Tillämpad matematisk statistik LMA201 (Elektro-programmet) Tentamen 2018-06-05

Tid: 14.00-18.00. **Tentamensplats:** Lindholmen
Hjälpmedel: Kursboken **Matematisk Statistik** av Ulla Dahlbom. Formelsamlingen **Tabell- och formelsamling i matematisk statistik, försöksplanering och kvalitetsstyrning** av Håkan Blomqvist. Boken och formelsamlingen får ej innehålla extra anteckningar, men understrykningar, sticks och markeringar är tillåtna. **Chalmersgodkänd räknare.**
Examinator: Reimond Emanuelsson/Johan Tykesson
Telefonvakt/jour: Reimond, 0708948456 / Johan, 0703182096
Till varje uppgift skall fullständig lösning lämnas!
OBS: text på TRE sidor!
Betygsgränser: För betyg 3, 4 resp. 5 krävs minst 20, 30 resp. 40 poä ng.

1. (2+4 poäng) Antag att mätvärdena 4.1, 4.2, 4.1 och 4.3 kommer från en normalfördelning med väntevärde μ och standardavvikelse σ. Antag också att mätningarna är gjorda oberoende av varandra.

   (a) Beräkna ett 95% konfidensintervall för μ om $\sigma = 0.1$.

   (b) Beräkna ett 95% konfidensintervall för μ om σ okänt.

   Lösning:

   (a)
   $$\bar{x} \pm 1.96\sigma/\sqrt{n} = 4.175 \pm 1.96 \times 0.1/2 = 4.175 \pm 0.098 = [4.077, 4.273].$$

   (b)
   $$\bar{x} \pm 3.18s/\sqrt{n} = 4.175 \pm 3.18 \times 0.096/2 = 4.175 \pm 0.153 = [4.022, 4.328].$$

2. (4+4 poäng) Antag att livslängden för en viss typ av smart-phone är exponentialfördelad med väntevärde 2 år. Antag att vi har 100 smartphones av denna modell, och att deras livslängder kan antas oberoende av varandra.

   (a) Beräkna (approximativt) sannolikheten att summan av de 100 telefonernas livslängder är mer än 210 år.

   (b) Beräkna (approximativt) sannolikheten att antalet telefoner som fungerar efter 2 år är mindre än eller lika med 53.

   Lösning:

   (a) Låt ξ_i vara livslängden för smartphone nummer i, $i = 1, \dots, 100$. Det gäller att $E(\xi) = 2$ och $\sigma(\xi) = 2$. Enligt CGS gäller att $T = \sum_{i=1}^{100}$ är approximativt $N(100 \times 2, \sqrt{100} \times 2) = N(200, 20)$. Alltså blir

   $$P(T > 210) = 1 - P(T \leq 210) = 1 - P(\frac{T-200}{20} \leq \frac{210-200}{20})$$
   $$\approx 1 - \Phi(0.5) = 1 - 0.691 = 0.309.$$

(b)

$$P(\xi \geq 2) = \int_2^\infty \frac{1}{2} e^{-x/2}\, dx = [-e^{-x/2}]_2^\infty = e^{-1} \approx 0.368.$$

Alltså är antalet hela smartphones efter 2 år binomialfördelat med $n = 100$ och $p = 0.368$. Beteckna detta antal med η. Eftersom $np(1-p) \approx 23.25 > 10$ så är η approx $N(np, \sqrt{np(1-p)}) = N(36.8, 4.82)$. Därfor blir

$$P(\eta \leq 53) = P(\frac{\eta - 36.8}{4.88} \leq \frac{53 - 36.8}{4.82}) \approx \Phi(3.36) \approx 0.9996. \quad (1)$$

3. (4 poäng) Antag att A, B och C är händelser, och att de alla är oberoende av varandra. Antag också att det gäller att $P(A) = P(B) = P(C) = 0.4$.

(a) Vad är sannolikheten att alla händelserna inträffar?

(b) Vad är sannolikheten att exakt en av händelserna inträffar?

Lösning:

(a) Oberoende ger att

$$P(A \cap B \cap C) = 0.4^3 = 0.064.$$

(b) Notera att

$$P(\text{ endast } A) = P(A \cap B^c \cap C^c) = P(A)P(B^c)P(C^c) = 0.4 \times 0.6^2 = 0.144,$$

där oberoende användes i andra likheten. På samma sätt fås att $P(\text{ endast } B) = 0.144$ och $P(\text{ endast } C) = 0.144$. Så

$$P(\text{exakt en av händelserna inträffar}) = 3 \times 0.144 = 0.432.$$

4. (2+2+2 poäng) Antag att ξ är en kontinuerlig stokastisk variabel med frekvensfunktion

$$f(x) = \begin{cases} 5x^4 & \text{för } 0 \leq x \leq 1 \\ 0 & \text{för övrigt} \end{cases}$$

(a) Visa att $f(x)$ är en frekvensfunktion.

