

Profissão: Cientista de Dados



GLOSSÁRIO



Probabilidade



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

- **Conheça modelos discretos**
- **Faça inferência sobre os parâmetros**
- **Diferencie qualidade de complexidade**
- **Selecione modelos**
- **Regularização**



Conheça modelos discretos



Conheça modelos discretos

● Bernoulli (Bern)

É um modelo probabilístico que descreve um evento que pode ter sucesso ou fracasso. Por exemplo, jogar uma moeda para cima, onde cara é sucesso e coroa é fracasso, cada um com uma probabilidade de 50%.

● Binomial (Bial)

É uma extensão do modelo de Bernoulli que descreve a probabilidade de obter um certo número de sucessos em um número fixo de tentativas independentes. Por exemplo, retirar uma bola de uma urna, anotar sua cor, e então devolver a bola à urna. Este processo é repetido e o número de sucessos é contado.



Conheça modelos discretos

• Distribuição Uniforme

É uma distribuição de probabilidade onde todos os resultados são igualmente prováveis. É usada para gerar variáveis aleatórias em simulações.

• Distribuição Normal

É uma distribuição de probabilidade que descreve dados que se distribuem de maneira simétrica em torno de uma média. A distribuição das variáveis em uma simulação de lançamento de moeda se aproxima de uma distribuição normal à medida que o número de lançamentos aumenta.



Calcule probabilidades



Calcule probabilidades

• Distribuição Condicional

É a probabilidade de um evento ocorrer dado que outro evento já ocorreu. É usado em regressão e em várias outras situações para identificar padrões de dependência e fazer previsões.

• Independência

No contexto de probabilidade, duas variáveis são consideradas independentes se a probabilidade de uma não é afetada pelo valor da outra.

• Distribuição Conjunta

Refere-se a todas as possíveis combinações de resultados de dois ou mais eventos. No exemplo da aula, são todas as combinações possíveis de resultados ao jogar dois dados, onde cada combinação tem uma probabilidade igual de ocorrer.



Função de probabilidade



Função probabilidade

• Distribuição Acumulada

É uma função que retorna a probabilidade de uma variável aleatória ser menor ou igual a um valor específico.

• Função de Probabilidade

É uma função que associa um valor a uma probabilidade.

• Função Densidade de Probabilidade

É a versão da função de probabilidade para variáveis contínuas. Ela associa uma densidade, pois a probabilidade de uma variável contínua ser igual a um número específico é conceitualmente zero.



Distribuição Normal



Distribuição Normal

• Densidade

Em estatística, a densidade se refere à concentração de valores em uma distribuição. Na distribuição normal, a densidade é maior em torno da média.

• Teorema Central do Limite

Teorema que afirma que a média de uma variável com um grande número de observações se aproxima de uma distribuição normal.

• Desvio Padrão (sigma)

Medida de dispersão que indica o quanto os valores de um conjunto de dados variam em torno da média. Na distribuição normal, o desvio padrão determina a largura da curva.



Distribuição Média



Distribuição Média

• Distribuição Média

É a distribuição das médias de uma amostra de variáveis independentes e identicamente distribuídas (i.i.d). De acordo com o Teorema Central do Limite, essa distribuição tende a ser normal à medida que o tamanho da amostra aumenta.

• Distribuição T

É uma distribuição de probabilidade que é usada quando a variância da população é desconhecida. Ela tem um parâmetro a mais, o número de graus de liberdade, que é o número de observações na amostra menos um. A distribuição T tem mais massa nas caudas da distribuição em comparação com a distribuição normal, o que significa que é mais provável observar valores raros.



Bons estudos!

