SESIÓN TEST DE SIGNIFICANCIA

CONTENIDOS:

- ¿Qué es una prueba de hipótesis?
- Hipótesis nula y alternativa.
- Significancia estadística.
- El valor p.
- Pruebas sobre una proporción de la población.
- Pruebas sobre una media poblacional.
- Errores tipo i y tipo ii.
- Pruebas de hipótesis.
- Diseño de experimento.
- Poder de experimento.

¿QUÉ ES UNA PRUEBA DE HIPÓTESIS?

Las pruebas de hipótesis son una herramienta esencial en la ciencia de datos para tomar decisiones basadas en datos. Una prueba de hipótesis es un procedimiento estadístico que permite evaluar la validez de una afirmación sobre una población basada en datos muestrales. Este proceso se basa en la comparación entre dos hipótesis: la hipótesis nula (H0) y la hipótesis alternativa (H1).

Objetivo de una Prueba de Hipótesis

El objetivo principal de una prueba de hipótesis es determinar si hay suficiente evidencia estadística en los datos muestrales para rechazar la hipótesis nula (H0) a favor de la hipótesis alternativa (H1). Si no hay suficiente evidencia, no se rechaza H0, lo que no implica que H0 sea verdadera, sino que no se puede descartar con los datos disponibles.

Ejemplo: Probar si la media de una población es igual a un valor específico, como determinar si el ingreso promedio mensual de una población es de $50,000.

HIPÓTESIS NULA Y ALTERNATIVA

Hipótesis Nula (H0):

Representa la afirmación que se quiere contrastar. Suele ser una afirmación de "no efecto" o "no diferencia". Es la hipótesis que se asume verdadera hasta que los datos muestrales proporcionen suficiente evidencia para rechazarla.

- Ejemplo: H0 : $\mu=50$ (La media poblacional es igual a 50).

Hipótesis Alternativa (H1):

Representa la afirmación que se quiere probar. Puede ser unilateral (una cola) o bilateral (dos colas), dependiendo de la naturaleza del problema.

- Ejemplo unilateral: H1: $\mu > 50$ (La media poblacional es mayor que 50).
- Ejemplo bilateral: H1: $\mu \neq 50$ (La media poblacional no es igual a 50).

Pasos para Realizar una Prueba de Hipótesis

1. Formular las hipótesis (H0 y H1):
   Definir claramente la afirmación que se quiere probar y su contraparte.
2. Seleccionar el nivel de significancia (α):
   El nivel de significancia (α) es la probabilidad de cometer un error tipo I (rechazar H0 cuando es verdadera). Valores comunes son $\alpha=0.05$, $\alpha=0.01$, o $\alpha=0.10$.
3. Calcular el estadístico de prueba:
   El estadístico de prueba mide qué tan alejados están los datos muestrales de lo que se esperaría bajo H0. Ejemplos incluyen:
   - Para medias: Estadístico Z o t.
   - Para proporciones: Estadístico Z.
4. Determinar la región crítica o calcular el valor-p:
   - La región crítica es el rango de valores del estadístico de prueba que llevan al rechazo de H0.
   - El valor-p es la probabilidad de observar un resultado tan extremo como el obtenido (o más extremo) bajo H0.
     Si $p<\alpha$, se rechaza H0.
5. Tomar una decisión:

- Si el valor-p es menor que α, se rechaza H0.
- Si el valor-p es mayor o igual que α, no se rechaza H0.

6. Interpretar los resultados:
   Explicar la decisión en términos del contexto del problema.

Ejemplo Práctico

Supongamos que queremos probar si la media de una población es igual a 50. Los datos muestrales son:

- Media muestral ($\bar{X}$): 52,
- Desviación estándar poblacional (σ): 10,
- Tamaño de la muestra (n): 36,
- Nivel de significancia (α): 0.05.

