

Aspectos generales

Título:	Introducción al Análisis Estadístico con R (CUPO LLENO)
Programas de posgrado o planes de estudio en donde se ofertará adicionalmente:	
PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS MÉDICAS ODONTOLÓGICAS Y DE LA SALUD, UNAM POSGRADO CIENCIAS BIOLÓGICAS, UNAM.	
Área del conocimiento:	Farmacología, toxicología y salud ambiental
Semestre:	2025-2
Modalidad:	Tópico selecto
Horario:	Miércoles de 17:00 A 19:30 h
No. sesiones:	16
Horas por sesión:	2.5
Total alumnos PDCB:	5
Total alumnos:	10
Videoconferencia:	Si
Lugar donde se imparte:	SALA DE JUNTAS DEL DEPARTAMENTO DE FARMACOLOGÍA FACULTAD DE MEDICINA, UNAM
Informes:	molina_ja2007@yahoo.com.mx

Métodos de evaluación

MÉTODO	PORCENTAJE	NOTAS
Examen	30%	El examen es único, se realiza en 10 minutos y es frente al profesor o mostrando su pantalla en zoom
Participación en clase	50%	La asistencia y participación se evalúan cada clase y se asigna un porcentaje proporcional a las 15 sesiones del curso
Trabajos y ejercicios	20%	Durante el curso, se asignan 4 tareas

Contribución de este curso/tópico en la formación del alumnado del PDCB:

El análisis de datos es fundamental en la investigación clínica y biomédica. La representación adecuada de los datos y la elección de la prueba de inferencia adecuada son primordiales en la investigación y representan el inicio de análisis de mayor complejidad. Este curso contribuirá a que el alumno pueda realizar el adecuado manejo de cualquier tipo de datos, representación e inferencia, con el objetivo de obtener información y realizar discusión y conclusiones basados en datos existentes y/o recabados.

Profesor (a) responsable

Nombre:	Molina Guarneros Juan Arcadio
Teléfono:	(55) 5611092747
Email:	molina_ja2007@yahoo.com.mx

Profesores (as) participantes

PARTICIPANTE	ENTIDAD O ADSCRIPCIÓN	SESIONES
MOLINA GUARNEROS JUAN ARCADIO Responsable	Facultad de Medicina	1.-INTRODUCCIÓN A R 2.-TIPO DE DATOS 3.-TALLER DE MANEJO DE DATOS 6.-INTRODUCCIÓN AL ANÁLISIS DESCRIPTIVO: VARIABLES CUANTITATIVAS. 7.-INTRODUCCIÓN AL ANÁLISIS DESCRIPTIVO: VARIABLES CUANTITATIVAS. 8.-INTRODUCCIÓN AL ANÁLISIS DESCRIPTIVO: VARIABLES CUALITATIVAS. 14.-ANÁLISIS DE REGRESIÓN LINEAL 15.-ANÁLISIS DE REGRESIÓN LOGÍSTICA

ORTEGA AYALA ADIEL
Integrante

DEPARTAMENTO DE FARMACOLOGIA, FACULTAD
DE MEDICINA, UNAM

4.-LISTAS Y DATA FRAME.
5.-ESTRUCTURAS DE CONTROL Y MANEJO DE DATOS.
9.-INTRODUCCION A GGLOT2
10.-INFERENCIA CON PRUEBAS PARAMÉTRICAS PARA 1
Y 2 MUESTRAS
11.-INFERENCIA CON PRUEBAS PARAMÉTRICAS PARA 3
O MAS MUESTRAS
12.-INFERENCIA CON PRUEBAS NO PARAMÉTRICAS
PARA 2 MUESTRAS
13.-INFERENCIA CON PRUEBAS NO PARAMÉTRICAS
PARA MAS DE 3 MUESTRAS.
16.-EXAMEN

Introducción

Dudemos de los datos hasta que los datos no dejen lugar a dudas, fue una frase que dejó Henri Poincaré, responsable de la Conjetura de Poincaré, uno de los 7 problemas del milenio en matemáticas y, recientemente demostrado por el matemático ruso Grigori Perelman. El manejo y análisis de los datos es un proceso crucial en la generación de conocimiento y elección de análisis y algoritmos de mayor complejidad. Este curso representa una vista introductoria para el manejo y análisis de los datos en investigación médica y biomédica.

Temario

DR. MOLINA GUARNEROS JUAN ARCADIO
Responsable

- 1.-INTRODUCCIÓN A R. (Miércoles 29 de enero de 2025).
- 2.-TIPO DE DATOS. (Miércoles 5 de febrero de 2025).
- 3.-TALLER DE MANEJO DE DATOS. (Miércoles 12 de febrero de 2025).
- 4.-LISTAS Y DATA FRAME. (Miércoles 19 de febrero de 2025).
- 5.-ESTRUCTURAS DE CONTROL Y MANEJO DE DATOS. (Miércoles 26 de febrero de 2025).
- 6.-INTRODUCCIÓN AL ANÁLISIS DESCRIPTIVO. (Miércoles 5 de marzo de 2025).

