ESPERANZA DE VARIABLES ALEATORIAS

Introducción

Uno de los conceptos básicos de la teoría de la probabilidad es el de esperanza de una variable aleatoria. Su importancia es comparable con la del concepto mismo de probabilidad de un evento. De hecho, el concepto de esperanza y el de probabilidad surgieron en forma paralela cuando, a mediados del siglo $XVII$, se inició el Cálculo de Probabilidades. Fue Christiaan Huygens quien introdujo este concepto en su libro *Du calcul dans les jeux de hasard*, publicado en el año 1657. En esa época, Blaise Pascal y Pierre de Fermat habían resuelto algunos problemas de probabilidad, con métodos que sentarían las bases para el desarrollo de una nueva disciplina matemática, el Cálculo de Probabilidades. Huygens resolvió, con sus propios métodos, los problemas que antes habían resuelto Pascal y Fermat y algunos otros más.

Uno de los aspectos interesantes de la metodología utilizada por Huygens es que todas las soluciones están basadas en el cálculo de esperanzas, lo cual hacía ver ya que este concepto podía tomarse como primario, previo incluso al de probabilidad de un evento, y a partir de él desarrollar la nueva disciplina. La historia no fue de ese modo pues el concepto que prevaleció como primario fue el de probabilidad. Sin embargo, la historia misma mostraría más adelante que esta dualidad de importancia, entre el concepto de esperanza y el de probabilidad, que se dio al inicio del desarrollo de la teoría de la probabilidad como disciplina matemática, tenía fuertes raíces pues al evolucionar el concepto de probabilidad, hasta fusionarse con el de medida en los primeros años del siglo XX, resultó palpable la estrecha relación entre ambos conceptos, de tal manera que, efectivamente, cualquiera de los dos puede tomarse como punto de partida, quedando inmersos uno dentro del otro. Esto último ya no únicamente dentro del contexto de la teoría de la probabilidad, sino dentro del contexto más amplio de la teoría de la medida, donde el concepto de probabilidad corresponde al de medida y el de esperanza al de integral con respecto a una medida.

Recordemos que si X es una variable aleatoria discreta cuyos posibles valores son $x_1, x_2, \ldots$, se define su esperanza de la siguiente manera:

$$E\left[X\right] = \sum_k x_k P\left[X = x_k\right]$$

En el año 1654 Pascal y Fermat resolvieron los siguientes dos problemas:

Problema 1. *¿Cuántas veces se necesita lanzar un par de dados para que sea más favorable obtener par de seises por lo menos en una ocasión?*

Problema 2 (Problema de la división de apuestas). *¿Cómo deben repartirse las apuestas en un juego que se interrumpe? Por ejemplo, suponiendo que dos jugadores, P y Q, apuestan* 32 *pesos cada uno en un juego que consiste de partidas consecutivas, en cada una de las cuales cada jugador tiene la misma posibilidad de ganarla y quien la gane acumula un punto, de tal manera que el juego es ganado por quien obtenga primero cuatro puntos, ¿cómo deben de repartirse las apuestas en caso de que el juego se interrumpa cuando el jugador P ha ganado dos puntos y Q un punto?*

En su libro, Huygens resolvió estos problemas con su propio método, tomando como punto de partida la siguiente hipótesis, la cual es básicamente una definición:

En un juego, la posibilidad que se tiene de ganar alguna cosa tiene un valor tal que, si se posee ese valor, se puede uno procurar la misma posibilidad en un juego equitativo.

Por un juego equitativo, Huygens entendía un juego que no va en detrimento de ninguno de los jugadores. Incluye el caso de un juego entre un número cualquiera de jugadores en el cual, o bien todos los jugadores tienen la misma posibilidad de ganar cierta cantidad, o bien cada uno de los jugadores tiene la misma posibilidad de ganar cierta cantidad que de perderla.

De su hipótesis, Huygens dedujo tres proposiciones, de las cuales, las dos primeras son un caso particular de la siguiente:

Proposición 1. *Tener iguales posibilidades de obtener a_1, a_2, a_3, ..., a_n tiene un valor de $\frac{a_1+a_2+a_3+\cdots+a_n}{n}$.*

Demostración

Llamemos P a quien tiene iguales posibilidades de obtener a_1, a_2, a_3, ..., a_n. Consideremos entonces un juego entre P y otros $n-1$ jugadores, que llamaremos Q_2, Q_3, Q_4, ..., Q_n, en el cual los n tienen las mismas posibilidades de ganar y cada uno apuesta $\frac{a_1+a_2+a_3+\cdots+a_n}{n}$. Quien gane se lleva todas las apuestas, es decir, $a_1 + a_2 + a_3 + \cdots + a_n$. Evidentemente, éste es un juego equitativo. Si, además, para cada $k \in \{2, 3, \ldots, n\}$, P acuerda con Q_k que, si alguno de los dos gana el juego, el ganador le dará al otro la cantidad a_k. Cada uno de estos acuerdos es equitativo y no va en detrimento de ninguno de los otros jugadores; así que el juego continúa siendo equitativo. Si P gana el juego, una vez que entrega lo acordado con cada uno de los otros jugadores, obtiene a_1.Si el juego lo gana Q_k, P obtiene a_k. Así que P tiene iguales posibilidades de obtener a_1, a_2, a_3, ..., a_n, en un juego equitativo.

