El siguiente material se encuentra en etapa de corrección y no deberá ser considerado una versión final.
Alejandro D. Zylberberg <alejandro@probabilidad.com.ar>
Versión Actualizada al: 4 de mayo de 2004

Probabilidad

No es que hayamos estado evadiéndola, pero era necesario definir algunos conceptos y recordar ciertas cuestiones de la teoría de conjuntos antes poder responder la pregunta: **¿Qué es la probabilidad?**

• La probabilidad expresa el **grado de certeza** de que ocurrirá un determinado **suceso** al hacer un determinado **experimento aleatorio**.
• Cuanto más alta es la probabilidad de un suceso, mayor es el grado de certeza de que ocurrirá al hacer el experimento aleatorio.
• Dado un suceso A, escribimos su probabilidad como P(A).

Daremos a continuación cuatro definiciones de probabilidad:

Definición informal

Informalmente, la probabilidad de un suceso es un número real entre 0 y 1.
Dicho número se puede expresar por ejemplo como 0.2, aunque también se lo puede representar como fracción ($^1/_5$), o bien como porcentaje (20%).
Si la probabilidad es 0, se **sabe** que el suceso no ocurrirá.
Si la probabilidad es 1, se **sabe** que el suceso ocurrirá.
Es decir, el 0 y el 1 son los casos **límite**.
Para valores intermedios, el suceso puede o no ocurrir. En general diremos que una probabilidad cercana a 0 es baja, y que una probabilidad cercana a 1 es alta.
Si por ejemplo la probabilidad de que mañana llueva es 0.9 significa que mañana es **altamente probable** que llueva. Si en cambio la probabilidad de que un avión se caiga es 0.000000001 significa que viajar en avión es bastante seguro.

¿Cuándo es alta una probabilidad? ¿Cuándo es baja? Eso es subjetivo. Por ejemplo si al despertarnos a la mañana el pronosticador del tiempo dice que hay 90% de probabilidades de lluvia, seguramente consideraremos que es un número alto, o por lo menos lo suficientemente alto como para tomarnos la molestia de llevar un paraguas al salir. En cambio si la probabilidad de que un avión complete un viaje sin caerse fuera ese mismo 0.9, dudo mucho que alguien quiera viajar en ese avión. Entonces cuándo una probabilidad es o no alta o baja depende en gran medida del contexto. Es decir, **a qué esté asociada** esa probabilidad.

Ejemplos:

1) Si el suceso A consiste en obtener cara al tirar una moneda, entonces intuitivamente podemos decir que si la moneda no está cargada, entonces $P(A) = {}^{1}/_{2}$.

2) Si el suceso A consiste en obtener un 3 al tirar un dado honesto (no cargado) entonces intuitivamente podemos decir que $P(A) = {}^{1}/_{6}$.

3) Si el experimento consiste en tomar a la primera persona que veamos y preguntarle el día de la semana en que nació (supongamos que no la conocemos) entonces si el suceso A es que la persona haya nacido durante un fin de semana, diríamos intuitivamente que $P(A) = {}^{2}/_{7}$.

Esto nos lleva a la segunda definición que daremos de probabilidad:

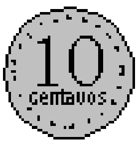
Definición de Laplace

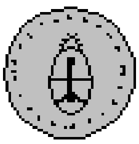
En los 3 ejemplos anteriores lo que hicimos intuitivamente fue contar la cantidad de casos posibles, y luego contar la cantidad de casos contenidos en el suceso A, y responder que P (A) era el cociente entre la cantidad de casos favorables a A y la cantidad de casos totales. Es decir:

$$P(A) = \frac{\text{cantidad de resultados contenidos en A}}{\text{cantidad total de resultados}}$$

Esto hace parecer que siempre que sepamos la cantidad de resultados posibles de un experimento y la cantidad de resultados englobados por el suceso A podemos calcular P (A). Sin embargo, esto es falso.

