



XXXIII SIC SALÃO INICIAÇÃO CIENTÍFICA

Evento	Salão UFRGS 2021: SIC - XXXIII SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS
Ano	2021
Local	Virtual
Título	Modelos para dados de Sobrevida com Censura Informativa - Comparação dos Softwares R e Python
Autor	NICOLAS MATHIAS HAHN
Orientador	SILVANA SCHNEIDER

Modelos para dados de sobrevivência com censura informativa - comparando softwares R e Python

Bolsista: Nicolas Mathias Hahn
Orientado por Prof^a. Dr^a. Silvana Schneider
UFRGS

01 de setembro de 2020 a 05 de março de 2021

Resumo

O presente trabalho visa a desenvolver um modelo de Análise de Sobrevivência que seja capaz de lidar com dados censura informativa e comparar, por meio do estudo de simulações, a acurácia e a velocidade dos softwares R e Python na estimação dos parâmetros. A presença de censura é comum em dados de sobrevivência, podendo ser informativa ou não. Caso utilize-se um modelo de sobrevivência que não assume que a censura é informativa e ela for, as estimações podem ficar viesadas. Devido a isso, é interessante averiguar se a censura é informativa ou não. Após a revisão teórica (definições de sobrevivência, censura, modelo de Cox, fragilidade compartilhada, distribuições de probabilidade Weibull e Gama), foram realizados os cálculos necessários para desenvolver o modelo e encontrar sua função de verossimilhança, fundamental para as estimações. Posteriormente, com o intuito de avaliar se o modelo proposto era capaz de lidar com os problemas aos quais fora criado - bem como para comparar os softwares de interesse -, foram feitas implementações computacionais utilizando os softwares R e Python. Resultados promissores nas simulações: ambos os softwares estimaram satisfatoriamente os parâmetros de interesse com boa precisão. No entanto, foi identificada uma diferença significativa nos entre os tempos de execução de cada software, sendo o R claramente mais rápido que o Python nas simulações realizadas neste estudo. Portanto, extensões futuras da pesquisa podem englobar a otimização das sintaxes, comparação de outras linguagens e o porquê da diferença expressiva nos tempos de execução.