



Análisis estadístico sobre una base de datos de béisbol.

Luis Riera
Universidad Simón Bolívar
Caracas, Venezuela
16-10976@usb.ve

Eduardo Gavazut
Universidad Simón Bolívar
Caracas, Venezuela
13-10524@usb.ve

Miguel Cordero
Universidad Simón Bolívar
Caracas, Venezuela
15-10326@usb.ve

8 de abril de 2022

RESUMEN: En este documento se realiza una prueba de hipótesis sobre la variable X2 o tasa de bateo para determinar si el promedio es menor que 0.3 .

Palabras clave: Prueba de hipótesis, tasa de bateo, variable X2

1. PRUEBAS SOBRE LA TASA DE BATEO

Se desea probar con un nivel de significancia de $\alpha = 0,05$, que el promedio de bateo es inferior a 0,300.

Como hipótesis nula H_0 , supongamos que la media de bateo, \bar{X}_1 , es igual a 0,3. Y como hipótesis alternativa, H_a , que el promedio de bateo es superior a 0,3, $\bar{X}_1 > 0,3$.

Suponiendo que los datos presentan una distribución normal, podemos aplicar el comando `t.test`.

Con esta función, se obtuvo que el valor para el estadístico t es $-23,811$, con 44 grados de libertad. Como el p -valor es bastante alto, de hecho es igual 0,9976 (que representa un 99,76 %), se cumple que $\alpha = 0,05 < 99,76$ y por lo tanto la hipótesis alternativa se rechaza, mas aún, se rechaza para todo nivel de significancia porque se necesita un valor para α más alto que el p -valor para rechazar la hipótesis nula.

Se afirma entonces, con total seguridad, que la tasa de bateo es inferior a 0,300.

```
# Inicializamos la librería que permite leer archivos xlsx
library(readxl)
# Asignamos a una variable la información almacenada en el archivo
Baseball <- read_excel("~/GitHub/data/Baseball.xlsx")
```



```
# Mostramos las primeras 5 entradas
```

```
head(Baseball, n=5)
```

```
## # A tibble: 5 x 6
##       X1      X2      X3      X4      X5      X6
##   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 0.283 0.144 0.049 0.012 0.013 0.086
## 2 0.276 0.125 0.039 0.013 0.002 0.062
## 3 0.281 0.141 0.045 0.021 0.013 0.074
## 4 0.328 0.189 0.043 0.001 0.03  0.032
## 5 0.29  0.161 0.044 0.011 0.07  0.076
```

```
X1<- Baseball$X1
```

1.1. PRUEBAS SOBRE LA TASA DE BATEO

Se desea probar con un nivel de significancia de $\alpha = 0,05$, que el promedio de bateo es inferior a 0,300.

Como hipótesis nula H_0 , supongamos que la media de bateo, $\overline{X1}$, es igual a 0,3. Y como hipótesis alternativa, H_a , que el promedio de bateo es superior a 0,3, $\overline{X1} > 0,3$.

Suponiendo que los datos presentan una distribución normal, podemos aplicar el comando `t.test`.

1.2.

```
t.test(X1, alternative = "greater", mu=0.3, conf.level = 0.95)
```

```
##
## One Sample t-test
##
## data:  X1
## t = -2.9779, df = 44, p-value = 0.9976
## alternative hypothesis: true mean is greater than 0.3
## 95 percent confidence interval:
##  0.2694453      Inf
## sample estimates:
## mean of x
## 0.2804667
```

Con este función, se obtuvo que el valor para el estadístico t es $-23,811$, con 44 grados libertad. Como el p - valor es bastante alto, de hecho es igual 0,9976 (que representa un 99,76 %), se cumple que $\alpha = 0,05 < 99,76$ y por lo tanto la hipótesis alternativa se rechaza, mas aún, se rechaza para todo nivel de significancia porque se necesita un valor para α más alto que el p - valor para rechazar la hipótesis nula.



Se afirma entonces, con total seguridad, que la tasa de bateo es inferior a 0,300.