Actividad 6

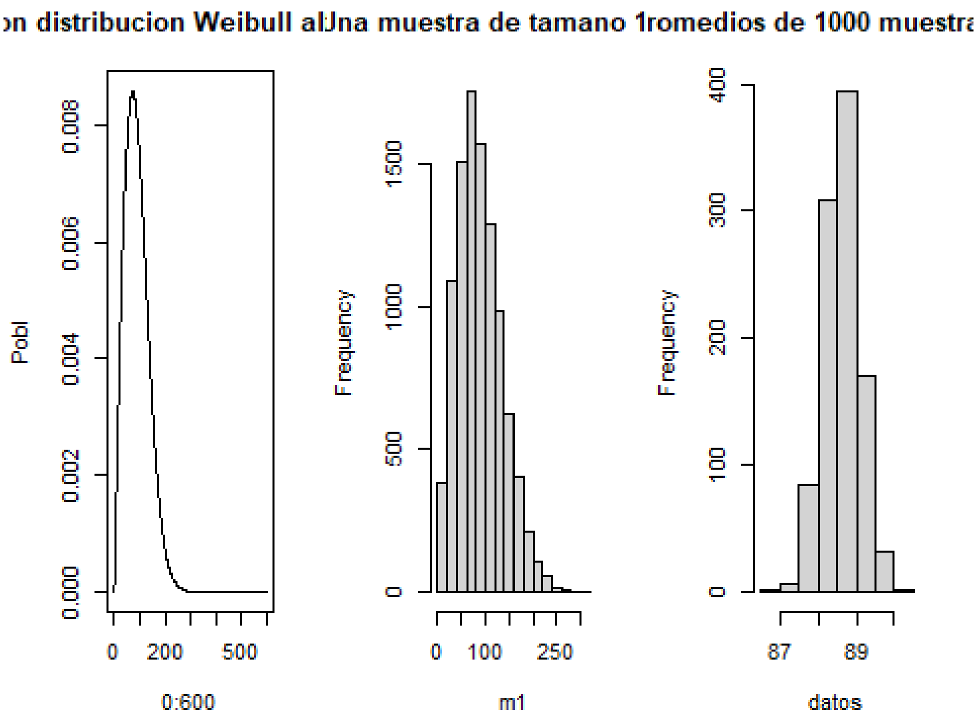
Oskar Arturo Gamboa Reyes

2024-08-16

1. Ensayando Distribuciones

Inciso A.

```
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")
```



Inciso B.

Calcular Sesgo y Curtosis

```
library(e1071)
cat("Curtosis de 10000 muestras =", kurtosis(m1))

## Curtosis de 10000 muestras = 0.1122584

cat("Sesgo de 10000 muestras =", skewness(m1))

## Sesgo de 10000 muestras = 0.6084586
```

Prueba de normalidad

```
library(nortest)
ad.test(m1)

##
##  Anderson-Darling normality test
##
## data:  m1
## A = 59.59, p-value < 2.2e-16
```

Los datos muestra no tienen una distribución normal ya que el p-value es demasiado pequeño y tiene demasiado sesgo.

Inciso C.

Calcular Sesgo y Curtosis

```r
library(e1071)
cat("Curtosis de 10000 muestras =", kurtosis(datos))

## Curtosis de 10000 muestras = 0.1248902

cat("Sesgo de 10000 muestras =", skewness(datos))

## Sesgo de 10000 muestras = -0.0157849
```

