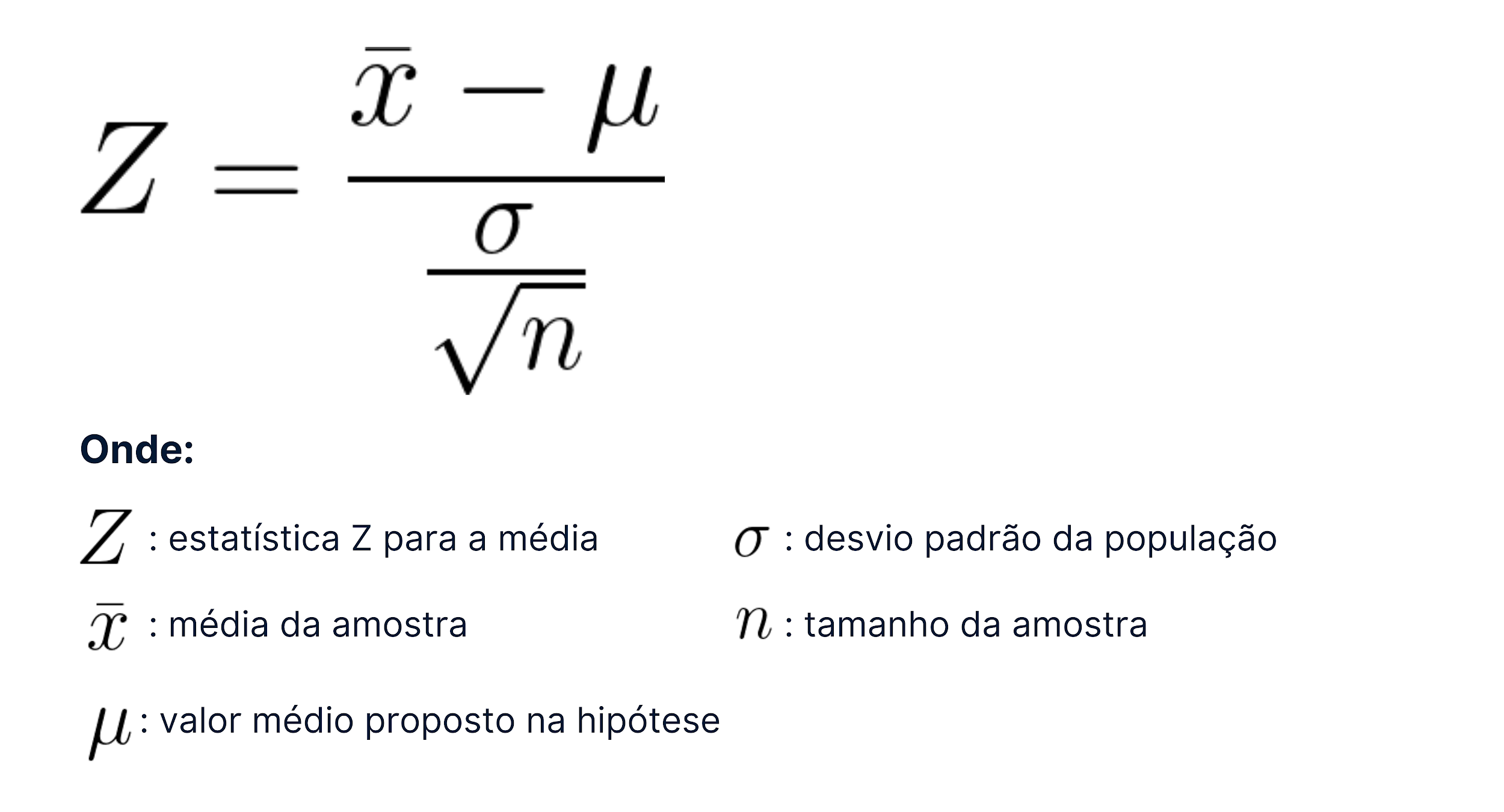
Durante as aulas utilizamos o Teste Z para fazer inferências sobre a média do nosso problema. Vamos expandir ainda mais nossos estudos sobre esse teste?

O teste Z é usado para testar hipóteses quando você tem condições específicas que justificam seu uso. Ele é especialmente aplicável em situações onde:

1. **Tamanho da amostra é grande**: Geralmente, o teste Z é usado quando o tamanho da amostra (n) é maior que 30. Isso se deve ao Teorema Central do Limite, que afirma que, para tamanhos de amostra suficientemente grandes, a distribuição das médias amostrais tende a ser normal, independentemente da forma da distribuição populacional.
2. **Desvio padrão populacional é conhecido**: Para aplicar o teste Z, você precisa conhecer o desvio padrão da população (σ). Isso raramente acontece na prática, mas quando conhecido, permite o uso do teste Z para avaliar diferenças entre médias ou proporções.
3. **A amostra é aleatória e as observações são independentes**: Isso garante que os dados não sejam enviesados e que o modelo estatístico aplicado seja válido.

