



Ministerio de Cultura y Educación
Universidad Nacional de San Luis
Facultad de Ciencias Físico Matemáticas y Naturales
Departamento: Física
Area: Area II: Superior y Posgrado

(Programa del año 2017)
(Programa en trámite de aprobación)
(Presentado el 29/08/2017 15:56:57)

I - Oferta Académica

Materia	Carrera	Plan	Año	Período
(MATERIA OPTATIVA I) ANALISIS DE DATOS	LIC.EN FISICA	015/06	2017	2° cuatrimestre

II - Equipo Docente

Docente	Función	Cargo	Dedicación
VALLADARES, DIEGO LEONARDO	Prof. Responsable	P.Adj Exc	40 Hs

III - Características del Curso

Credito Horario Semanal				
Teórico/Práctico	Teóricas	Prácticas de Aula	Práct. de lab/ camp/ Resid/ PIP, etc.	Total
7 Hs	4 Hs	3 Hs	Hs	7 Hs

Tipificación	Periodo
C - Teoria con prácticas de aula	2° Cuatrimestre

Duración			
Desde	Hasta	Cantidad de Semanas	Cantidad de Horas
07/08/2017	24/11/2017	16	112

IV - Fundamentación

El conocimiento de métodos de tratamiento de datos es esencial para la labor científica. Durante las últimas dos décadas se han producido avances importantes en el tema del análisis de datos, ya que ha crecido considerablemente la disponibilidad, cantidad y calidad de datos vinculados al trabajo en todas las disciplinas científicas. Además se han hecho accesibles programas de computadora en los que se pueden implementar de forma simple y eficiente métodos de análisis de datos que anteriormente resultaban prohibitivos, dado por ejemplo, el volumen de operaciones que requiere su implementación. Es por ello que en esta asignatura se pretende impartir conocimientos sobre algunas de dichas métodos. Se explicará el la descripción estadística cualitativa y cuantitativa de los datos. Se exponen las principales pruebas estadísticas que permiten valorar las hipótesis hechas sobre relaciones y comportamiento de los datos. Se exponen los conceptos básicos sobre el análisis de series temporales. Se da especial importancia a la implementación de los métodos de análisis a través de algoritmos en el lenguaje de programación R.

V - Objetivos / Resultados de Aprendizaje

- * Lograr que los alumnos adquieran conocimientos sobre métodos de análisis de datos y que sean capaces de implementarlos para una aplicación concreta.
- * Lograr que los alumnos se familiaricen con la escritura de algoritmos y con el lenguaje de programación R.

VI - Contenidos

Tema I: Tipos de datos. Parámetros estadísticos de descripción. Software de tratamiento y visualización. Introducción a R.

Tema II: Análisis exploratorio de datos. Medidas estadísticas cuantitativas de descripción.
 Tema III: Análisis de datos bivariados numéricos y no-numéricos. Correlación.
 Tema IV: Inferencia Estadística. Principales distribuciones de probabilidad. Teorema Central del Límite.
 Tema V: Intervalos de confianza a través de métodos paramétricos y no paramétricos.
 Tema VI: El concepto de prueba de hipótesis y su implementación en la estimación de estadísticos y comparación de muestras.
 Tema VII: Bondad de un ajuste. Pruebas de Pearson, Kolmogorov-Smirnov, Shapiro.
 Tema VIII: Inferencia estadística en la regresión lineal.
 Tema IX: Series temporales. Descomposición. Autocorrelación. Diferentes modelos de análisis.
 Tema X: Análisis exploratorio de datos multivariantes. Gráficas. Análisis de la componente principal. Análisis discriminante.

VII - Plan de Trabajos Prácticos

Cada tema del programa posee un práctico a realizar por los alumnos. Cada práctico se compone de problemas en que se plantea una situación o inferencia necesaria a realizar sobre un conjunto de datos, la cual se debe resolver o responder implementando un algoritmo en R.

VIII - Regimen de Aprobación

La asignatura es de carácter promocional y se aprueba realizando todos los prácticos de problemas de manera correcta, los cuales son corregidos por el equipo de la asignatura.

