

Universidad Nacional del Altiplano
Facultad de Ingeniería Estadística e Informática
Docente: Fred Torres Cruz
Estudiante: Anderson Brian Flores Suaña

REPOSITORIO: https://github.com/andersonfloress/Trabajos_Est_Computacional

ENLACE APLICACIÓN: <https://anderson-fls.shinyapps.io/MiApp/>

TAREA 03

Documentación de la Aplicación Shiny para Análisis Estadístico Interactivo

Introducción:

Esta aplicación desarrollada en Shiny permite realizar pruebas estadísticas básicas y visualizar datos de forma interactiva. Los usuarios pueden cargar archivos Excel, seleccionar variables para análisis, y elegir entre diferentes tipos de gráficos para representar los resultados. Es especialmente útil para estudiantes, investigadores y profesionales que requieren un análisis ágil y sencillo.

Características Principales:

- **1. Carga de Datos**

Formato de archivo: Se aceptan archivos Excel en formato `.xlsx`.

Requisitos del dato: Los datos deben ser numéricos si se desea que la tabla sea pivotada completamente. Caso contrario, se aceptan tipos variados. Cada fila debe representar una observación y cada columna una variable.

Pivoteo automático: Si se detectan varias columnas numéricas, la aplicación las transforma automáticamente para facilitar el análisis.

- **2. Selección de Prueba Estadística**

Prueba t: Comparación de medias entre dos grupos.

Requisito: Debe haber exactamente dos grupos en la variable seleccionada.

ANOVA + Tukey: Comparación de medias entre tres o más grupos. Permite identificar diferencias significativas.

Requisito: Al menos tres grupos presentes en la variable de agrupación.

- **3. Selección de Variables**

Variable de Agrupación: Define los grupos (por ejemplo: “Tratamiento” con A, B, C).

Variable de Valor: Contiene los valores numéricos a analizar (por ejemplo: “Resultados”).

- **4. Filtrado de Grupos (para ANOVA)**

Si existen más de tres grupos, se puede elegir cuáles incluir en el análisis ANOVA.

- **5. Tipos de Gráficos**

Boxplot: Representa la distribución de los datos por grupo.

Barras con Error Estándar: Muestra las medias por grupo con su error estándar.

Gráfico de Puntos: Muestra cada observación como un punto.

Gráfico de Violín: Une la densidad de datos con un boxplot.

Histogramas Agrupados: Histograma por cada grupo con sus respectivas frecuencias.

- **6. Ejecución de la Prueba**

Luego de seleccionar las opciones, hacer clic en **“Ejecutar Prueba”** para obtener los resultados.

Modo de Uso:

1. Hacer clic en **“Subir archivo Excel”** y seleccionar el archivo.
2. Si existen varias columnas numéricas, estas se pivotarán automáticamente.
3. Elegir la prueba deseada: **t-test (2 grupos)** o **ANOVA (más de 2 grupos) + Tukey**.
4. Seleccionar las variables de **Agrupación** y **Valor**.
5. (Opcional) Para ANOVA, elegir los grupos específicos a analizar.
6. Escoger el tipo de gráfico.
7. Pulsar **“Ejecutar Prueba”** para visualizar resultados y gráficos.

Resultados:

Pruebas estadísticas:

Se mostrarán los resultados del t-test o ANOVA, incluyendo el test de Tukey si corresponde. El valor-p indicará si las diferencias son estadísticamente significativas.

Gráficos:

Se mostrará el gráfico seleccionado, ayudando a comprender visualmente los datos.

Consideraciones:

- Verificar que se cumplan los requisitos: t-test necesita dos grupos; ANOVA, mínimo tres.
- En caso contrario, la app mostrará un mensaje de advertencia.
- Los datos deben estar correctamente estructurados para un análisis válido.

Ejemplos de Uso

1. Análisis de Experimento Científico

Escenario: Un investigador desea comparar el efecto de dos tratamientos sobre una variable. Puede cargar los datos, elegir la variable de agrupación (Tratamiento A y B) y la variable de valor, y ejecutar un t-test.

Visualización: Boxplot o gráfico de barras con error estándar.

2. Comparación de Grupos en Ciencias Sociales

Escenario: Un estudio social analiza la media de tres o más grupos. Se selecciona ANOVA, se eligen las variables correspondientes y se ejecuta la prueba. Luego, el test de Tukey identifica los pares con diferencias significativas.

Visualización: Gráfico de violín o histogramas agrupados.

Interpretación de Resultados

1. Prueba t

Hipótesis:

H: Las medias de los dos grupos son iguales.

H: Las medias son diferentes.

Interpretación:

■ Valor-p:

- $p < 0,05$: Se rechaza H — hay diferencia significativa.
- $p \geq 0,05$: No se rechaza H — no hay evidencia suficiente.

■ Valor-t:

- Valor alto (positivo o negativo): fuerte evidencia contra H.
- Valor cercano a cero: diferencias poco significativas.

■ Casos posibles:

- **Diferencia significativa** ($p < 0,05$):
 - t positivo: el primer grupo tiene mayor media.
 - t negativo: el segundo grupo tiene mayor media.
- **No significativa** ($p \geq 0,05$): No hay diferencia.

2. ANOVA + Tukey

Hipótesis:

H: Todas las medias son iguales.

H: Al menos una media es diferente.

Interpretación:

■ Valor-p de ANOVA:

- $p < 0,05$: Rechazar H — al menos un grupo difiere.
- $p \geq 0,05$: No se rechaza H — no hay diferencia clara.

■ Test de Tukey:

- Pares con $p < 0,05$: diferencia significativa.
- Pares con $p \geq 0,05$: no significativa.
- **Casos posibles:**
 - **Diferencias significativas (ANOVA $p < 0,05$):** Tukey identifica los pares distintos.
 - **Sin diferencias (ANOVA $p \geq 0,05$):** Todos los grupos son similares.

Limitaciones de la Aplicación

1. Formatos de archivo

Limitación: Solo se aceptan archivos Excel (.xlsx, .xls). No se permiten CSV o JSON.

Solución futura: Incluir soporte para otros formatos.

2. Tipos de pruebas estadísticas

Limitación: Solo se incluyen t-test y ANOVA con Tukey. No hay regresión lineal o ANOVA multivariado.

Solución futura: Ampliar con más métodos estadísticos.

3. Tipo de datos

Limitación: Solo se analiza información numérica. No se soportan variables categóricas o mixtas.

Solución futura: Incluir soporte para variables no numéricas.

4. Interfaz de usuario

Limitación: Puede resultar compleja para usuarios sin experiencia en R o estadística.

Solución futura: Mejorar la interfaz para mayor usabilidad.

5. Escalabilidad

Limitación: El rendimiento puede verse afectado con datos muy grandes.

Solución futura: Optimizar la aplicación para datasets de gran tamaño.

Evidencia:

