HM1

Pol Ribó León;1840853

13 de mayo de 2019

a) Show how it is possible to simulate from a standard Normal distribution using pseudo-random deviates from a standard Cauchy and the A-R algorithm

The esssence of the Accept_reject Algorithm is based on choosing a easy-to-sample distribution **g(x)** and find a coefficient **k** such that “envelopes” the target distribution **f(x)**. Then, sample from g(x) and for each draw, **xi**, also sample a **u** from a standard uniform distribution(U(u|0,1).

The sample **xi** is accepted if it is $kg(xi)u \leq f(xi)$, or rejected otherwise.

Having stated this, in this case our target distribution is a Normal(0,1), and a standard Cauchy is used.

- $X \sim N(0,1)|fx = \frac{1}{sqrt(2\pi)}e^{-1/2y^2}$
- $Y \sim Cauchy(1,0)|fy = \frac{1}{\pi(1+x^2)}$

Now, as we want to envelope the Normal distribution, we need to bound **fx** by **kfy** where **k>=1**. To do this, we know that the optimal **k** is the minnimum maximmum of **fx/fy**.

After, we form a new random variable,**E**, where $E|y \sim Bernoulli(fx(y)/kfy(y))$. With this, we are able to represent the algorithm accepting a draw from **Y**, which takes 1 with a determined acceptance probability and 0 otherwise.

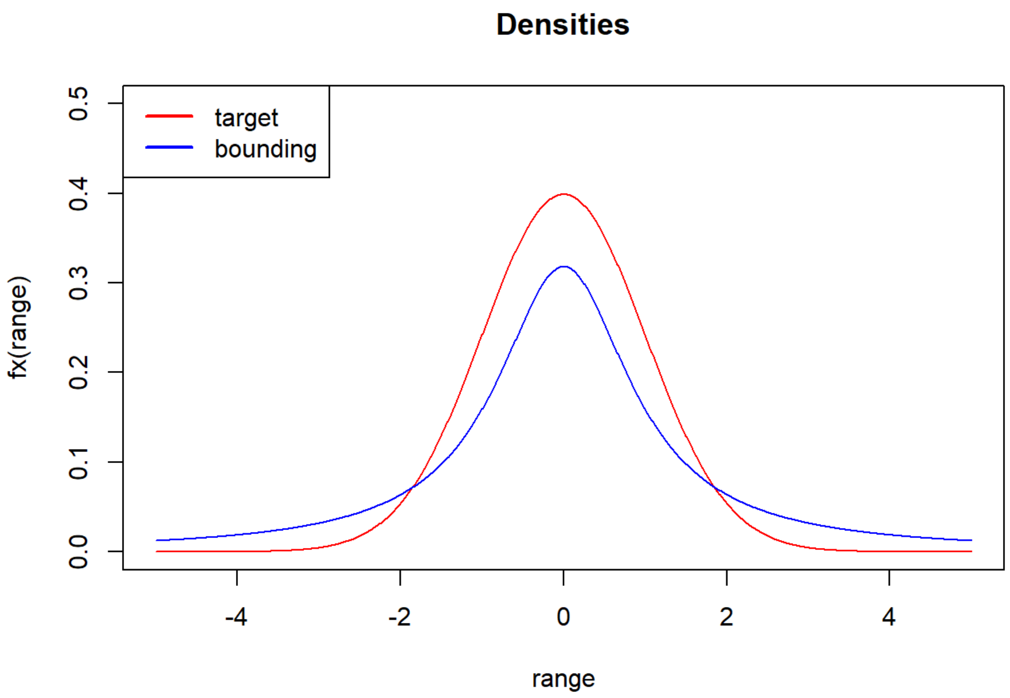
So, the A-R algorithm will take an average of iterations to obtain a sample. All the draws accepted are collected into a the random variable X=Y|E=1.

b) R code provided

```r
#normal standard dist
fx=function(x){
  1/(sqrt(pi*2))*exp((-1/2)*x^2)
}
range=seq(-5,5,by=0.01)

#cauchy standard dist
fy=function(x){
  1/(3.141593*(1+x^2))
}
#normal
plot(range,fx(range), type='l', col="red", ylim=c(0,.5))
#cauchy
lines(range, fy(range), col='blue')
```

```
legend(x="topleft",lty=1,lwd=2.4,col=c("red","blue"),legend=c("target","bounding"))
title(main="Densities")
```

