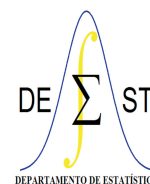




UNIVERSIDADE FEDERAL DE OURO PRETO  
INSTITUTO DE CIÊNCIAS EXATAS E BIOLÓGICAS  
DEPARTAMENTO DE ESTATÍSTICA  
BACHARELADO EM ESTATÍSTICA



# Análise de Dados de Cessação de Tabagismo

Andressa de Souza Freitas

Iara Mendes Oliveira

Joziani Mota Vieira

Samuel Sabino Freitas

Yuri Marcos Tomáz de Sousa

Ouro Preto-MG

Outubro/2020

**Andressa de Souza Freitas**  
**Iara Mendes Oliveira**  
**Joziani Mota Vieira**  
**Samuel Sabino Freitas**  
**Yuri Marcos Tomáz de Sousa**

## **Análise de Dados de Cessação de Tabagismo**

Trabalho final da disciplina de Análise de Sobrevivência EST023.

Professor(a)  
Dr. Rivert Paulo Braga de Oliveira

UNIVERSIDADE FEDERAL DE OURO PRETO – UFOP  
DEPARTAMENTO DE ESTATÍSTICA – DEEST

Ouro Preto-MG  
Outubro/2020

# Análise de Dados de Cessação de Tabagismo

## RESUMO

Milhões de pessoas morrem todos os anos devido ao consumo de tabaco. Diante desse fator, a busca por terapias mais eficazes se torna importante objeto de estudo. Este projeto investigou, por meio de técnicas de análise de sobrevivência, dois tipos de terapias contra o tabagismo (combinada de três tratamentos contra terapia apenas com adesivos), tendo como evento de interesse (falha) o tempo até a reincidência ao uso do tabaco. Os dados utilizados nesse projeto são provenientes de um estudo clínico aleatorizado de terapia tripla contra o adesivo para cessação do tabagismo. O modelo utilizado para descrever o comportamento dos dados diante das terapias foi o modelo de Cox e todas as análises foram geradas a partir do software R.

*Palavras-chave:* Tabagismo, Análise de Sobrevivência, Modelo de Cox

# Lista de figuras

1	Probabilidade sobrevivência por faixa etária. . . . .	p. 10
2	Probabilidade sobrevivência por gênero. . . . .	p. 11
3	Probabilidade sobrevivência por raça. . . . .	p. 11
4	Probabilidade sobrevivência por período tratamento. . . . .	p. 12
5	Probabilidade sobrevivência por tratamento. . . . .	p. 12
6	Ajustes do modelo . . . . .	p. 13
7	Análise Schoenfeld. . . . .	p. 14
8	Função de Sobrevivência . . . . .	p. 15

# Sumário

<b>1</b>	<b>Introdução</b>	p. 5
<b>2</b>	<b>Material e Métodos</b>	p. 6
2.1	Material . . . . .	p. 6
2.1.1	Bancos de Dados . . . . .	p. 6
2.1.2	Variáveis Estudadas . . . . .	p. 6
2.1.3	Softwares . . . . .	p. 7
2.2	Métodos . . . . .	p. 7
2.2.1	Modelo de Cox . . . . .	p. 7
2.2.2	O Modelo de Cox e Riscos Proporcionais . . . . .	p. 8
2.2.3	Interpretação dos Coeficientes . . . . .	p. 8
<b>3</b>	<b>Resultados</b>	p. 10
3.1	Análise descritiva . . . . .	p. 10
3.2	Análise não paramétrica . . . . .	p. 13
3.2.1	Modelo estimado . . . . .	p. 13
3.2.2	Razão de Risco . . . . .	p. 15
3.2.3	Sobrevivência e Risco Basal . . . . .	p. 15
<b>4</b>	<b>Considerações finais</b>	p. 17
	<b>Referências</b>	p. 18

# 1 Introdução

Segundo a WORLD HEALTH ORGANIZATION (OMS), o tabagismo é uma toxicomania que tira a vida de mais de oito milhões de pessoas por ano no mundo. Além disso, o tabagismo interfere na economia substancialmente pelos custos significativos da saúde para os tratamentos das doenças provenientes do uso do tabaco.

