Problem Set #2

MACS 40200, Dr. Evans
Bobae Kang

1. Health claim amounts and the GB family of distributions

(a) Calculate and report the mean, median, maximum, minimum, and standard deviation of monthly health expenditures for these data. Plot two histograms of the data in which the y-axis gives the percent of observations in the particular bin of health expenditures and the x-axis gives the value of monthly health expenditures.

The descriptive statistics of the data are as follows:

Table 1: The descriptive statistics of the data

Mean	720.277975327
Median	172.21
Maximum	227967.25
Minimum	0.01
Standard Deviation	3972.66375639

The following two graphs are the hisograms of the data

Figure 1: A histogram of the health claims from a fictitious sample households

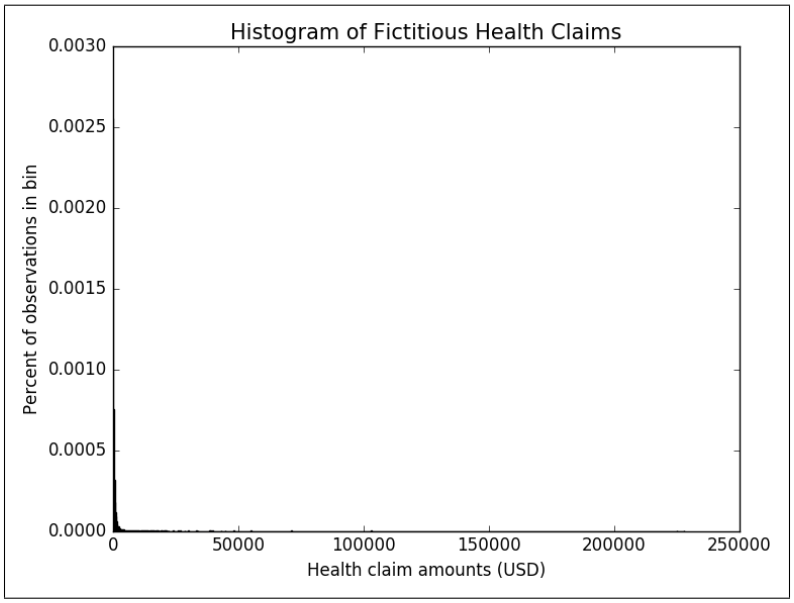
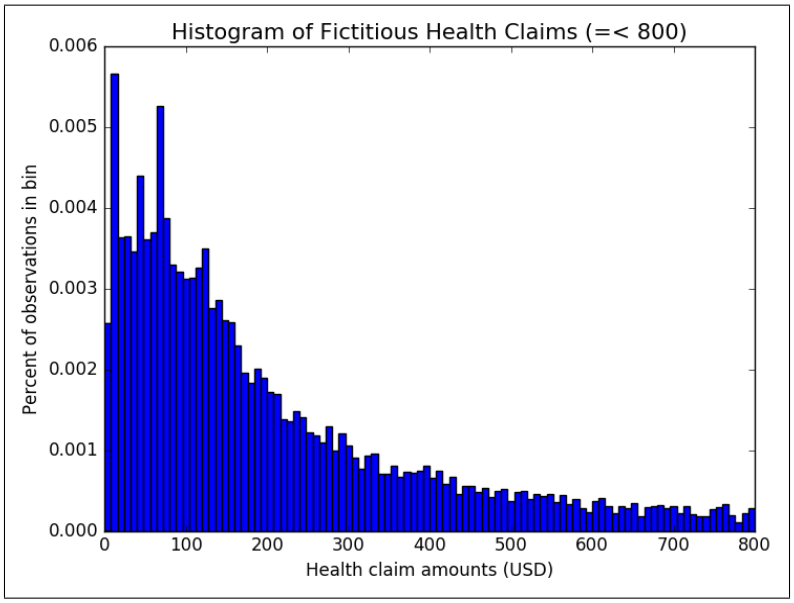


Figure 2: A histogram of the health claims from a fictitious sample households, $\leq$ \$800