VARIABLES CUANTITATIVAS.

- 7.-INTRODUCCIÓN AL ANÁLISIS DESCRIPTIVO. VARIABLES CUANTITATIVAS. (Miércoles 12 de marzo de 2025).
- 8.-INTRODUCCIÓN AL ANÁLISIS DESCRIPTIVO. (Miércoles 19 de marzo de 2025).
- 9.-INTRODUCCION A GGLOT2. (Miércoles 26 de marzo de 2025).
- 10.-INFERENCIA CON PRUEBAS PARAMÉTRICAS PARA 1 Y 2 MUESTRAS. (Miércoles 02 de abril de 2025).
- 11.-INFERENCIA CON PRUEBAS PARAMÉTRICAS PARA 3 O MAS MUESTRAS. (Miércoles 09 de abril de 2025).

VARIABLES CUALITATIVAS.

- 12.-INFERENCIA CON PRUEBAS NO PARAMÉTRICAS PARA 2 MUESTRAS. (Miércoles 23 de abril de 2025)
- 13.-INFERENCIA CON PRUEBAS NO PARAMÉTRICAS PARA MAS DE 3 MUESTRAS. (Miércoles 30 de abril de 2025).
- 14.-ANÁLISIS DE REGRESIÓN LINEAL. (Miércoles 07 de mayo de 2025)
- 15.-ANÁLISIS DE REGRESIÓN LOGÍSTICA. (Miércoles 14 de mayo de 2025).
- 16.-EXAMEN. (Miércoles 21 de mayo de 2025)

UNIDAD 1. INTRODUCCIÓN.

Introducción a R.

Instalación de R y Rstudio.

Características de R.

Ayuda en R.

Instalación de paquetes.

UNIDAD 2. TIPOS DE DATOS.

Tipos de datos.

Manejo de los datos.

Vectores

Acceso a elementos de un vector.

Operaciones con vectores.

Matrices.

Acceso a elementos de una matriz.

Factores y vectores caracteres.

Listas.

Acceso a elementos de una lista.

Data frames.

Importación y exportación de bases de datos en Excel.

Acceso a porciones o subconjuntos de datos.

UNIDAD 3. ESTRUCTURAS DE CONTROL Y MANEJO DE DATOS.

Tipos de operadores.

Uso de operadores.

Manipulación de matrices y data frames.

Introducción al uso de Tidyverse.

Manipulación de una base de datos: El uso de dplyr.

UNIDAD 4. INTRODUCCIÓN AL ANÁLISIS DESCRIPTIVO CON R: VARIABLES CUANTITATIVAS

Funciones estadísticas.

Medidas de tendencia central y dispersión.

Características de la curva normal.

¿Qué es la libre distribución?

Medidas de tendencia central y de dispersión.

Pruebas de normalidad.

Prueba de Shapiro-Wilk, Prueba de Kolmogorov-Smirnov con corrección de Lilliefors.

Prueba de hipótesis.

UNIDAD 5. INTRODUCCIÓN AL ANÁLISIS ESTADÍSTICO DESCRIPTIVO CON R: VARIABLES CUALITATIVAS. Cálculo de frecuencias y porcentajes Tablas de contingencia

Funciones gráficas

Gráficos de mosaico

Pruebas de inferencia.

Prueba Chi cuadrada de Pearson.

Prueba exacta de Fisher.

UNIDAD 6. INTRODUCCIÓN A GGLOT2.

Funciones gráficas.

Gráfico de barras.

Gráfico de cajas.

Gráfico de violín.

Gráfico tipo streapchart.

Gráfico de puntos.

Gráfico de correlación.

Uso de una paquetería de interfaz gráfica.

UNIDAD 7. PRUEBAS DE INFERENCIA ESTADÍSTICA. Inferencia con pruebas paramétricas: Prueba de t para muestras independientes.

Prueba de t para muestras relacionadas.

Prueba ANOVA.

Pruebas posthoc.

Inferencia con pruebas no paramétricas.

Prueba U de Mann Whitney.

Prueba de Kruskal Wallis.

UNIDAD 8. ANÁLISIS DE REGRESIÓN.

Regresión lineal.

Regresión logística.

Bibliografía

El arte de programar en R: un lenguaje para la estadística / Julio Sergio Santana Sepúlveda y Efraín Mateos Farfán 2014. Introducción al análisis de datos con R. Alfonso Urquía.

https://es.r4ds.hadley.nz/?fbclid=IwAR2ixKq5_wEeT0Pd782XPwLAyprC5SpAuID_Dw1Npry0oD_WHaKOgEoVKdU <http://www.cookbook-r.com/Graphs/>

<https://es.r4ds.hadley.nz/transform.html>

Observaciones

Desglose de sesiones a impartir.