O estudo de Malta et al. (2017), revelou que houve redução, na maioria das vezes, de fumantes no Brasil pelos fatores sexo, escolaridade, grandes regiões, e na maioria das faixas etárias. A prevalência de ex-fumantes passou de 22,2% (2006) para 21,2% (2014). Explicam também que a tendência da prevalência de fumantes é declinante, para ambos os sexos, níveis de escolaridade e grandes regiões, em quase todas as faixas etárias. Isso aponta que a meta global de redução de 30% do tabagismo até 2025 tem potencial para ser alcançada no País.

Investigações da Agência Americana de Pesquisa e Qualidade em Saúde indicam que a maioria (70%) dos tabagistas gostaria parar de fumar, 34% tentam anualmente e somente 2,5% o conseguem, Focchi e Braun (2005).

Neste trabalho confrontamos dados de duas terapias, uma amplamente conhecida, que são os adesivos de reposição da nicotina e outra terapia combinada de três fármacos. O objetivo principal do presente trabalho é encontrar um modelo que explique a recaída do ex-fumante. Com base nisso iremos investigar o tempo que os pacientes, em média, voltem a fumar após fazer os tratamentos. Dentre outros objetivos temos como a verificação da diferença entre os tipos de tratamento e se alguma das covariáveis influência no retorno à prática de fumar.

## 2 Material e Métodos

### 2.1 Material

#### 2.1.1 Bancos de Dados

O banco de dados foi disponibilizado pelo prof. Rivert Paulo Braga de Oliveira, que é proveniente de um estudo clínico aleatorizado de terapia tripla contra o adesivo para cessação do tabagismo. O conjunto de dados conta com 125 observações e 12 variáveis.

#### 2.1.2 Variáveis Estudadas

id: Identificador do paciente;

time: tempo em dias até reincidência do hábito de fumar;

status: indicador de reincidência (retorno ao fumo);

grp: grupo de tratamento (combination: combinação de 3 tratamentos; patchonly: tratamento somente com adesivo);

age: idade na época do início do estudo;

gender: gênero feminino (Female) ou masculino (Male);

race: negro(a) (black), hispânico(a) (hispanic), branco(a) (white), outro (other);

employment: ft (full-time- tempo integral), pt (part-time-meio período), ou other(outro);

yearsSmoking: número de anos que o paciente fuma;

levelSmoking: heavy (pesado) ou light (leve);

priorAttempts: número de tentativas de parar de fumar;

longestNoSmoke: o maior período de tempo, em dias, que previamente o paciente já conseguiu ficar sem fumar.

### 2.1.3 Softwares

Utilizou-se o software R, ferramenta de análise estatística. Este software está disponível no sitio <http://www.r-project.org/>.

A edição do texto foi realizada através do editor de textos científicos Latex (Miktex + TexStudio). Estes softwares estão disponíveis nos sítios <http://miktex.org> e <http://www.texstudio.org/>, respectivamente.

## 2.2 Métodos

### 2.2.1 Modelo de Cox

Segundo (COLOSIMO; GIOLO, 2016), o modelo de regressão de Cox permite analisar dados provenientes de estudos de tempo de vida em que a resposta é o tempo até ocorrer um evento de interesse (Falha).

Suponha um estudo controlado para comparar pacientes selecionados aleatoriamente para receber um tratamento padrão (grupo 0) ou um novo tratamento (grupo 1).

