

# Distribuciones Muestrales y TCL

Catherine Rojas

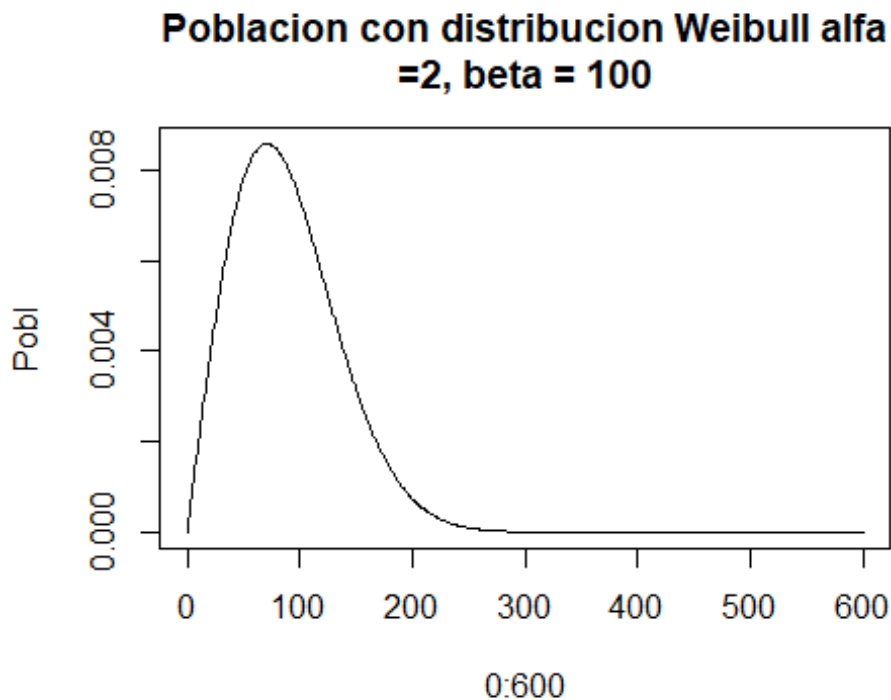
2024-08-16

## Ensayando Distribuciones

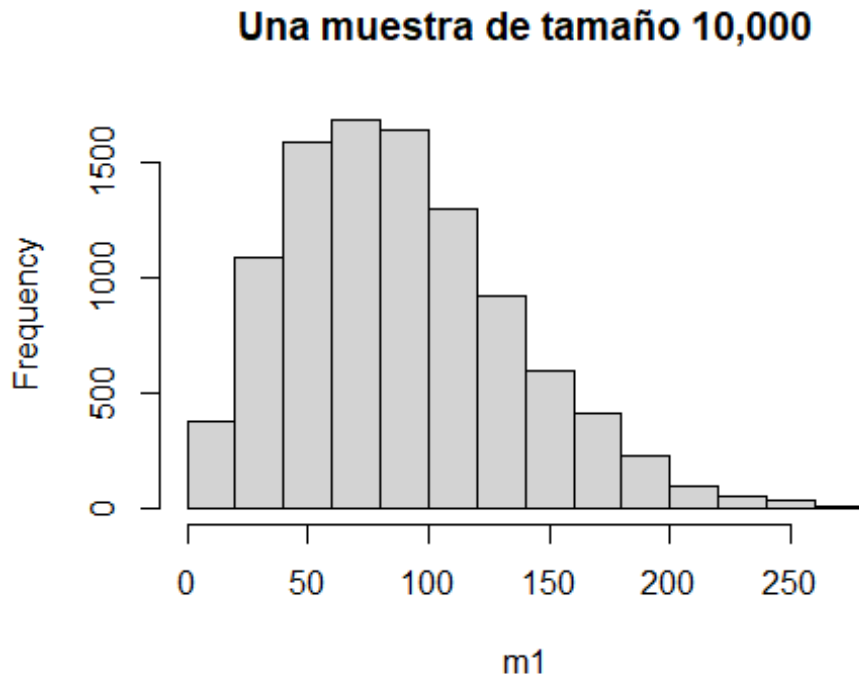
Gráfica la Distribución de una variable aleatoria, la de una muestra elegida al azar y la de la Distribución de las medias de 10000 muestras:

a) Ejercutar 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  $\alpha = 2$  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")
```

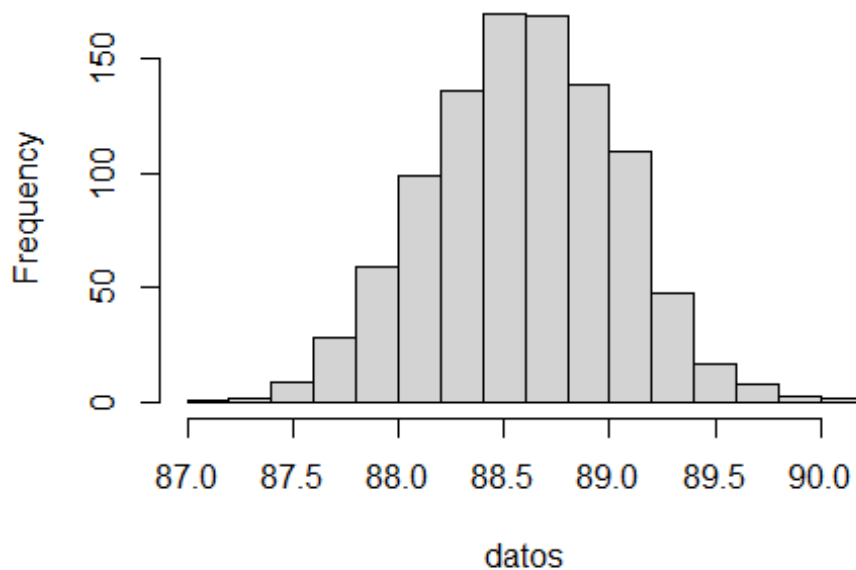


```
# Tomando una muestra de 10,000 elementos tomados al azar
m1 = rweibull(10000, 2, 100) # muestra 1
hist(m1, main = "Una muestra de tamaño 10,000")
```



