Inferência em Análise de Sobrevivência

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

Departamento de Matemática Aplicada e Estatística
Universidade de São Paulo

2023

Inferência em Análise de Sobrevivência

$T_1, \ldots, T_n$ é uma amostra aleatória de uma população com f.d.p. $f(t; \boldsymbol{\theta})$, em que $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^p$.

Inferência em Análise de Sobrevivência

$T_1, \ldots, T_n$ é uma amostra aleatória de uma população com f.d.p. $f(t; \boldsymbol{\theta})$, em que $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^p$. A distribuição de $f(t; \boldsymbol{\theta})$ é completamente determinada por $\boldsymbol{\theta}$. Por exemplo, se T tem distribuição log-normal, então a f.d.p. é dada por

$$f(t; \boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi}\sigma t} \exp\{-\frac{(\log(t) - \mu)^2}{\sigma^2}\}$$

e $\boldsymbol{\theta} = (\mu, \sigma^2)^\top$ é o vetor de parâmetros da distribuição log-normal.

Inferência em Análise de Sobrevivência

$T_1, \ldots, T_n$ é uma amostra aleatória de uma população com f.d.p. $f(t;\boldsymbol{\theta})$, em que $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^p$. A distribuição de $f(t;\boldsymbol{\theta})$ é completamente determinada por $\boldsymbol{\theta}$. Por exemplo, se T tem distribuição log-normal, então a f.d.p. é dada por

$$f(t;\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi}\sigma t} \exp\{-\frac{(\log(t)-\mu)^2}{\sigma^2}\}$$

e $\boldsymbol{\theta} = (\mu, \sigma^2)^\top$ é o vetor de parâmetros da distribuição log-normal. A função verossimilhança de $\boldsymbol{\theta}$ dada a amostra observada $\boldsymbol{t} = (t_1, \ldots, t_n)^\top$ é

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma t_i} \exp\{-\frac{(\log(t_i)-\mu)^2}{\sigma^2}\}$$

Inferência em Análise de Sobrevivência

Em geral, a função verossimilhança é dada por

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} f(t_i; \boldsymbol{\theta}).$$

Inferência em Análise de Sobrevivência

Em geral, a função verossimilhança é dada por

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} f(t_i; \boldsymbol{\theta}).$$

A função log-verossimilhança é dada por

$$\ell(\boldsymbol{\theta}) = \log(L(\boldsymbol{\theta})) = \sum_{i=1}^{n} \log(f(t_i; \boldsymbol{\theta})).$$

Inferência em Análise de Sobrevivência

Os estimadores de máxima verossimilhança (EMV) de μ e σ^2 resultam de maximizar a função log-verossimilhança, $\ell(\boldsymbol{\theta})$. Em várias situações, a maximização equivale a resolver as equações de estimação

$$\boldsymbol{U}(\boldsymbol{\theta}) = \frac{\partial \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \sum_{i=1}^{n} \frac{\partial \log(f(x_i; \boldsymbol{\theta}))}{\partial \boldsymbol{\theta}} = \mathbf{0}, \tag{1}$$

em que $\boldsymbol{U}(\boldsymbol{\theta}) = (U_1(\boldsymbol{\theta}), \ldots, U_p(\boldsymbol{\theta}))^\top$ é o vetor escore com componentes $U_j(\boldsymbol{\theta}) = \frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_j}$, $j = 1, \ldots, p$.

Inferência em Análise de Sobrevivência

Usualmente, o EMV de $\widehat{\boldsymbol{\theta}}$, denotado por $\widehat{\boldsymbol{\theta}}$, não tem forma fechada. Nesse caso, é necessário algum procedimento iterativo, tal como o algoritmo de Newton-Rapson. O EMV de $\boldsymbol{\theta}$ depende de $\boldsymbol{T} = (T_1, \ldots, T_n)^\top$, e, portanto é uma estatística que tem uma distribuição amostral.

Inferência em Análise de Sobrevivência

Usualmente, o EMV de $\widehat{\boldsymbol{\theta}}$, denotado por $\widehat{\boldsymbol{\theta}}$, não tem forma fechada. Nesse caso, é necessário algum procedimento iterativo, tal como o algoritmo de Newton-Rapson. O EMV de $\boldsymbol{\theta}$ depende de $\boldsymbol{T} = (T_1, \ldots, T_n)^\top$, e, portanto é uma estatística que tem uma distribuição amostral. Para amostras suficientemente grandes, o EMV $\widehat{\boldsymbol{\theta}}$ tem distribuição normal com vetor de médias $\boldsymbol{\theta}$ e matriz de covariâncias $\boldsymbol{C}$, isto é,

$$\widehat{\boldsymbol{\theta}} \sim N_p(\boldsymbol{\theta}, \boldsymbol{C}),$$

em que $\boldsymbol{C} = \boldsymbol{I}^{-1}(\boldsymbol{\theta})$ ou $\boldsymbol{C} = \boldsymbol{I}_0^{-1}(\boldsymbol{\theta})$ e

$$\boldsymbol{I}(\boldsymbol{\theta}) = -E\left[\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}\right] = \sum_{i=1}^{n} -E\left[\frac{\partial^2 \log(f(x_i;\boldsymbol{\theta}))}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}\right]$$

é a matriz de informação de Fisher (ou matriz de informação esperada).

Inferência em Análise de Sobrevivência

e
$$\boldsymbol{I}_0(\boldsymbol{\theta}) = -\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} = -\sum_{i=1}^{n} \frac{\partial^2 \log(f(x_i; \boldsymbol{\theta}))}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T}$$

é a matriz de informação observada. Podemos utilizar $\boldsymbol{I}(\boldsymbol{\theta})$ e/ou $\boldsymbol{I}_0(\boldsymbol{\theta})$ na construção de regiões de confiança aproximadas para os parâmetros de interesse.

e

$$\boldsymbol{I}_0(\boldsymbol{\theta}) = -\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T} = -\sum_{i=1}^{n} \frac{\partial^2 \log(f(x_i;\boldsymbol{\theta}))}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}$$

é a matriz de informação observada. Podemos utilizar $\boldsymbol{I}(\boldsymbol{\theta})$ e/ou $\boldsymbol{I}_0(\boldsymbol{\theta})$ na construção de regiões de confiança aproximadas para os parâmetros de interesse.
Tomamos $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \boldsymbol{\theta}_2^\top)^\top$, em que $k \leq p$ é a dimensão de $\boldsymbol{\theta}_1$.
Estamos interessados em testar

$$H_0 : \boldsymbol{\theta}_1 = \boldsymbol{\theta}_{10} \text{ versus } H_1 : \boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_{10}.$$

Sob certas condições de regularidade, as seguintes estatísticas podem ser usadas para testar H_0.

