Stat 221 Problem Set 4

Albert Young and Marco Gentili

4 November 2014

1 Question 1.1

$$p(\lambda, \theta) = p(\theta, \mu) \propto \frac{1}{\lambda} = \frac{1}{\mu\theta}$$

To determine the induced distribution on (N, θ), we marginalize over μ:

$$\begin{aligned}
p(N, \theta) &= p(N|\theta)p(\theta) \\
&= \int_0^\infty p(N|\theta, \mu)p(\theta, \mu)\, d\mu \\
&\propto \int_0^\infty \frac{e^{-\mu}\mu^N}{N!} \frac{1}{\theta\mu}\, d\mu \\
&= \frac{1}{N!\theta} \int_0^\infty e^{-\mu}\mu^{N-1}\, d\mu \\
&= \frac{\Gamma(N)}{N!\,\theta} \\
&= \frac{(N-1)!}{N(N-1)!\,\theta} \\
&= \frac{1}{N\theta}
\end{aligned}$$

...giving us a nice polynomial form.

2 Question 1.2

No, $p(\lambda, \theta)$ is not a proper distribution. $\int_0^\infty \frac{1}{\lambda}\, d\lambda$ does not converge.

3 Question 1.3

$$\begin{aligned}
Y_i &\sim Pois(\theta \cdot \mu) \\
p(Y_i|\theta, \mu) &= \frac{(\theta\mu)^{Y_i}}{(Y_i)!} \exp(-\theta\mu) \\
\log p(Y_i|\theta, \mu) &= -\theta\mu + Y_i \log(\theta\mu) + C
\end{aligned}$$

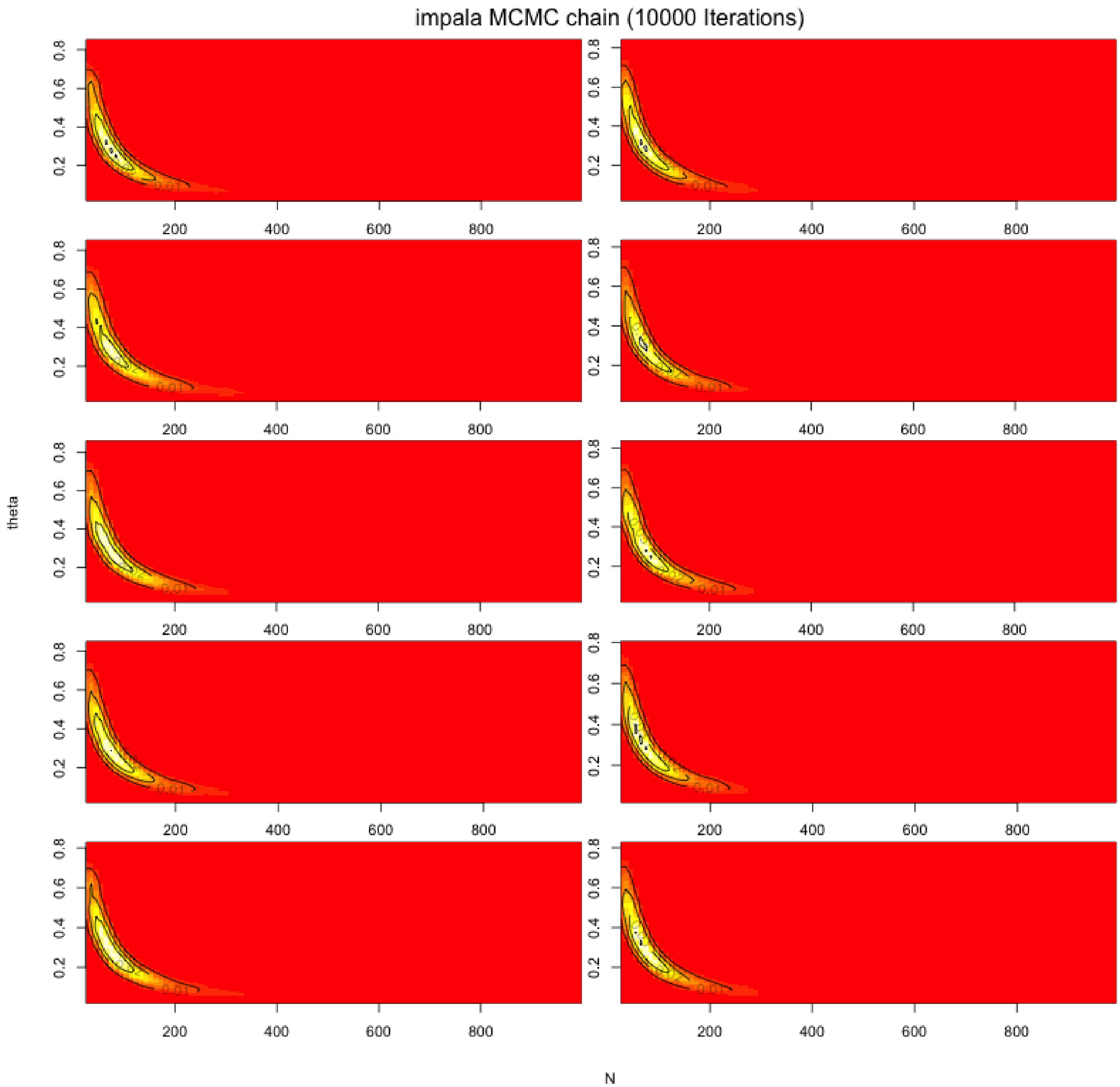
where C does not depend on θ or μ. Calculating the Fisher Information:

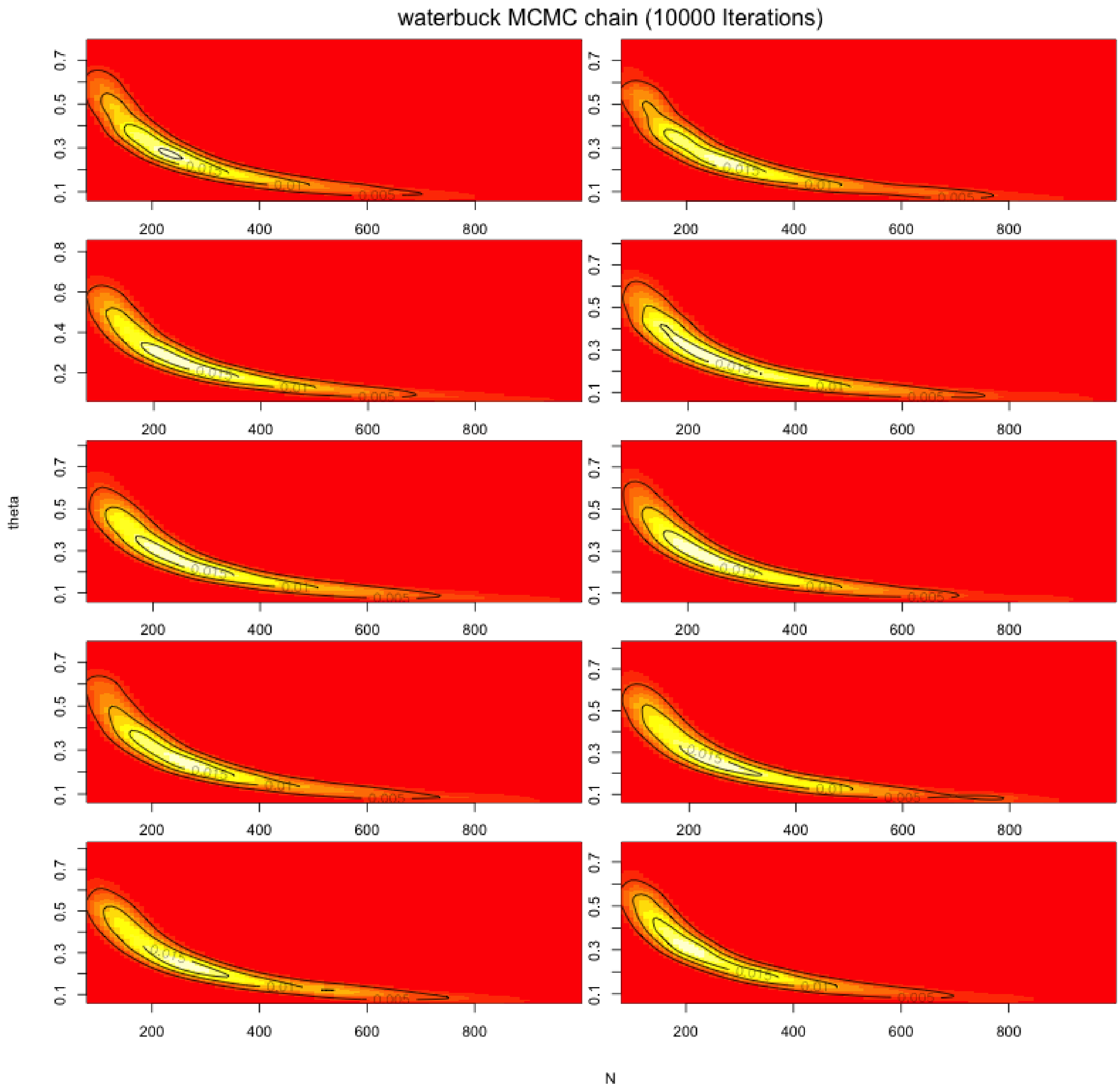
$$\begin{aligned}
\frac{\partial^2}{\partial \mu^2} \log p(Y_i|\mu,\theta) &= -\frac{Y_i}{\mu^2} \\
-E\left[\frac{\partial^2}{\partial \mu^2} \log p(Y_i|\mu,\theta)\right] &= \frac{\theta}{\mu} \\
\frac{\partial^2}{\partial \theta^2} \log p(Y_i|\mu,\theta) &= -\frac{Y_i}{\theta^2} \\
-E\left[\frac{\partial^2}{\partial \theta^2} \log p(Y_i|\mu,\theta)\right] &= \frac{\mu}{\theta} \\
\frac{\partial^2}{\partial \mu \partial \theta} \log p(Y_i|\mu,\theta) &= -1 \\
-E\left[\frac{\partial^2}{\partial \mu \partial \theta} \log p(Y_i|\mu,\theta)\right] &= 1 \\
det(I(\theta,\mu)) &= \frac{\theta}{\mu}\frac{\mu}{\theta} - 1 = 0
\end{aligned}$$