```
# Tomando 1000 promedios de las 1,000 muestras como la anterior
m = rweibull(10000, 2, 100) # muestra aleatoria
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 1,000 muestras de tamaño 10,000")
```

## Grafica de los promedios de 1,000 muestras de tamaño 10,000



### Interpretación de las graficas

- **Gráfica de la Distribución Weibull Teórica** Esta gráfica muestra la función de densidad de probabilidad de una distribución Weibull con parámetros  $\alpha = 2$  y  $\beta = 100$ . La curva tiene un pico que indica la mayor concentración de probabilidad cerca de un valor alrededor de 70 y 100, y luego desciende gradualmente.

La curva es asimétrica hacia la derecha, esto indica que la mayoría de los valores se encuentran cerca del pico, pero hay una cola a la derecha que representa eventos menos probables con valores más altos.

- **Histograma de una Muestra Aleatoria** Este histograma muestra la distribución de una muestra aleatoria de tamaño 10,000 extraída de la distribución Weibull obtenida anteriormente.

La distribución de la muestra refleja la distribución teórica. En este caso, existe un mayor número de observaciones aproximadamente entre 50 y 100, coincidiendo con el pico de la distribución teórica. Sin embargo, dado que es una muestra aleatoria, esta la forma llega a ser similar pero no idéntica a la teórica.

**Histograma de los Promedios de 1000 Muestras** Esta gráfica muestra la distribución de los promedios de 1,000 muestras de tamaño 10,000. Cada barra en el histograma representa la frecuencia con la que se repiten ciertos valores medios de las muestras.

Este histograma se aproxima a una distribución normal. La mayoría de los promedios están concentrados entre 88 y 89, con una menor variabilidad comparada con las

distribuciones individuales de las muestras. Esto es porque, al promediar grandes muestras, las fluctuaciones extremas se cancelan entre sí, lo que resulta en una distribución más concentrada alrededor de la media.

**b) Cálcula el sesgo y la curtosis de la muestra de tamaño 10,000. Aplica una prueba de hipótesis de normalidad. Concluye sobre la normalidad de los datos de la muestra.**

```
library(e1071)

## Warning: package 'e1071' was built under R version 4.3.3

# Calcular el sesgo
sesgo_m1 <- skewness(m1)
cat("\nSesgo de la muestra de tamaño 10,000:\n", sesgo_m1)

##
## Sesgo de la muestra de tamaño 10,000:
## 0.6604592

# Calcular la curtosis
curtosis_m1 <- kurtosis(m1)
cat("\nCurtosis de la muestra de tamaño 10,000:\n", curtosis_m1)

##
## Curtosis de la muestra de tamaño 10,000:
## 0.2155275
```

### Interpretación de resultados

El sesgo positivo indica que la distribución de la muestra tiene una cola más larga a la derecha de la distribución. Sin embargo, como se observó en las gráficas, la mayoría de los datos se encuentran a la izquierda de la media, pero hay valores más altos que extienden la cola hacia la derecha.

La curtosis obtenida es positiva e indica que los datos muestran más valores atípicos extremos que una distribución normal.

$H_0$  = Los datos siguen una distribución normal

```
library(nortest)
# Para concluir sobre la normalidad

# Anderson-Darling
ad_test_m1 <- ad.test(m1)
ad_test_m1

##
## Anderson-Darling normality test
##
## data: m1
## A = 67.286, p-value < 2.2e-16
```

```

if (ad_test_m1$p.value < 0.05) {
  conclusion <- "Los datos no siguen una distribución normal."
} else {
  conclusion <- "No se puede rechazar la hipótesis de que los datos
siguen una distribución normal."
}
conclusion

## [1] "Los datos no siguen una distribución normal."

library(tseries)

## Warning: package 'tseries' was built under R version 4.3.3

## Registered S3 method overwritten by 'quantmod':
##   method      from
##   as.zoo.data.frame zoo

# Jarque-Bera
jb_test_m1 <- jarque.bera.test(m1)
jb_test_m1

##
##  Jarque Bera Test
##
## data:  m1
## X-squared = 746.7, df = 2, p-value < 2.2e-16

if (jb_test_m1$p.value < 0.05) {
  conclusion <- "Los datos no siguen una distribución normal."
} else {
  conclusion <- "No se puede rechazar la hipótesis de que los datos
siguen una distribución normal."
}
conclusion

## [1] "Los datos no siguen una distribución normal."

```

## Interpretación de resultados

Ambas pruebas de normalidad, indican que la muestra de tamaño 10,000 no sigue una distribución normal al obtener p-value extremadamente bajos.

**c) Calcula el sesgo y la curtosis de las medias de las 1,000 muestras. Aplica la misma prueba de normalidad que aplicaste a la muestra de tamaño 10,000. Concluye sobre la normalidad de las medias de las muestras.**

```

library(e1071)
# Calcular el sesgo
sesgo_m <- skewness(m)
cat("\nSesgo de la muestra de tamaño 10,000:\n", sesgo_m)

```

```
##
## Sesgo de la muestra de tamaño 10,000:
## 0.6617976

# Calcular la curtosis
curtosis_m <- kurtosis(m)
cat("\nCurtosis de la muestra de tamaño 10,000:\n", curtosis_m)

##
## Curtosis de la muestra de tamaño 10,000:
## 0.2733707
```

$H_0$  = Los datos siguen una distribución normal

*# Para concluir sobre la normalidad*

*# Anderson-Darling*

```
ad_test_m <- ad.test(m)
ad_test_m
```

```
##
## Anderson-Darling normality test
##
## data: m
## A = 65.689, p-value < 2.2e-16

if (ad_test_m$p.value < 0.05) {
  conclusion <- "Los datos no siguen una distribución normal."
} else {
  conclusion <- "No se puede rechazar la hipótesis de que los datos
siguen una distribución normal."
}
conclusion

