

# TEORÍA DE LA PROBABILIDAD

---

## Variables Aleatorias

Desde el inicio del Cálculo de Probabilidades se trataba ya con cantidades que podían tomar diferentes valores, cada uno con una determinada probabilidad. En un principio se referían a la ganancia que un jugador podía obtener en un juego de azar. Más tarde se trataron otro tipo de problemas donde también intervenían cantidades que podían tomar distintos valores. A principios del siglo XX se hablaba simplemente de cantidades o variables. En su libro de 1925, Paul Lévy se refería a esas cantidades como variables eventuales. Markov se refería a ellas como variables aleatorias. Una vez que se formula axiomáticamente la Teoría de la Probabilidad, las variables aleatorias quedan identificadas con las funciones medibles.

**En lo que sigue asumimos que tenemos definido un espacio de probabilidad completo  $(\Omega, \mathfrak{S}, P)$ .**

Dada una función  $X : \Omega \rightarrow \overline{\mathbb{R}}$  y un conjunto  $B \subset \mathbb{R}$ , la notación  $[X \in B]$  será una manera abreviada de representar al conjunto  $\{\omega \in \Omega : X(\omega) \in B\}$ . Si  $a, b \in \mathbb{R}$ , las notaciones  $[a < X < b]$ ,  $[a \leq X \leq b]$ ,  $[a \leq X < b]$ ,  $[a < X \leq b]$ ,  $[X < b]$ ,  $[X \leq b]$ ,  $[X > a]$  y  $[X \geq a]$  se entenderán en un sentido similar.

Obsérvese que, si  $B \subset \mathbb{R}$ , se tiene:

$$[X \in B] = X^{-1}(B)$$

**Definición 1.** *Una variable aleatoria real es una función  $X : \Omega \rightarrow \overline{\mathbb{R}}$  tal que  $[X \leq x] \in \mathfrak{S}$  para cualquier  $x \in \mathbb{R}$ .*

Lo que nos interesa de una variable aleatoria son las probabilidades con las que toma sus diferentes posibles valores. En principio lo que nos gustaría conocer son todas las probabilidades de la forma  $P[X \in A]$ , donde  $A$  es cualquier subconjunto de  $\overline{\mathbb{R}}$ ; sin embargo, esto no siempre es posible obtenerlo, pero la condición  $[X \leq x] \in \mathfrak{S}$  para cualquier  $x \in \mathbb{R}$  implica que  $[X \in B] \in \mathfrak{S}$  para cualquier conjunto boreliano  $B$  de  $\overline{\mathbb{R}}$ , lo cual se muestra a continuación.

**Proposición 1.** *Sea  $X : \Omega \rightarrow \overline{\mathbb{R}}$  una variable aleatoria real, entonces  $[X \in B] \in \mathfrak{S}$  para cualquier conjunto  $B \in \mathfrak{B}(\overline{\mathbb{R}})$ .*

### Demostración

Definamos:

$$\mathbb{H} = \{B \in \mathfrak{B}(\overline{\mathbb{R}}) : [X \in B] \in \mathfrak{S}\}$$

Vamos a probar que  $\mathbb{H}$  es una  $\sigma$ -álgebra de subconjuntos de  $\overline{\mathbb{R}}$ .

i)  $\overline{\mathbb{R}} \in \mathbb{H}$  ya que  $[X \in \overline{\mathbb{R}}] = \Omega \in \mathfrak{S}$ .

ii) Si  $B \in \mathbb{H}$ , entonces  $[X \in B] \in \mathfrak{S}$ , así que  $[X \in B^c] = [X \in B]^c \in \mathfrak{S}$ . Por lo tanto,  $B^c \in \mathbb{H}$ .

iii) Si  $(B_n)_{n \in \mathbb{N}}$  es una sucesión de elementos de  $\mathbb{H}$ , entonces  $[X \in B_n] \in \mathfrak{S}$  para cualquier  $n \in \mathbb{N}$ , así que:

$$[X \in \cup_{n=1}^{\infty} B_n] = \cup_{n=1}^{\infty} [X \in B_n] \in \mathfrak{S}$$

Por lo tanto,  $\cup_{n=1}^{\infty} B_n \in \mathbb{H}$ .

Así que, efectivamente,  $\mathbb{H}$  es una  $\sigma$ -álgebra de subconjuntos de  $\mathbb{R}$ .

Además, por la definición de variable aleatoria real, los conjuntos de la forma  $[-\infty, x]$ , donde  $x \in \mathbb{R}$ , pertenecen a  $\mathbb{H}$ . Por lo tanto,  $\mathbb{H}$  contiene a la  $\sigma$ -álgebra generada por esos conjuntos; es decir, a los borelianos en  $\overline{\mathbb{R}}$ . Así que:

$$\mathfrak{B}(\overline{\mathbb{R}}) \subset \mathbb{H}$$

Finalmente, por la definición de  $\mathbb{H}$ , todo elemento de  $\mathbb{H}$  pertenece a  $\mathfrak{B}(\overline{\mathbb{R}})$ , así que:

$$\mathbb{H} \subset \mathfrak{B}(\overline{\mathbb{R}})$$

Por lo tanto:

$$\mathbb{H} = \mathfrak{B}(\overline{\mathbb{R}})$$

Es decir,  $[X \in B] \in \mathfrak{S}$  para cualquier conjunto  $B \in \mathfrak{B}(\overline{\mathbb{R}})$ . ■

**Proposición 2.** Sea  $X$  una variable aleatoria real. La función  $\mu_X : \mathfrak{B}(\overline{\mathbb{R}}) \mapsto \mathbb{R}$  definida por:

$$\mu_X(B) = P[X \in B]$$

es una medida de probabilidad.

### Demostración

Son 4 las propiedades que tenemos que verificar:

i)  $\mu_X(B) \geq 0$  para cualquier  $B \in \mathfrak{B}(\overline{\mathbb{R}})$ .

ii)  $\mu_X(\emptyset) = 0$ .

iii)  $\mu_X(\overline{\mathbb{R}}) = 1$ .

iv) Si  $(B_n)_{n \in \mathbb{N}}$  es una sucesión de elementos de  $\mathfrak{B}(\overline{\mathbb{R}})$ , ajenos por parejas, entonces:

$$\mu_X(\cup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mu_X(B_n)$$

Las propiedades i, ii y iii son inmediatas.

Demostremos la propiedad iv.

Sea  $(B_n)_{n \in \mathbb{N}}$  una sucesión de elementos de  $\mathfrak{B}(\overline{\mathbb{R}})$ , ajenos por parejas, entonces la sucesión  $([X \in B_n])_{n \in \mathbb{N}}$  está formada por conjuntos en  $\mathfrak{S}$ , ajenos por parejas; así que:

$$\mu_X(\cup_{n=1}^{\infty} B_n) = P[X \in \cup_{n=1}^{\infty} B_n] = P(\cup_{n=1}^{\infty} [X \in B_n]) = \sum_{n=1}^{\infty} P[X \in B_n] = \sum_{n=1}^{\infty} \mu_X(B_n)$$

■

**Definición 2.** Si  $X$  es una variable aleatoria real, la medida  $\mu_X$  será llamada la distribución de la variable aleatoria  $X$ .

En realidad es esta medida  $\mu_X$  lo que nos interesa de una variable aleatoria  $X$ , ya que es ésta la que nos da las probabilidades de que  $X$  tome sus diferentes valores. En otras palabras:

**Una variable aleatoria real representa una medida de probabilidad sobre los conjuntos borelianos de  $\overline{\mathbb{R}}$ .**

## Independencia de variables aleatorias

**Definición 3.** Diremos que las variables aleatorias de una familia no vacía cualquiera  $\{X_\gamma\}$ , finita o infinita, son independientes, si dada cualquier subcolección finita de ellas,  $X_{\gamma_1}, \dots, X_{\gamma_n}$  y cualquier colección de subconjuntos borelianos de  $\overline{\mathbb{R}}$ ,  $A_1, \dots, A_n$ , donde  $n \in \mathbb{N}$ , se tiene:

$$P[X_{\gamma_1} \in A_1, \dots, X_n \in X_{\gamma_n}] = P[X_{\gamma_1} \in A_1] \cdots P[X_{\gamma_n} \in A_n]$$

**Proposición 3.**  $n$  variables aleatorias  $X_1, \dots, X_n$ , son independientes si y sólo si para cualquier colección de subconjuntos borelianos de  $\overline{\mathbb{R}}$ ,  $A_1, \dots, A_n$ , se tiene:

$$P[X_1 \in A_1, \dots, X_n \in A_n] = P[X_1 \in A_1] \cdots P[X_n \in A_n].$$

### Demostración

Si las  $n$  variables aleatorias son independientes, la relación se sigue de la definición.

