

TEMA 8

Inferencia estadística

Métodos Numéricos y Estadísticos
Grado en Ingeniería Informática

Curso 2021-2022

Eva María Mazcuñán Navarro



Eva María Mazcuñán Navarro
Departamento de Matemáticas
Universidad de León
E-mail: emmazn@unileon.es



Esta obra está bajo una [licencia de Creative Commons Reconocimiento 4.0 Internacional](https://creativecommons.org/licenses/by/4.0/)

Contenidos

	Página
Introducción	1
1 Estadística descriptiva	3
1.1 Estimación puntual	3
1.2 Gráficos	4
2 Contrastes de hipótesis	6
2.1 Planteamiento de las hipótesis	6
2.2 Cálculo del p -valor	7
2.3 Conclusión	9
3 Intervalos de confianza	11
Problema <i>tipo examen</i>	13
Bibliografía	17

Introducción



Figura 1: Viñeta de la tira cómica *Dilbert*, de Scott Adams.

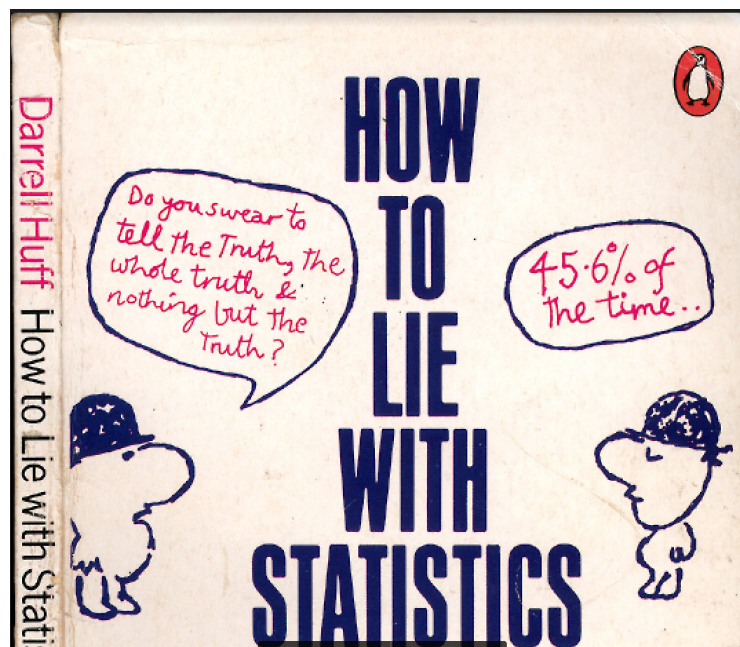


Figura 2: Portada del libro *How to Lie with Statistics*, de Darrell Huff.

Hasta ahora hemos realizado un estudio teórico de las variables aleatorias. Y hemos aprendido a resolver problemas como el siguiente:

Sabemos que, para un determinado jugador de baloncesto, la probabilidad con que acierta un lanzamiento de tiro libre es 0.75.

Calcula la probabilidad de que, en una tanda de 50 lanzamientos, acierte como máximo 30 de ellos.

Para resolver el problema anterior, consideramos la variable aleatoria

X = número de lanzamientos acertados.

Identificamos que X sigue una distribución binomial de parámetros $n = 50$ y $p = 0.75$, abreviadamente

$$X \sim \text{Binom}(50, 0.75),$$

y calculamos la probabilidad pedida como

$$P(X \leq 30) = F(30) = \text{pbinom}(30, 50, 0.75) = 0.0139.$$

Ahora, en una situación real, lo más frecuente es que no conozcamos el valor $p = 0.75$ para la probabilidad de acierto del jugador en el lanzamiento de tiros libres. Porque este valor p es un parámetro teórico, que interpretamos como la proporción de aciertos del jugador cuando el número de lanzamientos tiende a infinito. Para obtener el valor de p deberíamos observar los infinitos potenciales lanzamientos que podría efectuar el lanzador, situación imposible en la práctica. En consecuencia, el tipo de problema que se plantea en la práctica es más parecido al siguiente:

Problema 1

Nos planteamos investigar el valor del parámetro

p = probabilidad de que el jugador acierte un tiro libre.

