Stat 6045/5145

Homework **2**

**Directions** Please submit (on canvas) your R code along with a brief write-up of the solutions (do not submit raw output).

1. Create a binary operator that will concatenate two strings with a space separating the terms. Write an operator where quotations around the terms are optional. That is create an operator like %&% that gives the following output for quoted and/or unquoted arguments,

> Hello %&%World [1] "Hello World"

> "Hello" %&%"World" [1] "Hello World"

1. Linear models have the form , where is the resonse vector of length , is the vector containing the parameters, is the design matrix (i.e. the matrix with the predictors as columns) of dimension and represents gaussian noise with mean zero and variance . The parameter estimates are ( stands for the matrix multiplication of the transpose of with , and is the inverse of that product), an estimate for the variance is , an estimate for the covariance matrix of the parameter estimates is , the projection (“hat”) matrix is the fitted values are , and the residuals are .

Write a function that takes the response vector and the matrix of covariates as input, and

returns a list of the following:

* **beta**, the vector of least squares estimates,
* **sigma**, the residual standard error,
* **varbeta**, the covariance matrix of the least squares estimates,
* **fitted**, the vector of fitted values,
* **residuals**, the vector of residuals.

Further, put in an option to return **hat**, the projection matrix, upon request. The default should be to not return it. To fit an intercept, the elements in the first column of have to be equal to one, so your function should also have an option to add a vector of ones to the matrix with the predictors. Further, your function should check whether or not is invertible, and stop if it is not.

1. Write a function that generates random normal distribution from the Box-Muller method attached in blackboard HW (the method uses two uniformly random variables to generate two independent standardized normal random variables). Use this function to generate 1000 normally distributed random data with mean 10 and standard deviation 4. Construct a histogram from these 1000 generated data.

You can find information from Wikipedia:

<https://en.wikipedia.org/wiki/Box–Muller_transform>

and/or the attached pdf file.