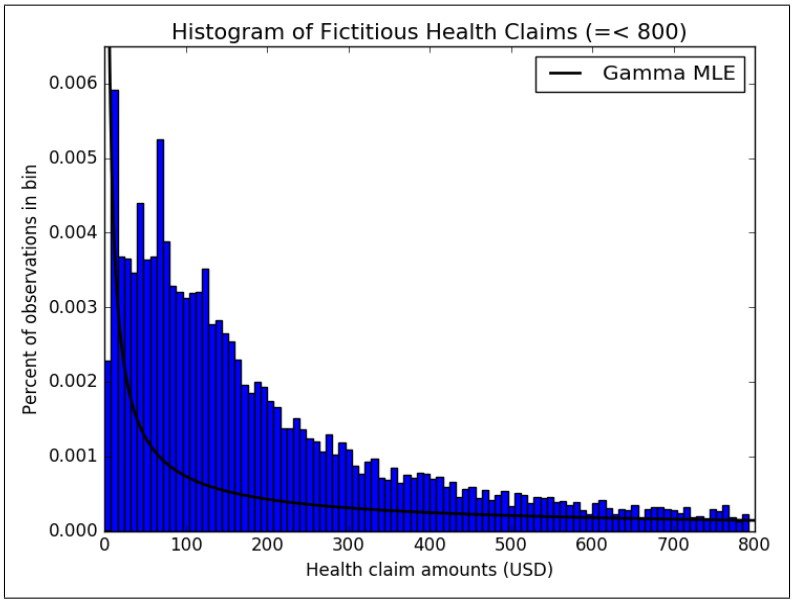
(b) Using MLE, fit the gamma distribution to the individual observation data. Report your estimated values for parameters, as well as the value of the maximized log likelihood function. Plot the second histogram from part (a) overlayed with a line representing the implied histogram from your estimated gamma distribution.

Table 2: MLE results using the gamma distribution

$\hat{\alpha}_{MLE}$	0.221755308612
$\hat{\beta}_{MLE}$	21911.0646993
Log-likelihood	-82076.4516057

VCV = [[4.37717820e-06 9.32885080e-05]
[9.32885080e-05 1.00000060e+00]]

Figure 3: Fitting the gamma distribution using MLE, $\leq$ \$800



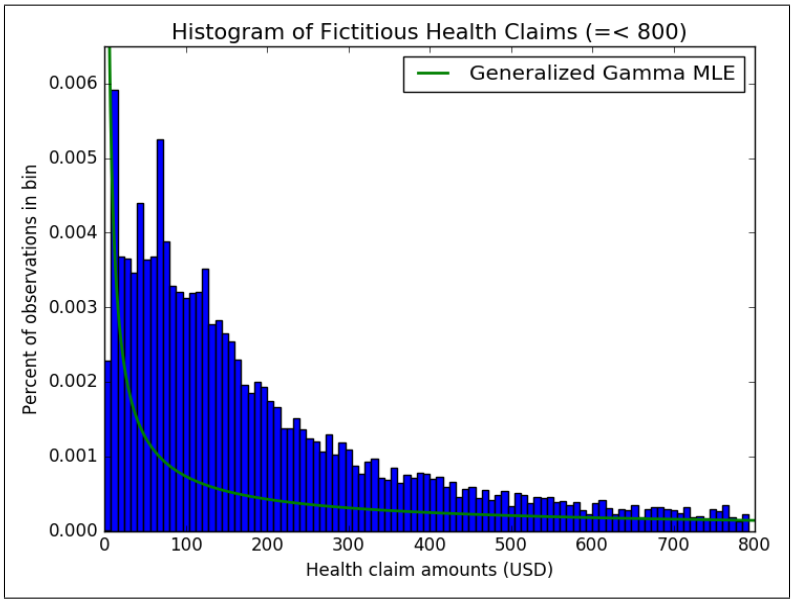
(c) Using MLE, fit the generalized gamma distribution to the individual observation data. Use your estimates for α and β from part (b), as well as m = 1, as your initial guess. Report your estimated values for parameters, as well as the value of the maximized log likelihood function. Plot the second histogram from part (a) overlayed with a line representing the implied histogram from your estimated generalized gamma distribution.

