Variables aleatorias discretas

Problema

Lanzamos dos dados al aire. Nos interesa encontrar probabilidades tal como la probabilidad de que la suma de los puntos en los dados es menor que 8.

Si asumimos que todos los resultados observados al tirar dos dados son equiprobables entonces el espacio muestral del experimento, con treinta y seis posibles resultados es:

		Dado Verde					
		1	2	3	4	5	6
Dado Rojo	1	1, 1	1, 2	1, 3	1, 4	1, 5	1, 6
	2	2, 1	2, 2	2, 3	2, 4	2, 5	2, 6
	3	3, 1	3, 2	3, 3	3, 4	3, 5	3, 6
	4	4, 1	4, 2	4, 3	4, 4	4, 5	4, 6
	5	5, 1	5, 2	5, 3	5, 4	5, 5	5, 6
	6	6, 1	6, 2	6, 3	6, 4	6, 5	6, 6

Tabla 1 Espacio muestral resultante al tirar dos dados

Como nos interesa la suma de los puntos observados, si obtenemos el resultado (3, 5) le asignamos el valor 8, correspondiente a la suma de 3 y 5. Podemos calcular la probabilidad de que la suma sea igual a 8, contando todos los resultados donde la suma es ocho. El evento de que la suma es ocho contiene 5 resultados: {(2,6), (3,5), (4,4), (5, 3), (6,2)}; por lo tanto la probabilidad deseada es 5/36. Podemos repetir este proceso con cada uno de los resultados para obtener la siguiente tabla:

Resultado	2	3	4	5	6	7	8	9	10	11	12
Probabilidad	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Tabla 2 Distribución de probabilidad del total observado al tirar dos dados.

Hemos encontrado la distribución de probabilidad de los valores posibles de la suma al tirar dos dados. Si R representa el resultado observado en el dado rojo y G el resultado que se observará en el dado verde, podemos expresar el valor que nos interesa así: X = R + G. Antes de lanzar los dados no sabemos qué valores observaremos para R y G, por lo tanto tampoco lo sabemos para X.

El valor que X asumirá puede variar de tirada en tirada sujeto a la distribución especificada en la tabla de arriba. Así X es una variable, que asume un número finito de valores sujeto a una distribución de probabilidad. Este es un ejemplo de una variable aleatoria discreta. Otros ejemplos son las variables R y G. En general, si S es un espacio muestral con una medida de probabilidad P, definimos una **variable aleatoria** como una función que asigna un número real a cada uno de los elementos de S. Es decir X es una función cuyo dominio es el espacio muestral S y su codominio es el conjunto de números reales Â, en la notación usual X: S ®Â.

Interpretamos, por ejemplo X = 8 como el evento de que se observó el resultado 8 al tirar los dos dados, es decir el evento { (2,6), (3,5), (4,4), (5, 3), (6,2)} ocurrió. También asignamos a X = 8 la probabilidad de ese evento. Así vemos que P(X = 8) = P({ (2,6), (3,5), (4,4), (5, 3), (6,2)}) = 5/36. Nota que a pesar de que X es una función, usualmente no se escribe el argumento de la función, es decir, si s es un elemento del espacio muestral S, en vez de escribir X(s), sencillamente escribimos X. Es usual denotar las variables aleatorias por letras mayúsculas y los valores que puede asumir por letras minúsculas.

En este caso la variable X puede asumir un valor de estre un conjunto finito de valores posibles. Cualquier variable que pueda asumir un número finito de valores decimos es una **variable aleatoria discreta**. También son variables aleatorias discretas aquellas que pueden asumir un número muy grande o infinito de valores que potencialmente podrían ser contados, tal como el número de estrellas en el firmamento, el número de granos de arena en el planeta o el número de hojas en los árboles.

Experimenta y resuelve
1. Da ejemplos de variables aleatorias discretas. Indica cuales pueden asumir un número finito de valores distintos y cuales un número infinito.
2. Da ejemplos de variables aleatorias que no son discretas.

