

UNIVERSIDADE ESTADUAL DE CAMPINAS  
INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E COMPUTAÇÃO CIENTÍFICA  
DEPARTAMENTO DE ESTATÍSTICA

**Estimação de Máxima Verossimilhança  
para Modelos de Sobrevivência  
com Efeitos Aleatórios**

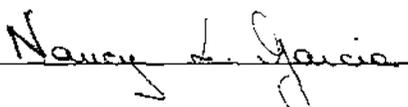
**ANDRÉA FERREIRA**

**Dissertação de Mestrado orientada pela  
Profa. Dra. Nancy Lopes Garcia**

**Estimação de Máxima Verossimilhança  
para Modelos de Sobrevivência  
com Efeitos Aleatórios**

Este exemplar corresponde à redação final da dissertação devidamente corrigida e defendida por ANDRÉA FERREIRA e aprovada pela comissão julgadora.

Campinas, 25 de março de 1999



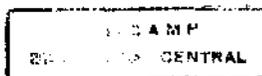
Profa. Dra. Nancy Lopes Garcia.

Orientadora

Banca Examinadora:

- 1 Profa. Dra. Nancy Lopes Garcia (orientadora)
- 2 Profa. Dra. Cicilia Yuko Wada
- 3 Profa. Dra. Aída Verdugo Lazo

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica, UNICAMP, como requisito parcial para obtenção do Título de MESTRE EM ESTATÍSTICA.



UNIDADE	BC
N.º CHAMADA	UNICAMP
	F413e
V.	
TÍTULO	37832
PP.	229/99
	R\$ 11,00
	10/06/99
N.º CPD	

CM-00124521-B

**FICHA CATALOGRÁFICA ELABORADA PELA  
BIBLIOTECA DO IMECC DA UNICAMP**

Ferreira, Andréa

F413e      Estimação de máxima verossimilhança para modelos de sobrevivência com efeitos aleatórios / Andréa Ferreira -- Campinas, [S.P. :s.n.], 1999.

Orientador : Nancy Lopes Garcia

Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Análise de sobrevivência. 2. Bioestatística. I. Garcia, Nancy Lopes. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

Dissertação de Mestrado defendida em 25 de março de 1999 e aprovada  
pela Banca Examinadora composta pelos Profs. Drs.

  
\_\_\_\_\_

Prof (a). Dr (a). NANCY LOPES GARCIA

  
\_\_\_\_\_

Prof (a). Dr (a). CÍCILIA YUKO WADA

  
\_\_\_\_\_

Prof (a). Dr (a). AIDA CECÍLIA GRACIELA VERDUGO LAZO

## Resumo

O objetivo desta dissertação foi verificar a consistência e a normalidade assintótica dos estimadores da variância dos **efeitos aleatórios** (fragilidade), assumindo três diferentes distribuições paramétricas para a fragilidade, gama, lognormal e normal. É de grande interesse verificar a consistência do estimador da variância da fragilidade, pois sendo ele consistente é possível testar se a população estudada é homogênea ou não.

Através das simulações realizadas, verificamos que o estimador da variância da fragilidade é sensível quanto a escolha da distribuição da fragilidade. Para o caso da fragilidade assumir a distribuição gama o estimador se mostrou consistente e assintoticamente normal, mas para as distribuições lognormal e normal ele não foi consistente.

Foi também verificado a estimação da função risco acumulada, mesmo sendo os estimadores da variância da fragilidade viciados, e podemos observar que a sua estimação foi comprometida com relação as distribuições assumidas para a fragilidade.

# Abstract

The goal of this work is to verify the consistency and asymptotic normality of one estimator suggested in the literature regarding the variance of the random effect (frailty), assuming that the distribution of the frailty follows one of the three models: gamma, lognormal and normal. It is of great interest to verify consistency and asymptotic normality of an estimator since these are the properties we would need if we were interested in testing whether the population is homogeneous or not.

Through the simulation results we verify that the estimator of the variance of the random effect is very sensitive to the choice of the distribution. If the frailty is gamma-distributed, the estimator is consistent and asymptotically normal, however for the lognormal distribution it underestimate the variance and under the normal model it overestimate the variance. Moreover, we verify that the estimation of the cumulative risk function is not good for all the models.

Dedico aos meus pais,  
à minha irmã e  
ao meu noivo.

# Índice

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Apresentação e Motivação . . . . .	1
1.2	Conceitos Básicos . . . . .	4
<b>2</b>	<b>Modelos de riscos com heterogeneidade não observada</b>	<b>8</b>
2.1	Estimação de Máxima Verossimilhança . . . . .	9
2.2	Estimação da função risco e do parâmetro de fragilidade, utilizando o algoritmo EM . . . . .	11
2.2.1	Modelos Paramétricos . . . . .	11
2.2.2	Modelos Semi-Paramétricos . . . . .	13
2.3	Convergência do estimador do parâmetro de fragilidade . . . . .	15
2.4	Cálculo da variância assintótica $\sigma^2$ . . . . .	20
<b>3</b>	<b>Simulações</b>	<b>26</b>
3.1	Resultados das Simulações . . . . .	29
3.1.1	Fragilidade Gama . . . . .	30
3.1.2	Fragilidade Lognormal . . . . .	49
3.1.3	Fragilidade Normal . . . . .	54
<b>4</b>	<b>Cálculo da Função Risco Acumulada estimada, através da fragilidade</b>	<b>60</b>
4.1	Comparação da função risco acumulada estimada com a verdadeira função risco acumulada . . . . .	61

<b>5 Considerações Finais</b>	<b>77</b>
<b>Bibliografia</b>	<b>79</b>
<b>Anexo Listagem do Programa</b>	<b>83</b>

# Capítulo 1

## Introdução

### 1.1 Apresentação e Motivação

Nas duas últimas décadas tem ocorrido um grande avanço em métodos para análise de duração<sup>1</sup> de eventos. Seja na engenharia ou nas ciências médicas existe um crescente interesse no desenvolvimento de tais métodos. Por exemplo, um dos principais interesses nas ciências biológicas, é estudar o tempo de vida de pacientes submetidos a transplantes e/ou tratamentos, através de métodos conhecidos como análise de sobrevivência. Na engenharia, a análise do tempo de vida de componentes elétricos e vários tipos de equipamentos industriais também é conhecida como análise de confiabilidade. Este desenvolvimento tem sido aplicado não só nas áreas médicas e na engenharia, mas também em outras áreas como demografia, ciências sociais e economia, sendo que foi na demografia que surgiu o primeiro e o mais antigo método para analisar dados sobre duração de eventos, conhecido como tábua de vida.

Um dos problemas na análise de dados de duração de eventos é que, muitas vezes não é possível observar-se o fenômeno até a ocorrência do evento e o tempo de falha é dito ser censurado. Isto é, sabe-se que o tempo de ocorrência do evento excede certo valor, mas não

---

<sup>1</sup>tempo de espera para algum evento ocorrer

se conhece o tempo exato da ocorrência. No entanto, existem métodos de estimação que permitem incorporar no modelo de risco observações censuradas, produzindo (sob certas condições) estimativas não-viciadas dos parâmetros de interesse. Um dos métodos mais utilizados na estimação dos modelos paramétricos e semi-paramétricos é o de máxima verossimilhança.

Frequentemente estes dados são analisados utilizando-se o modelo de riscos proporcionais semi-paramétrico de Cox(1972), que possibilita estudar a distribuição do tempo de ocorrência dos eventos, sem especificar a sua forma, pois em geral não temos conhecimento a priori da sua distribuição. Mas este modelo não permite incorporar na análise covariáveis não observáveis que representam a heterogeneidade populacional, que pode existir mas que a princípio não pode ser determinada.

Um outro problema na análise de duração dos eventos está justamente na influência da heterogeneidade populacional sobre o tempo de ocorrência do evento. É relevante considerar a população como uma mistura de indivíduos com riscos diferentes, sendo que cada indivíduo na população tem a sua própria fragilidade, ou seja, a ocorrência de um evento varia de um indivíduo para o outro, de acordo com a propensão individual de cada um.

A maioria das análises feitas com modelos de risco estão baseadas na suposição de que toda heterogeneidade da população de interesse é explicada por covariáveis observadas e que estas foram incluídas no modelo. Considerando-se que nem toda heterogeneidade possa ser explicada pelas covariáveis observadas, o modelo estimado pode estar sendo prejudicado. A heterogeneidade pode ser causada por uma covariável não observada que se não for incluída no modelo pode conduzir a diferentes conclusões, produzindo estimativas viciadas dos parâmetros de interesse e portanto reduzindo a eficiência do modelo (Heckman e Singer, 1982; Vaupel e Yashin, 1985; Trussel e Rodriguez, 1990).

Muitas análises em epidemiologia e em estudos prognósticos requerem métodos que permitam incorporar covariáveis não-observáveis no modelo. Por exemplo, em eventos que resultam de estudos sobre incidência de doenças, pode existir uma tendência para

a ocorrência da doença dentro de um determinado grupo familiar, pode ser uma característica genética não-observável ou um efeito ambiental, se os indivíduos são irmãos, familiares ou residem no mesmo domicílio. O curso natural de várias doenças varia de pessoa para pessoa, bem como o efeito do tratamento ou a influência de vários outros fatores de risco. Esta heterogeneidade é frequentemente chamada de variação biológica sendo reconhecida como uma das mais importantes fontes de variação na medicina e na biologia (Aalen, 1988).

O modelo de fragilidade é uma generalização do modelo de riscos proporcionais de Cox, que permite incluir no modelo de risco, um efeito aleatório representando a heterogeneidade não observada de cada indivíduo. Os modelos de fragilidade foram introduzidos por Vaupel, Manton e Stallard (1979) que apresentaram pela primeira vez a idéia de fragilidade e o impacto da heterogeneidade não observada na análise de duração dos eventos.

A estimação em modelos de fragilidade tem recebido particular atenção e está sendo desenvolvida e aplicada em várias áreas de estudo. Por exemplo, em análise de sobrevivência encontram-se entre outros Aalen(1988), Clayton (1991), Klein (1992), Nielsen *et al.* (1992), Murphy (1994 e 1995). Em demografia pode-se citar Trussel e Rodrigues(1990), Manton, Singer e Woodbury(1992). Na área de ciências sociais situam-se Allison(1982), Namboodiri e Suchindran(1987), Blossfeld e Hamerle(1992) e em econometria Heckman e Singer(1982).

No modelo de riscos proporcionais usual, os estimadores de máxima verossimilhança paramétricos e não paramétricos (estimadores de Nelson Aalen) do risco padrão acumulado mostram-se consistente e assintoticamente eficiente (GreenWood e Wefelmeyer,1990). Nielsen *et al.* (1992) propuseram o algoritmo EM para estimar o risco padrão acumulado e a variância do efeito aleatório. Recentemente Murphy(1994 e 1995) provou a consistência e a normalidade assintótica do estimador da variância da fragilidade, assumindo que a fragilidade tem distribuição gama.

A suposição de fragilidade com distribuição gama tem sido usada por vários autores. A justificativa para esta escolha é baseada na habilidade da distribuição assumir uma varie-

dade de formas e na simplicidade analítica. Mas muitas outras distribuições paramétricas podem ser utilizadas, tais como normal, log-normal, beta e outras (Heckman e Singer, 1982; Hougaard, 1984 e 1986).

Neste trabalho inicialmente são apresentados alguns conceitos importantes no contexto de análise de duração de eventos, que serão utilizados posteriormente nesta dissertação. A seguir, no capítulo 2 serão apresentados e discutidos estratégias paramétricas e não paramétricas para modelar a heterogeneidade não-observada em modelos de riscos, utilizando o algoritmo EM, conforme desenvolvido por Nielsen *et al.*(1992). Foram também verificados a consistência e a normalidade assintótica do estimador do parâmetro de fragilidade proposto nesta dissertação, quando a fragilidade assume a distribuição gama. Baseado em Murphy (1994 e 1995), o principal objetivo desta dissertação foi verificar quão sensíveis são as estimativas dos parâmetros do modelo com relação à escolha da distribuição para a fragilidade. Portanto, no capítulo 3 foi realizado um estudo de simulação, considerando-se além da distribuição gama para a fragilidade, as distribuições lognormal e normal. Através desta simulação foi possível fazer uma comparação com o estudo de simulação realizado por Nielsen(1992) e verificar o comportamento do estimador do parâmetro de interesse com relação a escolha da distribuição da fragilidade. No capítulo 4, utilizando os dados simulados no capítulo anterior, foi feita uma comparação entre a função risco acumulada verdadeira e a função risco acumulada estimada para as distribuições de fragilidade gama, lognormal e normal. Finalizando, no capítulo 5 é apresentado um resumo geral e algumas conclusões a respeito do trabalho realizado.

## 1.2 Conceitos Básicos

Suponha que a população estudada seja homogênea, isto é, o tempo de falha de todas as observações são independentes e identicamente distribuídas. Considere uma amostra de  $n$  observações independentes de pares  $(X_i, \delta_i)$ ,  $i = 1, 2, \dots, n$ , onde  $X_i$  é o tempo de falha ou censura da  $i$ -ésima observação. Para definir se o tempo é de falha ou censura, é

criada uma variável dicotômica:  $\delta_i = 1$  se ocorreu a falha na  $i$ -ésima observação e  $\delta_i = 0$  caso contrário. Denota-se  $T_i$  como a variável aleatória do tempo de falha da  $i$ -ésima observação e  $C_i$ , a variável aleatória do tempo de censura. Então  $X_i = \min(T_i, C_i)$  e

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

Suponha  $0 = t_0 < t_1 < \dots < t_m = t$  seja uma partição do intervalo  $[0, t]$ , e  $d_l$  e  $y_l$  sejam, respectivamente, o número de falhas em  $[t_{l-1}, t_l]$  e o número de casos conhecidos que não falharam antes do tempo  $t_{l-1}$ . O valor de  $y_l$  é também chamado o número de casos em risco até o tempo  $t_{l-1}$ .

A probabilidade de um indivíduo sobreviver até o tempo  $t$  é dada pela função de sobrevivência, definida por

$$S(t) = P(T \leq t).$$

Se a distribuição de probabilidade de  $T$  é conhecida, é possível calcular o estimador paramétrico de  $S(t)$ , apenas estimando os parâmetros da distribuição. Caso a distribuição de probabilidade de  $T$  seja desconhecida, pode-se calcular o estimador não paramétrico de  $S(t)$ , definido como

$$\hat{S}(t) = \prod_{l: t_l \leq t} \left[ 1 - \frac{d_l}{y_l} \right],$$

também conhecido como estimador produto limite (PL) da função de sobrevivência, ou estimador de Kaplan-Meier.

Pode-se estimar a função risco acumulada  $A(t) = -\log[S(t)]$  (Lawless 1982), ou seja, o risco de uma pessoa morrer ou de um componente falhar, ao longo do tempo. Um estimador natural é dado por

$$A^*(t) = -\log[\hat{S}(t)] = - \sum_{l: t_l \leq t} \log \left( 1 - \frac{d_l}{y_l} \right),$$

onde  $\hat{S}(t)$  é o estimador da função de sobrevivência. A função risco acumulada também pode ser escrita como  $A(t) = \int_0^t \alpha(u) du$ , sendo  $\alpha(u)$  a função risco básica. Outro estimador

sugerido de  $A(t)$  tem a seguinte forma

$$\tilde{A}(t) = \sum_{i: t_i \leq t} \frac{d_i}{y_i}$$

que também é chamado de função risco acumulada empírica. Para modelos contínuos  $A^*(t)$  e  $\tilde{A}(t)$  são assintoticamente equivalentes e não diferem grandemente em muitas situações, exceto para valores grandes de  $t$ . Neste caso o estimador da função risco acumulada no tempo  $t$  será o mesmo para toda a população sob estudo.

Considere os processos  $N_i$  e  $Y_i$

$$N_i(t) = I_{\{X_i \leq t, \delta_i = 1\}}$$

e

$$Y_i(t) = I_{\{X_i \geq t\}}$$

que equivalentemente contém as mesmas informações contidas nos pares  $(X_i, \delta_i)$ , onde  $N_i(t) = 1$  indica que a  $i$ -ésima observação falhou antes ou exatamente no tempo  $t$ , e  $Y_i = 1$  indica que até o tempo  $t$  a  $i$ -ésima observação ainda está em risco de falhar ou ser censurada. Esse tipo de notação será útil posteriormente para definir a função risco e a função risco acumulado de cada indivíduo da população.

Um outro estimador  $\hat{A}$  frequentemente referido como o estimador da função risco acumulada, foi proposto por Nelson(1969). Se  $\bar{N} = \sum_i N_i$  e  $\bar{Y} = \sum_i Y_i$ , então este estimador pode ser escrito como

$$\hat{A}(t) = \int_0^t \frac{I_{\{\bar{Y}(u) > 0\}}}{\bar{Y}(u)} d\bar{N}(u).$$

## Modelo de Riscos

Suponha que  $T$  seja a variável de interesse e que existam covariáveis que são representadas pelo vetor  $\mathbf{Z}' = (z_1, z_2, \dots, z_p)$ , e foram coletadas no tempo  $t_0$ . Pode ser de interesse estimar a função risco acumulada, sob certas condições iniciais e experimentais dos indivíduos, de acordo com os objetivos do estudo.

Em geral, não se tem conhecimento da distribuição da variável  $T$ , e portanto, a escolha não adequada da distribuição pode comprometer as conclusões com respeito a inferências sobre a população em estudo.

O propósito dos modelos de regressão é determinar como  $T$  varia com  $\mathbf{Z}$  usando um modelo apropriado para a dependência de  $T$  e  $\mathbf{Z}$ , com dados do tipo  $(T_i, \mathbf{Z}_i)$ ,  $i = 1, 2, \dots, n$ , tendo por finalidade estudar o efeito desses fatores na função de sobrevivência e na função risco acumulada.

Um modelo de regressão paramétrica conhecido é o modelo de riscos proporcionais. O modelo é dito ser de riscos proporcionais se a razão de riscos entre dois indivíduos com níveis de uma covariável  $Z$ , por exemplo  $z_1$  e  $z_2$ , não varia com o tempo, ou seja

$$\lambda(t|z_1) = \text{constante} \times \lambda(t|z_2)$$

Se isso ocorre, a função risco de  $T$  dado  $\mathbf{Z}$  pode ser escrita como

$$\lambda(t|\mathbf{Z}) = \alpha_0(t) \varphi(\boldsymbol{\beta}, \mathbf{Z})$$

onde  $\alpha_0(t)$  é uma função risco básica e  $\varphi(\boldsymbol{\beta}, \mathbf{Z})$  é uma função que depende de  $\mathbf{Z}' = (z_1, z_2, \dots, z_p)$ , com  $\boldsymbol{\beta}' = (\beta_1, \beta_2, \dots, \beta_p)$  o vetor de parâmetros desconhecidos associado a  $\mathbf{Z}$ . O modelo de riscos proporcionais assume, portanto, que  $\mathbf{Z}$  tem efeito multiplicativo na função risco básica.

Quando  $\varphi(\boldsymbol{\beta}, \mathbf{Z})$  é igual a  $\exp(\boldsymbol{\beta}' \mathbf{Z})$  então o modelo é considerado como modelo de riscos proporcionais de Cox(1972) semi-paramétrico, que possui a vantagem de tomar somente valores positivos e não exigir nenhuma restrição aos valores  $\boldsymbol{\beta}$  para garantir que  $\varphi(\boldsymbol{\beta}, \mathbf{Z}) > 0$  para todo  $\mathbf{Z}$ . Portanto o modelo de riscos proporcionais de Cox é escrito como

$$\lambda(t|\mathbf{Z}) = \alpha_0(t) \exp(\boldsymbol{\beta}' \mathbf{Z}).$$

O modelo de riscos proporcionais é frequentemente usado para estimar o efeito de uma ou mais covariáveis sobre a distribuição do tempo de falha  $T$ .

## Capítulo 2

# Modelos de riscos com heterogeneidade não observada

Até o momento, considerou-se que todas as covariáveis incluídas no modelo de risco eram observáveis e que a população sob estudo era homogênea. Mas ignorar a heterogeneidade entre os indivíduos de uma população, pode levar a conclusões errôneas à respeito da população amostrada.

Em um experimento de sobrevivência onde indivíduos dentro de um certo subconjunto da população têm uma participação comum não observável, pode-se dizer que os indivíduos dentro desta subpopulação são associados. Por exemplo, em um estudo de fatores de risco para um evento de doenças cardiovasculares é razoável acreditar que entre irmãos que partilham de um mesmo código genético ou de uma mesma condição ambiental (dieta alimentar, e outras características de estilo de vida), podem ter tempos de eventos mais semelhantes do que outros grupos ( Klein, 1992).

Se a dependência entre os eventos é ignorada, as estimativas dos efeitos das covariáveis são suspeitas. Portanto, para incorporar estas associações entre os indivíduos de uma subpopulação dentro da análise dos riscos é introduzido, no modelo, um efeito aleatório comum chamado de fragilidade.

O modelo de fragilidade permite incluir um efeito aleatório na função risco, que representará a heterogeneidade não observada de cada indivíduo. Sendo assim, cada indivíduo da população estará sujeito a uma determinada função risco, que pode ser diferente de uma pessoa para outra.

## 2.1 Estimação de Máxima Verossimilhança

Nielsen *et al.* (1992) propuseram um método para a estimação de parâmetros em modelos de fragilidade, assumindo que o efeito aleatório segue uma distribuição gama. A estimação de máxima verossimilhança dos parâmetros fixados e dos efeitos aleatórios foram obtidos usando o algoritmo EM.

Usando a notação de Nielsen *et al.* (1992) denota-se  $N_i$  como sendo um processo de contagem<sup>1</sup> total para o grupo  $i$ , assim cada  $N_i$  pode ter mais do que uma omissão. Os valores de  $\Delta N_i(t)$  serão dados pela expressão  $\Delta N_i(t) = N_i(t) - N_i(t-)$ , que irá indicar o tempo exato de falha, ou seja, quando  $\Delta N_i(t) = 1$  sabe-se que o indivíduo (i) falhou no tempo  $t$ . Os membros do grupo  $i$  partilham da mesma variável de fragilidade  $Z_i$ . A intensidade de  $N_i$  é denotada por  $\lambda_i$ . Considere  $Y_i$  um processo preditivo observável e não-negativo e finalmente denote por  $\alpha(t)$  a função risco básica não conhecida e  $Z_i$  são variáveis independentes e identicamente distribuídas segundo a distribuição gama, com parâmetro de escala  $\eta$  desconhecido e o parâmetro de forma  $\nu$  desconhecido. Obviamente, esta não é a única possibilidade, existem outras famílias de distribuições para  $Z_i$ , que podem ser utilizadas e serão estudadas posteriormente.

Neste trabalho são considerados tanto os modelos paramétricos como os semi-paramétricos. Nos modelos paramétricos o risco básico  $\alpha(t)$  dependerá apenas de um parâmetro desconhecido, que será estimado. No caso de modelos semi-paramétricos  $\alpha(t)$

---

<sup>1</sup>Observações ou dados experimentais reunidos sobre um período de tempo são frequentemente modelados com processos estocásticos. Em particular, dados de contagem do número de diferentes tipos de eventos, que ocorrem ao longo do tempo, podem ser modelados com processos de contagem (Fleming e Harrington, 1991).

será completamente desconhecida.

O modelo inicial  $N = (N_i : i = 1, 2, \dots, n)$  é um processo de contagem multivariado com intensidade  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$  dada por

$$\lambda_i(t) = Z_i Y_i(t) \alpha(t). \quad (2.1)$$

Quando os parâmetros da distribuição gama são desconhecidos surge o problema de identificabilidade (Hoem, 1990). Entretanto fazendo-se a restrição  $\nu = \eta$  o problema se restringe somente a um único parâmetro  $\eta$ . Isto significa que a distribuição de  $Z_i$  com esta restrição terá média  $\nu/\eta = 1$  e a variância igual a  $\nu/\eta^2 = \eta^{-1} = \theta$ .

