



Colóquio Internacional Educação e Contemporaneidade

www.coloquioeducon.com
27 a 29 de setembro de 2023



Anais, Volume XVII, n. 8, set. 2023
ISSN: 1982-3657 | Prefixo DOI: 10.29380

Eixo 8

Educação, Inovação e Tecnologias

**PREVISIBILIDADE DA EVASÃO ESCOLAR NUM MODELO DE
REGRESSÃO BETA INFLACIONADO**

PREDICTABILITY OF SCHOOL LEADERSHIP IN AN INFLATIONED
BETA REGRESSION MODEL

Carlos Alberto Silva

DOI: <http://dx.doi.org/10.29380/2023.17.08.01>

Recebido em: 17/08/2023

Aprovado em: 16/09/2023

Editores responsáveis:

Veleida Anahi Capua da Silva Charlot e Bernard Charlot



Colóquio Internacional Educação e Contemporaneidade

www.coloquioeducon.com
27 a 29 de setembro de 2023



*PREVISIBILIDADE DA EVASÃO ESCOLAR NUM MODELO DE REGRESSÃO BETA
INFLACIONADO*

*PREDICTABILITY OF SCHOOL LEADERSHIP IN AN INFLATIONED BETA REGRESSION
MODEL*

RESUMO

A evasão escolar é um fenômeno presente em menor ou maior grau nos sistemas educacionais de vários países, sendo mais acentuado nas populações de maior vulnerabilidade econômica ou social e, as vezes com estas associadas. Muito se discute a respeito do fenômeno, pouco se percebe na efetiva busca de ferramentas eficazes de seu combate. No presente trabalho constata-se de que as instituições escolares brasileiras, independente de natureza privada ou pública, independente do nível de ensino, não dispõem de um controle de dados consistentes para uma efetiva análise com rigor técnico, portanto sem condições de previsibilidade destes índices de evasão, extenuando a atenção à preditivos recorrentes, como nota e frequência sem a cultura da inclusão das variáveis comportamentais do indivíduo, da instituição e do contexto geosocial. Conseguira-se nesta busca, um banco de dados de duas escolas portuguesas com variáveis consistentes as quais consideradas por especialistas, pedagogos e na literatura educacional, como explicativas do fenômeno da evasão escolar. Realizada a estatística descritiva, gráficos e testes adequados, aplicaram-se os modelos de regressão logísticos exponenciais. Após os ajustes e comparação dos modelos pelos testes; lrtest, Wald, correlação; observando as características de melhor ajuste obtiveram-se dois modelos para a modelagem nas famílias GAMLSS e betareg.

Palavras-chave: Modelo de regressão beta inflacionado. modelagem de evasão escolar. previsão da evasão escolar.

ABSTRACT



Colóquio Internacional Educação e Contemporaneidade

www.coloquioeducon.com
27 a 29 de setembro de 2023



School dropout, a phenomenon present to a lesser or greater degree in the educational systems of several countries, accentuated in populations with greater economic or social vulnerability, and sometimes associated with them. Much is discussed about the phenomenon, little is noticed in the effective search for effective tools to combat it. The work finds that Brazilian school institutions, regardless of whether they are private or public, of the level of education, do not have consistent data control for effective analysis with technical rigor, therefore without predictability of these dropout rates, exhausting attention to recurrent predictors, such as grade and frequency without the culture of inclusion of behavioral variables of the individual, the institution and the geosocial context. In this search, a database of two Portuguese schools was obtained with consistent variables which are considered by specialists, pedagogues and in the educational literature, as explanatory of the phenomenon of school dropout. After performing descriptive statistics, graphs and appropriate tests, exponential logistic regression models were applied. Comparing models by tests; lrtest, Wald, correlation; observing the best fit characteristics, two models were obtained for modeling in the GAMLSS and betareg families.

Keywords: Inflated beta regression model. dropout modeling. dropout prediction.

INTRODUÇÃO

1. Introdução

As causas do abandono e evasão escolar relacionam-se a questões sociais, geográficas, econômicas, educacionais e no contexto do indivíduo, como emocionais, raciais e de saúde. Barros (2021), descreve que a probabilidade de um indivíduo ter uma ocupação formal aos 35 anos é de 77% para aqueles que completaram os estudos e 50% para quem evadiu do curso na universidade, no nordeste brasileiro. A renda mensal nessa mesma idade é outro aspecto que tem uma diferença expressiva: 1,37 contra 0,62 salários mínimos. Em resumo, completar o ensino médio proporciona que o indivíduo tenha um nível de qualidade de vida em torno de 83% do ideal; 1,7 vezes maior do que aqueles que não concluem essa etapa. Neste sentido, a concepção de um modelo de previsão eficaz, através de uma análise de dados usando técnicas estatísticas adequadas, torna-se necessária como uma ferramenta de cooperação para que gestores educacionais possam tomar decisões que vizam reduzir os altos índices da evasão escolar numa região ou país, começando pela unidade escolar.



Colóquio Internacional Educação e Contemporaneidade

www.coloquioeducon.com
27 a 29 de setembro de 2023



Como primeira proposta, suscita-se que haja uma mudança na forma de registro, armazenamento e análise de dados pelos especialistas educacionais, um hábito ausente nas unidades escolares e órgãos gestores pelo fato característico de suas formações acadêmicas. Como uma dentre outras justificativas para relevância desta proposta de criação do modelo, considera-se que a minimização da evasão tem que ser tratada como um fator de proteção do estudante jovem, na faixa dos 10 aos 17 anos, prioritariamente, a permanência no ambiente escolar se torna para a política pública uma ferramenta de redução de risco, com intuito de baixar os efeitos da pobreza, do impacto maternidade ou paternidade na adolescência e da exposição em relação a violência urbana. Sousa et al (2018), relata que em 2014 aproximadamente um terço das jovens brasileiras inclusas nesta faixa etária que abandonaram a escola já eram mães, não trabalhavam nem estudavam e que apenas 2% daquelas que permaneceram no ambiente escolar tinham filhos.

