

### 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. PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS BIOLÓGICAS, UNAM.
Área del conocimiento:	Genética, genómica y bioinformática
Semestre:	2027-1
Modalidad:	Curso fundamental
Horario:	Miércoles de 17:00 a 19: 30h
No. sesiones:	16
Horas por sesión:	2.5
Total alumnos PDCB:	6
Total alumnos:	10
Videoconferencia:	Si
Lugar donde se imparte:	Sala de juntas del Departamento de Farmacología de la Facultad de Medicina de la UNAM.
Informes:	ad.ortega@unam.mx

### Métodos de evaluación

MÉTODO	PORCENTAJE	NOTAS
Evaluación de la participación en clase	50%	Durante cada clase, el profesor explica los scripts de R y asigna ejercicios del tema o de dominio del lenguaje R. Los alumnos foráneos y aquellos que acuden de forma presencial, deberán unirse a una sesión de Zoom para poder enviar en el chat, los script
Examen	30%	Se realizará un examen en la sesión 16. Se realiza en 10 minutos y es frente al profesor y/o mostrando su pantalla en zoom. La calificación es numérica entre 0 y 10. Son 10 elementos para evaluar y cada uno tiene valor de 1 punto. El script utilizado para
Trabajos y ejercicios	20%	Durante el curso, se asignan 4 tareas: 1. Manejo de datos perdidos. 2. Elaboración de un modelo de ANOVA de medidas repetidas. 3. Elaboración de un modelo de Riesgos proporcionales de Cox. 4. Aplicación de un modelo avanzado (clustering o de reducción de

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

El análisis de inferencia es fundamental para el aprendizaje del análisis científico de los datos, sin embargo, el análisis de varias variables contribuye a un mejor entendimiento de los fenómenos en las ciencias biomédicas. Este curso contribuirá a que las y los alumnos, aumenten su criterio estadístico a partir de pruebas descriptivas y de análisis bivariados para poder elegir y ejecutar la o las pruebas adecuadas de análisis múltiple para el óptimo análisis de sus datos y la obtención de conclusiones.

### Profesor (a) responsable

Nombre:	Ortega Ayala Adiel
Teléfono:	
Email:	ad.ortega@unam.mx

### Profesores (as) participantes

PARTICIPANTE	ENTIDAD O ADSCRIPCIÓN	SESIONES
--------------	-----------------------	----------

<b>ORTEGA AYALA ADIEL</b> Responsable	Facultad de Medicina	ESTRUCTURAS DE CONTROL Y MANEJO DE DATOS INFERENCIA CON PRUEBAS NO PARAMÉTRICAS PARA 2 MUESTRAS INFERENCIA CON PRUEBAS NO PARAMÉTRICAS PARA 3 O MÁS MUESTRAS INFERENCIA CON PRUEBAS PARAMÉTRICAS PARA 1 Y 2 MUESTRAS INFERENCIA CON PRUEBAS PARAMÉTRICAS PARA 3 O MÁS MUESTRAS INTRODUCCIÓN A GGLOT2 INTRODUCCIÓN AL ANÁLISIS DESCRIPTIVO: VARIABLES CUALITATIVAS LISTAS Y DATA FRAMES
<b>MOLINA GUARNEROS JUAN ARCADIO</b> Integrante	Facultad de Medicina	ANÁLISIS DE REGRESIÓN LINEAL ANÁLISIS DE REGRESIÓN LOGÍSTICA EXAMEN Miércoles 25 de noviembre de 2026 INTRODUCCIÓN A R INTRODUCCIÓN AL ANÁLISIS DESCRIPTIVO: VARIABLES CUANTITATIVAS INTRODUCCIÓN AL ANÁLISIS DESCRIPTIVO: VARIABLES CUANTITATIVAS TALLER DE MANEJO DE DATOS TIPOS DE DATOS en R y TIPOS DE VARIABLE EN ESTADÍSTICA

