

Homework Example

Bayesian Statistical Methods

Mevin Hooten

Department of Statistics and Data Sciences

The University of Texas at Austin

Email: mevin.hooten@austin.utexas.edu

Zhendong Wang

Department of Statistics and Data Sciences

The University of Texas at Austin

Email: zhendong.wang@utexas.edu

January 20, 2022

1 Write a Bayesian model with an exponential likelihood, assuming conditional independence among eruptions, and assume an exponential distribution for the rate parameter in the data model

We developed the following Bayesian model for wait times between Old Faithful eruptions:

$$y_i \sim \text{Exp}(\lambda) , \quad (1)$$

$$\lambda \sim \text{Exp}(\mu) , \quad (2)$$

where, y_i are the wait times (i.e., response) in minutes between individual eruptions for $i = 1, \dots, n$ and λ is the unknown rate parameter that is assumed to arise from another exponential distribution *a priori* with mean $1/\mu$.

2 Derive the posterior distribution for the rate parameter analytically

The posterior distribution associated with the above model can be written as

$$[\lambda|\mathbf{y}] \propto [\mathbf{y}|\lambda][\lambda] , \quad (3)$$

$$\propto \left(\prod_{i=1}^n \lambda \exp(-\lambda y_i) \right) \mu \exp(-\mu\lambda) , \quad (4)$$

$$\propto \lambda^n \exp \left(-\lambda \left(\sum_{i=1}^n y_i + \mu \right) \right) , \quad (5)$$

$$\propto \lambda^{(n+1)-1} \exp \left(- \left(\sum_{i=1}^n y_i + \mu \right) \lambda \right) . \quad (6)$$

Thus, the simplified expression above is proportional to a Gamma probability density function parameterized as $\text{Gamma}(n+1, \sum_{i=1}^n y_i + \mu)$. We note that this is a conj-

gate posterior distribution considering that the exponential prior is also a simplified Gamma distribution.

With an analytically tractable posterior, we can either use its known properties directly for inference (e.g., the posterior mean of λ is $[\lambda|\mathbf{y}] = \frac{n+1}{\sum_{i=1}^n y_i + \mu}$), or we can obtain a Monte Carlo sample from it directly that can be used to compute derived quantities such as the posterior mean of $1/\lambda$ as presented in what follows. Finally, another correct, but less computationally efficient, approach to obtain a sample from the posterior distribution is using a Markov Chain Monte Carlo (MCMC) algorithm, as we describe next.

3 Develop an MCMC algorithm to fit the model above that uses a Metropolis-Hastings update for the rate parameter. Use an exponential distribution $\lambda^{(*)} \sim \text{Exp}(1/\lambda^{(k-1)})$ for the proposal

To construct an MCMC algorithm for our model with Metropolis-Hastings updates for λ based on the exponential proposal distribution, we used the starting value $\lambda^{(0)} = 1/\bar{\mathbf{y}}$ and the following Metropolis-Hastings ratio for sequential updates

$$mh = \frac{(\prod_{i=1}^n [y_i|\lambda^{(*)}]) [\lambda^{(*)}][\lambda^{(k-1)}|\lambda^{(*)}]}{(\prod_{i=1}^n [y_i|\lambda^{(k-1)}]) [\lambda^{(k-1)}][\lambda^{(*)}|\lambda^{(k-1)}]}, \quad (7)$$

for MCMC iterations $k = 1, \dots, K$. We let $\lambda^{(k)} = \lambda^{(*)}$ with probability $\min(mh, 1)$ and let $\lambda^{(k)} = \lambda^{(k-1)}$ otherwise. This MCMC algorithm did not require tuning because the proposal was an exponential random walk based on the most recent value of the rate parameter. See the Appendix for the full MCMC algorithm written in R.

4 Specify and justify the hyperparameter associated with the prior for the rate parameter

We specified the hyperparameter as $\mu = 75$ because the average wait time is approximately 75 minutes for a similar geyser (i.e., Lion geyser) according to the National Park Service (<https://www.nps.gov/yell/planyourvisit/geyser-activity.htm>).

5 Fit the model to the Old Faithful data set using the MCMC algorithm and report the following:

We fit the exponential Bayesian model to the geyser data using the prior specified as above and our MCMC algorithm (as well as the analytical posterior) and $K = 100000$ MCMC iterations. Based on visual inspection of the trace plot (below), the chain converged immediately and did not require a burn-in period.

