THE UNIVERSITY OF BRITISH COLUMBIA
DEPARTMENT OF STATISTICS
Stat 305 Introduction to Statistical Inference
Lab 5 – Solution

The goal of this laboratory is to practice the main procedures and concepts involved in Maximum Likelihood Estimation.

1. (2 marks) Let Y_i, for $i = 1, ..., n$, be independent random variables such that:

$$Y_i \sim N(\theta x_i, 1)$$

Assume that $\{x_i\}$ are **known constants** and they are not all 0, while θ is an **unknown parameter** that we want to estimate. Notice that the Y_i are **NOT** identically distributed.

   (a) Show that the joint likelihood function can be written as $f_{Y_1,...,Y_n}(y_1, ..., y_n|\theta) = ce^{\sum -\frac{1}{2}(y_i-\theta x_i)^2}$, where c is a constant that does not depend on θ.

   Solution:
   $f_{Y_1,...,Y_n}(y_1, ..., y_n|\theta) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y_i-\theta x_i)^2} = ce^{\sum -\frac{1}{2}(y_i-\theta x_i)^2}$

   (b) Hence, argue that we can write the joint log likelihood function as:

$$l(\theta) = lnf_{Y_1,...,Y_n}(y_1, ..., y_n|\theta) = c' + -\frac{1}{2}\sum(y_i - \theta x_i)^2,$$

   where c' is a constant that does not depend on θ.

   Solution:
   Take ln of result in 1. (Write in detail later)

2. (2 marks) Let $\hat{\theta}$ be the Maximum Likelihood **estimate** of θ. Show that:

$$\hat{\theta} = \frac{\sum x_i y_i}{\sum x_i^2}$$

Using the **appropriate notation** seen in class, write down the Maximum Likelihood **Estimator** (MLE) of θ.

Solution:
$l'(\theta) = \sum_{i=1}^n (y_i - \theta x_i)x_i$
$l'(\theta) = 0 \implies \sum_{i=1}^n (y_i - \theta x_i)x_i = 0 \implies \hat{\theta} = \frac{\sum x_i y_i}{\sum x_i^2}$
$\tilde{\theta} = \frac{\sum x_i Y_i}{\sum x_i^2}$
$l''(\theta) = -\sum_{i=1}^n x_i^2 < 0$

3. (2 marks) The MLE of θ, denoted $\tilde{\theta}$ is a random variable and, as such, has properties such as expected value and variance. We are interested in computing variance of $\tilde{\theta}$ in this question.

   (a) Find Var($\tilde{\theta}$). You can use basic properties of the variance operator to find it.

Solution:
$\text{Var}(\tilde{\theta}) = \text{Var}(\frac{\sum x_i Y_i}{\sum x_i^2}) = \frac{1}{(\sum x_i^2)^2} \sum x_i^2 Var(Y_i) = \frac{\sum x_i^2}{(\sum x_i^2)^2} = \frac{1}{\sum_{i=1}^n x_i^2}$

(b) Now use the Fisher Information to find an approximation for $\text{Var}(\tilde{\theta})$. Compare the approximation for $\text{Var}(\tilde{\theta})$ from the Fisher information and the exact value for $\text{Var}(\tilde{\theta})$ obtained in the previous question. Do they coincide *in this case*?

Solution:
$I(\theta) = -E(l''(\theta)) = E(\sum_{i=1}^n x_i^2) = \sum_{i=1}^n x_i^2$
An approximation for $\text{Var}(\tilde{\theta})$ is then: $\frac{1}{I(\theta)} = \frac{1}{\sum_{i=1}^n x_i^2}$
The approximation and the exact value for $\text{Var}(\tilde{\theta})$ coincide in this case.

4. (1 mark) Find a 95% CI for θ. Denote $z_{0.975}$ the 97.5% quantile of a standard normal distribution.

   Solution:

5. (3 marks) In this particular case, we can find more about $\tilde{\theta}$.

   (a) Show that $\tilde{\theta}$ is an unbiased estimator for θ. Is it always the case that MLEs are unbiased?

   (b) Use a result seen in class to show that, *in this particular case*, we can find the **exact** distribution of $\tilde{\theta}$. Provide its distribution with parameter(s). What does it say about the CI previously obtained?

6. (optional. 2 bonus marks) Use a property of MLE's to find the MLE for θ^2. Name the property used. Is this estimator unbiased?

7. (optional. discuss with your lab partners if time allows) We have seen in an clicker question in class that the recipe above isn't the only way to compute MLE's. In fact, taking derivatives is often not possible or practical. Given $X_i \sim U(0, \theta)$, try to do the same we did above to find the MLE for θ. What happens when you attempt to proceed as instructed in question 2? How would you find the MLE in this case?