Probabilidades y Estadística

Clase 5
Variable Aleatoria

Nicolás Araya Caro

Universidad Diego Portales
Escuela de Informática y Telecomunicaciones

14 de septiembre de 2023

Contenidos

Definición

- Para un ***EM*** dado, una ***Variable Aleatoria (V.A)*** es cualquier regla que asocia un número con cada resultado en el ***EM***.
- En otras palabras, una **VA** es una función cuyo dominio es el espacio muestral y cuyo rango es el conjunto de los números reales.
- Se utiliza letras en mayúscula para denotar una variable aleatoria. Por ejemplo la variable $X(s) = x$ significa "x es el valor asociado con el resultado s por medio de la variable aleatoria X".

Ejemplo

Considere el experimento en el cual se marca un número telefónico en cierto código de área con un marcador de números aleatorio (tales dispositivos los utilizan de forma extensa en organizaciones encuestadoras) y se define la variable aleatoria Y como:

$$Y = \begin{cases} 1, & \text{si el número no aparece en el directorio} \\ 0, & \text{si el número aparece en el directorio} \end{cases}$$

Por ejemplo, si el teléfono 56709843 aparece en el directorio, entonces $Y(56709843) = 0$. En cambio, si 569333445 no aparece, entonces $Y(569333445) = 1$.

Nota:

Cualquier variable aleatoria cuyos únicos valores posibles son 0 y 1 se llama **Variable aleatoria de Bernoulli**.

Tipos de variable aleatoria

Existen 2 tipos de Variables Aleatorias:

- **Discretas:** Los valores posibles constituyen un conjunto finito secuencial o una lista infinita contable (1,2,3,4...).
- **Continua:** Una variable aleatoria es continua si ambas de las siguientes condiciones aplican:
  1. Su conjunto de valores posibles se compone de o todos los números que hay en un solo intervalo sobre la línea de numeración ($-\infty$ a ∞). o todos los números en una unión excluyente de dichos intervalos (p. ej., $[0, 5] \cup [30, 60]$).
  2. Ningún valor posible de la variable aleatoria tiene probabilidad positiva, esto es, $P(X = c) = 0$ con cualquier valor posible de c.

Contenidos

Distribución de Probabilidad

Las probabilidades asignadas a varios resultados en EM determinan a su vez las probabilidades asociadas con los valores de cualquier variable aleatoria X particular. La *distribución de probabilidad* de X dice cómo está distribuida (asignada) la probabilidad total de 1 entre los varios posibles valores de X.

Función de masa

La **Distribución de probabilidad** o **Función de masa de probabilidad** (fmp) de una variable discreta se define para cada número x como $p(x) = P(X = x) = P(\text{todas las } s \in EM.\ X(s) = x)$.

Ejemplo

Considere si la siguiente persona que compre una computadora en una librería universitaria comprará un modelo portátil o uno de escritorio. Sea:

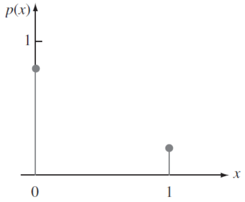
$$X = \begin{cases} 1, & \text{si el cliente compra una computadora portátil} \\ 0, & \text{si el cliente compra una computadora de escritorio} \end{cases}$$

Si 20 % de todas las compras durante esa semana seleccionan una portátil, la función masa de probabilidad de X es:

- $p(0) = \mathrm{P}(X = 0) = \mathrm{P}(\text{el siguiente cliente compra un modelo de escritorio}) = 0.8$
- $p(1) = \mathrm{P}(X = 1) = \mathrm{P}(\text{el siguiente cliente compra un modelo portátil}) = 0.2$
- $p(x) = \mathrm{P}(X = x) = 0$ con $x \neq 0$ o 1

Una descripción equivalente es:

$$p(x) = \begin{cases} 0.8, & \text{si } x = 0 \\ 0.2, & \text{si } x = 1 \\ 0, & \text{si } x \neq 0 \text{ o } 1 \end{cases}$$



Gráfica lineal de la función de masa de probabilidad en el ejemplo

Ejemplo II

Considere seleccionar al azar un estudiante de entre los 15000 inscritos en el semestre actual en la Universidad UDP. Sea X el número de cursos en los cuales el estudiante seleccionado está inscrito y suponga que X tiene la siguiente función masa de probabilidad.

x	1	2	3	4	5	6	7
$p(x)$	0.01	0.03	0.13	0.25	0.37	0.17	0.02

continuación

- Una forma de ver esta situación es pensar en la población como compuesta de 15.000 individuos, cada uno con su propio valor X; la proporción con cada valor de X está dada por p(x).
- Otra alternativa es olvidarse de los estudiantes y pensar en la población como compuesta de los valores X: Existen algunos 1 en la población, algunos 2, ... y finalmente algunos 7. La población se compone de los números, 1, 2, ... ,7 (Población discreta) y p(x) da un modelo para la distribución de los valores de población.

