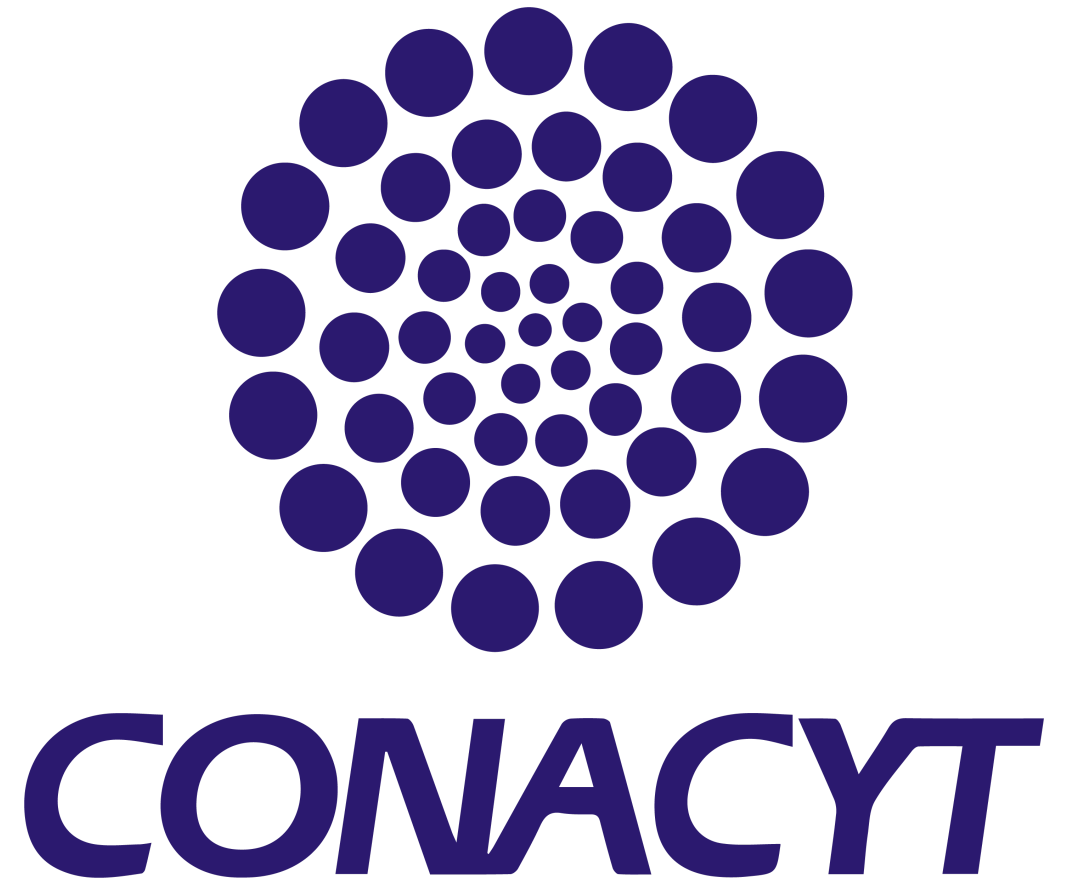


Introducción a la estadística

Dra. Stephanie Hereira Pacheco

CTCB, UATx

23-02-2023



Contenido

- Pruebas paramétricas y supuestos estadísticos.
 - t de student
- Pruebas no paramétricas y supuestos estadísticos.
 - U de Mann Whitney
 - Prueba de Wilcoxon.

Pruebas paramétricas y supuestos estadísticos

Se conoce como estadística paramétrica a aquella que se basa en el muestreo de una población con una distribución conocida y con parámetros fijos.

Los supuestos de las pruebas paramétricas en general son:

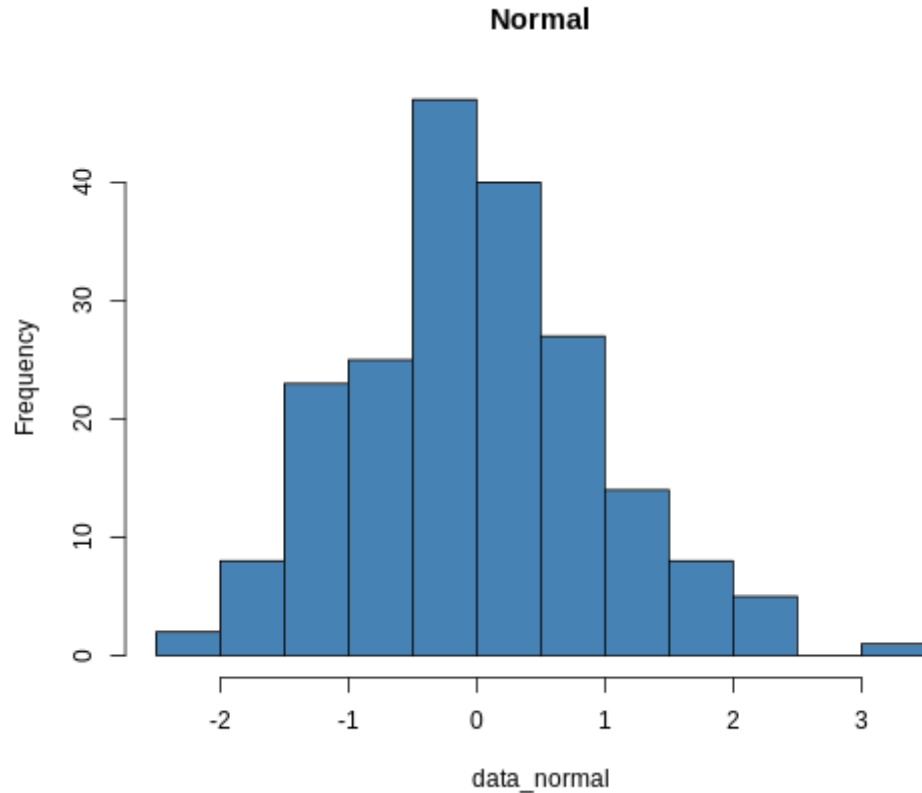
- Distribución conocida (**normal**): visual y pruebas numéricas.
- Homocedasticidad: visual y pruebas numéricas.
- Otros: tamaño de la muestra, variables cuantitativas o continuas, outliers, aleatoriedad, independencia de las observaciones, linealidad.

