Tentamentsskrivning i **Matematisk statistik TMA321, 4.5 hp.**

Tid: Måndagen den 15 januari, 2009 kl 08:30-12:30
Examinator och jour: Erik Broman, tel. 772-3514, mob. 073 7320791, MV-huset rum L3080.
Hjälpmedel: valfri räknare, egenskriven formelsamling (4 sidor på 2 blad A4) samt utdelade tabeller.

———————————————

Tentamen består av 8 frågor om sammanlagt 30 poäng. Preliminära betygsgränser är satta till:
betyg "3": 12 till 17 poäng
betyg "4": 18 till 23 poäng
betyg "5": 24 eller fler poäng.

———————————————

1. (3 poäng) Vi betraktar två händelser A, B sådana att $\mathbb{P}(A) = 0.3$ och $\mathbb{P}(B) = 0.5$.

a. Antag att händelserna A, B är oberoende, beräkna $\mathbb{P}(A \cup B)$, $\mathbb{P}(A \cap B)$ och $\mathbb{P}(A|B)$.

b. Antag att A, B är beroende med $\mathbb{P}(A|B) = 0.6$, beräkna $\mathbb{P}(A \cup B)$, $\mathbb{P}(A \cap B)$ och $\mathbb{P}(B|A)$.

c. Skulle det gå att hitta en händelse C så att $\mathbb{P}(C) = 0.6$, $A \cap C = \emptyset$ och $B \cap C = \emptyset$?

Lösning uppg 1.: a. Vi har att $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) = 0.3 * 0.5 = 0.15$, vidare är $\mathbb{P}(A|B) = \mathbb{P}(A) = 0.3$ och $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) = 0.3 + 0.5 - 0.15 = 0.65$.

b.Vi har att $\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B) = 0.6 * 0.5 = 0.3$, vidare är $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) = 0.3 + 0.5 - 0.3 = 0.5$ och $\mathbb{P}(B|A) = \mathbb{P}(B \cap A)/\mathbb{P}(A) = \mathbb{P}(A|B)\mathbb{P}(B)/\mathbb{P}(A) = 0.6 * 0.5/0.3 = 1$.

c. Nej. Om $B \cap C = \emptyset$ gäller att $\mathbb{P}(B \cup C) = \mathbb{P}(B) + \mathbb{P}(C) = 0.5 + 0.6 > 1$.

2. (3 poäng)
Antag att X, Y är två oberoende $N(0, 1)$-fördelade variabler.
a. Visa med hjälp av momentgenererande funktioner att $U = X + Y$ är normalfördelad.

b. Låt Z anta värdena $\{-1, 1\}$ med sannolikhet $1/2$ vardera. Antag att Z är oberoende av X och låt $V = ZX$. Vad har V för fördelning? (motivera)

c. Är $W = X + V$ också normalfördelad? (motivera)

Lösning uppg 2.:

a. Dom momentgenererande funktionerna är $M_X(t) = M_Y(t) = e^{t^2/2}$. Därmed blir $M_U(t) = M_X(t)M_Y(t) = e^{t^2}$ som är den momentgenererande funktionen för en $N(0,2)$ fördelad variabel.

b. Vi har att $\mathbb{E}[e^{tV}] = \frac{1}{2}\mathbb{E}[e^{tX}] + \frac{1}{2}\mathbb{E}[e^{-tX}] = \mathbb{E}[e^{tX}]$, därför är V ocksa $N(0,1)$ fördelad. Detta kan inses på en massa andra olika sätt. T.ex. är täthetsfunktionen symmetrisk kring origo.

c. Nej! Vi har att $W = X + V = X + ZX = (1+Z)X$. Därför gäller att

$$\mathbb{P}(W = 0) = \mathbb{P}(Z = 1) = \frac{1}{2}.$$

Detta kan inte vara sant om W skulle vara normalfördelad.

3. (4 poäng) Ett taxibolag i en stad numrerar sina taxibilar med $1, 2, \ldots, N$. Emil är ny i staden och under en promenad så observerar han 7 taxibilar från detta bolag. Dessa har nummer

$$97,\ 234,\ 166,\ 7,\ 65,\ 17,\ 4.$$

Emil vill nu uppskatta antalet bilar i detta bolag.

a. Vilka antaganden på Emils observationer måste han göra för att skatta bolagets totala antal bilar N?

b. Tag fram MLE (maximum likelihood estimate) skattaren för N.

c. Tag fram MME (method of moments estimate) skattaren för N.

d. Varför är MME olämplig för att uppskatta N i detta fall?