•

Sesión 1: INTRODUCCIÓN A R (Dr. Juan Arcadio Molina Guarneros).

Sesión 2: TIPOS DE DATOS (Dr. Juan Arcadio Molina Guarneros).

Sesión 3: TALLER DE MANEJO DE DATOS (Dr. Juan Arcadio Molina Guarneros).

Sesión 4: LISTAS Y DATA FRAMES (Dr. en C. Adiel Ortega Ayala).

Sesión 5: ESTRUCTURAS DE CONTROL Y MANEJO DE DATOS (Dr. en C. Adiel Ortega Ayala).

Sesión 6: INTRODUCCIÓN AL ANÁLISIS DESCRIPTIVO: VARIABLES CUANTITATIVAS (Dr. Juan Arcadio Molina Guarneros).

Sesión 7: INTRODUCCIÓN AL ANÁLISIS DESCRIPTIVO: VARIABLES CUANTITATIVAS (Dr. Juan Arcadio Molina Guarneros).

Sesión 8: INTRODUCCIÓN AL ANÁLISIS DESCRIPTIVO: VARIABLES CUALITATIVAS (Dr. Juan Arcadio Molina Guarneros).

Sesión 9: INTRODUCCIÓN A GGLOT2 (Dr. en C. Adiel Ortega Ayala).

Sesión 10: INFERENCIA CON PRUEBAS PARAMÉTRICAS PARA 1 Y 2 MUESTRAS (Dr. en C. Adiel Ortega Ayala).

Sesión 11: INFERENCIA CON PRUEBAS PARAMÉTRICAS PARA 3 O MÁS MUESTRAS (Dr. en C. Adiel Ortega Ayala).

Sesión 12: INFERENCIA CON PRUEBAS NO PARAMÉTRICAS PARA 2 MUESTRAS (Dr. en C. Adiel Ortega Ayala).

Sesión 13: INFERENCIA CON PRUEBAS PARAMÉTRICAS PARA 3 O MÁS MUESTRAS (Dr. en C. Adiel Ortega Ayala).

Sesión 14: ANÁLISIS DE REGRESIÓN LINEAL (Dr. Juan Arcadio Molina Guarneros).

Sesión 15: ANÁLISIS DE REGRESIÓN LOGÍSTICA (Dr. Juan Arcadio Molina Guarneros).

Sesión 16: EXAMEN.

El alumno deberá contar con equipo de cómputo propio y transportable (laptop). Las computadoras que los alumnos poseen, suelen cubrir los requisitos mínimos para el curso, los cuales son:

1. Laptop con cualquier sistema operativo al que esté acostumbrado el alumno (Windows, Mac OS, GNU/Linux, etc).
2. Memoria RAM mínimo 8 GB.
3. Procesador: A partir de Intel Core i5, AMD Ryzen 5 o equivalente.
4. Disco duro: Al menos de 128 GB SSD o mecánico.
5. Pantalla: Idealmente superior a 14 pulgadas.

Observaciones sobre la evaluación:

Participación en clase:

Durante cada clase, el profesor explica scripts en el lenguaje R y asigna ejercicios correspondientes a cada clase. Los alumnos deben enviar su script vía chat de Zoom para corroborar que el ejercicio esté realizado de manera correcta. En caso de que algún alumno tenga problemas, se le facilita la acción "compartir pantalla" y presenta su problema. El profesor invita a los alumnos a solucionar el problema del o la compañera que está compartiendo pantalla. La participación culmina cuando el alumno que presenta pantalla escribe el script correcto y muestra resultados frente al grupo.

La asistencia y participación se evalúan cada clase y se asigna un porcentaje proporcional a las 15 sesiones del curso.

La sesión 16 es el examen.

Trabajos y ejercicios:

Durante el curso, se asignan 4 tareas:

- 1.- Análisis descriptivo de variables.
- 2.- Análisis inferencial de variables entre 2 grupos independientes.
- 3.- Análisis inferencial y post hoc de variables entre 3 grupos independientes.
- 4.- Modelo de regresión (lineal o logístico). En cada una de estas tareas, se solicitan los siguientes resultados:
 - 1.- Tabla con estadísticos descriptivos, inferencia y descripción de la tabla.
 - 2.- Script utilizado en R.

Los criterios de evaluación son:

1. Elección de los estadísticos descriptivos correctos de acuerdo a las pruebas de normalidad.
2. Elección de la prueba de inferencia adecuada según la distribución de las variables por grupo.
3. Interpretación de resultados/significancia estadística.
4. Descripción detallada de la tabla.
5. Script utilizado en R para la elaboración de la tarea.

Cada uno de estos puntos tiene un valor de 2.0 puntos.

Examen:

El examen es único, se realiza en 10 minutos y es frente al profesor o mostrando su pantalla en zoom.