Table 3: MLE results using the generalized gamma distribution

$\hat{\alpha}_{MLE}$	0.222276613384
$\hat{\beta}_{MLE}$	21911.0644724
$\hat{m}_{MLE}$	0.997648454581
Log-likelihood	-82076.446472

VCV = [[3.08655089e-05 -1.52641448e-05 -1.51784572e-04]
[-1.52641448e-05 1.00903661e+00 5.29737394e-03]
[-1.51784572e-04 5.29737394e-03 4.58901429e-02]]

Figure 4: Fitting the generalized gamma distribution using MLE, $\leq$ \$800



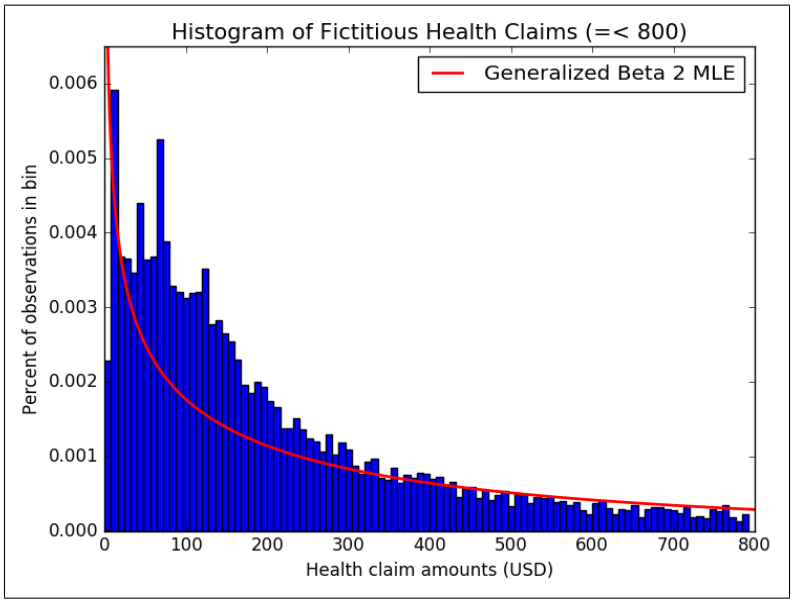
(d) Using MLE, fit the generalized beta 2 distribution to the individual observation data. Use your estimates for α, β and m from part (c), as well as $q = 10,000$, as your initial guess. Report your estimated values for parameters, as well as the value of the maximized log likelihood function. Plot the second histogram from part (a) overlayed with a line representing the implied histogram from your estimated generalized beta 2 distribution.

Table 4: MLE results using the generalized gamma distribution

$\hat{a}_{MLE}$	0.699528434728
$\hat{b}_{MLE}$	223919455.619
$\hat{p}_{MLE}$	0.995695905023
$\hat{q}_{MLE}$	10001.8955911
Log-likelihood	-76461.5898932

VCV = [[8.47578124e-03 0.00000000e+00 -1.03428896e-01 -1.15211264e+00]
[0.00000000e+00 1.00000000e+00 0.00000000e+00 0.00000000e+00]
[-1.03428896e-01 0.00000000e+00 1.26223521e+00 1.40607231e+01]
[-1.15211264e+00 0.00000000e+00 1.40607231e+01 1.56664117e+02]]

Figure 5: Fitting the generalized beta 2 distribution using MLE, $\leq$ \$800



(e) Perform a likelihood ratio test for each of the estimated in parts (b) and (c), respectively, against the GB2 speci

cation in part (d). Report the chi-square(4) values from the likelihood ratio test for the estimated GA and the estimate GG.

χ^2 of $H_0 : ga = gb2$ with 4 degrees of freedom p-value = 1.0
χ^2 of $H_0 : gg = gb2$ with 4 degrees of freedom p-value = 1.0

(f) Using the estimated GB2 distribution from part (d), how likely am I to have a monthly health care claim of more than $\$1,000$? How does this amount change if I use the estimated GA distribution from part (b)?

Likelihood of having a monthly healthcare claim of more than \$1000 according to the Generalized Beta 2 Distribution: 0.16267648717870076

Likelihood of having a monthly healthcare claim of more than \$1000 according to the Gamma Distribution: 0.45195972238946325