Supongamos ahora que para cualquier colección de subconjuntos borelianos de  $\overline{\mathbb{R}}$ ,  $A_1, \dots, A_n$ , se tiene:

$$P[X_1 \in A_1, \dots, X_n \in A_n] = P[X_1 \in A_1] \cdots P[X_n \in A_n]$$

Sea  $\{X_{i_1}, \dots, X_{i_k}\}$  un subconjunto de la familia  $\{X_1, \dots, X_n\}$  y  $A_{i_1}, \dots, A_{i_k}$  subconjuntos borelianos de  $\overline{\mathbb{R}}$ . Para  $j \in \{1, \dots, n\} - \{i_1, \dots, i_k\}$ , definamos  $A_j = \overline{\mathbb{R}}$ , entonces:

$$\begin{aligned} P[X_{i_1} \in A_{i_1}, \dots, X_{i_k} \in A_{i_k}] &= P[X_1 \in A_1, \dots, X_n \in A_n] \\ &= P[X_1 \in A_1] \cdots P[X_n \in A_n] = P[X_{i_1} \in A_{i_1}] \cdots P[X_{i_k} \in A_{i_k}] \end{aligned}$$

Así que  $X_1, \dots, X_n$  son independientes. ■

**Corolario 1.** Las variables aleatorias de una familia infinita numerable,  $\{X_1, X_2, \dots\}$ , son independientes si y sólo si para cualquier  $n \in \mathbb{N}$  y cualquier colección de subconjuntos borelianos en  $\overline{\mathbb{R}}$ ,  $A_1, \dots, A_n$ , se tiene:

$$P[X_1 \in A_1, \dots, X_n \in A_n] = P[X_1 \in A_1] \cdots P[X_n \in A_n]$$

**Definición 4.** Diremos que una función  $f : \overline{\mathbb{R}} \mapsto \overline{\mathbb{R}}$  es boreliana si  $f^{-1}(B) \in \mathfrak{B}(\overline{\mathbb{R}})$  para cualquier conjunto  $B \in \mathfrak{B}(\overline{\mathbb{R}})$ .

**Teorema 1.** Sean  $X_1, \dots, X_n$   $n$  variables aleatorias independientes y  $f_1, \dots, f_n$   $n$  funciones borelianas de  $\overline{\mathbb{R}}$  en  $\overline{\mathbb{R}}$ . Entonces las variables aleatorias  $f_1(X_1), \dots, f_n(X_n)$  son independientes.

### Demostración

Sean  $A_1, \dots, A_n$  subconjuntos borelianos de  $\overline{\mathbb{R}}$ , entonces  $f_1^{-1}(A_1), \dots, f_n^{-1}(A_n)$  son subconjuntos borelianos de  $\overline{\mathbb{R}}$ , así que:

$$\begin{aligned} P[f_1(X_1) \in A_1, \dots, f_n(X_n) \in A_n] &= P[X_1 \in f_1^{-1}(A_1), \dots, X_n \in f_n^{-1}(A_n)] \\ &= P[X_1 \in f_1^{-1}(A_1)] \cdots P[X_n \in f_n^{-1}(A_n)] = P[f_1(X_1) \in A_1] \cdots P[f_n(X_n) \in A_n] \end{aligned}$$
■

El gran impulso para el desarrollo de una teoría de la probabilidad, que le haría ganar un lugar dentro de las matemáticas, proviene de los llamados teoremas límite, los cuales se refieren al comportamiento a largo plazo de sucesiones de variables aleatorias. El primero de estos resultados, que para algunos autores marca verdaderamente el inicio de la historia de la teoría de la probabilidad, se debe a Jacques Bernoulli, quien dedicó 20 años de su vida a la búsqueda de una prueba matemática de la relación que existe entre la probabilidad de un evento y la frecuencia relativa con la que éste ocurre en una serie grande de repeticiones del correspondiente experimento aleatorio. El resultado, conocido como teorema de Bernoulli, se publicó en el año 1713, ocho años después de la muerte de su autor.

### Teorema de Bernoulli

Sea  $\mathcal{E}$  un experimento aleatorio que admite  $t$  posibles resultados equiprobables y  $A$  un evento relativo a ese experimento, para el cual hay  $r$  resultados que favorecen su ocurrencia. Consideremos un nuevo experimento aleatorio consistente en la repetición indefinida del experimento  $\mathcal{E}$ , de tal manera que cada repetición es independiente de las otras. Sean  $X_n$  el número de veces que ocurre el evento  $A$  en las primeras  $n$  repeticiones del experimento y  $\varepsilon$  un número positivo arbitrario, entonces:

$$\lim_{n \rightarrow \infty} P \left[ \left| \frac{X_n}{n} - \frac{r}{t} \right| > \varepsilon \right] = 0$$

Puede decirse que, a partir de la publicación del teorema de Bernoulli, el motor de desarrollo de la teoría de la probabilidad fue la búsqueda de resultados que permitieran mejorar y generalizar ese teorema.

La publicación del teorema de Bernoulli hizo renacer el interés por el Cálculo de Probabilidades, el cual, después de la publicación del trabajo de Christiaan Huygens, había quedado relegado, siendo visto únicamente como una curiosidad que tenía que ver exclusivamente con los juegos de azar.

En la búsqueda de mejorar el resultado de Bernoulli, Abraham de Moivre demostró, en 1733, un resultado que también sería de gran importancia en el desarrollo del Cálculo de Probabilidades. En ese año se publicó su artículo titulado *Approximatio ad Summam Terminorum Binomii  $(a + b)^n$  in Seriem expansis*, en el cual expuso un resultado que conduciría a lo que ahora se conoce como el Teorema del Límite Central. El artículo fue publicado en latín y circuló en forma privada. En el año 1738 ese artículo fue incluido en la segunda edición de su libro *The Doctrine of Chances* con el título *A Method of approximating the Sum of the Terms of the Binomial  $(a + b)^n$  expanded into a Series, from whence are deduced some practical Rules to estimate the Degree of Assent which is to be given to Experiments*. Con terminología y notación moderna, el resultado de de Moivre puede enunciarse de la siguiente manera:

### Teorema de de Moivre

Si  $d$  es un número real positivo del orden de  $\sqrt{n}$  y, para cada  $n \in \mathbb{N}$ ,  $X_n$  es una variable aleatoria con distribución binomial de parámetros  $n$  y  $p = \frac{1}{2}$ , entonces:

$$\lim_{n \rightarrow \infty} P \left[ -x \leq \frac{X_n - \frac{1}{2}n}{\frac{1}{2}\sqrt{n}} \leq x \right] = \frac{2}{\sqrt{2\pi}} \int_0^x e^{-\frac{1}{2}z^2} dz$$

Mencionaba de Moivre que su resultado se puede generalizar facilmente para cualquier valor de  $p$ . Esta generalización fue expuesta por Pierre Simon Laplace en su libro *Théorie Analytique des Probabilités*, publicado en el año 1812. En notación moderna, el resultado de Laplace puede escribirse de la siguiente manera:

### Teorema de de Moivre-Laplace

Si  $d$  es un número real positivo del orden de  $\sqrt{n}$  y, para cada  $n \in \mathbb{N}$ ,  $X_n$  es una variable aleatoria con distribución binomial de parámetros  $n$  y  $p$ , entonces:

$$\lim_{n \rightarrow \infty} P \left[ -x \leq \frac{X_n - np}{\sqrt{np(1-p)}} \leq x \right] = \frac{2}{\sqrt{2\pi}} \int_0^x e^{-\frac{1}{2}z^2} dz$$

Este proceso continuaría desarrollándose y recibiría un gran impulso, entre 1870 y 1900, con los trabajos de la llamada escuela rusa, representada por Pafnuty Lvovich Chebyshev, Andrei Andreyevich Markov y Aleksandr Mikhailovich Lyapunov.

