Due Date: October 10th, 2019

Part of this includes R work, so the assignment may be given in two parts : pdf file and paper copy in class.

1. Use `babiesI.data` from CHDS study in the Datasets folder in Dropbox shared with you.
   (a) Plot the `density` (estimate) for weights of babies for Maternal smokers and non-smokers.
   (b) Using the `t.test` decide with a level of significance of 5% whether or not to reject the null hypothesis that the means of weights of babies are the same.
2. The following data is from Kalyani Ramachandran's Lecture.
   (a) The data from human cells and four treatment. The responses for each treatments are given below.

A	B	C	D
606.5	506.5	133	418
713.5	503.5	228	459

   Assume variances are not equal.
      i. Write a function in R called `twosamplettest` that will take two data sets `x` and `y`, perform the t-test for equality of means. The output of the function should be the p- value.
      ii. Using the above `twosamplettest`, perform (three) equality of means test whether: mean of B is the same as mean of A; mean of C is the same as mean of A; and mean of D is the same as mean of A.
   (b) The data is from mice model. The responses for each treatments are given below.

Pre-Control	Pre-A	Post-Control	Post-A
0.37315	0.86272	0.13301	0.523392
0.39038	1.1056	0.67946	0.551693
0.40688	1.03096	0.96005	0.231634
0.09202	1.31211	0.30404	0.630685
0.51352	1.04229	0.63041	0.52417

   Assume variances are not equal.
      i. Enter the above data as a data frame[1] using the `data.frame` command.
      ii. plot the data using `dotplot` in the library `lattice`.
      iii. Using the above `twosamplettest`, perform (three) equality of means test whether: mean of Pre-Control is the same as mean of Pre-A; mean of Post-Control is the same as mean of Pre-A; and mean of Post-A is the same as mean of Pre-A.
      iv. Using the inbuilt R function `oneway.test`, perform a test to see the equality of means across Pre-Control, Post-Control and Post-A.
3. Suppose $X_1, \ldots, X_n$ is an i.i.d. sample with Uniform(a, b) distribution for some unknown a and b. Let m_1 and m_2 be the empirical realisation of the first and second moments of the $X_1, \ldots, X_n$ data. Find an expression for the estimates of a and b given by the method of moments in terms of the quantities m_1 and m_2. Furthe, prove that the method of moments produces estimates such that $a = b$ if an only if every data point in the empirical realisation has exactly the same value.
4. Suppose $X_1, \ldots, X_4$ is an i.i.d. sample with Binomial(n, p) distribution for some unknown N and p. Suppose the empirical realisation of these variables is $1, 2, 5, 12$. Show that the method of moments for estimating N and p gives negative (and therefore meaningless) results.
5. Let $\mu \in \mathbb{R}$ and $\sigma > 0$ and $(X_1, X_2, \ldots, X_n)$ be from a population distributed as Normal with mean p and variance σ^2. Then find the M.L.E. for μ, σ.

[1]Data frames are rectangular sets of data with each column being a variable and each row representing a case. To construct one from the variables themselves we can use `data.frame(variable1= VECTOR1, variable2= VECTOR2)`.