TEMA 4 VARIABLES ALEATORIAS

Pablo Buenestado

Curso 2020-2021 Otoño

Departamento de Matemáticas (UPC)

Índice

Esquema

Empezamos con la definición de una variable aleatoria discreta.

Definición

Una variable aleatoria es discreta si sus valores posibles constituyen un conjunto discreto.

Lo anterior significa que si los valores posibles se ordenan, hay una separación entre cada valor y el próximo.

El conjunto de valores posibles podría ser infinito; por ejemplo, el conjunto de todos los enteros o el conjunto de todos los enteros positivos.

Es común que los valores posibles de una variable aleatoria discreta sean un conjunto de enteros.

Para cualquier variable aleatoria discreta, si se especifica la lista de sus valores posibles junto con la probabilidad que tiene la variable aleatoria en cada uno de estos valores, entonces hemos descrito completamente la población a partir de la cual se seleccionó la variable aleatoria.

Esto se ilustra con un ejemplo.

El número de fallas en un alambre de cobre de 1 centímetro de longitud, fabricado en un proceso específico, varía de un alambre a otro.

En conjunto, el 48% de los alambres producidos no tiene falla, el 39% presenta una, el 12% contiene dos y el 1% tiene tres.

Sea X el número de fallas en una pieza de alambre seleccionada aleatoriamente. Entonces

$$P(X = 0) = 0.48,\ P(X = 1) = 0.39,\ P(X = 2) = 0.12,\ P(X = 3) = 0.01$$

La lista de valores posibles $\{0, 1, 2, 3\}$ junto con las probabilidades para cada uno de ellos, proporciona una descripción completa de la población de la que se tomó a X.

Esta descripción se conoce como función de masa de probabilidad.

Definición

La función de masa de probabilidad de una variable aleatoria discreta X es la función $p(x) = P(X = x)$.

A veces a la función de masa de probabilidad se le llama distribución de probabilidad.

Por tanto, para la variable aleatoria X que representa el número de fallas en una longitud de alambre, la función masa de probabilidad viene dada por $p(0) = 0.48$, $p(1) = 0.39$, $p(2) = 0.12$, $p(3) = 0.01$ y $p(x) = 0$ para cualquier valor de x diferente de 0, 1, 2 o 3.

La tabla correspondiente es

X	$p(x)$
0	0.48
1	0.39
2	0.12
3	0.01

Observemos que si se suman los valores de la función de masa de probabilidad sobre todos los valores posibles de X, la suma es igual a 1.

Esto es cierto para cualquier función de masa de probabilidad.

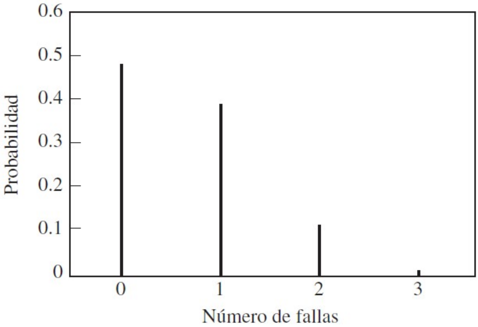
La razón es que al sumar los valores de una función de masa de probabilidad sobre todos los valores posibles de la variable aleatoria correspondiente, se obtiene la probabilidad de que la variable aleatoria es igual a uno de sus valores posibles y esta probabilidad es siempre igual a 1.

La función de masa de probabilidad se puede representar por un diagrama en el cual se dibuja una recta vertical para cada uno de los valores posibles de la variable aleatoria.

Las alturas de las rectas son iguales a las probabilidades de los valores correspondientes.

La interpretación física de este diagrama es que cada recta representa una masa igual a su altura.

En la figura se muestra un diagrama de la función de masa de probabilidad de la variable aleatoria X.



Función de distribución acumulativa de una variable aleatoria discreta

La función de masa de probabilidad especifica la probabilidad de que una variable aleatoria sea igual a un valor determinado.

La función de distribución acumulativa especifica la probabilidad de que una variable aleatoria sea menor o igual a un valor dado.

La función de distribución acumulativa de la variable aleatoria X es la función $F(x) = P(X \leq x)$.

