

Planos amostrais eficientes na estimação em modelos de mistura de escala com caudas pesadas

Renato Monteiro Pinha Gomes



Universidade Federal do Rio de Janeiro
Instituto de Matemática
Departamento de Métodos Estatísticos

2023

Planos amostrais eficientes na estimação em modelos de mistura de escala com caudas pesadas

Renato Monteiro Pinha Gomes

Tese de Doutorado submetida ao Programa de Pós-Graduação em Estatística do Instituto de Matemática da Universidade Federal do Rio de Janeiro, UFRJ, como parte dos requisitos necessários à obtenção do grau de Doutor em Estatística.

Orientadores: Fernando Antônio da Silva Moura, João Batista de Moraes Pereira.

Rio de Janeiro, RJ - Brasil

2023

Planos amostrais eficientes na estimação em modelos de mistura de escala com caudas pesadas

Renato Monteiro Pinha Gomes

Tese de Doutorado submetida ao Programa de Pós-Graduação em Estatística do Instituto de Matemática da Universidade Federal do Rio de Janeiro, UFRJ, como parte dos requisitos necessários à obtenção do grau de Doutor em Estatística.

Aprovada por:

Prof. Fernando A. S. Moura, Ph.D, UFRJ.

Prof. João Batista de Morais Pereira, Dr.Sc, UFRJ.

Prof^a. Kelly Cristina Mota Gonçalves, Dr.Sc, UFRJ.

Prof^a. Marina Silva Paes, Dr.Sc, UFRJ.

Prof. Gustavo da Silva Ferreira, Dr.Sc, ENCE.

Prof. Leonardo Soares Bastos, Ph.D, Fiocruz.

Prof. Pedro Luis do Nascimento Silva, Ph.D, ENCE.

Rio de Janeiro, RJ - Brasil

2023

FICHA CATALOGRÁFICA

Aos meus pais e irmãos.

“Scream for me Brazil!”

Bruce Dickinson.

Agradecimentos

Agradeço a Deus pelo seu infinito amor.

Aos meus pais Celso e Olga que sempre estiveram ao meu lado e que me estimularam a cursar o doutorado.

Aos meus irmãos César, Letícia e Livia pelo carinho e paciência que sempre tiveram comigo.

Ao meu cunhado Pedro e minha cunhada Roberta.

As minhas sobrinhas Maria Alice e Beatriz.

Aos meus tios, em especial ao Silvio Pinha que me ajudou e me estimulou a cursar o curso de Estatística.

Aos meus avôs Rubens e Acyr e às minhas avós Sylvia e Zaira.

Aos meus professores da graduação, do mestrado e do doutorado, em especial: José Francisco, Julio Siqueira, Ricardo Accioly, Eduardo Campos, Maria Elvira, Carlos Abanto-Valle e Flávia Landim.

Aos meus orientadores Fernando e João pelo auxílio fornecido para elaboração da minha tese. Minha eterna gratidão pelo auxílio e orientações dadas.

A todos os meus amigos, em especial: Allan, Ayrton, Gabriel, Humberto, Isabela, Luiz, Marcel, Marcus, Pedro, Rafael, Raíra, Rebecca, Roberta, Rodrigo, Victor Archanjo e Wagner.

À CAPES pelo apoio financeiro dos meus estudos.

Por fim aos professores que aceitaram fazer parte da banca.

Resumo

Em várias áreas de estudo há diversos conjuntos de dados que apresentam *outliers*, ou ainda, concentrações significativas de valores extremos nas caudas das suas respectivas distribuições. No caso particular da regressão T-Student, o parâmetro graus de liberdade é que determina o grau de robustez. Quanto menor for o número de graus de liberdade, mais pesadas serão as caudas. Entretanto, tal parâmetro é de difícil estimação, uma vez que para bem estimá-lo é necessário observar valores extremos da distribuição.

No contexto de populações finitas, geralmente o pesquisador não tem acesso a todos os elementos populacionais para a modelagem. Um plano amostral deve ser adotado para obtenção da amostra. O presente trabalho propõe, no contexto de modelos de superpopulação que apresentam o comportamento de caudas pesadas e que podem ser descritos no formato de misturas de escala de distribuições normais, adotar planos amostrais informativos que proporcionem uma maior probabilidade de se selecionarem valores mais extremos daqueles que seriam obtidos por amostragem aleatória simples. A ideia é melhorar a estimação do vetor paramétrico associado as caudas.

Dois planos amostrais foram propostos no presente trabalho. Para a implementação do primeiro, é necessário ter uma variável auxiliar, conhecida para toda a população e que tenha relação com a variável de interesse. No segundo, é necessário ter conhecimento de uma variável indicadora do estrato a que cada unidade amostral pertence.

Os estudos simulados com a distribuição T-Student e o estudo considerando dados reais mostram que os planos amostrais propostos proporcionam uma melhora apreciável nas estimativas do parâmetro graus de liberdade, tanto para inferência sob enfoque clássico quanto sob enfoque bayesiano, sem grande impacto nas estimativas dos demais parâmetros.

Palavras-Chave: modelos de superpopulação; populações finitas; distribuição T-Student; amostragem informativa; inferência bayesiana.

Abstract

In several areas of study there are several data sets that present *outliers*, or even significant concentrations of extreme values in the tails of their respective distributions. In the particular case of the T-Student regression, the degrees of freedom parameter is what determines the degree of robustness. The smaller the number of degrees of freedom, the more robust the tail. However, this parameter is difficult to estimate, since it is necessary to observe extreme values of the distribution.

In the context of finite populations, the researcher usually does not have access to all population elements for modeling. A sampling plan must be adopted to obtain the sample. The present work proposes, in the context of superpopulation models with heavy tails and that can be described in the format of scale mixtures of normal distributions, to adopt informative sampling plans that provide a greater probability of selecting more extreme values than those that would be obtained by simple random sampling. The idea is to improve the estimation of the parametric vector associated with the thickness of the tails.

Three sampling plans were proposed in the present work. For the implementation of the first, it is necessary to have an auxiliary variable, known to the entire population and related to the variable of interest. In the second, it is necessary to have knowledge of an indicator variable of the stratum to which each sample unit belongs.

The simulated studies with the T-Student distribution and the study considering real data show that the sampling plans proposed provide an appreciable improvement in the estimates of the degrees of freedom parameter, both for inference under the classical and the Bayesian approaches, without great impact on the estimates of the other parameters.

Keywords: Keywords: Superpopulation models; finite populations; T-Student distribution; informative sampling; bayesian inference.

