Capítulo 4

Distribuciones Conjuntas e Independencia

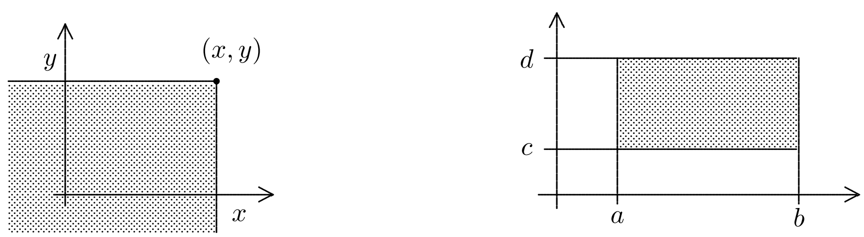
4.1. Distribución Conjunta de Dos Variables Aleatorias.

Sean X e Y dos variables aleatorias sobre un espacio de probabilidad común $(\Omega, \mathcal{A}, P)$. Llamaremos *función de distribución conjunta*, o simplemente *distribución conjunta*, de X e Y, a la función

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

En algunas ocasiones usaremos $F_{X,Y}(x,y)$ en lugar de $F(x,y)$ para destacar que se trata de la distribución conjunta de X e Y.

La definición anterior indica que $F(x,y)$ es la probabilidad de que el punto (X,Y) pertenezca al cuadrante que queda "abajo y a la izquierda"del punto (x,y), incluyendo el borde, indicado en la figura 4.1 (a).



(a) (b)

Figura 4.1

De esta manera

$$F(x,y) = P(\{\omega : X(\omega) \leq x\} \cap \{\omega : Y(\omega) \leq y\})$$

A partir de la definición obtenemos (ver figura 4.1 (b))

$$\begin{aligned} P(a < X \leq b,\ c < Y \leq d) &= P(X \leq b,\ Y \leq d) - P(X \leq b,\ Y \leq c) \\ &\quad - P(X \leq a,\ Y \leq d) + P(X \leq a,\ Y \leq c) \\ &= F(b,d) - F(b,c) - F(a,d) + F(a,c). \end{aligned} \tag{4.1}$$

La distribución conjunta de dos variables tiene además las siguientes propiedades:

1. $F(x,y)$ es creciente en cualquiera de las dos variables. Por ejemplo, si $x < x'$ entonces

$$\{\omega : X(\omega) \leq x\} \subset \{\omega : X(\omega) \leq x'\}$$

y por lo tanto

$$\begin{aligned} F(x,y) &= P(\{\omega : X(\omega) \le x\} \cap \{\omega : Y(\omega) \le y\}) \\ &\le P(\{\omega : X(\omega) \le x'\} \cap \{\omega : Y(\omega) \le y\}) \\ &= F(x', y). \end{aligned}$$

2. $\lim_{x\to-\infty} F(x,y) = \lim_{y\to-\infty} F(x,y) = 0;\ \lim_{\substack{x\to+\infty\\ y\to+\infty}} F(x,y) = 1.$

   y como la función es creciente en ambas variables se deduce que, para cualesquiera x, y,

$$0 \le F(x,y) \le 1.$$

3. $F(x,y)$ es continua por la derecha en cualquiera de las variables.

En contraste con el caso de funciones de distribución unidimensionales, para que una función $F(x,y)$ sea la distribución conjunta de un par de variables aleatorias X e Y, no es suficiente que tenga las tres propiedades que hemos considerado. Por ejemplo, la función

$$F(x,y) = \begin{cases} 0 & \text{si } x+y<0 \\ 1 & \text{si } x+y \ge 0 \end{cases}$$

toma el valor 0 en los puntos que están debajo de la recta $y = -x$, y el valor 1 para los puntos sobre y por encima de la recta (ver figura 4.2).

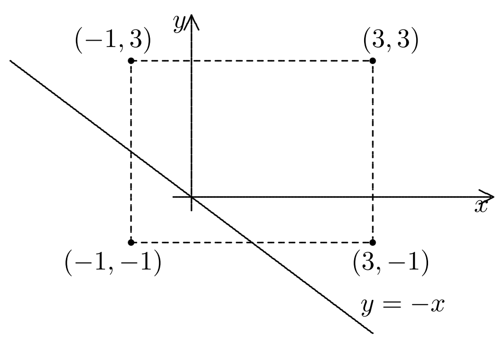


Figura 4.2

La función es creciente, continua por la derecha y satisface la propiedad 2. Sin embargo, si aplicamos la fórmula (4.1) para calcular la probabilidad de que el punto (X, Y) esté en el rectángulo de vértices $(3,3);\ (3,-1);\ (-1,3);\ (-1,-1)$, obtenemos

$$P(-1 < X \le 3, -1 < Y \le 3) = F(3,3) - F(3,-1) - F(-1,3) + F(-1,-1) = -1$$

lo cual es imposible ya que una probabilidad no puede ser negativa. Por lo tanto es necesario añadir la condición de que el segundo miembro de la relación (4.1) no sea negativo para ninguna colección de números $a < b,\ c < d$.

Teorema 4.1 *Una función $F(x,y)$ es la distribución conjunta de un par de variables aleatorias si y sólo si satisface las propiedades 1, 2 y 3 y además para cualesquiera $a < b,\ c < d$,*

$$F(b,d) - F(a,d) - F(b,c) + F(a,c) \ge 0.$$

A partir de la función de distribución conjunta $F_{X,Y}$ de dos variables aleatorias es posible obtener las funciones de distribución F_X y F_Y correspondientes a las variables X e Y. En efecto, para cualquier $x \in \mathbb{R}$ tenemos

$$\begin{aligned} F_X(x) &= P(X \leq x) = P(X \leq x,\ Y < \infty) \\ &= \lim_{y\to\infty} P(X \leq x, Y \leq y) \\ &= \lim_{y\to\infty} F(x,y) \end{aligned}$$

y de manera similar, para cualquier $y \in \mathbb{R}$

$$F_Y(y) = \lim_{x\to\infty} F(x,y).$$

Las funciones F_X y F_Y se conocen como las *funciones de distribución marginales* de X e Y, respectivamente.

4.2. Variables Aleatorias Independientes.

Se dice que las variables X e Y son *independientes* si cualesquiera que sean los intervalos $(a,b]$ y $(c,d]$, se verifica que los eventos

$$\{X \in (a,b]\} \qquad \text{y} \qquad \{Y \in (c,d]\}$$

son independientes, es decir que

$$P(a < X \leq b,\ c < Y \leq d) = P(a < X \leq b)P(c < Y \leq d). \tag{4.2}$$

En términos menos precisos, de acuerdo a lo que hemos visto sobre independencia de eventos en el Capítulo 3, esta relación dice que el saber que el valor de X está en cierto intervalo no arroja información alguna sobre la probabilidad de que Y esté en otro intervalo.

Es fácil ver que la condición (4.2) es equivalente a la condición

$$F_{X,Y}(x,y) = F_X(x)F_Y(y) \quad \text{para todo } x, y \in \mathbb{R}. \tag{4.3}$$

En efecto, si se cumple (4.2) basta poner $b = x,\ d = y$ y hacer tender $a \to -\infty,\ c \to -\infty$, para obtener (4.3). Recíprocamente, si se cumple (4.3), poniendo $F_{X,Y} = F$ tenemos

$$\begin{aligned} P(a < X \leq b,\ c < Y \leq d) &= F(b,d) - F(a,d) - F(b,c) + F(a,c) \\ &= F_X(b)F_Y(d) - F_X(a)F_Y(d) - F_X(b)F_Y(c) + F_X(a)F_Y(c) \\ &= (F_X(b) - F_X(a))(F_Y(d) - F_Y(c)) \\ &= P(a < X \leq b)P(c < Y \leq d) \end{aligned}$$

o sea que (4.3) implica (4.2), cualesquiera sean $a,\ b,\ c$ y d.

Las relaciones (4.2) y (4.3) dicen que los eventos $\{X \in B_1\}$ y $\{Y \in B_2\}$ son independientes cuando B_1 y B_2 son intervalos semiabiertos, en el caso de (4.2), y semirectas cerradas a la derecha, en el caso de (4.3). Es posible probar, aunque no lo haremos en este texto, que (4.3), o equivalentemente (4.2), implica que los eventos $\{X \in B_1\}$ y $\{Y \in B_2\}$ son independientes para cualesquiera conjuntos de Borel B_1 y B_2.

Si tenemos una colección de variables aleatorias $\{X_s, s \in S\}$, donde S es cualquier conjunto de índices, diremos que son independientes si para cualquier subconjunto finito de índices $K \subset S$ y cualesquiera intervalos $I_k = (a_k, b_k]$, $k \in K$, se tiene que

$$P(\cap_{k\in K}\{X_k \in I_k\}) = \prod_{k\in K} P(X_k \in I_k).$$

Ejemplo 4.1
Si tenemos tres variables aleatorias X, Y y Z, para verificar que son independientes es necesario verificar que son independientes a pares y además que se se satisface

$$P(X \in I_1, Y \in I_2, Z \in I_2) = P(X \in I_1)P(Y \in I_2)P(Z \in I_2)$$

Para ver que independencia a pares no es suficiente para garantizar que las variables sean independientes tenemos el siguiere ejemplo. Consideramos el lanzamiento de dos monedas. X vale 1 si la primera moneda es Aguila y 0 si es Sol. Y se define de manera similar para la segunda moneda mientras que Z vale 1 si exactamente una de las dos monedas es Aguila, y vale 0 en otro caso. Vemos que el vector (X, Y, Z) toma los valores $(0,0,0); (0,1,1); (1,0,1); (1,1,0)$ cada uno con probabilidad $1/4$.