Lösning uppg 3.:

a. Att bilnumren han ser är oberoende.

b. Vi har att $f(X_1, \ldots, X_7|N) = \frac{1}{N^7}1_{X_i \leq N}$. Denna funktion maximeras uppenbarligen av att välja $\hat{N} = \max(X_1, \ldots, X_7) = 234$.

c. Låt $\bar{X} = (X_1 + \cdots + X_7)/7$, vi har att $\mathbb{E}[\bar{X}] = \mathbb{E}[\bar{X_1}] = (N+1)/2$, därmed skall vi välja $\hat{N} = 2\bar{X} - 1 \approx 168$.

d. Svaret är uppenbarligen fel då vi vet att $N \geq 234$.

4. (4 poäng) Emilia har fått en sommarpraktikplats på "elstängsel'R'US". Hon får i uppgift att testa voltstyrkan på 20 stängsel som nyligen har satts upp av företaget. Resultaten ges av följande tabell:

stängsel nr:	1	2	3	4	5	6	7	8	9	10
voltstyrka:	37.73	38.50	48.56	43.70	49.15	49.78	49.26	52.15	49.04	51.65
stängsel nr:	11	12	13	14	15	16	17	18	19	20
voltstyrka:	39.69	45.60	44.92	37.68	46.99	40.42	52.78	41.67	48.34	46.80

Emilia kan anta att voltstyrkan är normalfördelat med parametrar μ, σ.

a. Skatta medelvärdet (μ) och variansen (σ^2) av voltstyrkan ur givna data.

b. Ange ett 95 % konfidensintervall för μ. Motivera ditt val av metod!

c. Voltstyrkan får inte overstiga 65 volt då detta anses orsaka djuren onödigt lidande. Vad kan man ur Emilias mätningar dra för slutsats om sannolikheten för att detta skall ske i någon av "elstängsel'R'US" elstängsel? Bör Emilia slå larm tilll djurskyddsmyndigheten?

Lösning uppg 4.:

a. Medelvärdet blir 45.72. Variansen skattas med hjälp av formeln

$$\hat{s}^2 = \frac{1}{19}\sum_{i=1}^{20}(X_i - \bar{X})^2 \approx 24.32$$

b. Vi använder oss av t-fördelning med 19 frihetsgrader. Det relevanta värdet läses ur tabell att vara 2.093. Därför blir ett 95 % CI

$$\bar{X} \pm s_{\bar{X}} 2.093 = 45.72 \pm 2.093 * 1.10 = 45.72 \pm 2.30.$$

Här är $s_{\bar{X}} = \hat{s}/\sqrt{20}$.

c. Vi ser att medelvärdet är ca 45 och variansen ca 5, därför är värdet 65 ungefär fyra standardavvikelser bort från medelvärdet. Detta är oerhört osannolikt och därmed kan hon med gott samvete låta bli att larma djurskyddsmyndigheten.

5. (4 poäng) En person anses besitta övernaturliga krafter. Detta ifrågasätts av några forskare. I ett experiment får därför denna person 100 gånger gissa utfallet av ett tärningskast med en vanlig sexsidig tärning.

a. Sätt upp nollhypotes och alternativ hypotes för detta test.

b. Hur många rätt måste personen gissa för att nollhypotesen skall förkastas på 5 % nivån? Använd gärna normalapproximation.

c. Antag att personen verkligen besitter övernaturliga krafter och därför har en sannolikhet att gissa rätt som är 0.25. Vad är sannolikheten att testet ovan kommer till rätt slutsats? Använd gärna normalapproximation.

Lösning uppg 5.:

a. Gissningarna är uppenbarligen oberoende med sannolikhet p att lyckas, dvs antalet korrekta gissningar är $Bin(100, p)$. Vår nollhypotes H_0 måste bli att personen är en bluff och därmed att $p = 1/6$. Vår alternativa hypotes H_1 blir då att $p > 1/6$.

b. Låt X beteckna antalet rätt. Nollhypotesen säger att $X \sim Bin(100, 1/6)$. Normalapproximation ger att $X \sim N(100/6, 500/36)$. Detta ger att $Z = (X - 100/6)/\sqrt{500/36} \sim N(0,1)$. Ur tabell läses att $0.95 = \mathbb{P}(Z \geq 1.645)$. Detta ger i sin tur att

$$X \geq \frac{100}{6} + \sqrt{\frac{500}{36}} * 1.645 = 22.8,$$

Vi drar slutsatsen att vi förkastar nollhypotesen på 5%-nivån om $X \geq 23$.

c. Om i själva verket $p = 0.25$ så blir frågan: Vad är sannolikheten att $X \geq 23$ om $X \sim Bin(100, 0.25)$? Som ovan blir nu $Z = (X - 25)/\sqrt{100 * 0.25 * 0.75} \sim N(0,1)$. Vi får beräkna

$$\mathbb{P}\left(\frac{X - 25}{\sqrt{100 * 0.25 * 0.75}} \geq \frac{23 - 25}{\sqrt{100 * 0.25 * 0.75}}\right) = \mathbb{P}(Z \geq -0.4619) \approx 0.677.$$

6. (4 poäng) Antag att (X, Y) är likformigt fördelad på enhetscirkeln (dvs en cirkel med radie 1, centrerad runt origo).

a. Ange den gemensamma pdf:en (sannolikhetsfördelningen) av (X, Y).

b. Ange marginalfördelningarna för X och Y.

c. Ange dom betingade fördelningarna for X givet Y och vice versa.

