Probabilidad

Introducción a la Estadística

Fiona Franco Churruarín
fionafch96@gmail.com

Universidad Torcuato Di Tella

Febrero 2022

Outline

Problema del cumpleaños

- Las reglas simples de probabilidad usualmente no se usan en forma correcta en cálculos mentales.
- En un grupo de 30 personas, ¿cuál es la probabilidad de que dos personas cumplan años el mismo día? (Supongan que no es año bisiesto y no hay personas gemelas)

Juego del cumpleaños

Para un grupo de 30 personas, la probabilidad es de 0,71!!!

Algunos comentarios

- Las reglas fundamentales de probabilidad usualmente no se usan en forma correcta en cálculos mentales.
- En general, ignoramos las reglas de probabilidad.
- Tendemos a ignorar la cantidad de casos posibles sobre los cuáles queremos computar una probabilidad.
- Tendemos a ignorar la influencia del tamaño muestral cuando nos basamos en estimaciones.

Definiciones

Para definir la probabilidad necesitamos algunos conceptos de base para describir la probabilidad de distintos "fenómenos".

- **Experimento aleatorio**: proceso con resultado incierto. Puede repetirse en condiciones uniformes para la obtención de "datos".
- **Resultados**: realizaciones del experimento aleatorio. El resultado es incierto, pero cada resultado particular tiene una probabilidad de ocurrencia. Existe una distribución regular de repeticiones de cada resultado después de un gran número de realizaciones del experimento.
- **Espacio muestral** (Ω): conjunto que tiene todos los posibles resultados de un experimento aleatorio.
- **Evento**: algún subconjunto de Ω.

Pensar estos conceptos en el experimento "tirar una moneda dos veces".

Definiciones

Usualmente, en un curso de Probabilidad y Estadística se presentan varias definiciones de "probabilidad".

1. Probabilidad clásica.
2. Probabilidad como frecuencia relativa.
3. Probabilidad subjetiva.
4. Definición axiomática de la probabilidad.

No vamos a entrar en (3) en este curso.

Definiciones

Probabilidad clásica

En una muestra, la de un evento es la cantidad de veces que ocurrió ese evento, $n(A)$, respecto del tamaño de la muestra, n:

$$f_A = \frac{n(A)}{n} = \frac{\text{casos favorables}}{\text{casos posibles}}$$

Por ejemplo recordemos el caso de la moneda arrojada 5 veces y la muestra $\{$cara, cruz, cruz, cara, cara$\}$. Si estudiamos la situación incierta 'se arroja una moneda al aire' entonces tenemos un tamaño de muestra $n = 5$, y la frecuencia relativa del evento A='sale cara' es:

$$f_A = \frac{n(A)}{n} = \frac{3}{5}$$

Definiciones

La definición clásica de probabilidad suele utilizarse con los llamados **problemas de conteo**, dado que permiten calcular de una forma sencilla la cantidad de casos favorables y posibles.

- Permutaciones.
- Combinaciones.
- Variaciones.

Definiciones

Esta definición es útil para una gran cantidad de casos, pero puede fallar. ¿Qué ocurre cuando el número de casos posibles es infinito?
En una muestra, la **frecuencia relativa** de un evento es la cantidad de veces que ocurrió ese evento, $n(A)$, respecto del tamaño de la muestra, n:

$$f_A = \frac{n(A)}{n}$$

El enfoque **frecuentista** de la probabilidad establece lo siguiente respecto de que es exactamente la probabilidad de un evento:

$$\mathrm{Prob}(A) = \lim_{n\to\infty} f_A = \lim_{n\to\infty} \frac{n(A)}{n}$$

Donde $\mathrm{Prob}(A)$ o también $P(A)$ denota la probabilidad del evento A.

Definiciones

Es decir, la probabilidad de un evento es la frecuencia relativa de ese evento sobre el total de realizaciones de una situación incierta, cuando la cantidad de realizaciones tiende a infinito.

En nuestro ejemplo, la probabilidad de obtener cara en una tirada es la cantidad de caras sobre el total de tiros, si tiráramos la moneda infinitas veces.

Otros ejemplos: productos con fallas, tasa de letalidad de una enfermedad.

