Problem Set #4

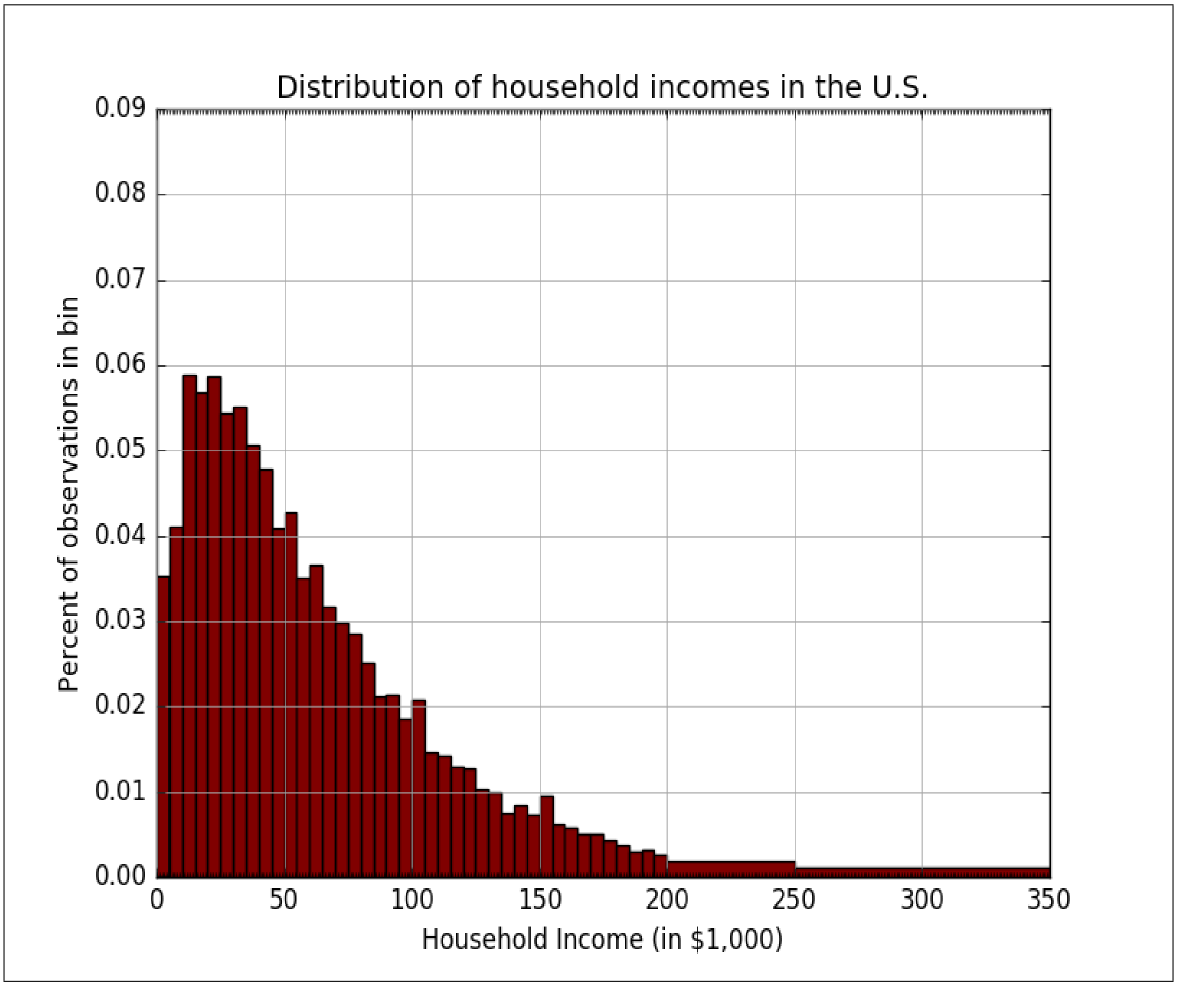
MACS 40200, Dr. Evans

Dongping Zhang

Problem 1: **Matching the U.S. income distribution by GMM**

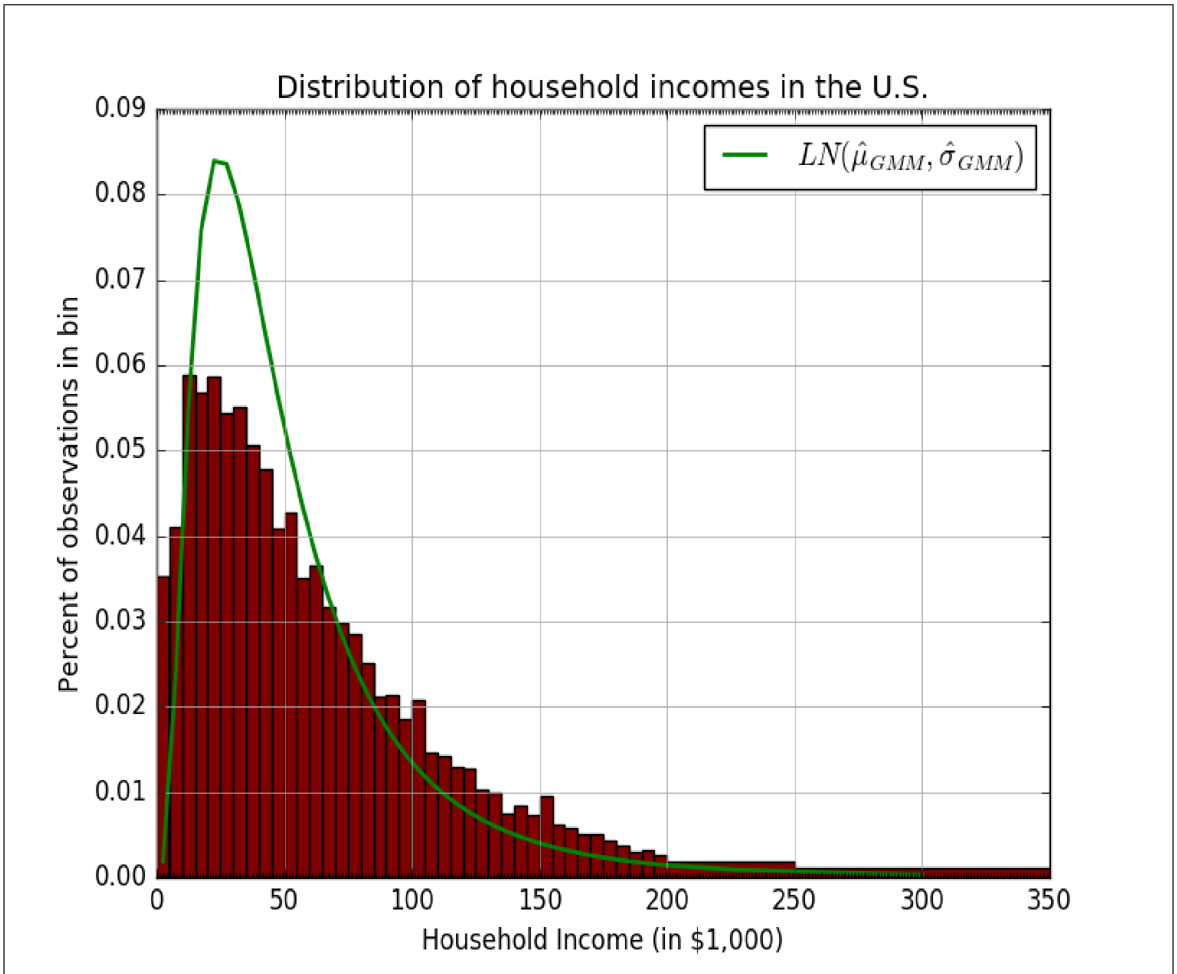
Part (a). Plot a histogram of percentages using *usincmoms.txt* data.

Figure 1: U.S. household incomes distribution



Part (b). Using GMM, fit a lognormal LN $(x; \mu, \sigma)$ distribution.

