Métodos Não Paramétricos

Adaptado de material do
Prof. Vicente G. Cancho (ICMC/USP)

Departamento de Matemática Aplicada e Estatística
Universidade de São Paulo

2023

Métodos não paramétricos

Considere n indivíduos em estudo. Sejam $t_1, t_2, \ldots, t_n$ tempos de vida distintos e sejam $t_{(1)} < t_{(2)} < \cdots < t_{(n)}$ as respectivas observações ordenadas (estatísticas de ordem).

Métodos não paramétricos

Considere n indivíduos em estudo. Sejam $t_1, t_2, \ldots, t_n$ tempos de vida distintos e sejam $t_{(1)} < t_{(2)} < \cdots < t_{(n)}$ as respectivas observações ordenadas (estatísticas de ordem).

Um estimador da função sobrevivência $S(t)$ é a função sobrevivência empírica ou amostral denotada por $\widehat{S}(t)$ e é dada por

$$\widehat{S}(t) = \frac{\text{N}^{\underline{\text{o}}} \text{ de indivíduos que não falharam até o tempo } t}{n} = \frac{\text{N}^{\underline{\text{o}}} \text{ de indivíduos com tempo de falha} > t}{n}, t \geq 0,$$

Métodos não paramétricos

Considere n indivíduos em estudo. Sejam $t_1, t_2, \ldots, t_n$ tempos de vida distintos e sejam $t_{(1)} < t_{(2)} < \cdots < t_{(n)}$ as respectivas observações ordenadas (estatísticas de ordem).

Um estimador da função sobrevivência $S(t)$ é a função sobrevivência empírica ou amostral denotada por $\widehat{S}(t)$ e é dada por

$$\begin{aligned}\widehat{S}(t) &= \frac{\text{Nº de indivíduos que não falharam até o tempo } t}{n} \\ &= \frac{\text{Nº de indivíduos com tempo de falha} > t}{n}, t \geq 0,\end{aligned}$$

ou equivalentemente,

$$\widehat{S}(t) = \frac{n-i}{n} = 1 - \frac{i}{n}, \text{ se } t_{(i)} < t \leq t_{(i+1)}. \qquad (1)$$

Observe que $\widehat{S}(0) = 1 - \frac{0}{n} = 1$,

Métodos não paramétricos

Considere n indivíduos em estudo. Sejam $t_1, t_2, \ldots, t_n$ tempos de vida distintos e sejam $t_{(1)} < t_{(2)} < \cdots < t_{(n)}$ as respectivas observações ordenadas (estatísticas de ordem).
Um estimador da função sobrevivência $S(t)$ é a função sobrevivência empírica ou amostral denotada por $\widehat{S}(t)$ e é dada por

$$\begin{aligned}\widehat{S}(t) &= \frac{\text{N}^{\text{o}} \text{ de indivíduos que não falharam até o tempo } t}{n} \\ &= \frac{\text{N}^{\text{o}} \text{ de indivíduos com tempo de falha} > t}{n}, t \geq 0,\end{aligned}$$

ou equivalentemente,

$$\widehat{S}(t) = \frac{n-i}{n} = 1 - \frac{i}{n}, \text{ se } t_{(i)} < t \leq t_{(i+1)}. \qquad (1)$$

Observe que $\widehat{S}(0) = 1 - \frac{0}{n} = 1$, $\widehat{S}(0) = 1$, para $0 < t \leq t_{(1)}$,

Métodos não paramétricos

Considere n indivíduos em estudo. Sejam $t_1, t_2, \ldots, t_n$ tempos de vida distintos e sejam $t_{(1)} < t_{(2)} < \cdots < t_{(n)}$ as respectivas observações ordenadas (estatísticas de ordem).
Um estimador da função sobrevivência $S(t)$ é a função sobrevivência empírica ou amostral denotada por $\widehat{S}(t)$ e é dada por

$$\begin{aligned}\widehat{S}(t) &= \frac{\text{Nº de indivíduos que não falharam até o tempo } t}{n}\\ &= \frac{\text{Nº de indivíduos com tempo de falha} > t}{n}, t \geq 0,\end{aligned}$$

ou equivalentemente,

$$\widehat{S}(t) = \frac{n-i}{n} = 1 - \frac{i}{n}, \text{ se } t_{(i)} < t \leq t_{(i+1)}. \tag{1}$$

Observe que $\widehat{S}(0) = 1 - \frac{0}{n} = 1$, $\widehat{S}(0) = 1$, para $0 < t \leq t_{(1)}$, $\widehat{S}(t_{(n)}) = 1 - \frac{n}{n} = 0$,

Métodos não paramétricos

Considere n indivíduos em estudo. Sejam $t_1, t_2, \ldots, t_n$ tempos de vida distintos e sejam $t_{(1)} < t_{(2)} < \cdots < t_{(n)}$ as respectivas observações ordenadas (estatísticas de ordem).
Um estimador da função sobrevivência $S(t)$ é a função sobrevivência empírica ou amostral denotada por $\widehat{S}(t)$ e é dada por

$$\begin{aligned}\widehat{S}(t) &= \frac{\text{Nº de indivíduos que não falharam até o tempo } t}{n} \\ &= \frac{\text{Nº de indivíduos com tempo de falha} > t}{n}, t \geq 0,\end{aligned}$$

ou equivalentemente,

$$\widehat{S}(t) = \frac{n-i}{n} = 1 - \frac{i}{n}, \text{ se } t_{(i)} < t \leq t_{(i+1)}. \tag{1}$$

Observe que $\widehat{S}(0) = 1 - \frac{0}{n} = 1$, $\widehat{S}(0) = 1$, para $0 < t \leq t_{(1)}$, $\widehat{S}(t_{(n)}) = 1 - \frac{n}{n} = 0$, $\widehat{S}(t) = 0$, para $t > t_{(n)}$

Métodos não paramétricos

Considere n indivíduos em estudo. Sejam $t_1, t_2, \ldots, t_n$ tempos de vida distintos e sejam $t_{(1)} < t_{(2)} < \cdots < t_{(n)}$ as respectivas observações ordenadas (estatísticas de ordem).
Um estimador da função sobrevivência $S(t)$ é a função sobrevivência empírica ou amostral denotada por $\widehat{S}(t)$ e é dada por

$$\begin{aligned}\widehat{S}(t) &= \frac{\text{N}^{\underline{\text{o}}} \text{ de indivíduos que não falharam até o tempo } t}{n} \\ &= \frac{\text{N}^{\underline{\text{o}}} \text{ de indivíduos com tempo de falha} > t}{n}, t \geq 0,\end{aligned}$$

ou equivalentemente,

$$\widehat{S}(t) = \frac{n-i}{n} = 1 - \frac{i}{n}, \text{ se } t_{(i)} < t \leq t_{(i+1)}. \qquad (1)$$

Observe que $\widehat{S}(0) = 1 - \frac{0}{n} = 1$, $\widehat{S}(0) = 1$, para $0 < t \leq t_{(1)}$, $\widehat{S}(t_{(n)}) = 1 - \frac{n}{n} = 0$, $\widehat{S}(t) = 0$, para $t > t_{(n)}$ e $\widehat{S}(t)$ é contínua à esquerda.

Exemplo (Dados sem censura)

Em seguida apresentamos a distribuição de frequência dos tempos de falha de $n = 54$ indivíduos.