## [1] "Los datos no siguen una distribución normal."
```

*# Jarque-Bera*

```
jb_test_m <- jarque.bera.test(m)
jb_test_m
```

```
##
## Jarque Bera Test
##
## data: m
## X-squared = 761.47, df = 2, p-value < 2.2e-16

if (jb_test_m$p.value < 0.05) {
  conclusion <- "Los datos no siguen una distribución normal."
} else {
  conclusion <- "No se puede rechazar la hipótesis de que los datos
siguen una distribución normal."
```

```
}
conclusion

## [1] "Los datos no siguen una distribución normal."
```

### Interpretación de resultados

Los resultados obtenidos son similares a los del inciso b), por lo que se determina que los datos de las medias de las 1,000 muestras no siguen una distribución normal.

**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.**

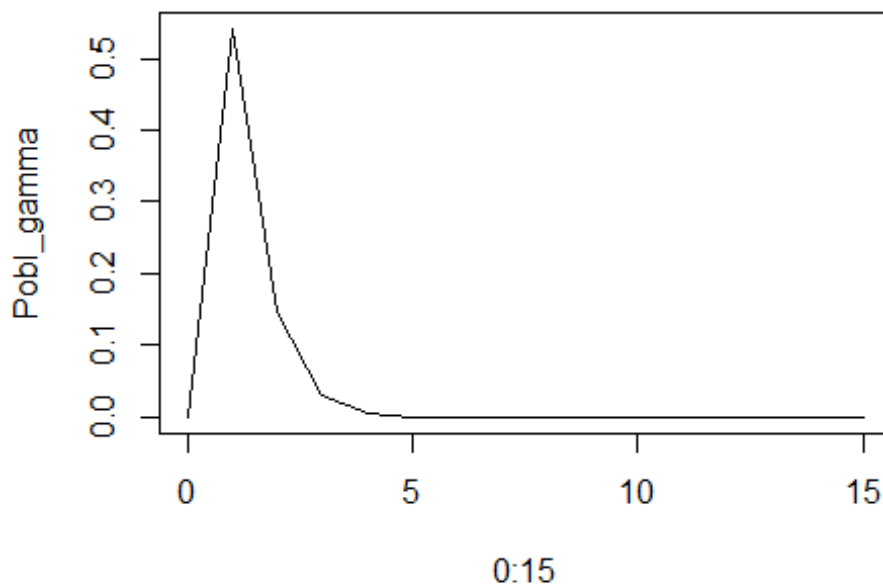
### Distribución Gamma

#### a) Procedimiento con Distribución Gamma

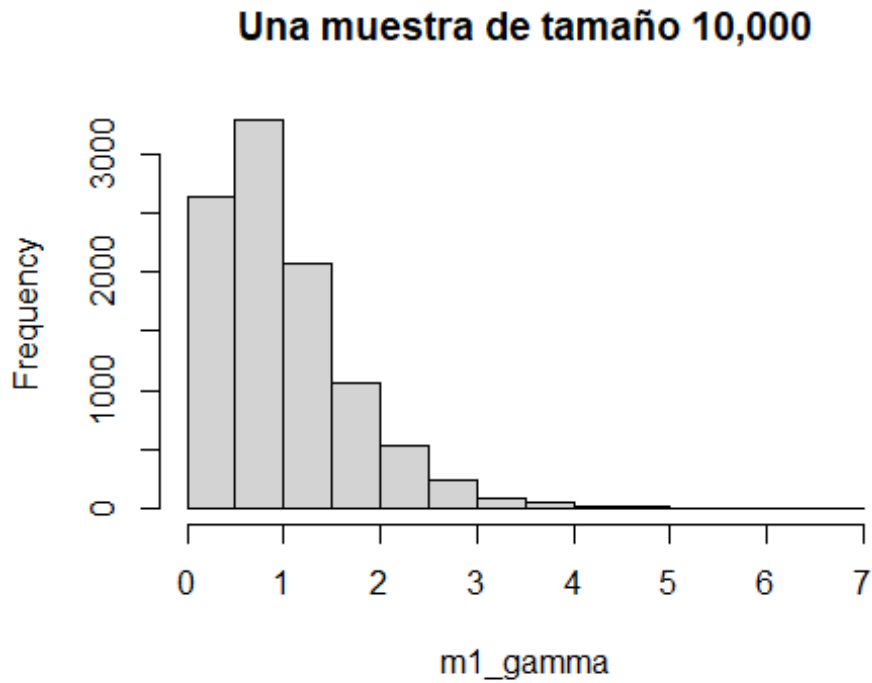
```
#par(mfrow=c(1,3))

# Graficando una distribucion Gamma de shape = 2, scale = 2
Pobl_gamma = dgamma(0:15, 2, 2)
plot(0:15, Pobl_gamma, type="l", main = "Poblacion con distribucion Gamma
shape=2, scale=2")
```

#### Poblacion con distribucion Gamma shape=2, scale



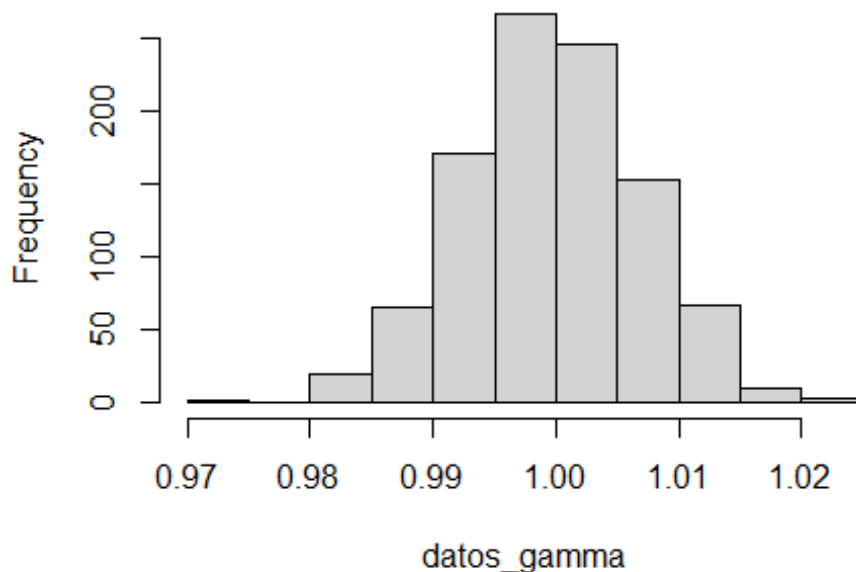
```
# Tomando una muestra de 10,000 elementos tomados al azar
m1_gamma = rgamma(10000, 2, 2)
hist(m1_gamma, main = "Una muestra de tamaño 10,000")
```



