Variables aleatorias

PID_00253303

Ana Escudero
Alícia Miralles
Alícia Vila

Tiempo mínimo de dedicación recomendado: 4 horas

Universitat Oberta de Catalunya

Índice

Introducción

Con mucha frecuencia, es necesario relacionar el resultado de una experiencia con un número. Imaginad, por ejemplo, que queremos evaluar la señal de salida de un circuito electrónico o saber cuál es el tiempo de servicio en el que se procesan peticiones de usuario que llegan a un servidor. Una primera aproximación a estos problemas sería considerar que los valores que estamos buscando son deterministas y que, por lo tanto, se pueden definir perfectamente con unos parámetros que nos permiten obtener valores exactos a lo largo del tiempo. En la práctica, no obstante, sabemos que hay muchos factores que hacen que la respuesta de los sistemas de telecomunicación tenga una cierta variabilidad. Así, por ejemplo, en el caso del circuito electrónico que acabamos de mencionar, deberíamos tener en cuenta la presencia de ruido y otras interferencias que hacen que la señal de salida no sea exactamente la esperada. También debemos tener en cuenta que los dispositivos electrónicos no son ideales y que pueden introducir errores. Las variables aleatorias nos permiten tener en cuenta esta variabilidad y modelizar los diferentes resultados que obtenemos en cada experiencia, de forma que podemos prever cuál será el comportamiento de nuestro sistema con una cierta probabilidad.

En el apartado 1 de este módulo, introducimos formalmente el concepto de variable aleatoria. Veremos que hay dos tipos básicos de variable aleatoria: la variable aleatoria discreta (apartado 2), que nos da un conjunto numerable de resultados posibles, y la variable aleatoria continua (apartado 3), que nos da como resultado cualquier número de un intervalo definido dentro de los números reales. Veremos cómo las podemos estudiar y mostraremos los casos de variables aleatorias que aparecen más habitualmente. Trabajaremos, en particular, con las distribuciones más importantes que están relacionadas con las telecomunicaciones. Definiremos los conceptos de valor medio, $\mathrm{E}(X)$ y varianza, $\mathrm{Var}(X)$, de una variable aleatoria X. En el apartado 4 veremos el teorema central del límite, de gran importancia en el campo de la estadística y que nos permite relacionar variables aleatorias continuas y discretas.

Objetivos

Los objetivos de este módulo son los siguientes:

1. Entender qué es una variable aleatoria, diferenciar dos tipos de las mismas y poner ejemplos: las discretas y las continuas.
2. Conocer cuatro tipos de variables aleatorias discretas y saber poner ejemplos: la distribución de Bernoulli, la distribución binomial, la distribución geométrica y la distribución de Poisson.
3. Entender los conceptos de función de distribución y función de densidad, y caracterizar las distribuciones aleatorias con estas funciones.
4. Conocer tres tipos de variables aleatorias continuas y saber poner ejemplos: la distribución uniforme, la distribución exponencial y la distribución normal o de Gauss.
5. Entender los conceptos de valor medio y varianza y caracterizar las distribuciones aleatorias con estos dos parámetros.
6. Eelegir el tipo de distribución aleatoria más adecuada para modelizar un fenómeno determinado.
7. Comprender el sentido del teorema central del límite y sus aplicaciones.

1. Concepto de variable aleatoria

A partir de una experiencia aleatoria, se puede definir el espacio muestral, Ω (omega), como el conjunto de todos los resultados posibles asociados a esta experiencia. Una variable aleatoria, X, asigna un número a cada uno de estos resultados. Veamos algunos ejemplos de ello.

Véase también

El espacio muestral se define en el módulo «Introducción a la probabilidad».

Ejemplo 1.1

Consideremos la experiencia de tirar una moneda. El espacio muestral es $\Omega = \{\text{cara}, \text{cruz}\}$. Observad que el espacio muestral Ω incluye todos los resultados posibles de este experimento. Podemos asignar a cada uno de estos resultados los valores 0 o 1, dependiendo de que el resultado de la experiencia sea cara o cruz. Escribimos, pues, que $X(\text{cara}) = 0$ y $X(\text{cruz}) = 1$. La variable X puede tomar los valores $\{0, 1\}$.

Observación

La variable X del ejemplo 1.1 toma los valores $\{0, 1\}$. Es una variable aleatoria discreta.
La variable X del ejemplo 1.2 toma los valores a $[0, 2]$. Es una variable aleatoria continua.

Ejemplo 1.2

Supongamos que un aparato eléctrico emite una señal aleatoria cada segundo. Esta señal aleatoria se expresa en milivoltios (mV) y toma valores dentro del intervalo $[0, 2]$. En este caso, el espacio muestral está formado por valores numéricos. Podemos definir la variable aleatoria como la aplicación identidad. A cada resultado de la experiencia, le asigna el mismo valor. La variable X puede tomar un valor cualquiera del intervalo $[0, 2]$.

Definición 1.1. Una **variable aleatoria**, X, es una función que asigna un número real a cada elemento del espacio muestral.

En los ejemplos 1.1 y 1.2 vemos la diferencia entre una variable aleatoria discreta y una continua. En el primer caso, tenemos un número determinado de resultados posibles, podemos obtener o bien cara o bien cruz. En el segundo caso, nuestro aparato eléctrico puede emitir un valor cualquiera dentro del intervalo $[0, 2]$.

Conjunto discreto

Un conjunto discreto es aquel que está formado por un número finito de elementos o bien por un número infinito de elementos que son numerables (es decir, que se pueden enumeran de manera que hay un primer elemento, un segundo elemento, etc.). Por ejemplo, los conjuntos de números naturales y enteros ($\mathbb{N}$ y $\mathbb{Z}$) son discretos. El conjunto de los números reales, $\mathbb{R}$, no es un conjunto discreto.

Definición 1.2. Una **variable aleatoria discreta** toma valores de un conjunto finito $\{a_1, a_1, \ldots, a_n\}$ o bien numerable infinito $\{a_1, a_2, \ldots\}$.

Definición 1.3. Una **variable aleatoria continua** puede tomar valores en conjuntos no numerables, como por ejemplo en un intervalo de $\mathbb{R}$ o en todo $\mathbb{R}$.

Esto hace que el tratamiento matemático de las variables aleatorias discretas y continuas sea muy diferente.

2. Variable aleatoria discreta

En este apartado, definiremos las variables aleatorias discretas, veremos las distribuciones más importantes y calcularemos para cada una dos parámetros: el valor medio y la varianza.

En el apartado 1 hemos visto que una variable aleatoria, X, nos da un valor numérico para el resultado de una experiencia. Para cada elemento del espacio muestral Ω, tenemos definido un valor numérico real que es el que toma la variable X cuando el resultado del experimento es este elemento. Al conjunto de valores que puede tomar X lo denominamos Ω_X, que es, por lo tanto, un subconjunto de $\mathbb{R}$. La variable aleatoria X es discreta cuando el conjunto Ω_X es discreto, es decir, finito o infinito numerable.

De manera natural, la probabilidad que tenemos definida en el espacio Ω se traslada a los valores que toma X.

En el ejemplo 1.1, en que el experimento es tirar una moneda al aire, escribimos $P(X=0) = P(\text{cara}) = \frac{1}{2}$ y $P(X=1) = P(\text{cruz}) = \frac{1}{2}$.

Dado que el resultado de la variable aleatoria X varía con cada repetición del experimento, no podemos definir el valor de X, pero sí podemos describir la probabilidad para cada uno de los resultados posibles de X, es decir, la probabilidad de que X tome un valor determinado. Esto se describe como $P(X = a_i)$, donde a_i es un valor posible de X.

Observación

Escribimos $X = 0$ cuando queremos indicar que X toma el valor 0 y $X = 1$ cuando X toma el valor 1. En el ejemplo 1.1, asociamos el valor 0 a sacar cara y el valor 1 a sacar cruz:

$$P(X=0) = \frac{1}{2}$$

$$P(X=1) = \frac{1}{2}$$

Definición 2.1. Se denomina **función de probabilidad** al conjunto de valores $P(X=a_i)$. Esta función asigna una probabilidad a cada valor posible de X. También se escribe como $P_X(a_i)$.

Por ejemplo, en el caso de la experiencia de lanzar una moneda al aire, hemos definido $\Omega = \{\text{cara}, \text{cruz}\}$ y hemos asignado los valores de 0 y 1 a este espacio muestral. En este caso, $\Omega_X = \{0, 1\}$ y la función de probabilidad nos dice que la probabilidad de que salga cara, $P(X = 0)$, es igual a $\frac{1}{2}$ y la probabilidad de que salga cruz, $P(X = 1)$, es igual a $\frac{1}{2}$.

Propiedades de la función de probabilidad de una variable aleatoria discreta:

1) La probabilidad $P(X=a_i)$ es un valor que está siempre entre 0 y 1, es decir, $0 \leq P(X=a_i) \leq 1$.

2) La suma de todas las probabilidades tiene que ser 1, puesto que los sucesos $X=a_i, \forall i$ forman una partición del espacio muestral. Es decir,

$$\sum_i P(X=a_i) = 1. \tag{1}$$

Definición 2.2. X es una **variable aleatoria discreta uniforme** si toma los valores $\{a_1, a_2, \ldots, a_n\}$ con probabilidades

$$P(X=a_i) = \frac{1}{n}, \quad i = 1, 2, \ldots, n.$$

Es decir, todos los elementos a_i tienen la misma probabilidad asignada.

El valor $1/n$ se deduce teniendo en cuenta que, según (1), si la probabilidad de cada resultado es un valor común p, la suma de estas probabilidades es $\sum_{i=1}^{n} p = np = 1$.

La uniformidad suele ir vinculada a la simetría de ciertos experimentos. Por ejemplo, el uso de monedas y dados simétricos. La variable que vale 0 o 1 según salga cruz o cara al lanzar una moneda al aire, o la variable que da el resultado obtenido al lanzar un dado, son ejemplos de variables aleatorias discretas uniformes. En el caso de la moneda, la probabilidad de obtener 0 o 1 es $\frac{1}{2}$. En el caso del dado, la probabilidad de obtener cualquiera de los resultados es $\frac{1}{6}$.

2.1. Variables aleatorias discretas más importantes

2.1.1. Variable aleatoria de Bernoulli: B(p)

Este es el tipo más sencillo de variable aleatoria discreta, y se utiliza para representar experiencias en las que solo podemos tener dos resultados posibles.

Partimos de una experiencia aleatoria y distinguimos entre los resultados $\Omega = \{A, A^c\}$. El resultado A se denomina **éxito** y definimos $X(A) = 1$. El resultado A^c se denomina **no éxito** y definimos $X(A^c) = 0$. La variable aleatoria toma solo los dos valores $\{0, 1\}$. Solo hay que dar la probabilidad asignada a uno de estos valores y la distribución queda definida completamente. Si $P(A) = p$, entonces $P(X=1) = p$ y $P(X=0) = 1 - p$.

A y A^c

Los conjuntos A y A^c son complementarios. Recordad lo que habíamos visto en el módulo «Introducción a la probabilidad» (apartado 2.2, fórmula 12): si $P(A) = p$, la probabilidad del conjunto complementario, $P(A^c)$, es $1 - p$.

Decimos que X es una **variable aleatoria de Bernoulli** con probabilidad de éxito p cuando esta variable puede tomar los valores $X = 1$ (éxito) con probabilidad p y $X = 0$ (no éxito) con probabilidad $(1 - p)$.

Se escribe: $X \sim \mathrm{B}(p)$, donde p indica la probabilidad de éxito y $p \in [0, 1]$. Diremos que se ha producido éxito cuando el resultado obtenido esté dentro del conjunto A, es decir, si pasa A, y no éxito cuando el resultado obtenido esté en el conjunto A^c (complementario de A), es decir, si no sucede A.

La variable del ejemplo 1.1 de este módulo sigue una distribución $\mathrm{B}(\frac{1}{2})$. Volvamos al caso de la moneda. Si definimos éxito, A, que salga cara, y el no éxito, A^c, que salga cruz, nuestra variable aleatoria, X, sigue una distribución de Bernoulli $\mathrm{B}(\frac{1}{2})$. Notad que también podríamos haber definido A como cruz y A^c como cara.

Ejemplo 2.1

En comunicaciones binarias, X puede indicar el error en la transmisión de un bit. El espacio muestral de la experiencia es determinado por $\Omega = \{\text{error}, \text{no error}\}$ y la variable aleatoria toma los valores $X = 1$ si hay error y $X = 0$ si no lo hay. $P(X=1) = P(\text{error})$ y $P(X=0) = P(\text{no error})$. Observad que en este ejemplo hemos definido éxito, A, como la presencia de error en la transmisión.

Otro ejemplo similar de aplicación de la distribución de Bernoulli en las telecomunicaciones se da en los sistemas de radar. Podemos definir éxito, A (es decir, $X = 1$), cuando el radar detecta la presencia de un objeto y A^c $(X = 0)$ cuando el radar no detecta ningún objeto.

Ejemplo 2.2

Al lanzar un dado, nos fijamos en la máxima puntuación posible y definimos $A = \{\text{sale un } 6\}$, entonces $A^c = \{\text{no sale un } 6\}$. Definimos $X(A) = 1$ y $X(A^c) = 0$. La variable X sigue una distribución $\mathrm{B}(\frac{1}{6})$, puesto que la probabilidad de lo que hemos definido como éxito es $\frac{1}{6}$.

Observación

En el ejemplo 2.2:

$$P(X=0) = \frac{5}{6}$$

$$P(X=1) = \frac{1}{6}$$

2.1.2. Variable aleatoria binomial: Bin(n, p)

La **variable aleatoria binomial**, Bin(n, p), se da cuando repetimos n veces y de manera independiente una experiencia B(p) de Bernoulli. A cada resultado (secuencia de n valores 0 o 1), la variable aleatoria X le asigna el número de éxitos que han salido. Así, X toma los valores $\{0, 1, 2, \ldots, n\}$.

La función de probabilidad (probabilidad que tiene cada uno de los valores que toma la variable X) es determinada por:

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{con} \quad k \in \{0, 1, 2, \ldots, n\}. \qquad (2)$$

Decimos que X es una variable aleatoria binomial y escribimos $X \sim$ Bin(n, p), donde n es el número de veces que repetimos la experiencia de Bernoulli B(p), de probabilidad de éxito p.

La fórmula (2) corresponde al hecho de que tenemos que poner un factor p para cada éxito resultante (p^k, por lo tanto) y un factor $(1-p)$ para cada no éxito ($(1-p)^{n-k}$, por lo tanto). El factor $\binom{n}{k}$ se debe a que fijar el número k de éxitos todavía deja libertad para situar estos éxitos en la secuencia de resultados n. El número de maneras de elegir las k posiciones de los éxitos es el anterior número combinatorio, puesto que la selección la hacemos sin orden y sin repetición.

Recordad que $\binom{n}{k}$ (se lee n sobre k) se calcula como $\frac{n!}{k!(n-k)!}$.

En el ejemplo 2.3, X sigue una distribución Bin$(10, 0{,}1)$, en la que 10 es el número de veces que se repite el experimento y $0{,}1$, la probabilidad de éxito de cada experimento.

