1 Write a function that takes as an argument :

- a realization $(x_1,\dots,x_n)$ of a sample of the $\mathcal{N}(\mu,\sigma^2)$
- a level of confidence

and that outputs the corresponding observed confidence interval for μ (when σ^2 is unknown).

2 Repeat exercise 1 for the mean of a sample of an unknown distribution.

3 Repeat exercise 1 for the variance of a $\mathcal{N}(\mu,\sigma^2)$ distribution with unknown expected value μ

4 Simulate a sample of the $\mathcal{N}(\mu,\sigma^2)$ of size 20 with randomly chosen parameters μ and σ^2 and build the confidence interval from exercise 1 : is μ in the confidence interval?

5 Build a "large number" of confidence intervals using the functions from exercises 1, 2 and 3. In average, how many times did you give a confidence interval containing the true value of the parameter?

6 At an election, 49% of the electors are going to vote for candidate A and 51% for candidate B. A survey is conducted in order to give a good approximation of those values without consulting everybody. We will denote by :

$$X_i = \begin{cases} 1 & \text{if the } i\text{-th person votes for candidate } A, \\ 0 & \text{otherwise.} \end{cases}$$

We will assume that the X_i's are mutually independent random variables with the same distribution $\mathscr{B}(0.49)$.

1. Write a function that simulates a sample $(x_1,\dots,x_n)$ of voters.
2. Write a function that gives an estimation of the proportion of persons that are going to vote for candidate A for a given sample.
3. Usually, pollsters take samples of size $1,000$: simulate a "large number" of estimations using the previous question and calculate the number of times that your sample gave candidate A as a winner instead of candidate B.

7 *The placebo effect : how to cheat customers*

Introduction

You are working for a pharmaceutical company who wants to release a new anti-aging skin cream. They know that such products strongly rely on the placebo effect and they want to take advantage of it. Thus, the product they are trying to sell will be a simple moisturizing cream.

Advertising is an important matter when it comes to cosmetics. Your pharmaceutical company's goal is to be able to advertise that "their product showed a significative improvement after one week for 95% of the women who tested the cream".

In order to do so, they will make N different products by slightly changing the concentration of water in each one of them. They will then send each product to a group of 30 women. Each woman in every group will be asked to use the cream and auto-evaluate its effects after one week.

Your job is to tell the company how to chose N so that the probability for at least one of their product to satisfy 95% of the women in at least one group is greater than 99%.

The model

We will assume that the probability for a woman to give a good review to a moisturizing cream is equal to 0.5. We will denote by $X_{i,j}$ the evaluation of the product by the j-th woman in the i-th group. To simplify, we will write :

$$X_{i,j} = \begin{cases} 1 & \text{if the review is good,} \\ 0 & \text{otherwise.} \end{cases}$$

What values will the integers i and j take ? How can one write that 95% of the women from group 1 are satisfied ?

Finally, we will denote by Y_j the variable :

$$Y_i = \begin{cases} 1 & \text{if } \sum_j X_{i,j} \geqslant n_0, \\ 0 & \text{otherwise,} \end{cases}$$

How can you chose n_0 such that Y_i equals one when the group i had a 95% satisfying rate? Using a simulation, find N.