El tercer teorema límite fue publicado en el año 1909 por Émile Borel, el cual se enuncia a continuación:

### Teorema de Borel

Sea  $\mathcal{E}$  un experimento aleatorio y  $A$  un evento relativo a ese experimento, de probabilidad igual a  $p$ . Consideremos un nuevo experimento aleatorio consistente en la repetición indefinida del experimento  $\mathcal{E}$ , de tal manera que cada repetición es independiente de las otras. Sea  $X_n$  el número de veces que ocurre el evento  $A$  en las primeras  $n$  repeticiones del experimento, entonces:

$$P \left[ \lim_{n \rightarrow \infty} \frac{X_n}{n} = p \right] = 1$$

La forma general de los teoremas límite se dio entre 1900 y 1930, con la formulación de las leyes de los grandes números y el teorema central del límite, tanto en su forma clásica, relativa a la convergencia a la distribución normal, como en su forma moderna, relativa a la convergencia a cualquier otro tipo de distribución, sobresaliendo en ese periodo los trabajos de Aleksandr Yakovlevich Khintchine, Andrey Nikolaevich Kolmogorov, J. W. Lindeberg, William Feller y Paul Pierre Lévy, entre otros.

Como puede verse, fueron más de 200 años de historia de la teoría de la probabilidad guiada por el estudio de los teoremas límite.

Una vez culminada esta etapa, el estudio de las variables aleatorias siguió siendo central en la teoría de la probabilidad ya que, a principios del siglo XX, paralelamente al estudio de los teoremas límite se inició el estudio de los procesos estocásticos; es decir, de familias infinitas de variables aleatorias no independientes.

## Funciones de distribución

**Definición 5 (Función de distribución).** Si  $X$  es una variable aleatoria real, la función  $F_X : \mathbb{R} \mapsto \mathbb{R}$ , definida por  $F_X(x) = P[X \leq x]$ , es llamada la función de distribución de  $X$ .

Introducimos antes una medida de probabilidad  $\mu_X$  asociada a una variable aleatoria  $X$  y denominamos a esa medida la distribución de  $X$ . Se puede mostrar que ésta queda determinada por la función de distribución de  $X$ ; de ahí la importancia de esta última.

En la exposición de lo que sigue, en esta parte 2, vamos a considerar únicamente variables aleatorias reales  $X$  tales que  $P[X \in \mathbb{R}] = 1$ .

**Proposición 4.** Sea  $X$  una variable aleatoria y  $F_X$  su función de distribución, entonces:

1.  $F_X$  es una función no decreciente y continua por la derecha.
2.  $\lim_{x \rightarrow \infty} F_X(x) = 1$
3.  $\lim_{x \rightarrow -\infty} F_X(x) = 0$
4.  $F_X(x-) = P[X < x]$  para cualquier  $x \in \mathbb{R}$ .

### Demostración

1. Sea  $(x_n)_{n \in \mathbb{N}}$  una sucesión decreciente tal que  $\lim_{n \rightarrow \infty} x_n = x$ , entonces:

$$\begin{aligned} F_X(x+) &= \lim_{n \rightarrow \infty} F_X(x_n) = \lim_{n \rightarrow \infty} P[X \leq x_n] = P\left(\bigcap_{n=1}^{\infty} [X \leq x_n]\right) \\ &= P[X \leq x] = F_X(x) \end{aligned}$$

2. Sea  $(x_n)$  una sucesión creciente tal que  $\lim_{n \rightarrow \infty} x_n = \infty$ , entonces:

$$\begin{aligned} \lim_{n \rightarrow \infty} F_X(x_n) &= \lim_{n \rightarrow \infty} P[X \leq x_n] = P\left(\bigcup_{n=1}^{\infty} [X \leq x_n]\right) \\ &= P[\Omega] = 1 \end{aligned}$$

3. Sea  $(x_n)$  una sucesión decreciente tal que  $\lim_{n \rightarrow \infty} x_n = -\infty$ , entonces:

$$\lim_{n \rightarrow \infty} F_X(x_n) = \lim_{n \rightarrow \infty} P[X \leq x_n] = P\left(\bigcap_{n=1}^{\infty} [X \leq x_n]\right) = P[\emptyset] = 0$$

4. Sea  $x \in \mathbb{R}$  y  $(x_n)$  una sucesión creciente tal que  $\lim_{n \rightarrow \infty} x_n = x$ , entonces:

$$F_X(x-) = \lim_{n \rightarrow \infty} F_X(x_n) = \lim_{n \rightarrow \infty} P[X \leq x_n] = P\left(\bigcup_{n=1}^{\infty} [X \leq x_n]\right) = P[X < x]$$

■

**Definición 6.** *Llamaremos función de distribución a cualquier función  $F : \mathbb{R} \mapsto \mathbb{R}$  que satisfaga las propiedades 1, 2 y 3 de la proposición 4.*

Sea  $X$  una variable aleatoria y definamos:

$$D = \{x \in \mathbb{R} : P[X = x] > 0\}$$

$$C = \{x \in \mathbb{R} : P[X = x] = 0\}$$

$$p = P[X \in D]$$

Clasificaremos a las variables aleatorias de acuerdo al valor de  $p$ . Si  $p = 0$ , diremos que la variable aleatoria es continua, si  $p = 1$  diremos que es discreta. Cuando  $0 < p < 1$ , se tiene  $P[X \in D] > 0$  y  $P[X \in C] > 0$ , de manera que se puede decir que, en ese caso, la variable aleatoria tiene una parte discreta (la que corresponde al conjunto  $D$ ) y una parte continua (la que corresponde al conjunto  $C$ ).

Un subconjunto importante del conjunto de variables aleatorias continuas está formado por las variables aleatorias  $X$  con valores en  $\mathbb{R}$  cuya función de distribución  $F_X$  es absolutamente continua; es decir, aquellas para las cuales existe una función Lebesgue medible no negativa  $f_X : (\mathbb{R}, \mathcal{L}(\mathbb{R}), \lambda) \mapsto (\overline{\mathbb{R}}, \mathfrak{B}(\overline{\mathbb{R}}))$  tal que:

$$F_X(x) = \int_{-\infty}^x f_X d\lambda$$

para cualquier  $x \in \mathbb{R}$ .

Diremos que una variable aleatoria de este tipo es absolutamente continua. Bajo determinadas condiciones,  $f_X$  es la derivada de  $F_X$  y es así como en muchas ocasiones se encuentra la función de densidad.

**Definición 7.** *Si  $X$  es una variable aleatoria discreta, llamaremos función de densidad de  $X$ , y la denotaremos por  $f_X$ , a la función  $f_X : \mathbb{R} \rightarrow [0, 1]$  definida mediante la relación:*

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

**Definición 8.** *Si  $X$  es una variable aleatoria absolutamente continua, llamaremos función de densidad de  $X$ , y la denotaremos por  $f_X$ , a cualquier función Lebesgue medible no negativa  $f : (\mathbb{R}, \mathcal{L}(\mathbb{R}), \lambda) \mapsto (\overline{\mathbb{R}}, \mathfrak{B}(\overline{\mathbb{R}}))$  tal que:*

$$F_X(x) = \int_{-\infty}^x f d\lambda$$

para cualquier  $x \in \mathbb{R}$ .

## Algunos ejemplos de distribuciones

Cualquier medida de probabilidad sobre  $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$  es una distribución de probabilidad, es decir, una distribución de alguna variable aleatoria. La demostración de esta afirmación es inmediata, ya que si  $\mu$  es una medida de probabilidad sobre  $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ , podemos tomar como espacio de probabilidad a la terna  $(\mathbb{R}, \mathfrak{B}(\mathbb{R}), \mu)$ . Entonces la variable aleatoria  $X : \Omega \rightarrow \mathbb{R}$  definida por  $X(\omega) = \omega$  tiene como distribución a la medida  $\mu$ . Sin embargo, es importante explicitar algunas de ellas, ya sea por su interés histórico o por presentarse con frecuencia en diferentes problemas. Así que a continuación haremos un listado de algunas distribuciones discretas y algunas absolutamente continuas, dando el nombre de la distribución y la función de densidad que la determina.

