Universidad
Rey Juan Carlos

GRADO EN MATEMÁTICAS

PROBABILIDAD

TEMA 3

VARIABLES ALEATORIAS

Sonia Hernández Alonso
Área de Estadística e Investigación Operativa (URJC)

Esquema

- Introducción
- Definición de variable aleatoria
- Tipos de variables aleatorias
- Función de distribución
- Variables aleatorias discretas
- Esperanza y varianza de variables aleatorias discretas.
- Modelos discretos especiales
  - Distribución binomial
  - Distribución geométrica
  - Distribución de Poisson

Esquema (continuación)

- Variables aleatorias continuas:
  - Función de densidad. Propiedades
  - Función de distribución de variables aleatorias continuas
  - Esperanza y varanza de variables aleatorias continuas
- Modelos continuos especiales:
  - Distribución uniforme
  - Distribución exponencial
  - Distribución normal
- Distribución de transformaciones de variables aleatorias
- Esperanza de transformaciones de variables aleatorias
- Desigualdades de Markov y Chebyshev

Introducción

Variables aleatorias: introducción

- Muchos experimentos aleatorios tienen un espacio muestral, Ω, cuyos elementos no son numéricos.

- Por ejemplo, para el experimento aleatorio consistente en lanzar al aire una moneda tres veces, el espacio muestral es

$$\Omega = \{\mathcal{CCC}, \mathcal{CCX}, \mathcal{CXC}, \mathcal{CXX}, \mathcal{XCC}, \mathcal{XCX}, \mathcal{XXC}, \mathcal{XXX}\}\,.$$

- Matemáticamente, resulta más útil **cuantificar** los resultados de un espacio muestral, ya que de esta forma se pueden emplear medidas numéricas para analizar las principales características del experimento.

- Además, en muchas ocasiones, no interesa estudiar todos los detalles del experimento, sino que el interés se centra en ciertas magnitudes numéricas. Por ejemplo, cuando se lanzan varios dados, podemos estar interesados en conocer cuál es la suma obtenida, y no en los resultados concretos de cada lanzamiento. O, al escoger una persona al azar de una población, puede que sólo nos interese saber su altura.

- Las **variables aleatorias** asignan un valor numérico a cada uno de los resultados posibles de un experimento.

Definición de variable aleatoria

Sonia Hernández Alonso
Probabilidad-Grado en Matemáticas (URJC)

¿Qué es una variable aleatoria (v.al.)?

- **Definición**: Sea $(\Omega, \mathcal{A}, P)$ un espacio probabilístico. Una **variable aleatoria** es cualquier función $X : \Omega \longrightarrow \mathbb{R}$ que para todo $B \in \mathcal{B}(\mathbb{R})$ verifica
$$X^{-1}(B) \in \mathcal{A},$$
o lo que es lo mismo,
$$\{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{A}.$$

- Una función que cumple esta condición se dice que es una función **medible** con respecto al espacio $(\Omega, \mathcal{A}, P)$.

- Dicha condición permitirá calcular probabilidades del tipo $P(X \in B)$ para cualquier boreliano B.

- Cuando el espacio muestral Ω es finito o infinito numerable, lo habitual es tomar como $\sigma-$álgebra $(\mathcal{A})$ el conjunto de partes de Ω, $\mathcal{P}(\Omega)$. En estos casos, la condición de ser medible siempre se verifica.

Ejemplo de variable aleatoria

- Para el experimento aleatorio consistente en lanzar al aire una moneda tres veces, el espacio muestral es

$$\Omega = \{\mathcal{CCC}, \mathcal{CCX}, \mathcal{CXC}, \mathcal{CXX}, \mathcal{XCC}, \mathcal{XCX}, \mathcal{XXC}, \mathcal{XXX}\}.$$

- Los elementos de Ω no son de tipo numérico, pero podemos definir alguna variable aleatoria que los transforme en números.

- Consideremos, por ejemplo, la variable

$$X = \text{ número de caras}$$

- Esta variable aleatoria,

$$X : \Omega \longrightarrow \mathbb{R},$$

hace la siguiente asignación de números a cada elemento de Ω:

$$\begin{array}{cccc}
X(\mathcal{XXX}) = 0 & X(\mathcal{XXC}) = 1 & X(\mathcal{CCX}) = 2 & X(\mathcal{CCC}) = 3 \\
 & X(\mathcal{XCX}) = 1 & X(\mathcal{CXC}) = 2 & \\
 & X(\mathcal{CXX}) = 1 & X(\mathcal{XCC}) = 2 &
\end{array}$$

¿Qué es lo aleatorio de la v.al. X?

- Observemos que la asignación numérica que la variable aleatoria X establece para cada suceso elemental $\omega \in \Omega$ es determinista.
- La aleatorioridad reside en el hecho de que, antes realizar el experimento, no sabemos cúal de los elementos $\omega \in \Omega$ va a ocurrir.

Probabilidad inducida por una variable aleatoria

- Sobre los elementos de Ω existe una **distribución de probabilidad**, y la variable aleatoria X traslada esta estructura probabilística a $\mathbb{R}$.

- Si partimos del espacio probabilístico $(\Omega, \mathscr{A}, P)$ y consideramos una variable aleatoria

$$X : \Omega \longrightarrow \mathbb{R},$$

podemos definir

$$P_X : \mathbb{R} \longrightarrow [0,1]$$

de la forma siguiente: para cada $B \in \mathcal{B}$,

$$P_X(B) = P(X \in B) = P(\{\omega \in \Omega : X(\omega) \in B\}) = P\left(X^{-1}(B)\right)$$

- **Teorema 1:** P_X es una probabilidad sobre $(\mathbb{R}, \mathcal{B})$.

demostración:.........pizarra

Ejemplo: probabilidad inducida por una v.al.

- **Ejercicio 1:** Para el experimento aleatorio consistente en lanzar una moneda equilibrada tres veces, y la variable aleatoria

$$X = \text{ número de caras,}$$

determinar la distribución de probabilidad inducida, P_X.

resolución:.........pizarra

- **Observación:** P es una probabilidad sobre el espacio probabilizable $(\Omega, \mathscr{A})$, mientras que P_X es una probabilidad sobre $(\mathbb{R}, \mathcal{B})$.

Solución al Ejercicio 1

- Come se ha dicho, sobre los elementos de Ω existe una **distribución de probabilidad**, y la variable aleatoria X traslada esta estructura probabilística a $\mathbb{R}$.

- Para el triple lanzamiento de una moneda equilibrada y la variable aleatoria

$$X = \text{ número de caras}$$

se tiene que

$$\begin{aligned}
P_X(0) &= P\left[X^{-1}(0)\right] = P(X = 0) = P(\mathcal{XXX}) = \frac{1}{8} = 0.125,\\
P_X(1) &= P\left[X^{-1}(1)\right] = P(X = 1) = P(\mathcal{XXC}, \mathcal{XCX}, \mathcal{CXX}) = \frac{3}{8} = 0.375,\\
P_X(2) &= P\left[X^{-1}(2)\right] = P(X = 2) = P(\mathcal{XCC},\ \mathcal{CXC},\ \mathcal{CCX}) = \frac{3}{8} = 0.375,\\
P_X(3) &= P\left[X^{-1}(3)\right] = P(X = 3) = P(\mathcal{CCC}) = \frac{1}{8} = 0.125.
\end{aligned}$$

Tipos de variables aleatorias

Sonia Hernández Alonso
Probabilidad-Grado en Matemáticas (URJC)

Variables aleatorias discretas y continuas

- Se llama **soporte** de una variable aleatoria X al conjunto de valores que puede tomar la variable. Denotaremos dicho conjunto por S_X.

- Dependiendo cómo sea su soporte, las variables aleatorias se clasifican discretas y continuas:

  - Una **variable aleatoria X es discreta** si y sólo si existe algún conjunto $A \subset \mathbb{R}$ finito o infinito **numerable** tal que

    $$P(X \in A) = 1$$

    En consecuencia, el soporte de una variable discreta está formado **puntos aislados**, es decir, valores puntuales.

  - Una **variable aleatoria X es continua** si y sólo si su soporte es no numerable, ya que **incluye a todos los números de algún intervalo de $\mathbb{R}$**.

- Las distribuciones de probabilidad de las variables discretas y de las variables continuas difieren mucho en varios aspectos.

Ejemplos de variables aleatorias discretas

- X= número de caras en tres lanzamientos de una moneda,
- N=número de lanzamientos de un dado hasta que sale el primer 5,
- S=suma de las puntuaciones de tres lanzamientos de un dado,
- L= número de llamadas atendidas en una central telefónica durante un mes,
- H= número de crías en por camada en gatas,
- etcetera

Ejemplos de variables aleatorias continuas

- $T=$ tiempo que transcurre entre dos llamadas telefónicas consecutivas en una centralita,
- $P=$ peso de un ciervo escogido al azar,
- $T=$ temperatura en Móstoles en un instante elegido aeatoriamente,
- $D=$ duración de una bombilla,
- $L=$ Longitud de un triple salto olímpico,
- etcetera

Función de distribución

Sonia Hernández Alonso
Probabilidad-Grado en Matemáticas (URJC)

¿Qué es la función de distribución de una v.al?

- Toda variable aleatoria tiene asociada una función llamada función de distribución que la caracteriza completamente.

- **Definición**: Dado un espacio probabilístico $(\Omega, \mathscr{A}, P)$ y una variable aleatoria X, se define su **función de distribución**,

$$F_X : \mathbb{R} \longrightarrow \mathbb{R},$$

  como

$$\begin{aligned} F_X(x) &= P_X(-\infty, x] \\ &= P\left(X^{-1}(-\infty, x]\right) \\ &= P\left(\omega \in \Omega : X\left(\omega\right) \leq x\right) \\ &= P(X \leq x) \end{aligned}$$

- Es decir, la función de distribución es una función de **probabilidad acumulada**.

Ejemplo: función de distribución de una v.al.

- **Ejercicio 2:** Determinar la función de distribución de la v. al.

$$X = \text{ número de caras,}$$

en el triple lanzamiento de una moneda equilibrada.

Representar gráficamente esta función y analizar sus propiedades.

<u>*resolución*</u>:.........pizarra

Solución al Ejercicio 2

- La expresión de la función de distribución de la variable

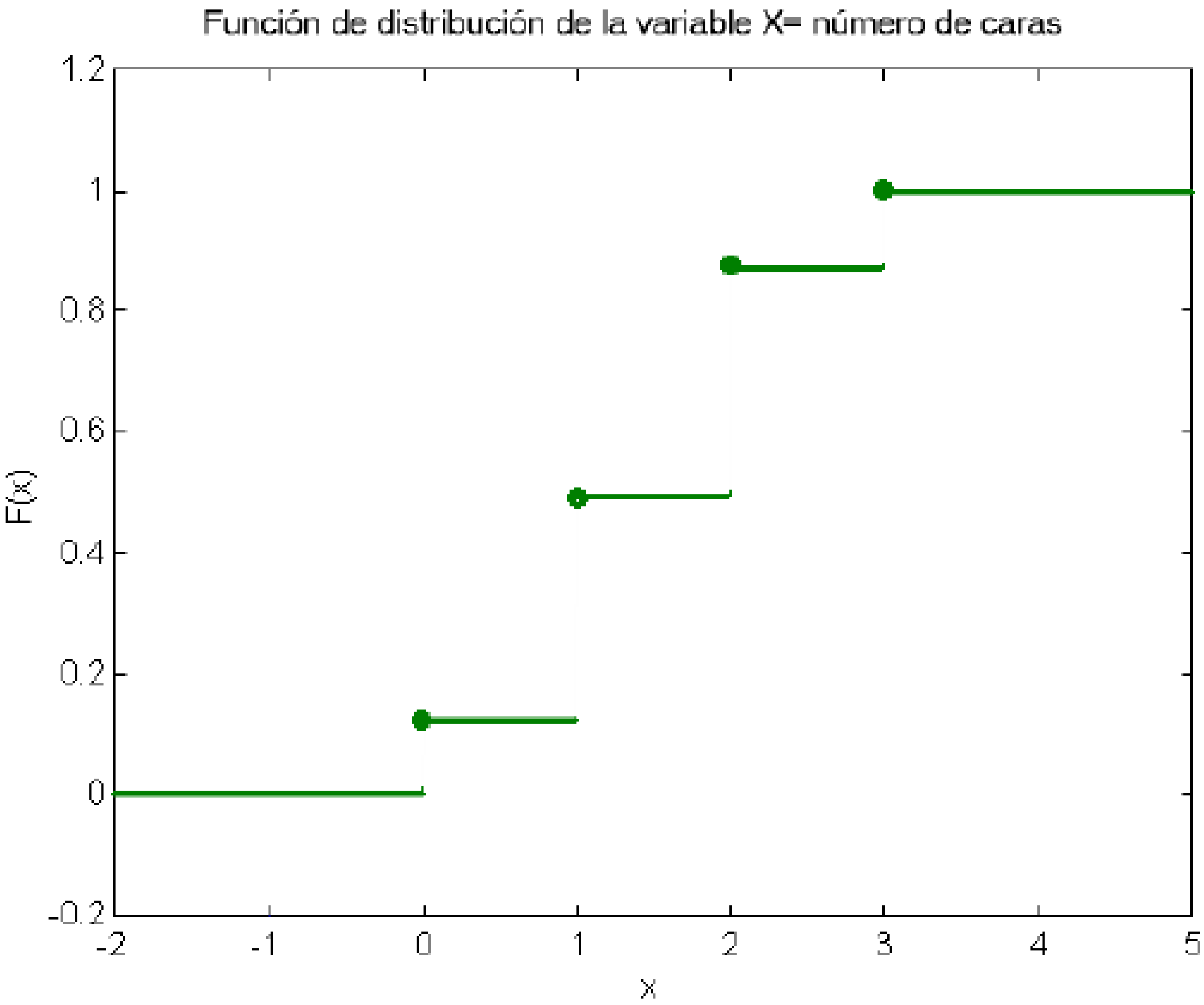
$$X = \text{ número de caras}$$

es

$$F_X(t) = P(X \leq t) = \begin{cases} 0 & \text{si } t < 0, \\ 0.125 & \text{si } 0 \leq t < 1, \\ 0.5 & \text{si } 1 \leq t < 2, \\ 0.875 & \text{si } 2 \leq t < 3, \\ 1 & \text{si } t \geq 3. \end{cases}$$

URJC

Solución al Ejercicio 2 (continuación)



Solución al Ejercicio 2: observaciones sobre F_X

- En este ejemplo observamos que F_X:
  - Comienza en 0, ya que $F_X(t) = 0$ para todo $t < 0$.
  - Termina en 1, ya que $F_X(t) = 1$ para todo $t \geq 3$.
  - Es continua por la derecha, pero presenta discontinuidades en los puntos 0, 1, 2, y 3, que son los puntos del soporte de X.
  - Es monótona no decreciente.
  - Crece a saltos, los puntos de salto son los puntos del soporte, y la amplitud de cada salto es el valor de la función de masa en ese punto.
- Como veremos, este es el aspecto que presentan en general las funciones de distribución de las variables aleatorias discretas.

Ejemplo: cálculo de probabilidades con F_X (discreta)

- **Ejercicio 3:** Estudios recientes han confirmado que las gaviotas que anidan en zonas costeras cercanas a las fábricas de acero están experimentando sorprendentes mutaciones.

  Entre otras anomalías, el número de dedos que tienen (contando ambas patas) varía de unos ejemplares a otros, siendo una variable aleatoria, D, con función de distribución

$$F_D(x) = \begin{cases} 0 & \text{si } x < 6, \\ 0.1 & \text{si } 6 \leq x < 8, \\ 0.5 & \text{si } 8 \leq x < 9, \\ 0.8 & \text{si } 9 \leq x < 10, \\ 1 & \text{si } x \geq 10. \end{cases}$$

  Un equipo investigador se instala en una playa situada en las proximidades de una fábrica de acero en la cual hay una **enorme cantidad** de gaviotas y las selecciona aleatoriamente para estudiar el número de dedos que presentan.

Ejercicio 3 (continuación)

1. Hallar la probabilidad de que una gaviota elegida al azar en esta zona tenga una cantidad par de dedos.

   Solución: *0.7*

2. Se ha comprobado que una de las gaviotas capturadas en esta playa tiene menos de 10 dedos. ¿Cuál es la probabilidad de que tenga un número par de dedos?

   Solución: *0.625*

3. Hasta ahora los investigadores han capturado en este área un total de 10 gaviotas. Calcular la probabilidad de que 2 de ellas tengan una cantidad impar dedos y el resto una cantidad par.

   Solución: *0.0014*

Ejercicio 3 (continuación)

4. Tras finalizar el estudio anterior, los miembros del equipo de investigación necesitan analizar las características de una gaviota que tenga un número impar de dedos, para lo cual seleccionan gaviotas de manera aleatoria, cuentan sus dedos y, en caso de que sean una cantidad par, las dejan en libertad.

   ¿Cuál es la probabilidad de que tengan que capturar exactamente 5 gaviotas hasta dar con una que tenga un número impar de dedos ?

   Solución: *0.07203*

Ejemplo: cálculo probabilidades con F_X (continua)

- **Ejercicio 4:** En cierto experimento químico, la temperatura de reacción (medida en grados centígrados) es una variable aleatoria, T, con función de distribución

$$F_T(x) = \begin{cases} 0 & \text{si } x < -1, \\ \dfrac{x^3+1}{9} & \text{si } -1 \leq x < 2, \\ 1 & \text{si } x \geq 2. \end{cases}$$

Calcular las siguientes probabilidades:

1. $P(T > 0)$
2. $P(0 \leq T \leq 1)$
3. $P(T = 0.6)$
4. $P(1 < T < 4)$
5. $P\left(T \in (-0.8, 0.2]\right)$
6. $P(|T| < 0.5)$

Solución al Ejercicio 4

1. $P(T > 0) = \dfrac{8}{9} = 0.8889$
2. $P(0 \leq T \leq 1) = \dfrac{1}{9} = 0.1111$
3. $P(T = 0.6) = 0$
4. $P(1 < T < 4) = \dfrac{7}{9} = 0.7778$
5. $P\left(T \in (-0.8, 0.2]\right) = 0.05778$
6. $P(|T| < 0.5) = 0.02778$

Observación: propiedades función distribución

- Más adelante (cuando tengáis os suficientes conocimientos sobre **Cálculo**) demostraremos la siguiente proposición. Por el momento vamos a razonar que se trata de un resultado muy acorde a la intuición.

- **Proposición *:** Sea $(\Omega, \mathscr{A}, P)$ un espacio probabilístico, y $X : \Omega \longrightarrow \mathbb{R}$ cualquier variable aleatoria.

  La función de distribución de X, F_X, verifica las siguientes propiedades:

  1. F_X es monótona no decreciente.
  2. F_X es continua por la derecha.
  3. $\lim\limits_{x \to \infty} F_X(x) = 1$.
  4. $\lim\limits_{x \to -\infty} F_X(x) = 0$.

  demostración:.........pendiente

Notación: límites por la derecha y por la izquierda

- Dada una función G, es común denotar $G(a+)$ a su **límite por la derecha** en el punto a, es decir,

$$G(a+) = \lim_{x \to a^+} G(a) = \lim_{h \downarrow 0} G(a+h).$$

- De manera análoga, $G(a-)$ denota el **límite por la izquierda** de G en el punto a, esto es,

$$G(a-) = \lim_{x \to a^-} G(a) = \lim_{h \downarrow 0} G(a-h).$$

- Utilizando esta notación, podemos escribir que G es **continua por la derecha** en un punto z si y sólo si

$$G(z+) = G(z),$$

  y que es **continua por la izquierda** en dicho punto si y sólo si

$$G(z-) = G(z).$$

- Recordemos que para que G sea una **función continua** en z es condición necesaria que sea continua por la derecha y por la izquierda.

