

Aula 12 - Noções de estatística básica e probabilidade aplicadas à IA.

Profa. Gabrielly Queiroz

Noções de Estatística Básica e Probabilidade

- Essenciais para Ciência de Dados e Inteligência Artificial.
- Fundamentos matemáticos para análise, identificação de padrões e previsões.
- Estatística: coleta, representação e análise de dados.
- Probabilidade: estudo de eventos aleatórios e incertos.

Divisão da Estatística

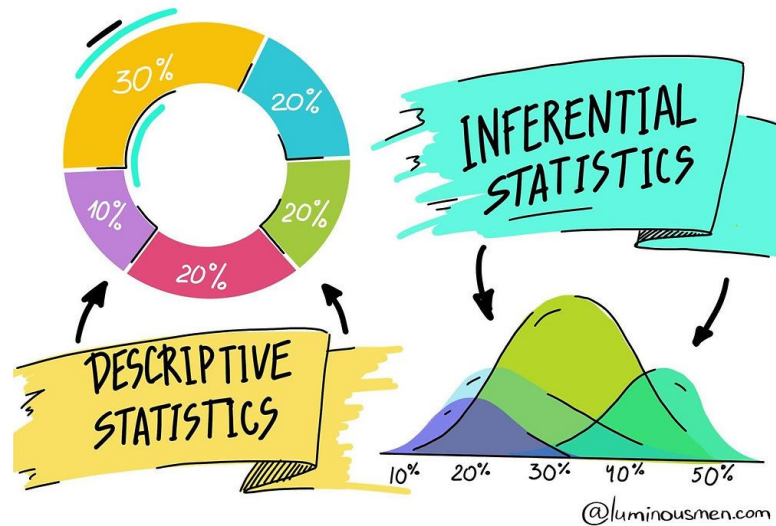
Estatística Básica

- **Estatística Descritiva**

- Organiza, resume e descreve dados.
- Principais medidas:
 - Média: Valor médio.
 - Mediana: Valor central.
 - Moda: Valor mais frequente.
 - Medidas de dispersão: Desvio padrão e variância.

- **Estatística Inferencial**

- Conclusões sobre uma população com base em uma amostra.
- Técnicas: testes de hipóteses, intervalos de confiança, análise de regressão.



Probabilidade

- Mede a chance de eventos ocorrerem (valor entre 0 e 1).
 - Exemplo: Probabilidade de tirar um número par em um dado justo = 0,5.
- Essencial para modelagem preditiva e análise de incertezas.

Distribuições de Probabilidade

- Representam como valores de uma variável aleatória estão distribuídos.
- Principais distribuições:
 - Normal
 - Binomial
 - Poisson

Distribuição Normal

Características

- Modelo de dados contínuos.
- Exemplo: Altura de pessoas.
- Formato de sino: valores próximos à média são mais comuns.

Gráfico

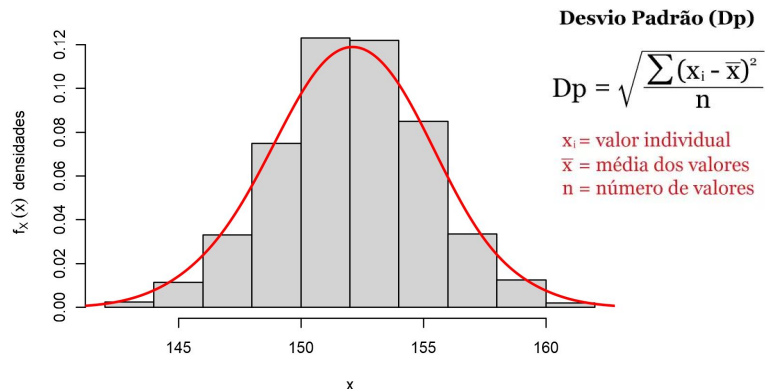
- Larga no meio (média).
- Estreita nas extremidades (valores raros).

Z-score: quão próximo da média.

x : O valor observado (dado que você quer analisar).

μ : A média da distribuição.

σ : O desvio padrão da distribuição.



$$Z = \frac{(x - \mu)}{\sigma}$$

Média (μ) de uma turma em um teste: 70 pontos.

Desvio padrão (σ): 10 pontos.

Aluno tirou $x = 85$.

$$Z = \frac{85 - 70}{10} = 1,5$$

O Z-score é **1,5**, ou seja, o aluno está **1,5 desvios padrão acima da média**.