O interesse é estimar o parâmetro  $\theta$  e a função risco padrão acumulada  $A(t) = \int_0^t \alpha(u) du$ , baseado nas observações de  $(N, Y)$  unicamente, e via estimação de máxima verossimilhança. Uma maneira de determinar a função de verossimilhança é escrevê-la como produto da função densidade de  $(N, Y)$ . Assim, para obtermos a função de densidade conjunta de  $(N, Y)$ , primeiro escrevemos a função densidade de  $(\mathbf{N}, \mathbf{Y})$  dado  $\mathbf{Z}$  como

$$f_{\mathbf{N}, \mathbf{Y} | \mathbf{Z}}(\mathbf{n}, \mathbf{y} | \mathbf{z}) = \prod_t (z_i Y_i(t) \alpha(t))^{\Delta N_i(t)} \exp \left[ -z_i \int_0^T Y_i(u) dA(u) \right], \quad (2.2)$$

multiplicando-se a função densidade de  $N, Y | Z$ , pela densidade de  $Z$  dada por

$$f_Z(z) = \frac{\left(\frac{1}{\theta}\right)^{\frac{1}{\theta}} (z_i)^{\frac{1}{\theta}-1}}{\Gamma\left(\frac{1}{\theta}\right)} \exp \left[ -\frac{1}{\theta} z_i \right], \quad (2.3)$$

obtêm-se a função de densidade completa

$$f_{N, Y, Z}(n, y, z) = \prod_t (z_i Y_i(t) \alpha(t))^{\Delta N_i(t)} \exp \left[ -z_i \int_0^T Y_i(u) dA(u) \right] \frac{\left(\frac{1}{\theta}\right)^{\frac{1}{\theta}} (z_i)^{\frac{1}{\theta}-1}}{\Gamma\left(\frac{1}{\theta}\right)} \exp \left[ -\frac{1}{\theta} z_i \right]. \quad (2.4)$$

Integrando sobre  $Z$  a função de densidade completa escrita acima, obtemos a função de densidade conjunta de  $(N, Y)$

$$f_{N, Y}(n, y) = \frac{\left(\frac{1}{\theta}\right)^{\frac{1}{\theta}}}{\Gamma\left(\frac{1}{\theta}\right)} \prod_t (Y_i(t) \alpha(t))^{\Delta N_i(t)} \frac{\Gamma\left(\sum_t \Delta N_i(t) + \frac{1}{\theta}\right)}{\left[\frac{1}{\theta} + \int_0^T Y_i(u) dA(u)\right]^{\sum_t \Delta N_i(t) + \frac{1}{\theta}}}. \quad (2.5)$$

Dividindo a função de densidade completa pela função de densidade de  $(N, Y)$ , tem-se a função de densidade de  $Z$  dado  $(N, Y)$

$$f_{Z|N,Y}(z|n, y) = \frac{z^{(\sum_t \Delta N_i(t) + \frac{1}{\theta} - 1)}}{\Gamma(\sum_t \Delta N_i(t) + \frac{1}{\theta})} \left[ \frac{1}{\theta} + \int_0^\tau Y_i(u) dA(u) \right]^{\sum_t \Delta N_i(t) + \frac{1}{\theta}} \times \exp \left[ -z \left( \frac{1}{\theta} + \int_0^\tau Y_i(u) dA(u) \right) \right]. \quad (2.6)$$

Como pode-se verificar, a distribuição de  $Z$  dado  $(N, Y)$  é gama com média

$$\frac{\sum_t \Delta N_i(t) + \frac{1}{\theta}}{\frac{1}{\theta} + \int_0^\tau Y_i(u) dA(u)},$$

e variância

$$\frac{\sum_t \Delta N_i(t) + \frac{1}{\theta}}{\left[ \frac{1}{\theta} + \int_0^\tau Y_i(u) dA(u) \right]^2}.$$

Note-se que se  $\theta = 0$ , então o efeito aleatório  $Z$  é identificável e igual a 1, e não induz a dependência entre os membros do grupo  $i$ .

## 2.2 Estimação da função risco e do parâmetro de fragilidade, utilizando o algoritmo EM

### 2.2.1 Modelos Paramétricos

Nos modelos paramétricos, a função risco básica  $\alpha(t)$  dependerá apenas de um parâmetro  $\omega$ , desconhecido. Considerando-se que os parâmetros da distribuição da fragilidade são fixos e conhecidos, pode-se estimar a função risco de cada indivíduo por

$$\lambda_i(t) = Z_i Y_i(t) \alpha(t),$$

usando o algoritmo EM (Dempster, 1977). Este algoritmo permite utilizar a verossimilhança da função de densidade conjunta de  $(N, Y, Z)$ .

O primeiro passo **E** consiste em estimar os valores de  $Z_i$  utilizando a esperança condicional de  $Z_i$  dado  $N$  e  $Y$ .

passo **E** : 
$$\hat{Z}_i = \frac{\sum_t \Delta N_i(t) + \frac{1}{\theta}}{\frac{1}{\theta} + \int_0^t Y_i(u) dA(u)}.$$

O passo seguinte **M** calcula o estimador de máxima verossimilhança de  $\omega$ , baseado no modelo de intensidade.

passo **M** : 
$$\hat{\lambda}_i(t) = \hat{Z}_i Y_i(t) \alpha(t).$$

A verossimilhança da função de densidade conjunta de  $(N, Y, Z)$  será dada através da seguinte expressão

$$L_{N,Y,Z}(n, y, z) = \prod_{i=1}^n f_{N,Y,Z}(n, y, z)$$

onde  $f_{N,Y,Z}(n, y, z)$  é dada em (2.4), portanto, escrita como

$$L_{N,Y,Z}(n, y, z) = \prod_{i=1}^n \left\{ \prod_t (z_i Y_i(t) dA(t))^{\Delta N_i(t)} \exp \left[ -z_i \int_0^t Y_i(u) dA(u) \right] \frac{\left(\frac{1}{\theta}\right)^{\frac{1}{\theta}} (z_i)^{\frac{1}{\theta}-1}}{\Gamma\left(\frac{1}{\theta}\right)} \exp \left[ -\frac{1}{\theta} z_i \right] \right\}$$

Aplicando o logaritmo na equação acima, obtêm-se

$$\begin{aligned} \log L_{N,Y,Z}(n, y, z) &= \sum_{i=1}^n \left\{ \frac{1}{\theta} \log \left( \frac{1}{\theta} \right) + \left( \frac{1}{\theta} - 1 \right) \log [z_i] - \log \left[ \Gamma \left( \frac{1}{\theta} \right) \right] \right. \\ &\quad + \sum_t \Delta N_i(t) [\log(z_i) + \log(Y_i(t)) + \log(dA(t))] \\ &\quad \left. - z_i \left[ \frac{1}{\theta} + \int_0^t Y_i(u) dA(u) \right] \right\}. \end{aligned}$$

Substituindo na log-verossimilhança  $dA(t)$  por  $\alpha(t)$  escrita em termos de  $\omega$  e escrevendo os valores de  $Z_i$  como no passo E, a log-verossimilhança irá depender apenas dos parâmetros  $\omega$  e  $\theta$ .

Este algoritmo converge para um ponto crítico da  $\log L(\omega, \theta)$ , isto é, a solução de

$$\begin{aligned}\frac{\partial}{\partial \omega} \log L(\omega, \theta) &= 0, \\ \frac{\partial}{\partial \theta} \log L(\omega, \theta) &= 0.\end{aligned}$$

É necessário um método numérico iterativo para calcular  $\hat{\theta}$  e  $\hat{\omega}$  em cada passo do algoritmo EM. Para rodar o algoritmo é preciso uma estimativa inicial dos  $Z_i^{(0)}$ . Substituindo os valores iniciais dos  $Z_i^{(0)}$  na log-verossimilhança, é possível calcular os estimadores  $\hat{\theta}$  e  $\hat{\omega}$ .

Com os primeiros valores de  $\hat{\omega}^{(1)}$  e de  $\hat{\theta}^{(1)}$ , calculamos através do passo E os valores de  $\hat{Z}_i^{(1)}$ .

Assim, no passo M com os valores de  $\hat{Z}_i^{(1)}$  calcula-se os estimadores  $\hat{\omega}^{(2)}$  e  $\hat{\theta}^{(2)}$ . E assim continua até os valores estimados convergirem.

Com os estimadores  $\hat{Z}_i$  e a estimativa de  $\hat{\alpha}(t)$  baseada no estimador  $\hat{\omega}$ , é possível estimar a função risco de cada indivíduo

$$\hat{\lambda}_i(t) = \hat{Z}_i Y_i(t) \hat{\alpha}(t)$$

e a função risco acumulada individual

$$\hat{\Lambda}_i(t) = \int_0^t \hat{Z}_i Y_i(u) \hat{\alpha}(u) du.$$

## 2.2.2 Modelos Semi-Paramétricos

Na estimação de tais modelos,  $\alpha(t)$  é completamente desconhecida. Portanto precisa-se estimar  $\alpha(t)$  através do estimador da função risco padrão acumulada  $A(t) = \int_0^t \alpha(u) du$ .

Neste caso  $\hat{A}(t)$  é calculada pelo estimador de Nelson-Aalen, dado a seguir

**passo M:** 
$$\hat{A}(t) = \int_0^t \frac{d\tilde{N}(u)}{\sum_i \hat{Z}_i Y_i(u)}.$$

O passo E será o mesmo dado anteriormente. No começo considera-se inicialmente um valor inicial para os  $Z_i^{(0)}$ , pelo passo M é possível calcular o estimador  $\hat{A}(t)^{(1)}$ .

A partir dos valores de  $\hat{Z}_i^{(1)}$  e de  $\hat{A}(t)^{(1)}$  é calculado, por verossimilhança, o estimador  $\hat{\theta}^{(1)}$ .

O esquema será o mesmo feito anteriormente e irá parar quando os valores dos estimadores convergirem.

Da mesma maneira, com os estimadores  $\hat{Z}_i$  e  $\hat{A}(t)$ , pode-se estimar a função risco e a função risco acumulada de cada indivíduo, através da seguinte expressão

$$\begin{aligned}\hat{\Lambda}_i(t) &= \hat{Z}_i \int_0^t Y_i(u) \hat{\alpha}(u) du \\ &= \hat{Z}_i \hat{A}(\max \{t : Y_i(t) = 1\}).\end{aligned}\tag{2.7}$$

Em geral a estimação dos parâmetros em modelos de fragilidade, apresenta melhores resultados trabalhando com os indivíduos em grupos, por esta razão foram formados dois estratos ( $h = 1, 2$ ), onde cada grupo de dois indivíduos participavam da mesma variável de fragilidade  $Z_i$ .

Foi utilizado a verossimilhança da função densidade conjunta de  $(N, Y)$  e o algoritmo EM, para obter os estimadores de máxima verossimilhança das funções risco de cada estrato  $A_1$  e  $A_2$ , os quais foram inseridos dentro da função de verossimilhança para obter a estimativa de  $\theta$ .

$$L(\theta) = \prod_{i=1}^n \prod_{h=1}^2 \left\{ \frac{\left(\frac{1}{\theta}\right)^{\frac{1}{\theta}}}{\Gamma\left(\frac{1}{\theta}\right)} \prod_t (Y_{ih}(t) dA_h(t))^{\Delta N_{ih}(t)} \frac{\Gamma\left(N_{ih}(\tau) + \frac{1}{\theta}\right)}{\left[\frac{1}{\theta} + \int_0^\tau Y_{ih}(u) dA_h(u)\right]^{N_{ih}(\tau) + \frac{1}{\theta}}} \right\}.$$

Em Nielsen *et al.* (1992) foi realizado o mesmo procedimento, sendo que a estimação de  $\theta$  foi considerada irrestrita, ou seja, foi permitido que os valores estimados de  $\theta$  tivessem valores negativos, o que na realidade não pode acontecer, pois  $\theta$  representa a variância da distribuição gama escolhida para representar a fragilidade. Este estimador foi denotado por  $\hat{\theta}$ .

Neste trabalho utilizamos o estimador de máxima verossimilhança restrito ao espaço paramétrico  $[0, \infty)$ , isto é

$$\tilde{\theta} = \arg \max_{\{\theta \geq 0\}} L(\theta).$$

É fácil verificar que

$$\tilde{\theta} = \hat{\theta} I_{[\hat{\theta} \geq 0]}.$$

## 2.3 Convergência do estimador do parâmetro de fragilidade

Nielsen *et al.* (1992) mostrou que as estimativas de  $\theta$  quando o número de indivíduos aumenta, tem distribuição normal com média aproximadamente igual ao verdadeiro valor do parâmetro  $\theta$ .

Murphy (1994, 1995) utilizando a teoria assintótica demonstrou a convergência do parâmetro da fragilidade e a sua distribuição, obtendo-se o seguinte resultado em forma de teorema.

### Teorema 1

- i)  $\hat{\theta} \xrightarrow{P} \theta_0$  ;
- ii)  $\sqrt{n} (\hat{\theta} - \theta_0) \xrightarrow{D} N(0, \sigma^2)$ .

onde  $\sigma^2$  é calculado através da matriz de informação de Fisher, definida por

$$\sigma^2 = \left[ E \left( - \frac{\partial^2 L(\theta, A)}{\partial \theta^2} \Big|_{(\theta_0, A_0)} \right) \right]^{-1}.$$

Para se obter a segunda derivada parcial de  $L(\theta, A)$  com respeito a  $\theta$ , foi utilizada a seguinte definição equivalente para  $L(\theta, A)$

$$L(\theta, A) = \frac{1}{n} \sum_{i=1}^n \int_0^\tau \ln(1 + \theta N_i(u-)) dN_i(u) - (\theta^{-1} + N_i(\tau)) \ln \left( 1 + \theta \int_0^\tau Y_i(u) dA(u) \right) \\ + \int_0^\tau \ln(Y_i(u) \Delta A(u)) dN_i(u),$$

esta definição é a mesma utilizada por Murphy(1994).

Desenvolvendo a primeira derivada de  $L(\theta, A)$  com relação a  $\theta$  obtemos

$$\frac{\partial L(\theta, A)}{\partial \theta} = \frac{1}{n} \sum_{i=1}^n \int_0^\tau \frac{N_i(u-)}{1 + \theta N_i(u-)} dN_i(u) + \frac{1}{\theta^2} \ln \left( 1 + \theta \int_0^\tau Y_i(u) dA(u) \right) \\ - \frac{1}{\theta} \frac{\int_0^\tau Y_i(u) dA(u)}{1 + \theta \int_0^\tau Y_i(u) dA(u)} - N_i(\tau) \frac{\int_0^\tau Y_i(u) dA(u)}{1 + \theta \int_0^\tau Y_i(u) dA(u)}.$$

A partir da primeira derivada parcial de  $L(\theta, A)$  com relação a  $\theta$ , desenvolvemos a segunda derivada parcial de  $L(\theta, A)$ , dada por

$$-\frac{\partial^2 L(\theta, A)}{\partial \theta^2} \Big|_{(\theta_0, A_0)} = n^{-1} \sum_{i=1}^n \int_0^\tau \left( \frac{N_i(t-)}{1 + \theta_0 N_i(t-)} \right)^2 dN_i(t) - N_i(\tau) \left( \frac{\int_0^\tau Y_i dA_0}{1 + \theta_0 \int_0^\tau Y_i dA_0} \right)^2 + \\ 2\theta_0^{-3} \left[ \log \left( 1 + \theta_0 \int_0^\tau Y_i dA_0 \right) - \frac{\theta_0 \int_0^\tau Y_i dA_0}{1 + \theta_0 \int_0^\tau Y_i dA_0} - \frac{1}{2} \left( \frac{\theta_0 \int_0^\tau Y_i dA_0}{1 + \theta_0 \int_0^\tau Y_i dA_0} \right)^2 \right],$$

quando  $\theta_0 = 0$ , o último termo é definido pelo seu limite  $\frac{2}{3} (\int_0^\tau Y_i dA_0)^3$ .

O interesse neste caso será demonstrar que o estimador restrito  $\tilde{\theta} = \hat{\theta} I_{[\hat{\theta} \geq 0]}$ , também convergirá para  $\theta_0$ . Iremos demonstrar este resultado, utilizando o teorema 1 demonstrado por Murphy e o teorema 2 apresentado a seguir

**Teorema 2:** Se  $Y_n$  é uma sequência de variáveis aleatórias, tais que  $Y_n \xrightarrow{P} \theta_0$ ,  $\theta_0 \geq 0$  então

$$Y_n I_{[Y_n \geq 0]} \xrightarrow{P} \theta_0.$$

**Demonstração:** Queremos mostrar que

$$P(|Y_n I_{[Y_n \geq 0]} - \theta_0| > \epsilon) \longrightarrow 0, \quad n \rightarrow \infty.$$

Note que

$$\begin{aligned} P(|Y_n I_{[Y_n \geq 0]} - \theta_0| > \epsilon) &= P(|Y_n I_{[Y_n \geq 0]} - Y_n + Y_n - \theta_0| > \epsilon) \\ &\leq P(|Y_n I_{[Y_n \geq 0]} - Y_n| + |Y_n - \theta_0| > \epsilon) \\ &\leq P(|Y_n I_{[Y_n \geq 0]} - Y_n| > \frac{\epsilon}{2}) + P(|Y_n - \theta_0| > \frac{\epsilon}{2}) \\ &= P(|Y_n - \theta_0| > \frac{\epsilon}{2}) + P(|Y_n I_{[Y_n \geq 0]} - Y_n| > \frac{\epsilon}{2}, Y_n \geq 0) \\ &\quad + P(|Y_n I_{[Y_n \geq 0]} - Y_n| > \frac{\epsilon}{2}, Y_n < 0) \\ &= P(|Y_n - Y_n| > \frac{\epsilon}{2}, Y_n \geq 0) + P(|Y_n I_{[Y_n \geq 0]} - Y_n| > \frac{\epsilon}{2}, Y_n < 0) \\ &\quad + P(|Y_n - \theta_0| > \frac{\epsilon}{2}). \end{aligned}$$

Quando  $n \rightarrow \infty$ , temos

$$P(|Y_n - \theta_0| > \frac{\epsilon}{2}) \longrightarrow 0$$

$$\begin{aligned} P(|Y_n I_{[Y_n \geq 0]} - Y_n| > \frac{\epsilon}{2}, Y_n < 0) &= P(|-Y_n| > \frac{\epsilon}{2}, Y_n < 0) \\ &= P(Y_n < -\frac{\epsilon}{2}, Y_n < 0) \\ &= P(Y_n < -\frac{\epsilon}{2}) \longrightarrow 0, \end{aligned}$$

pois  $Y_n \xrightarrow{P} \theta_0$ , por hipótese.

Portanto, demonstramos que

$$Y_n I_{[Y_n \geq 0]} \xrightarrow{P} \theta_0.$$

Com base nos Teoremas 1 e 2, podemos escrever o seguinte corolário

**Corolário 1:**  $\tilde{\theta} = \hat{\theta}_{I_{[\hat{\theta} \geq 0]}} \xrightarrow{P} \theta_0.$

Prova: Pelo Teorema 1  $\hat{\theta} \xrightarrow{P} \theta_0$ , então pelo Teorema 2  $\tilde{\theta} \xrightarrow{P} \theta_0.$

Utilizando o resultado assintótico demonstrado por Murphy(1994) da convergência em distribuição de  $\hat{\theta}$  (Teorema 1) e o Teorema 3 escrito a seguir, também é possível determinar a função de distribuição de  $\tilde{\theta}$ .

**Teorema 3:** Se  $Y_n$  é uma sequência de variáveis aleatórias, tais que  $\sqrt{n} (Y_n - \theta_0) \xrightarrow{D} N(0, \sigma^2)$ , então

$$\sqrt{n} [Y_n I_{[Y_n \geq 0]} - \theta_0] \xrightarrow{D} G \quad n \rightarrow \infty$$

onde  $G$  é uma variável aleatória com função de distribuição dada por

$$F_G(g) = \begin{cases} 0, & g < 0, \\ \frac{1}{2} I_{[\theta_0=0]}, & g = 0, \\ \Phi\left(\frac{g}{\sigma}\right), & g > 0. \end{cases}$$

Demonstração: Sabemos que

$$Y_n I_{[Y_n \geq 0]} = \begin{cases} 0, & Y_n \leq 0, \\ Y_n, & Y_n > 0. \end{cases}$$

e queremos provar que  $\sqrt{n} (Y_n I_{[Y_n \geq 0]} - \theta_0) \xrightarrow{D} G$ . Para isso precisamos demonstrar que quando  $n \rightarrow \infty$  a função de distribuição de  $\sqrt{n} (Y_n I_{[Y_n \geq 0]} - \theta_0)$  converge para a função de distribuição de  $G$ . Para uma melhor compreensão, iremos dividir esta demonstração em dois casos.

**1° caso ( $\theta_0 = 0$ )**

Sabemos por hipótese que  $\sqrt{n} Y_n \xrightarrow{D} N(0, \sigma^2)$ , utilizando este resultado iremos calcular a função de distribuição de  $\sqrt{n} Y_n I_{[Y_n \geq 0]}$

$$P(\sqrt{n} Y_n I_{[Y_n \geq 0]} = 0) = P(\sqrt{n} Y_n \leq 0) \rightarrow \Phi(0) = \frac{1}{2}.$$

$$\begin{aligned} P(\sqrt{n} Y_n I_{[Y_n \geq 0]} \leq g) &= P(\sqrt{n} Y_n I_{[Y_n \geq 0]} = 0) + P(0 < \sqrt{n} Y_n I_{[Y_n \geq 0]} \leq g) \\ &= P(\sqrt{n} Y_n \leq g) \rightarrow \Phi\left(\frac{g}{\sigma}\right). \end{aligned}$$

**2° caso ( $\theta_0 > 0$ )**

Neste caso iremos provar que  $\sqrt{n} (Y_n I_{[Y_n \geq 0]} - \theta_0)$  é equivalente a  $\sqrt{n} (Y_n - \theta_0)$  quando  $n \rightarrow \infty$

$$\begin{aligned} \sqrt{n} [Y_n I_{[Y_n \geq 0]} - \theta_0] &= \sqrt{n} [Y_n I_{[Y_n \geq 0]} + Y_n - Y_n - \theta_0] \\ &= \sqrt{n} [Y_n I_{[Y_n \geq 0]} - Y_n] + \sqrt{n} [Y_n - \theta_0] \\ &= \sqrt{n} Y_n [I_{[Y_n \geq 0]} - 1] + \sqrt{n} [Y_n - \theta_0] \end{aligned}$$

Note que

$$\begin{aligned} P\left(\sqrt{n}\left|I_{[Y_n \geq 0]} - 1\right| > \epsilon\right) &= P\left(\left|I_{[Y_n \geq 0]} - 1\right| > \frac{\epsilon}{\sqrt{n}}\right) \\ &= P\left(I_{[Y_n \geq 0]} = 0\right) \longrightarrow 0 \quad n \longrightarrow \infty \end{aligned}$$

pois

$$Y_n \xrightarrow{P} \theta_0 > 0.$$

Portanto demonstramos que  $\sqrt{n}\left[I_{[Y_n \geq 0]} - 1\right] \xrightarrow{P} 0$ , e como sabemos por hipótese que  $\sqrt{n}(Y_n - \theta_0) \xrightarrow{\mathcal{D}} N(0, \sigma^2)$  então podemos afirmar que  $\sqrt{n}(Y_n I_{[Y_n \geq 0]} - \theta_0)$  é equivalente a  $\sqrt{n}(Y_n - \theta_0)$  quando  $n \longrightarrow \infty$ .

Com base neste resultado, podemos escrever o seguinte corolário.

**Corolário 2:**  $\sqrt{n}[\tilde{\theta} - \theta_0] \xrightarrow{\mathcal{D}} G$

Prova: Pelo Teorema 1  $\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{\mathcal{D}} N(0, \sigma^2)$ , portanto pelo Teorema 3

$$\sqrt{n}(\hat{\theta} I_{[\hat{\theta} \geq 0]} - \theta_0) \xrightarrow{\mathcal{D}} G.$$

## 2.4 Cálculo da variância assintótica $\sigma^2$

No caso que  $\theta_0 = 0$ , todos os  $Z_i$  serão iguais a um. Portanto não existe heterogeneidade não observada. Sendo assim  $\sigma^2$  será determinado através da esperança

$$\begin{aligned}
E \left\{ n^{-1} \sum_{i=1}^n \sum_{h=1}^2 \left[ \int_0^\tau (N_{ih}(t-))^2 dN_{ih}(t) - N_{ih}(\tau) \left( \int_0^\tau Y_{ih} dA_{0h} \right)^2 + \frac{2}{3} \left( \int_0^\tau Y_{ih} dA_{0h} \right)^3 \right] \right\} \\
= n^{-1} \sum_{i=1}^n \sum_{h=1}^2 \left\{ E \left[ \int_0^\tau (N_{ih}(t-))^2 dN_{ih}(t) \right] - E \left[ N_{ih}(\tau) \left( \int_0^\tau Y_{ih} dA_{0h} \right)^2 \right] \right. \\
\left. + \frac{2}{3} E \left[ \left( \int_0^\tau Y_{ih} dA_{0h} \right)^3 \right] \right\} \quad (2.8)
\end{aligned}$$

desenvolvendo cada termo em separado, obtemos

$$\begin{aligned}
E \left[ \left( \int_0^\tau Y_{ih} dA_{0h} \right)^3 \right] &= E \left[ \left( \int_0^{\max\{t: Y_{ih}(t)=1\}} dA_{0h} \right)^3 \right] \\
&= E \left[ \left( \int_0^{\max\{t: Y_{ih}(t)=1\}} \alpha_{0h}(t) dt \right)^3 \right] \quad \alpha_{0h}(t) = \lambda \\
&= E \left[ (\lambda \max\{t: Y_{ih}(t) = 1\})^3 \right], \quad (2.9)
\end{aligned}$$

$$\begin{aligned}
E \left[ N_{ih}(\tau) \left( \int_0^\tau Y_{ih} dA_{0h} \right)^2 \right] &= E \left[ \left( \int_0^\tau Y_{ih} dA_{0h} \right)^2 \right] \quad N_{ih}(\tau) = 1 \\
&= E \left[ \left( \int_0^{\max\{t: Y_{ih}(t)=1\}} \alpha_{0h}(t) dt \right)^2 \right] \\
&= E \left[ (\lambda \max\{t: Y_{ih}(t) = 1\})^2 \right], \quad (2.10)
\end{aligned}$$

$$E \left[ \int_0^\tau (N_{ih}(t-))^2 dN_{ih}(t) \right] = E \left[ \sum_{t>0}^{\tau} (N_{ih}(t-))^2 \Delta N_{ih}(t) \right] = 0, \quad (2.11)$$

Nesta última expressão,  $N_{ih}(t-) = 0$ , pois antes do tempo  $t$  ainda não ocorreu a falha, ou seja  $N_{ih}(t)$  é uma função escada, que possui somente um salto no tempo  $t$ , portanto no instante  $t-$  ela ainda assume o valor 0.