Teste de Wald

Suponha a partição de $\widehat{\boldsymbol{\theta}}$ e $\boldsymbol{C}$

$$\widehat{\boldsymbol{\theta}} = \begin{pmatrix} \widehat{\boldsymbol{\theta}}_1 \\ \widehat{\boldsymbol{\theta}}_2 \end{pmatrix} \text{ e } \boldsymbol{C} = \begin{bmatrix} \boldsymbol{C}_{11} & \boldsymbol{C}_{12} \\ \boldsymbol{C}_{12}^\top & \boldsymbol{C}_{22} \end{bmatrix}.$$

Teste de Wald

Suponha a partição de $\widehat{\boldsymbol{\theta}}$ e $\boldsymbol{C}$

$$\widehat{\boldsymbol{\theta}} = \begin{pmatrix} \widehat{\boldsymbol{\theta}}_1 \\ \widehat{\boldsymbol{\theta}}_2 \end{pmatrix} \text{ e } \boldsymbol{C} = \begin{bmatrix} \boldsymbol{C}_{11} & \boldsymbol{C}_{12} \\ \boldsymbol{C}_{12}^\top & \boldsymbol{C}_{22} \end{bmatrix}.$$

Sob H_0,

$$W = \left(\widehat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_{10}\right)^T \boldsymbol{C}_{11}^{-1} \left(\widehat{\boldsymbol{\theta}}_1 - \boldsymbol{\theta}_{10}\right) \xrightarrow{D} \chi_k^2.$$

Portanto, rejeitamos H_0 se $W > \chi^2_{k,1-\alpha}$, em que $\chi^2_{k,1-\alpha}$ é $(1-\alpha)$-ésimo quantil da distribuição χ_k^2.

Teste escore

Considere a partição do vetor escore $\boldsymbol{U}(\boldsymbol{\theta}) = (\boldsymbol{U}_1(\boldsymbol{\theta})^\top, \boldsymbol{U}_2(\boldsymbol{\theta})^\top)^\top$ e seja $\widetilde{\boldsymbol{\theta}}_2$ o EMV de $\boldsymbol{\theta}_2$ sob $H_0 : \boldsymbol{\theta}_1 = \boldsymbol{\theta}_{10}$.

Teste escore

Considere a partição do vetor escore $\boldsymbol{U}(\boldsymbol{\theta}) = (\boldsymbol{U}_1(\boldsymbol{\theta})^\top, \boldsymbol{U}_2(\boldsymbol{\theta})^\top)^\top$ e seja $\widetilde{\boldsymbol{\theta}}_2$ o EMV de $\boldsymbol{\theta}_2$ sob $H_0 : \boldsymbol{\theta}_1 = \boldsymbol{\theta}_{10}$. Sob H_0,

$$S = \boldsymbol{U}_1^\top \boldsymbol{C}_{11} \boldsymbol{U}_1 \xrightarrow{D} \chi_k^2,$$

em que $\boldsymbol{U}_1$ e $\boldsymbol{C}_{11}$ são calculados com $\boldsymbol{\theta} = (\boldsymbol{\theta}_{10}^\top, \widetilde{\boldsymbol{\theta}}_2^\top)^\top$. Rejeita-se H_0 se $S > \chi^2_{k,1-\alpha}$.

Teste da razão de verossimilhanças

Sob H_0, a estatística

$$RV = -2\left(\ell(\widetilde{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}})\right) \xrightarrow{D} \chi^2_k,$$

em que $\widetilde{\boldsymbol{\theta}} = (\boldsymbol{\theta}_{10}^\top, \widetilde{\boldsymbol{\theta}}_2^\top)^\top$ é o EMV de $\boldsymbol{\theta}$ sob H_0 e $\widehat{\boldsymbol{\theta}}$ é o EMV de $\boldsymbol{\theta}$ irrestrito ($\boldsymbol{\theta} \in \Theta$). Rejeita-se H_0 se $RV > \chi^2_{k,1-\alpha}$

Exemplo (Teste escore)

Os dados abaixo são referentes aos tempos de sobrevivência (em dias) de 25 pacientes com câncer de pulmão terminal.

139	*304*	*193*	*248*	*27*	*210*	*134*	*203*	*724*
50	*511*	*22*	*561*	*580*	*136*	*308*	*224*	*482*
191	*144*	*68*	*683*	*256*	*441*	*101*		

A distribuição Weibull com função de risco, $h(t) = \alpha \lambda t^{\alpha - 1}$, $\alpha > 0$ e $\lambda > 0$, será utilizada

A distribuição Weibull com função de risco, $h(t) = \alpha\lambda t^{\alpha-1}$, $\alpha > 0$ e $\lambda > 0$, será utilizada (como verificar se a distribuição Weibull faz um bom ajuste?).

Temos interesse em testar as seguintes hipóteses: $H_0 : \alpha = 1$ *versus* $H_1 : \alpha \neq 1$.

A função verossimilhança de $\boldsymbol{\theta} = (\alpha, \lambda)^\top$ é dada por

$$\begin{aligned}
L(\boldsymbol{\theta}) &= \prod_{i=1}^{n} \alpha\lambda t_i^{\alpha-1} \exp(-\lambda t_i^{\alpha}) \\
&= \alpha^n \lambda^n (\prod_{i=1}^{n} t_i)^{\alpha-1} \exp(-\lambda \sum_{i=1}^{n} t_i^{\alpha}).
\end{aligned}$$

A distribuição Weibull com função de risco, $h(t) = \alpha\lambda t^{\alpha-1}$, $\alpha > 0$ e $\lambda > 0$, será utilizada (como verificar se a distribuição Weibull faz um bom ajuste?).
Temos interesse em testar as seguintes hipóteses: $H_0 : \alpha = 1$ *versus* $H_1 : \alpha \neq 1$.
A função verossimilhança de $\boldsymbol{\theta} = (\alpha, \lambda)^\top$ é dada por

$$
\begin{aligned}
L(\boldsymbol{\theta}) &= \prod_{i=1}^{n} \alpha\lambda t_i^{\alpha-1} \exp(-\lambda t_i^{\alpha}) \\
&= \alpha^n \, \lambda^n \, (\prod_{i=1}^{n} t_i)^{\alpha-1} \, \exp(-\lambda \sum_{i=1}^{n} t_i^{\alpha}).
\end{aligned}
$$

Portanto, a função log-verossimilhança é dada por

$$
\ell(\boldsymbol{\theta}) = n \log(\lambda) + n \log(\alpha) - \lambda \sum_{i=1}^{n} t_i^{\alpha} + (\alpha - 1) \sum_{i=1}^{n} \log(t_i).
$$

Assim, os componentes do vetor escore são

$$\begin{aligned} \frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha} &= \frac{n}{\alpha} - \lambda \sum_{i=1}^{n} t_i^{\alpha} log(t_i) + \sum_{i=1}^{n} log(t_i) \quad \text{e} \\ \frac{\partial \ell(\boldsymbol{\theta})}{\partial \lambda} &= \frac{n}{\lambda} - \sum_{i=1}^{n} t_i^{\alpha}. \end{aligned} \tag{2}$$

Assim, os componentes do vetor escore são

$$
\begin{aligned}
\frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha} &= \frac{n}{\alpha} - \lambda \sum_{i=1}^{n} t_i^{\alpha} \log(t_i) + \sum_{i=1}^{n} \log(t_i) \quad \text{e} \\
\frac{\partial \ell(\boldsymbol{\theta})}{\partial \lambda} &= \frac{n}{\lambda} - \sum_{i=1}^{n} t_i^{\alpha}.
\end{aligned}
\tag{2}
$$