Parámetro de una distribución

Supongase que p(x) depende de la cantidad que puede ser asignada a cualquiera de un número de valores posibles, y cada valor determina una distribución de probabilidad diferente. Tal cantidad se le conoce como **Parámetro** de distribución.

Ejemplo:

un parámetro α donde $0<\alpha<1$:

$$p(x;\alpha) = \begin{cases} 1-\alpha, & \text{si } x = 0 \\ \alpha, & \text{si } x = 1 \end{cases}$$

El ejemplo de arriba es la generalización de la función de masa de Bernuolli. A muchas de estas se le conoce como *familia de distribuciones de Bernoulli.*

A partir de un tiempo fijo, se observa el sexo de cada niño recién nacido en un hospital hasta que nace un varón (B). Sea p = P(B) y suponga que los nacimientos sucesivos son independientes y defina la variable aleatoria X como X número de nacimientos observados. Entonces:

$$p(1) = P(X = 1) = P(B) = p$$
$$p(2) = P(X = 2) = P(GB) = P(G)P(B) = (1 - p)p$$
$$p(3) = P(X = 3) = P(GGB) = P(G)P(G)P(B) = (1 - p)^2 p$$

Ejemplo - Continuación

Continuando de esta manera, emerge una fórmula general:

$$p(x) = \begin{cases} (1-p)^{x-1}p, & x = 1, 2, 3, \dots \\ 0, & \text{en caso contrario} \end{cases}$$

La cantidad p en la expresión anterior representa un número entre 0 y 1 y es un parámetro de la distribución de probabilidad.

- Para algún valor fijo x, a menudo se desea calcular la probabilidad de que el valor observado de X será cuando mucho x.
- Por ello se la **función de distribución acumulativa** (fda) F(x) de una variable aleatoria discreta X con función masa de probabilidad *p(x)* se define para cada número x como:

$$F(x) = P(X \leq x) = \sum_{y:y \leq x} p(y)$$

Para cualquier número x, F(x) es la probabilidad de que el valor observado de X será cuando mucho x.

Ejemplo

Considere la función masa de probabilidad de Y (el número de determinaciones de solicitudes 500 en un servicio API REST) es:

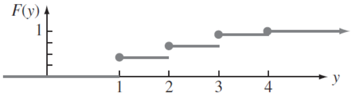
y	1	2	3	4
$p(y)$	0.4	0.3	0.2	0.1

Primero se determina F(y) para cada uno de los valores posibles del conjunto (1, 2, 3, 4):

$$F(1) = P(Y \leq 1) = P(Y = 1) = 0.4$$
$$F(2) = P(Y \leq 2) = P(Y = 1 \text{ o } 2) = p(1) + p(2) = 0.7$$
$$F(3) = P(Y \leq 3) = P(Y = 1 \text{ o } 2 \text{ o } 3) = p(1) + p(2) + p(3) = 0.9$$
$$F(4) = P(Y \leq 4) = P(Y = 1 \text{ o } 2 \text{ o } 3 \text{ o } 4) = p(1) + p(2) + p(3) + p(4) = 1$$

Ahora con cualquier otro número y, F(y) será igual al valor de F con el valor más próximo posible de Y a la izquierda de y. Por ejemplo, F(2,7) $P(Y \leq 2.7) = P(Y \leq 2) = 0.7$ y $\text{F}(3.999) = \text{F}(3) = 0.9$. La función de distribución acumulativa es por lo tanto:

$$F(y) = \begin{cases} 0 \text{ si}, & y < 1 \\ 0.4 \text{ si}, & 1 \leq y < 2 \\ 0.7 \text{ si}, & 2 \leq y < 3 \\ 0.9 \text{ si}, & 3 \leq y < 4 \\ 1 \text{ si}, & 4 \leq y \end{cases}$$



Gráfica de la función de distribución acumulativa del ejemplo

Para una variable aleatoria discreta X, la gráfica de F(x) mostrará un salto con cada valor posible de X y será plana entre los valores posibles. Tal gráfica se conoce como función escalonada.

Proposición, de fmp a fda

Para dos números cualesquiera a y b con $a \leq b$.

$P(a \leq b) = F(b) - F(a-)$

donde "a-" representa el valor posible de X más grande que es estrictamente menor que a. En particular, si los únicos valores posibles son enteros y si a y b son enteros, entonces

$P(a \leq b) = P(X = a \text{ o } a + 1 \text{ o}... \text{ o } b) = F(b) - F(a - 1)$

Con a - b se obtiene $P(X = a) = F(a) - F(a - 1)$ en este caso.

