



UNIVERSIDAD DE COSTA RICA
ESCUELA DE BIOLOGÍA
Tópicos en Bioestadística - R
B-0
I-2013

Requisitos	B304 y B305.
Profesor	Dr. Eric J. Fuchs eric.fuchs@ucr.ac.cr Of. 44. Escuela de Biología
Horas Lectivas	Teoría y práctica: 3 horas.
Horario	L: 15:00 - 18:00
Créditos	4

Descripción

El curso presenta un repaso de los temas en bioestadística más utilizados por los investigadores en ciencias biológicas. Sin embargo, el objetivo principal del curso es que los estudiantes se familiaricen con el lenguaje de programación en bioestadística R. Dicho programa es de distribución libre y además varios colaboradores han desarrollado múltiples métodos y librerías para realizar análisis en diferentes áreas de la ciencia, que van desde análisis de datos cuantitativos, cualitativos, imágenes, datos ecológicos y genéticos. Dado que es de distribución libre (i.e., open source) y la gran diversidad de métodos que realiza, se ha convertido en el paquete de más uso en biología. R es también un lenguaje de programación por lo que permite desarrollar procedimientos imposibles de realizar en programas tradicionales. No obstante, al ser un lenguaje de programación basado en funciones, su uso requiere una introducción a la sintaxis del lenguaje y el uso y creación de las funciones.

Objetivos

Objetivo general

Guiar a los estudiantes en el uso del lenguaje de programación bioestadística R (cran.r-project.org) con ejemplos de pruebas bioestadísticas.

1. Conocer el lenguaje de programación bioestadística R
2. Aprender a realizar pruebas estadísticas básicas en R
3. Aprender a analizar modelos lineales en R
- 4.

Metodología para cumplir los objetivos

Se darán clases magistrales en el cuarto de cómputo, sobre las pruebas estadísticas más utilizadas en bioestadística. Las prácticas se realizarán utilizando las computadoras del cuarto, donde los estudiantes aprenderán de forma interactiva a utilizar el lenguaje de programación R. Se presentará material de apoyo donde se resuman los comandos e instrucciones aprendidas en cada sesión, además se dará atención individualizada a los estudiantes con el uso de R.

Evaluación

Tareas	30%
Presentación oral tema de investigación	30%
Presentación escrita de tema de investigación	40%
Total	100%

Tareas

Todas las semanas se otorgarán tareas para que los estudiantes practiquen las pruebas estadísticas y los procedimientos o comandos en R. Las tareas deberán entregarse de forma electrónica una semana después de asignada. Las tareas son individuales.

Tema de Investigación

La investigación dirigida tiene como objetivo que los estudiantes exploren pruebas y procedimientos en R que no fueron cubiertos en el curso. Específicamente se pretende que los estudiantes, habiendo completado una introducción al uso de R, se familiaricen con los múltiples procedimientos que este lenguaje incluye en diversas ramas de la biología. Los estudiantes escogerán un tema específico de interés en biología (e.g. genética, ecología, análisis de imágenes, etc.) y buscarán cuáles son las librerías en R que tienen procedimientos estadísticos en el tema.

Los estudiantes deberán preparar una presentación oral que introduzca la librería, citando brevemente los análisis que dicha librería puede emplear. Finalmente, deberán escoger un procedimiento o prueba estadística dentro del paquete y presentarle en detalle a los estudiantes del curso. Para la presentación deberán incluir un juego de datos (original de preferencia) y deberán realizar el análisis completo en forma demostrativa. La interpretación del análisis deberá exponerse claramente.

De igual forma, deberán presentar un resumen escrito que incluya: el nombre y objetivo de la librería, un resumen de los procedimientos que incluye, el procedimiento detallado para realizar la prueba que expusieron, incluyendo los resultados, figuras e interpretación. Los datos deberán hacerse disponible a los compañeros vía la red del cuarto de cómputo, para que los participantes del curso puedan replicar los análisis realizados por el expositor. El resumen deberá también incluir una introducción de las pruebas estadísticas realizadas y de su utilidad e importancia. Se debe incluir una bibliografía de las pruebas (accesible por internet) y ejemplos de su uso e interpretación.

Cronograma y temas

Semana	Contenido
1	Introducción al curso. Introducción a R
2	Importar y exportar datos en R
3	Manipulación de datos en R.
4	Medidas de posición central y Medidas de dispersión
5	Probabilidad y distribuciones. Figuras QQ y normal
6	Prueba de hipótesis. Prueba t-student de una muestra Prueba t-student para dos muestras
7	Pruebas no paramétricas para dos muestras Prueba de supuestos Prueba Wilcoxon Prueba U-Mann-Whitney
8	Análisis de Varianza (ANDEVA) de una vía Modelos lineares Prueba de supuestos
9	Análisis de Varianza de dos vías Modelo factorial balanceado Estudio de la interacción
10	Análisis de Varianza de dos vías Modelo factorial desbalanceado Optimización de modelo (model parsimonioso) Sumas de cuadrado tipo III
11	Análisis de Regresión (LM) Modelo lineal Regresión múltiple
12	Modelos lineares generalizados (GLM) GLM con distribución Poisson GLM con datos binomiales
13	Análisis de frecuencias Análisis de tablas Análisis de Chi-cuadrado Análisis de log-linear
14	Presentación oral de temas de investigación
15	Presentación oral de temas de investigación

Bibliografía

Fuentes recomendables en Internet