Tempo de vida (em horas)	Número de indivíduos
$0 \vdash\!\dashv 100$	2
$100 \dashv 200$	5
$200 \dashv 300$	10
$300 \dashv 400$	16
$400 \dashv 500$	9
$500 \dashv 600$	7
$600 \dashv 700$	4
$700 \dashv 800$	1

Uma estimativa da probabilidade de sobrevivência no tempo $t = 400$ horas é dada por

$$\begin{aligned}\widehat{S}(400) &= \frac{\text{N}^{\underline{\text{o}}} \text{ de indivíduos que não falharam até o tempo } t = 400}{n} \\ &= \frac{9+7+4+1}{54} = \frac{21}{54} = 0,389.\end{aligned}$$

Similarmente,

$$\widehat{S}(500) = \frac{7+4+1}{54} = \frac{12}{54} = 0,222.$$

Estimador de Kaplan-Meier (1958)

n unidades experimentais no estudo, dados com censura à direita.

- $t_{(1)} < t_{(2)} < \cdots < t_{(k)}$: k $(k \leq n)$ tempos de falha distintos e ordenados,

Estimador de Kaplan-Meier (1958)

n unidades experimentais no estudo, dados com censura à direita.

- $t_{(1)} < t_{(2)} < \cdots < t_{(k)}$: k $(k \leq n)$ tempos de falha distintos e ordenados,
- d_j: número de falhas em $t_{(j)}$, $j = 1, \ldots, k$ e
- n_j : o número de indivíduos em risco no instante $t_{(j)}$, ou seja, os indivíduos que não falharam e não foram censurados até o instante imediatamente anterior a $t_{(j)}$.

O estimador do limite do produto ou de Kaplan-Meier (K-M) é dado por

$$\widehat{S}(t) = \prod_{j:t_{(j)}<t} \frac{n_j - d_j}{n_j} = \prod_{j:t_{(j)}<t} \left(1 - \frac{d_j}{n_j}\right). \qquad (2)$$

Propriedades do estimador K-M

Resultados assintóticos significam que $n \to \infty$.

- É um estimador consistente de $S(t)$ sob certas condições.

Propriedades do estimador K-M

Resultados assintóticos significam que $n \to \infty$.

- É um estimador consistente de $S(t)$ sob certas condições.
- É um estimador de máxima verossimilhança de $S(t)$.

Propriedades do estimador K-M

Resultados assintóticos significam que $n \to \infty$.

- É um estimador consistente de $S(t)$ sob certas condições.
- É um estimador de máxima verossimilhança de $S(t)$.
- Possui distribuição normal assintótica.

Propriedades do estimador K-M

Resultados assintóticos significam que $n \to \infty$.

- É um estimador consistente de $S(t)$ sob certas condições.
- É um estimador de máxima verossimilhança de $S(t)$.
- Possui distribuição normal assintótica.
- A variância de $\widehat{S}(t)$ é estimada por

$$\widehat{Var}(\widehat{S}(t)) = (\widehat{S}(t))^2 \sum_{j:t_{(j)}<t} \frac{d_j}{n_j(n_j - d_j)}. \text{ (Greenwood)}$$

Propriedades do estimador K-M

Resultados assintóticos significam que $n \to \infty$.

- É um estimador consistente de $S(t)$ sob certas condições.
- É um estimador de máxima verossimilhança de $S(t)$.
- Possui distribuição normal assintótica.
- A variância de $\widehat{S}(t)$ é estimada por

$$\widehat{Var}(\widehat{S}(t)) = (\widehat{S}(t))^2 \sum_{j:t_{(j)}<t} \frac{d_j}{n_j(n_j - d_j)}. \text{ (Greenwood)}$$

- Intervalo de $100(1-\alpha)\%$ de confiança para $S(t)$ tem limites

$$\widehat{S}(t) - z_{\alpha/2}\sqrt{\widehat{Var}(\widehat{S}(t))} \quad \text{e} \quad \widehat{S}(t) + z_{\alpha/2}\sqrt{\widehat{Var}(\widehat{S}(t))},$$

em que $z_{\alpha/2}$ denota o quantil $1-\alpha/2$ da distribuição $N(0,1)$.

Intervalo de confiança de Kalbfleish & Prentice (1980)

Considere a seguinte transformação:

- $$\widehat{U}(t) = \log[-\log(\widehat{S}(t))] \in \mathbb{R}.$$

Intervalo de confiança de Kalbfleish & Prentice (1980)

Considere a seguinte transformação:

- $$\widehat{U}(t) = \log[-\log(\widehat{S}(t))] \in \mathbb{R}.$$
- Variância assintótica:

$$\widehat{Var}(\widehat{U}(t)) = \frac{\sum_{j:t_{(j)}<t} \frac{d_j}{n_j(n_j-d_j)}}{\left[\sum_{j:t_{(j)}<t} \log\left(\frac{n_j-d_j}{n_j}\right)\right]^2} = \frac{\sum_{j:t_{(j)}<t} \frac{d_j}{n_j(n_j-d_j)}}{(\log(\widehat{S}(t)))^2}.$$

Intervalo de confiança de Kalbfleish & Prentice (1980)

Considere a seguinte transformação:

- $$\widehat{U}(t) = \log[-\log(\widehat{S}(t))] \in \mathbb{R}.$$
- Variância assintótica:

$$\widehat{Var}(\widehat{U}(t)) = \frac{\sum_{j:t_{(j)}<t} \frac{d_j}{n_j(n_j-d_j)}}{\left[\sum_{j:t_{(j)}<t} \log\left(\frac{n_j-d_j}{n_j}\right)\right]^2} = \frac{\sum_{j:t_{(j)}<t} \frac{d_j}{n_j(n_j-d_j)}}{(\log(\widehat{S}(t)))^2}.$$

- Intervalo de $100(1-\alpha)\%$ de confiança para $S(t)$ tem limites

$$[\widehat{S}(t)]^{\exp\left\{z_{\alpha/2}\sqrt{\widehat{Var}(\widehat{U}(t))}\right\}} \quad \text{e} \quad [\widehat{S}(t)]^{\exp\left\{-z_{\alpha/2}\sqrt{\widehat{Var}(\widehat{U}(t))}\right\}}.$$

Exemplo

Considere os tempos de reincidência de 10 pacientes com certo tipo de tumor, 6 pacientes apresentam reincidência aos 3; 6,5; 6,5; 10 ; 12 e 15 meses. Um deles perdeu-se contato aos 8,4 meses e 3 pacientes permanecem sem reincidência no final do estudo após 4; 5,7 e 10 meses de observação

Observação

- *As observações censuradas entram na ordenação,*
- *Se censura e falha ocorreram ao mesmo tempo convenciona-se ordenar o tempo censurado depois do tempo de falha,*
- *Considera tantos intervalos de tempo quantos forem o número de falhas distintas,*
- *Os limites dos intervalos de tempo são os tempos de falha.*

Ordenando os dados

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15
$t_{(1)}$	-	-	$t_{(2)}$	$t_{(2)}$	-	$t_{(3)}$	-	$t_{(4)}$	$t_{(5)}$

Tabela 1: Estimativa de Kaplan-Meier do exemplo.

Tempos	Intervalo	n_j	d_j	$(1 - \frac{d_j}{n_j})$	
0	[0,3)				
3	[3,6,5)				
6,5	[6,5,10)				
10	[10,12)				
12	[12,15)				
15	[15,∞)				

Ordenando os dados

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15
$t_{(1)}$	-	-	$t_{(2)}$	$t_{(2)}$	-	$t_{(3)}$	-	$t_{(4)}$	$t_{(5)}$

Tabela 1: Estimativa de Kaplan-Meier do exemplo.

