

Repositório de Estatística e Probabilidade

UFSC Joinville - EMB5010

Artur Gemaque

15 de setembro de 2024

Resumo

Este documento tem como principal funcionalidade registrar os conteúdos ensinados em sala de aula pelo professor Jaimes, ademais servirá como fonte de estudo para as provas referentes à Matéria.

1 Medidas Descritivas:

$$S = \sqrt[2]{S^2}$$

1.1 Valores de cálculo

Vamos começar com as definições mais simplórias da matéria que são resultado do cálculo dos dados.

1.1.1 Média

A média aritmética de um conjunto de n números é a soma desses números dividida por n .

$$\bar{x} = \sum_{i=1}^n \frac{X_i}{n}$$

1.1.2 Variância

A variância é um conceito fundamental em estatística que mede a dispersão de um conjunto de dados em relação à sua média.

$$(S)^2 = \frac{\sum_{i=1}^n (X_i - \bar{x})^2}{n-1}$$

1.1.3 Desvio padrão

O desvio padrão é uma medida de dispersão estatística que indica o quão afastados os dados de um conjunto estão da sua média. Ele é a raiz quadrada da variância, o que o torna mais interpretável.

1.2 Valores de análise

Partindo agora para valores que são obtidos a partir dos dados ordenados em formato crescente. Observação que as seguintes fórmulas serão para que possamos obter as posições dos referidos valores.

1.2.1 Mediana

Mediana é o número no centro de um grupo de números.

$$Md = X_{(\frac{n+1}{2})}$$

1.2.2 Quartis

Os quartis são valores que dividem uma amostra de dados em quatro partes iguais e são usados para avaliar a dispersão e a tendência central de um conjunto de dados. Eles são os valores contidos nas posições de $n * 25\%$ entre outras porcentagens, caso "n" dê um valor quebrado o Quartil vai ser a média entre os dois valores.

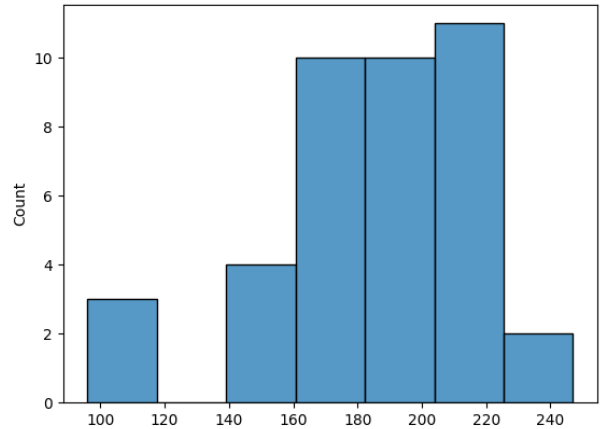
$$Q_1 = X_{(\frac{n+1}{4})}$$
$$Q_3 = X_{(\frac{3(n+1)}{4})}$$

1.2.3 Moda

A moda é o valor que mais aparece em um conjunto de dados, ou seja, o valor que tem maior frequência.

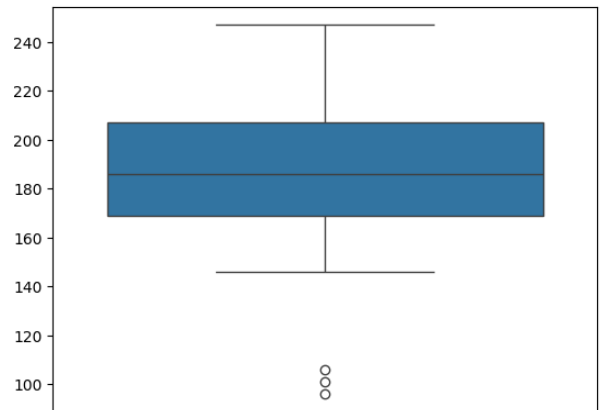
1.2.4 Amplitude

A amplitude de um conjunto de dados é a diferença entre o maior e o menor valor. Para calcular a amplitude, subtrai-se o menor valor do maior.



2.0.2 Boxplot

O Box Plot, que estudamos no curso Green Belt, é uma ferramenta gráfica que ajuda a identificar a existência de possíveis outliers no conjunto de dados. Em um boxplot são apresentadas 5 estatísticas: o mínimo, o primeiro quartil (Q1), a mediana, o terceiro quartil (Q3) e o máximo.



2 Gráficos

Partindo para a seção de Gráficos...

2.0.1 Histograma

Um histograma é uma espécie de gráfico de barras que demonstra uma distribuição de frequências. No histograma, a base de cada uma das barras representa uma classe e a altura representa a quantidade ou frequência absoluta com que o valor de cada classe ocorre.