Sumário

1	Introdução	1
2	Inferência em populações finitas	4
2.1	Conceitos básicas e terminologia	4
2.2	Modelos de superpopulação	6
2.2.1	Amostragem informativa	6
3	Planos amostrais para modelos de misturas com caudas pesadas	9
3.1	Modelos de misturas com caudas pesadas	10
3.2	Amostragem aleatória simples	11
3.3	Planos amostrais propostos	12
3.3.1	Amostragem quadrática simétrica com variável auxiliar	12
3.3.2	Amostragem estratificada informativa	16
4	Amostragem eficiente num modelo de regressão T-Student	20
5	Estudo de simulação	29
5.1	Primeiro estudo de simulação	30
5.2	Segundo estudo de simulação	34
5.2.1	Comparação entre os modelos AQS, AASX e AAS, e comparação entre os modelos AQSC e AASXC	35
5.2.2	Comparação entre os modelos AEI, AQS e AAS	43
6	Aplicação	47

7	Conclusões e trabalhos futuros	54
7.1	Trabalhos futuros	54
A	Resultados referentes aos planos amostrais apresentados	56
A.1	Resultados referentes aos modelos AASX, AQ, AQS e AQC	56
A.2	Informação de Fisher dos modelos AASX e AQ	58
A.3	Informação de Fisher do plano AQS - aproximação pela Lei dos grandes números	64
A.4	Função de verossimilhança referente ao plano AQS - considerando que \bar{X} não seja uma constante	65
B	Análise de sensibilidade da distribuição <i>a priori</i> do parâmetro graus de liberdade (ν)	67
C	Resultados da aplicação aos dados reais	73

Lista de Tabelas

4.1	Siglas referentes aos modelos considerados para inferência.	21
5.1	Resultados das simulações para os modelos AQ, AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,2$ e $n = 20, 60, 100$ (abordagem clássica).	31
5.2	Resultados das simulações para os modelos AQ, AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,5$ e $n = 20, 60, 100$ (abordagem clássica).	32
5.3	Resultados das simulações para os modelos AQ, AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,8$ e $n = 20, 60, 100$ (abordagem clássica).	32
5.4	Resultados das simulações para os modelos AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,2$ e $n = 20, 60, 100$ (abordagem bayesiana).	37
5.5	Resultados das simulações para os modelos AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,5$ e $n = 20, 60, 100$ (abordagem bayesiana).	38
5.6	Resultados das simulações para os modelos AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,8$ e $n = 20, 60, 100$ (abordagem bayesiana).	39
5.7	Comparação dos modelos AQSC e AASXC com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,8$ e $n = 20$ (abordagem bayesiana).	42
5.8	Tamanhos das amostras n_h dos estratos referente ao plano amostral amostragem estratificada informativa.	43

5.9	Resultados das simulações para os modelos AEI1, AEI2, AQS e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,7$ e $n = 20, 60, 100$ (abordagem bayesiana).	45
6.1	Resultados do ajuste da regressão de T-Student considerando os dados populacionais.	50
B.1	Resultados das simulações para os modelos AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,2$ e $n = 20,60,100$ (abordagem bayesiana).	69
B.2	Resultados das simulações para os modelos AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,5$ e $n = 20,60,100$ (abordagem bayesiana).	70
B.3	Resultados das simulações para os modelos AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,8$ e $n = 20,60,100$ (abordagem bayesiana).	71
C.1	Resultados obtidos para os planos amostrais AEI1, AEI2, AASX e AAS com $n = 20, 60$ e 100 (abordagem bayesiana).	76

Lista de Figuras

1.1	(a) Comparação das curvas da distribuição Normal(0,1) com a distribuição T-Student(0,1) graus de liberdade $\nu = 5, 10, 15, 20, 25$. (b) Comparação das caudas da distribuição Normal(0,1) com a distribuição T-Student(0,1) graus de liberdade $\nu = 5, 10, 15, 20, 25$	2
3.1	(a) Histograma da amostra selecionada considerando a amostragem quadrática com variável auxiliar. (b) Histograma da amostra selecionada considerando a amostragem aleatória simples com variável auxiliar	12
5.1	Boxplots das estimativas de máxima verossimilhança de ν obtidas nas simulações, considerando-se os modelos AQ, AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,2; 0,5; 0,8$ e $n = 20, 60, 100$	33
5.2	Boxplots das estimativas de máxima verossimilhança de μ obtidas nas simulações, considerando-se os modelos AQ, AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,2; 0,5; 0,8$ e $n = 20, 60, 100$	33
5.3	Boxplots das estimativas de máxima verossimilhança de σ obtidas nas simulações, considerando-se os modelos AQ, AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,2; 0,5; 0,8$ e $n = 20, 60, 100$	34
5.4	Boxplots das estimativas bayesianas pontuais de ν obtidas nas simulações, considerando-se os modelos AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,2; 0,5; 0,8$ e $n = 20, 60, 100$	40

5.5	Boxplots das estimativas bayesianas pontuais de μ obtidas nas simulações, considerando-se os modelos AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,2; 0,5; 0,8$ e $n = 20, 60, 100$	40
5.6	Boxplots das estimativas bayesianas pontuais de σ obtidas nas simulações, considerando-se os modelos AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,2; 0,5; 0,8$ e $n = 20, 60, 100$	41
5.7	Boxplots das estimativas bayesianas pontuais de ν , μ e σ obtidas nas simulações, considerando-se os modelos AEI1, AEI2, AQS e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,7$ e $n = 20, 60, 100$	46
6.1	Gráfico de dispersão das variáveis área dedicada às fazendas no ano de 1992 e área dedicada às fazendas no ano de 1987.	48
6.2	Histograma dos resíduos do modelo de regressão linear simples considerando como variável dependente a área dedicada às fazendas no ano de 1992 e covariável a área dedicada às fazendas no ano de 1987.	48
6.3	Gráfico de dispersão das variáveis área dedicada às fazendas no ano de 1992 e área dedicada às fazendas no ano de 1982.	49
6.4	Histograma das áreas dedicadas às fazendas no ano de 1982.	49
6.5	Boxplots referentes as amostras das distribuições <i>a posteriori</i> dos parâmetros ν , β_0 , β_1 e σ , considerando-se os modelos AEI1, AEI2, AASX e AAS e $n = 20,60,100$	51
6.6	Intervalos HPDs de 95% dos parâmetros ν , β_0 , β_1 e σ , considerando-se os modelos AEI1, AEI2, AASX e AAS e $n = 20,60,100$	52
B.1	Boxplots das estimativas bayesianas pontuais de ν obtidas nas simulações, considerando-se os modelos AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,2; 0,5; 0,8$ e $n = 20, 60, 100$	68
C.1	Amostras obtidas das distribuições <i>a posteriori</i> dos parâmetros ν , β_0 , β_1 , β_2 e σ em relação ao ajuste considerando todos os dados populacionais.	73

C.2	Amostras obtidas das distribuições <i>a posteriori</i> dos parâmetros ν , β_0 , β_1 , β_2 e σ em relação ao modelo AEI1.	74
C.3	Amostras obtidas das distribuições <i>a posteriori</i> dos parâmetros ν , β_0 , β_1 , β_2 e σ em relação ao modelo AEI2.	74
C.4	Amostras obtidas das distribuições <i>a posteriori</i> dos parâmetros ν , β_0 , β_1 , β_2 e σ em relação ao modelo AASX.	75
C.5	Amostras obtidas das distribuições <i>a posteriori</i> dos parâmetros ν , β_0 , β_1 , β_2 e σ em relação ao modelo AAS.	75