1. Distribución Bernoulli con parámetro  $p$ , donde  $p \in [0, 1]$ .

$$f_X(x) = \begin{cases} p & \text{si } x = 1 \\ 1 - p & \text{si } x = 0 \\ 0 & \text{en otro caso} \end{cases}$$

Esta distribución se presenta al realizar un ensayo de Bernoulli con probabilidad de éxito  $p$  y definiendo  $X = 1$  si se obtiene éxito y  $X = 0$  si se obtiene fracaso.

2. Distribución binomial con parámetros  $n$  y  $p$ , donde  $n \in \mathbb{N}$  y  $p \in [0, 1]$ .

$$f_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & \text{si } x \in \{0, 1, \dots, n\} \\ 0 & \text{en otro caso} \end{cases}$$

Esta distribución se presenta al realizar  $n$  ensayos de Bernoulli independientes, con probabilidad de éxito  $p$  en cada ensayo, y definiendo  $X$  como el número de éxitos que se obtienen en los  $n$  ensayos.

3. Distribución geométrica con parámetro  $p$ , donde  $p \in [0, 1]$ .

$$f_X(x) = \begin{cases} p(1-p)^x & \text{si } x \in \{0, 1, \dots\} \\ 0 & \text{en otro caso} \end{cases}$$

Esta distribución se presenta al realizar una sucesión indefinida de ensayos de Bernoulli independientes, con probabilidad de éxito  $p$  en cada ensayo, y definiendo  $X$  como el número de fracasos que se obtienen antes del primer éxito.

4. Distribución binomial negativa con parámetros  $n$  y  $p$ , donde  $n \in \mathbb{N}$  y  $p \in [0, 1]$ .

$$f_X(x) = \begin{cases} \binom{r+x-1}{x} p^r (1-p)^x & \text{si } x \in \{0, 1, \dots\} \\ 0 & \text{en otro caso} \end{cases}$$

Esta distribución se presenta al realizar una sucesión indefinida de ensayos de Bernoulli independientes, con probabilidad de éxito  $p$  en cada ensayo, y definiendo  $X$  como el número de fracasos que se obtienen antes del  $r$ -ésimo éxito.

En efecto:

Para  $k \in \{0, 1, \dots\}$ , definamos los siguientes eventos:

$A_k$ : Se obtienen  $r - 1$  éxitos en los primeros  $r + k - 1$  ensayos.

$B$ : Se obtiene éxito en el ensayo número  $r + k$ .

Entonces:

$$P[X = k] = P(A_k \cap B) = P(A_k)P(B) = \left[ \binom{r+k-1}{r-1} p^{r-1} (1-p)^k \right] p = \binom{r+k-1}{k} p^r (1-p)^k$$

5. Distribución Poisson con parámetro  $\lambda$ , donde  $\lambda$  es un número real positivo.

$$f_X(x) = \begin{cases} \frac{\lambda^x e^{-\lambda}}{x!} & \text{si } x \in \{0, 1, \dots\} \\ 0 & \text{en otro caso} \end{cases}$$

Esta distribución se obtiene como límite de una distribución binomial; de manera específica, se tiene el siguiente resultado:

Supongamos que, para cada  $n \in \mathbb{N}$ ,  $X_n$  es una variable aleatoria con distribución binomial de parámetros  $n$  y  $p \in (0, 1)$  de tal manera que  $\lambda = np$  es constante, entonces, para cualquier  $k \in \{0, 1, \dots\}$ , se tiene:

$$\lim_{n \rightarrow \infty} P[X_n = k] = \frac{\lambda^k e^{-\lambda}}{k!}$$

En efecto:

$$\begin{aligned} P[X = k] &= \binom{n}{k} p^k (1-p)^{n-k} = \frac{n(n-1)\dots(n-k+1)}{k!} \frac{\lambda^k}{n^k} \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \frac{\lambda^k}{k!} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{k-1}{n}\right) \left(1 - \frac{\lambda}{n}\right)^{-k} \left(1 - \frac{\lambda}{n}\right)^n \end{aligned}$$

Así que:

$$\lim_{n \rightarrow \infty} P[X = k] = \frac{\lambda^k}{k!} \lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n = \frac{\lambda^k e^{-\lambda}}{k!}$$

6. Distribución hipergeométrica con parámetros  $r$ ,  $s$  y  $n$ , donde  $r, s, n \in \mathbb{N}$  y  $n \leq r + s$ .

$$f_X(x) = \begin{cases} \frac{\binom{r}{x} \binom{s}{n-x}}{\binom{r+s}{n}} & \text{si } x \in \{0, 1, \dots, n\} \\ 0 & \text{en otro caso} \end{cases}$$

Esta distribución se presenta al considerar un muestreo sin reemplazo de la siguiente manera: Supongamos que se tiene una población formada por dos tipos de elementos, los cuales llamaremos de tipo I y II, de tal manera que hay  $r$  elementos de tipo I y  $s$  de tipo II. Si se toma una muestra sin reemplazo de tamaño  $n \leq r + s$  de esa población y llamamos  $X$  al número de elementos de tipo I que se obtienen en la muestra, entonces  $X$  tiene distribución hipergeométrica con parámetros  $r, s$  y  $n$ . En efecto, el número total de posibles muestras que pueden obtenerse es igual a  $\binom{r+s}{n}$  y todas ellas son equiprobables. De éstas, para  $k \in \{\text{máx}\{0, n-s\}, \dots, \text{mín}\{n, r\}\}$ , el total de muestras en las cuales se obtienen  $k$  objetos de tipo I y  $n-k$  objetos de tipo II es igual a  $\binom{r}{k} \binom{s}{n-k}$ , de lo cual se sigue el resultado.

7. Distribución uniforme discreta en el conjunto  $A$ , donde  $A$  es un conjunto finito de números reales.

$$f_X(x) = \begin{cases} \frac{1}{N} & \text{si } x \in A \\ 0 & \text{en otro caso} \end{cases}$$

donde  $N$  es el número de elementos de  $A$ .

Esta distribución se presenta al considerar el experimento consistente en elegir al azar un elemento del conjunto  $A$  y definiendo  $X$  como el elemento seleccionado.

8. Distribución uniforme continua en el conjunto  $A$ , donde  $A$  es un conjunto Lebesgue medible de medida positiva.

$$f_X(x) = \begin{cases} \frac{1}{\lambda(A)} & \text{si } x \in A \\ 0 & \text{en otro caso} \end{cases}$$

Esta distribución se presenta al considerar el experimento consistente en elegir al azar un elemento del conjunto  $A$  y definiendo  $X$  como el elemento seleccionado. Obsérvese que, en este caso,  $A$  no es un conjunto finito.

9. Distribución normal estándar.

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}$$

Aunque esta distribución se presenta en una diversidad de situaciones, por el momento la vamos a considerar como el límite de una sucesión de variables aleatorias independientes, todas ellas con distribución binomial. De manera específica, el resultado es el siguiente:

Sea  $(X_n)_{n \in \mathbb{N}}$  una sucesión de variables aleatorias independientes, todas ellas con distribución binomial de parámetros  $n$  y  $p$ , entonces:

$$\lim_{n \rightarrow \infty} P \left[ \frac{X_n - np}{\sqrt{np(1-p)}} \leq x \right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}z^2} dz$$

10. Distribución normal con parámetros  $\mu$  y  $\sigma^2$ , donde  $\mu, \sigma \in \mathbb{R}$  y  $\sigma > 0$ .

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$$

Esta distribución es una generalización de la distribución normal estándar y se tiene el siguiente resultado:

Sea  $X$  una variable aleatoria con distribución normal estándar,  $\mu \in \mathbb{R}$  y  $\sigma \in (0, \infty)$ . Entonces la variable aleatoria  $Y = \mu + \sigma X$  tiene distribución normal con parámetros  $\mu$  y  $\sigma^2$ .

