INTEGRACIÓN ESTOCÁSTICA
Sesión 03/16
Construcción de Espacios de Probabilidad

Introducción

La Teoría de la Probabilidad surgió del planteamiento de problemas teóricos, los cuales provenían de algún problema práctico. El problema de la división de apuestas, por ejemplo, surgió al buscar determinar cómo deberían de repartirse las apuestas en un juego de azar que se interrumpe antes de que alguno de los participantes gane el juego de acuerdo con las reglas establecidas.

Sin embargo, una vez que la Teoría de la Probabilidad se formula en forma axiomática, los elementos que la componen no requieren de una interpretación práctica. Más aún, en la formulación axiomática, no es admisible definir un concepto en términos de un determinado fenómeno aleatorio, o demostrar algún teorema utilizando propiedades de un fenómeno aleatorio.

Es posible que se utilice la Teoría de la Probabilidad para modelar algún fenómeno aleatorio y que los conceptos que se introduzcan o las propiedades que se demuestren provengan de las características de dicho fenómeno, o que estemos interesados en estudiar las propiedades del fenómeno en consideración basándonos en el modelo matemático, pero, si bien éste puede ser el caso, los elementos mismos del fenómeno o sus propiedades no forman parte del cuerpo teórico; lo que se observa del fenómeno podría motivar introducir algún concepto o buscar algún resultado dentro del modelo matemático que nos ayude a entender el fenómeno o darnos una idea de sus propiedades, pero las observaciones mismas no forman parte del modelo.

Por ejemplo, el lanzamiento de un dado n veces consecutivas lo podemos modelar utilizando como espacio muestral al conjunto de todos los posibles resultados de los n lanzamientos, es decir, al conjunto Ω formado por todas las colecciones ordenadas de n números naturales que pertenecen al conjunto $\{1, 2, 3, 4, 5, 6\}$; la familia de eventos la podemos considerar como el conjunto potencia de Ω, es decir la familia formada por todos los subconjuntos de Ω; finalmente, como medida de probabilidad podemos tomar la que asigna a cada subconjunto de Ω el cociente que resulta de dividir el número de elementos de ese subconjunto entre 6^n, el cual es el número de elementos de Ω.

Tendríamos así definido nuestro espacio de probabilidad, el cual podemos pensarlo como modelo matemático del lanzamiento de n dados en forma consecutiva, pero el modelo mismo es una abstracción, no requerimos de referirnos al lanzamiento del dado para definirlo.

Podríamos decir: un ejemplo de espacio de probabilidad es el siguiente: Definamos $\Omega = \{1,2,3,4,5,6\}$, $\Im$ como la familia de todos los subconjuntos de Ω y como medida de probabilidad tomemos a la función $P : \Im \to \mathbb{R}$ definida, para cada $B \in \Im$, mediante la relación $P(B) = \sum_{\{\omega \in B\}} p(\omega)$, donde $p(\omega) = \frac{1}{6^n}$ para cualquier $\omega \in \Omega$.

El espacio de probabilidad $(\Omega, \Im, P)$ así definido está formado por elementos matemáticos abstractos que se tendrían que tratar como independientes de cualquier fenómeno aleatorio.

En algunos problemas de probabilidad se hace referencia al concepto de experimento aleatorio y se construye un espacio de probabilidad ad hoc para ese experimento, de manera similar a como lo hicimos con el lanzamiento del dado, pero nuevamente, el espacio de probabilidad que se construye tiene que tratarse como independiente de cualquier fenómeno aleatorio.

Un caso particular de experimento aleatorio es lo que se conoce como ensayo de Bernoulli, el cual se define como un experimento aleatorio que admite únicamente dos posibles resultados, a uno de los cuales se le llama éxito y al otro fracaso. Utilizando este concepto, podemos, por ejemplo, definir algunas variables aleatorias de interés o plantearnos algunos problemas de probabilidad que historicamente fueron importantes para el desarrollo de la Teoría de la Probabilidad.

Por, ejemplo, podemos considerar al experimento aleatorio que consiste en la realización consecutiva de n ensayos de Bernoulli, cada uno de ellos independiente de los demás y tal que la probabilidad de obtener éxito es igual a un número $p \in [0,1]$, el cual es el mismo para cada uno de los ensayos. Si definimos X como el número de éxitos que se obtienen al realizar los n ensayos, X resulta ser una variable aleatoria con distribución binomial de parámetros n y p, lo cual podemos mostrar sin necesidad de definir formalmente algún espacio de probabilidad.