Es fácil ver que las funciones de probabilidad marginales de las tres variables son iguales:

$$P(X=0) = P(Y=0) = P(Z=0) = \frac{1}{2}; \qquad P(X=1) = P(Y=1) = P(Z=1) = \frac{1}{2}$$

y también es sencillo ver que son independientes dos a dos. Veamos que Y y Z son independientes. Para esto calculamos la función de probabilidad conjunta:

$$\begin{aligned} P(Y=0, Z=0) = P(0,0,0) = \frac{1}{4}; &\qquad P(Y=0, Z=1)) = P(1,0,1) = \frac{1}{4}; \\ P((Y=1, Z=0) = P(1,1,0) = \frac{1}{4}; &\qquad P(Y=1, Z=1) = P(0,1,1) = \frac{1}{4}. \end{aligned}$$

Sin embargo las tres variables no son independientes pues la función de probabilidad conjunta de (X, Y, Z) no es el producto de las marginales. Por ejemplo, $P(X=1, Y=1, Z=1) = 0$ pero $P(X=1)P(Y=1)P(Z=1) = 1/8$.

4.3. Distribución Conjunta de Variables Aleatorias Discretas.

Si X e Y son variables aleatorias discretas, con funciones de probabilidad respectivas

$$\begin{aligned} P(X = x_i) = p_i &\qquad (i = 1, 2, \dots) \\ P(Y = y_j) = q_j &\qquad (j = 1, 2, \dots) \end{aligned}$$

donde

$$p_i \geq 0, \quad q_j \geq 0, \quad \sum_i p_i = \sum_j q_j = 1,$$

la función de distribución conjunta queda definida por los números

$$r_{ij} = P(X = x_i,\ Y = y_j) \quad (i = 1, 2, \dots; j = 1, 2, \dots).$$

Estas probabilidades deben satisfacer las condiciones

$$p_i = \sum_j r_{ij} \quad (i = 1, 2, \dots) \tag{4.4}$$

$$q_j = \sum_i r_{ij} \quad (j = 1, 2, \dots) \tag{4.5}$$

ya que, por ejemplo,

$$\begin{aligned} p_i = P(X = x_i) &= P\left(\cup_j \{X = x_i,\ Y = y_j\}\right) \\ &= \sum_j P(X = x_i,\ Y = y_j) = \sum_j r_{ij}. \end{aligned}$$

En este caso $\{r_{ij}\}$ se llama la *función de probabilidad conjunta* y $\{p_i\}$, $\{q_j\}$ son las *funciones de probabilidad marginales*. A partir de $\{r_{ij}\}$, $F_{X,Y}$ se determina mediante

$$F_{X,Y}(x,y) = \sum_{\substack{i:x_i \le x \\ j:y_j \le y}} r_{ij}$$

y las variables X e Y son independientes si y sólo si, para todo $i,\ j$ se tiene que

$$r_{ij} = P(X = i,\ Y = j) = P(X = i)P(Y = j) = p_i q_j.$$

Supongamos que con probabilidad 1 las variables aleatorias X e Y toman un número finito de valores n y m respectivamente. La situación queda descrita por el siguiente cuadro:

		Y				
		y_1	y_2	$\cdots\cdots$	y_m	
	x_1	r_{11}	r_{12}	$\cdots\cdots$	r_{1m}	p_1
	x_2	r_{21}	r_{22}	$\cdots\cdots$	r_{2m}	p_2
X	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
	x_n	r_{n1}	r_{n2}	$\cdots\cdots$	r_{nm}	p_n
		q_1	q_2	$\cdots\cdots$	q_m	

Tabla 4.1

En la última columna aparecen $p_1,\ p_2, \dots, p_n$, que son las sumas respectivas de las filas (condición (4.4)) y en la última fila $q_1, q_2,\ \dots, q_m$, que son las sumas respectivas de las columnas (condición (4.5)).

Una situación de este tipo aparece en diversos problemas de aplicación. Supongamos que tenemos el proceso de producción de un objeto en el que nos interesan dos magnitudes, por ejemplo, el diámetro y la longitud de un cilindro, la densidad de un producto y la concentración de un componente del mismo, etc. Los valores obtenidos para estas dos magnitudes durante el proceso de producción fluctúan en virtud de diversas causas, algunas de ellas incontrolables, y otras cuyo control requeriría un costo elevado.

Sean α y β dichas magnitudes, procedemos a dividir el rango de variación de ambas en un número finito de secciones que numeramos ordenadamente, de 1 a n para la magnitud α y de 1 a m para la magnitud β. En la figura 4.4 hemos tomado $n = 5,\ m = 4$.

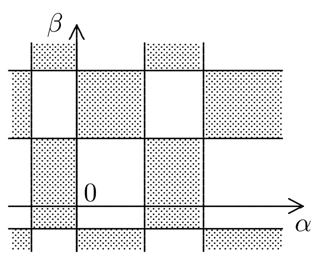


Figura 4.4

Llamemos X a la sección en la cual cae la magnitud α de un objeto e Y a la sección en la cual cae β. Para cada objeto tendremos entonces

$$r_{ij} = P(X = i,\ Y = j) \quad (i = 1, \dots, n;\ j = 1, \dots, m)$$

y tenemos definido un cuadro de doble entrada como el anterior con sus funciones de probabilidad (4.4) y (4.5).

Para estudiar la regulación de un proceso de producción, se extrae una muestra de N objetos producidos y se clasifica como se ha indicado:

$$
\begin{array}{cc|cccc|c}
 & & & & Y & & \\
 & & 1 & 2 & \cdots\cdots & m & \\
\hline
 & 1 & N_{11} & N_{12} & \cdots\cdots & N_{1m} & P_1 \\
 & 2 & N_{21} & N_{22} & \cdots\cdots & N_{2m} & P_2 \\
X & \vdots & \vdots & \vdots & & \vdots & \vdots \\
 & \vdots & \vdots & \vdots & & \vdots & \vdots \\
 & n & N_{n1} & N_{n2} & \cdots\cdots & N_{nm} & P_n \\
\hline
 & & Q_1 & Q_2 & \cdots\cdots & Q_m &
\end{array}
$$

Tabla 4.2

N_{ij} es el número de objetos de la muestra tales que el valor de α está en la i-ésima sección y el de β en la j-ésima.

A partir de una muestra de este tipo es posible inferir resultados sobre la Tabla 4.1, cuyos elementos en general son desconocidos. Por ejemplo, es interesante saber si las magnitudes consideradas fluctúan independientemente, es decir si

$$P(X = i,\ Y = j) = P(X = i)P(Y = j)$$

o sea

$$r_{ij} = p_i\, q_j$$

para $i = 1, \dots, n;\ j = 1, \dots, m$. Es claro que si ambas magnitudes son independientes, se puede regular el valor de una sin afectar el de la otra, o más precisamente, su distribución de probabilidad, mientras que, por el contrario, cuando no hay independencia, se debe esperar que al regular el valor de una de las variables se modifique la distribución de probabilidad de la otra.

4.4. La Distribución Multinomial.

Consideremos el siguiente ejemplo: se lanza un dado n veces y se cuenta el número X_1 de veces que se obtiene 1 y el número X_2 de veces que se obtiene 2. Supongamos que el dado es simétrico (es decir, que cada cara tiene probabilidad 1/6 en cada lanzamiento) y que los lanzamientos son independientes, entonces la distribución conjunta de X_1 y X_2 está dada por

$$r_{ij} = P(X_1 = i,\ X_2 = j) \quad \text{con} \quad i + j \le n$$

que se calcula mediante

$$
\begin{aligned}
r_{ij} &= \frac{n!}{i!\ j!\ (n-i-j)!} \left(\frac{1}{6}\right)^i \left(\frac{1}{6}\right)^j \left(\frac{4}{6}\right)^{n-i-j} \\
&= \frac{n!}{i!\ j!\ (n-i-j)!} \frac{4^{n-i-j}}{6^n}.
\end{aligned}
\tag{4.6}
$$

Una manera de obtener esta expresión es la siguiente: el número de resultados ordenados posibles en los n lanzamientos es 6^n, ya que en cada lanzamiento tenemos 6 resultados posibles, y estos 6^n resultados son igualmente probables. Por lo tanto, basta calcular el número de veces que obtenemos i *unos*, j *dos* y $n-i-j$ caras que no son ni *uno* ni *dos*. Para ello procedemos así: elegimos los i lugares en que ponemos los *unos*, lo cual podemos hacer de $\binom{n}{i}$ formas; entre los $n-i$ lugares que nos quedan, elegimos j lugares donde colocar los *dos*, lo cual podemos hacer de $\binom{n-i}{j}$ maneras, y en los $n-i-j$ lugares que nos quedan, colocamos de todas las maneras posibles caras que no son ni 1 ni 2, lo cual podemos hacer de 4^{n-i-j} formas. En total tendremos

$$\binom{n}{i}\binom{n-i}{j}4^{n-i-j} = \frac{n!}{i!\,(n-i)!}\frac{(n-i)!}{j!\,(n-i-j)!}4^{n-i-j}$$

casos favorables. Dividiendo por el número de casos posibles, 6^n, obtenemos (4.6).

Esta distribución es un caso particular de la distribución multinomial. En cuanto a las distribuciones marginales tenemos para $i = 0, 1, \dots, n$,

$$\begin{aligned} p_i = P(X_n = i) &= \sum_{j=0}^{n-i} r_{ij} \\ &= \sum_{j=0}^{n-i} \frac{n!}{i!\,j!\,(n-i-j)!}\frac{4^{n-i-j}}{6^n} \\ &= \frac{n!}{i!(n-i)!}\sum_{j=0}^{n-1}\frac{(n-i)!}{j!(n-i-j)!}\frac{4^{n-i-j}}{6^n} \\ &= \binom{n}{i}\frac{1}{6^n}(1+4)^{n-i} = \binom{n}{i}\left(\frac{1}{6}\right)^i\left(\frac{5}{6}\right)^{n-i} \end{aligned}$$

la función de probabilidad binomial $b(n, 1/6)$, que sabemos que corresponde a la variable aleatoria X_1.

