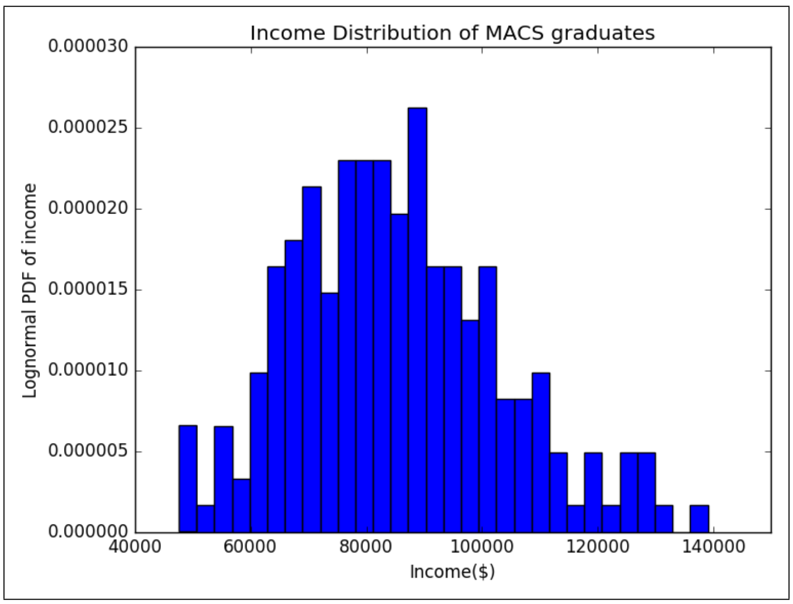
Problem Set #3

MACS 30100, Dr. Evans
Yinxian Zhang

Problem 1 Lognormal Distribution of Some Income Data
Part (a).
The histogram of percentages of the annual income of MACS graduates is plotted in Figure 1.

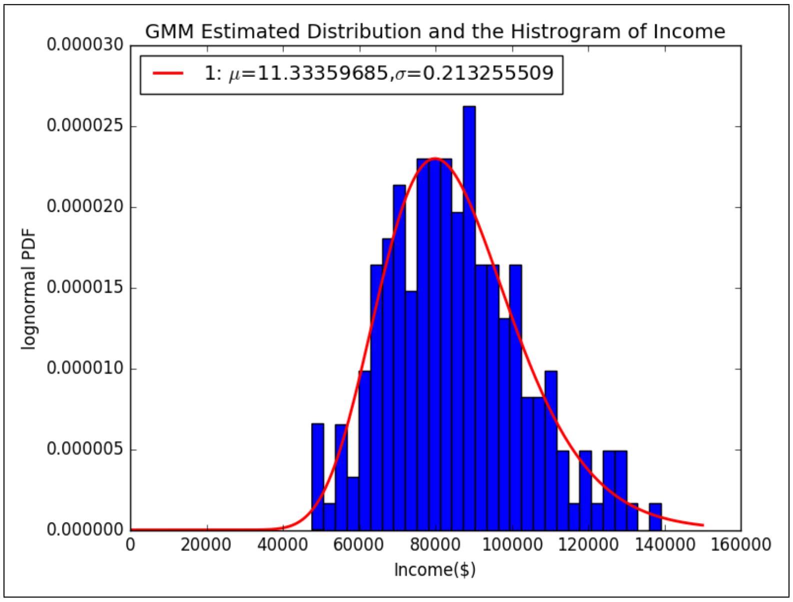
Figure 1:



Part (b).
Using the average income and standard deviation of income as two moments and estimating the parameters by GMM methods with an identity matrix $\hat{W}$, we can get a GMM criterion function value of 6.757763e-13, and the estimated parameters are μ = 11.333596851 and σ = 0.2132555.

While the two data moments are average income \$85276.82361 and standard deviation \$17992.54213, the two model moments are \$85276.87995 and \$17992.55093, respectively. The GMM estimated lognormal PDF is plotted against the histogram in Figure 2.

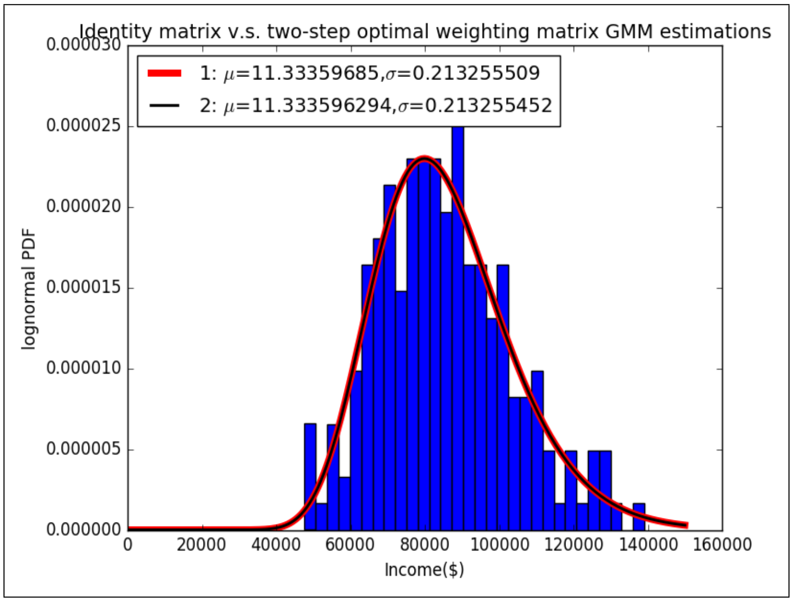
Figure 2:



Part (c).
Performing the two-step GMM estimator by using the estimates from part (b)and replacing the indentity martix with a optimal weighting matrix $\hat{W}_2 step$, we can get a GMM criterion function value of 0.00855143, and the estimated parameters are $\mu =$ 11.333596294 and $\sigma = 0.21325545$.

