E6_JaimeRiquelme_209647087

Jaime Riquelme

2023-12-10

```r
if(!requireNamespace('BSDA', quietly = TRUE)){
  install.package('BSDA')
}
if(!requireNamespace('EnvStats', quietly = TRUE)){
  install.package('EnvStats')
}
library(EnvStats)
```

```
## Warning: package 'EnvStats' was built under R version 4.3.2

##
## Attaching package: 'EnvStats'

## The following objects are masked from 'package:stats':
##
##     predict, predict.lm
```

```r
library(BSDA)
```

```
## Warning: package 'BSDA' was built under R version 4.3.2

## Loading required package: lattice

##
## Attaching package: 'BSDA'

## The following object is masked from 'package:datasets':
##
##     Orange
```

Enunciado 1:

Una empresa de tecnología quiere saber si la duración promedio de la batería de su nuevo modelo de teléfono móvil supera las 24 horas. Se sabe que la duración promedio de la batería de modelos anteriores es de 24 horas, y la desviación estándar poblacional es de 2 horas. Se selecciona una muestra aleatoria de 40 teléfonos del nuevo modelo y se registra la duración de la batería. La muestra arroja una duración media de 25 horas. Utilizando un nivel de significancia del 5 %, ¿existe suficiente evidencia para afirmar que la duración media de la batería del nuevo modelo es superior a las 24 horas? Realiza un z-test para responder esta pregunta. Para trabajar con una muestra, puede generarla con rnorm utilizando los parámetros indicados en el enunciado.

Definición de Hipótesis

Hipótesis Nula (H_0): La duración media de la batería del nuevo modelo es menor o igual a 24 horas. Hipótesis Alternativa (H_1): La duración media de la batería del nuevo modelo es mayor a 24 horas.

Nivel de significancia

Se utilizará un nivel de significancia del 5% ($\alpha = 0.05$).

Datos de la Muestra

```r
set.seed(123)  # Para reproducibilidad
datos_muestra = rnorm(n = 40, mean = 25, sd = 2)  # Muestra simulada: media 25, desviación estándar 2
```

Realizamos la prueba:

```r
resultado_test = z.test(datos_muestra, mu = 24, sigma.x = 2, alternative = "greater")
print(resultado_test)
```

```
##
##  One-sample z-Test
##
## data:  datos_muestra
## z = 3.448, p-value = 0.0002823
## alternative hypothesis: true mean is greater than 24
## 95 percent confidence interval:
##  24.57022       NA
## sample estimates:
## mean of x
##  25.09037
```

Conclusión

Los resultados del análisis estadístico proporcionan evidencia convincente para rechazar la hipótesis nula. Con un valor p de 0.0002823, que es significativamente menor que el umbral de significancia de 0.05, se concluye con un 95% de confianza que la duración media de la batería del nuevo modelo de teléfono móvil es mayor a 24 horas. Este bajo valor p indica que la probabilidad de observar un resultado tan extremo como el obtenido, bajo la suposición de que la hipótesis nula es cierta, es mínima. Por lo tanto, la duración promedio de la batería que supera las 24 horas no es una ocurrencia aleatoria, sino más bien una indicación de que el nuevo modelo tiene efectivamente una duración de batería mejorada en comparación con los modelos anteriores.

Además, el valor Z obtenido en el test refuerza esta conclusión, indicando que la media de la muestra está significativamente alejada de la media poblacional hipotética de 24 horas. Esto refuerza la afirmación de que el nuevo modelo muestra una mejora en la duración de la batería comparado con los estándares anteriores.

Enunciado 2:

Un restaurante introduce un nuevo plato y desea saber si el tiempo de preparación promedio de este plato es diferente del tiempo estándar de 30 minutos. El restaurante no tiene datos previos sobre la variabilidad del tiempo de preparación de este plato.Se toma una muestra de 25 preparaciones del nuevo plato, obteniendo un tiempo promedio de preparación de 32 minutos. Utilizando un nivel de significancia del 5 %, ¿existe suficiente evidencia para afirmar que el tiempo de preparación promedio del nuevo plato es diferente de 30 minutos?.

