CAPÍTULO 1

Variables aleatorias y sus distribuciones

1.1. VARIABLE ALEATORIA UNIDIMENSIONAL

En el curso anterior se estudiaron los conceptos básicos de probabilidad sobre los resultados o sucesos de un experimento aleatorio. Pero los experimentos aleatorios son tales que los resultados a que dan lugar pueden ser de naturaleza cualitativa o cuantitativa. Así por ejemplo, serían resultados cualitativos los derivados de los siguientes experimentos aleatorios:

- El lanzamiento de una moneda: cara o cruz.
- La calidad de las piezas fabricadas en una planta: buenas o defectuosas.
- La preferencia de una persona sobre tres tipos de coches: prefiere el coche A, el coche B o el coche C, etc.

Otros ejemplos de experimentos aleatorios cuyos resultados son cuantitativos serían:

- El número de accidentes de automóvil en una ciudad en un mes dado.
- El número de clientes que llegan a un comercio durante una hora.
- El número de errores detectados en la contabilidad de una empresa.
- La suma de los puntos que aparecen cuando se lanzan simultáneamente dos dados, etc.

Pero trabajar con los resultados cualitativos de un experimento aleatorio introduce ciertas complicaciones, siendo de gran utilidad el cuantificar los resultados cualitativos del experimento aleatorio, o lo que es lo mismo asignar un valor numérico a cada suceso del espacio muestral correspondiente al experimento aleatorio considerado. Esta relación entre los sucesos del espacio muestral y el valor numérico que se les asigna la establecemos mediante la **variable aleatoria**.

Ejemplo 1.1

Supongamos un experimento aleatorio que consiste en lanzar dos monedas al aire, designamos con F la aparición de cara y por C la aparición de cruz. Los posibles sucesos elemetales del espacio muestral serán:

$$E = \{(FC), (CF), (CC), (FF)\}$$

Si definimos la variable aleatoria X como el número de cruces que aparecen al lanzar las dos monedas, podemos establecer la siguiente correspondencia entre los sucesos del espacio muestral y los valores posibles de la variable aleatoria X.

Sucesos		Función		Valores de la variable aleatoria
$A_1 = (FC)$		Variable aleatoria		$X(A_1) = X(FC) = 1$
$A_2 = (CF)$	$\rightarrow$	$X(A_i)$	$\rightarrow$	$X(A_2) = X(CF) = 1$
$A_3 = (CC)$		Número de cruces		$X(A_3) = X(CC) = 2$
$A_4 = (FF)$				$X(A_4) = X(FF) = 0$

Observamos que la variable aleatoria X es una función, pues a cada elemento del conjunto origen, sucesos del espacio muestral, le hace corresponder un solo elemento en el conjunto imagen, de valores de la variable aleatoria. Sin embargo, a cada valor de la variable aleatoria le puede corresponder uno o más sucesos.

Estas funciones cuyos valores dependen de los resultados del experimento aleatorio las llamaremos **variables aleatorias,** y son funciones del espacio muestral E en $\mathbb{R}$, es decir $X\colon E \rightarrow \mathbb{R}$

Definición 1.1. Variable aleatoria.

Una **variable aleatoria** es una función que asigna un valor numérico a cada suceso elemental del espacio muestral.

También podemos decir, aunque de forma menos rigurosa, que una **variable aleatoria** es una variable cuyo valor numérico está determinado por el resultado de un experimento aleatorio. Notaremos con las letras mayúsculas X, Y,… la variable aleatoria y con las letras minúsculas, x, y,… sus valores.

La variable aleatoria puede tomar un número numerable o no numerable de valores posibles; dando lugar a dos tipos principales de variables aleatorias: **discretas** y **continuas**.

Definición 1.2. Variable aleatoria discreta.

> Se dice que una variable aleatoria X es **discreta** si puede tomar un número finito o infinito, pero numerable, de posibles valores.

Así pues, sería una variable aleatoria de tipo discreto el número de piezas defectuosas que aparecen en un proceso de fabricación, el número de llamadas telefónicas que son recibidas en una centralita durante un período de tiempo, o el número de depósitos de una entidad bancaria; y si designamos la correspondiente variable aleatoria por X, entonces X puede tomar cualquiera de los valores $x = 0, 1, 2, 3, \ldots$.

Definición 1.3. Variable aleatoria continua.

> Se dice que una variable aleatoria X es **continua** si puede tomar un número infinito (no numerable) de valores, o bien, si puede tomar un número infinito de valores correspondientes a los puntos de uno o más intervalos de la recta real.

Es decir, la variable aleatoria de tipo continuo puede tomar cualquier valor en uno o más intervalos de la recta real. Así pues, situaciones típicas de variables aleatorias continuas serían las que hacen referencia a medidas tales como tiempo, peso o longitud. Por ejemplo, si representamos por X el tiempo que dedica un alumno a hacer un examen cuya duración máxima es de dos horas (120 minutos); tendríamos una variable aleatoria de tipo continuo que puede tomar infinitos valores en el intervalo $0 \leqslant x \leqslant 120$.

Otro ejemplo de variable aleatoria continua sería la cantidad de lluvia caída cada día en un determinado punto geográfico, pues un equipo de medida de mucha precisión nos daría cada día la cantidad de lluvia registrada que podríamos hacer corresponder con un punto único del intervalo delimitado por el valor mínimo y el máximo de lluvia prevista. Cada uno de los infinitos puntos del intervalo podría ser un valor de la variable aleatoria X continua correspondiente a la cantidad de lluvia caída en ese punto geográfico durante un día.

1.1.1. Distribución de probabilidad de variables aleatorias discretas

Sea X una variable aleatoria discreta que toma un número finito de valores, r en total, que indicaremos por $x_1, x_2, \ldots, x_i, \ldots x_r$.

La probabilidad de que la variable aleatoria X tome un valor particular x_i la designaremos por:

$$p_i = P(x_i) = P(X = x_i)$$

Definición 1.4. Distribución de probabilidad.

> La **distribución de probabilidad, función de probabilidad** o **función de cuantía** de una variable aleatoria discreta X, la notaremos por $P(x)$, y es una función que asigna las probabilidades con que la variable aleatoria toma los posibles valores, de tal manera que las probabilidades verifiquen las dos condiciones siguientes:
>
> I. $0 \leqslant P(x_i) \leqslant 1, \quad i = 1, 2, ..., r$
>
> II. $\sum_{i=1}^{r} P(x_i) = 1.$

La condición I indica que la probabilidad no puede ser negativa, y de la condición II se deduce que los sucesos, $X = x$, para todos los valores posibles x, son mutuamente excluyentes y exhaustivos.

Ejemplo 1.2

Un lote de 100 lámparas contiene 10 lámparas defectuosas. El minorista decide tomar dos lámparas aleatoriamente, y si ninguna de las dos es defectuosa, entonces acepta el lote. Definimos la variable aleatoria X como el número de lámparas defectuosas en una selección aleatoria de dos lámparas. Obtener la distribución de probabilidades de la variable aleatoria X.

Solución:

El experimento aleatorio de selección de dos lámparas al azar define los sucesos elementales:

L_1: Primera lámpara defectuosa.

L_2: Segunda lámpara defectuosa.

$\overline{L}_1$: Primera lámpara no defectuosa.

$\overline{L}_2$: Segunda lámpara no defectuosa.

Al considerar la selección conjunta de las dos lámparas, el espacio muestral generado por este experimento aleatorio contiene cuatro sucesos mutuamente excluyentes y forman un sistema completo de sucesos.

En la tabla 1.1 aparecen los sucesos elementales del espacio muestral, las correspondientes probabilidades y los valores que puede tomar la variable aleatoria *X*.

Tabla 1.1. *Sucesos elementales, probabilidades de los mismos y valor de la variable aleatoria.*

Sucesos	Porbabilidades	Valor de la varable aleagoria X
$L_1 \cap L_2$	0,0091	2
$L_1 \cap \overline{L}_2$	0,0909	1
$\overline{L}_1 \cap L_2$	0,0909	1
$\overline{L}_1 \cap \overline{L}_2$	0,8091	0

Las probabilidades para cada suceso según la regla de multiplicación de probabilidades, serían:

$$P(L_1 \cap L_2) = P(L_1) \cdot P(L_2/L_1) = \frac{10}{100} \cdot \frac{9}{99} = 0{,}0091$$

$$P(L_1 \cap \overline{L}_2) = P(L_1) \cdot P(\overline{L}_2/L_1) = \frac{10}{100} \cdot \frac{90}{99} = 0{,}0909$$

$$P(\overline{L}_1 \cap L_2) = P(\overline{L}_1) \cdot P(L_2/\overline{L}_1) = \frac{90}{100} \cdot \frac{10}{99} = 0{,}0909$$

$$P(\overline{L}_1 \cap L_2) = P(\overline{L}_1) \cdot P(\overline{L}_2/\overline{L}_1) = \frac{90}{100} \cdot \frac{89}{99} = 0{,}8091$$

Y consecuentemente las probabilidades con que la variable aleatoria *X* toma sus diferentes valores serían:

$$P_1 = P(X = 0) = P(0) = 0{,}8091$$
$$P_2 = P(X = 1) = P(1) = 0{,}0909 + 0{,}0909 = 0{,}1818$$
$$P_3 = P(X = 2) = P(2) = 0{,}0091$$

La probabilidad de que el minorista acepte el lote de las 100 lámparas será:

$$P(X = 0) = P(0) = 0{,}8091$$

La distribución de probabilidad $P(x)$, de la variable aleatoria X aparece en la tabla 1.2, en donde aparecen los diferentes valores x de la variable aleatoria y sus probabilidades $P(x)$.

Tabla 1.2. *Distribución de probabilidad de la variable aleatoria X.*

Valores de la variable aleatoria x	Probabilidades $P(x)$
0	0,8091
1	0,1818
2	0,0091

Ya que los cuatro sucesos del experimento forman un conjunto mutuamente excluyente y exhaustivo de sucesos, las probabilidades correspondientes a estos sucesos suman la unidad.

La función de probabilidad de una variable aleatoria discreta se puede expresar de tres maneras:

- Tabular.
- Algebraica o funcional.
- Gráfica.

La forma **tabular** consiste en expresar en forma de tabla una lista de los valores posibles de la variable aleatoria X y las correspondientes probabilidades $P(x)$, como aparece en la tabla 1.2.

La forma **algebraica o funcional** se presentan cuando es posible escribir una función en la forma de una ecuación, de tal manera que para cada valor de la variable aleatoria se pueda obtener su probabilidad. Así pues, para el ejemplo 1.2 la forma funcional de la distribución de probabilidad sería[1]:

$$P(x) = P(X = x) = \frac{\binom{10}{x}\binom{100-10}{2-x}}{\binom{100}{2}}; \quad x = 0,\ 1,\ 2$$

[1] Esta función de probabilidad corresponde a la de una variable aleatoria hipergeométrica. No es necesario que intente obtener esta forma funcional pues solo se pretende que sepamos que se puede utilizar para expresar la distribución de probabilidad de una variable aleatoria discreta.

