When proposal density and target density have different support

Nov 20 2019

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com (http://rmarkdown.rstudio.com).

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

1. Metropolis random walk sampler for a target distribution that is known

We are interested in sampling θ from $p(\theta|y)$. For Metropolis sampler, usually we cannot directly sample from posterior $p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} \propto p(y|\theta)p(\theta)$, which is the target distribution or the target distribution up to a multiplicative constant.

As defined in Metropolis-Hastings algorithm, the acceptance ratio for a new proposed value of θ is: $r = \frac{p(\theta^*|y)}{p(\theta^s|y)} = \frac{p(y|\theta^*)p(\theta^*)/J(\theta^*|\theta^s)}{p(y|\theta^s)p(\theta^s)/J(\theta^s|\theta^*)} = \frac{\pi(\theta^*)/J(\theta^*|\theta^s)}{\pi(\theta^s)/J(\theta^s|\theta^*)}$, where $\pi(\theta)$ is the target distribution for θ and proposal $J(\theta^*|\theta^s)$ denotes the proposal density for θ^* given θ^s.

When using a symmetric proposal density (note it should be a valid probability density function (pdf), otherwise the sampler will approximate to some other distributions), the terms $J(.)$ will cancel out, acceptance ratio is simplified to $r = \frac{\pi(\theta^*)}{\pi(\theta^s)}$

In the following example, the target distribution is $\pi(x) = expon(1) = e^{-x}, x > 0$

2. When proposal density has a different support than the target density

The support for target density is $(0, +\infty)$.

Consider a Gaussian proposal $J(x^*|x^s) = N(x^s, 1^2)$, we may use it directly as proposal to approximate $\pi(x)$, and the target distribution $expon(1)$ will assign 0 to proposed values that are negative- this is because the support for Gaussian is from $(-\infty, +\infty)$.

Since the support for the target distribution is $(0, +\infty)$, we may instead consider to restrict proposal values to $(0, +\infty)$. In this case, the proposal $X^* \sim N(x^s, 1^2), x^* > 0$ is $J'(x^*|x^s) = N(x^*; x^s, 1^2)I(x^* > 0) = C\frac{1}{\sqrt{2\pi}}exp(-\frac{1}{2}(x^* - x^s)^2)I(x^* > 0)$.

The normalizing constant C can be identified as follows: $\int_0^{+\infty} \frac{1}{\sqrt{2\pi}}exp(-\frac{1}{2}(x^* - x^s)^2)dx^* = P(X^* > 0) = P(X^* - x_s > -x^s) = P(Z > -x^s) = P(Z \le x^s) = \Phi(x^s)$ $=> C = \frac{1}{\Phi(x^s)}(\because X^* \sim N(x^s, 1^2) => Z = X^* - x^s \sim N(0, 1^2))$

Therefore, the truncated proposal $J(x^*|x^s)$ with $x^* > 0$ is $J'(x^*|x^s) = \frac{1}{\Phi(x^s)}\frac{1}{\sqrt{2\pi}}exp(-\frac{1}{2}(x^* - x^s)^2), x^* > 0$

Using this new proposal, by Metropolis-Hastings the acceptance ratio is

$$r = \frac{dexp(x^*;1)/J'(x^*|x^s)}{dexp(x^s;1)/J'(x^s|x^*)} = \frac{dexp(x^*;1)/(\frac{1}{\Phi(x^s)}N(x^*;x^s,1^2))}{dexp(x^s;1)/(\frac{1}{\Phi(x^*)}N(x^s;x^*,1^2))} = \frac{dexp(x^*;1)\times\Phi(x^s)}{dexp(x;1)\times\Phi(x^*)}$$

With this, the random walk sampler can approximate the original target distribution, $\pi(x) = expon(\lambda = 1)$