Figure 2: U.S. household incomes distribution against the estimated lognormal

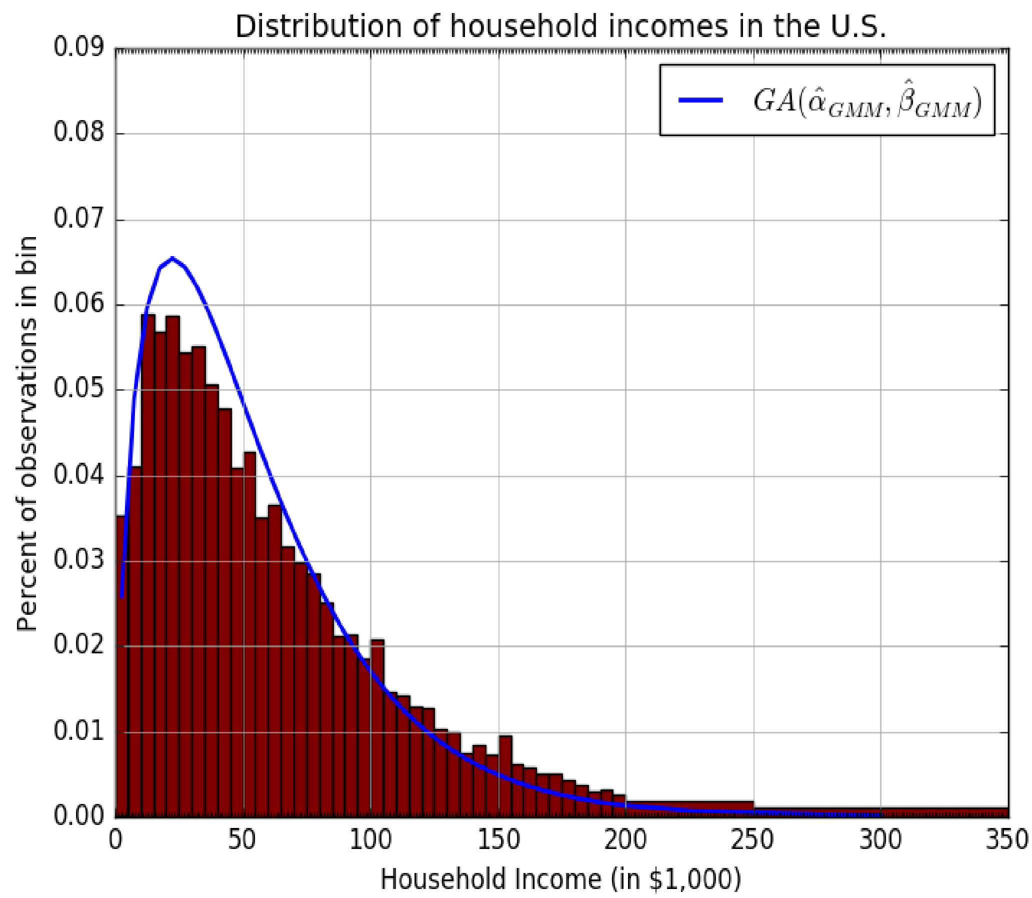


The estimated parameters of the lognormal distribution by GMM are:

- $\hat{\mu}_{2M}^{GMM}$ = **41.95197663**
- $\hat{\sigma}_{2M}^{GMM}$ = **0.73392374**
- Value of GMM criterion function at using GMM parameters obtained is: **0.180702396891**

Part (c). Using GMM, fit a Gamma GA (x; α, β) distribution.