Tempos	Intervalo	n_j	d_j	$(1 - \frac{d_j}{n_j})$	
0	[0,3)	10	0	$1 - \frac{0}{10} = 1$	
3	[3,6,5)				
6,5	[6,5,10)				
10	[10,12)				
12	[12,15)				
15	$[15,\infty)$				

Ordenando os dados

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15
$t_{(1)}$	-	-	$t_{(2)}$	$t_{(2)}$	-	$t_{(3)}$	-	$t_{(4)}$	$t_{(5)}$

Tabela 1: Estimativa de Kaplan-Meier do exemplo.

Tempos	Intervalo	n_j	d_j	$(1 - \frac{d_j}{n_j})$	
0	[0,3)	10	0	$1 - \frac{0}{10} = 1$	
3	[3,6,5)	10	1	$1 - \frac{1}{10} = \frac{9}{10}$	
6,5	[6,5,10)				
10	[10,12)				
12	[12,15)				
15	[15,∞)				

Ordenando os dados

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15
$t_{(1)}$	-	-	$t_{(2)}$	$t_{(2)}$	-	$t_{(3)}$	-	$t_{(4)}$	$t_{(5)}$

Tabela 1: Estimativa de Kaplan-Meier do exemplo.

Tempos	Intervalo	n_j	d_j	$(1 - \frac{d_j}{n_j})$	
0	[0,3)	10	0	$1 - \frac{0}{10} = 1$	
3	[3,6,5)	10	1	$1 - \frac{1}{10} = \frac{9}{10}$	
6,5	[6,5,10)	7	2	$1 - \frac{2}{7} = \frac{5}{7}$	
10	[10,12)				
12	[12,15)				
15	$[15,\infty)$				

Ordenando os dados

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15
$t_{(1)}$	-	-	$t_{(2)}$	$t_{(2)}$	-	$t_{(3)}$	-	$t_{(4)}$	$t_{(5)}$

Tabela 1: Estimativa de Kaplan-Meier do exemplo.

Tempos	Intervalo	n_j	d_j	$(1 - \frac{d_j}{n_j})$	
0	[0,3)	10	0	$1 - \frac{0}{10} = 1$	
3	[3,6,5)	10	1	$1 - \frac{1}{10} = \frac{9}{10}$	
6,5	[6,5,10)	7	2	$1 - \frac{2}{7} = \frac{5}{7}$	
10	[10,12)	4	1	$1 - \frac{1}{4} = \frac{3}{4}$	
12	[12,15)				
15	[15,∞)				

Ordenando os dados

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15
$t_{(1)}$	-	-	$t_{(2)}$	$t_{(2)}$	-	$t_{(3)}$	-	$t_{(4)}$	$t_{(5)}$

Tabela 1: Estimativa de Kaplan-Meier do exemplo.

Tempos	Intervalo	n_j	d_j	$(1 - \frac{d_j}{n_j})$	
0	[0,3)	10	0	$1 - \frac{0}{10} = 1$	
3	[3,6,5)	10	1	$1 - \frac{1}{10} = \frac{9}{10}$	
6,5	[6,5,10)	7	2	$1 - \frac{2}{7} = \frac{5}{7}$	
10	[10,12)	4	1	$1 - \frac{1}{4} = \frac{3}{4}$	
12	[12,15)	2	1	$1 - \frac{1}{2} = \frac{1}{2}$	
15	[15,∞)				

Ordenando os dados

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15
$t_{(1)}$	-	-	$t_{(2)}$	$t_{(2)}$	-	$t_{(3)}$	-	$t_{(4)}$	$t_{(5)}$

Tabela 1: Estimativa de Kaplan-Meier do exemplo.

Tempos	Intervalo	n_j	d_j	$(1 - \frac{d_j}{n_j})$	
0	[0,3)	10	0	$1 - \frac{0}{10} = 1$	
3	[3,6,5)	10	1	$1 - \frac{1}{10} = \frac{9}{10}$	
6,5	[6,5,10)	7	2	$1 - \frac{2}{7} = \frac{5}{7}$	
10	[10,12)	4	1	$1 - \frac{1}{4} = \frac{3}{4}$	
12	[12,15)	2	1	$1 - \frac{1}{2} = \frac{1}{2}$	
15	[15,∞)	1	1	$1 - \frac{1}{1} = 0$	

Ordenando os dados

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15
$t_{(1)}$	-	-	$t_{(2)}$	$t_{(2)}$	-	$t_{(3)}$	-	$t_{(4)}$	$t_{(5)}$

Tabela 1: Estimativa de Kaplan-Meier do exemplo.

Tempos	Intervalo	n_j	d_j	$(1 - \frac{d_j}{n_j})$	$\widehat{S}(t)$	$\widehat{F}(t)$
0	[0,3)	10	0	$1 - \frac{0}{10} = 1$		
3	[3,6,5)	10	1	$1 - \frac{1}{10} = \frac{9}{10}$		
6,5	[6,5,10)	7	2	$1 - \frac{2}{7} = \frac{5}{7}$		
10	[10,12)	4	1	$1 - \frac{1}{4} = \frac{3}{4}$		
12	[12,15)	2	1	$1 - \frac{1}{2} = \frac{1}{2}$		
15	$[15,\infty)$	1	1	$1 - \frac{1}{1} = 0$		

Ordenando os dados

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15
$t_{(1)}$	-	-	$t_{(2)}$	$t_{(2)}$	-	$t_{(3)}$	-	$t_{(4)}$	$t_{(5)}$

Tabela 1: Estimativa de Kaplan-Meier do exemplo.

Tempos	Intervalo	n_j	d_j	$(1 - \frac{d_j}{n_j})$	$\widehat{S}(t)$	$\widehat{F}(t)$
0	[0,3)	10	0	$1 - \frac{0}{10} = 1$	1	0
3	[3,6,5)	10	1	$1 - \frac{1}{10} = \frac{9}{10}$		
6,5	[6,5,10)	7	2	$1 - \frac{2}{7} = \frac{5}{7}$		
10	[10,12)	4	1	$1 - \frac{1}{4} = \frac{3}{4}$		
12	[12,15)	2	1	$1 - \frac{1}{2} = \frac{1}{2}$		
15	$[15,\infty)$	1	1	$1 - \frac{1}{1} = 0$		

Ordenando os dados

3	4^{+}	$5{,}7^{+}$	6,5	6,5	$8{,}4^{+}$	10	10^{+}	12	15
$t_{(1)}$	-	-	$t_{(2)}$	$t_{(2)}$	-	$t_{(3)}$	-	$t_{(4)}$	$t_{(5)}$

Tabela 1: Estimativa de Kaplan-Meier do exemplo.

Tempos	Intervalo	n_j	d_j	$(1 - \frac{d_j}{n_j})$	$\widehat{S}(t)$	$\widehat{F}(t)$
0	[0,3)	10	0	$1 - \frac{0}{10} = 1$	1	0
3	[3,6,5)	10	1	$1 - \frac{1}{10} = \frac{9}{10}$	$1 \times \frac{9}{10} = 0,9$	0,10
6,5	[6,5,10)	7	2	$1 - \frac{2}{7} = \frac{5}{7}$		
10	[10,12)	4	1	$1 - \frac{1}{4} = \frac{3}{4}$		
12	[12,15)	2	1	$1 - \frac{1}{2} = \frac{1}{2}$		
15	[15,∞)	1	1	$1 - \frac{1}{1} = 0$		

Ordenando os dados

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15
$t_{(1)}$	-	-	$t_{(2)}$	$t_{(2)}$	-	$t_{(3)}$	-	$t_{(4)}$	$t_{(5)}$

Tabela 1: Estimativa de Kaplan-Meier do exemplo.

