act6

Facundo Colasurdo Caldironi

2024-08-16

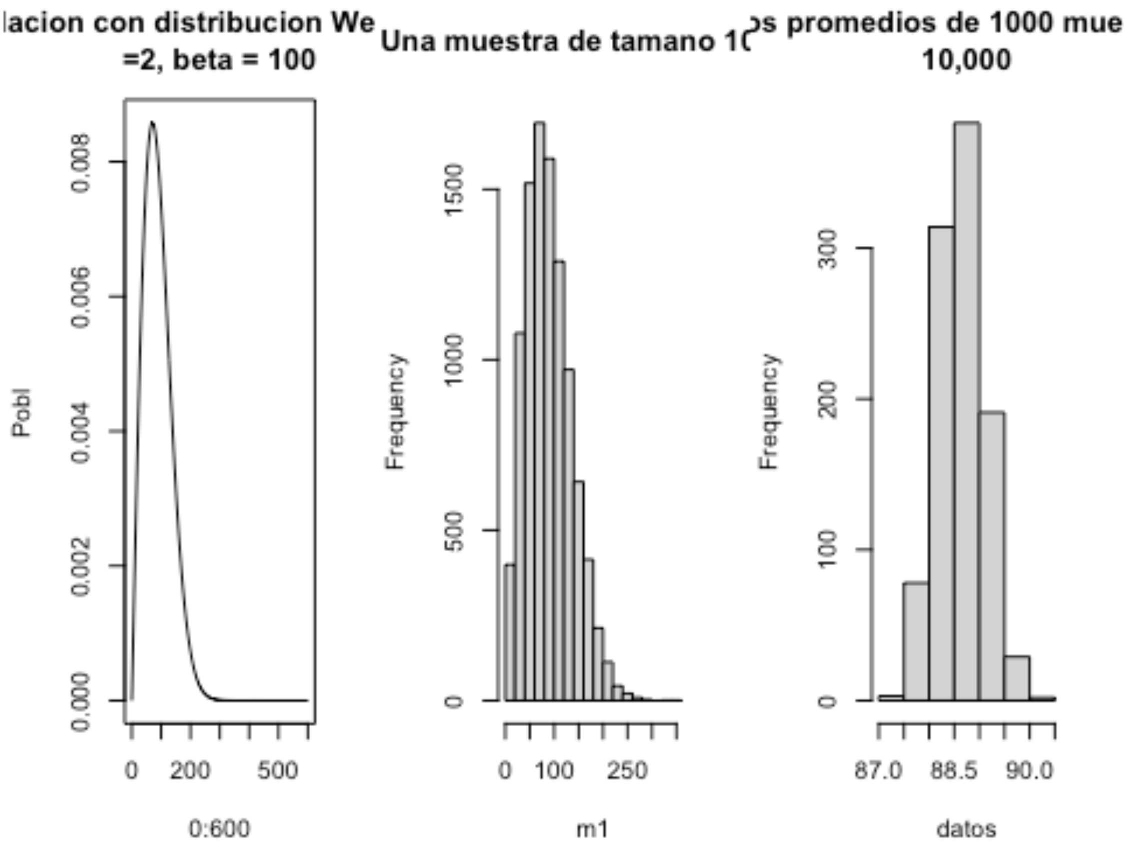
1. Ensayando Distribuciones

Gráfica la Distribución de una variable aleatoria, la de una muestra elegida al azar y la de la Distribución de las medias de 10000 muestras:

Inciso A

Ejecutar el siguiente código de R: DistrsM_enR.txt Download DistrsM_enR.txt. Se esperan tres gráficas, interprete cada una de ellas. Se usa una distribución Weibull, con parámetros alfa = 2 y beta = 100.

```
par(mfrow=c(1,3))
# Graficando una distribucion Weibull de alfa =2, beta = 100
Pobl = dweibull(0:600,2, 100)
plot(0:600,Pobl, type="l", main = "Poblacion con distribucion Weibull
alfa
=2, beta = 100")
# Tomando una muestra de 10000 elementos tomados al azar
m1 = rweibull(10000, 2, 100)
hist(m1, main = "Una muestra de tamano 10000")
# Tomando 1000 promedios de las 1000 muestras como la anterior
m =rweibull(10000,2,100)
prom=mean(m)
datos=prom
for(i in 1:999) {
m =rweibull(10000,2,100)
prom=mean(m)
datos=rbind(datos,prom) }
hist(datos, main="Grafica de los promedios de 1000 muestras de tamano
10,000")
```