Logo neste caso a variância assintótica  $\sigma^2$  do parâmetro da fragilidade quando os dados são não censurados será dada pela seguinte expressão

$$\sigma^2 = \left\{ n^{-1} \sum_{i=1}^n \sum_{h=1}^2 \left[ \frac{2}{3} E [(\lambda \max \{t : Y_{ih}(t) = 1\})^3] - E [(\lambda \max \{t : Y_{ih}(t)\})^2] \right] \right\}^{-1}. \quad (2.12)$$

Quando existir a censura, haverá uma modificação na esperança (2.10), que será calculada da seguinte maneira

$$E \left[ N_{ih}(\tau) \left( \int_0^\tau Y_{ih} dA_{0h} \right)^2 \right] = E \left[ E \left[ N_{ih}(\tau) \left( \int_0^\tau Y_{ih} dA_{0h} \right)^2 \mid N_{ih}(\tau) = w \right] \right] \quad (2.13)$$

$$\begin{aligned} E \left[ N_{ih}(\tau) \left( \int_0^\tau Y_{ih} dA_{0h} \right)^2 \mid N_{ih}(\tau) \right] &= N_{ih}(\tau) E \left[ \left( \int_0^\tau Y_{ih} dA_{0h} \right)^2 \right] \\ &= N_{ih}(\tau) E [(\lambda \max \{t : Y_{ih}(t) = 1\})^2]. \end{aligned} \quad (2.14)$$

Como sabemos  $N_{ih}(\tau)$  é uma função escada que assume o valor 0 até um determinado ponto e depois tem um salto para o valor 1 se ocorrer a falha, se não ocorrer a falha o dado é censurado, e a função permanece com o valor 0, ou seja

$$N_{ih}(\tau) = \begin{cases} 0, & \text{se ocorreu a censura,} \\ 1, & \text{se ocorreu a falha.} \end{cases}$$

portanto, substituindo em (2.13)

$$\begin{aligned} E \left[ N_{ih}(\tau) E [(\lambda \max \{t : Y_{ih}(t) = 1\})^2] \right] &= E [N_{ih}(\tau)] E [(\lambda \max \{t : Y_{ih}(t) = 1\})^2] \\ &= E [(\lambda \max \{t : Y_{ih}(t) = 1\})^2] P[\text{ocorrer a falha}] \end{aligned}$$

logo neste caso, quando houver censura, a variância assintótica será definida como

$$\sigma^2 = \left\{ n^{-1} \sum_{i=1}^n \sum_{h=1}^2 \left[ \frac{2}{3} E \left[ (\lambda \max \{t : Y_{ih}(t) = 1\})^3 \right] - E \left[ (\lambda \max \{t : Y_{ih}(t) = 1\})^2 \right] P(\text{ocorrer a falha}) \right] \right\}^{-1} \quad (2.15)$$

Para os outros casos, quando existe heterogeneidade não observada, o cálculo da esperança é mais complicado, sendo mais simples utilizar a matriz de informação de Fisher observada, definida por

$$\begin{aligned} I(\hat{\theta}) = - \frac{\partial^2 L(\theta, A)}{\partial \theta^2} \Big|_{\theta = \hat{\theta}} &= n^{-1} \sum_{i=1}^n \sum_{h=1}^2 \left\{ \int_0^\tau \left( \frac{N_{ih}(t-)}{1 + \hat{\theta} N_{ih}(t-)} \right)^2 dN_{ih}(t) \right. \\ &\quad - N_{ih}(\tau) \left( \frac{\int_0^\tau Y_{ih} dA_{0h}}{1 + \hat{\theta} \int_0^\tau Y_{ih} dA_{0h}} \right)^2 + 2\hat{\theta}^{-3} \left[ \log \left( 1 + \hat{\theta} \int_0^\tau Y_{ih} dA_{0h} \right) \right. \\ &\quad \left. \left. - \frac{\hat{\theta} \int_0^\tau Y_{ih} dA_{0h}}{1 + \hat{\theta} \int_0^\tau Y_{ih} dA_{0h}} - \frac{1}{2} \left( \frac{\hat{\theta} \int_0^\tau Y_{ih} dA_{0h}}{1 + \hat{\theta} \int_0^\tau Y_{ih} dA_{0h}} \right)^2 \right] \right\}. \end{aligned}$$

Desenvolvendo as integrais a seguir, obtemos

$$\begin{aligned} \int_0^\tau Y_{ih} dA_{0h} &= \int_0^\tau Y_{ih}(t) \alpha_{0h}(t) dt = \int_0^{\max\{t: Y_{ih}(t)=1\}} \alpha_{0h}(t) dt \\ &= A_{0h}(\max\{t : Y_{ih}(t) = 1\}) \end{aligned} \quad (2.16)$$

$$\int_0^\tau \left( \frac{N_{ih}(t-)}{1 + \hat{\theta} N_{ih}(t-)} \right)^2 dN_{ih}(t) = \sum_{t>0} \left( \frac{N_{ih}(t-)}{1 + \hat{\theta} N_{ih}(t-)} \right)^2 \Delta N_{ih}(t) = 0. \quad (2.17)$$

Utilizando os resultados acima (2.16) e (2.17), a matriz de informação de Fisher observada para o caso dos dados não censurados, pode ser escrita como

$$\begin{aligned}
I(\hat{\theta}) = & n^{-1} \sum_{i=1}^n \sum_{h=1}^2 \left\{ - \left( \frac{A_{0h} (\max \{t : Y_{ih}(t) = 1\})}{1 + \hat{\theta} A_{0h} (\max \{t : Y_{ih}(t) = 1\})} \right)^2 \right. \\
& + 2(\hat{\theta})^{-3} \left[ \log \left( 1 + \hat{\theta} A_{0h} (\max \{t : Y_{ih}(t) = 1\}) \right) - \frac{\hat{\theta} A_{0h} (\max \{t : Y_{ih}(t) = 1\})}{1 + \hat{\theta} A_{0h} (\max \{t : Y_{ih}(t) = 1\})} \right. \\
& \left. \left. - \frac{1}{2} \left( \frac{\hat{\theta} A_{0h} (\max \{t : Y_{ih}(t) = 1\})}{1 + \hat{\theta} A_{0h} (\max \{t : Y_{ih}(t) = 1\})} \right)^2 \right] \right\}. \quad (2.18)
\end{aligned}$$

Quando  $\theta_0 = 0$  também podemos calcular a matriz de informação de Fisher observada, dada por

$$\begin{aligned}
I(\hat{\theta}) = & n^{-1} \sum_{i=1}^n \sum_{h=1}^2 \left\{ 1 - (A_{0h} (\max \{t : Y_{ih}(t) = 1\}))^2 \right. \\
& \left. + \frac{2}{3} (A_{0h} (\max \{t : Y_{ih}(t) = 1\}))^3 \right\}. \quad (2.19)
\end{aligned}$$

Neste caso a expressão acima pode ser usada tanto para os dados censurados como também para os não censurados.

Para os dados censurados onde o valor de  $\theta$  difere de zero, a matriz de informação de Fisher observada será definida por

$$\begin{aligned}
I(\hat{\theta}) = & n^{-1} \sum_{i=1}^n \sum_{h=1}^2 \left\{ -N_{ih}(\tau) \left( \frac{A_{0h} (\max \{t : Y_{ih}(t) = 1\})}{1 + \hat{\theta} A_{0h} (\max \{t : Y_{ih}(t) = 1\})} \right)^2 \right. \\
& + 2(\hat{\theta})^{-3} \left[ \log \left( 1 + \hat{\theta} A_{0h} (\max \{t : Y_{ih}(t) = 1\}) \right) - \frac{\hat{\theta} A_{0h} (\max \{t : Y_{ih}(t) = 1\})}{1 + \hat{\theta} A_{0h} (\max \{t : Y_{ih}(t) = 1\})} \right. \\
& \left. \left. - \frac{1}{2} \left( \frac{\hat{\theta} A_{0h} (\max \{t : Y_{ih}(t) = 1\})}{1 + \hat{\theta} A_{0h} (\max \{t : Y_{ih}(t) = 1\})} \right)^2 \right] \right\}. \quad (2.20)
\end{aligned}$$

A matriz de informação de Fisher observada, obedecendo a certas condições de regularidade, é um estimador consistente em probabilidade da matriz de informação de

Fisher.

Assim, obtidos os valores dos  $\sigma^2$  de cada amostra através da matriz de informação de Fisher, usaremos a sua média como estimativa de  $\sigma^2$ .

É importante lembrar que toda a demonstração feita para a convergência do parâmetro da fragilidade, quanto o cálculo da variância assintótica  $\sigma^2$  é válida apenas para a distribuição de fragilidade gama com média igual a 1 e variância  $\theta$ . Não foi desenvolvido o cálculo para outras distribuições.

# Capítulo 3

## Simulações

O objetivo deste capítulo é verificar se o estimador da variância da fragilidade  $\theta$  é consistente e assintoticamente normal, para outras distribuições de fragilidade diferentes da gama. Foi realizado um estudo de simulação com os dados gerados a partir das distribuições gama, lognormal e normal, para verificar o quanto o estimador da variância da fragilidade é sensível com relação à escolha da distribuição para a fragilidade. Estas distribuições também podem ser utilizadas como distribuições de fragilidade, e já foram estudadas por alguns autores tais como Heckman e Singer (1982), Hougaard (1984 e 1986). Nesta simulação trabalhamos apenas com os modelos semi-paramétricos.

Paralelamente, também foi feita uma comparação entre a simulação realizada nesta dissertação com a simulação realizada por Nielsen *et. al.* (1992).

A fragilidade de cada indivíduo  $Z_i$  foi gerada segundo as distribuições gama, lognormal e normal, todas com média igual a 1 e variância igual a  $\theta$ .

$$Z_i, i = 1, \dots, n \text{ iid}$$

Como existe a necessidade da fragilidade assumir só valores positivos, no caso da distribuição normal os valores negativos foram substituídos pelos seus valores em módulo,

conservando a média e a variância estipuladas anteriormente.

Foram gerados  $M$  conjuntos com  $n$  pares independentes  $(t_{i1}, t_{i2})$  de tempos de sobrevivência. Trabalhamos com grupos de dois indivíduos, onde ambos compartilhavam da mesma fragilidade, formando dessa maneira dois estratos. Os estratos ( $h$ ) foram gerados da seguinte maneira

$$Y_{ih}, i = 1, \dots, n \text{ iid } \exp(1) \quad h = 1, 2 \quad \text{sendo que}$$

$$T_{ih} = \frac{Y_{ih}}{Z_i}.$$

Os dados foram simulados de duas formas: primeiro foram considerados os dados não censurados e depois os censurados. No caso da censura foi gerada uma variável aleatória  $C_{ih}$  segundo a distribuição uniforme(0,8), representando o tempo de censura de cada indivíduo. Neste caso o tempo de sobrevivência foi considerado igual ao  $\min(T_{ih}, C_{ih})$ .

Como o interesse é determinar a distribuição dos tempos  $T$ , e sabemos como estes foram gerados, é possível calcular sua distribuição. Como as variáveis  $Y$  e  $Z$  são independentes, a função densidade conjunta de  $(Y, Z)$  é dada pela multiplicação de suas densidades

$$f_{Y,Z}(y, z) = f_Y(y)f_Z(z).$$

Como  $T = Y/Z$ , a função densidade de  $T$  será dada usando-se a seguinte definição

$$f_T(t) = \int_0^{\infty} |z| f_{Y,Z}(tz, z) dz.$$

No caso da distribuição da fragilidade ser gama

$$\begin{aligned} f_{Y,Z}(y, z) &= \exp(-y) \frac{\left(\frac{1}{\theta}\right)^{\frac{1}{\theta}}}{\Gamma\left(\frac{1}{\theta}\right)} z^{\frac{1}{\theta}-1} \exp\left(-\frac{z}{\theta}\right) \\ &= \frac{\left(\frac{1}{\theta}\right)^{\frac{1}{\theta}}}{\Gamma\left(\frac{1}{\theta}\right)} z^{\frac{1}{\theta}-1} \exp\left[-\left(y + \frac{z}{\theta}\right)\right]. \end{aligned}$$

Portanto

$$f_T(t) = \left(\frac{1}{1 + \theta t}\right)^{\frac{1}{\theta}+1}.$$

A partir da função densidade de  $T$ , pode-se obter a função distribuição de  $T$

$$F_T(t) = P(T \leq t) = 1 - \left(\frac{1}{1 + \theta t}\right)^{\frac{1}{\theta}},$$

e conseqüentemente é possível determinar a função de sobrevivência

$$S(t) = 1 - F_T(t) = \left(\frac{1}{1 + \theta t}\right)^{\frac{1}{\theta}}.$$

A função risco acumulada será dada por

$$A_0(t) = -\log(S(t)) = -\frac{1}{\theta} \log\left(\frac{1}{1 + \theta t}\right). \quad (3.1)$$

Como a função risco é derivada da função risco acumulada, também é possível determiná-la

$$\alpha_0(t) = \frac{d}{dt}A(t) = \frac{1}{1 + \theta t}.$$

Desta maneira é possível calcular os verdadeiros valores da função risco, no caso da distribuição da fragilidade ser gama, o que posteriormente será utilizado para calcular a variância do estimador de  $\theta$  e para uma futura comparação com a função risco estimada através da fragilidade.

No caso de outras distribuições para a fragilidade vale o mesmo procedimento, sendo diferente apenas o cálculo da função densidade de  $T$ , que irá diferenciar-se de uma distribuição para outra.

### 3.1 Resultados das Simulações

As tabelas apresentadas a seguir mostram a média e o desvio padrão dos valores estimados do parâmetro  $\theta_0$ . Em cada caso os tamanhos de amostra ( $n$ ) utilizados foram 100, 200, 500 e 1000. O estimador  $\tilde{\theta}$  resulta das simulações realizadas neste trabalho. Para efeito de comparação foram acrescentados às Tabelas 1 e 2 os valores estimados de  $\theta_0$  por Nielsen *et al.* (1992), denotados por  $\hat{\theta}$ .

Em todos os casos foram simuladas 200 ( $M$ ) repetições do experimento, com exceção apenas dos estimadores  $\hat{\theta}$ , quando  $\theta_0 = 0$ , em cujo caso foi utilizada uma amostra de tamanho igual a 500.

As simulações foram realizadas utilizando-se S-plus em microcomputador Pentium 200, 32MB. No procedimento de maximização e interação foi utilizado o programa MATLAB for Windows.(em anexo é apresentado o programa utilizado na simulação)

### 3.1.1 Fragilidade Gama

Como era esperado segundo os Teoremas 1 e 2, ambos estimadores ( $\hat{\theta}$  e  $\tilde{\theta}$ ) se aproximam do verdadeiro valor de  $\theta_0$ , conforme o número de indivíduos aumenta, sendo que no caso dos dados censurados essa aproximação é um pouco melhor, apesar da variância também aumentar. (ver Tabelas 1 e 2) O estimador de Nielsen  $\hat{\theta}$  apresenta uma pequena vantagem em relação ao estimador  $\tilde{\theta}$  em média, quando  $\theta_0$  assume os valores 0.2 e 0.4, mas no caso de  $\theta_0 = 0$  a média do estimador de Nielsen apresenta valores negativos para  $\theta_0$ , contradizendo o fato de  $\theta_0$  representar a variância da fragilidade.

**Tabela 1: Média e desvio-padrão das estimações do parâmetro  $\theta_0$ , utilizando a fragilidade gama e dados não censurados**

$n$	$\theta_0$	média( $\tilde{\theta}$ )	dp( $\tilde{\theta}$ )	média( $\hat{\theta}$ )	dp( $\hat{\theta}$ )
100	0.0	0.0201	0.0471	-0.0524	0.0905
	0.2	0.1476	0.1188	0.1364	0.1368
	0.4	0.3133	0.1634	0.3058	0.1578
200	0.0	0.0170	0.0320	-0.0290	0.0715
	0.2	0.1639	0.0883	0.1702	0.0956
	0.4	0.3460	0.1181	0.3623	0.1147
500	0.0	0.0110	0.0201	-0.0170	0.0429
	0.2	0.1816	0.0592	0.1836	0.0570
	0.4	0.3713	0.0749	0.3811	0.0709
1000	0.0	0.0101	0.0160	-0.0047	0.0311
	0.2	0.1884	0.0468	0.1934	0.0421
	0.4	0.3756	0.0585	—	—

Nota: Os estimadores de  $\tilde{\theta}$  são resultados de simulações realizadas neste trabalho, os estimadores de  $\hat{\theta}$  são resultados de Nielsen *et. al.* (1992)

**Tabela 2: Média e desvio-padrão das estimações do parâmetro  $\theta_0$ , utilizando a fragilidade gama e dados censurados**

$n$	$\theta_0$	média( $\tilde{\theta}$ )	dp( $\tilde{\theta}$ )	média( $\hat{\theta}$ )	dp( $\hat{\theta}$ )
100	0.0	0.0276	0.0503	-0.0227	0.1133
	0.2	0.1614	0.1398	0.1711	0.1635
	0.4	0.3200	0.1739	0.3588	0.1822
200	0.0	0.0175	0.0366	-0.0136	0.0830
	0.2	0.1653	0.0917	0.1937	0.1040
	0.4	0.3519	0.1316	0.4031	0.1310
500	0.0	0.0119	0.0219	-0.0086	0.0497
	0.2	0.1855	0.0719	0.1955	0.0665
	0.4	0.3832	0.0877	0.3982	0.0831

Nota: Os estimadores de  $\tilde{\theta}$  são resultados de simulações realizadas neste trabalho, os estimadores de  $\hat{\theta}$  são resultados de Nielsen *et. al.* (1992)

As figuras 3.1 a 3.6 correspondem à distribuição empírica do estimador  $\tilde{\theta}$ . As figuras 3.7 a 3.10 mostram somente a distribuição da parte absolutamente contínua de  $\tilde{\theta}$ .

Note-se que  $\tilde{\theta}$  é uma variável aleatória mista pois  $P_{\theta_0}(\tilde{\theta} = 0) > 0$ , sendo que

$$\lim_{n \rightarrow \infty} P_{\theta_0}(\tilde{\theta} = 0) = 0.5$$

$$\text{e } \lim_{n \rightarrow \infty} P_{\theta_0}(\tilde{\theta} = 0) = 0 \quad \text{para } \theta_0 > 0.$$

Podemos verificar, tanto nos casos censurados como nos não censurados, que quando o número de indivíduos aumenta, a distribuição do estimador  $\tilde{\theta}$  se aproxima da distribuição normal, com média próxima do verdadeiro valor do parâmetro.

Figura 3.1: Distribuição empírica do estimador  $\tilde{\theta}$ , utilizando fragilidade gama, dados não censurados e  $\theta_0 = 0$

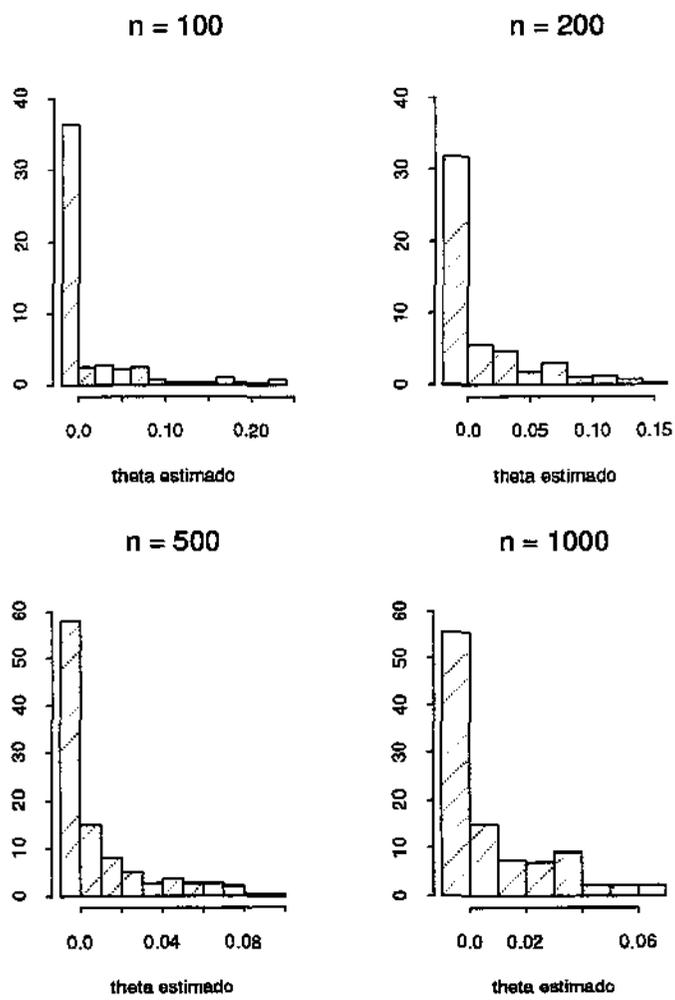


Figura 3.2: Distribuição empírica do estimador  $\tilde{\theta}$ , utilizando fragilidade gama, dados não censurados e  $\theta_0 = 0.2$

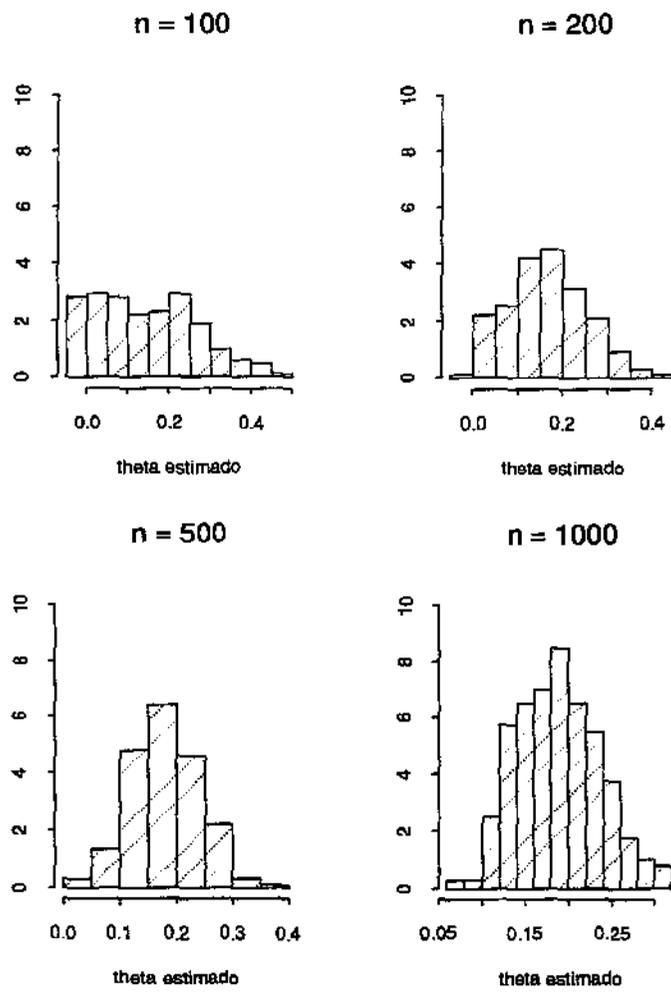


Figura 3.3: Distribuição empírica do estimador  $\tilde{\theta}$ , utilizando fragilidade gama, dados não censurados e  $\theta_0 = 0.4$

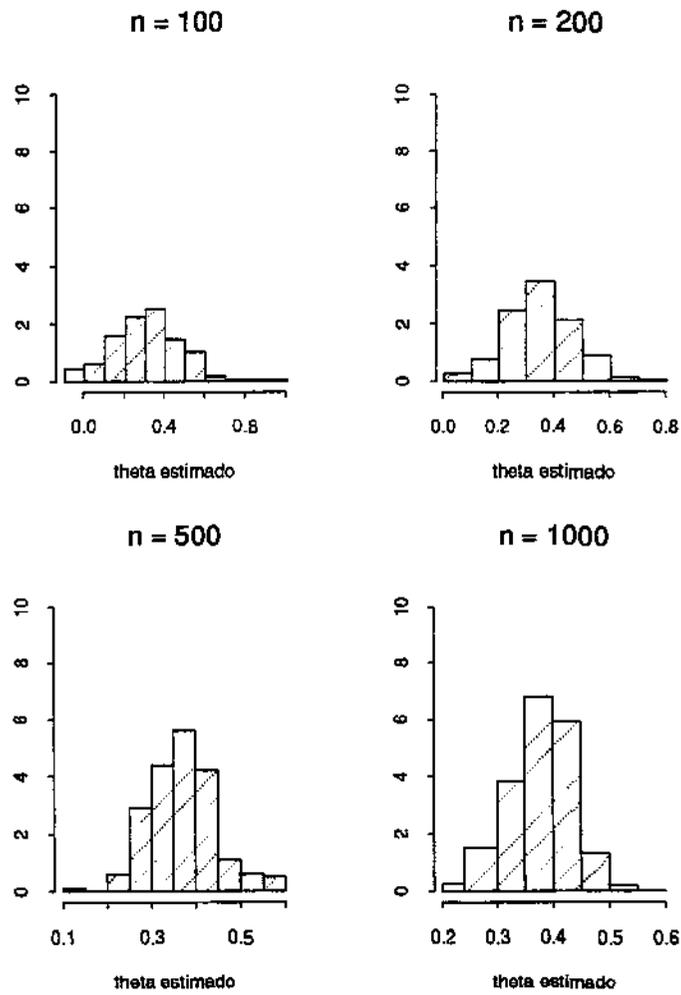


Figura 3.4: Distribuição empírica do estimador  $\tilde{\theta}$ , utilizando fragilidade gama, dados censurados e  $\theta_0 = 0$