Para ello, observamos el resultado de 50 lanzamientos efectuados por el jugador en cuestión, anotando en cada caso si acierta o falla. Se obtienen 28 aciertos y 22 fallos.

El objetivo es realizar inferencias sobre el parámetro p basándonos en los resultados de los lanzamientos observados. Queremos en concreto responder a las siguientes cuestiones:

- El jugador afirma que su probabilidad de acertar un tiro libre es 0.75 ¿resultan los datos compatibles con esta afirmación?
- A la vista de los datos ¿cuál sería un rango de valores plausibles para la probabilidad p ?

En este tema estudiaremos los procedimientos de inferencia estadística que permitirán responder a las preguntas anteriores. Responderemos a la primera pregunta utilizando un procedimiento de inferencia estadística denominado **contraste de hipótesis** (sección 2) y abordaremos la segunda cuestión construyendo un **intervalo de confianza** (sección 3).

En general, un procedimiento de inferencia estadística es un método o técnica que permite inferir, a partir de la información empírica proporcionada por una muestra (en nuestro ejemplo los aciertos observados en 50 lanzamientos), el comportamiento de una determinada población (en nuestro ejemplo la probabilidad de acierto p). Las conclusiones de un procedimiento de inferencia estadística nunca serán afirmaciones categóricas, sino que estarán sujetas a un riesgo de error, que se cuantifica en términos de probabilidades (ver figuras 1 y 2).

Presentaremos el comando de R `binom.test()` para realizar el contraste de hipótesis y obtener el intervalo de confianza para la proporción p que responderán a las preguntas planteadas antes. Basándonos en nuestro conocimiento de las variables binomiales estudiadas en el tema anterior, seremos capaces de reproducir los cálculos que realiza este comando. De esta forma, podremos entender los fundamentos generales de los contrastes de hipótesis y los intervalos de confianza.

En la última práctica de ordenador del curso, estudiaremos más comandos de R para realizar contrastes de hipótesis y calcular intervalos de confianza para otros parámetros, como la media de una variable continua. Gracias a lo aprendido en este tema, sabremos interpretar las salidas de dichos comandos, aún sin conocer el detalle de los cálculos realizados para obtenerlas.

1. Estadística descriptiva

Todo estudio estadístico comienza con una exploración inicial de los datos obtenidos, que habitualmente incluye el cálculo determinados valores de interés que resumen o describen los datos, y la visualización de los datos mediante los gráficos adecuados en cada caso.

Para nuestro ejemplo calcularemos la llamada proporción muestral y realizaremos un diagrama de barras.

Comenzamos cargando el paquete `tidyverse` para tener acceso a todas las funciones de R que usaremos a lo largo del tema.

```
library("tidyverse")
```

1.1. Estimación puntual

Supongamos, como se indicó en el problema planteado en la introducción, que en los 50 lanzamientos observados, el jugador acertó 28 de ellos (y falló los 22

restantes).

Recordemos que nuestro objetivo es estimar el valor teórico de la probabilidad o proporción de aciertos p del jugador. Una forma natural de hacerlo es calcular la proporción de aciertos en nuestra muestra de 50 lanzamientos:

$$\hat{p} = \frac{28}{50} = 0.56.$$

Este valor se denomina **proporción muestral** y se denota \hat{p} , porque nos da una idea aproximada de cuál puede ser el verdadero valor del valor teórico p , que se llama **proporción poblacional**.

Decimos que la proporción muestral $\hat{p} = 0.56$ es una **estimación puntual** de la proporción poblacional p . En general, una estimación puntual de un parámetro es un valor numérico calculado a partir de la muestra que nos orienta sobre el verdadero valor del parámetro.

