

# ANÁLISE DE SOBREVIVÊNCIA APLICADA

Modelo de Regressão de Cox - Cap. 5

Enrico A. Colosimo/UFMG

Depto. Estatística - ICEx - UFMG

- ▶ Modelo de Tempo de Vida Acelerado ou Modelo Paramétrico.
- ▶ Modelo de Taxas de Falhas Proporcionais ou Modelo Semiparamétrico de Cox.

## Modelo de Regressão de Cox

O modelo de taxas de falhas proporcionais ou, também chamado simplesmente, de regressão de Cox (Cox, 1972):

- abriu uma nova fase na modelagem de dados de sobrevivência;
- atualmente é o modelo mais popular na análise de dados de sobrevivência na área clínica/biológica;
- permite incorporar facilmente covariáveis dependentes do tempo, que ocorrem com frequência em várias áreas de aplicação.

## MODELO DE REGRESSÃO DE COX

- ▶ O modelo de Cox assume a seguinte forma para a função de taxa de falha:

$$\lambda(t) = \lambda_0(t) g(\mathbf{x}'\boldsymbol{\beta}),$$

em que  $g$  é uma função não-negativa tal que  $g(0) = 1$ .

- O componente não-paramétrico,  $\lambda_0(t)$ , não é especificado e é uma função não-negativa do tempo.
- $\lambda(t) = \lambda_0(t)$ , função de base, pode ser obtido para  $\mathbf{x} = \mathbf{0}$ .
- O componente paramétrico é frequentemente usado na seguinte forma multiplicativa:

$$g(\mathbf{x}'\boldsymbol{\beta}) = \exp\{\mathbf{x}'\boldsymbol{\beta}\} = \exp\{\beta_1 x_1 + \dots + \beta_p x_p\},$$

em que  $\boldsymbol{\beta}$  é o vetor de parâmetros associado às covariáveis.

## MODELO DE REGRESSÃO DE COX

- ▶ O modelo semi-paramétrico de Cox assume que:
  - as covariáveis atuam multiplicativamente na taxa de falha pela relação  $g(\mathbf{x}'\beta) = \exp\{\mathbf{x}'\beta\}$  e
  - nenhuma forma paramétrica para  $\lambda_0(t)$ .
- ▶ A suposição de taxas de falhas proporcionais significa que, para dois indivíduos diferentes  $i$  e  $j$ , temos que,

$$\frac{\lambda_i(t)}{\lambda_j(t)} = \frac{\lambda_0(t) \exp\{\mathbf{x}'_i\beta\}}{\lambda_0(t) \exp\{\mathbf{x}'_j\beta\}} = \exp\{\mathbf{x}'_i\beta - \mathbf{x}'_j\beta\},$$

não depende do tempo.

- ▶ Ou seja, um indivíduo que no início do estudo tem taxa de morte igual a duas vezes a taxa de um segundo indivíduo, então esta razão é a **mesma para todo o período de acompanhamento**.

## ESTIMAÇÃO NO MODELO DE COX

Modelo de Cox:

$$\lambda(t|\mathbf{x}) = \lambda_0(t) \exp\{\mathbf{x}'\boldsymbol{\beta}\}.$$

- ▶ O modelo de regressão de Cox é caracterizado pelos coeficientes  $\beta$ 's, que medem os efeitos das covariáveis sobre a função de taxa de falha.
- ▶ Queremos fazer inferência nos coeficientes  $\beta$ 's a partir das observações amostrais.
- ▶ A presença do componente não-paramétrico  $\lambda_0(t)$  na função de verossimilhança, traz dificuldades no processo inferencial.

## ESTIMAÇÃO NO MODELO DE COX

Considere uma amostra  $(t_1, \delta_1, \mathbf{x}_1), \dots, (t_n, \delta_n, \mathbf{x}_n)$ , a função de verossimilhança é dada por:

$$\begin{aligned} L(\beta) &= \prod_{i=1}^n [f(t_i | \mathbf{x}_i)]^{\delta_i} [S(t_i | \mathbf{x}_i)]^{1-\delta_i} \\ &= \prod_{i=1}^n [\lambda(t_i | \mathbf{x}_i)]^{\delta_i} S(t_i | \mathbf{x}_i) \end{aligned}$$

- ▶ Como no modelo de Cox,

$$S(t_i | \mathbf{x}_i) = \exp \left\{ - \int_0^{t_i} \lambda_0(u) \exp\{\mathbf{x}'_i \beta\} du \right\} = [S_0(t_i)]^{\exp\{\mathbf{x}'_i \beta\}}.$$

- ▶ Segue, aplicando este resultado, que:

$$L(\beta) = \prod_{i=1}^n [\lambda_0(t_i) \exp\{\mathbf{x}'_i \beta\}]^{\delta_i} [S_0(t_i)]^{\exp\{\mathbf{x}'_i \beta\}}$$

que é função do componente não-paramétrico  $\lambda_0(t)$ .

## FUNÇÃO DE VEROSSIMILHANÇA PARCIAL

- ▶ Cox (1975) propôs então uma solução alternativa.
- ▶ A ideia do Cox (1975) foi condicionar a construção da função de verossimilhança ao conhecimento da história passada de falhas e censuras.
- ▶ Desta forma, o componente não-paramétrico é eliminado da expressão da função de verossimilhança.
- ▶ Esta função foi denominada de **verossimilhança parcial**.



## FUNÇÃO DE VEROSSIMILHANÇA PARCIAL

- ▶ Considere uma amostra de  $n$  indivíduos existam  $k \leq n$  falhas distintas nos tempos  $t_1 < t_2 \dots < t_k$ .
- ▶ A probabilidade condicional da  $i$ -ésima observação vir a falhar no tempo  $t_i$  conhecendo quais observações estão sob risco em  $t_i$  é dada por:

$$\frac{\lambda_i(t | \mathbf{x}_i)}{\sum_{j \in R(t_i)} \lambda_j(t | \mathbf{x}_j)} = \frac{\lambda_0(t) \exp\{\mathbf{x}'_i \beta\}}{\sum_{j \in R(t_i)} \lambda_0(t) \exp\{\mathbf{x}'_j \beta\}} = \frac{\exp\{\mathbf{x}'_i \beta\}}{\sum_{j \in R(t_i)} \exp\{\mathbf{x}'_j \beta\}}, \quad (*)$$

em que  $R(t_i)$  é o conjunto dos índices das observações sob risco no tempo  $t_i$ .

- ▶ Observe que condicional à história de falhas e censuras até o tempo  $t_i$ , o componente não-paramétrico  $\lambda_0(t)$  desaparece.

