Exercises

E1. Your favourite auntie has given you a huge bag of assorted chocolates (thousands of them, in fact) as a Christmas present. Your favourite flavor is sweet almond, and you decide to investigate the proportion of such chocolates in the population represented by the bag. You pick randomly 10 pieces out of the bag and eat them to see if they contain sweet almond. Your observations are: 0,1,0,0,0,0,0,0,0,0, where '0' represents the case with no almond, and '1' the case with sweet almond filling in the chocolate.

Write down the likelihood and log-likelihood functions for the observed data, present them graphically and investigate their behavior in the neighborhood of the maximum likelihood estimate of the proportion of sweet almond flavored chocolates in the populations.

E2. You decide to make prediction about the number of sweet almond flavored chocolates in a future sample using the maximum likelihood estimate obtained in the previous exercise. Assume you would pick again 10 chocolates out of the bag as a future sample. Calculate the expected value and variance for the number of sweet almond flavored chocolates in the sample, using the previously obtained maximum likelihood estimate. Calculate also the probability that the number of sweet almond flavored chocolates would exceed 3. Think about the quantities you just have calculated, how well do you expect them to reflect the uncertainty you are facing?

E3. Consider a box that contains θ tickets labeled with the integers $1, ..., \theta$. You draw n tickets with replacement and observe the labels (integers). Formulate the likelihood function for the sample and identify a sufficient statistic (that is not equal to the complete sample). Consider maximum likelihood estimation in the case where the sample contains the integers 23, 10, 15, 2, 30. Investigate the behavior of the likelihood function.

E4. Continue with the distribution in E3. Consider a population with a relatively large θ. Create repeated random samples (e.g. of sizes 10 and 50) from the distribution and illustrate the behavior of the maximum likelihood estimate over the replicates (e.g. with a histogram and some numerical descriptives). Hint. A sample of size n is easily created in Matlab by the command ceil(rand(n,1)*θ).

E5. Let $x_1, ..., x_n$ be conditionally independent realizations from the Exponential(λ)-distribution. Write down the likelihood and log-likelihood functions, and derive a sufficient statistic for λ, as well as the maximum likelihood estimate. Investigate how the log-likelihood function behaves as a function of λ and n, for a range of suitable values of the sufficient statistic.

E6. Derive the Fisher information for the statistical models: Normal(μ,1), Normal(0,σ^2), Normal(μ, σ^2), assuming that in each case a sample of n values will be obtained from the corresponding distribution.