MSEL
Modelling in Science Education and Learning

Modelling in Science Education and Learning
Volume 13 (1), 2020 doi: 10.4995/msel.2020.12115.
Instituto Universitario de Matemática Pura y Aplicada
Universitat Politècnica de València

Probabilidades de sucesos independientes desde la simulación

Probabilities of independent events from the simulation

F.J. Boigues, V.D. Estruch, A. Vidal
Universitat Politècnica de València
fraboipl@mat.upv.es, vdestruc@mat.upv.es, avidal@mat.upv.es

Abstract

En este trabajo se describe una metodología para el estudio de las probabilidades de sucesos independientes en la asignatura Instrumentos de Estadística y Simulación, de segundo curso del Grado en Ciencias Ambientales de la Universitat Politècnica de València. Dicha metodología se basa en la utilización de un formalismo basado en grafos para la representación de los sucesos involucrados y en introducir la simulación como herramienta para el cálculo de probabilidades.

This paper describes a methodology for the study of the probabilities of independent events in the subject Instruments of Statistics and Simulation, of the second year of the Degree in Environmental Sciences of the Universitat Politècnica de València. This methodology is based on the use of a graph-based formalism for the representation of the involved events and the introduction of simulation as a tool for the calculation of probabilities.

Palabras clave: Sucesos independientes, probabilidad, grafos, simulación
Keywords: Independent events, probability, graphs, simulation

1. La probabilidad en la enseñanza superior

El pensamiento aleatorio y estadístico es importante, para el ingeniero y para el científico, desde la perspectiva del ejercicio de la responsabilidad de tener que plantear soluciones a muchos problemas. Por ello la universidad continúa con la tarea, iniciada en la educación secundaria, de preparar a los futuros científicos e ingenieros en el pensamiento probabilístico (Angarita y Parra, 2017). Así pues, la estadística es considerada como parte de las materias básicas en la formación de los estudiantes universitarios para la adquisición de competencias de cara a la investigación o el ejercicio profesional.

Según investigaciones recientes, el aprendizaje significativo de la estadística en niveles universitarios necesita sustentarse en la comprensión de conceptos fundamentales, en base a los cuales se sigue una construcción progresiva de representaciones mentales implícitas o explícitas más complejas (Escalante, 2008). Este hecho justifica el abordar con rigor el aprendizaje de las bases teóricas que sustentan las técnicas de análisis estadístico, y muy especialmente la teoría de la probabilidad. Behar et al. (2002) considera que es muy importante resolver problemas relacionados con la cuantificación del riesgo, lo cual implica conocer elementos básicos de probabilidad. Posteriormente se han de adquirir las competencias necesarias para que, partiendo de sucesos simples, se pueda valorar el riesgo de sucesos más complejos.

1.1. Modelos probabilísticos y problemas reales

La probabilidad es un concepto básico que se utiliza para estudiar los fenómenos aleatorios que siguen ciertas reglas de comportamiento. Dichas reglas suelen deducirse al estudiar los fenómenos de forma extensiva, llegándose entonces a la relación directa del concepto de probabilidad con el de frecuencia relativa, ya que las frecuencias relativas se estabilizan al repetirse sucesivamente el experimento. Este planteamiento frecuencial de la probabilidad facilita, posteriormente, establecer el nexo evidente entre la teoría de la probabilidad y la estadística aplicada, el cual se basa en la noción de variable aleatoria (Boigues y Estruch, 2017). El planteamiento frecuencial de la probabilidad también será fundamental a la hora de justificar el llegar a resultados prácticos mediante métodos de simulación de Monte Carlo.

