ANÁLISE DE SOBREVIVÊNCIA APLICADA

Modelos Paramétricos - Caps 3 e 4

Enrico A. Colosimo/UFMG

Depto. Estatística - ICEx - UFMG

- Estudos clínicos/industriais/sociais, em geral, envolvem covariáveis que podem estar relacionadas com o tempo de sobrevivência.
- Usualmente, o objetivo do estudo está relacionado à estas covariáveis.
- As covariáveis podem ser:
  - em estudos clínicos: gênero, idade, pressão arterial, presença de diabetes, etc;
  - em estudo industriais: temperatura, dureza do material, voltagem, etc.

EXEMPLO: Estudo sobre aleitamento materno

- Este estudo foi realizado no Centro de Saúde São Marcos, localizado em Belo Horizonte, por professores do Depto de Pediatria da UFMG. Este Centro de Saúde é um ambulatório municipal que atende essencialmente a população de baixa renda.
- Objetivos: conhecer a prática do aleitamento materno de mães que utilizam este centro, assim como os possíveis fatores de risco ou de proteção para o desmame precoce.
- Variável resposta: tempo máximo de aleitamento materno, ou seja, o desmame completo da criança.
- Algumas crianças não foram acompanhadas até o desmame e, portanto, registra-se a presença de censuras.

EXEMPLO: Estudo sobre aleitamento materno

- Foram registradas 11 covariáveis e a variável resposta.

Código	Descrição	Categorias
V1	Experiência anterior de amamentação	0 se sim e 1 se não
V2	Número de filhos vivos	0 se dois ou menos e 1 se mais de dois
V3	Conceito materno sobre o tempo ideal de amamentação	0 se $>$ 6 meses e 1 se $\leq$ 6 meses
V4	Dificuldades para amamentar nos primeiros dias pós-parto	0 se não e 1 se sim
V5	Tipo de serviço em que realizou o pré-natal	0 se público e 1 se privado/convênios
V6	Recebeu exclusivamente leite materno na maternidade	0 se sim e 1 se não
V7	A criança teve contato com o pai	0 se sim e 1 se não
V8	Renda per capita (em SM/mês)	0 se $\geq$ 1 SM e 0 se $<$ 1 SM
V9	Peso ao nascimento	0 se $\geq$ 2,5kg e 1 se $<$ 2,5kg
V10	Tempo de separação mãe-filho pós-parto	0 se $\leq$ 6 horas e 1 se $>$ 6 horas
V11	Permanência no berçário	0 se não e 1 se sim

Técnicas Não-Paramétricas vs Modelos Paramétricos

- As técnicas não-paramétricas são limitadas na presença de covariáveis.
- De forma a utilizar técnicas não-paramétricas necessitamos dividir em estratos de acordo com as categorias dessas covariáveis. Isto gera um número grande de estratos que podem conter poucas, ou talvez nenhuma observação e, portanto, impossibilita a incorporação de covariáveis, por exemplo, contínuas.
- Covariáveis são acomodadas, naturalmente, em um modelo com estrutura regressão.
- Modelos do tipo GAM (Generalized Additive Models) pode ser uma alternativa não-paramétrica para estas limitações.

Modelos (semi) Paramétricos

- Duas classes de modelos estão disponíveis para análise de dados de sobrevivência:
  - Modelos paramétricos ou de tempos de vida acelerados (Caps. 3 e 4);

$$T = \exp(X'\beta)T'$$

  - Modelo semi-paramétrico ou de modelo de taxas de falha proporcionais ou, simplesmente, modelo de Cox (Caps. 5 e 6).

$$\lambda(t; X) = \exp(X'\beta)\lambda'(t; X)$$

Modelos de Regressão Paramétricos

$$Y = \log T = X'\beta + \sigma\nu.$$

A distribuição de ν pode ser:

- Valor Extremo;
- Normal;
- log-Gama;
- Outras.

Modelos de Regressão Paramétricos

- População Homogênea - Cap. 3

$$Y = \log T = \mu + \sigma\nu.$$

ou

$$T \sim \text{distribuição}(\alpha, \gamma),$$

- População Heterogênea (incluindo covariáveis) - Cap. 4

$$Y = \log T = X'\beta + \sigma\nu.$$

- População Homogênea

$$Y = \log T = \mu + \sigma\nu.$$

  - μ: parâmetro de locação
  - σ: parâmetro de escala

ou

$$T \sim \text{distribuição}(\alpha, \gamma),$$

  - α: parâmetro de escala
  - γ: parâmetro de forma

- Existe uma relação determinística entre (μ, σ) e (α, γ).

Modelos Paramétrico para População Homogênea - Cap. 3.

- Principais Modelos para $\log T$
  - Valor Extremo;
  - Normal;
  - log-Gama;
  - log-gama generalizada;
  - Outras.
- Inferência Estatística
  - Função de Verossimilhança;
  - Estatísticas Teste (Wald, ERV)
  - Técnicas de Adequação dos modelos.
- Aplicações.

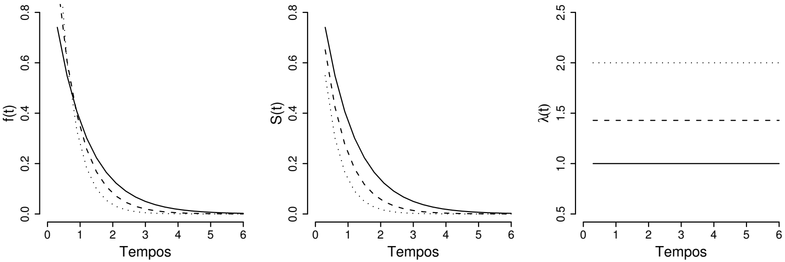
Modelo Exponencial

- O modelo mais simples.
- O modelo exponencial não tem memória.
- Único modelo com função de taxa de falha constante.
- $$f(t) = \frac{1}{\alpha} \exp(-(t/\alpha)), \quad t \geq 0$$

$$\lambda(t) = \frac{1}{\alpha}$$

Descrição do Modelo Exponencial

- linha contínua: $\alpha = 1,0$.
- linha tracejada: $\alpha = 0,7$.
- linha pontilhada: $\alpha = 0,5$.



Modelo Weibull

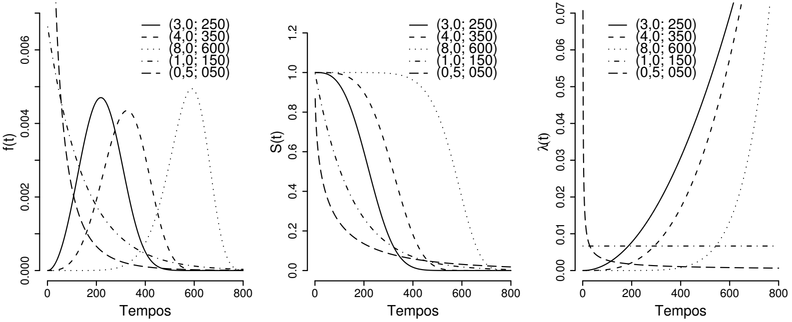
- O modelo foi proposto por Weibull (1951 e 1954).
- O modelo foi proposto em situações de engenharia envolvendo fadiga de metais.
- A distribuição de Weibull está relacionada com a de valores extremos. Isto significa que a falha de um produto complexo (muitas partes) ocorre quando falha a primeira parte.
- Este modelo é caracterizado por taxas de falha monótonas.
- O modelo exponencial é um caso especial.

$$f(t) = \frac{\gamma}{\alpha^\gamma} t^{\gamma-1} \exp(-(t/\alpha)^\gamma), \quad t \geq 0$$

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

Modelo Weibull

- primeira figura: $f(t)$.
- segunda figura: $S(t)$.
- terceira figura: $\lambda(t)$.



