Tarea 5 Inferencia Estadística

Code

Hairo Ulises Miranda Belmonte

6 de Noviembre de 2018

Code

EJERCICIO 6

c. Determine mú de MV númericamente para el caso de que la media muestral sea 3.2

Code

El estimador de maxima verosimilitud tiene unvalor de:

Code

```
##
##
##                 Estimador MV
## ------------  -------------
## Mu estimado       3.024529
```

EJERCICIOS DIAPOSITIVAS 145

Diapositiva 145, inclusión a la variabilidad

Ejemplo: En una rama de la industria alimentaria, se realizan en forma rutinaria mediciones del contenido de calcio en comida para animales (mascotas). El m□etodo est□andar utiliza precipitaci□on de oxalato de calcio seguida de tritanio; es una t□ecnica que consume tiempo. Los resultados de 118 muestras (Heckman 1960), se dan en el archivo calcio.txt (columna 1). Ahora, con lo que hemos discutido hasta aqu□□, qu□e podemos decir a partir del comportamiento de estos valores muestrales?

Tomando los datos

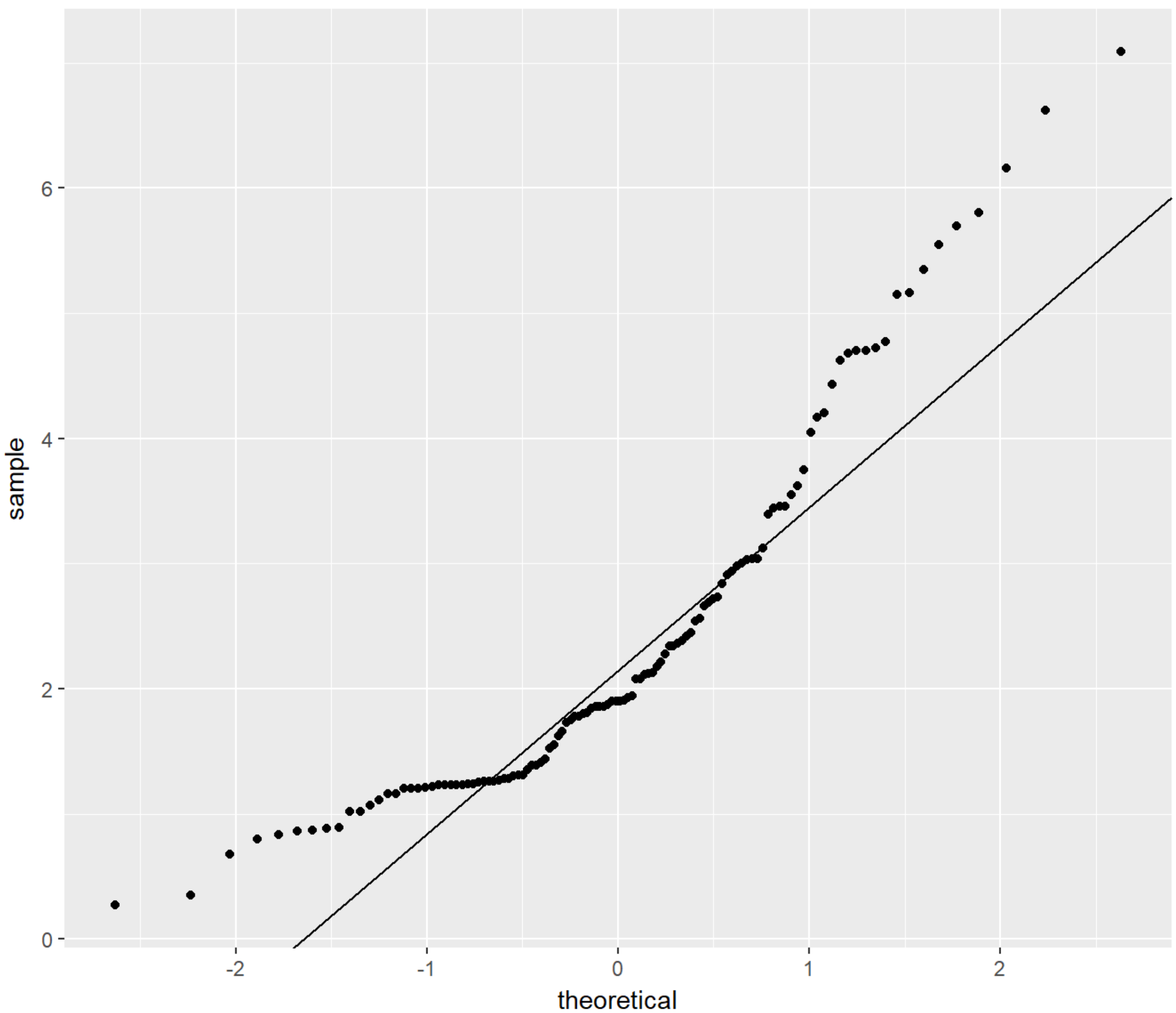
Code

Para iniciar, observaremos la distribución empirica contra una distribución normal teoríca. Para esto, utilizamos un qqplot.

Code

Medición de contenido de calcio en comida de mascotas con el método estándar el cual utiliza precipitaci□on de oxalato de calcio seguida de tritanio.