Capítulo 1

Introdução

Em diversas áreas de estudo é comum que conjuntos de dados contínuos apresentem *outliers* ou frequências altas de valores mais extremos. Por exemplo, na área financeira, [Abanto-Valle et al. \(2010\)](#) argumentam que alguns retornos de ativos têm distribuições com caudas pesadas. Neste caso, geralmente a distribuição T-Student é aplicada para modelar os erros ao invés de se assumir normalidade. Os autores propuseram ainda a utilização de uma classe de distribuições de misturas na escala normal para modelar esses erros. Alguns exemplos de distribuições presentes nesta classe são: T-Student, *Slash* e *Variance Gama*. Aplicações em outras áreas podem ser citadas em que as distribuições dos dados apresentam caudas pesadas: educação ([Fonseca et al., 2008](#)), hidrologia ([Anderson e Meerschaert, 1998](#)), meio ambiente ([Gómez et al., 2019](#)), saúde ([Gardiner et al., 2014](#)), entre outras.

No contexto de populações finitas, sob a abordagem dos modelos de superpopulação, [Kim et al. \(1996\)](#) propuseram a utilização de uma distribuição *a priori* com caudas pesadas usando modelos de misturas de escala normal, em particular a distribuição T-Student, para a obtenção de um estimador bayesiano para a média de um modelo de superpopulação com caudas pesadas. [Lee et al. \(2006\)](#) também propuseram a utilização da distribuição *a priori* com caudas pesadas utilizando uma variável auxiliar conhecida para todos os elementos populacionais.

A modelagem de dados contínuos que apresentam *outliers* ou frequências significativamente altas de valores mais extremos e a estimação eficiente dos parâmetros envolvidos

é a principal motivação do presente trabalho. Especificamente, o foco será em modelos de superpopulação. Num modelo de superpopulação, assume-se que o vetor de observações da variável resposta na população é gerado de acordo com um certo modelo proposto, para maiores detalhes, ver por exemplo [Bolfarine e Zacks \(2012\)](#). Em particular, no contexto de populações finitas, o pesquisador não tem acesso a todos os elementos populacionais. Um plano amostral deve ser adotado para obtenção da amostra.

Existe uma dificuldade natural na estimação dos parâmetros associados às caudas em modelos de caudas pesadas. A título de ilustração, a Figura 1.1 mostra a comparação das caudas da função de densidade de probabilidade de uma distribuição normal padrão com uma distribuição T-Student com diferentes graus de liberdade.

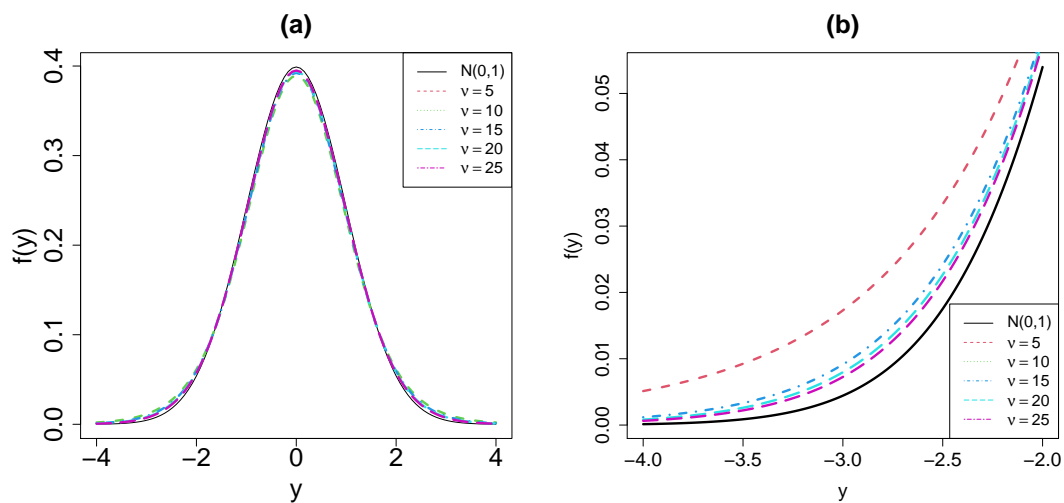


Figura 1.1: (a) Comparação das curvas da distribuição Normal(0,1) com a distribuição T-Student(0,1) graus de liberdade $\nu = 5,10,15,20,25$. (b) Comparação das caudas da distribuição Normal(0,1) com a distribuição T-Student(0,1) graus de liberdade $\nu = 5,10,15,20,25$

É importante notar que os graus de liberdade estão associados às caudas da distribuição. Quanto menores, mais pesadas são as caudas. Observe que as curvas são muito próximas, com exceção das caudas, região de relativa baixa massa de probabilidade. Por este motivo, muitas vezes a amostra obtida pelo pesquisador não contempla dados presentes nas caudas e assim a qualidade das estimativas obtidas destes parâmetros pode ficar

comprometida.

O presente trabalho propõe o emprego de planos amostrais que permitem amostrar valores mais extremos da variável resposta, presentes nas caudas, visando melhorar a estimação dos parâmetros relacionados com as caudas. Espera-se que a estimação dos demais parâmetros do modelo não seja prejudicada com a introdução desta proposta.

Dois planos amostrais foram propostos no presente trabalho. Para a implementação do primeiro, é necessário se ter uma variável auxiliar, X , conhecida para toda população e que seja relacionada com a variável resposta Y . Por meio desta variável auxiliar será possível construir um plano amostral em que a probabilidade de seleção de cada unidade populacional seja uma função de X , e portanto, relacionada com Y . A escolha propícia dessa função permite com que valores mais extremos de Y tenham maior probabilidade de fazerem parte da amostra. No segundo, é necessário ter conhecimento de uma variável indicadora do estrato a que cada unidade amostral pertence. A ideia é selecionar mais valores extremos da distribuição através da estratificação.

O trabalho está organizado em sete capítulos. No Capítulo 2 uma breve revisão sobre inferência em populações finitas é apresentada. Os planos amostrais para os modelos de misturas com caudas pesadas propostos são descritos no Capítulo 3. No Capítulo 4 são descritos os planos amostrais apresentados no Capítulo 3 para o caso particular do modelo de regressão T-Student. O Capítulo 5 apresenta resultados do estudo de simulação para os modelos propostos. No Capítulo 6 uma aplicação será apresentada. Por fim, no Capítulo 7 são apresentadas as conclusões finais e possíveis trabalhos futuros.

Capítulo 2

Inferência em populações finitas

Neste capítulo, apresentam-se os conceitos e definições importantes referentes a inferência em populações finitas. No contexto de inferência em populações finitas destacam-se duas abordagens: (i) aleatorização - baseada somente no plano amostral, onde a população é considerada fixa, ver [Bolfarine \(2005\)](#); e (ii) modelo de superpopulação - supõe-se que os valores das variáveis observadas na população são realizações de variáveis aleatórias de acordo com um certo modelo proposto, ver [Ericson \(1969\)](#); [Valliant et al. \(2000\)](#); [Bolfarine e Zacks \(2012\)](#).