Ejemplo

Sea $F(x)$ la función de distribución acumulativa de la variable aleatoria X que representa el número de fallas en un alambre elegido aleatoriamente. Determinemos $F(2)$ y $F(1.5)$.

Solución

Puesto que $F(2) = P(X \leq 2)$ se necesita encontrar $P(X \leq 2)$. Con dicho propósito se suman las probabilidades de los valores de X que son menores o iguales a 2, que en este ejemplo lo cumplen los valores 0, 1 y 2. Por tanto,

$$F(2) = P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$$
$$F(2) = 0.48 + 0.39 + 0.12 = 0.99$$

Ahora $F(1.5) = P(X \leq 1.5)$. Por tanto, para calcular $F(1.5)$ se debe sumar las probabilidades para los valores de X que son menores o iguales a 1.5, que son 0 y 1. Por lo que,

$$F(1.5) = P(X \leq 1.5) = P(X = 0) + P(X = 1) = 0.48 + 0.39 = 0.87$$

En general, para cualquier variable aleatoria discreta X, la función de distribución acumulativa $F(x)$ se puede calcular sumando las probabilidades de todos los valores posibles de X que son menores o iguales a x.

Observemos que $F(x)$ está definido para cualquier número x, no solo para los valores posibles de X.

Resumen

Sea X una variable aleatoria discreta. Entonces

- La función de masa de probabilidad de X es la función

$$p(x) = P(X = x)$$

- Si sumamos las probabilidades de todos los valores posibles de X:

$$\sum_x p(x) = \sum_x P(X = x) = 1$$

- La función de distribución acumulativa de X es la función

$$F(x) = P(X \leq x)$$

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

Ejemplo

Dibujamos la función de distribución acumulativa $F(x)$ de la variable aleatoria X que representa el número de fallas en un alambre elegido aleatoriamente.

Solución

Primero calculamos $F(x)$ para cada uno de los valores posibles de X, que son 0, 1, 2 y 3.

$$F(0) = P(X \leq 0) = 0.48$$

$$F(1) = P(X \leq 1) = 0.48 + 0.39 = 0.87$$

$$F(2) = P(X \leq 2) = 0.48 + 0.39 + 0.12 = 0.99$$

$$F(3) = P(X \leq 3) = 0.48 + 0.39 + 0.12 + 0.01 = 1$$

Solución (Continuación)

Para cualquier valor x, calculamos $F(x)$ sumando las probabilidades de todos los valores posibles de X que son menores o iguales a x.

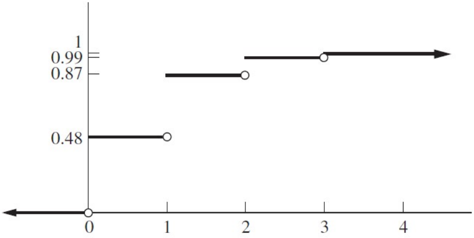
Por ejemplo, si $1 \leq x < 2$, los valores posibles de X que son menores o iguales a x son 0 y 1, por lo que

$$F(x) = P(X = 0) + P(X = 1) = F(1) = 0.87$$

$$F(x) = \begin{cases} 0 & x < 0 \\ 0.48 & 0 \leq x < 1 \\ 0.87 & 1 \leq x < 2 \\ 0.99 & 2 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$

Solución (Continuación)

En la figura se presenta la gráfica de $F(x)$.



Para una variable aleatoria discreta, la gráfica de $F(x)$ consta de una serie de rectas horizontales (llamadas "pasos") con saltos en cada uno de los valores posibles de X.

Observemos que el tamaño del salto en cualquier punto x es igual al valor de la función de masa de probabilidad $p(x) = P(X = x)$.

Media y varianza para variables aleatorias discretas

En Estadística Descriptiva se presenta la media de una muestra ($\bar{x}$), que es aproximadamente igual a la componente horizontal del centro de masas de un histograma de la muestra, que es el punto en el eje x en el cual el histograma mantendría el equilibrio si se sostuviera desde allí.

Por analogía, se define la media poblacional como la componente horizontal del centro de masas de la gráfica de su función de masa de probabilidad de una variable aleatoria discreta.

La media poblacional de una variable aleatoria X también se puede llamar esperanza, o valor esperado, de X y se denota por μ_X, por $E(X)$, o simplemente por μ.