Cálculo de probabilidades sobre X a partir de F_X

URJC

- Sea X una variable aleatoria y F_X su función de distribución. Para cualquier par de números reales $a \leq b$ se verifica:

1. $P(X < a) = F_X(a-)$
2. $P(X = a) = F_X(a) - F_X(a-)$
3. $P(X > a) = 1 - F_X(a)$
4. $P(X \geq a) = 1 - F_X(a-)$
5. $P\left(X \in (a, b]\right) = P(a < X \leq b) = F_X(b) - F_X(a)$
6. $P\left(X \in [a, b]\right) = P(a \leq X \leq b) = F_X(b) - F_X(a-)$
7. $P\left(X \in (a, b)\right) = P(a < X < b) = F_X(b-) - F_X(a)$
8. $P\left(X \in [a, b)\right) = P(a \leq X < b) = F_X(b-) - F_X(a-)$

Variables aleatorias discretas

Sonia Hernández Alonso
Probabilidad-Grado en Matemáticas (URJC)

Definición de variable aleatoria discreta

- **Definición**: Se dice que una variable aleatoria, X, es **discreta**, si para su función de distribución, F_X, existe un **conjunto numerable de puntos**, $\{k_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$, tal que

$$\sum_{n\in\mathbb{N}} [F_X(k_n) - F_X(k_n-)] = 1,$$

o, equivalentemente, tal que

$$\sum_{n\in\mathbb{N}} P(X = k_n) = 1.$$

Soporte de una variable aleatoria discreta

- Para describir una variable aleatoria discreta hay que especificar los valores que puede tomar y las probabilidades de que aparezca cada uno de ellos.

- Recordemos que se llama **soporte** o rango de una variable aleatoria X al conjunto de valores que puede tomar la variable. Denotamos dicho conjunto por S_X.

- El soporte de una variable aleatoria discreta viene dado por

$$S_X = \{k \in \mathbb{R} : P(X = k) > 0\}.$$

  o, lo que es lo mismo, por

$$S_X = \{k \in \mathbb{R} : F_X(k) - F_X(k-) > 0\}.$$

- Para las variables aleatorias discretas el soporte es siempre un conjunto **numerable**.

- Conociendo el soporte de una variable discreta quedan especificados los valores que ésta puede tomar.

Función de masa de probabilidad

- Se define la **función de masa de probabilidad** de una variable aleatoria X como la función, f_X, que indica la probabilidad de que X tome cada uno de sus valores, es decir,

$$\begin{aligned} f_X : \quad & \mathbb{R} \longrightarrow (0,1) \\ & k \longrightarrow f_X\,(k) = P\,(X = k) = P\,(\{\omega \in \Omega \text{ tal que } X(\omega) = k\,\}\,)\,. \end{aligned}$$

- Obviamente, si k no es uno de los valores que puede tomar X, entonces $f_X(k) = 0$, es decir,

$$k \notin S_X \Longrightarrow f_X(k) = P(X = k) = 0.$$

- La función de masa de probabilidad se puede representar gráficamente mediante un **diagrama de barras**.

Ejemplo: variable aleatoria discreta

- Consideremos de nuevo el triple lanzamiento de una moneda y la variable aleatoria X= "número de caras ".

- El soporte de la variable X

$$S_X = \{0, 1, 2, 3\}.$$

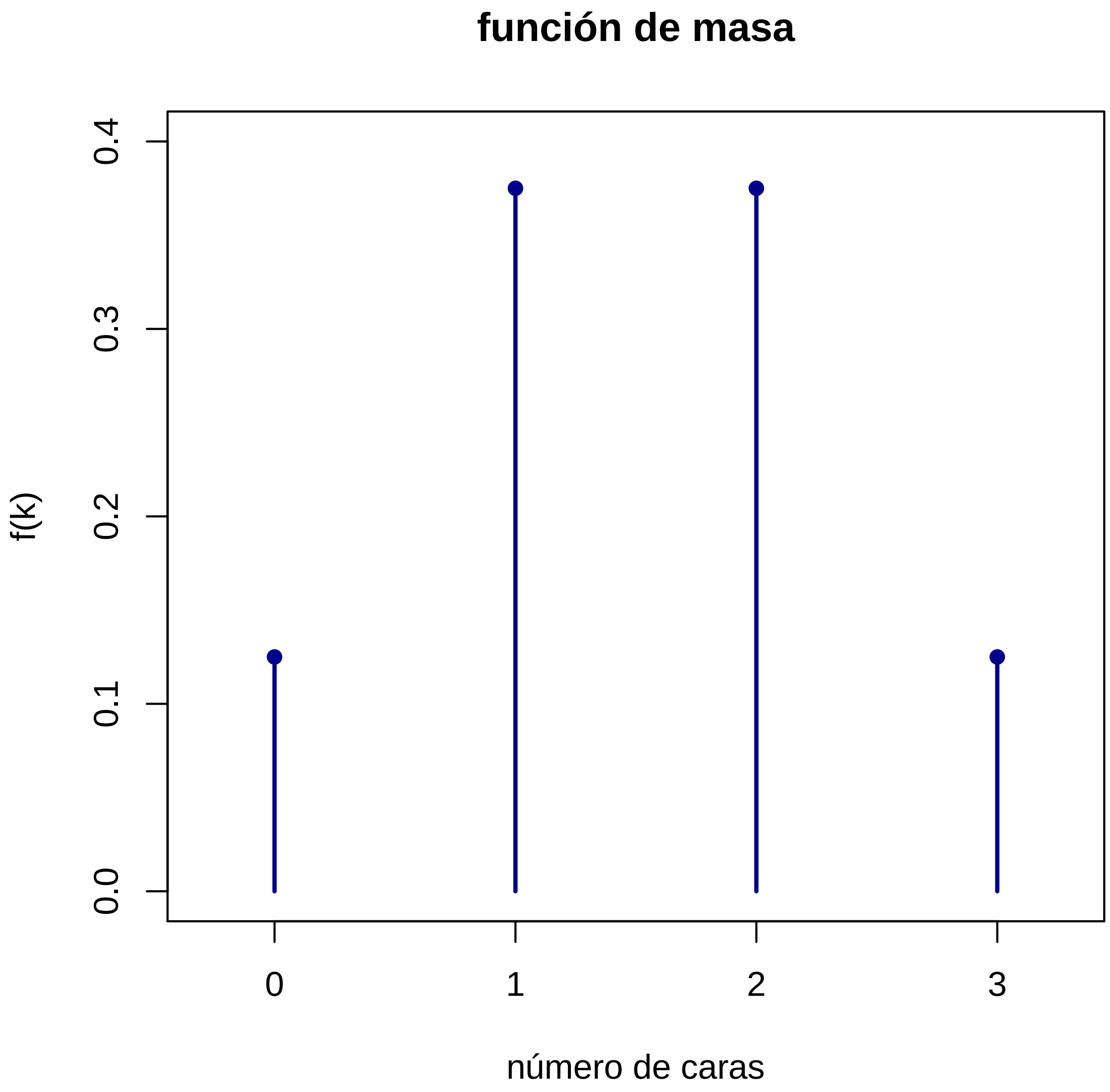
Puesto que S_X es un conjunto finito, X es una variable de tipo discreto.

- La función de masa de probabilidad de X es

$$\begin{aligned}
f(0) &= P_X(0) = P(X = 0) = P(\{XXX\}) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8},\\
f(1) &= P_X(1) = P(X = 1) = P(\{XXC, XCX, CXX\}) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8},\\
f(2) &= P_X(2) = P(X = 2) = P(\{XCC, CCX, CXC\}) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8},\\
f(3) &= P_X(3) = P(X = 3) = P(\{CCC\}) = \frac{1}{8},\\
f(k) &= 0 \text{ para } k \notin \{0, 1, 2, 3\}.
\end{aligned}$$

Ejemplo: función de masa (continuación)

- El diagrama de barras de esta función de masa es:



Notación matricial del soporte y la función de masa

- Para facilitar los cálculos, resulta útil expresar el soporte y la función de masa de una variable discreta mediante una matriz de dos filas. En la fila de arriba se coloca el soporte de la variable con sus elementos ordenados de menor a mayor, y en la fila de abajo la probabilidad de cada punto.

- Por ejemplo, para la variable aleatoria

$$X = \text{ número de caras}$$

en el triple lanzamiento de la moneda equilibrada, podemos expresar resumidamente su soporte (S_X) y su función de masa (f_X) mediante la siguiente matriz:

$$X \equiv \begin{pmatrix} 0 & 1 & 2 & 3 \\ \\ \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{1}{8} \end{pmatrix}$$

Propiedades de la función de masa

- La función de masa de probabilidad de una variable aleatoria discreta se caracteriza por dos propiedades, que son consecuencia directa de las propiedades de la probabilidad.

- **Proposición 1:**

1. La función de masa es **no negativa**, es decir,

$$f_X(k) \geq 0$$

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

2. La suma de la función de masa de todos los valores del soporte de una variable aleatoria discreta es 1, esto es,

$$\sum_{k \in S_X} f_X(k) = 1.$$

demostración:.........inmediata

Cálculo de F_X a partir de f_X

- Recordemos que la **función de distribución** de una variable aleatoria X es la función F_X que asigna a cada $t \in \mathbb{R}$ la probabilidad de que X tome un valor menor o igual que t:

$$\begin{aligned} F_X : \quad & \mathbb{R} \longrightarrow (0,1) \\ & t \longrightarrow F_X(t) = P(X \leq t) = P(\omega \in \Omega : X(\omega) \leq t) \end{aligned}$$

- En el caso discreto, la forma de calcular esta función de probabilidad acumulada es sumando la probabilidad de todos los puntos del soporte menores o iguales que t:

$$F_X(t) = P(X \leq t) = \sum_{k \in S_X, k \leq t} P(X = k) = \sum_{k \in S_X, k \leq t} f_X(k)$$

Repaso: función de distribución

- Recordemos cómo calcular la función de distribución de la variable aleatoria

$$X = \text{ número de caras}$$

en el triple lanzamiento de la moneda equilibrada.

Observemos que

$$\begin{aligned}
F_X(0) &= P(X=0) = f_X(0) = \frac{1}{8} = 0.125 \\
F_X(1) &= f_X(0) + f_X(1) = \frac{1}{8} + \frac{3}{8} = \frac{4}{8} = 0.5 \\
F_X(2) &= f_X(0) + f_X(1) + f_X(2) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} = \frac{7}{8} = 0.875 \\
F_X(3) &= f_X(0) + f_X(1) + f_X(2) + f_X(3) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = \frac{8}{8} = 1
\end{aligned}$$

Repaso: función de distribución (continuación)

- También podemos calcular la función de distribución en puntos que no están en el soporte de X. Por ejemplo

$$
\begin{aligned}
F_X(-1) &= P(X \leq -1) = 0 \\
F_X(0.8) &= P(X \leq 0.8) = f_X(0) = 0.125 \\
F_X(1.5) &= P(X \leq 1.5) = f_X(0) + f_X(1) = 0.5 \\
F_X(5.1) &= P(X \leq 5.1) = f_X(0) + f_X(1) + f_X(2) + f_X(3) = 1 \\
F_X(2.9) &= P(X \leq 2.9) = f_X(0) + f_X(1) + f_X(2) = 0.875 \\
&\vdots \\
&\vdots
\end{aligned}
$$

Repaso: función de distribución (continuación)

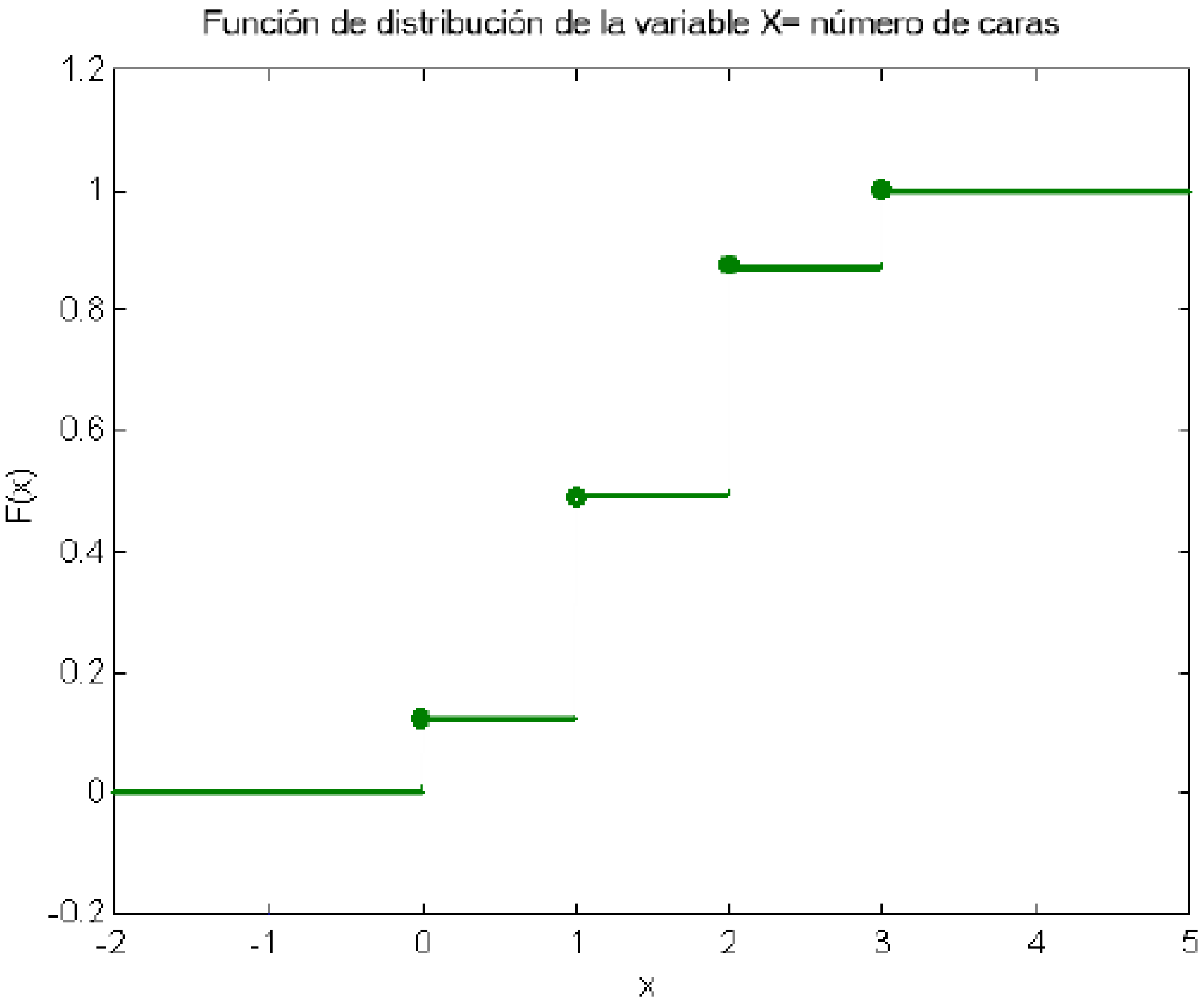
- La expresión completa de la función de distribución de la variable

$$X = \text{ número de caras}$$

es

$$F_X(t) = P(X \leq t) = \begin{cases} 0 & \text{si } t < 0, \\ 0.125 & \text{si } 0 \leq t < 1, \\ 0.5 & \text{si } 1 \leq t < 2, \\ 0.875 & \text{si } 2 \leq t < 3, \\ 1 & \text{si } t \geq 3. \end{cases}$$

Repaso: función de distribución (continuación)



Propiedades de F_X para variables discretas

- En el ejemplo anterior observamos que F_X:
  - Comienza en 0, ya que $F_X(t) = 0$ para todo $t < 0$.
  - Termina en 1, ya que $F_X(t) = 1$ para todo $t \geq 3$.
  - Es continua por la derecha, pero presenta discontinuidades en los puntos 0, 1, 2, y 3, que son los puntos del soporte de X.
  - Es monótona no decreciente.
  - Crece a saltos, los puntos de salto son los puntos del soporte, y la amplitud de cada salto es el valor de la función de masa en ese punto.
- Como probaremos más adelante, este es el aspecto que presentan en general las funciones de distribución de las variables aleatorias discretas.

Esperanza y varianza

de variables aleatorias discretas

Sonia Hernández Alonso
Probabilidad-Grado en Matemáticas (URJC)

Esperanza de una variable discreta

- **Definición**: La **esperanza**, o **media**, o **valor esperado** de una variable aleatoria aleatoria discreta X se define como

$$E(X) = \sum_{k \in S_X} k \times P(X = k)$$

o lo que es lo mismo

$$\boxed{E(X) = \sum_{k \in S_X} k \times f_X(k)}$$

- Si pensamos en la probabilidad como en una masa total de 1 repartida entre los puntos del soporte, $E(X)$ es el punto donde se encuentra el **centro de gravedad** o punto de equilibrio de la distribución de probabilidad de X.

- Es habitual denotar $E(X)$ por μ_X, o simplemente por μ.

Ejemplo: esperanza de una variable discreta

- Vamos a calcular el valor esperado de la variable aleatoria

$$X = \text{ número de caras}$$

en el triple lanzamiento de la moneda equilibrada.

- Recordemos que el soporte y la función de masa de X son

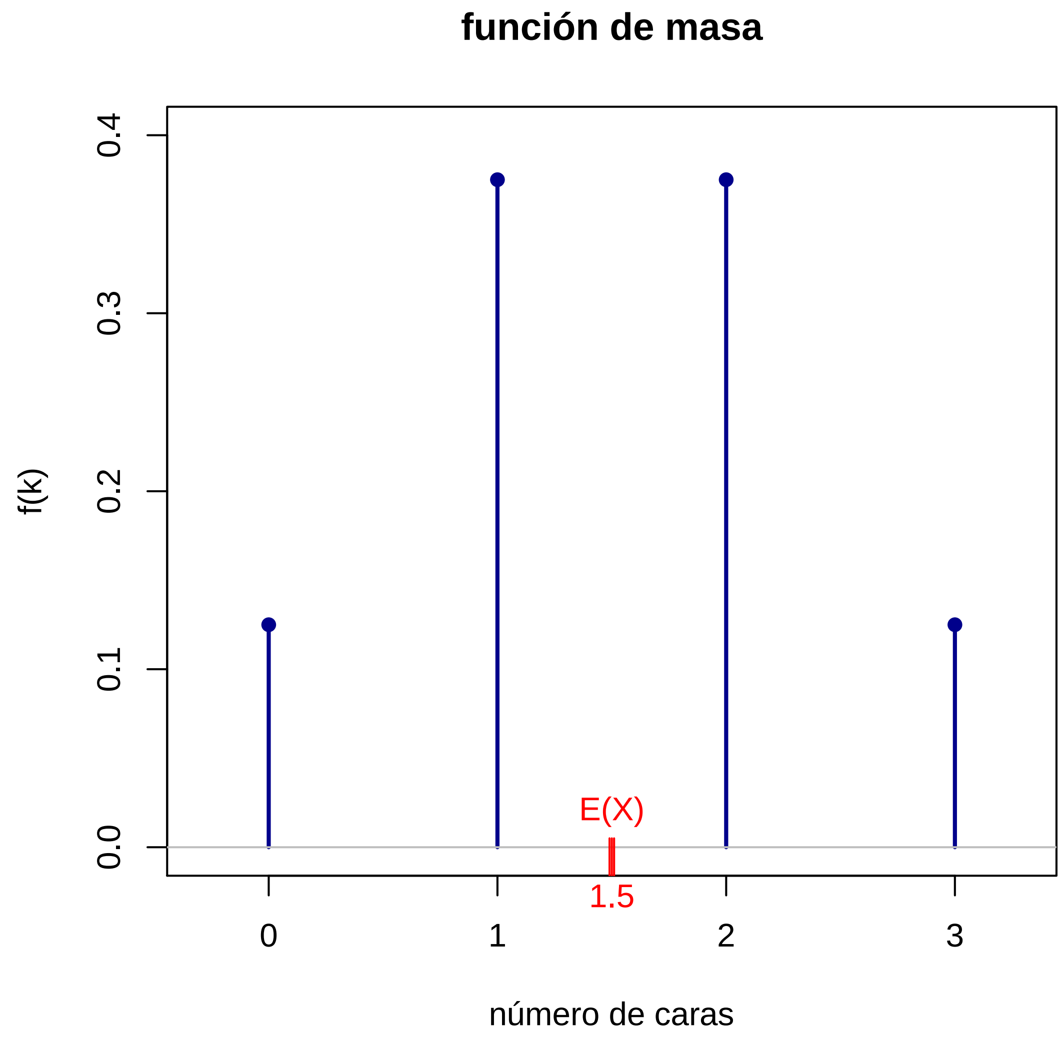
$$X \equiv \begin{pmatrix} 0 & 1 & 2 & 3 \\ \\ \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{1}{8} \end{pmatrix}$$

- Por consiguiente la esperanza de X es

$$E(X) = 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8} = \frac{12}{8} = 1.5 \text{ caras}$$

Gráfico: esperanza de una variable discreta

- Este valor esperado indica que, si se lanzan 3 monedas muchas veces, el número medio de caras sobre todos los lanzamientos será 1.5:



Esperanza de transformaciones lineales de variables

- **Proposición 2:** Si X es una variable aleatoria y $a, b \in \mathbb{R}$ son constantes, entonces se verifica

$$E(aX) = a\,E(X)$$

y

$$E(X+b) = E(X) + b$$

y por consiguiente

$$E(aX+b) = a\,E(X) + b$$

o, escrito abreviadamente,

$$\boxed{\mu_{aX+b} = a\mu_X + b}$$

demostración:.........pizarra

- La esperanza es por tanto un **operador lineal**.

