

Modelos Birnbaum-Saunders para sobrevivência com dados longitudinais

Diana Carolina Franco Soto

TESE APRESENTADA
AO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA
OBTENÇÃO DO TÍTULO
DE
DOUTOR EM CIÊNCIAS

Programa: Estatística

Orientador: Prof. Dr. Antonio Carlos Pedroso de Lima

Coorientador: Prof. Dr. Julio da Motta Singer

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro da
CAPES/CNPq

São Paulo, fevereiro 14 de 2014

Resumo

Em muitos estudos de diferentes áreas do conhecimento, medições repetidas de uma ou mais variáveis (respostas longitudinais), tempos até a ocorrência de um ou mais eventos (respostas de sobrevivência) e informações adicionais de variáveis explicativas são coletadas nas unidades analisadas com o intuito de caracterizar as relações existentes entre as respostas e as variáveis explicativas. Geralmente dados com essa natureza são analisados considerando as respostas longitudinais e de sobrevivência separadamente. No entanto, em muitos casos, uma modelagem conjunta pode facilitar a compreensão dos mecanismos subjacentes ao fenômeno estudado, além de melhorar as propriedades dos estimadores dos parâmetros de interesse. Nesse contexto, a maioria das metodologias propostas na literatura utiliza o modelo de Cox para descrever os tempos de sobrevivência, mesmo quando existem evidências da adequabilidade de algum modelo paramétrico. Adicionalmente, os métodos usuais para a modelagem conjunta utilizam apenas as informações de indivíduos que têm pelo menos uma observação da resposta longitudinal, embora possam existir unidades em estudo nas quais não são realizadas tais medições. Conseqüentemente, os estimadores correspondentes podem apresentar vícios e diminuição de precisão. Nesta tese propõe-se uma metodologia para a modelagem conjunta de dados longitudinais e de sobrevivência que além de utilizar um modelo linear de efeitos mistos para explicar a resposta longitudinal, considera um modelo paramétrico Birnbaum-Saunders para descrever a resposta de sobrevivência e leva em conta a informação de sobrevivência dos indivíduos para os quais a resposta longitudinal não foi observada. Após a obtenção dos resultados sobre aspectos inferenciais, o desempenho do modelo conjunto proposto é avaliado por meio de um exaustivo estudo de simulação. Finalmente, a metodologia em menção é aplicada ao conjunto de dados reais que motivou o presente trabalho, o qual é proveniente de um estudo desenvolvido no Instituto do Coração (InCor), São Paulo, Brasil, em que é feito o acompanhamento de uma coorte de pacientes portadores de Insuficiência Cardíaca Congestiva (ICC).

Palavras-chave: Dados longitudinais, dados de sobrevivência, modelagem conjunta, modelo Birnbaum-Saunders, modelo linear de efeitos mistos.

Capítulo 1

Introdução

Em pesquisas desenvolvidas em diferentes áreas do conhecimento, são comuns estudos longitudinais, em que medições repetidas de uma ou mais variáveis respostas, tempos até a ocorrência de um ou mais eventos e informações adicionais de variáveis explicativas que podem variar ou não com o tempo, são coletadas em unidades selecionadas de uma ou mais populações. O principal objetivo desse tipo de estudos é caracterizar as relações existentes entre as respostas e as variáveis explicativas. Embora o caso mais geral possa envolver a observação de múltiplas respostas longitudinais e tempos até a ocorrência de vários eventos do mesmo tipo ou de tipos diferentes ¹, neste trabalho a atenção será concentrada no caso em que nas unidades em estudo observam-se apenas medições repetidas de uma variável resposta contínua e o tempo até a ocorrência de um único evento de interesse. Esse é o caso de estudos clínicos destinados a identificar fatores prognósticos ou a analisar o efeito de um determinado tratamento sobre o tempo até a ocorrência de um evento (progressão para um nível específico da condição médica estudada, morte, abandono do estudo, etc.). Nos estudos do Vírus da Imunodeficiência Humana (HIV), por exemplo, os pacientes infectados são monitorados até que desenvolvam a Síndrome da Imunodeficiência Adquirida (AIDS) ou morram. Similarmente, em estudos oncológicos, o evento analisado é a ocorrência de metástase ou o óbito e, em estudos sobre Insuficiência Cardíaca Congestiva (ICC), o interesse concentra-se no tempo até o transplante de coração ou a morte.

Nesses estudos, uma resposta que reflete o estado de saúde do paciente é coletada em diferentes visitas, além de variáveis explicativas como características demográficas, clínicas, laboratoriais ou provenientes de exames complementares, medidas inicialmente ou em diferentes instantes. Nos estudos de HIV, por exemplo, observa-se uma resposta do estado

¹Cada unidade em estudo pode estar sujeita a múltiplos eventos do mesmo tipo (eventos recorrentes) ou a eventos de tipos diferentes. Nestes casos, mais de um tempo de sobrevivência é observado para cada unidade.

imunológico (contagem de células CD4 ou contagem de células CD8) ou uma resposta do estado virológico (número de cópias de RNA viral); em estudos de câncer de próstata, mede-se uma resposta do processo de carcinogênese (níveis de Antígeno Prostático Específico (PSA) no sangue) e, em estudos de ICC, coleta-se uma resposta que reflete o estado do funcionamento do coração (níveis sanguíneos de Peptídeo Natriurético Tipo-B (BNP)). O tempo até a ocorrência do evento de interesse (tempo de falha) também é registrado para cada participante, embora possam ocorrer situações em que o acompanhamento de alguns pacientes é interrompido porque o paciente abandonou o estudo, porque se perdeu o contato com ele, porque o estudo terminou antes da ocorrência do evento ou por outro motivo. Nesses casos, a informação relativa à resposta corresponde ao fato de o tempo de falha ser superior ao observado (tempo de censura) ², apesar de ser possível que o evento nunca ocorra para um ou mais pacientes pela causa estudada.

Uma das finalidades da coleta de informação da resposta longitudinal é utilizá-la como um marcador substituto do tempo de sobrevivência antes da ocorrência do evento, permitindo monitorar a evolução do estado de saúde do paciente e fornecendo informações que podem ser úteis para a tomada de medidas adequadas. Mais detalhes sobre este tipo de substituição podem ser obtidos em Prentice (1989).

Geralmente dados como os descritos anteriormente são analisados considerando as respostas longitudinais e de sobrevivência separadamente. No entanto, em muitos casos, uma modelagem conjunta pode facilitar a compreensão dos mecanismos subjacentes ao fenômeno estudado, além de melhorar as propriedades dos estimadores dos parâmetros de interesse [Ver Tsiatis e Davidian (2004), Yu et al. (2004) e Rizopoulos (2010), entre outros]. Nesse contexto, a maioria das metodologias propostas na literatura utiliza o modelo de Cox para descrever os tempos de sobrevivência, mesmo quando existem evidências da adequabilidade de algum modelo paramétrico. Adicionalmente, os métodos usuais para a modelagem conjunta utilizam apenas as informações de indivíduos que têm pelo menos uma observação da resposta longitudinal, embora possam existir unidades em estudo nas quais não são realizadas tais medições. Conseqüentemente, os estimadores correspondentes podem apresentar vícios e diminuição de precisão.

Nesta tese propõe-se uma metodologia para a modelagem conjunta de dados longitudinais e de sobrevivência que além de utilizar um modelo linear de efeitos mistos para explicar a resposta longitudinal, considera um modelo paramétrico Birnbaum-Saunders para descrever a resposta de sobrevivência e leva em conta a informação de sobrevivência dos indivíduos para os quais a resposta longitudinal não foi observada.

²A presença de censura, que se refere à observação parcial do tempo de falha é uma característica comum de dados de sobrevivência que deve ser levada em conta na respectiva análise.

O trabalho foi motivado pela análise de um conjunto de dados reais proveniente de um estudo de ICC desenvolvido no Instituto do Coração (InCor) do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (FMUSP), São Paulo, Brasil.

A ICC é uma doença crônica caracterizada pela perda da capacidade do coração de bombear o sangue com eficiência. Pelo fato de o volume de sangue bombeado ser insuficiente para suprir as demandas normais de oxigênio e de nutrientes do corpo, portadores de ICC padecem de sintomas limitantes como falta de ar e fraqueza generalizada, entre outros, além de ser mais suscetíveis a complicações, que podem levar à morte.

Considera-se o estudo desenvolvido no InCor, cuja principal finalidade é avaliar o valor prognóstico de características demográficas, clínicas, laboratoriais e provenientes de exames complementares (ecocardiograma e eletrocardiograma) de pacientes portadores de ICC, com relação ao tempo até a morte.

Num conjunto de $n = 1609$ pacientes portadores de ICC atendidos no Ambulatório Geral do InCor no período de 17 de janeiro de 2003 a 13 de dezembro de 2011, foram coletadas:

- i) uma sequência de medições longitudinais dos níveis de Peptídeo Natriurético Tipo-B (BNP) no sangue (pg/mL)³;
- ii) um registro do tempo entre a admissão do paciente no protocolo do Ambulatório Geral do InCor e a data de óbito (T_i) ou da censura (C_i), em meses. A censura pode ser ocasionada por perda de contato com o paciente (nessa circunstância, a data final é aquela em que as últimas informações sobre o paciente foram obtidas) ou pela realização de um transplante de coração (nesse caso, a data final é a correspondente à do procedimento cirúrgico);
- iii) um valor de cada uma das seguintes covariáveis coletadas na data de admissão do paciente no estudo:
 - Covariáveis demográficas informadas pelo paciente ou referidas pelo profissional responsável pelo diagnóstico:

³O BNP, também denominado Peptídeo Natriurético Cerebral, é um hormônio produzido no miocárdio atrial ou ventricular e um indicador de pressões intracardíacas aumentadas, cuja síntese e liberação são ativadas por estiramento ventricular diastólico. Desta forma, o BNP é estimulado por uma ampla gama de doenças cardíacas estruturais, como a ICC. Quando um indivíduo é portador de ICC, os níveis sanguíneos de BNP são afetados de forma considerável, de modo que a sua observação constitui uma medida da deterioração do funcionamento do coração causada pela doença. Em particular, quanto maior o nível de BNP no plasma sanguíneo, maior a probabilidade de um pior prognóstico da evolução da doença. Por outro lado, o BNP apresenta alta variabilidade biológica ao longo do tempo (variabilidade inerente). Além disso, a medição dos níveis de BNP no plasma sanguíneo é feita por meio de imunoensaio e depende da precisão dos instrumentos utilizados e, em alguns casos, de habilidades próprias do observador, pois as amostras devem ser analisadas em períodos de tempo muito próximos de sua coleta por ter a característica de se degradar rapidamente (variabilidade associada às condições de medição).

- ★ Idade (em anos),
- ★ Gênero (masculino, feminino).
- Covariáveis clínicas coletadas com o auxílio de instrumentos ou com questionamento ao paciente sobre sua condição frente a algumas condições impostas pelo profissional na coleta dos dados:
 - ★ Índice de massa corporal (kg/m^2),
 - ★ Pressão arterial sistólica ($mmHg$),
 - ★ Pressão arterial diastólica ($mmHg$),
 - ★ Frequência cardíaca (batimentos por minuto),
 - ★ Diagnóstico: causa da ICC, segundo o código CID10 (*Classificação Estatística Internacional de Doenças e Problemas Relacionados com a Saúde*) (cardiomiopatia da doença de Chagas, cardiomiopatia dilatada, cardiomiopatia isquêmica, cardiomiopatia hipertensiva, cardiomiopatia alcoólica, outras cardiomiopatias),
 - ★ Classe funcional: classificação dos indivíduos com ICC de acordo com suas limitações quanto à realização de atividades físicas, segundo o critério da NYHA (*New York Heart Association*)⁴,
 - ★ Hipertensão arterial (não, sim),
 - ★ Etilismo (não, sim, ex-etilista),
 - ★ Tabagismo (não, sim, ex-tabagista),
 - ★ Diabetes melito (não; sim, depende da insulina; sim, não depende da insulina).
- Covariáveis laboratoriais coletadas a partir da realização de exames de sangue:
 - ★ Sódio sérico: nível de sódio no sangue em relação à quantidade de água no corpo (mEq/L),
 - ★ Potássio sérico: concentração de potássio no sangue (mEq/L),
 - ★ Hemoglobina sérica: concentração de hemoglobina no sangue (mg/dL),
 - ★ Hematócrito sérico: relação entre o volume de células vermelhas e o volume total de sangue (%),
 - ★ Quantidade de leucócitos no sangue (mm^3),
 - ★ Taxa de triglicérides no sangue (mg/dL),
 - ★ Taxa de colesterol total no sangue (mg/dL),

⁴ I: não apresenta sintomas ou dificuldade para a realização de atividades físicas, II: apresenta sintomas leves e limitações em atividades rotineiras, III: com limitações em exercer atividades físicas, o exercício de atividades menos intensas que as rotineiras produzem sintomas e há sensação de conforto somente quando em repouso, IV: apresenta limitações severas e sintomas mesmo quando em estado de repouso.

-
- ★ Fração de colesterol HDL no sangue (mg/dL),
 - ★ Fração de colesterol LDL no sangue (mg/dL),
 - ★ Glicemia: taxa de glicose no sangue, em jejum (mg/dL),
 - ★ Creatinina sérica: taxa de creatinina no sangue (mg/dL).
- Covariáveis provenientes de exames complementares, referentes a dados coletados por meio de exames solicitados pelo cardiologista, que visam relatar a fisiologia e funcionamento das diversas estruturas do coração:
 - a) Ecocardiograma ⁵, gerando as seguintes medidas:
 - ★ Espessura do septo interventricular (mm),
 - ★ Espessura da parede posterior do ventrículo esquerdo (mm),
 - ★ Diâmetro diastólico do ventrículo esquerdo (mm),
 - ★ Diâmetro sistólico do ventrículo esquerdo (mm),
 - ★ Fração de ejeção do ventrículo esquerdo (%) ⁶,
 - ★ Diâmetro do átrio esquerdo (mm).
 - b) Eletrocardiograma ⁷, gerando a seguinte medida:
 - ★ Ritmo cardíaco (sinusal, fibrilação atrial, outros).

Observaram-se pacientes de ambos os gêneros, com idades entre 18 e 93 anos, com diferentes diagnósticos e classificados em classes funcionais de I a IV segundo o critério da NYHA. A presença de dados omissos nas covariáveis basais (exceto idade, gênero e diagnóstico) foi levada em consideração.

Na amostra avaliada 63% são homens, 55% têm entre 55 e 93 anos de idade, 58% apresentam menor gravidade (classes funcionais I e II) e existe uma concentração de pacientes diagnosticados com cardiomiopatia dilatada, isquêmica ou hipertensiva (79%). Além disso, somente 15% dos pacientes são etilistas e 20% são tabagistas; a proporção de pacientes hipertensos é de 73%, a proporção de indivíduos diabéticos é de 22% e segundo os resultados do eletrocardiograma, 62% dos pacientes são portadores de um ritmo cardíaco sinusal.

Foram calculados os coeficientes de correlação entre as covariáveis basais encontrando-se altas correlações positivas entre algumas delas, como apresentado na Tabela 1.1. Na análise subsequente, de acordo com as indicações dos pesquisadores, foram descartadas as variáveis pressão arterial diastólica, sódio sérico, hematócrito sérico, taxa de colesterol total no sangue, fração de colesterol HDL no sangue e espessura do septo interventricular, por causa de sua

⁵Exame que utiliza ondas sonoras de modo a construir imagens, permitindo obter dados sobre a estrutura e funcionamento do coração.

⁶Representa o percentual do volume ventricular diastólico final que foi ejetado na sístole.

⁷Exame no qual é registrada a variação dos potenciais elétricos gerados pela atividade elétrica do coração.

alta correlação com as variáveis pressão arterial sistólica, potássio sérico, hemoglobina sérica, fração de colesterol LDL no sangue e a espessura da parede posterior do ventrículo esquerdo, respectivamente.

A Figura 1.1 mostra a estimativa de Kaplan-Meier para a proporção de sobreviventes

Tabela 1.1: *Variáveis quantitativas com alta correlação.*

Variáveis	Correlação
Pressão arterial sistólica e pressão arterial diastólica	0,77
Sódio sérico e potássio sérico	0,85
Hemoglobina sérica e hematócrito sérico	0,96
Taxa de colesterol total e fração de colesterol HDL no sangue	0,68
Taxa de colesterol total e fração de colesterol LDL no sangue	0,89
Fração de colesterol HDL e fração de colesterol LDL no sangue	0,61
Espessura do septo interventricular e espessura da parede posterior do ventrículo esquerdo	0,94

no conjunto de 1609 pacientes portadores de ICC. Aproximadamente 80% dos pacientes não morreram no período de observação. Apesar da incerteza com respeito à causalidade de muitas das mortes neste conjunto de dados, isto é, se elas são devidas à ICC ou a outros motivos, todos os óbitos observados serão assumidos como falhas, tendo em vista as considerações de Lauer et al. (1999), Freitas et al. (2005) e Giolo et al. (2012), entre outros.

Embora o alto percentual de censuras possa ser um indício da presença de curados (Maller e Zhou, 1996), é razoável supor que todos os pacientes diagnosticados com ICC sejam suscetíveis ao evento de interesse. Neste caso, é provável que o alto percentual de censuras se deva ao fato de o período de acompanhamento não ser suficientemente longo para observar o evento em todos os pacientes, porque: i) alguns entraram tardiamente no estudo (os pacientes iniciaram o seguimento em instantes diferentes e aleatórios no período de 17 de janeiro de 2003 a 05 de junho de 2009), ii) alguns se retiraram precocemente do estudo, ou, iii) apesar de o quadro clínico da ICC apresentar um agravamento no decorrer do tempo, a evolução desta doença geralmente é lenta e os doentes podem sobreviver muitos anos. Em particular, a doença pode ser controlada por meio de tratamento, de forma que pacientes tratados podem ter prognóstico favorável e apresentar uma sobrevida muito além da esperada para pacientes não tratados.

Com o intuito de identificar as variáveis que mais influenciam o tempo de vida dos pacientes em estudo, considera-se uma análise unidimensional baseada em testes Log-Rank [Ver Colosimo e Giolo (2006)]. As variáveis quantitativas foram divididas em categorias pré-estabelecidas de acordo com critérios clínicos (Ver Apêndice A).

No Apêndice B são apresentados os valores-p para os testes Log-Rank, considerando as categorizações mencionadas anteriormente. Encontraram-se efeitos significativos na so-

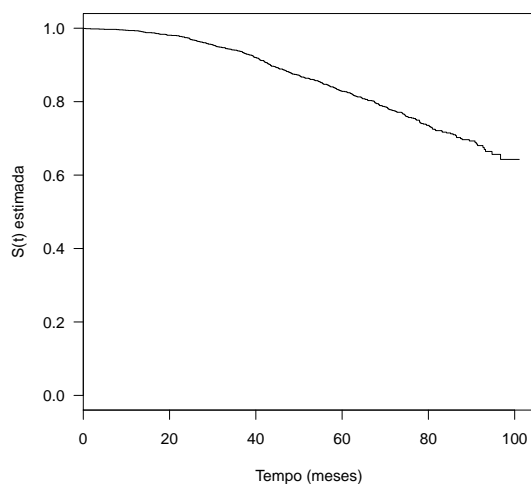


Figura 1.1: *Estimativa de Kaplan-Meier para a proporção de sobreviventes no conjunto de 1609 pacientes portadores de ICC.*

brevidade dos pacientes (em nível de 10%) para as variáveis: idade, gênero, índice de massa corporal, pressão arterial sistólica, diagnóstico, classe funcional, diabetes melito, potássio sérico, hemoglobina sérica, fração de colesterol LDL no sangue, creatinina sérica, fração de ejeção do ventrículo esquerdo, diâmetro do átrio esquerdo e ritmo cardíaco.

Para descrever o tempo até a falha, inicialmente foram ajustados modelos de regressão com as variáveis explicativas significativas segundo os testes Log-Rank. Consideraram-se três das distribuições de probabilidade mais comumente usadas na solução de problemas desta natureza (exponencial, Weibull e log-normal), além da distribuição Birnbaum-Saunders (BS). A utilização da distribuição BS é apropriada no contexto da ICC, pois supõe-se que em doenças cardíacas crônicas, um dano cumulativo causado por vários fatores de risco pode produzir uma degradação que leva a um processo de fadiga cujos tempos de propagação podem ser adequadamente modelados por dita distribuição [Ver Leiva, Barros e Paula (2009), entre outros]. Para avaliar a adequação desses modelos foi utilizado o critério da análise de resíduos de Cox-Snell (1968). Maiores detalhes podem ser consultados em Klein e Moeschberger (1997) e Colosimo e Giolo (2006).

Na Figura 1.2 são apresentados os gráficos dos resíduos de Cox-Snell obtidos a partir dos ajustes dos modelos exponencial, Weibull, log-normal e Birnbaum-Saunders versus a estimativa da função de taxa de falha acumulada calculada a partir deles. O modelo Birnbaum-Saunders parece ser o mais adequado.

A observação do BNP somente começou a ser considerada como critério de diagnóstico, evolução e tratamento da ICC há alguns anos e sua medição vem-se estabelecendo na

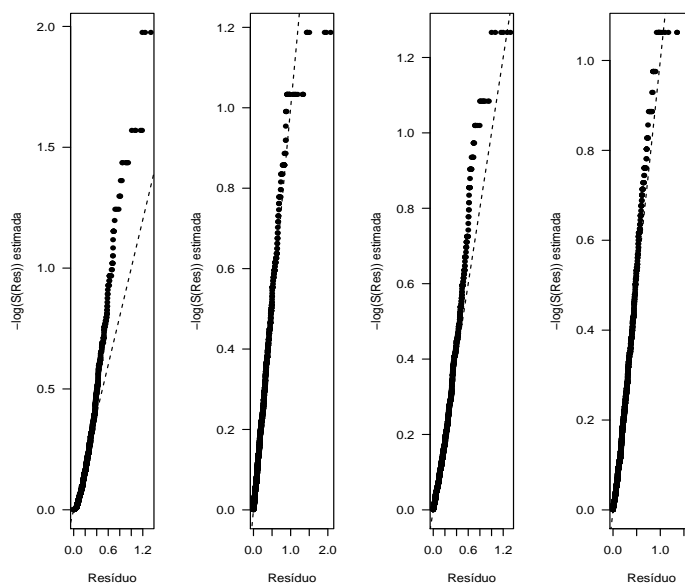


Figura 1.2: Gráficos dos resíduos de Cox-Snell obtidos a partir dos ajustes dos modelos exponencial, Weibull, log-normal e Birnbaum-Saunders versus estimativas da função de taxa de falha acumulada calculada a partir deles.

prática clínica de forma gradual, como consequência das conclusões de diferentes investigações multidisciplinares conduzidas para esclarecer o papel deste peptídeo no estudo de doenças cardíacas crônicas [Ver, por exemplo, Maisel (2002), Cowie et al. (2003) e Weber e Hamm (2006)]. Assim, a medição do BNP tem sido realizada no InCor desde 08 de junho de 2006, mas ainda não faz parte de um exame de rotina e os instantes de observação não são especificados pelo protocolo do estudo. Por essa razão, somente 33% dos indivíduos têm pelo menos uma medição de BNP ao longo do período de acompanhamento. No entanto, de acordo com os resultados apresentados em Bettencourt (2004) e Bettencourt et al. (2004), entre outros, as medições de BNP disponíveis podem fornecer informações de utilidade para a tomada de medidas adequadas no tratamento dos pacientes.

Na Figura 1.3 são apresentadas as estimativas de Kaplan-Meier para a proporção de sobreviventes nos conjuntos de pacientes portadores de ICC para os quais o BNP foi observado pelo menos uma vez ($n_{CBNP}=529$) e aqueles sem observações de BNP ($n_{SBNP}=1080$). Pode-se observar que embora os pacientes sem dados da resposta longitudinal apresentem uma evolução semelhante à dos pacientes que têm pelo menos uma observação dessa resposta, nem em todos os instantes o comportamento é o mesmo. É factível que em diferentes situações práticas as sobrevivências dos pacientes dos dois grupos sejam significativamente diferentes.

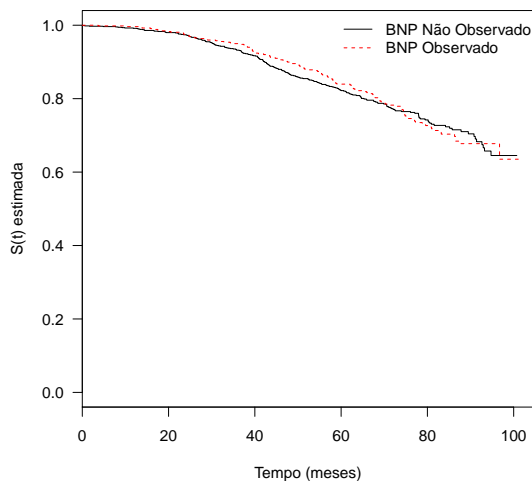


Figura 1.3: *Estimativas de Kaplan-Meier para a proporção de sobreviventes nos conjuntos de pacientes portadores de ICC com e sem observações da resposta longitudinal.*

O objetivo deste estudo é analisar o comportamento do tempo até o óbito dos pacientes portadores de ICC, condicionalmente às medições longitudinais de BNP, por meio de uma modelagem conjunta. Os métodos propostos até agora para a modelagem conjunta em questão utilizam apenas as informações de indivíduos que têm pelo menos uma medição da resposta longitudinal ao longo do tempo. Além de descartar parte da informação disponível, é provável que a análise baseada apenas nos pacientes que têm pelo menos uma medição da resposta longitudinal leve à obtenção de estimadores viesados e menos eficientes dos parâmetros de interesse, em comparação com aqueles que poderiam ser obtidos considerando um maior tamanho da amostra. O anterior torna evidente a importância do desenvolvimento de uma metodologia como a proposta nesta tese, que além de considerar na modelagem as informações dos indivíduos com pelo menos uma medição da resposta longitudinal, incorpore a informação proveniente daqueles sem observações dessa resposta.

Não há evidências de propostas para modelar conjuntamente dados longitudinais e de sobrevivência nas quais seja considerado o uso do modelo paramétrico Birnbaum-Saunders para descrever a resposta de sobrevivência, nem de propostas que incorporem as informações de indivíduos sem observações da resposta longitudinal na modelagem.

O Capítulo 2 é dedicado à apresentação de alguns conceitos preliminares da modelagem de dados longitudinais e de sobrevivência e à revisão de algumas das propostas de modelos conjuntos. A metodologia para a modelagem conjunta proposta nesta tese é descrita no Capítulo 3, no qual, além de se apresentarem detalhes sobre aspectos inferenciais, descreve-se o método para o cálculo das previsões dinâmicas das probabilidades de sobrevivência, com

base no ajuste do modelo conjunto considerado e no uso de todas as informações disponíveis até o momento em que se deseja realizar a previsão. No Capítulo 4 são apresentados os resultados de um estudo de simulação realizado com os objetivos de avaliar a robustez do modelo conjunto proposto relativamente a uma má especificação da distribuição paramétrica da resposta de sobrevivência, de analisar o comportamento das predições dinâmicas das probabilidades de sobrevivência descritas no capítulo anterior, calculadas considerando os diferentes instantes em que são coletadas as observações da resposta longitudinal e, de avaliar o desempenho do processo de estimação dos parâmetros nos cenários em que o modelo de sobrevivência utilizado para gerar os dados é igual ao considerado no modelo conjunto ajustado. Além disso, nesse capítulo aplica-se a metodologia em questão ao conjunto de dados reais provenientes do estudo de ICC que motivou o presente trabalho. Finalmente, as conclusões e as sugestões para trabalhos futuros são discutidas no Capítulo 5.

Capítulo 2

Modelagem de dados longitudinais e dados de sobrevivência

Suponha que um conjunto de n indivíduos seja acompanhado durante um intervalo de tempo $[0, \tau)$, $\tau > 0$, e que no i -ésimo indivíduo ($i = 1, \dots, n$) sejam observados: i) uma sequência de observações de uma resposta longitudinal $y_{ij} = \{y_i(t_{ij}), j = 1, \dots, n_i\}$ representadas no vetor $\mathbf{y}_i = (y_{i1}, \dots, y_{in_i})^T$, nos tempos t_{ij} representados no vetor $\mathbf{t}_i = (t_{i1}, \dots, t_{in_i})^T$; ii) um registro do tempo entre a admissão do indivíduo no estudo e a ocorrência do evento de interesse (T_i) ou da censura (C_i)¹ e, iii) valores das variáveis explicativas medidas no início ou em diferentes instantes do estudo. Os dados observados para cada indivíduo são, portanto, o vetor de respostas $(\mathbf{y}_i^T, Z_i, \delta_i)^T$, o vetor de tempos \mathbf{t}_i e, os vetores dos valores das p_h variáveis explicativas denotados por $\mathbf{x}_{hi}(t) = (1, x_{hi1}, \dots, x_{hia_h}, x_{hia_h+1}(t), \dots, x_{hip_h}(t))^T$, em que apenas as primeiras a_h não dependem do tempo e $h = 1, 2$ indica se corresponde ao modelo para o componente longitudinal ou ao modelo para o componente de sobrevivência, respectivamente.

2.1 Modelagem do componente longitudinal

A resposta longitudinal para o i -ésimo indivíduo no instante t , pode ser modelada por

$$\begin{aligned} y_i(t) &= m_i(t) + e_i(t) \\ &= \mu_{1i}(t) + w_{1i}(t) + e_i(t), \end{aligned} \tag{2.1}$$

¹Para levar em conta a presença de censuras, consideram-se a variável resposta $Z_i = \min(T_i, C_i)$ para o i -ésimo indivíduo e a variável indicadora do evento $\delta_i = I(T_i \leq C_i)$, isto é, $\delta_i = 1$ se uma falha foi observada e $\delta_i = 0$ se houve censura.

em que $m_i(t) = \mu_{1i}(t) + w_{1i}(t)$ denota o verdadeiro valor (não observável) da resposta longitudinal para o i -ésimo indivíduo no instante $t \geq 0$, especificado como uma função da resposta média $\mu_{1i}(t)$ e do processo $w_{1i}(t)$, que geralmente é caracterizado em termos dos traços latentes (não observáveis) do i -ésimo indivíduo e $e_i(t)$ denota o erro de medida correspondente ao instante t , que pode levar em conta tanto a variabilidade inerente à resposta, quanto a variabilidade associada às condições de medição. Em geral supõe-se que $e_i(t) \sim N(0, \sigma_e^2)$.

Muitas vezes, a resposta média em (2.1) é modelada de forma linear como

$$\mu_{1i}(t) = \mathbf{x}_{1i}(t)^T \boldsymbol{\beta}_1, \quad (2.2)$$

em que $\boldsymbol{\beta}_1 = (\beta_{10}, \beta_{11}, \dots, \beta_{1p_1})^T$ representa um vetor que contém os efeitos fixos correspondentes às p_1 variáveis explicativas consideradas no vetor $\mathbf{x}_{1i}(t)$. Em alguns casos, a resposta média $\mu_{1i}(t)$ pode ser modelada de forma não-linear ou de forma não-paramétrica. O leitor poderá consultar, por exemplo, Diggle et al. (2002), para maiores detalhes sobre o assunto.

Além disso, de acordo com Tsiatis e Davidian (2004), a caracterização do processo $w_{1i}(t)$ no modelo (2.1) geralmente é feita em termos de um vetor de $r + 1$ efeitos aleatórios específicos do i -ésimo indivíduo, considerados invariantes no tempo. Este vetor é denotado por $\mathbf{b}_i = (b_{0i}, b_{1i}, \dots, b_{ri})^T$, $r \geq 0$.

Em geral, considera-se

$$w_{1i}(t) = \mathbf{f}(t)^T \mathbf{b}_i, \quad (2.3)$$

em que $\mathbf{f}(t)$ é um vetor de funções do tempo t que pode incluir splines ². O modelo (2.3), considerado entre outros por Brown e Ibrahim (2003a) e Guo and Carlin (2004), especifica que o processo $w_{1i}(t)$ segue uma trajetória determinada por um pequeno número de efeitos aleatórios específicos do i -ésimo indivíduo. Em particular, o modelo (2.3) com

$$\mathbf{f}(t) = (1, t)^T, \quad \mathbf{b}_i = (b_{0i}, b_{1i})^T, \quad (2.4)$$

foi utilizado por DeGruttola e Tu (1994), Tsiatis, DeGruttola e Wulfsohn (1995), Faucett e Thomas (1996), Wulfsohn e Tsiatis (1997), Bycott e Taylor (1998), Dafni e Tsiatis (1998), Song, Davidian e Tsiatis (2002b), Zeng e Cai (2005), Elashoff, Li e Li (2008) e Albert e Shin (2010), com o intuito de descrever as trajetórias da resposta longitudinal de interesse, em termos de um intercepto b_{0i} e de uma inclinação b_{1i} específicos do i -ésimo indivíduo.

No contexto dos estudos de HIV, de câncer de próstata ou de ICC, por exemplo, o

²Em princípio também poderiam ser consideradas funções não-lineares de \mathbf{b}_i . Neste trabalho a atenção será restrita a funções de tipo linear.

comportamento a curto prazo das respostas longitudinais coletadas em cada caso pode ser descrito de forma razoável pela tendência $\mathbf{f}(t)^T \mathbf{b}_i$, mas a longo prazo é muito provável que os pacientes passem por períodos bons e maus conforme sua condição médica varia. Nesse caso parece ser mais plausível descrever o comportamento da resposta longitudinal por intermédio de um modelo que leve em conta tais flutuações aleatórias e considere a correlação serial intraunidades amostrais. Isto pode ser incluído em modelos da forma

$$w_{1i}(t) = \mathbf{f}(t)^T \mathbf{b}_i + u_i(t), \quad (2.5)$$

em que $u_i(t)$ é um processo estocástico com média zero, que descreve as flutuações aleatórias em torno da tendência $\mathbf{f}(t)^T \mathbf{b}_i$. Este tipo de variação aleatória, às vezes, pode resultar em uma estrutura de correlação serial entre as medidas intraunidades amostrais. O modelo (2.5) derivado das propostas de Laird e Ware (1982) e Diggle (1988), foi utilizado por Taylor, Cumberland e Sy (1994), Zeger e Diggle (1994), LaValley e DeGruttola (1996), Bycott e Taylor (1998), Zhang et al. (1998), Henderson, Diggle e Dobson (2000), Wang e Taylor (2001), Xu e Zeger (2001a), Diggle, Souza e Chetwynd (2008) e Ye, Lin e Taylor (2008b).

Segundo Tsiatis e Davidian (2004), a escolha de um modelo para o processo $w_{1i}(t)$ é em certa medida uma questão filosófica, ditada pela crença sobre os mecanismos biológicos subjacentes e pelo grau em que o modelo proposto pretende ser uma representação realista do resultado desses mecanismos. O modelo (2.1) com $w_{1i}(t)$ da forma (2.5), engloba as três fontes de variação aleatória em dados longitudinais consideradas por Diggle (1988) e Diggle et al. (2002): i) variação resultante dos efeitos aleatórios, que representam a variação estocástica entre os indivíduos, que formam uma amostra aleatória da população em estudo, ii) variação devida à correlação serial entre as medições repetidas observadas no i -ésimo indivíduo e, iii) variação proveniente de erros de medida. Diggle et al. (2002) argumentam que, em diferentes aplicações, o efeito da correlação serial é muitas vezes dominado por uma combinação dos efeitos aleatórios incorporados no vetor \mathbf{b}_i e dos erros de medição. Dependendo da forma de $w_{1i}(t)$, o erro de medida $e_i(t)$ é considerado independente de \mathbf{b}_i e de $u_i(t)$, para todo $t \geq 0$.

Adicionalmente, como discutido por Tsiatis e Davidian (2004), sob a perspectiva do modelo (2.1) e (2.5), os erros de medida $e_{ij} = \{e_i(t_{ij}), j = 1, \dots, n_i\}$ agrupados no vetor $\mathbf{e}_i = (e_{i1}, \dots, e_{in_i})^T$ representam desvios causados por erros de medição e pela variação biológica de curto prazo, de modo que podem ser considerados independentes através de j . No contexto do modelo (2.1) e (2.3), Tsiatis, DeGruttola e Wulfsohn (1995) e Song, Davidian e Tsiatis (2002ab) argumentam que se os tempos t_{ij} estão suficientemente espaçados para que a autocorrelação entre as medidas intraunidades amostrais seja praticamente de-

sprezível, ou se o erro de medição for grande em relação às flutuações biológicas, a suposição de que os erros em \mathbf{e}_i são independentes através de j ainda é razoável. Pelo contrário, se as observações são suficientemente próximas, a matriz de covariâncias do vetor de erros \mathbf{e}_i deve incluir a autocorrelação.

Em (2.3), (2.4) e (2.5) geralmente assume-se que o vetor \mathbf{b}_i considerado na caracterização do processo $w_{1i}(t)$ é normalmente distribuído. Dependendo da situação, tanto a média quanto a matriz de covariâncias de \mathbf{b}_i podem ser tomadas como funcionalmente dependentes das componentes de $\mathbf{x}_{1i}(t)$. Por exemplo, se $\mathbf{x}_{1i}(t)$ contém um indicador do tratamento, o intercepto e a inclinação em (2.4) podem ter uma média e uma matriz de covariâncias diferente para cada grupo de tratamento. No entanto, uma especificação comum é que os \mathbf{b}_i são funcionalmente independentes de $\mathbf{x}_{1i}(t)$, de modo que $\mathbf{b}_i \sim N(0, \boldsymbol{\Sigma}_1)$. Diferentes propostas que não assumem a normalidade para \mathbf{b}_i têm sido desenvolvidas, como em Zhang e Davidian (2001). Rizopoulos, Verbeke e Molenberghs (2008) mostraram que no contexto da modelagem conjunta, na medida em que o número de medições repetidas por indivíduo n_i aumenta, uma má especificação da distribuição do vetor de efeitos aleatórios tem um efeito mínimo sobre as estimativas dos parâmetros do modelo e seus respectivos erros padrão.

Em relação à especificação da distribuição do processo estocástico $u_i(t)$, Henderson et al. (2000) e Xu and Zeger (2001a) consideram $u_i(t)$ em (2.5) como um processo Gaussiano estacionário com média zero, variância σ_u^2 e função de correlação $R(s; \varpi) = \text{Corr}[u_i(s), u_i(t - s)] = \exp[-|s|^c/\varpi]$ para c fixo (com c igual a 0.5, 1 ou 2) e $\varpi > 0$. O leitor poderá consultar Verbeke, Lessafre e Brant (1998), Verbeke e Molenberghs (2000) e Diggle et al. (2002), para maiores detalhes sobre a detecção e incorporação da correlação serial em modelos para dados longitudinais.

Vários métodos de estimação para os parâmetros envolvidos no modelo (2.1) estão disponíveis na literatura estatística. Neste trabalho é considerado o *método de máxima verossimilhança restrita* (REML) (Harville, 1974 e Harville, 1977), considerando que em pequenas amostras pode produzir estimativas não enviesadas dos componentes de variância do modelo, ao contrário do que acontece com o *método de máxima verossimilhança* (ML). O leitor poderá consultar Laird e Ware (1982), Lindstrom e Bates (1988), Verbeke e Molenberghs (2000) e Diggle et al. (2002), entre outros, para maiores detalhes.

2.2 Modelagem do componente de sobrevivência

2.2.1 Abordagem semi-paramétrica

O risco de ocorrência do evento de interesse pode ser modelado por meio do modelo semi-paramétrico de Cox (1972), com função de taxa de falha para o i -ésimo indivíduo dada por

$$\lambda_i(t|\mathbf{x}_{2i}(t)) = \lambda_0(t) \exp\{\mathbf{x}_{2i}(t)^T \boldsymbol{\beta}_2\}, \quad (2.6)$$

em que $\lambda_0(t)$ denota o componente não-paramétrico, conhecido como função de risco basal, $\exp\{\mathbf{x}_{2i}(t)^T \boldsymbol{\beta}_2\}$ é o componente paramétrico, com $\boldsymbol{\beta}_2 = (\beta_{20}, \beta_{21}, \dots, \beta_{2p_2})^T$ representando um vetor que contém os efeitos fixos correspondentes às p_2 variáveis explicativas, consideradas no vetor $\mathbf{x}_{2i}(t)$. Os vetores $\mathbf{x}_{2i}(t)$ e $\boldsymbol{\beta}_2$ em (2.6) podem ter elementos em comum com $\mathbf{x}_{1i}(t)$ e $\boldsymbol{\beta}_1$ em (2.1). Detalhes sobre a estimação dos parâmetros envolvidos no modelo (2.6) podem ser encontrados em Klein e Moeschberger (1997), Kalbfleisch e Prentice (2002) e Colosimo e Giolo (2006), entre outros.

Uma das suposições dos modelos tradicionais para dados de sobrevivência, como o considerado acima, é a homogeneidade da população sob investigação com relação ao risco de ocorrência do evento de interesse. Deste modo, os modelos de sobrevivência básicos assumem que os tempos até a ocorrência do evento de interesse de indivíduos diferentes com o mesmo padrão de variáveis explicativas têm a mesma distribuição e são independentes (Wienke, 2011). No entanto, há situações em que essas suposições são inválidas, porque existem alguns fatores de risco latentes (não observáveis) que implicam que os tempos de sobrevivência de indivíduos diferentes sejam desiguais ou estejam correlacionados. No contexto médico, por exemplo, alguns indivíduos podem ter algumas características particulares ou uma predisposição genética compartilhada com os membros de sua família, que intensificam o risco de desenvolvimento de alguma doença, em comparação com outros que não as possuem. Nesse caso, um modelo muito utilizado é o modelo de fragilidade, em que a possível heterogeneidade ou a correlação entre os tempos de sobrevivência de indivíduos diferentes é incorporada através da introdução de um efeito aleatório latente (fragilidade). Tais modelos são também denominados modelos de sobrevivência com efeitos aleatórios [Ver Duchateau e Jensen (2008) e Wienke (2011), entre outros].

