Tarea 5 Inferencia Estadística

Code

Hairo Ulises Miranda Belmonte

6 de Noviembre de 2018

EJERCICIO 1

Ejercicios de las Notas

EJERCICIO 1.1 Diap. 283

Calcula las probabilidades mencionadas arriba.

Code

Se observa que la probabilidad de que aparezca el valor de 80.3 u otro, es muy baja con varianza de .1

Code

```
## [1] 0.1713909
```

La probabilidad de que aparezca el valor de 80.3 u otro es muy baja con varianza de 0.20 no se ve tan imposible

Code

```
## [1] 0.2887343
```

caso 1 varianza de 0.1

Code

```
## [1] TRUE
```

Con un nvel de significancia del 95% se tiene evidencia suficiente de rechazar Ho

Code

```
## [1] TRUE
```

Utilizando el p-valor, se tiene evidencia suficiente para rechazar la H_0

Caso 2 varianza de .29

Code

```
## [1] TRUE
```

A un nvel de significancia del 95 se tiene evidencia suficiente de rechazar H_0. Sin embargo, no es muy contundente la diferencia del valor z observado.

Ahora, con un nivel de confiabilidad mayor, se tiene el siguiente resultado:

Code

```
## [1] FALSE
```

Al 97.5 de significancia, no se tiene evidencia suficiente para rechazar la H_0. Utilizando el p-valor:

Code

```
## [1] TRUE
```

Como se puede observar, se tiene evidencia suficiente para rechazar la H_0. Sin embargo si no se es muy estricto, se podría no rechazar la hipótesis nula. Con menor nivel de confiabilidad, se tiene:

Code

```
## [1] TRUE
```

Con .1 de insignificancia ya se tiene evidencia suficiente para no rechazar la hipótesis nula.

Construcción de intervalos de confianza; caso 1:

Code

Intervalo de confianza para el caso 2:

Code

Por lo tanto, el primer intervalo no contiene al valor del parámetro (bajo la hipótesis nula) y el segundo intervalo de confianza sí lo contiene.

EJERCICIO 1.2 Diap. 346

Construir esas gráficas para este caso particular:

Refresquemos las hip□otesis del problema: $H_0 : \mu \geq 35$ vs $H_A : \mu < 35$ y hagamos la prueba al nivel de 10%.

Code

Prueba de hipótesis:

Code

```
## [1] TRUE
```

Tenemos evidencia suficiente para reazar la nule aun nivel del 10%

Con $\alpha = 0.05$

Code

```
## [1] FALSE
```

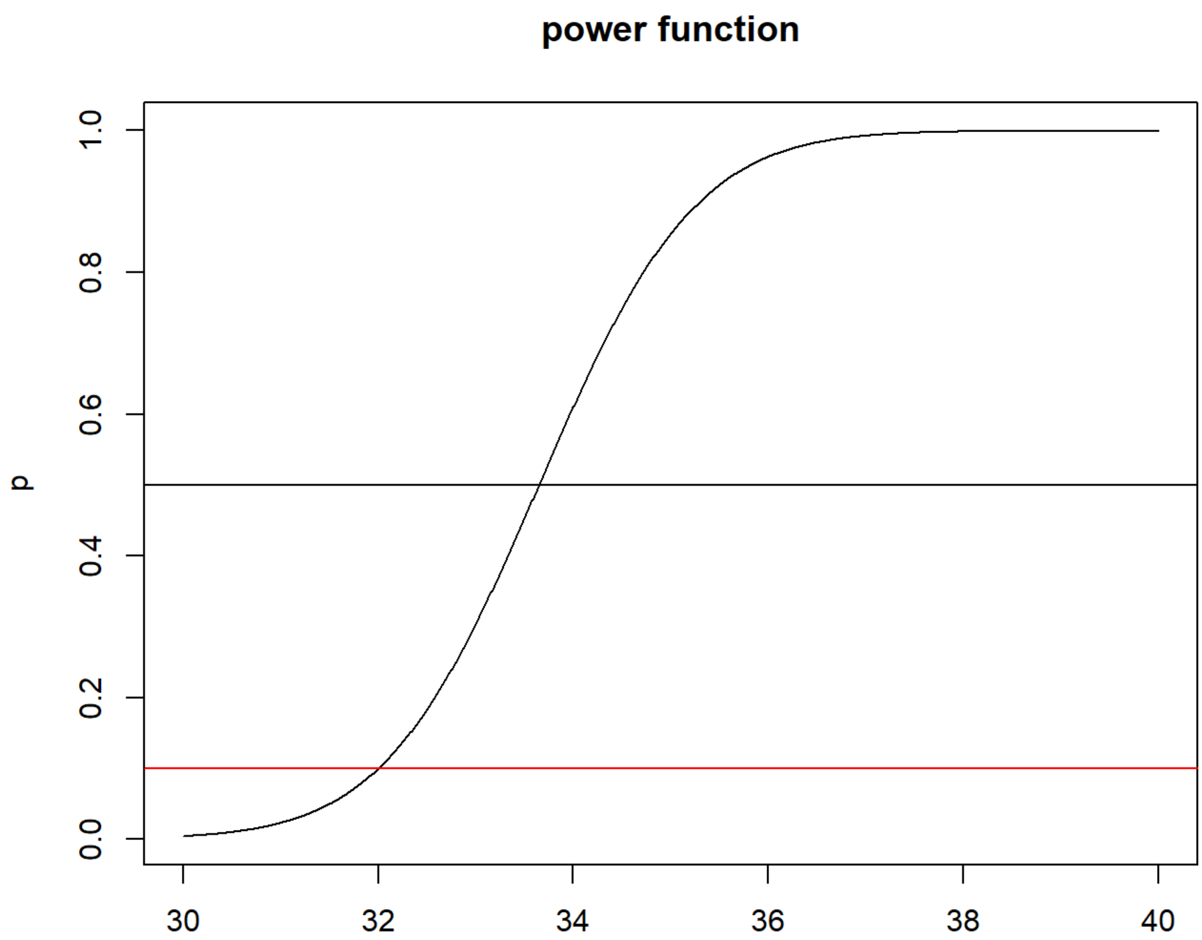
No se puede rechazar la hipótesis nula al 0.05 de insignificancia.

Potencia de los valores

Code

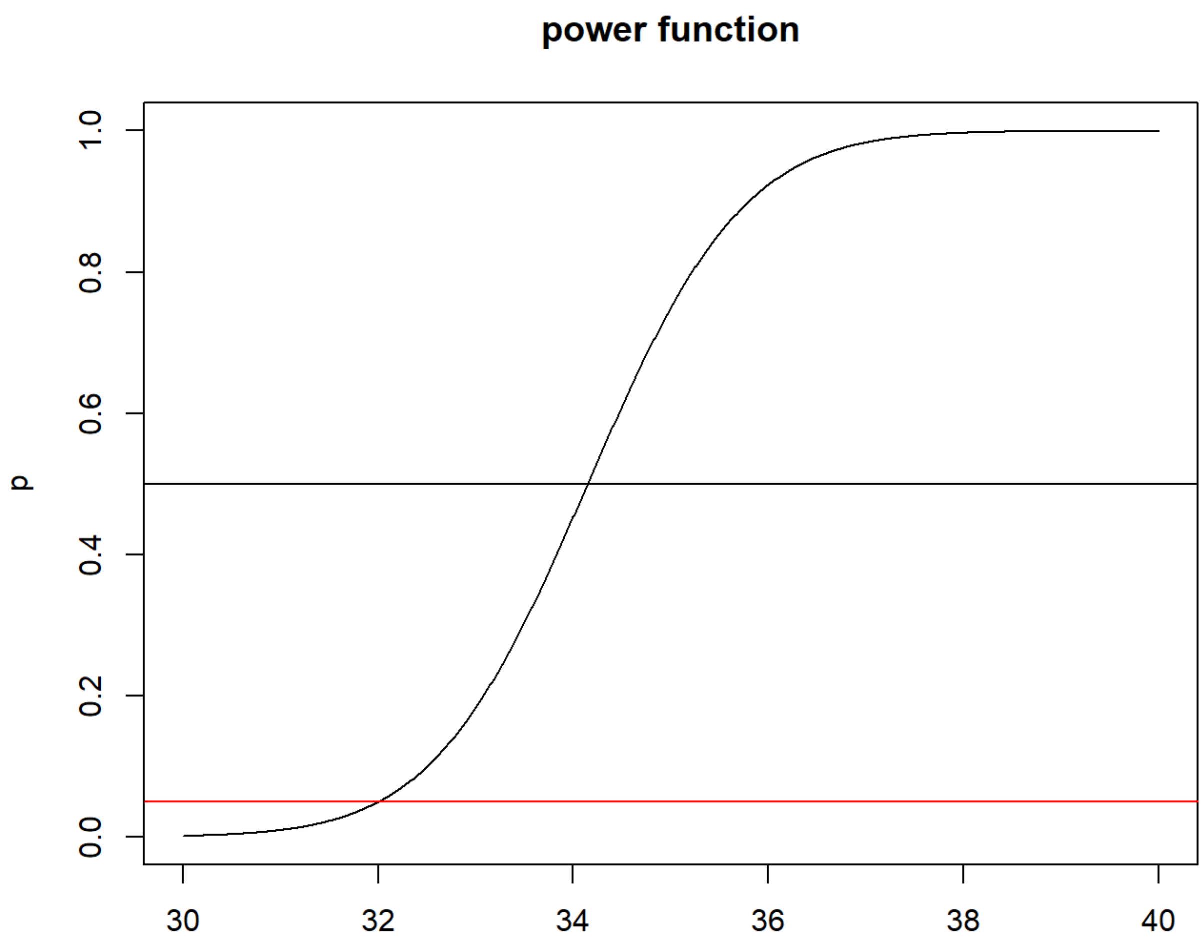
Potencia de la prueba con $\alpha = 0.1$

Code



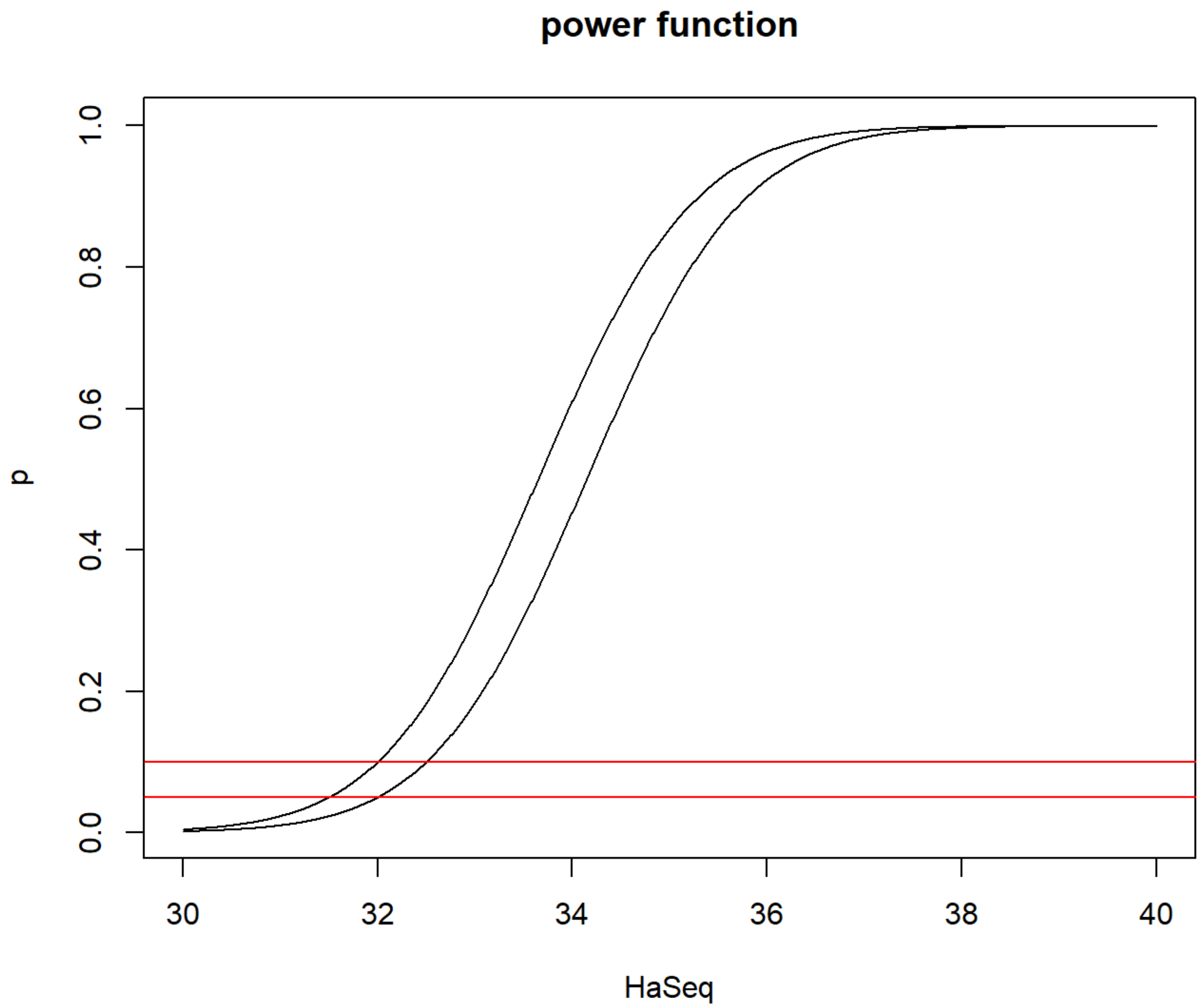
Potencia de la prueba con $\alpha = 0.05$

Code



Comparando funciones poder respecto a variaciones en el nivel de significancia.

Code



EJERCICIO 1.3 Diap. 352

Ejercicio: Realiza una prueba de nivel $\alpha = 0.01$ para las siguientes hipotesis: $_0 : \sigma^2 = 8$ vs $H_A : \sigma^2 < 8$ a partir del archivo datvar.txt.

a. Verifica inicialmente normalidad y concluye sobre la plausibilidad de lamisma en este conjunto de datos.

Code

PASO 1. REVISAR QUE LA MUESTRA PROVENGA DE UNA NORMAL

Code

```
##
##  Shapiro-Wilk normality test
##
## data:  datavar$V2
## W = 0.94952, p-value = 0.2087
```

Se observa el estadistico se acerca a uno, de esta manera, se puede decir que la muestra proviene de una normal.

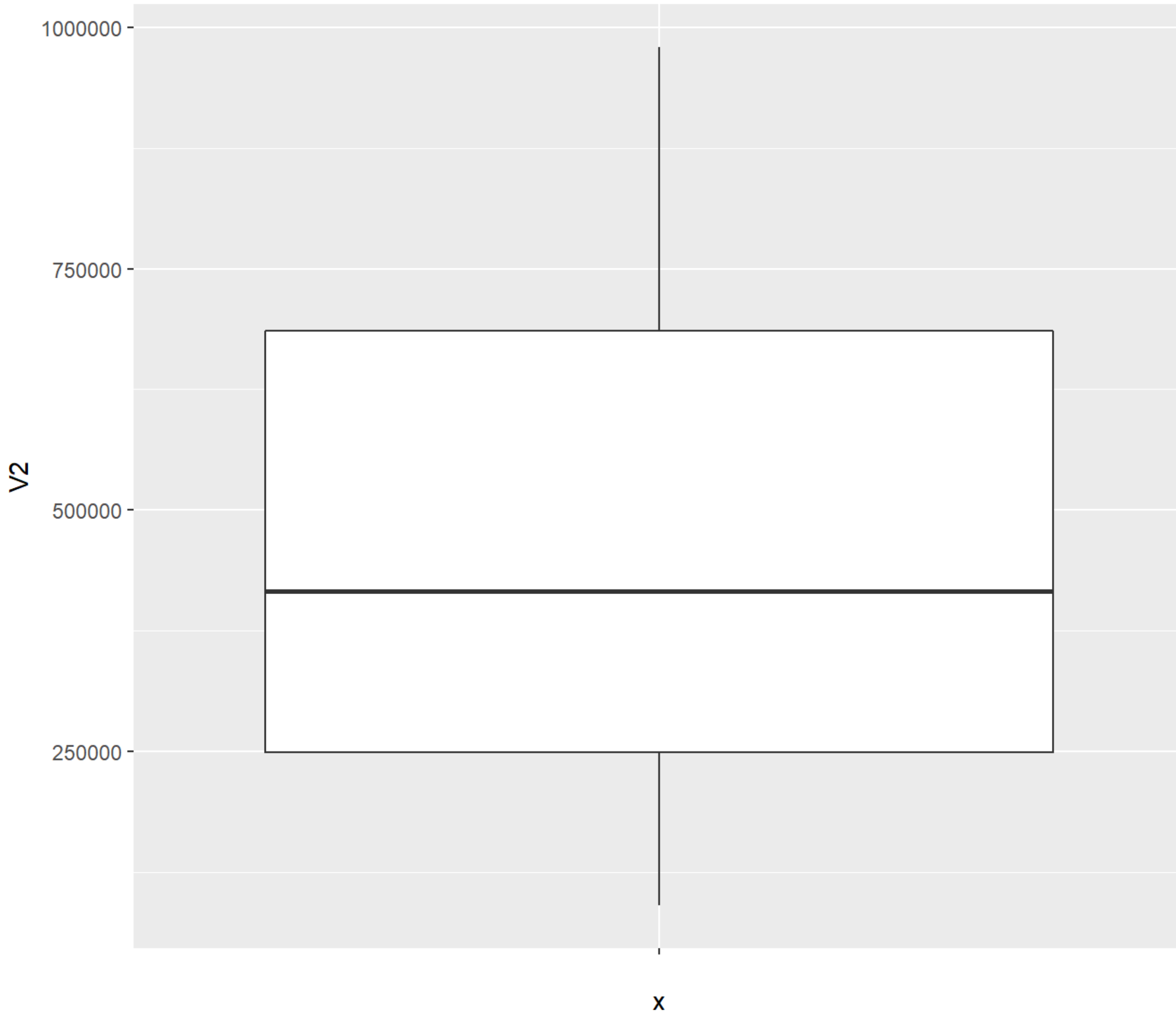
Contrastando contra la prueba de jarquer bera en normalidad sobre la hipótesis nula.