Varianza de una variable aleatoria discreta

- **Definición**: La **varianza** de una variable aleatoria se define como

$$V(X) = E\left([X - E(X)]^2\right)$$

- $V(X)$ es una **medida de la dispersión** de X alrededor su centro de gravedad, $E(X)$.

- Es común denotar $V(X)$ por σ_X^2 , o simplemente por σ^2 :

$$\boxed{\sigma_X^2 = E\left([X - \mu_X]^2\right)}$$

- Para las variables aleatorias discretas, la varianza puede calcularse mediante la fórmula

$$\boxed{V(X) = \sum_{k \in S_X} (k - \mu_X)^2 f_X(k)}$$

- Nótese que las unidades de la varianza son el **cuadrado de las unidades** de la variable aleatoria.

Fórmula alternativa para la varianza

- **Proposición 3:** Para cualquier variable aleatoria X, una expresión alternativa para su varianza es

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

o, escrito abreviadamente,

$$\boxed{\sigma_X^2 = E(X^2) - \mu_X^2}$$

demostración:.........pizarra

Fórmula alternativa para la varianza: observaciones

- Por tanto, **la varianza de una variable aleatoria es la esperanza de su cuadrado menos el cuadrado de su esperanza**.

- Luego, para las variables discretas, la varianza también puede calcularse como

$$\boxed{\sigma_X^2 = \sum_{k \in S_X} k^2 f_X(k) - \mu_X^2}$$

- Habitualmente el cálculo de la varianza resulta más sencillo si se utiliza esta segunda definición.

Ejemplo: varianza de una variable discreta

- Calculemos la varianza de la variable aleatoria

$$X = \text{ número de caras}$$

en el triple lanzamiento de la moneda equilibrada:

$$X \equiv \begin{pmatrix} 0 & 1 & 2 & 3 \\ \\ \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{1}{8} \end{pmatrix}$$

$$E(X) = 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8} = \frac{12}{8} = 1.5 \text{ caras}$$

$$E\left(X^2\right) = 0^2 \times \frac{1}{8} + 1^2 \times \frac{3}{8} + 2^2 \times \frac{3}{8} + 3^2 \times \frac{1}{8} = \frac{24}{8} = 3 \text{ caras}^2$$

y por tanto la varianza de esta variable es

$$V(X) = E(X^2) - E^2(X) = 3 - 1.5^2 = \mathbf{0.75} \textbf{ caras}^2$$

Varianza de transformaciones lineales de variables

- **Proposición 4:** Si X es una variable aleatoria y $a, b \in \mathbb{R}$ son constantes, entonces se verifica

$$V(aX + b) = a^2 V(X)$$

o, escrito abreviadamente,

$$\boxed{\sigma^2_{aX+b} = a^2 \sigma^2_X}$$

demostración:.........pizarra

Ejemplo: varianza transformaciones lineales

- **Ejemplo:** Si la varianza de una variable aleatoria X es

$$V(X) = 9$$

y definimos

$$Y = 10X + 50,$$

entonces

$$V(Y) = V(10X + 50) = 10^2\, V(X) = 100 \times 9 = 900$$

Desviación típica de una variable aleatoria

- La **desviación típica** de una variable aleatoria X, σ_X, es la raiz cuadrada de su varianza, esto es,

$$\sigma_X = \sqrt{\sigma_X^2} = \sqrt{V(X)}$$

- Las unidades de la varianza son el cuadrado de las unidades en las que esté medida la variable aleatoria. En cambio **la desviación típica tiene las mismas unidades que la variable aleatoria** a la que corresponde.

- **Ejemplo:** la desviación típica de la variable aleatoria

$$X = \text{ número de caras}$$

en tres lanzamientos de una moneda equilibrada, es

$$\sigma_X = \sqrt{V(X)} = \sqrt{0.75} = \mathbf{0.866 \text{ caras}}$$

Ejemplo: número de muertes de abejas obreras

- **Ejercicio 5:** Se ha comprobado que el número muertes de obreras que se producen diariamente en una colmena de abejas (*Apis mellifera*) es una variable aleatoria, M, con función de probabilidad:

$$M \equiv \begin{pmatrix} 6 & 7 & 8 & 9 & 10 \\ \\ 0.10 & 0.15 & 0.25 & 0.30 & 0.20 \end{pmatrix}$$

1. Hallar la probabilidad de que ayer muriesen al menos 8 obreras.
2. Calcular la probabilidad de que mañana mueran más de 8 obreras.
3. ¿Cúal es la probabilidad de que la cantidad de defunciones que se produzcan dentro de tres días sea un número impar?
4. Se sabe que anteayer fallecieron al menos 8 abejas obreras; ¿cúal es la probabilidad de que el número de defunciones fuese impar?
5. ¿Cuál es el número esperado de defunciones diarias?
6. Hallar la desviación típica del número de defunciones por día.

Solución al Ejercicio 5

1.

$$P(M \geq 8) = P(M = 8) + P(M = 9) + P(M = 10) = 0.25 + 0.3 + 0.2 = \mathbf{0.75}$$

2.

$$P(M > 8) = P(M = 9) + P(M = 10) = 0.3 + 0.2 = \mathbf{0.5}$$

3.

$$P(M \text{ impar}) = P(M = 7) + P(M = 9) = 0.15 + 0.3 = \mathbf{0.45}$$

4.

$$P(M \text{ impar} | M \geq 8) = \frac{P([M \text{ impar}] \cap [M \geq 8])}{P(M \geq 8)} = \frac{P(M = 9)}{P(M \geq 8)} = \frac{0.3}{0.75} = \mathbf{0.4}$$

Solución al Ejercicio 5 (continuación)

5. El número esperado de defunciones diarias es

   $$E(M) = 6 \times 0.1 + 7 \times 0.15 + 8 \times 0.25 + 9 \times 0.3 + 10 \times 0.2 = \mathbf{8.35 \text{ muertes}}$$

   Este valor medio aparace representado en la transparencia siguiente sobre el diagrama de barras de la función de probabilidad.

6. La varianza del número de defunciones por día viene dada por

   $$V(M) = E(M^2) - E^2(M)$$

   Se tiene que

   $$E(M^2) = 36 \times 0.1 + 49 \times 0.15 + 64 \times 0.25 + 81 \times 0.3 + 100 \times 0.2 = 71.25$$

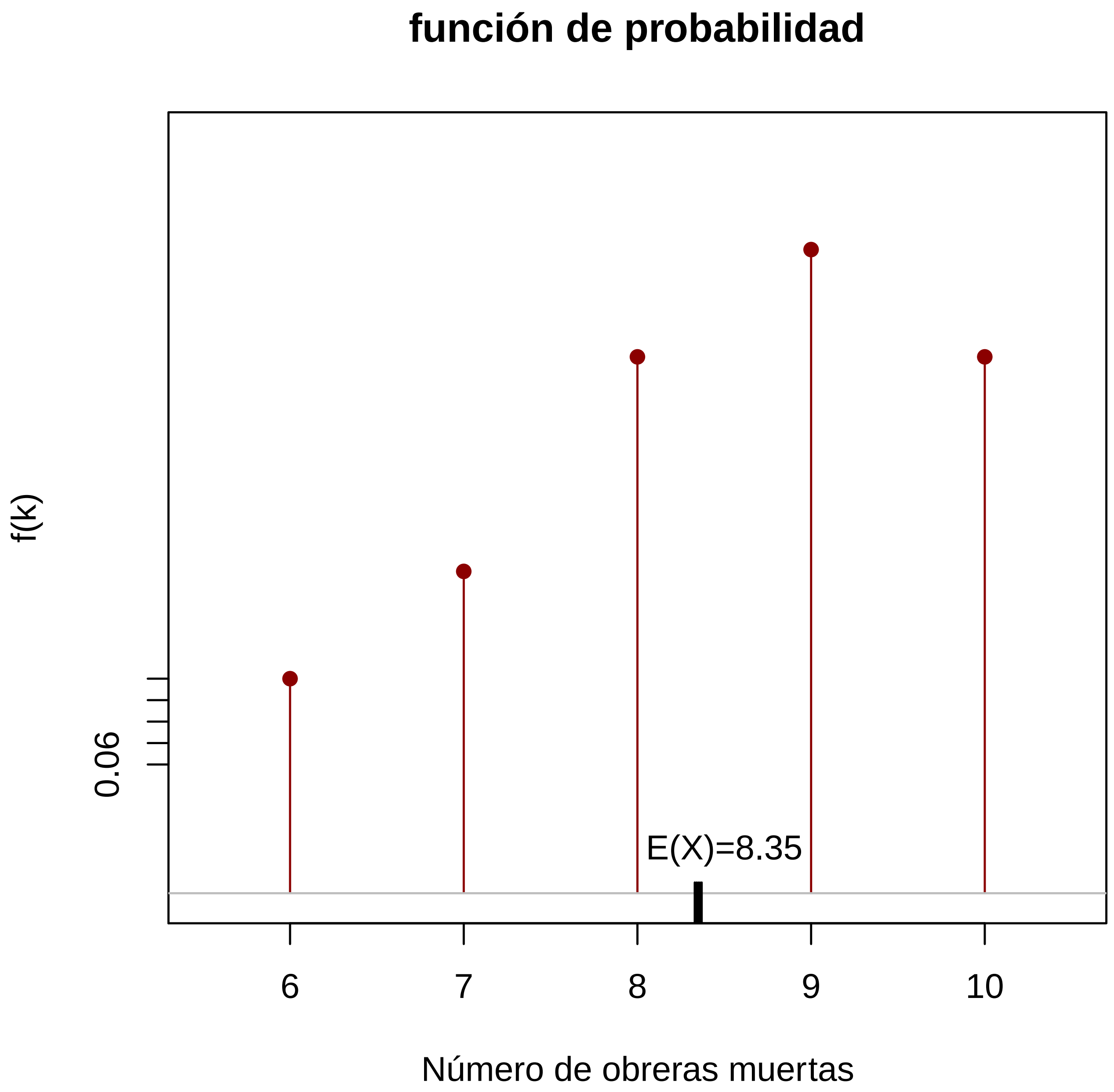
   y por tanto

   $$V(M) = 71.25 - 8.35^2 = 1.5275$$

   Luego la desviación típica del número de muertes diarias es

   $$\sigma_M = \sqrt{1.5275} = \mathbf{1.24 \text{ muertes}}$$

Gráfico: número de muertes de abejas obreras



Modelos discretos especiales

Sonia Hernández Alonso
Probabilidad-Grado en Matemáticas (URJC)

Familias de modelos discretos

- Algunos modelos de variables aleatorias discretas se repiten con mucha frecuencia.
- En este tema vamos a analizar algunas de las distribuciones discretas que se repiten con mayor frecuencia.
- Las distribuciones binomial y geométrica están relacionadas con el llamado **Proceso de Bernoulli**, que consiste repeticiones independientes de un experimento con dos resultados posibles, a los que nos referiremos como éxito y fracaso, en los que la probabilidad de éxito es la misma en todas las repeticiones.
- La distribución de Poisson está relacionada con otro tipo de proceso estocástico conocido como **Proceso de Poisson**.

Distribución binomial

Sonia Hernández Alonso
Probabilidad-Grado en Matemáticas (URJC)

Procesos de Bernoulli

- Consideremos un experimento aleatorio con **dos resultados posibles**, a los que nos referiremos como **éxito** y **fracaso**. Llamemos
  - $p \in (0, 1)$ a la **probabilidad de éxito**,
  - $q = 1 - p \in (0, 1)$ a la **probabilidad de fracaso**.
- La repetición de forma independiente de este experimento constituye lo que se conoce como un **proceso de Bernoulli**. Estos procesos permiten **modelar muchos fenomenos** de la vida real.
- La distribución binomial cuenta el número de éxitos ocurridos en n realizaciones, independientes y con la misma probabilidad de éxito, de un experimento de Bernoulli.

Situaciones que modela la distribución binomial

- La situación que modela la distribución binomial es la siguiente:
  - Se considera un experimento con dos resultados posibles: éxito o fracaso.
  - El experimento se repite, de manera independiente, n veces.
  - La probabilidad de éxito, p, es la misma en cada repetición.
  - La variable de interés es

$$X = \textbf{número de éxitos en las } n \textbf{ repeticiones}$$

- La distribución que sigue la variable aleatoria X recibe el nombre de **distribución binomial con parámetros n y p**.
- Lo denotaremos

$$\boxed{X \sim Bin(n, p)}$$

Ejemplos de la distribución binomial

- La variable aleatoria

$$X = \text{número de caras obtenidas en 3 lanzamientos de una moneda}$$

  sigue una distribución $X \sim Bin(3, 1/2)$.

- La variable aleatoria

$$T = \text{número de 3's obtenidos en 5 lanzamientos de un dado}$$

  sigue una distribución $T \sim Bin(5, 1/6)$.

- El 15 % de la población activa de cierta comarca de gran tamaño trabaja por cuenta propia, el 60 % lo hace por cuenta ajena, y el resto son funcionarios. La variable aleatoria

$$F = \text{número de funcionarios en una muestra de 4 personas seleccionadas al azar}$$

  sigue una distribución $F \sim Bin(4,\ 0.25)$.

Ejemplo: nº de 3's en 5 lanzamientos de un dado

- Supongamos que se lanza un dado equilibrado cinco veces y queremos analizar el número de 3's que aparecen en total.
- Llamemos T a número de 3's obtenidos entre los cinco lanzamientos.
- La distribución de esta variable aleatoria es binomial con 5 repeticiones y probabilidad de éxito 1/6:

$$T \sim Bin\left(n = 5, p = \frac{1}{6}\right)$$

- ¿Cuáles son los valores que puede tomar T?

  Es claro que

$$S_T = \{0, 1, 2, 3, 4, 5\}$$

- Y, ¿cuál es la probabilidad de cada uno de estos valores? Es decir, ¿cuál es la función de probabilidad de T?

Ejemplo: nº de 3's en 5 lanztos (continuación)

- Observemos que

$$P(T=0) = \left(\frac{5}{6}\right)^5 = 0.4019$$

$$P(T=1) = 5 \times \frac{1}{6} \times \left(\frac{5}{6}\right)^4 = 0.4019$$

$$P(T=2) = \binom{5}{2} \times \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^3 = 10 \times \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^3 = 0.1608$$

$$P(T=3) = \binom{5}{3} \times \left(\frac{1}{6}\right)^3 \times \left(\frac{5}{6}\right)^2 = 10 \times \left(\frac{1}{6}\right)^3 \times \left(\frac{5}{6}\right)^2 = 0.0321$$

$$P(T=4) = \binom{5}{4} \times \left(\frac{1}{6}\right)^4 \times \frac{5}{6} = 5 \times \left(\frac{1}{6}\right)^4 \times \frac{5}{6} = 0.0032$$

$$P(T=5) = \left(\frac{1}{6}\right)^5 = 0.0001$$

URJC

Ejemplo: nº de 3's en 5 lanztos (continuación)

- Nótese que todas las probabilidades anteriores responden a la fórmula

$$P(T=k)=\binom{5}{k}\times\left(\frac{1}{6}\right)^{k}\times\left(\frac{5}{6}\right)^{5-k}$$

- De forma análoga a la de este ejemplo, se puede razonar cómo son el soporte y la función de probabilidad de cualquier distribución binomial.

Soporte y función de masa de $X \sim Bin(n,p)$

- Sea X una variable aleatoria con distribución binomial, con n repeticiones y probabilidad de éxito p,

$$X \sim Bin(n,p)$$

Entonces:

• El **soporte** de X es

$$S_X = \{0, 1, 2, 3, \ldots n-1, n\}$$

• La **función de masa de probabilidad** de la variable aleatoria X es

$$f_X(k) = P\left(X = k\right) = \binom{n}{k} \times p^k \times (1-p)^{n-k} = \binom{n}{k} \times p^k \times q^{n-k}$$

- **Ejercicio 6:** Demostrar que f_X cumple las propiedades de las funciones de masa de probabilidad.

resolución:.........pizarra

Esperanza de la distribución binomial

- **Proposición 5:** Sea X una variable aleatoria con distribución binomial,

$$X \sim Bin(n,p).$$

Entonces su esperanza es

$$\boxed{E\,(X) = \sum_{k=0}^{n} \binom{n}{k} \times k \times p^k \times (1-p)^{n-k} = n \times p}$$

demostración:.........pizarra

Varianza de la distribución binomial

- **Proposición 6:** Sea X una v. al. con distribución binomial,

$$X \sim Bin(n, p).$$

Entonces su varianza es

$$V(X) = \sum_{k=0}^{n} \binom{n}{k} \times k^2 \times p^k \times (1-p)^{n-k} - n^2 p^2 = n \times p \times (1-p)$$

o lo que es lo mismo

$$\boxed{V(X) = n \times p \times q}$$

Ejemplo: esperanza y varianza de binomial

- Para la variable T que recoge el número de 3's en 5 lanzamientos de un dado, cuya distribución es

$$T \sim Bin(5, \frac{1}{6}),$$

se tiene que

$$E(T) = 5 \times \frac{1}{6} = 0.8333,$$

$$V(T) = 5 \times \frac{1}{6} \times \frac{5}{6} = 0.6944.$$

Ejemplo: número de piezas defectuosas

- **Ejercicio 7:** El porcentaje de piezas defectuosas que produce una máquina es del 25 %. Para un control rutinario se han seleccionado de forma aleatoria 8 de las piezas producidas.

  1. ¿Cuál es la probabilidad de que al menos una de las piezas seleccionadas sea defectuosa?
  2. ¿Cuál es la probabilidad de que como mucho dos de las piezas sean defectuosas?
  3. ¿Cuál es la probabilidad de que tres de las piezas seleccionadas sean defectuosas y el resto sean correctas?
  4. ¿Cuantas de las piezas seleccionadas se espera que tengan defectos?
  5. ¿Cuál es la varianza del número de piezas defectuosas entre las 8 seleccionadas?

