Assignment

Roll No.: CS23B1047
Name : Dhage Pratik Bhishmacharya
course : Probability and statistics

Q.1. Taking $l = 6$.

Solution: surgery certainity $= (l+3)\% 10 \times 10$
$= 9 \% 10 \times 10$
$= 90\%$

Initial disease certainity $= (l+1)\% 10 \times 10$
$= 7 \% 10 \times 10$
$= 70\%$

Events : D = disease
T = Test A positive
B = diabetic

$P(D) = 0.7 \qquad P(D^c) = 1 - 0.7 = 0.3$

$P(T|D,B) = 1 \qquad P(T|D,B^c) = 1$

$P(T|D^c,B^c) = 0 \qquad P(T|D^c,B) = 0.3$

$$P(D|T,B) = \frac{P(T|D,B)\,P(D|B)}{P(T|B)}$$

$$= \frac{P(T|D,B)\,P(D|B)}{P(T|D,B)P(D|B) + P(T|D^c,B)P(D^c|B)}$$

$$= \frac{1 \times P(D)}{1 \times P(D) + 0.3 \times 0.3}$$

$$= \frac{0.7}{0.79}$$

$= 0.886 < 0.9$ $\therefore$ Recommend Additional tests

Q.2. Taking $\ell = 7$.

Solution: X : No. of students on the bus carrying randomly selected student

Y : No. of students on the bus carrying randomly selected driver

(1) Which of $E[\ell X]$ or $E[\ell Y]$ is larger? Why?

selecting a student gives probability proportional to bus size $\left(\frac{K}{148}\right)$, favouring larger bus size.

selecting a driver gives equal probability $\left(\frac{1}{4}\right)$ for each bus, regardless of size.

Thus, $E[X] > E[Y]$

We already know, $E[\ell X] = \ell E[X]$ and $E[\ell Y] = \ell E[Y]$

and Hence $E[\ell X] > E[Y]$.

(2) compute $E[\ell X]$ and $E[\ell Y]$

$$P(X=K) = \frac{K}{148}, \text{ for } K = 40, 33, 25, 50$$

$$E[X] = \sum K \cdot \frac{K}{148} = \frac{40^2 + 30^2 + 25^2 + 50^2}{148} = \frac{2907}{74}$$

$$E[\ell X] = E[7X] = \frac{7 \cdot 2907}{74} = \frac{20349}{74} = 274.98$$

$$P(Y=K) = \frac{1}{4}, \text{ for } K = 40, 33, 25, 50$$

$$E[Y] = \frac{40 + 33 + 25 + 50}{4} = \frac{148}{4} = 37$$

$$E[\ell Y] = E[7Y] = 7 \cdot 37 = 259.$$

Q.3. Taking $\ell = 7$.

Solution: $\text{dist}(A,B) = 7 \times 100 = 700$ miles

$X \sim \text{Unif}(0, 700)$

current service centers: A - 0 miles, center - 350 miles
B - 700 miles

proposed service centers:

$2l = 2\times7 = 14$ miles   midpoints $= \frac{14+35}{2} = 24.5$

$5l = 5\times7 = 35$ miles

$7l = 7\times7 = 49$ miles   midpoint $= \frac{35+49}{2} = 42$

current scenario: stations: $\{0, 350, 700\}$

$[0,175)$: Nearest 0, distance $= x$.

$[175,350)$: Nearest 350, distance $= 350-x$

$[350,525)$: Nearest 350, distance $= x-350$

$[525,700]$: Nearest 700, distance $= 700-x$.

Density: $f(x) = \frac{1}{700}$ ($\because$ Uniform).

Each interval length $= 175$.

$$E[D] = 4\cdot\int_0^{175} x\cdot\frac{1}{700}\,dx = 4\cdot\frac{1}{700}\cdot\frac{175^2}{2} = 87.5 \text{ miles}$$

