Libro 1: Introducción a la teoría de la probabilidad

Sitio:	Agencia de Habilidades para el Futuro
Curso:	Estadística y probabilidades para el desarrollo del soft 1º D
Libro:	Libro 1: Introducción a la teoría de la probabilidad
Imprimido por:	Eduardo Moreno
Día:	lunes, 17 de marzo de 2025, 16:37

Tabla de contenidos

¿De que se trata el material de lectura?

- Introducción a la probabilidad y ¿matemática para qué?
- ¿Eventos aleatorios o experimentos aleatorios?
    - Para pensar juntos: ¿Lanzaste alguna vez una moneda?
    - Definimos: Espacio muestral
- Definimos probabilidad
    - El axiomática de probabilidad
- Cálculo de la probabilidad
    - Definimos: Espacios muestrales infinitos y equiprobables
- Si de propiedades se trata....Entre la teoría de conjuntos y la probabilidad
- Para pensar juntos: Experimentos aleatorios
        - Que una moneda caiga dos veces de cara
        - Que una moneda se lance 3 veces y que caiga una vez de cara
        - Simpatizantes de un candidato en las elecciones
        - Estudiantes de TDS, ¿probablemente cursan 3 materias?
- Probabilidad condicional. ¿Cuál es la probabilidad de que salga una bolita roja?
    - Propiedades de la probabilidad condicional
        - Para pensar juntos: ¿Cuál es la probabilidad de que la suma de los dos dados dé 8?
    - Definimos: Regla del producto
    - Definimos: Partición del espacio muestral
- Teorema de la probabilidad total
- Teorema de Bayes
- Resumimos: Probabilidad condicional, total y el teorema de Bayes
    - Para pensar juntos: La probabilidad de que el individuo esté enfermo si el test dio positivo
- Independencia de eventos aleatorios
    - Propiedades de la independencia de eventos aleatorios
    - Para pensar juntos: Volvamos a lanzar monedas. ¿Son independientes los 3 eventos?

Orientaciones generales

En este material se introducen los elementos que hacen a la teoría de la probabilidad para, luego, comenzar a calcular las primeras probabilidades en escenarios, posiblemente, muy sencillos y que se irán complejizando en las próximas semanas.

Vas a encontrar varios ejemplos del libro que están explicados y desarrollados “a mano”. Estos ejemplos los vas a volver a encontrar en el material de estudio de la semana 2. Durante la segunda semana se hace foco en cómo implementar muchos de esos ejemplos en Python e, incluso, implementar las primeras simulaciones que nos permitan estimar probabilidades cuando no sea posible hacer el cómputo manualmente.

Pero antes de introducirnos en el mundo de las probabilidades, te invitamos a leer "Matemática, ¿para qué?"

El autor del artículo es Pablo Groisman. Dr. en Matemática de la Universidad de Buenos Aires, quien se dedica a la investigación en teoría de probabilidades. Pablo es autor de la reciente obra *Te regalo un Teorema* en la que, notablemente, aparecen varios resultados que estudiaremos en esta materia (¡y otros que no, pero son igual de interesantes!).

Comenzá a leer haciendo clic en el título del artículo: **"Matemática, ¿para qué?"** El autor al inicio menciona que es “Más importante que aprender a sumar fracciones es el hecho de que estar en contacto con la matemática ayuda a distinguir un argumento correcto de uno falaz. Sin esta habilidad, una sociedad se transforma en una manada fácilmente manipulable”.

¿Qué es un evento aleatorio?

¡Comencemos con una definición!

Definición Un experimento aleatorio es aquel cuyos posibles resultados son conocidos, pero no se puede establecer de antemano cuál de los posibles valores toma al ser ejecutado.

Para pensar juntos: Lanzamiento de una moneda o de un dado

- Al lanzar una moneda, sabemos que puede salir "cara" o "ceca" pero no sabemos si caerá de cara o de ceca hasta que no la lanzamos, se realiza el experimento y vemos de qué lado efectivamente cae.
- Lo mismo ocurre al lanzar un dado: sabemos que puede salir un número entero entre el 1 y 6, pero, hasta no lanzarlo, no sabemos cuál saldrá.

Otros ejemplos de experimentos aleatorios pueden ser:

- la extracción de una carta de un mazo,
- la selección de una persona de un conjunto de personas en un lugar,
- las condiciones climáticas en un día y lugar determinados,
- la duración de una lámpara o de una pieza electrónica en un sistema, entre otros.

Definición **¿Cómo se llama al conjunto de posible resultados de un experimento aleatorio?**

Se denomina **espacio muestral** al conjunto de los posibles resultados de un experimento aleatorio. Por lo general, al espacio muestral se lo denota con la letra griega omega Ω o S.

Siguiendo con los ejemplos: ¿Cuál sería el espacio muestral?

- Si se trata del lanzamiento de una moneda, el espacio muestral es

$$\Omega = \{\text{cara, ceca}\}.$$

- En el caso del lanzamiento de un dado, el espacio muestral es

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

Definición **¿Cómo se llama al número de elementos del espacio muestral?**

El cardinal de un conjunto, ($\#$), es el número de elementos que posee dicho conjunto.

- En el ejemplo del lanzamiento de un dado, el cardinal del espacio muestral es

$$\#\Omega = 6.$$

Definición **Entonces, un evento aleatorio, ¿qué es?**

Es cualquier **subconjunto** del espacio muestral.

¿Cuál podría ser un evento aleatorio en el experimento del lanzamiento de una moneda o de un dado?

En el lanzamiento de una moneda, un ejemplo de evento podría ser que la moneda caiga de cara,

$$E : \text{la moneda cae de} \quad \text{cara},$$

que se trata de un subconjunto de Ω y tiene un solo elemento:

$$E = \{\text{cara}\},$$

Por lo que su cardinal es:

$$\#E = 1.$$

En el lanzamiento de un dado podríamos tomar como evento que salga un número impar:

$$E = \{1, 3, 5\},$$

Y su cardinal es:

$$\#E = 3.$$

¿Cómo cuantificar la incertidumbre?

La probabilidad surge, históricamente, como una forma de cuantificar la incertidumbre. Por ejemplo, la probabilidad de que ocurra el evento E nos da una medida para la certidumbre de que ese evento ocurra.

Sigamos pensando juntos con el ejemplo de lanzar una moneda

-¿Cuál es la probabilidad de E en el ejemplo del lanzamiento de una moneda?

Aún no sabemos calcularla ni dimos una definición formal de ello; sin embargo, con seguridad ya estarás pensando en que es "un medio", o 0.5 o 50% . Esto es así porque, de hecho, la definición formal de probabilidad veremos que está en sintonía con algo que sí sabemos hacer (¡y hemos hecho muchas veces!): ¡estimarla empíricamente!

