Eric Fischer
UID: 303 759 361
emfischer712@ucla.edu
Stat 202C - Project 2

Metropolis-Hastings: Versus Accept-Reject and Utilized in Logistic Regression

Comparison of Metropolis-Hastings with Accept-Reject

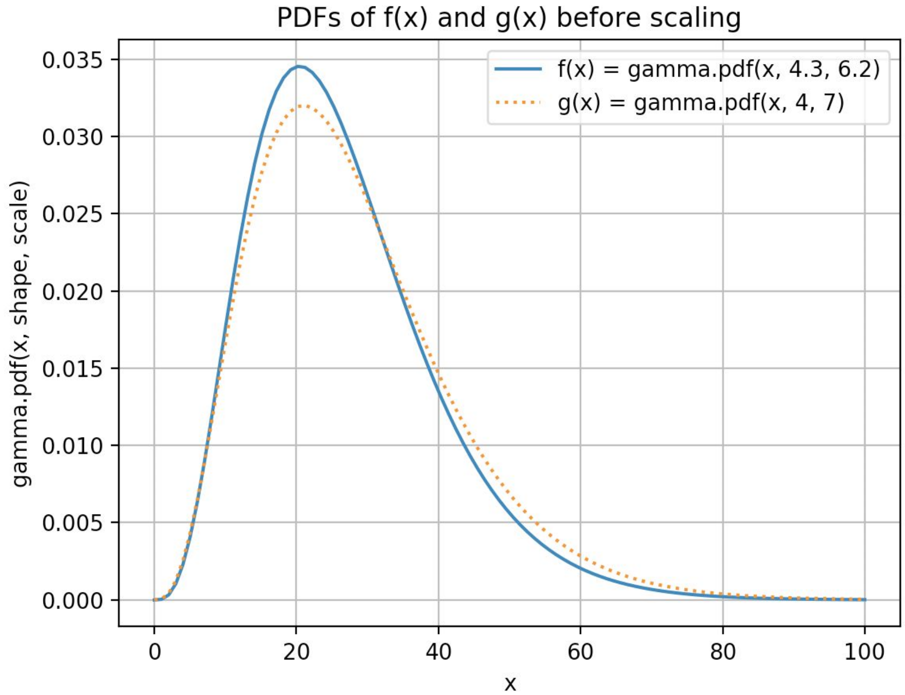
Problem 7.1: Here we will estimate the mean of a Gamma(4.3, 6.2) random variable using the following algorithms and gamma distribution parameters: **a)** Accept-Reject with a Gamma(4, 7) candidate, **b)** Metropolis-Hastings with a Gamma(4, 7) candidate, and **c)** Metropolis-Hastings with a Gamma(5, 6) candidate. For each algorithm and given Gamma parameters, we will monitor the convergence to the mean of the Gamma(4.3, 6.2) random variable.

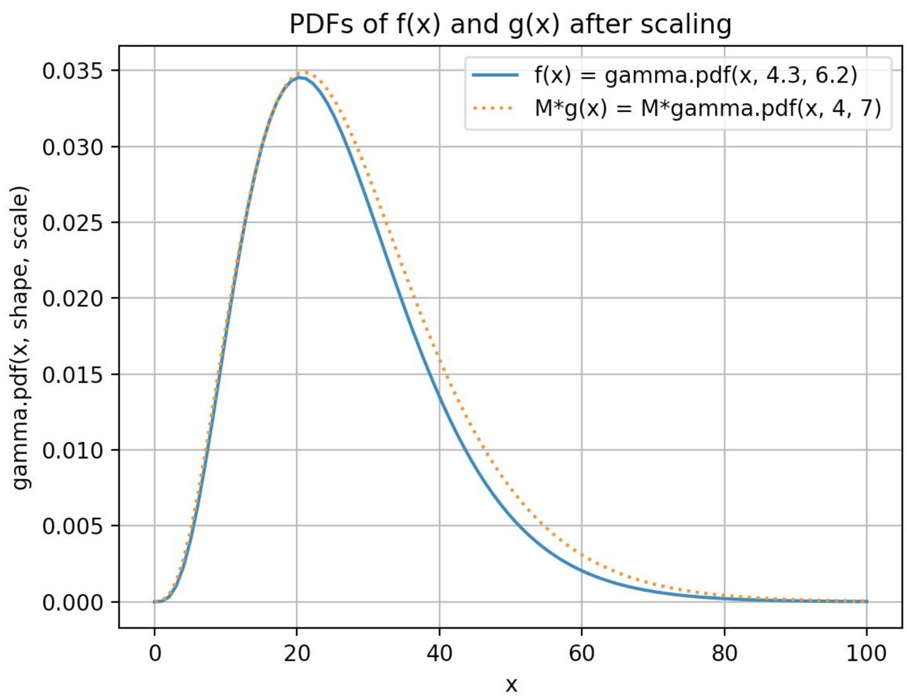
Problem 7.1a: In this subproblem we evaluate the Accept-Reject algorithm with a Gamma(4, 7) candidate. The Accept-Reject algorithm is shown below, taken from *Monte Carlo Statistical Methods* by Robert and Casella.