1. Formulación de las hipótesis:

$$H_0 : \mu = 50 \quad \text{(La media poblacional es igual a 50).}$$
$$H_1 : \mu \neq 50 \quad \text{(La media poblacional no es igual a 50).}$$

Ilustración 1 Formulación de las hipótesis

2. Cálculo del estadístico de prueba (Z):

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

Ilustración 2 Cálculo del estadístico de prueba

Sustituyendo los valores:

$$Z = \frac{52 - 50}{10/\sqrt{36}} = \frac{2}{10/6} = \frac{2}{1.67} \approx 1.20$$

Ilustración 3 Sustituyendo valores

3. Determinar el valor-p:

Para una prueba bilateral, el valor-p es el área bajo la curva normal estándar más allá de $|Z|=1.20$. Usando una tabla o software:

$$p = 2 \cdot P(Z > 1.20) \approx 2 \cdot 0.1151 = 0.2302$$

Ilustración 4 Determinar valor-p

4. Tomar una decisión:

Dado que $p=0.2302>\alpha=0.05$, no se rechaza $H0$.

5. Interpretación:

No hay suficiente evidencia para concluir que la media poblacional es diferente de 50.

Implementación en Python

A continuación, te muestro cómo realizar este ejemplo en Python utilizando scipy.stats.

```python
from scipy.stats import norm

# Parámetros
mu_0 = 50       # Media bajo la hipótesis nula
sigma = 10      # Desviación estándar poblacional
n = 36          # Tamaño de la muestra
X_bar = 52      # Media muestral
alpha = 0.05    # Nivel de significancia

# Paso 1: Calcular el estadístico Z
Z = (X_bar - mu_0) / (sigma / np.sqrt(n))

# Paso 2: Calcular el valor-p (prueba bilateral)
p_valor = 2 * (1 - norm.cdf(abs(Z)))

# Resultados
print(f"Estadístico Z: {Z:.2f}")
print(f"Valor-p: {p_valor:.4f}")

# Decisión
if p_valor < alpha:
    print("Se rechaza H0.")
else:
    print("No se rechaza H0.")
```

Salida esperada:

```
Estadístico Z: 1.20
Valor-p: 0.2302
No se rechaza H0.
```

Ilustración 5 Salida esperada

SIGNIFICANCIA ESTADÍSTICA

La significancia estadística mide la probabilidad de que los resultados observados en un estudio o experimento se deban únicamente al azar, en lugar de reflejar una relación o efecto real en la población. Es una herramienta clave para evaluar si los hallazgos son lo suficientemente robustos como para ser considerados significativos desde un punto de vista estadístico.

Nivel de Significancia (α):

El nivel de significancia (α) es el umbral predefinido que determina cuándo rechazar la hipótesis nula (H0). Representa la probabilidad máxima aceptable de cometer un error tipo I (rechazar H0 cuando es verdadera).

Valores comunes para α son:

- $\alpha=0.05$: Se utiliza en la mayoría de los estudios.
- $\alpha=0.01$: Se emplea cuando se requiere mayor certeza (menor riesgo de error tipo I).
- $\alpha=0.10$: Se usa en estudios exploratorios donde se permite más flexibilidad.

Regla de Decisión:

- Si el valor-p (probabilidad asociada a los datos bajo H0) es menor que α, se concluye que los resultados son estadísticamente significativos y se rechaza H0.
- Si el valor-p es mayor o igual que α, no hay suficiente evidencia para rechazar H0.

EL VALOR P

El valor-p es una medida que cuantifica la fuerza de la evidencia en contra de la hipótesis nula (H0). Representa la probabilidad de obtener un resultado al menos tan extremo como el observado en los datos, asumiendo que H0 es verdadera.

Interpretación del Valor-p:

Valor-p < α:

- Hay suficiente evidencia para rechazar H0.
- Los resultados son estadísticamente significativos.
- Indica que los datos son poco probables bajo H0, sugiriendo que el efecto observado podría ser real.

Valor-p ≥ α:

- No hay suficiente evidencia para rechazar H0.
- Los resultados no son estadísticamente significativos.
- Esto no prueba que H0 sea verdadera, sino que los datos no proporcionan suficiente evidencia para descartarla.

Ejemplo:

Supongamos que estamos realizando una prueba de hipótesis con los siguientes parámetros:

- Nivel de significancia (α): 0.05,
- Valor-p obtenido: 0.03.

Decisión:

Dado que valor-p=0.03<α=0.05, se rechaza H0.

Conclusión: Los resultados son estadísticamente significativos, lo que sugiere que el efecto observado en los datos es improbable que se deba al azar.

