Probabilidad

Introducción. Fenómenos aleatorios

Durante siglos la Ciencia se ha constituido mediante el establecimiento de leyes que predecían que la ocurrencia o presencia de unas causas conduce invariablemente a la aparición de determinados efectos. Este esquema determinístico ha dominado, hasta fechas recientes, las Ciencias de la Naturaleza, consideradas como la única Ciencia verdadera, o al menos el arquetipo de Ciencia al cual debían amoldarse todas ellas.

Dentro de este esquema de relaciones fijas, las fluctuaciones observadas en las distintas determinaciones de una magnitud eran atribuidas a los errores de medida, los cuales eran tratados de forma separada y sin afectar a la naturaleza fija de las leyes y al carácter determinístico de las relaciones entre las verdaderas magnitudes de las variables consideradas.

Sin embargo, siempre han existido situaciones, como los juegos de azar, que escapaban a este esquema determinístico. Estos fenómenos en los que unas mismas causas podían dar origen a la ocurrencia de distintos efectos eran denominados aleatorios y mantenidos al margen de la construcción del edificio teórico de la Ciencia.

La dicotomía entre fenómenos determinísticos y aleatorios no creó, durante siglos, disonancias ya que las relaciones analíticas entre las variables parecía ser una herramienta suficiente para la construcción de las teorías científicas y las cuestiones donde claramente se producían situaciones aleatorias parecían ser materias sin interés para la Ciencia.

La incorporación al conjunto de la Ciencia de materias que se ocupaban de los aspectos sociales y humanos que comportan el estudio de fenómenos con una fuerte componente aleatoria y la aparición dentro de las Ciencias de la Naturaleza de teorías que conllevan una cierta indeterminación e incertidumbre sobre la presencia de determinados efectos, hicieron necesario el estudio de los fenómenos aleatorios y su incorporación al esquema general de la explicación científica de la Naturaleza.

La Psicología en concreto es una ciencia en la cual es difícil establecer leyes o teorías que no conlleven un componente aleatorio y ello por el propio objeto de su estudio, la conducta y los procesos que la ocasionan, se presenta a nuestros ojos siempre con cierto grado de aleatoriedad e incertidumbre que no podemos eliminar.

Dos tipos de razones se han expuesto para explicar la aleatoriedad del comportamiento de los seres vivos. Para unos la aleatoriedad de las leyes que rigen la conducta vendría dada por su propia naturaleza que es aleatoria en si misma como parte del mecanismo adaptativo que habría favorecido el comportamiento aleatorio frente a las respuestas fijas a los estímulos del medio. Para otros las relaciones entre las variables conductuales serían aleatorias por nuestra incapacidad para controlar todas las variables e interacciones que determinan la conducta. El conjunto de los efectos de todas las variables incontroladas daría origen a la componente aleatoria que nos vemos obligados a incluir en nuestros modelos.

Sea cual sea la explicación a la que nos adscribamos, debemos admitir la imposibilidad de establecer, en el momento actual, leyes determinísticas para predecir la conducta de los sujetos. De ahí la importancia que cobra para el psicólogo el estudio de los fenómenos aleatorios y las leyes probabilísticas que los rigen.

Damos a continuación las definiciones de algunos de los términos que utilizaremos:

Situaciones aleatorias

Son aquellas en que las mismas condiciones o causas, pueden dar lugar a diferentes resultados o efectos.

Experimentos aleatorios

Son aquellos en que se presentan situaciones aleatorias, por ejemplo el lanzamiento de un dado es una experiencia que cada vez que se realiza puede dar origen a un resultado diferente, aunque aparentemente se haya ejecutado en la misma forma.

Sucesos elementales

Son cada uno de los posibles resultados del experimento aleatorio, indescomponibles en otros más simples. Por ejemplo, al lanzar un dado que salga la cara marcada con el 5, es un suceso elemental.

