# *Clase 6: Variables aleatorias discretas*

*Matías Carrasco*

*5 de agosto de 2019*

**Resumen** Definimos variables aleatorias, y nos centramos en las variables discretas. Veremos varias formas en las que uno puede describir la distribución de una o más variables discretas.

## Índice



## *¿Qué es una variable aleatoria?*

El cuadro de abajo muestra los datos recolectados en un experimento biológico en el cual se midió la altura (en cm) de $n = 60$ especímenes de *Onobrychis viciifolia*, una planta herbácea (i.e. un yuyo), luego de cultivarlos durante seis meses[1]:

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 21 | 21 | 23 | 22 | 23 | 29 | 24 | 21 | 18 | 23 | 19 | 18 |
| 20 | 24 | 20 | 20 | 19 | 19 | 22 | 21 | 18 | 20 | 23 | 17 |
| 20 | 25 | 23 | 21 | 14 | 18 | 29 | 28 | 28 | 14 | 28 | 26 |
| 22 | 22 | 22 | 29 | 19 | 26 | 16 | 17 | 23 | 18 | 25 | 22 |
| 20 | 22 | 18 | 32 | 26 | 21 | 20 | 27 | 20 | 19 | 19 | 18 |

[1] Los datos son reales, ver V. Rousson *Statistique appliqué aux sciences de la vie*, Capítulo 1.

Estos números parecen elegidos al azar, pero ¿de dónde viene ese azar? ¿Qué determina finalmente la altura de cada planta? Seguramente, aspectos genéticos, condiciones del suelo, el clima, la biodiversidad del lugar en donde se realizó la plantación, y un sin número de otros factores. Es prácticamente imposible determinar un espacio muestral cuyos elementos correspondan a las diferentes “historias” que cada planta pueda tener. Peor aún, si por un milagro de astucia logramos describir el espacio muestral, ¿cómo determinamos las probabilidades? Seguramente los diferentes factores influyen de forma particular en la altura final de la planta.

¿Es imposible hacer un modelo para este tipo de experimentos? Increíblemente no, y la gran invención que lo permite es el concepto de *variable aleatoria*.

La idea es la siguiente: no importa demasiado determinar el espacio muestral $\Omega$, lo que realmente nos interesa en estos casos es conocer la variable altura, una función $X : \Omega \to \mathbb{R}$ que a cada elemento $\omega \in \Omega$ le asigna una altura $X(\omega) \in \mathbb{R}$. La variable $X$ se dice aleatoria porque su valor depende del resultado aleatorio del experimento.

Veamos más detenidamente lo que esta idea quiere decir. Cada elemento $\omega$ del espacio muestral representa una “historia” posible para una planta; contiene la información sobre la genética, el suelo, el clima, etc, que “sufrirá” la planta. La función $X$ es una “oráculo”, que dada esa información, nos dice la altura que finalmente tendrá la planta.

La clave está en que no precisamos conocer la probabilidad de todos los eventos de $\Omega$, simplemente la probabilidad de los valores posibles de $X$. Esto se conoce como la *distribución* de $X$.

Los datos del cuadro nos proveen algo de información sobre la distribución de $X$. Por ejemplo, sabemos que $X$ puede tomar los valores 14, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 32. Pero no solo eso, tenemos una idea aproximada de con qué probabilidad toma estos valores. Esto es así pues podemos contar cuántas veces toma cada valor sobre el total de datos. Recordar que esto es la frecuencia relativa de cada valor. Ver la Tabla 1.

| Valor | Frec. | Frec. relativa |
|---|---|---|
| 14 | 2 | 0.0333 |
| 16 | 1 | 0.0167 |
| 17 | 2 | 0.0333 |
| 18 | 7 | 0.1167 |
| 19 | 6 | 0.1000 |
| 20 | 8 | 0.1333 |
| 21 | 6 | 0.1000 |
| 22 | 7 | 0.1167 |
| 23 | 6 | 0.1000 |
| 24 | 2 | 0.0333 |
| 25 | 2 | 0.0333 |
| 26 | 3 | 0.0500 |
| 27 | 1 | 0.0167 |
| 28 | 3 | 0.0500 |
| 29 | 3 | 0.0500 |
| 32 | 1 | 0.0167 |

Tabla 1: Frecuencias relativas de las alturas de las plantas. La segunda columna es la frecuencia absoluta.

