Capítulo 5

Distribuciones Conjuntas e Independencia

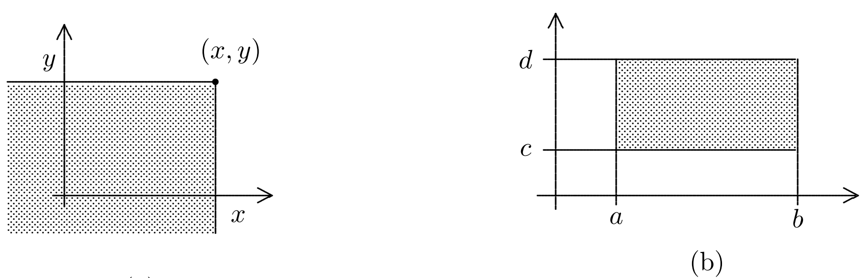
5.1. Distribución Conjunta de Dos Variables Aleatorias.

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 5.1 (a).



(a)

(b)

Figura 5.1

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 5.1 (b))

$$\begin{aligned} P(a < X \le b,\ c < Y \le d) &= P(X \le b,\ Y \le d) - P(X \le b,\ Y \le c) \\ &\quad - P(X \le a,\ Y \le d) + P(X \le a,\ Y \le c) \\ &= F(b,d) - F(b,c) - F(a,d) + F(a,c) \end{aligned} \tag{5.1}$$

5.1.1. Propiedades

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

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

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

y como la función F 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 5.2).

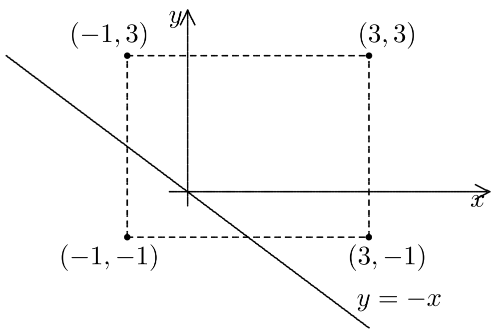


Figura 5.2

La función es creciente, continua por la derecha y satisface la propiedad 2. Sin embargo, si aplicamos la fórmula (5.1) 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 (5.1) no sea negativo para ninguna colección de números $a < b,\ c < d$.

Teorema 5.1 *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 $a < b,\ c < d$,*

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

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.

5.2. Variables Aleatorias Independientes.

Definición 5.1 Se dice que las variables X e Y son *independientes* si cualesquiera 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

$$P(a < X \leq b,\ c < Y \leq d) = P(a < X \leq b)P(c < Y \leq d). \tag{5.2}$$

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

Es fácil ver que la condición (5.2) es equivalente a la condición

$$F_{X,Y}(x, y) = F_X(x)F_Y(y) \quad \text{para todo } x, y \in \mathbb{R}. \tag{5.3}$$

En efecto, si se cumple (5.2) basta poner $b = x,\ d = y$ y hacer tender $a \to -\infty,\ c \to -\infty$, para obtener (5.3). Recíprocamente, si se cumple (5.3), 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 (5.3) implica (5.2), cualesquiera sean $a,\ b,\ c$ y d.

Las relaciones (5.2) y (5.3) 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 (5.2), y semirectas cerradas a la derecha, en el caso de (5.3). Es posible probar, aunque no lo haremos en este texto, que (5.3), o equivalentemente (5.2), 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.

5.3. 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

$$p_i = \sum_j r_{ij} \quad (i = 1, 2, \dots) \tag{5.4}$$

$$q_j = \sum_i r_{ij} \quad (j = 1, 2, \dots) \tag{5.5}$$

ya que, por ejemplo,

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

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

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

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:

$$\begin{array}{cc|cccc|c} & & & & Y & & \\ & & y_1 & y_2 & \cdots\cdots & y_m & \\ \hline & x_1 & r_{11} & r_{12} & \cdots\cdots & r_{1m} & p_1 \\ & x_2 & r_{21} & r_{22} & \cdots\cdots & r_{2m} & p_2 \\ X & \vdots & \vdots & \vdots & & \vdots & \vdots \\ & x_n & r_{n1} & r_{n2} & \cdots\cdots & r_{nm} & p_n \\ \hline & & q_1 & q_2 & \cdots\cdots & q_m & \end{array}$$

Tabla 5.1

En la última columna aparecen $p_1,\ p_2, \dots, p_n$, que son las sumas respectivas de las filas (condición (5.4)) y en la última fila $q_1, q_2,\ \dots, q_m$, que son las sumas respectivas de las columnas (condición (5.5)).