¿Qué es eso? Algo tan simple como lo que sigue:

- repetir muchas veces el experimento de lanzar una moneda y
- ver cuántas veces ocurre el evento de interés, por ejemplo, que salga cara.

Intuitivamente, la probabilidad de que ocurra ese evento parece ser la cantidad de veces que efectivamente vimos que ocurrió sobre el total de veces que repetimos ese experimento en igualdad de condiciones. Y si hacemos eso 1000 veces, por ejemplo, quizás no obtengamos exactamente 500 caras, pero de seguro sea un número que ronde esa cantidad.

En conclusión: La probabilidad es, como todo en matemática, algo abstracto.

Esa abstracción es necesaria para que los objetos definidos no sean ambiguos, recuperen mucho de lo que la intuición nos dice y, a la vez, sean lo suficientemente robustos como para que puedan deducirse resultados a partir de estos.

Veamos, a continuación, la definición axiomática de la probabilidad.

Definición axiomática de probabilidad

Dado un experimento aleatorio y un espacio muestral asociado Ω, a cada evento A se le asociará un número $P(A)$ que satisface estos tres axiomas:

1. $P(A) \geq 0$ para todo evento $A \subset \Omega$.
2. $P(\Omega) = 1$.
3. Si $A_1, A_2, \ldots, A_n, \ldots$ es una colección de eventos mutuamente excluyentes (es decir, disjuntos de a pares: $A_i \cap A_j = \emptyset, \forall i \neq j$) entonces:

$$P\left(\bigcup_{i=1}^{n \text{ ó } \infty} A_i\right) = \sum_{i=1}^{n \text{ ó } \infty} P\left(A_i\right).$$

Esta definición, aunque abstracta, recupera algo de la intuición detrás de lo que queremos modelar:

- la probabilidad es un número siempre positivo (o a lo sumo cero),
- la probabilidad de todo el espacio muestral es 1 (y claro, ¡no hay más posibles resultados que los que tiene el propio espacio muestral!) y
- la probabilidad de eventos disjuntos es igual a la suma de las probabilidades: la probabilidad de que al lanzar un dado salga 1 o 3 es igual a la suma de las probabilidades de que salga 1 o de que salga 3.

Sigamos con el lanzamiento del dado...

Consideremos el lanzamiento de un dado y el evento que salga un número impar

$$E = \{1, 3, 5\}, \quad \#E = 3.$$

Si acordamos que la probabilidad de obtener cualquier número (del 1 al 6) en el dado es la misma y es igual a $1/6$, entonces podemos calcular la probabilidad del evento E usando el tercer axioma de la definición de probabilidad:

$$P(E) = P(\{1, 3, 5\}) = P(\{1\} \cup \{3\} \cup \{5\}) = P(\{1\}) + P(\{3\}) + P(\{5\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2}.$$

Esto que acabamos de hacer es siempre posible bajo dos condiciones del espacio muestral:

- que sea finito y
- que sea equiprobable.

Esto es, precisamente, lo que ocurrió en el caso del dado; pero también ocurre en el caso de una moneda, o cuando se elige a una persona al azar de un conjunto de n personas, ¡y en muchos otros casos más!

En estos casos, entonces, la probabilidad se calcula a través de lo que se conoce como regla o fórmula de Laplace.

Regla de Laplace en espacios muestrales equiprobables

Diremos que un espacio muestral finito de n elementos $\Omega=\left\{e_1, \ldots, e_n\right\}$, es equiprobable si los n eventos elementales e_i tienen igual probabilidad, es decir

$$P\left(\left\{e_i\right\}\right)=\frac{1}{n}.$$

Por ejemplo, consideremos $\Omega=\left\{e_1, \ldots, e_8\right\}$ un espacio equiprobable y $A=\left\{e_3, e_8, e_1\right\} \subset \Omega$. Recordemos que en conjuntos, el orden en que se escriben los elementos no importa. Luego, $$P(A)=P\left(\left\{e_3\right\}\right)+P\left(\left\{e_8\right\}\right)+P\left(\left\{e_1\right\}\right)=\frac{3}{8}=\frac{\# A}{\# \Omega}.$$

Si Ω es un espacio equiprobable y $A \subset \Omega$ un evento cualquiera y $\#$ la notación para el cardinal de un conjunto, entonces $$P(A)=\frac{\# A}{\# \Omega}=\frac{\text { casos favorables }}{\text { casos totales }}.$$

Propiedades de la teoría de conjuntos para las propiedades de la probabilidad

De la definición de probabilidad, de la caracterización del espacio muestral y los eventos y de la fórmula de cálculo para probabilidades en espacios finitos y equiprobables, seguramente ya hayas notado que hay una estrecha relación con la teoría de conjuntos que estudiaste en EAM y en el curso de preingreso.

Recordemos algunas cosas antes de dar propiedades sobre la probabilidad.

- Ω : es el suceso cierto o seguro.
- $\emptyset$: es el suceso imposible.
- $A \cap B=\{e \in \Omega: e \in A \text{ y } e \in B\}$: ocurre A y ocurre B.
- $A \cup B=\{e \in \Omega: e \in A \text{ o } e \in B\}$: ocurre A u ocurre B.
- $A^c=\{e \in \Omega: e \notin A\}$: no ocurre A.
- $A-B=\{e \in \Omega: e \in A \text{ y } e \notin B\}$: ocurre A y no ocurre B.
- $A \subset B$: si ocurre A , entonces también ocurre B. Es decir, todo elemento de A es un elemento de B.
- Dos sucesos A y B se dicen disjuntos si $A \cap B=\emptyset$.

Tomá nota de lo siguiente, ya que que algunas de estas propiedades las usaremos mucho en lo sucesivo.

- $\left(\Omega^c\right)^c=\Omega$
- $A \cap \emptyset=\emptyset$
- $A \cup \emptyset=A$
- Si $A \subset B$, entonces $A \cap B=A$ y $A \cup B=B$.
- $A-B=A \cap B^c$
- $A=(A \cap B) \cup\left(A \cap B^c\right)$

Operaciones y propiedades con conjuntos

Por último, recordemos operaciones y propiedades con conjuntos (se generalizan a más eventos, pero por simplicidad escribámoslas para dos).