Relación entre Significancia Estadística y Valor-p

La significancia estadística y el valor-p están estrechamente relacionados. Ambos permiten evaluar si los resultados de un estudio son consistentes con la hipótesis nula o si, por el contrario, indican la presencia de un efecto significativo. Sin embargo, es importante recordar que:

1. Significancia estadística ≠ Importancia práctica:

   Un resultado puede ser estadísticamente significativo, pero tener un impacto práctico irrelevante, especialmente en estudios con muestras grandes.

2. Dependencia del nivel de significancia (α):
   La decisión de rechazar o no H0 depende del umbral elegido para α. Un valor-p cercano a α puede llevar a conclusiones diferentes según el nivel de significancia seleccionado.
3. Contexto importa:
   El valor-p debe interpretarse siempre en el contexto del problema y considerando otros factores, como el diseño del estudio, el tamaño de la muestra y las limitaciones de los datos.

Ejemplo Práctico en Python

A continuación, se muestra cómo calcular el valor-p y tomar una decisión basada en el nivel de significancia en Python:

```python
from scipy.stats import norm

# Parámetros
mu_0 = 50       # Media bajo la hipótesis nula
sigma = 10      # Desviación estándar poblacional
n = 36          # Tamaño de la muestra
X_bar = 52      # Media muestral
alpha = 0.05    # Nivel de significancia

# Paso 1: Calcular el estadístico Z
Z = (X_bar - mu_0) / (sigma / np.sqrt(n))

# Paso 2: Calcular el valor-p (prueba bilateral)
p_valor = 2 * (1 - norm.cdf(abs(Z)))

# Resultados
print(f"Estadístico Z: {Z:.2f}")
print(f"Valor-p: {p_valor:.4f}")

# Decisión
if p_valor < alpha:
    print("Se rechaza H0. Los resultados son estadísticamente significativos.")
else:
    print("No se rechaza H0. No hay suficiente evidencia para concluir que los resultados son significativos.")
```

Ilustración 6 Implementación en Python

Salida esperada

```
Estadístico Z: 1.20
Valor-p: 0.2302
No se rechaza H0. No hay suficiente evidencia para concluir que los resultados son significativos.
```

Ilustración 7 Salida esperada

PRUEBAS SOBRE UNA PROPORCIÓN DE LA POBLACIÓN

Las pruebas sobre proporciones evalúan si la proporción poblacional (p) es igual a un valor específico. Estas pruebas son útiles en situaciones donde los datos son categóricos, como encuestas de opinión o estudios de éxito/fracaso.

Fórmula:

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Ilustración 8 Fórmula prueba de proporciones

Donde:

$\hat{p}$: Proporción muestral.

p0: Proporción bajo la hipótesis nula (H0).

Ejemplo:

Probar si la proporción de votantes a favor de un candidato es 0.5 (H0: p=0.5) con los siguientes datos:

- $\hat{p}$ = 0.55 (proporción muestral),
- n=100 (tamaño de la muestra).

El cálculo del estadístico Z sería:

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.55 - 0.50}{\sqrt{\frac{0.5 \cdot 0.5}{100}}} = \frac{0.05}{\sqrt{0.0025}} = \frac{0.05}{0.05} = 1.0$$

Ilustración 9 Calculo estadístico Z

Valor-p:

Para una prueba bilateral, el valor-p es el área bajo la curva normal estándar más allá de $|Z|=1.0$. Usando tablas o software:

$$p = 2 \cdot P(Z > 1.0) \approx 2 \cdot 0.1587 = 0.3174$$

Ilustración 10 Valor p

Decisión:

Si $\alpha=0.05$, dado que $p=0.3174>\alpha$, no se rechaza H0.

Conclusión:

No hay suficiente evidencia para concluir que la proporción de votantes a favor del candidato sea diferente de 0.5.

Implementación en Python

Aquí tienes un ejemplo en Python para realizar esta prueba:

```python
from scipy.stats import norm

# Parámetros
p_hat = 0.55    # Proporción muestral
p_0 = 0.50      # Proporción bajo H0
n = 100         # Tamaño de la muestra
alpha = 0.05    # Nivel de significancia

# Paso 1: Calcular el estadístico Z
Z = (p_hat - p_0) / np.sqrt((p_0 * (1 - p_0)) / n)

# Paso 2: Calcular el valor-p (prueba bilateral)
p_valor = 2 * (1 - norm.cdf(abs(Z)))

# Resultados
print(f"Estadístico Z: {Z:.2f}")
print(f"Valor-p: {p_valor:.4f}")

# Decisión
if p_valor < alpha:
    print("Se rechaza H0. La proporción es significativamente diferente de 0.5.")
else:
    print("No se rechaza H0. No hay suficiente evidencia para concluir que la proporción es diferente de 0.5.")
```

