Estudo do Tempo de Internação dos Pacientes Usando o Modelo de Regressão Weibull

Rafaela Gil[(a)]
Carlos Aparecido dos Santos[(b)]

Departamento de Estatística
Universidade Estadual de Maringá
Maringá-PR, Brasil
e-mail: (a) rafauem_gil@hotmail.com
e-mail: (b) casantos@uem.br

1 Introdução

A análise de sobrevivência pode ser definida, de forma ampla, como a análise do tempo até a ocorrência de um dado evento. Este tempo é denominado tempo de falha e pode ser o tempo até a morte do paciente, bem como até a cura ou recidiva de uma doença.

A principal característica de dados de sobrevivência é a presença de censuras, que é a observação parcial da resposta. Isto se refere a situações em que, por alguma razão, o acompanhamento do paciente foi interrompido, seja porque o mesmo mudou de cidade, o estudo terminou para a análise dos dados ou, o paciente morreu de causa diferente da estudada. Sem a presença de censuras, técnicas estatísticas clássicas usuais podem ser usadas para analisar os dados, (ver por exemplo, Kalbfleisch & Prentice, 2002).

Motivados pelos dados fornecidos do Hospital Municipal de Maringá Dra. Thelma Villanova Kasprowicz, que apresenta os tempos de internação dos pacientes, utilizamos de técnicas de análise de sobrevivência para solucionar problemas encontrados com estes tempos.

O Hospital Municipal, que foi inaugurado em abril de 2002, contando inicialmente com 15 leitos de pediatria e 15 leitos de clínica médica, os quais representam apenas 4% do total de leitos disponíveis no sistema hospitalar de Maringá-PR. Por não possuir infra-estrutura completa, possuindo apenas clínica médica e pediátrica, o Hospital Municipal é considerado de pequeno porte.

Assim, um estudo sobre o tempo de internação dos pacientes é de suma importância para o município, uma vez que, através do mesmo se faz possível uma melhor administração dos recursos no hospital.

A rede hospitalar no município era na época, predominantemente privada, contando com dez hospitais, dos quais apenas dois pertenciam a rede pública de saúde, sendo eles o Hospital Municipal de Maringá Dra. Thelma Villanova Kasprowicz e o Hospital Universitário.

Em 2006, os dez hospitais do município ofereceram juntos 1128 leitos, porém, os dois hospitais da rede pública ofereceram juntos apenas 120 leitos, ou seja, $10,64\%$ dos leitos em relação ao total.

Ao visualizar a situação da saúde pública do município, um estudo sobre o tempo de permanência dos diversos pacientes é de suma importância para a administração destes poucos leitos.

Outra motivação para este estudo, além da citada inicialmente, foi a publicação feita pelo SIH/SUS no ano de 2000, onde este mostra as principais causas de internação em hospitais no país. Neste *ranking*, tem-se em primeiro lugar os partos, os quais representam 24% das internações. Em seguida tem-se, respectivamente, as doenças do aparelho respiratório $(16,2\%)$, doenças do aparelho circulatório $(9,5\%)$, doenças do aparelho digestivo $(8,5\%)$ e doenças infecciosas e parasitárias $(7,4\%)$.

2 Modelo de Regressão Weibull

A distribuição de Weibull foi proposta originalmente por Weibull (1951) e sua ampla aplicabilidade foi também discutida por este mesmo autor em anos posteriores. A sua popularidade em aplicações práticas, se deve ao fato dela apresentar uma grande variedade de formas, todas com uma propriedade em comum: a sua função de taxa de falha é monótona, isto é, ela é crescente, decrescente ou constante.

Na maioria dos estudos voltados a área médica, há características associadas à cada paciente que são representadas através de variáveis explicativas (covariáveis). Assim se tivermos associado a cada item um vetor de variáveis explicativas $x_1 = x_{i1}, x_{i2}, \ldots, x_{ip}$; e estamos interessados em determinar a relação entre o tempo até a ocorrência de falha e o vetor de covariáveis, fazemos isso através do uso de um modelo de regressão (Lawless, 1982).

Para uma variável aleatória T com distribuição de Weibull, temos a função de densidade de probabilidade dada por

$$f(t) = \frac{\gamma}{\alpha^{\gamma}} t^{\gamma-1} \exp\left\{-\left(\frac{t}{\alpha}\right)^{\gamma}\right\} \tag{1}$$

em que γ é o parâmetro de forma, α o de escala e t a variável aleatória positiva que indica os tempos de internação. O parâmetro α tem a mesma unidade de medida de t e γ não tem unidade, apenas indica se a função de risco é crescente $(\gamma > 1)$, decrescente $(\gamma < 1)$ ou constante $(\gamma = 1)$.

