

Lección 4: Variables aleatorias

Juan Antonio Barceló

Israel Herraiz

9 de enero de 2014

1. Introducción y algunas definiciones

La Estadística Descriptiva considera métodos que generalmente se utilizan para describir conjuntos de datos. Algunos de estos conjuntos representan el total de la población estudiada; otros constituyen una muestra extraída de una población más grande. En este último caso solo podemos intentar sacar conclusiones aproximadas acerca de la población de la que procede la muestra. Hemos visto que la Teoría de Probabilidad puede ser utilizada para resolver muchas cuestiones de naturaleza moderadamente complicada. Sin embargo, todavía no hemos empezado a mostrar cómo puede emplearse la Teoría de Probabilidad para obtener conclusiones acerca de una población, con base en una muestra extraída de ella. Para hacer esto, primero hemos de dirigir nuestra atención hacia el concepto de **variable aleatoria**, eslabón entre la Teoría de Probabilidades y la Estadística aplicada.

En un espacio probabilístico (Ω, \mathcal{A}, P) los elementos del espacio muestral Ω no tienen por qué ser números. En la tirada de una moneda al aire, los sucesos elementales, cara y cruz, no son valores numéricos. No obstante, siempre podemos hacer corresponder el número 1 a la cara, y el 0 a la cruz. Esta asignación de valores numéricos a los sucesos elementales de un espacio de probabilidades es la base para definir el concepto de variable aleatoria.

1.1. Concepto de variable aleatoria

Definición 1. Sea (Ω, \mathcal{A}, P) un espacio probabilístico. Una **variable aleatoria** X es una función

$$\begin{aligned} X : \Omega &\longrightarrow \mathbb{R} \\ \omega &\longrightarrow X(\omega). \end{aligned}$$

tal que para todo $x \in \mathbb{R}$ se cumple que

$$\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{A}. \quad (1)$$

Observación 1. Si trabajamos con $\mathcal{A} = \mathcal{P}(\Omega)$, la ecuación (1) siempre se va a cumplir.

La variable aleatoria asociada a un experimento aleatorio, “viene a **cuantificar** en algún sentido” el experimento aleatorio. Otra manera de decirlo, en una función real que actúa como un **aparato de medida** sobre los sucesos elementales.

Ejemplo 1. Una compañía de bebidas anuncia premios en los tapones asegurando que en cada 1000 tapones hay 500 con “inténtelo otra vez”, 300 con premio de 5 euros, 150 con premio de 10 euros, 40 con premio de 50 euros y 10 con premio de 100 euros. Un individuo, al que no le gusta esa bebida, decide comprar una bebida cuyo coste es de 10 euros. Vamos a caracterizar la ganancia por una variable aleatoria.

Experimento aleatorio: Comprar una botella y ver si tenemos premio.

Espacio muestral, los posibles resultados del experimento aleatorio

$$\Omega = \{\text{inténtelo otra vez}, 5 \text{ euros}, 10 \text{ euros}, 50 \text{ euros}, 100 \text{ euros}\}.$$

La variable aleatoria

$$X : \Omega \mapsto \mathbb{R} \text{ definida por } \begin{cases} X(\text{inténtelo otra vez}) = 0 - 10 = -10 \\ X(5 \text{ euros}) = 5 - 10 = -5 \\ X(10 \text{ euros}) = 10 - 10 = 0 \\ X(50 \text{ euros}) = 50 - 10 = 40 \\ X(100 \text{ euros}) = 100 - 10 = 90, \end{cases}$$

está cuantificando la ganancia del experimento aleatorio.

□

Definición 2. Sea (Ω, \mathcal{A}, P) un espacio probabilístico, $X : \Omega \mapsto \mathbb{R}$ una variable aleatoria y $A \subset \mathbb{R}$, definimos

$$P(A) = P(\{\omega \in \Omega : X(\omega) \in A\}).$$

Ejemplo 2.

En el Ejemplo 1, el espacio muestral contiene 5 sucesos elementales que no son igualmente probables.

$$P(\text{inténtelo otra vez}) = \frac{500}{1000} = \frac{1}{2}, \quad P(5 \text{ euros}) = \frac{300}{1000} = \frac{3}{10},$$

$$P(10 \text{ euros}) = \frac{150}{1000} = \frac{3}{20}, \quad P(50 \text{ euros}) = \frac{40}{1000} = \frac{1}{25},$$

$$P(100 \text{ euros}) = \frac{10}{1000} = \frac{1}{100}.$$

- Supongamos que queremos saber la probabilidad que tenemos de perder dinero al comprar una botella. Como la variable aleatoria cuantifica la pérdida-ganancia que tenemos al comprar una botella, perderemos dinero cuando

$X < 0$, más formalmente, si al comprar la botella

el resultado del experimento aleatorio $\omega \in \Omega$

satisface que $X(\omega) < 0 \iff -\infty < X(\omega) < 0$.

Y esto lo cuantificamos en términos de la función de probabilidad P asociada a la variable aleatoria X por

$$\begin{aligned} P(-\infty, 0) &= P(\{\omega \in \Omega : X(\omega) \in (-\infty, 0)\}) \\ &= P(\{\omega \in \Omega : -\infty < X(\omega) < 0\}) = P(\{\text{inténtelo otra vez}\} \cup \{5 \text{ euros}\}) \\ &= P(\{\text{inténtelo otra vez}\}) + P(\{5 \text{ euros}\}) = \frac{1}{2} + \frac{3}{10} = \frac{8}{10}. \end{aligned}$$

- $P(-7, 3) = P(\{\omega \in \Omega : -7 < X(\omega) < 3\})$, esto vendría a cuantificar la probabilidad de que al comprar una botella no tuviésemos pérdidas menores o iguales que 7 euros y ganancia mayores o iguales que 3 euros.

$$\begin{aligned} P(-7, 3) &= P(\{\omega \in \Omega : -7 < X(\omega) < 3\}) = P(\{5 \text{ euros}\} \cup \{10 \text{ euros}\}) \\ &= P(5 \text{ euros}) + P(10 \text{ euros}) = \frac{3}{10} + \frac{3}{20} = \frac{9}{20}. \end{aligned}$$

- $P[40, \infty)$, vendría a cuantificar la probabilidad de que al comprar una botella ganásemos una cantidad mayor o igual que 40 euros.

$$\begin{aligned} P[40, \infty) &= P(\{\omega \in \Omega : X(\omega) \geq 40\}) = P(\{50 \text{ euros}\} \cup \{100 \text{ euros}\}) \\ &= P(50 \text{ euros}) + P(100 \text{ euros}) = \frac{1}{25} + \frac{1}{100} = \frac{1}{20}. \end{aligned}$$

- $P(-\infty, 20)$ cuantifica la probabilidad de que al comprar una botella o tengamos pérdidas o una ganancia inferior a 20 euros.

$$\begin{aligned} P(-\infty, 20) &= P(\{\omega \in \Omega : X(\omega) < 20\}) \\ &= P(\{\text{inténtelo otra vez}\} \cup \{5 \text{ euros}\} \cup \{10 \text{ euros}\}) \\ &= P(\{\text{inténtelo otra vez}\}) + P(\{50 \text{ euros}\}) + P(\{100 \text{ euros}\}) \\ &= \frac{1}{2} + \frac{3}{10} + \frac{3}{20} = \frac{11}{20}. \end{aligned}$$

- $P(6)$ cuantifica la probabilidad de que al comprar una botella tengamos una ganancia de 6 euros.

$$P(6) = P\{\omega \in \Omega : X(\omega) = 6\} = P(\emptyset) = 0.$$

- $P(40)$ cuantifica la probabilidad de que al comprar una botella tengamos una ganancia de 40 euros.

$$P(6) = P\{\omega \in \Omega : X(\omega) = 40\} = P(\{50 \text{ euros}\}) = \frac{1}{25}.$$

Ejemplo 3. El lanzamiento de un dado nos cuesta 2 euros. Si sacamos un múltiplo de tres, nos devuelven el dinero mientras que si obtenemos un múltiplo de 5 nos dan 5 euros.

Calcúlese la probabilidad de que al jugar una vez

1. obtengamos una ganancia de 3 euros;
2. no perdamos dinero;
3. perdamos dinero.

Experimento aleatorio: tirar un dado y ver el número que sale.

Espacio muestral, posibles resultados del experimento aleatorio,

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

Los sucesos elementales son equiprobables.

Vamos a cuantificar la ganancia o pérdida al tirar una vez el dado por la siguiente variable aleatoria:

$$\begin{aligned} X : \Omega &\longrightarrow \mathbb{R} \\ 1 &\longrightarrow X(1) = 0 - 2 = -2 \\ 2 &\longrightarrow X(2) = 0 - 2 = -2 \\ 3 &\longrightarrow X(3) = 2 - 2 = 0 \\ 4 &\longrightarrow X(4) = 0 - 2 = -2 \\ 5 &\longrightarrow X(5) = 5 - 2 = 3 \\ 6 &\longrightarrow X(6) = 2 - 2 = 0. \end{aligned}$$

1. Si designamos por P la función de probabilidad asociada a la variable aleatoria X , $P(3)$ nos dará la probabilidad de que al tirar el dado obtengamos una ganancia de 3 euros.

$$P(3) = P\{\omega \in \Omega : X(\omega) = 3\} = P(\{5\}) = \frac{1}{6}.$$

2. Para no perder dinero, ya que X cuantifica la ganancia o pérdida de dinero, tendremos que imponer que $X \geq 0$, y la probabilidad de no perder dinero será

$$\begin{aligned} P(X \geq 0) &= P(\{\omega \in \Omega : X(\omega) \geq 0\}) = P(\{3, 5, 6\}) \\ &= P(\{3\}) + P(\{5\}) + P(\{6\}) = \frac{3}{6} = \frac{1}{2}. \end{aligned}$$

3.

$$\begin{aligned} P(X < 0) &= P(\{\omega \in \Omega : X(\omega) < 0\}) = P(\{1, 2, 4\}) \\ &= \frac{\text{Cardinal de } \{1, 2, 4\}}{\text{Cardinal de } \Omega} = \frac{3}{6} = \frac{1}{2}. \end{aligned}$$

□

1.2. Operaciones con variables aleatorias

Sea $(\Omega, \mathcal{A}; P)$ un espacio probabilístico y $\alpha \in \mathbb{R}$. X_α definida por

$$\begin{aligned} X_\alpha : \Omega &\longmapsto \mathbb{R} \\ \omega &\longrightarrow X_\alpha(\omega) = \alpha \end{aligned}$$

es una variable aleatoria que suele designarse simplemente por la letra α .

Si $X, Y : \Omega \longmapsto \mathbb{R}$ son dos variables aleatorias, entonces:

$$\begin{aligned} X + Y : \Omega &\longmapsto \mathbb{R} \\ \omega &\longrightarrow (X + Y)(\omega) = X(\omega) + Y(\omega), \end{aligned}$$

$$\begin{aligned} XY : \Omega &\longmapsto \mathbb{R} \\ \omega &\longrightarrow (XY)(\omega) = X(\omega) \cdot Y(\omega), \end{aligned}$$

$$\begin{aligned} \frac{X}{Y} : \Omega &\longmapsto \mathbb{R} \\ \omega &\longrightarrow \left(\frac{X}{Y}\right)(\omega) = \frac{X(\omega)}{Y(\omega)}, \quad \text{con } Y(\omega) \neq 0 \quad \forall \omega \in \Omega, \end{aligned}$$

$$\begin{aligned} \max\{X, Y\} : \Omega &\longmapsto \mathbb{R} \\ \omega &\longrightarrow \max\{X, Y\}(\omega) = \max\{X(\omega), Y(\omega)\}, \end{aligned}$$

$$\begin{aligned} \min\{X, Y\} : \Omega &\longmapsto \mathbb{R} \\ \omega &\longrightarrow \min\{X, Y\}(\omega) = \min\{X(\omega), Y(\omega)\}, \end{aligned}$$

son variables aleatorias.

1.3. Función de distribución

Proposición 1. Sea (Ω, \mathcal{A}, P) un espacio probabilístico y $X : \Omega \mapsto \mathbb{R}$ una variable aleatoria. Si I_1 e I_2 son subconjuntos de \mathbb{R} , **no son sucesos**, y $I_1 \subseteq I_2$, entonces

$$P(I_2 - I_1) = P(I_2) - P(I_1) \quad (2)$$

Demostración

$$\begin{aligned} P(I_2 - I_1) &= P(\{\omega \in \Omega : X(\omega) \in I_2 \text{ y } X(\omega) \notin I_1\}) \\ &= P(\{\omega \in \Omega : X(\omega) \in I_2\} - \{\omega \in \Omega : X(\omega) \in I_1\}), \\ \text{Como } I_1 \subseteq I_2 &\implies \{\omega \in \Omega : X(\omega) \in I_1\} \subseteq \{\omega \in \Omega : X(\omega) \in I_2\}, \text{ y} \\ &P(\{\omega \in \Omega : X(\omega) \in I_2\} - \{\omega \in \Omega : X(\omega) \in I_1\}) \\ &= P(\{\omega \in \Omega : X(\omega) \in I_2\}) - P(\{\omega \in \Omega : X(\omega) \in I_1\}) = P(I_2) - P(I_1). \end{aligned}$$

■

Definición 3. Sea (Ω, \mathcal{A}, P) un espacio probabilístico y $X : \Omega \mapsto \mathbb{R}$ una variable aleatoria. Definimos la función de distribución asociada a la variable aleatoria X como la función

$$\begin{aligned} F_X : \mathbb{R} &\mapsto [0, 1] \\ x &\longrightarrow F_X(x) = P(-\infty, x] = P(\{\omega \in \Omega : X(\omega) \leq x\}) \end{aligned}$$

Observación 2. La función de distribución se introduce para conocer cómo se reparte la probabilidad de los valores que toma la variable aleatoria. Obsérvese de que si $\{\omega \in \Omega : X(\omega) \leq x\}$ no fuera un elemento de \mathcal{A} , $F_X(x)$ no estaría bien definida. Escribiremos F en lugar de F_X cuando no haya confusión posible, y también $P(X \leq x)$ en lugar de $P(\{\omega \in \Omega : X(\omega) \leq x\})$.

Es importante saber distinguir los conceptos de variable aleatoria y función de distribución. Dada una variable aleatoria, tenemos los valores reales asignados a cada uno de los elementos del espacio muestral, mientras que dada una función de distribución, tenemos únicamente cuáles son estos valores reales y cómo se reparten, o sea, tenemos la distribución de estos valores. Al pasar de una variable aleatoria a su distribución se pierde la información relacionada con los objetos que dan lugar a estos valores reales y que se recoge en el espacio de probabilidades. Es importante observar que dos variables aleatorias distintas pueden tener la misma función de distribución, como muestra el siguiente ejemplo.

Ejemplo 4. Dado el espacio probabilístico (Ω, \mathcal{A}, P) , de manera que $\Omega = \{\omega_1, \omega_2\}$, $\mathcal{A} = \mathcal{P}(\Omega)$ y P viene dada por

$$P(\omega_1) = P(\omega_2) = \frac{1}{2},$$

Consideramos las variables aleatorias

$$X(\omega) = \begin{cases} 0 & \text{si } \omega = \omega_1 \\ 1 & \text{si } \omega = \omega_2 \end{cases} \quad \text{y} \quad Y(\omega) = \begin{cases} 1 & \text{si } \omega = \omega_1 \\ 0 & \text{si } \omega = \omega_2 \end{cases}$$

Está claro que $X \neq Y$. Por otra parte

$$\{\omega \in \Omega : X(\omega) \leq x\} = \begin{cases} \emptyset & x < 0 \\ \{\omega_1\} & 0 \leq x < 1 \\ \Omega & x \geq 1 \end{cases} \quad \text{y} \quad \{\omega \in \Omega : Y(\omega) \leq y\} = \begin{cases} \emptyset & y < 0 \\ \{\omega_2\} & 0 \leq y < 1 \\ \Omega & y \geq 1 \end{cases}$$

y, por tanto,

$$F_X(x) = P(\{\omega \in \Omega : X(\omega) \leq x\}) = \begin{cases} 0 & x < 0 \\ \frac{1}{2} & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

y

$$F_Y(y) = P(\{\omega \in \Omega : Y(\omega) \leq y\}) = \begin{cases} 0 & y < 0 \\ \frac{1}{2} & 0 \leq y < 1 \\ 1 & y \geq 1 \end{cases}$$

es decir, las dos variables aleatorias tienen la misma función de distribución. \square

Ejemplo 5.

En el Ejemplo 1,

- $F_X(2)$ nos indica la probabilidad de que al comprar una botella tengamos pérdidas o una ganancia menor o igual que dos euros.

$$\begin{aligned} F_X(2) &= P(-\infty, 2] = P(\{\omega \in \Omega; X(\omega) \leq 2\}) = P(\{\text{otra vez}, 5 \text{ euros}, 10 \text{ euros}\}) \\ &= P(\{\text{otra vez}\}) + P(\{5 \text{ euros}\}) + P(\{10 \text{ euros}\}) = \frac{1}{2} + \frac{3}{10} + \frac{3}{20} = \frac{11}{20}. \end{aligned}$$

- $F_X(-3)$ nos indica la probabilidad de que al comprar una botella tengamos pérdidas menores o iguales que 3 euros.

$$\begin{aligned} F_X(-3) &= P(-\infty, -3] = P(\{\omega \in \Omega; X(\omega) \leq -3\}) = P(\{\text{otra vez}, 5 \text{ euros}\}) \\ &= P(\{\text{otra vez}\}) + P(\{5 \text{ euros}\}) = \frac{1}{2} + \frac{3}{10} = \frac{8}{10} = \frac{4}{5}. \end{aligned}$$

- $F_X(-37)$ nos indica la probabilidad de que al comprar una botella tengamos pérdidas menores o iguales que 37 euros.

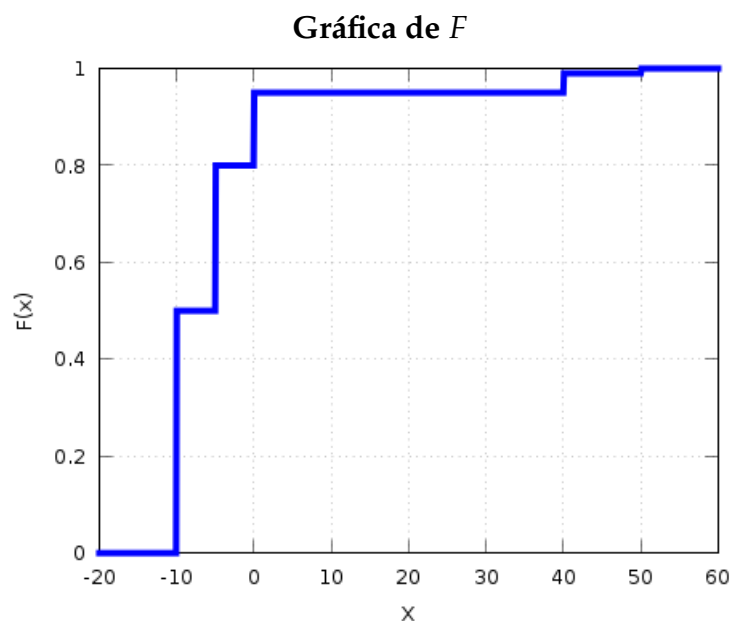
$$F_X(-37) = P(-\infty, -37] = P(\{\omega \in \Omega; X(\omega) \leq -37\}) = P(\emptyset) = 0.$$

- $F_X(82)$ nos indica la probabilidad de que al comprar una botella tengamos pérdidas o una ganancia menor o igual que 82 euros.

$$\begin{aligned} F_X(82) &= P(-\infty, 92] = P(\{\omega \in \Omega; X(\omega) \leq 92\}) \\ &= P(\{\text{otra vez}, 5 \text{ euros}, 10 \text{ euros}, 50 \text{ euros}, 100 \text{ euros}\}) \\ &= P(\{\text{otra vez}\}) + P(\{5 \text{ euros}\}) + P(\{10 \text{ euros}\}) + P(\{50 \text{ euros}\}) \\ &\quad + P(\{100 \text{ euros}\}) = \frac{1}{2} + \frac{3}{10} + \frac{3}{20} + \frac{1}{25} + \frac{1}{100} = 1. \end{aligned}$$

- En general

$$\begin{aligned} F(x) &= P(-\infty, x] \\ &= \begin{cases} P(\emptyset) = 0, & x < -10 \\ P(\text{otra vez}) = \frac{1}{2}, & -10 \leq x < -5 \\ P(\text{otra vez}) + P(5 \text{ euros}) = \frac{1}{2} + \frac{3}{10} = \frac{4}{5}, & -5 \leq x < 0 \\ P(\text{otra vez}) + \dots + P(10 \text{ euros}) = \frac{1}{2} + \dots + \frac{3}{20} = \frac{19}{20}, & 0 \leq x < 40 \\ P(\text{otra vez}) + \dots + P(50 \text{ euros}) = \frac{1}{2} + \dots + \frac{1}{25} = \frac{99}{100}, & 40 \leq x < 50 \\ P(\text{otra vez}) + \dots + P(100 \text{ euros}) = \frac{1}{2} + \dots + \frac{1}{100} = 1, & 50 \leq x. \end{cases} \end{aligned}$$



Ejercicio 1. Sea $(\Omega, \mathcal{A}; P)$ un espacio probabilístico. Calcule la función de distribución de la variable aleatoria

$$\begin{aligned} X_2 : \Omega &\longrightarrow \mathbb{R} \\ \omega &\longrightarrow X_2(\omega) = 2 \end{aligned}$$

Ejemplo 6. Consideremos el experimento aleatorio que consiste en tirar tres veces una moneda al aire. En este caso

$$\Omega = \{ccc, ccx, cxc, cxx, xcc, xc x, xxc, xxx\},$$

donde por ejemplo cxc significa "salir cara, cruz y cara en las tres tiradas".

Indicamos por X "número de caras obtenidas en las tres tiradas". Es claro que X es una variable aleatoria cuando $\mathcal{A} = \mathcal{P}(\Omega)$ y se cumple

$$\begin{aligned} X(ccc) &= 3 \\ X(ccx) &= X(cxc) = X(xcc) = 2 \\ X(cxx) &= X(xcx) = X(xxc) = 1 \\ X(xxx) &= 0 \end{aligned}$$

y

$$\{\omega \in \Omega : X(\omega) \leq x\} = \begin{cases} \emptyset & x < 0 \\ \{xxx\} & 0 \leq x < 1 \\ \{xxx, cxx, xc x, xxc\} & 1 \leq x < 2 \\ \{ccx, cxc, cxx, xcc, xc x, xxc, xxx\} & 2 \leq x < 3 \\ \Omega & x \geq 3 \end{cases}$$

La función de distribución de la variable aleatoria X es

$$F_X(x) = P(\{\omega \in \Omega : X(\omega) \leq x\}) = \begin{cases} 0 & x < 0 \\ \frac{1}{8} & 0 \leq x < 1 \\ \frac{4}{8} & 1 \leq x < 2 \\ \frac{7}{8} & 2 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$

Proposición 2. Propiedades de la Función de Distribución. Sea (Ω, \mathcal{A}, P) un espacio probabilístico, $X : \Omega \mapsto \mathbb{R}$ una variable aleatoria y $F \equiv F_X$ la función de distribución asociada a la variable aleatoria X . Se verifica

1. $\lim_{x \rightarrow -\infty} F(x) = 0, \quad \lim_{x \rightarrow \infty} F(x) = 1.$
 - $P(\emptyset) = \lim_{x \rightarrow -\infty} P(-\infty, x] = \lim_{x \rightarrow -\infty} F(x) = 0,$
 - $P(\mathbb{R}) = \lim_{x \rightarrow \infty} P(-\infty, x] = \lim_{x \rightarrow \infty} F(x) = 1.$
2. F es creciente, $x_1 \leq x_2 \implies F(x_1) \leq F(x_2).$
3. F es continua por la derecha, $\lim_{h \rightarrow 0, h > 0} F(x+h) = F(x).$

Observación 3. En general, dada una función $F : \mathbb{R} \mapsto [0, 1]$ que satisface

1. $\lim_{x \rightarrow -\infty} F(x) = 0, \quad \lim_{x \rightarrow \infty} F(x) = 1,$
2. F es creciente, .
3. F es continua por la derecha.

se dice **función de distribución**, ya que siempre es posible diseñar un espacio muestral, con un probabilidad y una variable aleatoria cuya función de distribución sea F .

Observación 4. La probabilidad de ciertos intervalos puede calcularse en términos de la función de distribución.

- Por ejemplo, si queremos calcular la probabilidad de que al comprar una botella no tengamos perdidas menores o iguales que 3 y unas ganancias mayores a 6 euros, tendremos que calcular $P(-3, 6]$.

Observamos que

$$(-3, 6] = (-\infty, 6] - (-\infty, -3], \quad (-\infty, -3] \subset (-\infty, 6],$$

entonces

$$\begin{aligned} P(-3, 6] &= P((-\infty, 6] - (-\infty, -3]) = P(-\infty, 6] - P(-\infty, -3] \\ &= F(6) - F(-3) = \frac{19}{20} - \frac{1}{2} = \frac{9}{20}. \end{aligned}$$

- Probabilidad de que al comprar una botella tengamos una ganancia mayor que 42 euros.

Tenemos que calcula

$$P(42, \infty) = P(\{\omega \in \Omega : X(\omega) > 42\}).$$

Observamos que

$$(42, \infty) = \mathbb{R} - (-\infty, 42], \quad \text{y se verifica} \quad (-\infty, 42] \subset \mathbb{R},$$

entonces

$$P(-\infty, 42] = P(\mathbb{R} - (-\infty, 42]) = P(\mathbb{R}) - P(-\infty, 42] = 1 - F_X(42) = 1 - \frac{99}{100} = \frac{1}{100}.$$

□

Ejercicio 2. En una universidad se ha observado que el 60 % de los estudiantes que se matriculan lo hacen en una carrera de Ciencias, mientras que el otro 40 % lo hacen en una carrera de Humanidades. Si un determinado día se realizan 20 matrículas, calcular la probabilidad de que:

1. haya el mismo número de matrículas en Ciencias que en Humanidades;
2. el número de matrícula en Ciencias sea menor que en Humanidades;
3. haya al menos ocho matrículas en Ciencias;
4. no haya mas de doce matrículas en Ciencias.
5. Si las cinco primeras matrículas son de Humanidades, calcular la probabilidad de que:
 - a) en total haya igual número de matrículas en Ciencias que en Humanidades;
 - b) en total haya al menos seis más en Ciencias que en Humanidades.

Escogemos un día en el que se han realizado 20 matriculas.

Experimento aleatorio: de entre los 20 matriculados, ver el número que se ha matriculado en Ciencias. (Si sabemos esto, también sabemos el número que se ha matriculado en Humanidades).

Un espacio muestral posible es

$$\Omega = \{\omega = (\omega_1, \omega_2, \dots, \omega_{20}) : \omega_i \in \{C, H\}, i = 1, 2, \dots, 20\}.$$

Un suceso elementales

$$\omega = (C, H, H, H, C, C, H, C, C, C, C, H, C, C, H, H, H, C, C, C),$$

y nos está indicando que el primer estudiante matriculado lo hizo en ciencias; los tres siguientes en Humanidades, el quinto y el sexto en Ciencias,....En total se han matriculado 12 estudiantes en Ciencias. Consideramos los sucesos

$$C_1 = (C, \omega_2, \omega_3, \dots, \omega_{20}),$$

este suceso indica que el primer matriculado del día lo ha hecho en ciencias, y los otros 19 no importa donde se hayan matriculado. El enunciado nos está diciendo que $P(C_1) = \frac{60}{100} = \frac{3}{5}$.

$$C_2 = (\omega_1, C, \omega_3, \dots, \omega_{20}),$$

este suceso indica que el segundo matriculado del día lo ha hecho en Ciencias, y los otros 19 no importa donde se hayan matriculado. El enunciado nos está diciendo que $P(C_2) = \frac{60}{100} = \frac{3}{5}$. Está claro que los suceso C_1 y C_2 son independientes.

$$C_{20} = (\omega_1, \omega_2, \omega_3, \dots, C),$$

este suceso indica que el último matriculado del día lo ha hecho en Ciencias, y los otros 19 no importa donde se hayan matriculado. El enunciado nos está diciendo que $P(C_{20}) = \frac{60}{100} = \frac{3}{5}$.

$$H_1 = (H, \omega_2, \omega_3, \dots, \omega_{20}),$$

este suceso indica que el primer matriculado del día lo ha hecho en Humanidades, y los otros 19 no importa donde se hayan matriculado. El enunciado nos está diciendo que $P(H_1) = \frac{40}{100} = \frac{2}{5}$.

$$H_2 = (\omega_1, H, \omega_3, \dots, \omega_{20}),$$

este suceso indica que el segundo matriculado del día lo ha hecho en Humanidades, y los otros 19 no importa donde se hayan matriculado. El enunciado nos está diciendo que $P(H_2) = \frac{40}{100} = \frac{2}{5}$. Está claro que los suceso H_1 y H_2 son independientes.

$$H_{20} = (\omega_1, \omega_2, \omega_3, \dots, H),$$

este suceso indica que el último matriculado del día lo ha hecho en Humanidades, y los otros 19 no importa donde se hayan matriculado. El enunciado nos está diciendo que $P(H_{20}) = \frac{60}{100} = \frac{3}{5}$.

