Matematisk statistik LKT325
Lösningar Tentamen 2017-04-12

1. (a) Låt μ vara det riktiga medelvärdet på tidsåtgången för Mias algoritm.

   Steg 1 $H_0 : \mu \geq 9.1$, $H_1 : \mu < 9.1$.

   Steg 2 Bestäm signifikansnivå, t ex $\alpha = 0.05$.

   Steg 3 Vi får ingen information om variansen, alltså måste vi göra ett t-test. Testvariabeln vi använder är således

   $$T = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}.$$

   Steg 4 I vårt fall är $\mu_0 = 9.1$, $\bar{x} = 7.99$, $s = 1.297391$ och $n = 10$. Vårt värde på testvariabeln blir då tillslut $T = -2.7055$.

   Steg 5 Det kritiska värdena tas fram genom att titta i t-tabellen med 9 frihetsgrader ($\alpha = 0.05$): $t_c = -1.83$. Eftersom vår observerade testvariabel $T = -2.7055$ är mindre än de kritiska värdet så kan inte förkasta H_0 på signifikansnivå $\alpha = 0.05$.

   Slutsats: Vi kan dra slutsatsen att Mias algoritm i genomsnitt är snabbare än den gamla för att att lösa detta problem.

   (b) Eftersom t-testet bygger på att datan är normalfördelad så måste vi anta att den är det.

2. Låt $A = \{\text{skyldig}\}$ och $B = \{\text{döms}\}$. Vi vet att $P(A) = 0.7$, $P(B|A) = 0.59$, $P(B|A^c) = 0.004$. Vi söker $P(A^c|B)$. Bayes sats ger att

$$\begin{aligned} P(A^c|B) = \frac{P(B|A^c)P(A^c)}{P(B)} &= \frac{P(B|A^c)P(A^c)}{P(B|A)P(A) + P(B|A^c)P(A^c)} \\ &= \frac{0.004 \times 0.3}{0.59 \times 0.7 + 0.004 \times 0.3} \approx \boxed{0.0029.} \end{aligned}$$

3. (a) Vi får att $\bar{x} = 23.8$, $s = 1.9975$, antalet frihetsgrader $n - 1 = 2$, $\alpha = 0.05$ samt att $t_{0.025}(2) = 4.3$ (enligt rad 2, kolumn 0.05 i t-tabellen). Konfidensintervallet blir

   $$\bar{x} \pm \frac{t_{0.025}(2)s}{\sqrt{3}} = \boxed{23.8 \pm 4.959.}$$

   (b) χ^2-tabellerna (rad 2, kolumn 0.025 respektive 0.975) ger $\chi^2_{0.025} = 7.378$ och $\chi^2_{0.975} = 0.0506$. Intervallet blir

   $$\left[\sqrt{\frac{2 \times 1.9975^2}{7.378}}, \sqrt{\frac{2 \times 1.9975^2}{0.0506}}\right] = \boxed{[1.04, 12.5582]}.$$

4. (a) Additions-satsen och oberoende ger att

   $$\begin{aligned} &P(A \cup B) = P(A) + P(B) - P(A \cap B) \\ &\quad = P(A) + P(B) - P(A)P(B) = 0.4 + 0.5 - 0.4 \times 0.5 = \boxed{0.7} \end{aligned}$$

(b) Eftersom A och C är disjunkta inser man (rita Venn-diagram!) att

$$A \cup B \cup C = (A \cup B) \cup (B^c \cap C).$$

Eftersom $A \cup B$ och $B^c \cap C$ är disjunkta får vi att

$$P(A \cup B \cup C) = P(A \cup B) + P(B^c \cap C),$$

vilket ger

$$P(B^c \cap C) = P(A \cup B \cup C) - P(A \cup B) = 0.9 - 0.7 = \boxed{0.2}$$

5. Den färdiga ANOVA-tabellen bör se ut såhär:

Variationskälla	SS	df	MS	F
Mellan grupper	0.52532	3	0.175108	8.734
Inom grupper	0.24059	12	0.020049	
Totalt	0.76591	15		

Vi använder de fem stegen.

Steg 1. Låt μ_1, μ_2, μ_3 och μ_4 beteckna de genomsnittliga livslängderna för de tre modellerna. Våra hypoteser blir $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ och $H_1 : \exists i, j : \mu_i \neq \mu_j$.

Steg 2. Vi väljer signifikansnivå $\alpha = 0.05$. (Man är fri att välja signifikansnivå men det är bara 0.05 som finns i tabellen så om man väljer någon annan nivå får man problem.)