5.1 An MCMC trace plot for the rate parameter

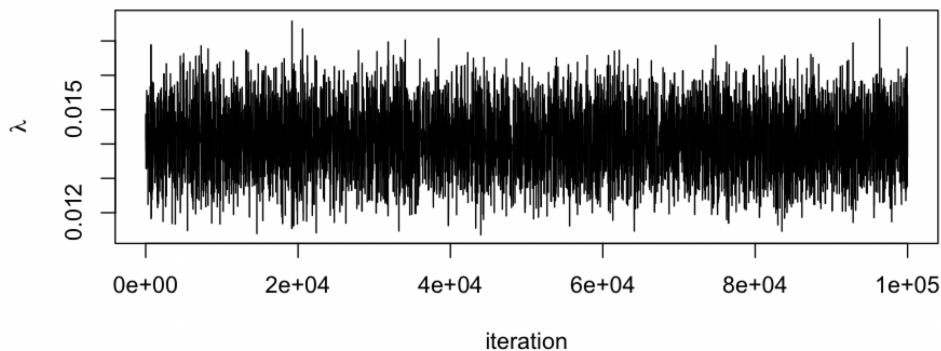


Figure 1: MCMC trace plot for λ based on $K = 100000$ iterations.

5.2 Metropolis-Hastings acceptance probability for the rate parameter

Using our MCMC algorithm with Metropolis-Hastings updates, we accepted proposals for λ 7101 times (out of $K = 100000$ total iterations) for an acceptance probability of 0.071.

5.3 Posterior mean, standard deviation, and 95% equal-tailed credible interval of the rate parameter

Using our MCMC sample for λ , the posterior mean was $E(\lambda|\mathbf{y}) = 0.014$, the posterior standard deviation was $\sqrt{\text{Var}(\lambda|\mathbf{y})} = 0.000849$, and posterior 95% credible interval is: (0.0125,0.0158).

5.4 Posterior mean, standard deviation, and 95% equal-tailed credible interval of the scale (reciprocal of rate) parameter

Using our MCMC sample for $1/\lambda$ (a transformation or derived quantity of λ), the posterior mean was $E(1/\lambda|\mathbf{y}) = 71.18$, the posterior standard deviation was $\sqrt{\text{Var}(1/\lambda|\mathbf{y})} = 4.314$, and posterior 95% credible interval is: (63.317,80.333).

6 What can we infer about the average wait time between Old Faithful eruptions under our assumed model?

Based on results presented above regarding the posterior summary for $1/\lambda$, we can infer that the point estimate for the model-based average wait time between geyser eruptions is approximately 71 minutes but that the average wait time falls between approximately 63.3 and 80.3 minutes with 95% probability.

7 Create a graphic comparing an MCMC-based posterior histogram and the posterior density function for the rate parameter based on the analytical solution

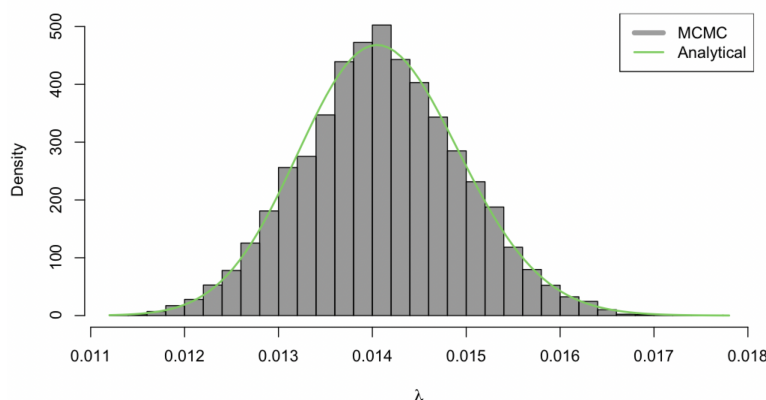


Figure 2: Comparison of MCMC-based posterior histogram and analytical posterior probability density function.

8 Qualitatively, discuss the appropriateness of the chosen likelihood for the Old Faithful wait time data

In principle, the exponential data model is appropriate for modeling time-to-event data such as wait times. However, the exponential distribution itself is “memoryless” which loosely means that after we have waited a certain amount of time, the probability of waiting that amount of time again is the same. This may not be an appropriate characteristic of the geyser mechanisms which depend on sufficient heat and water pressure accumulating inside an underground chamber. A more flexible time-to-event distribution (e.g., Weibull) may be more appropriate in that setting. Furthermore, the Old Faithful geyser data clearly indicate a bimodal wait time distribution (Fig-

ure 3) which suggests that a mixture distribution may be more appropriate as a model for these data.

