

### Optional coding homework 1

A .csv file with a list of 4000 zeros and ones was created using a Python 3 code. The list is a sample from the following parametric statistical model. The model has 3 unknown parameters  $(n, p, q)$ , where  $n$  is an integer  $10 \leq n \leq 20$  and  $p \in [0, 1/2)$  and  $q \in [1/2, 1)$ . Consider i.i.d. Bernoulli sequences  $X = (X_1, X_2, \dots)$  and  $Y = (Y_1, Y_2, \dots)$  with parameters  $p$  for  $X$  and  $q$  for  $Y$ . We define an independent sequence  $Z$  as follows:

$$Z = (X_1, \dots, X_n, Y_1, \dots, Y_n, X_{n+1}, \dots, X_{2n}, Y_{n+1}, \dots, Y_{2n}, X_{2n+1}, \dots).$$

The sample provided in the .csv file consists of the first 4000 entries of  $Z$ . Your task is to estimate  $n, p, q$ . Furthermore, to know how good your estimates are, assuming  $\hat{n}$  has the right value, give approximate 95% confidence intervals for  $\hat{p}$  and  $\hat{q}$ .

How would you modify the estimators if the parameter space was changed to  $p, q \in [0, 1]$ ?