



# ME-3004

## Probabilidad y Estadística Aplicados a la Ciencia de Datos e Inteligencia Artificial

### Duración

Cuatro semanas

### Descripción del Curso

Este curso proporciona una introducción a los conceptos de probabilidad y estadística fundamentales en el análisis de datos y su aplicación en inteligencia artificial. Los temas cubren desde la definición de probabilidad, probabilidad condicional y el Teorema de Bayes hasta aplicaciones prácticas en árboles de decisión y métodos bayesianos. Al finalizar, los estudiantes habrán desarrollado una base sólida para aplicar conceptos estadísticos en el desarrollo y optimización de modelos de Machine Learning y análisis de datos.

### Objetivos

1. Entender los conceptos básicos de probabilidad y estadística y su importancia en ciencia de datos.
2. Aplicar el Teorema de Bayes para inferencia en modelos de inteligencia artificial.
3. Utilizar probabilidades y estadísticas en la construcción y evaluación de árboles de decisión.
4. Introducir los fundamentos de los métodos bayesianos y su uso en machine learning.
5. Desarrollar habilidades para aplicar métodos probabilísticos en el análisis de datos y modelos predictivos.

## Metodología

El curso combina teoría con aplicaciones prácticas en R y Python, asegurando que los estudiantes puedan implementar y analizar modelos probabilísticos. Las herramientas y recursos incluyen:

- 6. Videoconferencias semanales:**  
Sesiones en Zoom que serán grabadas para su revisión posterior.
- 7. Trabajos prácticos semanales:**  
Ejercicios aplicados en R y Python para reforzar los conceptos de probabilidad y estadística.
- 8. Foros de discusión:**  
Espacio en Moodle para discutir temas del curso y resolver dudas.
- 9. Aula virtual en Moodle:**  
Aula virtual en Moodle: Acceso a materiales, ejemplos de código y recursos adicionales.

## Luego de este curso el estudiante será capaz de:

- 10.** Definir y calcular probabilidades en diferentes contextos.
- 11.** Aplicar el Teorema de Bayes para realizar inferencias y tomar decisiones en modelos de datos.
- 12.** Implementar árboles de decisión y analizar su rendimiento en problemas de clasificación.
- 13.** Comprender y aplicar métodos bayesianos en el desarrollo de modelos predictivos.
- 14.** Evaluar modelos de machine learning usando conceptos estadísticos fundamentales.

## Contenido

- **Semana 1:**  
**Análisis Combinatorio (aprendiendo a contar)**
  - a. Técnicas de Conteo.
  - b. Permutaciones con y sin repetición.
  - c. Combinaciones con y sin repetición.
  - d. Regla de la multiplicación.
  - e. Implementación en R y Python.
- **Semana 2:**  
**Introducción a la Probabilidad y Definición de Conceptos Básicos**
  - a. Introducción a la probabilidad y su importancia en ciencia de datos.
  - b. Definición de probabilidad, eventos y espacio muestral.
  - c. Conceptos básicos de probabilidad condicionada.

d. Teorema de Bayes.

e. Ejemplo práctico en Python y R: Aplicación del Teorema de Bayes a la Analítica Predictiva.

- **Semana 3:**  
Distribuciones de Probabilidad

a. Variables Aleatorias.

b. Distribución Normal.

c. Distribución Uniforme.

d. Otras distribuciones de probabilidad.

e. Esperanza matemática y desviación estándar.

f. Ejemplo práctico en R y Python: Implementación las distribuciones de probabilidad.

- **Semana 4:**  
Distribuciones de Probabilidad discretas y continuas

a. Distribución Binomial y su aproximación por la normal.

b. Distribución Poisson y su aproximación por la normal.

c. Distribución hipergeométrica y su aproximación por la normal.

d. Otras distribuciones de probabilidad.

e. Esperanza matemática y desviación estándar.

f. Ejemplo práctico en R y Python: Implementación las distribuciones de probabilidad discretas y continuas.

g. Ejemplo práctico en R y Python: Aplicaciones en Árboles de Decisión y otros métodos basados en árboles, como Bosques Aleatorios.

## Evaluación

La evaluación consistirá en 4 tareas prácticas con igual valor, enfocadas en aplicar conceptos de probabilidad y estadística en proyectos de ciencia de datos.

## Bibliografía

Wasserman, L. (2004). All of Statistics: A Concise Course in Statistical Inference. Springer.

Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Springer.

Murphy, K. P. (2012). Machine Learning: A Probabilistic Perspective. MIT Press.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). Bayesian Data Analysis. CRC Press.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning with Applications in R. Springer.

VanderPlas, J. (2016). Python Data Science Handbook. O'Reilly Media.

Silver, N. (2012). The Signal and the Noise: Why So Many Predictions Fail—but Some Don't. Penguin Press.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer.

Feller, W. (1968). An Introduction to Probability Theory and Its Applications. Wiley.

McGrayne, S. B. (2011). The Theory That Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines, and Emerged Triumphant from Two Centuries of Controversy. Yale University Press.

McGrayne, S. B. (2011). The Theory That Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines, and Emerged Triumphant from Two Centuries of Controversy. Yale University Press.



Programa Iberoamericano de  
Formación en Minería de Datos