Leyes conmutativas

- Para la unión: $A \cup B=B \cup A$
- Para la intersección: $A \cap B=B \cap A$

Leyes asociativas

- Para la unión: $(A \cup B) \cup C=A \cup(B \cup C)$
- Para la intersección: $(A \cap B) \cap C=A \cap(B \cap C)$

Leyes distributivas

- $(A \cup B) \cap C=(A \cap C) \cup(B \cap C)$
- $(A \cap B) \cup C=(A \cup C) \cap(B \cup C)$

Leyes de Morgan

- $(A \cup B)^c=A^c \cap B^c$

- $(A \cap B)^c=A^c \cup B^c$

Propiedades de la probabilidad

Las que siguen son algunas propiedades que cumple la probabilidad y que se deducen de los axiomas de su definición y de las propiedades que conocemos para la teoría de conjuntos:

- $P\left(A^c\right)=1-P(A)$.
- $P(\emptyset)=0$.
- Si $A \subseteq B$ entonces $P(B-A)=P(B)-P(A)$ y $P(A) \leq P(B)$.
- $0 \leq P(A) \leq 1$

Cualquiera sean A y B:

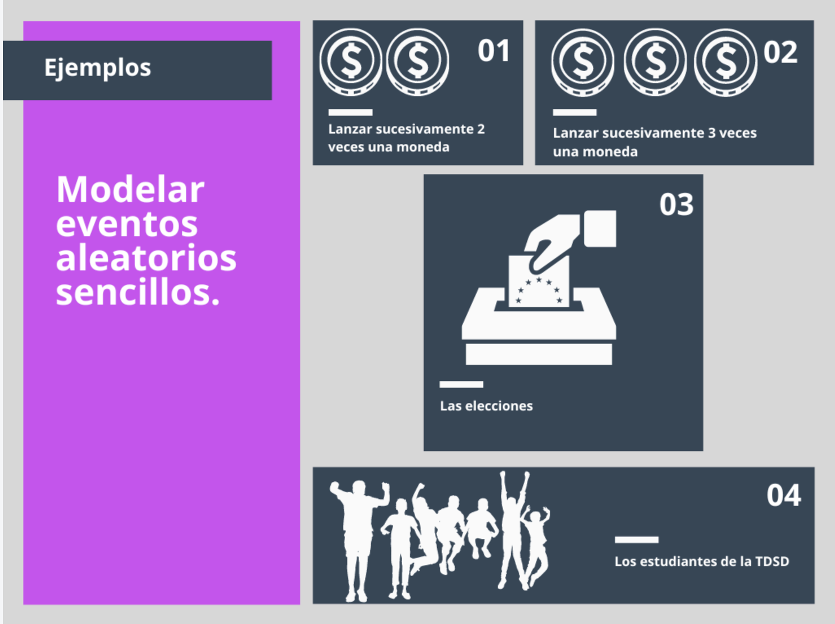
- $P(B-A)=P(B)-P(A \cap B)$
- $P(A \cup B)=P(A)+P(B)-P(A \cap B)$
- $P(A \cup B) \leq P(A)+P(B)$ (esto es directo, sale de la anterior)

Y, más generalmente, dados 3 sucesos A_1, A_2, A_3:

$$\begin{aligned} P\left(A_1 \cup A_2 \cup A_3\right) & =P\left(A_1\right)+P\left(A_2\right)+P\left(A_3\right)- \\ & -P\left(A_1 \cap A_2\right)-P\left(A_1 \cap A_3\right)-P\left(A_2 \cap A_3\right)+ \\ & +P\left(A_1 \cap A_2 \cap A_3\right) . \end{aligned}$$

Ejemplos para pensar juntos

En el siguiente apartado te presentamos algunos ejemplos para empezar a estudiar y modelar algunos eventos aleatorios sencillos.



Ejemplo 1: Lanzar sucesivamente dos veces una moneda

El **experimento aleatorio** consiste en lanzar sucesivamente dos veces una moneda. ¿Qué hay que hacer?

- **Comprender** el experimento aleatorio.
- **Dar un espacio muestral equiprobable** para este experimento.
- **Calcular la probabilidad** de algún evento de interés, por ejemplo, que la moneda caiga las dos veces de cara.

Ejemplo comentado

Como la moneda se lanza dos veces, los registros de la salida del experimento son duplas, por ejemplo, “cara-cara”.

Por lo tanto el espacio muestral es: $$\Omega=\{\text {cara-cara, cara-ceca, ceca-cara, ceca-ceca}\}.$$
Dicho espacio puede codificarse de diferentes formas, por ejemplo: $$\Omega=\{(C,C),(C,X),(X,C),(X,X)\}$$ O bien:

$$\Omega=\{(x,y) \text{ tales que } x,y\in\{C,X\}\}.$$
Cualquiera de estas elecciones vale, sin embargo, otras elecciones de codificación como esta $$\Omega=\{\text{las dos caras, las dos cecas, una cara y una ceca}\}$$

no conducirán a espacios equiprobables y, por lo tanto, no nos serán de utilidad en esta instancia ya que no sabríamos cómo calcular probabilidades a partir de ellos.

Luego, el evento $E:$ “salen dos caras” está dado por: $$E=\{\text {cara-cara}\},$$ Y como Ω es un espacio finito y equiprobable, la probabilidad de E está dada por: $$P(E)=\frac{\#E}{\#\Omega}=\frac{1}{4}=0.25.$$

En conclusión: Con este ejemplo vemos que no cualquier espacio muestral que propongamos al estudiar un experimento aleatorio será un espacio equiprobable, por lo tanto, tendremos que hacer un esfuerzo por hallar uno que sí lo sea. Una vez obtenido un espacio muestral equiprobable, podremos calcular probabilidades de eventos que nos interesen.

Ejemplo 2: Lanzar sucesivamente 3 veces una moneda

Consideremos el **experimento aleatorio** de lanzar sucesivamente tres veces una moneda. Para este experimento las consignas son:

- **Registrar la salida** en las tres ocasiones.
- Dar **un espacio muestral** para este experimento.
- **Calcular la probabilidad** de que salga al menos una cara.

Ejemplo comentado

Consideremos una codificación del espacio como la que sigue $$ \Omega=\{(x,y,z) \text{ tales que } x,y,z\in\{C,X\}\}. $$

El cardinal de Ω es 8, pues cada posición de la terna puede tomar dos posibles valores.