Los sucesos

$$C_1, C_2, \dots, C_{20}, H_1, H_2, \dots, H_{20},$$

son independientes.

El suceso

$$\omega = (C, H, H, H, C, C, H, C, C, C, C, H, C, C, H, H, H, C, C, C),$$

lo podemos expresar como

$$\omega = C_1 \cap H_2 \cap H_3 \cap H_4 \cap C_5 \cap C_6 \cap H_7 \cap C_8 \cap C_9 \cap C_{10}$$

$$\cap C_{11} \cap H_{12} \cap C_{13} \cap C_{14} \cap H_{15} \cap H_{16} \cap H_{17} \cap C_{18} \cap C_{19} \cap C_{20},$$

al ser independientes los sucesos

$$\begin{aligned} P(\omega) &= P(C_1) \cdot P(H_2) \cdot P(H_3) \cdot P(H_4) \cdot P(C_5) \cdot P(C_6) \cdot P(H_7) \cdot P(C_8) \\ &\cdot P(C_9) \cdot P(C_{10}) \cdot P(C_{11}) \cdot P(H_{12}) \cdot P(C_{13}) \cdot P(C_{14}) \cdot P(H_{15}) \cdot P(H_{16}) \\ &\cdot P(H_{17}) \cdot P(C_{18}) \cdot P(C_{19}) \cdot P(C_{20}) = \left(\frac{3}{5}\right)^{12} \cdot \left(\frac{2}{5}\right)^8. \end{aligned}$$

El espacio probabilístico no es equiprobable. Tomemos dos sucesos elementales

$$\omega^1 = (C, C, H, H, H, H, H, H, H, H, H, H, H, H, H, H, H, H, H, H),$$

y

$$\omega^2 = (C, H, H, H, H, C, H, H, C, C, H, H, H, H, H, H, H, H, H, H),$$

se tiene que

$$P(\omega^1) = \left(\frac{3}{5}\right)^2 \cdot \left(\frac{2}{5}\right)^{18} \neq \left(\frac{3}{5}\right)^4 \cdot \left(\frac{2}{5}\right)^{16} = P(\omega^2).$$

Definimos la variable aleatoria

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

$$\omega \longrightarrow X(\omega) = \text{número de matriculados en Ciencias que hay en } \omega,$$

si

$$\omega = (C, H, H, H, C, C, H, C, C, C, C, H, C, C, H, H, H, C, C, C),$$

entonces $X(\omega) = 12$ y $X(\omega^1) = 2, X(\omega^2) = 4$.

Vamos a calcular la función de probabilidad asociada a la variable aleatoria X , esto es, queremos calcular

$$P(i) = P(X = i) = \text{probabilidad de que se hayan matriculado en Ciencias } i \text{ persona.}$$

Empezamos con algunos ejemplos. Empecemos calculando la probabilidad de que se hayan matriculado en ciencias solo una persona.

$$P(X = 1) = P(\{\beta \in \Omega : X(\beta) = 1\}),$$

$$\{\beta \in \Omega : X(\beta) = 1\} = \bigcup_{i=1}^{20} \beta^i,$$

donde

$$\beta^1 = (C, H, H, H, H, H, H, H, H, H, H, H, H, H, H, H, H, H, H, H),$$

$$\beta^2 = (H, C, H, H, H, H, H, H, H, H, H, H, H, H, H, H, H, H, H, H),$$

...

$$\beta^{20} = (H, H, H, H, H, H, H, H, H, H, H, H, H, H, H, H, H, H, H, C),$$

entonces, como los β^i , $i = 1, 2, \dots, 20$ son disjuntos dos a dos,

$$P(X = 1) = P(\{\beta \in \Omega : X(\beta) = 1\})$$

$$= \sum_{i=1}^{20} P(\beta^i) = 20 \left(\frac{3}{5}\right)^1 \cdot \left(\frac{2}{5}\right)^{19} = \binom{20}{1} \cdot \left(\frac{3}{5}\right)^1 \cdot \left(\frac{2}{5}\right)^{19}.$$

Calculamos ahora la probabilidad de que se hayan matriculado solo dos personas en Ciencias. Un caso sería:

$$\gamma = (H, H, H, H, H, H, H, H, C, H, H, H, H, H, H, H, H, H, H, C),$$

y

$$P(\gamma) = \left(\frac{3}{5}\right)^2 \cdot \left(\frac{2}{5}\right)^{18}.$$

En γ el décimo y el último se han matriculado en Ciencias, y el resto en Humanidades. También podría darse el caso de que se hubiesen matriculado en Ciencias el sexto y noveno y el resto en Humanidades. El total de casos en el que se examinan dos en Ciencias y 18 2n Humanidades es $\binom{20}{2}$, y

$$P(X = 2) = \binom{20}{2} \cdot \left(\frac{3}{5}\right)^2 \cdot \left(\frac{2}{5}\right)^{18}.$$

En el caso general

$$P(X = j) = \binom{20}{j} \cdot \left(\frac{3}{5}\right)^j \cdot \left(\frac{2}{5}\right)^{20-j}, \quad j = 1, 2, \dots, 20.$$

Vamos a calcular la función de distribución $F_X \equiv F$ asociada a la variable aleatoria X .

$$F(x) = P(-\infty, x] = P(\{\omega \in \Omega : X(\omega) \leq x\})$$

$$= \begin{cases} 0, & x < 0 \\ \left(\frac{2}{5}\right)^{20} = \binom{20}{0} \cdot \left(\frac{2}{5}\right)^{20}, & 0 \leq x < 1 \\ \binom{20}{0} \cdot \left(\frac{2}{5}\right)^{20} + \binom{20}{1} \cdot \left(\frac{3}{5}\right)^1 \cdot \left(\frac{2}{5}\right)^{19} = \sum_{j=0}^1 \binom{20}{j} \cdot \left(\frac{3}{5}\right)^j \cdot \left(\frac{2}{5}\right)^{20-j}, & 1 \leq x < 2 \\ \sum_{j=0}^2 \binom{20}{j} \cdot \left(\frac{3}{5}\right)^j \cdot \left(\frac{2}{5}\right)^{20-j}, & 2 \leq x < 3 \\ \vdots & \vdots \\ \sum_{j=0}^{18} \binom{20}{j} \cdot \left(\frac{3}{5}\right)^j \cdot \left(\frac{2}{5}\right)^{20-j}, & 18 \leq x < 19 \\ \sum_{j=0}^{19} \binom{20}{j} \cdot \left(\frac{3}{5}\right)^j \cdot \left(\frac{2}{5}\right)^{20-j}, & 19 \leq x < 20 \\ \sum_{j=0}^{20} \binom{20}{j} \cdot \left(\frac{3}{5}\right)^j \cdot \left(\frac{2}{5}\right)^{20-j} = \left(\frac{3}{5} + \frac{2}{5}\right)^{20} = 1^{20} = 1, & x \geq 20. \end{cases}$$

1. Como hay 20 matriculados, tendremos que calcular la probabilidad de que 10 se hayan matriculado en Ciencias.

$$P(X = j) = \binom{20}{10} \cdot \left(\frac{3}{5}\right)^{10} \cdot \left(\frac{2}{5}\right)^{10} = 0,11714.$$

2. Tendremos que calcular $P(-\infty, 10)$. Sea 10^- un número tan próximo a 10 como queramos, mayor que 9 y menor que 10. Entonces

$$P(-\infty, 10) = F(10^-) = \sum_{j=0}^9 \binom{20}{j} \cdot \left(\frac{3}{5}\right)^j \cdot \left(\frac{2}{5}\right)^{20-j}.$$

3. Tenemos que calcular

$$P[8, \infty) = P(X \geq 8) = P(\{\omega \in \Omega : X(\omega) \geq 8\}).$$

El complementario de $\{\omega \in \Omega : X(\omega) \geq 8\}$ es $\{\omega \in \Omega : X(\omega) \leq 7\}$, entonces

$$\begin{aligned} P[8, \infty) &= 1 - P(\{\omega \in \Omega : X(\omega) \leq 7\}) = 1 - P(-\infty, 7] = 1 - F(7) \\ &= 1 - \sum_{j=0}^7 \binom{20}{j} \cdot \left(\frac{3}{5}\right)^j \cdot \left(\frac{2}{5}\right)^{20-j}. \end{aligned}$$

4. Tenemos que calcular $P(-\infty, 12]$.

$$P(-\infty, 12] = \sum_{j=0}^{12} \binom{20}{j} \cdot \left(\frac{3}{5}\right)^j \cdot \left(\frac{2}{5}\right)^{20-j}.$$

5. Consideramos el suceso:

$$5H = \text{"las cinco primeras matrículas son de Humanidades"} = H_1 \cap H_2 \cap H_3 \cap H_4 \cap H_5.$$

Ya que los sucesos son independientes, tenemos

$$P(5H) = P(H_1 \cap H_2 \cap H_3 \cap H_4 \cap H_5) = P(H_1) \cdot P(H_2) \cdot P(H_3) \cdot P(H_4) \cdot P(H_5) = \left(\frac{2}{5}\right)^5.$$

a) Nos encontramos ante una probabilidad condicionada:

$$P(X = 10|5H) = \frac{P(\{X = 10\} \cap 5H)}{P(5H)}.$$

El suceso $\{X = 10\} \cap 5H$, los cinco primeros matriculados son de Humanidades y entre los quince restantes 10 tiene que haberse matriculado en Ciencias.

$$\begin{aligned} P(X = 10|5H) &= \frac{P(\{X = 10\} \cap 5H)}{P(5H)} = \frac{\left(\frac{2}{5}\right)^5 \cdot \binom{15}{10} \cdot \left(\frac{3}{5}\right)^{10} \cdot \left(\frac{2}{5}\right)^5}{\left(\frac{2}{5}\right)^5} \\ &= \binom{15}{10} \cdot \left(\frac{3}{5}\right)^{10} \cdot \left(\frac{2}{5}\right)^5. \end{aligned}$$

b)

$$\begin{aligned} P(\{X \geq 6\}|5H) &= \frac{P(\{X \geq 6\} \cap 5H)}{P(5H)} = \frac{P(\{\cup_{j=6}^{20} \{X \geq j\}\} \cap 5H)}{P(5H)} \\ &= \frac{P(\cup_{j=6}^{20} \{\{X \geq j\} \cap 5H\})}{P(5H)} = \frac{\sum_{j=6}^{15} P(\{X \geq j\} \cap 5H)}{P(5H)} \\ &= \sum_{j=6}^{15} \binom{15}{j} \cdot \left(\frac{3}{5}\right)^j \cdot \left(\frac{2}{5}\right)^{15-j}. \end{aligned}$$

□

1.4. Concepto de esperanza matemática

Dada una variable aleatoria X definida sobre un espacio muestral Ω numerable, su esperanza matemática es el número $E(X)$ dado por la fórmula

$$E(X) = \sum_{\omega \in \Omega} X(\omega) \cdot P(\{\omega\}),$$

siempre que esta serie sea absolutamente convergente, es decir, siempre que

$$\sum_{\omega \in \Omega} |X(\omega)| \cdot P(\{\omega\}),$$

sea convergente. Se dice entonces que existe la esperanza de la variable aleatoria X .

Cuando se habla de *esperanza matemática* es corriente omitir el adjetivo. Se utilizan también otros términos, como *valor esperado*, *media* o *valor medio*. Para que un valor medio muestral se parezca a la esperanza matemática, es necesario que el tamaño de la muestra sea suficientemente grande.

Por ejemplo, si se lanza una moneda, y se gana 1 euro cuando sale cara y nada cuando sale cruz, nunca se obtendrá el valor esperado de 0,5 euros. Sin embargo, lanzando la moneda un gran número n de veces, si se puede esperar conseguir alrededor de $n/2$ euros con buena probabilidad.

Evidentemente, cuando Ω sea un espacio finito, siempre existirá la esperanza de la variable aleatoria X , pero cuando Ω sea infinito numerable, la condición $\sum_{\omega \in \Omega} |X(\omega)| \cdot P(\{\omega\}) < \infty$, es esencial, porque permite calcular la esperanza por cualquier procedimiento de reordenación y asociación de sumandos de $\sum_{\omega \in \Omega} X(\omega) \cdot P(\{\omega\})$, sin temor a que se obtengan resultados contradictorios.

La esperanza matemática es una generalización del concepto de media aritmética. Dada una muestra de valores observados x_1, x_2, \dots, x_n de una variable aleatoria con sus respectivas frecuencias f_1, f_2, \dots, f_n , sabemos que la media de la muestra es

$$\bar{x} = \frac{1}{N} \sum_{i=1}^n x_i \cdot f_i = \sum_{i=1}^n x_i \cdot \frac{f_i}{N},$$

donde $N = \sum_{i=1}^n f_i$.

Las frecuencias relativas f_i/N se pueden considerar como las probabilidades que tienen los valores x_i de presentarse en la muestra total de tamaño N . Poniendo entonces

$$P(X = x_i) = P(\{\omega \in \Omega : X(\omega) = x_i\}) = \frac{f_i}{N}, \quad i = 1, 2, \dots, n,$$

tenemos

$$\bar{x} = \sum_{i=1}^n x_i \cdot P(X = x_i),$$

que coincide con el caso de que Ω sea finito.

Al pensar que el concepto de esperanza está estrechamente relacionado a los juegos de azar, y por tanto, con la misma definición de probabilidad, quizás sorprenda al estudiante que suprimiendo la condición $\sum_{\omega \in \Omega} |X(\omega)| \cdot P(\{\omega\}) < \infty$, puedan producirse sorpresas. El siguiente ejemplo pone de manifiesto este hecho.

Consideramos el siguiente juego: lanzamos una moneda la aire hasta que sale cara. Si sale cara en la primera tirada, la banca paga al jugador 2 euros. Si sale cara en la segunda tirada el jugador recibe 2^2 euros. Si sale cara por primera vez en la tirada n -ésima el jugador gana 2^n euros. Es claro que la cantidad de euros que puede ganar el jugador es una variable aleatoria X . ¿Cuál es el valor esperado de X ? ¿Estaríamos dispuestos a pagar este valor por avanzado para participar en este juego?

Los valores que puede tomar X son $x_1, x_2, \dots, x_n, \dots$ con $x_n = 2^n$. Además, la probabilidad del valor x_n corresponde a la probabilidad del suceso de que en n tiradas salga cara por primera vez en el último lanzamiento

$$P(X = x_n) = \left(\frac{1}{2}\right)^n.$$

Por lo tanto la esperanza de X viene dada por

$$E(X) = \sum_{n=1}^{\infty} x_n \cdot P(X = x_n) = \sum_{n=1}^{\infty} 2^n \cdot \frac{1}{2^n} = 1 + 1 + \dots + 1 + \dots = \infty,$$

de este modo hemos encontrado una variable aleatoria para la que no existe la esperanza matemática o que su valor esperado es infinito.

1.5. Clasificación de las variables aleatorias

Sea (Ω, \mathcal{A}, P) un espacio probabilístico y $X : \Omega \mapsto \mathbb{R}$ una variables aleatoria. Esta puede ser:

- **Discreta:** $X(\Omega)$ es un conjunto finito o numerable.
- **Continua:** $X(\Omega)$ puede se una unión de intervalos de \mathbb{R} , que pueden ser acotados o no acotados.
- **Mixta:** es una combinación de las dos anteriores.

1.6. Ejercicios

Ejercicio 3. Un dado es lanzado 5 veces. Consideramos como espacio muestral asociado al experimento aleatoria

$$\Omega = \{\omega = (\omega_1, \omega_2, \omega_3, \omega_4, \omega_5) : \omega_i \in \{1, 2, 3, 4, 5, 6\}, i = 1, 2, 3, 4, 5\}.$$

Consideramos la variable aleatoria que considerara la suma de los valores de las caras de los 5 dados ($X(\omega) = \sum_{i=1}^5 \omega_i$). Determine los conjuntos

- a) $\{\omega \in \Omega : X(\omega) = 4\};$
- b) $\{\omega \in \Omega : X(\omega) = 6\};$
- c) $\{\omega \in \Omega : X(\omega) = 30\};$
- d) $\{\omega \in \Omega : X(\omega) \geq 29\}.$

Ejercicio 4. Sea (Ω, \mathcal{A}, P) un espacio probabilístico y $X : \Omega \mapsto \mathbb{R}$ una variable aleatoria. Sea

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

la función de distribución asociada a la variable aleatoria X .

- a) Calcule k .
- b) Calcule la probabilidad de $[1/2, 1]$, ($= \{\omega \in \Omega : X(\omega) \in [1/2, 1]\}$), condicionada por $[1/2, 2]$, ($= \{\omega \in \Omega : X(\omega) \in [1/2, 2]\}$).

Solución: 1. k cualquier número real en $(0, 1]$; 2. 1.

2. Variables aleatorias discretas

Supongamos que realizamos un experimento aleatorio con espacio muestral Ω , y sea $X : \Omega \mapsto \mathbb{R}$ una variable aleatoria. Diremos que X es una **variable aleatoria discreta** si solo toma un número finito o numerables de valores. Esto significa que existe una sucesión de números reales diferentes, puede ser finita o no,

$$\{x_1, x_2, \dots, x_n, \dots\},$$

tal que

$$\text{rango de } X = \{x_1, x_2, \dots, x_n, \dots\}.$$

Definición 4. Sea (Ω, \mathcal{A}, P) un espacio probabilístico y $X : \Omega \mapsto \mathbb{R}$ una variable aleatoria discreta con

$$\text{rango de } X = \{x_1, x_2, \dots, x_n, \dots\}.$$

Se llama **función de probabilidad** asociada a la variable aleatoria X a

$$P_X(x_k) = P_X(X = x_k) \equiv P(X = x_k) = P(\{\omega \in \Omega : X(\omega) = x_k\}), \quad k \in \{1, 2, \dots\}$$

Observación 5. Algunos autores la llaman función de densidad de probabilidad, o simplemente función de densidad.

Ejemplo 7.

1. La variable aleatoria considerada en el Ejemplo 1,

$$X : \Omega \mapsto \mathbb{R} \text{ definida por } \begin{cases} X(\text{inténtelo otra vez}) = 0 - 10 = -10 \\ X(5 \text{ euros}) = 5 - 10 = -5 \\ X(10 \text{ euros}) = 10 - 10 = 0 \\ X(50 \text{ euros}) = 50 - 10 = 40 \\ X(100 \text{ euros}) = 100 - 10 = 90, \end{cases}$$

solamente toma 5 valores, es decir $X(\Omega) = \{-10, -5, 0, 40, 90\}$ es un conjunto finito de 5 elementos, por lo tanto X es una variable aleatoria discreta.

Ejemplos de "elementos" de la función de probabilidad asociada a la variable aleatoria X :

$$P_X(-5) = P_X(X = -5) = P(\{\omega \in \Omega : X(\omega) = -5\}) = P(\{5 \text{ euros}\}) = \frac{3}{10},$$

$$P_X(37) = P_X(X = 37) = P(\{\omega \in \Omega : X(\omega) = 37\}) = P(\emptyset) = 0,$$

en general

$$P_X(X = k) = \begin{cases} 0 & \text{si } k \notin \{-10, -5, 0, 40, 90\} \\ \frac{1}{2} & \text{si } k = -10 \\ \frac{3}{10} & \text{si } k = -5 \\ \frac{3}{20} & \text{si } k = 0 \\ \frac{1}{25} & \text{si } k = 40 \\ \frac{1}{100} & \text{si } k = 90. \end{cases}$$

2. En Ejemplo 3, la variable aleatoria

$$\begin{aligned} X : \Omega &\mapsto \mathbb{R} \\ 1 &\longrightarrow X(1) = 0 - 2 = -2 \\ 2 &\longrightarrow X(2) = 0 - 2 = -2 \\ 3 &\longrightarrow X(3) = 2 - 2 = 0 \\ 4 &\longrightarrow X(4) = 0 - 2 = -2 \\ 5 &\longrightarrow X(5) = 5 - 2 = 3 \\ 6 &\longrightarrow X(6) = 2 - 2 = 0. \end{aligned}$$

toma 3 valores $X(\Omega) = \{-2, 0, 3\}$, y en consecuencia, también será discreta.

La función de probabilidad es

$$P_X(X = k) = \begin{cases} 0 & \text{si } k \notin \{-2, 0, 3\} \\ \frac{1}{2} & \text{si } k = -2 \\ \frac{1}{3} & \text{si } k = 0 \\ \frac{1}{6} & \text{si } k = 3. \end{cases}$$

3. En Ejercicio 2, la variable aleatoria

$$\begin{aligned} X : \Omega &\mapsto \mathbb{R} \\ \omega &\longrightarrow X(\omega) = \text{número de matriculados en Ciencias que hay en } \omega, \end{aligned}$$

toma un número finito de valores (21), $X(\Omega) = \{0, 1, 2, \dots, 19, 20\}$, y por lo tanto también es discreta.

La función de probabilidad es

$$P_X(X = k) = \begin{cases} 0 & \text{si } k \notin \{0, 1, 2, \dots, 19, 20\} \\ \binom{20}{j} \cdot \left(\frac{3}{5}\right)^j \cdot \left(\frac{2}{5}\right)^{20-j} & \text{si } k \in \{0, 1, 2, \dots, 19, 20\}. \end{cases}$$

2.1. Distribución de Bernoulli y binomial

Supongamos que realizamos el experimento aleatorio de tirar un dado. El espacio muestral es

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

Consideramos los sucesos

$E = \text{Éxito} = \text{"obtener una puntuación de 2 o 3 puntos"},$

$F = \text{Fracaso} = \Omega - E = \text{"obtener una puntuación de 1, 4, 5 o 6 puntos"}.$

Definimos la variable aleatoria

$$\begin{aligned} X : \Omega &\longmapsto \mathbb{R} \\ \omega &\longrightarrow X(\omega) = \begin{cases} 1 & \text{si obtenemos éxito} \\ 0 & \text{si obtenemos fracaso} \end{cases} \end{aligned}$$

La distribución de probabilidad de esta variable aleatoria es

$$P(X = 1) = \frac{2}{6} = \frac{1}{3} = p, \quad P(X = 0) = \frac{4}{6} = \frac{2}{3} = 1 - p,$$

y se tiene

$$E[X] = 1 \cdot P(X = 1) + 0 \cdot P(X = 0) = 1 \cdot \frac{1}{3} = \frac{1}{3} = p,$$

$$V(X) = 1^2 \cdot P(X = 1) + 0^2 \cdot P(X = 0) - p^2 = p - p^2 = p(1 - p).$$

Consideremos ahora una situación más general.

Experimento aleatorio: Tirar cinco veces un dado y vamos viendo en cada tirada el resultado que sale.

Espacio muestral $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Nos planteamos la probabilidad de tener 3 éxitos.

Definimos las variables aleatorias

$$\begin{aligned} X_1 : \Omega &\longmapsto \mathbb{R} \\ \omega &\longrightarrow X_1(\omega) = \begin{cases} 1 & \text{si obtenemos éxito en la primera tirada} \\ 0 & \text{si obtenemos fracaso en la primera tirada} \end{cases} \end{aligned}$$

$$\begin{aligned} X_2 : \Omega &\longmapsto \mathbb{R} \\ \omega &\longrightarrow X_2(\omega) = \begin{cases} 1 & \text{si obtenemos éxito en la segunda tirada} \\ 0 & \text{si obtenemos fracaso en la segunda tirada} \end{cases} \end{aligned}$$

... ..

$$\begin{aligned} X_5 : \Omega &\longmapsto \mathbb{R} \\ \omega &\longrightarrow X_5(\omega) = \begin{cases} 1 & \text{si obtenemos éxito en la quinta tirada} \\ 0 & \text{si obtenemos fracaso en la quinta tirada} \end{cases} \end{aligned}$$

Definimos la variable aleatoria

$$X = X_1 + X_2 + X_3 + X_4 + X_5,$$

y consideramos los sucesos $S = \text{“obtener tres éxitos”}$

$$= \{\omega \in \Omega : X(\omega) = X_1(\omega) + X_2(\omega) + X_3(\omega) + X_4(\omega) + X_5(\omega) = 3\},$$

$E_i = \text{"obtener éxito en la } i\text{-ésima tirada"}$

$$= \{\omega \in \Omega : X_i(\omega) = 1\}, \quad i = 1, 2, 3, 4, 5.$$

(En este suceso, nos da igual el resultado obtenido en la j -ésima tirada, cuando $i \neq j$.)

$F_i = \text{"obtener fracaso en la } i\text{-ésima tirada"}$

$$= \{\omega \in \Omega : X_i(\omega) = 0\}, \quad i = 1, 2, 3, 4, 5.$$

. (En este suceso, nos da igual el resultado obtenido en la j -ésima tirada, cuando $i \neq j$).

Se verifica

$$P(E_i) = p = \frac{1}{3}, \quad P(F_i) = 1 - p = \frac{2}{3}, \quad i = 1, 2, 3, 4, 5.$$

Los sucesos E_i , $i = 1, 2, 3, 4, 5$, son independientes. F_i , $i = 1, 2, 3, 4, 5$, también son independientes. E_2 , F_1 , F_3 , F_4 y F_5 son independientes. E_2 y F_2 son dependientes.

$S_{245} = \text{"obtener éxito solo en la segunda, cuarta y quinta tirada"}$

$$P(S_{245}) = P(F_1 \cap E_2 \cap F_3 \cap E_4 \cap E_5).$$

Ya que los sucesos F_1 , E_2 , F_3 , E_4 y E_5 son independientes

$$\begin{aligned} P(S_{245}) &= P(F_1) \cdot P(E_2) \cdot P(F_3) \cdot P(E_4) \cdot P(E_5) \\ &= (1 - p) \cdot p \cdot (1 - p) \cdot p \cdot p = p^3(1 - p)^2. \end{aligned}$$

Es claro que si tomamos otra posibilidad, por ejemplo el suceso S_{135} , se tiene que

$$P(S_{245}) = P(S_{135}).$$

El número de posibilidades de obtener tres éxitos es

$$C_5^3 = \frac{V_5^3}{P_3} = \frac{5 \cdot 4 \cdot 2}{3 \cdot 2 \cdot 1} = \binom{5}{3}.$$

$$P(S) = P(X_1 + X_2 + X_3 + X_4 + X_5 = 3) = \binom{5}{3} P(S_{245}) = \binom{5}{3} p^3(1 - p)^2.$$

Definición 5. Sea (Ω, \mathcal{A}, P) un espacio probabilístico asociado a un cierto experimento aleatorio y $E \in \mathcal{A}$ un suceso con $P(E) = p$. Al suceso E lo vamos a llamar suceso "Éxito" a su complementario $F = E^c = \Omega - E = \text{suceso "Fracaso"}$, el cual satisface que $P(F) = 1 - p$. Realizamos el experimento aleatorio n veces y para $i \in \{1, 2, \dots, n\}$ definimos las variables aleatorias

$$\begin{aligned} X_i : \Omega &\longrightarrow \mathbb{R} \\ \omega &\longrightarrow X_i(\omega) = \begin{cases} 1 & \text{si obtenemos éxito en la prueba } i\text{-ésima} \\ 0 & \text{si obtenemos fracaso en la prueba } i\text{-ésima.} \end{cases} \end{aligned}$$

La distribución asociada a la variable aleatoria

$$X = X_1 + X_2 + \cdots + X_n,$$

es conocida como la **distribución binomial, de parámetros** $n \in \mathbb{N}$ y $p \in [0,1]$, y designada por $B(n, p)$.

Observación 6. $B(1, p)$ es conocida como la **distribución de Bernoulli**.

Proposición 3. Sea (Ω, \mathcal{A}, P) un espacio probabilístico asociado a un cierto experimento aleatorio, $E = \text{"Éxito"} \in \mathcal{A}$ un suceso con $P(E) = p$ y $F = E^c = \Omega - E = \text{"Fracaso"}$ con $P(F) = 1 - p$. Consideramos las variables aleatorias

$$\begin{aligned} X_i : \Omega &\longrightarrow \mathbb{R} & i &\in \{1, 2, \dots, n\} \\ \omega &\longrightarrow X_i(\omega) = \begin{cases} 1 & \text{si obtenemos éxito en la prueba } i\text{-ésima} \\ 0 & \text{si obtenemos fracaso en la prueba } i\text{-ésima.} \end{cases} \end{aligned}$$

La distribución binomial asociada a la variable aleatoria

$$X = X_1 + X_2 + \cdots + X_n$$

- Es una variable aleatoria discreta.
- Su función de probabilidad es

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, 2, \dots, n \quad (3)$$

$P(X = k)$ lo interpretamos como **la probabilidad de que al realizar n veces el experimento aleatorio obtengamos k éxitos**.