Solución al Ejercicio 7

- Llamemos D al número de piezas defectuosas de entre las 8 seleccionadas para el control.

  Tenemos que

  $$D \sim Bin\,(8, 0.25)$$

  Por tanto:

  1.

  $$P(D \geq 1) = 1 - P(D = 0) = 1 - 0.75^8 = 1 - 0.1001 = \mathbf{0.8999}$$

  2.

  $$\begin{aligned} P(D \leq 2) &= P(D = 0) + P(D = 1) + P(D = 2) \\ &= 0.75^8 + 8 \times 0.25 \times 0.75^7 + \binom{8}{2} \times 0.25^2 \times 0.75^6 \\ &= \mathbf{0.6785} \end{aligned}$$

Solución al Ejercicio 7 (continuación)

3. La probabilidad de que sean exactamente tres las piezas defectuosas es

$$P(D = 3) = \binom{8}{3} \times 0.25^3 \times 0.75^5 = 56 \times 0.25^3 \times 0.75^5 = \mathbf{0.2076}$$

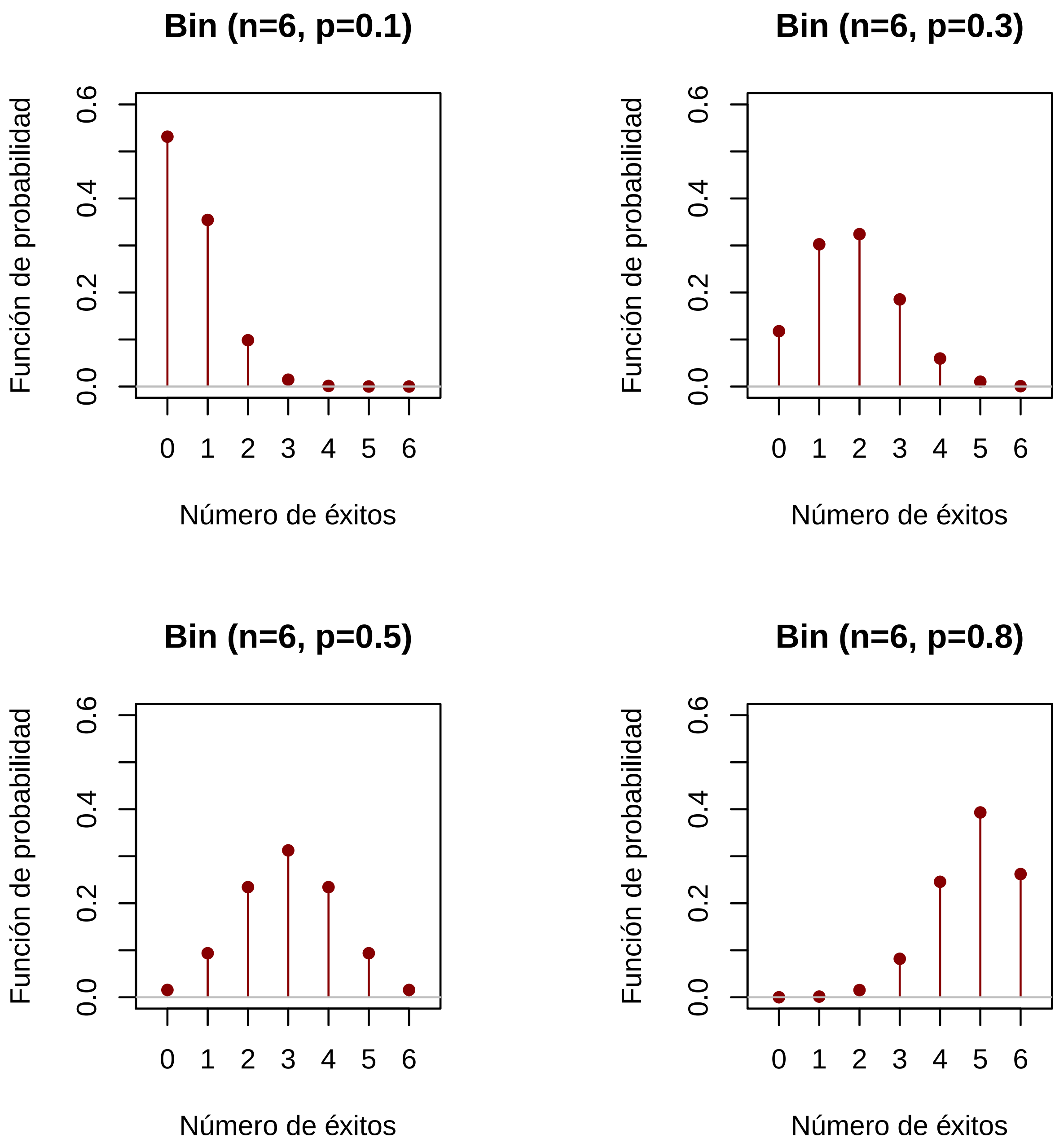
4. La cantidad de piezas seleccionadas que se espera que tengan algún defecto es

$$E(D) = 8 \times 0.25 = \mathbf{2\ \textbf{piezas}}$$

5. La varianza del número de piezas que tienen algún defecto de entre las 8 seleccionadas para el control es

$$V(D) = 8 \times 0.25 \times 0.75 = \mathbf{1.5\ \textbf{piezas}^2}$$

¿Cómo cambia la distribución al variar p?



Distribución geométrica

Sonia Hernández Alonso
Probabilidad-Grado en Matemáticas (URJC)

Distribución geométrica

URJC

- La distribución **geométrica** modela el número de veces que es necesario repetir un experimento de Bernoulli hasta obtener el primer éxito.

- Ahora, en lugar de estar interesados por el número de éxitos en una cantidad fija de repeticiones n estamos interesados en predecir el instante en el que se produce el primer éxito.

- La situación a modelar es la siguiente:

  - Se consideran repeticiones **independientes** de un experimento.
  - El experimento tiene dos resultados posibles: éxito o fracaso.
  - La probabilidad de éxito, p, es la misma en cada repetición.
  - La variable de interés es

    $X =$ **"número de repeticiones del experimento hasta obtener el primer éxito"**.

- Diremos que la variable aleatoria X sigue una distribución **Geométrica** con parámetro p, y lo denotaremos por $X \sim Ge(p)$.

Ejemplos de la distribución geométrica

- La variable aleatoria

$$L = \text{número de lanzamientos de un dado equilibrado hasta obtener el primer cinco}$$

  sigue una distribución $Ge(1/6)$.

- La variable aleatoria

$$M = \text{número de lanzamientos de una moneda equilibrada hasta obtener la primera cara}$$

  sigue una distribución $Ge(1/2)$.

- En una fábrica se analizan todas las piezas fabricadas con el fin de encontrar las posibles piezas defectuosas. La probabilidad de que una pieza esté en perfecto estado es 0.95.

  La variable aleatoria

$$D = \text{número de piezas analizadas hasta encontrar la primera defectuosa}$$

  sigue una distribución $Ge(0.05)$.

Función de masa de la distribución geométrica

- El **soporte** de cualquier variable X con distribución geométrica es

$$S_X = \mathbb{N}^+ = \{1, 2, 3, 4, \ldots\ldots\}$$

- La **función de masa de probabilidad** de una variable geométrica con probabilidad de éxito p es

$$f_X(k) = \begin{cases} q^{k-1}\, p & \text{para } k = 1, 2, 3, 4, \ldots\ldots \\ \\ 0 & \text{para } k \notin \mathbb{N}^+. \end{cases}$$

- **Ejercicio 8:** Demostrar que f_X cumple las propiedades de las funciones de masa de probabilidad.

  resolución:.........pizarra

Esperanza y varianza de la distribución geométrica

- **Proposición 7:** Sea X una variable aleatoria con distribución geométrica,

$$X \sim Ge(p).$$

Entonces su **esperanza** es

$$E(X) = \sum_{k=1}^{\infty} k \times q^{k-1} \times p = \frac{1}{p}.$$

demostración:.........pizarra

- **Proposición 7*:** Sea X una variable aleatoria con distribución geométrica,

$$X \sim Ge(p).$$

Entonces su **varianza** es

$$V(X) = \sum_{k=1}^{\infty} k^2 \times q^{k-1} \times p - \frac{1}{p^2} = \frac{1-p}{p^2} = \frac{q}{p^2}.$$

Forma de la distribución geométrica

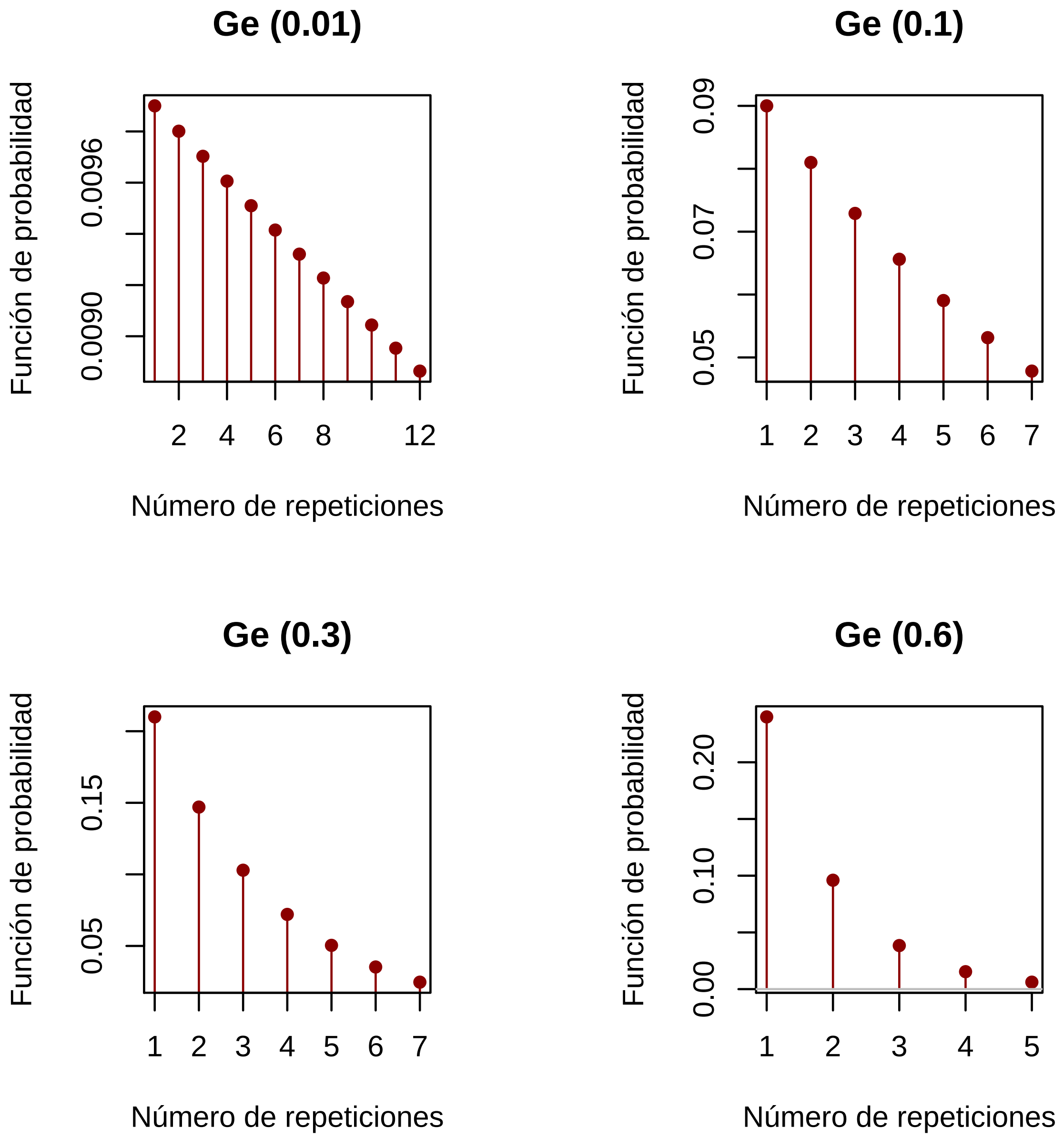
- La función de masa de la distribución geométrica,

$$f_X(k) = P(X = k) = (1 - p)^{k-1}\, p = q^{k-1} p,$$

  alcanza su máximo en $k = 1$ y decrece con k.

- El decrecimiento es más rápido cuanto mayor sea p.

Gráficos de la distribución geométrica



Ejemplo: distribución geométrica

- **Ejercicio 9:** Una pareja decide que tendrán descendientes hasta que se produzca el nacimiento de la primera niña.

  Suponiendo que, en cada parto, la probabilidad de que nazca un niño es 0.6,

  1. Hallar la probabilidad de que tengan en total 3 descendientes.
  2. Calcular la probabilidad de que tengan más de 4 hij@s.
  3. Supongamos que la pareja ha tenido ya dos hijos varones.
     a) Hallar la probabilidad de que tengan en total 5 descendientes.
     b) Calcular la probabilidad de que (en total) tengan más de 6 hij@s.

Solución al Ejercicio 9

- La distribución de variable aleatoria

  $N=$ nº descendientes hasta la primera niña,

  es $N \sim Ge(0.4)$. Por consiguiente:

  1. $P(N=3)=0.6^2\times 0.4=0.144$.
  2. $P(N>4)=0.6^4=0.1296$.
  3. Si se sabe que la pareja ha tenido ya dos hijos varones,

     a) $P(N=5|N>2)=\dfrac{P(N=5)}{P(N>2)}=\dfrac{0.6^4\times 0.4}{0.6^2}=0.6^2\times 0.4=0.144$.

     Nótese que este valor coincide con la probabilidad de que la pareja tenga un total de 3 descendientes.

     b) $P(N>6|N>2)=\dfrac{P(N>6)}{P(N>2)}=\dfrac{0.6^6}{0.6^2}=0.6^4=0.1296$.

     Nótese que este valor coincide con la probabilidad de que la pareja tenga más de 4 hij@s.

Falta de memoria de la distribución geométrica

- En el ejemplo anterior ilustra una de las propiedades más importantes de la distribución geométrica: para este tipo de variables aleatorias, **las probabilidades sobre lo que ocurra en el futuro no dependen de lo que haya ocurrido en el pasado**.

- **Proposición 8:** Si $X \sim Ge(p)$, y $k, a \in \mathbb{N}^+$, entonces,

  1. $P(X = k + a | X > a) = P(X = k)$
  2. $P(X > k + a | X > a) = P(X > k)$
  3. $P(X \geq k + a | X > a) = P(X \geq k)$
  4. $P(X \leq k + a | X > a) = P(X \leq k)$
  5. $P(X < k + a | X > a) = P(X < k)$

- Se dice por ello que **la distribución geométrica no tiene memoria**.

- **La distribución geométrica es la única de las distribuciones discretas que verifica esta propiedad de falta de memoria**.

Demostración de la Proposición 8 (apartado 1)

- A modo de ejemplo demostraremos los apartados 1 y 4:

(1) Sabemos que

$$P(X = k) = q^{k-1} \times p,$$

y por otra parte tenemos que

$$\begin{aligned} P(X = k + a|X > a) &= \frac{P[(X = a + k) \cap (X > a)]}{P(X > a)} \\ &= \frac{P(X = a + k)}{P(X > a)} \\ &= \frac{q^{a+k-1} \times p}{q^a} = \frac{q^a \times q^{k-1} \times p}{q^a} \\ &= q^{k-1} \times p \end{aligned}$$

Luego, en efecto, se verifica

$$P(X = k + a|X > a) = P(X = k) = q^{k-1} \times p.$$

c.q.d.

Demostración de la Proposición 8 (apartado 4)

URJC

(4) Tenemos por una parte que,

$$P(X \leq k) = 1 - P(X > k) = 1 - q^k,$$

y por otra que

$$\begin{aligned} P(X \leq k + a | X > a) &= \frac{P[(X \leq k + a) \cap (X > a)]}{P(X > a)} \\ &= \frac{P[(a < X \leq k + a}{P(X > a)} \\ &= \frac{P(X > a) - P(X > k + a)}{P(X > a)} \\ &= \frac{q^a - q^{k+a}}{q^a} = \frac{q^a \times (1 - q^k)}{q^a} \\ &= 1 - q^k \end{aligned}$$

Luego

$$P(X \leq k + a | X > a) = P(X \leq k) = 1 - q^k$$

c.q.d.

Distribución de Poisson

Siméon Denis Poisson (1781-1840; Francia)

Sucesos puntuales a lo largo del tiempo

- En ocasiones nos encontramos con variables que representan el número de sucesos que ocurren en un determinado periodo de tiempo, como por ejemplo el número de cebras que acuden a beber a un arroyo en una hora, el número de visitas diarias a una página web, o el número averías anuales de un ascensor.

- Muchas veces estos sucesos van apareciendo a lo largo del tiempo de manera independiente y con una intensidad constante.

- La distribución de **Poisson** proporciona un modelo adecuado para el **número de ocurrencias** independientes de un suceso en situaciones en las que la intensidad de ocurrencias se mantiene estable.

- En muchas ocasiones los que se quiere modelar son **procesos de llegada**, como por ejemplo las llegadas de e-mails a un servidor, o de cientes a una sucursal bancaria.

- Esta distribución depende de un único parámetro: el número medio de ocurrencias del suceso en ese intervalo de tiempo. Denotaremos por λ dicho parámetro.

Soporte de la distribución de Poisson

- Llamemos X al número de apariciones del suceso (por ejemplo llegadas de coches a una gasolinera, o de trabajos a un servidor) en el intervalo de tiempo que estemos considerando(por ejemplo 1 hora).
- Como es natural, el soporte de una variable de este tipo es

$$S_X = \mathbb{N} = \{0, 1, 2, 3, \ldots\ldots$$

Función de masa de la distribución de Poisson

- La distribución de Poisson se obtiene como límite de la binomial cuando el número repeticiones (n), tiende a infinito, la probabilidad de éxito (p) tiende a cero y el número medio de éxitos (np) se estabiliza alrededor de un número λ.

- Calculando dicho límite se obtiene:

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \qquad \text{para } k = 0, 1, 2, \ldots$$

- Esta es la función de masa de probabilidad de una **distribución de Poisson** con parámetro λ.

- Abreviadamente lo expresaremos $X \sim Po(\lambda)$

¿Es esta una función de masa?

- **Ejercicio 10:** Demostrar que f_X cumple las propiedades de las funciones de masa de probabilidad.

  resolución:.........pizarra

  Ayuda: Recordar que

$$\sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x$$

¿Qué indica el parámetro λ?

- El **parámetro** λ de la distribución de Poisson es el número medio (o esperado) de ocurrencias del suceso en el periodo considerado.

- λ es siempre un número positivo,

$$\lambda > 0,$$

  y mide la **intensidad** con la que se producen las ocurrencias del suceso en el intervalo.

- Esta intensidad cambia en función de la longitud del intervalo que se esté considerando, y es proporcional a la longitud del intervalo.

- Para calcular probabilidades sobre una distribución de Poisson es importante fijarse antes de nada en el intervalo de tiempo que se va a considerar.

- **Observación:** La distribución de Poisson aparece como límite de la binomial cuando n tiene a infinito y p tiende a 0 de manera que el producto $\lambda = np$ de mantiene constante.

Esperanza y varianza de la distribución de Poisson

- Evidentemente, la **esperanza** de una variable $X \sim Po(\lambda)$ es λ:

$$E\left(X\right) = \lambda$$

- Además, puede probarse que su **varianza** es

$$V\left(X\right) = \lambda$$

- Observamos que se verifica

$$\boxed{E(X) = V(X) = \lambda}$$

- Las variables aleatorias de la familia de Poisson son las únicas en las que la esperanza y la varianza coinciden.

Ejemplo: nº de ataques de leonas

- **Ejercicio 11:** El número de ataques de leonas que sufre una manada de gacelas sigue una distribución de Poisson con una media de 2 ataques diarios.