.

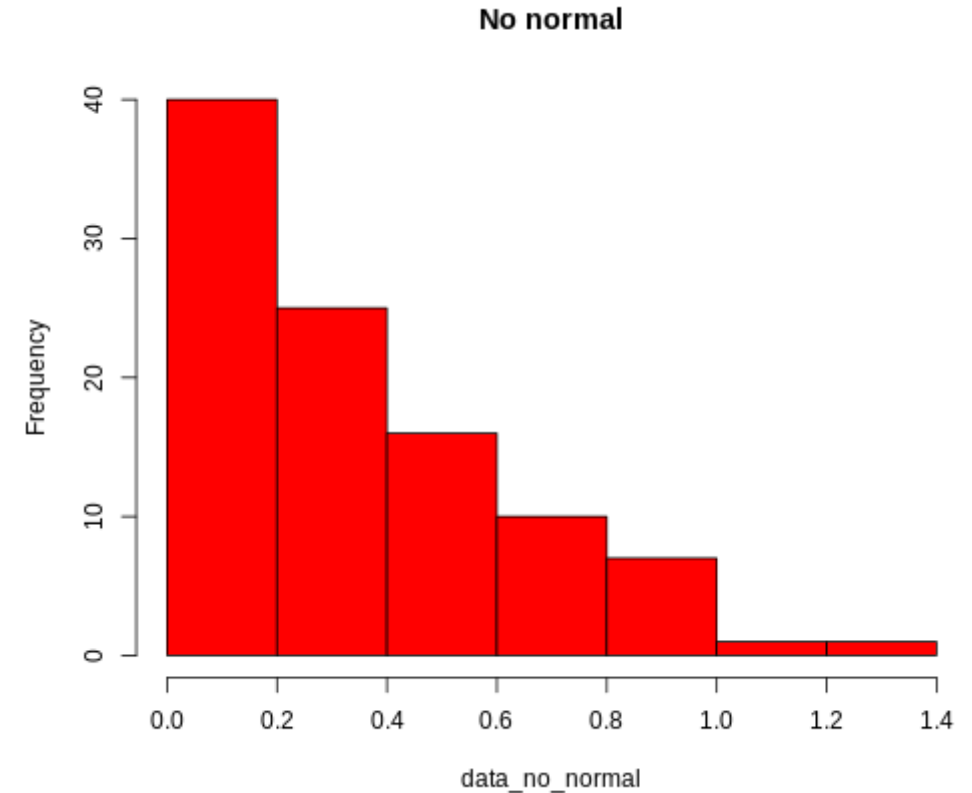
Cada tipo de prueba paramétrica tiene sus propios supuestos

Distribución normal: métodos visuales

```
data_normal<- rnorm(200)  
hist(data_normal, col='steelblue', main='Normal')
```

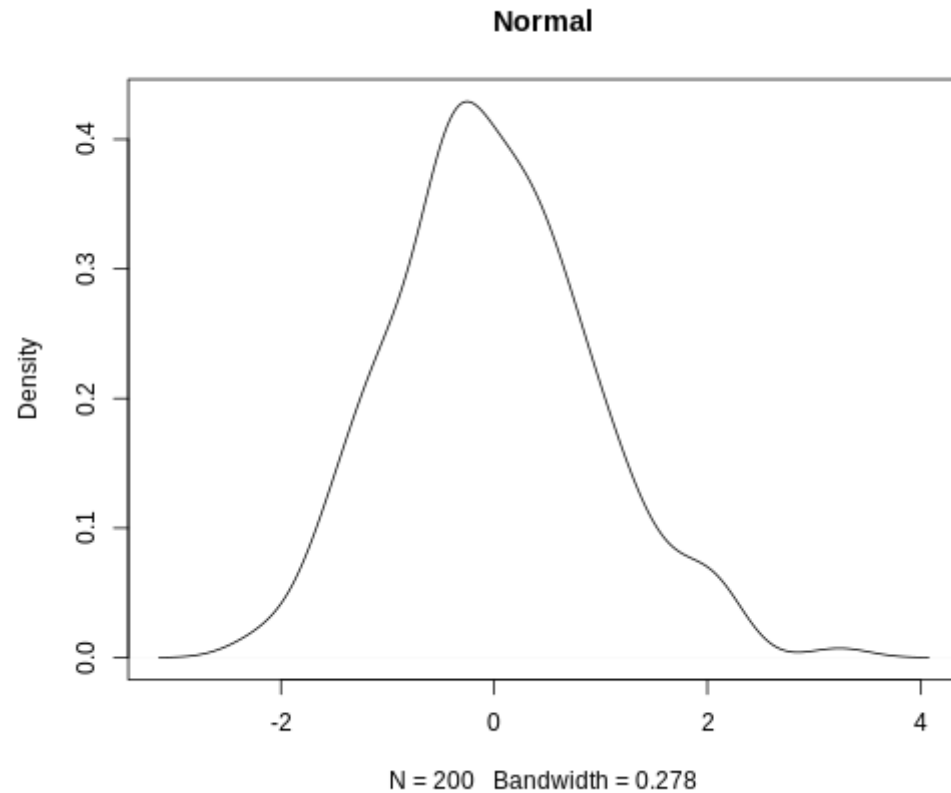


```
data_no_normal<- rexp(100, rate=3)  
hist(data_no_normal, col='red', main='No normal')
```

