

# Variables aleatorias

Carles Rovira Escofet

P08/75057/02305



Universitat Oberta  
de Catalunya

[www.uoc.edu](http://www.uoc.edu)



# Índice

## Sesión 1

### Introducción a las variables aleatorias

<b>Variables aleatorias discretas</b> .....	5
1. Introducción a las variables aleatorias .....	5
2. Variables aleatorias discretas .....	6
2.1. Función de masa de probabilidad y función de distribución .....	7
2.2. Independencia .....	9
3. Resumen .....	10
Ejercicios .....	11

## Sesión 2

<b>Esperanza y varianza</b> .....	15
1. Definiciones .....	15
2. Algunas propiedades de la esperanza .....	19
3. Algunas propiedades de la varianza .....	20
4. La desigualdad de Tchebichev .....	21
5. Resumen .....	22
Ejercicios .....	23

## Sesión 3

<b>Algunas distribuciones discretas</b> .....	27
1. Distribución de Bernoulli .....	27
2. Distribución binomial .....	28
3. Distribución geométrica .....	30
4. Distribución de Poisson .....	33
5. Resumen .....	35
Ejercicios .....	36

## Sesión 4

<b>Variables aleatorias continuas</b> .....	39
1. Función de densidad .....	39
2. Relación entre las funciones de distribución y de densidad. Cálculo de probabilidades .....	42
3. Independencia .....	44
4. Esperanza y varianza .....	45
5. Resumen .....	46
Ejercicios .....	47

## Sesión 5

<b>Algunas leyes continuas. La ley normal</b> .....	50
1. Distribución uniforme .....	50

2. Distribución exponencial .....	52
3. Distribución normal .....	54
3.1. El papel de los parámetros .....	55
3.2. Estandarizar .....	57
3.3. Cálculo de probabilidades utilizando la ley normal estándar .....	60
4. Resumen .....	62
Ejercicios .....	63

## Sesión 6

<b>Procesos estocásticos.....</b>	<b>65</b>
1. Definiciones .....	65
2. Procesos aleatorios a tiempo discreto: el paseo aleatorio .....	67
3. Procesos a tiempo continuo.....	69
3.1. El proceso de Poisson.....	69
3.2. El proceso de Wiener o movimiento browniano .....	71
4. Resumen.....	72

## Introducción a las variables aleatorias.

### Variables aleatorias discretas

En el módulo anterior hemos visto el concepto de probabilidad y cómo podíamos utilizarlo para modelar el resultado de un experimento.

Consideremos, por ejemplo, el espacio muestral que obtenemos si lanzamos una moneda tres veces. En este caso el espacio muestral tiene ocho elementos y es el siguiente:

$$\Omega = \{(C,C,C), (C,C,+), (+,C,C), (C,+,C), (C,+,+), (+,C,+), (+,+,C), (+,+,+)\}$$

donde C indica que ha salido una cara y + indica que ha salido una cruz. Tenemos también una probabilidad sobre este espacio muestral, en concreto sabemos que la probabilidad de cada elemento del espacio muestral es de  $1/8$ .

#### Espacio muestral

Es el espacio con todos los posibles resultados de nuestro experimento.

### 1. Introducción a las variables aleatorias

Sin embargo, podría interesarnos estudiar el número de caras que han salido. Obtenemos así una función, que llamaremos  $X$ , que a cada elemento del espacio muestral le hace corresponder un valor numérico, el correspondiente al número de caras, por ejemplo,  $X((C,C,+)) = 2$ .

Una vez que tenemos definida esta función  $X$ , podemos plantearnos preguntas como “¿cuál es la probabilidad de que salgan dos caras?” que podemos formalizar como el cálculo de la probabilidad del suceso  $\{X = 2\}$ . Fijaos en que:

$$\{X = 2\} = \{(C,C,+), (C,+,C), (+,C,C)\}$$

Así,

$$P(X = 2) = \frac{3}{8}$$

esta función  $X$  será lo que llamaremos una *variable aleatoria*.

Una **variable aleatoria** es una función que asocia un valor numérico a cada elemento del espacio muestral.

#### Otros ejemplos de variables aleatorias

Si consideramos como espacio muestral a los estudiantes de esta asignatura, la función que asocia a cada individuo con su peso es una variable aleatoria.

Si consideramos como espacio muestral el resultado del lanzamiento de cinco dados, podemos considerar diversas variables aleatorias como: el máximo de los cinco dados, el mínimo de los cinco dados o la suma de los cinco dados.

#### Alea jacta est

La palabra *aleatorio* proviene del hecho de que utilizamos las variables aleatorias para formalizar fenómenos en los que interviene el azar. Recordad que en las películas de romanos dicen *alea jacta est*, que en español sería ‘la suerte está echada’.

Estudiaremos básicamente dos tipos de variables aleatorias, las variables aleatorias discretas y las variables aleatorias continuas (existen otros tipos que no estudiaremos). Como idea básica, podemos diferenciarlas de la manera siguiente: cuando una variable aleatoria toma valores sólo sobre un conjunto finito, es una variable aleatoria discreta, y cuando toma valores sobre todo un intervalo de los números reales, es una variable aleatoria continua. Esta definición no es rigurosa, pero nos sirve para hacer una primera aproximación.

### Nuestro entorno...

... está lleno de variables aleatorias. Sólo un ejemplo: el número de letras de cada página de estas notas.

## 2. Variables aleatorias discretas

Una **variable aleatoria discreta** es una variable aleatoria que puede tomar un número finito o numerable de valores, en los que cada uno de estos valores tiene una probabilidad mayor que cero de ser observado.

### Conjunto numerable de valores

Es aquel que podemos numerar utilizando los números naturales.

Fijaos en los ejemplos que hemos dado de variables aleatorias. Algunas son variables aleatorias discretas y otras no lo son. El ejemplo de los dados es un caso claramente discreto, en el que la variable aleatoria que nos da el máximo de los cinco resultados sólo puede alcanzar uno de los seis valores siguientes: 1, 2, 3, 4, 5 y 6. Además, somos capaces de calcular la probabilidad de que la variable tome cada uno de estos valores.

### Ejemplo de las tres monedas

Si lanzamos tres veces una moneda, el conjunto de resultados posibles es:

$$\Omega = \{(c,c,c), (c,c,+), (+,c,c), (c,+,c), (c,+,+), (+,c,+), (+,+,c), (+,+,+)\}$$

donde  $c$  indica una cara y  $+$  indica una cruz. En este espacio muestral podemos definir muchas variables aleatorias diferentes. Llamaremos  $X$  al número de caras que han salido. Esta variable puede tomar los valores siguientes:

Valores posibles $X$	Resultados del lanzamiento del dado
0	$(+,+,+)$
1	$(c,+,+), (+,c,+), (+,+,c)$
2	$(c,c,+), (+,c,c), (c,+,c)$
3	$(c,c,c)$

Podemos ver cuál es la probabilidad asociada a cada uno de los posibles valores de la variable  $X$ . De la tabla anterior y sabiendo que los resultados son equiprobables, resulta muy claro que obtenemos las probabilidades siguientes:

Valores posibles $X = x$	$P(X = x)$
0	1/8
1	3/8
2	3/8
3	1/8

Utilizando esta tabla, es fácil calcular la probabilidad de cualquier acontecimiento que se pueda escribir en términos de la variable  $X$ . Por ejemplo, podemos contestar a la pregunta “¿cuál es la probabilidad de que salga más de una cara?”.

$$P(X > 1) = P(\{X = 2\} \cup \{X = 3\}) = P(X = 2) + P(X = 3) = \frac{4}{8}$$

Fijaos en que podemos tener sucesos con probabilidad cero. Por ejemplo, “¿cuál es la probabilidad de que salgan cuatro caras?”. ¡Claro que la probabilidad es cero!

## 2.1. Función de masa de probabilidad y función de distribución

Una variable aleatoria discreta quedará determinada si conocemos cuál es la probabilidad de que tome cada uno de los posibles valores. La función que nos da la probabilidad de cada valor se conoce como la *función de masa de probabilidad*.

La **función de masa de probabilidad** de una variable aleatoria es la función que nos da, para cada valor  $x$ , la probabilidad de que la variable tenga este valor,  $P(X = x)$ .

Una propiedad fundamental de esta función es que la suma de la función de masa de probabilidad en todos los posibles valores vale siempre 1, es decir:

$$\sum_x P(X = x) = 1$$

Recordad que la probabilidad del conjunto total era 1. Fijaos, por tanto, en que en la propiedad anterior sólo recordamos este resultado, ya que cuando sumamos las probabilidades de todos los valores posibles, obtenemos la probabilidad del conjunto total.

En definitiva, la función de masa de probabilidad nos explica cómo se distribuye una probabilidad total igual a la unidad en los resultados que toma la variable  $X$ .

Otra herramienta fundamental para el estudio de las variables aleatorias es la llamada *función de distribución*.

La **función de distribución** de una variable aleatoria es la función que nos da, para cada valor  $x$ , la probabilidad de que la variable tenga un valor inferior o igual a  $x$ , es decir:

$$F(x) = P(X \leq x)$$

### Sumatorios

El símbolo  $\sum_x f(x)$  indica hacer un sumatorio de los valores de  $f(x)$  para todos los posibles valores de la  $x$ .

### CDF

En inglés la función de distribución se denomina *cumulative distribution function*. Esto hace que en el *software* estadístico se denomine usualmente CDF.

Para las variables discretas tenemos:

$$F(x) = \sum_{y \leq x} P(X = y)$$

Esta función a veces resulta muy útil. Conociendo la función de masa de probabilidad, podemos calcular la función de distribución, lo cual resulta fácil de comprobar. Cuando conozcamos una de estas dos funciones, diremos que conocemos la **ley** o la **distribución** de la variable aleatoria discreta.

Conocer la ley o la distribución de una variable aleatoria  $X$  significa ser capaces de calcular cualquier probabilidad que haga referencia al comportamiento de la variable aleatoria  $X$ .

Así, diremos que tanto la función de masa de probabilidad como la función de distribución determinan la ley de la variable aleatoria. Observad que dos variables diferentes pueden tener la misma ley. Por ejemplo, si tiramos dos dados y consideramos las dos variables que nos dan el resultado de cada dado, estas dos variables aleatorias tienen la misma ley, pero no son la misma (corresponden a dos tiradas diferentes; en la primera ha podido salir un 1 y en la segunda ha podido salir un 2).

### Ejemplo de las tres monedas

Retomemos el ejemplo anterior de las tres monedas. A partir de la función de masa de probabilidad podemos calcular fácilmente la función de distribución. Por ejemplo:

$$P(X \leq 1) = P(X = 0) + P(X = 1) = \frac{4}{8}$$

Podemos recoger la función de distribución en la tabla siguiente:

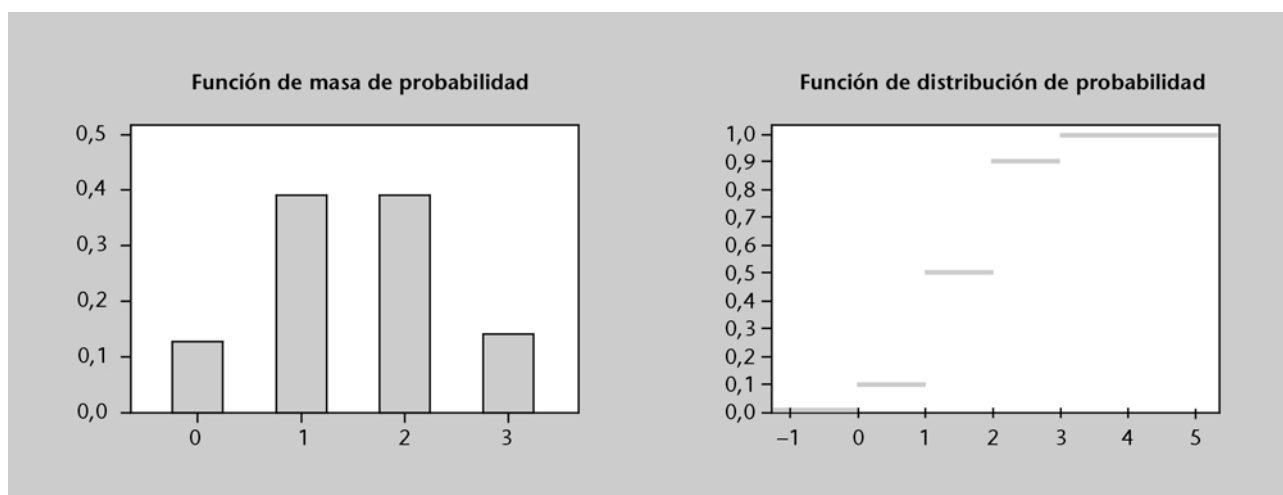
Valores $x$	$P(X \leq x)$
0	1/8
1	4/8
2	7/8
3	1

Fijaos en que podemos calcular la función de distribución para cualquier valor de la  $x$ . Por ejemplo:

$$F(2,7) = P(X \leq 2,7) = P(X \leq 2) = \frac{7}{8}$$

Podemos dibujar también los gráficos de la función de masa de probabilidad y de la función de distribución.





La función de distribución presenta algunas propiedades importantes que debemos conocer:

- Arranca siempre desde cero y llega hasta uno.
- Es siempre una función creciente.

#### Histograma de frecuencias

Si repetimos un experimento muchas veces, el histograma de las frecuencias relativas para cada valor se parece al gráfico de la función de masa de probabilidad.

## 2.2. Independencia

Para finalizar esta sesión, introducimos el concepto de **independencia**.

En el ejemplo de las tres monedas, la variable  $X$  nos daba el número de caras que han salido. Consideremos también la variable  $Y$ , que valdrá 1 si al tirar la moneda la primera vez sale cara y valdrá 0 si sale cruz. Finalmente definimos la  $Z$ , que valdrá 1 si al lanzar la moneda la segunda vez sale cara y valdrá 0 si sale cruz.

Está claro que si conocemos el valor de la variable  $Y$ , ya tenemos cierta información sobre los valores que puede tomar la variable  $X$ . Por ejemplo, si  $Y = 1$ , ya sabemos que  $X$  no puede ser cero. Esto nos indica que  $X$  e  $Y$  no son independientes.

En cambio, aunque conozcamos el valor que toma  $Y$ , no tenemos ninguna información sobre qué puede pasar con la variable  $Z$ . En este caso diremos que  $Y$  y  $Z$  son independientes.

La definición rigurosa de independencia entre variables aleatorias, que nos permitirá tratar fácilmente con parejas de variables aleatorias, es muy parecida al concepto de sucesos independientes que vimos en el módulo de “Cálculo de probabilidades”.

Las variables aleatorias discretas  $X$  e  $Y$  son **independientes** si:

$$P(X = x, Y = y) = P(X = x)P(Y = y)$$

para todo posible par de valores  $x$  e  $y$ .

La definición de independencia implica que para todo  $y$  con  $P(Y = y) > 0$ ,

$$P(X = x | Y = y) = P(X = x)$$

para todo  $x$ .

#### Notación

Observad que utilizamos la notación:

$$P(\{X = x\} \cap \{Y = y\}) = P(X = x, Y = y).$$

Y recordad que:  $P(X = x | Y = y)$  indica la probabilidad de que  $X = x$  si sabemos que  $Y = y$ .

#### Ejemplo de las tres monedas

Recordad que hemos definido las variables  $X$ ,  $Y$  y  $Z$  (inicio del apartado 1.2.2). Observad que:

$$P(X = 0, Y = 1) = 0 \text{ y } P(X = 0) = \frac{1}{8}, P(Y = 1) = \frac{1}{2}$$

Por tanto, está claro que  $X$  e  $Y$  no son independientes, ya que:

$$P(X = 0, Y = 1) \neq P(X = 0)P(Y = 1)$$

Por otro lado, si  $z$  e  $y$  son 0 ó 1:

$$P(Y = y, Z = z) = \frac{1}{4}, \text{ y } P(Z = z) = P(Y = y) = \frac{1}{2}$$

Por tanto, está claro que  $Z$  e  $Y$  son independientes, ya que siempre se cumple:

$$P(Z = z, Y = y) = P(Z = z)P(Y = y)$$

### 3. Resumen

En esta sección hemos aprendido el concepto de variable aleatoria y hemos estudiado un tipo particular: las variables aleatorias discretas. Hemos presentado los conceptos de función de masa de probabilidad y función de distribución. Finalmente, hemos extendido el concepto de independencia de sucesos a variables aleatorias.

## Ejercicios

1. Lanzamos dos veces un dado y consideramos las variables aleatorias siguientes:  $Y$  es el resultado del primer dado,  $Z$  es el resultado del segundo y  $X$  es la suma de ambos.

- Determinad el espacio muestral asociado al experimento.
- Determinad las funciones de masa de probabilidad y de distribución de  $X$ . Dibujadlas.
- Calculad  $P(X = 2, Y = 3)$ ,  $P(X = 8)$ ,  $P(X = 7, 5)$ ,  $P(X > 7)$ ,  $P(X \leq 7, 5)$ .
- ¿Son independientes las variables  $X$  e  $Y$ ? ¿Son independientes las variables  $Z$  e  $Y$ ?

2. Un estudiante debe contestar un cuestionario con cuatro preguntas. Todas las preguntas tienen dos posibles respuestas. El estudiante, que no se ha preparado para el cuestionario, contesta las preguntas al azar. Sea  $X$  el número total de respuestas correctas del estudiante e  $Y$  el número de respuestas incorrectas a las tres primeras preguntas.

- Calculad la función de masa de probabilidad de  $X$ .
- Calculad la probabilidad de que el estudiante conteste al menos una pregunta bien.
- Calculad  $P(X = 2, Y = 3)$ .
- ¿Son independientes las variables  $X$  e  $Y$ ?

## Solucionario

1.

- Si lanzamos dos veces un dado, los resultados posibles son todos los pares de números del 1 al 6. Fijaos en que el conjunto de estos pares forma el espacio muestral y tiene 36 elementos (tenemos en cuenta el orden):

$$\Omega = \{(1,1), (1,2), (1,3), \dots, (6,5), (6,6)\}$$

- Para la variable  $X$  tenemos:

Valores posibles $X$	Resultados lanzamiento del dado
2	(1,1)
3	(1,2), (2,1)
4	(1,3), (2,2), (3,1)
5	(1,4), (2,3), (3,2), (4,1)
6	(1,5), (2,4), (3,3), (4,2), (5,1)
7	(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)
8	(2,6), (3,5), (4,4), (5,3), (6,2)
9	(3,6), (4,5), (5,4), (6,3)
10	(4,6), (5,5), (6,4)
11	(5,6), (6,5)
12	(6,6)

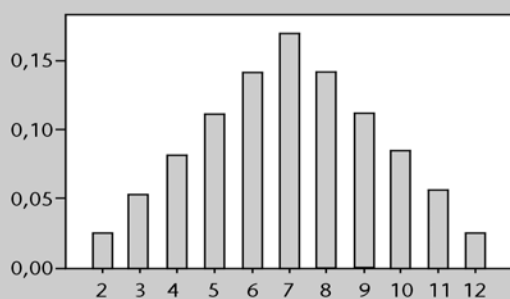
La función de masa de probabilidad es, por tanto:

Valores posibles $X = x$	$P(X = x)$
2	1/36
3	2/36
4	3/36
5	4/36
6	5/36
7	6/36
8	5/36
9	4/36
10	3/36
11	2/36
12	1/36

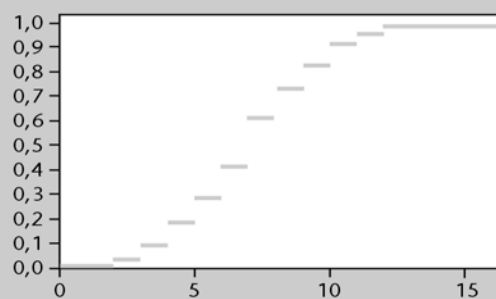
Y la función de distribución:

$x$	$P(X \leq x)$
2	1/36
3	3/36
4	6/36
5	10/36
6	15/36
7	21/36
8	26/36
9	30/36
10	33/36
11	35/36
12	36/36

**Función de masa de probabilidad**



**Función de distribución de probabilidad**



c) Utilizando estas tablas resulta fácil calcular la probabilidad de los acontecimientos que se pueden escribir en términos de la variable  $X$ . Así:

$$P(X = 8) = \frac{5}{36}$$

$$P(X = 7,5) = 0$$

$$P(X > 7) = 1 - P(X \leq 7) = 1 - F(7) = 1 - \frac{21}{36} = \frac{15}{36}$$

$$P(X \leq 7,5) = P(X \leq 7) = \frac{21}{36}$$

También está claro que el suceso  $\{X = 2, Y = 3\}$  es el suceso vacío, así que:

$$P(X = 2, Y = 3) = 0$$

d) Observad:

$$P(X = 8, Y = 1) = 0$$

y en cambio:

$$P(X = 8) = \frac{5}{36} \quad \text{y} \quad P(Y = 1) = \frac{1}{6}$$

En este caso está claro que  $X$  e  $Y$  no son independientes.

