Introducción a la probabilidad y estadística CM274

César Lara Avila

28 de agosto de 2017

https://github.com/C-Lara

4. Variables aleatorias discretas

Variables aleatorias

Este tema trata en gran medida de introducir alguna terminología útil, basándose en las nociones de espacio muestral y función de probabilidad. Las palabras clave son:

1. Variable aleatoria.
2. Función de masa de probabilidad (PMF).
3. Función de distribución acumulativa (CDF).

Recordar:

- Un espacio muestral discreta Ω es finito o es un conjunto listable de salidas $\{\omega_1, \omega_2, \ldots\}$. La probabilidad de una salida ω es denotada por $\mathbb{P}(\omega)$.
- Un evento E es un subconjunto de Ω. La probabilidad de un evento E es

$$\mathbb{P}(E) = \sum_{\omega \in E} \mathbb{P}(\omega).$$

Ejemplo de variables aleatorias

- **Ejemplo 1** Un juego con 2 dados.
  Tirar un dado dos veces y registrar los resultados como (i, j), donde i es el resultado de la primera tirada y j el resultado de la segunda. Podemos tomar el espacio muestral:

$$\Omega = \{(1,1), (1,2), (1,3), \ldots, (6,6)\} = \{(i,j) | i,j = 1, \ldots, 6\}.$$

  La función de probabilidad es $\mathbb{P}(i,j) = 1/36$.

  En este juego, tu ganas 500 soles si la suma es 7 y perdemos 100 en otro caso. Damos a esta función de pago (payoff) el nombre de X, que describe esto formalmente como:

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

Ejemplo de variables aleatorias(1)

- **Ejemplo 2**
  Podemos cambiar el juego usando una función de pago diferente. Por ejemplo:

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

  En este ejemplo, si tira $(6,2)$ entonces gana 2 soles. Si tira $(2,3)$ entonces ganas -4 (es decir, pierdes 4 soles).

  Pregunta: ¿Qué juego es la mejor apuesta?.

  Observación importante :

  Las variables aleatorias asignan un número a cada salida en un espacio muestral.

Definición

Sea Ω un espacio muestral. Una **variable aleatoria discreta** es una función:

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

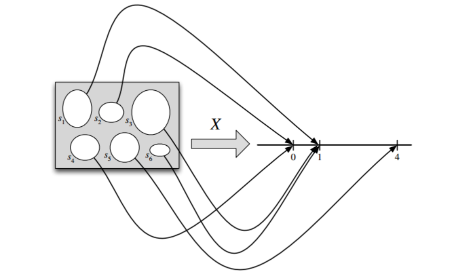
que toma un conjunto de valores discretos.

¿Por qué se llama X a una variable aleatoria?

Es `aleatoria` porque su valor depende de un resultado aleatorio de un experimento. Y tratamos a X como lo haríamos con una variable usual: podemos añadirla a otras variables aleatorias, elevarla al cuadrado y así sucesivamente.

Definición(1)

Dado un experimento y el correspondiente conjunto de posibles respuestas o salidas (el espacio muestral), una variable aleatoria asocia un número particular con cada respuestas o salidas, como se indica en la figura siguiente:



donde la variable aleatoria X representada, es definida sobre un espacio muestral con 6 elementos y los valores posibles $1, 2$ y 4. La aleatoriedad viene de la elección de un punto aleatorio de acuerdo con la función de probabilidad $\mathbb{P}$ para el espacio muestral.

Eventos y variables aleatorias

Para algún valor a, escribimos $X = a$, para indicar que el evento consiste de todas las salidas ω con $X(\omega) = a$.

- **Ejemplo 3**

  En el Ejemplo 1, lanzamos dos dados y X fue una variable aleatoria

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

  El evento $X = 500$ es el conjunto $\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$, esto es, el conjunto de todas las salidas que suman 7. Así $\mathbb{P}(X = 500) = 1/6$.

Permitimos que a sea un valor, incluso valores que X nunca toma. En el ejemplo 1, podríamos mirar el evento $X = 1000$. Puesto que X nunca es igual a 1000 esto es solo el evento vacío (o conjunto vacío).

