Clase popurri

2025-04-01

La Catadora de te (Ejercio 5 de la guia 3 Teorica)

Una señora afirma que al probar una taza de té con leche puede distinguir qué fue lo primero que se echó en la taza: el té o la leche.

¿Cómo podríamos hacer un experimento para determinar si la señora sabe distinguir?

Procuramos realizar un experimento que indique evidencia a favor de la señora.

Para ello, procedemos de la siguiente manera: le damos $n = 12$ tazas preparadas de diferentes maneras y contamos el número de veces que la mujer acierta el orden en el que fue preparado el té.

Sea X la variable aleatoria que cuenta el número de aciertos en los $n = 12$ ensayos.

Distribución de la variable aleatoria X

Dado que contamos cantidad de aciertos sobre un total conocido (n) sin considerar el orden, X será correctamente modelada con una distribución Binomial de parámetros $n = 12$ y probabilidad de acierto p:

$$X \sim \text{Bin}(12, p)$$

Con función de probabilidad puntual:

$$P(X = k) = \binom{12}{k} p^k (1 - p)^{12-k}$$

Probabilidad de acertar todas las tazas al azar

Si la señora responde al azar ($p = 0.5$), la probabilidad de acertar todas las 12 tazas es:

```r
f_binomial <- function(n=12, p=0.5, aciertos){
  choose(n, aciertos) * p^aciertos * (1 - p)^(n - aciertos)
}

prob_12 <- f_binomial(n=12, p=0.5, aciertos=12)
prob_12
```

```
## [1] 0.0002441406
```

O bien usando la funcion de R

```r
dbinom(12, size=12, p=0.5)
```

```
## [1] 0.0002441406
```

La probabilidad de que acierte las 12 tazas si responde al azar es de aproximadamente 0.0244%.

Esto implica que si acierta las 12 tazas, entonces:

- Es muy poco probable (0.0244%) que haya elegido sus respuestas al azar.
- Muy probablemente (99.975%) **NO** sea una respuesta al azar.

$$P(X = 12) = 2.441 \times 10^{-4}$$

Evaluación de distintos criterios de decisión

Supongamos que la señora responde al azar. Para cada uno de los criterios propuestos, calculamos la probabilidad de que consideremos que la señora sabe.

Definimos los eventos:

- C_0: "12 aciertos en los 12 ensayos (Criterio 0)"
- C_1: "al menos 11 aciertos en los 12 ensayos (Criterio 1)"
- C_2: "al menos 10 aciertos en los 12 ensayos (Criterio 2)"
- C_3: "al menos 9 aciertos en los 12 ensayos (Criterio 3)"

Calculamos las probabilidades utilizando la función de distribución acumulada de una Binomial $(12, 0.5)$:

```
prob_C0 <- dbinom(12, size=12, p=0.5)
prob_C1 <- sum(dbinom(11:12, size=12, p=0.5))
prob_C2 <- sum(dbinom(10:12, size=12, p=0.5))
prob_C3 <- sum(dbinom(9:12, size=12, p=0.5))

c(prob_C0, prob_C1, prob_C2, prob_C3)
```

```
## [1] 0.0002441406 0.0031738281 0.0192871094 0.0729980469
```

Estos valores nos permiten evaluar qué tan estrictos o flexibles son los diferentes criterios a la hora de aceptar que la señora posee la habilidad que afirma.

```
X <- data.frame(
  Criterio = c(0, 1, 2, 3),
  `Proba Decidir que sabe` = c(prob_C0, prob_C1, prob_C2, prob_C3),
  NoSabe = 1 - c(prob_C0, prob_C1, prob_C2, prob_C3)
)
X
```

```
##   Criterio Proba.Decidir.que.sabe    NoSabe
## 1        0           0.0002441406 0.9997559
## 2        1           0.0031738281 0.9968262
## 3        2           0.0192871094 0.9807129
## 4        3           0.0729980469 0.9270020
```

Las probabilidades calculadas para cada criterio modelan nuestra confianza en cada evento:

- Si acierta 12 de 12, nuestra confianza es muy alta, pues la probabilidad de que este evento ocurra al azar es muy baja.
- Acertar 9 de 12 es un evento más probable si se actúa al azar, por lo que confiamos menos en que acierte a causa de sus habilidades y más en que sea producto del azar.

Buscamos los criterios que indican una probabilidad menor o igual a 0.02:

```
X[X$`Proba de equivocarnos` <= 0.02, ]
```

```
## [1] Criterio               Proba.Decidir.que.sabe NoSabe
## <0 rows> (or 0-length row.names)
```

El criterio 3 resulta demasiado laxo para nuestra exigencia.

Podemos elegir entre los criterios 0, 1 y 2, siendo 0 el más estricto (muy bajas probabilidades de equivocarnos, pero demasiado exigente) y el criterio 2 el más cercano al margen de error máximo requerido.