Ilustración 11 Implementación en Python

Salida esperada:

```
Estadístico Z: 1.00
Valor-p: 0.3173
No se rechaza H0. No hay suficiente evidencia para concluir que la proporción es diferente de 0.5.
```

Ilustración 12 Salida esperada

PRUEBAS SOBRE UNA MEDIA POBLACIONAL

Las pruebas sobre medias evalúan si la media poblacional (μ) es igual a un valor específico. Dependiendo de si la desviación estándar poblacional (σ) es conocida o desconocida, se utiliza la distribución normal estándar (Z) o la distribución t de Student (t).

Fórmulas

1. Desviación Estándar Conocida:

$$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

Ilustración 13 Fórmula desviación estándar conocida

2. Desviación Estándar Desconocida:

$$t = \frac{X - \mu_0}{\frac{s}{\sqrt{n}}}$$

Ilustración 14 Fórmula desviación estándar desconocida

Donde:

- $\hat{X}$: Media muestral.
- M0: Media bajo la hipótesis nula (H0).
- σ: Desviación estándar poblacional (si es conocida).
- s: Desviación estándar muestral (si es desconocida).
- n: Tamaño de la muestra.

Ejemplo:

Probar si la media de una población es 50 (H0:μ=50) con los siguientes datos:

- $\hat{X}$ = 52 (media muestral),
- s=10 (desviación estándar muestral),
- n=36 (tamaño de la muestra).

El cálculo del estadístico t sería:

$$t = \frac{X - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{52 - 50}{\frac{10}{\sqrt{36}}} = \frac{2}{\frac{10}{6}} = \frac{2}{1.67} \approx 1.2$$

Ilustración 15 Calculo estadístico t

Valor-p:

Usando tablas t o software, con $n-1=35$ grados de libertad y una prueba bilateral:

$$p \approx 0.238$$

Ilustración 16 Valor-p

Decisión:

Si $\alpha=0.05$, dado que $p=0.238>\alpha$, no se rechaza H0.

Conclusión:

No hay suficiente evidencia para concluir que la media poblacional sea diferente de 50.

Implementación en Python

Aquí un ejemplo en Python para realizar esta prueba:

```python
from scipy.stats import t

# Parámetros
X_bar = 52      # Media muestral
mu_0 = 50       # Media bajo H0
s = 10          # Desviación estándar muestral
n = 36          # Tamaño de la muestra
alpha = 0.05    # Nivel de significancia

# Paso 1: Calcular el estadístico t
grados_libertad = n - 1
t_estadistico = (X_bar - mu_0) / (s / np.sqrt(n))

# Paso 2: Calcular el valor-p (prueba bilateral)
p_valor = 2 * (1 - t.cdf(abs(t_estadistico), df=grados_libertad))

# Resultados
print(f"Estadístico t: {t_estadistico:.2f}")
print(f"Valor-p: {p_valor:.4f}")

# Decisión
if p_valor < alpha:
    print("Se rechaza H0. La media es significativamente diferente de 50.")
else:
    print("No se rechaza H0. No hay suficiente evidencia para concluir que la media es diferente de 50.")
```

Ilustración 17 Implementación en Python

Salida esperada:

```
Estadístico t: 1.20
Valor-p: 0.2380
No se rechaza H0. No hay suficiente evidencia para concluir que la media es diferente de 50.
```

Ilustración 18 Salida esperada

ERRORES TIPO I Y TIPO II

En el contexto de las pruebas de hipótesis, es fundamental entender los posibles errores que pueden ocurrir al tomar una decisión basada en los datos muestrales. Estos errores se clasifican en Error Tipo I y Error Tipo II, y cada uno tiene implicaciones importantes en la interpretación de los resultados.

Error Tipo I:

- Definición: Ocurre cuando rechazamos la hipótesis nula ($H0$) siendo esta verdadera.
- Probabilidad: La probabilidad de cometer un Error Tipo I está dada por el nivel de significancia (α).
- Ejemplo práctico:
  Concluir que un medicamento es efectivo cuando en realidad no lo es. Esto podría llevar a la aprobación de un tratamiento ineficaz, con riesgos para los pacientes y costos innecesarios.

