

MODELO DE REGRESSÃO DE COX

- ❖ Os modelos de regressão paramétricos vistos anteriormente exigem que se suponha uma distribuição estatística para o tempo de sobrevivência.
- ❖ Contudo esta suposição, caso não seja adequada, pode fazer com que as estimativas sejam pouco confiáveis.
- ❖ Com o objetivo de encontrar um modelo mais flexível, Cox propôs em 1972 um modelo, denominado modelo de risco proporcional de Cox.
- ❖ Esse modelo passou a ser o mais utilizado na análise de dados de sobrevivência por sua versatilidade.

- Como em análise de sobrevivência o interesse também pode estar no risco de falha o modelo proposto por Cox modela diretamente a função de risco.
- O princípio básico deste modelo para estimar o efeito das covariáveis é a proporcionalidade dos riscos ao longo de todo o tempo de observação.
- Suponha o caso simples em que uma única covariável, que é um indicador de grupo, é considerada.
- Considere, por exemplo, que pacientes são aleatorizados para receber um tratamento padrão ou um novo tratamento.
- Seja $h_1(t)$ e $h_0(t)$ as funções de risco no tempo t para pacientes no tratamento novo e no tratamento padrão, respectivamente.



- De acordo com o princípio da proporcionalidade o risco no tempo t para pacientes no novo tratamento é proporcional ao risco, no mesmo tempo, para pacientes sobre o tratamento padrão.
- O modelo de riscos proporcionais pode ser expresso na forma
$$h_1(t) = \psi h_0(t)$$
- Uma implicação da suposição de riscos proporcionais é que as correspondentes funções de sobrevivência para indivíduos no novo e no tratamento padrão são razoavelmente paralelas ao longo de todo tempo.
- Um cruzamento das curvas ou uma variação nas distâncias entre as curvas de diferentes categorias podem indicar ausência de proporcionalidade.

- O valor de ψ é uma taxa de risco ou risco relativo.
- Se $\psi < 1$, o risco de falha em t é menor para um indivíduo sobre o novo tratamento, relativo ao indivíduo no tratamento padrão.
- Por outro lado, se $\psi > 1$, o risco de falha em t é maior para um indivíduo no novo tratamento, ou seja o tratamento padrão indica uma melhor alternativa.
- Considere agora um estudo com n indivíduos e denote a função de risco para o i -ésimo indivíduo por $h_i(t)$, $i = 1, 2, \dots, n$.
- Seja $h_0(t)$ a função de risco para um tratamento padrão. A função de risco para o novo tratamento é então $\psi h_0(t)$.
- Como o risco relativo, ψ , não pode ser negativo é conveniente considerar $\psi = \exp(\beta)$.

- O parâmetro β é então o logaritmo do risco relativo e qualquer valor de β definido em $(-\infty, +\infty)$ leva a um valor positivo de ψ .
- Note que valores positivos de β são obtidos quando o risco relativo é maior do que 1, que é quando o novo tratamento é inferior ao padrão.
- Seja X uma variável indicadora de grupo que assume o valor 0 para indivíduos no tratamento padrão e 1 para indivíduos no tratamento novo.
- Se x_i é o valor de X para o i -ésimo indivíduo no estudo, a função de risco para este indivíduo pode ser escrita por

$$h_i(t) = h_0(t) \exp \{ \beta x_i \}$$

- Este modelo é o modelo de risco proporcional de Cox para a comparação de dois tratamentos.
- 

- De forma genérica, considere p covariáveis, de forma que x seja um vetor da forma $x = (x_1, x_2, \dots, x_p)'$. A função de risco para o i -ésimo indivíduo é então escrita por

$$h_i(t) = h_0(t) \exp \{ \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi} \} = h_0(t) \exp \{ x' \beta \}$$

- Este modelo é composto pelo produto de dois componentes, um não-paramétrico e o outro paramétrico.
- O componente não-paramétrico, $h_0(t)$, não é especificado e é uma função não-negativa do tempo.
- Este componente é geralmente chamado de função de base ou função básica pois $h(t) = h_0(t)$ quando $x = 0$.
- O componente paramétrico, ou componente linear é frequentemente usado na forma multiplicativa garantindo que $h(t)$ seja sempre não-negativa.

