

Teoría de probabilidad

Martín Andrés Macías Quintero

Ejercicio 12.7

Sean (X, Y) uniformes sobre la bola unitaria, es decir,

$$f_{(X,Y)}(x, y) = \begin{cases} \frac{1}{\pi} & \text{si } x^2 + y^2 \leq 1 \\ 0 & \text{si } x^2 + y^2 > 1 \end{cases}$$

Encuentre la distribución de $R = \sqrt{X^2 + Y^2}$. (*Pista:* Introduzca una variable aleatoria auxiliar definida como $S = \text{Arctan}\left(\frac{Y}{X}\right)$.) [Respuesta: $f_R(r) = 2r\mathbb{K}_{(0,1)}(r)$.]

Demostración:

Se introduce la variable aleatoria auxiliar $S = \tan^{-1}\left(\frac{Y}{X}\right)$. Sea $g(x, y) = \left(\sqrt{x^2 + y^2}, \tan^{-1}\left(\frac{y}{x}\right)\right)$. Si se escribe $r = \sqrt{x^2 + y^2}$; $0 \leq r \leq 1$ y $\tan^{-1}\left(\frac{y}{x}\right)$, las inversas serían $g_1^{-1}(r, \theta) = (r \cos \theta, r \sin \theta)$ y $g_2^{-1} = (-r \cos \theta, -r \sin \theta)$ puesto que g no es inyectiva. Sus correspondientes jacobianos serían:

$$\left|J_{g_1^{-1}}\right| = \left|\det \begin{bmatrix} \cos \theta & \sin \theta \\ -r \sin \theta & r \cos \theta \end{bmatrix}\right| = |r| = r \quad (1)$$

$$\left|J_{g_2^{-1}}\right| = \left|\det \begin{bmatrix} -\cos \theta & -\sin \theta \\ r \sin \theta & -r \cos \theta \end{bmatrix}\right| = |r| = r \quad (2)$$

De esta forma, se tiene que:

$$\begin{aligned} f_{Z,W}(z, w) &= r\{f_{X,Y}(r \cos \theta, r \sin \theta) + f_{X,Y}(-r \cos \theta, -r \sin \theta)\}\mathbb{K}_{(0,1)}(r)\mathbb{K}_{\left(\frac{-\pi}{2}, \frac{\pi}{2}\right)}(\theta) \\ &= r\mathbb{K}_{(0,1)}(r) \left(\frac{1}{\pi} + \frac{1}{\pi}\right) \mathbb{K}_{\left(\frac{-\pi}{2}, \frac{\pi}{2}\right)}(\theta) \\ &= \frac{2r\mathbb{K}_{(0,1)}(r)\mathbb{K}_{\left(\frac{-\pi}{2}, \frac{\pi}{2}\right)}(\theta)}{\pi} \end{aligned} \quad (3)$$

La densidad marginal sería:

$$f_R(r) = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{2r\mathbb{K}_{(0,1)}(r)}{\pi} d\theta = 2r\mathbb{K}_{(0,1)}(r) \quad (4)$$

Simulación:

Sean X y Y dos variables aleatorias con distribución $U(0, 1)$

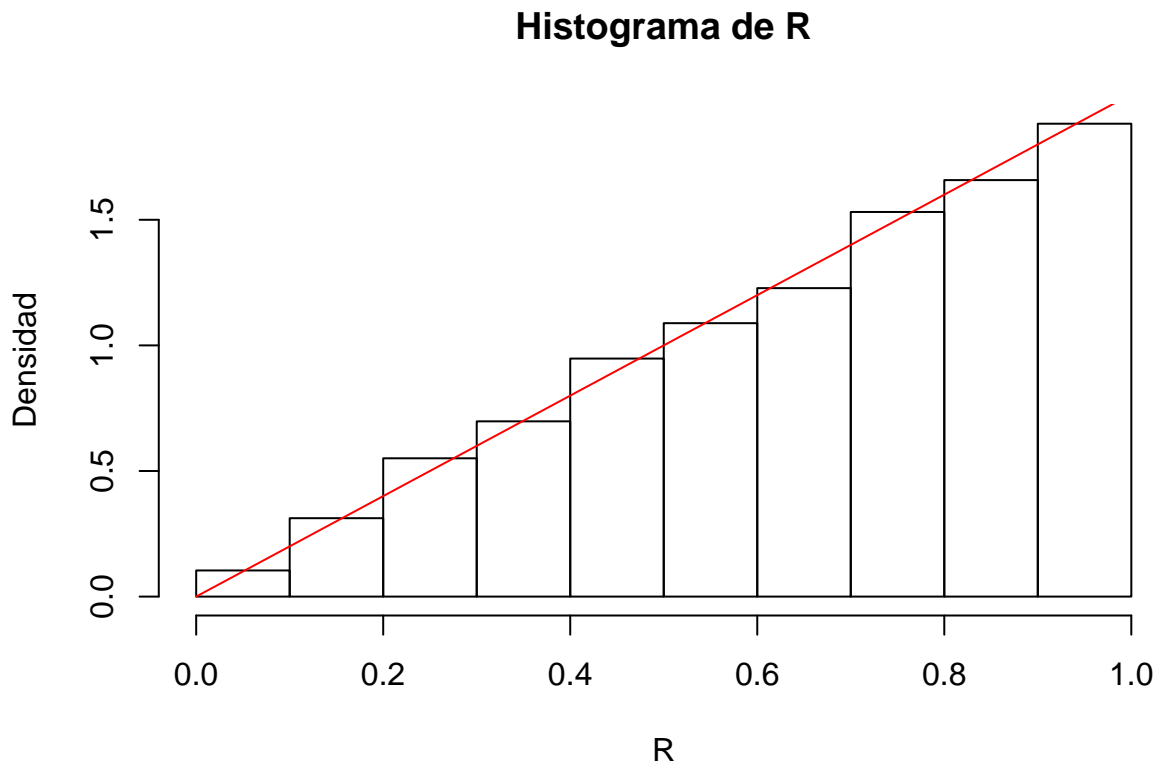
```

set.seed(163053)
X<-runif(10000)
Y<-runif(10000)

R<-sqrt(X^2+Y^2)
R<-R[R<=1]

hist(R,freq = F, main = "Histograma de R", xlab="R", ylab = "Densidad")
curve(2*x,from = 0,to = 1,col=2,add=T)

```



Gracias a la simulación, es posible que los datos de R simulados a partir de dos variables uniformes, se ajustan a la distribución teórica.

Ejercicio 12.11

Sean (X, Y) normales independientes, ambas con media $\mu = 0$ y varianza σ^2 . Sea

$$Z = \sqrt{X^2 + Y^2} \quad \text{y} \quad W = \text{Arctan}\left(\frac{Y}{X}\right), \quad -\frac{\pi}{2} < W \leq \frac{\pi}{2}.$$

Demuestre que Z tiene distribución Rayleigh, que W es uniforme sobre $(-\frac{\pi}{2}, \frac{\pi}{2})$ y que Z y W son independientes.

Demostración:

Sea $g(x, y) = \left(\sqrt{x^2 + y^2}, \arctan\left(\frac{y}{x}\right) \right)$

En este caso las inversas serían $g_1^{-1}(z, w) = (z \sin w, z \cos w)$ y $g_2^{-1}(z, w) = (z \sin w, -z \cos w)$ puesto que g no es inyectiva. Sus correspondientes jacobianos serían:

$$\left| J_{g_1^{-1}} \right| = \left| \det \begin{bmatrix} \sin w & z \cos w \\ \cos w & -z \sin w \end{bmatrix} \right| = |z| = z \quad (5)$$

$$\left| J_{g_2^{-1}} \right| = \left| \det \begin{bmatrix} \sin w & z \cos w \\ -\cos w & z \sin w \end{bmatrix} \right| = |z| = z \quad (6)$$

De esta forma, se tiene que:

$$\begin{aligned} f_{Z,W}(z, w) &= \left(\frac{1}{2\pi\sigma^2} e^{-\frac{z^2}{2\sigma^2}} z + \frac{1}{2\pi\sigma^2} e^{-\frac{z^2}{2\sigma^2}} z \right) \mathbb{K}_{(-\frac{\pi}{2}, \frac{\pi}{2})}(w) \mathbb{K}_{(0, \infty)}(z) \\ &= \frac{1}{\pi} \mathbb{K}_{(-\frac{\pi}{2}, \frac{\pi}{2})}(w) \cdot \frac{z}{\sigma^2} e^{-\frac{z^2}{2\sigma^2}} \mathbb{K}_{(0, \infty)}(z) \end{aligned} \quad (7)$$

Simulación:

Sean X y Y dos variables aleatorias independientes con distribución normal con media $\mu = 0$ y varianza $\sigma^2 = 4$

```
library(VGAM)

set.seed(163053)

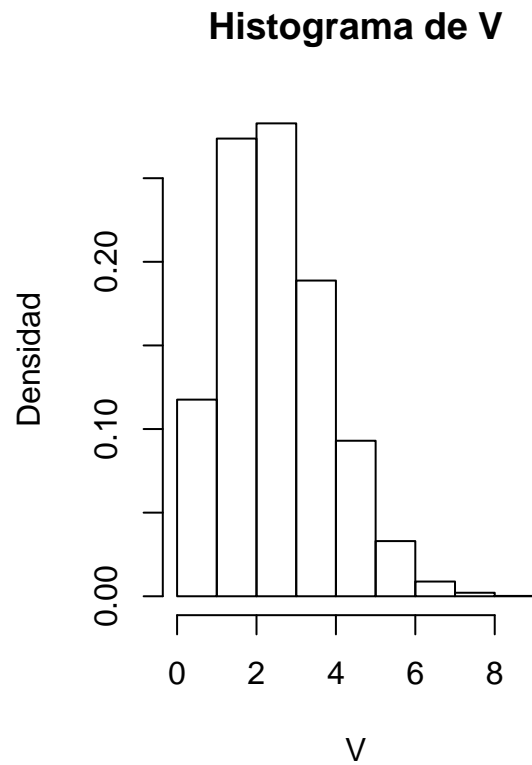
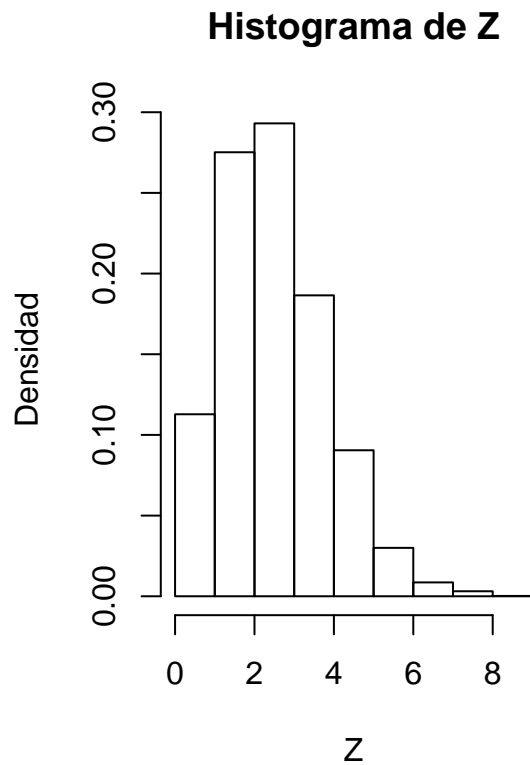
X<-rnorm(10000,0,2)
Y<-rnorm(10000,0,2)
Z<-sqrt(X^2+Y^2)
```

A continuación se simulan datos de una distribución Rayleigh con $\sigma = 2$:

```
V<-rrayleigh(10000,2)
```

Al comparar los dos histogramas de los datos simulados a partir de X y Y con los de Z :

```
par(mfrow=c(1,2))
hist(Z,freq = F,breaks = 10, main = "Histograma de Z", xlab="Z", ylab = "Densidad")
hist(V,freq = F,breaks = 10, main = "Histograma de V", xlab="V", ylab = "Densidad")
```