Figure 3: U.S. household incomes distribution against the estimated Gamma

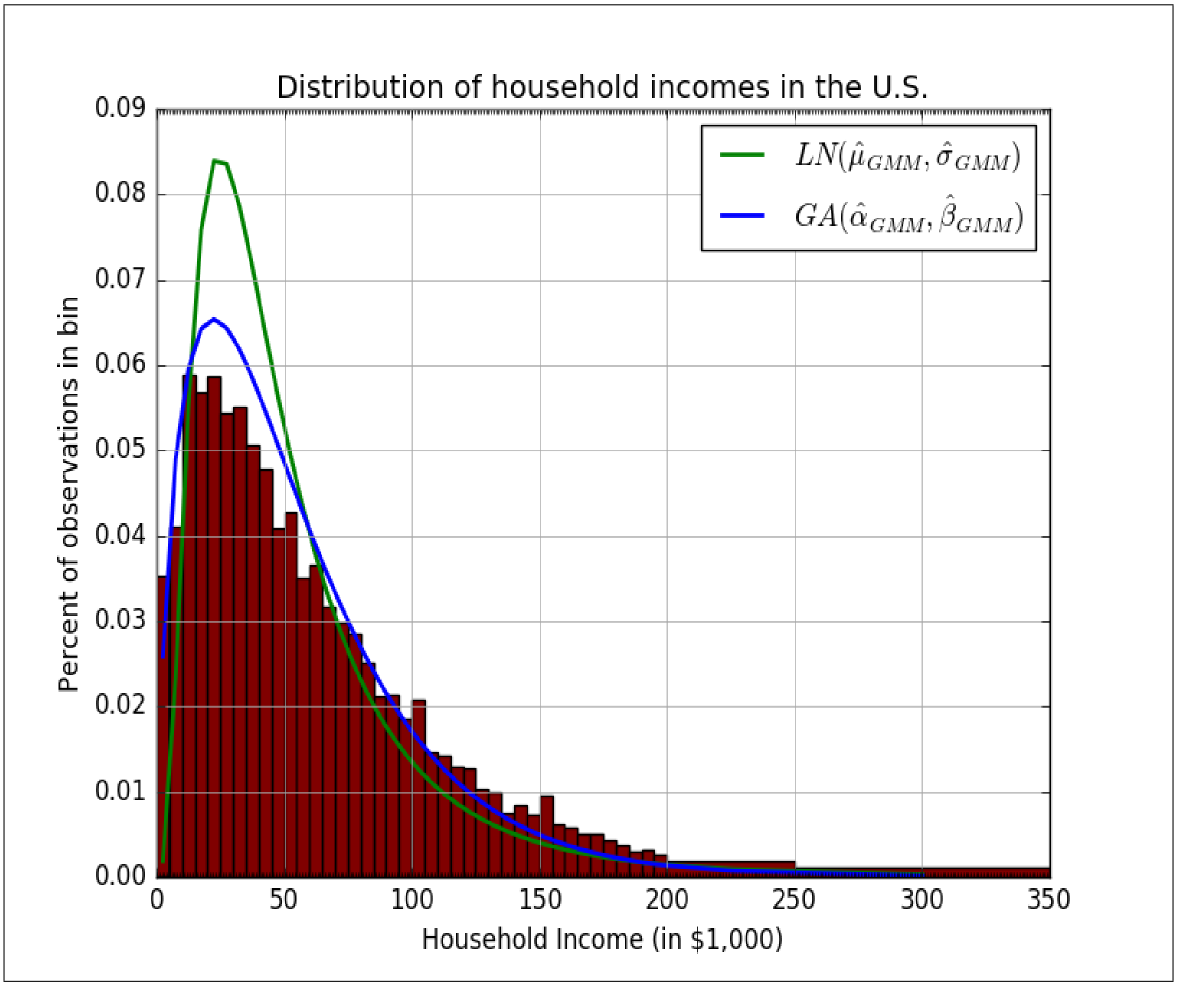


The estimated parameters of the lognormal distribution by GMM are:

- $\hat{\alpha}^{GMM}$ = **1.66738828**
- $\hat{\beta}^{GMM}$ = **33.12260221**
- Value of GMM criterion function at using GMM parameters obtained is: **0.0455036142137**

Part (d). Plot the histogram from part (a) overlayed with the line represent- ing the implied histogram from your estimated lognormal (LN) distribution from part (b) and the line representing the implied histogram from your estimated gamma (GA) distribution from part (c). What is the most pre- cise way to tell which distribution fits the data the best? Which estimated distribution, LN or GA, fits the data best?