Algorithm A.4 –Accept–Reject Method–

1. Generate $X \sim g$, $U \sim \mathcal{U}_{[0,1]}$;
2. Accept $Y = X$ if $U \leq f(X)/Mg(X)$; [A.4]
3. Return to 1. otherwise.

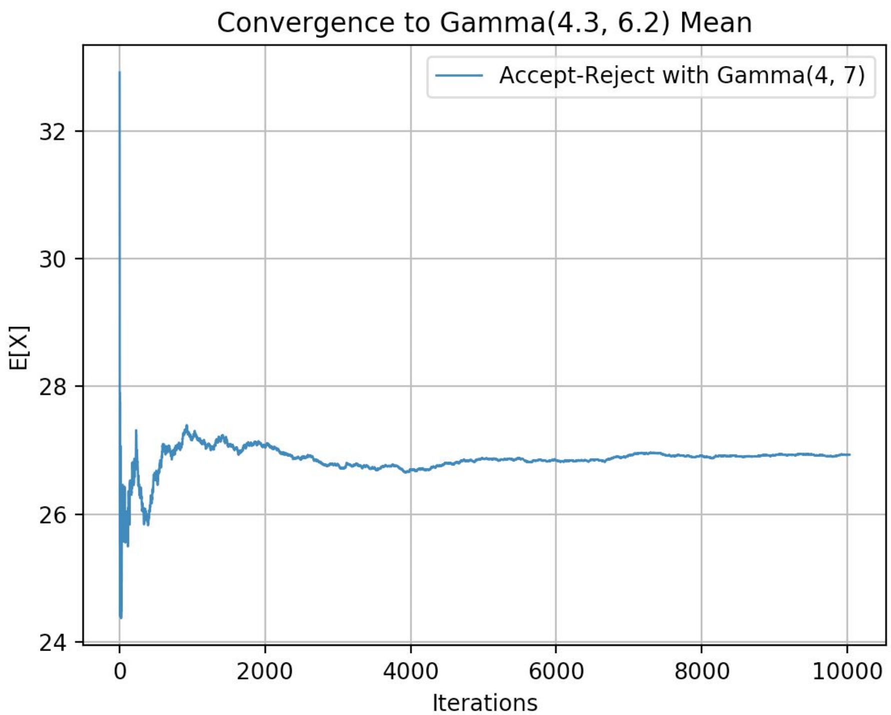
The two plots below show the target distribution *f(x)* and trial distribution *g(x)* before and after scaling *g(x)* by the value *M*. *M* satisfies the inequality Gamma(x; 4.3, 6.2) ≤ *M* · Gamma(x; 4, 7) and is the smallest value such that *f(x)*/*g(x)* < *M*. In this case, as we can evaluate a parameterized Gamma(*a*, *b*) distribution using the python scipy.stats method gamma.pdf(*x*, *a*, *b*), the maximized value *M* = **1.0895** was found using the scipy.optimize function minimize.