1.2. La probabilidad en las propuestas docentes

En gran parte de los cursos introductorios a la estadística, y en los libros que los apoyan, el tema dedicado a la teoría de la probabilidad presenta unos contenidos similares: Definición de suceso y espacio muestral de sucesos, probabilidad de un suceso, probabilidades de la unión e intersección de sucesos, probabilidad condicional e independencia y teoremas de la probabilidad total y de Bayes. La forma de desarrollar este tema a nivel práctico suele reducirse a la resolución de problemas que, aunque en muchos casos se contextualizan en situaciones reales, no permiten más margen de maniobra que la simple identificación de sucesos y la aplicación de las reglas del cálculo de probabilidades aprendidas. Este enfoque limita la posibilidad de entender la probabilidad como elemento estructural para el desarrollo de modelos que permitan afrontar situaciones reales complejas. Tendría que ser, precisamente, el tema dedicado a la teoría y cálculo de probabilidades el que marcara el punto de partida para introducir técnicas de simulación. Enseñar/aprender la probabilidad como una actividad de modelización y no como un conjunto estático de resultados matemáticos que se deducen de una serie de axiomas no es fácil (Batanero, 2003). Sin embargo el enfoque de la modelización es, posiblemente, el más adecuado para los futuros profesionales o investigadores que necesitarán recursos dinámicos y adaptables para resolver problemas reales.

2. Objetivos

En este trabajo, se presenta una propuesta didáctica que se concreta en una actividad práctica cuyo objetivo es que los alumnos adquieran las siguientes competencias:

- Trabajar con la unión e intersección de sucesos independientes.
- Construir modelos complejos a partir de estructuras relativamente sencillas asentadas en sistemas de sucesos.
- Calcular probabilidades de sucesos relativamente complejos mediante simulación.

En los apartados 3.1, 3.2, 3.3 y 3.4, se resume el desarrollo teórico de la actividad práctica propuesta, cuyo contenido está orientado a la resolución de problemas de probabilidad con sucesos independientes mediante simulación. En el apartado 3.5 se propone un caso realista que el estudiante debe modelar mediante un sistema y resolver mediante simulación.

3. Propuesta didáctica. Sucesos y sistemas

En lo que sigue, utilizaremos una formalización de los sucesos y sus probabilidades mediante grafos que representan sistemas.

3.1. Suceso simple y simulación

Sea un suceso simple, A, con probabilidad de suceder $P(A) = p$. A dicho suceso se le puede asociar un grafo o sistema (Figura 1), en el que interpretaremos que la pretensión es desplazarnos de I a II y A es una puerta que permite el paso (está abierta) con probabilidad $P(A) = p$.

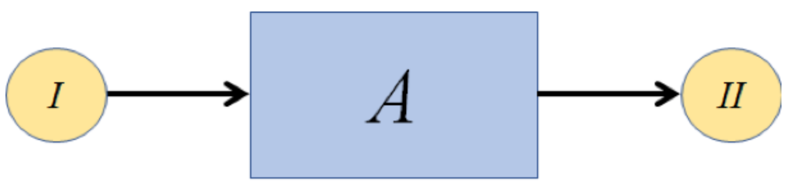

Figura 1: Esquema de un grafo

Simular si sucede un suceso simple, A, es sencillo. Basta con generar un valor entre 0 y 1, que llamaremos $rand$. Si $rand < p$ entonces sucede A. En caso contrario, no sucede A. La repetición de dicha simulación, si esta es correcta, debe llevarnos a que el porcentaje de casos en que sucede A será aproximadamente $p \times 100\,\%$.

A continuación, se introducirá la unión e intersección de sucesos y su probabilidad, relacionándola con circuitos en paralelo y en serie, respectivamente.

3.2. Sistema en paralelo y probabilidad de la unión de sucesos

Dado dos sucesos A y B se define la unión, $A \cup B$, como aquel suceso para el que podemos decir que ocurre $A \cup B$ si y sólo si sucede alguno de los dos, A ó B. En cambio, la intersección de A y B, $A \cap B$, sería aquel suceso que ocurre si ocurre A y también ocurre B. Se tiene que

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \tag{1}$$

Y si los sucesos son independientes se tiene que

$$P(A \cap B) = P(A) \cdot P(B). \tag{2}$$

La unión de sucesos puede representarse mediante un grafo que representa un sistema en paralelo Figura 2(a):