Ejemplo 2.3

Una persona, emisor, tiene que mandar un mensaje de 10 elementos, elegidos del conjunto $\{0, 1\}$ y ordenados. Un mensaje de este tipo podría ser la palabra 0011111101 (formada con 10 bits). Suponemos que cada vez que la persona elige un bit para formar la palabra, la probabilidad de que sea un 0 es 0,1 y, por lo tanto, la de que sea un 1 es 0,9. En este ejemplo, consideraremos que se da la condición de éxito, A, cuando se transmite un 0 y la condición de no éxito, A^c, cuando se transmite un 1. Con esta idea, nos vienen a la cabeza toda una serie de preguntas, como por ejemplo:

1) ¿Cuál es la probabilidad de que el emisor envíe exactamente la palabra 0011111101?

2) ¿Cuál es la probabilidad de que el emisor envíe una palabra que tenga exactamente tres ceros?

3) ¿Cuál es la probabilidad de que el emisor envíe una palabra que tenga exactamente k ceros?

4) ¿Cuál es la probabilidad de que el emisor envíe una palabra que tenga como máximo tres ceros?

Las cuestiones anteriores las podemos resolver aplicando lo que habéis aprendido en el tema anterior. Se obtienen los resultados siguientes:

1) $0{,}1^3 \cdot 0{,}9^7 = 0{,}00048$.

   La palabra para enviar tiene tres ceros y, por lo tanto, la probabilidad de que en 10 experiencias obtengamos 3 éxitos es: $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ En este caso, nos piden una combinación concreta de todas las posibles que podrían incluir 3 ceros y, por lo tanto, no tenemos en cuenta el término $\binom{n}{k}$ y $P = 0{,}1^3 \cdot 0{,}9^{10-3}$.

2) El resultado anterior nos da la probabilidad para una palabra que tiene 3 ceros y 7 unos. Vimos en el apartado anterior que el número de palabras que se pueden formar con 3 ceros y 7 unos es $\binom{10}{3}$. Entonces, la respuesta es $\binom{10}{3} 0{,}1^3 \cdot 0{,}9^7 = 0{,}0574$.

   En este caso, nos piden la probabilidad de una secuencia con 3 ceros. Estos ceros pueden estar en cualquier posición y, por lo tanto, debemos considerar todas las combinaciones posibles de palabras de 10 bits que pueden contener estos ceros.

3) Está claro que una palabra de tamaño 10 puede tener entre 0 y 10 ceros. Así, $0 \le k \le 10$. Haciendo el mismo razonamiento que en el apartado anterior, tenemos $\binom{10}{k} 0{,}1^k \cdot 0{,}9^{10-k}$. Observad que k es el número de ceros que puede contener la palabra, no el número de experiencias, que para este caso es 10.

4) Ahora debemos tener en cuenta aquellos casos en los que el número de ceros sea menor o igual que 3. De la expresión anterior, sumamos los casos en los que k toma los valores 0, 1, 2 y 3. Obtenemos:

$$\sum_{k=0}^{3} \binom{10}{k} 0{,}1^k \cdot 0{,}9^{10-k} = \binom{10}{0} 0{,}1^0 \cdot 0{,}9^{10-0} + \binom{10}{1} 0{,}1^1 \cdot 0{,}9^{10-1}$$

$$+ \binom{10}{2} 0{,}1^2 \cdot 0{,}9^{10-2} + \binom{10}{3} 0{,}1^3 \cdot 0{,}9^{10-3}$$

$$= 0{,}3487 + 0{,}3874 + 0{,}1937 + 0{,}0574 = 0{,}9872.$$

Si a la variable X le asignamos el número de ceros que tiene cada palabra, los apartados anteriores los podemos escribir utilizando la $X \sim \text{Bin}(10,\ 0{,}1)$:

- $P(X = 3) = \binom{10}{3} 0{,}1^3 \cdot 0{,}9^7$.
- $P(X = k) = \binom{10}{k} 0{,}1^k \cdot 0{,}9^{10-k}$.
- $P(X \le k) = \sum_{i=0}^{k} \binom{10}{i} 0{,}1^i \cdot 0{,}9^{10-i}$.

Ejemplo 2.4

Enviamos una palabra de n bits en la que cada bit puede llevar error o no, independientemente de los otros. La variable $X \sim \text{Bin}(n, p)$ toma el valor del número de bits erróneos que hay en la palabra y, por lo tanto, los valores posibles son $\{0, 1, 2, \ldots, n\}$. p es la probabilidad de que un bit sea erróneo.*

* Recordad que hemos definido el éxito A como el hecho de transmitir un bit erróneo.

2.1.3. Variable aleatoria geométrica: Geom(p)

La **variable aleatoria geométrica**, Geom(p), se da cuando repetimos, de manera independiente, una experiencia B(p), hasta obtener el primer éxito. X cuenta el número de veces que hay que hacer la experiencia para obtener el primer éxito. Por lo tanto, X toma los valores $\{1, 2, 3, \ldots\}$.

Su distribución de probabilidades es determinada por:

$$P(X=k) = (1-p)^{k-1}p \quad \text{amb} \quad k \in \{1,2,3,\ldots\}. \tag{3}$$

en la que p es la probabilidad de éxito y k, el número de intentos que necesitamos hasta obtener el éxito. Decimos que X es una variable aleatoria geométrica con probabilidad de éxito p, y escribimos $X \sim \text{Geom}(p)$.

La fórmula (3) se obtiene considerando que para que el primer éxito pase en la posición k se deben obtener $k-1$ no éxitos seguidos (factor $(1-p)^{k-1}$) y, a continuación, un éxito (factor p).

En el ejemplo que veremos a continuación, X sigue una distribución Geom(0,2). Observad que la expresión de la distribución geométrica es similar a la expresión de la distribución binomial, pero sin el término $\binom{n}{k}$, puesto que en este caso estamos fijando la secuencia de resultados como A^c,A^c,...,A^c,A.

Ejemplo 2.5

Para enviar mensajes por internet, los mensajes se dividen en paquetes y después se envían por la red. Si la red está congestionada, los paquetes se pueden perder. Supongamos que en una red muy congestionada, la probabilidad de perder un paquete es 0,8. Esto significa que el paquete no se pierde en una transmisión con una probabilidad de 0,2. El paquete se transmite repetidamente hasta que el receptor lo recibe. Nos hacemos las preguntas siguientes:

1) ¿Cuál es la probabilidad de que el paquete tenga que ser enviado al menos tres veces?
2) ¿Cuál es la probabilidad de que tengamos que enviar el paquete como máximo 5 veces para que el receptor lo reciba?

Si la variable aleatoria X cuenta el número de veces que hay que enviar un paquete, toma valores del conjunto $\{1,2,3,\ldots\}$ (X es una variable discreta infinita). Podemos escribir:

1) $P(X \geq 3) = 1 - P(X \leq 2) = 1 - (P(X=1) + P(X=2)) = 1 - (0{,}2 + 0{,}8 \cdot 0{,}2) = 0{,}64.$
2) $P(X \leq 5) = \sum_{k=1}^{5} 0,2 \cdot 0,8^{k-1} = 0{,}2(1 + 0{,}8 + 0{,}8^2 + 0{,}8^3 + 0{,}8^4) = 0{,}67232.$

2.1.4. Variable aleatoria de Poisson: Poiss(α)

La variable aleatoria de Poisson se utiliza para modelizar algunos fenómenos, como los siguientes:

- El número de accidentes en un cruce dado y para un intervalo de tiempo fijado.
- El número de llamadas que llegan a una centralita en un cierto intervalo de tiempo.

- El número de peticiones que llegan a un servidor en un cierto intervalo de tiempo.
- El número de electrones o agujeros que atraviesan una barrera de potencial.
- El número de defectos de fabricación de un producto de unas dimensiones determinadas.
- Teoría de colas en redes de comunicaciones de voz y datos.

Contamos sucesos que se producen en posiciones aleatorias de un cierto intervalo de tamaño T (típicamente, sucesos que pasan en instantes aleatorios a lo largo de un tiempo total T). La variable aleatoria X da el número total de sucesos y toma los valores $\{0, 1, 2, 3, \ldots\}$

La función de probabilidad para una **variable aleatoria de Poisson** es:

$$P(X=k) = \frac{\alpha^k}{k!}e^{-\alpha} \quad \text{amb} \quad k \in \{0, 1, 2, 3, \ldots\}. \tag{4}$$

Donde $\alpha = \lambda T$ es el número medio de sucesos en el intervalo T. Por lo tanto, λ es el número medio de sucesos por unidad de tiempo (tasa).

Decimos que X es una variable aleatoria de Poisson de parámetro α, y escribimos $X \sim \text{Poiss}(\alpha)$.

Ejemplo 2.6

Sabemos que a un servidor llegan de media 5 peticiones por segundo. ¿Cuál es la probabilidad de que en un segundo no llegue ninguna petición? ¿Cuál es la probabilidad de que en un segundo lleguen una o más peticiones?

Según el enunciado, $\lambda = 5$ y $T = 1$; así, $\alpha = 5$. Ahora ya podemos dar las respuestas. En el primer caso, tenemos que calcular la probabilidad de que el número de llegadas, k, sea cero; por lo tanto:

$$P(X=0) = \frac{5^0}{0!}e^{-5} = e^{-5} = 0{,}0067.$$

En este segundo caso, tenemos que calcular la probabilidad de que lleguen una o más peticiones al servidor:

$$P(X \geq 1) = 1 - P(X=0) = 1 - 0{,}0067 = 0{,}9933.$$

2.2. Parámetros: valor medio y varianza

Hasta ahora, en este apartado hemos definido qué es una variable aleatoria discreta y hemos visto algunas de las distribuciones más importantes: la distribución de Bernoulli, la binomial, la geométrica y la de Poisson. En este subapartado, veremos dos parámetros muy utilizados que nos permitirán, de manera muy global, evaluar y comparar las diferentes variables aleatorias. Estos parámetros son el **valor medio** y la **varianza**.

Definición 2.3. Sea X una variable aleatoria discreta que toma los valores $\{a_1, a_2, \ldots, a_n\}$ (para el caso de que X tome infinitos valores se hace de manera parecida, pero en lugar de sumas finitas, tenemos series numéricas).

El **valor medio, esperanza** o **momento de orden 1** de X:

$$\mathrm{E}(X) = \sum_{i=1}^{n} a_i P(X = a_i). \tag{5}$$

El **momento de orden 2**:

$$\mathrm{E}(X^2) = \sum_{i=1}^{n} a_i^2 P(X = a_i). \tag{6}$$

El **momento de orden** k, $(k = 0, 1, 2, 3, \ldots)$:

$$\mathrm{E}(X^k) = \sum_{i=1}^{n} a_i^k P(X = a_i). \tag{7}$$

La esperanza de X es un número que nos da la posición en torno a la cual se concentra la variable. Un significado más preciso es el siguiente. Si repetimos el experimento un número grande de veces N, obtenemos valores para la variable X: $X_1, X_2, \ldots, X_N$. En esta lista, aparece cada valor posible a_i un número N_i de veces. Si hacemos la media aritmética de todos los resultados:

$$\frac{1}{N}(X_1 + X_2 + \cdots + X_N) = \frac{1}{N}(a_1 N_1 + a_2 N_2 + \cdots + a_n N_n) = \sum_i a_i \frac{N_i}{N}.$$

Cuando N es muy grande, las frecuencias relativas $\frac{N_i}{N}$ se estabilizan en los valores $P(a_i)$, de forma que la media aritmética anterior tiende a la esperanza de la variable X.

Definición 2.4. La **varianza** de una variable aleatoria discreta X es:

$$\operatorname{Var}(X) = \operatorname{E}((X - \operatorname{E}(X))^2) = \sum_{i=1}^{n} (a_i - \operatorname{E}(X))^2 P(X = a_i). \tag{8}$$

La **desviación típica** de X es:

$$\sigma(X) = \sqrt{\operatorname{Var}(X)}. \tag{9}$$

La varianza es una media de las distancias de los valores de X al valor medio. El cuadrado se pone para que todas las desviaciones se cuenten con signo positivo. Con la desviación típica, recuperamos las dimensiones originales de X aplicando una raíz cuadrada a la varianza. La varianza o la desviación dan una medida de la dispersión de X. Si las probabilidades se concentran mucho en torno a $\operatorname{E}(X)$, la dispersión será pequeña.

Observación

La desviación típica se representa con la letra σ, que se lee «sigma».

Varianza y desviación típica

Notad que la varianza es una media de las diferencias al cuadrado, mientras que la desviación típica es el mismo parámetro pero dado en las mismas unidades que la variable aleatoria. La relación entre ellas es: $\sigma(X) = \sqrt{\operatorname{Var}(X)}$.

Propiedades de la varianza.

$$\operatorname{Var}(X) \geq 0. \tag{10}$$

$$\operatorname{Var}(X) = \operatorname{E}(X^2) - \operatorname{E}(X)^2. \tag{11}$$

En efecto, en (8) vemos que la suma que define $\operatorname{Var}(X)$ solo contiene términos positivos (números al cuadrado y probabilidades) de donde se obtiene (10). Desarrollando el cuadrado en (8):

$$\operatorname{Var}(X) = \sum_i (a_i - \operatorname{E}(X))^2 P(X = a_i) = \sum_i (a_i^2 - 2\operatorname{E}(X) a_i + \operatorname{E}(X)^2) P(X = a_i)$$

$$= \sum_i a_i^2 P(X = a_i) - 2\operatorname{E}(X) \sum_i a_i P(X = a_i) + \operatorname{E}(X)^2 \sum_i P(X = a_i)$$

$$= \operatorname{E}(X^2) - 2\operatorname{E}(X) \cdot \operatorname{E}(X) + \operatorname{E}(X)^2 = \operatorname{E}(X^2) - \operatorname{E}(X)^2,$$

obtenemos (11), que nos dice que la varianza es la esperanza del cuadrado de X menos el cuadrado de la esperanza de X. Esta es la manera más habitual de calcular varianzas, puesto que el segundo momento suele ser más fácil de calcular que la suma en (8).

Ejemplo 2.7

Sean X_1 y dos X_2 variables aleatorias equiprobables que toman los valores $\{4, 5, 6\}$ y $\{0, 5, 10\}$, respectivamente. Estos valores podrían ser las tres notas obtenidas en una determinada asignatura por dos alumnos diferentes. Suponiendo que las tres notas tienen el mismo peso, podemos estimar la esperanza y varianza de cada una de las variables:

$$\mathrm{E}(X_1) = 4 \cdot \frac{1}{3} + 5 \cdot \frac{1}{3} + 6 \cdot \frac{1}{3} = 5.$$

$$\mathrm{E}(X_2) = 0 \cdot \frac{1}{3} + 5 \cdot \frac{1}{3} + 10 \cdot \frac{1}{3} = 5.$$

$$\mathrm{Var}(X_1) = (4-5)^2 \cdot \frac{1}{3} + (5-5)^2 \cdot \frac{1}{3} + (6-5)^2 \cdot \frac{1}{3} = 0{,}66, \quad \sigma_{X_1} = 0{,}82.$$

$$\mathrm{Var}(X_2) = (0-5)^2 \cdot \frac{1}{3} + (5-5)^2 \cdot \frac{1}{3} + (10-5)^2 \cdot \frac{1}{3} = 16{,}67, \quad \sigma_{X_2} = 4{,}08.$$

Lo que podemos decir es que los dos alumnos tienen la misma nota de media $\mathrm{E}(X_1) = \mathrm{E}(X_2)$, pero el segundo alumno presenta más dispersión en sus notas, puesto que $\sigma_{X_2} > \sigma_{X_1}$ (es decir, el segundo alumno es menos regular).

