Statistical Data Analysis Problem sheet 4 Due Monday 7 November 2022

Please turn in a copy of the parts of the source code you wrote or modified and relevant output. Please do not turn in any large parts of source code that you use but did not modify.

**Exercise 1(a)** [3 marks] Let  $x_N = \sum_{i=1}^N r_i$ , where the  $r_i$  are independent and uniformly distributed between 0 and 1. Find the mean  $\mu_N$  and standard deviation  $\sigma_N$  of  $x_N$  as a function of N. Using these ingredients show that the standardized variable  $y_N = (x_N - \mu_N)/\sigma_N$  is

$$y_N = \sqrt{\frac{12}{N}} \left( \sum_{i=1}^N r_i - \frac{N}{2} \right) .$$

1(b) [7 marks] Using the simpleMC program (either C++ or Python) from problem sheet 3 as a starting point, write a computer program to make and plot histograms of 10000 values of  $y_N$  as defined above for N = 1, 2, 4, 12. Make sure to set the limits of the histogram such that the entire distribution is plotted.

In the program, find the mean and standard deviation of  $y_N$  and verify that these are close to the values you expect for N=1,2,4,12. In python, you can put the values in a numpy array and use the functions mean and std. If you are using C++ with ROOT, you can use the GetMean and GetStdDev functions of the TH1D class (see https://root.cern.ch/doc/master/classTH1.html).

Comment on the connection between your histograms and the central limit theorem.

**Exercise 2** Consider the pdf  $f(x) = 3x^2$ ,  $0 \le x \le 1$ .

- **2(a)** [5 marks] Use the transformation method to find the function x(r) to generate random numbers according f(x). Implement the method in a short computer program and make a histogram with 10000 values.
- **2(b)** [5 marks] Write a program to generate random numbers according to f(x) using the acceptance-rejection technique. Plot a histogram of the results.