2.1 Conceitos básicas e terminologia

Define-se população finita como o conjunto de N unidades com pelo menos uma característica em comum, que pode ser representada pelo conjunto de rótulos $U = \{1, 2, \dots, N\}$ ([Cassel et al., 1977](#)). Tal característica determina se uma unidade pertence ou não à população alvo. Associado às unidades $i = 1, 2, \dots, N$, tem-se o vetor $\mathbf{y} = [y_1, y_2, \dots, y_N]'$, que denota os valores da característica de interesse (variável resposta do modelo).

Em geral, deseja-se estimar funções relacionadas ao vetor \mathbf{y} , como por exemplo o total populacional $T = \sum_{i \in U} y_i$; a média populacional $\bar{Y} = N^{-1} \sum_{i \in U} y_i$; ou a variância populacional $\sigma^2 = N^{-1} \sum_{i \in U} (y_i - \bar{Y})^2$. Entretanto, seja por dificuldades logísticas ou pelos custos elevados, geralmente não se dispõe de todos os valores do vetor \mathbf{y} . Com o intuito de serem estimadas as funções desejadas, seleciona-se da população alvo um

subconjunto de elementos de tamanho $n < N$, denominado amostra.

No contexto de populações finitas, uma das abordagens para se selecionar a amostra e depois estimar as funções de interesse é considerar a aleatorização baseada no plano amostral, sendo que nesse caso, o vetor \mathbf{y} é considerado fixo. Nesta abordagem, o mecanismo probabilístico de seleção amostral define um procedimento predeterminado de aleatorização, denominado desenho amostral.

Considere s uma amostra ordenada, de tamanho n , dada por $s = \{i_1, i_2, \dots, i_n\}$. Denomina-se plano amostral (ou desenho amostral) uma função de probabilidade $P(s) = P(S = s)$, definida para todo $s \in F$, onde F é o conjunto das partes de U , ou seja, o conjunto de todas as possíveis amostras s de U , que satisfaz as seguintes propriedades: $P(s) \geq 0, \forall s \in F; \sum_{s \in F} P(s) = 1$. Nesta abordagem, qualquer amostra $s \in F$ é uma realização da variável aleatória S . A probabilidade da unidade $k \in U$ ser selecionada (π_k) é definida como o somatório das probabilidades de todas as amostras s que contêm o elemento k .

Depois da obtenção da amostra s , o objetivo principal agora é obter estimadores adequados das funções relacionadas ao vetor \mathbf{y} . Dentre o conjunto de todos os estimadores possíveis, destaca-se a classe dos estimadores lineares. Um estimador é dito linear se ele puder ser escrito da seguinte forma:

$$\hat{\theta} = \sum_{i \in s} \omega_i y_i.$$

No caso da amostragem aleatória simples, dois casos particulares de estimadores lineares para a média e o total populacional são obtidos considerando os pesos $\omega_i = n^{-1}, \forall i \in s$ e $\omega_i = N/n, \forall i \in s$, respectivamente.

Horvitz e Thompson (1952) propuseram estimadores lineares não viciados para o total e média para qualquer plano amostral sem reposição que seja probabilístico. O estimador de Horvitz-Thompson para o total populacional é:

$$\hat{T}_{HT} = \sum_{i \in s} \frac{1}{\pi_i} y_i = \sum_{i \in U} \frac{1}{\pi_i} y_i I_i,$$

onde π_i representa a probabilidade de inclusão das unidades i e I_i é uma variável indicadora, tal que $I_i = 1$ se $i \in S$ e $I_i = 0$ se $i \notin S$. A variância do estimador é apresentada

abaixo:

$$V\left(\hat{T}_{HT}\right) = V\left(\sum_{i \in U} \frac{1}{\pi_i} y_i I_i\right) = \sum_{i \in U} \left(\frac{1 - \pi_i}{\pi_i} y_i^2\right) + 2 \sum_{i \in U} \sum_{i < j} \left(\frac{\pi_{ij} - \pi_i \pi_j}{\pi_i \pi_j} y_i y_j\right),$$

uma vez que $V(I_i) = \pi_i(1 - \pi_i)$ e $cov(I_i, I_j) = \pi_{ij} - \pi_i \pi_j$; onde π_{ij} representa a probabilidade de inclusão das unidades i e j conjuntamente na amostra. Destaca-se que o π_i depende do plano amostral adotado.

2.2 Modelos de superpopulação

Uma segunda abordagem de inferência para populações finitas é baseada nos modelos de superpopulação. Diferente da abordagem anterior, em que o vetor \mathbf{y} é considerado fixo e que a inferência é realizada através da aleatorização do plano amostral, considera-se que o vetor \mathbf{y} é uma realização do vetor de variáveis aleatórias $\mathbf{Y} = [Y_1, Y_2, \dots, Y_N]'$.

O modelo de superpopulação é definido por uma função de densidade ou probabilidade conjunta, $f(\mathbf{y}|\boldsymbol{\theta}, \mathbf{z})$, que pode depender de um vetor desconhecido de parâmetros $\boldsymbol{\theta} \in \Theta$ e de covariáveis $\mathbf{z} = [z_1, \dots, z_N]$, onde $\mathbf{z}_i = (z_{i1}, \dots, z_{ip})'$; e de variáveis adicionais associadas à estrutura da população $\mathbf{X} = [\mathbf{X}_1, \dots, \mathbf{X}_N]$ de dimensão $N \times K$, onde $\mathbf{X}_i = (X_{i1}, \dots, X_{ik})'$, que podem depender de um vetor paramétrico $\boldsymbol{\phi} \in \Phi$ e de covariáveis $\mathbf{m} = [\mathbf{m}_1, \dots, \mathbf{m}_N]$, onde $\mathbf{m}_i = (m_{i1}, \dots, m_{il})'$. A distribuição conjunta das variáveis adicionais será denotada por $f(\mathbf{x}|\boldsymbol{\phi}, \mathbf{m})$.

2.2.1 Amostragem informativa

Um plano amostral é dito informativo quando o mecanismo de seleção amostral está relacionado às variáveis de interesse na população, ou seja, existe relação das variáveis resposta Y com as probabilidades de inclusão das unidades amostrais, mesmo após o condicionamento das covariáveis do modelo. Nesta situação, deve-se levar em conta o plano amostral na inferência dos parâmetros populacionais desejados.

