

Tentamen i Tillämpad matematisk statistik LMA521 för EPI och MI
den 14 dec 2011

Tentamen består av åtta uppgifter om totalt 50 poäng. Det krävs minst 20 poäng för betyg 3, minst 30 poäng för 4 och minst 40 poäng för 5.

Examinator: Ulla Blomqvist, ankn 5886

Hjälpmedel: Chalmersgodkänd miniräknare, Matematisk statistik, Ulla Dahlbom och Håkan Blomqvists formelsamling. Formelsamlingen och boken får inte innehålla egna anteckningar.

Lycka till!

Uppgift 1: Anta att vi har två urnor, U_1 och U_2 . Urna U_1 innehåller 4 vita och 3 svarta kulor medan urna U_2 innehåller 4 vita och 6 svarta kulor. En symmetrisk tärning kastas. Om man får upp ett udda antal ögon så flyttas en slumpmässigt vald kula från urna U_1 till urna U_2 . Om man får upp ett jämnt antal ögon så flyttas en slumpmässigt vald kula från urna U_2 till urna U_1 . Vad är sannolikheten att man efter 2 kast har 4 vita kulor i urna U_2 ?

(6 poäng)

Uppgift 2: Ur ett parti med CD-skivor packat i lådor väljer man slumpmässigt ut 4 lådor för kontroll.

- a) Hur stor är sannolikheten att man i de utvalda lådorna hittar några repiga CD-skivor om hela partiet består av 10 lådor, varav 2 innehåller repiga CD-skivor?
- b) Anta att man upprepar denna kontroll varje gång ett sådant parti köps in. Anta också att kvaliteten är samma varje gång d.v.s. 2 av 10 lådor innehåller repiga CD-skivor. Under en månad genomförs 5 sådana kontroller. Vad är sannolikheten att man vid minst ett tillfälle finner repiga CD-skivor?

(6 poäng)

Uppgift 3: En dator uppgraderas med ett nytt operativsystem. Ändringen av tiden, i sekunder, att koppla upp sig mot Internet kan beskrivas av nedanstående frekvensfunktion:

$$f(x) = \begin{cases} c e^{-6x} & x \geq 0 \\ -c x & -1 < x \leq 0 \\ 0 & x \leq -1 \end{cases}$$

- a) Bestäm konstanten c .
- b) Bestäm fördelningsfunktionen
- c) Vad är sannolikheten att uppkopplingstiden med hjälp av den nya systemet förändras med minst en halv sekund?

(8 poäng)

Uppgift 4: Anta att du spelar ett spel där du kan vinna eller förlora 10 kr varje gång med sannolikheten 0.5 för vardera alternativet. Anta att du spelar detta spel 1000 gånger. Vad är sannolikheten att du vinner minst 100 kr totalt?

(6 poäng)

Uppgift 5: Ett återkommande problem är att retur fjädern i fartreglaget på låglyftande truckar går av. Vid inspektion av fjädrar som gått av har det visat sig att dessa har saknat fett. Man beslöt vid ett tillfälle att genomföra ett faktorförsök för att undersöka vilka faktorer som var kritiska för livslängden. Önskvärt vore om man kunde få fram en fjäder med tillfredsställande livslängd, även utan infettning. De faktorer respektive nivåer som användes var:

	-	+
A: Bockningsradie	3 mm	6 mm
B: Material	Vanlig fjädertråd	Rostfri tråd
C: Smörjmedel	Ej infettat	Infettat

Som resultatvariabel valdes "antal cykler till brott" i tusental. Provet avbröts vid 3 miljoner cykler, vilket motsvarar minst 15 års normaldrift. Följande resultat erhöles:

A	B	C	fjäder nr 1	fjäder nr 2	fjäder nr 3	fjäder nr 4	fjäder nr 5	Medel- värde	Standard- avvikelse
-	-	-	223	3000	3000	150	3000	1875	1541.2
+	-	-	150	1465	3000	3000	562	1635	1333.4
-	+	-	150	150	150	150	150	150	0.0
+	+	-	150	223	150	150	150	165	32.6
-	-	+	3000	3000	3000	3000	3000	3000	0.0
+	-	+	3000	3000	3000	3000	3000	3000	0.0
-	+	+	3000	2766	3000	562	382	1942	1346.8
+	+	+	562	3000	3000	3000	3000	2512	1090.3

Följande effekter beräknades

$$\ell_A = 86.25 \quad \ell_B = -1185.25 \quad \ell_{AB} = 206.25 \quad \ell_{ABC} = 78.75$$

- Använd ovanstående för att bestämma de resterande effekterna.
- Beräkna ett 99%-igt referensintervall för effekterna.
- Ange en matematisk modell för ovanstående situation där enbart signifikanta effekter ingår.

