

# Nível de Significância e Erro Tipo I

- Objetivo: Demonstrar computacionalmente o significado do nível de significância ( $\alpha$ ) em testes estatísticos.
- Metodologia: Simulação de 10.000 testes-t independentes (Testes A/B) utilizando Python.
- Cenário: Comparação entre dois grupos extraídos de populações idênticas ( $H_0$  é verdadeira).

```
import numpy as np
from scipy.stats import ttest_ind

num_simulations = 10000
sample_size = 30
true_mean = 50
true_std = 10
significance_level = #Usar valores diferentes

significant_results = 0

for _ in range(num_simulations):

    group1 = np.random.normal(true_mean, true_std, sample_size)
    group2 = np.random.normal(true_mean, true_std, sample_size)

    t_stat, p_value = ttest_ind(group1, group2)

    if p_value < significance_level:
        significant_results += 1

proportion_significant = significant_results / num_simulations

print(f"Number of simulations: {num_simulations}")
print(f"Significance level (alpha): {significance_level}")
print(f"Proportion of significant results (Type I error rate): {proportion_significant:.4f}")
```

# A Construção da Simulação (Cenário H<sub>0</sub> Verdadeira)

```
• # Parâmetros fixos: Mesma média e desvio padrão para ambos  
• true_mean = 50  
• true_std = 10  
•  
• # Loop de Simulação (Repetido 10.000 vezes)  
• group1 = np.random.normal(true_mean, true_std, sample_size)  
• group2 = np.random.normal(true_mean, true_std, sample_size)  
•  
• # Teste T  
t_stat, p_value = ttest_ind(group1, group2)
```

Nesta etapa, o código constrói intencionalmente um cenário onde não existe diferença real entre os grupos:

1. As variáveis group1 e group2 são geradas usando exatamente a mesma média (true\_mean = 50) e o mesmo desvio padrão (true\_std = 10).
2. Isso significa que a Hipótese Nula H<sub>0</sub> é verdadeira por definição.
3. Qualquer teste estatístico que aponte uma diferença significativa entre esses grupos estará cometendo, necessariamente, um erro.
4. O código repete esse processo 10.000 vezes para medir a frequência com que esse erro ocorre aleatoriamente.

# Impacto do Nível de Significância ( $\alpha$ ) nos Resultados

Parâmetro Definido ( <code>significance_level</code> )	Resultado da Simulação ( <code>proportion_significant</code> )	Interpretação
<b>0.05 (5%)</b>	~ 0.0502 (aprox. 5%)	Taxa de Falsos Positivos Padrão
<b>0.01 (1%)</b>	~ 0.0104 (aprox. 1%)	Teste mais rigoroso, menos erros

Ao executar o código com dois valores diferentes para o `significance_level`, observamos que:

- Cenário 1 ( $\alpha = 0.05$ ): O código detectou "diferença significativa" em aproximadamente 5% das 10.000 simulações, mesmo os grupos sendo idênticos.
- Cenário 2 ( $\alpha = 0.01$ ): Ao tornarmos o critério mais rigoroso, a taxa de detecção caiu para aproximadamente 1%.
- Conclusão empírica: A proporção de resultados significativos convergiu quase perfeitamente para o valor de  $\alpha$  escolhido no código.

# O que é o Erro Tipo I?

```
if p_value < significance_level:  
    significant_results += 1  
# Proportion reflects Type I Error rate
```

O experimento comprova a definição estatística de Erro Tipo I:

- Definição: O Erro Tipo I ocorre quando rejeitamos a Hipótese Nula ( $H_0$ ), mas ela é verdadeira (um "Falso Positivo").
- O papel do Código: O código demonstrou que, ao definirmos `significance_level = 0.05`, estamos aceitando explicitamente que, se repetirmos o experimento muitas vezes, erraremos em 5% dos casos.
- Interpretação Final: O p-value baixo nessas simulações não indicou uma descoberta real, mas sim uma flutuação estatística natural (ruído) que cruzou o limiar de significância estabelecido. Portanto, o nível de significância é a nossa "taxa de tolerância" para cometer Erros do Tipo I.