Sejam  $\lambda_0(t)$  e  $\lambda_1(t)$  as funções risco de falha dos grupos 0 e 1, respectivamente. Assuma que elas são proporcionais, isto é

$$\frac{\lambda_1(t)}{\lambda_0(t)} = K \quad (2.1)$$

Se  $x$  é a variável indicadora de grupo, em que:

$$x = \begin{cases} 0, & \text{se grupo 0} \\ 1, & \text{se grupo 1} \end{cases} \quad K = \exp\{\beta x\}, \text{ então} \quad (2.2)$$

a equação abaixo é o modelo de Cox para uma única covariável:

$$\lambda_1(t) = \lambda_0(t) \exp\{\beta x\} \quad (2.3)$$

Generalizando o modelo acima, considere um vetor de covariáveis  $x = (x_1, x_2, \dots, x_p)'$  e um vetor de coeficientes  $\beta = (\beta_1, \beta_2, \dots, \beta_p)$ , logo a expressão geral do modelo de Cox fica:



$$\lambda(t) = \lambda_0(t)g(x'\beta) \quad (2.4)$$

em que  $g$  é uma função não negativa que deve ser especificada de modo que  $g(0) = 1$   $g(x'\beta)$  é o componente paramétrico do modelo.

$\lambda(t)$  é o componente não paramétrico do modelo, conhecido como função basal ( $x = 0 \Rightarrow \lambda(t) = \lambda_0(t)$ ). É este termo que absorve a parcela de variação que seria devida ao termo  $\beta_0$ .

### 2.2.2 O Modelo de Cox e Riscos Proporcionais

O componente paramétrico é frequentemente utilizado na forma multiplicativa:

$$g(x'\beta) = \exp\{\beta x'\} = \exp\{\beta_1 x_1 + \dots + \beta_p x_p\} \quad (2.5)$$

conhecido como modelo de riscos proporcionais devido ao fato de que a razão dos riscos de falha para itens amostrais  $i$  e  $j$  não dependem do tempo. Ou seja o modelo de riscos proporcionais supõe que as taxas de falha (também acumuladas) são proporcionais. Veja a equação abaixo:

$$\frac{\lambda_i(t)}{\lambda_j(t)} = \frac{\lambda_0(t)\exp\{x'_i\beta\}}{\lambda_0(t)\exp\{x'_j\beta\}} = \exp\{x'_i\beta - x'_j\beta\} \quad (2.6)$$

Para mais entendimento das equações consulte o livro (COLOSIMO; GIOLO, 2016).

### 2.2.3 Interpretação dos Coeficientes

O efeito das covariáveis é de acelerar/desacelerar a função de risco. Toma-se a razão de riscos de falha de dois itens amostrais  $i$  e  $j$ , que possuem os mesmos valores para as covariáveis, com exceção da  $l$ -ésima que pode ser interpretada como a razão de riscos instantânea no tempo  $t$ , ou simplesmente como razão de riscos, uma vez que os riscos são proporcionais.

$$\frac{\lambda_i(t)}{\lambda_j(t)} = \exp\{\beta_l(x_{il} - x_{jl})\} \quad (2.7)$$

Para uma variável dicotômica indicando pacientes hipertensos teríamos que o risco de morte entre hipertensos seria  $\exp\{\beta_l\}$  vezes o risco de morte de pacientes com pressão nor-

mal (fixas as demais variáveis) e o estimador de  $\exp\{\beta_l\}$  é obtido usando-se propriedades de invariância dos EMVs.

Toda a metodologia descrita acima pode ser encontrada em (COLOSIMO; GIOLO, 2016) e nos materiais do Prof. Dr. Rivert Paulo Braga de Oliveira.

## 3 Resultados

### 3.1 Análise descritiva

Podemos notar que na Figura 1, pessoas idosas (com mais de 51 anos) tendem a ter menos propensão à reincidência ao fumo com relação às outras faixas etárias: Jovens e adultos. Se tratando de gênero observando a Figura 2, pessoas do sexo masculino parecem demorar a ter reincidência após 50 dias, em relação as mulheres.