## FUNÇÃO DE VEROSSIMILHANÇA PARCIAL

- ▶ A função de verossimilhança é formada pelo produto de todos os termos representados por (\*) associados aos tempos distintos de falha, de modo que:

$$L(\beta) = \prod_{i=1}^k \frac{\exp\{\mathbf{x}'_i\beta\}}{\sum_{j \in R(t_i)} \exp\{\mathbf{x}'_j\beta\}} = \prod_{i=1}^n \left( \frac{\exp\{\mathbf{x}'_i\beta\}}{\sum_{j \in R(t_i)} \exp\{\mathbf{x}'_j\beta\}} \right)^{\delta_i}.$$

- ▶ O EMVP (Estimador de Máxima Verossimilhança Parcial) é o valor de  $\beta$  que maximiza  $L(\beta)$ .
- ▶ EMVP é obtido resolvendo-se o sistema de equações definido por  $U(\beta) = 0$ , em que  $U(\beta)$  é o vetor escore de primeiras derivadas da função  $l(\beta) = \log(L(\beta))$ . Isto é,

$$U(\hat{\beta}) = \sum_{i=1}^n \delta_i \left[ x_i - \frac{\sum_{j \in R(t_i)} x_j \exp\{\mathbf{x}'_j\hat{\beta}\}}{\sum_{j \in R(t_i)} \exp\{\mathbf{x}'_j\hat{\beta}\}} \right] = 0.$$

## FUNÇÃO DE VEROSSIMILHANÇA PARCIAL COM EMPATES

- ▶ A função de verossimilhança parcial,  $L(\beta)$  estabelecida anteriormente não pressupõe a possibilidade de empates nos tempos observados de falha.
- ▶ Empates podem, contudo, ocorrer nos tempos de falhas devido a medições imprecisas.
- ▶ Aproximações para  $L(\beta)$ , quando ocorrem empates, foram propostas, dentre outros, por Breslow (1972), Peto (1972), Efron (1977). A aproximação de Breslow considera:

$$L(\beta) = \prod_{i=1}^k \frac{\exp(\mathbf{s}'_i \beta)}{\left[ \sum_{j \in R(t_i)} \exp(\mathbf{x}'_j \beta) \right]^{d_i}},$$

em que  $\mathbf{s}_i$  é a soma das  $p$  covariáveis dos indivíduos que falham no mesmo tempo  $t_i$  ( $i = 1, \dots, k$ ) e  $d_i$  é o nº de falhas neste mesmo tempo.

## FUNÇÃO DE VEROSSIMILHANÇA PARCIAL COM EMPATES

Observações:

- Caso haja empate entre falha e censura, assumimos que o tempo de censura é maior que o de falha.
- A função `coxph` do pacote `survival` do R assume de Efron (1977) como default.

## Propriedades assintóticas do EMVP

- ▶ Cox(1975), Tsiatis(1981), Andersen e Gill(1982) estudaram as propriedades assintóticas dos estimadores de máxima verossimilhança parcial. Sob certas condições de regularidade:
  - são consistentes e
  - assintoticamente normais.
- ▶ Para fazer inferências no modelo de Cox é possível, então, usar as estatísticas de Wald, da razão de verossimilhança e Escore.
  - i) **Teste de Wald:** é geralmente o mais usado para testar hipóteses relativas a um único parâmetro, isto é,

$$H_0: \beta_j = \beta_{0j}, \quad j = 1, \dots, p$$

## Propriedades assintóticas do EMVP

$$W = \frac{(\hat{\beta}_j - \beta_{0j})^2}{\widehat{Var}(\hat{\beta}_j)} \sim \chi_1^2$$

$$Var(\hat{\beta}) = I^{-1}(\beta)$$

em que  $I(\beta)$  é a matriz de informação observada.

- ii) **Teste da Razão de Verossimilhança:** envolve a comparação dos valores do logaritmo da função de verossimilhança maximizada sem restrição e sob  $H_0$ .

$$TRV = -2 \log \left[ \frac{L(\hat{\beta}_0)}{L(\hat{\beta})} \right] \sim \chi_q^2$$

$q$ : número de restrições impostas por  $H_0$ . Valores de  $TRV > \chi_{q,1-\alpha}^2$  indicam a rejeição de  $H_0$ .

## Propriedades assintóticas do EMVP

- iii) **Teste Escore:** obtido a partir da função escore em que a estatística de teste é dada por:

$$S = U'(\hat{\beta}_0)[I(\hat{\beta}_0)]^{-1}U(\hat{\beta}_0),$$

com  $U(\hat{\beta}_0)$  a função escore e  $I(\hat{\beta}_0)$  a matriz de informação observada de  $\hat{\beta}$  avaliada em  $\hat{\beta}_0$ .

Para amostras grandes,  $H_0$  é rejeitada, a um nível  $100\alpha\%$  de significância, para valores de  $S > \chi_{q,1-\alpha}^2$ .

## INTERPRETAÇÃO DOS PARÂMETROS

- ▶ No modelo de Cox, o efeito das covariáveis é de acelerar ou desacelerar a função de taxa de falha. A propriedade de taxas de falhas proporcionais do modelo é utilizada para interpretar os coeficientes estimados.
- ▶ Tomando a razão das taxas de falha de dois indivíduos  $i$  e  $j$  que têm os mesmos valores para as covariáveis com exceção da  $l$ -ésima, tem-se

$$\frac{\lambda_i(t)}{\lambda_j(t)} = \exp \left\{ \beta_l (x_{il} - x_{jl}) \right\},$$

que é interpretado como a razão de taxas de falhas.



## INTERPRETAÇÃO DOS PARÂMETROS

- ▶ Por exemplo, suponha que  $x_l$  seja uma covariável dicotômica indicando pacientes hipertensos. A taxa de morte entre os hipertensos é  $\exp(\beta_l)$  vezes a taxa daqueles com pressão normal, mantida fixas as outras covariáveis.
- ▶ Uma interpretação similar é obtida para covariáveis contínuas. Se, por ex., o efeito de idade é significativo e  $e^{\hat{\beta}} = 1,05$  para este termo, tem-se com o aumento de 1 ano na idade, que a taxa de morte fica aumentada em 5%.
- ▶ Estimativa para  $\exp(\beta_l)$  é obtida utilizando a propriedade de invariância do estimador de máxima verossimilhança parcial. O intervalo de 95% de confiança é dado por:  $\exp(\hat{\beta} \pm 1,96 \times \widehat{EP}(\hat{\beta}))$ .