Como otro ejemplo, podemos considerar al experimento aleatorio que consiste en la realización consecutiva de una infinidad de ensayos de Bernoulli, cada uno de ellos independiente de los demás y tal que la probabilidad de obtener éxito es igual a un número $p \in [0,1]$, el cual es el mismo para cada uno de los ensayos. Si definimos Y como el número de fracasos que se obtienen antes de obtener éxito por primera vez al realizar el experimento, Y resulta ser una variable aleatoria con distribución geométrica de parámetro p, lo cual, como en el caso anterior, podemos mostrar sin necesidad de definir formalmente algún espacio de probabilidad.

Como tercer ejemplo, podemos, nuevamente, considerar al experimento aleatorio que consiste en la realización consecutiva de una infinidad de ensayos de Bernoulli, cada uno de ellos independiente de los demás y tal que la probabilidad de obtener éxito es igual a un número $p \in [0,1]$, el cual es el mismo para cada uno de los ensayos. Pero esta vez podemos preguntarnos por la probabilidad de obtener un número finito de éxitos al realizar el experimento.

En el caso del primer ejemplo, no hay problema para tratarlo sin necesidad de recurrir a la formulación axiomática de la teoría de la probabilidad. Pero, en el segundo ejemplo, y sobre todo en el tercero, podemos observar un problema si no se recurre a una formulación matemática (abstracta). El problema consiste en que la realización de una infinidad de ensayos de Bernoulli es imposible. En el segundo ejemplo podría argumentarse que no hay tal problema ya que casi con seguridad se obtendría el primer éxito en un número finito de ensayos; sin embargo, sí hay problema, ya que siendo independiente cada ensayo de los demás, en cualquiera de ellos es posible que el resultado sea fracaso; de manera que es posible que no podamos determinar el número de fracasos que se obtienen antes de obtener éxito por primera vez. Podríamos obtener la distribución de X imaginando que es posible la realización del experimento, pero estaríamos partiendo de algo que es falso. En el tercer ejemplo es completamente claro que no es posible determinar si se obtiene un número finito de éxitos al realizar el experimento ya que para determinarlo es necesario conocer la infinidad de resultados que se obtienen, lo cual no es posible.

Lo anterior muestra uno de los problemas a los que se enfrenta uno al no contar con un modelo matemático abstracto que permita formular los problemas de una manera distinta, dentro de un marco teórico donde no haya necesidad de realizar experimentos. Esto tiene relación con el planteamiento de Poincaré cuando, en el año 1896, decía: "No se puede dar una definición satisfactoria de la probabilidad." y agregaba: "La definición completa de la probabilidad es una especie de petición de principio... deberemos, en cada aplicación, hacer convenciones." Obsérvese que Poincaré incluye la frase "en cada aplicación", lo cual nos dice que un problema de probabilidad lo pensaba vinculado a un determinado problema práctico. Es decir, la teoría y la aplicación estaban mezcladas en una sola cosa. Esto ilustra una de las razones por las cuales el Cálculo de Probabilidades no era considerado en esa época como una rama de las Matemáticas, sino de la Física, tal como lo expresó Hilbert en el año 1900: "Las investigaciones sobre los principios fundamentales de la geometría nos conducen a plantear este problema: Tratar con base en ese modelo las ramas de la Física donde las Matemáticas juegan actualmente un papel preponderante; esas ramas de la ciencia son, antes que cualesquiera otras, el Cálculo de Probabilidades y la Mecánica."

Con la formulación axiomática de la Teoría de la Probabilidad, la teoría y las aplicaciones quedan separadas, aunque entrelazadas de alguna manera. La teoría puede ser desarrollada independientemente de las aplicaciones que se hagan de ella, sin quedar éstas por fuera completamente ya que son las aplicaciones las que hacen surgir y alimentan las teorías, complementándose unas con la otras.

Construcción de espacios de probabilidad

Definición 1 (**Espacio de probabilidad**). *Llamaremos espacio de probabilidad a una terna* $(\Omega, \Im, P)$*, donde* Ω *es un conjunto,* $\Im$ *una* σ*-álgebra de subconjuntos de* Ω *y* P *una medida sobre* $\Im$ *tal que* $P(\Omega) = 1$*, a la cual llamaremos medida de probabilidad. A* Ω *lo llamaremos el espacio muestral, a los elementos de* $\Im$ *eventos y a la medida* P *de un evento* A *la probabilidad de* A*.*

En la formulación moderna de la teoría de la probabilidad, el primer problema a resolver cuando se tiene un problema consiste en modelarlo mediante un espacio de probabilidad. Una vez hecho eso, se trabaja con ese modelo utilizando las propiedades que tiene la medida de probabilidad; esto último ya es puramente matemático; lo que se infiera a través del modelo puede o no ser válido en el problema que queremos resolver.

