Homework 6

11/23/22

Homework 6 is due 12/02/2022 at 11:59 PM. Submit your homework on Canvas as one PDF document.

The PDF version of this assignment can be found here.

You will need to install and load the following pacakges and load the data sets:

```
# install.packages("mlbench") # Uncomment line if you need to install the package.
library(mlbench)
data("BostonHousing")
```

1. Let $X_1, X_2, \ldots, X_n \overset{iid}{\sim} N(\mu_1, \sigma^2)$ and $Y_1, Y_2, \ldots, Y_m \overset{iid}{\sim} N(\mu_2, \sigma^2)$. Find the test statistic for the following hypothesis tests: $H_0 : \mu_1 = \mu_2 = \mu$ and $H_0 : \mu_1 \neq \mu_2$. Use the likelihood ratio test to show that the test statistic is $(\bar{X} - \bar{Y})/(\sqrt{s_p^2(1/n + 1/m)})$ with $s_p^2 = \frac{(n-1)s_x^2 + (m-1)s_y^2}{m+n-2}$.

2. For data $(X_i, Y_i)_{i=1}^n$ and $\hat{\beta} = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2}$, show that Variance of $\hat{\beta}_1$ is $\frac{\sigma^2}{\sum(X_i - \bar{X})^2}$.

3. The `BostonHousing` data set contains housing information for 506 census tracts to described in 14 variables. We are intersted in analyzing the relationship between median house value (`medv`) and the following variables:

- `crim`: per capita crime rate
- `zn`: proportion of residential land zoned
- `indus`: proportion non-retail business acres
- `nox`: nitric oxide concentrations
- `rad`: accessibility to radial highways
- `tax`: property-tax rate

4. From problem 3, fit a model without the variables that were not significantly associated with median house value.

5. Using the model from problem 4, interpret how access to radial highways is associated with median house value.