Exercise 1 Binomial distribution

Recall the example (mean number of recovered patients $\hat{p} = \frac{k}{n}$) from the lecture.

A total of n = 20 patients are examined to test whether or not a new drug yields a probability of recovery higher than $p = 0.4$. The number k of recovered patients follows a binomial distribution with $n = 20$ (`number of trials`, i.e. patients) and $p = 0.4$ (`probability of success`, i.e. probability of recovery).

1. What is the probability of observing $k = 13$ recoveries?
   Hint: Look at the help pages `?dbinom`.
2. Save the probabilities for each possible outcome (`k <- 0:20`) in a vector `probs` and plot the resulting probabilities using the `plot` command.
   Hint: Pay attention to the x-axis and try the argument `type="h"`.
3. What is the probability $\text{P}(k \geq 13)$ of observing 13 or more recoveries?
   Hint: You can do it manually or use the function `pbinom(q, size, prob)`.
4. How does $\text{P}(k \geq 13)$ change when the recovery probability is $p = 0.1$ or $p = 0.9$?

Exercise 2 Law of large numbers

Consider a random variable with a finite expected value which is repeatedly sampled. The law of large numbers states that as the number of these observations increases, the sample mean will tend to approach and stay close to the expected value (the average for the population).

Consider the flip of a coin. Given repeated flips of a fair coin ($p = 0.5$), the frequency of heads (or tails) will approach 50% over a large number of trials.

1. Draw $N = 10$ realizations x_i of a binomial distribution with `size` = 1 and `prob` = 0.5 and save them in a vector `x`.
   Hint: Use the function `rbinom(N, size=1, prob=0.5)`.
2. Repeat 1. a few times and look at the average $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$. What can you see?
3. Plot the outcomes `x` and draw a horizontal line at the expected value 0.5 using `abline(h=0.5)`.
4. Add a line which shows the running average $\frac{1}{nn}\sum_{i=1}^{nn} x_i$ for $nn = 1, \ldots, N$. To do so, create a vector `nn <- 1:N` and compute the cumulative sum of x_i with `csum <- cumsum(x)`. Add the line to the existing plot with `lines(csum/nn, col="red")`.
5. Repeat the above steps with $N = 100$ and $N = 1000$ realizations.
6. What happens to the average if we start with a sequence of 10 times heads (i.e. 1), and then flip the coin another $N = 1000$ times?