Ejemplo 8. Suponiendo que la probabilidad de que un recién nacido sea niña es 0,51, hallar la probabilidad de que una familia de 6 hijos tenga:

1. por lo menos un niña,
2. por lo menos una niño,
3. por lo menos dos niños y un niña.

Experimento aleatorio: Ver si el recién nacido es niño o niña.

$\Omega = \{\text{niño}, \text{niña}\}$.

Consideramos los sucesos

$E = \text{Éxito} = \text{niña}$, con probabilidad $P(\text{niña}) = p = 0,51$,
 $F = \text{Fracaso} = \text{niño}$, con probabilidad $P(\text{niño}) = 1 - p = 0,49$.
 Para $i \in 1, 2, 3, 4, 5, 6$ definimos la variable aleatoria

$$\begin{aligned} X_i : \Omega &\longrightarrow \mathbb{R} \\ \omega &\longrightarrow X_i(\omega) = \begin{cases} 1 & \text{el } i\text{-ésimo hijo es una niña} \\ 0 & \text{el } i\text{-ésimo hijo es un niño} \end{cases} \end{aligned}$$

y sea $X = X_1 + X_2 + X_3 + X_4 + X_5 + X_6$.

La distribución de probabilidad se distribuye según una $B(6, p = 0,51)$.

1.

$$\begin{aligned} P(X \geq 1) &= P(\{X = 0\}^c) = 1 - P(X = 0) = 1 - \binom{6}{0} p^0 (1-p)^6 \\ &= 1 - (0,49)^6 = 0,986 \end{aligned}$$

2.

$$\begin{aligned} P(X \leq 5) &= P(\{X = 6\}^c) = 1 - P(X = 6) = 1 - \binom{6}{6} p^6 (1-p)^0 \\ &= 1 - (0,51)^6 = 0,982 \end{aligned}$$

3.

$$\begin{aligned} P(1 \leq X \leq 4) &= P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) \\ &= \binom{6}{1} (0,51)^1 (0,49)^5 + \binom{6}{2} (0,51)^2 (0,49)^4 + \binom{6}{3} (0,51)^3 (0,49)^3 \\ &\quad + \binom{6}{4} (0,51)^4 (0,49)^2 \end{aligned}$$

Ejercicio 5. El gerente de un restaurante que solo da servicio mediante reservas, sabe por experiencia, que el 20 % de las personas que reservan una mesa no asistirán. Si el restaurante acepta 25 reservas pero solo dispone de 20 mesas, ¿cuál es la probabilidad de que a todas las personas que asistan al restaurante se le asigne una mesa?

Indicación: Considere las variables aleatorias

$$X_i(\omega) = \begin{cases} 1 & \text{la reserva } i\text{-ésima ha asistido} \\ 0 & \text{en caso contrario} \end{cases}, \quad i = 1, 2, \dots, 25,$$

y $X = X_1 + X_2 + \dots + X_{25}$.

Solución: 0,58

Ejercicio 6. Cada muestra de aire tiene 10 % de posibilidades de contener una molécula rara particular. Suponga que las muestras son independientes con respecto a la presencia de la molécula rara. Encuentre la probabilidad de que en las siguientes 18 muestras, exactamente dos contengan la molécula rara.

Indicación: Considere las variables aleatorias

$$X_i(\omega) = \begin{cases} 1 & \text{contiene molécula rara e la } i\text{-ésima muestra} \\ 0 & \text{en caso contrario} \end{cases}, \quad i = 1, 2, \dots, 18,$$

y $X = X_1 + X_2 + \dots + X_{18}$.

Solución: 0,284

2.2. Medidas de centralización y dispersión

En estadística descriptiva se definen unas medidas de centralización y dispersión con el objetivo de obtener una idea de la muestra. Vamos a definir los conceptos análogos para variables aleatorias, los cuales nos van a permitir resumir el reparto de probabilidad correspondiente a la variable aleatoria.

Definición 6. Sea (Ω, \mathcal{A}, P) un espacio probabilístico y $X : \Omega \mapsto \mathbb{R}$ una variable aleatoria discreta. Supongamos que $X(\Omega) = \{x_1, x_2, \dots, x_n, \dots\}$ consta de n elementos. **La media o esperanza** de la variable aleatoria X viene definida por

$$\mu = E[X] = \sum_i x_i P(X = x_i)$$

En **Ejemplo 1** la variable aleatoria toma los valores $\{x_1, x_2, x_3, x_4, x_5\} = \{-10, -5, 0, 40, 90\}$,

$$E[X] = \sum_{i=1}^5 x_i P(x_i)$$

$$\begin{aligned} &= (-10) \cdot P(-10) + (-5) \cdot P(-5) + 0 \cdot P(0) + 40 \cdot P(40) + 90 \cdot P(90) \\ &= -10 \cdot \frac{1}{2} + -5 \cdot \frac{3}{10} + 0 \cdot \frac{3}{20} + 40 \cdot \frac{1}{25} + 90 \cdot \frac{1}{100} = -4. \end{aligned}$$

Ejemplo 9. Realizamos el experimento aleatorio de tirar un dado y ver la puntuación que obtenemos. Sea X la variable aleatoria que indica la puntuación que hemos obtenido en la tirada. Calcule la esperanza de X .

Solución

El espacio muestral es $\Omega = \{1, 2, 3, 4, 5, 6\}$ y la variable aleatoria X está definida por

$$\begin{aligned} X : \Omega &\longrightarrow \mathbb{R} \\ j &\longrightarrow X(j) = j \end{aligned}$$

X es una variable aleatoria discreta, toma los valores $\{1, 2, 3, 4, 5, 6\}$, y su función de probabilidad es

$$P(X = j) = \frac{1}{6}, \quad j \in \{1, 2, 3, 4, 5, 6\}.$$

La esperanza es

$$E[X] = \sum_{j=1}^6 jP(X = j) = \frac{1}{6} \sum_{j=1}^6 j = \frac{1 + 2 + 3 + 4 + 5 + 6}{6} = \frac{21}{6} = \frac{7}{2}.$$

□

Ejercicio 7. Calcule la esperanza de la variable aleatoria de Bernoulli.

Solución

Sea (Ω, \mathcal{A}, P) un espacio probabilístico asociado a un cierto experimento aleatorio y $E \in \mathcal{A}$ un suceso con $P(E) = p$. Al suceso E lo vamos a llamar suceso "Éxito" a su complementario $F = E^c = \Omega - E$ = suceso "Fracaso", el cual satisface que $P(F) = 1 - p$. Realizamos el experimento aleatorio y definimos la variable aleatoria

$$\begin{aligned} X : \Omega &\longrightarrow \mathbb{R} \\ \omega &\longrightarrow X(\omega) = \begin{cases} 1 & \text{si obtenemos éxito} \\ 0 & \text{si obtenemos fracaso} \end{cases} \end{aligned}$$

La distribución de probabilidad de la variable aleatoria X , **distribución de Bernoulli**, es

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

Tenemos entonces que

$$E[X] = 0 \cdot P(X = 0) + 1 \cdot P(X = 1) = p.$$

□

Proposición 4. Propiedades de la esperanza matemática de variables aleatorias discretas

Sea (Ω, \mathcal{A}, P) un espacio probabilístico y X, Y dos variables aleatorias discretas para las que existe $E[X]$ y $E[Y]$. Se verifica:

1. Si $\alpha \in \mathbb{R}$ y X_α es la variable aleatoria que toma el valor α en todo elemento de Ω , se tiene $E[X_\alpha] = \alpha$.

2. $E[\alpha X] = E[X_\alpha X] = \alpha E[X]$.
3. $|E[X]| \leq E[|X|]$.
4. $E[X + Y] = E[X] + E[Y]$.
5. Si $X \leq Y$ entonces $E[X] \leq E[Y]$, en particular si $X \geq 0$ se tiene que $E[X] \geq 0$.

Ejercicio 8. Calcule la esperanza matemática de una distribución binomial $B(5, p)$.

Ejemplo 10. En el juego de la ruleta se hace girar una bola encima de una rueda circular dividida en 37 arcos de la misma longitud, numerados del 0 al 36. Suponemos que la probabilidad de que ocurra un arco es la misma para todos y, por tanto, la bola puede caer en cualquier número del 0 al 36 con una probabilidad de $1/37$. Supongamos que jugamos a números impares y que la apuesta se hace a dos por uno, es decir, si aportamos 1 euro y sale impar, recibimos 2 euros (incluida la apuesta), y no cobramos nada si sale par. ¿Qué esperamos ganar si apostamos continuamente a número impares?

Solución

Sea X la variable aleatoria que indica la cantidad que uno puede ganar o perder si apuesta un euro. Es claro que X puede tomar dos valores: $X = 1$ si sale impar; y $X = -1$ si sale par o cero (hay que recordar que la banca se queda con la apuesta si sale cero). Además

$$P(X = 1) = \frac{18}{37}, \quad P(X = -1) = \frac{19}{37}.$$

Como consecuencia tenemos

$$E[X] = 1 \cdot P(X = 1) + (-1) \cdot P(X = -1) = -\frac{1}{37}.$$

La esperanza de X representa en este caso la ganancia o pérdida media por apuesta. Si hacemos n apuestas de 1 euro al número impar, la ganancia o pérdida media sería la esperanza de la variable aleatoria

$$Y = \frac{X_1 + X_2 + \cdots + X_n}{n},$$

cunado n tiende a infinito, donde para $i = 1, 2, \dots, n$

$$X_i = \begin{cases} 1 & \text{ha salido impar en la apuesta } i\text{-ésima} \\ -1 & \text{ha salido par o cero en la apuesta } i\text{-ésima} \end{cases}$$

$$E[Y] = \frac{E[X_1] + E[X_2] + \cdots + E[X_n]}{n} = \frac{-\frac{1}{37}n}{n} = -\frac{1}{37}.$$

□

Definición 7. Sea (Ω, \mathcal{A}, P) un espacio probabilístico y $X : \Omega \mapsto \mathbb{R}$ una variable aleatoria discreta. Supongamos que $X(\Omega) = \{x_1, x_2, \dots, x_n\}$ consta de n elementos. **La mediana** de la variable aleatoria X es el valor o valores de $\{x_1, x_2, \dots, x_n\}$ que dejan a izquierda y derecha la misma cantidad de probabilidad.

En el caso del **Ejemplo 1** la mediana es -10.

La esperanza o valor esperado de una variable aleatoria describe donde está centrada su distribución de probabilidad, y por lo tanto es un valor que caracteriza a dicha distribución de probabilidad. Sin embargo, la esperanza por si sola no da una descripción adecuada de la forma de la distribución, es necesario saber como se dispersan los valores de la variable aleatoria respecto al valor esperado.

Dada una muestra de valores observados x_1, x_2, \dots, x_n de una variable aleatoria X con sus respectivas frecuencias f_1, f_2, \dots, f_n , la dispersión de un valor x_i respecto a la media $\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$ se puede definir por

$$(x_i - \bar{x})^2,$$

y la media de esta dispersión viene dada por

$$s^2 = \sum_{i=1}^n (x_i - \bar{x})^2 \cdot \frac{f_i}{N},$$

donde $N = \sum_{i=1}^n f_i$. Las frecuencias relativas $\frac{f_i}{N}$ se pueden considerar como las probabilidades que tienen los valores x_i de representarse en la muestra total de tamaño N . Si escribimos entonces

$$P(X = x_i) = \frac{f_i}{N}, \quad i = 1, 2, \dots, n,$$

tenemos

$$s^2 = \sum_{i=1}^n (x_i - \bar{x})^2 \cdot P(X = x_i),$$

que se llama varianza de la muestra. Extendemos este concepto al caso de una variable aleatoria

Definición 8. Sea (Ω, \mathcal{A}, P) un espacio probabilístico y $X : \Omega \mapsto \mathbb{R}$ una variable aleatoria discreta. Supongamos que $X(\Omega) = \{x_1, x_2, \dots, x_n, \dots\}$. **La varianza** de la variable aleatoria X viene definida por

$$\sigma^2 = V(X) = \sum_i (x_i - \mu)^2 P(X = x_i)$$

Observación 7. Si $X(\Omega) = \{x_1, x_2, \dots, x_n\}$, es fácil ver que

$$\sigma^2 = V(X) = \sum_{i=1}^n x_i^2 P(X = x_i) - \mu^2.$$

En Ejemplo 1

$$\begin{aligned} \sigma^2 &= \sum_{i=1}^5 x_i^2 P(x_i) - \mu^2 \\ &= (-10)^2 \cdot P(-10) + (-5)^2 \cdot P(-5) + (0)^2 \cdot P(0) + (40)^2 \cdot P(40) + (90)^2 \cdot P(90) - (-4)^2 \\ &= 100 \cdot \frac{1}{2} + 25 \cdot \frac{3}{10} + 0 \cdot \frac{3}{20} + 160 \cdot \frac{1}{25} + 810 \cdot \frac{1}{100} - 16 = 50. \end{aligned}$$

Proposición 5. Propiedades de la varianza Sea (Ω, \mathcal{A}, P) un espacio probabilístico y X e Y dos variables aleatorias discreta con varianza finita. Se tiene

1. $V(X) = E[X^2] - (E[X])^2$.
2. Si $\alpha \in \mathbb{R}$ y $X_\alpha(\omega) = \alpha$ para todo $\omega \in \Omega$, se verifica que $V(X_\alpha) = V(\alpha) = 0$.
3. $V(\alpha X) = \alpha^2 V(X)$, con $\alpha \in \mathbb{R}$.
4. $V(\alpha X + \beta) = \alpha^2 V(X)$, con $\alpha, \beta \in \mathbb{R}$.

Ejemplo 11. Si X designa el número de puntos obtenidos en una tirada de un dado, calcule $V(X)$.

Solución

La variable aleatoria X es discreta y toma los valores $1, 2, \dots, 6$ con probabilidad $1/6$. Entonces

$$\begin{aligned} E[X] &= \sum_{i=1}^6 i \cdot P(X = i) = \frac{1 + 2 + \dots + 6}{6} = \frac{7}{2}, \\ E[X^2] &= \sum_{i=1}^6 i^2 \cdot P(X = i) = \frac{1^2 + 2^2 + \dots + 6^2}{6} = \frac{91}{6}, \end{aligned}$$

y

$$V(X) = E[X^2] - (E[X])^2 = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}.$$

□

Definición 9. Sea (Ω, \mathcal{A}, P) un espacio probabilístico y $X : \Omega \rightarrow \mathbb{R}$ una variable aleatoria discreta. **La desviación típica** σ de la variable aleatoria X se define como la raíz cuadrada de la varianza σ^2 .

En Ejemplo 1

$$\sigma = \sqrt{50} \cong 7,07.$$

2.3. Distribución de Poisson

La distribución de Poisson también se conoce como distribución de los sucesos raros. Inicialmente la distribución de Poisson surgió como un límite de una distribución binomial, pero hoy en día se utiliza para caracterizar ciertos sucesos aleatorios.

Empecemos viéndola como un límite de $B(n; p)$ cuando $n \rightarrow \infty$, $p \rightarrow 0$ y $np \rightarrow \lambda \in (0, \infty)$.

No es difícil comprobar que

$$\lim_{n \rightarrow \infty, p \rightarrow 0, np \rightarrow \lambda} \binom{n}{k} p^k (1-p)^{n-k} = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

Definición 10. La distribución de Poisson de parámetro λ , es la que tiene como distribución de probabilidad

$$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k = 0, 1, 2, \dots$$

¿Cuándo nos vamos a encontrar con el Modelo de Poisson? Lo hemos obtenido al hacer $n \rightarrow \infty$ y $p \rightarrow 0$, por lo tanto cuando nos encontremos con un modelo binomial donde n es muy grande y p es pequeño, lo sustituimos por el correspondiente modelo de Poisson de parámetro $\lambda = np$.

Desde el punto de vista práctico, podemos pensar en utilizarlo cuando

$$n \geq 30 \quad \text{y} \quad p \leq 0,1$$

Ejemplo 12. La probabilidad de que un individuo tenga una reacción alérgica al inyectarle un suero es $p = 0,001$. Hallar la probabilidad de que, entre 2000 individuos, tengan reacción alérgica:

1. exactamente tres,
2. más de dos.

Experimento aleatorio: inyectarle suero a una persona de una cierta población y ver si tiene reacción alérgica.

Los sucesos elementales son

s = "tiene reacción alérgica", con una probabilidad $P(s) = p = 0,001$,

n = "no tiene reacción alérgica", con una probabilidad de $P(n) = 1 - p = 0,999$.

Tomamos 2000 individuos, y para $i = 1, 2, \dots, 2000$ definimos la variable aleatoria

$$X_i : \Omega \longrightarrow \mathbb{R}$$

$$\omega \longrightarrow X_i(\omega) = \begin{cases} 1 & \text{el individuo } i\text{-ésimo tiene reacción alérgica} \\ 0 & \text{el individuo } i\text{-ésimo no tiene reacción alérgica} \end{cases}$$

Consideramos la variable aleatoria

$$X = \sum_{i=1}^{2000} X_i$$

1.

$$P(X = 3) = \binom{2000}{3} p^3 (1-p)^{1997} = \frac{2000 \cdot 1999 \cdot 1998}{3 \cdot 2 \cdot 1} (0,001)^3 (0,999)^{1997}$$

$$= 0,18053$$

Vamos a tomar un modelo de Poisson de parámetro $\lambda = n \cdot p = 2000 \cdot 0,001 = 2$.

$$P(X = 3) = \frac{e^{-2} 2^3}{3!} = 0,18044$$

2. Este apartado lo hacemos utilizando el modelo de Poisson.

$$P(X > 2) = P(\{X = 0, 1, 2\}^c) = 1 - P(\{X = 0, 1, 2\})$$

$$= 1 - P(X = 0) - P(X = 1) - P(X = 2)$$

$$= 1 - 0,1352 - 0,2707 - 0,2707 = 0,3234$$

2.3.1. Esperanza y varianza de la distribución de Poisson

Proposición 6. Sea (Ω, \mathcal{A}, P) un espacio probabilístico y $X : \Omega \longrightarrow \mathbb{R}$ una variable aleatoria que sigue una distribución de Poisson de parámetro λ .

■ **Esperanza y varianza.**

$$E[X] = \lambda, \quad V(X) = \lambda. \quad (4)$$

■ Sean X_1, X_2, \dots, X_n n variables aleatorias linealmente independientes. Si X_i sigue una distribución de Poisson de parámetro λ_i , $i = 1, 2, \dots, n$, la variable aleatoria

$$X = X_1 + X_2 + \dots + X_n,$$

sigue una distribución de Poisson de parámetro $\lambda = \lambda_1 + \lambda_2 + \dots + \lambda_n$, $P(\lambda)$.

Ejemplo 13. El número de erratas por página en un libro se supone que sigue una distribución de Poisson. En una muestra de 95 páginas se han observado las siguientes frecuencias

Número de erratas:	0	1	2	3	4	5
Frecuencia:	40	30	15	7	2	

Hallar la probabilidad de que en una página tomada al azar haya alguna errata.

Nuestra variable aleatoria que sigue una distribución de Poisson es:

$X = \text{"número de erratas por página"}$

Tenemos un problema, y es que no conocemos el parámetro de la distribución de Poisson. Ya que $E[X] = \lambda$, parece normal tomar como parámetro la media de la muestra que nos han dado.

$$\lambda \cong \frac{40 \cdot 0 + 30 \cdot 1 + 15 \cdot 2 + 7 \cdot 3 + 2 \cdot 4 + 1 \cdot 5}{95} = 0,989 \cong 1.$$

Entonces

$$P(X \geq 1) = 1 - P(X = 0) = 0,6321$$

Ejercicio 9. Un canal de comunicación recibe impulsos independientes a razón de 200 impulsos por microsegundo. Calcular las probabilidades de los siguientes sucesos:

1. No hay ningún error en un microsegundo.
2. Hay exactamente un error en un microsegundo.
3. Hay al menos un error en un microsegundo.
4. Hay exactamente dos errores en un microsegundo.

En los siguientes ejemplos vamos a ver como las distribuciones de Poisson van a caracterizar a las variables aleatorias que miden el número de veces que se da un evento en una unidad de tiempo o espacio.

Ejercicio 10. A una gasolinera llegan, en media, 3 coches por minuto. Calcular la probabilidad de que:

1. En 1 minuto lleguen exactamente dos coches.

2. En 1 minuto lleguen al menos 2 coches.
3. En 5 minutos lleguen exactamente 12 coches.

En este ejercicio estamos ante los que se conoce como un **proceso de Poisson**.

El experimento aleatorio: **contar el número de eventos que ocurre a lo largo de un continuo (unidad de tiempo o espacio)**.

En nuestro ejercicio, el evento es que llegue a una gasolinera un coche en el apartado 1 y 2 en **unidad de tiempo de 1 minuto**.

Otras situaciones:

- Podíamos haber tomado como unidad de tiempo el día, el evento: que llegue a una gasolinera en un día. En este caso el experimento aleatorio: **contar el número de coches que llega a una gasolinera en un día**.

- **Evento:** llamada telefónica que entra a una centralita en 1 hora. La unidad de tiempo aquí es la hora.

Experimento aleatorio: contar el número de llamadas de teléfono que entran a una centralita en 1 hora.

- **Evento:** estrella en una cuadrícula del firmamento. Aquí la unidad de espacio es una cuadrícula del firmamento.

Experimento aleatorio: contar el número de estrellas que hay en una cuadrícula del firmamento.

- **Evento:** defecto en una plancha de metal. Aquí la unidad de espacio es la plancha de metal.

Experimento aleatorio: contar el número de defecto que hay en una plancha de metal.

Un experimento aleatorio de este tipo es un **proceso de Poisson**, si los sucesos cumplen las condiciones:

1. Los sucesos ocurren de una manera aleatoria e independiente.

El suceso S_4 = "que entren exactamente dos coches en 1 minuto"

2. El número medio de eventos que ocurre en una unidad de tiempo o espacio es constante.

En el ejercicio es 3.

El espacio muestral que podemos tomar asociado a un experimento aleatorio de este tipo es

$$\Omega = \{0, 1, 2, 3, \dots\}.$$

Consideramos la variable aleatoria

$$\begin{aligned} X : \Omega &\longrightarrow \mathbb{R} \\ n &\longrightarrow X(n) = n = \text{el evento ha ocurrido } n \text{ veces.} \end{aligned}$$

Si λ el número medio de eventos que ocurren en una unidad de tiempo o espacio en un proceso de Poisson, (en nuestro ejercicio $\lambda = 3$), la variable aleatoria X sigue una **distribución de Poisson de parámetro λ** , que designamos por $P(\lambda)$, y cuya función de probabilidad es

$$P(X = n) = \frac{\lambda^n}{n!} e^{-\lambda}, \quad n = 0, 1, 2, \dots \quad (5)$$

1. En este y siguiente apartado, la variable aleatoria X que cuenta el número de coches que llega a la gasolinera en un minuto sigue una $P(3)$, y la probabilidad de que lleguen exactamente dos coches en un minuto es

$$P(X = 2) = \frac{3^2}{2!} e^{-3} = 0'22.$$

2. Que en 1 minuto lleguen al menos dos coches es el suceso complementario de que en 1 minuto lleguen menos de 2 coches.

$$\begin{aligned} P(X \leq 2) &= 1 - P(X < 1) = 1 - P(\{X = 0\} \cup \{X = 1\}) = 1 - P(X = 0) - P(X = 1) \\ &= 1 - \frac{3^0}{0!} e^{-3} - \frac{3^1}{1!} e^{-3} = 0'8. \end{aligned}$$

3. Supongamos que $k > 0$ y definimos la variable aleatoria

$Y =$ " número de veces que ocurre el evento en k unidades de tiempo o espacio.

(En nuestro caso 5 minutos son 5 unidades de tiempo de 1 minuto, $k = 5$).

Si el número medio de veces que ocurre el evento en una unidad de tiempo (o espacio) es λ , el número medio de veces que ocurre el evento en k unidades de tiempo es $k\lambda$. La variable aleatoria Y **sigue una distribución de Poisson $P(k\lambda)$** .

En el ejercicio, Y es la variable aleatoria que cuenta el número de coches que llegan a la gasolinera en 5 minutos. Esta variable aleatoria seguirá una $P(15)$, y la probabilidad de que en 5 minutos lleguen 12 coches es

$$P(Y = 12) = \frac{15^{12}}{12!} e^{-15} = 0'08.$$

□

Ejercicio 11. Una empresa electrónica observa que el número de componentes que fallan en un cierto aparato antes de cumplir 100 horas de funcionamiento es una variable aleatoria de Poisson. Si el número medio de estos fallos es ocho,

1. ¿cual es la probabilidad de que falle una componente en 25 horas?;
2. ¿y de que fallen no más de dos componentes en 50 horas?;
3. ¿cual es la probabilidad de que fallen al menos diez en 125 horas?

Se X la variable aleatoria que cuenta el número de componentes que fallan en 100 horas (unidad de tiempo). Como el número medio de fallos es $\lambda = 8$, esta variable aleatoria va a seguir una distribución de Poisson $P(8)$.

1. Sea Y la variable aleatoria que mide el número de componentes que fallan en 25 horas de funcionamiento. 25 horas es un cuarto de la unidad de tiempo de 100 horas, $25 = \frac{1}{4} \cdot 100$, el número medio de fallos en 25 horas debe de ser $\frac{1}{4} \cdot 8 = 2$. La variable aleatoria Y sigue una distribución de Poisson $P(2)$ y

$$P(Y = 1) = \frac{2^1}{1!} e^{-2} = 0'27.$$

2. Solución: Si Z es la variable aleatoria que mide el número de componentes que fallan en 50 horas

$$P(Z \leq 2) = P(Z = 0) + P(Z = 1) + P(Z = 2) = \frac{4^0}{0!} e^{-4} + \frac{4^1}{1!} e^{-4} + \frac{4^2}{2!} e^{-4} = 0'24.$$

3. Solución: Si W es la variable aleatoria que mide el número de componentes que fallan en 125 horas

$$P(W \geq 10) = 1 - P(W \leq 9) = 1 - \sum_{i=0}^9 \frac{10^i}{i!} e^{-10} = 0'42.$$

□

Ejercicio 12. La contaminación constituye un problema en la fabricación de discos de almacenamiento óptico. El número de partículas de contaminación que ocurre en un disco óptico por centímetro cuadrado de superficie del disco, en media, es de 0,1. El área de un disco bajo estudio es de 100 centímetros cuadrados.

1. Encuentre la probabilidad de que ocurran 12 partículas en el área del disco bajo estudio.
2. La probabilidad de que ocurran cora partículas en el área del disco.
3. Determine la probabilidad de que ocurran 12 o menos partículas en el área del disco.

Si $\lambda = 0,1$ en el número medio de partículas en la unidad de superficie de un centímetro cuadrado, $100\lambda = 10$ será el número medio de partículas en la unidad de superficie de 100 centímetros cuadrados, o en el disco.

Consideramos la variable aleatoria

$X =$ número de partículas de contaminación en el disco.

La variable aleatoria X sigue una distribución de Poisson $P(10)$.

1. $P(X = 12) = \frac{10^{12}}{12!} e^{-10} = 0,095$
2. $P(X = 0) = \frac{10^0}{0!} e^{-10} = 0,0000454$
3. $P(x \leq 12) = P(X = 0) + P(X = 1) + \dots + P(X = 12) = \sum_{i=0}^{12} \frac{10^i}{i!} e^{-10}$

□

Ejercicio 13. En un libro de 500 páginas se distribuyen aleatoriamente 300 erratas de imprenta. Hallar la probabilidad de que en una página haya:

1. exactamente dos erratas,
2. dos o mas erratas.

Sea $X =$ "número de erratas que hay en una página del libro". Aquí, la unidad de espacio es la página del libro.

Como hemos se han distribuido de manera aleatoria 300 erratas, la media de erratas por página es de $\frac{300}{500} = 0,6$ erratas. La variable aleatoria X sigue una distribución de Poisson de parámetro $\lambda = \frac{300}{500} = 0,6$.

1. $P(X = 2) = \frac{e^{-0,6} \cdot (0,6)^2}{2!} = 0,098$.
2. $P(X \geq 2) = 1 - P(X < 1) = 1 - P(X = 0) - P(X = 1) = 0,1219$.

□

Ejercicio 14. Los mensajes que llegan a una computadora utilizada como servidor lo hacen con una tasa promedio de 0.1 mensajes por minuto.

1. ¿Cuál es la probabilidad de que lleguen como mucho 2 mensajes en una hora?
2. Determine el intervalo de tiempo necesario para que la probabilidad de que no llegue ningún mensaje en ese tiempo sea de 0.8

La variable aleatoria X = "mensajes que llegan en la unidad de tiempo de 1 minuto" sigue una distribución de Poisson de parámetro $\lambda = 0,1$, $P(0,1)$.