Una situación análoga a la anterior se plantea cuando consideramos el siguiente caso de muestreo. Tenemos una población de N individuos clasificados en 3 grupos de tamaños respectivos $N_1,\ N_2,\ N_3$. Los grupos son disjuntos y $N_1 + N_2 + N_3 = N$.

Si tomamos una muestra de tamaño n al azar y con reposición, y denotamos por $X_1,\ X_2,\ X_3$ respectivamente el número de elementos de la muestra que están en la clase 1, 2 ó 3, tenemos

$$P(X_1 = i, X_2 = j, X_3 = n-i-j) = \frac{n!}{i!\,j!\,(n-i-j)!}p_1^i\,p_2^j\,p_3^{n-i-j}$$

para $i+j \leq n$, donde $p_k = N_k/N,\ (k = 1, 2, 3)$ indica la fracción de la población que está en la i-ésima clase. El cálculo es enteramente similar al del ejemplo anterior.

Si en lugar de tres clases tenemos m, con una notación análoga se tiene

$$P(X_1 = i_1, X_2 = i_2, \dots, X_m = i_m) = \frac{n!}{i_1!\,i_2!\cdots i_m!}p_1^{i_1}p_2^{i_2}\cdots p_m^{i_m}$$

con $i_1 + i_2 + \cdots + i_m = n$. La distribución correspondiente se conoce como la *distribución multinomial*.

Si el muestreo se hace sin reposición, en lugar de la probabilidad anterior tenemos

$$P(X_1 = i_1, X_2 = i_2, \dots, X_m = i_m) = \frac{\binom{N_1}{i_1}\binom{N_2}{i_2}\cdots\binom{N_m}{i_m}}{\binom{N}{n}}$$

donde $N_1 + N_2 + \cdots + N_m = N,\ n_1 + n_2 + \cdots + n_m = n$, para la función de probabilidad conjunta de $X_1,\ X_2, \dots, X_m$. En este caso, las distribuciones marginales son hipergeométricas en lugar de binomiales.

4.5. Densidades.

Sean X e Y dos variables aleatorias y $F(x, y)$ su función de distribución conjunta. Se dice que esta distribución tiene densidad si existe una función $f \geq 0$, conocida como la *densidad de la distribución*, tal que

$$F(x, y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u, v)\,du\,dv \tag{4.7}$$

para todo punto (x, y) del plano. La fórmula anterior indica que $F(x, y)$ es la integral de la función f en el cuadrante que queda "abajo y a la izquierda"del punto (x, y).

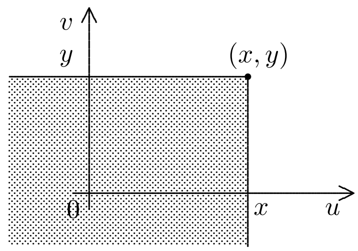


Figura 4.5

Es claro a partir de (4.7) que si F tiene densidad f, cualquiera sea el rectángulo A, se tiene

$$P((X, Y) \in A) = \iint_A f(u, v)\, du\, dv = \int_a^b \int_c^d f(u, v)\, du\, dv$$

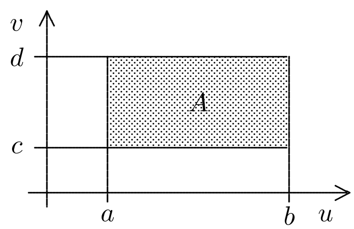


Figura 4.6

Se puede probar también que se verifica

$$P((X, Y) \in B) = \iint_B f(u, v)\, du\, dv \tag{4.8}$$

para cualquier conjunto de Borel B en el plano.

Observación 4.1 Análogamente a lo que ocurre en el caso de distribuciones unidimensionales, para que estas definiciones tengan sentido es necesario precisar cuál es la clase de funciones f (de densidad) que interviene, y qué se entiende por integral doble en (4.7) y (4.8). El contexto adecuado para responder estas cuestiones es el de la integral de Lebesgue. Sin embargo, en los casos que aparecerán en este texto, las funciones f, y también los conjuntos B en (4.8), serán lo suficientemente simples como para que el lector pueda asignar a las integrales dobles el significado que conoce de los cursos de cálculo, y pueda calcularlos por los procedimientos que ha estudiado. En estos casos, ambos tipos de integral coinciden.

Observamos además que, también del mismo modo que en el caso unidimensional, la definición de densidad muestra que la función f que satisface (4.7) no es única. Por ejemplo, si modificamos f en un solo punto (o en un conjunto finito o numerable de puntos) obteniendo una nueva función f_1, la igualdad (4.7) se sigue verificando si en lugar de f ponemos f_1, dado que la modificación realizada no cambia las integrales de las funciones que intervienen. Por lo tanto, f_1 también es una densidad para F.

Si f es una función de densidad, entonces

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(u, v)\, du\, dv = 1 \tag{4.9}$$

Esto es consecuencia de la propiedad 2 de la sección 4.1. Recíprocamente, si se verifica (4.9) y f no toma valores negativos, la función $F(x, y)$ definida por (4.7) es la función de distribución conjunta correspondiente a las variables definidas de la manera siguiente: como espacio de probabilidad tomamos $\Omega = \mathbb{R}^2$, $\mathcal{A} =$ familia de conjuntos de Borel en el plano y P definida de tal modo que

$$P(B) = \iint_B f(u,v)\,du\,dv$$

Las variables X, Y están definidas sobre este espacio de probabilidad $(\Omega, \mathcal{A}, P)$ por

$$X((u,v)) = u$$
$$Y((u,v)) = v$$

Observamos que en este caso los elementos de Ω son los puntos (u,v).

A continuación, sea F la función de distribución conjunta de las variables aleatorias X e Y, y denotemos por G y H las distribuciones (marginales) de X e Y, respectivamente, es decir que

$$G(x) = P(X \le x)$$
$$H(y) = P(Y \le y)$$

Teorema 4.2 *Supongamos que F tiene densidad f. Tenemos las siguientes propiedades que utilizaremos con frecuencia:*

1. *Las funciones*
$$g(u) = \int_{-\infty}^{\infty} f(u,v)\,dv \qquad h(v) = \int_{-\infty}^{\infty} f(u,v)\,du \tag{4.10}$$
son densidades de G y H respectivamente. Se denominan densidades marginales de la distribución conjunta.

2. *X e Y son variables independientes si y sólo si*
$$f(u,v) = g(u)h(v) \tag{4.11}$$
es una función de densidad para F.

3. *Si f es continua en (s,t) y existe la derivada cruzada*
$$\frac{\partial^2 F(s,t)}{\partial x \partial y}$$
en ese punto, entonces
$$\frac{\partial^2 F(s,t)}{\partial x \partial y} = f(s,t). \tag{4.12}$$

Demostración.

1.
$$\begin{aligned} G(x) = P(X \le x) &= \lim_{y \to \infty} F(x,y) \\ &= \lim_{y\to\infty} \int_{-\infty}^{x} \int_{-\infty}^{y} f(u,v)\,du\,dv = \int_{-\infty}^{x} g(u)\,du. \end{aligned}$$

Es decir que la función g definida por (4.10) es una densidad para G. En la misma forma se prueba que h es una densidad para H.

2. Si se verifica (4.11)

$$F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,du\,dv = \int_{-\infty}^{x} g(u)\,du \int_{-\infty}^{y} h(v)\,dv = G(x)\ H(y)$$

y X e Y son variables aleatorias independientes. Recíprocamente, si X e Y son independientes

$$\begin{aligned} F(x,y) = G(x)\ H(y) &= \int_{-\infty}^{x} g(u)\,du \int_{-\infty}^{y} h(v)\,dv \\ &= \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,v)\,du\,dv \end{aligned}$$

y $f(u,v) = g(u)h(v)$ es una densidad para F.

3. Sea A_{hk} el rectángulo de la figura 4.7 $(h \neq 0, k \neq 0)$.

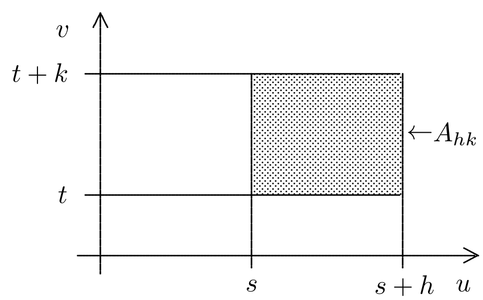


Figura 4.7

Se tiene

$$\begin{aligned} P((X,Y) \in A_{hk}) &= F(s+h,t+k) - F(s+h,t) - F(s,t+k) + F(s,t) \\ &= \iint_{A_{hk}} f(u,v)\,du\,dv = hk(f(s,t) + \varepsilon) \end{aligned}$$

donde $\varepsilon \to 0$ en el último miembro cuando $h \to 0,\ k \to 0$ en virtud de que f es continua en (s,t). Entonces

$$\frac{1}{hk}(F(s+h,t+k) - F(s,t+k) - F(s+h,t) + F(s,t)) = f(s,t) + \varepsilon$$

Si hacemos tender $h \to 0$ y luego $k \to 0$, el segundo miembro tiende a $f(s,t)$, en tanto que el primero

$$\begin{aligned} &\lim_{k\to 0}\frac{1}{k}\lim_{h\to 0}\left(\frac{F(s+h,t+k) - F(s,t+k)}{h} - \frac{F(s+h,t) - F(s,t)}{h}\right) \\ &= \lim_{k\to 0}\frac{1}{k}\left(\frac{\partial F(s,t+k)}{\partial x} - \frac{\partial F(s,t)}{\partial x}\right) \\ &= \frac{\partial^2 F(s,t)}{\partial x \partial y}. \end{aligned}$$

■

4.6. La Distribución Normal Bivariada.

Entre las distribuciones continuas de dos variables, una de las más importantes es la distribución normal bivariada, cuya densidad está dada por

$$\begin{aligned} f(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y(1-\rho^2)^{1/2}} \exp\Bigg\{-\frac{1}{2(1-\rho^2)}\bigg(&\frac{(x-\mu_X)^2}{\sigma_X^2} \\ &- \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\ \sigma_Y} + \frac{(y-\mu_Y)^2}{\sigma_Y^2}\bigg)\Bigg\}. \end{aligned}$$

El significado de los parámetros $\mu_X,\ \mu_Y,\ \sigma_X,\ \sigma_Y$ y ρ será estudiado en el próximo capítulo. Esta función tiene las siguientes propiedades:

1. Un plano paralelo al plano xy que intersecta a $f(x,y)$, la intersecta en una elipse.
2. Un plano perpendicular al plano xy intersecta a $f(x,y)$ en una curva que tiene la forma de la densidad de una variable normal.

