Statistics and Numerics

Lecture Prof. Dr. Jens Timmer

Exercises Helge Hass, Mirjam Fehling-Kaschek

Exercise Sheet Nr. 2

Exercise 1: Cauchy distribution

(i) Plot the Cauchy distribution. Supposed your programming package does not include an implementation of a random number generator for Cauchy distributed random variables: how can you generate Cauchy-distributed random numbers? Generate two histograms: one with an in-build function and one with your own.

(ii) Test the central limit theorem for the sum of Cauchy-distributed random variables, similar to sheet 1, exercise 2. This means, you calculate 1000 sums of $M = 1, \ldots, 1000$ (in reasonable spacing) realizations, normalize each to mean 0 and variance 1, sort them in ascending order and plot the empirical cumulative density function and compare it to the cumulative density function of the standard normal distribution.

(iii) Calculate the variance of the empirical mean for normal-distributed random variables: Generate M normal-distributed data sets of N data points each. For each data set, compute the empirical mean. Then compute the variance of the M means. Plot the variance against the size of N. What do you observe? Which law or theorem is illustrated by this exercise?

Hint: You should take numbers in the range of $M = 100$ and $N = 1, 2, \ldots, 100$.

(iv) Repeat (iii) for Cauchy-distributed data. Discuss the difference of the results from (iii) and (iv). How does the result relate to exercise (ii)?

Exercise 2: Empirical variance

(i) Draw 1000 replicates of normally distributed numbers with mean 5 and standard deviation 2, with N realizations each. Thereby, $N = 2, \ldots, 10000$ (in reasonable spacing). Calculate the empirical variance of each of the 1000 distributions

- with normalization $\frac{1}{N}$, and
- with normalization $\frac{1}{N-1}$ (Bessel correction).

Compare the variances in an appropriate plotting style to the variance that you specified for drawing realizations. What do you observe?

(ii) Draw 1000 replicates of a $N(5,1)$ distribution, with $r = 3, 5, 10, 20, 50, 200$ points each. Calculate the quadratic sum of each of the 1000 distributions, given by $QS = \sum_{i=1}^{r}(x_i - \bar{x})^2$. Plot a histogram of QS and compare it to a χ^2 distribution with degrees of freedom equal to the number of replicates. How does a normal distribution with mean r and variance $2r$ fit in the picture and what limit theorem do you observe?