IX - Bibliografía Básica

- [1] Verzani J. Using R for Introductory Statistics, Chapman & Hall/CRC (2005).
- [2] Maindonald J. y Braun J. W. Using R: an Example-Based Approach, Cambridge University Press (2010).
- [3] Cowpertwait P. S. P. y Metcalfe A. V. Introductory Time Series with R. Springer (2008).
- [4] Reimann C., Filzmoser P., Garrett R. G., y Dutter R. Statistical data analysis explained: Applied Environmental Statistics with R, John Wiley & Sons Ltd (2008).

X - Bibliografía Complementaria

[1]

XI - Resumen de Objetivos

XII - Resumen del Programa

Tema I: Lenguaje R.
 Introducción a R. ¿Qué es R?. Características de R. Diseño del sistem. Fuentes para aprender R. Otras fuentes. Instalación del sistema base y librerías. Uso de la consola y las GUIs. Rstudio.
 Programación en R. Ingresando datos por consola. Algunos comandos básicos. Objetos. Atributos. Creando vectores. Secuencias. Objetos mixtos y coerción. Coerción explícita. Factores. Matrices. cbind y rbind. Listas. Valores no definidos. Data Frames. Nombrado objetos. Arrays. Series temporales.

Tema II: Análisis exploratorio de datos
 ¿Qué es el análisis de datos? Tipos de datos. Datos geolocalizados. Revisión de algunos conceptos de Estadística. Medidas estadísticas de datos numéricos. Medidas de tendencia central. Medidas de dispersión de los datos. Análisis Exploratorio de Datos univariados. Histograma. La función densidad(). Características de las distribuciones. El diagrama de caja o boxplot.

Tema III: Regresión
 Datos bivariados. Datos no numéricos. Datos numéricos. La correlación entre dos conjuntos de datos. Regresión de modelos. Regresión lineal simple. Hipótesis del modelo. La función lm(). Valores predichos. Coeficiente de determinación. Regresión lineal simple con intercepción nula. Residuos. Regresión lineal simple con fórmulas no lineales. Regresión no lineal.

Tema IV: Inferencia

Revisión de conceptos relacionados con variables aleatorias. Distribuciones de probabilidad. Valor esperado de una variable aleatoria. Muestreo. Distribuciones de probabilidad. Distribuciones de probabilidad discretas: Bernoulli, Binomial y de Poisson. Distribuciones de probabilidad continuas: Exponencial, Normal o de Gauss, Lognormal. Teorema central del límite. Simulación de un proceso aleatorio.

Tema V: Intervalos de confianza

Intervalos de confianza. Intervalo de confianza para una proporción. Intervalo de confianza para la media de la población. Gráficas Q-Q. Intervalo de confianza para diferencias. Diferencia de proporciones. Diferencia de medias en muestras independientes. Diferencia de medias en muestras pareadas. Métodos no paramétricos. Intervalo de confianza para la mediana basado en la distribución binomial.

Tema VI: Pruebas de hipótesis

Prueba de hipótesis. Prueba de hipótesis para una proporción. Prueba de hipótesis para la media. Prueba de hipótesis para la comparación de muestras. Prueba de hipótesis para comparar proporciones. Prueba de hipótesis para comparar medias.

Tema VII: Bondad de un ajuste

Bondad de un ajuste. La prueba Chi cuadrado de Pearson. La distribución multinomial. El estadístico de Pearson. Prueba de hipótesis para distribuciones continuas. Prueba Kolmogorov-Smirnov. Prueba Shapiro-Wilk. Parámetros de una distribución.

Tema VIII: Inferencia estadística en la regresión lineal

Inferencia estadística en la regresión lineal. Prueba gráfica de las hipótesis del modelo. Gráficas comparativa del modelo. Verificación de las hipótesis del modelo. La distancia de Cook. Intervalo de confianza de los parámetros de la regresión.

Tema IX: Series temporales

Series temporales. Análisis exploratorio mediante gráficas. Funciones y repositorios de R para series temporales. Series estocásticas. Descomposición de series. Tendencia, variación estacional y componente aleatoria. Modelos aditivo y multiplicativo. Correlación. Estacionariedad y ergodicidad de la serie. Autocorrelación. Modelos.

Tema X: Análisis Multivariante

Análisis exploratorio de datos multivariantes. Scatterplots. Biplots. Análisis de la componente principal. Análisis discriminante.

XIII - Imprevistos

--

XIV - Otros

--

ELEVACIÓN y APROBACIÓN DE ESTE PROGRAMA	
	Profesor Responsable
Firma:	
Aclaración:	
Fecha:	