Na Figura 3, não houve diferença significativa entre os grupos. Podemos notar também que não houve diferença significativa entre os períodos de tratamento, podendo ser notados na Figura 4.

Utilizando a estimação de Kaplan-Meier e observando a Figura 5, podemos observar que pode existir diferença entre os tempos de vida dos dois grupos: Combinação de tratamentos e adesivo. A Figura 5 mostra que, a combinação de 3 tratamentos tende a ter uma probabilidade menor de reincidência comparado ao grupo do tratamento com adesivo.

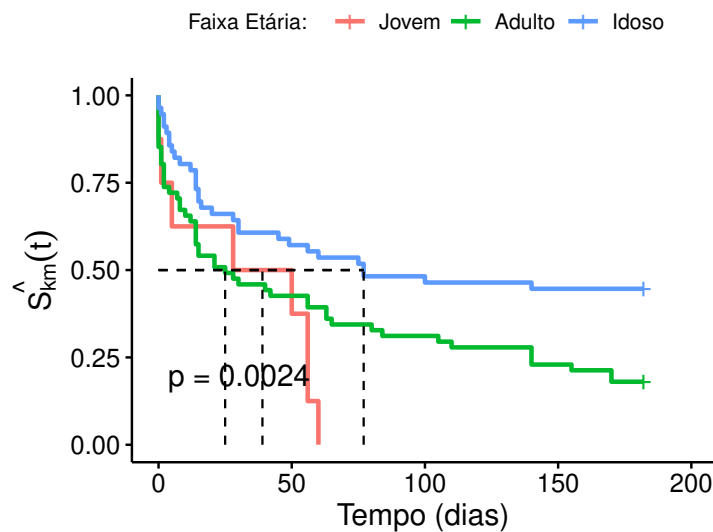


Figura 1: Probabilidade sobrevivência por faixa etária.

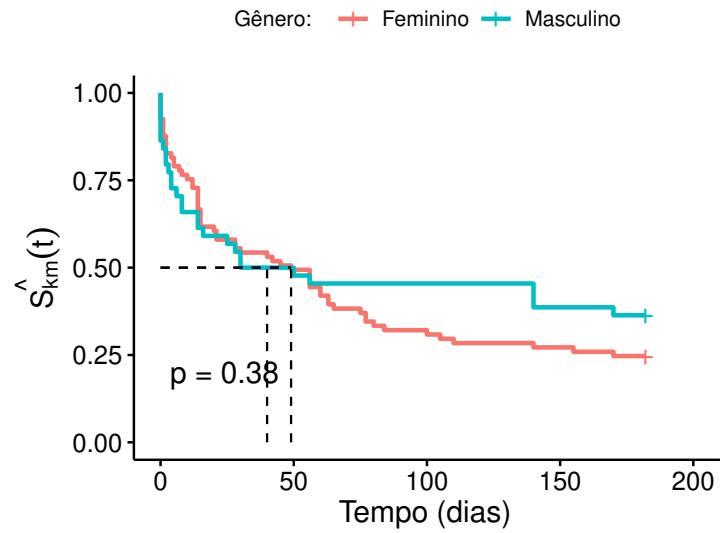


Figura 2: Probabilidade sobrevivência por gênero.

