total grade: 4.8/6.0

good job!

Ex-1 [0.9/1.4]

(a)

here for given sample;

$n = 20$

$\sigma = 10$

Sample mean $\bar{X} = \dfrac{\sum_{i=1}^{n} x_i}{n}$

$= \dfrac{(34.40 + 37.70 + \ldots \ldots + 35.27)}{20}$

$= \dfrac{828.33}{20}$

$\therefore \bar{X} = 41.4165$ ✓

Margin of error $= Z_{99} \cdot \dfrac{\sigma}{\sqrt{n}}$

$= \dfrac{(2.576)(10)}{\sqrt{20}}$

$= 5.76$

$CI = \bar{X} \pm M.E. = 41.4165 \pm 5.76$

$\boxed{\therefore CI = [35.65, 47.176]}$ ✓

(b) when σ is unknown we use t-distribution.

we first calculate sample s.d;

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{20}(x_i-\bar{x})^2}$$

$$= \sqrt{\frac{(20.33-41.41)^2+\dots\dots+(57.87-41.41)^2}{\cancel{20}\ 19}}$$

$$= \cancel{\sqrt{89.8892}}\ \sqrt{94.6202}$$

$$\therefore \sigma = 9.7272$$

t-value for 99% confidence and $df = n-1 = 19$;

$$t_c = 2.861$$

So; $M.E. = t_c \cdot \frac{\sigma}{\sqrt{n}}$

$$= \frac{(2.861)\,(\cancel{9.48}\ 9.7272)}{\sqrt{20}}$$

$$\boxed{\therefore M.E. = 6.22}$$

$C.I. = \bar{X} \pm M.E. = 41.41 \pm 6.22$

$CI = [35.19, 47.63]$

R Notebook

1 -(c)

```
# Given sample
data <- c(34.40, 37.70, 55.59, 40.71, 41.29, 57.15, 44.61, 27.35, 33.13, 35.54,52.24, 43.60,
44.01, 41.11, 34.44, 57.87, 44.98, 20.33, 47.01, 35.27)

sample_mean <- mean(data)
sample_sd <- sd(data)
df <- length(data) - 1

#t-score for a 99% confidence interval
t_score <- qt(0.995, df)

# Margin of error
margin_of_error <- t_score * (sample_sd / sqrt(length(data)))
#Confidence interval
CI <- c(sample_mean - margin_of_error, sample_mean + margin_of_error)

cat("Sample Mean:", sample_mean, "\n")
```

```
## Sample Mean: 41.4165
```

```
cat("Sample Standard Deviation:", sample_sd, "\n")
```

```
## Sample Standard Deviation: 9.727295
```

```
cat("Degrees of Freedom:", df, "\n")
```

```
## Degrees of Freedom: 19
```

```
cat("t-score:", t_score, "\n")
```

```
## t-score: 2.860935
```

```
cat("Margin of Error:", margin_of_error, "\n")
```

```
## Margin of Error: 6.222788
```

```
cat("99% Confidence Interval:", CI, "\n")
```

```
## 99% Confidence Interval: 35.19371 47.63929
```

(d) we want length of CI to be less than 5:

$\therefore 2 \times M.E. \leq 5$

$\therefore 2 \times t\left(\frac{\sigma}{\sqrt{n}}\right) \leq 5$

$\therefore n \geq \left(\frac{2 \times \boxed{t \times} \sigma}{5}\right)^2$

but t alos depends on n
you need to improve the approx.

putting values for t_c & σ

$\therefore n \geq \left(\frac{2 \times 2.861 \times 9.7272}{5}\right)^2$

$\therefore n \geq (11.13)^2$

$\therefore n \geq 123.91$

so sample size must be atleast 124.

Ex.-2

[0.2/0.2]

here;

Length of CI. $\leq 0.01\sigma$

$\therefore 2 \times M.E. \leq 0.01\sigma$

$\therefore 2 \times z_{95} \cdot \frac{\sigma}{\sqrt{n}} \leq 0.01\sigma$

$\therefore 2 \times (1.96) \cdot \frac{\cancel{\sigma}}{\sqrt{n}} \leq 0.01\cancel{\sigma}$

$\therefore n \geq \left(\frac{2 \times 1.96}{0.01}\right)^2$

$\boxed{\therefore n \geq 153664}$ ✓

so sample size should be atleast 153664.

Ex.-4

[0.4/0.4] from chapter 4.4.

$$\hat{\lambda}_n = \bar{X}_n = 2.5$$

for 98% confidence;

$$C.I. = \left[\hat{\lambda}_n - 2.33 \cdot \sqrt{\frac{\hat{\lambda}_n}{n}},\ \hat{\lambda}_n + 2.33 \sqrt{\frac{\hat{\lambda}_n}{n}} \right]$$

$$\therefore CI = \left[2.5 - 2.33 \sqrt{\frac{2.5}{150}},\ 2.5 + 2.33 \sqrt{\frac{2.5}{150}} \right]$$

$$\therefore CI = [2.5 - 0.300,\ 2.5 + 0.300]$$

$$\boxed{\therefore C.I. = [2.199, 2.800]}$$ ✓

Note

[0.5/1.0]

06) b) Cauchy distribution PDF $(v=1)$

$$f(x:v=1) = \frac{1}{\pi(1+x^2)}$$ ✓

CDF,

$$F(x:v=1) = \int_{-\infty}^{x} f(t:v=1)\,dt$$

$$F(x:v=1) = \int_{-\infty}^{x} \frac{1}{\pi(1+t^2)}\,dt$$

$$\int \frac{1}{1+t^2} = \tan^{-1}(t)$$

$$F(x:v=1) = \frac{1}{\pi}\int_{-\infty}^{x} \frac{1}{1+t^2}\,dt$$

$$= \frac{1}{\pi}\left[\tan^{-1}(t)\right]_{-\infty}^{x}$$

$$= \frac{1}{\pi}\left[\tan^{-1}(x) - \tan^{-1}(-\infty)\right]$$

Note

$$= \frac{1}{\pi}\left[\tan^{-1}(x) + \frac{\pi}{2}\right]$$ ✓

The CDF for the Cauchy di (v=1) is given by

$$\frac{1}{\pi}\left[\tan^{-1}(x) + \frac{\pi}{2}\right]$$

Note
[1/1]