La primera grafica muestra la distribucion teórica de la poblacion de Weibull

La segunda grafica enseña como se diistribuyen los datos de una muestra de 10000 extraida de Weibull

La tercera grafica muestra la distribución de promedios de 1000 muestras de tamaño 10000

Inciso B

Cálcula el sesgo y la curtosis de la muestra de tamaño 10000. Aplica una prueba de hipótesis de normalidad. Concluye sobre la normalidad de los datos de la muestra.

```
library(moments)
library(nortest)

# Generar la muestra de 10,000 elementos de la distribución Weibull
```

```r
m =rweibull(10000,2,100)

sesgo <- skewness(m)
cat("Sesgo de la muestra: ", sesgo, "\n")
```

```
## Sesgo de la muestra:  0.6440949
```

```r
kurtosis <- kurtosis(m)
cat("Curtosis de la muestra: ", kurtosis, "\n")
```

```
## Curtosis de la muestra:  3.283268
```

```r
prueba_ad <- ad.test(m)
cat("Prueba de Anderson:\n")
```

```
## Prueba de Anderson:
```

```r
print(prueba_ad)
```

```
##
##  Anderson-Darling normality test
##
## data:  m
## A = 60.616, p-value < 2.2e-16
```

Inciso C

Calcula el sesgo y la curtosis de las medias de las 1000 muestras. Aplica la misma prueba de normalidad que aplicaste a la muestra de tamaño 10000. Concluye sobre la normalidad de las medias de las muestras.

```r
library(moments)

medias <- numeric(1000)
for(i in 1:1000) {
  muestra <- rweibull(10000, 2, 100)
  medias[i] <- mean(muestra)
}

SesgoMedia <- skewness(medias)
cat("Sesgo de las medias de las muestras: ", SesgoMedia, "\n")
```

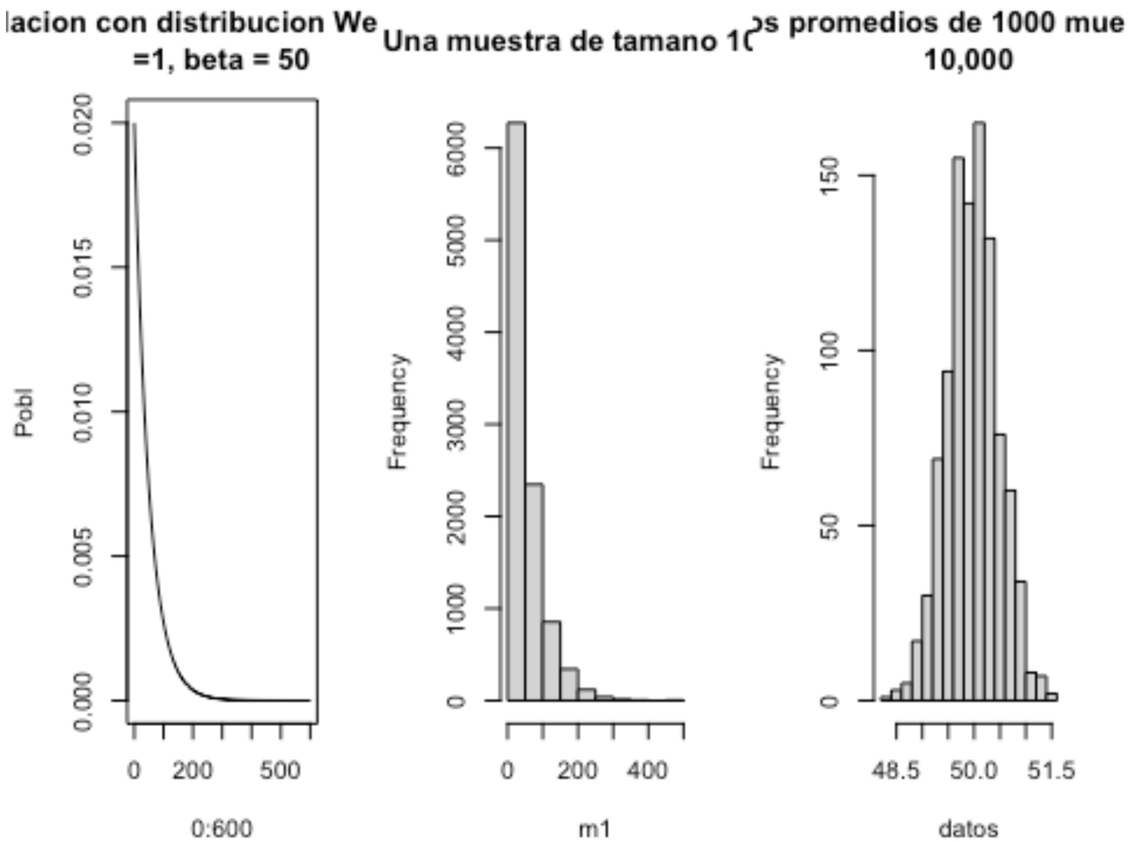
```
## Sesgo de las medias de las muestras:  0.03292805
```

```
CurtoMedia <- kurtosis(medias)
cat("Curtosis de las medias de las muestras: ", CurtoMedia, "\n")

## Curtosis de las medias de las muestras:  2.906709

pruebaMedia <- ad.test(medias)
cat("Prueba de Anderson:\n")

## Prueba de Anderson:

print(pruebaMedia)

##
##  Anderson-Darling normality test
##
## data:  medias
## A = 0.16642, p-value = 0.939
```