## FUNÇÕES RELACIONADAS A $\lambda_0(t)$

- ▶ As funções relacionadas a  $\lambda_0(t)$  referem-se basicamente a:

$$\Lambda_0(t) = \int_0^t \lambda_0(u) du \quad \text{e} \quad S_0(t) = \exp \{ - \Lambda_0(t) \}.$$

- ▶ Estas funções são importantes para avaliar a adequação do modelo e para fazer predições.
- ▶ Como  $\lambda_0(t)$  não é especificado parametricamente, os estimadores para estas quantidades são de natureza não-paramétrica.
- ▶ Um estimador simples proposto para a função de taxa de falha basal acumulada  $\Lambda_0(t)$ , referenciado como estimador de Nelson-Aalen-Breslow. Ele é uma função escada, com saltos nos distintos tempos de falha, e é expresso por:

## FUNÇÕES RELACIONADAS A $\lambda_0(t)$

$$\hat{\Lambda}_0(t|\mathbf{x}) = \sum_{j: t_j < t} \frac{d_j}{\sum_{l \in R_j} \exp\{\mathbf{x}'_l \hat{\beta}\}},$$

em que  $d_j$  é o número de falhas em  $t_j$ .

- ▶ Conseqüentemente, as funções de sobrevivência  $S_0(t|\mathbf{x})$  e  $S(t|\mathbf{x})$  podem ser estimadas por, respectivamente,

$$\hat{S}_0(t|\mathbf{x}) = \exp\{-\hat{\Lambda}_0(t)\} \quad \text{e} \quad \hat{S}(t|\mathbf{x}) = [\hat{S}_0(t)]^{\exp\{\mathbf{x}'\hat{\beta}\}}.$$

- ▶ Note que na ausência de covariáveis tem-se:

$$\hat{\Lambda}_0(t) = \sum_{j: t_j < t} \left( \frac{d_j}{n_j} \right),$$

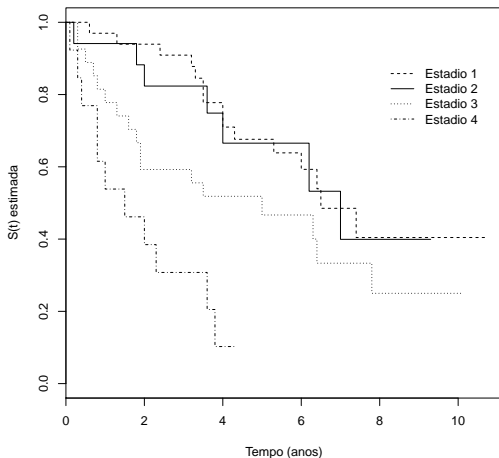
que é o estimador de Nelson-Aalen.

## EXEMPLO 1: Estudo sobre câncer de laringe

- ▶ Neste exemplo os dados considerados referem-se a um estudo, descrito em Klein e Moeschberger (1997).
- ▶ 90 pacientes do sexo masculino na Dinamarca foram diagnosticados no período de 1970-1978 com câncer de laringe e foram acompanhados até 01/01/1983.
- ▶ Para cada paciente foram registrados, no diagnóstico, a idade (em anos) e o estadio da doença (I = tumor primário, II = envolvimento de nódulos, III = metástases e IV = combinações dos 3 estadios anteriores) bem como a data do óbito ou do último acompanhamento (censura). Os tempos foram registrados em anos. Foram registrados 50 eventos (óbitos).
- ▶ Os estadios encontram-se ordenados pelo grau de seriedade da doença (menos sério para mais sério).

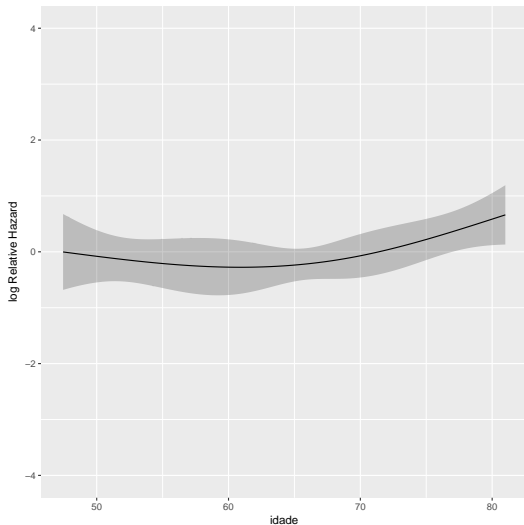
## EXEMPLO 1: Estudo sobre câncer de laringe

Kaplan-Meier para a covariável Estadio.



## EXEMPLO 1: Estudo sobre câncer de laringe

Avaliando a forma funcional de idade



## EXEMPLO 1: Estudo sobre câncer de laringe

- Utilizando-se o modelo de Cox para a análise desses dados, foram ajustados diversos modelos sequências.

Modelo	Covariáveis no modelo	Estimativas	Log verossimilhança parcial
1	nenhuma	-	$l_1 = -197,2129$
2	$X_1$ : estadio (II) (III) (IV)	$\hat{\beta}_1 = 0,0658$ $\hat{\beta}_2 = 0,6121$ $\hat{\beta}_3 = 1,7228$	$l_2 = -189,0812$
3	$X_1$ : estadio (II) (III) (IV) $X_2$ : idade	$\hat{\beta}_1 = 0,1386$ $\hat{\beta}_2 = 0,6383$ $\hat{\beta}_3 = 1,6931$ $\hat{\beta}_4 = 0,0189$	$l_3 = -188,1794$
4	$X_1$ : estadio (II) (III) (IV) $X_2$ : idade $X_1 * X_2$ (II* id) (III* id) (IV* id)	$\hat{\beta}_1 = -7,9461$ $\hat{\beta}_2 = -0,1225$ $\hat{\beta}_3 = 0,8470$ $\hat{\beta}_4 = -0,0026$ $\hat{\beta}_5 = 0,1203$ $\hat{\beta}_6 = 0,0114$ $\hat{\beta}_7 = 0,0137$	$l_4 = -185,0775$

## EXEMPLO 1: Estudo sobre câncer de laringe

- ▶ Teste da razão de verossimilhanças parciais para o termo de interação entre estadio e idade:

$$TRV = 2 \times (188,18 - 185,08) = 6,20$$

$$p = 0,10, \text{ g.l.} = 3.$$

- ▶ No entanto, os resultados mostram evidências de que pelo menos um dos  $\beta$ 's associados à referida interação difere significativamente de zero, no caso  $\beta_5$  com p-valor = 0,022.

parâmetro	estimativa	erro padrão	z	p-valor
$\beta_5 = \text{idade:est2}$	0,1203	0,0523	2,2990	<b>0,022</b>
$\beta_6 = \text{idade:est3}$	0,0114	0,0374	0,3031	0,760
$\beta_7 = \text{idade:est4}$	0,0137	0,0360	0,3802	0,700



## ADEQUAÇÃO DO MODELO DE COX

- ▶ O modelo de Cox não se ajusta a qualquer situação clínica e, como qualquer outro modelo estatístico, requer o uso de técnicas para avaliar a sua adequação.
- ▶ A violação da suposição básica, que é a de proporcionalidade das taxas de falhas, pode acarretar em sérios vícios na estimação dos coeficientes do modelo (Struthers e Kalbfleisch, 1986).
- ▶ As técnicas de avaliação do modelo são baseadas em resíduos, como em outros modelos. O resíduo de Schoenfeld é o mais utilizado atualmente. Os resíduos podem ser utilizados via teste de hipóteses assim como em técnicas gráficas.

