STA 601 - Homework 3

Kedar Prabhudesai

September 11, 2013

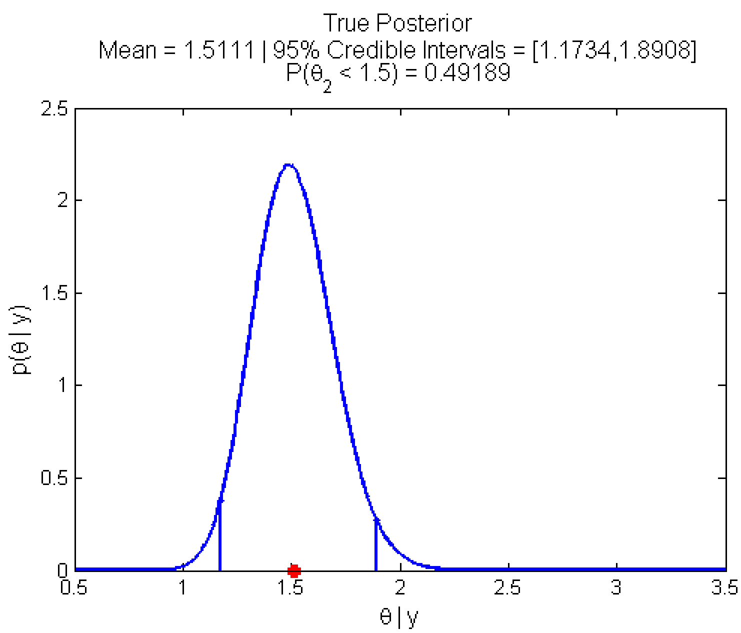
Gamma-Poisson Model:

Data Given: $n = 44, \sum y = 66$.

Prior: $\theta \sim Gamma(a, b)$. $a = 2, b = 1$.

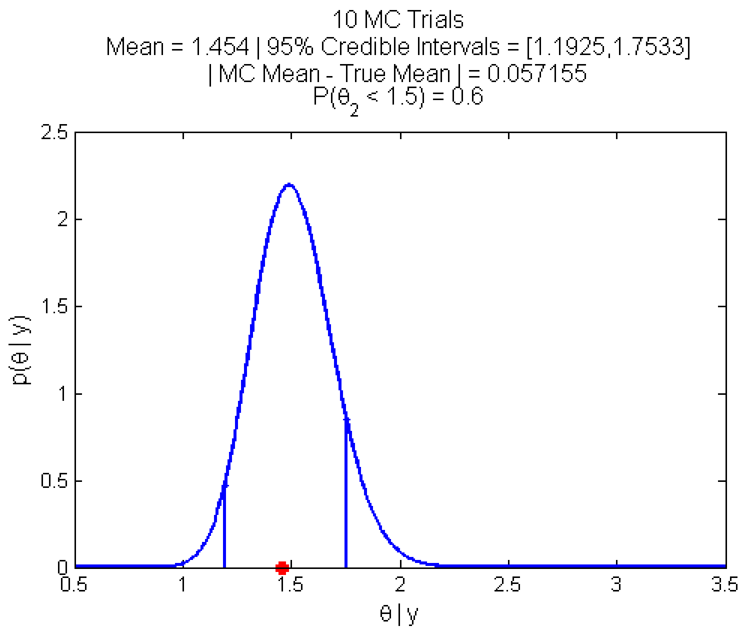
Posterior: $\theta|y \sim Gamma(a + \sum y, b + n)$

First we plot the Posterior Density function. (All the required information is provided in the title.)

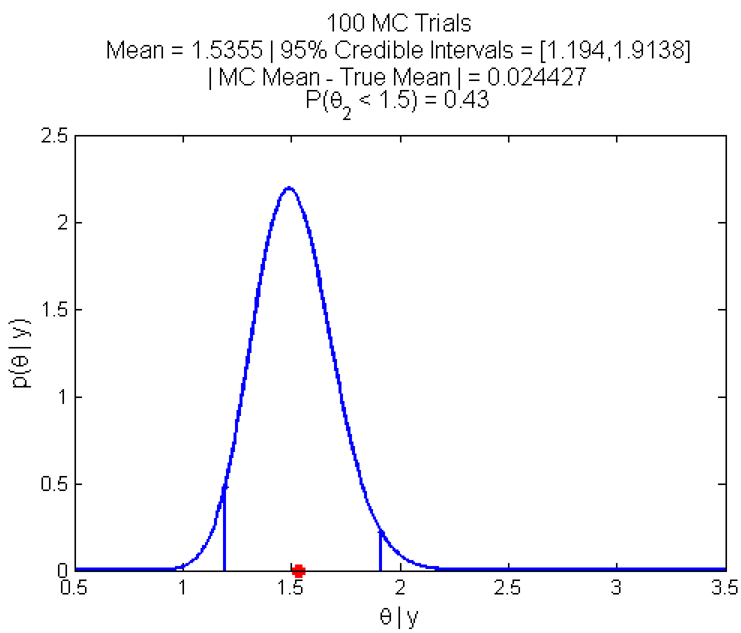


Next we will use different number of samples for MC Simulations and see the results for the same.