Tempos	Intervalo	n_j	d_j	$(1 - \frac{d_j}{n_j})$	$\widehat{S}(t)$	$\widehat{F}(t)$
0	[0,3)	10	0	$1 - \frac{0}{10} = 1$	1	0
3	[3,6,5)	10	1	$1 - \frac{1}{10} = \frac{9}{10}$	$1 \times \frac{9}{10} = 0,9$	0,10
6,5	[6,5,10)	7	2	$1 - \frac{2}{7} = \frac{5}{7}$	$0,9 \times \frac{5}{7} = 0,6429$	0,3571
10	[10,12)	4	1	$1 - \frac{1}{4} = \frac{3}{4}$		
12	[12,15)	2	1	$1 - \frac{1}{2} = \frac{1}{2}$		
15	[15,∞)	1	1	$1 - \frac{1}{1} = 0$		

Ordenando os dados

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15
$t_{(1)}$	-	-	$t_{(2)}$	$t_{(2)}$	-	$t_{(3)}$	-	$t_{(4)}$	$t_{(5)}$

Tabela 1: Estimativa de Kaplan-Meier do exemplo.

Tempos	Intervalo	n_j	d_j	$(1 - \frac{d_j}{n_j})$	$\widehat{S}(t)$	$\widehat{F}(t)$
0	[0,3)	10	0	$1 - \frac{0}{10} = 1$	1	0
3	[3,6,5)	10	1	$1 - \frac{1}{10} = \frac{9}{10}$	$1 \times \frac{9}{10} = 0,9$	0,10
6,5	[6,5,10)	7	2	$1 - \frac{2}{7} = \frac{5}{7}$	$0,9 \times \frac{5}{7} = 0,6429$	0,3571
10	[10,12)	4	1	$1 - \frac{1}{4} = \frac{3}{4}$	$0,6429 \times \frac{3}{4} = 0,4822$	0,5178
12	[12,15)	2	1	$1 - \frac{1}{2} = \frac{1}{2}$		
15	[15,∞)	1	1	$1 - \frac{1}{1} = 0$		

Ordenando os dados

3	4^{+}	$5{,}7^{+}$	6,5	6,5	$8{,}4^{+}$	10	10^{+}	12	15
$t_{(1)}$	-	-	$t_{(2)}$	$t_{(2)}$	-	$t_{(3)}$	-	$t_{(4)}$	$t_{(5)}$

Tabela 1: Estimativa de Kaplan-Meier do exemplo.

Tempos	Intervalo	n_j	d_j	$(1 - \frac{d_j}{n_j})$	$\widehat{S}(t)$	$\widehat{F}(t)$
0	[0,3)	10	0	$1 - \frac{0}{10} = 1$	1	0
3	[3,6,5)	10	1	$1 - \frac{1}{10} = \frac{9}{10}$	$1 \times \frac{9}{10} = 0,9$	0,10
6,5	[6,5,10)	7	2	$1 - \frac{2}{7} = \frac{5}{7}$	$0,9 \times \frac{5}{7} = 0,6429$	0,3571
10	[10,12)	4	1	$1 - \frac{1}{4} = \frac{3}{4}$	$0,6429 \times \frac{3}{4} = 0,4822$	0,5178
12	[12,15)	2	1	$1 - \frac{1}{2} = \frac{1}{2}$	$0,4822 \times \frac{1}{2} = 0,2411$	0,7589
15	[15,∞)	1	1	$1 - \frac{1}{1} = 0$		

Ordenando os dados

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15
$t_{(1)}$	-	-	$t_{(2)}$	$t_{(2)}$	-	$t_{(3)}$	-	$t_{(4)}$	$t_{(5)}$

Tabela 1: Estimativa de Kaplan-Meier do exemplo.

Tempos	Intervalo	n_j	d_j	$(1 - \frac{d_j}{n_j})$	$\widehat{S}(t)$	$\widehat{F}(t)$
0	[0,3)	10	0	$1 - \frac{0}{10} = 1$	1	0
3	[3,6,5)	10	1	$1 - \frac{1}{10} = \frac{9}{10}$	$1 \times \frac{9}{10} = 0,9$	0,10
6,5	[6,5,10)	7	2	$1 - \frac{2}{7} = \frac{5}{7}$	$0,9 \times \frac{5}{7} = 0,6429$	0,3571
10	[10,12)	4	1	$1 - \frac{1}{4} = \frac{3}{4}$	$0,6429 \times \frac{3}{4} = 0,4822$	0,5178
12	[12,15)	2	1	$1 - \frac{1}{2} = \frac{1}{2}$	$0,4822 \times \frac{1}{2} = 0,2411$	0,7589
15	[15,∞)	1	1	$1 - \frac{1}{1} = 0$	$0,2411 \times 0 = 0$	1

Assim, temos que

$$\widehat{S}(t) = \begin{cases} 1, & \text{se } 0 \leq t < 3 \\ 0,9000, & \text{se } 3 \leq t < 6,5 \\ 0,6429, & \text{se } 6,5 \leq t < 10 \\ 0,4822, & \text{se } 10 \leq t < 12 \\ 0,2411, & \text{se } 12 \leq t < 15 \\ 0, & \text{se } t \geq 15 \end{cases}$$

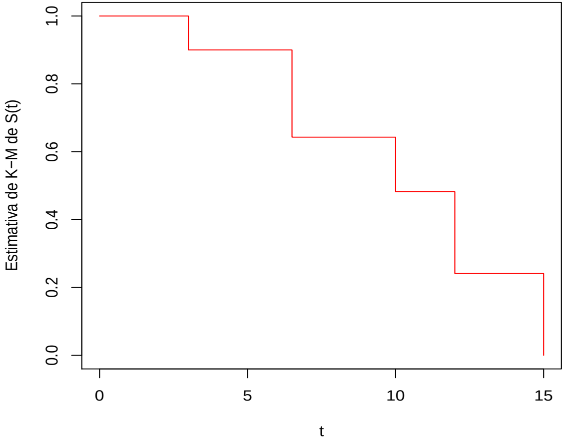


Figura 1: Representação gráfica da estimativa de K-M de $S(t)$.

Qual é a probabilidade de o paciente sobreviver após 6,5 meses?

Qual é a probabilidade de o paciente sobreviver após 6,5 meses?

$$\widehat{S}(6,5) = 0,6429(64,29\%)$$

- A estimativa da variância de $\widehat{S}(6,5)$

$$\begin{aligned}\widehat{Var}(\widehat{S}(t)) &= (\widehat{S}(6,5))^2 \sum_{j:t_{(j)}<6,5} \frac{d_j}{n_j(n_j - d_j)} \\ &= (0,6429)^2 \left(\frac{1}{(10)(9)} + \frac{2}{(7)(6)} \right) = 0,0282\end{aligned}$$

- Um intervalo de 95% de confiança para $S(6,5)$ é dado por

$$(0,314; 0,972).$$

- Intervalo de confiança de Kalbfleish & Prentice (1980)

$$(0,2740; 0,8600)$$

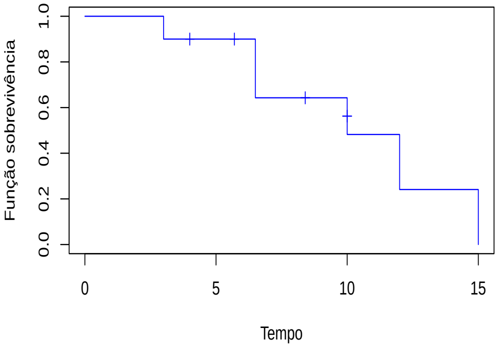
Estimativa de Kaplan-Meier em R

- library(survival)
- tempo = c(3,4,5.7,6.5,6.5,8.4,10,10,12,15)
- censura = c(1,0,0,1,1,0,1,0,1,1)
- mKM = survfit(Surv(tempo, censura) ˜ 1, se.fit = FALSE)
- summary(mKM)

time	n.risk	n.event	survival
3.0	10	1	0.900
6.5	7	2	0.643
10.0	4	1	0.482
12.0	2	1	0.241
15.0	1	1	0.000