3. The distribution using truncated proposal density but using Metropolis acceptance ratio $r = \frac{\pi(x^*)}{\pi(x^s)}$

$$r = \frac{\pi(x^*)}{\pi(x^s)} = \frac{\frac{\pi(x^*)}{\frac{1}{\Phi(x^s)}N(x^*;x^s,1^2)\times\Phi(x^s)}}{\frac{\pi(x^s)}{\frac{1}{\Phi(x^*)}N(x^s;x^*,1^2)\times\Phi(x^*)}} = \frac{\pi(x^*)\Phi(x^*)/J'(x^*|x^s)}{\pi(x^s)\Phi(x^s)/J'(x^s|x^*)} = \frac{\pi(x^*)\Phi(x^*)}{\pi(x^s)\Phi(x^s)},$$

where $J'(x^*|x^s) = \frac{1}{\Phi(x^s)}\frac{1}{\sqrt{2\pi}}exp(-\frac{1}{2}(x^* - x^s)^2), x^* > 0$, $\Phi(.)$ is the cumulative distribution function of a standard normal random variable.

From the above, it can be seen that the actual target distribution that the sampler approximates is $\pi'(x) \propto \pi(x)\Phi(x)$

The normalizing constant for $\pi'(x)$ is:

$$C_2 = \int_0^{+\infty} \pi'(x)dx = \int_0^{+\infty} \pi(x)\Phi(x)dx = \int_0^{+\infty} e^{-x}\Phi(x)dx = -\int_0^{+\infty} \Phi(x)de^{-x}$$

$$= -[\Phi(x)e^{-x}|_0^{+\infty} - \int_0^{+\infty} e^{-x}\phi(x)dx] = -[(0 - \frac{1}{2}\times 1) - \int_0^{+\infty} e^{-x}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx]$$

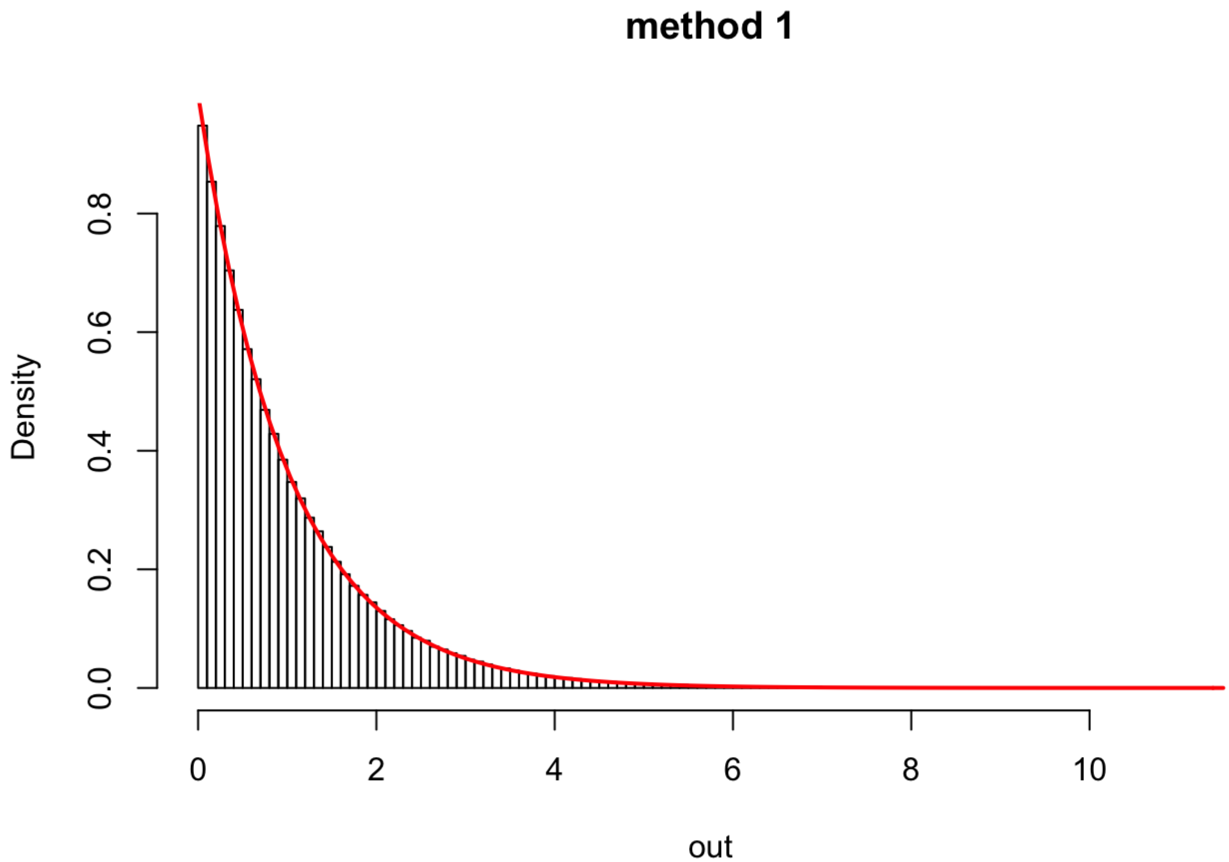
$$= \frac{1}{2} + \int_0^{+\infty} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x^2+2x+1-1)}dx$$