So $p(\lambda, \theta)$ is not non-informative in the sense of Jeffreys.

4 Question 1.4

We ran MCMC for 10,000 iterations. We sampled $S = N\theta \sim Beta(\sum Y_i - 1, Nn - \sum Y_i)$. Then, $N \sim Geom(\frac{\theta}{S})$. Our acceptance ratio is around 55% for both datasets.


impala MCMC chain (10000 Iterations)
theta
N
0.2
0.4
0.6
0.8
200
400
600
800
0.01


waterbuck MCMC chain (10000 Iterations)
theta
N
0.015
0.01
0.005

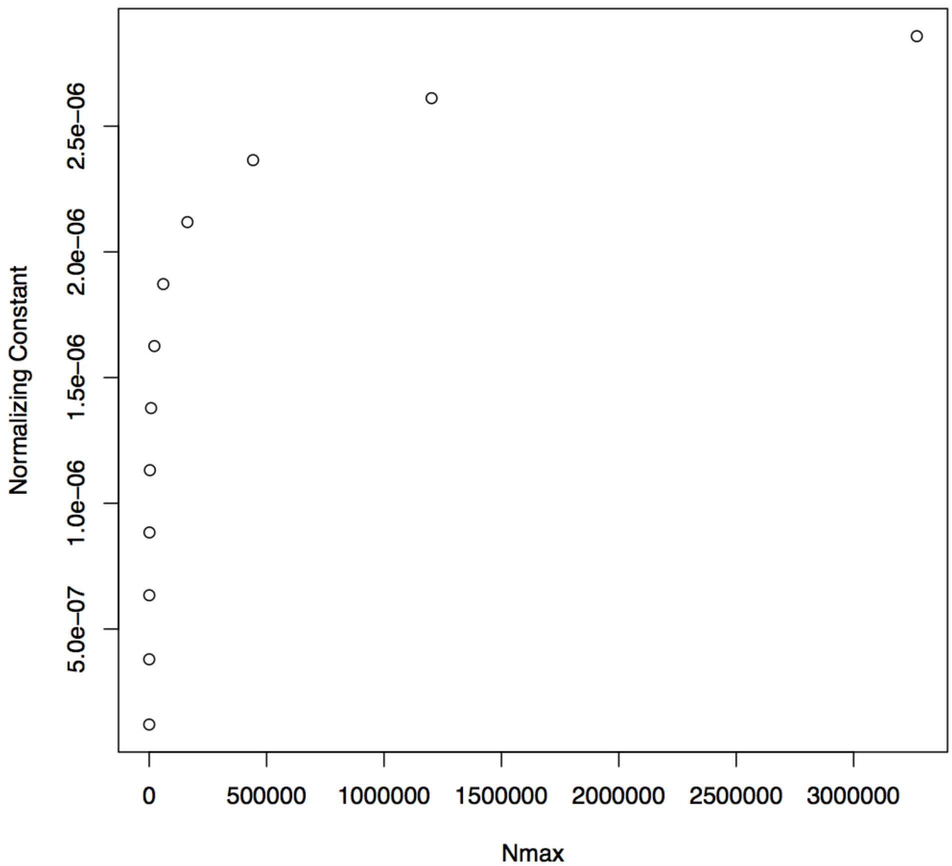
5 Question 1.5

$$
\begin{aligned}
p(N,\theta|Y) &= \frac{p(Y|N,\theta)p(N,\theta)}{\int p(Y|N,\theta)dNd\theta} \\
p(N|Y) &= C\int_0^1 p(Y|N,\theta)p(N,\theta)d\theta \\
&= \frac{C}{N}\prod_{i=1}^{n}\binom{N}{Y_i}\int_0^1 \theta^{\sum Y_i-1}(1-\theta)^{Nn-\sum Y_i}d\theta \\
&= \frac{C}{N}\prod_{i=1}^{n}\binom{N}{Y_i}\frac{\Gamma(\sum Y_i)\Gamma(Nn-\sum Y_i+1)}{\Gamma(Nn+1)} \\
&= \frac{C}{N}\prod_{i=1}^{n}\binom{N}{Y_i}\left(\frac{\sum Y_i(\sum Y_i-1)!(Nn-\sum Y_i)!}{(Nn)!\sum Y_i}\right) \\
&= \frac{C}{N}\frac{1}{\binom{Nn}{\sum Y_i}\sum Y_i}\prod_{i=1}^{n}\binom{N}{Y_i} \\
\sum_N p(N|Y) &= \sum_N\left(\frac{C}{N}\left(\frac{1}{\binom{Nn}{\sum Y_i}\sum Y_i}\right)\prod_{i=1}^{n}\binom{N}{Y_i}\right) = \frac{1}{C}
\end{aligned}
$$