Otra forma de presentar la función de probabilidad sería mediante un **gráfico** o **diagrama de barras**, en el que sobre el eje de abscisas llevaríamos los diferentes valores x, de la variable aleatoria X y sobre el eje de ordenadas, las probabilidades $P(x)$. Así pues, para el ejemplo 1.2 tendríamos el gráfico 1.1.

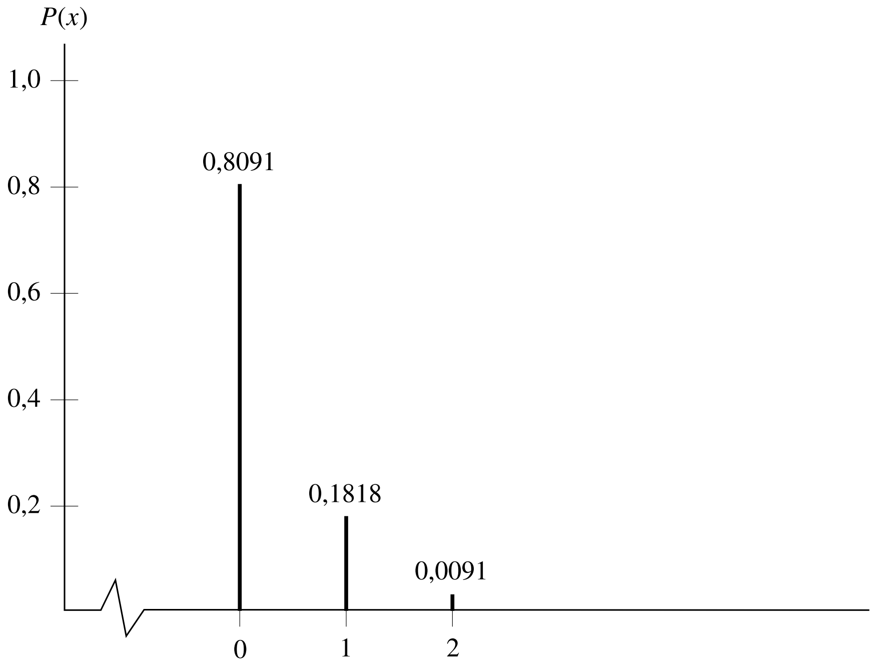


Gráfico 1.1. Función de probabilidad de la variable aleatoria X asociada al experimento aleatorio del ejemplo 1.2.

Una consideración importante sobre variables aleatorias discretas y sus distribuciones es la siguiente: si X es una variable aleatoria discreta y deseamos determinar la probabilidad de que X sea mayor o igual que un cierto valor a y menor o igual que un valor b, siendo $a \leqslant b$, solamente tenemos que sumar las probabilidades correspondientes a valores de X comprendidos entre a y b; es decir,

$$P(0 \leqslant X < 2) = P(0 \leqslant X \leqslant 1) = \sum_{x=0}^{x=1} P(X = x)$$

Así pues, para el ejemplo 1.2 la

$$P(0 \leqslant X < 2) = P(0 \leqslant X \leqslant 1) = \sum_{x=0}^{x=1} P(X = x)$$

$$= P(X = 0) + P(X = 1)$$

$$= 0{,}8091 + 0{,}1818 = 0{,}9909$$

Definición 1.5. Función de distribución.

Sea una variable aleatoria X de tipo discreto que toma un número finito o infinito numerable de valores $x_1, x_2, \ldots, x_r$ y cuya distribución de probabilidad es $P(x)$. Se define la **función de distribución** acumulativade la variable aleatoria X, que notaremos por $F(x)$, como la probabilidad de que la variable aleatoria X tome valores menores o iguales que x, es decir:

$$F(x) = P(X \leqslant x) = \sum_{x_i \leqslant x} P(x_i) = \sum_{x_i = x_1}^{x} P(X = x_i)$$

y representa la suma de las probabilidades puntuales hasta el valor x inclusive de la variable aleatoria X.

Para el ejemplo 1.2 la función de distribución $F(x)$ viene dada en la tabla 1.3.

Tabla 1.3. *Función de distribución de la variable aleatoria del ejemplo 1.2.*

Valores de X x	Función de probabilidad $P(x)$	Función de distribución $F(x) = P(X \leqslant x)$
0	0,8091	$F(0) = P(X \leqslant 0) = 0{,}8$
1	0,1818	$F(1) = P(X \leqslant 1) = 0{,}9$
2	0,0091	$F(2) = P(X \leqslant 2) = 1{,}0$

Así pues la función de distribución para cualquier valor de x quedaría de la siguiente forma:

$$F(x) = \begin{cases} 0 & , \quad x < 0 \\ 0{,}8091 & , \quad 0 \leqslant x < 1 \\ 0{,}9909 & , \quad 1 \leqslant x < 2 \\ 1 & , \quad x \geqslant 2 \end{cases}$$

Se deduce, de la propia definición, que la función de distribución de una variable aleatoria discreta es una función que verifica:

I. $0 \leqslant F(x) \leqslant 1, \quad \forall\, x$

II. Es no decreciente, es decir: si $x_i < x_j$, entonces $F(x_i) \leqslant F(x_j)$

La representación gráfica de la función de distribución viene dada por el gráfico 1.2.

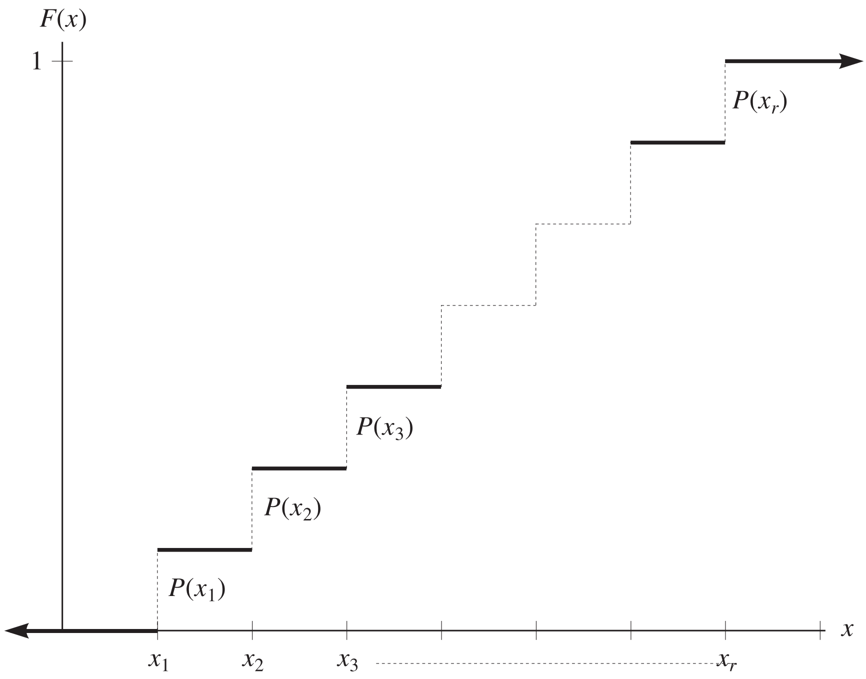


Gráfico 1.2. Función de distribución de una variable aleatoria discreta.

Vemos pues que una variable aleatoria discreta X, está caracterizada por su función de probabilidad o distribución de probabilidad $P(x)$ y también por su función de distribución $F(x)$.

1.1.2. Distribución de probabilidad de variables aleatorias continuas

En el apartado anterior estudiábamos las distribuciones de probabilidad para variables aleatorias de tipo discreto, las cuales podían tomar un número finito o infinito numerable de valores. Pero en muchas aplicaciones estadísticas, la variable aleatoria puede tomar cualquier valor sobre un intervalo de la recta real, es decir, puede tomar un número infinito de valores, siendo la correspondiente variable aleatoria de tipo continuo.

La variable aleatoria de tipo continuo se tratará de forma diferente a como hacíamos con la variable aleatoria discreta, pues en el caso discreto se asignaban probabilidades positivas a todos los valores puntuales de la variable aleatoria, y la suma de todas ellas era igual a la unidad. En el caso continuo no es posible asignar una probabilidad a cada uno de los infinitos posibles valores de la variable aleatoria y que estas probabilidades sumen la unidad como en el caso discreto, siendo necesario utilizar una aproximación diferente para llegar a obtener la distribución de probabilidad de una variable aleatoria continua. Así pues la aproximación que adoptaremos utiliza el histograma de frecuencia relativa, como veremos, con detalle, a continuación.

Supongamos una variable aleatoria X que nos mide los tiempos que transcurren entre dos llegadas consecutivas, de 100 vehículos a una gasolinera. Los tiempos entre dos llegadas consecutivas son medidos con un cronómetro de alta precisión que es capaz de medir el tiempo hasta una milésima de segundo, de tal manera que si un vehículo llega en el instante 54,325 seg. y el próximo en el instante 54,335 seg. el punto medio del intervalo sería 54,330 seg. pero la probabilidad correspondiente se le asigna al intervalo y no al valor medio 54,330 seg. pues en el caso continuo la probabilidad de que una variable aleatoria tome un valor concreto es cero.

Supongamos que la distribución de frecuencia relativas de los tiempos entre dos llegadas consecutivas de 100 vehículos a una gasolinera viene dada por la tabla 1.4, en donde consideramos 10 intervalos de longitud un minuto.

Tabla 1.4. *Distribución relativa de los tiempos entre dos llegadas consecutivas, de 100 vehículos a una gasolinera.*

Intervalo de tiempo	N.º de vehículos que llegan	Frecuencia relativa
$0 < x \leqslant 1$	3	0,03
$1 < x \leqslant 1$	4	0,04
$2 < x \leqslant 1$	6	0,06
$3 < x \leqslant 1$	18	0,18
$4 < x \leqslant 1$	23	0,23
$5 < x \leqslant 1$	17	0,17
$6 < x \leqslant 1$	16	0,16
$7 < x \leqslant 1$	8	0,08
$8 < x \leqslant 1$	3	0,03
$9 < x \leqslant 1$	2	0,02

Haciendo la representación gráfica de la distribución de frecuencias relativas correspondiente a los tiempos de llegadas consecutivas de los 100 vehículos tenemos el histograma de frecuencias relativas dado en el gráfico 1.3.

