EL PROBLEMA DE LA MEDIDA

Miguel Ángel García Álvarez

Consideremos el experimento aleatorio consistente en la elección al azar de un punto en el intervalo $[0,1]$. Si cualquier subconjunto del espacio muestral representa un evento, la construcción de un modelo matemático para este experimento consiste en la definición de una función de probabilidad P definida sobre la familia $\mathcal{A}$ formada por todos los subconjuntos del intervalo $[0,1]$. La condición de que la elección se realiza al azar determina de manera única la probabilidad de cada subintervalo del intervalo $[0,1]$, la cual es igual a la longitud de dicho subintervalo. Sin embargo, aún no tenemos el modelo completo pues no hemos asignado una probabilidad a cada subconjunto del intervalo $[0,1]$. Por otra parte, la función de probabilidad que queremos definir debe de tener la propiedad de que dos subconjuntos del intervalo $[0,1]$ que se puedan sobreponer geométricamente uno sobre el otro deben de tener asignada la misma probabilidad (ésta sería una manera de interpretar que la elección se realiza al azar).

En el año 1902, Henri Léon Lebesgue se planteó un problema similar al que queremos resolver. De manera específica, bautizó como **el problema de la medida** al consistente en encontrar una función m definida sobre todos los subconjuntos acotados de números reales y satisfaciendo las siguientes condiciones:

1. m es no negativa.
2. m es σ-aditiva.
3. $m([0,1]) = 1$
4. m es invariante bajo traslaciones.

Este problema fue atacado por diferentes personas, en particular por Giuseppe Vitali, quien demostró en 1905 que el problema de la medida de Lebesgue no tiene solución.

El resultado de Vitali provocó algo de insatisfacción pues de lo que se trata es de definir una función, llamada medida, que permita extender el concepto de longitud de un intervalo a todos los subconjuntos acotados de números reales, de tal manera que se satisfagan ciertas propiedades que razonablemente podrían esperarse para esta función. El resultado de Vitali planteó una disyuntiva, o bien se aceptan como razonables las condiciones *i*, *ii*, *iii* y *iv* y entonces se restringe la familia de conjuntos a los cuales se les puede asignar una medida, o bien se buscan condiciones menos restrictivas para la función m de tal manera que pueda definirse para cualquier subconjunto acotado de números reales. Cabe aclarar que, para mostar la existencia de conjuntos no medibles, Vitali utilizó el axioma de elección.

En la búsqueda de alternativas para el problema de la medida, otros matemáticos lo estudiaron considerando diferentes modificaciones. Los resultados en este sentido son los siguientes:

Felix Hausdorf propuso en 1914 el problema de la medida en sentido amplio, el cual consiste en encontrar una función m definida sobre todos los subconjuntos acotados de números reales y que satisfaga las siguientes condiciones:

1. m es no negativa.
2. m es finitamente aditiva.

3. $m([0,1]) = 1$
4. m es invariante bajo traslaciones.

Stefan Banach demostró en 1923 que este problema, así como el problema análogo en dos dimensiones, sí tiene solución. Sin embargo, el mismo Hausdorf había demostrado que el problema análogo en tres o más dimensiones no la tiene.

Banach y Kazimierz Kuratowski se plantearon en 1929 el problema de encontrar una función m definida sobre todos los subconjuntos el intervalo $[0,1]$ y satisfaciendo las siguientes condiciones:

1. m es no negativa.
2. m es σ-aditiva.
3. Si I es un intervalo, entonces $m(I)$ es igual a la longitud de I.

El resultado de Banach y Kuratowski fue que tal problema no tiene solución. Esto sorprendió a muchos, por ejemplo a Paul Pierre Lévy, quien pensaba que, al quitar a la medida la condición de ser invariante bajo traslaciones, es posible asignar una medida a todos los subconjuntos de números reales.

Alfred Tarski atacó en 1930 el problema de la medida en sentido amplio planteándose el problema de encontrar una función m definida sobre todos los subconjuntos del intervalo $[0,1]$ y satisfaciendo las siguientes condiciones:

1. m es no negativa.
2. m es finitamente aditiva.
3. Si I es un intervalo, entonces $m(I)$ es igual a la longitud de I.

Tarski mostró que tal problema, así como el análogo en dos o más dimensiones, sí tiene solución. Sin embargo, mostró también que la solución no es única.

Volviendo a nuestro problema de asignar una probabilidad a cada subconjunto del intervalo $[0,1]$ de tal manera que dos subconjuntos del intervalo $[0,1]$ que se puedan sobreponer geométricamente uno sobre el otro tengan asignada la misma probabilidad, vemos, a la luz de los resultados anteriores, que éste no tiene solución si se quiere una función de probabilidad que sea σ-aditiva, en cambio, si únicamente se quiere una función de probabilidad finitamente aditiva, el problema sí tiene solución, aunque, aún en ese caso, el problema análogo en tres dimensiones no lo tiene. De esta manera se puede concluir que la invariancia geométrica es muy restrictiva. Sustituyendo ésta por la condición de que un intervalo tenga asignada una probabilidad igual a su longitud, vemos que, si se quiere una función de probabilidad que sea σ-aditiva, el problema tampoco tiene solución. Por otra parte, si únicamente se quiere una función de probabilidad finitamente aditiva, el problema sí tiene solución, pero ésta no es única.