Sea X el número de días de ausencia por enfermedad tomados por un empleado seleccionado al azar de una gran compañía durante un año particular. Si el número máximo de días de ausencia por enfermedad permisibles al año es de 14, los valores posibles de X son 0,1,...,14. Con $F(0) = 0.58$, $F(1) = 0.72$, $F(2) = 0.76$, $F(3) = 0.81$, $F(4) = 0.88$ y $F(5) = 0.94$. si calculamos $P(2 \leq X \leq 5)$ entonces:

$$P(2 \leq X \leq 5) = P(X = 2, 3, 4, \text{o } 5) = F(5) - F(1) = 0.22$$

y para $P(X = 3)$:

$$P(X = 3) = F(3) - F(2) = 0.05$$

Contenidos

Sea X una variable aleatoria discreta con un conjunto de valores posibles D y una función masa de probabilidad p(x). El valor esperado o valor medio de X, denotado por E(X) o μ_X, es:

$$E(X) = \mu_X = \sum_{x \in D} x \cdot p(x)$$

Ejemplo

Sea $X = 1$ si un componente seleccionado al azar necesita servicio de garantía y 0 si no. Entonces X es una variable aleatoria de Bernoulli con función masa de probabilidad:

$$p(x) = \begin{cases} (1-p) \text{ si}, & x = 0 \\ p \text{ si}, & x = 1 \\ 0 \text{ si}, & x \neq 0, 1 \end{cases}$$

a partir de la cual $E(X) = 0 \cdot p(0) + 1 \cdot p(1) = 0(1-p) + 1(p) = p$. Es decir, el valor esperado de X es exactamente la probabilidad de que X tome el valor 1. Si se conceptualiza una población compuesta de ceros en la proporción de $1 - p$ y unos en la proporción de p, entonces el promedio de la proporción es $\mu = p$.

Valor esperado de una función

Si la variable aleatoria X tiene un conjunto de posibles valores D y una función masa de probabilidad p(x), entonces el valor esperado de cualquier función *h(X)*, denotada por E[*h(X)*] o *h(X)*, se calcula con:

$$E[h(X)] = \sum_D h(x) \cdot p(x)$$

Esto es, E[h(X)] se calcula del mismo modo que E(X), excepto que h(x) sustituye a x.

Ejemplo

Una tienda de computadoras adquirió tres computadoras de un tipo a \$500 cada una. Las venderá a \$1000 cada una. El fabricante se comprometió a readquirir cualquier computadora que no se haya vendido después de un periodo especificado a \$200 cada una.

Sea X el número de computadoras vendidas y suponga que $p(0) = 0.1$, $p(1)=0.2$, $p(2)=0.3$ y $p(3)=0.4$. Con $h(X)$ denotando la utilidad asociada con la venta de X unidades, la información dada implica que $h(X) =$ ingreso - costo $= 1000X + 200(3 - X) - 1500 = 800X - 900$.

Continuación

La utilidad esperada es entonces:

$$
\begin{aligned}
E[h(X)] &= h(0)\cdot p(0) + h(1)\cdot p(1) + h(2)\cdot p(2) + h(3)\cdot p(3) \\
&= (900)(0.1) + (100)(0.2) + (700)(0.3) + (1500)(0.4) = \$700
\end{aligned}
$$

Reglas de Valor esperado

La función de interés h(X) en muchas ocasiones es una función lineal $aX + b$. En este caso, E[h(X)] es fácil de calcular a partir de E(X):

$$E(aX + b) = a \cdot E(X) + b$$

Contenidos

Varianza de X

Sea p(x) la función masa de probabilidad de X y μ su valor esperado. En ese caso **la varianza de X,** denotada por V(X) o σ^2 es:

$$\sigma^2 = V(X) = \sum_{D} (x-\mu)^2 \cdot p(x) = E[(x-\mu)^2]$$

La desviación estándar (DE) de X es: $\sigma_X = \sqrt{\sigma_X^2}$

Forma Abreviada

Con un poco de álgebra, El número de operaciones aritméticas necesarias para calcular σ^2 pueden reducirse:

$$V(X) = \sigma^2 = \sum_{D} [x^2 \cdot p(x)] - \mu^2 = E[(X^2)] - [E(X)]^2$$

Reglas de Varianza

Existe una abreviación similar a lo que ocurre con el valor esperado:

$$V(aX+b) = \sigma^2_{aX+b} = a^2 \cdot \sigma^2_X \text{ and } \sigma_{aX+b} = |a| \cdot \sigma_x$$

El valor absoluto es necesario porque a podría ser negativa, no obstante una desviación estándar no puede serlo. Casi siempre la multiplicación por a corresponde a un cambio de la unidad de medición (p. ej., kg a lb o dólares a euros). De acuerdo con la primera relación, la desviación estándar en la nueva unidad es la desviación estándar original multiplicada por el factor de conversión. La segunda relación dice que la adición o sustracción de una constante no impacta la variabilidad; simplemente desplaza la distribución a la derecha o izquierda (a nivel gráfico).