We can see the trial distribution *g(x)* successfully encompasses or encloses the target distribution *f(x)* after *g(x)* is scaled by the optimized value *M*.

Below we plot the convergence of the Accept-Reject algorithm to the true estimated mean value of a Gamma(4.3, 6.2) random variable over 10,000 iterations. The mean of a Gamma distribution is **E**[X] = kθ, where k and θ are the shape and scale parameters. Hence the true estimated mean value of a Gamma(4.3, 6.2) random variable is **26.66**. The final value for the Accept-Reject estimator with a Gamma(4, 7) candidate is **26.9331**.



Problem 7.1b, 7.1c: Now we evaluate the Metropolis-Hastings algorithm with two sets of Gamma parameters: the same Gamma(4, 7) candidate given to the Accept-Reject algorithm and a Gamma(5, 6) candidate. The Metropolis-Hastings algorithm is shown below, taken again from *Monte Carlo Statistical Methods* by Robert and Casella.

Algorithm A.24 –Metropolis–Hastings–

Given $x^{(t)}$,

1. Generate $Y_t \sim q(y|x^{(t)})$.
2. Take

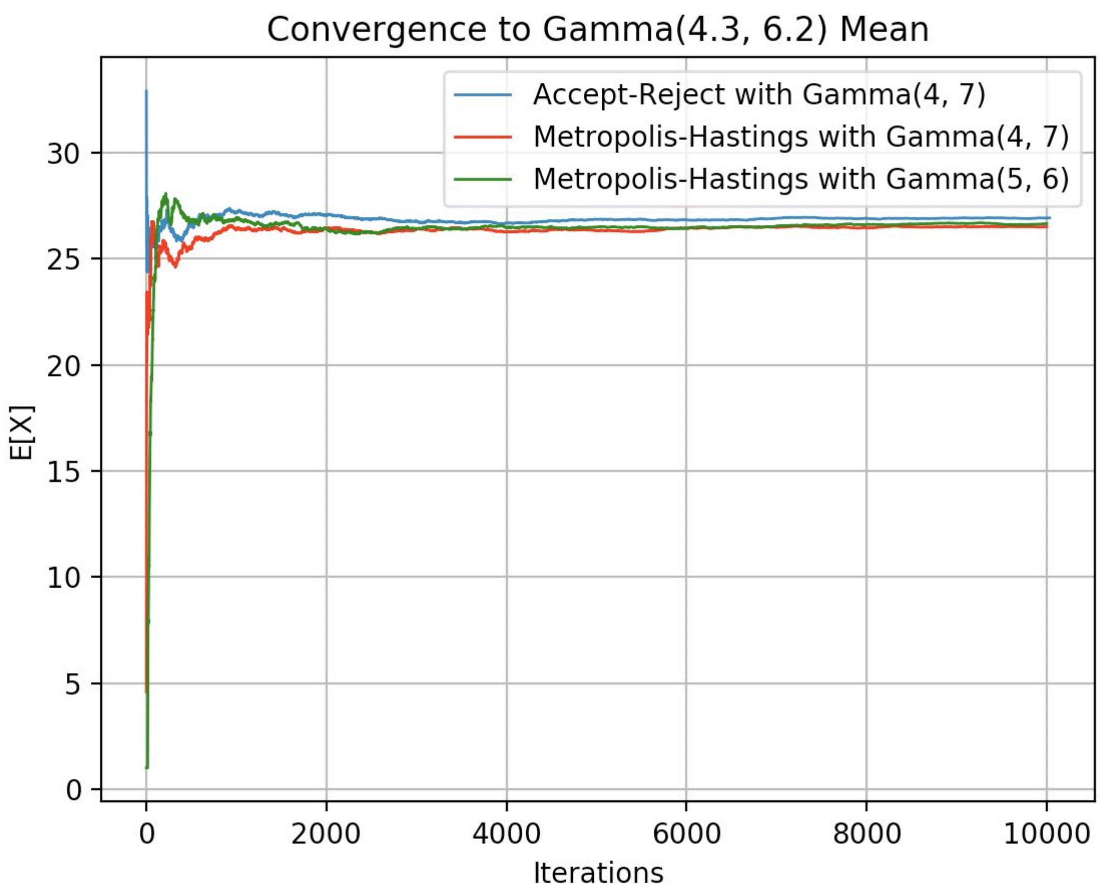
$$X^{(t+1)} = \begin{cases} Y_t & \texttt{with probability} \quad \rho(x^{(t)}, Y_t), \\ x^{(t)} & \texttt{with probability} \quad 1 - \rho(x^{(t)}, Y_t), \end{cases}$$