Por otro lado, si  $z$  e  $y$  son dos valores entre 1 y 6:

$$P(Y = y, Z = z) = \frac{1}{36}, \text{ y } P(Z = z) = P(Y = y) = \frac{1}{6}$$

Por tanto,  $Z$  e  $Y$  son independientes.

2.

a) Si decimos que la respuesta a cada pregunta es independiente de las otras, podemos calcular fácilmente:

$$P(X = 0) = \left(\frac{1}{2}\right)^4$$

$$P(X = 1) = 4\left(\frac{1}{2}\right)^4$$

$$P(X = 2) = 6\left(\frac{1}{2}\right)^4$$

$$P(X = 3) = 4\left(\frac{1}{2}\right)^4$$

$$P(X = 4) = \left(\frac{1}{2}\right)^4$$

Los coeficientes 4, 6 y 4 salen del número de combinaciones en las que podemos hacer 1, 2 ó 3 respuestas correctas respectivamente. Este cálculo está relacionado con la llamada *distribución binomial*, que se explica en la sesión “Algunas distribuciones discretas”.

b) El acontecimiento “contestar al menos alguna pregunta bien” es el acontecimiento “contestar una pregunta o más bien”. Así, tenemos que calcular:

$$P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0) = 1 - \frac{1}{16}$$

c) Está claro que el acontecimiento  $\{X = 2, Y = 3\}$  no se puede dar nunca. Por tanto:

$$P(X = 2, Y = 3) = 0$$

d) Ya hemos visto que la  $P(X = 2) = 6/16$ . Con un razonamiento análogo es fácil comprobar que:

$$P(Y = 3) = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$$

Por tanto,

$$P(X = 2, Y = 3) \neq P(X = 2)P(Y = 3)$$

y las variables no son independientes.

## Esperanza y varianza

En esta sesión estudiaremos un par de características importantes de las leyes de las variables aleatorias discretas: la esperanza y la varianza. Se trata de una medida numérica por el centro y una por la dispersión, respectivamente, de la distribución de la variable aleatoria. Dos variables aleatorias con la misma ley tendrán, por tanto, la misma esperanza y la misma varianza.

No serán conceptos radicalmente nuevos, ya que están relacionados con los conceptos de media muestral y de varianza muestral que estudiamos anteriormente. Las medidas de centralidad y de dispersión de un conjunto de datos tienen su equivalente en las distribuciones de probabilidad.

### Representabilidad de la esperanza y la varianza

La esperanza y la varianza nos permiten conocer un poco cómo es una variable aleatoria.

### 1. Definiciones

Suponed que lanzamos un dado veinte veces y que obtenemos los valores siguientes:

2, 2, 6, 1, 1, 6, 2, 5, 3, 2, 4, 6, 4, 4, 4, 5, 2, 6, 6, 5, 4

La media de estos datos, utilizando las frecuencias relativas, tal como hemos visto en el módulo “Estadística descriptiva”, se puede calcular como:

$$\bar{x} = \sum_{i=1}^6 i f_i = 1\left(\frac{2}{20}\right) + 2\left(\frac{5}{20}\right) + 3\left(\frac{1}{20}\right) + 4\left(\frac{4}{20}\right) + 5\left(\frac{3}{20}\right) + 6\left(\frac{5}{20}\right) = 3,8$$

donde cada  $f_i$  indica la frecuencia relativa de  $i$ . De la misma manera podríamos calcular la varianza muestral de los datos como:

$$s^2 = \sum_{i=1}^6 (i - \bar{x})^2 f_i \approx 3,06$$

Por otra parte, podemos pensar que si tenemos un dado (no trucado) y lo lanzamos muchas veces, las frecuencias esperadas de cada uno de los 6 posibles valores se irá acercando cada vez más a  $1/6$  (que es la probabilidad de cada valor). Así, la media de los datos cuando hiciéramos muchas tiradas sería:

$$\sum_{i=1}^6 i \frac{1}{6} = \frac{21}{6} = 3,5$$

### Recordad cómo...

... calculábamos la media y la varianza muestral.

y la varianza muestral:

$$\sum_{i=1}^6 (i - 3,5)^2 \frac{1}{6} \approx 2,916$$

Estos dos valores, equivalentes a la media y a la varianza muestral, serán la esperanza y la varianza de la variable aleatoria que nos da el resultado del lanzamiento de un dado. Hemos sustituido las frecuencias relativas de cada valor por la probabilidad de que se dé este valor.

### ¿Qué esperamos?

Como su propio nombre indica, la esperanza de una variable aleatoria es el valor que “esperamos” que salga. Pero ¿qué significa que “esperamos” que salga? Pensad en el ejemplo del dado.

La **esperanza** de una variable aleatoria discreta  $X$ , que denotaremos por  $E(X)$ , se define como:

$$E(X) = \sum_x xP(X = x)$$

### Notación

En algunos libros la esperanza se denomina *valor esperado*.

La esperanza de una variable aleatoria  $X$  se puede mirar también como la media ponderada de los valores que toma la variable aleatoria.

### Ejemplo de la ruleta

Un ejemplo muy fácil es el de la ruleta. Podemos plantearnos la pregunta siguiente. Si vamos al casino a jugar a la ruleta a rojo o negro, ¿qué ganancias esperamos tener?

Para aquellos que no tengáis este vicio, recordemos cómo funciona una ruleta. La ruleta tiene 18 casillas rojas, 18 negras y 2 blancas. Si jugamos 1.000 u.m. a rojo y la bola cae en una casilla roja, ganamos 1.000 u.m., y si cae en una negra o en una blanca, perdemos las 1.000 u.m. que hemos jugado. Si llamamos  $X$  a las ganancias que esperamos obtener, tenemos la tabla siguiente:

Valores posibles $X = x$	$P(X = x)$
1.000	18/38
-1.000	20/38

Y podemos calcular la esperanza:

$$E(X) = 1.000 \frac{18}{38} - 1.000 \frac{20}{38} = -52,63$$

Es decir, en cada jugada “esperamos” perder 52,63 u.m. A largo plazo, un mal negocio para el jugador y un buen negocio para el casino.

### Ejemplo de la bolsa

Veamos ahora un ejemplo en los mercados financieros. Compramos una acción en bolsa de la compañía A por 1.000 u.m. y creemos que con probabilidad 1/3 la semana siguiente valdrá 1.200 u.m., con probabilidad 1/2 el precio no variará y con probabilidad 1/6 valdrá sólo 900 u.m. ¿Os parece que hemos hecho una buena inversión?

Para decidirlo, deberíamos calcular el valor que “esperamos” que tendrá la semana que viene. Podemos considerar que tenemos una variable aleatoria  $X$ , que es el precio de la acción la próxima semana y que sabemos que tiene función de masa de probabilidad:

Valores posibles $X = x$	$P(X = x)$
1.200	1/3
1.000	1/2
900	1/6



Y podemos calcular la esperanza:

$$E(X) = 1.200 \frac{1}{3} + 1.000 \frac{1}{2} + 900 \frac{1}{6} = 1.050$$

Por tanto, esperamos obtener unos beneficios de 50 u.m.

Consideremos también la compañía B. Podemos comprar sus acciones también por 1.000 u.m., y sabemos que con probabilidad  $1/3$  la semana siguiente valdrán 1.500 u.m., con probabilidad  $1/2$  el precio no variará y con probabilidad  $1/6$  valdrá sólo 300 u.m. (¡qué desastre!).

Podemos calcular la esperanza de la variable  $Y$ , que nos dará el precio de la acción la semana siguiente:

$$E(X) = 1.500 \frac{1}{3} + 1.000 \frac{1}{2} + 300 \frac{1}{6} = 1.050$$

que es la misma que teníamos si comprábamos una acción de la compañía A.

Por tanto, parece que es indiferente comprar acciones de la compañía A o de la compañía B. Pero ya veremos que esto **no** es cierto.

Una vez calculada la esperanza, podemos calcular la dispersión para cada valor  $x$  que puede tomar la variable aleatoria  $X$ , como la distancia entre el valor  $x$  y la esperanza  $E(X)$ , es decir,  $x - E(X)$ . La varianza, medida de la dispersión, es en realidad el valor esperado de las desviaciones al cuadrado. El hecho de que las elevemos al cuadrado es para evitar que si tenemos desviaciones de diferentes signos, éstas se pudieran cancelar.

La **varianza** de una variable aleatoria  $X$ , que denotaremos por  $Var(X)$ , se define como:

$$Var(X) = E[(X - E(X))^2]$$

Si se utiliza la definición de esperanza que hemos dado, es fácil comprobar que para una variable aleatoria discreta podemos calcular la varianza como:

$$Var(X) = \sum_x (x - E(X))^2 P(X = x)$$

Tenemos también una expresión alternativa de la varianza:

$$Var(X) = E[X^2] - (E[X])^2$$

Como en el estudio poblacional (estadística descriptiva), podemos definir también la desviación típica.

La **desviación típica** de una variable aleatoria  $X$ , que denotaremos por  $s(X)$ , se define como:

$$s(X) = \sqrt{Var(x)}$$

#### Medida de la dispersión de una variable aleatoria

Como en el caso poblacional, la desviación típica es muy útil, ya que se da con las mismas unidades que la esperanza, de manera que nos puede dar una idea más clara de la medida de la dispersión de una variable aleatoria.

Como ya hemos comentado, la varianza es una medida de dispersión. Así, mientras que la esperanza nos da el centro de la distribución (hace el papel de la media muestral en “Estadística descriptiva”), la varianza nos indica lo dispersa que es esta distribución.

Veamos el ejemplo siguiente:

### Ejemplo de la bolsa

En el ejemplo de los mercados financieros, para la compañía A tenemos las dispersiones siguientes:

Valores posibles $X = x$	Dispersión $x - E(x)$	$P(X = x)$
1.200	150	1/3
1.000	-50	1/2
900	-150	1/6

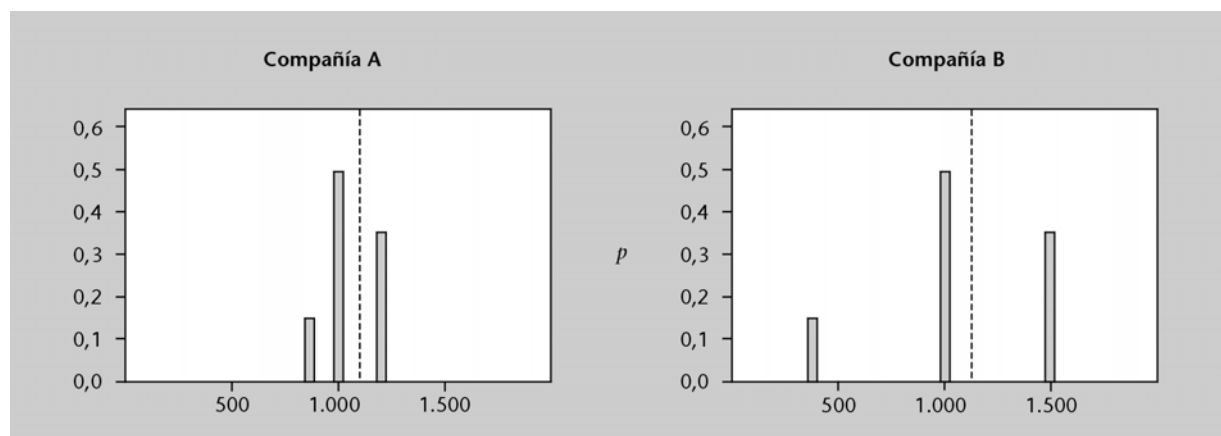
Y, por tanto, la varianza será ésta:

$$\text{Var}(X) = 150^2 \frac{1}{3} + (-50)^2 \frac{1}{2} + (-150)^2 \frac{1}{6} = 12.500$$

y la desviación típica es de 111,80. En cambio, si calculamos la varianza para la variable Y (compañía B):

$$\text{Var}(Y) = (1.500 - 1.050)^2 \frac{1}{3} + (1.000 - 1.050)^2 \frac{1}{2} + (300 - 1.050)^2 \frac{1}{6} = 162.500$$

mientras que la desviación típica es de 403,11. Como podemos ver, la varianza de la variable Y es bastante mayor que la de la variable X. Así, X e Y tienen la misma esperanza (el centro de las dos variables está en el mismo punto), pero las varianzas son bastante diferentes. Podemos verlo en los gráficos de las correspondientes funciones de masa de probabilidad.



¿Cómo podemos interpretarlo? La ganancia que esperamos al comprar una acción de la compañía A o una de la compañía B es la misma, pero en la compañía B la varianza es mayor, de manera que tanto podemos ganar más dinero, como podemos perder más. El caso extremo sería una compañía en la que compráramos una acción por 1.000 u.m. y supiéramos que a la semana siguiente podremos venderla por 1.050. La esperanza será la misma pero, en cambio, la varianza sería cero.

Tenemos todavía una última compañía llamada C. En este caso, el precio es también de 1.000 u.m. y sabemos que con probabilidad 1/3 la semana siguiente valdrá 1.300 u.m., con probabilidad 1/2 el precio será de 800 y con probabilidad 1/6 valdrá sólo 100 u.m.

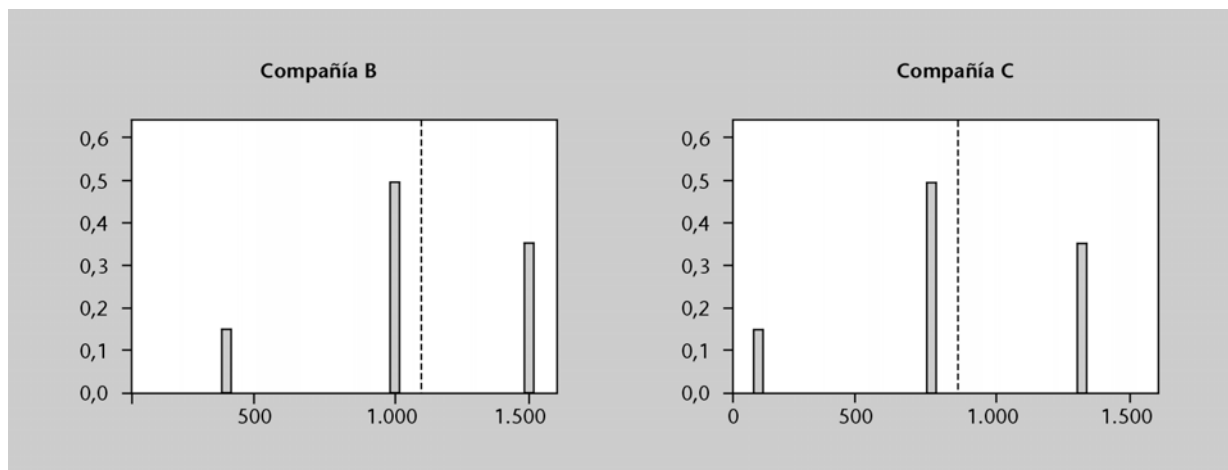
Podemos calcular la esperanza de la variable  $Z$ , indicada por una línea discontinua, que nos da el precio de la acción para la semana siguiente:

$$E(Z) = 1.300 \cdot \frac{1}{3} + 800 \cdot \frac{1}{2} + 100 \cdot \frac{1}{6} = 850$$

que es menor que la de las otras compañías. La varianza, en cambio, vale:

$$\text{Var}(Z) = (1.300 - 850)^2 \cdot \frac{1}{3} + (800 - 850)^2 \cdot \frac{1}{2} + (100 - 850)^2 \cdot \frac{1}{6} = 162.500$$

que es la misma que para la compañía B. Así,  $Z$  e  $Y$  tienen la misma varianza y esperanzas diferentes. Podemos verlo también comparando los gráficos de las funciones de masa de probabilidad.



## 2. Algunas propiedades de la esperanza

La mayoría de las propiedades de la esperanza podemos obtenerlas de una manera bastante intuitiva. Si en el ejemplo de la ruleta sabíamos que cuando jugábamos una vez “esperábamos” perder 52,63 u.m., si jugamos dos veces “esperaremos” perder el doble. En términos de esperanzas, estamos diciendo que la esperanza de la suma de dos variables aleatorias es la suma de las esperanzas.

También nos encontramos con las propiedades siguientes:

1) La esperanza de una constante es la misma constante. Una constante es una variable aleatoria que toma un único valor  $k$  con probabilidad 1. Así:

$$E(X) = kP(X = k) = k \cdot 1 = k$$

2) La esperanza de  $kX$ , donde  $X$  es una variable aleatoria y  $k$  es una constante, es igual a la esperanza de  $X$  multiplicada por  $k$ :

$$E(kX) = kE(X)$$

La demostración de esta propiedad se obtiene directamente de la definición de esperanza:

$$E(kX) = \sum_x kxP(X = x) = k \sum_x xP(X = x) = kE(X)$$

Observad que si tomamos  $k = -1$ , encontramos que  $E(-X) = -E(X)$ .

3) Si tenemos dos variables aleatorias  $X$  e  $Y$ , la esperanza de la suma de las dos es la suma de las dos esperanzas. Es decir:

$$E(X + Y) = E(X) + E(Y)$$

No demostraremos esta propiedad. Veremos sólo un caso particular: si sumamos una constante a la variable  $X$ , la esperanza aumenta también en la misma constante, es decir:

$$E(X + k) = E(X) + k$$

Para demostrarlo, sólo hay que utilizar la definición de esperanza y pensar que la nueva variable  $X + k$  toma los valores  $x + k$ , donde las  $x$  son los valores que toma la variable  $X$  con probabilidad  $P(X = x)$ . Veámoslo:

$$\begin{aligned} E(X + k) &= \sum_x (x + k) P(X = x) = \sum_x x P(X = x) + \sum_x k P(X = x) = \\ &= E(X) + k \sum_x P(X = x) = E(X) + k \end{aligned}$$

#### Extensión de $n$ variables

Esta propiedad se extiende fácilmente a  $n$  variables, de manera que la esperanza de la suma de  $n$  variables aleatorias es la suma de sus esperanzas.

#### Recordad que...

$$\dots \sum_x P(X = x) = 1$$

### 3. Algunas propiedades de la varianza

Teniendo en cuenta que la varianza es una medida de dispersión de la variable, es fácil justificar algunas de sus propiedades, aunque no otras. Veámoslas:

1) La varianza de una constante es cero. Una constante es una variable aleatoria que toma un único valor  $k$  con probabilidad 1. Así:

$$\text{Var}(X) = E((k - E(k))^2) = E(0) = 0$$

Este resultado es el esperado, ya que la varianza es una medida de la dispersión de la variable, de manera que, si la variable es constante, esta medida debe ser cero.

2) La varianza de  $kX$ , donde  $X$  es una variable aleatoria y  $k$  es una constante, es igual a la varianza de  $X$  multiplicada por  $k^2$ :

$$\text{Var}(kX) = k^2 \text{Var}(X)$$

Por tanto,  $s(kX) = k s(X)$ :

La demostración de esta propiedad se obtiene de la definición y de las propiedades de la esperanza:

$$\text{Var}(kX) = E((kX - E(kX))^2) = E(k^2(X - E(X))^2) = k^2 \text{Var}(X)$$

Observad que, si tomamos  $k = -1$ , obtenemos  $\text{Var}(-X) = \text{Var}(X)$ .