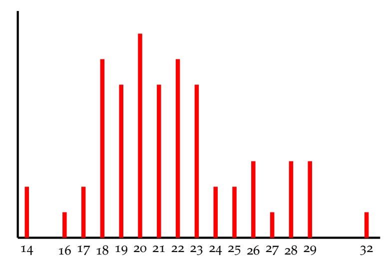



Figura 1: Distribución de plantas según su altura en cm.

En este ejemplo estamos considerando a la altura de las plantas como una variable *discreta*, ya que hemos redondeado a valores enteros de cm. Esto no tiene por qué ser así, podríamos haber utilizado medidas más precisas que tomen cualquier valor real (con varias cifras decimales). En este segundo caso diríamos que la variable es *continua*. Más adelante veremos definiciones precisas.

Sin un modelo mejor, toda la información queda resumida en la tabla anterior. Podemos visualizar mejor la distribución de $X$ si graficamos las frecuencias relativas, como hicimos en la Figura 1.

Estos son los datos observados. Un modelo *teórico* del experimento sería una fórmula o algoritmo que nos permita calcular la distribución de $X$ (la probabilidad de cada valor posible) a partir de ciertos principios. Esas fórmulas teóricas pueden depender de varios parámetros, y contrastando los datos con el modelo, podemos ver cuáles son los parámetros que mejor lo *ajustan* a la realidad.

De este modo, un estadístico al ver los datos de la tabla y la gráfica de frecuencias relativas, propondría una (a veces complicada) fórmula para la distribución de $X$:

$$P(X = k) = p(k;\theta) \tag{1}$$

en donde $\theta$ es un parámetro a determinar que aparece en la fórmula[2].

[2] Podrían ser varios parámetros $\theta_1, \ldots, \theta_r$. De hecho, el estadístico podría proponer una "formula" con un número infinito de parámetros, o modelos aún más complicados que no queremos detallar ahora.

Notar que la fórmula (1) pretende ser un modelo que explique las "regularidades" o "patrones" que presentan los datos a través de sus frecuencia relativas. Muchas veces los parámetros tienen interpretaciones como cantidades físicas importantes sobre la población, como puede ser la altura promedio. Una vez que se ha definido el modelo, se ha elegido el parámetro, el investigador puede sacar conclusiones, comparar varias poblaciones de plantas, etc. Todo esto sin saber cuál es el espacio muestral.

Los modelos de variables aleatorias no requieren conocer el espacio muestral, ni todas las probabilidades en él definidas. Solamente necesitan especificar la distribución de $X$. En esto radica su utilidad.

**Definición de variable aleatoria**

Una variable aleatoria es una función

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

que a cada elemento del espacio muestral $\omega$ asigna un número real $X(\omega)$.

La distribución de $X$ queda determinada por los valores posibles que puede tomar y las probabilidades con que efectivamente lo hace.

## *Variables aleatorias discretas*

Como dijimos antes, hay dos grandes tipos de variables aleatorias: las discretas y las continuas. Esencialmente, cuando estamos interesados en *contar* casos tratamos con variables discretas, y cuando *medimos* cantidades tratamos con variables continuas. También obtenemos variables discretas si redondeamos una continua, o si tenemos en cuenta la precisión de los aparatos de medición.

Vamos a comenzar estudiando variables discretas pues son más sencillas desde el punto de vista matemático. Más adelante volveremos sobre las variables continuas.

Si $X$ es una variable aleatoria, su recorrido (o imagen) es el conjunto de valores que puede tomar. Lo notaremos por $R_X$.