where $\qquad [A.24]$

$$\rho(x, y) = \min\left\{ \frac{f(y)}{f(x)} \frac{q(x|y)}{q(y|x)}, 1 \right\}.$$

Recall the final value for the Accept-Reject estimator with a Gamma(4, 7) candidate is **26.9331**. The final value for the Metropolis-Hastings estimator with a Gamma(4, 7) candidate is **26.5114**. The final value for the Metropolis-Hastings estimator with a Gamma(5, 6) candidate is **26.6621**. Note that these final estimated values are after 10,000 iterations. With more iterations, they continually better approximate the true value **26.66**, with the Metropolis-Hastings estimator with a Gamma(5, 6) candidate generally producing the closest approximation.

Below we plot the convergence for each of the three algorithms and Gamma distributions: **a)** Accept-Reject with a Gamma(4, 7) candidate, **b)** Metropolis-Hastings with a Gamma(4, 7) candidate, and **c)** Metropolis-Hastings with a Gamma(5, 6) candidate.



As we can see, the Metropolis-Hastings algorithm with a Gamma(5, 6) candidate, resulting in a final estimate of **26.9331**, best approximated the true estimated mean value **26.66** of a Gamma(4.3, 6.2) random variable over the first 10,000 iterations.

Metropolis-Hastings for Logistic Regression

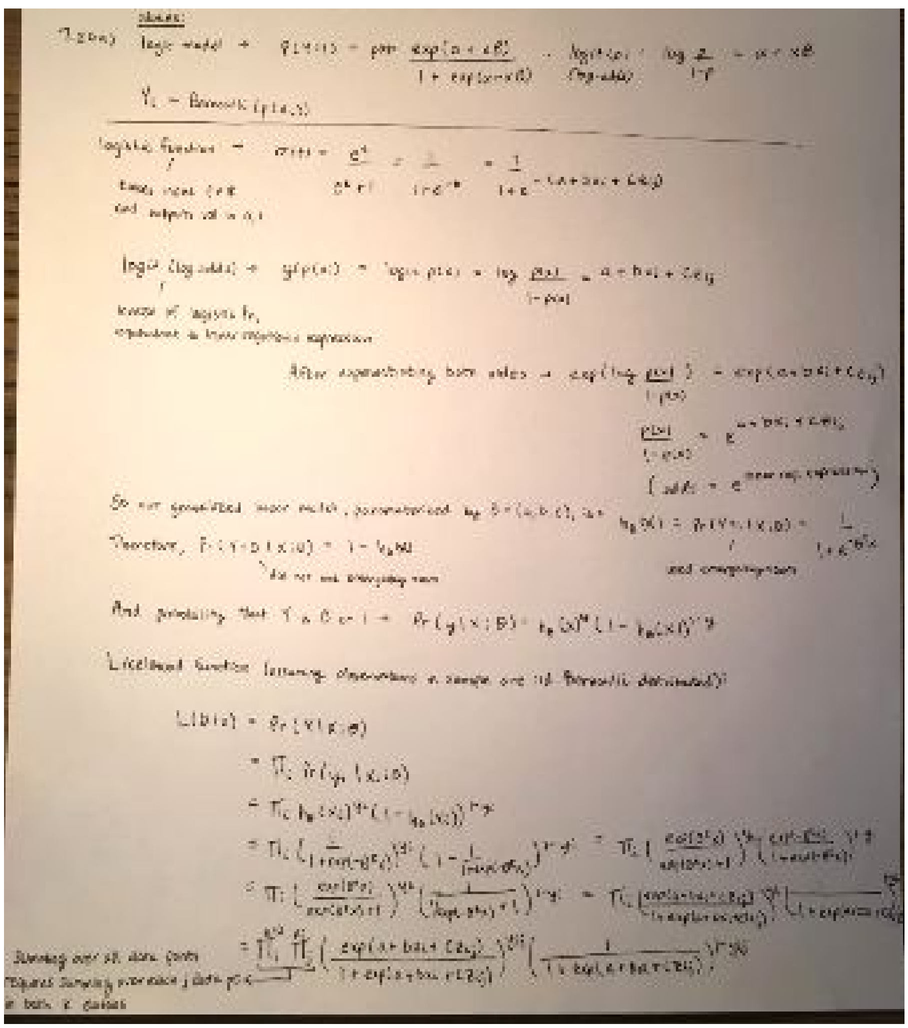
Problem 7.20: Here we use data from the Institute for Child Health Policy at the University of Florida to answer a question of interest: whether the status of the HMO affects the emergency room choice. An appropriate model is the logistic regression model,