Os elementos da matriz de derivadas segundas são dados por

$$
\begin{aligned}
\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \alpha^2} &= -\frac{n}{\alpha^2} - \lambda \sum_{i=1}^{n} t_i^{\alpha} \log(t_i)^2, \\
\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \alpha \partial \lambda} &= -\sum_{i=1}^{n} t_i^{\alpha} \log(t_i) \quad \text{e} \\
\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \lambda^2} &= -\frac{n}{\lambda^2}.
\end{aligned}
\tag{3}
$$

O EMV de λ sob H_0 é dado por $\tilde{\lambda} = n / \sum_{i=1}^{n} t_i$.
No exemplo, temos $n = 25$, $\sum_{i=1}^{n} t_i = 6940$,
$\sum_{i=1}^{n} log(t_i) = 132,2484$, $\sum_{i=1}^{n} t_i log(t_i) = 40870,27$ e
$\sum_{i=1}^{n} t_i [log(t_i)]^2 = 243502,91$ Portanto, a estimativa de máxima verossimilhança é dada por $\tilde{\lambda} = 25/6940 = 0,0036023$, de modo que

$$
\begin{aligned}
U_1 &= n - \tilde{\lambda} \sum_{i=1}^{n} t_i log(t_i) + \sum_{i=1}^{n} log(t_i) \\
&= 25 - (0,0036023)(40870,27) + 132,2484 \\
&= 10.
\end{aligned}
$$

A matriz de informação observada e sua inversa são dadas por

$$\boldsymbol{I}_0 = \begin{bmatrix} 902,17186 & 40870,268 \\ 40870,268 & 1926544 \end{bmatrix} \text{ e } \boldsymbol{C} = \boldsymbol{I}_0^{-1} = \begin{bmatrix} 0,0284592 & 0,000604 \\ 0,000604 & 0,0000133 \end{bmatrix}.$$

Assim, a estatística escore resulta em

$$S = U_1 C_{11} U_1 = 10 \times 0,0284592 \times 10 = 2,8$$

e o valor-p do teste é

$$P(\chi^2_1 > 2,8) = 0,094.$$

A matriz de informação observada e sua inversa são dadas por

$$\boldsymbol{I}_0 = \begin{bmatrix} 902,17186 & 40870,268 \\ 40870,268 & 1926544 \end{bmatrix} \text{ e } \boldsymbol{C} = \boldsymbol{I}_0^{-1} = \begin{bmatrix} 0,0284592 & 0,000604 \\ 0,000604 & 0,0000133 \end{bmatrix}.$$

Assim, a estatística escore resulta em

$$S = U_1 C_{11} U_1 = 10 \times 0,0284592 \times 10 = 2,8$$

e o valor-p do teste é

$$P(\chi_1^2 > 2,8) = 0,094.$$

Com 5% de nível de significância, não rejeitamos H_0, o que implica que os dados podem ser ajustados pela distribuição exponencial.

A função verossimilhança para dados censurados à direita

- Suponha que temos uma amostra aleatória de indivíduos de tamanho n de uma população específica cujos tempos de sobrevivência são, $T_1, \ldots, T_n$.

[1] Também chamado de indicador de censura.

A função verossimilhança para dados censurados à direita

- Suponha que temos uma amostra aleatória de indivíduos de tamanho n de uma população específica cujos tempos de sobrevivência são, $T_1, \ldots, T_n$.
- No entanto, devido à censura à direita, o tempo de sobrevivência pode não ser observado. Denote por $C_1, \ldots, C_n$ os (potenciais) tempos de censura.

[1] Também chamado de indicador de censura.

A função verossimilhança para dados censurados à direita

- Suponha que temos uma amostra aleatória de indivíduos de tamanho n de uma população específica cujos tempos de sobrevivência são, $T_1, \ldots, T_n$.
- No entanto, devido à censura à direita, o tempo de sobrevivência pode não ser observado. Denote por $C_1, \ldots, C_n$ os (potenciais) tempos de censura.
- Dados observados: $t_i = \min(T_i, C_i)$ e o indicador de falha[1] δ_i,

$$\delta_i = \begin{cases} 1, & \text{se } T_i \leq C_i, \\ 0, & \text{se } T_i > C_i. \end{cases}$$

[1] Também chamado de indicador de censura.

- Se a f.d.p de T é $f(t;\boldsymbol{\theta})$, com $\boldsymbol{\theta} \in \Theta \subset R^p$ e a função sobrevivência de T é $S(t;\boldsymbol{\theta})$, a função de verossimilhança de $\boldsymbol{\theta}$ é dada por

- Se a f.d.p de T é $f(t;\boldsymbol{\theta})$, com $\boldsymbol{\theta} \in \Theta \subset R^p$ e a função sobrevivência de T é $S(t;\boldsymbol{\theta})$, a função de verossimilhança de $\boldsymbol{\theta}$ é dada por

$$L(\boldsymbol{\theta}) \propto \prod_{i=1}^{n} [f(t_i;\boldsymbol{\theta})]^{\delta_i} [S(t_i;\boldsymbol{\theta})]^{1-\delta_i} \qquad (4)$$

Alternativamente,

$$L(\boldsymbol{\theta}) \propto \prod_{i=1}^{n} [h(t_i;\boldsymbol{\theta})]^{\delta_i} S(t_i;\boldsymbol{\theta})$$

em que $h(t;\boldsymbol{\theta}) = \frac{f(t;\boldsymbol{\theta})}{S(t;\boldsymbol{\theta})}$.

- Se a f.d.p de T é $f(t;\boldsymbol{\theta})$, com $\boldsymbol{\theta} \in \Theta \subset R^p$ e a função sobrevivência de T é $S(t;\boldsymbol{\theta})$, a função de verossimilhança de $\boldsymbol{\theta}$ é dada por

$$L(\boldsymbol{\theta}) \propto \prod_{i=1}^{n} [f(t_i;\boldsymbol{\theta})]^{\delta_i} [S(t_i;\boldsymbol{\theta})]^{1-\delta_i} \tag{4}$$

  Alternativamente,

$$L(\boldsymbol{\theta}) \propto \prod_{i=1}^{n} [h(t_i;\boldsymbol{\theta})]^{\delta_i} S(t_i;\boldsymbol{\theta})$$

  em que $h(t;\boldsymbol{\theta}) = \frac{f(t;\boldsymbol{\theta})}{S(t;\boldsymbol{\theta})}$.

- A função de verossimilhança dada em (4) pode ser representando por

$$L(\boldsymbol{\theta}) \propto \prod_{i \in F} f(t_i;\boldsymbol{\theta}) \prod_{i \in C} S(t_i;\boldsymbol{\theta}),$$

  em que F é o conjunto de tempos não censurados (falhas) e C é o conjunto de tempos censurados.

Exemplo 1. Dados censurados

Os dados abaixo correspondem aos tempos de sobrevivência (em meses) de pacientes com certa doença.

