LA AXIOMÁTICA DE KOLMOGOROV: FUNDAMENTO DE LA TEORÍA DE LA PROBABILIDAD

Natividad Jiménez Saavedra

La palabra *probabilidad* afecta hoy en día a un amplio abanico de situaciones en nuestra vida, lo cual se desprende del abundante uso que se hace de esta palabra en todos los medios de comunicación, en nuestras casas y en todo el mundo científico. Y si bien tenemos una idea de lo que es la *probabilidad*, muchas veces resulta difícil definirla de una manera que se aplique a cada situación donde usamos el término y en la que estemos de acuerdo la mayoría de las personas.

La definición que Kolmogorov establece en 1933 resuelve el problema desde un punto de vista formal y práctico y es, sin duda alguna, lo que ha configurado el gran desarrollo de la teoría de la probabilidad a partir de ese momento. El problema que nos concierne es cómo hacer llegar este bello concepto, altamente abstracto en su expresión formal, a personas de otras disciplinas que no tienen una base matemática para entenderlo.

Kolmogorov

Para ello basta recordar que muchos modelos matemáticos parten de un simple hecho de la vida que deseamos explicar. Y muchas veces sucede que se desarrolla toda una teoría a partir de ese modelo abstracto, y luego resulta aplicable a muchas y muy variadas situaciones. Este es el caso de la *probabilidad*. Quizás, en algún momento, también ha sucedido

37

que nos hemos quedado atrapados en los "modelos", y hemos perdido la belleza subyacente tras ellos y, posiblemente en muchos casos, hemos perdido la capacidad de comunicar la idea a personas que no hablan nuestro lenguaje, pero que, con toda certeza, pueden beneficiarse de su comprensión y de su utilización. Este también puede ser el caso de la *probabilidad*. Como P. Chebyshev ha dicho:

> La aproximación entre la teoría y la práctica asegura los resultados más favorables, y no sólo la práctica se beneficia con esto; las ciencias se desarrollan bajo el influjo de la práctica que descubre nuevos objetos a investigar o facetas desconocidas de los objetos ya conocidos... si la teoría gana mucho aplicando un método viejo o sus modificaciones, gana más creando métodos nuevos, y en este caso la ciencia cuenta con un guía certero en su quehacer.

Con el objeto de clarificar la definición de *probabilidad,* vamos a relacionarla con las nociones intuitivas e interpretaciones fundamentales que constituyen la base para las aplicaciones.

Concepto intuitivo de probabilidad

En el pensamiento popular abundan nociones imprecisas, pero intuitivas, de probabilidad. En este sentido, la palabra *probabilidad* indica una apreciación de la facilidad que se atribuye a que ocurra cierto acontecimiento aleatorio (suceso aleatorio), partiendo de una tendencia (más o menos inconsciente) a pensar que unos hechos son más verosímiles que otros, y de un deseo de medir esa verosimilitud. Incluso en este tipo de pensamiento, la *probabilidad*
37 se considera como un número p que se le asigna a un suceso A, que puede ocurrir como resultado de algún experimento.

Esta idea de que la probabilidad del suceso aleatorio A admite, bajo ciertas condiciones, una estimación cuantitativa mediante un número, fue desarrollada en el siglo XVII, en las obras de P. Fermat, B. Ch. Huygens y, en especial, de J. Bernoulli. Y es precisamente este deseo de querer medir de una forma más objetiva la probabilidad, lo que llevaría a Kolmogorov a establecer su axiomática varios siglos después.

Concepto frecuentista de probabilidad

La evidencia experimental muestra cómo, en experimentos que se pueden repetir bajo idénticas condiciones, la frecuencia relativa de un suceso concreto se aproxima a un valor, y a este valor se le denomina la *probabilidad* de que ocurra ese determinado suceso. Este concepto frecuentista está apoyado por la Ley empírica del azar (teorema de Bernoulli), publicada en 1713 en su obra póstuma "Ars Conjectandi":

> Cuando el número de realizaciones de un experimento aleatorio crece mucho, la frecuencia relativa del suceso asociado se va acercando cada vez más y más hacia un cierto valor.

Esta definición no pudo darse por válida durante mucho tiempo porque presentaba varios inconvenientes:

1. No da ningún indicio del grado de aproximación entre la frecuencia relativa y la probabilidad.

2. Tampoco da indicios acerca del número de experimentos que se deben realizar para conseguir una aproximación previamente establecida.