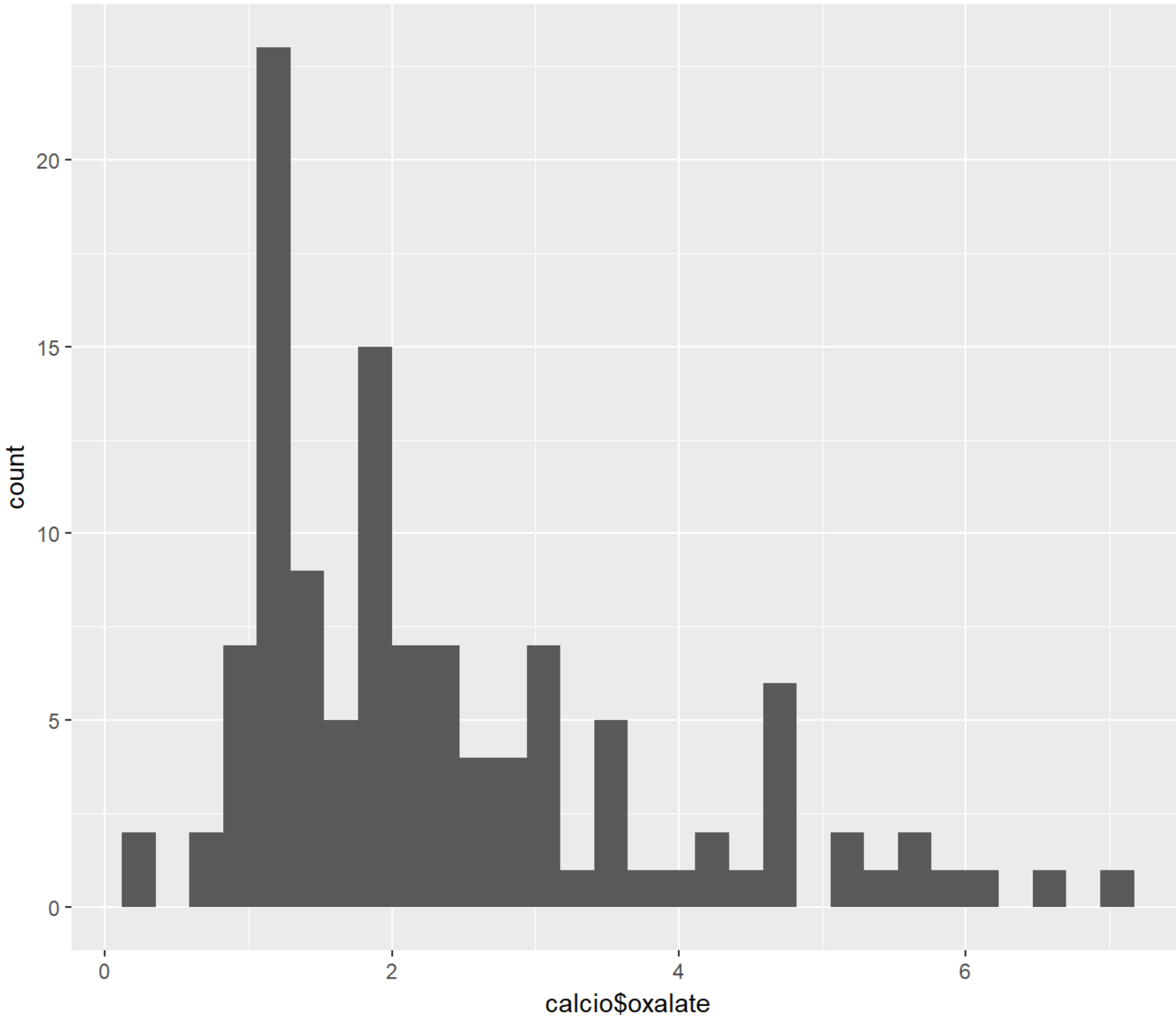
Code



Como se observa en la gráfica, la distribución muestral respecto a una teorica normal, tiene una distribución sesgada a la derecha (izquierda para quien observa), con algunos valores