3	5	6^+	8	10^+	11^+	15	20^+	22	23
29	32	35	40	26	28	33^+	21	24^+	27^+

$+$ indica o dado censurado.
Assumindo que os tempos de sobrevivência tem taxa de falha constante, $\lambda > 0$. Estimar pontualmente e por intervalo: (a) o tempo mediano, (b) a proporção de pacientes com tempos de sobrevivência mais de 10 meses.

- A função de taxa de falha e a função de sobrevivência são dados por $h(t;\lambda) = \lambda$ e $S(t;\lambda) = e^{-\lambda t}$ respectivamente.
- A função de sobrevivência de λ dado os dados observados é dado por

$$L(\lambda) = \prod_{i=1}^{n} \lambda^{\delta_i} e^{-\lambda t_i} = \lambda^r e^{-\lambda \sum_{i=1}^{n} t_i}$$

  onde $r = \sum_{i=1}^{n} \delta_i$ é número observações de tempos de sobrevivência.
- De modo que a função log-verossimilhança é dado por

$$\ell(\lambda) = r \log(\lambda) - \lambda \sum_{i=1}^{n} t_i.$$

- A equação de estimação é

$$U(\lambda) = \frac{r}{\lambda} - \sum_{i=1}^{n} t_i = 0$$

- Portanto, o EMV de λ é dado por

$$\widehat{\lambda} = \frac{r}{\sum_{i=1}^{n} t_i}$$

Um vez que

$$\frac{\partial U(\lambda)}{\partial \lambda} = -\frac{r}{\lambda^2},$$

a informação de Fisher é

$$I(\lambda) = E\left[-\frac{\partial U(\lambda)}{\partial \lambda}\right] = \frac{E(r)}{\lambda^2} = \frac{\sum_{i=1}^{n} E(\delta_i)}{\lambda^2} = \frac{\sum_{i=1}^{n} P(T_i \leq C_i)}{\lambda^2}.$$

- Sob censura tipo I, a informação de Fisher é dada por

$$I(\lambda) = \frac{\sum_{i=1}^{n}(1 - e^{-\lambda c_i})}{\lambda^2}$$

- A informação observada é dada por

$$I_0(\lambda) = \frac{r}{\lambda^2}$$

- A estimativa da variância de $\widehat{\lambda}$ é

$$\widehat{Var(\widehat{\lambda})} = I_0^{-1}(\widehat{\lambda}) = \frac{r}{[\sum_{i=1}^{n} t_i]^2} = \frac{\widehat{\lambda}^2}{r}$$

- Assintoticamente,

$$\widehat{\lambda} \overset{a}{\sim} N(\lambda, \frac{\widehat{\lambda}^2}{r})$$

- Um intervalo de 95% de confiança para λ é dado por:

$$(\widehat{\lambda} - 1,96 \times \frac{\widehat{\lambda}}{\sqrt{r}}; \widehat{\lambda} + 1,96 \times \frac{\widehat{\lambda}}{\sqrt{r}})$$

- Algumas vezes a distribuição exponencial é parametrizada na média, $E(T) = \theta = 1/\lambda$. O EMV de θ é

$$\widehat{\theta} = \frac{\sum_{i=1}^{n} t_i}{\sum_{i=1}^{n} \delta_i},$$

  Assintoticamente,

$$\widehat{\theta} \overset{a}{\sim} N(\theta, \frac{\widehat{\theta}^2}{r})$$

- Se os dados são não censurados, $\delta_i = 1,\ i = 1, dots, n$ então

$$\widehat{\theta} = \frac{\sum_{i=1}^{n} t_i}{n} = \bar{t},$$

- Para os dados do exemplo, temos: $r = 13$ e $\sum_{i=1}^{n} t_i = 418$. Portanto,

$$\widehat{\lambda} = \frac{r}{\sum_{i=1}^{n} t_i} = \frac{13}{408} = 0,0311,$$

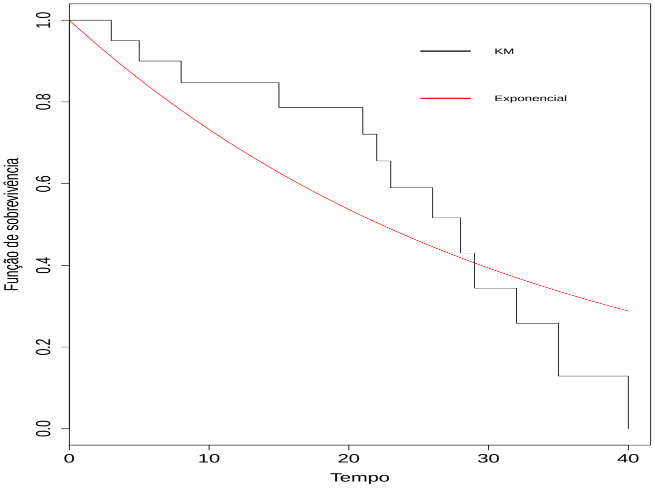
- O erro padrão

$$EP(\widehat{\lambda}) = \frac{\widehat{\lambda}}{\sqrt{r}} = \frac{0,0311}{\sqrt{13}} = 0,0086$$

- Um intervalo de 95% de confiança para λ é dado por:

$$(\widehat{\lambda} - 1,96 \times EP(\widehat{\lambda}); \widehat{\lambda} + 1,96 \times EP(\widehat{\lambda})) =$$
$$(0,0311 - 1,96 \times 0,0086; 0,0311 + 1,96 \times 0,0086) =$$
$$(0,0142; 0,0480)$$

Gráfico da estimativa da função de sobrevivência



Obviamente, a distribuição exponencial é um ajuste inadequado. Neste caso, podemos escolher um dos seguintes opções

1. Escolha um modelo mais flexível, como o modelo de Weibull.
2. Considerar o estimador de Kaplan-Meier que não faz qualquer hipótese quanto à forma da distribuição.
3. Na maioria das aplicações biomédicas, é considerado o estimador de Kaplan-Meier.

Modelo Weibull

Uma v.a. T tem a distribuição Weibull parâmetros $\lambda > 0$ (escala) e $\alpha > 0$ (forma) se sua f.d.p dada por

$$f(t;\alpha,\beta) = \alpha\lambda t^{\alpha-1} e^{-\lambda t^{\alpha}},\ t > 0 \tag{5}$$

A função de sobrevivência e função de riscos são dada por,

$$S(t;\alpha,\beta) = e^{-\lambda t^{\alpha}},\ t > 0 \tag{6}$$

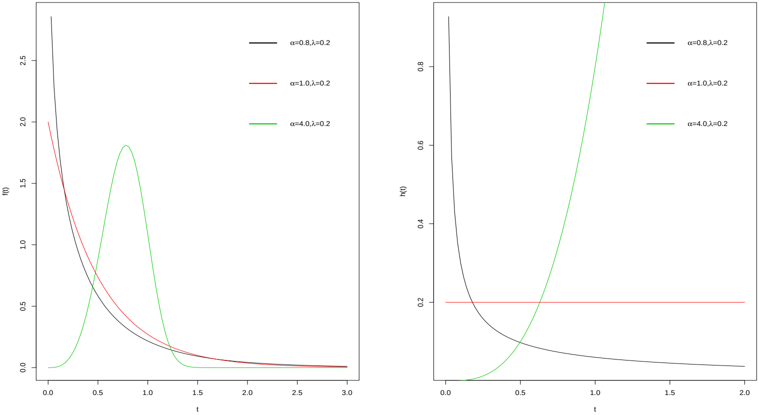
e

$$h(t;\alpha,\beta) = \alpha\lambda t^{\alpha-1},\ t > 0 \tag{7}$$

A forma da função de risco tem as seguintes formas:

- Se $\alpha = 1$ a função de risco é constante (distribuição exponencial)
- Se $\alpha < 1$ a função de risco é decrescente,
- Se $\alpha > 1$ a função de risco é crescente.