Distribuição Binominal

Características

- Experimentos com dois resultados possíveis (sucesso ou falha).
- Exemplo: Jogar uma moeda 10 vezes.
 - Resultados: "cara" ou "coroa".
- Aplicações: Aprovação em provas, jogos, etc.

$$P(X = k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k}$$

$P(X = k)$: Probabilidade de obter exatamente k sucessos.

$\binom{n}{k}$: Número de combinações possíveis de k sucessos em n tentativas. Calculado como:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

p : Probabilidade de sucesso em uma única tentativa.

$1 - p$: Probabilidade de falha.

k : Número de sucessos desejados.

n : Número total de tentativas.

$$P(X = 3) = \binom{5}{3} \cdot (0,5)^3 \cdot (1 - 0,5)^{5-3} \qquad \binom{5}{3} = \frac{5!}{3!(5-3)!} = \frac{5 \cdot 4}{2 \cdot 1} = 10$$

$$P(X = 3) = 10 \cdot (0,5)^3 \cdot (0,5)^2 = 10 \cdot 0,125 \cdot 0,25 = 0,3125$$

A probabilidade de obter exatamente 3 "caras" é **31,25%**.

Distribuição de Poisson

$$P(X = k) = \frac{\lambda^k \cdot e^{-\lambda}}{k!}$$

Características

- Modela eventos raros em intervalos de tempo ou espaço.
- Exemplo: Quantas pessoas entram em uma loja por hora.
- Gráfico:
 - Poucos eventos ocorrem com frequência.
 - Valores muito altos ou baixos são raros.

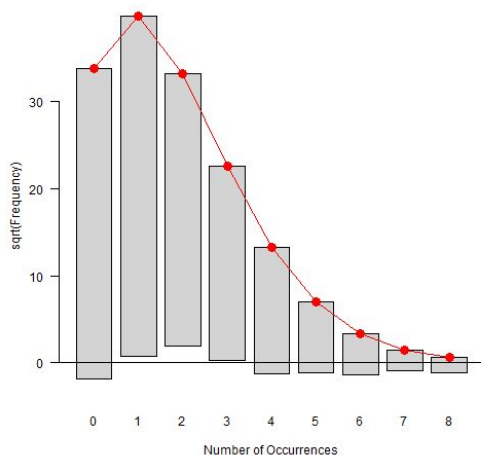
$P(X = k)$: Probabilidade de ocorrer exatamente k eventos no intervalo.

λ : Média esperada de eventos por intervalo (valor esperado).

k : Número de eventos reais que você quer calcular a probabilidade.

e : Número de Euler (aproximadamente 2,718).

$k!$: Fatorial de k ($k! = k \cdot (k - 1) \cdot (k - 2) \cdot \dots \cdot 1$).



Uma pizzeria recebe, em média, **3 pedidos por hora** ($\lambda=3$). Qual é a probabilidade de receber exatamente **5 pedidos** em uma hora ($k=5$)?

$$P(X = 5) = \frac{3^5 \cdot e^{-3}}{5!}$$

A probabilidade de receber exatamente **5 pedidos** em uma hora é **10,09%**.

Importância para Ciências de Dados e IA

Estatística e Probabilidade na Prática

- Transformam dados brutos em conhecimento útil.
- Estatística Básica:
 - Organização e resumo de dados.
 - Identificação de padrões e variabilidade.
 - Exemplo: Análise de tendências de consumo.
- Estatística Inferencial:
 - Conclusões baseadas em amostras.
 - Validação de modelos.

Probabilidade:

- Previsão de eventos futuros.
- Aplicações:
 - Algoritmos de aprendizado de máquina.
 - Classificação e tomada de decisões.

Importância para Ciências de Dados e IA

Distribuição normal: para dados contínuos que variam em torno de uma média, como altura ou erro de previsão.

Distribuição binomial: para contar quantos acertos ou erros acontecem em várias tentativas, como prever se uma IA acerta ou erra.

Distribuição de Poisson: para contar quantas vezes algo raro acontece em certo tempo, como quantas falhas ocorrem por hora.

Atividade

Um hospital de pronto-atendimento observa que, em média, 4 pacientes chegam para atendimento de emergência por hora durante o período da noite (das 18h às 6h).

a) Suponha que, em determinada hora, chegam 12 pessoas ao hospital, mas nem todas são emergências. Cada pessoa tem $1/3$ de chance de ser classificada como emergência. Qual é a probabilidade de que exatamente 6 dessas 12 pessoas sejam casos de emergência?

b) Usando a distribuição de Poisson, qual a probabilidade de exatamente 6 pacientes de emergência chegarem ao hospital em uma hora qualquer durante o período da noite? (Use $\lambda=4$)