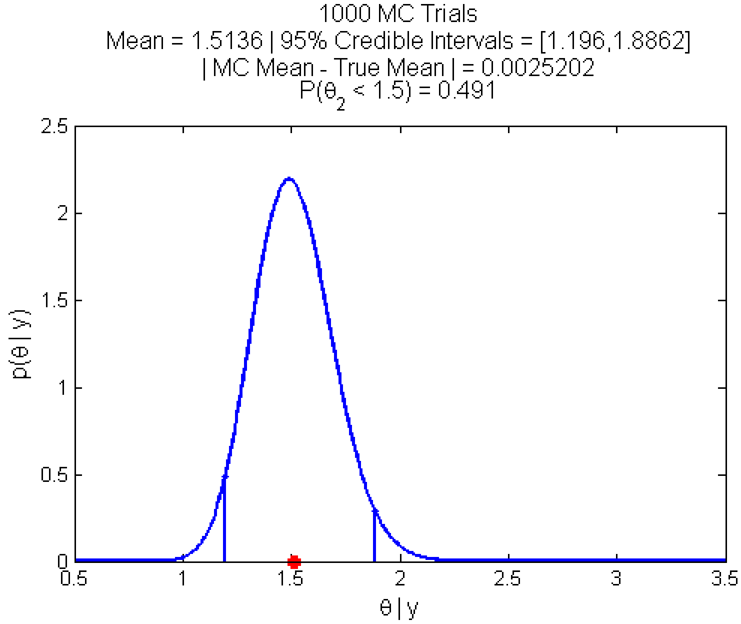
10 Monte-Carlo Trials:



100 Monte-Carlo Trials:



1000 Monte-Carlo Trials:



We observe that as the number of samples increase, the absolute error between the estimated mean and the mean from MC Simulations goes down.

The Hoff Book provides a method to estimate the number of MC Samples to use by calculating the Monte-Carlo Standard Error. For $S = 1000$, our observed $Var[\theta|y_1, ..., y_n] = 0.03226$. For the difference between $E[\theta|y_1, ..., y_n]$ and its MC-Estimate to be ≤ 0.001 (with very high probability), the following equation should hold. $2\sqrt{0.03226/S} \leq 0.001$. Using this equation we get, $S \geq 129,078$!!!

Appendix:

```
%% STA 601 — Homework 3
% Author: Kedar Prabhudesai
% Created on: 9/11/2013

close all;
clear all;

%% Make Distributions
% Gamma Prior Parameters
a = 2; b = 1;
% Data for Women with Bachelors degree
n = 44;
sumy = 66;
% Posterior
Posterior = makedist('Gamma','a',a+sumy,'b',1/(b+n));
% True mean of the Posterior
MeanFromPosterior = Posterior.mean();
% Probability of Interest from Distribution P(theta2 < 1.5)
ProbFromPosterior = Posterior.cdf(1.5);
% Lower and Upper 95% Bounds from Posterior
PostCredIntervals = Posterior.icdf([0.025 0.975]);

% Support for pdf
x = 0:0.01:4;
% Plot stuff
figure;
plot(x,Posterior.pdf(x),'Linewidth',2);hold on;
stem(PostCredIntervals,Posterior.pdf(PostCredIntervals),'Linewidth',2,'Marker','.','Color','b');
stem(MeanFromPosterior,0,'Linewidth',4,'Marker','*','Color','r');hold off;
title({'True Posterior',...
    ['Mean = ',num2str(MeanFromPosterior),' | 95% Credible Intervals = [',num2str(PostCredIntervals(1)),',',num2s
    ['P(\theta_2 < 1.5) = ',num2str(ProbFromPosterior)]},...
    'FontSize',12);
xlabel('\theta | y','FontSize',12);
ylabel('p(\theta | y)','FontSize',12);
xlim([0.5 3.5]);

%% Do Monte—Carlo Simulations
nTrials = [10 100 1000];
ProbFromMC = zeros(size(nTrials));
MeanFromMC = zeros(size(nTrials));
CIFromMC = zeros(numel(nTrials),2);
rng('shuffle');

for iTrials = 1:numel(nTrials)
    % Draw Random Samples from Posterior
    PostSamples = Posterior.random(1,nTrials(iTrials));
    % Estimate the mean
    MeanFromMC = mean(PostSamples);
    % Estimate P(theta2 < 1.5)
    ProbFromMC = sum(PostSamples < 1.5)/nTrials(iTrials);
    % Find 95% Credible Intervals
    CredIntervalFromMC = quantile(PostSamples,[0.025 0.975]);

    % Plot stuff
    figure;
    plot(x,Posterior.pdf(x),'Linewidth',2);hold on;
    stem(CredIntervalFromMC,Posterior.pdf(CredIntervalFromMC),'Linewidth',2,'Marker','.');
    stem(MeanFromMC,0,'Linewidth',4,'Marker','*','Color','r');hold off;
```

```
    title({[num2str(nTrials(iTrials)),' MC Trials'],...
        ['Mean = ',num2str(MeanFromMC),' | 95% Credible Intervals = [',num2str(CredIntervalFromMC(1)),',',num2str
        ['| MC Mean - True Mean | = ',num2str(abs(MeanFromPosterior-MeanFromMC))],...
        ['P(\theta_2 < 1.5) = ',num2str(ProbFromMC)]},...
        'FontSize',12);
    xlabel('\theta | y','FontSize',12);
    ylabel('p(\theta | y)','FontSize',12);
    xlim([0.5 3.5]);
end
```