a tipicos, y colas pesadas. Por ultimo, la distribución empirica no se aproxima a una normal en el centro de su distribución.

Code

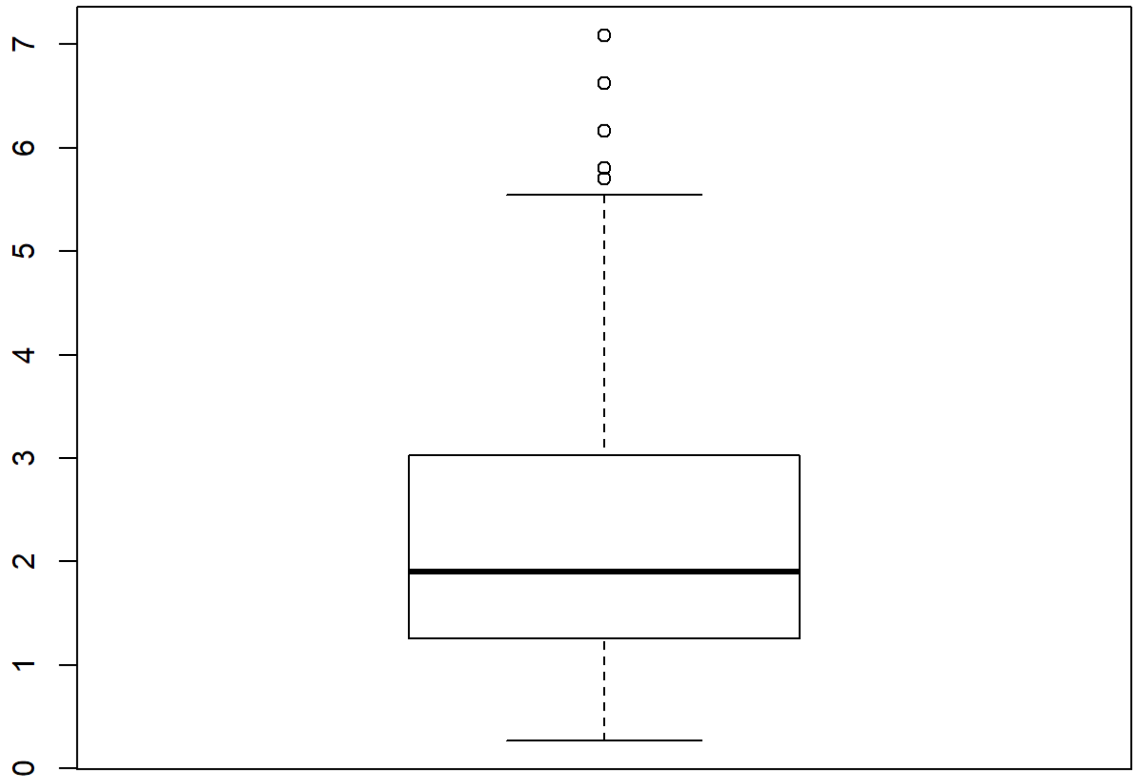
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