Hay que tener claro que la proporción muestral $\hat{p} = 0.56$ nos proporciona una estimación puntual de la proporción poblacional p , pero que el verdadero valor de p sigue siendo desconocido. El verdadero valor de p es la probabilidad de que el jugador acierte un tiro libre, que identificamos con la proporción de aciertos en el conjunto ideal de los infinitos lanzamientos que podría efectuar el jugador. Por su parte, el valor $\hat{p} = 0.56$ es tan sólo la proporción de aciertos en los 50 lanzamientos que hemos observado. Podría ser que los 50 lanzamientos observados sean una “mala racha”, y que el verdadero valor de la proporción de aciertos p sea superior a $\hat{p} = 0.56$.

1.2. Gráficos

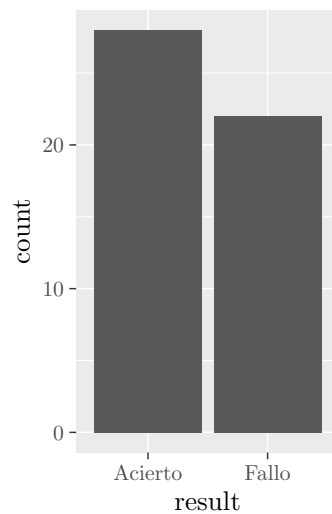
Ahora vamos a representar gráficamente los datos mediante un diagrama de barras.

En primer lugar creamos una tabla con el recuento de aciertos y fallos:

```
basket <- tibble(  
  result = c("Acierto", "Fallo"),  
  count = c(28, 22)  
)
```

Y usamos la tabla creada para crear un diagrama de barras con la función `geom_col()`:

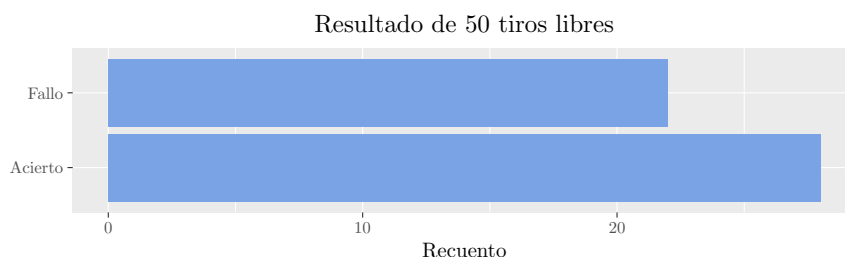

```
ggplot(  
  data = basket,  
  mapping = aes(  
    x = result,  
    y = count  
  )  
) +  
  geom_col()
```



Mejoramos un poco el aspecto inicial de nuestro gráfico de barras añadiendo rótulos, cambiando el color de las barras, y girándolo con `coord_flip()`:

```
ggplot(  
  data = basket,  
  mapping = aes(  
    x = result,  
    y = count  
  )  
) +  
  geom_col(  
    fill = "#7AA3E5"  
  ) +  
  coord_flip() +  
  labs(  
    title = "Resultado de 50 tiros libres",  
    x = "",  
    y = "Recuento")
```

```
) +
theme(plot.title = element_text(hjust = 0.5)) # centrar
título
```



2. Contrastes de hipótesis

Hasta ahora hemos descrito los datos numéricamente y gráficamente. Concretamente, hemos calculado la proporción de aciertos en la muestra, obteniendo el valor $\hat{p} = 0.56$ para la proporción muestral, como estimación de la proporción poblacional p , y hemos realizado un diagrama de barras de los recuentos de aciertos y fallos.

Nos planteamos ahora responder a la primera cuestión planteada en el problema de la introducción:

Supongamos que el jugador afirma que su probabilidad de acertar un tiro libre es 0.75. El objetivo es contrastar si los datos observados resultan compatibles con la afirmación del jugador.