## Resíduos para o Modelo de Cox

- ▶ Definir resíduo para o modelo de Cox não foi tarefa simples.
- ▶ O modelo de Cox foi proposto para  $\lambda(t|x)$  e não para a resposta  $T$ , como em outros modelos.
- ▶ Cox e Snell (1968) apresentam uma definição geral de resíduos.
- ▶ No entanto, os resíduos de Schoenfeld (1982) são atualmente os mais utilizados para verificar a adequação do modelo de Cox, em especial, a suposição de proporcionalidade das taxas de falhas.

## Avaliação da Proporcionalidade das Taxas

Existem disponíveis na literatura, técnicas gráficas e testes estatísticos para avaliar esta suposição.

Técnica Gráfica:

- dividir os dados em  $m$  estratos, usualmente de acordo com alguma covariável;
- em seguida, estima-se  $\hat{\Lambda}_{0j}(t)$  para o  $j$ -ésimo estrato usando o estimador de Breslow;
- analisa-se as curvas do logaritmo de  $\hat{\Lambda}_{0j}(t) \times t$ , ou  $\log(t)$ . Curvas não paralelas significam desvios da suposição de taxas de falhas proporcionais. Situações extremas de violação da suposição ocorrem quando as curvas se cruzam.
- As curvas de KM nos fornecem indicações nesta direção.

## Avaliação da Proporcionalidade das Taxas

### ■ Resíduos de Schoenfeld (1982).

- Para o  $i$ -ésimo indivíduo, correspondente a um evento, com covariáveis  $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})'$ , o vetor de resíduos de Schoenfeld  $\mathbf{r}_i = (r_{i1}, \dots, r_{ip})$  é definido para cada componente  $r_{iq}$ ,  $q = 1, \dots, p$ , por:

$$r_{iq} = x_{iq} - \frac{\sum_{j \in R(t_i)} x_{jq} \exp\{\mathbf{x}'_j \hat{\boldsymbol{\beta}}\}}{\sum_{j \in R(t_i)} \exp\{\mathbf{x}'_j \hat{\boldsymbol{\beta}}\}}.$$

- Os resíduos padronizados de Schoenfeld são dados por:

$$\mathbf{s}_i^* = [I(\hat{\boldsymbol{\beta}})]^{-1} \times \mathbf{r}_i$$

em que  $I(\hat{\boldsymbol{\beta}})$  é a matriz de informação observada.

## Avaliação da Proporcionalidade das Taxas - Gráfico

- ▶ Grambsch e Therneau (1994) sugerem a utilização de  $\mathbf{s}_i^*$  para avaliar a suposição de TFP. Considere o modelo de Cox dinâmico:

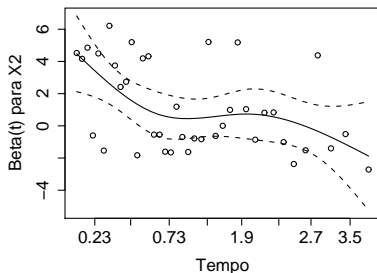
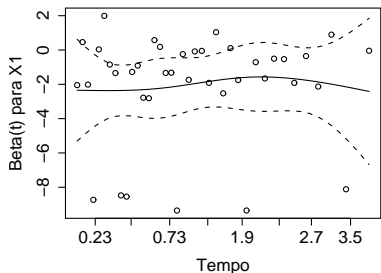
$$\lambda(t) = \lambda_0(t) \exp \{ \mathbf{x}'\beta(t) \},$$

a restrição  $\beta(\mathbf{t}) = \beta$  corresponde a proporcionalidade das taxas.

- ⇒ Se a suposição de TFP é válida, o gráfico de  $\beta_q(t) \times t$  deve ser uma linha horizontal.
- ⇒ Sugestão: usar o gráfico de  $(\mathbf{s}_{iq}^* + \hat{\beta}_q) \times t$  ou alguma função do tempo,  $g(t)$ , ( $q = 1, \dots, p$ ).
- ⇒ Inclinação zero indica evidências a favor da proporcionalidade das taxas.

## Avaliação da Proporcionalidade das Taxas - Gráfico

- ▶ Uma curva suavizada, com bandas de confiança, é usualmente adicionada a este gráfico para facilitar sua interpretação.
- ▶ As figuras abaixo ilustram estes gráficos (primeira, adequada e a segunda, inadequada).



## Avaliação da Proporcionalidade das Taxas - Testes

- Testes de hipóteses associado aos resíduos de Schoenfeld.
- ▶ O coeficiente de correlação de Pearson ( $\rho$ ) entre os resíduos padronizados de Schoenfeld e  $g(t)$  para cada covariável é uma dessas medidas.
- ▶ Valores de  $\rho$  próximos de zero mostram evidências a favor da suposição de TFP.
- ▶ Um hipótese global de proporcionalidade das taxas sobre todas as covariáveis no modelo de Cox, é obtido a partir da forma quadrática:

$$T = \frac{(g - \bar{g})' S^* I S^{*'} (g - \bar{g})}{k \sum_i (g_i - \bar{g})^2} \sim \chi_{(p)}^2$$

sob a hipótese de proporcionalidade, comparamos com uma  $\chi_{(p)}^2$ .

## Avaliação da Proporcionalidade das Taxas - Testes

- ▶ Para testar a hipótese de taxas proporcionais associado a  $q$ -ésima covariável ( $q = 1, \dots, p$ ) utiliza-se a estatística de teste:

$$T_q = \frac{k \left( \sum_i (g_i - \bar{g}) s_{qi}^* \right)^2}{l_q^{-1} \sum_i (g_i - \bar{g})^2},$$

em que  $l_q^{-1}$  é o  $q$ -ésimo elemento da diagonal de  $l^{-1}$ . Sob a hipótese nula, de taxas proporcionais para a  $j$ -ésima covariável,  $T_{qs}$  tem aproximadamente distribuição qui-quadrado com 1 g.l.

