

Métodos não-paramétricos em análise de sobrevivência - uma aplicação a dados de evasão

Ana Carolina Andrade Gonçalves¹
Graziela Dutra Rocha Gouvêa²

Introdução

Sabe-se que a evasão nas universidades são problemas complexos, decorrentes de várias causas e que devem ser colocados em discussão de modo permanente, para solucioná-los ou evitar que esses eventos continuem aumentando. O Instituto de Ciências Exatas e Biológicas (ICEB) da Universidade Federal de Ouro Preto (UFOP) possui a maior taxa de evasão desta universidade, alcançando um índice de 37% divulgado pela Pró-Reitoria de Graduação recentemente.

Este projeto tem como objetivo estudar os fatores que influenciam no tempo até a evasão/trancamento dos alunos dos cursos oferecidos pelo ICEB/UFOP de modo que tais resultados possam ajudar a todos envolvidos com o desenvolvimento da instituição a solucionar ou pelo menos minimizar tal problema.

Metodologia

Para analisar o tempo até a evasão dos alunos dos cursos oferecidos pelo ICEB/UFOP foram utilizadas técnicas não paramétricas em Análise de Sobrevivência.

A variável resposta neste estudo é o tempo até a evasão/trancamento do estudante em horas cursadas. Os estudantes que não evadiram ou trancaram sua matrícula forma considerados dados censurados. As covariáveis analisadas foram: sexo (feminino ou masculino), modo de admissão (vestibular, transferência ou obtenção de novo título), turnos (noturno, vespertino ou integral), raça (branca, negra ou outras) e estado de origem dos estudantes (Minas Gerais ou outra origem).

Foram geradas curvas de Sobrevivência por meio do Estimador de Kaplan-Meier (Kaplan e Meier, 1958) e também por meio do Estimador de Nelson-Aalen (Nelson, 1972 e Aalen, 1978). Foi utilizado o teste de log-rank e o teste de Wilcoxon para testar a hipótese nula de igualdade das curvas de sobrevivência (Colosimo e Giolo, 2006).

Resultados e Discussões

Os resultados preliminares do estudo sobre o tempo até a evasão via métodos não paramétricos em Análise de Sobrevivência, utilizando o Estimador de Kaplan-Meier e o Estimador de Nelson-Aalen para a função de sobrevivência, apontam que o tempo médio até a evasão/trancamento é de 1961,7 horas cursadas e que o tempo mediano é de 1015,28 horas cursadas. A Figura 1 apresenta a curva de Sobrevivência estimada via Kaplan-Meier.

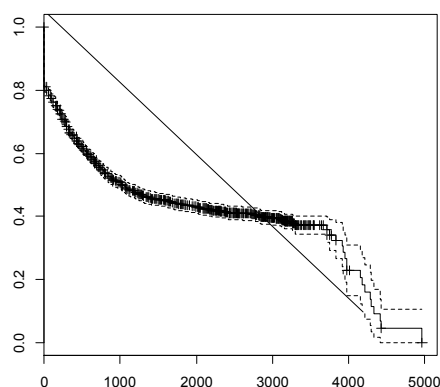


Figura 1: Curva de Sobrevivência estimada via Kaplan-Meier

A Figura 2 apresenta as curvas de sobrevivência estimadas para as covariáveis sexo e modo de admissão.

¹Estudante do curso de Bacharelado em Estatística da Universidade Federal de Ouro Preto

²Professora Adjunta da Universidade Federal de Ouro Preto - UFOP

Quanto ao sexo, verificou-se que estudantes do sexo feminino sobrevivem mais à evasão/trancamento do que os do sexo masculino. Os p-valores resultantes dos testes de log-rank e Wilcoxon apresentaram valores $p < 0,00$.

Não foram encontradas diferenças significativas entre as curvas de sobrevivência referentes ao modo de admissão (vestibular ou outra), uma vez que as curvas de sobrevivência se cruzam. Além disso, os p-valores resultantes dos testes de log-rank e Wilcoxon apresentaram valores $p = 0,68$ e $p = 0,72$ respectivamente.