As funções de sobrevivência e de risco são, respectivamente,

$$S(t) = \exp\left\{-\left(\frac{t}{\alpha}\right)^{\gamma}\right\} \tag{2}$$

e

$$\lambda(t) = \frac{\gamma}{\alpha^{\gamma}} t^{\gamma-1} \tag{3}$$

Observe que, quando $\gamma = 1$, tem-se a distribuição exponencial, ou seja, a distribuição explonencial é um caso particular da distribuição Weibull quando a taxa de falha é constante.

Assim, o modelo de regressão para o modelo Weibull é dado por

$$f(t) = \frac{\gamma}{(\alpha_0 + \alpha_1 X_1 + \ldots + \alpha_n X_n)^{\gamma}} t^{\gamma-1} \exp\left\{-\left(\frac{t}{\alpha_0 + \alpha_1 X_1 + \ldots + \alpha_n X_n}\right)^{\gamma}\right\} \tag{4}$$

2.1 Função de Verossimilhança

A função de verossimilhança contemplando dados censurados à direita, é dada por (Colosimo & Giolo, 2006)

$$L(\theta) = \prod_{i=1}^{n} [f(t_i \mid \theta)]^{\delta_i} [S(t_i \mid \theta)]^{1-\delta_i} \tag{5}$$

Para o caso Weibull, a função de verossimilhança é dada por

$$L(\theta) = \prod_{i=1}^{n} \left[\frac{\gamma}{\alpha^{\gamma}} t^{\gamma-1} \exp\left\{-\left(\frac{t}{\alpha}\right)^{\gamma}\right\}\right]^{\delta_i} \left[\exp\left\{-\left(\frac{t}{\alpha}\right)^{\gamma}\right\}\right]^{1-\delta_i} \tag{6}$$

e o logaritmo desta função é

$$l(\theta) = \sum_{i=1}^{n} \delta_i \log\left(\frac{\gamma}{\alpha^{\gamma}}\right) + (\gamma - 1) \sum_{i=1}^{n} \delta_i \log t_i - \sum_{i=1}^{n} \left(\frac{t_i}{\alpha}\right)^{\gamma} \tag{7}$$

Para estimação dos parâmetros desta distribuição, consideremos a reparametrização onde,

$$\begin{aligned} \alpha &= \exp(\alpha_0 + \alpha_1 X_1 + \ldots + \alpha_n X_n) \\ \gamma &= \exp(\gamma_0). \end{aligned}$$

2.2 Estimador Não-Paramétrico

O Estimador não-paramétrico de Kaplan-Meier, proposto por Kaplan & Meier (1958), estima a função de sobrevivência para um estudo em que nem todas as observações falharam, ou seja, existiram censuras. A observação censurada informa que o tempo até a falha é maior que o observado.

O Estimador de Kaplan-Meier é definido como:

$$\hat{S}(t) = \prod_{j:t_j<t} \left(\frac{n_j - d_j}{n_j} \right) = \prod_{j:t_j<t} \left(1 - \frac{d_j}{n_j} \right) \tag{8}$$

Onde,

- $t_1 < ... < t_k$, os k tempos distintos e ordenados de falha,
- d_j o número de falhas em t_j, $j = 1, ..., k$, e
- n_j o número de indivíduos sob risco em t_j, ou seja, os indivíduos que não falharam e não foram censurados até o instante imediatamente anterior a t_j.

3 Resultados

Neste trabalho, a variável de interesse são os tempos de internação dos pacientes no Hospital Municipal. O mesmo disponibilizou para o estudo, todos os dados de internação, desde sua inauguração em abril de 2002 até março de 2003, totalizando 1732 pacientes.

Outra variável também disponibilizada, foi os tipos de doenças, classificados em cinco categorias:

- doenças do aparelho respiratório (DR);
- doenças do aparelho circulatório (DC);
- doenças do aparelho digestivo (DD);
- doenças infecciosas e parasitárias (DI) e
- outras doenças, sendo estas, doenças que não oferecem risco de morte aos pacientes (DO).

Além dos tipos de doenças citados anteriormente, temos disponível as variáveis sexo e idade desses 1732 pacientes.

Inicialmente, para os dados apresentados, obtivemos as estimativas dos parâmetros para o modelo Weibull sem covariáveis, o que esta apresentado na Tabela 1.