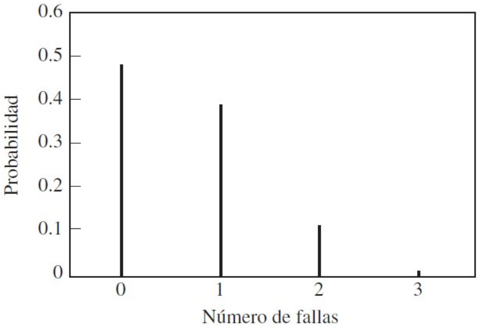
A veces se puede eliminar la palabra "poblacional" y sólo se hará referencia a la media poblacional como la media.

Ejemplo

Determinemos la media de la variable aleatoria X que representa el número de fallas en una pieza de alambre elegida aleatoriamente.

Solución

La media es el centro de masas de la gráfica de la función de masa de probabilidad.



Solución (Continuación)

El centro de masas se calcula multiplicando la altura de cada recta por su valor en el eje horizontal y después se suman los productos.

Los valores de la variable aleatoria son 0, 1, 2 y 3.

Las alturas son $P(X=0)=0.48$, $P(X=1)=0.39$, $P(X=2)=0.12$ y $P(X=3)=0.01$.

La media es, por tanto,

$$\mu_X = 0 \cdot P(X=0) + 1 \cdot P(X=1) + 2 \cdot P(X=2) + 3 \cdot P(X=3)$$

$$\mu_X = 0 \cdot 0.48 + 1 \cdot 0.39 + 2 \cdot 0.12 + 3 \cdot 0.01 = 0.66$$

En general, la media de una variable aleatoria discreta se encuentra al multiplicar cada valor posible de la variable aleatoria por su probabilidad y después se suman.

Definición

Sea X una variable aleatoria discreta con función de masa de probabilidad $p(x) = P(X = x)$. La media de X es

$$\mu_X = \sum_x x \cdot P(X = x)$$

donde el sumatorio se hace sobre todos los valores posibles de X.

A veces la media de X se llama esperanza, o valor esperado, de X y también se denota por $E(X)$ o por μ.

En Estadística Desriptiva se mostró que la varianza de una muestra es aproximadamente igual al momento de inercia del histograma de la muestra alrededor de la media de la muestra.

Se define a la varianza poblacional de una variable aleatoria discreta como el momento de inercia de la gráfica de su función de masa de probabilidad con respecto a la media poblacional μ.

La varianza poblacional de una variable aleatoria X con frecuencia se denomina simplemente varianza de X.

Se puede denotar por σ_X^2, por $V(X)$, $Var(X)$, o simplemente por σ^2.

Para calcularla se multiplica la altura de cada recta del diagrama de la función de masa de probabilidad por el cuadrado de su distancia horizontal a la media poblacional y después se suman los productos.

Es más fácil de comprender esto cuando se presenta en una fórmula:

$$\sigma_X^2 = \sum_x (x - \mu_X)^2 P(X = x)$$

Realizando un poco de álgebra se puede obtener una fórmula alternativa

$$\sigma_X^2 = \sum_x x^2 \cdot P(X = x) - \mu_X^2$$

También se define la desviación estándar poblacional como la raíz cuadrada de la varianza poblacional.

Se denota a la desviación estándar de la población de una variable aleatoria X por σ_X o simplemente por σ.

Como con la media, a veces se eliminará la palabra "poblacional" y sólo se denominará a la varianza poblacional y a la desviación estándar poblacional como la varianza y la desviación estándar, respectivamente.

Resumen

Sea X una variable aleatoria discreta con función de masa de probabilidad $p(x) = P(X = x)$. Entonces

- La varianza de X está dada por

$$\sigma_X^2 = \sum_x (x - \mu_X)^2 P(X = x)$$

- Una fórmula alternativa para la varianza es

$$\sigma_X^2 = \sum_x x^2 \cdot P(X = x) - \mu_X^2$$

- La varianza de X también se puede denotar por $V(X)$, $Var(X)$ o por σ^2.
- La desviación estándar es la raíz cuadrada de la varianza: $\sigma_X = \sqrt{\sigma_X^2}$

Ejemplo

Determinemos la varianza y la desviación estándar para la variable aleatoria X que representa el número de fallas de una pieza de alambre elegida aleatoriamente.

