TEMA 4 VARIABLES ALEATORIAS

Pablo Buenestado

Curso 2020-2021 Otoño

Departamento de Matemáticas (UPC)

Índice

Esquema

En muchos casos deseamos asignar un valor numérico a cada resultado de un experimento.

Esta asignación se llama variable aleatoria.

Para aclarar la idea presentamos un ejemplo:

Supongamos que un ingeniero eléctrico tiene seis resistores a mano.

Tres de ellos tienen etiqueta de 10 Ω y los otros tres tienen etiqueta de 20 Ω.

El ingeniero quiere conectar un resistor de 10 Ω y un resistor de 20 Ω en serie, para crear una resistencia de 30 Ω.

Ahora suponemos que, en efecto, los tres resistores etiquetados con 10 Ω tienen las resistencias reales de 9, 10 y 11 Ω y que los tres resistores etiquetados con 20 Ω tienen las resistencias reales de 19, 20 y 21 Ω.

El proceso para seleccionar un resistor de cada tipo es un experimento cuyo espacio muestral consta de nueve resultados igualmente probables.

El espacio muestral se presenta en la tabla siguiente.

Resultado	Probabilidad
$\{9, 19\}$	$1/9$
$\{9, 20\}$	$1/9$
$\{9, 21\}$	$1/9$
$\{10, 19\}$	$1/9$
$\{10, 20\}$	$1/9$
$\{10, 21\}$	$1/9$
$\{11, 19\}$	$1/9$
$\{11, 20\}$	$1/9$
$\{11, 21\}$	$1/9$

Ahora lo que es importante para el ingeniero de este experimento es la suma de las dos resistencias, en vez de sus valores individuales.

Por tanto, asignamos a cada resultado un número igual a la suma de las dos resistencias seleccionadas.

Esta asignación la representamos con la letra X y la exponemos en la tabla siguiente.

Resultado	X	Probabilidad
$\{9, 19\}$	28	$1/9$
$\{9, 20\}$	29	$1/9$
$\{9, 21\}$	30	$1/9$
$\{10, 19\}$	29	$1/9$
$\{10, 20\}$	30	$1/9$
$\{10, 21\}$	31	$1/9$
$\{11, 19\}$	30	$1/9$
$\{11, 20\}$	31	$1/9$
$\{11, 21\}$	32	$1/9$

La función X, que asigna un valor numérico a cada resultado en el espacio muestral, es una variable aleatoria.

Resumen

Una variable aleatoria asigna un valor numérico a cada resultado en un espacio muestral.

Acostumbramos a denotar a las variables aleatorias con letras mayúsculas. Las letras X, Y y Z se usan con más frecuencia.

En muchas ocasiones, podemos calcular las probabilidades de las variables aleatorias de una manera obvia.

En el ejemplo anterior, el evento $X = 29$ corresponde al evento $\{\{9, 20\}, \{10, 19\}\}$ del espacio muestral.

Por tanto, $P(X = 29) = P\left(\{\{9, 20\}, \{10, 19\}\}\right) = 1/9 + 1/9 = 2/9$.

Ejemplo

Hagamos una lista de los valores posibles de la variable aleatoria X definida en el ejemplo previo y determinemos la probabilidad $P(X = x)$ para cada uno de ellos[1].

Solución

Los valores posibles de X son 28, 29, 30, 31 y 32.

Para encontrar la probabilidad de uno de estos valores, se suman las probabilidades de los resultados en el espacio muestral que corresponden al valor.

En la tabla anterior observamos que tenemos un caso para el valor 28, dos para el 29, tres para el 30, dos para el 31 y uno para el 32.

[1] Notación: X (en mayúscula) hace referencia al conjunto de valores de la variable aleatoria, x (en minúscula) hace referencia a un valor concreto de la variable aleatoria.

Solución (Continuación)

Los resultados se organizan en la tabla siguiente.

X	$P(X = x)$
28	$1/9$
29	$2/9$
30	$3/9$
31	$2/9$
32	$1/9$

La tabla de probabilidades del ejemplo anterior contiene toda la información necesaria para calcular cualquier probabilidad que consideremos sobre la variable aleatoria X.

Observemos que los resultados del espacio muestral no se presentan en la tabla.

Cuando se conocen las probabilidades en relación con una variable aleatoria, generalmente no se piensa en el espacio muestral: sólo nos concentramos en las probabilidades.

Hay dos tipos importantes de variables aleatorias: discretas y continuas.

Una variable aleatoria discreta es aquella cuyos valores posibles forman un conjunto discreto; en otras palabras, los valores se pueden ordenar y existen separaciones entre los valores adyacentes. La variable aleatoria X, que se acaba de describir, es discreta.