Función de masa de probabilidad

Resulta difícil de leer y escribir $\mathbb{P}(X = a)$ para la probabilidad que $X = a$. Cuando sabemos que estamos hablando de X simplemente escribiremos $p(a)$. Si queremos hacer explícita X escribiremos $p_X(a)$.

Explicamos esto con una definición.

Definición: La función de masa de probabilidad (PMF) de una variable aleatoria discreta es la función $p(a) = \mathbb{P}(X = a)$. Se debe notar que:

- Siempre se tiene $0 \leq p(a) \leq 1$.
- Permitimos que a sea un número. Si a es un valor que no toma X, entonces $p(a) = 0$.

Ejemplo explicativo

- **Ejemplo 4** Sea Ω el espacio muestral anterior de tirar 2 dados. Definamos la variable aleatoria M como el valor máximo de los dos dados:

$$M(i, j) = \max(i, j).$$

Podemos describir una variable aleatoria enumerando sus posibles valores y las probabilidades asociadas a estos valores. Para el ejemplo anterior tenemos:

valor a:	0	1	2	3	5	6
PMF p(a):	1/36	3/36	5/36	7/36	9/36	11/36

Por ejemplo $p(2) = 3/36$.

Eventos y desigualdades

Las desigualdades con variables aleatorias describen eventos. Por ejemplo, $X \leq a$ es el conjunto de todos los resultados ω tales que $X(\omega) \leq a$.

- **Ejemplo 5**
  Si el espacio muestral es el conjunto de todos los pares de (i, j) procedentes de tirar dos dados y $Z(i, j) = i + j$ es la suma de los dados entonces:

$$Z \leq 4 = \{(1,1),(1,2),(1,3),(2,1),(2,2),(3,1)\}$$

Función de distribución acumulativa

Definición: La función de distribución acumulativa (CDF) de una variable aleatoria X es la función F dada por $F(a) = \mathbb{P}(X \leq a)$. Con frecuencia acortaremos esto a la función de distribución.

Tenga en cuenta que la definición de $F(a)$ usa el símbolo menor o igual. Esto será importante para obtener los cálculos exactamente .

- **Ejemplo 6** Continuando con el ejemplo M, tenemos:

valor a:	1	2	3	4	5	6
PMF p(a):	1/36	3/36	5/36	7/36	9/36	11/36
CDF F(a):	1/36	4/36	9/36	16/36	25/36	36/36

$F(a)$ se denomina función de distribución acumulativa porque $F(a)$ da la probabilidad total que se acumula sumando las probabilidades $p(b)$ cuando b va de $-\infty$ a a.

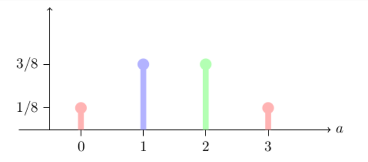
Por ejemplo, en la tabla anterior, la entrada 16/36 en la columna 4 para el CDF es la suma de los valores de PMF desde la columna 1 a la columna 4. En la notación:

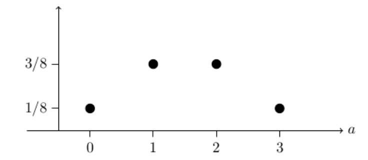
Como eventos:$'M \leq 4' = \{1, 2, 3, 4\}$; $\quad F(4) = \mathbb{P}(M \leq 4) = 1/36+3/36+5/36+7/36 = 16/36.$

Gráficos de $p(a)$ y $F(a)$

Podemos visualizar el PMF y el CDF con gráficos. Por ejemplo, sea X el número de caras en 3 lanzamientos de una moneda:

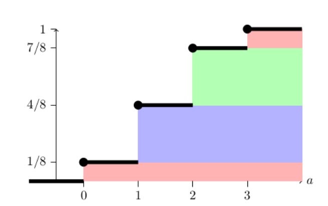
valor a :	0	1	2	3
PMF p(a):	1/8	3/8	3/8	1/8
CDF F(a):	1/8	4/8	7/8	1

