ESTADÍSTICA
DIAPOSITIVAS DE EJEMPLO

Unidad 3 – *Características, Esperanza y Varianza*

Ingeniería en Informática

Año 2022

Prof. Juan Pablo Taulamet

consultas: taulamet@unl.edu.ar

Caso Aplicado

La proporción de errores de "deploy" en un repositorio de control de versiones Git sigue la siguiente ley de probabilidad:

$$f(x) = \begin{cases} 3x^2 & \text{si } 0 \leq x \leq 1 \\ 0 & \text{en otro caso} \end{cases}$$

¿Qué proporción de errores de este tipo se espera tener y con qué variabilidad?

Esperanza

$$E(X) = \mu = \int_{-\infty}^{+\infty} x f(x) dx$$

$$f(x) = \begin{cases} 3x^2 & \text{si } 0 \leq x \leq 1 \\ 0 & \text{en otro caso} \end{cases}$$

$$E(X) = \int_{-\infty}^{0} 0 dx + \int_{0}^{1} x 3x^2 dx + \int_{1}^{\infty} 0 dx =$$

$$\int_{0}^{1} 3x^3 dx = 3 \int_{0}^{1} x^3 dx = 3[1^4/4 - 0] = 3/4$$

Caso Aplicado

¿Qué proporción de errores de este tipo se espera tener?

Respuesta:

Para la variable que representa la proporción de errores de “deploy” en un repositorio de control de versiones Git, se espera tener un 0.75 de errores, lo cual implica una proporción muy alta.

Varianza

$$Var(X) = \sigma^2 = E(X - E(X))^2$$

$$Var(X) = \int_{-\infty}^{+\infty} (x - E(x))^2 f(x) dx$$

Existen 2 formas de cálculo

$$Var(X) = \sigma^2 = E(X^2) - E(X)^2$$

$$E(X)^2 = \left[\int_{-\infty}^{+\infty} x f(x) dx \right]^2$$

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x) dx$$

$$f(x) = \begin{cases} 3x^2 & \text{si } 0 \leq x \leq 1 \\ 0 & \text{en otro caso} \end{cases}$$

$$Var(X) = E(X^2) - E(X)^2$$

$$E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x) dx$$

$$\int_{-\infty}^{0} x^2.0dx + \int_{0}^{1} x^2.3x^2 dx + \int_{1}^{\infty} x^2.0dx =$$

$$\int_{0}^{1} 3x^4 dx = 3 \int_{0}^{1} x^4 dx = 3[1^5/5 - 0] = 3/5$$

$$Var(X) = 3/5 - (3/4)^2 = 3/80$$

Caso Aplicado

*¿Qué proporción de errores de este tipo se espera tener y **<u>con qué variabilidad?</u>***

Respuesta:

Para la variable que representa la proporción de errores de "deploy" en un repositorio de control de versiones Git, la varianza es de 3/80.

¿Unidades?

ESTADÍSTICA
DIAPOSITIVAS DE EJEMPLO

Unidad 3 – *Características*

Ingenierías en Recursos Hídricos / Ambiental / Agrimensura

Año 2022

Prof. Juan Pablo Taulamet

consultas: taulamet@unl.edu.ar

Para rellenar una zona baja aledaña a un río, se utilizan dos camiones: A y B. La distribución de la carga diaria (en tn) transportada por cada camión se puede adaptar a la siguiente función:

$$f(x) = \begin{cases} \frac{x}{52} & \text{si } 11 < x < 15 \\ 0 & \text{e.o.c} \end{cases}$$

¿Se podrá decir cuál es el valor medio de carga de cada camión?¿Alcanza con sólo este valor para caracterizar a las cargas? ¿Qué más se podría calcular?

consultas: taulamet@unl.edu.ar

Antes que nada...