1. La variable aleatoria X_{60} = "número de mensajes que llegan en la unidad de tiempo de una hora", sigue una distribución de Poisson de parámetro $\lambda = 60 \times 0,1 = 6$, $P(6)$.

La probabilidad de que lleguen, como mucho, dos mensajes en una hora es

$$\begin{aligned} P(X_{60} \leq 2) &= P(X_{60} = 0) + P(X_{60} = 1) + P(X_{60} = 2) \\ &= \frac{e^{-6}6^0}{0!} + \frac{e^{-6}6^1}{1!} + \frac{e^{-6}6^2}{2!} = 0,062. \end{aligned}$$

2. Consideremos la variable aleatoria X_t = "número de mensajes que llegan en la unidad de tiempo t minutos". Esta sigue una distribución de Poisson de parámetro $\lambda = t \times 0,1 = 0,1t$, $P(0,1t)$. Si la probabilidad para que no llegue ningún mensaje en t minutos es 0,8, se tendrá que verificar

$$\begin{aligned} P(X_t = 0) &= 0,8 \iff \frac{e^{-0,1t}(0,1t)^0}{0!} = 0,8 \iff e^{-0,1t} = 0,8 \\ &\iff -0,1t = \log 0,8 \implies t = -10 \log 0,8 = 2,231 \text{ minutos.} \end{aligned}$$

□

Ejercicio 15. El número de pinchazos medio en los neumáticos de cierto vehículo industrial es de 0.3 por cada 50.000 km. Si el vehículo recorre 100.000 km, se pide:

1. probabilidad de que no tenga pinchazos,
2. probabilidad de que tenga menos de tres pinchazos,
3. número de km. recorridos para que la probabilidad de que no tenga ningún pinchazo sea 0.4066.

Solución: 1. 0,5488; 2. 0,9767; 3. 150,000km.

2.4. Ejercicios

Ejercicio 16. Sea (Ω, \mathcal{A}, P) un espacio probabilístico, $A \in \mathcal{A}$ un suceso y I_A la variable aleatoria definida, indicador de A , por

$$I_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$$

Calcule la esperanza de I_A .

Solución $P(A)$.

Ejercicio 17. La función de probabilidad de una variable aleatoria discreta está dada por

$X=$	1	3	5	7	9
$P(X=)$	0.2	0.1	0.3	0.3	0.1

1. Función de distribución.
2. Esperanza matemática.
3. Desviación típica y varianza.

Solución: 2. 5; 3. varianza 6,4.

Ejercicio 18. Una urna contiene 10 bolas de las que 8 son blancas. Se sacan dos al azar. Sea X la variable aleatoria que designa el número de bolas blancas obtenido. Calcule la función de probabilidad y distribución de la variable aleatoria X .

Solución 1. $P(X=0) = \frac{1}{45}$, $P(X=1) = \frac{16}{45}$, $P(X=2) = \frac{28}{45}$.

Ejercicio 19. Sea X la variable aleatoria que indica el número total de puntos obtenidos en n tiradas de un dado. Calcule la esperanza de la variable aleatoria X .

Solución $\frac{7n}{2}$.

Ejercicio 20. Sea (Ω, \mathcal{A}, P) un espacio probabilístico con $\Omega = \{1, 2, 3, 4, 5, 6, 7, 8\}$ y X una variable aleatoria que tiene por función de probabilidad

$$P(X=j) = kj, \quad j \in \Omega.$$

1. Calcule k .
2. Función de distribución de la variable aleatoria X .
3. $P(X \leq 5)$ y $P(2 < X \leq 5)$.

Solución 1. $k = \frac{1}{36}$; 3. $\frac{15}{36}$, $\frac{12}{36}$.

Ejercicio 21. Se considera una moneda trucada, para la que la probabilidad de salir cara es $\frac{1}{3}$ y la de salir cruz es $\frac{2}{3}$. Se lanza la moneda tres veces y sea X la variable aleatoria que representa las veces que sale cara. Calcúlese

1. $P(X = i)$, $i = 0, 1, 2, 3$.
2. Media de X .
3. Desviación típica de X .

Solución: 1. $P(X = 0) = \frac{8}{27}$, $P(X = 1) = \frac{4}{9}$, $P(X = 2) = \frac{2}{9}$, $P(X = 3) = \frac{1}{27}$; 2. 1; 3. $\sigma = \sqrt{\frac{2}{3}}$.

Ejercicio 22. En una urna U_1 hay dos bolas blancas y tres negras y en otra urna U_2 hay dos bolas negras. Seleccionamos una urna al azar de la que extraemos dos bolas sin reemplazamiento. Sea X la variable aleatoria "número de bolas negras extraídas". Calcular la función de probabilidad de X , su valor esperado y la probabilidad de que hayamos escogido la urna U_1 si el número de bolas negras extraídas ha resultado ser 2.

Ejercicio 23. En un puesto de feria se ofrece la posibilidad de lanzar a ciegas un dardo a unos globos. Si se consigue reventar un globo, se recibe un premio igual a una cantidad oculta tras el globo. Supongamos que la probabilidad de acertar con algún globo es $1/3$.

Los premios se distribuyen de la siguiente manera:

- 40 % de premios de 0,50 euros
- 30 % de premios de 1 euros
- 20 % de premios de 2 euros
- 10 % de premios de 6 euros

Si cada lanzamiento cuesta 1 euro, ¿cuál es la "ganancia esperada del dueño del puesto en cada lanzamiento?

Experimento aleatorio: Lanzar un dardo, y si reventamos un globo ver el premio que hemos conseguido.

Espacio muestral:

$$\{0, 0.5, 1, 2, 6\},$$

donde

0: no hemos conseguido premio.

0.5 : hemos obtenido un premio de 0.5 euros.

1 : hemos obtenido un premio de 1 euros.

2 : hemos obtenido un premio de 2 euros.

6 : hemos reventado un globo y obtenido un premio de 6 euros.

Si definimos el suceso

R: hemos reventado un globo,
el enunciado nos dice

$$P(R) = \frac{1}{3}, \quad P(0,5|R) = \frac{4}{10}, \quad P(1|R) = \frac{3}{10}, \quad P(2|R) = \frac{2}{10}, \quad P(6|R) = \frac{1}{10}.$$

Calculamos las probabilidades de los sucesos elementales.

$$P(0) = P(R^c) = 1 - P(R) = 1 - \frac{1}{3} = \frac{2}{3},$$

$$P(0,5) = P(0,5 \cap R) = P(0,5|R) \cdot P(R) = \frac{4}{10} \cdot \frac{1}{3} = \frac{4}{30},$$

$$P(1) = P(1 \cap R) = P(1|R) \cdot P(R) = \frac{3}{10} \cdot \frac{1}{3} = \frac{3}{30},$$

$$P(2) = P(2 \cap R) = P(2|R) \cdot P(R) = \frac{2}{10} \cdot \frac{1}{3} = \frac{2}{30},$$

$$P(6) = P(6 \cap R) = P(6|R) \cdot P(R) = \frac{1}{10} \cdot \frac{1}{3} = \frac{1}{30}.$$

Definimos la variable aleatoria $X = \text{"ganancia del dueño"}$

$$X : \Omega \mapsto \mathbb{R}, \quad \text{definida por} \quad \begin{cases} X(0) = 1 - 0 = 1 \\ X(0,5) = 1 - 0,5 = 0,5 \\ X(1) = 1 - 1 = 0 \\ X(2) = 1 - 2 = -1 \\ X(6) = 1 - 6 = -5. \end{cases}$$

La variable aleatoria X es discreta y toma los valores

$$\{x_1, x_2, x_3, x_4, x_5\} = \{1, 0,5, 0, -1, -5\}.$$

Su función de probabilidad $P_X \equiv P$ es

$$P(k) = P(X = k) = \begin{cases} 0, & \text{si } k \notin \{-5, -1, 0, 0,5, 1\} \\ \frac{1}{30}, & \text{si } k = -5 \\ \frac{2}{30}, & \text{si } k = -1 \\ \frac{3}{30}, & \text{si } k = 0 \\ \frac{4}{30}, & \text{si } k = 0,5 \\ \frac{2}{30}, & \text{si } k = 1. \end{cases}$$

La ganancia esperada por el dueño es la esperanza de la variable aleatoria X .

$$\begin{aligned} E[X] &= 1 \cdot P(X = 1) + 0,5 \cdot P(X = 0,5) + 0 \cdot P(X = 0) - 1 \cdot P(X = -1) - 5 \cdot P(X = -5) \\ &= 1 \cdot \frac{2}{3} + 0,5 \cdot \frac{4}{30} + 0 \cdot \frac{3}{30} - 1 \cdot \frac{2}{30} - 5 \cdot \frac{1}{30} = \frac{1}{2}. \end{aligned}$$

La ganancia esperada por el dueño en cada tiro es de 0,5 euros.

□

3. Variables aleatorias continuas

Consideramos un experimento aleatorio con espacio muestral asociado Ω . Decimos que una variable aleatoria $X : \Omega \mapsto \mathbb{R}$ es **continua** si X puede tomar cualquier valor de una unión de intervalos finitos o infinitos.

Ejemplo 14. Para llegar a un centro comercial hay que tomar un tren en una estación A, llegar a otra estación B y luego andar durante 7 minutos hasta el centro comercial. De la estación A a la B, parte un tren cada 20 minutos, y emplea 13 minutos en hacer el trayecto. Realizamos el siguiente experimento aleatorio: elegir una persona al azar, que va al centro comercial, en el momento que llega a la estación A, y medir el tiempo que espera en tomar el tren.

El espacio muestral sería el conjunto de los posibles minutos, segundos,... que se emplea hasta tomar el tren. Es lógico que puede ser cualquier tiempo entre 0 y 20 minutos. $\Omega = [0, 20]$.

Definimos la variable aleatoria $X =$ tiempo empleado desde que se entra en la estación A hasta que se llega al centro comercial,

$$\begin{aligned} X : \Omega = [0, 20] &\mapsto \mathbb{R} \\ \omega &\longrightarrow X(\omega) = \omega + 13 + 7 = \omega + 20. \end{aligned}$$

La variable aleatoria X puede tomar cualquier valor en el intervalo $[20, 40]$, por lo tanto, es una variable aleatoria continua.

Para esta variable aleatoria, no va a tener sentido preguntarnos cual es la probabilidad de que $X = 28,77$, es decir, cual es la probabilidad de que una persona emplee exactamente 28,77 minutos desde que entró en la estación A y llegó al centro comercial, y si lo tuviese, la probabilidad sería nula, pues exactamente 28,77 no emplearía ninguna persona. En este problema tiene mas sentido preguntarnos por la probabilidad de que una persona espere menos de 24 minutos, o emplee entre 27 y 31 minutos,..... Necesitaremos disponer de una expresión que nos permita calcular este tipo de probabilidades.

□

Ejemplo 15. Experimento aleatorio: número de horas transcurridas entre la percepción de temblores de tierra y la próxima erupción del volcán Kilauea. En este caso, podemos tomar como espacio muestral $\Omega = (0, \infty)$. Definimos la variable aleatoria $T =$ "número de horas transcurridas entre la percepción de temblores de tierra y la próxima erupción del volcán Kilauea".

$$\begin{aligned} T : \Omega = (0, \infty) &\mapsto \mathbb{R} \\ \omega &\longrightarrow X(\omega) = \omega. \end{aligned}$$

T es una variable aleatoria continua. Nos gustaría, por ejemplo, calcular $P(T < 24)$, probabilidad de que la erupción del volcán Kilauea se produzca en las 24 horas que siguen desde que empezamos a notar los temblores de tierra. En el caso continuo, el cálculo de probabilidades puede hacerse geoméricamente igualando probabilidades a áreas.

□

Entre las variables aleatorias continuas, las más fáciles de estudiar, y las que nosotros vamos a tratar son las **absolutamente continuas**. Nos referiremos a estas simplemente como continuas.

Definición 11. Sea (Ω, \mathcal{A}, P) un espacio probabilístico y $X : \Omega \mapsto \mathbb{R}$ una variable aleatoria continua. Asociada a la variable aleatoria X , tenemos una función f , llamada función de densidad de la variable aleatoria, definida en \mathbb{R} verificando:

1. $f(x) \geq 0$, (f es no negativa).
2. El área comprendida entre la gráfica de f y el eje x es igual a 1, $\left(\int_{-\infty}^{\infty} f(x)dx = 1\right)$.
3. Dados dos números reales $a \leq b$,

$$P[a, b] = P(\{\omega \in \Omega : a \leq X(\omega) \leq b\}) = \int_a^b f(x)dx.$$

Ejemplo 16. Consideremos un cierto experimento aleatorio con espacio muestral asociado Ω . Sea $X : \Omega \mapsto \mathbb{R}$ una variable aleatoria con función de densidad asociada

$$f(x) = \begin{cases} k(1+x^2) & \text{si } x \in (0, 3) \\ 0 & \text{si } x \notin (0, 3) \end{cases}$$

1. Calcular la constante k .
2. Probabilidad de que X esté comprendida entre 1 y 2.
3. Probabilidad de que X sea menor que 1.
4. Sabiendo que X es mayor que 1, probabilidad de que sea menor que 2.

1. Como f tiene que ser no negativa, se tendrá que verificar que $k > 0$. Para calcular la constante, vamos a imponer $\int_{-\infty}^{\infty} f(x)dx = 1$.

$$\int_{-\infty}^{\infty} f(x)dx = k \int_0^3 (1+x^2)dx = k \left(x + \frac{x^3}{3} \Big|_0^3 \right) = 12k.$$

Entonces

$$\int_{-\infty}^{\infty} f(x)dx = 1 \iff \boxed{k = \frac{1}{12}}.$$

y la función de densidad es

$$f(x) = \begin{cases} \frac{1+x^2}{12} & \text{si } x \in (0,3) \\ 0 & \text{si } x \notin (0,3) \end{cases}$$

$$2. P[1,2] = P(\{\omega \in \Omega : 1 \leq X(\omega) \leq 2\}) = \int_1^2 f(x)dx$$

$$= \frac{1}{12} \int_1^2 (1+x^2)dx = \frac{1}{12} \left(x + \frac{x^3}{3} \Big|_1^2 \right) = \frac{1}{12} \left(2 + \frac{2^3}{3} - 1 - \frac{1^3}{3} \right) = \frac{5}{18}.$$

$$3. P(-\infty, 1) = P(\{\omega \in \Omega : X(\omega) < 1\}) = \int_{-\infty}^1 f(x)dx$$

$$= \frac{1}{12} \int_0^1 (1+x^2)dx = \frac{1}{12} \left(x + \frac{x^3}{3} \Big|_0^1 \right) = \frac{1}{12} \left(1 + \frac{1^3}{3} \right) = \frac{1}{9}.$$

$$4. P((-\infty, 2)|(1, \infty))$$

$$= \frac{P((-\infty, 2) \cap (1, \infty))}{P(1, \infty)} = \frac{P(1, 2)}{1 - P(-\infty, 1)} = \frac{\frac{5}{18}}{1 - \frac{1}{9}} = \frac{5}{16}.$$

□

La definición de función de densidad asociada a una variable aleatoria parece maravillosa, pero en realidad no lo es tanto. La primera tarea de un científico experimental es la de determinar la densidad adecuada de la variable aleatoria que esté considerando. Si esta variable aleatoria ha sido ampliamente estudiada en el pasado, se supone que su densidad puede haber sido ya desarrollada por otros y podemos utilizarla para contestar a las preguntas que se nos plantean. En cambio, si la variable es estudiada por primera vez, su densidad debe ser hallada a partir de datos experimentales.

3.1. Función de distribución

Definición 12. Sea (Ω, \mathcal{A}, P) un espacio probabilístico y $X : \Omega \mapsto \mathbb{R}$ una variable aleatoria continua con densidad f . Definimos la función de **distribución** $F_X \equiv F$ asociada a la variable aleatoria continua X por

$$F(x) = P(-\infty, x] = \int_{-\infty}^x f(t) dt .$$

Observación 8. La función de distribución asociada a una variable aleatoria continua no tiene discontinuidades y, por lo tanto, el conjunto de valores que toma con probabilidad no nula es no numerable.

Observación 9. La clasificación de variables aleatorias en discretas y continuas no implica que toda distribución de probabilidad haya de ser discreta o bien continua. Las distribuciones discretas y las distribuciones continuas son dos pequeñas clases disjuntas de distribuciones; son las más fáciles de estudiar, pero es importante saber que hay muchas funciones de distribución que no son discretas ni continuas.

Proposición 7. Propiedades Si f es la función de densidad de una variable aleatoria continua X y F su función de distribución, se cumple:

1. F es continua.
2. $P(X = a) = 0$, para todo $a \in \mathbb{R}$.
3. Si f es continua en $a \in \mathbb{R}$ se verifica

$$F'(a) = f(a).$$

4.

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

Ejercicio 24. Sea $([0, 1], \mathcal{A}, P)$ un espacio probabilístico con $P([a, b]) = b - a$ para $[a, b] \subset [0, 1]$, y

$$X : [0, 1] \mapsto \mathbb{R}$$

$$\omega \mapsto X(\omega) = \begin{cases} \omega & \omega \leq \frac{1}{2} \\ \omega - \frac{1}{2} & \omega > \frac{1}{2} \end{cases}.$$

Dibújese la gráfica de X y hállese la función de distribución de X y su función de densidad.

$$F_X(x) \equiv F(x) = P(\{\omega \in [0, 1] : X(\omega) \leq x\}) = \begin{cases} 0, & x < 0 \\ 2x, & 0 \leq x \leq \frac{1}{2} \\ 1, & x > \frac{1}{2}. \end{cases}$$

Para calcular la densidad $f(x)$, recordamos el Teorema fundamental del cálculo:

$$F(x) = \int_a^x f(t)dt \implies F'(x) = f(x).$$

(a puede ser $-\infty$).

Como

$$F(x) = \int_{-\infty}^x f(t)dt \implies f(x) = \begin{cases} 0, & x < 0 \\ 2, & 0 \leq x \leq \frac{1}{2} \\ 0, & x > \frac{1}{2}. \end{cases}$$

□

1. Distribución uniforme

Una variable aleatoria continua tiene una distribución uniforme en un intervalo $[a, b]$ si su función de densidad viene dada por

$$f(x) = \begin{cases} \frac{1}{b-a} & x \in [a, b] \\ 0 & x \notin [a, b] \end{cases}$$

La función de distribución vendrá dada por

$$F(x) = \int_{-\infty}^x f(t)dt.$$

Si $x < a$, entonces $F(x) = 0$ ya que $f(x) = 0$. Si $a \leq x < b$, entonces

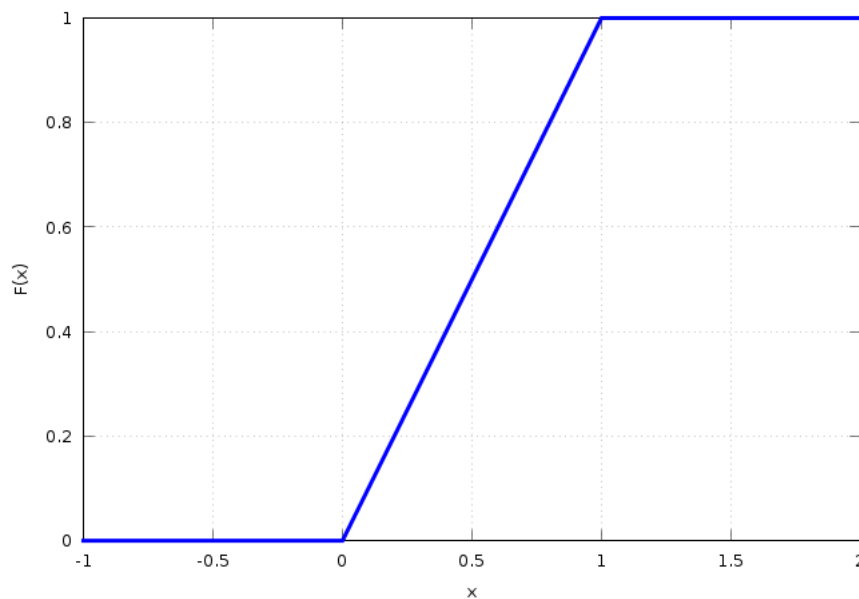
$$F(x) = \int_{-\infty}^x f(t)dt = \int_{-\infty}^a f(t)dt + \int_a^x f(t)dt = \int_a^x \frac{1}{b-a}dt = \frac{x-a}{b-a}.$$

Finalmente, si $x \geq b$, entonces

$$\begin{aligned} F(x) &= \int_{-\infty}^x f(t)dt = \int_{-\infty}^a f(t)dt + \int_a^b f(t)dt + \int_b^x f(t)dt \\ &= \int_a^b \frac{1}{b-a}dt = \frac{b-a}{b-a} = 1. \end{aligned}$$

por tanto

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



2. Distribución exponencial

Sea X una variable aleatoria continua, decimos que tiene una distribución exponencial de parámetro $\lambda > 0$ si su función de densidad viene dada por

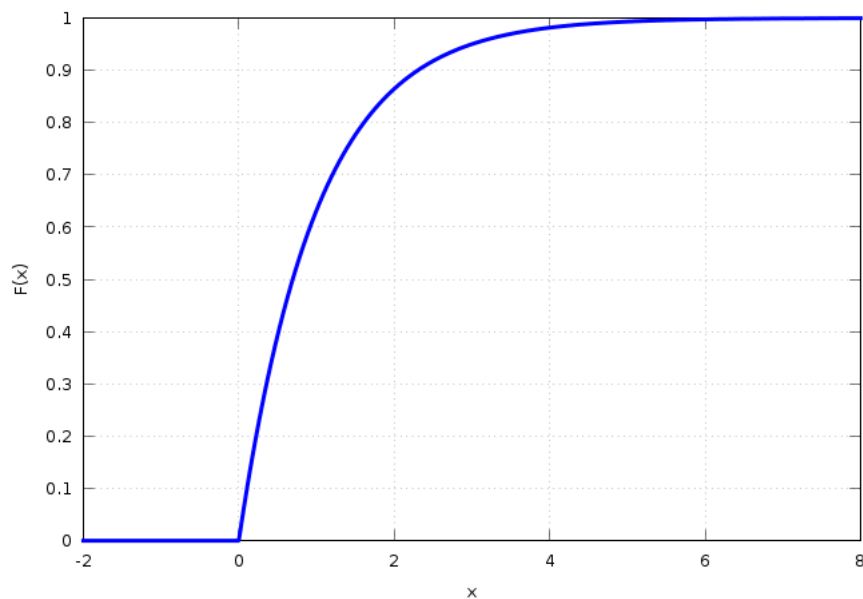
$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Si $x < 0$, la función de distribución $F(x) = 0$ ya que $f(x) = 0$. Si $x \geq 0$ entonces

$$F(x) = \int_{-\infty}^x f(t) dt = \int_{-\infty}^0 f(t) dt + \int_0^x f(t) dt = \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}.$$

por tanto

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

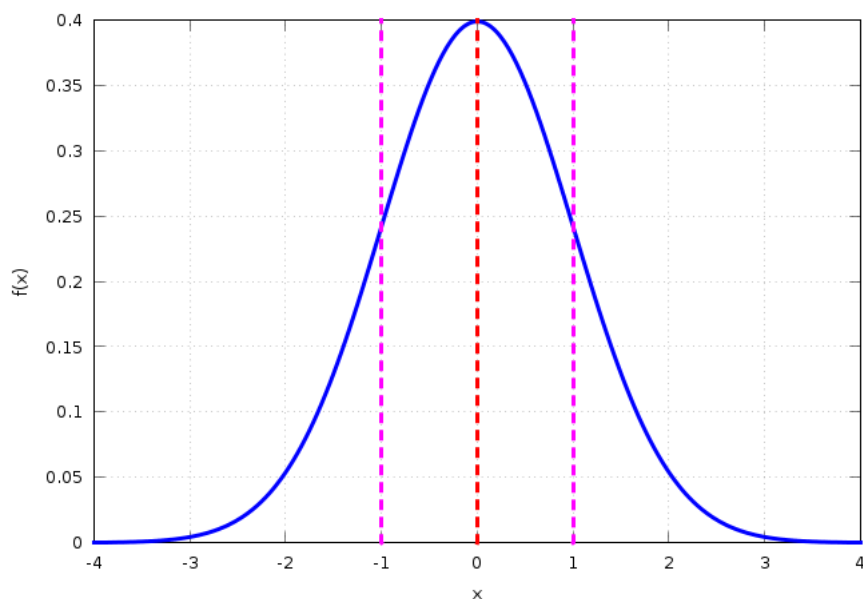


3. Distribución normal o de Gauss

Sea X una variable aleatoria continua, decimos que tiene una distribución normal o de Gauss si la función de densidad es

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

donde μ y $\sigma > 0$ son parámetros de la distribución. La gráfica de f es la conocida campana de Gauss con un máximo situado en $x = \mu$ y dos puntos de inflexión en $x = \mu \pm \sigma$.



Para calcular su función de distribución procedemos de la siguiente manera:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt,$$

hacemos el cambio de variable $u = \frac{t-\mu}{\sigma}$

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 e^{-\frac{u^2}{2}} du + \frac{1}{\sqrt{2\pi}} \int_0^{\frac{x-\mu}{\sigma}} e^{-\frac{u^2}{2}} du,$$

al ser $e^{-u^2/2}$ una función par

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 e^{-\frac{u^2}{2}} du = \frac{1}{\sqrt{2\pi}} \int_0^{\infty} e^{-\frac{u^2}{2}} du,$$

usando que

$$\int_0^{\infty} u^{2p-1} e^{-au^2} du = \frac{\Gamma(p)}{2a^p}, \quad \text{donde } \Gamma(p) = \int_0^{\infty} x^{p-1} e^{-x} dx,$$

tenemos

$$\int_0^{\infty} e^{-\frac{u^2}{2}} du = \frac{\Gamma(1/2)}{2\sqrt{1/2}} = \frac{\sqrt{2\pi}}{2},$$

y obtenemos

$$F(x) = \frac{1}{\sqrt{2\pi}} \cdot \frac{\sqrt{2\pi}}{2} + \frac{1}{\sqrt{2\pi}} \int_0^{\frac{x-\mu}{\sigma}} e^{-\frac{u^2}{2}} du = \frac{1}{2} + \Phi\left(\frac{x-\mu}{\sigma}\right),$$

donde

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{u^2}{2}} du$$

se llama la integral de probabilidad.

4. Transformación de variables aleatorias

Sea (Ω, \mathcal{A}, P) un espacio probabilístico, X una variable aleatoria y $g : \mathbb{R} \rightarrow \mathbb{R}$ una función. Tiene sentido hablar de la composición de X y g

$$\begin{array}{ccc} X & & g \\ \Omega & \longmapsto & \mathbb{R} \longmapsto \mathbb{R} \end{array}$$

$$\omega \longrightarrow X(\omega) \longrightarrow g(X(\omega))$$

Entonces $Y = g \circ X \equiv g(X)$ es una nueva función real definida sobre Ω tal que

$$Y(\omega) = g(X(\omega)).$$

Si queremos que Y sea una variable aleatoria debemos imponer la condición de que

$$\{\omega \in \Omega : Y(\omega) \leq y\} \in \mathcal{A}, \quad \text{para todo } y \in \mathbb{R}.$$

Admitiremos en lo que sigue esta suposición, de forma que Y va a ser una variable aleatoria.

Nos planteamos el encontrar la función de densidad (función de probabilidad) y de distribución de Y en términos de la función de densidad (función de probabilidad) y de distribución de X . En el caso de variables aleatorias discretas tenemos **Proposición 8.** Sea X una variable aleatoria discreta e Y definida por $Y = g(X)$. Y es una variable aleatoria discreta y su función de distribución de distribución viene dada por

$$F_Y(y) = \sum_{g(x) \leq y} P(X = x).$$

Demostración

$$P(Y = y) = \sum_{g(x)=y} P(X = x),$$

y por lo tanto

$$F_Y(y) = P(Y \leq y) = \sum_{g(x) \leq y} P(X = x).$$

■

Ejemplo 17. Consideramos la variable aleatoria X discreta definida por

$x_i =$	0	2	5	7
$P(X = x_i) =$	0.3	0.2	0.4	0.1

Determine la función de distribución de la variable aleatoria $Y = 3X + 2$.

Solución

Primero observamos que

$x_i =$	0	2	5	7
$y_i =$	2	8	17	23

y entonces

$y_i =$	2	8	17	23
$P(X = x_i) =$	0.3	0.2	0.4	0.1

La función de distribución de Y viene dada por

$$F(y) = \sum_{3x_i+2 \leq y} P(X = x_i) = \begin{cases} 0 & y < 2 \\ 0,3 & 2 \leq y < 8 \\ 0,5 & 8 \leq y < 17 \\ 0,9 & 17 \leq y < 23 \\ 1 & 23 \leq y \end{cases}$$

□

En el caso de que las variables sean continuas, hay dos procedimientos. Uno consiste en primero calcular la función de distribución F_Y de $Y = g(X)$ a partir de la de X y, después, si F_Y es derivable, determinar la función de densidad por $f_Y(y) = F'_Y(y)$.