$$= \frac{1}{2} + e^{\frac{1}{2}} \times \int_0^{+\infty} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x+1)^2}dx$$

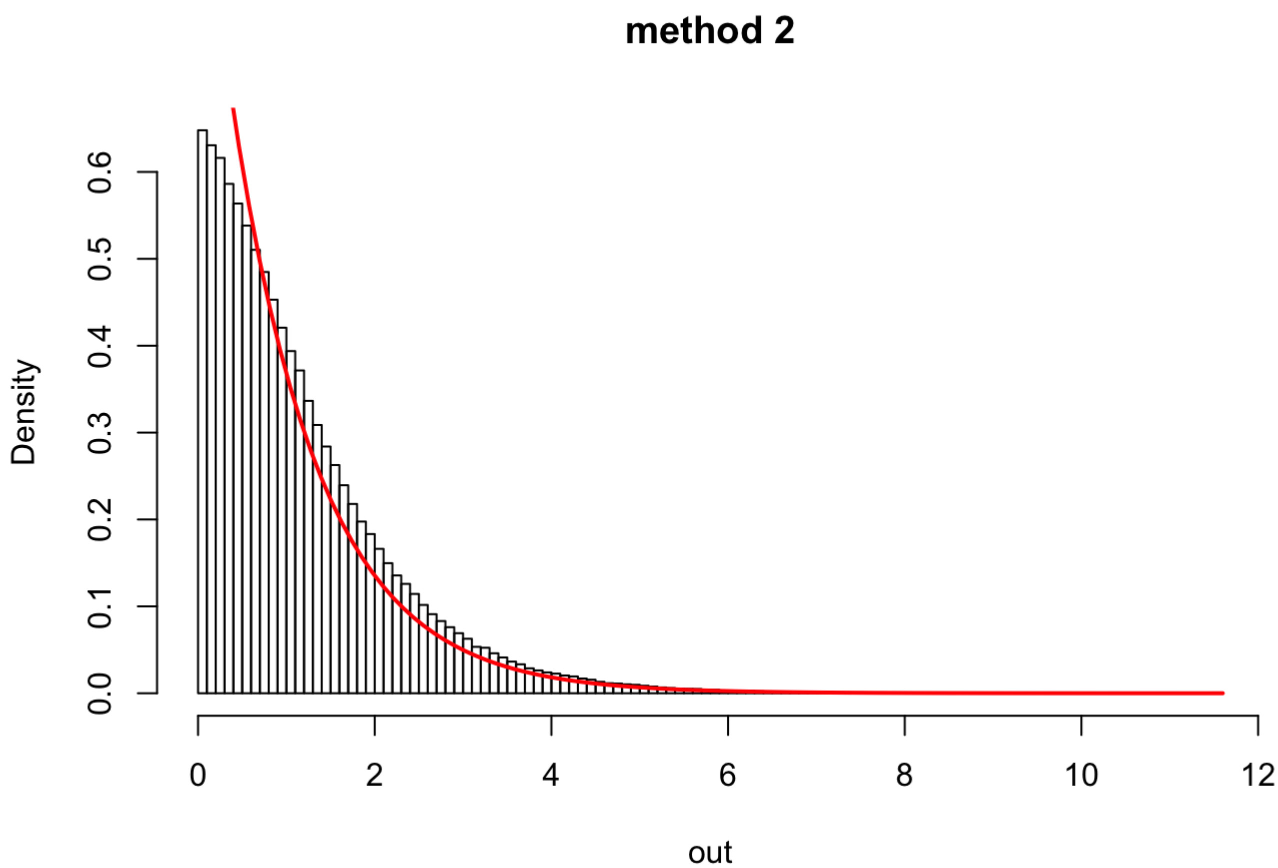
$$= \frac{1}{2} + e^{\frac{1}{2}} \times pnorm(0, mean = -1, sd = 1^2, lower.tail = F)$$