¿Es una función de densidad válida?:

$$f(x) = \begin{cases} \frac{x}{52} & \text{si } 11 < x < 15 \\ 0 & \text{e.o.c} \end{cases}$$

- ¿Definida para todo posible x?
- ¿No negativa para todo posible x?
- ¿Cumple la condición de cierre?

$$f(x) = \begin{cases} \frac{x}{52} & \text{si } 11 < x < 15 \\ 0 & \text{e.o.c} \end{cases}$$

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

$$\int_{\infty}^{11} 0dx + \int_{11}^{15} \frac{x}{52} dx + \int_{15}^{\infty} 0dx =$$

$$0 + \frac{x^2}{52 * 2}\bigg|_{11}^{15} dx + 0 = 1$$

Lo último que se pierde...

$$f(x) = \begin{cases} \frac{x}{52} & \text{si } 11 < x < 15 \\ 0 & \text{e.o.c} \end{cases}$$

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

$$\int_{-\infty}^{11} x0dx + \int_{11}^{15} \frac{x^2}{52}dx + \int_{15}^{\infty} x0dx =$$

$$0 + \frac{x^3}{52 * 3}\bigg|_{11}^{15} dx + 0 => E(X) \approx 13.10$$

¿Variabilidad?

$$D(X) = \sqrt{V(X)}; V(X) = E(X^2) - E(X)^2$$

$$E(X^2) = \int_{11}^{15} x^2 \cdot \frac{x}{52} dx = \int_{11}^{15} \frac{x^3}{52} dx =$$

$$\left.\frac{x^4}{52 * 4}\right|_{11}^{15} dx = 173) => V(X) = 173 - 13.10^2 =$$

$$V(X) \approx 1.39 => D(X) \approx 1.18$$

consultas: taulamet@unl.edu.ar

Si se supone que las cargas transportadas por los dos camiones son independientes; ¿Se podría decir cuál es el valor medio de carga total transportada por los camiones? ¿Alcanza con sólo este valor para caracterizar a la carga total? ¿Qué más se podría calcular?

consultas: taulamet@unl.edu.ar

Pensando en 2 camiones

Suponiendo que X mide la carga del camión A e Y la carga del camión B, pasamos en limpio las funciones de densidad de cada variable:

$$f(x) = \begin{cases} \frac{x}{52} & \text{si } 11 < x < 15 \\ 0 & \text{e.o.c} \end{cases}$$

$$f(y) = \begin{cases} \frac{y}{52} & \text{si } 11 < y < 15 \\ 0 & \text{e.o.c} \end{cases}$$

Carga Total Media

Luego si definimos la carga total T como la suma de las cargas de los camiones y aplicamos la propiedad de la esperanza de una suma:

$$T = X + Y$$

$$E(T) = E(X + Y) = E(X) + E(Y)$$

$$=> E(T) = 13.10 + 13.10 = 26.20$$

Desvío Total Medio

Análogamente, por la propiedad de la varianza de una suma, necesitaremos ***calcular la covarianza:***

$$V(T) = V(X+Y) =$$

$$V(X) + V(Y) + 2COV(X,Y)$$

$$COV(X,Y) = E(X,Y) - E(X) * E(Y)$$

Covarianza

$$COV(X,Y) = E(X,Y) - E(X) \cdot E(Y)$$

$$E(X,Y) = \int_{11}^{15} \int_{11}^{15} xyf(x,y) \cdot dx \cdot dy =$$

$$\int_{11}^{15} \int_{11}^{15} xy \frac{x}{52} \frac{y}{52} \cdot dx \cdot dy =$$

$$\int_{11}^{15} \int_{11}^{15} \frac{x^2y^2}{2704} \cdot dx \cdot dy \approx 171.68$$

Covarianza: Interpretación

$$COV(X,Y) = E(X,Y) - E(X) \cdot E(Y)$$

$$=> 171.68 - (13.10^2) \approx 0$$

$$=> COV(X,Y) = 0$$

Recordemos que si la covarianza es cero, podemos interpretar que no existe relación lineal entre las dos variables, lo cual es coherente con lo supuesto en el enunciado del problema.

Desvío Total Medio

$$V(T) = V(X+Y) =$$

$$V(X) + V(Y) + 2COV(X,Y)$$

$$=> V(T) = 1.3 + 1.3 = 2.6$$

$$=> D(T) \approx 1.61$$

consultas: taulamet@unl.edu.ar