Figura 1. Notas obtenidas por los dos estudiantes

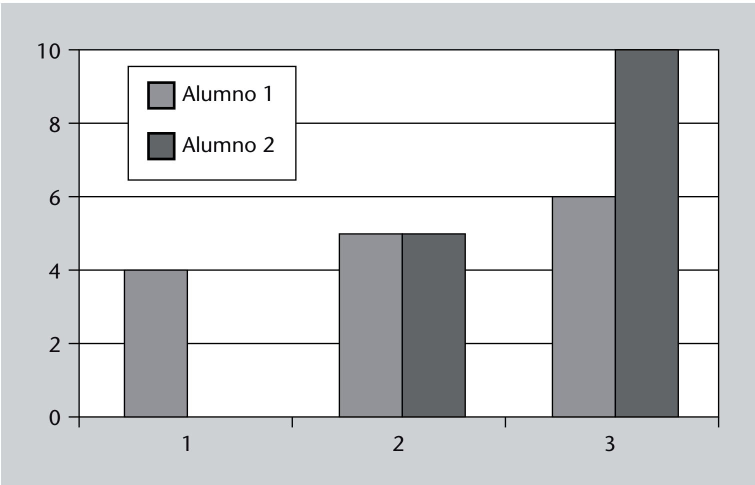


Figura 1

Distribución de las notas obtenidas por los dos estudiantes. En el eje horizontal se representa el número de prueba, y en el eje vertical, la nota obtenida. Los dos estudiantes obtienen la misma nota media, pero observad que el primero es más regular (tiende a sacar notas en un intervalo menor) que el segundo, que obtiene resultados más dispersos.

Parámetros de las principales variables aleatorias discretas

Para cada una de las distribuciones vistas, se obtienen los valores de la esperanza y la varianza de la tabla siguiente.

Distribuciones de variables aleatorias discretas

$X \sim$	k	$P(X=k)$	$\mathrm{E}(X)$	$\mathrm{Var}(X)$
$\mathrm{B}(p)$	$\{0,1\}$	$P(X=1)=p$ $P(X=0)=1-p$	p	$p(1-p)$
$\mathrm{Bin}(n,p)$	$\{0,1,2,\ldots,\ n\}$	$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$	np	$np(1-p)$
$\mathrm{Geom}(p)$	$\{1,2,3,\ldots\}$	$P(X=k)=(1-p)^{k-1}p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
$\mathrm{Poiss}(\alpha)$	$\{0,1,2,\ldots\}$	$P(X=k)=\frac{\alpha^k}{k!}e^{-\alpha}$	α	α

Comprobamos los resultados de la tabla anterior, que corresponden a $X \sim \mathrm{B}(p)$:

- El valor medio, $\mathrm{E}(X) = 1 \cdot p + 0 \cdot (1-p) = p$.
- El momento de orden 2, $\mathrm{E}(X^2) = 1^2 \cdot p + 0^2 \cdot (1-p) = p$.
- La varianza, $\mathrm{Var}(X) = p - p^2 = p(1-p)$.

2.3. Función de distribución

Una manera de dar el valor de las probabilidades acumuladas es a partir de la función de distribución.

Definición 2.5. La **función de distribución** de una variable aleatoria X se define como

$$F_X(x) = P(X \leq x), \quad \forall x \in \mathbb{R}. \tag{12}$$

Es decir, la función de distribución nos da la probabilidad de que nuestra variable aleatoria X tome un valor igual o menor que un valor x determinado. Veamos un ejemplo de ello.

Ejemplo 2.8

Para el caso de $X \sim \mathrm{B}(p)$, la función de distribución $F_X(x)$ presenta una discontinuidad de salto en $X = 0$ y en $X = 1$:

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1-p & \text{si } 0 \leq x < 1 \\ 1 & \text{si } 1 \leq x \end{cases}$$

Consideramos aquí que cero es no éxito.

Figura 2. Función de distribución de $X \sim \mathrm{B}(p)$

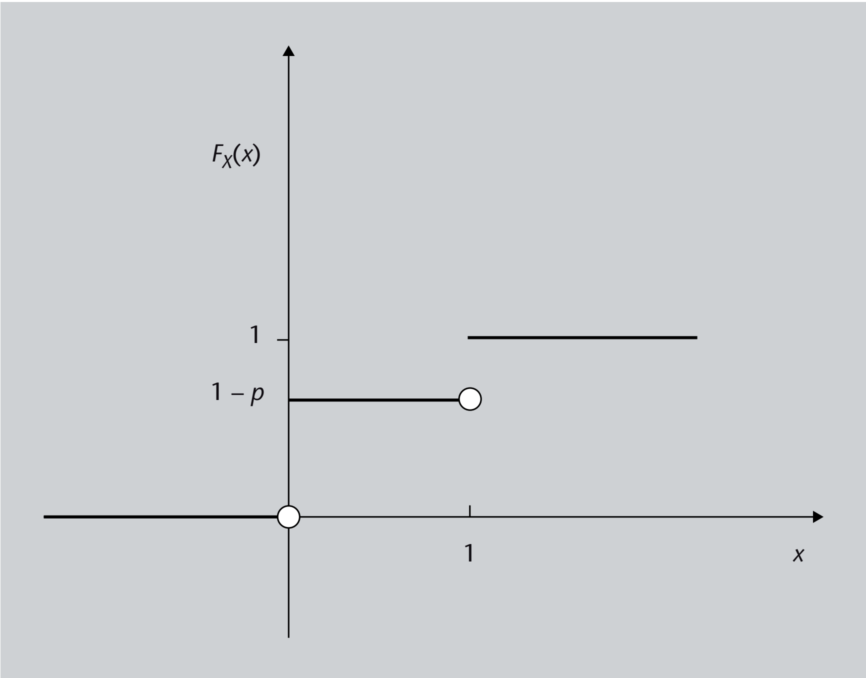


Figura 2

Recordad del subapartado 2.1.1 de este módulo que el espacio muestral de la distribución de Bernoulli se definía como $\Omega = \{A^c, A\} = \{0, 1\}$. En la figura, se ha definido $1-p$ como la probabilidad de no éxito, $P(X=0)$, y p como la probabilidad de éxito $P(X=1)$.

Observad los intervalos definidos en la figura 2. El resultado del experimento aleatorio nos tiene que dar no éxito o éxito, es decir, 0 o 1. Por lo tanto, la probabilidad de obtener un número menor que 0 como resultado es nula. En el segundo intervalo de la gráfica, tenemos la probabilidad de obtener un cero (no éxito), y esta probabilidad es $1-p$. Para la última parte de la gráfica ($x \geq 1$), estamos considerando la probabilidad acumulada de obtener o bien 0 o bien 1. Puesto que sabemos seguro que obtendremos uno de los dos resultados, la función de distribución vale 1 a partir de este punto. En los casos de variables aleatorias discretas, la función de distribución es escalonada. La función experimenta un salto en cada número real que corresponda a un valor que toma X.

Función de distribución

La función de distribución nos da la probabilidad de que una variable aleatoria, X, tenga un valor menor o igual que una x determinada. Notad que, por esta razón, también se denomina **probabilidad acumulada**, es decir, probabilidad de todos los valores hasta x.

Veamos ahora un segundo ejemplo: la función de distribución por la variable aleatoria binomial.

Ejemplo 2.9

En el caso $X \sim \mathrm{Bin}(4, 1/2)$, $F_X(x)$ presenta una discontinuidad de salto en los valores del conjunto $\{0, 1, 2, 3, 4\}$. Teniendo en cuenta que $P(X=k) = \binom{4}{k}(\frac{1}{2})^k(\frac{1}{2})^{4-k}$ y que

$$F_X(x) = P(X \leq k) = \sum_{i=0}^{k} \binom{4}{i} \left(\frac{1}{2}\right)^i \left(\frac{1}{2}\right)^{4-i} = \sum_{i=0}^{k} \binom{4}{i} \left(\frac{1}{2}\right)^4,$$

obtenemos los valores más significativos de la función de distribución:

$$\begin{aligned}
F_X(0) &= 0{,}0625.\\
F_X(1) &= 0{,}0625+0{,}250=0{,}3125.\\
F_X(2) &= 0{,}0625+0{,}250+0{,}375=0{,}6875.\\
F_X(3) &= 0{,}0625+0{,}2500+0{,}3750+0{,}2500=0{,}9375.\\
F_X(4) &= 0{,}0625+0{,}2500+0{,}3750+0{,}2500+0{,}0625=1.
\end{aligned}$$

Distribución binomial

Recordad, como hemos visto en el subapartado 2.1.2 de este módulo, que la distribución binomial consiste en repetir n veces un experimento de Bernoulli. En este caso, nuestra variable aleatoria contabiliza el número de éxitos que obtenemos al hacer el experimento n veces.

$$F_X(x)=\begin{cases} 0 & \text{si } x<0\\ 0{,}0625 & \text{si } 0\leq x<1\\ 0{,}3125 & \text{si } 1\leq x<2\\ 0{,}6875 & \text{si } 2\leq x<3\\ 0{,}9375 & \text{si } 3\leq x<4\\ 1 & \text{si } 4\leq x \end{cases}$$

Figura 3. Función de distribución de $X \sim \text{Bin}(4, \frac{1}{2})$

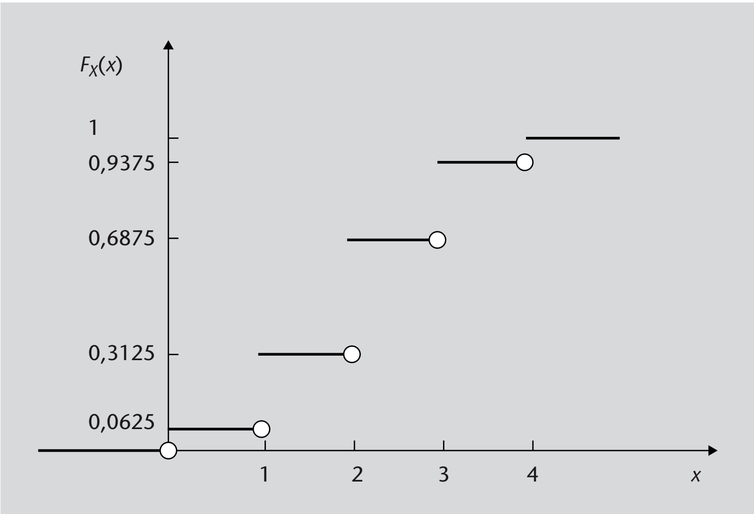


Figura 3

Notad cómo la función de distribución es un valor que va acumulando la probabilidad de obtener un cierto número de éxitos iguales o menores que x.

Observad la figura. La probabilidad de obtener un número de éxitos menor que cero es nula, porque si hacemos el experimento de Bernoulli 4 veces obtendremos o bien 0 éxitos o bien 1, 2, 3 o 4, pero en ningún caso un valor negativo. Por lo tanto, $F(x<0)=0$.

La probabilidad de obtener 0 éxitos es la siguiente:

$$P(X=0)=\binom{4}{0}\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^{4-0}=\frac{1}{16}=0{,}0625.$$

Por lo tanto, en $x=0$ tenemos un salto de la función de distribución, $F(x)$, que pasa de valer cero a tener el valor de 0,0625.

Ahora calculamos la probabilidad de tener un éxito en los cuatro experimentos. Esta probabilidad está determinada por la expresión siguiente:

$$P(X=1)=\binom{4}{1}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^{4-1}=\frac{1}{16}=0{,}25.$$

Así pues, la función de distribución, $F(x)$, en $x = 1$ es la probabilidad acumulada de obtener cero éxitos o un éxito, es decir, $F(1) = 0{,}0625 + 0{,}25 = 0{,}3125$.

Si hacemos los cálculos para el resto de los valores de x, en los que x es el número de éxitos para los cuatro experimentos, obtenemos los valores que se muestran en la figura 3.

Notamos que $F_X(x)$ es una probabilidad, de forma que toma valores entre 0 y 1. Además, dados $a < b$, de la descomposición disjunta $(-\infty, b] = (\infty, a] \cup (a, b]$ obtenemos $P(X \leq b) = P(X \leq a) + P(a < X \leq b)$, es decir,

$$F_X(b) = F_X(a) + P(a < X \leq b).$$

Puesto que $P(a < X \leq b) \geq 0$, tenemos que $F_X(b) \geq F_X(a)$, así que F_X es una función creciente. Además, si $x \to -\infty$, $F_X(x)$ da la probabilidad de $\varnothing$, es decir, 0. Si $x \to \infty$, $F_X(x)$ da la probabilidad de todo $\mathbb{R}$, es decir, 1. Por lo tanto, se verifican las siguientes propiedades:

Propiedades de la función de distribución

1) $0 \leq F_X(x) \leq 1$.
2) $\lim\limits_{x\to\infty} F_X(x) = 1$.
3) $\lim\limits_{x\to-\infty} F_X(x) = 0$.
4) $F_X(x)$ es creciente, es decir, si $a < b$ entonces $F_X(a) \leq F_X(b)$.
5)

$$P(a < X \leq b) = F_X(b) - F_X(a). \tag{13}$$

Las cuatro primeras propiedades de la función de distribución, F_X, se observan fácilmente en la figura 3. Los valores de las probabilidades acumuladas los obtenemos directamente de la función de distribución. Damos algunos valores de esto como ejemplo:

$$\begin{aligned}
P(X \leq 0{,}5) &= F_X(0{,}5) = 0{,}0625. \\
P(X \leq 1{,}7) &= F_X(1{,}7) = 0{,}3125. \\
P(X \leq 2{,}4) &= F_X(2{,}4) = 0{,}6875. \\
P(1,7 < X \leq 2{,}4) &= F_X(2{,}4) - F_X(1{,}7) = 0{,}375.
\end{aligned}$$

En el caso de variables discretas, solo nos interesa conocer la función de distribución en los valores que puede tomar la variable, puesto que F_X se mantiene constante entre uno de estos valores y el siguiente.

Consideramos la probabilidad de un punto aislado cualquiera a. Para calcular la probabilidad $P(X = a)$, calculamos la de un intervalo $(a - \epsilon, a]$ utilizando (13) y hacemos $\epsilon \to 0^+$:

$$P(X = a) = \lim_{\epsilon \to 0^+} P(a - \epsilon < X \leq a) = \lim_{\epsilon \to 0^+} (F_X(a) - F_X(a - \epsilon)). \qquad (14)$$

El anterior límite es el salto que da la función F_X en el punto a. Como hemos visto en los ejemplos anteriores, F_X tiene una discontinuidad de salto en cada punto de los posibles. La magnitud del salto es igual a la probabilidad del punto. En cambio, los puntos en los que F_X es continua tienen probabilidad cero.

Si una variable aleatoria discreta X puede tomar los valores a_i ordenados de manera creciente según el índice i:

$$F_X(a_i) = \sum_{j \leq i} P(X = a_j). \qquad (15)$$

En el caso de una variable $X \sim \mathrm{Bin}(n, p)$, tenemos:

$$F_X(k) = \sum_{l=0}^{k} \binom{n}{l} p^l (1-p)^{n-l}.$$

Para una variable $X \sim \mathrm{Poiss}(\alpha)$, tenemos:

$$F_X(k) = \sum_{l=0}^{k} \frac{\alpha^l}{l!} e^{-\alpha}.$$

En los dos casos anteriores, no es posible expresar en forma compacta el resultado de los sumatorios. Tenemos que recurrir a software matemático o hacer la suma cuando el número de términos es pequeño.