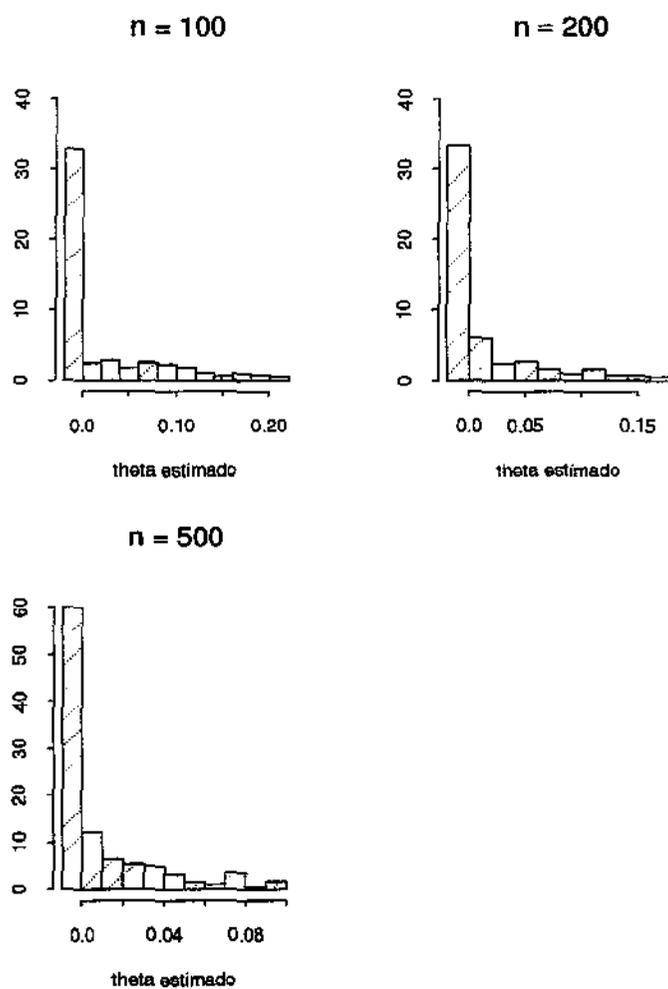


Figura 3.5: Distribuição empírica do estimador  $\tilde{\theta}$ , utilizando fragilidade gama, dados censurados e  $\theta_0 = 0.2$

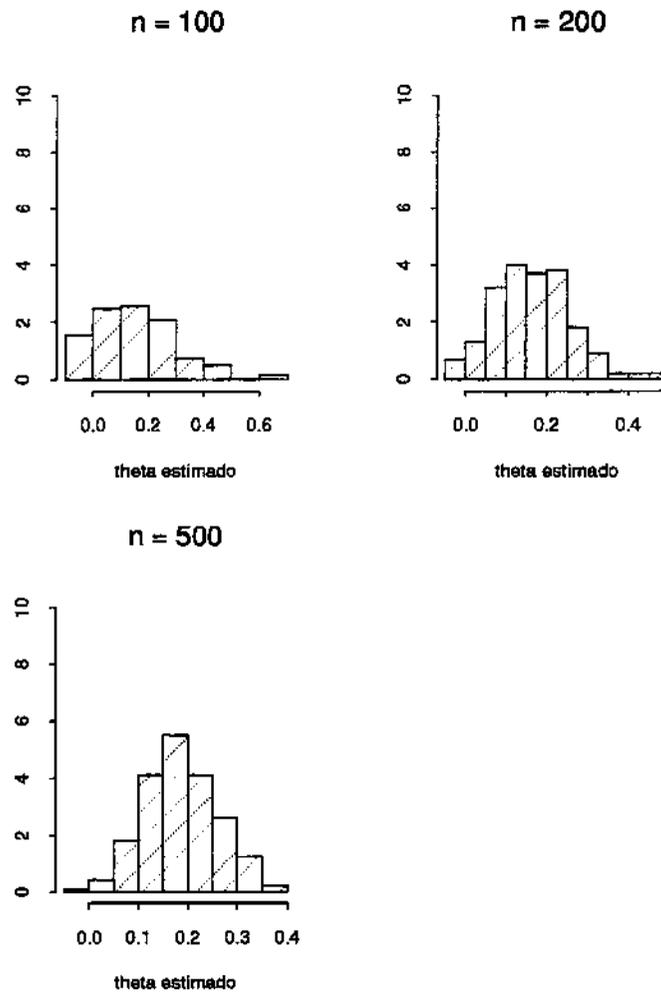
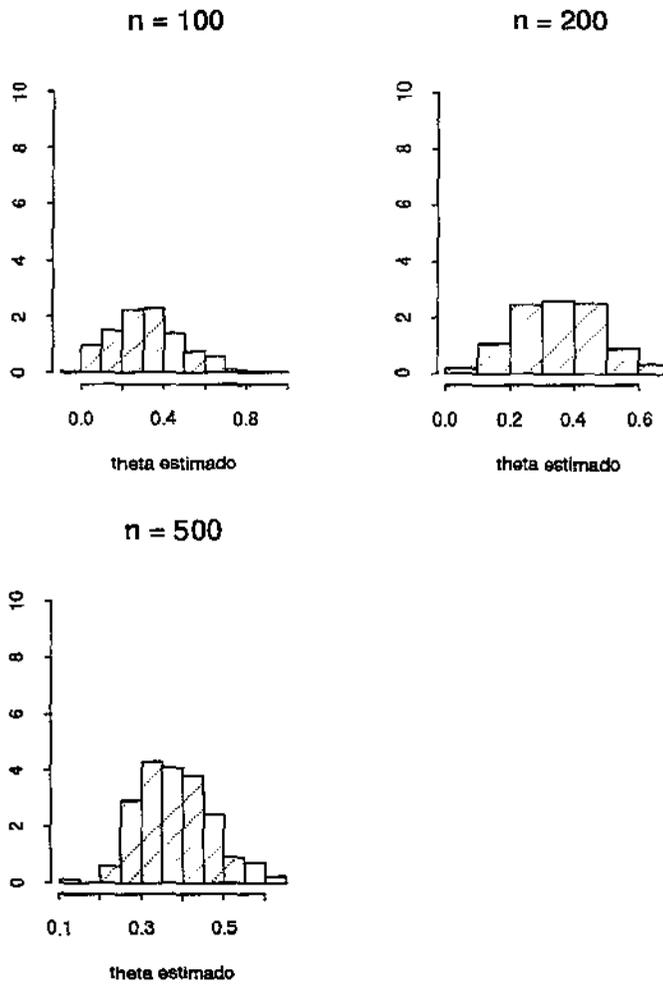


Figura 3.6: Distribuição empírica do estimador  $\tilde{\theta}$ , utilizando fragilidade gama, dados censurados e  $\theta_0 = 0.4$



Para obter-se a variância assintótica de  $\hat{\theta}$  utilizamos a variância assintótica de  $\hat{\theta}$  através da teoria apresentada no capítulo anterior e os dados simulados nesta seção. No caso de  $\theta_0 = 0$ , as variâncias assintóticas foram calculadas a partir das expressões (2.12) e (2.15) do número de informação de Fisher para os dados não censurados e censurados, respectivamente, e pela expressão (2.19) do número de informação observada utilizando-se os dados simulados. Como neste caso não existe heterogeneidade não observada isto é,  $Z_i = 1$ , a distribuição do tempo de sobrevivência é exponencial(1) e a função risco acumulada será definida como  $A_{0h}(t) = t$ .

Resolvendo as esperanças a seguir, obtemos

$$E[(\lambda \max\{t : Y_{ih}(t) = 1\})^3] = E[T^3] = 6. \quad \text{pois } T \sim \exp(1). \quad (3.2)$$

$$E[(\lambda \max\{t : Y_{ih}(t) = 1\})^2] = E[T^2] = 2. \quad (3.3)$$

A partir destes resultados, voltando a expressão (2.12), a variância assintótica de  $\hat{\theta}$  para os dados não censurados, a partir da matriz de informação de Fisher, quando  $\theta_0 = 0$  será

$$\sigma^2 = \left[ n^{-1} \sum_{i=1}^n \sum_{h=1}^2 \left[ \left( \frac{2}{3} \right) 6 - 2 \right] \right]^{-1} = 0.250.$$

Para os dados censurados, primeiro foi calculado a probabilidade de ocorrer a falha, ou seja  $P(T < C)$  onde  $T$  é o tempo de falha com distribuição exponencial(1) e  $C$  é o tempo de censura com distribuição uniforme(0,8). Esta probabilidade é igual a 0.8750, portanto com este resultado é possível calcular a variância assintótica de  $\hat{\theta}$  a partir da expressão (2.15), que será igual a

$$\sigma^2 = \left[ n^{-1} \sum_{i=1}^n \sum_{h=1}^2 \left[ \left( \frac{2}{3} \right) 6 - 2(0.8750) \right] \right]^{-1} = 0.2222.$$

Calculando a matriz de informação de Fisher observada dada pela expressão (2.19), através dos dados simulados, também foram obtidas as variâncias assintóticas de  $\hat{\theta}$ .

Para se obter a variância assintótica de  $\tilde{\theta}$ , é preciso relembrar que a distribuição assintótica de  $\hat{\theta}$  é  $N\left(\theta_0, \frac{\sigma^2}{n}\right)$  e  $\tilde{\theta} = \hat{\theta}I_{[\hat{\theta} \geq 0]}$ , portanto

$$\begin{aligned} Var(\tilde{\theta}) &= Var\left(\hat{\theta}I_{[\hat{\theta} \geq 0]}\right) \\ &= E\left[\hat{\theta}^2 I_{[\hat{\theta} \geq 0]}^2\right] - \left(E\left[\hat{\theta} I_{[\hat{\theta} \geq 0]}\right]\right)^2 \end{aligned}$$

sendo que

$$E\left[\hat{\theta} I_{[\hat{\theta} \geq 0]}\right] \cong \int_0^{\infty} \hat{\theta} \frac{1}{\sqrt{2\pi \frac{\sigma^2}{n}}} \exp\left(-\frac{(\hat{\theta} - \theta_0)^2}{2\frac{\sigma^2}{n}}\right) d\hat{\theta}.$$

Calculando-se estas esperanças, pode-se observar que as variâncias assintóticas de  $\tilde{\theta}$  e  $\hat{\theta}$  são equivalentes, quando o tamanho da amostra ( $n$ ) cresce. Portanto, iremos assumir neste trabalho que a variância de  $\tilde{\theta}$  é praticamente igual a variância de  $\hat{\theta}$ .

As Tabelas (3 a 6) a seguir, mostram os resultados obtidos das variâncias assintóticas de  $\tilde{\theta}$ , calculadas através da matriz de informação de Fisher  $I(\theta)$  e da matriz de informação de Fisher observada  $I(\hat{\theta})$ , para o caso de  $\theta_0 = 0$ .

**Tabela 3: Variância assintótica de  $\tilde{\theta}$ , utilizando  $I(\theta)$  e dados não censurados**

$n$	$Var(\tilde{\theta}) = \frac{\sigma^2}{n}$	$dp(\tilde{\theta})$
100	0.00250	0.0500
200	0.00125	0.0353
500	0.00050	0.0223
1000	0.00025	0.0158

**Tabela 4: Variância assintótica de  $\tilde{\theta}$ , utilizando  $I(\theta)$  e dados censurados**

$n$	$Var(\tilde{\theta}) = \frac{\sigma^2}{n}$	$dp(\tilde{\theta})$
100	0.002222	0.0471
200	0.001111	0.0333
500	0.000444	0.0210

**Tabela 5: Variância assintótica de  $\tilde{\theta}$ , utilizando  $I(\hat{\theta})$  e dados não censurados**

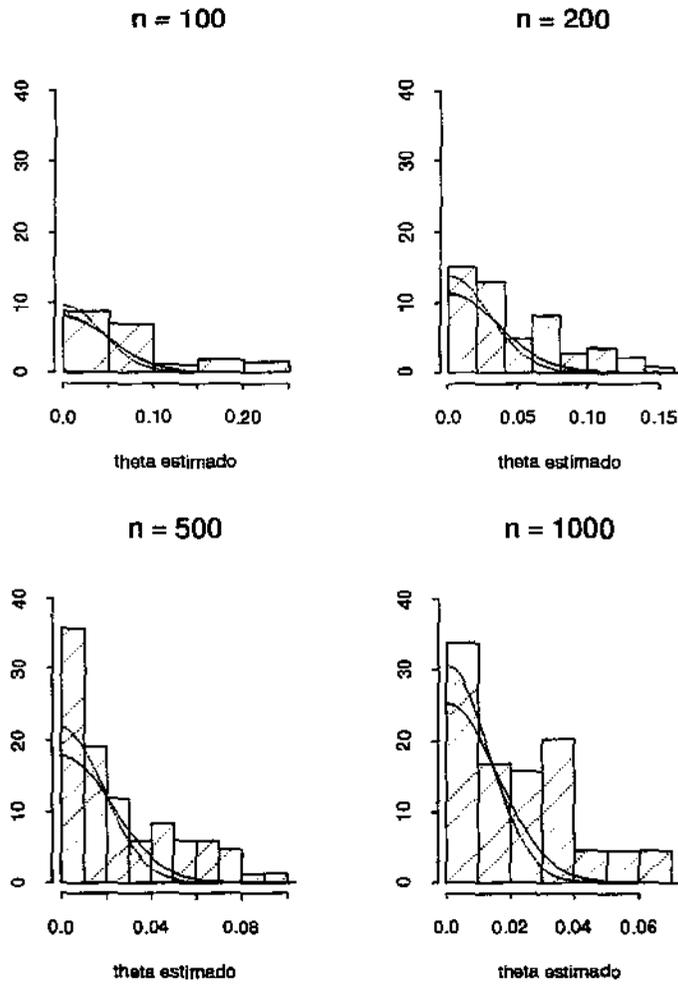
$n$	$I(\hat{\theta})$	$Var(\tilde{\theta}) = \frac{I(\hat{\theta})^{-1}}{n}$	$dp(\tilde{\theta})$
100	5.7321	0.001723157	0.0415
200	5.9614	0.000838729	0.0289
500	5.9231	0.000337661	0.0183
1000	5.9742	0.000169799	0.0130

**Tabela 6: Variância assintótica de  $\tilde{\theta}$ , utilizando  $I(\hat{\theta})$  e dados não censurados**

$n$	$I(\hat{\theta})$	$Var(\tilde{\theta}) = \frac{I(\hat{\theta})^{-1}}{n}$	$dp(\tilde{\theta})$
100	3.9772	0.002514331	0.0501
200	4.0133	0.001245857	0.0352
500	4.0563	0.000493060	0.0222

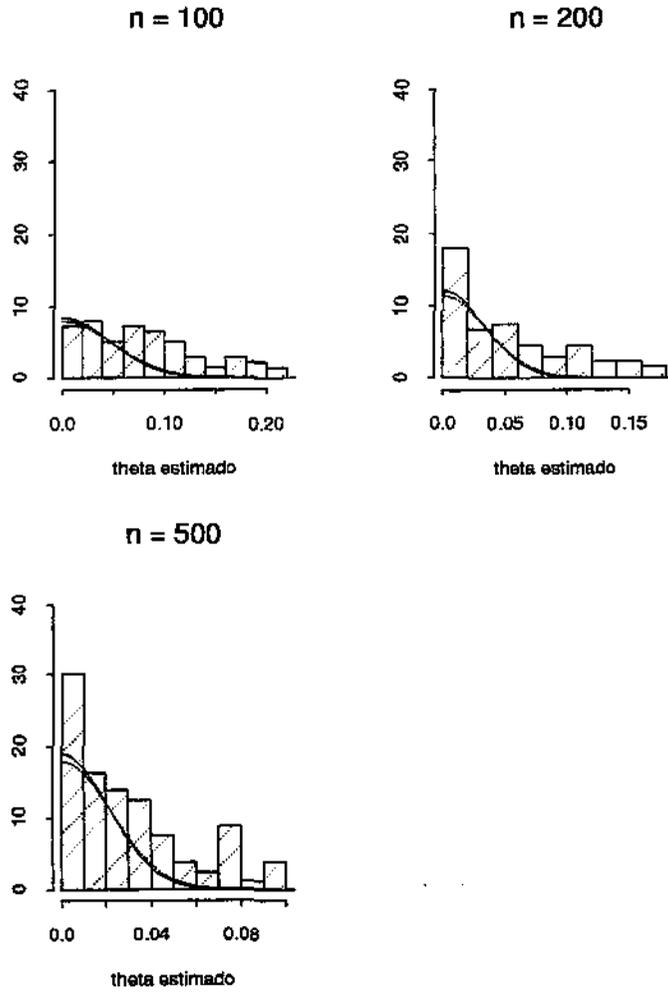
Com base nestes resultados, a seguir são mostrados os gráficos relacionados ao parâmetro  $\theta_0 = 0$ , Figuras 3.7 e 3.8, nos quais foi traçada a função densidade da normal correspondente ao número de indivíduos utilizado em cada amostra, onde a média é exatamente o valor do parâmetro  $\theta_0 = 0$  e a variância foi calculada a partir da matriz de informação de Fisher e através da matriz de informação de Fisher observada, descrita nas Tabelas (3 a 6).

Figura 3.7: Distribuição da parte absolutamente contínua de  $\tilde{\theta}$ , utilizando fragilidade Gama, dados não censurados e  $\theta_0 = 0$



Obs: linha completa corresponde a  $I(\theta)$  e linha pontilhada corresponde a  $I(\hat{\theta})$

Figura 3.8: Distribuição da parte absolutamente contínua de  $\tilde{\theta}$ , utilizando fragilidade Gama, dados censurados e  $\theta_0 = 0$



Obs: linha completa corresponde a  $I(\theta)$  e linha pontilhada corresponde a  $I(\hat{\theta})$

Nas figuras 3.7 e 3.8 as linhas completas correspondem à função densidade normal com variância calculada a partir da matriz de informação de Fisher; já as linhas pontilhadas correspondem à função densidade normal com variância calculada a partir da matriz de informação de Fisher observada. Podemos verificar a partir das figuras citadas acima, que a aproximação da função densidade da normal foi razoavelmente boa com relação às amostras obtidas em ambos os casos, censurados e não censurados.

Para os casos onde o valor de  $\theta_0$  difere de zero, foi utilizada apenas a matriz de informação de Fisher observada para o cálculo da variância assintótica, dada pelas expressões (2.18) no caso dos dados não censurados, e (2.20) no caso dos dados censurados. As duas tabelas a seguir mostram os resultados das variâncias para ambos os casos.

**Tabela 7: Variância assintótica de  $\tilde{\theta}$ , utilizando a  $I(\hat{\theta})$  e os dados não censurados**

$n$	$\theta_0$	$I(\hat{\theta})$	$Var(\tilde{\theta}) = \frac{I(\hat{\theta})^{-1}}{n}$	$dp(\tilde{\theta})$
100	0.2	1.0687	0.0093571	0.0967
	0.4	0.5594	0.0178762	0.1337
200	0.2	1.1530	0.0043365	0.0658
	0.4	0.3155	0.0158478	0.1258
500	0.2	0.8845	0.0022611	0.0475
	0.4	0.2175	0.0091932	0.0958
1000	0.2	0.8018	0.0012471	0.0353
	0.4	0.2077	0.0048134	0.0693

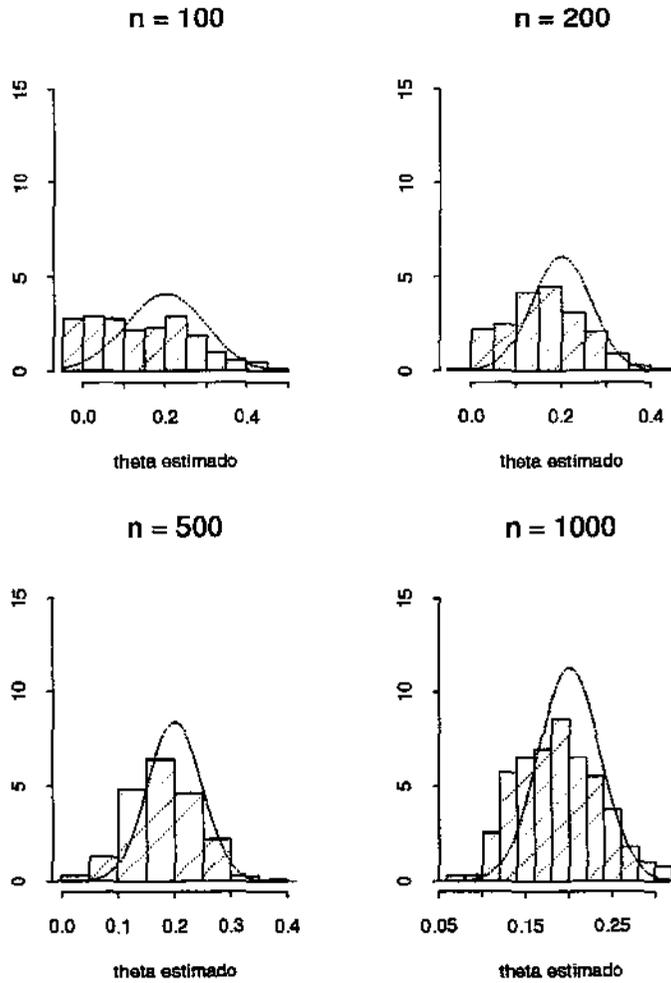
Tabela 8: Variância assintótica de  $\tilde{\theta}$ , utilizando a  $I(\hat{\theta})$  e os dados censurados

$n$	$\theta_0$	$I(\hat{\theta})$	$Var(\tilde{\theta}) = \frac{I(\hat{\theta})^{-1}}{n}$	$dp(\tilde{\theta})$
100	0.2	0.6413	0.0155933	0.1248
	0.4	0.3547	0.0281928	0.1679
200	0.2	0.6723	0.0074366	0.0862
	0.4	0.2714	0.0184229	0.1357
500	0.2	0.6195	0.0032284	0.0568
	0.4	0.2087	0.0095831	0.0978

Também foram feitos os gráficos relacionados às Tabelas 7 e 8 (Figuras 3.9 a 3.12), traçando a curva da função densidade da normal com média dada a partir do verdadeiro valor do parâmetro e variância obtida através da matriz de informação de Fisher observada  $I(\hat{\theta})$ . As Figuras 3.9 e 3.10 correspondem aos dados não censurados com parâmetros de fragilidade  $\theta$  iguais a 0.2 e 0.4, respectivamente. As Figuras 3.11 e 3.12 correspondem aos dados censurados.

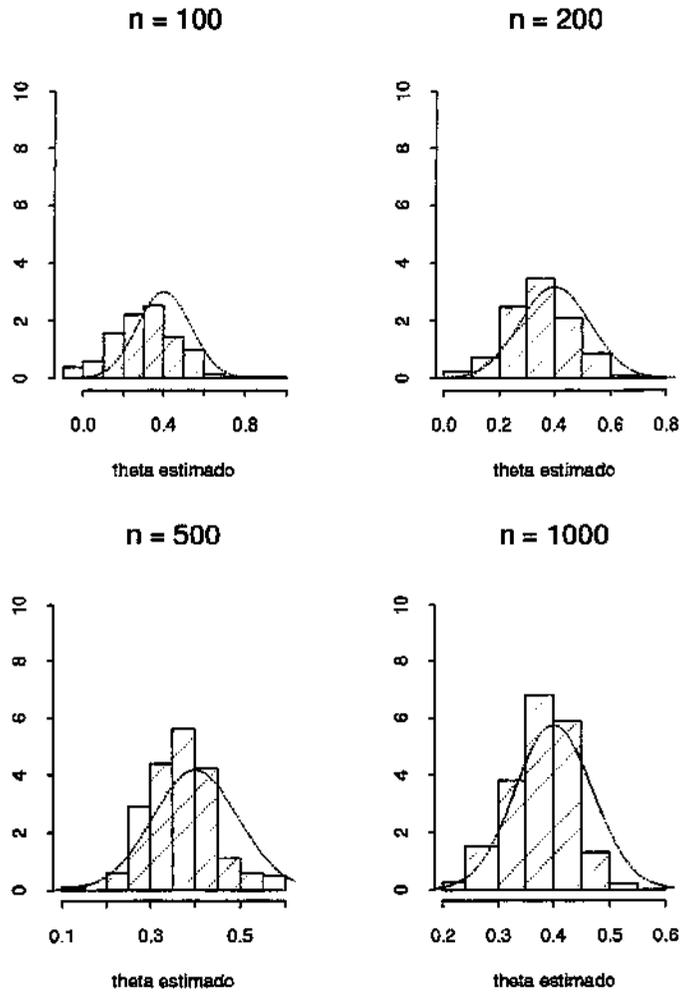
Observando as Figuras 3.9 a 3.12 é possível verificar que quanto maior é o número de indivíduos utilizados em cada amostra, melhor é a aproximação da função densidade da normal com relação à amostra e a estimação do parâmetro da fragilidade  $\tilde{\theta}$  fica mais perto do verdadeiro valor do parâmetro (ver Tabelas 1 e 2).