En particular, esta última propiedad dice que las densidades marginales, y por lo tanto las distribuciones marginales, son normales. Veamos la demostración de este hecho.

La densidad marginal de X se define por

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\ dy.$$

Hacemos la sustitución $v = (y-\mu_Y)/\sigma_Y$. El exponente de la densidad normal bivariada se transforma en

$$\frac{-1}{2(1-\rho^2)}\left(\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - 2\rho v\left(\frac{x-\mu_X}{\sigma_X}\right) + v^2\right).$$

Ahora completamos el cuadrado en v, que transforma la expresión anterior en

$$\begin{aligned} &\frac{-1}{2(1-\rho^2)}\left(\left(\frac{x-\mu_X}{\sigma_X}\right)^2 + v^2 - 2\rho v\left(\frac{x-\mu_X}{\sigma_X}\right) + \rho^2\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - \rho^2\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right) \\ &\quad = \frac{-1}{2(1-\rho^2)}\left[(1-\rho^2)\left(\frac{x-\mu_X}{\sigma_X}\right)^2 + \left(v - \rho\frac{x-\mu_X}{\sigma_X}\right)^2\right] \\ &\quad = \frac{-1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - \frac{1}{2(1-\rho^2)}\left(v-\rho\frac{x-\mu_X}{\sigma_X}\right)^2. \end{aligned}$$

Por lo tanto tenemos

$$f_X(x) = \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_X(1-\rho^2)^{1/2}} \exp\left\{\frac{-1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - \frac{1}{2(1-\rho^2)}\left(v - \frac{\rho(x-\mu_X)}{\sigma_X}\right)^2\right\} dv$$

Hacemos la sustitución

$$w = \frac{1}{(1-\rho^2)^{1/2}}\left(v - \frac{\rho(x-\mu_X)}{\sigma_X}\right), \quad dw = \frac{dv}{(1-\rho^2)^{1/2}}$$

obteniendo

$$f_X(x) = \frac{1}{2\pi\sigma_X}\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_x}\right)^2 - \frac{1}{2}w^2\right\}\ dw \tag{4.13}$$

$$= \frac{1}{\sqrt{2\pi}\sigma_X}\exp\left(-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)\right)\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}w^2\right)\ dw \tag{4.14}$$

y como el integrando en el último miembro es la densidad normal típica $n(0,1)$, sabemos que la integral vale 1 y por lo tanto

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma_X} \exp\left(-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2\right) = n(x;\mu_X,\sigma_X).$$

4.7. Suma de Variables Aleatorias Independientes.

Sean X e Y dos variables aleatorias discretas con función de probabilidad conjunta

$$r_{ij} = P(X = x_i, Y = y_j) \qquad (i = 1,2,\dots; j = 1,2,\dots)$$

Si queremos calcular la función de probabilidad de la variable aleatoria discreta

$$S = X + Y$$

es inmediato que la misma se obtiene mediante

$$P(S = s_k) = \sum_{x_i+y_j=s_k} r_{ij} \tag{4.15}$$

donde la suma se extiende a todas las parejas de índices (i,j) tales que $x_i + y_j = s_k$. En particular, si X e Y son independientes con funciones de probabilidad respectivas

$$P(X = x_i) = p_i \qquad P(Y = y_j) = q_j$$

la fórmula (4.13) se reduce a

$$P(S = s_k) = \sum_{x_i+y_j=s_k} p_i\, q_j.$$

Consideremos, por ejemplo, el caso de la distribución binomial, ejemplificada mediante el modelo más sencillo de control de calidad, consistente en sucesivas extracciones independientes con reposición y con probabilidad p de extraer un objeto defectuoso en cada ocasión.

Ponemos $X_i = 0$ ó 1 respectivamente, si extraemos un objeto bueno o defectuoso en la i-ésima extracción, y denotamos por d_n el número de defectuosos al cabo de n extracciones. Entonces

$$d_n = d_{n-1} + X_n \tag{4.16}$$

donde las variables aleatorias discretas d_{n-1} y X_n son independientes. La función de probabilidad de d_n se obtiene usando (4.15) y (4.16):

$$p_{n,k} = P(d_n = k) = \sum_{j=0}^{k} P(d_{n-1} = j)P(X_n = k - j).$$

En la suma intervienen a lo sumo dos términos significativos, ya que

$$P(X_n = 1) = p, \quad P(X_n = 0) = 1 - p \quad \text{y} \quad P(X_n \neq 0 \text{ ó } 1) = 0.$$

Se obtiene

$$\begin{aligned}
&p_{1,1} = p \qquad p_{1,0} = 1 - p\\
&p_{n,n} = p\ p_{n-1,n-1} \qquad p_{n,0} = (1-p)p_{n-1,0}\\
&p_{n,k} = p\ p_{n-1,k-1} + (1-p)p_{n-1,k} \quad (k = 1,2,\dots,n-1)
\end{aligned}$$

Utilizando estas igualdades y procediendo por inducción se obtiene la fórmula conocida para la función de probabilidad binomial

$$p_{n,k} = \binom{n}{k} p^k (1-p)^{n-k}, \quad (k = 0,1,\dots,n).$$

4.8. Variables Absolutamente Continuas.

Sean ahora X e Y dos variables aleatorias con densidad conjunta f. La función de distribución de la suma $S = X + Y$ es

$$A(t) = P(S \leq t) = P((X, Y) \in \pi_t)$$

donde π_t es el semiplano sombreado en la figura 4.8.

Por lo tanto

$$A(t) = \iint_{x+y\leq t} f(x,y)\; dx\; dy = \int_{-\infty}^{\infty}\int_{-\infty}^{t-x} f(x,y)\; dy\; dx$$

Efectuando el cambio de variables $y = z - x$ obtenemos

$$\begin{aligned} A(t) &= \int_{-\infty}^{\infty}\int_{-\infty}^{t} f(x, z-x)\, dz\; dx = \int_{-\infty}^{t}\int_{-\infty}^{\infty} f(x, z-x)\; dx\, dz \\ &= \int_{-\infty}^{t} a(z)\, dz \end{aligned}$$

con

$$a(z) = \int_{-\infty}^{\infty} f(x, z-x)\, dz \tag{4.17}$$

que resulta ser una densidad para la variable aleatoria S.

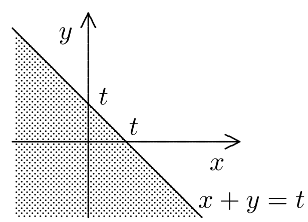


Figura 4.8

En particular, si las variables aleatorias X e Y son independientes, y sus densidades respectivas son f_X y f_Y, se tiene

$$f(x,y) = f_X(x) f_Y(y)$$

y por lo tanto, la variable aleatoria $S = X + Y$ tiene la densidad

$$a(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\; dx = \int_{-\infty}^{\infty} f_X(z-x) f_Y(x)\; dx \tag{4.18}$$

Las integrales que aparecen en el segundo miembro se llaman la *convolución* de las funciones f_X y f_Y.

Como aplicación, probemos que si X e Y tienen ambas distribución normal y son variables aleatorias independientes, entonces $X + Y$ también tiene distribución normal. Sean

$$f_X(x) = \frac{1}{\sqrt{2\pi}\;\sigma_1}\exp\left(-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right); \quad f_X(x) = \frac{1}{\sqrt{2\pi}\;\sigma_2}\exp\left(-\frac{(x-\mu_2)^2}{2\sigma_2^2}\right)$$

las densidades respectivas de X e Y. Es decir, que ambas tienen distribución normal con parámetros $(\mu_1,\ \sigma_1^2)$ y $(\mu_2,\ \sigma_2^2)$ respectivamente. De acuerdo a (4.15), la densidad de la variable aleatoria $S = X + Y$ es

$$\begin{aligned} a(t) &= \int_{-\infty}^{\infty} f_X(x) f_Y(t-x)\; dx \\ &= \int_{-\infty}^{\infty} \frac{1}{2\pi\sigma_1\sigma_2}\exp\left(-\frac{1}{2}\left(\frac{(x-\mu_1)^2}{\sigma_1^2} + \frac{(t-x-\mu_2)^2}{\sigma_2^2}\right)\right)\; dx. \end{aligned}$$

Agrupando adecuadamente, el exponente que aparece en el integrando se puede escribir como

$$-\frac{\sigma_1^2+\sigma_2^2}{2\sigma_1^2\sigma_2^2}\left(\left(x-\frac{\mu_1\sigma_2^2+(t-\mu_2)\sigma_1^2}{\sigma_1^2+\sigma_2^2}\right)^2+\frac{\sigma_1^2\sigma_2^2(t-\mu_1-\mu_2)^2}{(\sigma_1^2+\sigma_2^2)^2}\right)$$

y por lo tanto

$$a(t)=\frac{1}{2\pi\sigma_1\sigma_2}\exp\left(-\frac{1}{2}\frac{(t-\mu_1-\mu_2)^2}{\sigma_1^2+\sigma_2^2}\right)\int_{-\infty}^{\infty}\exp\left(-\frac{\sigma_1^2+\sigma_2^2}{2\sigma_1^2\sigma_2^2}\left(x-\frac{\mu_1\sigma_2^2+(t-\mu_2)\sigma_1^2}{\sigma_1^2+\sigma_2^2}\right)^2\right)dx.$$

En esta integral hacemos el siguiente cambio de variables:

$$u=\frac{\sqrt{\sigma_1^2+\sigma_2^2}}{\sqrt{2}\ \sigma_1\sigma_2}x-\frac{\mu_1\sigma_2^2+(t-\mu_2)\sigma_1^2}{\sigma_1^2+\sigma_2^2}$$

y recordando que

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

resulta

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

donde $\mu=\mu_1+\mu_2$ y $\sigma=\sqrt{\sigma_1^2+\sigma_2^2}$. En consecuencia S tiene distribución normal con parámetros (μ,σ^2).