Como dijimos antes, cuando se tiene un problema de probabilidad se busca definir un experimento aleatorio que lo represente; el cual tiene como característica principal que, si bien admite diferentes posibles resultados, el que se obtendría al realizarlo es sólo uno de ellos.

En general, el conjunto Ω se toma como el conjunto de posibles resultados del experimento aleatorio; sin embargo, esto no tiene que ser así necesariamente, aunque sí se requiere que, de alguna manera, Ω contenga a ese conjunto de posibles resultados.

Dependiendo de la complejidad del problema que queremos atacar, el conjunto Ω puede ser muy simple o muy complicado. Por ejemplo, si el experimento aleatorio consiste en elegir al azar un elemento del conjunto $A = \{1, 2, \dots, N\}$, donde N es un número natural fijo, el conjunto Ω podría ser precisamente el conjunto A. Otro ejemplo, si el experimento aleatorio consiste en medir día a día, durante 30 días, la altura de un bambú leñoso, en etapa de crecimiento, que mide al inicio x_0 metros, el espacio muestral Ω podría ser el conjunto

$$\{(h_1, h_2, \dots, h_{30}) : h_k \in [x_0, x_0 + 25] \text{ para cualquier } k \in \{1, 2, \dots, 30\}\}$$

La construcción de espacios de probabilidad con características que permitan su fácil manejo es un problema importante a resolver, sobre todo en la teoría de los procesos estocásticos.

Pasemos a la construcción de un espacio de probabilidad para cada uno de los 3 ejemplos descritos en la sección anterior.

Cuando se trata de problemas que tienen que ver con la realización de ensayos de Bernoulli, lo que se hace es introducir lo que se conoce como distribución Bernoulli. Para esto se representa un éxito con el número 1 y un fracaso con el número 0. Entonces decimos que una variable aleatoria X tiene distribución Bernoulli, con parámetro $p \in [0, 1]$, si $P[X = 1] = p$ y $P[X = 1] = 1 - p$.

Al considerar al experimento aleatorio que consiste en la realización consecutiva de n ensayos de Bernoulli, cada uno de ellos independiente de los demás y tal que la probabilidad de obtener éxito es igual a un número $p \in [0,1]$, tenemos que construir un espacio de probabilidad $(\Omega, \Im, P)$ en el cual se puedan definir n variables aleatorias independientes, cada una con distribución Bernoulli de parámetro p. Este espacio podría ser el siguiente:

Sea $\Omega_n = \{(s_1, s_2, \ldots, s_n) : s_j \in \{0,1\} \text{ para cualquier } j \in \{1, 2, \ldots, n\}\}$

y, para cada $\omega = (s_1, \ldots, s_n) \in \Omega_n$ y cada subconjunto A de Ω, definamos:

$$p_n(\omega) = \prod_{j=1}^{n} [ps_j + (1-p)(1-s_j)]$$

$$P(A) = \sum_{\{\omega \in A\}} p_n(\omega)$$

Para hacer ver que P es una medida de probabilidad definida sobre el conjunto potencia de Ω basta con demostrar que $\sum_{\{\omega \in \Omega_n\}} p_n(\omega) = 1$, lo cual se prueba a continuación.

Lema 1. *Para cada $n \in \mathbb{N}$, sea*

$$\Omega_n = \{(s_1, s_2, \ldots, s_n) : s_j \in \{0,1\} \text{ para cualquier } j \in \{1, 2, \ldots, n\}\}$$

Si $p \in [0,1]$ y $\omega = (s_1, \ldots, s_n) \in \Omega_n$, definamos:

$$p_n(\omega) = \prod_{j=1}^{n} [ps_j + (1-p)(1-s_j)]$$

Entonces, $\sum_{\{\omega \in \Omega_n\}} p_n(\omega) = 1$.

