Tentamenskrivning för TMS063, Matematisk Statistik.
Tisdag fm den 14 april, 2015, V-huset.
Examinator: Marina Axelson-Fisk. Tel: 070-2288113
Tillåtna hjälpmedel: typgodkänd miniräknare, tabell- och formelhäfte (utdelas).
****

Tentan är på totalt 50 poäng: 40 poäng på matstat-delen, 10 poäng på flervariabel-delen. För att bli godkänd krävs godkänt på båda delarna. Betygsgränser för betyg 3, 4 och 5 är 40% (20p), 60% (30p) och 80% (40p), respektive. Givet att båda delarna är godkända, ges betyget av den totala poängsumman.
Fullständiga och välmotiverad lösningar skall ges till varje uppgift.

Del I: Matematisk Statistik

1. Antag att $P(A) = 0.01$, $P(B|A) = 0.40$ och $P(B) = 0.30$.
   (a) Vad är $P(A \cap B)$? (2p)
   (b) Vad är $P(A|B)$? (2p)
   (c) Vad är $P(A \cup B)$? (2p)

2. En villaägare köper en stor påse blandade lökar. Enligt förpackningen är en fjärdedel av lökarna påskliljor, en fjärdedel krokusar och resten tulpaner. Alla påskliljor och hälften av tulpanerna är gula, resten är röda. Bland krokusarna är en tredjedel lila och övriga gula.
   (a) Villaägaren grävde ner en slumpmässigt vald lök utanför köksföstret. Vad är sannolikheten att blomman är gul? (3p)
   (b) Om blomman är gul, vad är sannolikheten att det är en påsklilja? (3p)
   (c) Villaägaren grävde också ner 10 lökar vid ytterdörren. Vad är sannolikheten att precis hälften av blommorna är gula? (3p)

3. De två oberoende stokastiska variablerna X och Y har samma väntevärde $E[X] = E[Y] = 2$ men olika varians $V(X) = 1$, $V(Y) = 3$. Låt nu $Z = 2X - 3Y$ och $W = 2XY$. Vad är kovariansen $Cov(Z, W)$? (6p)

4. Antag att olyckor på en given vägsträcka sker oberoende av varandra och att tiden mellan två olyckor är exponentialfördelad med väntevärde 0.5 (år). Vad är sannolikheten att
   (a) det dröjer mer än 9 månader mellan två olyckor? (3p)
   (b) det inte sker någon olycka under ett givet år? (3p)

5. I ett recept på äppelmos står det att man ska ta "14 lagom stora äpplen eller 2 kg". Man har slumpmässigt valt ut 14 hemodlade äpplen av sorten Ingrid Marie och anser att vikterna kan vara $N(120, 40^2)$-fördelade (enhet: gram).
   (a) Hur stor är sannolikheten att de 14 äpplenas sammanlagda vikt räcker till att få ihop minst 2 kg? (3p)
   (b) Hur många hemodlade Ingrid Marie-äpplen måste man använda för att sannolikheten att de tillsammans ska väga minst 2 kg skall vara minst 0.90? (4p)

Var god vänd!

6. Ett sätt att mäta ett reningsverks kapacitet är att mäta värdet på Biochemical Oxygen Demand (BOD) i det renade vattnet, och ju lägre BOD-värde desto renare vatten. Tillåtet utsläpp brukar vara maximalt 10 mg/l. I ett reningsverk ville man testa en ny reningsmetod och mätte vid sex olika tillfällen BOD-värdena för den gamla och den nya metoden.

Tillfälle	1	2	3	4	5	6
Gammal metod (BOD)	5.6	12.4	7.3	10.3	2.6	9.5
Ny metod (BOD)	4.7	9.3	6.6	8.5	1.1	8.2

Antag att BOD-värdena kommer från två oberoende normalfördelningar och testa på nivå 5% om det är någon skillnad mellan de båda metoderna. Är det någon skillnad? (6p)

Del II: Flervariabelanalys

7. Låt $x(t) = \ln t$, $y(t) = t^2$ och $f(x, y) = xe^{xy} + y$
Beräkna $\dfrac{d}{dt} f(x(t), y(t))$ genom att använda kedjeregeln för funktioner av flera variabler. (3p)

8. Beräkna trippelintegralen $\displaystyle\iiint_K \sqrt{x^2 + y^2 + z^2}\, dxdydz$,
då K är halvklotet $x^2 + y^2 + z^2 \leq 4$, $z \geq 0$ (4p)

9. En partikel rör sig längs kurvan $\mathbf{r}(t) = t^2\mathbf{i} + (1 + t)\mathbf{j}$, $0 \leq t < \infty$.
Skissa partikelbanan och indikera färdriktningen. I vilken punkt befinner sig partikeln då dess fart är 5 (l.e./t.e.) ? (3p)