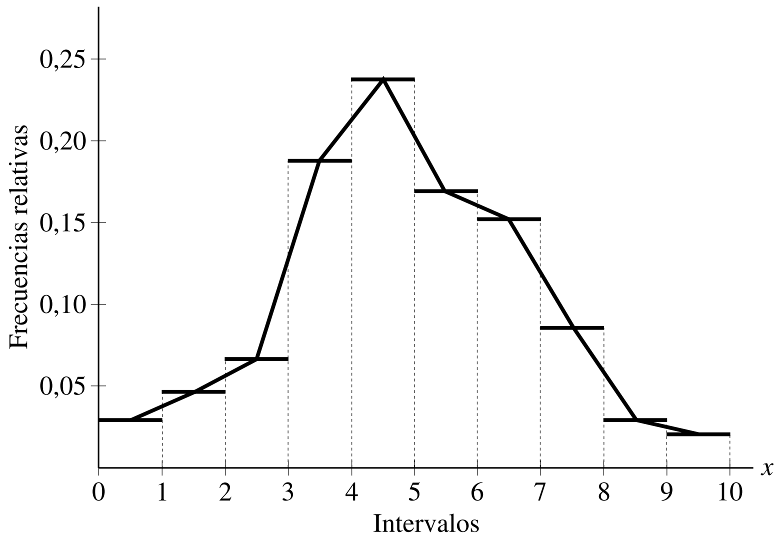


Gráfico 1.3. Histograma de frecuencias relativas correspondiente a los tiempos de llegadas consecutivas de 100 vehículos a una gasolinera.

Hemos de tener en cuenta que la frecuencia se refiere a cada intervalo y no al punto medio del intervalo, y además como todos los intervalos tienen la misma longitud, una unidad, el área del rectángulo construida sobre cada intervalo es igual a la frecuencia relativa del correspondiente intervalo y consecuentemente la suma de todas las áreas de todos los rectángulos es igual a la unidad.

Supongamos que en lugar de observar los tiempos entre dos llegadas consecutivas para 100 vehículos, lo haremos para 500 vehículos y consideramos 20 intervalos de medio minuto cada uno, entonces obtendríamos un gráfico parecido al gráfico 1.3, pero con 20 intervalos y consecuentemente 20 rectángulos de base (o amplitud) la mitad que los anteriores y continuando con este proceso de aumentar el número de observaciones y el número de intervalos aunque de menos amplitud se llegara a obtener un histograma de frecuencias relativas con muchos intervalos de amplitud muy pequeña, que en el límite dará lugar a una representación del tipo de la del gráfico 1.4.

La curva límite $f(x)$ del histograma de frecuencias relativas se obtiene para un número grande de observaciones, en nuestro ejemplo, vehículos que llegan a la gasolinera, y para unas amplitudes de intervalos de tiempo muy pequeñas.

La función $f(x)$, cuya representación gráfica es la curva límite correspondiente al histograma de frecuencias relativas es la **función de densidad de probabilidad** o simplemente la **función de densidad** de la variable aleatoria continua X. El área bajo la función $f(x)$ es igual a la unidad, de la misma manera que el área en el correspondiente histograma de frecuencias relativas era igual a la unidad.

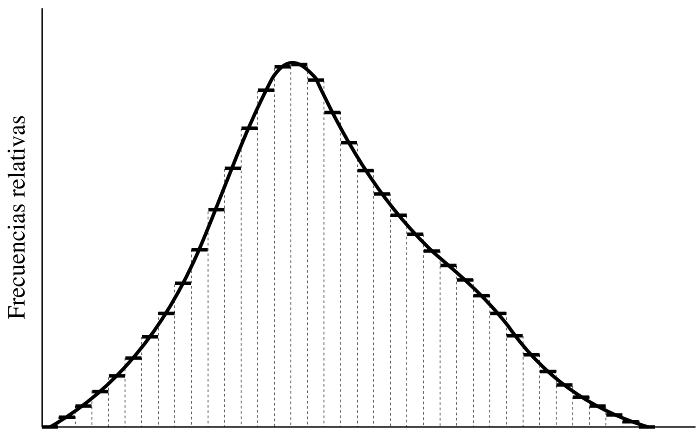


Gráfico 1.4. Curva límite del histograma de frecuencias relativas correspondiente a los tiempos de llegadas consecutivas de un gran número de vehículos a una gasolinera.

La representación gráfica de la función de densidad viene dada por el gráfico 1.5.

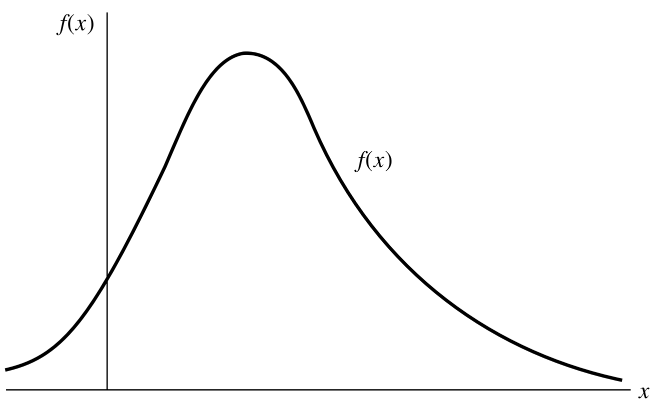


Gráfico 1.5. Función de densidad de la variable aleatoria X continua.

Definición 1.6. Función de densidad.

Sea X una variable aleatoria de tipo continuo, y si existe una función $f(x)$ tal que verifica:

$$\text{I.} \quad f(x) \geqslant 0, \quad -\infty < x < +\infty.$$

$$\text{II.} \quad \int_{-\infty}^{+\infty} f(x)\, dx = 1.$$

diremos que $f(x)$ es la **función de densidad** de la variable aleatoria continua X.

Observemos la analogía entre las condiciones que tenía que verificar $P(x)$ para ser una función de probabilidad de una variable aleatoria discreta, dadas en la definición 1.4 y las condiciones que debe verificar $f(x)$ para ser una función de densidad de una variable aleatoria continua, dadas en la definición 1.6. En el caso discreto, la probabilidad o masa de probabilidad se asigna a cada valor de la variable aleatoria, mientras que en el caso continuo, la probabilidad es considerada como una densidad de probabilidad sobre el rango de valores de la variable aleatoria, es decir en el caso discreto existen probabilidades puntuales, sin embargo, en el caso continuo la $P(X = x) = 0$, y la probabilidad no se asigna a un punto sino a un intervalo. En el caso discreto, la suma de las probabilidades puntuales es igual a la unidad, mientras que en el caso continuo la suma de las densidades de probabilidad o el área bajo la curva $f(x)$ debe ser igual a la unidad, y consecuentemente la probabilidad de que la variable aleatoria continua X tome valores en el intervalo $[a, b]$ será igual al área bajo la curva $f(x)$ acotada entre a y b, es decir:

$$P(a \leqslant X \leqslant b) = \int_a^b f(x)\, dx \qquad [1.1]$$

o bien gráficamente como se indica en el gráfico 1.6.

Hay una diferencia significativa en el cálculo de las probabilidades para una variable aleatoria discreta y una variable aleatoria continua. Si la variable aleatoria X es discreta, sabemos que existen las probabilidades puntuales $P(X = x_i)$, sin embargo, en el caso continuo, si x_i es un punto interior al intervalo de definición de la variable aleatoria continua X, entonces la $P(X = x) = 0$, en efecto:

$$P(X = x_i) = P(x_i \leqslant X \leqslant x_i) = 0$$

ya que utilizando la expresión [1.1] vemos que el área entre x_i y x_i es nula.

Es interesante observar que, puesto que en el caso continuo la $P(X = x) = 0$, entonces se verifica:

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

ya que

$$P(X = a) = 0 \quad \text{y} \quad P(X = b) = 0$$

lo cual no sucede en el caso discreto.

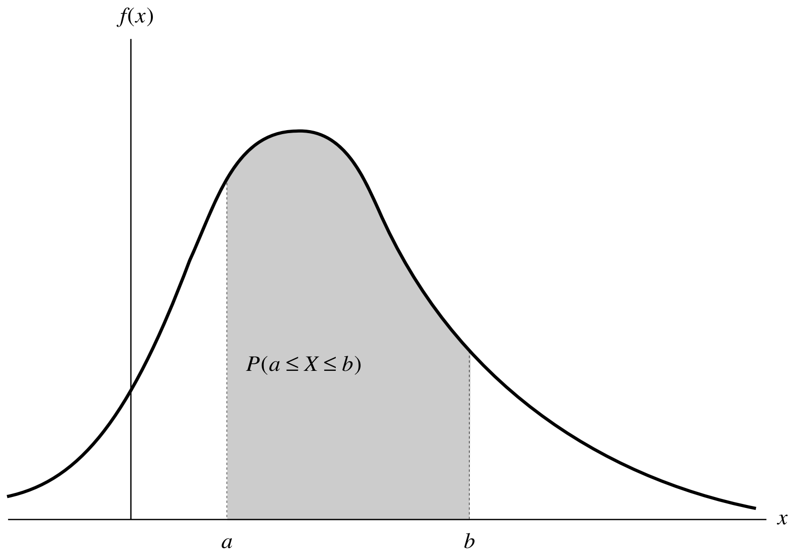


Gráfico 1.6. Área bajo la curva f(x) correspondiente a la $P(a \leqslant X \leqslant b)$.

Definición 1.7. Función de distribución.

Sea una variable aleatoria X de tipo continuo que toma un número infinito de valores sobre la recta real y cuya función de densidad es $f(x)$. Se define la **función de distribución** acumulativa de la variable aleatoria X, que notaremos por $F(x)$, como la probabilidad de que la variable aleatoria continua X tome valores menores o iguales a x, es decir:

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

y representa el área limitada por la curva función de densidad, $f(x)$, y a la izquierda de la recta $X = x$, como se indica en el gráfico 1.7.