■

Proposición 2. *Tener r posibilidades de obtener a y s posibilidades de obtener b, las posibilidades siendo equivalentes, tiene un valor de $\frac{ra+sb}{r+s}$.*

Demostración

Llamemos P a quien tiene r posibilidades de obtener a y s posibilidades de obtener b. Consideremos entonces un juego entre P y otros $r+s-1$ jugadores, en el cual los $r+s$ tienen las mismas posibilidades de ganar y cada uno apuesta $\frac{ra+sb}{r+s}$. Quien gane se lleva todas las apuestas, es decir, $ra+sb$. Evidentemente, éste es un juego equitativo. Supongamos, además, que P acuerda con cada uno de s jugadores que, si alguno de los dos gana el juego, el ganador le dará al otro la cantidad b; llamaremos a este conjunto de s jugadores, el grupo 1. Finalmente, supongamos también que, con cada uno del resto de los jugadores, P acuerda que, si alguno de los dos gana el juego, el ganador le dará al otro la cantidad a; llamaremos a este conjunto de $r-1$ jugadores, el grupo 2. Cada uno de estos acuerdos es equitativo y no va en detrimento de ninguno de los otros jugadores; así que el juego continúa siendo equitativo. Si P gana el juego, una vez que entrega lo acordado con cada uno de los otros jugadores, obtiene a. Si el juego lo gana algún jugador del grupo 1, P obtiene b, mientras que si lo gana alguno del grupo 2, obtiene a. Como los $r+s$ jugadores tienen las mismas posibilidades de ganar, P tiene entonces r posibilidades de obtener a y s posibilidades de obtener b, en un juego equitativo.

■

Podemos ver que la hipótesis de Huygens es básicamente la definición del concepto de esperanza de lo que se obtiene sobre lo que está en juego; es decir, la esperanza de una variable aleatoria definida como lo que obtiene el jugador.

El estudio de problemas con dados lo comenzó Huygens haciendo ver que con un dado se pueden obtener 6 resultados diferentes, todos igualmente posibles porque se supone que el dado tiene la forma de un cubo perfecto. Como puede verse, aquí está la idea de equiprobabilidad. En seguida hacía ver que con dos dados se pueden obtener $6 \times 6 = 36$ resultados diferentes, también todos igualmente posibles; con tres dados se tienen $36 \times 6 = 216$ y así sucesivamente. Por último, encontró el número de resultados, de entre todos los posibles, con los cuales se obtiene cada una de las diferentes sumas que pueden obtenerse con dos y tres dados.

Para resolver el problema 1 Huygens razonó de la manera siguiente:

Llamemos A a la apuesta.

Paso 1. Quien juega a un sólo lanzamiento (de dos dados) tiene 1 posibilidad de ganar A y 35 posibilidades de no ganar nada, así que le corresponde $\frac{1}{36}A$.

Paso 2. Quien juega a dos lanzamientos tiene 1 posibilidad de ganar A y 35 posibilidades de obtener $\frac{1}{36}A$ (por el paso 1), lo que le vale:

$$\frac{A+(35)\left(\frac{1}{36}A\right)}{36}=\frac{71}{(36)^2}A$$

Lo que se obtiene como valor es una fracción $\frac{r}{s}$ multiplicada por la apuesta A. Lo que decía Huygens es que esa fracción es un cociente de resultados favorables entre el total de posibles resultados, es decir, es una probabilidad.

Es decir, en el razonamiento de Huygens está implícita la idea de que sus proposiciones 1 y 2 pueden expresarse no sólo en términos de esperanzas sino también en términos de probabilidades. En otras palabras, dentro del concepto de Esperanza que definió Huygens está contenido implícitamente el concepto clásico de probabilidad. De aquí se refuerza la idea que hemos dado ya en el sentido de que dentro del concepto de esperanza que definió Huygens está contenido implícitamente el concepto clásico de probabilidad.