3) Si sumamos una constante a la variable  $X$ , la varianza no varía:

$$\text{Var}(k + X) = \text{Var}(X)$$

Este resultado es también intuitivo si pensamos que la varianza es una medida de dispersión y que, al sumar una constante, la dispersión de la variable no se modifica. La demostración es parecida a la de la propiedad anterior:

$$\begin{aligned}\text{Var}(k + X) &= E((k + X - E(k + X))^2) = E((k + X - k - E(X))^2) = \\ &= E((X - E(X))^2) = \text{Var}(X)\end{aligned}$$

4) Si tenemos dos variables aleatorias  $X$  e  $Y$  independientes, la varianza de la suma de las dos es la suma de las varianzas. Es decir:

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$$

No demostraremos esta propiedad. La independencia entre las dos variables resulta fundamental para obtener esta propiedad. Por ejemplo, si tomamos ( $Y = -X$ ), sabemos que la  $\text{Var}(X + Y) = \text{Var}(0) = 0$ , mientras que si la propiedad fuese cierta, tendríamos  $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(-X) = 2\text{Var}(X) \neq 0$ .

#### Obsevad que...

... si tenemos dos variables aleatorias  $X$  e  $Y$  que no son independientes, entonces en general:

$$\text{Var}(X + Y) \neq \text{Var}(X) + \text{Var}(Y).$$

Aunque dos variables aleatorias  $X$  e  $Y$  sean independientes, en general:

$$s(X + Y) \neq s(X) + s(Y).$$

#### Extensión a $n$ variables

Esta propiedad se extiende fácilmente a  $n$  variables, de manera que la varianza de la suma de  $n$  variables aleatorias independientes es la suma de las varianzas.

## 4. La desigualdad de Tchebichev

La desigualdad de Tchebichev nos permite interpretar la información conjunta que nos dan la esperanza y la varianza de una variable aleatoria indicando la masa de probabilidad mínima en intervalos de longitud  $2n\sqrt{\text{Var}(X)}$  centrados en  $E(X)$ .

Sea  $X$  una variable aleatoria, entonces la desigualdad de Tchebichev nos indica que para todo  $k > 0$ :

$$P(|X - E(X)| \geq k) \leq \frac{1}{k^2} \text{Var}(X)$$

#### Demostración de la desigualdad de Tchebichev

Por la definición de la varianza tenemos:

$$\begin{aligned}\text{Var}(X) &= \sum_x (x - E(X))^2 P(X = x) = \\ &= \sum_{x, |x - E(X)| \geq k} (x - E(X))^2 P(X = x) + \sum_{x, |x - E(X)| < k} (x - E(X))^2 P(X = x)\end{aligned}$$

en la que en el último paso hemos partido el sumatorio en dos fragmentos. En el primero sumamos las  $x$ , donde  $|x - E(X)| \geq k$ ; en el segundo sumamos las que  $|x - E(X)| < k$ . Puesto que el segundo sumando es positivo, obtenemos:

$$\text{Var}(X) \geq \sum_{x, |x - E(X)| \geq k} (x - E(X))^2 P(X = x) \geq \sum_{x, |x - E(X)| \geq k} k^2 P(X = x) = k^2 P(|X - E(X)| \geq k)$$

que nos prueba la desigualdad de Tchebichev.

Esta desigualdad es importante, ya que nos indica que si cogemos  $k = n\sqrt{\text{Var}(X)}$ , es decir,  $n$  veces la desviación típica, tendremos:

$$P(|X - E(X)| \geq n\sqrt{\text{Var}(x)}) \leq \frac{1}{n^2}$$

Es decir, que tendremos como mínimo una masa de probabilidad de  $1 - \frac{1}{n^2}$  en el intervalo:

$$(E(X) - n\sqrt{\text{Var}(x)}, E(X) + n\sqrt{\text{Var}(x)})$$

### Ejemplo

Consideremos una variable aleatoria  $X$  con esperanza 2 y varianza 4, en cuyo caso:

$$P(|X - 2| \geq 6) \leq \frac{1}{9}$$

es decir:

$$P(-4 < X < 8) \geq 1 - \frac{1}{9} = \frac{8}{9}$$

### Probabilidad para $n = 2$

Para cualquier variable  $X$  en el intervalo:

$(E(X) - 2s, E(X) + 2s)$   
se concentra como mínimo  
una probabilidad de 0,75.  
Recordad que:

$$s = \sqrt{\text{Var}(X)}$$

### Observación

Fijaos en que  $6 = 3\sqrt{4}$

## 5. Resumen

En esta sesión hemos aprendido los conceptos de esperanza y de varianza de una variable aleatoria. Hemos estudiado sus propiedades básicas y hemos visto cómo calcularlas para variables aleatorias discretas. Para acabar, hemos visto cómo utilizar la desigualdad de Tchebichev para relacionar los dos conceptos.

## Ejercicios

1. Lanzamos dos veces un dado y consideramos la variable aleatoria que nos da el máximo de los dos valores obtenidos.

- Calculad la función de masa de probabilidad de esta variable aleatoria.
- Calculad su esperanza y su varianza.

2. En la compañía ACBXX tienen cuatro tipos de trabajadores: directivos, con un sueldo de 15.000.000 u.m.; especialistas, con un sueldo de 5.000.000; obreros, con un sueldo de 2.000.000, y aprendices, con un sueldo de 1.000.000. Sabemos que la compañía tiene cuatro directivos, cuarenta especialistas, doscientos obreros y cien aprendices. Cogemos a un trabajador al azar y llamamos  $X$  a la variable aleatoria que nos da el sueldo de este trabajador.

- Calculad la esperanza y la varianza de esta variable.
- Si en la compañía se hace un aumento lineal de 200.000 a todos los trabajadores, ¿cómo varía la esperanza y la varianza de la variable  $X$ ?
- ¿Y si hacemos un aumento proporcional del 10%?

3. Tenemos una variable aleatoria  $X$  con esperanza 10 y varianza 1.

- Calculad  $E(3X)$ ,  $Var(2X + X)$ .
- Utilizando la desigualdad de Tchebichev, encontrad una cota inferior para  $P(8 < X < 12)$ .

## Solucionario

1.

a) Para calcular la función de masa de probabilidad, observad el cuadro siguiente con todos los posibles resultados (fijaos en que hemos tenido en cuenta el orden):

(1,1)	(2,1)	(3,1)	(4,1)	(5,1)	(6,1)
(1,2)	(2,2)	(3,2)	(4,2)	(5,2)	(6,2)
(1,3)	(2,3)	(3,3)	(4,3)	(5,3)	(6,3)
(1,4)	(2,4)	(3,4)	(4,4)	(5,4)	(6,4)
(1,5)	(2,5)	(3,5)	(4,5)	(5,5)	(6,5)
(1,6)	(2,6)	(3,6)	(4,6)	(5,6)	(6,6)

Hemos marcado la tabla con dos colores para distinguir mejor qué casos nos da cada valor. Por ejemplo, el máximo igual a 1 sólo lo obtendremos con el resultado (1,1). En cambio, obtenemos que el máximo es igual a 2 en tres situaciones (1,2), (2,2) y (2,1). Si utilizamos esta tabla y los resultados son equiprobables, podemos construir fácilmente la función de masa de probabilidad.

Valores posibles $X = x$	$P(X = x)$
1	1/36
2	3/36
3	5/36
4	7/36
5	9/36
6	11/36

b) Podemos calcular fácilmente la esperanza:

$$E(X) = 1\left(\frac{1}{36}\right) + 2\left(\frac{3}{36}\right) + 3\left(\frac{5}{36}\right) + 4\left(\frac{7}{36}\right) + 5\left(\frac{9}{36}\right) + 6\left(\frac{11}{36}\right) = \frac{161}{36} \approx 4,47$$

Para calcular la varianza, haremos una tabla con las desviaciones:

Valores posibles $X = x$	$P(X = x)$	Dispersión $x - E(X)$	$ x - E(X) ^2$
1	1/36	-125/36	15.625/(36) <sup>2</sup>
2	3/36	-89/36	7.921/(36) <sup>2</sup>
3	5/36	-53/36	2.809/(36) <sup>2</sup>
4	7/36	-17/36	289/(36) <sup>2</sup>
5	9/36	19/36	361/(36) <sup>2</sup>
6	11/36	55/36	3.025/(36) <sup>2</sup>

Y si utilizamos la fórmula:

$$Var(X) = \sum_x (x - E(X))^2 P(X = x)$$

obtenemos fácilmente  $Var(X) = \frac{91.980}{36^3} \approx 1,971$

2.

a) Tenemos una variable  $X$  que tiene función de masa de probabilidad:

Valores posibles $X = x$	$P(X = x)$
15.000.000	4/344
5.000.000	40/344
2.000.000	200/344
1.000.000	100/344



Podemos calcular la esperanza:

$$E(X) = 15.000.000\left(\frac{4}{344}\right) + 5.000.000\left(\frac{40}{344}\right) + 2.000.000\left(\frac{200}{344}\right) + 1.000.000\left(\frac{100}{344}\right) \approx 2.209.302,33$$

Y también la varianza:

$$\begin{aligned} Var(X) &= (15.000.000 - E(X))^2\left(\frac{4}{344}\right) + (5.000.000 - E(X))^2\left(\frac{40}{344}\right) + \\ &+ (2.000.000 - E(X))^2\left(\frac{200}{344}\right) + (1.000.000 - E(X))^2\left(\frac{100}{344}\right) \approx 3.258.518.117.902 \end{aligned}$$

Como podéis ver, la varianza no nos da una idea clara del nivel de dispersión y, en cambio, si miramos la desviación típica, 1.805.137 u.m., lo vemos más claro.

**b)** No haremos cálculos. Fijaos en que si llamamos  $Y$  a la variable que nos da el sueldo después del aumento lineal, tenemos que  $Y = X + 200.000$ . Por las propiedades de la esperanza, tendremos  $E(Y) = E(X) + 200.000$ , mientras que la varianza no variará.

**c)** Tampoco haremos cálculos. Si ahora llamamos  $Z$  a la variable que nos da el sueldo después del aumento proporcional, tenemos que  $Z = (1,1)X$ . Por las propiedades de la esperanza y de la varianza, tendremos  $E(Z) = 1,1E(X)$ , mientras que la varianza cumple  $Var(Z) = (1,1)^2 Var(X) = (1,21)Var(X)$ , de manera que representa un incremento del 21%.

Como podéis ver en este caso, la varianza de los sueldos aumenta, de manera que parece más justo hacer un aumento lineal que hacer uno proporcional.

**3.**

**a)** Si utilizamos las propiedades que hemos estudiado, obtenemos:

$$\begin{aligned} E(3X) &= 3E(X) = 30 \\ Var(2X + X) &= Var(3X) = 9Var(X) = 9 \end{aligned}$$

**b)** Tenemos:

$$P(8 < X < 12) = P(-2 < X - 10 < 2) = P(|X - 10| < 2) = 1 - P(|X - 10| \geq 2)$$

Utilizando la desigualdad de Tchebichev, tenemos:

$$P(|X - 10| \geq 2) \leq \frac{1}{2^2} Var(V) = \frac{1}{4} \cdot 1 = \frac{1}{4}$$

Por tanto, la desigualdad pedida quedará:

$$P(8 < X < 12) = 1 - P(|X - 10| \geq 2) \geq 1 - \frac{1}{4} = \frac{3}{4}$$

## Algunas distribuciones discretas

Ahora estudiaremos algunas de las distribuciones discretas más usuales. Imaginemos la situación siguiente. Sabemos que cada vez que intentemos conectarnos a Internet utilizando la línea telefónica, cabe la posibilidad de que la conexión falle y no podamos conectarnos. Además, sabemos que la probabilidad de que esto suceda, es decir, de que falle la conexión, es de 0,1. Finalmente, supondremos también que el resultado de cada llamada que intentemos es independiente de las otras llamadas que hayamos hecho. Indicaremos finalmente con una  $C$  el acontecimiento “llamada que obtiene la conexión” y con  $\bar{C}$ , el acontecimiento complementario, “llamada que no obtiene la conexión”.

### 1. Distribución de Bernoulli

Consideremos ahora una variable aleatoria  $X$  que, cuando intentemos conectarnos una vez a Internet, vale 1 si podemos hacerlo y 0 si la conexión falla y no podemos. Esta variable tiene la función de masa de probabilidad siguiente:

$$\begin{aligned}P(X = 0) &= 0,1 \\P(X = 1) &= 0,9\end{aligned}$$

La ley de esta variable se conoce como una distribución de Bernoulli de parámetro 0,9.

#### Bernoulli

El nombre de Bernoulli proviene de un importante matemático del siglo XVII.

Una variable aleatoria  $X$  sigue a una **distribución de Bernoulli** de parámetro  $p$ ,  $p \in (0,1)$ . Escribiremos  $B(p)$  si sólo toma los valores 0 y 1 con probabilidades:

$$P(X = 1) = p, \quad P(X = 0) = 1 - p$$

#### Como podéis ver...

... la distribución de Bernoulli es la distribución discreta más sencilla, ya que sólo puede tomar dos valores: cero y uno.

Podemos calcular fácilmente la esperanza y la varianza de una variable  $X$  con una distribución  $B(p)$ :

$$\begin{aligned}E(X) &= 0P(X = 0) + 1P(X = 1) = p \\Var(X) &= (0 - p)^2P(X = 0) + (1 - p)^2P(X = 1) = p(1 - p)\end{aligned}$$

Si  $X$  es una variable aleatoria con ley  $B(p)$ , entonces  $E(X) = p$  y  $Var(X) = p(1 - p)$ .

En general, este tipo de distribución se puede utilizar para modelar cualquier “experimento” en el que tengamos una probabilidad  $p$  de tener éxito y una probabilidad  $1 - p$  de tener un fracaso. Se trata sólo de codificar el éxito como un “1” y el fracaso, como un “0”.

Por ejemplo, cuando compramos un billete de lotería, existe cierta probabilidad  $p$  de que nos toque el premio (y eso sería un “éxito”) y una probabilidad  $1 - p$  de que no nos toque nada (sería el fracaso). ¡Ya sabréis por experiencia que, en este caso, la probabilidad  $p$  de éxito es bastante reducida!

#### Concepto de “éxito”

El concepto de “éxito” es relativo. Por ejemplo, en un proceso de control de calidad, podemos considerar un “éxito” la producción de una pieza defectuosa.

## 2. Distribución binomial

Volvamos a nuestro ejemplo de la conexión a Internet. En general no intentamos conectarnos sólo una vez. Así, imaginemos que, durante todo un día, nos intentamos conectar  $n$  veces y llamamos  $Y$  a la variable aleatoria que cuenta el número de veces que nos hemos podido conectar. Los valores que puede tomar esta nueva variable aleatoria son los números naturales entre cero (si todas las veces que hemos intentado conectarnos la conexión ha fallado) y  $n$  (si nos hemos podido conectar todas las veces que lo hemos intentado).

Veamos cómo podemos calcular la función de masa de probabilidad asociada a la variable  $Y$ :

- La probabilidad,  $P(Y = 0)$ , que no nos hayamos podido conectar ninguna vez, se calcula fácilmente. El acontecimiento  $\{Y = 0\}$  es en realidad:

$$\overline{C}\overline{C}\dots\overline{C}$$

Si decimos que cada llamada es independiente de las otras, tenemos que calcular el producto de las probabilidades de que cada llamada haya fallado:

$$P(Y = 0) = (0,1) \dots (0,1) = (0,1)^n$$

- La probabilidad  $P(Y = 1)$  de que la conexión haya funcionado bien exactamente una vez vendrá dada por la fórmula siguiente:

$$P(Y = 1) = n(0,9)(0,1)^{n-1}$$

Veámoslo: si suponemos que la llamada en la que nos hemos podido conectar ha sido la primera, el acontecimiento de que haya exactamente una única conexión y que ésta se haya producido en la primera llamada es el suceso:

$$C\overline{C}\overline{C}\dots\overline{C}$$

#### Condiciones del experimento

Durante buena parte de esta sesión supondremos que estamos repitiendo un experimento bajo las condiciones siguientes: el resultado de cada prueba puede ser éxito o fracaso, la probabilidad de éxito (bien, de lo que entendemos por éxito) siempre es  $p$  y cada repetición de la prueba es independiente del resto.

Siguiendo un razonamiento análogo al utilizado para calcular  $P(Y = 0)$ , la probabilidad de este suceso es igual a:

$$(0,9)(0,1) \dots (0,1) = (0,9) * (0,1)^{n-1}$$

Pero es posible que la conexión buena se haya producido en el segundo intento, en el tercero o en el  $n$ -ésimo. Esto hace que para calcular  $P(Y = 1)$ , tengamos que multiplicar por  $n$  la cantidad anterior.

- Ya podemos calcular la expresión general de  $P(Y = k)$ , donde la  $k$  será un número natural entre cero y  $n$ . El razonamiento que utilizaremos será parecido al que hemos utilizado para calcular la  $P(Y = 1)$ , aunque deberemos utilizar también algunas de las nociones de combinatoria que vimos en el módulo dedicado al cálculo de probabilidades.

Si suponemos que hemos logrado las  $k$  conexiones en los primeros  $k$  intentos, la probabilidad de este resultado particular es:

$$(0,9)^k (0,1)^{n-k}$$

Sin embargo, las  $k$  conexiones se han podido obtener de muchas formas diferentes entre los  $n$  intentos que hemos hecho. Tal como vimos en el módulo correspondiente, el número de maneras que podemos elegir  $k$  elementos diferentes dentro de un grupo de  $n$  es:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Así, tenemos que:

$$P(Y = k) = \binom{n}{k} (0,9)^k (0,1)^{n-k}$$

donde  $k = 0, \dots, n$ .

La ley de la variable  $Y$  se conoce como la distribución **binomial** de parámetros  $n$  y  $0,9$ .

#### Observación

Evidentemente, si  $k < 0$  o  $k > n$ , entonces  $P(Y = k) = 0$ .

Una variable aleatoria  $Y$  sigue una distribución **binomial** de parámetros  $n$  y  $p$ ,  $n$  natural y  $p \in (0,1)$  –escribiremos  $B(n, p)$ –, si toma valores sobre los naturales entre  $0$  y  $n$  con función de masa de probabilidad:

$$P(Y = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

donde  $k = 0, \dots, n$ .

En el ejemplo hemos utilizado  $p = 0,9$ .

Recordad que la distribución  $B(p)$  se puede interpretar como la realización de un experimento con probabilidad  $p$  de tener éxito y una probabilidad  $1 - p$  de no tenerlo. Ahora lo que hacemos es repetir el experimento  $n$  veces, de manera independiente, y contar el número de éxitos. Es decir, obtenemos la distribución  $B(n, p)$  como suma de  $n$  Bernoullis independientes de parámetro  $p$ .

Si  $X_1, \dots, X_n$  son variables aleatorias independientes con distribución  $B(p)$ , entonces  $X_1 + \dots + X_n$  sigue una distribución  $B(n, p)$ .

#### Obsevación

En particular la distribución de Bernoulli  $B(p)$  es una binomial  $B(1, p)$ .

El cálculo de la esperanza y la varianza de la binomial se obtiene de la esperanza y la varianza de una distribución de Bernoulli:

$$E(X_1 + \dots + X_n) = E(X_1) + \dots + E(X_n) = np$$

$$Var(X_1 + \dots + X_n) = Var(X_1) + \dots + Var(X_n) = np(1 - p)$$

Si  $Y$  es una variable aleatoria con ley  $B(n, p)$ , entonces  $E(Y) = np$  y  $Var(Y) = np(1 - p)$ .

#### Valor de $p$

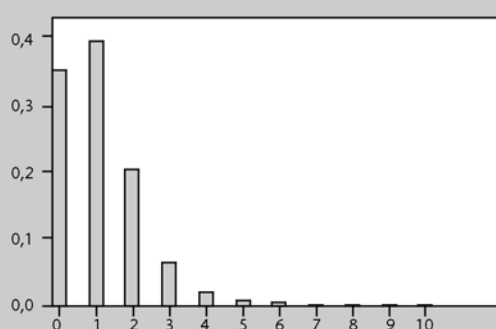
Observad que si  $p$  es pequeña, la probabilidad de lograr pocos éxitos es grande.