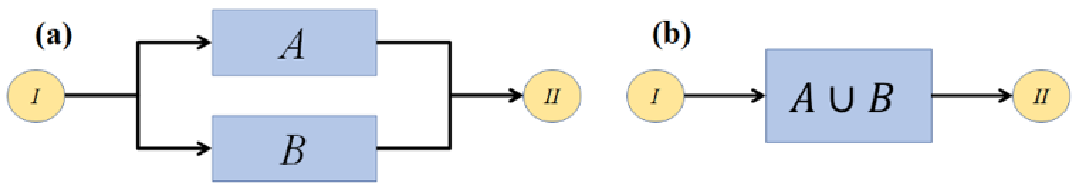

Figura 2: Sistema en paralelo

Obviamente, el sistema de la Figura 2(a) es equivalente al de un suceso simple Figura 2(b).

Interpretaremos la Figura 2(a) de forma que el paso de I a II será posible si la puerta A está abierta (ocurre A) o está abierta B (ocurre B). Si A tiene una probabilidad $P(A) = p_1$ y B tiene una probabilidad $P(B) = p_2$, entonces la probabilidad de poder pasar de I a II será igual a $P(A \cup B) = P(A) + P(B) - P(A \cap B) = p_1 + p_2 - p_1 \cdot p_2$

A este mismo resultado podemos aproximarnos mediante simulación. En primer lugar generamos un suceso del sistema:

1. Consideramos el vector de probabilidades de A y B, $P = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}$, y generamos una matriz con dos valores aleatorios $S = \begin{pmatrix} rand(1) \\ rand(2) \end{pmatrix}$.

2. Formamos la matriz T a partir de la operación lógica $T = (S < P)$. T será un vector con 2 componentes, que serán unos ó ceros: el elemento i de T valdrá 1 si $rand(i) < p_i$ y valdrá 0 si $rand(i) \geq p_i$, $i = 1, 2$.

3. Calculamos la suma de los elementos de T, $Sum(T)$. Si el resultado es mayor que 0 es porque hay alguna puerta abierta, con lo que podemos pasar de I a II, o lo que es lo mismo, ocurre $A \cup B$. En cambio, si $Sum(T)$= 0 las dos puertas estarían cerradas y por tanto no ocurre $A \cup B$.

Si repetimos el proceso descrito muchas veces, n, y contabilizamos que hay k ocasiones en que $Sum(T) > 0$ entonces se tendrá que $P(A \cup B) \approx k/n$.

3.3. Sistema en serie y probabilidad de la intersección

La intersección de sucesos puede representarse mediante un grafo que representa un sistema en paralelo (Figura 3(a)). El sistema de la Figura 3(a) sería equivalente a un sistema correspondiente a un suceso simple (Figura 3(b)):

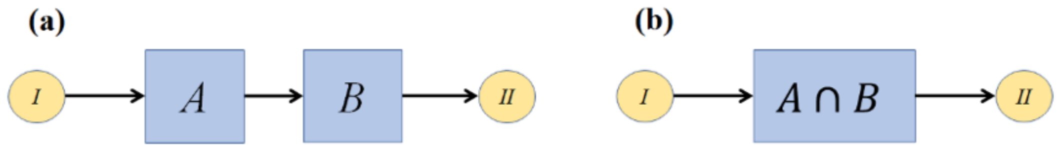

Figura 3: Sistema en serie

Interpretaremos la Figura 3(a) de forma que el paso de I a II será posible si la puerta A está abierta (ocurre A) y también está abierta B (ocurre B). Si $P(A) = p_1$ y $P(A) = p_2$, entonces la probabilidad de poder pasar de I a II será igual a $P(A \cap B) = p_1 \cdot p_2$.

A este mismo resultado podemos aproximarnos también mediante simulación. En primer lugar generamos un suceso del sistema:

1. Considerando el vector de probabilidades de A y B, $P = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}$, y generamos una matriz S como en el caso anterior.
2. Formamos la matriz T como en el caso anterior, con $T = (S < P)$.
3. Calculamos el producto de los elementos de $T, Prod(T)$. Si $Prod(T) > 0$ ($Prod(T) = 1$), es porque todas las puertas están abiertas, con lo que podemos pasar de I a II, o lo que es lo mismo, ocurre $A \cap B$. De lo contrario, $Prod(T) = 0$, que significa que alguna de las dos puertas está cerrada y no se podría pasar de I a II.