En la Tabla 2 arriba vemos que a cada valor posible de X, le asignamos un número correspondiente a su probabilidad. Así podemos definir otra función: f(x) = P(X = x), para cada número x en el campo de valores de la variable X. Esta función se llama la **función de probabilidad** o **distribución de probabilidad** de la variable X. Para el ejemplo de la suma de los puntos al tirar dos dados, los valores de esta función están dados en la Tabla 2, la cual se puede reescribir usando los conceptos estudiados.

x	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
f(x)	**1/36**	**2/36**	**3/36**	**4/36**	**5/36**	**6/36**	**5/36**	**4/36**	**3/36**	**2/36**	**1/36**

Tabla 3 Distribución de probabilidad del total observado al tirar dos dados.

Experimenta y resuelve
Examina la Tabla 3 y usa la definición de f(x) para deducir algunas propiedades de esta función.

Vemos que f(x) nunca adquiere un valor menor de cero. Esto se debe a que f(x) representa una probabilidad, la cual nunca puede ser menor de cero. De igual manera f(x) nunca puede ser menor de 1. Si sumamos todos lo valores que puede tener f(x) obtenemos 1, debido a que estamos sumando las probabilidades de que la variable aleatoria asuma uno de los valores establecidos. Por su definición, la función de probabilidad tiene las siguientes características:

1. f(x) ³ 0 para todo valor x en su dominio
2. $\sum_{x} f(x) = 1$ donde la sumatoria se extiende sobre todos los valores x en el dominio de f.

Experimenta y resuelve
1. Coteja que la función $f(x) = \frac{x}{15}$ es una función de probabilidad para x = 1, 2, 3, 4, 5. Indica su dominio y su campo de valores.
2. Lanza cuatro monedas al aire. Define la variable aleatoria Y como el número de caras observadas. Construye la función de probabilidad de Y.
3. Lanza dos dados al aire. Define la variable aleatoria X como la diferencia de los puntos observados en los dados. Construye la función de probabilidad de X.

Los valores de la función de probabilidad se pueden representar en una gráfica como la siguiente:

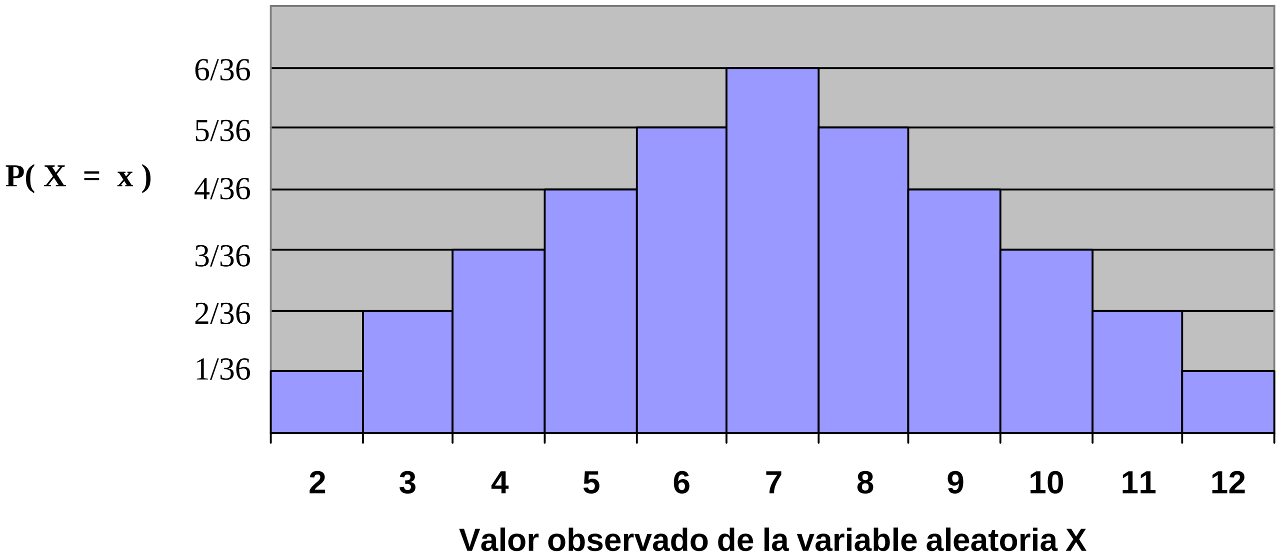

Figura 1 Histograma de probabilidad de X

La probabilidad de observar un valor particular de la variable aleatoria, digamos X = 3 está dado por la altura de la barra sobre el 3, es decir P(X = 3) = 2/36. De igual manera, en vez de asociar la altura de la barra con la probabilidad, podemos ver que el área de la barra sobre el 3 es $2/36 \times 1 = 2/36$, ya que la altura de la barra es 2/36 y su ancho es 1. Usar el área de las barras para representar la probabilidad es muy útil para extender la noción de probabilidad a otras variables.