Através desta é possível verificar que todos os parâmetros estimados para o modelo são significativos. Também, através da Figura 1, temos que o modelo Weibull se ajusta bem aos dados, o que é possível concluir dado que as curvas de sobrevivência empírica e estimada estão muito próximas.

Tabela 1: Estimativas dos parâmetros para o modelo Weibull sem covariáveis.

Parâmetros	Estimativas	Erro Padrão	P-Valor	Reparametrização
α_0	1.706	0.020	< 0.0001	5.507
γ_0	0.302	0.019	< 0.0001	1.352

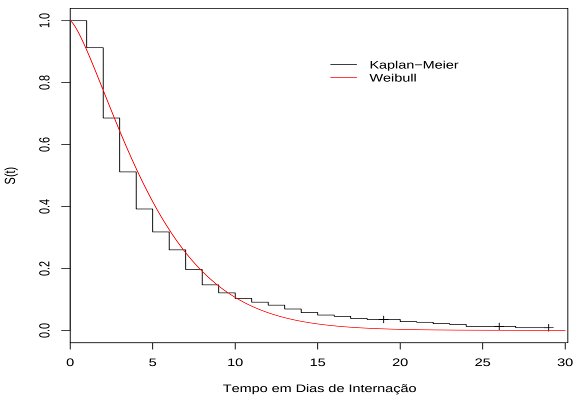

Figura 1: Curva empírica versus estimada pelo modelo Weibull para os tempos de internação.

Em um segundo momento, estimamos os parâmetros do modelo Weibull considerando um modelo de regressão como parâmetro de escala dado por $\alpha = \exp(\alpha_0 + \alpha_1 X_{Idade} + \alpha_2 X_{DC} + \alpha_3 X_{DI} + \alpha_4 X_{DD} + \alpha_5 X_{DR} + \alpha_6 X_{Sexo})$. Ao estimar este modelo, verificamos que a covariável sexo é não significativa ao nível de 5%. Deste modo, reestimamos os parâmetros deste modelo excluindo o regressor α_6. As estimativas estão apresentadas na Tabela 2 que segue.

Tabela 2: Estimativas dos parâmetros para o modelo Weibull com covariáveis.

Parâmetros	Estimativas	Erro Padrão	P-Valor	Reparametrização
α_0	0.913389	0.047117	< 0.0001	–
α_1	0.298326	0.015086	< 0.0001	–
α_2	0.123979	0.090035	0.168687	–
α_3	0.222249	0.062599	0.000395	–
α_4	0.586737	0.164589	0.000374	–
α_5	0.063956	0.038197	0.094236	–
γ_0	0.43733	0.018804	< 0.0001	1.548567

onde $\alpha = \exp(\alpha_0 + \alpha_1 X_{Idade} + \alpha_2 X_{DC} + \alpha_3 X_{DI} + \alpha_4 X_{DD} + \alpha_5 X_{DR})$ e $\gamma = \exp(\gamma_0)$ é o parâmetro de forma do modelo Weibull, indicando que o risco é crescente.

Através destas estimativas, podemos verificar que o tipo de doença que mais influencia na variável resposta, tempo de internação, são as doença do aparelho digestivo. Por outro lado, a doença que dispende menor tempo de internação são as do aparelho respiratório.

Também verificamos que, quanto maior a idade do paciente, maior é o tempo de internação.

4 Conclusões

Motivados pelos dados referentes aos tempos de internação dos pacientes do Hospital Municipal de Maringá, um estudo inicial em análise de sobrevivência foi realizado.

Neste estudo buscamos técnicas para, além de modelar os tempos, incluir covariáveis que, de alguma forma, influenciassem nos tempos de internação finais.

Dentro deste contexto, a utilização do modelo Weibull, com e sem covariáveis, parece ser adequada para modelar os tempos de internação dos pacientes do hospital mencionado.

Referências

Colosimo, E. A. & Giolo, S. R. (2006). *Análise de Sobrevivência Aplicada*. Projeto Fisher - ABE. Edgard Blücher Ltda., São Paulo.

Kalbfleisch, J. D. & Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data*. John Wiley, New York.

Kaplan, E. L. & Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, (53), 457–481.

Lawless, J. F. (1982). *Statistical Models and Methods for Lifetime Data*. John Wiley, New York.

Weibull, W. (1951). A Statistical distribution function of wide applicability. *Journal of Applied Mechanics*, pages 292–297.