Tabea Rebafka
October 2018

Feuille d'exercice N° 1

Descriptive statistics & point estimation

Exercice 1. *Data visualization*

Figure 1 shows four i.i.d. datasets with sample size 100. For each dataset

(i) deduce the characteristics of the the distribution of the data,

(ii) propose a distribution (or family of distributions) that may have generated the data.

Exercice 2. *MM et MLE*
We observe $\mathbf{x} = (x_1, \dots, x_n)$ that we consider as a realization of the random vector $\mathbf{X} = (X_1, \dots, X_n)$, where X_i are i.i.d. random variables with distribution $\mathbb{P}_{\theta_0}$. Compute the estimator by the method of moments and the maximum likelihood estimator θ_0 (if they exist) in the following statistical models $\{\mathbb{P}_\theta, \theta \in \Theta\}$.

1. $\mathbb{P}_\theta$ is the Poisson distribution $\mathcal{P}(\theta)$ with parameter $\theta > 0$.
2. $\mathbb{P}_\theta$ is the geometric distribution $\mathrm{Geo}(\theta)$ with parameter $0 < \theta < 1$ with probabilities
$$\mathbb{P}_\theta(X = \ell) = \theta(1-\theta)^\ell , \quad \text{pour } \ell = 0, 1, \dots .$$
You may use that $\sum_{k=0}^{\infty} k a^k = \frac{a}{(1-a)^2}$ for $0 < a < 1$.
3. $\mathbb{P}_\theta$ is the normal distribution $\mathcal{N}(\mu, \sigma^2)$ with $\mu \in \mathbb{R}$ and $\sigma^2 > 0$ and $\theta = (\mu, \sigma^2)$.
4. $\mathbb{P}_\theta$ has density $f_\theta(x) = \mathrm{e}^{-(x-\theta)} \mathbb{1}_{\{x \geq \theta\}}$ with parameter $\theta \in \mathbb{R}$.