**Definición de variable aleatoria discreta**

Una variable aleatoria $X$ es discreta si su recorrido es numerable. Es decir, si podemos ordenar en una sucesión

$$R_X = \{x_1, x_2, \dots\}$$

los valores posibles que puede tomar. El recorrido de $X$ puede ser tanto finito como infinito.

*Ejemplo 1* Al lanzar un dado dos veces, podemos registrar los resultados mediante el par $(i, j)$, en donde $i$ es el resultado del primer lanzamiento, y $j$ el del segundo. Podemos tomar como espacio muestral

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

La probabilidad de cada resultado es $P\,(i, j) = 1/36$.

En un juego de apuestas con estos dos dados, se gana \$500 si la suma es 7 y se pierde \$100 en caso contrario. Llamando a las ganancias

$X$, podemos describirla formalmente como

$$X(i,j) = \begin{cases} 500 & \text{si } i+j=7; \\ -100 & \text{si } i+j \neq 7. \end{cases}$$

$X$ es un ejemplo de variable aleatoria discreta.

Podemos cambiar la apuesta. Por ejemplo

$$Y(i,j) = ij - 10.$$

En este caso, si sacas $(6,2)$ entonces ganas \$2. Si sacas $(2,3)$ perdes \$4. ■

*Responder intuitivamente*: ¿Qué apuesta preferís jugar? Volveremos más tarde sobre este asunto.

## *Distribución de una variable discreta*

Las variables aleatorias determinan eventos. Para cada $x \in \mathbb{R}$, escribimos $\{X = x\}$ para indicar el evento que consiste de aquellos resultados $\omega$ tales que $X(\omega) = x$. Del mismo interpretamos eventos como $\{X \leq x\}$, $\{X < x\}$, $\{X \geq x\}$, etc.

*Ejemplo 2* Tomemos $X$ como en el ejemplo anterior. El evento $\{X = 500\}$ consiste de los resultados

$$(1,6),\ (2,5),\ (3,4),\ (4,3),\ (5,2),\ (6,1),$$

es decir, de todos los pares $(i,j)$ que suman 7 (cuadros azules en el dibujo de la Figura 2).

En particular $P(X = 500) = 1/6$. Cuando un valor de $x$ no está en el recorrido de $X$, la probabilidad de $\{X = x\}$ es cero. Por ejemplo, $P(X = 1000) = 0$ en este caso. ■

| | | | | | |
|---|---|---|---|---|---|
| (1,6) | (2,6) | (3,6) | (4,6) | (5,6) | (6,6) |
| (1,5) | (2,5) | (3,5) | (4,5) | (5,5) | (6,5) |
| (1,4) | (2,4) | (3,4) | (4,4) | (5,4) | (6,4) |
| (1,3) | (2,3) | (3,3) | (4,3) | (5,3) | (6,3) |
| (1,2) | (2,2) | (3,2) | (4,2) | (5,2) | (6,2) |
| (1,1) | (2,1) | (3,1) | (4,1) | (5,1) | (6,1) |

Figura 2: El evento $\{X = 500\}$ del Ejemplo 2.

La distribución de una variable discreta queda determinada entonces por las probabilidades con las que toma cada uno de los valores de su recorrido.

**Función de probabilidad puntual**

La función de probabilidad puntual (f.p.p.) de una variable aleatoria discreta $X$ es la función $p : \mathbb{R} \to [0,1]$ definida por

$$p(x) = P(X = x).$$

La distribución de $X$ queda entonces determinada por su f.p.p..

Si hay riesgo de confusión, escribiremos $p_X(x)$ en lugar de $p(x)$ para indicar que se trata de la f.p.p. de $X$.

La función $p(x)$ está definida para todo $x \in \mathbb{R}$, pero si $x$ no es un valor que $X$ pueda tomar, entonces $p(x) = 0$. Es claro que $0 \leq p(x) \leq$

1, y que

$$\sum_{x \in R_X} p(x) = 1,$$

pues en $R_X$ están todos los valores posibles de $X$.