Solución

Anteriormente calculamos la media de X con un valor de $\mu_X = 0.66$.

Se encuentra la varianza a partir de la ecuación:

$$\sigma_X^2 = \sum_x (x - \mu_X)^2 P(X = x)$$

$$\sigma_X^2 = (0-0.66)^2 \cdot P(X=0) + (1-0.66)^2 \cdot P(X=1) + (2-0.66)^2 \cdot P(X=2) +$$
$$+ (3-0.66)^2 \cdot P(X=3)$$

$$\sigma_X^2 = 0.4356 \cdot 0.48 + 0.1156 \cdot 0.39 + 1.7956 \cdot 0.12 + 5.4756 \cdot 0.01 = 0.5244$$

La desviación estándar es $\sigma_X = \sqrt{0.5244} = 0.724$.

Solución (Continuación)

Usamos la fórmula alternativa para calcular la varianza de X:

$$\sigma_X^2 = \sum_x x^2 \cdot P(X = x) - \mu_X^2$$

$$\sigma_X^2 = 0^2 \cdot P(X = 0) + 1^2 \cdot P(X = 1) + 2^2 \cdot P(X = 2) + 3^2 \cdot P(X = 3) - (0.66)^2$$

$$\sigma_X^2 = 0 \cdot 0.48 + 1 \cdot 0.39 + 4 \cdot 0.12 + 9 \cdot 0.01 - 0.66^2 = 0.5244$$

Histogramas

Un histograma es una gráfica que da una idea de la ”forma” de una muestra, indicando las regiones donde los puntos de la muestra están concentrados y las regiones donde son escasos.

Construimos un histograma para las emisiones EP de 62 vehículos, presentadas en la tabla siguiente:

7.59	6.28	6.07	5.23	5.54	3.46	2.44	3.01	13.63	13.02
23.38	9.24	3.22	2.06	4.04	17.11	12.26	19.91	8.50	7.81
7.18	6.95	18.64	7.10	6.04	5.66	8.86	4.40	3.57	4.35
3.84	2.37	3.81	5.32	5.84	2.89	4.68	1.85	9.14	8.67
9.52	2.68	10.14	9.20	7.31	2.09	6.32	6.53	6.32	2.01
5.91	5.60	5.61	1.50	6.46	5.29	5.64	2.07	1.11	3.32
1.83	7.56								

El rango de la muestra va desde un valor mínimo de 1.11 a uno máximo de 23.38, en unidades de gramos de emisiones por galón de combustible.

El primer paso es construir la tabla de frecuencias, que se muestra en la tabla siguiente:

Intervalo de clase (g/gal)	Frecuencia	Frecuencia Relativa	Densidad
$[1\,,3)$	12	0.194	0.0970
$[3\,,5)$	11	0.177	0.0885
$[5\,,7)$	18	0.290	0.1450
$[7\,,9)$	9	0.145	0.0725
$[9\,,11)$	5	0.081	0.0405
$[11\,,15)$	3	0.048	0.0120
$[15\,,25)$	4	0.065	0.0065

Los intervalos de la columna de la izquierda se llaman intervalos de clase.

Los intervalos de clase dividen la muestra en grupos.

La notación $[1\,,3)$, $[3\,,5)$ y así sucesivamente, indica que un punto que está en el límite entrará en la clase de su derecha.

Por ejemplo, un valor de la muestra igual a 3 entrará en la clase $[3\,,5)$, y no en la $[1\,,3)$.

En la columna "Frecuencia", de la tabla, se presentan los números de puntos de los datos que están en cada uno de los intervalos de clase.

En la columna "Frecuencia relativa" se presentan las frecuencias divididas entre el número total de puntos de los datos, que para este caso es de 62.

La frecuencia relativa de un intervalo de clase es la proporción de puntos de los datos que están en el intervalo.

Observemos que debido a que cada punto de los datos está exactamente en un intervalo de clase, las frecuencias relativas deben sumar 1.

Por último, en la columna "Densidad" se presenta la frecuencia relativa dividida entre el ancho de clase.

Por ejemplo, en el primer renglón la frecuencia relativa es 0.194 y el ancho de clase es 2 (3 - 1= 2). Por tanto, la densidad es $0.194/2 = 0.0970$.