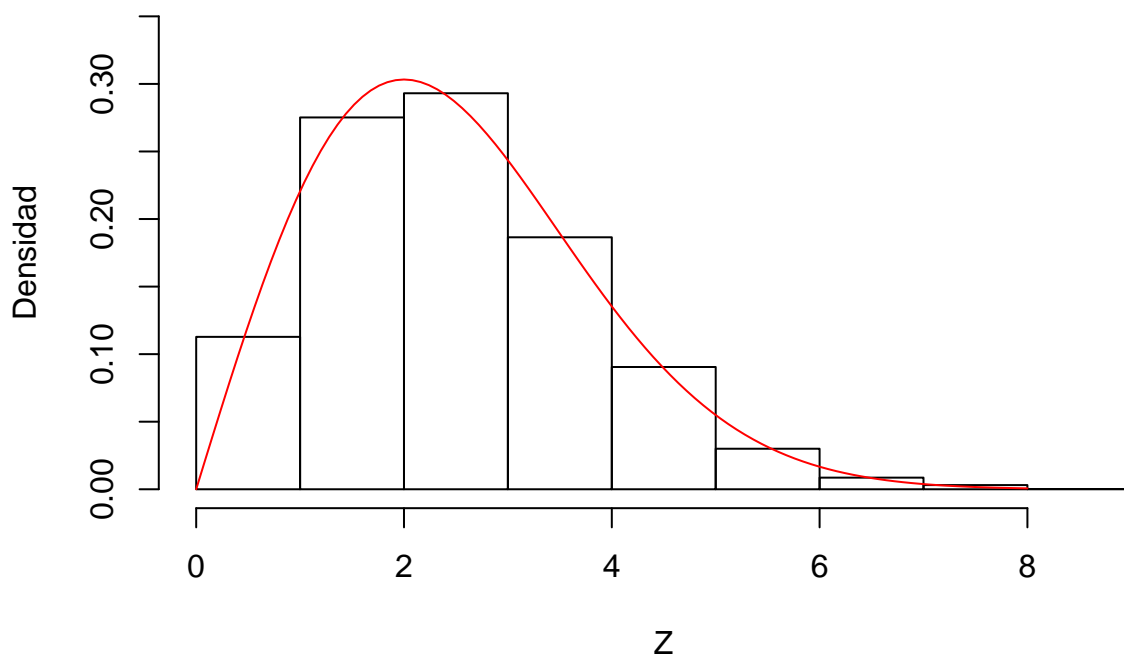


Puede verse que los dos histogramas son similares, es decir, que tanto los datos simulados (Z) a partir de dos normales como los obtenidos (V) son equivalentes. Si se comparan con la función de densidad teórica:

```
rayleigh<-function(x,y){
  (x/y^2)*exp((-x^2)/(2*y^2))
}

par(mfrow=c(1,1))
hist(Z,freq = F,breaks = 10,ylim = c(0,0.35), main = "Histograma de Z", xlab="Z", ylab = "Densidad")
curve(rayleigh(x,2),from = 0,to = 8,add = T,col=2)
```

Histograma de Z



De esta forma, es posible confirmar que los datos de la distribución Rayleigh simulados a partir de dos normales coinciden con la función de densidad de una distribución Rayleigh.

Ahora, sean X y Y dos variables normales independientes con media $\mu = 0$ y varianza $\sigma^2 = 4$

```
set.seed(163053)
```

```
X<-rnorm(10000,0,2)
```

```
Y<-rnorm(10000,0,2)
```

```
U<-X/Y
```

```
W<-atan(U)
```

Al simular datos Z de una distribución uniforme $U(-\pi/2, \pi/2)$, se tiene:.

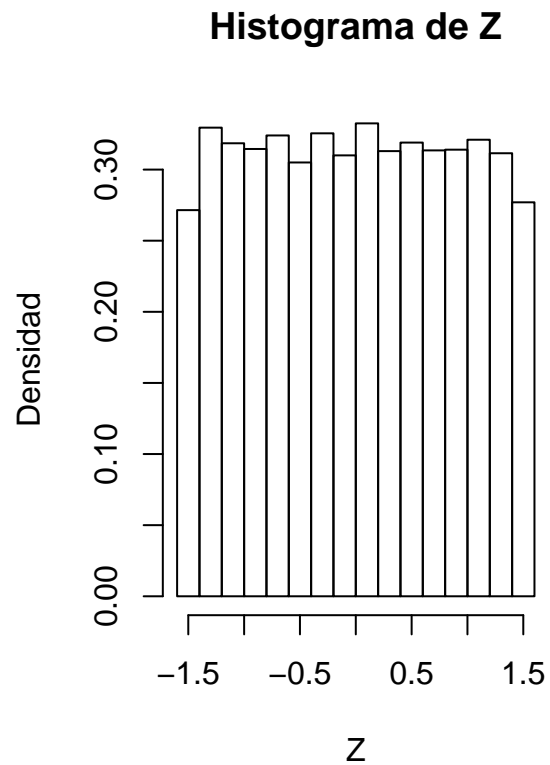
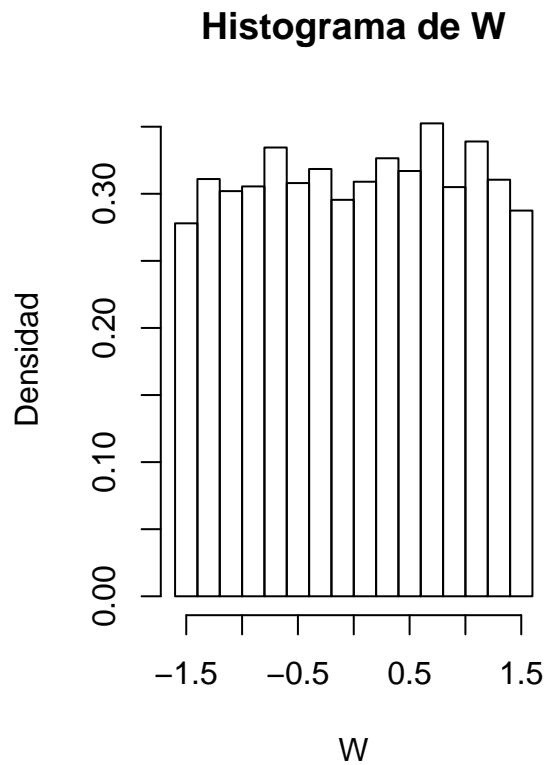
```
Z<-runif(10000,-pi/2,pi/2)
```

