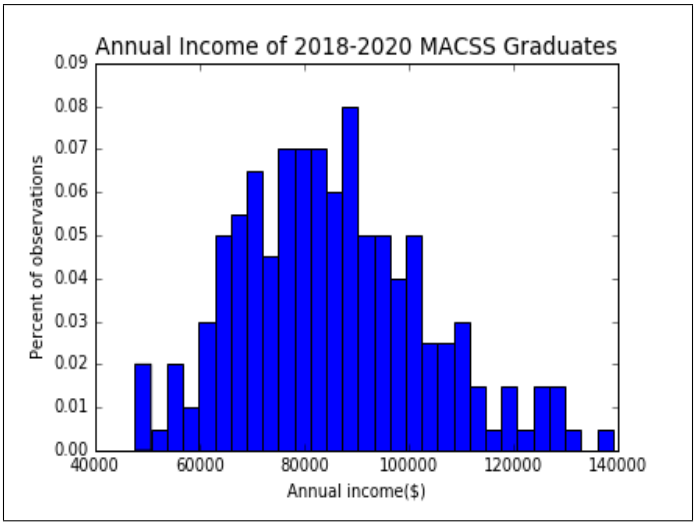
Problem Set #2

MACS 30100, Dr. Evans
Yiqing Zhu

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

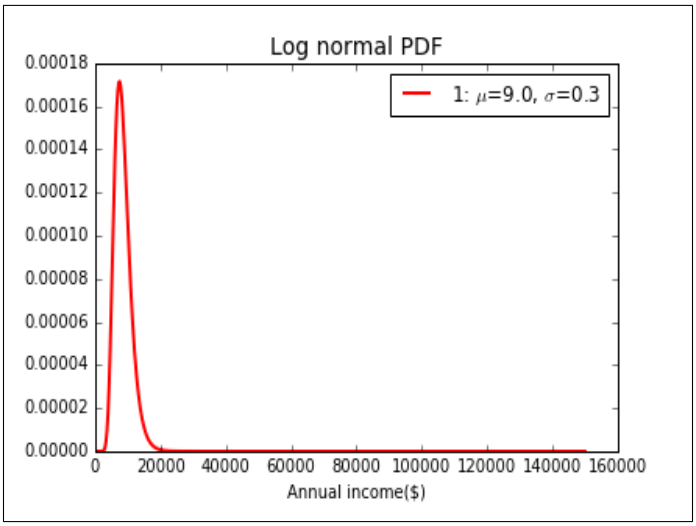
Part (a).

Figure 1: Histogram of the data



Part (b).

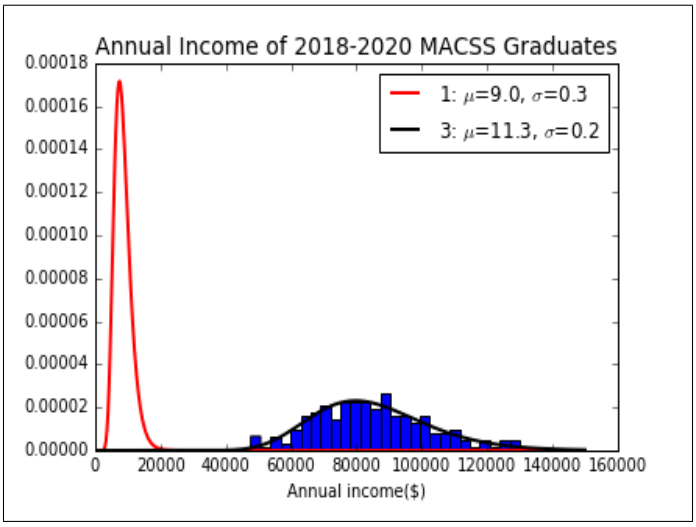
Figure 2: Plot for an estimated distribution



The log likelihood value for this parameterization of the distribution is -8298.64.

Part (c).

Figure 3: Plot for MLE distribution



The ML estimates for μ is 11.33, and for σ is 0.21.

The value of the likelihood function is -2239.53.

The variance-covariance matrix is $\begin{bmatrix} 2.23621817e-04 & -2.06177504e-07 \\ -2.06177504e-07 & 1.11785736e-04 \end{bmatrix}$

Part (d).
Chi squared of H0 with 2 degrees of freedom = 0.0
It is very unlikely that the data in incomes.txt came from the distribution in part (b) since the p-value is 0.

Part (e).
The probability that I will earn more than $100,000 after graduation: 19.58%;
The probability that I will earn less than $75,000 after graduation: 30.77%.

Problem 2 Linear regression and MLE

Part (a). ML estimates parameter: β_0: 0.25, β_1: 0.01, β_2: 0.40, β_3: -0.01, σ: 0.003.

The value of the likelihood function is 876.87.

The variance-covariance matrix is $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

Part (b).

Chi squared of H0 with 5 degrees of freedom = 0.0

It is very unlikely that age, number of children, and average winter temperature have no effect on sick days since the p-value is 0.