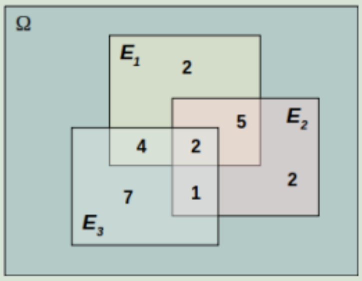
Luego, consideremos el evento $E:$ sale al menos una cara. Calcular su probabilidad implicaría contar todos los elementos de Ω en los que hay, por lo menos, una C en la terna. Esa tarea de conteo puede ser pesada. Sin embargo, podemos considerar el evento $E^c:$ no sale ninguna cara, ya que es fácil saber cuántos elementos tiene este evento: ¡solo uno: (X,X,X) ! Luego, como: $$ P(E^c)=\frac{\#E^c}{\#\Omega}=\frac{1}{8}, $$ y como vale que: $$ P(E^c)=1-P(E), $$ es fácil mostrar que la probabilidad buscada es: $$ P(E)=1-\frac{1}{8}=\frac{7}{8}. $$

En conclusión: En este ejemplo vimos cómo extender aquello que habíamos desarrollado en el ejemplo anterior (lanzar una moneda dos veces). Más aún, con muy poco esfuerzo podríamos extenderlo a más lanzamientos.

Ejemplo 3: Las elecciones

En este ejemplo vamos a calcular probabilidades que pongan en práctica algunas de las propiedades que vimos que se heredan de la teoría de conjuntos.

Un conjunto de personas simpatizan con los candidatos E_1, E_2 y E_3 para las próximas elecciones. A partir del siguiente diagrama de Venn queremos determinar:



- el número total de personas de este universo,
- el número de simpatizantes de cada candidato,
- el número de personas que siguen a los candidatos E_1 y E_2,
- el número de personas que siguen a los candidatos E_2 y E_3,
- el número de personas que siguen a los candidatos E_1 y E_2 pero no a E_3 y
- la probabilidad de que una persona elegida al azar de este conjunto simpatice con los candidatos E_1 y E_2 pero no con E_3 .

Ejemplo comentado

El número total de personas en el conjunto es la suma de las personas que muestra el diagrama para cada uno de los subconjuntos. Es decir, si denotamos por Ω al conjunto formado por todas las personas de este experimento, entonces: $$\#\Omega=2+4+2+5+7+1+2=23.$$ Del diagrama también podemos extraer más información. Por ejemplo, el número de personas que simpatizan con el candidato E_1 es: $$\# E_1=2+4+2+5=13,$$ El número de personas que simpatizan con el candidato E_2 es: $$\# E_2=2+5+1+2=10,$$ Y el número de personas que simpatizan con el candidato E_3 es: $$\#E_3=4+2+7+1=14.$$ También es posible hallar el cardinal de las inetrsecciones. Por ejemplo, el número de personas que simpatizan con los candidatos E_1 y E_2 es: $$\#\left(E_1 \cap E_2\right)=5+2=7,$$ Y para los candidatos E_2 y E_3 es: $$\#\left(E_2 \cap E_3\right)=2+1=3.$$ De esta forma, y usando las propiedades que conocemos para las operaciones con conjuntos, es fácil ver que las personas que simpatizan con los candidatos E_1 y E_2 pero no con E_3 son: $$\#\left(\left(E_1 \cap E_2\right)-E_3\right)=7-2=5.$$ Así, la probabilidad del evento $A:$ "que una persona elegida al azar de este conjunto simpatice con los candidatos E_1 y E_2 pero no con E_3" está dada por: $$P(A)=\frac{\#A}{\#\Omega}=\frac{\#\left(\left(E_1 \cap E_2\right)-E_3\right)}{\#\Omega}=\frac{5}{23},$$ pues Ω es un espacio muestral finito y equiprobable, formado por cada una de las 23 personas de este conjunto, es decir $$\Omega=\{s_i, 1\leq i \leq 23\}$$ siendo s_i el $i-$ésimo sujeto del experimento.

En conclusión: En este ejemplo vimos que un diagrama de Venn puede ser un recurso útil para calcular probabilidades cuando se tienen dos o tres eventos de interés, porque nos permite representar gráficamente los conjuntos y sus intersecciones. Eso ayuda a obtener los cardinales sin necesidad de tanta escritura simbólica. Cuando se tienen cuatro o más eventos, este diagrama, en general, deja de ser útil.

Ejemplo 4: Los estudiantes de la TDSD

En este último ejemplo vamos a resolver un problema similar al ejemplo anterior, pero el diagrama de Venn no va a estar dado desde el inicio. Como vimos en el ejemplo 3, este diagrama puede ser un recurso muy útil si es que no tenemos muchos eventos que representar.

- Imaginemos un contexto en el que 17 estudiantes cursan la Tecnicatura en Desarrollo de Software.
- De ellos se sabe que:
  - 8 de ellos/as cursan Probabilidad y Estadística,
  - 3 cursan: Probabilidad y Estadística y Elementos del Análisis Matemático,
  - 12 cursan: Elementos del Análisis Matemático y Modelado de Software,
  - 6 cursan: Probabilidad y Estadística y Modelado de Software; y
  - 3 cursan: las tres materias (Probabilidad y Estadística, Elementos de Análisis Matemático y Modelado de Software)
- Si se elige un/a estudiante al azar:
  - ¿Cuál es la probabilidad de que curse cada una de esas materias?
  - ¿Cuál es la probabilidad de que curse Elementos del Análisis Matemático y Probabilidad y Estadística?
  - ¿Cuál es la probabilidad de que curse alguna de las tres materias?

Ejemplo comentado

Para ordenarnos, llamemos:

- E_1 al conjunto de los estudiantes que cursan Elementos del Análisis Matemático,
- E_2 al conjunto de los estudiantes que cursan Probabilidad y Estadística, y
- E_3 al conjunto de los estudiantes que cursan Modelado de Software.

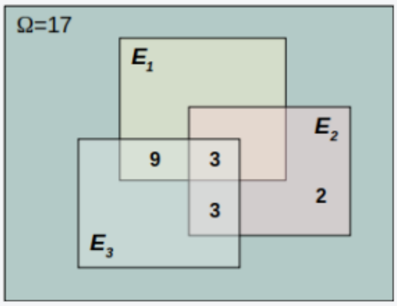
De los datos del enunciado sabemos que:

- $\#(E_2)=8$
- $\#(E_1\cap E_2)=3$
- $\#(E_1\cap E_3)=12$
- $\#(E_2\cap E_3)=6$
- $\#(E_1\cap E_2\cap E_3)=3$

Con esta información puede obtenerse que:

- $\#\left(\left(E_1 \cap E_3\right)-\left(E_1 \cap E_2 \cap E_3\right)\right)=9=$, que es el cardinal de $\left(E_1 \cap E_3 \cap E_2^c\right)$.
- $\#\left(\left(E_2 \cap E_3\right)-\left(E_1 \cap E_2 \cap E_3\right)\right)=3$, que es el cardinal de $\left(E_2 \cap E_3 \cap E_1^c\right)$.
- $\#\left(\left(E_1 \cap E_2\right)-\left(E_1 \cap E_2 \cap E_3\right)\right)=0$, que es el cardinal de $\left(E_1 \cap E_2 \cap E_3^c\right)$.