1. Calcular la probabilidad de que se produzcan 3 ataques en un día.
2. Hallar la probabilidad de que no ocurra ningún ataque en 12 horas.
3. ¿Cuál es la probabilidad de que en dos días consecutivos haya menos de 3 ataques?
4. Determinar la probabilidad de que ocurra algún ataque en 36 horas.
5. ¿Cuál es la probabilidad de que en una semana se produzcan 12 ataques?
6. Si se acaba de producir un ataque, ¿cuál es la probabilidad de que transcurran más de 24 horas hasta el siguiente?

Solución al Ejercicio 11

1. Llamemos N_1 al número de ataques de leonas que sufre la manada a lo largo de un día.

   Sabemos que N_1 sigue una distribucion de Poisson y también que su media o esperanza es $E(N_1) = 2$, luego

   $$N_1 \sim Po\,(2)$$

   Por tanto, la probabilidad de que se produzcan 3 ataques en un día es

   $$P(N_1 = 3) = e^{-2} \times \frac{2^3}{3!} = \mathbf{0.1804}$$

Solución al Ejercicio 11 (continuación)

2. En este caso la probabilidad no se refiere a un día, sino a medio.

   La distribución del número de ataques en medio día es diferente que en un día entero. Sigue también el modelo de Poisson pero la intensidad de los ataques (λ) es la mitad.

   Es decir, si denotamos por $N_{\frac{1}{2}}$ el número de ataques de leonas en medio día, se tiene que

$$N_{\frac{1}{2}} \sim Po(1)$$

   Por tanto, la probabilidad de que no se produzca ningún ataque en 12 horas es

$$P(N_{\frac{1}{2}} = 0) = e^{-1} \times \frac{1^0}{0!} = e^{-1} = \mathbf{0.3679}$$

Solución al Ejercicio 11 (continuación)

3. La distribución del número de ataques en dos días es

$$N_2 \sim Po(4)$$

Luego la probabilidad de que se produzcan menos de 3 ataques en dos días es

$$\begin{aligned} P(N_2 < 3) &= P(N_2 = 0) + P(N_2 = 1) + P(N_2 = 2) \\ &= e^{-4} \times \frac{4^0}{0!} + e^{-4} \times \frac{4^1}{1!} + e^{-4} \times \frac{4^2}{2!} \\ &= \mathbf{0.2381} \end{aligned}$$

4. La distribución del número de ataques en 36 horas (día y medio) es

$$N_{1.5} \sim Po(3),$$

y por tanto la probabilidad de que se produzca algún ataque en 36 horas (es decir, al menos 1 ataque en un día y medio) es

$$P(N_{1.5} \geq 1) = 1 - P(N_{1.5} = 0) = 1 - e^{-3} \times \frac{3^0}{0!} = 1 - e^{-3} = \mathbf{0.9502}$$

Solución al Ejercicio 11 (continuación)

5. La distribución del número de ataques en siete días es

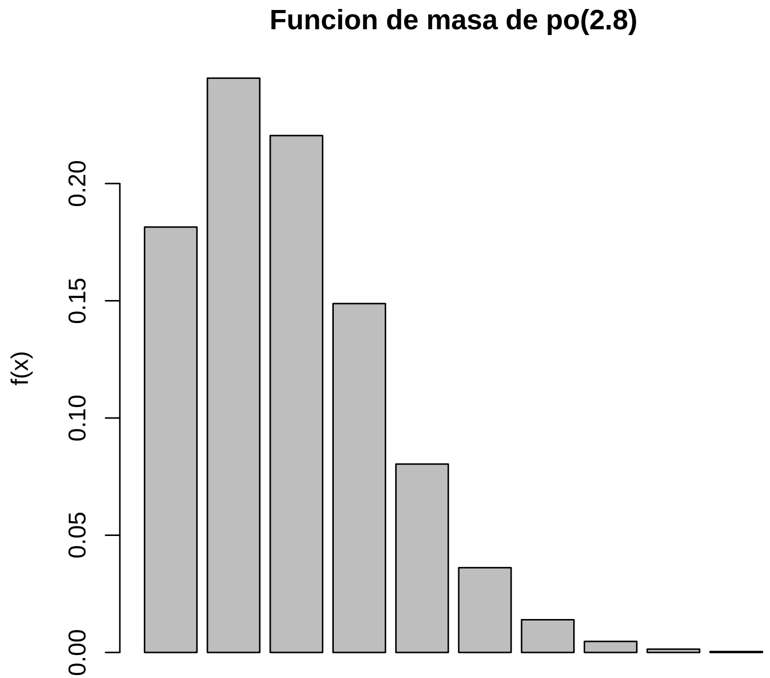
$$N_7 \sim Po(14)$$

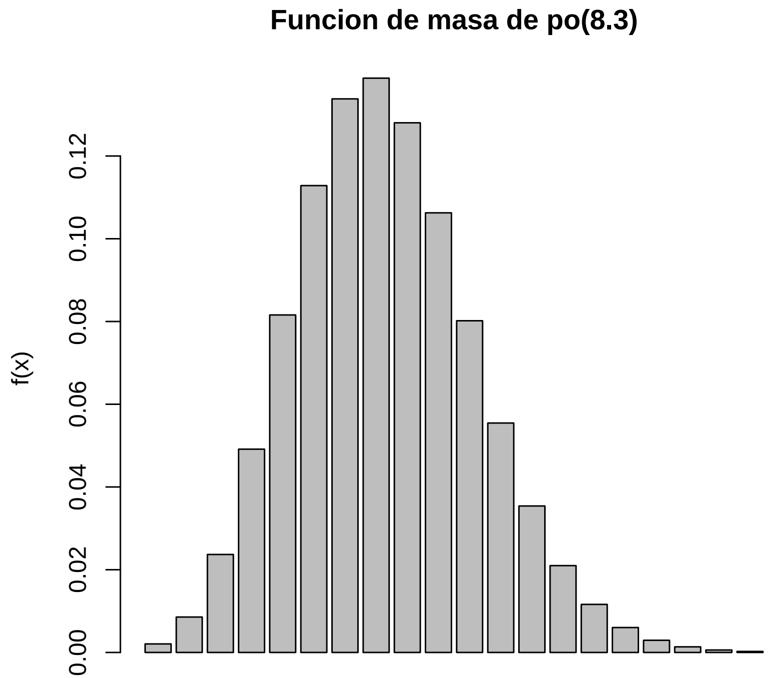
En consecuencia, la probabilidad de que se produzcan exactamente 12 ataques en una semana es

$$P(N_7 = 12) = e^{-14} \times \frac{14^{12}}{12!} = \mathbf{0.0984}$$

Forma de la distribución de Poisson

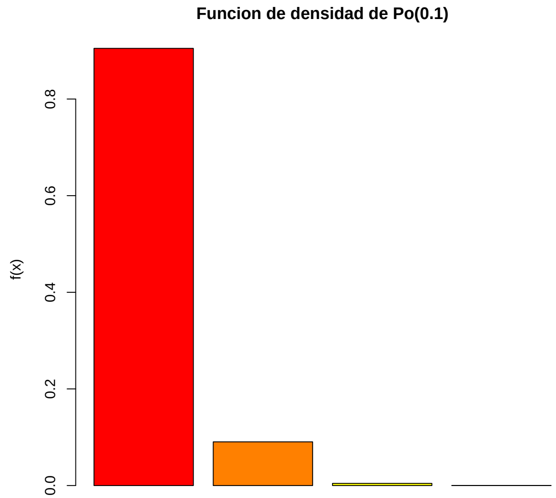
- La función de masa de la distribución de Poisson es **asímetrica a la derecha**.
- La cola hacia la derecha de esta distribución es tanto más marcada cuanto mayor sea la intensidad λ.

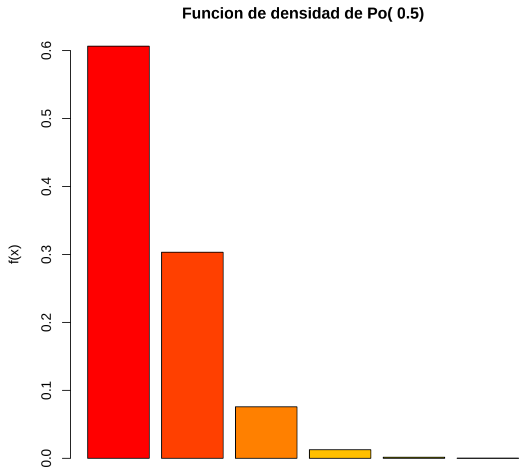


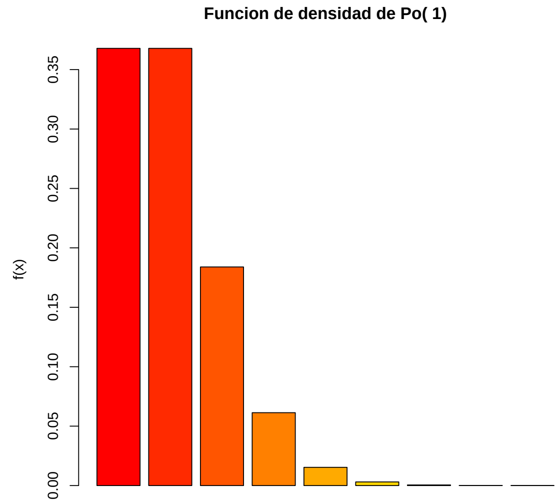


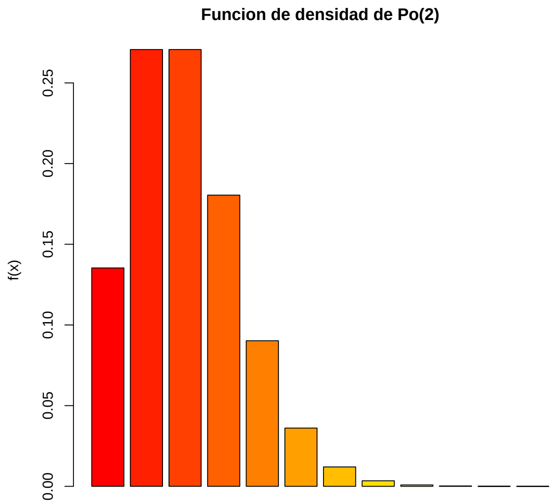
- Si $\lambda \in \mathbb{N}$, la distribución tiene dos modas: λ y λ-1.
- Si $\lambda \notin \mathbb{N}$, la única moda es $[\lambda] - 1$.

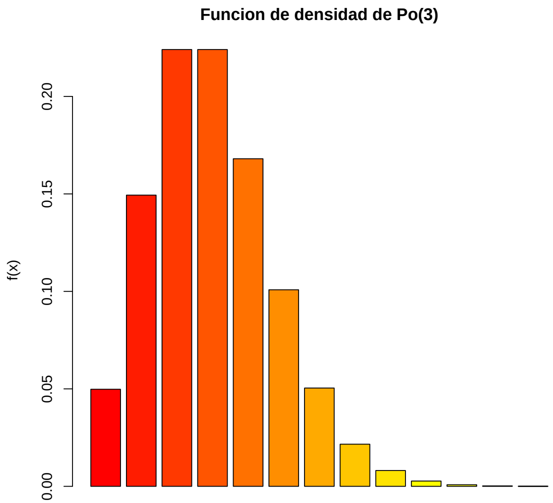
Gráficos de la distribución de Poisson

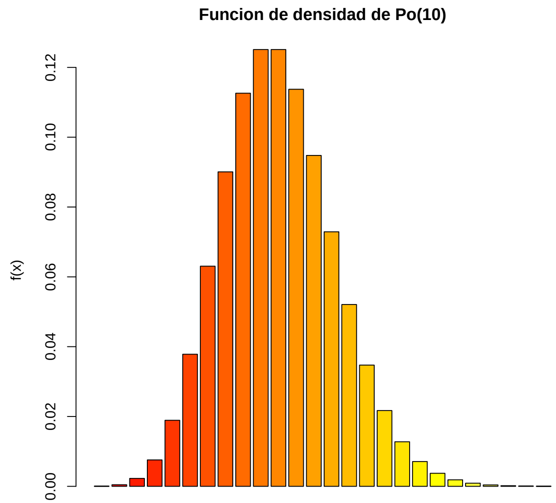












Sucesos puntuales a lo largo del espacio

- Hemos descrito la Poisson como una distribución que modeliza el número de ocurrencias a lo largo del tiempo, pero también sirve como modelo para contar las ocurrencias que se producen a lo largo de un espacio, o de otro soporte continuo.
- Por ejemplo, podemos modelizar mediante una distribución de Poisson el número de puntos de óxido por metro de alambre, el número de taras por cada 10 m^2 de tela, o el número de pasas en una cucharada de cereales para el desayuno.

Variables aleatorias continuas

URJC

Sonia Hernández Alonso
Probabilidad-Grado en Matemáticas (URJC)

Soporte de una variable aleatoria continua

- Una **variable aleatoria** es **continua** si la cantidad de valores que puede tomar es no numerable.
- El **soporte** de una variable aleatoria continua está formado por **uno o varios intervalos** de la recta real.

Valores de una variable continua

- Al contrario de lo que ocurre con las variables discretas, no se puede conocer el valor exacto de una variable continua.
- Tomemos como ejemplo el peso de una persona elegida al azar. El peso es una variable aleatoria continua, ya que, en principio, puede tomar cualquier valor del intervalo $(0, \infty)$.
- Si medimos los pesos con una balanza que sea capaz de distinguir hasta los kilogramos, y el resultado de la medición es por ejemplo 62 kilos, entonces todo lo que podemos afirmar es que el peso de la persona está entre 61.5 kilos y 62.5 kilos.
- Si usamos una balanza más precisa, capaz de medir hasta los hectogramos, y el resultado de la medición es por ejemplo 62.3 kilos, lo máximo que se puede afirmar es que el peso de la persona está entre 62.25 kilos y 62.35 kilos...

Probabilidades sobre una variable continua

- Dado que no es posible determinar el peso exacto de una persona, ¿tiene sentido plantearse cuál es la probabilidad de que la persona pese exactamente una determinada cantidad?

- Claramente, no. Pero sí tiene sentido tratar de calcular, por ejemplo, la probabilidad de que el peso de la persona esté entre 60 y 65 kilos.

- Para las variables aleatorias continuas, lo interesante no va a ser calcular probabilidades sobre puntos aislados, sino **probabilidades de que la variable** se encuentre en el entorno de un punto, es decir, que **tome valores dentro de un determinado intervalo.**

- En las variables continuas, la probabilidad de cada punto aislado es 0, esto es, se verifica

$$P(X = k) = 0$$

para todo $k \in \mathbb{R}$.

Función de densidad. Propiedades

Sonia Hernández Alonso
Probabilidad-Grado en Matemáticas (URJC)

Función de densidad de una variable continua

- Para describir una variable aleatoria continua, X, hay que especificar cómo se distribuye la probabilidad a lo largo de su soporte.

- Puesto que

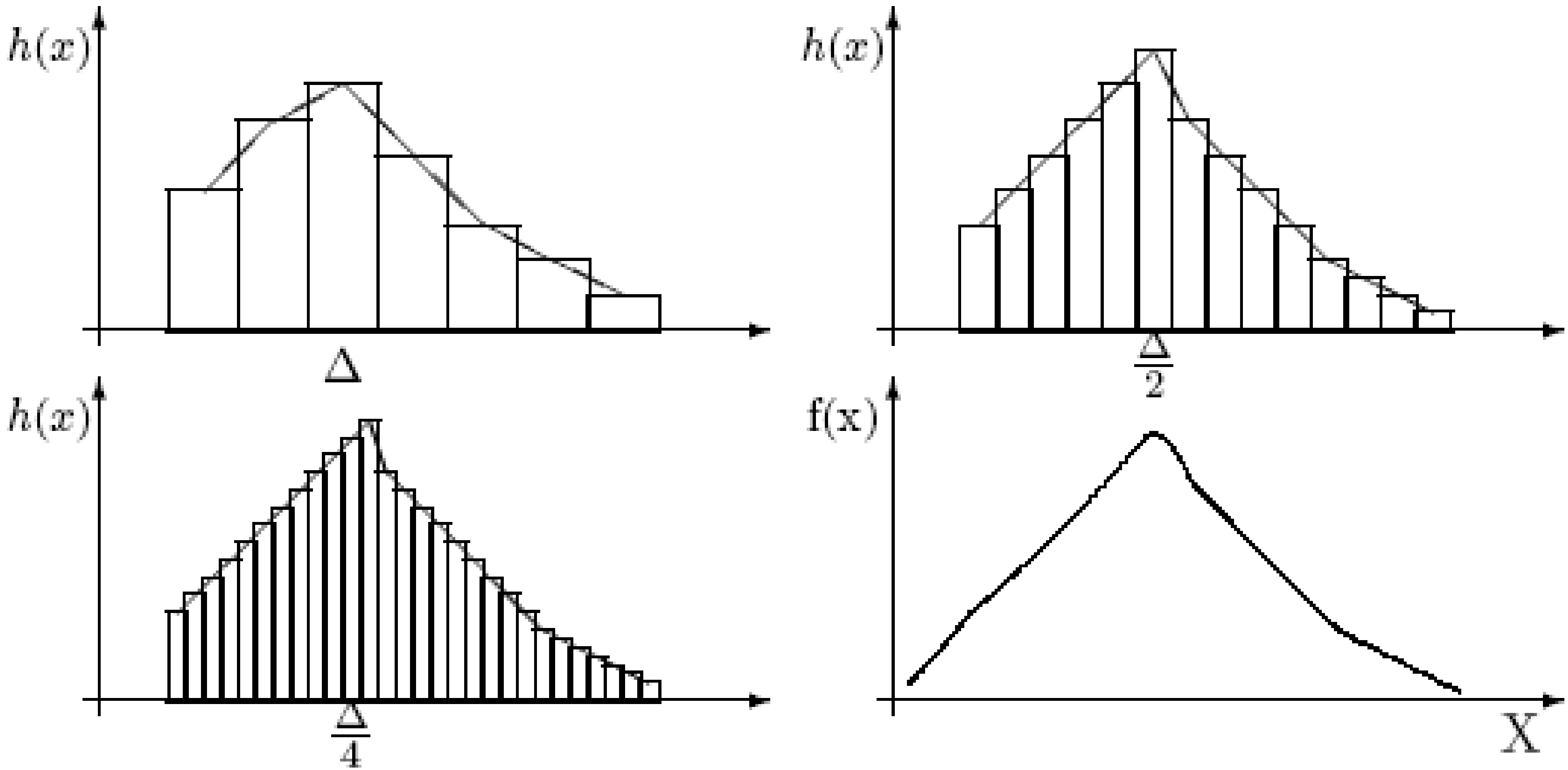
$$P(X = k) = 0$$

  para todo $k \in \mathbb{R}$, no podemos describir la distribución de probabilidades de X mediante una función de masa.

- Supongamos que se dispone de n realizaciones de una variable aleatoria continua X. Estos valores se pueden representar en un **histograma de frecuencias relativas**.

- Si se hace crecer n, es decir, si se toman cada vez más realizaciones de X, y se representan en un histograma con intervalos cada vez más pequeños, el histograma tiende a una curva suave que describe la distribución de la variable. Esta **curva límite** recibe el nombre de **función de densidad**, y la denotaremos por f_X.