Definición de Hipótesis

Hipótesis Nula (H_0): El tiempo promedio de preparación del plato es igual a 30 minutos. Hipótesis Alternativa (H_1): El tiempo promedio de preparación del plato no es igual a 30 minutos.

Nivel de significancia

Se utilizará un nivel de significancia del 5% ($\alpha = 0.05$).

Datos de la Muestra

```
set.seed(123)
tiempo_preparacion = rnorm(n=25, mean=32, sd=5)  # Generación de una muestra simulada de 25 preparacion
```

Realizamos la prueba:

```
resultado = t.test(tiempo_preparacion, mu = 30, alternative="two.sided")
print(resultado)
```

```
##
##  One Sample t-test
##
## data:  tiempo_preparacion
## t = 1.9365, df = 24, p-value = 0.06467
## alternative hypothesis: true mean is not equal to 30
## 95 percent confidence interval:
##  29.87939 33.78731
## sample estimates:
## mean of x
##  31.83335
```

Conclusión

El valor p obtenido es 0.06467, lo cual es mayor que el nivel de significancia de 0.05. Esto indica que no hay suficiente evidencia estadística para rechazar la hipótesis nula. Por lo tanto, no podemos afirmar con un 95% de confianza que el tiempo promedio de preparación del nuevo plato sea diferente de 30 minutos. El resultado sugiere que, aunque el tiempo promedio de la muestra es de 32 minutos, esta diferencia no es estadísticamente significativa para concluir que el tiempo promedio de preparación ha cambiado respecto al estándar de 30 minutos.

Enunciado 3:

Una empresa de fabricación de neumáticos desea evaluar si el proceso de producción de neumáticos nuevos está bajo control en términos de variabilidad. La empresa tiene un estándar de calidad que establece que la variabilidad en el espesor de los neumáticos no debe exceder una varianza de 4 mm². Se toma una muestra aleatoria de 30 neumáticos y se mide el espesor de cada uno, resultando en una varianza muestral de 5 mm². Utilizando un nivel de significancia del 5 %, ¿existe suficiente evidencia para afirmar que la variabilidad en el proceso de producción es mayor que el estándar establecido?.

Definición de Hipótesis

Hipótesis Nula (H_0): La varianza en el espesor de los neumáticos es igual a 4mm^2. Hipótesis Alternativa (H_1): La varianza en el espesor de los neumáticos es mayor que 4mm^2.

Nivel de significancia

Se utilizará un nivel de significancia del 5% ($\alpha = 0.05$).

Datos de la Muestra

```r
n = 30  # Tamaño de la muestra
varianza_muestral = 5  # Varianza muestral en mm^2
varianza_estandar = 4  # Varianza estándar en mm^2
```

Conclusión

El valor chi-cuadrado obtenido es 36.25 y el p-value es aproximadamente 0.1664. Estos resultados indican que no hay suficiente evidencia estadística para rechazar la hipótesis nula, ya que el p-value es mayor que el nivel de significancia de 0.05. Esto implica que con un 95% de confianza, no se puede concluir que la variabilidad en el proceso de producción de neumáticos sea significativamente mayor que el estándar de 4mm^2. La variabilidad observada en la muestra no muestra una diferencia significativa en comparación con el valor estándar especificado.

Realizamos la prueba:

```r
# Cálculo del estadístico chi-cuadrado y p-value
chi_cuadrado = (n - 1) * varianza_muestral / varianza_estandar
grados_libertad = n - 1
p_valor = pchisq(chi_cuadrado, grados_libertad, lower.tail = FALSE)
resultado_test = list(estadistico = chi_cuadrado, p_valor = p_valor)
print(resultado_test)
```

```
## $estadistico
## [1] 36.25
##
## $p_valor
## [1] 0.1663945
```

Enunciado 4:

Una cadena de cines está interesada en aumentar la venta de palomitas de maíz. Un estudio previo mostró que aproximadamente el 40 % de los clientes compran palomitas. Tras implementar una nueva estrategia de marketing, la cadena desea saber si el porcentaje ha aumentado. Después de la implementación de la nueva estrategia, se selecciona una muestra aleatoria de 200 clientes, y se encuentra que 96 de ellos compraron palomitas de maíz. Utilizando un nivel de significancia del 5 %, ¿existe suficiente evidencia para afirmar que el porcentaje de clientes que compran palomitas ha aumentado?.