```
# Tomando 1000 promedios de las 1,000 muestras como la anterior
m_gamma = rgamma(10000, 2, 2)
prom_gamma = mean(m_gamma)
datos_gamma = prom_gamma
for(i in 1:999) {
  m_gamma = rgamma(10000, 2, 2)
  prom_gamma = mean(m_gamma)
  datos_gamma = rbind(datos_gamma, prom_gamma)
}
hist(datos_gamma, main="Grafica de los promedios de 1,000 muestras de
tamaño 10,000")
```



## Gráfica de los promedios de 1,000 muestras de tamaño



### Interpretación de gráficas

- **Gráfica de la Distribución Gamma Teórica** La primera gráfica confirma que la distribución Gamma es asimétrica con una cola hacia la derecha, lo cual es esperado.

- **Histograma de una Muestra Aleatoria** Muestra que esta asimetría también está presente en una muestra grande extraída de la distribución, reflejando la forma teórica.

**Histograma de los Promedios de 1000 Muestras** Indica que, al promediar un gran número de muestras, la distribución de los promedios se aproxima a una normalidad.

Los resultados confirman la idea de que, aunque las distribuciones individuales pueden ser altamente asimétricas, los promedios de un gran número de muestras tienden a ser normales.

## b) Calculo de sesgo, curtosis y pruebas de normalidad para la muestra Gamma

```
# Calcular el sesgo
sesgo_m1_gamma <- skewness(m1_gamma)
cat("\nSesgo de la muestra de tamaño 10,000:\n", sesgo_m1_gamma)

##
## Sesgo de la muestra de tamaño 10,000:
## 1.444515
```

```

# Calcular La curtosis
kurtosis_m1_gamma <- kurtosis(m1_gamma)
cat("\nCurtosis de la muestra de tamaño 10,000:\n", kurtosis_m1_gamma)

##
## Curtosis de la muestra de tamaño 10,000:
## 3.292643

# Anderson-Darling para La muestra Gamma
ad_test_m1_gamma <- ad.test(m1_gamma)
ad_test_m1_gamma

##
## Anderson-Darling normality test
##
## data: m1_gamma
## A = 225.95, p-value < 2.2e-16

if (ad_test_m1_gamma$p.value < 0.05) {
  conclusion <- "Los datos no siguen una distribución normal."
} else {
  conclusion <- "No se puede rechazar la hipótesis de que los datos
siguen una distribución normal."
}
conclusion

## [1] "Los datos no siguen una distribución normal."

# Jarque-Bera para La muestra Gamma
jb_test_m1_gamma <- jarque.bera.test(m1_gamma)
jb_test_m1_gamma

##
## Jarque Bera Test
##
## data: m1_gamma
## X-squared = 7999.5, df = 2, p-value < 2.2e-16

if (jb_test_m1_gamma$p.value < 0.05) {
  conclusion <- "Los datos no siguen una distribución normal."
} else {
  conclusion <- "No se puede rechazar la hipótesis de que los datos
siguen una distribución normal."
}
conclusion

## [1] "Los datos no siguen una distribución normal."

```

**Interpretación de resultados** Los resultados indican que la muestra de tamaño 10,000, extraída de la distribución Gamma, es altamente asimétrica y tiene colas pesadas, lo que se refleja en el alto sesgo y la curtosis. A diferencia de incisos anteriores, estas gráficas se ven mucho más picudas.

Las pruebas de normalidad confirman que los datos no siguen una distribución normal. Esto es consistente con la naturaleza de la distribución Gamma, que es asimétrica y no normal con los parámetros utilizados.

### c) Calculo de sesgo, curtosis y pruebas de normalidad para los promedios Gamma

```
# Calcular el sesgo
sesgo_m_gamma <- skewness(datos_gamma)
cat("\nSesgo de las medias de las 1,000 muestras:\n", sesgo_m_gamma)

##
## Sesgo de las medias de las 1,000 muestras:
## 0.005694203

# Calcular la curtosis
curtosis_m_gamma <- kurtosis(datos_gamma)
cat("\nCurtosis de las medias de las 1,000 muestras:\n",
    curtosis_m_gamma)

##
## Curtosis de las medias de las 1,000 muestras:
## -0.1623425

# Anderson-Darling para Los promedios Gamma
ad_test_m_gamma <- ad.test(datos_gamma)
ad_test_m_gamma

##
## Anderson-Darling normality test
##
## data:  datos_gamma
## A = 0.28685, p-value = 0.6217

if (ad_test_m_gamma$p.value < 0.05) {
  conclusion <- "Los datos no siguen una distribución normal."
} else {
  conclusion <- "No se puede rechazar la hipótesis de que los datos
siguen una distribución normal."
}
conclusion

## [1] "No se puede rechazar la hipótesis de que los datos siguen una
distribución normal."

# Jarque-Bera para Los promedios Gamma
jb_test_m_gamma <- jarque.bera.test(datos_gamma)
jb_test_m_gamma

##
## Jarque Bera Test
```

```
##
## data:  datos_gamma
## X-squared = 1.028, df = 2, p-value = 0.5981

if (jb_test_m_gamma$p.value < 0.05) {
  conclusion <- "Los datos no siguen una distribución normal."
} else {
  conclusion <- "No se puede rechazar la hipótesis de que los datos
siguen una distribución normal."
}
conclusion

