

II Workshop y Seminarios de estadística aplicada

23-26 Enero 2018

Este evento consta de dos secciones, un workshop y un set de seminarios. Los objetivos de cada sección se detallan a continuación. Importantemente, los cupos para ambas secciones son limitados. Se puede postular a cupo para una o ambas secciones.

**Objetivos Generales**

El objetivo principal de este evento es acercar de manera práctica la aplicación de herramientas estadísticas a publico general.

**Objetivos del Seminario**

Presentar técnicas de análisis avanzadas implementadas en problemas concretos haciendo hincapié en la implementación y contexto de aplicación más que en los detalles técnicos.

**Objetivos del workshop**

Este curso tiene por objetivo entregar una formación práctica y básica del análisis de datos. El curso se centra en la aplicación de pruebas estadísticas, su interpretación, herramientas de visualización y conceptos básicos relevantes para una correcta interpretación de resultados. Asimismo, se espera que los participantes terminen no sólo con la habilidad de aplicar test básicos estadísticos, sino que también sepan distinguir las herramientas más apropiadas para desarrollar sus análisis. Este curso tiene un fuerte énfasis práctico, donde los aspectos teóricos y matemáticos se abordaran sólo de acuerdo a las necesidades puntuales de los participantes. Por el carácter práctico de este curso, no se verán contenidos ligados a cómo generar pruebas estadísticas ni la matemática que subyace a las pruebas que se mostrarán.

**Público Objetivo**

Este evento es abierto y gratuito para todo aquel que quiera participar. No obstante, se priorizaran personas que estén analizando datos para el desarrollo de sus investigaciones y con mayor experiencia en investigación. Además se priorizará la diversidad de origen de los estudiantes. La intención de estos criterios es que podamos maximizar el número de equipos de investigación que pueden tener acceso al workshop, así como priorizar gente que dirige y apoya activamente los miembros de su laboratorio.

**Requisitos**

Saber programar y haber tenido cursos de estadística previos **es deseable, pero no obligatorio** para tomar este workshop. Las personas que deseen asistir deben contar con su propio computador personal, **con Project R y R Studio instalados**. A los estudiantes seleccionados se les enviará un correo con las instrucciones para poder instalar estos programas en caso de no tenerlos.

**Modo de trabajo**

El workshop consta de un breve módulo teórico seguido de los módulos prácticos correspondientes, donde se analizarán e interpretaran diversas bases de datos. Durante las tardes se espera que los estudiantes apliquen lo aprendido sobre sus datos. Se darán 30 minutos a partir del segundo día del workshop para que los estudiantes compartan sus dudas y dificultades para aplicar los conocimientos a sus datos, así como dudas de interpretación.

**Profesores**

Dr. Rodrigo Vergara

Dr. (c) Miguel Concha

**Programa Workshop**

|  |  |  |  |
| --- | --- | --- | --- |
|  | Teórico | Taller | Fecha y Hora |
| Introducción a análisis de datos | Bienvenida al workshop  Aspectos generales del análisis de datos |  | Martes 23  9:00-9:30 |
| Aproximación a R | Manejo de Variables  Conversión de variables  Funciones descriptivas básicas  Funciones gráficas básicas. | Conceptos mínimos de programación para usar R.  Obtención de estadística descriptiva. (medidas centrales y dispersiones). | Martes 23  9:30-10:50 |
| Manejo de datos | Selección de casos  Tipos de variables  Construcción de Bases de Datos  Tratamiento de medidas repetidas  Variables categóricas vs continuas (y ordinales) | **Base de datos: Nutrición.csv**  Construcción de bases de datos.  Importación de datos.  Subselección de datos  Gráficos de distribuciones.  Gráficos para detección de outliers.  Boxplots  Scatter plots | Martes 23  11:00-11:50 |
| Conceptos básicos  Y comparación de dos muestras | Distribuciones  Medidas de centralidad  Medidas de dispersión  Poder y tamaño del efecto  Pareadas  No paramétrica | **Bases de datos: BlinkAvg.csv**  Aplicación de t-test (pareado y no pareado)  Aplicación de Mann-Whitney U Test  Aplicación de Wilcoxon Test  Presentación de resultados  Interpretación de resultados | Martes 23  12:00-13:00 |
| Comparando muchas muestras | ANOVA de una vía y Post Hoc | **Base de datos: BlinkAvg.csv**  Aplicación de ANOVA  Aplicación de post hoc | Miércoles 24  9:30-10:50 |
| Comparando muchas muestras | Pareadas  Mixtas  No paramétricas  Interpretación de interacciones  Comparaciones post hoc | **Base de datos: framingham.csv**  Aplicación de ANOVA de medidas repetidas  paramétricos  Aplicación de Kruskal Wallis Test  Aplicación de Friedman Test  Aplicación de post hoc no paramétricos | Miércoles 24  11:00-12:20 |
| **Resolución de dudas en la aplicación a datos traídos por los estudiantes** | | | Miércoles 24  12:30-13:00 |
| Correlaciones | Introducción al tratamiento de variables continúas.  Interpretación de correlaciones.  Visualización y comprensión de hipótesis nula y obtención del R cuadrado | **Base de datos: Penguin.sav**  Aplicación de correlaciones (spearman y Pearson)  Matrices de correlación  Visualización de correlaciones | Jueves 25  9:30-10:50 |
| Regresiones simples | Similitudes de las regresiones y las correlaciones.  Interpretación de coeficientes. | **Base de datos: vino.csv**  Aplicación de modelos de regresión simple  Revisión de supuestos. | Jueves 25  11:00-12:20 |
| **Resolución de dudas en la aplicación a datos traídos por los estudiantes** | | | Jueves 25  12:30-13:00 |
| Análisis de tendencias | Interpretación de p valores  Cálculos de muestra necesaria  Evaluación de incremento de muestra  Evaluación de tamaño de efecto y poder | **Base de datos: vino.csv**  Estimación de tamaño muestral  Estimación de tamaños de efecto  Estimación de poder  (Usaremos G-power) | Viernes 26  9:30-10:50 |
| **Resolución de dudas en la aplicación a datos traídos por los estudiantes** | | | Viernes 26  11:15-13:00 |

**Horarios y Presentaciones de los seminarios**

|  |  |  |  |
| --- | --- | --- | --- |
| **Horario** | **Martes 23** | **Miércoles 24** | **Jueves 25** |
| **14:00-15:30** | Regresiones multiples  Dr. Rodrigo Vergara | Introducción a inferencia Bayesiana  Ismael Jaras | Análisis de senderos  Dr. Rodrigo Vergara |
| **15:45-17:15** | Árboles de regresión y clasificación  Dr. Rodrigo Assar | Redes Neuronales  Ismael Jaras | Reducción de dimensiones  PhD (c) Miguel Concha |