Inciso D

Repite el procedimiento A, B y C para otras dos distribuciones que no sean simétricas. Puedes cambiar los valores de alfa y beta para lograr sesgo diferente o puedes ensayar con otra distribución, como la uniforme (punif y runif). Interpreta los resultados.

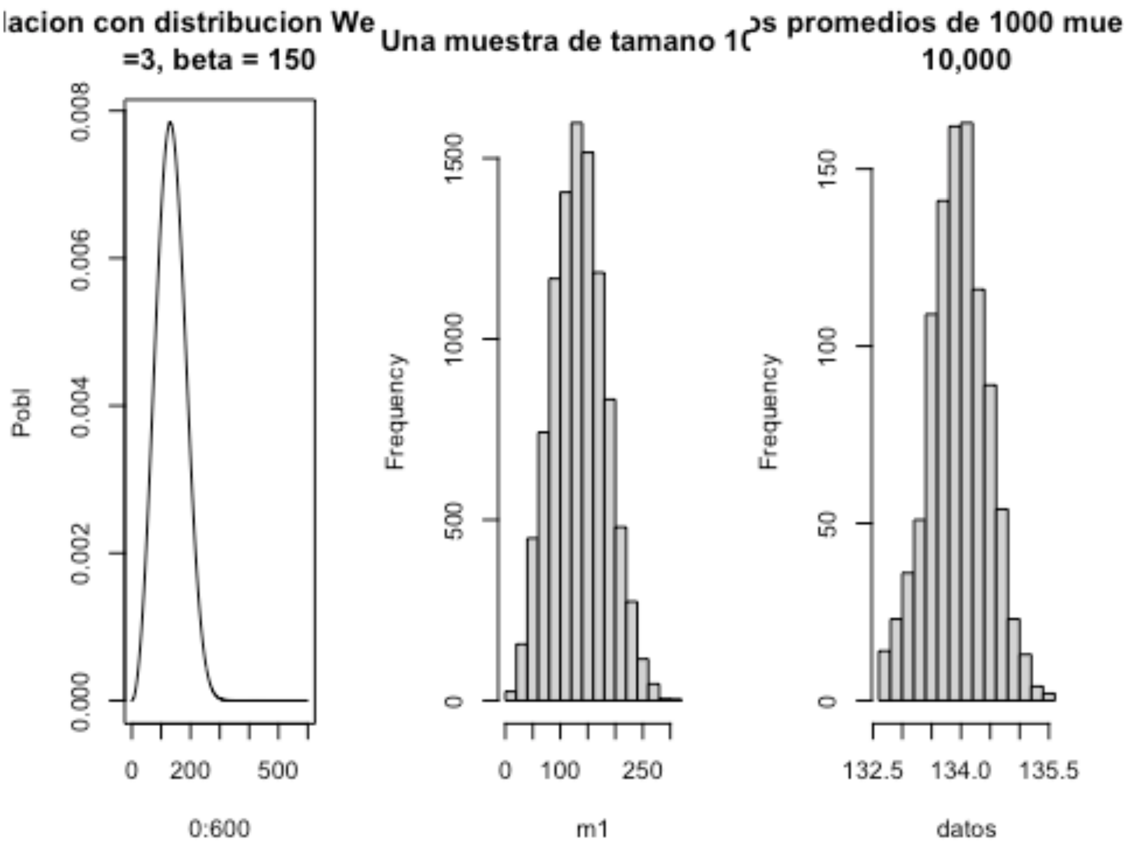
```
par(mfrow=c(1,3))
# Graficando una distribucion Weibull de alfa =1, beta = 50
Pobl = dweibull(0:600,1, 50)
plot(0:600,Pobl, type="l", main = "Poblacion con distribucion Weibull
alfa
=1, beta = 50")
# Tomando una muestra de 10000 elementos tomados al azar
m1 = rweibull(10000, 1, 50)
hist(m1, main = "Una muestra de tamano 10000")
# Tomando 1000 promedios de las 1000 muestras como la anterior
m =rweibull(10000,1,50)
prom=mean(m)
datos=prom
for(i in 1:999) {
m =rweibull(10000,1,50)
prom=mean(m)
datos=rbind(datos,prom) }
hist(datos, main="Grafica de los promedios de 1000 muestras de tamano
10,000")
```




```r
library(moments)
library(nortest)

# Generar la muestra de 10,000 elementos de la distribución Weibull
m =rweibull(10000,1,50)

sesgo <- skewness(m)
cat("Sesgo de la muestra: ", sesgo, "\n")

## Sesgo de la muestra:  2.244425

kurtosis <- kurtosis(m)
cat("Curtosis de la muestra: ", kurtosis, "\n")

## Curtosis de la muestra:  12.20163
```

```r
prueba_ad <- ad.test(m)
cat("Prueba de Anderson:\n")

## Prueba de Anderson:

print(prueba_ad)

##
##  Anderson-Darling normality test
##
## data:  m
## A = 456.25, p-value < 2.2e-16

library(moments)

medias <- numeric(1000)
for(i in 1:1000) {
  muestra <- rweibull(10000, 1, 50)
  medias[i] <- mean(muestra)
}

SesgoMedia <- skewness(medias)
cat("Sesgo de las medias de las muestras: ", SesgoMedia, "\n")

## Sesgo de las medias de las muestras:  0.03953932

CurtoMedia <- kurtosis(medias)
cat("Curtosis de las medias de las muestras: ", CurtoMedia, "\n")

## Curtosis de las medias de las muestras:  2.845178

pruebaMedia <- ad.test(medias)
cat("Prueba de Anderson:\n")

