

PROGRAMA DEL CURSO: "OBTENCION Y ANALISIS DE DATOS"

CURSO OBLIGATORIO PARA LAS MAESTRIAS DE CIENCIAS AMBIENTALES Y DEL PEDECIBA BIOLOGIA (Botánica, C. Fisiológicas, Ecología, Genética, Microbiología, Zoología) Y

Coordinador y Docente Responsable: Dr. Walter Norbis

Duración: semestral (clases de dos horas de duración que se dictarán los días martes, miércoles y jueves de 8 a 10 horas). Prácticos (Comienzan después del primer mes de teórico): lunes, jueves y viernes sala de informática, Facultad de Ciencias (horario: 8 a 10 horas) (Se coordinaran días y horas con los estudiantes).

Evaluación: Cada estudiante debe presentar a lo largo del curso y de manera obligatoria, un seminario teórico basado en la separata de un trabajo o un tema previamente seleccionado. Este seminario tendrá una nota de calificación. Al final del curso el estudiante deberá realizar un trabajo final escrito que deberá contener una breve introducción del problema a abordar, los objetivos del trabajo, el planteo de hipótesis (y aquellas que puedan ser estadísticamente verificables), una descripción fundamentada de los métodos a utilizar, como serán o fueron obtenidos los datos y un desarrollo completo del análisis de datos y sus resultados. Plazo de presentación del trabajo: 30 de Junio de 2017 sin excepciones. Además: Asistencia al 75% de las clases teóricas y 90% de las clases practicas.

Fecha de comienzo: Agosto (se fijara dia), 8.30 AM (salón 210).

MODULO 1

TEMA 1. Niveles de clasificación y ordenación. Atributos y variables continuas, discontinuas y discretas. Ordenación y clasificación. Niveles de información: agrupación, clasificación, adición, sustracción (variables discretas y continuas).

TEMA 2. Estadística descriptiva. Distribuciones de frecuencias (continuas y discontinuas). Discretización de variables continuas (intervalos de clase). Curvas de frecuencia acumulada. Medidas de posición: moda, mediana, media (aritmética, geométrica, armónica) y cuantiles. Principales propiedades. Medidas de dispersión: rango, desviación media, varianza, desvío estándar, coeficiente de variación. Asimetría y kurtosis. Propiedades. Análisis exploratorio de datos: objetivos. Box and Whisker Plot. Interpretación. Otros métodos.

TEMA 3. Experimentos aleatorios y espacio de muestreo: concepto. Variables aleatorias. Definición. Funciones de probabilidad y distribución. Esperanza y varianza. Funciones: Binomial, Poisson y Normal. Funciones de densidad generadas por la normal: Chi-cuadrado, Student y F (Snedecor).

MODULO 2

TEMA 4. Teoría y Técnicas de muestreo. Obtención de datos: objetivos e hipótesis. Tipos de datos. Variabilidad espacial y temporal. Población o universo. Censo. Muestra. Conveniencia y ventajas del muestreo. Muestreo probabilístico: aleatorio, estratificado (tipos), sistemático, conglomerados. Otros diseños. Ventajas y desventajas. Factores que afectan el muestreo. Tamaño de muestra y costo. Diseño optimo. Métodos de re - muestreo: "Jackknife y Bootstrap".

MODULO 3

TEMA 5. Estimación y test de hipótesis. Teorema central del límite. Parámetros y estimadores: estimador de la media y varianza. Propiedades. Estimador del cociente de varianzas. Hipótesis nula. Errores tipo I y II. Hipótesis alternativa. Potencia de un test.

TEMA 6. Supuestos básicos en estadísticas paramétricas. Aleatoriedad, Independencia, normalidad, homocedasticidad. Test. Transformaciones. Violación de los supuestos y supuestos básicos en estadística no-paramétrica. Ventajas y desventajas.

