

Prueba t para muestras relacionadas e independientes usando Rstudio, para que sirva y cómo aplicarlo.

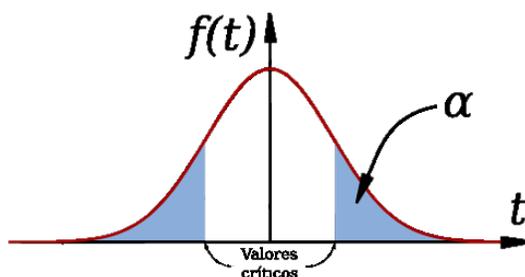
Mao Queen Garzón Quiroz¹

Wellington Remigio Villota Oyarvide²

Introducción

El procedimiento prueba T sirve para comparar de forma estadística a través de una distribución t de Student a una hipótesis que sea para variables cuantitativas. Para establecer una comparación entre variables continuas se puede utilizar el promedio de ambas, sin embargo, eso no quiere decir que estadísticamente exista una diferencia significativa (Amat, 2016). En palabras más sencillas, la prueba t indica cuán significativas son las diferencias entre dos grupos, es decir, si esas diferencias podrían haber acontecido por casualidad.

Figura 1. Distribución t-Student. Adaptado de “Valores críticos de una distribución t en R Commander”, por Modesto-Mata, 2016.



Cuando se extraen muestras de una población N con un tamaño de muestra de n , la distribución de la media de la muestra \bar{x} debe constituirse como una distribución normal. Bajo la hipótesis nula $\mu = \mu_0$, la distribución de estadísticas z debe también constituirse como una distribución normal. En el caso de que la varianza de la población sea desconocida, se puede reemplazar con la varianza de la muestra s^2 . Si es así, se da lugar a una distribución t ($n-1$ grados de libertad).

¹Mao Queen Garzón Quiroz. Universidad Católica de Santiago de Guayaquil, Guayaquil. Ecuador. mao.garzon@cu.ucsg.edu.ec

²Wellington Remigio Villota Oyarvide. Universidad Católica de Santiago de Guayaquil, Guayaquil. Ecuador. wellington.villota@cu.ucsg.edu.ec

La prueba t al ser paramétrica, las muestras deben cumplir con condiciones previas que se estudiarán en este ensayo (Seltman, 2018). Por consiguiente, se concluye que un requisito indispensable para la aplicación de esta prueba, es que los datos deben contar con una distribución normal.

A medida que el tamaño de la muestra aumenta, se aproxima a una distribución normal estándar. Por otro, se utiliza cuando la desviación estándar de la población es desconocida y suele emplearse en análisis discriminante. Generalmente, las pruebas T son utilizadas cuando los sujetos a quienes se les realizará la investigación experimental se dividen en dos grupos, a los que pueden denominarse grupo A y B para una mejor comprensión.

Como resumen se observa que la prueba T puede tener varios escenarios:

- Cuando se tiene una muestra relacionada o apareada.
- Cuando las muestras son independientes o desapareadas.

Bajo el escenario que posee dos muestras independientes existen dos subdivisiones que difieren entre varianzas homogéneas, es decir que son idénticas; o varianzas heterogéneas. Si los datos de las dos muestras poseen homocedasticidad, se debe usar el marco de muestras con varianzas homogéneas, por el contrario si no poseen esta característica de las varianzas, se trata de las heterogéneas (Siegle, 2019). El concepto de homocedasticidad la varianza de los errores permanece constante (Poo, 2015). La prueba estadística T depende de los grados de libertad, que es el número que nos permite saber a variabilidad de una muestra.

El investigador puede obtener dos diferentes resultado por grupo. Por ejemplo, en el caso de un tratamiento médico, un resultado antes del procedimiento y después del mismo, se obtendría como consecuencia: “preA” y “postA”, “preB” y “postB”.

En el escenario de prueba T independiente se utiliza una comparación entre los dos grupos A y B, después del tratamiento “postA” y “postB” o los cambios de cada grupo individual después del tratamiento. Por el contrario, las pruebas t relacionadas con usadas en diversos entornos experimentales. Por ejemplo, lamuestra no se divide en dos grupos, sólo se hace medición de los cambios en los mismos sujetos experimentales (Kim, 2015).

