Introducción a la probabilidad y estadística CM274

César Lara Avila

9 de agosto de 2017

https://github.com/C-Lara

1. Introducción a la probabilidad

Introducción a la teoria de conjuntos

Un conjunto es una colección desordenada de elementos.

Un conjunto A con un número finito de elementos se denomina conjunto finito y su tamaño (número de elementos) se denota como $|A|$. Un conjunto con un número infinito de elementos se denomina conjunto infinito.

En muchos casos consideramos conjuntos indexados por un conjunto finito o infinito. Por ejemplo $U_a, a \in A$ representa múltiples conjuntos, un conjunto para cada elemento de A. Para múltiples conjuntos $U_a, a \in A$, definimos las operaciones de unión e intersección, como sigue

$$\bigcup_{a\in A} U_a = \{u : u \in U_a \text{ para uno o más } a \in A\}$$

$$\bigcap_{a\in A} U_a = \{u : u \in U_a \text{ para todo } a \in A\}.$$

Introducción a la teoria de conjuntos(1)

Ejemplo Si $A_1 = \{1\}, A_2 = \{1,2\}, A_3 = \{1,2,3\}$, tenemos:

$$\{A_i : i \in \{1,2,3\}\} = \{\{1\},\{1,2\},\{1,2,3\}\}$$
$$\bigcup_{i\in\{1,2,3\}} = \{1,2,3\}$$
$$\bigcap_{i\in\{1,2,3\}} = \{1\}$$

Los conjuntos $U_a, a \in A$ son mutualmente disjuntos si $\bigcap_{a\in A} U_a = \emptyset$ y son disjuntos dos a dos si $a \neq b$ implica que $U_a \cap U_b = \emptyset$. La unión de conjuntos disjuntos dos a dos $U_a, a \in A$ es denotado algunas veces como $\biguplus_{a\in A} U_a$.

Introducción a la teoria de conjuntos(2)

El <u>conjunto potencia</u> de un conjunto A, es el conjunto de todos los subconjuntos de A, incluyendo el conjunto vacio $\emptyset$ y A. Este conjunto es denotado por 2^A. Si A es un conjunto finito entonces

$$|2^A| = 2^{|A|}$$

El conjunto $U_a, a \in A$ forma una <u>partición</u> de U si

$$\biguplus_{\alpha \in A} U_\alpha = U.$$

La generalización del tamaño de un conjunto A a conjuntos infinitos, puede hacerse notando que dos conjuntos finitos A, B tienen el mismo tamaño si y sólo si existe una biyección entre ellos, y generalizando esta noción a conjuntos infinitos.

Introducción a la teoria de conjuntos(3)

Ley distributiva de conjuntos: Sean los conjuntos $U_a, a \in A$ y un conjunto C,

$$\left(\bigcup_{a\in A} U_a \bigcap C\right) = \bigcup_{a\in A}\left(U_a \bigcap C\right)$$
$$\left(\bigcap_{a\in A} U_a \bigcup C\right) = \bigcap_{a\in A}\left(U_a \bigcup C\right)$$

(Ley de Morgan) Sean los conjuntos $U_a, a \in A$,

$$\left(\bigcup_{a\in A} U_a\right)^c = \bigcap_{a\in A} U_a^c$$
$$\left(\bigcap_{a\in A} U_a\right)^c = \bigcup_{a\in A} U_a^c$$

Definiciones básicas

Probabilidad es el lenguaje para cuantificar la incertidumbre.

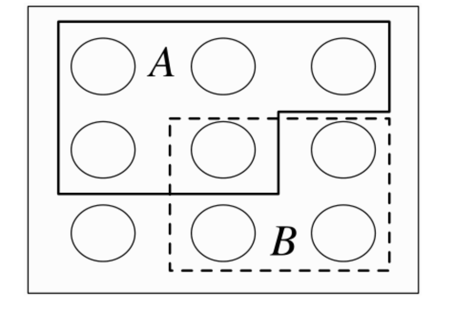
Conceptos:

- El **espacio muestral** Ω de un experimento es el conjunto de todos los resultados o salidas del experimento.
- Puntos ω en Ω son llamados: **elementos, realizaciones o resultados muestrales**.
- Subconjuntos de Ω son llamados **eventos**.

