# Correction – Decision Statistics Exam (IDO 2025)

## Exercise 1 (5 points)

Population: $\{3, 5, 7, 10, 15\}$, $N = 5$, samples of size $n = 2$.

### 1. Population mean and standard deviation

$$\mu = \frac{3+5+7+10+15}{5} = 8 \qquad \text{(0.75 pt)}$$

$$\sigma^2 = \frac{(3-8)^2 + (5-8)^2 + (7-8)^2 + (10-8)^2 + (15-8)^2}{5} = \frac{88}{5} = 17.6$$

$$\sigma = \sqrt{17.6} \approx 4.195 \qquad \text{(0.75 pt)}$$

### 2. Mean of the sampling distribution of $\bar{X}$

$$E(\bar{X}) = \mu = 8 \qquad \text{(1 pt)}$$

### 3. Standard deviation of the sampling distribution of $\bar{X}$

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{4.195}{\sqrt{2}} \approx 2.97 \qquad \text{(1 pt)}$$

### 4. Mean of the sampling distribution of $S^2$

$$E(S^2) = \frac{n-1}{n}\sigma^2 = \frac{1}{2} \times 17.6 = 8.8 \qquad \text{(1 pt)}$$

### 5. Exhaustive sampling case

$$E(S^2) = \frac{N-}{N-1}\frac{n-1}{n}\sigma^2 = 11$$

The non-exhaustive case underestimates the population variance.

(0.5 pt)

# Exercise 2 (5 points)

Given: support proportion $p = 0.46$, $q = 0.54$.

## 1.(a) Definition of the random variable

Let $F$ be the random variable representing the **frequency of voters supporting the candidate** in a random sample of size $n$. **1 pt**

For $n \geq 30$ and $np \geq 5$, the sampling distribution of $F$ is approximately normal:

$$F \sim \mathcal{N}\left(p, \frac{pq}{n}\right), \qquad E(F) = p, \quad \sigma_F = \sqrt{\frac{pq}{n}}.$$

## 1.(b) Sample of 200 voters – majority probability

**1.5 pts**

$$E(F) = 0.46, \qquad \sigma_F = \sqrt{\frac{0.46 \times 0.54}{200}} = 0.035.$$

Majority condition: $F \geq 0.5$.

$$P(F \geq 0.5) = 1 - \Phi\left(\frac{0.5 - 0.46}{0.035}\right) = 1 - \Phi(1.14) = 0.1271.$$

$$\boxed{P = 12.71\%}$$

## 2. Sample of 1000 voters – majority probability

**1.5 pts**

$$E(F) = 0.46, \qquad \sigma_F = \sqrt{\frac{0.46 \times 0.54}{1000}} = 0.016.$$

$$P(F \geq 0.5) = 1 - \Phi\left(\frac{0.5 - 0.46}{0.016}\right) = 1 - \Phi(2.5) = 0.0062.$$

$$\boxed{P = 0.62\%}$$

## 3. Less than 40% supporters (sample of 200)

**1 pt**

$$P(F < 0.4) \approx P\left(Z < \frac{0.4 - 0.46}{0.035}\right) = P(Z < -1.77) = 0.0384.$$

Expected number among 1000 samples:

$$1000 \times 0.0384 \approx \boxed{38}.$$

# Exercise 3 (5 points)

## 1. Sample size (Chebyshev inequality)

Since the distribution of the variable is unknown, the Bienaymé–Chebyshev inequality is used: **3 pts**

$$n = \frac{\sigma^2}{\alpha\,\varepsilon^2}.$$

With $\sigma = 5000$, $\varepsilon = 1000$, and $\alpha = 0.05$, we obtain:

$$n = \frac{5000^2}{0.05 \times 1000^2} = 500.$$

$$\boxed{n = 500}$$

## 2. Effect of increasing the confidence level to 99%

Increasing the confidence level from 95% to 99% means decreasing $\alpha$ from 0.05 to 0.01. Since $n$ is inversely proportional to $\alpha$, the required sample size **increases**. **2 pts**

# Exercise 4 (5 points)

## 1. Likelihood function

$$L(m, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(X_i - m)^2}{2\sigma^2}\right) \qquad \text{(1 pt)}$$

## 2. MLE of $m$

$$\hat{m} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad \text{(1.5 pts)}$$

## 3. MLE of $\sigma^2$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 \qquad \text{(1.5 pts)}$$

## 4. Bias

$$E(\hat{\sigma}^2) = \frac{n-1}{n}\sigma^2$$

Thus, the estimator is biased.

(1 pt)