proposed scenario: stations: $\{14, 35, 49\}$

$[0,24.5)$: Nearest 14, distance $= |x-14|$

$[24.5,42)$: Nearest 35, distance $= |x-35|$

$[42,700]$: Nearest 49, distance $= |x-49|$

$[0,14)$: distance $= 14-x$, length $= 14-0 = 14$

$$\int_0^{14}(14-x)\frac{1}{700}\,dx = \frac{14^2}{1400} = 0.14$$

$[14,24.5)$: distance $= x-14$, length $= 24.5-14 = 10.5$

$$\int_{14}^{24.5}(x-14)\frac{1}{700}\,dx = \frac{10.5^2}{1400} = 0.078$$

$[24.5,35)$: distance $= 35-x$, length $= 35-24.5 = 10.5$

$$\int_{24.5}^{35}(35-x)\frac{1}{700}\,dx = \frac{10.5^2}{1400} = 0.078$$

$[35,42)$: distance $= x-35$, length $= 42-35 = 7$

$$\int_{35}^{42} (x-35)\frac{1}{700}\,dx = \frac{7^2}{1400} = 0.035$$

$[42,49)$: distance $= 49-x$, length $= 49-42 = 7$

$$\int_{42}^{49} (49-x)\frac{1}{700}\,dx = \frac{7^2}{1400} = 0.035$$

$[49,700]$: distance $= x-49$, length $= 700-49 = 651$

$$\int_{49}^{700} (x-49)\frac{1}{700}\,dx = \frac{1}{700}\cdot\frac{651^2}{2} = 303.42$$

Total:

$E[D] = 0.14 + 0.078 + 0.078 + 0.035 + 0.035 + 303.42$

$= 303.76$ miles

current: $E[D] = 87.5$ miles

proposed: $E[D] = 303.76$ miles

$\therefore$ The proposed configuration is <u>less</u> efficient, as it significantly increases the expected distance to the nearest station.

Q.4. Taking $l = 7$.

<u>Solution:</u> No. of policy holders. $(n) = 70{,}000$

$E[X_i] = 240$ rupees

$SD[X_i] = 800$ rupees

$Var[X_i] = 800^2 = 640{,}000$ rupees

Total claim: $S = X_1 + X_2 + \ldots + X_{70{,}000}$

since n is large, use central limit Theorem

Mean: $E[S] = n \cdot E[X_i] = 70{,}000 \times 240 = 16{,}800{,}000$ rupees

Variance: $Var(S) = n \cdot Var(X_i) = 70{,}000 \times 640{,}000 = 44{,}800{,}000{,}000$

Standard deviation: $SD(S) = \sqrt{Var(S)} = 211{,}661.94$

By CLT,

$$P(S > 2{,}700{,}000) = P\left(\frac{S-E[S]}{SD(S)} > \frac{2{,}700{,}000 - 16{,}800{,}000}{211{,}661.94}\right)$$

$$= P\left(\frac{S-E[S]}{SD(S)} > -66.614\right)$$

$$= P(z > -66.614)$$

$$= 1 - \phi(-66.614)$$

$$= \phi(66.614)$$

$$\approx 1$$

∴ The probability ~~of~~ that the total yearly claim exceeds 2.7 million rupees is approx. 1

Q.5.

(I) Taking $l = 6$.

<u>solution:</u>

(a) Null Hypothesis $H_0 : p = 0.9$

(b) Alternative Hypothesis $H_1 : p \neq 0.9$ (two tailed)

(c) Fix Z_α. significance $= 5\% = 0.05$

$$Z_\alpha = \pm 1.96$$

(d) Test statistics :

$$\hat{p} = \frac{112}{200} = 0.56, \quad \hat{Q} = 1 - \hat{P} = 0.44$$

$$nP_0 = 200 \times 0.9 = 180$$

$$z = \frac{\hat{p} - P_0}{\sqrt{\frac{P_0(1-P_0)}{n}}} = \frac{0.56 - 0.9}{\sqrt{\frac{0.9(0.1)}{200}}} = \frac{\cancel{-0.03}\ -0.34}{0.02121}$$

$$z \approx -16.03$$

(e) conclusion : $|z| = 16.03 > 1.96$ ∴ Reject H_0

The proportion quitting is significancly different from 90%.

(ii) Taking $l = 7$.

Solution: $(l+2)8\% = (7+2)8\% = 98\%$

(a) Null Hypothesis $H_0: p \geq 0.08$

(b) Alternative Hypothesis: $H_1: p < 0.98$ (one-tailed)

(c) fix Z_α

(i) $\alpha = 0.05$, critical value: $Z_{0.05} = -1.645$

(ii) $\alpha = 0.01$, critical value: $Z_{0.01} = -2.326$

(d) Test statistics

$$\hat{p} = \frac{470}{500} = 0.94$$

$$z = \frac{\hat{p} - P_0}{\sqrt{\frac{P_0(1-P_0)}{n}}} = \frac{0.94 - 0.98}{\sqrt{\frac{0.98 \times 0.02}{500}}} = -6.39$$

(e) conclusion.

(i) $\alpha = 0.05$

$z = -6.39 < -1.645$; Reject H_0

(ii) $\alpha = 0.01$

$z = -6.39 < -2.326$; Reject H_0

$\therefore$ The proportion of conforming pipes is significantly less than 98%