

Let  $\theta$  be a random variable with density  $p(\theta)$ —the target density—and  $\theta^*$  be a random variable with density  $g(\theta^*)$ —the proposal density. Further, let  $M$  be some finite, nonnegative value such that  $p(\theta) \leq Mg(\theta)$ . Prove that the rejection sampler (or accept/reject algorithm) results in samples from the target density. Hint: It is useful to begin by working with the CDF of  $\theta$ ,  $P(\theta \leq a)$ , and noting that conditional on the acceptance rule,  $P(\theta \leq a) = P(\theta^* \leq a | \text{accept } \theta^*)$ .

Let  $Y \sim f_Y(y)$  and  $V \sim f_V(V)$ , where  $f_Y$  and  $f_V$  have common support with  $M = \sup_y \frac{f_Y(y)}{Mf_V(y)} < \infty$

To generate a random variable  $Y \sim f_Y$ :

- a. Generate  $U \sim \text{Unif}(0, 1)$  and  $V \sim f_V$ , independent.
- b. If  $U < \frac{f_Y(V)}{Mf_V(V)}$ , set  $Y = V$ ; otherwise, return to step (a)

The generated random variable  $Y$  has cdf  $P(Y \leq y)$

$$\begin{aligned}
 P(Y \leq y) &= P(V \leq y | \text{accept}) \\
 &= P\left(V \leq y \mid U < \frac{f_Y(V)}{Mf_V(V)}\right) \\
 &= \frac{P\left(V \leq y, U < \frac{f_Y(V)}{Mf_V(V)}\right)}{P\left(U < \frac{f_Y(V)}{Mf_V(V)}\right)} \\
 &= \frac{\int_{-\infty}^y \int_0^{\frac{f_Y(v)}{Mf_V(v)}} du f_V(v) dv}{\int_{-\infty}^{\infty} \int_0^{\frac{f_Y(v)}{Mf_V(v)}} du f_V(v) dv} \\
 &= \int_{-\infty}^y f_Y(v) dv
 \end{aligned}$$

which is the desired cdf.

Let  $y_i \stackrel{iid}{\sim} N(\mu, \alpha\sigma^2)$  and place the following priors on the model components:  $\pi(\mu) \propto 1$ ,  $\alpha \sim IG(a, b)$ , and  $\sigma^2 \sim IG(c, d)$  for constants  $a, b, c$ , and  $d$ . This is an example of a parameter expanded normal model. Find the posterior and determine the full conditionals for the model. Using these, write the steps necessary to use a Gibbs sampler to draw samples from  $\mu$ ,  $\alpha$ , and  $\sigma^2$ .

$$\mathcal{L}(y_i|\mu, \sigma^2, \alpha) \propto (\alpha\sigma^2)^{-\frac{n}{2}} \exp\left[-\frac{1}{2\alpha\sigma^2} \sum_i (y_i - \mu)^2\right]$$

$$\begin{aligned} p(\mu, \alpha, \sigma^2|y) &\propto (\alpha\sigma^2)^{-\frac{n}{2}} \exp\left[-\frac{1}{2\alpha\sigma^2} \sum_i (y_i - \mu)^2\right] \frac{b^a}{\Gamma(a)} \alpha^{-(a+1)} \exp\left[-\frac{b}{\alpha}\right] \frac{d^c}{\Gamma(c)} (\sigma^2)^{-(c+1)} \exp\left[-\frac{d}{\sigma^2}\right] \\ &\propto (\alpha\sigma^2)^{-\frac{n}{2}} \exp\left[-\frac{1}{2\alpha\sigma^2} \sum_i (y_i - \mu)^2\right] \alpha^{-(a+1)} \exp\left[-\frac{b}{\alpha}\right] (\sigma^2)^{-(c+1)} \exp\left[-\frac{d}{\sigma^2}\right] \end{aligned}$$

$$\begin{aligned} p(\alpha|\mu, \sigma^2, y) &\propto \alpha^{-\frac{n}{2}} \exp\left[-\frac{1}{2\alpha\sigma^2} \sum_i (y_i - \mu)^2\right] \alpha^{-(a+1)} \exp\left[-\frac{b}{\alpha}\right] \\ &\propto \alpha^{-(\frac{n}{2}+a+1)} \exp\left[-\frac{1}{\alpha} \left(b + \frac{1}{2\sigma^2} \sum_i (y_i - \mu)^2\right)\right] \end{aligned}$$

$$\alpha|\mu, \sigma^2, y \sim IG\left(\frac{n}{2} + a, b + \frac{1}{2\sigma^2} \sum_i (y_i - \mu)^2\right)$$