## 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 Perelmán. 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

- 1 Sesión 1: INTRODUCCIÓN A R (Dr. Juan Arcadio Molina Guarneros, 2.5h). Miércoles 05 de agosto de 2026.
  - 1.1 Introducción a R.
  - 1.2 Instalación de R y Rstudio.
  - 1.3 Características de R.
  - 1.4 Ayuda en R.
  - 1.5 Instalación de paquetes.
- 2 Sesión 2: TIPOS DE DATOS en R y TIPOS DE VARIABLE EN ESTADÍSTICA (Dr. Juan Arcadio Molina Guarneros). Miércoles 12 de agosto de 2026.
  - 2.1 Tipos de datos en R
    - 2.1.1 Numéricos
    - 2.1.2 Enteros
    - 2.1.3 Carácter
    - 2.1.4 Factor
    - 2.1.5 Lógico
  - 2.2 Tipos de variable en estadística
    - 2.2.1 Cuantitativas
    - 2.2.2 Cualitativas
- 3 Sesión 3: TALLER DE MANEJO DE DATOS (Dr. Juan Arcadio Molina Guarneros). Miércoles 19 de agosto de 2026.
  - 3.1 Manejo de los datos.
    - 3.1.1 Vectores
    - 3.1.2 Acceso a elementos de un vector.
    - 3.1.3 Operaciones con vectores.
    - 3.1.4 Matrices.
    - 3.1.5 Acceso a elementos de una matriz.
    - 3.1.6 Factores y vectores caracteres.
    - 3.1.7 R como calculadora
- 4 Sesión 4: LISTAS Y DATA FRAMES (Dr. en C. Adiel Ortega Ayala). Miércoles 26 de agosto de 2026.
  - 4.1 Listas.
  - 4.2 Acceso a elementos de una lista.
  - 4.3 Data frames.
  - 4.4 Importación y exportación de bases de datos en Excel.
  - 4.5 Acceso a porciones o subconjuntos de datos.
- 5 Sesión 5: ESTRUCTURAS DE CONTROL Y MANEJO DE DATOS (Dr. en C. Adiel Ortega Ayala). Miércoles 02 de septiembre de 2026.
  - 5.1 Tipos de operadores.
  - 5.2 Uso de operadores.
  - 5.3 Manipulación de matrices y data frames.
  - 5.4 Introducción al uso de Tidyverse.
  - 5.5 Manipulación de una base de datos: El uso de dplyr y tidyr
- 6 Sesión 6: INTRODUCCIÓN AL ANÁLISIS DESCRIPTIVO: VARIABLES CUANTITATIVAS (Dr. Juan Arcadio Molina Guarneros). Miércoles 09 de septiembre de 2026.

- 6.1 Funciones estadísticas.
- 6.2 Medidas de tendencia central y dispersión.
- 6.3 Características de la curva normal.
- 6.4 ¿Qué es la libre distribución?
- 6.5 Medidas de tendencia central y de dispersión.
- 6.6 Pruebas de normalidad.
- 6.7 Prueba de Shapiro-Wilk y Prueba de Kolmogorov-Smirnov con corrección de Lilliefors.
- 6.8 Pruebas de hipótesis.
- 7 Sesión 7: INTRODUCCIÓN AL ANÁLISIS DESCRIPTIVO: VARIABLES CUANTITATIVAS (Dr. Juan Arcadio Molina Guarneros). Miércoles 23 de septiembre de 2026.
- 7.1 Medidas de tendencia central y de dispersión
- 7.2 Media, mediana y moda
- 7.3 Desviación estándar
- 7.4 Error estándar
- 7.5 ¿Qué es el intervalo de confianza al 95%?
- 7.6 Percentiles
- 7.7 Rango intercuartilar
- 7.8 Descripción de variables con distribución normal y libre distribución
- 7.9 Paquetería numSummary
- 8 Sesión 8: INTRODUCCIÓN AL ANÁLISIS DESCRIPTIVO: VARIABLES CUALITATIVAS (Dr. Juan Arcadio Molina Guarneros). Miércoles 30 de septiembre de 2026.
- 8.1 Cálculo de frecuencias y porcentajes Tablas de
- 8.2 contingencia
- 8.3 Funciones gráficas
- 8.4 Gráficos de mosaico
- 8.5 Pruebas de inferencia.
- 8.6 Prueba Chi cuadrada de Pearson.
- 8.7 Prueba exacta de Fisher.
- 9 Sesión 9: INTRODUCCIÓN A GGLOT2 (Dr. en C. Adiel Ortega Ayala). Miércoles 07 de octubre de 2026.
- 9.1 Funciones gráficas.
- 9.2 Gráfico de barras.
- 9.3 Gráfico de cajas.
- 9.4 Gráfico de violín.
- 9.5 Gráfico tipo streapchart.
- 9.6 Gráfico de puntos.
- 9.7 Gráfico de correlación.
- 9.8 Uso de una paquetería de interfaz gráfica.
- 10 Sesión 10: INFERENCIA CON PRUEBAS PARAMÉTRICAS PARA 1 Y 2 MUESTRAS (Dr. en C. Adiel Ortega Ayala). Miércoles 14 de octubre de 2026.
- 10.1 Prueba de t para una muestra.
- 10.2 Prueba t de Student para muestras independientes.
- 10.3 Prueba t para muestras relacionadas.
- 10.4 Prueba t de 2 colas.
- 10.5 Prueba t de una cola.
- 11 Sesión 11: INFERENCIA CON PRUEBAS PARAMÉTRICAS PARA 3 O MÁS MUESTRAS (Dr. en C. Adiel Ortega Ayala). Miércoles 21 de octubre de 2026.
- 11.1 Prueba de análisis de varianza de 1 vía.
- 11.2 Evaluación de supuestos de ANOVA.
- 11.3 Pruebas post hoc
- 12 Sesión 12: INFERENCIA CON PRUEBAS NO PARAMÉTRICAS PARA 2 MUESTRAS (Dr. en C. Adiel Ortega Ayala). Miércoles 28 de octubre de 2026.
- 12.1 Prueba U de Mann Whitney
- 12.2 Prueba de la suma de rangos con signo de Wilcoxon
- 13 Sesión 13: INFERENCIA CON PRUEBAS NO PARAMÉTRICAS PARA 3 O MÁS MUESTRAS (Dr. en C. Adiel Ortega Ayala). Miércoles 04 de noviembre de 2026.
- 13.1 Prueba de Kruskal Wallis
- 13.2 Pruebas post hoc
- 14 Sesión 14: ANÁLISIS DE REGRESIÓN LINEAL (Dr. Juan Arcadio Molina Guarneros). Miércoles 11 de noviembre de 2026.
- 14.1 ¿Qué es una regresión lineal?
- 14.2 El coeficiente de correlación en regresión lineal.
- 14.3 Evaluación de supuestos de la regresión lineal.
- 14.4 Ajuste de valores predichos en una regresión lineal.
- 15 Sesión 15: ANÁLISIS DE REGRESIÓN LOGÍSTICA (Dr. Juan Arcadio Molina Guarneros). Miércoles 18 de noviembre de 2026.
- 15.1 Evaluación de supuestos de una regresión logística.
- 15.2 ¿Por qué no una regresión lineal?
- 15.3 Medidas de asociación.
- 15.4 Interpretación del OR en regresión logística.
- 15.5 El p valor y el estadístico de Wald.
- 16 Sesión 16: EXAMEN Miércoles 25 de noviembre de 2026.