Figura 3.9: Distribuição da parte absolutamente contínua do estimador de  $\hat{\theta}$ , utilizando fragilidade gama, dados não censurados e  $\theta_0 = 0.2$



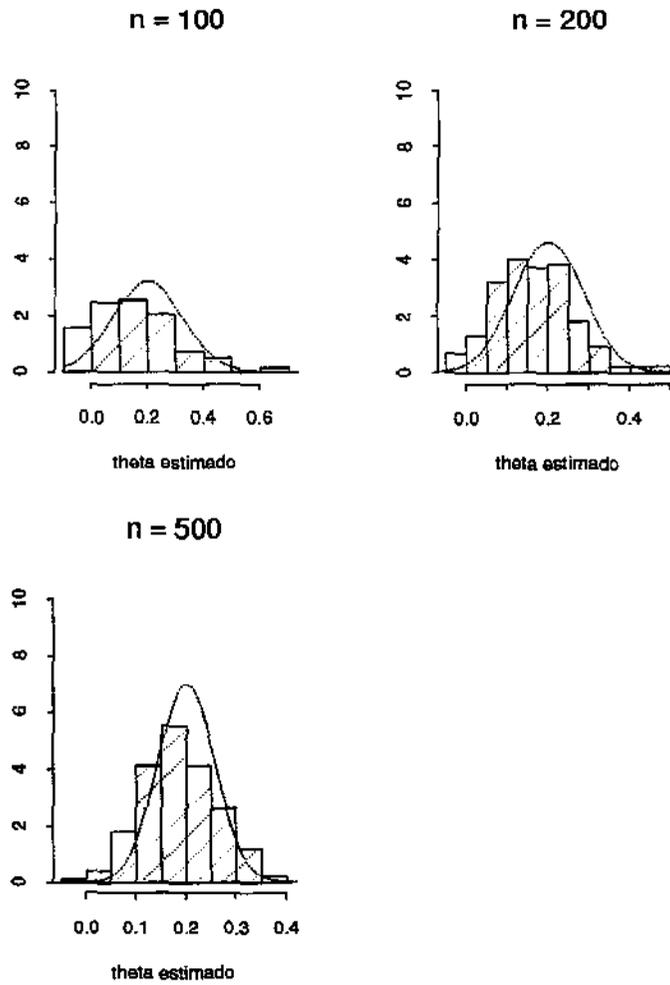
Obs: linha pontilhada corresponde a  $I(\hat{\theta})$

Figura 3.10: Distribuição da parte absolutamente contínua do estimador de  $\tilde{\theta}$ , utilizando fragilidade gama, dados não censurados e  $\theta_0 = 0.4$



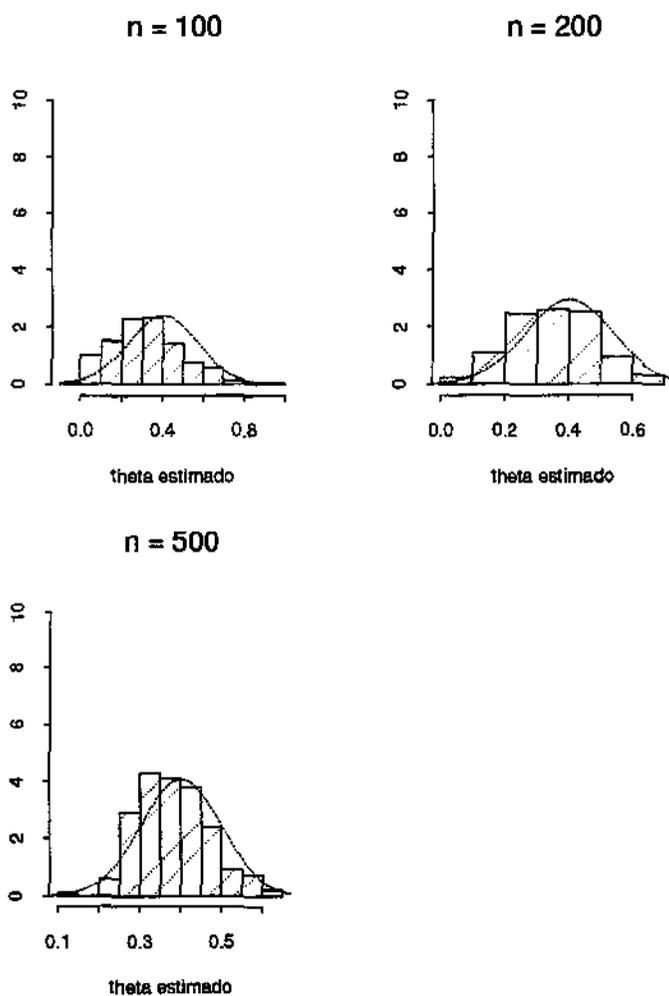
Obs: linha pontilhada corresponde a  $I(\hat{\theta})$

Figura 3.11: Distribuição da parte absolutamente contínua do estimador de  $\tilde{\theta}$ , utilizando fragilidade gama, dados censurados e  $\theta_0 = 0.2$

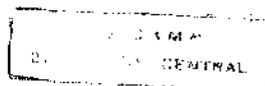


Obs: linha pontilhada corresponde a  $I(\hat{\theta})$

Figura 3.12: Distribuição da parte absolutamente contínua do estimador de  $\tilde{\theta}$ , utilizando fragilidade gama, dados censurados e  $\theta_0 = 0.4$



Obs: linha pontilhada corresponde a  $I(\hat{\theta})$



### 3.1.2 Fragilidade Lognormal

Os valores de  $\theta$  estimados ( $\tilde{\theta}$ ) utilizando a distribuição lognormal para a fragilidade estão resumidos nas Tabelas 9 e 10. Em ambos casos (não censurados e censurados) os estimadores de  $\theta$  subestimaram o parâmetro.

**Tabela 9: Média e desvio padrão das estimações do parâmetro de  $\theta_0$ , utilizando a fragilidade lognormal e dados não censurados**

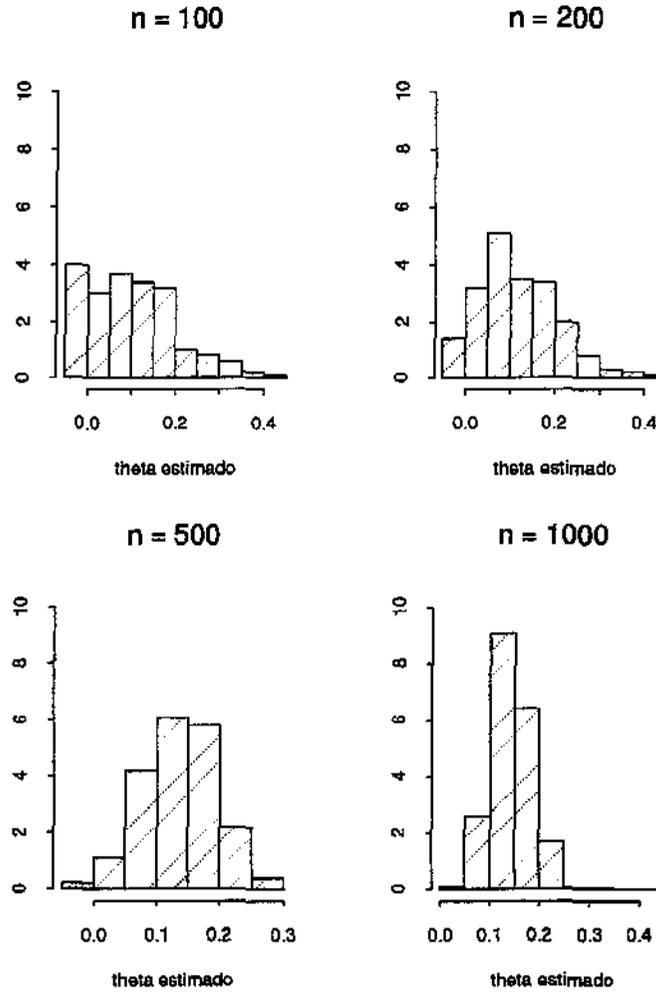
n	$\theta_0$	média( $\tilde{\theta}$ )	dp( $\tilde{\theta}$ )
100	0.2	0.1035	0.0938
	0.4	0.1862	0.1326
200	0.2	0.1195	0.0866
	0.4	0.2280	0.1063
500	0.2	0.1361	0.0574
	0.4	0.2407	0.0624
1000	0.2	0.1441	0.0374
	0.4	0.2481	0.0469

**Tabela 10: Média e desvio padrão das estimações do parâmetro de  $\theta_0$ , utilizando a fragilidade lognormal e dados censurados**

n	$\theta_0$	média( $\tilde{\theta}$ )	dp( $\tilde{\theta}$ )
100	0.2	0.1091	0.1074
	0.4	0.2136	0.1568
200	0.2	0.1292	0.0950
	0.4	0.2533	0.1225
500	0.2	0.1388	0.0644
	0.4	0.2573	0.0722

Um outro ponto de interesse, neste caso, foi verificar a distribuição dos estimadores de  $\theta$ .

Figura 3.13: Distribuição do estimador de  $\tilde{\theta}$ , utilizando fragilidade lognormal, dados não censurados e  $\theta_0 = 0.2$



A partir das figuras 3.13 a 3.16 verificamos, graficamente, que as estimativas de  $\tilde{\theta}$  se aproximam da distribuição normal conforme o número de indivíduos aumenta, mas a média das estimativas de  $\tilde{\theta}$  subestimam o verdadeiro valor do parâmetro (ver Tabelas 9 e 10). Portanto, quando assumimos a distribuição lognormal para a fragilidade, as estimativas de  $\theta$  se tornam viciadas com relação ao parâmetro de interesse.

Figura 3.14: Distribuição do estimador de  $\tilde{\theta}$ , utilizando fragilidade lognormal, dados não censurados e  $\theta_0 = 0.4$

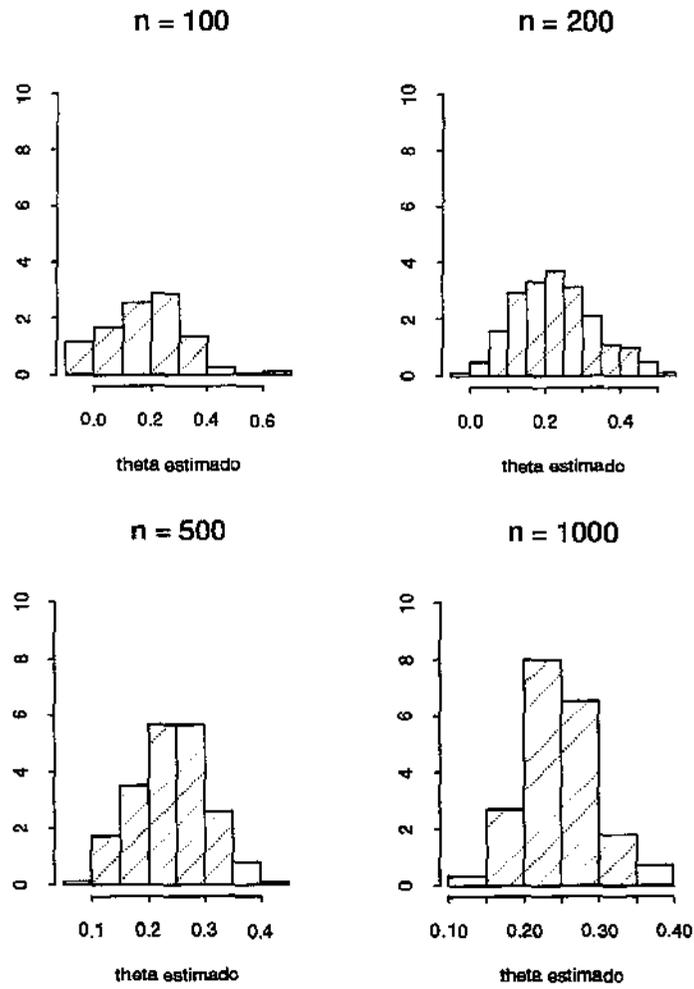


Figura 3.15: Distribuição do estimador de  $\tilde{\theta}$ , utilizando fragilidade lognormal, dados censurados e  $\theta_0 = 0.2$

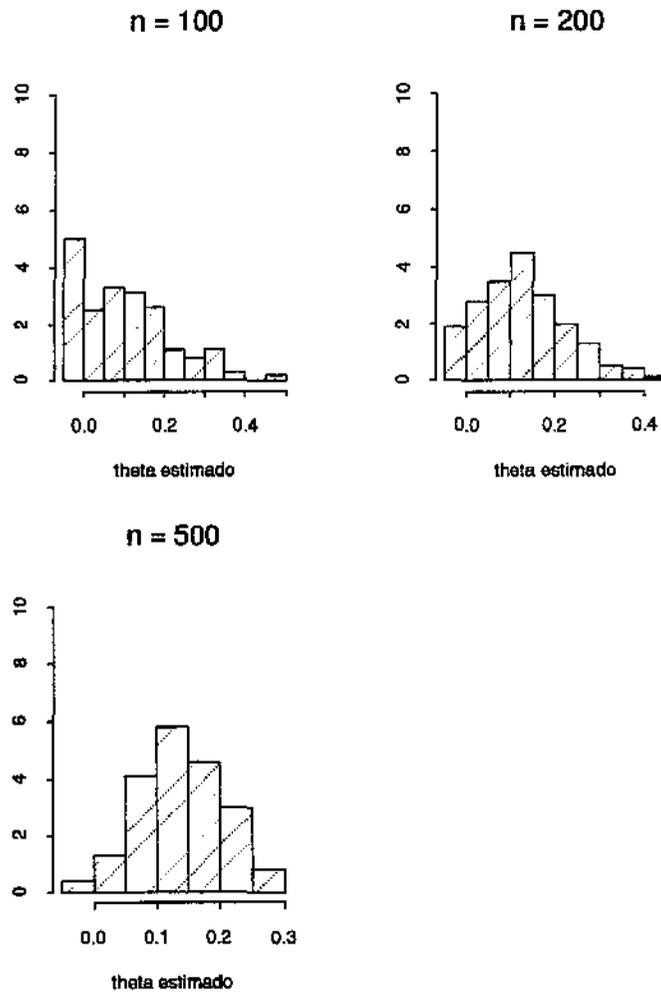
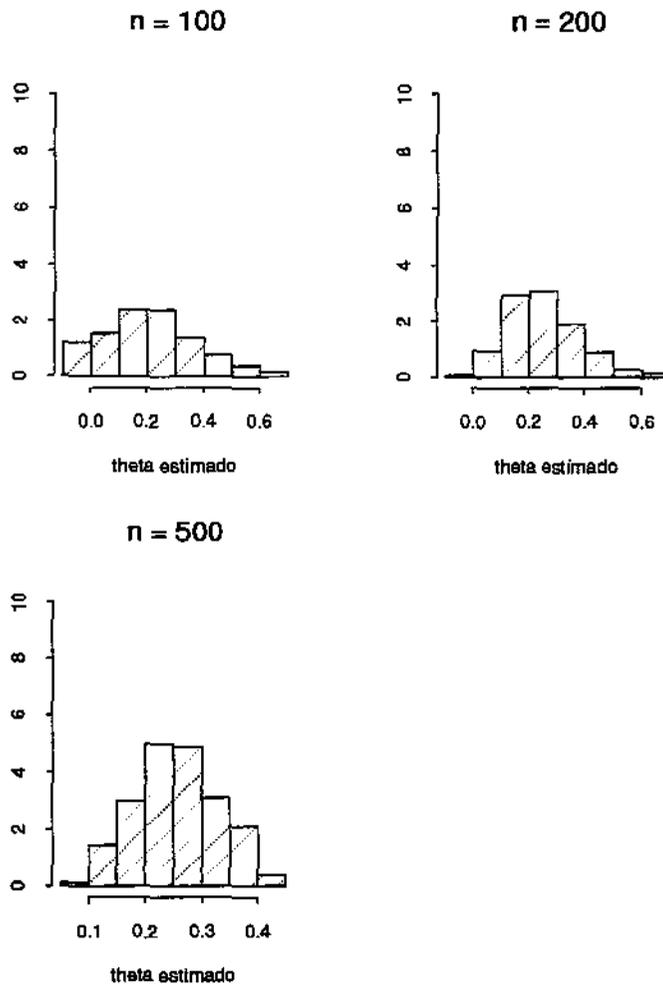


Figura 3.16: Distribuição do estimador de  $\tilde{\theta}$ , utilizando fragilidade lognormal, dados censurados e  $\theta_0 = 0.4$



### 3.1.3 Fragilidade Normal

Os estimadores de  $\theta$  ( $\tilde{\theta}$ ) utilizando a distribuição normal superestimaram o parâmetro, tanto no caso de dados não censurados, como no caso de dados censurados (ver Tabelas 11 e 12).

**Tabela 11: Média e desvio-padrão das estimações do parâmetro de  $\theta_0$ , utilizando a fragilidade normal e dados não censurados**

n	$\theta_0$	média( $\tilde{\theta}$ )	dp( $\tilde{\theta}$ )
100	0.2	0.3526	0.1844
	0.4	0.5395	0.1817
200	0.2	0.3255	0.1135
	0.4	0.5749	0.1340
500	0.2	0.3470	0.0680
	0.4	0.6076	0.0907
1000	0.2	0.3546	0.0515
	0.4	0.6275	0.0675

**Tabela 12: Média e desvio-padrão das estimações do parâmetro de  $\theta_0$ , utilizando a fragilidade normal e dados censurados**

n	$\theta_0$	média( $\tilde{\theta}$ )	dp( $\tilde{\theta}$ )
100	0.2	0.3028	0.1809
	0.4	0.5199	0.1977
200	0.2	0.2931	0.1221
	0.4	0.5454	0.1430
500	0.2	0.3072	0.0705
	0.4	0.5691	0.0976

Observando as figuras 3.17 a 3.20 vemos que, em ambos os casos, as estimativas de  $\theta$  se aproximam da distribuição normal quando o número de indivíduos aumenta, mas como no caso anterior, o estimador do parâmetro ( $\tilde{\theta}$ ) é viciado e superestima o verdadeiro parâmetro de  $\theta$ .

Portanto podemos afirmar que quando assumimos outra distribuição para a fragilidade, como no caso da distribuição lognormal e normal, podemos produzir estimativas viciadas do parâmetro de fragilidade, ou seja, os estimadores  $\tilde{\theta}$  são sensíveis quanto à escolha da distribuição da fragilidade, podendo desta maneira produzir resultados diferentes para cada tipo de distribuição assumida, comprometendo assim a eficiência do modelo final.

Figura 3.17: Distribuição do estimador de  $\tilde{\theta}$ , utilizando fragilidade normal, dados não censurados e  $\theta_0 = 0.2$

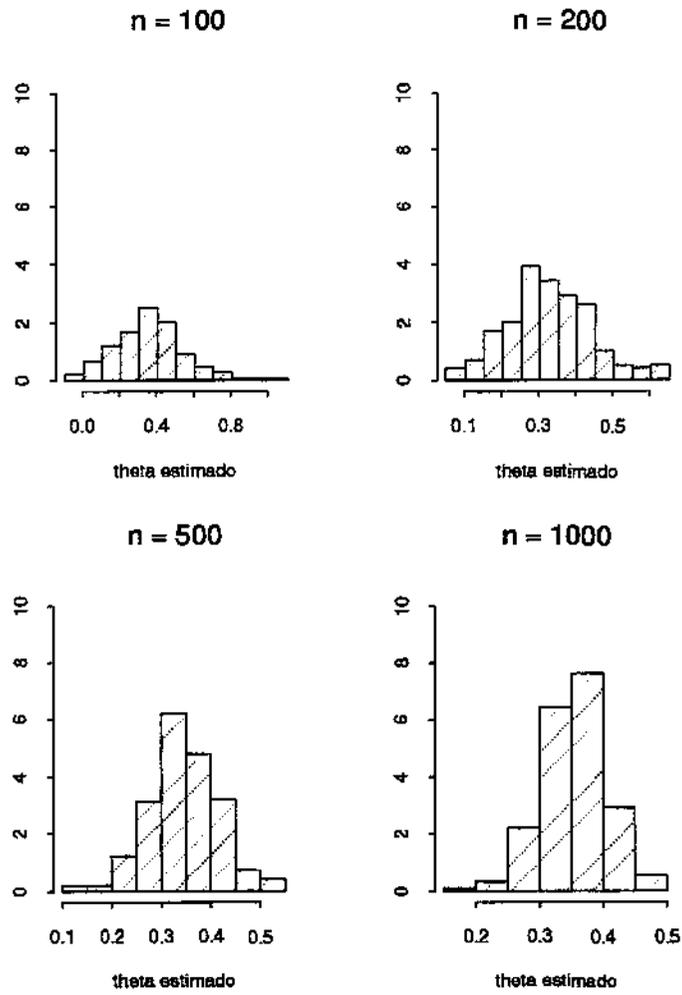


Figura 3.18: Distribuição do estimador de  $\tilde{\theta}$ , utilizando fragilidade normal, dados não censurados, e  $\theta_0 = 0.4$

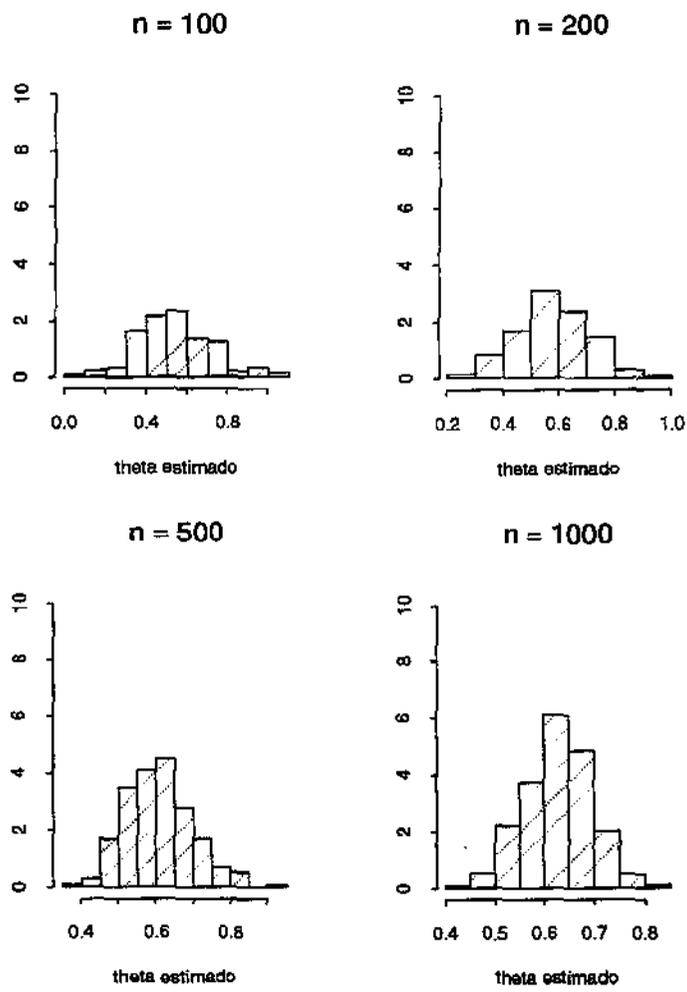


Figura 3.19: Distribuição do estimador de  $\tilde{\theta}$ , utilizando fragilidade normal, dados censurados e  $\theta_0 = 0.2$

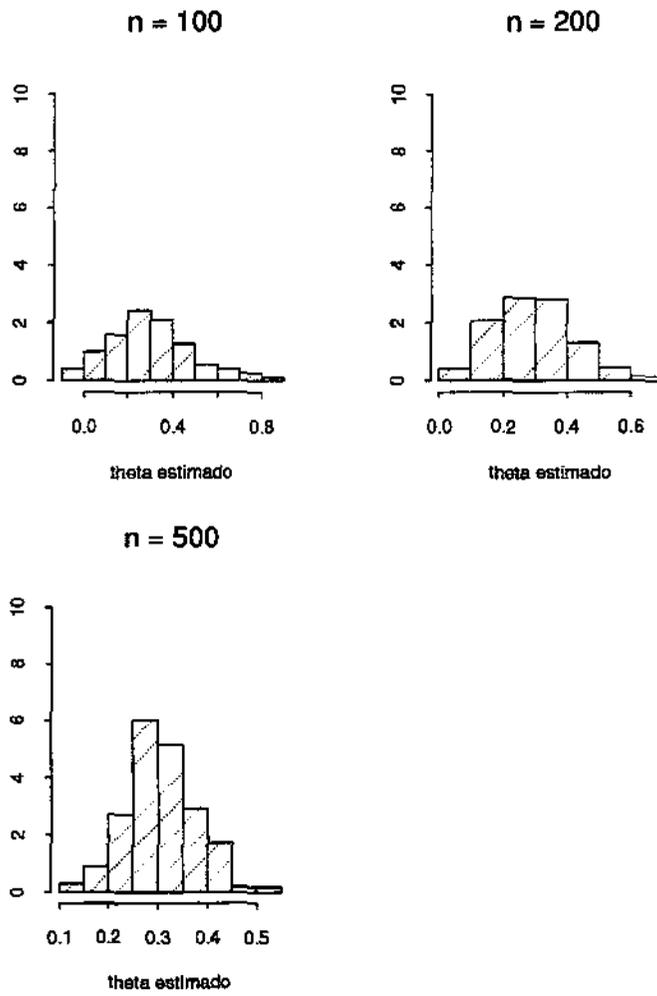
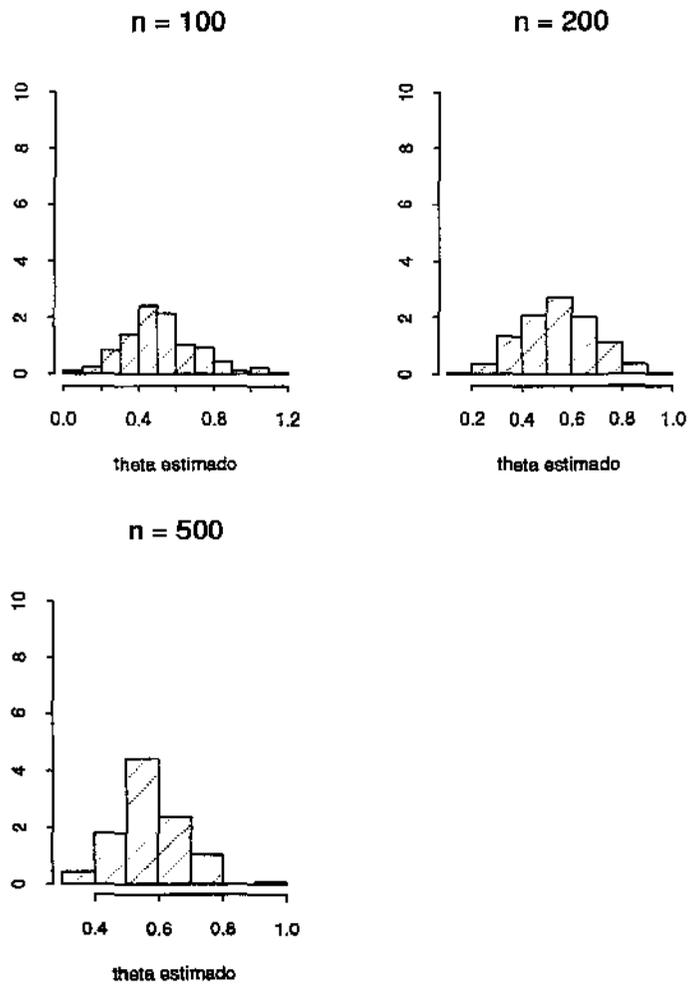


Figura 3.20: Distribuição do estimador de  $\tilde{\theta}$ , utilizando fragilidade normal, dados censurados e  $\theta_0 = 0.4$



## Capítulo 4

# Cálculo da Função Risco Acumulada estimada, através da fragilidade

Como vimos no capítulo anterior, a estimação do parâmetro da fragilidade pode se tornar viciada conforme a escolha da sua distribuição. O interesse é verificar se, mesmo com a estimativa viciada do parâmetro, podemos utilizar este resultado, sem maiores preocupações, para calcular a função risco acumulada estimada. Isto será feito comparando graficamente a função risco acumulada estimada com a verdadeira função risco acumulada.