TEMA 7. Análisis de muestras y test de hipótesis. Comparación de dos muestras. Planteos básicos y supuestos. Test de Student. Comparación de varias muestras: análisis de varianza (ANVA). Supuestos. ANVA factorial y anidado. Test "a priori" y "a posteriori".

TEMA 8. Relación de dos variables. Correlación. Regresión lineal. Regresión funcional. Ajuste y validación del modelo. Análisis de varianza del modelo. Regresión potencial y exponencial. Análisis de covarianza. Regresión logística.

TEMA 9. Análisis no – paramétrico: Test Binomial; Chi – cuadrado; Mac Nemar; Test de los signos; Wilcoxon; Fisher; Mann – Whitney; Kolmogorov – Smirnov; Kruskal – Wallis; Friedman; Correlación no - paramétrica: Spearman y Kendall. Otros test.

MODULO 4

TEMA 10. Relación de tres o más variables. Repaso sobre matrices y determinantes: adición y producto de matrices. Determinantes. Rango de una matriz. Matriz inversa y traspuesta. Valores propios y vectores propios. Propiedades. Correlación y regresión múltiple. Modelos lineales generales (GLM) y modelos aditivos generales (GAM).

TEMA 11. Análisis multidimensional. Análisis numérico de datos múltiples: matriz básica de datos. Datos cualitativos y cuantitativos. Transformaciones. Clasificación: índices de similitud. Propiedades.

TEMA 12. Análisis Factorial. Matrices de dispersión y de correlación. Análisis de Componentes Principales. Interpretación. Análisis Factorial de Correspondencias. Interpretación. Otros análisis derivados. Análisis multidimensional métrico y no métrico. Ventajas y desventajas.

BIBLIOGRAFIA BASICA RECOMENDADA

(Durante el curso se brinda una lista de bibliografía y separatas que amplían cada uno de los temas que forman parte de cada módulo).

- Cochran, W. (1986) Métodos de muestreo. CECSA, México. (La versión original del año 1977 es en inglés).
- Conover, W. (1980) Practical nonparametric statistic. Second Edition. John Wiley and Sons.
- Crisci, J. y M.F. Armengol (1983) Introducción a la teoría y práctica de la taxonomía numérica. OEA, Serie de Biología. Monografía número 26.
- Cuadras, C. (1993) Análisis Multivariante. EUNIBAR, Barcelona.
- Daniels, W. (1989) Bioestadística. Métodos aplicados a las Ciencias Biológicas y de la Salud.
- Dobson A. 2002. An introduction to generalized linear models. Chapman & Hall. New York.
- Kent, J. and D. Cooker (1994). Vegetation Data Analysis.
- Legendre, P. and L. Legendre. 2012. Numerical ecology. Elsevier
- Manly, B.F.J., 1994. Multivariate Statistical Methods: A Primer. Second Edition. Chapman & Hall/CRC, London, New York
- Plá, L. (1986) Análisis multivariado: método de componentes principales. OEA, Serie de matemáticas. Monografía número 27.
- Santaló, L. (1975) Probabilidad e inferencia estadística. OEA, Serie de Matemática. Monografía número 11.
- Sneath, P. and R. Sokal (1973) Numerical taxonomy. The principles and practice of numerical classification. Freeman, San Francisco.
- Siegel, S. (1956) Non – parametric Statistic for the Behavioral Sciences. Mc Graw – Hill. New York.
- Sokal, R. y F. Rohlf (1979) Biometría. Principios y métodos estadísticos en la investigación biológica. Editorial BLUME. Barcelona. (Hay ediciones en inglés más recientes, de los años 1981 y 1995 y 1998).
- Steel, R. y J. Torrie (1985) Bioestadística: principios y procedimientos. Editorial MacGraw - Hill.
- Snedecor, G. y W. Cochran (1984) Métodos Estadísticos. CECSA, México.
- Zar, J. (1994) Biostatistical analysis. Prentice - Hall International Editions.(Hay ediciones mas recientes)