Así como se comento con el qqplot, la distribución empirica presenta sesgo, colas pesadas y, valores atipicos. Lo anterior se puede apreciar mejor con el histograma.

También, se realiza un boxplot.

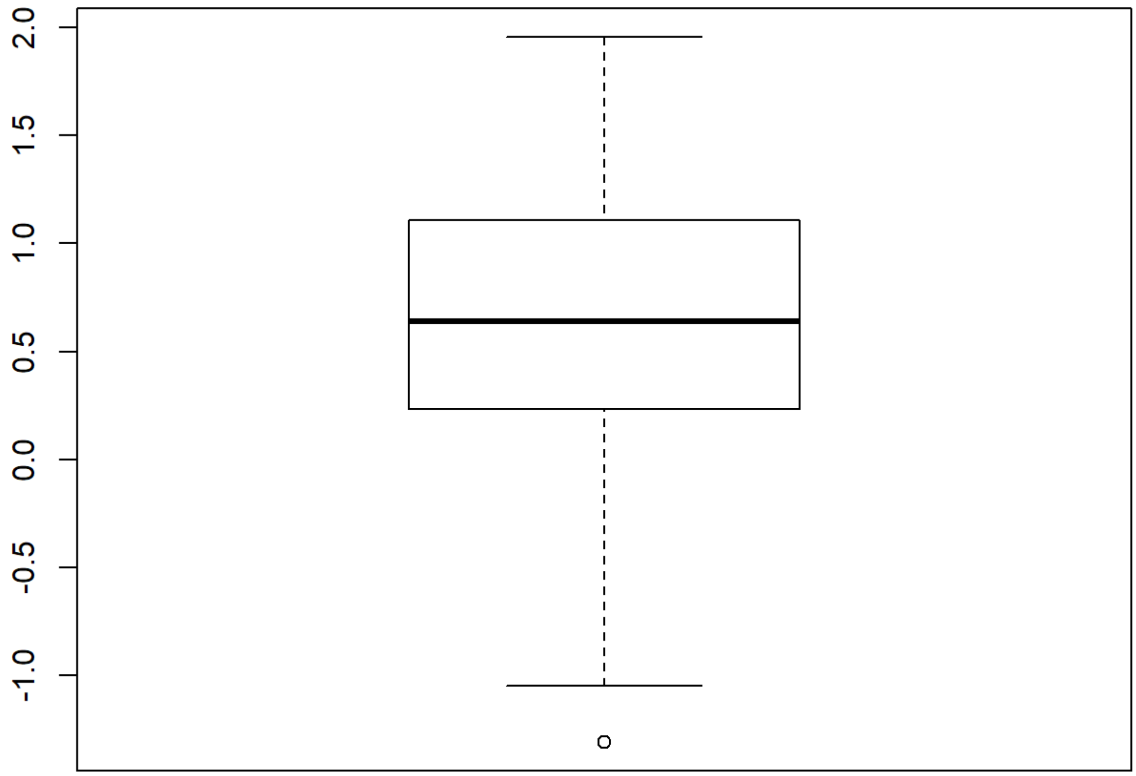
Code



El boxplot de la distribución empirica, es decir, los datos sobre el método estándar, argumenta lo anterior dicho sobre los valores a tipicos. Tambien se ve de manera marcada el sesgo a la derecha (izquiera para quien ve), y como su mediana se tira a la derecha