Steg 3. Testvariabeln vi vill använda är

$$F = \frac{MSB}{MSE} = \frac{(SSB)/(k-1)}{(SSE)/(n-k)}$$

som är F-fördelad med 3 och 12 frihetsgrader.

Steg 4. Vi vill ta fram vårt observerade värde på F. För att göra det så fyller vi i ANOVA-tabellen succesivt. Vi vet att $SST = SSB + SSE$ så vi vet att $SSE = .76591 - .52532 = .24059$. Eftersom det är fyre grupper och fyra observationer i varje grupp blir frihetsgraderna 3 och 12. Vi vet också att $MSB = SSB/df = .175108$ och att $MSE = SSE/df = .020049$. Tillslut får vi att $F = MSB/MSE = 8.734$.

Steg 5. Vi tittar i F-fördelningstabellen med 3 och 12 frihetsgrader och ser att det kritiska värdet är 3.49. Eftersom vårt observerade värde är större än det kritiska värdet så förkastar vi H_0 på signifikansnivå $\alpha = 0.05$.

Slutsats: Kaffet i de fyra kaféerna har i genomsnitt olika pH-värden.

6. (a)

$$E(\xi) = \int_0^1 x(6x - 6x^2)dx = ... = \boxed{1/2.}$$

$$E(\xi^2) = \int_0^1 x^2(6x - 6x^2)dx = ... = 3/10.$$

$$S(\xi) = \sqrt{\mathrm{Var}(\xi)} = \sqrt{3/10 - (1/2)^2} \approx \boxed{0.2236.}$$

(b)

$$E(\eta) = 9.75 + 0.5E(\xi) = \boxed{10.}$$

$$S(\eta) = 0.5S(\xi) = \boxed{0.1118}.$$

(c)

$$P(\eta \leq 10.1) = P(9.75 + 0.5\xi \leq 10.1) = P(\xi \leq 0.7)$$

$$\int_0^{0.7} f(x)dx = ... = \boxed{0.784.}$$

7. (a)

$$l_A = \frac{54 + 78 + 55 + 79 - 53 - 77 - 53 - 77}{4} = \boxed{1.5}$$

$$l_{ABC} = \frac{54 + 77 + 53 + 79 - 53 - 78 - 55 - 77}{4} = \boxed{0}$$

(b) $I_1 = ABCD$, $I_2 = BCE$, $I_3 = ACF$, $I_4 = I_1I_2 = ADE$, $I_5 = I_1I_3 = BDF$, $I_6 = I_2I_3 = ABEF$, $I_7 = I_1I_2I_3 = CDEF$. Eftersom kortaste order har längd 3 blir upplösningen $\boxed{3}$. Alias till A blir $\boxed{BCD, ABCE, CF, DE, ABDF, BEF, ACDEF}$.

8. (a) Låt ξ =restiden.

$$\begin{aligned} P(\text{kommer i tid}) &= P(\xi \leq 25) \\ &= P\left(\frac{\xi - 21}{3} \leq \frac{25 - 21}{3}\right) = \Phi(4/3) = \boxed{0.9082}. \end{aligned}$$

(b) Låt η = antalet dagar hon kommer i tid. Då är η Bin(n=200,p=0.9082). Eftersom $np(1-p) = 16.6746 > 10$ är η approximativt $N(np, \sqrt{np(1-p)}) = N(181.64, 4.08345)$. Vi får att

$$P(\eta \leq 190) = P\left(\frac{\eta - 181.64}{4.0835} \leq \frac{190 - 181.64}{4.0835}\right) = \Phi(2.05) = \boxed{0.9798}.$$

9. Laat $\gamma = \xi\eta$. Vi ser att γ är en diskret stokastisk variabel som kan anta värdena 0 eller 1, och att $\gamma = 1$ om och endast om $\xi = \eta = 1$. Vi får att

$$P(\gamma = 1) = P(\xi = \eta = 1) = P(\xi = 1)P(\eta = 1) = 0.5 \times 0.5 = 0.25$$

eftersom ξ och η är oberoende. Alltså blir $P(\gamma = 0) = 1 - P(\gamma = 1) = 0.75$. Till slut får vi

$$\begin{aligned} E(2^{\xi\eta}) = E(2^{\gamma}) &= 2^0 \times P(\gamma = 0) + 2^1 \times P(\gamma = 1) \\ &= 1 \times 0.75 + 2 \times 0.25 = \boxed{1.25.} \end{aligned}$$

Lycka till!