Paso 3. Quien juega a cuatro lanzamientos, gana A si obtiene par de seises en los primeros dos lanzamientos, si no lo obtiene, gana, por el paso 2, $\frac{71}{(36)^2}A$; pero, también por el paso 2, hay 71 posibilidades de obtener par de seises en los dos primeros lanzamientos y $(36)^2-71=1,225$ posibilidades de no obtenerlo; por lo tanto, en 4 lanzamientos le corresponde:

$$\frac{71A+(1{,}225)\left(\frac{71}{(36)^2}A\right)}{(36)^2}=\frac{178{,}991}{(36)^4}A$$

Paso 4. Quien juega a ocho lanzamientos, gana A si obtiene par de seises en los primeros cuatro lanzamientos, si no lo obtiene, gana, por el paso 3, $\frac{178{,}991}{(36)^4}A$; pero, también por el paso 3, hay $178,991$ posibilidades de obtener par de seises en los cuatro primeros lanzamientos y $(36)^4-178,991=1,500,625$ posibilidades de no obtenerlo; por lo tanto, en 8 lanzamientos le corresponde:

$$\frac{178{,}991A+1{,}500{,}625\left(\frac{178{,}991}{(36)^4}A\right)}{(36)^4}=\frac{569{,}234{,}516{,}831}{(36)^8}A$$

Continuando este procedimiento, se llega a que quien juega a 24 lanzamientos le corresponde

$$\frac{11{,}033{,}126{,}465{,}283{,}976{,}852{,}912{,}127{,}963{,}392{,}284{,}191}{(36)^{24}}A\approx 0{,}4914A.$$

Finalmente, quien juega a veinticinco lanzamientos, gana A si obtiene par de seises en el primer lanzamiento, si no lo obtiene, gana, por el paso anterior:

$$\frac{11{,}033{,}126{,}465{,}283{,}976{,}852{,}912{,}127{,}963{,}392{,}284{,}191}{(36)^{24}}A$$

Pero, tiene 1 posibilidad de obtener par de seises en el primer lanzamiento y 35 posibilidades de no obtenerlo; por lo tanto, en 25 lanzamientos le corresponde

$$\frac{A+35\left(\frac{11{,}033{,}126{,}465{,}283{,}976{,}852{,}912{,}127{,}963{,}392{,}284{,}191}{(36)^{24}}A\right)}{36}$$

$$= \frac{408{,}611{,}683{,}992{,}293{,}747{,}092{,}011{,}689{,}842{,}522{,}621{,}501}{(36)^{25}}A \approx 0{,}5055A$$

Así que, quien juega a 24 lanzamientos tiene todavía una ligera desventaja, mientras que quien juega a 25 lanzamientos tiene una ligera ventaja.

Se desconoce si Huygens realizó alguna simplificación para llegar a la solución del problema planteado o si hizo los cálculos anteriores de la manera descrita, lo cual, en su época, era bastante laborioso. Una posible simplificación del método de Huygens es la siguiente:

Paso 1. Quien juega a un sólo lanzamiento (de dos dados) tiene 1 posibilidad de ganar A y 35 posibilidades de no ganar nada, así que,le corresponde:

$$\tfrac{1}{36}A = \tfrac{36-35}{36}$$

Paso 2. Quien juega a dos lanzamientos tiene $36 - 35$ posibilidades de ganar A y 35 posibilidades de obtener $\frac{36-35}{36}A$ (por el paso 1), lo que le vale:

$$\frac{(36-35)A+(35)\left(\frac{36-35}{36}A\right)}{36} = \frac{(36)^2-(35)^2}{(36)^2}A$$

Paso 3. Quien juega a tres lanzamientos, tiene $36 - 35$ posibilidades de ganar A y 35 posibilidades de obtener $\frac{(36)^2-(35)^2}{(36)^2}A$ (por el paso 2), lo que le vale:

$$\frac{(36-35)A+(35)\left(\frac{(36)^2-(35)^2}{(36)^2}A\right)}{(36)^2} = \frac{(36)^3-(35)^3}{(36)^3}A$$

Continuando con este procedimiento se llega a que, lo que le vale el juego a quien juega a n lanzamiento está dado por:

$$\frac{(36)^n-(35)^n}{(36)^n}A = \left[1 - \left(\tfrac{35}{36}\right)^n\right]A$$

Para $n = 24$ y $n = 25$, se obtiene, respectivamente:

$$1 - \left(\tfrac{35}{36}\right)^{24} \approx 0{,}4914$$

$$1 - \left(\tfrac{35}{36}\right)^{25} \approx 0{,}5055$$

Este mismo resultado puede obtenerse de una manera más simple, utilizando el mismo método de Huygens, considerando lo que corresponde al contrario en lugar de lo que corresponde al jugador. Este cálculo sería como sigue:

En un lanzamiento, el contrario tiene 35 posibilidades de obtener A y 1 posibilidad de no obtener nada; le corresponde entonces:

$$\tfrac{35}{36}A$$

En dos lanzamientos tiene 1 posibilidad de no obtener nada (si el jugador obtiene par de seises en el primer lanzamiento) y 35 posibilidades de obtener $\frac{35}{36}A$ por el primer paso; le corresponde entonces:

$$\frac{35\left(\frac{35}{36}A\right)}{36} = \left(\tfrac{35}{36}\right)^2 A$$

Con el mismo razonamiento se encuentra entonces que en 3 lanzamientos le corresponde $\left(\frac{35}{36}\right)^3 A$; en 4 lanzamientos $\left(\frac{35}{36}\right)^4 A$ y así sucesivamente; es decir, en n lanzamientos le corresponde $\left(\frac{35}{36}\right)^n A$. Por lo tanto, al jugador que trata de obtener par de seises en n lanzamientos le corresponde:

$$A - \left(\tfrac{35}{36}\right)^n A = \left[1 - \left(\tfrac{35}{36}\right)^n\right] A$$

Este hecho muestra que no siempre la solución más simple es la primera que se ocurre e incluso puede no ser evidente. Lo inmediato o simple de una solución a un problema requiere, a veces, de ensayos de solución y de maduración de determinados conceptos.

Haciendo uso de sus proposiciones, Huygens resolvió el problema 2 para algunos casos particulares, siguiendo un método recursivo (como lo había hecho antes Pascal).

Primero observó, al igual que Fermat y Pascal, que la repartición de las apuestas debe estar en función de los puntos que le faltan a cada jugador para ganar el juego.

Llamando A al total de las apuestas, cuando al jugador P le falta una partida para ganar y a Q dos, al jugar la siguiente partida hay dos posibilidades, la primera es que P la gane, en cuyo caso gana el juego y por lo tanto le corresponde toda la apuesta, la segunda es que Q la gane en cuyo caso P y Q quedan en igualdad de condiciones y debe entonces tocar a cada uno la mitad de las apuestas, es decir $\frac{1}{2}A$. Entonces en el primer caso a P le corresponde A y en el segundo, $\frac{1}{2}A$.

Por lo tanto, si el juego se detiene en las condiciones dadas inicialmente, a P le corresponde, por la proposición 1, $\frac{A+\frac{1}{2}A}{2} = \frac{3}{2^2}A$; y a Q, $\frac{0+\frac{1}{2}A}{2} = \frac{1}{2^2}A$.

Cuando a P le falta un punto y a P tres, si se juega la siguiente partida, P puede ganar toda apuesta o bien $\frac{3}{2^2}A$ por el primer caso. Por lo tanto a P le corresponde $\frac{A+\frac{3}{2^2}A}{2} = \frac{7}{2^3}A$ y a Q $\frac{1}{2^3}A$.

Cuando a P le faltan dos puntos y a Q tres (que son las condiciones del problema 2), si se juega la siguiente partida, P puede quedar faltándole un punto y tres a P, en cuyo caso le corresponde $\frac{7}{2^3}A$ por el segundo caso; o bien, si Q gana esta partida, quedan en igualdad de circunstancias y toca a cada uno $\frac{1}{2}A$. Por lo tanto a P le corresponde $\frac{\frac{7}{2^3}A+\frac{1}{2}A}{2} = \frac{11}{2^4}A$ y a P, $\frac{5}{2^4}A$.

Así que, si el juego se interrumpe cuando el jugador P ha ganado dos puntos y Q un punto, a P le corresponde $\frac{11}{2^4}(64) = 44$ y a P, 20.

Huygens resolvió el problema análogo cuando se trata de 3 jugadores. Por ejemplo, supongamos que 3 personas están jugando a cierto número de partidas y que a la primera le falta una partida para ganar, una también a la segunda y dos a la tercera. Utilizando la proposición 1 diríamos:

Si el juego continúa y la siguiente partida es ganada por la primera persona, entonces a ésta le corresponde toda la apuesta, llamémosla A, y a las otras no les toca nada; si es la segunda persona la que la gana, entonces a ésta le toca A y a las otras nada; finalmente, si es la tercera persona las que la gana entonces quedan en igualdad de circunstancias, tocándole a cada una $\frac{1}{3}A$. Por lo tanto, si el juego se detiene en las condiciones dadas inicialmente, a la primera persona le corresponde, por la proposición 1, $\frac{A+0+\frac{1}{3}A}{3} = \frac{4}{9}A$; a la segunda lo mismo y a la tercera $\frac{0+0+\frac{1}{3}A}{3} = \frac{1}{9}A$.