Modelos Weibull e Valor Extremo

- Se T tem distribuição Weibull, então $Y = \log T$ tem distribuição do valor extremo.

- Ou seja, o modelo Valor Extremo tem a seguinte forma:

$$f(y) = \frac{1}{\sigma} \exp((y - \mu)/\sigma) - \exp(-(y - \mu)/\sigma)), \quad t \in \mathcal{R}$$

$$\lambda(y) = \frac{1}{\sigma} \exp(-(y - \mu)/\sigma))$$

- $\mu = \log(\alpha)$ e $\sigma = 1/\gamma$.

Modelo Weibull

- O modelo Weibull possui as seguintes duas propriedades importantes em análise de sobrevivência.
  - Taxas de falha proporcionais;
  - Tempo de vida acelerado.
- Desta forma, podemos ter duas interpretações ao ajustarmos um modelo Weibull.
- No entanto, os parâmetros para as duas interpretações são distintos
$$\mu_{TP} = -\frac{\mu_{TA}}{\sigma}$$
- a saída do R (survreg) vai fornecer $\widehat{\mu}_{TA}$. A obtenção de $\widehat{\mu}_{TP}$ é imediata pela expressão acima mas o Erro Padrão necessita do uso do método delta.

Modelo log-normal

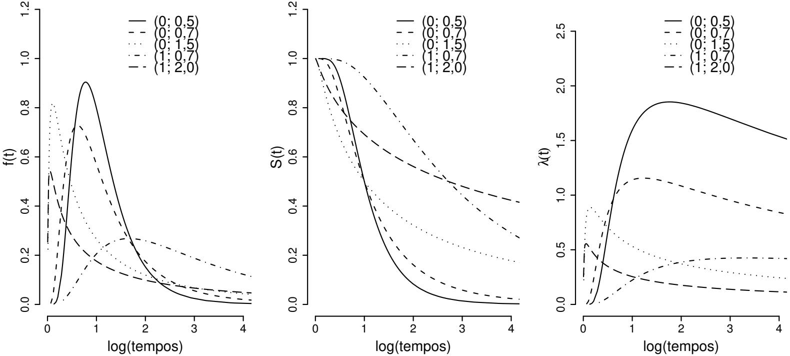
- O modelo log-normal é, juntamente com o Weibull, um modelo importante em análise de sobrevivência.
- Este modelo apresenta taxas de falhas não monótonas.

$$f(t) = \frac{1}{\sqrt{2\pi}t\sigma} \exp\left\{-\frac{1}{2\sigma^2}(\log(t) - \mu)^2\right\}, \quad t \geq 0$$

$$S(t) = \Phi\left(\frac{-\log(t) + \mu}{\sigma}\right)$$

$$\lambda(t) = \frac{f(t)}{S(t)}$$

Modelo Log-normal



Modelos log-normal e Normal

- Se T tem distribuição log-normal, então $Y = \log T$ tem distribuição normal ou Gaussiana.

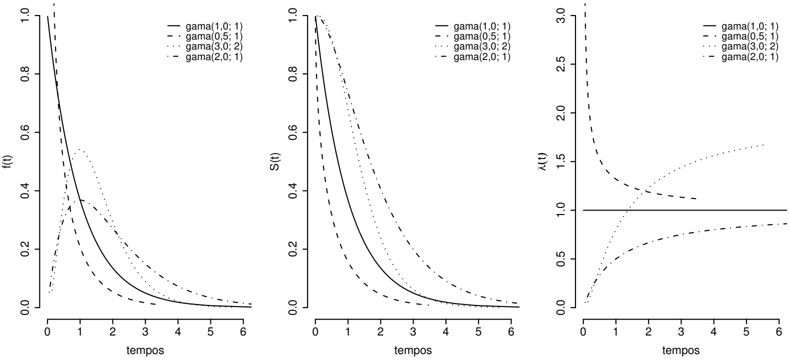
- Observe que o modelo log-normal é definido em termos dos parâmetros da distribuição geradora normal. Ou seja,

    - μ (parâmetro de escala da log-normal) é o parâmetro de locação da distribuição normal.

    - σ (parâmetro de forma da log-normal) é o parâmetro de escala da distribuição normal.

Modelo Gama

- O gama é outro modelo importante em análise de sobrevivência.
- Mostramos a seguir as formas de $f(t)$, $S(t)$ e $\lambda(t)$ para o modelo gama.



Modelo Gama Generalizado

- O modelo gama generalizado tem um parâmetro de escala e dois de forma.
- $$f(t) = \frac{\gamma}{\Gamma(k)\alpha^{\gamma k}} t^{\gamma k - 1} \exp\left\{-(\frac{t}{\alpha})^2\right\}, \quad t \geq 0$$
- Os principais modelos em análise de sobrevivência são casos particulares da gama generalizada:
  - para $k = 1$ e $\gamma = 1$ tem-se $T \sim \mathrm{Exp}(\alpha)$.
  - para $k = 1$ tem-se $T \sim \mathrm{Weibull}(\gamma, \alpha)$.
  - para $\gamma = 1$ tem-se $T \sim \mathrm{Gama}(k, \alpha)$.
  - para $k \to \infty$ tem-se $T \sim$ log-normal.

Modelo Gama Generalizado

- O modelo gama generalizado é complexo mas útil na seleção de modelos.
- Os parâmetros, ou uma função deles, não tem interpretação.
- Difícil de ajustar computacionalmente. É comum obtermos falta de convergência.
- O R ajusta o modelo gama generalizado no pacote flexsurv. Maiores informações podem ser obtidas em https://docs.ufpr.br/ giolo/Livro/ (contribuições extras).

Inferência para os Modelos Paramétricos

- Se corretamente especificado, os modelos paramétricos são bastante eficientes.
- Inferência para as quantidades desconhecidas dos modelos é baseada na função de verossimilhança e suas propriedades assintóticas.
- Cuidado na incorporação de censuras na função de verossimilhança.
- Má especificação de um modelo paramétrico acarreta em vício na estimação das quantidades de interesse.
- Técnicas de adequação, via resíduos, são fundamentais para verificar a adequação dos modelos paramétricos

Inferência no Modelo Paramétrico

- Estimador de Máxima de Verossimilhança e suas propriedades assintóticas.
- Inferência exata: praticamente não tem disponível.

Inferência Exata: Modelo Exponencial e Censura Tipo II

r falhas (fixo) e $n - r$ censuras em $t_{(r)}$.

RESULTADO: Considere a seguinte estatística

$$W = \sum_{i=1}^{r} t_{(i)} + (n-r)t_{(r)} = \sum_{i=1}^{n} t_i \qquad \text{(tempo total sob teste)}$$

- A estatística $2W/\alpha$ tem uma distribuição qui-quadrado com $2r$ graus de liberdade, em que α é o parâmetro do modelo exponencial.
- A partir deste resultado é possível construir intervalos de confiança e testes de hipóteses exatos para α.
- Um intervalo de $(1-\alpha)100\%$ de confiança para α é

$$(2W/\chi^2_{2r;1-\alpha/2}; 2W/\chi^2_{2r;\alpha/2})$$

Construção da Função de Verossimilhança

- Considere θ o vetor de parâmetros do modelo paramétrico; $f(t;\theta)$.
- Amostra aleatória

$$(t_1, \delta_1), \ldots, (t_n, \delta_n)$$

- Suposição: censura não-informativa. Ou seja, T e C são independentes.

Função de Verossimilhança (censura aleatória não-informativa)

- T tem função de densidade $f(.)$ e de sobrevivência $S(.)$.
- C tem função de densidade $g(.)$ e de sobrevivência $G(.)$.
- Observamos: $t = min(T, C)$ e $\delta = I(T < C)$ em uma amostra de tamanho n.