Un caso en el que sí podemos calcular el sumatorio que da F_X es el de la variable aleatoria geométrica.

Sí $X \sim \mathrm{Geom}(p)$:

$$F_X(k) = 1 - (1-p)^k \quad \text{con} \quad k \in \{1, 2, 3, \ldots\}. \qquad (16)$$

Este se demuestra haciendo el sumatorio $\sum_{l=1}^{k}(1-p)^l p$ (suma de tipo geométrico) o con el siguiente razonamiento: $F_X(k) = P(X \leq k) = 1 - P(X > k) = 1 - (1-p)^k$, puesto que $X > k$ equivale a que en las primeras k realizaciones se ha obtenido no éxito.

Ejemplo 2.10

Se lanzan dos dados repetidamente, hasta que se obtiene el doble seis. Sea X la variable que cuenta el número de lanzamientos.

1) Calculamos las siguientes probabilidades:

- Que hagan falta 10 tiradas o menos.
- Que hagan falta 20 tiradas o más.
- Que hagan falta entre 30 y 40 tiradas.

La probabilidad del doble seis es 1/36. El número de tiradas es $X \sim \text{Geom}(1/36)$. Su función de distribución, según (16), es $F_X(k) = 1 - (\frac{35}{36})^k$.
La primera probabilidad pedida es:

$$P(X \leq 10) = F_X(10) = 1 - \left(\frac{35}{36}\right)^{10} = 0{,}2455.$$

La segunda probabilidad es:

$$P(X \geq 20) = 1 - P(X \leq 19) = 1 - F_X(19) = \left(\frac{35}{36}\right)^{19} = 0{,}5855.$$

La tercera probabilidad es:

$$P(30 \leq X \leq 40) = P(29 < X \leq 40) = F_X(40) - F_X(29)$$

$$= \left(\frac{35}{36}\right)^{29} - \left(\frac{35}{36}\right)^{40} = 0{,}1177.$$

2) ¿Cuál es el número medio de tiradas que hay que hacer?
Se trata de $\mathrm{E}(X)$. En la tabla de parámetros, vemos que una variable $\text{Geom}(p)$ tiene esperanza $1/p$. Así, en nuestro caso $\mathrm{E}(X) = 36$.

3) ¿Cuál es el número mínimo de tiradas que nos asegura una probabilidad de al menos un 90 % de obtener el doble seis?
La probabilidad de que el doble seis salga como muy tarde en la jugada N-ésima es $P(X \leq N) = F_X(N)$. Entonces, queremos $F_X(N) \geq 0{,}9$. La ecuación es:

$$1 - \left(\frac{35}{36}\right)^{N} \geq 0{,}9.$$

que implica

$$N \geq \frac{\ln(0{,}1)}{\ln(35/36)} = 81{,}7.$$

Así, son necesarias como mínimo 82 tiradas.

3. Variable aleatoria continua

En el apartado 2 de este módulo, hemos visto qué es una variable aleatoria discreta y hemos estudiado cuatro de las distribuciones más utilizadas: la distribución de Bernoulli, la binomial, la geométrica y la de Poisson. Hemos visto también los parámetros valor medio y varianza y, finalmente, hemos visto qué es la función de distribución de una variable aleatoria discreta. La estructura de este apartado es muy similar a la del anterior. Aquí veremos los conceptos anteriores aplicados al caso de las variables aleatorias continuas.

En el ejemplo 1.2 del apartado 1, cuando hemos definido qué es una variable aleatoria, y hemos visto que la variable aleatoria continua X puede tomar un valor cualquiera del intervalo $[0, 2]$. En este caso, si suponemos que ninguno de los valores dentro de $[0, 2]$ tiene preferencia, podríamos encontrar los resultados siguientes de manera intuitiva:

1) ¿Cuál es la probabilidad de que la señal emitida se encuentre entre 0 y 1 mV, es decir, $P(0 \leq X \leq 1)$? Todo nos hace pensar que es $\frac{1}{2}$, puesto que estamos jugando con la mitad de posibilidades.

2) ¿Cuál es la probabilidad de que la señal emitida se encuentre entre 3 y 4 mV, es decir, $P(3 \leq X \leq 4)$? Dado que sabemos que esto no sucederá nunca, porque el generador solo nos da una señal en el intervalo [0,2], decimos que es 0.

3) ¿Cuál es la probabilidad de que la señal emitida sea exactamente de 1 mV? Este caso nos caracteriza las distribuciones de variables aleatorias continuas. Decimos que $P(X = 1) = 0$. En una distribución de variable aleatoria continua, la probabilidad en cualquier punto x es cero.

3.1. Función de distribución y función de densidad

La función de distribución se define del mismo modo que para una variable aleatoria discreta, tal y como lo habíamos definido en el apartado 2.3.

Definición 3.1. La **función de distribución de una variable aleatoria** X se define como

$$F_X(x) = P(X \leq x), \quad \forall x \in \mathbb{R}. \tag{17}$$

Observad las similitudes en la definición de **función de distribución** para variables aleatorias discretas y continuas.

La función de distribución $F_X(x)$ verifica las propiedades siguientes:

1) $0 \leq F_X(x) \leq 1$.

2) $\lim\limits_{x \to \infty} F_X(x) = 1$.

3) $\lim\limits_{x \to -\infty} F_X(x) = 0$.

4) $F_X(x)$ es creciente, es decir, si $a < b$ entonces $F_X(a) \leq F_X(b)$.

5) $P(a < X \leq b) = F_X(b) - F_X(a)$.

Observad que son exactamente las mismas propiedades que habíamos visto para el caso de las variables aleatorias discretas, pero aplicadas, en este caso, a las variables continuas.

Definición 3.2. Una **variable aleatoria** X es **continua** si $F_X(x)$ es continua en todo $\mathbb{R}$ y derivable en $\mathbb{R}$ (salvo, quizá, un número finito de puntos).

Esto implica que la probabilidad de un punto aislado vale cero. Aplicando (14), tenemos $P(X = a) = 0$.

Definición 3.3. Si X es una variable aleatoria continua con función de distribución $F_X(x)$, la **función de densidad** se define como:

$$f_X(x) = \frac{dF_X(x)}{dx}, \quad \forall x \in \mathbb{R}. \tag{18}$$

Funciones continuas y derivables

Decimos que una función $f(x)$ es **continua** si a medida que nos vamos desplazando por el eje de la variable independiente, x, no se producen saltos o cambios bruscos. Intuitivamente, son funciones que podríamos dibujar sobre un papel sin levantar el lápiz. Decimos que una función es **derivable** o **diferenciable** en un punto si existe su derivada en aquel punto. Recordad que todas las funciones derivables son continuas.

Es decir, la función de densidad, $f_X(x)$, es la derivada en función de x (variable independiente que nos dice qué valores puede tomar nuestra variable aleatoria X) de la función de distribución, $F_X(x)$. Recordad que la función de distribución nos daba la probabilidad acumulada a medida que íbamos considerando los valores posibles de la variable aleatoria. La función de densidad, $f_X(x)$, nos da una idea de cómo varía la función de distribución de una variable aleatoria. E inversamente, la función de distribución es la integral sobre x de la función de distribución.

Veamos un ejemplo. Como podéis ver en la figura 4, el área por debajo de la curva de $f(x)$ corresponde a un punto de $F(x)$.

Figura 4. Función de densidad $f(x)$ y función de distribución $F(x)$

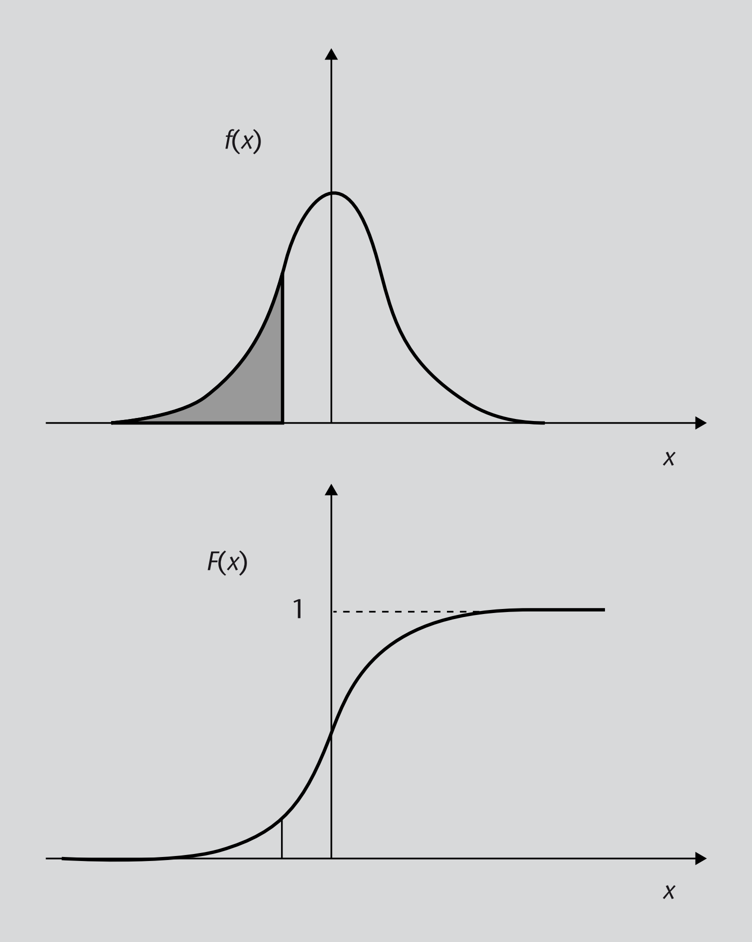


Figura 4

En la figura, podéis observar la relación entre la función de distribución, $F_X(x)$, y la función de densidad, $f_X(x)$.

A continuación, veamos qué relaciones hay entre la **función de distribución**, $F_X(x)$, y la **función de densidad**, $f_X(x)$, de una variable aleatoria continua.

Propiedades de la función de densidad

Si X es una variable aleatoria continua con función de distribución $F_X(x)$ y función de densidad $f_X(x)$, entonces:

1)

$$f_X(x) \geq 0. \tag{19}$$

Esto está claro si observamos la ecuación (18) de la definición 3.3 y pensamos que $F_X(x)$ es creciente, puesto que una función creciente siempre tiene pendiente positiva.

Observación

Si X es una variable aleatoria continua, entonces $P(a \leq X \leq b) = P(a < X < b)$.
En efecto, dado que la probabilidad de un punto es cero, la probabilidad de un intervalo no varía, incluyamos o no sus puntos extremos.

2)

$$P(a \leq X \leq b) = F_X(b) - F_X(a) = \int_a^b f_X(x)\,dx. \qquad (20)$$

Es decir, la probabilidad entre dos puntos a y b la obtenemos integrando la función de densidad entre estos dos puntos, es decir, el área por debajo de la curva de la función de densidad (el resultado se deduce del hecho de que F_X es una primitiva de f_X y es continua).

3)

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt \qquad (21)$$

Es decir, la probabilidad acumulada también la podemos pensar como un área por debajo de la función de densidad $f_X(x)$. Esto se deduce de (20) con $a = -\infty$ y $b = x$, ya que $X \leq x$ corresponde al intervalo $(-\infty, x]$.

4)

$$\int_{-\infty}^{\infty} f_X(x)\,dx = 1. \qquad (22)$$

Es decir, el área total por debajo de la curva $f_X(x)$ es 1. Se obtiene de (21) haciendo $x \to \infty$.

Llegados a este punto, nos podríamos preguntar: ¿por qué es necesario definir la función de densidad de una variable aleatoria continua si ya tenemos el concepto de función de distribución, como habíamos visto con las variables aleatorias discretas? La respuesta es que para el caso de las variables continuas, no siempre es posible expresar la función de distribución de una manera sencilla y cerrada. Además, muchas de las propiedades de estas variables se ven más claramente cuando utilizamos la función de densidad en vez de la función de distribución.

3.2. Variables aleatorias continuas más importantes

Como acabamos de ver en el subapartado anterior, cuando trabajamos con variables aleatorias continuas, estas se pueden caracterizar con su función de densidad. Veamos a continuación las más importantes.

3.2.1. Variable aleatoria uniforme: U(a, b)

La variable X puede tomar un valor cualquiera del intervalo (a, b) y de manera uniforme. En este caso, decimos que X es una **variable aleatoria uniforme** en (a, b). Esto lo indicamos con la función de densidad:

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{si } \ a < x < b \\ \\ 0 & \text{de otro modo} \end{cases} \tag{23}$$

La función de distribución será, pues:

$$F(x) = \int_{-\infty}^{x} f_X(t)\,dt = \begin{cases} 0 & \text{si } \ x < a \\ \\ \dfrac{1}{b-a}(x-a) & \text{si } \ a \leq x < b \\ \\ 1 & \text{si } \ b \leq x \end{cases} \tag{24}$$

Escribimos $X \sim \mathrm{U}(a, b)$.

En la figura 5, podéis ver un ejemplo de variable uniforme y sus funciones de densidad y de distribución. Como podéis ver, el área indicada por debajo de $f(x)$ corresponde a un punto de $F(x)$.

Figura 5. Función de densidad $f(x)$ y función de distribución $F(x)$ de la variable aleatoria $X \sim \mathrm{U}(a, b)$

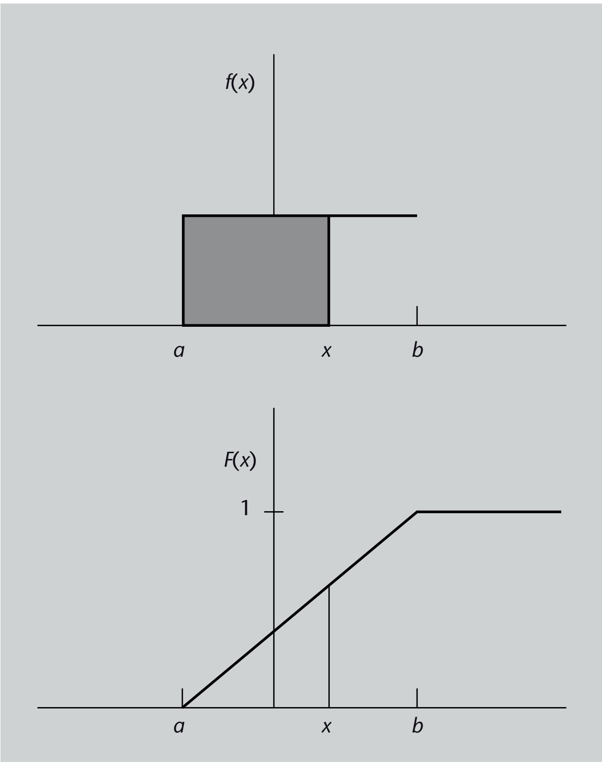


Figura 5

Función de densidad $f(x)$ y función de distribución $F(x)$ de una variable aleatoria uniforme $X \sim \mathrm{U}(a, b)$. Observad que la probabilidad acumulada en el punto x, $F(x)$ corresponde al área bajo la función de densidad $f(x)$.