Se todos os fatores de risco fossem conhecidos e observáveis, idealmente eles poderiam ser incluídos na análise, usando o modelo (2.6). No entanto, muitas vezes é impossível incluir todos os fatores importantes na análise, por existirem muitas covariáveis a considerar no modelo ou, em outros casos, porque não se tem conhecimento da existência de algumas

ou, ainda, por não ser possível medir todas. Conforme Therneau e Grambsch (2000), além de explicar a heterogeneidade dos indivíduos, a fragilidade introduzida na função de risco, também serve para avaliar o efeito de covariáveis que por algum motivo não foram observadas no estudo.

Vários autores têm mostrado que ignorar a fragilidade pode gerar estimativas enviesadas dos efeitos das covariáveis, o que em alguns casos conduz à subestimação do risco de o evento acontecer [Ver Wienke (2011), por exemplo]. Entre outros autores, Henderson e Oman (1999) quantificaram esse viés.

O modelo univariado de fragilidade introduzido por Vaupel, Manton e Stallard (1979), é uma extensão do modelo de Cox, dada por

$$\begin{aligned}\lambda_i(t|\mathbf{x}_{2i}(t), s_i) &= \lambda_0(t) \exp\{\mathbf{x}_{2i}(t)^T \boldsymbol{\beta}_2 + s_i\} \\ &= l_i \lambda_0(t) \exp\{\mathbf{x}_{2i}(t)^T \boldsymbol{\beta}_2\},\end{aligned}\tag{2.7}$$

em que $\lambda_0(t)$, $\mathbf{x}_{2i}(t)$ e $\boldsymbol{\beta}_2$ têm as mesmas interpretações que no modelo (2.6) e, $l_i = \exp\{s_i\}$ representa a fragilidade, que diminui ($l_i < 1$) ou aumenta ($l_i > 1$) o risco para o i -ésimo indivíduo. As variáveis aleatórias l_i são consideradas independentes e identicamente distribuídas, com distribuição de probabilidade conhecida com média 1 e variância desconhecida, e, geralmente, não são consideradas dependentes do tempo, o que pode ser uma limitação em algumas situações [Ver Colosimo e Giolo (2006), por exemplo]. Para contornar esse problema, diferentes modelos com variáveis de fragilidade que variam com o tempo têm sido propostos, como aqueles apresentados em Yau e McGilchrist (1998).

Um problema importante neste contexto é a escolha da distribuição de probabilidade da fragilidade no modelo. A mais utilizada é a distribuição Gama (Clayton, 1978; Vaupel, Manton e Stallard, 1979; Clayton e Cuzick, 1985; Nielsen et al., 1992; Klein, 1992; Murphy, 1994; Murphy, 1995; Andersen et al., 1997), embora outros autores tenham considerado a distribuição normal inversa (Hougaard, 1984), a família de distribuições de três parâmetros proposta por Hougaard (Hougaard, 1986a), a distribuição estável positiva (Hougaard, 1986b) e, a distribuição log-normal (McGilchrist e Aisbett, 1991; McGilchrist, 1993), entre outras. Apesar de a escolha da distribuição da fragilidade ter efeito sobre as conclusões, na prática ela é feita por razões computacionais; de modo que a distribuição Gama é a mais comumente usada. O leitor interessado pode consultar comparações de diferentes distribuições da fragilidade em Hougaard (1984, 1991, 2000).

Modelos de fragilidade também podem ser utilizados em outras situações em que a suposição de independência dos tempos até a ocorrência do evento de interesse pode não ser válida. Isso ocorre, por exemplo, quando cada indivíduo em estudo está sujeito a vários even-

tos do mesmo tipo (eventos recorrentes), como ataques de asma ou infartos do miocárdio; ou quando em cada indivíduo em estudo são observados vários eventos de tipos diferentes, por exemplo, múltiplas sequelas em pacientes com doenças crônicas como a diabetes ou o câncer. Nestes casos, mais de um tempo de sobrevivência é observado em cada indivíduo e tais tempos podem estar correlacionados, o que caracteriza dados de sobrevivência multivariados (Hougaard, 2000; Therneau e Grambsch, 2000; Kalbfleisch e Prentice, 2002). Em situações como essas o uso de modelos de risco padrão, como o modelo de Cox, pode gerar estimativas enviesadas dos efeitos das covariáveis envolvidas no modelo (Wei et al., 1989).

Detalhes sobre os modelos de fragilidade no contexto de dados de sobrevivência multivariados podem ser encontrados em Clayton (1978), Petersen (1998), Vaida e Xu (2000), Ripatti e Palmgren (2000), Ripatti, Larsen e Palmgren (2002) e Colosimo e Giolo (2006). Além disso, o leitor interessado nos métodos de estimação associados, poderá consultar Therneau e Grambsch (2000), Ibrahim, Chen e Sinha (2001), Duchateau e Jensen (2008) e Wienke(2011), entre outros.

2.2.2 Abordagem paramétrica

Outra abordagem utilizada com frequência na análise de dados de sobrevivência é aquela em que são considerados modelos probabilísticos baseados em distribuições como exponencial, Weibull, log-normal, gama e Gaussiana inversa, entre outras. Embora estes modelos têm ocupado uma posição de destaque devido à sua adequação a diversos problemas práticos, existem certas situações em que o mecanismo de falha tem características específicas que requerem modelos mais flexíveis, como aquele baseado na distribuição Birnbaum-Saunders (BS). Essa distribuição foi proposta por Birnbaum e Saunders (1969a) com o intuito de descrever a variação aleatória de tempos de falha de materiais expostos à fadiga resultante de diferentes padrões cíclicos de estresse e tensão, considerando que tais tempos estão associados a algum tipo de dano cumulativo.

2.2.2.1 Distribuição Birnbaum-Saunders

A distribuição BS (Birnbaum e Saunders, 1969a) tem sido amplamente utilizada na modelagem de processos de fadiga de materiais e em análises de confiabilidade [Ver Rieck e Nedelman (1991), Achcar (1993), Owen e Padgett (1999, 2000), Tsionas (2001), Galea, Leiva e Paula (2004), Xie e Wei (2007), Desmond, Rodríguez-Yam e Lu (2008), Lemonte e Cordeiro (2009), Cancho, Ortega e Paula (2010), Lemonte, Ferrari e Cribari-Neto (2010),

Balakrishnan et al. (2011), Villegas, Paula e Leiva (2011), Desmond et al. (2012), Ortega, Cordeiro e Lemonte (2012) e, Lemonte (2013), entre outros]. Os modelos de regressão BS também têm sido aplicados com sucesso em outras áreas como indicam Johnson, Kots e Balakrishnan (1995) e Leiva, Barros e Paula (2009). Em problemas relacionados com doenças cardíacas crônicas e diversos tipos de câncer, por exemplo, um dano cumulativo causado por vários fatores de risco pode causar uma degradação que conduz a um processo de fadiga cujos tempos de propagação podem ser adequadamente modelados pela distribuição BS. Levando isso em conta, Leiva et al. (2007), Barros, Paula e Leiva (2008) e, Cancho, Louzada e Barriga (2012) utilizam a distribuição BS para modelar tempos de sobrevivência de pacientes com melanoma múltiplo, com câncer de pulmão e com melanoma, respectivamente. No contexto das ciências naturais e ambientais encontram-se as aplicações consideradas em Leiva et al. (2008a, 2011), Podlaski (2008), Leiva, Sanhueza e Angulo (2009) e Vilca et al. (2010). Para outras aplicações da distribuição BS os leitores podem consultar Balakrishnan, Leiva e López (2007), Leiva et al. (2008b) e Paula et al. (2011).

Para caracterizar a distribuição BS considera-se uma variável aleatória T que representa o tempo de falha, de modo que a função de distribuição acumulada de T é

$$F_T(t) = P(T \leq t) = \Phi(a_t(\alpha, \eta)), \quad t > 0,$$

em que $\Phi(\cdot)$ é a função de distribuição acumulada de uma variável aleatória $N(0, 1)$ e,

$$a_t(\alpha, \eta) = \frac{1}{\alpha} \left[\sqrt{\frac{t}{\eta}} - \sqrt{\frac{\eta}{t}} \right]. \quad (2.8)$$

Simbolicamente $T \sim BS(\alpha, \eta)$, com $\alpha > 0$ e $\eta > 0$ representando os parâmetros de forma e escala, respectivamente.

A função densidade de probabilidade de T é

$$f_T(t) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2\alpha^2} \left[\frac{t}{\eta} + \frac{\eta}{t} - 2 \right] \right\} \frac{t^{-3/2}[t + \eta]}{2\alpha\sqrt{\eta}}, \quad t > 0, \alpha > 0, \eta > 0,$$

alternativamente expressa como

$$f_T(t) = \phi(a_t(\alpha, \eta))A_t(\alpha, \eta), \quad t > 0, \alpha > 0, \eta > 0,$$

em que $\phi(\cdot)$ é a função densidade de probabilidade de uma variável aleatória $N(0, 1)$ e

$$A_t(\alpha, \eta) = \frac{d}{dt}a_t(\alpha, \eta) = \frac{t^{-3/2}[t + \eta]}{2\alpha\sqrt{\eta}},$$

com $a_t(\alpha, \eta)$ dada em (2.8). A correspondente função de sobrevivência é

$$S_T(t) = P(T > t) = 1 - F_T(t) = 1 - \Phi(a_t(\alpha, \eta)) = \Phi(-a_t(\alpha, \eta)), \quad t > 0.$$

Além disso, a variável T com distribuição BS está relacionada com a distribuição normal padrão mediante

$$T = \eta \left[\frac{\alpha Z}{2} + \sqrt{\left(\frac{\alpha Z}{2}\right)^2 + 1} \right]^2,$$

em que Z é uma variável aleatória tal que

$$Z = \frac{1}{\alpha} \left[\sqrt{\frac{T}{\eta}} - \sqrt{\frac{\eta}{T}} \right] \sim N(0, 1), \quad T > 0.$$

Birnbaum e Saunders (1969b) obtiveram os estimadores de máxima verossimilhança de α e η . Posteriormente, Engelhardt, Bain e Wright (1981) propuseram intervalos de confiança e testes de hipóteses para os parâmetros da distribuição. Desmond (1985) reforçou a justificativa física para o uso desta distribuição, flexibilizando algumas das suposições feitas por Birnbaum e Saunders (1969a), além de derivar uma forma mais geral baseada em um modelo biológico. Desmond (1986) estabeleceu uma conexão entre a distribuição Gaussiana inversa e a distribuição BS. Rieck e Nedelman (1991) propuseram um modelo log-linear para a distribuição BS e desenvolveram métodos de estimação dos parâmetros de interesse. Resultados inferenciais posteriores foram obtidos por Achcar e Espinosa (1992), Achcar (1993), Lu e Chang (1997), Dupuis e Mills (1998) e Rieck (1999). Algumas generalizações e extensões da distribuição BS podem ser consultadas em Owen e Padgett (1999), Díaz-García e Leiva (2005), Owen (2006), Vilca-Labra e Leiva (2006), Sanhueza, Leiva e Balakrishnan (2008), Balakrishnan et al. (2009), Gomez, Pacheco e Bolfarine (2009), Guiraud, Leiva e Fierro (2009), Cordeiro e Lemonte (2011), Villegas, Paula e Leiva (2011), Desmond et al. (2012), Cancho, Louzada e Barriga (2012), Ortega, Cordeiro e Lemonte (2012) e, Lemonte (2013). Outros trabalhos recentes sobre a distribuição BS podem ser encontrados em From e Li (2006), Ng, Kundu e Balakrishnan (2006), Balakrishnan, Leiva e Lopez (2007), Lemonte, Cribari-Neto e Vasconcellos (2007), Cysneiros, Cribari-Neto e Araújo (2008), Kundu, Kannan e Balakrishnan (2008), Leiva et al. (2008c), Lemonte, Simas e Cribari-Neto (2008), Lemonte e Cordeiro (2009), Vilca et al. (2011) e Santos-Neto et al. (2012), entre outros.

2.2.2.2 Modelo de regressão log-linear Birnbaum-Saunders

Rieck e Nedelman (1991) consideraram o modelo de regressão $T = \exp\{X\beta\}\epsilon$, em que $T \sim BS(\alpha, \eta)$ e analisaram o modelo log-linear associado dado por $V = \log(T) = X\beta + \log(\epsilon) = X\beta + \epsilon^*$, com $\epsilon^* = \log(\epsilon)$, em que a distribuição de V é um caso particular da distribuição seno hiperbólico normal (SenhN) descrita a seguir.

A distribuição SenhN também é definida através de uma transformação da distribuição normal padrão por meio da variável aleatória

$$V = \operatorname{arcsenh}\left(\frac{\alpha Z}{2}\right)\sigma + \psi,$$

em que Z é uma variável aleatória dada por

$$Z = \frac{2}{\alpha} \operatorname{senh}\left(\frac{V - \psi}{\sigma}\right) \sim N(0, 1),$$

e $\alpha > 0$, $\psi \in \Re$ e $\sigma > 0$ representam os parâmetros de forma, localização e escala, respectivamente.

Pode-se dizer que $V \sim \operatorname{SenhN}(\alpha, \psi, \sigma)$ com função densidade de probabilidade da forma

$$f_V(v) = \left(\frac{2}{\alpha\sigma\sqrt{2\pi}}\right) \cosh\left(\frac{v - \psi}{\sigma}\right) \exp\left\{-\frac{2}{\alpha^2} \operatorname{senh}^2\left(\frac{v - \psi}{\sigma}\right)\right\}, \quad v \in \Re, \quad (2.9)$$

e funções de distribuição acumulada e de sobrevivência dadas, respectivamente, por

$$F_V(v) = P(V \leq v) = \Phi\left(\frac{2}{\alpha} \operatorname{senh}\left[\frac{v - \psi}{\sigma}\right]\right), \quad v \in \Re,$$

e,

$$S_V(v) = P(V > v) = 1 - F_V(v) = 1 - \Phi\left(\frac{2}{\alpha} \operatorname{senh}\left[\frac{v - \psi}{\sigma}\right]\right), \quad v \in \Re. \quad (2.10)$$

Rieck e Nedelman (1991) desenvolveram a distribuição SenhN e provaram que se $T \sim BS(\alpha, \eta)$, então, $V = \log(T) \sim \operatorname{SenhN}(\alpha, \psi, \sigma = 2)$, em que $\psi = \log(\eta)$. Por este motivo, a distribuição SenhN aparece naturalmente na formulação do modelo de regressão log-linear Birnbaum-Saunders, a qual é feita considerando algumas suposições sobre os parâmetros α e η da seguinte forma:

Sejam T_1, T_2, \dots, T_n variáveis aleatórias independentes tais que $T_i \sim BS(\alpha_i, \eta_i)$, para $i = 1, \dots, n$. Supondo que a distribuição de cada T_i dependa de um conjunto de p variáveis explicativas ($p < n$) consideradas no vetor $\mathbf{x}_i = (1, x_{i1}, \dots, x_{ip})^T$ e assumindo que:

- i) $\eta_i = \exp\{\mathbf{x}_i^T \boldsymbol{\beta}\}$, em que $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_p)^T$ é um vetor de parâmetros desconhecidos e,
- ii) o parâmetro de forma é independente das variáveis explicativas, isto é, $\alpha_i = \alpha$ para todo $i = 1, \dots, n$,

então, $T_i \sim BS(\alpha, \exp\{\mathbf{x}_i^T \boldsymbol{\beta}\})$.

Como a distribuição BS é fechada sob transformações de escala, isto é, se $T \sim BS(\alpha, \eta)$, $cT \sim BS(\alpha, c\eta)$ para $c > 0$, e considerando que $\epsilon_i \sim BS(\alpha, 1)$, então,

$$T_i = \exp\{\mathbf{x}_i^T \boldsymbol{\beta}\} \epsilon_i \sim BS(\alpha, \exp\{\mathbf{x}_i^T \boldsymbol{\beta}\}).$$

Portanto, fazendo $V_i = \log(T_i)$ obtem-se o modelo de regressão log-linear Birnbaum-Saunders

$$V_i = \mathbf{x}_i^T \boldsymbol{\beta} + \epsilon_i^*, \tag{2.11}$$

em que $\boldsymbol{\beta}$ e \mathbf{x}_i são definidos como anteriormente e, as $\epsilon_i^* = \log(\epsilon_i)$ são variáveis aleatórias independentes e identicamente distribuídas, tais que $\epsilon_i^* \sim \text{SenhN}(\alpha, 0, 2)$ para $i = 1, \dots, n$.

Detalhes sobre a obtenção dos estimadores de máxima verossimilhança dos parâmetros em $\boldsymbol{\beta}$ e α no modelo (2.11) podem ser consultados em Galea, Leiva e Paula (2004) para o caso de observações não-censuradas e em Leiva et al. (2007) para o caso de observações censuradas.

2.3 Modelagem conjunta de dados longitudinais e dados de sobrevivência

Muitas vezes a resposta longitudinal e o tempo de sobrevivência estão relacionados por meio de algumas variáveis latentes. Neste caso, modelar os dados separadamente pode produzir estimativas enviesadas e/ou menos eficientes para os parâmetros. Uma modelagem conjunta, em que as potenciais relações subjacentes entre os processos das respostas longitudinal e de sobrevivência sejam explicitamente reconhecidas, pode ser mais adequada. De acordo com Tsiatis e Davidian (2004) e Yu et al. (2004), diferentes autores têm mostrado que a utilização de modelos conjuntos neste contexto leva à correção do viés e ao potencial aumento

da eficiência na estimação dos parâmetros.

Rizopoulos (2010) indica que existem basicamente dois tipos de objetivos atingíveis por meio de uma modelagem conjunta: i) analisar o comportamento da resposta longitudinal, considerando uma eventual dependência do tempo até o abandono do estudo, potencialmente informativo (quando o interesse está na resposta longitudinal) e, ii) analisar o comportamento do tempo até a ocorrência do evento, condicionalmente às medições da resposta longitudinal (quando o interesse está na resposta de sobrevivência).

Existem situações em que a resposta longitudinal está relacionada com o abandono do estudo, potencialmente informativo, no sentido descrito por Rubin (1976), isto é, situações em que o abandono pode ser preditivo da resposta longitudinal que teria sido observada em caso contrário. Conforme sugerem Wu e Carroll (1988), o fato de não se considerar este fenômeno pode gerar estimativas viesadas dos parâmetros que descrevem o comportamento da resposta longitudinal. Em estudos de HIV, câncer ou ICC, por exemplo, o perfil médio da resposta longitudinal observada em cada caso pode tanto refletir a verdadeira evolução desta resposta ao longo do tempo ou ser um resultado do abandono seletivo dos pacientes com valores desfavoráveis. Nesses casos, é mais provável que pacientes com essas características apresentem o evento de interesse ou se retirem do estudo antes que indivíduos saudáveis. Aqui, o abandono prematuro do estudo fornece informação que deve ser levada em conta na modelagem da resposta longitudinal e uma das opções é considerar uma modelagem conjunta da resposta longitudinal de interesse e do tempo até o abandono do estudo, tratado como a resposta de sobrevivência. Dentre os diferentes autores que têm-se focado nesse problema, destacam-se Wu e Carroll (1988), Schluchter (1992), Little (1995), Hogan e Laird (1997ab), Dobson e Henderson (2003), Hogan, Lin e Herman (2004), Vonesh, Greene e Schluchter (2006) e Diggle, Farewell e Henderson (2007).

Por outro lado, em estudos com o objetivo de identificar fatores prognósticos ou de analisar o efeito de algum tratamento sobre o tempo de falha, condicionalmente às medições da resposta longitudinal, geralmente considera-se o modelo de Cox com função de taxa de falha para o i -ésimo indivíduo dada por

$$\lambda_i(t|M_i(t), \mathbf{x}_{2i}(t)) = \lambda_0(t) \exp\{\mathbf{x}_{2i}(t)^T \boldsymbol{\beta}_2 + \gamma m_i(t)\}, \quad (2.12)$$

em que $M_i(t) = \{m_i(u), 0 \leq u \leq t\}$ representa a história do verdadeiro processo longitudinal (não observável) até o tempo t , γ é o parâmetro que mede o efeito do verdadeiro valor da resposta longitudinal no instante t ($m_i(t)$) no risco, ou em outras palavras, γ é o parâmetro que mede a associação entre os processos das respostas longitudinal e de sobrevivência, e, $\lambda_0(t)$, $\mathbf{x}_{2i}(t)$ e $\boldsymbol{\beta}_2$ são definidos como em (2.6).

Uma abordagem ingênua para o ajuste do modelo (2.12) foi considerar a resposta longitudinal como uma covariável dependente do tempo e maximizar a verossimilhança parcial com base nos valores observados dessa resposta. No entanto, a formulação do modelo (2.12) exige que os valores das covariáveis dependentes do tempo sejam conhecidos exatamente em cada instante de falha observado e, além disso, que tais covariáveis sejam externas no sentido descrito por Kalbfleisch and Prentice (2002), isto é, que o seu valor no instante t não deve ser afetado pela ocorrência do evento no instante de tempo s , com $s < t$. Na prática, a implementação desta abordagem é complicada pelos seguintes motivos:

- i) Em geral, a resposta longitudinal somente está disponível intermitentemente nos instantes $t_{ij} \leq Z_i$, $i = 1, \dots, n$, $j = 1, \dots, n_i$, que podem ser diferentes para cada indivíduo, mesmo que os momentos de observação tenham sido especificados pelo protocolo do estudo. Além disso, os valores observados da resposta longitudinal comumente são medidos com erro. É possível que ao invés dos verdadeiros valores $m_i(t)$, apenas os valores $y_i(t)$ da forma dada em (2.1) possam ser observados.
- ii) Geralmente a resposta longitudinal não pode ser considerada como uma covariável dependente do tempo externa por ser resultado de um processo estocástico gerado pelo indivíduo e diretamente relacionado com o mecanismo de falha; em outras palavras, a resposta longitudinal está frequentemente relacionada com o tempo de sobrevivência.

As opções iniciais para contornar a primeira de tais dificuldades envolveram algum tipo de imputação ingênua dos valores requeridos da resposta longitudinal, utilizando, por exemplo, o *método da última observação levada adiante* (*last value carried forward - LVCF*), no qual, o verdadeiro valor não observado da resposta longitudinal para o i -ésimo indivíduo, é substituído pelo valor mais recente. Prentice (1982) mostrou que essa substituição gera estimativas enviesadas para os parâmetros do modelo. Posteriormente, Raboud et al. (1993) trabalharam na análise do viés potencial devido ao uso da técnica LVCF e mostraram que métodos simples de suavização para imputar o valor da resposta longitudinal em (2.12) reduzem o viés em relação à abordagem ingênua.

Além disso, o fato de os erros de medição na resposta longitudinal não serem devidamente considerados no ajuste do risco em (2.12) pode levar a estimativas enviesadas dos parâmetros que descrevem a relação entre essa resposta (considerada como uma covariável dependente do tempo) e o risco de falha (Prentice, 1982). Assim, para corrigir o possível viés nas estimativas dos parâmetros do modelo quando as covariáveis estão medidas com erro, diferentes alternativas foram propostas por Nakamura (1992), Raboud et al. (1993), Hu, Tsiatis e Davidian (1998), Kong e Gu (1999), Hu e Lin (2002), Augustin (2004) e Greene e Cai (2004), entre outros.

Com relação à segunda limitação, diferentes estudos têm mostrado a existência de uma associação importante entre as respostas longitudinais observadas em cada caso e o tempo de falha [Ver Guo e Carlin (2004) e Rizopoulos (2010), entre outros]. Em particular, se o evento analisado é o óbito ou o abandono do estudo devido à falta de uma resposta favorável, é impossível medir a resposta longitudinal no instante em que aconteceu ou após a sua ocorrência, de modo que, nestes casos, a resposta longitudinal não pode ser considerada como uma covariável dependente do tempo externa.

Outros autores consideraram uma abordagem alternativa, na qual a estimação dos parâmetros é feita em duas etapas. Na primeira, escolhe-se um modelo para descrever o comportamento do processo longitudinal, de modo que em cada instante em que o evento ocorreu, obtêm-se estimativas dos valores omissos da resposta longitudinal com base nos dados de todos os indivíduos que ainda estão em risco. Na segunda etapa, os valores modelados na primeira são tratados como os verdadeiros valores da resposta longitudinal, considerada como covariável dependente do tempo no modelo (2.12), e são substituídos na verossimilhança parcial de Cox em cada instante de falha. Essa abordagem foi considerada por Tsiatis, DeGruttola e Wulfsohn (1995), que, na primeira etapa, propuseram um modelo de efeitos mistos padrão para descrever a resposta longitudinal. Em cada instante em que o evento ocorreu, os efeitos aleatórios específicos do i -ésimo indivíduo são estimados usando os estimadores de Bayes empíricos. Posteriormente, Bycott e Taylor (1998) consideraram duas técnicas de suavização para obter estimativas da resposta longitudinal em cada tempo de falha na primeira etapa. Dafni e Tsiatis (1998) investigaram o desempenho da proposta de Tsiatis, DeGruttola e Wulfsohn (1995) por intermédio de simulações e mostraram que esse método aproximado reduz, mas não elimina completamente, o viés nas estimativas dos parâmetros em relação aos métodos ingênuos, confirmando os resultados de Wulfson e Tsiatis (1997).

Outra limitação da abordagem em duas etapas, frequentemente discutida na literatura, é que a informação não é usada de forma tão eficiente quanto poderia, pois a informação proveniente do componente de sobrevivência não é considerada na modelagem do processo longitudinal. No entanto, essa abordagem não requer técnicas de estimação complexas e é facilmente implementada computacionalmente, através da utilização de *software estatístico padrão*, como considerado em Ye, Lin e Taylor (2008b) e Albert e Shin (2010).

Diferentes autores consideraram uma abordagem baseada na especificação de uma verossimilhança correspondente à distribuição conjunta das respostas longitudinal e de sobrevivência. A principal vantagem desta abordagem é que utiliza a informação disponível de uma forma ótima no processo de estimação dos parâmetros do modelo, na medida em que os dados longitudinais e de sobrevivência são usados simultaneamente.

Neste caso, o vetor de efeitos aleatórios invariantes no tempo associados ao i -ésimo in-

divíduo, nomeadamente \mathbf{b}_i ($i = 1, \dots, n$), é incluído no modelo para contemplar tanto a associação entre as respostas longitudinal e de sobrevivência, quanto a correlação entre as medidas repetidas da resposta longitudinal. Em geral, assume-se que as respostas longitudinal e de sobrevivência do i -ésimo indivíduo são independentes condicionalmente ao seu vetor \mathbf{b}_i , formalmente

$$p(Z_i, \delta_i, \mathbf{y}_i | \mathbf{b}_i; \boldsymbol{\theta}) = p(Z_i, \delta_i | \mathbf{b}_i; \boldsymbol{\theta})p(\mathbf{y}_i | \mathbf{b}_i; \boldsymbol{\theta}), \quad (2.13)$$

e, além disso, supõe-se independência entre as medições longitudinais dado \mathbf{b}_i , isto é

$$p(\mathbf{y}_i | \mathbf{b}_i; \boldsymbol{\theta}) = \prod_{j=1}^{n_i} p(y_i(t_{ij}) | \mathbf{b}_i; \boldsymbol{\theta}), \quad (2.14)$$

em que $\boldsymbol{\theta} = (\boldsymbol{\theta}_z^T, \boldsymbol{\theta}_y^T, \boldsymbol{\theta}_b^T)^T$ com $\boldsymbol{\theta}_z$, $\boldsymbol{\theta}_y$ e $\boldsymbol{\theta}_b$ indicando os vetores dos parâmetros associados à resposta de sobrevivência, à resposta longitudinal e à distribuição dos efeitos aleatórios, respectivamente.

Sob as suposições de independência condicional (2.13) e (2.14) e considerando que a censura e os tempos de observação da resposta longitudinal são não informativos [Ver Rizopoulos (2012b) para maiores detalhes], a contribuição do i -ésimo indivíduo à verossimilhança conjunta é

$$\begin{aligned} p(Z_i, \delta_i, \mathbf{y}_i; \boldsymbol{\theta}) &= \int p(Z_i, \delta_i, \mathbf{y}_i, \mathbf{b}_i; \boldsymbol{\theta}) d\mathbf{b}_i \\ &= \int p(Z_i, \delta_i | \mathbf{b}_i; \boldsymbol{\theta}_z, \boldsymbol{\beta}_1) \left[\prod_{j=1}^{n_i} p(y_i(t_{ij}) | \mathbf{b}_i; \boldsymbol{\theta}_y) \right] p(\mathbf{b}_i; \boldsymbol{\theta}_b) d\mathbf{b}_i, \end{aligned} \quad (2.15)$$

em que, sob independência entre os tempos de sobrevivência e de censura

$$p(Z_i, \delta_i | \mathbf{b}_i; \boldsymbol{\theta}_z, \boldsymbol{\beta}_1) = [f_i(Z_i | M_i(Z_i); \boldsymbol{\theta}_z, \boldsymbol{\beta}_1)]^{\delta_i} [S_i(Z_i | M_i(Z_i); \boldsymbol{\theta}_z, \boldsymbol{\beta}_1)]^{(1-\delta_i)}, \quad (2.16)$$

com $f_i(\cdot)$ e $S_i(\cdot)$ denotando as funções densidade de probabilidade e de sobrevivência, respectivamente, cujas formas dependem da abordagem (semi-paramétrica ou paramétrica) escolhida para modelar o componente de sobrevivência, e $M_i(\cdot)$ é definida como em (2.12). Além disso, considerando que a resposta longitudinal é modelada como em (2.1) e sob normalidade, tem-se

$$p(y_i(t_{ij}) | \mathbf{b}_i; \boldsymbol{\theta}_y) = \frac{1}{(2\pi\sigma_e^2)^{1/2}} \exp \left\{ -\frac{(y_i(t_{ij}) - m_i(t_{ij}))^2}{2\sigma_e^2} \right\}, \quad (2.17)$$

e,

$$p(\mathbf{b}_i; \boldsymbol{\theta}_b) = \frac{1}{(2\pi|\Sigma_1|)^{1/2}} \exp\left\{-\frac{\mathbf{b}_i^T \Sigma_1^{-1} \mathbf{b}_i}{2}\right\}. \quad (2.18)$$

Uma característica desta abordagem é que a obtenção das estimativas de máxima verossimilhança dos parâmetros do modelo é um processo geralmente exigente em termos computacionais, pois além de necessitar o uso de algoritmos de otimização numérica, exceto em alguns casos especiais, as integrais envolvidas não têm solução analítica. Por isso, esse processo é realizado por meio da maximização direta da função de log-verossimilhança conjunta usando algoritmos numéricos padrão como o algoritmo de Newton-Raphson ou alguma de suas variações (Lange, 2004). Outra alternativa é a utilização de um algoritmo EM (Dempster, Laird e Rubin, 1977), considerando os efeitos aleatórios como dados omissos. O passo E do algoritmo envolve uma integral analiticamente intratável, de modo que a aproximação da esperança condicional requer a utilização de métodos de integração numérica, enquanto que, no passo M, o método usualmente empregado para a maximização é o de Newton-Raphson. A principal motivação para o uso do algoritmo EM no contexto da modelagem conjunta é que as atualizações no passo M têm forma fechada para certos parâmetros do modelo [Ver Wulfsohn e Tsiatis (1997) e Henderson, Diggle e Dobson (2000)]. Uma desvantagem é que, geralmente, o método tem uma convergência lenta em torno do máximo. Considerando que uma maximização direta da log-verossimilhança dos dados longitudinais e de sobrevivência utilizando um algoritmo quase-Newton requer computação muito semelhante à requerida pelo algoritmo EM, Rizopoulos, Verbeke e Lesaffre (2009) sugerem que para estimar os parâmetros de interesse poderia ser usada uma mistura de métodos de otimização que comece com um algoritmo EM e continue com uma maximização direta da função em questão. Além disso, a aproximação das integrais requeridas em cada caso é feita por meio de diferentes técnicas de integração numérica, tais como a quadratura de Gauss-Hermite ou os métodos Monte Carlo [Ver Song, Davidian e Tsiatis (2002b)]. Outra opção discutida por Ye, Lin e Taylor (2008a) e Rizopoulos, Verbeke e Lesaffre (2009) é a utilização de aproximações de Laplace, que são especialmente úteis nos casos em que é necessário estimar um grande número de efeitos aleatórios. Mais recentemente, Rizopoulos (2012a) propôs uma aproximação no âmbito dos modelos conjuntos por meio de uma quadratura pseudo-adaptativa, cujos benefícios estão descritos em Rizopoulos (2012b).

Schluchter (1992), Pawitan e Self (1993) e DeGruttola e Tu (1994) consideraram modelos conjuntos nos quais o processo de sobrevivência é descrito de forma paramétrica. No primeiro e no último destes trabalhos, os autores desenvolveram um algoritmo EM com base na especificação de uma distribuição Normal Multivariada para os dados longitudinais e de sobrevivência, enquanto que no segundo, foram consideradas distribuições normal

e Weibull, respectivamente. Wulfson e Tsiatis (1997) propuseram um algoritmo EM para maximizar a log-verossimilhança conjunta, baseada no modelo de Cox, sob a suposição de que os efeitos aleatórios são normalmente distribuídos e constantes ao longo do tempo. Esta última proposta tem sido considerada por diversos autores como uma das mais satisfatórias para combinar informações dos processos das respostas longitudinal e de sobrevivência, entre os quais destacam-se Hsieh, Tseng e Wang (2006), que verificaram a robustez do modelo quanto à violação da normalidade dos efeitos aleatórios.

Henderson, Diggle e Dobson (2000) propuseram um modelo flexível, que acomoda uma ampla gama de formas da dependência entre as respostas longitudinal e de sobrevivência. A idéia chave é unir as respostas longitudinal e de sobrevivência por meio de um processo Gaussiano bivariado latente $w_i(t) = \{w_{1i}(t), w_{2i}(t)\}$ com média zero, que se realiza de forma independente para indivíduos diferentes. Os processos para as respostas longitudinal e de sobrevivência são considerados independentes, condicionalmente ao processo latente de vinculação $w_i(t)$. Desta forma, com a finalidade de considerar a associação latente entre os processos, o risco de o evento de interesse acontecer é descrito por meio do modelo de Cox, com função de taxa de falha para o i -ésimo indivíduo dada por

$$\lambda_i(t|w_{2i}(t), \mathbf{x}_{2i}(t)) = \lambda_0(t) \exp\{\mathbf{x}_{2i}(t)^T \boldsymbol{\beta}_2 + w_{2i}(t)\}, \quad (2.19)$$

em que $w_{2i}(t)$ é um processo que envolve as variáveis latentes comuns às respostas longitudinal e de sobrevivência e, em alguns casos inclui um termo de fragilidade e $\lambda_0(t)$ e $\mathbf{x}_{2i}(t)$ são definidos como em (2.6). A associação entre os processos das respostas longitudinal e de sobrevivência é resultante da dependência entre os processos $w_{1i}(t)$ no modelo (2.1) e $w_{2i}(t)$ no modelo (2.19). Por exemplo, se o processo $w_{1i}(t)$ no modelo (2.1) é caracterizado em termos de (2.4), o processo $w_{2i}(t)$ em (2.19) pode ser modelado como

$$w_{2i}(t) = \gamma w_{1i}(t) + s_i = \gamma(b_{0i} + b_{1i}t) + s_i, \quad (2.20)$$

em que $\mathbf{b}_i = (b_{0i}, b_{1i})^T \sim N_2(0, \boldsymbol{\Sigma}_1)$, $s_i \sim N(0, \sigma_s^2)$ é um termo de fragilidade independente de \mathbf{b}_i e, γ é o parâmetro que mede a associação entre os dois submodelos. Na ausência de associação entre os processos, o modelo (2.19) com $w_{2i}(t)$ da forma dada em (2.20) reduz-se ao modelo de fragilidade univariada (2.7).

Tsiatis e Davidian (2001) motivados pelo método do score condicional de Stefanski e Carroll (1987) desenvolveram um conjunto de equações de estimação que sob certas condições de regularidade geram estimadores consistentes e assintoticamente normais para os parâmetros de um modelo conjunto no qual não se faz nenhuma suposição sobre a distribuição dos efeitos aleatórios subjacentes, \mathbf{b}_i . Esta proposta foi generalizada para o caso de mais de

uma resposta longitudinal por Song, Davidian e Tsiatis (2002a). Song, Davidian e Tsiatis (2002b) propuseram uma alternativa com base no modelo de Wulfsohn e Tsiatis (1997), na qual a suposição usual de normalidade do vetor \mathbf{b}_i é relaxada, exigindo somente que a sua densidade faça parte da classe de densidades suaves estudadas por Gallant e Nychka (1987).

Lin et al. (2002) analisaram um modelo conjunto no qual a dependência entre os processos longitudinal e de sobrevivência baseada em efeitos aleatórios comuns é substituída por uma dependência compartilhada por meio de uma variável de classe latente, em que cada classe representa uma subpopulação que tem seu próprio padrão de respostas longitudinal e de sobrevivência. Ratcliffe, Guo e Ten Have (2004) propuseram um modelo conjunto para dados agrupados, no qual as respostas longitudinal e de sobrevivência estão ligadas através das fragilidades comuns aos membros de cada grupo. Zeng e Cai (2005a) derivaram alguns resultados assintóticos para os estimadores de máxima verossimilhança obtidos na análise conjunta. Wang (2006), considerando o método do escore corrigido de Nakamura (1992), propôs um método para estimar os parâmetros de um modelo conjunto que não requer nenhuma suposição sobre a distribuição do vetor \mathbf{b}_i e, adicionalmente, gera estimadores consistentes e assintoticamente normais sob certas condições de regularidade.

Faucett e Thomas (1996) adotaram uma abordagem bayesiana em que o *método Monte Carlo baseado em cadeias de Markov* (MCMC) é utilizado para estimar a distribuição conjunta a posteriori dos parâmetros desconhecidos envolvidos nos processos longitudinal e de sobrevivência. Faucett, Schenker e Elashoff (1998) consideraram uma abordagem semelhante no caso especial em que se tem uma resposta longitudinal binária.

Xu e Zeger (2001a) adotaram a abordagem considerada por Faucett e Thomas (1996) no caso em que o interesse é identificar se a resposta longitudinal pode ser considerada como um marcador substituto do tempo de sobrevivência antes da ocorrência do evento, no sentido considerado por Prentice (1989). Esta abordagem foi estendida por Xu e Zeger (2001b), que consideraram situações nas quais existem múltiplas variáveis potencialmente substitutas do tempo de falha. Wang e Taylor (2001) desenvolveram um algoritmo MCMC para ajustar o modelo conjunto, que incorpora um processo Ornstein-Unlenbeck Integrado para monitorar as flutuações biológicas sobre a tendência longitudinal. Brown e Ibrahim (2003a) propuseram um modelo bayesiano semiparamétrico que utiliza processos Dirichlet a priori para a distribuição dos parâmetros que definem o modelo da resposta longitudinal, que posteriormente foi considerado por Rizopoulos e Ghosh (2011) para realizar a modelagem conjunta, em casos em que são observadas várias respostas longitudinais.

Guo e Carlin (2004) desenvolveram uma versão bayesiana da proposta de Henderson, Diggle e Dobson (2000), usando métodos MCMC. Ibrahim, Chen e Sinha (2004) propuseram uma alternativa bayesiana para a modelagem conjunta, na qual o modelo para dados longi-

tuais é estendido para o caso multivariado. Faucett, Schenker e Taylor (2002) utilizaram modelos conjuntos para fazer inferências sobre a distribuição marginal do tempo de falha, incorporando a resposta longitudinal como informação auxiliar para recuperar informações de observações censuradas por meio de um processo de imputação múltipla.

Law, Taylor e Sandler (2002), Brown e Ibrahim (2003b), Yu et al. (2004), Chen, Ibrahim e Sinha (2004), Chi e Ibrahim (2006), Yu, Taylor e Sandler (2008), Abu Bakar et al. (2009) e Zhang, Müller e Do (2010) desenvolveram modelos conjuntos com fração de cura na população estudada. Em particular, Chen, Ibrahim e Sinha (2004) apresentaram uma generalização para o caso de dados longitudinais multivariados e Chi e Ibrahim (2006) trabalharam o caso de dados longitudinais e de sobrevivência multivariados.

Tseng, Hsieh e Wang (2005) utilizaram modelos de tempos de vida acelerada, nos quais os parâmetros desconhecidos são estimados pela maximização da probabilidade conjunta de dados longitudinais e de sobrevivência utilizando o *algoritmo EM Monte Carlo* (MCEM). Zeng e Cai (2005b) propuseram uma abordagem para a modelagem conjunta de algumas medidas de qualidade de vida e tempos de sobrevivência, na qual a estimativa dos parâmetros é realizada através do algoritmo EM. Hsieh, Tseng e Wang (2006) levando em conta que no contexto da modelagem conjunta a partir de uma perspectiva semiparamétrica, o fato de não especificar a função de risco basal pode levar a subestimação de erros padrão dos parâmetros de interesse, propuseram a utilização de um método bootstrap para estimá-los.

