

Modelo de Tempo de Vida Acelerado Log-Logístico Aplicado a Dados de câncer de pulmão

Lucas Pereira de Moura¹, Alex Leal Mota² (PQ)

^{1,2}Universidade Federal do Amazonas, Departamento de Estatística, Av. Rodrigo Otávio, 6200, Coroado I, 69067-005, Manaus AM, Brasil.

¹lucas.moura@ufam.edu.br, ²alexmota@ufam.edu.br

Palavras-Chave: Análise de sobrevivência, Dados de câncer de pulmão, Distribuição log-logística.

Introdução

O câncer de pulmão é um dos principais tipos de câncer com causas de mortalidade no Brasil. De acordo com Instituto Nacional de Câncer (INCA), foram estimados 28.220 novos casos em 2016, configurando-se como o segundo tipo de câncer mais comum entre os homens e o quarto entre as mulheres. Frequentemente diagnosticado em estágios avançados (III ou IV), o que contribui para baixas taxas de sobrevivência¹.

Neste contexto, análise de sobrevivência é amplamente utilizada em estudos oncológicos para modelar o tempo até a ocorrência de eventos como óbito ou progressão da doença. A escolha da distribuição probabilística é crucial, uma vez que modelos inadequados podem levar a estimativas viesadas³.

Neste trabalho, o modelo de Tempo de Falha Acelerado Log-Logístico (AFT-Log-Log) foi aplicado à análise de sobrevivência de pacientes com câncer de pulmão, devido à flexibilidade de sua função de risco, que pode assumir formas decrescentes ou unimodais.

Material e Métodos

O conjunto de dados sobre câncer de pulmão é proveniente de uma pesquisa retrospectiva com 25.971 registros de pacientes diagnosticados com câncer de pulmão no estado de São Paulo, Brasil, entre 2000 e 2014. O acompanhamento desses pacientes foi realizado até 2018. Todos os pacientes foram diagnosticados com neoplasia maligna do brônquio e do pulmão (C34 – código de diagnóstico CID-10), e casos dos estágios clínicos III e IV (metastáticos) foram incluídos na amostra. O conjunto de dados foi fornecido pela Fundação Oncocentro de São Paulo (FOSP), que é responsável pela coordenação do Registro Hospitalar de Câncer do Estado de São Paulo (<http://fosp.saude.sp.gov.br>)⁵.

Neste contexto, foi aplicado o modelo AFT-Log-Log, afim de verificar os efeitos das covariáveis na sobrevivência de pacientes acometidos por câncer de pulmão. O modelo de Tempo de Falha Acelerado (AFT) é definido por

$$Y = \log T = \beta^T \mathbf{X} + \sigma \epsilon,$$

em que ϵ é uma variável aleatória com distribui-

ção conhecida (dependendo da distribuição escolhida para T). Assim $T = \exp(\beta^T \mathbf{X}) W$, onde $W = \exp(\sigma \epsilon)$ tem função de risco $h_0(w)$ independente de β . Segue que as funções de risco e de sobrevivência de T são expressas, respectivamente, por $h(t | \mathbf{X}) = \exp(-\beta^T \mathbf{X}) h_0(\exp(-\beta^T \mathbf{X}) t)$ e $S(t | \mathbf{X}) = \exp(-H_0(\exp(-\beta^T \mathbf{X}) t))$, onde $H_0(\cdot)$ denota a função de risco acumulado⁴. Para $j \geq 1$, $\exp(\beta_j)$ representa o fator de aceleração associado à covariável X_j , mantendo as demais constantes. Assim:

- Se $\exp(\beta_j) > 1$: X_j está associado a um aumento no tempo até o evento (o evento demora mais para acontecer);
- Se $\exp(\beta_j) < 1$: X_j está associado a uma redução no tempo até o evento (o evento acontece mais rapidamente).