El ejemplo 1.2 sigue una distribución $X \sim \mathrm{U}(0,2)$. De este modo, la probabilidad de que la señal emitida se encuentre entre 0 y 1 mV nos la da el área por debajo de la curva de la función de densidad, que en este caso corresponde al área de un rectángulo de base 1 y altura $\frac{1}{2}$. Es decir, $P(0 \leq X \leq 1) = \frac{1}{2}$.

Otro ejemplo de sistema que utiliza las distribuciones uniformes son los generadores de números aleatorios. Estos dispositivos generan números dentro de un intervalo determinado de manera uniforme, de tal manera que todos los números tienen la misma probabilidad de ser generados. Veamos un ejemplo a continuación.

Ejemplo 3.1

Elegimos al azar un número, X, en el intervalo $(0,5)$. La función de densidad es:

$$f(x) = \begin{cases} \dfrac{1}{5} & \text{si } x \in (0,5) \\ 0 & \text{de otro modo} \end{cases}$$

Calculamos algunas probabilidades.

1) Probabilidad de que el número sea menor que 3, $P(X < 3) = \int_0^3 \frac{1}{5}\,dx = \frac{3}{5}$.

2) Sabiendo que el número es mayor que 2, probabilidad de que sea menor que 3,

$$P(X < 3 \,|\, X > 2) = \frac{P(2 < X < 3)}{P(X > 2)} = \frac{1/5}{3/5} = \frac{1}{3}.$$

Véase también

Recordad el subapartado 2.3 del módulo «Introducción a la probabilidad», en el que vimos la probabilidad condicionada, es decir, la probabilidad de un suceso sabiendo que se ha producido otro suceso conocido.

3.2.2. Variable aleatoria exponencial: Exp(λ)

La distribución exponencial se suele utilizar para modelizar experiencias en las que interviene un tiempo de espera, como:

- Tiempo de espera en una consulta sin cita previa.
- Tiempo de espera en un servidor para recibir respuesta a una petición enviada.
- La vida de un componente electrónico.

La distribución de Poisson que hemos visto en el apartado de las variables aleatorias discretas está muy relacionada con la distribución exponencial. Si un proceso es de Poisson (suceso aleatorio en el tiempo), la variable tiempo, t, que pasa hasta que tiene lugar el primer suceso, es exponencial. Hay que destacar que el parámetro de la variable de Poisson vale $\alpha = \lambda T$, donde T es el intervalo en el que contamos los acontecimientos que suceden.

La **variable aleatoria exponencial** tiene por función de densidad:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geq 0 \\ \\ 0 & x < 0 \end{cases} \tag{25}$$

Obtenemos la función de distribución integrando. Así,

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{si } x \geq 0 \\ \\ 0 & x < 0 \end{cases} \tag{26}$$

Escribimos $X \sim \text{Exp}(\lambda)$.

En la figura 6 vemos la representación de la función de densidad, para tres valores diferentes de λ. (No se ha representado el eje negativo de abscisas, en el que la función es 0.)

Figura 6. Funciones densidad de $X \sim \text{Exp}(1)$, $X \sim \text{Exp}(2)$ y $X \sim \text{Exp}(3)$

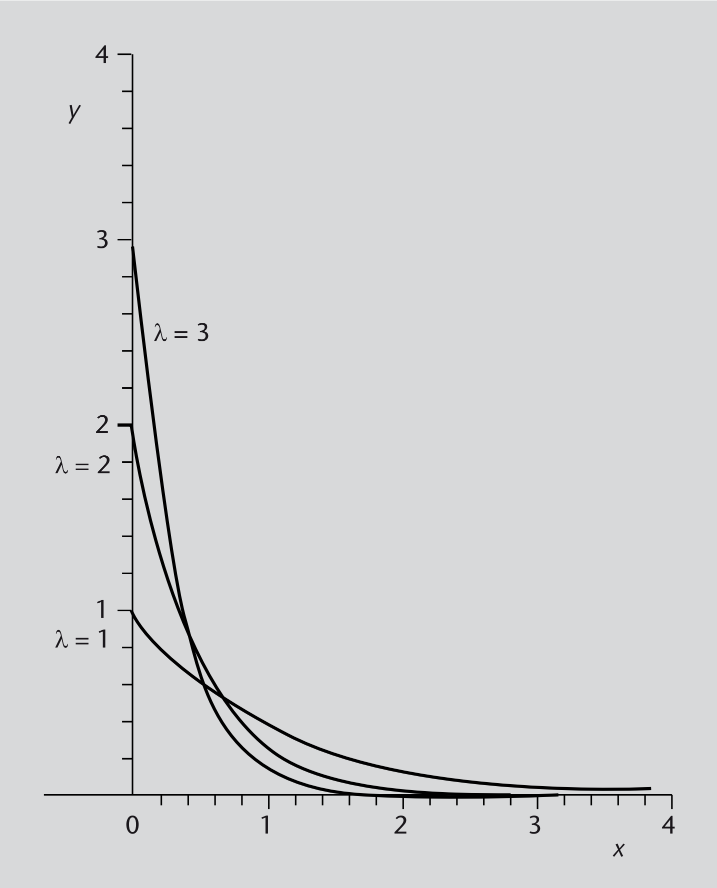


Figura 6

Representación de la función de densidad de una variable exponencial para los valores de λ 1, 2 y 3. Cuanto mayor es λ (más llamadas por unidad de tiempo, por ejemplo), es más probable que tengamos que esperar poco tiempo hasta que llegue una llamada.

Observad que, tal y como λ crece, el pico de la función en torno a $x = 0$ se acentúa. La causa de esto es que, si llegan muchos sucesos por unidad de tiempo, la mayor parte de la probabilidad de tener una llegada se concentra en valores de x pequeños. Por el contrario, si tenemos pocos sucesos por unidad de tiempo (α o λ pequeños), la probabilidad de llegada es más uniforme.

Ejemplo 3.2

Supongamos que el tiempo, en horas, que se necesita para arreglar un cierto tipo de avería telefónica es una variable aleatoria, T, que sigue una ley exponencial de parámetro $\lambda = 0{,}5$. En este caso, tenemos $f(t) = 0{,}5e^{-0{,}5t}$ y $F(t) = 1-e^{-0{,}5t}$ para $x \geq 0$. Calculamos algunas probabilidades:

1) Probabilidad de que el tiempo de reparación pase de las 2 horas. Es decir,

$$P(T > 2) = 1 - P(T < 2) = 1 - F(2) = e^{-1} = 0,368.$$

2) Sabiendo que el tiempo de reparación ya ha sobrepasado las 9 horas, ¿cuál es la probabilidad de que la reparación tarde al menos 10 horas? En este caso, se trata de encontrar una probabilidad condicionada. Escribimos:

$$P(T > 10 \,|\, T > 9) = \frac{P(\{T > 10\} \cap \{T > 9\})}{P(T > 9)} = \frac{P(T > 10)}{P(T > 9)}$$

$$= \frac{1 - F(10)}{1 - F(9)} = e^{-0{,}5 \cdot 1} = 0{,}606.$$

Véase también

Aquí volvemos a utilizar la noción de probabilidad condicionada del subapartado 2.3 del módulo «Introducción a la probabilidad».

3.2.3. Variable aleatoria normal o de Gauss: $\mathrm{N}(m, \sigma)$

Es una de las distribuciones de probabilidad más utilizadas. Muchos fenómenos físicos que afectan a circuitos y aparatos de telecomunicaciones se modelizan utilizando la distribución normal o de Gauss. También se utiliza con mucha frecuencia para el control de calidad estadístico de componentes electrónicos. Depende de dos parámetros, m y σ, que veremos en el subapartado siguiente.

Parámetros de la distribución normal

La distribución normal o de Gauss se caracteriza por dos parámetros: el valor medio m (parámetro de posición), y la desviación típica σ, parámetro que nos mide la dispersión de la variable aleatoria respecto a m.

Una particularidad que presenta esta distribución es que se trata de la forma límite de algunas distribuciones discretas cuando se aumenta indefinidamente el número de repeticiones de un experimento. Muchas variables aleatorias como pesos, alturas, tallas, consumos de gas, etc. siguen una distribución normal porque cada una es la suma de un gran número de variables aleatorias independientes. De este modo, la altura de una persona es la suma de muchos factores, hereditarios, alimentación, tipo de vida, etc.

Los errores, denominados *aleatorios*, que se presentan en observaciones astronómicas, pesadas de una balanza, el ruido generado en los aparatos de telecomunicación, etc. y, en general, en la mayoría de las medidas con algún aparato, son la suma de un gran número de errores elementales independientes como corrientes de aire, vibraciones, error de apreciación, etc. Por eso, los errores aleatorios siguen una distribución normal.

La **variable aleatoria normal** tiene por función de densidad:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}, \quad \forall x \in \mathbb{R}. \tag{27}$$

$m \in \mathbb{R}$ es la esperanza de X y $\sigma > 0$ es la desviación típica de X. Escribimos $X \sim \mathrm{N}(m, \sigma)$.

Cálculo de probabilidades

Para calcular probabilidades, utilizaremos tablas estadísticas o bien algún software matemático de tipo Scilab, Excel, Wiris, Minitab, SPSS, R, etc.

En este caso, la función de distribución no se puede encontrar integrando analíticamente de manera sencilla, como lo hemos hecho antes. Por esta razón, nos será más útil trabajar con la función de densidad.

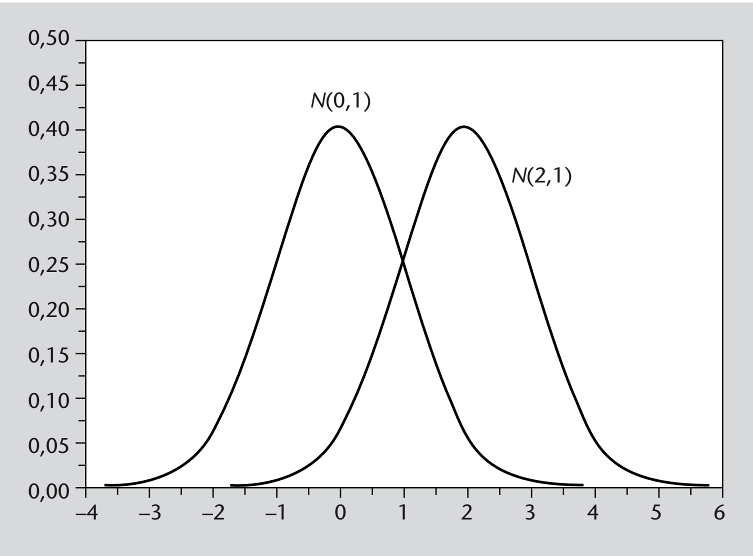
Veamos algunos gráficos de la función de densidad al variar los parámetros m y σ.

En la figura 7, hemos representado $\mathrm{N}(0,1)$ y $\mathrm{N}(2,1)$. El primer parámetro de la función $\mathrm{N}(x,y)$ (0 y 2 en este ejemplo) hace referencia al valor medio. El segundo parámetro (1 en los dos ejemplos) es la desviación estándar. Observad que, dado que las dos variables aleatorias tienen la misma desviación típica, σ (sigma), la forma de la función no ha variado. $\mathrm{N}(2,1)$ está trasladada dos unidades a la derecha respecto de $\mathrm{N}(0,1)$, puesto que el valor medio es diferente en cada caso.

Figura 7

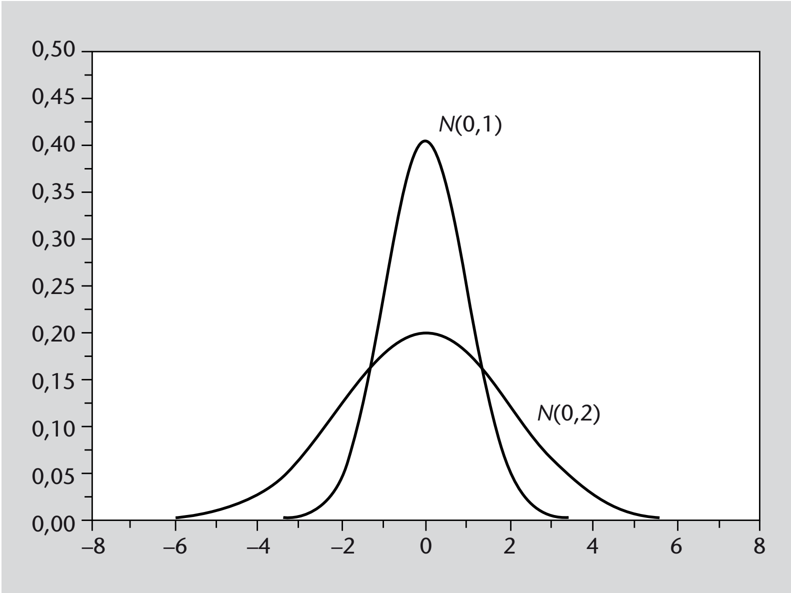
La curva de la izquierda representa la distribución $\mathrm{N}(0,1)$, centrada en 0 y con una desviación típica de valor 1. La curva de la derecha representa la distribución $\mathrm{N}(2,1)$, con valor medio 2 y desviación estándar 1.

Figura 7. Funciones de densidad de $X \sim \mathrm{N}(0,1)$ y $X \sim \mathrm{N}(2,1)$



En la figura 8, fijamos el valor a $m = 0$ y modificamos σ. Observamos que para un valor menor que σ hay menos dispersión. Para $\sigma = 2$ tenemos más dispersión y, por lo tanto, la función de densidad es menos puntiaguda.

Figura 8. Funciones de densidad de $X \sim \mathrm{N}(0,1)$ y $X \sim \mathrm{N}(0,2)$



Observad que aunque cambie la forma de la función, el área total por debajo de la curva es 1, la probabilidad total.

Figura 8

La función $\mathrm{N}(0,1)$ es una distribución gaussiana con media 0 y desviación estándar 1. En la función $\mathrm{N}(0,2)$ hemos aumentado la desviación estándar, y dado que tenemos más dispersión, esta función es más plana, menos puntiaguda. Notad, sin embargo, que el área total bajo las dos curvas ha de ser la misma e igual a 1.

Puesto que no podemos integrar analíticamente $f(x)$, para encontrar probabilidades hay que utilizar tablas o bien algún software matemático o estadístico.

Figura 9. Función de densidad de $X \sim \mathrm{N}(0,1)$

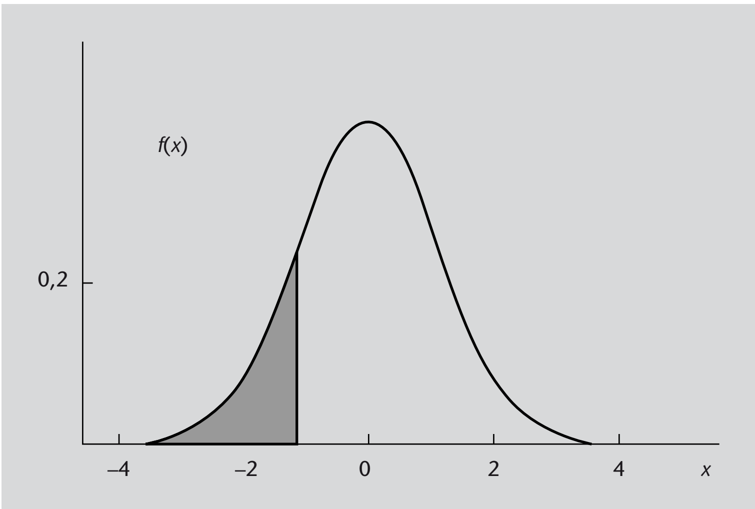


Figura 9

En este ejemplo, se ha calculado la probabilidad de que $x < -1$. Para hacerlo, hemos calculado el área bajo la distribución de Gauss desde $-\infty$ hasta $x = 1$. Utilizando software matemático, se ha encontrado que $P(x < -1) = 0,158655$.