1. O i-ésimo indivíduo é uma censura:

$$P(t_i = t, \delta_i = 0) = P(C_i = t, T_i > C_i) = P(C_i = t, T_i > t) = g(t)S(t)$$

2. O i-ésimo indivíduo é um evento:

$$P(t_i = t, \delta_i = 1) = P(T_i = t, T_i < C_i) = P(T_i = t, C_i > t) = f(t)G(t)$$

Então a Função de Verossimilhança é dada por:

$$L(\theta, \nu) = \prod_{i=1}^{n} (f(t_i; \theta) \times G(t_i; \nu))^{\delta_i} (S(t_i; \theta) \times g(t_i; \nu))^{1-\delta_i}$$

Função de Verossimilhança (censura aleatória não-informativa)

$$
\begin{aligned}
L(\theta,\nu) &= \prod_{i=1}^{n} (f(t_i;\theta) \times G(t_i;\nu))^{\delta_i} (S(t_i;\theta) \times g(t_i;\nu))^{1-\delta_i} \\
&= \prod_{i=1}^{n} f(t_i;\theta)^{\delta_i} S(t_i;\theta)^{1-\delta_i} \times \prod_{i=1}^{n} G(t_i;\nu)^{\delta_i} g(t_i;\nu)^{1-\delta_i} \\
L(\theta,\nu) &= L(\theta)L(\nu)
\end{aligned}
$$

Função de Verossimilhança (censura não-informativa)

Então a Função de Verossimilhança para θ é dada por:

$$
\begin{aligned}
L(\theta) &= \prod_{i=1}^{n} f(t_i;\theta)^{\delta_i} S(t_i;\theta)^{1-\delta_i} \\
&= \prod_{i=1}^{n} \lambda(t_i;\theta)^{\delta_i} S(t_i;\theta),
\end{aligned}
$$

em que δ_i é o indicador de falha para a i-ésima observação e $i = 1, \ldots, n$.

Estimador de Máxima Verossimilhança

O vetor Escore é dado por:

$$
\begin{aligned}
U(\theta) &= \frac{\partial \log L(\theta)}{\partial \theta} \\
&= \frac{\partial l(\theta)}{\partial \theta}.
\end{aligned}
$$

O Estimador de Máxima Verossimilhança (EMV) é a solução do seguinte sistema de equações:

$$U(\widehat{\theta}) = 0$$

Usualmente não existe solução analítica para este sistema de equações.

Quantidades Importantes

- Vetor Escore:

$$U(\theta) = \frac{\partial \log L(\theta)}{\partial \theta} = \frac{\partial l(\theta)}{\partial \theta}.$$

- Matriz de Informação Observada:

$$I(\theta) = -\frac{\partial^2 \log L(\theta)}{\partial \theta^2} = -\frac{\partial^2 l(\theta)}{\partial \theta^2}.$$

- Matriz de Informação de Fisher.

$$\mathcal{I}(\theta) = E(I(\theta))$$

Solução do Sistema de Equações

$$\widehat{\theta}^{k+1} = \widehat{\theta}^{k} + I^{-1}(\widehat{\theta}^{k})U(\widehat{\theta}^{k})$$

em que U é o vetor escore de primeiras derivadas de $l(\theta)$ e I é a matriz de informação observada.

Usualmente o sistema é inicializado com $\theta^0 = 1$.

Propriedades Assintóticas EMV

- O EMV tem, assintoticamente, distribuição normal;
- O EMV é consistente;
- A estatística da Razão de Verossimilhança (RV)

$$-2\log(L(\theta)/L(\widehat{\theta}))$$

  tem, assintoticamente, uma distribuição qui-quadrado com gl igual a dimensão de θ.
- Invariância: Se $\widehat{\theta}$ é o EMV de θ, então $g(\widehat{\theta})$ é o EMV de $g(\theta)$.

Referências:

- Cox e Hinkley (1974), Theoretical Statistics
- Cordeiro (1992), Teoria da Verossimilhança.

Resultados Importantes (sob certas condições de regularidade).

- Propriedade Importante: $E(U(\theta)) = 0$

- $$Var(\widehat{\theta}) \approx \mathcal{I}(\theta)^{-1}$$

  e

  $$Var(U(\theta)) \approx \mathcal{I}(\theta).$$

- Para $\mathcal{I}(\theta)$ ser obtida é necessário especificar uma distribuição para os tempos de censura. Isso não é razoável em análise de sobrevivência. No entanto, $\mathcal{I}(\theta)$ é bem estimada por $I(\theta)$.

- $$\widehat{Var}(\widehat{\theta}) \approx I(\widehat{\theta})^{-1}$$

Estatísticas relacionadas ao EMV

1. WALD

   Exemplos:

   - Exponencial:
   $$\widehat{\alpha} \approx N(\alpha, \widehat{Var}(\widehat{\alpha}) \approx \widehat{\alpha}^2/r)$$
   em que r é o número de eventos/falhas.
   - Weibull
   $$\widehat{\theta} \approx N_2(\theta, \widehat{Var}(\widehat{\theta}) \approx I(\widehat{\theta})^{-1})$$

2. Ou escrevendo como forma quadrática:

$$W = (\widehat{\theta} - \theta)' \mathcal{I}(\theta)(\widehat{\theta} - \theta)$$

tem, para amostra grande, uma distribuição qui-quadrado com gl igual a dimensão de θ.

Estatísticas relacionadas ao EMV

2 Razão de Verosssimilhança (RV)

$$-2\ \log(L(\theta)/L(\widehat{\theta})) = 2(l(\widehat{\theta}) - l(\theta))$$

3 Estatística Escore S (Estatística de Rao)

$$U(\theta)'\mathcal{I}^{-1}(\theta)U(\theta)$$

para amostra grande, ambas estatísticas tem uma distribuição qui-quadrado com gl igual a dimensão de θ

Em resumo......

- Considere que a dimensão de θ é k.
- As três estatísticas podem ser utilizadas para testar hipóteses e construir intervalos de confiança.
- As três estatísticas tem assintoticamente uma distribuição qui-quadrado com k (dimensão de θ) graus de liberdade.
- $\mathcal{I}(\theta)$ vai ser estimada por $I(\widehat{\theta})$.
- As estatísticas da RV e S devem ser preferidas à de Wald.

Seleção/Adequação de Modelos

- Método Gráfico.
    - Modelo candidato vs Kaplan-Meier.
    - Linearização do modelo.
- Teste de hipóteses: utilizar o TRV para comparar o modelo proposto com a gama generalizada.

Seleção de Modelos: Gráficos

1. Comparação direta da $S(t)$ estimada do modelo proposto com o Kaplan-Meier (no mesmo gráfico). Duas formas: (1) $\widehat{S}_{KM}$ vs $\widehat{S}_{MP}$; (2) $\widehat{S}_{KM}$ vs tempo e $\widehat{S}_{MP}$ vs tempo.

2. Linearização da $S(t)$ para comparação com uma reta.

   - Modelo Weibull:

   $$\log - \log S(t) = \gamma \log(\alpha) + \gamma \log(t)$$

   - Modelo log-normal:

   $$\Phi^{-1} S(t) = \frac{-\log(t) + \mu}{\sigma}$$

Seleção de Modelos: Teste de Hipóteses (modelo gama generalizado)

Comparação direta do modelo proposto com o modelo gama generalizada via teste da razão de verossimilhança.

$$H_0 : \text{o modelo proposto é adequado}$$

$$TRV = 2(l(\widehat{\theta}_{GG}) - l(\widehat{\theta}_{MP}))$$

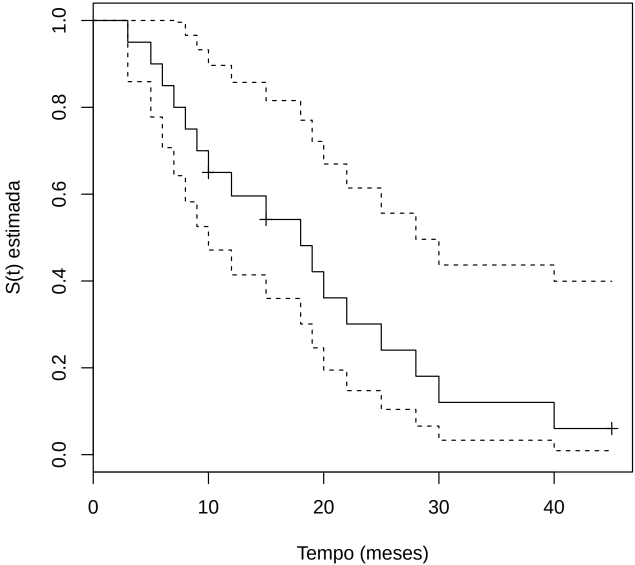
sob H_0 tem, para amostras grandes, uma distribuição qui-quadrado com o número de graus de liberdade igual a diferença do número de parâmetros dos modelos.