Outras propostas recentes para a modelagem conjunta são as de Ding e Wang (2008), que ao invés de utilizar um modelo linear de efeitos mistos paramétrico para o processo longitudinal, consideraram um modelo não paramétrico com efeitos aleatórios multiplicativos e estimaram os parâmetros desconhecidos por meio do algoritmo MCEM. Diggle, Sousa e Chetwynd (2008) estudaram um modelo completamente paramétrico, que assume distribuições normal e lognormal para os processos das respostas longitudinal e de sobrevivência, respectivamente.

Song e Wang (2008) consideraram um modelo de Cox com coeficientes variando ao longo do tempo para o processo da resposta de sobrevivência e propuseram dois estimadores para os parâmetros de interesse com base no método de verossimilhança parcial local de Cai e Sun (2003), que, sob certas condições de regularidade, são consistentes e assintoticamente normais. Elashoff, Li e Li (2007), Elashoff, Li e Li (2008), Williamson et al. (2008), Hu, Li e Li (2009), Huang, Li e Elashoff (2010), Li et al. (2010) e Huang et al. (2011) desenvolveram modelos conjuntos considerando riscos competitivos no processo de sobrevivência. Em especial, Li et al. (2010) consideraram uma resposta longitudinal ordinal.

Wu, Hu e Wu (2008) e Wu, Liu e Hu (2010) desenvolveram metodologias para lidar com problemas de dados omissos no contexto da modelagem conjunta, utilizando algoritmos

MCEM e aproximações de Laplace, respectivamente. Rizopoulos et al. (2008) forneceram uma alternativa para modelar conjuntamente dados de sobrevivência e dados longitudinais binários com excesso de zeros. Liu e Huang (2009) consideraram um modelo conjunto para situações em que são observados eventos recorrentes e um evento terminal nos indivíduos em estudo. Li, Hu e Greene (2009), motivados pelo método do escore corrigido de Nakamura (1990), propuseram um modelo conjunto em que não é feita suposição alguma sobre a distribuição do vetor \mathbf{b}_i e sob certas condições de regularidade conduz a estimadores consistentes e assintoticamente normais. Tapsoba, Lee e Wang (2011) considerando que os perfis individuais do processo longitudinal podem apresentar pontos de mudança em algumas situações, incorporaram uma função de transição suave para o modelo dos dados longitudinais e desenvolveram estimadores do escore corrigido e do escore condicional que não requerem nem suposições sobre a distribuição dos efeitos aleatórios nem sobre aquela dos pontos de mudança.

Rizopoulos, Verbeke e Molenberghs (2010), desenvolveram algumas ferramentas de diagnóstico para modelos conjuntos de dados longitudinais e de sobrevivência, considerando que a principal dificuldade do uso de métodos diagnósticos tradicionais nos modelos conjuntos é o abandono do estudo, potencialmente informativo. Garre et al. (2008), Proust-Lima e Taylor (2009), Hanson, Branscum e Johnson (2011), Rizopoulos (2011) e Njeru-Njagi et al. (2013) consideraram metodologias para realizar a predição das probabilidades de sobrevivência dos indivíduos analisados com base no ajuste de um modelo conjunto de dados longitudinais e de sobrevivência.

Rizopoulos (2012b) forneceu uma detalhada e cuidadosa revisão da teoria e das aplicações dos modelos conjuntos, quando o uso de efeitos aleatórios é a abordagem escolhida para captar a associação entre as respostas longitudinal e de sobrevivência.

Com relação ao software para a modelagem conjunta, Guo e Carlin (2004) consideraram a implementação de modelos conjuntos por meio do uso dos programas SAS e WinBUGS, Rizopoulos (2010) propôs recentemente o pacote JM desenvolvido no ambiente R [Ver Rizopoulos (2012c) para maiores detalhes] e, Philipson et al. (2012) propuseram o pacote `joiner` desenvolvido no ambiente R. Tais pacotes são muito flexíveis e acomodam diferentes formas da dependência entre os processos longitudinal e de sobrevivência em situações em que o último desses processos é descrito utilizando o modelo de Cox.

Capítulo 3

Modelagem conjunta de dados longitudinais normais e tempos de sobrevivência Birnbaum-Saunders

Conforme discutido no Capítulo 2, no contexto da modelagem de dados longitudinais e dados de sobrevivência, a informação observada no i -ésimo indivíduo ($i = 1, \dots, n$) é: i) um vetor com observações das respostas longitudinal e de sobrevivência $(\mathbf{y}_i^T, Z_i, \delta_i)^T$, com $\mathbf{y}_i = (y_{i1}, \dots, y_{in_i})^T$, $Z_i = \min(T_i, C_i)$ e $\delta_i = I(T_i \leq C_i)$; ii) um vetor com os tempos em que a resposta longitudinal foi obtida $\mathbf{t}_i = (t_{i1}, \dots, t_{in_i})^T$ e, iii) dois vetores contendo os valores das variáveis explicativas, nomeadamente $\mathbf{x}_{hi}(t) = (1, x_{hi1}, \dots, x_{hip_n}(t))^T$, com $h = 1$ indicando o vetor correspondente ao modelo para o componente longitudinal e $h = 2$, o vetor correspondente ao modelo para o componente de sobrevivência.

A resposta longitudinal para o i -ésimo indivíduo no instante $t \geq 0$, $y_i(t)$, é descrita por meio do modelo (2.1), com a resposta média modelada como em (2.2). O processo $w_{1i}(t)$ é caracterizado em termos de um intercepto aleatório específico para o i -ésimo indivíduo, considerado invariante no tempo e com distribuição $b_{0i} \sim N(0, \sigma_0^2)$. Assume-se que o erro de medida leva em conta tanto a variabilidade inerente à resposta quanto a variabilidade associada às condições de medição e, é tal que $e_i(t) \sim N(0, \sigma_e^2)$ e é independente de b_{0i} , para todo $t \geq 0$.

Adicionalmente, o logaritmo do tempo de sobrevivência ou de censura observado para o i -ésimo indivíduo, $v_i = \log(Z_i)$, é modelado por meio do modelo de regressão log-linear Birnbaum-Saunders

$$v_i = \mathbf{x}_{2i}^T \boldsymbol{\beta}_2 + \varepsilon_i, \quad (3.1)$$

em que β_2 representa um vetor que contém os efeitos fixos correspondentes às p_2 variáveis explicativas contidas no vetor \mathbf{x}_{2i} ¹ e os termos ε_i são definidos de forma semelhante à considerada em (2.11), como variáveis aleatórias independentes e identicamente distribuídas, tais que $\varepsilon_i \sim \text{SenhN}(\alpha, 0, 2)$ para $i = 1, \dots, n$.

Além de utilizar o modelo paramétrico Birnbaum-Saunders para descrever a resposta de sobrevivência, tem-se interesse em incorporar as informações daqueles indivíduos para os quais a resposta longitudinal não foi observada.

Neste caso, as suposições de independência condicional (2.13) e (2.14) podem ser expressas como

$$p(v_i, \delta_i, \mathbf{y}_i | b_{0i}; \boldsymbol{\theta}) = p(v_i, \delta_i | b_{0i}; \boldsymbol{\theta}) p(\mathbf{y}_i | b_{0i}; \boldsymbol{\theta}),$$

e,

$$p(\mathbf{y}_i | b_{0i}; \boldsymbol{\theta}) = \prod_{j=1}^{n_i} p(y_i(t_{ij}) | b_{0i}; \boldsymbol{\theta}),$$

em que b_{0i} denota o intercepto aleatório do i -ésimo indivíduo e, $\boldsymbol{\theta} = (\boldsymbol{\theta}_z^T, \boldsymbol{\theta}_y^T, \theta_b)^T$ com $\boldsymbol{\theta}_z = (\boldsymbol{\theta}_{z0}^T, \gamma)^T$ e $\boldsymbol{\theta}_{z0} = (\alpha, \beta_2^T)^T$ representando os vetores que contêm os parâmetros associados à resposta de sobrevivência para os subconjuntos de indivíduos que possuem ou não medições da resposta longitudinal, respectivamente; $\boldsymbol{\theta}_y = (\beta_1^T, \sigma_e^2)^T$ denotando o vetor que contêm os parâmetros associados à resposta longitudinal e $\theta_b = \sigma_0^2$ igual à variância de b_{0i} .

Sob as suposições de independência condicional e considerando que a censura e os tempos de observação da resposta longitudinal são não informativos [Ver Rizopoulos (2012b)], a contribuição à verossimilhança conjunta de um indivíduo com pelo menos uma medição da resposta longitudinal é

$$\begin{aligned} p(v_i, \delta_i, \mathbf{y}_i; \boldsymbol{\theta}) &= \int p(v_i, \delta_i, \mathbf{y}_i, b_{0i}; \boldsymbol{\theta}) db_{0i} \\ &= \int p(v_i, \delta_i | b_{0i}; \boldsymbol{\theta}_z, \beta_1) \left[\prod_{j=1}^{n_i} p(y_i(t_{ij}) | b_{0i}; \boldsymbol{\theta}_y) \right] p(b_{0i}; \theta_b) db_{0i}, \end{aligned} \quad (3.2)$$

em que, sob independência entre os tempos de sobrevivência e de censura

$$p(v_i, \delta_i | b_{0i}; \boldsymbol{\theta}_z, \beta_1) = [f_i(v_i | M_i(v_i); \boldsymbol{\theta}_z, \beta_1)]^{\delta_i} [S_i(v_i | M_i(v_i); \boldsymbol{\theta}_z, \beta_1)]^{(1-\delta_i)}, \quad (3.3)$$

¹Ao contrário do modelo semi-paramétrico de Cox, os modelos de sobrevivência paramétricos não têm sido estendidos para permitir a inclusão de covariáveis dependentes do tempo externas, no sentido descrito no Capítulo 2, de modo que neste modelo serão consideradas apenas variáveis explicativas que não variam com o tempo.

com $f_i(\cdot)$ e $S_i(\cdot)$ denotando as funções densidade de probabilidade e de sobrevivência da distribuição SenhN explicitadas em (2.9) e (2.10), com parâmetro de forma $\alpha > 0$, parâmetro de escala $\sigma = 2$ e parâmetro de localização igual a

$$\begin{aligned}\psi_{\gamma i} &= \log(\eta_{\gamma i}) = \mathbf{x}_{2i}^T \boldsymbol{\beta}_2 + \gamma m_i(v_i) \\ &= \mathbf{x}_{2i}^T \boldsymbol{\beta}_2 + \gamma (\mathbf{x}_{1i}(v_i))^T \boldsymbol{\beta}_1 + b_{0i}, \quad i = 1, \dots, n,\end{aligned}\tag{3.4}$$

em que \mathbf{x}_{2i} e $\boldsymbol{\beta}_2$ são definidos como em (3.1) e, $M_i(\cdot)$ é definida como em (2.12) especificando o verdadeiro valor (não observável) da resposta longitudinal como

$$m_i(t) = \mathbf{x}_{1i}(t)^T \boldsymbol{\beta}_1 + b_{0i},\tag{3.5}$$

com $\mathbf{x}_{1i}(t)$ e $\boldsymbol{\beta}_1$ definidos como em (2.2). Adicionalmente γ mede a associação entre os processos das respostas longitudinal e de sobrevivência. Explicitamente as funções envolvidas em (3.3) são

$$\begin{aligned}f_i(v_i | M_i(v_i); \boldsymbol{\theta}_z, \boldsymbol{\beta}_1) &= \left(\frac{1}{\alpha \sqrt{2\pi}} \right) \cosh \left(\frac{v_i - \psi_{\gamma i}}{2} \right) \exp \left\{ -\frac{2}{\alpha^2} \sinh^2 \left(\frac{v_i - \psi_{\gamma i}}{2} \right) \right\} \\ &= \exp \left\{ \log(\kappa_{i1}(v_i)) - \frac{1}{2} \left(\log(8\pi) + \kappa_{i2}^2(v_i) \right) \right\},\end{aligned}\tag{3.6}$$

e,

$$\begin{aligned}S_i(v_i | M_i(v_i); \boldsymbol{\theta}_z, \boldsymbol{\beta}_1) &= 1 - \Phi \left(\frac{2}{\alpha} \sinh \left[\frac{v_i - \psi_{\gamma i}}{2} \right] \right) \\ &= 1 - \Phi(\kappa_{i2}(v_i)),\end{aligned}\tag{3.7}$$

em que

$$\kappa_{i1}(v_i) = \frac{2}{\alpha} \cosh \left(\frac{v_i - \psi_{\gamma i}}{2} \right) \quad \text{e} \quad \kappa_{i2}(v_i) = \frac{2}{\alpha} \sinh \left(\frac{v_i - \psi_{\gamma i}}{2} \right),\tag{3.8}$$

com $\psi_{\gamma i}$ definido em (3.4).

Além disso, as funções (2.17) e (2.18) podem ser expressas como

$$\begin{aligned}p(y_i(t_{ij}) | b_{0i}; \boldsymbol{\theta}_y) &= \frac{1}{(2\pi\sigma_e^2)^{1/2}} \exp \left\{ -\frac{(y_i(t_{ij}) - m_i(t_{ij}))^2}{2\sigma_e^2} \right\} \\ &= \frac{1}{(2\pi\sigma_e^2)^{1/2}} \exp \left\{ -\frac{(y_i(t_{ij}) - (\mathbf{x}_{1i}(t_{ij}))^T \boldsymbol{\beta}_1 + b_{0i})^2}{2\sigma_e^2} \right\},\end{aligned}\tag{3.9}$$

e,

$$p(b_{0i}; \theta_b) = \frac{1}{(2\pi\sigma_0^2)^{1/2}} \exp\left\{-\frac{b_{0i}^2}{2\sigma_0^2}\right\}. \quad (3.10)$$

Por outro lado, considerando que a censura é não informativa e assumindo independência entre os tempos de sobrevivência e de censura, a contribuição à verossimilhança conjunta de um indivíduo sem medições da resposta longitudinal é

$$p(v_i, \delta_i; \boldsymbol{\theta}_{z0}) = [f_i(v_i; \boldsymbol{\theta}_{z0})]^{\delta_i} [S_i(v_i; \boldsymbol{\theta}_{z0})]^{(1-\delta_i)}, \quad (3.11)$$

em que $f_i(\cdot)$ e $S_i(\cdot)$ têm interpretações semelhantes às consideradas em (3.3), com parâmetro de forma $\alpha > 0$, parâmetro de escala $\sigma = 2$ e parâmetro de localização

$$\psi_{0i} = \log(\eta_{0i}) = \mathbf{x}_{2i}^T \boldsymbol{\beta}_2, \quad (3.12)$$

com \mathbf{x}_{2i} e $\boldsymbol{\beta}_2$ definidos como em (3.1). As funções envolvidas em (3.11) podem ser escritas como

$$\begin{aligned} f_i(v_i; \boldsymbol{\theta}_{z0}) &= \left(\frac{1}{\alpha\sqrt{2\pi}} \right) \cosh\left(\frac{v_i - \psi_{0i}}{2}\right) \exp\left\{-\frac{2}{\alpha^2} \sinh^2\left(\frac{v_i - \psi_{0i}}{2}\right)\right\} \\ &= \exp\left\{\log(\xi_{i1}(v_i)) - \frac{1}{2}\left(\log(8\pi) + \xi_{i2}^2(v_i)\right)\right\}, \end{aligned} \quad (3.13)$$

e,

$$\begin{aligned} S_i(v_i; \boldsymbol{\theta}_{z0}) &= 1 - \Phi\left(\frac{2}{\alpha} \sinh\left[\frac{v_i - \psi_{0i}}{2}\right]\right) \\ &= 1 - \Phi(\xi_{i2}(v_i)), \end{aligned} \quad (3.14)$$

em que

$$\xi_{i1}(v_i) = \frac{2}{\alpha} \cosh\left(\frac{v_i - \psi_{0i}}{2}\right) \quad \text{e} \quad \xi_{i2}(v_i) = \frac{2}{\alpha} \sinh\left(\frac{v_i - \psi_{0i}}{2}\right), \quad (3.15)$$

com ψ_{0i} definido em (3.12).

De (3.2) e (3.11) conclui-se que a função de verossimilhança conjunta é

$$\begin{aligned}
 L(\boldsymbol{\theta}) &= \prod_{i=1}^n [p(v_i, \delta_i, \mathbf{y}_i; \boldsymbol{\theta})]^{\omega_i} [p(v_i, \delta_i; \boldsymbol{\theta}_{z_0})]^{(1-\omega_i)} \\
 &= \prod_{i=1}^n \left[\int p(v_i, \delta_i | b_{0i}; \boldsymbol{\theta}_z, \boldsymbol{\beta}_1) \left[\prod_{j=1}^{n_i} p(y_i(t_{ij}) | b_{0i}; \boldsymbol{\theta}_y) \right] p(b_{0i}; \boldsymbol{\theta}_b) db_{0i} \right]^{\omega_i} [p(v_i, \delta_i; \boldsymbol{\theta}_{z_0})]^{(1-\omega_i)},
 \end{aligned} \tag{3.16}$$

com cada um dos componentes definidos como nas expressões (3.3)-(3.15) e $\omega_i = I(n_i > 0)$, isto é, $\omega_i = 1$ se foi realizada pelo menos uma medição da resposta longitudinal e $\omega_i = 0$ em caso contrário, em que n_i denota o número total de observações longitudinais coletadas no indivíduo i no período de acompanhamento, $i = 1, \dots, n$.

3.1 Estimação e testes de hipóteses sobre os parâmetros do modelo conjunto

Estimativas de máxima verossimilhança dos parâmetros do modelo são obtidas por meio da maximização direta da log-verossimilhança correspondente à função (3.16), utilizando um algoritmo quase-Newton (rotinas PORT do otimizador `nlm` do R) [Ver Gay (1990)], que é um dos algoritmos considerados no pacote JM (Rizopoulos, 2012c). Além disso, como as integrais com respeito aos efeitos aleatórios b_{0i} , $i = 1, \dots, n$, não têm solução analítica, a sua aproximação é feita por meio do *método de integração numérica da quadratura de Gauss-Hermite* (QGH), de acordo com a abordagem adotada por Wulfson e Tsiatis (1997), Henderson, Diggle e Dobson (2000) e Rizopoulos (2010), entre outros, em casos como este, em que o vetor de efeitos aleatórios de cada indivíduo é de baixa dimensão. Em particular no pacote JM, este é o método usado por omissão para aproximar as integrais necessárias nas diferentes opções de ajuste disponíveis, exceto no caso em que explicitamente é utilizada uma aproximação de Laplace [Ver Rizopoulos, Verbeke e Lesaffre (2009)], recomendável em situações em que se tem um grande número de efeitos aleatórios para cada indivíduo. Detalhes sobre o método QGH unidimensional são apresentados no Apêndice C.

3.1.1 Aproximação da função de verossimilhança conjunta

Neste caso, tem-se interesse em aproximar integrais da forma

$$\int_{-\infty}^{\infty} g(b_0) db_0 = \int_{-\infty}^{\infty} w(b_0) h(b_0) db_0, \tag{3.17}$$

com $b_0 \sim N(0, \sigma_0^2)$. Considerando a função de densidade de b_0 dada por

$$w(b_0) = \frac{1}{\sqrt{2\pi\sigma_0}} \exp\left\{-\frac{b_0^2}{2\sigma_0^2}\right\},$$

como função de ponderação e a mudança de variáveis $z_0 = b_0/(\sqrt{2}\sigma_0)$, a integral (3.17) pode ser reescrita como

$$\begin{aligned} \int_{-\infty}^{\infty} g(b_0) db_0 &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_0}} \exp\{-z_0^2\} h(\sqrt{2}\sigma_0 z_0) \sqrt{2}\sigma_0 dz_0 \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}} \exp\{-z_0^2\} h(\sqrt{2}\sigma_0 z_0) dz_0, \end{aligned}$$

e pode ser aproximada por meio da QGH como

$$\int_{-\infty}^{\infty} g(b_0) db_0 \approx \sum_{k=1}^Q \frac{w_k}{\sqrt{\pi}} h(\sqrt{2}\sigma_0 s_k), \quad (3.18)$$

em que s_k é o k -ésimo zero do polinômio de Hermite de ordem Q e w_k é o peso associado, definidos no Apêndice C.

O logaritmo de (3.16) é

$$\begin{aligned} l(\boldsymbol{\theta}) &= \sum_{i=1}^n \omega_i \log[p(v_i, \delta_i, \mathbf{y}_i; \boldsymbol{\theta})] + \sum_{i=1}^n (1 - \omega_i) \log[p(v_i, \delta_i; \boldsymbol{\theta}_{z_0})] \\ &= \sum_{i=1}^n \omega_i \log \left[\int p(v_i, \delta_i | b_{0i}; \boldsymbol{\theta}_z, \boldsymbol{\beta}_1) \left[\prod_{j=1}^{n_i} p(y_i(t_{ij}) | b_{0i}; \boldsymbol{\theta}_y) \right] p(b_{0i}; \theta_b) db_{0i} \right] \\ &\quad + \sum_{i=1}^n (1 - \omega_i) \log[p(v_i, \delta_i; \boldsymbol{\theta}_{z_0})], \end{aligned} \quad (3.19)$$

e considerando (3.18), esta função pode ser aproximada como

$$l(\boldsymbol{\theta}) \approx \sum_{i=1}^n \omega_i \log \left[\sum_{k=1}^Q \frac{w_k}{\sqrt{\pi}} A_{ik}(\boldsymbol{\theta}) B_{ik}(\boldsymbol{\theta}) C_{ik}(\boldsymbol{\theta}) \right] + \sum_{i=1}^n (1 - \omega_i) \log[E_i(\boldsymbol{\theta}) F_i(\boldsymbol{\theta})], \quad (3.20)$$

em que

3.1 Estimação e testes de hipóteses sobre os parâmetros do modelo conjunto 39

$$\begin{aligned}
 A_{ik}(\boldsymbol{\theta}) &= \exp \left\{ \delta_i \left[\log(\kappa_{ik1}(v_i)) - \frac{1}{2} \left(\log(8\pi) + \kappa_{ik2}^2(v_i) \right) \right] \right\}, \\
 B_{ik}(\boldsymbol{\theta}) &= \exp \left\{ (1 - \delta_i) \log[1 - \Phi(\kappa_{ik2}(v_i))] \right\}, \\
 C_{ik}(\boldsymbol{\theta}) &= \exp \left\{ \sum_{j=1}^{n_i} \log \left[\frac{1}{\sqrt{2\pi}\sigma_e} \exp \left\{ -\frac{(y_i(t_{ij}) - (\mathbf{x}_{1i}(t_{ij})^T \boldsymbol{\beta}_1 + \sqrt{2}\sigma_0 s_k))^2}{2\sigma_e^2} \right\} \right] \right\},
 \end{aligned} \tag{3.21}$$

com

$$\begin{aligned}
 \kappa_{ik1}(v_i) &= \frac{2}{\alpha} \cosh \left(\frac{v_i - (\mathbf{x}_{2i}^T \boldsymbol{\beta}_2 + \gamma(\mathbf{x}_{1i}(v_i)^T \boldsymbol{\beta}_1 + \sqrt{2}\sigma_0 s_k))}{2} \right), \\
 \kappa_{ik2}(v_i) &= \frac{2}{\alpha} \sinh \left(\frac{v_i - (\mathbf{x}_{2i}^T \boldsymbol{\beta}_2 + \gamma(\mathbf{x}_{1i}(v_i)^T \boldsymbol{\beta}_1 + \sqrt{2}\sigma_0 s_k))}{2} \right),
 \end{aligned} \tag{3.22}$$

s_k , w_k e Q definidos como em (3.18),

$$\begin{aligned}
 E_i(\boldsymbol{\theta}) &= \exp \left\{ \delta_i \left[\log(\xi_{i1}(v_i)) - \frac{1}{2} \left(\log(8\pi) + \xi_{i2}^2(v_i) \right) \right] \right\} \text{ e}, \\
 F_i(\boldsymbol{\theta}) &= \exp \left\{ (1 - \delta_i) \log[1 - \Phi(\xi_{i2}(v_i))] \right\},
 \end{aligned} \tag{3.23}$$

em que $\xi_{i1}(v_i)$ e $\xi_{i2}(v_i)$ são definidos como em (3.15).

A função escore aproximada é

$$\mathbf{U}(\boldsymbol{\theta}) = (U_\alpha(\boldsymbol{\theta}), \mathbf{U}_{\boldsymbol{\beta}_2}(\boldsymbol{\theta})^T, U_\gamma(\boldsymbol{\theta}), \mathbf{U}_{\boldsymbol{\beta}_1}(\boldsymbol{\theta})^T, U_{\sigma_e^2}(\boldsymbol{\theta}), U_{\sigma_0^2}(\boldsymbol{\theta}))^T,$$

com componentes detalhados no Apêndice D.

O estimador de máxima verossimilhança de $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}_2^T, \gamma, \boldsymbol{\beta}_1^T, \sigma_e^2, \sigma_0^2)^T$, $\hat{\boldsymbol{\theta}}$, é obtido como a solução do sistema de equações $\mathbf{U}(\boldsymbol{\theta}) = \partial l(\boldsymbol{\theta}) / \partial \boldsymbol{\theta} = \mathbf{0}$. Quando a abordagem baseia-se na maximização de uma função de verossimilhança conjunta dos dados longitudinais e de sobrevivência, como a adotada aqui, os aspectos inferenciais geralmente apoiam-se em resultados empíricos e embora não sejam realizadas demonstrações formais, assume-se que $\hat{\boldsymbol{\theta}}$ tem distribuição aproximada $N(\boldsymbol{\theta}, \{\mathbf{I}(\boldsymbol{\theta})\}^{-1})$, com $\mathbf{I}(\boldsymbol{\theta}) = -\partial^2 l(\boldsymbol{\theta}) / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T$ representando a matriz de informação observada [Ver Rizopoulos (2012b)]. Tanto as estimativas dos componentes de $\boldsymbol{\theta}$, quanto as dos elementos de $\mathbf{I}(\boldsymbol{\theta})$ precisam ser calculadas de forma numérica. É importante considerar que, neste caso, as distribuições empíricas dos componentes de $\hat{\boldsymbol{\theta}}$,

obtidas por meio do estudo de simulação descrito no Capítulo 4, não contradizem a normalidade aproximada assumida anteriormente.

3.1.2 Testes de hipóteses

Se o interesse é testar as hipóteses

$$\{H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0 \text{ versus } H_a : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0\}, \quad (3.24)$$

é possível fazer uso de:

1. **Teste de Wald**, com a estatística de teste definida como

$$T_W = (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)^T \mathbf{I}(\hat{\boldsymbol{\theta}}) (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0),$$

em que $\hat{\boldsymbol{\theta}}$ representa a estimativa de máxima verossimilhança de $\boldsymbol{\theta}$ sob a hipótese alternativa,

2. **Teste da Razão de Verossimilhanças**, com a estatística de teste dada por

$$T_{RV} = -2\{l(\boldsymbol{\theta}_0) - l(\hat{\boldsymbol{\theta}})\},$$

ou,

3. **Teste de Escore**, com a estatística de teste igual a

$$T_E = \mathbf{U}(\boldsymbol{\theta}_0)^T \{\mathbf{I}(\boldsymbol{\theta}_0)\}^{-1} \mathbf{U}(\boldsymbol{\theta}_0),$$

em que $\mathbf{U}(\cdot)$ e $\mathbf{I}(\cdot)$ representam a função escore e a matriz de informação observada sob a hipótese alternativa.

Sob a hipótese nula, a distribuição assintótica de cada um destes testes é qui-quadrado com q graus de liberdade, com q denotando a dimensão do vetor de parâmetros $\boldsymbol{\theta}$. O leitor interessado em testar hipóteses mais gerais do que as especificadas em (3.24) no contexto dos modelos de conjuntos em questão, pode consultar Rizopoulos (2012b) para detalhes.

3.2 Predição das probabilidades de sobrevivência

Com base no ajuste de um modelo conjunto a uma amostra aleatória $D_n = \{Z_i, \delta_i, \mathbf{y}_i; i = 1, \dots, n\}$, tem-se interesse em prever as probabilidades de sobrevivência para um novo indivíduo i associado a um conjunto de medições longitudinais $Y_i(t) = \{y_i(s); 0 \leq s \leq t\}$ e valores das variáveis explicativas basais contidas no vetor \mathbf{x}_i . Aqui é importante salientar que uma das características da natureza endógena da resposta longitudinal é que ela está diretamente relacionada com o mecanismo de falha e, portanto, o fato de um indivíduo ter medições longitudinais até o tempo t implica a sua sobrevivência até esse instante.

Neste caso, deseja-se calcular a probabilidade condicional de sobreviver até um tempo $u > t$, dada a sobrevivência até o tempo t , isto é

$$\pi_i(u|t) = P(T_i \geq u | T_i > t, Y_i(t), \mathbf{x}_i, D_n; \boldsymbol{\theta}), \quad t > 0. \quad (3.25)$$

Utilizando (2.13), pode-se reescrever (3.25) como ²

$$\begin{aligned} & P(T_i \geq u | T_i > t, Y_i(t); \boldsymbol{\theta}) \\ &= \int P(T_i \geq u | T_i > t, Y_i(t), \mathbf{b}_i; \boldsymbol{\theta}) p(\mathbf{b}_i | T_i > t, Y_i(t); \boldsymbol{\theta}) d\mathbf{b}_i, \\ &= \int P(T_i \geq u | T_i > t, \mathbf{b}_i; \boldsymbol{\theta}) p(\mathbf{b}_i | T_i > t, Y_i(t); \boldsymbol{\theta}) d\mathbf{b}_i, \\ &= \int \frac{S_i\{u | M_i(u, \mathbf{b}_i, \boldsymbol{\theta}); \boldsymbol{\theta}\}}{S_i\{t | M_i(t, \mathbf{b}_i, \boldsymbol{\theta}); \boldsymbol{\theta}\}} p(\mathbf{b}_i | T_i > t, Y_i(t); \boldsymbol{\theta}) d\mathbf{b}_i, \end{aligned} \quad (3.26)$$

em que $S_i\{t | M_i(t, \mathbf{b}_i, \boldsymbol{\theta}); \boldsymbol{\theta}\} = P\{T_i > t | M_i(t, \mathbf{b}_i, \boldsymbol{\theta}); \boldsymbol{\theta}\}$ denota a função de sobrevivência e $M_i(t)$ é definida como em (2.12), especificando $m_i(t)$ como uma função tanto dos parâmetros quanto dos efeitos aleatórios da forma considerada em (2.1).

Com base em (3.26) e no preditor de Bayes empírico de \mathbf{b}_i , Rizopoulos (2011) propôs o estimador para $\pi_i(u|t)$ dado por

$$\tilde{\pi}_i(u|t) = \frac{S_i\{u | M_i(u, \tilde{\mathbf{b}}_i^{(t)}, \hat{\boldsymbol{\theta}}); \hat{\boldsymbol{\theta}}\}}{S_i\{t | M_i(t, \tilde{\mathbf{b}}_i^{(t)}, \hat{\boldsymbol{\theta}}); \hat{\boldsymbol{\theta}}\}}, \quad (3.27)$$

em que $\tilde{\mathbf{b}}_i^{(t)}$ denota a predição do efeito aleatório \mathbf{b}_i no tempo t , que neste caso é a moda do logaritmo da distribuição condicional

²Com o intuito de simplificar as expressões, o condicionamento nas covariáveis contidas em \mathbf{x}_i e na amostra D_n é assumido, mas omitido da notação, conforme considerado por Rizopoulos (2011).

$$\begin{aligned}
p(\mathbf{b}|T_i > t, Y_i(t); \hat{\boldsymbol{\theta}}) &= \frac{P(T_i > t, Y_i(t), \mathbf{b}; \hat{\boldsymbol{\theta}})}{P(T_i > t, Y_i(t); \hat{\boldsymbol{\theta}})} \\
&= \frac{P(T_i > t|\mathbf{b}; \hat{\boldsymbol{\theta}})p(Y_i(t)|\mathbf{b}; \hat{\boldsymbol{\theta}})p(\mathbf{b}; \hat{\boldsymbol{\theta}})}{\int P(T_i > t|\mathbf{b}_i; \hat{\boldsymbol{\theta}})p(Y_i(t)|\mathbf{b}_i; \hat{\boldsymbol{\theta}})p(\mathbf{b}_i; \hat{\boldsymbol{\theta}})d\mathbf{b}_i},
\end{aligned}$$

em que $P\{T_i > t|\mathbf{b}_i; \hat{\boldsymbol{\theta}}\} = S_i\{t|M_i(t, \mathbf{b}_i; \hat{\boldsymbol{\theta}}); \hat{\boldsymbol{\theta}}\}$ denota a função de sobrevivência, $M_i(t)$ é definida como em (3.26) e,

$$p(Y_i(t)|\mathbf{b}_i; \hat{\boldsymbol{\theta}}) = \prod_{j=1}^{n_i(t)} p(y_i(t_{ij})|\mathbf{b}_i; \hat{\boldsymbol{\theta}}),$$

com $n_i(t)$ denotando o número de observações da resposta longitudinal do indivíduo i até o tempo t . O desempenho do estimador (3.27) em amostras finitas vai depender da qualidade das estimativas de máxima verossimilhança dos componentes de $\boldsymbol{\theta}$ e da predição do efeito aleatório \mathbf{b}_i .

Uma característica importante das predições das probabilidades de sobrevivência (3.25), é a sua natureza dinâmica, ou seja, elas podem ser atualizadas à medida que novos dados são coletados nos indivíduos durante o período de acompanhamento, o que pode ser muito útil na prática. No contexto médico, por exemplo, especialistas podem estar interessados em utilizar toda a informação disponível sobre um determinado paciente (basal e longitudinal) em um instante específico do seguimento, a fim de prever a sua sobrevivência, obter uma melhor compreensão da evolução da doença e tomar medidas adequadas naquele momento em particular.

Finalmente, é importante ressaltar que a metodologia proposta neste trabalho foi totalmente implementada no ambiente R (R Development Core Team, 2013). O leitor poderá consultar o Apêndice C para maiores detalhes.

Capítulo 4

Estudo de simulação e aplicação da metodologia proposta

4.1 Estudo de simulação

Foi realizado um estudo de simulação com três objetivos principais. O primeiro é avaliar a robustez do modelo conjunto proposto relativamente a uma má especificação da distribuição paramétrica da resposta de sobrevivência. O segundo é analisar o comportamento das predições dinâmicas das probabilidades de sobrevivência, calculadas com base no ajuste do modelo conjunto e no uso de todas as informações disponíveis até momento em que se deseja realizar a previsão, considerando os diferentes instantes em que são coletadas as observações da resposta longitudinal; a finalidade aqui é verificar se um aumento na quantidade de informação longitudinal disponível melhora a qualidade das predições tanto dos efeitos aleatórios quanto das probabilidades de sobrevivência em menção. O terceiro é avaliar o desempenho do processo de estimação dos parâmetros em termos do viés e do erro quadrático médio das estimativas obtidas por meio do ajuste do modelo conjunto, nos cenários em que o modelo de sobrevivência utilizado para gerar os dados é igual ao considerado no modelo ajustado.

O estudo foi baseado no conjunto de dados reais que motivou o presente trabalho, considerando as informações referentes a:

- i) uma sequência de medições longitudinais do logaritmo natural dos níveis de BNP no sangue (LBNP) ¹, contidas no vetor $\mathbf{y}_i = (y_{i1}, \dots, y_{in_i})^T$, observadas nos tempos es-

¹Levando em conta que o níveis de BNP apresentam alta variabilidade ao longo do tempo, além de ter uma distribuição assimétrica com concentração em valores baixos, aqui e nas análises posteriores será considerado o logaritmo natural do BNP como a variável resposta longitudinal, tendo em vista as considerações de Azevedo, Bettencourt e Barros (2007) e Hogenhuis et al. (2007).

pecificados no vetor $\mathbf{t}_i = (t_{i1}, \dots, t_{in_i})^T$, $i = 1, \dots, n$;

- ii) um registro do tempo entre a admissão do paciente no estudo e a data de óbito (T_i) ou da censura (C_i) em meses, realizado por meio do uso das variáveis $Z_i = \min(T_i, C_i)$ e $\delta_i = I(T_i \leq C_i)$;
- iii) valores de algumas das variáveis explicativas coletadas na data de admissão do paciente no estudo, contidos nos vetores $\mathbf{x}_{1i}(t)$ e \mathbf{x}_{2i} , em que $\mathbf{x}_{1i}(t) = (1, x_{1i1}(t), x_{1i2})^T$ representa o vetor das covariáveis consideradas no modelo para o componente longitudinal, com $x_{1i1}(t)$ denotando o tempo de observação da resposta longitudinal e x_{1i2} denotando o diagnóstico da ICC. Além disso, $\mathbf{x}_{2i} = (1, x_{2i1})^T$ representa o vetor das covariáveis consideradas no modelo para o componente de sobrevivência, com x_{2i1} denotando o diagnóstico da ICC. Nestes modelos foi considerado o diagnóstico da ICC pelo fato de ser o fator mais significativo nos ajustes dos modelos marginais para as respostas longitudinal e de sobrevivência ².

No conjunto de dados há uma grande proporção de pacientes sem informações da resposta longitudinal (LBNP), de modo que para o estudo de simulação considerou-se somente o subconjunto dos $n = 529$ pacientes com pelo menos uma medição dessa resposta.

A resposta longitudinal para o i -ésimo indivíduo no instante $t \geq 0$, $y_i(t)$, foi descrita por meio do modelo (2.1)-(2.2), com $\boldsymbol{\beta}_1 = (\beta_{10}, \beta_{11}, \beta_{12})^T$. O processo $w_{1i}(t)$ foi caracterizado em termos de um intercepto aleatório específico do i -ésimo indivíduo, considerado invariante no tempo e com distribuição $b_{0i} \sim N(0, \sigma_0^2)$ e, assumiu-se que o erro de medida é tal que $e_i(t) \sim N(0, \sigma_e^2)$ e é independente de b_{0i} , para todo $t \geq 0$.

Além disso, o logaritmo do tempo de sobrevivência ou de censura observado para o i -ésimo indivíduo, $v_i = \log(Z_i)$, foi modelado por meio dos modelos de regressão:

- i) log-linear Birnbaum-Saunders (3.1);
- ii) log-linear Weibull (descrito no Apêndice E),

em que $\boldsymbol{\beta}_2 = (\beta_{20}, \beta_{21})^T$.

Para a modelagem conjunta foi considerada a abordagem tradicional, na qual somente são incluídos os indivíduos que têm pelo menos uma medição da resposta longitudinal, de modo a que a função de verossimilhança conjunta (3.16) reduz-se a

²Embora esta covariável originalmente tenha mais de duas categorias, para efeitos ilustrativos, aqui será considerada como uma variável binária com as categorias (0 = cardiomiopatia da doença de Chagas, 1 = outras cardiomiopatias).

$$\begin{aligned}
L(\boldsymbol{\theta}) &= \prod_{i=1}^n [p(v_i, \delta_i, \mathbf{y}_i; \boldsymbol{\theta})] \\
&= \prod_{i=1}^n \left[\int p(v_i, \delta_i | b_{0i}; \boldsymbol{\theta}_z, \boldsymbol{\beta}_1) \left[\prod_{j=1}^{n_i} p(y_i(t_{ij}) | b_{0i}; \boldsymbol{\theta}_y) \right] p(b_{0i}; \theta_b) db_{0i} \right],
\end{aligned} \tag{4.1}$$

com cada um dos componentes definidos como nas expressões (3.2)-(3.10) quando é considerado o modelo log-linear Birnbaum-Saunders para a explicação do componente de sobrevivência. Quando o modelo utilizado é o modelo log-linear Weibull, as funções $f_i(\cdot)$ e $S_i(\cdot)$ a ser consideradas em (3.3) correspondem às funções densidade de probabilidade e de sobrevivência da distribuição do valor extremo (VE) especificadas no Apêndice E, com parâmetro de escala $\zeta > 0$ e parâmetro de localização

$$\begin{aligned}
\varphi_{\gamma_i} &= \log(\zeta_{\gamma_i}) = \mathbf{x}_{2i}^T \boldsymbol{\beta}_2 + \gamma m_i(v_i) \\
&= \mathbf{x}_{2i}^T \boldsymbol{\beta}_2 + \gamma (\mathbf{x}_{1i}(v_i))^T \boldsymbol{\beta}_1 + b_{0i}, \quad i = 1, \dots, n,
\end{aligned}$$

em que todos os termos têm a mesma interpretação que em (3.4).

4.1.1 Detalhes do processo de simulação

1. As observações da resposta longitudinal para cada indivíduo foram geradas considerando a distribuição Normal, nos tempos $t = 0, 24, 48, 72, 96$ (meses). Uma vez geradas as medições longitudinais, foram gerados os tempos de sobrevivência considerando as distribuições Birnbaum-Saunders ou Weibull, envolvendo uma medida das observações da resposta longitudinal de cada indivíduo. Os valores dos parâmetros foram considerados iguais às estimativas obtidas sob o ajuste dos modelos conjuntos: i) log-linear Birnbaum-Saunders para o componente de sobrevivência e linear de efeitos mistos Normal para o componente longitudinal (BS-Normal) ou, ii) log-linear Weibull para o componente de sobrevivência e linear de efeitos mistos Normal para o componente longitudinal (Wei-Normal), com os dados dos $n = 529$ pacientes portadores de ICC com pelo menos uma medição da resposta longitudinal. No entanto, devido ao alto percentual de censuras neste conjunto de dados (aproximadamente 76%), os valores dos interceptos para o componente de sobrevivência foram fixados conforme indicado na Tabela 4.1, a fim de ter a possibilidade de considerar cenários com percentuais de

censura mais baixos.