Una situación de este tipo aparece en diversos problemas de aplicación. Supongamos que en el proceso de producción de un objeto 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 5.4 hemos tomado $n = 5,\ m = 4$.

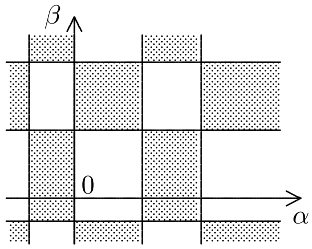


Figura 5.4

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 (5.4) y (5.5).

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

		Y				
		1	2	$\cdots\cdots$	m	
	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
		Q_1	Q_2	$\cdots\cdots$	Q_m	

Tabla 5.2

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 5.1, 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.

5.4. 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

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

que se calcula mediante

$$\begin{aligned} r_{ij} &= \frac{n!}{i!\ j!\ (n-i-j)!} \left(\frac{1}{6}\right)^i \left(\frac{1}{6}\right)^j \left(\frac{4}{6}\right)^{n-i-j} \\ &= \frac{n!}{i!\ j!\ (n-i-j)!} \frac{4^{n-i-j}}{6^n}. \end{aligned} \tag{5.6}$$

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 (5.6).

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

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

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 $N_1 + N_2 + \cdots + N_m = N$, $i_1 + i_2 + \cdots + i_m = n$, 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.

5.5. Funciones de Variables Aleatorias Independientes

Observamos que si X e Y son variables aleatorias independientes, y g y h son funciones medibles (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.

Esto es cierto si, por ejemplo, las funciones g y h son continuas o monótonas o tienen una cantidad numerable de discontinuidades.

5.6. 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

$$P(S = s_k) = \sum_{x_i + y_j = s_k} r_{ij} \tag{5.7}$$

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 (5.7) se reduce a

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

Ejemplo.

Consideremos 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

$$D_n = D_{n-1} + X_n \tag{5.8}$$

donde las variables aleatorias discretas D_{n-1} y X_n son independientes. La función de probabilidad de D_n se obtiene usando (5.7) y (5.8):

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

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

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

◀

En el caso especial en el cual X, Y son variables con valores i, j, donde $i, j \in \mathbb{Z}$. Entonces la ecuación (5.7) es

$$P(S = k) = \sum_{i+j=k} r_{i,j} = \sum_i r_{i,k-i}.$$

En particular, si X, Y son independientes,

$$P(S = k) = \sum_i p_i q_{k-i},$$

y si las variables X, Y son no-negativas tenemos

$$P(S = k) = \sum_{i=0}^{k} p_i q_{k-i}.$$

Ejemplos.

1. **Suma Binomial.** Si X tiene distribución binomial $b(n,p)$ e Y tiene distribución binomial $b(m,p)$, entonces la distribución de $S = X + Y$ es

$$\begin{aligned}
P(S = k) &= \sum_{i=0}^{k} \binom{n}{i} p^i q^{n-i} \binom{m}{k-i} p^{k-i} q^{m-k+i} \\
&= p^k q^{m+n-k} \sum_{i=0}^{k} \binom{n}{i} \binom{m}{k-i} \\
&= \binom{m+n}{k} p^k q^{m+n-k},
\end{aligned}$$

de modo que $S = X + Y$ tiene distribución binomial $b(m + n, p)$.

2. **Suma Poisson.** Si X, Y son variables de Poisson independientes, con parámetros λ y μ respectivamente, entonces

$$
\begin{aligned}
P(S = k) &= \sum_{i=0}^{k} \frac{\lambda^i}{i!} e^{-\lambda} \frac{\mu^{k-i}}{(k-i)!} e^{-\mu} = \frac{e^{-(\lambda+\mu)}}{k!} \sum_{i=0}^{k} \binom{k}{i} \lambda^i \mu^{k-i} \\
&= \frac{e^{-(\lambda+\mu)}}{k!} (\lambda + \mu)^k,
\end{aligned}
$$

de modo que S tiene distribución de Poisson con parámetro $\lambda + \mu$.