Para poder hacer un razonamiento intuitivo sobre esta definición de la función de distribución, consideramos una masa unidad, distribuida sobre todo el intervalo de definición de la variable aleatoria, que podría ser el intervalo $(-\infty, +\infty)$, y la función de distribución $F(x)$ para cada valor x de la variable aleatoria nos da la cantidad de masa que hay en el intervalo $(-\infty, x]$, es decir la masa que hay en el punto x y a la izquierda de x, pero como en el caso continuo sabemos que no existen masas puntuales, nos dará lo mismo considerar el intervalo abierto por la derecha $(-\infty, x)$, de tal manera que:

$$F(x) = P(X \leqslant x) = P(X < x)$$

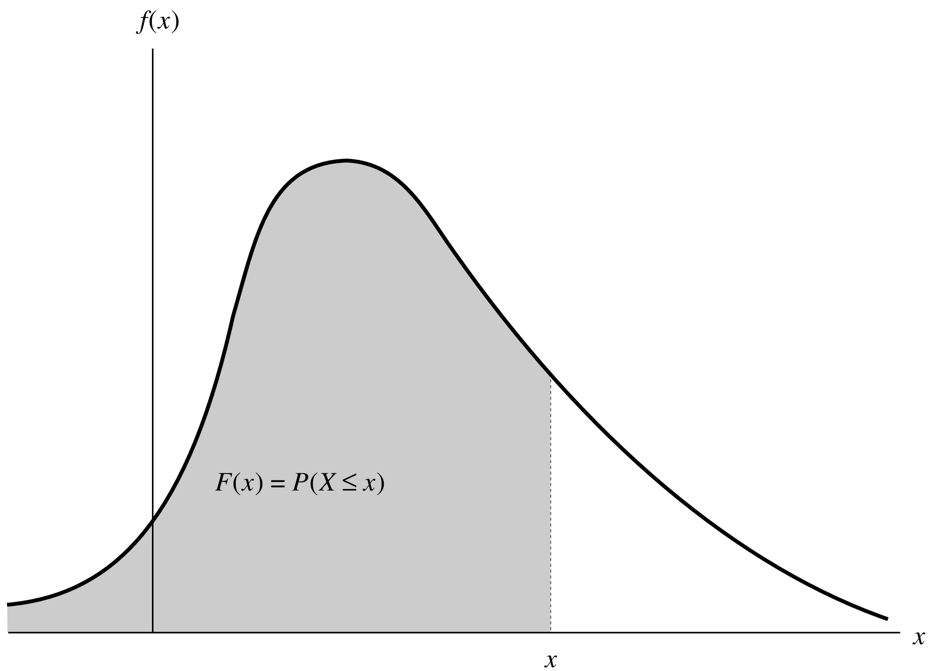


Gráfico 1.7. Función de distribución de la variable aleatoria continua X.

pues la

$$P(X = x) = 0$$

y también,

$$1 - F(x) = P(X \geqslant x) = P(X > x)$$

La función de distribución de una variable aleatoria continua es una función que verifica:

1. $F(-\infty) = 0$.
2. $F(+\infty) = 1$.
3. Es no decreciente, es decir: si $x_i < x_j$ entonces $F(x_i) \leqslant F(x_j)$.
4. $P(a \leqslant X \leqslant b) = \int_a^b f(x)\, dx = F(b) - F(a)$.
5. La derivada de la función de distribución es la función de densidad:

$$\frac{dF(x)}{dx} = f(x)$$

Como la $\int_a^b f(x)\, dx$ representa gráficamente el área encerrada bajo la curva $f(x)$ y los valores a y b del eje de abscisas, entonces a la probabilidad:

$$P(a \leqslant X \leqslant b) = F(b) - F(a)$$

se le puede dar la misma interpretación, y podemos hacer la siguiente representación, gráfico 1.8.

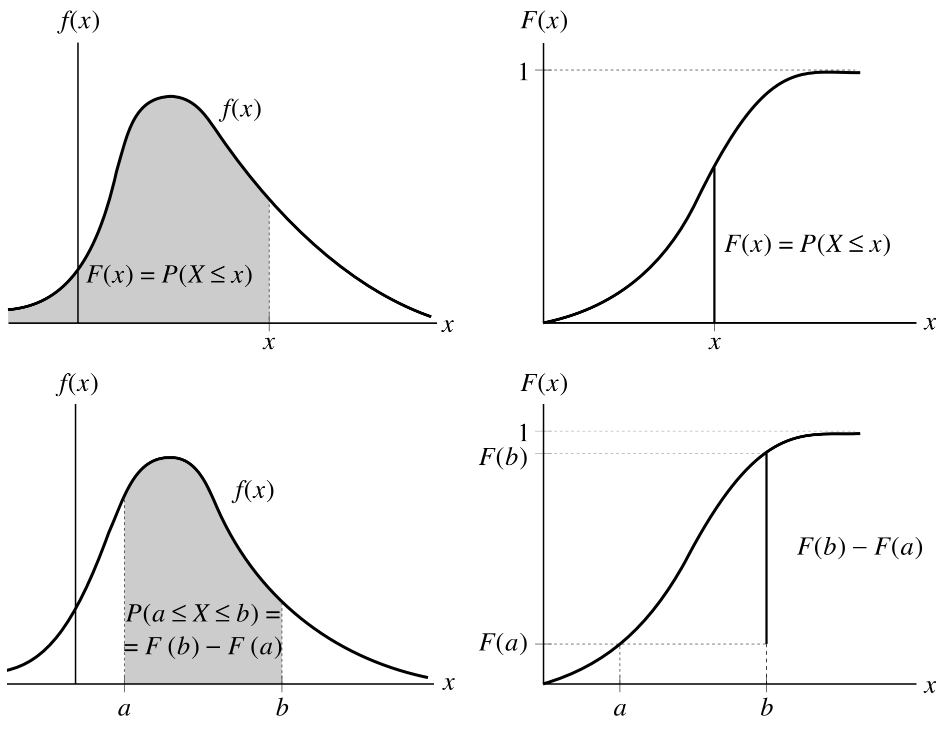


Gráfico 1.8. Representación gráfica de la función de distribución.

Ejemplo 1.3

Sea una variable aleatoria continua X cuya función de densidad $f(x)$ es:

$$f(x) = \begin{cases} \dfrac{1}{2}, & 15 \leqslant x \leqslant 17 \\ 0, & \text{en el resto} \end{cases}$$

Se pide:

1.º Probar que efectivamente es una función de densidad.

2.º La representación gráfica de la función de densidad.

3.º Obtener la $P(15 \leqslant X \leqslant 16)$ y $P(16{,}25 \leqslant X \leqslant 16{,}75)$.

4.º Obtener la $P(X \leqslant 16)$, $P(X \leqslant 17)$ y la $P(X \leqslant k)$ para $15 \leqslant k \leqslant 17$.

5.º Obtener la función de distribución y el valor que toma en el punto $x = 16$.

Solución:

1.º Para ver que efectivamente es una función de densidad se tienen que verificar las dos condiciones siguientes:

$$f(x) \geqslant 0, \quad -\infty < x < +\infty$$

y

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

en nuestro caso será:

$$f(x) \geqslant 0, \quad -\infty \leqslant x \leqslant +\infty,$$

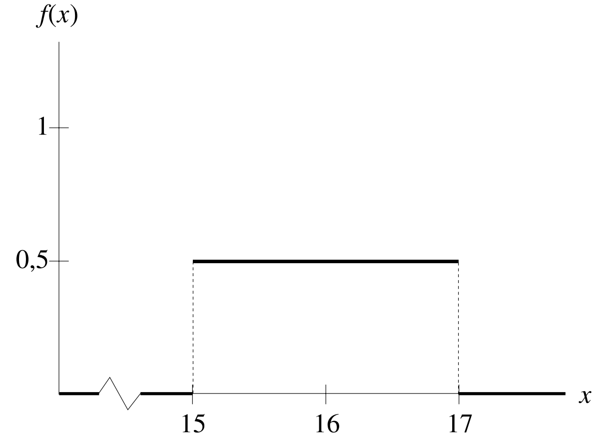
y

$$\int_{15}^{17} \frac{1}{2}\, dx = 1$$

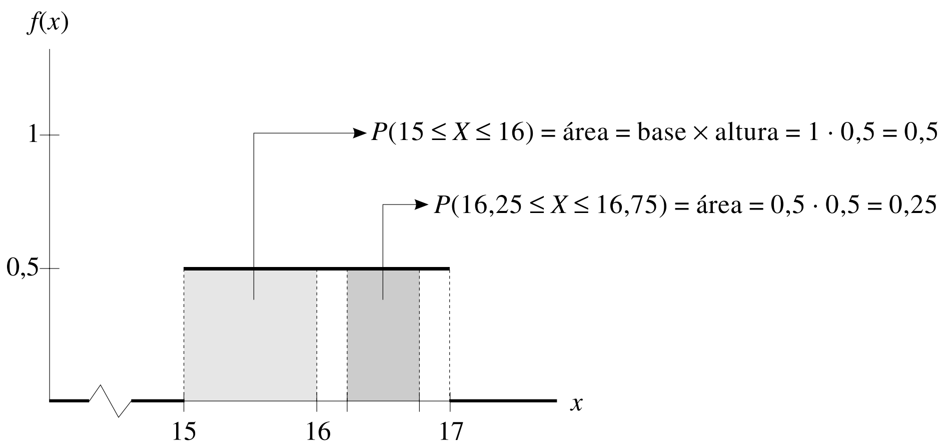
en efecto:

$$\int_{15}^{17} \frac{1}{2}\, dx = \frac{1}{2}\,[x]_{15}^{17} = \frac{1}{2}\,(17 - 15) = \frac{1}{2} \cdot 2 = 1$$

2.º La representación gráfica de la función de densidad será:



3.º Representando gráficamente ambas probabilidades tenemos:



4.º $P(X \leqslant 16) = P(-\infty < X \leqslant 16) = P(15 \leqslant X \leqslant 16) =$

$$= \int_{15}^{16} \frac{1}{2}\, dx = \frac{1}{2}\, [x]_{15}^{16} = \frac{1}{2}\, (16 - 15) = 0{,}5$$

$P(X \leqslant 17) = P(-\infty < X \leqslant 17) = P(15 \leqslant X \leqslant 17) =$

$$= \int_{15}^{17} \frac{1}{2}\, dx = \frac{1}{2}\, [x]_{15}^{17} = \frac{1}{2}\, (17 - 15) = 1$$

$P(X \leqslant k) = P(-\infty < X \leqslant k) = P(15 \leqslant X \leqslant k) =$

$$= \int_{15}^{k} \frac{1}{2}\, dx = \frac{1}{2}\,[x]_{15}^{k} = \frac{1}{2}\,(k - 15) \quad , \quad \text{si } 15 \leqslant k \leqslant 17$$

5.º La función de distribución $F(x)$ será:

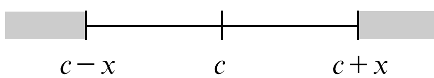
$$F(x) = P(X \leqslant x) = \begin{cases} 0, & x < 15 \\ \dfrac{x - 15}{2}, & 15 \leqslant x \leqslant 17 \\ 1, & x > 17 \end{cases}$$

$$F(16) = P(X \leqslant 16) = \frac{16 - 15}{2} = \frac{1}{2} = 0{,}5$$

1.1.3. Distribución simétrica

Diremos que una variable aleatoria X es **simétrica** respecto de un punto c, cuando se verifica:

$$P(X \leqslant c - x) = P(X \geqslant c + x) \quad \forall\ x$$



O bien, si expresamos esta definición por medio de la función de distribución, entonces diremos que la variable aleatoria X tiene una distribución simétrica respecto del punto c, si se verifica:

$$F(c - x) = 1 - F(c + x) + P(X = c + x), \quad \forall\ x$$

lo cual nos permite decir que la función de distribución $F(x)$ es simétrica respecto del centro de simetría c.

Si la variable aleatoria X es de tipo continuo, entonces tendremos:

$$F(c - x) = 1 - F(c + x)$$

ya que

$$P(X = c + x) = 0$$

y en consecuencia, la función de densidad será:

$$f(c - x) = f(c + x)$$

Si $c = 0$, diremos que la variable aleatoria X es simétrica respecto del punto cero si se verifica:

$$f(-x) = f(x)$$

o lo que es lo mismo, diremos que $F(x)$ es simétrica.