Espacio muestral

Es el conjunto formado por todos los posibles sucesos elementales. En el caso del lanzamiento de un dado ordinario de seis caras el espacio muestral estaría constituido por los seis sucesos elementales, es decir que salga un uno, que salga un dos, etc. Simbólicamente lo representamos

$$E = \{ \{1\} \ \{2\} \ \{3\} \ \{4\} \ \{5\} \ \{6\} \}$$

Suceso

Un suceso es cualquier subconjunto del espacio muestral. Por ejemplo el suceso de que al lanzar un dado salga un número par sería el subconjunto $\{ \{2\} \{4\} \{6\} \}$. El suceso de que salga un número mayor que tres sería el subconjunto $\{ \{4\} \{5\} \{6\} \}$.

Muestra

Es el conjunto de resultados que se obtienen al repetir un cierto número de veces n, un experimento aleatorio. Por ejemplo si lanzamos un dado diez veces y obtenemos los siguientes resultados (6, 5, 1, 5, 3, 3, 3, 1, 4, 1) esta secuencia constituye una muestra de tamaño diez del lanzamiento de un dado.

Frecuencia

Se define la frecuencia de un suceso en una muestra, como el cociente entre el número de veces que ha ocurrido el suceso en la muestra y el tamaño de la misma. Por ejemplo, en la muestra anterior la frecuencia del suceso de que salga un cinco sería: $2/10 = 0{,}2$. Análogamente la frecuencia del suceso de que salga un número mayor que tres sería: $4/10 = 0{,}4$.

Se ha observado empíricamente que al ir aumentando el tamaño de una muestra, la frecuencia de los sucesos tiende a estabilizarse alrededor de un número fijo, a este hecho se le ha denominado ley de estabilidad de las frecuencias o ley única del azar y ese número ideal, límite que alcanzaría la frecuencia de un suceso si obtuviésemos una muestra infinita del experimento, fue el primer concepto de probabilidad de un suceso que se manejó.

Esta interpretación, llamada frecuentista, sigue siendo la más ampliamente utilizada y la que ha inspirado el desarrollo formal de la teoría del cálculo de probabilidades. Desde un punto de vista formal, la probabilidad de un suceso va a ser una medida del subconjunto del espacio muestral que corresponde a dicho suceso, indicadora de la posibilidad de ocurrencia de tal suceso. Las interpretaciones de esa medida constituirán los distintos modelos probabilísticos de la realidad.

Operaciones con sucesos

Para poder construir esta medida que es la probabilidad, necesitamos dotar al espacio de los sucesos de una determinada estructura y para ello vamos a definir distintas operaciones y relaciones con los sucesos.

Sea ***S*** un experimento aleatorio al cual se le asocia una familia de sucesos $\{ A, B, C, \ldots \}$ tales que fijado uno de ellos, por ejemplo él A, al realizar una prueba particular del experimento aleatorio ***S***, podemos decir, al observar el resultado de la prueba, si se ha verificado, o no, el suceso A.

Implicación

Decimos que el suceso A implica el B, y lo notamos $A \subseteq B$ o $B \supseteq A$, si siempre que ocurre A ocurre B. Por ejemplo, en el lanzamiento de un dado el suceso de que salga un dos $\{2\}$, implica el suceso de que salga un número par $\{ 2, 4, 6 \}$, pues obviamente siempre que sale un dos sale un número par.

Igualdad

Si el suceso A implica el B, $A \subseteq B$, y también el suceso B implica él A, $B \subseteq A$, entonces decimos que el suceso A y el suceso B son iguales $A = B$.

Suceso contrario

Definimos el suceso contrario de A, $\tilde{A}$ como el suceso que ocurre cuando no ocurre A. Por ejemplo, en el lanzamiento de un dado el suceso contrario de que salga un número par, es el suceso de que salga un número impar, pues siempre que no sale un número par sale uno impar.

Unión de sucesos

Se define la unión de dos sucesos A y B, como el suceso C que se verifica siempre que se verifican A o B y se nota $C = A \cup B$.