*Ejemplo 3* Consideremos nuevamente el espacio muestral $\Omega$ asociado al lanzamiento de dos dados. Pero sea ahora $M$ la variable aleatoria igual al máximo de los dos resultados:

$$M(i, j) = \text{máx}\{i, j\}.$$

Por ejemplo, si sacas $(3, 5)$ el máximo es 5, i.e. $M(3, 5) = 5$.

Podemos describir entonces la f.p.p. de $M$ con una tabla que enumere los posibles valores y las probabilidades de los mismos. Como en este caso $R_M = \{1, 2, 3, 4, 5, 6\}$, tenemos

| Valor | $x$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| f.p.p. | $p(x)$ | $1/36$ | $3/36$ | $5/36$ | $7/36$ | $9/36$ | $11/36$ |

A veces es mejor visualizar la f.p.p. con un gráfico, como el de la Figura 3. Rigurosamente deberíamos dibujar puntos en lugar de barras, pero colocamos una barra entera para visualizar mejor. Notar que en aquellos $x$ fuera del recorrido de $M$ no hemos graficado nada.

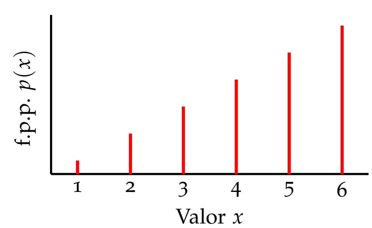



Figura 3: Gráfico de la función de probabilidad puntual de $X$.

Otra forma útil de representar la distribución de una variable es con la función de distribución acumulada (f.d.a.).

**Función de distribución acumulada**

La función de distribución acumulada de una variable aleatoria $X$ es la función $F : \mathbb{R} \to [0, 1]$ definida por $F(x) = P(X \leq x)$.

Con frecuencia diremos simplemente función de distribución. Igual que antes, si hay riesgo de confusión, escribiremos $F_X$ para indicar que se trata de la f.d.a. de $X$.

Observar que $F(x)$ usa el signo de *menor o igual*. Esto será importante para no equivocarse en los cálculos.

*Ejemplo 4* Siguiendo con el ejemplo anterior, tenemos

| Valor | $x$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| f.p.p. | $p(x)$ | $1/36$ | $3/36$ | $5/36$ | $7/36$ | $9/36$ | $11/36$ |
| f.d.a. | $F(x)$ | $1/36$ | $4/36$ | $9/36$ | $16/36$ | $25/36$ | $36/36$ |

Se llama función de distribución *acumulada* porque $F(x)$ da la probabilidad acumulada al sumar las probabilidades $p(y)$ con $y \leq x$. Por ejemplo, en la tabla de arriba, la entrada $16/36$ para la f.d.a. de la columna 4 es la suma de los valores de la f.p.p. desde la columna 1 hasta la 4. Formalmente:

$$F(4) = P(M \leq 4) = \frac{1}{36} + \frac{3}{36} + \frac{5}{36} + \frac{7}{36} = \frac{16}{36}.$$

Igual que la f.p.p. la f.d.a. está definida para todo $x$. Pero a diferencia de aquella, $F(x)$ no es cero fuera de $R_M$, sino que es constante. Por ejemplo, $F(4.5) = 16/36$.

El gráfico de $F(x)$ para una variable discreta siempre se parece a una escalera como es el caso en este ejemplo (ver la Figura 4). ■

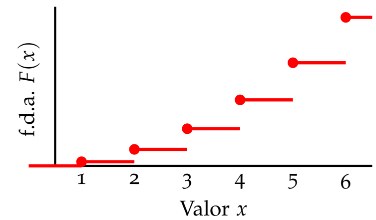
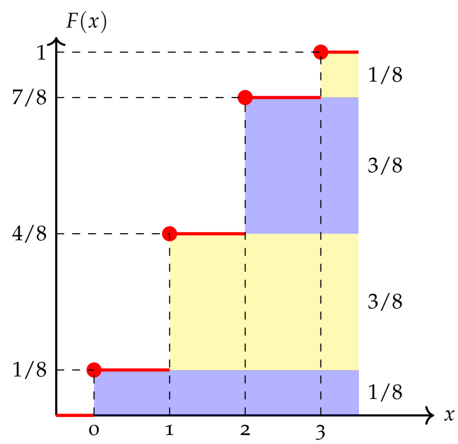