Al comparar los dos histogramas de los datos simulados a partir X y Y y los de Z :

```
par(mfrow=c(1,2))
```

```
hist(W,freq = FALSE, main = "Histograma de W", xlab="W", ylab = "Densidad")
```

```
hist(Z,freq = FALSE, main = "Histograma de Z", xlab="Z", ylab = "Densidad")
```

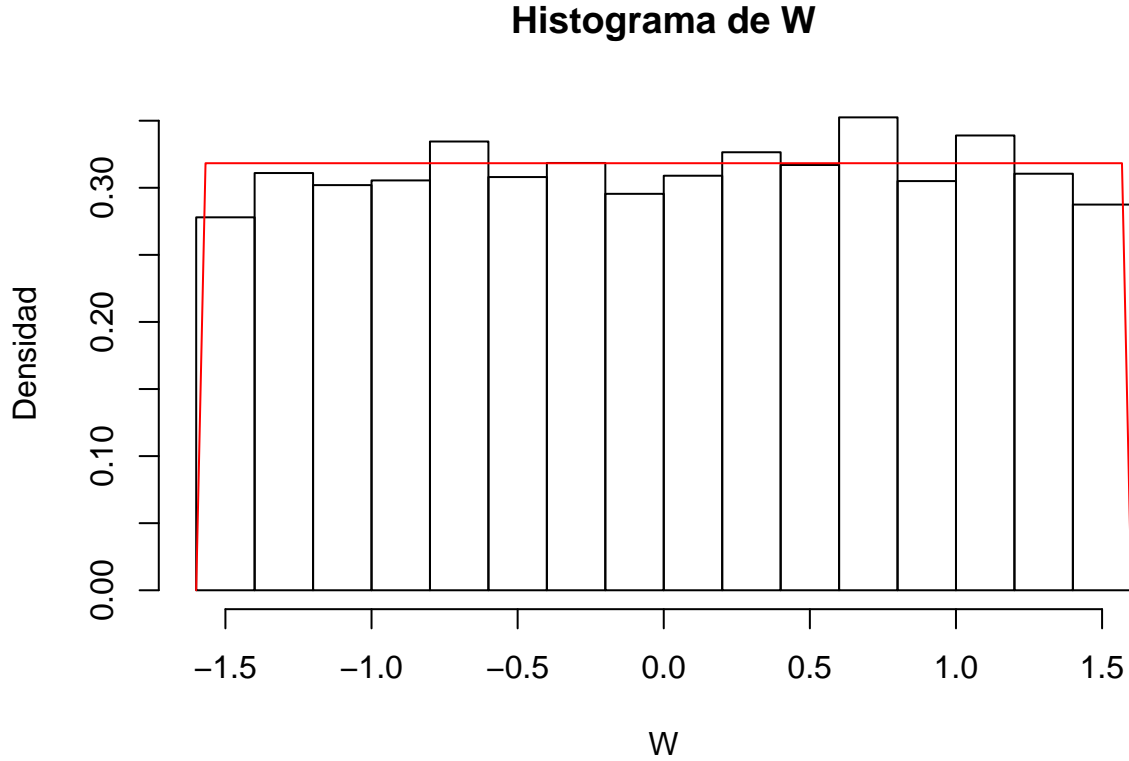


Se observa que los dos histogramas son similares, es decir, que tanto los simulados W a partir de dos normales y los obtenidos (Z) son equivalentes. Al comparar con la función de densidad teórica:

```
unifor<-function(x,y,z){
  if(x>=y&x<=z){1/(z-y)}
  else{0}
}

unif<-Vectorize(unifor)

par(mfrow=c(1,1))
hist(W,freq = FALSE, main = "Histograma de W", xlab="W", ylab = "Densidad")
curve(unif(x,-pi/2,pi/2),add = T,col=2)
```



Se confirma que los datos de la distribución uniforme entre $-\pi/2$ y $\pi/2$ simulados a partir de dos normales coinciden con la función de densidad de la una distribución uniforme.

Ejercicio 12.15

(Simulación de variables aleatorias normales) Sean U_1, U_2 dos variables aleatorias uniformes independientes sobre $(0, 1)$. Sean $\theta = 2\pi U_1$ y $S = -\ln(U_2)$.

(Pista: Para la parte (a) recuerde que una exponencial es un caso especial de una distribución Gamma: de hecho, esto es χ_2^2 . Para la parte (b) invierta el procedimiento del ejercicio 12.11).

Observación: El ejercicio 12.15 es conocido como el método Box-Muller para simular variables aleatorias normales.