Como se ve por su esperanza y su varianza, la forma de una distribución binomial depende mucho de los valores de los parámetros. Podéis verlo también en los gráficos de las funciones de masa de probabilidad de una  $B(10, 0,1)$  y de una  $B(20, 0,5)$ .

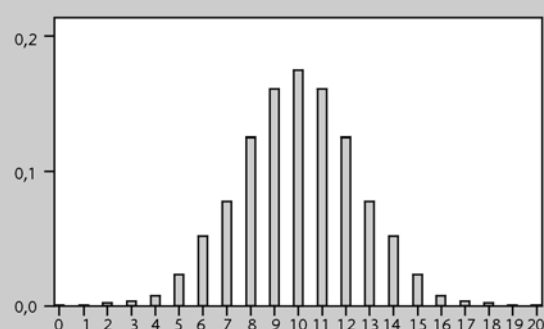
#### Simetría

Observad que la función de masa de probabilidad es simétrica cuando  $p = 0,5$ .

Función de masa de probabilidad binomial de parámetros 10 y 0,1



Función de masa de probabilidad binomial de parámetros 20 y 0,5



### 3. Distribución geométrica

Volvamos a la situación en la que estábamos: continuemos intentando conectarnos a Internet. Lo más común es que lo intentemos hasta que consigamos conectarnos, de manera que podemos contar el número de veces que hay que

intentar la conexión hasta conseguirla. Es decir, ahora miramos el problema desde otra óptica. En la situación anterior intentábamos  $n$  conexiones, en las que  $n$  era un valor fijado desde el principio, y queríamos determinar en cuántos de estos intentos la conexión se había podido establecer. Ahora lo hacemos al revés, intentaremos conectarnos hasta que consigamos la primera conexión y contaremos cuántas veces hemos tenido que intentarlo para conseguirla.

Denominamos  $Z$  a la variable aleatoria que cuenta el número de veces que hay que intentar la conexión hasta que consigamos conectarnos por primera vez.

Podemos calcular ahora la función de masa de probabilidad asociada a la variable  $Z$ :

- La probabilidad  $P(Z = 1)$  de que nos podamos conectar al primer intento es claramente la probabilidad de  $C$ , que es:

$$P(Z = 1) = 0,9$$

- La probabilidad  $P(Z = 2)$  de que nos podamos conectar por primera vez al segundo intento es la probabilidad de que en la primera llamada la conexión haya fallado y que en la segunda la hayamos podido establecer, es decir, la probabilidad de  $\bar{C}C$ . Por tanto:

$$P(Z = 2) = (0,1)(0,9)$$

- Y la expresión general para  $P(Z = k)$ , donde  $k$  será un número natural mayor que 1, vendrá dada por:

$$P(Z = k) = (0,1)^{k-1}(0,9)$$

Esta expresión sale fácilmente del hecho de que el acontecimiento  $\{Z = k\}$  significa que en los  $k - 1$  primeros intentos la conexión ha fallado y que en  $k$ -ésimo hemos conseguido conectarnos.

La ley de la variable  $Z$  se conoce como la distribución geométrica de parámetro 0,9.

#### Observación

Los valores que puede tomar esta nueva variable aleatoria son los números naturales desde 1 (si nos conectamos al primer intento) en adelante (sin límite superior).

Una variable aleatoria  $Z$  sigue una distribución **geométrica** de parámetro  $p$ ,  $p \in [0,1]$  –escribiremos  $geom(p)$ – si toma valores sobre todos los naturales estrictamente positivos con función de masa de probabilidad:

$$P(Z = k) = (1 - p)^{k-1}p$$

donde  $k \geq 1$  natural.

Si volvemos a la interpretación de la variable a partir de la realización de un experimento con probabilidad  $p$  de tener éxito y una probabilidad  $1 - p$  de te-

ner un fracaso, ahora se trata de contar cuántas veces hemos repetido el experimento hasta el primer éxito.

Para esta distribución el cálculo de la esperanza y la varianza es un poco más complicado, así que sólo daremos el resultado.

Si  $Z$  es una variable aleatoria con ley  $geom(p)$ , entonces  $E(Z) = \frac{1}{p}$  y  $Var(Z) = \frac{1-p}{p^2}$ .

#### Disparidad de experimentos

Recordad que el concepto de experimento es muy amplio: puede incluir lanzar una moneda al aire y mirar si sale cara, poner en marcha un coche y mirar si se pone en funcionamiento, etc.

Volviendo a la situación de la conexión a Internet, podemos preguntarnos la probabilidad de que consigamos la conexión con cinco intentos como máximo. Tenemos que calcular:

$$P(Z \leq 5) = \sum_{i=1}^5 P(Z = i) = \sum_{i=1}^5 (0,1)^{i-1} 0,9$$

que, utilizando la fórmula para sumar series geométricas, es:

$$1 - (0,1)^5 = 0,99999$$

Se trata de una probabilidad bastante elevada. Pero lo que nos interesa destacar es que la respuesta a esta pregunta se obtiene fácilmente si conocemos la función de distribución de la variable  $Z$ . ¡Y es fácil calcularla! Si  $k$  es un número natural, partiendo de que sabemos sumar una serie geométrica, tenemos:

$$F(k) = P(Z \leq k) = \sum_{i=1}^k P(Z = i) = \sum_{i=1}^k (0,1)^{i-1} 0,9 = 1 - (0,1)^k$$

Si  $Z$  es una variable aleatoria con ley  $geom(p)$ , entonces su función de distribución es  $F(k) = 1 - (1-p)^k$ , para los  $k > 1$  naturales.

#### Ejemplo del control de calidad

Una de las situaciones más habituales en las que podemos utilizar estas leyes es un proceso de control de calidad. Imaginemos que tenemos una máquina que fabrica al día 100 unidades de una determinada pieza A y sabemos que la probabilidad de que una pieza sea defectuosa es de 0,05. Supongamos, además, que lo que pasa en una pieza es independiente de las otras piezas.

Si llamamos  $X$  al número de piezas defectuosas fabricadas durante un día, está claro que sigue una ley binomial de parámetros 100 y 0,05. Y podemos responder a preguntas del tipo siguiente:

a) Probabilidad de que fabriquemos exactamente diez piezas defectuosas:

$$P(X = 10) = \binom{100}{10} (0,05)^{10} (0,95)^{90}$$



b) Probabilidad de que fabriquemos más de una pieza defectuosa:

$$\begin{aligned} P(X > 1) &= 1 - P(X \leq 1) = 1 - [P(X = 0) + P(X = 1)] = \\ &= 1 - \binom{100}{0} (0,05)^0 (0,95)^{100} - \binom{100}{1} (0,05)^1 (0,95)^{99} \\ &= 1 - (0,95)^{100} - 100(0,05)(0,95)^{99} \approx 0,9629 \end{aligned}$$

c) Número esperado de piezas defectuosas fabricadas durante un día, que corresponderá a la esperanza:

$$E(X) = 100(0,05) = 5$$

Si llamamos  $Z$  a la variable que nos da el número de la primera pieza defectuosa, está claro que seguirá una distribución geométrica de parámetro 0,05. También podemos responder preguntas del tipo:

a) Probabilidad de que antes de diez piezas salga alguna defectuosa:

$$P(Z < 10) = P(Z \leq 9) = F(9) = 1 - (0,95)^9 \approx 0,3698$$

b) Cuándo esperamos que salga la primera pieza defectuosa. En este caso corresponde a la esperanza, que será:

$$E(Z) = \frac{1}{0,05} = 20$$

#### 4. Distribución de Poisson

La última distribución discreta que estudiaremos es la llamada *distribución de Poisson*. Este tipo de distribución se utiliza para modelar varias clases de situaciones, como el número de coches que pasan por un peaje durante una hora, el número de clientes que recibe una peluquería durante un día, el número de llamadas que pasan por una centralita durante una hora o el número de terremotos que hay en un año, etc.

Una variable aleatoria  $X$  sigue una distribución de Poisson de parámetro  $\lambda$ ,  $\lambda > 0$  –escribiremos  $\text{pois}(\lambda)$ – si toma valores sobre todos los naturales con función de masa de probabilidad:

$$P(X = k) = \frac{\lambda^k}{k!} \exp(-\lambda)$$

donde  $k \geq 0$ .

La distribución de Poisson tiene la propiedad de que tanto la esperanza, como la varianza coinciden con el valor del parámetro.

Si  $X$  es una variable aleatoria con ley  $\text{pois}(\lambda)$ , entonces  $E(X) = \lambda$  y  $\text{Var}(X) = \lambda$ .

##### Ejemplo del peaje

Sabemos que el número de coches que pasan durante un minuto por un determinado peaje de la autopista sigue una distribución de Poisson de parámetro 5. Llamamos  $X$  a la

##### Primeras utilizaciones de la Poisson

Uno de los primeros ejemplos de la utilización de la distribución de Poisson que recoge la literatura (siglo XIX) trataba del número de soldados prusianos muertos por una cox de caballo.

##### Observación

Fijaos en que para valores de  $k$  grandes el valor de la probabilidad  $P(X = k)$  será pequeño.

Recordad  $\exp(-\lambda) = e^{-\lambda}$

variable aleatoria que cuenta el número de vehículos que pasan por el peaje. Podemos responder preguntas del tipo siguiente:

a) Probabilidad de que pasen exactamente tres coches por el peaje durante un minuto:

$$P(X = 3) = \frac{5^3}{3!} \exp(-5) = 0,1404$$

b) Probabilidad de que pase algún coche:

$$P(X > 0) = 1 - P(X = 0) = 1 - \exp(-5) = 0,9933$$

c) Número esperado de coches que pasan por el peaje durante un minuto:

$$E(X) = 5$$

La distribución de Poisson se utiliza en ciertas condiciones como una aproximación de la distribución binomial. Se considera que si tenemos una variable aleatoria  $X$  con distribución  $B(n, p)$  con  $n$  grande y  $p$  pequeña, podemos aproximarla por una variable  $Y$  con distribución de Poisson de parámetro  $np$ . Por ejemplo, fijado  $k > 0$ , en las condiciones anteriores, podemos aproximar la probabilidad de que una binomial de parámetros  $n$  y  $p$  tome el valor  $k$ :

$$\binom{n}{k} p^k (1-p)^{n-k}$$

**Condiciones para hacer la aproximación**

Para poder hacer la aproximación, en algunos libros dan la condición  $p < 0,1$  y  $np \leq 5$ .

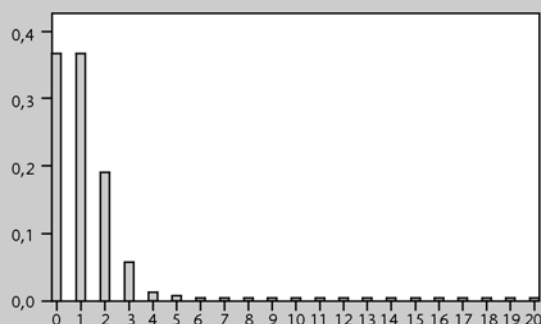
por la probabilidad correspondiente a la ley de Poisson de parámetro  $np$ , es decir:

$$\frac{(np)^k}{k!} \exp(-np)$$

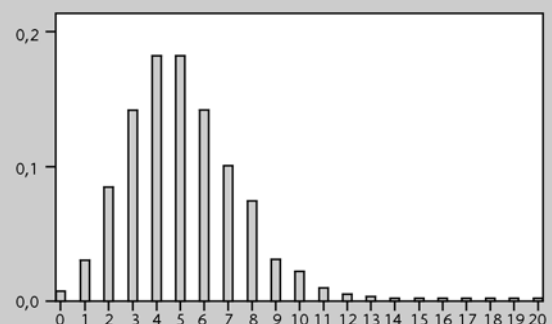
que es más fácil de calcular.

Observad que estamos aproximando una distribución binomial, que sólo puede tomar un número finito de valores, por una distribución de Poisson, que toma valores sobre todos los naturales. Esto es posible porque las probabilidades de la distribución de Poisson para valores grandes son prácticamente nulas. Veamos algunos ejemplos de la función de masa de probabilidades:

**Función de masa de probabilidad de Poisson de parámetro 1**



**Función de masa de probabilidad de Poisson de parámetro 5**



### Ejemplo de los dos dados

Lanzamos dos dados sesenta veces y llamamos  $X$  a la variable aleatoria que cuenta el número de veces que han salido dos seises. Tal como hemos visto,  $X$  sigue una distribución:

$$B\left(60, \frac{1}{6^2}\right)$$

Así, podemos calcular la probabilidad de que salgan exactamente dos seises quince veces, que será:

$$P(X = 15) = \binom{60}{15} \left(\frac{1}{6^2}\right)^{15} \left(1 - \frac{1}{6^2}\right)^{45}$$

Este cálculo no es fácil de hacer. En cambio, si utilizamos la aproximación por la distribución de Poisson, podemos aproximar la probabilidad anterior:

$$\frac{\left(\frac{10}{6}\right)^{15}}{15!} \exp\left(-\frac{10}{6}\right)$$

que es más fácil de calcular.

## 5. Resumen

Hemos presentado las distribuciones discretas más habituales: la de Bernoulli, la binomial, la geométrica y la de Poisson. Hemos aprendido a calcular probabilidades expresadas a partir de estas leyes y hemos visto cuáles son sus esperanzas y sus varianzas.

## Ejercicios

1. En la carretera C-969 hay una única gasolinera. Los automóviles que pasan por delante deciden repostar con una probabilidad de 0,1. Sabemos que entre las 3 y las 4 de la tarde han pasado diez coches por la carretera.

- a) ¿Cuál es la probabilidad de que como mínimo hayan decidido poner gasolina ocho coches?
- b) ¿Cuántos coches esperamos que se hayan detenido a repostar?

Supongamos ahora que el número de coches que pasan entre las 4 y las 5 de la tarde por la carretera C-969 sigue una ley de Poisson de parámetro 10.

- c) ¿Cuál es la probabilidad de que no pase ningún coche?
- d) ¿Cuál es la probabilidad de que al menos pasen tres?
- e) ¿Cuántos automóviles esperamos que pasen?

2. Por problemas de nuestro servidor de correo, sabemos que cada vez que recibimos un correo electrónico hay una probabilidad de 0,05 de que no podamos leerlo correctamente.

- a) ¿Cuál es la probabilidad de que no podamos leer correctamente el tercer mensaje?
- b) ¿Cuál es la probabilidad de que el tercer mensaje sea el primero que no podamos leer correctamente?
- c) ¿Cuál es la probabilidad de que podamos leer correctamente más de tres mensajes antes de que llegue alguno que no se lea bien?
- d) ¿Qué mensaje esperamos que sea el primero que no podamos leer correctamente?

## Solucionario

1. Llamamos  $X$  a la variable aleatoria que nos da el número de coches que se han detenido en la gasolinera entre las 3 y las 4. Según el enunciado, está claro que la variable  $X$  sigue una distribución  $B(10, 0,1)$ .

a) Entonces, dado que el acontecimiento {al menos ocho coches se detengan a repostar} es el suceso  $\{X \geq 8\}$ , debemos calcular:

$$\begin{aligned}
 P(X \geq 8) &= P(X = 8) + P(X = 9) + P(X = 10) = \\
 &= \binom{10}{8} (0,1)^8 (0,9)^2 + \binom{10}{9} (0,1)^9 (0,9)^1 + \binom{10}{10} (0,1)^{10} (0,9)^0 = \\
 &= 45(0,1)^8 (0,9)^2 + 10(0,1)^9 (0,9)^1 + (0,1)^{10} (0,9)^0 \approx \\
 &\approx 0,0000003736
 \end{aligned}$$

Como veis, es muy pequeña.

b) El número de coches que esperamos que se detengan a repostar es la esperanza de  $X$ . Puesto que conocemos la esperanza de una binomial, tenemos:

$$E(X) = 10(0,1) = 1$$

es decir, un coche.

Ahora llamamos  $Y$  a la variable aleatoria que nos da el número de coches que se detienen en la gasolinera entre las 4 y las 5. Está claro que la variable  $Y$  sigue una distribución  $pois(10)$ .

c) El suceso {no pase ningún coche} es el suceso  $\{Y = 0\}$ , de modo que debemos calcular:

$$P(Y = 0) = \frac{10^0}{0!} \exp(-10) = \exp(-10) \approx 0,000045$$

d) De forma análoga, calcularemos:

$$\begin{aligned} P(X \geq 3) &= 1 - P(Y < 3) = 1 - [P(Y = 0) + P(Y = 1) + P(Y = 2)] = \\ &= 1 - \left[ \frac{10^0}{0!} \exp(-10) + \frac{10^1}{1!} \exp(-10) + \frac{10^2}{2!} \exp(-10) \right] = \\ &= 1 - ([1 + 10 + 50] \exp(-10)) \approx 0,9973 \end{aligned}$$

e) El número de coches que esperamos que pasen vendrá dado por la esperanza de  $Y$ . Puesto que conocemos la esperanza de una ley de Poisson, tenemos que:

$$E(Y) = 10$$

es decir, diez coches.

2.

a) El tercer mensaje es como todos los otros y, por tanto, existe una probabilidad de 0,05 de que no podamos leerlo correctamente.

Consideraremos ahora la variable aleatoria  $X$  que nos dará el número del primer mensaje que no podemos leer correctamente. La variable  $X$  seguirá una distribución  $geom(0,05)$ .

b) El suceso {el tercer mensaje es el primero que no podemos leer correctamente} es el acontecimiento  $\{X = 3\}$ . Utilizando la función de masa de probabilidad de una ley geométrica, tenemos:

$$P(X = 3) = (0,95)^2 (0,05) \approx 0,0451$$

c) Ahora debemos calcular:

$$P(X \geq 4) = 1 - P(X < 4) = 1 - P(X \leq 3)$$

Y utilizando la función de distribución para una ley geométrica, tendremos:

$$P(X \geq 4) = 1 - [1 - (0,95)^3] = (0,95)^3 \approx 0,8574$$

d) El número del primer mensaje que esperamos no poder leer correctamente corresponde a la esperanza de  $X$ , que será:

$$E(X) = \frac{1}{0,05} = 20$$

es decir, el vigésimo.

## Variables aleatorias continuas

A continuación estudiaremos las distribuciones continuas. Hasta ahora hemos visto las distribuciones discretas que, recordémoslo, sólo podían tomar un número finito o numerable de valores.

Hablando sin demasiado rigor, podemos decir que las **variables aleatorias continuas** serán aquellas que pueden tomar cualquier valor en un intervalo de los reales. Podemos dar muchos ejemplos de ello:

- a) El peso de una persona adulta: podemos pensar que puede tomar cualquier peso entre 30 kg y 300 kg. Además, si tenemos las herramientas de medida necesarias, podemos dar el peso con tanta precisión como haga falta.
- b) La velocidad a la que va un coche determinado: en este caso podrá coger cualquier velocidad entre 0 y 400 km/h.

### Terminología

Nosotros hablamos de variables aleatorias continuas, pero en muchos libros se denominan *variables aleatorias absolutamente continuas*.

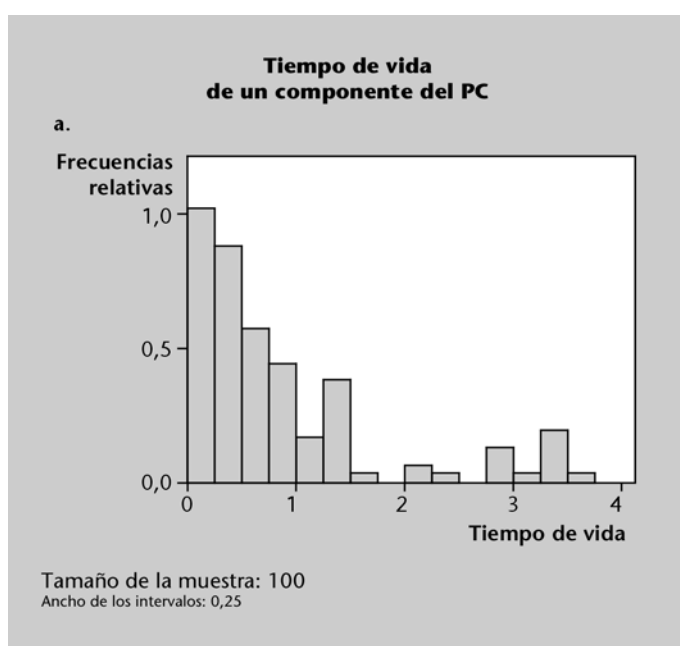
### 1. Función de densidad

La herramienta básica para el estudio de este tipo de distribuciones es la llamada *función de densidad*. Para entenderla bien, primero la relacionaremos con los histogramas de densidades.

