

Prova Prática

Pontuação total: 10

Prazo: 02/05/2024 - A ser marcado em aula

NÚMERO DO GRUPO:

Nome:	Matrícula:
Nome:	Matrícula:
Nome:	Matrícula:
Nome:	Matrícula:
Nome:	Matrícula:

Na resolução da prova, cada grupo (1 ao 14) deve analisar os dados correspondentes a seu grupo.

Os dados estão disponíveis no formato [csv: dados_p2.csv] ou [txt: dados_p2.txt] e também na [página da disciplina](#) no assunto [**Provas**], no final da página.

As variáveis aleatórias de interesse são: Y1 (medida em un), Y2 (medida em un) e Sexo. Adicionalmente, considere que Y1 e Y2 não podem assumir valores reais negativos. Os dados são fictícios e tem finalidades exclusivamente didáticas para fins de avaliação prática em análise exploratória de dados.

Realizar a análise dos dados com respostas às seguintes questões:

1 AED: Apresentações tabulares e gráficas (2.0)

1.1 Diagrama de caixa (boxplot) para Y1 e Y2 (1.0)

- 1) (0.5) Antes e após a eliminação de possíveis outliers¹;
- 2) (0.5) Após a eliminação de possíveis outliers².

1.2 Para Y1 (1.0)

- 1) (0.5) Uma apresentação tabular contendo apenas as frequências: absoluta (F_i), relativa (Fr , %) e acumulada (Fac , %), nessa ordem²;
- 2) (0.5) Histograma e o polígono de frequência acumulada dos dados².

2 AED: Medidas estatísticas básicas (5.0)

2.1 AED: Medidas determinadas a partir dos dados (3.0)

Para as variáveis Y1 e Y2 elaborar apresentações tabulares² contendo as seguintes estimativas:

- 1) (1.0) Tendência central: média, mediana e moda;
- 2) (1.0) Posição: quartis e decis;
- 3) (1.0) Dispersão: amplitude total, variância, desvio padrão e coeficiente de variação.

¹Não distinguindo sexo

²Para cada sexo: M seguido de F

2.2 AED: Medidas determinadas a partir de apresentações tabulares (2.0)

No arquivo [tabelas_p2.txt] são encontradas tabelas de distribuição de frequências do tipo comum, dessas que se encontra em publicações. Considere que estas tabelas descrevem um assunto de seu interesse - publicado - e que é necessário determinar as medidas estatísticas básicas com finalidades de entendimento e comparações.

Elabore uma apresentação tabular contendo:

- 1) (0.8) Tendência central: média, mediana e moda;
- 2) (0.4) Posição: quartis;
- 3) (0.8) Dispersão: amplitude total, variância, desvio padrão e coeficiente de variação.

3 AED: Medidas estatísticas de associação(2.0)

- 1) (0.8) Estimativas: covariância e correlação linear simples²;
- 2) (0.8) Diagramas de dispersão dos dados^{2,3};
- 3) (0.4) Um estudo semelhante foi realizado em um outro município, por outras pessoas. Contudo, as unidades de medida usadas foram: $Y1 (100 * un)$ e $Y2 (100 * un)$.

Para comparar associações entre as variáveis de ambos os estudos, qual seria a medida estatística recomendada? Justifique.

4 Contextualização (1.0)

Localize um artigo científico - (periódico Qualis A ou B) em área de seu interesse e preferencialmente ligada às Ciências Sociais - no qual a análise exploratória de dados (AED) - possivelmente com medidas de associação, teve papel preponderante. Discuta o artigo com ênfase nos recursos da AED usados e também na adequação das normas básicas das apresentações gráficas e tabulares adotada pelo autor e pelo periódico.

Observações:

- As provas deverão ser entregues impressas.
- Para possibilitar devida a correção, anexe esta prova devidamente preenchida com o número do grupo e os alunos que o compõe na primeira página das respostas.
- As normas para apresentações gráficas e tabulares são obrigatórias, serão observadas e corrigidas.
- Sugere-se (mas não é obrigatório) o uso do ambiente R na resolução das questões propostas.
- Cada hora de atraso na entrega da avaliação implica na perda de 25%. Portanto, após 4 horas não entregue.

³Considere $Y2$ no eixo das ordenadas e $Y1$ no eixo das abscissas