Neste capítulo utilizamos os resultados obtidos através do algoritmo EM para estimar a função risco acumulada, pois para cada replicação do algoritmo, também era estimada iterativamente a função risco padrão acumulada através do estimador de Nelson-Aalen e os valores estimados dos  $Z_i$ .

Foram consideradas as mesmas distribuições utilizadas anteriormente para a fragilidade, a saber distribuição gama, lognormal e normal.

Uma expressão analítica para a função risco acumulada verdadeira no caso da fragilidade gama foi obtida através da expressão (3.1). Já nos casos da fragilidade lognormal e normal a expressão analítica para a função risco acumulada verdadeira não é possível de ser obtida e esta foi aproximada através de métodos numéricos.

No caso da fragilidade lognormal e normal a função densidade de  $T$  é dada por

$$f_T(t) = \int_0^\infty f_{T,Z}(t, z) dz$$

e não pode ser calculada explicitamente. Portanto, utilizando integração numérica obtemos uma aproximação para a função distribuição  $F_T(t)$  e portanto para a função risco acumulada  $A_T(t)$ .

No caso da distribuição lognormal a função densidade de  $T$  torna-se

$$f_T(t) = \int_0^\infty \frac{1}{\sqrt{2\pi \ln(\theta + 1)}} \exp \left\{ -tz - \frac{\ln \left[ \frac{z}{\exp \frac{-\ln(\theta+1)}{z}} \right]^2}{2 \ln(\theta + 1)} \right\} dz$$

e no caso da distribuição normal

$$f_T(t) = \int_0^\infty z \frac{1}{\sqrt{2\pi\theta}} \exp \left[ -tz + \frac{(-z-1)^2}{2\theta} \right] dz + \int_0^\infty z \frac{1}{\sqrt{2\pi\theta}} \exp \left[ -tz + \frac{(z-1)^2}{2\theta} \right] dz.$$

## 4.1 Comparação da função risco acumulada estimada com a verdadeira função risco acumulada

Primeiro será analisado o caso da fragilidade gama. Foi utilizado a amostra com o número de indivíduos igual a 1000 para o caso dos dados não censurados e para os dados censurados utilizamos a amostra com o número de indivíduos igual a 500.

As figuras 4.1 a 4.6 apresentam as funções de risco acumulada para os dados não censurados e censurados, com os parâmetros de fragilidade iguais a 0, 0.2 e 0.4 respectivamente.

As funções de risco acumulada se diferenciam quanto ao valor do estimador de  $\theta_0$  e com relação ao estrato utilizado na estimação (1, 2).

Observando as Figuras 4.1 a 4.6, podemos verificar que a estimação da função risco acumulada não foi considerada boa, exceto quando  $\theta_0 = 0$ , ou seja, quando não existe a fragilidade (Figuras 4.1 e 4.4). Para os outros casos, apesar da estimação não ter sido muito ruim no tempo inicial de observação, conforme o tempo aumenta começa a existir um aumento na variação entre a função risco acumulada verdadeira e as funções estimadas, principalmente no caso em que o parâmetro de fragilidade é maior, ou seja, no caso de  $\theta_0 = 0.4$ . Isto se explica pois há menos observações para tempos grandes.

As Figuras 4.7 a 4.10 mostram as funções de risco acumulada para o caso da fragilidade lognormal, correspondentes aos valores de  $\theta_0$  iguais a 0.2 e 0.4, ambos para os casos não censurados e censurados.

As Figuras 4.11 a 4.14 apresentam as funções de risco acumuladas para o caso da fragilidade normal. Foram assumidos os mesmos valores para a variância da fragilidade, ou seja,  $\theta_0 = 0.2$  e  $0.4$ , também considerando-se os dados não censurados e censurados.

Podemos observar, em ambos os casos da fragilidade lognormal e normal, que as funções risco acumulada estimadas diferem consideravelmente da verdadeira função risco acumulada, principalmente quando o tempo aumenta.

Portanto, baseado nas figuras descritas acima, podemos dizer que a estimação da função risco acumulada em todos os casos, ou seja, da fragilidade assumir as distribuições gama, lognormal e normal, não foi considerada satisfatória.

Figura 4.1: Funções riscos acumulada com fragilidade gama, utilizando as estimações de  $\bar{\theta}$ , os estratos (1 e 2), dados não censurados e  $\theta_0 = 0.0$

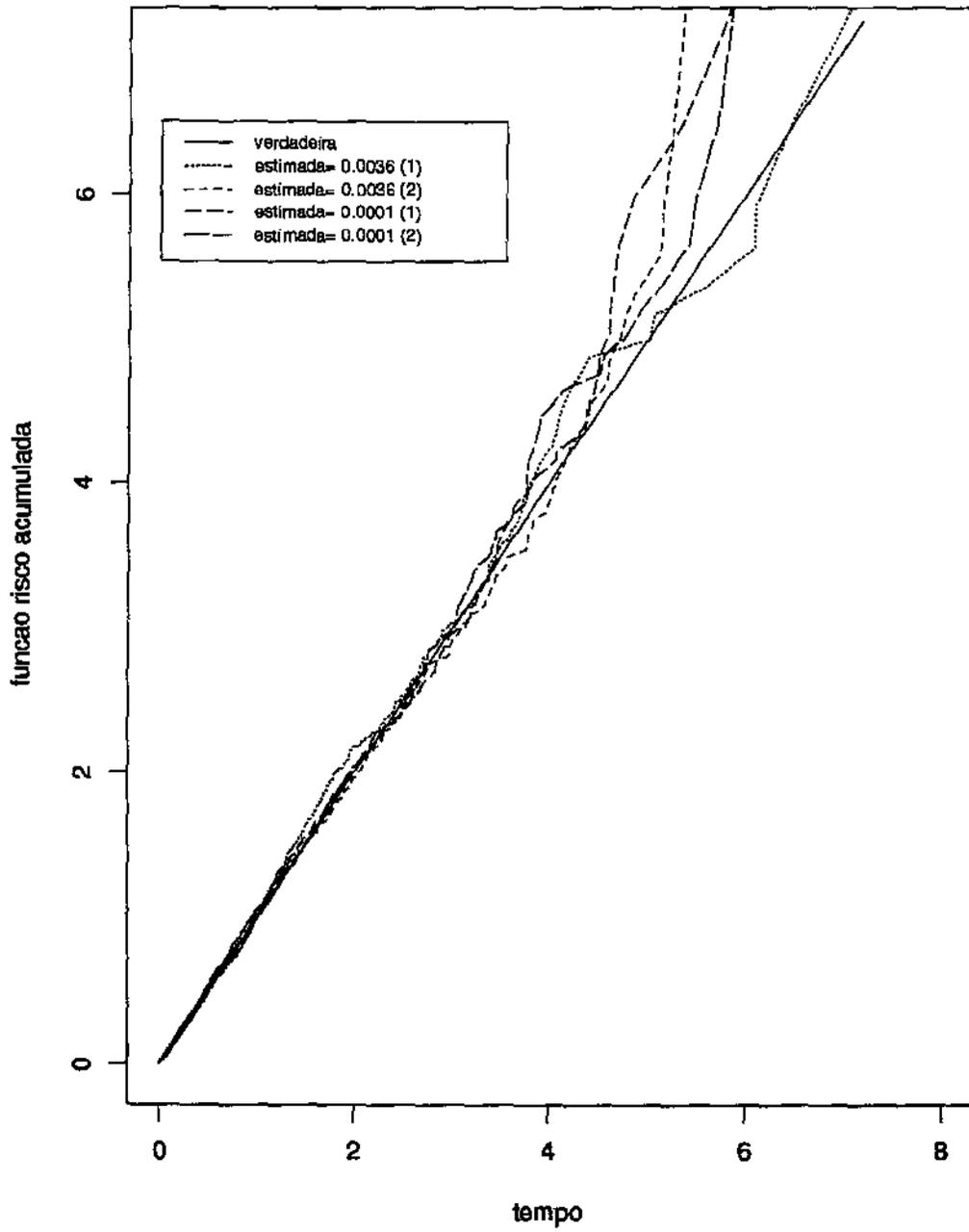


Figura 4.2: Funções riscos acumulada com fragilidade gama, utilizando as estimações de  $\bar{\theta}$ , os estratos (1 e 2), dados não censurados,  $\theta_0 = 0.2$

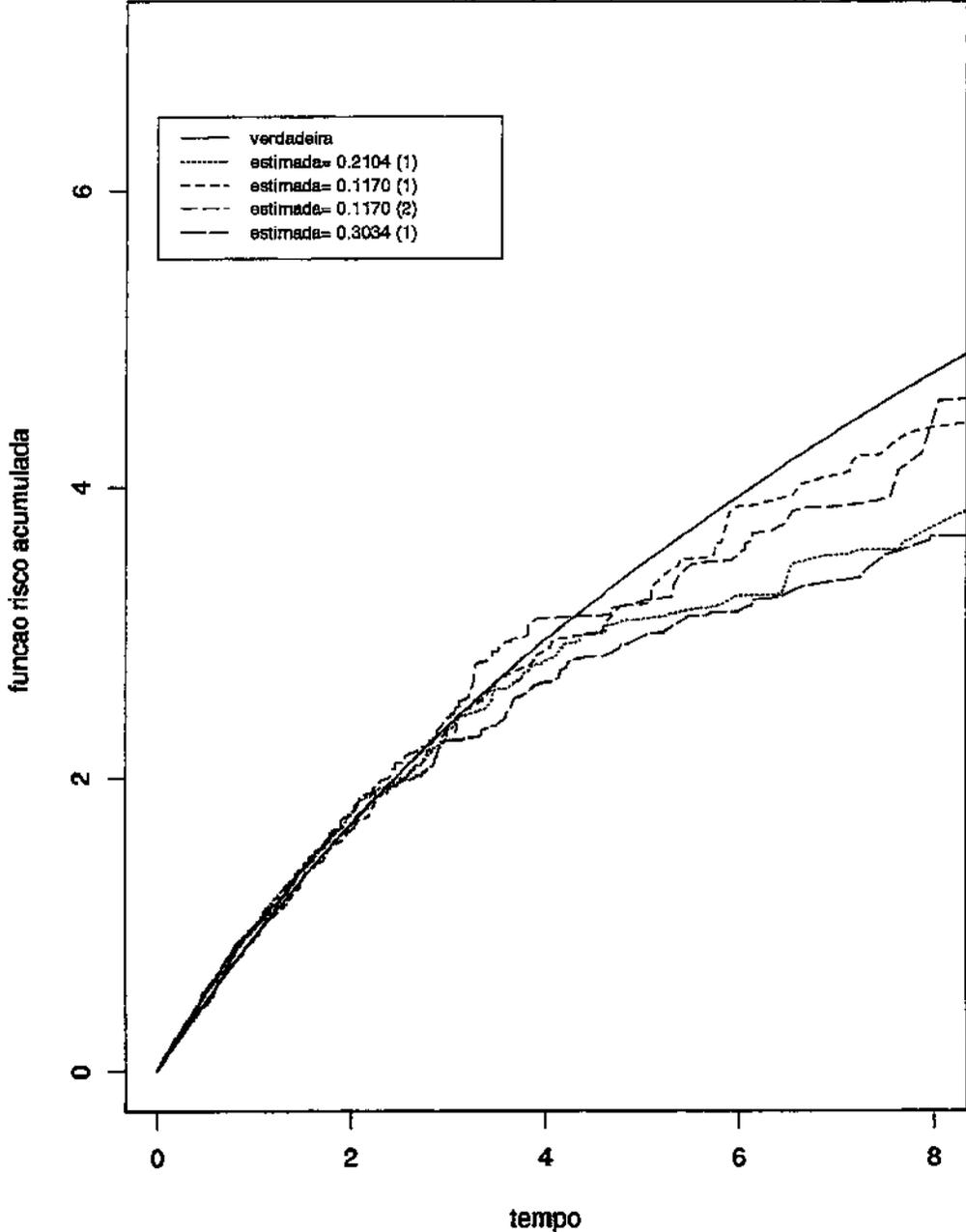


Figura 4.3: Funções riscos acumulada com fragilidade gama, utilizando as estimações de  $\tilde{\theta}$ , os estratos (1 e 2), dados não censurados,  $\theta_0 = 0.4$

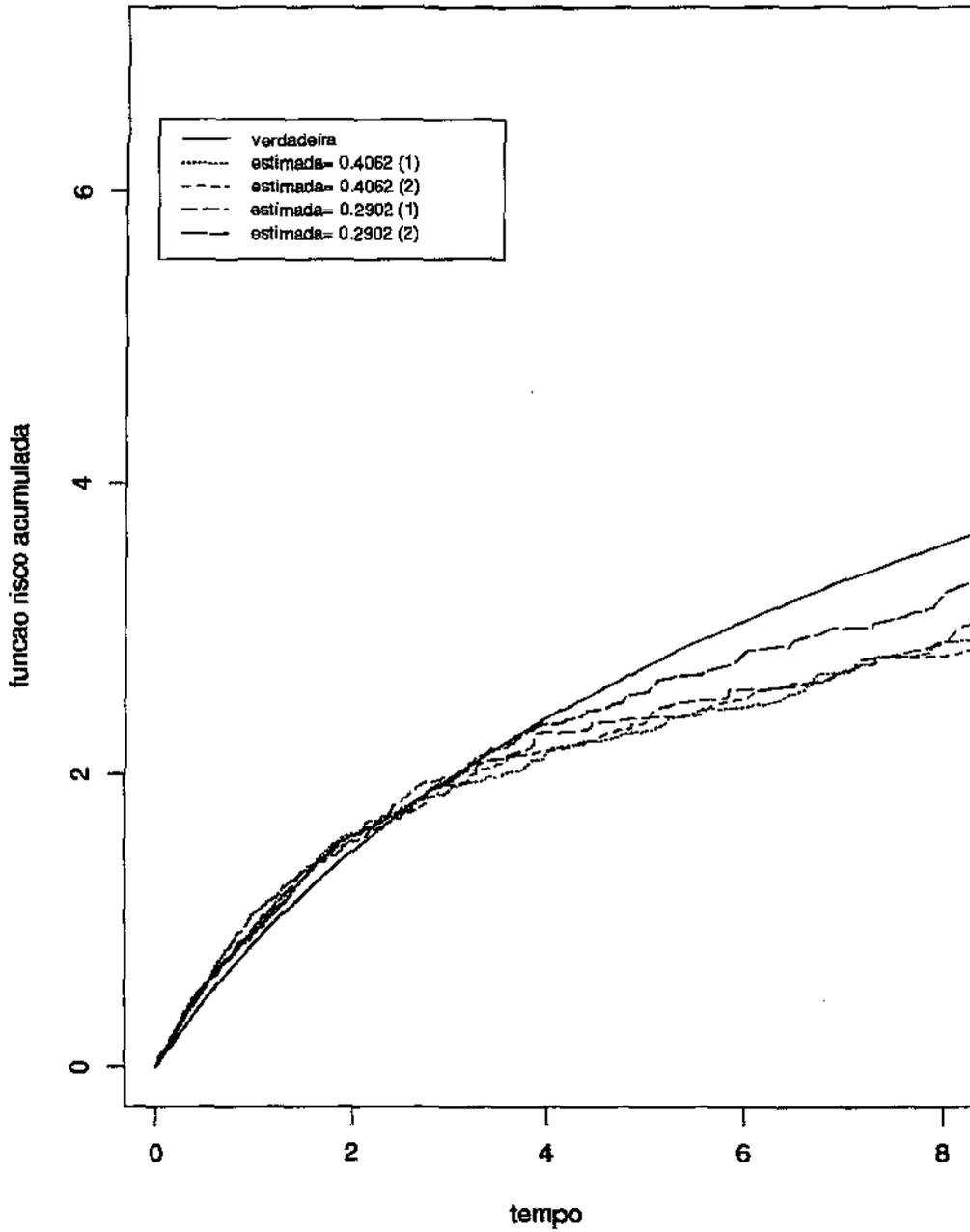


Figura 4.4: Funções riscos acumulada com fragilidade gama, utilizando as estimações de  $\bar{\theta}$ , os estratos (1 e 2), dados censurados,  $\theta_0 = 0.0$

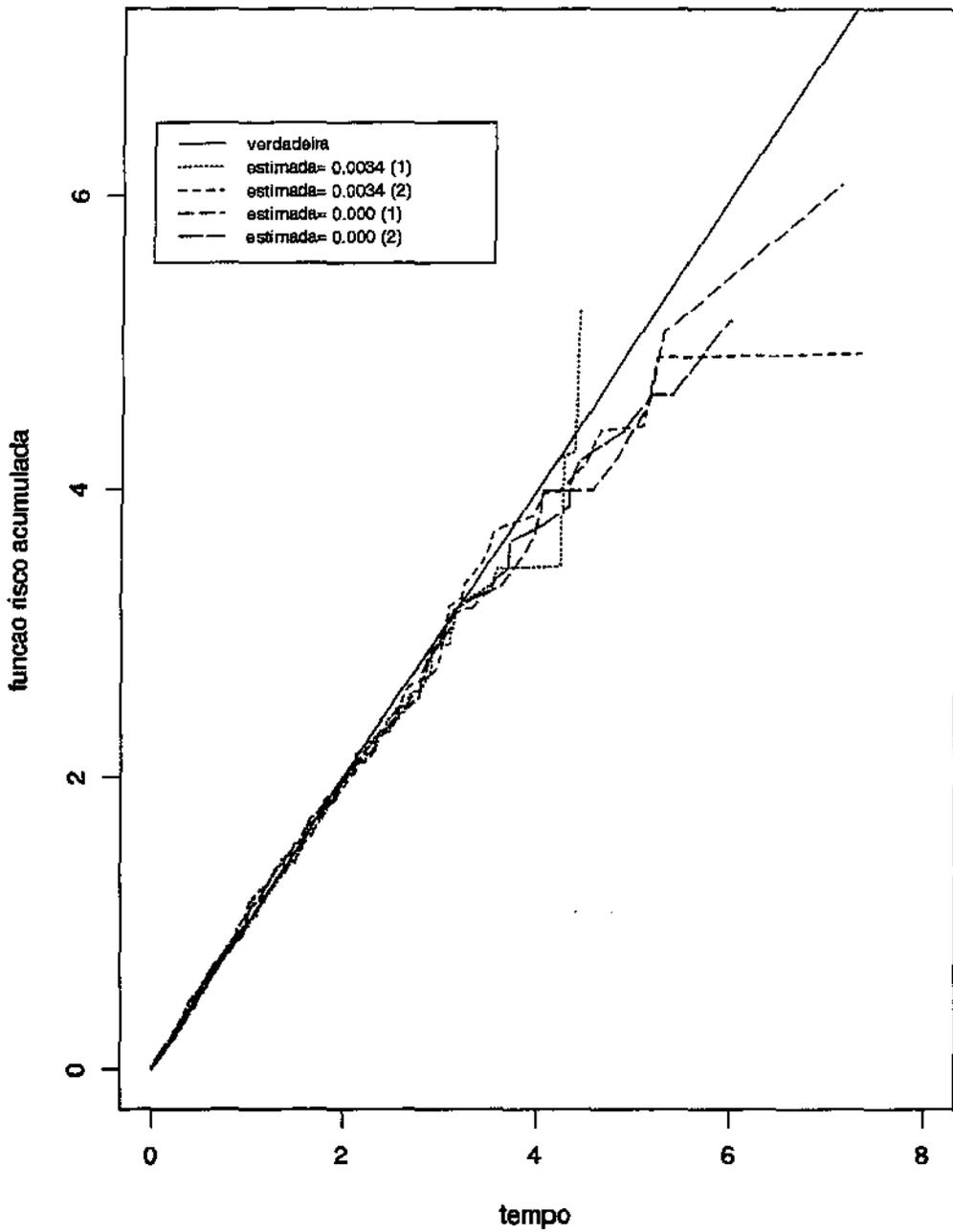


Figura 4.5: Funções riscos acumulada com fragilidade gama, utilizando as estimações de  $\bar{\theta}$ , os estratos (1 e 2), dados censurados,  $\theta_0 = 0.2$

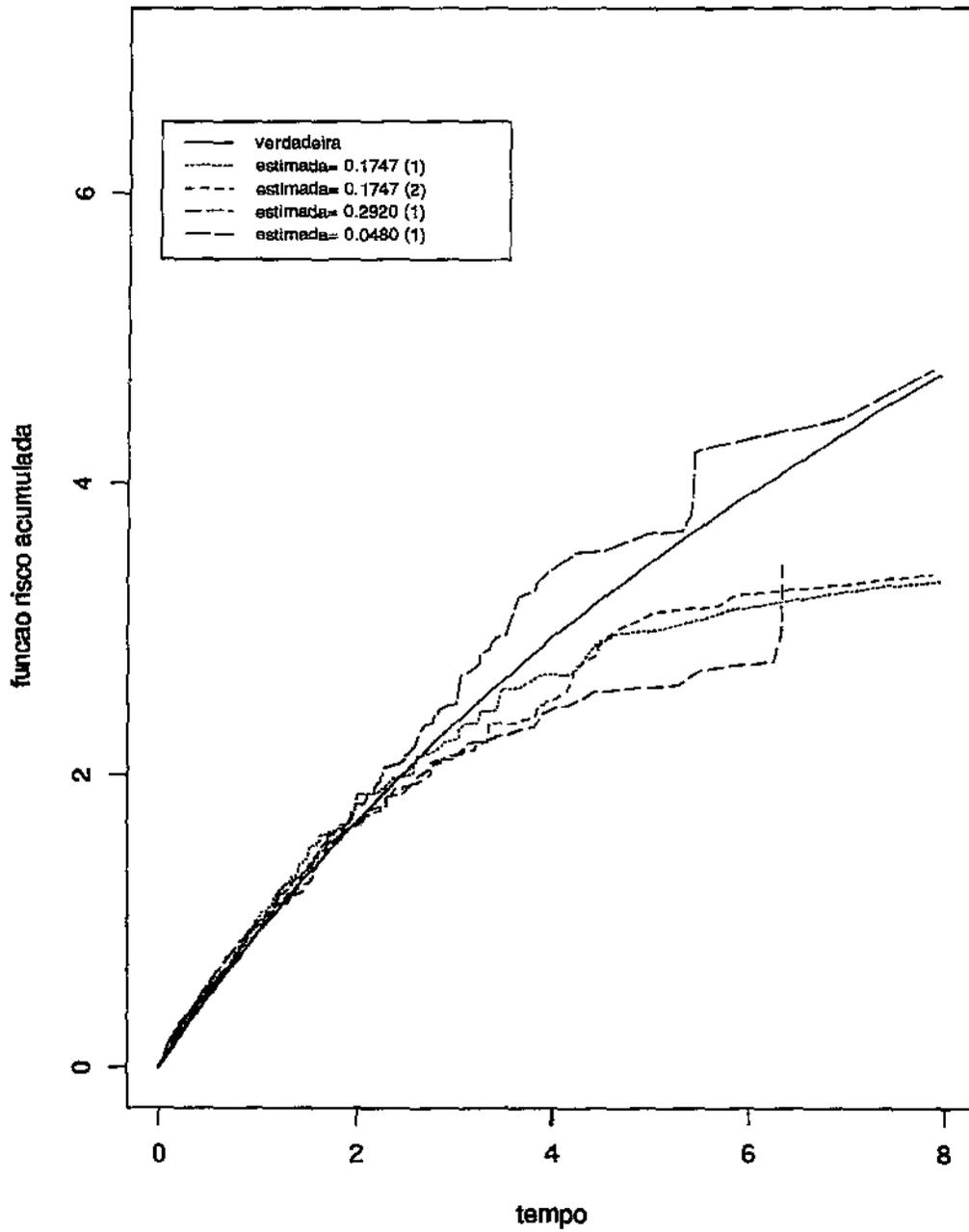


Figura 4.6: Funções riscos acumulada com fragilidade gama, utilizando as estimações de  $\tilde{\theta}$ , os estratos (1 e 2), dados censurados,  $\theta_0 = 0.4$