Gráfico da f.d.p e função de sobrevivência da distribuição Weibull



Momentos e a função quantil do modelo Weibull

- O k-ésimo momento é dada por

$$E[T^k] = \lambda^{-k/\alpha}\Gamma(\frac{k}{\alpha} + 1)$$

- Média e variância

$$E(T) = \lambda^{-1/\alpha}\Gamma(\frac{1}{\alpha}+1),\ Var(T) = \lambda^{-2/\alpha}[\Gamma(\frac{2}{\alpha}+1)-\Gamma^2(\frac{1}{\alpha}+1)]$$

- O $p-$ésimo quantil, denotado por t_p

$$t_p = \lambda^{-1/\alpha}[-\log(1-p)]^{1/\alpha},\ 0 < p < 1.$$

- A mediana, $md = t_{0,5}$

$$t_{0,5} = \lambda^{-1/\alpha}[\log(2)]^{1/\alpha}.$$

Adequação do modelo Weibull

- O $\log(-\log(S(t)))$ é linear em relação a $\log(t)$.

$$
\begin{aligned}
S(t) = e^{-\lambda t^{\alpha}}, &\Longrightarrow -\log(S(t)) = \lambda t^{\alpha} \\
&\Longrightarrow \log(-\log(S(t))) = \log(\lambda) + \alpha \log(t).
\end{aligned}
$$

- Esta propriedade permite uma avaliação gráfica da adequação de um modelo Weibull plotando

$$\log(-\log(\widehat{S(t)})) \ \ vs \ \ \log(t)$$

onde $\widehat{S(t)}$ é a estimativa K-M da função de sobrevivência.

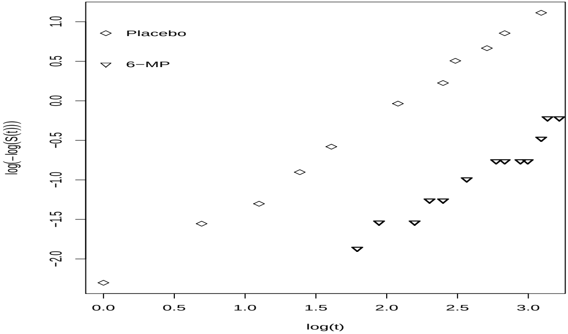
Exemplo 2: Dados leucemia

Os dados abaixo representam tempos de remissão (em meses) do câncer (leucemia aguda) de 42 crianças tratadas com:6-mercaptopurine (6-MP) ou placebo.

Placebo	1	22	3	12	8	17	2	11	8	12	2
	5	4	15	8	23	5	11	4	1	8	
6-MP	10	7	32^+	23	22	6	16	34^+	32^+	25^+	11^+
	20^+	19^+	6	17^+	35^+	6	13	9^+	6^+	10^+	

$+$: indica tempo censurado.

Gráfico de $\log(-\log(\widehat{S(t)}))$ *versus* $\log(t)$ para os dados de leucemia.



- ▷ Linhas retas: modelo Weibull.
- ▷ Retas paralelas: riscos proporcionais (RP).

Modelo Weibull-RP

- Relembre que $h(t) = \alpha\lambda t^{\alpha}$.
- Modelo Weibull-RP
  - Tomamos $\lambda = \exp(\beta_0 + \beta_1 x)$ (modelo para λ).
  - A razão de riscos ($x = 1$ para 6-MP *versus* $x = 0$ para placebo) é

$$RR(t) = \frac{\exp(\beta_0 + \beta_1)\alpha t^{\alpha}}{\exp(\beta_0)\alpha t^{\alpha}} = \exp(\beta_1),$$

    que indica que a propriedade de RP é satisfeita.
- Neste resultado, α é o mesmo valor para 6-MP e placebo.

Inferência no modelo Weibull com dados censurados

A função verossimilhança de $\boldsymbol{\theta} = (\alpha, \lambda)^\top$ é dada por

$$
\begin{aligned}
L(\boldsymbol{\theta}) &= \prod_{i=1}^{n} [\alpha \lambda t_i^{\alpha-1}]^{\delta_i} \exp\{-\lambda t_i^{\alpha}\} \\
&= \alpha^r \lambda^r \exp\{-\lambda \sum_{i=1}^{n} t_i^{\alpha} + (\alpha - 1) \sum_{i=1}^{n} \delta_i \log(t_i)\}
\end{aligned}
$$

em que $t_i = \min(T_i, C_i)$ $i = 1, \ldots, n$ e δ_i é o indicador de falha, isto é,

$$
\delta_i = \begin{cases} 1, & \text{se } T_i \leq C_i, \\ 0, & \text{se } T_i > C_i. \end{cases}
$$

Portanto, a função log-verossimilhança é dada por

$$
\ell(\boldsymbol{\theta}) = r \log(\lambda) + r \log(\alpha) - \lambda \sum_{i=1}^{n} t_i^{\alpha} + (\alpha - 1) \sum_{i=1}^{n} \delta_i \log(t_i).
$$

Assim, os componentes do escore são

$$\begin{aligned}\frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha} &= \frac{r}{\alpha} - \lambda \sum_{i=1}^{n} t_i^{\alpha} log(t_i) + \sum_{i=1}^{n} \delta_i log(t_i) \text{ e} \\ \frac{\partial \ell(\boldsymbol{\theta})}{\partial \lambda} &= \frac{r}{\lambda} - \sum_{i=1}^{n} t_i^{\alpha},\end{aligned} \tag{8}$$

e os elementos da matriz de derivadas segundas são dados por

$$\begin{aligned}\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \alpha^2} &= -\frac{r}{\alpha^2} - \lambda \sum_{i=1}^{n} t_i^{\alpha} log(t_i)^2, \\ \frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha \partial \lambda} &= -\sum_{i=1}^{n} t_i^{\alpha} log(t_i) \text{ e} \\ \frac{\partial \ell(\boldsymbol{\theta})}{\partial \lambda^2} &= -\frac{r}{\lambda^2}.\end{aligned} \tag{9}$$

- As estimativas de MV são $\widehat{\alpha} = 2,14943$ e $\widehat{\lambda} = 0,00072$.
- A matriz de informação observada e sua inversa são dadas por

$$\boldsymbol{I}_0 = \begin{bmatrix} 149,3758 & 60466,06 \\ 60466,0604 & 25148299,99 \end{bmatrix}$$

e

$$\boldsymbol{C} = \boldsymbol{I}_0^{-1} = \begin{bmatrix} 0,2504761623 & -0,0006022398 \\ -0,0006022398 & 1,487777 \times 10^{-6} \end{bmatrix}.$$