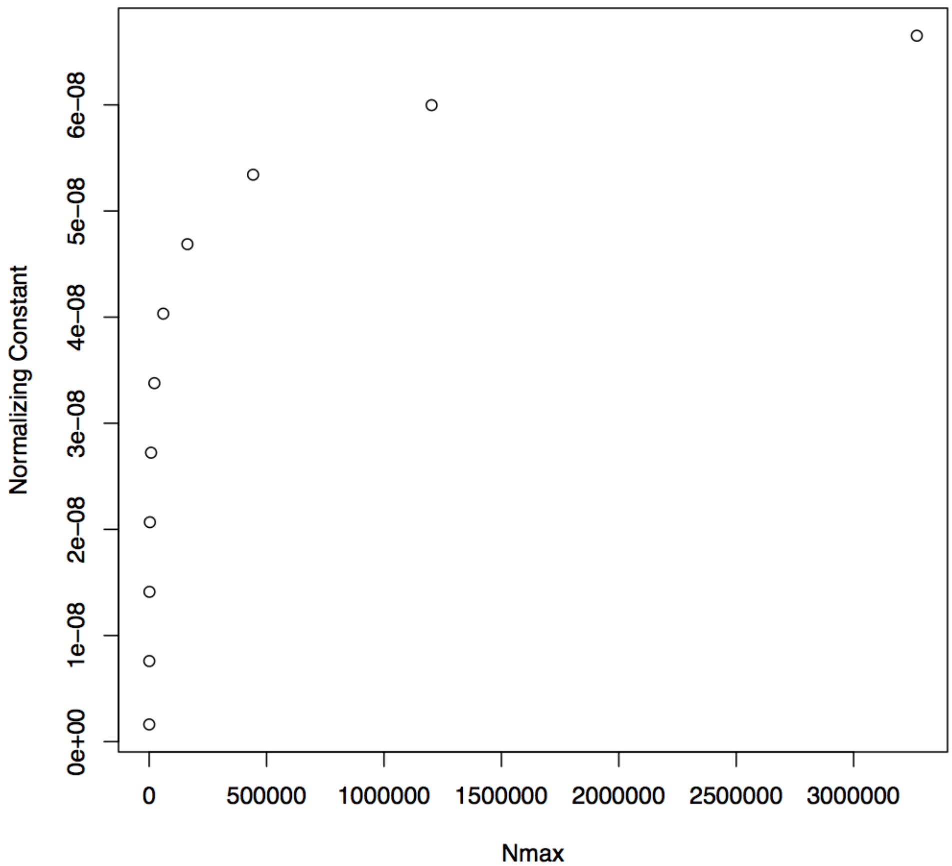
Taking advantage of $\theta^{\sum Y_i-1}(1-\theta)^{Nn-\sum Y_i} \sim Beta(\sum Y_i, Nn-\sum Y_i+1)$. $\frac{1}{C}$ is our normalizing constant.

We can't sum can't sum across Ns to infinity so we have an upper cutoff N_{max}. Summing from $N=1$ to $N_{max} = e^{15} = 3269017$, we obtain a normalizing constant of $4.687545e-08$ for the impala data and $6.651808e-08$ for the waterbuck data.

IMPALA DATA:



WATERBUCK DATA:



6 Question 1.6

For the posterior probability that $N > 100$, we obtained .4443 for the impala and .9734 for the waterbuck datasets, respectively.

The posterior analytics for IMPALA for 10 chains:

mean of N	theta mean	N sd	theta sd
562.6458	0.277963067889716	7734.6927528985	0.192755573974767
606.7675	0.283222893222428	10994.9857642763	0.196222534258168
529.8685	0.266034858270702	7004.90443723719	0.192651141217575
310.7725	0.271150421612974	1604.4739139038	0.190313579905066
299.0939	0.273036193335137	1470.02884024835	0.19483950133021
364.636	0.275585642568672	4773.39335350804	0.188345061657908
343.8988	0.275394386473951	2230.53458126306	0.194183567944313
237.8746	0.281194063195724	748.629962102009	0.189712764848506
249.0048	0.275086596973311	817.206296102764	0.189637013466521
284.8015	0.279820325894997	1368.82568191589	0.192675396725995

The posterior analytics for WATERBUCK for 10 chains:

N mean	theta mean	N sd	theta sd
1808.7721	0.252534661435681	34864.3725616162	0.176124672730363
3650.5406	0.230168067949625	64985.6590189971	0.176697277434169
1733.8773	0.23790001011818	21642.9758183035	0.179266390848925
1032.0534	0.238284848648184	4478.15529903069	0.174011580648649
1761.6458	0.230391570960453	17656.2429029136	0.173322859129494
834.8959	0.245838955571044	3210.58914284086	0.17751790617653
2060.0502	0.236212304212799	24943.6691832118	0.176469790227135
1592.9866	0.241322634819646	12971.1625989168	0.179918973072355
1358.5298	0.238231696529153	11611.0746579799	0.174486573154321
1560.9429	0.242356503626048	10597.3512770243	0.176531597736815