$$\begin{aligned} p(\sigma^2|\mu, \alpha, y) &\propto (\sigma^2)^{-\frac{n}{2}} \exp\left[-\frac{1}{2\alpha\sigma^2} \sum_i (y_i - \mu)^2\right] (\sigma^2)^{-(c+1)} \exp\left[-\frac{d}{\sigma^2}\right] \\ &\propto (\sigma^2)^{-(\frac{n}{2}+c+1)} \exp\left[-\frac{1}{\sigma^2} \left(d + \frac{1}{2\alpha} \sum_i (y_i - \mu)^2\right)\right] \end{aligned}$$

$$\sigma^2|\mu, \alpha, y \sim IG\left(\frac{n}{2} + c, d + \frac{1}{2\alpha} \sum_i (y_i - \mu)^2\right)$$

$$\begin{aligned}
p(\mu|\sigma^2, \alpha^2, y) &\propto \exp \left[ -\frac{1}{2\alpha\sigma^2} \sum_i (y_i - \mu)^2 \right] \\
&\propto \exp \left[ -\frac{1}{2\frac{\alpha\sigma^2}{n}} \sum_i (\mu - \bar{y})^2 \right] \\
\mu|\sigma^2, \alpha, y &\sim N \left( \bar{y}, \frac{\alpha\sigma^2}{n} \right)
\end{aligned}$$

Let  $x_i \stackrel{iid}{\sim} N(\mu, \sigma^2)$  and place the flat prior on  $\mu$ ,  $\pi(\mu) \propto 1$ . Next assume  $\sigma^2 \sim IG(v/2, v/\alpha)$  and  $\alpha \sim IG(1/2, 1/A^2)$  for fixed values  $v$  and  $A$ . When  $v = 1$ , this prior and hyperprior combination are equivalent to placing a Half-Cauchy prior with scale  $A$  on  $\sigma^2$ —a useful prior for variances, particularly when they are near zero. Express the posterior and determine the full conditionals for the model. Then write out the steps for the Gibbs sampler to draw samples from  $\mu$ ,  $\sigma^2$ , and  $\alpha$ .

$$\mathcal{L}(x_i|\mu, \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}} \exp \left[ -\frac{1}{2\sigma^2} \sum_i (x_i - \mu)^2 \right]$$

$$\begin{aligned} p(\mu, \sigma^2, \alpha|x) &\propto (\sigma^2)^{-\frac{n}{2}} \exp \left[ -\frac{1}{2\sigma^2} \sum_i (x_i - \mu)^2 \right] \frac{\left(\frac{v}{\alpha}\right)^{\frac{v}{2}}}{\Gamma\left(\frac{v}{2}\right)} (\sigma^2)^{-\left(\frac{v}{2}+1\right)} \exp \left[ -\frac{v}{\sigma^2} \right] \frac{\frac{1}{A^2}}{\Gamma\left(\frac{1}{2}\right)} \alpha^{-\left(\frac{1}{2}+1\right)} \exp \left[ -\frac{1}{A^2\alpha} \right] \\ &\propto (\sigma^2)^{-\left(\frac{n+v}{2}+1\right)} \exp \left[ -\frac{1}{2\sigma^2} \sum_i (x_i - \mu)^2 \right] \left(\frac{v}{\alpha}\right)^{\frac{v}{2}} \exp \left[ -\frac{v}{\alpha\sigma^2} \right] \alpha^{-1.5} \exp \left[ -\frac{1}{A^2\alpha} \right] \end{aligned}$$

$$\begin{aligned} p(\alpha|\mu, \sigma^2, y) &\propto \left(\frac{v}{\alpha}\right)^{\frac{v}{2}} \exp \left[ -\frac{v}{\alpha\sigma^2} \right] \alpha^{-1.5} \exp \left[ -\frac{1}{A^2\alpha} \right] \\ &\propto \alpha^{-\left(\frac{v+1}{2}+1\right)} \exp \left[ -\frac{1}{\alpha} \left( \frac{v}{\sigma^2} + \frac{1}{A^2} \right) \right] \end{aligned}$$

$$\alpha|\mu, \sigma^2, y \sim IG \left( \frac{v+1}{2}, \frac{v}{\sigma^2} + \frac{1}{A^2} \right)$$

$$\begin{aligned} p(\sigma^2|\mu, \alpha, y) &\propto (\sigma^2)^{-\left(\frac{n+v}{2}+1\right)} \exp \left[ -\frac{1}{2\sigma^2} \sum_i (x_i - \mu)^2 \right] \exp \left[ -\frac{v}{\alpha\sigma^2} \right] \\ &\propto (\sigma^2)^{-\left(\frac{n+v}{2}+1\right)} \exp \left[ -\frac{1}{\sigma^2} \left( \frac{v}{\alpha} + \frac{1}{2} \sum_i (x_i - \mu)^2 \right) \right] \end{aligned}$$