Estimativa de Kaplan-Meier com R

plot(mKM, mark.time = TRUE, xlab = "Tempo", ylab = "Função sobrevivência", col = "blue")



Estimativa do tempo médio de vida (TMV)

A estimativa do TMV consiste em obter a área sob a curva de Kaplan-Meier, que por ser uma função escada é dada por

$$\widehat{TMV} = t_{(1)} + \sum_{j=1}^{k-1} \widehat{S}(t_{(j)})[t_{(j+1)} - t_{(j)}].$$

Estimativa do tempo médio de vida (TMV)

A estimativa do TMV consiste em obter a área sob a curva de Kaplan-Meier, que por ser uma função escada é dada por

$$\widehat{TMV} = t_{(1)} + \sum_{j=1}^{k-1} \widehat{S}(t_{(j)})[t_{(j+1)} - t_{(j)}].$$

Kaplan & Meier (1958) mostraram que a variância de $\widehat{TMV}$ pode ser estimada por

$$\widehat{Var}(\widehat{TMV}) = \frac{k}{k-1} \sum_{j=1}^{k-1} \frac{A_j^2}{n_j(n_j - d_j)},$$

em que $A_j^2 = \widehat{S}(t_{(j)})[t_{(j+1)-t_{(j)}}] + \cdots + \widehat{S}(t_{(k-1)})[t_{(k)} - t_{(k-1)}]$, $j = 1, \ldots, k-1$.

Exemplo

No exemplo anterior temos que

$$\begin{aligned}\widehat{TMV} &= 3 + 0.9(6,5-3) + 0,6429(10-6,5) \\ &\quad +0,4822(12-10) + 0,2411(15-12) = 10,09 \text{ } meses.\end{aligned}$$

A estimativa de $Var\widehat{TMV}$ é

$$\begin{aligned}\widehat{Var}(\widehat{TMV}) &= \frac{6}{5}\left[\frac{A_1^2}{(10)(9)} + \frac{A_2^2}{(7)(6)} + \frac{A_3^2}{(6)(5)} + \frac{A_4^2}{(4)(3)} + \frac{A_5^2}{(2)(1)}\right] \\ &= 2,33\,\text{meses}^2.\end{aligned}$$

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15
$t_{(1)}$	-	-	$t_{(2)}$	$t_{(3)}$	-	$t_{(4)}$	-	$t_{(5)}$	$t_{(6)}$

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15
$t_{(1)}$	-	-	$t_{(2)}$	$t_{(3)}$	-	$t_{(4)}$	-	$t_{(5)}$	$t_{(6)}$

- $$A_1 = \widehat{S(t_{(1)})}[t_{(2)} - t_{(1)}] + \cdots + \widehat{(S)(t_{(5)})}[t_{(6)} - t_{(5)}]$$

$$\begin{aligned} A_1 &= (0,9)[6,5-3] + (0,6429)[6,5-6,5] + (0,6429)[10-6,5] \\ & 0,4822[12-10] + 0,2411[15-12] = 7,0877 \end{aligned}$$

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15
$t_{(1)}$	-	-	$t_{(2)}$	$t_{(3)}$	-	$t_{(4)}$	-	$t_{(5)}$	$t_{(6)}$

- $$A_1 = \widehat{S(t_{(1)})}[t_{(2)} - t_{(1)}] + \cdots + (\widehat{S)(t_{(5)}})[t_{(6)} - t_{(5)}]$$

$$\begin{aligned} A_1 &= (0,9)[6,5-3] + (0,6429)[6,5-6,5] + (0,6429)[10-6,5] \\ & 0,4822[12-10] + 0,2411[15-12] = 7,0877 \end{aligned}$$

- $A_2 = (0,6429)[6,5-6,5] + (0,6429)[10-6,5] + 0,4822[12-10] + 0,2411[15-12] = 3,9377 = A_3,$

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15
$t_{(1)}$	-	-	$t_{(2)}$	$t_{(3)}$	-	$t_{(4)}$	-	$t_{(5)}$	$t_{(6)}$

- $$A_1 = \widehat{S(t_{(1)})}[t_{(2)} - t_{(1)}] + \cdots + \widehat{(S)(t_{(5)})}[t_{(6)} - t_{(5)}]$$

$$\begin{aligned} A_1 &= (0,9)[6,5-3] + (0,6429)[6,5-6,5] + (0,6429)[10-6,5] \\ & 0,4822[12-10] + 0,2411[15-12] = 7,0877 \end{aligned}$$

- $A_2 = (0,6429)[6,5-6,5] + (0,6429)[10-6,5] + 0,4822[12-10] + 0,2411[15-12] = 3,9377 = A_3,$
- $A_4 = +0,4822[12-10] + 0,2411[15-12] = 1,6877$

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15
$t_{(1)}$	-	-	$t_{(2)}$	$t_{(3)}$	-	$t_{(4)}$	-	$t_{(5)}$	$t_{(6)}$

- $$A_1 = \widehat{S(t_{(1)})}[t_{(2)} - t_{(1)}] + \cdots + (\widehat{S)(t_{(5)})}[t_{(6)} - t_{(5)}]$$

$$\begin{aligned} A_1 &= (0,9)[6,5-3] + (0,6429)[6,5-6,5] + (0,6429)[10-6,5] \\ & \quad 0,4822[12-10] + 0,2411[15-12] = 7,0877 \end{aligned}$$

- $A_2 = (0,6429)[6,5-6,5] + (0,6429)[10-6,5] + 0,4822[12-10] + 0,2411[15-12] = 3,9377 = A_3$,
- $A_4 = +0,4822[12-10] + 0,2411[15-12] = 1,6877$
- $A_5 = 0,2411[15-12] = 0,7233.$

Estimativa do tempo médio de vida residual (VMR)

O tempo médio de vida restante é de interesse daqueles pacientes que se encontram livres do evento em um determinado tempo t. Quanto tempo, em média, eles ainda têm de vida após esse tempo? No exemplo, estima-se, por exemplo, para os pacientes que sobreviveram até o tempo $t = 10$ meses, o tempo médio de vida restante.
Da definição de VMR, uma estimativa baseada na estimativa de K-M de $S(t)$ é dada por

$$\begin{aligned}
\widehat{VMR} &= \frac{\text{Área sob a curva } \widehat{S}(\text{t}) \text{ à direita de } t = 10}{\widehat{S}(10)} \\
&= \frac{0,4822(12-10) + 0,2411(15-12)}{0,4822} = 3,5 \text{ meses.}
\end{aligned}$$

Estimador atuarial ou tábua de vida

- Suponha que o eixo do tempo seja dividido em s intervalos definidos pelos pontos de corte $t_0 < t_1 < t_2 < \cdots < t_s$, com intervalos $I_j = [t_{j-1}, t_j)$, $j = 1, \ldots, s$, em que $t_0 = 0$ e $t_s = \infty$.
- d_j : o número de falhas no intervalo I_j, $j = 1, \ldots, s$.
- $$n_j^* = \underbrace{\begin{array}{c}\text{númeo de indivíduos}\\ \text{em risco em } t_{j-1}\end{array}}_{n_j} - 0{,}5 \times \underbrace{\begin{array}{c}\text{númeo de indivíduos}\\ \text{censurados em } I_j\end{array}}_{c_j}$$
- Defina
$$q_j = \frac{d_j}{n_j^*}$$
- O estimador tábua de vida é dado por
$$\widehat{S}(t) = \prod_{i=1}^{j} (1 - q_{i-1}), \; t \in I_j, j = 1, \ldots, s, \; e \; q_0 = 0.$$

Estimador atuarial ou tábua de vida

Considere os dados do exemplo anterior.