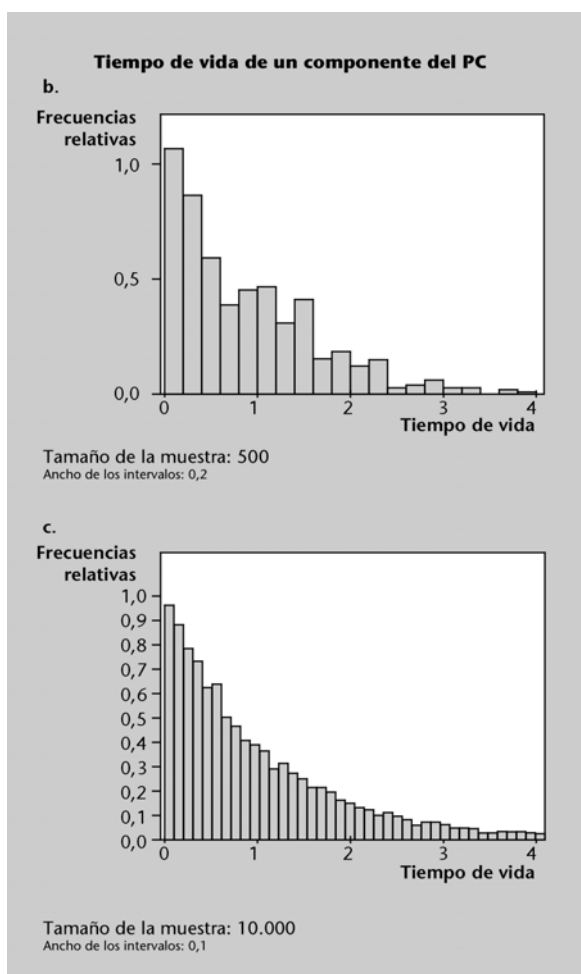
Supongamos que queremos estudiar el tiempo de vida de un componente determinado de los PC. Observemos primero los tiempos de vida de una muestra de cien discos duros, después de quinientos y finalmente, de diez mil. Representamos los datos en los histogramas de densidades siguientes.

### Tiempo de vida

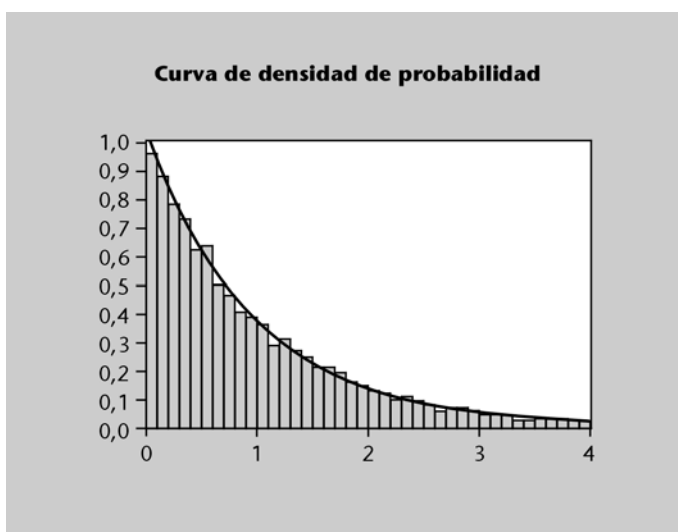
Cuando hablamos del tiempo de vida de un disco duro, nos referimos al tiempo que durará hasta que se estropee.



Puesto que cada vez tomamos muestras de mayores dimensiones, podemos aumentar el número de clases y coger intervalos de menor longitud, de manera que obtendremos los histogramas de densidad de probabilidad siguientes:



Si cogemos cada vez muestras más grandes y hacemos un histograma de densidades con un número más elevado de clases, el perfil del histograma de densidades se acerca cada vez más a la siguiente curva, que llamaremos **curva de densidad de probabilidad**:





A partir de las propiedades del histograma es fácil comprobar que esta curva también satisface las condiciones siguientes:

- 1) El área total bajo la curva de densidad es 1.
- 2) La probabilidad de que el tiempo de vida de un componente esté entre dos valores  $a$  y  $b$  es el área bajo la curva de densidad entre estos puntos.
- 3) La función de densidad es siempre positiva.

Podemos definir así la función de densidad.

La **función de densidad** de una variable aleatoria  $X$  continua es una función  $f(X)$  positiva tal que para todo  $a \leq b$ :

$$P(a < X < b) = \int_a^b f(X) dx$$

#### Signo de la función de densidad

Observad que la función de densidad siempre debe ser positiva.

De la definición es fácil deducir que, dada  $f$  una función de densidad,  $f$  verifica la propiedad siguiente:

$$\int_{-\infty}^{+\infty} f(X) dx = 1$$

#### Recordad que...

... el área bajo una curva se podía calcular utilizando integrales.

### Ejemplos de funciones de densidad

Aquí tenéis algunos ejemplos de funciones de densidad. Es fácil comprobar que cumplen las condiciones:

$$f_1(X) = \begin{cases} 0, & x \leq 0 \\ \exp(-x), & x > 0 \end{cases}$$

$$f_2(X) = \begin{cases} x - \frac{9}{2}, & 5 < x \leq 6 \\ 0, & \text{de otro modo} \end{cases}$$

$$f_3(X) = \begin{cases} 2, & -\frac{3}{2} \leq x \leq -1 \\ 0, & \text{de otro modo} \end{cases}$$

En todos los casos es fácil comprobar que se trata de funciones positivas y que su integral vale 1.

Ahora, si nos preguntan cuál es la probabilidad de que nuestro disco duro haya tenido un tiempo de vida de exactamente tres años (ni un segundo más, ni un segundo menos), la respuesta deberá ser cero. Puede parecer una paradoja. Lo que nos está diciendo es que, dado que la medida teórica del tiempo puede ser tan precisa como queramos, la posibilidad de que un componente dure exactamente tres años es prácticamente despreciable. Podríamos pensar la probabilidad de que el tiempo de vida estuviese entre tres años menos un segundo y tres años más un segundo. La probabilidad sería muy pequeña. Y podemos

hacerla todavía más pequeña. Pensad en una centésima de segundo o una millonésima, etc.

En una variable aleatoria continua  $X$ , la probabilidad de que  $X = x$  es siempre cero, para cualquier  $x$ . Por este motivo, en este tipo de variables sólo tiene sentido hablar de la probabilidad de que la  $X$  pertenezca a cierto intervalo.

Si  $X$  es una variable aleatoria continua, entonces:  
 $P(X = x) = 0$ , para todo  $x$ .

A partir de la función de densidad podemos recuperar, para las variables continuas, algunos conceptos que ya habíamos visto para las variables discretas.

## 2. Relación entre las funciones de distribución y de densidad.

### Cálculo de probabilidades

La definición de función de distribución es la misma que la que utilizamos para las variables discretas. Sin embargo, fijaos en que utilizando la función de densidad podemos calcular fácilmente sus valores.

Dado  $x$  real, podemos calcular la función de distribución en este punto como:

$$F(x) = \int_{-\infty}^x f(y) dy$$

#### Recordatorio

La función de distribución se define como:

$$F(X) = P(X \leq x)$$

y arranca siempre desde cero y llega hasta uno. Además, es siempre una función creciente.

Las propiedades de la función de distribución son las mismas que ya habíamos comentado cuando estudiamos las variables discretas.

Ya hemos visto cómo podemos obtener la función de distribución a partir de la función de densidad. También podemos hacer el proceso inverso, ya que si calculamos la función de distribución integrando la función de densidad, podemos obtener la función de densidad derivando la función de distribución.

Si  $F$  es la función de distribución de una variable aleatoria y  $F$  es una función derivable, entonces  $F$  es la función de distribución de una variable continua con densidad  $F'$ .

#### Por las variables discretas...

... no se puede derivar la función de distribución. Este tipo de variables no tiene asociada una función de densidad.

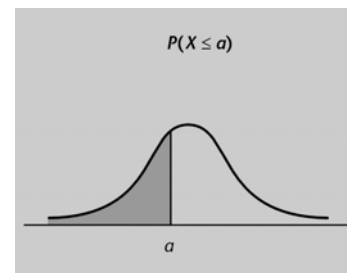
Evidentemente, tanto la función de distribución como la función de densidad caracterizan la ley de una variable aleatoria continua, ya que permiten calcular cualquier probabilidad que se pueda describir con la variable aleatoria  $X$ .

Acabamos este apartado viendo cómo podemos calcular varias probabilidades a partir de la función de densidad. También veremos cómo hacerlo con la función de distribución utilizando la relación que conocemos entre las dos funciones.

1) Cálculo  $P(X \leq a)$ .

Observad que la probabilidad corresponde al área más oscura. En términos de la función de densidad o de la función de distribución, tendremos:

$$P(X \leq a) = F(a) = \int_{-\infty}^a f(x) dx$$



Además, puesto que la probabilidad de que la variable tome un valor concreto es cero, tenemos:

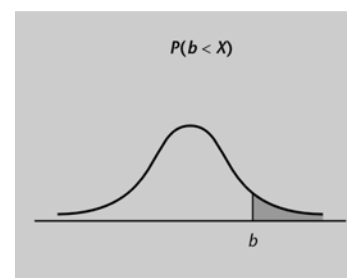
$$P(X \leq a) = P(X < a)$$

A veces no es fácil o no es posible calcular una integral.

2) Cálculo  $P(b < X)$ .

Ahora, en términos de la función de densidad:

$$P(b < X) = \int_b^{\infty} f(x) dx$$



Además:

$$P(b < X) = 1 - P(X \leq b) = 1 - \int_{-\infty}^b f(x) dx = 1 - F(b)$$

Y como antes:

$$P(b < X) = P(b \leq X)$$

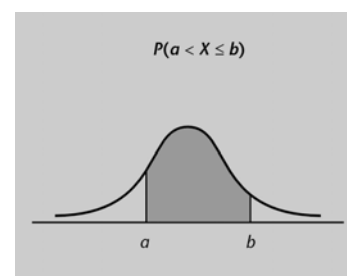
3) Cálculo de  $P(a < X \leq b)$ .

Fijaos en que, utilizando:

$$\{a < X \leq b\} \cup \{X \leq a\} = \{X \leq b\}$$

tenemos:

$$P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a) = \int_a^b f(x) dx$$



Además, tenemos:

$$P(a < X \leq b) = P(a < X < b) = P(a \leq X \leq b) = P(a \leq X < b)$$

### Ejemplo del tiempo de procesamiento

El tiempo que tarda (en minutos) nuestro ordenador en procesar una determinada cantidad de datos sigue una variable aleatoria con función de densidad:

$$f(x) = \begin{cases} \exp(-(x-2)), & x > 2 \\ 0 & x \leq 2 \end{cases}$$

Llamamos  $X$  a esta variable aleatoria. Entonces, utilizando la función de densidad, podemos calcular la probabilidad de toda una serie de acontecimientos:

- a) Probabilidad de que tarde como mucho cinco minutos.
- b) Probabilidad de que tarde más de ocho minutos y menos de diez.
- c) Probabilidad de que tarde como mínimo ocho minutos.
- d) Probabilidad de que tarde más de cuatro minutos y menos de diez si sabemos que como mínimo tarda ocho.

$$\begin{aligned} \text{a) } P(X \leq 5) &= \int_2^5 \exp(-(x-2)) dx = \exp(2) \int_2^5 \exp(-x) dx = \\ &= \exp(2) [-\exp(-x)]_2^5 = \exp(2) (-\exp(-5) + \exp(-2)) = 1 - \exp(-3) \end{aligned}$$

$$\begin{aligned} \text{b) } P(8 < X < 10) &= \int_8^{10} \exp(-(x-2)) dx = \exp(2) \int_8^{10} \exp(-x) dx = \\ &= \exp(2) [-\exp(-x)]_8^{10} = \exp(2) (-\exp(-10) + \exp(-8)) = \exp(-6) - \exp(-8) \end{aligned}$$

$$\begin{aligned} \text{c) } P(X \geq 8) &= \int_8^{\infty} \exp(-(x-2)) dx = \exp(2) \int_8^{\infty} \exp(-x) dx = \\ &= \exp(2) [-\exp(-x)]_8^{\infty} = \exp(2) (\exp(-8)) = \exp(-6) \end{aligned}$$

$$\begin{aligned} \text{d) } P(4 < X < 10 | X \geq 8) &= \frac{P(8 \leq X < 10)}{P(X \geq 8)} = \frac{\exp(-6) - \exp(-8)}{\exp(-6)} = \\ &= 1 - \exp(-2) \end{aligned}$$

Fijaos en que hay algunos acontecimientos que tienen claramente probabilidad cero, como: probabilidad de que tarde como mucho menos cuatro minutos (fijaos en que no tiene sentido) o la probabilidad de que tarde exactamente diez minutos.

También podemos hacer los cálculos anteriores utilizando la función de distribución. Calculémosla:

$$\begin{aligned} F(x) &= \int_{-\infty}^x f(y) dy = \int_2^x \exp(-(y-2)) dy = \int_2^x e^{-(y-2)} dy = \int_2^x e^{-y} e^2 dy = e^2 \int_2^x e^{-y} dy = -e^2 [e^{-y}]_2^x \\ &= -e^2 [e^{-y}]_2^x = -e^2 (e^{-x} - e^{-2}) = e^2 (-e^{-x} + e^{-2}) = -e^{-x} e^2 + e^2 e^{-2} = -e^{2-x} + 1 \\ &= 1 - e^{-(x-2)} \end{aligned}$$

si  $x \geq 2$  y  $F(x) = 0$  si  $x < 2$ .

Observad que si derivamos la función de distribución, recuperamos la función de densidad.

Ahora damos las respuestas utilizando la función de distribución:

- a)  $P(X \leq 4) = F(4) = 0$
- b)  $P(X \leq 5) = F(5) = 1 - \exp(-3)$
- c)  $P(8 < X < 10) = P(X \leq 10) - P(X \leq 8) = F(10) - F(8) = \exp(-6) - \exp(-8)$
- d)  $P(X \geq 8) = 1 - P(X < 8) = 1 - P(X \leq 8) = \exp(-6)$

### 3. Independencia

El concepto de independencia se extiende también a las variables continuas. Aunque la idea es la misma que para las variables discretas, debemos dar otra

caracterización, ya que ahora la condición  $P(X = x, Y = y) = P(X = x)P(Y = y)$  no tiene sentido.

Las variables aleatorias continuas  $X$  e  $Y$  son **independientes** si:

$$P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$$

para todo posible par de valores  $x$  e  $y$ .

Esta caracterización es equivalente a la que vimos para las variables aleatorias discretas. En todo caso, el estudio detallado de la independencia para este tipo de variable supera los objetivos de este curso.

#### 4. Esperanza y varianza

La esperanza y la varianza se pueden calcular también utilizando la función de densidad.

Recordemos que para calcular la función de distribución para una variable aleatoria discreta  $X$ , teníamos que calcular el sumatorio:

$$F(x) = \sum_{y \leq x} P(X = y)$$

mientras que en el caso continuo debemos calcular una integral de la función de densidad del tipo:

$$F(x) = \int_{-\infty}^x f(y) dy$$

Así, siempre que pasemos a trabajar con variables aleatorias absolutamente continuas, lo que tenemos que hacer es convertir los sumatorios en integrales.

Para las variables discretas, podíamos calcular la esperanza y la varianza:

$$E(X) = \sum_x xP(X = x)$$

$$Var(X) = \sum_x (x - E(X))^2 P(X = x)$$

Las definiciones para el caso continuo son las siguientes:

La **esperanza** de una variable aleatoria continua  $X$ , que denotaremos por  $E(X)$ , se define como:

$$E(X) = \int_{-\infty}^{\infty} xf(x) dx$$

**Para pasar de variables discretas a continuas...**

... deberemos convertir los sumatorios en integrales.

La **varianza** de una variable aleatoria continua  $X$ , que denotaremos por  $Var(X)$ , se define como:

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

Las propiedades de la esperanza y la varianza y la desigualdad de Tchebichev que vimos para variables discretas en la sesión “Esperanza y varianza” (apartados 2, 3 y 4) continúan siendo ciertas.

#### Recordatorio

La varianza se podía calcular como:

$$E(X^2) - (E(X))^2$$

#### Ejemplo del tiempo de procesamiento

Volvamos al ejemplo del tiempo de procesamiento. Queremos calcular ahora el tiempo que esperamos que tarde el proceso. Por tanto, tendremos que calcular la esperanza de la variable  $X$ . Si utilizamos la fórmula y hacemos integración por partes, tendremos:

$$\begin{aligned} E(X) &= \int_2^{\infty} x \exp(-(x-2)) dx = \exp(2) \int_2^{\infty} x \exp(-x) dx = \\ &= \exp(2) \left\{ [-x \exp(-x)]_2^{\infty} + \int_2^{\infty} \exp(-x) dx \right\} = \\ &= \exp(2) \{ 2\exp(-2) + \exp(-2) \} = 3 \end{aligned}$$

Por tanto, esperamos que tarde tres minutos.

Consideremos ahora una variable con función de densidad:

$$f(x) = \begin{cases} \frac{x}{2}, & 0 < x < 2 \\ 0 & \text{de otro modo} \end{cases}$$

Queremos calcular la esperanza y la varianza de esta variable que llamaremos  $X$ :

$$E(X) = \int_0^2 x \frac{x}{2} dx = \frac{1}{2} \left[ \frac{x^3}{3} \right]_0^2 = \frac{4}{3}$$

y puesto que:

$$E(X^2) = \int_0^2 x^2 \frac{x}{2} dx = \frac{1}{2} \left[ \frac{x^4}{4} \right]_0^2 = 2$$

tenemos:

$$Var(X) = E(X^2) - (E(X))^2 = 2 - \frac{16}{9} = \frac{2}{9}$$

## 5. Resumen

En esta sesión hemos aprendido el concepto de variable aleatoria continua. Hemos estudiado la función de densidad y hemos extendido a esta clase de variables los conceptos de función de distribución, independencia, esperanza y varianza que habíamos estudiado para las variables discretas.

## Ejercicios

1.

Sea  $X$  una variable aleatoria continua con función de densidad  $f(x) = \frac{1}{9}x^2$ , si  $0 \leq x \leq 3$ .

- Comprobad que es una función de densidad.
- Calculad la función de distribución.
- Calculad:  $P(X > 2)$  y  $P(X = 5)$ .
- Calculad la esperanza y la varianza de  $X$ .

2.

Sea  $X$  una variable aleatoria continua con función de distribución:

$$F(x) = \begin{cases} 0, & x \leq 1 \\ \frac{(x-1)^2}{9}, & 1 < x \leq 4 \\ 1 & 4 < x \end{cases}$$

- Calculad la función de densidad.
- Calculad:  $P(X < 2)$ ,  $P(2 < X \leq 3)$ ,  $P(X < 5)$  y  $P(X > 5)$ .
- Calculad la esperanza y la varianza de  $X$ .

## Solucionario

1.

- Claramente se trata de una función de densidad, ya que es una función positiva, y

$$\int_{-\infty}^{+\infty} f(x) dx = \int_0^3 \frac{1}{9} x^2 dx = \left[ \frac{x^3}{27} \right]_0^3 = 1$$

- Para calcular su función de distribución, para  $x \in (0,3)$  debemos hacer la integral:

$$\int_0^x \frac{1}{9} y^2 dy = \left[ \frac{y^3}{27} \right]_0^x = \frac{x^3}{27}$$

Por tanto, la función de distribución es:

$$F(x) = \begin{cases} 0, & x \leq 0 \\ \frac{x^3}{27}, & 0 < x \leq 3 \\ 1 & 3 < x \end{cases}$$

- La probabilidad  $P(X = 5)$  es cero.