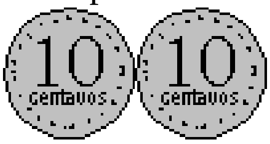
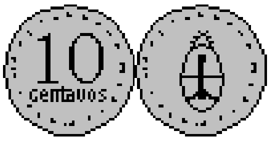
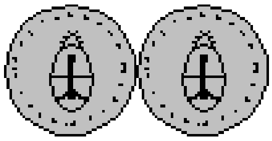
Volvamos al ejemplo de las monedas:
1) ¿Cuál es la probabilidad de sacar cara al tirar una moneda no cargada?
De acuerdo al razonamiendo intuitivo anterior, los resultados posibles son:

E = { 

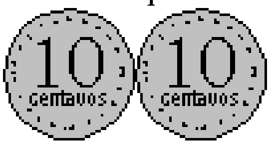
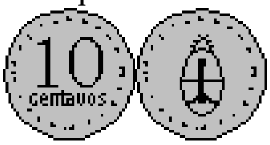
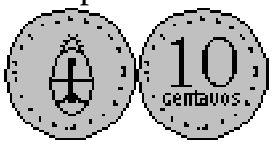
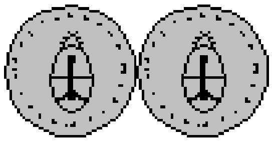
,  }

Luego, si el suceso A consiste en sacar cara, constituye 1 entre 2 resultados posibles, y en consecuencia P(A) = 1/2.

2) ¿Cuál es la probabilidad de sacar dos caras al tirar dos monedas iguales?
L resultados posibles son:

E = {  ,  ,  }

Entonces si A es "sacar dos caras", deberíamos decir que sacar dos caras es 1 entre 3 resultados posibles, y entonces P(A) = $^1/_3$. Pero ese resultado es incorrecto, ya que intuitivamente sabemos (o deberíamos saber) que el resultado correcto es $^1/_4$, y que el error se debió a que tendríamos que haber usado el espacio muestral:

E = {  ,  ,  ,  }

que tiene 4 resultados posibles en vez de 3. Luego diremos correctamente que P(A) = $^1/_4$.

Pero... ¿Cuál es la razón por la cual el espacio muestral que escribimos al final es apropiado y el anterior no? ¿Por qué la cantidad de resultados "correcta" es 4 y no 3, si según los que dijimos antes, ambas son formas perfectamente válidas de escribir el espacio muestral?

Y la respuesta es: porque los 4 resultados de la última expresión para E son equiprobables, mientras que los 3 de la expresión anterior no lo son.

¿Qué significa que los resultados de E sean equiprobables?
Que tienen todos la misma probabilidad.

¿Y cómo se sabe si los resultados que componen una determina expresión de E son equiprobables?
No se sabe. Se supone.

Lamentablemente, en los problemas reales no existe una forma idónea de determinar si una determinada expresión de E está compuesta por sucesos equiprobables.
En el ejemplo de las 2 monedas, *suponemos* intuitivamente que el 4 resultados que se obtienen al diferenciar las dos monedas son equiprobables y los 3 resultados que obtienen sin distinguiarlas no son equiprobables, porque el suceso "1 cara y 1 ceca" tiene 2 formas distintas de ocurrir, mientras que "2 caras" y "2 cecas" tienen solamente una forma de ocurrir cada una.

Es aceptable suponer equiprobabilidad cuando no se tiene absolutamente ningún conocimiento acerca de las probabilidades de los resultados, y eso incluye no solamente no conocer ninguna de las probabilidades sino también no tener razones que hagan pensar que algunos resultados pueden ser más probables que otros. Eso fue lo que hicimos en el ejemplo de preguntarle a la persona el día de la semana en que nació: como no conocemos a la persona, no tenemos forma de saber qué día de la semana nació, y tampoco conocemos nada que nos pueda dar una idea de cuáles días pueden ser más probables que otros. En cambio si la pregunta fuera sobre el año de nacimiento, ya no sería tan aceptable suponer equiprobabilidad, porque no todos los años posibles tienen la misma probabilidad: por ejemplo si la persona parece ser adulta, los años recientes tienen menos probabilidad de ser el año de nacimiento de la persona que los años no-tan-recientes.