Neste trabalho consideramos que $\epsilon \sim \text{Logística}(0, 1)$. É possível provar que $T | \mathbf{X} \sim \text{Log-Logística}(\lambda(\mathbf{X}), \gamma)$ em que $\gamma = 1/\sigma$ e $\lambda(\mathbf{X}) = \exp\{\beta^T \mathbf{X}\}$. As funções de sobrevivência e de risco do modelo de regressão AFT Log-logístico são dadas, respectivamente, por

$$S(t | \mathbf{X}) = \frac{1}{1 + \left(\frac{t}{\lambda(\mathbf{X})}\right)^\gamma}$$

e

$$h(t | \mathbf{X}) = \frac{\left(\frac{\gamma}{\lambda(\mathbf{X})}\right) \left(\frac{t}{\lambda(\mathbf{X})}\right)^{\gamma-1}}{1 + \left(\frac{t}{\lambda(\mathbf{X})}\right)^\gamma}.$$

A distribuição Log-Logística é flexível, podendo acomodar riscos decrescente ($\gamma \leq 1$) e unimodal ($\gamma > 1$).

Resultados e Discussão

Um total de 24.279 eventos (93,49%) ocorreram durante o período de acompanhamento. O tempo máximo de observação foi de aproximadamente 18,75 anos, enquanto o tempo mediano de acompanhamento foi de 6,28 anos. A população do estudo foi majoritariamente do sexo masculino (64,01%), com idade superior a 60 anos (59,59%) e em estágio clínico IV (64,58%). Em relação às interven-

ções terapêuticas, 12,10% dos pacientes foram submetidos à cirurgia, 40,46% receberam radioterapia e 30,40% realizaram quimioterapia.

Com base no gráfico TTT (Fig. 2), observa-se que, na origem do tempo, a curva se sobrepõe à reta identidade, indicando um comportamento unimodal da função de risco².

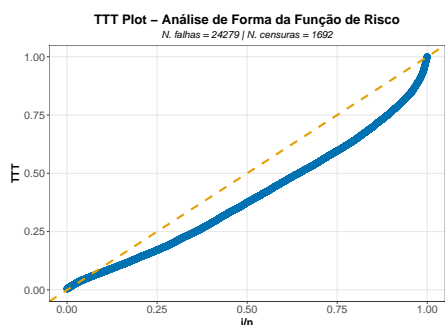


Figura 1: TTT Plot - Dados de câncer de pulmão.

Após a indicação pelo gráfico na Fig. 1, os dados foram submetidos à modelagem por dois diferentes modelos de tempo de falha acelerado (AFT): o modelo Gama Generalizado e o modelo Log-Logístico. Os critérios de comparação (AIC e Log-verossimilhança) revelaram que o modelo AFT-Log-Logístico apresentou resultados superiores em relação ao modelo AFT Gama Generalizado, justificando sua escolha para a análise subsequente.

A Tab. 1 apresenta as estimativas do modelo AFT-Log-Log para os dados de câncer de pulmão (AIC = 41.936,02). Nota-se que a covariável idade não foi significativa. A Tab. 2 mostra o modelo sem idade, que apresentou melhor ajuste pelo AIC (41.934,11).

Tabela 1: Resultados do modelo AFT-Log-Log.

Parâmetro	Estimativas	EP	IC 95%
Forma (γ)	1,575	0,0084	(1,559; 1,592)
β_0	0,197	0,0053	(0,187; 0,207)
Sexo:Fem	0,187	0,0144	(0,160; 0,215)
Radioterapia:Sim	0,417	0,0140	(0,389; 0,444)
Quimioterapia:Sim	1,358	0,0154	(1,328; 1,388)
Estágio Clínico:Sim	-0,375	0,0146	(-0,404 -0,347)
Cirurgia:Sim	0,653	0,0220	(0,610; 0,697)
Idade:> 60	0,006	0,0206	(-0,034; 0,047)
Log-verossimilhança = -20.960,01 AIC = 41.936,02			

Tabela 2: Resultados do modelo Log-Logístico AFT (sem idade).