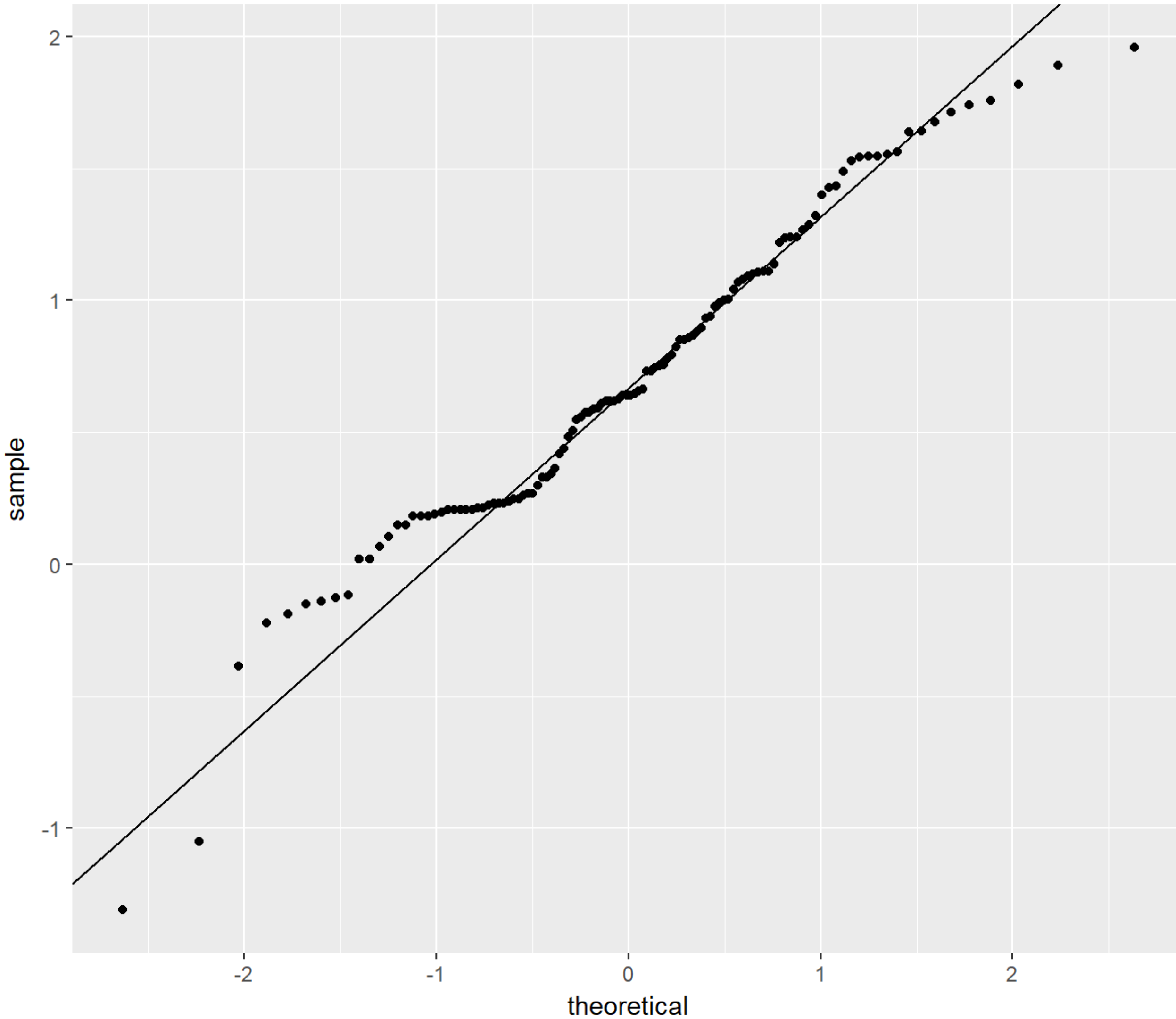
En las notas se indica que posiblemente la variable aleatoria sea una log normal. Utilizando la serie con logaritmos veremos si la distribución muesrral en logaritmos se distribuye como una normal.