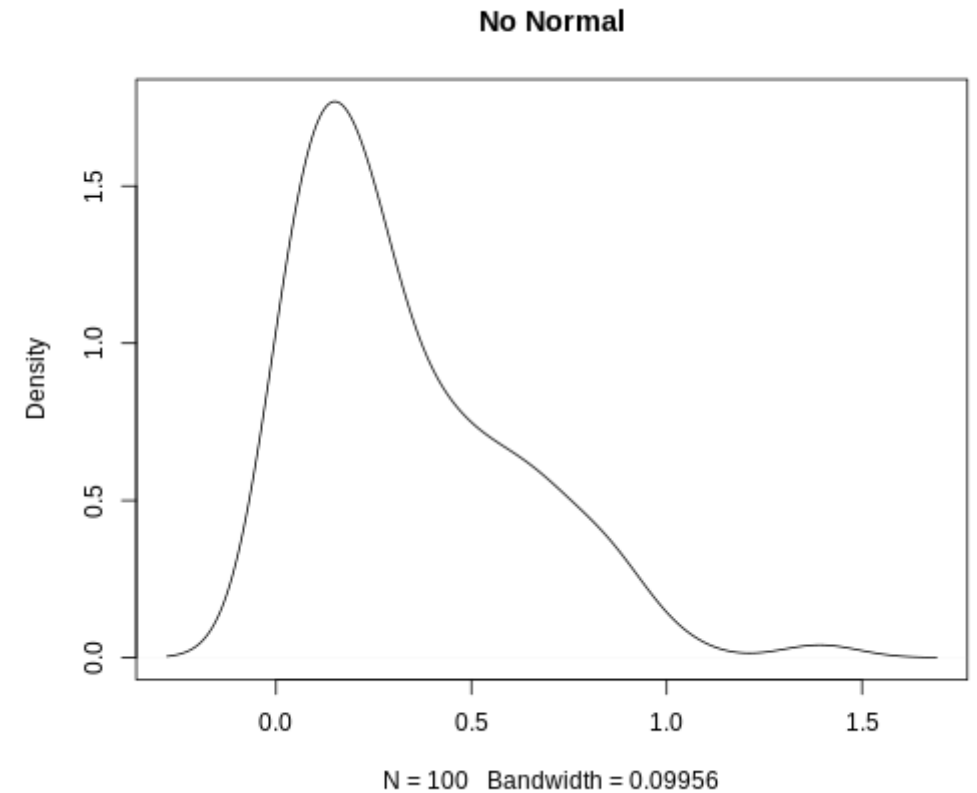


Distribución normal: métodos visuales

```
plot(density(data_normal), main="Normal")
```



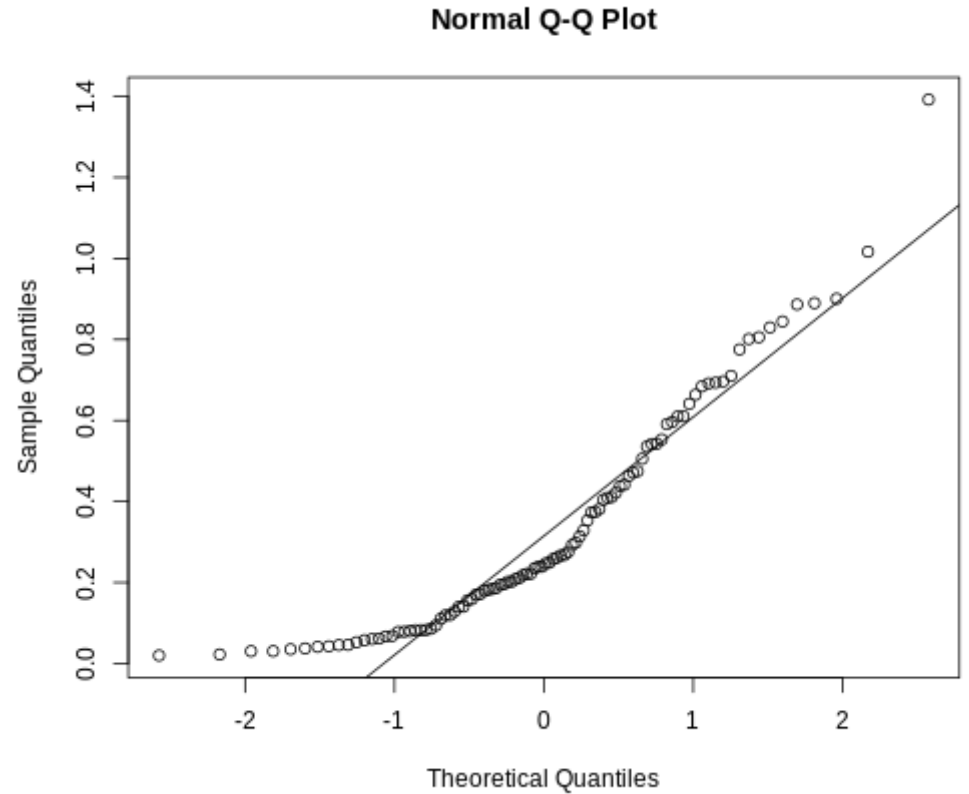
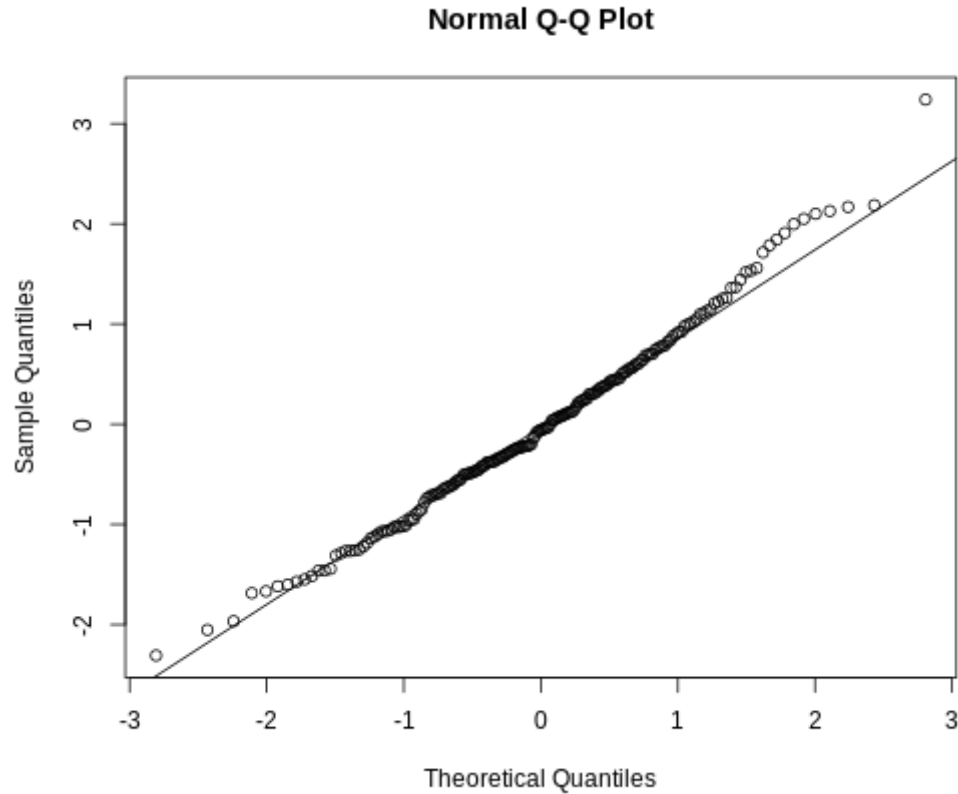
```
plot(density(data_no_normal), main="No Normal")
```