$$\text{logit}(p_{ij}) = a + bx_i + cz_{ij}, \quad i = 1, \ldots, k, \quad j = 1, \ldots, n_i$$

where x_i is the HMO type, z_{ij} is the health status of the child, and p_{ij} is the probability of using an emergency room.

Problem 7.20a: First we verify the following likelihood function for the logistic regression model:

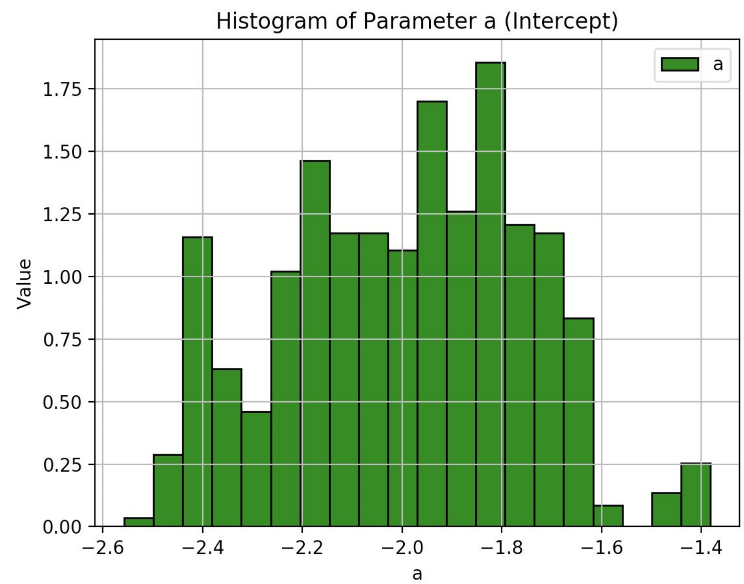
$$\prod_{i=1}^{k}\prod_{j=1}^{n_i}\left(\frac{\exp(a+bx_i+cz_{ij})}{1+\exp(a+bx_i+cz_{ij})}\right)^{y_{ij}}\left(\frac{1}{1+\exp(a+bx_i+cz_{ij})}\right)^{1-y_{ij}}$$