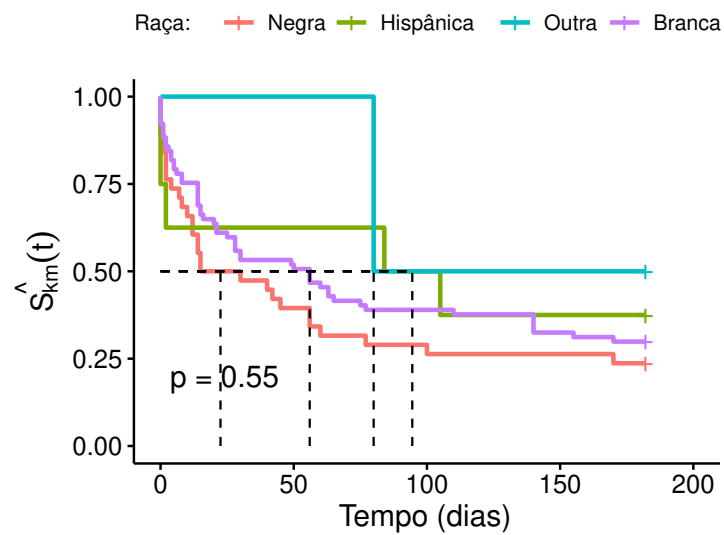


Figura 3: Probabilidade sobrevivência por raça.

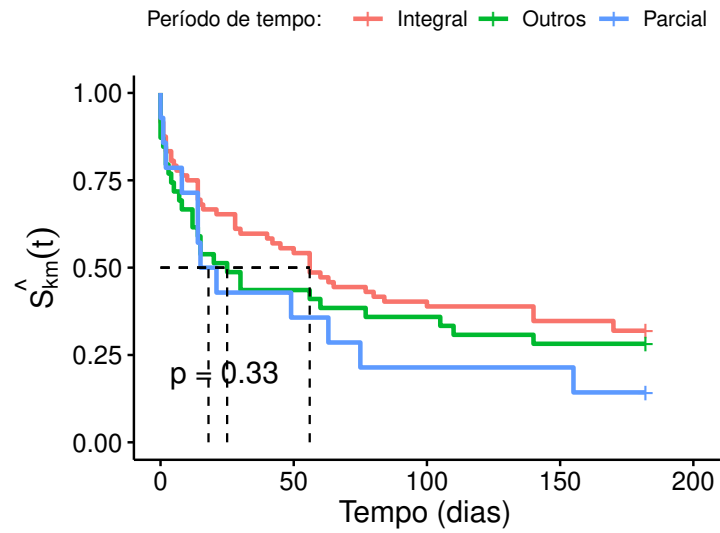


Figura 4: Probabilidade sobrevivência por período tratamento.

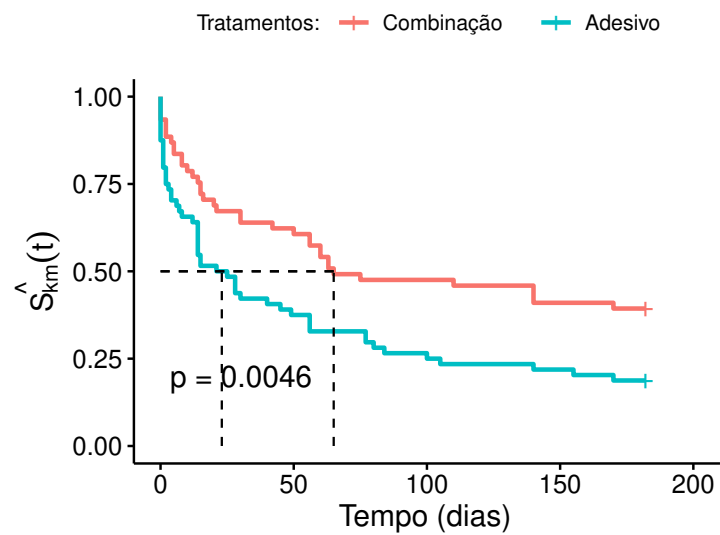


Figura 5: Probabilidade sobrevivência por tratamento.

## 3.2 Análise não paramétrica

### 3.2.1 Modelo estimado

Abordagens paramétricas, como ajustes Exponenciais, de Weibull, Lognormal e Gama Generalizada, foram utilizadas para tentar explicar os dados, porém nenhuma delas foram capazes de explicá-los. Sendo assim, optamos pela abordagem não paramétrica.

