I.3.Probabilidad Condicionada. Independencia. Verosimilitud

February 21, 2019

I.3 Probabilidad condicionada. Independencia de sucesos. Verosimilitud

Probabilidad condicionada

Consideremos que tenemos un espacio de probabilidad $(\Omega, \mathscr{F}, P)$. Recordemos que un **suceso** $A \in \mathscr{F}$, $A \subset \Omega$ es un subconjunto formado por algunos resultados posibles del espacio muestral. También que el suceso acontece si lo hace cualquiera de los resultados que lo componen.

Supongamos ahora que sabemos que acontece otro suceso B del espacio de probabilidad. ¿Podemos utilizar este conocimiento para mejorar nuestro conocimineto de A? Por ejemplo, si se lanza un dado bueno sabemos que las probabilidades de todas sus caras son idénticas e iguales a $\frac{1}{6}$. Si nos dicen que el resultado que ha salido es par, esto es el suceso $\{par\}$, ello nos permite saber que las probabilidades de todos los resultados impares es nula, y que el de los pares pasa a ser de $\frac{1}{3}$ cada una.

Consideremos que estamos interesados en el suceso A, que tiene una probabilidad $P(A)$. Supongamos que observamos un suceso relacionado B. ¿Aumenta de algún modo la certidumbre que tenemos en relación al suceso A por haber acontecido otro relacionado B?

Definimos la **probabilidad del suceso** A **condicionada por el suceso** B como

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

Como sabemos que B, el **suceso condicionante**, acontece, podemos interpretarlo como el espacio muestral en un nuevo espacio de probabilidad con una nueva asignación de probabilidades $P(\bullet/B)$. Adviértase que:

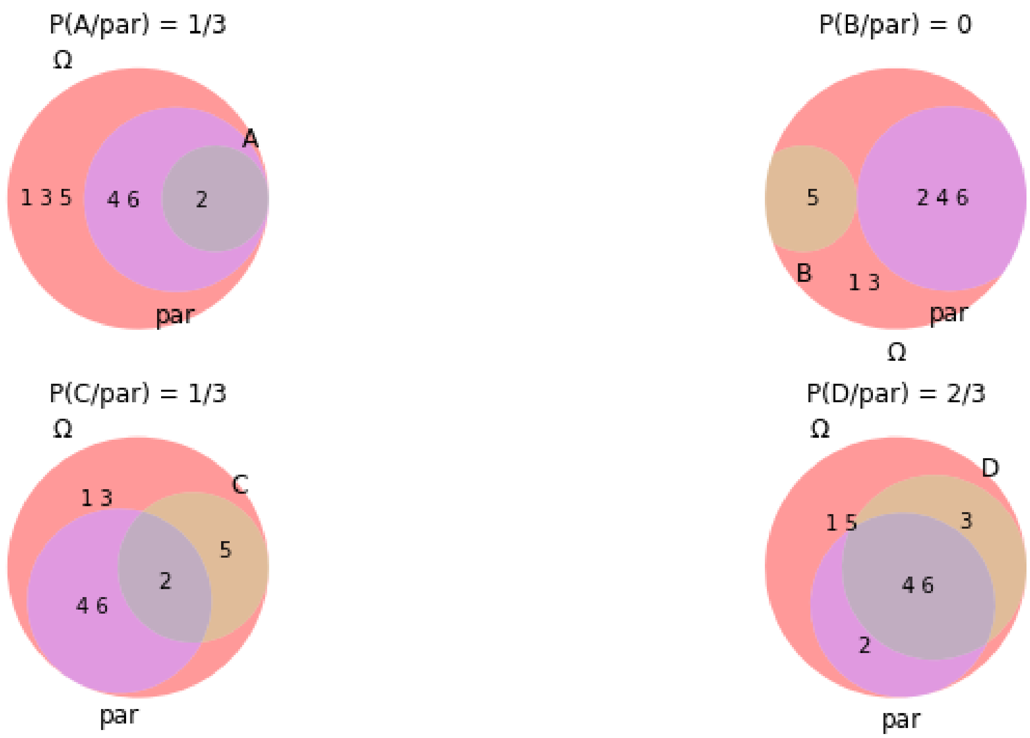
- B se convierte en el suceso seguro al saber que ha acontecido, por lo que $P(B/B) = 1$. Dado que $B \cap B = B$, la división en la definión por $P(B)$ nos asegura que, efectivamente, el nuevo suceso seguro tiene probabilidad uno.
- Si $A \subset B$ entonces $P(A/B) = \frac{P(A)}{P(B)}$.
- Sin embargo A, el suceso cuya probabilidad condicionada queremos calcular, puede tener elementos que no están en B aunque sí en el espacio muestral original. $A \cap B$ excluye los elementos de A que no están en B, razón por la que se considera $P(A \cap B)$.