Proposición 9. Sea X una variable aleatoria continua con función de densidad f_X y sea $Y = g(X)$. Entonces la función de distribución de Y es

$$F_Y(y) = \int_D f_X(x) dx, \quad \text{donde } D = \{x \in \mathbb{R} : g(x) \leq y\}.$$

Demostración Tenemos que

$$F_Y(y) = P(Y \leq y) = P(g(X) \leq y) = P(D),$$

ya que de la condición $g(X(\omega))$ se deduce que $X(\omega) \in D$. Por tanto

$$P(X \in D) = \int_D f_X(x) dx \implies F_Y(y) = \int_D f_X(x) dx.$$

■

Ejemplo 18. Dada la variable aleatoria continua X con función de densidad

$$f(x) = \begin{cases} 2x & x \in (0, 1) \\ 0 & x \notin (0, 1) \end{cases}$$

consideramos la siguientes transformaciones:

1. $Y = 3X + 5$,
2. $Y = X^2$,
3. $Y = 2X^2 + 5$.

En cada uno de los casos, calcule la función de distribución y densidad.

Solución La función de distribución de X es

$$F(x) = \begin{cases} 0 & x < 0 \\ x^2 & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

1. Consideremos la transformación $Y = 3X + 5$. Tenemos

$$F_Y(y) = P(Y \leq y) = P(3X + 5 \leq y) = P\left(X \leq \frac{y-5}{3}\right) = F\left(\frac{y-5}{3}\right).$$

Como

$$\begin{aligned}\frac{y-5}{3} < 0 &\iff y < 5 \\ 0 \leq \frac{y-5}{3} < 1 &\iff 5 \leq y < 8 \\ \frac{y-5}{3} \geq 1 &\iff y \geq 8\end{aligned}$$

tenemos

$$F_Y(y) = F\left(\frac{y-5}{3}\right) = \begin{cases} 0 & y < 5 \\ \frac{(y-5)^2}{9} & 5 \leq y < 8 \\ 1 & y \geq 8 \end{cases}$$

Ya que F_Y es derivable, tenemos que la función de densidad es

$$f_Y(y) = \begin{cases} \frac{2(y-5)}{9} & y \in (5, 8) \\ 0 & y \notin (5, 8) \end{cases}$$

2. Consideramos la transformación $Y = X^2$. Si $y < 0$ tenemos

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y) = P(\emptyset) = 0,$$

supongamos entonces que $y \geq 0$,

$$F_Y(y) = P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = F(\sqrt{y}) - F(-\sqrt{y}) = F(\sqrt{y}),$$

ya que $F(-\sqrt{y}) = 0$. Como

$$\begin{aligned}0 \leq \sqrt{y} < 1 &\iff 0 \leq y < 1 \\ \sqrt{y} \geq 1 &\iff y \geq 1\end{aligned}$$

tenemos

$$\begin{cases} 0 & y < 0 \\ y & 0 \leq y < 1 \\ 1 & y \geq 1 \end{cases}$$

La función de densidad es

$$f_Y(y) = \begin{cases} 1 & y \in (0, 1) \\ 0 & y \notin (0, 1) \end{cases}$$

3. Se deja al alumno como ejercicio.

$$F_Y(y) = \begin{cases} 0 & y < 5 \\ \frac{y-5}{2} & 5 \leq y < 7 \\ 1 & y \geq 7 \end{cases} \quad f_Y(y) = \begin{cases} \frac{1}{2} & y \in (5, 7) \\ 0 & y \notin (5, 7) \end{cases}$$

□

4.1. Medidas de centralización y dispersión

Definición 13. Sea (Ω, \mathcal{A}, P) un espacio probabilístico y $X : \Omega \mapsto \mathbb{R}$ una variable aleatoria continua con densidad f . La **media o esperanza** de la variable aleatoria continua X se define

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

Observación 10. Normalmente se suele exigir que la integral que de la esperanza de X sea absolutamente convergente, es decir, que la integral

$$\int_{-\infty}^{\infty} |x|f(x)dx < \infty.$$

Como en el caso de las variables aleatorias discretas, existen variables aleatorias continuas para las que no existe su esperanza.

Ejemplo 19. Sea X una variable aleatoria continua que se distribuye uniformemente en el intervalo (a, b) . Calcule la esperanza de X .

Solución

Al tratarse de una distribución uniforme en (a, b) se tiene que

$$f(x) = \begin{cases} \frac{1}{b-a} & x \in (a, b) \\ 0 & x \notin (a, b) \end{cases}$$

Tenemos

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx = \frac{1}{b-a} \int_a^b xdx = \frac{1}{b-a} \frac{x^2}{2} \Big|_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}.$$

□

Proposición 10. Propiedades de la esperanza matemática Sea (Ω, \mathcal{A}, P) un espacio probabilístico. X y Y dos variables aleatorias continuas y α y β dos números reales.

1. $E[\alpha X + \beta Y] = \alpha E[X] + \beta E[Y]$.
2. $E[X_\alpha] = E[\alpha] = \alpha$.
3. Si $g : \mathbb{R} \mapsto \mathbb{R}$ es derivable y $f(x)$ es la función de densidad de la variable aleatoria X , se tiene

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

$$4. |E[g(X)]| \leq E[|g(X)|].$$

Ejercicio 25. Una variable aleatoria X tiene como función de densidad

$$f(x) = \begin{cases} \frac{2x}{3} & x \in (1, 2) \\ 0 & x \notin (1, 2) \end{cases}$$

Calcule el valor esperado de la variable aleatoria $g(X) = 3X + X^2$.

Solución

$$\begin{aligned} E[g(X)] &= \int_{-\infty}^{\infty} g(x)f(x)dx = \int_1^2 (3x + x^2) \frac{2x}{3} dx \\ &= \int_1^2 \left(2x^2 + \frac{2x^3}{3} \right) dx = \frac{2x^3}{3} + \frac{x^4}{6} \Big|_1^2 = \frac{43}{6}. \end{aligned}$$

□

Definición 14. Sea (Ω, \mathcal{A}, P) un espacio probabilístico y X una variables aleatoria continua. La **mediana** de la variable aleatoria X se define como el valor M que satisface

$$F(M) = \frac{1}{2},$$

donde F es la función de distribución de X .

Definición 15. Sea (Ω, \mathcal{A}, P) un espacio probabilístico y X una variables aleatoria continua. La **varianza** de la variable aleatoria X se define

$$\sigma^2 = V(X) = E[(X - E[X])^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx = \int_{-\infty}^{\infty} x^2 f(x) dx - \mu^2.$$

La **desviación típica** σ de la variable aleatoria X se define como la raíz cuadrada positiva de la varianza.

Ejercicio 26. Sea X una variable aleatoria que tiene como función de densidad

$$f(x) = \begin{cases} 0, & x \leq 0 \\ a(1 + x^2), & 0 < x < 3 \\ 0, & x \geq 3 \end{cases}.$$

1. Hállese a y la función de distribución de X .
2. Hallar la probabilidad de que X esté comprendida entre 1 y 2.
3. Hallar $P(X < 1)$.
4. Hallar $P(X < 2 | X > 1)$.
5. Hallar la media μ y varianza σ^2 de X .
6. Calcule $P(|X - \mu| \geq 2\sigma)$.

Solución

1. Como f tiene que ser mayor o igual que 0, se tendrá que verificar que $a > 0$. Para calcular a vamos a imponerla condición de $\int_{-\infty}^{\infty} f(x)dx = 1$.

$$\int_{-\infty}^{\infty} f(x)dx = a \int_0^3 (1 + x^2)dx = a \left(x + \frac{x^3}{3} \Big|_0^3 \right) = 12a,$$

entonces

$$\int_{-\infty}^{\infty} f(x)dx = 1 \iff a = \frac{1}{12},$$

y la función de densidad es

$$f(x) = \begin{cases} 0, & x \leq 0 \\ \frac{1+x^2}{12}, & 0 < x < 3 \\ 0, & x \geq 3 \end{cases}.$$

Vamos a calcular ahora la función de distribución

$$F(x) = \int_{-\infty}^x f(t)dt.$$

Si $x \leq 0$, entonces $f(t) = 0$ si $t \in (-\infty, x]$, y

$$\begin{aligned} F(x) &= \int_{-\infty}^x f(t)dt = \lim_{z \rightarrow -\infty} \int_z^x 0dt = \lim_{z \rightarrow -\infty} \int_z^x 0dt \\ &= \lim_{z \rightarrow -\infty} (1|_z^x) = \lim_{z \rightarrow -\infty} (1 - 1) = 0, \quad x \leq 0. \end{aligned}$$

Supongamos ahora que $0 < x < 3$,

$$F(x) = \int_{-\infty}^x f(t)dt = \int_{-\infty}^0 f(t)dt + \int_0^x f(t)dt = \frac{1}{12} \int_0^x (1 + t^2)dt$$

$$= \frac{1}{12} \left(t + \frac{t^3}{3} \Big|_0^x \right) = \frac{1}{12} \left(x + \frac{x^3}{3} \right) = \frac{3x + x^3}{36}, \quad 0 < x < 3.$$

Si $x \geq 3$,

$$\begin{aligned} F(x) &= \int_{-\infty}^x f(t)dt = \int_{-\infty}^0 f(t)dt + \int_0^3 f(t)dt + \int_3^x f(t)dt = \frac{1}{12} \int_0^3 (1+t^2)dt \\ &= \frac{1}{12} \left(t + \frac{t^3}{3} \Big|_0^3 \right) = 1, \quad x \geq 3. \end{aligned}$$

Resumiendo:

$$F(x) = P((-\infty, x]) = \int_{-\infty}^x f(t)dt = \begin{cases} 0, & x \leq 0 \\ \frac{3x+x^3}{36}, & 0 < x < 3 \\ 0, & x \geq 3 \end{cases}$$

2.

$$\begin{aligned} P(1 < X < 2) &= \int_1^2 f(t)dt = \frac{1}{12} \int_1^2 (1+t^2)dt = \frac{1}{12} \left(t + \frac{t^3}{3} \Big|_1^2 \right) \\ &= \frac{1}{12} \left(\left(2 + \frac{8}{3} \right) - \left(1 + \frac{1}{3} \right) \right) = \frac{5}{18}. \end{aligned}$$

Ya que conocemos la función de distribución, no teníamos que haber trabajado tanto:

$$P((1, 2)) = P((1 < X < 2)) = P((-\infty, 2) - (-\infty, 1]) = P((-\infty, 2)) - P((-\infty, 1]).$$

$$\text{Como } P(\{2\}) = \int_2^2 f(t)dt = 0,$$

$$P((-\infty, 2]) = P((-\infty, 2) \cup \{2\}) = P((-\infty, 2)) + P(\{2\}) = P((-\infty, 2)),$$

y

$$P((1, 2)) = F(2) - F(1) = \frac{5}{18}.$$

3.

$$P(X < 1) = F(1) = \frac{3 \cdot 1 + 1^3}{36} = \frac{1}{9}.$$

4.

$$\begin{aligned} P(X < 2 | X > 1) &= \frac{P(\{X < 2\} \cap \{X > 1\})}{P(X > 1)} = \frac{P(1 < X < 2)}{P(\{X \leq 1\}^c)} \\ &= \frac{P(1 < X < 2)}{1 - P(X \leq 1)} = \frac{f(2) - F(1)}{1 - F(1)} = \frac{45}{144}. \end{aligned}$$

5.

$$\begin{aligned}
 E[X] = \mu &= \int_{-\infty}^{\infty} t f(t) dt = \frac{1}{12} \int_0^3 (t + t^3) dt \\
 &= \frac{1}{12} \left(\frac{t^2}{2} + \frac{t^4}{4} \Big|_0^3 \right) = \frac{1}{12} \left(\frac{9}{2} + \frac{81}{4} \right) = \frac{33}{16} = 2,0625 \\
 \sigma^2 = V(X) &= \int_{-\infty}^{\infty} t^2 f(t) dx - \mu^2 = \frac{1}{12} \int_0^3 (t^2 + t^4) dt - \mu^2 \\
 &= \frac{1}{12} \left(\frac{t^3}{3} + \frac{t^5}{5} \Big|_0^3 \right) - \mu^2 = \frac{1}{12} \left(9 + \frac{243}{5} \right) - \frac{33^2}{16^2} = 0,546
 \end{aligned}$$

6.

$$\begin{aligned}
 P(|X - \mu| \geq 2\sigma) &= P(|X - 2,0625| \geq 1,48) \\
 &= P(\{X - 2,0625 \geq 1,48\} \cup \{X - 2,0625 \leq -1,48\}) \\
 &= P(\{X \geq 3,5425\} \cup \{X \leq 0,5825\}) = P(\{X \geq 3,5425\}) + P(\{X \leq 0,5825\}) \\
 &= 1 - P(X < 3,5425) + P(X \leq 0,5825) \\
 &= 1 - F(3,5425) + F(0,5825) = 1 - 1 + F(0,5825) = F(0,5825).
 \end{aligned}$$

□

Ejercicio 27. La cantidad aleatoria de dinero ahorrado por una persona en un mes sigue una ley de probabilidad (función de distribución) dada por :

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

donde x viene expresado en cientos de euros. Determine la probabilidad de que en un mes, la cantidad de dinero ahorrado sea:

1. superior a 200 euros;
2. inferior a 450 euros;
3. superior a 50 euros y menor o igual que 250 euros.
4. Calcular el ahorro medio mensual.

Solución

1. $P(X > 2) = 1 - F(2) = 0,5$
2. $P(X < 4,5) = f(4,5) = 1.$
3. $P(0,5 < X \leq 2,5) = f(2,5) - f(0,5) = \frac{3}{8}.$
4. El ahorro medio mensual es la esperanza de la variable aleatoria continua X . No conocemos esta, solamente su función de distribución. Vamos a calcular la densidad $f(x)$ asociada a X .

Recordamos el teorema fundamental del cálculo:

$$F(x) = \int_a^x f(t)dt \implies F'(x) = f(x),$$

(a puede ser $-\infty$). Entonces

$$f(x) = F'(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{2}, & 0 \leq x < 1 \\ 0, & 1 \leq x < 2 \\ \frac{1}{4}, & 2 \leq x < 4 \\ 0, & x \geq 4 \end{cases}$$

Entonces

$$E[X] = \mu = \int_{-\infty}^{\infty} tf(t)dt = \int_0^1 \frac{t}{2}dt + \int_2^4 \frac{t}{4}dt = 1,75$$

□

Ejercicio 28. Sea $([-1, 1], \mathcal{A}, P)$ un espacio probabilístico y X una variable aleatoria con función de densidad

$$f(x) = \begin{cases} 0, & x < -1 \\ kx + \frac{1}{2}, & -1 \leq x \leq 1 \\ 0, & x > 1. \end{cases}$$

1. Determínese el valor (los valores) de k .
2. Calcule la esperanza, moda y mediana de X .
3. ¿Para qué valores de k se minimiza la varianza?

1. Se tiene que verificar que $f(x) \geq 0 \ x \in \mathbb{R} \iff x \in [-1, 1]$.

$$f(x) = kx + \frac{1}{2} \geq 0 \ x \in [-1, 1] \iff kx \geq -\frac{1}{2} \ x \in [-1, 1]$$

$$\iff \begin{cases} x \leq -\frac{1}{2x} & x \in [-1, 0) \\ x \geq -\frac{1}{2x} & x \in (0, 1] \end{cases} \iff \begin{cases} k \leq \frac{1}{2} \\ k \geq -\frac{1}{2} \end{cases} \iff k \in [-1, 1].$$

También tenemos que imponer que $\int_{-\infty}^{\infty} f(x)dx = 1$.

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)dx &= \int_{-\infty}^{\infty} \left(kx + \frac{1}{2}\right) dx = \left(k\frac{x^2}{2} + \frac{x}{2}\right) \Big|_{-1}^1 \\ &= \left(\left(k\frac{1}{2} + \frac{1}{2}\right) - \left(k\frac{1}{2} - \frac{1}{2}\right)\right) = 1, \end{aligned}$$

luego cualquiera que sea el número k , se verifica que $\int_{-\infty}^{\infty} f(x)dx = 1$. Luego $k \in [-1, 1]$.

2. Empecemos calculando la esperanza de X .

$$\begin{aligned} \mu = E[X] &= \int_{-\infty}^{\infty} xf(x)dx = \int_{-\infty}^{\infty} \left(kx^2 + \frac{x}{2}\right) dx \\ &= \left(k\frac{x^3}{3} + \frac{x^2}{4}\right) \Big|_{-1}^1 = \left(\left(k\frac{1}{3} + \frac{1}{4}\right) - \left(-k\frac{1}{3} + \frac{1}{4}\right)\right) = \frac{2k}{3}. \end{aligned}$$

La moda es el valor que maximiza la función de densidad.

- a) Si $k > 0$, $kx + \frac{1}{2}$ es creciente en $[-1, 1]$ y f es máxima si $x = 1$. En este caso, la moda es 1.
- b) Si $k = 0$, $\frac{1}{2}$ es constante en $[-1, 1]$ y f es en cualquier punto de $[-1, 1]$. En este caso, la moda es cualquier punto de $[-1, 1]$.
- c) Si $k < 0$, $kx + \frac{1}{2}$ es decreciente en $[-1, 1]$ y f es máxima si $x = -1$. En este caso, la moda es -1 .

Para obtener la mediana, vamos a calcular la función de distribución:

$$F(x) = \int_{-\infty}^x f(t)dt = \begin{cases} \int_{-\infty}^x 0dt = 0 & x < -1 \\ \int_{-\infty}^x \left(kt + \frac{1}{2}\right) dt = \frac{kx^2}{2} + \frac{x}{2} - \frac{k}{2} + \frac{1}{2}, & -1 \leq x \leq 1 \\ \int_{-\infty}^x f(t)dt = \int_{-1}^1 \left(kx + \frac{1}{2}\right) dt = 1 & x > 1. \end{cases}$$

$$\text{Si } k \neq 0, F(M) = \frac{1}{2} \iff M = \frac{-1}{2}$$

Ejemplo 20. De una estación parte un tren cada 20 minutos. Un viajero llega de improviso. Hallar:

1. Función de distribución de la variable aleatoria "tiempo de espera".

2. Probabilidad de que espere al tren menos de 7 minutos.
3. Esperanza y varianza de la variable aleatoria "tiempo de espera".
4. Probabilidad de que espere exactamente 12 minutos.

Experimento aleatorio: cuando llega un viajero de improviso a la estación, ver el tiempo de espera hasta la llegada del tren.

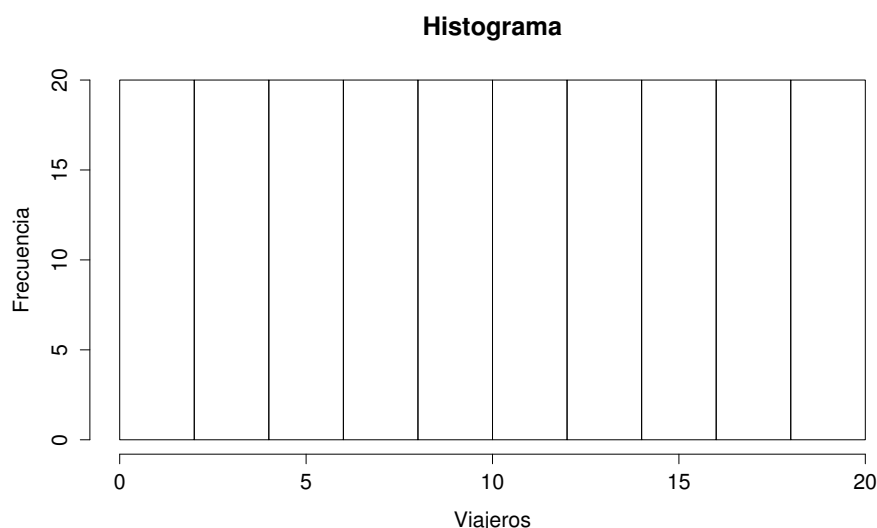
El espacio muestral, posibles resultados del experimento aleatorio, cualquier tiempo entre 0 y 20 minutos, esto es el intervalo $[0, 20]$.

La variable aleatoria "tiempo de espera" está definida por

$$\begin{aligned} X : \Omega = [0, 20] &\longrightarrow \mathbb{R} \\ \omega &\longrightarrow X(\omega) = \omega \text{ (tiempo que espera el viajero).} \end{aligned}$$

La variable aleatoria es claramente continua, puede tomar cualquier valor en el intervalo $[0, 20]$. Necesitamos una función de densidad y esta la tenemos que calcular nosotros.

Si tomamos el tiempo de espera de 1000 viajeros que llegan de improviso a la estación, la lógica nos dice que ningún viajero va a esperar menos de 0 minutos y más de 20 minutos, y que el número de viajeros que va a esperar entre 2 y 4 minutos, va a ser aproximadamente el mismo que el que va a esperar entre 12 y 14 minutos, y serán aproximadamente 100 viajeros. Tendríamos para esta muestra de 1000 viajeros el siguiente histograma



Esto nos sugiere que la función de densidad debe ser

$$f(x) = \begin{cases} k & x \in [0, 20] \\ 0 & x \notin [0, 20]. \end{cases}$$

para algún número real k positivo. Para calcular este, vamos a imponer que $\int_{-\infty}^{\infty} f(x)dx = 1$.

$$\int_{-\infty}^{\infty} f(x)dx = k \int_0^{20} dx = 20k,$$

entonces

$$\int_{-\infty}^{\infty} f(x)dx = 1 \iff k = \frac{1}{20},$$

y la función de densidad es

$$f(x) = \begin{cases} \frac{1}{20} & x \in [0, 20] \\ 0 & x \notin [0, 20]. \end{cases}$$

1. Vamos a calcular la función de distribución.

- Sea $x \leq 0$, $f(t) = 0$ si $t \in (-\infty, x)$ y $F(x) = \int_{-\infty}^x f(t)dt = \int_{-\infty}^x 0dt = 0$.
- Si $0 < x < 20$, $F(x) = \int_{-\infty}^x f(t)dt = \frac{1}{20} \int_0^x dt = \frac{x}{20}$.
- Si $20 \leq x$, $F(x) = \int_{-\infty}^x f(t)dt = \frac{1}{20} \int_0^{20} dt = 1$.

$$F(x) = \begin{cases} 0 & x \leq 0 \\ \frac{x}{20} & 0 < x < 20 \\ 1 & x \geq 20. \end{cases}$$

$$2. P(X < 7) = P(-\infty, 7) = \int_{-\infty}^7 f(x)dx = \frac{1}{20} \int_0^7 dx = \frac{7}{20}.$$

3.

$$\mu = E[X] = \int_{-\infty}^{\infty} xf(x)dx = \frac{1}{20} \int_0^{20} xdx = \frac{1}{20} \left(\frac{x^2}{2} \Big|_0^{20} \right) = \frac{1}{20} \cdot \frac{20^2}{2} = 10.$$

$$\begin{aligned} \sigma^2 = V(X) &= \int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2 = \frac{1}{20} \int_0^{20} x^2 dx - 100 = \frac{1}{20} \left(\frac{x^3}{3} \Big|_0^{20} \right) - 100 \\ &= \frac{1}{20} \cdot \frac{20^3}{3} - 100 = \frac{100}{3}. \end{aligned}$$

$$4. P(X = 12) = P[12, 12] = \int_{12}^{12} f(x)dx = 0.$$

Ejemplo 21. El tiempo de vida (en minutos) de un determinado virus es una variable aleatoria con función de densidad.

$$f(x) = \begin{cases} \frac{1}{1,000} e^{-\frac{x}{1,000}} & x > 0 \\ 0 & x \leq 0. \end{cases}$$

1. Probabilidad de que el tiempo de vida sea superior a 100 minutos.
2. Probabilidad de que el tiempo de vida sea superior a 100 minutos e inferior a 1.000 minutos.
3. Observamos el virus a los 500 minutos y comprobamos que ha muerto. ¿Cuál es la probabilidad de que estuviese vivo a los 100 minutos?

Experimento aleatorio: elegir al azar un virus y ver el tiempo de vida en minutos.

Un virus vive más de cero minutos, podemos pensar en el espacio muestral como el conjunto $\Omega = (0, \infty)$.

La variable aleatoria "tiempo de vida en minutos del virus." está definida por

$$\begin{aligned} X : \Omega = (0, \infty) &\longrightarrow \mathbb{R} \\ \omega &\longrightarrow X(\omega) = \omega \text{ (tiempo de vida en minutos del virus).} \end{aligned}$$

Es una variable aleatoria continua y su función de densidad es $f(x)$.

$$1. P(X > 100) = P(100, \infty) = \int_{100}^{\infty} f(x) dx = \lim_{k \rightarrow \infty} \int_{100}^k f(x) dx$$

$$\begin{aligned} &= \lim_{k \rightarrow \infty} \int_{100}^k \frac{1}{1,000} e^{-\frac{x}{1,000}} dx = \lim_{k \rightarrow \infty} \left(-e^{-\frac{x}{1,000}} \Big|_{100}^k \right) \\ &= \lim_{k \rightarrow \infty} \left(e^{-\frac{100}{1,000}} - e^{-\frac{k}{1,000}} \right) = e^{-\frac{1}{10}} \cong 0,904. \end{aligned}$$

$$2. P(100 < X < 1,000) = P(100, 1,000) = \int_{100}^{1,000} f(x) dx$$

$$= \int_{100}^{1,000} \frac{1}{1,000} e^{-\frac{x}{1,000}} dx = \left(-e^{-\frac{x}{1,000}} \Big|_{100}^{1,000} \right) = e^{-\frac{1}{10}} - e^{-1} \cong 0,537.$$

$$3. P(X > 100 | X < 500) = \frac{P((X > 100) \cap (X < 500))}{P(X < 500)} = \frac{P(100 < X < 500)}{P(X < 500)}$$

$$\begin{aligned} &= \frac{P(100, 500)}{P(-\infty, 500)} = \frac{\int_{100}^{500} \frac{1}{1,000} e^{-\frac{x}{1,000}} dx}{\int_0^{500} \frac{1}{1,000} e^{-\frac{x}{1,000}} dx} = \frac{\left(-e^{-\frac{x}{1,000}} \Big|_{100}^{500} \right)}{\left(-e^{-\frac{x}{1,000}} \Big|_0^{500} \right)} \\ &= \frac{e^{-\frac{1}{10}} - e^{-\frac{1}{2}}}{e^0 - e^{-\frac{1}{2}}} \cong 0,76. \end{aligned}$$

□

Ejemplo 22. Un fabricante debe elegir entre dos procesos de producción que dan lugar a que las longitudes (en cm.) de los elementos producidos se distribuyan según

■ Proceso 1: $f_1(x) = \begin{cases} \frac{3}{x^4} & x \geq 1 \\ 0 & x < 1 \end{cases}$

■ Proceso 2: $f_2(x) = \begin{cases} \frac{4}{x^5} & x \geq 1 \\ 0 & x < 1 \end{cases}$

1. Si los elementos aceptables deben de tener longitud entre 1 y 2 cm., ¿qué proceso produce un porcentaje mayor de elementos aceptables?
2. Si se elige al azar uno de los procesos, ¿cuál es la probabilidad de obtener una pieza aceptable?
3. ¿Cuál es la longitud media en cada proceso?

1. Experimento aleatorio: Tomar una pieza y medir su longitud.

Espacio muestral $\Omega = (0, 20)$.

Definimos la variable aleatoria

$$\begin{aligned} X : \Omega = (0, 20) &\longrightarrow \mathbb{R} \\ \omega &\longrightarrow X(\omega) = \omega \text{ (longitud de la pieza).} \end{aligned}$$

Es una variable aleatoria continua.

Proceso 1

$$\begin{aligned} P(1 \leq X \leq 2) &= \int_1^2 f_1(x) dx = 3 \int_1^2 \frac{dx}{x^4} \\ &= 3 \left(-\frac{1}{3x^3} \Big|_1^2 \right) = \left(-\frac{1}{x^3} \Big|_1^2 \right) = \left(-\frac{1}{8} + 1 \right) = \frac{7}{8}. \end{aligned}$$

Proceso 2

$$\begin{aligned} P(1 \leq X \leq 2) &= \int_1^2 f_2(x) dx = 4 \int_1^2 \frac{dx}{x^5} \\ &= 4 \left(-\frac{1}{4x^4} \Big|_1^2 \right) = \left(-\frac{1}{x^4} \Big|_1^2 \right) = \left(-\frac{1}{16} + 1 \right) = \frac{15}{16}. \end{aligned}$$

Produce un mayor porcentaje de elementos aceptables el Proceso 2.