La proposición 2 se puede utilizar para resolver el problema de las partidas en el caso en que cada jugador tenga distintas posibilidades de ganar cada partida, sin embargo Huygens no hizo esta aplicación. También la proposición 2 es inmediatamente generalizable al caso de cualquier número n de cantidades $a_1, \ldots, a_n$ con posibilidades iguales a $r_1, \ldots, r_n$, respectivamente, en cuyo caso el valor del juego resultaría $\frac{r_1a_1+\cdots+r_na_n}{r_1+\cdots+r_n}$; esta generalización tampoco la hizo Huygens.

Observemos por último que el método de Huygens involucra implícitamente el concepto de Esperanza Condicional.

Consideremos por ejemplo un juego a cuatro puntos que se detiene cuando P ha ganado dos puntos y Q uno. Nos preguntamos entonces por la proporción de las apuestas que corresponde a P, o bien, partiendo de la situación dada, por el valor del juego para P.

Sea a el total de las apuestas y x el valor del juego para P partiendo de la situación dada. Si supiéramos cuantas posibilidades tienen P y Q de ganar, partiendo de la situación dada, obtendríamos fácilmente lo que le corresponde a cada uno, pues si r es el primer número y s el segundo, entonces, por la proposición 2, al jugador P le corresponde $\frac{ra}{r+s}$ y al jugador Q $\frac{sa}{r+s}$. Es decir, el valor del juego para P es $x = \frac{ra}{r+s}$.

Como r y s no se conocen, lo que hacemos es, partiendo de la situación dada, suponer que se juega la siguiente partida, entonces llamando x_1 al valor del juego para P cuando P gana la siguiente partida y x_2 su valor cuando Q gana la siguiente partida, usando la proposición 1, obtenemos $x = \frac{x_1+x_2}{2}$.

Utilizando la terminología moderna, estamos considerando los eventos:

A_1: El jugador P gana la siguiente partida.

B_1: El jugador Q gana la siguiente partida.

Entonces, x es el valor del juego para P (i.e. el valor que espera recibir P), mientras que x_1 y x_2 son los valores del juego para P condicionados a la ocurrencia de A_1 y B_1 respectivamente (i.e. esperanzas condicionales). De manera que si llamamos X a lo que recibe P en este juego, la fórmula $x = \frac{x_1+x_2}{2}$ puede escribirse en la forma:

$$E\left[X\right] = E\left[X \mid A_1\right] P(A_1) + E\left[X \mid B_1\right] P(B_1)$$

la cual podríamos llamar la regla de la probabilidad total para esperanzas.

Formulación moderna

En la formulación moderna de la teoría de la probabilidad, cuando se tiene un fenómeno aleatorio, se trata de modelarlo mediante una terna $(\Omega, \Im, P)$, donde Ω es un conjunto, el cual, para un problema práctico, está formado por todas observaciones posibles que puedan hacerse de dicho fenómeno; $\Im$ es una familia de subconjuntos de Ω, cada uno de los cuales es llamado un evento; P es una función definida sobre $\Im$, la cual nos dice cuál es la probabilidad de ocurrencia de cada uno de los eventos.

La familia $\Im$ debe tener las siguientes propiedades:

a) $\Omega \in \Im$.

b) $\Im$ es cerrada bajo complementos.

c) $\Im$ es cerrada bajo uniones infinitas numerables.

Cuando se tienen estas 3 propiedades se dice que $\Im$ forma una $\sigma-$álgebra de subconjuntos de Ω.

La función P debe tener las siguientes propiedades:

a) $P(A) \geq 0$ para cualquier $A \in \Im$.

b) $P(\emptyset) = 0$.

c) P es $\sigma-$aditiva; es decir, si $(A_n)_{n\in\mathbb{N}}$ es una sucesión de elementos de $\Im$, ajenos por parejas, entonces:

$$P\left(\cup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

d) $P(\Omega) = 1$.

Cuando se tienen las 3 primeras propiedades se dice que P es una medida. Si además se cumple la cuarta, se dice que P es una medida de probabilidad.

A una terna $(\Omega, \Im, P)$ que satisface las propiedades anteriores se le llama un **espacio de probabilidad**.

Asumiendo que ya se tiene definido un espacio de probabilidad, se define una variable aleatoria real como una función $X : \Omega \to \mathbb{R}$ tal que:

$$[X \leq x] = \{\omega \in \Omega : X(\omega) \leq x\} \in \Im \text{ para cualquier } x \in \mathbb{R}$$

Si X es una variable aleatoria real, se define su función de distribución $F_X : \mathbb{R} \to \mathbb{R}$ de la siguiente manera:

$$F_X(x) = P[X \leq x]$$

Cualquier función de distribución F_X satisface las siguientes propiedades:

a) F_X es una función no decreciente.

b) $\lim_{x\to\infty} F_X(x) = 1$.

c) $\lim_{x\to-\infty} F_X(x) = 0$.

d) F_X es continua por la derecha en cualquier punto $x \in \mathbb{R}$.