1.2. VARIABLE ALEATORIA BIDIMENSIONAL

Como indicamos al iniciar el concepto de variable aleatoria unidimensional, veíamos que introducíamos las variables aleatorias con el fin de transformar los resultados elementales de un experimento aleatorio en números reales, y de esta forma aprovechamos las propiedades del nuevo espacio probabilístico para estudiar los sucesos relacionados con cantidades numéricas.

En algunas situaciones la información que se tiene del experimento aleatorio no se puede expresar mediante valores de un subconjunto de $\mathbb{R}$. Por ejemplo, si estamos interesados en estudiar la relación existente entre el peso y la estatura de los individuos de una población, el correspondiente suceso que se tendría en cada realización del experimento aleatorio sería: **el individuo extraído al azar pesa 85 kg y mide 1,85 m** y a este suceso no se le puede asignar un número en $\mathbb{R}$, sino que en este caso tendríamos que considerar dos variables aleatorias X e Y, una para el peso y otra para la estatura, las cuales darían lugar al espacio $\mathbb{R}^2$, sobre el que se tendría el correspondiente espacio probabilístico.

Esto pone de manifiesto que en ciertas situaciones hay que trabajar en espacios de más de una dimensión (en el ejemplo anterior en $\mathbb{R}^2$, pero en general sería en $\mathbb{R}^n$), estableciendo aplicaciones que transforman los resultados o sucesos elementales del experimento aleatorio en puntos del espacio n-dimensional $\mathbb{R}^n$, estas aplicaciones se hacen utilizando variables aleatorias bi-dimensionales o n-dimensionales en general. Nosotros aquí nos vamos a referir siempre al caso bidimensional.

En muchas ocasiones es necesario estudiar conjuntamente dos características del fenómeno aleatorio en cuestión, es decir, el comportamiento conjunto de dos variables aleatorias, intentando explicar la posible relación existente entre ellas. Así por ejemplo, nos puede interesar estudiar la renta y el consumo de las familias, los gastos de mantenimiento y el valor de la producción de una empresa, los salarios de las familias y el ahorro, etc. Pero para estudiar conjuntamente las dos variables aleatorias, es decir, la variable aleatoria bidimensional, será necesario conocer la distribución de probabilidad conjunta de ambas variables.

Para introducir las ideas básicas sobre la distribución conjunta de probabilidad consideramos el siguiente ejemplo:

Supongamos una caja que contiene cuatro bombillas buenas y una defectuosa. Realizamos extracciones sin reemplazamiento, es decir, se extrae aleatoriamente una bombilla de la caja se observa si es buena o defectuosa pero no se devuelve a la caja hasta después de que se ha extraído la segunda bombilla. Designamos por X la variable aleatoria, de tipo discreto, que indica si la primera bombilla extraída es defectuosa o no y por Y la variable aleatoria que indica si la segunda bombilla extraída es defectuosa o no. Ambas variables aleatorias X e Y sólo pueden tomar los valores cero o uno según que la bombilla extraída sea buena o defectuosa.

Este experimento aleatorio da lugar a 20 posibles resultados o sucesos, pues en la primera extracción son posibles 5 elecciones (las 5 bombillas que hay), y para la segunda extracción solo quedan 4 posibles elecciones, luego hay $5 \cdot 4 = 20$ **posibles resultados**. Pero como las extracciones se realizan de forma aleatoria el modelo probabilístico asociado a este experimento aleatorio asigna probabilidades iguales a 1/20 a cada uno de los veinte posibles sucesos. De estos veinte sucesos, doce corresponden a los valores de la variable aleatoria bidimensional ($X = 0$, $Y = 0$), pues en la caja hay cuatro bombillas buenas, una de las cuales se puede obtener en la primera extracción, y quedan tres bombillas buenas para la segunda extracción, luego hay $4 \cdot 3 = 12$ posibles resultados o sucesos en los que no se extrae bombilla defectuosa ni en la primera ni en la segunda extracción. Resultando que la probabilidad del suceso considerado será:

$$P(X = 0,\ Y = 0) = \frac{12}{20} = 0{,}6$$

Análogamente obtendríamos:

$$P(X = 1,\ Y = 0) = \frac{4}{20} = 0{,}2$$

$$P(X = 0,\ Y = 1) = \frac{4}{20} = 0{,}2$$

$$P(X = 1,\ Y = 1) = \frac{0}{20} = 0$$

Estas probabilidades se pueden resumir en la tabla 1.5, y nos indica la distribución de probabilidad conjunta de la variable aleatoria bidimensional (X, Y).

Tabla 1.5. *Distribución de probabilidad conjunta de la variable aleatoria bidimensional (X, Y).*

Y \ X	0	1	Probabilidad marginal de *Y*
0	0,6	0,2	0,8
1	0,2	0	0,2
Probabilidad marginal de *X*	0,8	0,2	1

En el gráfico 1.9 se muestra la representación gráfica de la distribución de probabilidad bidimensional del experimento aleatorio considerado.

La última fila y la última columna de la tabla 1.5 nos indican las probabilidades marginales de X yde Y respectivamente.

Además de las probabilidades conjuntas y marginales, nos pueden interesar las probabilidades condicionadas, tales como la probabilidad de que la segunda bombilla sacada sea buena cuando la primera también fue buena. Para obtener esta probabilidad condicionada en este experimento, basta con tener en cuenta que

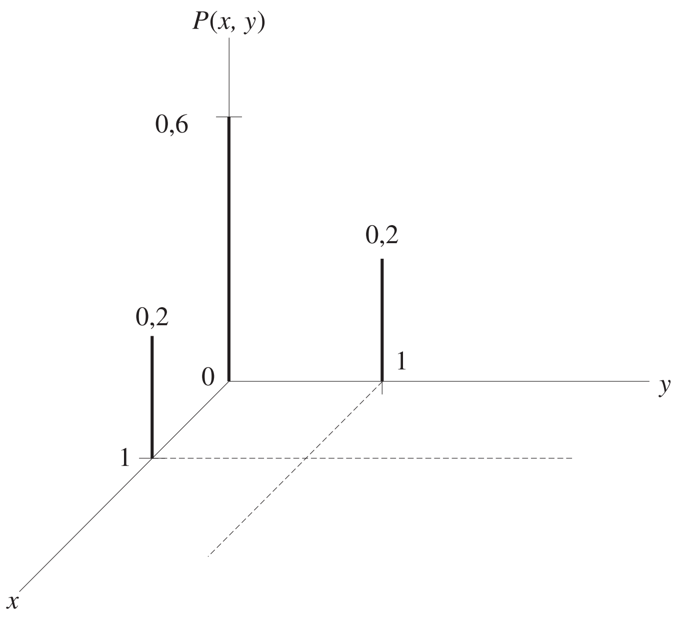


Gráfico 1.9. Diagrama de barras, tridimensional, de probabilidades relativas para los pares de observaciones (x, y).

después de que se sabe que la primera bombilla extraída es buena, hay solamente tres bombillas buenas entre las cuatro restantes, y tendríamos:

$$P(Y = 0/X = 0) = \frac{3}{4}$$

que coincide con la probabilidad obtenida a partir de la definición de probabilidad condicionada que daremos después de esta introducción intuitiva y que será:

$$P(Y = 0/X = 0) = \frac{P(X = 0,\ Y = 0)}{P(X = 0)} = \frac{0{,}6}{0{,}8} = \frac{3}{4}$$

1.2.1. Distribución de probabilidad bidimensional

De manera análoga a lo que hacíamos en el caso unidimensional, aquí daremos para la **variable aleatoria bidimensional discreta,** su distribución de probabilidades y su función de distribución y para la variable aleatoria bidimensional de tipo continuo daremos su función de densidad de probabilidad y la correspondiente función de distribución.

Sea la variable aleatoria X de tipo discreto que toma un número finito de valores $x_1, x_2, \ldots, x_r$, y sea la variable aleatoria Y, también de tipo discreto, que toma los valores $y_1, y_2, \ldots, y_r$.

La probabilidad de que la variable aleatoria X tome el valor x, y la variable aleatoria Y tome el valor y_j, la designaremos por:

$$p_{ij} = P(x_i, y_j) = P(X = x_i, Y = y_j)$$

Definición 1.8. Distribución de probabilidad bidimensional.

La **distribución de probabilidad bidimensional** o **distribución de probabilidad conjunta** de una variable aleatoria discreta bidimensional es una función $P(x, y)$ que asigna las probabilidades a los diferentes valores conjuntos de la variable aleatoria bidimensional (X, Y), de tal manera que se verifiquen las dos condiciones siguientes:

I. $0 \leqslant P(x_i, y_j) \leqslant 1, \quad i = 1, 2, ..., r; \quad j = 1, 2, ..., s.$

II. $\sum_{i=1}^{r} \sum_{j=1}^{s} P(x_i, y_j) = 1.$

Definición 1.9. Función de distribución bidimensional.

Sea una variable aleatoria bidimensional (X, Y) de tipo discreto cuya distribución de probabilidades es $p_{ij} = P(x_i, y_j)$, $i = 1, 2, \ldots, r$ y $j = 1, 2, \ldots, s$. Definimos la **función de distribución conjunta**, que notaremos por $F(x, y)$ como:

$$F(x, y) = P(X \leqslant x, Y \leqslant y) = \sum_{x_i \leqslant x} \sum_{y_j \leqslant y} P(x_i, y_i) =$$

$$= \sum_{x_i = x_1}^{x} \sum_{y_j = y_1}^{y} P(X = x_i, Y = y_j)$$

y representa la suma de las probabilidades puntuales $P(x, y)$ hasta el valor (x, y) inclusive de la variable aleatoria bidimensional (X, Y).

Consideramos ahora una variable aleatoria bidimensional de tipo continuo (X, Y).

Definición 1.10. Función de densidad bidimensional.

Sea (X, Y) una variable aleatoria bidimensional de tipo continuo, y si existe una función $f(x, y)$ tal que verifica:

I. $f(x, y) \geqslant 0, \quad -\infty < x < +\infty, \quad -\infty < y < +\infty.$

II. $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y)\, dx\, dy = 1.$

diremos que $f(x, y)$es la **función de densidad** de la variable aleatoria bidimensional continua (X, Y).

Esta función de densidad conjunta, de forma análoga al caso unidimensional, se puede interpretar como un histograma de frecuencias relativas conjuntas para X e Y, pues la función de densidad $f(x, y)$ representa una superficie de densidad de probabilidad en el espacio tridimensional, y el volumen, por debajo de esta superficie y por encima del rectángulo *x*m *X* m *x*e *y*m *Y* m *y*, es igual a la probabilidad de que las variables aleatorias tienen valores dentro del rectángulo indicado, es decir:

$$P(x_1 \leqslant X \leqslant x_2, y_1 \leqslant Y \leqslant y_2) = \int_{x_1}^{x_2} \int_{y_1}^{y_2} f(x, y)\, dx\, dy$$

y esta probabilidad viene representada por el volumen sombreado del gráfico 1.10.

