

# Teoría de la Información y Estadística

**Descripción.** En este curso se cubrirán los conceptos y herramientas básicas de teoría de la información necesarias para estudiar algunos problemas clásicos y contemporáneos en estadística. Más específicamente, se estudiarán algunos problemas relacionados con pruebas de hipótesis binarias, generalización en aprendizaje máquina y estimación.

**Prerrequisitos.** Se recomienda haber tomado los cursos de cálculo, análisis matemático, probabilidad y estadística al nivel de licenciatura. Si bien no es requisito, es deseable haber cursado al menos un curso de probabilidad, estadística o aprendizaje máquina al nivel de posgrado.

**Objetivos.** Al finalizar el curso los alumnos:

- aprenderán los conceptos y técnicas básicas de teoría de la información;
- plantearán problemas estadísticos desde un punto de vista de información;
- resolverán dichos problemas con las técnicas aprendidas.

## Contenido

### Módulo 1: Elementos de Teoría de la Información

1. Entropía y Divergencia
  - Definiciones y Propiedades Básicas
  - Submodularidad de la Entropía
  - Entropía Diferencial y sus Propiedades
2. Información Mutua
  - Desigualdad "Log-Sum"
  - Interpretación Geométrica de la Información Mutua
  - Información Mutua y Divergencia Condicionales
  - Desigualdades de Procesamiento de Datos (DPI)
  - DPI y Estadísticos Suficientes
  - Representaciones Variacionales de la Información Mutua
  - Convexidad de la Información Mutua
  - Comportamiento Local de la Información Mutua
3. f-Divergencias
  - Definición y Propiedades Básicas
  - Desigualdades de Procesamiento de Datos
  - Desigualdades entre f-Divergencias

### Módulo 2: Aplicaciones en Estadística

1. Pruebas de Hipótesis Binarias
  - Formulación de Neyman-Pearson
  - Prueba de Verosimilitud
  - Análisis Asintótico: Regímenes de Stein y Chernoff
  - Análisis Asintótico: Relación entre  $(E_0, E_1)$
2. Generalización y Estabilidad en Aprendizaje Máquina
  - Cotas de Generalización Básicas
  - Minimización del Riesgo y Adaptabilidad
  - Cotas Bayesianas para Aprendizaje PAC
  - Análisis Interactivo
3. Cotas Minimax en Estimación
  - Marco de Referencia
  - Transformando Estimación en Prueba de Hipótesis
  - Tensorización de Medidas de Divergencia
  - Métodos de LeCam, Fano y Assouad

### **Bibliografía Básica**

- Polyanskiy, Y. & Wu, Y. (2019). *Lecture Notes on Information Theory*. Lecture Notes for ECE563 (UIUC).
- Duchi, J. (2019). *Information Theory and Statistics*. Lecture Notes for STATS 311 / EE 377 (SU).
- Cover, T. M., & Thomas, J. A. (2012). *Elements of Information Theory*. John Wiley & Sons.

### **Acerca del Curso**

- Evaluación: tareas (60%) y proyecto final (40%)
- Horario: viernes de 12:30pm a 3:30pm
- Instructor: Mario Diaz, IIMAS - Oficina 115, [mario.diaz@sigma.iimas.unam.mx](mailto:mario.diaz@sigma.iimas.unam.mx)