Statistics and Exploratory Data Analysis

Marcin Chlebus, Ewa Cukrowska-Torzewska

Laboratory 1: Random variables and probability distributions

Exercises for students

Exercise 1:

Suppose you roll an unfair die 50 times with the probability of obtaining "6" equal to 0.6 (i.e. we deal with binomial distribution with n=50 and p=0.6).

a) Plot the probability density function for this distribution
b) What is the probability of obtaining more than 40 times "6"?

Exercise 2:

A study on birds collects information such as the length of their eggs (in mm). Assume that the length is normally distributed with μ = 42.1mm and $\sigma 2 = 20.8^2$.

a) What is the probability of finding an egg with a length greater than 50 mm?
b) What is the probability of finding an egg between 30 and 40 mm in length?
c) Generate probability distribution and cumulative distribution functions for this distribution and show graphically probabilities you calculated in points (a) and (b)

Exercise 3:

Generate 1000 random numbers from a standard normal distribution. Plot the empirical cumulative distribution function of your sample (use ecdf fuction). What is the value of the ecdf for the random number of 0? What does it mean?

Exercise 4:

Use the data on crime rates in the US that are available in R (USArrests).

a) Plot the distribution of murders and rapes in the US
b) Fit the distribution for these two variables. Try different distributions we covered during the class.