Practical n°1

stoehr@ceremade.dauphine.fr

Preliminary step and Objectives.
- Make sure to read the introduction to R and do the exercises before starting here.
- This sheet focuses on the use of probability and statistical functions in R.

About probability functions...

Exercise 1 *(Random Generators).*

1. Explain what is the function `set.seed` in R. Why would you use it?

2. Generate a random sample of size 500 from the following discrete distributions:

(a) Bernoulli distribution with parameter $p = 0.4$. Store it in a vector `x1`.

(b) Binomial distribution with parameters $n = 50$, $p = 0.4$. Store it in a vector `x2`.

(c) Poisson distribution with parameter $\lambda = 2$. Store it in a vector `x3`.

3. Generate a random sample of size 1000 from the following continuous distributions:

(a) Uniform distribution on $[2,5]$. Store it in a vector `y1`.

(b) Exponential distribution with mean $1/\lambda = 0.5$. Store it in a vector `y2`.

(c) Normal distribution with mean $\mu = 1$ and variance $\sigma^2 = 2$. Store it in a vector `y3`.

4. An urn contains 24 balls numbered from 1 to 24. Use R to draw at random 10 balls

(a) without replacement, each ball being equiprobable ;

(b) with replacement, each ball being equiprobable ;

(c) without replacement, balls with even numbers being twice as likely to be drawn.

Exercise 2 *(Probabilities and CDF).*

1. Assume X is following the exponential distribution $\varepsilon(2)$ (*i.e.*, with mean 0.5) and Z the Gaussian distribution $\mathcal{N}(0,2)$ (*i.e.*, with variance 2). Using R, compute the following quantities

(a) $\mathbb{P}[X > 0.5]$,

(b) $\mathbb{P}[X \leq 5]$,

(c) $\mathbb{P}[1 < X < 10]$,

(d) $\mathbb{P}[Z > 3]$,

(e) $\mathbb{P}[Z \leq -0.2]$,

(f) $\mathbb{P}[|Z| > 2]$.

2. Represent the cumulative distribution function of the Poisson distribution with mean $\lambda = 4$.

Exercise 3 *(Probability distributions).*

1. **Discrete variable.** Represent the mass function of a Poisson distrubtion with mean $\lambda = 2$.
2. **Continuous variable.** Represent the density function of a Chi-squared distribution with 3 degrees of freedom.

Exercise 4 *(Quantiles).* Consider X distributed according to a binomial distribution with parameters $n = 50$, $p = 0.4$, Y distributed according a Chi-squared distribution with 3 degrees of freedom and Z distributed according to the standard Normal distribution $\mathcal{N}(0,1)$.

1. Using R, find the value of the quantile q in what follows

(a) $\mathbb{P}\left[X \leq q\right] \geq 0.1$, **(b)** $\mathbb{P}\left[Y \leq q\right] = 0.95$, **(c)** $\mathbb{P}\left[Y \geq q\right] = 0.65$, **(d)** $\mathbb{P}\left[|Z| < q\right] = 0.95$.

About statistics...

Preliminary step.
- Download "dat.RData" into a folder "TP1" and set the working directory to the download location.
- Import the data from the file "dat.RData".

Exercise 5 *(Samples distributions and densities).*

1. **Discrete variable.** Represent the distribution of sample `dat1` and compare it to the mass function of a Poisson distrubtion with mean $\lambda = 2$.
2. **Continuous variable.** Represent the distribution of sample `dat2` and compare it to the density function of a Chi-squared distribution with 3 degrees of freedom.

Exercise 6 *(Moments and quantiles).*

1. Compute the deciles of samples `dat1` and `dat2`.
2. Compute the mean, the median and the variance of samples `dat1` and `dat2`.
3. Compare the output of the function `summary` for samples `dat1`, and `dat2` with the summaries for a Poisson distribution with mean $\lambda = 2$ and a Chi-squared distribution with 3 degrees of freedom.