Lösningar

Del I: Matematisk Statistik

1. (a) $P(A \cap B) = P(B|A)P(A) = 0.40 \cdot 0.01 = 0.004.$
   (b) $P(A|B) = P(A \cap B)/P(B) = 0.004/0.3 = 0.01333.$
   (c) $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.01 + 0.3 - 0.004 = 0.306.$

2. Definiera händelserna

$$\begin{aligned} A &= \text{blomman är en påsklilja} \\ B &= \text{blomman är en krokus} \\ C &= \text{blomman är en tulpan} \\ G &= \text{blomman är gul} \end{aligned}$$

Vi vet att $P(A) = P(B) = 1/4$ och $P(C) = 1/2$. Vidare vet vi att $P(G|A) = 1$, $P(G|B) = 2/3$ och $P(G|C) = 1/2$.

(a) Totala sannolikhetslagen ger att

$$P(G) = P(G|A)P(A) + P(G|B)P(B) + P(G|C)P(C) = 1 \cdot \frac{1}{4} + \frac{2}{3} \cdot \frac{1}{4} + \frac{1}{2} \cdot \frac{1}{2} = \frac{2}{3} = 0.667.$$

(b)

$$P(A|G) = \frac{P(A \cap G)}{P(G)} = \frac{P(G|A)P(A)}{P(G)} = \frac{3}{8} = 0.375.$$

(c) Låt X vara antalet gula blommor som planterats vid ytterdörren. Enlig (a) är sannolikheten för en gul blomma 2/3. Alltså är $X \sim Bin(10, 2/3)$ och vi får att

$$P(X = 5) = \binom{10}{5} \left(\frac{2}{3}\right)^5 \left(1 - \frac{2}{3}\right)^{10-5} = 0.1366$$

3. För $Z = 2X - 3Y$ och $W = 2XY$ har vi att

$$Cov(Z, W) = E[ZW] - E[Z]E[W]$$

där

$$\begin{aligned} E[Z] &= E[2X - 3Y] = 2E[X] - 3E[Y] = 2 \cdot 2 - 3 \cdot 2 = -2 \\ E[W] &= E[2XY] = 2E[X]E[Y] = 2 \cdot 2 \cdot 2 = 8 \end{aligned}$$

och eftersom

$$\begin{aligned} E[X^2] &= V(X) - (E[X])^2 = 1 - 2^2 = -3 \\ E[Y^2] &= V(Y) - (E[Y])^2 = 3 - 2^2 = -1 \end{aligned}$$

blir

$$\begin{aligned} E[ZW] &= E[(2X - 3Y) \cdot 2XY] = E[4X^2Y - 6XY^2] \\ &= 4E[X^2]E[Y] - 6E[X]E[Y^2] = 4 \cdot (-3) \cdot 2 - 6 \cdot 2 \cdot (-1) = -24 + 12 = -12 \end{aligned}$$

och

$$Cov(Z, W) = -24 - (-2) \cdot 8 = -8.$$

4. (a) Om vi låter

$$X = \text{tiden mellan två olyckor}$$

så är $X \sim Exp(\lambda)$ och med $E[X] = 1/\lambda = 0.5$ (år) har vi att $\lambda = 2$. Med 9 mån $= 0.75$ år får vi

$$\begin{aligned} P(X > 0.75) &= \int_{0.75}^{\infty} 2e^{-2x} dx \\ &= \left[-e^{-2x}\right]_{0.75}^{\infty} \\ &\approx 0 - (-0.223) = 0.223 \end{aligned}$$

(b) Om

$$N(t) = \text{antal olyckor i } [0, t]$$

så är $N(t)$ en Poisson-process med intensitet $\lambda = 2$. Dvs sannolikheten att det inte sker någon olycka under ett år ($t = 1$) blir

$$P(N(1) = 0) = e^{-2\cdot 1} \frac{(2 \cdot 1)^0}{0!} = e^{-2} \approx 0.135$$

5. (a) Låt

$$X_i = \text{vikten av äpple nr } i$$

så att $X_i \sim N(120, 40^2)$. Om

$$Y = \text{vikten av 14 äpplen}$$

får vi, eftersom äpplena är oberoende, att

$$Y = \sum_{i=1}^{14} X_i \sim N\left(\sum_{i=1}^{14} E[X_i], \sum_{i=1}^{14} V(X_i)\right) = N(14\cdot 120, 14\cdot 40^2) = N(1680, 22400)$$

Och

$$\begin{aligned} P(Y \geq 2000) &= 1 - P(Y \leq 2000) = 1 - \Phi\left(\frac{2000 - 1680}{\sqrt{22400}}\right) \\ &= 1 - \Phi(2.14) = 1 - 0.9837 = 0.0163 \end{aligned}$$

