

---

**TD n° 1 - Standard laws**


---

**Exercise 1.***First steps with R*

R is a free software environment for statistical computing and graphics (<https://www.r-project.org/>). Besides classical routines for statistics, it integrates a programming language that enables to develop customized tools. The R community is very active and many interesting or sophisticated libraries are available. It is already installed on ENS computers, and can be easily downloaded and installed on any computer and operating system. You can use the command line interface, but there exist several graphical front-ends such as Rstudio (<https://www.rstudio.com/>). You can directly type in the command line interface, but it is advised to save sequences of R commands or programs in a `.r` file. Any editor is suitable. Use the instruction `source("mypath/myfile.r")` to run the contents in the command line interface.

1. The assignment symbol is `<-` and data can be collected in vectors constructed with the concatenation operator `c()`.

Store in the variable `mylotto` your 6 favorite numbers (integers or not). Now consider the instruction `mysymlotto <- c(mylotto, 0, -mylotto)`, try to guess what is stored in `mysymlotto` and check it.

2. Data is also often stored in simple text files in tabular form. Consider the file `network-crash.txt` (<http://perso.ens-lyon.fr/etienne.moutot/static/teach/network-crash.txt>): the first line contains some computer names and each column contains the sequence of daily network crashes (1 = one crash during the day, 0 = no crash) over a large number of days (the same for all computers). This data can be imported in a variable with the command:

`mydata <- read.table("network-crash.txt")`. Find the number of computers and the number of days recorded. Compute the mean number of crashes for a computer.

Note that new functions can be implemented in a classical imperative style. Functions can be declared with the following syntaxes:

```
myfunction <- function(param1,param2,...) output expression
```

```
myfunction <- function(param1,param2,...) {program body ending by an output expression}
```

**Exercise 2.***Bernoulli law*

Let  $p \in [0, 1]$ , a random variable  $X$  follows a *Bernoulli law*  $\mathcal{B}(p)$ , if  $X$  takes its values  $\{0, 1\}$  with  $\mathbb{P}(X = 1) = p$  and  $\mathbb{P}(X = 0) = 1 - p$ .

1. By hand, compute the mean, the variance and the moment formal series for  $X$ .

2. Using R, fix a value for  $p$ , create a large sample of independent  $\{0, 1\}$  random values following the Bernoulli law  $\mathcal{B}(p)$  (*suggestion : use the `rbinom()` random function and remember that one Binomial experiment is the repetition of several Bernoulli experiments*). Compute the empirical average for your sample (sum of values / number of values). Could you expect the result and do you think the other students observe the same thing? What if you choose some even bigger sample?

**Exercise 3.***Poisson law*

Let  $\lambda > 0$ , a random variable  $X$  follows a *Poisson law*  $\mathcal{P}(\lambda)$  if  $X$  takes its values in  $\mathbb{N}$  and  $\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$ .

1. Using R, draw histograms for several Poisson law samples of different sizes and different parameters  $\lambda$  to visualize those kind of shapes, using to the `hist()` function.

2. Compute the mean, the variance and the moment formal series for  $X$ .

3. Show that if  $X$  and  $Y$  are two independent random variables with poisson laws of parameters  $\lambda$  and  $\mu$ , then  $X + Y$  is still a poisson law with parameter  $\lambda + \mu$ .

The *Poisson approximation principle* states that the sum  $S_n$  of many independent bernoulli random variables  $X_1, \dots, X_n$  with small parameters, tends to follow a Poisson law of parameter  $E(S_n)$ .

4. 1000 competitors attend a fishing competition, each one having independently a probability 0,0001 to hook a fish and win a victory medal. The organizing committee has bought 2 medals, is it enough to reward all the winners with probability at least 80 % ? Use R as a pocket calculator to help you.

#### Exercise 4.

*Geometric law*

Let  $p \in [0, 1]$ , a random variable  $X$  follows a *geometric law*  $\mathcal{G}(p)$  if  $X$  takes its values in  $\mathbb{N}^*$  and  $\mathbb{P}(X = k) = (1 - p)^{k-1} p$ .

1. Give an example of modelisation using a random variable with law  $\mathcal{G}(p)$ .
2. Compute the mean, the variance and the moment formal series for  $X$ .
3. Show that geometric laws have no memory : for all  $n, k \in \mathbb{N}$ ,  $\mathbb{P}(X = k + n | X > n) = \mathbb{P}(X = k)$ .

#### Exercise 5.

*Exponential law*

Let  $\lambda > 0$ , a random variable  $X$  follows an *exponential law*  $\text{Exp}(\lambda)$  if  $X$  takes its values in  $\mathbb{R}_+$  with density  $f(x) = \lambda e^{-\lambda x}$ .

1. Use R to plot this density for your favorite value of  $\lambda$ , try to play on the horizontal and vertical ranges of the figure.
2. Compute the mean and the variance of  $X$ .
3. Show that exponential laws have no memory: for all  $a, b \in \mathbb{R}_+$ ,  $\mathbb{P}(X \geq a + b | X \geq b) = \mathbb{P}(X \geq a)$ .
4. Show the converse: any continuous law with no memory is exponential.
5. Let  $X_1, \dots, X_n$  independent exponential random variables of parameters  $\lambda_1, \dots, \lambda_n$ , prove that  $\min(X_1, \dots, X_n)$  also follows an exponential law and find its parameter. Compute  $\mathbb{P}(\min(X_1, \dots, X_n) = X_i)$  for  $1 \leq i \leq n$ .  
*Suggestion* : start with  $n = 2$ .

#### Exercise 6.

*Normal law*

Let  $m, \sigma^2 \in \mathbb{R}_+$ , a random variable  $X$  follows a *normal law*  $\mathcal{N}(m, \sigma^2)$  if  $X$  takes its values in  $\mathbb{R}$  and has density  $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}$ .

1. Compute the mean and variance for the normal law  $\mathcal{N}(m, \sigma^2)$ .
2. Let  $X, Y$  be independent normal random variables of laws  $\mathcal{N}(m_1, \sigma_1^2)$  and  $\mathcal{N}(m_2, \sigma_2^2)$ , show that  $X + Y$  has law  $\mathcal{N}(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$ .

#### Exercise 7.

*Optional homework question (from P. Brémaud)*

A long time ago, a number  $p$  was chosen at random uniformly between 0 and 1, but this value was never revealed to mankind. Since this time, the sun rises every day with probability  $p$  (still unknown). What happened during the preceding days is independent of what happens today. You know that the sun has risen every day from the beginning, that is  $n$  times (and you know this number), what is the probability that it will rise tomorrow ?