El espacio muestral de un experimento puede ser finito, infinito contable o infinito no contable. En particular el conjunto vacio $\emptyset$ y el espacio muestra Ω son eventos.

Espacio muestral

Cuando el espacio muestral es finito, se puede visualizar como la figura siguiente.



Cada óvalo representa una salida y un evento es un conjunto de óvalos A o B.

Casos de espacios muestrales

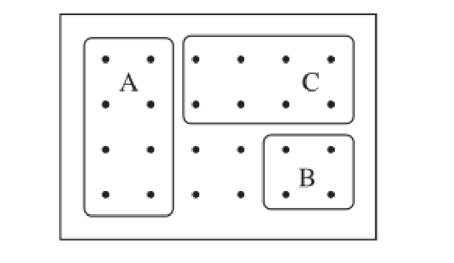
1. En el experimento aleatorio de lanzar una moneda tres veces y observar los resultados el espacio muestral es el conjunto $\Omega = \{CCC, CCS, CSC, CSS, SCC, SCS, SSC, SSS\}$. El evento $A = \{CCC, CCS, CSS, CSC\} \subset \Omega$ describe `una cara que se obtuvo en el lanzamiento de la primera moneda`. En este caso tanto el espacio muestral Ω como el evento A son conjuntos finitos.
2. Considere un experimento aleatorio de lanzar un dardo en un tablero redondo, sin salirse del tablero. Suponiendo que el radio del tablero es 1, el espacio muestral es el conjunto de todos los vectores bidimensionales dentro del círculo unitario
$$\Omega = \left\{(x, y) \,:\, x, y \in \mathbb{R},\ \sqrt{x^2 + y^2} < 1\right\}$$
. Un evento que describe un golpe en el tablero puede puede ser,
$$A = \left\{(x, y) \,:\, x, y \in \mathbb{R},\ \sqrt{x^2 + y^2} < 0{,}1\right\} \subset \Omega.$$
En ambos casos el espacio muestral Ω y el evento A son infinitos incontables.

Eventos

1. Para un evento A, el resultado del experimento aleatorio $\omega \in \Omega$ está en A ($\omega \in A$) o no estén $A(w \notin A)$. En el primer caso, decimos que el evento A ocurrió y en el segundo caso decimos que el evento A no ocurrió.
2. $A \cup B$ es el evento de que A o B ocurren y $A \cap B$ es el evento de que A y B ocurren. El complemento A^c (el complemento en el conjunto universal se toma como $\Omega : A^c = \Omega - A$) representa el evento de que A no ocurrió.
3. Si los eventos A, B son disjuntos ($A \cap B = \emptyset$), los dos eventos no pueden ocurrir al mismo tiempo, ya que ningún resultado del experimento aleatorio pertenece tanto a A como B. Si $A \subset B$ entonces si B ocurre implica que A también ocurre.

Más sobre eventos

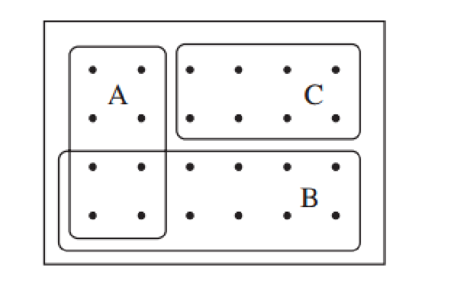
Una lista de eventos se dice mutualmente exclusivo si ningún elemento del espacio muestral pertenece a más de un evento. Como lo muestra la siguiente figura:



Esto implica que la ocurrencia de cualquiera de estos eventos implica la no ocurrencia de los restantes eventos

Más sobre eventos

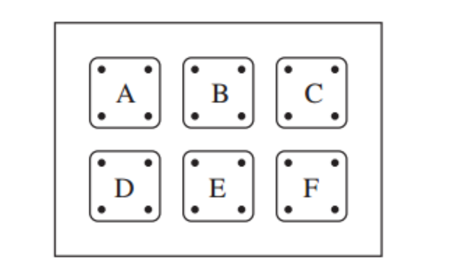
Cuando todos los elementos en un espacio muestral pueden ser encontrados en al menos un evento en una lista de eventos, entonces la lista de eventos se dice colectivamente exhaustivos.



