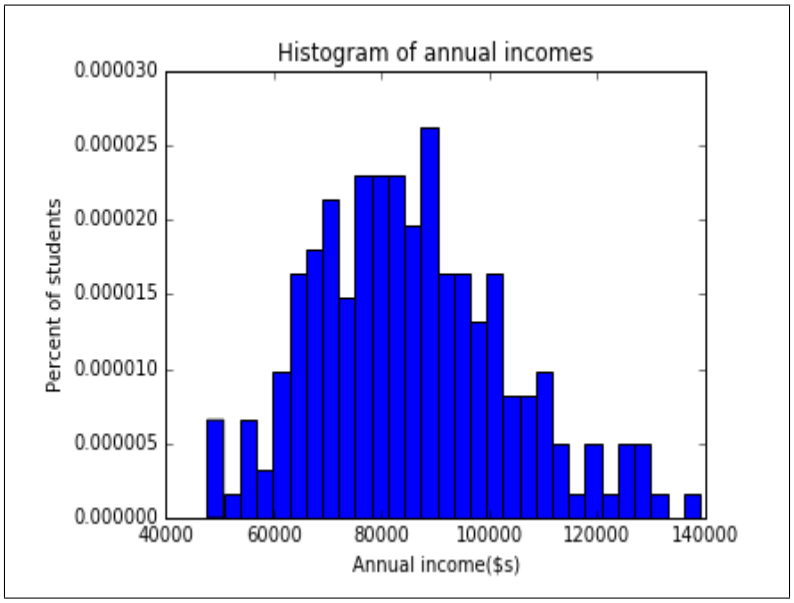
Problem Set #2

MACS 30100, Dr. Evans
Huanye Liu

Problem 1

Part (a). The histogram based on the given data is shown below.

Figure 1: Histogram of annual incomes



Part (b). The lognormal distribution pdf given the provided parameter is shown on the next page. The value of the log lilkelihood value for this parameterization of the distribution and given the data is -8298.63695601.

Part (c). The three overlaid plots as specified are shown on the next page. The ML estimates for μ and σ are 11.3314403742 0.211674571004 respectively, and the value of the log likelihood function is -2239.534744, and the variance-covariance matrix of MLE estimates is

$$\begin{bmatrix} 1.58680e-06 & 9.80377e-06 \\ 9.80377e-06 & 2.89832e-03 \end{bmatrix}$$

Part (d). After running the hypothesis testing we found the p-valune is practically zero, which means that it is not possible that the given data came from the distribution specified in part (b).

Part (e). The probability that I will earn more that $100,000 is 0.195618045405, and the probability that I will earn less than $70,000 is 0.307939532566.

Problem 2

Part (a). The MLE value of standard deviation is 0.0164687775806, the MLE values of $\beta 0$, $\beta 1$, $\beta 2$ and $\beta 3$ are 0.253676099571, 0.0129218826337, 0.400370277848 and -0.0100195586474 respectively. The log likehood function value using the maximum

Figure 2: Lognormal distribution pdf

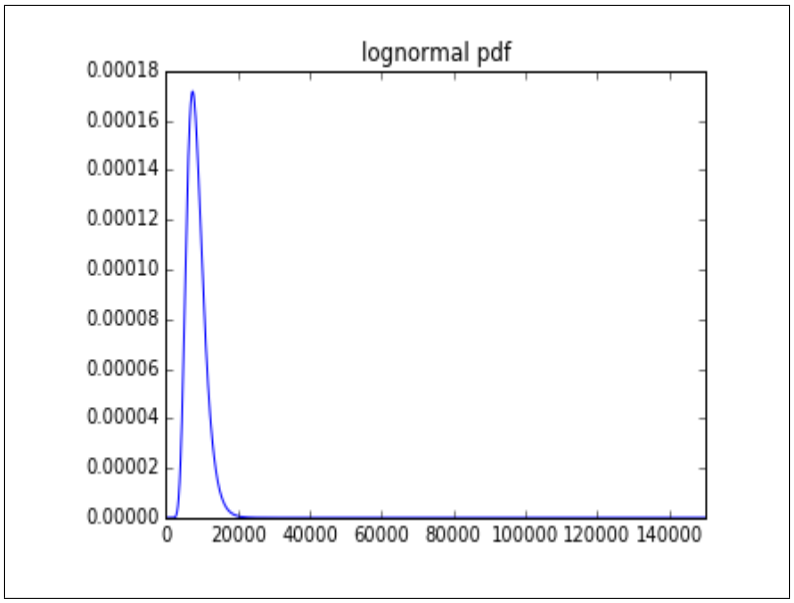
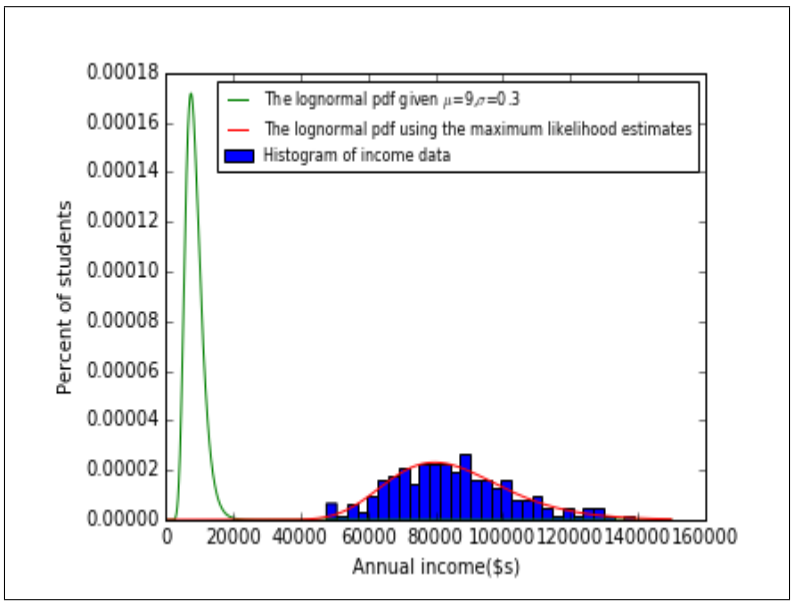


Figure 3: Three overlaid plots for comparison



likelihood estimates is 634.0363498426419, and the variance-covariance matrix of MLE estimates is

$$\begin{bmatrix} 1.0 & 0 & 0 & 0 & 0 \\ 0 & 1.0 & 0 & 0 & 0 \\ 0 & 0 & 1.0 & 0 & 0 \\ 0 & 0 & 0 & 1.0 & 0 \\ 0 & 0 & 0 & 0 & 1.0 \end{bmatrix}$$

Part (b). After running the hypothesis testing we found the p-valune is again practically zero, which means it is not possible that the given data came from the linear regression model given the parameters specified in part (b).