Exemplo: Pacientes com Câncer de Bexiga

- Dados provenientes da UFPR.
- 20 pacientes com câncer de bexiga submetidos a um procedimento cirúrgico a laser.
- Resposta: tempo da cirurgia até a reincidência da doença (meses).
- Objetivo: estimar o tempo médio e mediano de vida destes pacientes.
- Dados (em meses): 17 falhas e 3 censuras (ver pag. 101).

Câncer de Bexiga

Kaplan-Meier.

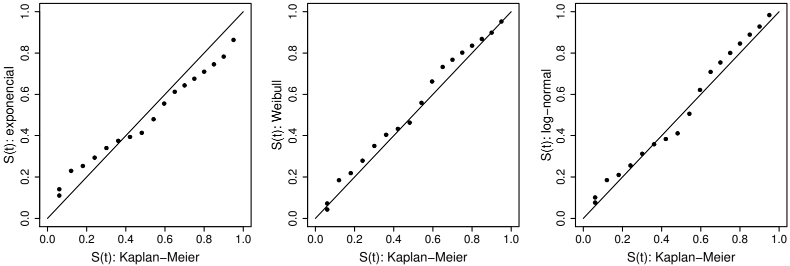


1. Tempo Médio de Vida
   - Maior tempo observado é uma censura em t=45 meses.
   - $\widehat{\mu} = 18,7$; (IC 95%(18, 73 $\pm$ 1, 96 $\times$ 2, 73) 13, 4; 24, 1) (mal estimado!!!)
2. Tempo Mediano de Vida
   - Inversa de KM e seu IC.
   - Usando interpolação.
   - $\widehat{t_{0,5}} = 17,1 \quad (9,5; 27,8)$

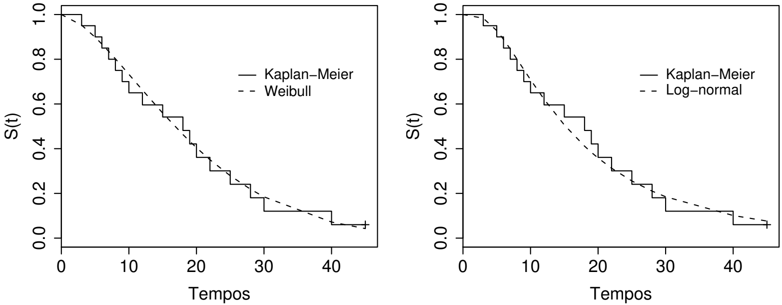
Exemplo: Pacientes com Câncer de Bexiga

- Exponencial: $\widehat{S}_e(t) = \exp(-t/20,41)$,
- Weibull: $\widehat{S}_w(t) = \exp(-(t/21,34)^{1,54})$,
- log-normal: $\widehat{S}_{ln}(t) = \Phi(-(\log(t) - 2,72)/0,76)$,

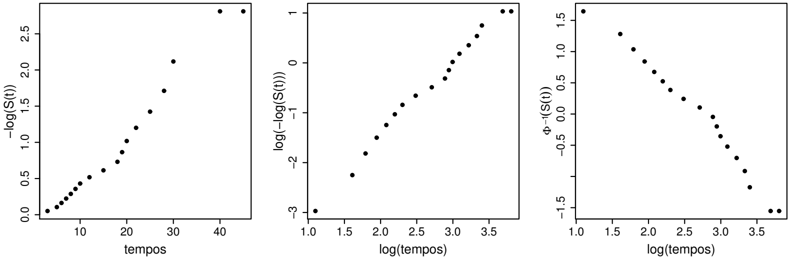
Câncer de Bexiga: Seleção de Modelos.
Modelo Proposto vs Kaplan-Meier.



Câncer de Bexiga: Modelo vs Kaplan-Meier.



Câncer de Bexiga: Gráficos de linearização.



Teste de Adequação de Modelos via Gama Generalizada.

Modelo	$l(\theta)$	TRV	valor-p
Gama Generalizada	-65,69		
Exponencial	-68,27	5,16	0,075
Weibull	-66,13	0,88	0,348
Log-normal	-65,74	0,10	0,752

Conclusão: podemos descartar o modelo exponencial mas não conseguimos diferenciar entre os modelos log-normal e Weibull. A provável razão é o pequeno tamanho de amostra.

Respondendo às Perguntas de Interesse. Modelo log-normal.

1. $E(T)$ (tempo médio de vida):

$$\widehat{E}(T) = \widehat{\alpha}\Gamma(1 + (1/\widehat{\gamma})) = 19,2 \text{ meses} \quad \text{(Weibull)};$$

$$\widehat{E}(T) = \exp(\widehat{\mu} + \widehat{\sigma}^2/2) = 20,3 \text{ meses} \quad \text{(log-normal)};$$

2. Tempo mediano de vida:

$$\widehat{t}_{0,5} = \widehat{\alpha}(-log(1 - 0,5))^{1/\widehat{\gamma}} = 16,8\text{meses} \quad \text{(Weibull)};$$

$$\widehat{t}_{0,5} = \exp(z_{0,5}\widehat{\sigma} + \widehat{\mu}) = 15,2\text{meses} \quad \text{(log-normal)};$$

3. $S(20)$: probabilidade de sobreviver a 20 meses:

$$\widehat{S}(20) = 0,40 \quad \text{(Weibull)}$$

$$\widehat{S}(20) = 0,36 \quad \text{(log-normal)};$$

.

Construção dos Intervalos de Confiança para $\phi = g(\theta)$.

Possíveis soluções:

- Encontrar a $Var(\widehat{\phi})$ (Método Delta).
  - Possivelmente usando uma reparametrização (de acordo com os resultados do R).
  - Exemplo (log-normal):Reparametrizar $E(T)$ em função de μ e $\log(\sigma)$.
- Usar reamostragem bootstrap.
- Inferência Bayesiana (Intervalo de Credibilidade).

Construção dos Intervalos de Confiança para $\phi = g(\theta)$.
Via Método Delta.

1. θ é um escalar:

$$Var(\widehat{\phi}) = Var(g(\widehat{\theta})) \approx Var(\widehat{\theta})(\frac{dg(\theta)}{d\theta})^2$$

2. $\theta = (\gamma, \alpha)'$:

$$\begin{aligned} Var(\widehat{\phi}) &\approx Var(\widehat{\alpha})(\frac{\partial\phi}{\partial\alpha})^2 + 2Cov(\widehat{\alpha},\widehat{\gamma})(\frac{\partial\phi}{\partial\alpha})(\frac{\partial\phi}{\partial\gamma}) \\ &+ Var(\widehat{\gamma})(\frac{\partial\phi}{\partial\gamma})^2 \end{aligned}$$

Resultado do R - modelo log-normal.

$$E(T) = \exp(\mu + \sigma^2/2);$$

Objetivo: encontrar $Var(\widehat{\phi})$ = $Var(\widehat{E}(T))$

$>$ ajust3$*var*

	(Intercept)	Log(scale)
(Intercept)	0.031061677	0.002706896
Log(scale)	0.002706896	0.030119031

Solução 1 - Melta Delta: modelo log-normal.

