0 \leq x \leq 1$

Symmetri i x och y ger $f_Y(y) = y + \frac{1}{2} \quad 0 \leq y \leq 1$

b) $E[X] = \int x \cdot f_X(x)\,dx = \int_{x=0}^{1} x \cdot (x + \frac{1}{2})\,dx = [\frac{x^3}{3} + \frac{x^2}{4}]_{x=0}^{1} = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}$

$E[Y] = \frac{7}{12}$ (symmetri)

c) C (X , Y) = E [X Y] - E [X] E [Y]

$E[X \cdot Y] = \int \int x \cdot y \cdot f(x,y)\,dxdy = \int_{x=0}^{1} \int_{y=0}^{1} x \cdot y \cdot (x+y)\,dxdy$

$= \int_{y=0}^{1} [\frac{x^3}{3} \cdot y + \frac{x^2}{2} \cdot y^2]_{x=0}^{1} = \int_{y=0}^{1} (\frac{y}{3} + \frac{y^2}{2})\,dy = [\frac{y^2}{6} + \frac{y^3}{6}]_{y=0}^{1} = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$

$C(X,Y) = \frac{1}{3} - (\frac{7}{12})^2 = \frac{1}{3} - \frac{49}{144} = -\frac{1}{144} \approx -0.0069$

(Anmärkning: V(X) = V(Y) kan beräknas till 11/144 vilket ger

korrelationen $\rho = \frac{\frac{-1}{144}}{\sqrt{\frac{11}{144} \cdot \frac{11}{144}}} = \frac{-1}{11}$)

6.

X är likformig på (0,1) dvs f(x) = 1 i intervallet (0,1)

Arean A = $\frac{X \cdot (1-X)}{2} = \frac{1}{2} \cdot (X - X^2)$

$E[A] = \frac{1}{2} E[X - X^2] = \frac{1}{2}(E[X] - E[X^2])$

$V[A] = E[A^2] - (E[A])^2$

$E[A^2] = E[(\frac{1}{2}(X - X^2))^2] = \frac{1}{4} \cdot (E[X^2] + E[X^4] - 2E[X^3])$

Vi behöver alltså $E[X], E[X^2], E[X^3], E[X^4]$

$E[X) = \int_0^1 x\,dx = [\frac{x^2}{2}]_{x=0}^{1} = \frac{1}{2} \quad E[X^2] = \int_0^1 x^2\,dx = [\frac{x^3}{3}]_{x=0}^{1} = \frac{1}{3}$

På samma sätt fås $E[X^3] = \frac{1}{4}$ $E[X^4] = \frac{1}{5}$

$E[A] = \frac{1}{2}(\frac{1}{2} - \frac{1}{3}) = \frac{1}{12} \quad E[A^2] = \frac{1}{4}(\frac{1}{3} + \frac{1}{5} - 2 \cdot \frac{1}{4}) = \frac{1}{120}$

$V[A] = \frac{1}{120} - (\frac{1}{12})^2 = \frac{1}{120} - \frac{1}{144} = \frac{1}{720}$