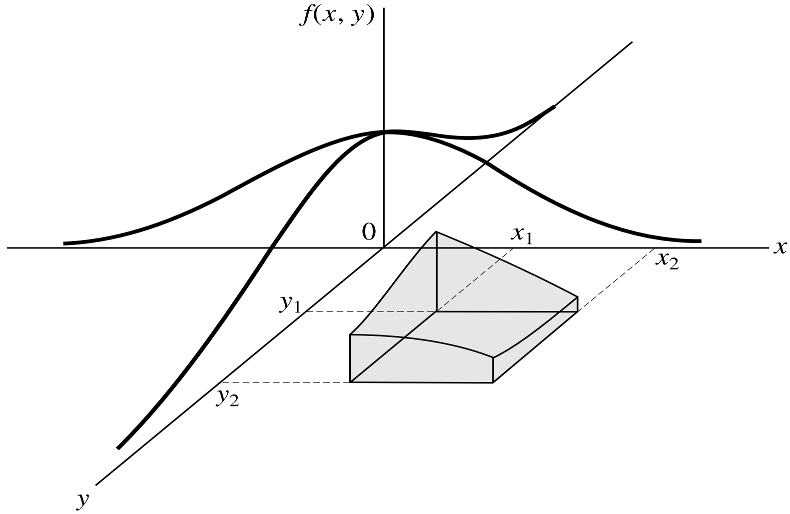


Gráfico 1.10. Función de densidad bidimensional.

Definición 1.11. Función de distribución bidimensional.

> Sea una variable aleatoria bidimensional (X, Y) de tipo continuo que toma valores sobre el espacio bidimensional $\mathbb{R}^2$ y cuya función de densidad es $f(x, y)$. Se define la función de distribución de la variable aleatoria bidimensional, que notaremos por $F(x, y)$, como:
>
> $$F(x, y) = P(X \leqslant x, Y \leqslant y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(x, y)\, dx\, dy$$

La función de distribución bidimensional satisface una serie de propiedades análogas a las indicadas en el caso unidimensional[2]:

1. $F(-\infty, -\infty) = F(-\infty, y) = F(x, -\infty) = 0.$

2. $F(+\infty, +\infty) = 1.$

3. Es monótona no decreciente respecto a las dos variables, es decir:

$$\text{Si} \quad x_1 < x_2 \quad \Rightarrow \quad F(x_1, y) \leqslant F(x_2, y)$$

$$\text{Si} \quad y_1 < y_2 \quad \Rightarrow \quad F(x, y_1) \leqslant F(x, y_2)$$

[2] Las propiedades 1, 2, 3 y 4 son válidas para el caso discreto y el caso continuo y la 5 sólo es válida para el caso continuo.

4. Si $a_1 < a_2$ y $b_1 < b_2$, siendo $a_1, a_2, b_2, b_2 \in \mathbb{R}$, entonces:

$$P(a_1 < X \leqslant a_2,\ b_1 < Y \leqslant b_2) \geqslant 0$$

5. La derivada parcial segunda respecto de x e y de la función de distribución $F(x, y)$ es la función de densidad:

$$\frac{\partial^2 F(x,\ y)}{\partial x \partial y} = f(x,\ y)$$

Ejemplo 1.4

Sea una variable aleatoria bidimensional (X, Y) cuya función de densidad es:

$$f(x,\ y) = \begin{cases} 1, & 0 \leqslant x \leqslant 1; \quad 0 \leqslant y \leqslant 1 \\ 0, & \text{en el resto} \end{cases}$$

Obtener:

1. La representación gráfica de la función de densidad.
2. La función de distribución.
3. La $P(X \leqslant 0{,}2,\ Y \leqslant 0{,}4)$.
4. La $P(0{,}1 \leqslant X \leqslant 0{,}2,\ 0 \leqslant Y \leqslant 0{,}4)$ y su interpretación.

Solución:

La representación gráfica de la función de densidad viene dada por el gráfico 1.11.

La función de distribución será:

$$F(x,\ y) = P(X \leqslant x,\ Y \leqslant y) = \int_0^x \int_0^y 1 \cdot dx\, dy =$$

$$= \begin{cases} xy & \text{si} \quad 0 \leqslant \text{x} < 1, \quad 0 \leqslant y < 1 \\ x & \text{si} \quad 0 \leqslant x < 1, \quad y \geqslant 1 \\ y & \text{si} \quad x \geqslant 1, \quad 0 \leqslant y < 1 \\ 1 & \text{si} \quad x \geqslant 1, \quad y \geqslant 1 \\ 0 & \text{en el resto} \end{cases}$$

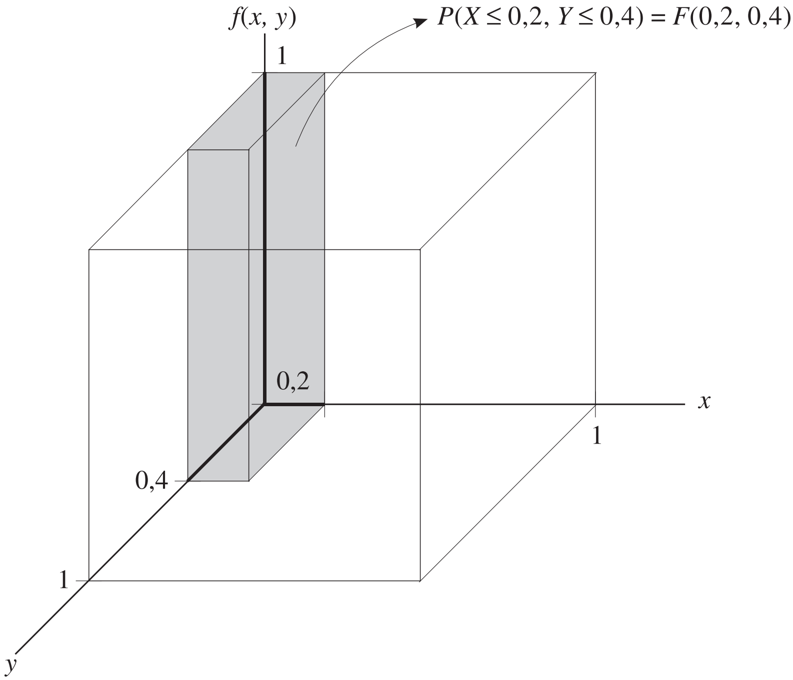


Gráfico 1.11. Representación gráfica de la función de densidad del ejemplo 1.4.

3. $P(X \leqslant 0{,}2,\ Y \leqslant 0{,}4) = F(0{,}2,\ 0{,}4) = \int_0^{0,2} \int_0^{0,4} 1 \cdot dx\, dy =$

$$= \int_0^{0,2} [y]_0^{0,4} dx = [y]_0^{0,4} \cdot [x]_0^{0,2} = 0{,}08$$

y representa el volumen bajo la $f(x, y) = 1$ en la zona sombreada del plano (x, y).

4. $P(0{,}1 \leqslant X \leqslant 0{,}2,\ 0 \leqslant Y \leqslant 0{,}4) = \int_{0,1}^{0,2} \int_0^{0,4} 1 \cdot dx\, dy =$

$$= \int_{0,1}^{0,2} [y]_0^{0,4} dx =$$

$$= [y]_0^{0,4} [x]_{0,1}^{0,2} = 0{,}04$$

Esta probabilidad representa el volumen bajo la curva función de densidad $f(x, y) = 1$, sobre la región del plano (x, y) delimitada por:

$$0{,}1 \leqslant x \leqslant 0{,}2 \quad \text{y} \quad 0 \leqslant y \leqslant 0{,}4$$

1.2.2. Distribuciones marginales

Anteriormente hemos estudiado las variables aleatorias bidimensionales y conocíamos el comportamiento conjunto de las dos variables aleatorias, es decir conocíamos la distribución conjunta de (X, Y). Ahora nos interesa conocer la distribución de alguna o de ambas variables aleatorias, por separado, a partir de la información que nos proporcionaba la distribución conjunta de (X, Y), y esto nos lleva al concepto de **distribuciones marginales**.

En el caso de una variable aleatoria bidimensional (X, Y), teniendo en cuenta que, todos los sucesos bivariantes, correspondientes a los diferentes valores que puede tomar la variable aleatoria bidimensional, es decir:

$$(X = x_i,\ Y = y_j); \quad i, j = 1, 2, 3, \ldots$$

eran sucesos mutuamente excluyentes, entonces podremos decir que el suceso univariante $X = x_i$ es la unión de todos los sucesos bivariantes $(X = x_i, Y = y_j)$ para todos los valores posibles de y_j.

Para el ejemplo de la Tabla 1.5, tenemos:

$$P(x_1) = P(X = x_1) = P(X = 0) = P(X = 0,\ Y = 0) + P(X = 0,\ Y = 1) = \\ = 0{,}6 + 0{,}2 = 0{,}8$$

$$P(x_2) = P(X = x_2) = P(X = 1) = P(X = 1,\ Y = 0) + P(X = 1,\ Y = 1) = \\ = 0{,}2 + 0 = 0{,}2$$

$$P(y_1) = P(Y = y_1) = P(Y = 0) = P(X = 0,\ Y = 0) + P(X = 1,\ Y = 0) = \\ = 0{,}6 + 0{,}2 = 0{,}8$$

$$P(y_2) = P(Y = y_2) = P(Y = 1) = P(X = 0,\ Y = 1) + P(X = 1,\ Y = 1) = \\ = 0{,}2 + 0 = 0{,}2$$

Definición 1.12. Distribución de probabilidad marginal.

Sea una variable aleatoria bidimensional (X, Y) de tipo discreto cuya distribución de probabilidad es $p_{ij} = P(x_i, y_j)$. Entonces las **distribuciones de probabilidad marginales**, de X e Y, respectivamente, vienen dadas por:

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

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

Observemos que el término **marginal** tiene un significado intuitivo y que la distribución de probabilidad marginal de X es la probabilidad de que $X = x$ con independencia del valor que tome Y, luego también podemos escribir que:

$$P_{i.} = P(X = x_i) = \sum_j p_{ij}$$

Análogamente para la distribución de probabilidad marginal de Y

$$P_{.j} = P(Y = y_j) = \sum_i p_{ij}$$

Definición 1.13. Función de distribución marginal.

> Sea una variable aleatoria bidimensional (X, Y) de tipo discreto cuya distribución de probabilidad es $p_{ij} = P(x, y)$. Entonces las **funciones de distribución marginales** de X e Y, respectivamente, se definen como:
>
> $$F_1(x) = F(x, +\infty) = P(X \leqslant x, Y < +\infty) = \sum_{x_i \leqslant x} \sum_{y_j} P(x_i, y_j) = \sum_{x_i \leqslant x} P_{i.}$$
>
> $$F_2(y) = F(+\infty, y) = P(X < +\infty, Y \leqslant y) = \sum_{y_j \leqslant y} \sum_{x_i} P(x_i, y_j) = \sum_{y_j \leqslant y} P_{.j}$$

Si ahora consideramos una variable aleatoria bidimensional de tipo continuo (X, Y), tendremos sus correspondientes funciones de densidad marginales y sus funciones de distribución.