R es un poderoso programa de análisis de datos, posee herramientas y funciones que permite evaluar datos estadísticos de carácter numérico y alfabético, es utilizado por científicos, estadísticos y analistas de datos (Figuroa Preciado & Montoya Laos, 2015). Además, R puede examinar el código de cualquier función, ejecutar gráficas, vectores, matrices, etc (Collatón, 2014). R puede extenderse con los paquetes que le dan una potencia más allá de lo que permite la estructura básica. Mediante este programa es posible extraer información y descubrir patrones que permitan tomar decisiones basadas en los datos. Los usuarios interactúan con r por medio de comandos en vez de interfaz gráfico. Al usar este programa, es esencial mantener el orden y la estructura para recoger los distintos tipos de archivos como datos, scripts, imágenes, historias del conjunto de comandos y anotaciones.

Para pruebas estadísticas, como en el caso de conocer si las medias se relacionan, la prueba más común es T. Una prueba t es un tipo de estadística inferencial, en otras palabras, se refiere a un análisis que va más allá de simplemente una descripción de los números provistos por los datos de una muestra, pero que tiene la intención de delinear conclusiones sobre estos números entre poblaciones. Para ello, la prueba t analiza la diferencia entre las dos promedios originarias de las diferentes puntuaciones del grupo. Esta indica al investigador si es casualidad o no que la diferencia entre dos promedios sea mayor de lo esperado.

Las hipótesis pueden ser:

- $H_0 : \mu = 0$ $H_A : \mu \neq 0$

Quiere decir que en la hipótesis nula, la diferencia es igual a 0, versus la alternativa es que indica que la diferencia no es igual a 0. Esta sería la opción de dos colas.

- $H_0 : \mu \leq 0$ $H_A : \mu > 0$
- $H_0 : \mu \geq 0$ $H_A : \mu < 0$

Las dos últimas opciones representan a prueba unilateral o de una cola. Para una prueba de hipótesis de este tipo, se debe hallar si el valor de la muestra se localiza en extremo superior o inferior de la distribución (Fallas, 2012).

La hipótesis nula es que ambas mediciones son iguales, y la hipótesis alternativa es lo contrario. En el marco de la hipótesis nula, se calcula un estadístico t que con “distribución t con $n_1 + n_2 - 2$ grados de libertad” (Spector, 2014).

Cuando se tiene una muestra relacionada o apareada.

Las pruebas T de muestra relacionada o apareada consisten típicamente en una muestra de pares de valores con similares unidades estadísticas. Se aplica para una sola muestra que contiene un conjunto de datos o un grupo de unidades que han sido evaluadas en dos ocasiones diferentes, pero son dependientes (una prueba T de mediciones repetitivas). En términos generales, que dentro de un grupo a cada sujeto se le aplique una evaluación, se realizó algún procedimiento intermedio y luego se le vuelve a aplicar la evaluación. Un ejemplo sería analizar si ha habido cambios significativos a un grupo de personas que sean evaluados antes y después de un tratamiento médico, es decir una sola muestra se realizan dos mediciones, entre otros ejemplos se resaltan la medición a gemelos; padre e hijo; calificaciones después de un refuerzo académico o prueba de diagnóstico; compara fuerza de flexión de una articulación de una extremidad dominante y una no dominante; hacer un paralelo entre la capacidad de mantener el equilibrio ante condiciones distintas como tapados los oídos y sin taparlos, cerrando los ojos o con ellos abiertos, entre muchos otros (Amat, 2016).

La fórmula original para realizar la prueba T para muestras relacionadas es:

$$t = \frac{X - \mu}{\frac{s_x}{\sqrt{n}}} \sim t_{p-1}$$

En el numerador se incluye el promedio de las diferencias menos μ que será igual a 0, y en el denominador se utiliza la desviación estándar de las diferencias sobre la raíz cuadrada de n (que no es de cada grupo sino el número de pares). Los grados de libertad corresponderán al número de pares, menos 1.