Figure 4: GMM using 3 percentile moments

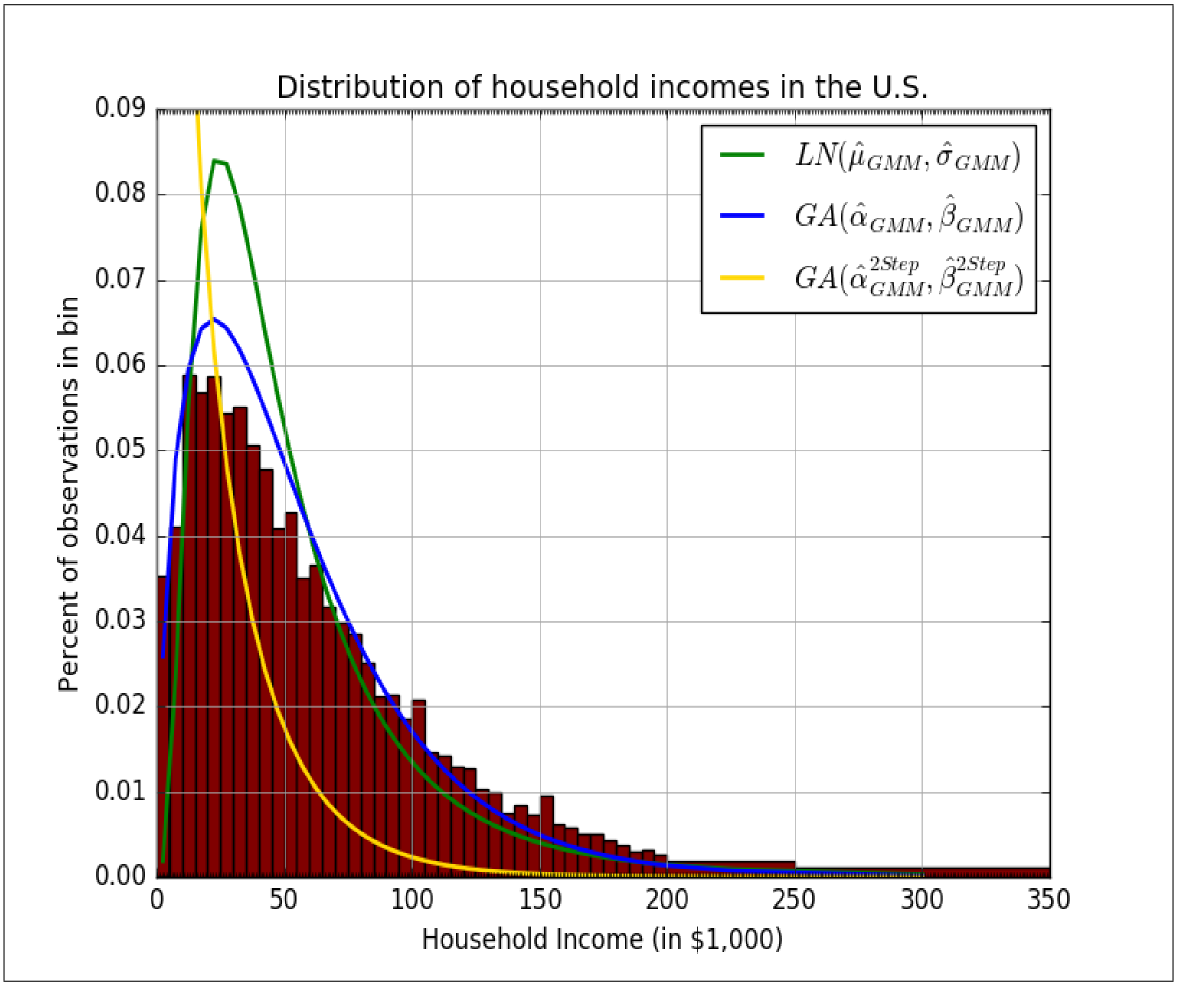


The most precise way to tell which distribution fits the data the best is by looking at the number of moments versus the number of estimated parameters. However, in the cases of lognormal and Gamma estimations, we both use 42 moments to estimate 2 parameters, thus, we could only count on the values obtained from the criterion function. We would prefer to have the estimation with smaller values returned by the criterion function using GMM parameters.

In conclusion, because the value returned by Gamma GMM criterion function is smaller than that of the lognormal, we could claim that Gamma distribution fits the data the better, and by observing the Figure 4, we could confirm that Gamma fits the data better.

Part (e). Repeat your estimation of the GA distribution from part (c), but use the two-step estimator for the optimal weighting matrix.

Figure 5: GMM_{2Step} using 3 percentile moments



The estimated parameters of the lognormal distribution by GMM are:

- $\hat{\alpha}^{GMM}_{2S}$ = **0.63947259**
- $\hat{\beta}^{GMM}_{2S}$ = **28.27613904**
- Value of GMM criterion function at using GMM parameters obtained is: **3.626061901297946** $\times$ **10^{-15}**

My estimated parameters for the Gamma distribution using 2-step GMM method seems to vary greatly from part (c). Although the value returned by GMM criterion function using 2-step GMM estimators and optimal variance-covariance matrix is $3.626061901297946 \times 10^{-15}$, which is significantly smaller, the goodness of fit by visualization does not seem to support this small value. In conclusion, I would still prefer the estimation from part (c).