Parâmetro	Estimativas	EP	IC 95%
Forma (γ)	1,575	0,0084	(1,559; 1,592)
β_0	0,198	0,0036	(0,191; 0,205)
Sexo:Fem	0,187	0,0144	(0,159; 0,215)
Radioterapia:Sim	0,417	0,0140	(0,389; 0,444)
Quimioterapia:Sim	1,358	0,0154	(1,328 1,388)
Estágio Clínico:IV	-0,375	0,0146	(-0,404; -0,347)
Cirurgia:Sim	0,653	0,0220	(0,610; 0,696)
Log-verossimilhança = -20.960,05 AIC = 41.934,11			

Levando em consideração os perfis de dois hipotéticos pacientes, um do sexo feminino (curva vermelha) e ou-

tro do sexo masculino (curva azul), mantendo-se fixos os mesmos prognósticos (ambos realizaram cirurgia, radioterapia, quimioterapia e encontram-se no estágio clínico III), a Figura 2 permite observar que a sobrevida é superior no sexo feminino ao longo de todo o período analisado. Por exemplo, aos 5 anos, a probabilidade de sobrevivência é de 17% para mulheres e 14% para homens, enquanto aos 10 anos, essas probabilidades caem para 7% e 5%, respectivamente. Esses resultados reforçam a influência do sexo no desfecho de sobrevida, mesmo quando comparados pacientes com características clínicas e terapêuticas semelhantes.

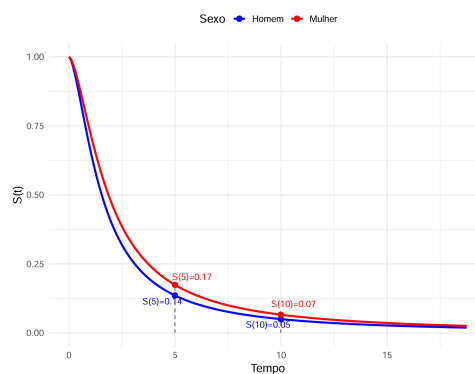


Figura 2: Comparação das sobrevivências estimadas para Homens e Mulheres com mesmo perfil clínico (tratamentos e estágio).

Conclusões

O modelo final indicou que os tratamentos tiveram efeito expressivo sobre o tempo de sobrevida dos pacientes. Aqueles que realizaram quimioterapia apresentaram tempo de vida aproximadamente 3,8 vezes maior ($\exp(\hat{\beta}) \approx 3,8$), enquanto a cirurgia aumentou esse tempo em cerca de 90% ($\exp(\hat{\beta}) \approx 1,9$) e a radioterapia, em aproximadamente 50% ($\exp(\hat{\beta}) \approx 1,5$). Mulheres apresentaram tempo de sobrevida cerca de 20% superior ao dos homens ($\exp(\hat{\beta}) \approx 1,2$). Por outro lado, pacientes diagnosticados em estágio clínico IV apresentaram tempo de vida aproximadamente 31% menor em comparação aos pacientes em estágio III ($\exp(\hat{\beta}) \approx 0,69$), indicando uma ocorrência mais rápida do óbito nesse grupo.

Referências

- [1] Araujo, L. H.; Baldotto, C.; Castro Jr, G. d.; Katz, A.; Ferreira, C. G.; Mathias, C.; Mascarenhas, E.; Lopes, G. d. L.; Carvalho, H.; Tabacof, J. et al. Lung cancer in Brazil. *Jornal Brasileiro de Pneumologia*, 44(01):55-64. 2018.
- [2] Barlow, R. E.; Campo, R. Total time on test processes and applications to failure data analysis. In *Reliability and fault tree analysis*. 1975.
- [3] Colosimo, E. A.; Giolo, S. R. *Análise de Sobrevida Aplicada*. Editora Edgard Blücher, São Paulo. 2006.
- [4] Kalbfleisch, J. D.; Prentice, R. L. *The statistical analysis of failure time data*. John Wiley & Sons. 2002.
- [5] Mota, A.; Milani, E. A.; Calsavara, V. F.; Tomazella, V. L.; Leao, J.; Ramos, P. L.; Ferreira, P. H.; Louzada, F. Weighted lindley frailty model: estimation and application to lung cancer data. *Lifetime Data Analysis*, 27(4):561-587. 2021.