Experimenta y resuelve
1. Encuentra las siguientes probabilidades: P(X = 6), P(X < 7), P(X ≥ 9), P(X = 6.5), P(X ≤ 4.4)

Podemos usar el histograma de probabilidad para calcular probabilidades tal como P(X ≤ 4). Vemos que P(X ≤ 4) = P(X =2 ó X =3 ó X =4) = P(X = 2) + P(X = 3) + P(X = 4) , ya que los eventos donde X = 2, X= 3 y X=4 son disyuntos. Entonces P(X ≤ 4) = 1/36 + 2/36 + 3/36 = 6/36, sumando las áreas de la barras que están sobre el 4 y a su izquierda. Debemos ser muy cuidadosos con las desigualdades, ya que P(X ≤ 4) = 6/36, mientras que P(X< 4) = 3/26.

Extendiendo esta idea de probabilidades acumulativas, podemos definir otra función partiendo de la distribución de probabilidad. Si X es una variable aleatoria discreta, definimos la **función de distribución de X** o **función de distribución acumulativa de X** de la siguiente manera:

F(x) = P(X £ x) = $\sum_{t \le x} f(t) \;\; para \;\; -\infty < x < \infty.$

Experimenta y resuelve
1. Calcula la función de distribución acumulativa de la suma de los puntos en dos dados.
2. Usa las propiedades de la función de probabilidad para encontrar algunas propiedades de la función de distribución acumulativa.

La Tabla 4 presenta la función de distribución acumulativa del resultado observado al tirar dos dados. De esa tabla podemos deducir algunas propiedades. Por ejemplo, vemos que F(3) < F(4), es decir si el valor en que se evalúa la función es mayor, el valor de la función también será mayor.

Experimenta y resuelve
¿Es cierta siempre esta propiedad? Examina qué ocurre con x = 5, con x = 6 y con x = 5.7.

x	2	3	4	5	6	7	8	9	10	11	12
F(x)	1/36	3/36	6/36	10/36	15/36	21/36	26/36	30/36	33/36	35/36	36/36

Tabla 4 Función de distribución acumulativa del total observado al tirar dos dados.

A pesar de que el valor de la función de distribución acumulativa para x = 5.7 no está incluida entre los valores en la tabla, podemos usar la definición para obtenerlo. F(x) = P(X ≤ x), así F(5.7) = P (X ≤ 5.7). Cuando escribimos esta última probabilidad nos preguntamos cuál es la probabilidad de observar que el total de puntos en dos dados es menor o igual a 5.7. Por la naturaleza del experimento, vemos que no es posible observar valores distintos a {2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12}, por esta razón los resultados que pueden observarse y que son menores o iguales a 5.7 son {2, 3, 4, 5}, tenemos pues que, F(5.7) = P (X ≤ 5.7) = P (X ≤ 5) = F(5) = 10/36.

Esto nos demuestra que la propiedad que habíamos visto antes, en la que establecíamos que si a y b son dos número reales con a < b entonces F(a) < F(b), no es cierta siempre. Lo que sí es cierta es que si tenemos dos números reales a, b tal que a < b, entonces F(a) ≤ F(b).
Por la definición de probabilidad y por esta misma propiedad vemos que el valor más grande que puede tener F(x) es 1 y el valor más pequeño de esa función es 0. Resumimos las propiedades encontradas:

La función de distribución acumulativa F(x) tiene las siguientes propiedades:
1. F(- ∞) = 0
2. F(∞) = 1
3. Si a, b son números reales, con a < b, entonces F(a) ≤ F(b). Esto quiere decir, en lenguaje matemático, que F es una función no decreciente.
4. F(x) es una función continua por la derecha: si a es un número real, entonces $\lim_{x \to a^+} F(x) = F(a)$.

