資料分析方法-HW1

經研一 江彥亨 R11323040

WAY TO WIN

1. Let X be the numbers of correct answers, then $X \sim \text{Binomial}(5, \frac{1}{4})$

$\text{Prob}(X \geq 4) = \text{Prob}(X=4) + \text{Prob}(X=5)$

$= C^5_4 \cdot (\frac{1}{4})^4 (\frac{3}{4}) + C^5_5 (\frac{1}{4})^5 = \frac{1}{64}$ #

2. Let X be the numbers of matchs.

$E(X) = 4 \times [P^4 + (1-P)^4] + 5 \times [P \cdot C^4_3 \cdot P^3 \cdot (1-P)$

$+ (1-P) \cdot C^4_1 \cdot P^1 \cdot (1-P)^3] + 6 \times [P \cdot C^5_3 \cdot P^3 \cdot (1-P)^2$

$+ (1-P) \cdot C^5_2 \cdot P^2 \cdot (1-P)^3] + 7 \times [P \cdot C^6_3 \cdot P^3 \cdot (1-P)^3$

$+ (1-P) \cdot C^6_3 \cdot P^3 \cdot (1-P)^3] = \sum_{X=4}^{7} X \times [C^{X-1}_{X-4} \times (P^4 (1-P)^{X-4} + P^{X-4}(1-P)^4]$

$P = \frac{1}{2} \Rightarrow E(X) = 4 \times \frac{1}{8} + 5 \times \frac{1}{4} + 6 \times \frac{5}{16} + 7 \times \frac{5}{16}$

$= \frac{93}{16} = 5.8125$ #

3. Let x be the time of arrival at the place, then $X \sim \text{Uniform}(0, 80)$.

$\text{Prob}(\text{I witness the show}) = \text{Prob}(X > 60)$

$= \frac{80-60}{80} = \frac{1}{4}$

4. a. Let $C_1 \sim \exp(\beta = \frac{1}{\mu_1})$ $C_2 \sim \exp(\beta = \frac{1}{\mu_2})$

$\text{Prob}(\text{Mary sees John}) = \text{Prob}(C_1 > 10)$

$= 1 - \text{Prob}(C_1 \leq 10) = 1 - 1 - e^{-10\mu_1} = e^{-10\mu_1}$

b. $\text{Prob}(\text{Mary finishes her service before John})$

$= \text{Prob}(C_1 > C_2 + 10 \mid C_1 > 10) (= \text{Prob}(C_1 > C_2)$

$= \int_0^\infty \int_{C_2}^\infty \mu_1 \mu_2 e^{-(\mu_1 C_1 + \mu_2 C_2)} dC_1 dC_2$

$= \mu_1 \mu_2 \int_0^\infty e^{-\mu_2 C_2} \int_{C_2}^\infty e^{-\mu_1 C_1} dC_1 dC_2$

$= \mu_1 \mu_2 \int_0^\infty e^{-\mu_2 C_2} \cdot (\frac{1}{\mu_1} e^{-\mu_1 C_2}) dC_2 = \mu_2 \int_0^\infty e^{-(\mu_1+\mu_2)C_2} dC_2$

$= \frac{\mu_2}{\mu_1 + \mu_2}$ #

5. Let X be the train is of a delay and Y be the bus departure time.

Prob(John will be late)

$= \text{Prob}(Y > 0 \cup X = 12) + \text{Prob}(Y < 0 \cap X = 10)$

$+ \text{Prob}(Y < -2 \cap X = 8) + \text{Prob}(Y < -4 \cap X = 6)$

$+ \text{Prob}(Y < -6 \cap X = 4)$

$= (\frac{1}{2} + \frac{1}{16} - \frac{1}{2} \times \frac{1}{16}) + (\frac{1}{2} \times \frac{1}{4}) + (\frac{0.32}{2} \times \frac{1}{2}) + (\frac{0.05}{2} \times \frac{1}{8})$

$+ (\frac{0.003}{2} \times \frac{1}{16}) \approx 0.74$ #

6. a. Prob(John really has the cancer | positive)

$$= \frac{\text{Prob(John has the cancer} \cap \text{positive)}}{\text{Prob(positive)}}$$

$$= \frac{0.01 \times 0.99}{0.01 \times 0.99 + 0.99 \times 0.01} = \frac{1}{2} \ \#$$

b. Prob(John really has the cancer | (positive, positive))

$$= \frac{\text{Prob(John has the cancer} \cap \text{(positive, positive))}}{\text{Prob((positive, positive))}}$$

$$= \frac{0.01 \times 0.99 \times 0.999}{0.01 \times 0.99 \times 0.999 + 0.99 \times 0.01 \times 0.001} = 0.999 \ \#$$