Prueba de normalidad

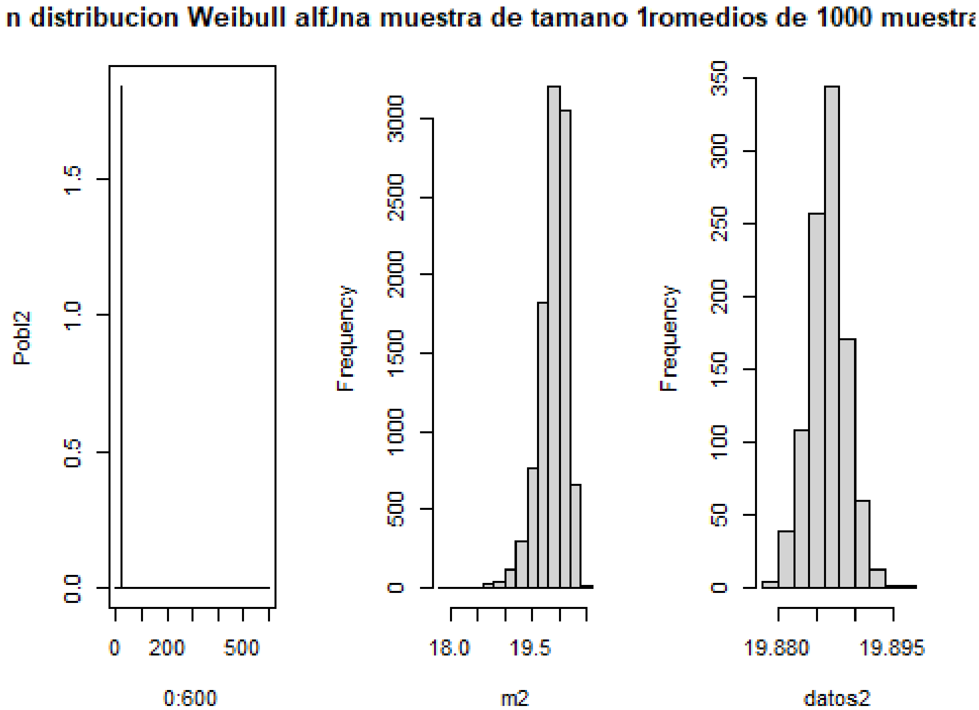
```r
library(nortest)
ad.test(datos)

##
##  Anderson-Darling normality test
##
## data:  datos
## A = 0.16485, p-value = 0.9416
```

Estos datos ya tienen un distribución normal, el sesgo es suficientemente bajo y el p-value es mayor a 0.05.

Inciso D.

Prueba 1

```r
par(mfrow=c(1,3))
# Graficando una distribucion Weibull de alfa =100, beta = 20
Pobl2 = dweibull(0:600,100, 20)
plot(0:600,Pobl2, type="l", main = "Poblacion con distribucion Weibull alfa
=100, beta = 20")

# Tomando una muestra de 10000 elementos tomados al azar
m2 = rweibull(10000, 100, 20)
hist(m2, main = "Una muestra de tamano 10000")

# Tomando 1000 promedios de las 1000 muestras como la anterior
    m =rweibull(10000,100,20)
    prom=mean(m)
    datos2=prom
for(i in 1:999) {
    m =rweibull(10000,100,20)
    prom=mean(m)
    datos2=rbind(datos2,prom)  }
hist(datos2, main="Grafica de los promedios de 1000 muestras de tamano
10,000")
```



Calcular Sesgo y Curtosis muestra 10000

```
library(e1071)
cat("Curtosis de 10000 muestras =", kurtosis(m2))

## Curtosis de 10000 muestras = 2.503802

cat("Sesgo de 10000 muestras =", skewness(m2))

## Sesgo de 10000 muestras = -1.136948
```

Prueba de normalidad promedio muestra 10000

```
library(nortest)
ad.test(m2)

##
##  Anderson-Darling normality test
##
## data:  m2
## A = 117.55, p-value < 2.2e-16
```

Calcular Sesgo y Curtosis muestras

```
library(e1071)
cat("Curtosis de 10000 muestras =", kurtosis(datos2))

## Curtosis de 10000 muestras = 0.2213299
```