Distribución normal: métodos visuales

```
qqnorm(data_normal)  
qqline(data_normal)
```

```
qqnorm(data_no_normal)  
qqline(data_no_normal)
```



Distribución normal: métodos numéricos

```
shapiro.test(data_normal)
```

```
##  
##      Shapiro-Wilk normality test  
##  
## data:  data_normal  
## W = 0.99076, p-value = 0.2298
```

```
ks.test(data_normal, "pnorm")
```

```
##  
##      One-sample Kolmogorov-Smirnov test  
##  
## data:  data_normal  
## D = 0.053155, p-value = 0.6243  
## alternative hypothesis: two-sided
```

```
shapiro.test(data_no_normal)
```

```
##  
##      Shapiro-Wilk normality test  
##  
## data:  data_no_normal  
## W = 0.89131, p-value = 5.641e-07
```

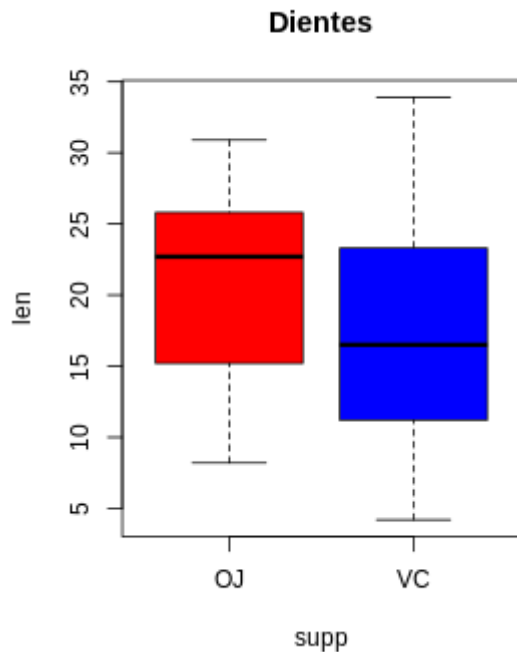
```
ks.test(data_no_normal, "pnorm")
```

```
##  
##      One-sample Kolmogorov-Smirnov test  
##  
## data:  data_no_normal  
## D = 0.50783, p-value < 2.2e-16  
## alternative hypothesis: two-sided
```

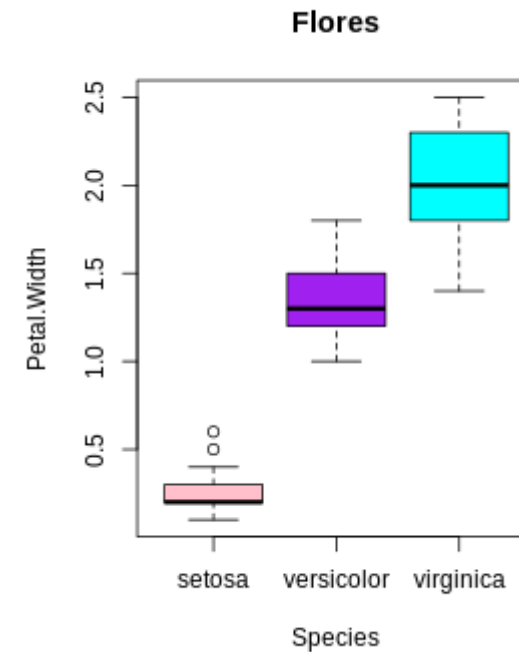

Probando heterocedasticidad

- Métodos visuales = Pruebas estadísticas de comparación y modelos lineales

```
data("ToothGrowth")  
boxplot(len ~ supp, data=ToothGrowth, col=c(
```



```
data("iris")  
boxplot(Petal.Width ~ Species, data=iris, co
```



Probando heterocedasticidad

- Métodos visuales = Pruebas estadísticas de comparación y modelos lineales

```
aggregate(len ~ supp, data = ToothGrowth, va
```

```
##      supp      len
## 1      OJ 43.63344
## 2      VC 68.32723
```

Ratio

```
8.26 / 6.60
```

```
## [1] 1.251515
```

```
aggregate(Petal.Width ~ Species, data = iris
```

```
##      Species Petal.Width
## 1      setosa  0.01110612
## 2 versicolor  0.03910612
## 3  virginica  0.07543265
```