Tomando $\sigma = \exp(\eta)$, em que $\eta = \log\sigma$ temos

$$Var(\widehat{\sigma}) \approx (\sigma)^2 Var(\widehat{\log\sigma});$$

e

$$\widehat{Var}(\widehat{\sigma}) \approx 0.7648167^2 0.030119031 = 0.01761796;$$

e

$$\widehat{Cov}(\widehat{\mu},\widehat{\sigma}) \approx \widehat{\sigma}\,\widehat{Cov}(\widehat{\mu},\widehat{log}\sigma) = 0.7648167 * 0.002706896 = 0.00207.$$

$$\begin{aligned}
\widehat{Var}(\widehat{E}(T)) &\approx \widehat{Var}(\widehat{\mu})(\exp(\widehat{\mu}+\widehat{\sigma}^2/2))^2 + \widehat{Var}(\widehat{\sigma})(\widehat{\sigma}\exp(\widehat{\mu}+\widehat{\sigma}^2/2))^2 \\
&+ 2\widehat{Cov}(\widehat{\mu},\widehat{\sigma})(\exp(\widehat{\mu}+\widehat{\sigma}^2/2))(\widehat{\sigma}\exp(\widehat{\mu}+\widehat{\sigma}^2/2)) \\
&= (0,031)(20,263)^2 + 2(0,00207)(0,76)(20,263)^2 \\
&+ (0,0176)((0,76)(20,263))^2 = 18,2
\end{aligned}$$

O intervalo de confiança para $E(T)$ é $20,263 \pm 1,96\sqrt{18,2}$, ou seja, $(11,9; 28,6)$ meses.

Solução 2 - Método Delta: modelo log-normal.

Se $\eta = \log \sigma$ significa que $\sigma = \exp(\eta)$. Então:

$$E(T) = \exp(\mu + \exp(2\eta)/2)$$

$$\begin{aligned}
\widehat{Var}(\widehat{E}(T)) &\approx \widehat{Var}(\widehat{\mu})(\exp(\widehat{\mu} + \widehat{\sigma}^2/2))^2 + \widehat{Var}(\widehat{\eta})(\widehat{\sigma}^2 \exp(\widehat{\mu} + \widehat{\sigma}^2/2))^2) \\
&+ 2\widehat{Cov}(\widehat{\mu}, \widehat{\sigma})(\exp(\widehat{\mu} + \widehat{\sigma}^2/2))(\widehat{\sigma}^2 \exp(\widehat{\mu} + \widehat{\sigma}^2/2)) \\
&= (0,031)(20,263)^2 + 2(0,00270)(0,76)^2(20,263)^2 \\
&+ (0,0301)((0,76)^4(20,263))^2 = 18,2
\end{aligned}$$

O intervalo de confiança para $E(T)$ é $20,263 \pm 1,96\sqrt{18,2}$, ou seja, $(11,9; 28,6)$ meses.

Resultado do R - modelo log-normal.

Encontrar um IC para $t_{0,5}$.

$$t_{0,5} = \exp(z_{0,5}\sigma + \mu) = e^{\mu}.$$

Utilizando o método delta temos:

$$\widehat{Var}(\widehat{t}_{0,5}) = \widehat{Var}(\widehat{\mu})\exp(2\widehat{\mu}) = 0,00311 e^{2\times 2,72} = 7,16.$$

Um intervalo de 95% de confiança para $t_{0,5}$ é dado por

$$15,2 \pm 1,96\sqrt{7,16} = (10,0; 20,4) \text{ meses}$$

Intervalo de Confiança Bootstrap

- Amostragem com reposição;
- Unidade (t_i, δ_i); $i = 1, \ldots, n$.
- B amostras de tamanho n.
- Resultados - log-normal, B=5000
  - tempo médio
    - Intervalo de Confiança percentílico: $20,3 \ \ (14,0; 30,1)$;
    - Intervalo de Confiança Gaussiano: $20,3 \ \ (12,0; 28,6)$.
  - tempo mediano
    - Intervalo de Confiança percentílico: $15,1 \ \ (9,8; 20,5)$
    - Intervalo de Confiança Gaussiano: $15,1 \ \ (10,8; 21,5)$;

Resumo dos resutados: Modelo log-normal.

Quantidade		KM	Assintótico	Boot perc.	Boot normal
Média	Est. Pontual	18,7	20,3	20,3	20,3
	Est. Intervalar	(13,4; 24,1)	(11,9; 28,6)	(14,0; 30,1)	(12,0; 28,6)
Mediana	Est. Pontual	17,1	15,2	15,1	15,1
	Est. Intervalar	(9,5; 27,8)	(10,0; 20,4)	(9,8; 20,5)	(10,8; 21,5)

Conclusões:

- as estimativas do KM para a média são viciadas (subestimada) tanto em termos pontuais quanto intervalares;
- as estimativas assintóticas e bootstrap são bastante similares;
- como esperado, as estimativas do KM para a mediana são menos precisas (maiores IC) do que as demais.

1. Incluir covariáveis na modelagem.
2. Usualmente covariáveis são incluídas no modelo via uma estrutura de regressão.
3. As distribuições usuais para o tempo até a falha é escala-forma.
4. Desta forma, modelamos $Y = \log T$ que é um modelo locação-escala.

Características Gerais do Modelo

- INTERPRETAÇÃO: O modelo paramétrico específica um efeito multiplicativo das covariáveis em T. O papel de X é acelerar ou desacelerar o tempo até a falha.

- ESTIMAÇÃO: Os parâmetros do modelo são estimados pelo método de máxima verossimilhança.

- ADEQUAÇÃO DO MODELO: a adequação do modelo ajustado é realizada utilizando os resíduos:

$$\widehat{\nu}_i = \frac{Y_i - x_i'\widehat{\beta}}{\widehat{\sigma}}, \quad i = 1, \ldots, n.$$

  Os resíduos são estimativas dos erros, que vêm de uma população homogênea.

Modelo de Tempos de Vida Acelerados

- Covariáveis agem multiplicativamente na escala do tempo.
- Modelo interpretado em termos da velocidade de progressão do tempo de acompanhamento.
- Forma do Modelo para uma única covariável representada por dois grupos (0 e 1):

$$S_1(t) = S_0(\phi t)$$

  - Se $\phi < 1$, o grupo 1 tem o tempo acelerado quando comparado com o 0.
  - Se $\phi < 1$ e o evento for morte, o grupo 1 é melhor que o 0 no sentido que caminha para o evento de forma mais devagar.

Modelo de Tempos de Vida Acelerados - Interpretação

- Forma do Modelo para uma única covariável representada por dois grupos (0 e 1):

$$S_1(t) = S_0(\phi t)$$

  - Interpretação em termos de percentis, em especial tempos medianos:

$$S_1(t^1_{.5}) = S_0(t^0_{.5}) = 0,50$$

    ou

$$S_1(t^1_{.5}) = S_0(\phi t^1_{.5})$$

    ou

$$t^0_{.5} = \phi t^1_{.5} \quad \rightarrow \quad \frac{t^0_{.5}}{t^1_{.5}} = \phi.$$

    Ou seja, razão de tempos medianos.

Modelo de Tempos de Vida Acelerados

$$Y = \log T = X'\beta + \sigma\nu$$

Exemplo: Modelo de Regressão Exponencial.

- Modelo é linear em $\log T$.
- Modelo é multiplicativo em T.

$$T = \exp(X'\beta)\eta$$

em que $\log(\eta) = \sigma\nu$.