Code

```
##
## Title:
##  Jarque - Bera Normalality Test
##
## Test Results:
##   STATISTIC:
##     X-squared: 1.5894
##   P VALUE:
##     Asymptotic p Value: 0.4517
##
## Description:
##  Mon Nov 26 16:24:47 2018 by user: h_air
```

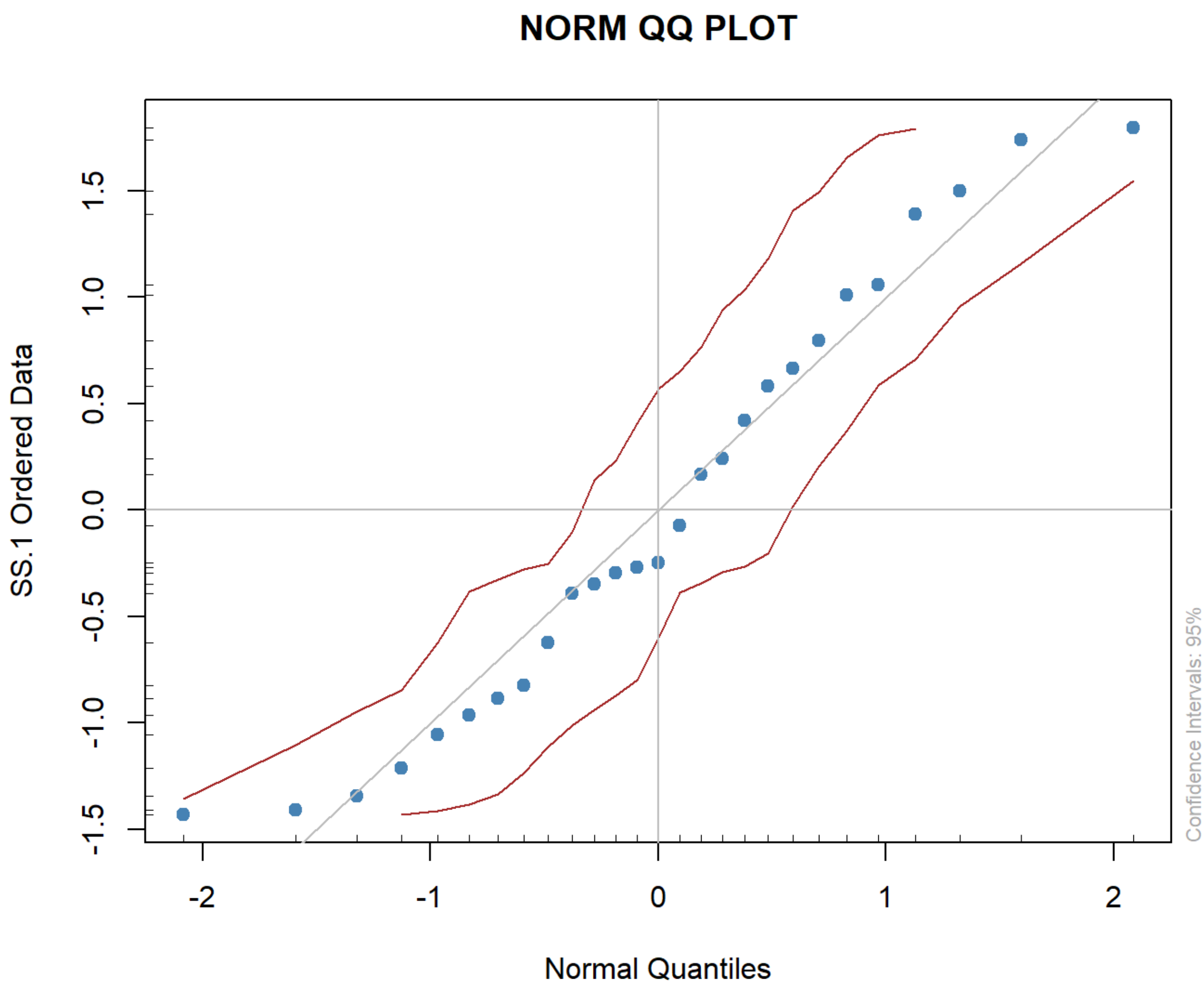
Ho Normalidad, No se tiene evidencia suficiente para rechazar la nula

Code

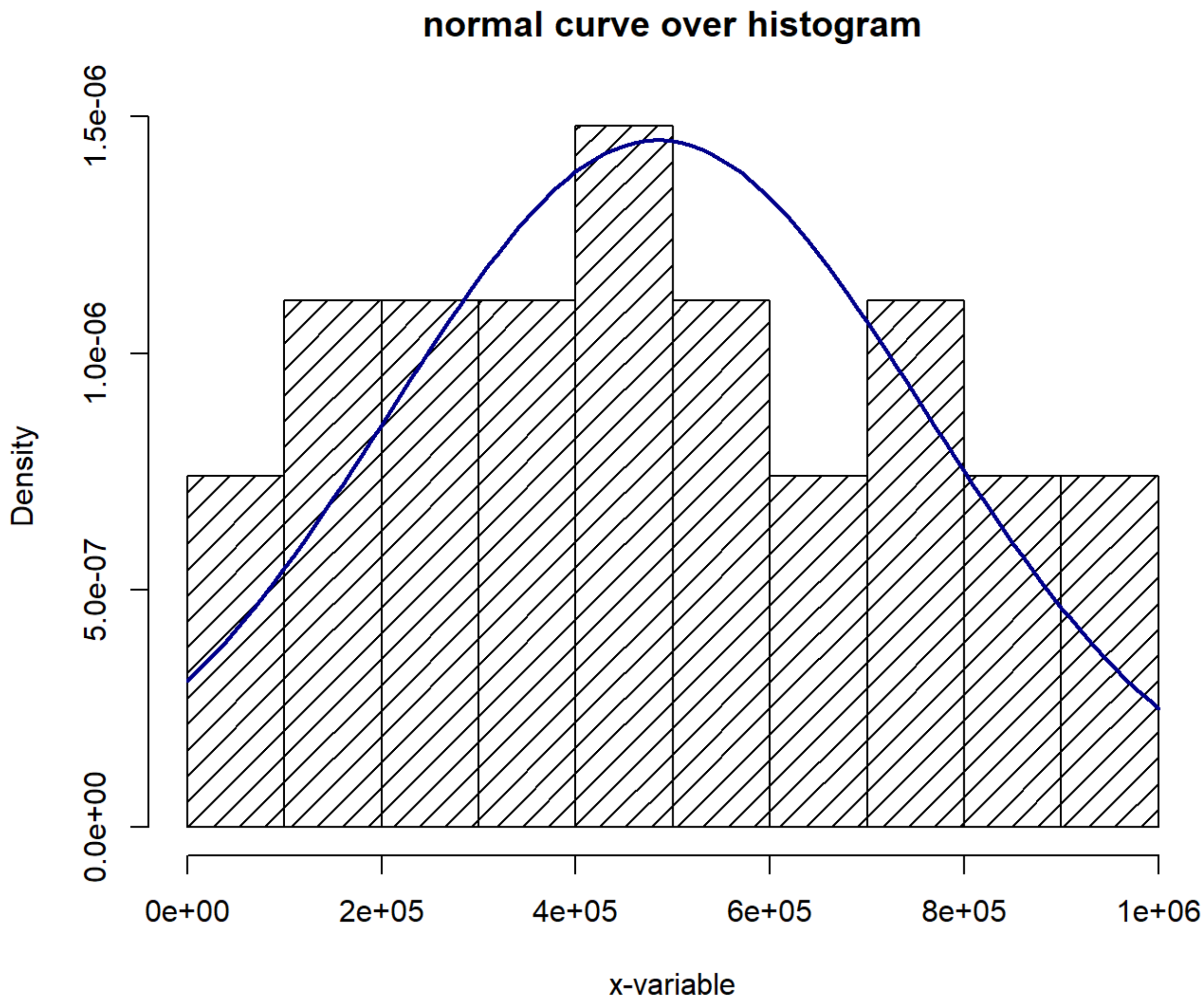


Se observa un poco de sesgo, cargada un poco a la derecha. También se puede ver, como su valor medio se encuentra un poco cargado.

Code


NORM QQ PLOT
SS.1 Ordered Data
Normal Quantiles
Confidence Intervals: 95%
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
-2
-1
0
1
2

Code



Los datos tratan de aproximar a la normal. Sin embargo tanto en colas, y ligeramente en el centro, no lo logran en su totalidad. Se concluye que se trabaja con el supuesto de normalidad. Siendo así, conocemos la distribución del estadístico de prueba. A su vez, podemos fijar regiones de rechazo y calcular p-valores dada las hipótesis que estemos considerando

Para el siguiente set de datos se realiza lo anterior.

Code

```
##
##  Shapiro-Wilk normality test
##
## data:  datavar$V1
## W = 0.81735, p-value = 0.0002776
```

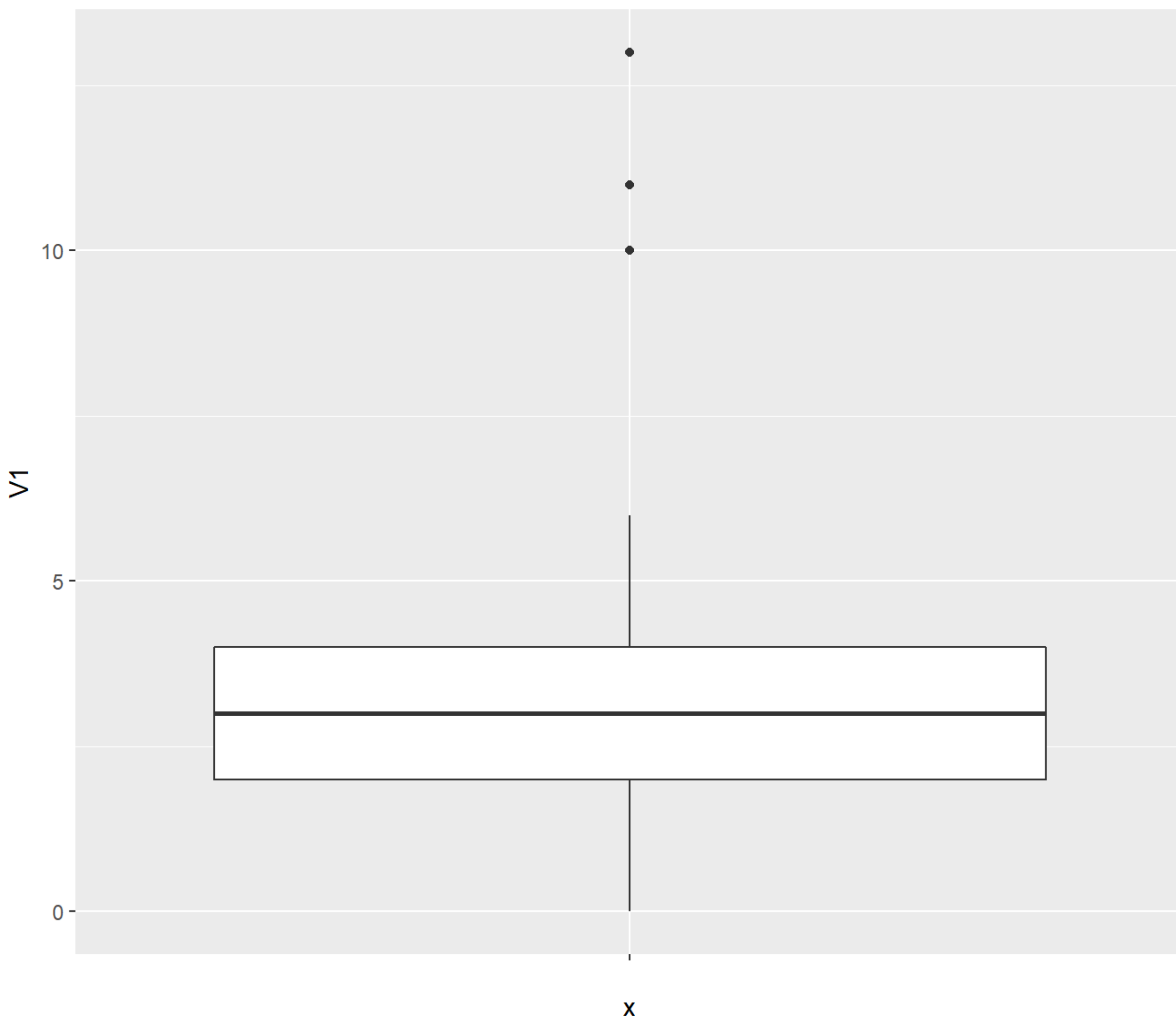
p value muy pequeño, pero puede deberce al tamaño de muestra el estadístico es cercano a uno, no estaría rechazando la nula. Sin embargo, no se tiene aún evidencia suficiente.

Code

```
##
## Title:
##  Jarque - Bera Normalality Test
##
## Test Results:
##   STATISTIC:
##     X-squared: 15.2634
##   P VALUE:
##     Asymptotic p Value: 0.0004848
##
## Description:
##  Mon Nov 26 16:24:49 2018 by user: h_air
```

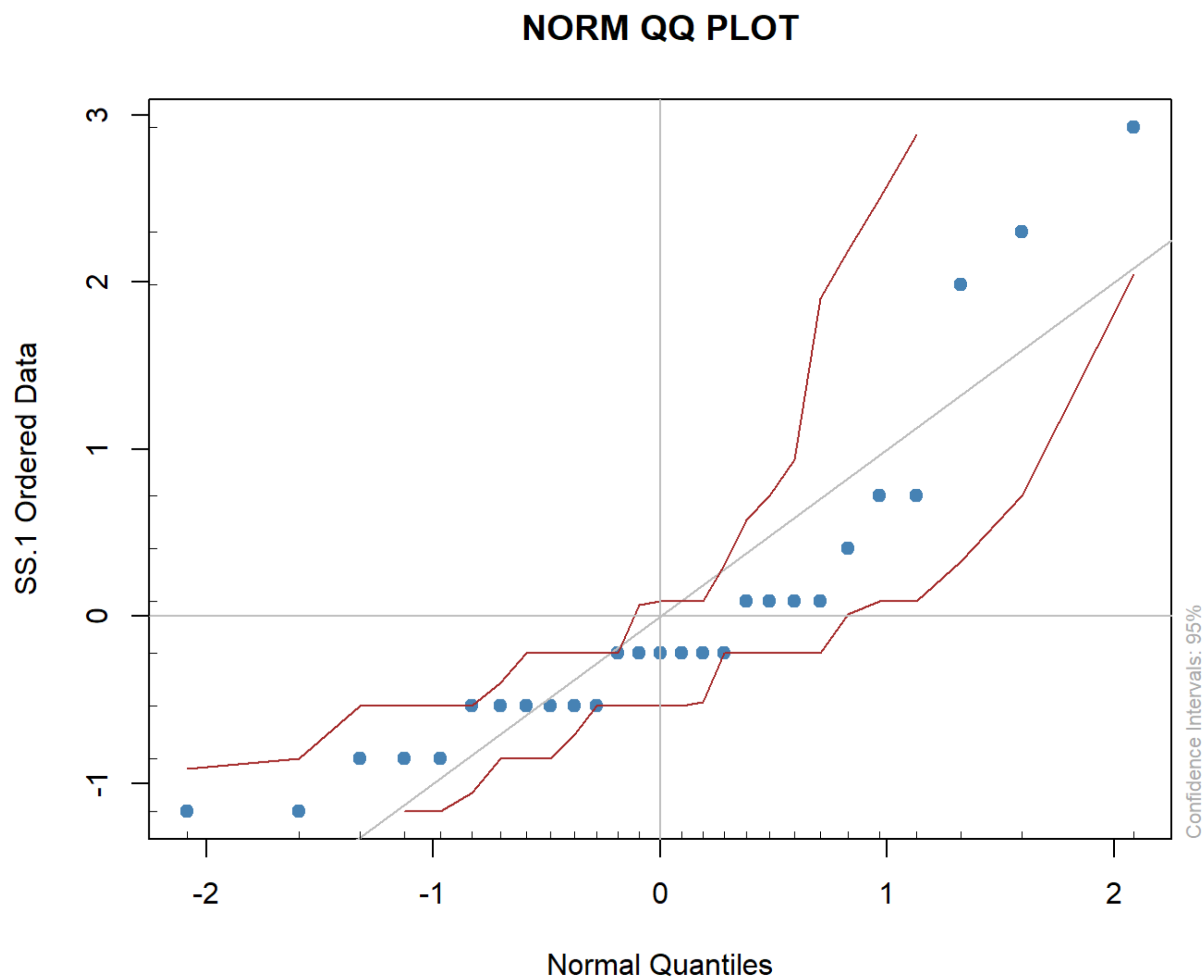
H_0 Normalidad, No se tiene evidencia suficiente para rechazar la H_0

Code

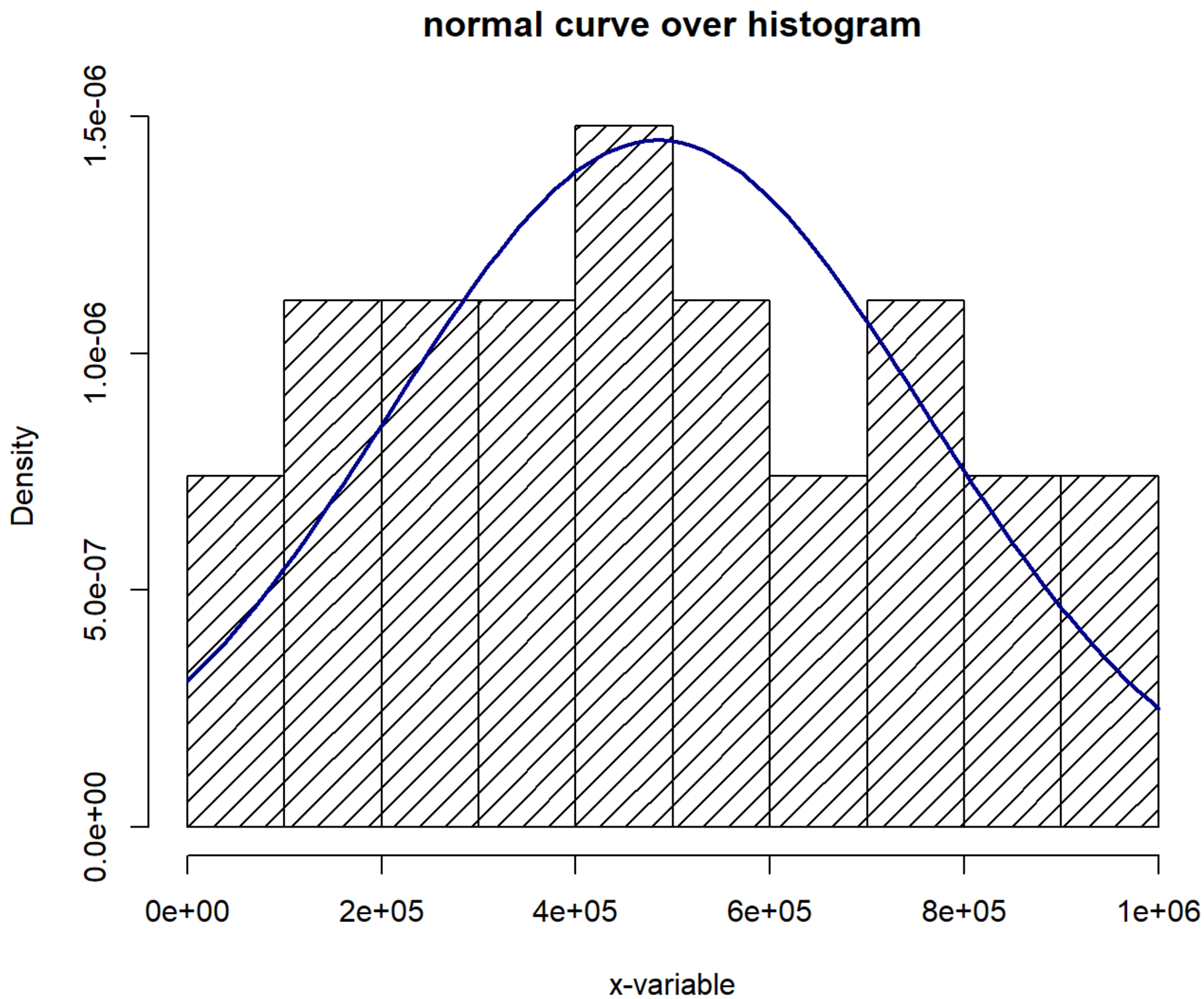


Se observa un poco de sesgo, y varios valores a tipicos, con una ediana centrada

Code


NORM QQ PLOT
SS.1 Ordered Data
Normal Quantiles
Confidence Intervals: 95%
-1
0
1
2
3
-2
-1
0
1
2

Code



Se concluye que se cumple el supuesto de normalidad.

b. Utiliza el estadístico discutido en este capítulo y el que se presenta además en el capítuloo anterior cuando la población no es normal.

