TP2 : Hypothesis tests

Elèves : IMAHO Aladin, HADJ SASSI Mahdi

Encadrant : YANG Shuo

ENSEA - 2024-2025

Table des matières

1 Preparation

1.1 Question 1

The skewness of a probability distribution is a measure of its asymmetry with respect to its mean. It is defined as:

$$\gamma_1 = \mathbb{E}\left[\left(\frac{X-\mu}{\sigma}\right)^3\right]$$

where μ is the mean and σ is the standard deviation.

- If $\gamma_1 > 0$, the distribution is skewed to the right (longer tail on the right).
- If $\gamma_1 < 0$, the distribution is skewed to the left.
- If $\gamma_1 = 0$, the distribution is symmetric.

Kurtosis measures the concentration of values around the mean:

$$\gamma_2 = \mathbb{E}\left[\left(\frac{X-\mu}{\sigma}\right)^4\right] - 3$$

A normal distribution has a kurtosis of 0. A positive kurtosis indicates a more peaked distribution than the normal one, whereas a negative kurtosis indicates a flatter distribution.
Proof that the skewness of a normal distribution is zero: Let $X \sim \mathcal{N}(\mu, \sigma^2)$. By definition, a standardized normal variable $Z = \frac{X-\mu}{\sigma}$ follows $\mathcal{N}(0, 1)$. For any standard normal variable:

$$\mathbb{E}[Z^3] = 0$$

Thus, by the definition of γ_1:

$$\gamma_1 = \mathbb{E}[Z^3] = 0.$$

This proves that a normal distribution is always symmetric.

1.2 Question 2A

The null and alternative hypotheses for the χ^2 goodness-of-fit test are:

$$H_0 : (X_1, \ldots, X_n) \sim \mathcal{N}(\mu, \sigma^2)$$

$$H_1 : (X_1, \ldots, X_n) \text{ does not follow a normal distribution } \mathcal{N}(\mu, \sigma^2)$$

1.3 Question 2B

The maximum likelihood estimators of μ and σ^2 are given by:

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

$$\sigma'^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu')^2$$

1.4 Question 2C

The theoretical expectation of the number of realizations x_i satisfying $x_i \in [a, b]$ is given by:

$$\mathbb{E}[N_{[a,b]}] = n \cdot (F(b) - F(a))$$

where F is the cumulative distribution function of $\mathcal{N}(\mu', \sigma'^2)$.

1.5 Question 2D

To apply the χ^2 test without merging classes, the expected theoretical count in each class must be greater than 5:

$$\forall i, \quad \mathbb{E}[N_i] \geq 5.$$

1.6 Question 2E

The type-I error risk α is the probability of rejecting H_0 when it is actually true. When α increases, the rejection zone enlarges, making the test more stringent:

$$\text{Critical threshold} = \chi^2_{1-\alpha,k}$$

where k is the number of degrees of freedom.