La otra se puede calcular utilizando la función de densidad:

$$P(X > 2) = \int_2^{+\infty} f(x) dx = \int_2^3 \frac{1}{9} x^2 dx = \left[ \frac{x^3}{27} \right]_2^3 = 1 - \frac{8}{27} = \frac{19}{27}$$

o la función de distribución:

$$P(X > 2) = 1 - P(X \leq 2) = 1 - F(2) = 1 - \frac{8}{27} = \frac{19}{27}$$

d) Primero calculamos la esperanza:

$$E(X) = \int_0^3 x \frac{x^2}{9} dx = \frac{1}{9} \left[ \frac{x^4}{4} \right]_0^3 = \frac{9}{4}$$

Y puesto que:

$$E(X^2) = \int_0^3 x^2 \frac{x^2}{9} dx = \frac{1}{9} \left[ \frac{x^5}{5} \right]_0^3 = \frac{27}{5}$$

tenemos:

$$Var(X) = E(X^2) - (E(X))^2 = \frac{27}{5} - \frac{81}{16} = \frac{432 - 405}{80} = \frac{27}{80}$$

2.

a) Para obtener la función de densidad, debemos derivar la función de distribución. Fijaos en que la derivada de la función:

$$\frac{(x-1)^2}{9}$$

es:

$$\frac{2(x-1)}{9}$$

Por tanto, la función de densidad es:

$$x) = \begin{cases} \frac{2(x-1)}{9}, & 1 < x \leq 4 \\ 0 & \text{de otro modo} \end{cases}$$

Como antes, es fácil comprobar que se trata de una función de densidad.



b) La probabilidad  $P(X > 5)$  es claramente cero.

La probabilidad  $P(X < 5)$  es claramente uno.

La probabilidad  $P(X < 2)$  se puede calcular utilizando la función de densidad:

$$P(X < 2) = \int_{-\infty}^2 f(x) dx = \int_1^2 \frac{2(x-1)}{9} dx = \left[ \frac{(x-1)^2}{9} \right]_1^2 = \frac{1}{9} - 0 = \frac{1}{9}$$

o la función de distribución:

$$P(X < 2) = P(X \leq 2) = F(2) = \frac{1}{9}$$

La probabilidad  $P(2 < X \leq 3)$  también se puede calcular utilizando la función de densidad:

$$P(2 < X \leq 3) = \int_2^3 f(x) dx = \int_2^3 \frac{2(x-1)}{9} dx = \left[ \frac{(x-1)^2}{9} \right]_2^3 = \frac{4}{9} - \frac{1}{9} = \frac{1}{3}$$

o la función de distribución:

$$P(2 < X \leq 3) = P(X \leq 3) - P(X \leq 2) = F(3) - F(2) = \frac{4}{9} - \frac{1}{9} = \frac{1}{3}$$

c) Una vez conocida la función de densidad de probabilidad, calculamos la esperanza y la varianza de  $X$  siguiendo el mismo procedimiento que en el ejercicio 1:

$$E(X) = \int_1^4 x \frac{2(x-1)}{9} dx = \left( \frac{2}{27} x^3 - \frac{1}{9} x^2 \right)_1^4 = 3$$

$$\text{Y puesto que: } E(X^2) = \int_1^4 x^2 \frac{2(x-1)}{9} dx = \left( \frac{1}{18} x^4 - \frac{2}{27} x^3 \right)_1^4 = \frac{19}{2}$$

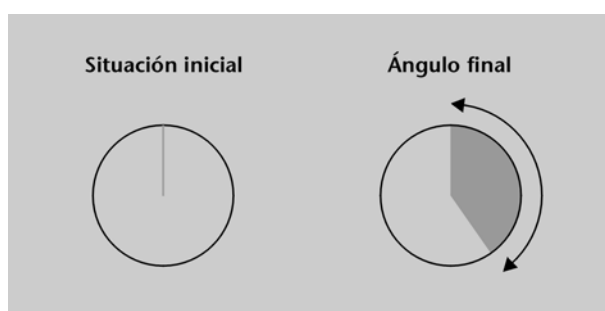
$$\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{19}{2} - 9 = \frac{1}{2}$$

## Algunas leyes continuas. La ley normal

Ahora estudiaremos algunas distribuciones continuas. La ley continua más importante y que aparece constantemente en diferentes contextos es la llamada *ley normal*. En esta sesión la presentaremos y la estudiaremos a fondo. Sin embargo, comencemos por el estudio de otras distribuciones importantes, como la distribución uniforme y la distribución exponencial.

### 1. Distribución uniforme

Comencemos con un ejemplo sencillo. Imaginad que tenemos una especie de ruleta y la hacemos girar alrededor de su eje. Después medimos el ángulo que hemos recorrido.



Suponemos que al hacer girar nuestra “ruleta”, ésta puede detenerse en cualquier punto de la circunferencia y que no hay puntos en los que tenga más tendencia a parar. Si llamamos  $X$  a la variable que nos da el ángulo, tendremos que tomará valores entre 0 y  $2\pi$ . Según nuestras hipótesis, si nos preguntan cuál es la probabilidad de que el ángulo esté entre cero y  $\frac{\pi}{2}$ , tenemos:

$$P\left(0 < X < \frac{\pi}{2}\right) = \frac{1}{4}$$

ya que el ángulo que hemos pedido corresponde a una cuarta parte de los ángulos posibles y hemos supuesto que la probabilidad está distribuida de una manera “uniforme”.

Tenemos así un primer ejemplo de una distribución uniforme, en este caso en el intervalo  $(0, 2\pi)$ .

Sin embargo, es habitual tomar una variable uniforme con valores entre 0 y 1. Puesto que la probabilidad de cualquier número debe ser la misma, la función de densidad deberá ser constante. Imponiendo que la integral de la función de densidad valga 1, obtenemos que la función de densidad ha de ser constante a 1.

**Como ya sabéis,...**

... el ángulo puede tomar un valor entre 0 y  $2\pi$ .

La función de densidad de la distribución **uniforme** (0,1) es:

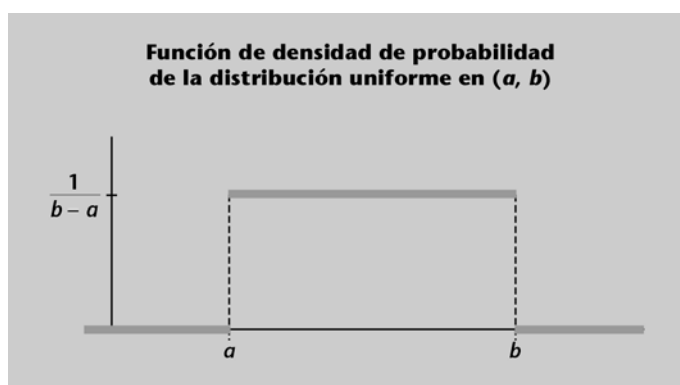
$$f(x) = \begin{cases} 1, & 0 \leq x \leq 1 \\ 0 & \text{de otro modo} \end{cases}$$

En general, podemos definir una distribución uniforme en un intervalo  $(a,b)$ . Siguiendo el mismo razonamiento anterior, podemos obtener ahora su función de densidad.

La función de densidad de la distribución **uniforme**  $(a,b)$  es:

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0 & \text{de otro modo} \end{cases}$$

La función de densidad presenta la forma siguiente.



Como ya hemos visto en la sesión anterior, a partir de la función de densidad podemos calcular las probabilidades que queramos, la función de distribución, la esperanza y la varianza. Veamos estos cálculos para la distribución uniforme.

Consideremos  $X$  una variable aleatoria con distribución uniforme  $(0,1)$ .

Calculemos primero la probabilidad de que  $X$  tome valores entre 0,4 y 0,6 utilizando la función de densidad:

$$P(0,4 \leq X \leq 0,6) = \int_{0,4}^{0,6} 1 dx = 0,6 - 0,4 = 0,2$$

Fijaos en que la probabilidad depende sólo del ancho del intervalo. Así, esta probabilidad será la misma que la probabilidad de obtener un valor entre 0,1 y 0,3.

La función de distribución es:

$$F(x) = \int_0^x 1 dy = x$$

Así, podemos calcular la probabilidad anterior utilizando la función de distribución:

$$P(0,4 \leq X \leq 0,6) = F(0,6) - F(0,4) = 0,6 - 0,4 = 0,2$$

Finalmente, mediante integrales, obtenemos la media y la varianza de la distribución:

$$E(X) = \int_0^1 x dx = \frac{1}{2} x^2 \Big|_0^1 = \frac{1}{2}$$

$$Var(X) = E[(X - 0,5)^2] = \int_0^1 (x - 0,5)^2 dx = \frac{1}{3} (x - 0,5)^3 \Big|_0^1 = \frac{1}{12}$$

Con cálculos similares podemos obtener la función de distribución, la esperanza y la varianza de una uniforme  $(a, b)$ . Por ejemplo, la función de distribución será:

$$F(x) = \int_a^x \frac{1}{b-a} dy = \frac{x-a}{b-a}$$

si  $a \leq x \leq b$ . Podemos calcular también la esperanza:

$$E(x) = \int_a^b x \frac{1}{b-a} dx = \frac{1}{2(b-a)} x^2 \Big|_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{1}{2}(b+a)$$

mientras que el cálculo de la varianza se hace de manera parecida.

Si  $X$  es una variable aleatoria con distribución uniforme, entonces su función de distribución es:

$$F(x) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a < x \leq b \\ 1, & b < x \end{cases}$$

$$\text{Además, } E(X) = \frac{1}{2}(b+a) \text{ y } Var(X) = \frac{1}{12}(b-a)^2$$

## 2. Distribución exponencial

Una distribución continua muy importante es la llamada *distribución exponencial*.

En muchas ocasiones está relacionada con la distribución de Poisson que vimos hace dos sesiones. Recordemos que la distribución de Poisson se utiliza para modelar el número de sucesos que se dan en una unidad de tiempo. Por

ejemplo, considerábamos que podíamos modelar con una Poisson el número de peticiones que recibe un servidor de Internet durante una hora o el número de clientes que entran en una oficina de una caja de ahorros durante un día. En estas circunstancias, la distribución exponencial sirve para determinar la distribución de probabilidad del tiempo entre dos sucesos consecutivos. Por ejemplo, en el caso del servidor de Internet comentado antes, utilizaremos la distribución exponencial para modelar el tiempo entre dos peticiones. Como en el caso de la Poisson, la distribución exponencial dependerá también de un parámetro que llamaremos  $\lambda$ .

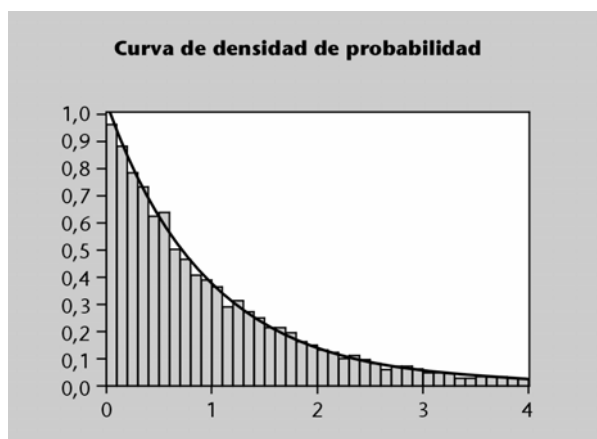
#### Consideremos el peaje de una autopista

Hemos visto que el número de coches que pasan por el peaje durante un minuto se puede representar con una variable aleatoria Poisson. En cambio, utilizamos una exponencial para modelar el tiempo que pasa entre el primer y el segundo coche (y entre el segundo y el tercero, y entre el tercero y el cuarto y etc.).

La función de densidad de la distribución exponencial de parámetro  $\lambda$  es:

$$f(x) = \begin{cases} 0, & x < 0 \\ \lambda e^{-\lambda x}, & 0 \leq x \end{cases}$$

Existe otra situación en la que aparece a menudo la distribución exponencial. En la sesión anterior hemos estudiado los tiempos de vida del disco duro de los PC. A fuerza de ir recogiendo cada vez más datos, llegábamos a ver que el perfil del histograma de frecuencias tenía la forma siguiente:



Para modelizar los tiempos de vida, también se utiliza la distribución exponencial. Fijaos en que, en una distribución exponencial, cuanto más tiempo pasa, menos probable es que se dé el suceso.

Calculemos ahora la función de distribución, la esperanza y la varianza. La función de distribución será ésta:

$$F(x) = \int_0^x \lambda e^{-\lambda y} dy = -e^{-\lambda y} \Big|_0^x = 1 - e^{-\lambda x}$$

si  $0 \leq x$ . Calculemos también la esperanza, en la que tendremos que hacer una integral por partes:

$$E(X) = \int_0^{\infty} x \lambda e^{-\lambda x} dx = -x e^{-\lambda x} \Big|_0^{\infty} + \int_0^{\infty} e^{-\lambda x} dx = -\frac{1}{\lambda} e^{-\lambda x} \Big|_0^{\infty} = \frac{1}{\lambda}$$

#### Valores de la variable

Esta distribución permite que la variable tome cualquier valor positivo. En realidad, en la práctica habrá un valor de manera que la probabilidad de que la variable sea mayor será despreciable.

Si  $X$  es una variable aleatoria con distribución exponencial de parámetro  $\lambda$ , entonces su función de distribución es:

$$F(x) = \begin{cases} 0, & x \leq 0 \\ 1 - e^{-\lambda x}, & 0 < x \end{cases}$$

Además,  $E(X) = \frac{1}{\lambda}$  y  $Var(X) = \frac{1}{\lambda^2}$

Fijaos en que tanto la esperanza como la varianza se pueden dar en términos del parámetro de la exponencial.

### Ejemplo del tiempo de vida

El tiempo de vida de una bombilla (en años) sigue una exponencial de parámetro 3. Llamamos  $X$  a la variable aleatoria que nos da este tiempo de vida. Entonces tenemos que su función de densidad será:

$$f(x) = \begin{cases} 0, & x < 0 \\ 3e^{-3x}, & 0 \leq x \end{cases}$$

Su función de distribución será:

$$F(x) = \begin{cases} 0, & x \leq 0 \\ 1 - e^{-3x}, & 0 < x \end{cases}$$

El tiempo de vida esperado de la bombilla será de  $1/3$  y su varianza, de  $1/9$ .

A partir de aquí también podemos calcular cualquier probabilidad:

a) Probabilidad de que dure menos de un año; si lo resolvemos con la función de densidad:

$$P(X < 1) = \int_0^1 3\exp(-3x)dx = [-\exp(-3x)] = 1 - \exp(-3) = 0,9502$$

b) Probabilidad de que dure más de dos años; si los resolvemos con la función de distribución:

$$P(X > 2) = 1 - F(2) = \exp(-6) = 0,0025$$

#### Notación

Recordad que a veces utilizamos la notación  $e^{-x}$  y a veces,  $\exp(-x)$ .

## 3. Distribución normal

La ley más utilizada en estadística es la llamada *ley normal*. La curva normal aparece en numerosos fenómenos naturales: en la altura de una especie de árboles, en el número de glóbulos rojos en la sangre y en la mayoría de las características biológicas que se pueden determinar. También la distribución de los errores en un proceso de fabricación se acerca a una ley normal, hecho que hace que sea un trabajo fundamental en los procesos de control de calidad.

Una variable aleatoria  $X$  sigue una ley normal con media  $\mu$  y varianza  $\sigma^2$ , que denotaremos por  $N(\mu, \sigma^2)$ , si su función de densidad es:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu)^2\right)$$

para todo  $x$  real.

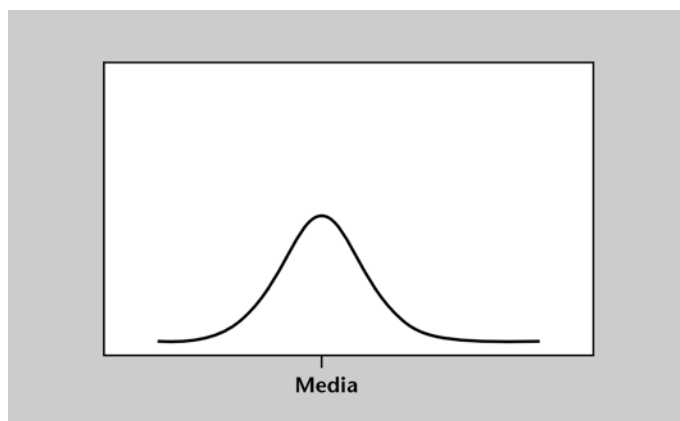
#### El nombre de normal...

... proviene del hecho de que se pensaba que era el patrón natural de las distribuciones.

En un primer vistazo, fijaos en dos cosas importantes:

- La función de densidad es un tanto complicada.
- La ley normal queda determinada por dos parámetros,  $\mu$  (será la esperanza) y  $\sigma^2$  (será la varianza).

Hablaremos primero de la expresión de la función de densidad. Seguramente no os dirá nada, pero si la dibujamos, nos encontramos con la forma siguiente:



#### Campana de Gauss

Esta forma es conocida con el nombre de *la campana de Gauss*. El nombre de Gauss se debe a que él fue uno de los primeros en deducirla cuando estudiaba los errores en observaciones astronómicas. Gauss fue un famoso matemático alemán del siglo XVIII.

En este caso no tenemos una expresión explícita para la función de distribución. Evidentemente, podemos darla como la integral de la densidad, es decir, si  $X$  es una variable aleatoria con ley  $N(\mu, \sigma^2)$ , su función de distribución es:

$$F(x) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(y-\mu)^2\right) dy$$

pero dado que no sabemos calcular esta integral, esta expresión no nos sirve de mucho y tendremos que utilizar las llamadas *tablas de la normal*.

### 3.1. El papel de los parámetros

Como ya hemos comentado, la ley normal queda determinada por los dos parámetros  $\mu$  y  $\sigma^2$ .

Si  $X$  es una variable aleatoria  $N(\mu, \sigma^2)$ , entonces  $E(X) = \mu$  y  $Var(X) = \sigma^2$ .

El cálculo de la esperanza y la varianza requiere resolver unas integrales bastante complicadas.

El valor de la  $\mu$  nos da el punto donde se encuentra el centro de la curva. En cambio, la  $\sigma^2$  afecta a la forma de la curva: si la  $\sigma^2$  es pequeña, la curva será más puntiaguda y, por tanto, se concentrará más masa en un entorno de la  $\mu$  (bien,

eso ya lo sabíamos, ya que precisamente la  $\sigma^2$  nos da la varianza), mientras que si la  $\sigma^2$  es grande, la dispersión será mayor y la curva será más plana.

En algunos libros (o *software*), en lugar de dar la varianza, nos dan la desviación típica y, por tanto, tendremos leyes normales  $N(\mu, \sigma)$ . Así, si nos dicen que tenemos una  $N(0,4)$ , debemos saber si el 4 se refiere a la desviación típica o a la varianza. Nosotros hacemos referencia siempre a la varianza.

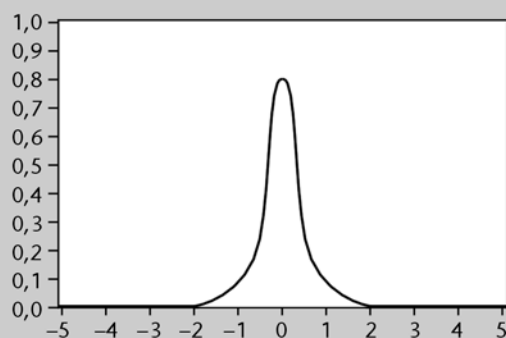
Veamos el papel de la desviación en los gráficos siguientes (dibujamos la densidad de leyes normales  $N(0, 0,25)$   $N(0,1)$  y  $N(0,4)$ , en el intervalo  $[-5,5]$ ).