Con la exploración inicial que hemos hecho de los datos, podemos concretar el objetivo planteando lo que en estadística se denomina un **contraste de hipótesis**. Dividiremos el procedimiento en tres pasos: planteamiento de las hipótesis, cálculo del p -valor y conclusión.

2.1. Planteamiento de las hipótesis

Según el jugador la proporción verdadera es $p = 0.75$. Y la estimación de p según nuestra muestra es $\hat{p} = 0.56$, inferior al valor de p según el jugador.

Como resaltamos antes, no podemos concluir con seguridad, basándonos en los datos obtenidos en la muestra de 50 lanzamientos, que el jugador mienta. Puede que el jugador diga la verdad, y que, por azar, hayamos observado una racha peor de lo habitual.

Pero nuestros datos levantan la sospecha de que el jugador exagera, y nos sugieren que $p < 0.75$. Esta hipótesis sugerida por los datos se denomina **hipótesis alternativa**, y se denota H_1 .

Por otra parte, a la igualdad $p = 0.75$, se le llama **hipótesis nula** y se denota H_0 .

Resumimos el contraste de hipótesis que planteamos escribiendo

$$\begin{cases} H_0 : p = 0.75 \\ H_1 : p < 0.75. \end{cases}$$

2.2. Cálculo del p -valor

Vamos a resolver el conflicto, entre la hipótesis nula H_0 , y la hipótesis alternativa H_1 que sugieren los datos, contestando a la pregunta siguiente:

Si la realidad fuera que $p = 0.75$ (la hipótesis nula es cierta, el jugador dice la verdad) ¿cuál es la probabilidad de observar una racha tan mala como la observada o peor?

Respondiendo a esta pregunta, cuantificaremos la sospecha que han levantado nuestros datos y que provocó la formulación de la hipótesis alternativa.

Si la respuesta es que la probabilidad es muy baja, tomaremos la decisión de acusar de mentiroso al jugador y nos decantaremos por la hipótesis alternativa de que $p < 0.75$. Si por el contrario la probabilidad no es muy baja, lo dejaremos estar y no acusaremos al jugador de mentiroso, pensaremos que los datos observados pueden deberse al azar y no a que el jugador mienta.

Adoptaremos una actitud conservadora, y fijaremos el umbral de lo que consideramos una probabilidad muy baja en 0.05. De esta manera, si los datos observados tienen una probabilidad superior a 0.05, no acusaremos de mentiroso al jugador. Esta probabilidad se llama **nivel de significación** del contraste y se denota α . Lo más frecuente en el ámbito científico es trabajar con este nivel de significación $\alpha = 0.05$, pero también pueden convenirse otros valores como $\alpha = 0.01$.

El procedimiento estadístico con el que vamos a resolver el contraste de hipótesis que hemos planteado y a calcular la probabilidad que buscamos se llama **test binomial exacto**. El comando de R que lo lleva a cabo es `binom.test()`. Concretamente para resolver nuestro contraste tenemos que ejecutar:

```
binom.test(  
  x = 28,  
  n = 50,  
  p = 0.75,  
  alternative = "less"  
)  
  
##  
## Exact binomial test  
##  
## data: 28 and 50  
## number of successes = 28, number of trials = 50, p-value =  
0.002618  
## alternative hypothesis: true probability of success is less  
than 0.75  
## 95 percent confidence interval:  
## 0.0000000 0.6801741  
## sample estimates:  
## probability of success  
## 0.56
```

Los argumentos $x = 28$ y $n = 50$ indican que hemos observado 28 aciertos en 50 lanzamientos. Con el argumento $p = 0.75$ indicamos que queremos comparar nuestros datos con el valor de referencia 0.75 de la hipótesis nula. Y el argumento `alternative = "less"` indica que nuestra hipótesis alternativa es que p es menor que el valor umbral 0.75.