Logo o modelo não paramétrico estimado foi o modelo de Cox:

Modelo =  $\exp(0.59225189 * \text{grppatchOnly} - 0.033848224 * \text{age} + 0.67592 * \text{employmentother} + 0.63344968 * \text{employmentpt})$

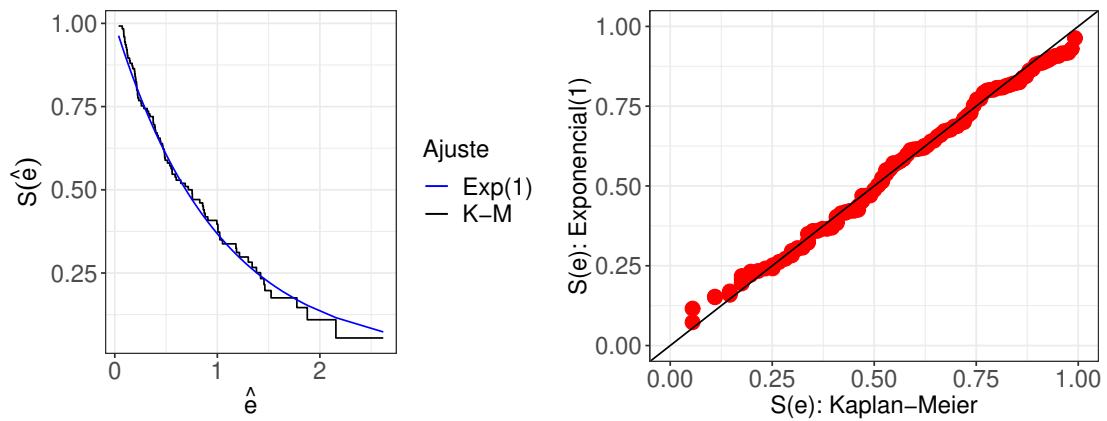


Figura 6: Ajustes do modelo

Podemos notar que na Figura 6, o modelo exponencial se ajustou bem ao dados.

Ao fazermos o Teste Global de Schoenfeld o p-valor deu 0.8647, o que significa que não temos razões estatísticas para rejeitar a hipótese nula de que os riscos são proporcionais.

Global Schoenfeld Test p: 0.8647

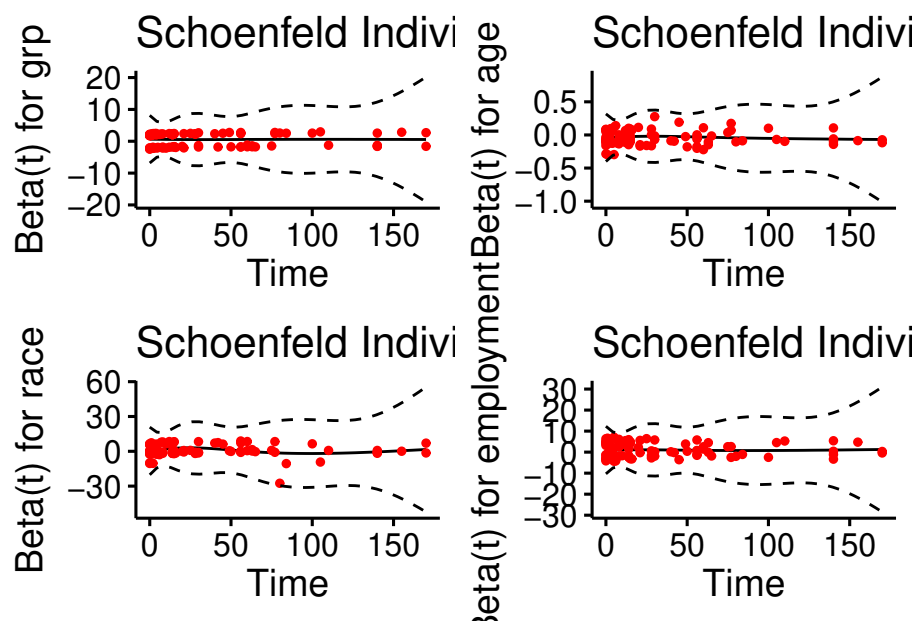


Figura 7: Análise Schoenfeld.

