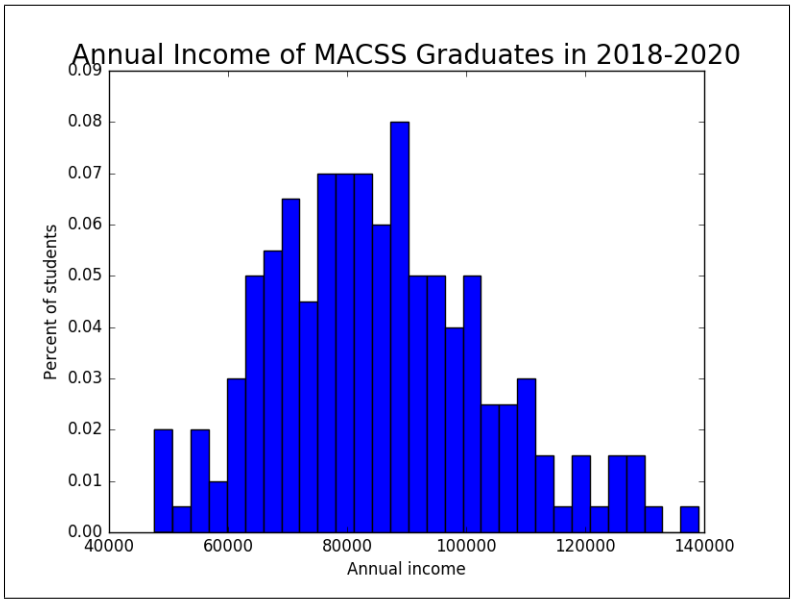
Problem Set #2

MACS 30100, Dr. Evans
Chih-Yu Chiang
Python Version: 3.5.2

Problem 1
Part (a). histogram

A histogram of annual incomes of students who graduated in 2018, 2019, and 2020 from the University of Chicago M.A. Program in Computational Social Science.

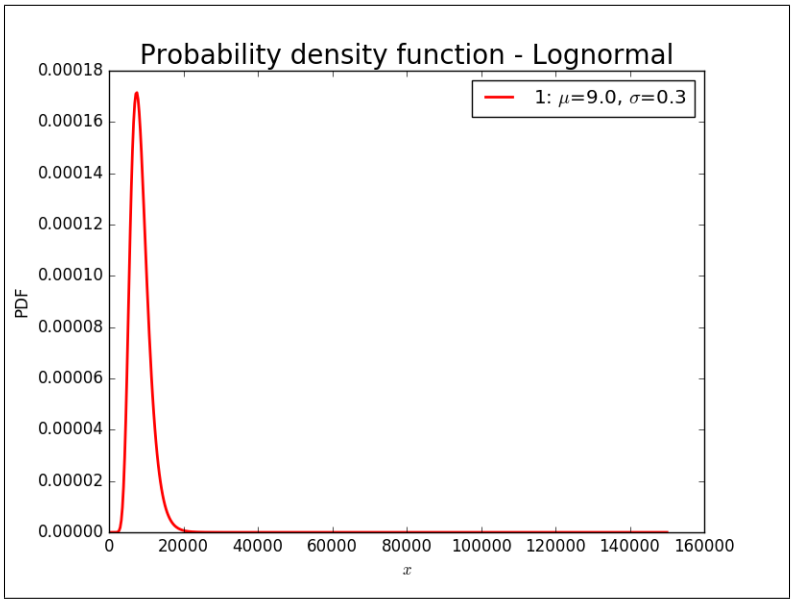
Figure 1:



Part (b). log likelihood pdf

A lognormal pdf with $\mu = 9.0$ and $\sigma = 0.3$, for $0 \leq x \leq 150{,}000$ is ploted.
The log-likelihood value is -8298.636956005032.

Figure 2:



Part (c). estimate distribution

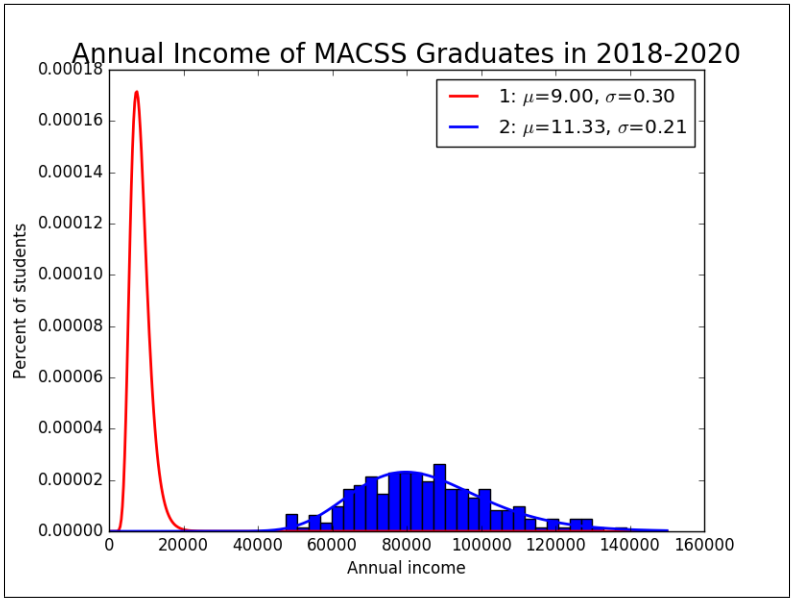
The estimated parameters are as follows:

$$\mu_{(MLE)} = 11.3314403048$$
$$\sigma_{(MLE)} = 0.211674604497$$

The maximized log-likelihood is -2239.5347439980105.

$$VCV_{(MLE)} = \begin{bmatrix} 0.01967716 & -0.0024631 \\ -0.0024631 & 0.00042405 \end{bmatrix}$$

Figure 3:



Part (d). chi squared test

χ^2 of H_0 with 2 degrees of freedom p-value = 0.0 .
That is, the likelihood that the income data came from the distribution is very low.

Part (e). infer proportion

19.58% I will earn more than $100,000.
30.77% I will earn less than $75,000.

Problem 2
Part (a). estimate linear model

The estimated parameters are as follows:

$$\beta_{0(MLE)} = 0.25164631958$$
$$\beta_{1(MLE)} = 0.012933347853$$
$$\beta_{2(MLE)} = 0.400502072911$$
$$\beta_{3(MLE)} = -0.00999167071918$$
$$\sigma_{(MLE)} = 0.00301768213759$$

The maximized log-likelihood is 876.8650464331619.

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

Part (b). chi squared test

χ^2 of H_0 with 5 degrees of freedom p-value = 0.0 .
That is, the likelihood that age, number of children, and average winter temperature have no effect on the sick days is very low.