3. **Suma Geométrica.** Sean X, Y variables independientes con distribución geométrica de parámero p. Entonces para $k \geq 2$,

$$
\begin{aligned}
P(S = k) &= \sum_{i=1}^{k-1} q^{i-1} p q^{k-i-1} p \\
&= (k-1) q^{k-2} p^2,
\end{aligned}
$$

que es una distribución binomial negativa.

5.7. Sucesiones de Variables de Bernoulli

Vamos a considerar en esta sección sucesiones (X_n) de variables aleatorias de Bernoulli independientes, todas con la misma probabilidad de éxito p. Estas variables satisfacen la propiedad de que para cualquier colección de índices $i_1, i_2, \dots, i_k$ y cualquier vector $\mathbf{j} = (j_1, j_2, \dots, j_k)$ de componentes iguales a 0 ó 1, se tiene

$$
P(X_{i_1} = j_1, X_{i_2} = j_2, \dots, X_{i_k} = j_k) = P(X_{i_1} = j_1) P(X_{i_2} = j_2) \cdots P(X_{i_k} = j_k).
$$

Como todas las variables tienen la misma probabilidad de éxito, si en el vector $\mathbf{j}$ hay m componente iguales a 1 y $k - m$ iguales a 0, esta probabilidad es igual a

$$
p^m (1-p)^{k-m}.
$$

Hemos estudiado algunas propiedades de estas sucesiones anteriormente. Por ejemplo, la distribución binomial corresponde al número de éxitos en una colección de n variables de este tipo: si $S_n = \sum_1^n X_i$ entonces S_n tiene distribución binomial $b(n, p)$. La distribución geométrica representa el tiempo de espera hasta el primer éxito mientras que la distribución binomial negativa representa el tiempo de espera hasta conseguir k éxitos. Veamos otras propiedades de estas sucesiones.

5.7.1. Intervalos entre Exitos

Llamemos T_k a la longitud del intervalo entre el $(k-1)$-ésimo y el k-ésimo éxito en la sucesión de Bernoulli. Vamos a incluir siempre al último éxito, pero no al anterior en nuestra contabilidad. Como ejemplo mostramos los valores de los T_k iniciales para un resultado particular:

$$
\underbrace{0001}_{T_1}\underbrace{001}_{T_2}\underbrace{\ 1\ }_{T_3}\underbrace{00001}_{T_4}\dots
$$

Ya conocemos la distribución de T_1, porque esta variable representa el tiempo de espera hasta el primer éxito, y tiene por lo tanto una distribución geométrica. Ahora bien, una vez que ocurre el primer éxito, la independencia implica que lo que ocurre después es como si iniciáramos el proceso de nuevo, y por lo tanto T_2 tiene la misma distribución que T_1 y además son variables independientes. Veamos esto en detalle.

Para i y j en $\mathbb{N}$ el evento $\{T_1 = i, T_2 = j\}$ ocurre si hay $(i-1)$ fracasos inicialmente, luego un éxito, luego $(j-1)$ fracasos más y finalmente un segundo éxito:

$$\underbrace{00\dots0}_{(i-1)}1\underbrace{00\dots0}_{(j-1)}1$$

y la probabilidad de que esto ocurra es

$$q^{i-1}pq^{j-1}p = P(T_1 = i)P(T_2 = j),$$

de modo que estas variables son independientes. De manera similar se puede demostrar que cualquier colección de variables $T_{i_1}, \dots, T_{i_k}$ son independientes.

Existe una relación fundamental entre estas variables T_k y las variables S_n. Supongamos que para cierta sucesión de ensayos de Bernoulli el k-ésimo éxito ha ocurrido antes de o en el n-ésimo experimento: $T_k \leq n$. Para que esto suceda, el número de éxitos en los primeros n ensayos debe ser mayor o igual a k: $S_n \geq k$. Recíprocamente, si $S_n \geq k$ entonces $T_k \leq n$.