Tabela 4.1: *Valores dos parâmetros.*

Parâmetro									
Modelo	$\log(\alpha)$	β_{20}	β_{21}	γ	β_{10}	β_{11}	β_{12}	$\log(\sigma_0)$	$\log(\sigma_e)$
B-S-Normal	-0,46	6,14	0,10	-0,46	5,80	0,01	-1,10	0,45	-0,32
Parâmetro									
Modelo	$\log(\varsigma)$	β_{20}	β_{21}	γ	β_{10}	β_{11}	β_{12}	$\log(\sigma_0)$	$\log(\sigma_e)$
Wei-Normal	-0,97	6,28	0,12	-0,43	5,83	0,01	-1,11	0,45	-0,33

Os tempos de sobrevivência foram censurados considerando dois tipos de censura à direita na seguinte ordem: i) censura Tipo I, especificando um tempo máximo de acompanhamento de $\tau = 180$ meses (o percentual de censura alcançado nesta fase é denotado por PcI) e, ii) censura aleatória, realizando a quantidade de censuras necessárias para atingir o percentual de censura à direita total pré-definido para cada cenário (Pc). Nesta fase um tempo de censura Uniforme($0, \tau$) foi gerado aleatoriamente e o indivíduo a ser censurado foi escolhido ao acaso do grupo dos não censurados com sobrevivência maior ao tempo de censura gerado (o percentual de censura atingido nesta fase é rotulado PcA , de modo que $Pc = PcI + PcA$).

2. Para o ajuste de cada um dos dois modelos conjuntos de interesse (BS-Normal) e (Wei-Normal) foram considerados 32 cenários resultantes da combinação dos níveis de: Distribuição para a resposta de sobrevivência (Birnbau-Saunders, Weibull), tamanho da amostra ($n = 100, 250, 500, 1000$) e percentual de censura à direita ($Pc = 0\%, 25\%, 50\%, 75\%$)³.
3. Em cada um dos cenários em questão foram geradas 500 amostras e para cada amostra foram selecionados aleatoriamente 95% dos indivíduos para ajustar o modelo conjunto de interesse, (BS-Normal) ou (Wei-Normal), e o 5% restante foi utilizado para estimar as probabilidades condicionais de sobrevivência $\pi_i(t + \Delta t|t)$ da forma dada em (3.25), especificadas como uma função dos tempos de observação da resposta longitudinal $t = 0, 24, 48, 72, 96$ (meses) e dos incrementos de tempo $\Delta t = 12, 24, 36$ (meses).

³Nos casos com o percentual de censura à direita mais baixo, $Pc = 0\%$, o percentual de censura aleatória é igual a zero e o percentual de censura Tipo I é o mais baixo possível dada a especificação do tempo máximo de acompanhamento de $\tau = 180$ meses. Nos outros casos, $Pc = PcI + PcA$ de acordo com o descrito no item 1.

Para garantir que os percentuais de pacientes diagnosticados com cardiomiopatia da doença de Chagas (13%) e outras cardiomiopatias (87%) observados no conjunto de dados original fossem respeitados nos conjuntos dos indivíduos escolhidos para os processos de estimação dos parâmetros e de previsão das probabilidades condicionais de sobrevivência, foi considerada uma seleção aleatória estratificada pelo diagnóstico da ICC.

4. As probabilidades condicionais de sobrevivência foram estimadas com base no estimador $\tilde{\pi}_i(t + \Delta t|t)$ definido em (3.27), além de ser calculados os seus verdadeiros valores $\pi_i(t + \Delta t|t)$ com os valores dos parâmetros e os efeitos aleatórios simulados. Foram calculadas as diferenças em valor absoluto entre as estimativas das probabilidades condicionais de sobrevivência e os seus verdadeiros valores (DAI). Para cada combinação dos valores de t e Δt foi gerado o histograma das diferenças DAI obtidas em todas as 500 amostras geradas no cenário analisado e calculado o percentil 95 de tais diferenças. Posteriormente, para cada uma das amostras geradas foi calculado o máximo das diferenças DAI e, em seguida, foram calculadas tanto a média quanto a mediana desses máximos (Máx (DAI)) provenientes de todas as 500 amostras. O comportamento das medidas obtidas em cada caso foi analisado considerando valores de t fixos e diferentes valores de Δt e vice-versa.
5. Nos cenários em que o modelo de sobrevivência utilizado para gerar os dados é igual ao considerado no modelo conjunto ajustado, foi realizada uma avaliação do desempenho do processo de estimação dos parâmetros em termos do viés, do viés relativo, da raiz do erro quadrático médio e da raiz do erro quadrático médio relativo das estimativas obtidas, definidos respectivamente como

$$VI(\hat{\theta}) = \hat{\theta} - \theta, \quad VIR(\hat{\theta}) = \frac{\hat{\theta} - \theta}{\theta}, \quad \text{com } \hat{\theta} = \frac{1}{R} \sum_{l=1}^R \hat{\theta}_l,$$

$$REQM(\hat{\theta}) = \sqrt{\frac{1}{R} \sum_{l=1}^R (\hat{\theta}_l - \theta)^2} \quad \text{e} \quad REQMR(\hat{\theta}) = \sqrt{\frac{1}{R} \sum_{l=1}^R \left(\frac{\hat{\theta}_l - \theta}{\theta} \right)^2},$$

em que $\hat{\theta}_l$ denota a estimativa do parâmetro θ obtida no ajuste realizado com a l -ésima amostra, $l = 1, \dots, 500$.

Para a realização do estudo de simulação descrito anteriormente foram utilizados recursos computacionais disponíveis tanto nos clusters Puma e Aguia administrados pelo Laboratório de Computação Científica Avançada (LCCA) da Universidade de São Paulo (USP), quanto

no projeto NUVEM da USP.

É importante destacar que em nenhuma das amostras simuladas nos diferentes cenários considerados o processo de estimação dos parâmetros apresentou problemas de convergência.

4.1.2 Resultados

Para a análise dos resultados do estudo de simulação serão considerados inicialmente os cenários com o maior tamanho da amostra ($n = 1000$) e o menor percentual de censura à direita ($Pc = 0\%$) nos casos:

- i) Distribuição Birnbaum-Saunders para gerar a resposta de sobrevivência e modelo conjunto ajustado BS-Normal (Dados BS - Modelo BS-Normal),
- ii) Distribuição Weibull para gerar a resposta de sobrevivência e modelo conjunto ajustado BS-Normal (Dados Wei - Modelo BS-Normal),
- iii) Distribuição Birnbaum-Saunders para gerar a resposta de sobrevivência e modelo conjunto ajustado Wei-Normal (Dados BS - Modelo Wei-Normal) e,
- iv) Distribuição Weibull para gerar a resposta de sobrevivência e modelo conjunto ajustado Wei-Normal (Dados Wei - Modelo Wei-Normal).

Na Figura 4.1 são apresentados os histogramas das diferenças DAI obtidas em todas as amostras geradas no caso Dados BS - Modelo BS-Normal para cada combinação dos valores de t e Δt e nas Figuras 4.2 e 4.3 são expostos os gráficos dos valores do percentil 95 de tais diferenças, considerando diferentes valores de Δt para t fixo e diferentes valores de t para Δt fixo, respectivamente. Tanto os valores representados quanto o número de diferenças calculadas em cada situação estão resumidos na Tabela 4.2.

Tabela 4.2: (Dados BS - Modelo BS-Normal): Percentil 95 e quantidade de diferenças (DAI), $n = 1000$, $PcI = 7\%$, $PcA = 0\%$.

	Percentil 95			Quantidade de diferenças
	$\Delta t = 12$	$\Delta t = 24$	$\Delta t = 36$	
$t = 0$	0,2509	0,3376	0,3463	25000
$t = 24$	0,1812	0,2411	0,2675	18354
$t = 48$	0,1237	0,1816	0,2125	11736
$t = 72$	0,0939	0,1467	0,1780	7761
$t = 96$	0,0753	0,1225	0,1528	5364

Quando a atenção é concentrada em valores diferentes de Δt para t fixo observa-se (tanto nos histogramas quanto nos valores do percentil 95) que a magnitude das diferenças DAI aumenta conforme Δt aumenta, o que é esperado pelo fato de se ter um maior espaçamento entre o momento em que se tem as últimas informações do indivíduo e o instante para o qual é feita a previsão. Além disso, para Δt fixo e diferentes valores de t há uma diminuição na magnitude das diferenças DAI conforme t aumenta. Considerando que esse resultado poderia ser devido a uma melhora nas previsões dos efeitos aleatórios envolvidos no estimador utilizado, calculou-se o percentil 95 das diferenças em valor absoluto entre as previsões dos efeitos aleatórios e os seus verdadeiros valores (simulados), obtendo-se uma diminuição de tal medida conforme t aumenta, o que sugere que a coleta de mais medidas longitudinais para cada indivíduo leva a uma melhora na previsão de seu efeito aleatório e ao mesmo tempo pode produzir uma melhora na estimação de sua probabilidade de sobrevivência.

Similarmente, em todos os cenários da simulação calculou-se o percentil 95 das diferenças em valor absoluto entre as previsões dos efeitos aleatórios e os seus verdadeiros valores e, em todos os casos apresentou-se uma diminuição de tal medida conforme t aumentou, corroborando o resultado anterior.

Adicionalmente, nas Figuras 4.4 - 4.7 são apresentados os gráficos dos valores da média e da mediana dos máximos ($\text{Máx}(\text{DAI})$) considerando diferentes valores de Δt para t fixo e diferentes valores de t para Δt fixo, respectivamente. Esses valores estão resumidos na Tabela 4.3. Pode-se observar que o comportamento é semelhante ao apresentado em termos do percentil 95 das diferenças DAI, reiterando as conclusões anteriores.

Tabela 4.3: (*Dados BS - Modelo BS-Normal*): Média e mediana ($\text{Máx}(\text{DAI})$), $n = 1000$, $PcI = 7\%$, $PcA = 0\%$.

	$\Delta t = 12$		$\Delta t = 24$		$\Delta t = 36$	
	Média	Mediana	Média	Mediana	Média	Mediana
$t = 0$	0,3031	0,2767	0,4050	0,3731	0,4317	0,4002
$t = 24$	0,2132	0,1918	0,2890	0,2632	0,3269	0,3012
$t = 48$	0,1410	0,1215	0,2048	0,1843	0,2415	0,2203
$t = 72$	0,1055	0,0881	0,1594	0,1411	0,1910	0,1749
$t = 96$	0,0801	0,0682	0,1265	0,1112	0,1552	0,1417

Seguindo a mesma estrutura anterior, na Figura 4.8 são apresentados os histogramas das diferenças DAI obtidas em todas as amostras geradas no caso Dados Wei - Modelo BS-Normal para cada combinação dos valores de t e Δt e nas Figuras 4.9 e 4.10 são expostos os

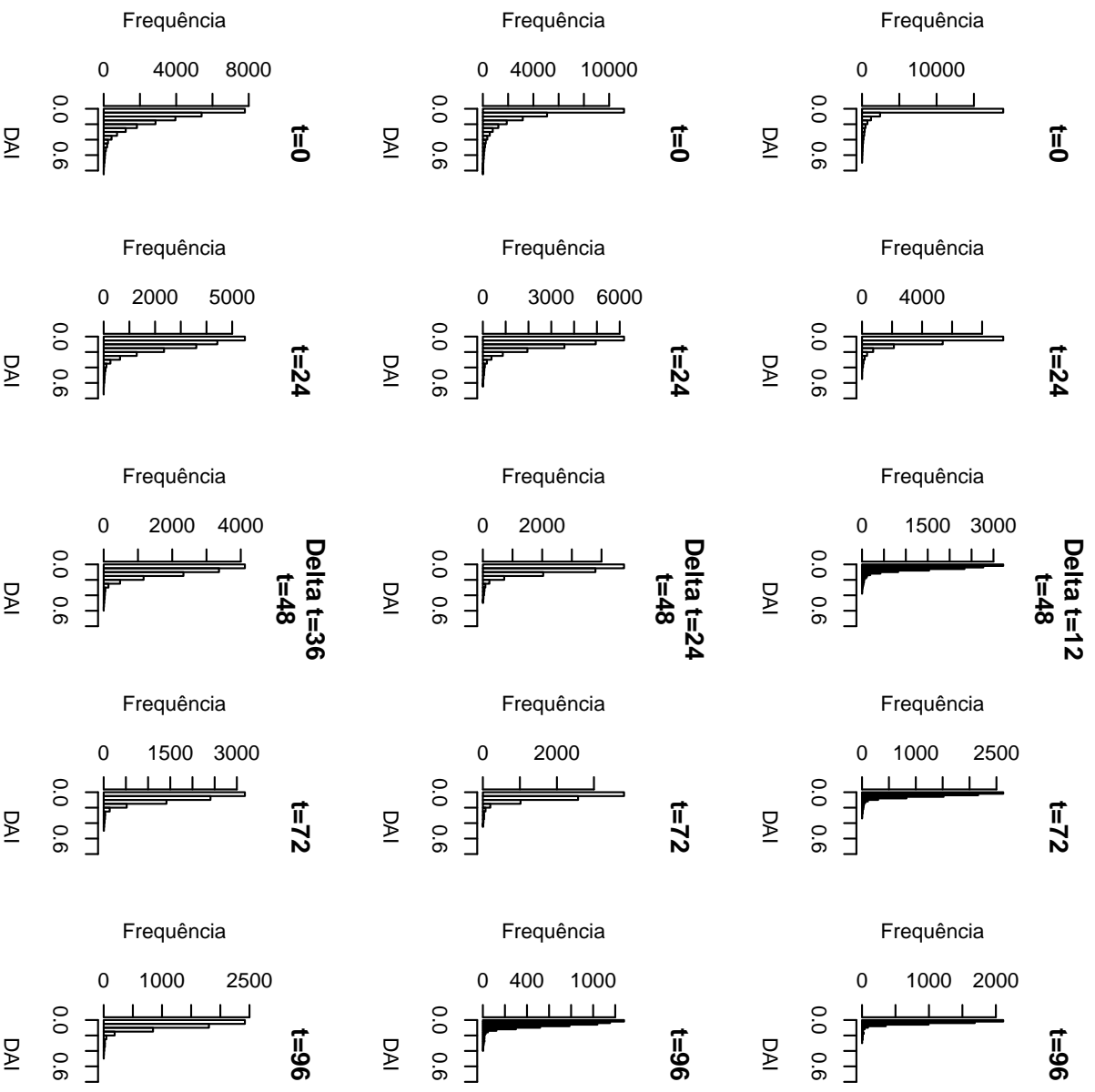


Figura 4.1: (Dados BS - Modelo BS-Normal): Histograma (ADI), $n = 1000$, $P_{cI} = 7\%$, $P_{cA} = 0\%$.

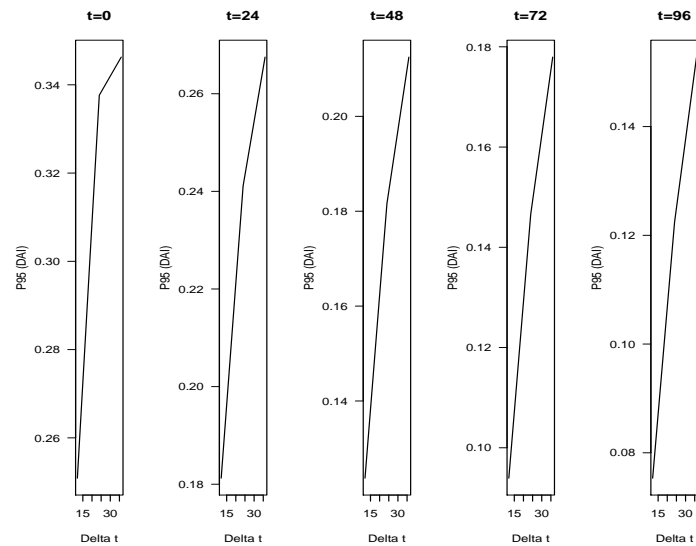


Figura 4.2: (Dados BS - Modelo BS-Normal): Percentil 95 (ADI), diferentes Δt para t fixo, $n = 1000$, $PcI = 7\%$, $PcA = 0\%$.

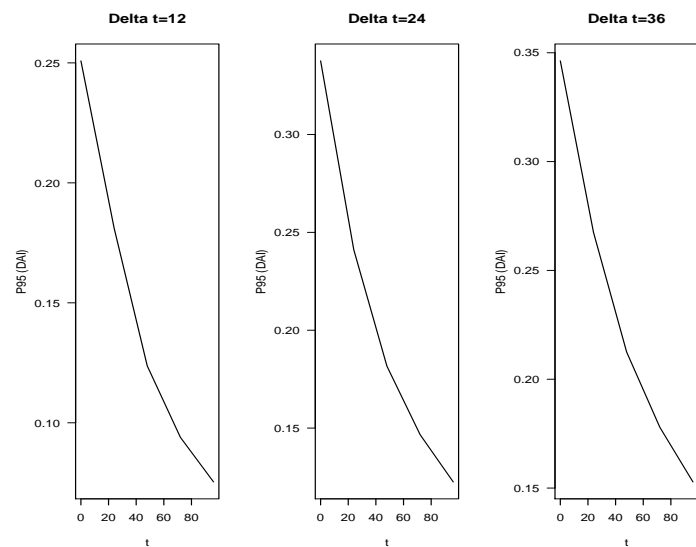


Figura 4.3: (Dados BS - Modelo BS-Normal): Percentil 95 (ADI), diferentes t para Δt fixo, $n = 1000$, $PcI = 7\%$, $PcA = 0\%$.

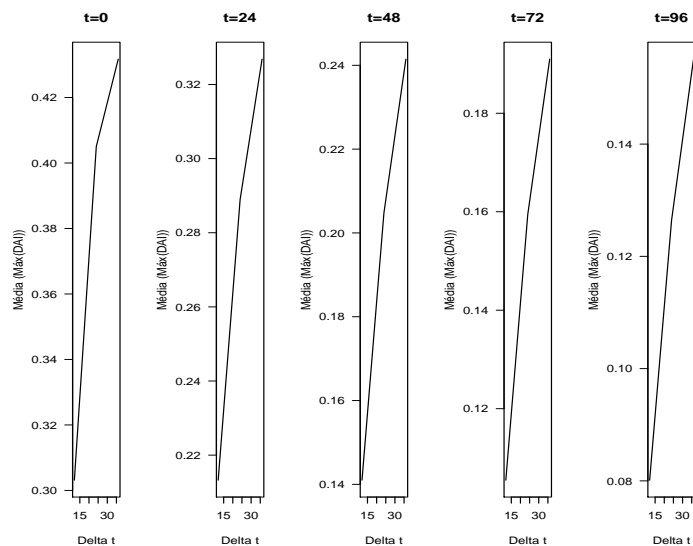


Figura 4.4: (Dados BS - Modelo BS-Normal): Médias Máx(ADI), diferentes Δt para t fixo, $n = 1000$, $PcI = 7\%$, $PcA = 0\%$.

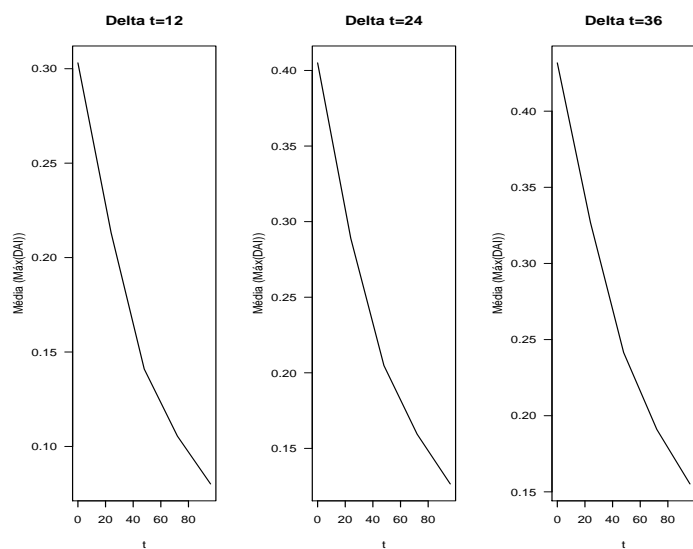


Figura 4.5: (Dados BS - Modelo BS-Normal): Médias Máx(ADI), diferentes t para Δt fixo, $n = 1000$, $PcI = 7\%$, $PcA = 0\%$.

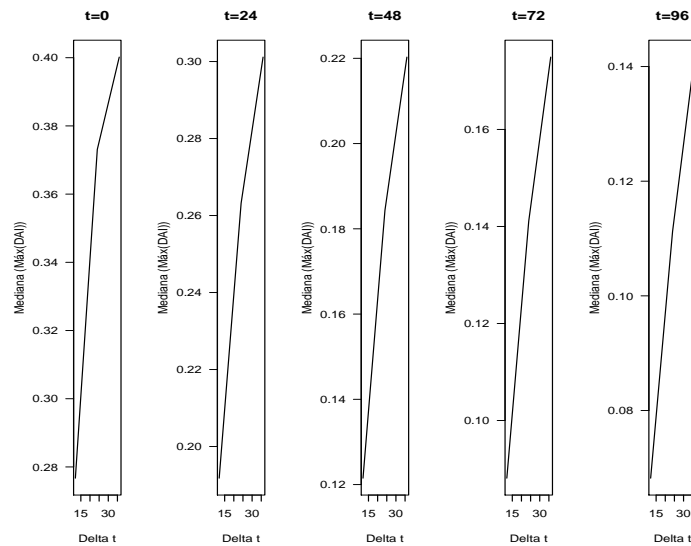


Figura 4.6: (Dados BS - Modelo BS-Normal): Medianas Máx(ADI), diferentes Δt para t fixo, $n = 1000$, $PcI = 7\%$, $PcA = 0\%$.

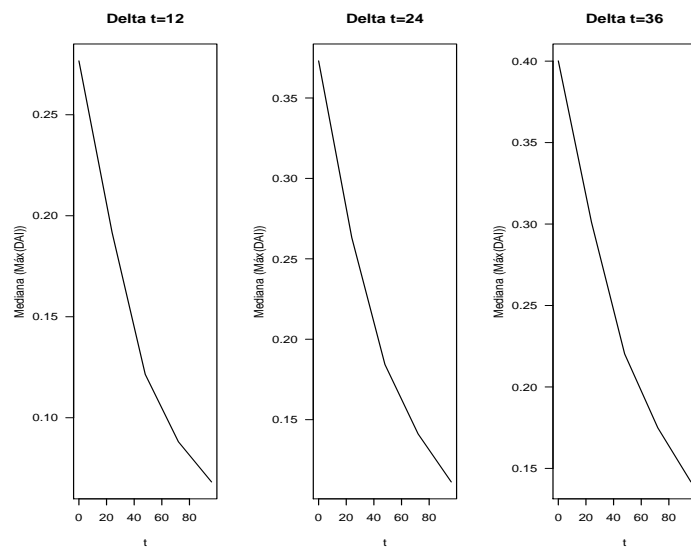


Figura 4.7: (Dados BS - Modelo BS-Normal): Medianas Máx(ADI), diferentes t para Δt fixo, $n = 1000$, $PcI = 7\%$, $PcA = 0\%$.

gráficos dos valores do percentil 95 de tais diferenças. Tanto os valores representados quanto o número de diferenças calculadas em cada situação estão resumidos na Tabela 4.4.

Tabela 4.4: (*Dados Wei - Modelo BS-Normal*): Percentil 95 e quantidade de diferenças (DAI), $n = 1000$, $PcI = 6\%$, $PcA = 0\%$.

	Percentil 95			Quantidade de diferenças
	$\Delta t = 12$	$\Delta t = 24$	$\Delta t = 36$	
$t = 0$	0,2901	0,4597	0,4615	25000
$t = 24$	0,3282	0,3656	0,3723	20096
$t = 48$	0,2571	0,3259	0,3293	13043
$t = 72$	0,2266	0,3076	0,3169	8342
$t = 96$	0,1998	0,2835	0,3016	5420

O comportamento das diferenças DAI para diferentes valores de Δt e t fixo é semelhante ao apresentado no caso anterior; no entanto, para $\Delta t = 12$ com t fixo observa-se que o percentil 95 das diferenças é menor quando $t = 0$ do que quando $t = 24$, talvez porque tanto o estimador (3.27) quanto o verdadeiro valor das probabilidades condicionais de sobrevivência (3.25) dependam apenas do numerador para os casos em que $t = 0$, uma vez que o denominador é igual a 1. No entanto, para valores de $t \neq 0$, ainda se observa uma diminuição na magnitude das diferenças DAI conforme t aumenta.

Nas Figuras 4.11 - 4.14 são expostos os gráficos dos valores da média e da mediana dos máximos (Máx(DAI)). Os valores estão resumidos na Tabela 4.5. O comportamento é semelhante ao apresentado em termos do percentil 95 das diferenças DAI.

Tabela 4.5: (*Dados Wei - Modelo BS-Normal*): Média e mediana (Máx(DAI)), $n = 1000$, $PcI = 6\%$, $PcA = 0\%$.

	$\Delta t = 12$		$\Delta t = 24$		$\Delta t = 36$	
	Média	Mediana	Média	Mediana	Média	Mediana
$t = 0$	0,2521	0,1889	0,4206	0,3881	0,4656	0,4227
$t = 24$	0,3292	0,3145	0,3882	0,3584	0,4222	0,3853
$t = 48$	0,2824	0,2593	0,3514	0,3334	0,3676	0,3357
$t = 72$	0,2316	0,2174	0,3170	0,3012	0,3352	0,3148
$t = 96$	0,1809	0,1411	0,2646	0,2299	0,2960	0,2755

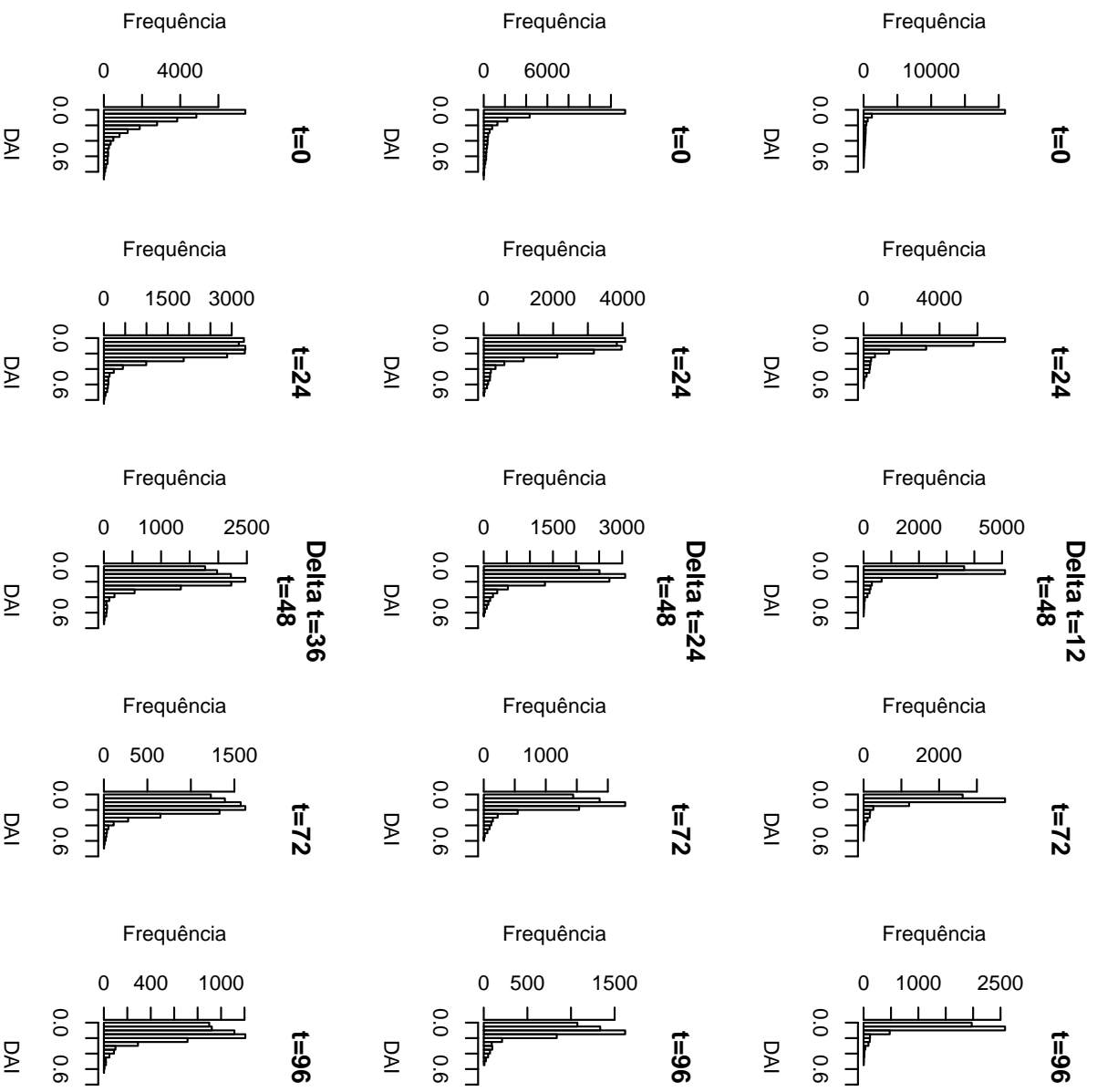


Figura 4.8: (Dados Wei - Modelo BS-Normal): Histograma (ADI), $n = 1000$, $PcI = 6\%$, $PcA = 0\%$

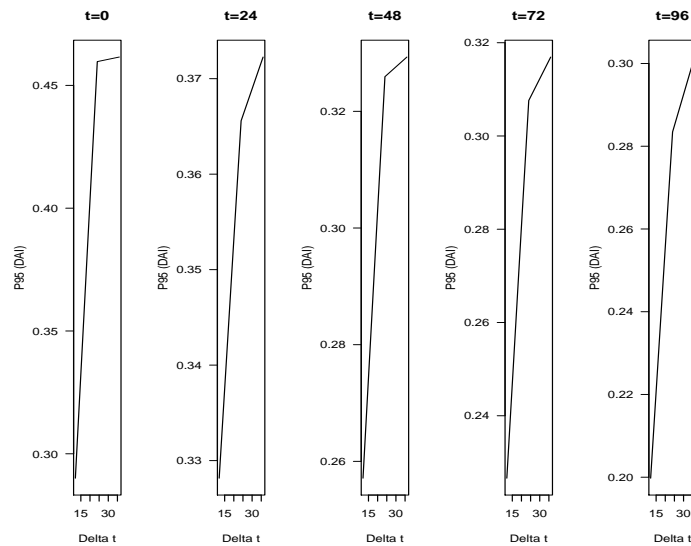


Figura 4.9: (Dados Wei - Modelo BS-Normal): Percentil 95 (ADI), diferentes Δt para t fixo, $n = 1000$, $PcI = 6\%$, $PcA = 0\%$.

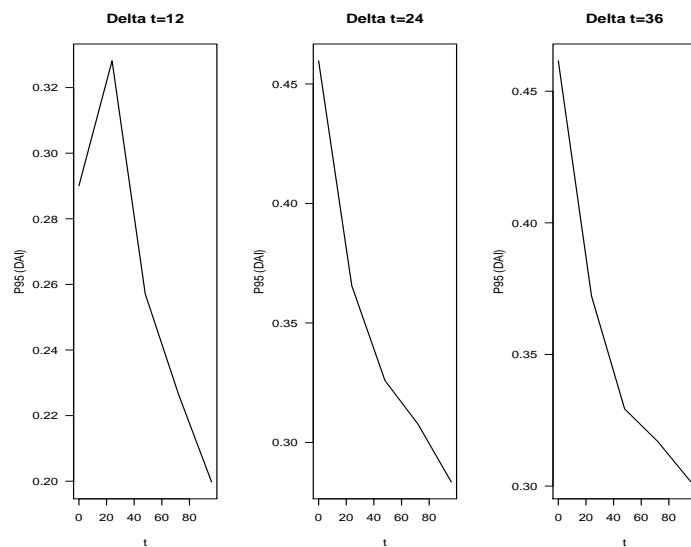


Figura 4.10: (Dados Wei - Modelo BS-Normal): Percentil 95 (ADI), diferentes t para Δt fixo, $n = 1000$, $PcI = 6\%$, $PcA = 0\%$.

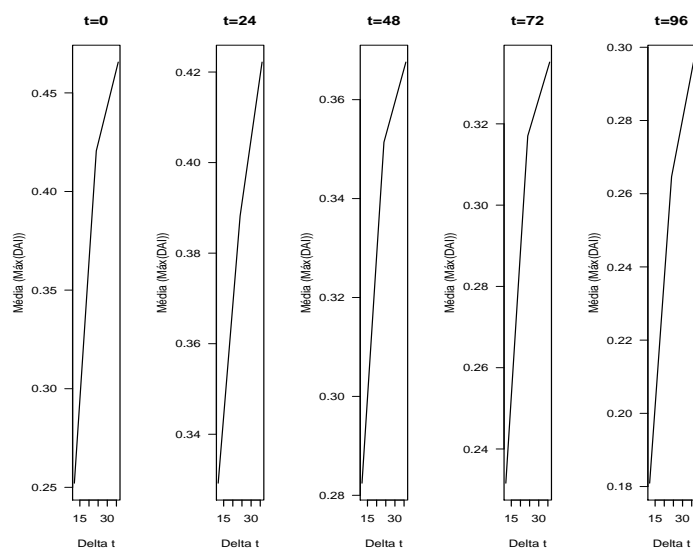


Figura 4.11: (Dados Wei - Modelo BS-Normal): Médias Máx(ADI), diferentes Δt para t fixo, $n = 1000$, $PcI = 6\%$, $PcA = 0\%$.

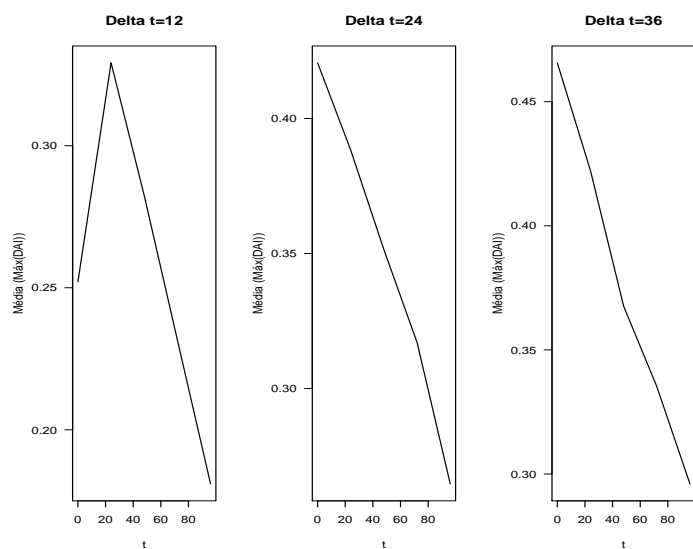


Figura 4.12: (Dados Wei - Modelo BS-Normal): Médias Máx(ADI), diferentes t para Δt fixo, $n = 1000$, $PcI = 6\%$, $PcA = 0\%$.

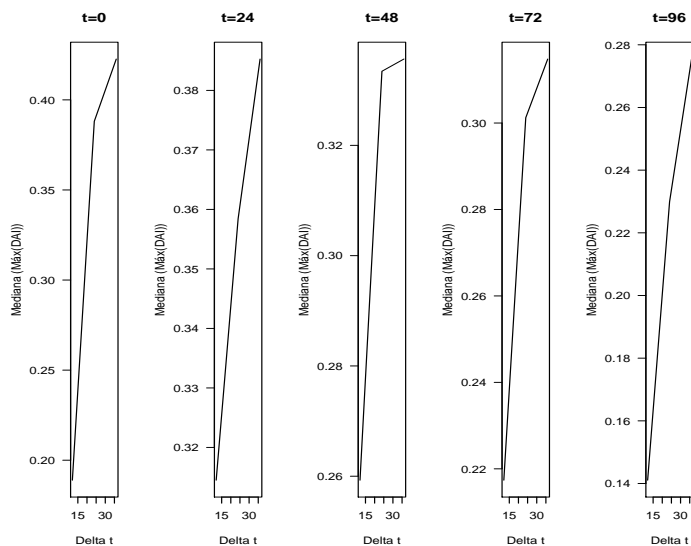


Figura 4.13: (Dados Wei - Modelo BS-Normal): Medianas Máx(ADI), diferentes Δt para t fixo, $n = 1000$, $PcI = 6\%$, $PcA = 0\%$.

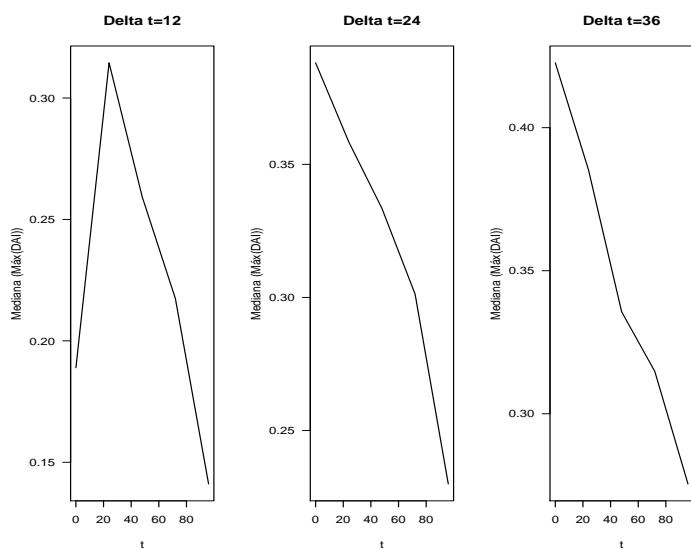


Figura 4.14: (Dados Wei - Modelo BS-Normal): Medianas Máx(ADI), diferentes t para Δt fixo, $n = 1000$, $PcI = 6\%$, $PcA = 0\%$.

Na Figura 4.15 são apresentados os histogramas das diferenças DAI obtidas em todas as amostras geradas no caso Dados BS - Modelo Wei-Normal para cada combinação dos valores de t e Δt e nas Figuras 4.16 e 4.17 são expostos os gráficos dos valores do percentil 95 de tais diferenças. Tanto os valores representados quanto o número de diferenças calculadas em cada situação estão resumidos na Tabela 4.6.

Tabela 4.6: (*Dados BS - Modelo Wei-Normal*): Percentil 95 e quantidade de diferenças (DAI), $n = 1000$, $PcI = 7\%$, $PcA = 0\%$.

	Percentil 95			Quantidade de diferenças
	$\Delta t = 12$	$\Delta t = 24$	$\Delta t = 36$	
$t = 0$	0,1663	0,2641	0,3050	25000
$t = 24$	0,1562	0,2163	0,2357	18372
$t = 48$	0,1550	0,2124	0,2216	11674
$t = 72$	0,1637	0,2257	0,2325	7664
$t = 96$	0,1658	0,2353	0,2478	5238

O comportamento das diferenças DAI para diferentes valores de Δt e t fixo é semelhante ao apresentado nos casos anteriores. No entanto, para Δt fixo e diferentes valores de t observa-se (tanto nos histogramas quanto nos valores do percentil 95) que essas diferenças diminuem de $t = 0$ a $t = 48$, mas aumentam a partir desse momento até $t = 96$. Esse resultado pode estar relacionado com uma competição de informações entre os processos da resposta longitudinal e da resposta de sobrevivência. Conforme são coletadas mais medidas longitudinais para cada indivíduo espera-se uma melhora nas predições dos efeitos aleatórios; no entanto, para etapas mais avançadas do acompanhamento há menos pacientes restantes no estudo, o que leva a que os tempos de sobrevivência estimados em tais momentos apresentem maior variabilidade.

Nas Figuras 4.18 - 4.21 são expostos os gráficos dos valores da média e da mediana dos máximos (Máx(DAI)). Os valores estão resumidos na Tabela 4.7. Observa-se que o comportamento dos valores da média e da mediana dos máximos (Máx(DAI)) é semelhante ao apresentado no caso exposto inicialmente.

