Tillämpad matematisk statistik LMA521
Lösningar Tentamen 20180109

- **Uppgift 1, 3+3 poäng**

(a)

$$E(\xi) = \int_0^2 x(\frac{3}{2}x - \frac{3}{4}x^2)dx = \ldots = \boxed{1}.$$

$$E(\xi^2) = \int_0^2 x^2(\frac{3}{2}x - \frac{3}{4}x^2)dx = \ldots = \frac{6}{5}.$$

$$\sigma = \sqrt{E(\xi^2) - E(\xi)^2} = \sqrt{\frac{6}{5} - 1} = \boxed{\frac{1}{\sqrt{5}} \approx 0.4472}.$$

(b)

$$E(\eta) = 700E(\xi) = 700.$$

$$S(\eta) = 700S(\xi) = \frac{700}{\sqrt{5}}.$$

Låt T vara utgifterna under ett år. Enligt centrala gränsvärdessatsen är T approximativt $N(365 \times 700, \frac{\sqrt{365 \times 700}}{\sqrt{5}}) \approx N(255500, 5980.8)$-fördelad. Nu fås att

$$\begin{aligned} P(T \leq 245000) &= P(\frac{T - 255500}{5980.8} \leq \frac{245000 - 255500}{5980.8}) \\ &\approx \Phi(-1.76) = 1 - \Phi(1.76) \approx 1 - 0.9608 = \boxed{0.0392.} \end{aligned}$$

- **Uppgift 2, 2+2+2 poäng**

Låt ξ vara antalet gröna bollar som dras och η antalet blå bollar som dras. Det gäller att både ξ och η är $Hyp(N = 14, n = 5, p = 1/2)$. Däremot är de förstås inte oberoende.

(a)

$$P(A) = P(\xi = 5) = \frac{\binom{7}{5}\binom{7}{0}}{\binom{14}{5}} = \ldots = \boxed{\frac{3}{286} \approx 0.010490}.$$

(b) Vi beräknar först $P(B)$. Komplementet till B är att $\eta = 5$ eller $\xi = 5$ vilket ger

$$P(B) = 1 - P(\xi = 5) - P(\eta = 5) = 1 - 2P(A) = 0.979021.$$

Ovan användes att η och ξ har samma fördelning. Eftersom A och B är disjunkta fås att

$$P(A \cup B) = P(A) + P(B) = 0.010490 + 0.979021 \approx \boxed{0.9895}.$$

(c) *Lösning via resonemang*: Att betinga på B^c betyder att vi betingar på att vi fick antingen bara gula bollar, eller bara gröna bollar. Eftersom det finns lika många gula som gröna bollar måste alltså

$$P(A|B^c) = \boxed{\frac{1}{2}}.$$

Alternativ lösning: Från deluppgift b vet vi att $P(B^c) = 2P(A)$. Dessutom gäller att $A \cap B^c = A$. Alltså fås att

$$P(A|B^c) = \frac{P(A \cap B^c)}{P(B^c)} = \frac{P(A)}{2P(A)} = \boxed{\frac{1}{2}}.$$

- **Uppgift 3, 4 poäng** Vi får att $\bar{x} = 34.7$ och $s = 3.90299$. Antal frihetsgrader är $4 - 1 = 3$. Från t-tabellen får vi med $\alpha = 0.05$ att $t(3) = 3.18$ (se rad 3, kolumn 0.05). Intervallet blir alltså

$$34.7 \pm \frac{3.18 \times 3.90299}{\sqrt{4}} = \boxed{34.7 \pm 6.2058} = \boxed{[28.4942, 40.9058]}.$$