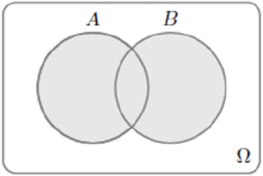
Conjuntos o eventos

Para definir correctamente la probabilidad se necesitan algunos resultados básicos de Teoría de Conjuntos.
Sean A y B dos eventos cualesquiera del espacio muestral Ω:

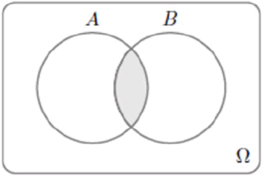
- $A \cup B$ es el evento 'A o B o ambos' y se denomina la **unión** de A y B.
- $A \cap B$ es el evento 'A y B' y se denomina la **intersección** de A y B.
- A^c o $\overline{A}$ es el evento 'no A' y se denomina la **complemento** de A.
- Si A y B son disjuntos (no tienen elementos en común), es decir, $A \cap B = \emptyset$ (donde $\emptyset$ es un conjunto vacío), se dice que los eventos son **mutuamente excluyentes**.

Ejemplos.

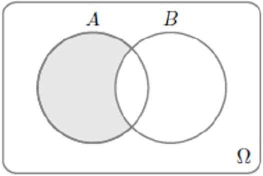
Diagrama de Venn



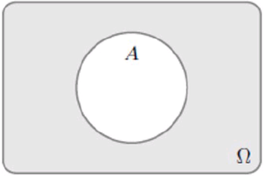
$A \cup B$



$A \cap B$



$A - B$



A^c

Definición Axiomática de la Probabilidad

Las probabilidades se definen a partir de estos 3 axiomas, entonces siempre deben cumplirse.

Axiomas de Kolmogorov

Sea A un evento y Ω el espacio muestral, una medida de probabilidad P satisface:

(a) **No negatividad.** $P(A) \geq 0, \ \forall A \subseteq \Omega$

(b) **Las probabilidades suman 1.** $P(\Omega) = 1$

(c) **Regla de la suma.** Si A y B son eventos mutuamente excluyentes, $P(A \cup B) = P(A) + P(B)$

Reglas adicionales

De estos axiomas se pueden derivar otras reglas que cumplen las probabilidades.

Regla de la suma

Para dos eventos A y B cualesquiera,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Si A y B son mutuamente excluyentes, entonces $P(A \cap B) = 0$, y por lo tanto,

$$P(A \cup B) = P(A) + P(B)$$

Reglas de probabilidad

Reglas adicionales

(d) **Las probabilidades están entre 0 y 1.**
(a) y (b) implican que $0 \leq P(A) \leq 1$

- Si la probabilidad es 0 hay certeza de que ese evento NO va a ocurrir
- Si la probabilidad es 1 hay certeza de que ese evento SI va a ocurrir

(e) **Probabilidad del complemento.** $P(A) = 1 - P(A^c)$

(f) **Probabilidad de un conjunto vacío.** $P(\emptyset) = 0$

(g) **Probabilidad de un subconjunto.** A, B eventos en Ω ($A, B \subseteq \Omega$), si $B \subseteq A$, entonces $P(B) \leq P(A)$

Espacios Equiprobables

Entonces... ¿Qué pasa con la defininción de casos favorables/casos posibles? Veamos que este caso contiene al otro.
Llamamos **espacios equiprobables** a aquellos espacios muestrales en los cuales rige un argumento de simetría entre las probabilidades de los distintos elementos, y bajo el cual no podemos favorecer a la ocurrencia de un elemento particular por sobre otro. Por ejemplo, si mezclamos un mazo de cartas y sacamos luego una al azar, ninguna carta se ve favorecida por sobre otra para ser escogida. Es decir, todas tienen la misma probabilidad de ser escogidas.