La última clase tiene un ancho de 10 y una frecuencia relativa de 0.065, por lo que su densidad es $0.065/10 = 0.0065$.

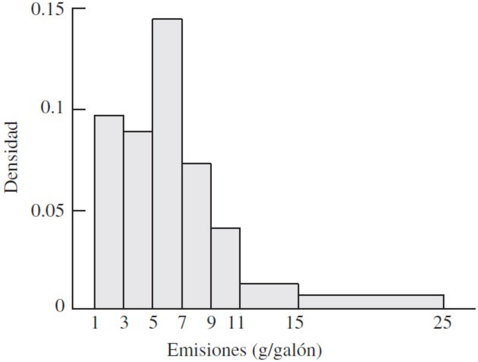
El propósito de la densidad es ajustar la frecuencia relativa con el ancho de la clase.

Sin que nada cambie, las clases anchas tienden a contener más elementos de la muestra que las clases más finas y, por consiguiente, tienden a tener frecuencias relativas más grandes.

Al dividir la frecuencia relativa entre el ancho de la clase se ajusta esta tendencia.

La densidad representa la frecuencia relativa por unidad de la variable aleatoria.

La figura presenta el histograma de la tabla anterior.



Las unidades en el eje horizontal son las unidades de los datos, en este caso g/galón.

Cada intervalo de clase se representa por un rectángulo.

La altura de cada rectángulo es la densidad de la muestra en ese intervalo de clase, que está dado en la cuarta columna de la tabla.

El área de cada rectángulo es, por tanto, la frecuencia relativa del intervalo de clase, que se encuentra en la tercera columna de la tabla.

Como las frecuencias relativas suman 1, el área bajo todo el histograma debe ser igual a 1.

Histograma de probabilidad

Cuando los valores posibles de una variable aleatoria discreta están espaciados uniformemente, la función de masa de probabilidad se puede representar por medio de un histograma, con rectángulos centrados en los valores posibles de la variable aleatoria.

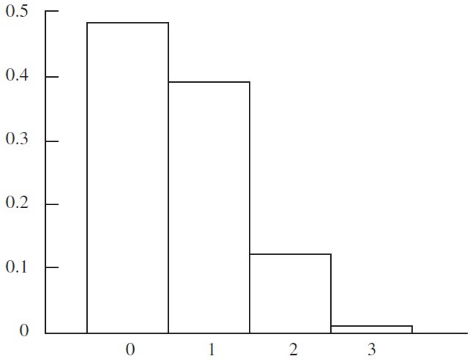
El área de un rectángulo centrado en un valor x es igual a $P(X = x)$.

Este histograma se llama histograma de probabilidad, ya que las áreas representan probabilidades.

Antes vimos el diagrama de la función de masa de probabilidad de una variable aleatoria X que representa el número de fallas en un alambre.

La figura siguiente presenta un histograma de probabilidad para esta variable aleatoria.

La figura presenta el histograma de probabilidad para esta variable aleatoria.



La probabilidad de que el valor de una variable aleatoria esté en un intervalo específico está dada por el área bajo el histograma de probabilidad.

Ejemplo

Determinemos la probabilidad de que un alambre elegido aleatoriamente tenga más de una falla.

Indiquemos esta probabilidad como un área bajo el histograma de probabilidad.

Solución

Se desea encontrar $P(X > 1)$.

Como ningún alambre tiene más de tres fallas, la proporción de que los alambres tengan más de una se puede encontrar al sumar la proporción de que tengan dos más la proporción de que tengan tres.

Simbólicamente, $P(X > 1) = P(X = 2) + P(X = 3)$.

La función de masa de probabilidad especifica que $P(X = 2) = 0.12$ y $P(X = 3) = 0.01$.

Por tanto, $P(X > 1) = 0.12 + 0.01 = 0.13$.

Solución (Continuación)

Esta probabilidad está dada por el área bajo el histograma de probabilidad que corresponde a esos rectángulos centrados en valores superiores a 1 (como vemos en la figura).

Hay dos de estos rectángulos; sus áreas son $P(X = 2) = 0.12$ y $P(X = 3) = 0.01$.

Ésta es otra manera de mostrar que $P(X > 1) = 0.12 + 0.01 = 0.13$.