(b) Beräkna väntevärde och standardavvikelse för ξ.

(c) Beräkna den betingade sannolikheten

$$P(0.2 \leq \xi \leq 0.6 \,|\, 0.4 \leq \xi \leq 0.8).$$

Lösning:

(a) För $0 \leq x \leq 1$ gäller att $5x^4 \geq 0$. Dessutom är $\int_0^1 5x^4\, dx = 1$.

(b)

$$E(\xi) = \int_0^1 5x^5\, dx = \frac{5}{6}$$

.

$$E(\xi^2) = \int_0^1 5x^6\, dx = \frac{5}{7}.$$

$$\sigma = \sqrt{5/7 - (5/6)^2} \approx 0.141.$$

(c)

$$P(0.2 \leq \xi \leq 0.6 \,|\, 0.4 \leq \xi \leq 0.8) = \frac{P(0.4 \leq \xi \leq 0.6)}{P(0.4 \leq \xi \leq 0.8)} = \frac{\int_{0.4}^{0.6} 5x^4\,dx}{\int_{0.4}^{0.8} 5x^4\,dx} \approx 0.213.$$

5. (2+1+2+1 poäng) Ett fullständigt faktorförsök har gjorts enligt försöksplanen i Tabell 1. Två olika faktorer användes och i varje försöksgrupp gjordes 20 mätningar. I tabellen kan man se stickprovsmedelvärdet och stickprovsvariansen från varje försöksgrupp.

Försöksgrupp nr	A	B	AB	Resultat $\hat{y}$	Resultat s^2
1	-	-	+	$\bar{y}_1 = 66.81$	$s_1^2 = 333.61$
2	+	-	-	$\bar{y}_2 = 48.87$	$s_2^2 = 303.71$
3	-	+	-	$\bar{y}_3 = 59.78$	$s_3^2 = 262.26$
4	+	+	+	$\bar{y}_4 = 57.04$	$s_4^2 = 441.22$

Tabell 1: Provtagningsplan och uträknade stickprovsmedelvärden och stickprovsvarianser.

(a) Beräkna huvudeffekterna, medelvärdet samt tvåfaktorsamspelet.

(b) Man har beräknat att ett 95% referensintervall ges av $[-8.15, 8.15]$. Vilka huvudeffekter anser du är signifikanta (med en signifikansgrad av 5%) givet detta referensintervall?

(c) Om vi antar att vi skulle vara intresserade av ytterligare en faktor, C, och vi väljer provtagningsplanen såsom i Tabell 2 nedan, vad blir alias för A, B och C?

(d) Vad är upplösningen för den reducerade försöksplanen?

Försöksgrupp nr	A	B	C
1	-	-	-
2	+	-	-
3	-	+	+
4	+	+	+

Tabell 2: Reducerad provtagningsplan.

Lösning:

(a)

$$\text{medelvärdet} = \frac{\bar{y}_2 + \bar{y}_4 + \bar{y}_1 + \bar{y}_3}{4} = \frac{48.87 + 57.04 + 66.81 + 59.78}{4} = 58.13 \tag{2}$$

$$l_A = \frac{\bar{y}_2 + \bar{y}_4}{2} - \frac{\bar{y}_1 + \bar{y}_3}{2} = \frac{48.87 + 57.04 - 66.81 - 59.78}{2} = -10.34 \tag{3}$$

$$l_B = \frac{\bar{y}_3 + \bar{y}_4}{2} - \frac{\bar{y}_1 + \bar{y}_2}{2} = \frac{59.78 + 57.04 - 66.81 - 48.87}{2} = 0.57 \tag{4}$$

$$l_{AB} = \frac{\bar{y}_1 + \bar{y}_4}{2} - \frac{\bar{y}_2 + \bar{y}_3}{2} = \frac{66.81 + 57.04 - 48.87 - 59.78}{2} = 7.6 \tag{5}$$

$$\tag{6}$$

(b) Huvudeffekten av A anser vi är signifikant eftersom den är utanför referensintervallet. Huvudeffekt B och tvåfaktorsamspelet AB är för nära 0.

(c) Kolumnerna för B och C är identiska. Alltså har vi generatorn $B = C$. Detta ger oss den definierande relationen $I = BC$ vilket i sin tur ger oss sammanblandningsmönstret:

$$A = ABC \tag{7}$$
$$B = C \tag{8}$$
$$AB = AC \tag{9}$$
$$BC = I \tag{10}$$

(d) Upplösningen är längden på det kortaste ordet, alltså 2.

6. (2+2+3 poäng) I en fabrik produceras skåpsluckor (av samma sort) på två band, A och B. På band A produceras 25% fler luckor än på band B (det vill säga, på samma tid som det produceras 100 luckor på band B, produceras 125 luckor på band A.) Andelen defekta luckor från band A är 2% och från band B 3.5%.

