Introducción a la probabilidad y estadística CM274

César Lara Avila

1 de septiembre de 2017

https://github.com/C-Lara

5. Esperanza de variables aleatorias discretas

Ejemplos iniciales

Ejemplo 1: Supongamos que tenemos un dado de seis lados marcado con cinco 3 y un 6. ¿Cuál seria el promedio de 6000 lanzamiento?

Si supiéramos el valor de cada lanzamiento, podríamos calcular el promedio sumando los 6000 valores y dividiéndolos por 6000. Sin conocer los valores, podemos calcular el promedio esperado como sigue.

Dado que hay cinco 3 y un 6 esperamos que aproximadamente $5/6$ de los lanzamientos dará 3 y $1/6$ dará 6. Suponiendo que esto sea exactamente cierto, tenemos la siguiente tabla de valores y conteos

valores:	3	6
conteo esperado:	5000	1000

El promedio de esos 6000 valores es entonces

$$\frac{5000 \cdot 3 + 1000 \cdot 6}{6000} = \frac{5}{6} \cdot 3 + \frac{1}{6} \cdot 6 = 3.5$$

Ejemplos iniciales (continuación)

Consideramos que este es el promedio esperado en el sentido de que `esperamos` que cada uno de los valores posibles ocurra con las frecuencias dadas.

Ejemplo 2: Lanzamos dos dados estándar de 6 caras. Ganamos $S/.1000$ si la suma es 2 y pierde $S/.100$ en caso contrario. ¿Cuánto se espera ganar en promedio por lanzamiento?.

La probabilidad de un 2 es $1/36$. Si tu juegas N veces, puedes `esperar` $\frac{1}{36} \cdot N$ de los lanzamientos da un 2 y $\frac{35}{36} \cdot N$ de los lanzamientos puede dar algo más. Así sus ganancias totales esperadas

$$1000 \cdot \frac{N}{36} - 100 \cdot \frac{35N}{36}.$$

Para obtener el promedio esperado por lanzamiento, dividimos el total por N:

$$\text{promedio esperado} = 1000 \cdot \frac{1}{36} - 100 \cdot \frac{5}{36} = -69.44.$$

Definición de esperanza

Se debe observar que en ambos ejemplos la suma para el promedio esperado consta de términos que son un valor de la variable aleatoria multiplicada por su probabilidad. Esto lleva a la siguiente definición.

Supongamos que X es una variable aleatoria discreta que toma valores $x_1, x_2, \ldots, x_n$ con probabilidades $p(x_1), p(x_2), \ldots, p(x_n)$. La esperanza o valor esperado de X es denotado por $\mathbb{E}(X)$ y definida como

$$\mathbb{E}(X) = \sum_{j=1}^{n} p(x_j)x_j = p(x_1)x_1 + p(x_2)x_2 + \ldots p(x_n)x_n.$$

Notas

- El valor esperado también se denomina media o promedio de X y a menudo se indica por μ.
- La esperanza proporciona una medida de la localización de la tendencia central de una variable aleatoria.
- Si todos los valores son igualmente probables entonces el valor esperado es sólo el promedio habitual de los valores.

Ejemplos

Ejemplo 3: Sea X una variable aleatoria Bernoulli(p). Como X toma los valores 0 y 1, con probabilidades p y $1-p$, la esperanza se calcula como

$$\mathbb{E}(X) = p \cdot 1 + (1-p) \cdot 0 = p.$$

Ejemplo 4: Sea X una variable aleatoria Binomial(n, p). Por definición, la media de la distribución binomial es dada por

$$\begin{aligned}
\mu = \mathbb{E}(X) = \sum_{k=0}^{n} kp(k) &= \sum_{k=1}^{n} k \frac{n!}{k!(n-k)!} p^k q^{n-k} \\
&= np \sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-k)!} p^{k-1} q^{n-k} = np \sum_{x=0}^{n-1} \binom{n-1}{x} p^x q^{n-1-x} \\
&= np(p+q)^{n-1} = np.
\end{aligned}$$