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 2: Estimativa atuarial (ou tábua de vida) do exemplo.

j	Intervalo	n_j	d_j	c_j	n_j^*	
1	[0,3)					
2	[3,6,5)					
3	[6,5,10)					
4	[10,12)					
5	[12,15)					
6	[15,∞)					

Estimador atuarial ou tábua de vida

Considere os dados do exemplo anterior.

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 2: Estimativa atuarial (ou tábua de vida) do exemplo.

j	Intervalo	n_j	d_j	c_j	n_j^*	
1	[0,3)	10	0	0		
2	[3,6,5)					
3	[6,5,10)					
4	[10,12)					
5	[12,15)					
6	[15,∞)					

Estimador atuarial ou tábua de vida

Considere os dados do exemplo anterior.

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 2: Estimativa atuarial (ou tábua de vida) do exemplo.

j	Intervalo	n_j	d_j	c_j	n_j^*	
1	[0,3)	10	0	0		
2	[3,6,5)	10	1	2		
3	[6,5,10)					
4	[10,12)					
5	[12,15)					
6	[15,∞)					

Estimador atuarial ou tábua de vida

Considere os dados do exemplo anterior.

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 2: Estimativa atuarial (ou tábua de vida) do exemplo.

j	Intervalo	n_j	d_j	c_j	n_j^*	
1	[0,3)	10	0	0		
2	[3,6,5)	10	1	2		
3	[6,5,10)	7	2	1		
4	[10,12)					
5	[12,15)					
6	[15,∞)					

Estimador atuarial ou tábua de vida

Considere os dados do exemplo anterior.

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 2: Estimativa atuarial (ou tábua de vida) do exemplo.

j	Intervalo	n_j	d_j	c_j	n_j^*	
1	[0,3)	10	0	0		
2	[3,6,5)	10	1	2		
3	[6,5,10)	7	2	1		
4	[10,12)	4	1	1		
5	[12,15)					
6	[15,∞)					

Estimador atuarial ou tábua de vida

Considere os dados do exemplo anterior.

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 2: Estimativa atuarial (ou tábua de vida) do exemplo.

j	Intervalo	n_j	d_j	c_j	n_j^*	
1	[0,3)	10	0	0		
2	[3,6,5)	10	1	2		
3	[6,5,10)	7	2	1		
4	[10,12)	4	1	1		
5	[12,15)	2	1	0		
6	$[15,\infty)$					

Estimador atuarial ou tábua de vida

Considere os dados do exemplo anterior.

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 2: Estimativa atuarial (ou tábua de vida) do exemplo.

j	Intervalo	n_j	d_j	c_j	n_j^*	
1	[0,3)	10	0	0		
2	[3,6,5)	10	1	2		
3	[6,5,10)	7	2	1		
4	[10,12)	4	1	1		
5	[12,15)	2	1	0		
6	[15,∞)	1	1	0		

Estimador atuarial ou tábua de vida

Considere os dados do exemplo anterior.

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 2: Estimativa atuarial (ou tábua de vida) do exemplo.

j	Intervalo	n_j	d_j	c_j	n_j^*	q_j	$1 - q_j$	$\widehat{S}(t)$
1	[0,3)	10	0	0				
2	[3,6,5)	10	1	2				
3	[6,5,10)	7	2	1				
4	[10,12)	4	1	1				
5	[12,15)	2	1	0				
6	[15,∞)	1	1	0				

Estimador atuarial ou tábua de vida

Considere os dados do exemplo anterior.

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 2: Estimativa atuarial (ou tábua de vida) do exemplo.

j	Intervalo	n_j	d_j	c_j	n_j^*	q_j	$1 - q_j$	$\widehat{S}(t)$
1	[0,3)	10	0	0	10-0,5(0)=10	0	1	1
2	[3,6,5)	10	1	2				
3	[6,5,10)	7	2	1				
4	[10,12)	4	1	1				
5	[12,15)	2	1	0				
6	[15,∞)	1	1	0				

Estimador atuarial ou tábua de vida

Considere os dados do exemplo anterior.

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 2: Estimativa atuarial (ou tábua de vida) do exemplo.

j	Intervalo	n_j	d_j	c_j	n_j^*	q_j	$1 - q_j$	$\widehat{S}(t)$
1	[0,3)	10	0	0	10-0,5(0)=10	0	1	1
2	[3,6,5)	10	1	2	10-0,5(2)=9	1/9	0,889	1× 0,889=0,889
3	[6,5,10)	7	2	1	7-0,5(1)=6,5	2/6,5	0.6923	1× 0,889× 0,6923=0,6155
4	[10,12)	4	1	1				
5	[12,15)	2	1	0				
6	[15,∞)	1	1	0				

Estimador atuarial ou tábua de vida

Considere os dados do exemplo anterior.

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 2: Estimativa atuarial (ou tábua de vida) do exemplo.

j	Intervalo	n_j	d_j	c_j	n_j^*	q_j	$1 - q_j$	$\widehat{S}(t)$
1	[0,3)	10	0	0	10-0,5(0)=10	0	1	1
2	[3,6,5)	10	1	2	10-0,5(2)=9	1/9	0,889	$1\times$ 0,889=0,889
3	[6,5,10)	7	2	1	7-0,5(1)=6,5	2/6,5	0.6923	$1\times$ 0,889$\times$ 0,6923=0,6155
4	[10,12)	4	1	1	4-0,5(1)=3,5	1/3,5	0.7143	$1\times$ 0,889$\times$ 0,6923$\times$ 0.7143=0,4397
5	[12,15)	2	1	0				
6	[15,∞)	1	1	0				

Estimador atuarial ou tábua de vida

Considere os dados do exemplo anterior.

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 2: Estimativa atuarial (ou tábua de vida) do exemplo.

j	Intervalo	n_j	d_j	c_j	n_j^*	q_j	$1 - q_j$	$\widehat{S}(t)$
1	[0,3)	10	0	0	10-0,5(0)=10	0	1	1
2	[3,6,5)	10	1	2	10-0,5(2)=9	1/9	0,889	$1\times$ 0,889=0,889
3	[6,5,10)	7	2	1	7-0,5(1)=6,5	2/6,5	0.6923	$1\times$ 0,889$\times$ 0,6923=0,6155
4	[10,12)	4	1	1	4-0,5(1)=3,5	1/3,5	0.7143	$1\times$ 0,889$\times$ 0,6923$\times$ 0.7143=0,4397
5	[12,15)	2	1	0	2-0,5(0)=2	1/2	0,5	$1\times$ 0,889$\times$ 0,6923$\times$ 0.7143$\times$ 0,5=0.2148
6	[15,∞)	1	1	0				

Estimador atuarial ou tábua de vida

Considere os dados do exemplo anterior.