Concepto gráfico de la función de densidad



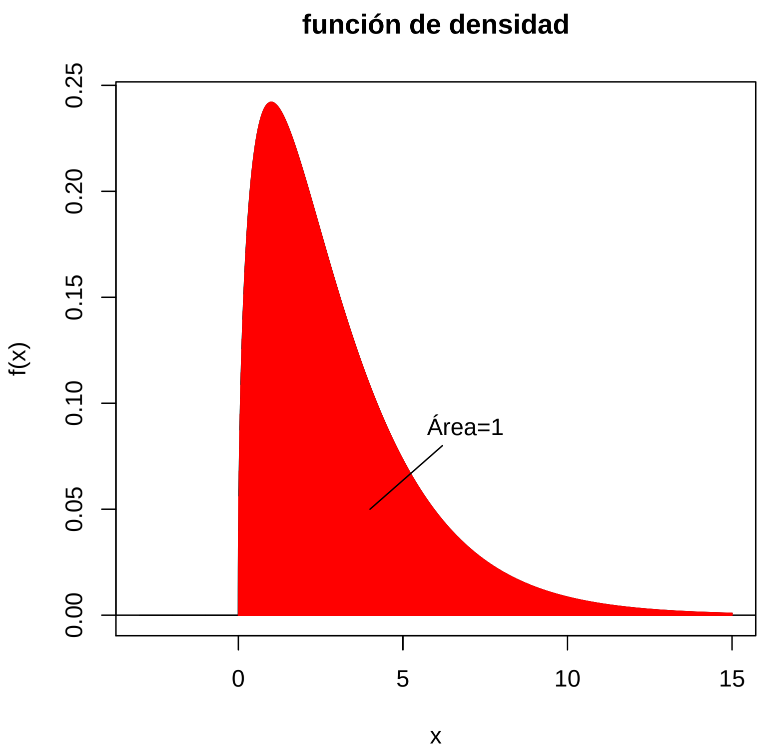
Propiedades de las funciones de densidad

- La función de densidad verifica las siguientes propiedades:

  1. Es una función **no negativa**, es decir, $f_X(x) \geq 0$ para todo x.
  2. La **integral** de f_X a lo largo de su soporte es **1**, es decir,

  $$\int_{S_X} f_X(x)\, dx = 1.$$

  Esto supone que el área que deja por debajo f_X siempre es 1:



Cálculo de probabilidades sobre variables continuas

- La función de densidad de una variable continua X permite calcular cualquier probabilidad acerca de dicha variable como un área:

$$P(X \in A) = \int_A f_X(x)\, dx.$$

- En particular, para $a \leq b$, se tiene que

$$P[a \leq X \leq b] = \int_a^b f_X(x)\, dx.$$

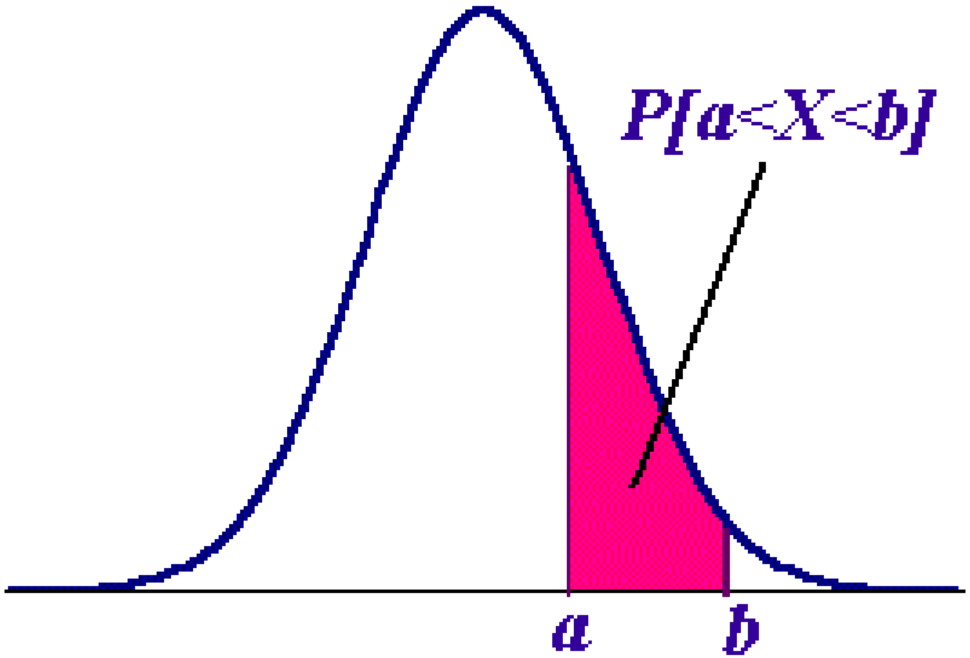
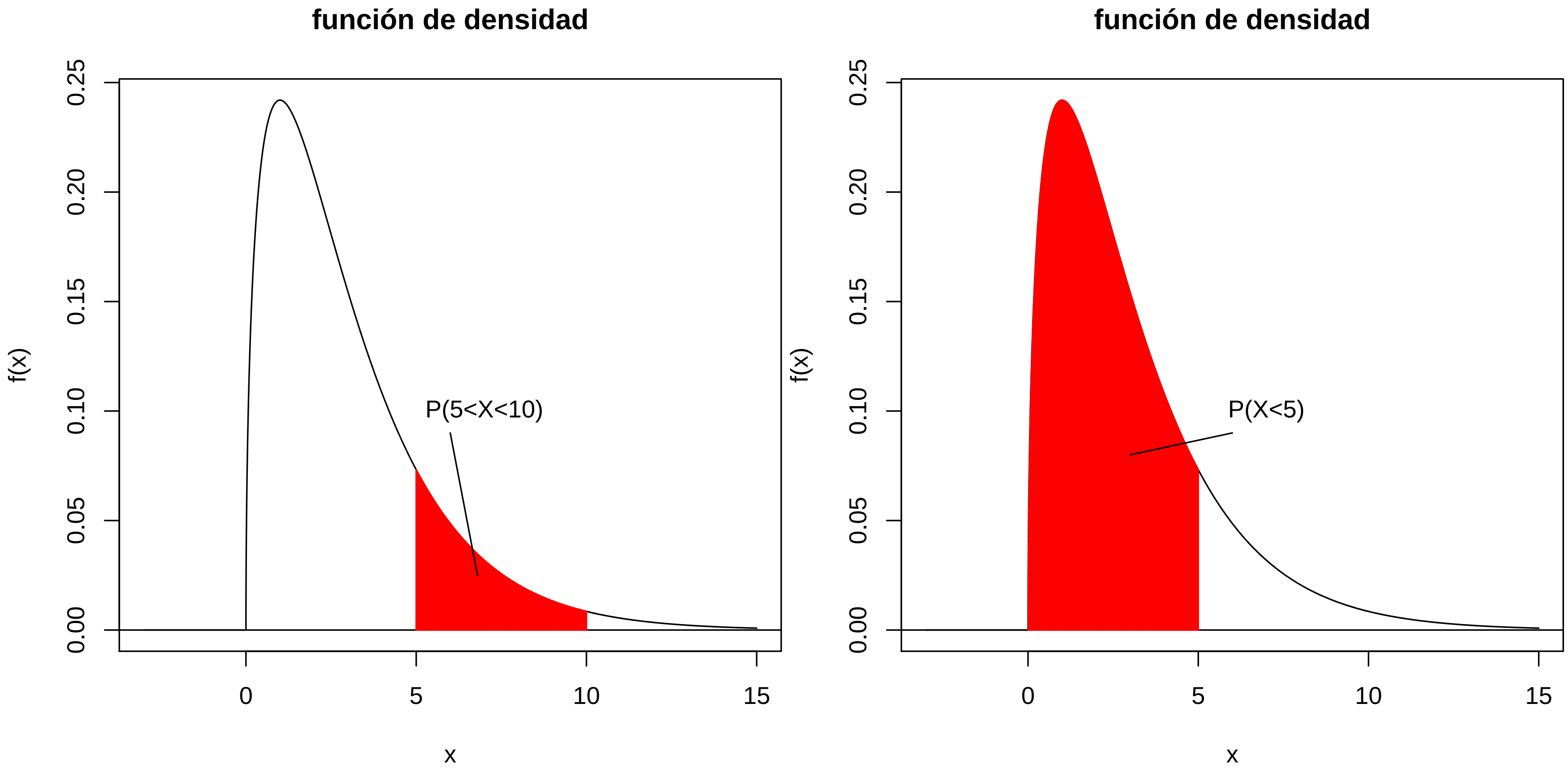


Gráfico: probabilidades sobre variables continuas

- Por ejemplo, en los diagramas siguientes aparecen señaladas las áreas que determinan las probabilidades

$$P(5 < X < 10) \qquad \text{y} \qquad P(X < 5)$$

para la variable aleatoria con la función de densidad del dibujo:



Probabilidad de intervalos abiertos/cerrados

- Para las variables continuas, la probabilidad de cualquier punto $a \in \mathbb{R}$ es cero, ya que

$$P[X = a] = P[a \leq X \leq a] = \int_a^a f_X(x)\,dx = 0.$$

- Por consiguiente, la probabilidad de un intervalo será la misma si es abierto que si es cerrado por cualquiera de sus extremos, pues los extremos son puntos y por tanto tienen probabilidad nula:

$$\begin{aligned} P[a \leq X \leq b] &= P[a \leq X < b] \\ &= P[a < X \leq b] \\ &= P[a < X < b] \\ &= \int_a^b f_X(x)\,dx. \end{aligned}$$

¿Están bien definidas estas probabilidades?

- Las probabilidades calculadas a partir de una función de densidad verifican los **axiomas de Kolmogorov**:

1. La probabilidad de cualquier suceso es **no negativa**:

$$P(X \in A) = \int_A f_X(x)\,dx \geq 0,$$

2. La **probabilidad del suceso seguro es 1**:

$$P(X \in \mathbb{R}) = P(-\infty < X < \infty) = \int_{-\infty}^{\infty} f_X(x)\,dx = \int_{S_X} f_X(x)\,dx = 1,$$

3. Si A_1 y A_2 son **disjuntos** (es decir, si $A \cap B = \emptyset$), entonces

$$\begin{aligned} P[X \in (A_1 \cup A_2)] &= \int_{A_1 \cup A_2} f_X(x)\,dx \\ &= \int_{A_1} f_X(x)\,dx + \int_{A_2} f_X(x)\,dx \\ &= P(X \in A_1) + P(X \in A_2). \end{aligned}$$

Ejemplo: variable aleatoria continua

- **Ejercicio 12:** Una máquina fabrica ejes cuyos diámetros (D), medidos en metros, se distribuyen según la función de densidad

$$f_D(x) = \begin{cases} \dfrac{x}{32} & \text{si } x \in [6, 10], \\ 0 & \text{en caso contrario.} \end{cases}$$

1. Comprobar que f_D es una función de densidad.

2. Calcular $P(7 < D < 9)$, $P(D < 8)$ y $P(D > 5)$.

Solución al Ejercicio 12

1. Empecemos por comprobar que f_D es, efectivamente, una función de densidad.

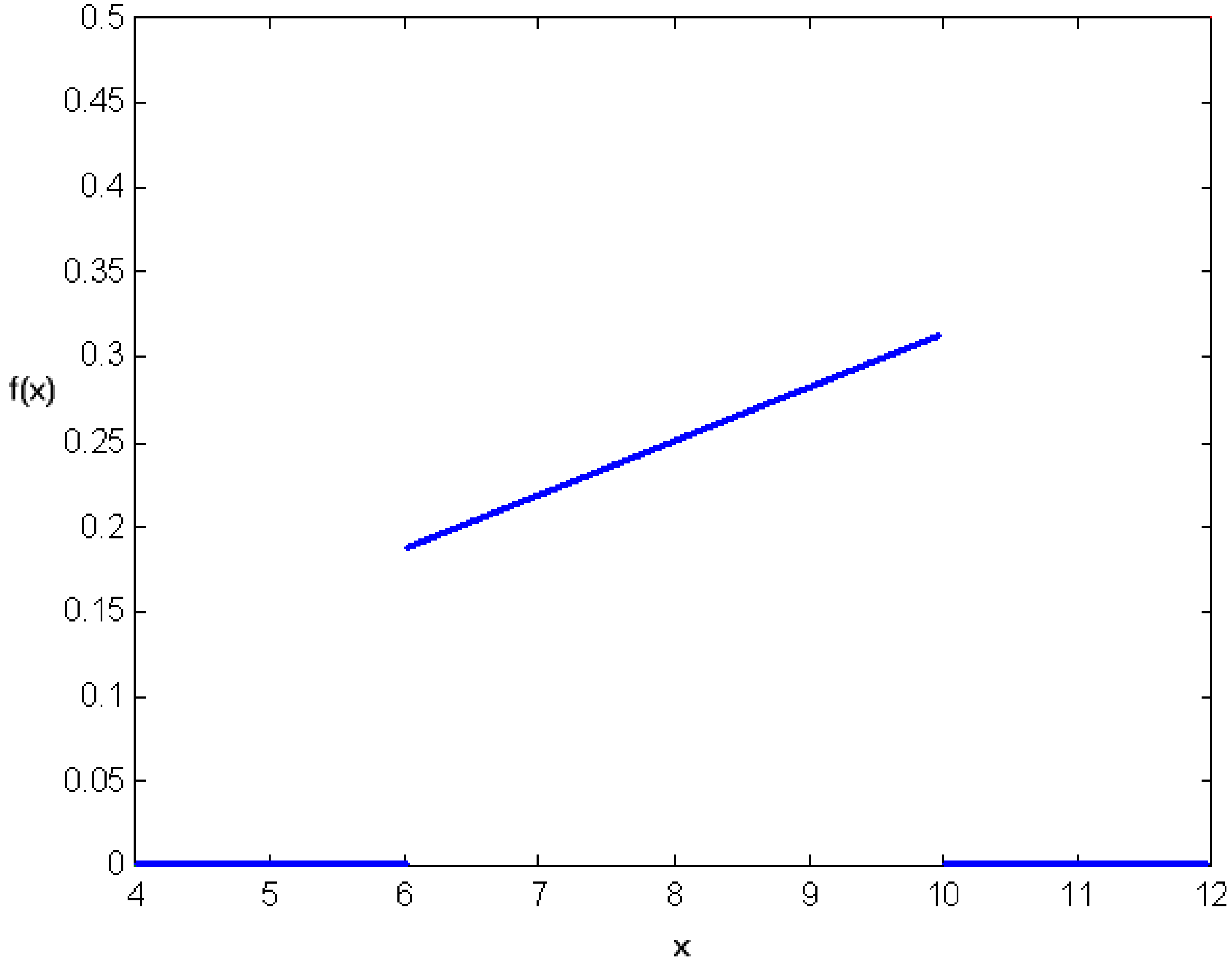
   Esto supone verificar que:

   a) $f_D(x) \geq 0$ para todo $x \in \mathbb{R}$

   b) $\displaystyle\int_6^{10} f_D(x)\,dx = 1$

Solución al Ejercicio 12 (continuación)

- El gráfico siguiente representa f_D:

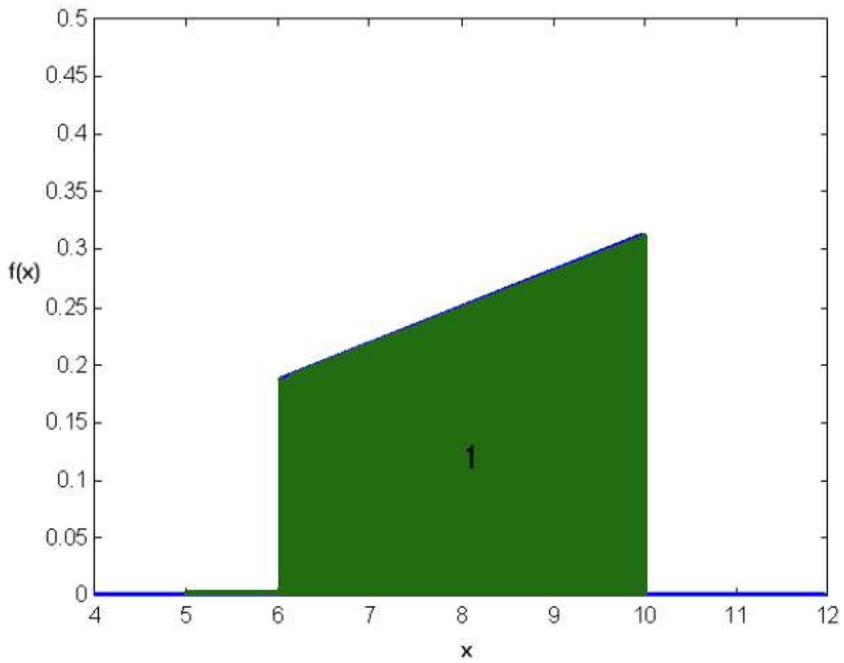


Solución al Ejercicio 12 (continuación)

a) Constatar que $f_D(x) \geq 0$ para todo $x \in \mathbb{R}$ es trivial.

b) La segunda propiedad supone comprobar que el área de la región del plano comprendida entre f_D y el eje x es 1:

$$\int_6^{10} \frac{x}{32}\,dx = \left[\frac{x^2}{64}\right]_6^{10} = \frac{10^2 - 6^2}{64} = \frac{100 - 36}{64} = 1.$$



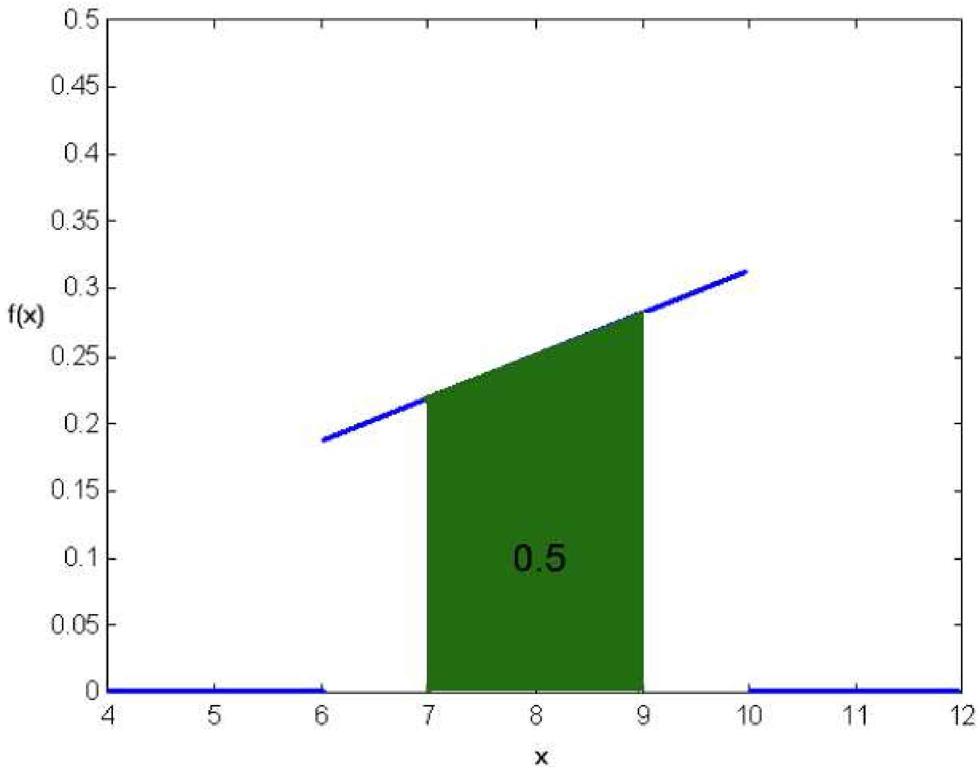
- Por tanto f_D es, en efecto, una función de densidad.

Solución al Ejercicio 12 (continuación)

2. Las probabilidades de que D tome un valor dentro de un determinado conjunto se calcularán integrando f_D en ese conjunto.