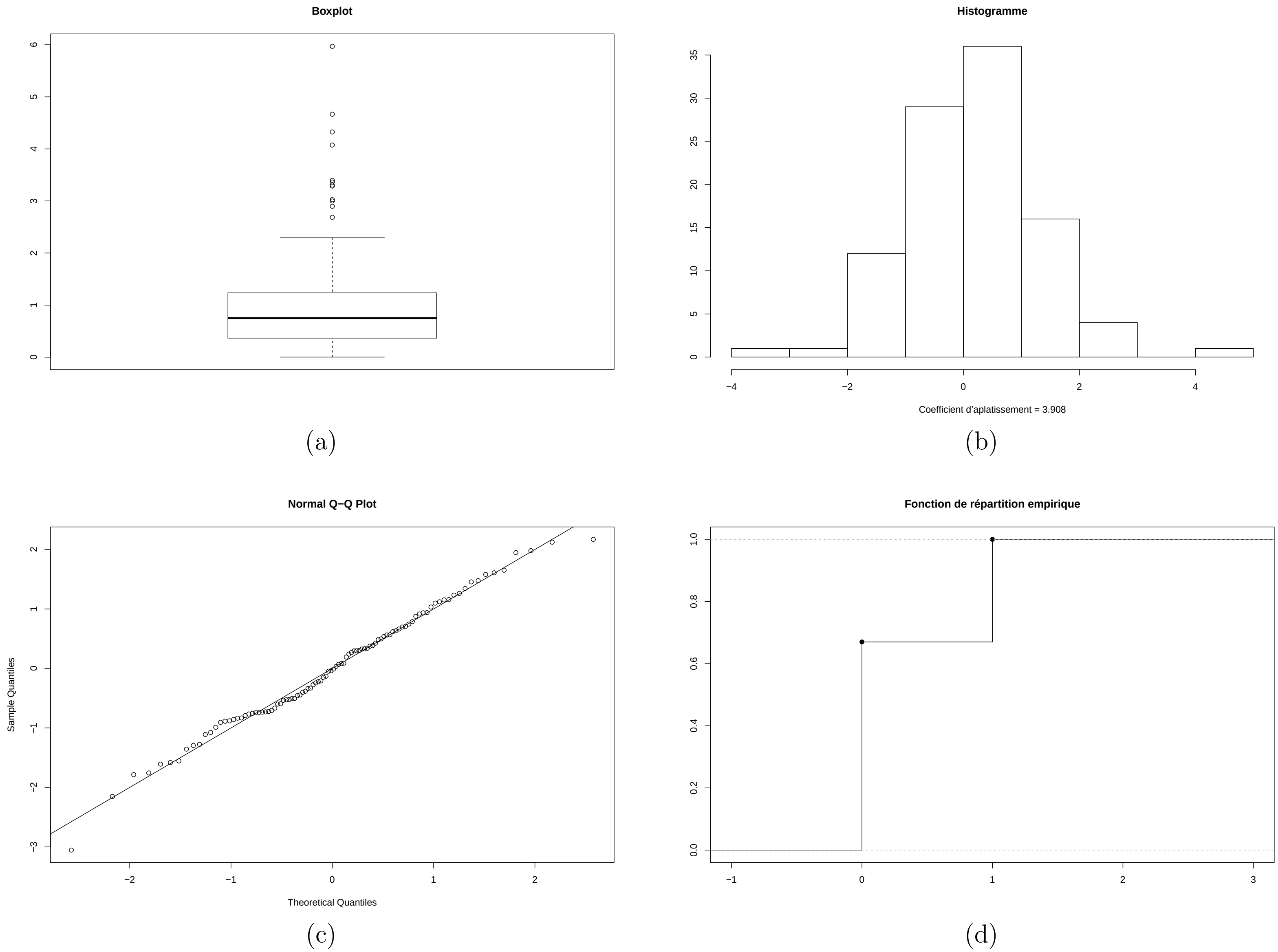

Figure 1: Graphical representation of four different datasets.

Exercice 3. *Bernoulli distribution*
Suppose that the observations $\mathbf{x} = (x_1, \dots, x_n)$ are n independent realizations a Bernoulli random variable X with unknown Bernoulli parameter $p \in (0, 1)$.

1. (a) Estimate p by the method of moments and the maximum likelihood estimator.
   (b) Compute the mean squared error of the MLE $\hat{p}_n$ of p. Show that $\hat{p}_n$ is consistent and determine its limit distribution.

2. (a) Denote $v = p(1-p)$ the variance of the Bernoulli distribution $\mathcal{B}(p)$. The statistic $\hat{v}_n = \bar{x}_n(1-\bar{x}_n)$ is an estimator of v. By which method have we found this estimator? Show that $\hat{v}_n$ is a consistent estimator of v.
   (b) Based on $\hat{v}_n$, propose an unbiased estimator $\tilde{v}_n$ of v.
   (c) Determine the limit distribution of $\tilde{v}_n$.

Exercice 4. *Double exponential distribution*

1. Let $\mathbf{x} = (x_1, \dots, x_n)$ be an i.i.d. sample from the *double exponential distribution* or *Laplace distribution*, with density given by

$$f_\theta(x) = \frac{1}{2\theta} \mathrm{e}^{-\frac{|x|}{\theta}}, \quad x \in \mathbb{R},$$

   with $\theta \in \Theta =]0, +\infty[$.

   (a) Draw the density f_θ for different values of θ.
   (b) Compute the MLE of θ. Is the MLE unique et and almost surely well defined? Is the MLE consistent and asymptotically normal?

2. We now consider an i.i.d. sample $\mathbf{x}$ with double exponential distribution with location parameter μ. More precisely, the density is given by

$$f_\mu(x) = \frac{1}{2} \mathrm{e}^{-|x-\mu|}, \quad x \in \mathbb{R},$$

   where $\mu \in \mathbb{R} = \Theta$.

   2.1 Draw the density f_μ for different values of μ.
   2.2 Compute the MLE of θ. Is the MLE unique et and almost surely well defined? Is the MLE consistent and asymptotically normal?