## Prueba de Anderson:

print(pruebaMedia)

##
##  Anderson-Darling normality test
##
## data:  medias
## A = 0.40393, p-value = 0.3543
```

```r
par(mfrow=c(1,3))
# Graficando una distribucion Weibull de alfa =3, beta = 150
Pobl = dweibull(0:600,3, 150)
plot(0:600,Pobl, type="l", main = "Poblacion con distribucion Weibull
alfa
=3, beta = 150")
# Tomando una muestra de 10000 elementos tomados al azar
m1 = rweibull(10000, 3, 150)
hist(m1, main = "Una muestra de tamano 10000")
# Tomando 1000 promedios de las 1000 muestras como la anterior
m =rweibull(10000,3,150)
prom=mean(m)
datos=prom
for(i in 1:999) {
m =rweibull(10000,3,150)
prom=mean(m)
datos=rbind(datos,prom) }
hist(datos, main="Grafica de los promedios de 1000 muestras de tamano
10,000")
```




```r
library(moments)
library(nortest)

# Generar la muestra de 10,000 elementos de la distribución Weibull
m =rweibull(10000,3,150)

sesgo <- skewness(m)
cat("Sesgo de la muestra: ", sesgo, "\n")

## Sesgo de la muestra:  0.1450677

kurtosis <- kurtosis(m)
cat("Curtosis de la muestra: ", kurtosis, "\n")

## Curtosis de la muestra:  2.759453
```

```r
prueba_ad <- ad.test(m)
cat("Prueba de Anderson:\n")

## Prueba de Anderson:

print(prueba_ad)

##
##  Anderson-Darling normality test
##
## data:  m
## A = 3.5285, p-value = 8.191e-09

library(moments)

medias <- numeric(1000)
for(i in 1:1000) {
  muestra <- rweibull(10000, 3, 150)
  medias[i] <- mean(muestra)
}

SesgoMedia <- skewness(medias)
cat("Sesgo de las medias de las muestras: ", SesgoMedia, "\n")

## Sesgo de las medias de las muestras:  0.04510495

CurtoMedia <- kurtosis(medias)
cat("Curtosis de las medias de las muestras: ", CurtoMedia, "\n")

## Curtosis de las medias de las muestras:  2.983144

pruebaMedia <- ad.test(medias)
cat("Prueba de Anderson:\n")

## Prueba de Anderson:

print(pruebaMedia)

##
##  Anderson-Darling normality test
##
## data:  medias
## A = 0.3343, p-value = 0.5079
```