Suceso seguro

Es evidente que la unión de un suceso y su contrario se verificará siempre, a este suceso le llamamos el suceso seguro y lo designaremos por E.

$$E = A \cup \tilde{A}$$

Intersección de sucesos

Se define la intersección de dos sucesos A y B como el suceso D, que ocurre siempre que ocurre él A y el B y lo notamos $D = A \cap B$.

Sucesos incompatibles

Dos sucesos A y B se dicen incompatibles si al verificarse uno de ellos, no se verifica el otro. Por ejemplo, al lanzar un dado son incompatibles los sucesos de que salga un número par y que salga un tres.

Suceso imposible

Es obvio que la intersección de dos sucesos incompatibles no se verificará nunca, por ello a este suceso intersección se le denomina suceso imposible y lo representaremos por ϕ.

Diferencia de sucesos

Se define la diferencia de dos sucesos A y B que se denota B-A, como el suceso que ocurre cuando ocurre B y no ocurre A, es decir:

$$B - A = B \cap \tilde{A}$$

Sigma álgebra

La probabilidad no puede ser definida sobre una colección cualquiera de sucesos, sino que la clase de subconjuntos del espacio muestral sobre la que puede establecerse una probabilidad debe tener una estructura mínima que se denomina sigma álgebra.

Una clase de sucesos $\boldsymbol{A} = \{ A_1, A_2, A_3, \ldots \}$ constituye un sigma álgebra si verifica las dos propiedades siguientes:

1. - La unión numerable de sucesos del sigma álgebra también pertenece al sigma álgebra.

$$\text{Si} \quad A_1,\ A_2,\ A_3,\ \ldots \in \boldsymbol{A} \quad \Rightarrow \quad \cup A_i \in \boldsymbol{A}$$

2.- Si un suceso pertenece al sigma álgebra entonces su contrario también pertenece.

$$\text{Si} \quad A \in \boldsymbol{A} \ \Rightarrow \ \tilde{A} \in \boldsymbol{A}$$

Es inmediato comprobar que el suceso seguro E y el suceso imposible ϕ, pertenecen a toda sigma álgebra.

Axiomas del cálculo de probabilidades

Una vez que tenemos el espacio muestral E, a partir del cual hemos construido un sigma álgebra ***A***, disponemos de una estructura sobre la cual es posible definir una medida P, a la que llamamos probabilidad. Esta probabilidad va a ser una función de conjunto que cumpla los siguientes axiomas:

<u>Axioma1</u>

A cada suceso A, perteneciente al sigma álgebra ***A***, le corresponde un número real positivo, llamado probabilidad de A y que notaremos P(A), es decir:

$$\forall A \in \boldsymbol{A} \quad \exists P(A) \in \Re \quad \text{tal que} \quad 0 \le P(A)$$

<u>Axioma2</u>

La probabilidad de la unión de un conjunto numerable de sucesos incompatibles es igual a la suma de las probabilidades de los sucesos, o dicho de otro modo, si los sucesos A_1, A_2, ... son incompatibles se verifica:

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

<u>Axioma3</u>

La probabilidad del suceso seguro E, es igual a la unidad, P(E) = 1.

Llamaremos espacio probabilístico a la terna (E, ***A***, P), es decir a un espacio muestral E, sobre el que se ha definido un sigma álgebra ***A***, donde se ha definido una función de probabilidad P.

Debe de notarse que dado un espacio muestral, es decir el conjunto de sucesos elementales asociados a los resultados de un experimento aleatorio, pueden definirse distintos espacios probabilísticos, o dicho de otro modo un mismo suceso puede tener distintas probabilidades. Esto es así porque el sistema de axiomas nos indica las condiciones que debe de cumplir la función de probabilidad, pero no da ninguna regla para asignar los valores de probabilidad a los sucesos elementales, en consecuencia las probabilidades de todos los sucesos están indeterminadas y deberá recurrirse a criterios adicionales para fijar las probabilidades de los sucesos elementales. Estos criterios pueden hacer referencia a propiedades físicas de los elementos materiales que intervienen en el experimento aleatorio, o a registros de aparición de los sucesos, o bien a elementos subjetivos acerca de la ocurrencia de los mismos.