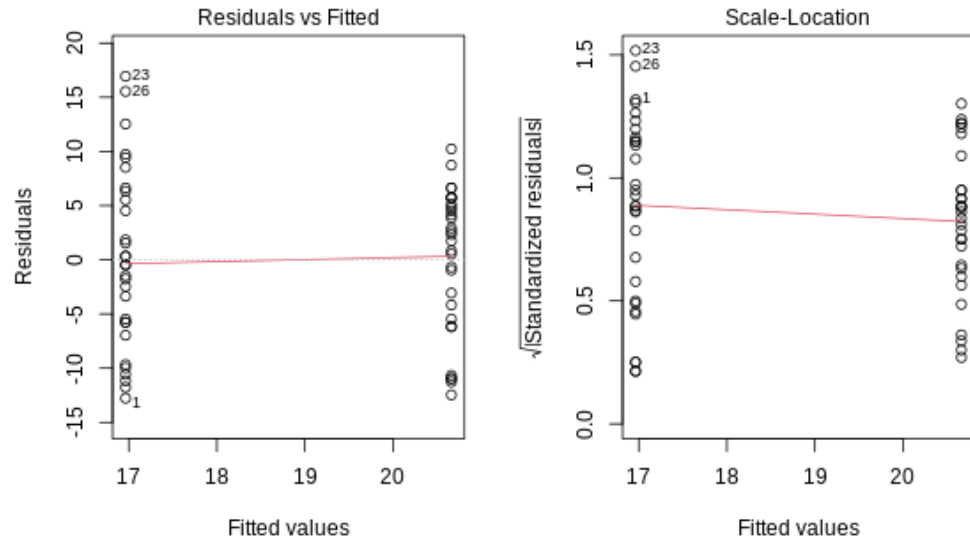
Ratio

```
r1<-0.03910612 / 0.01110612
r2<-0.07543265 / 0.1110612
r3<-0.07543264 / 0.03910612
cbind(r1,r2,r3)
```

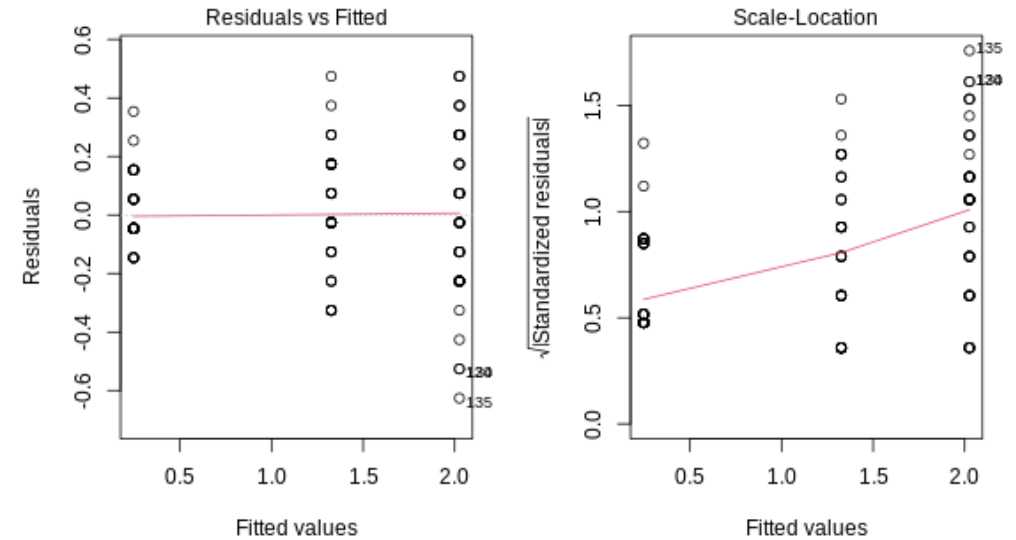
```
##      r1      r2      r3
## [1,] 3.535432 0.6791989 1.928922
```

Probando heterocedasticidad

```
m1<-lm(len ~ supp, data=ToothGrowth)
par(mfrow = c(1, 2))
plot(m1, which=c(1,3))
```



```
m2<-lm(Petal.Width ~ Species, data=iris)
par(mfrow = c(1, 2))
plot(m2, which=c(1,3))
```



Probando heterocedasticidad

Prueba para dos niveles = F

```
var.test(len ~ supp, data = ToothGrowth)
```

```
##  
##      F test to compare two variances  
##  
## data:  len by supp  
## F = 0.6386, num df = 29, denom df = 29, p-value = 0.2331  
## alternative hypothesis: true ratio of variances is not equal to 1  
## 95 percent confidence interval:  
##  0.3039488 1.3416857  
## sample estimates:  
## ratio of variances  
##           0.6385951
```

```
lmtest::bptest(m2) #sobre un modelo
```

```
##  
##      studentized Breusch-Pagan test  
##  
## data:  m2  
## BP = 25.099, df = 2, p-value = 3.546e-06
```

- Otras pruebas: Prueba de Levene, Goldfeld-Quand, White, entre otras.

PRUEBAS PARAMÉTRICAS

| VENTAJAS | DESVENTAJAS |
|---|--|
| - Sensibles a rasgos de los datos recolectados | - Más complicadas de calcular |
| - Estimaciones probabilísticas más exactas | - Solo se pueden aplicar si se cumplen sus supuestos |
| - Tienen una mayor eficiencia y poder estadístico | - Los datos que se pueden observar son limitados |

PRUEBAS PARAMÉTRICAS

| TIPO | PRUEBA |
|------------------------------|------------------------|
| Comparación de 2 grupos | t de Student/Welch |
| Comparación de >2 grupos | Anova |
| Correlación de dos variables | Coeficiente de Pearson |
| Variables cualitativas | Prueba de Z |

t de Student

- Desarrollada por William Sealy Gosset.
- La prueba t permite comparar las medias de una muestra con la media teórica de una población, o comparar dos poblaciones.
- Una de las características de la prueba de student, es que permite la hipótesis alternativa de ver si dos medias son diferentes o si, una media es mayor, o menor que otra.
- Esta prueba tiene variantes dependiendo de nuestros datos.
- Para ser utilizada es necesario que se cumplan los siguientes supuestos:
 - Muestreo aleatorio
 - Los datos y errores deben tener una distribución normal

t de Student una muestra

En la prueba t de Student para una muestra, se usa la siguiente fórmula:

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}}$$

Dónde t es el estadístico, \bar{x} es la media muestral, s es la desviación estándar muestral, y n es el tamaño de la muestra.

El estadístico t posee un valor de p asociado dependiendo de los grados de libertad de la prueba.

Mide qué tan inusual es la media de la muestra observada de la media de la muestra hipotética. Lo hace midiendo qué tan lejos está la media de la muestra observada de la media de la población hipotética en términos de errores estándar.

