TEMA 4 VARIABLES ALEATORIAS

Pablo Buenestado

Curso 2020-2021 Otoño

Departamento de Matemáticas (UPC)

Índice

Esquema

En la tabla relativa a las emisiones de una muestra de 62 vehículos se presentaron los intervalos de clase.

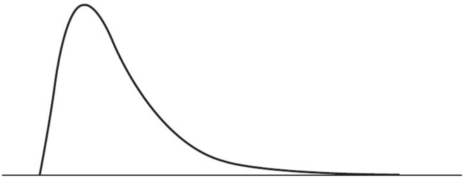
Observemos que las emisiones constituyen una variable continua, ya que sus valores posibles no están restringidos a algún conjunto discretamente espaciado.

Los intervalos de clase son elegidos para que cada intervalo contenga un número razonablemente grande de vehículos.

Si la muestra fuera más grande, se podrían hacer los intervalos más estrechos.

En particular, si se tuviera información sobre toda la población, que contiene millones de vehículos, se podrían hacer los intervalos extremadamente finos.

El histograma parecería entonces muy suave y se podría aproximar con una curva, como la que se muestra en la figura siguiente.



Si se hubiera elegido aleatoriamente un vehículo de esta población para medir sus emisiones, el nivel de emisiones X sería una variable aleatoria.

La probabilidad de que X esté entre cualesquiera dos valores a y b es igual al área bajo el histograma entre a y b.

Debido a que el histograma en este caso se representa por una curva, la probabilidad se encontraría mediante el cálculo de una integral.

La variable aleatoria X descrita aquí es un ejemplo de una variable aleatoria continua.

Ésta se define como una variable aleatoria cuyas probabilidades se representan por áreas bajo una curva.

Esta curva se llama función de densidad de probabilidad.

Como consecuencia de que la función de densidad de probabilidad es una curva, los cálculos de las probabilidades implican integrales, en vez de las sumatorias que se usan en el caso discreto.

Definición

Una variable aleatoria es continua si sus probabilidades están dadas por áreas bajo una curva.

La curva se llama función de densidad de probabilidad para la variable aleatoria.

A veces la función de densidad de probabilidad se llama distribución de probabilidad.

Cálculo de probabilidades con la función de densidad de probabilidad

Sea X una variable aleatoria continua.

Sea la función $f(x)$ la función de densidad de probabilidad de X.

Sean a y b cualesquiera dos números, con $a < b$.

La proporción de la población cuyos valores de X están entre a y b están dados por $\int_a^b f(x) \cdot dx$ y el área bajo la función de densidad de probabilidad entre a y b.

Ésta es la probabilidad de que la variable aleatoria X tome un valor entre a y b.

Observemos que el área bajo la curva no depende de si los puntos finales a y b estén incluidos en el intervalo.

Por tanto, las probabilidades que implican a X no dependen de si se incluyen a los puntos finales.

Resumen

Sea X una variable aleatoria continua con función de densidad de probabilidad $f(x)$.

Sean a y b cualesquiera dos números, con $a < b$. Entonces

$$P(a \leq X \leq b) = P(a \leq X < b) = P)(a < X \leq b) = P(a < X < b)$$

$$P(a \leq X \leq b) = \int_a^b f(x) \cdot dx$$

Además,

$$P(X \leq a) = P(X < a) = \int_{-\infty}^{a} f(x) \cdot dx$$

$$P(X \geq a) = P(X > a) = \int_{a}^{\infty} f(x) \cdot dx$$

Si $f(x)$ es la función de densidad de probabilidad de una variable aleatoria X, entonces el área bajo toda la curva desde $-\infty$ a ∞ es la probabilidad de que el valor de X esté entre $-\infty$ y ∞.

Esta probabilidad debe ser igual a 1, ya que el valor de X siempre está entre $-\infty$ y ∞.

Por tanto, el área bajo toda la curva $f(x)$ es igual a 1.

Resumen

Sea X una variable aleatoria continua con función de densidad de probabilidad $f(x)$. Entonces

$$\int_{-\infty}^{\infty} f(x) \cdot dx = 1$$

Ejemplo

Se perfora un hueco en un componente de una hoja de metal y después se inserta un eje a través del hueco.

La holgura del eje es igual a la diferencia entre el radio del hueco y el radio del eje.

Sea X la variable aleatoria que denota a la holgura, en milímetros.