Consideremos un dado bueno (sin trucar) con seis caras numeradas. Supongamos que estamos interesados en conocer la probabilidad de cuatro sucesos: $A = \{2\}$, $B = \{5\}$, $C = A \cup B = \{2, 5\}$, $D = \{3, 4, 6\}$:

- $P(A) = P(B) = \frac{1}{6}$
- $P(C) = P(A \cup B) = P(A) + P(B) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$
- $P(D) = 3 \times \frac{1}{6} = \frac{1}{2}$

Consideremos ahora que sabemos que ha salido *par*, que corresponde al suceso $par = \{2, 4, 6\}$, aunque no sabemos qué número. ¿Podemos mejorar el cálculo de probabilidades?

- $P(A/par) = \frac{P(A \cap par)}{P(par)} = \frac{P(A)}{P(par)} = \frac{P(2)}{P(par)} = \frac{1/6}{1/2} = \frac{1}{3}$
- $P(B/par) = \frac{P(B \cap par)}{P(par)} = \frac{P(\emptyset)}{P(par)} = \frac{0}{1/2} = 0$
- $P(C/par) = \frac{P(C \cap par)}{P(par)} = \frac{P(A)}{P(par)} = \frac{P(2)}{P(par)} = \frac{1/6}{1/2} = \frac{1}{3}$
- $P(D/par) = \frac{P(D \cap par)}{P(par)} = \frac{P(\{4,6\})}{P(par)} = \frac{2 \times 1/6}{1/2} = \frac{2}{3}$



Independencia de sucesos

Dos sucesos A y B son independientes cuando que uno de ellos acontezca no tiene ningún efecto en que lo haga o lo deje de hacer el otro. En tal caso:

$$P(A/B) = \frac{P(A \cap B)}{P(B)} = P(A) \quad \text{y} \quad P(B/A) = \frac{P(B \cap A)}{P(A)} = P(B)$$

y, consecuentemente, $P(A \cap B) = P(A)P(B)$ que suele tomarse también como definición de independencia.

Es muy importante **no confundir sucesos independientes y sucesos incompatibles**. Recordemos que éstos son sucesos disjuntos, esto es, su intersección es nula y, por tanto, no pueden acontecer simultáneamente. Por tanto, si uno sucede sabemos que el otro no puede suceder. Ello es radicalmente diferente de la independencia, que significa que si un suceso sucede no sabemos nada de qué va a pasar con el otro.

Consideremos de nuevo el experimento de probabilidad correspondiente al lanzamiento de un dado bueno.

Los sucesos *par* e *impar* son disjuntos ($A \cap B = \emptyset$) y, por ello, incompatibles. Si el resutado es par, ciertamente sabemos que no puede haber salido impar. Igualmente sucede con los sucesos *div*3 (divisible por 3) y *nodiv*3 (no divisible por 3) y, en general, con cualquier conjunto de sucesos disjuntos.

Pero, supongamos ahora que sabemos que el resultado ha sido par. £Afecta ello a que el resultado haya sido no divisible por 3? Si no lo hace, los sucesos *par* y *nodiv*3 son independientes. Veámoslo...