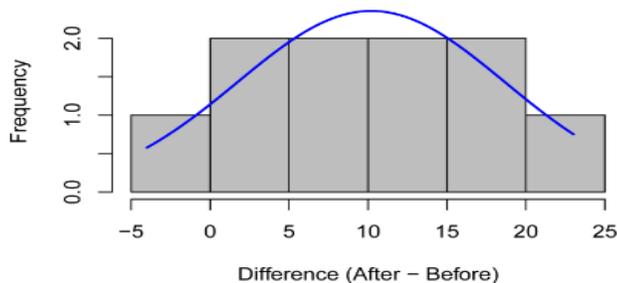
Es necesario que ambas muestras se distribuyan siguiendo una distribución normal y sus varianzas sean iguales. En R se puede cargar una base de datos para su correspondiente análisis. Para introducir los datos dentro del programa, una de las formas es indicarle que tome los datos a partir de una dirección específica dentro de un archivo. Lo primero que debe conocerse es si los datos las diferencias entre cada par de los primeros datos y los segundos siguen una

distribución normal. Para ello, se crea un histograma con la función “hist” o “plotNormalHistogram”, al reflejar el gráfico se debe mostrar que sea simétrico para que se cumpla el supuesto de que sea una distribución normal.

Figura 2. Histograma en R para comprobar distribución normal R. Tomado de “Paired t-test”, por R Handbook, 2016.

```
Before = Data$Score[Data$Time=="Before"]
After  = Data$Score[Data$Time=="After"]
Difference = After - Before

x = Difference
library(rcompanion)
plotNormalHistogram(x,lab="Difference (After - Before)")
```



Para contrastar la normalidad se usa el test de shapiro-wilk. Se debe de observar el valor de ambas variables y si se sitúan por encima de 0,05 significa que se acepta la hipótesis nula y se considera que ambas se distribuyen siguiendo una distribución normal. El siguiente paso es contrastar la homocedasticidad o igualdad de varianzas se usa el test f de Fisher y de la misma forma si el valor es mayor a 0,05 se acepta la hipótesis nula de igualdad de varianzas.

Para ejecutar una prueba para muestras relacionadas, se requiere dos vectores o los dos pares de datos, y_1 y y_2 . La prueba luego se ejecuta en R, una opción es utilizando la sintaxis “t.test (y_1 , y_2 , paired = True)”. Es importante especificar que los valores son pareados. El procedimiento interno es $y_1 - y_2$, es decir el promedio de las diferencias. Es así como se obtiene el estadístico t y además el p valor, que según la regla si es menor de 0,05 se rechaza la hipótesis nula y se acepta la alternativa, se aceptan los datos como válidos (STHDA, 2019a). Cada estadístico t tiene un p valor, este representa la probabilidad de que los resultados se hayan ocasionado por casualidad o no. Estos valores fluctúan entre 0 a 100% y normalmente se escriben en términos

decimales. Mientras más bajo es el p valor indican que los datos no fueron producidos como eventualidad (Manterola D & Pineda N, 2008).

La segunda alternativa es empleando una tercera columna en la base de datos que incluya la diferencia de y_1 y y_2 , este se convertiría en un tercer vector y_3 y es así como mediante el uso del comando “t.test (y3)” directamente el valor de t deber salir exactamente igual. Un ejemplo done se usa la prueba T con muestras relacionadas sería en un hospital donde se testea un nuevo medicamento que pretende reducir la presión arterial. A los pacientes se evalúa su presión antes y después del tratamiento y a través de la prueba T se evalúa si la mejoría es significativa o no (Peck & Devore, 2008).

Adicionalmente, se puede hacer la prueba con otra versión donde se coloca la alternativa menor “t.test (y2-y1, paired = True, alternative = ”less”)”. Esto se interpretaría que la diferencia es menor que 0. Finalmente se aplica la misma regla para aceptar o rechazar una hipótesis. R tiene una ventaja porque permite mayor rapidez cuando se calcula por una sola variable, es decir con el vector de las diferencias con el comando “t.test (y3, alternative=”less”)”

Ejemplo: Un catedrático de Finanzas realizó una serie de preguntas a sus estudiantes, luego de tal evaluación, llevó a cabo un taller sobre gestión financiera y volvió a tomar un cuestionario. Posteriormente, hizo una comparación de las calificaciones de cada alumno. Se asume que alumnos toman la prueba antes y después y que por lo tanto hay un primer y un segundo puntaje. Además en los datos se ordenará por letras a cada estudiante y en cuanto al momento si fue en la primera prueba o posterior (R Handbook, 2016).

```
Input = ("
Time Student Score
Before a 65
Before b 75
Before c 86
Before d 69
Before e 60
Before f 81
Before g 88
Before h 53
Before i 75
Before j 73
After a 77
After b 98
After c 92
After d 77
After e 65
After f 77
After g 100
After h 73
After i 93
After j 75
")
Data = read.table(textConnection(Input),header=TRUE)
```

