Laboratory 3 - Simulation of random variables.
(Illustrations of LLN and CLT)

I. Remarkable continuous distributions

RStudio. Don't forget to set the working directory: **Session** → **Set Working Directory** → **Choose Directory**.

Solved exercise. Graphically represent the density of the Exponential distribution, $Exp(\lambda)$ ($\lambda > 0$).

This distribution is null for negative values of the real axis, hence it will be enough to represent it on the positive semi-axis (we will use in fact an interval of type $[0, a]$).

```
density_exponential = function(lambda, n, a) {
   x = seq(0, a, n;
   y = dexp(x, lambda);
   plot(x, y, type = 'l');
}
```

Exercise to work.

I.1. Write three functions that graphically represent the densities of the following distributions:

(a) $Gamma(\alpha, \lambda)$.

(b) $Student(r)$.

(c) $N(\mu, \sigma^2)$.

II. The Law of Large Numbers (LLN).

Let X_i, $1 \leqslant i \leqslant n$, be a sequence of independent and identically distributed random variables, their mean is

$$\overline{x}_n = \frac{X_1 + X_2 + \ldots + X_n}{n},$$

Then, according to LLN, $\overline{x}_n \to \mu$, where $\mu = \mathbb{E}[X_i]$, $\forall 1 \leqslant i \leqslant n$.

Solved exercise. Verify the LLN using the sequence of random variables $X_i : Poisson(\lambda)$, $1 \leqslant i \leqslant n$. (We know that $\mathbb{E}[X_i] = \lambda$.)

```
LLN_Poisson = function(lambda, n) {
   sum = 0;
   for(i in 1:n) {
      u = rpois(1, lambda);
      sum = sum + u;
   }
   return(sum/n);
}
```

A more simple (and also faster) variant:

```
LLN_Poisson = function(lambda, n) {
   return(mean(rpois(n, lambda)));
}
```

Solved exercise. Verify the LLN using the sequence of random variables $X_i : Gamma(\alpha, \lambda)$, $1 \leqslant i \leqslant n$. (We know that $\mathbb{E}[X_i] = \alpha/\lambda$.)

We employ here just just the faster variant:

```
LLN_Gamma = function(alfa, lambda, n) {
  return(mean(rgamma(n, alfa, lambda)));
}
```

Exercises to work.

II.1. Write a function that has to verify the LLN using the sequence of random variables

(a) $X_i : Exponential(\lambda)$, $1 \leqslant i \leqslant n$. (We know that $\mathbb{E}[X_i] = 1/\lambda$.)

(b) $X_i : B(m, p)$, $1 \leqslant i \leqslant n$. (We know that $\mathbb{E}[X_i] = mp$.)

II.2 Solve the same exercise for the sequence of random variables $X_i : Student(r)$, $1 \leqslant i \leqslant n$. (We know that $\mathbb{E}[X_i] = 0$.) Compare the results with the exact values for the following parameters: $n \in \{1000, 10000, 100000, 1000000\}$ and $r \in \{2, 3, 4, 5\}$.

III. The Central Limit Theorem (CLT).

Let X_i, $1 \leqslant i \leqslant n$, be a sequence of independent and identically distributed random variables: $\mathbb{E}[X_i] = \mu$ and $Var[X] = \sigma^2$, $\forall 1 \leqslant I \leqslant n$. Let

$$\overline{x}_n = \frac{X_1 + X_2 + \ldots + X_n}{n},$$

be their mean, then, according to CLT, $\overline{x}_n$ (for large values of n) follows the distribution of $N(\mu, \sigma^2)$.

By using the standardization of the sample mean we get

$$\frac{\overline{x}_n - \mu}{\sigma/\sqrt{n}} : N(0, 1).$$

Solved exercise. Verify the CLT using the sequence of random variables $X_i : Poisson(\lambda)$, $1 \leqslant I \leqslant n$. (We know that $\mathbb{E}[X_i] = \lambda$ and $Var[X_i] = \lambda$.)