While the two data moments are average income \$85276.82361 and standard deviation \$17992.54213, the two model moments are \$85276.8334804 and \$17992.53915 respectively. The two-step GMM estimated lognormal PDF is plotted against the PDF from part (b) and the histogram from part (a) in Figure 3.

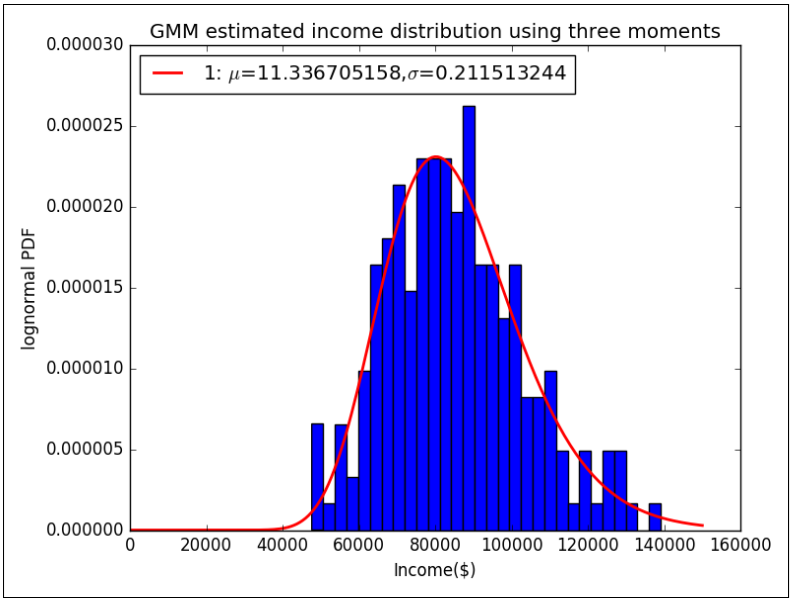
Figure 3:



Part (d).
Using percent of individuals who earn less than \$75,000, between \$75,000 and \$100,000, and more than \$100,000 as three moments, and using the identity matrix as the weighting matrix, we can a GMM criterion function value of 2.79639e-07, and the estimated parameters are $\mu = 11.3367052$ and $\sigma = 0.2115132$.

While the three data moments are 0.3, 0.5 and 0.2, the three model moments are 0.30000077, 0.50000215 and 0.20000034. The GMM estimated lognormal PDF is plotted in Figure 4.

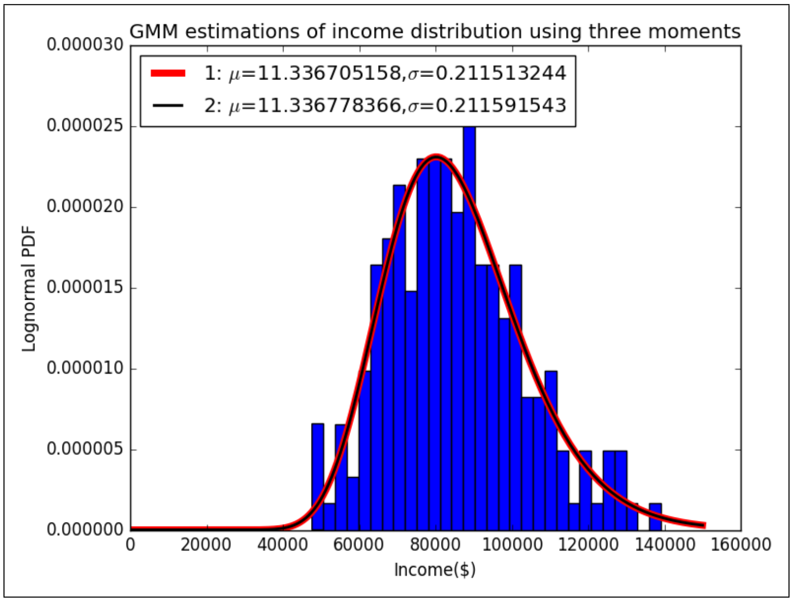
Figure 4:



Part (e).
Performing the two-step GMM estimator by using the estimates from part (d)and replacing the indentity martix with a optimal weighting matrix $\hat{W}_2 step$, we can get a GMM criterion function value of 1.67243008678e-09, and the estimated parameters are $\mu = 11.3367783658$ and $\sigma = 0.211591543$.

While the three data moments are 0.3, 0.5 and 0.2, the three model moments are 0.29995193, 0.49987588 and 0.20017546. The two-step GMM estimated lognormal PDF is plotted against the PDF from part (d) and the histogram from part (a) in Figure 5.

Figure 5:



Part (f).
Eyeballing Figure 2, 3, 4 and 5, we found that the model fit based on estimations in parts (b), (c), (d), and (e) are almost indistinguishable. The estimated distributions look very similar in four figures. However, examining the GMM criterion function values of the four models, we found that the model in part (b) somehow has the smallest value of 6.757763e-13, indicating that its weighted sum of squared errors is the smallest, whereas the two-step estimator in part (c) gives us the largest value(and thus the worst estimation). We therefore conclude that using the simple Indentity matrix and the two data moments of the mean and the standard deviation of income somehow fits the data best.

Problem 2 Linear regression and GMM
Part (a).
Treating each of the 200 values of the variable *sick* as 200 data moments and estimating the model parameters with the identity matrix, we can get a GMM criterion function value of 0.0018212898, and the estimated parameters are:
β_0 = 0.25164
β_1 = 0.01293
β_2 = 0.4005
β_3 = -0.00999