Tentamentsskrivning i **Matematisk Statistik TMA321**

Dag: Onsdagen den 29 augusti, 2018
Hjälpmedel: Typgodkänd miniräknare, egenhändigt skriven formelsamling om två A4 fram och bak (dvs 4 sidor), samt utdelade tabeller.

Tentamen består av 8 frågor om sammanlagt 50 poäng. Preliminära betygsgränser är satta till:
betyg "3": 20 till 29 poäng
betyg "4": 30 till 39 poäng
betyg "5": 40 eller fler poäng.

OBS! Alla lösningar skall vara väl redovisade och motiverade. Talen är ej ordnade efter svårighetsgrad.

1. Alexandra har en byrålåda med 6 röda, 4 blå och 4 göna strumpor.

   (a) Beräkna sannolikheten att Alexandra får två strumpor av samma färg om hon drar två strumpor slumpmässigt. (2p)

   (b) Om hon i uppgift a drar två strumpor av samma färg, vad är sannolikheten att dessa är röda? (2p)

   (c) Vad är sannolikheten att hon får ett färgmatchande par om hon istället drar tre strumpor? (2p)

2. Carl för statistik över hur långt han går varje dag. Carl kommer fram till att sträckan kan betraktas som en kontinuerlig slumpvariabel X med täthetsfunktion

$$f(x) = \frac{30x^2 - 3x^3}{2500} \text{ för } 0 \leq x \leq 10.$$

   Enheten är sjömil.

   (a) Beräkna väntevärde och standardavvikelse för X. (2p)

   (b) Beräkna sannolikheten att Carl sammanlagt går mindre än eller likamed 2200 sjömil under en tidsperiod om 365 dagar. Vi kan anta att gångsträckorna är oberoende mellan dagarna. (2p)

   (c) Betrakta återigen 365 dagar, och låt Y vara antalet av dessa då Carl går längre än 8 sjömil. Beräkna sannolikheten att Y är större än eller lika med 45. (2p)

3. I ett experiment mäts drogkoncentration i blodet som en funktion av läkemedelsdosen. De uppmätta data blev som följer:

dos (mg):	50	100	200	300	400	600	800
konc (μg/liter):	54.5	62.9	132.2	266.8	291.7	427.9	623.5

Data kan sammanfattas med att $S_{xx} = 445000, S_{yy} \approx 258562$ och $S_{xy} = 336890$.

Forskarna ansatta en linjär regressionsmodell $y = \beta_0 + \beta_1 x$ där x är dosen (i mg) och y är blodkoncentrationen (i μg per liter blod). Lös följande uppgifter:

(a) Skatta β_0 och β_1. (2p)

(b) Skapa ett 95% konfidensintervall för β_1 och testa huruvida $\beta_1 = 0.5$ på 95%-nivån. (2p)

(c) Ange förklaringsgraden och beräkna residualerna. Verkar modellen rimlig? (2p)

4. Låt (X, Y) vara likformigt fördelade på området

$$A = \{0 \leq x \leq 2, x \leq y \leq 2x\}.$$

(a) Hitta den gemensamma täthetsfunktionen för (X, Y) och även de två marginaltäthetsfunktionerna. (3p)

(b) Bestäm de betingade täthetsfunktionerna. (2p)

(c) Ange explicita uttryck för slumpvariablerna $\mathbb{E}[Y|X]$ och $\mathbb{E}[X|Y]$. (2p)

5. Låt X_n vara en diskret slumpvariabel med sannolikhetsfunktion

$$\mathbb{P}(X_n = k/n) = \begin{cases} \frac{1}{3n} & \text{om } k = 0, \ldots, n-1 \\ \frac{2}{3n} & \text{om } k = n, \ldots, 2n-1, \end{cases}$$

och låt X vara en kontinuerlig slumpvariabel med täthetsfunktion

$$f(x) = \begin{cases} \frac{1}{3} & \text{om } 0 \leq x \leq 1 \\ \frac{2}{3} & \text{om } 1 < x \leq 2. \end{cases}$$

(a) Hitta den momentgenererande funktionen $M_{X_n}(t)$ för X_n. Skriv den på enklaste form. (2p)

(b) Hitta den momentgenererande funktionen $M_X(t)$ för X. Skriv den på enklaste form. (2p)

(c) Visa att $M_{X_n}(t) \to M_X(t)$ (och därmed att $X_n \xrightarrow{d} X$). (2p)

6. Låt X vara en slumpvariabel med täthetsfunktion

$$f(x) = (2 + \theta)x^{1+\theta} \text{ för } 0 \leq x \leq 1,$$

där $\theta > -2$ är parametern vars värde skall skattas.

(a) Hitta momentskattaren (MME) för θ. (2p)

(b) Hitta maximum likelihoodskattaren (MLE) för θ. (3p)

7. Den galna kattladyn Cat har jättemånga katter. Tyvärr har de blivit feta av all utfodring, så Cat bestämmer sig för att sätta dem på diet. Cat testar dieten på nio av sina katter. Hon väger dem innan dieten börjar och sedan igen efter två månader. Hon fick följande data

katt nr:	1	2	3	4	5	6	7	8	9
vikt innan:	4.64	5.12	5.67	5.31	4.68	4.53	5.05	5.12	5.89
vikt efter:	4.3	5.19	5.43	4.99	4.42	4.16	4.63	5.43	6.02

(a) Skatta effekten av Cats diet, och skatta även standardavvikelsen för densamma. (2p)

(b) Ange ett 95% ensidigt konfidensintervall för viktskillnaden. Formulera lämpligt hypotestest motsvarande detta konfidensintervall. Verkar Cats diet ge effekt? (2p)

(c) Vad är p-värdet av ditt test? (2p)

(d) Vilket/vilka antaganden måste du göra för att kunna genomföra din analys? (1p)

8. Antag att vi har tre slumpvariabler $X \sim N(\mu, \sigma^2), Y \sim N(\gamma, \sigma^2)$ och $Z \sim N(\mu + \gamma, \sigma^2)$, och låt x, y, z vara oberoende observationer från respektive slumpvariabler. Om vi vill skatta μ kan vi t.ex. använda $\mu^* = x$, eller $\hat{\mu} = (x - y + z)/2$.

(a) Visa att båda skattarna är väntvärdesriktiga (man får förstås använda att väntevärdet av en normalfördelad slumpvariabel är känd). (2p)

(b) Vilken av skattarna är effektivast? (2p)

(c) Kan du hitta en egen väntevärdesriktig skattare som är effektivare än båda de två föreslagna? (3p)