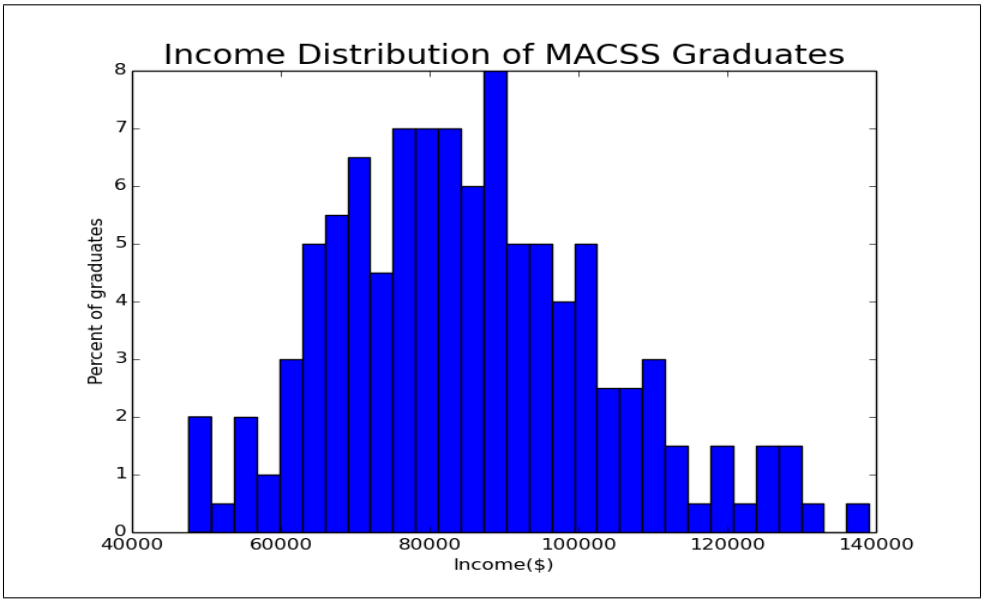
Problem Set #[2]

MACS, Dr. Evans
Yuqing Zhang

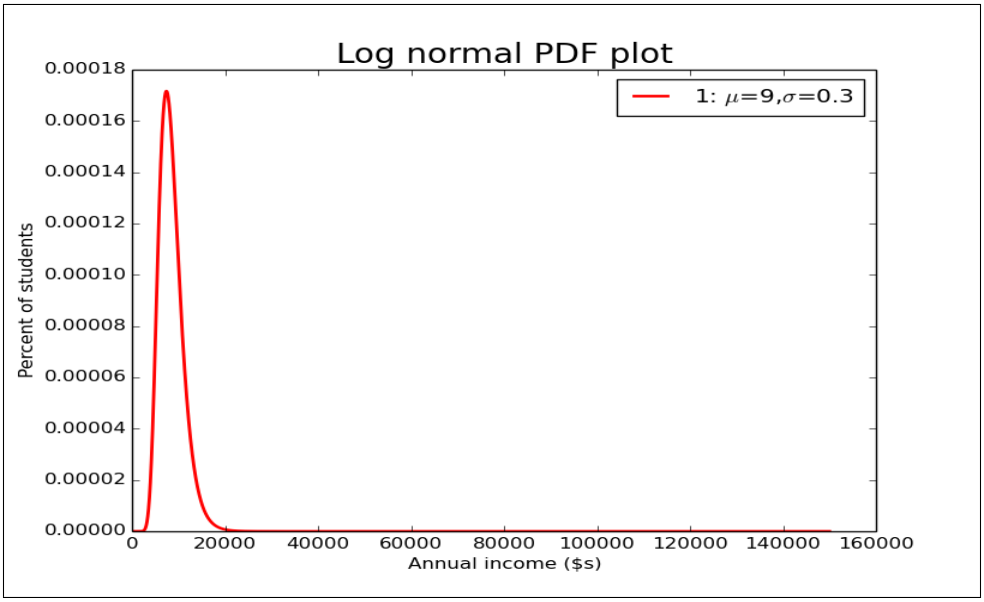
Problem 1 Income data, lognormal distribution and hypothesis testing
Part (1a).