**Atención**

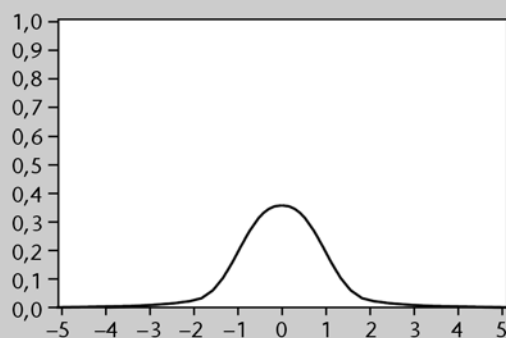
Tened siempre cuidado cuando os den una ley normal. Tenéis que saber si os están dando la varianza o la desviación típica.

**Ejemplos de diferentes desviaciones típicas**

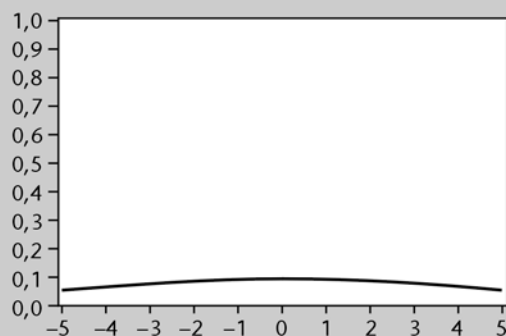
a. Función de densidad de probabilidad  $N(0, 0,25)$



b. Función de densidad de probabilidad  $N(0, 1)$

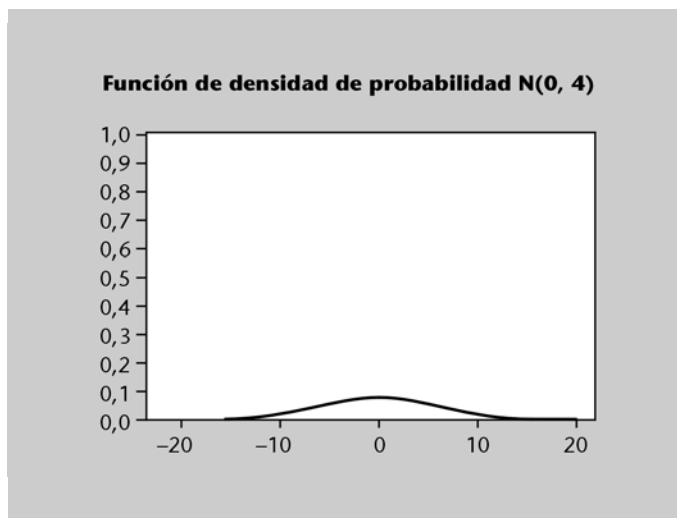


c. Función de densidad de probabilidad  $N(0, 4)$





Como podéis observar, las formas son diferentes. Para que veais bien la última, quizá será mejor verla con otra escala:



En las representaciones gráficas hemos visto dibujado qué sucede cuando varía el  $\sigma$ . Cuando varía el  $\mu$ , lo único que tenemos que hacer es una translación y mover la curva hacia la derecha o hacia la izquierda.

### 3.2. Estandarizar

Una propiedad especial de la ley normal nos permite reducir el estudio de cualquier probabilidad de cualquier variable normal al estudio de una ley  $N(0,1)$ . Llamaremos a esta ley una *distribución normal estándar*.

La distribución **normal estándar** es la distribución normal con esperanza 0 y varianza 1. Su función de densidad es:

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right)$$

Normalmente la denotaremos con la letra  $Z$ .

Suponed que tenéis  $X$ , una variable aleatoria con distribución  $N(\mu, \sigma^2)$ . Si utilizamos la función de densidad, tenemos:

$$P(x_1 \leq X \leq x_2) = \int_{x_1}^{x_2} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(x - \mu)^2\right) dx$$

Sin embargo, si hacemos el cambio de variable  $z = \frac{x-\mu}{\sigma}$ , podemos escribir esta última integral como:

Repasad cómo se hacía un cambio de variable.



$$\int_{\frac{x_1-\mu}{\sigma}}^{\frac{x_2-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right) dz$$

Fijaos en que de esta manera podemos escribir:

$$P\left(\frac{x_1-\mu}{\sigma} \leq Z \leq \frac{x_2-\mu}{\sigma}\right) = \int_{\frac{x_1-\mu}{\sigma}}^{\frac{x_2-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right) dz$$

#### Normal estándar

La ley normal estándar es importante porque podemos reducir cualquier otra ley normal a una de este tipo.

donde la  $Z$  indica la variable  $\frac{x-\mu}{\sigma}$ .

Así, la probabilidad de que la variable  $Z = \frac{x-\mu}{\sigma}$  tome valores en el intervalo  $\left(\frac{x_1-\mu}{\sigma}, \frac{x_2-\mu}{\sigma}\right)$  viene dada por la integral entre estos puntos de la función de densidad de una distribución normal estándar. Por tanto,  $Z$  deberá ser una ley normal estándar.

Si  $X$  es una variable aleatoria con distribución  $N(\mu, \sigma^2)$ , entonces la variable  $Z = \frac{x-\mu}{\sigma}$  seguirá una distribución normal estándar, es decir, será una  $N(0,1)$ .

La consecuencia fundamental de esta propiedad es que los cálculos del área bajo la curva en términos de la desviación son los mismos para todas las densidades normales.

Si  $X$  es una variable aleatoria con distribución  $N(\mu, \sigma^2)$  y  $Z$  una distribución  $N(0,1)$ , entonces para todo par  $k_1, k_2$  de números positivos:

$$P(\mu - k_2\sigma \leq X \leq \mu + k_1\sigma) = P(-k_2 \leq Z \leq k_1)$$

En particular, para todo  $k$ :

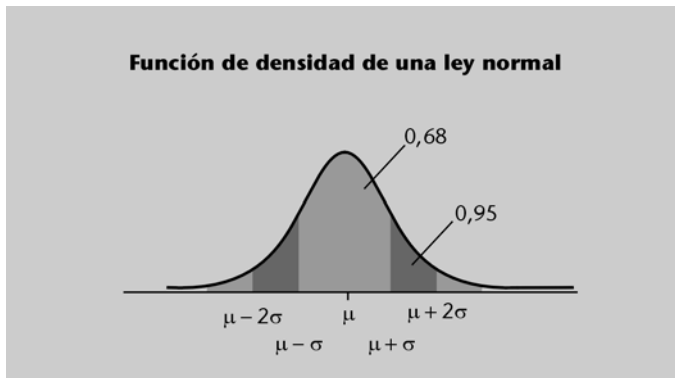
$$P(\mu - k\sigma \leq X \leq \mu + k\sigma) = P(-k \leq Z \leq k)$$

donde  $k$  no debe ser necesariamente entero.

Así, encontramos:

- la probabilidad del intervalo dado por una desviación estándar a un lado y otro de la media es siempre la misma, de 0,68. Podemos decir, por tanto, que en una distribución normal el 68% de los individuos toma valores entre la media menos una desviación estándar y la media más una desviación estándar.

- Si tomamos dos desviaciones estándar, la probabilidad es de 0,95.
- Si consideramos tres, la probabilidad es de 0,997.



- También podemos hacer el proceso inverso: si queremos tener una probabilidad del 0,90, debemos coger el intervalo dado por la media menos 1,645 desviaciones estándar y la media más 1,645 desviaciones estándar, es decir:

$$P(\mu - 1,645\sigma \leq X \leq \mu + 1,645\sigma) = 0,90$$

### Ejemplo de las resistencias

La fábrica XSE fabrica unas resistencias con un valor nominal de 100  $\Omega$  ( $\Omega$  es el símbolo del ohmio, unidad de resistencia), pero en realidad la resistencia de las resistencias sigue una distribución normal con media 100,6 y desviación estándar 3. Por tanto, tendremos:

- El 68% de las resistencias tendrá un valor real entre 97,6  $\Omega$  y 103,6  $\Omega$
- El 95% de las resistencias tendrá un valor real entre 94,6  $\Omega$  y 106,6  $\Omega$
- El 99,7% de las resistencias tendrá un valor real entre 91,6  $\Omega$  y 109,6  $\Omega$
- Si queremos un intervalo que contenga el valor real del 90% de las resistencias, debemos considerar el intervalo 95,6650  $\Omega$  y 105,5350  $\Omega$

Todo esto nos permite lo siguiente:

- Comparar valores de diferentes leyes normales para ver cuáles son más probables.
- Reducir los cálculos de probabilidades de cualquier variable normal en el estudio de la distribución normal estándar.

Veamos primero cómo comparar varios valores de diferentes poblaciones normales. Por ejemplo, imaginad que las alturas de los hombres de una determinada tribu de la selva de la Amazonia siguen una distribución normal con media 167 cm y desviación estándar 3, mientras que las alturas de las mujeres siguen también una distribución normal, pero con media 162 cm y desviación estándar 2. Cogemos a un hombre y una mujer de esta tribu y los medimos: el hombre mide 168 cm y la mujer, 163 cm. Claro que la mujer es más baja que el hombre, pero dentro de sus sexos, ¿cuál de los dos se puede considerar más alto?

Para responder a esta pregunta, lo que hacemos es estandarizar los dos valores, es decir, pasarlos a observaciones de una normal estándar y después compararlos. Así, el valor estandarizado para el hombre es:

$$\frac{168 - 167}{3} = 0,333$$

y para la mujer:

$$\frac{163 - 162}{2} = 0,5$$

Puesto que  $0,333 < 0,5$  y los datos hacen referencia a la misma distribución, la mujer es, dentro de su sexo, más alta que el hombre.

Dada  $x$  una observación de una ley normal con media  $\mu$  y varianza  $\sigma^2$ , su valor estandarizado es  $\frac{x - \mu}{\sigma}$ .

### Ejemplo de las resistencias

Sabemos que la fábrica XSE fabrica unas resistencias con un valor nominal de  $100 \Omega$ , pero en realidad siguen una distribución normal con media  $100,6$  y desviación estándar  $3$ . Por otra parte, la SXX fabrica también resistencias con un valor nominal de  $100 \Omega$ , pero ahora en realidad siguen una distribución normal con media  $100,2$  y desviación estándar  $4$ .

Compramos una resistencia a la fábrica SXX cuyo valor real es de  $106$ , y una resistencia a la fábrica XSE con un valor real también de  $106$ . Así, sus valores estandarizados son:

$$106 - 100,6/3 = 1,8$$

para la resistencia comprada a XSE y de:

$$106 - 100,2/4 = 1,45$$

para la comprada a la fábrica SXX. Por tanto, la de XSE está en un lado alto si la comparamos con la de SXX.

### 3.3. Cálculo de probabilidades usando la ley normal estándar

Mediante el proceso de estandarización, el cálculo de probabilidades para una ley normal  $N(\mu, \sigma^2)$  se reduce a la evaluación de la función de distribución de una normal estándar. Como ya hemos visto antes, si  $X$  es una variable aleatoria  $N(\mu, \sigma^2)$ , entonces:

$$P(X \leq x) = P\left(Z \leq \frac{x - \mu}{\sigma}\right)$$

donde  $Z$  es una distribución normal estándar. También podemos escribir:

$$P(X \leq x) = \Phi\left(\frac{x - \mu}{\sigma}\right)$$

#### Observación

Dado que se trata de una ley continua, es lo mismo  $P(X \leq x)$  que  $P(X < x)$ .

donde  $\Phi$  denota la función de distribución de una distribución normal estándar.

Para la normal estándar hasta hace poco se utilizaban unas tablas estadísticas, pero actualmente es mucho más sencillo y potente utilizar herramientas informáticas. De todos modos, para acabar la sesión calculamos algunas probabilidades utilizando las tablas (podéis repetir los ejemplos con el *software* estadístico que utilizéis). Fijaos primero en que en la tabla que utilizaremos nos dan la cola superior de la distribución, es decir, para cada  $a$  nos dan:

$$P(Z \geq a)$$

y tenemos:

$$P(Z \geq a) = 1 - P(Z \leq a) = 1 - \Phi(a)$$

donde recordemos que  $\Phi$  es la función de distribución de una distribución normal estándar.

Fijaos también en que la función de densidad de la distribución normal estándar es simétrica. Esto hace que tengamos las propiedades siguientes:

Si  $Z$  es una variable aleatoria con distribución normal estándar y  $x > 0$ , entonces  $P(Z < -x) = P(Z > x)$  y  $P(Z > -x) = P(Z < x)$ .

Esta última propiedad de simetría nos será muy útil para calcular probabilidades mediante las tablas.

#### Software y hojas de cálculo

Aparte del *software* estadístico, las hojas de cálculo también nos permiten calcular probabilidades, y no sólo por las leyes normales.

Repasad el cálculo de probabilidades de la sesión "Variables aleatorias continuas".

#### Simplificación

Reducimos todos los cálculos a expresiones del tipo  $P(Z \geq a)$ , que son los que salen en las tablas.

Consultad las tablas de la distribución normal.

#### Ejemplo del cálculo de probabilidades

Supongamos que tenemos una variable aleatoria  $X$  con distribución normal de media 6 y varianza 4; calculemos lo siguiente:

a) Probabilidad de que  $X$  tome valores menores que 8:

$$P(X \leq 8) = P\left(\frac{X-6}{2} \leq \frac{8-6}{2}\right) = P\left(\frac{X-6}{2} \leq 1\right) = 1 - P\left(\frac{X-6}{2} \geq 1\right)$$

Mirando la tabla, encontramos por el punto 1,00 (la fila nos da las unidades y las décimas –cogemos 1,0– y la columna la centésima –cogemos 0,00) el valor 0,1587. Así:

$$P\left(\frac{X-6}{2} \geq 1\right) = 0,1587$$

y, por tanto:

$$P(X \leq 8) = 1 - 0,1587 = 0,8413$$

b) Probabilidad de que  $X$  tome valores mayores que 3:

$$P(3 < X) = P\left(\frac{3-6}{2} \leq \frac{X-6}{2}\right) = P\left(-1,5 \leq \frac{X-6}{2}\right)$$

y por simetría de la normal:

$$P\left(-1,5 \leq \frac{X-6}{2}\right) = P\left(1,5 \geq \frac{X-6}{2}\right) = 1 - 0,0668$$

c) Valor  $a$  tal que la probabilidad de que  $X$  sea mayor que  $a$  sea 0,22:

$$0,22 = P(X \geq a) = P\left(\frac{X-6}{2} \geq \frac{a-6}{2}\right)$$

En la tabla no encontramos la probabilidad 0,22. Encontramos la 0,2207 (corresponde al punto 0,77) y la 0,2177 (corresponde al punto 0,78). Entonces cogeremos el punto intermedio, es decir, 0,775 (con un ordenador encontraríamos el valor exacto). Por tanto:

$$\frac{a-6}{2} = 0,775$$

y tendremos  $a = 7,55$ .

#### 4. Resumen

Hemos visto las distribuciones continuas más habituales: la uniforme, la exponencial y la normal. Hemos hecho un estudio a fondo de esta última y hemos visto el proceso de estandarización y la reducción de cualquier cálculo con una ley normal a un cálculo con una ley normal estándar. Hemos visto también cómo utilizar las tablas de la distribución normal.

## Ejercicios

1. Sabemos que el tiempo que transcurre entre que abrimos nuestra tienda de ordenadores y la llegada del primer cliente sigue una distribución exponencial de parámetro 0,1 (en minutos). Si abrimos a las nueve y media, calculad:

- a) La probabilidad de que llegue después de las diez menos cuarto.
- b) La probabilidad de que llegue exactamente a las 9.42 h.
- c) La probabilidad de que no nos deje ni los cinco minutos necesarios para fumar un cigarrillo.
- d) La hora a la que esperamos que llegue.

2. En 1972 la altura de los hombres en edad militar en España seguía una distribución normal con media 169 cm y desviación estándar de 6 cm. Calculad:

- a) Si eran rechazados los más bajos de 150 cm y los más altos de 200 cm, ¿qué proporción de hombres tenían la suerte de ser rechazados?
- b) Si un hombre en edad militar medía 175 cm y se encontraba, por ejemplo en un bar, a otro hombre en edad militar, ¿cuál es la probabilidad de que el primero fuese más alto que el segundo? ¿Y de que fuesen exactamente iguales?
- c) Calculad el valor estandarizado de vuestra altura y la probabilidad de que un hombre en edad militar del año 1972 fuese más alto que vosotros. ¿Creéis que este resultado todavía es cierto?
- d) Determinad un intervalo en el que se encuentren las alturas del 90% de los hombres en edad militar.

## Solucionario

1. Llamaremos  $X$  a la variable que nos da el tiempo entre que abrimos y la llegada del primer cliente. Entonces deberemos calcular:

- a) Realizaremos este apartado utilizando la función de densidad:

$$P(X > 15) = \int_{15}^{\infty} (0,1) \exp(-0,1x) dx = [-\exp(-0,1x)]_{15}^{\infty} = \exp(-1,5) = 0,2231$$

- b) Puesto que se trata de una ley continua, la probabilidad de que tome un valor concreto será cero, es decir:

$$P(X = 12) = 0$$

- c) Realizaremos este apartado utilizando la función de distribución:

$$P(X < 5) = P(X \leq 5) = F(5) = 1 - \exp(-0,5) = 0,3935$$

- d) La esperanza de la variable aleatoria  $X$  es  $1/0,1 = 10$ . Esperamos, por tanto, que el primer cliente llegue a las 9.40 h.

2. Haremos este problema utilizando la tabla de la distribución normal estándar. De todos modos, si después lo repetís utilizando un *software* estadístico, encontraréis unos resultados más precisos y más rápidamente.

Llamamos  $Y$  a la variable que nos da la altura de un hombre en edad militar en 1972.

a) Debemos calcular:

$$P(Y < 150) + P(Y > 200)$$

Tenemos:

$$P(Y < 150) = P\left(\frac{Y - 169}{6} < \frac{150 - 169}{6}\right) = P\left(\frac{Y - 169}{6} \leq -3,1667\right) = 0,0008, \text{ y}$$

$$P(Y > 200) = P\left(\frac{Y - 169}{6} > \frac{200 - 169}{6}\right) = P\left(\frac{Y - 169}{6} \geq 5,1667\right)$$

que mirando a la tabla es despreciable. Por tanto, nos sale una probabilidad de 0,0008.

b) Nos piden primero:

$$\begin{aligned} P(Y > 175) &= P(Y < 175) = P\left(\frac{Y - 169}{6} < \frac{175 - 169}{6}\right) = P\left(\frac{Y - 169}{6} < 1\right) = \\ &= P(Z < 1) = 1 - P(Z \geq 1) = 1 - 0,1587 = 0,8413 \end{aligned}$$

y la probabilidad de que midan exactamente igual será:

$$P(Y = 175) = 0$$

ya que la normal es una distribución continua.

c) Supongamos que un hombre mide 181 cm de altura. Su valor estandarizado será:

$$\frac{181 - 169}{6} = 2$$

y si miramos la tabla de la normal, tenemos que sólo el 2,28% de los hombres en edad militar eran más altos que nuestro hombre. Creemos que hoy en día esto no es cierto.

d) Debemos coger el intervalo dado por la media menos 1,645 desviaciones estándar y la media más 1,645 desviaciones estándar, es decir:

$$(169 - (1,645 \cdot 6); 169 + (1,645 \cdot 6)) = (159,13; 178,87)$$



## Procesos estocásticos

Acabaremos este módulo con una última sesión dedicada a introducir, de manera básicamente descriptiva, el concepto de proceso estocástico. Aunque la utilización de procesos estocásticos supera con creces el objetivo de este curso, a causa de su importancia (por ejemplo, se utilizan procesos estocásticos en el tratamiento de imágenes, el reconocimiento de la voz o el estudio de colas) es interesante que conozcáis algunas de sus ideas. Sin embargo, veréis que el tratamiento de esta sesión es un tanto diferente al de las otras, es bastante más corta y no encontraréis ejercicios.

### 1. Definiciones

En ciertos experimentos aleatorios el resultado que obtenemos no es un valor numérico concreto, sino una función que depende del tiempo. Estas situaciones no se pueden modelar utilizando variables aleatorias, deberemos utilizar los llamados *procesos estocásticos*.