Modelo de Regressão Exponencial

Desta forma:

$$S(y/x) = \exp(-\exp(y - x'\beta))$$

e

$$S(t/x) = \exp\left(-\frac{t}{\exp(x'\beta)}\right)$$

Em outras palavras:

- Y tem uma distribuição do valor extremo padrão com parâmetro de locação $X'\beta$ e

- T tem uma distribuição exponencial com parâmetro de escala

$$\alpha = \exp(X'\beta)$$

Modelos de Regressão Paramétricos

De forma equivalente podemos definir os modelos Weibull e log-normal. Por exemplo, para o modelo Weibull,

- T tem uma distribuição Weibull com parâmetro de escala

  $$\alpha = \exp(X'\beta)$$

  e de forma γ.

- Y tem uma distribuição do valor extremo padrão com parâmetro de locação $X'\beta$ e escala $\sigma = 1/\gamma$.

Inferência para $\theta = (\beta, \sigma)$ - Mecanismo de censura não-informativo

A função de verossimilhança para uma amostra de tamanho n é dada por:

$$
\begin{aligned}
L(\theta) &= \prod_{i=1}^{n} f(y_i; x_i, \theta)^{\delta_i} S(y_i; x_i, \theta)^{1-\delta_i} \\
&= \prod_{i=1}^{n} \lambda(y_i; x_i, \theta)^{\delta_i} S(y_i; x_i, \theta),
\end{aligned}
$$

em que δ_i é o indicador de falha para a i-ésima observação.

Valem todas as propriedades, para grandes amostras, do EMV e das estatísticas de teste.

Técnicas de Adequação do Modelo

1. Resíduos padronizados.

$$\widehat{\nu}_i = \frac{y_i - x_i'\widehat{\beta}}{\widehat{\sigma}}, \quad i = 1, \ldots, n,$$

   - Os resíduos vêm de uma população homogênea. Observe que os resíduos de observações censuradas, também são censurados.
   - Podemos utilizar as técnicas apresentadas para população homogênea (Cap. 3). Comparação dos resíduos do modelo proposto com o Kaplan-Meier.
   - Devemos usar os resíduos na escala original, $\exp(\widehat{\nu}_i)$, para compará-los com o Kaplan-Meier.

2. Teste RV - usando a Gama generalizada.

Outros Resíduos: Cox-Snell, martigalas, etc

- Definição Geral de resíduos (Cox e Snell, 1963).
- Forma do resíduo:

$$\widehat{e}_i = \widehat{\Lambda}(t_i/x_i)$$

- Se o modelo for adequado os $\widehat{e}_i$'s têm uma distribuição exponencial padrão ($\alpha = 1$).
- Cuidado: existem evidências de situações em que o modelo não é adequado e os $\widehat{e}_i$'s têm a forma de uma exponecial padrão.
- Estes resíduos, juntamente com os de martingala, são mais indicados para indicar a forma funcional de covariáveis contínuas.

Interpretação dos Coeficientes Estimados

- Importante interpretar na escala de T. Observe sempre que usamos uma escala transformada (logarítmica) na modelagem estatística.

- Lembre que:

$$E(\log T) \neq \log E(T)$$

- Interpretação: Razão de tempos medianos = $\exp(\beta)$.

- Exemplo: $\exp(\widehat{\beta}) = 2$.

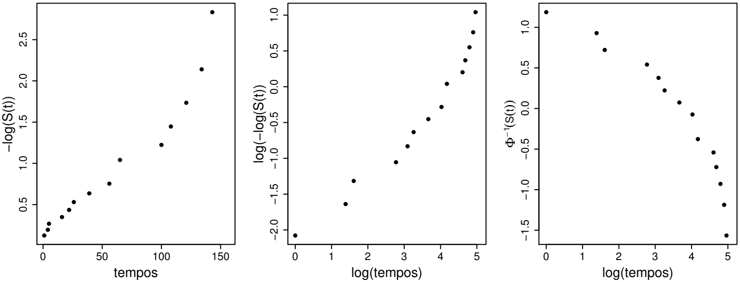
  Isto significa que o tempo mediano de um grupo é duas vezes o do outro grupo (mantendo fixa as demais covariáveis).

Estudo de Caso (Feigl e Zelen, 1965): Leucemia - p. 129

- Livro: Cox e Snell (1981, Applied Statistics: Principles and Examples.), p. 148.
- 17 pacientes com leucemia.
- Resposta: tempo (semanas) do diagnóstico até a morte do paciente.
- Objetivo: modelar a resposta em termos da contagem de glóbulos brancos (WBC) no diagnóstico.
- $x = \log_{10} \text{WBC}$

Estudo de Caso: Leucemia - p. 129

Preliminarmente, vamos ignorar a covariável.



Estudo de Caso: Leucemia

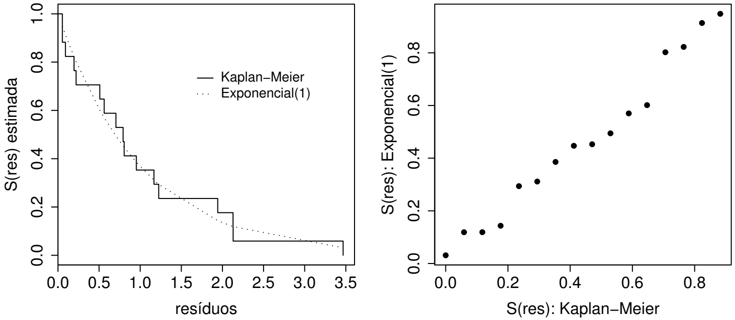
Este banco de dados foi exaustivamente analisado na literatura utilizando o modelo de regressão exponencial.

Termo	Exponencial	Weibull	log-normal
$\widehat{\beta}_0$	8,48	8,44	
$\widehat{\beta}_1$	-1,11	-1,10	
forma	fixo=1	1,02	
loglik	-83,88	-83,87	-83,76

TRV: Weibull vs Exponencial.

TRV = 2(83,88 - 83,87) = 0,013 (valor-p=0,915).

Análise de Resíduos - Modelo Exponencial



Interpretação do Ajuste

- X: $\log_{10}$ WBC;
- Interpretação na escala original WBC é multiplicativa;
- Defina p: proporção de aumento/redução em WBC;
- $\widehat{\beta}_1 = -1,11$

Desta forma:

$$\widehat{\mathrm{R}TM} = exp(\widehat{\beta}_1 \times \log_{10} p)$$

Interpretação:

- $p = 1,1$ (aumento em 10%) - $\widehat{\mathrm{R}TM} = exp(-1,11 \times \log_{10} 1,1) = 0,96$
- $p = 1,2$ (aumento em 20%) - $\widehat{\mathrm{R}TM} = 0,92$
- $p = 0,9$ (redução em 10%) - $\widehat{\mathrm{R}TM} = 1,05$

Curvas de Sobrevivência

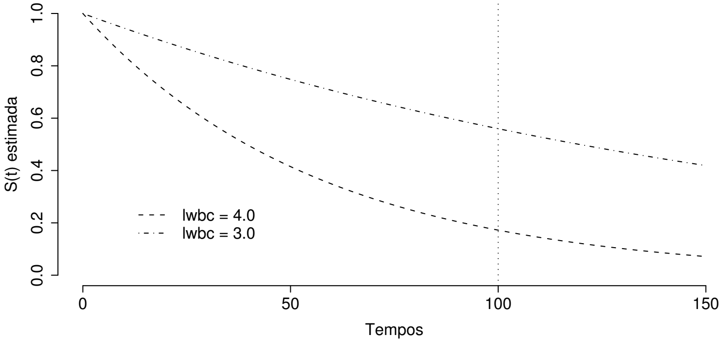
$$\widehat{S}(t/\log_{10}(WBC)) = \exp\left(\frac{t}{\exp(8,48 - 1,11\log_{10}(WBC))}\right)$$

$$\widehat{S}(100/\log_{10}(WBC) = 4) = 0,172$$

e

$$\widehat{S}(100/\log_{10}(WBC) = 3) = 0,559$$

Curvas de Sobrevivência - Modelo Exponencial



Estudo de Caso: Aleitamento Materno

- Estudo realizado pelos Profs. Eugênio Goulart e Cláudia Lindgren do Departamento de Pediatria da UFMG.
- O estudo foi realizado no Centro de Saúde São Marcos, ambulatório municipal de BH, que atende essencialmente a população de baixa renda.
- O objetivo principal é conhecer a prática do aleitamento materno de mães que utilizam este centro, assim como os possíveis fatores de risco ou de proteção para o desmame precoce.
- Um inquérito epidemiológico composto por questões demográficas e comportamentais foi aplicado a 150 mães de crianças menores de 2 anos de idade.
- A variável resposta de interesse foi o tempo máximo de aleitamento materno, ou seja, o tempo contado a partir do nascimento até o desmame completo da criança.

