

# A pensar

Alisson Rosa

## Resumo

teste de abstrac

```
import pandas as pd
import numpy as np
from lifelines import KaplanMeierFitter
from lifelines import CoxPHFitter
import matplotlib.pyplot as plt
plt.rcParams["figure.figsize"] = (13,5)
```

## 1 Introdução

As Crônicas de Gelo e Fogo (no original em inglês: A Song of Ice and Fire) é uma série de livros de fantasia épica escrita pelo romancista e roteirista norte-americano George R. R. Martin e publicada pela editora Bantam Spectra. Martin começou a desenvolvê-la em 1991 e o primeiro volume foi lançado em 1996. Originalmente concebida para ser uma trilogia, a saga agora consiste em cinco volumes publicados, com mais dois planejados. Há também três contos derivados e algumas novelas que consistem de resumos dos romances principais.

Há três argumentos principais na história, que se tornam cada vez mais interligados: a crônica de uma guerra civil dinástica entre várias famílias concorrentes pelo controle dos Sete Reinos; a ameaça crescente das criaturas sobrenaturais conhecidas como os Outros, que habitam além de uma imensa muralha de gelo ao Norte; e a ambição de Daenerys Targaryen, a filha exilada de um rei assassinado em uma outra guerra civil treze anos antes, prestes a voltar à sua terra e reivindicar seu trono de direito.

A série foi traduzida para quase 50 idiomas e vendeu mais de 90 milhões de exemplares mundialmente (até abril de 2019).[1] Todos os romances foram, em geral, muito bem recebidos pela crítica literária e pelo público, sendo indicados a diversos prêmios de fantasia e ficção científica, como o Prêmio Locus, o Prêmio Nebula e o Prêmio Hugo. No Brasil os livros são publicados pela editora Suma (anteriormente pela LeYa) e em Portugal pela Saída de Emergência.

Os livros de As Crônicas de Gelo e Fogo foram adaptados para diversos formatos, como jogos de videogame, histórias em quadrinhos, bonecos em miniatura e uma série de TV intitulada Game of Thrones. A atração televisiva apresentou a saga a um maior número de leitores e lhe trouxe maior notoriedade, fazendo com que os quatro primeiros volumes da série surgissem entre os dez primeiros colocados na referencial lista de mais vendidos do jornal norte-americano The New York Times em 2011.

Vamos nesse breve ensaio fazer uma análise de sobrevivência estatística dos personagens dos livros.

## 2 Análise de Sobrevivência

A análise de sobrevivência difere dos métodos estatísticos clássicos pois temos dados censurados, que são observações a qual o evento de interesse por algum motivo não pode ser medido. Aqui então o nosso evento de interesse é o tempo para morte de um personagem, e a seu respectivo tempo de censura será o último livro que apareceu. Sendo assim nossa variável tempo ( $T$ ) possui somente 5 valores possíveis, que são respectivamente a numeração dos livros.

### 3 Metodologia da criação do banco de dados

As crônicas de gelo e fogo possui uma quantidade enorme de personagens, assim fica claro que o número de personagens que morre é muito menor que os que não morrem, portanto foi removido da análise todos os personagens que não estavam associados a alguma família relevante.

### 4 Análise Básica

Como dito antes G possui muitos personagens, e também muitas associações, assim é interessante avaliarmos a quantidade de indivíduos em cada família, como também a quantidade de mortos respectivamente

Notamos portanto que proporção de personagens mortos segue a proporção de quantidade de membros da família, assim é interessante juntar algumas famílias que possuem poucos membros em um outro grupo para análise ser mais focada, assim vamos agrupar a casa Martell, Tully, Tyrell, Arryn em um grupo chamado de “outros”, vale ressaltar que os Targaryen também possuem poucos membros, porém são peça chave da obra.

### 5 Estimando quantidades básicas

A função de sobrevivência ( $S$ ) nos informa qual a probabilidade de uma observação vir a ter o evento de interesse a partir do tempo  $t$ , assim aqui podemos avaliar a função de sobrevivência estratificada por grupos, avaliando tanto em termos visuais e inferências a possibilidade de existir diferença na  $S$

#### 5.1 Estimando a sobrevivência

Sabemos que a proporção de personagens do sexo feminino é de apenas aproximadamente 16%, assim uma pergunta a se fazer é se existe diferença na sobrevivência entre os sexos

Assim o que a figura NN nos indica é que existe diferença no tempo de sobrevivência entre os sexos, pois a curva estimada de  $S$  para o sexo feminino está pontualmente superior que a do sexo masculino para todos os pontos, e mesmo em termos de intervalos de confiança existe pouca interseção.