11. Distribución exponencial con parámetro  $\lambda$ , donde  $\lambda$  es un número real positivo.

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x > 0 \\ 0 & \text{en otro caso} \end{cases}$$

Una distribución de este tipo se obtiene al considerar el tiempo que transcurre entre dos ocurrencias consecutivas de eventos que ocurren aleatoriamente en el tiempo.

Consideramos un experimento aleatorio consistente en la observación de eventos que ocurren aleatoriamente en el tiempo de tal manera que si, para  $t \geq 0$ ,  $X_t$  es el número de veces que ocurre el evento hasta el tiempo  $t$ , entonces, si  $X_0 = 0$  y si  $s < t$ , la variable aleatoria  $X_t - X_s$  tiene distribución Poisson con parámetro  $\lambda(t - s)$ , donde  $\lambda$  es una constante positiva.

Sea ahora  $T_1$  el tiempo que transcurre desde el instante inicial  $t = 0$  hasta que ocurre el primer evento. Obsérvese entonces que, para  $t > 0$ , el evento  $[T_1 > t]$  ocurre si y sólo si ocurre el evento  $[N_t = 0]$ . Por lo tanto:

$$P[T_1 > t] = P_0(t) = e^{-\lambda t}$$

Así que:

$$F_{T_1}(t) = 1 - e^{-\lambda t}$$

Y entonces, una función de densidad de  $T_1$  está dada por:

$$f_{T_1}(t) = \begin{cases} \lambda e^{-\lambda t} & \text{si } t > 0 \\ 0 & \text{en otro caso} \end{cases}$$

12. Distribución gama con parámetros  $\alpha$  y  $\lambda$ , donde  $\alpha$  y  $\lambda$  son números reales positivos.

$$f_X(x) = \begin{cases} \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)} & \text{si } x > 0 \\ 0 & \text{en otro caso} \end{cases}$$

donde  $\Gamma : (0, \infty) \rightarrow \mathbb{R}$  es la función gama, la cual está definida por:

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt.$$

Al igual que la distribución exponencial, una distribución de este tipo se obtiene al considerar el tiempo que transcurre entre cierto número de ocurrencias de eventos que ocurren aleatoriamente en el tiempo.

Consideremos un experimento aleatorio consistente en la observación de eventos que ocurren aleatoriamente en el tiempo de tal manera que si, para  $t \geq 0$ ,  $X_t$  es el número de veces que ocurre el evento hasta el tiempo  $t$ , entonces, si  $X_0 = 0$  y  $s < t$ , la variable aleatoria  $X_t - X_s$  tiene distribución Poisson con parámetro  $\lambda(t - s)$ , donde  $\lambda$  es una constante positiva.

Sabemos entonces que, para  $k$  entero no negativo y  $t > 0$ :

$$P[N_t = k] = \frac{(\lambda t)^k e^{-\lambda t}}{k!}$$

Sea ahora  $X$  el tiempo que transcurre desde el origen hasta la  $r$ -sima ocurrencia del evento. Para  $x > 0$ , se tiene:

$$P[X > x] = P[\text{ocurren menos de } r \text{ eventos en el intervalo de tiempo } [0, x]]$$

$$= \sum_{k=0}^{r-1} P[N_x = k] = \sum_{k=0}^{r-1} \frac{e^{-\lambda x} (\lambda x)^k}{k!}$$

$$\text{Así que: } F_X(x) = 1 - \sum_{k=0}^{r-1} \frac{e^{-\lambda x} (\lambda x)^k}{k!}$$

Por lo tanto, una función de densidad de  $X$  está dada por:

$$\begin{aligned} f_X(x) &= F'_X(x) = \lambda e^{-\lambda x} + \lambda \sum_{k=1}^{r-1} \frac{e^{-\lambda x} (\lambda x)^k}{k!} - \lambda \sum_{k=1}^{r-1} \frac{k e^{-\lambda x} (\lambda x)^{k-1}}{k!} \\ &= \lambda e^{-\lambda x} + \lambda \sum_{k=1}^{r-1} \frac{e^{-\lambda x} (\lambda x)^k}{k!} - \lambda \sum_{k=1}^{r-1} \frac{e^{-\lambda x} (\lambda x)^{k-1}}{(k-1)!} \\ &= \lambda \sum_{k=1}^{r-1} \frac{e^{-\lambda x} (\lambda x)^k}{k!} - \lambda \sum_{k=1}^{r-2} \frac{e^{-\lambda x} (\lambda x)^k}{k!} = \lambda \frac{e^{-\lambda x} (\lambda x)^{r-1}}{(r-1)!} = \frac{\lambda^r x^{r-1} e^{-\lambda x}}{(r-1)!} \end{aligned}$$

Por último, como  $\Gamma(r) = (r-1)!$ , la expresión anterior puede escribirse en la forma siguiente:

$$f_X(x) = \frac{\lambda^r x^{r-1} e^{-\lambda x}}{\Gamma(r)}$$

Ahora bien, como  $\Gamma(\alpha)$  está bien definida para cualquier  $\alpha > 0$ , se puede definir la distribución gama no únicamente para cuando  $\alpha$  es un número natural, sino para cualquier  $\alpha > 0$ .

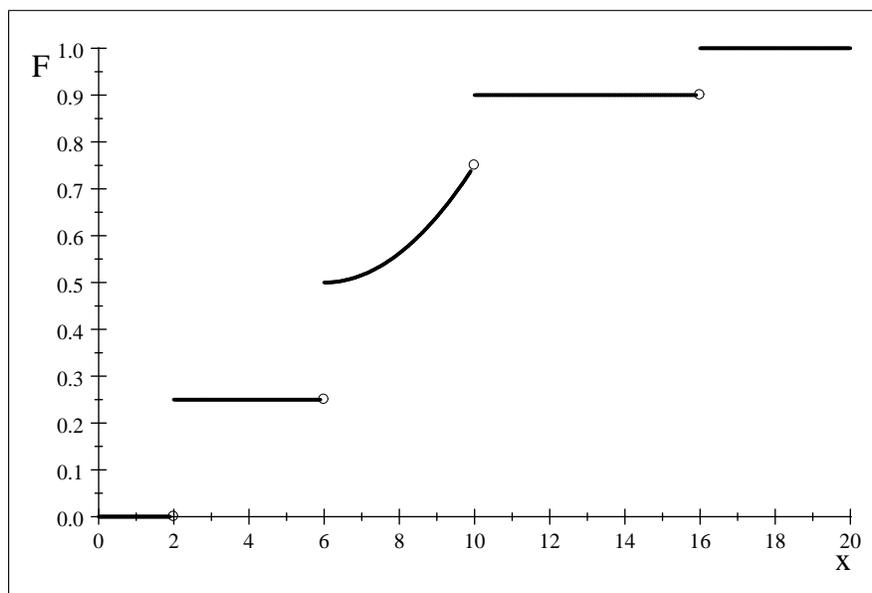
Un resultado interesante es que cualquier función de distribución se puede obtener mediante una transformación de una distribución uniforme continua, lo cual se expone a continuación.

Si  $X$  una variable aleatoria real, definamos la función  $d_X : (0, 1) \mapsto \mathbb{R}$  mediante la siguiente relación:

$$d_X(t) = \inf \{x \in \mathbb{R} : F_X(x) \geq t\}$$

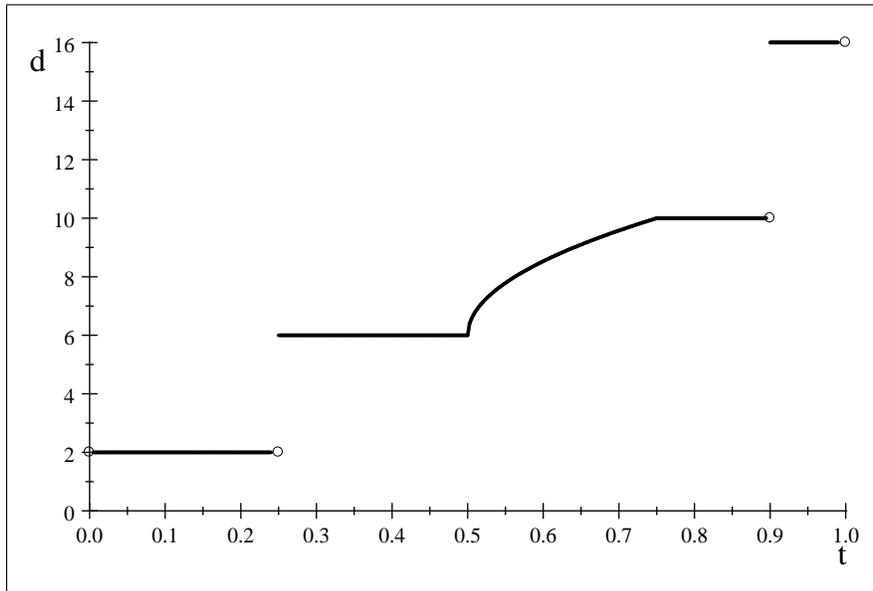
Para ilustrar esta definición, consideremos una variable aleatoria  $X$  cuya función de distribución está dada por la siguiente relación:

$$F_X(x) = \begin{cases} 0 & \text{si } x \in (-\infty, 2) \\ \frac{1}{4} & \text{si } x \in [2, 6) \\ \frac{1}{2} + \frac{1}{64}(x-6)^2 & \text{si } x \in [6, 10) \\ \frac{9}{10} & \text{si } x \in [10, 16) \\ 1 & \text{si } x \in [16, \infty) \end{cases}$$



Entonces:

$$d_X(t) = \begin{cases} 2 & \text{si } t \in (0, \frac{1}{4}] \\ 6 & \text{si } t \in (\frac{1}{4}, \frac{1}{2}] \\ 6 + 4\sqrt{4t - 2} & \text{si } t \in (\frac{1}{2}, \frac{3}{4}] \\ 10 & \text{si } t \in (\frac{3}{4}, \frac{9}{10}] \\ 16 & \text{si } t \in (\frac{9}{10}, 1) \end{cases}$$



**Teorema 2.** Sean  $X$  una variable aleatoria real con función de distribución  $F_X$  y  $U$  una variable aleatoria con distribución uniforme en el intervalo  $[0, 1]$ , ambas definidas sobre el mismo espacio de probabilidad  $(\Omega, \mathfrak{S}, P)$ . Entonces la función de distribución de la variable aleatoria  $d_X(U)$  es  $F_X$ .

### Demostración

Como  $F_X$  es continua por la derecha, se tiene  $F_X(d_X(t)) \geq t$ .

Tomemos  $z \in \mathbb{R}$ .

Si  $d_X(U(\omega)) \leq z$ , entonces  $F_X(d_X(U(\omega))) \leq F_X(z)$ ; así que:

$$U(\omega) \leq F_X(d_X(U(\omega))) \leq F_X(z)$$

Por lo tanto:

$$[d_X(U) \leq z] \subset [U \leq F_X(z)]$$

Si  $U(\omega) \leq F_X(z)$ , entonces  $d_X(U(\omega)) = \inf \{s \in \mathbb{R} : F_X(s) \geq U(\omega)\} \leq z$

Así que:

$$[U \leq F_X(z)] \subset [d_X(U) \leq z]$$

Por lo tanto:

$$P[d_X(U) \leq z] = P[U \leq F_X(z)] = F_X(z)$$

Así que:

$$F_{d_X(U)} = F_X$$

■

## Más sobre construcción de espacios de probabilidad

Lo expuesto hasta aquí presenta una dificultad. Por ejemplo, en el caso de la distribución normal estándar dijimos: Sea  $(X_n)_{n \in \mathbb{N}}$  una sucesión de variables aleatorias independientes, todas ellas con distribución binomial de parámetros  $n$  y  $p$ . Para poder considerar una sucesión de ese tipo, primero tendríamos que demostrar que hay un espacio de probabilidad en el cual se pueda definir formalmente, cosa que no hicimos. Algo similar puede decirse con respecto a las otras distribuciones.

En lo que sigue vamos a mostrar cómo se podría construir un espacio de probabilidad en el cual se pueden definir formalmente variables aleatorias con las distribuciones que mencionamos en los 12 puntos anteriores.

En particular, si denotamos por  $\mathfrak{L}((0, 1])$  a la familia de conjuntos Lebesgue medibles contenidos en el intervalo  $(0, 1]$ , la terna  $((0, 1], \mathfrak{L}((0, 1]), \lambda)$  constituye un espacio de probabilidad.

Vamos a ver que este espacio de probabilidad es bastante rico en el sentido de que podemos tomarlo para definir sucesiones de variables aleatorias independientes.

Recordemos que cualquier número real se puede expresar en base 2. Para los números reales  $x \in (0, 1)$  de la forma  $x = \frac{j}{2^n}$ , donde  $n \in \mathbb{N}$  y  $j \in \{1, 2, \dots, 2^n - 1\}$ , es decir, para los racionales diádicos, el desarrollo en base 2 no es único; por ejemplo  $\frac{5}{8}$  puede escribirse, en base 2, de las siguientes maneras:

0,101

0,10011111111111111111...

Para cada uno de estos puntos, elijamos como desarrollo en base 2 a la sucesión  $s_1, s_2, \dots$  para la cual existe  $N \in \mathbb{N}$  tal que  $s_k = 1$  para cualquier  $k \in \{N+1, N+2, \dots\}$ . El número 1 lo escribimos como  $0,1111111111111111\dots$ . De esta forma, para cada número real  $x \in (0, 1]$  tenemos un único desarrollo en base 2, a saber, el que contiene una infinidad de 1's.

Recordemos también que si  $x \in (0, 1]$  se desarrolla, en base 2, como  $0.s_1s_2s_3\dots$ , donde  $s_k \in \{0, 1\}$  para cualquier  $k \in \mathbb{N}$ , esto significa que se tiene la siguiente relación:

$$x = \sum_{k=1}^{\infty} \frac{s_k}{2^k}$$

Obsérvese que  $s_1, s_2, \dots, s_n$  son los primeros  $n$  términos del desarrollo de  $x$  en base 2 si y sólo si  $x \in \left(\sum_{i=1}^n \frac{s_i}{2^i}, \sum_{i=1}^n \frac{s_i}{2^i} + \frac{1}{2^n}\right]$ . Hay  $2^n$  diferentes secuencias  $s_1, s_2, \dots, s_n$ , así que hay  $2^n$  intervalos de ese tipo.

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  $s_k \in \{0, 1\}$  para cualquier  $k \in \{1, 2, \dots, n\}$ , es alguno de los intervalos de la forma  $\left(\frac{k-1}{2^n}, \frac{k}{2^n}\right]$ , donde  $k \in \{1, 2, \dots, 2^n\}$ , los cuales constituyen una partición del intervalo  $(0, 1]$ .

**Teorema 3.** Para cada  $n \in \mathbb{N}$  definamos  $X_n : (0, 1] \rightarrow \mathbb{R}$  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:

$$\begin{aligned} P[X_n = r] &= \sum_{\{(r_1, r_2, \dots, r_{n-1}) \in \mathbb{B}_{n-1}\}} P[X_1 = r_1, X_2 = r_2, \dots, X_{n-1} = r_{n-1}, X_n = r] \\ &= \sum_{\{(r_1, r_2, \dots, r_{n-1}) \in \mathbb{B}_{n-1}\}} P\left(\left(\sum_{k=1}^{n-1} \frac{r_k}{2^k} + \frac{r}{2^n}, \sum_{k=1}^{n-1} \frac{r_k}{2^k} + \frac{r}{2^n} + \frac{1}{2^n}\right)\right) \\ &= \sum_{\{(r_1, r_2, \dots, r_{n-1}) \in \mathbb{B}_{n-1}\}} \frac{1}{2^n} = \frac{1}{2} \end{aligned}$$

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  $(r_1, r_2, \dots, r_m) \in \mathbb{B}_m$ , 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  $X_1, X_2, \dots, X_m$  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  $p = \frac{1}{2}$ . ■

**Teorema 4.** Sea  $(\Omega, \mathfrak{F}, P)$  un espacio de probabilidad y  $X_1, X_2, \dots$  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  $P[X_k = s_k \text{ para toda } k \in \mathbb{N}] = 0$  y si  $x \in (0, 1]$  es un racional diádico,  $x$  tiene exactamente dos desarrollos en base 2, así que  $P[X \in x] = 0$ . Además,  $x = 0$  tiene únicamente un desarrollo en base 2, así que, también,  $P[X = 0] = 0$ . Por lo tanto,  $P[X \in x] = 0$  para cualquier racional diádico  $x \in [0, 1]$ .