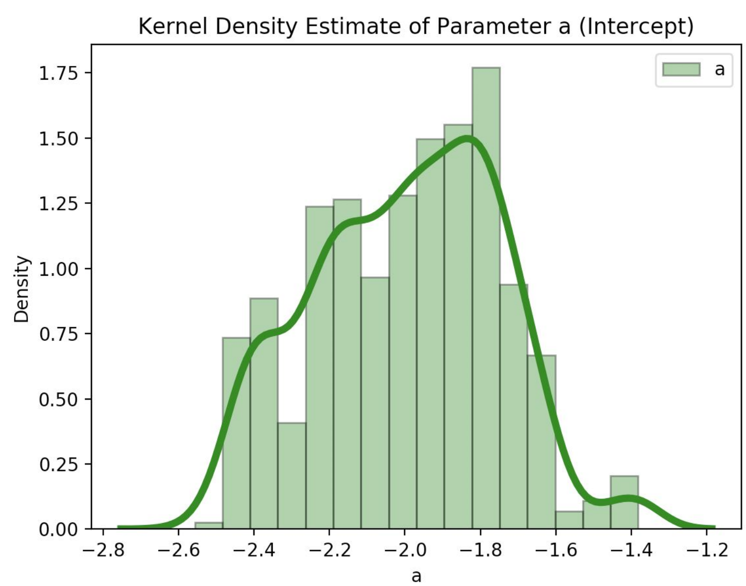

7.20a)
logistic function
logit (log odds)
Likelihood function

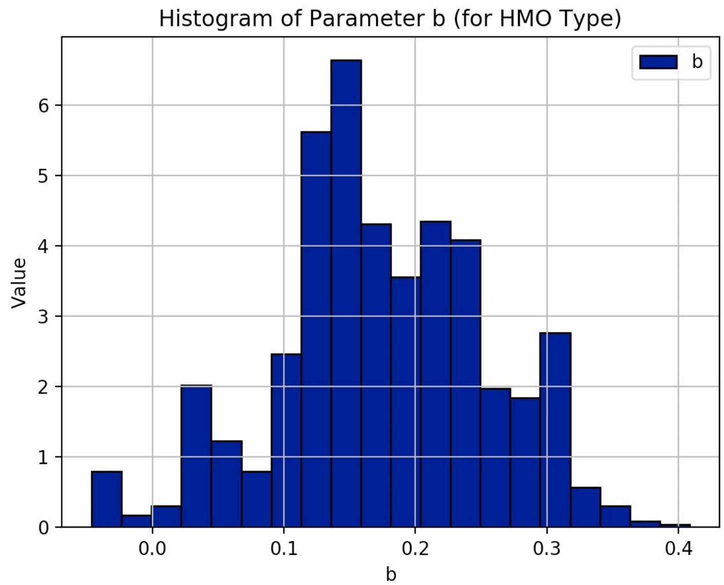
<u>Problem 7.20b</u>: We can run a standard GLM on the data and get the estimated mean, standard error, and variance of *a*, *b*, and *c*.

The mean, standard error, and variance of *a* are **-1.9739**, **0.221**, **0.0489**, respectively.
The mean, standard error, and variance of *b* are **0.1622**, **0.080**, **0.0064**, respectively.
The mean, standard error, and variance of *c* are **0.2844**, **0.093**, **0.0086**, respectively.

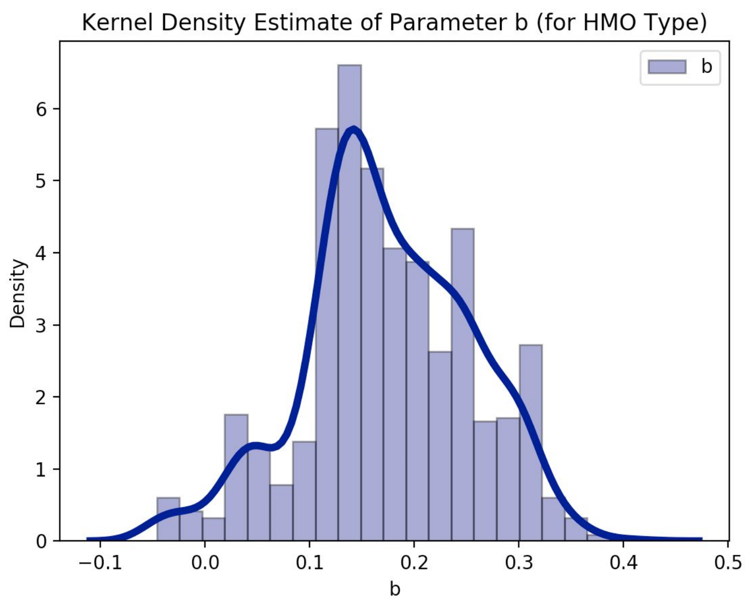
<u>Problem 7.20c</u>: Now we employ normal candidate densities, using the GLM estimates for mean and variance, in a Metropolis-Hastings algorithm that samples from the likelihood function. Below are histograms and kernel density estimates of the parameter values using this method.