Por otro lado, si $T_k = n$, el k-ésimo éxito ocurre en el experimento n. Esto quiere decir que en los primeros $n-1$ ensayos ocurrieron $k-1$ éxitos y el n-ésimo ensayo resultó en éxito: $S_{n-1} = k-1$ y $X_n = 1$. La relación recíproca también es cierta y es fácil de verificar. Resumimos estos resultados en la siguiente proposición.

Proposición 5.1 *Si $k \leq n$ tenemos*
a. $T_k \leq n$ si y sólo si $S_n \geq k$.
b. $T_k = n$ si y sólo si $S_{n-1} = k-1$, $X_n = 1$.

5.8. Paseo al Azar

Sea (X_n) una sucesión de variables de Bernoulli con probabilidad de éxito p. Definimos las variables $Y_n = 2X_n - 1$, que son independientes y toman valores 1 y -1 con probabilidades p y $1-p$. Si consideramos una serie de juegos en los que tenemos probabilidad p de ganar y $q = 1-p$ de perder y apostamos una unidad en cada uno, podemos representar el resultados de cada juego por una sucesión de variables de este tipo.

En esta situación la suma de las variables $Y_1 + Y_2 + \cdots + Y_n = S_n$ representa la evolución de la fortuna del jugador al cabo de n juegos, es decir, es la cantidad que el jugador gana o pierde luego de haber jugado n juegos. Si el capital inicial del jugador es A entonces su fortuna al cabo de n juegos es $S'_n = A + S_n$. La sucesión $(S_n)_{n\geq1}$ (o $(S'_n)_{n\geq1}$ si la fortuna inicial es distinta de 0) se conoce como el paseo al azar simple y en el caso $p = q = 1/2$ es el paseo simple simétrico. Podemos representarlo gráficamente como en la siguiente figura:

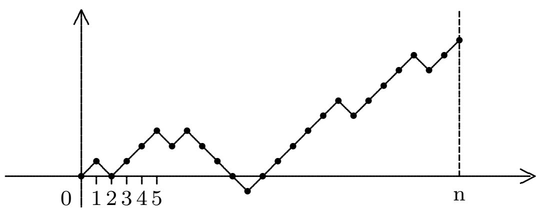


Figura 5.5

Es sencillo calcular la distribución de probabilidad de S_n.

Proposición 5.2 *Para $0 \leq k \leq n$, $P(S_n = k) = 0$ si k y n no tienen la misma paridad. En caso contrario tenemos*

$$P(S_n = k) = \binom{n}{\frac{n+k}{2}}p^{(n+k)/2}q^{(n-k)/2}.$$

Demostración. Sea s el número de $+1's$ en los n juegos y t el número de $-1's$. Claramente $s+t=n$. Para que $S_n = k$ es necesario que $s - t = k$. Resolviendo este sistema de ecuaciones tenemos

$$s = \frac{n+k}{2}, \qquad t = \frac{n-k}{2}$$

que son enteros si y sólo si n y k tienen la misma paridad. Si esto es cierto, tenemos $s = (n+k)/2$ éxitos en n ensayos de Bernoulli, que corresponde a una distribución binomial:

$$P(S_n = k) = \binom{n}{\frac{n+k}{2}} p^{(n+k)/2} q^{(n-k)/2}.$$

■

En el capítulo 3 estudiamos un problema asociado al paseo al azar: el problema de la ruina del jugador. En el próximo capítulo analizaremos el problema de la duración promedio del juego.

5.9. Muestreo Secuencial

Ya hemos visto en capítulos anteriores el muestreo sin reposición, en el cual tenemos una población de n individuos que muestreamos sucesivamente pero sin reponer a la población los individuos que vamos observando. En esta situación los sucesivos resultados que obtenemos no son independientes entre sí, porque con cada selección el espacio muestral cambia. Para poblaciones grandes el efecto es menor pero con poblaciones pequeñas puede ser considerable.