Figure 1: A Histogram of percentages of Income



Part (1b). The value of the log likelihhod value for μ=9.0, σ=0.3, is -8298.64

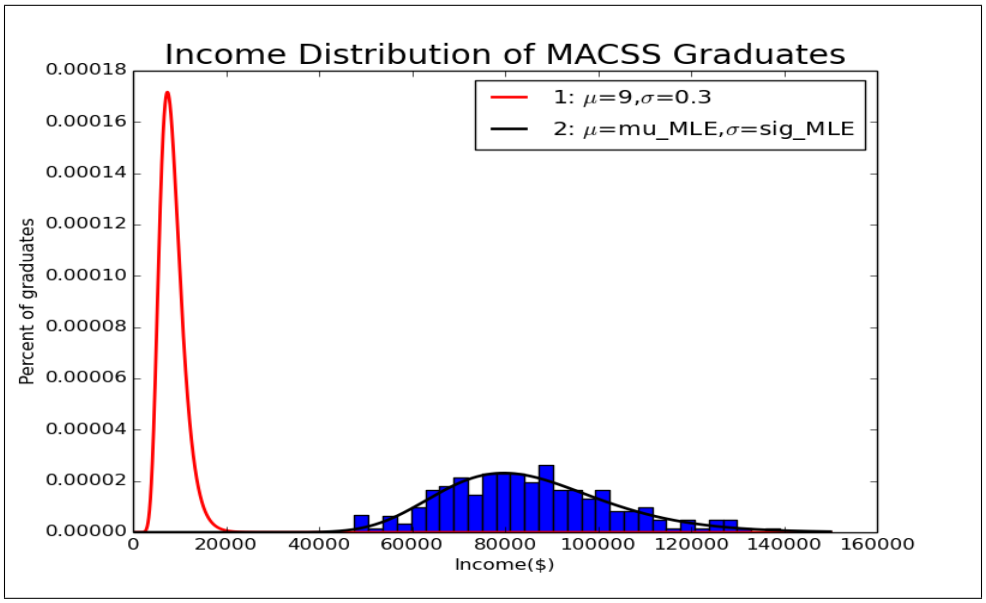
Figure 2: A log normal pdf plot



Part (1c).

ML estimate for μ is: 11.33
ML estimate for σ is 0.21
The value for the ML function is -2239.53
The variance-covariance matrix is :
[[3.15e-04 4.17e-06]
[4.17e-06 1.12e-04]]
The standard error for μ estimate is 0.02
The standard error for σ estimate is 0.01

Figure 3: Comparison between 3 plots



Part (1d).

Chi squared of H0 with 2 degrees of freedom p-value = 0.0. P-value equals to 0, which provides strong evidence against the null hypothesis. We can thus reject the hypothesis that the data in incomes.txt came from distribution in part b.

Part (1e).

The probability that I will earn more than $\$100,000$ is 19.58%
The probability that I will earn less than $\$75,000$ is 30.77%

Problem 2 Linear Regression and MLE

Part (2a).

The ML estimate for σ is 0.003
The ML estimate for β_0 is 0.25
The ML estimate for β_1 is 0.01
The ML estimate for β_2 is 0.4
The ML estimate for β_3 is -0.01
The value of the maximized log likelihood function is: 876.87
The estimated variable-covariance matrix is:

[[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.]]

Part (2b).

Chi squared of H0 with 2 degrees of freedom p-value = 0.0. P-value equals to 0, which provides strong evidence against the null hypothesis. We can thus reject the hypothesis that the age, number of children, the average temperature of winter have no effect on the number of sick days.