Estudo de Caso: Avaliação da Prática de Aleitamento Materno

Código	Covariável
V1	Experiência anterior de amamentação
V2	Número de filhos vivos (2)
V3	Conceito materno sobre o tempo ideal de amamentação (6 meses)
V4	Dificuldade para amamentar nos primeiros dias pós-parto
V5	Tipo de serviço em que realizou o (público) pré-natal
V6	Recebeu exclusivamente leite materno na maternidade
V7	A criança tem contato com o pai
V8	Renda per capita (1 SM)
V9	Peso ao nascimento (2,5 kgs)
V10	Tempo se separação mãe-filho pós-parto (6 horas)
V11	Permanência no berçário

UM ROTEIRO PARA A CONSTRUÇÃO DE MODELOS

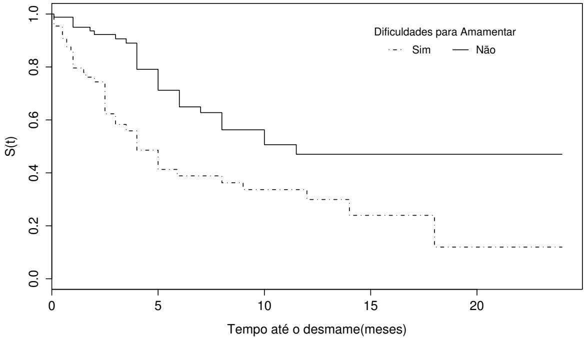
1. Entender o estudo: detalhes técnicos, clínicos e objetivos.
2. Desenho do Estudo: tipo, possíveis viéses, confundimento, etc.
3. Banco de Dados: arquivo, formato, dicionário, etc.
4. Exploração e Verificação da Consistência do Banco de Dados: cada variável separadamente, dados perdidos, etc.
5. Análise Descritiva:
    - Kaplan-Meier;
    - log-rank, Wilcoxon, etc.
6. Regra Empírica: excluir covariáveis com valor-p $> 0,25$

UM ROTEIRO PARA A CONSTRUÇÃO DE MODELOS

7. Construção de Modelos de Regressão:
   - Avaliar a forma funcional de covariáveis contínuas, eventualmente categorizá-las;
   - Utilizar alguma técnica de construção de modelos (por exemplo, "stepwise").
   - Utilizar de preferência o Teste da RV;
   - Dar um tratamento para os dados perdidos;
   - Investigar possíveis associações entre as covariáveis (colinearidade);
   - Obter um "Modelo Final "utilizando todos estes passos.
8. Incluir possíveis termos de interação.
9. Verificar a adequação do modelo ajustado.
10. Interpretar o modelo final apresentando intervalos de confiança para as quantidades de interesse.
11. Escrever o Relatório.

ANÁLISE EXPLORATÓRIA: Prática de Aleitamento Materno

- Curvas de Kaplan-Meier
- Testes de Wilcoxon e log-rank.



TESTES DE IGUALDADE DE CURVAS DE SOBREVIVÊNCIA

	Testes (valor-p)	
Covariável	logrank	Wilcoxon
Experiência de amamentação	$3,95\ (0,047)$	$6,73\ (0,010)$
Número de filhos vivos	$2,60\ (0,107)$	$2,02\ (0,155)$
Tempo de amamentação	$6,15\ (0,013)$	$8,54\ (0,004)$
Dificuldade para amamentar	$12,26\ (0,001)$	$15,45\ (< 0,001)$
Tipo de serviço do pré-natal	$1,38\ (0,241)$	$1,09\ (0,296)$
Recebeu leite materno	$7,47\ (0,006)$	$6,31\ (0,012)$
Contato com o pai	$1,84\ (0,175)$	$0,90\ (0,344)$
Renda per capita	$2,11\ (0,146)$	$2,60\ (0,107)$
Peso ao nascimento	$1,87\ (0,171)$	$2,59\ (0,108)$
Separação mãe-filho	$2,60\ (0,107)$	$0,97\ (0,325)$
Permanência no berçário	$2,93\ (0,087)$	$0,90\ (0,343)$

CRITÉRIO: Manter todas as covariáveis com pelo menos um valor-p $< 0,25$.

ESTRATÉGIA PARA A SELEÇÃO DE VARIÁVEIS (Collett, 1994)

1. Ajustar todos os modelos contendo uma única covariável. Incluir todas as covariáveis que forem significativas ao nível de $0,10$.
2. As covariáveis significativas no passo 1 são então ajustadas conjuntamente. Ajustamos modelos reduzidos, excluindo uma única covariável.
3. Neste passo as covariáveis excluídas no passo 2 retornam ao modelo para confirmar que elas não são estatisticamente significantes.
4. Neste passo retornamos com as covariáveis excluídas no passo 1 para confirmar que elas não são estatisticamente significantes.
5. Ajustamos um modelo incluindo as covariáveis significativas no passo 4. Neste passo testamos se alguma delas podem ser retiradas do modelo.
6. Ajustamos o modelo final para os efeitos principais. Devemos verificar a possibilidade de inclusão de termos de interação.

PONTOS IMPORTANTES:

- Incluir as informações clínicas no processo de decisão;
- Evitar ser muito rigoroso ao testar cada nível individual de significância. É recomendado um valor próximo de $0,10$.
- Utilizar o modelo Gama Generalizado no processo de seleção de variáveis.

Seleção de Variáveis - Utilizando o modelo gama generalizado.

Passos	Modelo	$-2 \log L(\theta)$	Estatística de teste (*TRV*)	Valor *p*
Passo 1	Nulo	335,540	—	—
	V1	330,235	5,305	0,0212
	V2	332,715	2,825	0,0933
	V3	329,746	5,794	0,0161
	V4	322,692	12,848	0,0003
	V5	333,756	1,784	0,1816
	V6	328,524	7,016	0,0080
	V7	333,291	2,249	0,1337
	V8	332,561	2,979	0,0843
	V9	332,592	2,948	0,0859
	V10	333,599	1,941	0,1635
	V11	333,449	2,091	0,1481
Passo 2	V1+V2+V3+V4+V6+V8+V9	304,038	—	—
	V2+V3+V4+V6+V8+V9	305,287	1,249	0,2639
	V1+V3+V4+V6+V8+V9	304,165	0,127	0,7226
	V1+V2+V4+V6+V8+V9	307,398	3,360	0,0667
	V1+V2+V3+V6+V8+V9	312,484	9,446	0,0021
	V1+V2+V3+V4+V8+V9	309,478	5,440	0,0201
	V1+V2+V3+V4+V6+V9	307,512	3,474	0,0623
	V1+V2+V3+V4+V6+V8	305,346	1.308	0,2527

Seleção de Variáveis - Utilizando o modelo gama generalizado.

