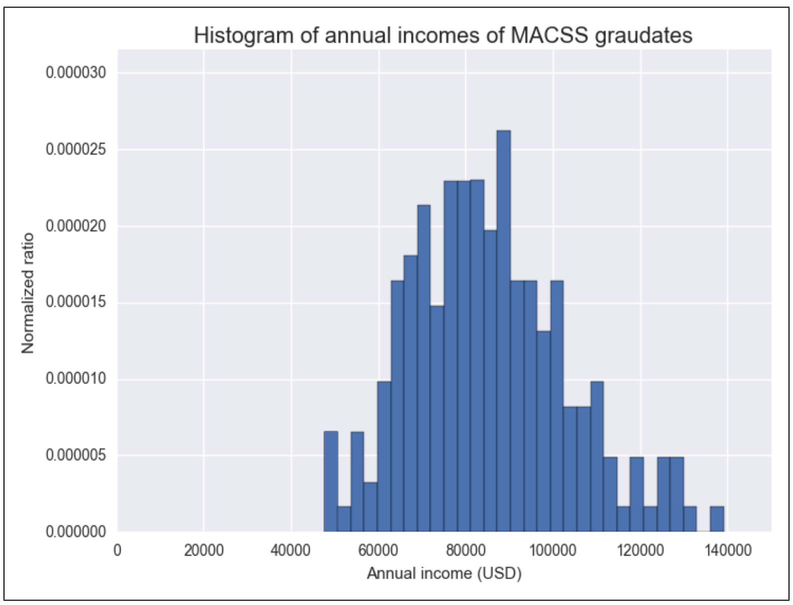
Problem Set #3

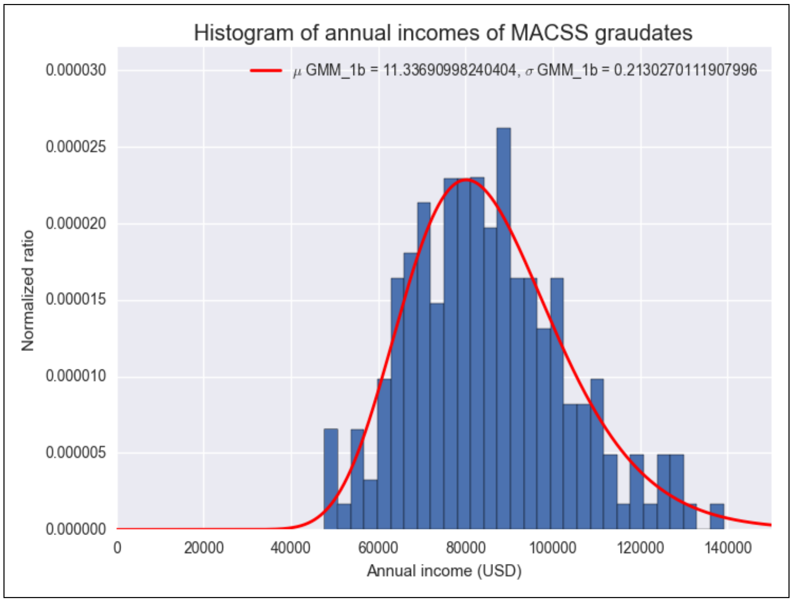
MACS 30100, Dr. Evans
Bobae Kang

1. **Some income data, lognormal distribution, and GMM.**

   (a) Plot a histogram of percentages of the ‘income.txt‘ data with 30 bins. Make sure that the bins are weighted using the ‘normed=True‘ option. Make sure your plot has correct x-axis and y-axis labels as well as a plot title.

   (b) Estimate the parameters of the lognormal distribution by generalized method of moments. Plot your estimated lognormal PDF against the histogram from part (a). Report the value of your GMM criterion function at the estimated parameter values. Report and compare your two data moments against your two model moments at the estimated parameter values.



