

Gráfica de las probabilidades del experimento del té con leche

Karla E. Juárez Contreras

29/10/2023

Distribución de probabilidad

El siguiente código está basado en el experimento descrito por R. Fisher en su libro de 1935 “The Design of experiments”. En él, R. Fisher pidió a una mujer probar 8 tazas de té. En 4 de ellas, el té se había adicionado antes de la leche (T), y en las otras 4, se había adicionado la leche antes que el té (L). El le pidió a la mujer que indicará las 4 tazas que había adicionado leche, y para su sorpresa, acertó en las 4 tazas. R. Fisher calculó que, si la mujer tomaba 4 de las 8 tazas preparadas, habría 70 posibles maneras de tomarlas, ya que la combinatoria de 4 en 8 es igual a 70. Ya que solo hay 1 combinación posible para escoger correctamente las 4 tazas a las que se les habría agregado leche (LLLL), entonces la probabilidad de obtener ese resultado sería $1/70$. También calculó la probabilidad de acertar por azar a 3 de las 4 tazas, 2 de las 4 tazas, 1 de las 4 tazas y 0 de las 4 tazas. A continuación se muestran los comandos para calcular esas probabilidades en R:

```
# Se muestran a continuación las operaciones
```

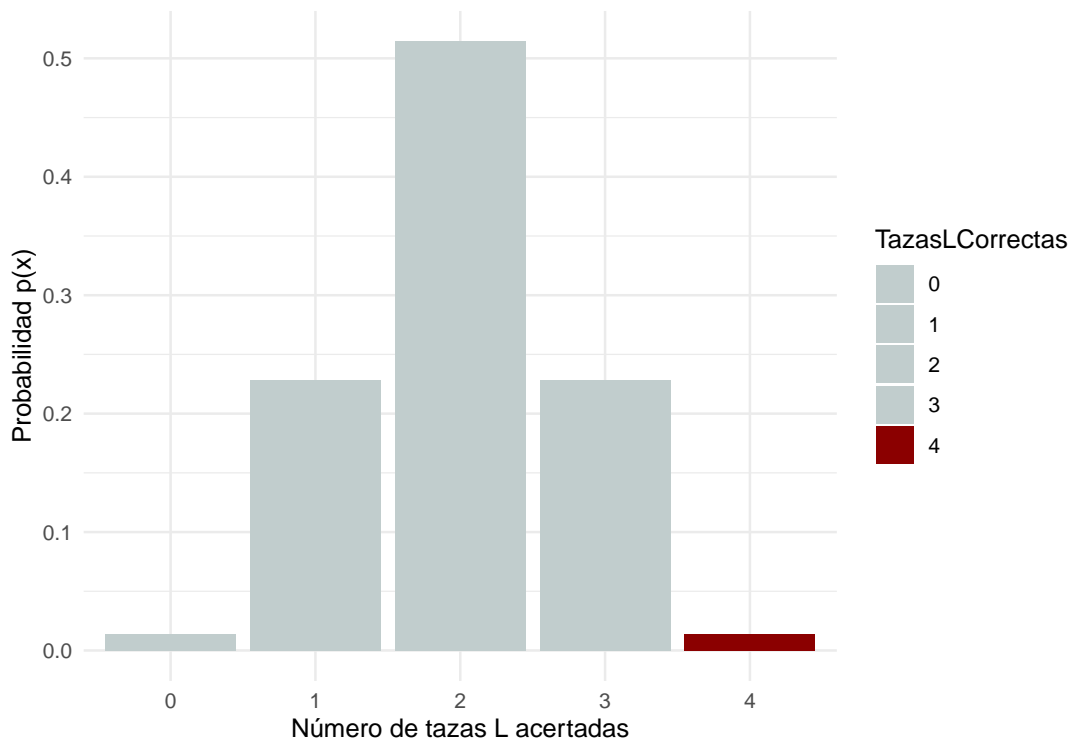
```
P_x0<-(choose(4, 0) * choose(4, 0))/choose(8, 4)
P_x1<-(choose(4, 1) * choose(4, 1))/choose(8, 4)
P_x2<-(choose(4, 2) * choose(4, 2))/choose(8, 4)
P_x3<-(choose(4, 3) * choose(4, 3))/choose(8, 4)
P_x4<-(choose(4, 4) * choose(4, 4))/choose(8, 4)
```

```
Probabilidades<- c(P_x0, P_x1, P_x2, P_x3, P_x4 )
TazasLCorrectas<- c(0,1,2,3,4)
datos<-data.frame("TazasLCorrectas"=TazasLCorrectas, "Probabilidades"=Probabilidades)
datos$TazasLCorrectas<-as.factor(datos$TazasLCorrectas)
datos
```

```
##   TazasLCorrectas Probabilidades
## 1                0      0.01428571
## 2                1      0.22857143
## 3                2      0.51428571
## 4                3      0.22857143
## 5                4      0.01428571
```

A continuación se grafica en barras la distribución de probabilidad:

```
library(ggplot2)
ggplot(datos, aes(x=TazasLCorrectas, y=Probabilidades, fill=TazasLCorrectas)) +
  geom_bar(stat="identity", position=position_dodge())+
  scale_fill_manual(values=c("azure3", "azure3", "azure3", "azure3", "darkred"))+
  labs(x="Número de tazas L acertadas", y="Probabilidad p(x)") +
  theme_minimal()
```



La hipótesis nula de la que él partió fue que no habría asociación entre la sustancia agregada primero a la taza (té o leche) y lo que indicara la señora (e.g. la señora estaba adivinando). Y la única manera de rechazar esa hipótesis nula era obtener el resultado tan extremo de acertar a las 4 tazas correctamente, pues llegar a ese resultado por azar era muy poco probable (0.01428). De esta manera, pudo rechazar la hipótesis nula. En amarillo se muestra esa probabilidad en la gráfica, la cual se conoce como el p-value de la prueba, definida como la probabilidad de obtener un resultado por azar igual o más extremo que el hallado en el estudio, considerando a la hipótesis nula como cierta.

““