La probabilidad que buscamos la encontramos en el fragmento de la salida
 $p\text{-value} = 0.002618$.

Esta probabilidad 0.002618 asociada a nuestro contraste de hipótesis se denomina **p-valor**. La letra p en p -valor es inicial de probabilidad, no tiene que ver con que estemos realizando un contraste de hipótesis sobre el parámetro p (en un contraste de hipótesis para una media también hablaríamos del p -valor de ese contraste).

En el apartado siguiente, utilizaremos el p -valor para decidir entre las hipótesis contrastadas, esto es, entre la hipótesis nula $p = 0.75$ y la hipótesis alternativa sugerida por los datos $p < 0.75$. Para entender el criterio que seguiremos, vamos a analizar cómo se calcula el p -valor que hemos obtenido con `binom.test()`.

El p -valor 0.002618 que hemos obtenido con la función `binom.test()` cuantifica

nuestra sospecha de que $p < 0.75$, en el sentido que se explica a continuación: Si fuera verdad que $p = 0.75$, la probabilidad de observar en 50 lanzamientos una racha tan mala o peor como la de nuestra muestra sería de 0.002618.

Dicho con otras palabras: si repetimos un gran número de veces el experimento (observar otros 50 lanzamientos y contabilizar los aciertos), en aproximadamente 26 de cada 10000 repeticiones observaríamos un resultado tan desfavorable como acertar tan sólo 28 de los 50 lanzamientos.

Con las explicaciones anteriores, que describen el significado del p -valor, y la teoría que conocemos de variables aleatorias, tenemos herramientas suficientes para saber cómo se ha calculado el p -valor 0.002618: se obtiene como

$$P(X \leq 28) = F(28) = \text{pbinom}(28, 50, 0.75) = 0.0026178,$$

donde hemos supuesto que $X \sim \text{Binom}(50, 0.75)$, es decir, hemos supuesto que p es igual a 0.75, el valor umbral en la hipótesis nula.

2.3. Conclusión

Recordemos que habíamos decidido adoptar una actitud conservadora, de manera que no acusaríamos al jugador de mentiroso a no ser que, en el supuesto de que diga la verdad, los datos obtenidos tuvieran una probabilidad inferior a $\alpha = 0.05$. La probabilidad obtenida ha sido 0.002618, mucho menor que nuestro “umbral de credulidad” o nivel de significación $\alpha = 0.05$. Así que nuestra conclusión es que rechazamos la afirmación del jugador en favor de nuestra hipótesis alternativa de que $p < 0.75$.

Ahora, nunca sabremos si nos hemos equivocado en nuestra decisión. Como hemos trabajado con un umbral de credulidad del 5%, podríamos decir que tenemos una confianza del 95% en que nuestra decisión es correcta. Esto es así porque sólo corremos el riesgo de equivocarnos en el 5% de las ocasiones en que, siendo verdad la afirmación del jugador, la mala racha se deba al azar. Como se indicó en la introducción esta incertidumbre acompaña a cualquier conclusión de un procedimiento de inferencia estadística (figuras 1 y 2).

Nota: Si en los 50 lanzamientos hubiéramos obtenido 40 aciertos, entonces la proporción muestral sería

$$\hat{p} = 40/50 = 0.8.$$

En ese caso plantearíamos el contraste de hipótesis

$$\begin{cases} H_0 : p = 0.75 \\ H_1 : p > 0.75. \end{cases}$$

puesto que ahora los datos sugieren como hipótesis alternativa que la probabilidad de acierto es superior al valor 0.75 indicado por el jugador. La hipótesis alternativa siempre es la sugerida por los datos. Y al realizar el contraste de hipótesis valoramos si los datos arrojan suficiente evidencia como para extrapolar lo observado en la muestra ($\hat{p} = 0.8 > 0.75$) a la población ($p > 0.75$).