Considere o vetor indicador de pertinência $\mathbf{I} = (I_1, \dots, I_N)$, tal que $I_i = 1$ se $i \in S$ e $I_i = 0$ se $i \notin S$. Um plano amostral é um sistema de regras que define uma distribuição conjunta do vetor \mathbf{I} . Este sistema pode depender de variáveis adicionais \mathbf{X} , da variável

resposta \mathbf{Y} e de um vetor paramétrico $\boldsymbol{\phi} \in \Phi$. A distribuição conjunta de $\mathbf{Y}, \mathbf{X}, \mathbf{I}$ pode ser escrita como:

$$f(\mathbf{y}, \mathbf{x}, \mathbf{i} | \boldsymbol{\theta}, \boldsymbol{\phi}, \mathbf{z}, \mathbf{m}) = f(\mathbf{i} | \mathbf{y}, \mathbf{x}) f(\mathbf{x} | \mathbf{y}, \boldsymbol{\phi}, \mathbf{m}) f(\mathbf{y} | \boldsymbol{\theta}, \mathbf{z}). \quad (2.1)$$

onde $f(\mathbf{i} | \mathbf{y}, \mathbf{x})$ é a função de probabilidade do vetor \mathbf{I} . Esta expressão será útil quando $\mathbf{X} = (X_1, X_2, \dots, X_N)'$ for conhecido e tiver uma relação com \mathbf{Y} . Um caso particular que será utilizado no presente trabalho é o modelo linear $X_i = b_0 + b_1 Y_i + \epsilon_i$, com $\epsilon_i \sim N(0, \sigma_\epsilon^2)$ independentes e identicamente distribuídos $\forall i \in U$. Observe que $\boldsymbol{\phi} = (b_0, b_1, \sigma_\epsilon)$ e $X_i | Y_i, b_0, b_1, \sigma_\epsilon \sim N(b_0 + b_1 y_i, \sigma_\epsilon^2); \forall i \in U$.

Note que a expressão (2.1) é válida para todos os valores da população. Na prática são apenas conhecidos os valores obtidos na amostra. A informação conhecida é oriunda de $(\mathbf{Y}_s, \mathbf{X}, \mathbf{I})$. Integrando a expressão (2.1) em relação aos valores não observados na amostra, tem-se a distribuição dos dados observados:

$$f(\mathbf{y}_s, \mathbf{x}, \mathbf{i} | \boldsymbol{\theta}, \boldsymbol{\phi}, \mathbf{z}, \mathbf{m}) = \int f(\mathbf{i} | \mathbf{y}, \mathbf{x}) f(\mathbf{x} | \mathbf{y}, \boldsymbol{\phi}, \mathbf{m}) f(\mathbf{y} | \boldsymbol{\theta}, \mathbf{z}) d\mathbf{y}_{\bar{s}}. \quad (2.2)$$

O caso discreto é análogo aos casos anteriores, basta substituir a integral pelo somatório. Os resultados posteriores terão como foco variáveis aleatórias contínuas.

No contexto bayesiano, o foco está na obtenção da distribuição *a posteriori* dos parâmetros $(\boldsymbol{\theta}, \boldsymbol{\phi})$, que pode ser obtida por através da expressão abaixo:

$$f(\boldsymbol{\theta}, \boldsymbol{\phi} | \mathbf{y}_s, \mathbf{x}, \mathbf{i}, \mathbf{z}, \mathbf{m}) \propto \pi(\boldsymbol{\theta}, \boldsymbol{\phi}) f(\mathbf{y}_s, \mathbf{x}, \mathbf{i} | \boldsymbol{\theta}, \boldsymbol{\phi}, \mathbf{z}, \mathbf{m}); \quad (2.3)$$

em que $\pi(\boldsymbol{\theta}, \boldsymbol{\phi})$ é a distribuição *a priori* conjunta do vetor paramétrico $\boldsymbol{\theta}$ e $\boldsymbol{\phi}$. A distribuição *a priori* pode ser escrita como $\pi(\boldsymbol{\theta}, \boldsymbol{\phi}) = \pi(\boldsymbol{\theta}) \pi(\boldsymbol{\phi} | \boldsymbol{\theta})$. Geralmente o foco principal está na distribuição *a posteriori* de $\boldsymbol{\theta}$. Integrando a expressão (2.3), tem-se:

$$f(\boldsymbol{\theta} | \mathbf{y}_s, \mathbf{x}, \mathbf{i}, \mathbf{z}, \mathbf{m}) \propto \int f(\boldsymbol{\theta}, \boldsymbol{\phi} | \mathbf{y}_s, \mathbf{x}, \mathbf{i}, \mathbf{z}, \mathbf{m}) d\boldsymbol{\phi}. \quad (2.4)$$

Caso o mecanismo de seleção de amostra seja ignorado na inferência, a distribuição *a posteriori* de $\boldsymbol{\theta}$ será dada por:

$$f(\boldsymbol{\theta} | \mathbf{y}_s, \mathbf{x}, \mathbf{z}, \mathbf{m}) \propto \int f(\boldsymbol{\theta}, \boldsymbol{\phi} | \mathbf{y}_s, \mathbf{x}, \mathbf{z}, \mathbf{m}) d\boldsymbol{\phi}. \quad (2.5)$$

Sob a ótica bayesiana, quando as equações (2.4) e (2.5) são iguais, diz-se que o desenho amostral é ignorável. Para garantir a ignorabilidade do plano amostral, duas condições são necessárias. A primeira condição é $f(\mathbf{i}|\mathbf{y},\mathbf{x}) = f(\mathbf{i}|\mathbf{y}_s,\mathbf{x})$; e a segunda é $\pi(\boldsymbol{\phi}|\boldsymbol{\theta}) = \pi(\boldsymbol{\phi})$. Como consequência das condições anteriores, tem-se que se o desenho amostral é ignorável com respeito ao vetor paramétrico $\boldsymbol{\theta}$, se

$$f(\boldsymbol{\theta}|\mathbf{y}_s,\mathbf{x},\mathbf{i}) = f(\boldsymbol{\theta}|\mathbf{y}_s,\mathbf{x}), \quad (2.6)$$

Na amostragem de populações finitas geralmente deseja-se estimar funções dos valores de \mathbf{y} na população. No presente estudo o foco principal da inferência será nos parâmetros dos modelos de superpopulação.

Capítulo 3

Planos amostrais para modelos de misturas com caudas pesadas

Neste capítulo serão apresentados os planos amostrais para modelos de misturas com caudas pesadas que serão considerados no presente estudo. Serão descritos os planos propostos e planos que são comumente aplicados na prática.

No processo de obtenção da amostra, dois mecanismos de seleção podem ser considerados: a amostragem com reposição e a amostragem sem reposição (Bolfarine, 2005). No primeiro, ao se sortear um elemento populacional, ele novamente pode ser sorteado. No segundo, um elemento sorteado é removido antes de se realizar o próximo sorteio.

A amostragem com reposição conduz a um tratamento teórico mais simples, pois ela implica a independência entre as unidades selecionadas. Além disso, na maioria das aplicações as frações de amostragem são desprezíveis, e a probabilidade de selecionar o mesmo elemento mais de uma vez é muito baixa. Por estas razões, todos os planos amostrais considerados no presente trabalho serão com reposição.