En contraparte, los valores posibles de una variable aleatoria continua siempre están contenidos en un intervalo; es decir, son todos los puntos entre dos números. Proporcionaremos definiciones precisas de estas clases de variables aleatorias más adelante.

Ejemplo

Con frecuencia los chips de computadora tienen imperfecciones en su superficie.

Para cierto tipo de chip de computadora, el 9% no tiene imperfecciones, el 22% contiene una imperfección, el 26% presenta dos imperfecciones, el 20% contiene tres imperfecciones, el 12% tiene cuatro imperfecciones y el 11% presenta cinco imperfecciones. No constan chips con seis imperfecciones o más.

Sea Y el número de imperfecciones en un chip elegido aleatoriamente, buscamos:

1. ¿Cuáles son los valores posibles de Y?
2. ¿Y es discreto o continuo?
3. Determinemos $P(Y = y)$ para cada valor posible y

Solución

1. Los valores posibles de Y son los enteros 0, 1, 2, 3, 4 y 5.
2. La variable aleatoria Y es discreta, ya que solamente toma valores enteros.
3. Al 9% de los resultados en el espacio muestral se le asigna el valor 0.

   Por tanto, $P(Y = 0) = 0.09$.

   Similarmente $P(Y = 1) = 0.22$, $P(Y = 2) = 0.26$, $P(Y = 3) = 0.20$, $P(Y = 4) = 0.12$ y $P(Y = 5) = 0.11$.

Y	$P(Y = y)$
0	0.09
1	0.22
2	0.26
3	0.20
4	0.12
5	0.11

Ejemplo

Cierto tipo de disco magnético debe funcionar en un ambiente donde está expuesto a gases corrosivos.

Se sabe que el 10% de estos discos tiene tiempos de vidas menores o iguales a 100 horas, el 50% lo tiene mayor a 100 horas, pero menor o igual a 500, y el 40% incluye tiempos superiores a 500 horas.

Sea Z el número de horas en tiempo de vida de un disco elegido aleatoriamente, buscamos

1. ¿Z es continua o discreta?
2. Determinemos $P(Z \leq 500)$.
3. ¿Se pueden calcular todas las probabilidades de Z?

Solución

1. El tiempo de vida de un componente no está limitado a una lista de valores discretamente espaciados; Z es continua.
2. De todos los componentes, el 60% tiene vidas menores o iguales a 500 horas. Por tanto, $P(Z \leq 500) = 0.60$.

Solución (Continuación)

3. No se tiene la información suficiente para calcular todas las probabilidades de Z.

   Se pueden calcular algunas de ellas; por ejemplo, $P(Z \leq 100) = 0.10$, $P(100 < Z \leq 500) = 0.50$ y $P(Z > 500) = 0.40$.

   Pero no se sabe, por ejemplo, la proporción de componentes que tiene tiempos de vida entre 100 y 200 horas, o entre 200 y 300 horas, así que no se puede encontrar la probabilidad de que la variable aleatoria Z tenga cualquiera de estos intervalos.

   Para calcular todas las probabilidades de Z, se necesitaría calcular la probabilidad para cada intervalo posible; por ejemplo, $P(200 < Z \leq 300)$, $P(200 < Z \leq 201)$, $P(200 < Z \leq 200.1)$ y así sucesivamente.

   Veremos cómo se puede hacer lo anterior más adelante, cuando se analicen las variables aleatorias continuas.

Variables aleatorias y poblaciones

Con frecuencia es útil pensar en un valor de una variable aleatoria como si seleccionara un elemento de una población.

Por ejemplo, consideremos la variable aleatoria Y descrita en el ejemplo de las imperfecciones de los chips.

Observar un valor para esta variable aleatoria es como seleccionar un valor de una población que consta de los enteros 0, 1, 2, 3, 4 y 5, en las proporciones siguientes:

0's un 9%; 1's un 22%; 2's un 26%; 3's un 20%; 4's un 12%; y 5's un 11%.

Para una variable aleatoria continua es adecuado imaginar una población infinita que contiene todos los valores posibles de la variable aleatoria.

Por ejemplo, para la variable aleatoria Z del ejemplo de los tiempos de vida de los discos magnéticos se podría imaginar una población que contenga a todos los números positivos, con un 10% de los valores de la población menores o iguales a 100, el 50% superiores a 100 pero menores o iguales a 500, y el 40% mayores a 500.

La proporción de valores de la población en cualquier intervalo sería igual a la probabilidad de que la variable Z está en ese intervalo.

Los métodos para trabajar con variables aleatorias son diferentes para variables aleatorias discretas y continuas.

Empezamos con las variables aleatorias discretas.