Definición 1.14. Función de densidad marginal.

> Sea una variable aleatoria bidimensional (X, Y) de tipo continuo, cuya función de densidad conjunta es $f(x, y)$. Entonces las **funciones de densidad marginales** de X e Y, respectivamente, se definen como:
>
> $$f_1(x) = \int_{-\infty}^{+\infty} f(x, y)\, dy$$
>
> $$f_2(y) = \int_{-\infty}^{+\infty} f(x, y)\, dx$$

Definición 1.15. Función de distribución marginal.

Sea una variable aleatoria bidimensional (X, Y) de tipo continuo, cuya función de densidad conjunta es $f(x, y)$. Entonces se definen las **funciones de distribución marginales** de X e Y como:

$$F_1(x) = F(x, +\infty) = P(X \leqslant x, Y < +\infty) = \int_{-\infty}^{x} \left(\int_{-\infty}^{+\infty} f(x, y)\, dy \right) dx =$$

$$= \int_{-\infty}^{x} f_1(x)\, dx$$

$$F_2(y) = F(+\infty, y) = P(X < +\infty, Y \leqslant y) = \int_{-\infty}^{y} \left(\int_{-\infty}^{+\infty} f(x, y)\, dx \right) dy =$$

$$= \int_{-\infty}^{y} f_2(y)\, dy$$

La distribución marginal de la variable aleatoria X nos da la probabilidad de que $X \leq x$, con independencia de los valores que toma la variable aleatoria Y. Físicamente sería la cantidad de masa de probabilidad que hay en el plano XY pero a la izquierda y sobre la recta $X = x$, ver gráfico 1.12.

Una forma posible de calcular la cantidad de masa de probabilidad que queda sobre la recta $X = x$ y a la izquierda de ella, podría ser acumularla sobre la recta $X = x$ y después asignarla al punto $(X = x, Y = 0)$. De aquí el nombre que recibe de distribución marginal.

Análogamente sería para la distribución marginal de Y.

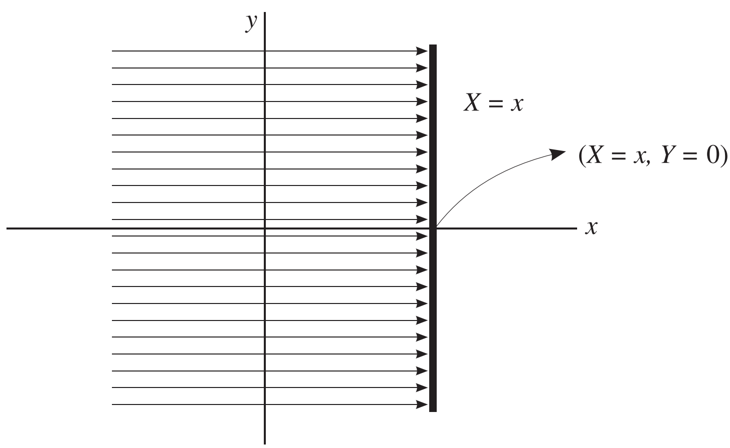


Gráfico 1.12. Distribución marginal de la variable aleatoria X.

1.2.3. Distribuciones condicionadas

En apartados anteriores hemos estudiado el comportamiento conjunto e individual de las variables aleatorias que componían la variable aleatoria bidimensional, obteniendo las distribuciones conjuntas y marginales. Ahora nos interesa conocer cómo se distribuye una de las variables cuando se imponen condiciones a la otra y esto nos llevará al estudio de las **distribuciones condicionadas.**

En el capítulo anterior estudiábamos la probabilidad condicionada $P(B/A)$, y nos indicaba el cambio que se producía en la probabilidad de un suceso como consecuencia de la ocurrencia de otro, y la definíamos como:

$$P(B/A) = \frac{P(A \cap B)}{P(A)}, \quad P(A) > 0$$

Por analogía con esta definición y considerando la variable aleatoria bidimensional discreta (X, Y), con su correspondiente distribución de probabilidades, podremos definir la **probabilidad condicionada** $P(X = x/Y = y)$, que indica la probabilidad de que la variable aleatoria $X = x$cuando la variable aleatoria $Y = y_j$, como:

$$P(X = x_i/Y = y_j) = \frac{P(X = x_i,\ Y = y_j)}{P(Y = y_j)}$$

siempre que $P(Y = y_j) > 0$.

Definición 1.16. Distribución de probabilidad condicionada de la variable aleatoria discreta X dado $Y = y_j$.

$$P(x_i/y_j) = P(X = x_i/Y = y_j) = \frac{P(X = x_i,\ Y = y_j)}{P(Y = y_j)} = \frac{p_{ij}}{P_{.j}}$$

siempre que

$$P(Y = y_j) > 0.$$

En esta expresión y_j es fijo y x_i varía sobre todos los posibles valores de la variable aleatoria X, obteniendo así la distribución de probabilidad de la variable aleatoria discreta X bajo la condición $Y = y_j$.

Análogamente tendríamos la **distribución de probabilidad condicionada de la variable aleatoria discreta *Y* dado $X = x_i$.**

$$P(y_j/x_i) = P(Y = y_j/X = x_i) = \frac{P(X = x_i,\ Y = y_j)}{P(X = x_i)} = \frac{p_{ij}}{P_{i.}}$$

Siempre que

$$P(X = x_i) > 0$$

Aquí x_i es fijo e y_j varía sobre todos los posibles valores de la variable aleatoria Y.

Es evidente que:

$$P(x_i/y_j) \geqslant 0$$

$$P(y_j/x_i) \geqslant 0$$

y además:

$$\sum_i P(X = x_i/Y = y_j) = \frac{\sum_i p_{ij}}{P_{.j}} = \frac{P_{.j}}{P_{.j}} = 1$$

$$\sum_j P(Y = y_j/X = x_i) = \frac{\sum_j p_{ij}}{P_{i.}} = \frac{P_{i.}}{P_{i.}} = 1$$

Luego son efectivamente distribuciones de probabilidad.

Definición 1.17. Función de distribución condicionada de la variable aleatoria discreta X dado $Y = y_j$.

$$F(x/y_i) = P(X \leqslant x/Y = y_j) = \frac{\sum\limits_{x_i \leqslant x} P(X = x_i,\ Y = y_j)}{P(Y = y_j)} = \frac{\sum\limits_{x_i \leqslant x} p_{ij}}{P_{.j}}$$

Siempre que:

$$P(Y = y_j) = P_{.j} > 0.$$

Análogamente tendríamos la **función de distribución condicionada de la variable aleatoria discreta *Y* dado $X = x_i$.**

$$F(y/x_i) = P(Y \leqslant y/X = x_i) = \frac{\sum\limits_{y_j \leqslant y} P(X = x_i,\ Y = y_j)}{P(X = x_i)} = \frac{\sum\limits_{y_j \leqslant y} p_{ij}}{P_{i.}}$$

Siempre que:

$$P(X = x_i) = P_{i.} > 0$$

En el **caso continuo** podemos obtener, por analogía, la función de densidad condicionada y su correspondiente función de distribución condicionada. Aquí no se pueden obtener de manera tan sencilla y directa a como sucedía en el caso discreto, pues en el caso continuo la

$$P(X = x_i) = 0 \quad \text{y la} \quad P(Y = y_j) = 0$$

y las probabilidades condicionadas:

$$P(X \leqslant x/Y = y) \quad \text{y} \quad P(Y \leqslant y/X = x)$$

no están definidas, pues los denominadores son nulos.

Definición 1.18. Función de densidad condicionada de la variable aleatoria continua X dado $Y = y$.

$$f(x/y) = \begin{cases} \dfrac{f(x,\ y)}{f_2(y)}, & f_2(y) > 0 \\ 0, & \text{en el resto} \end{cases}$$

Análogamente se tiene **la función de densidad condicionada de la variable aleatoria continua Y dado $X = x$.**

$$f(y/x) = \begin{cases} \dfrac{f(x,\ y)}{f_1(x)}, & f_1(x) > 0 \\ 0, & \text{en el resto} \end{cases}$$

Definición 1.19. Función de distribución condicionada de la variable aleatoria continua X dado $Y = y$.

$$F(x/y) = \int_{-\infty}^{x} f(x/y)\ dx = \frac{\displaystyle\int_{-\infty}^{x} f(x,\ y)\ dx}{f_2(y)},$$

De la misma forma tendríamos la **función de distribución condicionada de la variable aleatoria continua *Y* dada *X* = *x*.**

$$F(y/x) = \int_{-\infty}^{y} f(y/x)\,dy = \frac{\displaystyle\int_{-\infty}^{y} f(x,\ y)\ dy}{f_1(x)},$$

Ejemplo 1.5

Sea una urna con 21 bolas, numeradas desde el número 1 al 21, de manera que no hay dos bolas con el mismo número. Deseamos conocer dos características de estos números: su divisibilidad por 2 y por 3. Elegimos una bola al azar, y consideramos una variable aleatoria X, tal que, si aparece un número par (divisible por 2) le asignamos el **uno** y si aparece número impar le asignamos el **cero**. Entonces, la variable aleatoria X toma dos valores $x_1 = 1$ y $x_2 = 0$. Análogamente definimos una variable aleatoria Y, tal que si aparece un número divisible por 3, entonces le asignamos el **uno** y en los demás casos le asignamos el **cero**. Es decir, la variable aleatoria Y toma los valores $y_1 = 1$ e $y_2 = 0$. Obtener:

1.º La distribución de probabilidad.
2.º Las distribuciones marginales.
3.º La probabilidad de que el número elegido sea divisible por 3, siendo par.

Solución:

1.º Para la variable aleatoria X se tiene:

$$X \begin{cases} 1, & \text{si aparece un número divisible por 2} \\ 0, & \text{en los demás casos} \end{cases}$$

para la variable aleatoria Y se tiene: 1, si aparece un número divisible por 3

$$Y \begin{cases} 1, & \text{si aparece un número divisible por 3} \\ 0, & \text{en los demás casos} \end{cases}$$

Entre las 21 bolas tenemos los siguientes tipos de números:

divisibles por 2 y por 3 ——— 3
divisibles por 2 y no por 3 ——— 7
divisibles por 3 y no por 2 ——— 4
no divisibles por 2 ni por 3 ——— 7

Como todas las bolas tienen la misma probabilidad de ser extraídas, tenemos la siguiente distribución de probabilidades de la variable aleatoria bidimensional de tipo discreto (X, Y):

$$p_{11} = P[X = 1,\ Y = 1] = \frac{3}{21} \qquad p_{21} = P[X = 0,\ Y = 1] = \frac{4}{21}$$

$$p_{12} = P[X = 1,\ Y = 0] = \frac{7}{21} \qquad p_{22} = P[X = 0,\ Y = 0] = \frac{7}{21}$$

Observemos que:

$$\sum_i \sum_j p_{ij} = 1$$

2.º Las distribuciones marginales son:

Clasificamos los números según su divisibilidad por 2 o por 3.