Na Figura 4.22 são apresentados os histogramas das diferenças DAI obtidos em todas as amostras geradas no caso Dados Wei - Modelo Wei-Normal para cada combinação dos

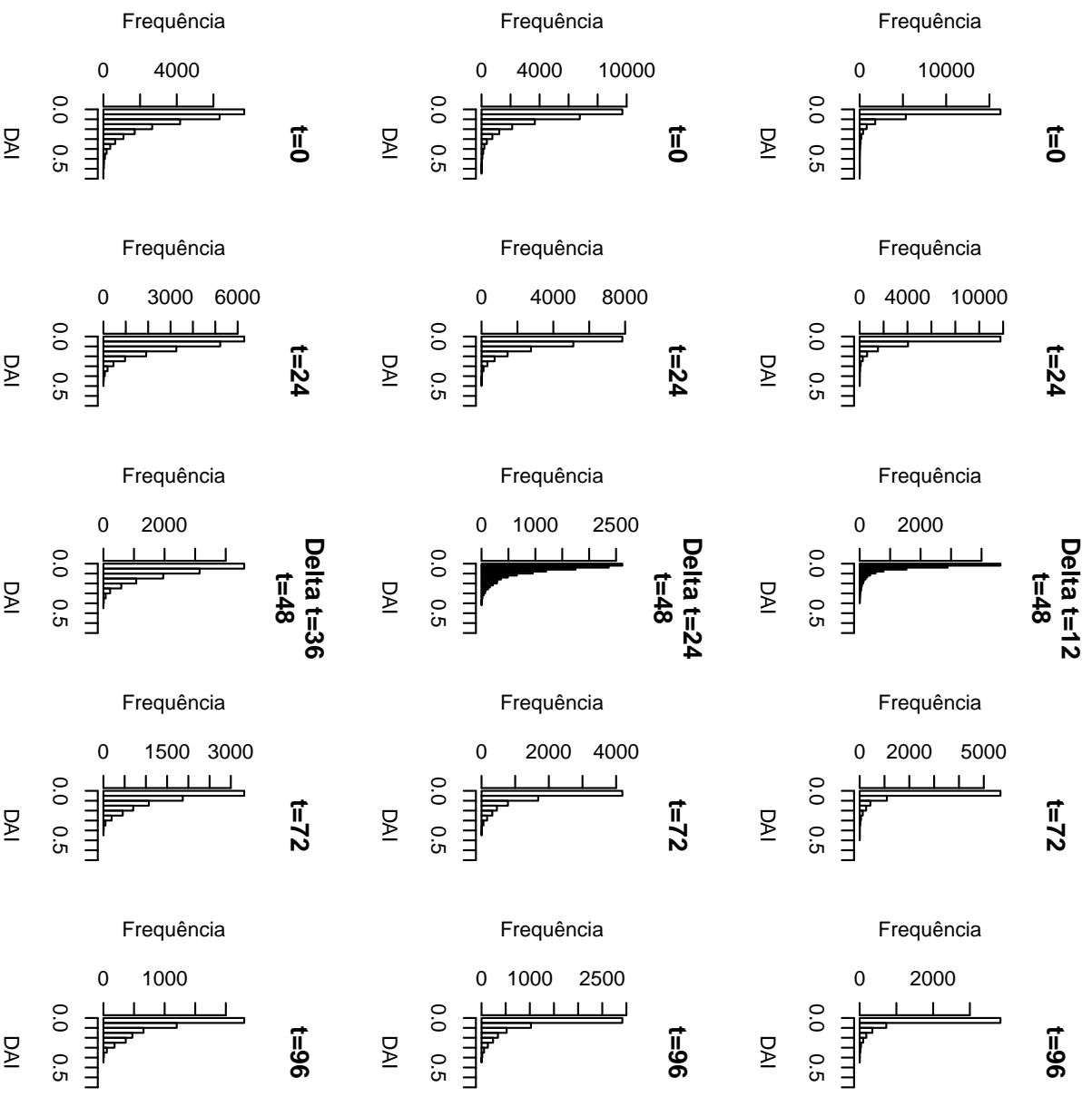


Figura 4.15: (Dados BS - Modelo Wei-Normal): Histograma (ADI), $n = 1000$, $PcI = 7\%$, $PcA = 0\%$

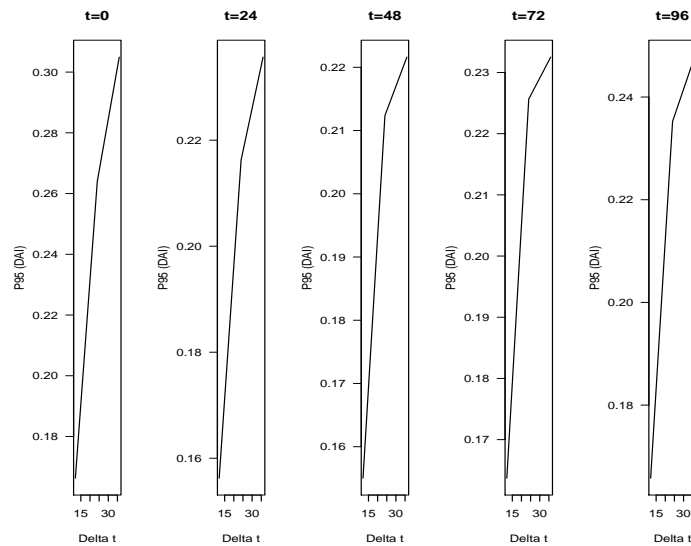


Figura 4.16: (Dados BS - Modelo Wei-Normal): Percentil 95 (ADI), diferentes Δt para t fixo, $n = 1000$, $PcI = 7\%$, $PcA = 0\%$.

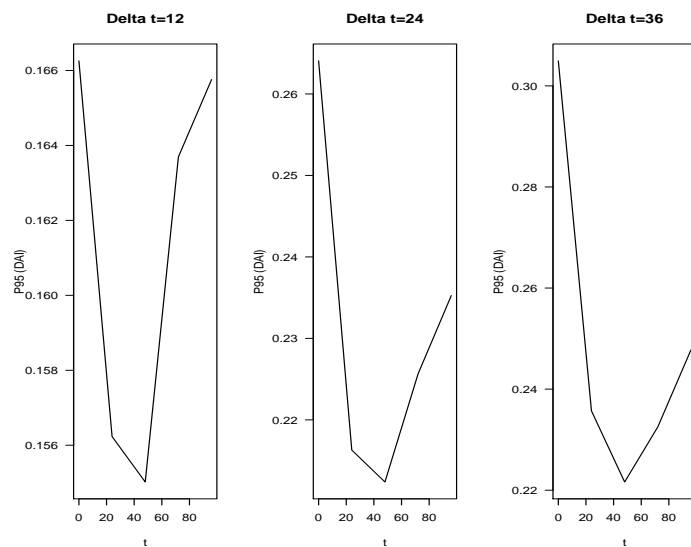


Figura 4.17: (Dados BS - Modelo Wei-Normal): Percentil 95 (ADI), diferentes t para Δt fixo, $n = 1000$, $PcI = 7\%$, $PcA = 0\%$.

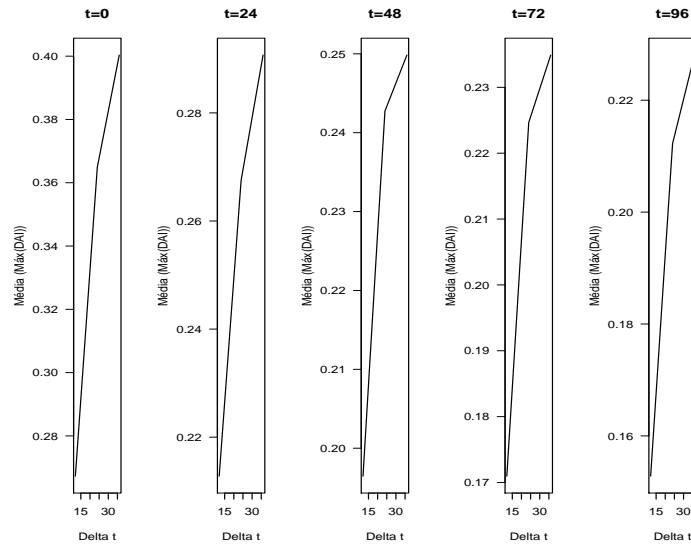


Figura 4.18: (Dados BS - Modelo Wei-Normal): Médias Máx(ADI), diferentes Δt para t fixo, $n = 1000$, $PcI = 7\%$, $PcA = 0\%$.

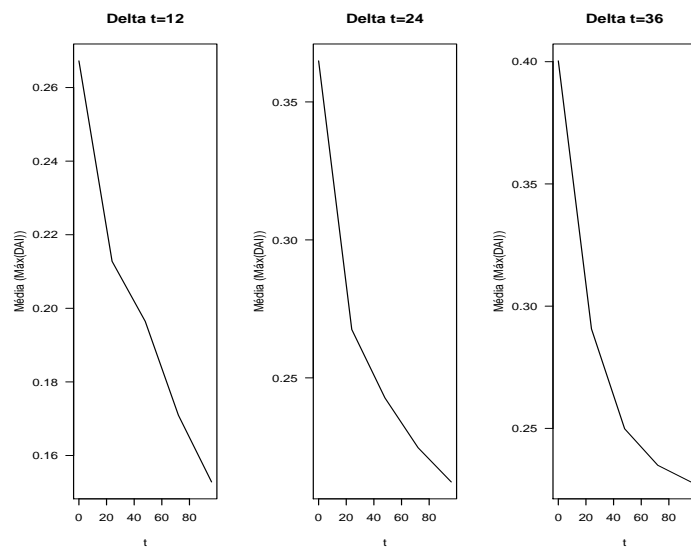


Figura 4.19: (Dados BS - Modelo Wei-Normal): Médias Máx(ADI), diferentes t para Δt fixo, $n = 1000$, $PcI = 7\%$, $PcA = 0\%$.

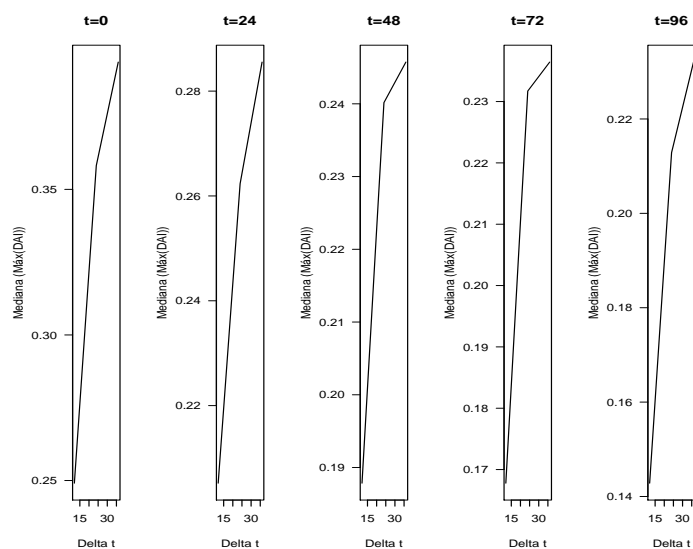


Figura 4.20: (Dados BS - Modelo Wei-Normal): Medianas Máx(ADI), diferentes Δt para t fixo, $n = 1000$, $PcI = 7\%$, $PcA = 0\%$.

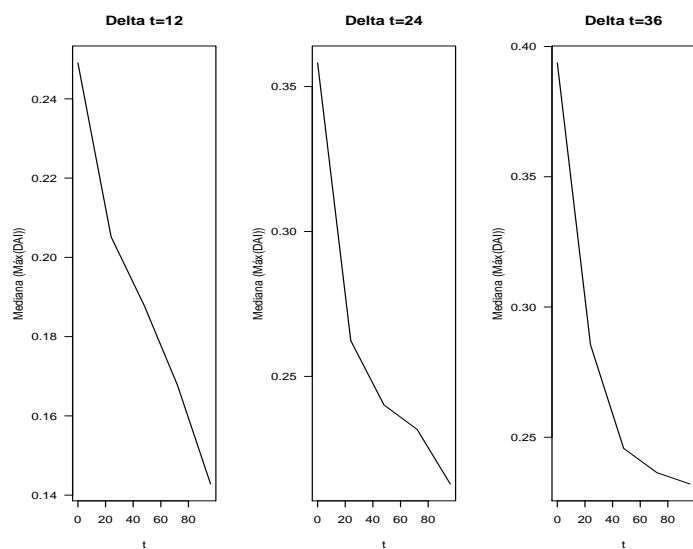


Figura 4.21: (Dados BS - Modelo Wei-Normal): Medianas Máx(ADI), diferentes t para Δt fixo, $n = 1000$, $PcI = 7\%$, $PcA = 0\%$.

Tabela 4.7: (*Dados BS - Modelo Wei-Normal*): Média e mediana ($Máx(DAI)$), $n = 1000$, $PcI = 7\%$, $PcA = 0\%$.

	$\Delta t = 12$		$\Delta t = 24$		$\Delta t = 36$	
	Média	Mediana	Média	Mediana	Média	Mediana
$t = 0$	0,2673	0,2491	0,3649	0,3582	0,4004	0,3937
$t = 24$	0,2128	0,2052	0,2675	0,2623	0,2907	0,2855
$t = 48$	0,1965	0,1878	0,2427	0,2402	0,2498	0,2457
$t = 72$	0,1709	0,1678	0,2247	0,2317	0,2349	0,2364
$t = 96$	0,1528	0,1428	0,2123	0,2129	0,2281	0,2321

valores de t e Δt e nas Figuras 4.23 e 4.24 são expostos os gráficos dos valores do percentil 95 de tais diferenças. Tanto os valores representados quanto o número de diferenças calculadas em cada situação estão resumidos na Tabela 4.8. O comportamento das diferenças DAI para diferentes valores de Δt e t fixo é semelhante ao apresentado nos casos anteriores; no entanto, para $\Delta t = 12$ com t fixo observa-se que o percentil 95 das diferenças é menor quando $t = 0$ do que quando $t = 24$, mas para valores de $t \neq 0$, ainda se observa uma diminuição nas diferenças DAI conforme t aumenta, de acordo com o que é considerado no segundo caso apresentado.

Nas Figuras 4.25 - 4.28 são expostos os gráficos dos valores da média e da mediana dos máximos ($Máx(DAI)$). Os valores estão resumidos na Tabela 4.9. Observa-se que o comportamento é semelhante ao apresentado em termos do percentil 95 das diferenças DAI.

Tabela 4.8: (*Dados Wei - Modelo Wei-Normal*): Percentil 95 e quantidade de diferenças (DAI), $n = 1000$, $PcI = 6\%$, $PcA = 0\%$.

	Percentil 95			Quantidade de diferenças
	$\Delta t = 12$	$\Delta t = 24$	$\Delta t = 36$	
$t = 0$	0,1085	0,3115	0,3956	25000
$t = 24$	0,2635	0,3544	0,3760	19948
$t = 48$	0,2329	0,3150	0,3418	13002
$t = 72$	0,1893	0,2721	0,3101	8178
$t = 96$	0,1703	0,2551	0,2837	5261

Finalmente, com o intuito de avaliar o desempenho do processo de estimação dos parâmetros nos cenários em que o modelo de sobrevivência utilizado para gerar os dados é igual ao considerado no modelo conjunto ajustado, serão expostos inicialmente os resultados dos cenários

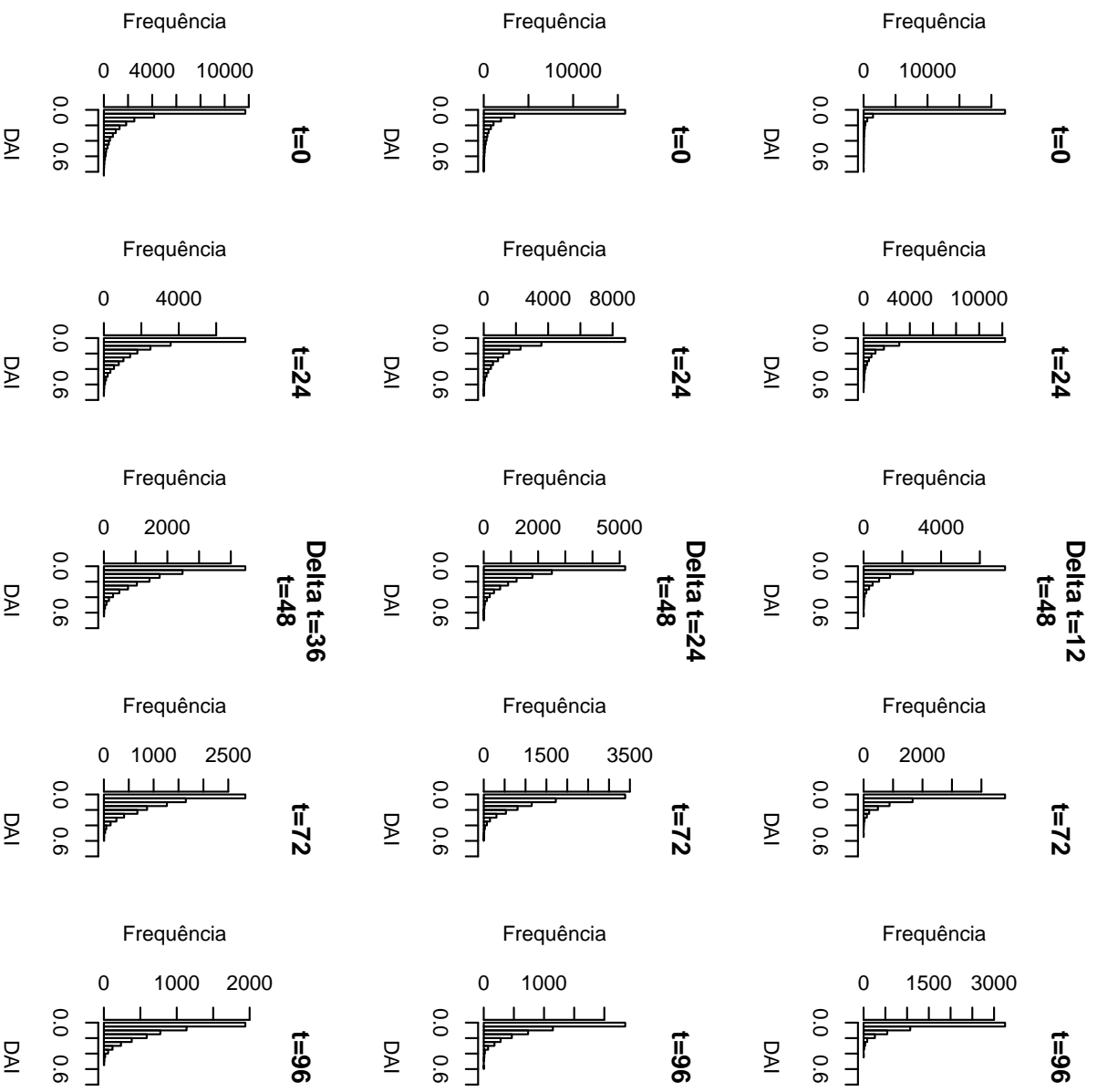


Figura 4.22: (Dados Wei - Modelo Wei-Normal): Histograma (ADI), $n = 1000$, $P_{cI} = 6\%$, $P_{cA} = 0\%$

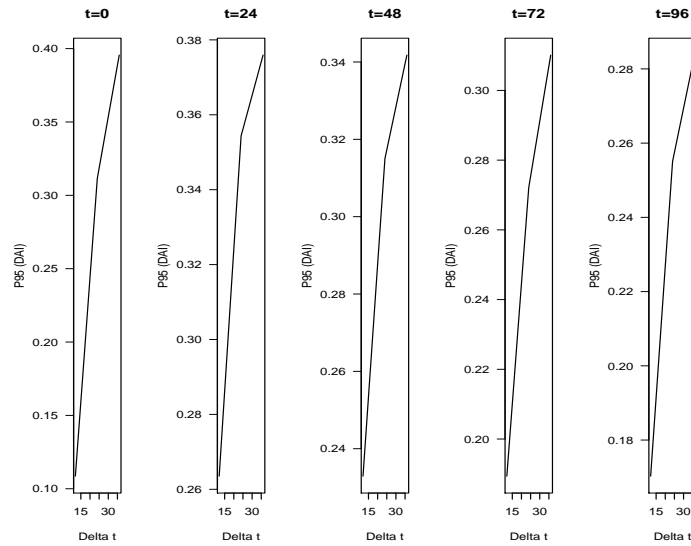


Figura 4.23: (Dados Wei - Modelo Wei-Normal): Percentil 95 (ADI), diferentes Δt para t fixo, $n = 1000$, $PcI = 6\%$, $PcA = 0\%$.

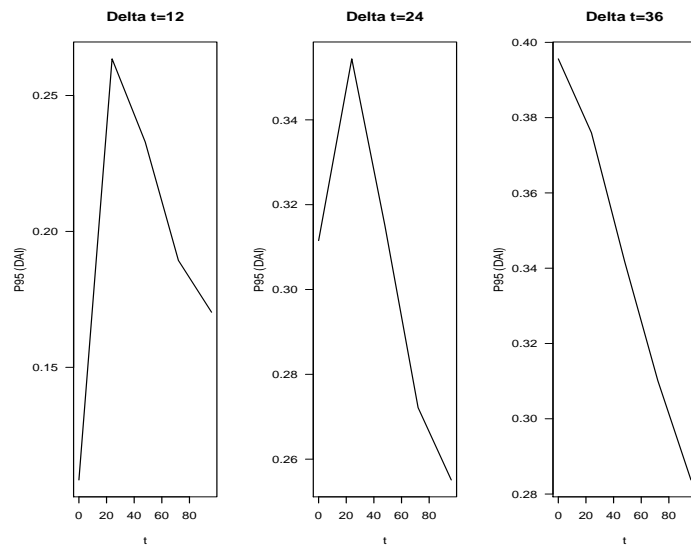


Figura 4.24: (Dados Wei - Modelo Wei-Normal): Percentil 95 (ADI), diferentes t para Δt fixo, $n = 1000$, $PcI = 6\%$, $PcA = 0\%$.

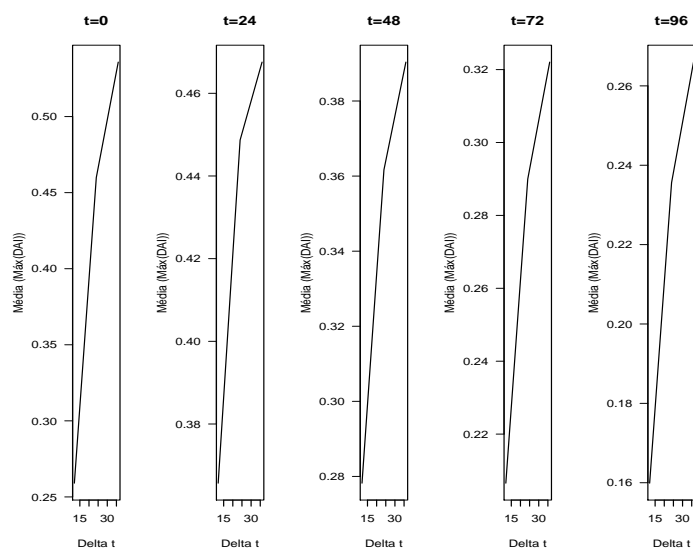


Figura 4.25: (Dados Wei - Modelo Wei-Normal): Médias Máx(ADI), diferentes Δt para t fixo, $n = 1000$, $PcI = 6\%$, $PcA = 0\%$.

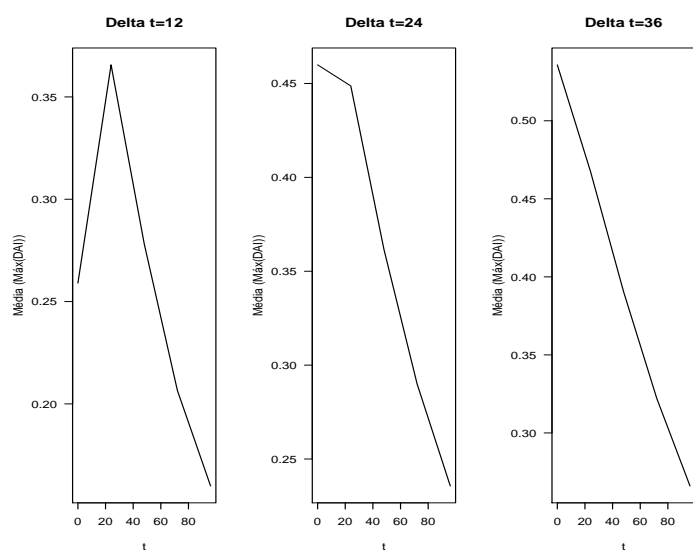


Figura 4.26: (Dados Wei - Modelo Wei-Normal): Médias Máx(ADI), diferentes t para Δt fixo, $n = 1000$, $PcI = 6\%$, $PcA = 0\%$.

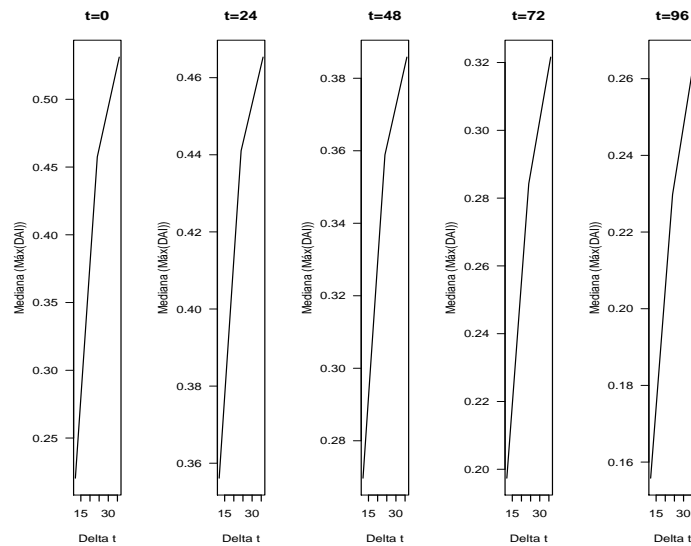


Figura 4.27: (Dados Wei - Modelo Wei-Normal): Medianas Máx(ADI), diferentes Δt para t fixo, $n = 1000$, $PcI = 6\%$, $PcA = 0\%$.

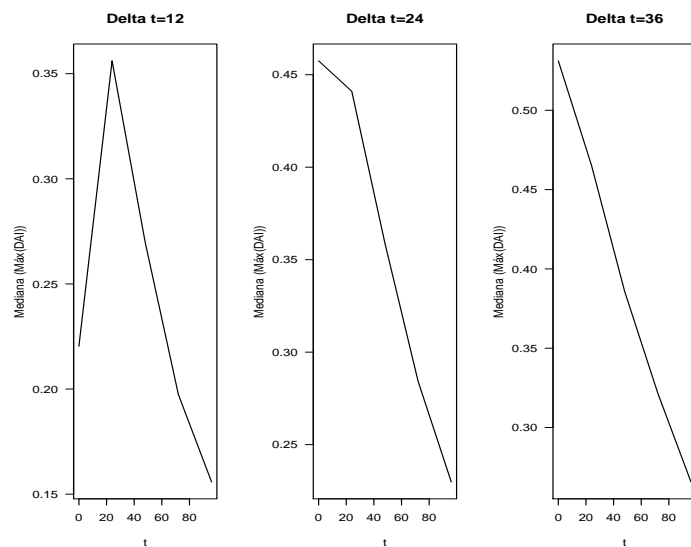


Figura 4.28: (Dados Wei - Modelo Wei-Normal): Medianas Máx(ADI), diferentes t para Δt fixo, $n = 1000$, $PcI = 6\%$, $PcA = 0\%$.

Tabela 4.9: (*Dados Wei - Modelo Wei-Normal*): Média e mediana ($Máx(DAI)$), $n = 1000$, $PcI = 6\%$, $PcA = 0\%$.

	$\Delta t = 12$		$\Delta t = 24$		$\Delta t = 36$	
	Média	Mediana	Média	Mediana	Média	Mediana
$t = 0$	0,2590	0,2203	0,4599	0,4575	0,5358	0,5312
$t = 24$	0,3657	0,3561	0,4487	0,4410	0,4676	0,4654
$t = 48$	0,2782	0,2697	0,3617	0,3589	0,3904	0,3859
$t = 72$	0,2065	0,1974	0,2900	0,2842	0,3221	0,3216
$t = 96$	0,1599	0,1558	0,2356	0,2298	0,2660	0,2656

com o maior tamanho da amostra ($n = 1000$) considerando os diferentes percentuais de censura à direita ($Pc = 0\%, 25\%, 50\%, 75\%$) nos casos: i) Dados BS - Modelo B-S-Normal e, ii) Dados Wei - Modelo Wei-Normal.

Tabela 4.10: (*Dados BS - Modelo BS-Normal*): Medidas das estimativas dos parâmetros, $n = 1000$, $PcI = 7\%$, $PcA = 0\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\alpha)$	-0,4600	-0,3958	0,0642	0,1395	0,1043	0,2267
β_{20}	6,1400	5,4800	-0,6600	-0,1075	0,9277	0,1511
β_{21}	0,1000	0,4792	0,3792	3,7924	0,5610	5,6099
γ	-0,4600	-0,4358	0,0242	0,0525	0,0651	0,1415
β_{10}	5,8000	5,6704	-0,1296	-0,0223	0,2178	0,0376
β_{11}	0,0100	0,0087	-0,0013	-0,1268	0,0015	0,1459
β_{12}	-1,1000	-1,0731	0,0269	0,0244	0,1834	0,1667
$\log(\sigma_0)$	0,4500	0,3816	-0,0684	-0,1519	0,0905	0,2011
$\log(\sigma_e)$	-0,3200	-0,2755	0,0445	0,1391	0,0613	0,1917

Tabela 4.11: (*Dados BS - Modelo BS-Normal*): Medidas das estimativas dos parâmetros, $n = 1000$, $PcI = 7\%$, $PcA = 18\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\alpha)$	-0,4600	-0,3963	0,0637	0,1384	0,0795	0,1728
β_{20}	6,1400	6,2931	0,1531	0,0249	0,4444	0,0724
β_{21}	0,1000	0,1870	0,0870	0,8701	0,3110	3,1095
γ	-0,4600	-0,5364	-0,0764	-0,1661	0,0842	0,1830
β_{10}	5,8000	5,6842	-0,1158	-0,0200	0,2173	0,0375
β_{11}	0,0100	0,0089	-0,0011	-0,1133	0,0013	0,1315
β_{12}	-1,1000	-1,0979	0,0021	0,0020	0,1999	0,1818
$\log(\sigma_0)$	0,4500	0,3860	-0,0640	-0,1423	0,0774	0,1719
$\log(\sigma_e)$	-0,3200	-0,2814	0,0386	0,1206	0,0485	0,1517

Tabela 4.12: (*Dados BS - Modelo BS-Normal*): Medidas das estimativas dos parâmetros, $n = 1000$, $PcI = 7\%$, $PcA = 43\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\alpha)$	-0,4600	-0,2179	0,2421	0,5264	0,2457	0,5340
β_{20}	6,1400	7,2791	1,1391	0,1855	1,1536	0,1879
β_{21}	0,1000	0,0534	-0,0466	-0,4657	0,1137	1,1372
γ	-0,4600	-0,6679	-0,2079	-0,4519	0,2092	0,4549
β_{10}	5,8000	5,7043	-0,0957	-0,0165	0,2063	0,0356
β_{11}	0,0100	0,0084	-0,0016	-0,1581	0,0017	0,1721
β_{12}	-1,1000	-1,0898	0,0102	0,0093	0,1977	0,1798
$\log(\sigma_0)$	0,4500	0,3932	-0,0568	-0,1263	0,0610	0,1355
$\log(\sigma_e)$	-0,3200	-0,2767	0,0433	0,1352	0,0501	0,1566

Tabela 4.13: (*Dados BS - Modelo BS-Normal*): Medidas das estimativas dos parâmetros, $n = 1000$, $PcI = 7\%$, $PcA = 68\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\alpha)$	-0,4600	7,4609	7,9209	17,2193	8,1851	17,7936
β_{20}	6,1400	22,2398	16,0998	2,6221	16,5363	2,6932
β_{21}	0,1000	0,0750	-0,0250	-0,2504	0,1298	1,2976
γ	-0,4600	-0,7224	-0,2624	-0,5704	0,2680	0,5826
β_{10}	5,8000	5,6933	-0,1067	-0,0184	0,2029	0,0350
β_{11}	0,0100	0,0068	-0,0032	-0,3228	0,0033	0,3313
β_{12}	-1,1000	-1,0796	0,0204	0,0185	0,1868	0,1699
$\log(\sigma_0)$	0,4500	0,3845	-0,0655	-0,1455	0,0692	0,1538
$\log(\sigma_e)$	-0,3200	-0,2747	0,0453	0,1416	0,0510	0,1595

Nas Tabelas 4.10-4.13 são apresentados os verdadeiros valores dos parâmetros e os valores da média $\hat{\theta}$ ($MED(\hat{\theta})$), do viés ($VI(\hat{\theta})$), do viés relativo ($VIR(\hat{\theta})$), da raiz do erro quadrático médio ($REQM(\hat{\theta})$) e da raiz do erro quadrático médio relativo ($REQMR(\hat{\theta})$) das estimativas obtidas nos ajustes realizados nas 500 amostras para os diferentes percentuais de censura à direita (de menor a maior) no caso Dados BS - Modelo BS-Normal. Em termos gerais, observou-se que:

- ★ O logaritmo do parâmetro de forma da distribuição Birnbaum-Saunders que foi subestimado, em média, para os casos de menor censura à direita, passou a ser superestimado quando se considerou o maior percentual de tal censura ($Pc = 75\%$), caso em que suas estimativas apresentaram o maior viés relativo e a maior raiz do erro quadrático médio relativo em relação às estimativas dos outros parâmetros do modelo conjunto.

- ★ A média das estimativas do intercepto do componente de sobrevivência aumentou à medida que o percentual de censura à direita foi maior, de acordo com o esperado. Esse intercepto tornou-se consideravelmente superestimado quando se considerou o maior percentual da censura em menção ($P_c = 75\%$).
- ★ As estimativas do coeficiente de regressão para o diagnóstico da ICC apresentaram, em média, o maior viés relativo e a maior raiz do erro quadrático médio relativo em comparação com os obtidos para os outros parâmetros do modelo conjunto, para os casos com menor censura. Uma possível explicação deste fato pode ser a diferença entre as porcentagens de indivíduos diagnosticados com cardiomiopatia chagásica (13%) e indivíduos diagnosticados com outras cardiomiopatias (87%), que pode levar à instabilidade na estimação do parâmetro em questão.
- ★ O parâmetro que mede a associação entre as respostas longitudinal e de sobrevivência foi superestimado, em média, à medida que o percentual de censura à direita aumentou.
- ★ As estimativas dos parâmetros do componente longitudinal mostraram melhores resultados em termos do viés relativo e da raiz do erro quadrático médio relativo em relação aos obtidos para os parâmetros do componente de sobrevivência, que, de acordo com o discutido anteriormente, foram muito afetados pelo percentual de censura à direita considerado em cada caso.

Os resultados obtidos nos cenários em que foram considerados outros tamanhos da amostra, isto é, $n = 500, 250, 100$ e diferentes porcentagens de censura encontram-se no Apêndice F. Em geral, o comportamento das estimativas dos parâmetros do modelo foi semelhante ao descrito previamente. No entanto, é importante notar que as estimativas dos parâmetros relativos ao componente longitudinal mostraram maior estabilidade nos diferentes cenários em comparação com as estimativas dos parâmetros do componente de sobrevivência, afetadas especialmente nos casos com o menor tamanho da amostra ($n = 100$), dentre os quais, o pior desempenho apresentou-se quando foi considerado o maior percentual de censura à direita ($P_c = 75\%$).

Nas Tabelas 4.14-4.17 são expostos os resultados para os diferentes percentuais de censura à direita (de menor a maior) no caso Dados Wei - Modelo Wei-Normal, observando que:

Tabela 4.14: (*Dados Wei - Modelo Wei-Normal*): Medidas das estimativas dos parâmetros, $n = 1000$, $PcI = 6\%$, $PcA = 0\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\varsigma)$	-0,9700	-1,0307	-0,0607	-0,0625	0,0701	0,0723
β_{20}	6,2800	6,2789	-0,0011	-0,0002	0,0765	0,0122
β_{21}	0,1200	0,0919	-0,0281	-0,2339	0,0519	0,4327
γ	-0,4300	-0,4660	-0,0360	-0,0838	0,0378	0,0880
β_{10}	5,8300	5,7232	-0,1068	-0,0183	0,2159	0,0370
β_{11}	0,0100	0,0090	-0,0010	-0,0967	0,0011	0,1115
β_{12}	-1,1100	-1,0907	0,0193	0,0174	0,2043	0,1841
$\log(\sigma_0)$	0,4500	0,4007	-0,0493	-0,1096	0,0540	0,1201
$\log(\sigma_e)$	-0,3300	-0,2970	0,0330	0,1001	0,0380	0,1151

Tabela 4.15: (*Dados Wei - Modelo Wei-Normal*): Medidas das estimativas dos parâmetros, $n = 1000$, $PcI = 6\%$, $PcA = 19\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\varsigma)$	-0,9700	-0,9838	-0,0138	-0,0142	0,0400	0,0413
β_{20}	6,2800	6,6008	0,3208	0,0511	0,3314	0,0528
β_{21}	0,1200	0,0847	-0,0353	-0,2945	0,0595	0,4962
γ	-0,4300	-0,5200	-0,0900	-0,2092	0,0910	0,2116
β_{10}	5,8300	5,7290	-0,1010	-0,0173	0,2215	0,0380
β_{11}	0,0100	0,0088	-0,0012	-0,1164	0,0013	0,1315
β_{12}	-1,1100	-1,0953	0,0147	0,0132	0,2115	0,1906
$\log(\sigma_0)$	0,4500	0,3964	-0,0536	-0,1191	0,0575	0,1279
$\log(\sigma_e)$	-0,3300	-0,2915	0,0385	0,1168	0,0443	0,1343

Tabela 4.16: (*Dados Wei - Modelo Wei-Normal*): Medidas das estimativas dos parâmetros, $n = 1000$, $PcI = 6\%$, $PcA = 44\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\varsigma)$	-0,9700	-0,8909	0,0791	0,0815	0,0919	0,0948
β_{20}	6,2800	7,2711	0,9911	0,1578	0,9986	0,1590
β_{21}	0,1200	0,0813	-0,0387	-0,3224	0,0766	0,6382
γ	-0,4300	-0,6208	-0,1908	-0,4438	0,1917	0,4458
β_{10}	5,8300	5,7159	-0,1141	-0,0196	0,2242	0,0385
β_{11}	0,0100	0,0084	-0,0016	-0,1626	0,0018	0,1752
β_{12}	-1,1100	-1,0841	0,0259	0,0233	0,2107	0,1898
$\log(\sigma_0)$	0,4500	0,3930	-0,0570	-0,1266	0,0610	0,1355
$\log(\sigma_e)$	-0,3300	-0,2859	0,0441	0,1335	0,0493	0,1493

★ A média das estimativas do intercepto do componente de sobrevivência aumentou à medida que o percentual de censura à direita foi maior. Esse intercepto tornou-se superestimado quando se considerou o maior percentual da censura ($Pc = 75\%$).

Tabela 4.17: (*Dados Wei - Modelo Wei-Normal*): Medidas das estimativas dos parâmetros, $n = 1000$, $PcI = 6\%$, $PcA = 69\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\varsigma)$	-0,9700	-0,7362	0,2338	0,2410	0,2431	0,2506
β_{20}	6,2800	8,6905	2,4105	0,3838	2,4204	0,3854
β_{21}	0,1200	0,0668	-0,0532	-0,4434	0,1168	0,9731
γ	-0,4300	-0,8072	-0,3772	-0,8772	0,3784	0,8800
β_{10}	5,8300	5,7144	-0,1156	-0,0198	0,2200	0,0377
β_{11}	0,0100	0,0073	-0,0027	-0,2665	0,0028	0,2750
β_{12}	-1,1100	-1,0754	0,0346	0,0311	0,2100	0,1892
$\log(\sigma_0)$	0,4500	0,3900	-0,0600	-0,1334	0,0645	0,1434
$\log(\sigma_e)$	-0,3300	-0,2821	0,0479	0,1450	0,0535	0,1621

- ★ As estimativas do coeficiente de regressão para o diagnóstico da ICC apresentaram, em média, o maior viés relativo em comparação com o obtido para os outros parâmetros do modelo conjunto, para os casos com menor censura e, a maior raiz do erro quadrático médio relativo em todos os cenários considerados, sendo o parâmetro correspondente sempre subestimado.
- ★ O parâmetro que mede a associação entre as respostas longitudinal e de sobrevivência foi superestimado em todos os casos, sendo que as suas estimativas apresentaram o maior viés relativo em relação aos correspondentes aos outros parâmetros do modelo conjunto em casos com maior censura à direita.
- ★ As estimativas dos parâmetros do componente longitudinal mostraram melhores resultados em termos do viés relativo e da raiz do erro quadrático médio relativo em relação aos obtidos para os parâmetros do componente de sobrevivência, que, foram muito afetados pelo aumento no percentual de censura à direita considerado em cada caso.

O comportamento observado nos cenários em que foram considerados outros tamanhos da amostra, isto é, $n = 500, 250, 100$ e diferentes porcentagens de censura (Ver Apêndice F) foi similar ao descrito anteriormente. No entanto, as estimativas dos parâmetros relativos ao componente longitudinal mostraram uma maior estabilidade nos diferentes cenários em comparação com as estimativas dos parâmetros do componente de sobrevivência, como aconteceu no caso Dados BS - Modelo BS-Normal.

4.2 Aplicação

A metodologia proposta foi aplicada ao conjunto de dados dos 1609 pacientes portadores de ICC atendidos no InCor, dos quais somente em 529 foram realizadas medições da res-

posta longitudinal. Para tal finalidade foram consideradas todas as informações descritas no Capítulo 1, exceto os registros sobre as covariáveis basais que apresentaram alta correlação com outras covariáveis observadas.