- Erros padrão das EMV:

$$\text{ep}(\widehat{\alpha}) = 0,502257493 \text{ e } \text{ep}(\widehat{\lambda}) = 0,001224176.$$

- Intervalo de 95% de confiança (aproximado):

$$\widehat{\alpha} \mp 1,96\text{EP}(\widehat{\alpha}) = (1,165; 3,134), \ \widehat{\lambda} \mp 1,96\text{EP}(\widehat{\lambda}) = (0; 0,00312).$$

Estimação pontual e intervalar do tempo mediano $t_{0,5}$

- Pelo principio de invariância, o EMV de $t_{0,5} = \lambda^{-1/\alpha}[\log(2)]^{1/\alpha}$ é dado por $\widehat{t_{0,5}} = \widehat{\lambda}^{-1/\widehat{\alpha}}[\log(2)]^{1/\widehat{\alpha}}$.
- A variância de $\widehat{t_{0,5}}$ é obtida pelo método delta, a qual resulta em

$$Var(\widehat{t_{0,5}}) = \begin{pmatrix} \frac{\partial t_{0,5}}{\partial \alpha} \\ \frac{\partial t_{0,5}}{\partial \lambda} \end{pmatrix}^{\top} \begin{bmatrix} Var(\widehat{\alpha}) & Cov(\widehat{\alpha}, \widehat{\lambda}) \\ Cov(\widehat{\alpha}, \widehat{\lambda}) & Var(\widehat{\lambda}) \end{bmatrix} \begin{pmatrix} \frac{\partial t_{0,5}}{\partial \alpha} \\ \frac{\partial t_{0,5}}{\partial \lambda} \end{pmatrix},$$

em que

$$\frac{\partial t_{0.5}}{\partial \alpha} = -\frac{1}{\alpha^2}\left(\frac{\log(2)}{\lambda}\right)^{\alpha^{-1}} \log\left(\frac{\log(2)}{\lambda}\right) \text{ e}$$

$$\frac{\partial t_{0.5}}{\partial \lambda} = -\frac{1}{\alpha\,\lambda}\left(\frac{\log(2)}{\lambda}\right)^{\alpha^{-1}}.$$

Estimação pontual e intervalar do tempo mediano $t_{0,5}$

Obtemos

$$Var(\widehat{t_{0.5}}) = Var(\widehat{\alpha})\left(\frac{\partial t_{0.5}}{\partial \alpha}\right)^2 + 2Cov(\widehat{\alpha}, \widehat{\lambda})\frac{\partial t_{0.5}}{\partial \alpha}\frac{\partial t_{0.5}}{\partial \lambda} + Var(\widehat{\lambda})\left(\frac{\partial t_{0.5}}{\partial \lambda}\right)^2.$$

Erro padrão:

$$ep(\widehat{t_{0.5}}) = \sqrt{Var(\widehat{t_{0.5}})|_{\widehat{\alpha},\widehat{\lambda}}}.$$

Intervalo de 95% de confiança para $t_{0.5}$ (aproximado):

$$\widehat{t_{0.5}} \mp 1,96 ep(\widehat{t_{0.5}}).$$

Estimação pontual e intervalar do tempo mediano $t_{0,5}$

- Para o Exemplo 1, a EMV do tempo mediano é

$$\widehat{t_{0,5}} = \widehat{\lambda}^{-1/\widehat{\alpha}}[\log(2)]^{1/\widehat{\alpha}} = 0,0072^{-1/2,14943}[\log(2)]^{1/2,14943} = 24,5.$$

- A estimativa da variância de $\widehat{t_{0,5}}$ é dada por

$$\widehat{Var}(\widehat{t_{0,5}}) = \begin{pmatrix} -36.36386 \\ -15816.09349 \end{pmatrix}^{\top} \begin{bmatrix} 0.2505 & -0.0006 \\ -0.0006 & 1.488 \times 10^{-6} \end{bmatrix} \begin{pmatrix} -36.36386 \\ -15816.09349 \end{pmatrix}$$
$$= 10,6707.$$

- Intervalo de 95% de confiança para $t_{0,5}$ é dado por

$$\widehat{t_{0,5}} \mp 1,96\, ep(\widehat{t_{0,5}}) = 24,5 \mp 1,96\sqrt{10,6707} = (18,048; 30,853).$$

Estimação pontual e intervalar de $S(t_0)$

- Pelo princípio de invariância, o EMV de $S(t_0) = \exp\{-\lambda t_0^{\alpha}\}$ é dado por $\widehat{S(t_0)} = \exp\{-\widehat{\lambda} t_0^{\widehat{\alpha}}\}$.
- A variância de $\widehat{S(t_0)}$ é obtida pelo método delta, que resulta em

$$Var(\widehat{S(t_0)}) = \begin{pmatrix} \frac{\partial S(t_0)}{\partial \alpha} \\ \frac{\partial S(t_0)}{\partial \lambda} \end{pmatrix}^{\top} \begin{bmatrix} Var(\widehat{\alpha}) & Cov(\widehat{\alpha}, \widehat{\lambda}) \\ Cov(\widehat{\alpha}, \widehat{\lambda}) & Var(\widehat{\lambda}) \end{bmatrix} \begin{pmatrix} \frac{\partial S(t_0)}{\partial \alpha} \\ \frac{\partial S(t_0)}{\partial \lambda} \end{pmatrix}$$

em que

$$\frac{\partial S(t_0)}{\partial \alpha} = -\lambda\, {t_0}^{\alpha} \log(t_0)\, \mathrm{e}^{-\lambda\, t^{\alpha}} \text{ e}$$
$$\frac{\partial S(t_0)}{\partial \lambda} = -{t_0}^{\alpha} \mathrm{e}^{-\lambda\, {t_0}^{\alpha}}.$$

Estimação pontual e intervalar de $S(t_0)$

Variância de $\widehat{S(t_0)}$:

$$Var(\widehat{S(t_0)}) = Var(\widehat{\alpha})\left[\frac{\partial S(t_0)}{\partial \alpha}\right]^2 + 2Cov(\widehat{\alpha},\widehat{\lambda})\frac{\partial S(t_0)}{\partial \alpha}\frac{\partial S(t_0)}{\partial \lambda} + Var(\widehat{\lambda})\left[\frac{\partial S(t_0)}{\partial \lambda}\right]^2.$$

Erro padrão:

$$ep(\widehat{S(t_0)}) = \sqrt{Var(\widehat{S(t_0)})|_{\widehat{\alpha},\widehat{\lambda}}}.$$