- a. Demuestre que S tiene distribución exponencial, y que $R = \sqrt{2S}$ tiene distribución Rayleigh.

Demostración:

$$S = -\ln(V) \Rightarrow S = h(v) = -\ln(V) \Rightarrow v = e^{-s} \Rightarrow v = h^{-1}(s) = e^{-s} \left| \frac{dh^{-1}(s)}{ds} \right| = |-e^{-s}| = e^{-s} \quad (8)$$

Como

$$f_V(v) = 1 \quad f_S(s) = f_V(v) \cdot \left| \frac{dv}{ds} \right| = 1 \cdot e^{-s} = e^{-s} \quad (9)$$

De tal suerte que puede concluirse que $S \sim \text{Exp}(1)$

Ahora,

$$R = \sqrt{2S} = \sqrt{-2\ln(V)} R = h(v) = \sqrt{-2\ln(V)} \quad (10)$$

Luego,

$$R = \sqrt{-2\ln(V)} R^2 = -2\ln(V) - \frac{R^2}{2} = \ln(V)V = e^{-\frac{R^2}{2}} \quad (11)$$

Así,

$$v = h^{-1}(r) = e^{-\frac{r^2}{2}}$$

De esta forma,

$$\frac{dv}{dr} = \frac{dh^{-1}(r)}{dr} = -e^{-\frac{r^2}{2}} \cdot r = r \cdot e^{-\frac{r^2}{2}}$$

Entonces,

$$f_R(r) = f_V(v) \cdot \left| \frac{dv}{dr} \right| = r \cdot e^{-\frac{r^2}{2}}$$

Que claramente se refiere a una variable aleatoria $R \sim \text{Rayleigh}(1)$

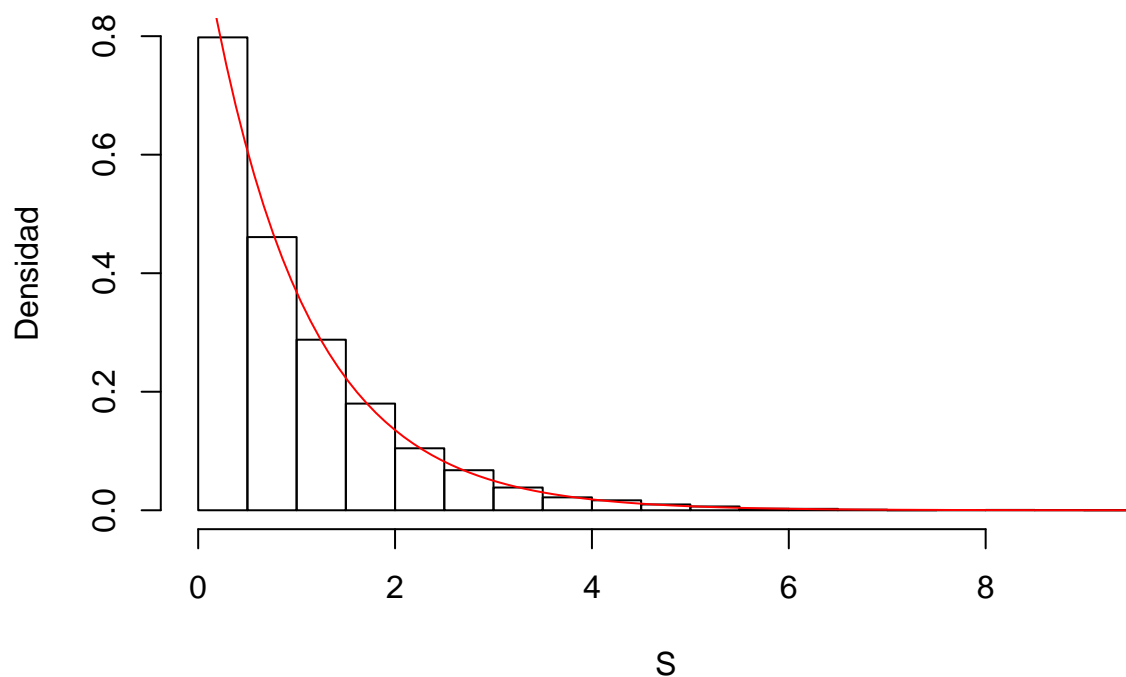
Simulación:

Sea U_2 una variable aleatoria con distribución uniforme en $(0, 1)$

```
set.seed(163053)
U2<-runif(10000)
S<-log(U2)
R<-sqrt(2*S)

hist(S,freq = FALSE, main = "Histograma de S", xlab="S", ylab = "Densidad")
curve(dexp(x,1),add=T,col=2)
```