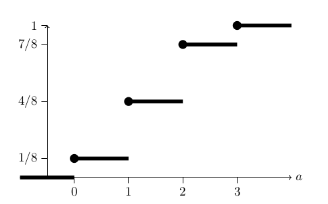




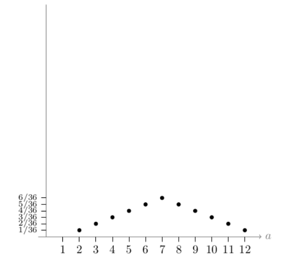
Los gráficos coloreados muestran cómo se construye la función de distribución acumulativa acumulando la probabilidad como un incremento. Los gráficos en blanco y negro son las presentaciones estándar.

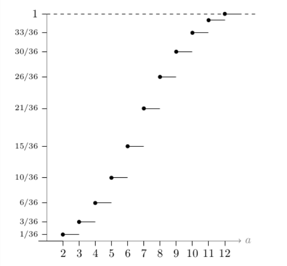
Gráficos de $p(a)$ y $F(a)$ (continuación)





Para el Ejemplo4, el PMF y el CDF, está dado por:





Propiedades de la función de distribución acumulativa

El CDF F de una variable aleatoria satisface varias propiedades:

- F es no decreciente. Es decir, su gráfico nunca cae, o simbólicamente si $a \leq b$ entonces $F(a) \leq F(b)$.
- $0 \leq F(a) \leq 1$.
- $\lim_{a \to \infty} F(a) = 1, \qquad , \lim_{a \to -\infty} F(a) = 0$.

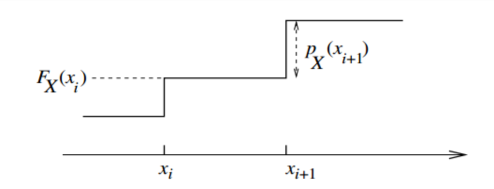
En palabras, (1) dice que la probabilidad acumulativa $F(a)$ aumenta o permanece constante a medida que a aumenta, pero nunca disminuye; (2) dice que la probabilidad acumulada está siempre entre 0 y 1; (3) dice que a medida que a se hace muy grande, se hace cada vez más cierto que $X \leq a$ y cuando a se vuelve muy negativo se hace cada vez más cierto que $X > a$.

Característica de la función de distribución acumulativa

Para una variable aleatoria discreta, $F_X(x)$ tiene la apariencia de escalera. En cada uno de los puntos x_i, tiene un salto positivo igual a $p_X(x_i)$. Entre estos puntos, es decir, en el intervalo $[x_i, x_i + 1)$, tiene un valor constante. En otras palabras, tenemos lo siguiente:

$$F_X(x) = F_X(x_i) \quad \text{para} \quad x_i \leq x < x_{i+1}$$
$$F_X(x_{i+1}) = F_X(x_i) + p_X(x_{i+1}).$$

Como se muestra en la siguiente figura:



Distribuciones de Bernoulli

Modelo: La distribución de Bernoulli modela ensayos en un experimento que puede resultar en éxito o fracaso. Esta es la distribución más importante y también es la más simple. Una variable aleatoria X tiene una distribución de Bernoulli con parámetro p si:

- X toma los valores 0 y 1.
- $\mathbb{P}(X=1)=p$ y $\mathbb{P}(X=0)=1-p$.

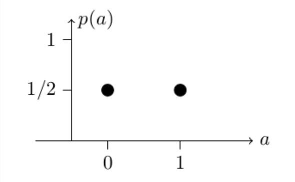
Escribiremos $X \sim \text{Bernoulli}(p)$ o $\text{Ber}(p)$, que se lee: X sigue una distribución de Bernoulli con el parámetro p. Un modelo simple para la distribución de Bernoulli consiste en lanzar una moneda con probabilidad p de caras, con $X=1$ en las caras y $X=0$ en los sellos.

