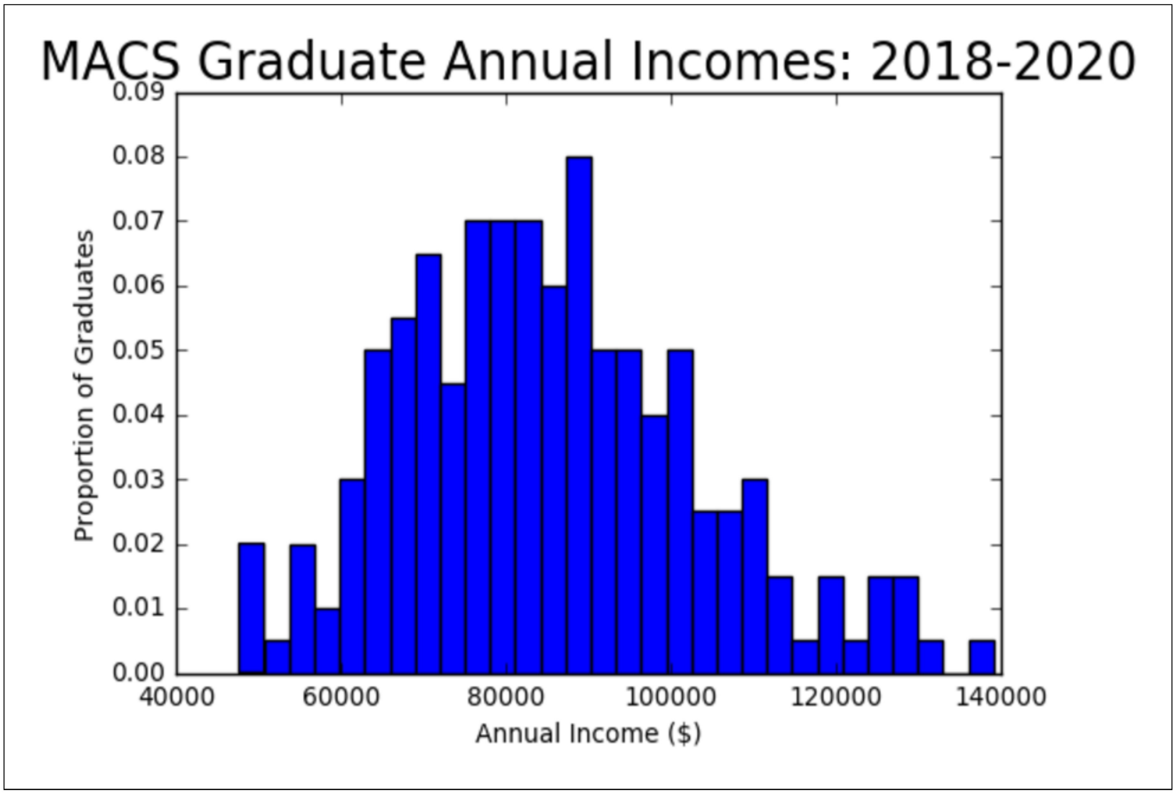
Problem Set #2

MACS 30100, Dr. Evans
Cheng Yee Lim

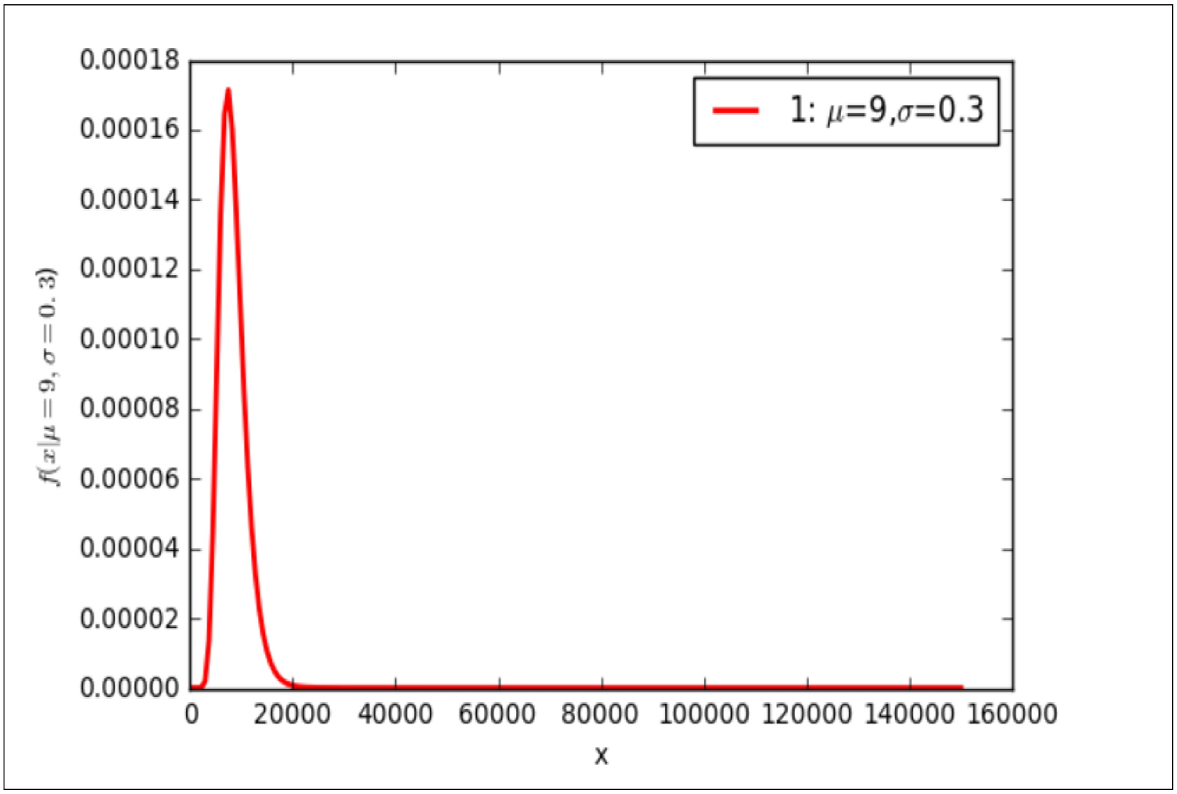
Problem 1
Part (a).

Figure 1: Density histogram of Annual Incomes of MACSS Graduates



Part (b).

Figure 2: Density histogram of Annual Incomes of MACSS Graduates



The log likelihood value for the parameters $\mu = 9$ and $\sigma = 0.3$ is -8298.64. The very negative log likelihood value suggests that this parameters are a bad fit for the data given.

Part (c).

