

Profissão: Cientista de Dados



GLOSSÁRIO



Inferência



Dica: para encontrar rapidamente a palavra que procura aperte o comando CTRL+F e digite o termo que deseja achar.

- **Lógica indutiva vs dedutiva**
- **O problema da inferência**
- **Teste de significância**
- **Teste de hipóteses no Python**
- **Teste T Pareado**



Lógica indutiva vs dedutiva



Lógica indutiva vs dedutiva

● Inferência Estatística

É o processo de extrapolar conclusões de uma amostra para uma população, com um certo nível de confiança. É um elemento crucial na ciência de dados, pois permite o uso de ferramentas estatísticas para fazer previsões e tirar conclusões.

● Pensamento Dedutivo

Segue uma lógica matemática, onde se parte de uma premissa geral para chegar a uma conclusão específica.

● Pensamento Indutivo

Parte de um fato particular para chegar a uma conclusão geral.



O problema da inferência



O problema da inferência

• Dados estocásticos

São dados gerados ao longo do tempo, muitas vezes de forma aleatória.

• Inferência estatística

É um método usado para estimar características de uma população com base em uma amostra dessa população.



Teste de significância



Teste de significância

• Distribuição t

É uma distribuição de probabilidade contínua que é usada quando a variância da população não é conhecida. Ela tem uma média de zero e um desvio padrão de um, mas tem um parâmetro adicional chamado graus de liberdade.

• Teste t

É um tipo de teste de hipóteses que é usado quando a variância da população não é conhecida. Ele é baseado na distribuição t.

• Graus de Liberdade

É um parâmetro da distribuição t que indica o número de valores independentes que podem variar em uma análise estatística.

• Estatística de Teste

É uma função dos dados que tem uma distribuição conhecida. É usada para determinar se a hipótese nula deve ser rejeitada ou não.



Teste T Pareado



Teste T Pareado

• Teste-T Pareado

É um teste estatístico usado para comparar duas amostras pareadas para determinar se elas são significativamente diferentes. Nesta aula, é usado para avaliar as diferenças entre a sola nova e a tradicional.



Poder do teste



Poder do teste

● **Análise do poder do teste**

Processo de avaliar a eficácia de um teste estatístico em termos de sua capacidade de evitar erros do tipo II. Um teste mais poderoso tem uma área menor sob a curva do erro tipo II.

● **Erro tipo I**

Probabilidade de rejeitar a hipótese nula (H_0) quando ela é verdadeira.

● **Diferença significativa**

Conceito subjetivo que depende do conhecimento do negócio. Refere-se a se a diferença identificada pelo teste de hipóteses tem um impacto significativo no contexto do negócio.

● **Erro tipo II**

Probabilidade de não rejeitar H_0 quando uma hipótese alternativa é verdadeira. Para calcular o erro tipo II, é necessário especificar a hipótese alternativa de forma bem específica.



Poder do teste

• Hipótese alternativa

Proposição que é aceita se a hipótese nula for rejeitada. Geralmente baseada no conceito de uma diferença significativa.

• Poder do teste

Definido como um menos o erro do tipo II, ou seja, é o complementar do erro do tipo II. Quando se diz que um teste é mais poderoso, significa que a probabilidade de não rejeitar H_0 , dada uma hipótese alternativa, é pequena.

• Hipótese nula (H_0)

Proposição que é testada para rejeição em um teste de hipóteses.



Bons estudos!

