

Homework 5

Estimation

Instructions

Please answer the following questions in the order in which they are posed. Add a few empty lines below each and write your answers there. **Focus on answering in complete sentences and show work whether we ask for it or not.** You will also need scratch paper/pen to work out the answers before typing it.

For help with formatting documents in RMarkdown, please consult R Markdown: The Definitive Guide. Another option is to search using Google.

Exercises

1. (MIAA basketball) The MIAA05 dataset from the **fastR2** package contains statistics on each of the 134 players from the 2004-2005 season. In this problem we will consider modeling the players' free throw shooting percentage. The variable is called **FTPct**. It is the ratio of free throws made (**FT**) to free throws attempted (**FTA**). Please type `?MIAA05` in the Console for a description of the variables in the dataset.

Since **FTPct** is a proportion, we will consider the Beta distribution:

$$f(x) = \frac{\Gamma(\alpha_0 + \beta_0)}{\Gamma(\alpha_0) \Gamma(\beta_0)} x^{\alpha_0-1} (1-x)^{\beta_0-1} \quad 0 < x < 1.$$

- a. Show that the method of moments estimators of α_0 and β_0 are:

$$\hat{\alpha}_0^{mom} = \bar{x} \left[\frac{\bar{x} - s}{s - \bar{x}^2} \right],$$
$$\hat{\beta}_0^{mom} = \hat{\alpha}_0^{mom} \frac{1 - \bar{x}}{\bar{x}}$$

where $s = \frac{1}{n} \sum_{i=1}^n x_i^2$.

Hint: You can use without proof from homework 1 key:

$$E[X] = \frac{\alpha_0}{\alpha_0 + \beta_0}$$
$$E[X^2] = \frac{\alpha_0(\alpha_0 + 1)}{(\alpha_0 + \beta_0)(\alpha_0 + \beta_0 + 1)}.$$

The remaining two parts involve coding. Be sure to show the code and output, however, suppress warnings and messages.

- b. Make a histogram and a QQplot to assess the goodness of fit.
 - c. Are there any players you should remove from the data before attempting the analysis? Decide on an elimination rule, apply it, and repeat the analysis. Do you like this fit better?
2. (Unbias your estimator) Let $X \sim \text{Binom}(n, \pi_0)$.

- a. Show, with justification, that

$$E \left[\frac{X}{n} \left(1 - \frac{X}{n} \right) \right] = \frac{(n-1) \pi_0 (1 - \pi_0)}{n}$$

- b. Suppose we want an unbiased estimator for $\pi_0 (1 - \pi_0)$. Use your answer from (a) to construct such an estimator.
3. (Bayes estimator) Let X_1, X_2, \dots, X_n be independent Bernoulli random variables drawn from the PMF:

$$f(x) = \begin{cases} (1 - \pi_0) & x = 0 \\ \pi_0 & x = 1 \end{cases}$$

Consider the Bayesian estimator of π_0 ¹:

$$\hat{\pi}_0^{bayes} = \frac{X + 1}{n + 2}$$

where $X = X_1 + X_2 + \dots + X_n$ is $Binom(n, \pi_0)$.

- a. Is $\hat{\pi}_0^{bayes}$ an unbiased estimator of π_0 ? If not, is it asymptotically unbiased?
- b. Is $\hat{\pi}_0^{bayes}$ a consistent estimator?
- c. Based on the sample 1, 0, 1, 0, 1, calculate the value of $\hat{\pi}_0^{bayes}$ for this sample. Also calculate its estimated standard error.
4. (Bias variance trade-off) In this problem we will continue working with the model described in problem 3. Our focus will be on comparing the mean square error (MSE) of $\hat{\pi}_0^{bayes}$ with the MSE of

$$\hat{\pi}_0^{mom} = \frac{X}{n}.$$

- a. Give expressions for the MSE of each estimator. Show your work clearly.
- b. For $n = 10$, plot the MSE of both estimators on the same graph as a function of π_0 . Describe (just visually) the π_0 values for which the MSE is smaller for the Bayes estimator. Repeat for $n = 1,000$. Don't forget to label the plot, and make it easy for the reader to know which estimator is being represented by which curve.

This example illustrates the bias variance tradeoff. Sometimes a biased estimator does better than an unbiased one, if it has a smaller variance!

¹don't worry about how to derive this estimator