Com a finalidade de identificar as covariáveis mais significativas para a descrição de cada um dos componentes (longitudinal e de sobrevivência) do modelo, foram consideradas inicialmente as 24 covariáveis quantitativas (com categorias especificadas no Apêndice A) e foi realizado um processo de seleção considerando a abordagem comumente utilizada na literatura, em que a escolha das covariáveis é efetuada separadamente para cada um dos componentes do modelo por meio do uso de métodos padrão de seleção e uma vez concluído o processo, a associação entre os dois componentes é especificada de acordo com a abordagem adotada para a modelagem conjunta [Ver Wu et al. (2012)]. O processo de seleção das covariáveis para o componente longitudinal foi baseado nas informações dos 529 pacientes com pelo menos uma observação da resposta longitudinal, enquanto que para o componente de sobrevivência foram consideradas as informações de todos os 1609 pacientes. O processo de seleção envolveu as seguintes etapas:

- Etapa 1: Ajustaram-se 24 modelos univariados em que sempre foi incluído o diagnóstico da ICC, nos quais foi avaliada a significância de cada uma das variáveis de interesse separadamente. A significância de cada variável (considerando conjuntamente todas as categorias diferentes à de referência (definida pelos pesquisadores do InCor)), foi avaliada pelo teste da razão de verossimilhanças com nível de significância $\alpha = 0,10$.
- Etapa 2: Para as variáveis significativas na etapa anterior, foi ajustado um modelo multivariado e utilizados testes de Wald para verificar a igualdade entre a categoria de referência e cada uma das outras categorias de cada variável, respectivamente. Para as variáveis com valor-p > 0,05 em todas as categorias, selecionou-se o menor deles e dentre os valores selecionados, foi identificado o maior valor-p e a significância da variável à qual pertencia foi verificada a um nível de $\alpha = 0,05$ (considerando conjuntamente todas as categorias diferentes à de referência), usando o teste da razão de verossimilhanças. Em caso de concluir-se que a variável em questão era não significativa, ela foi excluída, ajustou-se um novo modelo com as variáveis restantes e o procedimento anterior, repetido; caso contrário, a variável em menção foi mantida e o processo descrito previamente foi realizado considerando a variável com o seguinte maior valor-p. Foram ajustados modelos consecutivos até que todas as variáveis fossem significativas segundo o teste da razão de verossimilhanças e apresentassem pelo menos uma categoria (diferente à de faltantes) com valor-p do teste de Wald menor que $\alpha = 0,05$.

Após as duas etapas anteriores foram identificados efeitos significativos na descrição do componente longitudinal para as variáveis: diagnóstico da ICC, diâmetro sistólico do ventrículo esquerdo e diâmetro do átrio esquerdo e; efeitos significativos na explicação do componente de sobrevivência para as variáveis: diagnóstico da ICC, classe funcional, fração de colesterol LDL no sangue, glicemia, creatinina sérica, fração de ejeção do ventrículo esquerdo e ritmo cardíaco. Uma vez especificados os conjuntos de covariáveis significativas em cada caso e para efeitos ilustrativos, foram escolhidas as duas covariáveis que os pesquisadores do InCor julgaram de maior interesse para a descrição tanto do componente longitudinal quanto do componente de sobrevivência e com elas foi ajustado o modelo conjunto. Para o componente longitudinal foram considerados o diagnóstico da ICC e o diâmetro do átrio esquerdo, enquanto que para o componente de sobrevivência foram seleccionados o diagnóstico da ICC e a fração de ejeção do ventrículo esquerdo.

Desta forma, a modelagem conjunta foi realizada considerando as informações referentes a:

- i) uma sequência de medições longitudinais do logaritmo natural dos níveis de BNP no sangue (LBNP), contidas no vetor $\mathbf{y}_i = (y_{i1}, \dots, y_{in_i})^T$, observadas nos tempos especificados no vetor $\mathbf{t}_i = (t_{i1}, \dots, t_{in_i})^T$, $i = 1, \dots, n$;
- ii) um registro do tempo entre a admissão do paciente no estudo e a data de óbito (T_i) ou da censura (C_i) em meses, realizado utilizando as variáveis $Z_i = \min(T_i, C_i)$ e $\delta_i = I(T_i \leq C_i)$;
- iii) valores das variáveis explicativas especificadas anteriormente e contidos nos vetores $\mathbf{x}_{1i}(t)$ e \mathbf{x}_{2i} , em que $\mathbf{x}_{1i}(t) = (1, x_{1i1}(t), x_{1i2}, x_{1i3}, x_{1i4}, x_{1i5}, x_{1i6}, x_{1i7}, x_{1i8})^T$ representa o vetor das covariáveis consideradas no modelo para o componente longitudinal, com $x_{1i1}(t)$ denotando o tempo de observação da resposta longitudinal, $x_{1i2}, x_{1i3}, x_{1i4}, x_{1i5}$ e x_{1i6} denotando as variáveis dummy associadas às categorias (cardiomaopatia dilatada, cardiomiopatia isquêmica, cardiomiopatia hipertensiva, cardiomiopatia alcoólica e outras cardiomiopatias) da variável diagnóstico da ICC e, x_{1i7} e x_{1i8} denotando as variáveis dummy associadas às categorias (aumentado e faltante) da variável diâmetro do átrio esquerdo. Além disso, $\mathbf{x}_{2i} = (1, x_{2i1}, x_{2i2}, x_{2i3}, x_{2i4}, x_{2i5}, x_{2i6}, x_{2i7}, x_{2i8})^T$ representa o vetor das covariáveis consideradas no modelo para o componente de sobrevivência, com $x_{2i1}, x_{2i2}, x_{2i3}, x_{2i4}$ e x_{2i5} denotando as variáveis dummy associadas às categorias (cardiomaopatia dilatada, cardiomiopatia isquêmica, cardiomiopatia hipertensiva, cardiomiopatia alcoólica e outras cardiomiopatias) da variável diagnóstico da ICC e, x_{2i6}, x_{2i7} e x_{2i8} denotando as variáveis dummy associadas às categorias (muito baixa, baixa e faltante) da variável fração de ejeção do ventrículo esquerdo.

A resposta longitudinal para o i -ésimo indivíduo no instante $t \geq 0$, $y_i(t)$, foi descrita por meio do modelo (2.1)-(2.2), com $\beta_1 = (\beta_{10}, \beta_{11}, \beta_{12}, \beta_{13}, \beta_{14}, \beta_{15}, \beta_{16}, \beta_{17}, \beta_{18})^T$. O processo $w_{1i}(t)$ foi caracterizado em termos de um intercepto aleatório específico do i -ésimo indivíduo, considerado invariante no tempo e com distribuição $b_{0i} \sim N(0, \sigma_0^2)$ e, assumiu-se que o erro de medida é tal que $e_i(t) \sim N(0, \sigma_e^2)$ e é independente de b_{0i} , para todo $t \geq 0$.

Além disso, o logaritmo do tempo de sobrevivência ou de censura observado para o i -ésimo indivíduo, $v_i = \log(Z_i)$, foi modelado por meio do modelo de regressão log-linear Birnbaum-Saunders (3.1), em que $\beta_2 = (\beta_{20}, \beta_{21}, \beta_{22}, \beta_{23}, \beta_{24}, \beta_{25}, \beta_{26}, \beta_{27}, \beta_{28})^T$ e os termos ε_i são variáveis aleatórias independentes e identicamente distribuídas, tais que $\varepsilon_i \sim \text{SenhN}(\alpha, 0, 2)$ para $i = 1, \dots, n$.

Para a modelagem conjunta foi considerada tanto a abordagem tradicional, na qual somente são incluídos os indivíduos que têm pelo menos uma medição da resposta longitudinal, em que a que a função de verossimilhança conjunta é da forma dada em (4.1), quanto a abordagem proposta, que acomoda a informação de sobrevivência de todos os indivíduos, para a qual a função de verossimilhança conjunta é definida como em (3.16).

4.2.1 Resultados

As estimativas dos parâmetros e de seus erros padrão obtidas ao fazer o ajuste do modelo conjunto proposto via: i) a abordagem tradicional e ii) a abordagem proposta, foram comparadas com as obtidas ao realizar o ajuste dos respectivos modelos marginais (longitudinal e de sobrevivência). Os resultados da modelagem dos componentes longitudinal e de sobrevivência para cada uma das duas abordagens (estimativas dos parâmetros e de seus erros padrão e valores-p dos testes de Wald), são apresentados nas Tabelas 4.18 e 4.19 e, 4.20 e 4.21, respectivamente.

No caso i) encontrou-se que a associação entre as respostas longitudinal e de sobrevivência foi significativa ao nível de 5%, indicando, como esperado, que quanto maior o nível de LBNP no plasma sanguíneo, menor a sobrevivência do paciente. Além disso, comparando-se os ajustes dos modelos marginal e conjunto, em geral, não foram observadas grandes diferenças nas estimativas dos parâmetros do componente longitudinal, ao contrário dos resultados obtidos quando foi considerado o componente de sobrevivência. No entanto, a maioria dos erros padrão correspondentes, foram menores em ambos componentes quando foi utilizado o modelo conjunto.

No caso ii) verificou-se que a associação entre as duas respostas somente foi significativa ao nível de 10% e deixou de ser negativa, ao contrário do que aconteceu no caso i). Além

disso, quando foram comparados os ajustes dos modelos marginal e conjunto, não se observaram grandes diferenças entre as estimativas dos parâmetros para nenhum dos componentes e, embora os erros padrão para o componente longitudinal foram menores quando foi considerado o modelo conjunto, não foram observadas diferenças importantes nesses erros para o componente de sobrevivência.

Tabela 4.18: *Modelagem do componente longitudinal - Abordagem tradicional (n=529).*

Parâmetro	Modelo Marginal			Modelo Conjunto		
	Estimativa	Erro Padrão	Pr(z)	Estimativa	Erro Padrão	Pr(z)
Intercepto	5,0529	0,2710	<0,0001	4,7027	0,1744	<0,0001
Tempo Obs.	0,0040	0,0021	0,0586	0,0092	0,0018	<0,0001
Diag. C. Dilatada	-1,0525	0,2280	<0,0001	-0,8847	0,1283	<0,0001
Diag. C. Isquêmica	-0,9640	0,2450	0,0001	-0,7222	0,1505	<0,0001
Diag. C. Hipertensiva	-1,2735	0,2268	<0,0001	-1,1674	0,1694	<0,0001
Diag. C. Alcoólica	-0,2730	0,3861	0,4799	-0,2799	0,1975	0,1563
Diag. Outras C.	-0,4967	0,4000	0,2149	-0,1495	0,2116	0,4801
Di.At.Es. Aumentado	1,2543	0,2002	<0,0001	1,2418	0,1359	<0,0001
Di.At.Es. Faltante	1,1106	0,3080	0,0003	1,0863	0,2477	<0,0001
$\log(\sigma_0)$	0,3874	-	-	0,3882	0,0364	<0,0001
$\log(\sigma_e)$	-0,3267	-	-	-0,3266	0,0257	<0,0001

Tabela 4.19: *Modelagem do componente de sobrevivência - Abordagem tradicional (n=529).*

Parâmetro	Modelo Marginal			Modelo Conjunto		
	Estimativa	Erro Padrão	Pr(z)	Estimativa	Erro Padrão	Pr(z)
$\log(\alpha)$	0,0265	0,0881	0,7636	-0,4753	0,0802	<0,0001
Intercepto	5,0137	0,2959	<0,0001	7,1990	0,3427	<0,0001
Diag. C. Dilatada	0,3346	0,1689	0,0476	-0,0518	0,1251	0,6792
Diag. C. Isquêmica	0,3241	0,1786	0,0696	0,1722	0,1346	0,2007
Diag. C. Hipertensiva	0,6687	0,1726	0,0001	0,1698	0,1291	0,1883
Diag. C. Alcoólica	0,4770	0,2905	0,1006	0,3777	0,2148	0,0787
Diag. Outras C.	0,3225	0,2869	0,2609	0,2170	0,2193	0,3224
Fr.Ej.Ve.Es. Muito Baixa	-0,6077	0,2614	0,0201	-0,0620	0,2090	0,7668
Fr.Ej.Ve.Es. Baixa	-0,2235	0,2579	0,3860	0,0041	0,2073	0,9843
Fr.Ej.Ve.Es. Faltante	-0,3338	0,3977	0,4014	-0,1503	0,2962	0,6118
γ	-	-	-	-0,4569	0,0428	<0,0001

Tabela 4.20: *Modelagem do componente longitudinal - Abordagem proposta (n=1609).*

Parâmetro	Modelo Marginal			Modelo Conjunto		
	Estimativa	Erro Padrão	Pr(z)	Estimativa	Erro Padrão	Pr(z)
Intercepto	5,0529	0,2710	<0,0001	5,0808	0,2271	<0,0001
Tempo Obs.	0,0040	0,0021	0,0586	0,0037	0,0023	0,1100
Diag. C. Dilatada	-1,0525	0,2280	<0,0001	-1,0352	0,1561	<0,0001
Diag. C. Isquêmica	-0,9640	0,2450	0,0001	-0,9692	0,1923	<0,0001
Diag. C. Hipertensiva	-1,2735	0,2268	<0,0001	-1,1949	0,1720	<0,0001
Diag. C. Alcoólica	-0,2730	0,3861	0,4799	-0,1960	0,2427	0,4194
Diag. Outras C.	-0,4967	0,4000	0,2149	-0,4280	0,2408	0,0755
Di.At.Es. Aumentado	1,2543	0,2002	<0,0001	1,2047	0,1404	<0,0001
Di.At.Es. Faltante	1,1106	0,3080	0,0003	1,0539	0,2468	<0,0001
$\log(\sigma_0)$	0,3874	-	-	0,3747	0,0364	<0,0001
$\log(\sigma_e)$	-0,3267	-	-	-0,3273	0,0256	<0,0001

Tabela 4.21: *Modelagem do componente de sobrevivência - Abordagem proposta (n=1609).*

Parâmetro	Modelo Marginal			Modelo Conjunto		
	Estimativa	Erro Padrão	Pr(z)	Estimativa	Erro Padrão	Pr(z)
$\log(\alpha)$	1,2506	0,1329	<0,0001	1,2401	0,1355	<0,0001
Intercepto	7,1560	0,3460	<0,0001	7,0577	0,3530	<0,0001
Diag. C. Dilatada	-1,7624	0,1915	<0,0001	-1,6388	0,2011	<0,0001
Diag. C. Isquêmica	0,0945	0,1829	0,6056	0,1261	0,1836	0,4920
Diag. C. Hipertensiva	0,3879	0,1673	0,0204	0,4371	0,1695	0,0099
Diag. C. Alcoólica	0,4768	0,2841	0,0933	0,4898	0,2825	0,0829
Diag. Outras C.	0,1453	0,2734	0,5952	0,1788	0,2735	0,5133
Fr.Ej.Ve.Es. Muito Baixa	-0,9429	0,2135	<0,0001	-1,0043	0,2156	<0,0001
Fr.Ej.Ve.Es. Baixa	-0,1941	0,1991	0,3296	-0,2238	0,1995	0,2619
Fr.Ej.Ve.Es. Faltante	-0,3156	0,2626	0,2294	-0,3049	0,2620	0,2445
γ	-	-	-	0,0358	0,0191	0,0609

A partir dos resultados anteriores pode concluir-se que um aumento no número de indivíduos com medições da resposta longitudinal pode melhorar a evidência da associação entre as respostas longitudinal e de sobrevivência e além disso, pode levar a um aumento na precisão das estimativas dos parâmetros de interesse.

Capítulo 5

Conclusões e sugestões para trabalhos futuros

Nesta tese propõe-se uma metodologia para a modelagem conjunta de dados longitudinais e de sobrevivência, que diferentemente das metodologias propostas na literatura para modelar dados dessa natureza: i) utiliza um modelo paramétrico Birnbaum-Saunders para descrever a resposta de sobrevivência e, ii) leva em consideração a informação de sobrevivência dos indivíduos para os quais a resposta longitudinal não foi observada.

Considera-se uma abordagem baseada na especificação de uma verossimilhança correspondente à distribuição conjunta das respostas longitudinal e de sobrevivência, em que a associação entre as duas respostas é captada por meio de efeitos aleatórios. As estimativas de máxima verossimilhança dos parâmetros do modelo são obtidas por meio da maximização direta da log-verossimilhança correspondente utilizando um algoritmo quase-Newton e as integrais com respeito aos efeitos aleatórios são aproximadas por meio do método da quadratura de Gauss-Hermite. Além de fornecer as estimativas de máxima verossimilhança, a metodologia proposta permite o cálculo dos erros padrão de tais estimativas, necessários para a construção de intervalos de confiança e testes de hipóteses sobre os parâmetros envolvidos no modelo conjunto. A metodologia em questão foi totalmente implementada no ambiente R.

Os resultados obtidos tanto no estudo de simulação quanto na aplicação ao conjunto de dados motivador da pesquisa sugerem que a inclusão de medições longitudinais de uma resposta apropriada pode ser utilizada para melhorar a análise dos dados de sobrevivência. Em particular, um aumento no número de indivíduos com medições da resposta longitudinal pode melhorar a evidência da associação entre as respostas longitudinal e de sobrevivência e pode levar a um aumento na precisão das estimativas dos parâmetros. Além disso, um aumento no número de observações da resposta longitudinal coletadas em um indivíduo pode

melhorar a qualidade da previsão de sua sobrevivência.

Entre as sugestões para futuras pesquisas encontram-se estender a metodologia proposta neste trabalho para os casos em que são consideradas formas funcionais mais sofisticadas para descrever a resposta longitudinal, situações em que é necessária a adição de mais efeitos aleatórios no modelo para capturar a associação entre as respostas longitudinal e de sobrevivência e para estudos em que são coletadas observações de mais de uma variável longitudinal e/ou mais de um tempo de sobrevivência.

Apêndice A

Categorização das variáveis e caracterização da amostra

Tabela A.1: *Gênero.*

Categoria	Frequência	
	Absoluta	Relativa (%)
Maculino	1006	62,5
Feminino	603	37,5
Total	1609	100

Tabela A.2: *Diagnóstico.*

Categoria	Frequência	
	Absoluta	Relativa (%)
Chagásica	186	11,6
Dilatada	419	26,0
Isquêmica	302	18,8
Hipertensiva	554	34,4
Alcoólica	71	4,4
Outras	77	4,8
Total	1609	100

Tabela A.3: *Classe funcional.*

Categoria	Frequência	
	Absoluta	Relativa (%)
I	247	15,4
II	692	43,0
III	424	26,4
IV	169	10,5
Faltante	77	4,8
Total	1609	100

Tabela A.4: *Hipertensão arterial.*

Categoria	Frequência	
	Absoluta	Relativa (%)
Não	390	24,2
Sim	1169	72,7
Faltante	50	3,1
Total	1609	100

Tabela A.5: *Etilismo.*

Categoria	Frequência	
	Absoluta	Relativa (%)
Não	811	50,4
Sim	238	14,8
Ex-etilista	226	14,0
Faltante	334	20,8
Total	1609	100

Tabela A.6: *Tabagismo.*

Categoria	Frequência	
	Absoluta	Relativa (%)
Não	660	41,0
Sim	319	19,8
Ex-tabagista	486	30,2
Faltante	144	8,9
Total	1609	100

Tabela A.7: *Diabetes melito.*

Categoria	Frequência	
	Absoluta	Relativa (%)
Não	1165	72,4
Sim, depende da insulina	64	4,0
Sim, não depende da insulina	284	17,7
Faltante	96	6,0
Total	1609	100

Tabela A.8: *Ritmo cardíaco.*

Categoria	Frequência	
	Absoluta	Relativa (%)
Sinusal	1004	62,4
Fibrilação atrial	212	13,2
Outros	269	16,7
Faltante	124	7,7
Total	1609	100

Tabela A.9: *Idade (anos).*

Valor	Frequência	
	Absoluta	Relativa (%)
[18, 35)	88	5,5
[35, 45)	199	12,4
[45, 55)	439	27,3
[55, 65)	438	27,2
[65, 93]	445	27,7
Total	1609	100

Tabela A.10: *Índice de massa corporal (kg/m^2).*

Classificação	Valor	Frequência	
		Absoluta	Relativa (%)
Abaixo do peso ou normal	Menor que 25,0	527	32,8
Acima do peso	[25,0; 30,0)	479	29,8
Obeso	30,0 ou mais	352	21,9
Faltante		251	15,6
Total		1609	100

Tabela A.11: *Pressão arterial sistólica ($mmHg$).*

Classificação	Valor	Frequência	
		Absoluta	Relativa (%)
Ótima	Menor que 120	301	18,7
Normal	[120, 130)	249	15,5
Limítrofe	[130, 140)	228	14,2
Hipertensão	140 ou mais	774	48,1
Faltante		57	3,5
Total		1609	100

Tabela A.12: *Frequência cardíaca (em batimentos por minuto).*

Classificação	Valor	Frequência	
		Absoluta	Relativa (%)
Baixa ou Normal	Menor que 100	1195	74,3
Alta	100 ou mais	217	13,5
Faltante		197	12,2
Total		1609	100

Tabela A.13: *Potássio sérico (mEq/L).*

Classificação	Valor	Frequência	
		Absoluta	Relativa (%)
Baixo ou normal	5,0 ou menos	1269	78,9
Alto	Maior que 5,0	263	16,3
Faltante		77	4,8
Total		1609	100

Tabela A.14: *Hemoglobina sérico (mg/dL).*

Classificação	Valor	Frequência	
		Absoluta	Relativa (%)
Baixo	Menor que 13	338	21,0
Normal ou alto	13 ou mais	1199	74,5
Faltante		72	4,5
Total		1609	100

Tabela A.15: *Taxa de leucócitos (mm³).*

Classificação	Valor	Frequência	
		Absoluta	Relativa (%)
Baixo ou normal	11000 ou menos	1410	87,6
Alto	Maior que 11000	107	6,7
Faltante		92	5,7
Total		1609	100

Tabela A.16: *Taxa de triglicérides (mg/dL).*

Classificação	Valor	Frequência	
		Absoluta	Relativa (%)
Ótimo	Menor que 150	920	57,2
Limítrofe	[150, 200)	207	12,9
Alto	200 ou mais	241	15,0
Faltante		241	15,0
Total		1609	100

Tabela A.17: *Fração de colesterol LDL (mg/dL).*

Classificação	Valor	Frequência	
		Absoluta	Relativa (%)
Ótimo	Menor que 100	436	27,1
Desejável	[100, 130)	425	26,4
Limítrofe	[130, 160)	286	17,8
Alto	160 ou mais	182	11,3
Faltante		280	17,4
Total		1609	100

Tabela A.18: *Glicemia (mg/dL).*

Classificação	Valor	Frequência	
		Absoluta	Relativa (%)
Baixa ou normal	Menor que 100	725	45,1
Inapropriada	[100, 126)	477	29,6
Elevada	126 ou mais	275	17,1
Faltante		132	8,2
Total		1609	100

Tabela A.19: *Creatinina sérica (mg/dL).*

Classificação	Valor	Frequência	
		Absoluta	Relativa (%)
Baixa	Menor que 0,8	173	10,8
Normal	[0,8; 1,4)	1099	68,3
Alta ou muito alta	1,4 ou mais	281	17,5
Faltante		56	3,5
Total		1609	100

Tabela A.20: *Espessura da parede posterior do ventrículo esquerdo (mm).*

Classificação	Valor	Frequência	
		Absoluta	Relativa (%)
Normal	[6, 10)	867	53,9
Aumentada	10 ou mais	572	35,6
Faltante		170	10,6
Total		1609	100

Tabela A.21: *Diâmetro diastólico do ventrículo esquerdo (mm).*

Classificação	Valor	Frequência	
		Absoluta	Relativa (%)
Normal	[39, 54)	211	13,1
Aumentado	54 ou mais	1195	74,3
Faltante		203	12,6
Total		1609	100

Tabela A.22: *Diâmetro sistólico do ventrículo esquerdo (mm).*

Classificação	Valor	Frequência	
		Absoluta	Relativa (%)
Normal	[22, 38)	96	6,0
Aumentado	38 ou mais	947	58,9
Faltante		566	35,2
Total		1609	100

Tabela A.23: *Fração de ejeção do ventrículo esquerdo (%).*

Classificação	Valor	Frequência	
		Absoluta	Relativa (%)
Muito baixa	Menor que 30	528	32,8
Baixa	[30; 55)	830	51,6
Normal	55 ou mais	128	8,0
Faltante		123	7,6
Total		1609	100

Tabela A.24: *Diâmetro do átrio esquerdo (mm).*

Classificação	Valor	Frequência	
		Absoluta	Relativa (%)
Normal	[27, 39)	241	15,0
Aumentado	39 ou mais	1161	72,2
Faltante		207	12,9
Total		1609	100

Apêndice B

Resultados dos testes Log-Rank

Tabela B.1: *Resultados dos testes Log-Rank para as variáveis explicativas.*

Variável	Valor da estatística de teste	Valor-p
Idade	48,0	<0,0001
Gênero	5,3	0,0209
Índice de massa corporal	6,4	0,0946
Pressão arterial sistólica	17,8	0,0014
Frequência cardíaca	2,8	0,2440
Diagnóstico	25,9	<0,0001
Classe funcional	10,8	0,0283
Hipertensão arterial	2,9	0,2340
Etilismo	4,5	0,2120
Tabagismo	3,0	0,3910
Diabetes melito	10,1	0,0181
Potássio sérico	14,5	0,0007
Hemoglobina sérica	6,3	0,0426
Quantidade de leucócitos no sangue	1,6	0,4600
Taxa de triglicérides no sangue ¹	21,9	<0,0001
Fração de colesterol LDL no sangue	20,3	0,0004
Glicemia ¹	16,4	0,0009
Creatinina sérica	31,6	<0,0001
Espessura da parede posterior do ventrículo esquerdo ¹	8,8	0,0125
Diâmetro diastólico do ventrículo esquerdo	2,8	0,2420
Diâmetro sistólico do ventrículo esquerdo	1,4	0,5060
Fração de ejeção do ventrículo esquerdo	26,5	<0,0001
Diâmetro do átrio esquerdo	9,8	0,0073
Ritmo Cardíaco	9,1	0,0280

1

¹Variável com valor-p < 0,10 devido à diferença entre a curva de sobrevivência da categoria faltante e as demais e, portanto, é considerada pouco influente na sobrevivência dos pacientes em estudo.

Apêndice C

Aspectos computacionais

C.1 Ajuste dos modelos marginais

C.1.1 Ajuste do modelo para o componente longitudinal

O ajuste do modelo linear de efeitos mistos Normal (2.1) considerado para descrever a resposta longitudinal para o i -ésimo indivíduo no instante $t \geq 0$, $y_i(t)$, foi realizado utilizando a função `lme` do pacote `nlme` do R (Pinheiro et al., 2013). As estimativas obtidas por meio do ajuste deste modelo marginal são consideradas como os valores iniciais dos parâmetros correspondentes ao componente longitudinal para realizar o ajuste do modelo conjunto de interesse.

C.1.2 Ajuste do modelo para o componente de sobrevivência

Para o ajuste dos modelos de regressão considerados para descrever o logaritmo do tempo de sobrevivência ou de censura observado para o i -ésimo indivíduo, $v_i = \log(Z_i)$, tem-se que:

- i) O ajuste do modelo log-linear Birnbaum-Saunders (3.1) foi implementado no ambiente R fazendo uma maximização direta da respectiva função de log-verossimilhança (especificada em termos das funções de densidade de probabilidade e de sobrevivência da distribuição `SenhN` associada), por meio do uso do algoritmo quase-Newton implementado na função `nlminb` do pacote `stats` do R, utilizando como valor inicial do parâmetro de forma a estimativa obtida ao fazer o ajuste do modelo log-linear Birnbaum-Saunders para o caso de observações não-censuradas [Ver Galea, Leiva e Paula, 2004] e, como valores iniciais para os coeficientes de regressão as estimativas de mínimos quadrados, tendo em vista as considerações de Leiva et al. (2007);

- ii) O ajuste do modelo log-linear Weibull (descrito no Apêndice E) foi realizado fazendo uso da função `survreg` do pacote `survival` do R (Therneau e Lumley, 2012).

Dependendo de qual destes dois modelos marginais é escolhido para descrever a resposta de sobrevivência, as estimativas obtidas por meio do respectivo ajuste são consideradas como os valores iniciais dos parâmetros relativos ao componente de sobrevivência para realizar o ajuste do modelo conjunto correspondente.

C.2 Ajuste dos modelos conjuntos

Os ajustes dos modelos conjuntos: i) log-linear Birnbaum-Saunders para o componente de sobrevivência e linear de efeitos mistos Normal para o componente longitudinal (BS-Normal) ou, ii) log-linear Weibull para o componente de sobrevivência e linear de efeitos mistos Normal para o componente longitudinal (Wei-Normal), foram implementados no ambiente R fazendo uma maximização direta da respectiva função de log-verossimilhança conjunta, por meio do uso da função `mle2` do pacote `bbmle` do R (Bolker e R Development Core Team, 2013), especificando a opção `nlminb` como o otimizador a ser utilizado, considerando-se como valores iniciais dos parâmetros as estimativas obtidas por meio dos ajustes marginais dos modelos longitudinal e de sobrevivência respectivos e o valor zero para o parâmetro que mede a associação entre as duas respostas. A aproximação das integrais com respeito aos efeitos aleatórios b_{0i} , $i = 1, \dots, n$, é realizada utilizando o método de quadratura de Gauss-Hermite (QGH) unidimensional descrito a seguir, em que, de acordo com o considerado no pacote `JM` do R (Rizopoulos, 2012c) quando este é o método usado para aproximar as integrais em questão, o número de pontos de quadratura é igual a 15 para $N < 2000$ e 9 em caso contrário, em que N representa o número total de medições longitudinais no conjunto de dados analisado. Para a obtenção dos pontos e dos pesos da quadratura é utilizada a função `gauss.quad` do pacote `statmod` do R (Smyth et al., 2013).

C.2.1 Quadratura de Gauss-Hermite unidimensional

Os métodos de integração numérica utilizados para aproximar integrais que não podem ser resolvidas analiticamente, consistem em expressar a função de interesse $g(x) : \mathfrak{R} \rightarrow \mathfrak{R}$ como o produto de uma função de ponderação $w(x) : \mathfrak{R} \rightarrow \mathfrak{R}^+$ e outra função $h(x) : \mathfrak{R} \rightarrow \mathfrak{R}$, de modo que a integral seja aproximada por uma soma ponderada da forma

$$\int_{-\infty}^{\infty} g(x) dx = \int_{-\infty}^{\infty} w(x) h(x) dx \approx \sum_{k=1}^Q w_k h(s_k),$$

em que s_k e w_k denotam os nós e os pesos da regra de integração, respectivamente e, Q denota a ordem da aproximação. Ver, por exemplo, Davis e Rabinowits (1975) e Cohen (2011).

O método da quadratura de Gauss-Hermite (QGH) unidimensional aproxima integrais do tipo

$$\int_{-\infty}^{\infty} \exp\{-x^2\}h(x)dx,$$

por meio de

$$\int_{-\infty}^{\infty} \exp\{-x^2\}h(x)dx \approx \sum_{k=1}^Q w_k h(s_k), \quad (\text{C.1})$$

em que s_k é o k -ésimo zero do polinômio de Hermite de ordem Q , dado por

$$H_Q(x) = (-1)^Q \exp\{x^2\} \frac{d^Q}{dx^Q} (\exp\{x^2\}),$$

e,

$$w_k = \frac{2^{Q-1} Q! \sqrt{\pi}}{Q^2 (H_{Q-1}(s_k))^2}.$$

Uma característica da QGH é que os pontos da quadratura são escolhidos com base em $w(x)$ independentemente da forma da função $g(x)$, de modo que podem ou não estar dentro da região em que a função a ser aproximada tem valores mais altos. Assim, quanto mais próxima $g(x)$ for do núcleo da função de densidade de uma distribuição normal padrão, melhor será a aproximação obtida. Uma extensão da técnica QGH é a versão adaptativa (QGHA) (Liu e Pierce (1994) e Pinheiro e Bates (1995)), que consiste em centrar e escalonar os pontos de quadratura em torno da área de maior densidade da função $g(x)$ com o intuito de melhorar a aproximação da integral. Embora geralmente requeira uma menor quantidade de pontos para se obter um erro de aproximação da mesma ordem de grandeza daquele obtido com a QGH usual, o custo computacional da QGHA é maior, devido à otimização necessária para a transformação dos pontos da quadratura. No contexto dos modelos conjuntos de dados longitudinais e de sobrevivência, Rizopoulos (2012ab) propôs recentemente um método denominado quadratura de Gauss-Hermite pseudo-adaptativa (QGHPA), que da mesma forma que o método QGH, encontra-se implementado no pacote JM do R (Rizopoulos, 2012c) para casos em que a resposta de sobrevivência é descrita por meio do modelo de Cox.

Apêndice D

Função escore

Com base em (3.20) a função escore aproximada pode ser expressa como

$$\begin{aligned} U(\boldsymbol{\theta}) &= \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \sum_{i=1}^n \omega_i \frac{\partial \log D_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} + \sum_{i=1}^n (1 - \omega_i) \frac{\partial \log G_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \\ &= \sum_{i=1}^n \omega_i U_{i1}(\boldsymbol{\theta}) + \sum_{i=1}^n (1 - \omega_i) U_{i2}(\boldsymbol{\theta}), \end{aligned} \quad (\text{D.1})$$

em que

$$D_i(\boldsymbol{\theta}) = \sum_{k=1}^Q \frac{w_k}{\sqrt{\pi}} A_{ik}(\boldsymbol{\theta}) B_{ik}(\boldsymbol{\theta}) C_{ik}(\boldsymbol{\theta}), \quad G_i(\boldsymbol{\theta}) = E_i(\boldsymbol{\theta}) F_i(\boldsymbol{\theta}),$$

e,

$$U_{i1}(\boldsymbol{\theta}) = \frac{\partial \log D_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = D_i^{-1}(\boldsymbol{\theta}) \frac{\partial D_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}, \quad U_{i2}(\boldsymbol{\theta}) = \frac{\partial \log G_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = G_i^{-1}(\boldsymbol{\theta}) \frac{\partial G_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}.$$

Assim,

$$\frac{\partial D_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \sum_{k=1}^Q \frac{w_k}{\sqrt{\pi}} \frac{\partial (A_{ik}(\boldsymbol{\theta}) B_{ik}(\boldsymbol{\theta}) C_{ik}(\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}} = \sum_{k=1}^Q \frac{w_k}{\sqrt{\pi}} H_{ik}(\boldsymbol{\theta}),$$

com

$$H_{ik}(\boldsymbol{\theta}) = \frac{\partial (A_{ik}(\boldsymbol{\theta}) B_{ik}(\boldsymbol{\theta}) C_{ik}(\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}} = A_{ik}(\boldsymbol{\theta}) B_{ik}(\boldsymbol{\theta}) C_{ik}(\boldsymbol{\theta}) J_{ik}(\boldsymbol{\theta}),$$

$$\begin{aligned}
J_{ik}(\boldsymbol{\theta}) &= \delta_i \left(\kappa_{ik1}(v_i)^{-1} \frac{\partial \kappa_{ik1}(v_i)}{\partial \boldsymbol{\theta}} - \kappa_{ik2}(v_i) \frac{\partial \kappa_{ik2}(v_i)}{\partial \boldsymbol{\theta}} \right) \\
&+ (1 - \delta_i) \frac{1}{1 - \Phi(\kappa_{ik2}(v_i))} \left(-\Phi'(\kappa_{ik2}(v_i)) \frac{\partial \kappa_{ik2}(v_i)}{\partial \boldsymbol{\theta}} + \frac{\partial \kappa_{ik3}}{\partial \boldsymbol{\theta}} \right) \\
&= \delta_i \left(\kappa_{ik1}(v_i)^{-1} \frac{\partial \kappa_{ik1}(v_i)}{\partial \boldsymbol{\theta}} - \kappa_{ik2}(v_i) \frac{\partial \kappa_{ik2}(v_i)}{\partial \boldsymbol{\theta}} \right) - (1 - \delta_i) h(\kappa_{ik2}(v_i)) \frac{\partial \kappa_{ik2}(v_i)}{\partial \boldsymbol{\theta}} + \frac{\partial \kappa_{ik3}}{\partial \boldsymbol{\theta}},
\end{aligned} \tag{D.2}$$

em que

$$h(\kappa_{ik2}(v_i)) = \frac{\Phi'(\kappa_{ik2}(v_i))}{1 - \Phi(\kappa_{ik2}(v_i))}, \quad \Phi'(\kappa_{ik2}(v_i)) = \frac{d\Phi(\kappa_{ik2}(v_i))}{d\kappa_{ik2}(v_i)} = \phi(\kappa_{ik2}(v_i)),$$

e,

$$\begin{aligned}
\kappa_{ik3} &= \log \left[\prod_{j=1}^{n_i} \frac{1}{\sqrt{2\pi}\sigma_e} \exp \left\{ -\frac{(y_i(t_{ij}) - (\mathbf{x}_{1i}(t_{ij})^T \boldsymbol{\beta}_1 + \sqrt{2}\sigma_0 s_k))^2}{2\sigma_e^2} \right\} \right] \\
&= -\frac{n_i}{2} \log(2\pi) - \frac{n_i}{2} \log(\sigma_e^2) - \sum_{j=1}^{n_i} \left(\frac{(y_i(t_{ij}) - (\mathbf{x}_{1i}(t_{ij})^T \boldsymbol{\beta}_1 + \sqrt{2}\sigma_0 s_k))^2}{2\sigma_e^2} \right).
\end{aligned}$$

Além disso,

$$\frac{\partial G_i(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \frac{\partial (E_i(\boldsymbol{\theta}) F_i(\boldsymbol{\theta}))}{\partial \boldsymbol{\theta}} = E_i(\boldsymbol{\theta}) F_i(\boldsymbol{\theta}) M_i(\boldsymbol{\theta}),$$

com

$$\begin{aligned}
M_i(\boldsymbol{\theta}) &= \delta_i \left(\xi_{i1}(v_i)^{-1} \frac{\partial \xi_{i1}(v_i)}{\partial \boldsymbol{\theta}} - \xi_{i2}(v_i) \frac{\partial \xi_{i2}(v_i)}{\partial \boldsymbol{\theta}} \right) \\
&+ (1 - \delta_i) \frac{1}{1 - \Phi(\xi_{i2}(v_i))} \left(-\Phi'(\xi_{i2}(v_i)) \frac{\partial \xi_{i2}(v_i)}{\partial \boldsymbol{\theta}} \right) \\
&= \delta_i \left(\xi_{i1}(v_i)^{-1} \frac{\partial \xi_{i1}(v_i)}{\partial \boldsymbol{\theta}} - \xi_{i2}(v_i) \frac{\partial \xi_{i2}(v_i)}{\partial \boldsymbol{\theta}} \right) - (1 - \delta_i) h(\xi_{i2}(v_i)) \frac{\partial \xi_{i2}(v_i)}{\partial \boldsymbol{\theta}},
\end{aligned} \tag{D.3}$$

em que

$$h(\xi_{i2}(v_i)) = \frac{\phi(\xi_{i2}(v_i))}{1 - \Phi(\xi_{i2}(v_i))},$$

portanto,

$$U_{i1}(\boldsymbol{\theta}) D_i^{-1}(\boldsymbol{\theta}) = \sum_{k=1}^Q \frac{w_k}{\sqrt{\pi}} A_{ik}(\boldsymbol{\theta}) B_{ik}(\boldsymbol{\theta}) C_{ik}(\boldsymbol{\theta}) J_{ik}(\boldsymbol{\theta}),$$

e,

$$U_{i2}(\boldsymbol{\theta}) = G_i^{-1}(\boldsymbol{\theta})G_i(\boldsymbol{\theta})M_i(\boldsymbol{\theta}) = M_i(\boldsymbol{\theta}),$$

com $A_{ik}(\boldsymbol{\theta})$, $B_{ik}(\boldsymbol{\theta})$, $C_{ik}(\boldsymbol{\theta})$, $J_{ik}(\boldsymbol{\theta})$ e $M_i(\boldsymbol{\theta})$ definidos como em (3.21), (D.2) e (D.3), respectivamente.