El código de R para realizar el nuevo contraste y calcular su p -valor sería ahora:

```
binom.test(  
  x = 40,  
  n = 50,  
  p = 0.75,  
  alternative = "greater"  
)  
  
##  
## Exact binomial test  
##  
## data: 40 and 50  
## number of successes = 40, number of trials = 50, p-value =  
0.2622  
## alternative hypothesis: true probability of success is  
greater than 0.75  
## 95 percent confidence interval:  
## 0.6844039 1.0000000  
## sample estimates:  
## probability of success  
## 0.8
```

Al ser ahora el p -valor 0.2622 superior a $\alpha = 0.05$, no podríamos rechazar la hipótesis nula en favor de la hipótesis alternativa y concluir que la probabilidad de aciertos del jugador es superior a 0.75. Creeríamos simplemente que, por azar, hemos observado una racha de 50 lanzamientos un poco mejor de lo habitual para el jugador. No nos resulta extraña la racha observada porque la probabilidad de obtener una racha así de buena o mejor, siendo la verdadera probabilidad de acierto 0.75, es 0.2622, el p -valor.

La interpretación general del p -valor de un contraste de hipótesis es que es la

probabilidad de obtener unos datos tan extremos como el observado (extremos en el sentido de ir a favor de la hipótesis alternativa) siendo la hipótesis nula cierta. Por eso, cuando el p -valor es muy pequeño, menor que el valor de referencia $\alpha = 0.05$, rechazamos la hipótesis nula en favor de la alternativa, al parecernos muy extraño que los datos observados se deban al azar. Mientras que cuando el p -valor es superior a $\alpha = 0.05$, adoptamos la postura conservadora de no rechazar la hipótesis nula, asumiendo que hemos observado datos extremos por azar.

Una analogía útil para entender la filosofía de un contraste de hipótesis es pensar que se trata de un juicio, comparando la hipótesis nula H_0 con la presunción de inocencia, y la hipótesis alternativa H_1 con la acusación de culpabilidad sugerida por las pruebas. Solo cuando las pruebas arrojan la suficiente evidencia, abandonaremos la presunción de inocencia inicial para dictar una acusación de culpabilidad. Adoptamos la actitud conservadora de considerar que las pruebas arrojan suficiente evidencia de culpabilidad cuando su probabilidad fuera inferior a $\alpha = 0.05$ si la persona juzgada fuera inocente.

3. Intervalos de confianza

En esta sección nos ocuparemos de la siguiente parte en la salida del comando `binom.test()` que ejecutamos en la sección anterior para resolver la primera parte el problema planteado en la introducción:

```
95 percent confidence interval:
0.0000000 0.6801741
```

Esta parte de la salida nos da $(0, 0.6801741)$ como intervalo de confianza del 95% para el parámetro p , y responde a la segunda cuestión que planteamos al inicio de encontrar un conjunto de valores plausibles para p . A continuación se explica cómo ha de interpretarse este intervalo de confianza y cómo se ha obtenido.

Si la afirmación del jugador hubiera sido que su probabilidad de acierto es 0.6 hubiéramos planteado el contraste de hipótesis:

$$\begin{cases} H_0 : p = 0.6 \\ H_1 : p < 0.6. \end{cases}$$

Para calcular el p -valor de este nuevo contraste, supondríamos que

$$X \sim \text{Binom}(50, 0.6)$$

y efectuaríamos el cálculo

$$P(X \leq 28) = F(28) = \text{pbinom}(28, 50, 0.6) = 0.3299.$$

Puedes comprobarlo con el código

```
binom.test(  
  x = 28,  
  n = 50,  
  p = 0.6,  
  alternative = "less"  
)
```

Al ser el nuevo p -valor 0.3299 mayor que $\alpha = 0.05$ no podemos asegurar, con una confianza del 95%, que $p < 0.6$. Por eso, el extremo superior del intervalo, 0.6801741, es superior a 0.6.