Por otra parte, para cada  $n \in \mathbb{N}$ , consideremos un intervalo de la forma  $(\frac{j-1}{2^n}, \frac{j}{2^n}]$ , con  $j \in \{1, \dots, 2^n\}$ . Asociada a tal intervalo existe una única colección de 0's y 1's,  $(s_1, s_2, \dots, s_n)$ , tal que un punto  $x$  pertenece al intervalo  $(\frac{j-1}{2^n}, \frac{j}{2^n}]$  si y sólo si tiene un desarrollo en base 2 de la forma  $x = 0.s_1s_2 \dots s_n \dots$ . Por lo tanto:

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

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

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

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

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

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

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

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  $[0, 1]$ . ■

**Teorema 5.** *Se puede definir, sobre  $((0, 1], \mathfrak{L}((0, 1]), \lambda)$ , una sucesión de variables aleatorias independientes, cada una de ellas con distribución uniforme en el intervalo  $[0, 1]$ .*

### Demostración

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

De acuerdo con la proposición 4, si  $(X_{n_k})_{k \in \mathbb{N}}$  es cualquier subsucesión de la sucesión  $(X_n)_{n \in \mathbb{N}}$  y definimos  $Y : (0, 1] \rightarrow \mathbb{R}$  mediante la relación  $Y = \sum_{k=1}^{\infty} \frac{Y_k}{2^k}$ , entonces  $Y$  es una variable aleatoria con distribución uniforme en el intervalo  $[0, 1]$ .

Para definir una sucesión de variables aleatorias independientes, definidas sobre  $((0, 1], \mathfrak{L}((0, 1]), \lambda)$ , 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  $\{p_1, p_2, \dots\}$  es el conjunto de números primos mayores que 1 y, para cada  $n \in \mathbb{N}$ , definimos  $A_n = \{p_n, p_n^2, p_n^3, \dots\}$ , entonces los conjuntos  $A_n$  son ajenos por parejas, así que las subsucesiones  $(X_{p_1^k})_{k \in \mathbb{N}}$ ,  $(X_{p_2^k})_{k \in \mathbb{N}}$ ,  $(X_{p_3^k})_{k \in \mathbb{N}}$ ,  $\dots$  cumplen con la propiedad requerida.

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

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

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

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

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

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

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

$X_{28}, \dots$

$\vdots$

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

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

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  $\left( \left[ \sum_{k=1}^m \frac{X_k^{(1)}}{2^k} \leq x_1, \dots, \sum_{k=1}^m \frac{X_k^{(n)}}{2^k} \leq x_n \right] \right)_{m \in \mathbb{N}}$  es decreciente y la intersección de todos ellos es el evento  $\left[ \sum_{k=1}^{\infty} \frac{X_k^{(1)}}{2^k} \leq x_1, \dots, \sum_{k=1}^{\infty} \frac{X_k^{(n)}}{2^k} \leq x_n \right]$ ; así que:

$$\begin{aligned} P \left[ \sum_{k=1}^{\infty} \frac{X_k^{(1)}}{2^k} \leq x_1, \dots, \sum_{k=1}^{\infty} \frac{X_k^{(n)}}{2^k} \leq x_n \right] &= \lim_{m \rightarrow \infty} P \left[ \sum_{k=1}^m \frac{X_k^{(1)}}{2^k} \leq x_1, \dots, \sum_{k=1}^m \frac{X_k^{(n)}}{2^k} \leq x_n \right] \\ &= \left( \lim_{m \rightarrow \infty} P \left[ \sum_{k=1}^m \frac{X_k^{(1)}}{2^k} \leq x_1 \right] \right) \cdots \left( \lim_{m \rightarrow \infty} P \left[ \sum_{k=1}^m \frac{X_k^{(n)}}{2^k} \leq x_n \right] \right) \\ &= P \left[ \sum_{k=1}^{\infty} \frac{X_k^{(1)}}{2^k} \leq x_1 \right] \cdots P \left[ \sum_{k=1}^{\infty} \frac{X_k^{(n)}}{2^k} \leq x_n \right] \end{aligned}$$

Así que las variables aleatorias  $U_1, U_2, \dots$  son independientes. ■

**Teorema 6.** Sea  $(F_n)_{n \in \mathbb{N}}$  una sucesión de funciones de distribución. Se puede definir, sobre  $((0, 1], \mathcal{L}((0, 1]), \lambda)$ , una sucesión  $(X_n)_{n \in \mathbb{N}}$  de variables aleatorias independientes de tal manera que, para cada  $n \in \mathbb{N}$ , la función de distribución de  $X_n$  es  $F_n$ .

### Demostración

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

Para cada  $n \in \mathbb{N}$ , definamos las funciones  $d_n : (0, 1) \mapsto \mathbb{R}$  y  $X_n : (0, 1] \mapsto \mathbb{R}$  mediante las siguientes relaciones:

$$d_n(t) = \inf \{x \in \mathbb{R} : F_n(x) \geq t\}$$

$$X_n = d_n(U_n)$$

Por la proposición 2, la función de distribución de  $X_n$  es  $F_n$ .

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

### Ejercicios

**Ejercicio 1.** Demuestra que los términos  $t_k = \binom{n}{k} p^k (1-p)^{n-k}$ , de una distribución binomial, crecen con  $k$  hasta alcanzar su máximo valor cuando  $np + p - 1 \leq k \leq np + p$ , después de lo cual decrecen con  $k$ .

**Ejercicio 2.** Sea  $X$  una variable aleatoria con distribución geométrica. Demuestra que para cualquier par de números enteros no negativos  $n$  y  $k$ , se tiene:

$$P[X = k + n \mid X > n - 1] = P[X = k]$$

Obsérvese que este resultado dice que si en una sucesión de ensayos de Bernoulli independientes ya han ocurrido  $n$  fracasos, entonces la probabilidad de que haya  $k$  fracasos más antes del primer éxito se puede calcular como si se reiniciara el proceso. Este resultado se puede interpretar diciendo que una variable aleatoria con distribución geométrica no tiene memoria.

**Ejercicio 3.** Supongamos que, para cada  $r \in \mathbb{N}$ ,  $X_n$  es una variable aleatoria con distribución binomial negativa de parámetros  $r$  y  $p \in (0, 1)$  de tal manera que  $\lambda = r \frac{q}{p}$  es constante. Demuestra que, para cualquier  $k \in \{0, 1, \dots\}$ , se tiene:

$$\lim_{n \rightarrow \infty} P[X_n = k] = \frac{\lambda^k e^{-\lambda}}{k!}$$

**Ejercicio 4.** Supongamos que, para cada  $r, s, n \in \mathbb{N}$ ,  $X_n^{(r,s)}$  es una variable aleatoria con distribución hipergeométrica de parámetros  $r, s$  y  $n$ . Demuestra que, para cualquier  $k \in \{0, 1, \dots, n\}$ , se tiene:

$$\lim_{r,s \rightarrow \infty} \frac{P[X_n^{(r,s)}=k]}{\binom{n}{k} \left(\frac{r}{r+s}\right)^k \left(\frac{s}{r+s}\right)^{n-k}} = 1$$

**Ejercicio 5.** Sea  $X$  una variable aleatoria con distribución binomial de parámetros  $n$  y  $p$ . Si se sabe que el valor de  $X$  es  $k$ , ¿cuál es la probabilidad de que ocurra cada uno de los  $\binom{n}{k}$  arreglos de  $k$  éxitos y  $n - k$  fracasos?

**Ejercicio 6.** Sea  $X$  una variable aleatoria discreta que toma únicamente valores enteros no negativos y tal que, para cualquier entero no negativo  $n$ , se tiene:

$$P[X = n \mid X > n - 1] = P[X = 0]$$

Demuestra que  $X$  tiene distribución geométrica.

