Laboratory Session 03 : April 14, 2021

Exercises due on : May 2, 2021

Exercise 1

- The triangular distribution, in the interval (a, b), is given by the following:

$$f(X) = \begin{cases} \dfrac{2(x-a)}{(b-a)(c-a)} & a \le x < c \\[2ex] \dfrac{2(b-x)}{(b-a)(b-c)} & c \le x \le b \\[2ex] 0 & \text{otherwise} \end{cases}$$

  where $c \in [a, b]$.

a) plot the function, given the interval (a, b)

b) and write an algorithm to generate random numbers from the triangular distribution

c) generate 10^4 random number from the distribution, show them in an histogram and superimpose the analytical curve

Exercise 2 - Markov's inequality

- Markov's inequality represents an upper bound to probability distributions:

$$P(X \ge k) \le \frac{E[X]}{k} \text{ for } k > 0$$

- having defined a function

$$G(k) = 1 - F(k) \equiv P(X \ge k)$$

  plot $G(k)$ and the Markov's upper bound for

a) the exponential, $\text{Exp}(\lambda = 1)$, distribution function

b) the uniform, $\mathcal{U}(3, 5)$, distribution function

c) the binomial, $\text{Bin}(n = 1, p = 1/2)$, distribution function

d) a Poisson, $\text{Pois}(\lambda = 1/2)$, distribution function

Exercise 3 - Chebyshev's inequality

- Chebyshev's inequality tell us that

$$P(\left|X-\mu\right| \geq k\sigma) \leq \frac{1}{k^2}$$

- which can also be written as

$$P(\left|X-\mu\right| < k\sigma) \geq 1-\frac{1}{k^2}$$

- use R to show, with a plot, that Chebyshev's inequality is is an upper bound to the following distributions:

a) a normal distribution, $\mathrm{N}(\mu = 3, \sigma = 5)$

a) an exponential distribution, $\mathrm{Exp}(\lambda = 1)$

b) a uniform distribution $\mathcal{U}(1-\sqrt{2}, 1+\sqrt{2})$

d) a Poisson, $\mathrm{Pois}(\lambda = 1/3)$, distribution function

Exercise 4 - Community Mobility Open Data

- Community Mobility Reports have been created with the aim to provide insights into what has changed in response to policies aimed at combating COVID-19. Data can be found at `https://www.google.com/covid19/mobility/`

- Download and analyze the following data sets:

- `https://www.gstatic.com/covid19/mobility/Global_Mobility_Report.csv`

  and

- `https://www.gstatic.com/covid19/mobility/Region_Mobility_Report_CSVs.zip`

  The data show how visitors to (or time spent in) categorized places change compared to baseline days. A baseline day represents a normal value for that day of the week. The baseline day is the median value from the 5-week period Jan 3 – Feb 6, 2020.
  To make the reports useful, categories have been used to group some of the places with similar characteristics for purposes of social distancing guidance. The following categories are available:

- `retail_and_recreation`, i.e. places like restaurants,cafes, shopping centers, theme parks,museums, libraries, and movie theaters

- `grocery_and_pharmacy`, i.e. grocery markets, food warehouses, farmers markets, specialty food shops, drug stores, and pharmacies

- `parks`, i.e. national parks, public beaches, marinas, dog parks, plazas,and public gardens

- `transit_stations` i.e. all public transport hubs such as subway, bus, and train stations

- `workplaces`, i.e. places of work

- `residential`, i.e. people's residence

• Select a couple of European countries of your choice and analyze the trends in the previous variables over time:

- produce a plot of the data by averaging the observable over a period of one week (hint: convert the `data` field to `lubridate::week`) and one month and quantify the impact of COVID-19 restrictions on mobility sitations.