Histograma de S

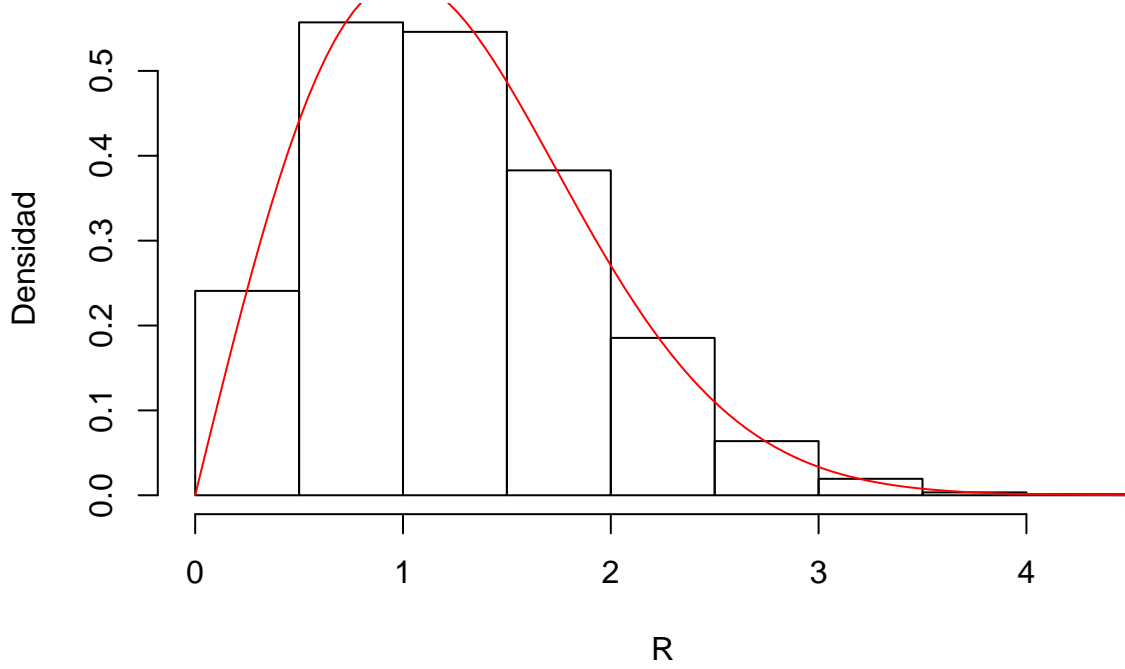


Se observa que los datos de S simulados a partir de una distribución uniforme, se ajusta a una distribución exponencial con parámetro 1.

Por otra parte, los datos simulados de R

```
hist(R,freq = FALSE, main = "Histograma de R", xlab="R", ylab = "Densidad")  
curve(drayleigh(x,scale = 1),add=T,col=2)
```

Histograma de R



Se observa que los datos de R simulados a partir de una distribución exponencial de parámetro 1, se ajusta a una distribución Rayleigh con parámetro 1.

b. Sean $X = R \cos \theta$, $Y = R \sin \theta$. Demuestre que X y Y son normales independientes.

Demostración:

Se sabe que $X = \sqrt{-2 \ln(V)} \cos(2\pi U)$ y $Y = \sqrt{-2 \ln(V)} \sin(2\pi U)$, así, $g(x, y) = (\sqrt{-2 \ln(V)} \cos(2\pi U), \sqrt{-2 \ln(V)} \sin(2\pi U))$

A continuación se calculan la inversa de g :

$$\frac{Y}{X} = \frac{\sqrt{-2 \ln(V)} \cos(2\pi U)}{\sqrt{-2 \ln(V)} \sin(2\pi U)} = \tan(2\pi U) \Rightarrow U = \frac{1}{2\pi} \tan^{-1} \left(\frac{Y}{X} \right) \quad (12)$$

Por otra parte,

$$\begin{aligned} X^2 + Y^2 &= (\sqrt{-2 \ln(V)} \cos(2\pi U))^2 + (\sqrt{-2 \ln(V)} \sin(2\pi U))^2 \\ &= (-2 \ln(V) \cos^2(2\pi U)) + (-2 \ln(V) \sin^2(2\pi U)) \\ &= -2 \ln(V) (\cos^2(2\pi U) + \sin^2(2\pi U)) \\ &= -2 \ln(V) \end{aligned} \quad (13)$$

Así,

$$V = e^{-\frac{X^2 + Y^2}{2}} \quad (14)$$

De esta forma, $g^{-1}(x, y) = \left(\frac{1}{2\pi} \tan^{-1} \left(\frac{Y}{X} \right), e^{-\frac{X^2 + Y^2}{2}} \right)$