Com a criação da Prova Brasil em 2005, e a formulação do Índice de Desenvolvimento da Educação Básica (IDEB) em 2007, Lacruz et al (2019) afirmam que o Sistema de Avaliação do Ensino Básico (SAEB), fora ampliado e redimensionado para se adequar às prerrogativas internacionais de “competitividade” e de “eficiência” na formação escolar. Desta forma, examina escolas e estudantes das redes públicas e privadas, em áreas rurais e urbanas, matriculados nos anos iniciais e finais do ensino básico, incluindo o ensino médio. Afirma Coelho (2008, p. 231), “a avaliação se firma cada vez mais como elemento da regulação, da administração gerencial e competitiva do Estado-avaliador no Brasil”, ou seja, como um instrumento que promove o aumento da interferência e do controle sobre a educação. Como podemos observar, o sistema proposto não visa mitigar aspectos da evasão e sim de controle e ranqueamento por unidades escolares, municípios e estados.

Visto que há uso indiscriminado do modelo de regressão logística, propostos na literatura para que a análise da relação entre a variável resposta e as demais covariáveis independentes seja feita da melhor maneira possível, o presente estudo busca identificar se a distribuição da variáveis em amostra composta por grupos de discentes do ensino médio, por meio do modelo de regressão beta, poderão ser avaliados numa escala de risco do fenômeno. Essa abordagem pressupõe o fato de que a variável resposta corresponde a uma proporção definida no intervalo (0,1). Ferrari & Cribari-Neto (2004) propuseram o modelo de regressão beta para este tipo de variável resposta, ou seja, dados como taxas ou proporções, equivalentes às prevalências no objeto de estudos. Com relação aos estimadores destas variáveis, estudos revelam que a presença de pontos de alavancagem na matriz dos regressores, deteriora o desempenho em amostras finitas de estimadores consistentes na matriz de covariâncias do estimador de mínimos quadrados ordinários (EMQO), conduzindo a testes quase-t associados (CRIBARI-NETO & GOIS, 2002 ; CRIBARI-NETO & ZARKOS, 2001).

O processo de seleção conjunta de variáveis regressoras para o estimador e para a seleção de modelos com o esquema proposto, pressupõe um menor custo computacional, visando melhor desempenho. Dentre os critérios investigados recomenda-se o critério de informação de Akaike (AIC). O AIC é dito, como mais conhecido e aplicado em diferentes classes de modelos. Baseados no AIC, diversos critérios têm sido propostos, dentre eles; o SIC, o HQ e o AICc; com o objetivo de estimar o valor esperado da log-verossimilhança, que é uma medida de discrepância entre o modelo verdadeiro e o



Colóquio Internacional Educação e Contemporaneidade

www.coloquioeducon.com
27 a 29 de setembro de 2023



modelo proposto estimado.

Observa-se, em pequenas amostras, ou quando o número de parâmetros do modelo é grande relativamente ao tamanho amostral, que o AIC se torna viesado e tende a selecionar modelos com alta dimensionalidade. Ao considerarmos uma estrutura de regressão também para o parâmetro de dispersão, introduzimos um maior número de parâmetros a serem estimados no modelo. Isso pode diminuir o desempenho dos critérios de seleção quando o tamanho amostral é pequeno ou moderado (BAYER, 2011).

Como *objetivo geral* a priori este trabalho busca a *elaboração de uma ferramenta de análise capaz de gerar um índice de previsibilidade para evasão escolar*, com o qual o gestor educacional possa agir em tempo hábil, intervindo nas variáveis significativas causadoras, minimizando recursos e contribuindo no combate de uma desigualdade social de oportunidades, efeito consequente do fenômeno.

Em outra frente, o *objetivo específico* refere-se a disseminação desta cultura de *criação da análise das informações via elaboração de um banco de dados consistente nas unidades escolares*. Para tal propósito, constará justificativa para cada variável considerada determinante da evasão escolar, acompanhada da análise do banco de dados utilizado na concepção do modelo regressivo beta inflacionado. Esta análise considera além do aspecto descritivo da variável, sua distribuição de probabilidades para adequação ao modelo, pois muitas são categóricas em virtude da face comportamental do estudante e descritores da unidade escolar. Todas as etapas de análise e concepção do modelo foram realizadas com recursos computacionais disponíveis no software RStudio e interfaces amigáveis.

Concluída a fase de escolha destas variáveis e suas características, a etapa seguinte fora dedicada à concepção do *modelo estatístico de regressão beta inflacionado*, comparando através dos ajustes e testes estatísticos, o poder de previsibilidade de cada variável em proporcionar um índice confiável para tomada de decisão do gestor escolar.

2. Desenvolvimento

2.1 Evasão: a face do problema.



Colóquio Internacional Educação e Contemporaneidade

www.coloquioeducon.com
27 a 29 de setembro de 2023



A evasão de estudantes no sistema educacional brasileiro tem ao longo do tempo desafiado especialistas, órgãos governamentais, dirigentes escolares e a comunidade escolar da unidade de ensino, na identificação e proposta de mitigação do problema, dada a diversidade presente nas diversas regiões do país. De acordo com Mendes (2013), o estudante, objeto da inclusão no sistema escolar, deve ser considerado como indivíduo que apresenta necessidade educativa. Abramovay & Castro (2003), afirmam que o principal problema em cada nível de ensino consiste no desinteresse dos alunos frente as atividades escolares, sendo este problema apontado por seis em cada dez estudantes do ensino médio.