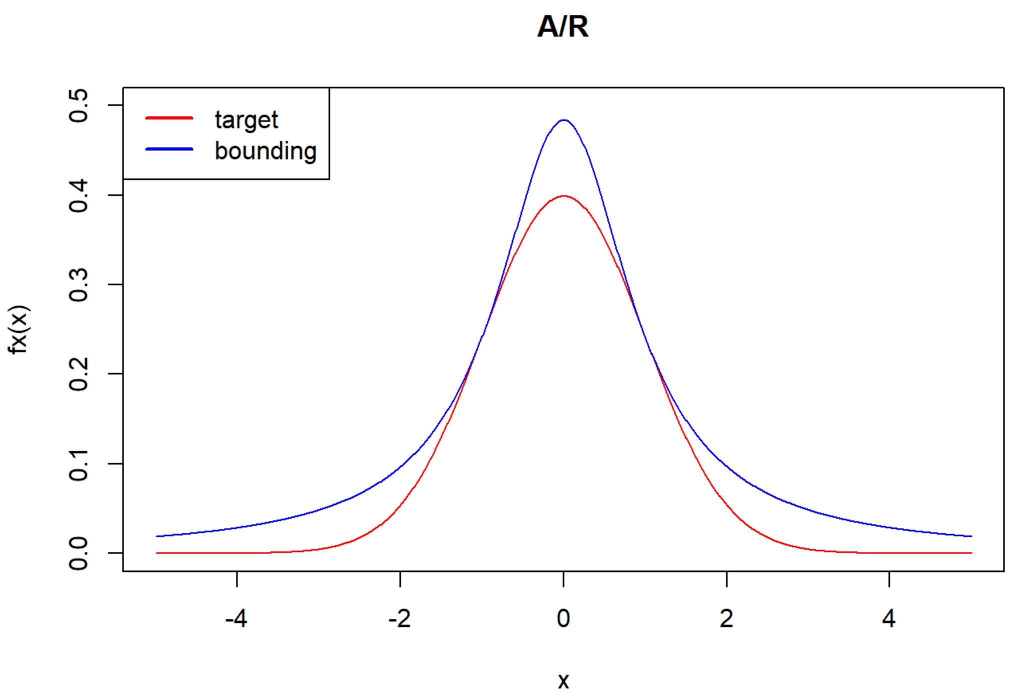



```
# optimal k_star(1)
girs=seq(0,1,length=100000)
k1=max(fx(girs)/fy(girs))
k1
```

```
## [1] 1.520347
```

```
plot(range,fx(range), type='l', col='red',ylim = c(0,.5), xlab = 'x', ylab = 'fx(x)')
lines(range,k1*fy(range),col='blue')
text(0.8,3.5,labels=expression(k~f[U](x)))
text(0.8,0.7,labels=expression(f[X](x)),col="red")

legend(x="topleft",lty=1,lwd=2.4,col=c("red","blue"),legend=c("target","bounding"))
title(main="A/R")
```




```
##SIMULATION

ef=function(x){
  fx(x)
}

q=function(x){
  fy(x)
}

k=2

n_sim_aux=10000

Y=rep(NA,n_sim_aux)
E=rep(NA,n_sim_aux)
for(i in 1:n_sim_aux){
  Y[i]=rcauchy(1)
  E[i]=rbinom(1,size=1,prob=ef(Y[i])/(k1*q(Y[i])))
```

```
}

X <- Y
X[E==0] <- NA

# Accepted Y[i]'s
X=Y[E==1]

sum(E)
```

```
## [1] 6629
```