Ejemplo 5: Sea X una variable aleatoria Geometrica(p). Calculemos $\mathbb{E}(X)$. De manera usual, la esperanza de la distribución geométrica es (donde $q = 1 - p$)

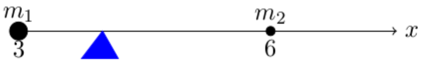
$$\mathbb{E}(X) = \sum_{k=0}^{\infty} kq^k p = pq \sum_{k=0}^{\infty} kq^{k-1} = pq \frac{1}{(1-q)^2} = \frac{q}{p}.$$

Media y centro o masa

Podriamos habernos preguntado por que usamos el nombre de función de masa de probabilidad. Esta es la razón: si colocamos un objeto de masa $p(x_j)$ en la posición x_j para cada j, entonces $\mathbb{E}(X)$ es la posición del centro de masa. Recordemos esta última noción a través de un ejemplo.

Ejemplo 6: Supongamos que tenemos dos masas a lo largo del eje x, la masa $m_1 = 500$ en la posición $x_1 = 3$ y la masa $m_2 = 100$ en la posición $x_2 = 6$. ¿Dónde se encuentra el centro de masa?.

Intuitivamente sabemos que el centro de masa está más cerca de la masa mayor.



Continuación . . .

De la física sabemos que el centro de masa es

$$\overline{x} = \frac{m_1 x_1 + m_2 x_2}{m_1 + m_2} = \frac{500 \cdot 3 + 100 \cdot 6}{600} = 3.5$$

Llamamos a esta fórmula un promedio ponderado de x_1 y x_2. Aquí x_1 se pesa más fuertemente porque tiene más masa.

Ahora veamos la definición de valor esperado $\mathbb{E}(X)$. Es una media ponderada de los valores de X siendo los pesos las probabilidades $p(x_i)$ en lugar de las masas!. Podríamos decir `que el valor esperado es el punto en el que la distribución se equilibraría`.

Propiedades algebraicas de $\mathbb{E}(X)$

El valor esperado $\mathbb{E}(X)$ de variables aleatorias es lineal.

1. Si X e Y son variables aleatorias de un espacio muestral Ω, entonces

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y).$$

2. Si a y b son constantes, entonces

$$\mathbb{E}(aX + b) = a\mathbb{E}(X) + b.$$

Ejemplo 7: Lanzamos dos dados y X sea la suma, sea X_1 el valor en el primer dado y X_2 sea el valor del segundo dado. Como $X = X_1 + X_2$ tenemos $\mathbb{E}(X) = \mathbb{E}(X_1) + \mathbb{E}(X_2)$. Se calcula que $\mathbb{E}(X_1) = \mathbb{E}(X_2) = 3.5$, por lo tanto $\mathbb{E}(X) = 7$.

Ejemplo 8: (Para las variables aleatorias infinitas la media no siempre existe.) Supongamos que X tiene un número infinito de valores según la siguiente tabla

valores x	2	2^2	2^3	$\dots$	2^k	$\dots$
PMF $p(x)$	$1/2$	$1/2^2$	$1/2^3$	$\dots$	$1/2^k$	$\dots$

Calculemos la media : $\mathbb{E}(X) = \sum_{k=1}^{\infty} 2^k \frac{1}{2^k} = \sum_{k=1}^{\infty} 1 = \infty$.

Pruebas de las propiedades algebráicas de $\mathbb{E}(X)$

La prueba de la propiedad (1) es simple, pero hay alguna sutileza en la comprensión de lo que significa añadir dos variables aleatorias. Recordemos que el valor de la variable aleatoria es un número determinado por el resultado de un experimento. Agregar X e Y significa agregar los valores de X e Y para el mismo resultado. En forma de tabla esto se parece a:

salida ω	ω_1	ω_2	ω_3	$\dots$	ω_n
valores de X	x_1	x_2	x_3	$\dots$	x_n
valores de Y	y_1	y_1	y_3	$\dots$	y_n
valores de $X+Y$	x_1+y_1	x_2+y_2	x_3+y_3	$\dots$	x_n+y_n
prob $\mathbb{P}(\omega)$	$\mathbb{P}(\omega_1)$	$\mathbb{P}(\omega_2)$	$\mathbb{P}(\omega_3)$	$\dots$	$\mathbb{P}(\omega_n)$

La prueba de (1) se sigue inmediatamente:

$$\mathbb{E}(X+Y) = \sum (x_i + y_i)\mathbb{P}(\omega_i) = \sum x_i \mathbb{P}(\omega_i) + \sum y_i \mathbb{P}(\omega_i) = \mathbb{E}(X) + \mathbb{E}(Y).$$

continuación ...