Ahora bien, una vez que hemos fijado los valores de probabilidad de los sucesos elementales, podemos calcular las probabilidades de todos los sucesos del sigma álgebra atendiendo a los axiomas y a ciertas propiedades que como veremos a continuación, se derivan de ellos. Esta es la razón por la que esta disciplina se denomina cálculo de probabilidades, por que nos da reglas de cálculo para determinar las probabilidades de los sucesos en función de los números que se han asignado a los sucesos elementales.

Propiedades de la Función de Probabilidad

Como consecuencia de los axiomas anteriores, la probabilidad posee las siguientes propiedades:

1.- Si $\tilde{A}$ es el suceso contrario de A, entonces:

$$P(\tilde{A}) = 1 - P(A)$$

2.- Puesto que el suceso imposible ϕ es el contrario del suceso seguro, aplicando la propiedad anterior se verifica:

$$P(\phi) = 0$$

3.- Si el suceso A implica el suceso B, $A \subseteq B$ entonces:

$$P(A) \leq P(B)$$

4.- Puesto que cualquier suceso A implica el suceso seguro E entonces:

$$\forall A \quad P(A) \leq 1$$

5.- Dados dos sucesos compatibles A y B , es decir tales que $A \cap B \neq \phi$, se tiene que:

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

6.- Como consecuencia inmediata de la propiedad anterior se tiene que:

$$P(A \cup B) \leq P(A) + P(B)$$

7.- Si el conjunto de sucesos elementales es numerable podemos definir la probabilidad de un suceso A como la suma de las probabilidades de los sucesos elementales que pertenecen al suceso A, es decir:

$$P(A) = \sum_{a_i \in A} P(a_i)$$

8.- Si el suceso seguro E puede descomponerse en k sucesos elementales equiprobables, entonces la probabilidad de cualquier suceso A es igual al número h de sucesos elementales que pertenecen a A dividido por k, es decir:

$$P(A) = h/k$$

Esta propiedad constituye la clásica regla de Laplace que definía la probabilidad de un suceso como el cociente entre los casos favorables y los casos posibles. Pero debe tenerse en cuenta que esta conocida regla sólo puede aplicarse cuando el espacio muestral es finito y además se acepta el postulado de indiferencia.

Interpretación de la Probabilidad

La teoría del Cálculo de Probabilidades es una teoría matemática, que como todas las teorías matemáticas, establece una serie de reglas de operación con unos términos abstractos o símbolos. De esta forma, una afirmación en Matemáticas es verdadera o falsa si se ha deducido correctamente, o no, a partir de los axiomas o hipótesis que definen el sistema, por consiguiente no puede ser sometida a verificación empírica. Por ello, cuando utilizamos las teorías matemáticas para construir modelos de la realidad, debemos dotar de significado a los símbolos que se manejan, es decir, interpretar en términos reales los términos abstractos. De esta forma, el modelo matemático, mediante la asignación de significado, se convierte en un modelo del mundo real cuyas afirmaciones y predicciones pueden someterse a contrastación con los datos obtenidos experimentalmente.

En este sentido, la probabilidad que en términos matemáticos es una función de conjunto que asigna a cada suceso un número real, debe de dotarse de un significado, una interpretación que haga que el modelo matemático del Cálculo de Probabilidades, se transforme en un modelo de la ocurrencia de los sucesos.

<u>Interpretación frecuentista</u>

La interpretación más difundida es la frecuentista, aquella que entiende la probabilidad de un suceso como el límite de la frecuencia relativa de aparición de un suceso, al repetirse indefinidamente el experimento aleatorio. Puede verse fácilmente que las frecuencias relativas de los sucesos verifican los axiomas de la probabilidad, y por tanto constituyen una interpretación válida de la misma.

