Stat 240 - Lab 07

Dr. Lloyd T. Elliott

March 16, 2020

Visualisation of Gene Expresssion Data

Question 1a, (10 points): This question is difficult, easier option below. Reproduce the shiny app that I live coded, except with the dataset *GSE21935* instead of the dataset *GSE83294* from NCBI Gene Expression Omnibus. The dataset *GSE21935* concerns gene expression data for samples with and without schizophrenia. In particular:

1. Download the dataset *GSE21935* from NCBI. Extract the gene expression data and the schizophrenia indicator. Select the gene expression for the first 10 genes listed in the file, and select the gene names.

2. In *R*, create a data frame x with column names given by the gene names, and with one row for each sample and entries given by the gene expression.

3. Create a vector y with coordinates giving the schizophrenia indicator, in the same order as the samples are listed in the data frame x. That is: y_i is 1 if the n-th row of x corresponds to a schizophrenic subject and is 0 if the n-th row of x corresponds to a non-schizophrenic subject.

4. Make a Shiny app with a drop down menu listing the column names of x, and two density plots. When a column of x is selected, the left density plot should show a kernel density estimate of the gene expressions for all schizophrenic subjects for the selected gene. The right density plot should show a kernel density estimate of the gene expressions for all non-schizophrenic subjects.

Deploy the app on heroku (using previous instructions), or on the *rcg* Shiny server (as demonstrated in the live code), or just in your *RStudio* terminal. Provide a screenshot of the App working, and of your code. Easier option: if the gene expression data is too hard, complete this question using another dataset. Choose x and y to indicate any dataset (downloaded by you) such that x is a data frame with real numbers (with around 10 columns), and y is binary (*i.e.*, taking values 0/1 or 'true' and 'false', or -1 and 1). If you need inspiration for the dataset, consider this source: `https://archive.ics.uci.edu/ml/datasets.php`

Your First App

Question 2a, (10 bonus points): Work from one of the Shiny App examples here: `https://shiny.rstudio.com/gallery/` and modify it to use a different dataset or modify it in any way of your chosing. Consult the Shiny App documentation to understand how to change the gallery example to implement your idea. Provide a *pdf* with a screenshot of the App running on *heroku* or the *rcg* Shiny server or on your local *RStudio*.