

# Notas: Inferencia

---

## Lunes 09 de Abril

---

### Temas vistos:

---

- Tema de la clase: Introducción a la inferencia

### Actividades:

---

- Examen diagnóstico.
- Criterios de calificación, se acuerda como vamos a obtener la nota, calendarios y medios de comunicación, puntos extras.

### Temas vistos

---

#### Estadísticos muestrales

##### Definición:

Se le llama *estadístico muestral* a una función  $g$  tal que, al obtener observaciones del vector  $\hat{X} = (X_1, \dots, X_n)$ , el valor  $g(\hat{X})$  estima un parámetro poblacional. También a la distribución de  $\hat{X}$  se le llama distribución muestral.

---

**Que entiendo:** Un estadístico es una función que recibe datos y nos devuelve un número que trata de inferir un parámetro poblacional.

**Que no entendí:** ¿A que se refiere con observaciones?

R= Se refiere a hacer un muestreo.

##### Ejemplos:

- Promedio.
- Otro Ejemplo (Referencia: [Estadístico muestral - Wikipedia, la enciclopedia libre](#))

**Investigación extra:** Otros de estadísticos conocidos (aplicaciones a la IA).....

# Viernes 14 de Abril

---

## Temas vistos:

---

- Media muestral.
- Varianza muestral.

## Actividades:

---

- Se deja la primera tarea: Temas a evaluar ----

### Investigación extra:

Aplicaciones de cierto tema

#### 1.9.1. Gaussian Naive Bayes

`GaussianNB` implements the Gaussian Naive Bayes algorithm for classification. The likelihood of the features is assumed to be Gaussian:

$$P(x_i | y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{(x_i - \mu_y)^2}{2\sigma_y^2}\right)$$

The parameters  $\sigma_y$  and  $\mu_y$  are estimated using maximum likelihood.

```
>>> from sklearn.datasets import load_iris
>>> from sklearn.model_selection import train_test_split
>>> from sklearn.naive_bayes import GaussianNB
>>> X, y = load_iris(return_X_y=True)
>>> X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.5, random_state=0)
>>> gnb = GaussianNB()
>>> y_pred = gnb.fit(X_train, y_train).predict(X_test)
>>> print("Number of mislabeled points out of a total %d points : %d"
...       % (X_test.shape[0], (y_test != y_pred).sum()))
Number of mislabeled points out of a total 75 points : 4
```

Acá explico como el tema se aplica.

Referencia : [1.9. Naive Bayes — scikit-learn 1.2.2 documentation](#)

## Ejercicios resueltos

---

Viernes 14 de abril

---

### Ejercicio:

Calcula la media de estos datos 1, 2, 3, 4, 5, 6, 77, 88, 88, 88

### Respuesta:

El promedio se calcula con la ecuación  $\frac{\sum_i^m k}{m}$ , usando esta fórmula,

```
# Esta lista contiene los datos que vamos a calcular el promedio
lista = [1,2,3,4,5,6,77,88,88,88]

'''
    Con esta función vamos obtener la cantidad de datos en la lista,
    que es contar la cantidad de elementos.
'''

total = len(lista)

'''
Vamos obtener la suma de los elementos contenidos en la lista
'''

suma_total = sum(lista)

'''
El siguiente código replica la ecuación del promedio
'''
promedio = suma_total/total
```

```
# Esta lista contiene los datos que vamos a calcular el promedio
lista <- c(1,2,3,4,5,6,7,8,8,8,8)

#Con esta función vamos obtener la cantidad de datos en la lista, que es contar la cantidad de
elementos.

total = length(lista)

# Vamos obtener la suma de los elementos contenidos en la lista

suma_total = sum(lista)

# El siguiente código replica la ecuación del promedio

promedio = suma_total/total

# Esta función pertenece a la librería base de R calcula la media de un vector
mean(lista)
```

## Lunes 17 de Abril

---

### Ejercicio:

Calcula el valor esperado de la variable aleatoria  $Y = \sum_{i=0}^M X_i$  si cada  $X_i$  tiene valor esperado  $E(X_i) = i$ .

### Respuesta:

De la propiedad lineal de la esperanza,

$$E(Y) = E\left(\sum_{i=0}^M X_i\right) = \sum_{i=0}^M E(X_i) = \sum_{i=0}^M i = \frac{M(M+1)}{2}$$

donde usamos la ecuación  $1 + 2 + 3 + 4 + \dots + n = \frac{n(n+1)}{2}$  para  $M = n$  y como  $E(X_0) = 0$  no afecta la ecuación.