(b), (c) y el argumento de espacio equiprobable implican algo muy particular. Notar por ejemplo si tiramos una moneda balanceada:

$$\begin{aligned}
1 = P(\Omega) &= P(cara \cup cruz) \\
&= P(cara) + P(cruz) \\
&= 2 \times P(cara) \quad \text{pues el espacio es equiprobable} \\
&\Rightarrow P(cara) = P(cruz) = 1/2
\end{aligned}$$

Espacios Equiprobables

De manera un poco mas general, consideremos tirar un dado balanceado (es decir todas las caras tienen la misma probabilidad), y calculemos la probabilidad del evento 'A=sale un número par':

$$\begin{aligned} 1 = P(\Omega) &= P(par \cup impar) \\ &= P(par) + P(impar) \\ &= P(1) + P(2) + P(3) + P(4) + P(5) + P(6) \\ &= 6 \times P(1) \quad \text{pues el espacio es equiprobable} \\ &\Rightarrow P(1) = P(2) ... = P(6) = \frac{1}{6} \end{aligned}$$

Entonces:

$$P(par) = P(1) + P(2) + P(3) = \frac{3}{6}$$

En general, si el espacio es equiprobable: $P(A) = \frac{\textit{casos favorables a A}}{\textit{casos posibles}}$

Problema del cumpleaños - Solución

- Un truco muy usado para calcular probabilidades de eventos que parecen engorrosas es reescribir el evento que estamos analizando. En este caso podemos comenzar por calcular la probabilidad de que las n personas cumplan años todas en distintos días.
- Si en el año hay 365 días y tenemos $n = 2$ personas, la probabilidad de que **no** hayan nacido el mismo día es:

$$(365/365) \times (364/365) = 0,99726$$

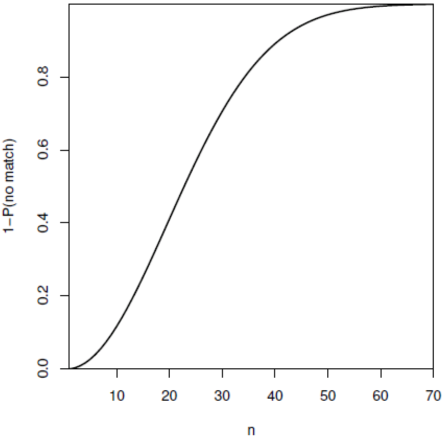
  Entonces, la probabilidad de que **sí** hayan nacido el mismo día es: $1 - 0.99726 = 0.00274$
- Si extendemos el análisis a n personas, todas ellas cumpliendo un día diferente:

$$\frac{365}{365} \cdot \frac{364}{365} \ldots \frac{365-n+1}{365} = \frac{365!/(365-n)!}{365^n}$$

- Hacer el cálculo para $n = 30$.

Problema del cumpleaños - análisis

Calculando en R las probabilidades a medida que $\uparrow n$:



Nro de personas	20	23	30	40	50	60
1-P(no match)	0.411	0.501	0.706	0.891	0.970	0.994

Ejercicio 1

Se extrae una carta de un mazo de 52 cartas (13 de cada palo):

- $A=$ extraer un as
- $R=$ extraer un rey
- $T=$ extraer trébol

(a) ¿Cuál es la probabilidad de obtener un as o un rey?
(b) ¿Cuál es la probabilidad de obtener un as o una carta de trébol?
(c) ¿Cuál es la probabilidad de obtener un as y una carta de trébol?
(d) ¿Cuál es la probabilidad de obtener un as y un rey?

Ejercicio 2

Para ingresar al HomeBanking necesita una contraseña de 5 caracteres, cada caracter puede ser una de las 26 letras (a-z) o uno de los 10 números (0-9). El primer caracter tiene que ser una letra.

Suponga que un hacker selecciona una contraseña al azar. Determine las probabilidades de que la contraseña seleccionada por el hacker:

- A: empiece con una vocal (a, e, i, o, u);
- B: termine con un número impar (1, 3, 5, 7, 9);
- $A \cup B$: empiece con una vocal o termine con un número impar

Probabilidad condicional

La probabilidad de un evento cuando se dispone de información parcial sobre el resultado de un experimento se la denomina **probabilidad condicional**.

Definición: probabilidad condicional

Sean A y B dos eventos. La probabilidad condicional del evento A dado que el evento B ha ocurrido (o la probabilidad A dado B) se denota con $P(A|B)$ y es

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Análogamente, la probabilidad condicional de B dado A se denota con $P(B|A)$ y es

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

Notar que siempre la probabilidad del denominador debe no ser nula.