Gattás (2018) cita no texto “ *O preço alto da evasão escolar*”, o fato de que em 2017, no ensino médio, 1,5 milhão sequer faziam matrícula, 700 mil abandonaram o ano letivo e outros 600 mil foram reprovados no final. A isto associam-se fatores divididos em três categorias, a primeira externa explicada pela condição socioeconômica, a segunda interna explicada condição do ambiente escolar e a terceira própria do indivíduo. Veja o detalhamento no Quadro 1.

Quadro 1. Classificação dos fatores determinantes para evasão escolar

Tipo	Contexto	Variáveis
Externo	Socioeconômico	Condição de acesso, Violência, Saúde, Trabalho
Interno	Qualidade do sistema escolar	Défict de aprendizado, Estrutura da escola, Qualidade do ensino, Clima escolar
Individual	Falta de informação	Currículo desmotivador, Repetência

Fonte: <http://gesta.org.br/o-preco-alto-da-evasio-escolar/> 2017

As variáveis pressupostas no Quadro 1, resultam em graves consequências tanto sociais, elevado custo de políticas públicas, quanto individuais, como a redução da perspectiva de renda para 35% em relação ao estudante que integraliza o ensino médio.

Considerar estas variáveis subdivididas em grupos, torna imperativo ao analisá-las, pois segundo pesquisa PNAD Educação (2015), 14,1% dos estudantes estavam fora da escola e ainda 8,1% deixaram de concluir o nível fundamental até os 14 anos. Dentre as principais causas, 39,1% por necessidade de trabalho, 29,2% por falta de interesse, 20,4% por assumirem maternidade/paternidade prematura e 11,3% por assumirem tarefas domésticas. Em Sergipe encontramos um caso, onde o caos da evasão mobilizou a comunidade escolar,



Colóquio Internacional Educação e Contemporaneidade

www.coloquioeducon.com
27 a 29 de setembro de 2023



ao chegar ao cargo, em 2010, o Prof. Allmir Pinto, Diretor da Escola Estadual Edélzio Vieira de Melo, Município de Santa Rosa de Lima em Sergipe, intrigou-se ao descobrir que a evasão escolar por ali beirava os 43%. "Aquilo me deixou muito incomodado. Estava claro que os alunos não enxergavam na escola um papel de transformação social na vida deles"(6º Relatório Bimestral / Execução Orçamentária / MEC / 2020). Nota-se que uma boa caracterização da variável será determinante na elaboração do modelo, pois na maioria dos bancos de dados pesquisados em se tratando de escolas brasileiras de ensino médio, não existem registros por desconhecimento ou omissão. O uso de um banco de dados externo ao país, no caso Portugal, foi uma decisão didática para a elaboração e construção de uma ferramenta estatística consistente.

Dados da Pesquisa Nacional por Amostra de Domicílio Contínua, do Instituto Brasileiro de Geografia e Estatística (IBGE), mostram que somente no segundo semestre de 2021 o número de alunos entre 6 e 14 anos fora da escola era de 244 mil, acumulando um aumento de 171% em relação ao mesmo período de 2019. Por essa perspectiva, estamos falhando enquanto sociedade. Estima-se que um em cada quatro jovens de 15 a 17 anos não estão frequentando a escola, o que tem reflexos negativos em suas vidas e em áreas estratégicas para o país, como economia, saúde e segurança. As informações são do estudo "Políticas públicas para redução do abandono e evasão escolar de jovens", realizado pela Fundação Brava em parceria com Instituto Ayrton Senna, Inesper e Instituto Unibanco, (GATTÁS, 2018).

Em 2019, eram quase 6 milhões de analfabetos com 60 anos ou mais, o que equivale a uma taxa de analfabetismo de 18,0% para esse grupo etário. Ao incluir gradualmente os grupos etários mais novos, observa-se uma queda no analfabetismo: para 11,1% entre as pessoas com 40 anos ou mais, 7,9% entre aquelas com 25 anos ou mais e 6,6% entre a população de 15 anos ou mais, conforme observado nas Tabelas 1.1 e 1.2.

Tabela 1.1. Taxa de analfabetismo, grupo 15-60 anos, segundo as grandes regiões (%)

Regiões	2016	2017	2018	2019
Brasil	7,19	6,96	6,77	6,60
Norte	8,51	8,00	7,98	7,60
Nordeste	14,79	14,48	13,87	13,90



Colóquio Internacional Educação e Contemporaneidade

www.coloquioeducon.com
27 a 29 de setembro de 2023



Sudeste	3,76	3,51	3,47	3,30
Sul	3,61	3,52	3,63	3,30
Centro Oeste	5,71	5,23	5,40	4,90

Fonte: IBGE, Diretoria de Pesquisa, Coordenação de Trabalho e Rendimento, Pesquisa Nacional por Amostra de Domicílios Contínua 2016-2019

Tabela 1.2. Taxa de analfabetismo, grupo acima de 60 anos, nas grandes regiões (%)

Regiões	2016	2017	2018	2019
Brasil	20,36	19,21	18,59	18,00
Norte	29,95	27,39	27,02	25,50
Nordeste	39,78	38,65	36,87	37,20
Sudeste	11,64	10,57	10,33	9,70
Sul	11,31	10,86	10,80	9,50
Centro Oeste	21,18	18,96	18,27	16,60

Fonte: IBGE, Diretoria de Pesquisa, Coordenação de Trabalho e Rendimento, Pesquisa Nacional por Amostra de Domicílios Contínua 2016-2019



Colóquio Internacional Educação e Contemporaneidade

www.coloquioeducon.com
27 a 29 de setembro de 2023



Esses resultados indicam que as gerações mais novas estão tendo um maior acesso à educação e sendo alfabetizadas ainda enquanto crianças, veja na Tabela 2. Por outro lado, os analfabetos continuam concentrados entre os mais velhos e mudanças na taxa de analfabetismo para esse grupo se dão, em grande parte, devido às questões demográficas como, por exemplo, o envelhecimento da população.

Tabela 2. Pessoas de 18 a 24 anos/ por indicadores de educação / grandes regiões (%).