## [1] "No se puede rechazar la hipótesis de que los datos siguen una
distribución normal."
```

**Interpretación de los resultados** Tanto el sesgo cercano a 0, (distribución de las medias aprximadamente simétrica en torno a la media) como la curtosis cercana a 3, junto con los resultados de las pruebas de normalidad, sugieren que la distribución de las medias de las 1,000 muestras es normal, pues la distribución de las medias de un número suficientemente grande de muestras, independientemente de la forma de la distribución original, tiende a ser normal como se menciona anteriormente.

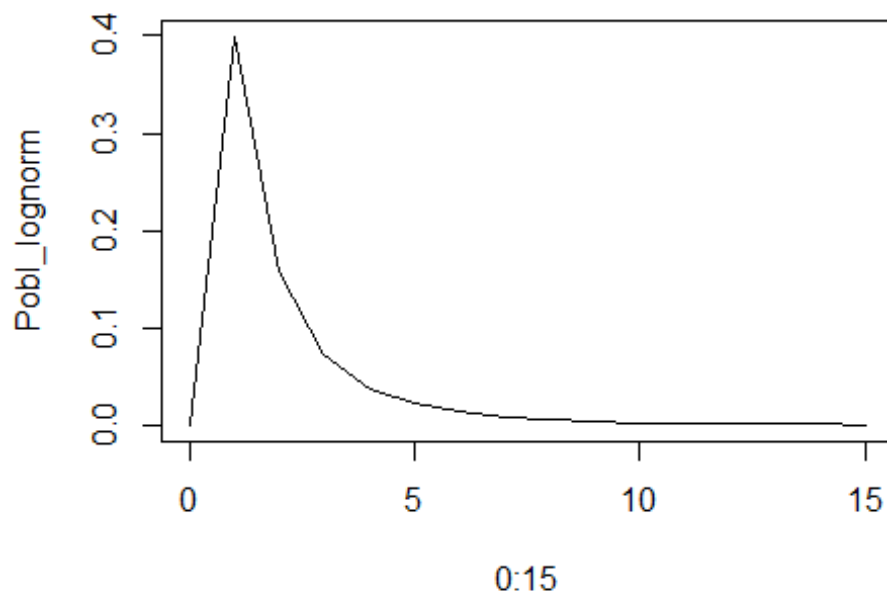
## Distribución Log-Normal

### a) Procedimiento con Distribución Log-Normal

```
#par(mfrow=c(1,3))

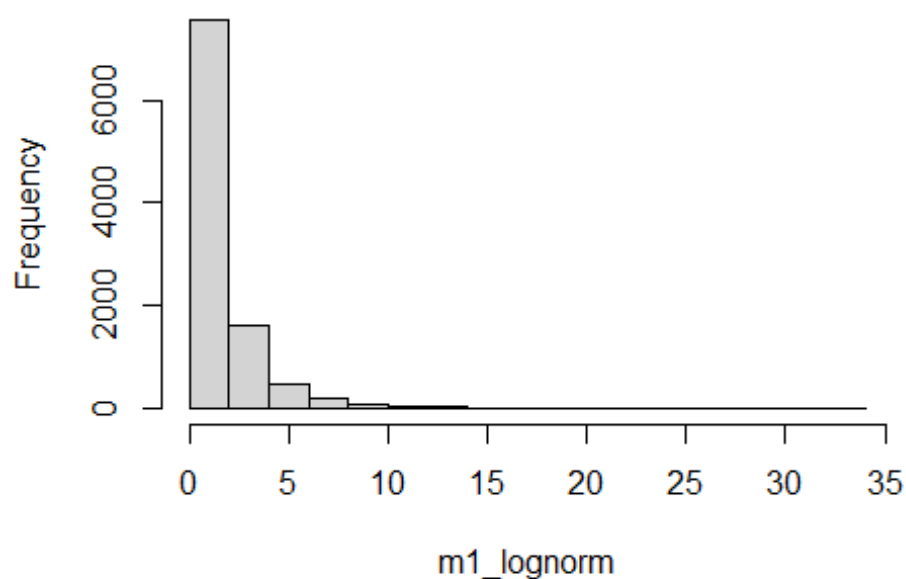
# Graficando una distribucion Log-Normal de meanLog = 0, sdlog = 1
Pobl_lognorm = dlnorm(0:15, 0, 1)
plot(0:15, Pobl_lognorm, type="l", main = "Poblacion con distribucion
Log-Normal meanlog=0, sdlog=1")
```

Poblacion con distribucion Log-Normal meanlog=0, sc



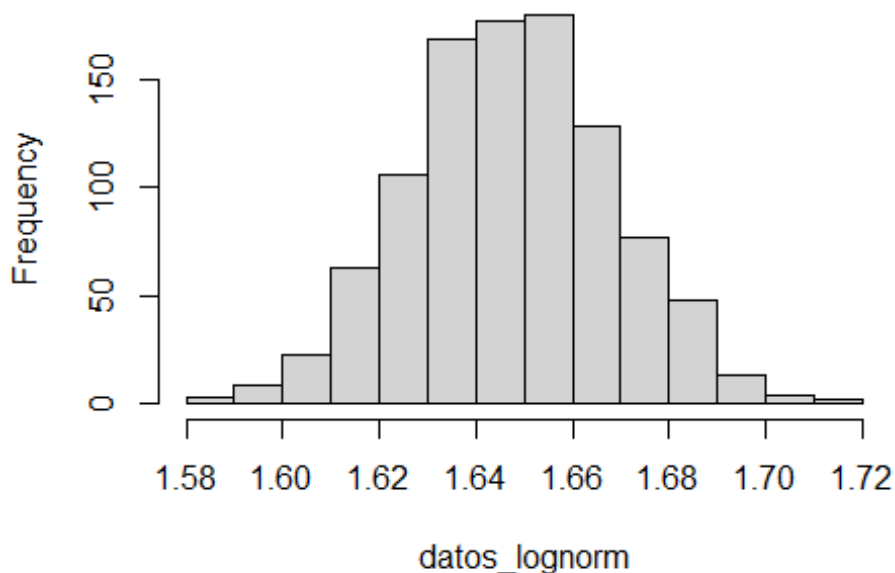
```
# Tomando una muestra de 10,000 elementos tomados al azar  
m1_lognorm = rlnorm(10000, 0, 1)  
hist(m1_lognorm, main = "Una muestra de tamaño 10,000")
```

Una muestra de tamaño 10,000