- ▶ Desta forma podemos seguir os seguintes passos: (1) fazer inicialmente o teste global; (2) caso este for rejeitado, fazemos os individuais para identificar a(s) covariável (is) que estão violando a suposição de proporcionalidade.



## Continuação Exemplo 1: Estudo sobre câncer de laringe

- ▶ Neste exemplo os dados considerados referem-se a um estudo, descrito em Klein e Moeschberger (1997).
- ▶ 90 pacientes do sexo masculino na Dinamarca foram diagnosticados no período de 1970-1978 com câncer de laringe e foram acompanhados até 01/01/1983.
- ▶ Para cada paciente foram registrados, no diagnóstico, a idade (em anos) e o estadio da doença (I = tumor primário, II = envolvimento de nódulos, III = metástases e IV = combinações dos 3 estadios anteriores) bem como a data do óbito ou do último acompanhamento (censura). Os tempos foram registrados em anos.
- ▶ Os estadios encontram-se ordenados pelo grau de seriedade da doença (menos sério para mais sério).

## Continuação Exemplo 1: Estudo sobre câncer de laringe

- Utilizando-se o modelo de Cox para a análise desses dados, foram ajustados diversos modelos sequências.

Modelo	Covariáveis no modelo	Estimativas	Log verossimilhança parcial
1	nenhuma	-	$l_1 = -197,2129$
2	$X_1$ : estadio (II) (III) (IV)	$\hat{\beta}_1 = 0,0658$ $\hat{\beta}_2 = 0,6121$ $\hat{\beta}_3 = 1,7228$	$l_2 = -189,0812$
3	$X_1$ : estadio (II) (III) (IV) $X_2$ : idade	$\hat{\beta}_1 = 0,1386$ $\hat{\beta}_2 = 0,6383$ $\hat{\beta}_3 = 1,6931$ $\hat{\beta}_4 = 0,0189$	$l_3 = -188,1794$
4	$X_1$ : estadio (II) (III) (IV) $X_2$ : idade $X_1 * X_2$ (II* id) (III* id) (IV* id)	$\hat{\beta}_1 = -7,9461$ $\hat{\beta}_2 = -0,1225$ $\hat{\beta}_3 = 0,8470$ $\hat{\beta}_4 = -0,0026$ $\hat{\beta}_5 = 0,1203$ $\hat{\beta}_6 = 0,0114$ $\hat{\beta}_7 = 0,0137$	$l_4 = -185,0775$

## Continuação Exemplo 1: Estudo sobre câncer de laringe

- ▶ Teste da razão de verossimilhanças parciais para o termo de interação entre estadio e idade:

$$TRV = 2 \times (188,18 - 185,08) = 6,20$$

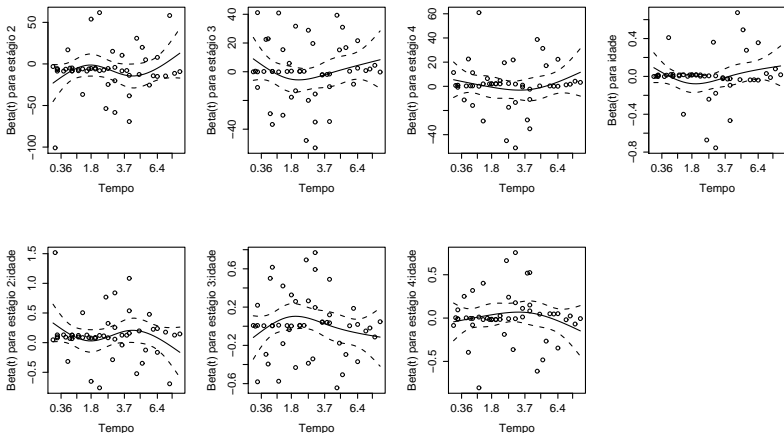
$$p = 0,10, \text{ g.l.} = 3.$$

- ▶ No entanto, os resultados mostram evidências de que pelo menos um dos  $\beta$ 's associados à referida interação difere significativamente de zero, no caso  $\beta_5$  com p-valor = 0,022.

parâmetro	estimativa	erro padrão	z	p-valor
$\beta_5 = \text{idade:est2}$	0,1203	0,0523	2,2990	<b>0,022</b>
$\beta_6 = \text{idade:est3}$	0,0114	0,0374	0,3031	0,760
$\beta_7 = \text{idade:est4}$	0,0137	0,0360	0,3802	0,700

## EXEMPLO 1: Estudo sobre câncer de laringe

- ▶ Vamos verificar a adequação dos modelos ajustados utilizando os resíduos de Schoenfeld.



## EXEMPLO 1: Estudo sobre câncer de laringe

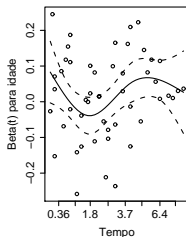
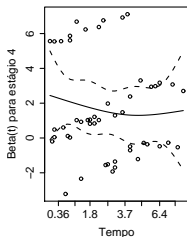
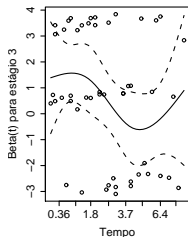
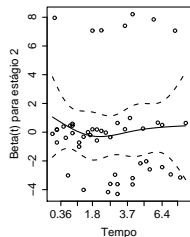
Covariável	rho ( $\rho$ )	Chisq	valor- $p$
estadio II	0,0958	0,5033	0,478
estadio III	0,0462	0,1577	0,691
estadio IV	0,0269	0,0421	0,837
idade	0,1082	0,9376	0,333
estadio II * idade	-0,0943	0,4929	0,483
estadio III * idade	-0,0768	0,4364	0,509
estadio IV * idade	-0,0443	0,1160	0,733
GLOBAL	–	5,7988	0,563

Conclusão: não existem evidências contra a violação da suposição de proporcionalidade das taxas de falha.

