Lab session 1: The Central Limit Theorem

MATH 350.011

Your task for this session is to write a set of R codes to:

- create a sample from a given probability distribution (either continuous or discrete)
- simulate the values of the mean of the various samples of size n and visualize them in a histogram

1 Part 1: Drawing samples from a distribution

Choose ONE of the following two examples

1.1 Example 1 (Continuous random variable)

Let a be the last non-zero digit of your student ID (either one if you are working in a group of 2).

Let X be a continuous random variable with the following probability density function

$$f(x) = \begin{cases} cx & \text{if } 0 \leq x \leq a \\ 0 & \text{elsewhere} \end{cases}$$

where c is some unknown constant.

- Compute the value of c and write down the formula for the distribution function $F(x)$ of X
- Write a set of R codes to simulate a sample of 2000 random draws from the distribution described above
- Compute the mean and produce a histogram of the sample

1.2 Example 2 (Discrete random variable)

Let X be a discrete random variable with the following probability mass function

$$p(x) = \begin{cases} cx & \text{if } x = 1, 2, 3, 4 \\ 0 & \text{elsewhere} \end{cases}$$

- Compute the value of c
- Write a set of R codes to simulate a sample of 2000 random draws from the distribution described above
- Compute the mean and produce a histogram of the sample

2 Part 2: Central Limit Theorem

- Write a set of R codes to repeat Part 1 (without doing histogram) 2000 times and record the mean of each sample in a vector v
- Produce a histogram of the mean

If you still have time, work one the other example in Part 1.

If you still have time and you're bored, move to Part 3.

3 Part 3: Simulate bivariate distribution by conditional probability

Assume that we are interested in sampling two continuous random variables X and Y with the joint probability density function

$$f(x,y) = \begin{cases} x+y & \text{if } 0 < x < 1,\ 0 < y < 1 \\ 0 & \text{elsewhere} \end{cases}$$

How do we do this?

First, compute the marginal probability distribution of X. Can you do that? Good. Use that marginal pdf to sample an instance of X, as in part 1.

Second, after obtain the value x, compute the conditional probability of Y given that $X = x$, that is

$$f_{Y|X}(y|X = x)$$

Use this conditional pdf to sample an instance of Y, as in part 1.

Repeat the task 2000 times to obtain 2000 pairs (x, y). Use the function *plot* to produce a scatter plot of the samples.