

Simulación en Ciencia de Datos

CARÁCTER: Optativa

DEDICACIÓN: Bimestral

MODALIDAD: Híbrida

CARGA HORARIA: 32 horas

Profesores: Rodrigo Del Rosso - Ezequiel Nuske

OBJETIVOS DE LA ASIGNATURA

La simulación es una herramienta fundamental para validar modelos, analizar fenómenos aleatorios complejos y generar datos sintéticos cuando los datos reales son insuficientes o sensibles. En el contexto de la ciencia de datos, su aplicación permite mejorar la inferencia estadística, estimar incertidumbre y robustecer la toma de decisiones basada en modelos predictivos.

En este marco, la presente asignatura tiene por objetivo:

- Comprender las distintas técnicas de simulación aplicables a problemas de ciencia de datos.
- Aplicar simulación Monte Carlo para análisis de riesgo, inferencia y optimización.
- Desarrollar simulaciones para series de tiempo y experimentos controlados.
- Generar y validar datos sintéticos para entrenamiento de modelos.
- Estimar parámetros e interpretar incertidumbre estadística mediante resultados simulados.

UNIDADES TEMÁTICAS

Unidad 1: Introducción a la Simulación

- Fundamentos de la simulación computacional.
- Generación de números aleatorios y variables aleatorias simuladas.
- **Actividades:** Análisis y codificación de experimentos simulados en R o Python.

Unidad 2: Simulación Monte Carlo

- Principios y fundamentos.
- Aplicaciones en inferencia, integración y riesgo financiero.
- **Actividades:** Estimación de áreas, precios de opciones y análisis de varianza mediante Monte Carlo.

Unidad 3: Simulación en Series de Tiempo

- Bootstrap para series temporales.
- Simulación de caminatas aleatorias y procesos AR.
- **Actividades:** Modelado estocástico de fenómenos económicos.

Unidad 4: Simulación de Experimentos

- A/B testing y ensayos clínicos sintéticos.
- Diseño experimental computacional.
- **Actividades:** Comparación de estrategias de marketing mediante simulación.

Unidad 5: Simulación Avanzada

- Cadenas de Markov y ecuaciones diferenciales estocásticas.
- Simulación de procesos dinámicos.
- **Actividades:** Simulación de inventarios, redes y sistemas complejos.

Unidad 6: Interpretación de Resultados Simulados

- Estimación de parámetros y evaluación de modelos.
- Cálculo de error cuadrático medio y varianza de predicción.
- **Actividades:** Validación cruzada con datos sintéticos y análisis de sensibilidad.

METODOLOGÍA GENERAL DE ENSEÑANZA

La asignatura será desarrollada mediante clases interactivas teórico-prácticas, donde se explicarán los conceptos fundamentales y se aplicarán en ejemplos reproducibles con herramientas estadísticas (R/Python). Cada clase incluirá una sección práctica con asistencia docente y actividades individuales o grupales.

DEDICACIÓN Y CARGA HORARIA

- **Clases presenciales:** 32 horas totales (8 clases + evaluación final)
-

MODALIDAD Y CRITERIOS DE APROBACIÓN Y EVALUACIÓN

La evaluación será continua a través de:

- **Trabajos prácticos:** entregables individuales o grupales de cada unidad (60%)
 - **Trabajo integrador final:** proyecto de simulación aplicado con notebook reproducible y exposición (40%)
-

BIBLIOGRAFÍA

- Efron & Tibshirani – *An Introduction to the Bootstrap*
- Robert & Casella – *Monte Carlo Statistical Methods*
- Ross, S. – *Simulation*
- Templ, M. – *Simulation for Data Science with R*
- Allen Downey – *Think Bayes / Think Stats*