Colóquio Internacional Educação e Contemporaneidade

www.coloquioeducon.com
27 a 29 de setembro de 2023



Regiões	Taxa de escolarização	Taxa ajustada de frequência escolar líquida	Frequência escolar adequada	Atraso escolar dos estudantes	Não frequenta escola concluiu etapa	Não frequenta escola concluiu etapa
Brasil	32,40	25,50	21,40	11,00	4,10	
Norte	33,30	21,00	18,00	15,20	3,00	
Nordeste	32,00	19,50	17,00	15,00	2,50	
Sudeste	31,20	28,10	23,20	8,00	4,90	
Sul	34,60	30,60	25,90	8,80	4,80	
C. Oeste	35,00	31,10	25,40	9,60	5,70	
Homens	30,70	21,50	18,40	12,30	3,10	
Mulheres	34,20	29,70	24,50	9,70	5,10	
Branco	37,90	35,70	29,70	8,20	6,00	
Negro/Pardo	28,80	18,90	16,10	12,70	2,80	

Fonte: IBGE, Diretoria de Pesquisas, Coordenação de Trabalho e Rendimento, Pesquisa Nacional por Amostra de Domicílios Contínua 2019.

2.2 Variáveis preditoras significativas no contexto da evasão escolar

Dados disponibilizados por Barros (2021), estudos e evidências científicas nacionais e internacionais, relatam os motivos do abandono e da evasão escolar em três eixos: contexto, motivação e compreensão. O contexto refere-se a impedimentos e variáveis fora do âmbito escolar, dentre estes a limitação do acesso ao ensino médio em pequenas cidades, comunidades distritais e rurais. Neste mesmo contexto é possível citar entre outras, a atividade laboral desta faixa etária por necessidade de



Colóquio Internacional Educação e Contemporaneidade

www.coloquioeducon.com
27 a 29 de setembro de 2023



complementação da renda familiar, a distância do domicílio à escola e atributos da escola.

A motivação refere-se ao déficit de aprendizagem acumulada do ensino fundamental, falta de significado prático do currículo escolar, inexistência de flexibilidade e sensibilidade em relação as reais necessidades do estudante, acolhimento, violência urbana e orientação ocupacional de seus problemas. Gera-se portanto frustração de seu próprio desempenho, instabilidade emocional na família com reflexos no círculo social e no ambiente escolar. A compreensão refere-se a desinformação dos valores que a educação pode proporcionar, criando uma percepção equivocada por parte do estudante na sua visão limitada do uso do conhecimento e também pelo trabalhador docente ao retransmitir uma visão não contemplada na sua formação acadêmica. Com este enfoque, nesta divisão dos fatores leva-se à reflexão do modus operandi adequado para o enfrentamento da evasão escolar, ou seja a solução não concentra num único foco mas em múltiplas variáveis que demandam ações e estratégias diferenciadas.

Utilizando dados do Saeb de 2001, Soares e Mendonça (2003); Jesus e Laros (2004) verificaram que alunos com melhores condições socioeconômicas alcançaram melhores desempenhos quando comparados aos demais alunos. Os resultados do estudo de Alves e Soares (2008) corroboraram as conclusões dos trabalhos previamente citados. De acordo com os autores, estudos com dados de avaliação dos sistemas de ensino conduzidos no país, mostram que a maior parte da variação nos resultados escolares pode ser explicada por fatores extra escolares associados, principalmente à origem social dos alunos. Esses conhecimentos proporcionam e demandam desenvolvimento do pensamento, da linguagem, das capacidades humanas complexas, dos processos funcionais complexos, que somente podem ser apropriados e desenvolvidos pelos alunos quando orientados pelo professor, que deve dominar esses conhecimentos, seus signos, sua linguagem e os modos, adaptando a pedagogia, para orientá-los (MARTINS, 2013; SAVIANI, 2000).

A evasão escolar e a defasagem idade por série parecem estar diretamente relacionadas à necessidade de complementação da renda familiar. De acordo com o IBGE (2000) dos jovens de 15 anos de idade, apenas 16,53% estão na escola, enquanto 22% trabalham e estudam, 8% só estudam, 7% estudam e estão a procura de emprego e 10% não estudam. Dentre os problemas citados, merecem destaque a desnutrição, tanto da educação como da saúde, e as disfunções neurológicas, referidas por 92,5% dos 40 professores e 100% dos 19 profissionais de saúde relatados numa pesquisa (MOYSÈS e COLLARES, 1997). Por outro ângulo fatores indicados pela formação educacional e no âmbito da ocupação dos pais são determinantes, pois o não êxito profissional dos seus responsáveis, instabiliza a inserção social pela educação, inferindo que em todas as classes, acentuando-se nas menos favorecidas representem alta porcentagem entre os evadidos do ensino médio.



Colóquio Internacional Educação e Contemporaneidade

www.coloquioeducon.com
27 a 29 de setembro de 2023



A evasão escolar constitui situação extrema, resultante do fracasso das ações nos sistemas educacionais, preocupadas com monitoramento de dados sem sua devida análise efetiva, para composição de políticas públicas, como parâmetro em 2015 haviam 1.049.837 estudantes nessa situação. Ao considerar o VAAF (Valor Aluno / Ano / FUNDEB) e o VAAT (Valor Aluno / Ano / Total), nos anos de 2021 e 2022 suas novas regras de distribuição aplicadas nos termos da EC nº 108/2020, definem $VAAF(2021)=R\$4.462,83$, $VAAT(2022)=R\$5.643,92$ considerando 2.049 redes de ensino (NT nº 23_2022 Fundeb VAAF VAAT 2021 2022). Desta forma, em um cálculo simplista são 5,93 bilhões de reais aplicados sem efeito comprovado de minoração do problema da evasão escolar no país.