Passo 3	V3+V4+V6+V8	307,485	—	—
	V3+V4+V6+V8+V1	305,529	1,956	0,1619
	V3+V4+V6+V8+V2	306,357	1,128	0,2882
	V3+V4+V6+V8+V9	306,382	1,103	0,2936
Passo 4	V3+V4+V6+V8	307,485	—	—
	V3+V4+V6+V8+V5	307,485	0,000	1,0000
	V3+V4+V6+V8+V7	305,725	1,759	0,1847
	V3+V4+V6+V8+V10	307,231	0,253	0,6149
	V3+V4+V6+V8+V11	307,322	0,163	0,6864
Passo 5	V3+V4+V6+V8	307,485	—	—
	V4+V6+V8	311.306	3,821	0,0506
	V3+V6+V8	320,594	13,109	0,0003
	V3+V4+V8	312,582	5,097	0,0239
	V3+V4+V6	312,999	5,514	0,0188
Passo 6	V3+V4+V6+V8	307,485	—	—
	V3+V4+V6+V8+V3*V4	306,777	0,708	0,4004
	V3+V4+V6+V8+V3*V6	305,678	1,807	0,1789
	V3+V4+V6+V8+V3*V8	307,206	0,279	0,5973
	V3+V4+V6+V8+V4*V6	306,735	0,750	0,3864
	V3+V4+V6+V8+V4*V8	306,740	0,745	0,3883
	V3+V4+V6+V8+V6*V8	307,200	0,285	0,5941
Etapa Final*	V3+V4+V6+V8	307,485		
	V1+V3+V4+V6	309,544		
Modelo Final	V1+V3+V4+V6	309,544		

* Escolha baseada em evidências clínicas e discussões com o pesquisador

Estudo de Caso: Avaliação da Prática de Aleitamento Materno

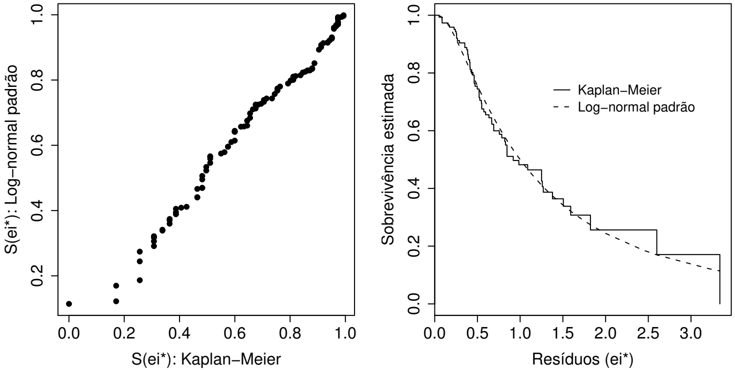
- MODELO MAIS SIMPLES
  - Os populares modelos Weibull e Log-normal são casos particulares do Gama Generalizado;
  - O teste da razão de verossimilhança é utilizado para selecionar os modelos.
  - Adequação do modelo Weibull: TRV = 5,35 (valor-p $= 0,021$) e;
  - Adequação do modelo log-normal: TRV = 0,218 (valor-p $= 0,641$).

- ADEQUAÇÃO DO MODELO LOG-NORMAL
  - Vamos utilizar os resíduos:

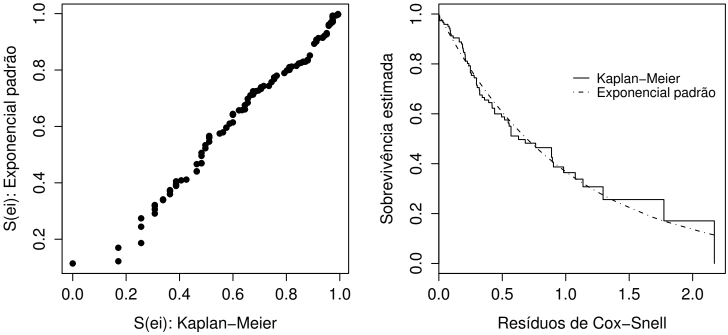
$$\widehat{\nu}_i = \frac{Y_i - x_i'\widehat{\beta}}{\widehat{\sigma}}.$$

  - Se o modelo log-normal para o tempo T estiver bem ajustado para estes dados, a distribuição dos resíduos deve estar bastante próxima de uma log-normal padrão.

Análise de Resíduos - Modelo Log-normal



Análise de Resíduos - Modelo Log-normal



- AJUSTE DO MODELO LOG-NORMAL

Covariável	Estimativa	Erro-Padrão	Valor-p
Constante	-0,584	0,309	0,059
Exper. Amam. (V1)	-0,572	0,301	0,057
Conc. Amam. (V3)	-0,631	0,270	0,029
Dific. Amam. (V4)	-0,824	0,302	0,006
Leite Excl. (V6)	-0,681	0,293	0,020
Par. Forma	1,439	0,130	0,001

- INTERPRETAÇÃO DOS COEFICIENTES:

  Os coeficiente estimados estão expressos na escala logaritímica do tempo. Tomando o exponencial dos coeficientes estaremos obtendo a razão dos tempos medianos de sobrevivência.

- EXEMPLO: uma covariável codificada $0/1$, esta razão compara os tempos medianos de sobrevivência do grupo 1 com relação ao grupo 0.

INTERPRETAÇÃO DOS RESULTADOS:

- o tempo mediano até o desmame de mães que não tiveram experiência anterior de amamentação é aproximadamente a metade (IC 90%: 0,34;0,93)daquele das mães que já tiveram essa experiência;
- as mães que acreditam que o tempo ideal de amamentação é superior a seis meses têm um tempo médiano até o desmame de aproximadamente duas vezes (IC 90%: 1,21;2,93) o tempo das mães que pensam ser esse tempo inferior ou igual a seis meses;
- o tempo mediano até o desmame das mães que não apresentaram dificuldades de amamentar nos primeiros dias após o parto é 2,3 vezes (IC 90%: 1,39; 3,75) o tempo das que sofreram essas dificuldades;
- as crianças que receberam exclusivamente leite materno na maternidade têm um tempo mediano de amamentação que é duas vezes (IC 90%: 1,22; 3,20) o tempo daquelas que receberam outra tipo de alimentação juntamente com o leite materno.

Métodos Modernos ou Nem Tão Modernos Assim.....

✓ Tratamento de Dados Perdidos:

- Imputação Múltipla: Rubin (1987)
- Ponderação: Robins (2000, 2002, 2005)
- Distribuição Preditiva (Bayesiano)

✓ Relaxar suposições de linearidade e aditividade

- Modelos semi-paramétricos (splines)
- Não-paramétricos: Classificação/Árvores/Aprendizado de Máquina.

✓ Computacionalmente Intensivos

- Bootstrap.
- Monte Carlo

✓ Penalizações/Suavizações/Encolhimento.

✓ Etc, etc, etc.

- Imputação Múltipla
  - Rubin (1987).
  - Marshall, Altman e Holder (BMJ, 2010)

Imputação Múltipla - Pacote MICE - R

1. Etapa 1- Imputar
   - Construir modelos para as variáveis com dados perdidos.
   - Imputar os valores perdidos, completando o banco de dados ("Chained Equations").
   - Completar mais de um banco de dados para estimar o erro de predição/imputação.
2. Etapa 2- Ajustar
   - Ajustar o modelo de Cox para cada um dos banco de dados imputados.
3. Etapa 3- Combinar
   - Usar as equações de Rubin para combinar as estimativas e erros padrões.

Imputação Múltipla - Observações Técnicas

1. MI é aceito nos mecanismos MCAR e MAR.
2. Modelos para Imputação
   - resposta (modelos de sobrevivência)
   - covariáveis: binária (logístico), contínua (regressão linear), etc;
3. Chained Equations: 20 ciclos (estabilizar os resultados).
4. Combinação dos Resultados
   - Estimativa dos coeficientes: médias dos coeficientes estimados nos bancos imputados (Rubin, 1987).
   - Variância das estimativas: combinar a variância intra-imputação com a variância entre-imputações.