Ejemplo de una muestra

Con R

```
data<- data.frame(Congelador=1:10,  
                  Temperatura=c(-2.14,-0.8,  
                                -2.75,-2.58,  
                                -2.26,-2.46,  
                                -1.33,-2.85,  
                                -0.93,-2.01))  
shapiro.test(data$Temperatura)$p.value
```

```
## [1] 0.1823153
```

Manualmente

```
prom = mean(data$Temperatura); desves = sd(data$Temperatura)  
t=(prom-mu)/(desves/sqrt(10))  
t;pt(t, df=9)
```

```
## [1] -8.572431
```

```
## [1] 6.345623e-06
```

```
t.test(data$Temperatura, mu = 0, alternative = "less")
```

```
##  
## One Sample t-test  
##  
## data: data$Temperatura  
## t = -8.5724, df = 9, p-value = 6.346e-06  
## alternative hypothesis: true mean is less than 0  
## 95 percent confidence interval:  
##      -Inf -1.580971  
## sample estimates:  
## mean of x  
##      -2.011
```

t de Student para medidas repetidas

En la prueba t de Student para medidas repetidas se usa la siguiente fórmula:

$$t = \frac{\bar{x}_d - \mu_d}{s_d / \sqrt{n}}$$

Dónde t es el estadístico, \bar{x}_d es la media muestral de la diferencia de las medidas, s_d es la desviación estándar muestral, y n_d es el tamaño de la muestra.

- Data ejemplo:

```
data_pareada<- data.frame(Pan=1:10,  
                          Temp_pre=c(20.83,19.72,19.64,20.09,22.25,  
                                     20.83,21.31,22.50,21.17,19.57),  
                          Temp_post=c(100.87,98.58,109.09,121.83,122.78,  
                                     111.41,103.96,121.81,127.85,115.17))  
  
data_pareada$Diferencia<- data_pareada$Temp_post - data_pareada$Temp_pre
```

Ejemplo muestras pareadas

| Pan | Temp_pre | Temp_post | Diferencia |
|-----|----------|-----------|------------|
| 1 | 20.83 | 100.87 | 80.04 |
| 2 | 19.72 | 98.58 | 78.86 |
| 3 | 19.64 | 109.09 | 89.45 |
| 4 | 20.09 | 121.83 | 101.74 |
| 5 | 22.25 | 122.78 | 100.53 |
| 6 | 20.83 | 111.41 | 90.58 |
| 7 | 21.31 | 103.96 | 82.65 |
| 8 | 22.50 | 121.81 | 99.31 |
| 9 | 21.17 | 127.85 | 106.68 |

Ejemplo muestras pareadas

Con R

```
shapiro.test(data_pareada$Diferencia)
```

```
##  
##      Shapiro-Wilk normality test  
##  
## data:  data_pareada$Diferencia  
## W = 0.93611, p-value = 0.5106
```

Manualmente

```
prom = mean(data_pareada$Diferencia); desves =  
t=(prom-mu)/(desves/sn)  
t;pt(t, df=9, lower.tail = FALSE)
```

```
## [1] 29.96903
```

```
## [1] 1.253345e-10
```

```
t.test(data_pareada$Temp_post, data_pareada$Temp_pre)
```

```
##  
##      Paired t-test  
##  
## data:  data_pareada$Temp_post and data_pareada$Temp_pre  
## t = 29.969, df = 9, p-value = 1.253e-10  
## alternative hypothesis: true difference in means is greater than 0  
## 95 percent confidence interval:  
##      86.88337      Inf  
## sample estimates:  
## mean of the differences  
##  
##      92.544
```

t de Student para dos grupos

En la prueba t de Student para dos grupos se usa la siguiente fórmula:

$$t_{obs} = \frac{\bar{x}_1 - \bar{x}_2 - \Delta}{\sqrt{s^2/n_1 + s^2/n_2}}$$

Dónde t_{obs} es el estadístico, $\bar{x}_{1,2}$ son los promedios de los grupos y Δ es el promedio teórico de la distribución nula de la diferencia de promedios (usualmente se usa 0), el s^2 viene dado por:

$$s^2 = \frac{\Sigma(x_1 - \bar{x}_1)^2 - \Sigma(x_2 - \bar{x}_2)^2}{n_1 + n_2 - 2} = \frac{s_1^2(n_1 - 1) - s_2^2(n_2 - 1)}{n_1 + n_2 - 2}$$

En este caso los grados de libertad se calculan :

$$df = (n_1 - 1) + (n_2 - 1) = n_1 + n_2 - 2$$

t de Student para dos grupos

- Debemos tener en cuenta si hay dependencia o no de los grupos y si la varianza es igual.
 - Cuando los grupos son independientes, se supone que proceden de diferentes poblaciones que tienen la misma varianza.
 - La t de muestras independientes es una medida de la diferencia de las medias de los dos grupos.
 - La distribución de la que se muestrea la diferencia de medias no es de una sola población sino de la combinación de las dos poblaciones.
- Por ejemplo, tengo una población con promedio de 10 y otra de 7, varianzas de 4 y 4, entonces mi combinación tendrá un promedio de 3 (diferencia) y varianza de 8.

Ejemplo para dos grupos

```
freidora<- data.frame(  
  Oster1=c(4.72,7.40,3.50,3.85,4.47,  
           4.09,6.34,3.30,7.13,4.99),  
  Oster2=c(9.19,11.44,9.64,12.09,10.80,  
           12.71,10.04,9.06,6.31,9.44))  
ro1<- freidora$Oster1-mean(freidora$Oster1)  
ro2<- freidora$Oster2-mean(freidora$Oster2)  
  
shapiro.test(c(ro1,ro2))
```

```
##  
##      Shapiro-Wilk normality test  
##  
## data:  c(ro1, ro2)  
## W = 0.9483, p-value = 0.342
```

| Oster1 | Oster2 |
|--------|--------|
| 4.72 | 9.19 |
| 7.40 | 11.44 |
| 3.50 | 9.64 |
| 3.85 | 12.09 |
| 4.47 | 10.80 |
| 4.09 | 12.71 |
| 6.34 | 10.04 |
| 3.30 | 9.06 |
| 7.13 | 6.31 |