Lösning uppg 6.:

a. Enhetscirkeln har area π därmed gäller att

$$f(x, y) = \begin{cases} \frac{1}{\pi} & \text{om } x^2 + y^2 \leq 1 \\ 0 & \text{annars.} \end{cases}$$

b. Vi har:

$$f_X(x) = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} f_{X,Y}(x,y)dy = \begin{cases} \frac{2\sqrt{1-x^2}}{\pi} & \text{om } -1 \leq x \leq 1 \\ 0 & \text{annars.} \end{cases}$$

På samma sätt:

$$f_Y(y) = \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} f_{X,Y}(x,y)dx = \begin{cases} \frac{2\sqrt{1-y^2}}{\pi} & \text{om } -1 \leq y \leq 1 \\ 0 & \text{annars.} \end{cases}$$

c. Vi får här:

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \begin{cases} \frac{1}{2\sqrt{1-y^2}} & \text{om } -1 \leq y \leq 1 \quad \text{för } -\sqrt{1-y^2} \leq x \leq \sqrt{1-y^2} \\ 0 & \text{annars.} \end{cases}$$

På samma sätt:

$$f_{Y|X}(y|x) = \frac{f_{Y,X}(y,x)}{f_X(x)} = \begin{cases} \frac{1}{2\sqrt{1-x^2}} & \text{om } -1 \leq x \leq 1 \quad \text{för } -\sqrt{1-x^2} \leq y \leq \sqrt{1-x^2} \\ 0 & \text{annars.} \end{cases}$$

7. (4 poäng)

Antag att det finns 10 bollar i en urna, vara 7 är blåa och 3 är vita. Bollarna dras ur urnan en efter en och placeras på rad.

a. Vad är sannolikheten att de tre vita bollarna hamnar på rad?

b. Antag att den tredje bollen som dras är vit, vad är sannolikheten att ingen av de två första bollarna är blåa?

c. Antag igen att den tredje bollen som dras är vit, vad är sannolikheten att den sista bollen är blå?

Lösning uppg 7.:

a. Dom tre vita kan hamna bredvid varandra i åtta olika positioner. Därför blir den efterfrågade sannolikheten

$$8 * 3! * 7!/10! \approx 0.0667$$

b. Om den tredje är vit så har vi två vita kvar att ta hänsyn till. Sannolikheten att ingen av de två första är blåa blir då samma som sannolikheten att de två första är vita, vilket blir

$$2/9 * 1/8 = 2/72.$$

c. Denna sannolikhet blir

$$7/9.$$

8. (4 poäng) Två kompisar, Kalle och Ada, står på ett hustak och kastar vattenballonger på förbipasserande. Sannolikheten att Kalle traffar kallar vi för p_K och sannolikheten att Ada träffar kallar vi för p_A. Vi antar också att sannolikheten för träff är oberoende mellan kasten. Låt X, Y vara antalet kast som Kalle respektive Ada behöver innan dom träffar första gången.

a. Vad har X och Y för fördelning?

b. Antag att Ada börjar kasta, vad är sannolikheten att Ada vinner?

c Antag att Kalle börjar kasta, vad är sannolikheten att Ada vinner?

Lösning uppg 8.:

a. Dom är Geometriskt fördelade med parameter p_K respektive p_A.

b. Sannolikheten att Ada vinner pa exakt kast nummer j är $p_A(1-p_A)^{j-1}(1-p_K)^{j-1}$. Vi får att sannnolikheten att Ada vinner blir

$$\sum_{j=1}^{\infty} p_A(1-p_A)^{j-1}(1-p_K)^{j-1} = p_A \sum_{j=0}^{\infty}(1-p_A)^j(1-p_K)^j = \frac{p_A}{1-(1-p_A)(1-p_K)}.$$

c. Sannolikheten att Ada vinner på exakt kast nummer j är nu $p_A(1-p_A)^{j-1}(1-p_K)^j$. Vi får att sannnolikheten att Ada vinner blir

$$\sum_{j=1}^{\infty} p_A(1-p_A)^{j-1}(1-p_K)^j = (1-p_K)\frac{p_A}{1-(1-p_A)(1-p_K)}.$$

Denna formel inses lättast genom att tänka att om Kalle missar första kastet (med sannolikhet $1-p_K$), så är vi genast i stuationen i b.