Error Tipo II:

- Definición: Ocurre cuando no rechazamos la hipótesis nula ($H0$) siendo esta falsa.
- Probabilidad: La probabilidad de cometer un Error Tipo II se denota como (β).
- Potencia de la prueba: La potencia de una prueba estadística es la probabilidad de rechazar $H0$ cuando es falsa, y se calcula como $1-\beta$. Una prueba con alta potencia tiene menor probabilidad de cometer un Error Tipo II.
- Ejemplo práctico:
  Concluir que un medicamento no es efectivo cuando en realidad sí lo es. Esto podría resultar en la pérdida de una oportunidad para implementar un tratamiento valioso.

Relación entre Errores Tipo I y Tipo II:

- Compromiso entre α y β: Reducir la probabilidad de un Error Tipo I (α) aumenta la probabilidad de un Error Tipo II (β), y viceversa. Por ello, es importante elegir un nivel de significancia adecuado según el contexto del problema.

- Tamaño de muestra (n): Aumentar el tamaño de la muestra reduce tanto α como β, mejorando la precisión de la prueba.

PRUEBAS DE HIPÓTESIS

Las pruebas de hipótesis pueden clasificarse según la dirección de la hipótesis alternativa (H1) en unilaterales o bilaterales. Esta clasificación influye en cómo se interpreta el estadístico de prueba y cómo se define la región crítica.

Pruebas Unilaterales:

La hipótesis alternativa (H1) especifica una dirección (mayor o menor). Esto implica que solo estamos interesados en detectar desviaciones en una dirección específica.

Ejemplo:

- $H0:\mu=50$
- $H1:\mu>50$ (la media poblacional es mayor que 50).
- En este caso, la región crítica se encuentra en una sola cola de la distribución (cola derecha si $H1:\mu>50$, o cola izquierda si $H1:\mu<50$).

Pruebas Bilaterales:

La hipótesis alternativa (H1) no especifica una dirección, sino que busca detectar cualquier desviación significativa de la hipótesis nula (H0).

Ejemplo:

- $H0:\mu=50$
- $H1:\mu \neq 50$ (la media poblacional es diferente de 50).
- En este caso, la región crítica se divide en ambas colas de la distribución.

Ejemplo Práctico

Supongamos que estamos evaluando la efectividad de un nuevo método de enseñanza en comparación con el método tradicional. Queremos probar si el puntaje promedio de los estudiantes (μ) mejora significativamente con el nuevo método.

Hipótesis:

- $H0: \mu=70$ (el puntaje promedio no cambia con el nuevo método).
- $H1: \mu>70$ (el puntaje promedio mejora con el nuevo método).

Esta es una prueba unilateral, ya que estamos interesados únicamente en detectar si el puntaje promedio es mayor que 70.

Si, en cambio, quisiéramos probar si el puntaje promedio cambia (ya sea aumentando o disminuyendo), las hipótesis serían:

- $H0:\mu = 70$
- $H1:\mu \neq 70$.

Esta sería una prueba bilateral, ya que estamos interesados en cualquier cambio significativo en el puntaje promedio.

Implementación en Python

Aquí un ejemplo en Python para realizar una prueba unilateral y una prueba bilateral:

```python
from scipy.stats import norm

# Parámetros
mu_0 = 70        # Media bajo H0
sigma = 10       # Desviación estándar poblacional
n = 36           # Tamaño de la muestra
X_bar = 72       # Media muestral
alpha = 0.05     # Nivel de significancia

# Paso 1: Calcular el estadístico Z
Z = (X_bar - mu_0) / (sigma / np.sqrt(n))

# Paso 2: Calcular el valor-p
# Prueba unilateral (H1: μ > 70)
p_valor_unilateral = 1 - norm.cdf(Z)

# Prueba bilateral (H1: μ ≠ 70)
p_valor_bilateral = 2 * (1 - norm.cdf(abs(Z)))

# Resultados
print(f"Estadístico Z: {Z:.2f}")
print(f"Valor-p (unilateral): {p_valor_unilateral:.4f}")
print(f"Valor-p (bilateral): {p_valor_bilateral:.4f}")

# Decisión
if p_valor_unilateral < alpha:
    print("Unilateral: Se rechaza H0.")
else:
    print("Unilateral: No se rechaza H0.")

if p_valor_bilateral < alpha:
    print("Bilateral: Se rechaza H0.")
else:
    print("Bilateral: No se rechaza H0.")
```

