QM1-B Assignment-4

Group-14
Bir Anmol Singh (PGPGC2022004~~13~~)
J. Shravani (PGPGC202200424)
Rohit Venkatesh (PGPGC202200

1) From the data file given,

$\overline{X_n} = 2.956637$ (using R & excel funct.)

Sample std $S = 1.053255$ (using R & excel stdev.s( ))

$n = 50$

Cofidence interval = 95%

$\Rightarrow 100(1-\alpha) = 95$

$\Rightarrow \alpha = 0.05 \Rightarrow \alpha/2 = 0.025$

Using excel $Z_{\alpha/2}$ = Norminv(1-α/2,0,1)

$Z_{\alpha/2} = 1.96 \quad (P(Z > Z_{\alpha/2}) = 0.025)$

$\Rightarrow CI = \left(\bar{X} - Z_{\alpha/2}\frac{S}{\sqrt{n}},\ \bar{X} + Z_{\alpha/2}\frac{S}{\sqrt{n}}\right)$

$CI = (2.665, 3.249)$

2) Using R

Sample median = 0.3694

Confidence Interval (95% Bootstrap Pivotal) = (0.2452, 0.4664)

3)

a) Using R at $\alpha = 0.05$

Sample skewness = 2.26

95% bootstrap Percentile CI for sample skewness = (0.950, 3.103)

b) H_0: Skewness = 0
H_a: Skewness ≠ 0 At 5% significance level

Since $0 \notin (0.95, 3.103) \Rightarrow H_0$ rejected $\Rightarrow H_a$ accepted
or ∴ Skewness ≠ 0

c) H_0 rejected in test at 5% level of significance
∴ there is only 5% chance that we rejected H_0 when the skewness was actually 0
OR there is only 5% chance that sample was from symmetric distribution.

Q4)

$P(X \le 10)$, $X \sim RV$

$$Z - \begin{cases} z_i = 1 , \text{ if } x_i \le 10 \\ z_i = 0 , \text{ if } x > 10 \end{cases}$$

Sample $\bar{p} = P(X \le 10) = (\text{Proportion of } z_i = 1) = 0.594$

Sample $s = \sqrt{\frac{n}{n-1} \cdot p \cdot (1-p)} = 0.491$

At 95% CI , $\alpha = 0.05$

$z_{\alpha/2} = 1.96$ $\quad (P(z > z_{\alpha/2}) = \alpha/2)$

$$CI = \left(\bar{p} - z_{\alpha/2} \frac{s}{\sqrt{n}} , \bar{p} + z_{\alpha/2} \frac{s}{\sqrt{n}} \right)$$

$$CI = (0.577, 0.612)$$

Each z_i is RV with $E[x_i] = p$ & $Var(x_i) = s^2$.
z_i are independent & identically distributed
∴ $\bar{p} = \bar{X}_n$ of z_i can be used in CLT
and $\bar{p} \sim ND(\mu, \sigma/\sqrt{n})$ can be applied.

Q5.

a) Using R

SD of sample = 0.193

95% Pivoted bootstrap CI on SD = (0.1732, 0.2158)

b) $H_0: \sigma = 0.2$

$H_a: \sigma \neq 0.2$

Since $H_0 = 0.2 \in (0.1732, 0.2158)$, we cannot reject H_0. The test $\sigma \neq 0.2$ is statistically insignificant at Confidence level of 95%.

Q6.

a) $D = |\text{machine} - \text{expert}|$

Sample mean $\bar{x}_D = 5.00$

the Bootstrap 95% Percentile interval for $\mu_D = (3.733, 6.200)$

b) $H_0: \mu_D = 10$

$H_a: \mu_D \neq 10$

Since $\mu_D = 10 \notin (3.733, 6.200) \Rightarrow H_0$ rejected

$\Rightarrow H_a$ accepted or $\mu_D \neq 10$

Q7. a)

FOR A:-

Sample avg delay = 18.75

95% Bootstrap Pivotal CI = (14.15, 22.75)

FOR B:-

Sample avg delay = 16.35

95% Bootstrap Pivotal CI = (12.05, 20.10)

For C:-

Sample Avg delay = 21.0

95% Bootstrap Pivotal CI = (16.9, 24.85)

For D:-

Sample Avg delay = 18.2

95% Bootstrap Pivotal CI = (13.0, 22.8)

b) Looking at all the CI ranges, D has the largest range (more variable delays) and C has the least (more consistent in terms of delay). C has its bell curve slightly on the higher side of values (most delay prone) and B on the least (least delay prone).

Q8)

a) $\lambda \to$ avg of number of dependents in families

$X_i \sim$ number of dependendents of a family

Since X_i is randomly chosen & is IID

$\therefore$ $\bar{X}_n \sim ND\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$ (CLT)

In R,

($\bar{x}_n$) Sample mean :- 1.94

(s) Sample SD :- 1.531

The 95% CI for λ (population mean)

$$CI = \left(\bar{X}_n - z_{\alpha/2}\frac{s}{\sqrt{n}},\ \bar{X}_n + z_{\alpha/2}\frac{s}{\sqrt{n}}\right)$$

$\alpha = 0.05$

$z_{\alpha/2} = 1.96$

$$CI = (1.516,\ 2.364)$$

b) Given Poisson dist

$$P(\text{no of dependents} > 5) = 1 - P(\text{no of dependents} < 5)$$

$$= 1 - (1 - e^{-\lambda 5})$$

$$= e^{-\lambda 5}$$

$\hat{p}$ be estimate of probability

$$\left[\hat{p} = e^{-5\hat{\lambda}}\right]$$

CI for $\hat{p}$ at 95% confidence

$$\hat{p} = \left[e^{-5 \times 2.364},\ e^{-5 \times 1.516}\right]$$

95% CI for $\hat{p}$ = $(0.0000074,\ 0.00051)$

Q2

a) Sample mean $\bar{x} = 179.44$

Sample SD. $s = 101.45$

Confidence level = 95% $\Rightarrow \alpha = 0.05$

$z_{\alpha/2} = 1.96$

CI for $\mu = (149.81, 209.084)$

b) 95% Bootstrap PIVOTAL CI for μ

$= (148.1, 207.8)$

c) 95% Bootstrap Pivotal CI for σ

$= (78.8, 127.8)$

d) $X \sim ND(\mu, \sigma)$

$p = P(X \geq 300) = 1 - P(X < 300) \Rightarrow \hat{p}$ be the estimate

$\hat{\mu}_L = 148.1 \rightarrow \hat{\mu}_{avg} = \frac{\mu_L + \mu_H}{2} = \underline{177.95}$

$\hat{\mu}_H = 207.8$

$\hat{\sigma}_L = 78.8 \rightarrow \hat{\sigma}_{avg} = \frac{\sigma_L + \sigma_H}{2} = \underline{103.3}$

$\hat{\sigma}_H = 127.8$

Using norm.dist in excel

$\hat{p} = 1 -$ norm.dist (300, Mavg, σavg, TRUE)

$\hat{p} = \boxed{0.1187}$