Code



Se aprecia una mediana un poco a la izquierda, como la distribución se jala o se carga para el lado opuesto al no tener logaritmos,y por último, un valor a tipico perdido en la distribución.

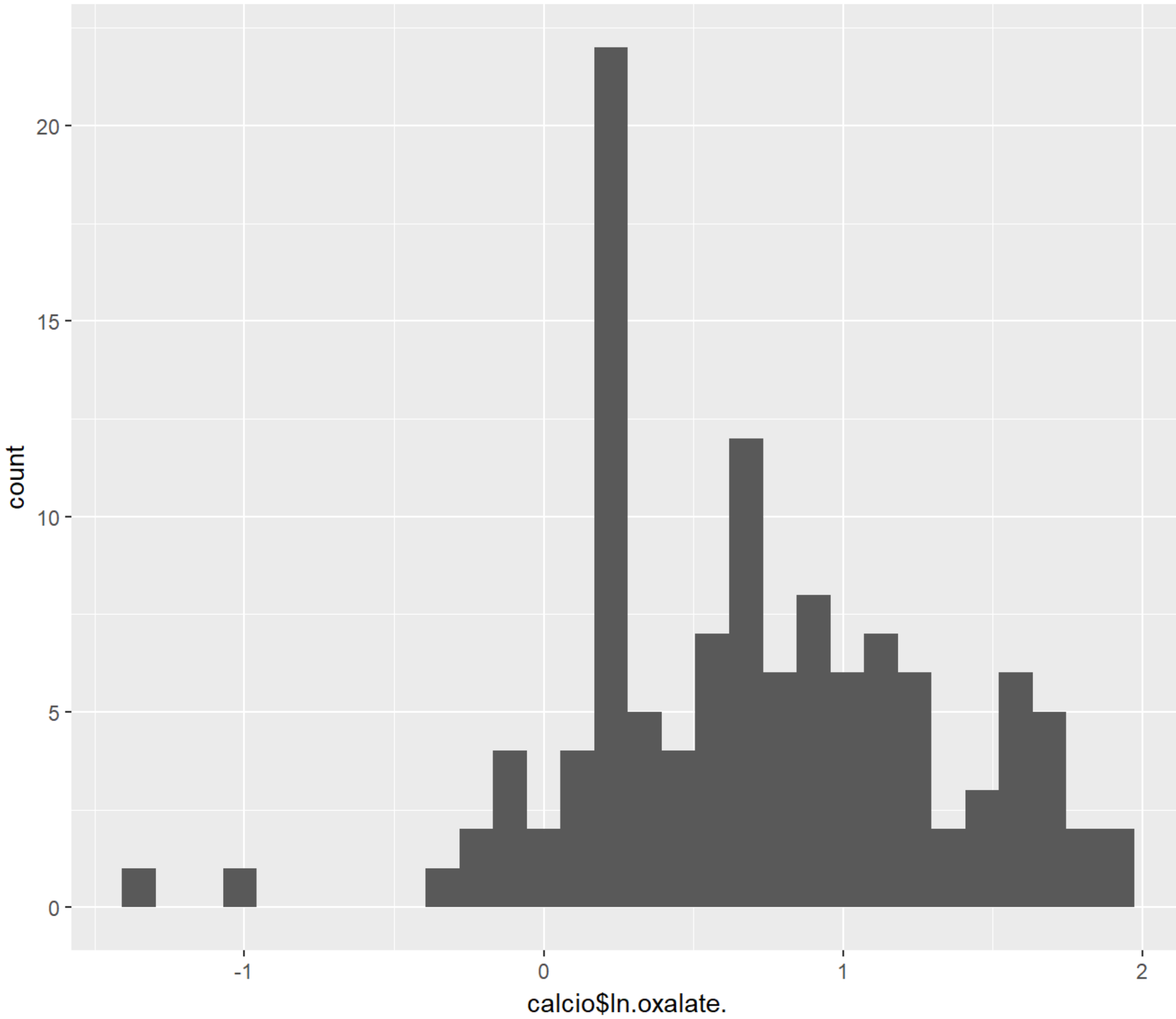
Code



Lo ya mencionado, se puede ver mejor con el qqplot. Los valores a tipicos en una de sus colas, y un poco de sesgo hacia el lado opuesto preevio de aplicar logaritmos. Es importante observar como el logaritmo de la muestra se distribuye como una normal.