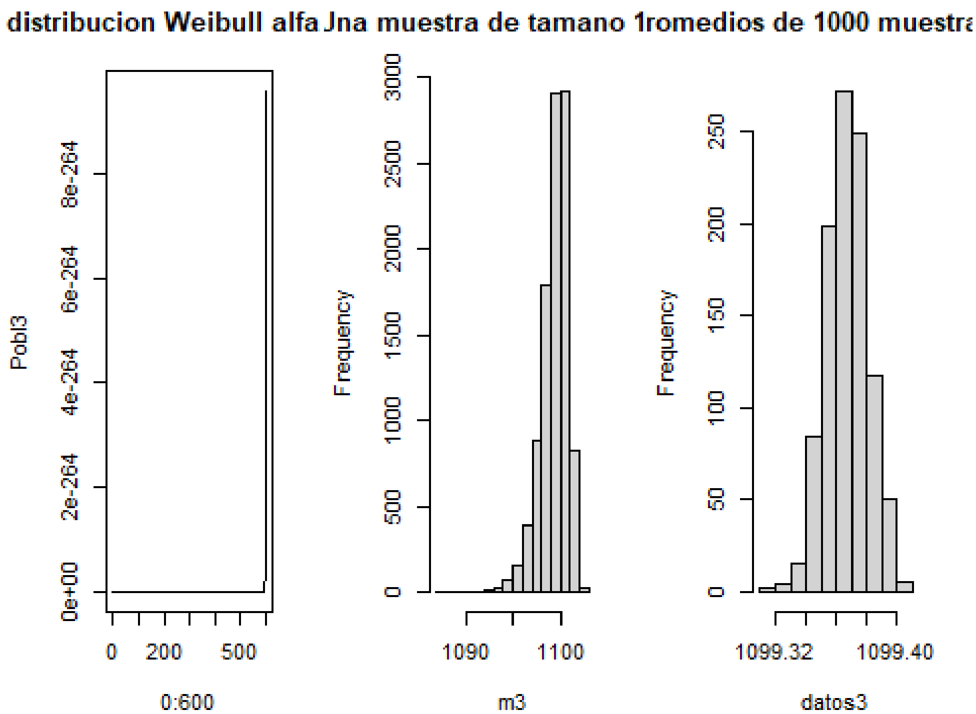
```
cat("Sesgo de 10000 muestras =", skewness(datos2))

## Sesgo de 10000 muestras = 0.0597257
```

Prueba de normalidad muestras

```
library(nortest)
ad.test(datos2)

##
##  Anderson-Darling normality test
##
## data:  datos2
## A = 0.61051, p-value = 0.1122
```

Prueba 2

```
par(mfrow=c(1,3))
# Graficando una distribucion Weibull de alfa =1000, beta = 1100
Pobl3 = dweibull(0:600,1000, 1100)
plot(0:600,Pobl3, type="l", main = "Poblacion con distribucion Weibull alfa
=1000, beta = 1100")

# Tomando una muestra de 10000 elementos tomados al azar
m3 = rweibull(10000, 1000, 1100)
hist(m3, main = "Una muestra de tamano 10000")

# Tomando 1000 promedios de las 1000 muestras como la anterior
    m =rweibull(10000,1000,1100)
    prom=mean(m)
    datos3=prom
for(i in 1:999) {
    m =rweibull(10000,1000,1100)
    prom=mean(m)
    datos3=rbind(datos3,prom)   }
hist(datos3, main="Grafica de los promedios de 1000 muestras de tamano
10,000")
```



Calcular Sesgo y Curtosis muestra 10000

```
library(e1071)
cat("Curtosis de 10000 muestras =", kurtosis(m3))

## Curtosis de 10000 muestras = 2.660203

cat("Sesgo de 10000 muestras =", skewness(m3))

## Sesgo de 10000 muestras = -1.156701
```

Prueba de normalidad promedio muestra 10000

```
library(nortest)
ad.test(m3)

##
##  Anderson-Darling normality test
##
## data:  m3
## A = 122.67, p-value < 2.2e-16
```