7. a. Let $X_i \overset{iid}{\sim} (\mu = 8.5, \ \sigma = 3.5)$. Since sample size is big enough, by Central Limit Theorem:

$$\sqrt{49}(\bar{X} - 8.5) \xrightarrow{d} N(0, 3.5^2) \quad \text{where } \bar{X} = \frac{\sum_{i=1}^{49} X_i}{49}$$

$$\text{Prob}(\bar{X} < 10) \approx \text{Prob}\left(\sqrt{49} \cdot \frac{\bar{X} - 8.5}{3.5} < \sqrt{49} \cdot \frac{10 - 8.5}{3.5}\right)$$

$$= \text{Prob}(Z < 3) \approx 0.999$$

b. $\text{Prob}(7 < \bar{X} < 10) \approx \text{Prob}(-3 < Z < 3)$

≈ 0.997

c. $\text{Prob}(\bar{X} < 7.5) \approx \text{Prob}(Z < -2) \approx 0.0228$

8. Let P_i be 1 if the person use iphone,

by Central Limit Theorem:

$$\sqrt{600}(\bar{P} - 0.4) \xrightarrow{d} N(0, 0.24) \quad \text{where } \bar{P} = \frac{\sum_{i=1}^{600} P_i}{600}$$

$H_0: P = 0.4$ against $H_1: P \neq 0.4$

Type I error probability $= \text{Prob}\left(\bar{P} < \frac{216}{600} \text{ or } \bar{P} > \frac{264}{600} \mid P = 0.4\right)$

$$= \text{Prob}\left(\sqrt{600}\frac{(\bar{P} - 0.4)}{\sqrt{0.24}} < \sqrt{600}\frac{(0.36 - 0.4)}{\sqrt{0.24}}\right)$$

$$+ \text{Prob}\left(\sqrt{600}\frac{(\bar{P} - 0.4)}{\sqrt{0.24}} > \sqrt{600}\frac{(0.44 - 0.4)}{\sqrt{0.24}}\right)$$

$$= \text{Prob}(Z < -2) + \text{Prob}(Z > 2) \approx 0.0456$$

9. Yes, we can assume that $X_i \overset{\text{iid}}{\sim} \text{Binomial}(24, P)$

where P is the probability that beer bottle is under-filled.

Thus, by Central Limit Theorem:

$$\sqrt{75}\frac{(\bar{X} - 24P)}{\sqrt{24P(1-P)}} \xrightarrow{d} N(0, 1) \quad \bar{X} = \frac{\sum_{i=1}^{75} X_i}{75} = \frac{3}{4}$$

By CMT and slutsky theorem:

$$\sqrt{75}\frac{\bar{X} - 24P}{\sqrt{24\hat{P}(1-\hat{P})}} \xrightarrow{d} N(0, 1) \quad \hat{P} = \frac{\bar{X}}{24} = \frac{1}{32}$$

Then, we can construct the confidence interval for "P"

$$1 - \alpha = 0.95 = P\left(\bar{X} - 1.96 \cdot \sqrt{\frac{24 \cdot \hat{P}(1-\hat{P})}{75}} \le P \le \bar{X} + 1.96 \cdot \sqrt{\frac{24\hat{P}(1-\hat{P})}{75}}\right)$$

$\Rightarrow$ Confidence interval $= [\, 0.5571,\ 0.9429 \,]$

10. H_0 : Prob and OR are independent
H_1 : Prob and OR are dependent

$$\psi = \sum_{i=1}^{3} \sum_{j=1}^{3} \frac{(O_{ij} - e_{ij})^2}{e_{ij}} \sim \chi^2(4)$$

$$\psi^* = \frac{(24-15.75)^2}{15.75} + \frac{(11-13.5)^2}{13.5} + \frac{(10-15.75)^2}{15.75} + \frac{(7-8.75)^2}{8.75} + \frac{(13-7.5)^2}{7.5}$$
$$+ \frac{(5-8.75)^2}{8.75} + \frac{(4-10.5)^2}{10.5} + \frac{(6-9)^2}{9} + \frac{(20-10.5)^2}{10.5} \approx 26.4931$$

$\alpha = 0.01$ $\quad RR = \{\psi \mid \psi > \chi^2_{0.01}(4) = 13.2767\}$

$\psi^* \in RR \Rightarrow$ reject H_0, Prob and OR are dependent