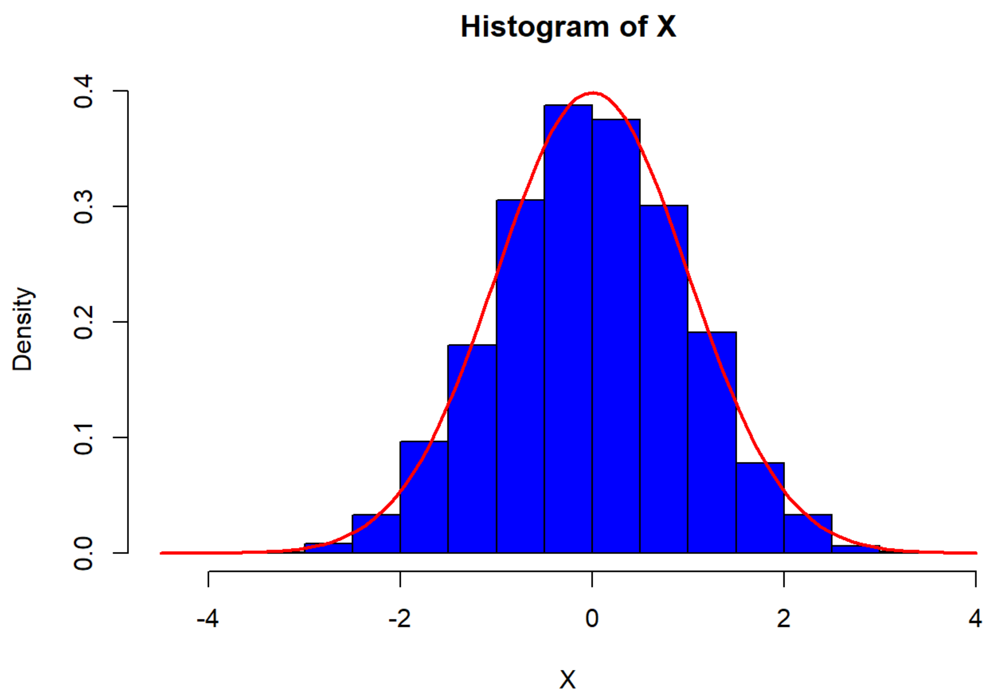
```
length(X)
```

```
## [1] 6629
```

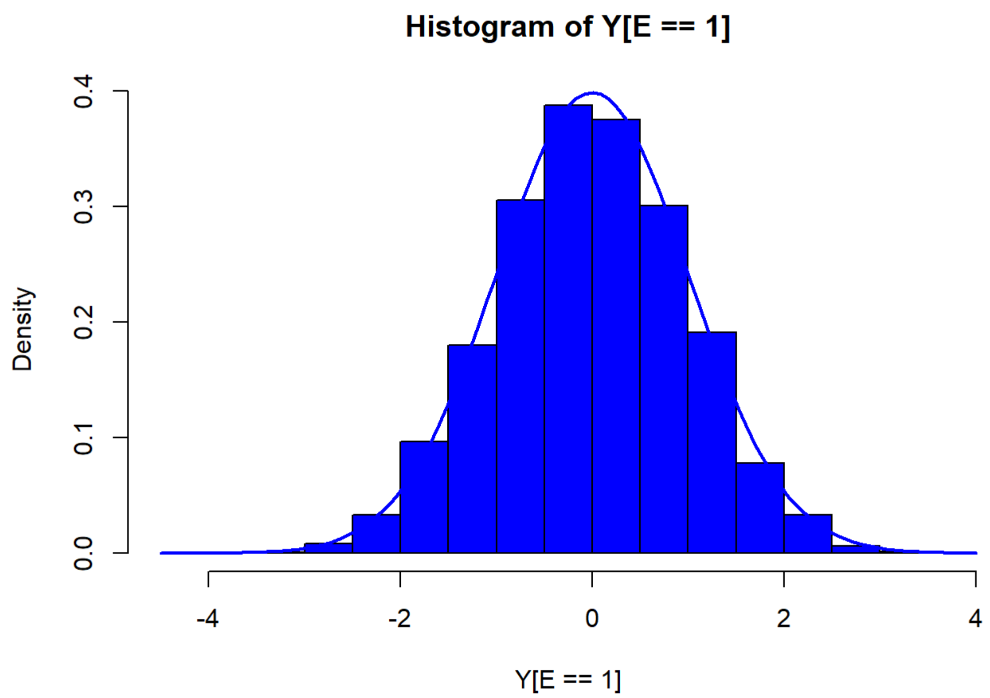
```
mean(E)
```

```
## [1] 0.6629
```

```
#distribution of accepted Y[i]'s
hist(X,prob=TRUE, col='blue')
curve(ef(x),add=TRUE,col="red",lwd=2)
```




```
#distribution of accepted Y[i]'s
hist(Y[E==1],prob=TRUE, col='blue')
curve(fx(x),col="blue",lwd=2,add=TRUE)
```