Code

Se realiza prueba de hipótesis contrastando con valores críticos y pvalores

Code

```
## [1] TRUE
```

Se rechaza la hipótesis nula de que la varianza sea igual a 8. Utilizando el estadístico de prueba que asume normalidad.

Con el otro estadistico de prueba para distribuciones arbitrarias, se tiene:

Code

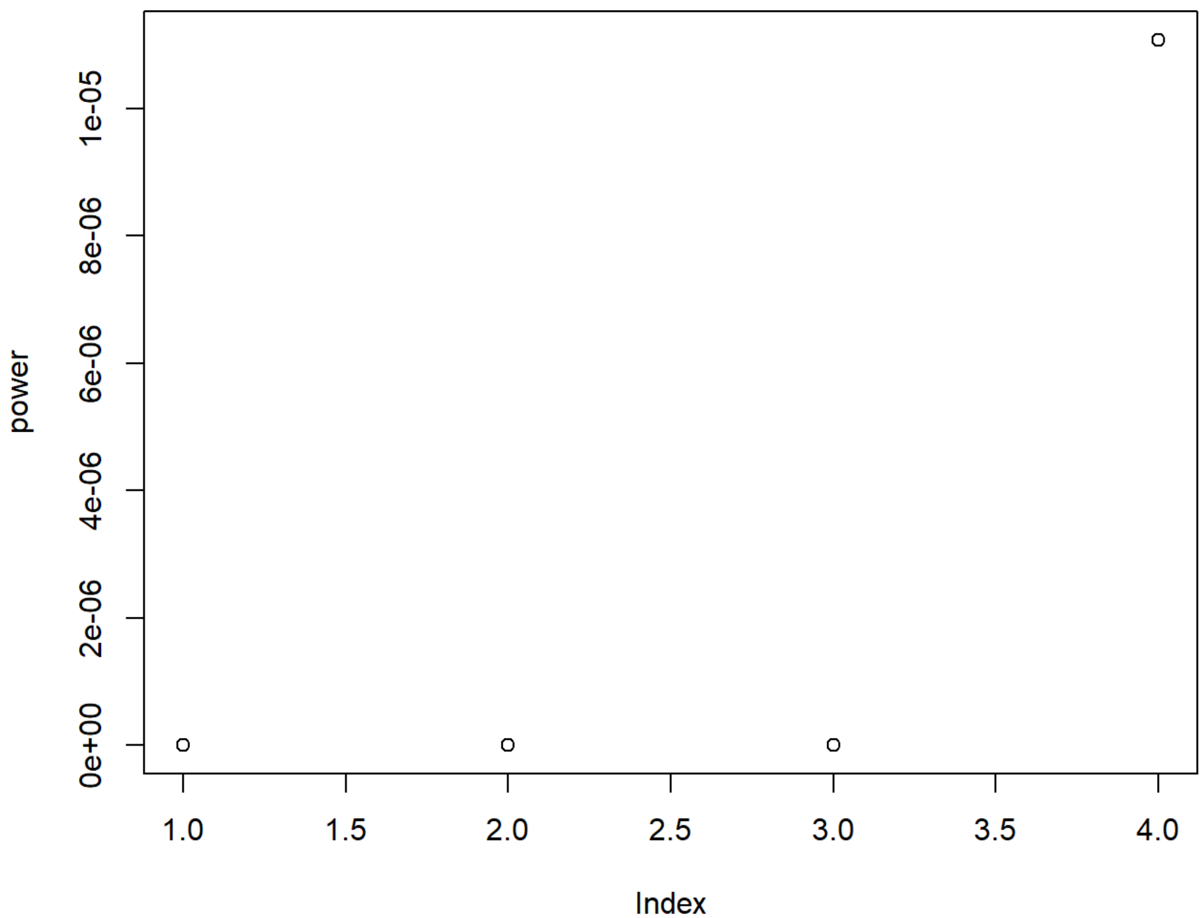
```
## [1] TRUE
```

También rechazando la hipótesis nula.

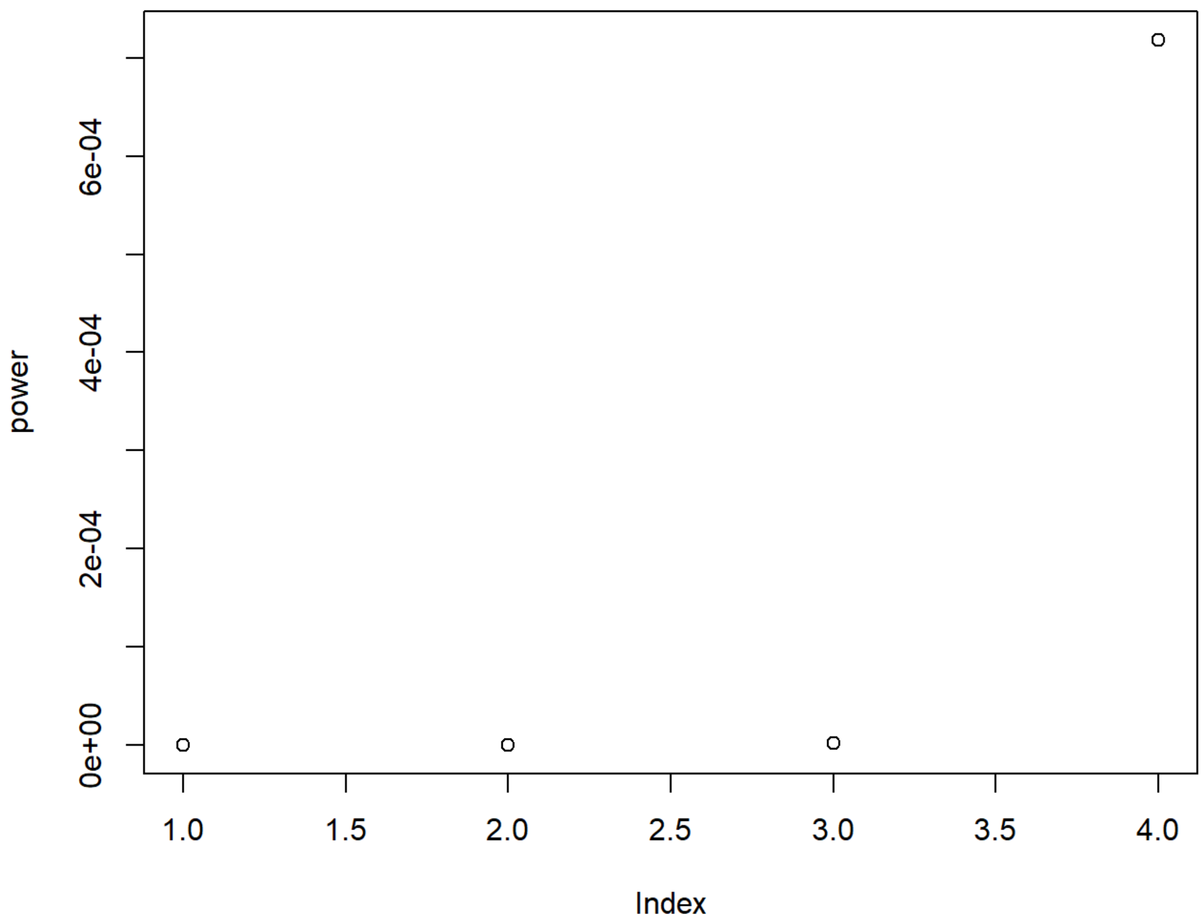
c. Calcula la potencia de la prueba para valores del cociente 1.5, 2, 2.5, 3.

Code

Code



Code



EJERCICIO 2

Se hicieron 27 mediciones de los rendimientos de dos procesos industriales, con los siguientes resultados:

EJERCICIO 2 DE PRACTICA

Code

Suponiendo que los rendimientos se distribuyen normalmente con la misma varianza, encuentre intervalos de confianza para las medias μ_1 y μ_2 y para la diferencia de medias μ_1 - μ_2

Asumiendo que las varianzas que se dan en el ejercicio son las de proceso (poblacionales). Al final de este ejercicio, se realiza de nuevo, sin embargo, ahora si se considera que son estimaciones y se realizan los intervalos correspondientes. Esto con el fin, de poder construir la mayor forma posibe de intervalos, con fines didacticos.

Intervalo de confianza al para μ_1 con varianza del proceso conocida

Code

```
##
##
##          Proceso 1 IC al 0.1   Proceso 1 IC al 0.05   Proceso 1 IC al 
0.01
## ------  --------------------  ---------------------  ------------------
---
## LOWER               4.35038               3.990294               3.286
528
## UPPER               8.10962               8.469706               9.173
472
```

Calculado intervalo de confianza para el parámetro μ_2

Code

```
##
##
##          IC al  0.1   IC al 0.05   IC al 0.01
## ------  -----------  -----------  -----------
## LOWER      11.02524     10.69674      10.0547
## UPPER      14.45476     14.78326      15.4253
```

Proceso 1 y 2 : Intervalos de confianza, para dos poblaciones con varianza del proceso conocida, pero diferente. A su vez, se realizará una prueba de hipótesis sobres si la discrepancia entre las muestras son muy grandes. En el sentido de si la probabilidad de que se observe la discrepancia es grande o caso contrario no sea.

De esta manera, estamos interesados en ver si las dos poblaciones son distintas.

Como la varianza de los procesos son conocidas, el pivote para la construcción del intervalo de confianza es el siguiente: $\mu_1 - \mu_2 = 0$

Code

```
##
##
##          IC al  0.1   IC al 0.05   IC al 0.01
## ------  -----------  -----------  -----------
## LOWER     -7.789576    -8.034709    -8.513807
## UPPER     -5.230424    -4.985291    -4.506193
```

Como el intervalo es integramente negativo, podemos estar seguros (con una con una confiabilidad tanto del 90%, 95% y 99% de significancia), que la diferencia entre las poblaciones es negativa. Significando que la media del proceso uno, es menor que la media del proceso dos. Por lo tanto no podemos aceptar que las dos poblaciones son iguales.

EJERCICIO 2 TAREA

Ahora, se realiza lo que en realidad se pide. Se desconocen varianzas poblacionales, y se toman las muestrales (pp Diapositiva 174 Caso 3). Caso en el que la población viene de una normal y sigma cuadrada es desconocida

Intervalo de confianza al para $\mu_1$con varianza del proceso conocida.

Code

```
##
##
##          Proceso 1 IC al 0.1   Proceso 1 IC al 0.05   Proceso 1 IC al
0.01
## ------  --------------------  ---------------------  ------------------
---
## LOWER               5.166122               4.922127               4.369
701
## UPPER               7.293878               7.537873               8.090
300
```

Calculado intervalo de confianza para el parámetro μ_1

Code

```
##
##
##          Proceso 1 IC al 0.1   Proceso 1 IC al 0.05   Proceso 1 IC al
0.01
## ------  --------------------  ---------------------  ------------------
---
## LOWER               11.84504               11.65186               11.23
566
## UPPER               13.63496               13.82814               14.24
434
```

Proceso 1 y 2 : Intervalos de confianza para dos poblaciones con varianza desconocida y diferente. Para este caso como la población se asume normal, una de las soluciones es la propuesta por Satterthwaite (1946) $\mu_1 - \mu_2 = 0$

Code

El pivote se va a distribuir como un t- student con n grados de libertad los cuales se utiliza la solución propuesta por Satterthwaite (1946) , para el númeor efectivo de grados de libertad.

Code

```
##
##
##          IC al  0.1   IC al 0.05   IC al 0.01
## ------  -----------  -----------  -----------
## LOWER     -7.845813    -8.123322    -8.702787
## UPPER     -5.174187    -4.896678    -4.317214
```

Como el intervalo es integramente negativo, podemos estar seguros (con una con una confiabilidad tanto del 90%, 95% y 99% de significancia), que la diferencia entre las poblaciones es negativa. Significando que la media del proceso uno, es menor que la media del proceso dos. Por lo tanto no podemos aceptar que las dos poblaciones son iguales.

EJERCICIO 3

Un experimento para determinar el efecto de una medicina en la concentraci´on de glucosa en la sangre de ratas diab´eticas dio los siguientes resultados: Grupo control: 2.051.822.001.942.12 Grupo tratamiento: 1.711.372.041.501.691.83

Analiza la hip´otesis de que el tratamiento no tiene efecto sobre la media de la concentraciónde la glucosa en la sangre. Mencione las hip´otesis bajo las cuales realiza el análsis.

Experimento en concentración de glucosa en la sangre de ratas diabéticas

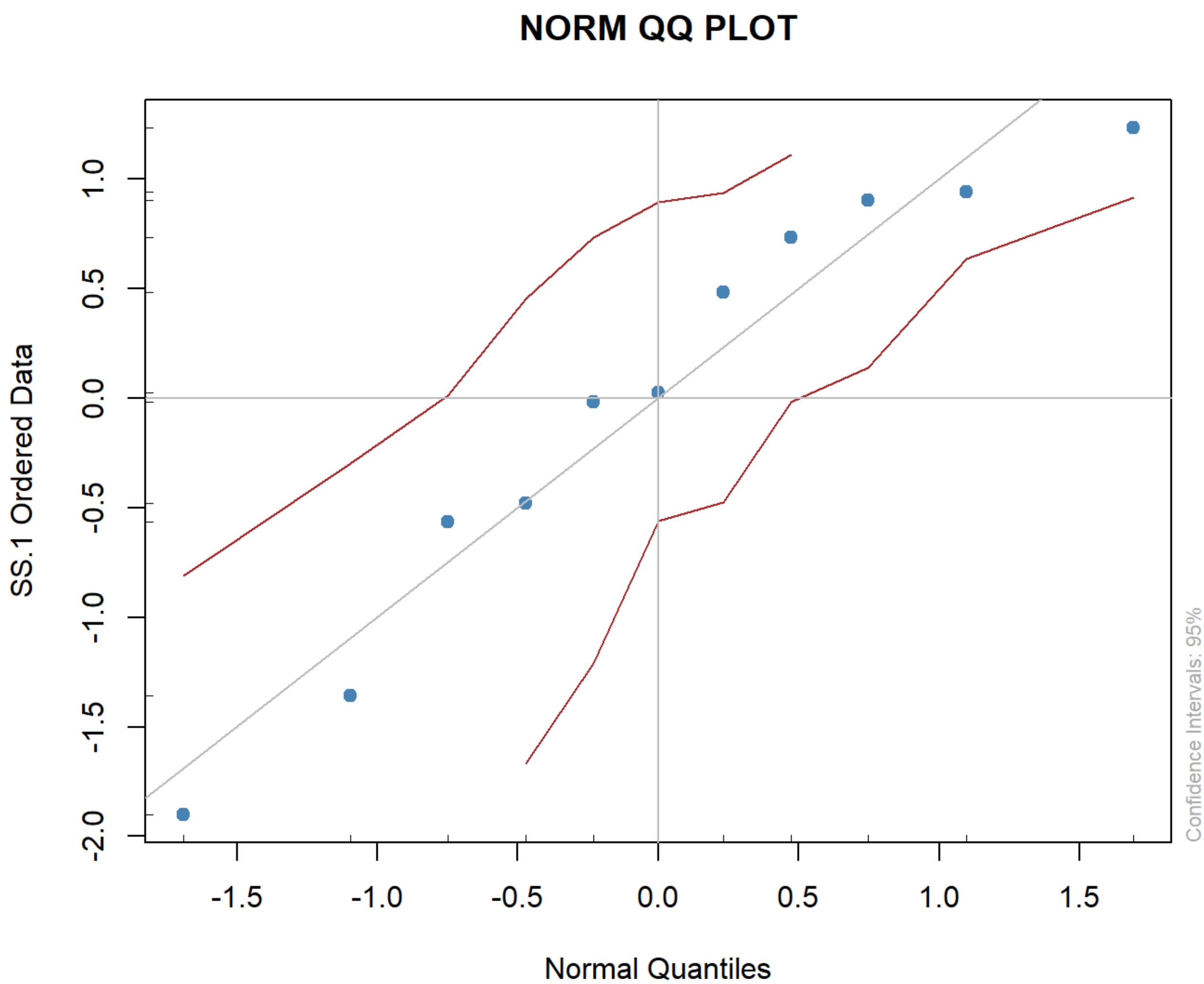
Code

Hipotesis de que el tratamiento no tiene efecto sobre la media de la concentración de la glucosa de sangre.