(8 poäng)

Uppgift 6: Anta att du skall genomföra ett 2^{5-2} -faktorförsök.

- Hur många faktorer kommer att undersökas?
- Hur många försök krävs?
- Hur många generatorer används?
- Hur många definierande relationer finns?

(4 poäng)

Uppgift 7: Anta att man vill bestämma en lämplig provtagningsplan som går genom punkterna ($p_1 = 0.01$, $L(p_1) = 0.95$) och ($p_2 = 0.05$, $L(p_2) = 0.10$) på OC-kurvan.

- a) Bestäm en enkel provtagningsplan som uppfyller detta villkor genom att använda det bifogade binomialfördelningsnomogrammet.
- b) Bestäm en dubbel provtagningsplan som uppfyller detta villkor.

(6 poäng)

Uppgift 8: För en process som är under statistisk kontroll är $p = 3\%$. Man tar dagligen 400 enheter för kontroll och använder ett p-diagram med 3-sigma-gränser.

- a) Vad är genomsnittligt antal inprickade punkter i kontrolldiagrammet fram till ett falsklarm, ARL?
- b) Vad är sannolikheten att en plötslig ändring till $p = 6\%$ upptäcks
 - 1) vid första provgruppen efter ändringen?
 - 2) vid någon av de tre första provgrupperna efter ändringen?

(6 poäng)

Lösningar till Tillämpad matematisk statistik den 14/12-11

Uppgift 1: Det finns 4 olika situationer som kan uppstå när man kastar tärningen:

(udda, udda) (udda, jämn) (jämn, udda) (jämn, jämn)

Vart och ett av fallen inträffar med sannolikheten $\frac{1}{4}$

Situation 1: (udda, udda) I denna situation flyttas en kula från urna U_1 efter både första och andra kastet. Det betyder att kulorna som flyttas båda måste vara svarta.

$$P(\text{svart efter kast 1} \cap \text{svart efter kast 2}) = \frac{3}{7} \cdot \frac{2}{6} = \frac{1}{7}$$

Situation 2: (udda, jämn) I denna situation flyttas en kula från urna U_1 efter första kastet och en kula från urna U_2 efter andra kastet. Nu kan 2 situationer uppkomma: (svart, svart) eller (vit, vit)

$$P(\text{svart efter kast 1} \cap \text{svart efter kast 2}) + P(\text{vit efter kast 1} \cap \text{vit efter kast 2}) =$$

$$= \frac{3}{7} \cdot \frac{7}{11} + \frac{4}{7} \cdot \frac{5}{11} = \frac{41}{77}$$

Situation 3: (jämn, udda) I denna situation flyttas en kula från urna U_2 efter första kastet och en kula från urna U_1 efter andra kastet. Även här kan 2 situationer uppkomma: (svart, svart) eller (vit, vit)

$$P(\text{svart efter kast 1} \cap \text{svart efter kast 2}) + P(\text{vit efter kast 1} \cap \text{vit efter kast 2}) =$$

$$= \frac{6}{10} \cdot \frac{4}{8} + \frac{4}{10} \cdot \frac{5}{8} = \frac{11}{20}$$

Situation 4: (jämn, jämn) I denna situation flyttas en kula från urna U_2 efter både första och andra kastet. Det betyder att kulorna som flyttas båda måste vara svarta.

$$P(\text{svart efter kast 1} \cap \text{svart efter kast 2}) = \frac{6}{10} \cdot \frac{5}{9} = \frac{1}{3}$$

$$P(4 \text{ vita i urna } U_2 \text{ efter 2 kast}) = \frac{1}{4} \left(\frac{1}{7} + \frac{41}{77} + \frac{11}{20} + \frac{1}{3} \right) = \frac{7201}{18480} \approx 0.3897$$

Uppgift 2 ξ = antal lådor med repiga CD-skivor i urvalet

$\xi = \text{Hyp}(N, Np, n) = \text{Hyp}(10, 2, 4)$

$$\text{a) } P(\xi > 0) = 1 - P(\xi = 0) = 1 - \frac{\binom{2}{0}\binom{8}{4}}{\binom{10}{4}} = 1 - \frac{8 \cdot 7 \cdot 6 \cdot 5}{10 \cdot 9 \cdot 8 \cdot 7} = 1 - \frac{1}{3} = \frac{2}{3}$$

b) η = antal tillfällen man finner repiga CD-skivor i urvalet

$$\eta = \text{Bin}(n, p) = \text{Bin}\left(5, \frac{2}{3}\right)$$

$$P(\eta > 1) = 1 - P(\eta = 0) = 1 - \binom{5}{0} \left(\frac{2}{3}\right)^0 \left(\frac{1}{3}\right)^5 \approx 1 - 0.00412 = 0.99588$$