Inciso E

Concluye sobre las semejanzas y diferencias entre los tres gráficos generados en cada una de las tres distribuciones teóricas.

Semejanzas: Las tres tienen la tercera gráfica, la gráfica de los promedios de 1000 muestras de tamaño 10,000 muy semejante ya que esta es una constante entre las tres distribuciones.

Diferencias: Las principales diferencias entre los tres gráficos se pueden apreciar en las gráficas Población con distribución Weibull, en donde se puede observar asimetría, por otra parte, también se aprecia diferencias en la gráfica de una muestra de tamaño 10000,

2. Remaches

La resistencia a la ruptura de un remache tiene un valor medio de 10,000 lb/pulg2 y una desviación estándar de 500 lb/pulg2. Si se sabe que la población se distribuye normalmente,

$X: La resistencia a la ruptura de un remache\ X \sim N(\mu_x = 10000, \sigma_x = 500)$

Inciso A

¿Cuál es la probabilidad de que al tomar un remache al azar de esa población, éste tenga una resistencia a la ruptura que esté a 100 unidades alrededor de la media? ¿a cuántas desviaciones estándar está de la media? $\p(9900 < \bar{x} < 10100)$

```
p1 = pnorm(10100, 10000, 500) - pnorm(9900, 10000, 500)
cat("p(9900 < X < 10100)=", p1)

## p(9900 < X < 10100)= 0.1585194

 z = (10100-10000)/500
cat("Desviaciones estandar lejos de la media)=", z)

## Desviaciones estandar lejos de la media)= 0.2
```

Inciso B

¿Cuál es la probabilidad de que la resistencia media a la ruptura de la muestra aleatoria de 120 remaches esté 100 unidades alrededor de su media? ¿a cuántas desviaciones estándar está de la media? $\p(9900 < \bar{x} < 10100)$ $\bar{X} \sim N\left(\mu_{\bar{X}} = 10000, \sigma_{\bar{X}} = \frac{500}{\sqrt{120}}\right)$

```r
p2 = pnorm(10100, 10000, 500/sqrt(120)) - pnorm(9900, 10000, 500/
sqrt(120))
cat("p(9900 < X_b < 10100)=", p2)

## p(9900 < X_b < 10100)= 0.9715403

 z1 = (10100-10000)/(500/sqrt(120))
cat("Desviaciones estandar lejos de la media)=", z1)

## Desviaciones estandar lejos de la media)= 2.19089
```

Inciso C

Si el tamaño muestral hubiera sido 15, en lugar de 120, ¿cuál es la probabilidad de que la resistencia media a la ruptura esté 100 unidades alrededor de la media? ¿a cuántas desviaciones estándar está de la media?