Code

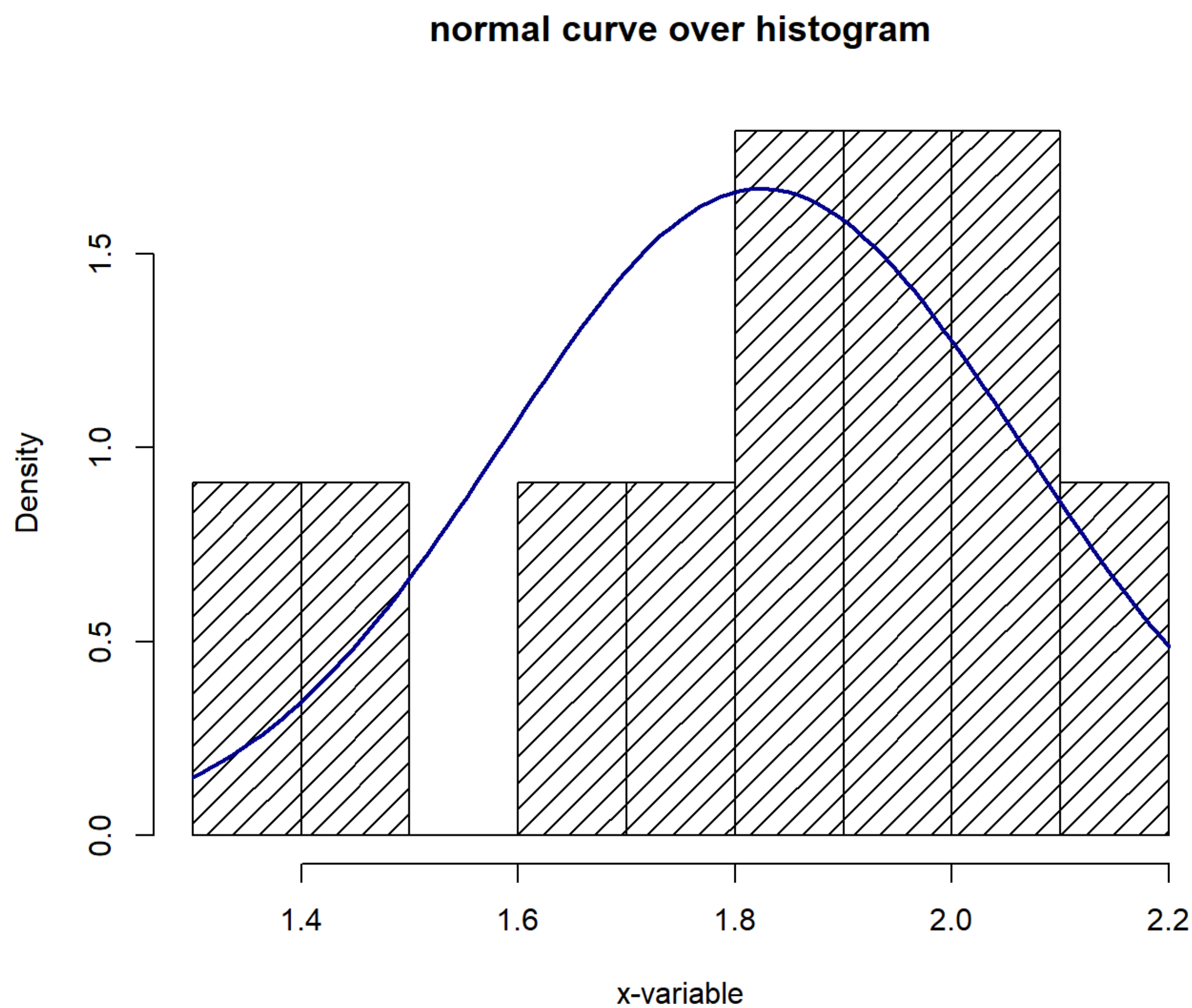
$H_0 : \mu = 0 \, \text{vs} \, H_A : \mu \neq 0$

Code



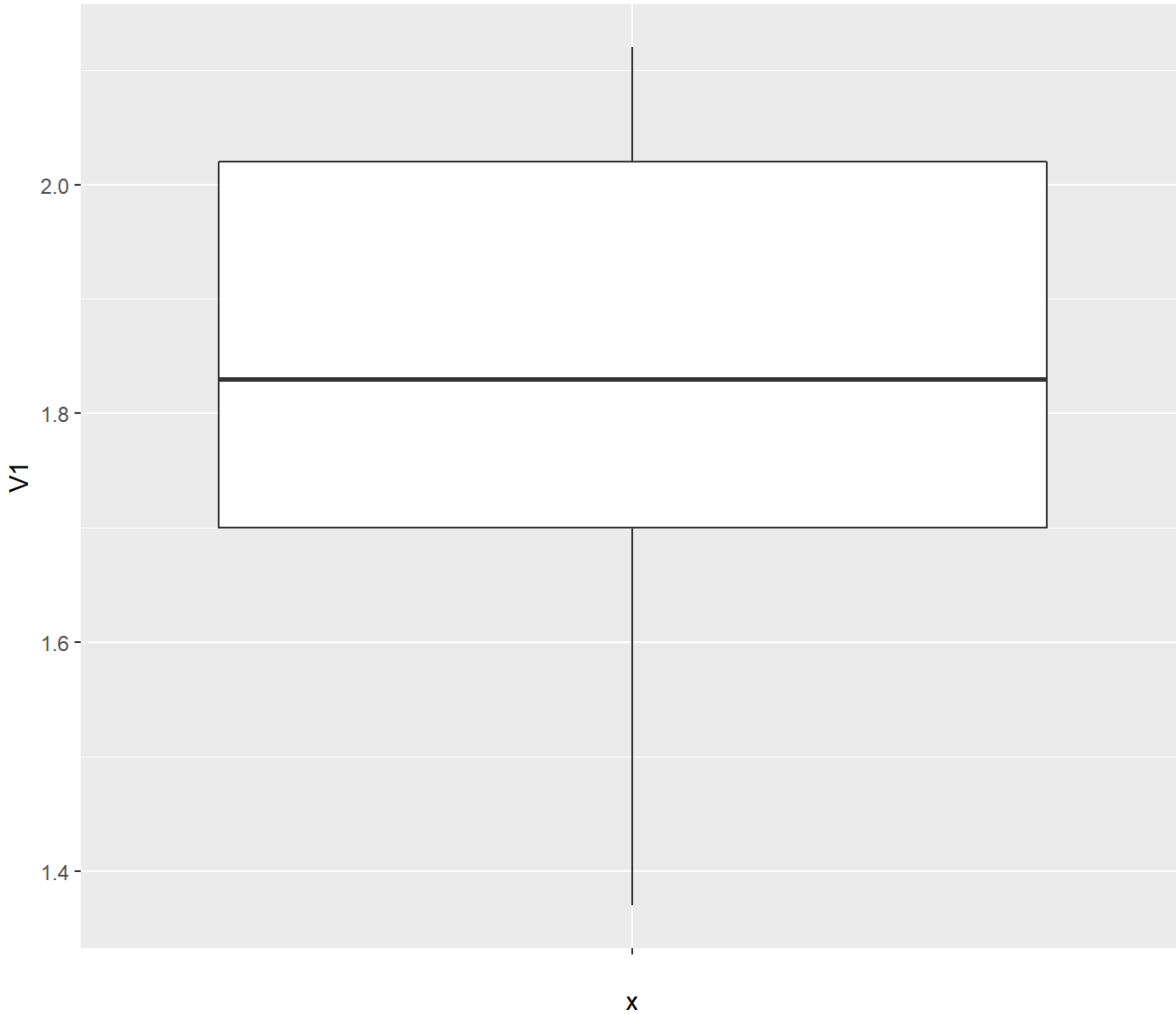
Se asume que la muestra es normal con media mu y varianza sigma cuadrara. No obstante, implementamos un qqplot y observamos que las observaciones se encuentran dentro del intervalo de confianza al 95% de significancia.

Code



Al trazar el histrograma se observa como la distribución se encuentra un poco sesgada a la izquierda (derecha para quien observa).

Code



Sin embargo, con el boxplot se observa de forma marcada el sesgo previamente mencionado. Entre las cosas que se puede observar es la mediana jalada un poco hacia el lado donde la distribución empirica se sesga.

Por último, aplicamos test de normalidad a la series

Code

```
##
##  Shapiro-Wilk normality test
##
## data:  Obs
## W = 0.93536, p-value = 0.4676
```

$p - value$ es mayor a 0.1, y el estadístico de shapiro-wilk se acerca a uno. Por lo tanto, bajo la hipótesis nula de normalidad, los datos muestran evidencia suficiente que se distribuyen como una normal.

Nota: se es conciente del limitante del tamaño de la muestra y que esto inclusive pueda que la distribución no sea una normal. Sin embargo, para manera de ejercicio, se asume la normalidad.

Una vez encontrada la distribución del las muestras, se realiza la prueba de hipótesis. Don de el estadístico de prueba de va a distribuir bajo la nula como una normal con media 0 (que no tenga efecto el test), y varianza sigma cuadrada.

Dado que no se conoce la varianza del proceso, se procede a estimarla, con la varianza muestral

Code

El estadistico de prueba esta dado por:

Code

De este modo, se rechaza la Ho de que la media es cero, es decir, que el tratamiento no tiene efecto sobre la media de la concentración de la glucosa de la sangre de las ratas.

Recordemos, que al utilizar la estimación de la varianza, el estadistico de prueba se distribuye como un t de Student con $n-1$ grados delibertad, y alpha grados de significancia.

Code

```
##
##
##                      Rechaza H0 de mu = 0
## ------------------  ---------------------
## 90% significancia   TRUE
## 95% significancia   TRUE
## 99% significancia   TRUE
```

Por lotanto, se tiene evidencia suficiente de rechazar la hipótesis nula, y con ello, decir que el tratamiento no tiene efecto sobre la media de la concentración de la glucosa en la sangre.

EJERCICIO 6

Una máquina de bebidas está diseñada para descargar, cuando opera apropiadamente, al menos 7 onzas de bebida por tasa con una desviación estándar de 0.2 onzas. Si un estadístico selecciona una muestra aleatoria de 16 tasa para examinar el servicio al cliente y este está dispuesto a tomar un riesgo $\alpha = 0.05$ de cometer un error de Tipo I, calcule la potencia de la prueba y la probabilidad (β) de tener un error del Tipo II si la media poblacional de la cantidad despachada es: a) 6.9 onzas por tasa. b) 6.8 onzas por tasa. Puede asumir que los datos son normales

Code

La hipótesis nula se plantea en terminos de que dicha maquina descarga 7 onzas por tasa, y en este caso se conoce la desviación estandar del proceso, que es de 2 desviaciones. De esta manera, se tiene una prueba de dos colas, o es 7 onzas por tasa, o es algo distinto a esa cantidad.

De acuerdo a la clase, pruebas de hipótesis con estás carácteristicas rechazará tanto valores pequeños del estadístico, como con valores grandes. Por lo tanto, se nos sugiere que se calcule la propabilidad conjunta de que estos dos eventos sucedan a un cierto nivel de acepatación del error tipo 1.

Code

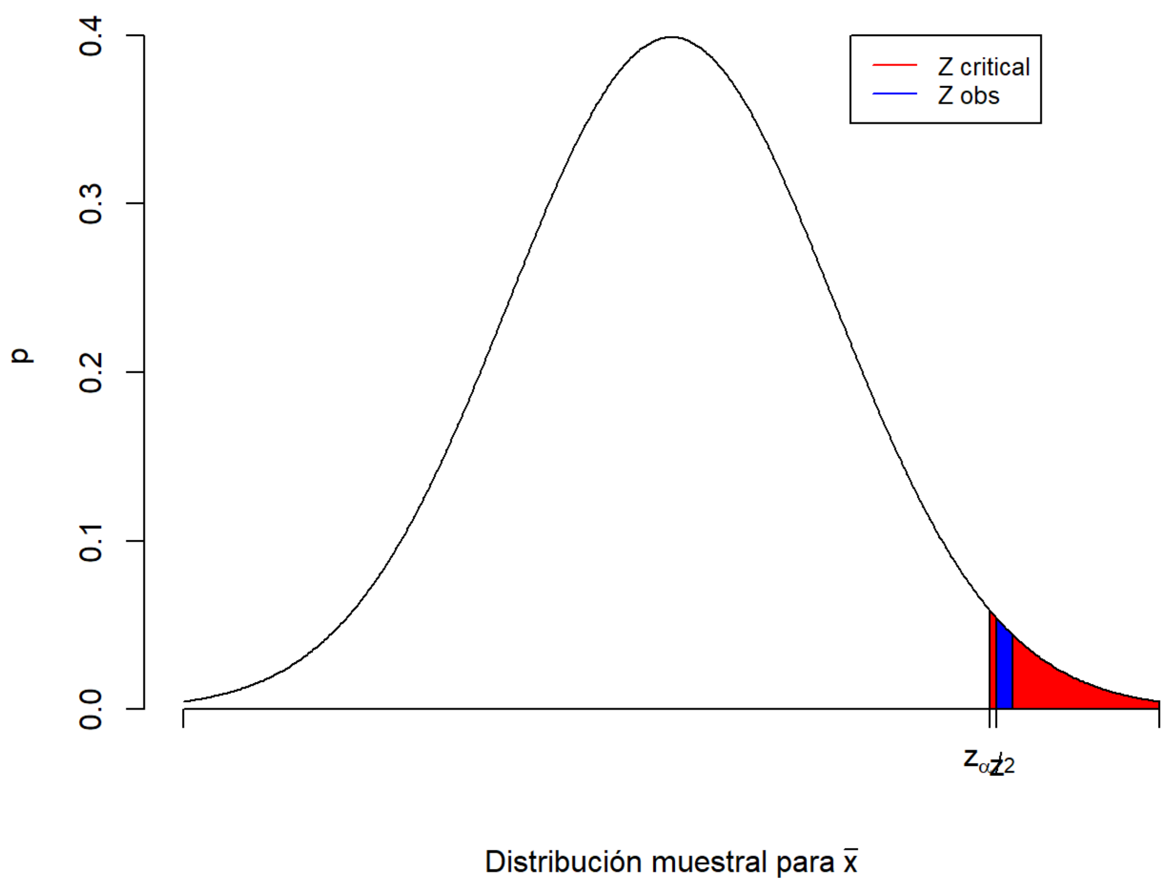
Asumiendo normalidad en la muestra, se calcula el valor z con un nivel de insignificancia de (alpha) .05. Que en este caso, será la probabilidad de incurrir en el error tipo 1, rechazar la hipótesis nula cuando era cierta.

a. Con media muestral de 6.9

Code

```
## [1] TRUE
```

Code



En este caso, se tiene evidencia suficiente para rechazar la nula

En terminos de p valores:

Code

Si se contrasta contra el valor de significancia, se tiene :

Code

```
##
##
##                      Rechaza H0
## ------------------  ------------
## 95% significancia   TRUE
```

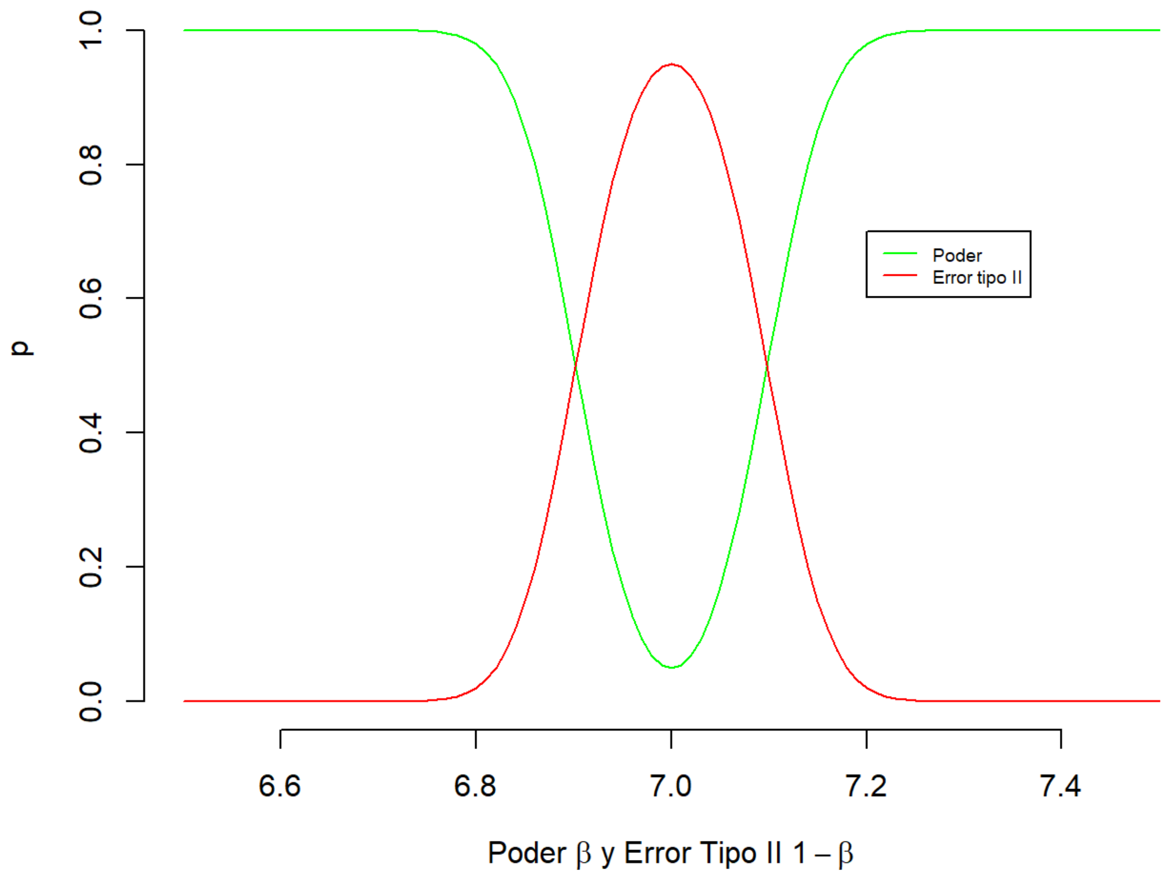
Code

```
##
##
##           Rechaza H0
## -------  ------------
## pvalue      0.0455003
```

Por lo tanto, Se tiene evidencia suficiente para rechazar ña hipótesis nula, sobre que la máquina de bebidas descarga en promedio 7 onzas por tasa.

Potencia Y error tipo 2 de la prueba a un cierto niver de significancia alpha

Code



En la gráfica anterior se observa la función poder y el error tipo II a un nivel de error tipo 1 del 0.05, cambiando los valores de la media sobre la hipótesis alternativa.

Una vez realizado lo anterior, ahora realizamos la segunda parte del ejercicio. En esta parte, la media muestra de de el número de onzas por tazas que descarga es de 6.8.

Primero realizamos una prueba de hipótesis de dos colas. Con la Hipótesis nula que es tener 7 onzas por taza con una desviación estandar conocida del proceso.