GMM crietron function value = 3.93982160e-13

GMM estimate for μ = 11.3369099824

GMM estimate for σ = 0.213027011191

Data moments:

Mean = 85276.8236063

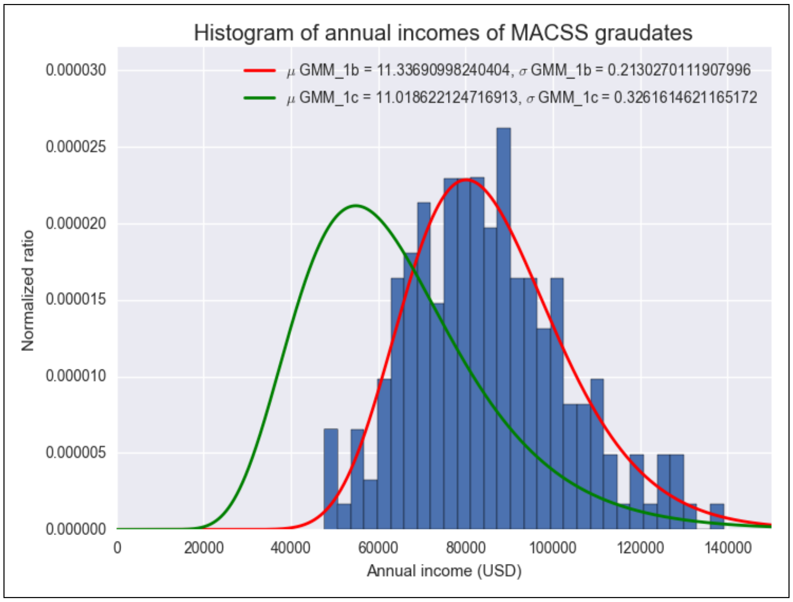
Standard Deviation = 17992.542128

Model moments:

Mean = 85276.79520510265

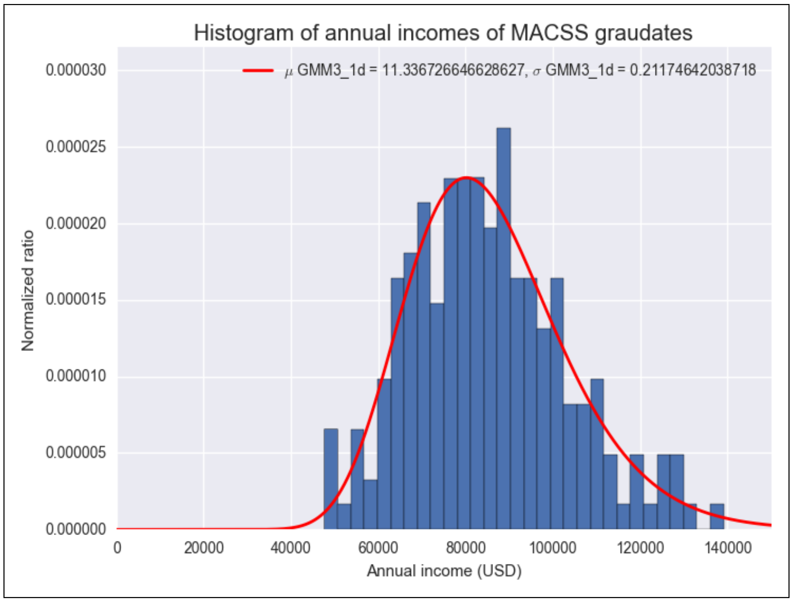
Standard Deviation = 17992.5325554

(c) Perform the two-step GMM estimator by using your estimates from part (b) with two moments to generate an estimator for the variance covariance matrix. Report your estimates as well as the criterion function value at these estimates. Plot your estimated lognormal PDF against the histogram from part (a) and the estimated PDF from part (b). Report and compare your two data moments against your two model moments at the estimated parameter values.