Nos queremos enfocar en la dependencia entre los sucesivos resultados. Supongamos que tenemos una caja con r bola rojas y b bolas blancas con $r + b = N$. Seleccionamos una bola al azar de la caja, vemos su color y la colocamos aparte. Repetimos este procedimiento n veces y para $1 \leq i \leq n$ definimos las variables

$$X_i = \begin{cases} 1 & \text{si la } i\text{-ésima bola es roja,} \\ 0 & \text{si la } i\text{-ésima bola es blanca.} \end{cases}$$

Consideremos inicialmente el resultado de la primera extracción. Es inmediato que

$$p_0 = P(X_1 = 0) = \frac{b}{N}, \qquad p_1 = P(X_1 = 1) = \frac{r}{N}.$$

Veamos ahora la distribución de la segunda variable, usando la Ley de la Probabilidad Total,

$$\begin{aligned} P(X_2 = 0) &= P(X_2 = 0|X_1 = 0)P(X_1 = 0) + P(X_2 = 0|X_1 = 1)P(X_1 = 1) \\ &= \frac{b-1}{N-1} \times \frac{b}{N} + \frac{b}{N-1} \times \frac{r}{N} \\ &= \frac{(b-1)b + br}{N(N-1)} \\ &= \frac{b(N-1)}{N(N-1)} = \frac{b}{N} \end{aligned}$$

y de manera similar tenemos que

$$P(X_2 = 1) = \frac{r}{N},$$

de modo que las variables X_1 y X_2 tienen la misma distribución. Este hecho resulta inesperado a primera vista, pues la situación cambia una vez que hacemos la primera extracción, así que esperaríamos que X_2 tuviese una distribución distinta a la de X_1. Sin embargo hay que observar que no estamos tomando en cuenta el resultado de la primera extracción, es decir, estamos calculando la distribución de X_2 sin conocer el resultado de X_1, y en estas condiciones, a falta de mayor información, el resultado nos dice que la segunda extracción tiene la misma distribución de probabilidad que bajo las condiciones iniciales.

Es posible demostrar, aunque es un poco más trabajoso, que esto es cierto para cualquiera de las variables X_i, $1 \leq i \leq n$: Todas tienen la misma distribución.

Para ver que estas variables no son independientes, calculemos la distribución conjunta de X_1 y X_2.

$$P(X_1 = 0, X_2 = 0) = P(X_2 = 0|X_1 = 0)P(X_1 = 0) = \frac{b(b-1)}{N(N-1)}.$$

De manera similar tenemos

$$P(X_1 = 0, X_2 = 1) = \frac{br}{N(N-1)}, \quad P(X_1 = 1, X_2 = 0) = \frac{br}{N(N-1)}, \quad P(X_1 = 1, X_2 = 1) = \frac{r(r-1)}{N(N-1)},$$

y vemos que la función de probabilidad conjunta no es el producto de las funciones de distribución individuales de modo que las variables no son independientes.

Usando de nuevo la ley de la probabilidad total no es difícil ver que si $i_1, i_2, \dots, i_n$ toman valores 0 ó 1, y hay j unos (y $n-j$ ceros) con $0 \leq j \leq r$, $0 \leq n-j \leq b$, entonces

$$P(X_1 = i_1, X_2 = i_2, \dots, X_n = i_n) = \frac{\binom{N-n}{r-j}}{\binom{N}{n}}$$

5.10. 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. Si realizamos n ensayos de Bernoulli y U y V son el número de éxitos y fracasos respectivamente, sabemos que estas variables son dependientes. Supongamos que n es una variable de Poisson con parámetro λ y realizamos N ensayos de Bernoulli, de modo que el número de ensayos es aleatorio. Ahora obtenemos X éxitos e Y fracasos. Demuestre que X e Y son independientes.

▶ Tenemos

$$\begin{aligned} P(X=i, Y=j) &= P(X=i, Y=j, N=i+j) \\ &= P(X=i, Y=j | N=i+j) P(N=i+j) \\ &= \frac{(i+j)!}{i!j!} p^i q^j e^{-\lambda} \frac{\lambda^{i+j}}{(i+j)!} \\ &= \frac{(\lambda p)^i}{i!} e^{-\lambda p} \frac{(\lambda q)^j}{j!} e^{-\lambda q} \end{aligned}$$

Esta función de probabilidad se factoriza para todos los valores de X e Y, de modo que estas variables son independientes y tienen distribución de Poisson con parámetros λp y λq, respectivamente.

◀

Ejercicios

1. Dada la función de probabilidad conjunta definida por

$$p_{ij} = C(i+j) \tag{5.9}$$

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. Sean X, Y variables aleatorias con valores en $\{1, 2, \dots, n\}$ y con función de probabilidad conjunta dada por (5.9). Halle el valor de C y las distribuciones marginales.