#### Estocástico o aleatorio

Podemos usar indistintamente los términos *procesos estocásticos* o *procesos aleatorios*.

#### Ejemplos de situaciones que no podemos modelar con una variable aleatoria

Son muchas. Veamos algunos ejemplos: la temperatura en el Tibidabo a lo largo de un día (temperatura cada 15 minutos), el comportamiento del consumo de electricidad en Palau de Plegamans durante un día (consumo a cada momento) o los correos electrónicos que nos llegan a lo largo de un día.

Consideremos, por ejemplo, el consumo de electricidad en la ciudad de Barcelona durante un día. Supongamos que a cada instante  $t$  ( $t = 0$  indicará el principio del día) sabemos que el consumo de electricidad sigue una distribución normal con media  $\mu_t$  y varianza  $\sigma_t^2$  (lógicamente, el consumo depende de la hora en que estemos, de manera que hemos supuesto que la media y la varianza dependen de  $t$ ). Este consumo es una variable aleatoria que denotaremos por  $X(t)$ . Es decir, definimos  $X(t)$  como el consumo de electricidad de la ciudad de Barcelona en el instante  $t$ .

Sin embargo, lo más probable es que no nos interese especialmente el consumo en un instante concreto, sino que queramos estudiar cómo se comporta el consumo de electricidad a lo largo de todo el día. Así, en realidad, nos interesa considerar la familia de variables aleatorias:

#### Unidad de tiempo

Daremos el tiempo en segundos: 24 horas = 86.400 segundos.

$$\{X(t), 0 \leq t \leq 86400\}$$

Esta familia de variables aleatorias es un proceso estocástico.

Un **proceso estocástico** es una familia de variables aleatorias

$$\{X(t), 0 \leq t \leq T\}$$

indexadas por el tiempo.

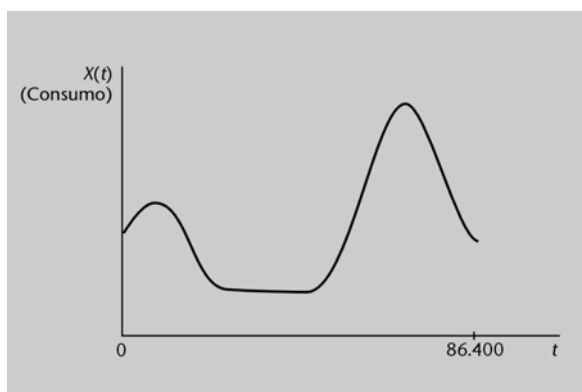
**Observación**

Para cada instante  $t$  fijado,  $X(t)$  es una variable aleatoria.

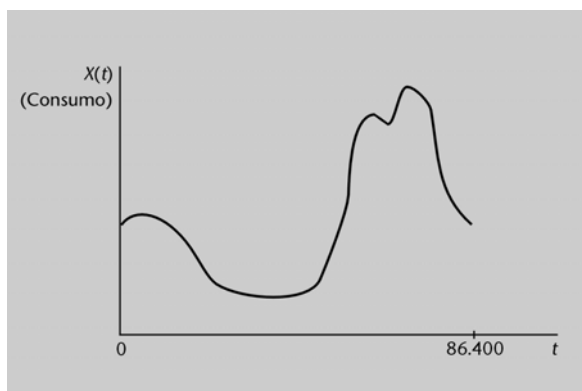
Como ya hemos comentado, el resultado de la realización de una variable aleatoria es un valor numérico. En cambio, el resultado de la realización de un proceso estocástico es una función que depende del tiempo y que llamaremos *trayectoria*.

Una **trayectoria** es el resultado de la realización de un proceso estocástico.

En el caso anterior, en el que estudiábamos el consumo de electricidad en Barcelona, si cogemos un día determinado, podemos obtener la trayectoria (realización) siguiente:



Mientras que para otro día podemos obtener otra (realización) diferente, como por ejemplo:



Veremos dos clases de procesos estocásticos: los **procesos a tiempo discreto** y los **procesos a tiempo continuo**:

En un proceso a tiempo discreto tenemos una familia de variables aleatorias indexada sólo sobre los naturales:

$$\{X(n), n \in N\}$$

En un proceso a tiempo continuo tenemos una familia de variables aleatorias indexada en un intervalo:

$$\{X(t), 0 \leq t \leq T\}$$

**Notación**

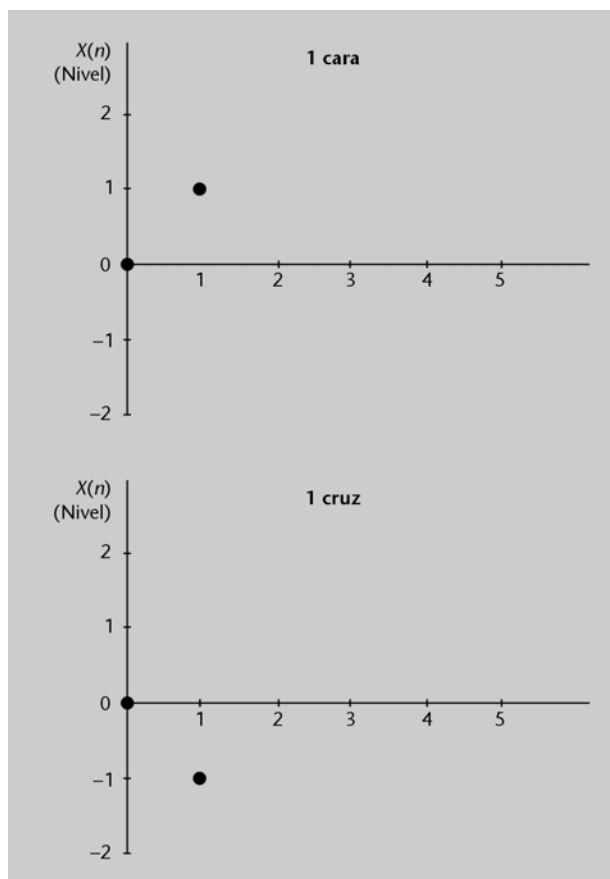
En los procesos a tiempo discreto utilizamos  $X(n)$  en lugar de  $X(t)$ .

Principalmente estudiaremos uno de cada tipo: el paseo aleatorio (proceso a tiempo discreto) y el proceso de Poisson (proceso a tiempo continuo). Por su importancia, también haremos una pequeña referencia al proceso de Wiener.

## 2. Procesos aleatorios a tiempo discreto: el paseo aleatorio

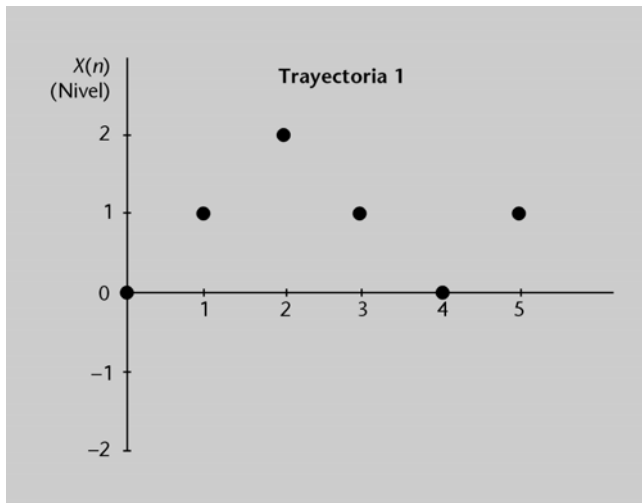
El paseo aleatorio es uno de los procesos aleatorios más sencillos. Explicaremos cómo podemos construirlo.

Comenzaremos ( $n = 0$ ) al inicio –punto  $(0,0)$ . Entonces lanzamos una moneda y, si sale cara, “C”, subimos un nivel, mientras que si sale cruz, “+”, bajamos uno. Llamamos  $X(1)$  a la variable aleatoria que nos da el número de nivel al que vamos a parar después del primer lanzamiento. Así, en el primer paso tenemos dos posibles trayectorias:

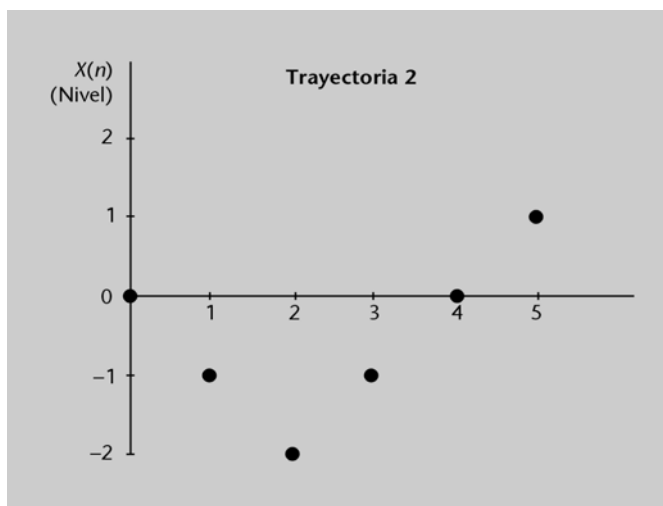


Podemos repetir este proceso indefinidamente. En el instante  $n$  lanzamos de nuevo una moneda y, según el resultado, subimos (cara) o bajamos (cruz). Ahora llamamos  $X(n)$  a la variable que nos da el nivel al que hemos ido a parar en el instante  $n$ .

Por ejemplo, si el resultado de los cinco primeros lanzamientos es C, C, +, +, C, tenemos la trayectoria:



en la que claramente  $X(5)$  vale 1, mientras que si el resultado de los cinco lanzamientos es +, +, C, C, C, llegamos al mismo punto pero con una trayectoria diferente.



Entonces el proceso  $\{X(n), n \geq 0\}$  es un **paseo aleatorio**.

Estudiaremos algunas propiedades del paseo aleatorio. Llamamos  $I(k)$  a la variable aleatoria que vale 1 si nos sale una cara cuando hacemos el  $k$ -ésimo lanzamiento de la moneda y que vale  $-1$  si nos sale una cruz. En este caso:

$$X(n) = \sum_{k=1}^n I(k).$$

\* En inglés el paseo aleatorio se denomina *random walk*.

Esta forma de expresar cada  $X(n)$  nos permite estudiar algunas de sus propiedades. Observad primero que es muy fácil comprobar que:

$$E(I(k)) = P(I(k) = 1) - P(I(k) = -1) = 0$$

$$\text{Var}(I(k)) = 1^2 P(I(k) = 1) + (-1)^2 P(I(k) = -1) = 1$$

Puesto que las  $I(k)$  son variables aleatorias independientes (el resultado de un lanzamiento de la moneda no afecta a los otros lanzamientos), utilizando las propiedades de la esperanza y de la varianza que hemos estudiado, tenemos:

$$E(X(n)) = E\left(\sum_{k=1}^n I(k)\right) = \sum_{k=1}^n E(I(k)) = 0$$

$$\text{Var}(X(n)) = \text{Var}\left(\sum_{k=1}^n I(k)\right) = \sum_{k=1}^n \text{Var}(I(k)) = n$$

Es decir, en cualquier paso la esperanza nos indica que deberíamos estar al nivel cero, pero con una varianza cada vez mayor.

### 3. Procesos a tiempo continuo

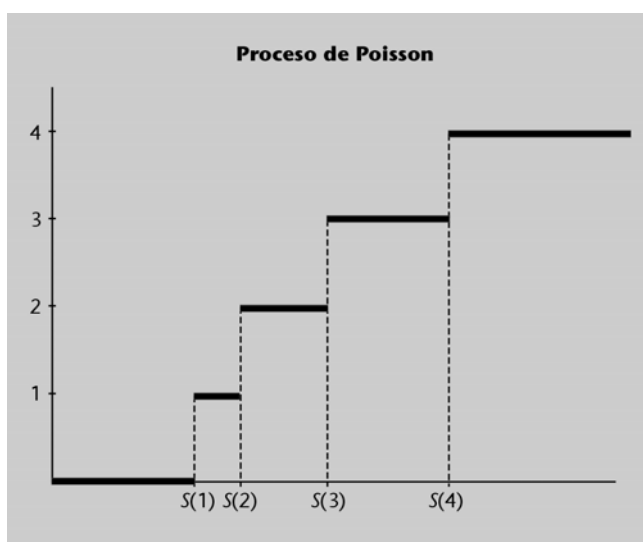
#### 3.1. El proceso de Poisson

El proceso de Poisson se utiliza básicamente para modelar los llamados *procesos de colas*. En ellos se pueden incluir muchos procesos: coches que llegan al peaje de una autopista, clientes que llegan a un banco, peticiones que llegan a un servidor de Internet, llamadas que pasan por una centralita, etc.

Consideremos, por ejemplo, el caso de un servidor de Internet. Queremos estudiar las peticiones que va recibiendo este servidor. Podemos suponer que las peticiones llegan de una en una. Si llamamos  $X(t)$  al número de peticiones que ha recibido el servidor hasta el instante  $t$ , encontramos que una posible realización de este proceso será del tipo:

#### Observación

No hay que confundir el proceso de Poisson con la distribución de Poisson.



donde  $S(1)$  indica el instante en que llega la primera petición,  $S(2)$  indica el instante en que llega la segunda y, en general,  $S(i)$  indica el instante en que llega la  $i$ -ésima.

Éste es un ejemplo de un proceso que se puede representar utilizando el proceso de Poisson. Tiene algunas características fundamentales:

- Para cada instante  $t$ ,  $X(t)$  seguirá una distribución de Poisson de parámetro  $\lambda t$ , donde  $\lambda$  es un parámetro que depende del proceso.
- Las diferencias entre los tiempos de llegada siguen una distribución exponencial de parámetro  $\lambda$ , es decir,  $S(i+1) - S(i)$  siguen una exponencial de parámetro  $\lambda$ .

Así, tenemos la propiedad siguiente:

Si  $\{X(t), t \geq 0\}$  sigue un proceso de Poisson de parámetro  $\lambda$ , entonces  $X(t)$  sigue una distribución de Poisson de parámetro  $\lambda t$ .

Evidentemente, la definición rigurosa del proceso de Poisson es bastante más complicada y del estudio de sus propiedades se han escrito libros enteros. En todo caso, hemos visto una pequeña introducción a algunas ideas básicas.

### Ejemplo de la centralita

Las llamadas que llegan a una centralita siguen un proceso de Poisson de parámetro 2 (daremos el tiempo en minutos). Podemos calcular las probabilidades siguientes:

a) Probabilidad de que en los cinco primeros minutos lleguen exactamente cuatro llamadas.

La variable  $X(5)$ , que corresponde al número de llamadas que hemos recibido hasta el minuto 5, sabemos que seguirá una distribución de Poisson de parámetro  $2 \cdot 5 = 10$ , por tanto:

$$P(X(5) = 4) = \frac{(10)^4}{4!} \exp(-10) = 0,0189$$

b) Probabilidad de que entre la primera y la segunda llamada pasen más de tres minutos.

Si llamamos  $Y$  a la variable que nos da el tiempo que transcurre entre la primera y la segunda llamada, por la definición del proceso de Poisson sabemos que seguirá una distribución exponencial de parámetro 2.

Así, debemos calcular:

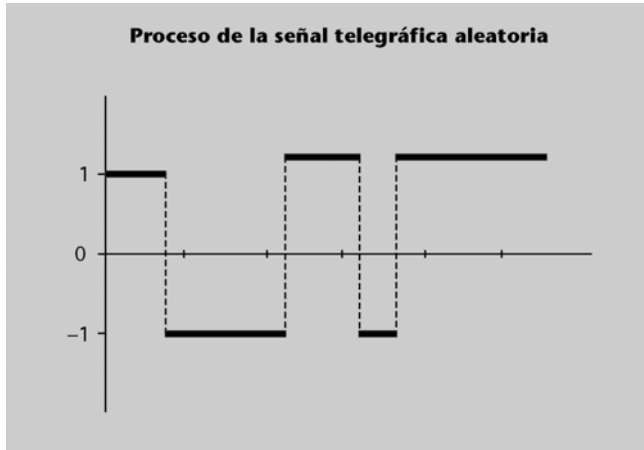
$$P(Y > 3) = \int_3^{\infty} 2 \exp(-2x) dx = [-\exp(-2x)]_3^{\infty} = \exp(-6) = 0,0025$$

Hay un proceso importante que es una derivación del proceso de Poisson. Es el llamado *proceso de la señal telegráfica aleatoria*. Consideremos un proceso  $X(t)$

#### Proceso de conteo

El proceso de Poisson es un proceso de conteo, ya que nos dedicamos a contar el número de veces que sucede un acontecimiento. El nombre de proceso de Poisson proviene del hecho de que las leyes que aparecen son del tipo Poisson.

que sólo puede tomar los valores 1 ó -1 y que al inicio vale 1 ó -1 con probabilidad 0,5. El proceso  $X(t)$  cambia de polaridad (pasa de 1 a -1 o a la inversa) en los puntos de salto del proceso de Poisson, de manera que obtendremos trayectorias del tipo siguiente:

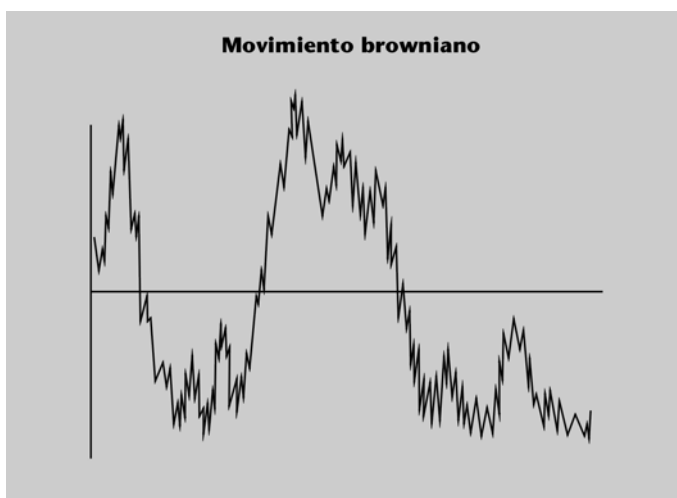


#### Proceso de la señal telegráfica

A partir del proceso de Poisson podemos construir fácilmente el proceso de la señal telegráfica.

### 3.2. El proceso de Wiener o movimiento browniano

El proceso de Wiener o movimiento browniano es quizá el proceso que más se ha estudiado. En 1827 el botánico Robert Brown observó el movimiento errático de granos de polen cuando estaban en suspensión en un líquido (su movimiento es muy rápido e irregular). Ahora sabemos que este tipo de movimiento está provocado por el impacto de las moléculas del fluido en el grano de polen. Como podéis imaginar, el movimiento que se produce es bastante caótico, de manera que se obtienen trayectorias muy irregulares, con un aspecto como el siguiente:



#### Terminología

Como podéis suponer, el nombre de *movimiento browniano* proviene del Sr. Brown. El nombre de Wiener es el del matemático que hizo la primera construcción rigurosa del proceso.

#### Finanzas

Últimamente han aparecido aplicaciones del movimiento browniano en las finanzas, especialmente en la bolsa, para modelar los precios de las acciones.

Este tipo de proceso se utiliza también para dar un modelo al movimiento de una molécula dentro de un gas, movimiento provocado y modificado continuamente a causa del impacto de dicha molécula con las otras del gas.

Entre muchas otras características, presenta una que lo hace muy útil: las leyes que aparecen son distribuciones normales.

Si  $\{X(t), t \geq 0\}$  sigue un proceso de Wiener, entonces  $X(t)$  sigue una distribución normal con media 0 y varianza  $t^2$ .

Como ya hemos dicho para el proceso de Poisson, la definición rigurosa del estudio de sus propiedades no es el objeto de esta sesión.

#### 4. Resumen

Hemos presentado el concepto de proceso estocástico para estudiar situaciones en las que no podemos utilizar una única variable aleatoria. Asimismo, hemos presentado algunos de los procesos más habituales: el paseo aleatorio, el proceso de Poisson y el proceso de Wiener.