Na Seção 3.1 é apresentado um breve resumo sobre misturas de escala de distribuições normais; na Seção 3.2 é apresentado o plano amostragem aleatória simples com ou sem variável auxiliar; na Seção 3.3 serão apresentados os planos amostrais propostos no presente trabalho.

3.1 Modelos de misturas com caudas pesadas

Misturas de escala de distribuições normais são originadas pela mistura de uma variável aleatória normal R com uma variável aleatória de escala não negativa λ ,

$$Y = \mu + \kappa(\lambda)^{1/2} R, \quad (3.1)$$

onde μ é um parâmetro de localização, $\kappa(\cdot)$ é uma função positiva, λ é uma variável aleatória de mistura positiva com função de densidade probabilidade $h(\lambda|\boldsymbol{\nu})$, independente de $R \sim N(0, \sigma^2)$ e $\boldsymbol{\nu}$ é um escalar ou um vetor paramétrico. Este vetor está relacionado às caudas. No presente trabalho, considerou-se somente casos em que ν é um escalar.

Assim como [Lange et al. \(1989\)](#), [Boris Choy e Chan \(2008\)](#) e [Abanto-Valle et al. \(2010\)](#), considerou-se o caso em que $\kappa(\lambda) = 1/\lambda$. De (3.1), tem-se que a distribuição condicional $Y|\lambda, \mu, \sigma$ será normal com média μ e variância $\lambda^{-1}\sigma^2$.

Integrando $f(y, \lambda|\mu, \sigma, \nu)$ em λ , tem-se a densidade condicional $Y|\mu, \sigma, \nu$:

$$f(y|\mu, \sigma, \nu) = \int_0^{\infty} N(y|\mu, \lambda^{-1}\sigma^2) h(\lambda|\nu) d\lambda; \quad (3.2)$$

onde $N(y|\mu, \lambda^{-1}\sigma^2)$ é função de densidade de probabilidade da distribuição normal com média μ e variância $\lambda^{-1}\sigma^2$. Pode-se considerar então os seguintes casos particulares para a distribuição $Y|\mu, \sigma, \nu$: T-student, quando $h(\lambda|\nu) = g(\lambda|\frac{\nu}{2}, \frac{\nu}{2})$; *Slash*, quando $h(\lambda|\nu) = Be(\lambda|\nu, 1)$; *Variance Gama* quando $h(\lambda|\nu) = Ig(\lambda|\frac{\nu}{2}, \frac{\nu}{2})$. Vale destacar que nestes casos as distribuições são unimodais. Pode-se escrever os modelos acima, respectivamente, de forma hierárquica:

$$\begin{aligned} Y|\mu, \sigma, \nu &\sim N(\mu, \lambda^{-1}\sigma^2); & \lambda|\nu &\sim g\left(\frac{\nu}{2}, \frac{\nu}{2}\right); \\ Y|\mu, \sigma, \nu &\sim N(\mu, \lambda^{-1}\sigma^2); & \lambda|\nu &\sim Be(\nu, 1); \\ Y|\mu, \sigma, \nu &\sim N(\mu, \lambda^{-1}\sigma^2); & \lambda|\nu &\sim Ig\left(\frac{\nu}{2}, \frac{\nu}{2}\right). \end{aligned}$$

No decorrer do trabalho, em relação às variáveis $Y_i, i = 1, 2, \dots, N$, considerou-se que $E(Y_i|\boldsymbol{\theta}, \mathbf{z}_i) = \mu_i = \mathbf{z}_i' \boldsymbol{\beta}, \forall i \in U$, com $\boldsymbol{\beta} = [\beta_0, \beta_1, \dots, \beta_k]'$ e $\mathbf{z}_i = [1, z_{i1}, \dots, z_{ik}]'$.

3.2 Amostragem aleatória simples

O primeiro plano amostral que será apresentado é a *amostragem aleatória simples*. Os pressupostos deste plano são que exista uma lista com todos elementos populacionais identificados e que ao sortear a amostra todos têm a mesma probabilidade de pertencerem a amostra (Singh, 2003).

Para uma amostra aleatória simples com reposição, a função de densidade de probabilidade dos valores amostrados \mathbf{y}_s é dada por:

$$f(\mathbf{y}_s|\boldsymbol{\theta}, \mathbf{z}) = \prod_{i \in s} f_p(y_i|\boldsymbol{\theta}, \mathbf{z}_i); \quad (3.3)$$

onde $f_p(y_i|\boldsymbol{\theta}, \mathbf{z}_i)$ é a função de densidade de probabilidade da variável Y_i .

Considere o cenário de dados que apresentem o comportamento de caudas pesadas e que dependam de um ou mais parâmetros que estão diretamente relacionados às caudas. Considere também o seguinte modelo linear entre as variáveis de interesse Y_i , $i = 1, 2, \dots, N$, e as variáveis auxiliares X_i , $i = 1, 2, \dots, N$:

$$X_i = b_0 + b_1 Y_i + \epsilon_i; \epsilon_i \sim N(0, \sigma_\epsilon^2), i = 1, 2, 3, \dots, N; \quad (3.4)$$

ou, equivalente,

$$X_i|Y_i = y_i, b_0, b_1, \sigma_\epsilon \sim N(b_0 + b_1 y_i, \sigma_\epsilon^2), i = 1, 2, 3, \dots, N.$$

Considerando o modelo linear em (3.4), a função de densidade de probabilidade dos valores amostrados \mathbf{y}_s e \mathbf{x}_s para uma amostra aleatória simples com reposição é dada por:

$$f(\mathbf{y}_s, \mathbf{x}_s|\boldsymbol{\theta}, b_0, b_1, \sigma_\epsilon, \mathbf{z}) = \prod_{i \in s} [f_p(y_i|\boldsymbol{\theta}, \mathbf{z}_i) f_p(x_i|y_i, b_0, b_1, \sigma_\epsilon)]. \quad (3.5)$$

Vale ressaltar que a expressão anterior só leva em consideração \mathbf{y}_s e \mathbf{x}_s , desconsiderando $\mathbf{x}_{\bar{s}}$.

Para fins de distinção, no decorrer do texto considerou-se o termo *amostragem aleatória simples com variável auxiliar* quando para inferência for considerada a função de densidade de probabilidade conjunta dos valores amostrados \mathbf{y}_s e \mathbf{x}_s .

3.3 Planos amostrais propostos

3.3.1 Amostragem quadrática simétrica com variável auxiliar

Admitindo novamente o modelo linear em (3.4), agora considera-se um plano amostral que proporcione uma maior probabilidade de se selecionar valores extremos das duas caudas de uma distribuição simétrica. A ideia é melhorar a estimação do parâmetro ν associado as caudas. Para se atingir este objetivo, adota-se então um plano amostral cuja probabilidade de seleção é $\pi_i \propto (x_i - c)^2$, onde c é uma contante, chamada constante de centralização. Este plano será chamado de *amostragem quadrática simétrica com variável auxiliar*. Adotando-se este plano, espera-se que as estimativas dos demais parâmetros não sejam afetadas e que haja uma melhora significativa na estimação de ν .

