Homework 1

Submission Instructions

- Homework is due on: **Sunday 11/08/19 23:55**.
- Homework should be done **only in pairs**. Each pair is to do their own work, separate from the other pairs.
- We prefer you type your submission, however, you may submit scanned handwritten material as long as it is **clear and readable**.
- Submit **only one** PDF file. Please **write your ID** on the top of the file.
- Submission is done via **Moodle** website.
- Homework can be done using either MATLAB or Python.

Question 1

A prisoner is confined in a cell that has three doors, behind each door there is a tunnel. One tunnel leads to freedom after a walk of three hours, a second tunnel leads back into prison after a walk of five hours and the third tunnel leads back to prison after a walk of seven hours. Find the expected time it will take the prisoner to obtain freedom, assume that the chance to choose any of the doors is always the same (the prisoner cannot remember what he had chosen last time).

Question 2

Consider a set of iid samples $D = \{x_i\}_{i=1}^n$ drawn from the following distribution

$$p(x|y) \sim \mathcal{N}(y, 1).$$

(a) Assuming y is unknown constant parameter, compute the MLE of y denoted by $\hat{y}_{\text{MLE}}$.

(b) Assume that $y \sim \mathcal{N}(z, 1/s)$ where s and z are parameters. Compute the MAP estimator $\hat{y}_1$. Express it as a function of $\hat{y}_{\text{MLE}}$.

(c) Now assume that $y \sim U[-1, 1]$. Compute the new MAP estimator $\hat{y}_2$.

Question 3

A discrete random variable X is said to have a Poisson distribution with parameter $\lambda > 0$, if, for an integer $k \geq 0$, the probability mass function of X is given by

$$p(k) = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Note that $E[X] = Var(X) = \lambda$.

(a) Consider a set of iid samples $D = \{x_i\}_{i=1}^n$ drawn according to the Poisson probability mass. Compute the estimator $\hat{\lambda}_{\text{MLE}}$. Is the estimator biased?

(b) Define $y = P(X = 0)^2 = e^{-2\lambda}$. Given a single sample ($n = 1$), compute the estimator $\hat{y}_{\text{MLE}}$. Show that the estimator is biased.

(c) Find an unbiased estimator y_U (Hint : Write the expression of the unbiased estimator and find the term that satisfies the equation).

(d) Recall that the MSE of an estimator $\hat{\theta}$ for parameter θ is defined by $E[(\hat{\theta} - \theta)^2]$ where the expectation is calculated with respect to the distribution of the samples that $\hat{\theta}$ is dependent on. Plot the MSE of the two estimators (MATLAB or Python). Which estimator has a lower MSE for all values of $\lambda \geq 0$?

Question 4

Consider a random Gaussian vector $X \sim \mathcal{N}(\mu, \Sigma)$. The ML estimators for the expectation and covariance matrix of X are given by

$$\hat{\mu}_{\text{MLE}} = \frac{1}{n}\sum_{i=1}^{n}, \quad \hat{\Sigma}_{\text{MLE}} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\mu}_{\text{MLE}})(x_i - \hat{\mu}_{\text{MLE}})^T.$$

Show that $\hat{\Sigma}_{\text{MLE}}$ is biased and propose an unbiased estimator of Σ.

Question 5

Consider k samples taken with equal probability without repetition from the range $[1, 2, ..., N]$ where $k < N$.

(a) Compute the ML estimator $\hat{N}_{\text{MLE}}$ of N. Notice that at each step the sampling is done from a smaller group since the sampling is done without repetition.

(b) The cumulative distribution function (CDF) of $\hat{N}_{\text{MLE}}$ is given by

$$P(\hat{N}_{\text{MLE}} \leq i) = \begin{cases} 0, & i < k, \\ \frac{\binom{i}{k}}{\binom{N}{k}}, & k \leq i \leq N, \\ 1, & N < i. \end{cases}$$

Prove that $\hat{N}_{\text{MLE}}$ is biased. You may use the following equality $\sum_{i=k}^{N}\binom{i}{k} = \binom{N+1}{K+1}$.

(c) Find an unbiased estimator $\hat{N}_U$ of N.

(d) Simulation exercise:

(1) Draw $k = 5$ samples without repetition from $[1, 2, ..., N]$ where $N = 300$. You may use *randperm* in MATLAB.

(2) Evaluate the value for both estimators $\hat{N}_{\text{MLE}}$ and $\hat{N}_U$.

(3) Repeat step (1) and (2) for $r = 10,000$ repetitions and compute empirical expectation (average) of each estimator.

Compare the results.