(b) Låt nu istället Y = vikten av n äpplen så att $Y \sim N(120n, 40^2 n)$. Vi vill hitta n så att $P(Y \geq 2000) \geq 0.90$. Dvs

$$P(Y \geq 2000) = 1 - P(Y \leq 2000) = 1 - \Phi\left(\frac{2000 - 120n}{40\sqrt{n}}\right) = 0.90 \Leftrightarrow$$

$$\Phi\left(\frac{2000 - 120n}{40\sqrt{n}}\right) = 0.1 \Leftrightarrow \frac{2000 - 120n}{40\sqrt{n}} = -1.28 \Leftrightarrow$$

$$\Leftrightarrow \frac{120n - 2000}{40\sqrt{n}} = 1.28$$

$$\Leftrightarrow n - \frac{1.28}{3} \cdot \sqrt{n} - \frac{50}{3} = 0$$

Det här är en kvadratisk ekvation: låt $x^2 = n$ så att

$$3x^2 - 1.28x - 50 = 0$$

Dvs

$$\begin{aligned} x &= \frac{1}{2 \cdot 3}\left(1.28 \pm \sqrt{1.28^2 - 4 \cdot 3 \cdot (-50)}\right) \\ &= \frac{1}{6}\left(0.213 \pm \sqrt{601.64}\right) \end{aligned}$$

och vi vill bara ha den positiva lösningen så

$$n = x^2 = \left(\frac{1}{6}(1.28 + 24.53)\right)^2 = 4.30^2 = 18.5$$

Dvs, det behövs minst $n = 19$ äpplen för att kommer över 2 kg med sannolikheten 0.9.

6. Vi har två oberoende stickprov där X_1 = gammal metod $\sim N(\mu_1, \sigma_1^2$ och X_2 = ny metod $\sim N(\mu_2, \sigma_2^2)$. Vi vill testa på nivå $\alpha = 0.05$

$$\begin{aligned} H_0 &: \mu_1 = \mu_2 \\ H_1 &: \mu_1 \neq \mu_2 \end{aligned}$$

σ_1 och σ_2 är okända, men vi får anta att de är lika (borde ha stått i uppgiften). Vår test-statistika blir

$$T_0 = \frac{\bar{X}_1 - \bar{X}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

där

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

och med

$$\begin{aligned} n_1 &= n_2 = 6 \\ \bar{X}_1 &= 7.95 \\ \bar{X}_2 &= 6.4 \\ s_1^2 &= 12.459 \\ s_2^2 &= 9.416 \\ s_p &= \sqrt{\frac{(6-1)\cdot 12.459 + (6-1)\cdot 9.416}{5+5-2}} = \sqrt{13.672} = 3.698 \end{aligned}$$

får vi att

$$T_0 = \frac{7.95 - 6.4}{3.698\sqrt{2/6}} = 0.726$$

Ur tabell får vi

$$t_{0.025,8} = 2.306$$

och eftersom $-2.306 < 0.726 < 2.306$ kan vi inte förkasta H_0. Vi ser alltså ingen signifikant skillnad mellan metoderna.

Del II: Flervariabelanalys

7. $\frac{d}{dt}\left(f(x,y)\right) \ = f_1'(x,y)\cdot x' + f_2'(x,y)\cdot y' =$

$= (e^{xy} + xye^{xy}){\cdot}x' {+} (x^2e^{xy}{+}1){\cdot}y' =$

$= e^{t^2\ln t}\left((1+t^2\ln t)\frac{1}{t} + (\ln t)^2 2t\right) + 2t$

8. $\iiint_K \sqrt{x^2+y^2+z^2}\,dxdydz =$

$= \int_0^{2\pi}\left(\int_0^{\pi/2}\left(\int_0^2 \rho\cdot\rho^2\sin\phi\,d\rho\right)d\phi\right)d\theta =$

$= 2\pi\left[-\cos\phi\right]_0^{\pi/2}\left[\frac{1}{4}\rho^4\right]_0^2 = 8\pi$

9. Partikelbanan går mot höger längs kurvan $y = \sqrt{x}{+}1$, med startpunkt i $(0,1)$. Hastigheten är $\mathbf{r}'(t) = (2t,1)$ så farten är $\sqrt{(2t)^2+1^2} = \sqrt{4t^2+1}$. Vidare är $\sqrt{4t^2+1} = 5$ då $t = \sqrt{6}$ och vid denna tidpunkt befinner sig partikeln i punkten $\mathbf{r}(\sqrt{6}) = ((\sqrt{6})^2, 1+\sqrt{6}) = (6, 1+\sqrt{6})$.