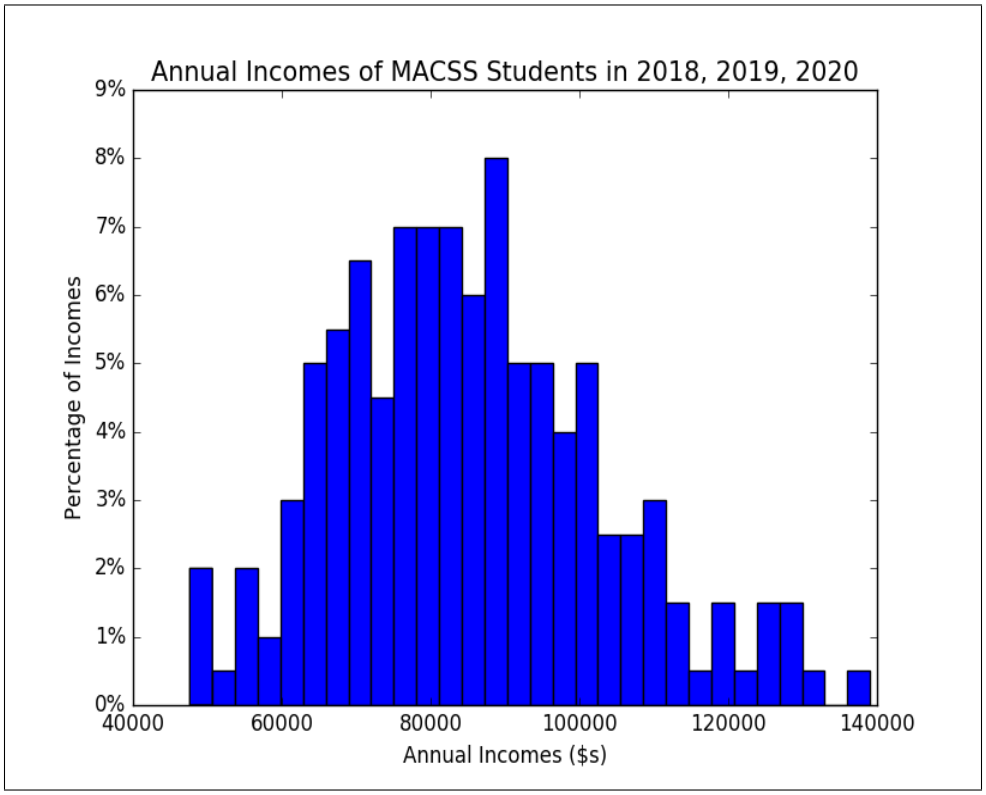
Problem Set #2

MACS 30100, Dr. Evans
Tong Ju

Problem 1 Some income data, lognormal distribution, and hypothesis testing

Part (1-a). Plot a histogram of percentages of the income.txt data with 30 bins.

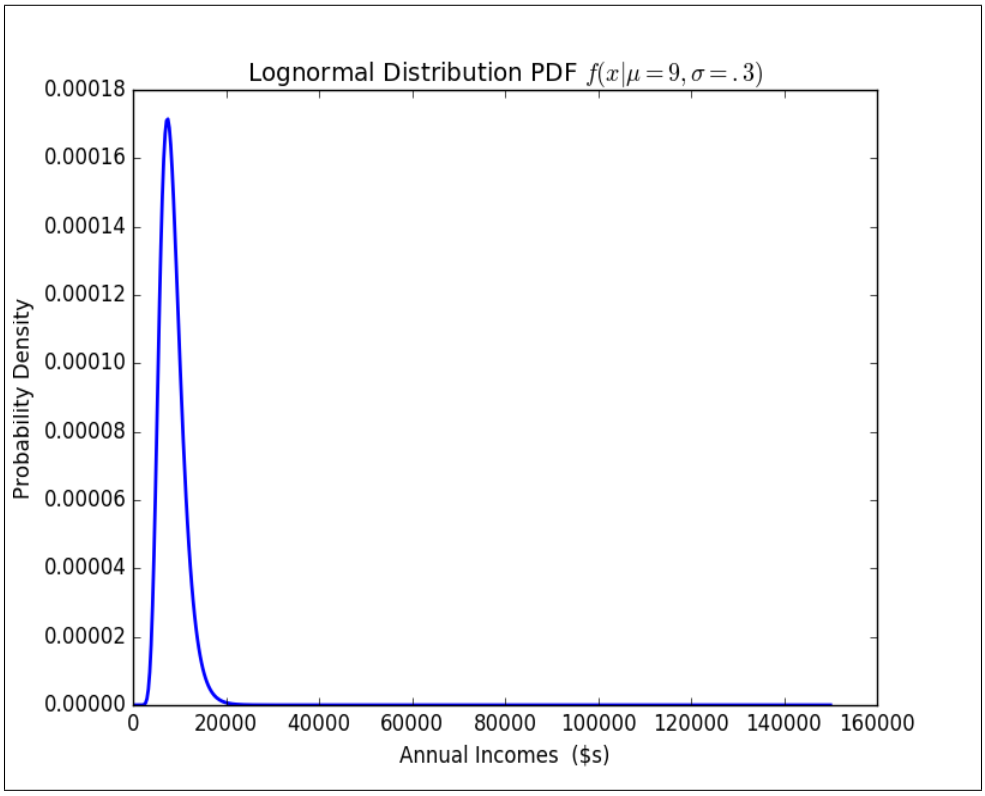
Figure 1: Question 1 part(a)



Part (1-b). Plot the lognormal PDF for the given parameters.