El extremo superior de nuestro intervalo de confianza del 95% es el mayor valor p_0 para el que el contraste de hipótesis

$$\begin{cases} H_0 : p = p_0 \\ H_1 : p < p_0. \end{cases}$$

resulta en la aceptación de la hipótesis nula a nivel de significación $\alpha = 0.05$.

Puedes comprobar que el p -valor del contraste anterior es justo 0.05 para $p_0 = 0.6801741$, superior a 0.05 para cualquier valor $p_0 < 0.6801741$ en el intervalo de confianza, e inferior a 0.05 para cualquier valor $p_0 > 0.6801741$ fuera del intervalo de confianza.

Con este procedimiento para construir el intervalo, sólo excluiríamos al verdadero valor del parámetro del intervalo de confianza del 95% cuando nos equivoquemos en nuestra decisión de rechazar la hipótesis nula. Y sabemos que, trabajando a nivel de significación $\alpha = 0.05$, dicha equivocación solo ocurre en un 5% de las ocasiones. Por tanto, un intervalo de confianza del 95% contendrá al verdadero valor del parámetro con probabilidad 0.95.

En general, un intervalo de confianza de nivel $1 - \alpha$ es el rango de valores del parámetro que no serían rechazados en un contraste de hipótesis a nivel de significación α . Esto garantiza que la proporción de intervalos de nivel $1 - \alpha$ que dejan fuera al verdadero valor del parámetro es α y en consecuencia que la proporción de los que contienen al verdadero valor del parámetro es $1 - \alpha$.

Nunca sabremos si nuestro intervalo de confianza $(0, 0.6801741)$ contiene o no al verdadero valor de p . Solo sabemos cuantificar que la probabilidad de que

lo contenga es 0.95, porque se ha calculado con un procedimiento que resulta acertado en el 95% de las ocasiones.

Problema *tipo examen*

A continuación se incluye un problema del estilo de los que se podrían plantear en el examen escrito para este tema.

Problema 2

Jugamos con un amigo al *Problema de Monty Hall*.

Nuestro amigo mantiene erróneamente que la probabilidad de ganar el coche es $1/3$, tanto si se cambia de puerta como si no.

Pero nosotros sabemos que la probabilidad de ganar el coche si no cambiamos de puerta es $1/3$, mientras que la probabilidad de ganarlo cambiando de puerta es $2/3$.

Sin embargo, no conseguimos convencer a nuestro amigo de su error utilizando argumentos teóricos. Por tanto, nos planteamos convencerlo de que el valor del parámetro

p = probabilidad de ganar el coche cambiando de puerta

es mayor que $1/3$, basándonos en los resultados de varias simulaciones del juego cambiando de puerta.

Se realizaron 66 simulaciones del juego cambiando de puerta y se ganó el coche en 48 de ellas (obteniendo una cabra en las 16 restantes).

- (a) Plantea el contraste de hipótesis para comparar p con $1/3$ y completa el siguiente código para realizar el test binomial exacto para resolverlo (a nivel de significación 0.05):

```
binom.test(  
  x =      ,  
  n =      ,  
  p =      ,  
  alternative =  
)
```

- (b) Al ejecutar el código del apartado anterior, con los valores adecuados para los argumentos, se obtiene la siguiente salida:

```
##
## Exact binomial test
##
## data:  x and n
## number of successes = 48, number of trials = 66,
p-value = 7.091e-11
## alternative hypothesis: true probability of success
is greater than 0.3333333
## 95 percent confidence interval:
##  0.6228693 1.0000000
## sample estimates:
## probability of success
##           0.7272727
```

Utiliza el p -valor obtenido para decidir si los datos conseguirán convencer a nuestro amigo de que p es mayor que $1/3$. Indica además cómo se interpreta y calcula dicho p -valor.