Faz-se então importante realizarmos um teste de hipótese para avaliar tal afirmação, aqui vamos utilizar o teste log rank, onde em termos gerais possui a seguinte hipótese

$$H_0 : \text{As funções de sobrevivências dos grupos são iguais.}$$

```
df = pd.read_csv("data/gotsurv.csv")
def kme_groups(df, variable, groups):
    with plt.style.context('seaborn-bright'):
        data = df[df[variable].isin(groups)]
        ax = plt.subplot(111)

        kmf = KaplanMeierFitter()

        for name, grouped_df in data.groupby(variable):
            kmf.fit(grouped_df["time_death"], grouped_df["status"], label=name)
            kmf.plot_survival_function(ax=ax)
            plt.xlabel('Tempo')
            plt.ylabel("Sobrevivência estimada")

kme_groups(df, 'Gender', [1, 0])
```

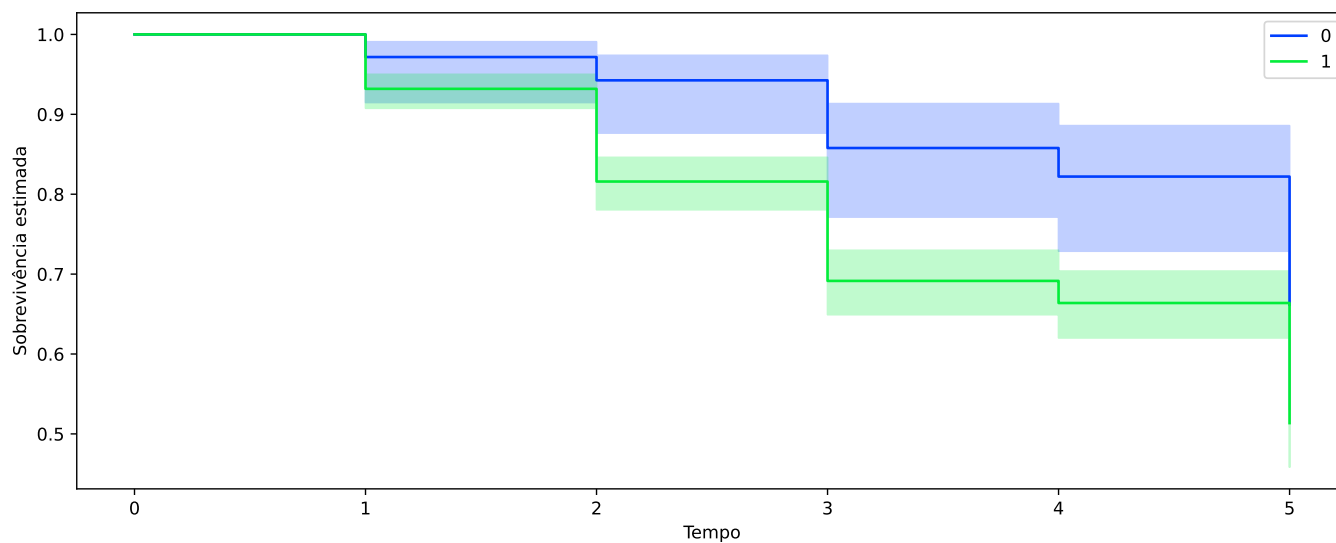


Figura 1: Função de Sobrevivência para ambos os sexos

```
kme_groups(df, 'Allegiances', ['House Lannister', 'House Greyjoy', 'House Stark'])
```

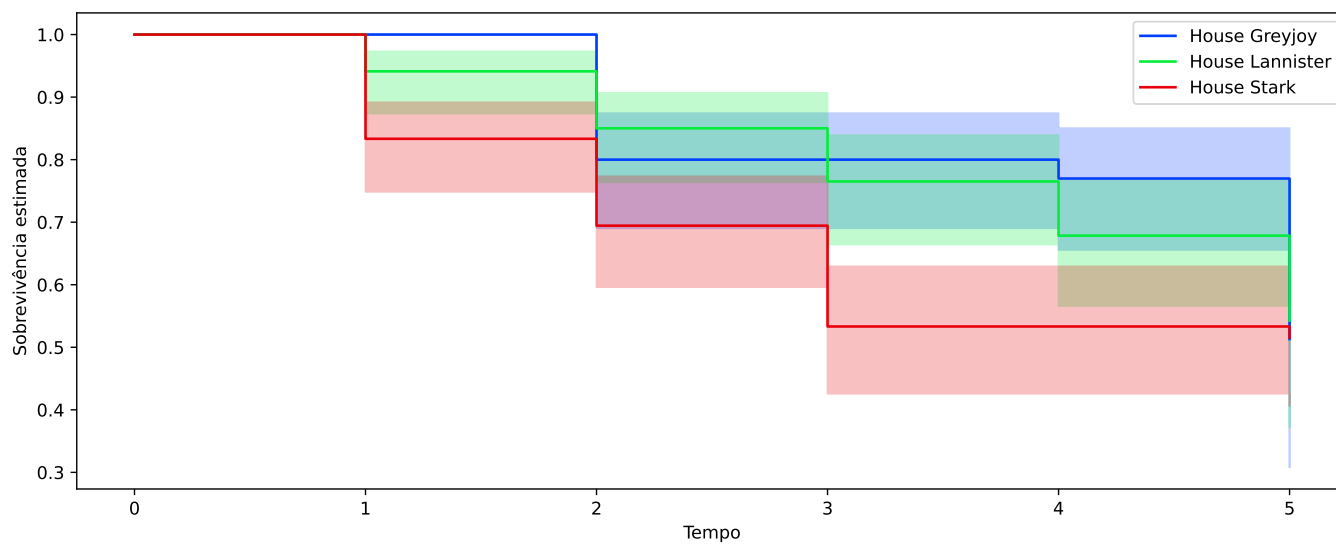


Figura 2: Função de Sobrevivência de cada família