Intervalo de 95% de confiança (aproximado) para $S(t_0)$:

$$\widehat{S(t_0)} \mp 1,96 ep(\widehat{S(t_0)}).$$

Estimação pontual e intervalar de $S(t_0)$

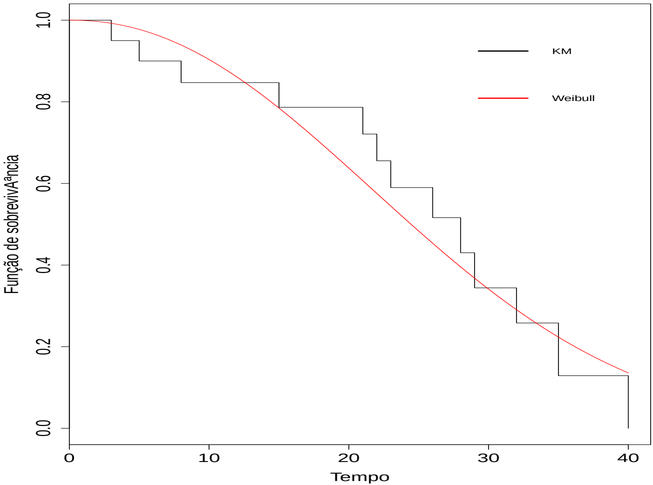
- Para o Exemplo 1, estimar, a proporção de pacientes com tempos de sobrevivência maiores do que10 meses, ou seja, $S(10) = \exp(-\lambda 10^{\alpha})$. A EMV de $S(10)$ é $\widehat{S(10)} = \exp(-\widehat{\lambda} 10^{\widehat{\alpha}}) = \exp(-0,0072 \times 10^{2,14943}) = 0,904.$
- A estimativa da variância de $\widehat{S(10)}$ é dada por

$$\begin{aligned}\widehat{Var}(\widehat{S(10)}) &= \begin{pmatrix} -0.2111 \\ -127.4430 \end{pmatrix}^{\top} \begin{bmatrix} 0.2505 & -0.0006 \\ -0.0006 & 1.488 \times 10^{-6} \end{bmatrix} \begin{pmatrix} -0.2111 \\ -127.4430 \end{pmatrix}. \\ &= 0.00294851.\end{aligned}$$

- Intervalo de 95% de confiança para $S(10)$ (aproximado):

$$\widehat{S(10)} \ \mp \ 1,96 ep(\widehat{S(10)}) = 0,904 \mp 1,96\sqrt{0.00294851} = (0,7972; 1,00).$$

Gráfico da estimativa da função de sobrevivência



Comparação entre a distribuição exponencial e distribuição Weibull

- Hipóteses"

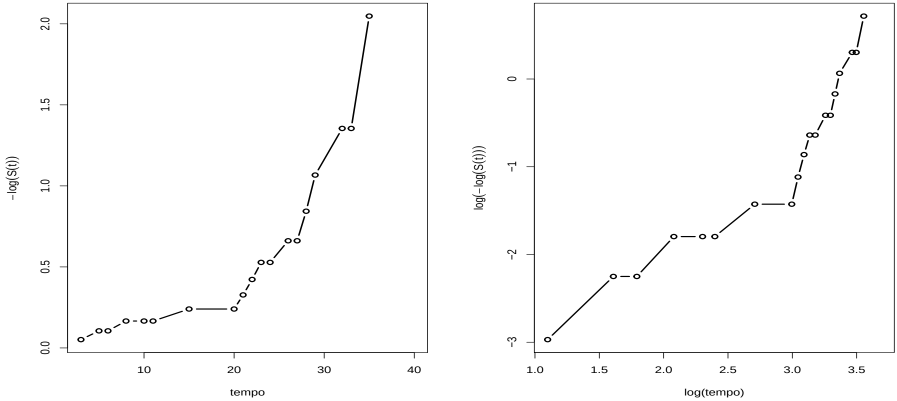
$$H_0 : \alpha = 1 \text{ (Exponencial) versus } H_1 : \alpha \neq 1 \text{ (Weibull).}$$

Estatística de Wald:

$$W = \left(\frac{\widehat{\alpha} - 1}{ep(\widehat{\alpha})} \right)^2 = \left(\frac{2,14943 - 1}{0.502257493} \right)^2 = 5,274667.$$

- Já que valor-p $= P(\chi^2_1 > 5,27) = 0,022 < 0,05$, rejeita-se $H_0 : \alpha = 1$, isto é, rejeita-se a distribuição exponencial ao nível de significância de 5%.

Adequabilidade do modelo Weibull



- A inadequação do ajuste exponencial também é demonstrada no primeiro gráfico da Figura(panel esquerdo).
- Se o modelo exponencial modela-se bem os dados, todos os pontos (t vs $-\log(\widehat{S(t)})$) estariam acima da linha reta passando pela origem.
- Por outro lado, o plot na Figura (panel direito) mostra a adequação do modelo de Weibull, uma vez que a maioria dos pontos ($\log(t)$ vs $\log[-\log(\widehat{S(t)})]$) passam pela linha reta.
- Aqui $\widehat{S(t)}$ é o estimador de K-M de $S(t)$.

A distribuição valor extremo

- Se T tem distribuição Weibull com parâmetros α e λ e f.d.p. dada em (5) então a v.a. $Y = \log(T)$ tem distribuição valor extremo com a seguinte f.d.p.

$$f(y) = \frac{1}{\sigma} \exp\left\{ \left(\frac{y-\mu}{\sigma} \right) - \exp\left\{ \frac{y-\mu}{\sigma} \right\} \right\}, \; y \in \mathbb{R}, \quad (10)$$

  em que $\mu \in \mathbb{R}$ e $\sigma > 0$.
- Se $\mu = 0$ e $\sigma = 1$, tem-se a distribuição valor extremo padrão.
- Os parâmetros μ e σ são denominados parâmetros de localização e escala, respectivamente.
- Os parâmetros das distribuição de Weibull e valor extremo apresentam as seguintes relações:

$$\alpha = 1/\sigma, \;\; e \;\; \lambda = \exp(-\mu/\sigma)$$

A distribuição valor extremo

- As funções de sobrevivência e de taxa de risco dq variável Y são dados por

$$S(y) = \exp\left\{-\exp\left\{\frac{y-\mu}{\sigma}\right\}\right\}, \tag{11}$$

e

$$h(y) = \frac{1}{\sigma}\exp\left\{\left(\frac{y-\mu}{\sigma}\right)\right\}. \tag{12}$$

- A média e variância são respectivamente,

$$E(Y) = \mu - \nu\sigma, \quad Var(Y) = \frac{\pi^2}{6}\sigma^2.$$

onde $\nu = 0,3772$ é a constante de Euler.

