Nombre: Introducción a la Estadística y Ciencia de Datos **(LCD)** Estadística **Lic (M) –**

Tópicos de Estadística **(Doc)** Curso IC 2025

Puntaje: 4 puntos para Doc. DOC8800931

Total HS: 128 horas CAT 2 arancelado CATEGORÍA BAJA

Correlatividades:

- **Para la carrera de Licenciatura en Ciencias de Datos, las correlativas son:**
  Laboratorio de Datos
  Probabilidades y Estadística (M)
  Álgebra Lineal Computacional

- **Para la carrera de Licenciatura en Ciencias Matemáticas, las correlativas son**
  Probabilidades y Estadística (M) (tp aprobados)
  Cálculo Avanzado (final)

Advertencia alumnos externos: Es una asignatura que tiene un enfoque matemático, asume sólidos conocimientos de Álgebra Lineal y utiliza conceptos vistos por los alumnos en el curso de Probabilidades y Estadística (Matematicos) y en Cálculo Avanzado o Análisis Avanzado, donde se enfatiza en distintas nociones de convergencia y otras herramientas teóricas.

Doctorado: Esta materia hace fuerte uso de las herramientas matemáticas. Está aprobada para el Doctorado en Matemática como optativa, con las mismas correlativas que para la carrera de grado. Para los restantes estudiantes de doctorado de la facultad interesados en cursarla, tengan en cuenta que esta materia tiene un alto enfoque matemático, asume sólidos conocimientos de Álgebra Lineal y utiliza conceptos vistos por los alumnos en Probabilidades y Estadística (M) y en Cálculo Avanzado o Análisis Avanzado, donde se enfatiza en distintas nociones de convergencia y otras herramientas teóricas.

Carga horaria: 6 horas semanales

Carreras:
Matemática: Nombre de la materia: Estadística Código SIU: MATE820015
LCD Nombre de la materia: Introducción a la Estadística y Ciencia de Datos Código SIU: LCDA210008
Doctorado Nombre de la materia: Tópicos de Estadística Código SIU: DOC8800931

PROGRAMA: Contenidos

Análisis exploratorio de datos: Visualización y resumen de datos. Función de distribución empírica. Estimación no paramétrica de la densidad. Estimación puntual: Estimación. Sesgo, varianza y Error Cuadrático Medio.

Compromiso sesgo-varianza. Estimación plug-in utilizando el enfoque funcional. Estimación en modelos paramétricos: máxima verosimilitud, momentos, M y Z estimadores. Propiedades asintóticas: consistencia y distribución asintótica. Método delta.

Regiones de confianza: Intervalos de confianza para la media de una distribución normal con varianza conocida. Intervalos de confianza para la media de una distribución normal con varianza desconocida:

Distribución t de Student. Intervalos de confianza de nivel asintótico basados en estadísticos asintóticamente normales. Intervalo para proporciones. Intervalos de confianza para dos muestras.

Técnicas de remuestreo: Métodos Bootstrap para la estimación de la varianza de un estimador asintóticamente normal. Métodos Bootstrap para la estimación de la distribución de un estimador. Intervalo de confianza Bootstrap: percentil y asintóticamente normal. Comparación con métodos clásicos. Bootstrap no paramétrico y paramétrico.

Tests de hipótesis: Presentación del problema de test de hipótesis. Hipótesis nula y alternativa. Tipos de errores. Nivel y potencia de un test. Valor "p". Test para la media de una población normal con varianza conocida y con varianza desconocida. Tests de Wald (basados en estadísticos asintóticamente normales). Tests e intervalos de confianza para dos muestras. Relación entre tests e intervalos de confianza. El problema de comparaciones múltiples o cubrimiento simultáneo.

Modelo lineal: Regresión lineal simple. Mínimos cuadrados. Supuestos. Inferencia para los parámetros del modelo: bajo normalidad y teoría asintótica. Regresión lineal múltiple. Predicción.

Modelos de Regresión: Función de regresión. Estimación de la función de regresión. Estimadores no paramétricos: Nadaraya, kNN. Regresión no lineal. Ajuste y sobreajuste. Métodos de regularización (Ridge, LASSO, etc.). Técnicas y métricas de evaluación de modelos-métodos (validación cruzada, etc.)

Clasificación: La regla de Bayes. Regresión logística, estimación de parámetros por máxima verosimilitud. Modelos generativos: LDA, QDA, Bayes Naive. Modelos discriminativos: kNN, clasificación logística. Técnicas y métricas de evaluación de modelos-métodos. Árboles de decisión.

BIBLIOGRAFÍA

Bickel, P. J. y Doksum. K. A. (2015). Mathematical Statistics: Basic Ideas and Selected Topics, Volume 1. 2da. Edición. Chapman and Hall.

Dalgaard, P. (2008). Introductory statistics with R. Springer Science & Business Media.

Hothorn, T., y Everitt, B. S. (2014). A handbook of statistical analyses using R. CRC press.

James, G., Witten, D., Hastie, T. y Tibshirani, R. (2013). An introduction to statistical learning. New York: Springer.

Lehmann, E. L. y G. Casella. (2003). Theory of Point Estimation. 2da. Edición. Springer.

Lock, R. H., Lock, P. F. y Morgan, K. L. (2012). Statistics: Unlocking the power of data.

Wiley Global Education.

Rice, J. (2006). Mathematical statistics and data analysis. Nelson Education.

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

Springer Science & Business Media.

Wasserman, L. (2013). All of statistics: a concise course in statistical inference. Springer Science & Business Media.

Fecha marzo 2026