Figura 3. Sintaxis para prueba t pareada en R. Tomado de “Paired t-test”, por R Handbook, 2016.

```
t.test(Score ~ Time,  
      data = Data,  
      paired = TRUE)  
  
Paired t-test  
t = 3.8084, df = 9, p-value = 0.004163  
alternative hypothesis: true difference in means is not equal to 0  
95 percent confidence interval:  
 4.141247 16.258753  
sample estimates:  
mean of the differences  
      10.2
```

Se concluye que se acepta la hipótesis alternativa porque el p valor es menor que 0,05 y se interpreta que sí existe diferencia significativa entre el primer cuestionario y la evaluación posterior al taller de finanzas.

Cuando las muestras son independientes o desapareadas

Las pruebas T desapareadas o de muestra independiente, se utiliza cuando se obtiene dos grupos de muestras aleatorias, independientes e idénticamente distribuidas a partir de las dos poblaciones a ser comparadas. Cuando se trata de muestras independientes, a diferencia de las pareadas, estas se emplean cuando se efectúa una comparación entre dos grupos que no se relacionan ni dependen de otros. En el momento que los participantes de ambos grupos no están relacionados entre sí, es decir son grupos totalmente diferentes y sin ningún vínculo, puede decirse que se trata de una prueba t para muestras independientes (REACH, 2019).

Otro ejemplo puede ser un grupo de 100 personas elegidos al azar, subdividiendo aleatoriamente a dos grupos de 50 personas cada uno donde el grupo A será evaluado con un tratamiento y el grupo B será determinado para un control esto sería una prueba de T independiente ya que se omite la edad y género, pero estos no difieren ya que los datos son observacionales (García Bellido, González Such, & Jornet Meliá, 2010)

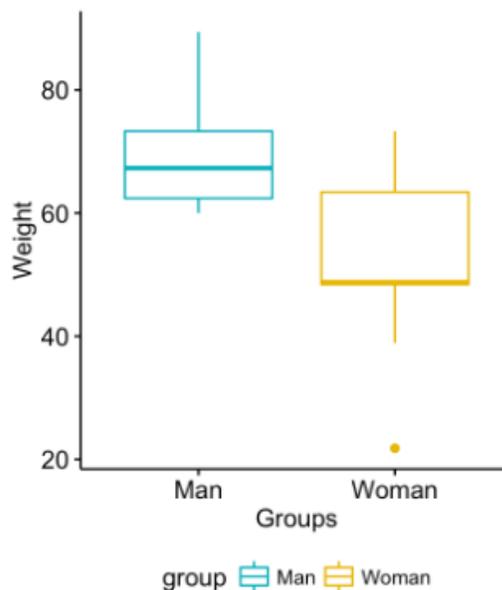
En el caso de tratarse de dos muestras independientes, es necesario valerse de un paquete denominado “tidyverse”. A partir de una base datos con múltiples campos e información y la creación de una variable que la almacena. Las características que deben cumplirse en los datos

para este tipo de prueba T son que los mismos sigan una distribución normal y que no existan valores atípicos en ellos. Se puede realizar esta verificación a través de expresiones gráficas mediante el paquete “ggplot” para que se revisen y relacionen los distintos campos para obtener un gráfico de densidad, si se observa que se crea una campana de Gauss, cumple la primera condición. También se puede usar otras herramientas como “ggnorm” que despliega todos los campos a elegir para su análisis posterior y “qqline” para descubrir que si están cercanas a la recta, cumple que se sigue una distribución normal (Wetherill, 2018). Adicionalmente, un estadístico que comprueba que las condiciones se satisfacen y que las variables sean iguales es el denominado “shapiro.test” o Shapiro-Wilk test de normalidad, que luego de seleccionar la variable se reflejan el estadístico y el p valor. Si el p valor es menor que 0,05 (es decir que el nivel de significancia) se rechaza la hipótesis nula y si se sitúan por encima de 0,05 se significa que se acepta la hipótesis nula y se considera que ambas tienen una distribución normal (MIT Mathematics, 2015).