- É importante citar que o componente não-paramétrico absorve o termo constante, β_0 , presente nos modelos paramétricos.
- Este modelo semiparamétrico torna-se mais flexível que o modelo paramétrico devido a presença da função de base.
- Existe outras formas possíveis para $\psi(x_i)$, mas essa é a mais comumente usada para modelos de dados de sobrevivência.
- Este modelo é também denominado modelo de riscos proporcionais pois a razão das taxas de falha de dois indivíduos diferentes é constante no tempo.
- Isto é, a razão das funções de risco para os indivíduos i e j é

$$\frac{h_i(t)}{h_j(t)} = \frac{h_0(t) \exp(x'_i \beta)}{h_0(t) \exp(x'_j \beta)} = \exp \{x'_i \beta - x'_j \beta\}$$

- Esta razão de riscos não depende do tempo.
- Se um indivíduo no início do estudo tem um risco de falha igual a duas vezes o risco de um outro indivíduo, esta razão de riscos será a mesma para todo o período de acompanhamento.
- O modelo de riscos proporcionais também pode ser escrito em termos da função de risco acumulada ou da função de sobrevivência.

$$H(t/x) = H_0(t) \exp\{x'\beta\}$$

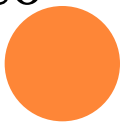
$$S(t/x) = [S_0(t)]^{\exp\{x'\beta\}}$$

$$H_0(t) = \int_0^t h_0(s) ds$$

$$\hat{S}_0(t) = \exp\{-\hat{H}_0(t)\}$$



ESTIMAÇÃO DOS PARÂMETROS

- O modelo de Cox é caracterizado pelos coeficientes β 's, que medem os efeitos das covariáveis sobre a função de risco.
 - Para que o modelo fique determinado, estas quantidades devem ser estimadas a partir dos dados amostrais.
 - Partindo do pressuposto de proporcionalidade, é possível estimar os efeitos das covariáveis sem ter que fazer qualquer suposição a respeito da distribuição do tempo de vida.
 - A função de risco básica e os coeficientes β 's podem ser estimados separadamente.
 - Os β 's são estimados primeiro e estas estimativas são então usadas para construir uma estimativa da função de risco básica.
- 

- Este é um resultado importante pois assim é possível fazer inferências sobre os efeitos das p variáveis explicativas no risco relativo sem precisar estimar a função de risco básica.
- Os coeficientes β 's podem ser estimados usando o método de máxima verossimilhança.
- Contudo, a presença do componente não-paramétrico ($h_0(t)$) na função de verossimilhança torna esse método inadequado.
- A solução proposta por Cox consiste em condicionar a construção da função de verossimilhança ao conhecimento da história passada de falhas e censuras para eliminar a função de risco básica.

- Este método é chamado de método de máxima verossimilhança parcial.
- Considere que em uma amostra de n indivíduos existam $k \leq n$ falhas distintas nos tempos $t_1 \leq t_2 \leq \dots \leq t_k$.
- A idéia básica deste método é considerar a probabilidade condicional da i -ésima observação vir a falhar no tempo t_i conhecendo quais observações estão sob risco em t_i .
- Esta probabilidade condicional, que é a razão entre o risco do indivíduo falhar em t_i e a soma dos riscos de falha de todos os indivíduos em risco, é a contribuição de cada indivíduo no tempo de falha t_i .
- Então a verossimilhança individual L_i será,


$$L_i = \frac{h_i(t_i)}{\sum_{j \in R(t_i)} h_j(t_j)} = \frac{h_0(t) \exp \{x'_i \beta\}}{\sum_{j \in R(t_i)} h_0(t) \exp \{x'_j \beta\}} = \frac{\exp \{x'_i \beta\}}{\sum_{j \in R(t_i)} \exp \{x'_j \beta\}}$$



- $R(t_i)$ é o conjunto dos índices das observações sob risco no tempo t_i .
- Assim, condicional a história de falhas e censuras até o tempo t_i , o componente não paramétrico desaparece da expressão de verossimilhança.
- A função de verossimilhança é dada por

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

- Os estimadores de máxima verossimilhança de β são obtidos a partir da verossimilhança parcial, $L(\beta)$.