El cálculo del jacobiano se presenta a continuación:

$$\begin{aligned}
 |J_{g^{-1}}| &= \left| \det \begin{bmatrix} \frac{y}{2\pi(x^2+y^2)} & -\frac{x}{2\pi(x^2+y^2)} \\ -x \cdot e^{-\frac{x^2+y^2}{2}} & -y \cdot e^{-\frac{x^2+y^2}{2}} \end{bmatrix} \right| \\
 &= \left| \frac{-y^2 \cdot e^{-\frac{x^2+y^2}{2}}}{2\pi(x^2+y^2)} - \frac{-x^2 \cdot e^{-\frac{x^2+y^2}{2}}}{2\pi(x^2+y^2)} \right| \\
 &= \left| \frac{-y^2 \cdot e^{-\frac{x^2+y^2}{2}} - x^2 \cdot e^{-\frac{x^2+y^2}{2}}}{2\pi(x^2+y^2)} \right| \\
 &= \left| \frac{-e^{-\frac{x^2+y^2}{2}}(x^2+y^2)}{2\pi(x^2+y^2)} \right| \\
 &= \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}}
 \end{aligned} \tag{15}$$

De esta forma, la densidad conjunta de X y Y es:

$$f_{X,Y} = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} = f(x)f(y) \tag{16}$$

Lo que demuestra que X y Y se distribuyen normales estándar y además son independientes.

Simulación:

Sea R una variable aleatoria con distribución Rayleigh y U_1 una variable aleatoria uniforme $(0, 1)$.

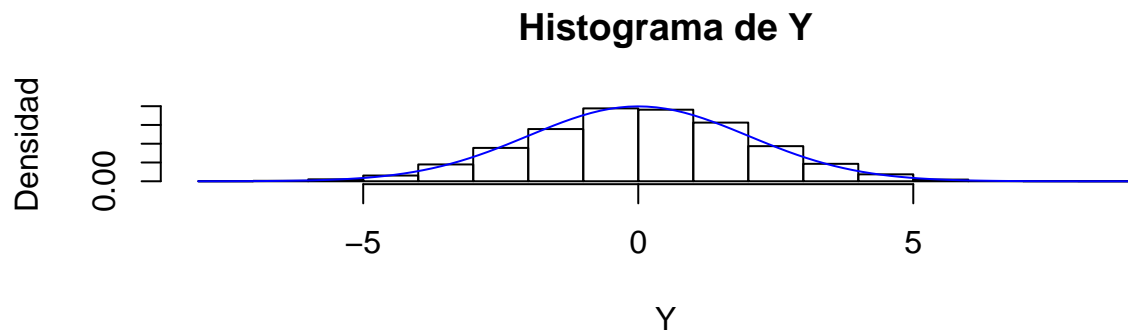
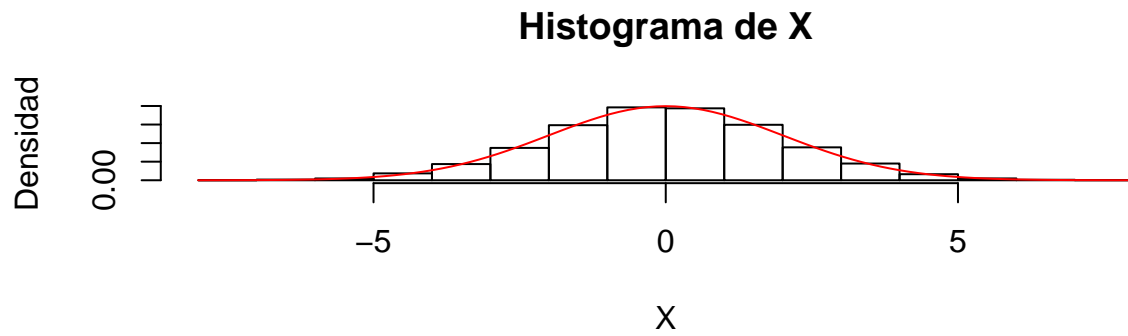
```

set.seed(163053)
R<-rrayleigh(10000,2)
U1<-runif(10000)
Theta<-2*pi*U1

X<-R*cos(Theta)
Y<-R*sin(Theta)

par(mfrow=c(2,1))
hist(X,freq = FALSE, main = "Histograma de X", xlab="X", ylab = "Densidad")
curve(dnorm(x,0,2),add=T,col=2)
hist(Y,freq = FALSE, main = "Histograma de Y", xlab="Y", ylab = "Densidad")
curve(dnorm(x,0,2),add=T,col=4)

```



Puede observarse que tanto las variables simuladas de X y Y , corresponden a variables normales con media 0 y varianza 4, considerando que la distribución de Rayleigh tiene parámetro 2.