Probabilidad condicional

Comentario: puede entenderse como que ahora miramos un nuevo espacio muestral que se reduce al evento B, y para que sigan valiendo los axiomas de probabilidad, reescalamos la probabilidad de la intersección con $\frac{1}{P(B)}$.

Ejercicio. El lugar donde la gente acude para obtener noticias varía de acuerdo a la edad. Se llevó a cabo un estudio al respecto[1] sobre una base de 200 encuestados entre 36 y 50 años de edad y 200 encuestados de más de 50 años. De los 200 encuestados entre 36 y 50, 82 obtenían las noticias principalmente de los diarios. De los 200 encuestados con más de 50 años, 104 obtenían las noticias principalmente de los diarios.

(a) Dado que el encuestado tiene más de 50, ¿cuál es la probabilidad de que obtenga las noticias principalmente de los diarios?

(b) Dado que el encuestado obtiene las noticias principalmente de los diarios, ¿cuál es la probabilidad de que tenga más de 50?

[1] Datos extraídos de P. Johnson "Young People Turn to the Web for News", USA Today, March 23, 2006.

Regla de la multiplicación

Notar que de la definición de probabilidad condicional con $P(B) \neq 0$:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Se desprende una regla para calcular la probabilidad de ocurrencia de dos eventos a la vez, es decir $P(A \cap B)$.

Regla de la multiplicación

Sean A y B dos eventos. Entonces, la probabilidad de la intersección de A y B se puede calcular como

$$P(A \cap B) = P(B)P(A|B) = P(A)P(B|A)$$

Nota: en la igualdad del medio se requiere $P(B) \neq 0$ y en la igualdad de la derecha se requiere $P(A) \neq 0$.

Ley de probabilidad total

La **ley de la probabilidad total** se utiliza para determinar la probabilidad incondicional de un evento, dadas las probabilidades condicionales.

Definición: ley de probabilidad total

Sea A un evento y sean los eventos $B_1, B_2, ..., B_N$ una 'partición' de Ω (*). La probabilidad total de A

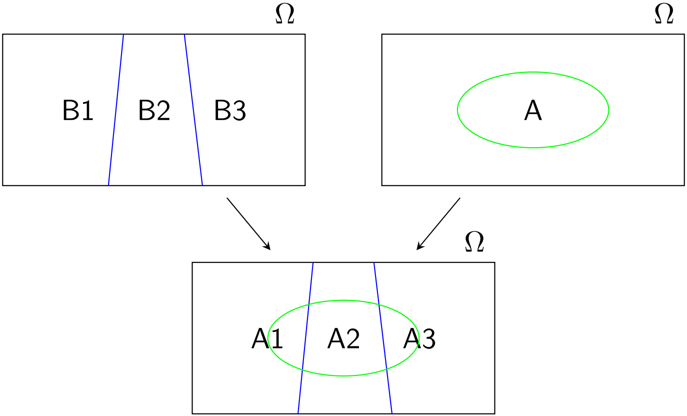
$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + ... + P(A|B_N)P(B_N)$$

(*): Es decir, los B_i, $i = 1, ..., n$ no tienen elementos en común, y la unión de todos ellos forma Ω. También se dice que los eventos $B_1, B_2, ..., B_N$ son mutuamente excluyentes y colectivamente exhaustivos.

Ley de la probabilidad total

Recordando la regla de la multiplicación y aplicándola en la definición de la ley de la probabilidad total.

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + ... + P(A \cap B_N)$$



Probabilidad total - Ejercicio

Van a reemplazar a tu jefe en el trabajo, los candidatos son:

- Pedro, con una probabilidad del 60%
- Juan, con una probabilidad del 30%
- Elena, con una probabilidad del 10%

En función de quién sea tu próximo jefe, la probabilidad de que te suban el sueldo es la siguiente:

- Si eligen a Pedro, del 5%
- Si eligen a Juan, del 20%
- Si eligen a Elena, del 60%

¿Cuál es la probabilidad de que te suban el sueldo?

Aplicación: preguntas sensibles

Supongamos que queremos aprender sobre la cantidad de gente que usa drogas ilegales. Ciertas preguntas pueden tener un costo social asociado a una respuesta:

- ¿Usás drogas ilegales?
- ¿Alguna vez hiciste trampa en un examen?
- ¿Alguna vez robaste?
- ¿Tenés sexo desprotegido?

