

Núcleo de Avaliação: Núcleo II

Área temática: Probabilidade e Estatística

Área do Conhecimento: Probabilidade e Estatística Aplicadas

Estudo de Decomposição em fatores paralelos para análise de dados educacionais.

Lara Carla Freitas Rodrigues, Luiza Helena Félix de Andrade, Sílvio Roberto Fernandes

Ao longo da história humana, diversas linhas de pensamento vêm sendo elaboradas para analisar a relação entre a educação e as desigualdades sociais. Ainda no século XX, Pierre Bourdieu defendia que o desempenho individual de cada estudante é profundamente afetado por sua realidade social, afirmação que foi comprovada pelo resultado de alguns relatórios do mesmo século, como o de Plowden e o de Coleman. Essa realidade não mudou com o passar dos anos, na verdade, segundo dados do Índice de Desenvolvimento da Educação Básica (Ideb) de 2023 e o Programme for International Student Assessment (PISA) de 2022, no Brasil, os alunos mais pobres continuam tendo um desempenho inferior quando comparado aos mais ricos. Sabendo disso, é evidente que não é possível falar sobre o desenvolvimento da educação do país deixando de lado questões socioeconômicas. Contudo, existem diversas variáveis que compõem essa dimensão e, para elaborar um plano assertivo contra a descontinuação das desigualdades sociais e educacionais no país, é relevante ter conhecimento acerca de quais fatores são mais relevantes nesse contexto, por isso, o presente estudo tem como objetivo apontar quais questões socioeconômicas tiveram maior peso no desempenho dos estudantes brasileiros, no contexto do ensino médio, nos anos de 2022 e 2023. Para isso, analisaram-se as respostas do questionário socioeconômico e desempenho dos candidatos do Exame Nacional do Ensino Médio (ENEM) nos anos de selecionados. A escolha do ENEM se deu devido à sua ampla cobertura nacional, riqueza de detalhes e fácil acesso através do portal do Instituto Nacional de Estudo e Pesquisas Educacionais Anísio Teixeira (INEP). Além disso, o período de análise adotado priorizou os anos mais recentes do exame (os dois últimos, disponíveis na época da pesquisa), a fim de garantir dados que refletissem o cenário atual. Após a aquisição dos dados, eles passaram por um tratamento que envolveu a exclusão dos participantes que não realizaram algum dia de prova ou estavam fazendo a prova somente por experiência, os “Treineiros”. Posteriormente, usando bibliotecas Python, foi calculada a média geral de cada ano, a qual foi usada para dividir os participantes no grupo de maior proficiência, com média igual ou superior à média geral, e de menor proficiência. Os dados foram organizados na forma de um tensor de ordem 3 (alunos x questões x proficiência). Para identificar quais questões socioeconômicas mais relacionadas à proficiência dos grupos, foi aplicada a técnica de decomposição em fatores paralelos (PARAFAC), isso porque o método oferece uma solução única, sem perda de informações relevantes e uma sintetização do volume de informações do banco de dados de uma vez só. Os resultados indicaram que a renda familiar e a escolaridade de ambos os pais foram os que mais fizeram diferença no desempenho dos estudantes, dialogando com estudos clássicos e recentes, oferecendo evidências robustas e atuais que podem fundamentar a formulação de políticas públicas que ajudem a combater o processo da transmissão intergeracional da pobreza e da continuação das desigualdades sociais.

Palavras-chave: Educação, PARAFAC, Desigualdade.

Agência financiadora: PICI-UFRSA.

Campus: Mossoró.