- O modelo de risco proporcional para dados de sobrevivência e sua função de verossimilhança parcial assumem que os tempos de sobrevivência são contínuos.
 - Sob esta suposição, não permitem empates nos valores observados.
 - Como o tempo de sobrevivência pode ser registrado em horas, dias, meses ou até anos podem ocorrer empates nos tempos de falha ou de censura.
 - Quando ocorrem empates entre falhas e censuras, usa-se a convenção de que a censura ocorreu após a falha, definindo assim as observações a serem incluídas no conjunto de risco em cada tempo de falha.
- 

- Para considerar empates entre tempos de falhas, a função de verossimilhança parcial pode ser modificada.
- Uma aproximação para a função de verossimilhança foi proposta por Breslow e Peto em 1972 e é freqüentemente usada em pacotes estatísticos pela sua forma simples.
- Esta aproximação é adequada quando o número de empates em qualquer tempo não é grande.
- Alguns autores provaram que os estimadores de máxima verossimilhança para o modelo de Cox são consistentes e assintoticamente normais sob certas condições de regularidade.



INTERPRETAÇÃO DOS COEFICIENTES

- O efeito das covariáveis no modelo de riscos proporcionais de Cox é de acelerar ou desacelerar a função de risco.
- Para interpretar os coeficientes estimados, a propriedade de riscos proporcionais do modelo deve ser usada.
- Considere 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.

$$\frac{h_i(t)}{h_j(t)} = \exp \{ \beta_l (x_{il} - x_{jl}) \}$$

- Considere que x_1 seja uma variável dicotômica indicando pacientes hipertensos.

- O risco de morte entre os hipertensos é $\exp\{\beta_1\}$ vezes o risco de pacientes com pressão normal, com as outras covariáveis mantida fixas.
- Seja $\psi = \exp\{\beta\}$, que é a taxa de falha relativa no tempo t , assim $\hat{\psi} = \exp\{\hat{\beta}\}$
- Para verificar a existência de diferenças significativas entre os grupos, basta observar se o valor 1 pertence ao intervalo de confiança estimado.
- Caso isto ocorra não há evidências de que os riscos dos pacientes nos dois grupos apresentam diferenças significativas.

- **EXEMPLO:** Considere uma covariável grupo com três níveis, representada por x_1 : grupo 1 e x_2 : grupo 2. As estimativas de máxima verossimilhança parcial com I.C. entre parênteses são:

$$\exp\{\hat{\beta}_1\} = 2,0(1,5;4,1)$$

$$\exp\{\hat{\beta}_2\} = 1,2(0,7;1,8)$$

- Existe diferença significativa entre o grupo controle e grupo 1, mas não existe diferença entre o grupo controle e grupo 2.
- O risco de falha para pacientes do grupo 1 é duas vezes o risco dos pacientes do grupo controle.
- Considere agora a covariável idade com efeito significativo e estimativa pontual dada por $\exp\{\hat{\beta}\} = 1.05$
- Temos então que se aumentarmos em um ano a idade, o risco de falha fica aumentado em 5%.


AVALIAÇÃO DA PROPORCIONALIDADE DOS RISCOS

- Uma avaliação inicial da proporcionalidade do efeito das covariáveis no tempo pode ser feita através da construção das curvas de Kaplan-Meier.
- A suposição de proporcionalidade ao longo do tempo, será aceita se não houver cruzamento entre as curvas de sobrevivência por categorias das variáveis.
- Uma outra forma de avaliar a suposição de proporcionalidade é através da análise de resíduos de Schoenfeld.
- Considere que se o i -ésimo indivíduo com vetor de covariáveis $\mathbf{x}_i = (x_{1i}, \dots, x_{pi})'$ é observado falhar.



- Tem-se para este indivíduo um vetor de resíduos de Schoenfeld $r_i = (r_{i1}, \dots, r_{ip})$ dado por

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

- Estes resíduos são interpretados como a diferença entre os valores observados de covariáveis de um indivíduo com tempo de ocorrência do evento t_i e os valores esperados em t_i dado o grupo de risco $R(t_i)$.
 - Estes resíduos são definidos apenas nos tempos de falha.
 - O número de vetores de resíduos é igual ao número de covariáveis ajustadas no modelo.
- 

- Dessa forma, através do gráfico dos resíduos padronizados de Schoenfeld contra o tempo é possível verificar a existência ou não de proporcionalidade.
- Isto é, se a suposição de riscos proporcionais for satisfeita não deverá existir nenhuma tendência sistemática no gráfico.
- É possível realizar um teste para verificar a hipótese de que não existe correlação entre o tempo de sobrevivência transformado e os resíduos padronizados.
- Isto equivale a testar a hipótese nula de que não existe tendência no tempo ($H_0: \rho=0$).