Podemos, también, traducir gráficamente esta información en un diagrama de Venn:



Llamemos Ω al espacio formado por cada uno de los 17 estudiantes de la Tecnicatura $$\Omega=\{s_i, 1\leq i \leq 17\}$$ con s_i el $i-$ésimo estudiante. Si consideramos el experimento aleatorio de elegir un estudiante al azar, este Ω es un espacio muestral finito y equiprobable, luego:

- $P(E_1)=\frac{\#E_1}{\#\Omega}=\frac{12}{17}$,
- $P(E_2)=\frac{\#E_2}{\#\Omega}=\frac{8}{17}$,
- $P(E_3)=\frac{\#E_3}{\#\Omega}=\frac{12}{17}$,
- $P(E_1\cap E_2)=\frac{\#(E_1\cap E_2)}{\#\Omega}=\frac{3}{17}$.

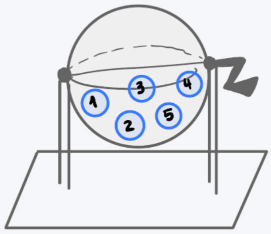
Por último, como $\Omega=E_1\cup E_2 \cup E_3$, la probabilidad de que curse alguna de las tres materias es: $$P(E_1\cup E_2\cup E_3)=\frac{\#(E_1\cup E_2\cup E_3)}{\#\Omega}=\frac{\#\Omega}{\#\Omega}=1.$$

En conclusión: En este ejemplo vimos que es importante definir con claridad los eventos aleatorios de interés, representarlos cuando sea posible e identificar, tanto en la escritura simbólica como en el gráfico, la información que tenemos, es decir, los datos del problema. Como siempre, no hay que olvidar definir adecuadamente un espacio muestral que resulte equiprobable, pues esa es la única manera que conocemos hasta el momento para calcular probabilidades.

Probabilidad condicional: Bolillas rojas

Para comprender la definición de **probabilidad condicional,** vamos a jugar primero con un experimento.

Consideremos un bolillero que contiene 4 bolillas rojas y 5 azules.



Ejemplo comentado

Extraemos una bolita y miramos su color.

-¿Cómo podríamos describir el espacio muestral?

Por lo que estuvimos estudiando: $$\Omega=\left\{r_1, r_2, r_3, r_4, a_1, a_2, a_3, a_4, a_5\right\}.$$

-¿Cuál es la probabilidad de que salga una bolita roja?

Llamemos R al evento ``la bolilla extraída es roja'', o bien: $R=\left\{r_1, r_2, r_3, r_4\right\}$. Luego,

$$P(R)=\frac{\# R}{\# \Omega}=\frac{4}{9} .$$

Ahora, supongamos que extraemos dos, una tras otra (sin reposición), y observamos su color.

-¿Cómo podríamos describir el espacio muestral?

Armamos pares ordenados con los posibles resultados de cada extracción, sin embargo, tenemos que considerar que, como el experimento es **sin reposición**, no puede haber pares ordenados con el mismo elemento en ambas coordenadas. Es decir,

$$\Omega=\left\{\left(x,y\right): x,y \in\left\{r_1, r_2, r_3, r_4, a_1, a_2, a_3, a_4, a_5\right\} \text{ con } x \neq y\right\} .$$

-¿Cuál es la probabilidad de que la primer bolita extraída sea roja?

Llamemos R_1 al evento ``la primera bolita es roja'', o bien:

$$R_1=\left\{\left(x,y\right) \in \Omega: x \in\left\{r_1, r_2, r_3, r_4\right\}\right\}$$ Luego,

$$P\left(R_1\right)=\frac{\# R_1}{\# \Omega}=\frac{4\times8}{9\times8}=\frac{4}{9}$$

¡Da lo mismo que antes! De alguna forma, cuando sacamos la primera bolita, no está importando qué pasa con la siguiente.

Probabilidad condicional: Bolillas azules

Ahora supongamos que nos pasan información sobre la primera extracción. Intuitivamente… ¿Cuál dirías que es la probabilidad de extraer una bolita azul, **sabiendo que** la primera fue roja? Animate a parar, pensarlo un rato, anotar y escribir tu respuesta.

Si ya respondiste lo anterior, sigamos.

Ejemplo comentado

A esta probabilidad la notaremos como: $$P\left(A_2 \mid R_1\right).$$ Es decir, usaremos la barra $\mid$ para indicar la información extra que tenemos, en este caso, que ocurrió R_1.

Al calcular $P\left(A_2 \mid R_1\right)$ nos están pidiendo **restringir el espacio muestral** a aquellas extracciones donde salió roja en la primera bolita. Ya no nos interesa todo el espacio muestral Ω, pues ya sabemos lo que ocurrió en la primera extracción, es decir, sabemos que ocurrió R_1. Podemos pensar: $$P\left(A_2 \mid R_1\right)=\frac{\# \{R_1 \cap A_2\}}{\# R_1}=\frac{4\times5}{4\times8}=\frac{20}{32}=\frac{5}{8} .$$

Como el espacio es equiprobable, podemos escribir: $$P\left(A_2 \mid R_1\right)=\frac{\frac{\# \{R_1 \cap A_2\}}{\# \Omega}}{\frac{\# R_1}{\# \Omega}}=\frac{P\left(R_1 \cap A_2\right)}{P\left(R_1\right)} .$$

¿Coincidió con lo que habías intuído? Comentalo en el foro o en la clase de consultas. En cualquier caso, este ejemplo sencillo motiva la definición de **probabilidad condicional**.

> Más precisamente, sean A y B eventos tales que $P(B)>0$, la probabilidad del evento A condicional a la ocurrencia del evento B es: $$P(A \mid B)=\frac{P(A \cap B)}{P(B)} .$$ Si el espacio es equiprobable, entonces, lo anterior es equivalente a: $$P(A \mid B)=\frac{\#(A \cap B)}{\# B} .$$

En resumen

Antes de continuar con otros ejemplos, resumamos algunas propiedades de la probabilidad condicional.

> Dado cualquier evento B tal que $P(B)>0, P(\, \circ \mid B)$ satisface los axiomas y propiedades de cualquier probabilidad.

Es decir, la probabilidad condicional ¡es una probabilidad! Por ejemplo:

- $P(\Omega \mid B)=1$,
- $P\left(A^c \mid B\right)=1-P(A \mid B)$,
- $P(A \cup C \mid B)=P(A \mid B)+P(C \mid B)-P(A \cap C \mid B)$.

