

# Análise Inferencial de Dados – Entrega 4

## Método:

Essa entrega é uma continuação das últimas 3 entregas, então o mesmo código usado para realizar as contas necessárias para o primeiro documento, também serão utilizadas aqui.

Os números de exemplo seguem o mesmo, e pegamos o valor do fechamento das ações da Apple entre Jan/2022 e Dez/2022.

As entregas e o código da análise na integra podem ser encontrados nestes links:

- [Entrega 1](#)
- [Entrega 2](#)
- [Entrega 3](#)
- [Análise na integra](#)

## Definir hipótese nula e hipótese alternativa:

Definimos nossa hipótese nula ( $H_0$ ) e nossa hipótese alternativa ( $H_1$ ) dessa forma:

$H_0$ : A proporção populacional é 50%

$H_1$ : A proporção populacional é diferente de  $p_0$

```
p0 = 0.5
```

## Nível de Significância: 0.05

O nível de significância estabelecido é de 5% (0,05). Este é o limiar abaixo do qual rejeitaríamos a hipótese nula.

```
nivel_significancia = 0.05
```

## Número de Observações (n): 251

Foram analisadas 251 observações, correspondentes aos dias de negociação no período em questão.

O valor foi obtido a partir do seguinte método:

```
n = len(df)
```

## Número de Sucessos: 117

Houve 117 dias em que o preço de fechamento aumentou em relação ao dia anterior, considerados como "sucessos" para o teste de hipótese.

O valor foi obtido a partir do seguinte método:

```
numero_sucessos = df['Sucesso'].sum()
```

### Proporção Amostral ( $\hat{p}$ ): 0.4661

A proporção amostral é de 0,4661 ou 46,61%, indicando que em 46,61% dos dias houve um aumento no preço de fechamento. Isso sugere que os aumentos de preço ocorreram em menos da metade dos dias analisados.

O valor foi obtido a partir do seguinte método:

```
proporcao_amostrual = numero_sucessos / n
```

### Erro Padrão da Proporção: 0.0316

O erro padrão da proporção é 0,0316, refletindo a variabilidade esperada da proporção amostral em relação à proporção populacional sob a hipótese nula. Um erro padrão menor indica estimativas mais precisas da proporção.

O valor foi obtido a partir do seguinte método:

```
erro_padrao_proporcao = np.sqrt((p0 * (1 - p0)) / n)
```

### Estatística z: -1.0730

A estatística z de -1,0730 quantifica o número de desvios padrão que a proporção amostral está distante da proporção hipotética sob a hipótese nula (50%). Um valor negativo indica que a proporção amostral é menor que a proporção esperada.

O valor foi obtido a partir do seguinte método:

```
z_estatistica = (proporcao_amostrual - p0) / erro_padrao_proporcao
```

### Valor-p: 0.2833

O valor-p de 0,2833 representa a probabilidade de obter uma proporção amostral tão extrema quanto 46,61%, assumindo que a verdadeira proporção populacional é 50%. Um valor-p alto sugere que a diferença observada pode ser atribuída ao acaso.

O valor foi obtido a partir do seguinte método:

```
valor_p = 2 * (1 - stats.norm.cdf(abs(z_estatistica)))
```

### Conclusão: Não rejeitamos a hipótese nula ( $H_0$ ).

Como o valor-p (0,2833) é maior que o nível de significância (0,05), não rejeitamos a hipótese nula. Isso significa que não há evidências estatisticamente significativas para concluir que a proporção de dias em que o preço de fechamento da Apple Inc. aumenta é diferente de 50%.

Em outras palavras, os dados não fornecem suporte suficiente para afirmar que a ação tende a subir ou descer em uma frequência diferente do que seria esperado por acaso.

Para validar nossa conclusão, usamos o seguinte método:

```
rejeitar_H0 = valor_p < nivel_significancia
```