Demostración

Para cada $n \in \mathbb{N}$, definamos $S(n) = \sum_{\{\omega \in \Omega_n\}} p_n(\omega)$. Entonces,

$$\Omega_1 = \{(0), (1)\}$$

Por lo tanto:

$$S(1) = p \times 0 + (1-p)(1-0) + p \times 1 + (1-p)(1-1) = (1-p) + p = 1$$

Además, para cualquier $n \in \mathbb{N}$, se tiene:

$S(n+1) = [p \times 0 + (1-p)(1-0) + p \times 1 + (1-p)(1-1)]\, S(n) = S(n)$

Así que, por el principio de inducción matemática, $S(n) = 1$ para cualquier $n \in \mathbb{N}$.

■

Definamos ahora, para cada $j \in \{1, 2, \ldots, n\}$, $X_j : \Omega_n \to \mathbb{R}$ mediante la relación:

$$X_j\left((s_1, s_2, \ldots, s_n)\right) = s_j$$

Entonces, para cualquier $k \in \{1, 2, \ldots, n\}$, dados $r_1, r_2, \ldots, r_k \in \{0, 1\}$, se tiene:

$P\left(\bigcap_{j=1}^{k}[X_j = r_j]\right)$

$= \sum_{\{(s_1,s_2,\ldots,s_n)\in\Omega_n : s_j = r_j \text{ para cualquier } j\in\{1,\ldots,k\}\}} \prod_{j=1}^{n}[ps_j + (1-p)(1-s_j)]$

$= \left(\prod_{j=1}^{k}[pr_j + (1-p)(1-r_j)]\right)$

$\left(\sum_{\{(s_{k+1},\ldots,s_n): s_j\in\{0,1\} \text{ para cualquier } j\in\{k+1,\ldots,n\}\}} \prod_{j=k+1}^{n}[ps_j + (1-p)(1-s_j)]\right)$

$= \prod_{j=1}^{k}[pr_j + (1-p)(1-r_j)]$

Así que, si $r \in \{0, 1\}$, se tiene:

$P[X_k = r]$

$= \sum_{\{(s_1,s_2,\ldots,s_{k-1}): s_j\in\{0,1\} \text{ para cualquier } j\in\{1,\ldots,k-1\}\}} P[X_1 = s_1, X_2 = s_2, \ldots, X_{k-1} = s_{j-1}, X_k = r]$

$= \sum_{\{(s_1,s_2,\ldots,s_{k-1}): s_j\in\{0,1\} \text{ para cualquier } j\in\{1,\ldots,k-1\}\}} [pr + (1-p)(1-r)] \prod_{j=1}^{k-1}[pr_j + (1-p)(1-r_j)]$

$= [pr + (1-p)(1-r)] \sum_{\{(s_1,s_2,\ldots,s_{k-1}): s_j\in\{0,1\} \text{ para cualquier } j\in\{1,\ldots,k-1\}\}} \prod_{j=1}^{k-1}[pr_j + (1-p)(1-r_j)]$

$= pr + (1-p)(1-r).$

Por lo tanto, para cualquier $k \in \{1, 2, \ldots, n\}$, X_k tiene distribución Bernoulli con parámetro p.

Además, para cualquier $(s_1, s_2, \ldots, s_n) \in \Omega_n$, se tiene:

$$P\left(\bigcap_{j=1}^{n}[X_j = s_j]\right) = \prod_{j=1}^{n}[ps_j + (1-p)(1-s_j)] = \prod_{j=1}^{n} P[X_j = s_j]$$

Así que las variables aleatorias $X_1, X_2, \ldots, X_n$ son independientes.

En este espacio de probabilidad podemos definir una variable aleatoria X con distribución binomial de parámetros n y p; en efecto, la variable aleatoria $X : \Omega_n \to \mathbb{R}$, definida mediante la relación $X = \sum_{j=1}^{n} X_j$, tiene distribución binomial con parámetros n y p.

Tenemos así construido un espacio de probabilidad para el primer ejemplo.

Para formular los otros dos ejemplos dentro del marco teórico de que disponemos, requerimos construir un espacio de probabilidad $(\Omega, \Im, P)$ en el cual se pueda definir una infinidad numerable de variables aleatorias independientes, cada una con distribución Bernoulli de parámetro p. Vamos a mostrar esta construcción para el caso $p = \frac{1}{2}$. En el caso general la construcción es similar, pero requiere de algunos pasos adicionales,

Primero un resultado que usaremos más adelante:

Teorema 1. *Sea $(\Omega, \Im, P)$ un espacio de probabilidad y $X_1, X_2, \ldots$ una sucesión de variables aleatorias independientes, definidas sobre ese espacio, cada una de ellas con distribución Bernoulli de parámetro $\frac{1}{2}$. Definamos $X : \Omega \mapsto \mathbb{R}$ mediante la relación $X = \sum_{k=1}^{\infty} \frac{X_k}{2^k}$. Entonces, X es una variable aleatoria con distribución uniforme en el intervalo $[0, 1]$.*

Demostración

X es el límite de la sucesión no decreciente de variables aleatorias $X_n = \sum_{k=1}^{n} \frac{X_k}{2^k}$, así que es ella misma una variable aleatoria.