Sabemos que una cadena de cines ha implementado una nueva estrategia de marketing para aumentar la venta de palomitas de maíz. Un estudio previo mostró que aproximadamente el 40% de los clientes compraban palomitas. Tras la implementación de la nueva estrategia, se desea saber si este porcentaje ha aumentado.

Se seleccionó una muestra aleatoria de 200 clientes después de implementar la estrategia, y se registró cuántos de ellos compraron palomitas de maíz. Para analizar si hay un aumento significativo en la proporción de ventas, se realizará una prueba de hipótesis para una proporción.

Definición de Hipótesis

- Hipótesis Nula (H_0): La proporción de clientes que compran palomitas es igual a 0.4 (40%).
- Hipótesis Alternativa (H_1): La proporción de clientes que compran palomitas es mayor que 0.4 (40%).

Nivel de Significancia

Se utilizará un nivel de significancia del 5% ($\alpha = 0.05$).

Datos de la Muestra

```r
# Definición de variables
probabilidad_anterior = 0.4  # Proporción anterior
tamano_muestra = 200         # Tamaño de la muestra
compras_palomitas = 96       # Número de clientes que compraron palomitas
```

Realizamos la prueba:

```r
# Realización del test de proporciones
probabilidad_resultado = prop.test(compras_palomitas, tamano_muestra, p = probabilidad_anterior, alterna
probabilidad_resultado
```

```
##
##  1-sample proportions test with continuity correction
##
## data:  compras_palomitas out of tamano_muestra, null probability probabilidad_anterior
## X-squared = 5.0052, df = 1, p-value = 0.01264
## alternative hypothesis: true p is greater than 0.4
## 95 percent confidence interval:
##  0.4200931 1.0000000
## sample estimates:
##    p
## 0.48
```

Conclusión

Con base en los resultados obtenidos de la prueba de proporciones con un valor-p de 0.02527 y considerando un nivel de significancia de 0.05, rechazamos la hipótesis nula. Por lo tanto, existe evidencia estadística para afirmar que la proporción de clientes que compran palomitas de maíz es diferente de 0.40. Dado que estábamos investigando si el porcentaje ha aumentado y la proporción de la muestra es de 0.48, se concluye que la proporción de clientes que compran palomitas de maíz ha aumentado significativamente después de la implementación de la nueva estrategia de marketing. El intervalo de confianza del 95% para la proporción verdadera va de aproximadamente 40.94% a 55.14%, lo que refuerza esta conclusión.

Enunciado 5

Una compañía de suplementos alimenticios está investigando el efecto de un nuevo producto en la mejora del rendimiento atlético. Se lleva a cabo un estudio con un grupo de deportistas para evaluar el impacto del suplemento. Un grupo de 15 deportistas participa en un experimento donde se mide su rendimiento en una prueba de resistencia. Se toman mediciones antes de iniciar el consumo del suplemento y después de cuatro semanas de uso continuo. Los investigadores están interesados en determinar si el consumo del suplemento ha tenido un efecto significativo en el rendimiento de los deportistas.

Definición de Hipótesis

- Hipótesis Nula (H_0): No hay diferencia en el rendimiento antes y después de consumir el suplemento.
- Hipótesis Alternativa (H_1): Existe una diferencia en el rendimiento antes y después de consumir el suplemento.

Nivel de Significancia

Se utilizará un nivel de significancia del 5% ($\alpha = 0.05$).