3. Su validez se fundamenta únicamente en la experiencia.

Sin embargo, es interesante observar tres de las propiedades que verifican la frecuencia relativa $fr(A)$ de un suceso A:

1. La frecuencia relativa de un suceso es un número comprendido entre cero y uno; esto es: $0 \leq fr(A) \leq 1$, para todo suceso A.

2. La frecuencia relativa del suceso seguro S vale uno; esto es: $fr(S) = 1$.

3. Si A y B son dos sucesos incompatibles, la frecuencia relativa de su unión es la suma de sus frecuencias relativas; es decir: si A y B son disjuntos, entonces $fr(A \cup B) = fr(A) + fr(B)$.

Concepto clásico de probabilidad

Laplace, a finales del siglo XVIII, introdujo la definición de probabilidad de un suceso A como "el cociente entre el número de casos favorables a que ocurra dicho suceso y el número de casos posibles, bajo el supuesto de que todos los casos posibles sean igualmente probables".

Desde un punto de vista formal, no obstante, esta definición es incorrecta, pues introduce el término definido en la definición. Y, atendiendo a un punto de vista práctico, sólo es aplicable a los casos de "equiprobabilidad", dejando al margen muchos casos de interés. Además, si el número de resultados posibles fuera infinito, esta definición sería inadecuada.

Sin embargo, en un gran número de aplicaciones es imposible determinar las probabilidades de los sucesos vía la definición frecuentista, repitiendo el experimento un número "suficiente" de veces. En tales casos, la única alternativa que nos queda es usar la definición clásica como una "hipótesis de trabajo", y aceptar esta hipótesis si las consecuencias observables validan la experiencia, y rechazarla en otro caso.

Por tanto, esta definición no permite la construcción y desarrollo de una teoría matemática de la probabilidad, lo que justifica el rechazo que tuvo durante mucho tiempo, y la necesidad de encontrar una definición satisfactoria.

Concepto axiomático de probabilidad

La definición axiomática de Kolmogorov surge a partir de la acumulación de diferentes hechos y del desarrollo de otras disciplinas científicas y matemáticas. Con esto, el autor coloca en su lugar natural los conceptos básicos de la teoría de la probabilidad, conceptos que hasta este momento se consideraron bastante peculiares. No obstante, es importante señalar que esta construcción axiomática arranca de las propiedades fundamentales de la probabilidad, que surgen del análisis de las definiciones frecuentista y clásica anteriormente descritas, y que, además, estas definiciones quedan totalmente justificadas a partir de estos axiomas (véase Ash, Robert B.)

El sistema de axiomas establecido por Kolmogorov, cuando tratamos con probabilidades de sólo un número finito de sucesos, es el siguiente:

> Sea Ω una colección de elementos, llamados *sucesos elementales*, y sea F un conjunto de subconjuntos de Ω; a los elementos del conjunto F se les llama "sucesos aleatorios". F es un *álgebra* de conjuntos.
>
> I. F contiene el conjunto $\varnothing$.
>
> II. A cada conjunto A de F se le asigna un número real no negativo $P(A)$. Este número $P(A)$ se llama "probabilidad del suceso A".
>
> III. $P(\Omega) = 1$
>
> IV. Si A y B son disjuntos, entonces $P(A + B) = P(A) + P(B)$.
>
> Un sistema de conjuntos F, con una asignación definida de números $P(A)$ que satisfaga los axiomas I - IV se llama *álgebra de probabilidad.*

Nota de Kolmogorov:

> Un sistema de conjuntos se llama *álgebra* si la suma, producto, y la diferencia de dos conjuntos del sistema también pertenece al mismo sistema. Por $A+B$ se entiende el conjunto compuesto por los elementos de Ω contenidos en A o B, o bien en A y B simultáneamente; por AB, el conjunto de los elementos de Ω pertenecientes a A y B al mismo tiempo; y, finalmente, por A^* (B^*) se entiende el conjunto de los elementos de Ω no contenidos en A (o en B).

[Literalmente de la traducción al inglés (1950) del trabajo original, en alemán, de Kolmogorov. Aquí también se puede consultar el caso en que disponemos de un número infinito de sucesos. Véase bibliografía].

En muchos experimentos aleatorios pueden surgir complicaciones técnicas cuando queremos asignar probabilidades a los diferentes sucesos, dado que, por ejemplo, puede no ser siempre posible considerar todos los subconjuntos de Ω como sucesos, pues podría pasar que descartemos o cometamos errores en medir alguna información que sea el resultado de alguno de los experimentos realizados, $\omega \in \Omega$, de modo que para un subconjunto A de Ω, pueda no ser posible dar una respuesta SÍ o NO a la pregunta ¿está ω en A?