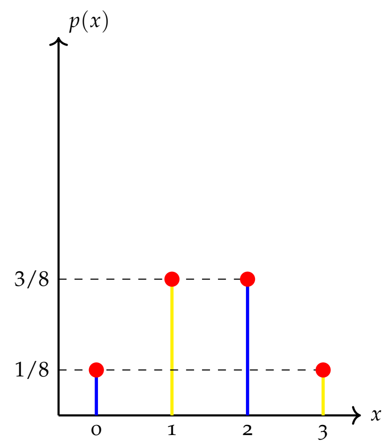

Figura 4: Gráfico de la fda de $X$.

*Ejemplo 5* Sea $X$ el número de caras en 3 lanzamientos de una moneda justa. Entonces

| Valor | $x$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|---|
| f.p.p. | $p(x)$ | $1/8$ | $3/8$ | $3/8$ | $1/8$ |
| f.d.a. | $F(x)$ | $1/8$ | $4/8$ | $7/8$ | $8/8$ |

y los gráficos son




Los colores muestran la relación entre ambas, y como la f.d.a. se obtiene acumulando las probabilidades a medida que $x$ crece. ■

Ciertas propiedades de la f.d.a. se hacen visibles en los ejemplos anteriores:

- **Monotonia**: $F$ es no-decreciente, esto es, el gráfico de $F$ nunca va hacia abajo. Formalmente, si $x \leq y$ entonces $F(x) \leq F(y)$, ya que el evento $\{X \leq x\}$ implica $\{X \leq y\}$, por lo que $P(X \leq x) \leq P(X \leq y)$.

  Esto explica porque $F(x)$ crece o se mantiene constante a medida que $x$ crece, pero nunca decrece.

- **Límites en infinito**:

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

  Esto es, a medida que $x$ crece sin límite, se hace más y más cierto que $\{X \leq x\}$, y del mismo modo, se hace menos probable a media que $x$ decrece sin límite.

La prueba se basa en la propiedad de continuidad de la probabilidad. Los eventos $\{X \leq n\}$ son crecientes, y su unión es todo $\Omega$, por lo que

$$\lim_{n\to+\infty} P\left(X \leq n\right) = P\left(\Omega\right) = 1.$$

El límite en $-\infty$ es análogo.

- **Continuidad por derecha**:

$$\lim_{y\to x^+} F(y) = F(x).$$

Los eventos $\{X \leq x + 1/n\}$ son decrecientes y su intersección es $\{X \leq n\}$. Nuevamente, por la propiedad de continuidad

$$\lim_{n\to+\infty} P\left(X \leq x + \frac{1}{n}\right) = P\left(X \leq x\right).$$

Esto explica porque ponemos un círculo en el gráfico de $F$, del lado derecho de cada punto en donde $F$ pega un salto.

- **Saltos**: El salto de $F$ en $x$ es igual a $P\left(X = x\right)$. Ya sabemos que el límite por derecha es igual a $F(x)$. El límite por izquierda es igual a

$$\lim_{y\to x^-} F(y) = \lim_{n\to\infty} P\left(X \leq x - \frac{1}{n}\right) = P\left(X < x\right),$$

ya que la unión de los eventos $\{X \leq x - 1/n\}$ es $\{X < x\}$. Luego, el salto es

$$F(x) - F(x^-) = P\left(X \leq x\right) - P\left(X < x\right) = P\left(X = x\right).$$

En particular, $F$ es continua en $x$ si, y solo si $P\left(X = x\right) = 0$.