Tené en cuenta que...

En general, $P\left(A \mid B^c\right) \neq 1-P(A \mid B)$.

Ejemplo 5: Arrojando dos dados equilibrados

Como introducción al tema de probabilidad condicional e independencia, vamos a resolver otro problema con dados, solo que ahora no vamos a trabajar con los resultados de cada dado, sino con la suma de los resultados que pueden obtenerse al lanzar dos dados equilibrados.

> Se arrojan dos dados equilibrados. Sabiendo que el primer resultado es un 4, ¿cuál es la probabilidad de que la suma de los dos dados dé 8?

A continuación escribimos el resumen de una posible resolución, pero te invitamos a que:

- describas el espacio muestral del experimento, y
- realices todas las cuentas y planteos de rigor.

$$\begin{aligned} P(\text {la suma es } 8 \mid \text {el primer resultado es } 4)= &\frac{P(\text {la suma es 8 y el primer resultado es 4})}{P(\text {el primer resultado es } 4)} \\ & =\frac{P(\text {ambos resultados son 4})}{P(\text {el primer resultado es 4})} \\ & =\frac{1 / 36}{1 / 6}=\frac{1}{6} . \end{aligned}$$

Importante

Para representar al espacio muestral podés usar un cuadro de doble entrada e identificar allí los eventos de interés. Los cuadros de doble entrada son muy útiles para reconocer probabilidades conjuntas, es decir, probabilidad de la intersección de dos eventos.

Este ejemplo motiva algo que se conoce como “regla del producto”.

Regla del producto

Pongámosle nombre a la intuición que quizás emergió.

De la definición de probabilidad condicional, se deduce que:

> Dados dos sucesos A y B, tales que $P(B)>0$ $$P(A \cap B)=P(A \mid B) P(B) .$$

Retomemos el ejemplo del bolillero con 4 bolillas rojas y 5 azules, de la que extraemos dos bolillas sin reposición.

Ejemplo comentado

-¿Cuál es la probabilidad de extraer una bolita roja y una azul, en ese orden?

En términos de la notación usada, queremos calcular $P\left(R_1 \cap A_2\right)$ y usando la regla del producto y la probabilidad condicional, veamos que es posible reconstruir la probabilidad conjunta pedida. $$P\left(R_1 \cap A_2\right)=P\left(A_2 \mid R_1\right) P\left(R_1\right)=\frac{5}{8} \cdot \frac{4}{9}=\frac{20}{72}=\frac{5}{18} .$$

-¿Cómo podemos obtener la probabilidad de que la segunda bolita extraída sea azul?

Es decir, queremos calcular $P\left(A_2\right)$. Para ello, notemos que: $$\Omega=R_1 \cup R_1^c,$$ y que $R_1^c=A_1=$ ``sale azul en la primera extracción''.

Entonces: $$\begin{aligned} P\left(A_2\right)=P\left(A_2 \cap\left(R_1 \cup R_1^c\right)\right) & =P\left(\left(A_2 \cap R_1\right) \cup\left(A_2 \cap R_1^c\right)\right) \\ & =P\left(A_2 \cap R_1\right)+P\left(A_2 \cap R_1^c\right) \\ & =P\left(A_2 \mid R_1\right) P\left(R_1\right)+P\left(A_2 \mid R_1^c\right) P\left(R_1^c\right) \\ & =\frac{5}{18}+\frac{4}{8} \cdot \frac{5}{9}=\frac{5}{18}+\frac{5}{18}=\frac{5}{9} . \end{aligned}$$

La partición del espacio muestral

En el ejemplo del bolillero, para calcular $P(A_2)$ notamos que fue conveniente pensar que todo el espacio muestral Ω se dividía en dos posibles escenarios: o salía roja en la primera extracción (R_1) o salía azul (R_1^c).

Ejemplo comentado

En efecto, R_1 y R_1^c forman lo que se conoce como una **partición** del espacio muestral pues, entre ambas "cubren" todo el espacio y, a la vez, entre ambas NO hay elementos en común.

Más precisamente:

> Una colección de eventos $E_1, E_2, \ldots, E_k$ constituye una **partición** del espacio muestral Ω si:
>
> - $E_i \cap E_j=\emptyset$ para todo $i \neq j$,
> - $P\left(E_i\right)>0$ para todo i,
> - $\bigcup_{i=1}^k E_i=\Omega$.

Es fácil verificar que en el ejemplo de las bolitas, $E_1=R_1$ y $E_2=R_1^c$ forman una partición de Ω .

Pero, ¿para qué son útiles las particiones? Cuando calculamos $P(A_2)$ fue necesario pensar que el evento "sale azul en la segunda extracción" es un evento que puede ocurrir en "dos posibles mundos": pues, o sale roja en la primera extracción, o sale azul. Y en cualquiera de esos "mundos", una vez que sabemos que ocurrieron, es posible calcular probabilidades para la segunda extracción.

En conclusión: Esto que hicimos casi de forma intuitiva (aunque lo hicimos usando propiedades de la probabilidad y de lo que sabemos sobre conjuntos) se resume en el resultado que motiva la siguiente sección: **el teorema de la probabilidad total**.

Teorema de la probabilidad total

Sea $E_1, E_2, \ldots, E_k$ una partición del espacio muestral Ω y B un suceso cualquiera. Entonces:

$$P(B)=\sum_{i=1}^k P\left(B \cap E_i\right) =\sum_{i=1}^k P\left(B \mid E_i\right) P\left(E_i\right).$$

Este teorema se conoce como el teorema de la probabilidad total.

Definición

La probabilidad total permite expresar la probabilidad de un evento como la suma de las probabilidades conjuntas de ese evento con todos los elementos de una partición del espacio. Y estas probabilidades conjuntas, a su vez, pueden expresarse como condicionales, por la “regla del producto” que enunciamos anteriormente.

Ejemplo comentado

Sigamos con el ejemplo del bolillero con 4 bolillas rojas y 5 azules, de la que extraemos dos bolillas sin reposición.

Calculemos la probabilidad de que la primera haya sido roja, sabiendo que la segunda fue azul. Es decir: $$P\left(R_1 \mid A_2\right)=\frac{P\left(R_1 \cap A_2\right)}{P\left(A_2\right)}.$$ Si en el numerador usamos la regla del producto y en el denominador el Teorema de la probabilidad total queda: $$\frac{P\left(A_2 \mid R_1\right) P\left(R_1\right)}{P\left(A_2 \mid R_1\right) P\left(R_1\right)+P\left(A_2 \mid R_1^c\right) P\left(R_1^c\right)}=\frac{1}{2} .$$ Notar que R_1 y R_1^c forman una partición de Ω.