4.9. Cambio de Variables.

Sean $X_1,\ X_2$ dos variables aleatorias con densidad conjunta $f_X(x_1,\ x_2)$. Consideremos las funciones de dos variables

$$\begin{cases} y_1 &= g_1(x_1,x_2)\\ y_2 &= g_2(x_1,x_2)\end{cases} \tag{4.19}$$

que supondremos continuas en un conjunto abierto $\mathcal{U}$ del plano, fuera del cual la función de densidad f_X se anula. Suponemos además que la función definida por (4.17), que lleva $\mathcal{U}$ en un subconjunto $\mathcal{V}$ del plano, es biyectiva, y por lo tanto invertible, que existen y son continuas las derivadas parciales

$$\frac{\partial g_i}{\partial x_j}\qquad i=1,2;\ j=1,2$$

y que el determinante jacobiano asociado cumple

$$\frac{\partial(y_1,y_2)}{\partial(x_1,x_2)}=\begin{vmatrix}\frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2}\\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2}\end{vmatrix}\neq 0\qquad\text{para todo }(x_1,x_2)\in U.$$

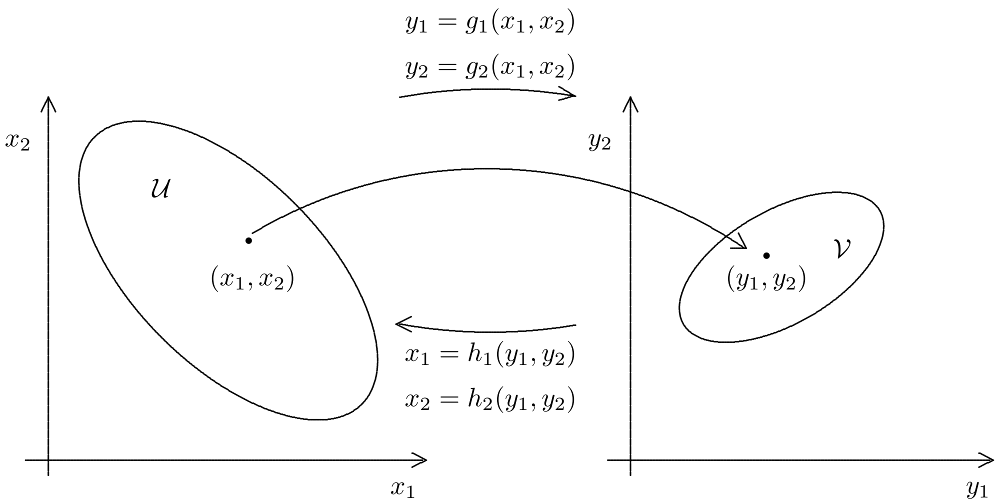


Figura 5-9

En estas condiciones, existe la función inversa, de $\mathcal{V}$ en $\mathcal{U}$, definida por

$$\begin{cases} x_1 & = h_1(y_1, y_2) \\ x_2 & = h_2(y_1, y_2) \end{cases} \tag{4.20}$$

donde h_1 y h_2 también son funciones continuas y con derivadas parciales continuas. Además el jacobiano

$$\left.\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)}\right|_{(y_1,y_2)} = \begin{vmatrix} \frac{\partial h_1}{\partial y_1} & \frac{\partial h_1}{\partial y_2} \\ \frac{\partial h_2}{\partial y_1} & \frac{\partial h_2}{\partial y_2} \end{vmatrix} = \left(\left.\frac{\partial(y_1, y_2)}{\partial(x_1, x_2)}\right|_{(x_1,x_2)} \right)^{-1} \tag{4.21}$$

donde los puntos (x_1, x_2) e (y_1, y_2) que aparecen en (4.21), están vinculados por la correspondencia definida en (4.19) (o equivalentemente en (4.20)), es decir que el jacobiano de la función inversa es el inverso del jacobiano de la función dada, con la precaución de calcular este último en (x_1, x_2) y el primero en (y_1, y_2).

Nos interesa la distribución conjunta de las variables aleatorias $Y_1,\ Y_2$ definidas por

$$\begin{aligned} Y_1 &= g_1(X_1, X_2) \\ Y_2 &= g_2(X_1, X_2) \end{aligned}$$

que resultan de aplicar el cambio de variables (4.17) a la pareja (X_1, X_2). Del mismo modo que para las variables aleatorias unidimensionales, se puede probar que Y_1, Y_2 tiene densidad conjunta $f_Y(y_1, y_2)$ dada por

$$f_Y(y_1, y_2) = \begin{cases} f_x(h_1(y_1, y_2), h_2(y_1, y_2)) \frac{\partial(x_1,x_2)}{\partial(y_1,y_2)} & \text{si } (y_1, y_2) \in \mathcal{V} \\ 0 & \text{si } (y_1, y_2) \notin \mathcal{V} \end{cases} \tag{4.22}$$

La demostración de este hecho se apoya en el teorema de cambio de variables para integrales múltiples. Una indicación sobre la manera de proceder es la siguiente: para simplificar la notación pongamos

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad \mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \quad \mathbf{Y} = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$$

y en lugar de (4.19),

$$\mathbf{y} = \mathbf{g}(\mathbf{x}) = \mathbf{g}(x_1, x_2) = \begin{pmatrix} g_1(x_1, x_2) \\ g_2(x_1, x_2) \end{pmatrix}.$$

Entonces, si A es un rectángulo,

$$\begin{aligned}
P((Y_1,Y_2)\in A) &= P(\mathbf{g}(\mathbf{X})\in A\cap\mathcal{V}) = P(\mathbf{X}\in\mathbf{g}^{-1}(A\cap\mathcal{V}))\\
&= \iint_{\mathbf{g}^{-1}(A\cap\mathcal{V})} f_X(x_1,x_2)\ dx_1\ dx_2\\
&= \iint_{A\cap\mathcal{V}} f_X(h_1(y_1,y_2),h_2(y_1,y_2))\left|\frac{\partial(x_1,x_2)}{\partial(y_1,y_2)}\right| dy_1\ dy_2\\
&= \iint_A f_Y(y_1,y_2)\ dy_1\ dy_2
\end{aligned}$$

definiendo f_Y como en (4.22). La penúltima igualdad se apoya en la fórmula de cambio de variables para integrales múltiples.

Ejemplo

Como aplicación de (4.22), supongamos que (X_1,X_2) es una variable aleatoria bidimensional con densidad conjunta $f_X(x_1,x_2)$ y que queremos hallar la densidad de la variable aleatoria $Z=X_1X_2$.

► Consideremos para ello la función de $\mathbb{R}^2$ en $\mathbb{R}^2$ definida por

$$\begin{aligned}
w &= w(x_1,x_2) = x_1\\
z &= x(x_1,x_2) = x_1x_2
\end{aligned}$$

entonces

$$\begin{aligned}
x_1 &= x_1(w,z) = w\\
x_2 &= x_2(w,z) = z/w
\end{aligned}$$

y

$$\frac{\partial(x_1,x_2)}{\partial(w,z)} = \begin{vmatrix} 1 & 0\\ \frac{-z}{w^2} & \frac{1}{w}. \end{vmatrix} = \frac{1}{w}$$

Eligiendo $\mathcal{U}=\{(x_1,x_2):\ x_1\neq 0\}$ y $\mathcal{V}=\{(w,z):\ w\neq 0\}$, estamos en las condiciones anteriores. Por lo tanto, la densidad conjunta de W y Z es

$$k(w,z) = f_X(w,\frac{z}{w})\frac{1}{|w|}$$

y la densidad de Z no es otra cosa que la densidad marginal correspondiente a la segunda variable, a partir de la densidad conjunta k, es decir

$$h(z) = \int_{-\infty}^{\infty} k(w,z)\ dw = \int_{-\infty}^{\infty} f_X(w,\frac{z}{w})\frac{1}{|w|}\ dw. \tag{4.23}$$

En particular, si X_1 y X_2 son independientes y con densidades respectivas f y g, teniendo en cuenta que en este caso $f_X(x_1,x_2)=f(x_1)g(x_2)$, y sustituyendo en (4.21) resulta

$$h(z) = \int_{-\infty}^{\infty} f(w)g(\frac{z}{w})\frac{1}{|w|}\ dw.$$

Un método similar al desarrollado aquí para hallar la densidad del producto de dos variables aleatorias puede emplearse para calcular la densidad de otras funciones. ◄

4.10. Funciones de Variables Aleatorias Independientes

Para terminar este capítulo observemos finalmente que si X e Y son variables aleatorias independientes, y g y h son funciones tales que la preimagen de un intervalo es un conjunto de Borel, entonces las variables aleatorias

$$X_1 = g(X) \qquad Y_1 = h(Y)$$

también son independientes. Esto se apoya en el hecho de que siendo $I,\ J$ intervalos

$$\{X_1 \in I\} = \{g(X) \in I\} = \{X \in g^{-1}(I)\} = X^{-1}(g^{-1}(I))$$

y del mismo modo

$$\{Y_1 \in J\} = Y^{-1}(h^{-1}(J)).$$

Como $g^{-1}(I),\ h^{-1}(J)$ son conjuntos de Borel y las variables X e Y son independientes, los eventos $\{X_1 \in I\},\ \{Y_1 \in J\}$ resultan ser independientes.