La interpretación frecuentista encuentra grandes dificultades cuando trata el problema de los casos aislados, es decir fenómenos aleatorios que ocurren una sola vez, de tal forma que el experimento no puede ser repetido en las mismas condiciones. Es claro que en esta situación, el concepto de frecuencia carece de significado.

Más aún, incluso en aquellos casos en que el experimento es susceptible de repetición, si el sujeto que tiene que juzgar la posibilidad de ocurrencia de un suceso, lo ve como un fenómeno único, en el que no volverá a estar implicado, difícilmente adecuará sus decisiones, o su comportamiento, a un concepto de probabilidad que es exacto cuando la muestra de realizaciones es infinitamente grande, pero que solo proporciona una indicación imprecisa sobre la ocurrencia del suceso, en una única realización. En estos casos la interpretación subjetiva puede ser mucho más fructífera.

Interpretación subjetiva

La idea alrededor de la cual gravita la interpretación subjetiva, es que ante un fenómeno aleatorio que puede verificarse o no, los individuos asignan, de forma implícita o explícita, una probabilidad que representa el grado de confianza que tienen en la ocurrencia de ese hecho. La diferencia fundamental con la interpretación frecuentista radica en que la asignación de valores, aunque condicionada por la información existente, es propia de cada observador particular, sin que las opiniones de varios sujetos tengan porqué coincidir. Así, mientras que en la concepción frecuentista la probabilidad es única y lo que puede variar son las estimaciones que hacemos de ella, en esta interpretación subjetiva la probabilidad no está determinada de forma previa, sino que se establece por los individuos en función de sus juicios previos acerca de la situación y de la forma en que valoran la información existente.

Para que esta interpretación pueda plasmarse en una teoría operativa, hay que exigir que los sujetos mantengan un cierto grado de racionalidad en la asignación de probabilidades, los valores que se determinen, aunque elegidos libremente, no pueden ser arbitrarios. Una forma de asegurar la racionalidad y que nuestras probabilidades subjetivas no nos hagan incurrir en situaciones contrarias a nuestros intereses, es que los grados de confianza que asignemos a la ocurrencia de los sucesos, verifiquen los axiomas de Kolmogorov.

Dentro de la probabilidad subjetiva hay distintas corrientes, como el intuicionismo que considera la probabilidad como algo directamente intuible por el individuo y previo a cualquier experiencia, o el decisionismo que trata de medir el grado de confianza, con que el sujeto espera la aparición de un suceso, a través de las decisiones que adopta en relación con la verificación de ese suceso.

Existen otras interpretaciones posibles de la probabilidad, como los distintos conceptos lógicos o relacionales, pero son de menos interés para nosotros.

Probabilidad condicionada

Sea un suceso B del sigma álgebra, tal que se verifique que $P(B) > 0$. Entendemos por "*probabilidad del suceso* A *condicionada por el suceso* B", y la notaremos $P(A / B)$, a la probabilidad de que ocurra A supuesto que ha ocurrido B. El valor de esta probabilidad viene dada por la expresión:

$$P(A / B) = \frac{P(A \cap B)}{P(B)}$$

Es inmediato ver que fijado B, la probabilidad condicionada $P(A / B)$ es una función de probabilidad, pues verifica los tres axiomas.

Sucesos dependientes e independientes

En general, ocurrirá que $P(A / B) \neq P(A)$, y en tal caso, diremos que A depende de B .

Si por el contrario, $P(A / B) = P(A)$, diremos que A es independiente de B. En tal caso sucede que:

$$P(A) = \frac{P(A \cap B)}{P(B)}$$

despejando se obtiene la siguiente expresión:

$$P(A \cap B) = P(A) \cdot P(B)$$

que puede ser utilizada como definición alternativa de sucesos independientes, o bien nos permite calcular la probabilidad de la intersección de sucesos independientes.