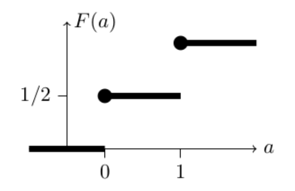
La terminología general es decir que X es 1 en los éxitos y 0 en los fracasos, con el éxito y el fracaso definido por el contexto.

Distribuciones de Bernoulli

Aquí se muestran la tabla y los gráficos del PMF y del CDF para la distribución de **Bernoulli**(1/2)

valor a :	0	1
PMF p(a):	1/2	1/2
CDF F(a):	1/2	1

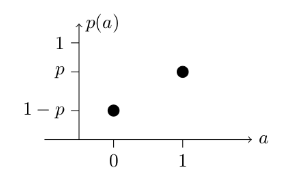


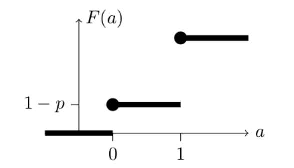


Distribuciones de Bernoulli(1)

En general los gráficos del PMF y el CDF para una distribución de Bernoulli, son de la forma:

valor a :	0	1
PMF p(a):	1 -p	p
CDF F(a):	1 -p	1





Distribuciones de binomiales

La distribución Binomial(n, p), o Bin(n, p), modela el número de éxitos en n ensayos independientes de Bernoulli(p).

Aquí hay una jerarquía. Un solo ensayo de Bernoulli es, digamos, un lanzamiento de una moneda. Un solo ensayo binomial consta de n ensayos de Bernoulli. Para el lanzamiento de una moneda, el espacio muestral para un ensayo de Bernoulli es $\{H, T\}$. El espacio muestral para un ensayo binomial son todas las secuencias de caras y sellos de longitud n. Del mismo modo, una variable aleatoria de Bernoulli toma los valores 0 y 1 y una variable aleatoria binomial toma los valores $0, 1, 2, \ldots n$.

- **Ejemplo 7** Binomial(1,p) es lo mismo que Bernoulli(p).
- **Ejemplo 8** El número de caras en n lanzamientos de una moneda con probabilidad p de caras sigue una distribución Binomial(n, p).

Distribuciones de Binomiales(1)

Describimos $X \sim \text{Binomial}(n, p)$ dándonos sus valores y probabilidades. Para la notación usaremos k para indicar un número arbitrario entre 0 y n.

Recordamos que `n elige k` $= \binom{n}{k}$, es el número de maneras de elegir k cosas de una coleccin de n cosas y tiene la fórmula

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

(También se denomina coeficiente binomial).

Aquí hay una tabla para la PMF de una variable aleatoria $\text{Binomial}(n, k)$.

valor a :	0	1	...	k	...	n
PMF p(a):	$(1-p)^n$	$\binom{n}{1} p^1 (1-p)^{n-1}$	...	$\binom{n}{k} p^1 (1-p)^{n-k}$	...	p^n

Ejemplo 9: ¿Cuál es la probabilidad de 3 o más caras en 5 lanzamientos de una moneda ?.

Explicación de las probabilidades binomiales

Para concretar, sean $n = 5$ y $k = 2$ (el argumento para valores arbitrarios n y k es idéntico). Entonces $X \sim \text{Binomial}(5, p)$ y queremos calcular $p(2)$. El largo camino para calcular $p(2)$ es enumerar todas las maneras de obtener exactamente 2 caras en 5 lanzamientos de monedas y sumar sus probabilidades. La lista tiene 10 entradas:

CCSSS, CSCSS, CSSCS, CSSSC, SCCSS, SCSCS, SSSSC, SSCCS, SSCSC, SSSCC

Cada entrada tiene la misma probabilidad de ocurrir, a saber:

$$p^2(1-p)^3.$$

Esto se debe a que cada una de las dos caras tiene probabilidad p y cada una de los 3 sellos tiene probabilidad $1 - p$. Debido a que los lanzamientos individuales son independientes, podemos multiplicar las probabilidades.

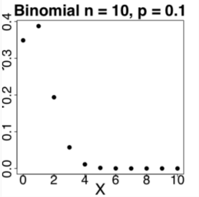
Por tanto, la probabilidad total de exactamente 2 caras es la suma de 10 probabilidades idénticas, es decir, $p(2) = 10p^2(1-p)^3$, como se muestra en la tabla.