Preguntas incómodas, pero que necesitamos formular para estimar, por ejemplo, la proporción de personas que tienen sexo desprotegido. Sólamente con formular dicha pregunta, no tendríamos una encuesta bien diseñada: existen incentivos a mentir.

Una posible solución

Analicemos una posible solución utilizando las herramientas que estudiamos hasta ahora.

La pregunta incómoda

- A los encuestados se les da una moneda y se les pide que la arrojen en privado, sin que nadie vea el resultado.
- Si toca "cara", responden la **pregunta de interés** (e.g. ¿Alguna vez ha consumido drogas ilegales?)
- Si toca " ceca", responden una **pregunta inocua** (e.g. ¿Naciste en la primera mitad del año, de enero a junio?)
- El encuestado reporta "**sí**" o "**no**", pero no dice qué pregunta es realmente contestada.
- A partir de una muestra de respuestas afirmativas se puede estimar el parámetro de interés, como la proporción de personas que alguna vez consumieron drogas ilegales. **¿Cómo funciona?**

La pregunta incómoda

Sean:

- S y N eventos que denotan las respuestas de "sí" y "no", respectivamente;
- P_{inc} el evento que denota la pregunta incómoda y P_{ino} la pregunta inocente o inocua.

El parámetro que nos interesa estimar es:

$$p = P(S|P_{inc})$$

El truco está en que la pregunta inocua o inocente sea: (a) fácil de contestar, (b) tiene una probabilidad conocida de "sí" y "no". En este caso asumamos que es del 50% en cada caso.

La pregunta incómoda

Consideren la probabilidad incondicional $P(S)$, por la **regla de la probabilidad total**:

$$\begin{aligned} P(S) &= P(S|P_{inc}) \cdot P(P_{inc}) + P(S|P_{ino}) \cdot P(P_{ino}) \\ &= p\left(\frac{1}{2}\right) + \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) \\ &= \frac{p}{2} + \frac{1}{4} \end{aligned}$$

La pregunta incómoda

Cuando la encuesta se realiza sobre n personas, la muestra final consiste en n sí's y no's. La proporción de sí's puede usarse como una estimación de $P(S)$. Así,

$$\frac{p}{2} + \frac{1}{4} = P(S) \approx \frac{\text{Nro de sí's en la muestra}}{n}$$

Despejando p:

$$p \approx 2\left(\frac{\text{Nro de sí's en la muestra}}{n} - \frac{1}{4}\right)$$

Lectura recomendada: Bolstad et al. (2001) "Sex, Drugs and Rock & Roll Survey in a First-Year Service Course in Statistics", *The American Statistician* 55, 145-149.

Independencia

Definición: Independencia de eventos

Diremos que dos eventos A y B son independientes si al conocer que ocurrió el evento B, la probabilidad de que ocurra A no se ve alterada. Es decir,

$$P(A|B) = P(A)$$

Por ejemplo, si A='El tipo de cambio nominal de depreciará mañana' y B='Esta clase durará mas de 3 horas', podemos suponer que A y B son independientes, y la probabilidad de que el tipo de cambio nominal se deprecie no depende de si esta clase dura o no mas de 3 horas, y viceversa.

Notar que de la regla del producto, si A y B son eventos independientes, obtenemos:

$$P(A \cap B) = P(B)P(A|B) = P(B) \times P(A)$$

Independencia - Ejemplo

Supongamos que arrojo un dado de 6 caras pero no le informo el resultado.

¿Cuál es la probabilidad de que haya salido 6?

Suponga que le informo que el número que salió es par, ¿cuál es la probabilidad de que haya salido 6?

Considere los eventos $A =$'Salió 6' $B =$'Salió un número par', ¿son A y B independientes?

Teorema de Bayes

Tal vez uno de los resultados mas importantes en teoría de probabilidad, y con una interpretación de inferencia estadística clave. Este teorema relaciona la probabilidad $P(A|B)$ con $P(B|A)$.

Teorema de Bayes

Sea A un evento y sean los eventos $B_1, B_2, ..., B_N$ una 'partición' de Ω (*). Entonces,

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$

Teorema de Bayes

El resultado surge de la definición de probabilidad condicional, la regla de la multiplicación y la ley de la probabilidad total.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A) \times P(A)}{P(B)}$$

Usando probabilidad total abajo:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$

¿Por qué es tan relevante este choclo? Veamos un ejemplo...