Debe de tenerse muy en cuenta que esta regla que nos indica que la probabilidad de la intersección de dos sucesos es igual al producto de las probabilidades de los sucesos, sólo es aplicable en el caso de que sean independientes. En el caso general, la expresión que proporciona la probabilidad de la intersección es:

$$P(A \cap B) = P(B) \cdot P(A / B)$$

Es fácil demostrar que la independencia es mutua, es decir que si A es independiente de B, también B es independiente de A.

Teorema de Bayes

Sean A_1, A_2, ... A_n , n sucesos mutuamente excluyentes que forman un sistema exhaustivo; es decir que $E = A_1 \cup A_2 \cup ... \cup A_n$, y de los cuales se saben (o pueden asignarse) sus probabilidades $P(A_i)$. Sea B un suceso para el cual se conocen las probabilidades condicionadas $P(B / A_i)$. En estas condiciones, el teorema de Bayes expresa que las $P(A_i / B)$ son:

$$P(A_i / B) = \frac{P(A_i) \cdot P(B / A_i)}{\sum_{i=1}^{n} P(A_i) \cdot P(B / A_i)}$$

En efecto, por la definición de probabilidad condicionada:

$$P(A_i / B) = \frac{P(A_i \cap B)}{P(B)}$$

pero el numerador es igual a:

$$P(A_i \cap B) = P(A_i) \cdot P(B / A_i)$$

por otra parte, el denominador es igual a:

$$P(B) \quad = \quad P(B \cap E) \quad = \quad P\left(B \cap \left(\bigcup_{i=1}^{n} A_i \right) \right)$$

y por la propiedad distributiva de la intersección respecto de la unión, tenemos que:

$$P(B) \quad = \quad P\left(\bigcup_{i=1}^{n} (B \cap A_i) \right) \quad = \quad \sum_{i=1}^{n} P(B \cap A_i)$$

por consiguiente:

$$P(B) \quad = \quad \sum_{i=1}^{n} P(A_i) \cdot P(B / A_i)$$

Sustituyendo numerador y denominador por las expresiones obtenidas, tendremos finalmente:

$$P(A_i / B) \quad = \quad \frac{P(A_i) \cdot P(B / A_i)}{\sum_{i=1}^{n} P(A_i) \cdot P(B / A_i)}$$

A las probabilidades $P(A_i)$ se les denomina *probabilidades a priori*. Se llama *verosimilitudes* a las $P(B / A_i)$ y a las $P(A_i / B)$ *probabilidades a posteriori*.

Como ha podido verse, el teorema de Bayes se obtiene como una deducción lógica, a partir de la definición de probabilidad condicionada, por lo que su expresión no puede estar sujeta a controversia, dentro del Cálculo de Probabilidades. Sin embargo, las condiciones necesarias para su empleo y su interpretación divide a los estadísticos en dos escuelas diferentes, y a menudo irreconciliables, los clásicos o frecuentistas, frente a los bayesianos.

El punto fundamental de discrepancia radica en la naturaleza de las *probabilidades a priori*, para un estadístico clásico estas probabilidades deben de ser conocidas objetivamente, a partir de unos datos empíricos, y no son modificables. Para un estadístico bayesiano las *probabilidades a priori* pueden ser objetivamente desconocidas, de forma que se les asigna valor en función de nuestro grado de conocimiento o convicciones previas; por ejemplo, cuando no tenemos ninguna información sobre la ocurrencia de los A_i se les asigna a todos ellos la misma probabilidad, dando lugar a una distribución uniforme de probabilidad, que es la menos informativa. La ocurrencia, o no ocurrencia, del suceso B proporciona evidencia en favor de uno u otro de los sucesos del sistema, de tal forma que las *probabilidades a posteriori* vendrían a sustituir a las probabilidades asignadas inicialmente.