- **Uppgift 4, 7 poäng**

  Låt A beteckna händelsen att man går till sjö A, och låt B beteckna händelsen att man går till sjö B. Låt ξ beteckna antalet fångade fiskar. Låt C beteckna händelsen $\{\xi \geq 3\}$. Vi söker $P(A|C)$, och vi kommer att använda Bayes sats. Om A inträffar blir ξ en Po(5)-fördelad stokastisk variabel och om B inträffar blir ξ en $Po(4)$-fördelad stokastisk variabel. Vi beräknar först $P(A)$ och $P(B)$. Om η betecknar tiden man väntar på bussen blir

$$P(A) = P(\eta \leq 9) = \int_0^9 \frac{1}{12} dx = 0.75.$$

  Eftersom han går till B om han inte går till A blir

$$P(B) = 1 - P(A) = 0.25.$$

  Enligt ovan har vi att

$$\begin{aligned} P(C|A) = P(\xi \geq 3|A) = 1 - P(\xi \leq 2|A) &= 1 - \left(\frac{e^{-5} \times 5^0}{0!} + \frac{e^{-5} \times 5^1}{1!} + \frac{e^{-5} \times 5^2}{2!} \right) \\ &= 1 - 0.124652 = 0.875348. \end{aligned}$$

  På liknande sätt fås att

$$P(C|B) = 1 - \left(\frac{e^{-4} \times 4^0}{0!} + \frac{e^{-4} \times 4^1}{1!} + \frac{e^{-4} \times 4^2}{2!} \right) = 0.761897.$$

  Nu ger Bayes sats att

$$\begin{aligned} P(A|C) &= \frac{P(C|A)P(A)}{P(C|A)P(A) + P(C|B)P(B)} \\ &\approx \frac{0.875348 \times 0.75}{0.875348 \times 0.75 + 0.761897 \times 0.25} \approx \boxed{0.7751}. \end{aligned} \quad (1)$$

- **Uppgift 5, 2+1+3+1 poäng**

  (a)

  $$l_{AB} = (42+52+41+55+41+51+43+59-54-43-56-44-53-44-57-48)/8 = -1.875$$

  (b) l_{AB} hamnar innanför intervallet, och är således inte signifikant på nivå 0.01.

  (c) Från generatorerna får vi att $I_1 = ABCDE$, $I_2 = ABCF$ och $I_3 = BCDG$. Sedan får vi att $I_4 = I_1 I_2 = ABCDEABCF = DEF$, $I_5 = I_1 I_3 = ABCDEBCDG = AEG$, $I_6 = I_2 I_3 = ABCFBCDG = ADFG$ och $I_7 = I_1 I_2 I_3 = ABCDEABCFBCDG = BCEFG$. Det kortaste ordet har längd 3, och upplösningen är därför 3 (III). Vi ser till exempel att $E = EI_4 = EDEF = DF$, så E kommer att blandas ihop med (bland annat) två-faktorsamspelet DF i detta reducerade faktorförsök.

  (d) Antalet tre-faktorsamspel ges av $\binom{8}{3} = \ldots = 56$.

- **Uppgift 6, 4+5 poäng**

  (a) Sannolikheten att en given lampa lyser efter 105 timmar ges av

  $$\int_{105}^{\infty} \frac{1}{100} e^{-x/100} = \ldots = e^{-1.05}.$$

  Låt η vara antalet lysande lampor efter 105 timmar. Då är η, eftersom lamporna antas oberoende, $Bin(n = 100, p = e^{-1.05})$. Eftersom $np(1-p) = 22.7481 > 10$ så blir η approximativt $N(np, \sqrt{np(1-p)}) = N(34.99, 4.7694)$. Vi får att

  $$\begin{aligned} P(\eta > 40) = 1 - P(\eta \leq 40) = 1 - P\left(\frac{\eta - 34.99}{4.7694} \leq \frac{40 - 34.99}{4.7694}\right) \\ \approx 1 - \Phi(1.05) = 1 - 0.8531 = \boxed{0.1469.} \end{aligned}$$

  (b) Vi skall beräkna $F(x) = P(\xi \leq x)$ för alla x. Om $x < 0$ blir förstås $F(x) = 0$. Så antag nu att $x \geq 0$. Låt η_x vara antalet lampor som gått sönder vid tiden x. Observera att

  $$F(x) = P(\xi \leq x) = P(\eta_x \geq 2).$$

  Sannolikheten att en given lampa gått sönder vid tiden x ges av

  $$\int_0^x \frac{1}{100} e^{-t/100} = 1 - e^{-x/100}.$$

  Eftersom lamporna är oberoende av varandra är η_x $Bin(n = 5, p = 1 - e^{-x/100})$-fördelad. Alltså blir,

  $$\begin{aligned} P(\eta_x \geq 2) = 1 - P(\eta_x < 2) = 1 - P(\eta_x = 0) - P(\eta_x = 1) \\ = 1 - \binom{5}{0}(1 - e^{-x/100})^0 (e^{-x/100})^5 - \binom{5}{1}(1 - e^{-x/100})^1 (e^{-x/100})^4 \\ = \ldots = \boxed{1 + 4e^{-x/20} - 5e^{-x/25}}. \end{aligned}$$