Uppgift 3:

$$\begin{aligned} \text{a) } \int_{-\infty}^{\infty} f(x) dx &= 1 \Rightarrow -c \int_{-1}^0 x dx + c \int_0^{\infty} e^{-6x} dx = -c \left[\frac{x^2}{2} \right]_{-1}^0 + c \left[\frac{e^{-6x}}{-6} \right]_0^{\infty} = \\ &= -c \left(0 - \frac{1}{2} \right) + c \left(0 - \left(-\frac{1}{6} \right) \right) = c \frac{4}{6} \Leftrightarrow c = 1.5 \end{aligned}$$

b)

$x \leq -1$:

$$F(x) = \int_{-\infty}^x f(t) dt = 0$$

$-1 < x < 0$:

$$F(x) = \int_{-\infty}^x f(t) dt = \int_{-\infty}^{-1} 0 dt + 1.5 \int_{-1}^x (-t) dt = 1.5 \left[-\frac{t^2}{2} \right]_{-1}^x = 1.5 \left(-\frac{x^2}{2} + \frac{1}{2} \right) = 0.75(-x^2 + 1)$$

$0 \leq x$:

$$F(x) = \int_{-\infty}^x f(t) dt = \int_{-\infty}^{-1} 0 dt + 1.5 \int_{-1}^0 (-t) dt + 1.5 \int_0^x e^{-6t} dt =$$

$$= 0 + 1.5 \left[-\frac{t^2}{2} \right]_{-1}^0 + 1.5 \left[\frac{e^{-6t}}{-6} \right]_0^x = 1.5 \left(0 + \frac{1}{2} \right) + 1.5 \left(\frac{e^{-6x}}{-6} + \frac{1}{6} \right) = -0.25e^{-6x} + 1$$

fortsättning uppgift 3b på nästa sida

fortsättning uppgift 3b

Fördelningsfunktionen:

$$F(x) = \begin{cases} 0 & \text{för } x \leq -1 \\ 0.75(-x^2 + 1) & \text{för } -1 < x < 0 \\ -0.25e^{-6x} + 1 & \text{för } x \geq 0 \end{cases}$$

$$\begin{aligned} \text{c) } P(0.5 < \xi) + P(\xi < -0.5) &= 1 - P(\xi < 0.5) + P(\xi < -0.5) = \\ &= 1 - [-0.25e^{-6x} + 1]_{x=0.5} + [0.75(-x^2 + 1)]_{x=-0.5} = 0.25e^{-3} + 0.5625 \approx 0.5749 \end{aligned}$$

Uppgift 4: ξ_i = vinsten efter ett spel

$\xi = x$	$P(\xi = x)$
-10	0.5
10	0.5

$$E(\xi) = -10 \cdot 0.5 + 10 \cdot 0.5 = 0 \quad \text{Var}(\xi) = (-10)^2 \cdot 0.5 + 10^2 \cdot 0.5 = 100$$

$$\eta = \text{total vinst efter 1000 spel} \quad \text{där} \quad \eta = \xi_1 + \xi_2 + \dots + \xi_{1000}$$

$$E(\eta) = 1000 E(\xi) = 0 \quad \text{Var}(\eta) = 1000 \text{Var}(\xi) = 1000 \cdot 100 = 100\,000$$

$n \geq 30$. använd normalapproximation (centrala gränsvärdessatsen)

$$P(\eta > 100) = 1 - P(\eta < 100) = 1 - P\left(Z < \frac{100 - 0}{\sqrt{100000}}\right) = 1 - P(Z < 0.32) \approx$$

$$\approx 1 - 0.6255 = 0.3745$$

Uppgift 5:

$$\text{a) } \ell_C = \frac{3000 + 3000 + 1942 + 2512}{4} - \frac{1875 + 1635 + 150 + 165}{4} = 1657.25$$

$$\ell_{AC} = \frac{1875 + 150 + 3000 + 2512}{4} - \frac{1635 + 165 + 3000 + 1942}{4} = 198.75$$

$$\ell_{BC} = \frac{1875 + 1635 + 1942 + 2512}{4} - \frac{150 + 165 + 3000 + 3000}{4} = 412.25$$

b) σ är okänd $\Rightarrow$ använd t-fördelningen med $5 \cdot 8 - 8 = 32$ df

Eftersom $N = 40 > 30$ kan man välja om man vill göra beräkningen exakt eller om man vill använda centrala gränsvärdessatsen och använda normalfördelningen.