## EXEMPLO 1: Estudo sobre câncer de laringe

- ▶ Para o modelo de Cox sem a interação, foram obtidos os resultados:

Covariável	rho ( $\rho$ )	chisq	valor- $p$
estadio II	-0,0107	0,00605	0,938
estadio III	-0,2440	2,83791	0,092
estadio IV	-0,1188	0,62202	0,430
idade	0,1328	1,16886	0,280
GLOBAL	—	4,56330	0,335



## EXEMPLO 1: Estudo sobre câncer de laringe

- ▶ Comparando-se os resultados de ambos os diagnósticos, decidiu-se pelo modelo com a presença da interação. As estimativas obtidas para este modelo foram:

Covariável	Estimativa	Erro-Padrão	Valor-p	RR
Estadio II	-7,946	3,678	0,031	0,0003
Estadio III	-0,123	2,468	0,960	0,8847
Estadio IV	0,847	2,426	0,730	2,3326
Idade	-0,003	0,026	0,920	0,9974
Estadio II*idade	0,120	0,052	0,022	1,1278
Estadio III*idade	0,011	0,037	0,760	1,0114
Estadio IV*idade	0,014	0,036	0,700	1,0138

- ▶ A decisão pelo modelo com interação é questionável. O termo da interação não é significativo a um nível de 0,10 e podemos aceitar a suposição de RTF proporcionais para o modelo sem interação.

$$\widehat{S}(t | \mathbf{x}) = \begin{cases} \left[ \widehat{S}_0(t) \right]^{\exp\{\hat{\beta}_4 x_2\}} & \text{se estadio I} \\ \left[ \widehat{S}_0(t) \right]^{\exp\{\hat{\beta}_1 + (\hat{\beta}_4 + \hat{\beta}_5) x_2\}} & \text{se estadio II} \\ \left[ \widehat{S}_0(t) \right]^{\exp\{\hat{\beta}_2 + (\hat{\beta}_4 + \hat{\beta}_6) x_2\}} & \text{se estadio III} \\ \left[ \widehat{S}_0(t) \right]^{\exp\{\hat{\beta}_3 + (\hat{\beta}_4 + \hat{\beta}_7) x_2\}} & \text{se estadio IV} \end{cases}$$

e

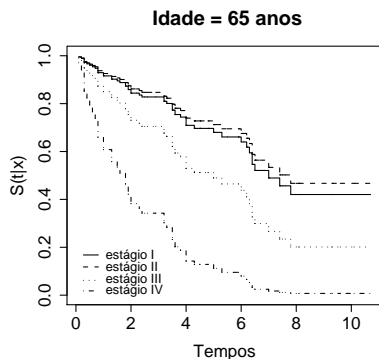
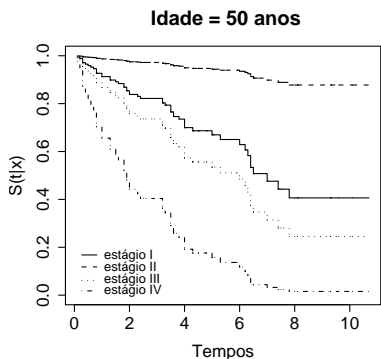
$$\widehat{\lambda}(t | \mathbf{x}) = \begin{cases} \widehat{\lambda}_0(t) \exp\{\hat{\beta}_4 x_2\} & \text{se estadio I} \\ \widehat{\lambda}_0(t) \exp\{\hat{\beta}_1 + (\hat{\beta}_4 + \hat{\beta}_5) x_2\} & \text{se estadio II} \\ \widehat{\lambda}_0(t) \exp\{\hat{\beta}_2 + (\hat{\beta}_4 + \hat{\beta}_6) x_2\} & \text{se estadio III} \\ \widehat{\lambda}_0(t) \exp\{\hat{\beta}_3 + (\hat{\beta}_4 + \hat{\beta}_7) x_2\} & \text{se estadio IV} \end{cases}$$

em que  $x_2$  é a idade.



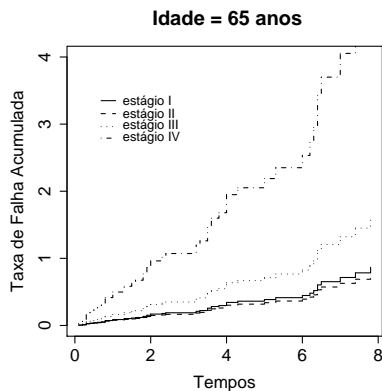
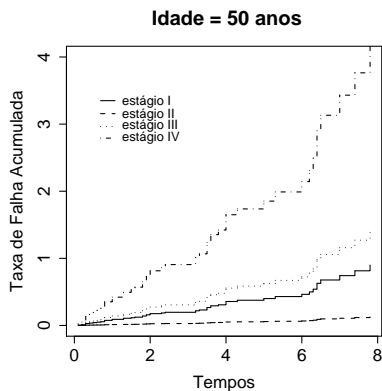
## Funções de sobrevivência e risco estimadas

- ▶ Curvas de sobrevivência estimadas para pacientes com idades de 50 e 65 anos, em cada um dos 4 estádios da doença.



## Funções de sobrevivência e risco estimadas

- ▶ Curvas de taxas de falhas acumuladas estimadas para pacientes com idades de 50 e 65 anos, em cada um dos 4 estádios da doença.



## Funções de sobrevivência e risco estimadas

- ▶ Assim, por exemplo, para os pacientes  $i$  e  $l$  em que ambos encontram-se no estadio II da doença, mas um deles apresenta idade de 65 anos e outro de 50 anos, tem-se que a razão de taxas de falha é:

$$\begin{aligned}\frac{\widehat{\lambda}(t | \mathbf{x}_i)}{\widehat{\lambda}(t | \mathbf{x}_l)} &= \frac{\exp \left\{ \widehat{\beta}_1 + (\widehat{\beta}_4 + \widehat{\beta}_5) * 65 \right\}}{\exp \left\{ \widehat{\beta}_1 + (\widehat{\beta}_4 + \widehat{\beta}_5) * 50 \right\}} \\ &= \exp \left\{ (\widehat{\beta}_4 + \widehat{\beta}_5) * (65 - 50) \right\} = 5,84,\end{aligned}$$

Isto significa que a taxa de óbito de pacientes com 65 anos de idade e no estadio II da doença, é de aproximadamente 6 vezes a taxa de óbito de pacientes com 50 anos e no mesmo estadio da doença.

## Funções de sobrevivência e risco estimadas

- ▶ Por outro lado, tem-se, por exemplo, para os pacientes  $j$  e  $k$  em que ambos têm 50 anos de idade mas um deles encontra-se no estadio IV da doença e outro no estadio III, que a razão de taxas de óbito é de:

$$\frac{\hat{\lambda}(t | \mathbf{x}_j)}{\hat{\lambda}(t | \mathbf{x}_k)} = \frac{\exp \left\{ \hat{\beta}_3 + (\hat{\beta}_4 + \hat{\beta}_7) * 50 \right\}}{\exp \left\{ \hat{\beta}_2 + (\hat{\beta}_4 + \hat{\beta}_6) * 50 \right\}} = 2,96.$$

Desse modo, tem-se que a taxa de óbito de pacientes com 50 anos de idade e no estadio IV da doença é de 3 vezes a taxa de óbito de pacientes também com 50 anos de idade mas que encontram-se no estadio III da doença.