Obsérvese que si vemos cada sucesión $(X_k(\omega))$ como el desarrollo en base 2 de un número real en el intervalo $[0, 1]$, $X(\omega)$ es precisamente ese número real.

Si $(s_n)_{n\in\mathbb{N}}$ es una sucesión de 0's y 1's, se tiene:

$$P\left[X_k = s_k \text{ para toda } k \in \mathbb{N}\right] = 0$$

y si $x \in (0, 1]$ es un racional diádico, x tiene exactamente dos desarrollos en base 2, así que $P[X = x] = 0$. Además, $x = 0$ tiene únicamente un desarrollo en base 2, así que, también, $P[X = 0] = 0$. Por lo tanto, $P[X = x] = 0$ para cualquier racional diádico $x \in [0, 1]$.

Como ya lo mencionamos, para los números reales $x \in (0, 1]$ de la forma $x = \frac{j}{2^n}$, donde $n \in \mathbb{N}$ y $j \in \{1, 2, \ldots, 2^n - 1\}$, es decir, para los racionales diádicos, el desarrollo en base 2 no es único. Para cada uno de estos puntos, elijamos como desarrollo en base 2 a la sucesión $s_1, s_2, \ldots$ para la cual existe $N \in \mathbb{N}$ tal que $s_k = 1$ para cualquier $k \in \{N + 1, N + 2, \ldots\}$.

Por otra parte, para cada $n \in \mathbb{N}$, consideremos un intervalo de la forma $\left(\frac{j-1}{2^n}, \frac{j}{2^n}\right]$, con $j \in \{1, \ldots, 2^n\}$. Asociada a tal intervalo existe una única colección de 0's y 1's, $(s_1, s_2, \ldots, s_n)$, tal que un punto x pertenece al intervalo $\left(\frac{j-1}{2^n}, \frac{j}{2^n}\right]$ si y sólo si tiene un desarrollo en base 2 de la forma $x = 0.s_1 s_2 \ldots s_n \ldots$. Por lo tanto:

$$P\left[X \in \left(\tfrac{j-1}{2^n}, \tfrac{j}{2^n}\right]\right] = P\left[X_1 = s_1, X_2 = s_2, \ldots, X_n = s_n\right] = \tfrac{1}{2^n}$$

Así que, si $k \in \{1, \ldots, 2^n\}$, se tiene:

$$P\left[X \in \left(0, \tfrac{k}{2^n}\right]\right] = \sum_{j=1}^{k} P\left[X \in \left(\tfrac{j-1}{2^n}, \tfrac{j}{2^n}\right]\right] = \sum_{j=1}^{k} \tfrac{1}{2^n} = \tfrac{k}{2^n}$$

Es decir, si $x \in (0,1]$ es un racional diádico, se tiene:

$$P\left[X \in (0,x]\right] = x$$

Además, la función $F_X : [0,1] \to \mathbb{R}$, definida por

$$F_X(x) = P\left[X \in [0,x]\right]$$

es continua por la derecha.

Por lo tanto:

$$F_X(x) = x \text{ para cualquier } x \in [0,1]$$

Así que X tiene distribución uniforme en el intervalo $[0,1]$.

■

En lo que sigue vamos a tomar como espacio de probabilidad a la terna $((0,1],\mathcal{L},P)$, donde P es la medida de Lebesgue en el intervalo $(0,1]$.

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

$$\mathbb{B}_n=\{(s_1,s_2,\ldots,s_n):s_j\in\{0,1\}\ \text{ para cualquier } j\in\{1,2,\ldots,n\}\}$$

donde $n\in\mathbb{N}$.

Sabemos que el desarrollo en base 2 de un número real $x\in(0,1]$ es único y, consiste de una sucesión $(s_n)_{n\in\mathbb{N}}$ tal que $s_k=1$ para una infinidad de índices k.