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 2: Estimativa atuarial (ou tábua de vida) do exemplo.

j	Intervalo	n_j	d_j	c_j	n_j^*	q_j	$1-q_j$	$\widehat{S}(t)$
1	[0,3)	10	0	0	10-0,5(0)=10	0	1	1
2	[3,6,5)	10	1	2	10-0,5(2)=9	1/9	0,889	$1\times$ 0,889=0,889
3	[6,5,10)	7	2	1	7-0,5(1)=6,5	2/6,5	0.6923	$1\times$ 0,889$\times$ 0,6923=0,6155
4	[10,12)	4	1	1	4-0,5(1)=3,5	1/3,5	0.7143	$1\times$ 0,889$\times$ 0,6923$\times$ 0.7143=0,4397
5	[12,15)	2	1	0	2-0,5(0)=2	1/2	0,5	$1\times$ 0,889$\times$ 0,6923$\times$ 0.7143$\times$ 0,5=0.2148
6	[15,∞)	1	1	0	1-0,5(0)=1	1	0	0

Estimador atuarial ou tábua de vida

Assim, temos que a estimativa atuarial da função sobrevivência para o exemplo é dada por

$$\widehat{S}(t) = \begin{cases} 1, & \text{se } 0 \leq t < 3 \\ 0,889, & \text{se } 5 \leq t < 6,5 \\ 0,6155, & \text{se } 6,5 \leq t < 10 \\ 0,4397, & \text{se } 10 \leq t < 12 \\ 0,2148, & \text{se } 12 \leq t < 15 \\ 0, & \text{se } t \geq 15 \end{cases}$$

Estimador de Nelson-Aalen

- O estimador de Nelson-Aalen baseia-se na função sobrevivência dada por

$$S(t) = \exp\{-H(t)\},$$

em que $H(t)$ é a função risco acumulado.

- Nelson (1972) e, posteriormente, Aalen (1978) mostraram propriedades assintóticas do estimador de $H(t)$ dado por

$$\widehat{H}(t) = \sum_{j:t_{(j)}<t} \frac{d_j}{n_j},$$

em que d_j: número de falhas em $t_{(j)}$ e n_j : número de indivíduos em risco em $t_{(j)}$.

Estimador de Nelson-Aalen

- A variância do estimador (Aalen, 1978) é dada por

$$Var(\widehat{H}(t)) = \sum_{j:t_{(j)}<t} \frac{d_j}{n_j^2}.$$

- Portanto, o estimador de N-A para a função sobrevivência é dado por

$$\widehat{S}(t) = \exp\{-\widehat{H}(t)\}$$

- A variância do estimador de Nelson-Aalen de $S(t)$ é dado por

$$Var(\widehat{S}(t)) = (\widehat{S}(t))^2 \sum_{j:t_{(j)}<t} \frac{d_j}{n_j^2}.$$

Estimador de Nelson-Aalen

Considere os dados do exemplo anterior

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 3: Estimativa de Nelson-Aalen do exemplo.

Tempos	Intervalo	n_j	d_j	$\frac{d_j}{n_j}$	
0	[0,3)				
3	[3,6,5)				
6,5	[6,5,10)				
10	[10,12)				
12	[12,15)				
15	[15,∞)				

Estimador de Nelson-Aalen

Considere os dados do exemplo anterior

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 3: Estimativa de Nelson-Aalen do exemplo.

Tempos	Intervalo	n_j	d_j	$\frac{d_j}{n_j}$	
0	[0,3)	10	0	$\frac{0}{10} = 0$	
3	[3,6,5)				
6,5	[6,5,10)				
10	[10,12)				
12	[12,15)				
15	[15,∞)				

Estimador de Nelson-Aalen

Considere os dados do exemplo anterior

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 3: Estimativa de Nelson-Aalen do exemplo.

Tempos	Intervalo	n_j	d_j	$\frac{d_j}{n_j}$	
0	[0,3)	10	0	$\frac{0}{10} = 0$	
3	[3,6,5)	10	1	$\frac{1}{10}$	
6,5	[6,5,10)				
10	[10,12)				
12	[12,15)				
15	[15,∞)				

Estimador de Nelson-Aalen

Considere os dados do exemplo anterior

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 3: Estimativa de Nelson-Aalen do exemplo.

Tempos	Intervalo	n_j	d_j	$\frac{d_j}{n_j}$	
0	[0,3)	10	0	$\frac{0}{10} = 0$	
3	[3,6,5)	10	1	$\frac{1}{10}$	
6,5	[6,5,10)	7	2	$\frac{2}{7} = \frac{2}{7}$	
10	[10,12)				
12	[12,15)				
15	[15,∞)				

Estimador de Nelson-Aalen

Considere os dados do exemplo anterior

3	4^{+}	$5{,}7^{+}$	6,5	6,5	$8{,}4^{+}$	10	10^{+}	12	15

Tabela 3: Estimativa de Nelson-Aalen do exemplo.

Tempos	Intervalo	n_j	d_j	$\frac{d_j}{n_j}$	
0	[0,3)	10	0	$\frac{0}{10} = 0$	
3	[3,6,5)	10	1	$\frac{1}{10}$	
6,5	[6,5,10)	7	2	$\frac{2}{7} = \frac{2}{7}$	
10	[10,12)	4	1	$\frac{1}{4} = \frac{1}{4}$	
12	[12,15)				
15	[15,∞)				

Estimador de Nelson-Aalen

Considere os dados do exemplo anterior

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 3: Estimativa de Nelson-Aalen do exemplo.

Tempos	Intervalo	n_j	d_j	$\frac{d_j}{n_j}$	
0	[0,3)	10	0	$\frac{0}{10} = 0$	
3	[3,6,5)	10	1	$\frac{1}{10}$	
6,5	[6,5,10)	7	2	$\frac{2}{7} = \frac{2}{7}$	
10	[10,12)	4	1	$\frac{1}{4} = \frac{1}{4}$	
12	[12,15)	2	1	$\frac{1}{2} = \frac{1}{2}$	
15	$[15,\infty)$				

Estimador de Nelson-Aalen

Considere os dados do exemplo anterior

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 3: Estimativa de Nelson-Aalen do exemplo.

Tempos	Intervalo	n_j	d_j	$\frac{d_j}{n_j}$	
0	[0,3)	10	0	$\frac{0}{10} = 0$	
3	[3,6,5)	10	1	$\frac{1}{10}$	
6,5	[6,5,10)	7	2	$\frac{2}{7} = \frac{2}{7}$	
10	[10,12)	4	1	$\frac{1}{4} = \frac{1}{4}$	
12	[12,15)	2	1	$\frac{1}{2} = \frac{1}{2}$	
15	[15,∞)	1	1	$\frac{1}{1} = 1$	

Estimador de Nelson-Aalen

Considere os dados do exemplo anterior

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 3: Estimativa de Nelson-Aalen do exemplo.

Tempos	Intervalo	n_j	d_j	$\frac{d_j}{n_j}$	$\widehat{H}(t)$	$\widehat{S}(t) = \exp\{-\widehat{H}(t)\}$
0	[0,3)	10	0	$\frac{0}{10} = 0$		
3	[3,6,5)	10	1	$\frac{1}{10}$		
6,5	[6,5,10)	7	2	$\frac{2}{7} = \frac{2}{7}$		
10	[10,12)	4	1	$\frac{1}{4} = \frac{1}{4}$		
12	[12,15)	2	1	$\frac{1}{2} = \frac{1}{2}$		
15	$[15,\infty)$	1	1	$\frac{1}{1} = 1$		

Estimador de Nelson-Aalen

Considere os dados do exemplo anterior

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 3: Estimativa de Nelson-Aalen do exemplo.