- O p-ésimo quantil é dado por:

$$y_p = \mu + \sigma \log[-log(1-p)], \ 0 < p < 1.$$

Observação

- A distribuição VE, tem a seguinte representação:

$$y = \mu + \sigma\varepsilon,$$

onde ε tem distribuição VE padrão, com f.d.p,

$$f(\varepsilon) = \exp\{\varepsilon - \exp\{\varepsilon\}\}$$

- A maioria dos pacotes computacionais tais como R ou SAS considera a distribuição Weibull com a seguinte função de sobrevivência

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

- Os parâmetros das distribuição de Weibull com parametrizaçao acima e o valor extremo apresentam as seguintes relações de igualdade:

$$\alpha = 1/\sigma, \quad e \quad \theta = \exp(\mu)$$

Estimação dos parâmetros da distribuição a partir da distribuição valor extremo

A função de verossimilhança de μ e σ dado os dados observados

$$L(\mu,\sigma)=\prod_{i=1}^{n}\left[\frac{1}{\sigma}\exp\left\{\left(\frac{y_i-\mu}{\sigma}\right)\right\}\right]^{\delta_i}\exp\left\{-\exp\left\{\frac{y_i-\mu}{\sigma}\right\}\right\},$$

onde $y_i=\log(t_i)$, $i=1,\ldots,n$.
A correspondente função log-verossimilhança é dada por

$$\ell(\mu,\sigma)=-r\log(\sigma)+\sum_{i=1}^{n}\delta_i z_i-\sum_{i=1}^{n}\exp\left\{z_i\right\}. \qquad (13)$$

onde $z_i=\frac{y_i-\mu}{\sigma}$

Assim, os componentes do escore são

$$\begin{aligned}
\frac{\partial \ell(\mu,\sigma)}{\partial \mu} &= -\frac{r}{\sigma} + \frac{1}{\sigma}\sum_{i=1}^{n} e^{z_i} \\
\frac{\partial \ell(\mu,\sigma)}{\partial \sigma} &= -\frac{r}{\sigma} - \frac{1}{\sigma}\sum_{i=1}^{n} \delta_i z_i + \frac{1}{\sigma}\sum_{i=1}^{n} z_i e^{z_i},
\end{aligned} \tag{14}$$

e os elementos da matriz de derivadas segundas são

$$\begin{aligned}
\frac{\partial^2 \ell(\mu,\sigma)}{\partial \mu^2} &= -\frac{1}{\sigma^2}\sum_{i=1}^{n} e^{z_i} \\
\frac{\partial^2 \ell(\mu,\sigma)}{\partial \mu \partial \sigma} &= \frac{r}{\sigma^2} - \frac{1}{\sigma^2}\sum_{i=1}^{n} e^{z_i} - -\frac{1}{\sigma^2}\sum_{i=1}^{n} z_i e^{z_i} \\
\frac{\partial \ell(\mu,\sigma)}{\partial \sigma^2} &= \frac{1}{\sigma^2}\left[r - \sum_{i=1}^{n} \delta_i z_i - 2\sum_{i=1}^{n} z_i e^{z_i} - \sum_{i=1}^{n} z_i^2 e^{z_i} \right]
\end{aligned} \tag{15}$$

Ajuste dos dados Exemplo 1

- Com o a função escore e a matriz de informação pode-se implementar o algoritmo de Newton-Rapson.
- O algoritmo está implementado na maioria dos pacotes estatísticos tais como R, Minitab, SAS, etc.
- A continuação o codigo no software R:
- $y = c(3, 5, 6, 8, 10, 11, 15, 20, 22, 23, 27, 29, 32, 35, 40, 26, 28, 33, 21, 24)$
- $delta = c(1, 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 1, 1, 1, 0, 1, 0)$
- fit1=survreg(Surv(y, delta) $\sim$ 1, dist='weibull')
- summary(fit1)

Ajuste dos dados Exemplo 1

Call:
survreg(formula = Surv(y, delta) $\sim$ 1, dist = "weibull")

	Value	Std. Error	z	p-value
(Intercept)	3.367	0.129	26.07	7.36e-150
Log(scale)	-0.765	0.234	-3.27	1.06e-03

Scale= 0.465
Weibull distribution
Loglik(model)= -54.1 Loglik(intercept only)= -54.1
Number of Newton-Raphson Iterations: 7
n= 20

Ajuste dos dados Exemplo 1

Os parâmetros de escala (σ) e forma (α) das funções `rweibull`, `pweibull`, etc são dados por $\sigma = \exp(\text{Intercept})$ e $\alpha = 1/\text{Scale}$.

Na parametrização da eq. (7): $\lambda = 1/\sigma^{\alpha}$.
Fazendo $\mu = \text{Intercept}$ e $\sigma_* = \text{Log(scale)}$, obtemos

$$\alpha = \frac{1}{\exp(\sigma_*)} \quad \text{e} \quad \lambda = e^{-\mu/\exp(\sigma_*)}.$$

Já que as EMV são $\widehat{\mu} = 3,36717$ e $\widehat{\sigma} = \exp(\widehat{\sigma_*}) = \exp(-0,765) = 0,46525$, as EMV de α e λ são

$$\widehat{\alpha} = \frac{1}{0,465} = 2,149 \text{ e } \widehat{\lambda} = e^{-3.36717/0,465} = 0.00072.$$

O método delta é usado para calcular os erros padrão das EMV de α e λ.

Ajuste dos dados Exemplo 1

- α é função de σ_*, isto é, $\alpha = \frac{1}{\exp(\sigma_*)} = g_1(\sigma_*)$.

$$Var(\widehat{\alpha}) = \left[\frac{\partial g_1(\sigma_*)}{\partial \sigma_*}\right]^2 Var(\widehat{\sigma_*}) = \frac{Var(\widehat{\sigma_*})}{\exp(2\sigma_*)}.$$

- Eo erro padrão da EMV de α:

$$ep(\widehat{\alpha}) = \frac{0,234}{\exp(-0,765)} = 0,50226.$$

- $\lambda = e^{-\mu/\exp\{\sigma_*\}} = g_2(\mu,\sigma_*)$ é função de μ e σ_*.

$$Var(\widehat{\lambda}) = \begin{pmatrix} \frac{\partial g_2(\mu,\sigma_*))}{\partial \mu} \\ \frac{\partial g_2(\mu,\sigma_*)}{\partial \sigma_*} \end{pmatrix}^\top \begin{bmatrix} Var(\widehat{\mu}) & Cov(\widehat{\mu},\widehat{\sigma}_*) \\ Cov(\widehat{\mu},\widehat{\sigma}_*) & Var(\widehat{\sigma}_*) \end{bmatrix} \begin{pmatrix} \frac{\partial g_2(\mu,\sigma_*))}{\partial \mu} \\ \frac{\partial g_2(\mu,\sigma_*)}{\partial \sigma_*} \end{pmatrix}.$$

Ajuste dos dados do Exemplo 1

- Derivadas:

$$\frac{\partial g_2(\mu, \sigma_*)}{\partial \mu} = -\mathrm{e}^{-\mu\, e^{-\sigma_*} - \sigma_*} \text{ e}$$
$$\frac{\partial g_2(\mu, \sigma_*)}{\partial \sigma_*} = \mu\, e^{-\mu\, \mathrm{e}^{-\sigma_*} - \sigma_*}.$$

- Estimativa da variância da EMV de λ (a matriz de covariâncias abaixo é obtida com `fit1$var`):

$$\widehat{Var}(\widehat{\lambda}) = \begin{pmatrix} -0,001546 \\ 0,00521 \end{pmatrix}^T \begin{bmatrix} 0,016678010 & 0,001221347 \\ 0,001221347 & 0,054603314 \end{bmatrix} \begin{pmatrix} -0,001545 \\ 0,00521 \end{pmatrix}.$$

- Erro padrão de $\widehat{\lambda}$:

$$ep(\widehat{\lambda}) = \sqrt{\widehat{Var}(\widehat{\lambda})} = 0,001224.$$