Datos de la Muestra

```r
# Creación de la tabla de datos
rendimiento <- data.frame(
  Atleta = 1:15,
  Antes = c(20.5, 18.7, 21.3, 19.5, 22.1, 17.8, 20.0, 23.4, 21.5, 18.0, 19.2, 22.6, 17.9, 21.7, 20.3),
  Despues = c(22.0, 19.1, 21.8, 20.0, 23.5, 18.2, 20.7, 25.1, 22.3, 28.5, 19.7, 23.0, 18.3, 22.2, 21.5)
)

print(rendimiento)
```

```
##    Atleta Antes Despues
## 1       1  20.5    22.0
## 2       2  18.7    19.1
## 3       3  21.3    21.8
## 4       4  19.5    20.0
## 5       5  22.1    23.5
## 6       6  17.8    18.2
## 7       7  20.0    20.7
## 8       8  23.4    25.1
```

```
## 9        9  21.5    22.3
## 10      10  18.0    28.5
## 11      11  19.2    19.7
## 12      12  22.6    23.0
## 13      13  17.9    18.3
## 14      14  21.7    22.2
## 15      15  20.3    21.5
```

Realizamos la prueba:

```r
# Realización del test de Wilcoxon
resultado_test = wilcox.test(rendimiento$Antes, rendimiento$Despues, paired = TRUE)
```

```
## Warning in wilcox.test.default(rendimiento$Antes, rendimiento$Despues, paired =
## TRUE): cannot compute exact p-value with ties
```

```r
print(resultado_test)
```

```
##
##  Wilcoxon signed rank test with continuity correction
##
## data:  rendimiento$Antes and rendimiento$Despues
## V = 0, p-value = 0.0007051
## alternative hypothesis: true location shift is not equal to 0
```

Conslusión

El p-value obtenido del test de Wilcoxon es 0.0007051, lo cual es menor que el nivel de significancia de 0.05. Por lo tanto, con un 95% de confianza, rechazamos la hipótesis nula, indicando que hay una diferencia significativa en el rendimiento de los deportistas antes y después de tomar el suplemento. Esto sugiere que el consumo del suplemento ha tenido un efecto positivo en el rendimiento atlético de los participantes.

Enunciado 6

Un restaurante popular ha decidido modificar la receta de su plato estrella para hacerlo más saludable, cambiando algunos ingredientes para reducir el contenido de grasa. Para evaluar la aceptación de la nueva receta, el restaurante realiza una encuesta con 20 de sus clientes habituales que han probado tanto la versión original como la nueva del plato. Cada cliente indica si prefiere la versión original o la nueva. El restaurante necesita saber si la nueva receta es tan buena como la original según la percepción de los clientes. Examina los resultados de la encuesta para determinar si hay una preferencia clara por alguna de las dos versiones. Sugerencia : puede señalar en un arreglo la preferencia por un plato u otro, como un vector de valores binarios, es decir, c(1,0,1,1,0,...) y puede hacerlo aleatorio.

Definición de Hipótesis

- Hipótesis Nula (H_0): No hay una preferencia clara por ninguna de las versiones del plato.
- Hipótesis Alternativa (H_1): Hay una preferencia clara por el plato nuevo.

Nivel de Significancia

Se utilizará un nivel de significancia del 5% ($\alpha = 0.05$).

Datos de la Muestra

```r
set.seed(123)  # Para reproducibilidad
preferencias <- c(0, 1, 0, 1, 1, 0, 0, 1, 1, 1, 0, 0, 0, 1, 1, 0, 0, 0, 1, 0)
```

Realizamos la prueba:

```r
resultado_test = binom.test(sum(preferencias), length(preferencias), p = 0.5)
print(resultado_test)
```

```
##
##  Exact binomial test
##
## data:  sum(preferencias) and length(preferencias)
## number of successes = 9, number of trials = 20, p-value = 0.8238
## alternative hypothesis: true probability of success is not equal to 0.5
## 95 percent confidence interval:
##  0.2305779 0.6847219
## sample estimates:
## probability of success
##                   0.45
```

Conclusión

Dado que el p-value es mucho mayor que 0.05, aceptamos la hipótesis nula con un 95% de confianza. Esto sugiere que no existe evidencia estadística de que haya una preferencia clara por el plato nuevo sobre el plato antiguo. La proporción estimada de clientes que prefieren el plato nuevo es del 45%, lo cual está bastante equilibrado y no muestra una inclinación significativa hacia ninguna de las dos opciones.

Por lo tanto, a partir de esta muestra y análisis, se puede inferir que la introducción del plato nuevo en el menú no ha generado una preferencia notablemente diferente en comparación con el plato antiguo entre los clientes. El restaurante podría considerar que ambos platos tienen un nivel de aceptación similar entre su clientela.