¿Qué aporta esto de nuevo?

A continuación te lo comentamos.

Del ejemplo de las elecciones y las bolitas

En los ejemplos anteriores, con las urnas y las bolitas, siempre es muy fácil obtener las probabilidades condicionales cuando condicionamos a extracciones previas: claro, es obvio, si me decís qué color salió en la extracción anterior, ya sé de qué color tengo una bolita menos y, por lo tanto, puedo calcular probabilidades sin problemas.

Es decir, en este tipo de ejemplos que son secuenciales siempre es directo conocer este tipo de probabilidades condicionales $P(R_{i+1}|R_i)$ que estas otras $P(R_{i}|R_{i+1})$.

Tanto es así, que hasta es anti-intuitivo que nos pidan la probabilidad de roja en la primera extracción sabiendo que salió roja en la segunda. Es tan fuerte la idea de que el experimento es secuencial, que nos parece zonzo que alguien se interese por una probabilidad como esa. Sin embargo, esa probabilidad está bien definida matemáticamente (¡aunque la intuición nos engañe!) y es posible calcularla.

Pero esto tiene aplicaciones mucho más poderosas que solamente ejemplos triviales como este, y los exploraremos luego de enunciar un teorema que resume lo que acabamos de notar: el teorema de Bayes.

Teorema de Bayes

Consideremos $E_1, E_2, \ldots, E_k$ una partición del espacio muestral Ω y B un suceso tal que $P(B)>0$. Entonces, el teorema de Bayes afirma que:

$$P\left(E_j \mid B\right)=\frac{P\left(B \mid E_j\right) P\left(E_j\right)}{\sum_{i=1}^k P\left(B \mid E_i\right) P\left(E_i\right)} .$$

En general

- $P\left(E_j\right)$ se llama probabilidad a priori,
- $P\left(E_j \mid B\right)$ se llama probabilidad a posteriori del evento E_j ,

y en lo sucesivo comentaremos el porqué de estos nombres.

Algunos comentarios

Este resultado, el t**eorema de Bayes**, resume lo que ya sabíamos acerca de:

- la probabilidad condicional,
- la regla del producto y
- la probabilidad total.

Es más, es posible demostrar ese teorema usando esos tres elementos. Entonces, ¿qué aporta de nuevo? Más precisamente, ¿en qué situaciones puede ser de interés calcular una probabilidad como la que se propone en el teorema?

En el ejemplo de la urna y las bolitas, ya dijimos que era anti-intuitivo que nos pidan calcular la probabilidad de roja en la primera extracción sabiendo que salió roja en la segunda. Se trata de un experimento fuertemente secuencial y parece zonzo que alguien se interese por una probabilidad como esa.
Sin embargo, y como vimos en el teorema de Bayes, esa probabilidad no solo está bien definida matemáticamente, sino que hay contextos en donde es **especialmente importante** calcular esas probabilidades condicionales.

El contexto por excelencia para ilustrar este resultado es de los tests para detectar enfermedades. Veamos un ejemplo.

Ejemplo 6: El test para saber si una persona está o no enferma

Seguimos trabajando con **probabilidad condicional**. Un contexto muy frecuente en el que aparece esto, así como el teorema de Bayes, es en la elaboración de pruebas de laboratorio para la detección de enfermedades. En un contexto ideal, uno desearía tener una prueba o test que sea capaz de dar positivo solo cuando la persona está enferma, y negativo, cuando no lo está. Sin embargo, esto es imposible desde las ciencias médicas: los tests arrojan falsos positivos y falso negativos. **Estos falsos positivos y negativos son, de hecho, probabilidades condicionales.** En este ejemplo vamos a ilustrar un posible escenario de aplicación de estas probabilidades condicionales.

Supongamos que una determinada enfermedad está presente en un 4% de la población. A esto se conoce como prevalencia de la enfermedad. Imaginemos que existe un test para detectar la presencia de dicha enfermedad y se sabe que la probabilidad de que el test dé positivo dado que se aplica a una persona enferma es 0.92; y que la probabilidad de que el test de positivo dado que se aplica a un persona sana es 0.06. Es decir, el test "acierta" dando positivo en el 92% de las veces cuando alguien está enfermo y "yerra" en el 6% de las veces dando positivo cuando una persona está sana.

Supongamos que se elige una persona al azar de la población y se le realiza el test. Hay una probabilidad que es la que realmente nos va a interesar: **la probabilidad de que el individuo esté enfermo si el test dio positivo**. ¿Por qué? Porque estamos usando el test para saber si la persona está o no enferma y, si el test es **bueno**, nos gustaría poder usarlo como herramienta para detectar rápidamente casos de personas potencialmente enfermas en una población. Basta recordar lo que ocurrió en épocas de pandemia por COVID−19: que un test diera positivo no era garantía de que la persona estuviera enferma (es decir, estuviera infectada), pero el testeo buscaba ser una herramienta rápida, económica y útil para poder tratar, aislar y hacer seguimiento de la población.

Ejemplo comentado

Definamos los eventos de interés.

- $T^+:$ el test aplicado a una persona elegida al azar da positivo; $T^-=(T^+)^c:$ el test … da negativo.
- $E:$ la persona elegida al azar está enferma; $S=(E)^c:$ la persona … está sana.

Por dato, sabemos que:

- $P(E)=0.04$,
- $P(T^+ | E)=0.92$,
- $P(T^+ | S)=0.06$.

Queremos hallar $P(E|T^+)$. Notemos, entonces, que conocemos UNA condicional, $P(T^+ | E)$, y queremos la otra, es decir $P(E|T^+)$. Esto debería hacernos intuir que Bayes es un resultado que podría asistirnos. En efecto, como E y $S=E^c$ forman una partición del espacio, vale que:

$$P(E|T^+)=\frac{P(T^+ | E) P(E)}{P(T^+ | E) P(E)+P(T^+ | S) P(S)}.$$ Como: $P(S)=P(E^c)=1-P(E)=1-0.04=0.96$, entonces:

$$P(E|T^+)=\frac{0.92 \cdot 0.04}{0.92 \cdot 0.04 + 0.06 \cdot 0.96}\approx 0.3898.$$ ¿Sorprendida/o? Si bien la probabilidad de que el test dé positivo en una persona enferma es alta (0.92), la probabilidad de que la persona **realmente esté enferma dado que el test le dio positivo** no parece serlo tanto (menos de 0.39). Esto ocurre, entre otras razones, porque la prevalencia de la enfermedad es baja (0.04). Basta con pensar qué hubiese ocurrido si la prevalencia hubiera sido del 50%, es decir, la mitad de la población tuviera la enfermedad realmente. ¡Calculá la probabilidad anterior con este nuevo dato y sorprendete con el resultado!

