



RETOS INTEGRADORES – PASO A PASO



Objetivo: Aplicar pruebas de hipótesis, inferencia estadística, y visualización de datos reales o simulados utilizando herramientas de Python como pandas, scipy.stats, matplotlib y seaborn.



Proyecto 1: Física – Tiempo de caída libre

Contexto: Se mide el tiempo de caída de una esfera desde una altura fija. Se quiere saber si la media observada coincide con el valor teórico (2.02 s).



Pasos a seguir:

1. Importar librerías necesarias (pandas, numpy, scipy.stats, matplotlib.pyplot, seaborn).
2. Crear o cargar los datos simulados como lista y convertirlos en un DataFrame.
3. Calcular:
 - o **Media muestral**
 - o **Desviación estándar**
 - o **Error estándar**
4. Formular hipótesis:
 - o $H_0: \mu = 2.02$
 - o $H_1: \mu \neq 2.02$
5. Aplicar **prueba t bilateral** con ttest_1samp de scipy.stats.
6. Calcular:
 - o **Intervalo de confianza al 95%**
 - o **Valor p**
 - o **Tamaño muestral necesario para un error de ± 0.02 s**
7. Simular 100 muestras (de tamaño n) y calcular medias.
8. Visualizar con:
 - o Histograma de los datos originales
 - o KDE de las 100 medias muestrales
9. ¿Se rechaza H_0 con el valor p obtenido?
¿Qué significa ese resultado en el contexto del problema?
¿La diferencia es significativa o atribuible al azar?



Proyecto 2: Forestal – Salud de los árboles

Contexto: Se quiere probar si más del 85% de los árboles están sanos.



Pasos a seguir:

1. Crear la lista binaria estado (1 = sano, 0 = enfermo).
2. Cargarla en un DataFrame.
3. Calcular la **proporción de árboles sanos**.
4. Formular hipótesis:
 - o $H_0: p \leq 0.85$
 - o $H_1: p > 0.85$

5. Realizar una prueba de proporciones:
 - o Usar fórmula Z para proporciones con `statsmodels.stats.proportion`
 6. Calcular:
 - o **Intervalo de confianza del 90%**
 - o **Valor p**
 - o **Tamaño muestral necesario para margen de $\pm 3\%$ con 95% de confianza**
 7. Visualizar:
 - o **Gráfico de barras** de árboles sanos/enfermos
 - o **Curva binomial teórica** para comparar la proporción observada vs. esperada
 8. ¿Se rechaza H_0 con el valor p obtenido?
¿Qué significa ese resultado en el contexto del problema?
¿La diferencia es significativa o atribuible al azar?
-

Proyecto 3: Geofísica - Magnitud de sismos

Contexto: Comparar la magnitud media reciente con el histórico (3.5), sabiendo que $\sigma = 0.3$

Pasos a seguir:

1. Cargar los datos de magnitudes en un DataFrame.
 2. Calcular:
 - o **Media muestral**
 - o **Error estándar** usando $\sigma = 0.3$
 3. Formular hipótesis:
 - o $H_0: \mu \leq 3.5$
 - o $H_1: \mu > 3.5$
 4. Calcular **estadístico Z** (como σ es conocida).
 5. Calcular:
 - o **Intervalo de confianza al 95%**
 - o **Valor p**
 - o **Tamaño muestral para detectar diferencia de 0.2**
 6. Visualizar:
 - o Histograma de magnitudes
 - o Curva normal con la zona crítica marcada
 7. ¿Se rechaza H_0 con el valor p obtenido?
¿Qué significa ese resultado en el contexto del problema?
¿La diferencia es significativa o atribuible al azar?
 - 8.
-

Proyecto 4: Petróleo - Presión de pozos

Contexto: Verificar si la media de presión en los pozos es diferente de 2500 psi.

Pasos a seguir:

1. Cargar las presiones simuladas en un DataFrame.
2. Calcular:

- **Media**
 - **Desviación estándar**
 - **Error estándar**
3. Formular hipótesis:
 - $H_0: \mu = 2500$
 - $H_1: \mu \neq 2500$
 4. Aplicar prueba t bilateral (σ desconocida).
 5. Calcular:
 - **IC al 95%**
 - **Valor p**
 - **Nivel de confianza necesario para detectar ± 20 psi**
 6. Visualizar:
 - Histograma de presiones
 - Curva t de Student con zona de rechazo
 7. ¿Se rechaza H_0 con el valor p obtenido?
¿Qué significa ese resultado en el contexto del problema?
¿La diferencia es significativa o atribuible al azar?

Proyecto 5: Energía Solar – Eficiencia de paneles

Contexto: Verificar si la eficiencia media de los paneles es al menos 16%.

Pasos a seguir:

1. Cargar los datos de eficiencia en un DataFrame.
2. Calcular:
 - **Media**
 - **Desviación estándar**
 - **Error estándar**
3. Formular hipótesis:
 - $H_0: \mu < 16$
 - $H_1: \mu \geq 16$
4. Aplicar prueba t unilateral (σ desconocida).
5. Calcular:
 - **IC al 95%**
 - **Valor p**
 - **Tamaño muestral necesario para margen $\pm 0.2\%$**
6. Visualizar:
 - KDE (seaborn)
 - Curva t con zona de rechazo
7. ¿Se rechaza H_0 con el valor p obtenido?
¿Qué significa ese resultado en el contexto del problema?
¿La diferencia es significativa o atribuible al azar?
- 8.

Proyecto 6: Ingeniería Civil – Resistencia del concreto

Contexto: Evaluar si la resistencia media de bloques cumple con mínimo de 250 kg/cm².



Pasos a seguir:

1. Cargar las resistencias simuladas en un DataFrame.
2. Calcular:
 - o **Media**
 - o **Desviación estándar**
 - o **Error estándar**
3. Formular hipótesis:
 - o $H_0: \mu < 250$
 - o $H_1: \mu \geq 250$
4. Aplicar prueba t unilateral (σ desconocida).
5. Calcular:
 - o **IC al 95%**
 - o **Valor p**
 - o **Tamaño muestral para margen ± 5 kg/cm²**
6. Visualizar:
 - o Histograma y KDE
 - o Curva t de Student y zonas críticas
7. ¿Se rechaza H_0 con el valor p obtenido?
¿Qué significa ese resultado en el contexto del problema?
¿La diferencia es significativa o atribuible al azar?
- 8.