Pero entonces, ¿Cómo se pueden calcular las probabilidades cuando no se puede suponer equiprobabilidad?
Hay dos formas: una consiste en aplicar alguno de los modelos que veremos a lo largo de esta obra. La otra, tiene que ver con la tercera definición:

Definición empírica

Esta definición consiste en asociar las probabilidades de los resultados con sus frecuencias relativas luego de repetir el experimento una determinada cantidad de veces. De ahí el nombre "empírica".
Es decir,

$$P(A) \approx fr_{rel}(A) = \frac{fr_{abs}(A)}{n}$$

donde $fr_{abs}(A)$ es la cantidad de veces que ocurrió A en las n veces que se llevó a cabo el experimento.
Cuanto más grande sea n, mejor será la aproximación de P(A) por $fr_{rel}(A)$.

Ejemplo:
Si se quiere tener una idea de cuál es la probabilidad de que eligiendo un alumno de la facultad al azar, éste tenga ojos claros, se puede tomar a 50 alumnos al azar y contar cuántos tienen ojos celestes. Luego si 13 de esos 50 tienen ojos claros, estimaremos que P (A) = $^{13}/_{50}$ = 0.26.
Si en vez de examinar a 50 alumnos hubiéramos examinado a 200, la exactitud esperable sería mayor. Por ejemplo quizás entre los 200 alumnos habría 53 con ojos claros, y entonces P(A) = 0.265.
Y si hubiera infinitos alumnos, y tomáramos muestras cada vez mayores, nos acercaríamos asintóticamente al resultado real, que podría ser, por ejemplo, 0.263.

Definición axiomática

Las tres definiciones que dimos hasta ahora cumplen con esta cuarta y última definición. La definición axiomática consta de los siguientes tres axiomas:

- **Axioma 1: $P(A) \geq 0$**

"La probabilidad no puede ser negativa"

- **Axioma 2: $P(E) = 1$**

"La probabilidad del espacio muestral es uno"

- **Axioma 3: $A \cap B = \varnothing \iff P(A \cup B) = P(A) + P(B)$**

"Dos sucesos son disjuntos si y sólo si la probabilidad de su unión es la suma de sus probabilidades".

De los tres axiomas, se deducen casi inmediatamente cinco consecuencias:

- **Consecuencia 1: $P(A) \leq 1$**

"La probabilidad tampoco puede ser mayor que uno"
Porque como $A \subset E$, si $P(A) > 1$ entonces necesariamente $P(E) > 1$, lo cual va en contra del segundo axioma.

- **Consecuencia 2: $P(A) + P(\overline{A}) = 1$**

"Las probabilidades de dos sucesos complementarios suman uno"
$P(E) = P(A \cup \overline{A})$ porque como vimos antes $A \cup \overline{A} = E$
$P(A \cup \overline{A}) = P(A) + P(\overline{A})$ por el tercer axioma, porque A y $\overline{A}$ son disjuntos.
y como $P(E) = 1$, $P(A) + P(\overline{A}) =1$
Esto es muy útil porque a menudo es más fácil calcular $P(\overline{A})$ que P(A), y entonces P(A) se obtiene de $P(A) = 1 - P(\overline{A})$

- **Consecuencia 3: $P(\varnothing) = 0$**

"La probabilidad de un suceso imposible es cero"
Intuitivamente, si un suceso es el conjunto vacío, es porque no contiene ningún resultado, y entonces nunca podría suceder (de ahí el nombre "imposible").
Como $\varnothing \cap \varnothing = \varnothing$, entonces por el tercer axioma:
$P(\varnothing \cup \varnothing) = P(\varnothing) + P(\varnothing)$
$P(\varnothing) = P(\varnothing) + P(\varnothing)$
$P(\varnothing) - P(\varnothing) = P(\varnothing)$
$P(\varnothing) = 0$

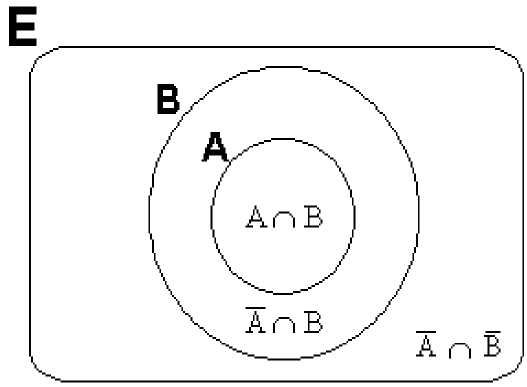
- **Consecuencia 4: $A \subset B \Rightarrow P(A) \leq P(B)$**

"Si un suceso está incluido en otro, su probabilidad es a lo sumo la de éste"

Partimos B en $A \cap B$ y $\overline{A} \cap B$ y aplicamos el tercer axioma:

$P((A \cap B) \cup (\overline{A} \cap B)) = P(A \cap B) + P(\overline{A} \cap B)$

$P(B) = P(A \cap B) + P(B \cap \overline{A})$



Partimos A en $A \cap B$ y $A \cap \overline{B}$ y aplicamos el tercer axioma:

$P((A \cap B) \cup (A \cap \overline{B})) = P(A \cap B) + P(A \cap \overline{B})$

$P(A) = P(A \cap B) + P(A \cap \overline{B})$

Pero como $A \subset B$, entonces $A \cap \overline{B} = \varnothing$, con lo cual $P(A \cap \overline{B}) = 0$, y entonces queda:

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

Y como, según calculamos antes, $P(B) = P(A \cap B) + P(B \cap \overline{A})$, queda:

$P(A) = P(B) - P(B \cap \overline{A})$

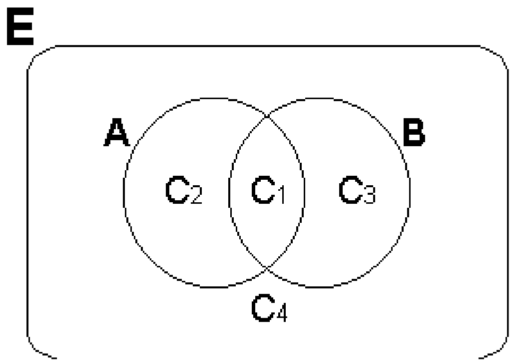
Y como $P(B \cap \overline{A}) \geq 0$, llegamos lo que queríamos demostrar.

Observemos que en el caso particular de que A no solamente esté incluido en B sino que sea igual a B (la igualdad de conjuntos es un caso particular de inclusión) entonces queda $P(B \cap \overline{A}) = 0$ y consecuentemente $P(A) = P(B)$.

• Consecuencia 5: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

La probabilidad de la unión de dos sucesos es la suma de sus probabilidades menos la probabilidad de la intersección.

Tomemos la siguiente partición de E: $\{C_1, C_2, C_3, C_4\}$ donde $C_1 = A \cap B$, $C_2 = A \cap \overline{B}$, $C_3 = \overline{A} \cap B$, $C_4 = \overline{A} \cap \overline{B}$



Luego:

$A = C_1 \cup C_2$ por propiedades de conjuntos

$B = C_1 \cup C_3$ por propiedades de conjuntos

$P(A) = P(C_1) + P(C_2)$ por el tercer axioma

$P(B) = P(C_1) + P(C_3)$ por el tercer axioma

$A \cup B = C_1 \cup C_2 \cup C_3$ por propiedades de conjuntos

$P(A \cup B) = P(C_1) + P(C_2) + P(C_3)$ por el tercer axioma dos veces

$A \cap B = C_1$ por propiedades de conjuntos

$P(A \cap B) = P(C_1)$ porque si $X = Y$ entonces $P(X) = P(Y)$

Juntando todo queda que:

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

es equivalente a:

$P(C_1) + P(C_2) + P(C_3) = P(C_1) + P(C_2) + P(C_1) + P(C_3) - P(C_1)$

Simplificando del lado derecho:

$P(C_1) + P(C_2) + P(C_3) = P(C_1) + P(C_2)\ P(C_3)$

Con lo cual la tercera consecuencia es válida.
Explicación intuitiva: Al construir $A \cup B$ "sumando" A y B estamos "contando" dos veces la intersección; por eso hay que restarla. $P(A \cup B) = P(A) + P(B)$ **$- P(A \cap B)$**
Cuando son disjuntos (el caso contemplado por el tercer axioma) la intersección es $\emptyset$, por eso en la expresión del axioma no hace falta que aparezca restando.