Teorema de Bayes - Ejemplo

Este teorema es fundamental porque podemos interpretar a los eventos A y A^c como estados (enfermo, sano, habil, no habil, etc) y al evento B como un síntoma (fiebre, nota de un examen). Típicamente, si supieramos que una persona esta sana o enferma, la probabilidad de que tenga gripe sería conocida. Es decir, dado un estado, conocemos la probabilidad de los síntomas.

La pregunta interesante es, si medimos la temperatura de una persona y determinamos que tiene fiebre, ¿cuál es la probabilidad de que tenga gripe? de esta forma decidiríamos si recetarle o no un antibiótico... Es decir, dado el síntoma, ¿cuál es la probabilidad de cada estado?

Teorema de Bayes - Ejemplo

Supongamos que una enfermedad muy rara afecta sólo al 1% de la población. Hay un estudio médico relativo a dicha enfermedad que tiene las siguientes propiedades:

- si la persona está enferma, el test dará positivo con probabilidad 0.98; y
- si la persona está sana, el test dará positivo con probabilidad 0.05.

Pregunta: dado que el test dió positivo, ¿cuál es la probabilidad de que la persona esté sana? Veamos...

Teorema de Bayes - Ejemplo

Definamos los eventos Pos='el test dió positivo' (Pos^c='el test dió negativo'), Sana='la persona está sana' ($Sana^c$='la persona está enferma'). Sabemos que:

$$P(Sana) = 0.99$$
$$P(Pos|Sana) = 0.05$$
$$P(Pos|Sana^c) = 0.98$$

Teorema de Bayes - Ejemplo

Definamos los eventos Pos='el test dió positivo' (Pos^c='el test dió negativo'), Sana='la persona está sana' ($Sana^c$='la persona está enferma'). Sabemos que:

$$P(Sana) = 0.99$$
$$P(Pos|Sana) = 0.05$$
$$P(Pos|Sana^c) = 0.98$$

Y estamos buscando $P(Sana|Pos)$. Por Bayes:

$$P(Sana|Pos) = \frac{P(Pos|Sana)P(Sana)}{P(Pos|Sana)P(Sana) + P(Pos|Sana^c)P(Sana^c)}$$

Teorema de Bayes - Ejemplo

Finalmente, reemplazando con los datos numéricos:

$$\begin{aligned} P(Sana|Pos) &= \frac{0.05 \times 0.99}{0.05 \times 0.99 + 0.98 \times 0.01} \\ &= \frac{0.0495}{0.0495 + 0.0098} \\ &= 0.8347 \end{aligned}$$

Es decir, en este ejemplo, si el test dió positivo, lo mas probable es que la persona esté sana!

Ejercicio

Su analista le dio un informe sobre Tesla, donde asigna una probabilidad de 80% a que Tesla anuncie la fecha de lanzamiento del nuevo Model Y tras la próxima reunión de directorio. Si se hace el anuncio, la probabilidad de que la acción de Tesla se incremente más de un 10% es de 0.9 y, en caso contrario, de solo 0.5.

(a) Si Tesla se incrementó más de un 10%, ¿cuál es la probabilidad de que se haya anunciado el lanzamiento del Model Y?

(b) Si no sube más de un 10%, ¿cuál es la probabilidad de que Tesla no haya hecho el anuncio?

Problema de Monty Hall

Un típico caso para pensar probabilísticamente es en el conocido problema de Monty Hall. Un presentador le hace elegir entre 3 puertas: detrás de 2 de ellas hay cabras, y detrás de la otra hay un auto, todo posicionado al azar. Una vez que usted elige una puerta, el presentador le muestra que detrás de una que no eligió hay una cabra, y le ofrece cambiar de puerta.

¿Desearía usted cambiar de puerta? ¿Por qué? Ayuda: piense en la misma situación pero con 100 puertas, y una vez que eligió una, el presentador le muestra 98 donde hay cabras.

Lectura: "*Behind Monty Hall's Doors: Puzzle, Debate and Answer?*", de John Tierney, New York Times, July 21, 1991.