4.11. Ejemplos.

1. Se extrae una muestra de tamaño dos con reposición de una bolsa que contiene dos bolas blancas, una negra y dos rojas. Definimos las variables aleatorias X_1 y X_2 de la siguiente manera: para $k = 1, 2,\ X_k = 1$ ó 0 según si la bola obtenida en la k-ésima extracción es blanca o no lo es.

   a. Describa la función de probabilidad conjunta de estas variables.

   b. Describa las funciones marginales de probabilidad.

   c. ¿Son independientes estas variables aleatorias?

   d. ¿Qué sucede si el muestreo se realiza sin reposición?

► a. Para el caso de muestreo con reposición tenemos

$$r_{00} = \frac{3}{5}\frac{3}{5} = \frac{9}{25} \qquad r_{01} = \frac{3}{5}\frac{2}{5} = \frac{6}{25}$$
$$r_{10} = \frac{2}{5}\frac{3}{5} = \frac{6}{25} \qquad r_{11} = \frac{2}{5}\frac{2}{5} = \frac{4}{25}$$

b. Las funciones marginales de probabilidad son

$$p_0 = P(X_1 = 0) = \frac{3}{5} \qquad p_1 = P(X_1 = 1) = \frac{2}{5}$$
$$q_0 = P(X_2 = 0) = \frac{3}{5} \qquad q_1 = P(X_2 = 1) = \frac{2}{5}$$

c. Es fácil verificar que para todo i, j se tiene

$$r_{ij} = p_i\ q_j$$

de modo que las variables X_1 y X_2 son independientes.

d. Si el muestreo es sin reposición, la función de probabilidad conjunta es

$$r_{00} = \frac{3}{5}\frac{2}{4} = \frac{6}{20} \qquad r_{01} = \frac{3}{5}\frac{2}{4} = \frac{6}{20}$$
$$r_{10} = \frac{2}{5}\frac{3}{4} = \frac{6}{20} \qquad r_{11} = \frac{2}{5}\frac{1}{4} = \frac{2}{20}$$

Las funciones marginales de probabilidad son

$$p_0 = P(X_1 = 0) = \frac{3}{5} \qquad p_1 = P(X_1 = 1) = \frac{2}{5}$$
$$q_0 = P(X_2 = 0) = \frac{3}{5} \qquad q_1 = P(X_2 = 1) = \frac{2}{5}$$

Las variables no son independientes en este caso ya que, por ejemplo, $r_{00} \neq p_0\, q_0$. ◀

2. Supongamos que las variables (X, Y) pueden tomar los valores $(0, 0)$; $(1, 1)$; $(-1, 1)$; $(1, -1)$ y $(-1, -1)$ cada uno con probabilidad $1/5$. Determine si estas variables son independientes.

▶ La función de probabilidad conjunta está resumida en la siguiente tabla

		X		
		-1	0	1
	-1	$1/5$	0	$1/5$
Y	0	0	$1/5$	0
	1	$1/5$	0	$1/5$

Las funciones marginales de probabilidad son

$$P(X = -1) = p_{-1} = r_{-1,-1} + r_{-1,0} + r_{-1,1} = \frac{1}{5} + \frac{1}{5} = \frac{2}{5}$$

y similarmente

$$P(X = 0) = p_0 = \frac{1}{5}, \qquad P(X = 1) = p_1 = \frac{2}{5}.$$

Se verifica fácilmente que Y tiene la misma función de probabilidad. Ahora bien,

$$P(X = 0,\ Y = -1) = 0 \neq \frac{1}{5} \times \frac{2}{5} = P(X = 0)P(Y = -1)$$

y las variables no son independientes. ◀

3. Si X e Y son variables aleatorias independientes con la misma función de distribución F, ¿Cuál es la función de distribución $G(z)$ de la variable aleatoria $Z = \text{máx}(X, Y)$?

▶

$$\begin{aligned} G(z) &= P(Z \leq z) = P(\text{máx}(X, Y) \leq z) \\ &= P(X \leq z,\ Y \leq z) = P(X \leq z)P(Y \leq z) \\ &= F^2(z). \end{aligned}$$

◀

4. Sean X e Y variables aleatorias independientes con función de probabilidad uniforme en $\{1, 2, \dots, N\}$ (es decir, $P(X = i) = P(Y = i) = 1/N,\ i = 1, \dots, N$). Calcule la función de probabilidad de $X + Y$.

▶ Es evidente que para $j < 2$ ó $j > 2N$ tenemos

$$P(X + Y = j) = 0.$$

Si $2 \leq j \leq N$

$$P(X + Y = j) = \sum_{i=1}^{j-1} P(X = i,\ Y = j - i) = \sum_{i=1}^{j-1} \frac{1}{N^2} = \frac{j-1}{N^2}$$

mientras que para $N+1 \leq j \leq 2N$, definiendo $i = j - N$ tenemos

$$\begin{aligned} P(X+Y=j) = P(X+Y=N+i) &= \sum_{k=i}^{N} P(X=k,\ Y=N+i-k) \\ &= \frac{N-i+1}{N^2} = \frac{2N-j+1}{N^2}. \end{aligned}$$

Resumiendo

$$P(X+Y=j) = \begin{cases} (j-1)/N^2 & \text{para } 2 \leq j \leq N \\ (2N-j+1)/N^2 & \text{para } N+1 \leq j \leq 2N \\ 0 & \text{en cualquier otro caso} \end{cases}$$

◀

5. * En dos lugares de una habitación se mide la intensidad del ruido. Sean X e Y las variables aleatorias que representan la intensidad del ruido en estos puntos y supongamos que la distribución conjunta de estas variables es continua y tiene densidad

$$f(x,y) = \begin{cases} xy \exp\left(-\frac{1}{2}(x^2+y^2)\right) & \text{si } x>0,\ y>0. \\ 0 & \text{en caso contrario.} \end{cases}$$

Calcular las densidades marginales de X e Y y las probabilidades $P(X \leq 1, Y \leq 1)$ y $P(X+Y \leq 1)$. ¿Son independientes estas variables aleatorias?

► Las densidades marginales son

$$\begin{aligned} g(x) &= \int_{-\infty}^{\infty} f(x,y)dy \\ &= \int_0^{\infty} x\, y \exp\left(-\frac{1}{2}(x^2+y^2)\right) dy \\ &= x \exp\left(-\frac{x^2}{2}\right) \int_0^{\infty} y \exp\left(-\frac{y^2}{2}\right) dy \\ &= x \exp\left(-\frac{x^2}{2}\right) \qquad \text{para } x > 0, \end{aligned}$$

y similarmente

$$h(y) = y \exp\left(-\frac{y^2}{2}\right) \qquad \text{para } y > 0.$$

Observamos que $f(x,y) = g(x)h(y)$, y por lo tanto las variables son independientes. Por otro lado,

$$\begin{aligned} P(X \leq 1,\ Y \leq 1) &= \int_{-\infty}^{1}\int_{-\infty}^{1} f(u,v)du\,dv \\ &= \int_0^1 u\,e^{-u^2/2}du \quad \int_0^1 v\,e^{-v^2/2}dv \\ &= 0{,}1548 \end{aligned}$$

$$\begin{aligned} P(X + Y \leq 1) &= \iint_{x+y\leq 1} f(u,v)du\,dv \\ &= \int_0^1 \int_0^{1-y} f(u,v)du\,dv \\ &= \int_0^1 u\,e^{-u^2/2}\left(\int_0^{1-y} v\,e^{-v^2/2}dv\right)du \\ &= 0{,}2433 \end{aligned}$$

◀

6. * Se escogen al azar e independientemente dos puntos X e Y sobre un segmento de longitud L. ¿Cuál es la probabilidad de que la distancia entre los puntos no sea mayor que ℓ?

► Supongamos que X se escoge en un intervalo $[0, L]$ sobre el eje x, mientras que Y se escoge en un intervalo $[0, L]$ sobre el eje y. La probabilidad que deseamos calcular es la de que un punto (X, Y) escogido al azar en el cuadrado $0 \leq x \leq L,\ 0 \leq y \leq L$ caiga en la región B acotada por las rectas $y = x - \ell,\ y = x + \ell$ (ver Figura 4.10).

Por hipótesis, las variables aleatorias X e Y son independientes y ambas tienen distribución uniforme en $[0, L]$, es decir, ambas tienen como densidad a la función

$$g(x) = \frac{1}{L}, \qquad 0 \leq x \leq L.$$

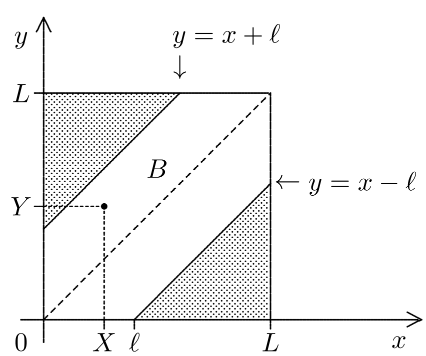


Figura 4.10

Por lo tanto, la densidad conjunta de las variables independientes X e Y es el producto de estas funciones:

$$f(x,y) = \frac{1}{L^2} \qquad 0 \leq x \leq L,\ 0 \leq y \leq L.$$

La probabilidad de que el punto (X, Y) caiga en la región B está dada por

$$P((X,Y) \in B) = \iint_B \frac{1}{L^2}\, dx\, dy = \frac{\text{Area de } B}{L^2}$$

y el área de B es $L^2 - 2\frac{1}{2}(L-\ell)^2 = 2L\ell - \ell^2$ de modo que

$$P((X,Y) \in B) = \frac{2L\ell - \ell^2}{L^2}.$$

◀

7. * *El Problema de la Aguja de Buffon.* Sobre una superficie plana, rayada con rectas paralelas que están a una distancia L entre si, se lanza al azar una aguja de longitud ℓ (una aguja quiere decir un segmento de recta). ¿Cuál es la probabilidad de que la aguja intersecte alguna de las rectas paralelas?