Se examina la existencia de valores atípicos, también se puede ejecutar a través de “ggplot” con un gráfico de cajas que al ubicarse en la por encima del límite superior y bajo el límite inferior representa valores atípicos. El siguiente paso es contrastar la homosedasticidad o igualdad de varianzas se usa el test f de Fisher con “var.test” y si el p valor es mayor que 0,05 no se rechaza la hipótesis nula. Para darle fuerza a la interpretación gráfica, se realiza la prueba T student para pruebas independientes, R lo entiende como “t.test” para descubrir si hay diferencia significativa entre las variables y se asume que las varianzas son iguales. Uno de los datos obtenidos es el p valor que es la probabilidad asociada y se aplica la misma regla, si es mayor que el nivel de significancia no se rechaza la hipótesis nula. Se realizan varios métodos de sintaxis para la prueba t en R, sin embargo depende del modo en que se introducen los datos (Santana & Nieves, 2016).

```
# Plot weight by group and color by group
library("ggpubr")
ggboxplot(my_data, x = "group", y = "weight",
          color = "group", palette = c("#00AFBB", "#E7B800"),
          ylab = "Weight", xlab = "Groups")
```

**Figura 4. Uso de ggplot para realizar gráfico de cajas para prueba t independiente en R.
Tomado de “Unpaired Two-Samples T-test in R”, por STHDA, 2019b.**



Ejemplo: Un especialista en deporte desea saber si existe una diferencia significativa en el peso de las mujeres y varones de sus equipos, para ellos se toma el peso individual de cada grupo Hombres y Mujeres.

	group	weight
1	Woman	38.9
2	Woman	61.2
3	Woman	73.3
4	Woman	21.8
5	Woman	63.4
6	Woman	64.6
7	Woman	48.4
8	Woman	48.8
9	Woman	48.5
10	Man	67.8
11	Man	60.0
12	Man	63.4
13	Man	76.0
14	Man	89.4
15	Man	73.3
16	Man	67.3
17	Man	61.3
18	Man	62.4

Figura 5. Sintaxis para prueba t para muestras independientes en R. Tomado de “Unpaired Two-Samples T-test in R”, por STHDA, 2019b.

```
# Compute t-test
res <- t.test(women_weight, men_weight, var.equal = TRUE)
res

Two Sample t-test
data: women_weight and men_weight
t = -2.7842, df = 16, p-value = 0.01327
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -29.748019  -4.029759
sample estimates:
mean of x mean of y
 52.10000  68.98889
```

Se concluye que se acepta la hipótesis alternativa porque el p valor es menor que 0,05 y se interpreta que sí existe diferencia significativa del peso entre hombres y mujeres de los equipos.

Conclusiones

En la actualidad se generan grandes conjuntos de datos cada vez mayores, relacionados con ciencias experimentales, medicina, economía, ciencias sociales, comunicación. Saber explorar, describir y analizar estos conjuntos de datos es necesario para comprender no sólo el mundo que nos rodea sino para la toma de decisiones en el ámbito científico. En las pruebas estadísticas, la distribución de probabilidad es importante, para identificar la probabilidad de cierto evento, si es significativo o no se utiliza la prueba t.

Se sugiere que antes de hacer un análisis en R Studio con la prueba t.test se realice algunos supuestos y gráficos para tener una idea más clara de lo que se busca, además se obtiene la seguridad de que efectivamente se determina estadísticamente la relación entre dos variables, puesto que aunque matemáticamente parecieran diferentes, este valor puede que no sea significativo. A su vez, el trabajo de R Studio se potencia cuando el tamaño de la muestra es considerable. Lo interesante de este recurso de análisis estadístico es que ofrece múltiples

paquetes, que se han desarrollado exponencialmente para análisis sofisticados y además para que muestre tanto de manera gráfica lo que se señala a través de comandos.

