Practical Session 3 in R — Sampling and Estimation

Instructor: Dr. Chennaf Bouchra

Objectives of the Lab

- Apply random sampling techniques using `R`.
- Compute sample statistics: mean, variance, and standard deviation.
- Construct confidence intervals for the population mean.

Exercise 1

Consider the monthly salaries (in \$) of 8 employees:

$$\text{population} = (2200, 2400, 2600, 2800, 2500, 2700, 2900, 3000)$$

1. Draw a random sample of size $n = 4$ without replacement.
2. Compute the sample mean and variance.
3. Construct a 95% confidence interval for the mean (σ unknown).

Solution and R Code

```
# Step 1: Random sample
population <- c(2200, 2400, 2600, 2800, 2500, 2700, 2900, 3000)
set.seed(101)  # reproducibility
n <- 4
sample_data <- sample(population, n, replace = FALSE)
sample_data

# Suppose the output is:
# sample_data = 2400, 2600, 2800, 2500

# Step 2: Sample mean and variance
x_bar <- mean(sample_data)       # 2575
v <- mean((sample_data - x_bar)^2)  # uncorrected variance 21875
S2 <- (n/(n-1)) * v              # corrected (unbiased) variance 29166.67
S <- sqrt(S2)                     # standard deviation 170.81

# Step 3: 95% confidence interval using Student t
t_val <- qt(0.975, df=n-1)       # t_{0.975,3}  3.182
IC_lower <- x_bar - t_val * S / sqrt(n)
IC_upper <- x_bar + t_val * S / sqrt(n)
c(IC_lower, IC_upper)            # [2303.3, 2846.7]
```

Step-by-step explanation:

- Sample drawn: 2400, 2600, 2800, 2500.
- Sample mean: $\bar{x} = 2575$.
- Sample variance (corrected): $S^2 \approx 29166.67$, standard deviation $S \approx 170.81$.
- 95% confidence interval for the mean:

$$\bar{x} \pm t_{0.975,3} \frac{S}{\sqrt{n}} = 2575 \pm 3.182 \frac{170.81}{2} = 2575 \pm 271.7 = [2303.3, 2846.7]$$

Conclusion: We are 95% confident that the average salary of the 8 employees lies between \$2303.3 and \$2846.7.

Exercise 2:

A small company records the daily working hours (in hours) of 10 employees for one week:

$$\text{working\_hours} = (8, 7.5, 9, 8.5, 7, 8, 8.5, 9, 7.5, 8)$$

1. Draw a random sample of size $n = 5$ without replacement.
2. Compute the sample mean, variance, and standard deviation.
3. Plot a bar chart of the sampled working hours.

Solution and R Code

```
# Step 3: Bar chart of sampled data
barplot(sample_hours, names.arg=1:n, col="lightblue",
        main="Sampled Working Hours",
        xlab="Employee Index", ylab="Hours")
```