▶ Sea θ el ángulo entre la aguja y la dirección de las rectas paralelas, y sea X la distancia entre el punto inferior de la aguja y la recta más próxima que pase por arriba de él (ver figura 4.11).

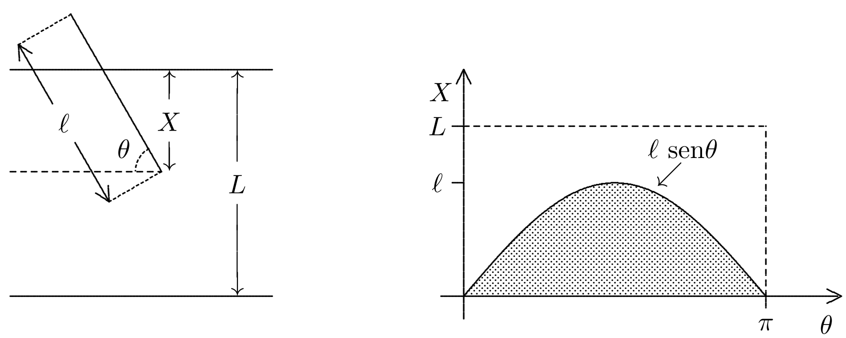


Figura 4.11

Por las condiciones del ejercicio sabemos que la variable θ tiene distribución uniforme en $[0, \pi]$, mientras que la variable X tiene distribución uniforme en $[0, L]$. Por lo tanto, si suponemos que las variables aleatorias θ y X son independientes, encontramos que su densidad conjunta es

$$f_{X,\theta}(u,v) = \frac{1}{\pi L}, \quad 0 \le u \le L,\ 0, \le v, \le \pi.$$

La aguja intersecta una de las rectas si y sólo si

$$X \le \ell \text{ sen } \theta$$

es decir, si y sólo si el punto (θ, X) cae en la región B, donde B es la parte del rectángulo $0 \le x \le \pi,\ 0 \le y \le L$ que está entre el eje x y la curva $y = \ell$ sen x (ver figura 4.11). Por lo tanto

$$\begin{aligned} P((\theta, X) \in B) &= \iint_B \frac{1}{\pi L} du\, dv = \frac{1}{\pi L} \int_0^\pi \int_0^{\ell \text{ sen } u} dv\, du \\ &= \frac{1}{\pi L} \int_0^\pi \ell \text{ sen } u\, du = \frac{2\ell}{\pi L}. \end{aligned}$$

◀

8. * Dé un ejemplo de dos variables aleatorias que no sean independientes pero tales que sus cuadrados si lo sean.

► Consideremos las variables (X,Y) cuya densidad conjunta vale 2 en los cuadrados $0 < x < 0,5,\ 0 < y < 0,5$ y $-0,5 < x < 0,\ -0,5 < y < 0$ (ver figura 4.12).

Para ver que estas variables no son independientes consideremos las probabilidades

$$P\left(0 < X < 0,5,\ 0 < Y < 0,5\right) = \frac{1}{2}$$

$$P\left(0 < X < 0,5\right) = \int_{0}^{0,5} \int_{-\infty}^{\infty} f(x,y)\ dy\ dx = \int_{0}^{0,5} \int_{0}^{0,5} 2\ dy\ dx = \frac{1}{2}$$

y de forma similar $P(0 < Y < 0,5) = 1/2$. Se ve que las variables no son independientes.

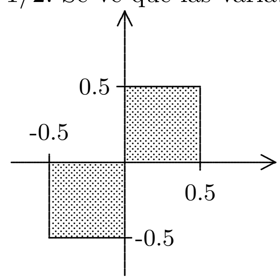


Figura 4.12

Sea ahora $G(x,y)$ la función de distribución conjunta de las variables X^2 e Y^2. Calculemos esta función: supongamos que $0 < x < 0,5,\ 0 < y < 0,5$

$$\begin{aligned} G(x,y) &= P(X^2 \leq x,\ Y^2 \leq y) = P(-\sqrt{x} \leq X \leq \sqrt{x},\ -\sqrt{y} \leq Y \leq \sqrt{y}) \\ &= \int_{-\sqrt{x}}^{0} \int_{-\sqrt{y}}^{\sqrt{y}} f(u,v)\, du\, dv + \int_{0}^{\sqrt{x}} \int_{-\sqrt{y}}^{\sqrt{y}} f(u,v)\, du\, dv \\ &= \int_{-\sqrt{x}}^{0} \int_{-\sqrt{y}}^{0} 2\, du\, dv + \int_{0}^{\sqrt{x}} \int_{0}^{\sqrt{y}} 2\, du\, dv \\ &= 4\sqrt{xy} \quad \text{para } 0 < x < 0,5,\ 0 < y < 0,5. \end{aligned}$$

De manera similar se puede obtener que

$$G(x,y) = \begin{cases} 0 & \text{para } x \leq 0 \text{ ó } y \leq 0 \\ 2\sqrt{x} & \text{para } 0 < x < 0,5,\ y \geq 0,5 \\ 2\sqrt{y} & \text{para } 0 < y < 0,5,\ x \geq 0,5 \\ 1 & \text{para } x \geq 0,5,\ y \geq 0,5 \end{cases}$$

Las funciones marginales de distribución correspondientes son

$$G_{X^2}(x) = \begin{cases} 0 & \text{para } x \leq 0 \\ 2\sqrt{x} & \text{para } 0 < x \leq 0,5 \\ 1 & \text{para } x \geq 0,5 \end{cases}$$

$$G_{Y^2}(y) = \begin{cases} 0 & \text{para } y \leq 0 \\ 2\sqrt{y} & \text{para } 0 < y \leq 0,5 \\ 1 & \text{para } y \geq 0,5 \end{cases}$$

y vemos que

$$G(x,y) = G_{X^2}(x) G_{Y^2}(y)$$

de modo que X^2 e Y^2 son independientes. ◄

9. * Sean X e Y variables aleatorias independientes con distribución uniforme en $(0,1)$. Halle la densidad de $X+Y$.

▶ La densidad común de las variables es

$$f_X(x) = f_Y(x) = \begin{cases} 1 & \text{si } 0 < x < 1 \\ 0 & \text{en cualquier otro caso.} \end{cases}$$

Por los resultados de la sección 4.8 sabemos que la densidad de la suma es

$$a(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\, dx.$$

Si $0 \leq z \leq 1$, entonces

$$a(z) = \int_0^z f_X(x) f_Y(z-x)\, dx = \int_0^z dx = z.$$

Si $1 < z \leq 2$, entonces

$$a(z) = \int_{z-1}^{1} dx = 2 - z.$$

Si $z < 0$ ó $z > 2$, el integrando es cero.

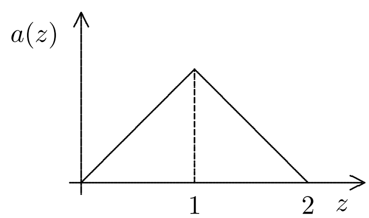


Figura 4.13

Resumiendo, obtenemos una distribución triangular con densidad

$$a(z) = \begin{cases} z & \text{si } 0 \leq z \leq 1 \\ 2-z & \text{si } 1 < z \leq 2 \\ 0 & \text{en cualquier otro caso.} \end{cases}$$

La gráfica de la densidad está dada en la figura 4.13. ◀

10. * Sean X e Y variables aleatorias con densidad conjunta f. Obtenga la densidad de la variable aleatoria $Z = Y/X$ en el caso general y en el caso de variables independientes.

▶ Sea B_z el conjunto

$$B_z = \{(x,y) : \frac{y}{x} \leq z\} = \{(x,y) : x < 0,\ y \geq xz\} \cup \{(x,y) : x > 0,\ y \leq xz\}.$$

Sea G la distribución de Z, tenemos

$$G(z) = \iint_{B_z} f(x,y)\, dx\, dy = \int_{-\infty}^{0} \int_{xz}^{\infty} f(x,y)\, dy\, dx + \int_0^{\infty} \int_{-\infty}^{xz} f(x,y)\, dy\, dx.$$

Hacemos el cambio de variables $y = xv$ obteniendo

$$\begin{aligned} G(z) &= \int_{-\infty}^{0} \int_{z}^{-\infty} x f(x, xv)\, dv\, dx + \int_{0}^{\infty} \int_{-\infty}^{z} x f(x, xv)\, dv\, dx \\ &= \int_{-\infty}^{0} \int_{-\infty}^{z} (-x) f(x, xv)\, dv\, dx + \int_{0}^{\infty} \int_{-\infty}^{z} x f(x, xv)\, dv\, dx \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{z} |x| f(x, xv)\, dv\, dx. \end{aligned}$$

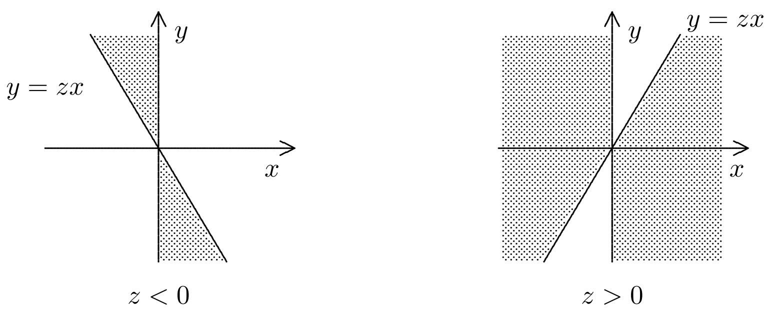


Figura 4.14

Intercambiando el order de integración, obtenemos

$$G(z) = \int_{-\infty}^{z} \int_{-\infty}^{\infty} |x| f(x, xv)\, dx\, dv$$

y por lo tanto, la densidad g de G está dada por

$$g(z) = \int_{-\infty}^{\infty} |x| f(x, xz)\, dx \qquad -\infty < z < \infty.$$

Si X e Y son independientes, entonces la densidad es

$$g(z) = \int_{-\infty}^{\infty} |x| f_X(x) f_Y(xz)\, dx.$$

◀

11. * Los tiempos de espera X e Y de dos clientes que entran a un banco en distintos instantes son variables aleatorias que supondremos independientes y con la misma distribución de densidad

$$f(x) = \begin{cases} e^{-x} & \text{si } x \geq 0 \\ 0 & \text{en otro caso} \end{cases}$$

Halle la densidad conjunta de la suma de sus tiempos de espera $U = X + Y$ y de la fracción de este tiempo que el primer cliente permanece en espera, $V = X/(X + Y)$. Halle también las densidades marginales de U y V y muestre que son independientes.