```
# Tomando 1000 promedios de las 1,000 muestras como la anterior
m_lognorm = rlnorm(10000, 0, 1)
prom_lognorm = mean(m_lognorm)
datos_lognorm = prom_lognorm
for(i in 1:999) {
  m_lognorm = rlnorm(10000, 0, 1)
  prom_lognorm = mean(m_lognorm)
  datos_lognorm = rbind(datos_lognorm, prom_lognorm)
}
hist(datos_lognorm, main="Grafica de los promedios de 1,000 muestras de
tamaño 10,000")
```

## Grafica de los promedios de 1,000 muestras de tamaño



### Interpretación de gráficos

- **Gráfica de la Distribución Log-Normal Teórica** La primera gráfica muestra la asimetría y cola larga hacia la derecha de una distribución Log-Normal.
- **Histograma de una Muestra Aleatoria** Existe la misma asimetría de la distribución teórica, con una alta concentración de valores bajos y una cola larga de valores más altos.

**Histograma de los Promedios de 1000 Muestras** Al promediar 1,000 muestras grandes, la distribución de los promedios tiende hacia la normalidad. Esto subraya cómo, a pesar de la asimetría inicial de la distribución original, el proceso de muestreo y promedio puede dar lugar a una distribución aproximadamente normal.

## b) Cálculo de sesgo, curtosis y pruebas de normalidad para la muestra Log-Normal

```
# Calcular el sesgo
sesgo_m1_lognorm <- skewness(m1_lognorm)
cat("\nSesgo de la muestra de tamaño 10,000:\n", sesgo_m1_lognorm)

##
## Sesgo de la muestra de tamaño 10,000:
## 4.158874

# Calcular la curtosis
curtosis_m1_lognorm <- kurtosis(m1_lognorm)
cat("\nCurtosis de la muestra de tamaño 10,000:\n", curtosis_m1_lognorm)

##
## Curtosis de la muestra de tamaño 10,000:
## 29.22346

# Anderson-Darling para La muestra Log-Normal
ad_test_m1_lognorm <- ad.test(m1_lognorm)
ad_test_m1_lognorm

##
## Anderson-Darling normality test
##
## data: m1_lognorm
## A = 914.09, p-value < 2.2e-16

if (ad_test_m1_lognorm$p.value < 0.05) {
  conclusion <- "Los datos no siguen una distribución normal."
} else {
  conclusion <- "No se puede rechazar la hipótesis de que los datos
siguen una distribución normal."
}
conclusion

## [1] "Los datos no siguen una distribución normal."

# Jarque-Bera para La muestra Log-Normal
jb_test_m1_lognorm <- jarque.bera.test(m1_lognorm)
jb_test_m1_lognorm

##
## Jarque Bera Test
##
## data: m1_lognorm
## X-squared = 384830, df = 2, p-value < 2.2e-16

if (jb_test_m1_lognorm$p.value < 0.05) {
  conclusion <- "Los datos no siguen una distribución normal."
} else {
```

```

    conclusion <- "No se puede rechazar la hipótesis de que los datos
siguen una distribución normal."
}
conclusion

## [1] "Los datos no siguen una distribución normal."

```

**Interpretación de resultados** Los resultados indican que la muestra de tamaño 10,000 tomada de la distribución Log-Normal es altamente asimétrica y tiene colas extremadamente pesadas, tal como lo reflejan el sesgo y la curtosis extremadamente altos.

Las pruebas de normalidad confirman que la muestra no sigue una distribución normal. Esto es consistente con la naturaleza de la distribución Log-Normal con los parámetros utilizados, que generan una gran concentración de valores bajos y una pequeña proporción de valores extremadamente altos.

### c) Calculo de sesgo, curtosis y pruebas de normalidad para los promedios Log-Normal

```

# Calcular el sesgo
sesgo_m_lognorm <- skewness(datos_lognorm)
cat("\nSesgo de las medias de las 1,000 muestras:\n", sesgo_m_lognorm)

##
## Sesgo de las medias de las 1,000 muestras:
## 0.02241084

# Calcular la curtosis
curtosis_m_lognorm <- kurtosis(datos_lognorm)
cat("\nCurtosis de las medias de las 1,000 muestras:\n",
curtosis_m_lognorm)

##
## Curtosis de las medias de las 1,000 muestras:
## -0.1433867

# Anderson-Darling para Los promedios Log-Normal
ad_test_m_lognorm <- ad.test(datos_lognorm)
ad_test_m_lognorm

##
## Anderson-Darling normality test
##
## data:  datos_lognorm
## A = 0.19919, p-value = 0.8852

if (ad_test_m_lognorm$p.value < 0.05) {
  conclusion <- "Los datos no siguen una distribución normal."
} else {

```



```

    conclusion <- "No se puede rechazar la hipótesis de que los datos
siguen una distribución normal."
}
conclusion

## [1] "No se puede rechazar la hipótesis de que los datos siguen una
distribución normal."

# Jarque-Bera para Los promedios Log-Normal
j_b_test_m_lognorm <- jarque.bera.test(datos_lognorm)
j_b_test_m_lognorm

##
## Jarque Bera Test
##
## data: datos_lognorm
## X-squared = 0.87361, df = 2, p-value = 0.6461

if (j_b_test_m_lognorm$p.value < 0.05) {
  conclusion <- "Los datos no siguen una distribución normal."
} else {
  conclusion <- "No se puede rechazar la hipótesis de que los datos
siguen una distribución normal."
}
conclusion