### 3.2.2 Razão de Risco

Pela expressão  $\exp(0.59225189 - 0.033848224 * (60 - 30) + 0.67592)$  podemos observar que: O risco de reincidência ao fumo de pacientes com 60 anos de idade, no tratamento com adesivo e que se tratam em períodos não integral é 1,28 vezes maior do que em pacientes de 30 no mesmo tratamento.

E  $\exp(0.59225189 - 0.033848224 * (60) + 0.67592) / (\exp(0.59225189 - 0.033848224 * (60) + 0.63345))$  mostra que: O risco de reincidência ao fumo de pacientes com 60 anos de idade, no tratamento com adesivo e que se tratam em outros períodos é 1,04 vezes maior do que em pacientes de 60 no mesmo tratamento que se tratam em período parcial.

### 3.2.3 Sobrevivência e Risco Basal

Podemos notar através da Figura 8 que:

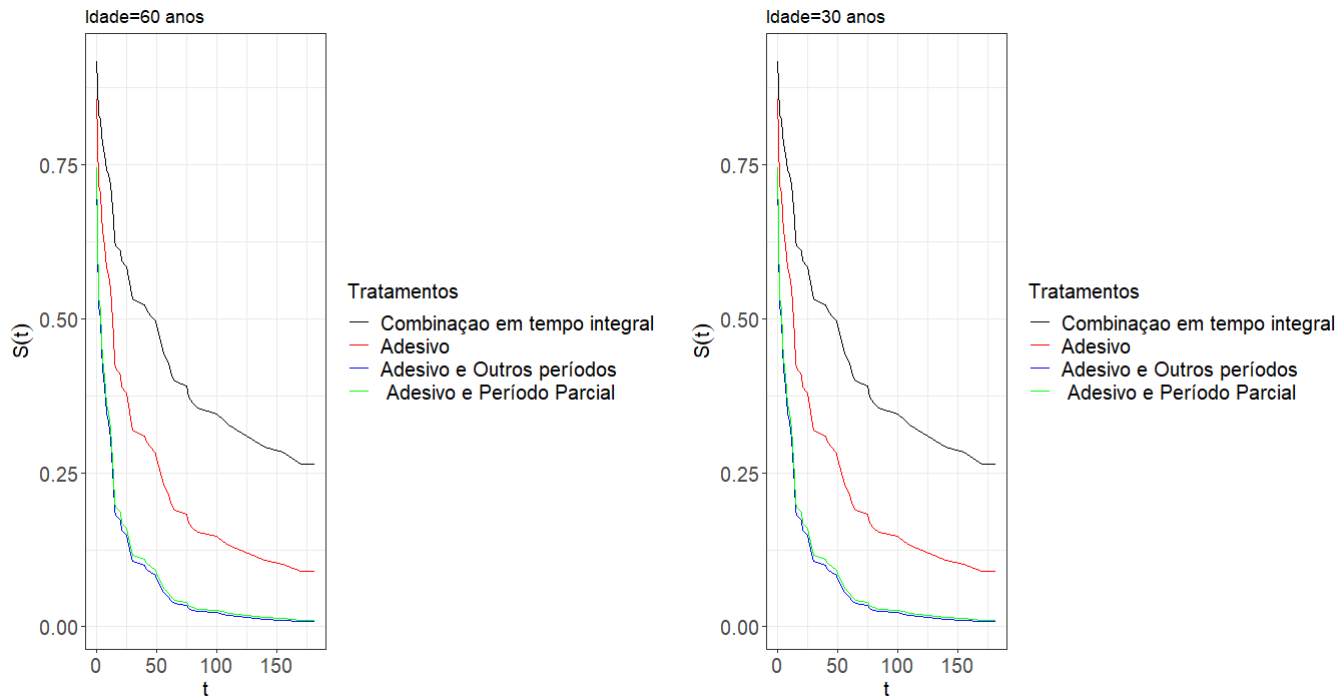


Figura 8: Função de Sobrevivência

A probabilidade de uma pessoa com 60 anos, que faz a combinação de 3 tratamentos em tempo integral, não voltar a fumar até 100 dias é de 68%. Uma pessoa de 30 nas mesmas condições tem probabilidade de 34,52%. Ou seja, pessoas mais novas tendem a voltar a fumar com menos tempo de tratamento.



A probabilidade de uma pessoa com 60 anos, que faz o tratamento com adesivo em período integral, não voltar a fumar em até 100 dias é de 49,83%, já uma pessoa de 30 nas mesmas condições é de 14,62%.

A probabilidade de uma pessoa com 60 anos, que faz o tratamento com adesivo em período parcial ou outros períodos, não voltar a fumar em até 100 dias é de cerca de 25,5%, já uma pessoa de 30 nas mesmas condições é de 2,5%.

Concluimos, então, que a combinação de 3 tratamentos é mais eficaz do que o tratamento com adesivo; tratamentos em período integral dão mais resultados positivos; períodos de tratamento parciais ou outros, com exceção do integral, geram o mesmo efeito, e são ineficazes; pessoas mais idosas tendem a ficar mais tempo sem ter recaída com relação às pessoas mais novas.