Calcular Sesgo y Curtosis muestras

```
library(e1071)
cat("Curtosis de 10000 muestras =", kurtosis(datos3))

## Curtosis de 10000 muestras = 0.04329233
```

```
cat("Sesgo de 10000 muestras =", skewness(datos3))

## Sesgo de 10000 muestras = -0.04409344
```

Prueba de normalidad muestras

```
library(nortest)
ad.test(datos3)

##
##  Anderson-Darling normality test
##
## data:  datos3
## A = 0.45552, p-value = 0.2672
```

Las 10000 muestras siguen una distribución parecida a la población total, por lo que no tiene una distribución normal. Mientras que la media de las muestras si sigue una distribución normal.

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: Resistencia a la repturra de un remache

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

Inciso A.

¿Cuál es la probabilidad de que la 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 < X < 10100)

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

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

Desviaciones lejos de la media (z)

```
z = 100/500
cat("z =", z)

## z = 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?

$$P(9900 < \bar{X} < 10100)$$

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

```
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 lejos de la media (z)

```
z2 = 100/(500/sqrt(120))
cat("z =", z2)

## z = 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?

$$P(9900 < \bar{X} < 10100)$$

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

```
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 lejos de la media (z)

```
z3 = 100/(500/sqrt(15))
cat("z =", z3)

## z = 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?.

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

```
## P(9800) = 5.88567e-06
```

```r
z4 = 200/(500/sqrt(120))
cat("z =", z4)
```

```
## z = 4.38178
```

Hizo lo correcto, la probabilidad de que una muestra de 120 remaches tenga una resistencia de 9800 es demasiado baja, además la desviación es demasiada alta lo que indica que son casos muy extremos.

Inciso E.

¿Qué decisión recomiendas al ingeniero si la media obtenida en la media hubiera sido 9925? ¿recomendarías rechazarlo?

```r
p5 = pnorm(9925, 10000, 500/sqrt(120))
cat("P(9800) =", p5)
```

```
## P(9800) = 0.05017412
```

```r
z5 = (10000-9925)/(500/sqrt(120))
cat("z =", z5)
```

```
## z = 1.643168
```

No recomendaría comprarlo, al igual que el inciso anterior las probabilidades son muy bajas y la desviación es muy alta, por lo que debe haber algo mal con la producción de estos remaches.

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.

X: Onzas por botella

Inciso 1.

¿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?

```r
Ndesv = qnorm(0.975,0,1)
cat("Número de desviaciones estandar =", Ndesv)

## Número de desviaciones estandar = 1.959964
```

Inciso 2.

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

```r
p6 = pnorm(16, 15, 1/sqrt(10))
cat("P(>16) =", 1-p6)

## P(>16) = 0.0007827011
```

Inciso 3.

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

Si se detendria ya que la probabilidad es menor que la distribución central.

```r
z6 = (1)/(1/sqrt(10))
cat("z =", z6)

## z = 3.162278
```

Inciso 4.

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

```r
p7 = pnorm(14.5, 15, 1/sqrt(10))
cat("P(<14.5) =", p7)

## P(<14.5) = 0.05692315
```

Inciso 5.

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?

```r
p8 = pnorm(15.5, 15, 1/sqrt(10))
cat("P(>15.5) =", 1-p8)

## P(>15.5) = 0.05692315

z7 = (0.5)/(1/sqrt(10))
cat("z =", z7)

## z = 1.581139
```

No se detiene porque queda dentro de la distribución central.

Inciso 6.

Hacer una gráfica del inciso 1.

```
media = 0
desviacion = 1
x <- seq(media - 4*desviacion, media + 4*desviacion, length=1000)
y <- dnorm(x, mean=media, sd=desviacion)
plot(x, y, type="l", lwd=2, col="blue", main="Distribución Normal",
     xlab="X", ylab="Densidad")


abline(v=media + 1.96*desviacion, col="red", lwd=2, lty=2)
abline(v=media - 1.96*desviacion, col="red", lwd=2, lty=2)
```