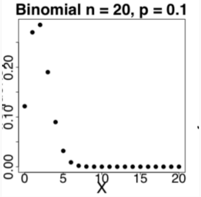
Explicación de las probabilidades binomiales(1)

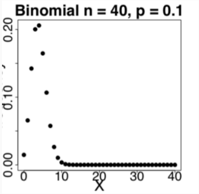
Esto nos guia, la manera más corta de hacer el cálculo. Tenemos que contar el número de secuencias con exactamente 2 caras. Para hacer esto necesitamos elegir 2 de los lanzamientos para ser caras y los 3 restantes para ser sellos. El número de tales secuencias es el número de formas de elegir 2 de 5 cosas, es decir, $\binom{5}{2}$.

Dado que cada secuencia tiene la misma probabilidad, $p^2(1-p)^3$, obtenemos la probabilidad de exactamente 2 caras $p(2) = \binom{5}{2}p^2(1-p)^3$.

Aquí algunos gráficos de la función de masa de probabilidad binomial







Distribuciones geométricas

Una distribución geométrica modela el número de sellos antes de la primera cara en una secuencia de lanzamientos de monedas (ensayos de Bernoulli).

Ejemplo 10:

- Lanzamos una moneda repetidamente. Sea X el número de sellos antes de la primera cara. Por lo tanto, X puede ser igual a 0, es decir, el primer lanzamiento es cara, $1, 2, \ldots$. En principio, toma cualquier valor entero no negativo.
- Damos un lanzamiento de sellos el valor 0, y de caras el valor 1. En este caso, X es el número de 0 antes del primer 1.
- Damos un lanzamiento de sellos el valor 1, y de caras el valor 0. En este caso, X es el número de 1 antes del primer 0.
- Llamamos a un lanzamiento de sellos un éxito y de caras un fracaso. Por lo tanto, X es el número de éxitos antes de la primer fracaso.
- Llamamos a un lanzamiento de sellos un fracaso y de caras un éxito. Por lo tanto, X es el número de fracasos antes de primer éxito.

Se puede ver estos modelos en muchos escenarios diferentes de este tipo. El lenguaje más neutral es el número de sellos antes de la primera cara.

Distribuciones geométricas(1)

La variable aleatoria X sigue una distribución geométrica con parámetro p si:

- X toma los valores $0, 1, 2, 3, \ldots$
- El PMF es dado por $p(k) = \mathbb{P}(X = k) = (1 - p)^k p$.

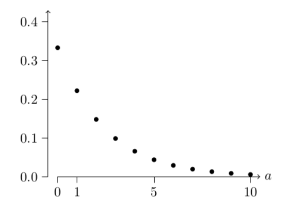
Denotamos esto por $X \sim$ Geometrica(p). En forma de tabla tenemos:

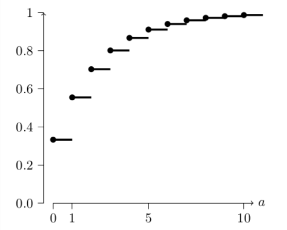
valor a :	0	1	2	3	...	k	...
PMF p(a):	p	$(1-p)p$	$(1-p)^2p$	$(1-p)^3p$	...	$(1-p)^kp$	...

La distribución geométrica es un ejemplo de una distribución discreta que toma un número infinito de valores posibles.

Distribuciones geométricas(2)

Las cosas pueden ser confusas cuando trabajamos con éxitos y fracasos, ya que tal vez queramos modelar el número de éxitos antes del primer fracaso o quizás queramos el número de fracasos antes del primer éxito. Para mantener las cosas rectas directamente se puede traducir al lenguaje neutral del número de sellos antes de la primeras caras.





El gráfico muestra el PMF y el CDF de la distribución `Geometrica(1/3)`.