▶ Tenemos que $X = UV$ e $Y = U - X = U - UV$. El jacobiano de la transformación es

$$\frac{\partial(X, Y)}{\partial(U, V)} = \begin{vmatrix} V & U \\ 1 - V & -U \end{vmatrix} = -UV - U + UV = -U.$$

Sea $f(x,y)$ la densidad conjunta de X e Y, dada por

$$f(x,y) = \begin{cases} exp(-x-y) & \text{para } x \geq 0, \text{ y } y \geq 0 \\ 0 & \text{en otro caso.} \end{cases}$$

En base a lo visto en la sección 4.9, la densidad conjunta de U y V es

$$g(u,v) = f(uv, u-uv)|u| = |u|e^{-u} \quad \text{para } u, v \in \mathcal{R}.$$

Falta determinar cuál es la región $\mathcal{R}$, que es imagen del primer cuadrante bajo la transformación

$$u = x+y \qquad v = \frac{x}{x+y}$$

Observamos que u puede tomar cualquier valor positivo mientras que $0 \leq \frac{x}{x+y} \leq 1$, de modo que $\mathcal{R}$ es la región que se muestra en la figura 4.15.

Resumiendo

$$g(u,v) = \begin{cases} ue^{-u} & \text{para } (u,v) \in \mathcal{R} \\ 0 & \text{en otro caso.} \end{cases}$$

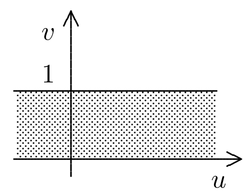


Figura 4.15

Las densidades marginales de U y V son

$$g_U(u) = \int_0^1 g(u,v)\,dv = \int_0^1 ue^{-u}dv = ue^{-u} \quad 0 \leq u < 0$$
$$g_V(v) = \int_0^\infty g(u,v)du = \int_0^\infty ue^{-u}du = 1 \quad 0 \leq v \leq 1$$

y concluimos que U y V son variables aleatorias independientes. ◀

Ejercicios

1. Dada la función de probabilidad conjunta definida por

$$r_{ij} = C(i+j)$$

en los puntos $(1,1)$; $(2,1)$; $(2,1)$ y $(3,1)$, donde C es una constante, determine en valor de C y obtenga la función de probabilidad marginal correspondiente a la primera variable.

2. Considere un grupo de cartas que consiste de J, Q, K y A de las cuatro pintas. Se extraen dos cartas del grupo sin reposición y llamamos X e Y al número de diamantes y corazones obtenidos, respectivamente. Obtenga la función de probabilidad conjunta y la función marginal correspondiente a X.

3. Una caja tiene 6 bolas numeradas del 1 al 6. Las bolas numeradas 1 y 2 son rojas mientras que las otras son blancas. Extraemos dos bolas al azar de la caja y sean X, Y las variables aleatorias que representan el número de bolas rojas y el número de bolas pares en la muestra, respectivamente. Halle la distribuciones de X e Y y su distribución conjunta. Determine si estas variables son independientes.

4. Una caja contiene ocho bolas numeradas del 1 al 8. Las primeras cuatro son rojas y las otras blancas. Seleccionamos dos bolas al azar de la caja y definimos las siguientes variables: X es el número de bolas blancas en la muestra, Y es el número de bolas pares y Z el número de bolas en la muestra cuyo número es menor que 6. Halle la distribución conjunta de las variables (X,Y); (X,Z); (Y,Z) y (X,Y,Z). Estudie la independencia de estas variables.

5. Considere dos variables aleatorias X e Y con distribución conjunta discreta definida por la siguiente tabla para la función de probabilidad conjunta, donde $h = 1/60$.

		X_1		
		0	1	2
	0	h	$2h$	$3h$
X_2	1	$2h$	$4h$	$6h$
	2	$3h$	$6h$	$9h$
	3	$4h$	$8h$	$12h$

Calcule

a. $P(X \leq 1,\ Y \leq 1)$	b. $P(X+Y \leq 1)$
c. $P(X+Y > 2)$	d. $P(X < 2Y)$
e. $P(X > 1)$	f. $P(X = Y)$
g. $P(X \geq Y \mid Y > 1)$	h. $P(X^2 + Y^2 \leq 1)$

6. Repita el ejercicio anterior para la siguiente función de probabilidad conjunta (de nuevo $h = 1/60$).

		X_1		
		0	1	2
	0	h	$6h$	$6h$
X_2	1	$2h$	$8h$	$9h$
	2	$3h$	$2h$	$12h$
	3	$4h$	$4h$	$3h$

7. Lanzamos un dado dos veces. Sea X el resultado del primer lanzamiento, Y el mayor de los resultados de los dos lanzamientos. Halle la distribución conjunta y las distribuciones marginales de estas variables. Determine si son independientes.

8. Lanzamos una moneda tres veces y definimos las siguientes variables aleatorias: X es el número de águilas, Y es la longitud de la mayor sucesión de águilas en la muestra. Por ejemplo $Y(A,S,A) = 1,\ Y(A,A,S) = 2$. Halle la distribución conjunta, las distribuciones marginales y determine si estas variables son independientes.

9. Lanzamos una moneda cuatro veces y definimos las siguientes variables aleatorias: X vale 1 si hay más águilas que soles y vale 0 si esto no es cierto. Por otro lado, Y representa la longitud de la mayor sucesión de águlas en la muestra. Hallar la distribución conjunta y las marginales. Determine si estas variables son independientes.

10. Consideremos un experimento que tiene resultados $\omega_1, \omega_2, \dots, \omega_8$ con probabilidades correspondientes $0.1; 0.1; 0.2; 0.2; 0.1; 0.1; 0.1; 0.1$. Sea X, Y y Z las variables aleatorias definidas por la siguiente tabla

	X	Y	Z
ω_1	1	1	1
ω_2	2	2	2
ω_3	1	3	3
ω_4	2	1	4
ω_5	1	2	1
ω_6	2	3	2
ω_7	1	1	3
ω_8	2	2	4

Halle las distribuciones de probabilidad de X, Y y Z y las distribuciones conjuntas de (X, Y); (X, Z); (Y, Z) y (X, Y, Z).

11. Considere dos eventos A y B tales que $P(A) = 1/4$, $P(B|A) = 1/2$ y $P(A|B) = 1/4$. Definimos las variables X e Y por $X = \mathbf{1}_A$, $Y = \mathbf{1}_B$, donde $\mathbf{1}_E(x)$ vale 1 si $x \in E$ y vale 0 si $x \notin E$. Diga si las siguientes proposiciones son ciertas o falsas.

    a. Las variables aleatorias X e Y son independientes.

    b. $P(X^2 + Y^2 = 1) = 1/4$.

    c. $P(XY = X^2Y^2) = 1$.

    d. La variable aleatoria X tiene distribución uniforme en el intervalo $(0, 1)$.

    e. Las variables X e Y tienen la misma distribución.

12. Considere las variables aleatorias X e Y con densidad conjunta dada por

$$f(x, y) = \begin{cases} \frac{1}{4} & \text{si } \ 0 \le x \le 2, \ 0 \le y \le 2 \\ 0 & \text{en cualquier otro caso.} \end{cases}$$

    Calcule

    a. $P(X \le 1, \ Y \le 1)$ b. $P(X + Y \le 1)$
    c. $P(X + Y > 2)$ d. $P(X < 2Y)$
    e. $P(X > 1)$ f. $P(X = Y)$
    g. $P(Y > 1, \ X \le 1)$ h. $P(X \ge Y | Y > 1)$

13. Repita el ejercicio anterior para la densidad

$$f(x, y) = \begin{cases} \exp(-(x + y)) & \text{si } x \ge 0, \ y \ge 0 \\ 0 & \text{en cualquier otro caso} \end{cases}$$

14. Sean X e Y variables aleatorias independientes con distribución uniforme en $[0, 1]$. Calcule

    a. $P(X + Y < 0,5)$ b. $P(X - Y < 0,5)$
    c. $P(XY < 0,5)$ d. $P(X/Y < 0,5)$
    e. $P(X^2 < 0,5)$ f. $P(X^2 + Y^2 < 0,5)$
    g. $P(e^{-X} < 0,5)$ h. $P(\cos \pi Y < 0,5)$.

15. Dada la densidad $f(x, y) = 8xy,\ 0 < x < 1,\ 0 < y < 1$, calcule

$$P(X < 0{,}5,\ Y < 0{,}5), \qquad P(X < 0{,}5), \qquad P(Y < 0{,}5).$$

A partir de estos cálculos, ¿qué se puede decir sobre la independencia de X e Y?

16. Dada la densidad $f(x, y) = xy\ e^{-(x+y)},\ x > 0,\ y > 0$, calcule $P(X > 1,\ Y > 1)$. ¿Son independientes estas variables aleatorias?