Si repetimos el proceso descrito muchas veces, n, y contabilizamos que hay k ocasiones en que $Prod(T) > 0$, entonces se tendrá que $P(A \cap B) \approx k/n$.

3.4. Probabilidades en sistemas complejos

A partir de sistemas en paralelo o en serie, podemos construir sistemas más complejos. En la Figura 4(a) se describe un sistema con cierta complejidad y, en la Figura 4(b) se visualiza como el sistema inicial se reduce fácilmente a un sistema en paralelo, más simple.

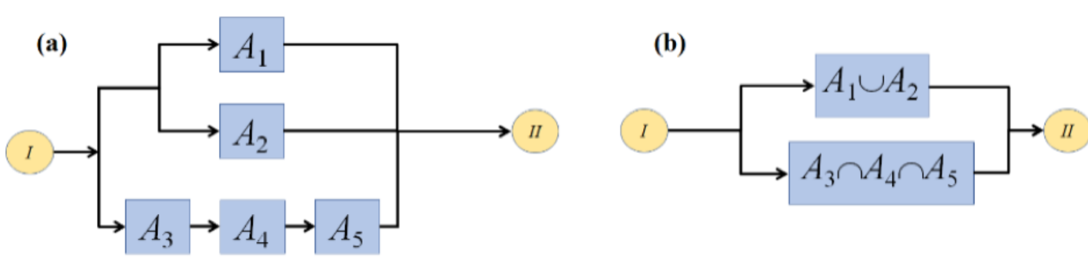

Figura 4: (a) Sistema inicial. (b) Reducción del sistema inicial a un sistema en paralelo.

Se supone que las componentes en la Figura 4, A_1, A_2, A_3, A_4 y A_5 son sucesos independientes. En este caso, el paso de I a II será posible si está abierto el paso por A_1 o por A_2 o, si a la vez, esta abierto el paso por A_3, A_4 y por A_5. Por lo tanto, para el caso de la Figura 4, el paso de I a II, corresponderá al suceso $(A_1 \cup A_2) \cup (A_3 \cap A_4 \cap A_5)$, cuya probabilidad llamaremos P . Obtener el valor de P presenta cierta complejidad, y se puede calcular de forma exacta, utilizando propiedades fundamentales del cálculo de probabilidades, obteniéndose $P = 0.9568$. Pero también podemos usar la simulación para obtener un valor aproximado de P. A continuación se presenta una implementación en MATLAB, donde el valor final `P_Fun` es un valor aproximado de P:

```
clear all;
P=[0.7,0.7,0.7,0.8,0.8];
n=10000;Fun=0; %Se repite 10000 veces el proceso
for i=1:n
    S = rand(1,5); %Se generan 5 valores U(0,1)
    T0 =(S < P); %Se obtiene qué puertas quedan abiertas
    T1 = sum(T0(1:2)); %Se obtiene si el grupo en paralelo A1,A2 queda abierto
    T2 = prod(T0(3:5)); %Se obtiene si el grupo en serie A3, A4, A5 queda abierto
    R = ((T1+T2)>0); %Se obtiene si el sistema globalmente queda abierto
    Fun = Fun+R;
end
P_Fun = Fun/n
```

Se ha obtenido que `P_Fun = 0.9499`, es decir $P \approx 0.95$

3.5. Modelización probabilística mediante sistemas

Exponemos dos motivos importantes para trabajar con sucesos independientes siguiendo el formalismo de los sistemas. En primer lugar el lenguaje gráfico facilita la modelización de situaciones reales con cierta complejidad. En segundo lugar, los sistemas permiten entender, abstraer y resolver más fácilmente el problema, facilitando también la implementación de las rutinas para el cálculo de las probabilidades mediante simulación. A continución se presenta un ejemplo ilustrativo que se ha propuesto al alumnado, después de realizar una práctica con ejemplos explicativos (Boigues, Estruch y Vidal, 2018).