Ejemplo para dos grupos

Con R

Manualmente

```
prom1 = mean(freidora$0ster1)
prom2=mean(freidora$0ster2)

s=((var(freidora$0ster1)*9)+(var(freidora$0ster2)*9))
t=(prom1-prom2-0)/(sqrt(s/10+s/10))
t; pt(-6.8527, df=18, lower.tail=TRUE)+pt(6.8527, df=18, lower.tail=FALSE)
```

```
## [1] -6.852696
```

```
## [1] 2.062577e-06
```

```
t.test(freidora$0ster1, freidora$0ster2, mu=0, paired=FALSE)
```

```
##
```

```
##      Two Sample t-test
```

```
##
```

```
## data:  freidora$0ster1 and freidora$0ster2
```

```
## t = -6.8527, df = 18, p-value = 2.063e-06
```

```
## alternative hypothesis: true difference in means is not equal to 0
```

```
## 95 percent confidence interval:
```

```
##  -6.654429 -3.531571
```

```
## sample estimates:
```

```
## mean of x mean of y
```

```
##      4.979      10.072
```


Prueba de Welch

Para los casos donde no se cumplen que las varianzas sean iguales u homogéneas.

$$t_{obs} = \frac{\bar{x}_1 - \bar{x}_2 - \Delta}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$$

Los grados de libertad se calculan:

$$df = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{s_1^4/(n_1 - 1) + s_2^4/(n_2 - 1)}$$

```
t.test(freidora$ster1, freidora$ster2, mu=0,
```

```
##  
##      Welch Two Sample t-test  
##  
## data:  freidora$ster1 and freidora$ster2  
## t = -6.8527, df = 17.268, p-value = 2.58e-06  
## alternative hypothesis: true difference in means  
## 95 percent confidence interval:  
##  -6.659188 -3.526812  
## sample estimates:  
## mean of x mean of y  
##      4.979      10.072
```

Pruebas no paramétricas y supuestos estadísticos

- Las pruebas no paramétricas no requieren que conozcamos la distribución de nuestros datos ni parámetros de la población de estudio como su media o varianza poblacional.
- Las pruebas no paramétricas nos permiten analizar datos en escala nominal u ordinal y en su mayoría los resultados se derivan a partir de procedimientos de ordenación y recuento, por lo que su base lógica es de fácil comprensión.
- Cuando trabajamos con muestras pequeñas ($n < 10$) en las que se desconoce si es válido suponer la normalidad de los datos, conviene utilizar pruebas no paramétricas
- Pocos supuestos: aleatoriedad e independencia.

PRUEBAS PARAMÉTRICAS

| VENTAJAS | DESVENTAJAS |
|--------------------------------------|--------------------------------------|
| - Mayor aplicabilidad | - Eficiencia estadística menor |
| - Se pueden usar variables ordinales | - Poder estadístico menor |
| - Son más fáciles de calcular | - No se pueden evaluar interacciones |

PRUEBAS PARAMÉTRICAS

| TIPO | PRUEBA |
|------------------------------|-------------------------|
| Comparación de 2 grupos | Wilcoxon/U Mann Withney |
| Comparación de >2 grupos | Kruskal-wallis/Friedman |
| Correlación de dos variables | Coeficiente de Spearman |
| Variables cualitativas | Chi-cuadrada/Fisher |

U Mann Whitney

- Este estadístico, introducido simultáneamente por Mann y Whitney en 1947, se utiliza para contrastar si dos muestras, extraídas independientemente, proceden de la misma población, no admite medidas repetidas.
- Compara las diferencias entre dos medianas, por lo que se basa en rangos en lugar de en los parámetros de la muestra (media, varianza).
- Requisitos: variable cuantitativa, independencia y aleatoriedad.
- Establece que existen diferencias entre las medianas de dos grupos sin suponer cuál de las dos es mayor o menor que la otra.

U Mann Whitney

- Se ordenan los datos (juntando los grupos) en orden creciente. El rango de cada dato será el número de orden que le corresponde.
- Se suman los rangos de cada uno de los grupos y se calcula la suma de los rangos de cada grupo (W_1, W_2).
- Se escoge el menor y considerando los df se puede usar una tablar de valores críticos para evaluar la significancia.

U Mann Whitney - Ejemplo

```
data_u<- data.frame(Control=c(12, 22, 38, 42,  
                             Tratamiento=c(14, 25, 40,
```

```
knitr::kable(data_u, align = "c")
```

| Control | Tratamiento |
|---------|-------------|
| 12 | 14 |
| 22 | 25 |
| 38 | 40 |
| 42 | 44 |
| 50 | 48 |

| Tipo | Valor |
|------|-------|
| C | 12 |
| E | 14 |
| C | 22 |
| E | 25 |
| C | 38 |
| E | 40 |
| C | 42 |
| E | 44 |
| E | 48 |

| | | | | | | | | | |
|----|----|----|----|----|----|----|----|----|----|
| C | E | C | E | C | E | C | E | E | C |
| 12 | 14 | 22 | 25 | 38 | 40 | 42 | 44 | 48 | 50 |

Para la condición de control, tomamos el número total de valores de la condición experimental que son más pequeños que cada valor de la condición de control. Hay 0 valores (12), 1 que es más pequeño que el segundo valor más pequeño (22), 2 que son más pequeños que el tercer valor más pequeño (38), 3 que son menores de 42, y 5 que son menores de 50 para un total de 11.

Para la condición del tratamiento, tomamos el número de valores que son más pequeños que cada valor de la condición experimental. Hay 1 valor de condición de control más pequeño que el valor de condición experimental más pequeño (14), 2 que son más pequeños que 25, 3 que son más pequeños que 40, 4 que son más pequeños que 44 y también 4 que son más pequeños que 48 para un Total de 14.

Tomamos el valor de U como el menor que sería 11.