Derivando $l(\boldsymbol{\theta})$ em relação a α , tem-se que

$$U_{i1\alpha}(\boldsymbol{\theta}) = D_i^{-1}(\boldsymbol{\theta}) \sum_{k=1}^Q \frac{w_k}{\sqrt{\pi}} A_{ik}(\boldsymbol{\theta}) B_{ik}(\boldsymbol{\theta}) C_{ik}(\boldsymbol{\theta}) J_{ik\alpha}(\boldsymbol{\theta}),$$

com

$$\begin{aligned} J_{ik\alpha}(\boldsymbol{\theta}) &= \delta_i \left(\kappa_{ik1}(v_i)^{-1} \frac{\partial \kappa_{ik1}(v_i)}{\partial \alpha} - \kappa_{ik2}(v_i) \frac{\partial \kappa_{ik2}(v_i)}{\partial \alpha} \right) \\ &\quad - (1 - \delta_i) h(\kappa_{ik2}(v_i)) \frac{\partial \kappa_{ik2}(v_i)}{\partial \alpha} + \frac{\partial \kappa_{ik3}}{\partial \alpha}, \end{aligned}$$

considerando

$$O_{ik} = \frac{v_i - (\mathbf{x}_{2i}(v_i)^T \boldsymbol{\beta}_2 + \gamma(\mathbf{x}_{1i}(v_i)^T \boldsymbol{\beta}_1 + \sqrt{2}\sigma_0 s_k))}{2} \quad \text{e,} \quad O_i = \frac{v_i - \mathbf{x}_{2i}(v_i)^T \boldsymbol{\beta}_2}{2},$$

então,

$$\kappa_{ik1}(v_i) = \frac{2}{\alpha} \cosh(O_{ik}), \quad \kappa_{ik2}(v_i) = \frac{2}{\alpha} \sinh(O_{ik}), \quad \xi_{i1}(v_i) = \frac{2}{\alpha} \cosh(O_i), \quad \xi_{i2}(v_i) = \frac{2}{\alpha} \sinh(O_i),$$

$$\frac{\partial \kappa_{ik1}(v_i)}{\partial \alpha} = -\frac{2}{\alpha^2} \cosh(O_{ik}) = -\frac{1}{\alpha} \kappa_{ik1}(v_i),$$

$$\frac{\partial \kappa_{ik2}(v_i)}{\partial \alpha} = -\frac{2}{\alpha^2} \sinh(O_{ik}) = -\frac{1}{\alpha} \kappa_{ik2}(v_i) \quad \text{e,} \quad \frac{\partial \kappa_{ik3}(v_i)}{\partial \alpha} = 0,$$

portanto,

$$\begin{aligned}
J_{ik\alpha}(\boldsymbol{\theta}) &= \delta_i \left(\kappa_{ik1}(v_i)^{-1} \left(-\frac{1}{\alpha} \kappa_{ik1}(v_i) \right) - \kappa_{ik2}(v_i) \left(-\frac{1}{\alpha} \kappa_{ik2}(v_i) \right) \right) \\
&\quad - (1 - \delta_i) h(\kappa_{ik2}(v_i)) \left(-\frac{1}{\alpha} \kappa_{ik2}(v_i) \right) \\
&= \delta_i \left(-\frac{1}{\alpha} + \frac{1}{\alpha} \kappa_{ik2}^2(v_i) \right) + (1 - \delta_i) h(\kappa_{ik2}(v_i)) \left(\frac{1}{\alpha} \kappa_{ik2}(v_i) \right) \\
&= \frac{1}{\alpha} \left[\delta_i (\kappa_{ik2}^2(v_i) - 1) + (1 - \delta_i) (h(\kappa_{ik2}(v_i)) \kappa_{ik2}(v_i)) \right],
\end{aligned}$$

$$\begin{aligned}
U_{i2\alpha}(\boldsymbol{\theta}) &= M_{i\alpha}(\boldsymbol{\theta}) = \delta_i \left(\xi_{i1}(v_i)^{-1} \frac{\partial \xi_{i1}(v_i)}{\partial \alpha} - \xi_{i2}(v_i) \frac{\partial \xi_{i2}(v_i)}{\partial \alpha} \right) \\
&\quad - (1 - \delta_i) h(\xi_{i2}(v_i)) \frac{\partial \xi_{i2}(v_i)}{\partial \alpha},
\end{aligned}$$

com

$$\frac{\partial \xi_{i1}(v_i)}{\partial \alpha} = -\frac{1}{\alpha} \xi_{i1}(v_i) \quad \text{e} \quad \frac{\partial \xi_{i2}(v_i)}{\partial \alpha} = -\frac{1}{\alpha} \xi_{i2}(v_i),$$

de modo que

$$\begin{aligned}
U_{i2\alpha}(\boldsymbol{\theta}) &= \delta_i \left(\xi_{i1}(v_i)^{-1} \left(-\frac{1}{\alpha} \xi_{i1}(v_i) \right) - \xi_{i2}(v_i) \left(-\frac{1}{\alpha} \xi_{i2}(v_i) \right) \right) \\
&\quad - (1 - \delta_i) h(\xi_{i2}(v_i)) \left(-\frac{1}{\alpha} \xi_{i2}(v_i) \right) \\
&= \delta_i \left(-\frac{1}{\alpha} + \frac{1}{\alpha} \xi_{i2}^2(v_i) \right) + (1 - \delta_i) h(\xi_{i2}(v_i)) \left(\frac{1}{\alpha} \xi_{i2}(v_i) \right) \\
&= \frac{1}{\alpha} \left[\delta_i (\xi_{i2}^2(v_i) - 1) + (1 - \delta_i) (h(\xi_{i2}(v_i)) \xi_{i2}(v_i)) \right],
\end{aligned}$$

assim,

$$U_{\alpha}(\boldsymbol{\theta}) = \sum_{i=1}^n \omega_i U_{i1\alpha}(\boldsymbol{\theta}) + \sum_{i=1}^n (1 - \omega_i) U_{i2\alpha}(\boldsymbol{\theta}).$$

Derivando $l(\boldsymbol{\theta})$ em relação a β_{2s} , $s = 1, \dots, p_2$, tem-se que

$$U_{i1\beta_{2s}}(\boldsymbol{\theta}) = D_i^{-1}(\boldsymbol{\theta}) \sum_{k=1}^Q \frac{w_k}{\sqrt{\pi}} A_{ik}(\boldsymbol{\theta}) B_{ik}(\boldsymbol{\theta}) C_{ik}(\boldsymbol{\theta}) J_{ik\beta_{2s}}(\boldsymbol{\theta}),$$

com

$$J_{ik\beta_{2s}}(\boldsymbol{\theta}) = \delta_i \left(\kappa_{ik1}(v_i)^{-1} \frac{\partial \kappa_{ik1}(v_i)}{\partial \beta_{2s}} - \kappa_{ik2}(v_i) \frac{\partial \kappa_{ik2}(v_i)}{\partial \beta_{2s}} \right) \\ - (1 - \delta_i) h(\kappa_{ik2}(v_i)) \frac{\partial \kappa_{ik2}(v_i)}{\partial \beta_{2s}} + \frac{\partial \kappa_{ik3}}{\partial \beta_{2s}},$$

em que

$$\frac{\partial \kappa_{ik1}(v_i)}{\partial \beta_{2s}} = \frac{2}{\alpha} \sinh(O_{ik}) \left(-\frac{x_{2is}(v_i)}{2} \right) = -\frac{x_{2is}(v_i)}{2} \kappa_{ik2}(v_i), \\ \frac{\partial \kappa_{ik2}(v_i)}{\partial \beta_{2s}} = \frac{2}{\alpha} \cosh(O_{ik}) \left(-\frac{x_{2is}(v_i)}{2} \right) = -\frac{x_{2is}(v_i)}{2} \kappa_{ik1}(v_i) \quad \text{e,} \quad \frac{\partial \kappa_{ik3}}{\partial \beta_{2s}} = 0,$$

logo

$$J_{ik\beta_{2s}}(\boldsymbol{\theta}) = \delta_i \left(\kappa_{ik1}(v_i)^{-1} \left(-\frac{x_{2is}(v_i)}{2} \kappa_{ik2}(v_i) \right) - \kappa_{ik2}(v_i) \left(-\frac{x_{2is}(v_i)}{2} \kappa_{ik1}(v_i) \right) \right) \\ - (1 - \delta_i) h(\kappa_{ik2}(v_i)) \left(-\frac{x_{2is}(v_i)}{2} \kappa_{ik1}(v_i) \right) \\ = \delta_i \left(\frac{x_{2is}(v_i)}{2} (\kappa_{ik1}(v_i) \kappa_{ik2}(v_i) - \kappa_{ik1}(v_i)^{-1} \kappa_{ik2}(v_i)) \right) \\ + (1 - \delta_i) \left(\frac{x_{2is}(v_i)}{2} h(\kappa_{ik2}(v_i)) \kappa_{ik1}(v_i) \right) \\ = \frac{1}{2} \left[\delta_i x_{2is}(v_i) (\kappa_{ik1}(v_i) \kappa_{ik2}(v_i) - \kappa_{ik4}(v_i)) + (1 - \delta_i) x_{2is}(v_i) h(\kappa_{ik2}(v_i)) \kappa_{ik1}(v_i) \right],$$

em que

$$\kappa_{ik4}(v_i) = \frac{\kappa_{ik2}(v_i)}{\kappa_{ik1}(v_i)} = \tanh(O_{ik}) \quad \text{e,}$$

$$\kappa_{ik1}(v_i) \kappa_{ik2}(v_i) = \left(\frac{2}{\alpha} \cosh(O_{ik}) \right) \left(\frac{2}{\alpha} \sinh(O_{ik}) \right) \\ = \frac{2}{\alpha^2} \left(2 \cosh(O_{ik}) \sinh(O_{ik}) \right) = \frac{2}{\alpha^2} \left(\sinh(2O_{ik}) \right),$$

$$U_{i2\beta_{2s}}(\boldsymbol{\theta}) = M_{i\beta_{2s}}(\boldsymbol{\theta}) = \delta_i \left(\xi_{i1}(v_i)^{-1} \frac{\partial \xi_{i1}(v_i)}{\partial \beta_{2s}} - \xi_{i2}(v_i) \frac{\partial \xi_{i2}(v_i)}{\partial \beta_{2s}} \right) \\ - (1 - \delta_i) \left(h(\xi_{i2}(v_i)) \frac{\partial \xi_{i2}(v_i)}{\partial \beta_{2s}} \right),$$

com

$$\frac{\partial \xi_{i1}(v_i)}{\partial \beta_{2s}} = -\frac{x_{2is}(v_i)}{2} \xi_{i2}(v_i) \quad \text{e}, \quad \frac{\partial \xi_{i2}(v_i)}{\partial \beta_{2s}} = -\frac{x_{2is}(v_i)}{2} \xi_{i1}(v_i),$$

de modo que

$$U_{i2\beta_{2s}}(\boldsymbol{\theta}) = \delta_i \left(\xi_{i1}(v_i)^{-1} \left(-\frac{x_{2is}(v_i)}{2} \xi_{i2}(v_i) \right) - \xi_{i2}(v_i) \left(-\frac{x_{2is}(v_i)}{2} \xi_{i1}(v_i) \right) \right) \\ - (1 - \delta_i) \left(h(\xi_{i2}(v_i)) \left(-\frac{x_{2is}(v_i)}{2} \xi_{i1}(v_i) \right) \right) \\ = \delta_i \left(\frac{x_{2is}(v_i)}{2} (\xi_{i1}(v_i) \xi_{i2}(v_i) - \xi_{i1}(v_i)^{-1} \xi_{i2}(v_i)) \right) + (1 - \delta_i) \left(\frac{x_{2is}(v_i)}{2} h(\xi_{i2}(v_i)) \xi_{i1}(v_i) \right) \\ = \frac{1}{2} \left[\delta_i x_{2is}(z_i) (\xi_{i1}(v_i) \xi_{i2}(v_i) - \xi_{i3}(v_i)) + (1 - \delta_i) x_{2is}(v_i) h(\xi_{i2}(v_i)) \xi_{i1}(v_i) \right],$$

em que

$$\xi_{i3}(v_i) = \frac{\xi_{i2}(v_i)}{\xi_{i1}(v_i)} = \tanh(O_i) \quad \text{e},$$

$$\xi_{i1}(z_i) \xi_{i2}(z_i) = \left(\frac{2}{\alpha} \cosh(O_i) \right) \left(\frac{2}{\alpha} \sinh(O_i) \right) \\ = \frac{2}{\alpha^2} (2 \cosh(O_i) \sinh(O_i)) = \frac{2}{\alpha^2} (\sinh(2O_i)),$$

assim,

$$U_{\beta_{2s}}(\boldsymbol{\theta}) = \sum_{i=1}^n \omega_i U_{i1\beta_{2s}}(\boldsymbol{\theta}) + \sum_{i=1}^n (1 - \omega_i) U_{i2\beta_{2s}}(\boldsymbol{\theta}), \quad s = 1, \dots, p_2.$$

Derivando $l(\boldsymbol{\theta})$ em relação a γ , tem-se que

$$U_{i1\gamma}(\boldsymbol{\theta}) = D_i^{-1}(\boldsymbol{\theta}) \sum_{k=1}^Q \frac{w_k}{\sqrt{\pi}} A_{ik}(\boldsymbol{\theta}) B_{ik}(\boldsymbol{\theta}) C_{ik}(\boldsymbol{\theta}) J_{ik\gamma}(\boldsymbol{\theta}),$$

com

$$J_{ik\gamma}(\boldsymbol{\theta}) = \delta_i \left(\kappa_{ik1}(v_i)^{-1} \frac{\partial \kappa_{ik1}(v_i)}{\partial \gamma} - \kappa_{ik2}(v_i) \frac{\partial \kappa_{ik2}(v_i)}{\partial \gamma} \right) \\ - (1 - \delta_i) h(\kappa_{ik2}(v_i)) \frac{\partial \kappa_{ik2}(v_i)}{\partial \gamma} + \frac{\partial \kappa_{ik3}}{\partial \gamma},$$

em que

$$\frac{\partial \kappa_{ik1}(v_i)}{\partial \gamma} = \frac{2}{\alpha} \sinh(O_{ik}) \left(- \left(\frac{\mathbf{x}_{1i}(v_i)^T \boldsymbol{\beta}_1 + \sqrt{2} \sigma_0 s_k}{2} \right) \right) \\ = - \left(\frac{\mathbf{x}_{1i}(v_i)^T \boldsymbol{\beta}_1 + \sqrt{2} \sigma_0 s_k}{2} \right) \kappa_{ik2}(v_i),$$

$$\frac{\partial \kappa_{ik2}(v_i)}{\partial \gamma} = \frac{2}{\alpha} \cosh(O_{ik}) \left(- \left(\frac{\mathbf{x}_{1i}(v_i)^T \boldsymbol{\beta}_1 + \sqrt{2} \sigma_0 s_k}{2} \right) \right) \\ = - \left(\frac{\mathbf{x}_{1i}(v_i)^T \boldsymbol{\beta}_1 + \sqrt{2} \sigma_0 s_k}{2} \right) \kappa_{ik1}(v_i)$$

e,

$$\frac{\partial \kappa_{ik3}}{\partial \gamma} = 0,$$

logo

$$J_{ik\gamma}(\boldsymbol{\theta}) = \delta_i \left[\kappa_{ik1}(v_i)^{-1} \left(- \left(\frac{\mathbf{x}_{1i}(v_i)^T \boldsymbol{\beta}_1 + \sqrt{2} \sigma_0 s_k}{2} \right) \right) \kappa_{ik2}(v_i) \right. \\ \left. - \kappa_{ik2}(v_i) \left(- \left(\frac{\mathbf{x}_{1i}(z_i)^T \boldsymbol{\beta}_1 + \sqrt{2} \sigma_0 s_k}{2} \right) \right) \kappa_{ik1}(v_i) \right] \\ - (1 - \delta_i) h(\kappa_{ik2}(v_i)) \left(- \left(\frac{\mathbf{x}_{1i}(v_i)^T \boldsymbol{\beta}_1 + \sqrt{2} \sigma_0 s_k}{2} \right) \right) \kappa_{ik1}(v_i) \\ = \frac{1}{2} \left[\delta_i (\mathbf{x}_{1i}(v_i)^T \boldsymbol{\beta}_1 + \sqrt{2} \sigma_0 s_k) (\kappa_{ik1}(v_i) \kappa_{ik2}(v_i) - \kappa_{ik4}(v_i)) \right. \\ \left. + (1 - \delta_i) (\mathbf{x}_{1i}(v_i)^T \boldsymbol{\beta}_1 + \sqrt{2} \sigma_0 s_k) h(\kappa_{ik2}(v_i)) \kappa_{ik1}(v_i) \right],$$

além disso

$$U_{i2\gamma}(\boldsymbol{\theta}) = M_{i2\gamma}(\boldsymbol{\theta}) = \delta_i \left(\xi_{i1}(v_i)^{-1} \frac{\partial \xi_{i1}(v_i)}{\partial \gamma} - \xi_{i2}(v_i) \frac{\partial \xi_{i2}(v_i)}{\partial \gamma} \right) - (1 - \delta_i) h(\xi_{i2}(v_i)) \left(\frac{\partial \xi_{i2}(v_i)}{\partial \gamma} \right),$$

com

$$\frac{\partial \xi_{i1}(v_i)}{\partial \gamma} = 0 \quad \text{e,} \quad \frac{\partial \xi_{i2}(v_i)}{\partial \gamma} = 0,$$

portanto

$$U_{i2\gamma}(\boldsymbol{\theta}) = M_{i2\gamma}(\boldsymbol{\theta}) = 0$$

assim,

$$U_{\gamma}(\boldsymbol{\theta}) = \sum_{i=1}^n \omega_i U_{i1\gamma}(\boldsymbol{\theta}).$$

Apêndice E

Distribuição de Weibull

Seja T uma variável aleatória que representa o tempo de falha. Supondo que T segue uma distribuição de Weibull (Weibull, 1939), simbolicamente $T \sim Wei(\varpi, \zeta)$, com $\varpi > 0$ e $\zeta > 0$ representando os parâmetros de forma e escala, respectivamente, as funções densidade de probabilidade e de sobrevivência de T são

$$f_T(t) = \frac{\varpi}{\zeta^\varpi} t^{\varpi-1} \exp\left\{-\left(\frac{t}{\zeta}\right)^\varpi\right\}, \quad t > 0, \varpi > 0, \zeta > 0,$$

e,

$$S_T(t) = P(T > t) = 1 - F_T(t) = \exp\left\{-\left(\frac{t}{\zeta}\right)^\varpi\right\}, \quad t > 0.$$

E.1 Modelo de regressão log-linear Weibull

Se $T \sim Wei(\varpi, \zeta)$, então, $V = \log(T)$ tem uma distribuição do valor extremo (VE), o que é representado por $V \sim VE(\varphi, \varsigma)$, com $\varphi = \log(\zeta) \in \Re$ e $\varsigma = \varpi^{-1} > 0$ denotando os parâmetros de localização e escala, respectivamente.

As funções densidade de probabilidade e de sobrevivência de V são

$$f_V(v) = \frac{1}{\varsigma} \exp\left\{\left(\frac{v - \varphi}{\varsigma}\right) - \exp\left\{\frac{v - \varphi}{\varsigma}\right\}\right\}, \quad v \in \Re, \varphi \in \Re, \varsigma > 0, \quad (\text{E.1})$$

e,

$$S_V(v) = P(V > v) = 1 - F_V(v) = \exp\left\{-\exp\left\{\frac{v - \varphi}{\varsigma}\right\}\right\}, \quad v \in \Re. \quad (\text{E.2})$$

Desta forma, o logaritmo do tempo de sobrevivência ou de censura observado para o i -ésimo indivíduo, $v_i = \log(Z_i)$, pode ser descrito por meio do modelo de regressão log-linear

Weibull

$$v_i = \mathbf{x}_{2i}^T \boldsymbol{\beta}_2 + \varsigma \vartheta_i, \quad (\text{E.3})$$

em que $\boldsymbol{\beta}_2$ e \mathbf{x}_{2i} são definidos de forma semelhante à considerada em (3.1), $\varsigma > 0$ representa o parâmetro de escala da distribuição VE e os termos ϑ_i são definidos como variáveis aleatórias independentes e identicamente distribuídas, tais que $\vartheta_i \sim \text{VE}(0, 1)$ para $i = 1, \dots, n$. Detalhes sobre a obtenção dos estimadores de máxima verossimilhança dos parâmetros em (E.3) podem ser consultados em Kalbfleisch e Prentice (2002) ou em Colosimo e Giolo (2006).

Apêndice F

Outros resultados do estudo de simulação

F.1 Desempenho do processo de estimação dos parâmetros

F.1.1 Dados B-S - Modelo B-S-Normal

Tabela F.1: (*Dados BS - Modelo BS-Normal*): Medidas das estimativas dos parâmetros, $n = 500$, $PcI = 7\%$, $PcA = 0\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\alpha)$	-0,4600	-0,3162	0,1438	0,3125	0,1683	0,3658
β_{20}	6,1400	4,6773	-1,4627	-0,2382	1,5368	0,2503
β_{21}	0,1000	0,7422	0,6422	6,4216	0,7102	7,1022
γ	-0,4600	-0,3241	0,1359	0,2954	0,1440	0,3131
β_{10}	5,8000	5,7485	-0,0515	-0,0089	0,2122	0,0366
β_{11}	0,0100	0,0094	-0,0006	-0,0636	0,0011	0,1096
β_{12}	-1,1000	-1,0710	0,0290	0,0264	0,2226	0,2023
$\log(\sigma_0)$	0,4500	0,4616	0,0116	0,0259	0,0624	0,1386
$\log(\sigma_e)$	-0,3200	-0,3332	-0,0132	-0,0414	0,0572	0,1786

Tabela F.2: (*Dados BS - Modelo BS-Normal*): Medidas das estimativas dos parâmetros, $n = 500$, $PcI = 7\%$, $PcA = 18\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\alpha)$	-0,4600	-0,2871	0,1729	0,3759	0,2088	0,4540
β_{20}	6,1400	5,3689	-0,7711	-0,1256	0,9721	0,1583
β_{21}	0,1000	0,6540	0,5540	5,5399	0,6848	6,8476
γ	-0,4600	-0,4308	0,0292	0,0634	0,0644	0,1399
β_{10}	5,8000	5,7455	-0,0545	-0,0094	0,2309	0,0398
β_{11}	0,0100	0,0094	-0,0006	-0,0618	0,0011	0,1118
β_{12}	-1,1000	-1,0541	0,0459	0,0417	0,2434	0,2212
$\log(\sigma_0)$	0,4500	0,4515	0,0015	0,0033	0,0623	0,1384
$\log(\sigma_e)$	-0,3200	-0,3273	-0,0073	-0,0230	0,0503	0,1571

Tabela F.3: (*Dados BS - Modelo BS-Normal*): Medidas das estimativas dos parâmetros, $n = 500$, $PcI = 7\%$, $PcA = 43\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\alpha)$	-0,4600	-0,2330	0,2270	0,4935	0,2406	0,5231
β_{20}	6,1400	6,6122	0,4722	0,0769	0,8814	0,1436
β_{21}	0,1000	0,3445	0,2445	2,4446	0,5245	5,2449
γ	-0,4600	-0,5890	-0,1290	-0,2805	0,1456	0,3165
β_{10}	5,8000	5,7649	-0,0351	-0,0061	0,2286	0,0394
β_{11}	0,0100	0,0094	-0,0006	-0,0580	0,0011	0,1130
β_{12}	-1,1000	-1,0840	0,0160	0,0145	0,2395	0,2177
$\log(\sigma_0)$	0,4500	0,4432	-0,0068	-0,0151	0,0490	0,1089
$\log(\sigma_e)$	-0,3200	-0,3196	0,0004	0,0013	0,0346	0,1083

Tabela F.4: (*Dados BS - Modelo BS-Normal*): Medidas das estimativas dos parâmetros, $n = 500$, $PcI = 7\%$, $PcA = 68\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\alpha)$	-0,4600	6,3946	6,8546	14,9014	7,5833	16,4855
β_{20}	6,1400	20,1135	13,9735	2,2758	15,2006	2,4757
β_{21}	0,1000	0,1038	0,0038	0,0381	0,1958	1,9577
γ	-0,4600	-0,7064	-0,2464	-0,5357	0,2570	0,5586
β_{10}	5,8000	5,7701	-0,0299	-0,0052	0,2221	0,0383
β_{11}	0,0100	0,0079	-0,0021	-0,2081	0,0023	0,2296
β_{12}	-1,1000	-1,0952	0,0048	0,0044	0,2345	0,2132
$\log(\sigma_0)$	0,4500	0,4237	-0,0263	-0,0583	0,0452	0,1004
$\log(\sigma_e)$	-0,3200	-0,3188	0,0012	0,0037	0,0297	0,0928

Tabela F.5: (Dados BS - Modelo BS-Normal): Medidas das estimativas dos parâmetros, $n = 250$, $PcI = 7\%$, $PcA = 0\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\alpha)$	-0,4600	-0,3517	0,1083	0,2354	0,1522	0,3309
β_{20}	6,1400	4,9195	-1,2205	-0,1988	1,3651	0,2223
β_{21}	0,1000	0,6591	0,5591	5,5907	0,6732	6,7324
γ	-0,4600	-0,3569	0,1031	0,2242	0,1207	0,2625
β_{10}	5,8000	5,7352	-0,0648	-0,0112	0,3016	0,0520
β_{11}	0,0100	0,0095	-0,0005	-0,0513	0,0013	0,1303
β_{12}	-1,1000	-1,0528	0,0472	0,0429	0,3262	0,2966
$\log(\sigma_0)$	0,4500	0,4601	0,0101	0,0224	0,0734	0,1632
$\log(\sigma_e)$	-0,3200	-0,3312	-0,0112	-0,0350	0,0606	0,1893

Tabela F.6: (Dados BS - Modelo BS-Normal): Medidas das estimativas dos parâmetros, $n = 250$, $PcI = 7\%$, $PcA = 18\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\alpha)$	-0,4600	-0,3175	0,1425	0,3097	0,1826	0,3969
β_{20}	6,1400	5,5364	-0,6036	-0,0983	0,9219	0,1501
β_{21}	0,1000	0,5745	0,4745	4,7453	0,6745	6,7454
γ	-0,4600	-0,4505	0,0095	0,0207	0,0651	0,1415
β_{10}	5,8000	5,7410	-0,0590	-0,0102	0,3209	0,0553
β_{11}	0,0100	0,0095	-0,0005	-0,0451	0,0013	0,1333
β_{12}	-1,1000	-1,0500	0,0500	0,0455	0,3429	0,3117
$\log(\sigma_0)$	0,4500	0,4479	-0,0021	-0,0046	0,0803	0,1784
$\log(\sigma_e)$	-0,3200	-0,3246	-0,0046	-0,0145	0,0577	0,1804

Tabela F.7: (Dados BS - Modelo BS-Normal): Medidas das estimativas dos parâmetros, $n = 250$, $PcI = 7\%$, $PcA = 43\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\alpha)$	-0,4600	-0,2296	0,2304	0,5009	0,2469	0,5367
β_{20}	6,1400	6,7145	0,5745	0,0936	0,9489	0,1545
β_{21}	0,1000	0,2821	0,1821	1,8206	0,5141	5,1410
γ	-0,4600	-0,5988	-0,1388	-0,3016	0,1551	0,3372
β_{10}	5,8000	5,7731	-0,0269	-0,0046	0,2904	0,0501
β_{11}	0,0100	0,0093	-0,0007	-0,0691	0,0014	0,1421
β_{12}	-1,1000	-1,0967	0,0033	0,0030	0,3028	0,2753
$\log(\sigma_0)$	0,4500	0,4329	-0,0171	-0,0379	0,0604	0,1343
$\log(\sigma_e)$	-0,3200	-0,3223	-0,0023	-0,0073	0,0424	0,1326

Tabela F.8: (*Dados BS - Modelo BS-Normal*): Medidas das estimativas dos parâmetros, $n = 250$, $PcI = 7\%$, $PcA = 68\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\alpha)$	-0,4600	5,6511	6,1111	13,2851	7,1745	15,5967
β_{20}	6,1400	18,8168	12,6768	2,0646	14,4746	2,3574
β_{21}	0,1000	0,1017	0,0017	0,0174	0,3187	3,1874
γ	-0,4600	-0,7256	-0,2656	-0,5775	0,2799	0,6084
β_{10}	5,8000	5,8024	0,0024	0,0004	0,3218	0,0555
β_{11}	0,0100	0,0079	-0,0021	-0,2108	0,0026	0,2575
β_{12}	-1,1000	-1,1138	-0,0138	-0,0126	0,3421	0,3110
$\log(\sigma_0)$	0,4500	0,4240	-0,0260	-0,0577	0,0601	0,1335
$\log(\sigma_e)$	-0,3200	-0,3215	-0,0015	-0,0047	0,0409	0,1277

Tabela F.9: (*Dados BS - Modelo BS-Normal*): Medidas das estimativas dos parâmetros, $n = 100$, $PcI = 7\%$, $PcA = 0\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\alpha)$	-0,4600	-0,3811	0,0789	0,1715	0,1383	0,3006
β_{20}	6,1400	5,1288	-1,0112	-0,1647	1,2087	0,1969
β_{21}	0,1000	0,6154	0,5154	5,1537	0,6782	6,7822
γ	-0,4600	-0,3893	0,0707	0,1537	0,0957	0,2080
β_{10}	5,8000	5,7392	-0,0608	-0,0105	0,5189	0,0895
β_{11}	0,0100	0,0095	-0,0005	-0,0479	0,0018	0,1805
β_{12}	-1,1000	-1,0708	0,0292	0,0265	0,5485	0,4987
$\log(\sigma_0)$	0,4500	0,4449	-0,0051	-0,0113	0,0998	0,2217
$\log(\sigma_e)$	-0,3200	-0,3368	-0,0168	-0,0526	0,0747	0,2334

Tabela F.10: (*Dados BS - Modelo BS-Normal*): Medidas das estimativas dos parâmetros, $n = 100$, $PcI = 7\%$, $PcA = 18\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\alpha)$	-0,4600	-0,3580	0,1020	0,2218	0,1561	0,3393
β_{20}	6,1400	5,6858	-0,4542	-0,0740	0,8492	0,1383
β_{21}	0,1000	0,4768	0,3768	3,7676	0,6369	6,3689
γ	-0,4600	-0,4662	-0,0062	-0,0135	0,0658	0,1431
β_{10}	5,8000	5,8077	0,0077	0,0013	0,4898	0,0845
β_{11}	0,0100	0,0096	-0,0004	-0,0368	0,0021	0,2131
β_{12}	-1,1000	-1,1191	-0,0191	-0,0174	0,5155	0,4687
$\log(\sigma_0)$	0,4500	0,4346	-0,0154	-0,0343	0,1030	0,2289
$\log(\sigma_e)$	-0,3200	-0,3284	-0,0084	-0,0264	0,0662	0,2067

Tabela F.11: (*Dados BS - Modelo BS-Normal*): Medidas das estimativas dos parâmetros, $n = 100$, $PcI = 7\%$, $PcA = 43\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\alpha)$	-0,4600	-0,2502	0,2098	0,4561	0,2478	0,5386
β_{20}	6,1400	6,7099	0,5699	0,0928	1,0183	0,1658
β_{21}	0,1000	0,2692	0,1692	1,6923	0,6083	6,0829
γ	-0,4600	-0,5951	-0,1351	-0,2936	0,1553	0,3376
β_{10}	5,8000	5,7728	-0,0272	-0,0047	0,4871	0,0840
β_{11}	0,0100	0,0094	-0,0006	-0,0620	0,0021	0,2135
β_{12}	-1,1000	-1,0822	0,0178	0,0161	0,5295	0,4813
$\log(\sigma_0)$	0,4500	0,4252	-0,0248	-0,0552	0,0883	0,1962
$\log(\sigma_e)$	-0,3200	-0,3264	-0,0064	-0,0201	0,0667	0,2085

Tabela F.12: (*Dados BS - Modelo BS-Normal*): Medidas das estimativas dos parâmetros, $n = 100$, $PcI = 7\%$, $PcA = 68\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\alpha)$	-0,4600	4,4423	4,9023	10,6573	6,3816	13,8730
β_{20}	6,1400	16,8852	10,7452	1,7500	13,4881	2,1968
β_{21}	0,1000	-0,0691	-0,1691	-1,6912	2,1336	21,3358
γ	-0,4600	-0,7558	-0,2958	-0,6430	0,3157	0,6863
β_{10}	5,8000	5,8576	0,0576	0,0099	0,5306	0,0915
β_{11}	0,0100	0,0080	-0,0020	-0,1994	0,0029	0,2931
β_{12}	-1,1000	-1,1717	-0,0717	-0,0652	0,5556	0,5051
$\log(\sigma_0)$	0,4500	0,4102	-0,0398	-0,0883	0,0974	0,2165
$\log(\sigma_e)$	-0,3200	-0,3249	-0,0049	-0,0152	0,0665	0,2078

F.1.2 Dados Wei - Modelo Wei-Normal

Tabela F.13: (*Dados Wei - Modelo Wei-Normal*): Medidas das estimativas dos parâmetros, $n = 500$, $PcI = 6\%$, $PcA = 0\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\varsigma)$	-0,9700	-1,0394	-0,0694	-0,0716	0,0855	0,0881
β_{20}	6,2800	6,1901	-0,0899	-0,0143	0,1352	0,0215
β_{21}	0,1200	0,1045	-0,0155	-0,1294	0,0628	0,5232
γ	-0,4300	-0,4497	-0,0197	-0,0458	0,0249	0,0580
β_{10}	5,8300	5,8377	0,0077	0,0013	0,2188	0,0375
β_{11}	0,0100	0,0098	-0,0002	-0,0160	0,0008	0,0785
β_{12}	-1,1100	-1,1245	-0,0145	-0,0130	0,2338	0,2106
$\log(\sigma_0)$	0,4500	0,4450	-0,0050	-0,0110	0,0353	0,0784
$\log(\sigma_e)$	-0,3300	-0,3315	-0,0015	-0,0045	0,0235	0,0713

Tabela F.14: (*Dados Wei - Modelo Wei-Normal*): Medidas das estimativas dos parâmetros, $n = 500$, $PcI = 6\%$, $PcA = 19\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\varsigma)$	-0,9700	-0,9955	-0,0255	-0,0263	0,0579	0,0597
β_{20}	6,2800	6,4844	0,2044	0,0325	0,2347	0,0374
β_{21}	0,1200	0,0954	-0,0246	-0,2050	0,0739	0,6160
γ	-0,4300	-0,4981	-0,0681	-0,1584	0,0701	0,1631
β_{10}	5,8300	5,8302	0,0002	0,0000	0,2132	0,0366
β_{11}	0,0100	0,0096	-0,0004	-0,0387	0,0009	0,0887
β_{12}	-1,1100	-1,1146	-0,0046	-0,0041	0,2335	0,2104
$\log(\sigma_0)$	0,4500	0,4339	-0,0161	-0,0358	0,0392	0,0870
$\log(\sigma_e)$	-0,3300	-0,3266	0,0034	0,0102	0,0253	0,0766

Tabela F.15: (*Dados Wei - Modelo Wei-Normal*): Medidas das estimativas dos parâmetros, $n = 500$, $PcI = 6\%$, $PcA = 44\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\zeta)$	-0,9700	-0,9002	0,0698	0,0720	0,0955	0,0985
β_{20}	6,2800	7,1100	0,8300	0,1322	0,8448	0,1345
β_{21}	0,1200	0,0947	-0,0253	-0,2111	0,0955	0,7960
γ	-0,4300	-0,5919	-0,1619	-0,3764	0,1635	0,3803
β_{10}	5,8300	5,8151	-0,0149	-0,0026	0,2146	0,0368
β_{11}	0,0100	0,0093	-0,0007	-0,0694	0,0011	0,1087
β_{12}	-1,1100	-1,1043	0,0057	0,0052	0,2417	0,2177
$\log(\sigma_0)$	0,4500	0,4310	-0,0190	-0,0422	0,0420	0,0933
$\log(\sigma_e)$	-0,3300	-0,3248	0,0052	0,0158	0,0276	0,0836

Tabela F.16: (*Dados Wei - Modelo Wei-Normal*): Medidas das estimativas dos parâmetros, $n = 500$, $PcI = 6\%$, $PcA = 69\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\zeta)$	-0,9700	-0,7424	0,2276	0,2346	0,2443	0,2519
β_{20}	6,2800	8,3661	2,0861	0,3322	2,1029	0,3349
β_{21}	0,1200	0,0967	-0,0233	-0,1944	0,1384	1,1529
γ	-0,4300	-0,7528	-0,3228	-0,7507	0,3249	0,7557
β_{10}	5,8300	5,8301	0,0001	0,0000	0,2220	0,0381
β_{11}	0,0100	0,0083	-0,0017	-0,1697	0,0019	0,1940
β_{12}	-1,1100	-1,1226	-0,0126	-0,0114	0,2376	0,2141
$\log(\sigma_0)$	0,4500	0,4307	-0,0193	-0,0430	0,0408	0,0908
$\log(\sigma_e)$	-0,3300	-0,3206	0,0094	0,0285	0,0300	0,0910

Tabela F.17: (*Dados Wei - Modelo Wei-Normal*): Medidas das estimativas dos parâmetros, $n = 250$, $PcI = 6\%$, $PcA = 0\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\zeta)$	-0,9700	-1,0408	-0,0708	-0,0730	0,0994	0,1025
β_{20}	6,2800	6,1880	-0,0920	-0,0146	0,1801	0,0287
β_{21}	0,1200	0,1121	-0,0079	-0,0662	0,0961	0,8008
γ	-0,4300	-0,4508	-0,0208	-0,0484	0,0301	0,0700
β_{10}	5,8300	5,8419	0,0119	0,0020	0,3123	0,0536
β_{11}	0,0100	0,0098	-0,0002	-0,0201	0,0011	0,1095
β_{12}	-1,1100	-1,1299	-0,0199	-0,0180	0,3459	0,3116
$\log(\sigma_0)$	0,4500	0,4394	-0,0106	-0,0236	0,0540	0,1200
$\log(\sigma_e)$	-0,3300	-0,3311	-0,0011	-0,0034	0,0339	0,1027

Tabela F.18: (*Dados Wei - Modelo Wei-Normal*): Medidas das estimativas dos parâmetros, $n = 250$, $PcI = 6\%$, $PcA = 19\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\varsigma)$	-0,9700	-1,0041	-0,0341	-0,0351	0,0826	0,0852
β_{20}	6,2800	6,4784	0,1984	0,0316	0,2582	0,0411
β_{21}	0,1200	0,1019	-0,0181	-0,1511	0,1021	0,8505
γ	-0,4300	-0,4984	-0,0684	-0,1591	0,0727	0,1690
β_{10}	5,8300	5,8154	-0,0146	-0,0025	0,3045	0,0522
β_{11}	0,0100	0,0096	-0,0004	-0,0407	0,0013	0,1251
β_{12}	-1,1100	-1,1009	0,0091	0,0082	0,3293	0,2967
$\log(\sigma_0)$	0,4500	0,4320	-0,0180	-0,0400	0,0524	0,1165
$\log(\sigma_e)$	-0,3300	-0,3306	-0,0006	-0,0017	0,0345	0,1046

Tabela F.19: (*Dados Wei - Modelo Wei-Normal*): Medidas das estimativas dos parâmetros, $n = 250$, $PcI = 6\%$, $PcA = 44\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\varsigma)$	-0,9700	-0,9217	0,0483	0,0498	0,1070	0,1103
β_{20}	6,2800	7,0971	0,8171	0,1301	0,8491	0,1352
β_{21}	0,1200	0,1016	-0,0184	-0,1534	0,1284	1,0698
γ	-0,4300	-0,5913	-0,1613	-0,3752	0,1650	0,3836
β_{10}	5,8300	5,7998	-0,0302	-0,0052	0,3110	0,0533
β_{11}	0,0100	0,0092	-0,0008	-0,0807	0,0014	0,1433
β_{12}	-1,1100	-1,0968	0,0132	0,0119	0,3302	0,2975
$\log(\sigma_0)$	0,4500	0,4338	-0,0162	-0,0361	0,0540	0,1199
$\log(\sigma_e)$	-0,3300	-0,3227	0,0073	0,0221	0,0404	0,1223

Tabela F.20: (*Dados Wei - Modelo Wei-Normal*): Medidas das estimativas dos parâmetros, $n = 250$, $PcI = 6\%$, $PcA = 69\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\varsigma)$	-0,9700	-0,7756	0,1944	0,2004	0,2367	0,2440
β_{20}	6,2800	8,3243	2,0443	0,3255	2,0817	0,3315
β_{21}	0,1200	0,1089	-0,0111	-0,0928	0,2091	1,7423
γ	-0,4300	-0,7497	-0,3197	-0,7434	0,3241	0,7538
β_{10}	5,8300	5,8039	-0,0261	-0,0045	0,3073	0,0527
β_{11}	0,0100	0,0084	-0,0016	-0,1643	0,0021	0,2090
β_{12}	-1,1100	-1,0936	0,0164	0,0148	0,3270	0,2946
$\log(\sigma_0)$	0,4500	0,4279	-0,0221	-0,0491	0,0528	0,1174
$\log(\sigma_e)$	-0,3300	-0,3245	0,0055	0,0166	0,0418	0,1266

Tabela F.21: (*Dados Wei - Modelo Wei-Normal*): Medidas das estimativas dos parâmetros, $n = 100$, $PcI = 6\%$, $PcA = 0\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\varsigma)$	-0,9700	-1,0633	-0,0933	-0,0962	0,1441	0,1485
β_{20}	6,2800	6,2234	-0,0566	-0,0090	0,2522	0,0402
β_{21}	0,1200	0,1086	-0,0114	-0,0953	0,1512	1,2604
γ	-0,4300	-0,4576	-0,0276	-0,0643	0,0451	0,1050
β_{10}	5,8300	5,7974	-0,0326	-0,0056	0,4971	0,0853
β_{11}	0,0100	0,0097	-0,0003	-0,0318	0,0017	0,1726
β_{12}	-1,1100	-1,0861	0,0239	0,0215	0,5267	0,4745
$\log(\sigma_0)$	0,4500	0,4294	-0,0206	-0,0457	0,0850	0,1889
$\log(\sigma_e)$	-0,3300	-0,3351	-0,0051	-0,0154	0,0542	0,1643

Tabela F.22: (*Dados Wei - Modelo Wei-Normal*): Medidas das estimativas dos parâmetros, $n = 100$, $PcI = 6\%$, $PcA = 19\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\varsigma)$	-0,9700	-1,0327	-0,0627	-0,0647	0,1477	0,1522
β_{20}	6,2800	6,4781	0,1981	0,0315	0,3344	0,0532
β_{21}	0,1200	0,0962	-0,0238	-0,1982	0,1661	1,3839
γ	-0,4300	-0,4992	-0,0692	-0,1608	0,0795	0,1850
β_{10}	5,8300	5,7770	-0,0530	-0,0091	0,5128	0,0880
β_{11}	0,0100	0,0098	-0,0002	-0,0192	0,0019	0,1857
β_{12}	-1,1100	-1,0521	0,0579	0,0521	0,5552	0,5002
$\log(\sigma_0)$	0,4500	0,4231	-0,0269	-0,0599	0,0879	0,1954
$\log(\sigma_e)$	-0,3300	-0,3376	-0,0076	-0,0232	0,0579	0,1756

Tabela F.23: (*Dados Wei - Modelo Wei-Normal*): Medidas das estimativas dos parâmetros, $n = 100$, $PcI = 6\%$, $PcA = 44\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\varsigma)$	-0,9700	-0,9587	0,0113	0,0116	0,1646	0,1697
β_{20}	6,2800	7,0714	0,7914	0,1260	0,8696	0,1385
β_{21}	0,1200	0,0898	-0,0302	-0,2518	0,2101	1,7506
γ	-0,4300	-0,5873	-0,1573	-0,3657	0,1657	0,3854
β_{10}	5,8300	5,8145	-0,0155	-0,0027	0,5166	0,0886
β_{11}	0,0100	0,0092	-0,0008	-0,0770	0,0021	0,2095
β_{12}	-1,1100	-1,1045	0,0055	0,0050	0,5514	0,4968
$\log(\sigma_0)$	0,4500	0,4209	-0,0291	-0,0647	0,0907	0,2015
$\log(\sigma_e)$	-0,3300	-0,3303	-0,0003	-0,0010	0,0599	0,1816

Tabela F.24: (*Dados Wei - Modelo Wei-Normal*): Medidas das estimativas dos parâmetros, $n = 100$, $PcI = 6\%$, $PcA = 69\%$.