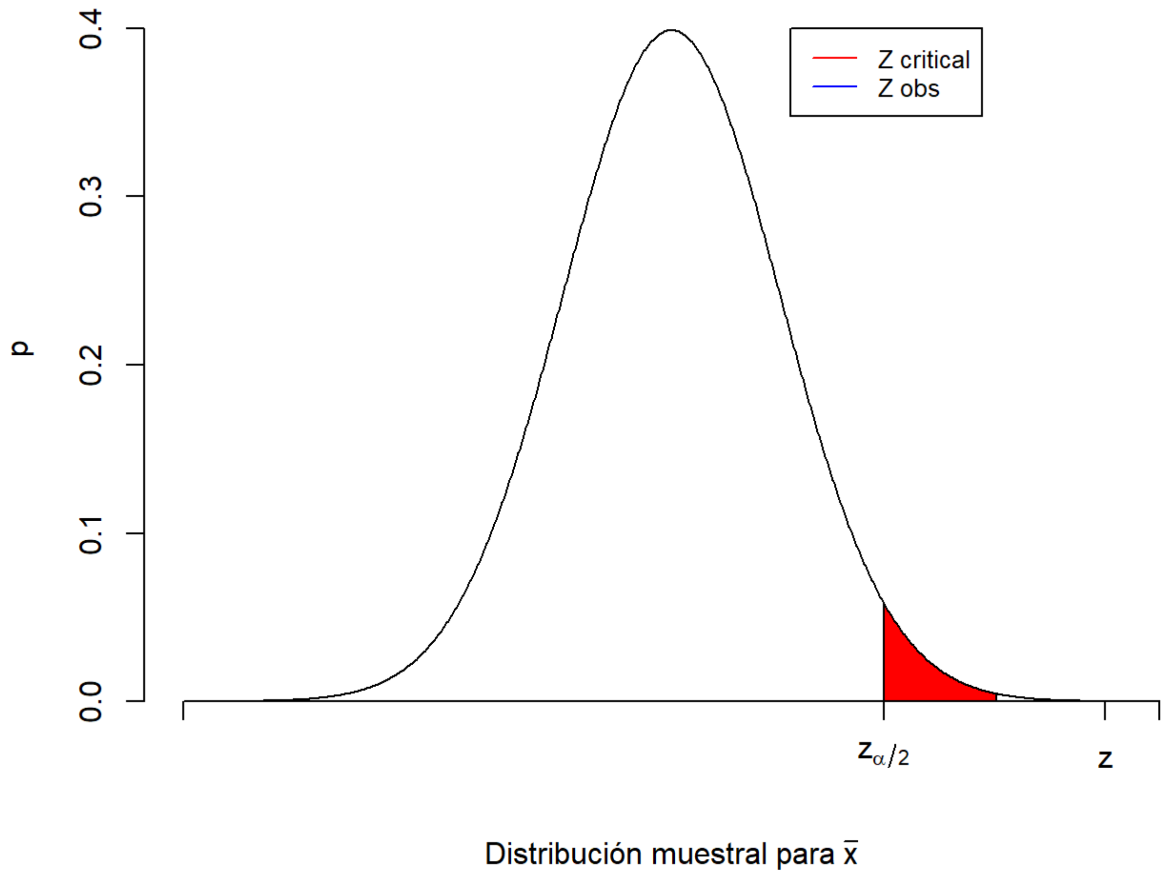
Puede asumir que los datos son normales. No, dado que el tamaño de la muestra (16) es pequeña. Sin embargo, se podría utiizar la distribución t para muestras pequeñas, pero lo cual implicaría colas más pesadad.

b. Con media muestral de 6.8.

Code

```
## [1] TRUE
```

Code



En este caso, se tiene evidencia suficiente para rechazar la nula ya que cae en la región de rechazo

En terminos de p valores:

Code

Si se contrasta contra el valor de significancia, se tiene :

Code

```
##
##
##                      Rechaza H0
## ------------------   ------------
## 95% significancia    TRUE
```

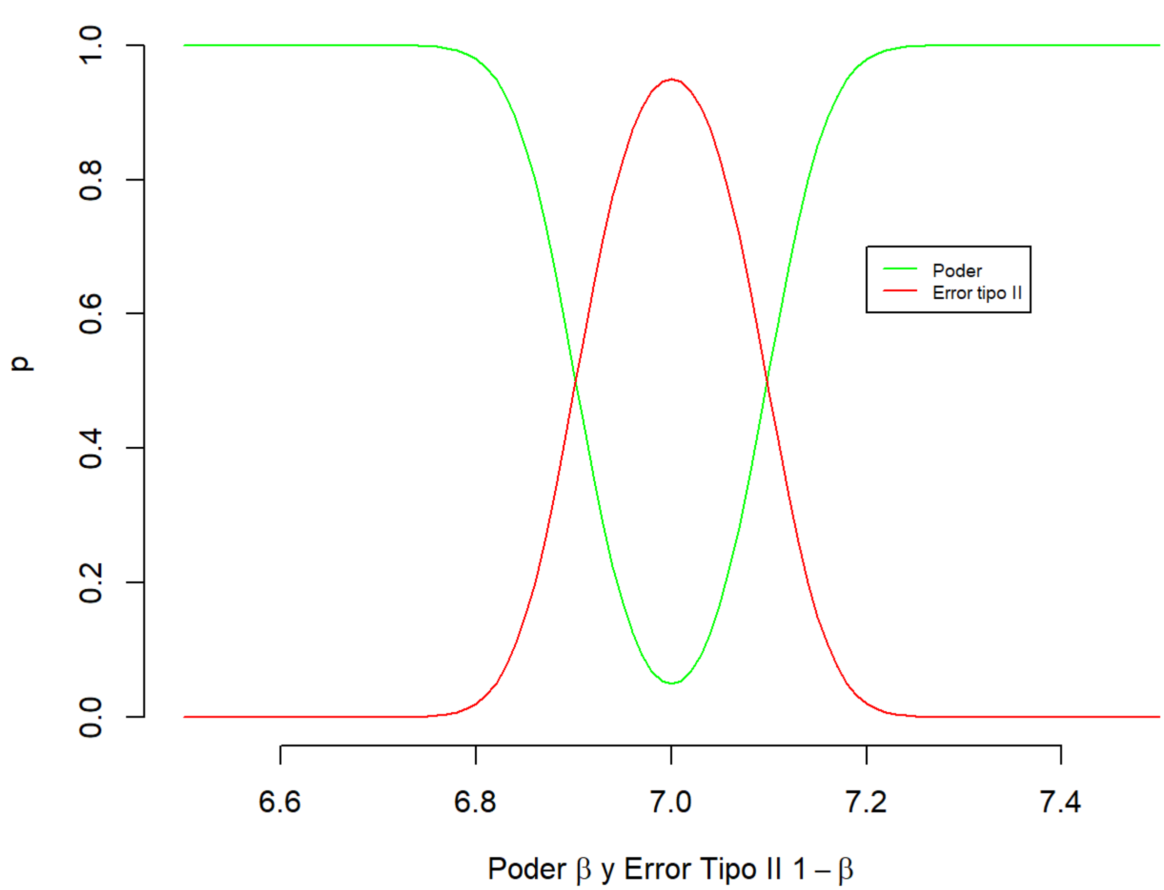
Code

```
##
##
##           Rechaza H0
## -------  ------------
## pvalue       6.33e-05
```

Por lo tanto, al igual que con una media de 6.9, se tiene evidencia suficiente para rechazar ña hipótesis nula, sobre que la máquina de bebidas descarga en promedio 7 onzas por taza.

Para responder la pregunta de si se puede asumir normalidad, se vuelve a gráficar la función de poder y el error tipo II.

Code



Lo que se puede decir sobre la operación de la máqina de bebidad, es:

Code

```
##
##
##                  Ha = 6.9    Ha = 7.1
## -------------  ----------  ----------
## Poder           0.5160053   0.5160053
## Error Tipo 2    0.4839947   0.4839947
```

Ante variaciones de 0.1 se tiene la misma probabilidad tanto de detectar un incremento (poder) o como de no detectarlo (error tipo 2) con esta prueba, dado a ese tamaño de la muestra. Posiblemente la desviación no es lo suficientemente pequeña, esto como para detectar una variación lo suficientemente pequeña y detectar dicha variación. De está manera, se recomendaría incrementar el tamaño de la muestra, para bajo ciertas condiciones se pueda aplicar el TLC, y que el estadístico de prueba se pueda convergen en distribución a una normal estandar. Por lo tanto, se concluye que no se puede asumir que estos datos son normales, a menos que asíntoticamente aproximaramos a una distribución normal estandar.

EJERCICIO 7

Construya un intervalo de confianza aproximado del 90% para el parámetro λ de una distribución de Poisson. Evalúe su intervalo si una muestra de tamaño 30 produce $\sum x_i = 240$

Code

```
##
##
##            Información
## -------  ------------
## Lambda            8.0
## N                30.0
## alpha             0.1
```

Si consideramos n como grande, por TLC, el pivote converge en distribución, bajo ciertas condiciones, a una normal estandar $N(0, 1)$.

Se busca que el estimador sea consistente, es decir, que cuando n sea grande, el estimador del parámetro conerga en probabilidad al parámetro poblacional. Entonces, se busca que la diferencia de medias sea cero.

Code

```
##
##
##          IC  lambda ML 95% significancia
## ------  --------------------------------
## Lower                          5.597537
## Upper                         10.402463
```

En conclusión, si la muestra fue grande, el intervalo construido por estimadores de maxima verosimilitud por la propiedad de invariabilidad ante transformaciones, hacen que el pivote sea eficiente, así como lo es el estimador de MV, cumpliendo con e críterio de mínima varianza y garantizando el menor error de estimación posible.

EJERCICIO 9

Sea $X_1, \ldots, X_n \sim$ Poisson(λ).

a. Dado a que es una prueba de dos colas y un estimador para lamda es la media muestral (ta sea por maxima verosimilitud u método de momentos), y como en el ejercicio nos indican que el tamaño de muestra es grande. Entonces, el estimador de lamda se puede aproximar asintoticamente a una $Normal(0,1)$. Siendo esto, la definición del test de Wald. Por lo tanto, el estadístico de prueba que se propone es el estadístico de Wald (ponerlo).

$W = (\hat{\theta} - \theta)/\hat{s}e$

b) Establezca la región de rechazo para alpha 0.05 c) Sea $\lambda_0 = 1$, $n = 20$,$\alpha = 0.05$...

Code

No se tiene evidencia suficiente para rechazar la hipótesis nula.

Code

```
##
##
##                     Rechaza H0
## ------------------  ------------
## 95% significancia   FALSE
```

Code

```
##
##
##          Rechaza H0
## -------  ------------
## pvalue      0.7452619
```

Inclusive con el $p - valor$ no se tiene evidencia suficiente para rechazar la Hipótesis nula.

Repita esto 10^4 veces y cuente que tan a menudo rechaza la hipótesis nula

Code

¿ Qué tan a menudo rechaza la hipótesis nula ?

Code

```
##
##
##                 (%) proporción de veces que rechaza Ho
## ------------  --------------------------------------
## 100%(1-.05)                                      6.94
```

6.9% de las veces se rechaza la hipótesis nula.

¿ Qué tan cercanaes la tasa de error tipo I de$ 0.05$?

Code

```
##
##
##                   (%) proporción de veces que rechaza Ho
## --------------  --------------------------------------
## Tasa de error                                    0.0694
## alpha                                            0.0500
```

Como se observa la tasa de error es cercana al nivel de significancia de 0.05. Con una diferencia mayor de 0.0194.

EJERCICIO 10

En la librería boot de R accesar los datos cd4 los cules son conteos de células CD4 en pacientes VIH-positivos antes y despúes de un año de tratamento con un antiviral

Code

```
##   baseline oneyear
## 1     2.12    2.47
## 2     4.35    4.61
## 3     3.39    5.26
## 4     2.51    3.02
## 5     4.04    6.36
## 6     5.10    5.93
```

Code

a. Construya un intervalo de conanza bootstrap para el coeciente de correlaci□on entre los conteos base y los conteos despu□es del tratamiento.

Code

```
##
## ORDINARY NONPARAMETRIC BOOTSTRAP
##
##
## Call:
## boot(data = data, statistic = fc, R = 1000)
##
##
## Bootstrap Statistics :
##      original      bias    std. error
## t1* 0.7231654 -0.00591788  0.09028391
```

Code

Ahora utilizamos Boostrap con función propia Algoritmo basado en Wasserman (2005)

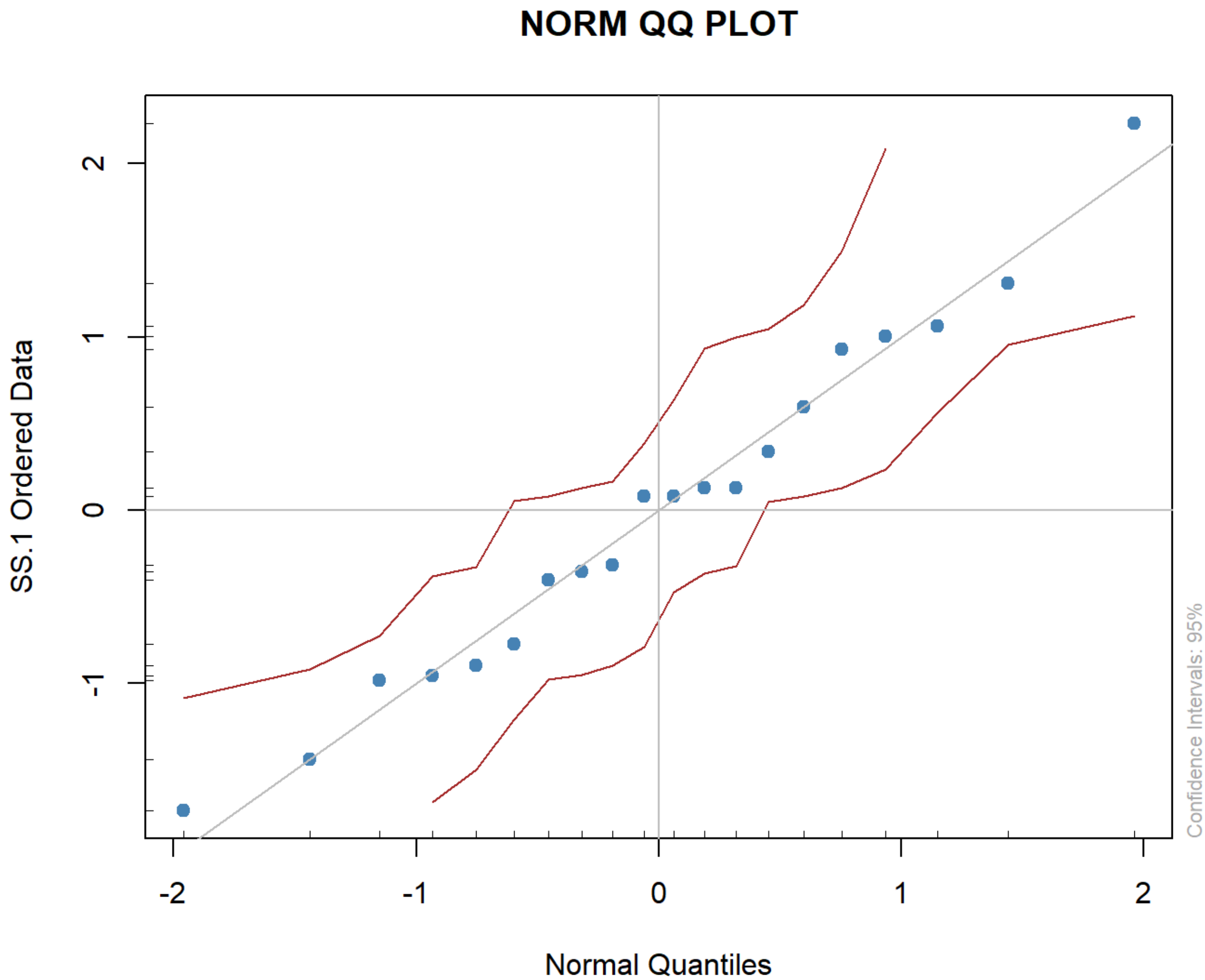
Code

Ahora, para construir el intervalo de confianza entre el coeficiente de correlación entre conteos base y conteos desopués del tratamiento, se proponen 3 intervalos de confianza. El primero es el Intervalo normal, el cual no es tan preciso si la distribución del estadístico no es una normal.

Para esto, se revisa a manera rápida un qqplot, para oservar si las muestras con las que el estadístico se construye son normales.

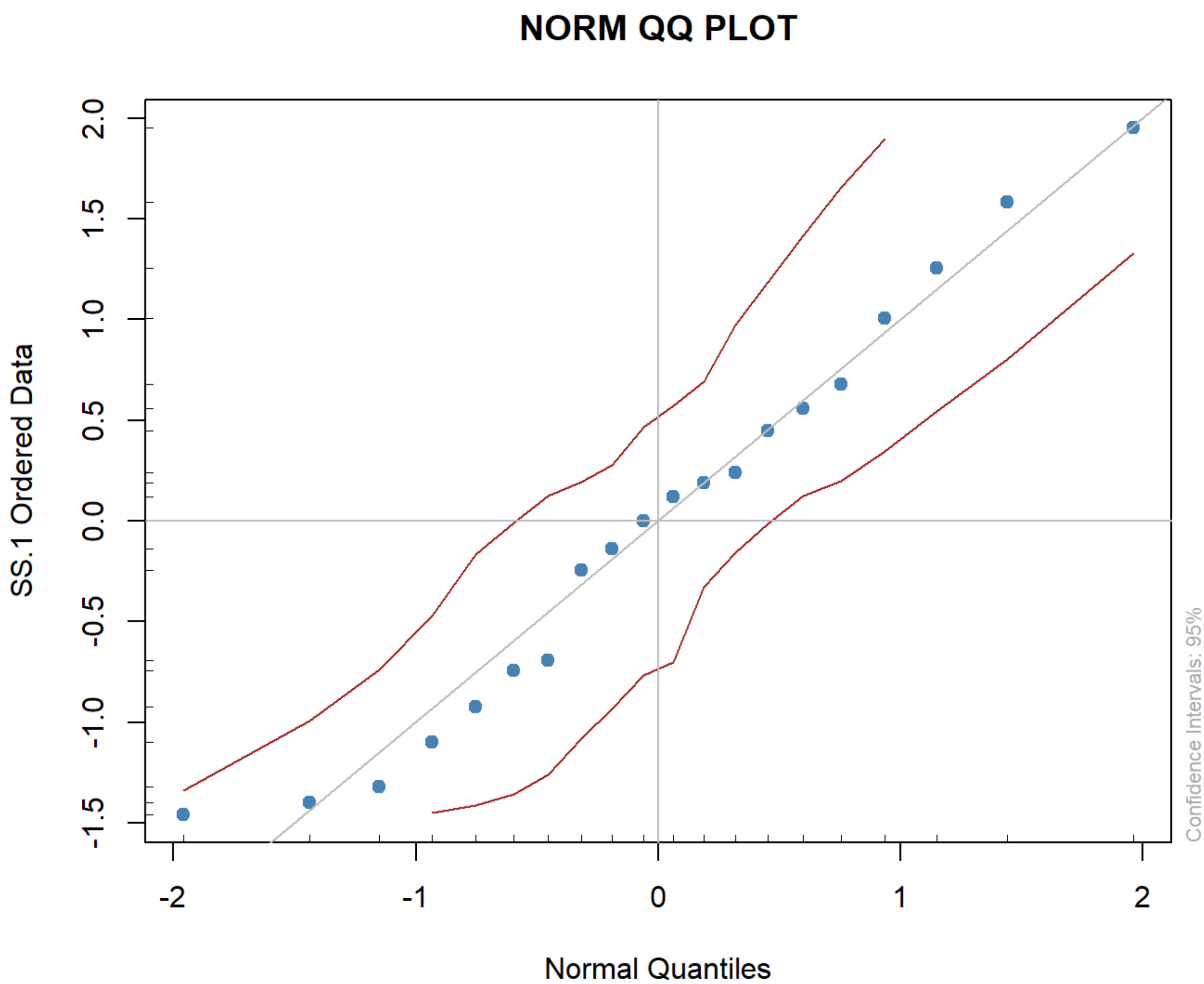
Primero los datos del conteo base:

Code



Seguido de los datos del conteo después del tratamiento:

Code



De manera general, se concluye que los datos vienen de una población normal, lo cual argumenta (no en gran medida) al utilizar un intervalo normal, con el error estandar estimado por boostrap (función propia).

Code

```
##
##
##          Intervalo de Confianza Normal al 95%
## ------  --------------------------------------
## Lower                            0.5462122
## Upper                            0.9043930
```

A su vez, se presentan intervalos de confianza alternos al Normal.