Al ser F_X una función no decreciente, el conjunto de puntos donde es discontinua es a lo más infinito numerable y todas sus discontinuidades son de salto.

Si $\{x_1, x_2, \ldots\}$ es el conjunto de puntos donde F_X es discontinua, se dice que la variable aleatoria X es discreta si se cumple la siguiente relación:

$$\sum_k P[X = x_k] = 1$$

En este caso, F_X crece únicamente mediante saltos y se define la función $f_X : \mathbb{R} \to [0, \infty)$, llamada la función de densidad de X, de la siguiente manera:

$$f_X(x) = P[X = x]$$

Si F_X es continua en cualquier punto $x \in \mathbb{R}$, se dice que la variable aleatoria X es continua.

Si X es una variable aleatoria continua, con función de distribución F_X, se dice que X es absolutamente continua si existe una función $f_X : \mathbb{R} \to [0, \infty)$, llamada función de densidad de X, tal que:

$$F_X(x) = \int_{-\infty}^{x} f_X(y)\, dy$$

En los cursos y libros de probabilidad de nivel licenciatura, se define la esperanza de una variable aleatoria únicamente para el caso discreto y el absolutamente continuo.

En el caso discreto, si la serie $\sum_{x\in\mathbb{R}} |x| f_X(x)$ es convergente, se define:

$$E[X] = \sum_{x\in\mathbb{R}} x f_X(x)$$

Mientras que en el caso absolutamente continuo, si la integral $\int_{-\infty}^{\infty} |x| f_X(x)\, dx$ es finita, se define:

$$E[X] = \int_{-\infty}^{\infty} x f_X(x)\, dx$$

Esto constituye una limitación y lleva a enunciar resultados que, rigurosamente, están mal planteados. Por ejemplo, Si X es una variable aleatoria absolutamente continua y $g : \mathbb{R} \to \mathbb{R}$ es una función tal que $g(X) = g \circ X$ es una variable aleatoria, se calcula su esperanza de la siguiente manera:

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx$$

El problema con esta fórmula consiste en que la variable aleatoria $g(X)$ podría no ser ni discreta, ni absolutamente continua y entonces su esperanza no estaría definida.

Otro problema que se presenta al contar únicamente con las definiciones anteriores de la Esperanza de una variable aleatoria consisten en que algunas demostraciones se complican bastante; por ejemplo, simplemente el demostrar que $E[X+Y] = E[X] + E[Y]$, para una pareja de variables aleatorias X y Y, no es una trivialidad en el caso absolutamente continuo ya que para evaluar $E[X+Y]$ se requeriría encontrar una función de densidad de $X+Y$.

Para trabajar con el concepto de esperanza es necesario contar con una definición que se aplique a cualquier variable aleatoria y no únicamente a las discretas o a las absolutamente continuas.

Es precisamente por problema de este tipo, y otros más relevantes, que resulta más conveniente y riguroso trabajar con la teoría de la probabilidad utilizando la teoría de la medida.

Definición general de la Esperanza

En esta parte asumimos que tenemos definido un espacio de probabilidad $(\Omega, \Im, P)$.

Como queremos tratar con el paso al límite, se hace necesario trabajar con el conjunto de los números reales extendido, agregándole a $\mathbb{R}$ dos elementos, ∞ y $-\infty$, con las siguientes convenciones:

Si $c \in \mathbb{R}$, entonces:

$-\infty < c < \infty$

$c - \infty = -\infty$

$c + \infty = \infty$

$c\,(\infty) = \infty$ si $c > 0$

$c\,(\infty) = -\infty$ si $c < 0$

$(0)\,(\infty) = (0)\,(-\infty) = 0$

$\frac{c}{\infty} = \frac{c}{-\infty} = 0$

$(\infty)\,(\infty) = \infty + \infty = \infty$

$\infty - \infty$ e $\frac{\infty}{\infty}$ no están definidos.

Al conjunto de los números reales extendido lo denotaremos por $\overline{\mathbb{R}}$.

Con este agregado, una variable aleatoria será entonces una función $X : \Omega \to \overline{\mathbb{R}}$ tal que $[X \leq x]$ para cualquier $x \in \overline{\mathbb{R}}$.

También, es conveniente completar el espacio de probabilidad $(\Omega, \Im, P)$ de tal manera que cualquier subconjunto de un conjunto de probabilidad cero sea parte de la $\sigma-$álgebra $\Im$.