- ▶ Razões de taxas de óbito para todas as demais comparações de interesse podem ser obtidas de forma análoga.

## Avaliação de outros Aspectos do Modelo de Cox

- ▶ forma funcional de covariável contínua;
- ▶ presença de potenciais indivíduos atípicos;

## Avaliação da Qualidade Geral de Ajuste do Modelo

- ▶ Resíduos de Cox-Snell (1968) definidos por:

$$\hat{e}_i = \hat{\Lambda}_0(t_i) \exp \left\{ \sum_{j=1}^p x_{ij} \hat{\beta}_j \right\}, \quad i = 1, \dots, n.$$

- ▶ Se o modelo for adequado, os  $\hat{e}_i$ 's devem ser tratados como uma amostra censurada de uma distribuição exponencial padrão e, então. O gráfico de, por exemplo,  $\hat{\Lambda}(\hat{e}_i)$  versus  $\hat{e}_i$  deve ser aproximadamente uma reta.
- ▶ A análise gráfica desses resíduos não fornece, contudo, informações sobre o tipo de problema que estaria ocorrendo caso o ajuste não se apresentar satisfatório.

## Avaliação da Qualidade Geral de Ajuste do Modelo

- ▶ Klein e Moeschberger (1997) observam que os resíduos de Cox-Snell deveriam ser usados com cuidado pois a distribuição exponencial dos mesmos mantém-se somente quando os verdadeiros valores dos parâmetros são usados em  $\hat{e}_i$ .
- ▶ Quando as estimativas dessas quantidades são usadas para o cálculo dos resíduos, como é feito aqui, falhas quanto à distribuição exponencial podem ocorrer devido, parcialmente, à incerteza no processo de estimação dos parâmetros  $\beta$ . Essa incerteza é maior na cauda direita da distribuição e para amostras pequenas.

### (a) Pontos Atípicos e Forma Funcional das Covariáveis Os

resíduos martingal são freqüentemente usados para tais finalidades. Quando os dados apresentam censuras à direita e todas as covariáveis são fixadas no início do estudo, estes são definidos, para  $i = 1, \dots, n$ , por:

$$\hat{m}_i = \delta_i - \hat{\Lambda}_0(t_i) \exp \left\{ \sum_{j=1}^p x_{ij} \hat{\beta}_j \right\} = \delta_i - \hat{e}_i.$$



## Avaliação de outros Aspectos do Modelo de Cox

- ▶ Para, por exemplo, a covariável contínua  $x_q$ , o gráfico de  $\hat{m}_i$  versus  $x_{iq}$  é utilizado para que se possa avaliar a forma funcional desta covariável. Na prática, as interpretações desses gráficos não são simples em razão da distribuição assimétrica desses resíduos.
- ▶ Outro resíduo usado, em geral, com o propósito de detectar pontos atípicos é o resíduo deviance, definido no modelo de Cox por:

$$\hat{d}_i = \text{sinal}(\hat{m}_i) \left[ -2 \left( \hat{m}_i + \delta_i \log(\delta_i - \hat{m}_i) \right) \right]^{1/2}.$$

Tais resíduos não são tão assimétricos como os resíduos martingal. O gráfico de  $\hat{d}_i$  versus o preditor linear  $\sum_{j=1}^p \mathbf{x}_{ij} \beta_j$ ,  $i = 1, \dots, n$ , é utilizado, nesse caso, para avaliar a presença de dados atípicos.

## Pontos Influentes

- ▶ Uma medida global de efeito das observações pode ser obtida por:

$$\Delta\beta_i = (\hat{\beta} - \hat{\beta}_{(i)})' [I(\hat{\beta})]^{-1} (\hat{\beta} - \hat{\beta}_{(i)}), \quad i = 1, \dots, n,$$

em que  $\Delta\beta$  a mudança no vetor de coeficientes estimados obtida pela remoção da  $i$ -ésima observação.

- ▶ O gráfico desses resíduos *versus*  $i$  pode ser útil na detecção de observações influentes. Esta medida é usualmente chamada de D-Cook na literatura.

## Pontos Influentes

- ▶ Para a  $q$ -ésima covariável ( $q = 1, \dots, p$ ) no modelo de Cox, tem-se, para  $i = 1, \dots, n$ ,

$$\Delta\beta_{i,q} = (\hat{\beta}_q - \hat{\beta}_{q,(i)})' [I(\hat{\beta})]_{qq}^{-1} (\hat{\beta}_q - \hat{\beta}_{q,(i)}).$$

Obtemos uma matriz  $n \times p$  que é chamada de resíduos *dfbetas*.

- ▶ Gráficos dos respectivos resíduos associados a cada covariável *versus* os valores da respectiva covariável, são usados para a identificação de pontos influentes.

## 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

- 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 $\geq 2,5$ kg e 1 se $< 2,5$ kg
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

# MODELO DE REGRESSÃO DE COX

Nome	Modelo	-2 log L	Estatística	Valor-p
Passo 1	Nulo	560,628	—	—
	V1	556,958	3,670	0,0554
	V2	557,922	2,706	0,1000
	V3	554,920	5,708	0,0169
	V4	549,455	11,173	0,0008
	V5	559,402	1,226	0,2682
	V6	554,008	6,620	0,0101
	V7	558,420	2,208	0,1373
	V8	558,617	2,011	0,1562
	V9	558,597	2,031	0,1541
	V10	558,137	2,491	0,1145
Passo 2	V11	557,872	2,756	0,0969
	V1+V2+V3+V4+V6+V11	536,196	—	—
	V2+V3+V4+V6+V11	538,771	2,575	0,2358
	V1+V3+V4+V6+V11	536,196	0,000	1,0000
	V1+V2+V4+V6+V11	541,104	4,908	0,0267
	V1+V2+V3+V6+V11	543,629	7,433	0,0064
	V1+V2+V3+V4+V11	540,242	4,046	0,0443
Passo 3	V1+V2+V3+V4+V6	536,346	0,150	0,6985
	V3+V4+V6	539,433	—	—
	V3+V4+V6+V1	536,347	3,086	0,0790
	V3+V4+V6+V2	538,823	0,610	0,4348
	V3+V4+V6+V11	539,359	0,074	0,7856