## Distribución conjunta

Cuando tenemos dos variables $X$ e $Y$, las distribuciones respectivas no nos proporcionas toda la información que precisamos para calcular probabilidades. Consideremos un ejemplo simple: la distribución aleatoria de tres bolas distinguibles en tres celdas también distinguibles.

El espacio muestral consiste de $3^3 = 27$ elementos, todos representados en la Tabla 2. Cada resultado tiene entonces probabilidad $1/27$.

Consideremos las siguientes variables: $N$ el número de celdas ocupadas, $X$ el número de bolas en la primer celda, e $Y$ el número de bolas en la segunda celda.

Las distribuciones respectivas de $N$, $X$ e $Y$ son:

| Valor | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| $p_N$ | 0 | 1/9 | 2/3 | 2/9 |
| $p_X$ | 8/27 | 12/27 | 6/27 | 1/27 |
| $p_Y$ | 8/27 | 12/27 | 6/27 | 1/27 |

| | | |
|---|---|---|
| $\{abc, -, -\}$ | $\{a, bc, -\}$ | $\{-, a, bc\}$ |
| $\{-, abc, -\}$ | $\{b, ac, -\}$ | $\{-, b, ac\}$ |
| $\{-, -, abc\}$ | $\{c, ab, -\}$ | $\{-, c, ab\}$ |
| $\{ab, c, -\}$ | $\{a, -, bc\}$ | $\{a, b, c\}$ |
| $\{ac, b, -\}$ | $\{b, -, ac\}$ | $\{a, c, b\}$ |
| $\{bc, a, -\}$ | $\{c, -, ab\}$ | $\{b, a, c\}$ |
| $\{ab, -, c\}$ | $\{-, ab, c\}$ | $\{b, c, a\}$ |
| $\{ac, -, b\}$ | $\{-, ac, b\}$ | $\{c, a, b\}$ |
| $\{bc, -, a\}$ | $\{-, bc, a\}$ | $\{c, b, a\}$ |

Tabla 2: Distribución aleatoria de tres bolas distinguibles en tres celdas distinguibles. Las bolas están representadas por las letras $a, b, c$ y las celdas por los lugares entre las comas.

Notar que $X$ e $Y$ tienen la misma distribución, hecho que es claro por simetría. Sin embargo, este cuadro no nos proporciona la información suficiente para calcular, por ejemplo, la probabilidad de que $\{X = 1\}$ e $\{Y = 2\}$. Para esto debemos hacer una tabla más completa, que representa la *distribución conjunta* de las variables. Es decir, debemos especificar la probabilidad de eventos del tipo $\{X = i\}$ e $\{Y = j\}$, lo cual hacemos mediante una *tabla de contingencia*. Esta información se muestra en la Tabla 3.

| $N\backslash X$ | 0 | 1 | 2 | 3 | Distribución de $N$ |
|---|---|---|---|---|---|
| 1 | 2/27 | 0 | 0 | 1/27 | 1/9 |
| 2 | 6/27 | 6/27 | 6/27 | 0 | 2/3 |
| 3 | 0 | 6/27 | 0 | 0 | 2/9 |
| Distribución de $X$ | 8/27 | 12/27 | 6/27 | 1/27 | 1 |

| $Y\backslash X$ | 0 | 1 | 2 | 3 | Distribución de $Y$ |
|---|---|---|---|---|---|
| 0 | 1/27 | 3/27 | 3/27 | 1/27 | 8/27 |
| 1 | 3/27 | 6/27 | 3/27 | 0 | 12/27 |
| 2 | 3/27 | 3/27 | 0 | 0 | 6/27 |
| 3 | 1/27 | 0 | 0 | 0 | 1/27 |
| Distribución de $X$ | 8/27 | 12/27 | 6/27 | 1/27 | 1 |

Tabla 3: Arriba: distribución conjunta de $N$ y $X$. Abajo: distribución conjunta de $X$ e $Y$.

Por ejemplo, la tabla contiene la información adicional de que $X$ e $Y$ no pueden ser igual a 3 simultáneamente.