3.3. Parámetros: valor medio (esperanza) y varianza

En el subapartado anterior, hemos visto tres de las distribuciones continuas más frecuentes: la distribución uniforme, la exponencial y la de Gauss. De manera parecida a como lo hemos hecho para el caso de las variables aleatorias discretas, en este subapartado veremos dos parámetros que definen estas distribuciones: el valor medio y la varianza. La diferencia fundamental entre las variables discretas y las continuas es que, en este caso, transformaremos los sumatorios que habíamos visto en el subapartado 2.2 en integrales. El significado de los parámetros es el mismo que en el caso discreto.

Definición 3.4. Sea X una variable aleatoria continua.

El **valor medio, esperanza** o **momento de orden 1** de X:

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x f(x)\, dx. \tag{28}$$

El **momento de orden 2**:

$$\mathrm{E}(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\, dx. \tag{29}$$

El **momento de orden *k***, $(k = 0, 1, 2, \ldots)$:

$$\mathrm{E}(X^k) = \int_{-\infty}^{\infty} x^k f(x)\, dx. \tag{30}$$

La **varianza**:

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} (x - \mathrm{E}(X))^2 f(x)\, dx = \mathrm{E}(X^2) - \mathrm{E}(X)^2. \tag{31}$$

La **desviación típica**:

$$\sigma = \sqrt{\mathrm{Var}(X)}. \tag{32}$$

Parámetros de las principales variables aleatorias continuas

Para cada una de las distribuciones vistas, se obtienen los valores de la esperanza y la varianza de la tabla siguiente.

Distribuciones de variables aleatorias continuas

$X \sim$	**Función de densidad**	$\mathrm{E}(X)$	$\mathrm{Var}(X)$
$\mathrm{U}(a,b)$	$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{si } x \in (a,b) \\ 0 & \text{de otro modo} \end{cases}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
$\mathrm{N}(m,\sigma)$	$f(x) = \dfrac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} \ \forall x \in \mathbb{R}$	m	σ^2
$\mathrm{Exp}(\lambda)$	$f(x) = \begin{cases} \lambda\, e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{de otro modo} \end{cases}$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$

Comprobamos algún resultado para el caso de $X \sim \mathrm{Exp}(\lambda)$:

$$\begin{aligned}
\mathrm{E}(X) &= \int_{-\infty}^{\infty} x f(x)\, dx = \int_{0}^{\infty} x\lambda e^{-\lambda x}\, dx = \left\{ \begin{array}{c} u = x \quad \text{i} \quad du = dx \\ dv = \lambda e^{-\lambda x} \quad \text{i} \quad v = -e^{-\lambda x} \end{array} \right\} = \\
&= \left[-xe^{-\lambda x}\right]_0^{\infty} + \int_0^{\infty} e^{-\lambda x}\, dx = \lim_{x\to\infty}(-xe^{-\lambda x}) - \lim_{x\to 0}(-xe^{-\lambda x}) + \frac{1}{\lambda} = \\
&= 0 + 0 + \frac{1}{\lambda}.
\end{aligned}$$

3.4. Variables aleatorias mixtas

Como hemos visto, las variables aleatorias discretas tienen una función de distribución constante a trozos, y las variables aleatorias continuas tienen una función de distribución continua en todo $\mathbb{R}$ y derivable a trozos. Esto no agota todas las posibilidades, puesto que podemos tener una función de distribución que sea derivable a trozos y que tenga algunas discontinuidades de salto. En este caso, hablamos de variables aleatorias mixtas. Para estas, tenemos algunos puntos con probabilidad no nula ($P(X = a) > 0$) y, además, pueden tomar todos los valores de conjuntos como intervalos reales.

Ejemplo 3.3

Consideramos una variable aleatoria con función de distribución

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ \dfrac{x+1}{x+2} & \text{si } x \geq 0 \end{cases}$$

La anterior función es creciente, tiende a 0 para $x \to -\infty$ y tiende a 1 por $x \to \infty$, así que es una función de distribución correcta.

La variable X no es continua, puesto que en $x = 0$ $F(x)$ salta de 0 a 1/2. Entonces $P(X = 0) = 1/2$, probabilidad que sería cero para una variable continua.

Tampoco es discreta, puesto que puede tomar cualquier valor positivo dado que $F(x)$ es estrictamente creciente para $x > 0$.

Se trata, pues, de una variable mixta. No tiene una densidad definida, debido a la discontinuidad en cero. Aun así, se puede definir una densidad utilizando funciones generalizadas (delta de Dirac), tal y como se verá en el apartado 3.6.

3.5. Funciones de densidad condicionadas

Del mismo modo que la probabilidad de un suceso, $P(A)$, cambia cuando sabemos que se ha producido otro suceso B, para pasar a ser $P(A \mid B)$, las características de las variables aleatorias también cambian en esta situación.

Ejemplo 3.4

Consideramos una variable X que da el tiempo que pasa hasta que llega una señal de comunicación. Si ha pasado un cierto tiempo, digamos que 2 segundos, y la señal no ha llegado, la probabilidad de que X tome ciertos valores queda condicionada por este hecho. En este caso, $B = \{X > 2\}$ y, dado B, la probabilidad de que, por ejemplo, $X > 5$ se tiene que calcular como una probabilidad condicionada: $P(X > 5 \,|\, B)$.

En esta situación, puede ser más práctico tener una densidad para la variable X que ya incorpore la condición. Esto nos lleva a definir la densidad condicionada.

Si B es un suceso que afecta a la variable aleatoria X, la densidad de X, $f(x)$, se modifica para pasar a ser $f(x \,|\, B)$, densidad de X condicionada al suceso B.

A partir de aquí, podemos calcular la esperanza de X condicionada al suceso B:

$$\mathrm{E}(X \,|\, B) = \int_{-\infty}^{\infty} x f(x \,|\, B) dx. \qquad (33)$$

La situación que consideramos se produce cuando el suceso B es de la forma $X \in A$, donde A es un subconjunto de $\mathbb{R}$. El resultado es el siguiente:

$$f(x \,|\, X \in A) = \begin{cases} \dfrac{f(x)}{P(A)} & \text{si } x \in A \\ 0 & \text{si } x \notin A \end{cases} \qquad (34)$$

En el anterior ejemplo, este sería el caso. Con $B = \{X > 2\}$ es $A = (2, \infty)$.

Ejemplo 3.5

El tiempo que tarda un servidor en procesar una petición es una variable aleatoria X de tipo exponencial con valor medio de 2 segundos. Suponemos que pasado 1 segundo, el proceso todavía no ha acabado. ¿Cuál es la densidad y el valor medio de X condicionados a este hecho?

$X \sim \mathrm{Exp}(1/2)$, puesto que el valor medio es $1/\lambda = 2$. La densidad de X es $f(x) = \frac{1}{2}e^{-x/2}$, $x \geq 0$ y la función de distribución de X es $F(x) = 1 - e^{-x/2}$, $x \geq 0$.

La condición que tenemos es $X > 1$, así que el conjunto A es $(1, \infty)$. Notamos que su probabilidad vale $P(A) = P(X > 1) = 1 - P(X \leq 1) = 1 - F(1) = e^{-1/2}$. Entonces:

$$f(x \,|\, X > 1) = \begin{cases} \dfrac{\frac{1}{2}e^{-x/2}}{e^{-1/2}} = \frac{1}{2}e^{-(x-1)/2} & \text{si } x > 1 \\ 0 & \text{si } x < 1 \end{cases}$$

La esperanza condicionada vale:

$$\mathrm{E}(X \,|\, X > 1) = \int_{1}^{\infty} x \frac{1}{2} e^{-\frac{x-1}{2}} dx = [-(x+2)e^{-\frac{x-1}{2}}]_1^{\infty} = 3.$$

Notamos que esta esperanza es mayor que la esperanza sin condicionar, $E(X) = 2$, puesto que la condición desplaza el peso de las probabilidades hacia valores mayores.

3.6. Delta de Dirac. Densidad en el caso discreto

Las variables aleatorias discretas y las mixtas tienen funciones de distribución con discontinuidades de salto. En estos casos, no existe la derivada y, por lo tanto, no hay una función de densidad en el sentido de la fórmula (18) o la (20). Esto es posible hacerlo utilizando las denominadas **funciones generalizadas**, más concretamente, la delta de Dirac. Su definición es:

Definición 3.5. La función generalizada **delta de Dirac**, $\delta(x)$, se define a través de la siguiente propiedad formal:

Para toda función continua $g(x)$ y números $\alpha < \beta$ se verifica:

$$\int_\alpha^\beta \delta(x)g(x)\,dx = \begin{cases} g(0) & \text{si } \alpha < 0 < \beta \\ 0 & \text{de otro modo} \end{cases} \tag{35}$$

Más generalmente, utilizamos la delta desplazada al punto a:

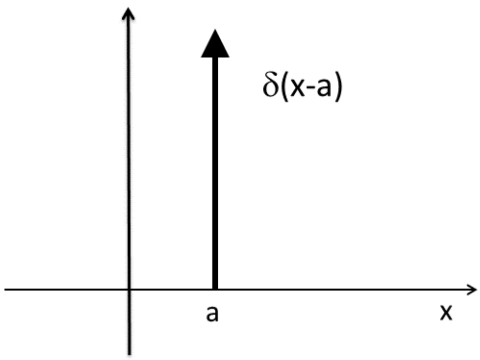
$$\int_\alpha^\beta \delta(x-a)g(x)\,dx = \begin{cases} g(a) & \text{si } \alpha < a < \beta \\ 0 & \text{de otro modo} \end{cases} \tag{36}$$

Notamos que $\int_\alpha^\beta \delta(x-a)\,dx = 1$ siempre que el intervalo de integración contenga el punto a, por pequeño que sea este intervalo. Esto no se podría conseguir con ninguna función ordinaria. Una visualización de la $\delta(x-a)$ sería una función que vale 0 para toda $x \neq a$ y vale ∞ en el punto a, de forma que la integral valga 1 (figura 10).

Dado que la probabilidad de una variable aleatoria discreta se concentra en puntos aislados, es razonable representarla con una densidad con deltas de Dirac. Si X es una variable aleatoria discreta que toma valores a_i con probabilidades $P_X(a_i) = P(X = a_i)$, le podemos asociar la siguiente densidad:

$$f_X(x) = \sum_i P_X(a_i)\delta(x - a_i). \tag{37}$$

Figura 10. Delta de Dirac centrada en el punto a



Una manera más precisa de llegar al anterior resultado es a través de la función escalón:

Definición 3.6. La **función escalón** o **función de Heaviside**, $u(x)$ es:

$$u(x) = \begin{cases} 1 & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases} \tag{38}$$

La función escalón se relaciona con la delta de la siguiente manera:

$$\frac{d}{dx}u(x-a) = \delta(x-a). \tag{39}$$

En efecto, de (36) deducimos que: $\int_{-\infty}^{x} \delta(y-a)\,dy = 0$ para $x < a$, mientras que
$\int_{-\infty}^{x} \delta(y-a)\,dy = 1$ para $x > a$, así que

$$\int_{-\infty}^{x} \delta(y-a)\,dy = u(x-a).$$

Derivando los dos lados de la anterior ecuación, sale (39).

Ahora, podemos expresar la función de distribución de una variable discreta o mixta con funciones escalón y derivarla utilizando (39).

Por una variable discreta, podemos expresar la función de distribución:

$$F_X(x) = \sum_i P_X(a_i)u(x - a_i),$$

puesto que F_X es una función escalonada y la altura de cada escalón es el factor $P_X(a_i)$. Derivando la anterior ecuación, se obtiene (37).

En el caso de variables mixtas, conviene tener en cuenta la siguiente propiedad. Para cualquier función continua $h(x)$:

$$\delta(x - a)h(x) = \delta(x - a)h(a). \tag{40}$$

La anterior igualdad se demuestra multiplicando los dos lados por cualquier función continua $g(x)$ y viendo que al integrar sobre cualquier intervalo, se obtiene el mismo resultado en los dos lados.

Ejemplo 3.6

Obtendremos la densidad por la variable mixta del ejemplo 3.3.

Podemos expresar la distribución de X como: $F(x) = u(x)\frac{x+1}{x+2}$

Derivando, y utilizando (40):

$$f(x) = u'(x)\frac{x+1}{x+2} + u(x)\left(\frac{x+1}{x+2}\right)' = \delta(x)\frac{x+1}{x+2} + u(x)\frac{1}{(x+2)^2}$$

$$= \frac{1}{2}\delta(x) + u(x)\frac{1}{(x+2)^2}.$$

4. Teorema central del límite. Aplicación

En este módulo, hemos visto las variables aleatorias discretas y continuas. Para cada variable aleatoria, hemos calculado el valor medio y la varianza. Estos parámetros nos dan una idea global del comportamiento de la variable aleatoria. También, para cada una de las variables estudiadas, hemos visto la función de distribución, que nos permite calcular la probabilidad acumulada dado un cierto valor x y la función de densidad, que nos dice cómo varía (es la derivada en función de x) la función de distribución.

En este último apartado de este módulo, veremos cómo podemos relacionar las variables aleatorias discretas vistas en el apartado 2 con la variable aleatoria continua normal o de Gauss que hemos visto en el apartado 3. Tal y como hemos dicho antes, la distribución normal es la forma límite de algunas distribuciones discretas, cuando se aumenta indefinidamente el número de repeticiones de un experimento.

Recordad el ejemplo 2.8 que hemos visto en el subapartado 2.2, en el que teníamos la distribución de notas de dos alumnos como ejemplo de distribución discreta. Observad que allí hemos calculado el valor medio y la desviación típica, justamente los parámetros que definen una distribución normal. ¿Qué pasaría si en nuestra distribución discreta aumentáramos el número de muestras y en vez de 3 resultados tomáramos muchas más notas? ¿O si dibujásemos la distribución de notas, no de 2, sino de muchos más alumnos? Esto es lo que nos permite el teorema central del límite: aproximar una distribución discreta a una normal cuando aumentamos el número de muestras y se dan una serie de condiciones que veremos a continuación. Enunciamos el teorema central del límite (TCL), que refleja este hecho.

Observación

El teorema central del límite nos permite aproximar una variable aleatoria discreta a una distribución de Gauss cuando repetimos un número lo bastante grande de veces un experimento.

Teorema central del límite

Sea $\{X_n\}$ con $n \geq 1$ una sucesión de variables aleatorias independientes, que siguen la misma ley de probabilidad, con una esperanza m y varianza σ^2. Consideramos la nueva variable aleatoria definida por:

$$Y_n = \frac{\dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i - m}{\dfrac{\sigma}{\sqrt{n}}}. \qquad (41)$$

Esta variable, tal y como la hemos definido, tiene un valor medio igual a cero y una varianza igual a 1. Se tiene que la variable Y_n converge hacia la distribución $\mathrm{N}(0,1)$ cuando n tiende a infinito. La distribución $\mathrm{N}(0,1)$ también se conoce como **distribución normal estándar**.