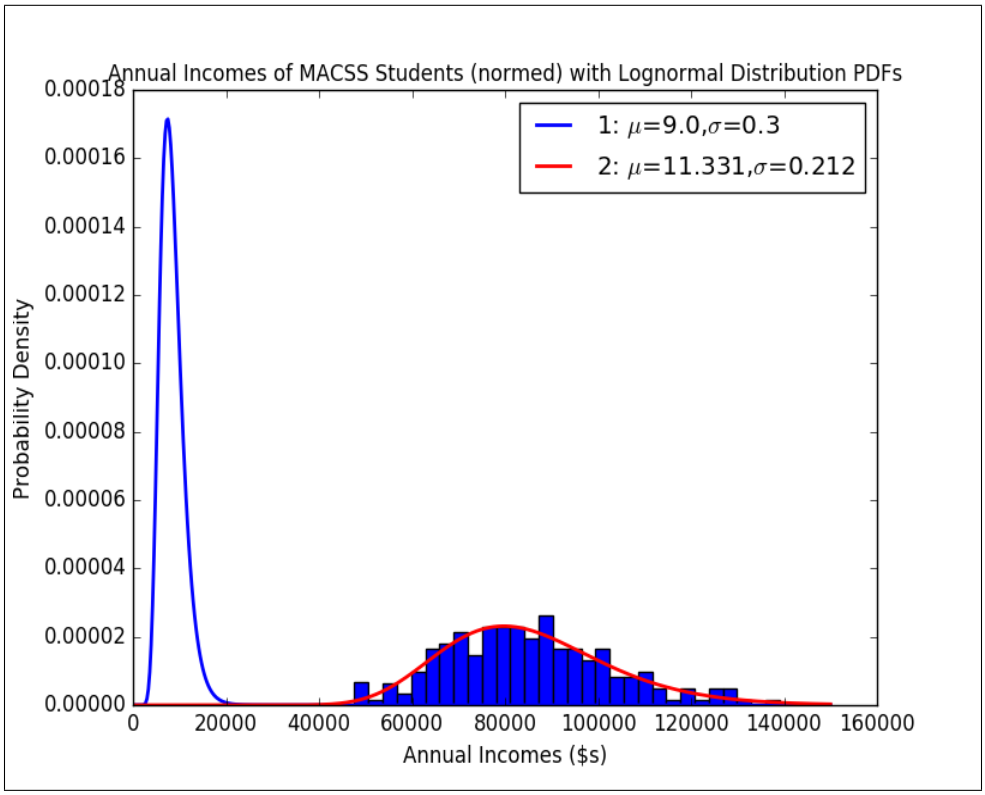
The value of the log likelihood value for this parameterization of the distribution and given this data is -8298.63695601.

Figure 2: Question 1 part(b)



Part (1-c). Histogram Plot with he lognormal distribution by maximum likelihood .

Figure 3: Question 1 part(c)



First, ML estimates for μ_{mle} is 11.3314404077 , and σ_{mle} is 0.211674582579 . The value of log likelihood function is -2239.534744 for the data given these parameters.
Second, the variance-covariance matrix is:

$$\begin{bmatrix} 0.0003153 & 0.0000042 \\ 0.0000042 & 0.0001123 \end{bmatrix}$$

In addition, the Standard error for μ_{mle} is 0.0177566431527; Standard error for σ_{mle} estimate is 0.0105962825978

Part (1-d). Likelihood Ratio Test.

χ^2 of H_0 with 2 degrees of freedom p-value in this likelihood ratio test is 0.0 .
Therefore, it is not likely that the data came from the distribution in part (b).

Part (1-e). Probability Estimation.

According to the estimated distribution of incomes, the probability that MACSS students will earn more than $1000,000 is 0.195766021972 and the probability that they will earn less than $75,000 is 0.307688434489.

Problem 2 Linear regression and MLE

Part (2-a). Estimate the parameters of the model by Maximum Likelihood
The estimated parameters are as follows:

$$\beta_{0(MLE)} = 0.251646226084$$
$$\beta_{1(MLE)} = 0.012933347781$$
$$\beta_{2(MLE)} = 0.400502030729$$
$$\beta_{3(MLE)} = -0.00999166733162$$
$$\sigma_{(MLE)} = 0.00301769300642$$
$$\sigma^2_{(MLE)} = 9.11 * 10^{-6}$$

The maximized log-likelihood is 876.8650474672079
the variance-covariance matrix is:

$$VCV_{(MLE)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Part (2-b). Likelihood Ratio Test
χ^2 of H_0 with 5 degrees of freedom p-value = 0.0 .
This result indicates it is not likely that age, number of children, and average winter temperature have no impact on the sick days.