La prueba de la propiedad (2) toma una línea.

$$\mathbb{E}(aX + b) = \sum p(x_i)(ax_i + b) = a \sum p(x_i)x_i + b \sum p(x_i) = a\mathbb{E}(X) + b.$$

El término b en la última expresión se debe a que $\sum p(x_i) = 1$.

Ejemplo 9: Michael Jordan, el mejor jugador de baloncesto de todos los tiempos, hizo el 80 % de sus tiros libres. En un juego cuál es el número esperado que haría antes de su primera falta.

Aquí hay un ejemplo donde queremos el número de éxitos antes de la primer fracaso. Usando el lenguaje neutral de caras y colas: el éxito es sello (probabilidad $1 - p$) y el fracaso es cara (probabilidad $= p$). Por lo tanto $p = .2$ y el número de sellos (lanzamientos libres) antes de la primera cara (tiro libre pérdido) es modelado por un $X \sim$ Geometrica(.2). Luego por el Ejemplo 5, tenemos,

$$\mathbb{E}(X) = \frac{1 - p}{p} = \frac{.8}{.2} = 4.$$

Variable aleatoria indicador

La variable aleatoria indicador I_A (o $I(A)$) para un evento A se define como 1 si A ocurre y 0 en caso contrario. Así A es una variable aleatoria de Bernoulli, donde el éxito se define como A `ocurre` y el fracaso se define como A `no ocurre`.

Algunas propiedades útiles del indicador se resumen a continuación.

- $(I_A)^k = I_A$ para algún entero k.
- $I_{A^c} = 1 - I_A$.
- $I_{A \cap B} = I_A I_B$.
- $I_{A \cup B} = I_A + I_B - I_A I_B$.

Hay una correspondencia uno a uno entre los eventos y la variable aleatoria indicador y la probabilidad de un evento A es el valor esperado de su variable aleatoria indicador I_A :

$$\mathbb{P}(A) = \mathbb{E}(I_A). \tag{1}$$

Un problema del Putnam

Sea $S = \{1, 2, \ldots, n\}$ para algún entero $n > 1$. Decimos que una permutación π de S tiene un máximo local en $j \in S$ si

- $\pi(j) > \pi(j+1)$ para $j = 1$;
- $\pi(j-1) < \pi(j)$ y $\pi(j) > \pi(j+1)$ para $1 < j < n$;
- $\pi(j-1) < \pi(j)$ para $j = n$.

Por ejemplo si $n = 5$ y π toma valores en $1, 2, 3, 4, 5$ de $2, 1, 4, 5, 3$, entonces π tiene un máximo local de 2 en $k = 1$ y un máximo local de 5 en $k = 4$. Para $n \geq 2$, ¿cuál es el número promedio de máximos locales de una permutación aleatoria de $1, 2, \ldots, n$, con todos las $n!$ permutaciones igualmente probables?

Sea $I_1, \ldots, I_n$ las variables aleatorias indicadoras, donde I_j es 1 si hay un máximo local en la posición j y 0 en otros casos. Estamos interesados en la esperanza de $\displaystyle\sum_{j=1}^{n} I_j$.