Primeiro considerou-se o caso particular em que $c = 0$. Recomenda-se aceitar essa suposição quando a distribuição de X for simétrica e tiver média próxima de zero. Este caso será chamado de *amostragem quadrática com variável auxiliar*.

Para fins de ilustração, simulou-se uma população de tamanho $N = 5000$ da distribuição T-Student com parâmetros $\mu = 0$, $\sigma = 1$ e $\nu = 5$. Considerou-se o modelo em (3.4), com $b_0 = 0$, $b_1 = 1$ e $\rho_{XY} = 0.5$. Duas amostras cada uma de tamanho $n = 100$ foram selecionadas, uma adotando o plano de amostragem aleatória simples com variável auxiliar e a outra o presente plano proposto. Os histogramas das amostras são apresentados abaixo:

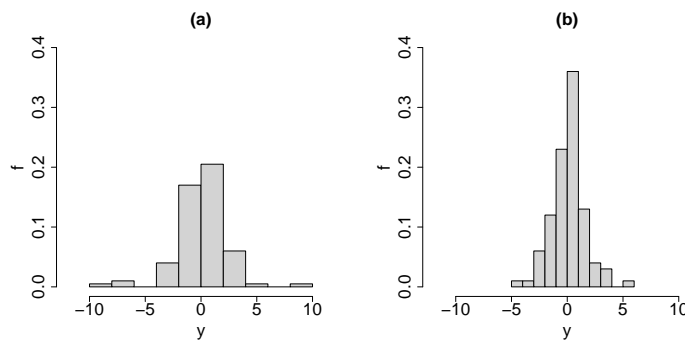


Figura 3.1: (a) Histograma da amostra selecionada considerando a amostragem quadrática com variável auxiliar. (b) Histograma da amostra selecionada considerando a amostragem aleatória simples com variável auxiliar

Analisando a Figura 3.1, nota-se que foram selecionados valores extremos da população ao adotar-se a amostragem quadrática com variável auxiliar.

A distribuição marginal de $[Y_i, X_i]$, $i \in S$, do plano proposto (caso $c = 0$) pode ser obtida adaptando-se a metodologia descrita em [Pfeffermann et al. \(1998\)](#):

$$\begin{aligned}
f_s(y_i, x_i | \boldsymbol{\theta}, b_0, b_1, \sigma_\epsilon, \mathbf{z}_i) &= f(y_i, x_i | \boldsymbol{\theta}, b_0, b_1, \sigma_\epsilon, \mathbf{z}_i, I_i = 1) \\
&= \frac{P(I_i = 1 | y_i, x_i, \boldsymbol{\theta}, b_0, b_1, \sigma_\epsilon, \mathbf{z}_i) f_p(y_i, x_i | \boldsymbol{\theta}, b_0, b_1, \sigma_\epsilon, \mathbf{z}_i)}{P(I_i = 1 | \boldsymbol{\theta}, b_0, b_1, \sigma_\epsilon, \mathbf{z}_i)} \\
&= \frac{f_p(y_i | \boldsymbol{\theta}, \mathbf{z}_i) f_p(x_i | y_i, b_0, b_1, \sigma_\epsilon) P(I_i = 1 | x_i, y_i, \boldsymbol{\theta}, b_0, b_1, \sigma_\epsilon, \mathbf{z}_i)}{\int \int P(I_i = 1, x_i, y_i | \boldsymbol{\theta}, \mathbf{z}_i, b_0, b_1, \sigma_\epsilon) dx_i dy_i} \\
&= \frac{f_p(y_i | \boldsymbol{\theta}, \mathbf{z}_i) f_p(x_i | y_i, b_0, b_1, \sigma_\epsilon) P(I_i = 1 | x_i, y_i, \boldsymbol{\theta}, b_0, b_1, \sigma_\epsilon, \mathbf{z}_i)}{\int \int f_p(y_i | \boldsymbol{\theta}, \mathbf{z}_i) f_p(x_i | y_i, b_0, b_1, \sigma_\epsilon) P(I_i = 1 | x_i, y_i, \boldsymbol{\theta}, b_0, b_1, \sigma_\epsilon, \mathbf{z}_i) dx_i dy_i} \\
&= \frac{f_p(y_i | \boldsymbol{\theta}, \mathbf{z}_i) f_p(x_i | y_i, b_0, b_1, \sigma_\epsilon) E_p(\pi_i | x_i, y_i, \boldsymbol{\theta}, b_0, b_1, \sigma_\epsilon, \mathbf{z}_i)}{\int \int f_p(y_i | \boldsymbol{\theta}, \mathbf{z}_i) f_p(x_i | y_i, b_0, b_1, \sigma_\epsilon) E_p(\pi_i | x_i, y_i, \boldsymbol{\theta}, b_0, b_1, \sigma_\epsilon, \mathbf{z}_i) dx_i dy_i}.
\end{aligned}$$

A última equação se deve ao fato de que:

$$\begin{aligned}
P(I_i = 1 | x_i, y_i, \boldsymbol{\theta}, b_0, b_1, \sigma_\epsilon, \mathbf{z}_i) &= \int P(I_i = 1 | x_i, y_i, \pi_i, \boldsymbol{\theta}, b_0, b_1, \sigma_\epsilon, \mathbf{z}_i) f_p(\pi_i | x_i, y_i) d\pi_i \\
&= E_p(\pi_i | x_i, y_i, \boldsymbol{\theta}, b_0, b_1, \sigma_\epsilon, \mathbf{z}_i),
\end{aligned}$$

já que $P(I_i = 1 | x_i, y_i, \pi_i, b_0, b_1, \sigma_\epsilon, \mathbf{z}_i) = \pi_i$, onde π_i é a probabilidade de seleção da unidade i .

Uma vez que $\pi_i \propto x_i^2$, tem-se:

$$E_p(\pi_i | x_i, y_i, \boldsymbol{\theta}, b_0, b_1, \sigma_\epsilon, \mathbf{z}_i) \propto E_p(x_i^2 | x_i, y_i, \boldsymbol{\theta}, b_0, b_1, \sigma_\epsilon, \mathbf{z}_i) = x_i^2.$$

