

# Experiencia 2

Camilo Valenzuela Carrasco

Departamento de informática  
Universidad Técnica Federico Santa María

Estadística Computacional

## 1 Noción Frecuencista de las probabilidades

# Noción Frecuentista de las probabilidades

Buscamos la probabilidad de que suceda A al realizar cierto experimento.  
Para eso tenemos.

N:  $N^o$  total de veces que se realiza un experimento

$N_A$  :  $N^o$  total de veces que ocurre A

$$P(A) = \lim_{N \rightarrow \infty} \frac{N_A}{N}$$

# Simulando un experimento

- Para simular un experimento utilizaremos la función **sample**(vector, cantidad, reemplazo)
- **vector** son los distintos resultados que puede tener nuestro experimento. Sample genera una lista de tamaño **cantidad**, con valores del vector.
- Si queremos calcular la probabilidad utilizando la noción frecuentista, necesitamos repetir, el experimento. Para esto lo más intuitivo es utilizar **un ciclo for** y repetir el experimento varias veces, pero esto es lento.

# Ejercicio 1

- Realice la simulación del lanzamiento de una moneda.
- ¿Cuál es la probabilidad de obtener una cara? (Realice 1000 experimentos).

# Introduciendo a uso matricial

Para acelerar el proceso de simulación se pueden utilizar matrices y vectores.

Como vimos en la experiencia 1, para crear un vector se utiliza la función `c()`.

Esta función también nos sirve para concatenar valores, y así poder agrandar nuestro vector.

## Example

- `vector <- c()`
- `vector <- c(vector, 1) → [1]`
- `vector <- c(vector, 2) → [1, 2]`

# Ejercicio 1.1

Realice el Ejercicio 1, pero ahora solo utilizando un vector.  
(**Hint:** Función sum y los valores T (True) valen 1, F vale 0).



# Creando una matriz

Para crear una matriz, existen dos funciones muy utilizadas.

- **cbind**: Parecido a **c()**, pero sirve para unir dos objetos de R, para generar una matriz unimos dos vectores

## Example

```
matriz <- cbind(vector1, vector2)
matriz <- cbind(matriz, vector)
```

- **matrix**: Se le entrega un vector de una dimensión, y luego se ingresan la cantidad de filas y columnas de la matriz.

## Example

```
matriz <- matrix(c(2, 4, 3, 1, 5, 7), nrow = 2, ncol = 3, byrow = TRUE)
```

## Ejercicio 1.2

- Ahora realice el mismo experimento pero con distinta cantidad de experimentos (100, 1000, 10000).
- Y cada muestra repítala 100 veces. Guarde las probabilidades de cada muestra en cada iteración.
- Junte todos sus resultados, y muéstrellos en un boxplot.
- Calcule la desviación estándar para cada muestra.  
(**Hint:** Vea información sobre función `apply`).

## Ejercicio 2

Una caja contiene 12 bolitas, de las cuales hay 5 blancas y 7 negras. Se sacan 2 bolitas y se vuelven a dejar en la caja. Se sacan otra vez 2 bolitas y se vuelven a dejar en la caja, y así se continúa hasta efectuar 5 extracciones.

¿Cual es la probabilidad de sacar 2 bolitas negras en cada uno de los 3 primeros lanzamientos? (No es necesario hacerlo matricial).

# Simulando una distribución

Para realizar una simulación de distribución existen varias funciones, en esta experiencia nos centraremos en la distribución Poisson.

- Para crear una muestra Poisson, utilizaremos la función **rpois(tamaño,lambda)**, que nos entrega un vector del tamaño especificado.
- Si queremos ver la frecuencia acumulada utilizamos la función **ppois(arreglo,lambda)**

# Ejercicio 3

- Realice varias muestras Poisson, de distinto tamaño. Y realice varias repeticiones de cada muestra.
- Calcule la media y desviación estandar de las muestras.
- Analice la convergencia de la distribución muestral utilizando el teorema de Glivenko-Cantelli<sup>1</sup>

---

<sup>1</sup>[https://en.wikipedia.org/wiki/Glivenko%E2%80%93Cantelli\\_theorem](https://en.wikipedia.org/wiki/Glivenko%E2%80%93Cantelli_theorem)