La conclusión que se deriva de estos resultados es que los experimentos descritos no pueden ser modelados dentro del marco que hemos considerado hasta este momento. Por lo tanto, se debe de tomar una de dos alternativas, o bien se renuncia a modelar este tipo de experimentos o bien se cambia alguna o algunas de las características del modelo.

Hemos visto que la modelación de experimentos consistentes en elecciones al azar, de los elementos de un conjunto finito, ha jugado un papel central para poder construir modelos probabilísticos en el caso general de experimentos aleatorios que admiten únicamente un conjunto finito de posibles resultados, e incluso para aquellos que admiten un conjunto infinito numerable. La misma situación se presenta en el caso de experimentos aleatorios que admiten una infinidad no numerable de posibles resultados, la modelación de experimentos consistentes en elecciones al azar juega también un papel central en la construcción de modelos probabilísticos en general. Además, se puede mostrar que el problema planteado para el caso de elecciones al azar se presenta en prácticamente todos los casos de experimentos aleatorios en donde el conjunto de posibles resultados es infinito no numerable. Por tal motivo, la renuncia a modelar tales tipos de experimentos empobrecería mucho la teoría desarrollada. De esta manera, la alternativa que queda consiste en relajar las características que se piden a la función de probabilidad.

Es posible tomar como modelo el que incluya una función de probabilidad que sea únicamente finitamente aditiva de tal manera que, para el experimento aleatorio planteado, se pueda asignar una probabilidad a cualquier subconjunto del intervalo $[0, 1]$. Con esta opción se estaría renunciando a la unicidad en la asignación de probabilidades. Esta opción es la que se toma, por ejemplo, en la Estadística Bayesiana. Sin embargo, en el enfoque clásico de la Teoría de la Probabilidad se busca que cada evento tenga asignada una probabilidad de manera única. Esto hace que la opción de tomar una función de probabilidad que sea únicamente finitamente aditiva tampoco resulte satisfactoria.

Abundando en el problema de la unicidad, se puede mostrar que si se quiere definir una función de probabilidad P σ-aditiva, esta función queda únicamente determinada sobre una familia $\mathcal{L}$ de subconjuntos del intervalo $[0, 1]$, la cual tiene las siguientes propiedades:

1. Si I es cualquier subintervalo del intervalo $[0, 1]$, entonces $I \in \mathcal{L}$.
2. Si A es un subconjunto del intervalo $[0, 1]$ tal que $A \in \mathcal{L}$, entonces $[0, 1] - A \in \mathcal{L}$.
3. Si $A_1, A_2, \ldots$ es una colección finita, o infinita numerable, de elementos de $\mathcal{L}$, entonces $\bigcup_k A_k \in \mathcal{L}$.

Por otra parte, si se quiere definir una función de probabilidad P finitamente aditiva, se puede mostrar que esta función queda únicamente determinada sobre una familia $\mathcal{J}$ de subconjuntos del intervalo $[0, 1]$, la cual tiene las siguientes propiedades:

1. Si I es cualquier subintervalo del intervalo $[0, 1]$, entonces $I \in \mathcal{J}$.
2. Si A es un subconjunto del intervalo $[0, 1]$ tal que $A \in \mathcal{J}$, entonces $[0, 1] - A \in \mathcal{J}$.

3. Si $A_1, \ldots, A_n$ es una colección finita de elementos de $\mathcal{J}$, entonces $\bigcup_{k=1}^{n} A_k \in \mathcal{J}$.

Evidentemente la familia $\mathcal{L}$ contiene a la familia $\mathcal{J}$ pues una función de probabilidad σ-aditiva también es finitamente aditiva. Tal contención es propia. En efecto, se puede mostrar que hay subconjuntos del intervalo $[0, 1]$ que pertenecen a $\mathcal{L}$ pero no a $\mathcal{J}$. Por ejemplo, la probabilidad de un subconjunto del intervalo $[0, 1]$ que sea infinito numerable no queda determinada de manera única si únicamente se cuenta con la propiedad de la aditividad finita, en cambio, asumiendo que la función de probabilidad es σ-aditiva, la probabilidad de tal subconjunto queda únicamente determinada y es igual a cero pues los conjuntos formados por un punto tienen probabilidad igual a cero.

Vemos así que el proceso de asignación de probabilidades, de manera única, a los subconjuntos del espacio muestral se puede llevar más lejos con una función de probabilidad σ-aditiva que con una que sea simplemente finitamente aditiva. Esta es una de las razones básicas por las cuales se prefiere adoptar como modelo matemático para los experimentos aleatorios el que incluye la σ-aditividad de la función de probabilidad.