En conclusión: Este ejemplo sirve para mostrar una situación en la que, conociendo una probabilidad *a priori* (la prevalencia, es decir la probabilidad de estar enfermo en una cierta población), queremos conocer una probabilidad a posteriori (es decir, la probabilidad de estar enfermo después de haber sido

> testeado como positivo para un test de esa enfermedad). Y aunque el evento “enfermo” sea el mismo, ¡su probabilidad cambia si consideramos el experimento de haber testeado a las personas!

Avancemos a la última definición de esta primera parte: la independencia de eventos aleatorios.

Independencia de eventos aleatorios

Decimos que los eventos A y B son si: $$P(A \cap B)=P(A) P(B).$$
Dos consecuencias inmediatas de la independencia de eventos son: Si $P(B) \neq 0$, los eventos A y B son independientes si $P(A \mid B)=P(A)$. Si $P(A) \neq 0$, los eventos A y B son independientes si $P(B \mid A)=P(B)$. ¿Por qué?
Basta ver que si $P(A \cap B)=P(A) P(B)$ y $P(B) \neq 0$, entonces: $$P(A \mid B)=\frac{P(A \cap B)}{P(B)}=\frac{P(A) P(B)}{P(B)}=P(A).$$ Y, análogamente, si $P(A \mid B)=P(A)$, entonces: $$P(A \cap B)=P(A \mid B) P(B)=P(A) P(B).$$

La definición probabilística de independencia de eventos recupera la intuición que uno/a naturalmente tiene: dos eventos son independientes cuando la probabilidad de que uno ocurra no afecta a la probabilidad de que ocurra el otro.

La definición es un poco más general que esto debido a tecnicismos matemáticos, pero lo interesante es notar que si $P(B) \neq 0$, los eventos A y B son independientes si $P(A \mid B)=P(A)$, es decir, saber que ocurrió B no cambia en nada a la probabilidad de que ocurra A.

En el ejemplo de Bayes y la enfermedad, está claro que los eventos “estar enfermo” y “el test da positivo” no son independientes, pues vimos que la probabilidad de estar enfermo cambiaba si se tenía información adicional acerca del resultado del test. ¡Y es esperable! Queremos hacer tests para detectar enfermedades que NO sean independientes de las enfermedades, ¡todo lo contrario!

Propiedades de la independencia

Algunas propiedades que se desprenden de la independencia:

- Si los sucesos A y B son disjuntos y tanto $P(A)>0$ como $P(B)>0$, entonces A y B NO son independientes.
- Si $P(B)=0$, entonces B es independiente de cualquier suceso A.
- Si A y B son sucesos independientes, A y B^c también lo son.
- Si A y B son sucesos independientes, A^c y B^c también lo son.

- Los eventos A_1, A_2, A_3 son independientes si se satisfacen al mismo tiempo que:
    - $P\left(A_1 \cap A_2 \cap A_3\right)=P\left(A_1\right) \cdot P\left(A_2\right) \cdot P\left(A_3\right)$,
        - $P\left(A_1 \cap A_2\right)=P\left(A_1\right) \cdot P\left(A_2\right)$,
        - $P\left(A_1 \cap A_3\right)=P\left(A_1\right) \cdot P\left(A_3\right)$ y
        - $P\left(A_2 \cap A_3\right)=P\left(A_2\right) \cdot P\left(A_3\right)$.

Importante

Si los sucesos A_1, A_2, A_3 son independientes, entonces son independientes de a pares, pero *la recíproca no es cierta*. Es decir, por más que probemos que todos los pares de eventos son independientes, eso no garantiza que los tres lo sean.

Ejemplo 7: Las monedas

¡Ahora sí, nuestro último ejemplo! En lo que sigue, vamos a estudiar un contexto de independencia que, de seguro, te sorprenda. **La independencia de eventos aleatorios no siempre es un concepto que podamos entender de forma intuitiva** (a veces sí, pero no siempre ocurre). De hecho, en este ejemplo vamos a ver que la independencia de eventos de a pares, no implica la independencia entre todos ellos, algo que uno intuitivamente podría pensar que valdría.

Se tiran dos monedas y se observan los resultados. Consideremos el espacio muestral: $$\Omega=\{(C, C),(C, X),(X, C),(X, X)\}$$

Los eventos son:

- A : ``sale cara en la primera moneda'',
- B: ``las dos monedas muestran la misma cara'',
- C: ``sale cara en la segunda moneda''

Calcular:

- $P(A), P(B), P(C)$.
- $P(A \cap B), P(A \cap C), P(B \cap C)$.
- $P(A \cap B \cap C)$.

Nos preguntamos: ¿Son independientes los 3 eventos? ¿Son independientes de a pares?

Ejemplo comentado

Calculemos las probabilidades pedidas. Como Ω es finito y equiprobable, basta con conocer los cardinales de cada uno de los eventos de interés. Notemos que: $$A=\{(C, C),(C, X)\}, \quad B=\{(C, C),(X, X)\}, \quad C=\{(C, C),(X, C)\},$$ y que: $$A\cap B=\{(C, C)\}, \quad A\cap C=\{(C, C)\}, \quad B\cap C=\{(C, C)\}.$$ Entonces:

- $P(A)=\frac{2}{4}$
- $P(B)=\frac{2}{4}$
- $P(C)=\frac{2}{4}$
- $P(A\cap B)=\frac{1}{4}$
- $P(A\cap C)=\frac{1}{4}$
- $P(B\cap C)=\frac{1}{4}$

Para estudiar la independencia de a pares, basta ver que:

$$P(A\cap B)=P(A)P(B), \quad P(A\cap C)=P(A)P(C), \quad P(B\cap C)=P(B)P(C).$$

Luego, A es independiente de B, A es independiente de C y C es independiente de B. Entonces, uno estaría tentado de pensar que los tres eventos son independientes, sin embargo, basta ver que: $$P(A\cap B \cap C)=\frac{\#(A\cap B \cap C)}{\#\Omega}=\frac{1}{4}$$,

pues los eventos A,B,C solo tienen en común el elemento (C,C), mientras que: $$P(A)P(B)P(C)=\frac{2}{4}\cdot \frac{2}{4}\cdot \frac{2}{4}=\frac{1}{8}.$$

Como $P(A\cap B \cap C)\neq P(A)P(B)P(C)$, entonces, los eventos A,B y C NO son independientes.