

## IDENTIFICACIÓN

<b>NOMBRE ESCUELA</b>	ESCUELA DE CIENCIAS
<b>NOMBRE DEPARTAMENTO</b>	Ciencias Matemáticas
<b>ÁREA DE CONOCIMIENTO</b>	MATEMATICAS, ESTADISTICA Y AFINES
<b>NOMBRE ASIGNATURA EN ESPAÑOL</b>	ESTADÍSTICA NO PARAMÉTRICA EN CIENCIAS DE LOS DATOS
<b>NOMBRE ASIGNATURA EN INGLÉS</b>	NONPARAMETRIC STATISTICS IN DATA SCIENCES
<b>CÓDIGO</b>	CM0896
<b>SEMESTRE DE UBICACIÓN</b>	20212
<b>INTENSIDAD HORARIA SEMANAL</b>	8 horas semanales
<b>INTENSIDAD HORARIA SEMESTRAL</b>	36 horas semestral
<b>CRÉDITOS</b>	3
<b>CARACTERÍSTICAS</b>	No suficientable

---

## 2. JUSTIFICACIÓN DEL CURSO

El desarrollo teórico de una gran variedad de métodos estadísticos ha sido principalmente soportado en diversos supuestos, concretamente, en el conocimiento de la distribución que genera las observaciones. Si los supuestos considerados se verifican, entonces las técnicas paramétricas son las metodologías más útiles para realizar estudios e inferencias sobre una población. Sin embargo, en los casos prácticos donde se violan este tipo de supuestos la eficiencia de la inferencia paramétrica se ve bastante comprometida para la obtención de conclusiones correctas. La estadística no paramétrica proporciona una alternativa en aquellas situaciones donde no es correcto la realización de contrastes de tipo paramétricos. De igual manera las técnicas estadísticas robustas son las metodologías llamadas a ser implementadas en casos donde las muestras están contaminadas o existen indicios de datos atípicos, ya que los estadísticos de uso común presentan demasiada sensibilidad a datos con estas características y por lo tanto las conclusiones obtenidas pueden ser incorrectas.

En este curso se proporciona a los estudiantes las principales herramientas y técnicas estadísticas no paramétricas y robustas de uso habitual de la teoría de estimación de parámetros. El enfoque del curso es desde un punto de vista formal pero también haciendo especial énfasis en la implementación computacional de todos los conceptos introducidos y de su aplicación. El curso concretamente va dirigido a estudiantes que están en formación de Maestría. Sin embargo, sería de notable utilidad para estudiantes de segundo año de Doctorado y de último año de ingeniería matemática

### 3. PROPÓSITO U OBJETIVO GENERAL DEL CURSO

3.1. Proporcionar a los estudiantes las principales herramientas y técnicas estadísticas no paramétricas y robustas de uso habitual de la teoría de estimación de parámetros

#### 3.2. OBJETIVOS ESPECÍFICOS

- 3.2.1. Conocer las versiones robustas de los estadísticos de localización, escala y forma
- 3.2.2. Implementar las técnicas de identificación de outliers univariantes y multivariantes
- 3.2.3. Identificar las ventajas del uso de técnicas estadísticas no paramétricas
- 3.2.4. Conocer y aplicar la metodología Bootstrap
- 3.2.5. Identificar e implementar algunos métodos de regresión no paramétrica
- 3.2.6. Identificar e implementar algunos métodos de regresión robusta
- 3.2.7. Desarrollar métodos basados en Kernels para estimación de densidades
- 3.2.8. Entender e implementar los test estadísticos basados en rangos
- 3.2.9. Implementar funciones de profundidad estadística y sus diversas aplicaciones
- 3.2.10. Desarrollar metodologías para la asignación de rangos en muestras multivariantes y funcionales

### 4. COMPETENCIAS BÁSICAS QUE EL ALUMNO ESTARÁ EN CONDICIONES DE LOGRAR:

Analiza los fundamentos de la estadística no paramétrica, su importancia, desarrollo y evolución, así como su aplicación en los diferentes casos que se le presentan, tanto en su profesión como en el ejercicio de la misma.

Adquirir la capacidad de adaptación a nuevas situaciones que puedan requerir la mejora o creación de técnicas estadísticas en términos de otras ya conocidas.

Analiza los fundamentos de la estadística no paramétrica.