The CLT says that

$$P\left(\frac{\overline{x}_n - \mu}{\sigma/\sqrt{n}} \leqslant z\right) \cong P(Z \leqslant z),$$

where $z \in \mathbb{R}$ and $Z : N(0, 1)$ - the standard normal law. The probability from the right is $pnorm(z)$, while the probability from the left can be approximated like this: take a large number, N, of independent simple random samples of the same size n, $\left(X_i^k\right)_{i=\overline{1,n}}^{k=\overline{1,N}}$, and compute

$$P^N(z) = \frac{|\{k \,:\, \overline{x}_n^k \leqslant z\sigma/\sqrt{n} + \mu\}|}{N}.$$

(X_i^k are variates of the given distribution - say Poisson, for our exercise). Then, compare this probability with $pnorm(z)$.

```
CLT_Poisson = function(lambda, n, N, z) {
  expectation = lambda;
  st_dev = lambda;
  upper_bound = z * st_dev/sqrt(n) + expectation;
  sum = 0;
  for(i in 1:N) {
    x_n = mean(rpois(n, lambda));
    if(x_n ⩽ upper_bound) {
      sum = sum + 1;
    }
  }
  return(sum/N);
}
```

Remark: n must be at least 30; e. g., $n = 30$, $N = 10000$, $z \in \{0, 1, 1.5, 2\}$.

Exercises to work.

III.1. Write a function that has to verify the CLT using the sequence of random variables $X_i : Exponential(\lambda)$, $1 \leqslant i \leqslant n$. (We know that $\mathbb{E}[X_i] = 1/\lambda, Var[X_i] = 1/\lambda^2$.)

III.2 Solve the same exercise for the sequence of random variables $X_i : Gamma(\alpha, \lambda)$, $1 \leqslant i \leqslant n$. (We know that $\mathbb{E}[X_i] = \alpha/\lambda$ and $Var[X_i] = \alpha/\lambda^2$.). Choose $n = 50$, $N \in \{5000, 10000, 20000\}$, and $z \in \{-1.5, 0, -1.5\}$.

IV. De Moivre-Laplace approximation

Let $X : B(n, p)$; the probabilities concerning this distribution can be approximated using the normal law (from the Central Limit Theorem - CLT).

$$P(X = k) = P(k - 0.5 \leqslant X \leqslant k + 0.5) =$$
$$= P\left(\frac{k - 0.5 - np}{\sqrt{np(1-p)}} \leqslant \frac{X - np}{\sqrt{np(1-p)}} \leqslant \frac{k + 0.5 - np}{\sqrt{np(1-p)}}\right) \cong$$
$$\cong \Phi\left(\frac{k + 0.5 - np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{k - 0.5 - np}{\sqrt{np(1-p)}}\right),$$

where $\Phi(\cdot)$ is the distribution function for the standard normal law, $N(0, 1)$, i. e., $\Phi(a) = P(Z \leqslant a)$, if $Z : N(0, 1)$. In R $\Phi(a) = pnorm(a)$.

Solved exercise. Let $X : B(30, 0.35)$; compute $P(X = 25)$.
$\mathbb{E}[X] = np = 30 \cdot 0.4 = 12$, $Var[X] = np(1-p) = 30 \cdot 0.4 \cdot 0.6 = 7.2$

$$P(X = 25) = P(24.5 \leqslant X \leqslant 25.5) = P\left(\frac{24.5 - 12}{\sqrt{7.2}} \leqslant \frac{X - 12}{\sqrt{7.2}} \leqslant \frac{24.5 + 12}{\sqrt{7.2}}\right) =$$
$$= P\left(4.658475 \leqslant \frac{X - 12}{\sqrt{7.2}} \leqslant 5.031153\right) \cong \Phi(4.658475) - \Phi(5.031153) =$$
$$= 0.9999998 - 0.9999984 = 0.0000004.$$

Solved exercise. Let $X : B(50, 0.3)$; compute $P(X > 10)$.
$\mathbb{E}[X] = np = 50 \cdot 0.3 = 15$, $Var[X] = np(1-p) = 50 \cdot 0.3 \cdot 0.7 = 10.5$

$$P(X > 10) = P(X \geqslant 11) = P(X \geqslant 10.5) = P\left(\frac{X - 15}{\sqrt{10.5}} \geqslant \frac{10.5 - 15}{\sqrt{10.5}}\right) =$$

$$= P\left(\frac{X - 15}{\sqrt{10.5}} \geqslant -1.38873\right) \cong 1 - \Phi(-1.38873) = \Phi(1.38873) = 0.9175426.$$

Solved exercise. Let $X : B(n,p)$; write a function which has to compute $P(X > k)$ $(0 \leqslant k < n)$.

```
binomial_probability = function(n, p, k) {
   expectation = n*p;
   variance = n*p*(1 - p);
   standard_deviation = sqrt(variance);
   q = (k + 0.5)/standard_deviation;
   return(1 - pnorm(q));
}
```

Exercises to work.

IV.1. Let $X : B(n,p)$; write a function which computes $P(X < k)$ $(0 < k \leqslant n)$.

IV.2. Let $X : B(n,p)$; write a function which has to compute $P(X > k)$ $(0 \leqslant k < n)$.