TUTORIAL SERIES
4
Sampling and Estimation

1 Exercise 1

We seek to estimate the mean age, the variance of ages, and the proportion of smokers in a population of five people labeled P1, P2, P3, P4, and P5. The answers to the two following questions are given below:

- "What is your age?"

- "Are you a smoker?"

Persons	Age	Smoker
P1	22 years	Yes
P2	20 years	No
P3	25 years	No
P4	24 years	Yes
P5	24 years	Yes

Based on all possible samples of size $n = 2$:

1. Define the random variables X_1 and X_2 of the sample (X_1, X_2).

2. Give the probability distributions of the random variables X_1 and X_2.

3. Give the probability distributions of the sample mean, the sample variance, and the sample proportion.

4. Study the properties of these different sampling variables.

Conclusion.

2 Exercise 2

We consider a population U of $N = 5$ individuals for which the values of variable Y are known:

$$y_1 = 3, \quad y_2 = 1, \quad y_3 = 0, \quad y_4 = 1, \quad y_5 = 5.$$

A simple random sample S of size 3 is selected from this population.

1. Determine the mean, median, and variance of Y in the population. List all possible samples of size $n = 3$. What is the probability of selection for each sample?

2. For a given sample, estimate the population mean (and the median). Compute these estimators for each sample and show that the mean estimator is unbiased, whereas the median estimator is biased.

3. For each sample, compute the estimator S_Y^2 of $S_Y^2(U)$ and deduce that this estimator is unbiased. (Recall that $S^2 = \dfrac{1}{n-1}\sum_i (X_i - \overline{X})^2$.)