$\p(9900 < \bar{x} < 10100)$ $\bar{X} \sim N\left(\mu_{\bar{X}} = 10000, \sigma_{\bar{X}} = \frac{500}{\sqrt{15}}\right)$

```r
p3 = pnorm(10100, 10000, 500/sqrt(15)) - pnorm(9900, 10000, 500/
sqrt(15))
cat("p(9900 < X_b < 10100)=", p3)

## p(9900 < X_b < 10100)= 0.561422

 z2 = (10100-10000)/(500/sqrt(15))
cat("Desviaciones estandar lejos de la media)=", z2)

## Desviaciones estandar lejos de la media)= 0.7745967
```

Inciso D

Un ingeniero recibió un lote muy grande de remaches. Antes de aceptarlo quiso verificar si efectivamente la media de la resistencia de los remaches es de 10 000 lb/pulg2. Para ello tomó una muestra de 120 remaches elegidos al azar tenía media de 9800 lb/pulg2 y rechazó el pedido, ¿hizo lo correcto? ¿por qué?.

$\p(9900 < \bar{x} < 9800)$ $\bar{X} \sim N\left(\mu_{\bar{X}} = 10000, \sigma_{\bar{X}} = \frac{500}{\sqrt{120}}\right)$

```r
p4 = pnorm(9800, 10000, 500/sqrt(120))
cat("p(< X_b < 9800)=", p4)
```

```
## p(< X_b < 9800)= 5.88567e-06

 z3 = (10000-9800)/(500/sqrt(120))
cat("Desviaciones estandar lejos de la media)=", z3)

## Desviaciones estandar lejos de la media)= 4.38178
```

Es correcto que no lo haya aceptado, la probabilidad es muy baja y debido a que la desviación estándar esta muy lejos de la media.

Inciso E

¿Qué decisión recomiendas al ingeniero si la media obtenida en la media hubiera sido 9925? ¿recomendarías rechazarlo? No es recomendable, ya que esta si se encuentra a una desviación estándar de la media, por lo tanto si valdría la pena conseguirla.

```
p5 = pnorm(9925, 10000, 500/sqrt(120))
cat("p(< X_b < 9925)=", p5)

## p(< X_b < 9925)= 0.05017412

 z4 = (10000-9925)/(500/sqrt(120))
cat("Desviaciones estandar lejos de la media)=", z4)

## Desviaciones estandar lejos de la media)= 1.643168
```

##3. Embotellando

Una máquina embotelladora puede ser regulada para que se descargue un promedio de μ onzas por botella. Se ha observado que la cantidad de líquido dosificado por una máquina embotelladora está distribuida normalmente con σ = 1 onza. La máquina embotelladora se calibra cuando la media de una muestra tomada al azar está fuera del 95% central de la distribución muestral. La media de la cantidad de líquido deseada requiere que μ sea de 15 onzas.

Inciso A

¿A cuántas desviaciones estándar alrededor de la verdadera media μ puede estar la media de una muestra para que esté dentro del estándar establecido del 95% central?

```
z6 = qnorm(0.025, 0 ,1)
cat("Desviaciones estandar lejos de la media)=", z6)

## Desviaciones estandar lejos de la media)= -1.959964
```

Inciso B

¿Cuál es la probabilidad de que en una muestra aleatoria de tamaño 10 botellas se obtenga una media mayor a 16 onzas?

```
z7 = pnorm(15, 16, 1/sqrt(10))
cat("Probabilidad=", z7)

## Probabilidad= 0.0007827011
```

Inciso C

Si en una muestra aleatoria de tamaño 10 botellas se obtuvo una media de 16 onzas, ¿se detendría la producción para calibrar la máquina?

```
 z8 = (16 - 15)/(1/sqrt(10))
cat("Desviaciones estandar lejos de la media)=", z8)

## Desviaciones estandar lejos de la media)= 3.162278
```

Se tendria que detener debido a que se encuentra muy lejos de la media

Inciso D

¿Cuál es la probabilidad de que en una muestra aleatoria de tamaño 10 botellas se obtenga una media menor a 14.5 onzas?

```
z9 = pnorm(14.5, 15, 1/sqrt(10))
cat("Probabilidad=", z9)

## Probabilidad= 0.05692315
```

Inciso E

Si en una muestra aleatoria de tamaño 10 botellas se obtuvo una media de 15.5 onzas, ¿se detendría la producción para calibrar la máquina?

```
 z10 = (15.5 - 15)/(1/sqrt(10))
cat("Desviaciones estandar lejos de la media)=", z10)

## Desviaciones estandar lejos de la media)= 1.581139
```

No se detendria la produccion, debido a que se encuentra dentro del 95% de los datos, al mismo tiempo que nollega al limite de 1.96

Inciso F

Hacer una gráfica del inciso 1.