Tempos	Intervalo	n_j	d_j	$\frac{d_j}{n_j}$	$\widehat{H}(t)$	$\widehat{S}(t) = \exp\{-\widehat{H}(t)\}$
0	[0,3)	10	0	$\frac{0}{10} = 0$	0	1
3	[3,6,5)	10	1	$\frac{1}{10}$		
6,5	[6,5,10)	7	2	$\frac{2}{7} = \frac{2}{7}$		
10	[10,12)	4	1	$\frac{1}{4} = \frac{1}{4}$		
12	[12,15)	2	1	$\frac{1}{2} = \frac{1}{2}$		
15	$[15,\infty)$	1	1	$\frac{1}{1} = 1$		

Estimador de Nelson-Aalen

Considere os dados do exemplo anterior

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 3: Estimativa de Nelson-Aalen do exemplo.

Tempos	Intervalo	n_j	d_j	$\frac{d_j}{n_j}$	$\widehat{H}(t)$	$\widehat{S}(t) = \exp\{-\widehat{H}(t)\}$
0	[0,3)	10	0	$\frac{0}{10} = 0$	0	1
3	[3,6,5)	10	1	$\frac{1}{10}$	$0 + \frac{1}{10} = \frac{1}{10}$	0,9048
6,5	[6,5,10)	7	2	$\frac{2}{7} = \frac{2}{7}$		
10	[10,12)	4	1	$\frac{1}{4} = \frac{1}{4}$		
12	[12,15)	2	1	$\frac{1}{2} = \frac{1}{2}$		
15	$[15,\infty)$	1	1	$\frac{1}{1} = 1$		

Estimador de Nelson-Aalen

Considere os dados do exemplo anterior

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 3: Estimativa de Nelson-Aalen do exemplo.

Tempos	Intervalo	n_j	d_j	$\frac{d_j}{n_j}$	$\widehat{H}(t)$	$\widehat{S}(t) = \exp\{-\widehat{H}(t)\}$
0	[0,3)	10	0	$\frac{0}{10} = 0$	0	1
3	[3,6,5)	10	1	$\frac{1}{10}$	$0 + \frac{1}{10} = \frac{1}{10}$	0,9048
6,5	[6,5,10)	7	2	$\frac{2}{7} = \frac{2}{7}$	$0 + \frac{1}{10} + \frac{2}{7}$	0,3857
10	[10,12)	4	1	$\frac{1}{4} = \frac{1}{4}$		
12	[12,15)	2	1	$\frac{1}{2} = \frac{1}{2}$		
15	[15,∞)	1	1	$\frac{1}{1} = 1$		

Estimador de Nelson-Aalen

Considere os dados do exemplo anterior

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 3: Estimativa de Nelson-Aalen do exemplo.

Tempos	Intervalo	n_j	d_j	$\frac{d_j}{n_j}$	$\widehat{H}(t)$	$\widehat{S}(t) = \exp\{-\widehat{H}(t)\}$
0	[0,3)	10	0	$\frac{0}{10} = 0$	0	1
3	[3,6,5)	10	1	$\frac{1}{10}$	$0 + \frac{1}{10} = \frac{1}{10}$	0,9048
6,5	[6,5,10)	7	2	$\frac{2}{7} = \frac{2}{7}$	$0 + \frac{1}{10} + \frac{2}{7}$	0,3857
10	[10,12)	4	1	$\frac{1}{4} = \frac{1}{4}$	$0 + \frac{1}{10} + \frac{2}{7} + \frac{1}{4} = 0,6357$	0,5229
12	[12,15)	2	1	$\frac{1}{2} = \frac{1}{2}$		
15	[15,∞)	1	1	$\frac{1}{1} = 1$		

Estimador de Nelson-Aalen

Considere os dados do exemplo anterior

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 3: Estimativa de Nelson-Aalen do exemplo.

Tempos	Intervalo	n_j	d_j	$\frac{d_j}{n_j}$	$\widehat{H}(t)$	$\widehat{S}(t) = \exp\{-\widehat{H}(t)\}$
0	[0,3)	10	0	$\frac{0}{10} = 0$	0	1
3	[3,6,5)	10	1	$\frac{1}{10}$	$0 + \frac{1}{10} = \frac{1}{10}$	0,9048
6,5	[6,5,10)	7	2	$\frac{2}{7} = \frac{2}{7}$	$0 + \frac{1}{10} + \frac{2}{7}$	0,3857
10	[10,12)	4	1	$\frac{1}{4} = \frac{1}{4}$	$0 + \frac{1}{10} + \frac{2}{7} + \frac{1}{4} = 0,6357$	0,5229
12	[12,15)	2	1	$\frac{1}{2} = \frac{1}{2}$	$0 + \frac{1}{10} + \frac{2}{7} + \frac{1}{4} + \frac{1}{2} = 1,1357$	0,3212
15	[15,∞)	1	1	$\frac{1}{1} = 1$		

Estimador de Nelson-Aalen

Considere os dados do exemplo anterior

3	4^+	$5{,}7^+$	6,5	6,5	$8{,}4^+$	10	10^+	12	15

Tabela 3: Estimativa de Nelson-Aalen do exemplo.

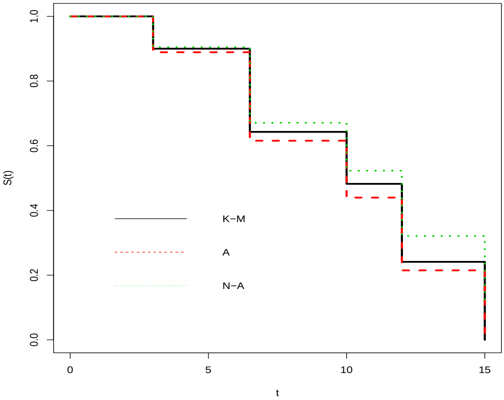
Tempos	Intervalo	n_j	d_j	$\frac{d_j}{n_j}$	$\widehat{H}(t)$	$\widehat{S}(t) = \exp\{-\widehat{H}(t)\}$
0	[0,3)	10	0	$\frac{0}{10} = 0$	0	1
3	[3,6,5)	10	1	$\frac{1}{10}$	$0 + \frac{1}{10} = \frac{1}{10}$	0,9048
6,5	[6,5,10)	7	2	$\frac{2}{7} = \frac{2}{7}$	$0 + \frac{1}{10} + \frac{2}{7}$	0,3857
10	[10,12)	4	1	$\frac{1}{4} = \frac{1}{4}$	$0 + \frac{1}{10} + \frac{2}{7} + \frac{1}{4} = 0,6357$	0,5229
12	[12,15)	2	1	$\frac{1}{2} = \frac{1}{2}$	$0 + \frac{1}{10} + \frac{2}{7} + \frac{1}{4} + \frac{1}{2} = 1,1357$	0,3212
15	[15,∞)	1	1	$\frac{1}{1} = 1$	2.1357	0.1182

Estimador de Nelson-Aalen

Assim, temos que o estimativa de Nelson-Aalen da função sobrevivência para o exemplo é dado por

$$\widehat{S}(t) = \begin{cases} 1, & \text{se } 0 \leq t < 3 \\ 0,9048, & \text{se } 5 \leq t < 6,5 \\ 0,6709, & \text{se } 6,5 \leq t < 10 \\ 0,5229, & \text{se } 10 \leq t < 12 \\ 0,3212, & \text{se } 12 \leq t < 15 \\ 0,1182, & \text{se } t \geq 15 \end{cases}$$

Gráfico da estimativa de K-M, atuarial (A) e Nelson-Aalen (N-A) da função sobrevivência para o exemplo



Observações

- O estimador K-M é consistente.
- O estimador atuarial é viesado e o víés diminui à medida que o comprimento dos intervalos diminui.
- Para amostras pequenas, existem evidências empíricas da superioridade do estimador de K-M.