- (c) Basándote en el intervalo de confianza en la salida del comando `binom.test()` dada en el apartado anterior, responde a la siguiente pregunta: Si planteamos el contraste de hipótesis

$$\begin{cases} H_0 : p = 0.65 \\ H_1 : p > 0.65 \end{cases}$$

¿cuál sería la conclusión a nivel de significación 0.05?

Solución:

- (a) En nuestra muestra se ganan $x = 48$ coches en un total de $n = 66$ juegos de forma que la proporción muestral es

$$\hat{p} = \frac{x}{n} = \frac{48}{66} = 0.73.$$

Puesto que la proporción muestral calculada es mayor que $1/3$, podemos formular el contraste de hipótesis

$$\begin{cases} H_0 : p = 1/3 \\ H_1 : p > 1/3. \end{cases}$$

El código de R para realizar el contraste planteado sería

```
binom.test(
  x = 48,
  n = 66,
  p = 1 / 3,
  alternative = "greater"
)
```

- (b) Al ser el p -valor obtenido 7.091×10^{-11} menor que el nivel de significación 0.05, rechazamos la hipótesis nula en favor de la alternativa, y concluimos que los datos arrojan suficiente evidencia para afirmar que $p > 1/3$ y convencer a nuestro amigo de su error.

En general, el p -valor de un contraste de hipótesis es la probabilidad de obtener datos tanto o más extremos que los observados en la muestra si la hipótesis nula fuera cierta. En este caso, el p -valor 7.091×10^{-11} indica la probabilidad de ganar 48 o más coches en 66 repeticiones del juego si la probabilidad de ganar fuera $1/3$, como cree nuestro amigo.

Para reproducir el cálculo del p -valor consideramos la variable aleatoria

X = Número de veces que se gana el coche en 66 juegos
y calculamos la probabilidad

$$P(X \geq 48) = 1 - F(47) = 1 - \text{pbinom}(47, 66, 1/3)$$

que resulta en el mismo valor de la salida de R. ❖

- (c) En general, un intervalo de confianza del 95% es el rango de valores del parámetro que no serían rechazados en un contraste de hipótesis a nivel de significación 0.05.

Así nuestro intervalo de confianza (0.6228693, 1) incluye a los valores p_0 que serían aceptados en el contraste

$$\begin{cases} H_0 : p = p_0 \\ H_1 : p > p_0. \end{cases}$$

En particular, como 0.65 pertenece al intervalo (0.6228693, 1), el contraste

$$\begin{cases} H_0 : p = 0.65 \\ H_1 : p > 0.65 \end{cases}$$

resultaría en la aceptación de la hipótesis nula, sin aportar los datos observados suficiente evidencia para rechazar que $p = 0.65$ y concluir que $p > 0.65$ a nivel de significación 0.05.

Nota: Puedes comprobar con el comando `binom.test()` que el p -valor del contraste propuesto es 0.1163, que significa que, siendo

$p = 0.65$, y no mayor, la probabilidad de ganar 48 coches o más en 66 juegos es 0.1163. Por tanto, trabajando a nivel de significación 0.05, resulta compatible creer que $p = 0.65$ y que el elevado número de coches ganados se debe al azar, y no a que $p > 0.65$.

Aunque sabemos que el valor real de p es $2/3 = 0.\bar{6}$ el contraste de hipótesis no ha tenido suficiente potencia para concluir que $p > 0.65$ a partir de nuestros datos. ❖

Bibliografía

Para ampliar conocimientos de este tema puedes consultar los Capítulos 6, 7, 8 y 9 de [1] y los Capítulos 6, 7, 8 y 9 de [2].

- [1] Janet Susan Milton y Jesse C. Arnold. *Introduction to probability and statistics: principles and applications for engineering and the computing sciences*. 2003. ISBN: 007246836X.
- [2] Douglas C. Montgomery y George C. Runger. *Applied Statistics and Probability for Engineers, 5th Edition*. Wiley, 2010, pág. 784. ISBN: 1118050177.