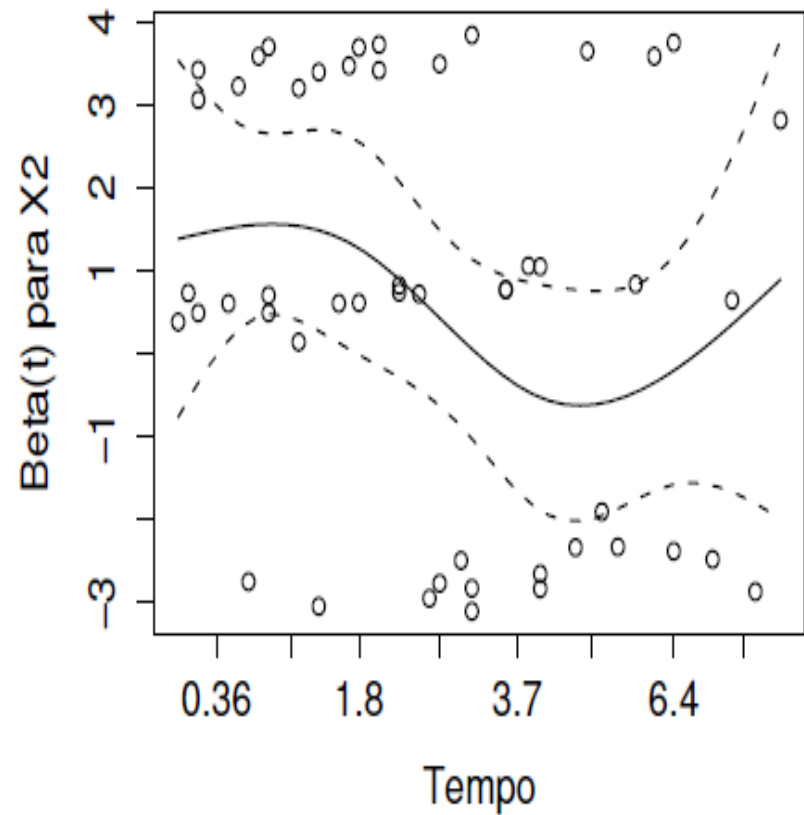
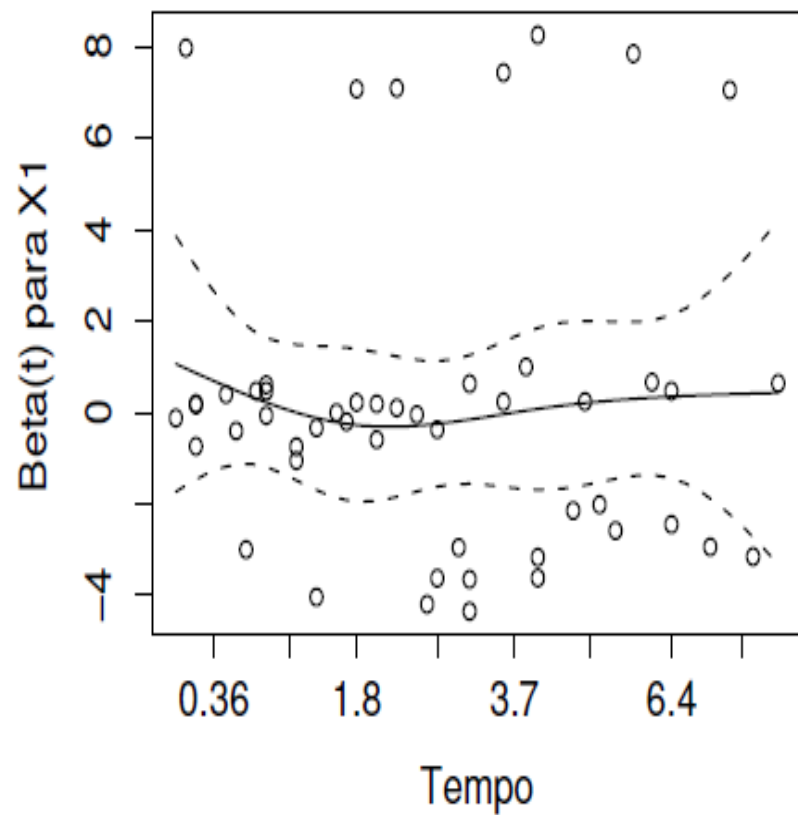


Figura 5.2: Resíduos escalonados de Schoenfeld *versus* tempo para as covariáveis X_1 (gráfico à esquerda) e X_2 (gráfico à direita).