2.3. Modelo de regressão beta inflacionado

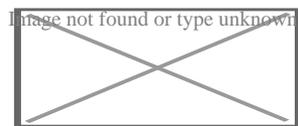
Simas et al. (2010) apresentam uma generalização do modelo de regressão beta, considerando uma estrutura de regressão para o parâmetro de precisão em modelos não-lineares. Também obtêm correções analíticas de viés para os estimadores de máxima verossimilhança, refazendo os resultados.

Uma discussão a respeito de modelagem pela regressão beta executada no R é revista por Cribari-Neto e Zeileis (2010), diz respeito aos modelos de regressão beta inflacionados, que acomodam dados que contêm zeros e/ou uns, são tratados como extensões do modelo de regressão beta (OSPINA et al. ,2006).

A função densidade beta é dada por:



em que $p > 0$ e $q > 0$ são parâmetros que indexam a distribuição e $\Gamma(p)$ é a função gama dada por:



A média e variância ficam assim definidas:

Para o modelo de regressão proposto por Ferrari e Cribari-Neto (2004), faz-se necessária uma reparametrização da densidade beta, sendo:

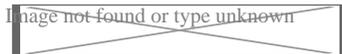


isto é:



Colóquio Internacional Educação e Contemporaneidade

www.coloquioeducon.com
27 a 29 de setembro de 2023



resultando que:



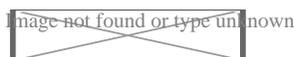
Assim μ representa a média da variável resposta e σ o parâmetro de precisão, se $\sigma = cte$, para valores maiores de μ resultarão valores menores da variância de y , implicando que a função densidade será dada por:



em que $0 < \sigma < 1$ e $\mu > 0$, dado que a esperança e a variância denotam por:



onde σ representa um parâmetro de precisão, com a função da variância:



A distribuição beta é adequada para modelar proporções e dependendo dos valores dos dois parâmetros que a indexam, a densidade assume formas bem variadas, acomodando distribuições simétricas, assimétricas, em formas de J e de J invertido (FERRARI & CRIBARI-NETO, 2004).

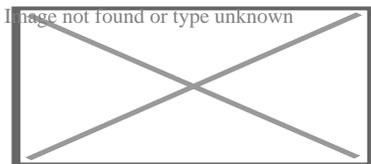


Colóquio Internacional Educação e Contemporaneidade

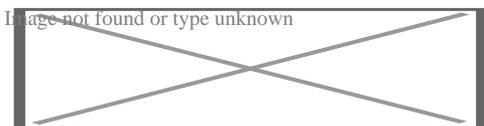
www.coloquioeducon.com
27 a 29 de setembro de 2023



Na construção do modelo de previsão para o fenômeno da evasão escolar, consideradas as variáveis com distribuições de probabilidades peculiares, os dois modelos, beta e beta inflacionado foram considerados e avaliados. A combinação da presença de covariáveis qualitativas associadas as quantitativas no modelo e a necessidade de transformação para contornar a não normalidade dos dados definiram quatro grupos de fatores para modelagem. Utilizou-se a Análise Fatorial Confirmatória, avaliada nos quatro grupos de similaridade. Para que a interpolação em porcentagens produza previsões confiáveis e represente a real variabilidade, a modelagem deve ser realizada com critérios estatísticos objetivos (STEEL et al., 2005). Ferrari & Cribari-Neto (2004) destacam que no modelo de regressão beta, para amostras grandes e sob condições de regularidade, os parâmetros não são ortogonais, resultam em aproximações satisfatórias para os estimadores de verossimilhança. Desta forma:



Da propriedade das inversas das matrizes particionadas, de acordo com RAO(1973), é obtida a inversa da matriz de Fisher, tal que:



2.4. Análise das variáveis significativas para a evasão escolar

Para a busca de um banco de dados na realidade escolar brasileira, observou-se que não faz parte da rotina das escolas nacionais, a organização de dados com as características necessárias para o estudo e modelagem do fenômeno. Encontram-se registros prioritários de notas, conceitos e frequência do aluno, sendo os demais registros executados de forma esporádica e sem conformação de um banco de dados.

Não foram encontrados dados que registram indicadores quanto a família do estudante, tempo de deslocamento, estado de saúde e parâmetros de avaliação institucional. Após a busca em diversos repositórios disponíveis na área educacional; foi encontrada uma base de dados localizada em (<https://archive.ics.uci.edu/ml/datasets/student+performance>).



Colóquio Internacional Educação e Contemporaneidade

www.coloquioeducon.com
27 a 29 de setembro de 2023



Executando uma primeira seleção das variáveis, pelo critério de relevância de acordo com os pressupostos obtidos nas referências teóricas, das trinta e três variáveis do arquivo original, foram consideradas as variáveis relacionadas no Quadro 1.

Quadro 1. Detalhamento das variáveis do banco de dados

Descrição	Sigla /SiglaOriginal	Tipo	Valores
Núm. de faltas	NFE / absences	Quantitativa discreta	Mín:0 a máx:(CH)
Idade do estudante	IDAD / age	Quantitativa discreta	Mín a máx
Núm. de falhas de aulas	NFAA / failures	Qualitativa binária	1<=3 ou >3
Relação familiar	RFAM / famrel	Qualitativa categ	1: MRU ;2: RU ;3: REG ;4: BOM ; 5: EXC
Tamanho da família	TFAM / famsize	Qualitativa binária	1,2,3: LE3 \ >3: GT3
Grau de instrução do pai	GIP / Fedu	Qualitativa categ	0: SI ; 1: FI ; 2: FC ;3: MED ; 4: SUP
Tempo livre após ativ. esc.	TLAE / freetime	Qualitativa categ	1: MBX; 2: BAX; 3: MED ;4: ALT; 5: MAL
Estado atual de saúde	EASE /health	Qualitativa categ	1: RU ;2: REG ;3: BOM ;4: MBOM ;5: EXC
Grau de instrução da mãe	GIM / Medu	Qualitativa categ	0: SI ; 1: FI ; 2: FC ;3: MED ; 4: SUP
Sexo do estudante	sex	Qualitativa binária	M: masc. \ F: fem.
Carga horária semanal	CHS / studytime	Qualitativa (horas)	1: [1,2) ;2: [2,5) ;3:[5,10);4: [>10)



Desloc. domicilio/escola

TDDE / traveltime

Qualitativa (min.)

