Tufts University
Department of Mathematics
Spring 2022

MA 166: Statistics

Homework 2 (v1.1) [1]

Assigned Monday 31 January 2022
Due Monday 7 February 2022 at 11:59 pm EDT.

Read the footnotes! They give important information about how to present your answers.

1. Suppose that you have a priori knowledge that the continuous random variable X is normally distributed. You make n experimental measurements of X, and you find that fn of the measurements are (miraculously) exactly equal to $+1$, and the other $(1-f)n$ measurements are exactly equal to -1. The fraction f is greater than $1/2$, so you think that X has a positive mean, but you are really not sure so you decide to do interval estimation to see with what confidence you can make that claim. You may assume that n is large enough that the Central Limit Theorem applies to a good approximation.

   (a) Use maximum likelihood estimators to find estimates of the mean, μ_e, and the standard deviation, σ_e.

   (b) We are going to use μ_e and σ_e as the basis for interval estimation in the remainder of this problem [2]. Suppose that you find you are able to conclude that the mean of X is positive with confidence probability $100(1-\alpha)\%$. Find an expression for n in terms of f and α [3].

   (c) To make the above concrete, suppose that $f = 0.51$. How large does n need to be to achieve 95% confidence that the mean of X is positive?

2. Suppose that you have a priori knowledge that the continuous random variable Y has the following two-parameter probability density function:

$$f_Y(y) = \begin{cases} \frac{1}{\mu-a}\exp\left(-\frac{y-a}{\mu-a}\right) & \text{for } y \geq a \\ 0 & \text{otherwise,} \end{cases}$$

   where it may be assumed that $a \geq 0$ and $\mu > a$.

   (a) Verify that the density function is normalized, and fine the theoretical mean and standard deviation of Y in terms of the parameters a and μ.

   (b) Find maximum likelihood estimators, $\hat{a}_{\text{mle}}$ and $\hat{\mu}_{\text{mle}}$, for the parameters a and μ. Justify your reasoning carefully, particularly for the calculation of $\hat{a}_{\text{mle}}$.

[2] As we shall see later, σ_e is not the optimum value to use for interval estimation, and the sample standard deviation for the normal distribution is usually defined in a different way. We have not yet covered that material however, so for the purposes of this problem you may proceed as instructed.

[3] Feel free to use the z_α notation described in the text on page 298

(c) Now suppose that you take n samples of this data, $y_1, \ldots, y_n$, where n is large enough that the Central Limit Theorem applies to a good approximation. (You may assume that the sample mean that you find is positive.) Find the $100(1-\alpha)\%$ confidence interval for $\mu = E(Y)$ centered at the sample mean. You may leave your answer in terms of $\hat{a}_{\text{mle}}$ and $\hat{\mu}_{\text{mle}}$ worked out in part (b).

3. Larsen & Marx, Problem 5.3.6, page 306

4. Larsen & Marx, Problem 5.3.15, page 307