U Mann Whitney - Ejemplo

En R

```
wilcox.test(data_u$Control, data_u$Tratamiento, paired = FALSE)
```

```
##  
##      Wilcoxon rank sum exact test  
##  
## data:  data_u$Control and data_u$Tratamiento  
## W = 11, p-value = 0.8413  
## alternative hypothesis: true location shift is not equal to 0
```

Prueba de Wilcoxon para dos muestras dependientes

- Esta es la opción no paramétrica para la prueba t de Student cuando los errores no se distribuyen de manera normal y se tiene una muestra o dos muestras apareadas.
- El estadístico W se calcula obteniendo las diferencias de los valores, ordenando de menor a mayor el valor absoluto de las diferencias las diferencias y asignando un rango, en caso de empates se obtiene la media de los rangos y se asigna el mismo rango a las dos observaciones.
- Funciona de la misma manera que la anterior pero es más potente porque tiene en cuenta la magnitud de los rangos de las diferencias. Es decir, una pequeña diferencia negativa o positiva significa menos o más en el cálculo de la estadística de prueba que una gran diferencia negativa o positiva.

Prueba de Wilcoxon

1. Arreglar las observaciones pareadas y obtener las diferencias de cada pareja.
2. Arreglar las diferencias en función de rangos como valores absolutos, sin importar el signo, pero de manera que los rangos conserven el signo correspondiente a la diferencia.
3. Obtener la sumatoria de los rangos en valores absolutos por signos.
4. Si se trata de muestras pequeñas, comparar el valor obtenido con los valores críticos de la tabla de Wilcoxon.
5. Distribuir las muestras mayores que 25 bajo la curva normal y, por tanto, calcular el valor Z.
6. Decidir si se acepta o rechaza la hipótesis

Prueba Wilcoxon - ejemplo

```
data_w<- data.frame(Observ=c(1:11),  
                    Antes=c(1,8,0,0,1,3,3,3,5,  
                             3,5,-2,-1,1,-5,-  
data_w$Diferencia<- data_w$Antes-data_w$Despué
```

| Observ | Antes | Después | Diferencia |
|--------|-------|---------|------------|
| 1 | 1 | 3 | -2 |
| 2 | 8 | 5 | 3 |
| 3 | 0 | -2 | 2 |
| 4 | 0 | -1 | 1 |
| 5 | 1 | 1 | 0 |
| 6 | 3 | -5 | 8 |
| 7 | 3 | -7 | 10 |
| 8 | 3 | -362 | 365 |
| 9 | 5 | -1 | 6 |

| Observ | Antes | Después | Diferencia |
|--------|-------|---------|------------|
| 8 | 3 | -362 | 365 |
| 11 | 5 | -10 | 15 |
| 7 | 3 | -7 | 10 |
| 6 | 3 | -5 | 8 |
| 9 | 5 | -1 | 6 |
| 10 | 6 | 0 | 6 |
| 2 | 8 | 5 | 3 |
| 3 | 0 | -2 | 2 |
| 4 | 0 | -1 | 1 |
| 5 | 1 | 1 | 0 |
| 1 | 1 | 3 | -2 |

| Observ | Antes | Después | Diferencia | Rango |
|--------|-------|---------|------------|-------|
| 8 | 3 | -362 | 365 | 10.0 |
| 11 | 5 | -10 | 15 | 9.0 |
| 7 | 3 | -7 | 10 | 8.0 |
| 6 | 3 | -5 | 8 | 7.0 |
| 9 | 5 | -1 | 6 | 5.5 |
| 10 | 6 | 0 | 6 | 5.5 |
| 2 | 8 | 5 | 3 | 4.0 |
| 3 | 0 | -2 | 2 | 2.5 |
| 4 | 0 | -1 | 1 | 1.0 |
| 5 | 1 | 1 | 0 | 0.0 |
| 1 | 1 | 3 | -2 | 2.5 |

| Observ | Antes | Después | Diferencia | Rango | Signo | Rango_Signo |
|--------|-------|---------|------------|-------|-------|-------------|
| 8 | 3 | -362 | 365 | 10.0 | + | +10 |
| 11 | 5 | -10 | 15 | 9.0 | + | +9 |
| 7 | 3 | -7 | 10 | 8.0 | + | +8 |
| 6 | 3 | -5 | 8 | 7.0 | + | +7 |
| 9 | 5 | -1 | 6 | 5.5 | + | +5.5 |
| 10 | 6 | 0 | 6 | 5.5 | + | +5.5 |
| 2 | 8 | 5 | 3 | 4.0 | + | +4 |
| 3 | 0 | -2 | 2 | 2.5 | + | +2.5 |
| 4 | 0 | -1 | 1 | 1.0 | + | +1 |
| 5 | 1 | 1 | 0 | 0.0 | = | =0 |
| 1 | 1 | 3 | -2 | 2.5 | - | -2.5 |

Wilcoxon -ejemplo

```
T_pos = 10+9+8+7+5.5+5.5+4+2.5+1  
T_neg = 2.5
```

```
T_pos; T_neg
```

```
## [1] 52.5
```

```
## [1] 2.5
```

```
Two_tailed_W = max(T_pos,T_neg)  
Two_tailed_W
```

```
## [1] 52.5
```

En R

```
wilcox.test(data_w$Antes, data_w$Después, pair
```

```
##
```

```
##      Wilcoxon signed rank test with continuity cor
```

```
##
```

```
## data:  data_w$Antes and data_w$Después
```

```
## V = 52.5, p-value = 0.0124
```

```
## alternative hypothesis: true location shift is no
```

RESUMEN

| ESTADÍSTICOS PARAMÉTRICOS | ESTADÍSTICOS NO PARAMÉTRICOS |
|---|--|
| Se conoce la población, parámetros y distribución | No se conoce la población, parámetros y distribución |
| Supuestos basados en población | Sin supuestos sobre la población |
| Aplicable para variables continuas | Aplicable para continuas y muchos tipos de discretas |
| Mayor poder estadístico | Menor poder estadístico |

Referencias y material suplementario

- Advanced Statistics I 2021 Edition
- Pruebas paramétricas y no paramétricas
- Estadística paramétrica y no paramétrica
- Prácticos de bioestadística 2

Tú: Los supuestos de esta prueba estadística no se cumplieron, por lo que vamos a usar una prueba no paramétrica.

El cliente:



¡MI TUTOR DICE QUE QUIERE ANOVA!

MUCHAS GRACIAS