Supongamos que la señora responde bien en 10 ocaciones. Los criterios que nos permiten decidir que la señora sabe son el 2 y el 3. Pero el criterio mas exigente para decidir que la señora sabe es el 2. En ese caso la probabilidad de decidir que sabe es

```
prob_C2
```

```
## [1] 0.01928711
```

Decidir que sabe coincide con la probabilidad de quivocarse. Ya que estamos asumiendo que responde al azar. Es decir estamos decidiendo que sabe cuando en realidad respondio al azar.

La probabilidad de observar un valor tan o más grande que el observado, asumiendo que responde al azar será la misma que la probabilidad de equivocarnos.

###El Don

Ahora si sabemos que la señora realmente tiene un don y acieta con porbabilidad 0.8. Se tiene que la variable aleatoria X ahora tendra distribución Binomial $(12, 0.8)$. Entonces podemos calcular la probabilida de cada criterio usando esta distribucion.

```
prob_C0 <- dbinom(12, size=12, p=0.8)
prob_C1 <- sum(dbinom(11:12, size=12, p=0.8))
prob_C2 <- sum(dbinom(10:12, size=12, p=0.8))
prob_C3 <- sum(dbinom(9:12, size=12, p=0.8))

c(prob_C0, prob_C1, prob_C2, prob_C3)
```

```
## [1] 0.06871948 0.27487791 0.55834575 0.79456895
```

```
X <- data.frame(
  Criterio = c(0, 1, 2, 3),
  `Proba de Decidir que sabe` = c(prob_C0, prob_C1, prob_C2, prob_C3),
  NoSabe = 1 - c(prob_C0, prob_C1, prob_C2, prob_C3)
)
X
```

```
##   Criterio Proba.de.Decidir.que.sabe    NoSabe
## 1        0                0.06871948 0.9312805
## 2        1                0.27487791 0.7251221
## 3        2                0.55834575 0.4416543
## 4        3                0.79456895 0.2054311
```

Muy distinto a los valores obtenidos suponiendo que la señora actuaba al azar, ahora con muy pocos errores, la probabilidad de equivocarnos aumenta considerablemente.

Y ahora si acierta con probabilidad $p = 0.9$

```
prob_C0 <- dbinom(12, size=12, p=0.9)
prob_C1 <- sum(dbinom(11:12, size=12, p=0.9))
prob_C2 <- sum(dbinom(10:12, size=12, p=0.9))
prob_C3 <- sum(dbinom(9:12, size=12, p=0.9))

c(prob_C0, prob_C1, prob_C2, prob_C3)
```

```
## [1] 0.2824295 0.6590023 0.8891300 0.9743625
```

```
X <- data.frame(
  Criterio = c(0, 1, 2, 3),
  `Proba de Decidir que sabe` = c(prob_C0, prob_C1, prob_C2, prob_C3),
  NoSabe = 1 - c(prob_C0, prob_C1, prob_C2, prob_C3)
```

```
)
X
```

```
##   Criterio Proba.de.Decidir.que.sabe     NoSabe
## 1        0                 0.2824295 0.71757046
## 2        1                 0.6590023 0.34099775
## 3        2                 0.8891300 0.11086998
## 4        3                 0.9743625 0.02563747
```

Similar a $p = 0.8$, nuestra confianza en la decisión que tomemos se reduce en gran medida con muy pocos fracasos observados.

Finalmente la probabilidad de decidir que la señora no sabe disminuye a medida que hacemos criterios mas laxos. Y disminuye a medida que aumenta p.

La Urna de Polya

Consideremos el siguiente experimento llamado **Esquema de Polya**. De un bolillero que contiene B bolillas blancas, $B \geq 1$ y R rojas, $R \geq 1$, se extraen sucesivamente y al azar n bolillas, $n \geq 2$, devolviendo cada bolilla extraída al bolillero junto con otras c bolillas del mismo color, $c \geq 1$.

Sea Rn salio una roja en la extraccion n

(1) Probabilidad de obtener una roja en la primera extracción

$P(R1) = \frac{R}{B+R}$

(2) Probabilidad de obtener una roja en la segunda extracción

$P(R2) = P(R2 \mid R1)P(R1) + P(R2 \mid B1)P(B1)$ $P(R2) = \frac{R+c}{R+B+c}\frac{R}{R+B} + \frac{R}{R+B+c}\frac{B}{R+B} = \frac{R}{R+B}$

(3) Probabilidad de obtener una roja en la n-ésima extracción

Se resuelve por induccion. La induccion es si tengo una unrna inicial con M rojas y K blancas la probabilidad de que en ela extraccion $n-1$ sea roja es $(M/(M+K)$

$P(Rn) = P(Rn \mid R1)P(R1) + P(Rn \mid B1)P(B1)$

aca hay que usar la hipotesis inductiva en $P(Rn \mid R1)$ y en $P(Rn \mid B1)$. $P(Rn) = \frac{R+c}{R+B+c}\frac{R}{R+B} + \frac{R}{R+B+c}\frac{B}{R+B} = \frac{R}{R+B}$