Q7) a) $\underline{X_n}$

$$X_n = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$X_n = \frac{1}{n}\sum_{i=1}^{n} (\mu + \sigma z_i)$$

$$X_n = \mu + \frac{\sigma}{n}\sum_{i=1}^{n} z_i$$

Since $\sum_{i=1}^{n} z_i = n \cdot Z_n$

$$X_n = \mu + \frac{\sigma}{n} \cdot n \cdot Z_n$$

$$X_n = \mu + \sigma \cdot Z_n$$ // ✓

CamScanner

Note

<u>S_x.</u>

$$S_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2}$$

$$S_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(\cancel{\mu} + \sigma z_i - \cancel{\mu} - \sigma \bar{z}_n\right)^2}$$

$$S_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(\sigma z_i - \sigma \bar{z}_n)^2}$$

$$S_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\sigma^2(z_i - \bar{z}_n)^2}$$

$$S_x = \sigma\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(z_i - \bar{z}_n)^2}$$

Since $\sum_{i=1}^{n}(z_i - \bar{z}_n)^2 = (n-1)S_z^2$

$$S_x = \sigma\sqrt{\frac{1}{\cancel{(n-1)}}\cancel{(n-1)}\cdot S_z^2}$$

$$S_x = \sigma \cdot S_z$$ //

✓

b) $$\frac{X_n - \mu}{S_x/\sqrt{n}} = \frac{Z_n}{S_z/\sqrt{n}}$$

$$L.H.S = \frac{(X_n - \mu)\sqrt{n}}{S_x}$$

$$X_n = \mu + \sigma Z_n \quad \text{and}$$

$$S_x = \sigma \cdot S_z$$

$$L.H.S = \frac{(\cancel{\mu} + \sigma \cdot Z_n - \cancel{\mu})\sqrt{n}}{\sigma \cdot S_z}$$

$$= \frac{\sqrt{n} \cdot Z_n}{S_z}$$

$$= \frac{Z_n}{S_z/\sqrt{n}}$$

$$L.H.S = R.H.S$$

Note

Explanation

The studentized mean. $\frac{X_n - \mu}{S_n/\sqrt{n}}$ and

it is equal to $\frac{Z_n}{S_Z/\sqrt{n}}$. (proved)

So, This implies that the distribution of the studentized mean depends only on the standardized values of the observations. not on the population mean μ or the standard daviation σ.

3] [0.8/1.0]

$X_n - Y_m$ follow normal distribution with mean $\mu_1 - \mu_2$
& variance $\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}$ why?

The formula for confindence interval :-

CI : point estimate $\pm$ margin of error

point estimate $= X_n - Y_m$
i.e $\quad = \bar{X}_n - \bar{Y}_m$

margin of error $= z \times$ standard deviation ... z is critical value

Critical value $Z_{\alpha/2}$, ~~$\alpha/2$ in the upper tail~~
assume $\alpha = 0.05$ for a 95% confidence interval
$\Rightarrow (1 - 0.05) = 0.95)$
The critical value $Z_{\alpha/2}$ for $\alpha/2 = 0.025$
which is approximately 1.96

& standard deviation $= \sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}$

The confidence interval is :-

$$\left(\bar{X} - \bar{Y} - 1.96\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}},\ \bar{X} - \bar{Y} + 1.96\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}\right)$$

✓ final answer correct

[1/1] 5] a)

by CLT,

the sample mean $\bar{X}_n$ converges to normal distribution as n increase.

$\therefore \bar{X}_n \xrightarrow{d} N(\lambda, \lambda/n)$

better to write that this random variable has approx. this distribution but not use convergence notation (when you write something converges to something the right-hand side must not depend on n !)

we have to approx. ~~var~~ variance of $Y_n := (\bar{X}_n)^2$

Delta formula is given by :

$$Var(g(\bar{X}_n)) \approx [g'(\lambda)]^2 \cdot Var(\bar{X}_n)$$

$$g(\bar{X}_n) = (\bar{X}_n)^2$$

Derivative of $g(\bar{X}_n)$ with resp. ~~to~~ to $\bar{X}_n$

$$g'(x_n) = 2\bar{X}_n$$

$$var[g(\bar{X}_n)) \approx (2\bar{X}_n)^2 \times \left(\frac{\bar{X}_n}{n}\right)$$

$$Var(Y_n) \approx (2\lambda)^2 \times \left(\frac{\lambda}{n}\right)$$

$$Var[Y_n] \approx \frac{4\lambda^3}{n}$$ ✓

R Notebook

```r
# Setting the parameters
n <- 30
lambda <- 2
num_simulations <- 10000

# Simulate Y_n for n = 30 and lambda = 2
set.seed(123)  #  for reproducibility
Yn <- replicate(num_simulations, mean(rpois(n, lambda))^2)

# Computing the sample variance
sample_variance <- var(Yn)

# Computing the true variance according to part a
true_variance <- 4 * lambda^3 / n

# Print the results
cat("True Variance of Y_n:", sample_variance, "\n")
```

```
## True Variance of Y_n: 1.095191
```

```r
cat("sample Variance of Y_n:", true_variance, "\n")
```

```
## sample Variance of Y_n: 1.066667
```