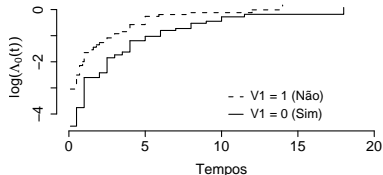
# MODELO DE REGRESSÃO DE COX

Passo 4	V3+V4+V6+V1	536,347	—	—
	V3+V4+V6+V1+V5	536,076	0,271	0,6027
	V3+V4+V6+V1+V7	534,108	2,239	0,1346
	V3+V4+V6+V1+V8	533,257	3,090	0,0788
	V3+V4+V6+V1+V9	535,012	1,335	0,2479
	V3+V4+V6+V1+V10	536,268	0,079	0,7787
Passo 5	V1+V3+V4+V6+V8	533,257	—	—
	V3+V4+V6+V8	534,492	1,235	0,2497
	V1+V4+V6+V8	538,540	5,283	0,0215
	V1+V3+V6+V8	542,136	8,879	0,0029
	V1+V3+V4+V8	538,172	4,915	0,0266
	V1+V3+V4+V6	536,347	3,090	0,0788
Passo 6	V1+V3+V4+V6	536,347	—	—
	V1+V3+V4+V6+V1*V3	535,922	0,425	0,5145
	V1+V3+V4+V6+V1*V4	536,123	0,224	0,6360
	V1+V3+V4+V6+V1*V6	536,005	0,342	0,5587
	V1+V3+V4+V6+V3*V4	535,136	1,211	0,2711
	V1+V3+V4+V6+V3*V6	534,673	1,674	0,1957
	V1+V3+V4+V6+V4*V6	535,873	0,474	0,4912
Modelo Final	V1+V3+V4+V6	536,347		

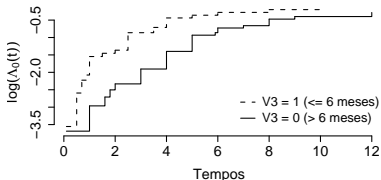
## VERIFICAÇÃO DA ADEQUAÇÃO DO MODELO

Curvas de  $\log \hat{H}(t)$  versus  $t$  para a Covariável Conceito Materno sobre o Tempo Ideal de Amamentação em dois Níveis (...  $\leq 6$  meses, —  $> 6$  meses).

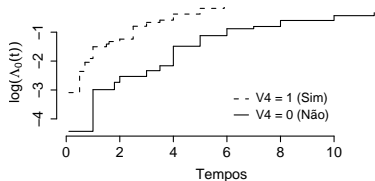
**V1: Experiência Amamentação**



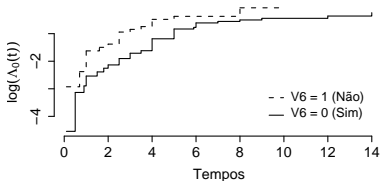
**V3: Conceito Amamentação**



**V4: Dificuldades Amamentar**



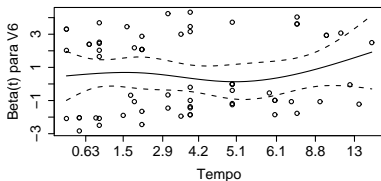
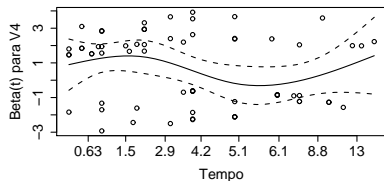
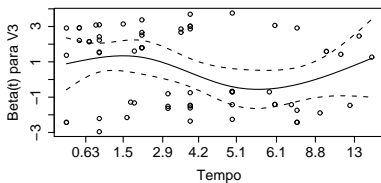
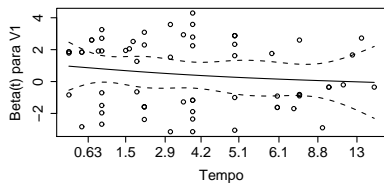
**V6: Leite Materno na Maternidade**





## Resíduos de Schoenfeld

Situações similares foram observadas para as outras curvas das covariáveis.



## Resíduos de Schoenfeld

Covariável	$\rho$	chisq	p
Exper. Amam. (V1)	-0.1098	0.754	0.385
Conc. Amam. (V3)	-0.1289	1.083	0.298
Dific. Amam. (V4)	-0.1047	0.653	0.419
Leite Excl. (V6)	0.0918	0.608	0.435
GLOBAL	NA	3.232	0.520

## RESULTADOS FINAIS E INTERPRETAÇÃO

Covariável	Estimativa	Valor-p	RR	IC(RR, 95%)
Exper. Amam. (V1)	0,471	0,079	1,601	(0,94; 2,71)
Conc. Amam. (V3)	0,579	0,027	1,785	(1,07; 2,99)
Dific. Amam. (V4)	0,716	0,007	2,046	(1,22; 3,43)
Leite Excl. (V6)	0,578	0,029	1,783	(1,06; 2,99)

## RESULTADOS FINAIS E INTERPRETAÇÃO

- A taxa de desmame precoce em mães que não tiveram experiência anterior de amamentação é 1,6 vezes a taxa das mães que tiveram essa experiência.
- A taxa de desmame precoce em mães que acreditam que o tempo ideal de amamentação é menor ou igual a 6 meses é aproximadamente 1,8 vezes a taxa das mães que acreditam que o tempo ideal de amamentação é superior a 6 meses.
- A taxa de desmame precoce em mães que apresentaram dificuldades de amamentar nos primeiros dias pós-parto é aproximadamente 2 vezes a taxa das mães que não apresentaram essas dificuldades.
- A taxa de desmame precoce em crianças que não receberam exclusivamente leite materno na maternidade é 1,8 vezes a taxa de desmame precoce em crianças que receberam exclusivamente o leite materno.

## Estudo de Caso: Avaliação da Prática de Aleitamento Materno

- 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

## COMPARAÇÕES DOS MODELOS

- Os dois modelos foram consistentes nos resultados;
- As estimativas para os parâmetros associados as covariáveis apontam na mesma direção. Entretanto, as interpretações são diferentes.
- No modelo de Cox um coeficiente positivo indica um aumento da taxa de falha e por consequência uma redução do tempo até a falha. A interpretação é feita em termos da razão de taxa de falhas.
- No modelo de tempo de vida acelerada um coeficiente positivo significa um aumento no tempo de vida. A interpretação é feita em termos de razão de tempos medianos de falha.
- Os modelos paramétricos, se bem ajustados, devem produzir resultados mais precisos do que os do modelo de Cox.
- O modelo de Cox utiliza o método de máxima verossimilhança parcial que exclui parte da informação da amostra pois baseia-se nos postos das observações.