Se o modelo de riscos proporcionais for apropriado, os gráficos dos resíduos r_i^* versus t_i , para cada uma das p covariáveis, não deveriam exibir tendências ao longo do tempo t . A Figura 5.3 ilustra tais gráficos em uma situação em que duas covariáveis (X_1 e X_2) são consideradas. O gráfico à esquerda, mostrado nesta figura, não apresenta nenhuma tendência acentuada ao longo do tempo. O mesmo não se pode concluir para o gráfico à direita. A suposição de riscos proporcionais parece, portanto, não ser apropriada e há evidências de que a covariável X_2 esteja gerando a violação desta suposição.



AVALIAÇÃO DO AJUSTE DO MODELO

- Os mesmos testes aplicados aos modelos paramétricos, também podem ser utilizados no modelo de Cox.
- A estatística de Wald pode ser utilizada tanto para testar a significância do parâmetro do modelo, como verificar o ajuste global do mesmo.
- O teste da razão de verossimilhança (análise da função desvio) compara modelos encaixados.
- Avalia se a inclusão de uma ou mais variáveis no modelo aumenta de modo significativo a verossimilhança de um modelo em relação ao modelo com menos parâmetros.
- A função desvio é assintoticamente semelhante a estatística de Wald quando o número de observações é grande. Caso esse número seja pequeno, a análise da função desvio é mais robusta.

AVALIAÇÃO DO AJUSTE DO MODELO

Para o modelo de Cox, os resíduos de Cox e Snell (1968) são definidos por:

$$\hat{e}_i = \hat{H}_0(t_i) \exp \left\{ \sum_{k=1}^p x_{ip} \hat{\beta}_k \right\}, \quad i = 1, \dots, n.$$

com $\hat{H}_0(t_i)$ estimado por

$$\hat{H}_0(t_i) = \sum_{j: t_j \leq t} \frac{d_j}{\sum_{l \in R_j} \exp\{\mathbf{x}_l' \hat{\beta}\}}$$

com d_j o número de falhas em t_j .

Se o modelo estiver bem ajustado, os \hat{e}_i 's podem ser olhados como uma amostra censurada de uma distribuição exponencial padrão e, então, o gráfico de, por exemplo, $\hat{H}(\hat{e}_i)$ versus \hat{e}_i deveria ser aproximadamente uma reta. Assim como nos modelos paramétricos, os resíduos de Cox-Snell são úteis para examinar o ajuste global do modelo de Cox.

AVALIAÇÃO DO AJUSTE DO MODELO

- PERGUNTA: Qual o poder explicativo de um modelo escolhido para avaliar os dados?
- Uma medida de qualidade de ajuste para modelos lineares é o R^2 .
- Poucas são as medidas estatísticas disponíveis para avaliar globalmente a qualidade de ajuste de um modelo de sobrevivência.
- A mais simples delas é uma medida baseada na razão de verossimilhanças e está disponível no R.