3. La función $p_{i,j}$ está dada por $p_{i,j} = C\alpha^i\beta^j$ para $i, j \in \mathbb{N}$ y $0 < \alpha, \beta < 1$. Halle el valor de C para que $p_{i,j}$ sea una función de probabilidad.

4. ¿Es $p_{i,j} = (0.5)^{i+j}$ para $i, j \in \{0, 1, 2, \dots\}$ una función de probabilidad? Si la respuesta es positiva, calcule $P\{1 \leq i \leq 3, j \geq 2\}$.

5. La función $p_{i,j}$ está dada por

$$p_{i,j} = C\binom{10}{j}\left(\frac{1}{2}\right)^{10}.$$

Halle el valor de C y determine las funciones de probabilidad marginales.

6. Sea X una variable aleatoria de Bernoulli con probabilidad de éxito p y sean $Y = 1 - X$, $Z = XY$. Halle la distribución conjunta de (X, Y); (X, Z) y (Y, Z) ¿Es independiente alguna de estas parejas?

7. Sean X, Y variables aleatorias con función de probabilidad conjunta
$$r_{ij} = C\frac{i + j}{i!j!}\theta^{i+j},$$
para $i, j \geq 0$ donde $\theta > 0$ es una constante. Halle C, las distribuciones marginales de X e Y y $P(X + Y = k)$. ¿Son independientes estas variables aleatorias?

8. Sea $\Omega = \{\omega_1, \omega_2, \omega_3\}$ y P la distribución uniforme en Ω (todos los puntos tienen igual probabilidad). Definimos las variables X, Y y Z de la siguiente manera: $X(\omega_1) = Y(\omega_2) = Z(\omega_3) = 1$, $X(\omega_2) = Y(\omega_3) = Z(\omega_1) = 2$, $X(\omega_3) = Y(\omega_1) = Z(\omega_2) = 3$. Demuestre que estas tres variables tienen la misma función de probabilidad. Halle las funciones de probabilidad de $X + Y$, $Y + Z$ y $X + Z$.

9. 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.

10. 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.

11. 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.

12. Sea X el número de aces en una mano de poker e Y el número de reinas. Halle la función de probabilidad conjunta para estas variables y sus funciones de probabilidad marginales. Determine si son independientes.

13. Lanzamos una moneda tres veces. Sea X el número de águilas en los dos primeros lanzamientos e Y el número de águilas en el tercer lanzamiento. Halle la distribución conjunta de X, Y, la distribución de $Z = X + Y$ y la de $W = X - Y$.

14. Sacamos cinco cartas con reposición de un paquete de 52 cartas. ¿Cuál es la probabilidad de obtener dos diamantes y un trébol? ¿Cuál es la probabilidad de obtener dos ases, dos reinas y un 10?

15. Un componente electrónico puede fallar de cuatro maneras distintas, y las probabilidades respectivas son $p_1 = 0.2; p_2 = 0.15; p_3 = 0.25; p_4 = 0.4$. Si examinamos 10 componentes, ¿Cuál es la probabilidad de que haya exactamente tres fallas de tipo 1, dos de tipo 2, dos de tipo 3 y y tres de tipo 4?

16. Las probabilidades de llenar una declaración de impuestos correctamente, con un error que favorezca al fisco, con un error que favorezca al declarante o con ambos tipos de errores son, respectivamente, $0.6; 0.2; 0.15$ y 0.05. Calcule la probabilidad de que entre 10 declaraciones de impuestos 5 estén correctas, 3 tengan errores a favor del declarante, 1 tenga un error a favor del fisco y una tenga ambos tipos de errores.

17. A través de un estudio se ha determinado que al llegar a cierto cruce, 50 % de los vehículos continua de frente, 30 % da vuelta a la derecha y el resto da vuelta a la izquierda. Calcule la probabilidad de que de los siguientes cinco automóviles, uno de vuelta a la izquierda, dos a la derecha y los otros dos sigan de frente. Calcule la probabilidad de que entre los siguientes siete automóviles a lo sumo dos den vuelta a la izquierda.