Em um contexto geral, a fórmula matemática do teste Z é igual à diferença entre o valor amostral calculado e o valor populacional proposto dividido pelo desvio padrão do parâmetro populacional. Mesmo tendo uma fórmula geral, o cálculo matemático pode ser diferente dependendo do parâmetro que está sendo testado.

Por exemplo, a fórmula do **Teste Z para a média**, usada para testar hipóteses sobre a média populacional, é definida pela diferença entre a média da amostra e a média proposta na hipótese, dividido pelo resultado da divisão entre o desvio padrão da população e a raiz quadrada do tamanho da amostra, como demonstrado na imagem abaixo.



De modo amplo, todas as fórmulas vão nos fornecer a informação que precisamos para rejeitar ou não rejeitar a hipótese nula de uma situação. Para que essa informação seja efetiva na nossa decisão, precisamos saber como interpretá-la. Buscando tornar a aprendizagem mais palpável, vamos seguir com um exemplo.

Imaginemos que estamos interessados em entender se os resultados de um experimento ou amostra são consistentes com o que esperávamos da população maior. Para isso, formulamos duas hipóteses:

* H0: Não há diferença significativa entre a amostra e a população.
* H1: Existe uma diferença significativa.

Coletamos os dados de uma amostra (por exemplo, médias, proporções) e queremos fazer afirmações sobre a população maior com base nessa amostra. Aplicamos o teste Z para investigar se a diferença observada entre a amostra e a população é grande o suficiente para ser considerada significativa. Do Teste Z, coletamos dois valores: Estatística Z, conhecido também como Z-score, e o **p-valor**.

O p-valor, como já aprendemos em aula, representa a probabilidade de obter os resultados observados se a hipótese nula fosse verdadeira. Em termos mais simples, é a medida de quão extremos são os resultados em relação ao que seria esperado, assumindo que não há efeito real. Ao utilizar o p-valor para tomar decisões estatísticas, comparamos esse valor com um nível de significância previamente escolhido, geralmente representado por alpha - α. Eis como interpretamos e agimos com base nessa comparação:

Para tomar uma decisão a partir do p-valor, comparamos esse valor com um nível de significância previamente escolhido:

* Se o **p-valor é menor que o nível de significância**, isso sugere que os resultados observados são altamente improváveis de ocorrer sob a premissa da hipótese nula. Em resposta a isso, **rejeitamos a hipótese nula** em favor da hipótese alternativa. Pois indica que há evidências estatísticas suficientes para suportar a hipótese alternativa.
* Se o **p-valor é maior que o nível de significância**, isso implica que os resultados observados são plausíveis mesmo considerando a hipótese nula. Em tal caso, **não rejeitamos a hipótese nula**. Isso indica uma falta de evidência estatística para sustentar a hipótese alternativa.

O valor da Estatística Z fornece uma medida da distância em termos de desvios padrão entre a média da amostra e a média da população. Essa medida é essencial para tomar decisões fundamentadas sobre a hipótese nula em um teste de hipóteses.

Ao realizar um teste de hipóteses, comumente comparamos a Estatística Z com um valor crítico. O valor crítico é um ponto específico na distribuição normal padrão que determina a região de rejeição da hipótese nula. Sua escolha está relacionada ao nível de significância.

A interpretação do valor crítico varia de acordo com a **natureza** do teste de hipóteses:

* **Teste Bilateral:** em um teste de hipóteses bilateral, usado quando estamos interessados em detectar qualquer diferença significativa na média, a hipótese nula é rejeitada se o valor absoluto da Estatística Z for maior que o valor crítico. Isso implica que a média da amostra é significativamente diferente da média da população, seja para mais ou para menos.
* **Teste Unilateral à direita:** em um teste de hipóteses unilateral à direita, usado quando estamos interessados em verificar se a média da amostra é significativamente maior que a média da população, a hipótese nula é rejeitada se a Estatística Z for maior que o valor crítico. Isso sugere uma diferença significativa na direção desejada.
* **Teste Unilateral à esquerda:** em um teste de hipóteses unilateral à esquerda, usado quando estamos interessados em verificar se a média da amostra é significativamente menor que a média da população, a hipótese nula é rejeitada se a Estatística Z for menor que o valor crítico. Indica uma diferença significativa na direção desejada.