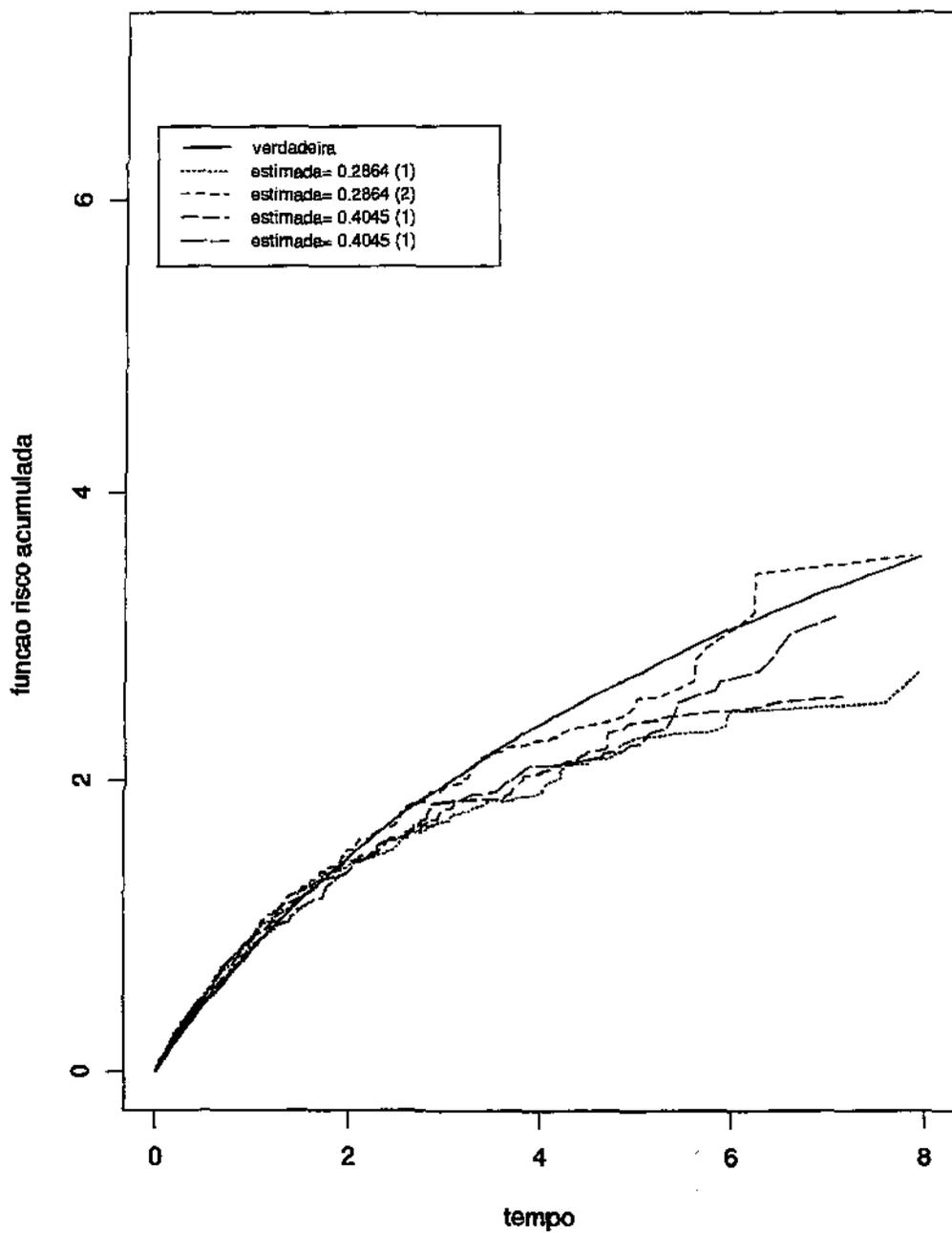


Figura 4.7: Funções riscos acumulada com fragilidade lognormal, utilizando as estimações de  $\bar{\theta}$ , os estratos (1 e 2), dados não censurados,  $\theta_0 = 0.2$

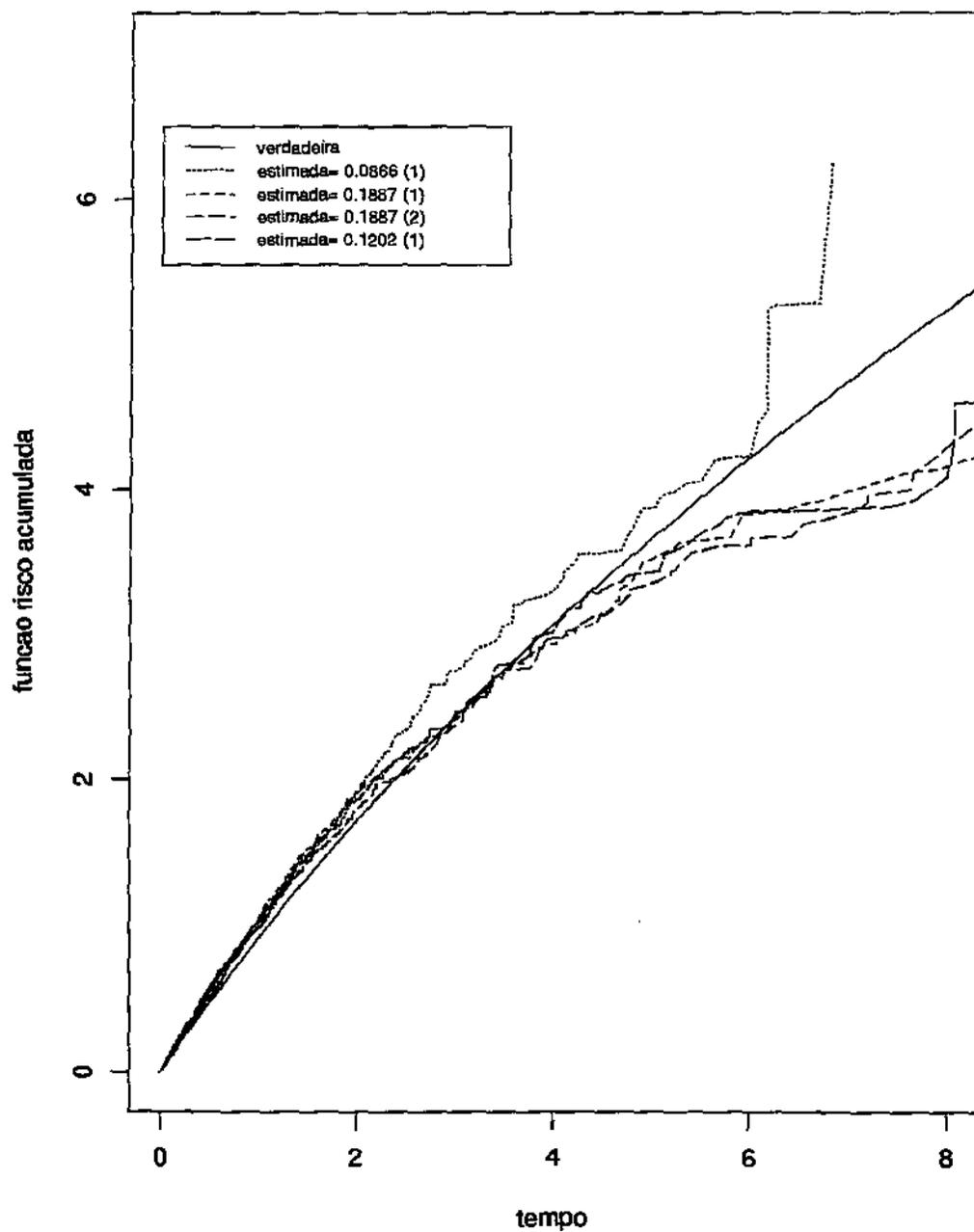


Figura 4.8: Funções riscos acumulada com fragilidade lognormal, utilizando as estimações de  $\tilde{\theta}$ , os estratos (1 e 2), dados não censurados,  $\theta_0 = 0.4$

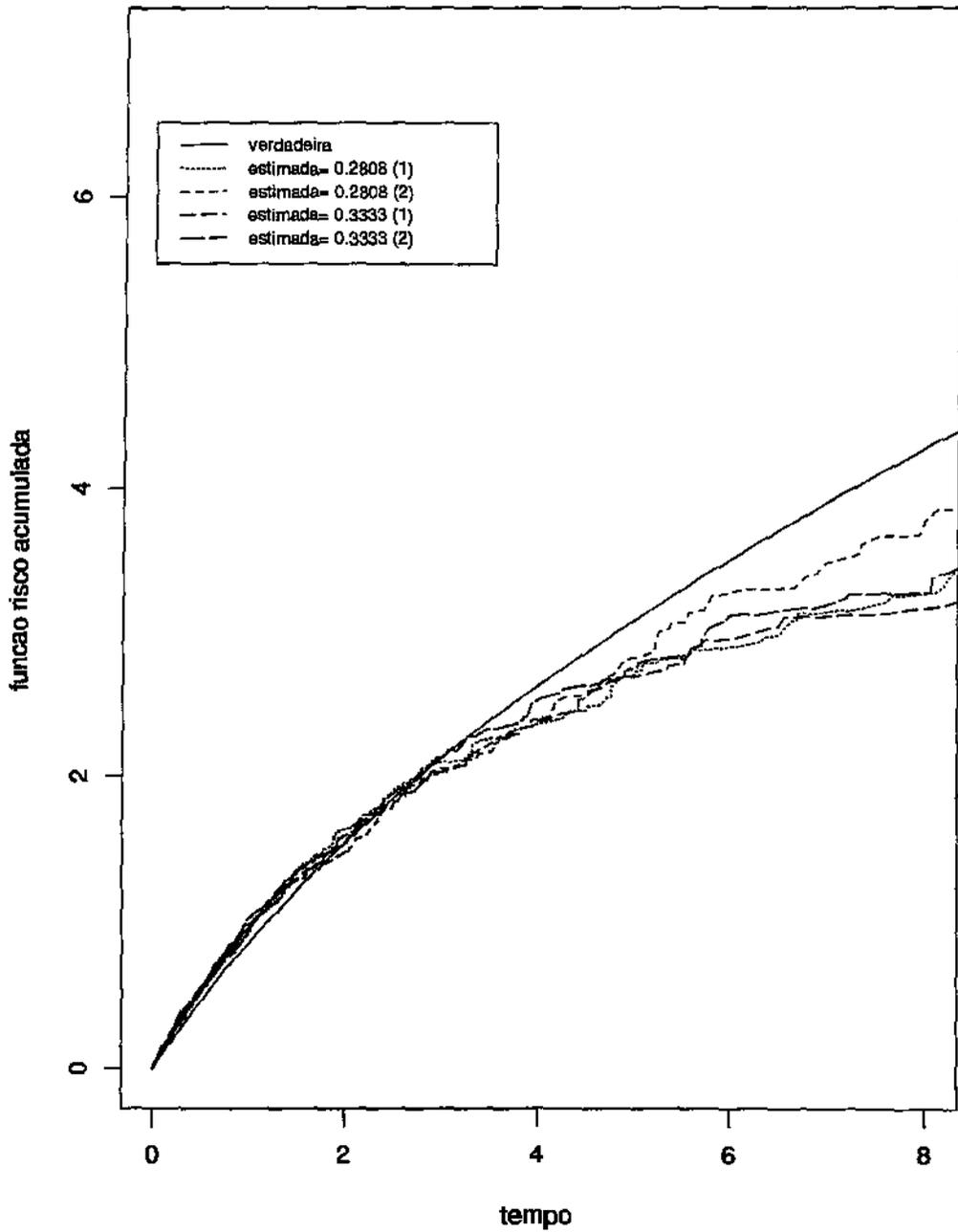


Figura 4.9: Funções riscos acumulada com fragilidade lognormal, utilizando as estimações de  $\bar{\theta}$ , os estratos (1 e 2), dados censurados,  $\theta_0 = 0.2$

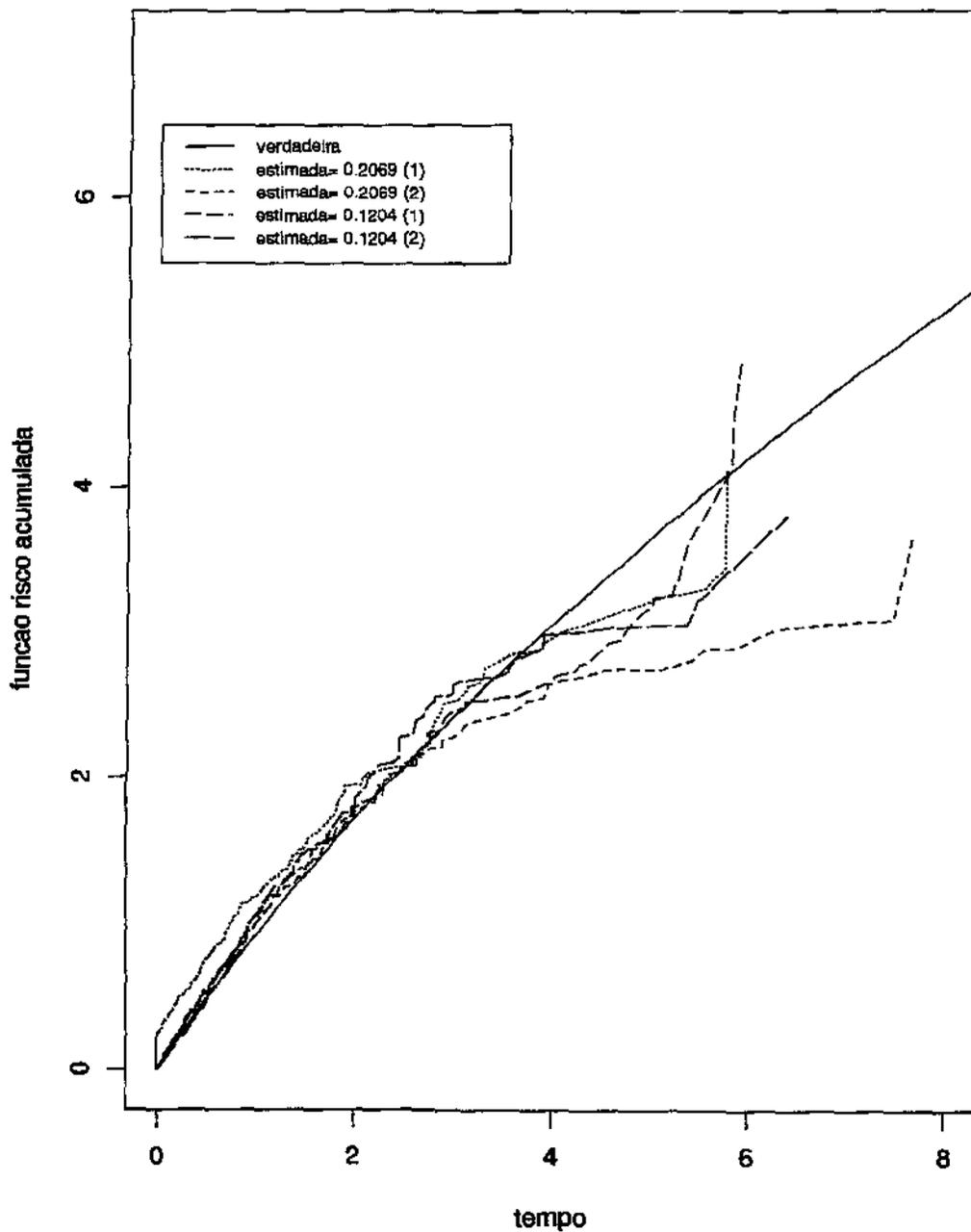


Figura 4.10: Funções riscos acumulada com fragilidade lognormal, utilizando as estimações de  $\tilde{\theta}$ , os estratos (1 e 2), dados censurados,  $\theta_0 = 0.4$

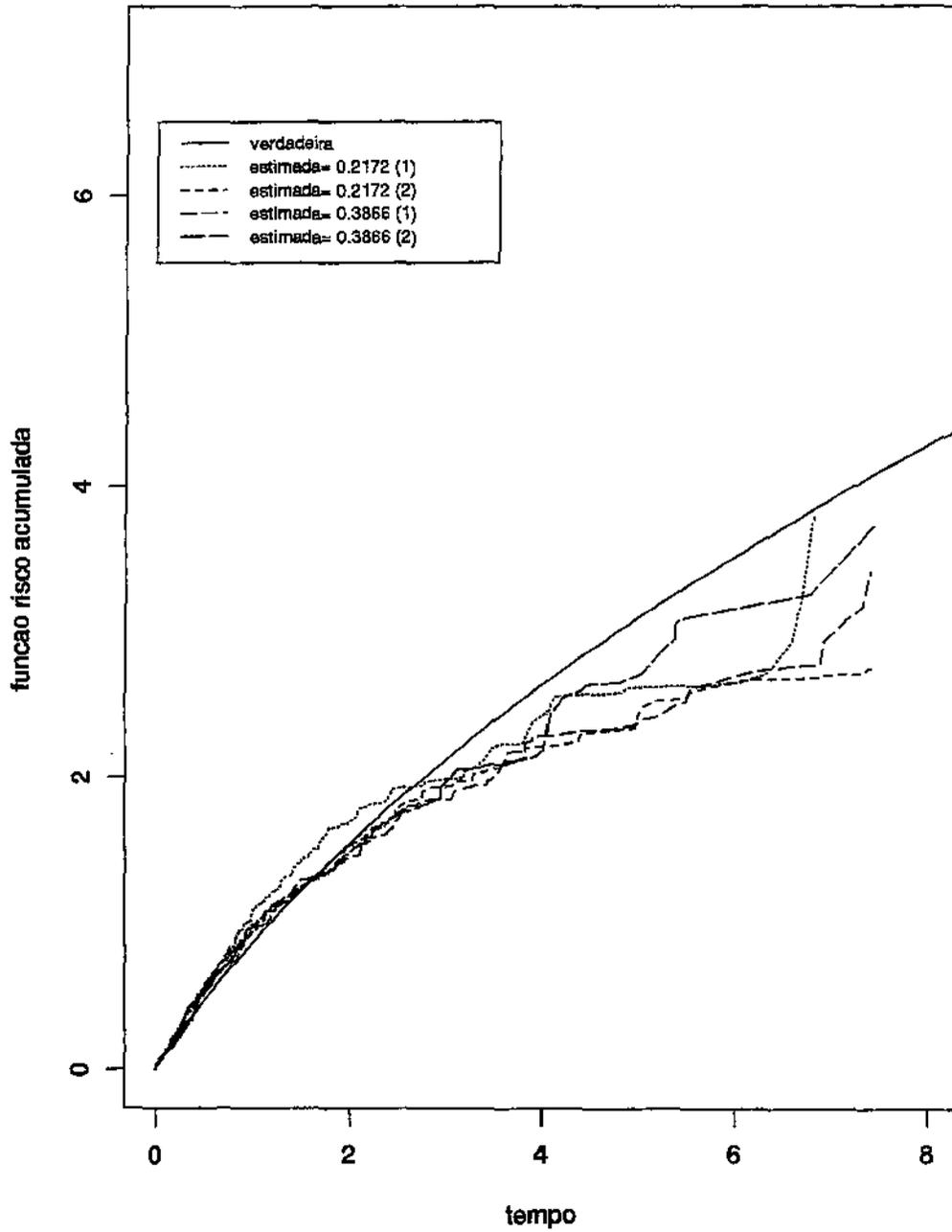


Figura 4.11: Funções riscos acumulada com fragilidade normal, utilizando as estimações de  $\tilde{\theta}$ , os estratos (1 e 2), dados não censurados,  $\theta_0 = 0.2$

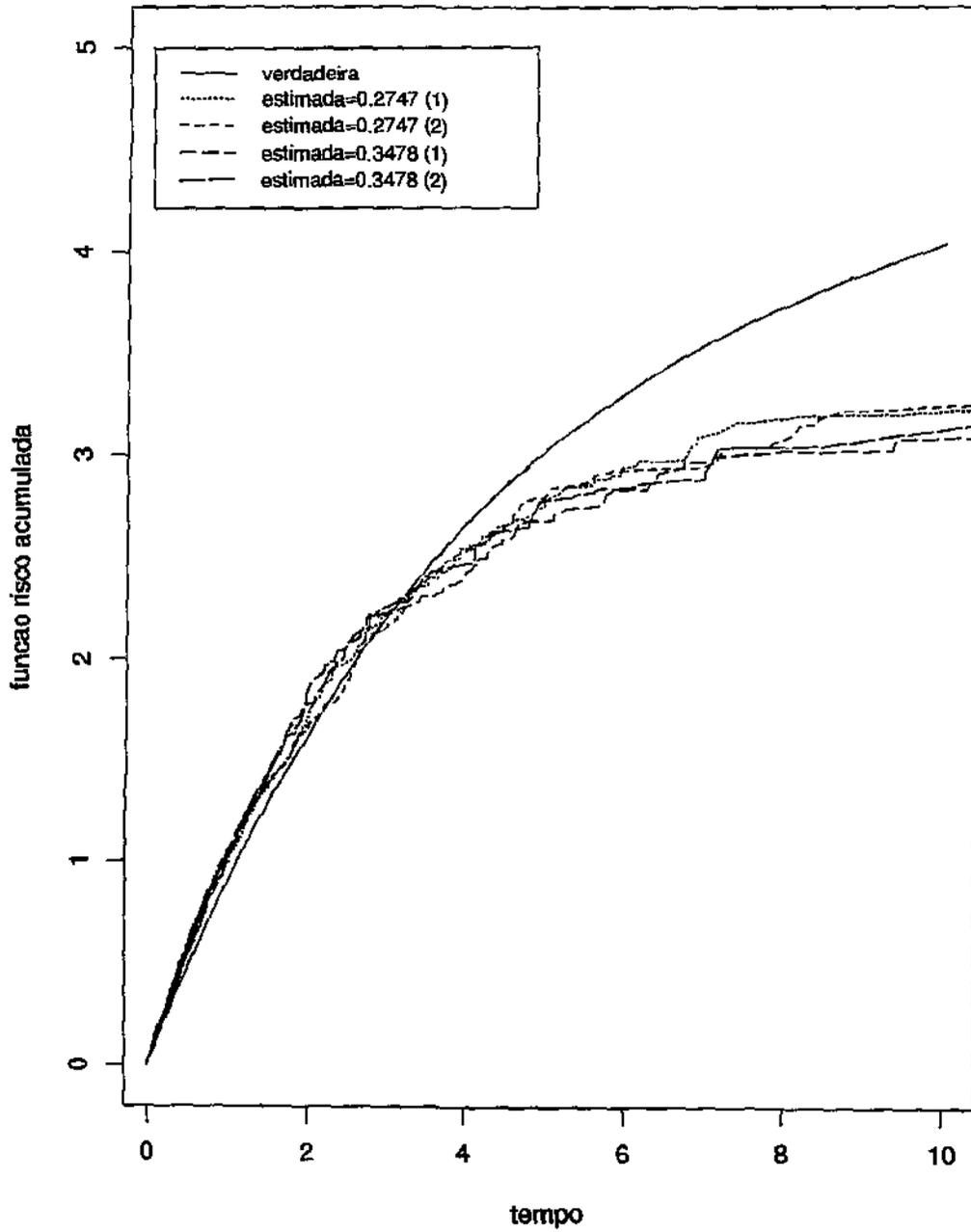


Figura 4.12: Funções riscos acumulada com fragilidade normal, utilizando as estimações de  $\tilde{\theta}$ , os estratos (1 e 2), dados não censurados,  $\theta_0 = 0.4$

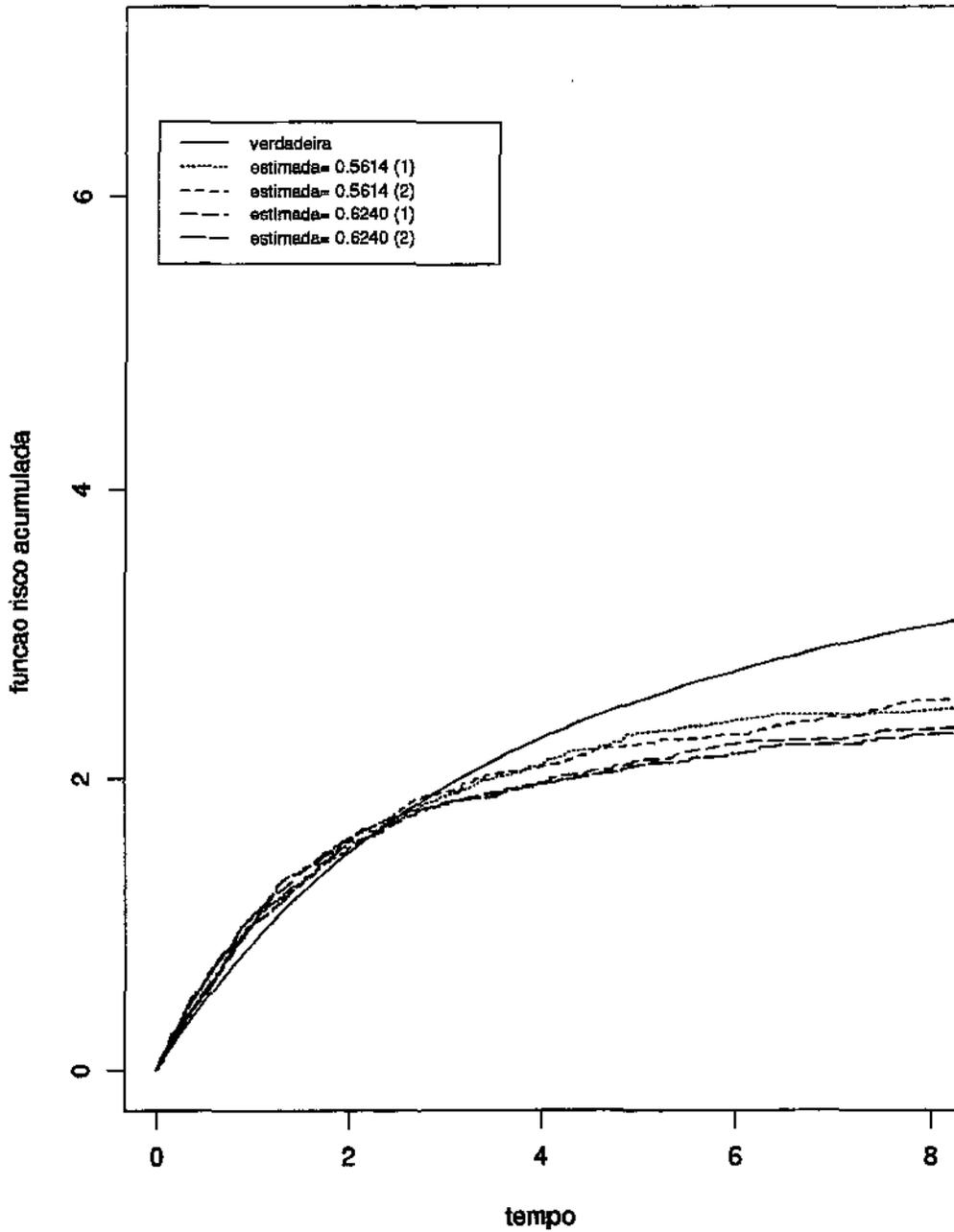


Figura 4.13: Funções riscos acumulada com fragilidade normal, utilizando as estimações de  $\tilde{\theta}$ , os estratos (1 e 2), dados censurados,  $\theta_0 = 0.2$