Notar que la distribución de $X$ (en cualquiera de las dos tablas) se obtiene sumando las filas de la tabla de contingencia. Lo mismo vale pata $N$ e $Y$, pero sumando las columnas. Las entradas de la tabla son las probabilidades conjuntas

$$P(N = i, X = j) \text{ (arriba) y } P(Y = i, X = j) \text{ (abajo).}$$

Más generalmente, la distribución conjunta de dos variables cualesquiera $X$ e $Y$ queda determinada por la *función de probabilidad conjunta*.

**Función de probabilidad conjunta**

La función de probabilidad conjunta de $X$ e $Y$ es la función $p : \mathbb{R}^2 \to [0,1]$ definida por

$$p(x,y) = P(X = x, Y = y).$$

Para indicar las variables $X$ e $Y$ a veces escribiremos $p_{X,Y}$.

Notar que $\sum_{x \in R_X, y \in R_Y} p(x,y) = 1$.

Las distribuciones de $X$ e $Y$ se pueden recuperar a partir de la conjunta sumando columnas y filas. Es decir

$$p_X(x) = P(X = x) = \sum_{y \in R_Y} P(X = x, Y = y) = \sum_{y \in R_Y} p_{X,Y}(x, y)$$
$$p_Y(y) = P(Y = y) = \sum_{x \in R_X} P(X = x, Y = y) = \sum_{x \in R_X} p_{X,Y}(x, y)$$

Cuando obtenemos las distribuciones de $X$ e $Y$ a partir de la distribución conjunta, éstas se llaman *distribuciones marginales* de $X$ e $Y$ respectivamente.

## *Variables independientes*

¿Se puede reconstruir la distribución conjunta $p_{X,Y}$ a partir de las distribuciones marginales $p_X$ y $p_Y$? La respuesta es no, pero comencemos por ver un par de ejemplos sencillos.

Supongamos que Ana y Beto disponen de tres pares de monedas. Dos de éstos pares son "mágicos" en el sentido de que el resultado del lanzamiento de una de las monedas influye sobre el resultado de la otra. El tercer par consiste de dos monedas normales.

Ana y Beto quieren saber cuál de los tres pares es el normal, y para ello se deciden a lanzar varias veces cada moneda. Primero eligen uno de los pares, Ana toma una moneda y Beto la otra, y las lanzan muchas veces para registrar las frecuencias relativas de caras y cruces.

Lo hacen con el primer par, y definen variables $X_1$ que vale uno si la moneda que lanza Ana sale cara y 0 si no, e igualmente definen $Y_1$ que valen 1 si la moneda que lanza Beto sale cara. Del mismo modo definen las variables $X_2, Y_2$ y $X_3, Y_3$ correspondientes a los otros pares de monedas.

En la Tabla 4 se muestran las probabilidades calculadas por Ana y Beto. Notar que en los tres casos, las marginales (que se muestran en rojo) son todas iguales. Más aún, son todas iguales a la distribución de una moneda justa, por lo que si miramos solamente los registros de caras y cruces que obtiene Ana (o Beto) individualmente, lo que percibimos son los resultados de una moneda normal perfectamente equilibrada.

Sin embargo, cuando miramos los resultados *en conjunto* vemos la magia entre las monedas. En el primer caso, si Ana obtiene una cara, la probabilidad de que Beto obtenga una cara es $2/3$ y no $1/2$. Lo opuesto ocurre en el tercer caso, en el cual la probabilidad es de $1/3$.