18. Un taller mecánico hace afinaciones para vehículos de 4, 6 y 8 cilindros. Hay dos tipos de afinación en cada caso, según los puntos que se revisen y los cambios que se efectúen. El precio P es $100 \times k$, donde k es el número de cilindros, para el afinamiento normal (N) y $200 \times k$ para el afinamiento especial (E). La siguiente tabla muestra la función de probabilidad conjunta para estas variables.

		K		
		4	6	8
P	N	0.3	0.2	0.1
	E	0.05	0.05	0.1

Halle las funciones de probabilidad marginales para las variables K y P. Si Z es el costo total del afinamiento para un carro que va al taller, halle su función de probabilidad.

19. 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. $P(X \geq Y | Y > 1)$ h. $P(X^2 + Y^2 \leq 1)$

20. 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	$4h$	$4h$	$3h$

21. Las funciones de probabilidad marginales de las variables aleatorias discretas X e Y están dadas en la siguiente tabla:

				X			
		1	2	3	4	5	
	1						5/14
	2						4/14
Y	3						3/14
	4						2/14
	5						1/14
		1/14	5/14	4/14	2/14	2/14	1

Para i, j entre 1 y 5 la probabilidad conjunta $P(X = i, Y = j)$ sólo puede tomar los valores 0 y 1/14. Determine la función de probabilidad conjunta para X e Y.

22. Si A es un conjunto medible hemos definido la función $\mathbf{1}_A(\omega)$ como la función que vale 1 si $\omega \in A$ y 0 si no. Demuestre que dos eventos A y B son independientes si y sólo si $\mathbf{1}_A$ y $\mathbf{1}_B$ son variables aleatorias independientes.

23. Demuestre que dos variables aleatorias X, Y son independientes si y sólo si para cualesquiera $x, y \in \mathbb{R}$,
$$P(X > x, Y > y) = P(X > x)P(Y > y).$$

24. Una caja contiene 5 bolas numeradas del 1 al 5. Sacamos dos bolas sucesivamente y sin reposición. Sea X el número en la primera bola e Y el número en la segunda. (a) Determine la distribución conjunta de X, Y. (b) Halle las distribuciones marginales y determine si las variables son independientes. (c) Calcule $P(X < Y)$.

25. 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.

26. 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.

27. 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 águilas en la muestra. Hallar la distribución conjunta y las marginales. Determine si estas variables son independientes.

28. Una función de probabilidad conjunta está dada por $p_{0,0} = a$, $p_{0,1} = b$, $p_{1,0} = c$, $p_{1,1} = d$, donde necesariamente $a + b + c + d = 1$. Demuestre que una condición necesaria para que haya independencia es que $ad = bc$.

29. Sean X, Y variables aleatorias independientes e idénticamente distribuidas que toman valores $1, 2, 3$ con igual probabilidad. Definimos $Z = X - Y$, $W = X + Y$. Halle la distribución conjunta de estas variables y sus distribuciones marginales. ¿Son independientes?

30. Sean X, Y variables aleatorias con valores en $\{1, 2, \dots, n\}$ y función de probabilidad conjunta $p_{ij} = 1/n^2$. Halle las funciones de probabilidad marginales y determine si las variables son independientes. Calcule $P(X > Y)$ y $P(X = Y)$.

31. En un vivero se siembran n semillas. Cada una de ellas germina de manera independiente con probabilidad α. Las X plantas germinadas son transplantadas a macetas y sobreviven de manera independiente con probabilidad β. sea Y el número de plantas que sobreviven. Halle la función de distribución conjunta de X, Y y las marginales.

32. 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

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

33. 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.

34. Sean X, Y variables aleatorias independientes, X con distribución geométrica de parámetro p e Y con distribución geométrica de parámetro r. Demuestre que $Z = \min\{X, Y\}$ tiene distribución geométrica de parámetro $p + r - pr$.

35. Sean X, Y variables aleatorias independientes, X con distribución geométrica de parámetro p e Y con distribución geométrica de parámetro r. Sean $U = \min\{X, Y\}, V = \text{máx}\{X, Y\}, W = V - U$. Halle la distribución conjunta de U y V y la de U y W. Demuestre que estas dos últimas variables son independientes.