## Bibliografía

1. Raykov, T., & Marcoulides, G. A. (2012). Basic statistics: An introduction with R. Routledge. ISBN 978-1-4422-1846-8 (cloth); ISBN 978-1-4422-1847-5 (pbk).
2. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning: with Applications in R (Vol. 103). Springer. <https://doi.org/10.1007/978-1-4614-7138-7>.
3. Fernández Casal, R., Costa Bouzas, J., & Oviedo de la Fuente, M. (2024). Métodos predictivos de aprendizaje estadístico. Universidade da Coruña. <https://doi.org/10.17979/spudc.9788497498937>
4. Crawley, M. J. (2015). Statistics: An introduction using R (2ª ed.). John Wiley & Sons. <https://doi.org/10.1002/9781118941126>

## Observaciones

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, las cuales son:

1. Laptop con cualquier sistema operativo al que esté acostumbrado el alumno (Windows, Mac OS, GNU/Linux, etc).
2. Memoria RAM mínimo 16 GB.
3. Procesador: A partir de Intel Core i5, AMD Ryzen 5 o equivalente.
4. Disco duro: Al menos de 512 GB SSD o mecánico.
5. Pantalla: Idealmente superior a 14 pulgadas.

Evaluación de la participación en clase.

Durante cada clase, el profesor explica los scripts de R y asigna ejercicios del tema o de dominio del lenguaje R. Los alumnos foráneos y aquellos que acuden de forma presencial, deberán unirse a una sesión de Zoom para poder enviar en el chat, los scripts de sus ejercicios realizados en clase.

En caso de que algún alumno tenga problemas, se le facilita la acción "compartir pantalla" y presenta su problema. El profesor invita 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 17 sesiones del curso.

Trabajos y ejercicios.

Durante el curso, se asignan 4 tareas:

1. Manejo de datos perdidos.
2. Elaboración de un modelo de ANOVA de medidas repetidas.
3. Elaboración de un modelo de Riesgos proporcionales de Cox.
4. Aplicación de un modelo avanzado (clustering o de reducción de dimensión).

Criterios de evaluación de los trabajos y ejercicios.

1. Presentación de estadísticos descriptivos.
2. Presentación de la inferencia estadística correspondiente.
3. Elaboración de gráficos.
4. Script utilizado en R.
5. Descripción escrita de las conclusiones.

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

Examen.

Se realizará un examen en la sesión 16. Se realiza en 10 minutos y es frente al profesor y/o mostrando su pantalla en zoom. La calificación es numérica entre 0 y 10. Son 10 elementos para evaluar y cada uno tiene valor de 1 punto. El script utilizado para el examen deberá ser enviado al profesor antes de finalizar el examen.