$$s_p^2 = \frac{1}{5 \cdot 8 - 8} \left[(5-1) \cdot 1541.2^2 + (5-1) \cdot 1333.4^2 + (5-1) \cdot 0^2 + (5-1) \cdot 32.6^2 + \right. \\ \left. + (5-1) \cdot 0^2 + (5-1) \cdot 0^2 + (5-1) \cdot 1346.8^2 + (5-1) \cdot 1090.3^2 \right] = 894617.5113$$

$$0 \pm t \cdot \frac{2 \cdot s_p}{\sqrt{N}} \Rightarrow 0 \pm 2.75 \cdot 2 \sqrt{\frac{894617.5113}{40}} \Rightarrow 0 \pm 822.529$$

c) Faktorerna B och C ger signifikanta effekter

$$\hat{y} = \ell_M + \frac{\ell_B}{2} \cdot x_B + \frac{\ell_C}{2} \cdot x_C \\ \Rightarrow \hat{y} = 1784.875 - \frac{1185.25}{2} \cdot x_B + \frac{1657.25}{2} \cdot x_C$$

$$\text{där } \ell_M = \frac{1875 + 1635 + 150 + 165 + 3000 + 3000 + 1942 + 2512}{8} = 1784.875$$

Uppgift 6:

- a) 5 faktorer, A, B, C, D och E
- b) 8 försök
- c) 2 generatorer, t.ex. $D = AB$ och $E = AC$
- d) 3 definierande relationer. Av ovanstående generatorer erhålls
 - $I_1 = ABD$
 - $I_2 = ACE$
 - $I_3 = I_1 \cdot I_2 = \cancel{ABD} \cdot \cancel{ACE} = BCDE$

Uppgift 7:

- a) Från binomialfördelningsnomogrammet fås ungefär $n = 140$ $c = 3$
- b) $\frac{p_2}{p_1} = \frac{0.05}{0.01} = 5$ Välj provtagningsplan 5 där $2n_1 = n_2$

Tabellen ger $c_1 = 1$ $c_2 = 4$ och $n_1 p = n_1 \cdot 0.01 = 0.77 \Leftrightarrow n_1 = 77$

En lämplig dubbel provtagningsplan blir $n_1 = 77$ $n_2 = 154$ $c_1 = 1$ $c_2 = 4$ $r_1 = r_2 = 5$

Uppgift 8: Kontrolldiagrammets styrgränser blir

$$S_{\bar{o}} = p + 3\sqrt{\frac{p \cdot (1-p)}{4}} = 0.03 + 3\sqrt{\frac{0.03 \cdot 0.97}{400}} = 0.056$$

$$S_u = p - 3\sqrt{\frac{p \cdot (1-p)}{4}} = 0.03 - 3\sqrt{\frac{0.03 \cdot 0.97}{400}} = 0.004$$

a) $ARL = \frac{1}{1-L(p)}$

$$\begin{aligned} L(p) &= P(0.004 < \hat{p} < 0.056) = P(\hat{p} < 0.056) - P(\hat{p} < 0.004) = \\ &= P\left(Z < \frac{0.056 - 0.03}{\sqrt{\frac{0.03 \cdot 0.97}{400}}}\right) - P\left(Z < \frac{0.004 - 0.03}{\sqrt{\frac{0.03 \cdot 0.97}{400}}}\right) = P(Z < 3.05) - P(Z < -3.05) = \\ &= P(Z < 3.00) - (1 - P(Z < 3.00)) = 2P(Z < 3.00) - 1 = 2 \cdot 0.9987 - 1 = 0.9974 \end{aligned}$$

$$ARL = \frac{1}{1-0.9974} = 384.61$$

b) p-värdet ändras till 0.06

$$\begin{aligned} P(0.004 < \hat{p} < 0.056) &= P(\hat{p} < 0.056) - P(\hat{p} < 0.004) = \\ &= P\left(Z < \frac{0.056 - 0.06}{\sqrt{\frac{0.06 \cdot 0.94}{400}}}\right) - P\left(Z < \frac{0.004 - 0.06}{\sqrt{\frac{0.06 \cdot 0.94}{400}}}\right) = P(Z < -0.34) - P(Z < -4.72) = \\ &= 1 - P(Z < 0.34) - 0.0000 = 1 - 0.6331 = 0.3669 \end{aligned}$$

1) $P(\text{upptäckt i första provgruppen efter ändring}) = 0.6331$

2) $P(\text{upptäckt i någon av de tre första provgrupperna efter ändring}) =$
 $P(\text{upptäckt i första provgruppen}) + P(\text{upptäckt i andra provgruppen}) +$
 $P(\text{upptäckt i tredje provgruppen}) =$
 $= 0.6331 + 0.3669 \cdot 0.6331 + 0.3669 \cdot 0.3669 \cdot 0.6331 = 0.9506$