Esta herramienta es importante saberla usar, puesto que puede servir para ramas como la medicina, en los negocios para saber si algún cambio empresarial realmente surtió efecto, en el ámbito educativo, etc. Para el caso de la prueba t relacionada es esencialmente útil cuando compara muestras relacionadas o dos mediciones en momentos temporales distintos para probar si la diferencia es estadísticamente diferente de 0, para realizar la prueba se asume que la distribución es normal y sin valores atípicos. La hipótesis nula indica que la media de las diferencias es igual a cero, mientras que la alternativa que no es igual a 0. En contraste, en el caso de las muestras independientes

En conclusión, tanto para muestras relacionadas como independientes, debe establecerse la hipótesis, estimar el estadístico con la diferencia de las medias, luego identificar la hipótesis alternativa, si la campana tiene una cola o dos. Aunque en su mayoría el nivel de significancia se recomienda que sea 0,05, se puede asumir otra probabilidad de error determinada por quien realiza el test. Sólo si se cumplen las condiciones de distribución normal, la inexistencia de valores atípicos y que la varianza sea igual en ambos casos, se puede proceder a calcular la prueba t obteniendo el p valor para la posterior interpretación. R Studio provee la herramienta de trabajo, sin embargo es el trabajo del científico de inferir correctamente para estimar valores correctos y llegar a las conclusiones que exija cada investigación de tal manera que se utilice este método efectivamente (Amat, 2016).

Bibliografía

- Collatón, R. (2014). Introducción al uso de R y R Commander para el análisis estadístico de datos en ciencias sociales. Recuperado de https://cran.r-project.org/doc/contrib/Chicana-Introduccion_al_uso_de_R.pdf
- Fallas, J. (2012). Prueba de Hipótesis. 54.
- Figuroa Preciado, G., & Montoya Laos, J. (2015). Introducción al Software Estadístico R. Recuperado de http://www.mat.uson.mx/sitio/lice/Intro_al_R.pdf
- García Bellido, R., González Such, J., & Jornet Meliá, J. M. (2010). Prueba T para muestras independientes. 6.

- Kim, T. K. (2015). T test as a parametric statistic. Korean journal of anesthesiology. <https://doi.org/10.4097/kjae.2015.68.6.540>
- Manterola D, C., & Pineda N, V. (2008). El valor de “p” y la “significación estadística”: Aspectos generales y su valor en la práctica clínica. Revista chilena de cirugía, 60(1). <https://doi.org/10.4067/S0718-40262008000100018>
- MIT Mathematics. (2015). The Shapiro-Wilk and related tests for normality. Recuperado de <https://math.mit.edu/~rmd/465/shapiro.pdf>
- Modesto-Mata, M. (2016, abril 4). Valores críticos de una distribución t en R Commander. Recuperado el 27 de junio de 2019, de Viva el Software Libre website: <https://vivaelssoftwarelibre.com/valores-criticos-de-una-distribucion-t/>
- Peck, R., & Devore, J. L. (2008). Statistics: The exploration and analysis of data (6th ed). Belmont, CA: Thomson Brooks/Cole.
- Poo, J. M. R. (2015). ECONOMETRÍA I - Tema 6: Heterocedasticidad. 23.
- R Handbook. (2016). Paired t-test. Recuperado el 28 de junio de 2019, de https://rcompanion.org/handbook/I_04.html
- REACH. (2019). Understanding t-Tests. Recuperado de http://gabrielschlomer.weebly.com/uploads/2/8/8/5/28853963/understanding_t_test_0.pdf
- Santana, A., & Nieves, C. (2016). Inferencia estadística con R. Recuperado el 21 de junio de 2019, de <http://www.dma.ulpgc.es/profesores/personal/stat/cursor4ULPGC/11-inferencia-MediaVar.html>
- Seltman, H. J. (2018). Experimental Design and Analysis. 428.
- Siegle, D. (2019). T Test. Recuperado el 21 de junio de 2019, de | Educational Research Basics website: <https://researchbasics.education.uconn.edu/t-test/#>
- Spector, P. (2014). Using t-tests in R. Recuperado el 21 de junio de 2019, de University of California, Berkeley - Department of Statistics website: <https://statistics.berkeley.edu/computing/r-t-tests>
- STHDA. (2019a). Paired Samples T-test in R. Recuperado el 27 de junio de 2019, de <http://www.sthda.com/english/wiki/paired-samples-t-test-in-r>
- STHDA. (2019b). Unpaired Two-Samples T-test in R - Easy Guides - Wiki - STHDA. Recuperado el 28 de junio de 2019, de <http://www.sthda.com/english/wiki/unpaired-two-samples-t-test-in-r>
- Wetherill, C. (2018). How to Perform T-tests in R. Recuperado el 21 de junio de 2019, de <https://datascienceplus.com/t-tests/>