Para $1 < j < n, \mathbb{E}(I_j) = 1/3$, desde que tener un máximo local en j es equivalente a a_j siendo el mayor de a_{j-1}, a_j, a_{j+1}, que tiene probabilidad $1/3$, desde que todos los órdenes son igualmente probables. Para $j = 1$ o $j = n$, tenemos $\mathbb{E}(I_j) = 1/2$, desde que hay solo un vecino. Así por linealidad,

$$\mathbb{E}\left(\sum_{j=1}^{n} I_j\right) = 2 \cdot \frac{1}{2} + (n-2) \cdot \frac{1}{3} = \frac{n+1}{3}.$$

Esperanza de funciones de una variable aleatoria

Si X es una variable aleatoria discreta, tomando los valores $x_1, x_2, \ldots$ y h es una función, $h(X)$ es una nueva variable aleatoria. La esperanza o valor esperado es

$$\mathbb{E}(h(X)) = \sum_j h(x_j)p(x_j).$$

Ejemplo 10: Sea X el valor de un lanzamiento de un dado y sea $Y = X^2$. Encuentra $E(Y)$.

Desde que hay un número pequeño de valores, podemos hacer una tabla.

X	1	2	3	4	5	6
Y	1	4	9	16	25	36
prob	1/6	1/6	1/6	1/6	1/6	1/6

Esperanza de funciones de una variable aleatoria (1)

Observe que la probabilidad para cada valor de Y es la mismo que el valor de X correspondiente. Asi que,

$$\mathbb{E}(Y) = \mathbb{E}(X^2) = 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + \cdots + 6^2 \cdot \frac{1}{6} = 15.167.$$

Ejemplo 11: Lanzamos dos dados y sea X su suma. Supongamos que la funcin de pago está dada por $Y = X^2 - 6X + 1$. ¿Es una buena apuesta?.

En efecto $\mathbb{E}(Y) = \sum_{j=2}^{12} (j^2 - 6j + 1)p(j)$, donde $p(j) = \mathbb{P}(X = j)$.

X	2	3	4	5	6	7	8	9	10	11	12
Y	-7	-8	-7	-4	1	8	17	28	41	56	73
prob	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Aquí el valor es $\mathbb{E}(Y) = 13.833$. Para responder a la pregunta anterior: ya que la ganancia esperada es positiva, parece una apuesta que vale la pena tomar.

Pregunta: Si $Y = h(X)$ hace que $\mathbb{E}(Y) = h(\mathbb{E}(X))$?.

Esperanza de funciones de una variable aleatoria (2)

Dado una variable aleatoria discreta X con un conjunto posible de valores A y la función de masa de probabilidad $p_X(x)$, la esperanza de un nueva variable aleatoria Y, que es una función de X, es decir $Y = h(X)$ se puede escribir como

$$\mathbb{E}(Y) = \mathbb{E}(h(X)) = \sum_{x \in A} h(x) p_X(x).$$

En efecto, sea Ω un espacio muestral. Sea $h : \mathbb{R} \to \mathbb{R}$, una función de variable real y $X : \Omega \to A \subseteq \mathbb{R}$ es una variable aleatoria con el conjunto de posibles de valores A. Como se sabe, $h(X)$, la composición de g y X es una función desde Ω al conjunto $h(A) = \{h(x) : x \in A\}$. Así $h(X)$ es una variable aleatoria con un posible conjunto de valores $g(A)$. Ahora por definición de esperanza

$$\mathbb{E}(h(X)) = \sum_{z \in h(A)} z \mathbb{P}(h(X) = z)$$

Un resultado

Sea X una variable aleatoria no negativa entera. Sea F el *CDF* de X y $G(x) = 1 - F(x) = \mathbb{P}(X > x)$. La función G es llamada función de supervivencia de X. Entonces

$$\mathbb{E}(X) = \sum_{n=0}^{\infty} G(n).$$

En efecto, para el caso en X es acotada (existe un número entero no negativo b , tal que X es siempre como máximo b), podemos representar como la suma de variables aleatorias indicadoras, $X = I_1 + I_2 + \cdots + I_b$, donde $I_n = I(X \geq n)$.

Entonces se tiene, por linealidad, la ecuación (1) y el hecho que $\{X \geq k\}$ es el mismo evento que $\{X > k - 1\}$.

$$\mathbb{E}(X) = \sum_{k=1}^{b} \mathbb{E}(I_k) = \sum_{k=1}^{b} \mathbb{P}(X \geq k) = \sum_{n=0}^{b-1} \mathbb{P}(X > n) = \sum_{n=0}^{\infty} G(n).$$