Distribuciones geométricas(2)

Ejemplo 11:

Supongamos que los habitantes de una isla planean a sus familias teniendo bebés hasta que nazca la primera niña. Supongamos que la probabilidad de tener una niña con cada embarazo es de 0,5 independiente de otros embarazos, que todos los bebés sobreviven y no hay nacimientos múltiples. ¿Cuál es la probabilidad de que una familia tenga k hijos?.

Otra definición común para la distribución geométrica es el número de lanzamientos hasta las primeras caras. En este caso X puede tomar el valor de 1, es decir, el primer lanzamiento es cara, $2, 3, \dots$. Esta es sólo nuestra variable aleatoria geométrica más 1.

Distribuciones uniforme discreta

La distribución uniforme discreta modela cualquier situacion en la que todos los resultados sean igualmente probables.

$$X \sim \text{Uniforme(N)}.$$

X toma los valores $1, 2, 3, \ldots, N$, cada una con probabilidad $1/N$. Ya hemos visto esta distribución muchas veces cuando modelamos monedas ($N = 2$), dados ($N = 6$), cumpleaños ($N = 365$) y manos de póquer ($N = \binom{52}{5}$).

Aritmética con variables aleatorias

Podemos hacer aritmética con variables aleatorias. Por ejemplo, podemos sumarlas, multiplicarlas o elevarlas al cuadrarlo.

Hay una idea simple, pero extremadamente importante para contar. Esta dice que si tenemos una secuencia de números que son 0 o 1 entonces la suma de la secuencia es el número de unos.

Ejemplo 12: Consideramos la secuencia con 5 unos

1, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 0, 1, 0, 0.

Es fácil ver que la suma de esta secuencia es 5 el número de unos.

Ilustramos esta idea contando el número de caras en n lanzamientos de una moneda.

Ejemplo 13:

Lanzamos una moneda n veces. Sea X_j es 1 si el j-ésimo lanzamiento es cara y 0 si es sello. Entonces, X_j es una variable aleatoria de Bernoulli(1/2).

Sea X el número total de caras en los n lanzamiento. Asumiendo que los lanzamientos son independientes, sabemos que $X \sim \text{Binomial}(n, 1/2)$. También podemos escribir:

$$X = X_1 + X_2 + X_3 + \cdots + X_n.$$

Aritmética con variables aleatorias(continuación)

De nuevo, esto se debe a que los términos de la suma a la derecha son todos 0 o 1. Por lo tanto, la suma es exactamente el número de X_j que son 1, es decir, el número de caras.

Lo importante a ver en el ejemplo anterior es que hemos escrito la variable aleatoria binomial X, más complicada, como la suma de variables aleatorias extremadamente simples X_j. Esto nos permitirá manipular X algebraicamente.

Ejemplo 14:

Supongamos X e Y son variables aleatorias independientes con las siguientes tablas:

Valores de X	x:	1	2	3	4	
PMF	$p_X(x)$:	1/10	2/10	3/10	4/10	
Valores de Y	y:	1	2	3	4	5
PMF	$p_Y(y)$:	1/15	2/15	3/15	4/15	5/15

Continuación del ejemplo

Comprueba que la probabilidad total para cada variable aleatoria es 1.
Construye una tabla para la variable aleatoria $X + Y$.

Respuesta:

Lo primero que debemos hacer es una tabla bidimensional para el espacio muestral del producto que consiste en pares (x, y), donde x es un valor posible de X e y uno de Y.

Para ayudar a hacer el cálculo, las probabilidades para los valores de X se ponen en la columna de la derecha y las de Y se encuentran en la fila inferior. Dado que X e Y son independientes, la probabilidad para el par (x, y) es sólo el producto de las probabilidades individuales.

Continuación del ejemplo(1)

En el siguiente gráfico, las rayas diagonales muestran conjuntos de cuadrados donde $X + Y$ es el mismo. Todo lo que tenemos que hacer para calcular la tabla de probabilidad para $X + Y$ es sumar las probabilidades para cada raya.

	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	

Valores de X + Y :	2	3	4	5	6	7	8	9
PMF	1/150	4/150	10/150	20/150	30/150	34/150	31/150	20/150