2. Experimento aleatorio: Tomar una pieza producida por uno de los dos procesos y ver si es aceptable.

Sucesos elementales

P1A= "pieza fabricada por el Proceso 1 y que es aceptable",

P1N= "pieza fabricada por el Proceso 1 y que no es aceptable",

P2A= "pieza fabricada por el Proceso 2 y que es aceptable",

P2n= "pieza fabricada por el Proceso 2 y que no es aceptable".

$$\Omega = \{P1A, P1N, P2A, P2N\}.$$

Sucesos

P1= "pieza producida por el Proceso 1",

P2= "pieza producida por el Proceso 2",

A= "pieza aceptable".

El apartado anterior nos dice que

$$P(A|P1) = \frac{7}{8}, \quad P(A|P2) = \frac{15}{16},$$

$$P(P1) = \frac{1}{2}, \quad P(P2) = \frac{1}{2}.$$

$$P(A) = P(A \cap (P1 \cup P2)) = P((A \cap P1) \cup (A \cap P2)) = P(A \cap P1) + P(A \cap P2)$$

$$= P(A|P1) \cdot P(P1) + P(A|P2) \cdot P(P2) = \frac{7}{8} \cdot \frac{1}{2} + \frac{15}{16} \cdot \frac{1}{2} = \frac{29}{32}.$$

3. Proceso 1

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x f_1(x) dx = 3 \int_1^{\infty} \frac{dx}{x^3} = 3 \lim_{k \rightarrow \infty} \int_1^k \frac{dx}{x^3} = 3 \lim_{k \rightarrow \infty} \left(-\frac{1}{2x^2} \Big|_1^k \right) \\ &= 3 \lim_{k \rightarrow \infty} \left(-\frac{1}{2k^2} + \frac{1}{2} \right) = \frac{3}{2}. \end{aligned}$$

Proceso 2

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x f_2(x) dx = 4 \int_1^{\infty} \frac{dx}{x^4} = 4 \lim_{k \rightarrow \infty} \int_1^k \frac{dx}{x^4} = 4 \lim_{k \rightarrow \infty} \left(-\frac{1}{3x^3} \Big|_1^k \right) \\ &= 4 \lim_{k \rightarrow \infty} \left(-\frac{1}{3k^3} + \frac{1}{3} \right) = \frac{4}{3}. \end{aligned}$$

4.2. Ejercicios

Ejercicio 29. Identificar la variable estadística como discreta o continua:

1. V: El volumen de orina producido por hora.
2. B: La cantidad de sangre perdida por un paciente durante el transcurso de una operación.
3. H: El número de horas de luz por día necesarias para que una planta florezca.
4. C: El número de abejas obreras en una colonia de abejas productoras de miel.
5. R: La cantidad de lluvias recibidas por día en una región concreta.
6. S: El nivel en suero de bilirrubina en un niño, en miligramos por decilitro.
7. W: El peso ganado por una mujer durante el embarazo.
8. X: El número de pruebas necesarias que permita conseguir el primer injerto realizado con éxito, de un tallo de cornejo rosa sobre un tronco de cornejo blanco.
9. C: Conde $C = 1$, donde el árbol muestra es de tamaño adecuado para madera, y $C = 0$ en caso contrario.
10. P: La tensión arterial sistólica de un paciente con hipertensión.
11. E: La altitud a la que se sitúa el límite de arbolado en una montaña.

Ejercicio 30. Determínese el valor de k para que las siguientes funciones sean de densidad:

$$1. f(x) = \begin{cases} 0 & x \leq 0 \\ ke^{-kx} & x > 0 \end{cases} . \quad \text{Solución: } k = 1.$$

$$2. f(x) = \begin{cases} 0 & x \leq 0 \\ \frac{k}{\sqrt{1-x}} & 0 < x < \frac{\sqrt{2}}{2} \\ 0 & x \geq \frac{\sqrt{2}}{2} \end{cases} . \quad \text{Solución: } k = 1,0898.$$

$$3. f(x) = \frac{k}{1+x^2}, \quad x \in \mathbb{R}. \quad \text{Solución: } k = \frac{1}{\pi}.$$

Ejercicio 31. Sea (Ω, \mathcal{A}, P) un espacio probabilístico y $X : \Omega \mapsto \mathbb{R}$ una variable aleatoria continua con función de densidad

$$f(x) = \begin{cases} 0, & x < -1 \\ 0,2, & -1 \leq x \leq 0 \\ 0,2 + kx, & 0 < x \leq 1 \\ 0, & x > 1 \end{cases}$$

1. Determine el valor de k . Solución: $k = 1,2$
2. Determine la función de distribución asociada a la variable aleatoria X .
Solución: $F(x) = \int_{-\infty}^x f(x) = \begin{cases} 0, & x < -1 \\ 0,2x + 0,2, & -1 \leq x \leq 0 \\ 0,6x^2 + 0,2x + 0,2, & 0 < x \leq 1 \\ 1, & x > 1 \end{cases}$
3. Calcule $P(0 \leq X \leq 0,5)$. Solución: $0,25$
4. Calcule $P(X > 0,5 | X > 0,1)$. Solución: $0,71$

Ejercicio 32. Se considera la duración de las bombillas fabricadas por una cierta empresa como una variable aleatoria X , cuya función de densidad es

$$f(x) = \begin{cases} 0 & x < 1 \\ \frac{k}{x^3} & x > 1 \end{cases}$$

1. Determínese la constante k .
2. Calcúlese la función de distribución.
3. Duración media de las bombillas.
4. Probabilidad de que una bombilla dure más de 50 horas.

Solución: 1. $k = 2$; 2. $F(x) = \begin{cases} 0 & x < 1 \\ -\frac{1}{x^2} + 1 & x > 1 \end{cases}$; 3. $E[X] = 2$; 4. $\frac{1}{2500}$

Ejercicio 33. El tiempo de vida (en años) de cierta especie es una variable aleatoria T con función de densidad

$$f(t) = \begin{cases} k(1-t)^2 t^2 & t \in (0, 1) \\ 0 & \text{en el resto} \end{cases}$$

1. Hallar el valor de k .
2. Hallar la esperanza de vida.
3. Probabilidad de que un ejemplar de esta especie viva menos de 9 meses.

Solución 1. $k = 30$; 2. 6 meses; 3. 0.8965

Ejercicio 34. En una empresa dedicada a la fabricación de tornillos se considera la dimensión de un tornillo como variable aleatorio, cuya función de densidad es

$$f(x) = \begin{cases} 0 & x < 1 \\ \frac{k}{x^2} & 1 \leq x \leq 8 \\ 0 & x > 8 \end{cases}$$

Determinése:

1. El valor de k .
2. La función de distribución.
3. Probabilidad de que la dimensión de un tornillo esté entre 3 y 5 cm..
4. Dimensión media de los tornillo y desviación con respecto a esta.
5. El valor de a , tal que el 90 % de los tornillos tenga su dimensión menor o igual que a .

Solución: 1. $k = \frac{8}{7}$; 2. $F(x) = \begin{cases} 0 & x < 1 \\ \frac{8}{7} \left(1 - \frac{1}{x}\right) & 1 \leq x \leq 8 \\ 1 & x > 8 \end{cases}$; 3. $\frac{16}{105}$; 4. $\frac{8 \log 8}{7}$; 5. $a = 4,706$

Ejercicio 35. Una fábrica produce una pieza en dos calidades diferentes:

- El 60 % de la producción es de calidad A . La duración (en años) de una pieza de esta calidad viene dada por la función de densidad

$$f_A(x) = \begin{cases} e^{-x} & x > 0 \\ 0 & x < 0 \end{cases}$$

- El 40 % de la producción es de calidad A. La duración (en años) de una pieza de esta calidad viene dada por la función de densidad

$$f_B(x) = \begin{cases} 2e^{-2x} & x > 0 \\ 0 & x < 0 \end{cases}$$

1. Probabilidad de que una pieza de calidad A dure más de un año.
2. Mediana de la distribución del tipo A.
3. Si tomamos una pieza al azar de toda la producción, ¿cuál es la probabilidad de que dure más de un año?
4. Si tomamos una pieza al azar de toda la producción, y observamos que dura más de un año, ¿cuál es la probabilidad de que fuera de calidad A?

Solución: 1. 0.3679; 2. M=0.6931; 3. 0.2749; 4. 0.8

Ejercicio 36. Los tiempos de vida, X e Y (en días); de una bacteria en dos medios distintos e independientes A y B, respectivamente, tienen las funciones de densidad

$$f(x) = \begin{cases} \frac{10-x}{50} & 0 < x < 10 \\ 0 & \text{en el resto} \end{cases} \quad g(y) = \begin{cases} \frac{1}{k}e^{-y/k} & 0 < y \\ 0 & \text{en el resto} \end{cases}$$

donde k es una constante positiva. La duración media de las bacterias en el medio B es de 5 días.

1. Calcular el valor de k y la esperanza de vida en el medio A.
2. Una bacteria tiene la misma probabilidad de estar en el medio A que el el B. Sabiendo que vivió más de 5 días, ¿cuál es la probabilidad de que se encontrara en el medio A?
3. Hallar la función de densidad conjunta de X e Y y $P(X > 5, Y > 5)$.

Solución: 1: $k = 5$, $E[X] = 3,33$ días; 2. 0.4; 3. 0.09

5. Vectores aleatorios

En la vida real es muy frecuente enfrentarnos a problemas en los que nos interesa analizar varias características simultáneamente, por ejemplo la velocidad de transmisión

de un mensaje y la proporción de errores. De esta manera seremos capaces de estudiar no solo el comportamiento de cada característica por separado, sino las relaciones que pudieran existir entre ellas. El objetivo de esta sección es elaborar un modelo matemático que permita analizar experimentos aleatorios en el que cada resultado experimental tiene asociado $p > 1$ valores numéricos.

El concepto de vector aleatorio nace como una generalización natural de la noción de variable aleatoria, al considerar simultáneamente el comportamiento aleatorio de varias características asociadas a un experimento.

Vamos a tratar, desde el punto de vista teórico, solo el caso de vectores aleatorios bidimensionales.

Supongamos que tenemos dos variables aleatorias X, Y sobre el mismo espacio probabilístico (Ω, \mathcal{A}, P) . De este modo para cada suceso elemental $\omega \in \Omega$ tenemos dos números reales $X(\omega)$ y $Y(\omega)$. Tenemos dos posibles interpretaciones

1. Considerar los números $X(\omega)$ y $Y(\omega)$ de forma separada, o
2. podemos considerar este par de números como las componentes de un vector $(X(\omega), Y(\omega)) \in \mathbb{R}^2$.

Esta segunda interpretación es la que va a conducir al concepto de **variable aleatoria bidimensional**.

Al ser X e Y variables aleatorias se tiene que

$$\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{A}, \quad \text{y} \quad \{\omega \in \Omega : Y(\omega) \leq y\} \in \mathcal{A},$$

para todo $x, y \in \mathbb{R}$. Por tanto

$$\{\omega \in \Omega : X(\omega) \leq x\} \cap \{\omega \in \Omega : Y(\omega) \leq y\} \in \mathcal{A},$$

y como consecuencia, estos sucesos tienen asignadas probabilidades. Podemos definir la función F_X por

$$F_{\vec{X}}(x, y) = P(\{\omega \in \Omega : X(\omega) \leq x\} \cap \{\omega \in \Omega : Y(\omega) \leq y\}),$$

para $(x, y) \in \mathbb{R}^2$. F_X es una función real de dos variables que se llama **función de distribución conjunta** o función de distribución de la "variable aleatoria bidimensional $\vec{X} = (X, Y)$ ".

Definición 16. Sea (Ω, \mathcal{A}, P) un espacio probabilístico. Un vector aleatorio bidimensional es una aplicación

$$\begin{aligned} \vec{X} &= (X, Y) : \Omega \longrightarrow \mathbb{R}^2 \\ \omega &\longrightarrow (X, Y)(\omega), \end{aligned}$$

tal que para todo A subconjunto de \mathbb{R}^2 conocemos $P(\{\omega \in \Omega : (X, Y)(\omega) \in A\})$.

En general, una variable aleatoria n -dimensional o vector aleatorio será una n -tupla $\vec{X} = (X_1, X_2, \dots, X_n)$ formada por variables aleatorias X_i , $i = 1, 2, \dots, n$ sobre el mismo espacio probabilístico (Ω, \mathcal{A}, P) . El vector $\vec{X} = (X_1, X_2, \dots, X_n)$ definirá una aplicación de Ω en \mathbb{R}^n que hace corresponder a cada suceso elemental $\omega \in \Omega$ un vector (x_1, x_2, \dots, x_n) de \mathbb{R}^n , siendo $x_i = X_i(\omega)$, $i = 1, 2, \dots, n$.

Definición 17. Dada una variable aleatoria bidimensional $\vec{X} = (X, Y)$ sobre el espacio probabilístico (Ω, \mathcal{A}, P) , se llama **función de distribución conjunta** a la función real de dos variables

$$F_{\vec{X}}(x, y) = P(X \leq x, Y \leq y) = P(\{\omega \in \Omega : X(\omega) \leq x\} \cap \{\omega \in \Omega : Y(\omega) \leq y\}).$$

Proposición 11. La función de distribución $F_{\vec{X}}$ de una variable aleatoria bidimensional $\vec{X} = (X, Y)$ satisface las siguientes propiedades:

1. $0 \leq F_{\vec{X}}(x, y) \leq 1$, para todo $(x, y) \in \mathbb{R}^2$.
2. $F_{\vec{X}}$ es monótona creciente en cada una de las variables
 - a) $x_1 \leq x_2 \implies F_{\vec{X}}(x_1, y) \leq F_{\vec{X}}(x_2, y)$,
 - b) $y_1 \leq y_2 \implies F_{\vec{X}}(x, y_1) \leq F_{\vec{X}}(x, y_2)$.
3. $F_{\vec{X}}(\infty, \infty) = 1$ y $F_{\vec{X}}(-\infty, y) = F_{\vec{X}}(x, -\infty) = 0$.
4. $P(a < X \leq b, c < Y \leq d) = F_{\vec{X}}(b, d) - F_{\vec{X}}(a, d) - F_{\vec{X}}(b, c) + F_{\vec{X}}(a, c)$.
5. $F_{\vec{X}}$ es continua por la derecha en cada una de las variables.
6. **Distribuciones marginales.**

$$\lim_{y \rightarrow \infty} F_{\vec{X}}(x, y) = F_X(x), \quad \lim_{x \rightarrow \infty} F_{\vec{X}}(x, y) = F_Y(y),$$

donde F_X y F_Y son las funciones de distribución de las variables aleatorias X e Y respectivamente. A estas dos funciones se les llama distribuciones marginales del vector aleatorio $\vec{X} = (X, Y)$.

5.1. Vectores aleatorios discretos o variables aleatorias bidimensionales discretas

Definición 18. Consideramos el espacio probabilístico (Ω, \mathcal{A}, P) , el vector aleatorio (X, Y) se llama discreto si X e Y son variables aleatorias discretas. Supongamos que

$$\text{rango de } X = \{x_1, x_2, x_3, \dots\} \quad \text{y} \quad \text{rango de } Y = \{y_1, y_2, y_3, \dots\},$$

definimos **la función de probabilidad conjunta del vector aleatorio** (X, Y) por

$$\{P(X = x_i, Y = y_j) = P(\{\omega \in \Omega : X(\omega) = x_i\} \cap \{\omega \in \Omega : Y(\omega) = y_j\})\}_{i,j=1,2,3,\dots}$$

Para $A \subset \mathbb{R}^2$ definimos

$$P(A) = P(\{\omega \in \Omega : (X(\omega), Y(\omega)) \in A\}).$$

La función de distribución conjunta viene dada por

$$\begin{aligned} F : \mathbb{R}^2 &\longrightarrow [0, 1] \\ (x, y) &\longrightarrow F(x, y) = P(X \leq x, Y \leq y) \end{aligned}$$

Observación 11. Por las propiedades que satisface la función de probabilidad, se tiene que

$$\sum_{i,j} P(X = x_i, Y = y_j) = 1.$$

Las funciones de probabilidad marginales del vector aleatorio $(X; Y)$ están dadas por

$$P(Y = y_j) = \sum_i P(X = x_i, Y = y_j), \quad j = 1, 2, \dots$$

$$P(X = x_i) = \sum_j P(X = x_i, Y = y_j), \quad i = 1, 2, \dots$$

Ejemplo 23. Consideramos el experimento aleatorio de tirar dos dados, uno de color rojo y otro azul.

Por el par $(2, 3)$ entendemos que con el dado rojo hemos obtenido 2 puntos y con el azul 3 puntos, el espacio muestral será

$$\Omega = \{(i, j) \mid i, j = 1, 2, 3, 4, 5, 6\}.$$

El espacio muestral está formado por 36 sucesos elementales igualmente probables y la probabilidad de cualquier suceso elemental $\{(i, j)\}$ es

$$P(\{(i, j)\}) = \frac{1}{36}, \quad i, j \in \{1, 2, \dots, 6\}.$$

Consideramos las variables aleatorias

$$\begin{aligned} X : \Omega &\longmapsto \mathbb{R} \\ (i, j) &\longrightarrow i \text{ (="número de puntos dado rojo),} \\ Y : \Omega &\longmapsto \mathbb{R} \\ (i, j) &\longrightarrow \max\{i, j\} \left(\begin{array}{l} \text{="número de puntos mayor} \\ \text{de los dos obtenidos} \end{array} \right). \end{aligned} \quad (6)$$

Por ejemplo

- $X(2, 4) = 2, X(5, 3) = 5.$
- $Y(2, 2) = 2, Y(1, 2) = 2, Y(2, 1) = 2, Y(3, 5) = 5, Y(5, 4) = 5.$

La variable aleatoria X es discreta y su rango, valores que toma, es

$$\{x_1, x_2, x_3, x_4, x_5, x_6\} = \{1, 2, 3, 4, 5, 6\}.$$

La variable aleatoria Y es discreta y su rango, valores que toma, es

$$\{y_1, y_2, y_3, y_4, y_5, y_6\} = \{1, 2, 3, 4, 5, 6\}.$$

Consideramos el vector aleatorio

$$\begin{aligned} (X, Y) : \Omega &\longmapsto \mathbb{R}^2 \\ (i, j) &\longrightarrow (X, Y)(i, j) = (X(i, j), Y(i, j)) = (i, \max\{i, j\}). \end{aligned}$$

El rango del vector aleatorio (X, Y) es el subconjunto de \mathbb{R}^2

$$\text{rango}(X, Y) = \{(X(i, j), Y(i, j)) = (i, \max\{i, j\}), \quad i, j \in \{1, 2, \dots, 6\}\}.$$

Vamos a determinar este conjunto:

Primeramente determinamos las imágenes por el vector aleatorio (X, Y) de los elementos del espacio muestral que son de la forma

$$\{(X, Y)(3, j), \quad j \in \{1, 2, \dots, 6\}\} = \{(3, 3), (3, 4), (3, 5), (3, 6)\}.$$

$$\begin{aligned} (X, Y)(3, 1) &= (X, Y)(3, 2) = (X, Y)(3, 3) = (3, 3), \\ (X, Y)(3, 4) &= (3, 4), \\ (X, Y)(3, 5) &= (3, 5), \\ (X, Y)(3, 6) &= (3, 6). \end{aligned}$$

El estudiante puede comprobar que

$$\text{rango}(X, Y) = \left\{ \begin{array}{cccccc} (1,1), & (1,2), & (1,3), & (1,4), & (1,5), & (1,6), \\ & (2,2), & (2,3), & (2,4), & (2,5), & (2,6), \\ & & (3,3), & (3,4), & (3,5), & (3,6), \\ & & & (4,4), & (4,5), & (4,6), \\ & & & & (5,5), & (5,6), \\ & & & & & (6,6), \end{array} \right\}$$

Es un conjunto finito, 21 elementos, de \mathbb{R}^2 . Como el rango es finito, decimos que el **vector aleatorio** (X, Y) es **discreto**.

Si A es un suceso elemental, por ejemplo $A = (2, 3)$, entonces $P(2, 3)$ suele designarse por $P(X = 2, Y = 3)$.

Como todos sucesos elementales son igualmente probables, podemos utilizar la regla de Laplace.

- $$\begin{aligned} P(2, 3) &\equiv P(X = 2, Y = 3) = P(\{(i, j) \in \Omega : X(i, j) = 2 \text{ y } Y(i, j) = 3\}) \\ &= \frac{\text{cardinal}\{\text{casos favorables}\}}{\text{cardinal}\Omega} = \frac{\text{cardinal}\{(2, 3)\}}{36} = \frac{1}{36}. \end{aligned}$$
- $$\begin{aligned} P(5, 2) &= P(X = 5, Y = 2) = P(\{(i, j) \in \Omega : X(i, j) = 5 \text{ y } Y(i, j) = 2\}) \\ &= \frac{\text{cardinal}\{\text{casos favorables}\}}{\text{cardinal}\Omega} = \frac{\text{cardinal}\emptyset}{36} = \frac{0}{36} = 0. \end{aligned}$$
- $$\begin{aligned} P(4, 4) &= P(X = 4, Y = 4) = P(\{(i, j) \in \Omega : X(i, j) = 4 \text{ y } Y(i, j) = 4\}) \\ &= \frac{\text{cardinal}\{\text{casos favorables}\}}{\text{cardinal}\Omega} = \frac{\text{cardinal}\{(4, 1), (4, 2), (4, 3), (4, 4)\}}{36} = \frac{4}{36} = \frac{1}{9}. \end{aligned}$$
- $$\begin{aligned} P(8, 3) &= P(\{(i, j) \in \Omega : X(i, j) = 8 \text{ y } Y(i, j) = 3\}) \\ &= \frac{\text{cardinal}\{\text{casos favorables}\}}{\text{cardinal}\Omega} = \frac{\text{cardinal}\emptyset}{36} = \frac{0}{36} = 0. \end{aligned}$$
- $$\begin{aligned} P(\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 25\}) &= P(\{(i, j) \in \Omega : (X(i, j), Y(i, j)) \in \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 25\}\}) \\ &= P(\{(i, j) \in \Omega : X^2(i, j) + Y^2(i, j) \leq 25\}) = \frac{\text{cardinal}\{\text{casos favorables}\}}{\text{cardinal}\Omega} \\ &= \frac{\text{cardinal}\{(1, 1), (1, 2), (1, 3), (1, 4), (2, 2), (2, 3), (2, 4), (3, 3), (3, 4)\}}{36} = \frac{9}{36} = \frac{1}{4}. \end{aligned}$$

Recordamos que si (X, Y) es un vector aleatorio discreto con

$$\text{rango de } (X, Y) = \{(x_i, y_j), i = 1, 2, \dots, n, j = 1, 2, \dots, m\},$$

la **distribución o función de probabilidad conjunta del vector aleatorio** (X, Y) viene dada por

$$\{P(X = x_i, Y = y_j) = P(\{\omega \in \Omega : X(\omega) = x_i \text{ e } Y(\omega) = y_j\})\}_{1 \leq i \leq n, 1 \leq j \leq m}.$$

En la siguiente tabla se muestra la distribución de probabilidad conjunta del vector aleatorio (X, Y) .

$P(X=i, Y=j)$	$Y=1$	$Y=2$	$Y=3$	$Y=4$	$Y=5$	$Y=6$
$X=1$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
$X=2$	0	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
$X=3$	0	0	$\frac{3}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
$X=4$	0	0	0	$\frac{4}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
$X=5$	0	0	0	0	$\frac{5}{36}$	$\frac{1}{36}$
$X=6$	0	0	0	0	0	$\frac{6}{36}$

Asociado al vector aleatorio (X, Y) , y dado que X e Y son variables aleatorias, existe **dos distribuciones o funciones de probabilidad marginales**:

■ **Distribución o función de probabilidad marginal de X**

Esta suele designarse usualmente por $P_X(x_i)$ o $P(X = x_i)$. Nosotros utilizaremos normalmente la última notación. $P(X = 2)$ nos da la probabilidad de que la variable aleatoria X tome el valor 2 independientemente del valor que tome la variable aleatoria Y .

$$P(X = x_i) = \sum_{j=1}^6 P(X = x_i, Y = y_j), \quad i = 1, 2, \dots, 6.$$

Ejemplos. Vamos a calcular $P(X = 2)$ y $P(X = 4)$.

$$\begin{aligned} P(X = 2) &= P\left(\bigcup_{j=1}^6 \{X = 2, Y = j\}\right) = \sum_{j=1}^6 P(X = 2, Y = j) \\ &= P(X = 2, Y = 1) + P(X = 2, Y = 2) + P(X = 2, Y = 3) \\ &\quad + P(X = 2, Y = 4) + P(X = 2, Y = 5) + P(X = 2, Y = 6) \\ &= 0 + \frac{2}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} = \frac{1}{6}. \end{aligned}$$

$$\begin{aligned} P(X = 4) &= P(X = 4, Y = 1) + P(X = 4, Y = 2) + P(X = 4, Y = 3) \\ &\quad + P(X = 4, Y = 4) + P(X = 4, Y = 5) + P(X = 4, Y = 6) \\ &= 0 + 0 + 0 + \frac{4}{36} + \frac{1}{36} + \frac{1}{36} = \frac{1}{6}. \end{aligned}$$

■ **Distribución o función de probabilidad marginal de Y**

$$P(Y = y_j) = \sum_{i=1}^6 P(X = x_i, Y = y_j), \quad j = 1, 2, \dots, 6.$$

Ejemplos. Vamos a calcular $P(Y = 4)$ y $P(Y = 5)$.

$$P(Y = 4) = \sum_{i=1}^6 P(X = i, Y = 4)$$

$$\begin{aligned} &P(X = 1, Y = 4) + P(X = 2, Y = 4) + P(X = 3, Y = 4) \\ &+ P(X = 4, Y = 4) + P(X = 5, Y = 4) + P(X = 6, Y = 4) \\ &= \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{4}{36} + 0 + 0 = \frac{7}{36}. \end{aligned}$$

$$\begin{aligned} P(Y = 5) &= P(X = 1, Y = 5) + P(X = 2, Y = 5) + P(X = 3, Y = 5) \\ &+ P(X = 4, Y = 5) + P(X = 5, Y = 5) + P(X = 6, Y = 5) \\ &= \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + 0 = \frac{5}{36}. \end{aligned}$$

En la siguiente tabla se muestra la distribución conjunta de probabilidad del vector aleatoria (X, Y) , $P(X = x_i, Y = y_j)$, $i, j \in \{1, 2, 3, 4, 5, 6\}$, y las distribuciones de probabilidad marginales, $P(X = x_i)$, $i \in \{1, 2, 3, 4, 5, 6\}$ y $P(Y = y_j)$, $j \in \{1, 2, 3, 4, 5, 6\}$.