La distribución marginal de la variable aleatoria X es:

$$P_{1.} = P[X = 1] = p_{11} + p_{12} = \frac{3}{21} + \frac{7}{21} = \frac{10}{21}$$

$$P_{2.} = P[X = 0] = p_{21} + p_{22} = \frac{4}{21} + \frac{7}{21} = \frac{11}{21}$$

$$\sum_i P_{i.} = P_{1.} + P_{2.} = \frac{10}{21} + \frac{11}{21} = 1$$

Análogamente la distribución marginal de la variable aleatoria Y es:

$$P_{.1} = P[Y = 1] = p_{11} + p_{12} = \frac{3}{21} + \frac{4}{21} = \frac{7}{21}$$

$$P_{.2} = P[Y = 0] = p_{12} + p_{21} = \frac{7}{21} + \frac{7}{21} = \frac{14}{21}$$

$$\sum_j P_{.j} = 1$$

Utilizando las frecuencias absolutas podemos obtener la distribución conjunta de frecuencias absolutas y las distribuciones marginales de frecuencias absolutas:

	N.^{os} divisibles por 2	N.^{os} no divisibles por 2	Total
N.^{os} divisibles por 3	3	4	7
N.^{os} no divisibles por 3	7	7	14
Total	10	11	21

Tanto la distribución de probabilidad conjunta como las distribuciones de probabilidad marginales las podemos resumir en la Tabla 1.6.

$Y \backslash X$	1	0	$P_{.j}$
1	$\frac{3}{21}$	$\frac{4}{21}$	$\frac{7}{21}$
0	$\frac{7}{21}$	$\frac{7}{21}$	$\frac{14}{21}$
$P_{i.}$	$\frac{10}{21}$	$\frac{11}{21}$	1

Tabla 1.6. *Distribución de probabilidad conjunta de la variable (X, Y).*

3.º Para obtener la probabilidad de que el número elegido sea divisible por 3, siendo par, tendremos que calcular la siguiente probabilidad condicionada:

$$P[Y = 1/X = 1] = \frac{P[Y = 1,\ X = 1]}{P[X = 1]} = \frac{p_{11}}{P_{1.}} = \frac{3}{10}$$

Análogamente tendríamos las demás probabilidades condicionadas:

$$P[Y = 0/X = 1] = \frac{P[Y = 0,\ X = 1]}{P[X = 1]} = \frac{p_{12}}{P_{1.}} = \frac{7}{10}$$

$$P[Y = 1/X = 0] = \frac{P[Y = 1,\ X = 0]}{P[X = 0]} = \frac{p_{21}}{P_{2.}} = \frac{4}{11}$$

$$P[Y = 0/X = 0] = \frac{P[Y = 0,\ X = 0]}{P[X = 0]} = \frac{p_{22}}{P_{2.}} = \frac{7}{11}$$

Ejemplo 1.6

Sea una variable aleatoria bidimensional (X, Y) de tipo continuo cuya función de densidad es:

$$f(x, y) = \begin{cases} k(x + y), & 0 \leqslant x \leqslant 2, \quad 0 \leqslant y \leqslant 1 \\ 0 & \text{en el resto} \end{cases}$$

Obtener:

1.º El valor de k para que $f(x, y)$ sea una función de densidad.

2.º La función de densidad marginal de la variable aleatoria X.

3.º La función de densidad condicionada de la variable aleatoria Y, dado $X = 1$.

4.º La función de distribución condicionada de la variable aleatoria Y, dado $X = x$.

Solución:

1.º Para que $f(x, y)$ sea una función de densidad se tiene que verificar:

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y)\, dx\, dy = 1$$

en nuestro caso será:

$$\int_0^2 \left(\int_0^1 k(x + y)\, dy \right) dx = 1$$

En efecto:

$$\int_0^2\left(\int_0^1 k(x+y)\,dy\right)dx = k\int_0^2\left[xy+\frac{y^2}{2}\right]_0^1 dx = k\int_0^2\left(x+\frac{1}{2}\right)dx =$$

$$= k\left[\frac{x^2}{2}+\frac{x}{2}\right]_0^2 = 3k = 1$$

Luego $k = \frac{1}{3}$, y la función de densidad será:

$$f(x,\ y) = \begin{cases} \frac{1}{3}\ (x+y), & 0 \leqslant x \leqslant 2, \quad 0 \leqslant y \leqslant 1 \\ 0, & \text{en el resto} \end{cases}$$

2.º La función de densidad marginal de la variable aleatoria X, según la definición 1.14, será:

$$f_1(x) = \int_0^1 \frac{1}{3}\ (x+y)\,dy = \frac{1}{3}\left[xy+\frac{y^2}{2}\right]_0^1 = \frac{1}{3}\left(x+\frac{1}{2}\right), \quad \text{si} \quad 0 \leqslant x \leqslant 2$$

3.º La función de densidad condicionada de la variable aleatoria Y, dado $X = x$, según la definición 1.18, será:

$$f(y/x) = \frac{f(x,\ y)}{f_1(x)} = \frac{\frac{\cancel{1}}{\cancel{3}}\ (x+y)}{\frac{\cancel{1}}{\cancel{3}}\left(x+\frac{1}{2}\right)} = \frac{2(x+y)}{2x+1}, \quad \begin{matrix} \text{si} & 0 \leqslant y \leqslant 1 \\ \text{y} & 0 \leqslant x \leqslant 2 \end{matrix}$$

y particularizando para $X = 1$, tendremos:

$$f(y/x=1) = \frac{2}{3}\ (1+y), \quad 0 \leqslant y \leqslant 1$$

4.º La función de distribución condicionada de la variable aleatoria Y, dado $X = x$, según la definición 1.19, será:

$$F(y/x) = \int_0^y f(y/x)\,dy = \int_0^y \frac{2(x+y)}{2x+1}\,dy =$$

$$= \left[\frac{2xy}{2x+1}\right]_0^y + \left[\frac{y^2}{2x+1}\right]_0^y = \frac{2xy+y^2}{2x+1}$$

Luego:

$$F(y/x) = \begin{cases} 0 & , \quad y < 0 \\ \dfrac{2xy + y^2}{2x + 1} & , \quad 0 \leqslant y < 1 \\ 1 & , \quad y \geqslant 1 \end{cases}$$

1.3. INDEPENDENCIA DE VARIABLES ALEATORIAS

En el curso anterior estudiábamos la independencia de sucesos y decíamos que dos sucesos A y B eran independientes si se verifica que:

$$P(A \cap B) = P(A) \cdot P(B)$$

De manera análoga podemos definir la independencia de variables aleatorias.

Sea (X, Y) una variable aleatoria bidimensional cuya función de distribución conjunta es $F(x, y)$, y sus funciones de distribución marginales $F_1(x)$ y $F_2(y)$ respectivamente.

Ejemplo 1.7

La distribución conjunta de la variable aleatoria X, número de tazas de café servidas por un camarero, e Y, número de sobres de azúcar por taza solicitadas, viene dada en la tabla 1.7. Determinar si ambas variables aleatorias son independientes.

Tabla 1.7. *Distribución de conjunto del número X de tazas de café servidas y el número Y de sobres de azúcar por taza solicitada.*

X \ Y	0	1	2	3	$P_{i.}$
1	0,03	0,04	0,02	0,01	0,10
2	0,09	0,12	0,06	0,03	0,30
3	0,09	0,12	0,06	0,03	0,30
4	0,06	0,08	0,04	0,02	0,20
5	0,03	0,04	0,02	0,01	0,10
$P_{.j}$	0,30	0,40	0,20	0,10	1

Solución

En la tabla 1.7 tenemos en la última columna y en la última fila las distribuciones marginales de X e Y, respectivamente.

Sabemos que para que ambas variables aleatorias sean independientes se tiene que verificar:

$$p_{ij} = P(x_i, y_i) = P(X = x_i, Y = y_i) = P[X = x_i] \cdot p[Y = y_j] = P_{i.} \cdot P_{.j}$$

para todo x_i e y_j. Podemos comprobar que esto ocurre, luego X e Y son independientes.

Definición 1.20. Independencia de variables aleatorias.

Diremos que las variables aleatorias X e Y son independientes solo si se verifica que la función de distribución conjunta $F(x, y)$ es igual al producto de sus distribuciones marginales:

$$F_1(x) \cdot F_2(y)$$

Es decir: X e Y independientes $\Leftrightarrow F(x, y) = F_1(x) \cdot F_2(y), \quad \forall (x, y)$

O bien, de manera equivalente, en el caso discreto, si:

$$P(x_i/Y = y_j) = P_{i.} \quad , \quad \forall\, j$$
$$P(y_j/X = x_i) = P_{.j} \quad , \quad \forall\, i$$

y en el caso continuo, si:

$$f(x/y) = f_1(x) \quad , \quad \forall\, y$$
$$f(y/x) = f_2(y) \quad , \quad \forall\, x$$

Si X e Y son variables aleatorias **discretas**, cuya distribución de probabilidad conjunta $p_{ij} = P(x_i, x_j)$ y sus distribuciones de probabilidad marginales $P_{i.} = P(x_i)$ y $P_{.i} = P(y_j)$ respectivamente, entonces diremos que **las variables aleatorias son independientes** si y solo si se verifica:

$$P(x_i, y_j) = P(x_i) \cdot P(y_j), \quad \forall\, (x_i, y_j)$$

o bien

$$p_{ij} = P_{i.} \cdot P_{.j}$$

Si X e Y son variables aleatorias **continuas**, con función de densidad conjunta $f(x, y)$ y funciones de densidad marginales $f_1(x)$ y $f_2(y)$ respectivamente, entonces diremos que las **variables aleatorias son independientes** si y solo si se verifica:

$$f(x, y) = f_1(x) \cdot f_2(y) \quad \forall \ (x, y)$$

Ejemplo 1.8

Considerando el ejemplo 1.5 y más concretamente la tabla 1.6 de la distribución de probabilidad conjunta de la variable aleatoria bidimensional (X, Y), probar que las variables X e Y no son independientes.

Solución:

Para que las variables aleatorias discretas X e Y sean independientes será necesario que:

$$P(x_i, y_i) = P(x_i) \cdot P(y_j), \quad \forall \ (x_i, y_i)$$

o lo que es lo mismo que:

$$P(X = x_i, Y = y_j) = P(X = x_i) \cdot P(Y = y_j), \quad \forall \ (x_i, y_j)$$

o también que:

$$p_{ij} = P_{i.} \cdot P_{.j}$$

Observando la Tabla 1.6 vemos que:

$$P(X = 1, Y = 1) = \frac{3}{21}$$

$$P(X = 1) = \frac{10}{21}$$

$$P(Y = 1) = \frac{7}{21}$$

y resulta que:

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

pues

$$\frac{3}{21} \neq \frac{10}{21} \cdot \frac{7}{21}$$

Luego las variables aleatorias X e Y no son independientes.