Generalización de la quinta consecuencia:

- **Para 3 sucesos:**

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

"La probabilidad de la unión de tres sucesos es:

las probabilidades individuales

menos las probabilidades de las intersecciones tomadas de a 2

más la probabilidad de la intersección tomada de a 3"

Análogamente:

- Para 4 sucesos:

"La probabilidad de la unión de cuatro sucesos es:

1) Las probabilidades individuales (sumando)
2) menos las probabilidades de las intersecciones tomadas de a 2
3) más las probabilidades de las intersecciones tomadas de a 3
4) menos la probabilidad de la intersección tomada de a 4"

Y así sucesivamente, alternando el signo se puede obtener la forma de calcular la probabilidad de la unión de cualquier número de sucesos.

Problemas típicos

1) Se tiran dos dados no cargados. Indique la probabilidad de que:

a) Salgan dos 3
b) Salgan dos 4
c) No salga ningún 5
d) Salga algún 5
e) No salga ningún 5 ni ningún 6
f) Salgan solamente números pares

Resolución
El espacio muestral es el siguiente:
E = { (1,1) , (1,2) , (1,3) , (1,4) , (1,5) , (1,6) , (2,1) , (2,2) , (2,3) , (2,4) , (2,5) , (2,6) , (3,1) , (3,2) , (3,3) , (3,4) , (3,5) , (3,6) , (4,1) , (4,2) , (4,3) , (4,4) , (4,5) , (4,6) , (5,1) , (5,2) , (5,3) , (5,4) , (5,5) , (5,6) , (6,1) , (6,2) , (6,3) , (6,4) , (6,5) , (6,6) }

Usamos este espacio muestral porque suponemos que sus elementos son equiprobables. Si hubiéramos considerado los dos dados no-distinguibles, entonces el suceso (1,2) tendría 2 formas posibles de ocurrir, y como vimos en el ejemplo de las monedas eso nos condujo a un espacio muestral no-equiprobable.
Queremos que el espacio muestral sea equiprobable para poder aplicar la definición de Laplace.
Hay 36 formas posibles de tirar los dos dados. Luego contando los resultados incluídos en cada suceso cuya probabilidad se pide, obtenemos:
a) 1/36
b) 1/36
c) 25/36
d) "salga algún 5" quiere decir "al menos un 5", es decir, 1 ó 2 cincos. En otras palabras, es el complemento del suceso a anterior. Su probabilidad es 11/36
e) 16/36
f) 9/36

2) En una determinada población, el 60% de las personas son mujeres, el 35% de la gente tiene ojos claros y el 25% de la gente es rubia. El 20% de la población son mujeres de ojos claros. El 10% de la población son mujeres rubias. El 15% de la población son personas rubias y de ojos claros. El 5% de la población son mujeres rubias de ojos claros. Calcule las probabilidades de que al elegir una persona al azar, esta:

a) sea mujer, sea rubia o tenga ojos claros (es decir, que tenga por lo menos una de esas 3 características.
b) tenga ojos oscuros
c) sea un hombre no rubio y de ojos oscuros
d) tenga cabello rubio o no tenga cabello rubio (alguna de las dos cosas).
e) tenga ojos claros y ojos oscuros (las dos cosas simultaneamente).
f) La probabilidad de encontrar a una mujer rubia, ¿es menor, igual, o mayor, que la de encontrar a una mujer rubia de ojos claros?

Resolución
Definiremos los sucesos:

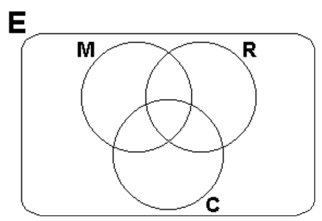
- M: la persona es mujer
- R: la persona es rubia
- C: la persona tiene ojos claros

Entonces los datos son:

$P(M) = 0.6$ $P(C) = 0.35$ $P(R) = 0.25$

$P(M \cap C) = 0.2$ $P(M \cap R) = 0.1$ $P(R \cap C) = 0.15$

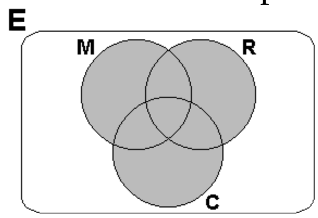
$P(M \cap C \cap R) = 0.05$



Vamos a resolver el ejercicio de 3 formas distintas.

• Forma 1: Aplicando los axiomas de la probabilidad y sus consecuencias para hallar las probabilidades pedidas.