Hay un resultado que resulta central para trabajar con variables aleatorias en general, el cual asegura que cualquier variable aleatoria no negativa se puede aproximar mediante una sucesión no decreciente de variables aleatorias discretas no negativas, cada una de las cuales toma valores en un conjunto finito.

Definición 1. *Diremos que una variable aleatoria discreta es simple si el conjunto de sus posibles valores es finito.*

Si $\{a_1, \ldots, a_n\}$ es el conjunto formado por todos los distintos posibles valores no nulos de una variable aleatoria simple φ y, para $k \in \{1, \ldots, n\}$, definimos:

$$E_k = \{\omega \in \Omega : \varphi(\omega) = a_k\}$$

entonces los conjuntos $E_1, \ldots, E_n$ pertenecen a la $\sigma-$ágebra $\Im$, son ajenos por parejas y $\varphi = \sum_{k=1}^{n} a_k I_{E_k}$. Esta última expresión será llamada la representación canónica de φ.

Teorema 1. *Sea $X : \Omega \to \mathbb{R}$ una variable aleatoria no negativa, entonces existe una sucesión no decreciente de variables aleatorias simples no negativas $\varphi_n : \Omega \to \mathbb{R}$ tales que $\lim_{n\to\infty} \varphi_n(\omega) = X(\omega)$ para cualquier $\omega \in \Omega$.*

Demostración

Para cada $n \in \mathbb{N}$, definamos:

$$\varphi_n(\omega) = \begin{cases} \frac{m-1}{2^n} & \text{si } \frac{m-1}{2^n} \leq X(\omega) < \frac{m}{2^n} \text{ y } m \in \{1, 2, \ldots, n2^n\} \\ n & \text{si } X(\omega) \geq n \end{cases}$$

Si $X(\omega) = \infty$, entonces $\varphi_n(\omega) = n$ para cualquier $n \in \mathbb{N}$, así que:

$\varphi_n(\omega) \leq \varphi_{n+1}(\omega)$ y $\lim_{n\to\infty} \varphi_n(\omega) = \infty = X(\omega)$.

Si $X(\omega) < n$ para alguna $n \in \mathbb{N}$, sea m el único número natural tal que $\frac{m-1}{2^n} \leq X(\omega) < \frac{m}{2^n}$. Entonces, como $\varphi_n(\omega) = \frac{m-1}{2^n}$, se tiene:

$X(\omega) - \frac{1}{2^n} < \varphi_n(\omega) \leq X(\omega)$, así que $\lim_{n\to\infty} \varphi_n(\omega) = X(\omega)$.

Ahora bien, como $\frac{2(m-1)}{2^{n+1}} \leq X(\omega) < \frac{2m}{2^{n+1}}$, se tiene que:

$$\frac{2m-2}{2^{n+1}} \leq X(\omega) < \frac{2m-1}{2^{n+1}}$$

o bien:

$$\frac{2m-1}{2^{n+1}} \leq X(\omega) < \frac{2m}{2^{n+1}}.$$

En el primer caso, se tiene:

$$\varphi_{n+1}(\omega) = \frac{2m-2}{2^{n+1}} = \frac{m-1}{2^n} = \varphi_n(\omega)$$

En el segundo, se tiene:

$$\varphi_{n+1}(\omega) = \frac{2m-1}{2^{n+1}} > \frac{2m-2}{2^{n+1}} = \varphi_n(\omega)$$

Así que, en cualquier caso, $\varphi_n(\omega) \leq \varphi_{n+1}(\omega)$.

Así que, φ_n es una sucesión no decreciente de variables aleatorias simples no negativas tales que $\lim_{n\to\infty} \varphi_n(\omega) = X(\omega)$ para cualquier $\omega \in \Omega$.

$\blacksquare$

Definición 2. *Si φ es una variable aleatoria simple no negativa con representación canónica $\varphi = \sum_{k=1}^{n} a_k I_{E_k}$, se define la integral de φ, $\int_\Omega \varphi dP$, de la siguiente manera:*

$$\int_\Omega \varphi dP = \sum_{k=1}^{n} a_k P(E_k)$$

Definición 3. *Si X es una variable aleatoria no negativa, se define la integral de X, con respecto a la medida P, $\int_\Omega X dP$, de la siguiente manera:*

$$\int_\Omega X dP = \sup \left\{ \int_\Omega \varphi dP : \varphi \text{ es una variable aleatoria simple y } 0 \leq \varphi \leq X \right\}$$

Obsérvese que el conjunto

$$\left\{ \int_\Omega \varphi dP : \varphi \text{ es una variable aleatoria simple y } 0 \leq \varphi \leq X \right\}$$

no necesariamente está acotado, de manera que estamos considerando que, en ese caso, el supremo del conjunto lo definimos como ∞.