Code

```
##
##
##           Intervalo de Confianza Percentil
## -------  ---------------------------------
## .2.5%                            0.4933814
## .97.5%                           0.8604793
```

Code

```
##
##
##           Intervalo de Confianza Pivotal
## -------  -------------------------------
## .2.5%                          0.9529493
## .97.5%                         0.5858514
```

b. Calcule el coeficiente estimado de correlaci□on corregido por sesgo usando Jackknife.

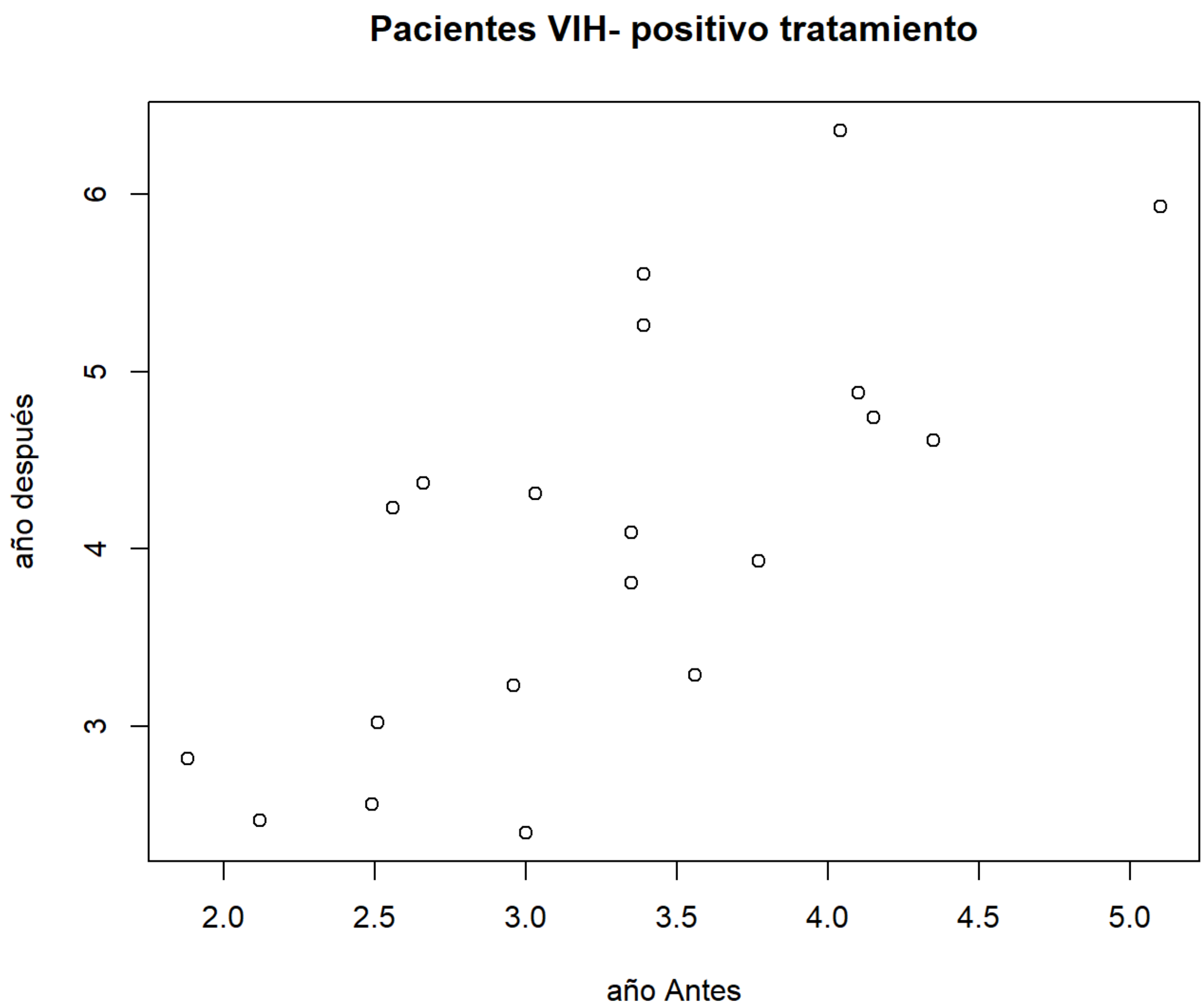
Code

```
##   baseline oneyear
## 1     2.12    2.47
## 2     4.35    4.61
## 3     3.39    5.26
## 4     2.51    3.02
## 5     4.04    6.36
## 6     5.10    5.93
```

Code

```
## [1] 20  2
```

Code



Como análisis previo se puede observarcomo la medida dependencia lineal, como lo es la correlación, entre el conteo de células CD4 en los pacientes con VIH-positivo antes y después de un año de tratamiento antiviral están corelacionados positivamente.

Jackknifing el coeficiente de correlación de corregido por sesgo

Primero calculamos el coeficiente de correlación de Pearson

Code

```
## Se observa el el coeficiente de correlación entre las serie es de:  0.7
231654
```

Jackknife coeficiente de correlación de Pearson Primero, vamos a utilizar la función jacknife de la paqueteria #library(bootstrap)

Una vez esto, calculamos el coeficiente de correlación corregido por el sesgo usando Jackknife

Code

```
##
##
##                              Jackknife
## -------------------------  -----------
## Sesgo jack                   -0.0067843
## Correlación - Sesgo jack      0.7299497
## Estimador JKf se(Corr)        0.0904822
## Correlación muestral          0.7231654
```

Ahora, se va a realizar el jacknife sin utilizar la libreria en R.

Code

```
##
##
##                              Jackknife
## -------------------------  -----------
## Sesgo jack                   -0.0067843
## Correlación - Sesgo jack      0.7299497
## Estimador JKf se(Corr)        0.0904822
## Correlación muestral          0.7231654
```

Calculando intervalo de confianza al .05

Code

```
##
##
##                              Intervalo de confianza al 95%
## -------------------------  ------------------------------
## Lower                                          0.5526045
## Correlación - Sesgo jack                       0.7299497
## Correlación muestral                           0.7231654
## Upper                                          0.9072948
```

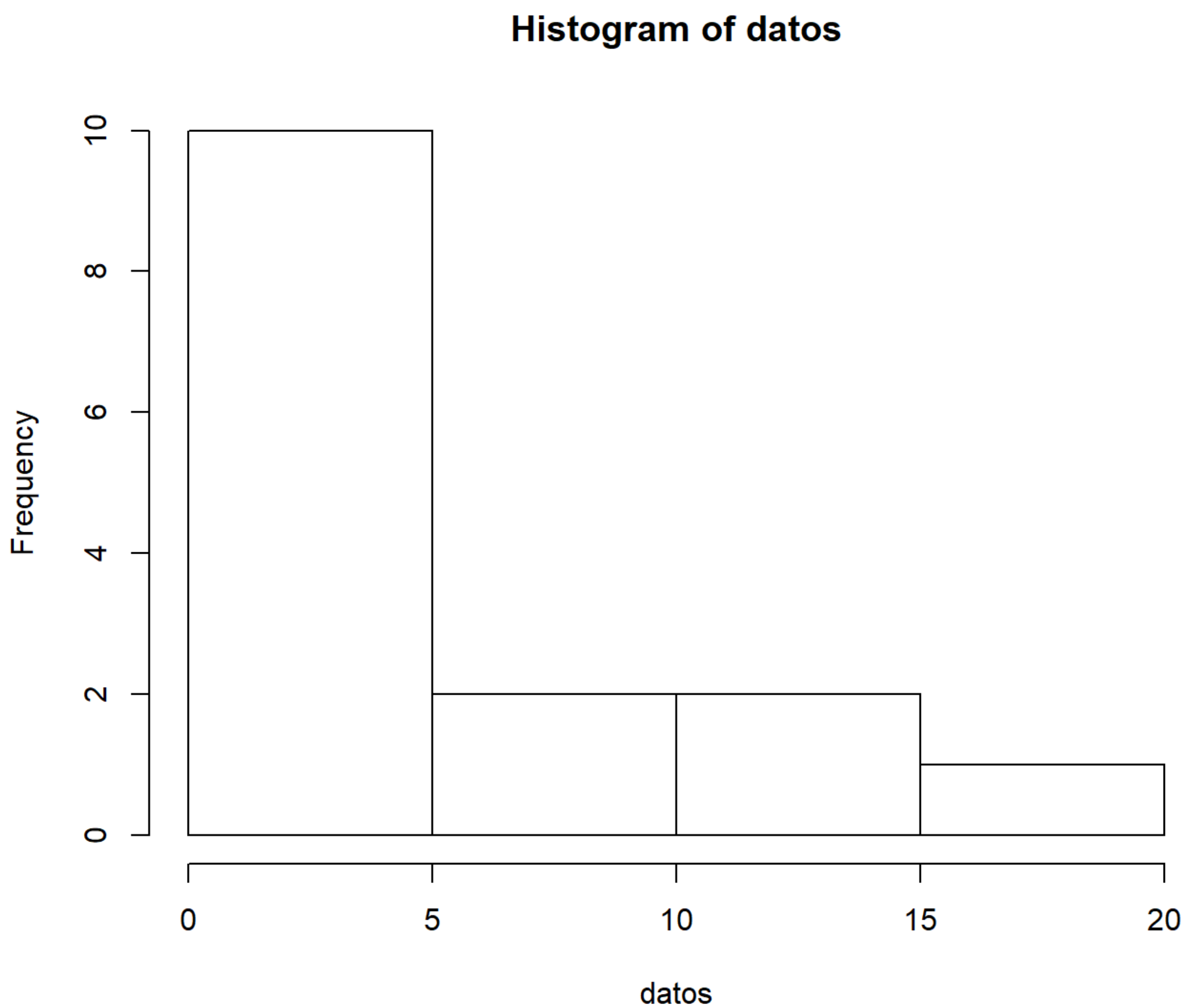
EJERCICIO 11

Los siguientes 15 datos forman una muestra aleatoria de una distribuci□on Gamma con par□ametro de forma lalpha = 3 y par□ametro de escala beta = 2 (la media es alpha*beta y la varianza alpha* beta^2)

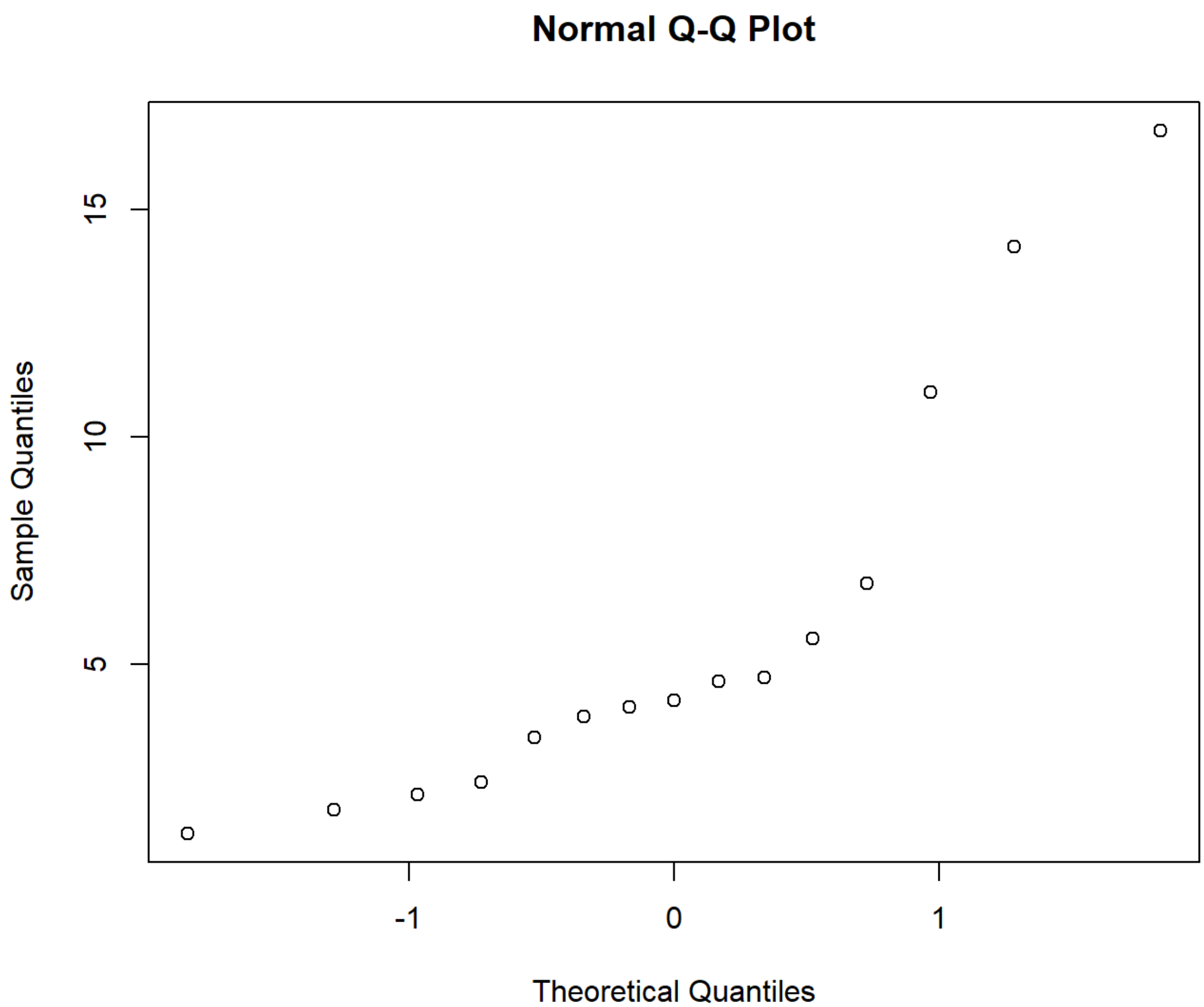
Code

Como se observa la distribución se encuentra sesgada, lo cual argumenta a foavor sobre la distribución de la muestra.

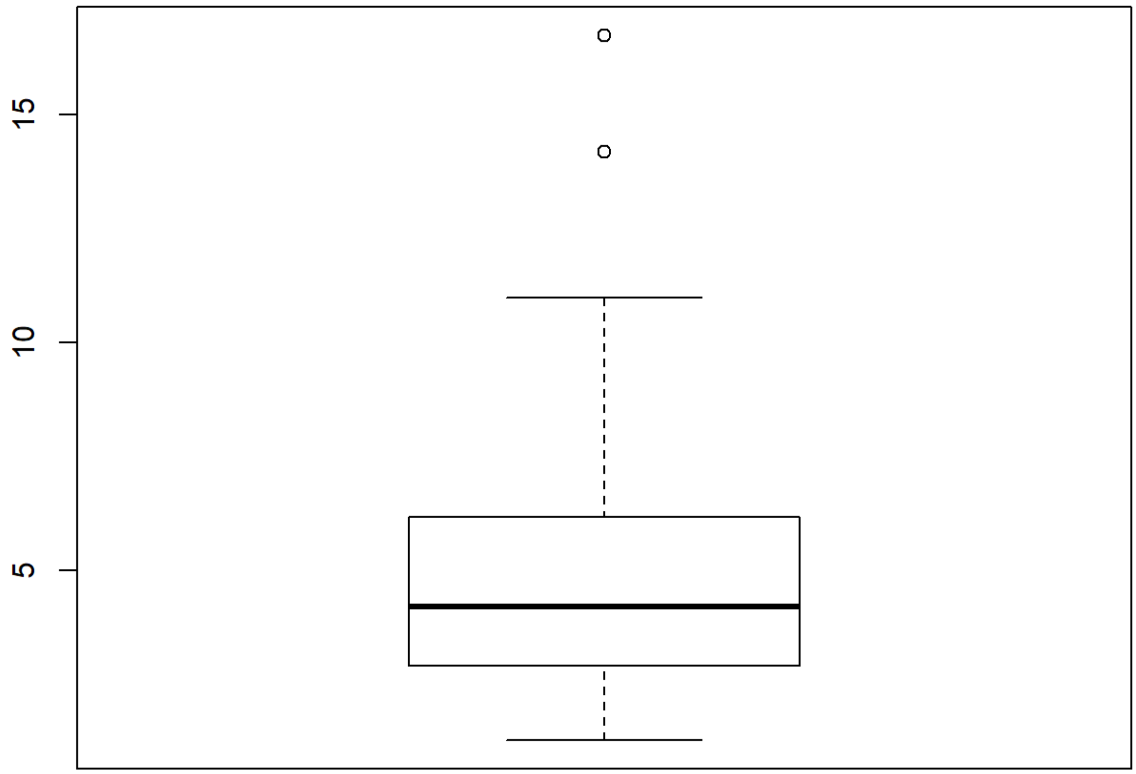
Code



Code


Normal Q-Q Plot
Sample Quantiles
5
10
15
-1
0
1
Theoretical Quantiles

Code



Encuentre un intervalo de conanza para la mediana de la distribuci□on. La mediana de la distribución no tiene una forma cerrada Por medio de Boostrap se estima el error estandar de la mediana para la construcción del intervalo de confianza.

Code

```
## [1] 4.19
```

Code

Construir un intervalo de confianza estaría mal, debido a que la muestra se distribuye gamma (3, 2). Por lo tanto, se utiliza un intervalo percentil y pivotal.

Code

A manera de ejercicio se calcula en intervalo Normal, sin embargo se sabe que no es bueno para este ejercicio.