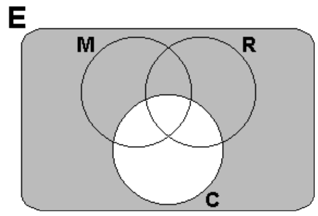
a) Nos piden $P(M \cup C \cup R)$. Por la generalización de la quinta consecuencia para 3 sucesos, sabemos que:



$P(M \cup C \cup R) = P(M) + P(C) + P(R) - P(M \cap C) - P(M \cap R) - P(C \cap R) + P(M \cap C \cap R)$

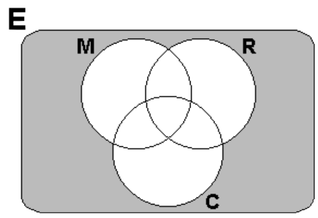
Y en este caso, todos los sumandos del lado derecho de la igualdad son dato. Entonces obtenemos:

$P(M \cup C \cup R) = 0.6 + 0.35 + 0.25 - 0.2 - 0.1 - 0.15 + 0.05 =$ **0.8**



b) El suceso "tener ojos oscuros" es la negación del suceso "tener ojos claros". Es decir, es el complemento de C. La segunda consecuencia nos dice que $P(A) + P(\overline{A}) = 1$, con lo cual:

$P(\overline{C}) = 1 - P(C) = 1 - 0.35 =$ **0.65**



c) Aquí el razonamiento es similar al del punto anterior. Si la persona elegida es hombre, no-rubio, y de ojos oscuros, no tiene ninguna de las 3 características M, C y R, y salió el complemento del conjunto $M \cup C \cup R$ (lo de afuera de los tres globlos del diagrama de Venn).

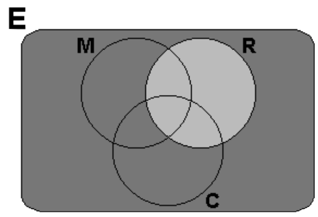
La segunda consecuencia dice que $P(A) + P(\overline{A}) = 1$, con lo cual si llamamos:

$A = M \cup C \cup R$

entonces lo que estamos buscando es $P(\overline{A})$, y como conocemos P(A), hacemos:

$P(\overline{A}) = 1 - P(A) = 1 - 0.8 =$ **0.2**

d) Estamos buscando $P(R \cup \overline{R})$. Como los sucesos complementarios son disjuntos (porque necesariamente $A \cap \overline{A} = \varnothing$), por el tercer axioma:



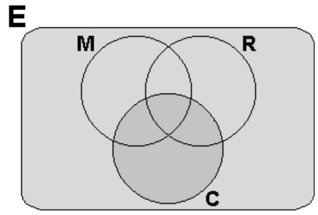
$P(R \cup \overline{R}) = P(R) + P(\overline{R})$.

Luego por la segunda consecuencia:

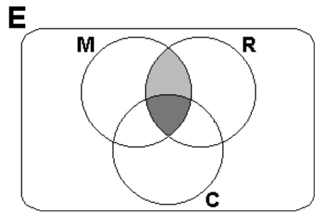
$P(R) + P(\overline{R}) =$ **1**

Este resultado era evidente, porque sólo se puede ser rubio o no-rubio. Sólo puede llover o no-llover. Por lo tanto la probabilidad de que suceda alguna de las dos cosas es necesariamente 1, porque siempre sucede alguna de las dos cosas.

e) Nos piden $P(C \cap \overline{C})$. C y su complemento no pueden ocurrir al mismo tiempo, porque una persona no puede tener ojos claros y ojos no-claros simultaneamente (supongamos que las personas tienen los dos ojos del mismo color). Entonces como las dos cosas no pueden ocurrir al mismo tiempo, la probabilidad de su intersección es necesariamente **cero**.



f) Las mujeres rubias pueden tener ojos claros u ojos oscuros. Siempre que una mujer sea rubia y de ojos claros, será necesariamente mujer rubia, pero no al revés, porque el hecho de que una mujer sea rubia no garantiza que además tenga ojos claros. Entonces la probabilidad de encontrar una mujer rubia que además tenga ojos claros es menor que la probabilidad de simplemente encontrar a una mujer rubia.