**Ejercicio 7.** Supongamos que, para  $n \geq 1$ , una pareja tiene  $n$  hijos con probabilidad igual a  $\alpha p^n$ , en donde  $0 < \alpha \leq \frac{1-p}{p}$  y  $0 < p < 1$ . Supongamos además que cada uno de los hijos es hombre o mujer con probabilidad igual a  $\frac{1}{2}$ . Para  $k \in \{1, 2, \dots\}$ , ¿cuál es la probabilidad de que una pareja tenga por lo menos  $k$  hijos, de los cuales  $k$  exactamente sean mujeres?

**Definición 9.** Elegir al azar un número real en el intervalo  $(0, 1)$  se interpreta diciendo que si tenemos dos subconjuntos  $A$  y  $B$ , contenidos en  $(0, 1)$  y tales que se pueda superponer geoméricamente uno sobre otro, entonces la probabilidad de que el punto seleccionado pertenezca al conjunto  $A$  es igual a la probabilidad de que pertenezca a  $B$ .

**Ejercicio 8.** Consideremos el experimento aleatorio consistente en la elección al azar de un número real en el intervalo  $(0, 1)$  y sea  $X$  el número que se obtiene.

- Construye un espacio de probabilidad  $(\Omega, \mathfrak{S}, P)$  para este experimento aleatorio.
- ¿Cuál es la probabilidad de que  $X$  sea un número racional?
- ¿Cuál es la probabilidad de que  $X$  pertenezca al conjunto de Cantor?

**Ejercicio 9.** Consideremos una partícula que se mueve a saltos sobre la recta real de tal manera que, comenzando en el origen, en cada intervalo de tiempo  $h$ , la partícula salta una distancia  $d$  hacia la derecha con probabilidad  $\frac{1}{2}$  o una distancia  $d$  hacia la izquierda, también con probabilidad  $\frac{1}{2}$ . Llamemos  $S_n$  a la posición de la partícula en el tiempo  $nh$ .

Demuestra que:

$$\lim_{n \rightarrow \infty} P \left[ a < \frac{S_n}{d\sqrt{n}} < b \right] = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{1}{2}x^2} dx$$

Es decir, cuando  $n$  es grande, la distribución de  $S_n$  es aproximadamente normal con parámetros  $\mu = 0$  y  $\sigma^2 = d^2n$ .

**Ejercicio 10.** Sea  $T$  una variable aleatoria con distribución exponencial de parámetro  $\lambda$ . Demuestra que, para cualquier par de números reales positivos  $s$  y  $t$ , se tiene:

$$P [T > t + s \mid T > t] = P [T > s]$$

Supongamos ahora que el tiempo, en años, que un radio funciona, tiene distribución exponencial con parámetro  $\lambda = \frac{1}{8}$ . Si una persona compra un radio usado, ¿cuál es la probabilidad de que ese radio funcionará durante 8 años más?

**Ejercicio 11.** Sean  $T_1$  y  $T_2$  dos variables aleatorias independientes, ambas con distribución exponencial,  $T_1$  con parámetro  $\lambda_1$  y  $T_2$  con parámetro  $\lambda_2$ . Demuestra que  $T = \min(T_1, T_2)$  tiene distribución exponencial con parámetro  $\lambda = \lambda_1 + \lambda_2$ .

**Ejercicio 12.** En una estación meteorológica se cuenta con un tipo de aparato de medición, el cual tiene un tiempo de vida que se distribuye exponencialmente, de tal manera que su tiempo promedio de vida es de 1,000 horas. Si se utilizan 10 de esos aparatos en forma consecutiva, uno de ellos después de que el anterior ya no funciona, a) ¿cuál es la probabilidad de que alguno de los aparatos estará funcionando después de 10,000 horas?, b) si después de 5,000 horas todavía está funcionando alguno de los aparatos, ¿cuál es la probabilidad de que alguno de ellos siga funcionando después de 10,000 horas más?

**Ejercicio 13.** Sea  $X$  una variable aleatoria con distribución normal estándar, muestra que  $X^2$  tiene distribución gama.

**Ejercicio 14.** En Mecánica Estadística, la velocidad  $V$  de una molécula de masa  $m$  en un gas, a temperatura  $T$ , se modela como una variable aleatoria con función de densidad dada por

$$f_V(v) = \begin{cases} \sqrt{\frac{2m^3}{\pi k^3 T^3}} v^2 e^{-\frac{m}{2kT} v^2} & \text{si } v > 0 \\ 0 & \text{en otro caso} \end{cases}$$

donde  $k$  es la constante de Boltzmann. Encuentra una función de densidad de la energía  $E = \frac{1}{2}mV^2$ .

**Ejercicio 15.** Sea  $X$  una variable aleatoria distribuida uniformemente en el intervalo  $(-\frac{\pi}{2}, \frac{\pi}{2})$ . Encuentra la función de distribución y una función de densidad de  $Y = \tan(X)$ .

**Ejercicio 16.** Sea  $T$  una variable aleatoria con distribución exponencial de parámetro  $\lambda$  y  $M > 0$ . Encuentra la función de distribución de  $X = \min\{T, M\}$ .

**Ejercicio 17.** Sea  $A$  un evento que puede ocurrir en algún momento entre los tiempos  $0$  y  $T$  de tal manera que la probabilidad de que ocurra es igual a  $p$  y, en caso de que ocurra, si  $X$  es el tiempo en el cual ocurre, entonces  $X$  tiene distribución uniforme en el intervalo  $[0, T]$ . Sabiendo que el evento  $A$  no ha ocurrido hasta el tiempo  $t \in (0, T)$ , encuentra la probabilidad de que  $A$  ocurra en el intervalo  $(t, T]$ .

**Ejercicio 18.** En una fábrica se producen componentes electrónicos cuyo voltaje tiene distribución normal de parámetros  $\mu$  y  $\sigma^2$ . El departamento de control de calidad checa los componentes y todos aquellos con un voltaje menor o igual a  $b$  son rechazados. Encuentra la distribución del voltaje de los componentes que pasan el control de calidad.

**Ejercicio 19.** Estima la probabilidad de que al lanzar 10,000 veces un dado se obtengan por lo menos 1680 1's.

**Ejercicio 20.** Estima el más pequeño número natural  $k$  tal que la probabilidad de que el número de soles que se obtienen en 1,000 lanzamientos de una moneda esté comprendido entre 480 y  $k$ , inclusive, sea mayor que 0,6.

**Ejercicio 21.** Sea  $T$  una variable aleatoria absolutamente continua con función de densidad continua y tal que, para cualquier par de números reales positivos  $s$  y  $t$ , se tiene:

$$P[T > t + s \mid T > t] = P[T > s]$$

Demuestra que  $T$  tiene distribución exponencial.

**Ejercicio 22.** Sea  $X$  una variable aleatoria con función de distribución  $F_X$ , continua y estrictamente creciente, y definamos la función  $d : (0, 1) \mapsto \mathbb{R}$  mediante la relación:

$$d(t) = \inf \{s \in \mathbb{R} : F_X(s) \geq t\}$$

Demuestra que  $d$  es la inversa de  $F_X$ .

**Ejercicio 23.** Demuestra que si  $Y$  tiene distribución uniforme en el intervalo  $(0, 1)$ , entonces, dado  $p \in (0, 1)$ ,  $\left\lceil \left[ \frac{\ln Y}{\ln(1-p)} \right] \right\rceil$  tiene distribución geométrica con parámetro  $p$ , donde  $\lceil x \rceil$  es el mayor entero menor o igual a  $x$ .

**Ejercicio 24.** Demuestra que si  $p \in (0, 1)$ ,  $\lambda = -\ln(1-p)$  y  $Z$  es una variable aleatoria con distribución exponencial de parámetro  $\lambda$ , entonces,  $\lceil Z \rceil$  tiene distribución geométrica con parámetro  $p$ , donde  $\lceil x \rceil$  es el mayor entero menor o igual a  $x$ .

**Ejercicio 25.** Demuestra que se puede definir, sobre  $((0, 1], \mathfrak{L}((0, 1]), \lambda)$ , una infinidad numerable de sucesiones de variables aleatorias independientes, todas ellas con distribución normal estándar, e independientes entre sí.