Ejercicio propuesto. En un incendio de una zona forestal se pretende que el fuego no pase del sector 1 al sector 2. El paso de un sector a otro se puede producir en dos ubicaciones A_1 y A_2. Para ello se establecen tres puntos de ataque. El primero se establece en una ubicación, A_1 , de forma que la probabilidad de que pase el fuego sería igual a 0.45. Por otra parte, en la ubicación A$_2$, se establecen dos puntos de ataque sucesivos, $A_{2,1}$ y $A_{2,2}$, de forma que la probabilidad de que pase el fuego sería 0.5 y 0.65, respectivamente. Con la estrategia de ataque establecida, ¿Cuál sería la probabilidad de que el fuego pase del sector 1 al sector 2? Calcúlese de forma teórica y mediante simulación.

Diseñamos un posible enfoque. El enunciado planteado sería fácilmente representable, de forma gráfica, mediante un sistema (Figura 5)

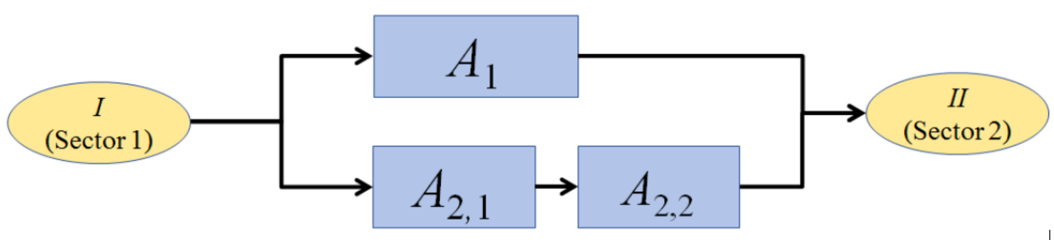

Figura 5: Sistema formal correspondiente al ejemplo Ejercicio propuesto

La probabilidad a calcular es

$$\begin{aligned} P &= P\big(A_1 \cup (A_{2,1} \cap A_{2,2})\big) \\ &= P(A_1) + P(A_{2,1} \cap A_{2,2}) - P(A_1 \cap A_{2,1} \cap A_{2,2}) \\ &= P(A_1) + P(A_{2,1}) \cdot P(A_{2,2}) - P(A_1) \cdot P(A_{2,1}) \cdot P(A_{2,2}) \\ &= 0.628, \end{aligned}$$

donde denotamos mediante A_1 y $A_{2,j}$ (para $j = 1, 2$) los sucesos *que pase el fuego por la ubicación* A_1 y *que pase el fuego por la ubicación* $A_{2,j}$, respectivamente.

El siguiente script de MATLAB nos proporcionaría una aproximación a la probabilidad buscada, a través de la simulación:

```
P=[0.45, 0.5,0.65];  n=10000;  Fun=0; %se consideran 10000 repeticiones
for i=1:n
    S=rand(1,3); % Se generan 3 valores U(0,1)
   T0=(S<P); % Se obtiene qué puertas quedan abiertas
   T1=T0(1); % Se obtiene si A1 queda abierta
   T2=prod(T0(2:3)); % Se obtiene si el grupo en serie A2,1, A2,2, queda abierto
   R=((T1+T2)>0); % Se obtiene si el sistema globalmente queda abierto
    Fun=Fun+R;
end
P_Fun=Fun/n
```

El resultado de una simulación es `P_Fun =0.6277`. Por lo tanto el suceso "que pase el fuego de un sector a otro", tendrá un probabilidad aproximada del 63 %, muy similar al teórico.

