



Universidad Tecnológica de la Mixteca

Clave DGP: 200089

Maestría en Ciencia de Datos

PROGRAMA DE ESTUDIOS

NOMBRE DE LA ASIGNATURA

Probabilidad y Estadística para Ciencia de Datos

SEMESTRE	CLAVE DE LA ASIGNATURA	TOTAL DE HORAS
Primer semestre	371013	35 Mediación docente 65 Estudio independiente

OBJETIVO(S) GENERAL(ES) DE LA ASIGNATURA

Al término del curso, el alumno será capaz de analizar situaciones con una medida de incertidumbre y etiquetarlas con comportamientos determinados. Al mismo tiempo será capaz de organizar, describir y analizar datos obtenidos a partir de situaciones no deterministas que observe en su entorno y realizar inferencia básica sobre los mismos.

TEMAS Y SUBTEMAS

1. Probabilidad

- 1.1 Introducción a la Probabilidad y sus interpretaciones.
- 1.2 Generación de datos, espacios muestrales y propiedades de la probabilidad.
- 1.3 Técnicas de conteo y su simulación: principio de adición, principio de multiplicación, permutaciones, ordenaciones y combinaciones.
- 1.4 Experimentos independientes, probabilidad condicional y regla de la multiplicación.
- 1.5 Teorema de la probabilidad total y teorema de Bayes.

2. Variables Aleatorias y su simulación.

- 2.1 Variables Aleatorias discretas, continuas, mixtas.
- 2.2 Propiedades de una variable aleatoria: función de densidad de probabilidad, función de distribución de probabilidad. Propiedades
- 2.3 Esperanza y varianza de variables aleatorias. Propiedades.

3. Distribuciones de probabilidad y su simulación

- 3.1 Casos discretos: Bernoulli, Binomial, Geométrica, Binomial Negativa, Hipergeométrica, Poisson.
- 3.2 Casos continuos: Uniforme, Normal, Gamma (Exponencial, Gamma, Chi-cuadrada), t-student y F-Fisher.

4. Distribuciones de muestreo y su aplicación

- 4.1 Distribución de la media muestral: datos normales varianza conocida; datos normales, varianza desconocida y muestra pequeña; misma distribución en los datos, varianza desconocida y muestra grande.
- 4.2 Distribución de la diferencia de medias muestrales: datos normales varianzas conocidas; datos normales, varianzas desconocidas y muestras pequeñas; misma distribución en los datos, varianzas desconocidas y muestras grandes.
- 4.3 Distribución muestra de la proporción y diferencia de proporciones.
- 4.4 Distribución de la varianza y comparación de varianzas.

5. Pruebas de hipótesis e intervalos de confianza.

- 5.1 Definiciones, tipos de errores y nivel de significancia.
- 5.2 Pruebas unilaterales y bilaterales.
- 5.3 Prueba de Hipótesis e Intervalo de Confianza para una muestra.
- 5.4 Prueba de Hipótesis e Intervalo de Confianza para comparación de muestras.

6. Análisis de regresión lineal simple y su aplicación.

- 6.1 Introducción.
- 6.2 Coeficiente de correlación y determinación.
- 6.3 El modelo de regresión lineal simple y su significancia.
- 6.4 Intervalos de confianza para los parámetros de regresión.
- 6.5 Modelo predictor.

ACTIVIDADES DE APRENDIZAJE

El profesor siempre buscará un balance entre la teoría matemática detrás del método, su aplicación a problemas prácticos y su implementación computacional. Introducir al alumno a un lenguaje computacional de preferencia con licencia libre, por ejemplo Python, R, entre otros.

CRITERIOS Y PROCEDIMIENTOS DE EVALUACIÓN Y ACREDITACIÓN

Exámenes parciales y final. Tareas Simulaciones en computadora. Proyectos. Esto tendrá una equivalencia del 100% en la calificación final del semestre

BIBLIOGRAFÍA (TIPO, TÍTULO, AUTOR, EDITORIAL Y AÑO)**Básica:**

1. Probabilidad y Estadística con aplicaciones para Ingeniería y Ciencias Computacionales (4ta. Ed.). J. Susan Milton & Jesse C. Arnold. Mc Graw-Hill 2005.
2. Estadística Matemática con aplicaciones (2da. Ed.). William Mendenhall, Dennis D. Wackerly & Richard L. Scheaffer. Grupo Editorial Iberoamericana 1994.
3. Practical Statistics for Data Scientists: 50+ Essential Concepts Using R and Python (2da. Ed.). Peter Bruce, Andrew Bruce & Peter Gedeck. O'Reilly 2020.

Consulta:

1. Probability and Statistics for Data Science: Math + R + Data (1ra. Ed.). Norman Matloff. Chapman & Hall 2019.
2. Essential Math for Data Science: Take Control of Your Data with Fundamental Linear Algebra, Probability, and Statistics (1ra. Ed.). Thomas Nield. O'Reilly 2022.
3. Probability and Statistics Essentials for Data Science and Machine Learning: 200+ examples and pictures. Simit Tomar & Ajay Thakur. Kindle Edition 2023.

PERFIL PROFESIONAL DEL DOCENTE

Maestría o Doctorado en Ciencias Matemáticas, Matemáticas Aplicadas, Estadística, Ciencia de Datos o afines, con conocimientos en el uso de software como Python, R, Minitab, SAS, SPSS, entre otros.

Vo.Bo

M.T.C.A. ERIK GERMÁN RAMOS PÉREZ
COORDINADOR DE LA UNIVERSIDAD VIRTUAL

AUTORIZÓ

L.I. MARIO ALBERTO MORENO ROCHA
VICE-RECTOR ACADÉMICO