Therefore, the actual target density $\pi'(x) = \frac{1}{C_2}\pi(x)\Phi(x), x > 0$, where $\Phi(.)$ is the cdf of a standard normal distribution.

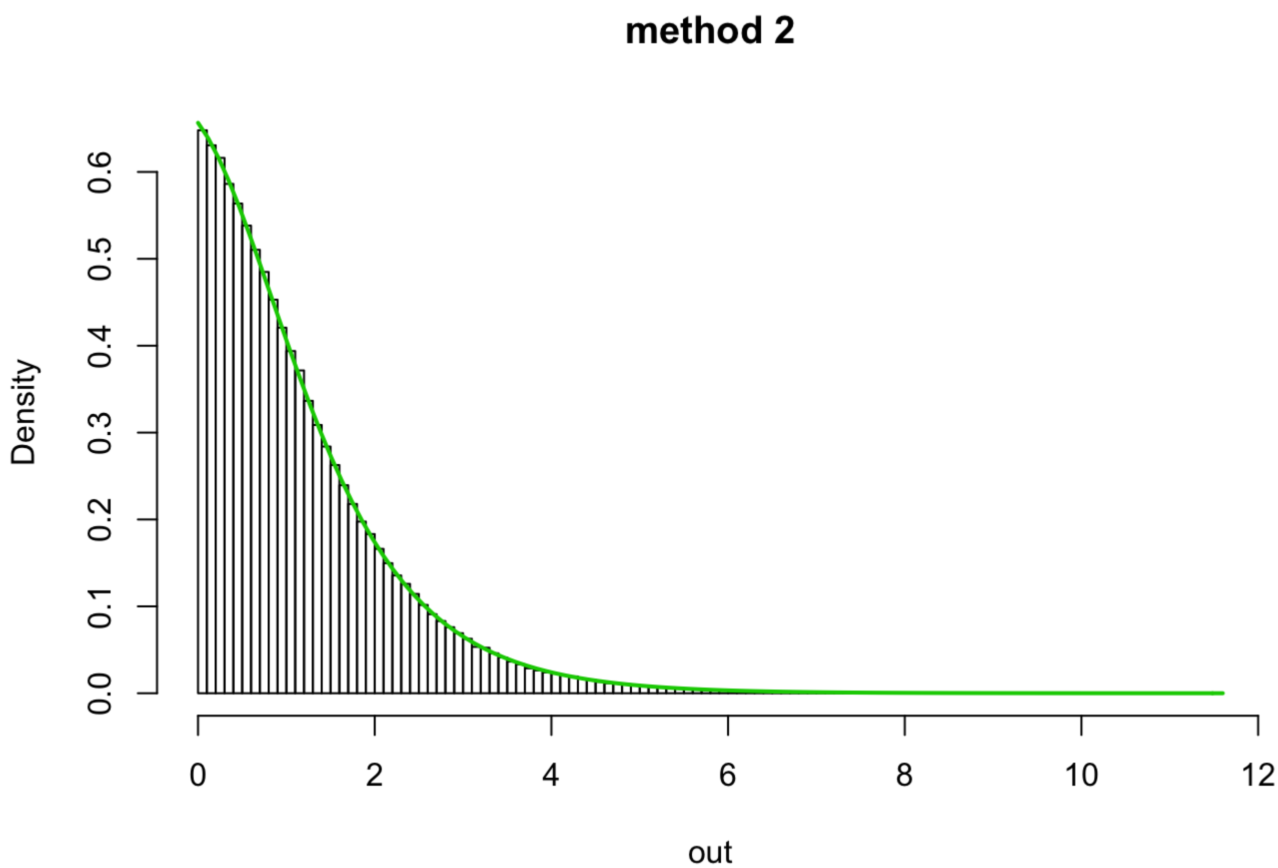
```r
S= 10^6
# the target dist. is expon(lambda= 1)
target = function(x){
    # return(ifelse(x<0, 0, exp(-x)))
    return(dexp(x= x, rate= 1))
}

method1= function(S) {
    xvec= numeric(S)
    x=1
    for (s in 1:S) {
        # note here we use a diff proposal
        xstar= x+ rnorm(n= 1, mean= 0, sd= 1)

        r= target(xstar)/ target(x)

        if (runif(n= 1, min= 0, max= 1)< r) {x = xstar}

        xvec[s]= x
    }
    return(xvec)
}

set.seed(21043)
out= method1(S)
hist(out,100,freq=FALSE,main="method 1")
curve(dexp(x, 1), add=TRUE,col=2,lwd=2) # this method works out fine
```