De aquí surge la importancia y la necesidad, como se manifiesta en el axioma I de Kolmogorov, de considerar una clase particular F de subconjuntos de Ω que, por razones de consistencia matemática, requeriremos que sea un álgebra (σ-álgebra, en el caso infinito). Así, por ejemplo, si la pregunta ¿ocurrió A_n? tiene una respuesta definida, para n = 1, 2, 3, ... k, también tendrán respuesta definida las preguntas: ¿ocurrió al menos uno de los A_n? y ¿ocurrieron todos los A_n?

Ahora, aplicando los axiomas de Kolmogorov, sí estaríamos en condiciones de asignar probabilidades a los elementos de *F*. Si, $A \in F$ la probabilidad *P(A)* debería reflejar la frecuencia relativa de del suceso *A* en un gran número de repeticiones independientes del experimento. Por tanto, *P(A)* debería ser un número entre 0 y 1, y $P(\Omega)$ debería ser 1.

Consecuentemente, definimos la probabilidad, con nuestra notación actual y para el caso en que Ω sea finito, como una función $P: F \rightarrow R$ que asigna un número *P(A)* a cada conjunto *A* del álgebra *F*, y que satisface las siguientes condiciones (obsérvese la total analogía entre estos tres axiomas y las propiedades de las frecuencias relativas):

1. $P(A) \geq 0, \quad \forall A \in F.$
2. $P(\Omega) = 1.$
3. Si $A_1, \ldots, A_n$ son conjuntos disjuntos de *F*, entonces:

$$P(A_1 \cup \ldots \cup A_n) = P(A_1) + \ldots + P(A_n).$$

Consideraciones finales

La dificultad básica con las definiciones *frecuentista* y *clásica* es que se encaminan a tratar de probar matemáticamente que, por ejemplo, la probabilidad de sacar un "oro" de una baraja española de 40 cartas perfectamente barajadas es $\frac{1}{4}$. Y esto no se puede hacer. Todo lo que podemos decir es que, si cogemos una carta de manera aleatoria y luego la reemplazamos, y repetimos el proceso una y otra vez, la relación entre el número de oros que se obtiene y el número total de extracciones estará próxima a $\frac{1}{4}$, de acuerdo a nuestra intuición y nuestra experiencia física. Por tanto, deberíamos asignar una probabilidad de $\frac{1}{4}$ al suceso de obtener un "oro". De esta manera, las consecuencias estarían de acuerdo con nuestra experiencia. Y si consideramos que existe alguna razón misteriosa por la que el "Rey de copas" sale más veces que cualquier otra carta, entonces podríamos incorporar este factor asignando una mayor probabilidad al "Rey de copas". Y así, el desarrollo matemático de la teoría no se vería afectado, aunque las conclusiones que obtengamos puedan discordar con la experiencia.

Nunca podemos usar realmente las matemáticas para probar un hecho específicamente físico. Por ejemplo, no podemos probar matemáticamente que existe una cantidad física llamada "velocidad". Lo que podemos hacer es postular una entidad matemática, que podemos llamar "velocidad", que satisfaga una cierta ecuación diferencial; es decir, podemos construir una colección de resultados matemáticos que, interpretados de una manera apropiada, dan una descripción razonable de un cierto fenómeno físico. Igualmente, en teoría de la probabilidad nos podemos encontrar con situaciones que nos sugieren ciertos resultados. La intuición y experiencia nos sirven para asignar probabilidades a los sucesos. Nuestra esperanza es desarrollar resultados matemáticos que, cuando son interpretados y relacionados con la experiencia física, nos ayudarán a hacer precisas ideas como "la relación entre el número de oros que se obtiene y el número total de extracciones, en un muy gran número de extracciones independientes, estará próxima a $\frac{1}{4}$".

Consecuentemente, la definición axiomática de Kolmogorov no descarta las definiciones frecuentista y clásica, sino más bien las valida y les confiere una forma más precisa.

Bibliografía

Ash, Robert B.: *Basic Probability Theory*. Wiley, 1970.

Gnedenko, B.: *Teoría de las probabilidades*. Euro Omega-Rubiños, 1995.

Kolmogorov, A. N.: *Foundations of the Theory of Probability*. Chelsea Publishing Company, 1950.

37