La función de densidad de probabilidad de X es

$$f(x) = \begin{cases} 1.25 \cdot (1 - x^4) & 0 < x < 1 \\ 0 & \text{de otro modo} \end{cases}$$

Los componentes con holguras superiores a 0.8 mm se deben desechar.

¿Cuál es la proporción de componentes que serán desechados?

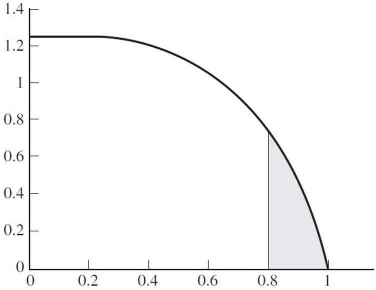
Solución

En la figura presentamos la función de densidad de probabilidad de X.

Observemos que la densidad $f(x)$ es 0 para $x \leq 0$ y para $x \geq 1$.

Esto indica que las holguras siempre están entre 0 y 1 mm.

La proporción de componentes que se debe desechar es $P(X > 0.8)$, que es igual al área bajo la función de densidad de probabilidad a la derecha de 0.8.



Solución (Continuación)

Esta área está dada por

$$P(X > 0.8) = \int_{0.8}^{\infty} f(x) \cdot dx = \int_{0.8}^{1} 1.25 \cdot (1 - x^4) \cdot dx$$

$$P(X > 0.8) = 1.25 \cdot \left(x - \frac{x^5}{5} \right) \Big]_{0.8}^{1} = 0.0819$$

Función de distribución acumulativa de una variable aleatoria continua

La función de distribución acumulativa de una variable aleatoria continua X es $F(x) = P(X \leq x)$, al igual que para una variable aleatoria discreta.

Para esta última, $F(x)$ se puede encontrar al sumar los valores de la función de masa de probabilidad.

Para una variable aleatoria continua, el valor de $F(x)$ se obtiene al integrar la función de densidad de probabilidad.

Puesto que $F(x) = P(X \leq x)$ se tiene que $F(x) = \int_{-\infty}^{x} f(t) \cdot dt$ donde $f(t)$ es la función de densidad de probabilidad.

Definición

Sea X una variable aleatoria continua con función de densidad de probabilidad $f(x)$.

La función de distribución acumulativa de X es la función

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t) \cdot dt$$

Ejemplo

Con referencia al ejemplo anterior, determinemos la función de distribución acumulativa $F(x)$ y dibujémosla.

Solución

La función de densidad de probabilidad X está dada por $f(x) = 0$ si $x \leq 0$, $f(x) = 1.25 \cdot (1 - x^4)$ si $0 < x < 1$ y $f(x) = 0$ si $x \geq 1$:

$$f(x) = \begin{cases} 0 & x \leq 0 \\ 1.25 \cdot (1 - x^4) & 0 < x < 1 \\ 0 & x \geq 1 \end{cases}$$

La función de distribución acumulativa está dada por $F(x) = \int_{-\infty}^{x} f(t) \cdot dt$.

Como que $f(x)$ está definida por partes en tres intervalos diferentes, el cálculo de la función de distribución acumulativa implica tres casos distintos.

Si $x \leq 0$:

$$F(x) = \int_{-\infty}^{x} f(t) \cdot dt = \int_{-\infty}^{x} 0 \cdot dt = 0$$

Solución (Continuación)

Si $0 < x < 1$:

$$F(x) = \int_{-\infty}^{x} f(t) \cdot dt = \int_{-\infty}^{0} f(t) \cdot dt + \int_{0}^{x} f(t) \cdot dt$$

$$F(x) = \int_{-\infty}^{0} 0 \cdot dt + \int_{0}^{x} 1.25 \cdot (1 - t^4) \cdot dt$$

$$F(x) = 0 + 1.25 \cdot \left(t - \frac{t^5}{5} \right) \Bigg]_0^x = 1.25 \cdot \left(x - \frac{x^5}{5} \right)$$

Solución (Continuación)

Si $x \geq 1$:

$$F(x) = \int_{-\infty}^{x} f(t) \cdot dt = \int_{-\infty}^{0} f(t) \cdot dt + \int_{0}^{1} f(t) \cdot dt + \int_{1}^{x} f(t) \cdot dt$$

$$F(x) = \int_{-\infty}^{0} 0 \cdot dt + \int_{0}^{1} 1.25 \cdot (1 - t^4) \cdot dt + \int_{1}^{x} 0 \cdot dt$$