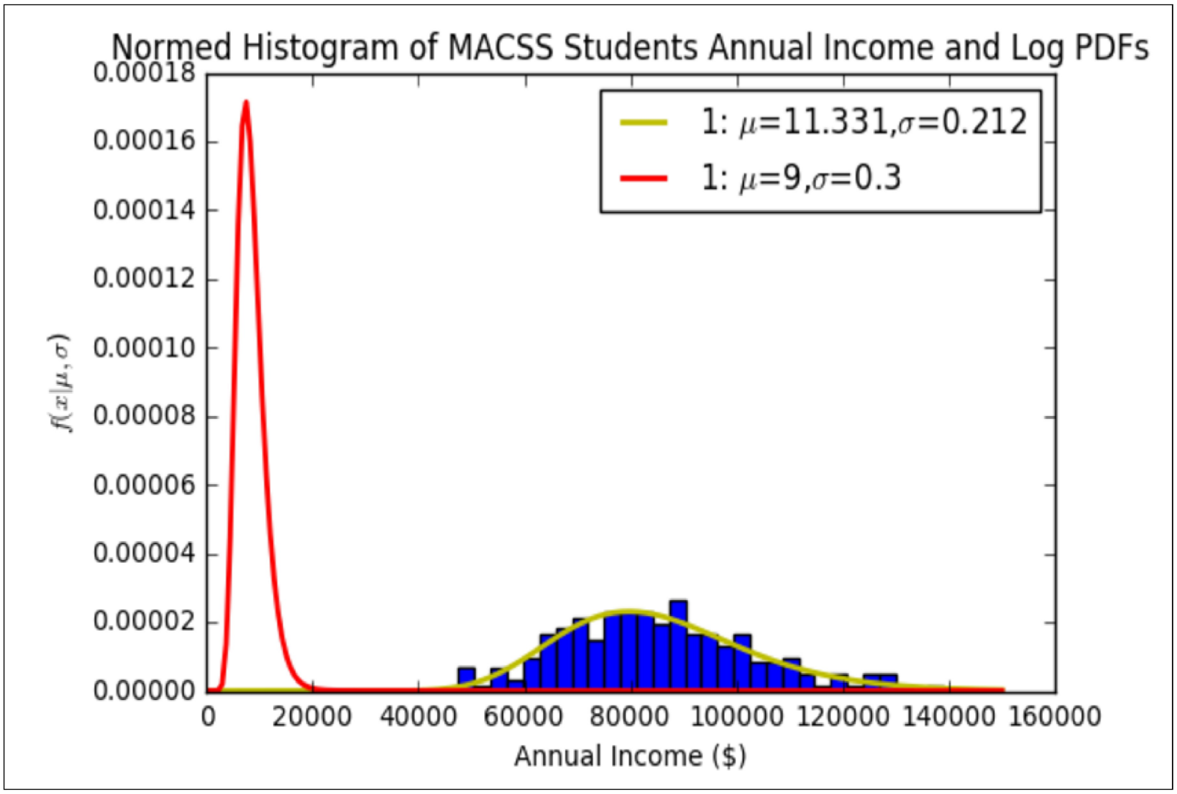
The ML estimates for μ and σ are 11.331 and 0.212.

The log likelihood value for $\mu = 11.331$ and $\sigma = 0.212$ is -2239.534744. The log likelihood value is much less negative than log likelihood value in part (b). This proves that the ML estimates for μ and σ are a better fit for the data given than the parameters in part (b).

The variance-covariance matrix, M_1, is reported below.

$$M_1 = \begin{bmatrix} 2.73 \times 10^{-4} & 1.55 \times 10^{-6} \\ 1.55 \times 10^{-6} & 1.11 \times 10^{-4} \end{bmatrix}$$

Figure 3: Normed Histogram of MACSS Students Annual Income and Log PDFs



Part (d).

According to the likelihood ratio test, there is no probability that the data in incomes.txt came from the distribution in part (b).

Part (e).

Using the estimated model from part (c), the probability of a MACSS graduate will earn more than $100,000 annually is 0.196 and the probability of a MACSS graduate earning less than $75,000 is 0.308.

Problem 2
Part (a).

$$sick_i = \beta_0 + \beta_1 age_i + \beta_2 children_i + \beta_3 temp\_winter_i + \varepsilon_i \tag{1}$$
$$where \quad \varepsilon_i \tilde{} N(0, \sigma^2)$$

To solve the parameters $(\beta_0, \beta_1, \beta_2, \beta_3, \sigma)$ of model (1), we leverage on the fact that the error term is normally distributed at $\mu = 0$ and solve the regression equation for ε_i

$$\varepsilon_i = sick_i - \beta_0 - \beta_1 age_i - \beta_2 children_i - \beta_3 temp\_winter_i \tag{2}$$

The estimates for the parameters of the model are as follows: β_0 = 0.25164638, β_1 = 0.01293335, β_2 = 0.40050205, β_3 = -0.00999167, σ = 0.00301768

The value of the log likelihood function is 876.865046288.

The variance-covariance matrix, M_2, is reported below.

$$M_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Part (b).

The probability that $\beta_0 = 1.0$, $\sigma^2 = 0.01$ and $\beta_1, \beta_2, \beta_3 = 0$ is 0.00. The likelihood that age, number of children, and average winter temperature have no effect on the number of sick days is 0. The conclusion is highly probable as people fall sick more easily when the winter is colder, and more children means more exposure to germs from people they interact with.