De manera alternativa, podemos considerar la variable:

$$S_n = \sum_{i=1}^{n} X_i.$$

Entonces, el TCL nos dice que para n grande $S_n \sim \mathrm{N}(nm, \sqrt{n}\sigma)$.

4.1. Aproximación de ley binomial a la normal

Si estamos trabajando con una variable discreta $X \sim \mathrm{Bin}(n,p)$, recordemos que X cuenta el número de éxitos en n repeticiones de un experimento de Bernoulli. Entonces podemos expresarla como $X = \sum_{i=1}^{n} X_i$, donde las X_i son variables de Bernoulli independientes. X_i es un indicador que vale 1 si la i-ésima vez que hacemos el experimento se obtiene éxito, y 0 en caso contrario. La suma de los indicadores da el número total de éxitos.

Siendo $X_i \sim \mathrm{B}(p)$, tenemos $\mathrm{E}(X_i) = p$ y $\mathrm{Var}(X_i) = p(1-p)$. Aplicando el TCL, si n es grande, X se aproxima a una normal $Y \sim \mathrm{N}(np, \sqrt{np(1-p)})$, en la que la variable Y tiene la misma esperanza y varianza que X. Observad que, en este caso, el valor medio de la distribución normal, m, es igual a np, y la desviación, σ, es $\sqrt{np(1-p)}$, en la que n y p son los parámetros de la distribución binomial.

Aunque en el teorema anterior se habla de aproximación cuando n tiende a infinito, en la práctica esta aproximación es válida cuando se cumple $np > 5$ y $n(1-p) > 5$.

Hay que tener en cuenta que pasamos de una distribución discreta que toma valores enteros entre 0 y n a una variable continua que toma valores en todo $\mathbb{R}$. Además, en el caso de la ley binomial, la probabilidad en un punto es diferente de cero, mientras que no es así en el caso de la ley normal, porque es una distribución continua.

Observación

Recordad que $X \sim \mathrm{Bin}(n,p)$ es el número de n experiencias de Bernoulli $\mathrm{B}(p)$ indexadas.

Por estas razones, cuando aproximamos una distribución binomial a una normal, hay que hacer una corrección de continuidad de la manera siguiente:

Si $X \sim \mathrm{Bin}(n,p)$ con $np > 5$ y $n(1-p) > 5$ y queremos calcular $P(a \leq X \leq b)$, consideramos la variable $Y \sim \mathrm{N}(np, \sqrt{np(1-p)})$ y calculamos $P(a - 0{,}5 < Y < b + 0{,}5)$.

Observación

Observad los factores $-0,5$ y $+0,5$, que hemos añadido a los límites a y b a causa de la corrección de continuidad.

Ejemplo 4.1

En un almacén, se ha analizado durante un año el porcentaje de piezas defectuosas y se ha detectado un 8 % de las mismas. Es decir, podemos considerar que la probabilidad de que una pieza sea defectuosa es de $0,08$. Se toma una muestra de piezas 100 y se define la variable aleatoria X como el número de piezas defectuosas dentro de la muestra de 100. La variable aleatoria X sigue una ley binomial $\text{Bin}(100, 0{,}08)$, puesto que repetiremos el experimento de tomar una pieza 100 veces, y la probabilidad de éxito (aquí definimos éxito como la obtención de una pieza defectuosa) es 0,08. Calculamos la probabilidad de que en las 100 piezas haya entre 10 y 20 defectuosas.

Primero lo calculamos sin hacer la aproximación. Puesto que $X \sim \text{Bin}(100,\ 0{,}08)$,

$$P(10 \leq X \leq 20) = \sum_{k=10}^{20} \binom{100}{k} 0{,}08^k 0{,}92^{100-k} = 0{,}2779.$$

Comprobamos a continuación que podemos hacer la aproximación de la distribución binomial a una normal:

- $np = 100 \cdot 0{,}08 = 8 > 5.$
- $n(1-p) = 100 \cdot 0{,}92 = 92 > 5.$

Una vez confirmado que podemos hacer la aproximación, tomamos una distribución normal con los parámetros siguientes: $m = np = 8$ y $\sigma = \sqrt{np(1-p)} = 2{,}713$. Por lo tanto, nuestra distribución normal es $Y \sim \text{N}(8,\ 2{,}713)$. La probabilidad de que en 100 piezas haya entre 10 y 20 defectuosas es:

$$P(10 \leq X \leq 20) = P(9{,}5 < Y < 20{,}5) = 0{,}2901.$$

Recordad que tenemos que corregir los límites del intervalo con el factor $0,5$. El último valor numérico se ha obtenido con la ayuda de un software matemático.

Resumen

En el apartado 1, hemos visto qué es una **variable aleatoria** y la hemos definido como una función que asigna un número a cada elemento del espacio muestral Ω. Por ejemplo, si lanzamos una moneda, podemos definir $X = 0$ si obtenemos cara y $X = 1$ si obtenemos cruz.

También hemos visto que hay dos tipos de variables aleatorias:

- Variables aleatorias discretas: los valores que puede tomar X se encuentran dentro de un conjunto finito o infinito numerable de elementos.

- Variables aleatorias continuas: X puede tomar cualquier valor en conjuntos no numerables.

El apartado 2 lo hemos dedicado a estudiar en detalle algunas de las variables aleatorias discretas más importantes. Hemos definido qué es la **función de probabilidad** y hemos visto las distribuciones siguientes:

- Variable aleatoria discreta **uniforme**.
- Distribución de **Bernoulli**, $\mathrm{B}(p)$, definida por el parámetro p (probabilidad de éxito).
- Distribución **binomial**, $\mathrm{Bin}(n, p)$, que consiste en un experimento de Bernoulli repetido n veces.
- Distribución **geométrica**, $\mathrm{Geom}(p)$, que se da cuando repetimos el experimento de Bernoulli hasta que obtenemos el primer éxito.
- Distribución de **Poisson**, $\mathrm{Poiss}(\alpha)$, caracterizada por el parámetro α (número medio de sucesos dentro de un cierto intervalo).

Hemos definido los momentos de orden n de una variable aleatoria discreta y, en particular, hemos visto la **esperanza** o **valor medio**, la **varianza** y la **desviación típica**. Definimos el valor medio o esperanza como se indica a continuación:

$$\mathrm{E}(X) = \sum_{i=1}^{n} a_i P(X = a_i).$$

El momento de orden 2 es:

$$\mathrm{E}(X^2) = \sum_{i=1}^{n} a_i^2 P(X = a_i).$$

La varianza y la desviación típica son:

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2, \qquad \sigma = \sqrt{\mathrm{Var}(X)}.$$

Finalmente, hemos visto el concepto de **función de distribución** como función de probabilidad acumulada que nos da la probabilidad de que la variable aleatoria X sea igual o menor que un cierto valor x:

$$F_X(x) = P(X \leq x), \qquad \forall x \in \mathbb{R}.$$

En la tabla siguiente, podéis ver los parámetros más importantes para las variables aleatorias discretas que hemos visto.

Distribuciones de variables aleatorias discretas

$X \sim$	k	$P(X=k)$	$\mathrm{E}(X)$	$\mathrm{Var}(X)$
$\mathrm{B}(p)$	$\{0,1\}$	$P(X=1) = p$ $P(X=0) = 1-p$	p	$p(1-p)$
$\mathrm{Bin}(n,p)$	$\{0,1,2,\ldots,\ n\}$	$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$
$\mathrm{Geom}(p)$	$\{1,2,3,\ldots\}$	$P(X=k) = (1-p)^{k-1} p$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
$\mathrm{Poiss}(\alpha)$	$\{0,1,2,\ldots\}$	$P(X=k) = \frac{\alpha^k}{k!} e^{-\alpha}$	α	α

Hemos dedicado el apartado 3 al estudio de las variables aleatorias continuas. Hemos empezado el apartado definiendo la **función de distribución**, que conceptualmente es la misma que para el caso de las variables discretas. Para el caso de las variables continuas, hemos definido una nueva función, la **función de densidad**, que es la derivada de la **función de distribución**. A continuación, hemos visto las variables aleatorias continuas más importantes:

- La distribución **uniforme**, $X \sim \mathrm{U}(a,b)$, caracterizada por el intervalo (a,b), en el que la **función de densidad** es constante y vale $\frac{1}{b-a}$ en este intervalo.

- La distribución **exponencial**, $\mathrm{Exp}\,(\lambda)$, en la que λ es la tasa de sucesos por unidad de tiempo.

- La distribución **normal** o de **Gauss**, $\mathrm{N}(m,\sigma)$, caracterizada por el valor medio y la desviación típica.

Del mismo modo que hemos hecho en el apartado 2 con las variables aleatorias discretas, hemos visto también en este apartado que para las variables aleatorias continuas, podemos definir los **momentos de orden** n, la **varianza** y la **desviación típica**. El valor medio o esperanza se expresa como sigue:

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x f(x) dx.$$

El momento de orden 2 es:

$$\mathrm{E}(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx.$$

La varianza y la desviación estándar son:

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2, \qquad \sigma = \sqrt{\mathrm{Var}(X)}.$$

La función de distribución y de densidad las definimos como sigue:

$$F_X(x) = P(X \leq x), \qquad f_X(x) = \frac{dF_X(x)}{dx}.$$

En la tabla siguiente, podéis ver los parámetros más importantes para las variables aleatorias continuas que hemos visto.

Distribuciones de variables aleatorias continuas

$X \sim$	**Función de densidad**	$\mathrm{E}(X)$	$\mathrm{Var}(X)$
$\mathrm{U}(a,b)$	$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in (a,b) \\ 0 & \text{de otro modo} \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
$\mathrm{N}(m,\sigma)$	$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} \ \forall x \in \mathbb{R}$	m	σ^2
$\mathrm{Exp}(\lambda)$	$f(x) = \begin{cases} \lambda\, e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{de otro modo} \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$

Finalmente, en el apartado 4, hemos visto el **teorema central del límite**. Este teorema nos permite aproximar la suma o la media de una sucesión de variables aleatorias independientes a una **distribución normal** bajo ciertas condiciones. En particular, hemos visto cómo podemos aplicar esta ley para aproximar una distribución binomial a una normal.

Actividades

1. Una empresa de fabricación de microchips observa que el número de componentes electrónicos que fallan antes de cumplir 100 horas de funcionamiento sigue una distribución de Poisson. De media, el número de microchips que fallan en este intervalo de tiempo es 8. Se pide:

a) Comprobad que la función de probabilidad correspondiente a la Poisson satisface la condición de que la suma de todas las probabilidades de valores posibles de la variable tiene un valor de 1, es decir, que $1 = \sum_{k=0}^{\infty} P(X=k)$.
Pista: Recordad la serie de Taylor de la función exponencial.

b) ¿Cuál es la probabilidad de que falle exactamente un microchip al cabo de 50 horas de funcionamiento?

c) ¿Cuál es la probabilidad de que no fallen más de dos microchips en 100 horas?

d) ¿Cuál es la probabilidad de que fallen al menos 10 microchips en 125 horas?

2. Una fuente binaria genera dígitos 1 y 0 de manera aleatoria con probabilidades $0,6$ y $0,4$, respectivamente. Se pide:

a) ¿Con cuál de las variables aleatorias vistas podríamos modelizar el comportamiento de esta fuente?

b) ¿Cuál es la probabilidad de que en una secuencia de 5 dígitos salgan dos 1 y tres 0?

c) ¿Cuál es la probabilidad de que en la secuencia de 5 dígitos se obtengan al menos tres 1?

3. Suponed que el tiempo (en segundos) que tarda un servidor de bases de datos en dar respuesta a una consulta SQL es una variable aleatoria exponencial con parámetro $\lambda = 1/10$. Si el servidor recibe otra consulta SQL justo antes de la vuestra, se pide:

a) ¿Qué valores puede tomar el tiempo de espera X para poder lanzar una consulta? ¿Se trata de una variable aleatoria discreta o continua?

b) ¿Cuál es la probabilidad de que el anterior tiempo de espera sea menos de 5 segundos?

c) ¿Cuál es la probabilidad de que el tiempo de espera esté entre 5 y 10 segundos?

4. En una red de telecomunicaciones, se ha calculado que la probabilidad de que un direccionador falle en una jornada de actividad extrema es de 0,04. Si se considera un total de 2.500 jornadas de actividad extrema, entonces:

a) Sea X el número de veces que falla el direccionador. ¿De qué tipo de variable aleatoria se trata?

b) ¿Cuál es la probabilidad de que el direccionador falle más de 120 veces en este periodo?

c) ¿Y de que falle entre 100 y 120 veces (los dos incluidos) en este periodo?

Pista: Gracias al TCL, es posible aproximar una binomial mediante una normal.

5. El número de consultas que un servidor de bases de datos procesa en un intervalo de 10 segundos es una variable aleatoria de Poisson, X, con tasa $\lambda = 0,5$ consultas por segundo. Se pide:

a) ¿Cuál es la probabilidad de que ninguna consulta sea procesada en un intervalo de 10 segundos?

b) ¿Cuál es la probabilidad de que al menos 2 consultas sean procesadas en un intervalo de 10 segundos?

6. Un emisor A transmite un mensaje a un receptor B. Sea p la probabilidad de que B reciba correctamente el mensaje. Para asegurarse de que el mensaje será recibido al menos una vez, A volverá a enviar el mensaje hasta un máximo de n intentos. Suponiendo que las n transmisiones son independientes, se pide:

a) Identificad cuál es la distribución estadística asociada a la variable $X=$ número de mensajes recibidos correctamente por B en los n intentos.

b) Si $p = 0{,}7$ y $n = 3$, calculad la probabilidad de que B acabe recibiendo el mensaje.

c) Si $p = 0{,}8$, ¿cuál es el valor mínimo de n que hace que la probabilidad de que el mensaje se reciba sea, como mínimo, de 0,95?

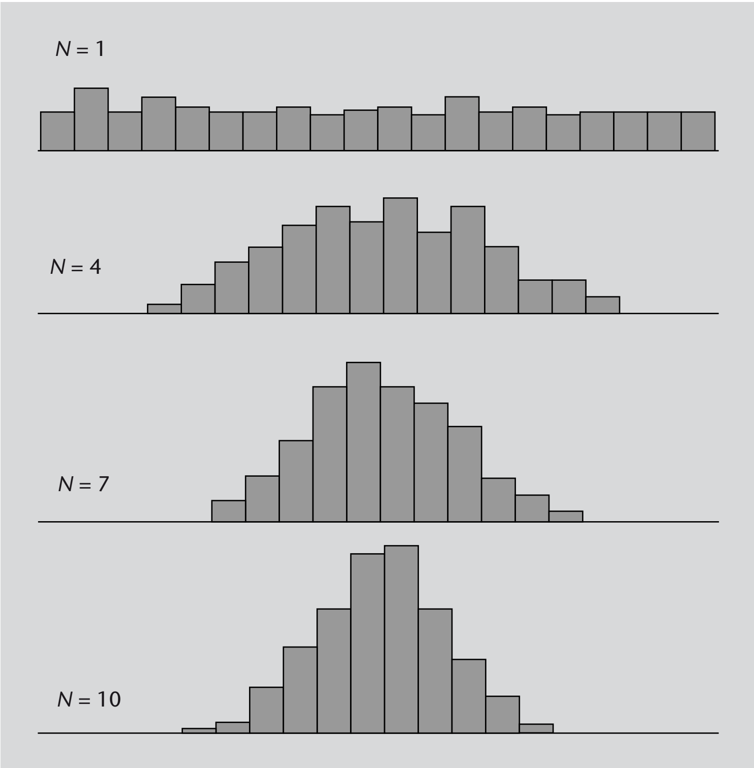
7. Suponed que la temperatura T a la que tiene que trabajar una sonda de medida durante una misión espacial es una variable aleatoria gaussiana (distribución normal) con media 85 grados Fahrenheit y desviación estándar de 10 grados Fahrenheit.

a) En un momento determinado, ¿cuál es la probabilidad de que la temperatura esté entre 75 y 95 °F?

b) ¿Y de que esté entre 65 y 105 °F?

c) Buscad en internet información sobre la regla 68/95/99 y justificad que los resultados anteriores son coherentes con esto.

d) Haced uso del teorema central del límite para explicar la gráfica siguiente.