Recordemos que que si $(s_1,s_{2,},\ldots,s_n)\in\mathbb{B}_n$ y $x\in(0,1]$, entonces, $s_1,s_{2,},\ldots,s_n$ son los primeros n términos del desarrollo de x en base 2 si y sólo si

$$x\in\left(\sum_{i=1}^{n}\frac{s_i}{2^i},\sum_{i=1}^{n}\frac{s_i}{2^i}+\frac{1}{2^n}\right]$$

Además cada uno de los 2^n intervalos de la forma

$$\left(\sum_{i=1}^{n}\frac{s_i}{2^i},\sum_{i=1}^{n}\frac{s_i}{2^i}+\frac{1}{2^n}\right]$$

donde $(s_1,s_{2,},\ldots,s_n)\in\mathbb{B}_n$, es alguno de los intervalos de la forma $\left(\frac{k-1}{2^n},\frac{k}{2^n}\right]$, donde $k\in\{1,2,\ldots,2^n\}$, los cuales constituyen una partición del intervalo $(0,1]$.

Teorema 2. *Para cada $n \in \mathbb{N}$ definamos $X_n : (0,1] \to \mathbb{R}$ mediante la relación $X_n(x) = s_n$, donde s_n es el n-simo término del desarrollo en base 2 de x. Entonces, las variables aleatorias de la familia $\{X_n : n \in \mathbb{N}\}$ son independientes y cada una de ellas tiene distribución Bernoulli con parámetro $p = \frac{1}{2}$.*

Demostración

Si $r \in \{0,1\}$, se tiene:

$P[X_n = r] = \sum_{\{(r_1,r_2,\ldots,r_{n-1})\in\mathbb{B}_{n-1}\}} P[X_1 = r_1, X_2 = r_2, \ldots, X_{n-1} = r_{n-1}, X_n = r]$

$= \sum_{\{(r_1,r_2,\ldots,r_{n-1})\in\mathbb{B}_{n-1}\}} P\left(\left(\sum_{k=1}^{n-1} \frac{r_k}{2^k} + \frac{r}{2^n}, \sum_{k=1}^{n-1} \frac{r_k}{2^k} + \frac{r}{2^n} + \frac{1}{2^n}\right]\right)$

$= \sum_{\{(r_1,r_2,\ldots,r_{n-1})\in\mathbb{B}_{n-1}\}} \frac{1}{2^n} = \frac{1}{2}$

Por lo tanto, para cualquier $n \in \mathbb{N}$, X_n tiene distribución Bernoulli con parámetro $p = \frac{1}{2}$.

Además, para cualquier $m \in \mathbb{N}$ y $(r_1, r_2, \ldots, r_m) \in \mathbb{B}_m$, se tiene:

$$P\left(\bigcap_{j=1}^{m} [X_j = r_j]\right) = P\left(\left(\sum_{k=1}^{m} \frac{r_k}{2^k}, \sum_{k=1}^{m} \frac{r_k}{2^k} + \frac{1}{2^m}\right]\right) = \frac{1}{2^m} = \prod_{j=1}^{m} P[X_j = r_j]$$

Así que las variables aleatorias $X_1, X_2, \ldots, X_m$ son independientes.

Por lo tanto, las variables aleatorias de la familia $\{X_n : n \in \mathbb{N}\}$ son independientes y cada una de ellas tiene distribución Bernoulli con parámetro $p = \frac{1}{2}$.

$\blacksquare$

A partir de lo anterior podemos demostrar los siguientes resultados:

Recordemos que si U y X son variables aleatorias reales, definidas sobre un espacio de probabilidad $(\Omega, \Im, P)$, se dice que U tiene distribución uniforme en un intervalo $[a, b]$, donde $a < b$, si su función de distribución $F_U : \mathbb{R} \to [0, 1]$ está dada por:

$$F_X(x) = \begin{cases} 0 & \text{si } x < a \\ \frac{x-a}{b-a} & \text{si } x \in [a, b] \\ 1 & \text{si } x \geq b \end{cases}$$

Obsérvemos que, como F_X es una función continua, se tiene:

$$P[X \in [a, b]] = F_X(b) - F(a) = \frac{b-a}{b-a} - \frac{a-a}{b-a} = 1$$

$$P[X = x] = 0 \text{ para cualquier } x \in \mathbb{R}$$

En particular:

$$P[X \notin (a, b)] = 0$$

Por otra parte, se dice que X tiene distribución normal estándar si su función de distribución $F_X : \mathbb{R} \to [0, 1]$ está dada por:

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy$$

Teorema 3. *Se puede definir, sobre* $((0,1], \mathcal{L}, P)$*, una sucesión de variables aleatorias independientes, cada una de ellas con distribución uniforme en el intervalo* $[0,1]$*.*

Demostración

Consideremos una sucesión $X_1, X_2, \ldots$ de variables aleatorias independientes, definidas sobre $((0,1], \mathcal{L}, P)$, cada una de ellas con distribución Bernoulli de parámetro $\frac{1}{2}$.