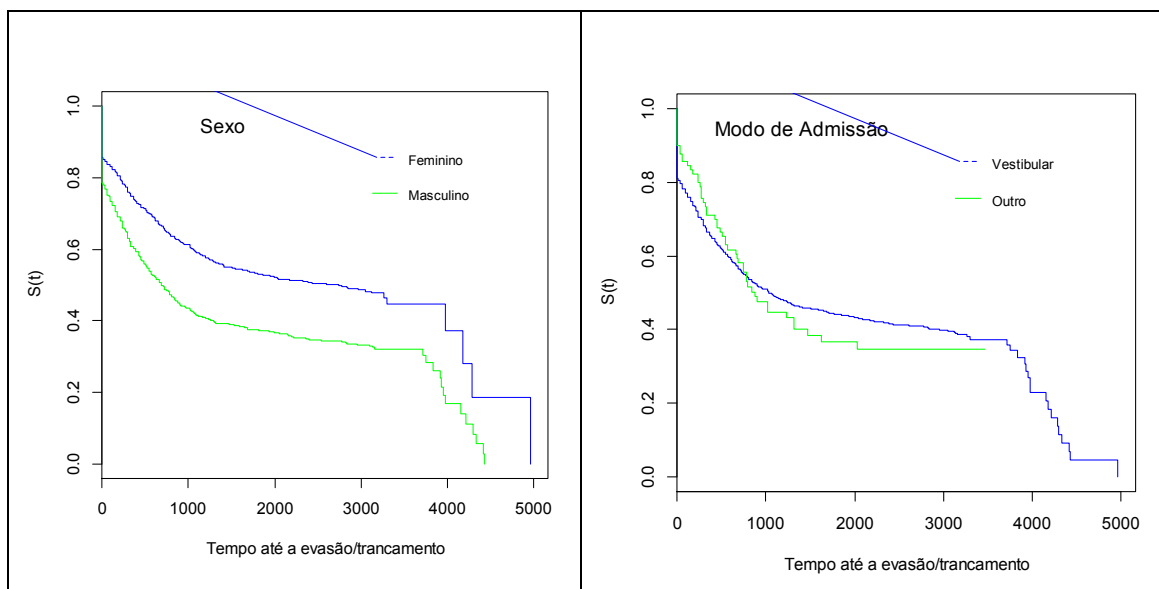


Figura 2: Comparação das Curvas de Sobrevivência para as covariáveis sexo e modo de admissão

A Figura 3 apresenta as curvas de sobrevivência estimadas para as covariáveis turno e estado de origem do estudante.

Não foram verificadas diferenças entre os tempos de ocorrência de evento para alunos dos turnos noturno, vespertino ou integral. Além disso, os p-valores resultantes dos testes de log-rank e Wilcoxon apresentaram valores $p = 0,70$ e $p = 0,84$ respectivamente.

Quanto ao estado de origem, estudantes mineiros evadem mais rapidamente do que aqueles de outros estados ou país. Os p-valores resultantes dos testes de log-rank e Wilcoxon apresentaram valores $p < 0,002$.

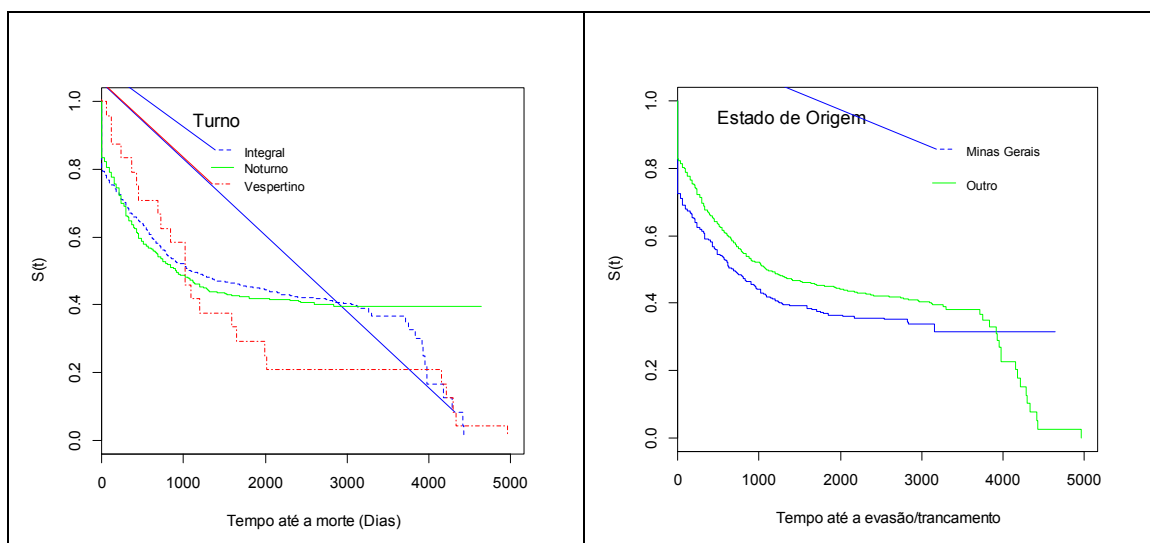
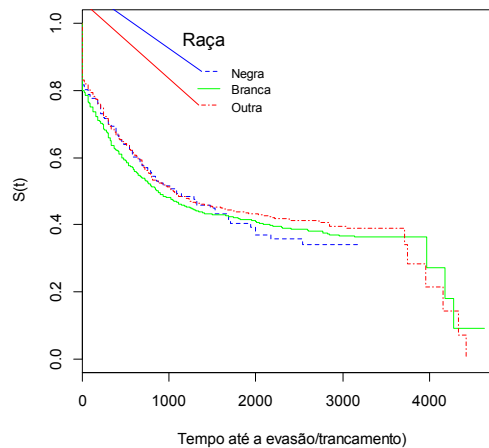


Figura 3: Comparação das Curvas de Sobrevivência para as covariáveis turno e estado de origem

A Figura 3 apresenta as curvas de sobrevivência estimadas para a covariável raça.

Quanto à raça declarada, também não foram encontradas diferenças no tempo até a evasão/trancamento por meio das curvas de sobrevivência estimadas. Pode-se observar que as curvas se cruzam. Além disso, os p-valores resultantes dos testes de log-rank e Wilcoxon apresentaram valores $p = 0,33$ e $p = 0,16$ respectivamente.



Todas as análises foram realizadas no software R.

Referências Bibliográficas

- [1] AALEM, O. O. Nonparametric Inference for a Family of Counting Processes. **Annals of Statistics**, 6, 701-726. 1978.
- [2] COLOSIMO, E. A.; GIOLO, S. R. **Análise de Sobrevivência Aplicada**. São Paulo: Edgard Blücher, 2006. 369p.
- [3] KAPLAN, E. L. ; MEIER, P. Nonparametric estimation from incomplete observations. **Journal of the American Statistical Association**, 53, 457-81. 1958.
- [4] NELSON, W. Theory and Applications of Hazard Plotting for Censored Failure Data. **Technometrics**, 14, 945-965. 1972