1: (<15] ;2: (15,30] ;3:(30,60] ; 4: (>60)

Fonte: Paulo Cortez, University of Minho, Guimarães, Portugal, <http://www3.dsi.uminho.pt/pcortez> * Modificações realizadas pelo autor.

2.4.1. Análise descritiva / Deslocamento do estudante do domicilio até a escola

Como variável importante no contexto da evasão escolar, o tempo de deslocamento foi auferido em minutos e agrupado conforme observa-se nos parâmetros descritos na Tabela 3. Nela está um esboço da distribuição deste tempo por sexo e por escola.

Tabela 3. Tempo de deslocamento / Domicilio até a escola / em minutos.

Grupo		Observações				Proporções			
		1	2	3	4	1	2	3	4
Parâmetros		[<15)	[15,30)	[30,60)	[>60)	[<15)	[15,30)	[30,60)	[>60)
GP	Masc	161	60	12	4	0,248	0,092	0,018	0,006
	Fem	122	49	8	7	0,188	0,076	0,012	0,011
MS	Masc	57	69	17	3	0,084	0,101	0,026	0,004
	Fem	26	35	17	2	0,046	0,056	0,026	0,006
	Totais	366	213	54	16	0,570	0,334	0,082	0,030

Fonte: Dados organizados pelo autor by <https://archive.ics.uci.edu/ml/datasets/Student+Performance> Paulo Cortez, University of Minho, Guimarães, Portugal, <http://www3.dsi.uminho.pt/pcortez>

Pela análise do Figura 1, verifica-se comportamento distinto entre estes tempos por escola, indicio de que esta variável poderá interferir no índice de evasão. Dado que a utilização do tempo é fator preponderante para o desempenho e permanência do estudante no sistema escolar, a Figura 1



Colóquio Internacional Educação e Contemporaneidade

www.coloquioeducon.com
27 a 29 de setembro de 2023



evidencia que o deslocamento diário, afeta estudantes das faixas etárias menores em sua maioria um tempo de deslocamento abaixo de 60 minutos. O fato é relevante pois sistematicamente maiores distâncias indicam maiores ocorrências tais como inexistência ou imprevisto no transporte gratuito ou público, atraso na entrada do horário da classe.

Figura 1. Quantidade de estudantes pelo tempo de deslocamento / Domicílio até escola.

F onte: Dados organizados pelo autor by <https://archive.ics.uci.edu/ml/datasets/Student+Performance>

1: <15min , 2: 15 a 30min , 3: 30 a 60min , 4: >60min

Desta forma em grandes centros urbanos é justificável o zoneamento da clientela estudantil, contornável no ensino diurno, mas que apresenta em muitos casos dificuldade desta prática no ensino noturno, principalmente na modalidade profissionalizante. Análise descritiva semelhante aplicada nas demais variáveis definidas com base nos pressupostos teóricos da área educacional foram executadas, para uma justificativa consistente do resultado final do modelo de regressão beta inflacionado.

CONSIDERAÇÕES FINAIS

3. Conclusão

Este trabalho tem como eixo central apontar quais são os principais fatores que influenciam o abandono escolar no ensino médio e propor um modelo de previsão para a taxa de incidência. Soares et al (2015), utilizou dados baseados em uma série histórica construída a partir de registros da PNAD com pesquisa realizada no estado de Minas Gerais, coletando informações de 3.418 entrevistados. Estimaram-se modelos de regressão de risco proporcional de Cox e um modelo exponencial logito , permitindo verificar a taxa do não abandono na trajetória do ensino médio. Entre os resultados encontrados, destacam-se alguns fatores expressivos na explicação do abandono, tais como: a dificuldade nas disciplinas, ânsia por uma escola diferente, percepção de melhores oportunidades de trabalho com a continuidade dos estudos e a importância atribuída na escolha à escola.

Com base neste pressuposto, as análises realizadas indicaram a adoção de um modelo de regressão beta inflacionado em zero, obtendo melhor qualificação nos testes de comparação aplicados, ou sejam: Critério de Informação de Akaike (AIC = -788.2006), Critério de Informação de Bayes (BIC = -627.0850), Correlação (cor = 0.2225068 > 0.2172773 /Ajustado), Likelihood ratio test [lrtest.default =



Colóquio Internacional Educação e Contemporaneidade

www.coloquioeducon.com
27 a 29 de setembro de 2023



$Pr(>Chisq) = 8.629e-07$ (significativo)] e teste de Breusch-Pagan, $BP = 33.796$, $df = 11$, $p\text{-value} = 0.0003906 < 0,05$.

No teste de Durbin-Watson o modelo beta mostrou mais variáveis significativas que o modelo beta inflacionado em zero, como todas as variáveis mostraram algum nível de explicação para o modelo, rejeitou-se o modelo beta pelo resultado dos critérios citados anteriores, principalmente AIC e BIC que representam a variância não explicada. Considerando as observações verificadas na análise gráfica, o modelo beta inflacionado em zero, imputou melhor ajuste para os resíduos padronizados (Ver Figura 2).

Figura 2. Análise dos resíduos considerando as variáveis significativas.