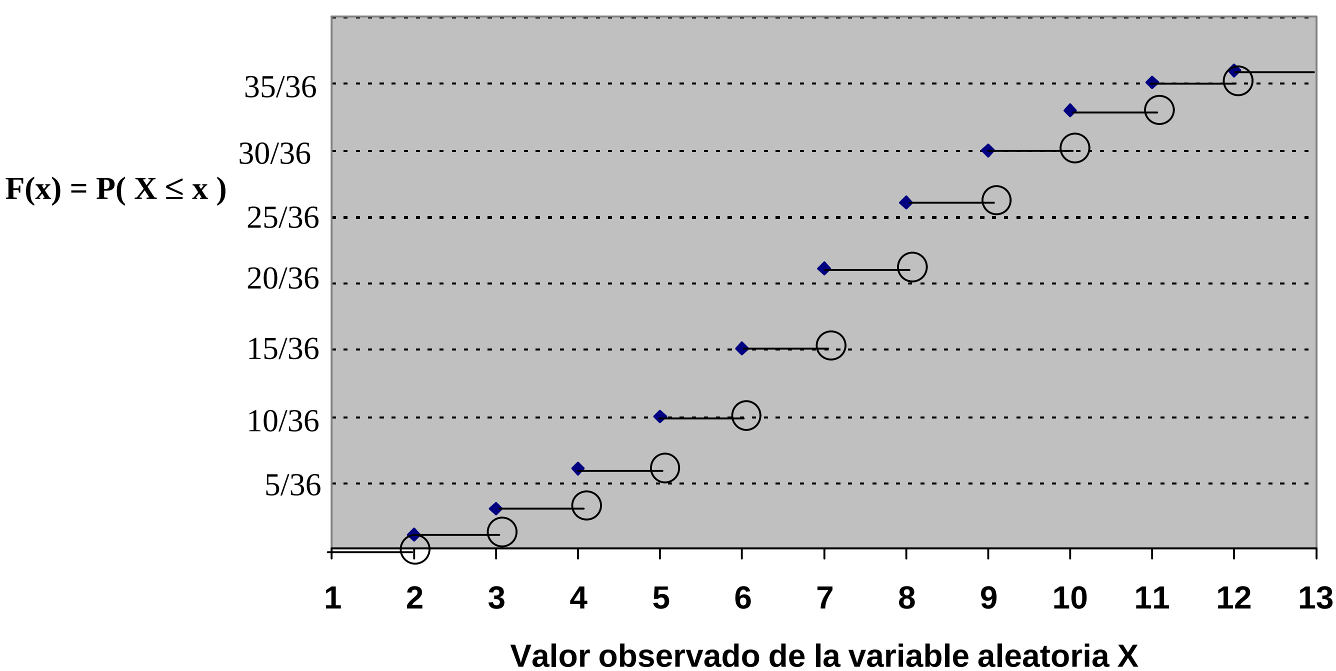

Figura 2 Gráfica de la función de distribución acumulativa del total de puntos al tirar dos dados

La gráfica de $F(x)$ parece una escalera. Podemos ver la razón por la cual esta gráfica debe ser de esta manera si examinamos los valores de la función de distribución en un intervalo tal como [3,4]. Vemos que $F(3) = 3/36$, si escogemos un número x mayor de 3, pero menor de 4, tenemos que $F(x) = 3/36$. Para todos los valores de x tal que $3 \leq x < 4$ tenemos que $F(x) = 3/36$. Sin embargo, al evaluar la función en $x = 4$ vemos que $F(4) = 6/36$, por esta razón la gráfica muestra un salto en ese punto.

También podemos notar que el tamaño del salto en $x = 4$ nos dice la probabilidad de que $X = 4$. Para valores de x entre 3 y 4 tenemos que $F(x) = 3/36$, como habíamos visto y luego $F(4) = 6/36$, así el tamaño del salto en $x = 4$ es $6/36 - 3/36 = 3/36$. Este último valor es la probabilidad de que el total de puntos en dos dados, X sea igual a 4, es decir, $P(X = 4) = 3/36$.

Visto de otro modo, $P(X = 4) = P(X \leq 4) - P(X < 4)$. Esto es igual a $P(X \leq 4) - P(X \leq 3) = 6/36 - 3/36= 3/36$. En general, el tamaño del salto de la función de distribución en un valor particular, nos da la probabilidad de que la variable aleatoria sea igual a ese valor.

Experimenta y resuelve
Usa la función de distribuición para encontrar $P(X = 5.7)$.