   Por ejemplo, la probabilidad de que el diámetro de un eje elegido al azar mida entre 7 y 9 metros es

$$P\,(7 < D < 9) = \int_7^9 \frac{x}{32}\,dx = \left[\frac{x^2}{64}\right]_7^9 = \frac{81-49}{64} = \frac{1}{2} = \mathbf{0.5}$$

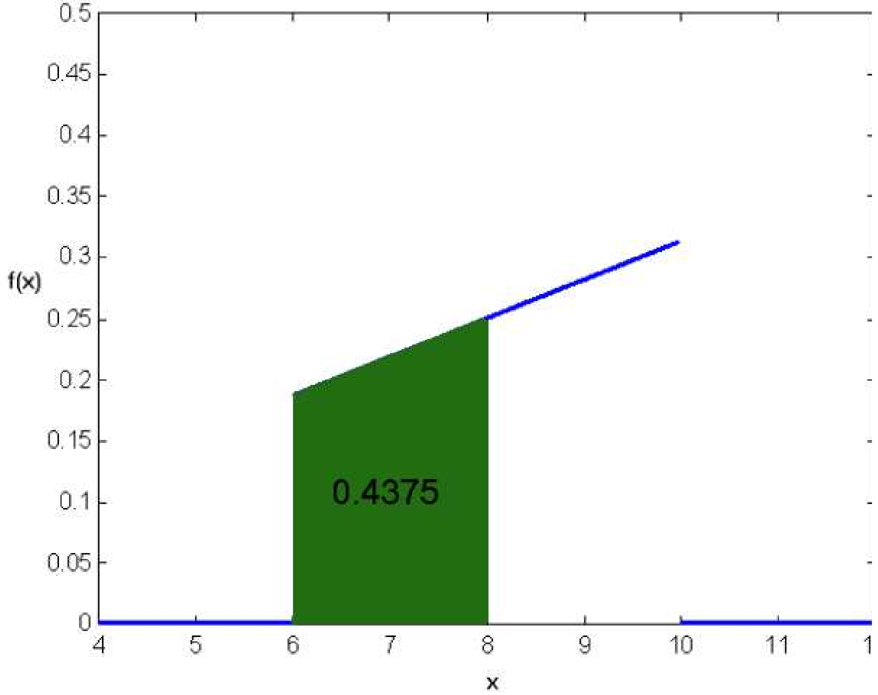


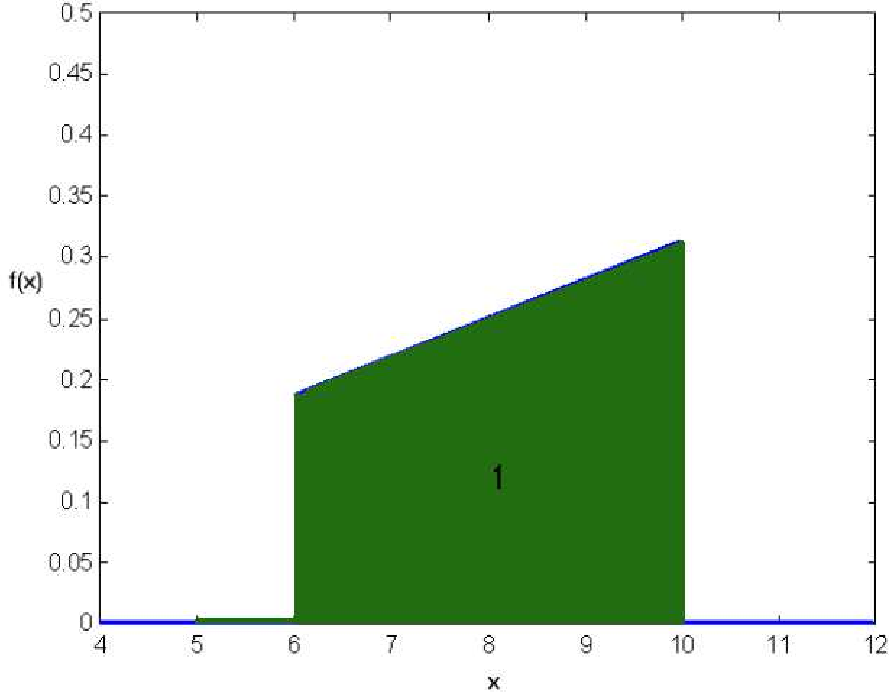
Solución al Ejercicio 12 (continuación)

- De manera análoga podemos calcular la probabilidad de que el eje mida menos de 8 metros o la de que mida más de 5 metros:

$$P(D<8) = \int_6^8 \frac{x}{32}\,dx = \left[\frac{x^2}{64}\right]_6^8 = \frac{64-36}{64} = \frac{28}{64} = \frac{7}{16} = \mathbf{0.4375}$$

$$P(D>5) = \mathbf{1}$$





Función de distribución

de variables aleatorias continuas

Sonia Hernández Alonso
Probabilidad-Grado en Matemáticas (URJC)

Función de distribución para X continua

- Recordemos que la **función de distribución** de una variable aleatoria X es la función F_X que asigna a cada $t \in \mathbb{R}$ la probabilidad de que X tome un valor menor o igual que t:

$$\begin{aligned} F_X : \quad & \mathbb{R} \longrightarrow (0,1) \\ & t \longrightarrow F_X(t) = P(X \leq t) \end{aligned}$$

- Para las variables continuas, la forma de calcular esta función de probabilidad acumulada es **integrando la función de densidad de la variable hasta el punto t:**

$$F_X(t) = P(X \leq t) = \int_{-\infty}^{t} f_X(x)\, dx.$$

Ejemplo: función de distribución

- **Ejercicio 13:** Retomemos la variable aleatoria D que indica el diámetro de los ejes, con función de densidad

$$f_D(x) = \begin{cases} \dfrac{x}{32} & \text{si } x \in [6, 10], \\ \\ 0 & \text{en caso contrario.} \end{cases}$$

¿Cuál es la función de distribución de D?

Solución al Ejercicio 13

- La función de distribución de la variable D es

$$F_D(t) = P[D \leq t] = \int_{-\infty}^{t} f_D(x) = \begin{cases} 0 & \text{si } t < 6, \\ \dfrac{t^2 - 36}{64} & \text{si } 6 \leq t \leq 10, \\ 1 & \text{si } t \geq 10. \end{cases}$$

Ejercicio 13: cálculo de probabilidades con F_D

- Conociendo la función de distribución de D se puede calcular cualquier probabilidad de esta variable sin necesidad de integrar.
- Así, por ejemplo,

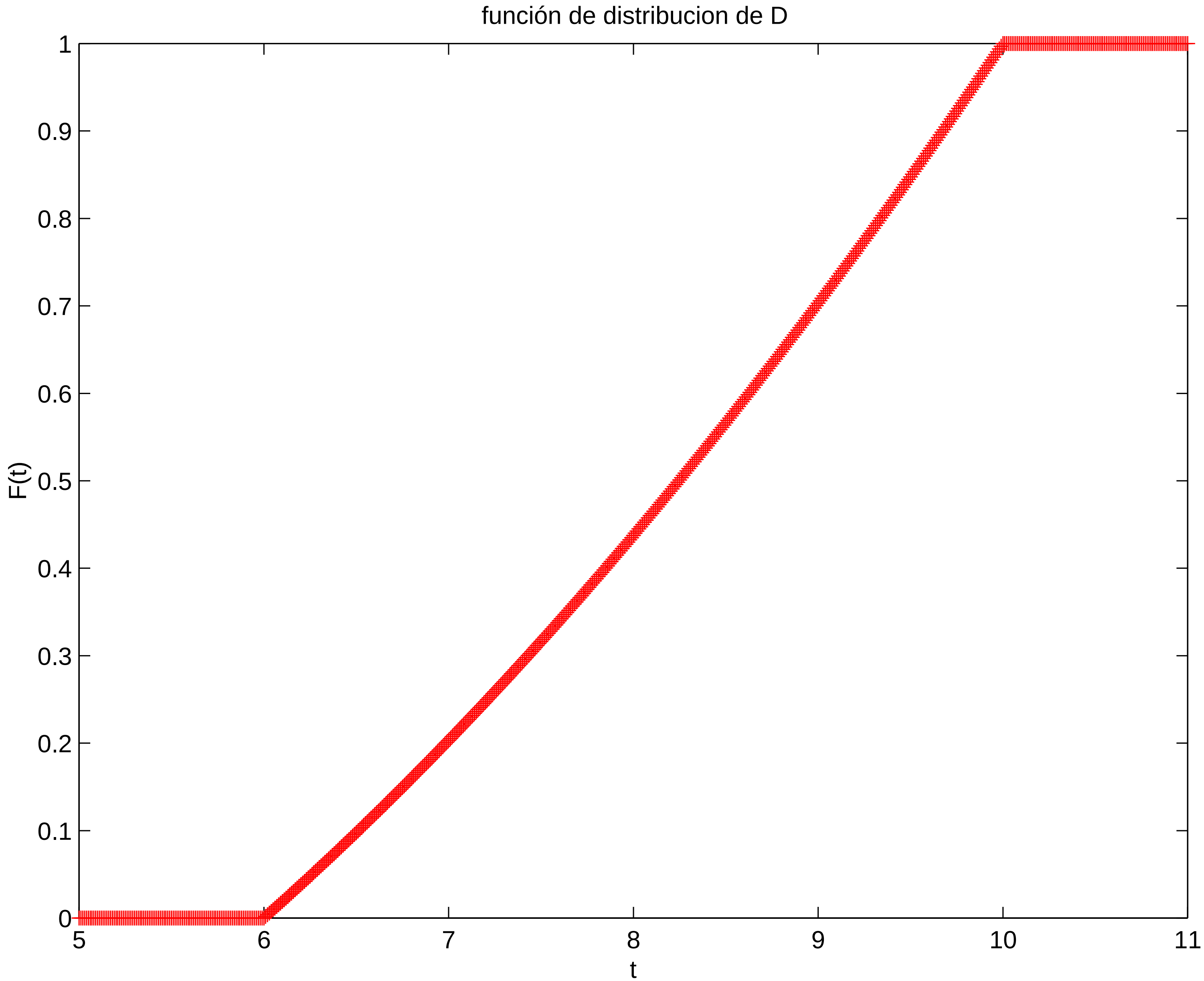
$$P(D < 8) = F_D(8) = \frac{8^2 - 36}{64} = 0.4375$$

$$P(D > 9) = 1 - F_D(9) = 1 - \frac{9^2 - 36}{64} = 1 - 0.7031 = 0.2969$$

$$P(7 < X < 9) = F_D(9) - F_D(7) = \frac{9^2 - 36}{64} - \frac{7^2 - 36}{64} = 0.5$$

$$P(8 < X < 11) = F_D(11) - F_D(8) = 1 - \frac{8^2 - 36}{64} = 0.5625$$

Gráfico: función de distribución de D (F_D)



Ejercicio 13: propiedades de F_D

- Observamos que F_D:
  - Comienza en 0, ya que $F_D(t) = 0$ para todo $t < 6$.
  - Termina en 1, ya que $F_D(t) = 1$ para todo $t \geq 10$.
  - Es monótona no decreciente.
  - Es una función continua.
- Este es el aspecto que presentan en general las funciones de distribución de las variables aleatorias continuas.

Relación entre F_X y f_X

- El **teorema fundamental del cálculo** implica que, para variables continuas,

$$\frac{dF_X(x)}{dx} = f_X(x).$$

- **Ejercicio 14:** Hallar la función de densidad de la variable aleatoria con función de distribución

$$F_H(t) = \begin{cases} 0 & \text{si } t < 0 \\ t^5 & \text{si } 0 \leq t \leq 1, \\ 1 & \text{si } t > 1. \end{cases}$$

resolución:.........pizarra

Esperanza y varianza

de variables aleatorias continuas

Sonia Hernández Alonso
Probabilidad-Grado en Matemáticas (URJC)

Esperanza de una variable aleatoria continua

- La **esperanza** o **media** de una variable aleatoria continua X viene dada por

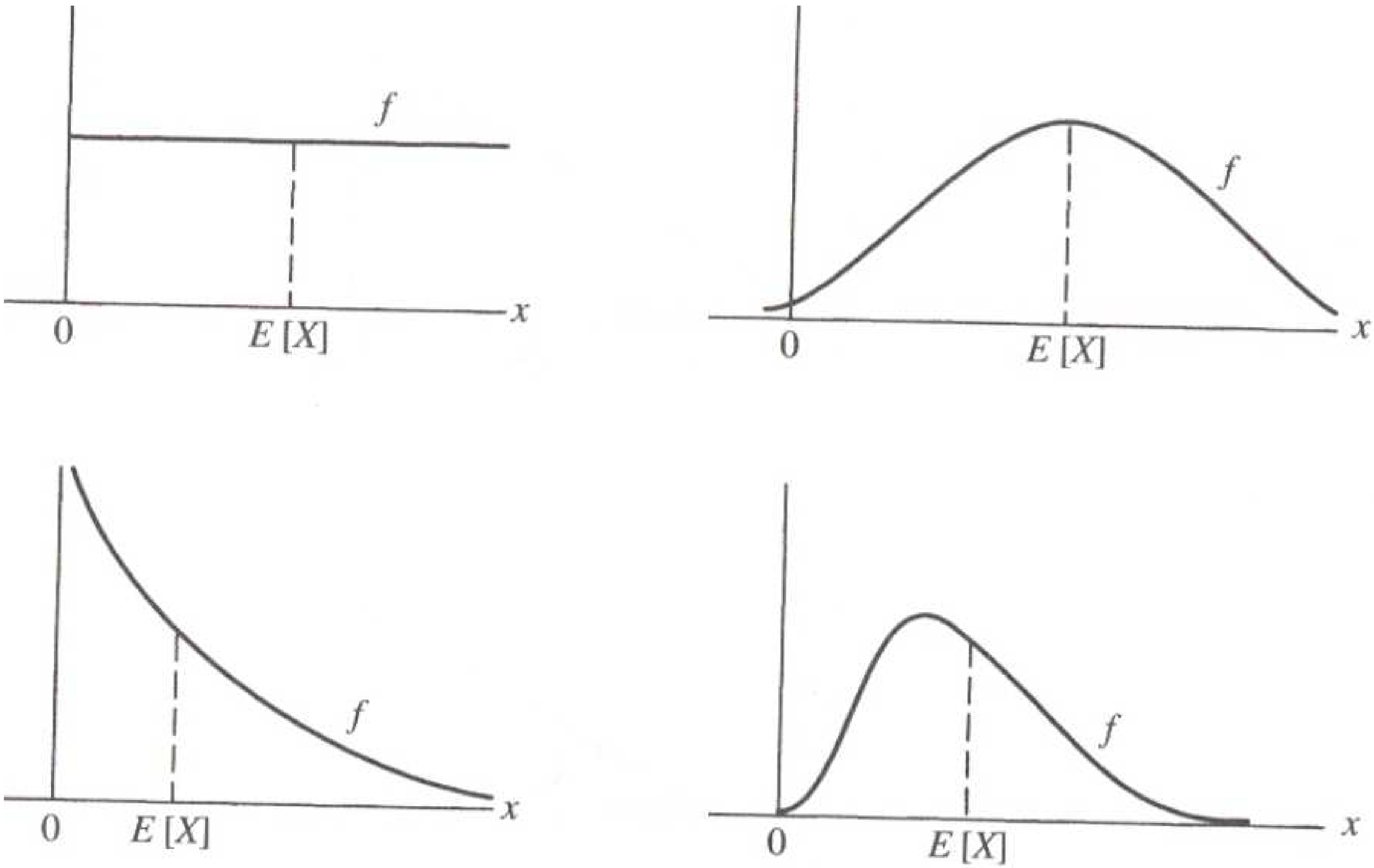
$$\boxed{E\left(X\right) = \int_{S_X} x f_X(x)\, dx}$$

- Al igual que en el caso de las variables discretas, es habitual denotar la esperanza de una variable continua X por μ_X, o por μ.

- μ_X indica dónde se encuentra el **centro de gravedad** o punto de equilibrio de la distribución de probabilidad de X.

Visualización gráfica de $E(X)$

- Excepto en los casos triviales, como por ejemplo las variables con densidad simétrica, no es posible hallar $E(X)$ sin recurrir a la integración. Pero sí es posible visualizar geométricamente la esperanza:



Ejemplo: esperanza de una variable continua

- **Ejercicio 15:** Consideremos de nuevo la variable aleatoria D que modela el diámetro de los ejes, cuya función de densidad era

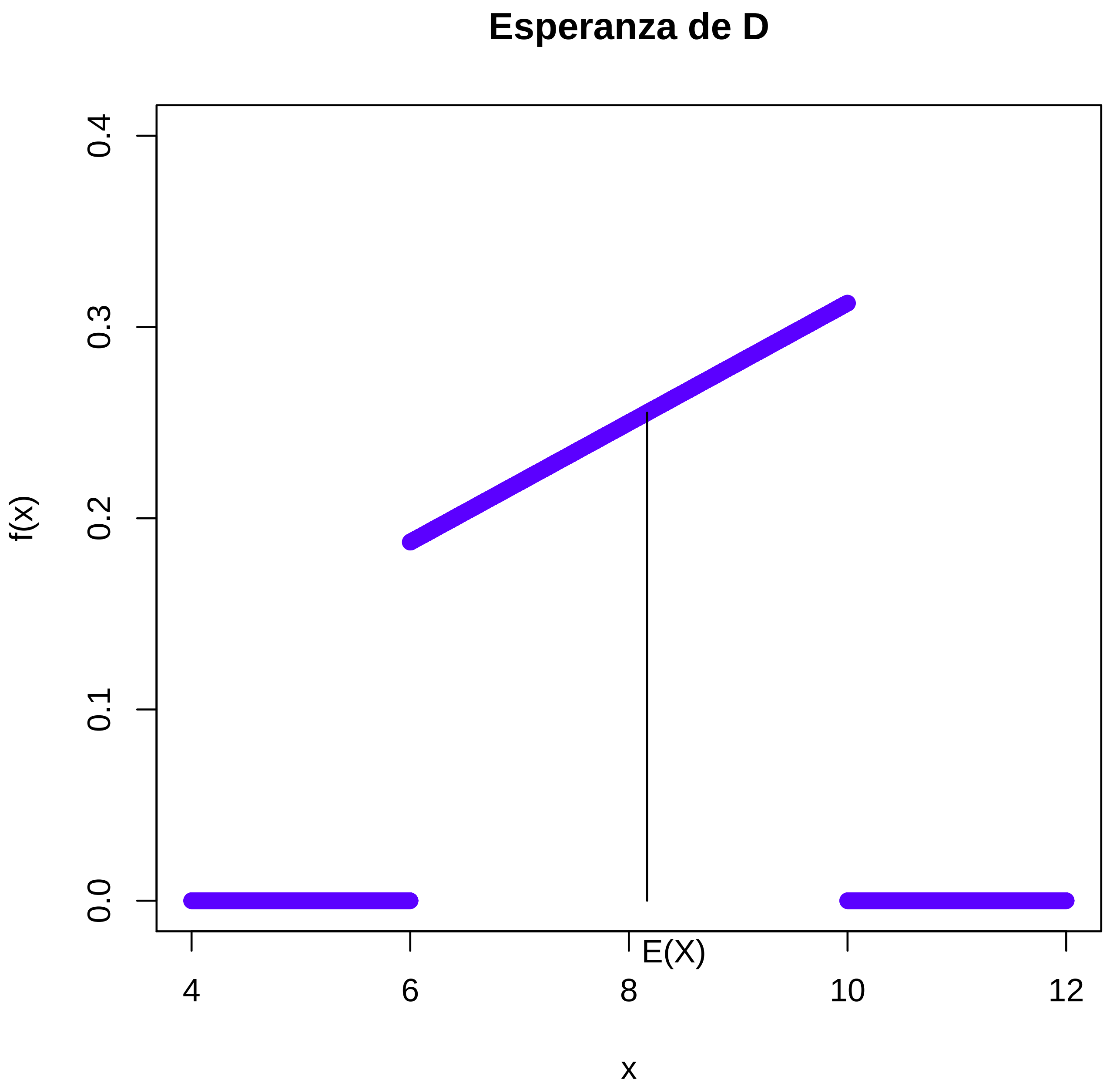
$$f_D(x) = \begin{cases} \dfrac{x}{32} & \text{si } x \in [6, 10], \\ 0 & \text{en caso contrario.} \end{cases}$$

- La esperanza de D es

$$\begin{aligned} E(D) &= \int_{S_D} x f_D(x)\, dx = \int_6^{10} x \frac{x}{32}\, dx = \int_6^{10} \frac{x^2}{32}\, dx = \\ &= \left[\frac{x^3}{32 \times 3}\right]_6^{10} = \frac{10^3 - 6^3}{96} = \frac{784}{96} = \mathbf{8.1667} \end{aligned}$$

- Luego la **longitud esperada** de los díametros de los ejes que fabrica esta máquina es **8,1667 metros**.

