

UNIVERSIDADE FEDERAL DE MINAS GERAIS

PROGRAMA DE DISCIPLINA

DEPARTAMENTO: Estatística						
DISCIPLINA: Fundamentos de Estatística e Ciências de Dados						
CÓDIGO: EST773						
CLASSIFICAÇÃO: OB						
PRÉ-REQUISITO: Cálculo I						
CARGA HORÁRIA:						
	TEÓRICA:	60h				
	PRÁTICA:					
CRÉDITOS: 4			-			

EMENTA: Introdução à Estatística e Ciência de Dados. Visualização de dados: tipos de variáveis, gráficos e tabelas, medidas de posição e variabilidade. Fundamentos de probabilidade e modelos probabilísticos: tipos de eventos, probabilidade condicional e independência de eventos, variáveis aleatórias discretas e contínuas. Tomada de decisão com base em evidências: estimação pontual e intervalar, conceitos de testes de hipóteses, testes para amostras

dependentes e independentes. Prática computacional.

CONTEÚDO PROGRAMÁTICO:

1- Introdução:

O papel da Estatística na Ciência de Dados. Exemplos de aplicação. Amostragem:

população x amostra, amostragem probabilística e não probabilística.

2- Visualização de dados:

Tipos de variáveis e níveis de mensuração (nominal, ordinal, intervalar). Visualização de dados em tabelas e gráficos (setores, barras e pareto, pontos, ramo-e-folhas, histograma, ogiva, polígonos de frequências, gráfico de linhas, gráfico de cascata, gráfico de radar, combinação de gráficos). Síntese de dados: medidas de posição (média, mediana, moda, percentis), de variabilidade (amplitude, variância, desviopadrão, coeficiente de variação) e de associação. Gráfico box-plot e gráfico de dispersão. Prática computacional: gráficos básicos e estatísticas descritivas.

3- Fundamentos de probabilidade e modelos probabilísticos:

Experimento e eventos aleatórios. Definição de probabilidade, probabilidade condicional, independência de eventos, regra de Bayes. Variáveis aleatórias discretas: função de probabilidade, esperança e variância, principais distribuições discretas. Variáveis aleatórias contínuas: função densidade, função de distribuição acumulada; média, variância e percentis, principais distribuições contínuas. Prática computacional: geração de dados de distribuições conhecidas, cálculo de probabilidades e quantis associados, gráficos de densidade.

4- Tomada de decisão com base em evidências:

Estimação de quantidades de interesse: estimação pontual e intervalar para a média, variância e proporção. Grandes amostras: Teorema central do limite. Conceitos básicos de testes de hipótese: hipótese nula e alternativa, nível de significância, valor-p. Comparação de dois grupos: amostras dependentes e amostras independentes. Testes para ajuste de distribuição. Prática computacional: funções básicas e pacotes.

5- Tópicos especiais:

Análise de variância envolvendo um fator fixo. Correlação. Regressão linear simples. Prática computacional.

BIBLIOGRAFIA GERAL RECOMENDADA:

A) Bibliografia básica:

- MONTGOMERY, Douglas C., RUNGER, George C., CALADO, Verônica. (2000). Estatística Aplicada e Probabilidade para Engenheiros. Grupo Gen-LTC.
- WICKHAM, H., GROLEMUND G. (2017) R for Data Science. O'Reilly.
- VANDERPLAS, J. (2017) Python Data Science Handbook: Essential Tools for Working with Data, 1st Edition,

1 of 2

eBook Kindle. O'Reilly.

• ALCOFORADO, L. F., LEVY, A. Visualização de Dados com Software R. LFA: Niterói, RJ, 2017. (e-book)

B) Bibliografia complementar:

- DEVORE, J. L.; Probabilidade e Estatística para Engenharia e Ciências. São Paulo: Cengage Learning, 2019. 630 p.
- CRAWLEY, Michael J. *The R Book*. John Wiley & Sons, 2012.



Documento assinado eletronicamente por **Glaura da Conceição Franco**, **Professora do Magistério Superior**, em 24/03/2023, às 13:43, conforme horário oficial de Brasília, com fundamento no art. 5º do <u>Decreto nº 10.543, de 13 de novembro de 2020</u>.



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2172794** e o código CRC **93B45A78**.

Referência: Processo nº 23072.216262/2023-41

SEI nº 2172794

2 of 2