Definición 4. *Si X es una variable aleatoria no negativa, se define su integral sobre un conjunto $E \in \Im$, $\int_E X dP$, de la siguiente manera:*

$$\int_E X dP = \int_\Omega X I_E dP$$

Definición 5. *Si X es una variable aleatoria no negativa, definimos la* ***esperanza*** *de X, $E[X]$, de la siguiente manera:*

$$E[X] = \int_\Omega X dP$$

Definición 6. *Diremos que la variable aleatoria X tiene esperanza finita si $E[|X|] < \infty$. En ese caso definimos su esperanza de la siguiente manera:*

$$E[X] = E[X^+] - E[X^-]$$

donde X^+ y X^- son la parte positiva y negativa de X, respectivamente.

Con esta definición de la esperanza de una variable aleatoria se pueden demostrar las siguientes propiedades:

Proposición 3. *Sean X y Y dos variables aleatorias no negativas, entonces:*

a) $E[aX + bY] = aE[X] + bE[Y]$ para cualesquiera números reales a y b no negativos.

b) Si $X \leq Y$ entonces $E[X] \leq E[Y]$.

c) $E[X] = 0$ si y sólo si $P\{\omega \in \Omega : X(\omega) > 0\} = 0$.

d) Si $E[X] < \infty$, entonces X es finita excepto a lo más en un conjunto de probabilidad cero.

Proposición 4. *Sea X una variable aleatoria de esperanza finita, entonces:*

a) Para cualquier número real c, cX tiene esperanza finita.

b) $|E[X]| \leq E[|X|]$

Si X y Y son dos variables aleatorias, $X+Y$ podría no estar definida para algunos elementos $\omega \in \Omega$ ya que toman valores en $\overline{\mathbb{R}}$; pero, si tanto X como Y tienen esperanza finita, entonces $X+Y$ está bien definida, excepto a lo más en un conjunto de probabilidad cero. En este caso, haremos la convención de definir $X(\omega) + Y(\omega) = 0$ cuando $X(\omega) + Y(\omega)$ no esté definida. Esto no altera de valor de las esperanzas de X y Y ya que son finitas excepto a lo más en un conjunto de probabilidad cero.

Proposición 5. *Sean X y Y dos variables aleatorias de esperanza finita, entonces:*

a) $E[X + Y] = E[X] + E[Y]$

b) Si $E[XI_E] \geq 0$ para cualquier conjunto $E \in \Im$, entonces:

$$P\{\omega \in \Omega : X(\omega) < 0\} = 0$$

c) Si $E[XI_E] \leq E[YI_E]$ para cualquier conjunto $E \in \Im$, entonces:

$$P\{\omega \in \Omega : X(\omega) > Y(\omega)\} = 0$$

d) Si $E[XI_E] \leq E[YI_E]$ para cualquier conjunto $E \in \Im$, entonces:

$$P\{\omega \in \Omega : X(\omega) \neq Y(\omega)\} = 0$$

e) Si $X_1, \ldots, X_n$ son n variables aleatorias independientes de esperanza finita, entonces $\prod_{k=1}^{n} X_k$ también tiene esperanza finita y:

$$E\left[\prod_{k=1}^{n} X_k\right] = \prod_{k=1}^{n} E\left[X_k\right]$$

Además, se tienen los teoremas de convergencia, los cuales permiten un manejo más simple de la Esperanza:

Teorema de la convergencia monótona. Si $(X_n)_{n\in\mathbb{N}}$ es una sucesión no decreciente de variables aleatorias no negativas, entonces:

$$E\left[\text{lím}_{n\to} X_n\right] = \text{lím}_{n\to\infty} E\left[X_n\right]$$

Teorema de la convergencia dominada. Si Y es una variable aleatoria no negativa de esperanza finita y $\{X_n\}_{n\in\mathbb{N}}$ es una sucesión de variables aleatorias tales que $|X_n| \leq Y$ y $\text{lím}_{n\to\infty} X_n$ existe excepto a lo más en un conjunto de probabilidad cero y X es una variable aleatoria tal que $X = \text{lím}_{n\to\infty} X_n$, donde este límite existe, entonces, X tiene esperanza finita y:

$$E\left[X\right] = \text{lím}_{n\to\infty} \text{lím}_{n\to\infty} E\left[X_n\right]$$

Finalmente, la definición de Esperanza dada en lo expuesto anteriormente es la que se utiliza para definir el concepto de esperanza condicional de una variable aleatoria X de esperanza finita dado que se conocen los valores de una familia de variables aleatorias $\{Y_\gamma : \gamma \in \Gamma\}$, donde Γ es un conjunto cualquiera, finito o infinito, concepto que es fundamental en la teoría de los procesos estocásticos.