Gráfico: esperanza de una variable continua



Varianza de una variable aleatoria continua

- La **varianza** de una variable aleatoria continua,

$$V(X) = E\left[(X-\mu)^2\right] = E\left(X^2\right) - E^2(X),$$

puede calcularse como

$$\boxed{V(X) = \int_{S_X} x^2 f_X(x)\, dx - \mu_X^2}$$

- $V(X)$ mide la **dispersión** de X alrededor de su punto de equilibrio, es decir, alrededor de $E(X)$.

- Al igual que en el caso de las variables discretas, es común denotar la varianza de una variable aleatoria continua σ_X^2, o simplemente σ^2.

- La **desviación típica** de X es la raiz cuadrada de su varianza, y se denota σ_X, o simplemente σ.

Ejemplo: varianza de una variable continua

- **Ejercicio 16:** Para los diámetros de los ejes de los ejemplos anteriores tenemos que

$$E(D) = \int_{S_D} x f_D(x)\,dx = \int_6^{10} \frac{x^2}{32}\,dx = 8.1667 \text{ m},$$

y

$$E(D^2) = \int_{S_D} x^2 f_D(x)\,dx = \int_6^{10} x^2 \frac{x}{32}\,dx = \int_6^{10} \frac{x^3}{32}\,dx = \frac{8704}{128} = 68 \text{ m}^2.$$

- Por tanto, la **varianza** de los diámetros de los ejes es

$$V(D) = E(D^2) - E^2(D) = 68 - 8.1667^2 = 1.3056 \text{ m}^2.$$

- La **desviación típica** de D es

$$\sigma_D = \sqrt{1.3056} = 1.1426 \text{ m}.$$

Modelos continuos especiales

Sonia Hernández Alonso
Probabilidad-Grado en Matemáticas (URJC)

Familias de modelos continuos

- Al igual que en el caso discreto, hay ciertos modelos de variables aleatorias continuas que se repiten con mucha frecuencia.
- Vamos a analizar tres de las familias de distribuciones continuas aparecen con mayor frecuencia: la distribución uniforme, la exponencial, y la distribución normal o gaussiana.

Distribución uniforme

Sonia Hernández Alonso
Probabilidad-Grado en Matemáticas (URJC)

Probabilidad uniformemente repartida

URJC

- Hay muchas situaciones en las cuales la probabilidad está repartida de manera uniforme entre los valores del soporte
- Consideremos por ejemplo una "rueda de la fortuna" consistente en dejar girar una aguja hasta que se para y medir el ángulo que forma esta aguja con el eje horizontal. En este caso, el resultado puede ser cualquier ángulo entre 0^0 y 360^0, y los ángulos de igual amplitud tienen la misma probabilidad de aparecer.
- Otro ejemplo consiste en considerar un vehículo que recorre con velocidad constante cierto tramo desde el principio al final y viceversa. La posición de ese vehículo en un instante elegido al azar, es una variable aleatoria que puede tomar cualquier valor dentro del tramo considerado, y los fragmentos de dicho tramo con la misma longitud tienen también la misma probabilidad.

Soporte y función de distribución de $X \sim U(a,b)$

- Consideremos dos números reales, a y b, con $a < b$.
- La variable aleatoria X sigue distribución **uniforme** en el intervalo (a,b) si su función de densidad es

$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & \text{si } a < x < b, \\ 0 & \text{en el resto.} \end{cases}$$

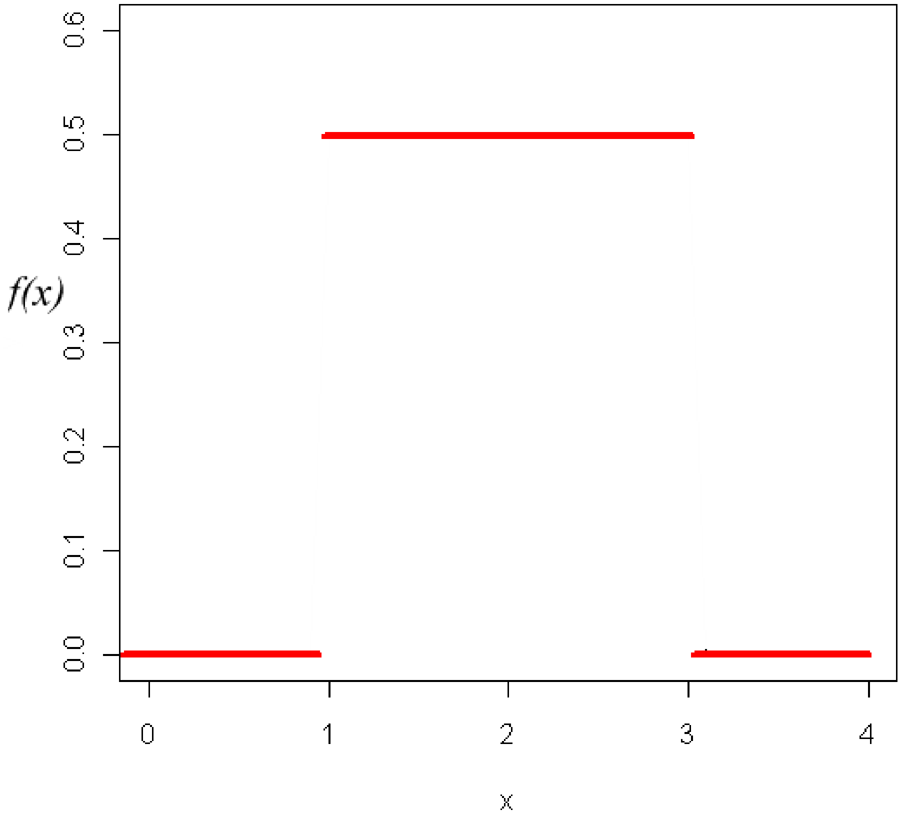
- De forma abreviada esto se denota

$$X \sim U\left(a,b\right)$$

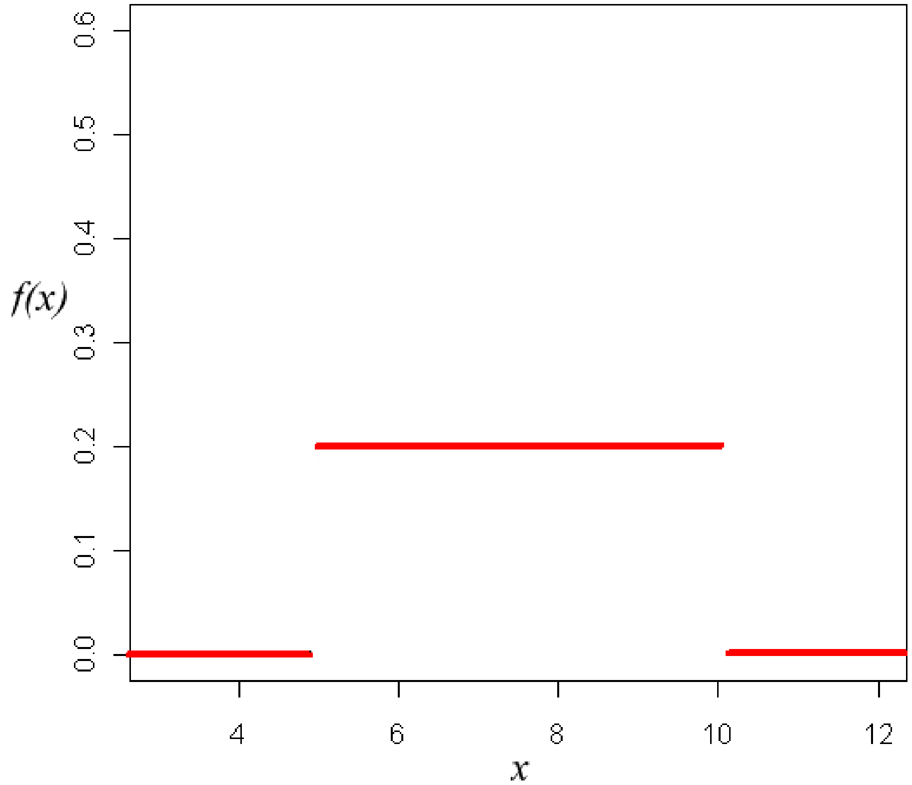
- La densidad de una variable uniforme es **constante a lo largo de su soporte**. Por consiguiente, la probabilidad de que el valor de X esté comprendido en un subintervalo de (a,b) **depende únicamente de la longitud del mismo, pero no de su posición.**

Gráficos de distribuciones uniformes

función de densidad de $U(1,3)$:



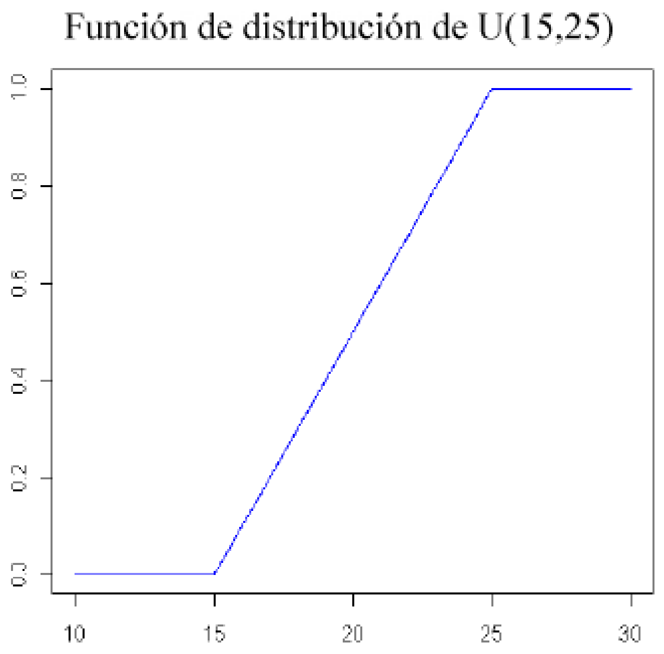
función de densidad de $U(5,10)$:



Función de distribución de $X \sim U(a,b)$

- Es fácil (por ejemplo utilizando áreas de rectángulos) comprobar que la **función de distribución** de una variable $X \sim U(a,b)$ es

$$F_X(t) = P[X \le t] = \int_{-\infty}^{t} \frac{1}{b-a}\,dx = \begin{cases} 0 & \text{si } x < a, \\ \dfrac{t-a}{b-a} & \text{si } a \le x \le b, \\ 1 & \text{si } x > b. \end{cases}$$

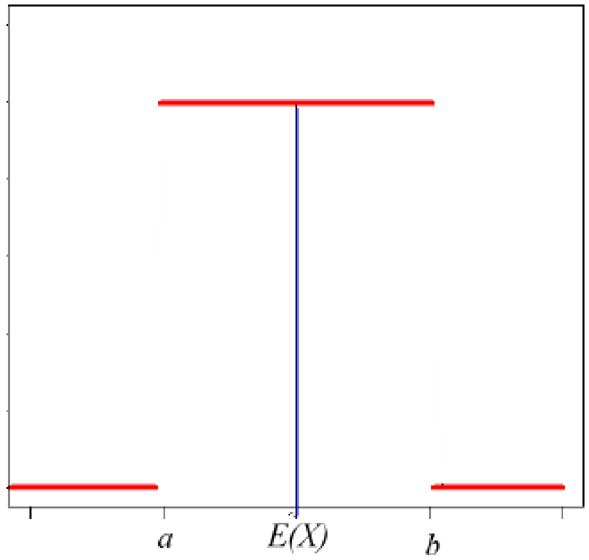


Esperanza de $X \sim U(a,b)$

- La **esperanza** de una variable $X \sim U(a,b)$ es

$$\boxed{E(X) = \int_a^b \frac{x}{b-a}\, dx = \frac{a+b}{2}}$$

es decir, el punto medio del intervalo (a,b):



- En cuanto a la varianza de la distribución uniforme viene dada por

$$V(X) = \int_a^b \frac{x^2}{b-a}\, dx - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}$$

Ejemplo: distribución uniforme

- **Ejercicio 17:** Un coche recorre con velocidad constante el tramo que va desde el kilómetro 15 hasta el kilómetro 25 de cierta carretera.

  Sea

  $$C = \text{posición del coche en un instante elegido al azar}$$

  expresada en kilómetros.

  1. ¿Qué distribución tiene la variable C? ¿Cuál es su función de densidad? ¿Y su función de distribución?
  2. ¿Cuál es la probabilidad de que el un instante aleatorio el coche se encuentre entre el kilómetro 18 y el kilómetro 22?
  3. ¿Cuál es valor esperado de C?

Solución al Ejercicio 17

1. Puesto el que el coche se despaza a velocidad constante y observamos su posición en un instante aleatorio, la distribución de P es uniforme a lo largo de su soporte, es decir, entre en kilómetro 15 y el 25:

$$C \sim U(15, 25)$$

Por tanto, su función de densidad es

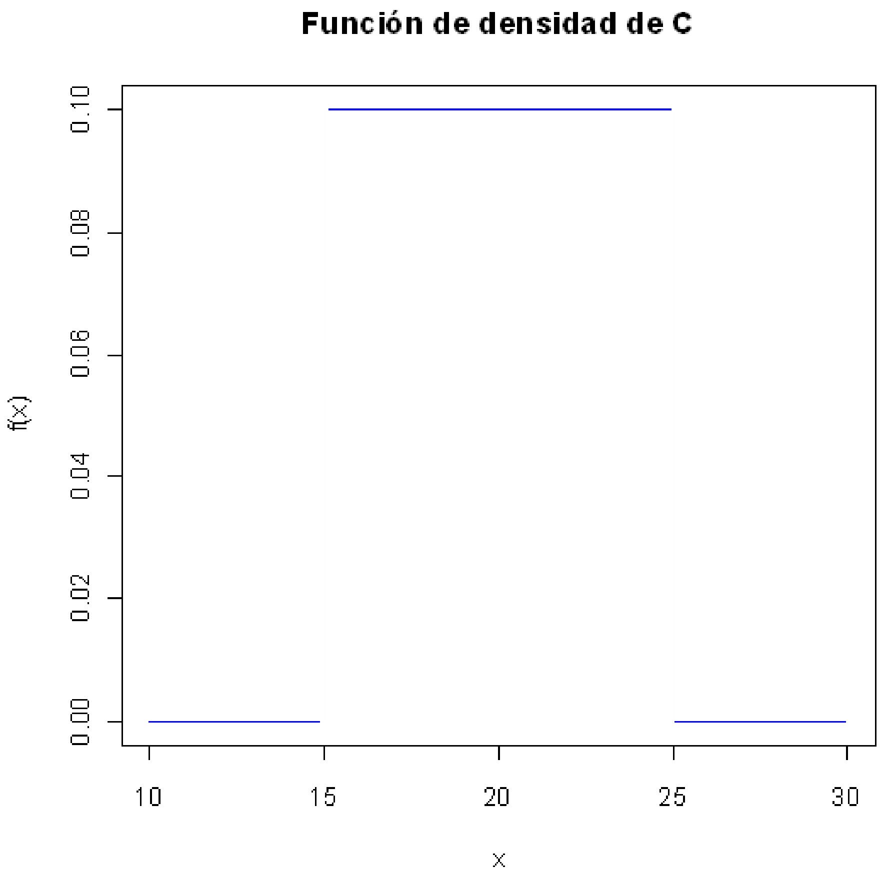
$$f_C(x) = \begin{cases} \frac{1}{10} & \text{si } 15 < x < 25, \\ 0 & \text{en el resto.} \end{cases}$$

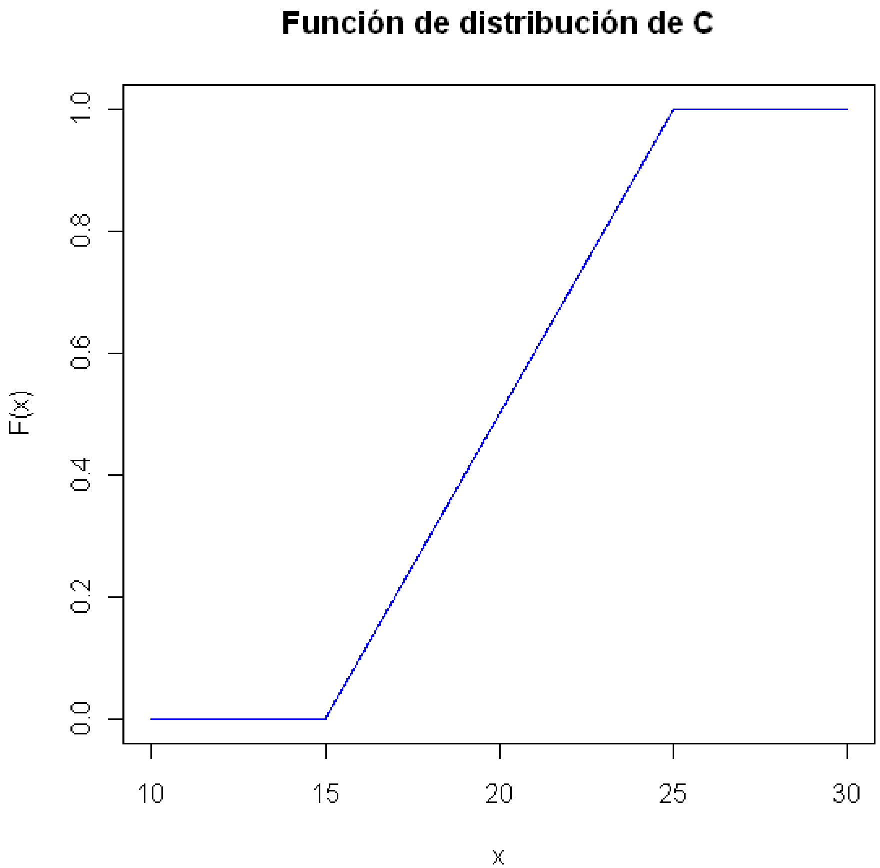
Integrando f_C se obtiene que la función de distribución de C es

$$F_C(x) = \begin{cases} 0 & x \leq 15, \\ \frac{x-15}{10} & 15 < x < 25, \\ 1 & x \geq 25. \end{cases}$$

Solución al Ejercicio 17 (continuación)

- Gráficos:





Solución al Ejercicio 17 (continuación)

2. La probabilidad de que P tome un valor entre 18 y 22 es

$$P\left[C \in (18,22)\right] = \int_{18}^{22} \frac{1}{10}\ dx = \frac{4}{10} = \mathbf{0.4},$$

que puede calcularde de forma más inmediata como

$$P\left[C \in (18,22)\right] = \frac{\text{área favorable a } C \in (18,22)}{\text{área posible}} = \frac{4}{10} = \mathbf{0.4}.$$

3. El valor esperado de C es

$$E\left(C\right) = \frac{15+25}{2} = \mathbf{20}$$

Esto indica que la posición en la que se espera que se encuentre el vehículo en un instante elegido al azar es el kilómetro 20 de esta carretera.