Es evidente que:

- $par \cap nodiv3 = \{2,4\} \implies P(par \cap nodiv3) = \frac{2}{6} = \frac{1}{3}$
- $P(nodiv3) = P(\{1,2,4,5\}) = \frac{4}{6} = \frac{2}{3}$

Por tanto:

- $P(nodiv3/par) = \frac{P(par \cap nodiv3)}{P(par)} = \frac{\frac{2}{6}}{\frac{3}{6}} = \frac{2}{3}$
- $P(nodiv3)P(par) = \frac{2}{3}\frac{1}{2} = \frac{1}{3}$

Vemos que, como dice la intuición si se reflexiona sobre ello, *nodiv*3 y *par* son independientes, ya que:

- $P(nodiv3/par) = P(nodiv3) = \frac{2}{3}$
- $P(par \cap nodiv3) = P(nodiv3)P(par) = \frac{1}{3}$

0.1 Verosimilitud

La **verosimilitud** (*likelihood*) corresponde a la probabilidad condicionada, dejándose variar el suceso condicionante. Si se satisface el suceso A, la verosimilitud evalúa cuanto varía la probabilidad condicionada de dicha observación con los sucesos condicionantes B_i, $i \in \mathbb{N}$:

$$L(B_i) \equiv P(A/B_i) \qquad i \in \mathbb{N}$$

A la vista del suceso A el principio de **"máxima verosimilitud"** (*maximum likelihood*) o **ML** considera que el el suceso condicionante **más verosímil** para la observacion es el que hace máxima la verosimilitud $L(B_i)$:

$$B_i \quad / \quad max(L(A/B_i)) \quad \forall i \in \mathbb{N}$$

Un sencillo ejemplo muestra que el principio de máxima verosimilitud corresponde al *sentido común*. Consideremos un dado, dos posibles sucesos condicionantes, *div*3 y *nodiv*3, y un suceso correspondiente a una observación $A = \{1,2,4,6\}$.

Si observamos el suceso A, esto es, si sabemos que ha salido 1 ó 2 ó 6 nos preguntamos si es **más versosímil** que el lanzamiento haya dado un resultado divisible por tres (*div*3) o no divisible por tres (*nodiv*3).

Adviértase que **no nos preguntamos si es más probable** que el lanzamiento haya sido *div*3 o *nodiv*3, pues ello supondría conocer $P(div3/A)$ y $P(nodiv3/A)$, **sino más verosímil**, esto es, conociendo $P(A/div3)$ y $P(A/nodiv3)$.

En estos momentos **la distinción entre más probable y más verosímil puede resultar artificiosa**, pues estamos utilizando ejemplos muy sencillos. Sin embargo, con frecuencia, conocemos la probabilidad condicionada en un sentido y no en el opuesto. Es más, **los sucesos condicionantes los estamos considerando aleatorios. Sin embargo esto no es estrictamente necesario, con lo que podría ni siquiera tener sentido la búsqueda de las probabilidades de tales sucesos condicionantes**.

Calculemos la verosimilitud, considerando el dado bueno:

$$L(div3) \equiv P(A/div3) = \frac{P(\{1,2,4,6\} \cap \{3,6\})}{P(\{3,6\})} = \frac{^1/_6}{^1/_3} = \frac{1}{2}$$

$$L(nodiv3) \equiv P(A/nodiv3) = \frac{P(\{1,2,4,6\} \cap \{1,2,4,5\})}{P(\{1,2,4,5\})} = \frac{^1/_2}{^2/_3} = \frac{3}{4}$$

Tal y como era fácil intuir, es más verosímil que el resultado no sea divisible por tres, ya que A se satisface con tres de cuatro sucesos elementales (todos menos el 6), mientras que si fuera divisible por tres, A tan solo se satisface en uno de dos (el 6).

Adviértase que **la verosimilitud NO es una probabilidad**, como queda puesto de manifiesto sin más que ver que $P(A/div3) + P(A/nodiv3) > 1$. Casualmente podría sumar uno (prúebese, por ejemplo, para $A = \{1,2,6\}$), pero seguiría sin ser una probabilidad.