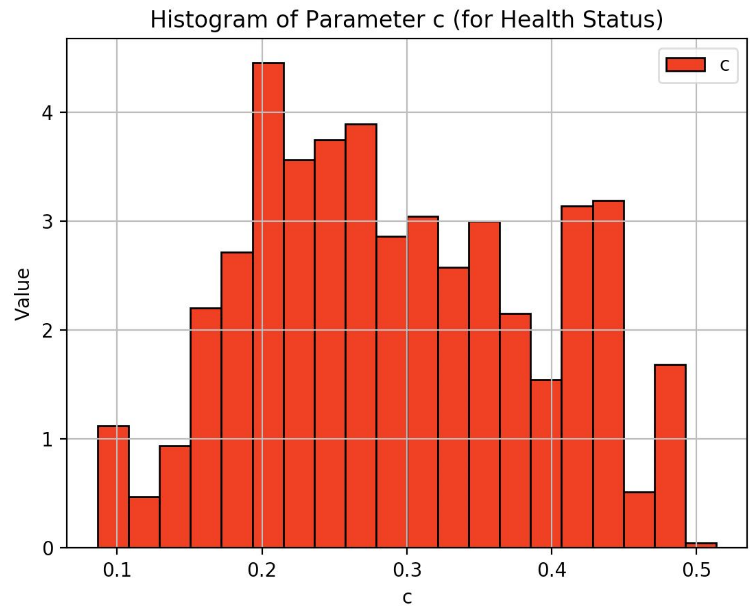


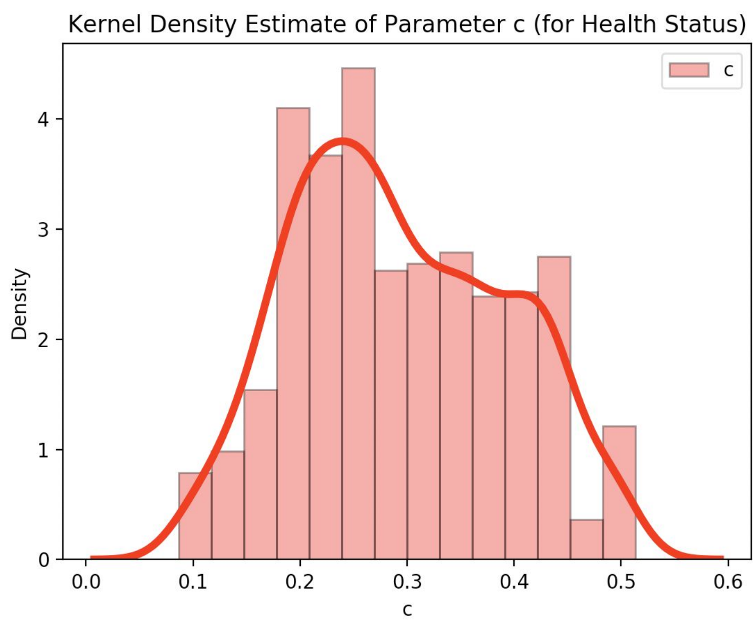


Histogram of Parameter b (for HMO Type)
b
Value
0
1
2
3
4
5
6
0.0
0.1
0.2
0.3
0.4
b



Kernel Density Estimate of Parameter b (for HMO Type)
b
Density
0
1
2
3
4
5
6
−0.1
0.0
0.1
0.2
0.3
0.4
0.5
b


Histogram of Parameter c (for Health Status)
c
Value
0
1
2
3
4
0.1
0.2
0.3
0.4
0.5
c


Kernel Density Estimate of Parameter c (for Health Status)
c
Density
0
1
2
3
4
0.0
0.1
0.2
0.3
0.4
0.5
0.6
c