4. Resultados. La experiencia en el aula

Los resultados obtenidos por 27 estudiantes, tras haber trabajado la propuesta didáctica expuesta en una clase de prácticas informáticas de la asignatura Instrumentos de Estadística y Simulación del Grado en Ciencias Ambientales de la Universitat Politècnica de València, han sido diversos. Por una parte, hay 9 estudiantes (33.3 %) que no calculan la probabilidad de forma teórica mientras que 15 (55.6 %) lo resolvieron bien. Por otra parte 12 estudiantes (44.4 %) no realizaron la simulación y 14 (51.9 %) la realizaron correctamente.Cabe indicar que algunos estudiantes optaron por una vía mixta, resolviendo teóricamente la parte del sistema que está en serie, para luego, mediante simulación, resolver un sistema en paralelo. En este último caso los pasos seguidos han sido:

1. Se calcula $P(A_{2,1} \cap A_{2,2}) = P(A_{2,1}) \cdot P(A_{2,2}) = 0.5 \cdot 0.65 = 0.325$.

2. Se simula el sistema en paralelo utilizando la rutina MATLAB

```
P=[0.45, 0.325]; n=10000; Fun=0;
for i=1:n
    S=rand(1,2);  %Se generan 2 valores U(0,1)
   T=(S<P);  %Se obtiene qué puertas quedan abiertas
   k= sum(T));  %Se obtiene si el grupo en paralelo queda abierto
   R=((k)>0);  %Se obtiene si el sistema globalmente queda abierto
    Fun=Fun+R;
end
P_Fun=Fun/n
```

Posteriormente se realizó un examen de prácticas en el que se propuso un sistema similar al del ejemplo, pero invirtiendo el orden de las componentes, apareciendo primero las dos componentes en serie. Se plantearon las mismas preguntas y los alumnos disponían de todo el material deseado para consultas. Un total de 32 estudiantes realizaron el examen. De nuevo 14 estudiantes (43.8 %) no hicieron el cálculo teórico y 12 (37.5 %) lo resolvieron correctamente. En cuanto a la simulación, 11 estudiantes (34.4 %) no la hicieron, y 12 (37.5 %) la hicieron correctamente. Entre los que la hicieron mal, 2 de ellos recurrieron a la rutina presentada en clase, sin tener en cuenta el cambio en el orden de las componentes.

5. Conclusiones

La experiencia en el desarrollo de la práctica ha evidenciado, en primer lugar, la necesidad de dedicar más tiempo al trabajo con grafos, de forma que el alumnado pudiese adquirir mayor habilidad a la hora de simplificar sistemas con cierta complejidad. Esto facilitaría poder proponer algún problema con un grado mayor de complejidad técnica, de forma que la opción de su resolución mediante simulación estuviese plenamente justificada. Por lo demás, la simulación se manifiesta como una potente herramienta que complementa la formación teórica del futuro ingeniero ya que le ofrece alternativas potentes y eficientes para resolver problemas complejos.

Referencias

Angarita, M.A.O., Parra, A.B.S. (2017).
Importancia de la probabilidad y la estadística en la formación del Ingeniero. I3+, 1(2), 26.
http://doi.org/10.24267/23462329.63

Batanero, C. (2003).
La simulación como Instrumento de modelización en probabilidad.
Revista Educación y Pedagogía, XV(35), 39–54.

Behar, R., Klinger, R., Olaya, J., Andrade, M., Mesa, E., Conde, G., Delgado J., Arbeláez, D., Trujillo, P., Solano, H., Díaz R.N. (2002).
El Rol de la Estadística en el trabajo del Ingeniero.
Ingeniería y Competitividad, 4(1), 47–54.

Boigues Planes, F.J., Estruch Fuster, V.D. (2018).
Aproximación frecuencialista de la Probabilidad.
RIUNET, UPV.
http://hdl.handle.net/10251/82991

Boigues Planes, F.J., Estruch Fuster, V.D., Vidal Meló, A. (2017).
Prácticas de estadística y simulación en el ámbito medio ambiental. Apuntes.
Universitat Politècnica de València.

Escalante Gómez, E. (2008).
Actitudes de alumnos de posgrado hacia la estadística aplicada a la investigación.
Encuentro 2010/ Año XLII, n° 85, 27–38.

Kay, S.M. (2006).
Intuitive probability and random processes using MATLAB.
New York, Springer.