$$\sigma^2|\mu, \alpha, y \sim IG \left( \frac{n+v}{2}, \frac{v}{\alpha} + \frac{1}{2} \sum_i (x_i - \mu)^2 \right)$$

$$\begin{aligned}
p(\mu|\sigma^2, \alpha, y) &\propto \exp \left[ -\frac{1}{2\sigma^2} \sum_i (x_i - \mu)^2 \right] \\
&\propto \exp \left[ -\frac{1}{2\frac{\sigma^2}{n}} \sum_i (\mu - \bar{x})^2 \right] \\
\mu|\sigma^2, \alpha, y &\sim N \left( \bar{x}, \frac{\sigma^2}{n} \right)
\end{aligned}$$

Suppose we wish to sample from the triangular distribution which has density function

$$p(\theta) = \begin{cases} 0 & \theta < 0 \\ 4\theta & 0 \leq \theta < 1/2 \\ 4 - 4\theta & 1/2 \leq \theta \leq 1 \\ 0 & \theta > 1 \end{cases}.$$

This density arises as the mean of two uniformly distributed random variables. One approach to sampling from  $p(\theta)$  is to use rejection sampling. For your sampler, consider two different forms of  $g(\theta)$ : first, let  $g(\theta)$  be the standard uniform density and second, let  $g(\theta)$  be the  $Beta(2, 2)$  density. For both proposal densities, you will need to determine the values of  $M$  such that  $p(\theta) \leq Mg(\theta)$ . For each  $g(\theta)$ , first write a rejection sampler to obtain 1000 samples from  $p(\theta)$ . Then compare the resulting sampled distributions as well as the acceptance rate. Is one  $g(\theta)$  more efficient than the other? Explain. For each sampler, set the seed to 52918.

```
fx <- function(x){ ifelse( x < 0 , 0,
  ifelse( 0 <= x & x < 0.5 , 4*x ,
  ifelse( 0.5 <= x & x <= 1, 4-4*x,
  ifelse( x < 1, 0, 0 )
)))}

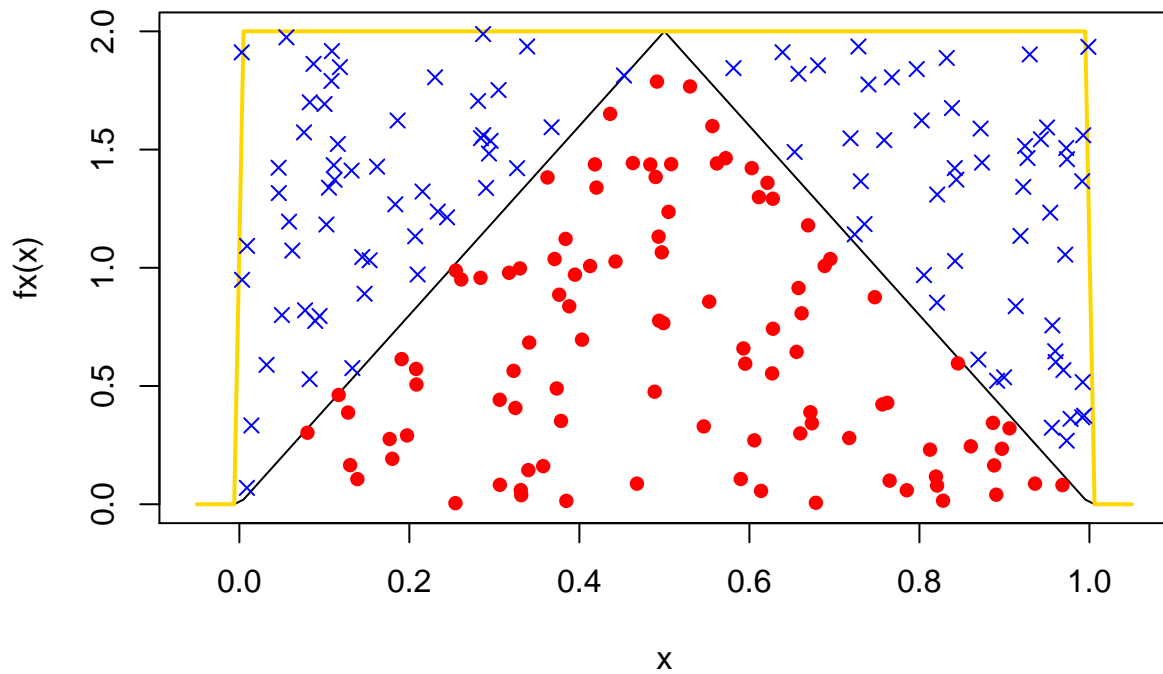
### U(0,1) ###
curve(fx, from = 0-.05, to = 1+.05)
M <- 2
curve(M *dunif(x, 0, 1), add=TRUE, col="gold", lwd=2)

theta <- vector(length = 1000)
arr <- NULL
t <- 1
count <- 1

while(t < 100){

  tb <- runif(1)
  U <- runif(1)

  r <- fx(tb) / (M*dunif(tb))
  if(U < r){
    theta[t] <- tb
    t <- t + 1
  }
  points( tb , M*U , col = "red" , pch = 16) } else {
  points( (tb) ,M*U , col = "blue" , pch = 4) }
  count <- count + 1 }
```



```
t/count
```

```
## [1] 0.4830918
```

```
mean(theta)
```

```
## [1] 0.05087764
```

```
### Beta(2,2) ###
```

```
curve(fx, from = 0-.05, to = 1+.05 )
```

```
M <- fx(.5) / ( (0.5)*(1-0.5) * (6) )
```

```
curve(M *dbeta(x, 2, 2), add=T, col="darkblue", lwd=2)
```

```
theta <- vector(length = 1000)
```

```
arr <- NULL
```

```
t <- 1
```

```
count <- 1
```

```
while(t < 100){
```

```
  tb <- rbeta(1,2,2)
```

```
  U <- runif(1)
```

```
  r <- fx(tb) / (M*dbeta(tb,2,2))
```

```
  if(U < r){
```