- **Uppgift 7, 3 poäng**

  Från mätvärdena beräknar vi $\bar{\bar{x}} = 6.32$ och $\bar{R} = 0.12$. Från tabellen rad 10 (eftersom provgruppsstorleken=10) hittar vi konstanterna $A_2 = 0.308$, $D_3 = 0.223$ och $D_4 = 1.777$. Gränserna för $\bar{x}$-diagrammet blir $\bar{\bar{x}} \pm A_2\bar{R} = 6.32 \pm 0.308 \times 0.12 = 6.32 \pm 0.03696 = [6.28304, 6.35696]$. Undre gräns R-diagrammet blir $D_3\bar{R} = 0.223 \times 0.12 = 0.02676$ och övre gräns blir $D_4\bar{R} = 1.777 \times 0.12 = 0.21324$. Eftersom till exempel provgruppsmedelvärdet i provgrupp 1 ligger ovanför den övre styrgränsen i $\bar{x}$-diagrammet så drar vi slutsatsen att processen ej är i statistisk kontroll.

- **Uppgift 8, 8 poäng**

  (a) Låt d_1 vara antalet defekta i urval 1 och d_2 antalet defekta i urval 2. Eftersom urvalens storlekar är tillräckligt små jämfört med partiets storlek $((50+50)/50000 = 0.002 < 0.1))$ så är d_1 och d_2 ungefärligen oberoende och de är båda approximativt $Bin(n = 50, p = 0.1)$.

  $$A_1 = P(\text{acceptera efter urval 1}) = P(d_1 = 0) + P(d_1 = 1) + P(d_1 = 2)$$
  $$= \binom{50}{0} 0.1^0 0.9^{50} + \binom{50}{1} 0.1^1 0.9^{49} + \binom{50}{2} 0.1^2 0.9^{48} = 0.111729.$$

  $$A_2 = P(\text{acceptera efter urval 2})$$
  $$= P(d_1 = 3)P(d_2 = 0) + P(d_1 = 3)P(d_2 = 1) + P(d_1 = 4)P(d_2 = 0)$$
  $$= \binom{50}{3} 0.1^3 0.9^{47} \binom{50}{0} 0.1^0 0.9^{50} + \binom{50}{3} 0.1^3 0.9^{47} \binom{50}{1} 0.1^1 0.9^{49} + \binom{50}{4} 0.1^4 0.9^{46} \binom{50}{0} 0.1^0 0.9^{50}$$
  $$= 0.00561388.$$

  Alltså blir $ATI(0.1) = 50A_1 + 100A_2 + 50000(1 - A_1 - A_2) = \boxed{44139}$.

  (b) Låt B vara händelsen att partiet accepteras. Vi söker $P(B|A)$. Enligt definition av betingad sannolikhet har vi

  $$P(B|A) = \frac{P(A \cap B)}{P(A)}.$$

  Händelsen $A \cap B$ är händelsen att partiet accepteras och att $d_1 \geq 1$. När vi beräknade A_1 och A_2 ovan har vi beräknat sannolikheterna för dessa utfall, förutom att vi där också hade med $P(d_1 = 0)$. Så

  $$P(A \cap B) = A_1 + A_2 - P(d_1 = 0)$$
  $$= 0.111729 + 0.00561388 - \binom{50}{0} 0.1^0 0.9^{50} = 0.112189. \quad (2)$$

  Vi har att $P(A) = 1 - P(d_1 = 0) = 1 - \binom{50}{0} 0.1^0 0.9^{50} = 0.994846$, vilket ger

  $$P(B|A) = \frac{0.1121189}{0.994846} \approx \boxed{0.1127}.$$