En este caso ningún elemento del espacio muestral, es omitido y un único elemento puede pertenecer a más de un evento, como se muestra en la anterior figura:

Más sobre eventos

Se dice que los eventos que son mutuamente exclusivos y colectivamente exhaustivos, forman una partición del espacio muestral. Como se muestra en la figura.



Cualquier conjunto de eventos mutuamente excluyentes y colectivamente exhaustivos se denomina espacio de eventos.

Caso explicativo

- Secuencias de bits se transmiten a través de un canal de comunicación en grupos de cinco. Cada bit puede ser recibido correctamente o puede ser modificado en tránsito, lo que ocasiona un error. Consideremos un experimento que consiste en observar los valores de bits a medida que llegan e identificarlos con la letra c si el bit es correcto y con la letra e si el bit está en error.

  El espacio muestral consta de 32 resultados de $ccccc$ a $eeeee$, de cero bits transmitidos incorrectamente a los cinco bits que están en error.

  Sea el evento $A_i, i = 0, 1, \ldots, 5$, consistente de todos los resultados en los cuales los i bits están en error. Así
  $A_0 = \{ccccc\}, A_1 = \{ecccc, ceccc, ccecc, cccec, cccce\}$ y así sucesivamente hasta $A_5 = \{eeeee\}$. Los sucesos $A_i, i = 0, 1, \ldots 5$, particionan el espacio muestral y por lo tanto constituyen un espacio de eventos.

Función probabilidad

Una función probabilidad $\mathbb{P}$ es una función que asigna un número real a cada evento $A \subset \Omega$ y que satisface los siguientes tres axiomas:

1. $\mathbb{P}(A) \geq 0, \quad$ para todo evento A.
2. $\mathbb{P}(\Omega) = 1$.
3. Sean $A_n, n \in \mathbb{N}$ es una secuencia de eventos disjuntos por pares $(A_i \cap A_j = \emptyset)$ siempre que $i \neq j$, entonces

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

Propiedades básicas de la función probabilidad

Podemos derivar algunas propiedades de $\mathbb{P}$ desde los anteriores axiomas.

1. $\mathbb{P}(\emptyset) = 0$.
2. $A \subset B \rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$.
3. $0 \leq \mathbb{P}(A) \leq 1$.
4. $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.
5. $A \cap B = \emptyset \rightarrow P(A \cup B) = P(A) + P(B)$.

Lectura: Chapter 1 Probability Probability, Markov Chains, Queues and Simulation Willian J.Stewart Princeton University Press 2009.

Importantes propiedades de la función probabilidad

- (Principio de inclusión-exclusión) Para dos eventos A y B,

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

En general el principio de inclusión-exclusión se puede escribir como,

$$\begin{aligned}
\mathbb{P}\left(\bigcup_{i}^{n} A_i\right) &= \sum_{i} \mathbb{P}(A_i) - \sum_{i<j} \mathbb{P}(A_i \cap A_j) \\
&+ \sum_{i<j<k} \mathbb{P}(A_i \cap A_j \cap A_k) \\
&\vdots \\
&+ (-1)^n \mathbb{P}(A_i \cap A_2 \dots A_n)
\end{aligned}$$

El signo que precede a una suma con la intersección de m conjuntos es $(-1)^{m+1}$. La razón para sumar sobre índices crecientes es evitar el doble conteo.

Se debe notar que si los conjuntos son disjuntos dos a dos, las intersecciones anteriores son todas vacías y por lo tanto tienen probabilidad cero y esto se reduce a la aditividad finita.

Importantes propiedades de la función probabilidad

- (Aditividad finita de la probabilidad) Para cada secuencia $A_1, \ldots A_n$ de eventos disjuntos dos a dos (por pares) $(A_i \cap A_j) = \emptyset$ siempre que $i \neq j$), entonces,

$$\mathbb{P}(A_1 \cup \cdots \cup A_n) = \mathbb{P}(A_1) + \cdots \mathbb{P}(A_n).$$

- Para una secuencia creciente o decreciente de eventos $\{E_n, n \geq 1\}$, se cumple que

$$\lim_{n \to \infty} \mathbb{P}(E_n) = \mathbb{P}(\lim_{n \to \infty} E_n).$$

- (Continuidad de la probabilidad) Si $A_n \to A$ entonces

$$\mathbb{P}(A_n) \to \mathbb{P}(A) \qquad \text{si } n \to \infty$$

- (Lema de Borel-Cantelli) Si se tiene $\{A_1, A_2, A_3, \ldots\}$ una secuencia de eventos. La serie $\displaystyle\sum_{n=1}^{\infty} \mathbb{P}(A_n)$ converge, entonces $\mathbb{P}\left(\bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty}\right) = 0$.