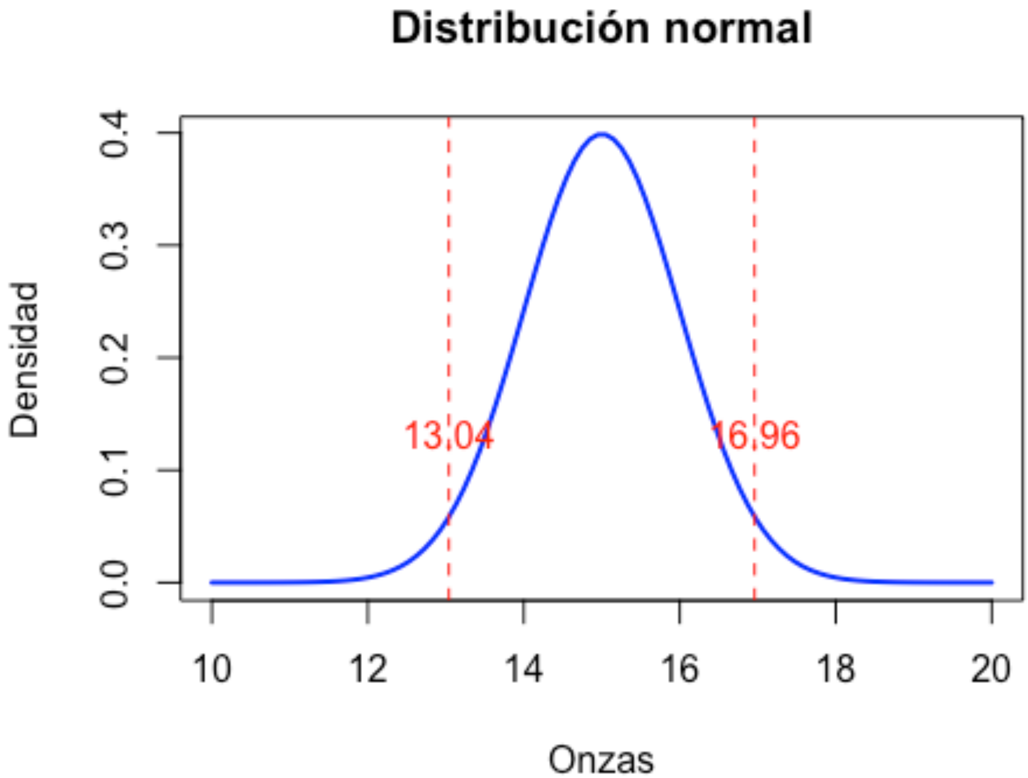
```
mu <- 15
sigma <- 1
z <- qnorm(0.975)

# Límites
lower_limit <- mu - z * sigma
upper_limit <- mu + z * sigma
cat("El límite inferior:", lower_limit, "\n")
```

```
## El límite inferior: 13.04004
```

```
cat("El límite superior:", upper_limit, "\n")
```

```
## El límite superior: 16.95996
```

```
# Graficar
x <- seq(10, 20, length = 100)
y <- dnorm(x, mean = mu, sd = sigma)

plot(x, y, type = "l", lwd = 2, col = "blue", ylab = "Densidad", xlab
= "Onzas", main = "Distribución normal")
abline(v = c(lower_limit, upper_limit), col = "red", lty = 2)
text(lower_limit, 0.1, round(lower_limit, 2), col = "red", pos = 3)
text(upper_limit, 0.1, round(upper_limit, 2), col = "red", pos = 3)
```



Distribución normal
Densidad
0.0
0.1
0.2
0.3
0.4
13,04
16,96
10
12
14
16
18
20
Onzas