



# **INSTITUTO TÉCNICO SUPERIOR “GAUSS KISH”**

**CURSO DE ESTADÍSTICA PRÁCTICA CON R Y  
PYTHON**

**AÑO 2024**



## CURSO DE ESTADÍSTICA PRÁCTICA CON R Y PYTHON

**Descripción:** Este curso te proporcionará las habilidades esenciales de estadística práctica necesarias para proyectos de ciencia de datos. Aprenderás a aplicar conceptos estadísticos utilizando R y Python, dos lenguajes fundamentales en el campo de la ciencia de datos.

### Objetivos del Curso

- Aplicar estadística de manera práctica en proyectos de ciencia de datos.
- Utilizar herramientas como R y Python para análisis estadístico.
- Desarrollar habilidades prácticas para la toma de decisiones basada en datos.

### Temas:

#### A - Análisis exploratorio de datos.

- **Elementos de datos estructurados**
- **Datos rectangulares**  
Marcos de datos e índices y Estructuras de datos no rectangulares.
- **Estimación de la localización**  
Media y Estimación de medianas robusta.
- **Estimación de la variabilidad**

Desviación estándar y estimaciones relacionadas y Estimación basada en percentiles.

- **Exploración de la distribución de datos**

Percentiles y diagramas de caja, Tablas de frecuencias e histogramas y Diagrama y estimación de la curva de densidad.

- **Exploración de datos binarios y categóricos**

Moda, Valor esperado y Probabilidad.

- **Correlación**

Diagramas de dispersión

- **Exploración de dos o más variables**

Agrupación hexagonal y contornos (representación numérica frente a datos numéricos), Dos variables categóricas, Datos categóricos y numéricos y Visualización de varias variables.

## **B - Distribuciones de datos y muestreo**

- **Muestreo aleatorio y sesgo de la muestra**

Sesgo, Selección aleatoria, Tamaño frente a calidad: ¿cuándo importa el tamaño?

Media muestral frente a media poblacional.

- **Sesgo de selección**

Regresión a la media.

- **Distribución muestral del estadístico**

Teorema del límite central y Error estándar.

- **Bootstrap**

Remuestreo frente a bootstrapping.

- **Intervalos de confianza**

- **Distribución normal**

Normal estándar y diagramas QQ.

- **Distribuciones de cola larga**

- **Distribución t de Student**

- **Distribución binomial**

- **Distribución chi cuadrado**
- **Distribución F**
- **La distribución de Poisson y distribuciones relacionadas**  
Distribución de Poisson, Distribución exponencial, Estimación de la tasa de fallos y Distribución de Weibull

## **C - Experimentos estadísticos y pruebas significativas**

- **Prueba A/B**  
¿Por qué tener un grupo de control? y ¿Por qué solo A/B? ¿Por qué no C, D, ...?
- **Pruebas de hipótesis**  
La hipótesis nula, Hipótesis alternativa, Pruebas de hipótesis unidireccionales o bidireccionales.
- **Remuestreo**  
Prueba de permutación, Ejemplo: adherencia de la web  
Pruebas de permutación exhaustiva y de bootstrap, Pruebas de permutación: el resultado final de la ciencia de datos.
- **Significación estadística y valores p**  
Valor p, Alfa, Errores de tipo 1 y 2 y Ciencia de datos y valores p.
- **Pruebas t**
- **Pruebas múltiples**
- **Grados de libertad**
- **ANOVA**  
Estadístico F y ANOVA bidireccional.
- **Prueba de chi cuadrado**  
Prueba de chi cuadrado: enfoque de remuestreo, Prueba de chi cuadrado: teoría estadística, Prueba exacta de Fisher y Relevancia para la ciencia de datos
- **Algoritmo Multi-Arm Bandit**
- **Potencia y tamaño de la muestra**  
Tamaño de la muestra