

Classwork 04

Instruções:

1. Este Classwork (CW) poderá ser resolvido em duplas;
2. O CW precisará conter o(s) **nome(s) completo(s)** e a(s) **matrícula(s)** do(s) discente(s);
3. **Serão considerados apenas** os Classworks resolvidos por discentes presentes;
4. O CW deverá ser enviado via **Google Classroom até as 09h51 do dia 29/04/2026** em **um único arquivo** e em **PDF** (*Portable Document Format*);
5. **Não serão considerados** Classworks com péssima qualidade visual ou em desacordo com as instruções anteriores.

Exercícios. Considere o conjunto de dados dos campeões do Campeonato Brasileiro de Futebol Masculino.

1. Assim como em Magalhães et al. (2025), considere apenas os campeonatos entre os anos de 2006 e 2023. A partir desses dados, ajuste um modelo de regressão linear no qual a variável resposta é o número de pontos obtidos (“Pts”), considerando como variáveis explicativas o saldo de gols (“Gd”) e o estado (“Homestate”).
 - (a) Realize uma análise exploratória dos dados. Apresente gráficos e medidas descritivas que permitam compreender o comportamento das variáveis. Discuta possíveis padrões, relações entre as variáveis e a presença de valores atípicos.
 - (b) Ajuste o modelo completo e avalie a significância estatística dos coeficientes (considere, por exemplo, um nível de significância de 5%). Caso existam variáveis não significativas, proponha um modelo reduzido, justificando a exclusão dessas variáveis.
 - (c) Interprete os resultados do modelo final, incluindo os coeficientes estimados. Discuta o significado dos efeitos das variáveis explicativas sobre a pontuação, mantendo as demais constantes.

- (d) A partir do modelo final, realize previsões pontuais e intervalares para o número de pontos dos times campeões dos anos de 2024 e 2025. Compare com os valores observados e interprete os resultados.
2. Considere todo o conjunto de dados e ajuste um modelo de regressão linear no qual a variável resposta é o aproveitamento de pontos (“Ratio”), considerando como variáveis explicativas o saldo de gols (“Gd”) e o estado (“Homestate”).
- (a) Realize uma análise exploratória dos dados. Apresente gráficos e medidas descritivas que permitam compreender o comportamento das variáveis. Discuta possíveis padrões, relações entre as variáveis e a presença de valores atípicos.
- (b) Ajuste o modelo completo e avalie a significância estatística dos coeficientes (considere, por exemplo, um nível de significância de 5%). Caso existam variáveis não significativas, proponha um modelo reduzido, justificando a exclusão dessas variáveis.
- (c) Verifique os principais pressupostos do modelo de regressão linear (normalidade dos resíduos, homocedasticidade e ausência de padrões sistemáticos), utilizando gráficos e/ou testes formais.
- (d) Interprete os resultados do modelo final, incluindo os coeficientes estimados. Discuta o significado dos efeitos das variáveis explicativas sobre o aproveitamento de pontos, mantendo as demais constantes.

Referências

Magalhães, T. M., D. I. Gallardo, and M. A. Diniz (2025). Inferential and predictive procedures for inverse gamma regression model. *Journal of Statistical Computation and Simulation* 95(11), 2327–2342.