Parâmetro	Valor	$MED(\hat{\theta})$	$VI(\hat{\theta})$	$VIR(\hat{\theta})$	$REQM(\hat{\theta})$	$REQMR(\hat{\theta})$
$\log(\varsigma)$	-0,9700	-0,8122	0,1578	0,1627	0,2917	0,3007
β_{20}	6,2800	8,2826	2,0026	0,3189	2,0985	0,3342
β_{21}	0,1200	0,1123	-0,0077	-0,0645	0,3236	2,6966
γ	-0,4300	-0,7468	-0,3168	-0,7368	0,3293	0,7658
β_{10}	5,8300	5,8006	-0,0294	-0,0050	0,4800	0,0823
β_{11}	0,0100	0,0084	-0,0016	-0,1555	0,0026	0,2571
β_{12}	-1,1100	-1,0920	0,0180	0,0162	0,5064	0,4562
$\log(\sigma_0)$	0,4500	0,4169	-0,0331	-0,0735	0,0881	0,1958
$\log(\sigma_e)$	-0,3300	-0,3300	0,0000	0,0001	0,0655	0,1986

Referências bibliográficas

- Abu Bakar, M. R., Salah, K. A., Ibrahim, N. A. and Haron, K. (2009). Bayesian approach for joint longitudinal and time-to-event data with survival fraction. *Bulletin of the Malaysian Mathematical Sciences Society* 32, 75-100.
- Achcar, J. A. and Espinosa, M. M. (1992). Use of bayesian methods in accelerated life tests considering a log-linear model for the Birnbaum-Saunders distribution. *Biometrics Bulletin* 10, 18-22.
- Achcar, J. A. (1993). Inference for Birnbaum-Saunders fatigue life model using Bayesian methods. *Computational Statistics and Data Analysis* 15, 367-380.
- Albert, P. S. and Shih, J. H. (2010). On estimating the relationship between longitudinal measurements and time-to-event data using a simple two-stage procedure. *Biometrics* 66, 983-991.
- Andersen, P. K., Klein, J. P., Knudsen, K. M. and Tabanera y Palacios, R. (1997). Estimation of variance in Cox's regression model with shared gamma frailties. *Biometrics* 53, 1475- 1484.
- Augustin, T. (2004). An exact corrected log-likelihood function for Cox's proportional hazards model under measurement error and some extensions. *Scandinavian Journal of Statistics* 31, 43-50.
- Azevedo, A., Bettencourt, P. and Barros, H. (2007). Demographic, clinical and echocardiographic determinants of B-type natriuretic peptide plasma concentration. *Revista Portuguesa de Cardiologia* 26, 105-113.
- Balakrishnan, N., Leiva, V. and López, J. (2007). Acceptance sampling plans from truncated life tests based on the generalized Birnbaum-Saunders distribution. *Communications in Statistics: Simulation and Computation* 36, 643-656.

- Balakrishnan, N., Leiva, V., Sanhueza, A. and Vilca, F. (2009). Estimation in the Birnbaum-Saunders distribution based on scale-mixture of normals and the EM-algorithm. *Statistics and Operations Research Transactions* 33, 171-192.
- Balakrishnan, N., Gupta, R., Kundu, D., Leiva, V. and Sanhueza, A. (2011). On some mixture models based on the Birnbaum-Saunders distribution and associated inference. *Journal of Statistical Planning and Inference* 141, 2175-2190.
- Barros, M., Paula, G. A. and Leiva, V. (2008). A new class of survival regression models with heavy-tailed errors: robustness and diagnostics. *Lifetime Data Analysis* 14, 316-332.
- Barros, M., Paula, G. A. and Leiva, V. (2009). An R implementation for generalized Birnbaum-Saunders distributions. *Computational Statistics and Data Analysis* 53, 1511-1528.
- Bettencourt, P. (2004). NT-proBNP and BNP: biomarkers for heart failure management. *The European Journal of Heart Failure* 6, 359-363.
- Bettencourt, P., Friões, F., Azevedo, A., Dias, P., Pimenta, J., Rocha-Gonçalves, F., Ferreira, A. (2004). Prognostic information provided by serial measurements of brain natriuretic peptide in heart failure. *International Journal of Cardiology* 93, 45-48.
- Birnbaum, Z. W. and Saunders, S. C. (1968). A probabilistic interpretation of Miner's rule. *Journal of Applied Mechanics* 16, 637-652.
- Birnbaum, Z. W. and Saunders, S. C. (1969a). A new family of life distributions. *Journal of Applied Probability* 6, 319-327.
- Birnbaum, Z. W. and Saunders, S. C. (1969b). Estimation for a family of life distributions with applications to fatigue. *Journal of Applied Probability* 6, 328-347.
- Bolker, B. and R Development Core Team (2013). *bbmle: Tools for General Maximum Likelihood Estimation*. R package version 1.0-15.
<http://cran.r-project.org/web/packages/bbmle/index.html>.
- Brown, E. R. and Ibrahim, J. G. (2003a). A Bayesian semiparametric joint hierarchical model for longitudinal and survival data. *Biometrics* 59, 221-228.
- Brown, E. R. and Ibrahim, J. G. (2003b). Bayesian approaches to joint cure-rate and longitudinal models with applications to cancer vaccine trials. *Biometrics* 59, 686-693.

- Bycott, P. and Taylor, J. (1998). A comparison of smoothing techniques for CD4 data measured with error in a time-dependent Cox proportional hazards model. *Statistics in Medicine* 17, 2061-2077.
- Cai, Z. and Sun, Y. (2003). Local linear estimation for time dependent coefficients in Cox's regression models. *Scandinavian Journal of Statistics* 30, 93-111.
- Cancho, V., Ortega, E. and Paula, G. (2010). On estimation and influence diagnostics for log-Birnbaum-Saunders Student-t regression models: full bayesian analysis. *Journal of Statistical Planning and Inference* 140, 2486-2496.
- Cancho, V. G., Louzada, F. and Barriga, G. D. C. (2012). The Geometric Birnbaum-Saunders regression model with cure rate. *Journal of Statistical Planning and Inference* 142, 993-1000.
- Chen, M. H., Ibrahim, J. and Sinha, D. (2004). A new joint model for longitudinal and survival data with a cure fraction. *Journal of Multivariate Analysis* 91, 18-34.
- Chi, Y. Y. and Ibrahim, J. G. (2006). Joint models for multivariate longitudinal and multivariate survival data. *Biometrics* 62, 432-445.
- Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* 65, 141-151.
- Clayton, D. and Cuzick, J. (1985). Multivariate generalizations of the proportional hazards model. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 148, 82-117.
- Cohen, H. (2011). *Numerical Approximation Methods*. Springer, New York.
- Colosimo, E. A. e Giolo, S. (2006). *Análise de Sobrevida Aplicada*. 1a edição. Edgard Blücher Ltda, São Paulo.
- Cordeiro, G. M. and Lemonte, A. J. (2011). The β -Birnbaum-Saunders distribution: An improved distribution for fatigue life modeling. *Computational Statistics and Data Analysis* 55, 1445-1461.

- Cowiea, M. R., Jourdainb, P., Maiselc, A., Dahlstromd, U., Follathe, F., Isnardf, R., Luchnerg, A. , McDonaghg, T., Mairi, J., Nieminenj, M. and Francisk, G. (2003). Clinical applications of B-type natriuretic peptide (BNP) testing. *European Heart Journal* 24, 1710-1718.
- Cox, D. R. (1972). Regression models and life tables (with Discussion). *Journal of the Royal Statistical Society: Series B (Methodological)* 34, 187-200.
- Cox, D. R. and Snell, E. J. (1968). A general definition of residuals. *Journal of the Royal Statistical Society: Series B (Methodological)* 30, 248-275.
- Cysneiros, A. H. M. A., Cribari-Neto, F. and Araújo Jr., C. A. G. (2008). On Birnbaum-Saunders inference. *Computational Statistics and Data Analysis* 52, 4939-4950.
- Dafni, U. G. and Tsiatis, A. A. (1998). Evaluating surrogate markers of clinical outcome when measured with error. *Biometrics* 54, 1445-1462.
- Davis, P. J. and Rabinowits, P. (1975). *Methods of Numerical Integration*. Academic Press, New York.
- DeGruttola, V. and Tu, X. M. (1994). Modeling progression of CD-4 lymphocyte count and its relationship to survival time. *Biometrics* 50, 1003-1014.
- Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)* 39, 1-38.
- Desmond A. (1985). Stochastic models of failure in random environments. *The Canadian Journal of Statistics* 13, 171-183.
- Desmond A. (1986). On the relationship between two fatigue life models. *IEEE Transactions on Reliability* 35, 167-169.
- Desmond, A. F., Rodríguez-Yam, G. A. and Lu, X. (2008). Estimation of the parameters for a Birnbaum-Saunders regression model with censored data. *Journal of Statistical Computation and Simulation* 11, 983-997.
- Desmond, A. F., Cíntora-González, C. L. , Singh, R. S. and Lu, X. (2012). A mixed effects log-linear model based on the Birnbaum-Saunders distribution. *Computational Statistics and Data Analysis* 56, 399-407.

- Díaz-García, J. A. and Leiva, V. (2005). A new family of life distributions based on elliptically contoured distributions. *Journal of Statistical Planning and Inference* 128, 445-457.
- Diggle, P. J. (1988). An approach to the analysis of repeated measurements. *Biometrics* 44, 959-971.
- Diggle, P. J., Heagerty, P. J., Liang, K. Y. and Zeger, S. L. (2002). *Analysis of Longitudinal Data*. 2nd edition. Oxford University Press, Oxford.
- Diggle, P., Farewell, D. and Henderson, R. (2007). Analysis of longitudinal data with drop-out: objectives, assumptions and a proposal. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 56, 499-550.
- Diggle, P. J., Sousa, I. and Chetwynd, A. G. (2008). Joint modeling of repeated measurements and time-to-event outcomes: the fourth Armitage lecture. *Statistics in Medicine* 27, 2981-2998.
- Ding, J. and Wang, J. L. (2008). Modeling longitudinal data with nonparametric multiplicative random effects jointly with survival data. *Biometrics* 64, 546-556.
- Dobson, A. and Henderson, R. (2003). Diagnostics for joint longitudinal and dropout time modeling. *Biometrics* 59, 741-751.
- Duchateau, L. and Jensen, P. (2008). *The Frailty Model*. 1st edition. Springer, New York.
- Dupuis, D. J. and Mills, J. E. (1998). Robust estimation of the Birnbaum-Saunders distribution. *IEEE Transactions on Reliability* 47, 88-95.
- Elashoff, R. M., Li, G. and Li, N. (2007). An approach to joint analysis of longitudinal measurements and competing risks failure time data. *Statistics in Medicine* 26, 2813-2835.
- Elashoff, R. M., Li, G. and Li, N. (2008). A joint model for longitudinal measurements and survival data in the presence of multiple failure types. *Biometrics* 64, 762-771.
- Engelhart, M., Bain, L. J. and , Wright, F. T. (1981). Inferences on the parameters of the Birnbaum-Saunders fatigue life distribution based on maximum likelihood estimation. *Technometrics* 23, 251-256.

- Faucett, C. L. and Thomas, D. C. (1996). Simultaneously modelling censored survival data and repeatedly measured covariates: a Gibbs sampling approach. *Statistics in Medicine* 15, 1663-1685.
- Faucett, C. L., Schenker, N. and Elashoff, R. M. (1998). Analysis of censored survival data with intermittently observed time-dependent binary covariates. *Journal of the American Statistical Association* 93, 427-437.
- Faucett, C. L., Schenker, N. and Taylor, J. M. G. (2002). Survival analysis using auxiliary variables via multiple imputation, with application to AIDS clinical trials data. *Biometrics* 58, 37-47.
- Freitas, H. F. G., Chizzola, P. R., Paes, A. T., Lima, A.C.P. and Mansur, A. J. (2005). Risk stratification in a brazilian hospital-based cohort of 1220 outpatients with heart failure: role of Chagas' heart disease. *International Journal of Cardiology* 102, 239-247.
- From, S. G. and Li, L. (2006). Estimation of the parameters of the Birnbaum- Saunders distribution. *Communications in Statistics: Theory and Methods* 35, 2157-2169.
- Galea, M., Leiva, V. and Paula, G.A. (2004). Influence diagnostics in log-Birnbaum-Saunders regression models. *Journal Applied Statistics* 31, 1049-1064.
- Gallant, A. R. and Nychka, D. W. (1987). Semi-nonparametric maximum likelihood estimation. *Econometrica* 55, 363-390.
- Garre, F., Zwinderman, A., Geskus, R. and Sijpkens, Y. (2008). A joint latent class changepoint model to improve the prediction of time to graft failure. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 171, 299-308.
- Giolo, S. R., Krieger, J. E., Mansur, A. J. and Pereira, A.C. (2012). Survival analysis of patients with heart failure: implications of time-varying regression effects in modeling mortality. *PLoS ONE* 7, e37392. doi:10.1371/journal.pone.0037392.
- Gomez, H., Pacheco, O. and Bolfarine, H. (2009). An extension of the generalized Birnbaum-Saunders distribution. *Statistics and Probability Letters* 79, 334-338.
- Guiraud, P., Leiva, V. and Fierro, R. (2009). A non-central version of the Birnbaum-Saunders distribution for reliability analysis. *IEEE Transactions on Reliability* 58, 152-160.

- Greene, W. F. and Cai, J. (2004). Measurement error in covariates in the marginal hazards model for multivariate failure time data. *Biometrics* 60, 987-996.
- Guo, X. and Carlin, B. P. (2004). Separate and joint modeling of longitudinal and event time data using standard computer packages. *The American Statistician* 58, 16-24.
- Hanson, T., Branscum, A. and Johnson, W. (2011). Predictive comparison of joint longitudinal-survival modeling: A case study illustrating competing approaches (with discussion). *Lifetime Data Analysis* 17, 3-28.
- Harville, D. A. (1974). Bayesian inference for variance components using only error contrast. *Biometrika* 61, 383-385.
- Harville, D. A. (1977). Maximum likelihood approaches to variance component estimation and to related problems. *Journal of the American Statistical Association* 72, 320-338.
- Henderson, R. and Oman, P. (1999). Effect of frailty on marginal regression estimates in survival analysis. *Journal of the Royal Statistical Society: Series B (Methodological)* 61, 367-379.
- Henderson, R., Diggle, P. and Dobson, A. (2000). Joint modeling of longitudinal measurements and event time data. *Biostatistics* 1, 465-480.
- Hogan, J. W. and Laird, N. M. (1997a). Mixture models for the joint distributions of repeated measures and event times. *Statistics in Medicine* 16, 239-257.
- Hogan, J. W. and Laird, N. M. (1997b). Model-based approaches to analysing incomplete longitudinal and failure time data. *Statistics in Medicine* 16, 259-272.
- Hogan, J. W., Lin, X. and Herman, B. (2004). Mixtures of varying coefficient models for longitudinal data with discrete or continuous nonignorable dropout. *Biometrics* 60, 854-864.
- Hogenhuis, J., Voors, A. A., Jaarsma, T., Hoes, A. W., Hillege, H. L., Kragten, J. A. and van Veldhuisen, D. J. (2007). Anaemia and renal dysfunction are independently associated with BNP and NT-proBNP levels in patients with heart failure. *The European Journal of Heart Failure* 9, 787-794.
- Hougaard, P. (1984). Life table methods for heterogeneous populations: Distributions describing the heterogeneity. *Biometrika* 71, 75 - 83.

- Hougaard, P. (1986a) . Survival models for heterogeneous populations derived from stable distributions. *Biometrika* 73, 387-396.
- Hougaard, P. (1986b). A class of multivariate failure time distributions. *Biometrika* 73, 671-678.
- Hougaard, P. (1991). Modeling heterogeneity in survival data. *Journal of Applied Probability* 28, 695-701.
- Hougaard, P. (2000). *Analysis of Multivariate Survival Data*. 1st edition. Springer-Verlag, New York.
- Hsieh, F., Tseng, Y. K. and Wang, J. L. (2006). Joint modeling of survival and longitudinal data: likelihood approach revisited. *Biometrics* 62, 1037-1043.
- Hu, P., Tsiatis, A. A. and Davidian, M. (1998). Estimating the parameters in the Cox model when covariate variables are measured with error. *Biometrics* 54, 1407-1419.
- Hu, C. and Lin, D. Y. (2002). Cox regression with covariate measurement error. *Scandinavian Journal of Statistics* 29, 637-655.
- Hu, W., Li, G. and Li, N. (2009). A bayesian approach to joint analysis of longitudinal measurements and competing risk failure time data. *Statistics in Medicine* 28, 1601-1619.
- Huang, X., Li, G. and Elashoff, R. M. (2010). A joint model of longitudinal and competing risk survival data with heterogeneous random effects and outlying longitudinal measurements. *Statistics and Its Interface* 3, 185-195.
- Huang, X., Li, G., Elashoff, R. and Pan, J. (2011). A general joint model for longitudinal measurements and competing risk survival data with heterogeneous random effects. *Lifetime Data Analysis* 17, 80-100.
- Ibrahim. J. G., Chen, M. H. and Sinha, D. (2001). *Bayesian Survival Analysis*. 1st edition. Springer, New York.
- Ibrahim. J. G., Chen, M. H. and Sinha, D. (2004). Bayesian methods for joint modeling of longitudinal and survival data with applications to cancer vaccine trials. *Statistica Sinica* 14, 863-883.
- Johnson, N. L., Kotz, S. and Balakrishnan, N. (1995). *Continuous Univariate Distributions*. Vol. 2. 2nd edition. John Wiley & Sons, New York.

- Kalbfleisch, J. D. and Prentice, R. L. (2002). *The Statistical Analysis of Failure Time Data*. 2nd edition. John Wiley & Sons, New Jersey.
- Klein, J. P. (1992). Semi-parametric estimation of random effects using the Cox model based on the EM algorithm. *Biometrics* 48, 795-806.
- Klein, J. P. and Moeschberger, M. L. (1997). *Survival Analysis: Techniques for Censored and Truncated Data*. 1st edition. Springer, New York.
- Kong, F. H. and Gu, M. (1999). Consistent estimation in Cox proportional hazards model with covariate measurement errors. *Statistica Sinica* 9, 953-969.
- Kundu, D., Kannan, N. and Balakrishnan, N. (2008). On the hazard function of Birnbaum-Saunders distribution and associated inference. *Computational Statistics and Data Analysis* 52, 2692-2702.
- Laird, N. M. and Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics* 38, 963-974.
- Lange, K. (2004). *Optimization*. 1st edition. Springer, New York.
- Lauer, M., Blackstone, E., Young, J. and Topol, E. (1999). Cause of death in clinical research: Time for a reassessment?. *Journal of the American College of Cardiology* 34, 618-620.
- Lavalley, M. P. and DeGruttola, V. (1996). Model for empirical Bayes estimators of longitudinal CD4 counts. *Statistics in Medicine* 15, 2289-2305.
- Law, N. J., Taylor, J. M. G. and Sandler, H. (2002). The joint modeling of a longitudinal disease progression marker and the failure time process in the presence of cure. *Biostatistics* 3, 547-563.
- Leiva, V., Barros, M. K., Paula, G. A. and Galea, M. (2007). Influence diagnostics in log-Birnbaum-Saunders regression models with censored data. *Computational Statistics and Data Analysis* 51, 5694-5707.
- Leiva, V., Barros, M., Paula, G. A. and Sanhueza, D. (2008a). Generalized Birnbaum-Saunders distributions applied to air pollutant concentration. *Environmetrics* 19, 235-249.

- Leiva, V., Riquelme, M., Balakrishnan, N. and Sanhueza, A. (2008b). Lifetime analysis based on the generalized Birnbaum-Saunders distribution. *Computational Statistics and Data Analysis* 52, 2079-2097.
- Leiva, V., Sanhueza, A., Sen, P. K., Paula, G. A. (2008c). Random number generators for the generalized Birnbaum-Saunders distribution. *Journal of Statistical Computation and Simulation* 78,1105-1118.
- Leiva, V., Barros, M., Paula, G. A. (2009). *Generalized Birnbaum-Saunders models using R*. XI Escola de Modelos de Regressão, Recife.
- Leiva, V., Sanhueza, A. and Angulo, J. M. (2009). A length-biased version of the Birnbaum-Saunders distribution with application in water quality. *Stochastic Environmental Research and Risk Assessment* 23, 299-307.
- Leiva, V., Athayde, E., Azevedo, C. and Marchant, C. (2011). Modeling wind energy flux by a Birnbaum-Saunders distribution with unknown shift parameter. *Journal of Applied Statistics* 38, 2819-2838.
- Lemonte, A., Cribari-Neto, F. and Vasconcellos, K. L. P. (2007). Improved statistical inference for the two-parameter Birnbaum-Saunders distribution. *Computational Statistics and Data Analysis* 51, 4656-4681.
- Lemonte, A., Simas, A., Cribari-Neto, F. (2008). Bootstrap-based improved estimators for the two-parameter Birnbaum-Saunders distribution. *Journal of Statistical Computation and Simulation* 78, 37-49.
- Lemonte, A. J. and Cordeiro, G. M. (2009). Birnbaum-Saunders nonlinear regression models. *Computational Statistics and Data Analysis* 53, 4441-4452.
- Lemonte, A. J., Ferrari, S. L. P. and Cribari-Neto, F. (2010). Improved likelihood inference in Birnbaum-Saunders regressions. *Computational Statistics and Data Analysis* 54, 1307-1316.
- Lemonte, A. J. (2013). A new extension of the Birnbaum-Saunders distribution. *Brazilian Journal of Probability and Statistics* 27, 133-149.
- Li, L., Hu, B. and Greene, T. (2009). A semiparametric joint model for longitudinal and survival data with application to hemodialysis study. *Biometrics* 65, 737-745.

- Li, N., Elashoff, R. M., Li, G. and Saver, J. (2010). Joint modeling of longitudinal ordinal data and competing risk survival times and analysis of the NINDS rt-PA stroke trial. *Statistics in Medicine* 29, 546-557.
- Lin, H., Turnbull, B. W., McCulloch, C. E. and Slate, E. H. (2002). Latent class models for joint analysis of longitudinal biomarker and event process data: application to longitudinal prostate-specific antigen readings and prostate cancer. *Journal of the American Statistical Association* 97, 53-65.
- Lin, H., Turnbull, B. W., McCulloch, C. E. and Slate, E. H. (2002). Latent class models for joint analysis of longitudinal biomarker and event process data: application to longitudinal prostate-specific antigen readings and prostate cancer. *Journal of the American Statistical Association* 97, 53-65.
- Lindstrom, M. J., and Bates, D. M. (1988). Newton-Raphson and EM algorithms for linear mixed-effects models for repeated measures data. *Journal of the American Statistical Association* 83, 1014-1022.
- Little, R. J. A. (1995). Modeling the drop-out mechanism in repeated-measures studies. *Journal of the American Statistical Association* 90, 1112-1121.
- Little, R. J. A. and Rubin, D. B. (2002). *Statistical Analysis with Missing Data*. 2nd edition. Wiley- Interscience, New Jersey.
- Liu, Q. and Pierce, D. A. (1994). A note on Gauss-Hermite quadrature. *Biometrika* 81, 624-629.
- Liu, L. and Huang, X. (2009). Joint analysis of correlated repeated measures and recurrent event processes in the presence of death, with application to a study on acquired immune deficiency syndrome. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 58, 65-81.
- Lu, M. and Chang, D. S. (1997). Bootstrap prediction intervals for the Birnbaum-Saunders distribution. *Microelectron Reliability* 37, 1213-1216.
- Maisel, A. (2002). B-type natriuretic peptide levels: diagnostic and prognostic in congestive heart failure: what's next?. *Circulation* 105, 2328-2331.
- Maller, R. and Zhou, X. (1996). *Survival Analysis with Long-Term Survivors*. 1st edition. John Wiley and Sons, Chichester.

- McGilchrist, C. A. and Aisbett, C. W. (1991). Regression with frailty in survival analysis. *Biometrics* 47, 461-466.
- McGilchrist, C. A. (1993). REML estimation for survival models with frailty. *Biometrics* 49, 221-225.
- Murphy, S. A. (1994). Consistency in a proportional hazards model incorporating a random effect. *The Annals of Statistics* 22, 712-731.
- Murphy, S. A. (1995). Asymptotic theory for the frailty model. *The Annals of Statistics* 23, 182-198.
- Nakamura, T. (1990). Corrected score function for errors-in-variables models: methodology and application to generalized linear models. *Biometrika* 77, 127-137.
- Nakamura, T. (1992). Proportional hazards model with covariates subject to measurement error. *Biometrics* 48, 829-838.
- Ng, H. K. T., Kundu, D. and Balakrishnan, N. (2006). Point and interval estimations for the two parameter Birnbaum-Saunders distributions based on type II censored samples. *Computational Statistics and Data Analysis* 50, 3222-3242.
- Nielsen, G. G., Gill, R. D., Andersen, P. K. and Sorensen, T. I. A. (1992). A counting process approach to maximum likelihood estimation in frailty models. *Scandinavian Journal of Statistics* 19, 25-43.
- Njeru Njagi, E., Rizopoulos, D., Molenberghs, G., Dendale, P. and Willekens, K. (2013). A joint survival-longitudinal modelling approach for the dynamic prediction of rehospitalization in telemonitored chronic heart failure patients. *Statistical Modelling* 13, 179-198.
- Ortega, E. M. M., Cordeiro, G. M. and Lemonte, A. J. (2012). A log-linear regression model for the β -Birnbaum-Saunders distribution with censored data. *Computational Statistics and Data Analysis* 56, 698-718.
- Owen, W. J. and Padgett, W. J. (1999). Accelerated test models for system strength based on Birnbaum-Saunders distribution. *Lifetime Data Analysis* 5, 133-147.
- Owen, W. J. and Padgett, W. J. (2000). A Birnbaum-Saunders accelerated life model. *IEEE Transactions on Reliability* 49, 224-229.

- Owen, W. J. (2006). A new three-parameter extension to the Birnbaum-Saunders distribution. *IEEE Transactions on Reliability* 55, 475-479.
- Paula, G. A., Leiva, V., Barros, M. and Liu, S. (2011). Robust statistical modeling using the Birnbaum-Saunders-t distribution applied to insurance. *Applied Stochastic Model in Business and Industry*, doi: 10.1002/asmb.887.
- Pawitan, Y. and Self, S. (1993). Modeling disease marker processes in AIDS. *Journal of the American Statistical Association* 83, 719-726.
- Petersen, J. H. (1998). An additive frailty model for correlated life times. *Biometrics* 54, 646-661.
- Philipson, P., Sousa, I., Diggle, P., Williamson, P., Kolamunnage-Dona, R. and Henderson, R. (2012). *joineR: Joint Modelling of Repeated Measurements and Time-to-Event Data*. R package version 1.0-3.
<http://cran.r-project.org/web/packages/joineR/index.html>.
- Pinheiro, J. C. and Bates, D. M (1995). Approximations to the log-likelihood function in the nonlinear mixed-effects model. *Journal of Computational and Graphical Statistics* 4, 12-35.
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. and R Development Core Team (2013). *nlme: Linear and Nonlinear Mixed-Effects Models*. R package version 3.1-113.
<http://cran.r-project.org/web/packages/nlme/index.html>.
- Podlaski, R. (2008). Characterization of diameter distribution data in near-natural forests using the Birnbaum-Saunders distribution. *Canadian Journal of Forest Research* 18, 518-526.
- Prentice, R. L. (1982). Covariate measurement errors and parameter estimates in a failure time regression model. *Biometrika* 69, 331-342.
- Prentice, R. L. (1989). Surrogate endpoints in clinical trials: definition and operation criteria. *Statistics in Medicine* 8, 431-440.
- Proust-Lima, C. and Taylor, J. (2009). Development and validation of a dynamic prognostic tool for prostate cancer recurrence using repeated measures of posttreatment PSA: A joint modeling approach. *Biostatistics* 10, 535-549.

- R Development Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna.
<http://www.r-project.org/>.
- Raboud, J., Reid, N., Coates, R. A. and Farewell, V. T. (1993). Estimating risks of progressing to AIDS when covariates are measured with error. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 156, 393-406.
- Ratcliffe, S. J., Guo, W. and Ten Have, T. R. (2004). Joint modeling of longitudinal and survival data via a common frailty. *Biometrics* 60, 892-899.
- Rieck, J. R. and Nedelman, J. R. (1991). A log-linear model for the Birnbaum-Saunders distribution. *Technometrics* 33, 51-60.
- Rieck, J. R. (1999). A moment-generating function with application to the Birnbaum-Saunders distribution. *Communications in Statistics: Theory and Methods* 28, 2213-2222.
- Ripatti, S. and Palmgren, J. (2000). Estimation of multivariate frailty models using penalized partial likelihood. *Biometrics* 56, 1016-1022.
- Ripatti, S., Larsen, K. and Palmgren, J. (2002). Maximum likelihood inference for multivariate frailty models using a Monte Carlo EM Algorithm. *Lifetime Data Analysis* 8, 349-360.
- Rizopoulos, D. (2010). JM: An R package for the joint modelling of longitudinal and time-to-event data. *Journal of Statistical Software* 35, 1-33.
- Rizopoulos, D. (2011). Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. *Biometrics* 67, 819-829.
- Rizopoulos, D. (2012a). Fast fitting of joint models for longitudinal and event time data using a pseudo-adaptative Gaussian quadrature rule. *Computational Statistics and Data Analysis* 56, 491-501.
- Rizopoulos, D. (2012b). *Joint Models for Longitudinal and Time-to-Event Data With Applications in R*. 1st edition. Chapman-Hall/CRC Biostatistics Series, Boca Raton.
- Rizopoulos, D. (2012c). *JM: Shared Parameter Models for the Joint Modelling of Longitudinal and Time-to-Event Data*. R package version 1.2-0.
<http://cran.r-project.org/web/packages/JM/index.html>.

- Rizopoulos, D. and Ghosh, P. (2011). A bayesian semiparametric multivariate joint model for multiple longitudinal outcomes and time-to-event. *Statistics in Medicine* 30, 1366-1380.
- Rizopoulos, D., Verbeke, G. and Lesaffre, E. (2009). Fully exponential Laplace approximations for the joint modelling of survival and longitudinal data. *Journal of the Royal Statistical Society: Series B (Methodological)* 71, 637-654.
- Rizopoulos, D., Verbeke, G. and Molenberghs, G. (2008). Shared parameter models under random effects misspecification. *Biometrika* 95, 63-74.
- Rizopoulos, D., Verbeke, G. and Molenberghs, G. (2010). Multiple-imputation-based residuals and diagnostic plots for joint models of longitudinal and survival outcomes. *Biometrics* 66, 20-29.
- Rizopoulos, D., Verbeke, G., Lesaffre, E. and Vanrenterghem, Y. (2008). A two-part joint model for the analysis of survival and longitudinal binary data with excess zeros. *Biometrics* 64, 611-619.
- Rubin, D. B. (1976). Inference and missing data. *Biometrika* 63, 581-592.
- Sanhueza, A., Leiva, V. e Balakrishnan, N. (2008). The generalized Birnbaum-Saunders distribution and its theory, methodology and application. *Communications in Statistics: Theory and Methods* 37, 645-670.
- Santos-Neto, M., Cysneiros F. J. A., Leiva, V. and Ahmed, S. E. (2012). On new parameterizations of the Birnbaum-Saunders distribution. *Pakistan Journal of Statistics* 28, 1-26.
- Schluchter, M. D. (1992). Methods for the analysis of informatively censored longitudinal data. *Statistics in Medicine* 11, 1861-1870.
- Smyth, G., Hu, Y., Dunn, P., Phipson, B. and Chen, Y. (2013). *statmod: Statistical Modelling*. R package version 1.4-18.
<http://cran.r-project.org/web/packages/statmod/index.html>.
- Song, X., Davidian, M. and Tsiatis, A. A. (2002a). An estimator for the proportional hazards model with multiple longitudinal covariates measured with error. *Biostatistics* 3, 511-528.

- Song, X., Davidian, M. and Tsiatis, A. A. (2002b). A semiparametric likelihood approach to joint modeling of longitudinal and time-to-event data. *Biometrics* 58, 742-753.
- Song, X. and Wang, C. V. (2008). Semiparametric approaches for joint modeling of longitudinal and survival data with time-varying coefficients. *Biometrics* 64, 557-566.
- Stefanski, L. A. and Carroll, R. (1987). Conditional scores and optimal scores in generalized linear measurement error models. *Biometrika* 74, 703-716.
- Tapsoba, J. D., Lee, S. M. and Wang, C. Y. (2011). Joint modeling of survival time and longitudinal data with subject-specific changepoints in the covariates. *Statistics in Medicine* 30, 232-249.
- Taylor, J. M. G., Cumberland, W. G. and Sy, J. P. (1994). A stochastic model for analysis of longitudinal data. *Journal of the American Statistical Association* 89, 727-736.
- Therneau, T. M. and Grambsch, P. M. (2000). *Modeling Survival Data: Extending the Cox Model*. 1st edition. Springer-Verlag, New York.
- Therneau, T. M. and Lumley, T. (2012). *survival: Survival Analysis*. R package version 2.36-12.
<http://cran.r-project.org/web/packages/survival/index.html>.
- Tseng, C. and Wong, W. (2011). *Analysis of a composite endpoint with longitudinal and time-to-event data*. *Statist. Medicine* 30, 1018-1027.
- Tseng, Y. K., Hsieh, F., and Wang, J. L. (2005). Joint modelling of accelerated failure time and longitudinal data. *Biometrika* 92, 587-603.
- Tsiatis, A. A., DeGruttola, V. and Wulfsohn, M. S. (1995). Modeling the relationship of survival to longitudinal data measured with error: applications to survival and CD4 counts in patients with AIDS. *Journal of the American Statistical Association* 90, 27-37.
- Tsiatis, A. A. and Davidian, M. (2001). A semiparametric estimator for the proportional hazards model with longitudinal covariates measured with error. *Biometrika* 88, 447-458.
- Tsiatis, A. A. and Davidian, M. (2004). Joint modeling of longitudinal and time-to-event data: an overview. *Statistica Sinica* 14, 809-834.

- Tsionas, E. G. (2001). Bayesian inference in Birnbaum-Saunders regression. *Communications in Statistics: Theory and Methods* 30, 179-193.
- Vaida, F. and Xu, R. (2000). Proportional hazards model with random effects. *Statistics in Medicine* 19, 3309-3324.
- Vaupel, J. W., Manton, K. G. and Stallard, E. (1979). The impact of heterogeneity in individual frailty and the dynamics of mortality. *Demography* 16, 439-454.
- Verbeke, G., Lesaffre, E. and Brant, L. J. (1998). The detection of residual serial correlation in linear mixed models. *Statistics in Medicine* 17, 1391-1402.
- Verbeke, G. and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*. 1st edition. Springer, New York.
- Vilca-Labra, F. and Leiva, V. (2006). A new fatigue life model based on the family of skew-elliptical distributions. *Communications in Statistics: Theory and Methods* 35, 229-244.
- Vilca, F., Sanhueza, A., Leiva, V. and Christakos, G. (2010). An extended Birnbaum-Saunders model and its application in the study of environmental quality in Santiago, Chile. *Stochastic Environmental Research and Risk Assessment* 24, 771-782.
- Vilca, F., Santana, L., Leiva, V. and Balakrishnan, N. (2011). Estimation of extreme percentiles in Birnbaum-Saunders distributions. *Computational Statistics and Data Analysis* 55, 1665-1678.
- Villegas, C., Paula, G. A. and Leiva, V. (2011). Birnbaum-Saunders mixed models for censored reliability data analysis. *IEEE Transactions on Reliability* 60, 748-758.
- Vonesh, E. F., Greene, T. and Schluchter, M. D. (2006). Shared parameter models for the joint analysis of longitudinal data and event times. *Statistics in Medicine* 25, 143-163.
- Wang, Y. and Taylor, J. M. G. (2001). Jointly modeling longitudinal and event time data with application to acquired immunodeficiency syndrome. *Journal of the American Statistical Association* 96, 895-905.
- Wang, C. Y. (2006). Corrected score estimator for joint modeling of longitudinal and failure time data. *Statistica Sinica* 16, 235-253.

- Weber, M. and Hamm, C. (2006). Role of B-type natriuretic peptide (BNP) and NT-proBNP in clinical routine. *Heart* 92, 843-849.
- Wei, L. J., Lin, D. Y. and Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure data by modeling marginal distributions. *Journal of the American Statistical Association* 84, 1065-1073.
- Wienke, A. (2011). *Frailty Models in Survival Analysis*. 1st edition. Chapman and Hall, Boca Raton.
- Williamson, P. R., Kolamunnage-Dona, R., Philipson, P. and Marson, A. G. (2008). Joint modelling of longitudinal and competing risk data. *Statistics in Medicine* 27, 6426-6438.
- Wu, M. C. and Carroll, R. J. (1988). Estimation and comparison of changes in the presence of informative right censoring by modeling the censoring process. *Biometrics* 44, 175-188.
- Wu, L. Hu, X. J. and Wu, H. (2008). Joint inference for nonlinear mixed-effects models and time to event at the presence of missing data. *Biostatistics* 9, 308-320.
- Wu, L. Liu, W. and Hu, X. J. (2010). Joint inference on HIV viral dynamics and immune suppression in presence of measurement errors. *Biometrics* 66, 327-335.
- Wu, L. Liu, W., Yi, G. Y. and Huang, Y. (2012). Analysis of longitudinal and survival data: Joint modeling, inference methods, and issues. *Journal of Probability and Statistics*, ID640153. doi:10.1155/2012/640153.
- Wulfsohn, M. S. and Tsiatis, A. A. (1997). A joint model for survival and longitudinal data measured with error. *Biometrics* 53, 330-339.
- Xie, F. C. and Wei, B. C. (2007). Diagnostics analysis for log-Birnbaum-Saunders regression models. *Computational Statistics and Data Analysis* 51, 4692-4706.
- Xu, J. and Zeger, S. L. (2001a). Joint analysis of longitudinal data comprising repeated measures and times to events. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 50, 375-387.
- Xu, J. and Zeger, S. L. (2001b). The evaluation of multiple surrogate endpoints. *Biometrics* 57, 81-87.

- Yau, K. K. W. and McGilchrist, C. A. (1998). ML and REML estimation in survival analysis with time dependent correlated frailty. *Statistics in Medicine* 17, 1201-1213.
- Ye, W., Lin, X. and Taylor J. M. G. (2008a). A penalized likelihood approach to joint modeling of longitudinal measurements and time-to-event data. *Statistics and Its Interface* 1, 33-45.
- Ye, W., Lin, X. and Taylor J. M. G. (2008b). Semiparametric modeling of longitudinal measurements and time-to-event data: a two-stage regression calibration approach. *Biometrics* 64, 1238-1246.
- Yu, M., Law, N. J., Taylor, J. M. G. and Sandler, H. M. (2004). Joint longitudinal-survival-cure models and their application to prostate cancer. *Statistica Sinica* 14, 835-862.
- Yu, M., Taylor. J. M. G. and Sandler, H. M. (2008). Individual prediction in prostate cancer studies using a joint longitudinal survival cure model. *Journal of the American Statistical Association* 103, 178-187.
- Zeger, S. L. and Diggle, P. J. (1994). Semiparametric models for longitudinal data with application to CD4 cell numbers, HIV seroconverters. *Biometrics* 5, 689-699.
- Zeng, D. and Cai, J. (2005a). Asymptotic results for maximum likelihood estimators in joint analysis of repeated measurements and survival time. *Annals of Statistics* 33, 2132-2163.
- Zeng, D. and Cai, J. (2005b). Simultaneous modelling of survival and longitudinal data with application to repeated quality of life measures. *Lifetime Data Analysis* 11, 151-174.
- Zhang, D., Lin, X., Raz, J., and Sowers, M. (1998). Semiparametric stochastic mixed models for longitudinal data. *Journal of the American Statistical Association* 93, 710-719.
- Zhang, D. and Davidian, M. (2001). Linear mixed models with flexible distributions of random effects for longitudinal data. *Biometrics* 57, 795-802.
- Zhang, S., Müller, P. and Do, K. A. (2010). A bayesian semiparametric survival model with longitudinal markers. *Biometrics* 66, 435-443.