Mecánica Estadística. Clase 1. Probabilidades y Estadística

Prof. Juan Mauricio Matera

06/03/2025

Clase 1. Nociones de teoría de Probabilidades y Estadística.

Nociones elementales

Para construir un modelo probabilístico del sistema de partículas que discutimos en la introducción, conviene introducir algunas nociones formales sobre lo que llamamos *teoría de probabilidades.*

En la literatura, encontramos dos nociones de "probabilidad": *frecuentista* y "bayesiana". Para la versión *frecuentista*, la probabilidad de un resultado corresponde a la "frecuencia relativa" con la que se obtendría ese resultado luego de realizar una gran cantidad de veces *el mismo experimento* en forma independiente. Dichas frecuencias están *determinadas* por un modelo subyacente, al que accedemos vía el análisis de los resultados de realizar *el mismo experimento* una *gran cantidade de veces.* Pero

- ¿Qué significa hacer el mismo experimento?
- ¿Cuántas veces son "una gran cantidad de veces?

La concepcion *bayesiana* propone evitar estas preguntas, definiendo a la probabilidad como una *expectativa* de observar una cierta frecuencia relativa, a partir de alguna hipótesis inicial (subjetiva), y de los datos observados hasta el momento. Trabajar con la formulación bayesiana requiere una formulación matemática más sofisticada que la requerida por la frecuentista. Por otro lado, como la formulación frecuentista será suficiente para las aplicaciones en este curso, nos enfocaremos en esta última.

El concepto frecuentista de probabilidad fue formalizado matemáticamente por el matemático ruso Andrey Kolmogorov (1903-1987) introduciendo los siguientes conceptos:

- Espacio muestral Ω al conjunto de resultados posibles en un experimento, o a un conjunto que lo contenga.
- Evento[1]: Todo subconjunto de Ω, que representa los resultados compatibles con cierto enunciado lógico es un evento. El conjunto de los eventos $\Sigma(\Omega)$ es cerrado frente a la unión, intersección y complemento de conjuntos. Además, el espacio muestral es un evento.
  - Ω: cualquier resultado cumple la condición.
  - $A \cap B$: Se deben cumplir ambas condiciones.
  - $A \cup B$: Se deben cumplir una de ambas condiciones.
  - $\overline{A}$: No se debe cumplir la condición A.
  - $\varnothing$: Es una condición que no se puede cumplir (un experimento siempre da algún resultado).

  Matemáticamente, el conjunto de los eventos posibles es una *sigma-álgebra.*
- Una *medida de probabilidad* es una función $P : \Sigma(\Omega) \to [0, 1] \subset \mathbb{R}$, es decir, una función que asigna a cada evento $A \in \Sigma(\Omega)$ un valor entre 0 y 1 y que cumple:
  - $P(\Omega) = 1$ y $P(\varnothing) = 0$
  - $P(A \cup B) = P(A) + P(B)$ si $A \cap B = \varnothing$
  - $P(\bigcup_{k=0}^{\infty} A_{n\to\infty}) = \lim_{n\to\infty} P(A_n)$

  Matemáticamente, $(\Omega, \Sigma(\Omega), P)$ representan lo que se conoce como *Espacio de medida 1.*

Observamos que si $\Omega^L = \Omega \times \Omega \times \ldots \Omega$ representa los resultados posibles de experimento que consiste en repetir el experimento Ω L veces. Si se obtiene el resultado $R = (m_1, \ldots, m_L) \in \Omega^L$, la frecuencia relativa para el evento A se define como

$$f_L(A|R) = \frac{1}{L} \left|\{i|m_i \in A, i = 1, \ldots, L\}\right|$$

[1]En esta discusión voy a seguir las definiciones del libro "Probabilidad y Estadística Elementales para estudiantes de ciencias" de J. Marona. Arovas llama *event* a los elementos Ω. Otros textos en inglés se refieren a los elementos de Ω como *outcomes* (resultados). Si revisamos la entrada en Wikipedia, la version el artículo en inglés sigue dicha convención, pero el artículo en castellano llama *eventos* a los elementos de Ω y *sucesos* a sus subconjuntos. Ténganlo en cuenta si consultan otras fuentes bibliográficas.

que cumple las mismas propiedades que P. Aquí utilizamos la notación $\|A\|$ para indicar el *tamaño* del conjunto A. Si Ω es *discreto* - esto es sus elementos se pueden contar-, el *tamaño* del conjunto es su *cardinalidad*, esto es, la cantidad de elementos que contiene. La teoría permite también trabajar con espacios de probabilidad *continuos*. Por ejemplo, si elegimos $\Omega = \mathbb{R}^d$, podemos definir una probabilidad sobre *regiones*. El *tamaño* de estas regiones viene dado por su *volumen*. En un experimento concreto, los resultados posibles son siempre *discretos*, ya que tanto la resolución como el alcance de los instrumentos que usamos para medir y contar son finitos. En general, la descripción continua es útil para establecer resultados que no dependen de un *contexto de observación* particular[2], o por una cuestión de conveniencia en el cálculo.