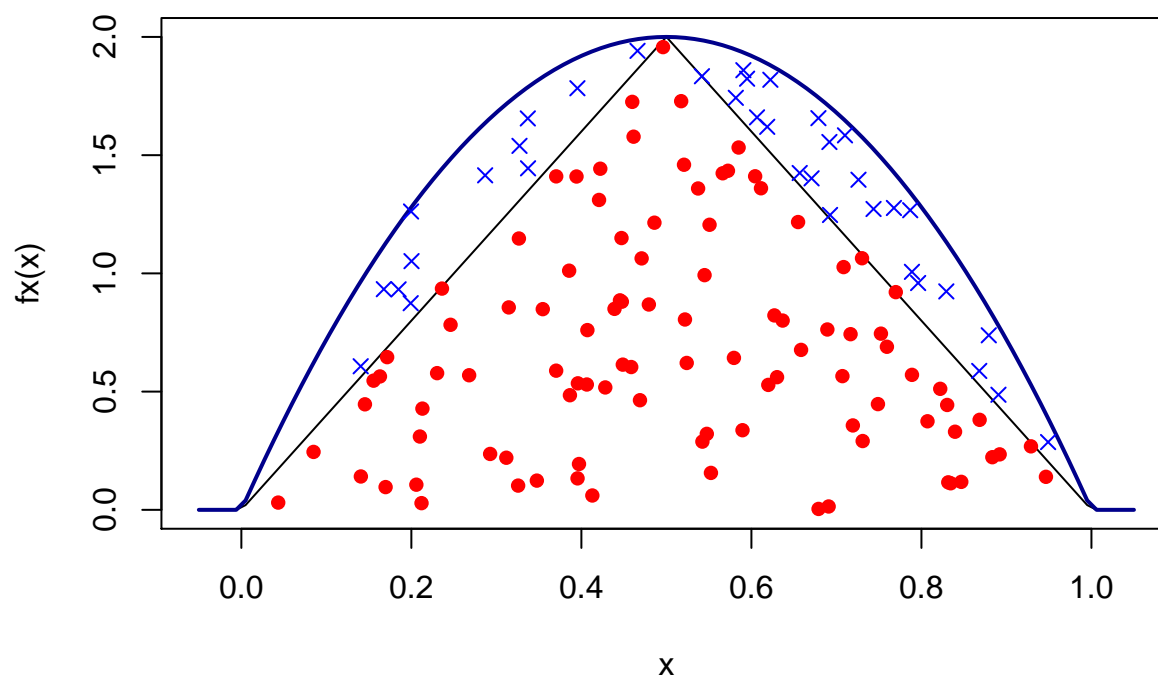
```
    theta[t] <- tb
```

```
    t <- t + 1
```

```

points( tb , M*U*dbeta(tb,2,2) , col = "red" , pch = 16) } else {
points( (tb) ,M*U*dbeta(tb,2,2) , col = "blue" , pch = 4) }
count  <- count + 1 }

```



```

t/count

```

```

## [1] 0.7352941

```

```

#plot(density(theta), lwd = 8, col = 'blue')
#curve(dbeta(x, 6, 3), add = TRUE, col = 'red', lwd = 8)
mean(theta)

```

```

## [1] 0.05058661

```