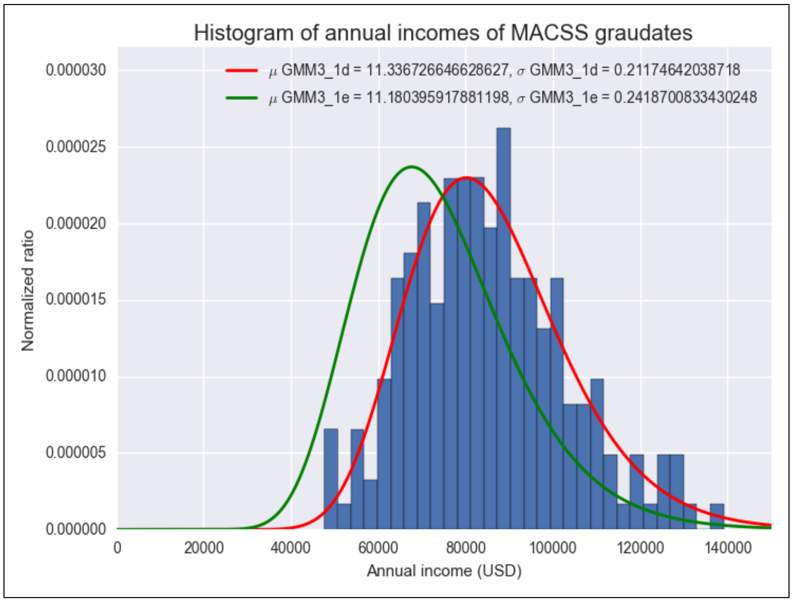
GMM crietron function value = 0.08985203
GMM estimate for μ = 11.0186221247
GMM estimate for σ = 0.326161462117
Data moments:
  Mean = 85276.8236063
  Standard Deviation = 17992.542128
Model moments:
  Mean = 63849.689103934565
  Standard Deviation = 20822.5632426

(d) Estimate the lognormal PDF to
t the data by GMM using different moments. Plot your estimated lognormal PDF against the histogram from part (a). Report the value of your GMM criterion function at the estimated parameter values. Report and compare your three data moments against your three model moments at the estimated parameter values.



GMM crietron function value = 0.23818173
GMM estimate for μ = 11.3367266466
GMM estimate for σ = 0.211746420387
Data moments:
% of inidivudals who earn less than \$75,000 = 0.3
% of inidivudals who earn between \$75,000 and \$100,000 = 0.5
% of inidivudals who earn more than \$100,000 = 0.2
Model moments:
% of inidivudals who earn less than \$75,000 = 0.299272453678561
% of inidivudals who earn between \$75,000 and \$100,000 = 0.4980574551520791
% of inidivudals who earn more than \$100,000 = 0.19966278330544504

(e) Perform the two-step GMM estimator by using your estimates from part (d) with three moments to generate an estimator for the variance covariance matrix. Report your estimates as well as the criterion function value at these estimates. Plot your estimated lognormal PDF against the histogram from part (a) and the estimated PDF from part (d). Report and compare your three data moments against your three model moments at the estimated parameter values.



GMM crietron function value = 2.89594050e-09
GMM estimate for μ = 11.1803959179
GMM estimate for σ = 0.241870083343
Data moments:
% of inidivudals who earn less than \$75,000 = 0.3
% of inidivudals who earn between \$75,000 and \$100,000 = 0.5
% of inidivudals who earn more than \$100,000 = 0.2
Model moments:
% of inidivudals who earn less than \$75,000 = 0.5735500650944066
% of inidivudals who earn between \$75,000 and \$100,000 = 0.34185740892053457
% of inidivudals who earn more than \$100,000 = 0.08345289282059601

(f) Which of the four estimations from parts (b), (c), (d), and (e) ts the data best? Justify your answer.

The graphs above suggest that, for the given choice of moments and initial parameter values, the GMM estimate using the identity matrix as the weight matrix (i.e., 1.(b) and 1.(d)) fits the data better than the other using the two-step weight matrix (i.e., 1.(c) and 1.(e)). It is more difficult to compare across estimates using different moments, e.g. between 1.(b) and 1.(d). Between 1.(b) and 1.(d), the former has a slightly thinker tail than the latter, captureing more of the middle part of the histogram. It is my view that choosing the proportions of individuals for different income groups may be suprerior to choosing mean and standard deviation as moments. This is because I expect the income distribution to be generally more fat-tailed than not.

2. **Linear regression and GMM**

(a) Estimate the parameters of the model $(\beta_0, \beta_1, \beta_2, \beta_3)$ by GMM by solving by solving the minimization problem of the GMM criterion function. Report your estimates and report the value of your GMM criterion function.

GMM estimate for β_0 = 0.251644953001
GMM estimate for β_1 = 0.0129334651063
GMM estimate for β_2 = 0.40050102133
GMM estimate for β_3= -0.00999170755539
GMM crietron function value = 0.00182128981749