De acuerdo con la primera proposición, si $(X_{n_k})_{k\in\mathbb{N}}$ es cualquier subsucesión de la sucesión $(X_n)_{n\in\mathbb{N}}$ y definimos $Y = \sum_{k=1}^{\infty} \frac{X_{n_k}}{2^k}$, entonces Y es una variable aleatoria con distribución uniforme en el intervalo $[0,1]$.

Para definir una sucesión de variables aleatorias independientes, definidas sobre $((0,1], \mathcal{L}, P)$, cada una de ellas con distribución uniforme en el intervalo $[0,1]$, basta con mostrar que existe una infinidad numerable de subsucesiones de $(X_n)_{n\in\mathbb{N}}$ de tal manera que cualquier par de ellas no tengan elementos en común. Una vez mostrado esto, cada una de esas subsucesiones genera una distribución uniforme en el intervalo $[0,1]$.

Existen diferentes maneras de tomar las subsucesiones de $(X_n)_{n\in\mathbb{N}}$ con la propiedad mencionada. Por ejemplo, si $\{p_1, p_2, \ldots\}$ es el conjunto de números primos mayores que 1 y, para cada $n \in \mathbb{N}$, definimos $A_n = \{p_n, p_n^2, p_n^3, \ldots\}$, entonces los conjuntos A_n son ajenos por parejas, así que las subsucesiones $\left(X_{p_1^k}\right)_{k\in\mathbb{N}}, \left(X_{p_2^k}\right)_{k\in\mathbb{N}}, \left(X_{p_3^k}\right)_{k\in\mathbb{N}}, \ldots$ cumplen con la propiedad requerida.

También podemos ordenar los elementos de la sucesión $(X_n)_{n\in\mathbb{N}}$ de la siguiente manera:

$X_1, X_2, X_4, X_7, X_{11}, X_{16}, X_{22}, \ldots$

$X_3, X_5, X_8, X_{12}, X_{17}, X_{23}, \ldots$

$X_6, X_9, X_{13}, X_{18}, X_{24}, \ldots$

$X_{10}, X_{14}, X_{19}, X_{25}, \ldots$

$X_{15}, X_{20}, X_{26}, \ldots$

$X_{21}, X_{27}, \ldots$

$X_{28}, \ldots$

$\vdots$

Cada renglón forma una subsucesión de $(X_n)_{n\in\mathbb{N}}$ y las subsucesiones de dos renglones diferentes no tienen elementos en común.

De manera general, si las sucesiones $\left(X_k^{(n)}\right)_{k\in\mathbb{N}}, \left(X_k^{(2)}\right)_{k\in\mathbb{N}}, \left(X_k^{(3)}\right)_{k\in\mathbb{N}}, \ldots$ son subsucesiones de $(X_n)_{n\in\mathbb{N}}$ tales que cualquier par de ellas no tienen elementos en común, definamos, para cada $n \in \mathbb{N}$:

$$U_n = \sum_{k=1}^{\infty} \frac{X_k^{(n)}}{2^k}$$

Entonces, para cada $n \in \mathbb{N}$, U_n tiene distribución uniforme en el intervalo $[0,1]$.

Para cada $m \in \mathbb{N}$, las sumas parciales $\sum_{k=1}^{m} \frac{X_k^{(1)}}{2^k}, \sum_{k=1}^{m} \frac{X_k^{(2)}}{2^k}, \sum_{k=1}^{m} \frac{X_k^{(3)}}{2^k}, \ldots$ forman una familia de variables aleatorias independientes, así que, para cualquier $n \in \mathbb{N}$ y cualquier $(x_1, \ldots, x_n) \in \mathbb{R}^n$, se tiene:

$$P\left[\sum_{k=1}^{m} \frac{X_k^{(1)}}{2^k} \leq x_1, \ldots, \sum_{k=1}^{m} \frac{X_k^{(n)}}{2^k} \leq x_n\right]$$

$$= P\left[\sum_{k=1}^{m} \frac{X_k^{(1)}}{2^k} \leq x_1\right] \cdots P\left[\sum_{k=1}^{m} \frac{X_k^{(n)}}{2^k} \leq x_n\right]$$