Utiliza datos de captura de forma adecuada en el software estadístico, con la finalidad de analizarlos posteriormente.

Adquirir la capacidad de comunicación con equipos multidisciplinares en los que el uso de la Estadística juega un papel relevante en la toma de decisiones.

Comprender y utilizar el lenguaje estadístico. Adquirir la capacidad para analizar y sintetizar los problemas de los distintos campos de aplicación de la Estadística.

## **5. DESCRIPCION ANALITICA DE CONTENIDOS: TEMAS Y SUBTEMAS**

### **5.1. ESTADÍSTICA ROBUSTA Y NO PARAMÉTRICA: UNA INTRODUCCIÓN A LAS TÉCNICAS BÁSICAS**

- 5.1.1. Definición de estadística no paramétrica
- 5.1.2. Versiones robustas y no paramétricas de los estadísticos descriptivos y propiedades
- 5.1.3. Identificación de outliers univariantes y multivariantes.
- 5.1.4. Estimación de funciones de distribución y estimadores de tipo Plug in.
- 5.1.5. Covarianzas y correlaciones no paramétricas y robustas (Kendall, Spearman y MAD y COMEDIAN)
- 5.1.6. Función influencia
- 5.1.7. Método Bootstrap y Jackknife y construcción no paramétrica de intervalos confidenciales.
- 5.1.8. Implementación computacional con datos reales y sintéticos

### **5.2. REGRESIÓN ROBUSTA, NO PARAMÉTRICA Y ESTIMACIÓN DE DENSIDADES.**

- 5.2.1. Conceptos generales de suavizamiento y suavizadores lineales
- 5.2.2. Revisión general de regresión lineal y logística
- 5.2.3. Selección del parámetro de suavizado, validación cruzada.
- 5.2.4. Regresión penalizada y splines
- 5.2.5. Estimación de varianza y bandas de confianza
- 5.2.6. Estimación de densidades y kernels
- 5.2.7. Métodos generales para regresión robusta
- 5.2.8. Implementación computacional con datos reales y sintéticos

### **5.3. TEST ESTADÍSTICOS BASADOS EN RANGOS.**

- 5.3.1. Pruebas estadísticas basadas en rangos para una y dos poblaciones (Wilcoxon, MannWhitney)
- 5.3.2. Funciones de profundidad estadística
- 5.3.3. Pruebas estadísticas basadas en funciones de profundidad
- 5.3.4. Técnicas estadísticas robustas basadas en funciones de profundidad (Muestras recortadas)
- 5.3.5. Implementación computacional con datos reales y sintéticos

## 6. ESTRATEGIAS METODOLÓGICAS Y DIDÁCTICAS:

Exposición magistral de los conceptos e implementación computacional de algunos de ellos.

Solución conjunta de problemas claves en los que se consideren los temas tratados en clase.

Trabajo independiente y consulta autónoma por parte del estudiante.

## 7. RECURSOS

### 7.1. Locativos

Aulas de clase y Salón de Cómputo

### 7.2. Tecnológicos

Matlab, Octave, R

## 8. CRITERIOS Y POLÍTICAS DE SEGUIMIENTO Y EVALUACIÓN ACADÉMICA

Examen Unidad 1 â 30%

Ejercicios Unidades 1,2 y 3 â 30%

Trabajo final â 40%

## 9. BIBLIOGRAFIA GENERAL

- 9.1. Wasserman, L. (2006) All of Nonparametric Statistics. Springer.
- 9.2. Maronna, R., Martin, R.D., Yohai, V.J. (2006). Robust Statistics, Theory and Methods. John Wiley & Sons Ltd

9.3. Sprent, P., Smeeton, N.C. (2001) Applied Nonparametric Statistical Methods. Chapman & Hall

**10. NOMBRE DEL PROFESOR COORDINADOR DE MATERIA Y NOMBRE DE PROFESORES DE LA MATERIA QUE PARTICIPARON EN LA ELABORACIÓN.**

**Coordinador**

Henry Laniado Rodas

**11. REQUISITOS DEL PROCESOS DE ASEGURAMIENTO DE LA CALIDAD**

**Versión número:**

1,0

**Fecha elaboración:**

2019/11/12

**Fecha actualización:**

2019/11/12

**Aprobación:**

CARLOS MARIO DE JESUS VELEZ SANCHEZ