method 1
Density
0.0
0.2
0.4
0.6
0.8
0
2
4
6
8
10
out

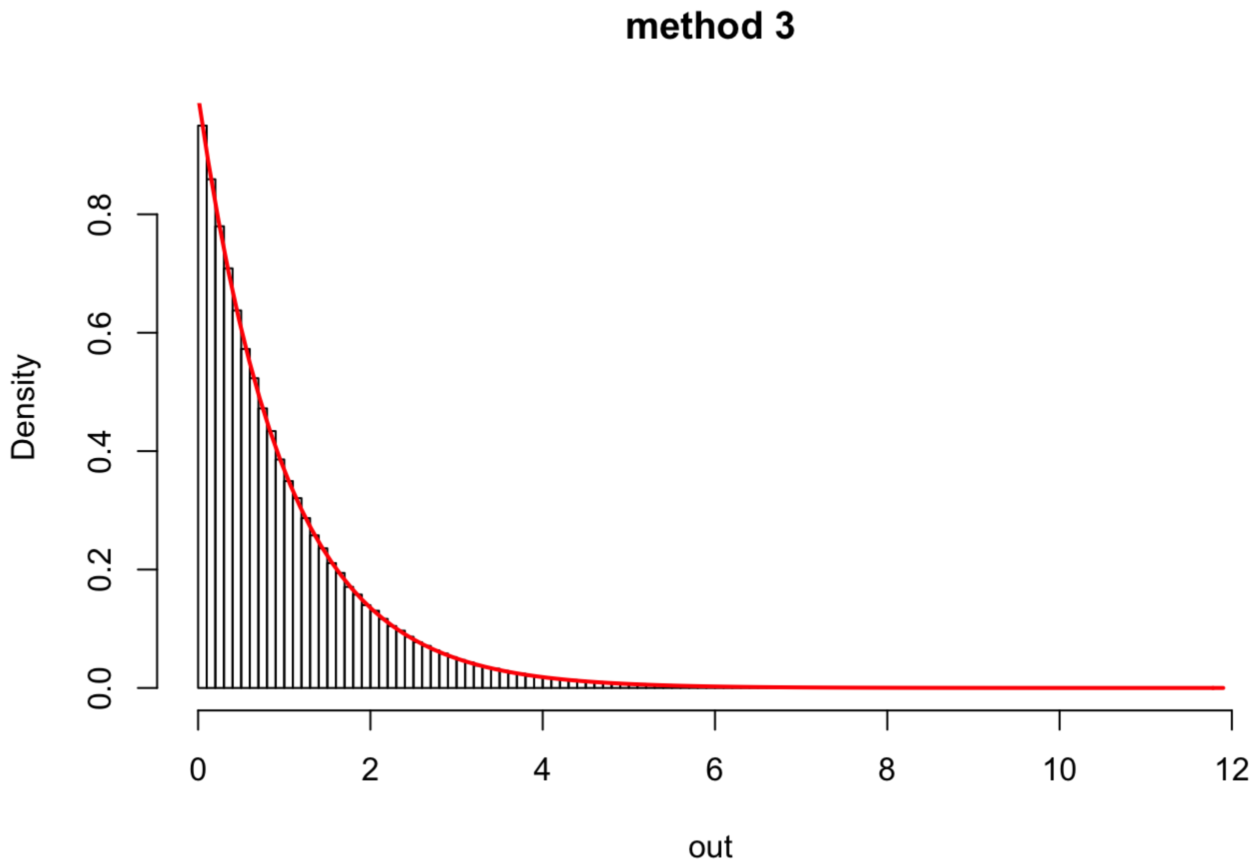
```r
method2= function(S) {
    xvec= numeric(S)
    x=1
    for (s in 1:S) {
        # note here we use a diff proposal
        # normal(xstar; mean= x, sd= 1)*I(xstar> 0)
        repeat {
            xstar= x+ rnorm(n= 1, mean= 0, sd= 1)
            if (xstar> 0)
                break
        }

        # the acceptance ratio is for another target dist.
        # the truncated proposal dist is no longer a valid density
        # which lead to a diff target dist. other than the original
        # from the sampler
        r= target(xstar)/ target(x)

        if (runif(n= 1)< r) {x = xstar}

        xvec[s]= x
    }
    return(xvec)
}

set.seed(21287)
out=method2(S)
hist(out,100,freq=FALSE,main="method 2")
curve(dexp(x, 1), add=TRUE,col=2,lwd=2)
```