Ejemplo explicativo

Un profesor distraido escribió n cartas y las selló en sobres antes de escribir las direcciones en los sobres. Luego escribió las n direcciones en los sobres al azar. ¿ Cuál es la probabilidad de que al menos una carta fue dirigida correctamente?.

El número total de formas en que uno puede escribir n direcciones en n sobres es $n!$. Por lo que el espacio muestral contiene $n!$ puntos. Ahora calculamos el número de resultados en el cual al menos un sobre se trata correctamente.

Para hacer esto, sea E_i el evento de que la i ésima carta es direccionada correctamente, entonces $E_1 \cup E_2 \cup \cdots \cup E_n$ es el evento de que al menos una carta es direccionada correctamente.

Para calcular $\mathbb{P}(E_1 \cup E_2 \cup \cdots \cup E_n)$, utilizamos el principio de inclusión-exclusión.

Ejemplo explicativo(1)

Desde que $\mathbb{P}(E_1 \cup E_2 \cup \cdots \cup E_n)$ hay n términos de la forma $\mathbb{P}(E_i)$, $\binom{n}{2}$ términos de la forma $\mathbb{P}(E_i \cap E_j)$, $\binom{n}{3}$ términos de la forma $\mathbb{P}(E_i \cap E_j \cap E_k)$ y así.

$$\mathbb{P}(E_1 \cup E_2 \cup \cdots \cup E_n) = n\frac{(n-1)!}{n!} - \binom{n}{2}\frac{(n-2)!}{n!} + \cdots + (-1)^{n-2}\binom{n}{n-1}\frac{[n-(n-1)]!}{n!} + (-1)^{n-1}\binom{n}{n}\frac{1}{n!}.$$

Esta expresión simplificada, es de la forma:

$$\mathbb{P}(E_1 \cup E_2 \cup \cdots \cup E_n) = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \cdots + \frac{(-1)^{n-1}}{n!}.$$

Función probabilidad de un espacio muestral discreto

Para un espacio muestral finito (finito contable) Ω un evento conteniendo un único elemento $A = \{\omega\}, \omega \in \Omega$ es llamado un <u>evento elemental</u>. Para un espacio muestral con n elementos, supongamos que se nos da un conjunto de n números no negativos $\{p_\omega : \omega \in \Omega\}$ que suman a uno, entonces existe una función de probabilidad única $\mathbb{P}$ sobre esos eventos tales que $\mathbb{P}(\{\omega\}) = p_\omega$.

Esta probabilidad se define para eventos arbitrarios a través de la propiedad de aditividad finita:

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\{\omega\}) = \sum_{\omega \in A} p_\omega .$$

En la clásica interpretación de probabilidad sobre espacios muestrales finitos, las probabilidades de todos los eventos elementales $\{\omega\}, \omega \in \Omega$ son iguales. Desde que la función de probabilidad debe satisfacer $\mathbb{P}(\Omega) = 1$, tenemos
$\mathbb{P}(\{\omega\}) = |\Omega|^{-1}, \text{para todo} \quad \omega \in \Omega$. Esto implica que bajo el modelo clásico sobre un finito Ω, tenemos

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|}.$$

Función probabilidad de un espacio muestral continuo

Para un espacio muestral continuo de dimensión n (por ejemplo $\Omega = \mathbb{R}^n$), definimos la función de probabilidad clásica como

$$\mathbb{P}(A) = \frac{\mathrm{vol}_n(A)}{\mathrm{vol}_n(\Omega)},$$

donde $\mathrm{vol}_n(S)$ es el volumen n-dimensional del conjunto S.

El volumen 1-dimensional de un conjunto $S \subset \mathbb{R}$ es su longitud. El volumen 2-dimensional de un conjunto $S \subset \mathbb{R}^2$ es su area. El volumen 3-dimensional de un conjunto $S \subset \mathbb{R}^3$ es su volumen. En general el volumen n-dimensional de A es la n-ésima integral de la función constante 1 sobre el conjunto A.