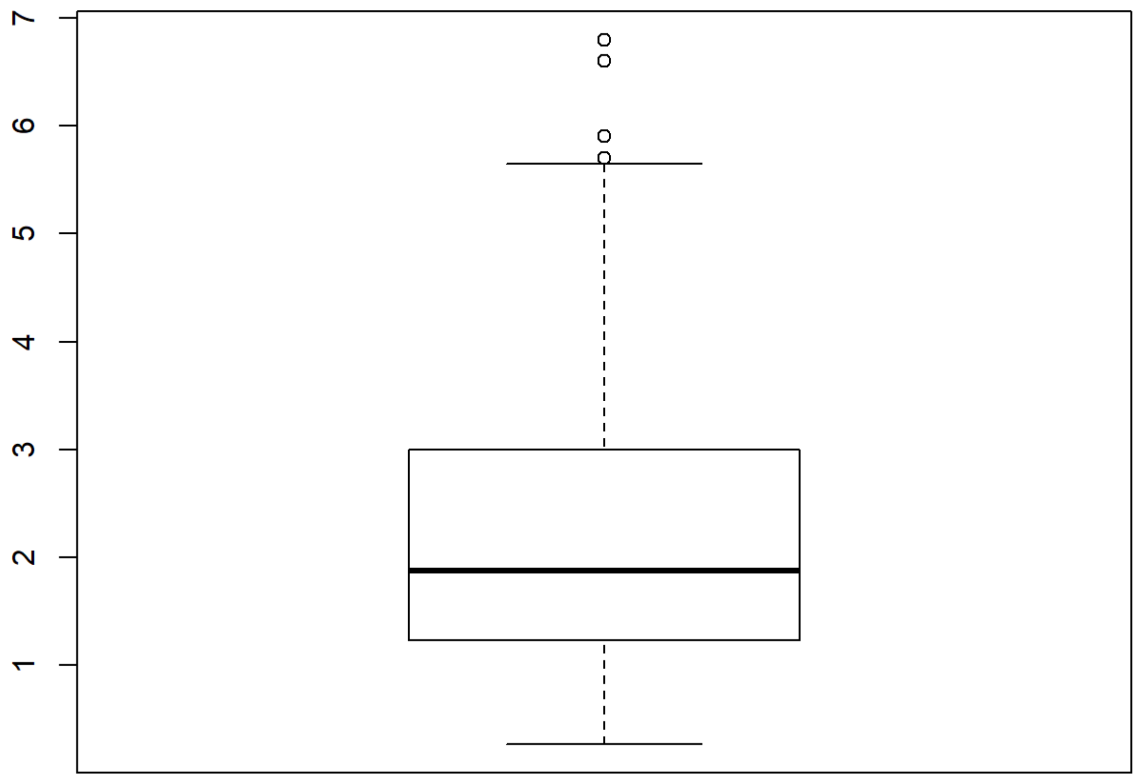
Code

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



El histograma muestra de mejor manera lo descito con el qq plot. Note que la cola de uno de los lados es más pesada de lo que se veía con el qqplot.

Code



Se observan datoa a tipico en el derecho de la distribución (cola), y unamediana cargada hacia el lado izquierdo

En conclusión la muestra sin logaritmos se distribuye como una lognnormal, pero la muestra con logaritmos es una normal.

EJERCICIOS DIAPOSITIVAS 179

IC PARA CALCIO PARA LOS DATOS DE CALCIO

Code

```
## 
## 
##          Intervalos de confianza
## ------  ------------------------
## LOWER                   2.125530
## UPPER                   2.649555
```

Considerando los logaritmos y consideramos pivote que sigue una t de student.