Ilustración 19 Implementación en Python

Salida esperada:

```
Estadístico Z: 1.20
Valor-p (unilateral): 0.1151
Valor-p (bilateral): 0.2302
Unilateral: No se rechaza H0.
Bilateral: No se rechaza H0.
```

Ilustración 20 Salida esperada

DISEÑO DE EXPERIMENTO

El diseño de experimentos es el proceso sistemático de planificar un experimento para obtener datos válidos y confiables que permitan responder preguntas específicas. Un buen diseño garantiza que los resultados sean interpretables, reproducibles y generalizables.

Componentes del Diseño de Experimentos:

1. Definir el objetivo: Especificar claramente la pregunta de investigación o el problema que se desea resolver. Por ejemplo, "¿Es efectivo un nuevo fertilizante en comparación con el fertilizante estándar?"
2. Seleccionar las variables:
   - Variable independiente (factor): Lo que se manipula o controla en el experimento (por ejemplo, el tipo de fertilizante).
   - Variable dependiente (respuesta): Lo que se mide como resultado (por ejemplo, el rendimiento del cultivo).
   - Variables de control: Factores que deben mantenerse constantes para evitar sesgos (por ejemplo, cantidad de agua, tipo de suelo).
3. Determinar el tamaño de la muestra: El tamaño de la muestra debe ser suficiente para detectar diferencias significativas si existen, minimizando errores estadísticos. Esto se relaciona directamente con el poder estadístico del experimento.

4. Asignar tratamientos: Distribuir aleatoriamente los tratamientos (por ejemplo, fertilizante nuevo vs. fertilizante estándar) entre las unidades experimentales (parcelas de cultivo) para reducir el impacto de factores externos.

Ejemplo Práctico:

Diseñar un experimento para probar la efectividad de un nuevo fertilizante:

- Objetivo: Determinar si el nuevo fertilizante mejora el rendimiento promedio de los cultivos en comparación con el fertilizante estándar.
- Variables:
  - Independiente: Tipo de fertilizante (nuevo vs. estándar).
  - Dependiente: Rendimiento del cultivo (kg/ha).
  - Controladas: Cantidad de agua, tipo de suelo, exposición al sol.
- Tamaño de la muestra: Seleccionar un número adecuado de parcelas para asegurar un poder estadístico suficiente.
- Asignación de tratamientos: Asignar aleatoriamente el fertilizante nuevo o estándar a cada parcela.

PODER DE EXPERIMENTO

El poder estadístico es la probabilidad de rechazar correctamente la hipótesis nula (H0) cuando esta es falsa. Un experimento con alto poder tiene mayor capacidad para detectar un efecto real si existe.

Fórmula:

$$\text{Poder} = 1 - \beta$$

Ilustración 21 Fórmula poder de experimento

Donde:

β: Probabilidad de cometer un Error Tipo II (no rechazar H0 cuando es falsa).

Factores que Afectan el Poder:

1. Tamaño de la muestra (n): Aumentar el tamaño de la muestra incrementa el poder del experimento, ya que reduce la variabilidad y mejora la precisión de las estimaciones.
2. Nivel de significancia (α): Un nivel de significancia más alto (α) aumenta el poder, pero también incrementa el riesgo de cometer un Error Tipo I.
3. Tamaño del efecto: Un efecto más grande (diferencia entre grupos o condiciones) es más fácil de detectar, lo que aumenta el poder del experimento.
4. Variabilidad de los datos: Menor variabilidad en los datos mejora el poder, ya que facilita identificar diferencias significativas.

Ejemplo Práctico:

Supongamos que queremos mejorar el poder de un experimento para probar la efectividad de un nuevo fertilizante:

- Si el tamaño de la muestra inicial es pequeño (por ejemplo, 20 parcelas), podríamos aumentarlo a 50 parcelas para mejorar el poder.
- Además, podríamos reducir la variabilidad controlando mejores factores externos, como el riego o la exposición al sol.