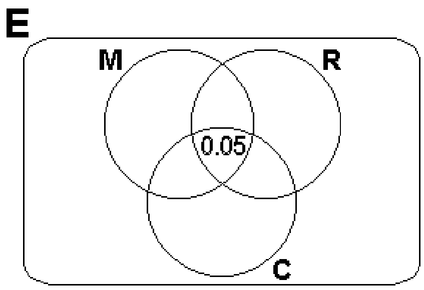
Si lo queremos pensar por la cuerta consecuencia:
$(M \cap R \cap C) \subset (M \cap R) \Rightarrow P(M \cap R \cap C) < P(M \cap R)$
(usamos < y no ≤ porque ≤ es para el caso particular en el cual un conjunto está incluido en otro porque ambos conjuntos son iguales (recordemos que $A = B \Rightarrow A \subset B$ y $B \subset A$)

- **Forma 2: Aplicando los axiomas de la probabilidad y sus consecuencias para hallar todas las probabilidades.**

Siendo los datos:

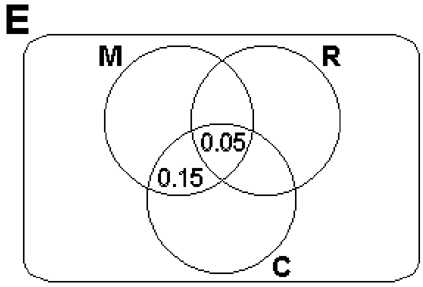
$P(M) = 0.6$	$P(C) = 0.35$	$P(R) = 0.25$
$P(M \cap C) = 0.2$	$P(M \cap R) = 0.1$	$P(R \cap C) = 0.15$
$P(M \cap C \cap R) = 0.05$		

1) En la intersección triple tenemos 0.05

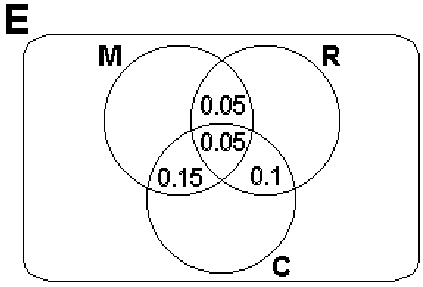


2) $(M \cap C)$ es la unión de los sucesos disjuntos:
$(M \cap C \cap R)$ y $(M \cap C \cap \overline{R})$.
Luego:
$P(M \cap C \cap R) + P(M \cap C \cap \overline{R}) = P(M \cap C)$

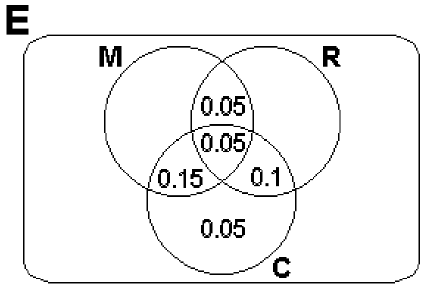
=> P(M ∩ C ∩ $\overline{R}$) = P(M ∩ C) - P(M ∩ C ∩ R) =
= 0.2 - 0.05 = 0.15



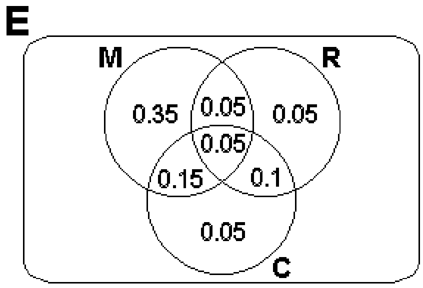
3) Análogamente aplicamos lo mismo para (M ∩ R) y para (R ∩ C). Es decir, sabemos que la probabilidad del "óvalo" (M ∩ R) debe dar en total 0.1, y que la probabilidad del "óvalo" (R ∩ C) debe dar en total 0.15.



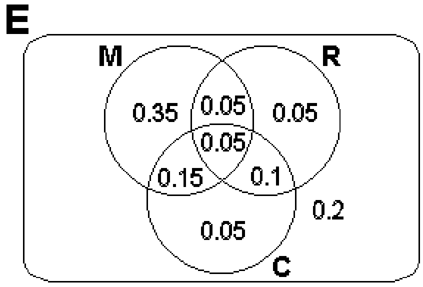
4) Sabemos que en total P(C) tiene que dar 0.35, por lo cual P($\overline{M}$ ∩ $\overline{R}$ ∩ C) debe dar 0.05.