Es decir, en el primer caso las variables están positivamente relacionadas, y en el tercero lo están negativamente. El par de monedas normales es el segundo, para el cuál el resultado del lanzamiento de Ana no influye sobre el resultado de Beto. En este caso decimos que $X_2$ e $Y_2$ son variables *independientes*.

| | | $X_1$ = 0 | $X_1$ = 1 | $p_{Y_1}$ |
|---|---|---|---|---|
| $Y_1$ | 0 | 1/3 | 1/6 | 1/2 |
| | 1 | 1/6 | 1/3 | 1/2 |
| | $p_{X_1}$ | 1/2 | 1/2 | 1 |

| | | $X_2$ = 0 | $X_2$ = 1 | $p_{Y_2}$ |
|---|---|---|---|---|
| $Y_2$ | 0 | 1/4 | 1/4 | 1/2 |
| | 1 | 1/4 | 1/4 | 1/2 |
| | $p_{X_2}$ | 1/2 | 1/2 | 1 |

| | | $X_3$ = 0 | $X_3$ = 1 | $p_{Y_3}$ |
|---|---|---|---|---|
| $Y_3$ | 0 | 1/6 | 1/3 | 1/2 |
| | 1 | 1/3 | 1/6 | 1/2 |
| | $p_{X_3}$ | 1/2 | 1/2 | 1 |

Tabla 4: Las monedas mágicas. Arriba: distribución conjunta de $X_1$ e $Y_1$. Centro: distribución conjunta de $X_2$ e $Y_2$. Abajo: distribución conjunta de $X_3$ e $Y_3$

Notar que en el segundo caso, la distribución conjunta de $X_2$ e $Y_2$ se obtiene *multiplicando las marginales*. Este es esencialmente el único caso en el cuál podemos recuperar la distribución conjunta a partir de las marginales.

**Variables discretas independientes**

Decimos que dos variables aleatorias discretas $X$ e $Y$ son independientes, si la distribución conjunta es igual al producto de las marginales:

$$p_{X,Y}(x,y) = p_X(x) \cdot p_Y(y),$$

para todo $x_i \in R_X$ e $y_j \in R_Y$.

## *Aritmética con variables aleatorias*

Podemos hacer aritmética con las variables aleatorias. Por ejemplo, podemos sumar, restar, multiplicar, o elevar al cuadrado. Una operación muy importante para nosotros será la suma de variables independientes.

*Ejemplo 6* Sean $X$ e $Y$ variables aleatorias independientes con las siguientes f.p.p.

| Valores de $X$ | 1 | 2 | 3 | 4 | |
|---|---|---|---|---|---|
| f.p.p. $p(x)$ | $1/10$ | $2/10$ | $3/10$ | $4/10$ | |

| Valores de $Y$ | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| f.p.p. $p(y)$ | $1/15$ | $2/15$ | $3/15$ | $4/15$ | $5/15$ |

Calculemos la f.p.p. de la suma $X + Y$.

| Valores de $X$ \ Valores de $Y$ | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| 1 | $1/150$ | $2/150$ | $3/150$ | $4/150$ | $5/150$ | $1/10$ |
| 2 | $2/150$ | $4/150$ | $6/150$ | $8/150$ | $10/150$ | $2/10$ |
| 3 | $3/150$ | $6/150$ | $9/150$ | $12/150$ | $15/150$ | $3/10$ |
| 4 | $4/150$ | $8/150$ | $12/150$ | $16/150$ | $20/150$ | $4/10$ |
| | $1/15$ | $2/15$ | $3/15$ | $4/15$ | $5/15$ | |

La primera cosa a hacer es una tabla de contingencia con la distribución conjunta de $X$ e $Y$. Como $X$ e $Y$ son independientes, la f.p.p. conjunta es simplemente el producto de las marginales $p(x, y) = p(x)p(y)$.

Las entradas en las diagonales corresponden a los casos con igual $X + Y$. Todo lo tenemos que hacer para calcular la f.p.p. de $X + Y$ es sumar las probabilidades de cada diagonal.

| Valores de $X + Y$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|
| f.p.p. | 1/150 | 4/150 | 10/150 | 20/150 | 30/150 | 34/150 | 31/150 | 20/150 |

Cuando las tablas sean demasiado grandes para poder escribirlas, vamos a tener que usar métodos puramente "algebraicos" para calcular las probabilidades de una suma. Aprenderemos cómo hacer esto en su debido tiempo. ■