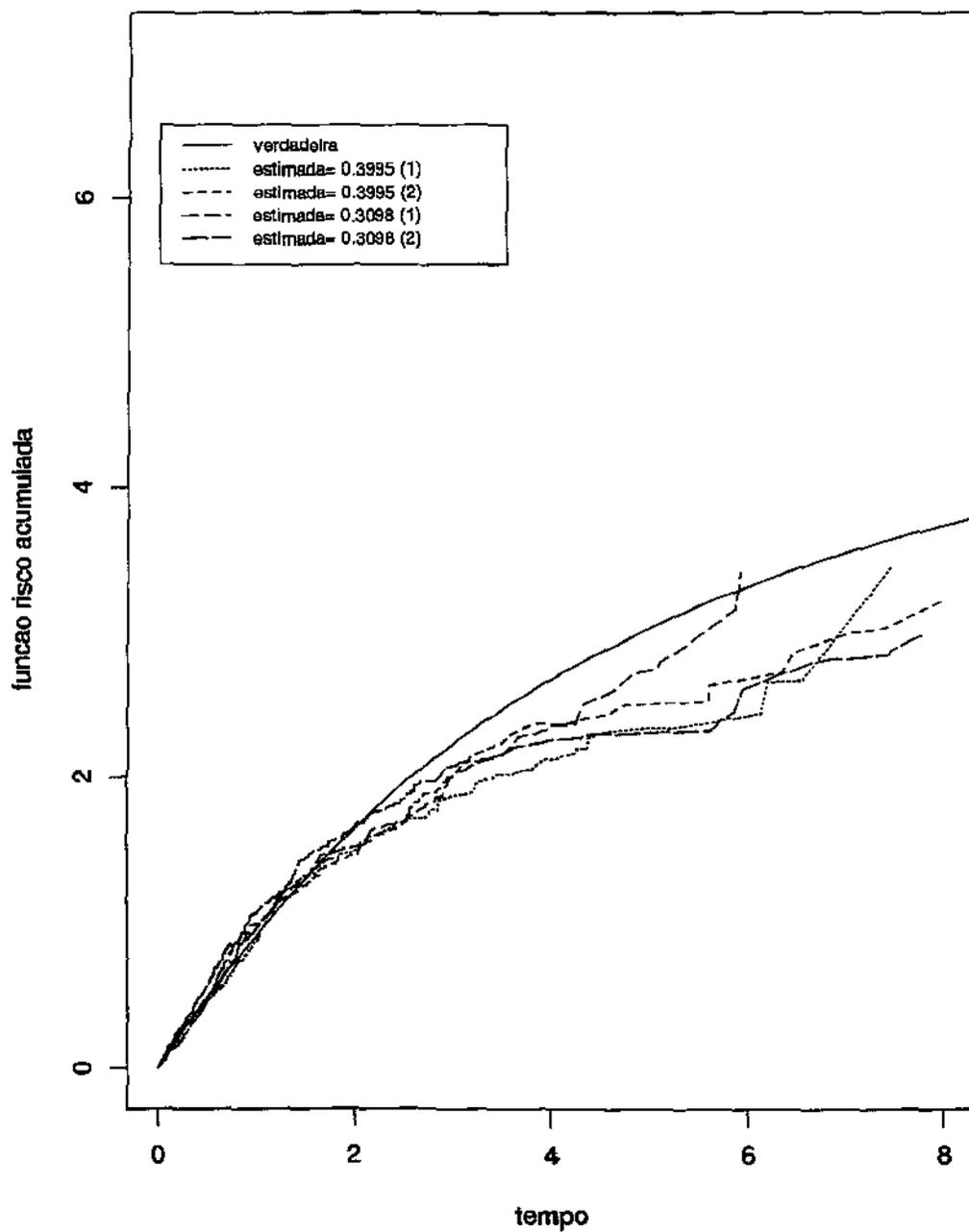
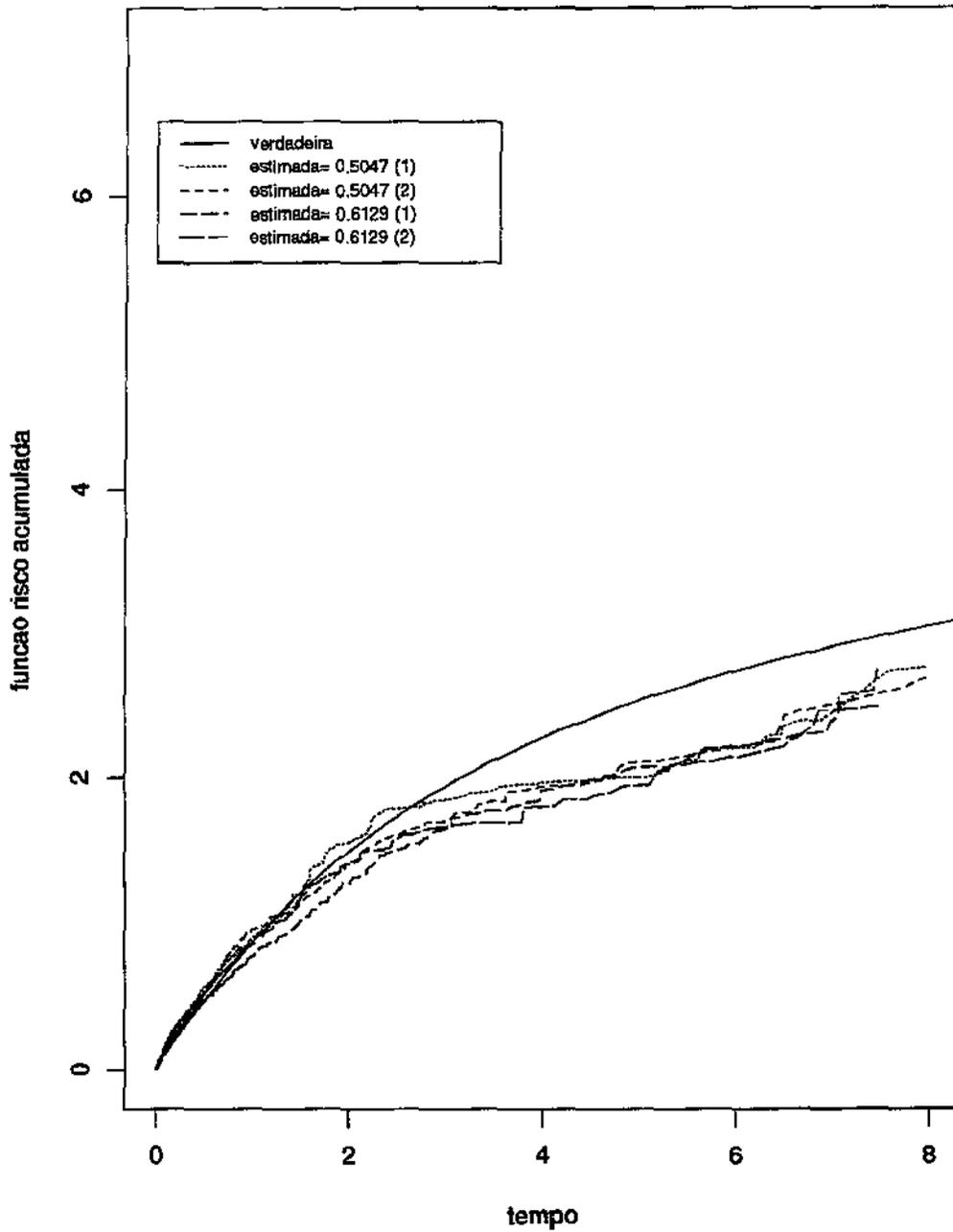


Figura 4.14: Funções riscos acumulada com fragilidade lognormal, utilizando as estimações de  $\tilde{\theta}$ , os estratos (1 e 2), dados censurados,  $\theta_0 = 0.4$



# Capítulo 5

## Considerações Finais

A parte inicial desta dissertação consistiu na apresentação de uma rápida revisão bibliográfica sobre os métodos para a análise de duração de eventos, em particular os métodos que permitem incluir covariáveis não observáveis no modelo, descrevendo o que tem sido feito e onde tem sido aplicado tais estudos. Foram também apresentados alguns conceitos básicos que se tornaram indispensáveis para a compreensão do trabalho proposto nesta dissertação.

No capítulo 2 é apresentado o modelo de fragilidade que permite incluir um efeito aleatório no modelo de risco, que representará a heterogeneidade não observada, ou seja, as covariáveis não observadas que não foram consideradas no modelo. Baseado no trabalho desenvolvido por Nielsen *et al.* (1992) foi proposto um outro estimador para o parâmetro da fragilidade  $\theta$ , que difere do estimador desenvolvido por Nielsen, por estar restrito ao espaço paramétrico  $[0, \infty)$ . Este estimador parece ser mais adequado, pois ele não permite que as estimativas de  $\theta$  assumam valores negativos, ao contrário do estimador sugerido por Nielsen.

Foi desenvolvida toda a parte da estimação dos parâmetros do modelo proposto e dos efeitos aleatórios, utilizando o algoritmo EM. Neste trabalho foi dada preferência aos modelos semi-paramétricos.

Demonstramos também a convergência do estimador do parâmetro de fragilidade proposto e a sua distribuição, utilizando os resultados provados por Murphy (1994, 1995) no caso da fragilidade ter distribuição gama. Como neste caso foi demonstrada a normalidade assintótica para a distribuição do estimador do parâmetro, também foi calculada sua variância assintótica, a partir da matriz de informação de Fisher. No capítulo seguinte foi verificado graficamente a aproximação da função densidade da distribuição normal com relação a cada amostra simulada através da fragilidade gama.

No capítulo 3 foi realizado um estudo de simulação considerando-se além da distribuição gama para a fragilidade, as distribuições lognormal e normal, para verificar se os estimadores da variância da fragilidade eram sensíveis à escolha dessas distribuições de fragilidade. Graficamente foi possível observar que a estimativa da variância da fragilidade, quando esta assume a distribuição gama, é razoavelmente boa se utilizamos um número grande de indivíduos no experimento. Quando assumimos outra distribuição para a fragilidade como no caso das distribuições lognormal e normal, a estimativa da variância da fragilidade se torna viciada, comprometendo assim o resultado final da estimação, pois para cada distribuição assumida para a fragilidade, obtemos resultados diferentes para a variância estimada.

Depois de verificado que os estimadores da variância da fragilidade são sensíveis quanto à escolha da distribuição da fragilidade, verificamos no capítulo 4 que no caso da estimação da função risco acumulada, houve diferença entre as funções estimadas e a verdadeira, exceto quando  $\theta = 0$ , ou seja, quando não existe fragilidade. Portanto, mesmo escolhendo a distribuição gama para a fragilidade, a estimação da função risco acumulada não foi considerada boa.

Note-se que existe, também, um grande interesse em verificar se o estimador da variância da fragilidade ( $\theta$ ) é consistente e assintoticamente normal, pois estamos interessados em testar a hipótese de  $\theta = 0$  versus  $\theta > 0$ , ou seja, testar se a população estudada é homogênea ou não.

No caso da fragilidade assumir a distribuição gama, foi verificada a consistência e a

normalidade assintótica do estimador da variância da fragilidade. Já quando a fragilidade assume as distribuições lognormal e normal o estimador da variância da fragilidade não se mostrou consistente, apesar de existir uma certa normalidade na distribuição das estimativas.

Como a fragilidade é uma variável não observável a utilização de modelos paramétricos na sua estimação não é aconselhável. Uma outra possibilidade para este caso seria utilizar a estimação e testes não paramétricos, mas o objetivo desta dissertação foi verificar a consistência do estimador da fragilidade com relação à escolha de outras distribuições paramétricas. Entretanto o desenvolvimento de um estimador não paramétrico pode ser de grande interesse e, possivelmente, estudado e aplicado em trabalhos futuros.

# Bibliografia

- [1] AALEN, O.O. (1988) Heterogeneity in survival analysis. *Statistics in Medicine* 7:1121-37.
- [2] ALLISON, P.D. (1982) Discrete-Time methods for the analysis of event histories. In: Leinhardt, S. (Ed.) *Sociological Methodology*. San Francisco, Jossey-Bass. pp. 61-98.
- [3] BLOSSFELD, H.-P.; HAMERLE, A. (1992) Unobserved heterogeneity in event history models. *Quality and Quantify* 26:157-68.
- [4] CLAYTON, D.G. (1991) A Monte Carlo Method for Bayesian Inference in Frailty Models. *Biometrics* 47:467-85.
- [5] COX, D.R. (1972) Regression models and life-tables. *Journal of the Royal Statistical Society B*, 34:187-202.
- [6] DEMPSTER, A.P.; LAIRD, N.M.; RUBIN, D.B. (1977) Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society B* 38:1-38.
- [7] FLEMING, T. R. ; HARRINGTON, D. P. (1991) *Counting Processes and Survival Analysis*. J. Wiley.

- [8] GREENWOOD, P.E.; WEFELMEYER, W. (1990) Efficiency of estimators for partially specified filtered models. *Stochastic Process. Appl.* **36**:353-370.
- [9] HECKMAN, J.; SINGER, B. (1982) Population heterogeneity in demographic models. In: Land, K.; Rogers, A. (Ed.) *Multidimensional Mathematical Demography*. New York, Academic Press. pp. 567- 599.
- [10] HOEM, J.M. (1990) Identifiability in hazards models with unobserved heterogeneity: The compatibility of two apparently contradictory results. *Theoretical Population Biology* **37**: 124-28.
- [11] HOUGAARD, P. (1986) Survival models for heterogeneous populations derived from stable distributions. *Biometrika* **73**:387-96.
- [12] KLEIN, J.P. (1992) Semiparametric Estimation of Random Effects Using the Cox Model Based on the EM Algorithm. *Biometrics* **48**:795-806.
- [13] LAWLESS, J.F. (1982) *Statistical Models and Methods for Lifetime Data*. New York, J. Wiley.
- [14] MANTON, K.G.; SINGER, B.; WOODBURY, M.A. (1992) Some issues in the quantitative characterization of heterogeneous populations. In: Trussel, J.; Hankinson, R.; Tilton, J. (Ed.) *Demographic Applications of Event History Analysis*. New York, Oxford University Press. pp. 9-37.
- [15] MATUDA, N.S. (1998) Heterogeneidade não observada na análise da história de eventos. Campinas. IMECC/UNICAMP. Dissertação de Mestrado.
- [16] Murphy, S.A. (1994) Consistency in a Proportional Hazards Model Incorporating a Random Effect. *The Annals of Statistics* **22**:712-731.
- [17] ———, (1995) Asymptotic Theory for the Frailty Model. *The Annals of Statistics* **23**:182-198.

- [18] NAMBOODIRI, K.; SUCHINDRAN, C.M. (1990) Life Table Techniques and their Applications. *Orlando, Academic Press.*
- [19] NIELSEN, G.G.; GILL, R.D.; ANDERSEN, P.K.; SORENSEN, T.I.A. (1992) A Counting Process Approach to Maximum Likelihood Estimation in Frailty Models. *Scandinavian Journal of Statistics* **19**:25-43.
- [20] TRUSSEL, J.; RODRIGUES, G. (1990) Heterogeneity in demographic research. In: Adams, J. *et. al.* (Ed.) *Convergent Questions in Genetics and Demography.* New York, Oxford University Press. pp. 111-132.
- [21] VAUPEL, J.W.; MANTON, K.G.; STALLARD, E. (1979) The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* **16**:439-54.
- [22] ——— ; YASHIN, I.A. (1985) Heterogeneity's ruses: some surprising effects of selection on population dynamics. *American Statistician* **39**:176-85.
- [23] WASSEL, J.T.; MOESCHBERGER, M.L. (1993) A bivariate survival model with modified gamma frailty for assessing the impact of interventions. *Statistics in Medicine* **12**:241-48.

# Anexo

## Listagem do Programa

O seguinte programa corresponde ao trabalho de simulação realizado nesta tese, onde os tempos de sobrevivência foram gerados no SPLUS e o processo de interação e maximização foram programados em MATLAB.

### **SPLUS (Dados não censurados)**

temp1 ~ Exponencial (1)

temp2 ~ Exponencial (1)

z ~ Gama

c1 = temp1 / z

c2 = temp2 / z

write(c1 , "c1.txt ") (cria um arquivo txt no diretório C:\ Spluswin \ Home \ c1.txt)

write(c2 , "c2.txt ")

transferir estes arquivos para C: \ Matlab \ Bin

### **MATLAB**

#### **Programa (principal)**

load c1.txt (lê o arquivo )

load c2.txt

dados5 (programa que organiza os dados )

alfa5 (programa que calcula a função risco padrão acumulada 1 e 2 )

rod5 (programa que roda a maximização de theta)

zest5 (programa que calcula os valores estimados dos  $Z_i$ )

rotina5 (programa que cria um loop até o parâmetro theta convergir)

### **Programa ( dados5 )**

```
t1=zeros(1,500);
t2=zeros(1,500);
for i=1:500;
t1(i)=c1(i);
t2(i)=c2(i);
end;
t1o= sort(t1);
t2o=sort(t2);
z=ones(1,500);
d1=zeros(1,500);
d2=zeros(1,500);
for i=1:500;
for j=1:500;
if t1(i) ≥ t1o(j);
d1(i) = d1(i) + 1; (vetor d1(i) irá indicar que o indivíduo(i) falhou no tempo t1o(d1(i)))
end;
if t2(i) ≥ t2o(j);
d2(i) = d2(i) +1;
end; end;
end;
```

### **Programa ( alfa5 )**

```
a1=zeros(1,500);
a2=zeros(1,500);
for j=1:500;
for i=1:500;
if j ≤ d1(i);
a1(j) = a1(j) + z(i);
```

```

end;
if j ≤ d2(i);
a2(j) = a2(j) + z(i);
end;
end;
a1(j) = 1 / a1(j);
a2(j) = 1 / a2(j);
end;
r1=zeros(1,500);
r2=zeros(1,500);
for i=1:500;
for j=1:d1(i);
r1(i) = r1(i) + a1(j); (função risco padrão acumulada estimada do estrato 1, individual)
end;
for j=1:d2(i);
r2(i) = r2(i) + a2(j); (função risco padrão acumulada estimada do estrato 2, individual)
end;
end;

```

### **Programa ( rod5 )**

```

theta1 = 0.0;
theta2 = 2.0;
options = foptions;
options(14) = 1000;
theta = fmin('the5',theta1,theta2,options,r1,r2,a1,a2) (calcula o máximo de theta usando

```

a function the5)

### **the5 ( subprograma)**

```

function somap= the5(theta,r1,r2,a1,a2);
somap = 0;

```

```

k2 = (1/theta) + 2;
for p=1:500;
k3 = r1(p) + r2(p);
somap = somap - k2*log( 1 + (theta*k3) ) + log(a1(p)) + log(a2(p));
end;
somap = somap + (500*(log( 1 + theta)));
somap = -somap;
Programa ( zest5 )
r1=zeros(1,500);
r2=zeros(1,500);
for i=1:500;
for j=1:d1(i);
r1(i) = r1(i) + a1(j);
end;
for j=1:d2(i);
r2(i) = r2(i) + a2(j);
end;
z(i) = ( (1/theta) + 2 ) / ( (1/theta) + r1(i) + r2(i) ); (  $Z_i$  estimados)
end;
Programa ( rotina5 )
e=theta;
d=theta;
b=0;
while ( d > 1.0000e-003)
zest5
alfa5
rod5
b = b+1

```

```
d = abs(e - theta)
e=theta;
end;
```

## **SPLUS ( Dados censurados )**

```
temp1 ~ Exponencial (1)
temp2 ~ Exponencial (1)
cens1 ~ Uniforme (0, 8)
cens2 ~ Uniforme (0, 8)
z ~ Gama
c1 = temp1 / z
c2 = temp2 / z
write(c1 , "c1.txt ") (cria um arquivo txt no diretório C:\ Spluswin \Home \ c1.txt)
write(c2 , "c2.txt ")
write(cens1 , "s1.txt ")
write(cens2 , "s2.txt ")
transferir estes arquivos para C: \ Matlab \ Bin
```

## **MATLAB**

### **Programa (principal)**

```
load c1.txt (lê o arquivo )
load c2.txt
load s1.txt
load s2.txt
dadcens5 (programa que organiza os dados )
alfcens5 (programa que calcula a função risco padrão acumulada 1 e 2 )
rodcens5 (programa que roda a maximização de theta)
zestcen5 (programa que calcula os valores estimados dos  $Z_i$ )
rotcens5 (programa que cria um loop até o parâmetro theta convergir)
```

## Programa ( dadcens5)

```
x1=zeros(1,500);
x2=zeros(1,500);
ss1=zeros(1,500);
ss2=zeros(1,500);
for i=1:500;
x1(i)=c1(i);
x2(i)=c2(i);
ss1(i)=s1(i);
ss2(i)=s2(i);
end;
t1=zeros(1,500);
t2=zeros(1,500);
b1=zeros(1,500);
b2=zeros(1,500);
for i=1:500;
if x1(i)<= ss1(i);
b1(i)=1;
t1(i)= x1(i);
else
t1(i)= ss1(i);
end;
if x2(i)<= ss2(i);
b2(i)=1;
t2(i)= x2(i);
else
t2(i)= ss2(i);
end;
end;
```

```

end;
t1o= sort(t1);
t2o=sort(t2);
z=ones(1,500);
d1=zeros(1,500);
d2=zeros(1,500);
for i=1:500;
for j=1:500;
if t1(i) >= t1o(j);
d1(i) = d1(i) + 1;
end;
if t2(i) >= t2o(j);
d2(i) = d2(i) + 1;
end;
end;
end;
end;
for i=1:500;
h(i)= b1(i) + b2(i);
end;

```

### **Programa ( alfcens5 )**

```

a1=zeros(1,500);
a2=zeros(1,500);
for j=1:500;
for i=1:500;
if j <= d1(i);
a1(j) = a1(j) + z(i);
end;
if j <= d2(i);

```

```

a2(j) = a2(j) + z(i);
end; end;
a1(j) = 1 / a1(j);
a2(j) = 1 / a2(j);
end;
for i=1:500;
if b1(i)==0;
a1(d1(i))=0;
end;
if b2(i)==0;
a2(d2(i))=0;
end;
end;
r1=zeros(1,500);
r2=zeros(1,500);
for i=1:500;
for j=1:d1(i);
r1(i) = r1(i) + a1(j); ( função risco padrão acumulada do estrato 1 , individual )
end;
for j=1:d2(i);
r2(i) = r2(i) + a2(j); ( função risco padrão acumulada do estrato 2 , individual )
end;
end;

```

### **Programa ( rodzens5 )**

```

theta1 = 0.0;
theta2 = 2.0;
options = foptions;
options(14) = 1000;

```

```
theta = fmin('thecens5',theta1,theta2,options,r1,r2,a1,a2,b1,b2,h)
```

(calcula o máximo de theta usando a function thecens5)

### **thecens5 ( subprograma )**

```
function somap= thecens5(theta,r1,r2,a1,a2,b1,b2,h);  
somap = 0;  
k2 = (1/theta) + 2;  
k4 = (1/theta) + 1;  
k5 = (1/theta);  
for p=1:500;  
k3 = r1(p) + r2(p);  
if h(p)==2;  
somap = somap + log(1+ theta) - k2*log( 1 + (theta*k3) );  
end;  
if h(p) ==1;  
somap = somap - k4*log( 1 + (theta*k3) );  
end;  
if h(p) ==0;  
somap = somap - k5*log( 1 + (theta*k3) );  
end;  
end;  
somap = -somap;
```

### **Programa ( zestcen5 )**

```
r1=zeros(1,500);  
r2=zeros(1,500);  
for i=1:500;  
for j=1:d1(i);  
r1(i) = r1(i) + a1(j);  
end;  
end;
```

```

for j=1:d2(i);
r2(i) = r2(i) + a2(j);
end;
if h(i)== 2;
z(i) = ( (1/theta) + 2 ) / ( (1/theta) + r1(i) + r2(i) );
end;
if h(i)== 1;
z(i) = ( (1/theta) + 1 ) / ( (1/theta) + r1(i) + r2(i) );
end;
if h(i)== 0;
z(i) = ( (1/theta) ) / ( (1/theta) + r1(i) + r2(i) );
end;
end;

```

### **Programa ( rotcen5 )**

```

e=theta;
d=theta;
b=0;
while ( d > 1.0000e-003)
zestcen5
alfcens5
rodcens5
b = b+1
d = abs(e - theta)
e=theta;
end;

```