5) Análogamente hacemos lo mismo para M y para R.



6) Como sabemos que P(E) debe dar en total 1, la probabilidad de la región que se encuentra afuera de los 3 conjuntos debe ser 0.2.

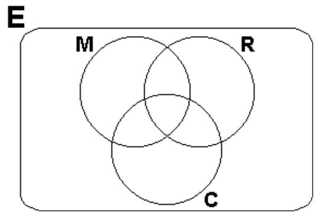


Luego las respuestas a las preguntas son inmediatas.

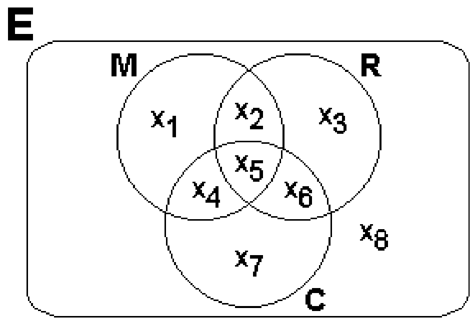
• **Forma 3: Planteando un sistema y resolviéndolo**

La tercera forma nos permite un mayor grado de automatización (que nos sería útil por ejemplo si fuéramos a desarrollar algún tipo de software que resolviera estas cuestiones).

Tomando los tres sucesos, el espacio muestral nos quedó dividido en $2^3 = 8$ regiones (el 2 porque al hacer el experimento puede pasar que ocurra o no ocurra (2 posibilidades) ese suceso, y el 3 porque eso lo aplicamos a cada uno de los 3 sucesos que estamos considerando). Tenemos entonces 8 incógnitas.



Comenzamos por ponerle nombre a cada una de las regiones. Si llamamos x_i a P(región i), entonces por ejemplo nos podría quedar como vemos en el gráfico. Luego escribimos ecuaciones a partir de los datos que tenemos:



Dato	Ecuación
P(M) = 0.6	$x_1 + x_2 + x_4 + x_5 = 0.6$
P(C) = 0.35	$x_4 + x_5 + x_6 + x_7 = 0.35$
P(R) = 0.25	$x_2 + x_3 + x_5 + x_6 = 0.25$
$P(M \cap C) = 0.2$	$x_4 + x_5 = 0.2$
$P(M \cap R) = 0.1$	$x_2 + x_5 = 0.1$
$P(R \cap C) = 0.15$	$x_5 + x_6 = 0.15$
$P(M \cap C \cap R) = 0.05$	$x_5 = 0.05$

Podría parecer que tenemos solamente 7 ecuaciones para las 8 incógnitas, pero también sabemos que la probabilidad del espacio muestral es 1, es decir:

$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 = 1$

El sistema ampliado queda:

$$\left(\begin{array}{cccccccc|c}
1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0.6 \\
0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0.35 \\
0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0.25 \\
0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0.2 \\
0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0.1 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0.15 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0.05 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1
\end{array}\right)$$

De donde por cualquier método, por ejemplo el de Gauss, obtenemos:
x_1=0.35 $x_2 = 0.05$ $x_3 = 0.05$ $x_4 = 0.15$
$x_5 = 0.05$ $x_6 = 0.1$ $x_7 = 0.05$ $x_8 = 0.2$

Con lo cual ya tenemos todo resuelto y estamos en condiciones de responder sobre las probabilidades de cualquiera de los 8 casos o uniones de ellos.
Para hallar las respuestas podemos sumar todas las probabilidades x_i de las regiones que cumplan con la condición. Si las regiones que cumplen con la condición son muchas, podemos hacer 1 - [las probabilidades de las regiones que NO cumplen con la condición].
Luego:

a) $1 - x_8 = 0.8$
b) $x_1 + x_2 + x_3 + x_8 = 0.65$
c) $x_8 = 0.2$
d) $x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + x_8 = 1$
e) 0
f) mujer rubia: $x_2 + x_5 = 0.1$
mujer rubia de ojos claros: $x_5 = 0.05$
$0.1 > 0.01$