.	Y=1	Y=2	Y=3	Y=4	Y=5	Y=6	P(X=i)=
X=1	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$P(X=1)=\frac{1}{6}$
X=2	0	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$P(X=2)=\frac{1}{6}$
X=3	0	0	$\frac{3}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$P(X=3)=\frac{1}{6}$
X=4	0	0	0	$\frac{4}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$P(X=4)=\frac{1}{6}$
X=5	0	0	0	0	$\frac{5}{36}$	$\frac{1}{36}$	$P(X=5)=\frac{1}{6}$
X=6	0	0	0	0	0	$\frac{6}{36}$	$P(X=6)=\frac{1}{6}$
P(Y=j)=	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$	$P(X=6)=\frac{1}{6}$

Probabilidad condicionada

En algunos casos, podemos poseer información sobre el resultado de una de las variables, información que puede ser útil para tener información sobre la otra variable. Así podemos hablar de $P(X = 3|Y = 3)$, que leemos como la probabilidad que la variable aleatoria X tome el valor 3 cuando sabemos que la variable aleatoria Y ha tomado el valor 3,

$$P(X = 3|Y = 3) = \frac{P(X = 3, Y = 3)}{P(Y = 3)} = \frac{\frac{3}{36}}{\frac{5}{36}} = \frac{3}{5},$$

$P(Y = 4|X = 5)$, que leemos como la probabilidad que la variable aleatoria Y tome el valor 4 cuando sabemos que la variable aleatoria X ha tomado el valor 5,

$$P(Y = 4|X = 5) = \frac{P(X = 5, Y = 4)}{P(X = 5)} = \frac{0}{\frac{1}{6}} = 0,$$

o $P(Y = 1|X \leq 3)$ que lo leemos como la probabilidad de que la variable aleatoria Y tome el valor 1 cuando sabemos que la variable aleatoria X ha tomado un valor menor o igual que 3,

$$\begin{aligned} P(Y = 1|X \leq 3) &= \frac{P(\{X \leq 3\} \cap \{Y = 1\})}{P(X \leq 3)} \\ &= \frac{P(\{X = 1, Y = 1\} \cup \{X = 2, Y = 1\} \cup \{X = 3, Y = 1\})}{P(\{X = 1\} \cup \{X = 2\} \cup \{X = 3\})} \\ &= \frac{P(X = 1, Y = 1) + P(X = 2, Y = 1) + P(X = 3, Y = 1)}{P(X = 1) + P(X = 2) + P(X = 3)} = \frac{\frac{1}{36} + 0 + 0}{\frac{1}{6} + \frac{1}{6} + \frac{1}{6}} = \frac{1}{2}. \end{aligned}$$

También podemos hablar de la variable aleatoria $Y|X > 2$. Para esto, siempre hemos obtenido con el dado rojo una puntuación mayor que dos, entonces el espacio muestral asociado a esta nueva situación es

$$\Omega_{X>2} = \{(3, j), (4, j), (5, j), (6, j), j \in \{1, 2, 3, 4, 5, 6\}\},$$

y la variable aleatoria $Y|X > 2$ está definida en $\Omega_{X>2}$ por

$$\begin{aligned} Y|X > 2 : \Omega_{X>2} &\mapsto \mathbb{R} \\ (3, j) &\longrightarrow (Y|X > 2)(3, j) = \max\{3, j\} = \begin{cases} 3 & j \leq 3 \\ j & j > 3 \end{cases} \\ (4, j) &\longrightarrow (Y|X > 2)(3, j) = \max\{4, j\} = \begin{cases} 4 & j \leq 4 \\ j & j > 4 \end{cases} \\ (5, j) &\longrightarrow (Y|X > 2)(3, j) = \max\{5, j\} = \begin{cases} 5 & j \leq 5 \\ j & j > 5 \end{cases} \\ (6, j) &\longrightarrow (Y|X > 2)(6, j) = \max\{6, j\} = 6, \end{aligned}$$

cuya distribución o función de probabilidad es el conjunto de valores

$$\begin{aligned} P(Y = 1|X > 2) &= \frac{P(\{X>2\} \cap \{Y=1\})}{P(X>2)} = 0, \\ P(Y = 2|X > 2) &= \frac{P(\{X>2\} \cap \{Y=2\})}{P(X>2)} = 0, \\ P(Y = 3|X > 2) &= \frac{P(\{X>2\} \cap \{Y=3\})}{P(X>2)} = \frac{1}{8}, \\ P(Y = 4|X > 2) &= \frac{P(\{X>2\} \cap \{Y=4\})}{P(X>2)} = \frac{5}{24}, \\ P(Y = 5|X > 2) &= \frac{P(\{X>2\} \cap \{Y=5\})}{P(X>2)} = \frac{7}{24}, \\ P(Y = 6|X > 2) &= \frac{P(\{X>2\} \cap \{Y=6\})}{P(X>2)} = \frac{1}{3}. \end{aligned}$$

Independencia de variables aleatorias.

Sin haber definido en concepto de variables aleatorias independientes, X e Y deben ser claramente **dependientes**, ya que el valor que toma la variable Y depende del que toma X . Si el dado rojo sale 4 y el azul 3, la variable aleatoria X toma el valor 4 y la variable aleatoria Y toma el valor 4, pero si el primer dado rojo sale 2 y el azul 3, la variable aleatoria X toma el valor 2 mientras que la Y toma el valor 3.

Definición 19. Sea (Ω, \mathcal{A}, P) un espacio probabilístico y X e Y dos variables aleatorias discretas. Sean A y B dos subconjuntos arbitrarios de \mathbb{R} , decimos que X e Y son variables aleatorias independientes si los sucesos

$$\{\omega \in \Omega : X(\omega) \in A\} \quad \text{y} \quad \{\omega \in \Omega : Y(\omega) \in B\}$$

son independientes, o dicho de otra modo, si

$$P(X \in A, Y \in B) = P(X \in A) \cdot P(Y \in B).$$

Si X e Y son independientes, tomando $A = (-\infty, x]$ y $B = (-\infty, y]$, donde x e y son dos números reales cualesquiera, tenemos

$$\begin{aligned} F(x, y) &= P(X \leq x, Y \leq y) = P(X \in A, Y \in B) = P(X \in A) \cdot P(Y \in B) \\ &= P(X \leq x) \cdot P(Y \leq y) = F_X(x) \cdot F_Y(y), \end{aligned}$$

donde F es la función de distribución del vector aleatorio (X, Y) y F_X, F_Y son las distribuciones marginales de X e Y respectivamente.

Es fácil demostrar el siguiente resultado que caracteriza la independencia de variables aleatorias discretas. **Proposición 12.** Sea (Ω, \mathcal{A}, P) un espacio probabilístico y X e Y dos variables aleatorias discretas con

$$\text{rango de } X = \{x_1, x_2, \dots\} \quad \text{y} \quad \text{rango de } Y = \{y_1, y_2, \dots\}.$$

X e Y son variables aleatorias independientes si y solamente si

$$P(X = x_i, Y = y_j) = P(X = x_i) \cdot P(Y = y_j), \quad x_i \in \text{rango de } X, \quad y_j \in \text{rango de } Y.$$

Observamos en el ejemplo con el que estamos trabajando, que

$$P(X = 3, Y = 5) = \frac{1}{36} \neq P(X = 3) \cdot P(Y = 5) = \frac{1}{6} \cdot \frac{9}{36} = \frac{1}{24}.$$

Medidas de centralización y dispersión marginales

Las distribuciones de probabilidad marginales de X e Y son distribuciones de probabilidad de variables aleatorias, y por lo tanto podemos considerar su esperanza, mediana, varianza y desviación típica.

La esperanza y varianza de X se conocen con el nombre de **esperanza y varianza marginal de X** y son designadas por μ_X y σ_X^2 respectivamente.

$$E[X] = \mu_X = \sum_{i=1}^6 x_i P(X = x_i) = 1 \cdot P(X = 1) + 2 \cdot P(X = 2) + 3 \cdot P(X = 3)$$

$$\begin{aligned}
 & +4 \cdot P(X=4) + 5 \cdot P(X=5) + 6 \cdot P(X=6) \\
 & = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{21}{6} = 3,5.
 \end{aligned}$$

$$E[Y] = \mu_Y = \sum_{j=1}^6 y_j P(Y=y_j) = 1 \cdot P(Y=1) + 2 \cdot P(Y=2) + 3 \cdot P(Y=3)$$

$$\begin{aligned}
 & +4 \cdot P(Y=4) + 5 \cdot P(Y=5) + 6 \cdot P(Y=6) \\
 & = 1 \cdot \frac{1}{36} + 2 \cdot \frac{3}{36} + 3 \cdot \frac{5}{36} + 4 \cdot \frac{7}{36} + 5 \cdot \frac{9}{36} + 6 \cdot \frac{11}{36} = \frac{161}{36} \cong 4,472.
 \end{aligned}$$

$$V(X) = \sigma_X^2 = \sum_{i=1}^6 x_i^2 P(X=x_i) - E[X]^2 = 1^2 \cdot P(X=1) + 2^2 \cdot P(X=2) + 3^2 \cdot P(X=3)$$

$$\begin{aligned}
 & +4^2 \cdot P(X=4) + 5^2 \cdot P(X=5) + 6^2 \cdot P(X=6) - 3,5^2 \\
 & = 1 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 9 \cdot \frac{1}{6} + 16 \cdot \frac{1}{6} + 25 \cdot \frac{1}{6} + 36 \cdot \frac{1}{6} - 12,25 = \frac{91}{6} - 12,25 \cong 2,916.
 \end{aligned}$$

$$V(Y) = \sigma_Y^2 = \sum_{j=1}^6 y_j^2 P(Y=y_j) - E^2 = 1^2 \cdot P(Y=1) + 2^2 \cdot P(Y=2) + 3^2 \cdot P(Y=3)$$

$$\begin{aligned}
 & +4^2 \cdot P(Y=4) + 5^2 \cdot P(Y=5) + 6^2 \cdot P(Y=6) - 4,472^2 \\
 & = 1 \cdot \frac{1}{36} + 4 \cdot \frac{3}{36} + 9 \cdot \frac{5}{36} + 16 \cdot \frac{7}{36} + 25 \cdot \frac{9}{36} + 36 \cdot \frac{11}{36} - 20 = \frac{791}{36} - 20 \cong 2.
 \end{aligned}$$

Suma de variables aleatorias

$X + Y$ es una variable aleatoria y recordamos que está definida por

$$\begin{aligned}
 X + Y : \Omega & \longrightarrow \mathbb{R} \\
 (i, j) & \longrightarrow X(i, j) + Y(i, j) = i + \max\{i, j\}.
 \end{aligned}$$

$$\begin{aligned}
 P(X + Y = 6) & = P(\{(i, j), i, j = 1, 2, 3, 4, 5, 6 : X(i, j) + Y(i, j) = 6\}) \\
 & = P(\{(3, 1), (3, 2), (3, 3), (2, 4), (1, 5)\}) = \frac{5}{36},
 \end{aligned}$$

$$\begin{aligned}
 P(Y = 3 | X + Y = 6) & = \frac{P(\{Y(i, j) = 3\} \cap \{X + Y = 6\})}{P(X + Y = 6)} \\
 & = \frac{P(\{((1, 3), (2, 3), (3, 3), (3, 1), (3, 2)) \cap \{(3, 1), (3, 2), (3, 3), (2, 4), (1, 5)\})}{P(\{(3, 1), (3, 2), (3, 3), (2, 4), (1, 5)\})} = \frac{\frac{3}{36}}{\frac{5}{36}} = \frac{3}{5}.
 \end{aligned}$$

Producto de variables aleatorias

XY es una variable aleatoria y recordamos que está definida por

$$\begin{aligned} XY : \Omega &\longrightarrow \mathbb{R} \\ (i, j) &\longrightarrow X(i, j) \cdot Y(i, j) = i \cdot \max\{i, j\}. \end{aligned}$$

$$P(XY = 12) = P(\{(i, j), i, j = 1, 2, 3, 4, 5, 6 : X(i, j) \cdot Y(i, j) = 12\})$$

$$P(\{(2, 6), (3, 4)\}) = \frac{2}{36} = \frac{1}{18}.$$

□

Ejercicio 37. Experimento aleatorio: Colocación aleatoria de tres bolas diferentes a, b, c en tres celdas c_1, c_2, c_3 .

Vamos a ver todas las posibilidades que tenemos:

1. Que en la primera celda no se coloque ninguna bola, las tres bolas las tenemos que colocar en 2 celdas $= 2^3 = 8$.
2. Que en la primera celda se coloque una bola, tenemos tres posibilidades, y fijada una posibilidad, para esta dos bolas se tendrían que colocar en dos celdas $3 \cdot 2^2 = 12$.
 $c_1(a), c_2(bc), c_3(\emptyset); \quad c_1(a), c_2(\emptyset), c_3(bc); \quad c_1(a), c_2(b), c_3(c); \quad c_1(a), c_2(c), c_3(b);$
 $c_1(b), c_2(ac), c_3(\emptyset); \quad c_1(b), c_2(\emptyset), c_3(ac); \quad c_1(b), c_2(a), c_3(c); \quad c_1(b), c_2(c), c_3(a);$
 $c_1(c), c_2(ab), c_3(\emptyset); \quad c_1(c), c_2(\emptyset), c_3(ab); \quad c_1(c), c_2(a), c_3(b); \quad c_1(c), c_2(c), c_3(a).$
3. Que dos bolas estén colocadas en la primera celda, tenemos $\binom{3}{2} = 3$ posibilidades, y para cada una de estas posibilidades, tendremos que colocar una bola en una de las dos celdas que quedan $= 2$ posibilidades. En total $= 3 \cdot 2 = 6$.
4. Que las tres bolas estén en la primera celda, solamente hay una posibilidad.

Espacio muestral Ω = El conjunto de todas las posibilidades. El cardinal es 27.

Vamos a suponer que todos los sucesos son igualmente probables.

Definimos las variables aleatorias

$$\begin{aligned} N : \Omega &\longrightarrow \mathbb{R} \\ \omega &\longrightarrow N(\omega) = \text{número de celdas ocupadas,} \end{aligned}$$

por ejemplo $N\{c_1(b), c_2(ac), c_3(\emptyset)\} = 2$, $N\{c_1(c), c_2(a), c_3(b)\} = 3$.

$$\begin{aligned} X : \Omega &\longrightarrow \mathbb{R} \\ \omega &\longrightarrow X(\omega) = \text{número de bolas en la primera celda,} \end{aligned}$$

por ejemplo $N\{c_1(b), c_2(ac), c_3(\emptyset)\} = 1$, $N\{c_1(cb), c_2(a), c_3(\emptyset)\} = 2$,
 $N\{c_1(\emptyset), c_2(abc), c_3(\emptyset)\} = 0$.

Vamos a calcular la distribución conjunta (N, X) .

$$P(N = 1, X = 0) = \frac{\text{cardinal de}\{\text{casos favorables}\}}{\text{cardinal de}\{\text{casos posibles}\}}$$

$$= \frac{\text{cardinal de}\{c_1(\emptyset), c_2(abc), c_3(\emptyset); c_1(\emptyset), c_2(\emptyset), c_3(abc)\}}{27} = \frac{2}{27}.$$

$N \setminus X$	0	1	2	3	Distribución de N
1	$\frac{2}{27}$	0	0	$\frac{1}{27}$	$\frac{1}{9}$
2	$\frac{6}{27}$	$\frac{6}{27}$	$\frac{6}{27}$	0	$\frac{2}{3}$
3	0	$\frac{6}{27}$	0	0	$\frac{2}{9}$
Distribución de X	$\frac{8}{27}$	$\frac{12}{27}$	$\frac{6}{27}$	$\frac{1}{27}$	

$N + X$ es una variable aleatoria.

$$P(N + X = 1) = P(N = 1; X = 0) = \frac{2}{27}.$$

$$P(N + X = 2) = P(N = 2; X = 0) + P(N = 1; X = 1) = \frac{6}{27} + 0 = \frac{6}{27}.$$

$$P(N + X = 3) = P(N = 3; X = 0) + P(N = 2; X = 1) + P(N = 1; X = 2)$$

$$= 0 + \frac{6}{27} + 0 = \frac{6}{27}.$$

$$P(N + X = 4) = P(N = 3; X = 1) + P(N = 2; X = 2) + P(N = 1; X = 3)$$

$$= \frac{6}{27} + \frac{6}{27} + \frac{1}{27} = \frac{13}{27}.$$

$$P(N + X = 5) = P(N = 3; X = 2) + P(N = 2; X = 3) = 0 + 0 = 0.$$

$$P(N + X = 6) = P(N = 3; X = 3) = 0.$$

$N \cdot X = NX$ es una variable aleatoria.

$$P(NX = 0) = P(N = 1, X = 0) + P(N = 2, X = 0) + P(N = 3, X = 0)$$

$$\frac{2}{27} + \frac{6}{27} + 0 = \frac{8}{27}.$$

$$P(NX = 1) = P(N = 1, X = 1) = 0.$$

$$P(NX = 2) = P(N = 2, X = 1) + P(N = 1; X = 2) = \frac{6}{27} + 0 = \frac{6}{27}.$$

$$P(NX = 3) = P(N = 3, X = 1) + P(N = 1; X = 3) = \frac{6}{27} + \frac{1}{27} = \frac{7}{27}.$$

$$P(NX = 4) = P(N = 2, X = 2) = \frac{6}{27}.$$

$$P(NX = 6) = P(N = 3, X = 2) + P(N = 2; X = 3) = 0 + 0 = 0.$$

$$P(NX = 9) = P(N = 3, X = 3) = 0.$$

□

Ejercicio 38. Sea (Ω, \mathcal{A}, P) un espacio probabilístico y (X, Y) un vector aleatorio discreto cuya función de probabilidad conjunta es

$Y \setminus X$	5	10	15
1	0.25	0.15	0.32
2	0.10	0.05	0.13

1. Calcule las distribuciones de probabilidad marginales.
2. Distribución de probabilidad de $X|Y = 2$.
3. Distribución de probabilidad de $Y|X = 5$.

Ejercicio 39. Se tiene una urna que contiene dos bolas blancas y tres negras. Efectuamos dos extracciones sucesivas con reemplazamiento y definimos las variables aleatorias

$$X(\omega) = \begin{cases} 1 & \text{si la primera bola extraída es blanca} \\ 0 & \text{si la primera bola extraída es negra,} \end{cases}$$

$$Y(\omega) = \begin{cases} 1 & \text{si la segunda bola extraída es blanca} \\ 0 & \text{si la segunda bola extraída es negra,} \end{cases}$$

Calcule

1. Distribuciones de probabilidad marginales.
2. ¿Son independientes?
3. Distribución de probabilidad de la variable $X|Y = 0$ e $Y|X = 1$.
4. $E[X]$, $E[Y]$ y σ_{XY} .

Ejercicio 40. Se lanzan dos dados, uno rojo y otro azul, y consideremos como espacio muestral del experimento aleatorio

$$\Omega = \left\{ (i, j), \begin{array}{l} i = \text{puntuación del dado rojo} \\ j = \text{puntuación del dado azul} \end{array}, i, j \in \{1, 2, 3, 4, 5, 6\} \right\}.$$

Consideramos las variables aleatorias:

$$\begin{array}{ll} X : \Omega \mapsto \mathbb{R} & Y : \Omega \mapsto \mathbb{R} \\ (i, j) \mapsto X(i, j) = i + j & (i, j) \mapsto Y(i, j) = |i - j| \end{array}$$

1. Rango de las variables X, Y y del vector aleatorio (X, Y) .
2. Calcule $P(X \leq 4, Y = 2)$.
3. Obtenga las distribuciones de probabilidad conjunta y marginales.
4. Calcule la función de probabilidad de la variable aleatoria $X|(Y \leq 2)$.
5. Calcule $E[X], E[Y], V[X]$ y $V[Y]$.

Solución: 1. rango de $X = \{2, 3, \dots, 12\}$, rango de $Y = \{0, 1, 2, 3, 4, 5\}$, rango de $(X, Y) = \{(i, j) : i \in \{2, 3, \dots, 12\}, j \in \{0, 1, 2, 3, 4, 5\}\}$;

2. $\frac{1}{18}$.

3.

Y \ X	2	3	4	5	6	7	8	9	10	11	12	P(Y=j)=
Y=0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	0	$\frac{1}{36}$	$P(Y=0)=\frac{1}{6}$
Y=1	0	$\frac{1}{18}$	0	$\frac{1}{18}$	0	$\frac{1}{18}$	0	$\frac{1}{18}$	0	$\frac{1}{18}$	0	$P(Y=1)=\frac{5}{18}$
Y=2	0	0	$\frac{1}{18}$	0	$\frac{1}{18}$	0	$\frac{1}{18}$	0	$\frac{1}{18}$	0	0	$P(Y=2)=\frac{4}{18}$
Y=3	0	0	0	$\frac{1}{18}$	0	$\frac{1}{18}$	0	$\frac{1}{18}$	0	0	0	$P(Y=3)=\frac{3}{18}$
Y=4	0	0	0	0	$\frac{1}{18}$	0	$\frac{1}{18}$	0	0	0	0	$P(Y=4)=\frac{2}{18}$
Y=5	0	0	0	0	0	$\frac{1}{18}$	0	0	0	0	0	$P(Y=5)=\frac{1}{18}$
P(X=i)	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	

4.

$X (Y \leq 2)$	2	3	4	5	6	7	8	9	10	11	12
$P(X = x_i Y \leq 2)$	$\frac{1}{24}$	$\frac{2}{24}$	$\frac{2}{24}$	$\frac{4}{24}$	$\frac{5}{24}$	$\frac{6}{24}$	$\frac{5}{24}$	$\frac{4}{24}$	$\frac{3}{24}$	$\frac{2}{24}$	$\frac{1}{24}$

5.2. Vectores aleatorios continuos

Definición 20. Sea (Ω, \mathcal{A}) un espacio probabilístico y (X, Y) un vector aleatorio bi-dimensional. Diremos que (X, Y) es continuo, si existe una función $f(x, y)$ integrable en \mathbb{R}^2 tal que

$$1. f(x, y) \geq 0, \quad (x, y) \in \mathbb{R}^2,$$

$$2. \int \int_{\mathbb{R}^2} f(x, y) dx dy = 1,$$

$$3. \text{ Si } A \subset \mathbb{R}^2, \quad P(\{\omega \in \Omega; (X, Y)(\omega) \in A\}) = \int \int_A f(x, y) dx dy.$$

f es conocida como la función de densidad asociada al vector aleatorio (X, Y) .

La **función de distribución conjunta** está dada por

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv,$$

y si $f(x, y)$ es continua, se tiene que

$$\frac{\partial^2 F}{\partial y \partial x}(x, y) = f(x, y).$$

Ejemplo 24. Para entender la definición anterior, como es habitual, vamos a trabajar con un ejemplo.

Queremos estudiar el porcentaje de magnesio y manganeso en las piezas de duraluminio que produce una cierta fábrica.

Experimento aleatorio: tomar al azar una pieza de las producidas durante el último mes y anotar su porcentaje de aluminio, cobre, zinc, magnesio y manganeso.

El espacio muestral asociado Ω está formado por todas las piezas producidas durante el último mes en la fábrica.

El porcentaje usual del manganeso en el duraluminio es entre el 0,25 % y el 1 %, mientras que del magnesio 0,5 % y el 2 %.

Consideramos las variables aleatorias:

$$\begin{aligned} X : \Omega &\longmapsto \mathbb{R} \\ \omega \text{ (pieza)} &\longrightarrow X(\omega) = \text{"porcentaje de manganeso"}, \end{aligned}$$

$$\begin{aligned} Y : \Omega &\longmapsto \mathbb{R} \\ \omega \text{ (pieza)} &\longrightarrow Y(\omega) = \text{"porcentaje de magnesio"} \end{aligned}$$

Ambas variables aleatorias son continuas, el rango de X , cantidad manganeso de la pieza, podemos suponer que puede ser cualquier valor ente 0 y 4/3, mientras que el

rango de Y , cantidad de magnesio de la pieza, vamos a suponer que es un valor entre 0 y $5/2$.

Consideramos el vector aleatorio (X, Y) , que es la función vectorial

$$\begin{aligned} (X, Y) : \Omega &\longrightarrow \mathbb{R}^2 \\ \omega \text{ (pieza)} &\longrightarrow (X, Y)(\omega) = (X(\omega), Y(\omega)) \\ &= (\text{"porcentaje de manganeso"}, \text{"porcentaje de magnesio"}). \end{aligned}$$

Conocemos la distribución conjunta de probabilidad del vector aleatorio (X, Y) , y viene dada por la función de densidad

$$f(x, y) = \begin{cases} kx(4 - 3x)y^2(5 - 2y) & (x, y) \in (0, \frac{4}{3}) \times (0, \frac{5}{2}) \\ 0 & (x, y) \notin (0, \frac{4}{3}) \times (0, \frac{5}{2}). \end{cases}$$

siendo k una constante que tendremos que determinar.

¿Qué significa que f se la función de densidad asociada al vector aleatorio (X, Y) . Si $A \subset \mathbb{R}^2$, entonces

$$P(A) = P(\{\omega \in \Omega : (X, Y)(\omega) = (X(\omega), Y(\omega)) \in A\}) = \int \int_A f(x, y) dx dy.$$

Por ejemplo, la probabilidad de que una pieza tenga un porcentaje de manganeso entre $3/4$ y 1 y un porcentaje de magnesio entre $3/2$ y 2 es

$$P((3/4, 1) \times (3/2, 2)) = \int \int_{(3/4, 1) \times (3/2, 2)} f(x, y) dx dy.$$

Para que f sea una función de densidad debe de verificas unas propiedades, que nos permitirán determinar la constante k .

1. $f(x, y) \geq 0$.

Esto nos dice que la constante k debes ser no negativa.

2. $\int \int_{\mathbb{R}^2} f(x, y) dx dy = 1$.

Vamos a determinar k para que se verifique esta condición.

$$\begin{aligned} \int \int_{\mathbb{R}^2} f(x, y) dx dy &= k \int \int_{(0, \frac{4}{3}) \times (0, \frac{5}{2})} x(4 - 3x)y^2(5 - 2y) dx dy \\ &= k \int_0^{4/3} x(4 - 3x) \left(\int_0^{5/2} y^2(5 - 2y) dy \right) dx \\ &= k \left(\int_0^{4/3} x(4 - 3x) dx \right) \cdot \left(\int_0^{5/2} y^2(5 - 2y) dy \right) \end{aligned}$$

$$= k \left(2x^2 - x^3 \Big|_0^{4/3} \right) \cdot \left(\frac{5y^3}{3} - \frac{y^4}{2} \Big|_0^{5/2} \right) = k \frac{625}{81}.$$

Entonces

$$\int \int_{\mathbb{R}^2} f(x, y) dx dy = 1 \iff \boxed{k = \frac{81}{625}}.$$

y la función de densidad es

$$f(x, y) = \begin{cases} \frac{81}{625} x(4 - 3x)y^2(5 - 2y) & (x, y) \in (0, \frac{4}{3}) \times (0, \frac{5}{2}) \\ 0 & (x, y) \notin (0, \frac{4}{3}) \times (0, \frac{5}{2}). \end{cases}$$

□

Pregunta 1. Una pieza de duraluminio se considera que es aceptable, si su porcentaje de manganeso está entre el 1/2 % y el 1 % y el de magnesio entre 1 % y el 2 %. Si se elige al azar una pieza, ¿cuál es la probabilidad de que la pieza sea aceptable?

$$\begin{aligned} P\left(\frac{1}{2} \leq X \leq 1, 1 \leq Y \leq 2\right) &= P\left(\left(\frac{1}{2}, 1\right) \times (1, 2)\right) \\ &= \frac{81}{625} \int \int_{(\frac{1}{2}, 1) \times (1, 2)} x(4 - 3x)y^2(5 - 2y) dx dy \\ &= \frac{81}{625} \left(\int_{1/2}^1 x(4 - 3x) dx \right) \cdot \left(\int_1^2 y^2(5 - 2y) dy \right) \\ &= \frac{81}{625} \left(2x^2 - x^3 \Big|_{1/2}^1 \right) \cdot \left(\frac{5y^3}{3} - \frac{y^4}{2} \Big|_1^2 \right) \\ &= \frac{81}{625} \cdot \left(2 - 1 - \frac{2}{4} + \frac{1}{8} \right) \cdot \left(\frac{40}{3} - \frac{16}{2} - \frac{5}{3} + \frac{1}{2} \right) \cong 0,33. \end{aligned}$$

La probabilidad es muy baja. La fábrica está perdiendo mucho dinero

□

Como en el caso discreto, podemos definir las **distribuciones de probabilidad marginales** del vector aleatorio (X, Y) .

- La distribución marginal de X es una variable aleatoria continua con función de densidad asociada dada por

$$\boxed{f_X(x) = \int_{\mathbb{R}} f(x, y) dy}$$

$$= \begin{cases} \text{Si } x \in (0, 4/3) & \begin{cases} = \frac{81}{625} \int_{\mathbb{R}} x(4-3x)y^2(5-2y)dy \\ = \frac{81}{625} x(4-3x) \int_0^{5/2} y^2(5-2y)dy = \frac{27}{32} x(4-3x) \end{cases} \\ \text{Si } x \notin (0, 4/3) & = 0. \end{cases}$$

$$f_X(x) = \begin{cases} \frac{27}{32} x(4-3x) & x \in (0, \frac{4}{3}) \\ 0 & x \notin (0, \frac{4}{3}) \end{cases}$$

- La distribución marginal de Y es una variable aleatoria continua con función de densidad asociada dada por

$$f_Y(y) = \int_{\mathbb{R}} f(x, y) dx$$

$$= \begin{cases} \text{Si } y \in (0, 5/2) & \begin{cases} = \frac{81}{625} \int_{\mathbb{R}} x(4-3x)y^2(5-2y)dx \\ = \frac{81}{625} y^2(5-2y) \int_0^{4/3} x(4-3x)dx = \frac{96}{625} y^2(5-2y) \end{cases} \\ \text{Si } y \notin (0, 5/2) & = 0. \end{cases}$$

$$f_Y(y) = \begin{cases} \frac{96}{625} y^2(5-2y) & y \in (0, \frac{5}{2}) \\ 0 & y \notin (0, \frac{5}{2}) \end{cases}$$

□

Definición 21. Sea (Ω, \mathcal{A}, P) un espacio probabilístico y $\vec{X} = (X, Y)$ un vector aleatorio continuo. Se define la función de densidad marginal con respecto a la primera componente del vector aleatorio \vec{X} , (normalmente designada por función de densidad marginal de X), como

$$f_X(x) = \int_{\mathbb{R}} f(x, y) dy,$$

y la función de densidad marginal de Y por

$$f_Y(y) = \int_{\mathbb{R}} f(x, y) dx.$$

Observación 12. Puede comprobarse que f_X y f_Y son funciones de densidad de variables aleatorias unidimensionales asociadas al espacio probabilístico (Ω, \mathcal{A}, P) .

