

# Numerical simulations

Christophe Pallier

Sept 2013

## Contents

The aim is to assess how well the proportion in a sample estimates the proportion in a population.

1. Create a population of  $N=1$  million voters out of which 52% support a candidate A, and the rest support candidate B. Use a numpy array.

```
import numpy as np
N = 1000000
p = .52

pop = np.zeros(N)
pop[0:int(N*p)] = 1
np.mean(pop)
```

2. Extract a random sample of  $s=100$  voters and compute the proportion of A supporters.

```
s = 100

import random
samp = random.sample(pop, s)
np.mean(samp)
```

3. Rerun the preceding process 1000 times and plot an histogram of the estimated proportions (using matplotlib.pyplot.hist).

```
sampmeans = [ np.mean(random.sample(pop, s)) for i in range(1000) ]

import matplotlib.pyplot as plt
plt.hist(sampmeans)
```

Exercise: Modify  $s$ , the size of the sample, and compute the standard deviation of the sampled means as a function of  $s$ .