Code

```
##
##
##                        Intervalo de confianza al 95% Mediana
## ------------------  --------------------------------------
## Lower                                              1.329952
## Upper                                              7.050048
## Mediana Muestral                                   4.190000
## Mediana -Boostrap                                  4.190000
```

Code

```
##
##
##                        Intervalo de Confianza Percentil
## ------------------  ----------------------------------
## .2.5%                                              2.39
## .97.5%                                             6.76
## Mediana Muestral                                   4.19
## Mediana -Boostrap                                  4.19
```

Code

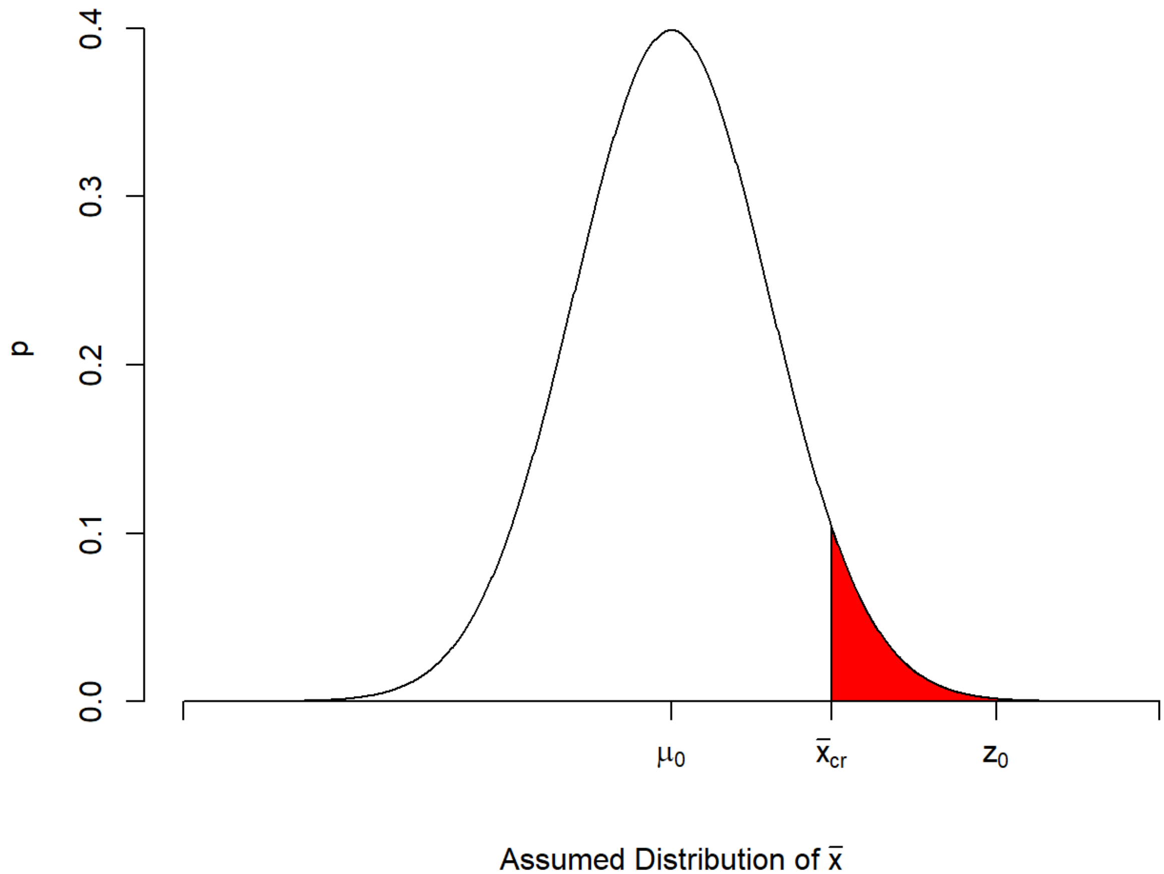
```
##
##
##                        Intervalo de Confianza Pivotal
## ------------------  --------------------------------
## .2.5%                                            5.99
## .97.5%                                           1.62
## Mediana Muestral                                 4.19
## Mediana -Boostrap                                4.19
```

EJERCICIO Extra 1 Diap. 291

En la sección de ejercicios extras no se pretende pedir puntuación extra, sólamente se hace por cuestiones morales ya que más de una vez la Dra. ha mencionado que es de buen ejercicio replicar las notas. De esta manera, solamente se replican resultados y gráficos.

Pruebas para Una Muestra. Gráfica y ejemplo de la clase.

Code



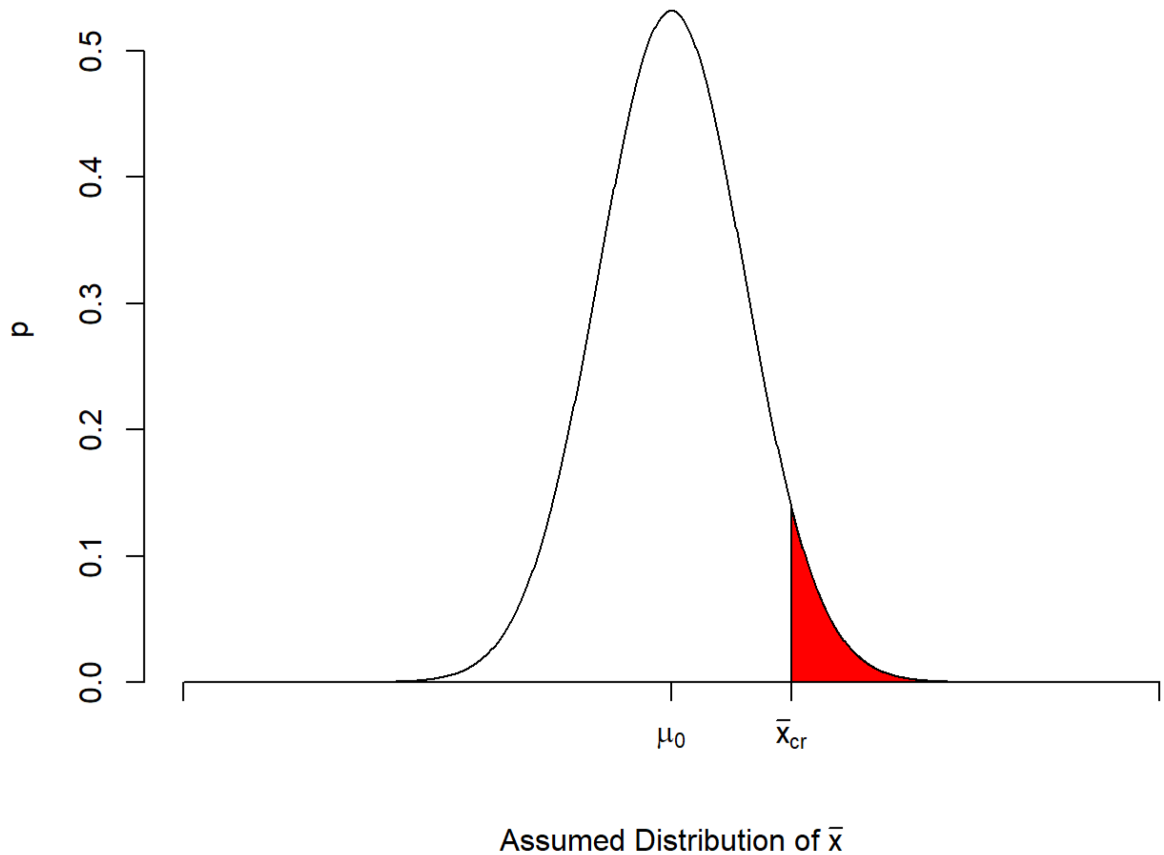
Cómo se puede observar, el valorór del piovote se encuentra en la zona de rechazo (rojo) a un nivel de significanca del 95.

Code

```
## [1] 0.0004376815
```

Incrementando la desviación estádar se tiene:

Code



EJERCICIO Extra 2 Diap. 291

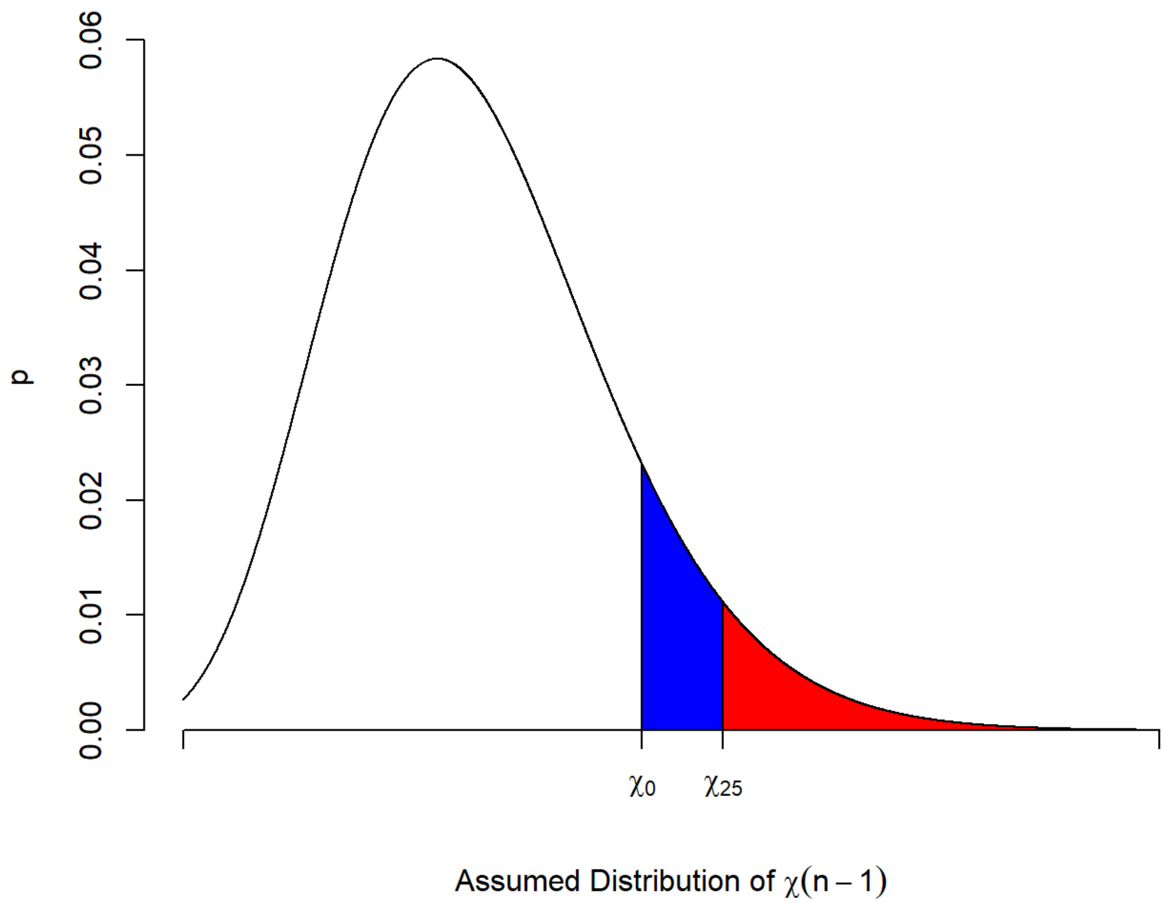
En la sección de ejercicios extras no se pretende pedir puntuación extra, sólamente se hace por cuestiones morales ya que más de una vez la Dra. ha mencionado que es de buen ejercicio replicar las notas. De esta manera, solamente se replican resultados y gráficos.

Problema: Una máquina llenadora de envases de bebida está bajo control si su promedio de llenado es de 12.2 onzas por envase, con una desviación est□andar de .05oz como máximo (vea diapositiva 291).

Code

Gráfica:

Code



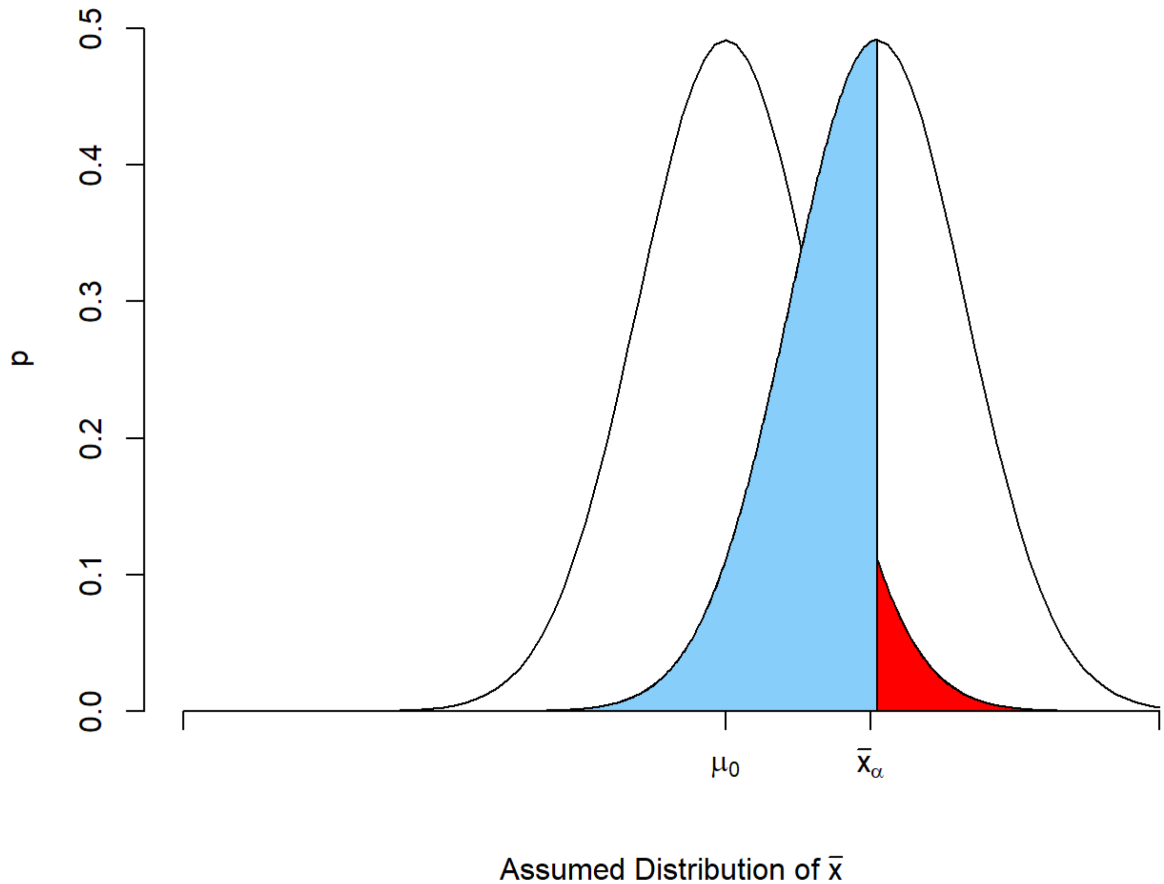
De está manera se concluye que no se tiene evidencia de rechazar la H_0.

EJERCICIO Extra 3 Diap. 321

En la sección de ejercicios extras no se pretende pedir puntuación extra, sólamente se hace por cuestiones morales ya que más de una vez la Dra. ha mencionado que es de buen ejercicio replicar las notas. De esta manera, solamente se replican resultados y gráficos.

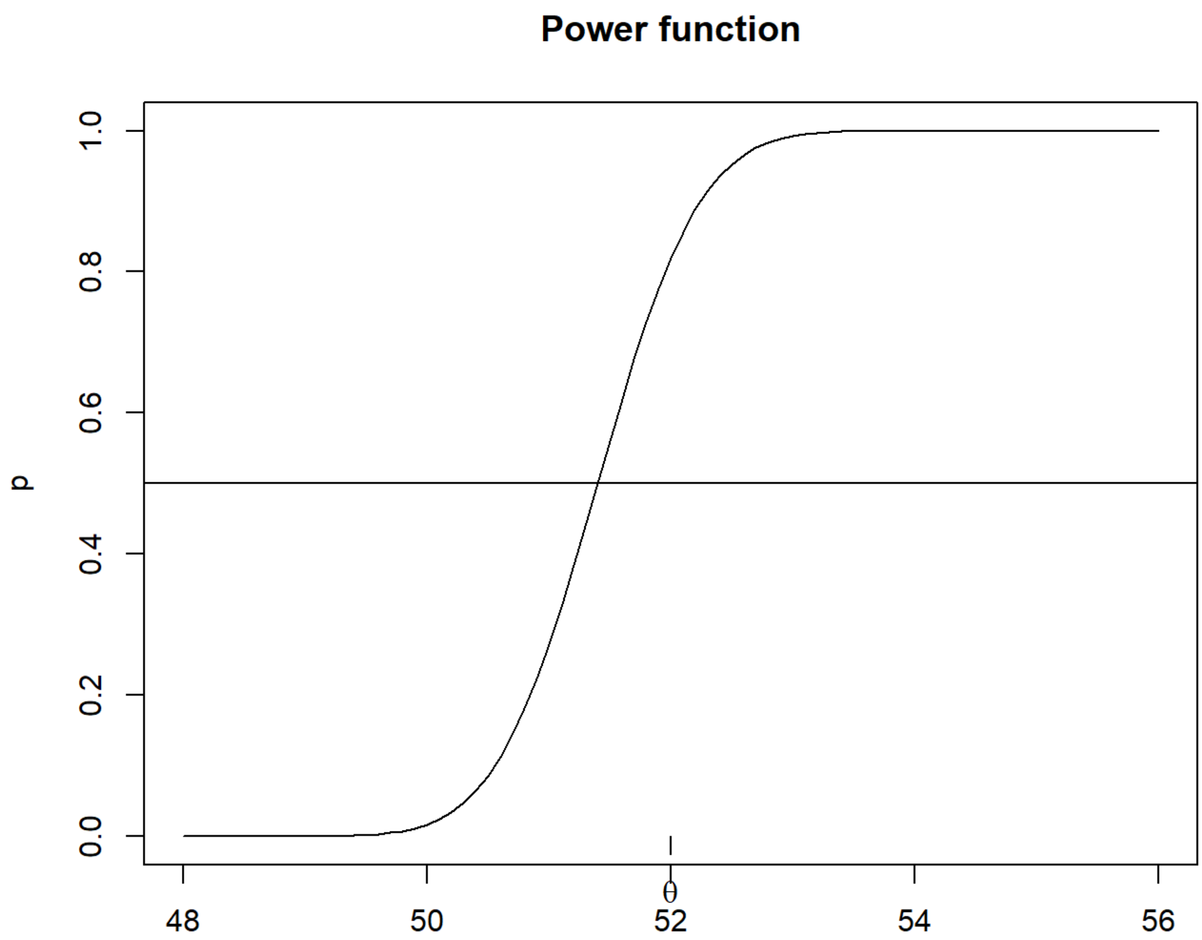
Observams que bajo un valor de la H_A la probabilidad de no rechzar dado que la alternativa es cierta es muy grande (área azul).