EXEMPLO: Aleitamento materno

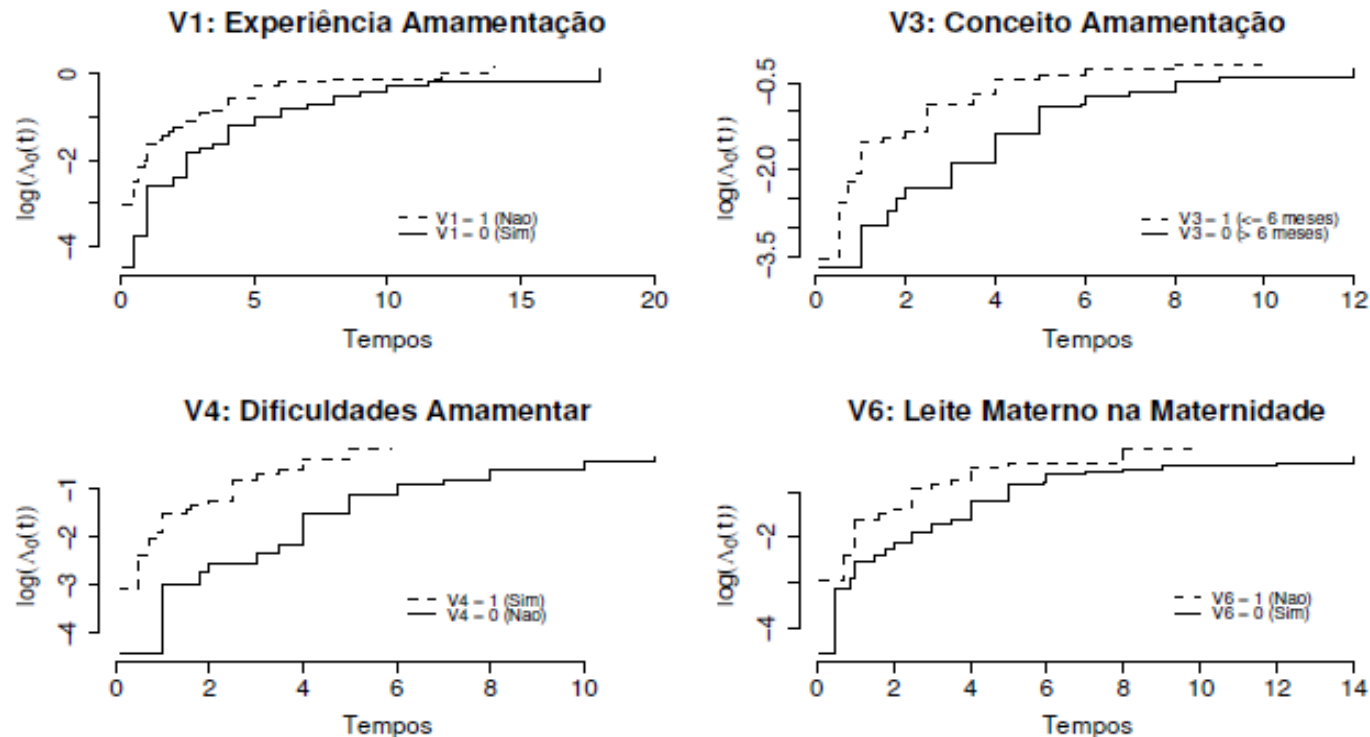


Figura 5.8: $\text{Log}(\hat{\Lambda}(t))$ versus tempo para as covariáveis V1, V3, V4 e V6.

Na Figura 5.8 encontram-se os gráficos envolvendo o logaritmo da função de risco acumulada para as covariáveis V1, V3, V4 e V6. Como pode ser observado desta figura, as curvas não indicam violação da suposição de riscos proporcionais. Embora as mesmas não sejam perfeitamente paralelas ao longo do eixo do tempo, não existem, em termos descritivos, afastamentos marcantes desta característica. A situação extrema de violação é caracterizada por curvas que se cruzam.

EXEMPLO: Aleitamento materno

Tabela 5.7: Resultado do ajuste do modelo de regressão de Cox para os dados de aleitamento materno.

Covariável	Estimativa	Erro-Padrão	Valor- <i>p</i>	RR	IC(RR, 95%)
V1: Exper. Amam.	0,471	0,268	0,079	1,601	(0,94; 2,71)
V3: Conc. Amam.	0,579	0,262	0,027	1,785	(1,07; 2,99)
V4: Dific. Amam.	0,716	0,264	0,007	2,046	(1,22; 3,43)
V6: Leite Excl.	0,578	0,264	0,028	1,783	(1,06; 2,99)