Independencia

Definición 22. Sea (Ω, \mathcal{A}, P) un espacio probabilístico y X e Y dos variables aleatorias continuas. Decimos que X e Y son independientes si para cada par de subconjuntos A y B de \mathbb{R} , los sucesos

$$\{\omega \in \Omega : X(\omega) \in A\} \quad \text{y} \quad \{\omega \in \Omega : Y(\omega) \in B\}$$

son independientes.

Como en el caso de las variables discretas, si X e Y son independientes, se verifica que

$$F(x, y) = F_X(x) \cdot F_Y(y),$$

donde F es la función de distribución conjunta de (X, Y) y F_X, F_Y son las funciones de distribución marginales de X e Y respectivamente.

El recíproco también es cierto. Si se cumple $F(x, y) = F_X(x) \cdot F_Y(y)$, se tiene que

$$\int_{-\infty}^x \int_{-\infty}^y f(u, v) du dv = \int_{-\infty}^x f_X(u) du \cdot \int_{-\infty}^y f_Y(v) dv,$$

y como consecuencia

$$f(x, y) = f_X(x) \cdot f_Y(y).$$

Sean A y B subconjuntos de \mathbb{R} , como

$$P(X \in A) = \int_A f_X(x) dx \quad \text{y} \quad P(Y \in B) = \int_B f_Y(y) dy,$$

tenemos

$$\begin{aligned} P(X \in A, Y \in B) &= \int \int_{A \times B} f(x, y) dx dy = \int \int_{A \times B} f_X(x) f_Y(y) dx dy \\ &= \int_A f_X(x) dx \cdot \int_B f_Y(y) dy = P(X \in A) \cdot P(Y \in B), \end{aligned}$$

lo que nos dice que las variables aleatorias X e Y son independientes. Establecemos el resultado:

Proposición 13. Sea (Ω, \mathcal{A}, P) un espacio probabilístico y X e Y dos variables aleatorias continuas. Las tres afirmaciones siguientes son equivalentes

1. X e Y son independientes,
2. $F(x, y) = F_X(x) \cdot F_Y(y)$, donde F es la función de distribución conjunta de (X, Y) y F_X, F_Y son las funciones de distribución marginales de X e Y respectivamente.
3. $f(x, y) = f_X(x) \cdot f_Y(y)$, donde f es la función de densidad conjunta de (X, Y) y f_X, f_Y son las funciones de densidad marginales de X e Y respectivamente.

Medidas de centralización y dispersión marginales

Definición 23. Sea (Ω, \mathcal{A}, P) un espacio probabilístico y (X, Y) un vector aleatorio continuo. Sea $f(x, y)$ su función de densidad conjunta y $f_X(x)$, $f_Y(y)$ las funciones de densidad marginales de las variables aleatorias X e Y respectivamente. Definimos

- **Esperanza marginal de X por**

$$\mu_X = E[X] = \int_{-\infty}^{\infty} x f_X(x) dx.$$

- **Varianza marginal de X por**

$$\sigma_X^2 = E[(X - \mu_X)^2] = \int_{-\infty}^{\infty} x^2 f_X(x) dx - \mu_X^2.$$

(análogamente definiríamos la esperanza y varianza marginal de Y)

Ejercicio 41. La función de densidad conjunta de dos variables aleatorias X e Y es

$$f(x, y) = \begin{cases} 6x & 0 < x < y < 1 \\ 0 & \text{en otro caso} \end{cases}$$

1. Calcule las funciones de densidad marginales.
2. Calcule la función de distribución conjunta y las marginales.
3. ¿Son independientes las variables X e Y ?
4. Calcule las esperanzas y varianzas marginales.

Solución

1. Empezamos calculando la función de densidad marginal de X .

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy = \begin{cases} \int_x^1 6x dy = 6x(1-x) & x \in (0, 1) \\ 0 & x \notin (0, 1) \end{cases}$$

Ahora la de Y .

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx = \begin{cases} \int_0^y 6x dx = 3y^2 & y \in (0, 1) \\ 0 & y \notin (0, 1) \end{cases}$$

2. Empezamos calculando F_X y F_Y .

$$F_X(x) = \int_{-\infty}^x f_X(t)dt = \begin{cases} 0 & x \leq 0 \\ \int_0^x 6t(1-t)dt = 3x^2 - 2x^3 & x \in (0, 1) \\ \int_0^1 6t(1-t)dt = 1 & x \geq 1 \end{cases}$$

$$F_Y(y) = \int_{-\infty}^y f_Y(t)dt = \begin{cases} 0 & y \leq 0 \\ \int_0^y 3t^2 dt = y^3 & y \in (0, 1) \\ \int_0^1 3t^2 dt = 1 & y \geq 1 \end{cases}$$

Ahora calculamos la conjunta

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v)dvdu = \begin{cases} \int_0^y \int_u^y 6udvdu = y^3 & 0 \leq y \leq 1, y \leq x \\ \int_0^x \int_u^y 6udvdu = 3yx^2 - 2x^3 & 0 < x < y < 1 \\ \int_0^x \int_u^1 6udvdu = 3x^2 - 3x^3 & 0 < x < 1, 1 \leq y \\ \int_0^1 \int_u^1 6udvdu = 1 & x \geq 1, y \geq 1 \\ 0 & \text{en el resto} \end{cases}$$

3. Calculamos el producto de las funciones de densidad marginales,

$$f_X(x) \cdot f_Y(y) = \begin{cases} 18x(1-x)y^2 & (x, y) \in (0, 1) \times (0, 1) \\ 0 & \text{en otros caso} \end{cases}$$

Claramente $f(x, y) \neq f_X(x) \cdot f_Y(y)$ y por lo tanto X e Y no son independientes.

4. Empezamos calculando las esperanzas marginales

$$\mu_X = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 6x^2(1-x) dx = \frac{1}{2},$$

$$\mu_Y = \int_{-\infty}^{\infty} y f_Y(y) dy = \int_0^1 3y^3 dy = \frac{3}{4}.$$

Calculamos ahora las varianzas marginales

$$\sigma_X^2 = \int_{-\infty}^{\infty} x^2 f_X(x) dx - \mu_X^2 = \int_0^1 6x^3(1-x) dx - \frac{1}{4} = \frac{1}{20},$$

$$\sigma_Y^2 = \int_{-\infty}^{\infty} y^2 f_Y(y) dy - \mu_Y^2 = \int_0^1 3y^4 dy - \frac{9}{16} = \frac{3}{5} - \frac{9}{16} = \frac{3}{80}.$$

□

Ejercicio 42. Una variable aleatoria bidimensional (X, Y) tiene una función de densidad uniforme en el recinto definido por las rectas $y = x, y = -x$ t $x = 1$.

1. Calcule la función de densidad de (X, Y) y las marginales.
2. Calcule la función de distribución de (X, Y) y las marginales.
3. ¿Son independientes las variables aleatorias X e Y .
4. Calcule las esperanzas y varianzas marginales.

5.3. Funciones y medidas características de un vector aleatorio

Sea (Ω, \mathcal{A}, P) un espacio probabilístico. Si a es un número real,

$$\begin{aligned} X_a : \Omega &\longmapsto \mathbb{R} \\ \omega &\longrightarrow X_a(\omega) = a, \end{aligned}$$

es una variable aleatoria (toma el valor a en cada suceso), y usualmente suele designarse solo con la letra a .

Supongamos que $X, Y : \Omega \longmapsto \mathbb{R}$ son dos variables aleatorias y $h : \mathbb{R}^2 \longmapsto \mathbb{R}$ una función continua. La variable aleatoria definida por la función h

$$\begin{aligned} h(X, Y) : \Omega &\longmapsto \mathbb{R} \\ \omega &\longrightarrow h(X, Y)(\omega), \end{aligned}$$

es una variable aleatoria.

Algunos ejemplos

■ $h(x, y) = x + y,$

$$\begin{aligned} X + Y : \Omega &\longmapsto \mathbb{R} \\ \omega &\longrightarrow (X + Y)(\omega) = X(\omega) + Y(\omega), \end{aligned}$$

■ $h(x, y) = x - y,$

$$\begin{aligned} X - Y : \Omega &\longmapsto \mathbb{R} \\ \omega &\longrightarrow (X - Y)(\omega) = X(\omega) - Y(\omega), \end{aligned}$$

■ $h(x, y) = xy,$

$$\begin{aligned} X \cdot Y : \Omega &\longmapsto \mathbb{R} \\ \omega &\longrightarrow (X \cdot Y)(\omega) = X(\omega) \cdot Y(\omega), \end{aligned}$$

- $h(x, y) = ax + by$, con a y b números reales,

$$\begin{aligned} aX + bY : \Omega &\longrightarrow \mathbb{R} \\ \omega &\longrightarrow (aX + bY)(\omega) = aX(\omega) + bY(\omega), \end{aligned}$$

- $h(x, y) = x^2 + 3xy - 2y^2$,

$$\begin{aligned} X^2 + 3XY - 2Y^2 : \Omega &\longrightarrow \mathbb{R} \\ \omega &\longrightarrow (X^2 + 3XY - 2Y^2)(\omega) = (X(\omega))^2 + 3X(\omega)Y(\omega) - 2(Y(\omega))^2, \end{aligned}$$

- $h(x, y) = |x - y|$,

$$\begin{aligned} |X - Y| : \Omega &\longrightarrow \mathbb{R} \\ \omega &\longrightarrow |X - Y|(\omega) = |X(\omega) - Y(\omega)|, \end{aligned}$$

- $h(x, y) = \max\{x, y\}$,

$$\begin{aligned} \max\{X, Y\} : \Omega &\longrightarrow \mathbb{R} \\ \omega &\longrightarrow \max\{X, Y\}(\omega) = \max\{X(\omega), Y(\omega)\}. \end{aligned}$$

□

Se puede demostrar que

- Si (X, Y) es una variable aleatoria discreta con

$$\text{rango de } (X, Y) = \{(x_i, y_j), \quad i, j = 1, 2, \dots\},$$

entonces

$$E[h(X, Y)] = \sum_i \sum_j h(x_i, y_j) P(X = x_i, Y = y_j).$$

- Si (X, Y) es una variable aleatoria continua con función de densidad $f(x, y)$, entonces

$$E[h(X, Y)] = \int \int_{\mathbb{R}} h(x, y) f(x, y) dx dy.$$

Ejercicio 43. La función de densidad conjunta del vector aleatorio (X, Y) está dada por

$$f(x, y) = \begin{cases} x + y & x, y \in (0, 1) \\ 0 & \text{en otro caso} \end{cases}$$

Calcule la esperanza de $Z = XY^2 + 2X$.

Solución

Consideramos la función $h(x, y) = xy^2 + 2x$. Entonces

$$\begin{aligned} E[Z] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) f(x, y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (xy^2 + 2x)(x + y) dx dy \\ &= \int_0^1 \left(\int_0^1 (x^2 y^2 + 2x^2 + xy^3 + 2xy) dx \right) dy = \int_0^1 \left(\frac{x^3}{3} y^2 + \frac{2x^3}{3} + \frac{x^2}{2} y^3 + x^2 y \Big|_0^1 \right) dy \\ &= \int_0^1 \left(\frac{y^2}{3} + \frac{2}{3} + \frac{y^3}{2} + y \right) dy = \frac{101}{72}. \end{aligned}$$

Proposición 14. Sea (ω, \mathcal{A}, P) un espacio probabilístico y (X, Y) un vector aleatorio. Si X e Y son independientes, entonces

$$E[XY] = E[X] \cdot E[Y] = \mu_X \cdot \mu_Y.$$

Demostración Vamos a hacerlo para el caso continuo. Sea $f(x, y)$ la función de densidad conjunta del vector aleatorio y f_X, f_Y las funciones de densidad marginales de las variables X e Y . Por ser independientes las variables X e Y , se verifica que

$$f(x, y) = f_X(x) f_Y(y).$$

Consideramos la función $h(x, y) = xy$, tenemos

$$\begin{aligned} E[XY] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) f(x, y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_X(x) f_Y(y) dx dy \\ &= \int_{-\infty}^{\infty} x f_X(x) dx \cdot \int_{-\infty}^{\infty} y f_Y(y) dy. \end{aligned}$$

■

Proposición 15. Sea (ω, \mathcal{A}, P) un espacio probabilístico y (X, Y) un vector aleatorio. Si X e Y son independientes, entonces

$$V(X + Y) = V(X) + V(Y) = \sigma_X^2 + \sigma_Y^2.$$

Demostración Ya que X e Y son independientes, vamos a usar que $E[XY] = E[X]E[Y]$. También usaremos que la varianza de una variable aleatoria Z es $V(Z) = E[Z^2] - (E[Z])^2$.

$$\begin{aligned} V(X + Y) &= E[(X + Y)^2] - (E[X + Y])^2 = E[X^2 + 2XY + Y^2] - (E[X] + E[Y])^2 \\ &= E[X^2] + E[2XY] + E[Y^2] - (E[X])^2 - 2E[X]E[Y] - (E[Y])^2 \\ &= E[X^2] + 2E[X]E[Y] + E[Y^2] - (E[X])^2 - 2E[X]E[Y] - (E[Y])^2 \\ &= E[X^2] - (E[X])^2 + E[Y^2] - (E[Y])^2 = V(X) + V(Y). \end{aligned}$$

■

Ejercicio 44. Sea (Ω, \mathcal{A}, P) un espacio probabilístico y X e Y dos variables aleatorias independientes, con varianzas finitas y tal que $E[X] = E[Y]$.

1. Demuestre que X y αY son independientes, siendo α un número real.
2. Compruebe que $E[(X - Y)^2] = V(X) + V(Y)$.
3. Si $V(X) = V(Y) = 3$, determine los valores de $V(X - Y)$ y $V(2X - 3Y + 1)$.

Solución

1. Sean A y B dos subconjuntos arbitrarios de \mathbb{R} . Para ver que X y αY son independientes, tendremos que comprobar que los sucesos

$$\{\omega \in \Omega : X(\omega) \in A\} \quad \text{y} \quad \{\omega \in \Omega : \alpha Y(\omega) \in B\}$$

son independientes.

Ya que

$$\{\omega \in \Omega : \alpha Y(\omega) \in B\} = \{\omega \in \Omega : Y(\omega) \in \frac{1}{\alpha}B\},$$

donde $\frac{1}{\alpha}B$ es el suceso

$$\frac{1}{\alpha}B = \left\{ \frac{b}{\alpha} : \text{con } b \in B \right\}$$

y X e Y son independientes, los sucesos

$$\{\omega \in \Omega : X(\omega) \in A\} \quad \text{y} \quad \{\omega \in \Omega : Y(\omega) \in \frac{1}{\alpha}B\}$$

son independientes.

2. Usaremos que

- a) $E[XY] = E[X]E[Y]$, (X e Y son independientes),
- b) Si Z es una variable aleatoria, se tiene $V(Z) = E[Z^2] - (E[Z])^2$,
- c) $E[X] = E[Y]$, (hipótesis).

$$\begin{aligned} E[(X - Y)^2] &= E[X^2 - 2XY + Y^2] = E[X^2] - 2E[X]E[Y] + E[Y^2] \\ &= E[X^2] - E[X]E[Y] + E[Y^2] - E[X]E[Y] \\ &= E[X^2] - (E[X])^2 + E[Y^2] - (E[Y])^2 = V(X) + V(Y). \end{aligned}$$

3. Usaremos

- a) Si X e Y son independientes, αX y βY también lo son, cualquiera que sean los números reales α y β .
- b) $V(Z_1 + Z_2) = V(Z_1) + V(Z_2)$ si Z_1 y Z_2 son variables independientes.
- c) $V(\alpha Z + \beta) = \alpha^2 V(Z)$, siendo Z una variable aleatoria y αX y βY dos números reales cualesquiera.

$$\begin{aligned} V(X - Y) &= V(X + (-Y)) = V(X) + V(-Y) \\ &= V(X) + (-1)^2 V(Y) = 3 + (-1)^2 \cdot 3 = 6. \\ V(2X - 3Y + 1) &= V(2X - 3Y) = V(2X + (-3Y)) \\ &= V(2X) + V(-3Y) = 2^2 V(X) + (-3)^2 V(Y) = 2^2 \cdot 3 + (-3)^2 \cdot 3 = 39. \end{aligned}$$

□

Covarianza de un vector aleatorio.

Sea (X, Y) un vector aleatorio, la **covarianza** del vector aleatorio (X, Y) , que "mide la manera de variar conjuntamente X e Y ", la definimos por:

- si el vector es discreto con

$$\text{rango de } (X, Y) = \{(x_i, y_j) \in \mathbb{R}^2; i \in \{1, 2, \dots, n\}, j \in \{1, 2, \dots, m\},$$

$$\sigma_{XY} = \text{cov}(X, Y) = \sum_{i=1}^n \sum_{j=1}^m (x_i - E[X])(y_j - E[Y])P(X = x_i, Y = y_j),$$

- si el vector es continuo con función de densidad $f(x, y)$

$$\sigma_{XY} = \text{cov}(X, Y) = \int \int_{\mathbb{R}^2} (x - E[X])(y - E[Y])f(x, y)dx dy.$$

Propiedades de la covarianza. Sea (Ω, \mathcal{A}, P) un espacio probabilístico y (X, Y) un vector aleatorio, se verifica:

1. $\text{con}(X, Y) = \text{con}(Y, X)$.
2. $\text{con}(X, Y) = \sum_{i=1}^n \sum_{j=1}^m x_i y_j P(X = x_i, Y = y_j) - E[X]E[Y]$, si (X, Y) es discreto.
3. $\text{con}(X, Y) = \int \int_{\mathbb{R}^2} xyf(x, y)dx dy - E[X]E[Y]$, si (X, Y) es continuo.
4. $\text{con}(aX + b, cY + d) = a \cdot c \cdot \text{con}(X, Y)$, con a, b, c y d en \mathbb{R} .
5. Si X e Y son independientes, entonces $\text{con}(X, Y) = 0$.

6. si $a, b \in \mathbb{R}$,

$$V(aX + bY) = a^2V(X) + b^2V(Y) + 2a \cdot b \cdot \text{cov}(X, Y), \quad V(aX) = a^2V(X), \quad (7)$$

$$V(aX + bY) = a^2V(X) + b^2V(Y), \quad \text{si } X \text{ e } Y \text{ independientes.} \quad (8)$$

El vector aleatorio considerado en el primer ejemplo de vectores discretos, y definido por (5.1), $\text{cov}(X, Y)$ está dada por

$$\text{con}(X, Y) = \sum_{i=1}^6 \sum_{j=1}^6 (x_i - E[X])(y_j - E[Y])P(X = x_i, Y = y_j)$$

$$= \sum_{i=1}^6 \sum_{j=1}^6 x_i y_j P(X = x_i, Y = y_j) - E[X]E[Y]$$

$$\begin{aligned} &= 1 \cdot 1 \cdot P(X = 1, Y = 1) + 1 \cdot 2 \cdot P(X = 1, Y = 2) + \cdots + 1 \cdot 6 \cdot P(X = 1, Y = 6) \\ &+ 2 \cdot 1 \cdot P(X = 2, Y = 1) + 2 \cdot 2 \cdot P(X = 2, Y = 2) + \cdots + 2 \cdot 6 \cdot P(X = 2, Y = 6) \\ &+ 3 \cdot 1 \cdot P(X = 3, Y = 1) + 3 \cdot 2 \cdot P(X = 3, Y = 2) + \cdots + 3 \cdot 6 \cdot P(X = 3, Y = 6) \\ &+ 4 \cdot 1 \cdot P(X = 4, Y = 1) + 4 \cdot 2 \cdot P(X = 4, Y = 2) + \cdots + 4 \cdot 6 \cdot P(X = 4, Y = 6) \\ &+ 5 \cdot 1 \cdot P(X = 5, Y = 1) + 5 \cdot 2 \cdot P(X = 5, Y = 2) + \cdots + 5 \cdot 6 \cdot P(X = 5, Y = 6) \\ &+ 6 \cdot 1 \cdot P(X = 6, Y = 1) + 6 \cdot 2 \cdot P(X = 6, Y = 2) + \cdots + 6 \cdot 6 \cdot P(X = 6, Y = 6) - E[X]E[Y] \\ &\quad \left(\frac{1}{36} + \frac{2}{36} + \frac{3}{36} + \frac{4}{36} + \frac{5}{36} + \frac{6}{36} \right) + \left(0 + \frac{6}{36} + \frac{6}{36} + \frac{8}{36} + \frac{10}{36} + \frac{12}{36} \right) \\ &\quad + \left(0 + 0 + \frac{27}{36} + \frac{12}{36} + \frac{15}{36} + \frac{18}{36} \right) + \left(0 + 0 + 0 + \frac{64}{36} + \frac{20}{36} + \frac{24}{36} \right) \\ &\quad + \left(0 + 0 + 0 + 0 + \frac{125}{36} + \frac{30}{36} \right) + \left(0 + 0 + 0 + 0 + 0 + \frac{216}{36} \right) - 3,5 \cdot 4,472 \\ &= \frac{21}{36} + \frac{42}{36} + \frac{72}{36} + \frac{108}{36} + \frac{155}{36} + \frac{216}{36} - 14,31 = 2,74. \end{aligned}$$

□

Ejercicio 45. La proporción en sangre de dos compuestos, X e Y , en una especie común de ratones es variable. Su distribución conjunta en toda la población se caracteriza por la función de densidad

$$f(x, y) = \begin{cases} k(1-x)y^2 & \text{si } 0 < x < 1, \quad 0 < y < 1 \\ 0 & \text{en el resto} \end{cases}$$

Un ratón se considera sano si ambas concentraciones son inferiores a $\frac{3}{4}$.

1. Hallar el valor de k . Decidir si X e Y son independientes.
2. Hallar la concentración media del compuesto Y en la especie.
3. Halla la probabilidad de que un ratón elegido al azar esté sano.

Experimento aleatorio: observación de un ratón al azar, perteneciente a cierta especie de ratones Ω , para obtener la proporción en sangre de dos compuestos, que designamos por X e Y .

El espacio muestral es Ω .

Consideramos dos variables aleatoria, que las vamos a designar también por X e Y .

$$\begin{aligned} X : \Omega &\longmapsto \mathbb{R} \\ \omega &\longrightarrow X(\omega) = \text{proporción en sangre en el} \\ &\quad \text{ratón } \omega \text{ del compuesto } X, \end{aligned}$$

$$\begin{aligned} Y : \Omega &\longmapsto \mathbb{R} \\ \omega &\longrightarrow Y(\omega) = \text{proporción en sangre en el} \\ &\quad \text{ratón } \omega \text{ del compuesto } Y. \end{aligned}$$

Si $f(x, y)$ es la función de densidad del vector aleatorio (X, Y) , se tendrá que verificar

$$1. f(x, y) \geq 0, \quad (x, y) \in \mathbb{R}^2,$$

$$2. \int_{\mathbb{R}^2} f(x, y) dx dy = 1.$$

1 se verifica si imponemos que $K \geq 0$.

$$\begin{aligned} \int_{\mathbb{R}^2} f(x, y) dx dy &= k \int \int_{[0,1] \times [0,1]} (1-x)y^2 dx dy = k \int_0^1 \left(\int_0^1 (1-x)y^2 dy \right) dx \\ &= k \int_0^1 (1-x) \left(\int_0^1 y^2 dy \right) dx = k \int_0^1 (1-x) \left(\frac{y^3}{3} \Big|_0^1 \right) dx \\ &= \frac{k}{3} \int_0^1 (1-x) dx = \frac{k}{3} \left(x - \frac{x^2}{2} \Big|_0^1 \right) = \frac{k}{6}. \end{aligned}$$

Si tomamos $k = 6$, se verifica 1 y 2.

Observación 13. Definimos la función característica del intervalo $(0, 1)$ como la función

$$\chi_{(0,1)}(x) = \begin{cases} 1 & \text{si } x \in (0, 1) \\ 0 & \text{si } x \notin (0, 1). \end{cases}$$

f puede escribirse en la forma

$$f(x, y) = 6(1-x)y^2 \chi_{(0,1)}(x) \chi_{(0,1)}(y). \quad (9)$$

Vamos a calcular las funciones de densidad marginales.

$$f_X(x) = \int_{\mathbb{R}} f(x, y) dy = 6(1-x)\chi_{(0,1)}(x) \int_0^1 y^2 dy = 6(1-x)\chi_{(0,1)}(x). \quad (10)$$

$$f_Y(y) = \int_{\mathbb{R}} f(x, y) dx = 6y^2\chi_{(0,1)}(y) \int_0^1 (1-x) dx = 3y^2\chi_{(0,1)}(y). \quad (11)$$

Después de (9)-(11), se verifica que

$$f(x, y) = f_X(x) \cdot f_Y(y),$$

y por lo tanto las variables aleatorias X e Y son **independientes**.

La concentración media del compuesto Y en la especie, vendrá dada por la media, o esperanza, de la variable aleatoria Y .

$$E[Y] = \int_{\mathbb{R}} y f_Y(y) dy = 3 \int_0^1 y^3 dy = \frac{3}{4}.$$

$$\begin{aligned} P(\text{ratón esté sano}) &= P\left(X < \frac{3}{4} \text{ y } Y < \frac{3}{4}\right) = \int \int_{(-\infty, \frac{3}{4}) \times (-\infty, \frac{3}{4})} f(x, y) dx dy \\ &= 6 \int_0^{\frac{3}{4}} (1-x) \left(\int_0^{\frac{3}{4}} y^2 dy \right) dx = 6 \left(\int_0^{\frac{3}{4}} (1-x) dx \right) \left(\int_0^{\frac{3}{4}} y^2 dy \right) \\ &= 6 \left(x - \frac{x^2}{2} \Big|_0^{\frac{3}{4}} \right) \left(\frac{y^3}{3} \Big|_0^{\frac{3}{4}} \right) = 0,40. \end{aligned}$$

□

Ejercicio 46. Sea (Ω, \mathcal{A}, P) un espacio probabilístico y (X, Y) un vector aleatorio con función de densidad

$$f(x, y) = \begin{cases} 4xye^{-x^2}e^{-y^2} & (x, y) \in (0, \infty) \times (0, \infty) \\ 0 & \text{en el resto.} \end{cases}$$

1. Calcule $P([1, 5] \times [2, \infty))$.
2. Densidades marginales.
3. $E[X]$ y $V(X)$.
4. $E[Y]$ y $V(Y)$.
5. ¿Son independientes las variables aleatorias?

Indicación: Utilícese que $\int_0^\infty e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$.

Solución: 2. $f_X(x) = \begin{cases} 2xe^{-x^2} & x > 0 \\ 0 & \text{en el resto} \end{cases}$, $f_Y(y) = \begin{cases} 2ye^{-y^2} & y > 0 \\ 0 & \text{en el resto} \end{cases}$; 3.

$$E[X] = E[Y] = \frac{\sqrt{\pi}}{2}, V(X) = V(Y) = 1 - \frac{\pi}{4}.$$

Ejercicio 47. Consideremos la variable aleatoria bidimensional (x, Y) con función de densidad conjunta

$$f(x, y) = \begin{cases} \frac{3}{4} & \text{si } x^2 \leq y \leq 1 \\ 0 & \text{en el resto} \end{cases}$$

Calcular las densidades marginales de cada variable. ¿Son independientes las variables? Calcular $P(X \geq Y)$.

Solución: $f_X(x) = \begin{cases} \frac{3}{4}(1 - x^2) & -1 \leq x \leq 1 \\ 0 & \text{en el resto} \end{cases}$, $f_Y(y) = \begin{cases} \frac{3}{2}\sqrt{y} & 0 \leq y \leq 1 \\ 0 & \text{en el resto} \end{cases}$.

No. $\frac{1}{8}$.

Ejercicio 48. Una junta de estudiantes está formada por 10 alumnos: tres de quinto, tres de cuarto, dos de tercero, uno de segundo y uno de primero. De los diez alumnos se seleccionan tres al azar para formar una comisión. Sea X la variable aleatoria que representa el número de alumnos de cuarto en la comisión e Y la variable aleatoria que representa el número de alumnos de quinto en la comisión.

1. Obtener la función de probabilidad conjunta de la variable aleatoria (X, Y) y las marginales.
2. Calcular $\text{Cov}(X, Y)$.
3. Calcular $E[X|Y = 2]$.

Solución: 1.

$Y \backslash X$	0	1	2	3	P_Y
0	$\frac{4}{120}$	$\frac{18}{120}$	$\frac{12}{120}$	$\frac{1}{120}$	$\frac{35}{120}$
1	$\frac{18}{120}$	$\frac{36}{120}$	$\frac{9}{120}$	0	$\frac{63}{120}$
2	$\frac{12}{120}$	$\frac{9}{120}$	0	0	$\frac{21}{120}$
3	$\frac{1}{120}$	0	0	0	$\frac{1}{120}$
P_X	$\frac{35}{120}$	$\frac{63}{120}$	$\frac{21}{120}$	$\frac{1}{120}$	

2. $-\frac{21}{100}$; 3. $\frac{9}{21}$.