Colóquio Internacional Educação e Contemporaneidade

www.coloquioeducon.com
27 a 29 de setembro de 2023



Figura 3. Análise gráfica com com controle para a Escola Gabriel Pereira



Colóquio Internacional Educação e Contemporaneidade

www.coloquioeducon.com
27 a 29 de setembro de 2023



Fonte: Dados organizados pelo autor by <https://archive.ics.uci.edu/ml/datasets/Student+Performance>

Figura 4. Controle considerando os dados da escola Mousinho Silveira

Fonte: Dados organizados pelo autor by <https://archive.ics.uci.edu/ml/datasets/Student+Performance>



Colóquio Internacional Educação e Contemporaneidade

www.coloquioeducon.com
27 a 29 de setembro de 2023



A pesquisa neste trabalho reporta a um instrumento que permita uma atitude de prevenção em relação ao abandono do estudante dadas as características do sistema educacional, desta forma as variáveis explicativas inputaram um novo olhar na elaboração do modelo de previsão. De acordo com Colombo (2010), os grupos de estudantes mais vulneráveis alijados do meio escolar, não se identificam numa escola alheia aos seus interesses e neste contexto os gestores carecem de ferramentas para reverter o processo de abandono, cujas variáveis comportamentais são complexas.

As diferenças entre as realidades em que as escolas estão inseridas, fundamentam-se sobre as características dos estudantes em risco de evasão ou que estão fora do sistema escolar evidenciam tipologias diversas. No relatório europeu, resultante do *Conseil de l'Emploi des Revenus et de la Cohésion Sociale*, 2008, a condição socioeconômica da família foi indicada como um alto fator de risco: jovens imigrantes, monoparentais, de grupos étnicos minoritários ou com pais desempregados por longo período de tempo têm maior probabilidade de abandonar a escola ou de se transformar em alguém que está fora da escola e do trabalho. O mesmo, segundo o relatório, vale para os jovens que vivem em bairros de risco ou em áreas economicamente desfavorecidas. Existem também aqueles que abandonam a escola porque os estudos competem com a atratividade do mercado de trabalho, nas áreas mais produtivas, e que veem no trabalho uma perspectiva imediata mais gratificante do que a dos estudos (MARGIOTTA et al, 2014).

Figura 5. Gráfico envelope meio normal dos resíduos para o modelo beta inflacionado em zero, para os dados das escolas Gabriel Pereira e Mousinho da Silveira.



Colóquio Internacional Educação e Contemporaneidade

www.coloquioeducon.com
27 a 29 de setembro de 2023



Fonte: Dados organizados pelo autor by <https://archive.ics.uci.edu/ml/datasets/Student+Performance>

Da Figura 5, os estudantes identificados como E144 e E315, são dois casos que acusam certo grau de atenção pois estão numa zona de não conformidade no ajuste residual do modelo de previsão, o detalhe do gráfico tem a explicação dada pela característica de suas variáveis, vejamos:

a) O estudante E144 é da escola *Gabriel Pereira*, do sexo feminino, com 16 anos, tem domicílio urbano, o *grupo familiar tem acima de 3 pessoas*, o pai e a mãe tem *grau de instrução fundamental completo*, a *guarda é da mãe*, utiliza até 15 minutos para o deslocamento até a escola, *estuda entre 2 e 5 horas semanais*, tem *1 reprovação anterior*, tem acesso a internet, tem uma regular relação familiar, tem um *tempo livre após as aulas considerado baixo*, apresenta um *estado de saúde excelente*, registra *12 faltas num total de 93 aulas previstas* e as notas no período *8, 10 e 10* em escala de *0 a 20*.

b) O estudante E315 é da escola *Gabriel Pereira*, do sexo masculino, com 17 anos, tem domicílio rural, o *grupo familiar tem acima de 3 pessoas*, o *pai não tem grau de instrução* e a *mãe tem grau de instrução fundamental incompleto*, a *guarda é da mãe*, utiliza até 15 minutos para o deslocamento até a escola, *estuda entre 2 e 5 horas semanais*, *não tem reprovação anterior*, tem acesso a internet, *tem regular relação familiar*, tem um *tempo livre após as aulas considerado baixo*, apresenta um *estado de saúde ruim*, registra *2 faltas num total de 93 aulas previstas* e as notas no período *16, 17 e 18* em escala de *0 a 20*.

A Figura 3 apresenta um diagnóstico da situação do índice G3 (média final) para a escola Gabriel Pereira, no qual há evidências de problemas com o rendimento escolar de 79 estudantes abaixo do limite crítico (LCL=7,00) para um valor central da escola em 11,90, desvio padrão de 1,63. Entre os limites (LCL=7,00 ; CL=11,90) para efeito da evasão considera-se um estudante em estado de alerta.

A Figura 4 apresenta um diagnóstico da situação do índice G3 (média final) para a escola Mousinho da Silveira, no qual há evidências de problemas com o rendimento escolar de 16 estudantes abaixo do limite crítico (LCL=3,68) para um valor central da escola em 11,90, desvio padrão de 2,74. Entre os limites (LCL=3,68 ; CL=11,90) para efeito da evasão considera-se um estudante em estado de alerta. Conclui-se que de fato o modelo é eficiente para mostrar que em diferentes escolas, os resultados acompanham o contexto das variáveis explicativas



Colóquio Internacional Educação e Contemporaneidade

www.coloquioeducon.com
27 a 29 de setembro de 2023



Do ponto de vista comparativo há que ater-se ao cuidado de que uma comparação visual entre escolas pode ser enganoso, pois observa-se que a escola Gabriel Pereira embora tenha maior número crítico maior que a escola Mousinho da Silveira ($79 > 16$), a mesma tem menor variância ($1,63 < 2,74$) e o limite crítico superior ($7,00 > 3,68$).

A representação estatística está detalhada, e define como:

Image not found or type unknown

Image not found or type unknown

Consideradas as análises e observações auferidas, evidencia-se que o modelo de regressão beta inflacionado em zero, é viável para a obtenção do índice de previsibilidade da evasão escolar.