Code

t de student con n-1 grados de libertad

Code

```
##
##
##                         IC (log muestra)
## --------------------  -----------------
## LOWER                         0.5894296
## UPPER                         0.8058099
## Error de estimación           0.1081901
```

Tomando transformación inversa a los logaritmos del intervalo de confianza

Code

```
##
##
##          IC (exponencial log muestra)
## ------  -----------------------------
## LOWER                        1.802960
## UPPER                        2.238509
```

En este □ultimo intervalo deber□□amos notar que la relaci□on entre la media de la variable original y la de su transformaci□on logar□□tmica est□a afectada por una constante que aqu□□ no aparece. Recuerda la relaci□on entre medias y varianzas que hemos establecido al nal del cap□□tulo 3, para la Normal y Lognormal. Trata de incorporar esta informaci□on para entender mejor la diferencia num□erica entre estos intervalos.

sabemos que los momentos de una normal y lognormal son distintos, ya que si una variable se distribuye como una lognormal, su generadora de momentos no existe. Pero sabemos como es la generadora de una normal, de esta forma la variable que se distribuye como una lognormal al sacarle la esperanza es como si plantearas la generadora pero con una e a la xt, sin t. lo cual el esperados es la generadora de momentos de una normal pero evaluando t en el momento que se quiera de la log normal.

Ahora comparamos las muestras del archivo txt. Calculamos la diferencia entre medias de los dos métodos distintos. El método con oxalte será el método estándar, el Flame el alternativo.

Code

```
##
##
##                        Dif. entre dos Poblaciones
## -------------------  ---------------------------
## Diff. Medias                           0.0355932
## Cociente Varianzas                     1.0430312
```

Se oserva que el estadístico de la diferencia de medias entre poblaciones, es relativamente pequeño. Lo cual, ya era de esperarse dado a la descripción de los datos (i.e, vienen de poblacioneslo más homogeneas posibles). Entonces, como se pretendia que la muestras fueran lo controladas pero independientes, se toma un críterio de variabilidad. El cociente entre la varianza de la muestra estandar y la alternativa es ligeramente mayor a uno. Sin embargo, decir que la muestra de la población estandar es más variable seria muy precipitado.

A continuación, se construye un intervalo de confianza para la diferencia de medias

Code

```
##
##
##              Dif. entre dos Poblaciones IC
## ---------  -----------------------------
## IC Lower                         1.710507
## IC upper                        -1.639321
```

Para seguir jugando un poco con los datos, se simula la proporción de veces que el intervalo de confianza contiene al valor de la hipótesis sobre la Nula. Recordemos que dicho valor es fijo,o dado, y son los intervalos de confianza la que es una variable aleatoria, la cual en cada simmulación estarán cambiando.

Code

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

Code

```
##
##
##                   prop.veces IC contiene Ho
## --------------  --------------------------
## 100%(1-alpha)                          100
```

El 98% de las veces el IC captura el valor de la nula.

A continuación, realizamos el mismo ejercicio pero utilizando el p-valor

Code

```
## [1] 0.99
```

Code

```
##
##
##                  p-valor
## --------------  --------
## 100%(1-alpha)         99
```

El 99 porciento de la veces el intervalo de confianza contiene al valor de la hipotesis nula.

EJERCICIOS BOOTSTRAP

Se pretende realizar un ejercicios computacional y estadístico utilizando bootstrap. Cuyo fin es estimar los errores estandar de algun estadístico. El caso de Intervalos de confienza con método bootstrap queda pediente parallel el siguiente tema.

Se utiliza los datos del archivo txt.

La media muestral para los datos del método estandar es:

Code

```
## [1] 1.9
```

Code

```
##
##
##                    Error estandar
## ---------------  ---------------
## Mediana                1.9250000
## Error estandar         0.4864468
```

Se observa como su media se aproxima a la verdadera cuando se utiliza em método de bootstra.