## 4 Considerações finais

Considerando os estudos desenvolvidos em cada etapa do trabalho, podemos concluir que pessoas idosas (com mais de 51 anos) tendem a ter menos propensão à reincidência ao fumo com relação às outras faixas etárias: Jovens e adultos. Se tratando de gênero, pessoas do sexo masculino parecem demorar a ter reincidência após 50 dias, em relação as mulheres.

Utilizando a estimação de Kaplan-Meier, podemos observar que pode existir diferença entre os tempos de vida dos dois grupos: Combinação de tratamentos e adesivo. sendo que a combinação de 3 tratamentos tende a ter uma probabilidade menor de reincidência comparado ao grupo do tratamento com adesivo.

Através da Razão de Risco, podemos afirmar que o risco de reincidência ao fumo de pacientes com 60 anos de idade, no tratamento com adesivo e que se tratam em períodos não integral é maior do que em pacientes de 30 anos no mesmo tratamento. E também que o risco de reincidência ao fumo de pacientes com 60 anos de idade no tratamento com adesivo e que se tratam em outros períodos é maior do que em pacientes de 60 no mesmo tratamento que se tratam em período parcial.

Por fim, concluímos que a combinação de 3 tratamentos se mostrou mais eficaz do que o tratamento com adesivo. Tratamentos em período integral dão mais resultados na cessação do tabagismo e períodos de tratamento parciais ou outros, com exceção do integral, têm o mesmo efeito, e são ineficazes. Pessoas com idade mais avançada tendem a ficar mais tempo sem ter recaída com relação às pessoas mais novas.

## Referências

COLOSIMO, E. A.; GIOLO, S. R. *Análise de Sobrevivência Aplicada*. 1. ed. São Paulo: Editora Blucher, 2016.

FOCCHI, G. R. de A.; BRAUN, I. M. Tratamento farmacológico do tabagismo. *Pharmacological Treatment of Tobacco Dependence*, v. 32, n. 5, p. 259–266, 2005.

MALTA, D. C. et al. Evolução de indicadores do tabagismo segundo inquéritos de telefone, 2006-2014. *Cad. Saude Pública*, v. 33, n. 3, p. 162–173, 2017.

WORLD HEALTH ORGANIZATION. *Tobacco*. Disponível em: <<https://www.who.int/news-room/fact-sheets/detail/tobacco#>>. Acesso em: 27 maio. 2020.