Como, para cada $j \in \mathbb{N}$, la sucesión $\left(\sum_{k=1}^{m} \frac{X_k^{(j)}}{2^k}\right)_{m\in\mathbb{N}}$ es no decreciente, la sucesión de eventos $\left(\left[\sum_{k=1}^{m} \frac{X_k^{(1)}}{2^k} \leq x_1, \ldots, \sum_{k=1}^{m} \frac{X_k^{(n)}}{2^k} \leq x_n\right]\right)_{m\in\mathbb{N}}$ es decreciente y la intersección de todos ellos es el evento

$$\left[\sum_{k=1}^{\infty} \frac{X_k^{(1)}}{2^k} \leq x_1, \ldots, \sum_{k=1}^{\infty} \frac{X_k^{(n)}}{2^k} \leq x_n\right]$$

Así que:

$$P\left[U_1 \leq x_1, U_n \leq x_2, \ldots, U_n \leq x_n\right]$$

$$P\left[\sum_{k=1}^{\infty} \frac{X_k^{(1)}}{2^k} \leq x_1, \ldots, \sum_{k=1}^{\infty} \frac{X_k^{(n)}}{2^k} \leq x_n\right]$$

$$= \lim_{m\to\infty} P\left[\sum_{k=1}^{m} \frac{X_k^{(1)}}{2^k} \leq x_1, \ldots, \sum_{k=1}^{m} \frac{X_k^{(n)}}{2^k} \leq x_n\right]$$

$$= \left(\lim_{m\to\infty} P\left[\sum_{k=1}^{m} \frac{X_k^{(1)}}{2^k} \leq x_1\right]\right) \cdots \left(\lim_{m\rightsquigarrow\infty} P\left[\sum_{k=1}^{m} \frac{X_k^{(n)}}{2^k} \leq x_n\right]\right)$$

$$= P\left[\sum_{k=1}^{\infty} \frac{X_k^{(1)}}{2^k} \leq x_1\right] \cdots P\left[\sum_{k=1}^{\infty} \frac{X_k^{(n)}}{2^k} \leq x_n\right]$$

$$= P\left[U_1 \leq x_1\right] \cdots P\left[U_n \leq x_n\right]$$

Por lo tanto, las variables aleatorias $U_1, U_2, \ldots$ son independientes.

∎

Teorema 4. *Se puede definir, sobre $((0,1], \mathcal{L}, P)$, una sucesión de variables aleatorias independientes $(X_n)_{n\in\mathbb{N}}$, tales que, para cualquier $n \in \mathbb{N}$, la distribución de X_n es normal estándar.*

Demostración

Consideremos una sucesión de variables aleatorias independientes $(U_n)_{n\in\mathbb{N}}$, definidas sobre $((0,1], \mathcal{L}, P)$, cada una de ellas con distribución uniforme en el intervalo $[0,1]$.

Para cada $n \in \mathbb{N}$, definamos la función $F_n : \mathbb{R} \mapsto \mathbb{R}$ de la siguiente manera:

$$F_n(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy$$

F_n es una función continua estrictamente creciente, así que tiene una inversa continua. Denotemos por c_n a la inversa de F_n; c_n es entonces una función continua definida en el intervalo $(0,1)$.

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

$$X_n = \begin{cases} c_n(U_n) & \text{si } U_n \in (0,1) \\ I_{(0,1)^c}(U_n) & \text{en otro caso} \end{cases}$$

Si $x \in \mathbb{R}$, se tiene:

$$[X_n \leq x] = [X_n \leq x, U_n \in (0,1)] \bigcup [X_n \leq x, U_n \notin (0,1)]$$

Así que, como $P[U_n \in (0,1)] = 1$, se tiene:

$$F_{X_n}(x) = P[X_n \leq x] = P[c_n(U_n) \leq x, U_n \in (0,1)] = P[U_n \leq c_n^{-1}(x), U_n \in (0,1)]$$

$$= P[U_n \leq c_n^{-1}(x)] = P[U_n \leq F_n(x)] = F_n(x)$$

Así que X_n tiene distribución normal estándar.

Finalmente, como las variables aleatorias $U_1, U_2, \ldots$ son independientes, también lo son $X_1, X_2, \ldots$.

■