c) evaluate numerically (approximately by MC) the acceptance probability

As shown in the Rcode, in this precise example with the Normal distribution and the Cauchy distribution, the ratio $\frac{fx(x)}{cfy(y)}$ was maximized by iterating, and so the value of **c** for the setup is **k1=1.520346** which implies an acceptance probability of about **0.6577446(1/k1)**. Also, applying Monte Carlo simulation, where we also got **0.6578605**

```
#c)evaluate numerically (approximately by MC)
#the acceptance probability

acceptance_prob=1/k1
acceptance_prob
```

```
## [1] 0.6577446
```

```
#By MC
accept_prob=c()
iter=1000
```

```
for(i in 1:iter){
  sim_data=rcauchy(1)
  p=ef(sim_data)/(k1*q(sim_data))
  accept_prob=c(accept_prob,p)
}
mean(accept_prob)
```

```
## [1] 0.6546529
```

d) write your theoretical explanation about how you have conceived your Monte Carlo estimate of the acceptance probability

One can find the acceptance probability by 1/c. Below the theory.

$$P(Xaccepted) = P(E \leq \frac{f(x)}{kg(x)}) = \int P(E \leq \frac{f(x)}{kg(x)}|X = x)g(x)dx = \int \frac{f(x)}{kg(x)}g(x)dx = \frac{1}{k}$$

If we take this into account , and also that the random variable E $\sim$ Ber(p), then by the strong law of Large Numbers that states that the sample mean converges almost surely to the population mean; we can derive;

$$\frac{\sum_{i=1}^{n}}{\lim_{x\to\infty} n} \to p = \frac{1}{k}$$

e) save the rejected simulations and provide a graphical representation of the empirical distribution (histogram or density estimation)

Rcode provided

```
ef=function(x){
  fx(x)
}

q=function(x){
  fy(x)
}

n_sim_aux=10000

Y=rep(NA,n_sim_aux)
E=rep(NA,n_sim_aux)
for(i in 1:n_sim_aux){
  Y[i]=rcauchy(1)
  E[i]=rbinom(1,size=1,prob=ef(Y[i])/(k1*q(Y[i])))
}

X <- Y

X[E==1] <- NA
```

```
#Rejected Y[i]'s
X=Y[E==0]

sum(E)
```

```
## [1] 6585
```

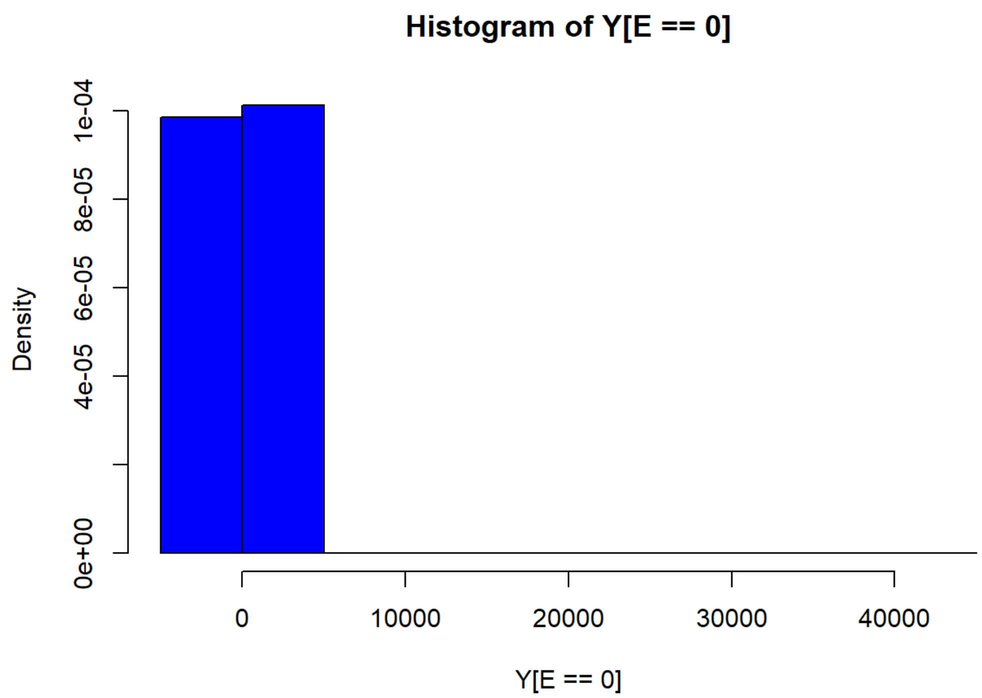
```
length(X)
```

```
## [1] 3415
```

```
mean(E)
```

```
## [1] 0.6585
```