8. Una centralita telefónica recibe 300 llamadas por hora. La centralita está dimensionada de tal manera que no se pueden establecer más de 12 conexiones por minuto. Con estos datos, nos piden lo siguiente:

a) ¿Cuál es la probabilidad de que la centralita quede saturada en un minuto determinado?

b) ¿Cuál es la probabilidad de que se reciba una única llamada en un minuto determinado?

9. De una estación parte un tren cada 20 minutos. Si llegamos a la estación en un momento cualquiera, nos piden determinar lo siguiente:

a) La función de distribución de la variable aleatoria «tiempo de espera».

b) La probabilidad de que tengamos que esperar en la estación menos de 7 minutos.

c) La esperanza, la varianza y la desviación de la variable aleatoria «tiempo de espera».

d) La probabilidad de que tengamos que esperar exactamente 12 minutos.

10. Un avión de alto rendimiento tiene una computadora central y dos más idénticas, preparadas por si falla alguna de las otras. Durante una hora de operación, la probabilidad de que falle la computadora principal o cualquiera de las otras es de 0,1. Suponiendo que cada hora representa un experimento independiente del resto:

a) ¿Cuál es el tiempo medio que pasa para que fallen las tres computadoras?

b) ¿Cuál es la probabilidad de que las 3 computadoras fallen en un vuelo de 5 horas?

Solucionario

1.

a) Se trata de comprobar que $1 = \sum_{k=0}^{\infty} P(X=k)$.

Teniendo en cuenta la serie de Taylor $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$:

$$\sum_{k=0}^{\infty} P(X=k) = e^{-\alpha} \sum_{k=0}^{\infty} \frac{\alpha^k}{k!} = e^{-\alpha} \cdot e^{\alpha} = 1.$$

b) Sabemos que el número medio de microchips que fallan antes de cumplir 100 horas de funcionamiento es 8. Tenemos que la tasa es $\lambda = 8/100 = 0,08$ microchips por hora. El número medio de microchips que fallan en un intervalo de 50 horas será $\alpha = \lambda \cdot 50 = 4$. La probabilidad de que fallen k microchips en este intervalo de tiempo es: $P(X=k) = \frac{\alpha^k}{k!} e^{-\alpha}$ Sustituyendo por los valores del enunciado:

$$P(X=1) = \frac{4^1}{1!} e^{-4} = 0{,}0732.$$

c) En este caso, el intervalo para considerar son 100 horas, el mismo que nos dan por enunciado. En este intervalo, $\alpha = 8$. La probabilidad de que fallen como mucho dos microchips es:

$$\sum_{k=0}^{2} \frac{\alpha^k}{k!} e^{-\alpha} = \sum_{k=0}^{2} \frac{8^k}{k!} e^{-8} = \left(\frac{8^0}{0!} + \frac{8^1}{1!} + \frac{8^2}{2!} \right) e^{-8} = 0{,}0137.$$

d) Considerando ahora el intervalo de 125 horas, podemos decir que de media el número de microchips que fallan es de $\alpha = 10$. La probabilidad de que fallen 10 o más microchips es igual a 1 menos la probabilidad de que falle un número menor que 10. Es decir:

$$P(X \geq 10) = 1 - P(X \leq 9) = 1 - \sum_{k=0}^{9} \frac{\alpha^k}{k!} e^{-\alpha} = 1 - \left(\sum_{k=0}^{9} \frac{10^k}{k!} \right) e^{-10} = 0{,}542.$$

2.

a) Podemos modelizar el comportamiento de este emisor con una variable aleatoria discreta binomial que cuente el número de 1 obtenidos, puesto que cada bit se considera una variable aleatoria de Bernoulli y el número de experiencias será el número de bits emitidos.

b) Si denominamos X la variable aleatoria que representa el número de observaciones 1 (éxitos) que se obtienen al generar la secuencia de 5 dígitos, X sigue una distribución binomial de parámetros $n = 5$ y $p = 0{,}6$, es decir, $X \sim \text{Bin}(5, 0{,}6)$. Por lo tanto, $P(X = 2) = \binom{5}{2} 0{,}6^2 \cdot 0{,}4^3 = 0{,}23$.

c) Ahora se pide $P(X \geq 3) = 1 - P(X \leq 2)$.

$$P(X \leq 2) = \sum_{k=0}^{2} \binom{5}{k} 0{,}6^k \cdot 0{,}4^{5-k} = 0{,}4^5 + 5 \cdot 0{,}6 \cdot 0{,}4^4 + 10 \cdot 0{,}6^2 \cdot 0{,}4^3 = 0{,}317.$$

Por lo tanto, $P(X \geq 3) = 1 - 0,317 = 0,683$.

3.

a) X toma cualquier valor entre 0 y ∞. Es una variable de tipo exponencial, puesto que coincide con el tiempo de respuesta descrito en el enunciado. Por lo tanto, es continua.

b) Como hemos visto, $X \sim \text{Exp}(1/10)$. Por lo tanto, su función de distribución es $F(x) = 1 - e^{-x/10}$ para $x \geq 0$.

$$P(X < 5) = F(5) = 1 - e^{-1/2} = 0{,}393.$$

c) De manera análoga:

$$P(5 < X < 10) = F(10) - F(5) = (1 - e^{-1}) - (1 - e^{-1/2}) = e^{-1/2} - e^{-1} = 0{,}239.$$

4.

a) Si representamos por X = «número de veces que falla el direccionador en las 2.500 jornadas», y entendemos por «éxito» el hecho de que el direccionador falle, entonces: $X \sim \text{Bin}(2.500,\ 0{,}04)$.

En este caso, podemos definir como A, éxito, el hecho de que el direccionador falle. Esto pasa con una probabilidad de 0,04. El número de veces que repetimos el experimento, según el enunciado, es de 2.500. Por lo tanto, podemos considerar que nuestra variable aleatoria sigue una distribución binomial con parámetros $n = 2.500$, $p = 0{,}04$.

Con la aproximación de una binomial a una normal, podemos trabajar con la variable Y:

$$Y \sim \text{N}(2.500 \cdot 0{,}04,\ \sqrt{2.500 \cdot 0{,}04 \cdot 0{,}96}) = \text{N}(100,\ 9{,}798).$$

b) Lo que se pide es:

$$P(X > 120) = 1 - P(X \leq 120) = 1 - P(Y < 120{,}5) = 0{,}98179 - 0{,}47965 = 0{,}0182.$$

c) Ahora lo que se pide es:

$$P(100 \leq X \leq 120) = P(99{,}5 < Y < 120{,}5) = 1 - 0{,}98179 = 0{,}502.$$

(En los apartados b y c se ha utilizado software matemático para calcular la función de distribución de la variable normal Y.)

5. X es una variable de Poisson de parámetro $\alpha = \lambda T = 0,5 \cdot 10 = 5$.

a) La función de probabilidad de X es $P(X = k) = e^{-\alpha}\frac{\alpha^k}{k!}$ y, por lo tanto, $P(X = 0) = e^{-5}\frac{5^0}{0!} = 0,0067$.

b) Ahora nos piden $P(X \geq 2)$:

$$P(X \geq 2) = 1 - P(X \leq 1) = 1 - (P(X=0) + P(X=1)) = 1 - (e^{-5} + 5e^{-5}) = 1 - 6e^{-5} = 0,9596.$$

6.

a) Si representamos por X la variable aleatoria que representa el número de éxitos (mensajes recibidos) en las n pruebas, tenemos que X sigue una distribución binomial de parámetros n y p, es decir, $X \sim \text{Bin}(n, p)$.

b) En este caso, $n = 3$ y $p = 0{,}7$, y por tanto:

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{3}{0} 0{,}7^0 \cdot 0{,}3^3 = 1 - 0{,}027 = 0{,}973.$$

c) Ahora n es desconocido y $p = 0{,}8$. Se pide encontrar n tal que $P(X \geq 1) \geq 0{,}95$:

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{n}{0} 0{,}8^0 \cdot 0{,}2^n = 1 - 0{,}2^n.$$

Planteamos la inecuación: $1 - 0{,}2^n \geq 0{,}95 \Rightarrow n \geq \ln 0{,}05 / \ln 0{,}2 = 1{,}86$, es decir, tomaremos $n = 2$.

7.

a) Lo que se pide es: $P(75 < X < 95) = 0,68269 \approx 0,68$ (coherente con la regla).

b) Ahora lo que se pide es: $P(65 < X < 105) = 0,9545 \approx 0,95$ (coherente con la regla).

c) Por ejemplo:

Figura 11. Distribución de las notas obtenidas por los dos estudiantes. En el eje horizontal, se representa el número de prueba, y en el eje vertical, la nota obtenida.

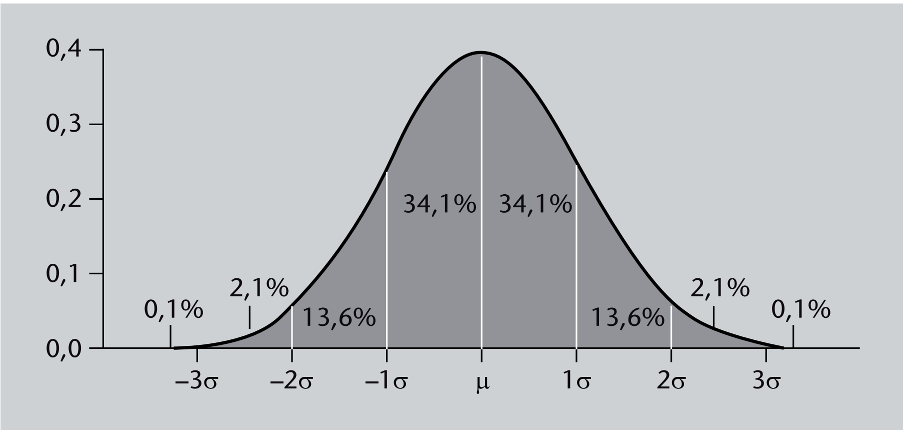


Figura 11

La parte en azul oscuro es menos de una desviación estándar desde la media. Para una distribución normal, esto representa el 68 % del conjunto (azul oscuro). La parte en azul menos oscuro está situada hasta dos desviaciones estándar y representa el 95 % del conjunto. La parte en azul claro (hasta tres desviaciones estándar) representa el 99,7 %.

d) La primera fila corresponde a la distribución original de la variable X. El resto de las filas corresponde a las distribuciones de las medias muestrales para diferentes tamaños de la muestra ($n = 4$, $n = 7$ y $n = 10$). Se observa que, con independencia de cómo se distribuya la variable X, las distribuciones de las medias muestrales se van aproximando a una distribución normal a medida que aumenta el tamaño de la muestra. Este resultado es conocido como teorema central del límite.

8. Para resolver este ejercicio, tendremos en cuenta que podemos modelizar la llegada de llamadas a una centralita mediante una distribución de Poisson, $X \sim \text{Poiss}(\alpha)$.

El parámetro que define la distribución, α, coincide con la esperanza de la variable, como hemos visto en el subapartado 2.2. El enunciado nos dice que llegan 300 llamadas en una hora. Por lo tanto, la tasa es $\lambda = \frac{300}{60} = 5$ llamadas/minuto.

a) La probabilidad de que el número de llamadas en un minuto sea mayor que 12 y de que, por tanto, la central quede saturada, es:

$$P(X > 12) = 1 - P(X \leq 12) = 1 - e^{-5} \sum_{k=0}^{12} \frac{5^k}{k!} = 0{,}00202.$$

b) La probabilidad de que se reciba una sola llamada en un minuto es:

$$P(X = 1) = e^{-5} \frac{5^1}{1!} = 0{,}0337.$$

9.

a) Si llegamos a la estación en cualquier momento y todos los momentos son equiprobables, podemos asumir que la variable aleatoria «tiempo de espera» sigue una distribución uniforme en el intervalo (0, 20). La función de distribución es la siguiente:

$$F(x) = P(X \leq x) = \begin{cases} 0 & \text{si } \leq 0, \\ \dfrac{x}{20} & \text{si } 0 \leq x < 20, \\ 1 & \text{si } \leq 20. \end{cases}$$

La función de densidad $f(x)$ es $\frac{1}{20}$ dentro del intervalo $(0, 20)$.

b) La probabilidad de que espere menos de 7 minutos es la siguiente:

$$P(X < 7) = F(7) = \frac{7}{20} = 0{,}35.$$

c) La esperanza de la variable «tiempo de espera» es:

$$\mathrm{E}(X) = \frac{0 + 20}{2} = 10.$$

Y la varianza es:

$$\mathrm{Var}(X) = \frac{(20-0)^2}{12} = \frac{100}{3} = 33{,}33.$$

La desviación es $\sigma = \sqrt{33,33} = 5{,}7$.

d) La probabilidad de que el tiempo de espera sea exactamente de 12 minutos es cero, ya que se trata de una variable aleatoria continua y, en este caso, la probabilidad de un valor determinado es cero.

10.

a) Representamos por X el número de horas que pasa hasta que fallan los 3 sistemas. X_1, X_2 y X_3 son el número de horas que pasa hasta que falla el primer, segundo y tercer sistema, respectivamente. Cada X_i es una variable geométrica de parámetro $p = 0{,}1$. Por lo tanto, $\mathrm{E}(X_1) = \mathrm{E}(X_2) = \mathrm{E}(X_3) = 1/p = 10$ horas. El tiempo total es $X = X_1 + X_2 + X_3$, de modo que $\mathrm{E}(X) = 3 \cdot 10 = 30$ horas.

b) A continuación, nos piden la probabilidad de que las 3 computadoras fallen en un vuelo de 5 horas, es decir:

$$P(X \leq 5) = P(X = 3) + P(X = 4) + P(X = 5).$$

$P(X = 3) = p^3 = 0{,}001$ ya que corresponde a $X_1 = X_2 = X_3 = 1$.

$P(X = 4) = 3p^3(1 - p) = 0,0027$ ya que corresponde a $X_1 = 2$, $X_2 = X_3 = 1$ o $X_2 = 2$, $X_1 = X_3 = 1$ o $X_3 = 2$, $X_1 = X_2 = 1$.

$P(X = 5) = 6p^3(1 - p)^2 = 0{,}00486$ ya que corresponde a que una valga 3 y las otras dos, 1 (3 maneras de hacerlo) o que dos valgan 2 y la otra, 1 (3 maneras de hacerlo).

Entonces

$$P(X \leq 5) = p^3 + 3p^3(1 - p) + 6p^3(1 - p)^2 = 0{,}00856.$$