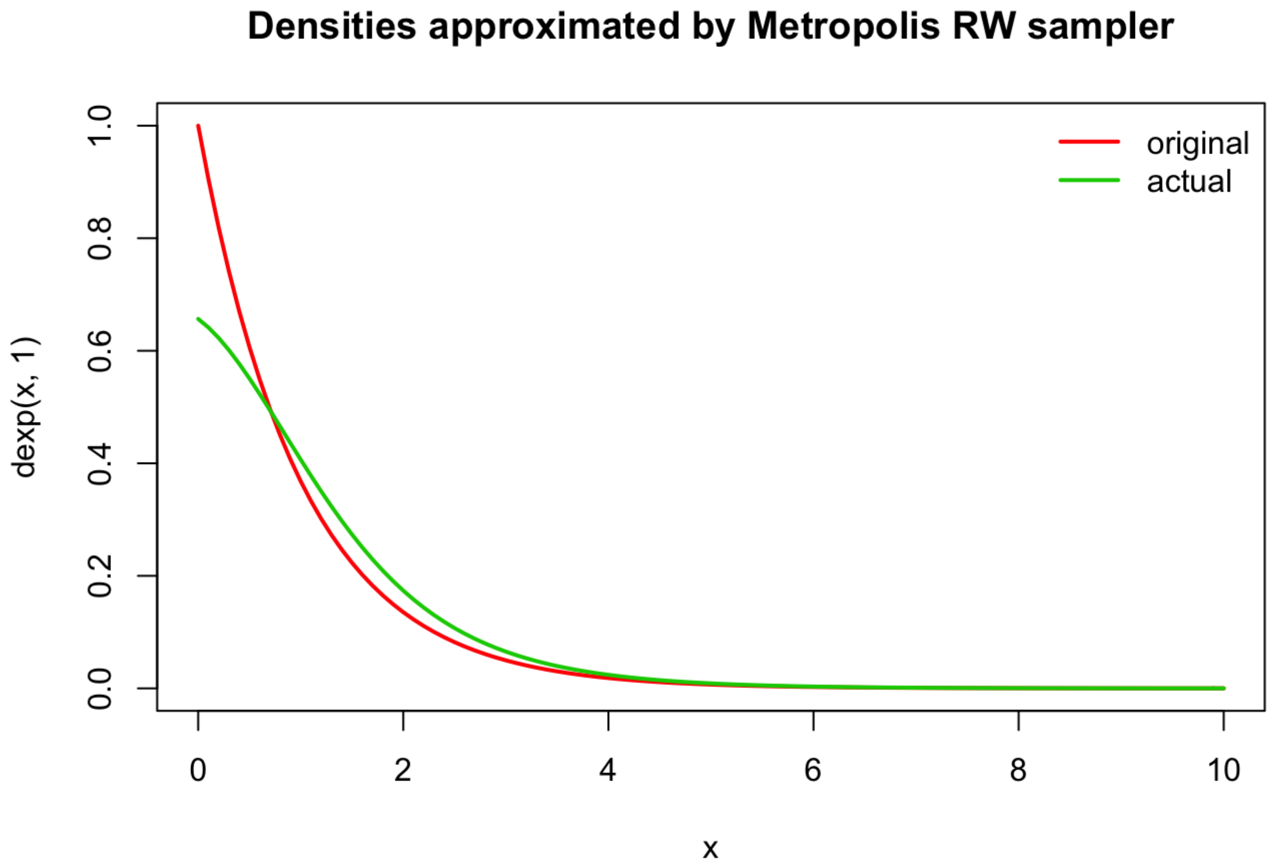

```
# result from truncated sampler is a bit off

# corrected w/ normalizing const.
constant= 1/(1/2+ exp(1/2)* pnorm(q=0, mean= -1, sd= 1, lower.tail= F))
hist(out,100,freq=FALSE,main="method 2")
curve(dexp(x, 1)* pnorm(x)* constant, add=TRUE,col=3,lwd=2) # now it overlaps well
```




```
method3= function(S) {
    xvec= numeric(S)
    x=1
    for (s in 1:S) {
        # note here we use a diff proposal
        # normal(xstar; mean= x, sd= 1)*I(xstar> 0)
        repeat {
            xstar= x+ rnorm(n= 1, mean= 0, sd= 1)
            if (xstar > 0)
                break
        }

        r= (target(xstar)* pnorm(x))/ (target(x)* pnorm(xstar))

        if (runif(n= 1)< r ) {x = xstar}

        xvec[s]= x
    }
    return(xvec)
}

set.seed(21203)
out=method3(S)
hist(out,100,freq=FALSE,main="method 3")
curve(dexp(x, 1), add=TRUE,col=2,lwd=2) # since we put correct density func. for truncat
ed proposal, the approx. target dist. is the original expon(1)
```




```
set.seed(21207)
curve(dexp(x, 1), 0, 10, col=2, lwd=2,
      main= "Densities approximated by Metropolis RW sampler")
curve(dexp(x, 1)* pnorm(x)* constant, add= TRUE,col=3,lwd=2)
legend("topright", legend= c("original", "actual"),
       col= c(2,3), lty= c(1,1), lwd= c(2,2), bty= 'n')
```



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.