

## INTERVALOS DE CONFIANZA Y TEST DE HIPOTESIS CON R

Estudiemos ahora la función en el lenguaje R, que nos ofrece tanto estimaciones puntuales y por intervalos de confianza como test de hipótesis, es:

```
> t.test (datos_x, datos_y =NULL, alternative = "two.sided", mu = 0,
paired =FALSE, var.equal = FALSE, conf.level= 0.95)
```

Las opciones indicadas son todas las ofrecidas por defecto. Podemos poner sólo un conjunto de datos para muestras unidimensionales (estimaciones puntuales) o dos conjuntos para comparación de muestras.

El argumento `alternative` indica el tipo de contraste, bilateral `two.sided`, si la hipótesis alternativa es mayor (Ho: menor o igual) se utiliza `greater`, si la hipótesis alternativa es menor (Ho: mayor o igual) entonces se usa `less`.

En `mu` indicamos el valor de la hipótesis nula.

En `paired=T` estamos ante una situación de datos no pareados, para indicar que estamos ante datos pareados poner `paired=F`.

Con `var.equal` estamos trabajando con los casos de igualdad o no de varianzas que sólo se emplean en comparación de dos poblaciones. Si `var.equal=T` las varianzas de las dos poblaciones son iguales si `var.equal=F` las varianzas de ambas poblaciones no se suponen iguales.

Por último tenemos el argumento `conf.level` en el que indicamos el nivel de confianza del test. Si deseáramos hacer el contraste para la igualdad de varianzas (cociente de varianzas=1) habríamos de emplear la función `var.test`:

```
var.test(x, y, ratio = 1, alternative = c("two.sided", "less",
"greater"), conf.level = 0.95, ...)
```

Los argumentos son los mismos que en la función `t.test`.

### Pruebas para una muestra

1. Los siguientes datos corresponden al rendimiento por hectárea de cierta nueva variedad de trigo, medido en 9 lotes experimentales: 3,35; 3,92; 4,26; 3,36; 3,72; 4,19; 3,42; 4,38; 4,5. Construya un intervalo de confianza al 90% de confianza para el rendimiento promedio de la nueva variedad de trigo si suponemos que el rendimiento por hectárea se distribuye aproximadamente normal.

En el lenguaje R procedemos como sigue

```
> T<-c(3.35, 3.92, 4.26, 3.36, 3.72, 4.19, 3.42, 4.38, 4.5)
> t.test(T, conf.level=0.90)
```

Construya:

El estadístico

Una prueba de hipótesis según la salida de R. Aceptela o recházela

¿Qué infiere? analice la salida de R

Verifique la hipótesis que el rendimiento medio es mayor que 4; aquí el test es ,

con el lenguaje R:

```
> t.test(T, conf.level=0.90, alternative="greater", mu=4)
```

2. Una operación de montaje en una fábrica manufacturera requiere aproximadamente un período de entrenamiento de un mes para que un nuevo operario alcance la máxima eficiencia. Se sugirió un nuevo método para el entrenamiento y se realizó una prueba para comparar el método nuevo con el método estándar. Se entrenaron dos grupos de 9 nuevos empleados durante un período de un mes; un grupo utilizó el método estándar y el otro grupo usó el método nuevo. Se midió el tiempo en minutos que necesito cada empleado en armar cierto dispositivo al final del período de entrenamiento; los resultados obtenidos fueron:

Método estándar	32	37	35	28	41	44	35	31	34
Método nuevo	35	31	29	25	34	40	27	32	31

Admitiendo que el tiempo de armado utilizado en ambos métodos son variables aleatorias independientes y distribuidas normalmente: ¿Tiene igual varianza? ¿Se puede aceptar la hipótesis de igualdad de tiempos de armado, en función de los datos y con un nivel de confianza del 95%?

Estime la diferencia real de las medias poblacionales con un nivel de confianza del 95%.

Con R procedemos como sigue:

```
E<-c(32,37,35,28,41,44,35,31,34)
N<-c(35,31,29,25,34,40,27,32,31)
```

```
###prueba de varianza
var.test(E,N)
```

```
## prueba de diferencias pareadas
t.test(E,N,paired=T,conf.level=0.95)
```

3. Se desea estudiar la proporción de individuos menores de 40 años en una determinada ciudad. Se tomó una muestra aleatoria de 800 individuos y se encontró que 600 de ellos tenían menos de 40 años. Determine la verdadera proporción de personas menores de 40 años en la mencionada ciudad mediante un intervalo de confianza al 99% de confianza.

Construya el intervalo de manera manual

Compare con la salida de R

Podemos determinar el intervalo de confianza con el lenguaje R usando la función `prop.test`.

```
prop.test(600, 800, conf.level = 0.99)
```

4. Dos grupos son considerados en un grupo sobre la efectividad de una nueva vacuna. El primer grupo, que recibe la vacuna contiene 200.745 individuos. El Segundo grupo, que recibe un placebo consiste de 201.229 individuos. Hubo 57 casos de enfermedad en el primer grupo y 142 casos en el segundo grupo

Asumamos un modelo binomial para cada grupo, con probabilidad de hacer la enfermedad en el primer grupo y con probabilidad de hacer la enfermedad en el segundo grupo; lo que deseamos saber es si la vacuna es o no eficiente, conocer las relaciones entre las proporciones  $p_1$  y  $p_2$ .

Podemos usar `prop.test` para obtener un intervalo de confianza para la diferencia de proporciones así como un test de hipótesis

Procedemos como sigue:

Los dos primeros argumentos de la función son vectores que contienen, respectivamente, el número de casos obtenidos y el segundo el tamaño de las muestras, así:

```
x <- c(57, 142)
n <- c(200745, 201229)
prop.test(x, n)
```

Realice una prueba unilateral que indique si la vacuna tiene efecto positivo, concluya

```
prop.test(x, n, alt = "l")
```

Halle un intervalo de confianza al 95% para la diferencia de proporciones. Que concluye

Halle un intervalo de confianza al 99% para la diferencia de proporciones. Que concluye

```
prop.test(x, n, conf.level = 0.99)
```