

Lista de Exercícios de Estatística para Data-Science

Responda as perguntas com base no conjunto de dados a seguir:

<https://drive.google.com/file/d/1Fkd92kLDCO-Bii6jKJ5V2cQi9II4gvAn/view?usp=sharing>

1. **O conjunto de dados contém mais de 2000 valores e a amplitude é maior que 0,2.**
☐ Verdadeiro
☐ Falso
2. **Não existem valores faltantes no conjunto de dados?**
☐ Verdadeiro
☐ Falso
3. **Qual é a afirmação correta sobre a média, mediana e moda do conjunto de dados fornecido?**
 - a) A média é maior que a mediana, indicando uma distribuição simétrica dos dados.
 - b) A média é menor que a mediana, o que sugere uma distribuição assimétrica à esquerda.
 - c) A média é maior que a mediana, sugerindo que a distribuição dos dados é assimétrica à direita.
 - d) A média e a mediana são praticamente iguais, indicando uma ausência de outliers significativos.
4. **Classifique as seguintes afirmações como verdadeiras (V) ou falsas (F) com base nas características do conjunto de dados:**
 - I - A média do conjunto de dados é superior a 0,07.
 - II - A mediana é um indicador mais robusto contra outliers do que a média, nesse conjunto de dados.
 - III - A moda do conjunto de dados é menor que a mediana.
5. **Qual das seguintes opções descreve corretamente os quartis do conjunto de dados?**
 - a) Todos os quartis estão acima de 0,1, refletindo uma distribuição de dados com valores majoritariamente altos.
 - b) O terceiro quartil (Q3) é inferior a 0,05, indicando que 75% dos dados estão concentrados nos valores mais baixos.
 - c) O primeiro quartil (Q1) é 0,05, o segundo quartil (mediana ou Q2) é 0,0688, e o terceiro quartil (Q3) é 0,0949.
 - d) A distância interquartilica (Q3-Q1) é menor que 0,02, mostrando uma baixa dispersão dos valores centrais do conjunto.

6. Qual das seguintes afirmações sobre desvio-padrão e variância é correta?

- a) O desvio-padrão é a raiz quadrada da variância e representa a média dos desvios em relação à média dos dados.
- b) A variância é uma medida de dispersão que mostra como os dados estão concentrados ao redor da média.
- c) Variância e desvio-padrão são medidas de centralidade que indicam o valor típico ao qual os dados tendem a agrupar-se.
- d) Desvio-padrão e variância são medidas independentes e não possuem relação matemática entre si.

7. Com base nas propriedades do conjunto de dados fornecido, qual das seguintes afirmações sobre a variância e o desvio padrão é correta?

- a) A variância do conjunto de dados é aproximadamente 0.0307, indicando que os valores estão muito dispersos em relação à média.
- b) O desvio padrão do conjunto de dados é aproximadamente 0.0307, o que sugere uma dispersão moderada dos dados em torno da média.
- c) O desvio padrão é superior a 0.1, demonstrando uma alta variabilidade dos dados.
- d) A variância é inferior a 0.01, mostrando que os dados estão muito concentrados ao redor da média.

8. Qual das seguintes afirmações sobre assimetria e curtose é correta?

- a) A assimetria é uma medida de quão simétrica é a distribuição dos dados, enquanto a curtose indica a "cauda" da distribuição em comparação com uma distribuição normal.
- b) A curtose é uma medida que indica o pico da distribuição dos dados, e uma curtose negativa implica uma distribuição mais achatada que a normal.
- c) Assimetria positiva indica uma cauda mais longa à esquerda da distribuição, enquanto uma curtose alta significa uma distribuição com picos mais acentuados que a normal.
- d) A assimetria mede a concentração de dados em torno da média, e a curtose é uma medida da variabilidade dos dados em relação à mediana.

9. Qual das seguintes opções descreve corretamente a assimetria e a curtose do conjunto de dados fornecido?

- a) A distribuição dos dados apresenta assimetria negativa e uma curtose indicativa de caudas mais pesadas que uma distribuição normal.
- b) A assimetria positiva da distribuição sugere uma cauda mais longa à esquerda, com uma curtose alta, indicando picos mais acentuados que uma distribuição normal.
- c) A distribuição é simétrica com uma curtose negativa, implicando que as caudas são mais pesadas e o pico menos acentuado que numa distribuição normal.
- d) A distribuição dos dados é moderadamente assimétrica para a direita e apresenta caudas mais leves que uma distribuição normal, conforme indicado por sua curtose.