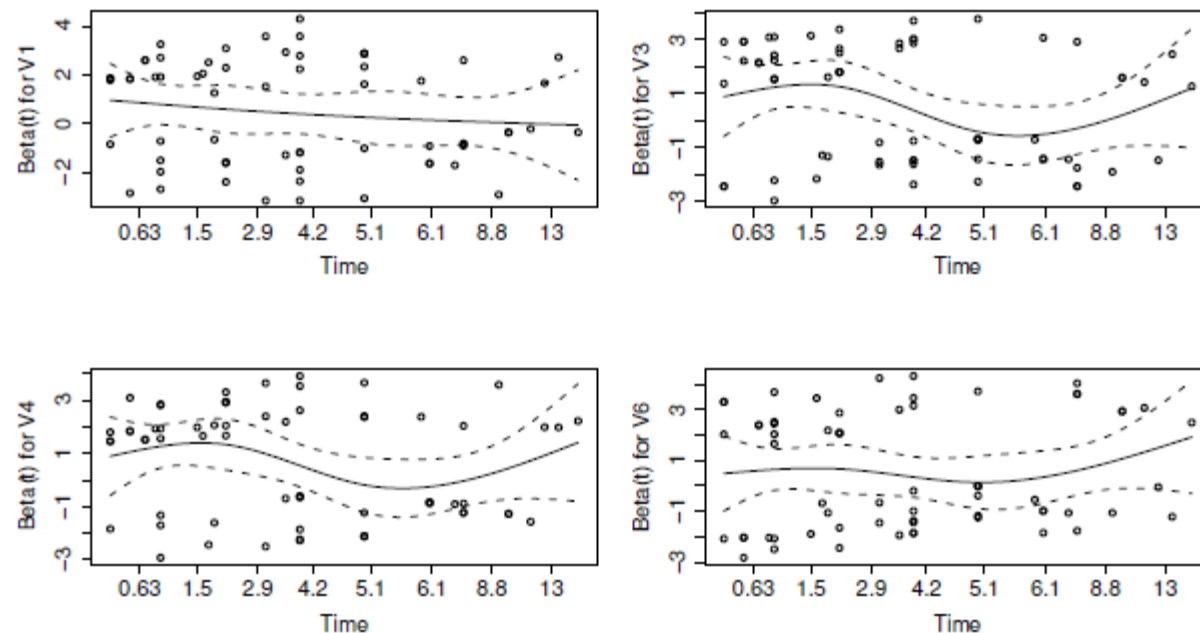


Figura 5.9: Suposição de riscos proporcionais para as covariáveis V1, V3, V4 e V6 fazendo uso dos resíduos escalonados de Schoenfeld.

EXEMPLO: Aleitamento materno

- i) O risco de desmame precoce em mães que não tiveram experiência anterior de amamentação é 1,6 vezes o risco das mães que tiveram essa experiência. Além disso, podemos afirmar com 95% de confiança que esse risco varia entre 0,95 e 2,71.
- ii) O risco 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 o risco das mães que acreditam que o tempo ideal de amamentação é superior a 6 meses. Além disso, podemos afirmar com 95% de confiança que esse risco varia entre 1,07 e 2,99.
- iii) O risco de desmame precoce em mães que apresentaram dificuldades de amamentar nos primeiros dias pós-parto é aproximadamente 2 vezes o risco das mães que não apresentaram essas dificuldades. Além disso, podemos afirmar com 95% de confiança que esse risco é superior a 1,22.
- iv) O risco de desmame precoce em crianças que não receberam exclusivamente leite materno na maternidade é 1,8 vezes o risco de desmame precoce em crianças que receberam exclusivamente o leite materno. Além disso, podemos afirmar com 95% de confiança que esse risco varia entre 1,06 e 2,99.

EXEMPLO: Leucemia Pediátrica

Tabela 5.8: Descrição das covariáveis utilizadas no estudo sobre leucemia pediátrica.

Código	Descrição	Categorias
LEUINI	Número de leucócitos no sangue periférico	0 se ≤ 75000 leucócitos/mm ³ 1 se > 75000 leucócitos/mm ³
IDADE	Idade em meses	0 se ≤ 96 meses 1 se > 96 meses
ZPESO	Peso padronizado pela idade e sexo	0 se ≤ -2 e 1 se > -2
ZEST	Altura padronizada pela idade e sexo	0 se ≤ -2 e 1 se > -2
PAS	Porcentagem de linfoblastos medulares que reagiram positivamente ao ácido periódico de Schiff	0 se $\leq 5\%$ e 1 se $> 5\%$
VAC	Porcentagem de vacúolos no citoplasma dos linfoblastos	0 se $\leq 15\%$ e 1 se $> 15\%$
RISK	Fator de risco obtido a partir de uma fórmula que é função dos tamanhos do fígado e do baço e do número de blastos	0 se $\leq 1,7\%$ e 1 se $> 1,7\%$
R6	Remissão na sexta semana de tratamento	0 se não e 1 se sim