## [1] "No se puede rechazar la hipótesis de que los datos siguen una
distribución normal."

```

**Interpretación de resultados** Tanto el sesgo cercano a 0 como la curtosis cercana a 3, junto con los resultados de las pruebas de normalidad, sugieren que la distribución de las medias de las 1,000 muestras es normal, pues la distribución de las medias de un número grande de muestras, independientemente de la forma de la distribución original, tiende a ser normal, justo como en el caso anterior.

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

### Semejanzas entre los Gráficos

- **Forma de las Distribuciones Teóricas** Las distribuciones teóricas de Weibull, Gamma y Log-Normal son todas asimétricas con colas largas hacia la derecha. Esto significa que en todas estas distribuciones, la mayor parte de los valores se encuentra cerca del extremo izquierdo de la gráfica, y solo unos pocos valores se extienden hacia la derecha.

- **Comportamiento de las Medias de las Muestras** En todas las distribuciones, los promedios de 1,000 muestras grandes (de tamaño 10,000) pareciera que se aproximan a una distribución normal. Los histogramas reflejan una forma de campana más simétrica y concentrada en torno a la media, pues las medias de muestras grandes de cualquier distribución tienden a distribuirse normalmente.

## Diferencias entre los Gráficos

### - Forma de las Distribuciones Teóricas

- **Weibull**: Presenta un pico más alto y una cola más suavemente decreciente en comparación con las otras dos distribuciones. La concentración de probabilidad es más cercana al pico, y la cola se extiende más lentamente.
- **Gamma**: La distribución Gamma tiene un pico menos pronunciado que la Weibull, pero todavía muestra una caída significativa después del pico, con una cola larga hacia la derecha.
- **Log-Normal**: La distribución Log-Normal muestra una cola mucho más larga y extendida hacia la derecha, con un pico más agudo que la distribución Gamma, pero con valores que pueden ser extremadamente grandes.

## Remaches

La resistencia a la ruptura de un remache tiene un valor medio de 10,000 lb/pulg 2 y una desviación estándar de 500 lb/pulg2. Si se sabe que la población se distribuye normalmente,

X: resistencia a la ruptura de un remache  $\mu_x = 10,000$  media  $\sigma_x = 500$  desviación estándar

$$X \sim N(\mu_x = 10000, \sigma_x = 500)$$

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?

$$P(9,900 < X < 10,100)$$

*# Determinar La probabilidad de que un remache tenga una resistencia a la ruptura X tal que esté a 100 unidades alrededor de la media*

```
p1 = pnorm(10100, 10000, 500) - pnorm(9900, 10000, 500)
cat("P(9900 < x < 10100)=", p1)
```

```
## P(9900 < x < 10100)= 0.1585194
```

*# Desviaciones estándar lejos de la media*

```
z_1 = (10100 - 10000) / 500
z_1
```

```
## [1] 0.2
```

**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?**

$P(9,900 < \bar{X} < 10,100)$   $n = 120$   $\sigma_{\bar{x}} = \frac{500}{\sqrt{120}}$

*# Encontrar La probabilidad de que La media muestral ((9,900 < \bar{X} < 10,100)) esté en el intervalo de 100 unidades alrededor de La media poblacional:*

```
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
```

*# Desviaciones estándar Lejos de La media*

```
z_2= 100/(500/sqrt(120))
z_2
```

```
## [1] 2.19089
```

**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?**

$n = 15$   $\sigma_{\bar{x}}$  tamaño de la muestra

*# Calcular La probabilidad de que La media muestral esté en el intervalo [9900, 10100]*

```
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
```

*# Desviaciones estándar Lejos de La media*

```
z_3= 100/(500/sqrt(15))
z_3
```

```
## [1] 0.7745967
```

**d) Un ingeniero recibió un lote muy grande de remaches. Antes de aceptarlo verificó si efectivamente tiene una media de 10,000 lb/pulg 2. 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?. Si la media hubiera sido 9925, ¿recomendarías rechazarlo?**

$h_0: \mu = 10,000$   $h_1: \mu < 10,000$

```
mu_0 <- 10000      # Media
sigma <- 500        # Desviación estándar
n <- 120            # Tamaño de La muestra
x_bar_1 <- 9800     # Media observada 1
```

$$Z = \frac{\bar{X} - \mu}{\sigma}$$

```
# Caso 1: Media observada = 9800
z_4 <- (x_bar_1 - mu_0) / (sigma / sqrt(n))
z_4

## [1] -4.38178

p4 <- pnorm(z_4)
p4

## [1] 5.88567e-06
```

Se rechaza  $h_0$  dado que el valor de la probabilidad asociada es muy baja. El ingeniero hizo lo correcto al rechazar el lote, ya que hay evidencia estadística significativa de que la media del lote no es 10,000 lb/pulg<sup>2</sup>, sino significativamente menor.

```
# Caso 2: Media observada = 9925
x_bar_2 <- 9925

# Calcular el estadístico z
z_5 <- (x_bar_2 - mu_0) / (sigma / sqrt(n))
z_5

## [1] -1.643168

p5 <- pnorm(z_5)
p5

## [1] 0.05017412
```

No se rechaza  $h_0$  dado que la probabilidad asociada está cerca del nivel de significancia típico (0.05), no hay suficiente evidencia para rechazar la hipótesis nula, es decir, no se recomienda rechazar el lote.

### 3. Embotellando

Una máquina embotelladora puede ser regulada para que se descargue un promedio de  $\mu$  onzas por botella. Se ha observado que la cantidad de líquido dosificado por una máquina embotelladora está distribuida normalmente con  $\sigma = 1$  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  $\mu$  sea de 15 onzas.

$$\sigma = 1$$

$\mu(\text{texto deseada}) = 15$

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

Para una distribución normal estándar, el 95% central de la distribución se encuentra entre los percentiles 2.5% y 97.5%.

```
# Encontrar Los valores de Z que encierran el 95% de La probabilidad.

z_6 <- qnorm(0.975) # 95% central, por lo tanto, 2.5% en cada cola

cat("La media de la muestra puede estar a", z_6, "desviaciones estándar
de la media para estar dentro del 95% central.\n")

## La media de la muestra puede estar a 1.959964 desviaciones estándar de
la media para estar dentro del 95% central.
```

f) Hacer una gráfica del inciso 1.