Más técnico:
https:mixtape.scunning.comch1.html#monty-hall-example

Un ejercicio para englobar todo lo que vimos hasta ahora

Considere el siguiente experimento aleatorio. Hay 2 dados: uno de 6 caras (con números del 1 al 6) y uno de 4 caras (con números del 1 al 4). Ambos dados están balanceados. Voy a realizar 2 tiradas. La primer tirada será con el dado de 6 caras. Luego, antes de realizar la segunda tirada, voy a arrojar una moneda balanceada al aire. Si la moneda cae cara, usaré nuevamente el dado de 6 caras, si cae cruz, usaré el dado de 4 caras.

- Hallar la probabilidad de que la suma de las tiradas de ambos dados sea 10.

Un ejercicio para englobar todo lo que vimos hasta ahora

- Suponga que la suma de los resultados de ambas tiradas dió 9. Considerando esta evidencia, compute la probabilidad de que en la segunda tirada haya usado el dado de 4 caras.
- Considere los eventos

$$A\text{: usé el dado de 4 caras en la segunda tirada}$$

$$B\text{: la suma de los resultados de ambas tiradas es 7}$$

Determine si A y B son independientes o no.

Distribuciones Bivariadas

La **función de probabilidad conjunta** es la probabilidad de que simultáneamente X tome un valor específico x e Y tome un valor específico y.

Función de probabilidad conjunta

Sean X e Y dos variables aleatorias.

(a) Si X e Y son discretas, Se llama función de probabilidad conjunta a la función

$$p(x, y) = P(X = x, Y = y)$$

que cumple

$$0 \leq p(x, y) \leq 1 \tag{1}$$

$$\sum_x \sum_y p(x, y) = 1 \tag{2}$$

Distribuciones Bivariadas

Función de densidad conjunta

Sean X e Y dos variables aleatorias.

(b) Si X e Y son continuas, se llama función de densidad conjunta a la función $f(x, y)$ que cumple

$$f(x, y) \geq 0 \tag{3}$$

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1 \tag{4}$$

Distribuciones Bivariadas

Dadas dos v.a. X e Y que tienen distribución conjunta, ¿podemos recuperar la probabilidad "por separado"?

Probabilidad marginal

Sean X e Y dos variables aleatorias discretas con probabilidad conjunta $p(x, y)$. Se defina la probabilidad marginal de X como

$$p(x) = P(X = x) = \sum_y p(x, y).$$

La probabilidad marginal de Y se define de forma análoga,

$$p(y) = P(Y = y) = \sum_x p(x, y).$$

Densidad marginal

Sean X e Y dos variables aleatorias continuas con densidad conjunta $f(x, y)$. Se defina la densidad marginal de X como

$$f(x) = \int_{-\infty}^{+\infty} f(x, y) dy.$$

La densidad marginal de Y se define de forma análoga,

$$f(y) = \int_{-\infty}^{+\infty} f(x, y) dx.$$

Distribuciones Bivariadas

Por ejemplo, si X e Y son el número de casa que las agencias inmobiliarias A y B venden por semana, respectivamente,

		X			
		0	**1**	**2**	**p(y)**
Y	**0**	0.12	0.42	0.06	**0.60**
	1	0.21	0.06	0.03	**0.30**
	2	0.07	0.02	0.01	**0.10**
	p(x)	**0.40**	**0.50**	**0.10**	**1.00**

Distribuciones Bivariadas

		X			
		0	**1**	**2**	**p(y)**
Y	**0**	0.12	0.42	0.06	**0.60**
	1	0.21	0.06	0.03	**0.30**
	2	0.07	0.02	0.01	**0.10**
	p(x)	**0.40**	**0.50**	**0.10**	**1.00**

En este cuadro de doble entrada todas las probabilidades que están en la parte central son probabilidades conjuntas. Por ejemplo, $P(X = 0, Y = 2) = 0.07$, es decir, la probabilidad de que la agencia A no venda nada y la B venda dos casas en una semana es del 7%. Mientras que la última columna y la última fila muestran las probabilidades marginales $P(X = x)$ y $P(Y = y)$.

Bibliografía sugerida

- Bertsekas et al. (2000) - Cap. 1
- Newbold et al. (2013) - Cap. 3