Distribuciones de probabilidad discretas

Si el espacio muestral es discreto, $\Omega = \{e_1, e_2, ...\}$ entonces la medida de probabilidad queda determinada por las probabilidades de los eventos $\omega_i = \{e_i\}$. Decimos entonces que $p_i = P(\omega_i)$ es la probabilidad de e_i. Naturalmente lo mismo aplica si e_i es una *sucesión de elementos* en Ω tal que $P(\{e_1, e_2, ...\}) = 1$. Podemos definir entonces un vector $p \equiv p_i$ de componentes no-negativas al que llamamos *vector de probabilidad*. Vía los axiomas de probabilidad, es posible entonces reconstruir la medida de probabilidad de cualquier evento *sumando* las probabilidades de sus elementos.

Cuando tal sucesión no existe, decimos que el espacio de probabilidad es *continuo*. En la práctica, la elección del uso de espacios de probabilidad discretos o continuos es más una cuestión de conveniencia en la descripción: Si el espacio muestral es $\mathbb{R}^n$, siempre podemos dividirlo convenientemente en *regiones* tan pequeñas como querramos, y construir un nuevo espacio muestral donde los elementos sean esas regiones. Este procedimiento se conoce como coarse-graining (granulado grueso).

Alternativamente, podemos *embeber* una distribución discreta en $\mathbb{R}^n$ asociando cada elemento del espacio muestral e_i en una región $\mathcal{R}_i \subset \mathbb{R}^n$ y definir una distribución de probabilidad tal que $p_i = P(\mathcal{R}_i)$.

Eventos equiprobables. Conteo.

En un espacio muestral es *discreto*, el caso opuesto a una distribución *pura* corresponde al cual ningún elemento de Ω es más *probable* que otro. Decimos entonces que la distribución es *equiprobable*, y

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

esto es, la probabilidad es proporcional a la *cantidad de elementos* (o *cardinalidad*) asociados al evento A. Conviene entonces disponer de algunas herramientas matemáticas para realizar este conteo.

- Si A consiste de una lista de l elementos en un orden definido, cada uno de ellos elegidos entre m posibles elementos, $|A| = m^l$.
- Si A consiste de todas las permutaciones posibles de los mismos l elementos $\{w_1, w_2, \ldots, w_l\}$, distintos entre sí, entonces $|A| = l \times (l-1) \times (l-2) \times \ldots 2 \times 1 = l!$. Se define además $0! = 1$, ya que hay una única forma de ordenar un conjunto vacío.
- Se llama*variaciones* de L en k al número de formas de elegir una lista ordenada de L elementos distintos tomados de un conjunto de N elementos
$$(N)_L = \frac{N!}{(N-L)!}$$
- La cantidad de subconjuntos (sin ordenar) de L elementos, tomados de un conjunto de N elementos es el *combinatorio* (o las combinaciónes) de N en L, y se expresa
$$\left(\begin{array}{c} N \\ L \end{array} \right) = \frac{N}{L!(N-L)!} = \frac{(N)_L}{L!}$$
- El número de formas de distribuir L elementos en S clases de tamaño m_k se expresa como el *coeficiente multinomial*
$$\left(\begin{array}{c} L \\ m_1 ... m_S \end{array} \right) = \frac{N!}{m_1! ... m_s!}$$

En particular, para un espacio muestral de las observaciones de L experimentos iguales, los eventos en los que las frecuencias relativas para cada elemento $\omega_i \in \Omega$ resultan ser $m_{i=1,\ldots,|\Omega|}/L$

$$A_{\{m_i\}} = \{R | f_L(\{\omega_i\}|R) = m_i/l\}$$

continen una cantidad de elementos dada por la multinomial

$$|A_{\{m_i\}}| = \left(\begin{array}{c} L \\ m_1 ... m_{|\Omega|} \end{array} \right)$$

[2] El contexto de observación tiene que ver con las limitaciones que tenemos al observar un fenómeno o medir cierta cantidad. Por ejemplo, el resultado de cualquier medida se encuentra restringuido a una cierta escala, y puede ser derminado con una cierta resolución finita, por lo que sólo tiene sentido expresarlo con una cantidad finita de dígitos. Volveremos a esto más adelante.