Implementación en Python

```python
from statsmodels.stats.power import TTestIndPower
import numpy as np

# Parámetros iniciales del experimento
efecto = 0.5  # Tamaño del efecto (diferencia estandarizada entre grupos)
alpha = 0.05  # Nivel de significancia
n_inicial = 20  # Tamaño inicial de la muestra por grupo
n_mejorado = 50  # Tamaño mejorado de la muestra por grupo
variabilidad_inicial = 1.0  # Variabilidad inicial (desviación estándar)
variabilidad_mejorada = 0.7  # Variabilidad reducida (mejora en control)

# Función para calcular el poder estadístico
def calcular_poder(n, efecto, alpha, variabilidad):
    # Calculamos el poder usando la clase TTestIndPower
    analisis = TTestIndPower()
    poder = analisis.power(
        effect_size=efecto / variabilidad,  # Tamaño del efecto ajustado por variabilidad
        nobs1=n,  # Tamaño de la muestra del primer grupo
        alpha=alpha,  # Nivel de significancia
        ratio=1.0  # Proporción entre los tamaños de los grupos (1:1)
    )
    return poder

# Calcular el poder para el diseño inicial
poder_inicial = calcular_poder(n_inicial, efecto, alpha, variabilidad_inicial)

# Calcular el poder para el diseño mejorado con mayor tamaño de muestra
poder_mejorado_n = calcular_poder(n_mejorado, efecto, alpha, variabilidad_inicial)

# Calcular el poder para el diseño mejorado con menor variabilidad
poder_mejorado_var = calcular_poder(n_inicial, efecto, alpha, variabilidad_mejorada)

# Resultados
print(f"Poder inicial (n={n_inicial}, variabilidad={variabilidad_inicial}): {poder_inicial:.2f}")
print(f"Poder mejorado (n={n_mejorado}, variabilidad={variabilidad_inicial}): {poder_mejorado_n:.2f}")
print(f"Poder mejorado (n={n_inicial}, variabilidad={variabilidad_mejorada}): {poder_mejorado_var:.2f}")
```

Ilustración 22 Implementación en Python

Explicación del Código

1. Parámetros Iniciales:
   - efecto: Representa la diferencia estandarizada entre los grupos (por ejemplo, la diferencia en rendimiento entre el fertilizante nuevo y el estándar).
   - alpha: Nivel de significancia (α), que define la probabilidad de cometer un Error Tipo I.
   - n_inicial y n_mejorado: Tamaños de muestra inicial y mejorado, respectivamente.
   - variabilidad_inicial y variabilidad_mejorada: Desviaciones estándar que representan la variabilidad en los datos antes y después de mejorar el control de factores externos.
2. Función calcular_poder:
   - Utiliza la clase TTestIndPower de statsmodels para calcular el poder estadístico de una prueba t de dos muestras independientes.
   - El tamaño del efecto se ajusta dividiendo el efecto absoluto por la variabilidad.
3. Cálculos:
   - Calculamos el poder inicial con un tamaño de muestra pequeño y alta variabilidad.
   - Mejoramos el poder aumentando el tamaño de la muestra.
   - Mejoramos el poder reduciendo la variabilidad.

Salida Esperada

Supongamos que ejecutamos el código con los parámetros proporcionados:

```
Poder inicial (n=20, variabilidad=1.0): 0.34
Poder mejorado (n=50, variabilidad=1.0): 0.70
Poder mejorado (n=20, variabilidad=0.7): 0.61
```

Ilustración 23 Salida esperada

Interpretación:

- Diseño inicial: Con n=20 y alta variabilidad, el poder es bajo (0.34), lo que significa que hay solo un 34% de probabilidad de detectar un efecto real si existe.

- Mejora en tamaño de muestra: Al aumentar n a 50, el poder aumenta a 0.70, lo que mejora significativamente la capacidad de detectar diferencias.
- Mejora en variabilidad: Al reducir la variabilidad (controlando mejor factores externos), el poder aumenta a 0.61, incluso con el tamaño de muestra inicial.

Los conceptos de diseño de experimentos y poder estadístico son fundamentales para garantizar que los estudios científicos y las pruebas de hipótesis sean válidas y confiables. Un diseño cuidadoso permite obtener datos útiles, mientras que un poder estadístico adecuado asegura que el experimento tenga la capacidad de detectar efectos reales. Estos principios son ampliamente aplicables en ciencia de datos, investigación científica, control de calidad y muchas otras áreas.