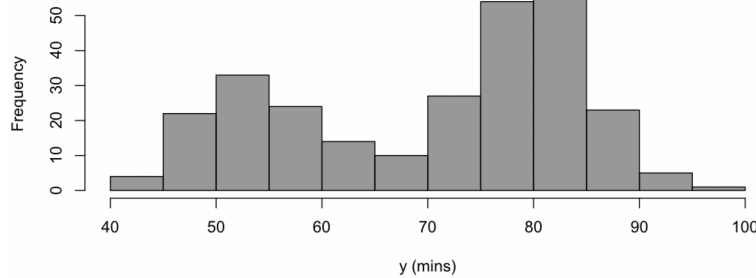


Figure 3: Histogram of Old Faithful geyser data.

9 Describe an alternative proposal distribution for the Metropolis-Hastings updates

We used an exponential distribution parameterized as a random walk for the proposal distribution in our analyses, but our acceptance probability was low at approximately 0.07. Instead, we could have used a more flexible proposal such as a truncated normal distribution with location parameter set to the previous value for the rate parameter $\lambda^{(k-1)}$ and truncated below at zero. This distribution includes a tuning parameter σ^2 that we could set to adjust the acceptance probability. Alternatively, we could have used a more flexible gamma distribution as the proposal which also includes two parameters, but we would have to moment-match it so that the mean of the gamma equals $\lambda^{(k-1)}$ and then tune the remaining function of parameters to achieve a suitable acceptance probability. We could have also used the prior as a proposal distribution which would simplify the Metropolis-Hastings ratio, but is very inflexible. There exist a variety of alternatives including the posterior distribution itself as the proposal in this case which would result in Gibbs updates and a Monte Carlo sample ultimately

because we only have a single parameter model. The main benefit of the exponential proposal we used is that it is capable of exploring the space of λ due to the random walk and does not require tuning.

Author Contributions

- MH developed and fit the model using MCMC, analyzed the data set, and wrote the homework solution document.
- ZW derived the posterior distribution, checked the MCMC code, analyzed the data set, and reviewed the homework solution document.

APPENDIX A: R Script

```
###
###  Read in Data
###

faithful.df=read.csv("faithful.csv",header=TRUE)
head(faithful.df)

y=faithful.df$Eruption.wait..mins.
n=length(y)
hist(y,col=rgb(0,0,0,.4),main="",xlab="y (mins)")

###
###  Fit Model using MCMC
###

source("exp.mcmc.R")
n.mcmc=100000
mu=75
out.1=exp.mcmc(y,mu,n.mcmc)

plot(out.1$lam.save,type="l",xlab="iteration",ylab=bquote(lambda)) # make trace plot
out.1$lam.acc/out.1$n.mcmc # MH acceptance prob

###
###  Make Inference
###

mean(out.1$lam.save)
sd(out.1$lam.save)
quantile(out.1$lam.save,c(0.025,0.975))

mean(1/out.1$lam.save)
sd(1/out.1$lam.save)
quantile(1/out.1$lam.save,c(0.025,0.975))

###
###  Compare Model Fit using Analytical Posterior
###

alpha.post=n+1
beta.post=sum(y)+mu

hist(out.1$lam.save,prob=TRUE,breaks=30,col=rgb(0,0,0,.4),main="",xlab=bquote(lambda))
curve(dgamma(x,alpha.post,beta.post),col=3,lwd=2,add=TRUE)
legend("topright",col=c(8,3),lwd=c(5,2),legend=c("MCMC","Analytical"))
```

APPENDIX B: MCMC Algorithm

```
exp.mcmc <- function(y,mu,n.mcmc){

###
### Setup Variables
###

n=length(y)
lam.save=rep(0,n.mcmc)
lam.acc=0

###
### Starting Value
###

lam=1/mean(y)

###
### Being MCMC Loop
###

for(k in 1:n.mcmc){
  if(k%%1000==0){cat(k," ")}

  ###
  ### Update lambda
  ###

  lam.star=rexp(1,1/lam)
  mh.1=sum(dexp(y,lam.star,log=TRUE))+dexp(lam.star,mu,log=TRUE)+dexp(lam,1/lam.star,log=TRUE)
  mh.2=sum(dexp(y,lam,log=TRUE))+dexp(lam,mu,log=TRUE)+dexp(lam.star,1/lam,log=TRUE)
  mh=exp(mh.1-mh.2)

  if(mh>runif(1)){
    lam=lam.star
    lam.acc=lam.acc+1
  }

  ###
  ### Save sample
  ###

  lam.save[k]=lam
};cat("\n")

###
### Write Output
###
```

```
list(lam.save=lam.save,lam.acc=lam.acc,mu=mu,n.mcmc=n.mcmc)
}
```