```
# Generar valores de Z para La gráfica
z_values <- seq(-4, 4, length = 1000)
y_values <- dnorm(z_values, 0, 1)

# Crear La gráfica
plot(z_values, y_values, type = "l", lwd = 2, col = "blue",
     xlab = "Z (Desviaciones estándar)", ylab = "Densidad",
     main = "95% Central de la Distribución Normal Estándar")

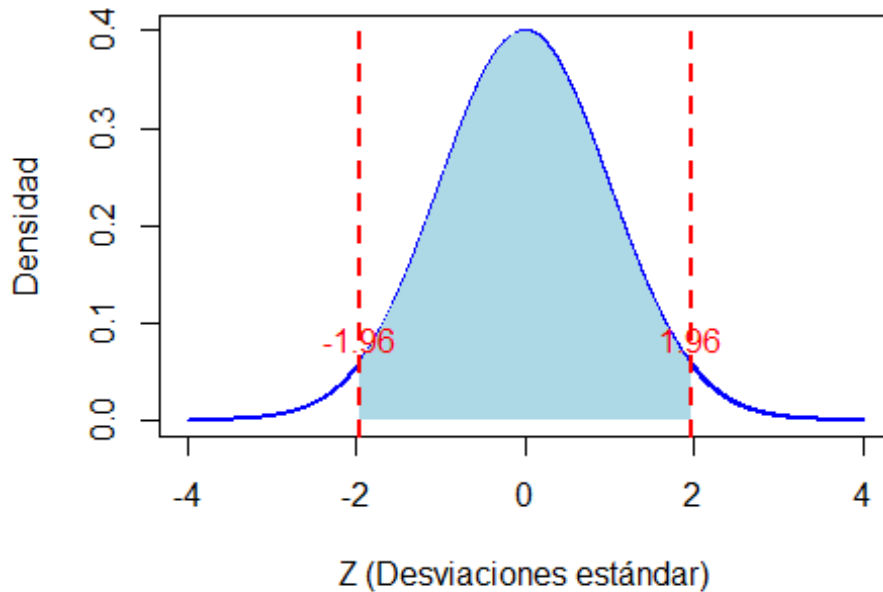
# Rellenar el área del 95% central
polygon(c(-z_6, seq(-z_6, z_6, length = 1000), z_6),
       c(0, dnorm(seq(-z_6, z_6, length = 1000)), 0),
       col = "lightblue", border = NA)

# Añadir líneas para Los valores críticos
```

```
abline(v = c(-z_6, z_6), col = "red", lwd = 2, lty = 2)
```

```
text(-z_6, 0.05, paste0("-1.96"), col = "red", pos = 3)  
text(z_6, 0.05, paste0("1.96"), col = "red", pos = 3)
```

### 95% Central de la Distribución Normal Estándar



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

Calcular la probabilidad de que la media muestral de una muestra de 10 botellas sea mayor a 16 onzas.

```
mu <- 15  
sigma <- 1  
n <- 10  
sigma_xbar <- sigma / sqrt(n) # Desviación estándar de la media muestral  
x_bar <- 16  
  
# Calcular el valor de Z para X_bar = 16  
z_7 <- (x_bar - mu) / sigma_xbar  
z_7  
  
## [1] 3.162278  
  
# Calcular la probabilidad de que la media muestral sea mayor a 16 onzas  
p_greater_16 <- 1 - pnorm(z_7)
```

```
cat("La probabilidad de que la media muestral sea mayor a 16 onzas es:",
p_greater_16, "\n")

## La probabilidad de que la media muestral sea mayor a 16 onzas es:
0.0007827011
```

**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?**

( $H_0$ ): La máquina está calibrada correctamente, es decir, la media de llenado es  $\mu = 15$  onzas.

( $H_1$ ): La máquina no está calibrada correctamente, es decir, la media de llenado es diferente de 15 onzas ( $\mu \neq 15$ ).

```
# Calcular el valor de Z
z_8 <- (x_bar - mu) / sigma_xbar
z_8

## [1] 3.162278

# Valor crítico para un nivel de significancia alfa = 0.05 (prueba
bilateral)
z_9 <- qnorm(0.975) # 1.96 aproximadamente
z_9

## [1] 1.959964

# Decisión basada en el valor de Z
if (abs(z_8) > z_9) {
  cat("Rechazamos H0. Se recomienda detener la producción para calibrar
la máquina.\n")
} else {
  cat("No se rechaza H0. No es necesario detener la producción.\n")
}

## Rechazamos H0. Se recomienda detener la producción para calibrar la
máquina.
```

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

calcular la probabilidad de que la media muestral de una muestra de 10 botellas sea menor a 14.5 onzas.

```
# Parámetros
mu <- 15 # Media poblacional
sigma <- 1 # Desviación estándar de La población
n <- 10 # Tamaño de La muestra
x_bar <- 14.5 # Media observada de La muestra
sigma_xbar <- sigma / sqrt(n) # Desviación estándar de La media muestral
```

```

# Calcular el valor de Z
z_10 <- (x_bar - mu) / sigma_xbar
z_10

## [1] -1.581139

# Calcular la probabilidad de que la media muestral sea menor a 14.5 onzas
p_less_14_5 <- pnorm(z_10)

cat("La probabilidad de que la media muestral sea menor a 14.5 onzas es:", p_less_14_5, "\n")

## La probabilidad de que la media muestral sea menor a 14.5 onzas es:
0.05692315

```

**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?**

```

# Parámetros
mu <- 15          # Media poblacional hipotética
sigma <- 1        # Desviación estándar de la población
n <- 10           # Tamaño de la muestra
x_bar <- 15.5     # Media observada de la muestra
sigma_xbar <- sigma / sqrt(n) # Desviación estándar de la media muestral

# Calcular el valor de Z
z_11 <- (x_bar - mu) / sigma_xbar
z_11

## [1] 1.581139

# Valor crítico para un nivel de significancia alfa = 0.05 (prueba bilateral)
z_12 <- qnorm(0.975) # 1.96 aproximadamente
z_12

## [1] 1.959964

# Decisión basada en el valor de Z
if (abs(z_11) > z_12) {
  cat("Rechazamos H0. Se recomienda detener la producción para calibrar la máquina.\n")
} else {
  cat("No se rechaza H0. No es necesario detener la producción.\n")
}

## No se rechaza H0. No es necesario detener la producción.

```