# Taller Formativo 3

Mario Castillo Sanhueza

Docente: Dr. Julio Rojas Mora

Departamento de Ingeniería Informática
Facultad de Ingeniería
Universidad Católica de Temuco

Agosto 08, 2022

# 1. Problemática.

- ¿Cuál es el efecto de un valor extremo en el promedio y la mediana?.
- Demuestre que para un conjunto de datos X =$\{xi\}, \forall i = 1, ..., n$ : $\bar{x} = arg_{x^*} min \sum_{i=0}^{n} (xi - x^*)^2$
  es decir, que el punto que minimiza la suma de las diferencias cuadráticas de los elementos de X y cualquier otro punto $x^*$ es la media aritmética.
- Para los siguientes datos de duración, en minutos, de llamadas a un call-center, obtenga la moda: 5, 5, 2, 1, 5, 2, 2, 3, 4, 4, 3 ¿Qué característica se presenta? ¿Que podría implicar esto?

# 2. Desarrollo.

- Tomando en cuenta que el cálculo del promedio o media aritmética utiliza todos los datos, un valor extremo puede distorsionar la regularidad o tendencia dada la alta sensibilidad a los cambios en esta producto de la importancia de cada punto para esta.
- Para desarrollar esta desmotracion comenzaremos derivando la expresion$(x_i - x^*)$ para luego realizar una igualdad a 0.

$$\frac{\partial}{\partial x}(x_i - x^*)^2$$

$$2 \cdot (x_i - x^*) \cdot \frac{\partial}{\partial x}(x_i - x^*)$$

$$2 \cdot (x_i - x^*) \cdot (\frac{\partial}{\partial x}x_i - \frac{\partial}{\partial x}x^*)$$

$$2 \cdot (x_i - x^*) \cdot (0 - 1)$$

$$2 \cdot (-x_i + x^*) \to 2 \cdot (x^* - x_i)$$

Realizada la derivación establecemos la igualdad.

$$2 \cdot (x^* - x_i) = 0$$

$$x^* - x_i = \frac{0}{2}$$

$$x^* - x_i = 0$$

$$x^* = x_i$$

De esta forma se demuestra que al despejar $x^*$ obtenemos una la igualdad $x^* = x_i$, dicho de otra forma, la suma de las diferencias al cuadrado es igual a 0, es decir, la varianza o media aritmética en un contexto donde esta media aritmética del cuadrado de las desviaciones de la media resulte ser 0 o mayor a este, de esta forma se puede demostrar que $x^*$ es esta media aritmética.

- comenzaremos ordenando nuestra muestra de la forma :

  {5, 5, 2, 1, 5, 2, 2, 3, 4, 4, 3} $\rightarrow \{1, 2, 2, 2, 3, 3, 4, 4, 5, 5, 5\}$ Dado que la moda por definición es el dato con la frecuencia más alta en todo el conjunto, al observar esta muestra podemos reconocer que se presentan 2 datos con la misma frecuencia, siendo estos los números 2 y 5 con una frecuencia de 3 cada uno.
  Esto implica que para esta serie de datos se nos presenta una serie bimodal dado que es una distribución de probabilidad continua con dos diferentes modos, una implicación de esto seria que si nosotros extendiéramos la muestra de estos datos, ya sea con una mayor recopilación de estos o mediante una regresión lineal podríamos encontrarnos con una distribución multimodal.