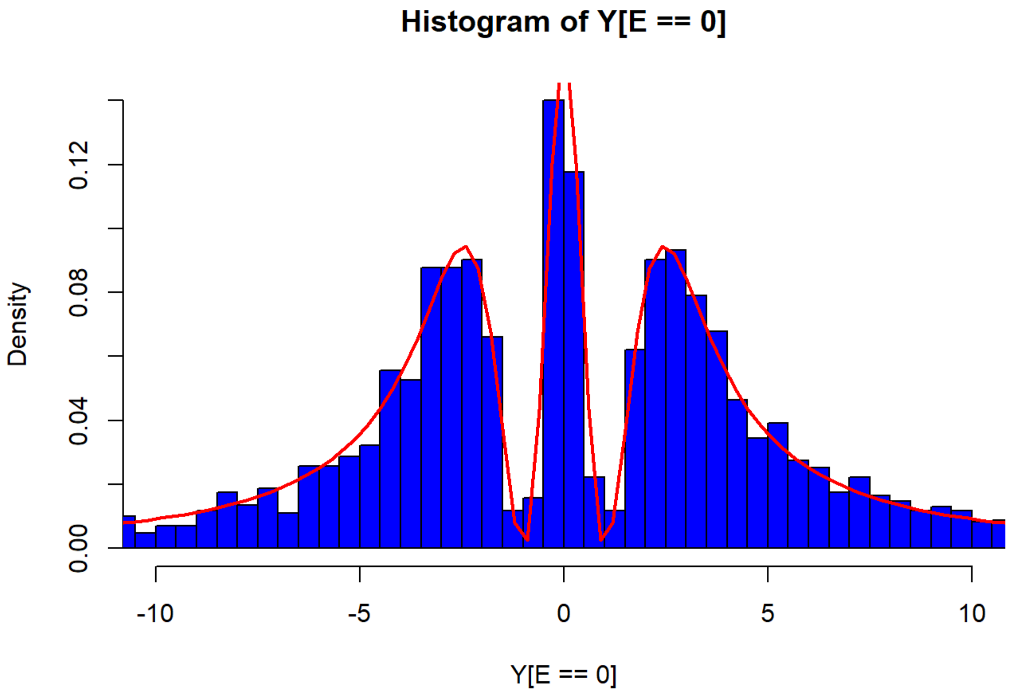
```
#Histogram of Y rejected
hist(Y[E==0],prob=TRUE, col='blue')
```




```r
#Histogram of Y rejected
hist(Y[E==0],prob=TRUE, breaks=100000, col='blue', xlim=c(-10,10))

rej_dist=function(x){
  (fy(x)*k1-dnorm(x))/(k1-1)
}

curve(rej_dist, add=T, lwd=2, xlim=c(-15,15), col='red')
```



f) As we are looking for the underlying density of the rejected random variables, this can be represented as:

$P(Y \leq x|E == 0) = \frac{P(Y \leq x, E=0)}{P(E=0)}$

So; $P(E = 0|Y = y) = 1 - P(E = 1|Y = y)$

FINDING THE DENSITY

NOTE:fy(y) is the standard cauchy dist as in R code provided

Solving the numerator

$P(Y \leq x, E = 0) = P(\in [0, x], E = 0) = \int_0^x P(E = 0|Y = y)fy(y)dy$

$\int_0^x [1 - P(E = 1)|Y = y)]fy(y)dy = \int_0^x fy(y)dy - \int_0^x P(E = 1|Y = y)fy(y)dy = \int_0^x fy(y)dy - \int_0^x \frac{f(y)}{kfy(y)} fy(y)dy = \int_0^x fy(y)dy - \frac{1}{k} \int_0^x fy(y)dy$

Solving the denominator $P(E = 0) = 1 - P(E = 1) = 1 - \frac{1}{k} = \frac{k-1}{k}$

Underlying Density

$$P(Y \leq x|E =)\frac{kG(y)-\frac{1}{k}F(y)}{1-k} \rightarrow \frac{1}{k-1}(kG(y) - F(y))$$

So, the empirical distribution corresponds to the subtraction between the bounding distribution(standard Cauchy) by the optimal k and the target distribution (standard Normal) up to $\frac{1}{k-1}$