(a) Vad är sannolikheten att en skåpslucka är defekt?

(b) Vad är betingade sannolikheten att en skåpslucka kommer från band A, givet att den är defekt?

(c) Vad är betingade sannolikheten att en skåpslucka kommer från band B, givet att den är korrekt (ej defekt)?

Lösning: Låt D beteckna händelsen att luckan defekt.

(a)

$$\begin{cases} P(A) + P(B) = 1 \\ P(A) = 1.25P(B) \end{cases} \iff \begin{cases} P(A) = 5/9 \\ P(B) = 4/9 \end{cases}$$

$$P(D) = P(D|A)P(A) + P(D|B)P(B) = 002 \cdot 5/9 + 0.035 \cdot 4/9 = 2/75 \approx 0.027.$$

(b)

$$P(A|D) = P(D|A) \cdot \frac{P(A)}{P(D)} = 0.02 \cdot \frac{5/9}{2/75} \approx 0.42.$$

Sannolikheten, givet att den är defekt, att den kommer från band A är 0.42.

(c)

$$P(B|D^c) = \frac{(1 - P(D|B))P(B)}{1 - P(D)} = \frac{0.965}{73/75} \cdot \frac{5}{9} \approx 0.44.$$

7. (2+2+2 poäng) En pluton bestående av 30 soldater delas helt slumpmässigt in i 10 grupper om vardera 3 personer. I varje grupp tilldelas en person rollen som befäl, en person rollen som spanare, och en person rollen som skytt. Detta sker också helt slumpmässigt. Tre av soldaterna heter Kurt, Sven och Veronika.

   (a) Vad är sannolikheten att Kurt, Sven och Veronika hamnar i samma grupp?

   (b) Vad är sannolikheten att Kurt blir spanare, Sven blir skytt och Veronika blir befäl? (De behöver inte hamna i samma grupp)

   (c) Vad är sannolikheten att Kurt och Sven hamnar i samma grupp, men Veronika hamnar i en annan grupp?

Lösning:

(a)

$$\begin{aligned} P(\text{K,S,V i samma grupp}) &= 10 \times P(\text{K,S,V i första gruppen}) \\ &= 10 \times \frac{3}{30} \times \frac{2}{29} \times \frac{1}{28} = \frac{1}{406} \end{aligned}$$

(b)

$$P(\text{K spanare, S skytt, V befäl}) \quad = \quad \frac{10}{30} \times \frac{10}{29} \times \frac{10}{28} \quad = \quad \frac{25}{609}.$$

(c)

$$P(\text{K,S i samma grupp, V i annan grupp})$$

$$= 10 \times P(\text{K,S i första gruppen, V i annan grupp}) = 10 \times \frac{3}{30} \times \frac{2}{29} \times \frac{27}{28} = \frac{27}{406}$$

8. (3+3+1 poäng) Antag att vi har en Markovkedja $X(n)$ i diskret tid (n heltal ≥ 0) med tillståndsrum $\{1, 2, 3\}$. Följande gäller för övergångar mellan de olika tillstånden. Om man befinner sig i tillstånd 1 går man till tillstånd 2 med sannolikhet 1. Om man är i tillstånd 2 går man till tillstånd 1 med sannolikhet 1/2 och tillstånd 3 med sannolikhet 1/2. Om man är i tillstånd 3 går man till tillstånd 2 med sannolikhet 1/2 och stannar kvar i tillstånd 3 med sannolikhet 1/2.

   (a) Bestäm den stationära fördelningen för Markovkedjan.

(b) Bestäm (ungefär) väntevärdet och standardavvikelsen för $X(n)$ om n är väldigt stort.

(c) Om man startar i tillstånd 2, vad är sannolikheten att man är i tillstånd 2 efter två tids-steg?

Lösning:

(a) Övergångsmatrisen ges av

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1/2 & 1/2 \end{pmatrix}.$$

Om vi betecknar stationära fördelningen med $\pi = (\pi_1, \pi_2, \pi_3)$ så får vi genom att lösa matrisekvationen $\pi P = \pi$ under villkoret att $\pi_1 + \pi_2 + \pi_3 = 1$ att $\pi = (1/5, 2/5, 2/5)$.

(b) Väntevärdet om n väldigt stort blir

$$E(X(n)) \approx 1 \cdot \frac{1}{5} + 2 \cdot \frac{2}{5} + 3 \cdot \frac{2}{5} = 2.2$$

Dessutom får vi

$$E(X(n)^2) \approx 1 \cdot \frac{1}{5} + 4 \cdot \frac{2}{5} + 9 \cdot \frac{2}{5} = 5.4.$$

Så standardavvikelsen

$$\sigma \approx \sqrt{5.4 - 2.2^2} \approx 0.75$$

(c) Den sökta sannolikheten ges av $0.5 \times 0.5 + 0.5 \times 1 = 0.75$.

Lycka till!