36. Sean X, Y variables aleatorias independientes, ambas con distribución geométrica de parámetro p. Demuestre que
$$P(X = k|X + Y = n) = \frac{1}{n+1}$$
para $k = 0, 1, \dots, n$.

37. Sean $X_1, X_2, \dots, X_n$ variables aleatorias independientes discretas con distribución uniforme en el conjunto $\{1, 2, \dots, k\}$. Halle la función de distribución conjunta de la variables M_n y m_n definidas por
$$M_n = \text{máx}\{X_1, X_2, \dots, X_n\}; \qquad m_n = \min\{X_1, X_2, \dots, X_n\}$$

38. Sean X, Y variables aleatorias discretas con función de probabilidad conjunta
$$p_{ij} = \beta^{i+j+2}, \quad i, j = 0, 1, \dots$$
¿Para cuáles valores de β es esta una función de probabilidad? Halle las funciones de probabilidad marginales y determine si las variables son independientes.

39. Responda la pregunta anterior con la restricción $0 \leq i < j < \infty$ sobre los valores posibles de X e Y.

40. Sea $X_1, \dots, X_n$ una colección de variables aleatorias con la propiedad de que para todo r, $1 \leq r \leq n-1$, la colección $\{X_1, \dots, X_r\}$ es independiente de X_{r+1}. Demuestre que las variable X_i, $1 \leq i \leq n$ son independientes.

41. Sean X, Y variables aleatorias discretas con función de probabilidad conjunta
$$p_{ij} = \frac{C}{(i+j-1)(i+j)(i+j+1)}, \qquad i, j \in \mathbb{N}.$$
Calcule C, halle las funciones de probabilidad marginales y determine si las variables son independientes.

42. Para las variables del ejercicio anterior halle la función de probabilidad de $U = X + Y$ y $V = X - Y$.

43. Sea X e Y variables aleatorias independientes, X con distribución uniforme en $\{1, 2, \dots, n\}$, Y con distribución uniforme en $\{1, 2, \dots, m\}$. Halle la distribución de probabilidad de $Z = X + Y$.

44. Sean X e Y variables aleatorias independientes con distribución de Poisson de parámetros λ y μ. Demuestre que la distribución condicional de X dado que $X+Y=n$ es binomial y halle sus parámetros.

45. Sean X e Y variables aleatorias independientes con distribución binomial, ambas de parámetros n y p y sea $Z=X+Y$. Demuestre que la distribución condicional de X dado que $Z=k$ es hipergeométrica.

46. Sea X,Y,Z variables aleatorias discretas con la probabilidad de que sus valores son distintos con probabilidad 1. Sea $a=P(X>Y, b=P(Y>Z), c=P(Z>X)$.

    a) Demuestre que $\min\{a,b,c\}\leq 2/3$ y dé un ejemplo donde esta cota se alcance.

    b) Demuestre que si X,Y,Z son independientes e idénticamente distribuidas, entonces $a=b=c=1/2$.

47. Sea X,Y variables aleatorias discretas. Demuestre que son independientes si y sólo si su función de probabilidad conjunta $P(X=i,Y=j)=r_{ij}$ se puede factorizar como el producto $s_i t_j$ de una funciíon de i por una función de j.

48. Sean $X_j, 1\leq j\leq n$ variables aleatorias independientes simétricas respecto a 0, es decir, X_j y $-X_j$ tienen la misma distribución. Demuestre que para todo x, $P(S_n\geq x)=P(S_n\leq -x)$, con $S_n=X_1+X_2+\cdots+X_n$.

    ¿Es cierta en general la conclusión si no suponemos independencia?

49. Considere un paseo al azar simétrico simple S con $S_0=0$. Sea $T=\min\{n\geq 1: S_n=0\}$ el instante del primer regreso al origen. Demuestre que

$$P(T=2n)=\frac{1}{2n-1}\binom{2n}{n}2^{-2n}.$$

50. Considere un paseo al azar simétrico simple S con $S_0=0$. Definimos $U=\min\{0\leq j\leq n: S_{2j}=S_{2n}\}$ el instante de la primera visita a la posición que ocupa en el instante $2n$. Demuestre que

$$P(U=2k)=P(S_{2k}=0)P(S_{2n-2k}=0),$$

para $0\leq k\leq n$.