Code



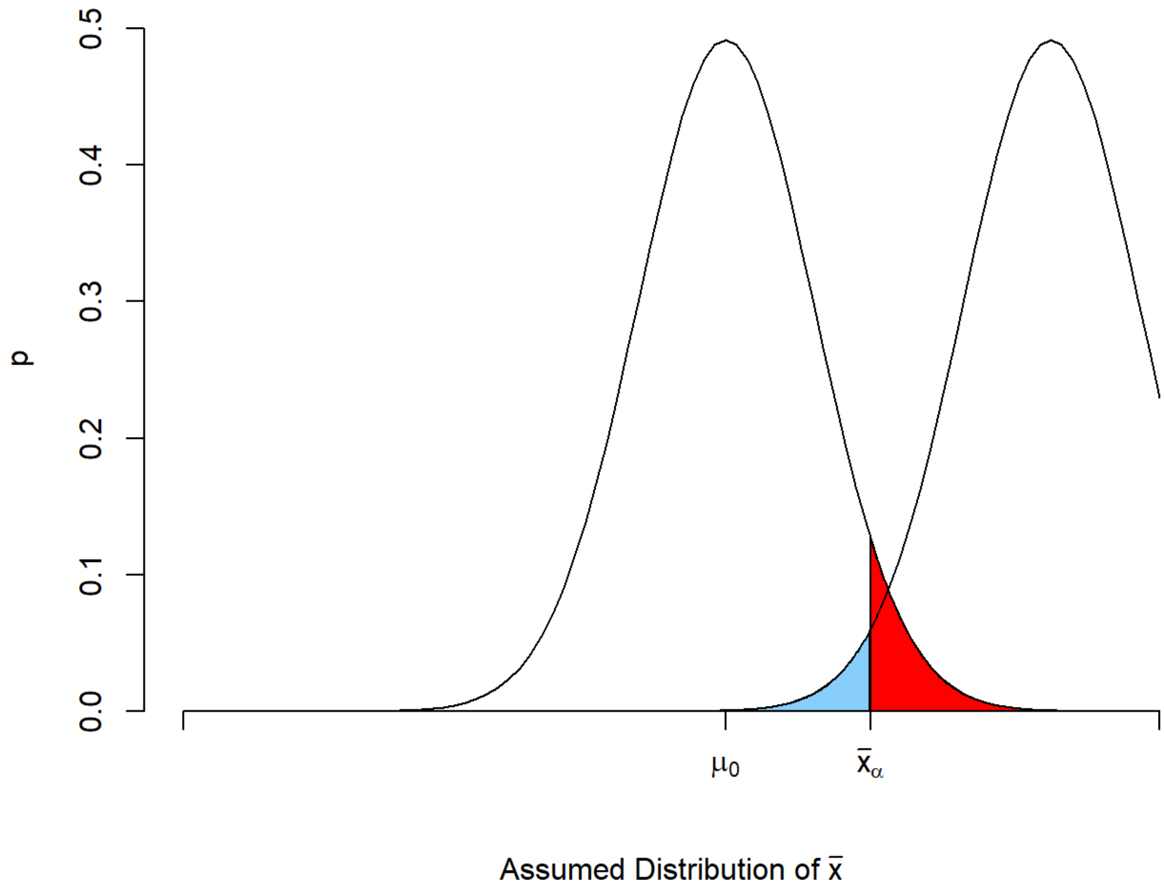
Power function: Se observa como bajo la nula toma probabilidades cercanas a cero, e inmediatamente bajo la alternativa su proabilidad va incrementando hasta tender a uno, i.e., la probabilidad de rechazar cuando la alternativa es cierta es cada vez mayor.

Code



Cambiando la H_A a 53, se observa como la probabilidad de incurrir al error tipo 2 disminuye.

Code

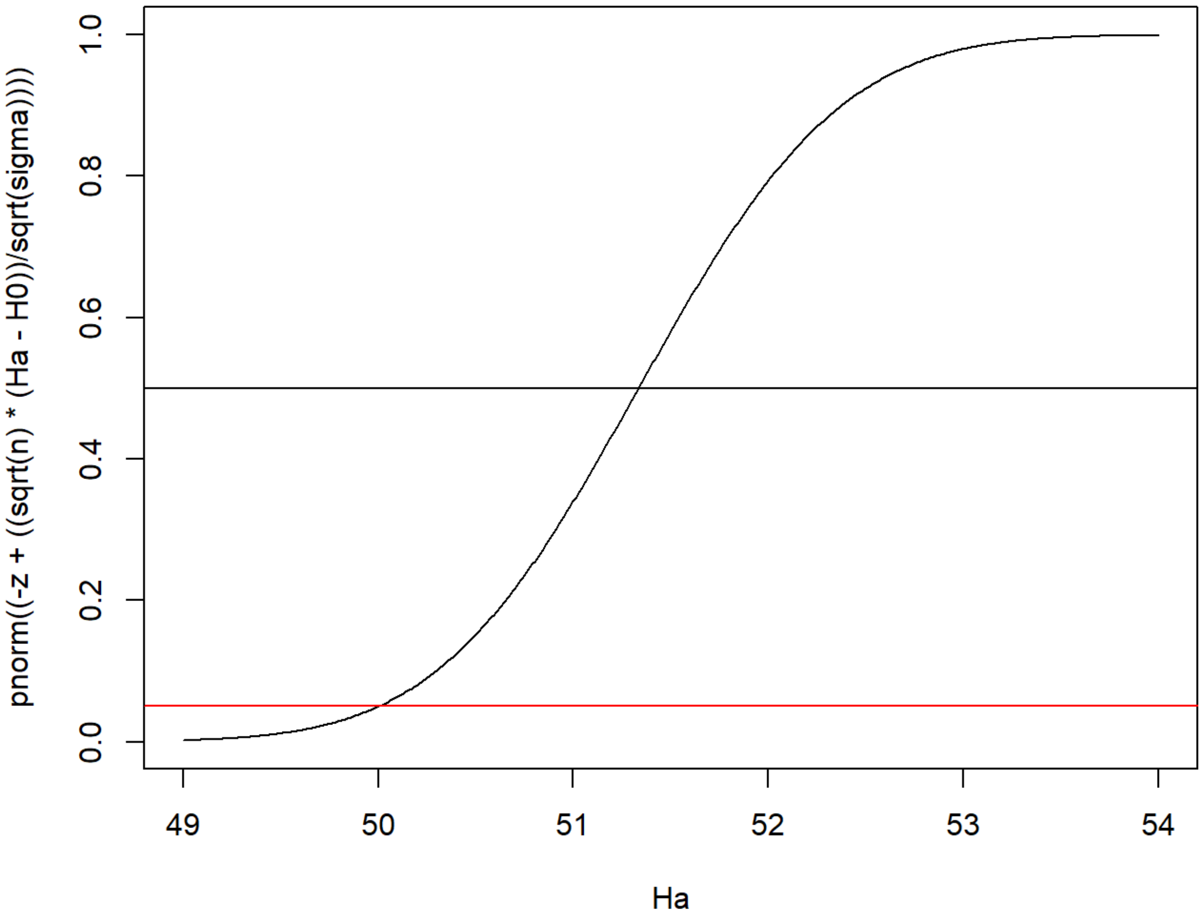


EJERCICIO Extra 4 Diap. 334 ———–

En la sección de ejercicios extras no se pretende pedir puntuación extra, sólamente se hace por cuestiones morales ya que más de una vez la Dra. ha mencionado que es de buen ejercicio replicar las notas. De esta manera, solamente se replican resultados y gráficos.

Gráfico de la potencia de la prueba En esta gráfica se muestra la potencia de la prueba para los distintos valores de μ tanto en la Hipótesis Alternativa como en la Nula. La cual nos indica cuando se rechaza la nula cuando no es plausible.

Code



Como se observa, para valores de μ menores al de la nula, la potencia es menor a .05.

EJERCICIO Extra 5 Diap. 340

En la sección de ejercicios extras no se pretende pedir puntuación extra, sólamente se hace por cuestiones morales ya que más de una vez la Dra. ha mencionado que es de buen ejercicio replicar las notas. De esta manera, solamente se replican resultados y gráficos.

Rendimiento de combustible y el deseo de los consumidores de verificar lo que dice el fabricante

Code

PASO 1. Revisar que la muestra provenga de una distribución normal.

Se realiza prueba de normalidad Shaphiro-Wilks,

Code

```
##
##  Shapiro-Wilk normality test
##
## data:  datavar$V2
## W = 0.94952, p-value = 0.2087
```

Se observa el esta´distico se acerca a uno, de esta manera, se puede decir que la muestra proviene de una normal.

Se contrasta con la prueba Jarquer-Bera

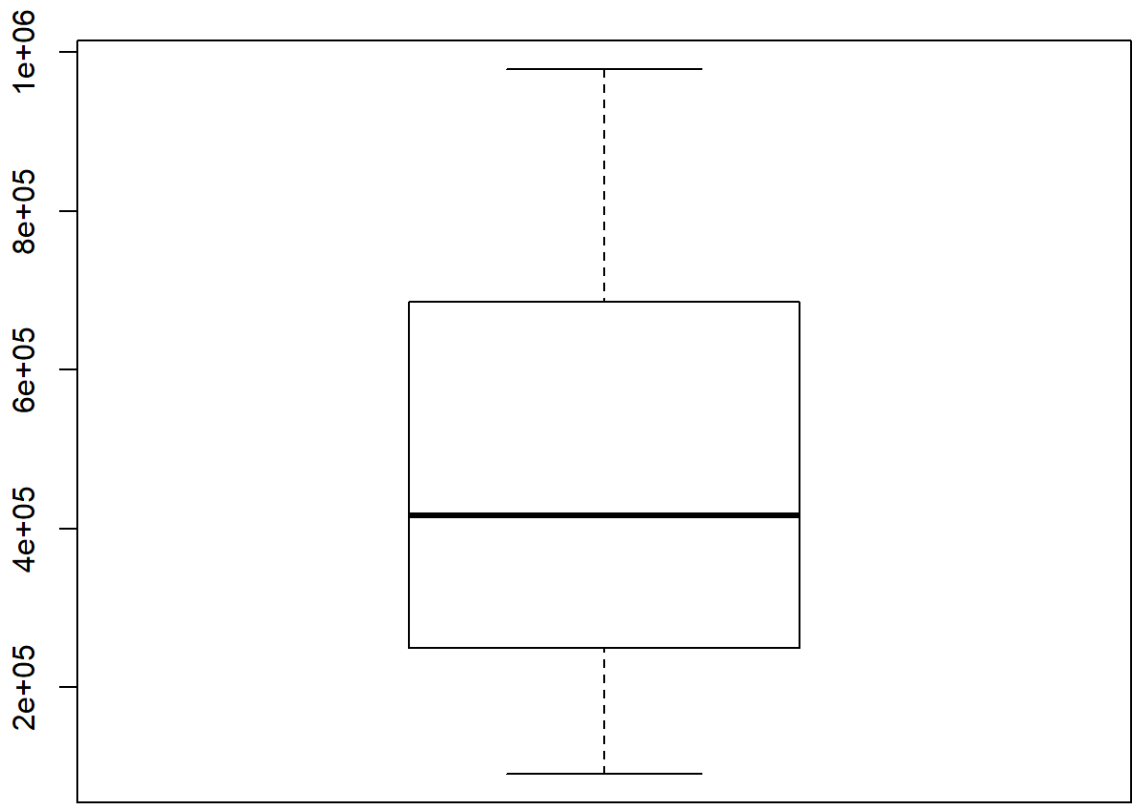
Code

```
##
## Title:
##  Jarque - Bera Normalality Test
##
## Test Results:
##   STATISTIC:
##     X-squared: 1.5894
##   P VALUE:
##     Asymptotic p Value: 0.4517
##
## Description:
##  Mon Nov 26 16:24:57 2018 by user: h_air
```

H_0 Normalidad, No se tiene evidencia suficiente para rechazar la H_0.

Ahora, se realiza un boxplot.

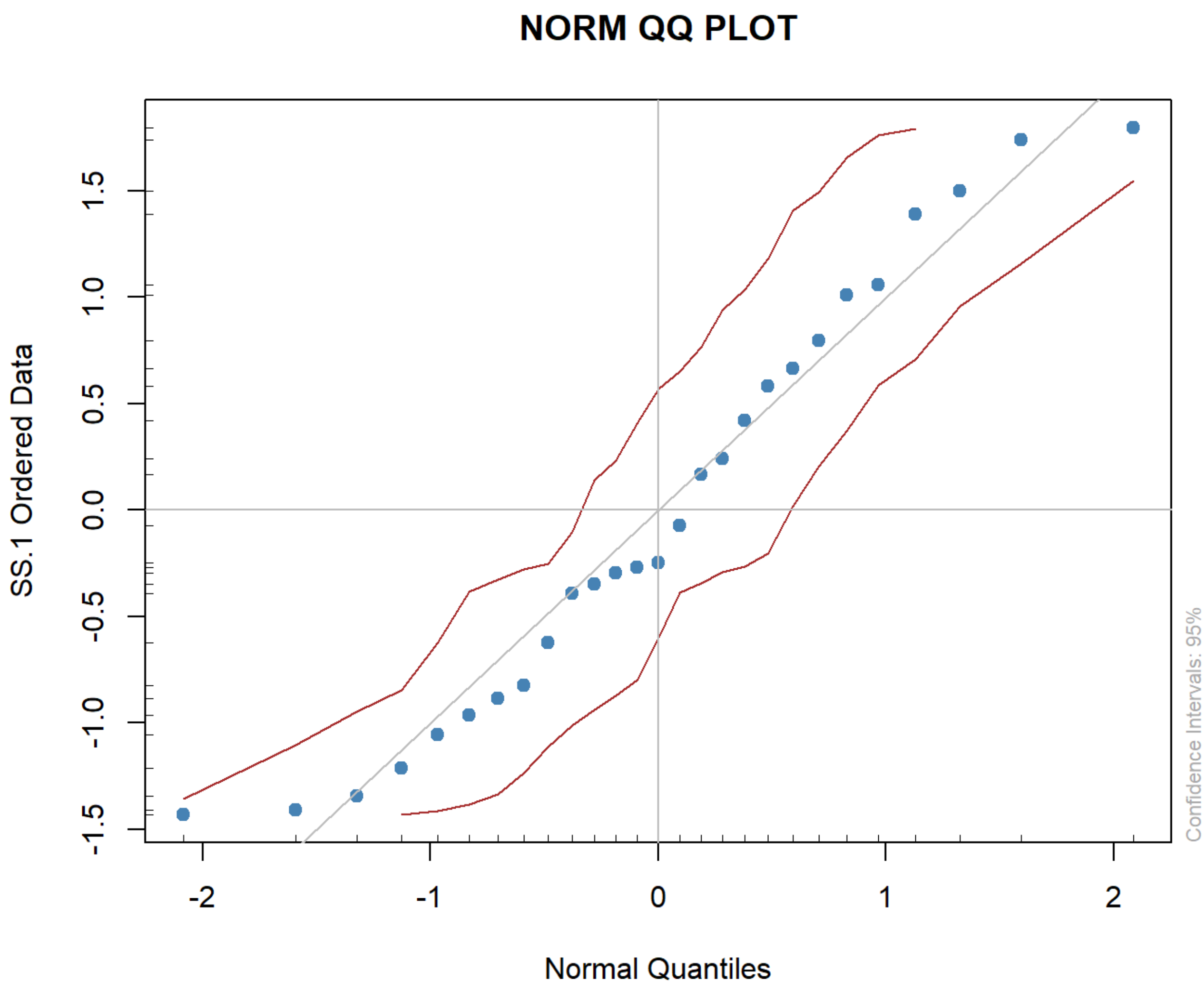
Code



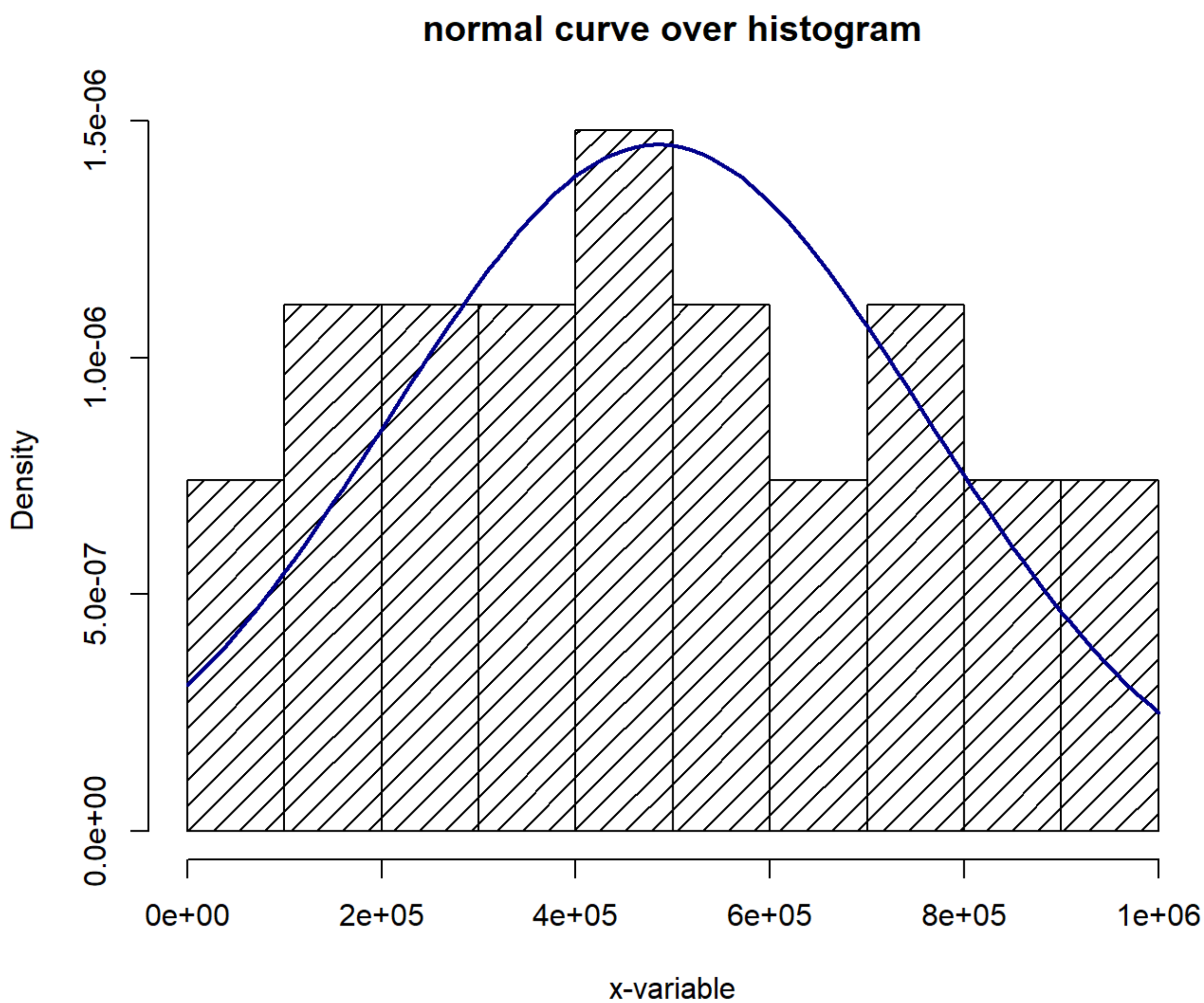
Se observa un poco de sesgo, cargada un poco a la derecha. Se observa como su valor medio se encuentra un poco cargado.

A continuación se reaiza un QQ-plot, para contrastar los cuantiles de la distribución empirica respecto a la teórica.

Code


NORM QQ PLOT
SS.1 Ordered Data
Normal Quantiles
Confidence Intervals: 95%
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
-2
-1
0
1
2

Code



Se observa que los datos tratan de aproximar a la normal, sin embargotanto en colas y ligeramente en el centro, no lo logran en su totalidad.

Se concluye que la muestra probiene de una distribución normal.De esta maner, conocemos la distribución del estadístico de prueba y podemos fijar regiones de rechazo y calcular p-valores dada las hipótesis que estemos considerando.

EJERCICIO Extra 6 Diap. 350

En la sección de ejercicios extras no se pretende pedir puntuación extra, sólamente se hace por cuestiones morales ya que más de una vez la Dra. ha mencionado que es de buen ejercicio replicar las notas. De esta manera, solamente se replican resultados y gráficos.

Potencia de la prueba para la varianza. $_0 : \sigma^2 \leq 0.0025$ vs $H_A : \sigma^2 > 0.0025$

Code

Variando cociente (DIAPOSITIVA 351)

Code