REFERÊNCIAS

4. Referências

ABRAMOVAY, Miriam; **CASTRO**, Mary Garcia & **WASELFISZ**, Júlio Jacobo. Juventudes na escola, sentidos e buscas: Por que frequentam?

ALVES, M. T. G.; **SOARES**, J. F. O efeito das escolas no aprendizado dos alunos: um estudo com dados longitudinais no ensino fundamental. *Educação e Pesquisa*, São Paulo, v. 34, n. 3, p. 527-544, 2008. DOI: <https://doi.org/bvc6px>

BARROS, Alessandra Menezes da Cunha et all. Fatores associados a retenção e intencionalidade de evasão nos cursos de Farmácia de uma universidade pública do nordeste brasileiro . <https://doi.org/10.1590/SciELOPreprints.3078>

BAYES, M. P. M. An essay towards solving a problem in the doctrine of chances. by the late rev. mr. bayes, f. r. s. communicated by mr. price, in a letter to john canton, a. m. f. r. s. *Philosophical Transactions* (1683-1775), The Royal Society, v. 53, p. 370–418, 1763. ISSN 02607085. 2011



Colóquio Internacional Educação e Contemporaneidade

www.coloquioeducon.com
27 a 29 de setembro de 2023



COLOMBO, M. Dispersione scolastica e politiche per il successo formativo: dalla ricerca sugli early school leaver alle proposte di innovazione. Trento. Erickson. 2010

COELHO, Francisco Antonio, ANDRADE, Junior Jairo Eduardo Borges .Uso do conceito de aprendizagem em estudos relacionados ao trabalho e organizações. Universidade de Brasília, Brasília-DF, Brasil. 2008

CRIBARI-NETO, F.; ZARKOS, S. G. Heteroskedasticity-consistent covariance matrix estimation: White's estimator and the bootstrap. Journal of Statistical Computation and Simulation, 68:391–411.2001

CRIBARI-NETO, F; GÓIS, Matheus Cabral de Araújo. Sumário. Revista Brasileira de Economia, Volume: 56, Número: 2, Publicado.2002

CRIBARI-NETO, F.; ZELEIS, A.; Beta regression in R. Journal of Statistical software 34(2), 1-24. 2010

FERRARI, S. L. P., CRIBARI-NETO, F. Beta regression for modeling rates and proportions. Journal of Applied Statistics software 31, 799-815. 2004

GATTÁS, Maria Lúcia Borges, FUREGATO, Antonia Regina Ferreira. Interdisciplinaridade: uma contextualização. Universidade de Uberaba, Uberaba, MG, Brasil. 2018

JESUS, G. R.; LAROS, J. A. Eficácia escolar: regressão multinível com dados de avaliação em larga escala. *Avaliação Psicológica*, Campinas, v. 3, n. 2, p. 93-106, 2004.

LACRUZ, Adonai José; AMÉRICO, Bruno Luiz & CARNIEL, Fagner. Indicadores de qualidade na educação: Análise discriminante dos desempenhos na Prova Brasil

<http://dx.doi.org/10.1590/S1413-24782019240002>. Revista Brasileira de Educação v. 24 e240002 2019



Colóquio Internacional Educação e Contemporaneidade

www.coloquioeducon.com
27 a 29 de setembro de 2023



MARGIOTTA, Umberto; **VITALE**, Gabriella; **SANTOS**, Jácia Soares dos. O FENÔMENO DO ABANDONO ESCOLAR NA EUROPA DO NOVO MILÊNIO: Dados, políticas, intervenções e perspectivas. Cad. Cedes, Campinas, v. 34, n. 94, p. 349-366, set.-dez., 2014

MARTINS, L. M. O desenvolvimento do psiquismo e a educação escolar: contribuições à luz da psicologia histórico-cultural e da pedagogia históricocrítica. Campinas: Autores Associados, 2013.

MENDES, Marcelo Simões. Da inclusão à evasão escolar: o papel da motivação no ensino médio. Estudos de Psicologia I Campinas I 30(2) I 261-265 I abril - junho 2013

MOYSÉS, Maria Aparecida Affonso; **COLLARES**, Cecília Azevedo Lima. **Inteligência Abstraída, Crianças Silenciadas: as Avaliações de Inteligência** Psicologia e Razão Instrumental. Psicologia. USP8(1).1997

OSPINA, R.; CRIBARI-NETO, F.; VASCONCELOS, K. L. P.; Improved point and interval for a beta regression model. Computational Statistics and Data Analysis 51, 960-981. 2006

RAO, C. R. Linear Statistical Inference and its Applications. 2nd. New York: Wiley, p.33.1973

SAVIANI, D. Pedagogia histórico-crítica: primeiras aproximações. 7. ed. Campinas: Autores Associados, 2000.

SIMAS, A. B.; **BARRETO-SOUZA**, W.; **ROCHA**, A. V. Improved estimators for a general class of beta regression models. Computational Statistics and Data Analysis 54, 348-366. 2010

SOARES, T. M. Influência do professor e do ambiente em sala de aula sobre a proficiência alcançada pelos alunos avaliados no Simave-2002. *Estudos em Avaliação Educacional*, São Paulo, n. 28, 2003. DOI: <https://doi.org/gxtg>

SOARES, T. M., **FERNANDES**, Neimar da Silva, **NÓBREGA**, Mariana Calife,



Colóquio Internacional Educação e Contemporaneidade

www.coloquioeducon.com
27 a 29 de setembro de 2023



SOUSA, Carolina Rodrigues de Oliveira; **GOMES**, Keila Rejane Oliveira & outros. Fatores preditores da evasão escolar entre adolescentes com experiência de gravidez. Cad. Saúde Colet., 2018, Rio de Janeiro, 26 (2): 160-169. 167

STEEL, R.G.D.; **TORRIE**, J.H.; **DICKEY**, D.A. Principles and procedures of statistics. Nova York: McGraw Hill Books, 3rd ed. 666p. 1997