ASSIGNMENT 6 — Advanced Techniques I

For all questions below, provide all programming code and plots in the report. Unless stated otherwise, assume $\alpha = 0.05$

Maximum Likelihood

1. You have collected the following data on a group of individuals: [8.453532, 10.025041, 11.495339, 9.367600, 8.333229, 9.788753, 10.883344, 10.543059, 9.869095, 10.799819]. (10 marks).

   a. Work through the math of calculating the Normal distribution MLE estimate for practice. No need to show this work.

   b. What is the MLE assuming a Normal distribution. (2 mark)

   c. Use a brute force method to find the maximum likelihood estimates. (4 mark).

      . vary μ from 5 to 15 and σ from 0.5 to 1.5, both in increments of 0.1.

   d. Use an optimizer to find the maximum likelihood estimates. Start your initial guesses as $\mu = 8$ and $\sigma^2 = 0.5$ (4 mark).

      . hint: https://stackoverflow.com/questions/40620277/r-how-can-the-nlm-function-be-used-for-optimization-with-multiple-variables

Bayesian Statistics

2. Point Probability (5 marks):

   a. As in the lecture example, let's say you are getting COVID tests and want to know the probability that you have caught it. Initial conditions: $p(+covid) = 0.5$, $p(+test|+covid) = 0.6$, $p(+test|-covid) = 0.35$. You observe 3 positive tests in row. List the probability that you have covid for each test (2 mark).

   b. Change your initial prior to 0.01. List the probability that you have covid for each test. What happens and why (0.5 mark)?

   c. Change your initial prior to 0.0. List the probability that you have covid for each test. What happens and why (0.5 mark)?

   d. Once again, lets assume $p(+covid) = 0.5$, $p(+test|+covid) = 0.6$, $p(+test|-covid) = 0.35$. This time, estimate the probability $p(+covid|-test)$ given you observe 3 negative tests in a row. Tip, $p(-test|+covid)$ is the complement of $p(+test|+covid)$, and $p(-test|-covid)$ is the complement of $p(+test|-covid)$ (2 mark).

3. Continuous Probability (5 marks):

a. The conjugate prior of a Normal distribution is a Normal distribution. This can be summarized as follows:
Prior = $\mathcal{N}(\mu_0, \sigma_0^2)$; Likelihood = $\mathcal{N}(\mu_1, \sigma_1^2)$; Posterior = $\mathcal{N}(\mu_2, \sigma_2^2)$

$$\mu_2 = \frac{1}{\frac{1}{\sigma_0^2} + \frac{n}{\sigma_1^2}} \cdot \left(\frac{\mu_0}{\sigma_0^2} + \frac{\sum_{i=1}^{n} x_i}{\sigma_1^2} \right)$$

.

$$\sigma_2^2 = \left(\frac{1}{\sigma_0^2} + \frac{n}{\sigma_1^2} \right)^{-1}$$

These formula allow us to get an estimate μ_2 (given σ_0, σ_1, and x_i are known). Use the same data from Question 1 above. Assume $\mu_0 = 11.0$, $\sigma_0 = 1.0$, and $\sigma_1 = 1.0$. Calculate the posterior after observing all ten individuals (2 points).

b. Plot the prior and posterior. Make sure you consider σ_2 vs. σ_2^2 when plotting (1 marks).

c. Why do our estimates of μ differ between Question 1 and 3 (1 mark)?

d. What does the posterior represent and how does this differ MLE (1 mark)?