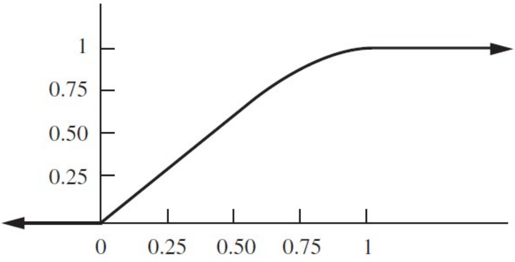
$$F(x) = 0 + 1.25 \cdot \left(t - \frac{t^5}{5} \right) \Big]_0^1 + 0 = 0 + 1 + 0 = 1$$

Por tanto,

$$F(x) = \begin{cases} 0 & x \leq 0 \\ 1.25 \cdot \left(x - \dfrac{x^5}{5} \right) & 0 < x < 1 \\ 1 & x \geq 1 \end{cases}$$

Solución (Continuación)

Aquí presentamos una gráfica de $F(x)$



Observemos que la función de distribución acumulativa $F(x)$ del ejemplo representa una función continua si su gráfica no presenta ningún salto.

Ésta es una característica de las variables aleatorias continuas.

La función de distribución acumulativa de una variable aleatoria continua será continua siempre, mientras que la función de distribución acumulativa de una variable aleatoria no continua nunca será continua.

Ejemplo

Con referencia al ejemplo anterior, usemos la función de distribución acumulativa para encontrar la probabilidad de que la holgura del eje sea menor a 0.5 mm.

Solución

Sea X la holgura del eje, se tiene que encontrar $P(X \leq 0.5)$.

Esta probabilidad es equivalente a $F(0.5)$, donde $F(x)$ es la función de distribución acumulativa.

Usando los resultados del ejemplo,

$$F(0.5) = 1.25 \cdot \left(0.5 - \frac{0.5^5}{5}\right) = 0.617$$

Media y varianza para variables aleatorias continuas

La media poblacional y la varianza poblacional de una variable aleatoria continua están definidas de la misma forma que para una variable aleatoria discreta, excepto que se usa la función de densidad de probabilidad en lugar de la función de masa de probabilidad.

Específicamente, si X constituye una variable aleatoria continua, su media poblacional se define como el centro de masas de su función de densidad de probabilidad y su varianza poblacional representa el momento de inercia con respecto a un eje vertical que pasa a través del centro de masas.

Las fórmulas son análogas a las ecuaciones respectivas de la variable aleatoria discreta, con los sumatorios reemplazados por integrales.

Como en el caso de las variables aleatorias discretas, a veces se eliminará la palabra "poblacional" y se hará referencia a la media poblacional, a la varianza poblacional y a la desviación poblacional sólo como la media, la varianza y la desviación estándar, respectivamente.

Definición

Sea X una variable aleatoria continua con función de densidad de probabilidad $f(x)$.

Entonces la media de X está dada por

$$\mu_X = \int_{-\infty}^{\infty} x \cdot f(x) \cdot dx$$

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

Definición

Sea X una variable aleatoria continua con función de densidad de probabilidad $f(x)$. Entonces

- La varianza de X está dada por

$$\sigma_X^2 = \int_{-\infty}^{\infty} (x - \mu_X)^2 \cdot f(x) \cdot dx$$

- Una fórmula alternativa para la varianza está dada por

$$\sigma_X^2 = \int_{-\infty}^{\infty} x^2 \cdot f(x) \cdot dx - \mu_X^2$$

- 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

Con referencia al ejemplo anterior, determinamos la media y la varianza de la holgura.

Solución

La media de la holgura está dada por

$$\mu_X = \int_{-\infty}^{\infty} x \cdot f(x) \cdot dx = \int_0^1 x \cdot \left[1.25\left(1 - x^4\right)\right] \cdot dx$$

$$\mu_X = 1.25 \cdot \left(\frac{x^2}{2} - \frac{x^6}{6}\right)\Big]_0^1 = 0.4167$$

Solución (Continuación)

Una vez que se ha calculado $\mu_X = 0.4167$, ahora se puede calcular σ_X^2.

Es más fácil usar la fórmula alternativa:

$$\sigma_X^2 = \int_{-\infty}^{\infty} x^2 \cdot f(x) \cdot dx - \mu_X^2 = \int_0^1 x^2 \cdot \left[1.25\left(1 - x^4\right)\right] \cdot dx - (0.4167)^2$$

$$\sigma_X^2 = 1.25 \cdot \left(\frac{x^3}{3} - \frac{x^7}{7}\right)\Bigg]_0^1 - (0.4167)^2 = 0.0645$$