Logo:

$$\begin{aligned}
f_s(y_i, x_i | \boldsymbol{\theta}, b_0, b_1, \sigma_\epsilon, \mathbf{z}_i) &= \frac{f_p(y_i | \boldsymbol{\theta}, \mathbf{z}_i) f_p(x_i | y_i, b_0, b_1, \sigma_\epsilon) x_i^2}{\int f_p(y_i | \boldsymbol{\theta}, \mathbf{z}_i) \int x_i^2 f_p(x_i | y_i, b_0, b_1, \sigma_\epsilon) dx_i dy_i} \\
&= \frac{f_p(y_i | \boldsymbol{\theta}, \mathbf{z}_i) f_p(x_i | y_i, b_0, b_1, \sigma_\epsilon) x_i^2}{\int [(b_0 + b_1 y_i)^2 + \sigma_\epsilon^2] f_p(y_i | \boldsymbol{\theta}, \mathbf{z}_i) dy_i} \\
&= \frac{f_p(y_i | \boldsymbol{\theta}, \mathbf{z}_i) f_p(x_i | y_i, b_0, b_1, \sigma_\epsilon) x_i^2}{b_1^2 E_p(Y_i^2 | \boldsymbol{\theta}, \mathbf{z}_i) + 2b_0 b_1 E_p(Y_i | \boldsymbol{\theta}, \mathbf{z}_i) + b_0^2 + \sigma_\epsilon^2} \\
&= \frac{f_p(y_i | \boldsymbol{\theta}, \mathbf{z}_i) f_p(x_i | y_i, b_0, b_1, \sigma_\epsilon) x_i^2}{[b_1 E_p(Y_i | \boldsymbol{\theta}, \mathbf{z}_i) + b_0]^2 + b_1^2 V_p(Y_i | \boldsymbol{\theta}, \mathbf{z}_i) + \sigma_\epsilon^2}. \quad (3.6)
\end{aligned}$$

A função de densidade de probabilidade (3.6) é válida para qualquer distribuição de Y , desde que exista $V_p(Y_i | \boldsymbol{\theta}, \mathbf{z}_i)$ e que se suponha o modelo (3.4). Note que a escolha $\pi_i \propto x_i^2$

possibilitou um tratamento analítico mais fácil para a obtenção da expressão (3.6) e a garantia de que as probabilidades de inclusão sejam sempre positivas. Além disso, a única exigência do plano amostral é a existência do momento de segunda ordem de Y . Portanto, pelos motivos citados, justifica-se a escolha da função quadrática.

A função de densidade de probabilidade conjunta dos valores amostrados \mathbf{y}_s e \mathbf{x}_s é dada por:

$$f(\mathbf{y}_s, \mathbf{x}_s | \boldsymbol{\theta}, b_0, b_1, \sigma_\epsilon, \mathbf{z}) = \prod_{i \in s} \left\{ \frac{f_p(y_i | \boldsymbol{\theta}, \mathbf{z}_i) f_p(x_i | y_i, b_0, b_1, \sigma_\epsilon) x_i^2}{[b_1 E_p(Y_i | \boldsymbol{\theta}, \mathbf{z}_i) + b_0]^2 + b_1^2 V_p(Y_i | \boldsymbol{\theta}, \mathbf{z}_i) + \sigma_\epsilon^2} \right\}. \quad (3.7)$$

Agora, considere a distribuição conjunta de $\mathbf{Y}, \mathbf{X}, \mathbf{I}$ do plano amostragem quadrática com variável auxiliar:

$$f(\mathbf{y}, \mathbf{x}, \mathbf{i} | \boldsymbol{\theta}, \boldsymbol{\phi}, \mathbf{z}) = f(\mathbf{i} | \mathbf{y}, \mathbf{x}) f(\mathbf{x} | \mathbf{y}, \boldsymbol{\phi}) f(\mathbf{y} | \boldsymbol{\theta}, \mathbf{z}), \quad (3.8)$$

onde $\boldsymbol{\phi} = [b_0, b_1, \sigma_\epsilon]'$. Integrando (3.8) em relação a $\mathbf{y}_{\bar{s}}$, tem-se que:

$$\begin{aligned} f(\mathbf{y}_s, \mathbf{x}, \mathbf{i} | \boldsymbol{\theta}, \boldsymbol{\phi}, \mathbf{z}) &= \prod_{i \in s} f_p(y_i | \boldsymbol{\theta}, \mathbf{z}_i) \prod_{i \in s} N(x_i | b_0 + b_1 y_i, \sigma_\epsilon^2) \\ &\times \int f(\mathbf{i} | \mathbf{y}, \mathbf{x}) \prod_{i \notin s} [N(x_i | b_0 + b_1 y_i, \sigma_\epsilon^2) f_p(y_i | \boldsymbol{\theta}, \mathbf{z}_i)] d\mathbf{y}_{\bar{s}}. \end{aligned}$$

Reescrevendo a função de densidade de probabilidade $f_p(y_i | \boldsymbol{\theta}, \mathbf{z}_i)$ como uma mistura de escala de distribuições normais, tem-se:

$$\begin{aligned} f(\mathbf{y}_s, \mathbf{x}, \mathbf{i} | \boldsymbol{\theta}, \boldsymbol{\phi}, \mathbf{z}) &= \int \prod_{i \in s} [h(\lambda_i | \boldsymbol{\nu}) N(y_i | \mathbf{z}'_i \boldsymbol{\beta}, \lambda_i^{-1} \sigma^2)] d\boldsymbol{\lambda}_s \prod_{i \in s} N(x_i | b_0 + b_1 y_i, \sigma_\epsilon^2) \\ &\times \int f(\mathbf{i} | \mathbf{y}, \mathbf{x}) \prod_{i \notin s} N(x_i | b_0 + b_1 y_i, \sigma_\epsilon^2) \int \prod_{i \notin s} [h(\lambda_i | \boldsymbol{\nu}) N(y_i | \mathbf{z}'_i \boldsymbol{\beta}, \lambda_i^{-1} \sigma^2)] d\boldsymbol{\lambda}_{\bar{s}} d\mathbf{y}_{\bar{s}}. \end{aligned}$$

Trocando a ordem de integração, tem-se:

$$\begin{aligned} f(\mathbf{y}_s, \mathbf{x}, \mathbf{i} | \boldsymbol{\theta}, \boldsymbol{\phi}, \mathbf{z}) &= \int \prod_{i \in s} [h(\lambda_i | \boldsymbol{\nu}) N(y_i | \mathbf{z}'_i \boldsymbol{\beta}, \lambda_i^{-1} \sigma^2)] d\boldsymbol{\lambda}_s \prod_{i \in s} N(x_i | b_0 + b_1 y_i, \sigma_\epsilon^2) \\ &\times \int \prod_{i \notin s} h(\lambda_i | \boldsymbol{\nu}) \int f(\mathbf{i} | \mathbf{y}, \mathbf{x}) \prod_{i \notin s} [N(x_i | b_0 + b_1 y_i, \sigma_\epsilon^2) N(y_i | \mathbf{z}'_i \boldsymbol{\beta}, \lambda_i^{-1} \sigma^2)] d\mathbf{y}_{\bar{s}} d\boldsymbol{\lambda}_{\bar{s}}. \end{aligned}$$

Reescrevendo $N(x_i | b_0 + b_1 y_i, \sigma_\epsilon^2) N(y_i | \mathbf{z}'_i \boldsymbol{\beta}, \lambda_i^{-1} \sigma^2)$, obtém-se:

$$\begin{aligned} f(\mathbf{y}_s, \mathbf{x}, \mathbf{i} | \boldsymbol{\theta}, \boldsymbol{\phi}, \mathbf{z}) &= \int \prod_{i \in s} [h(\lambda_i | \boldsymbol{\nu}) N(y_i | \mathbf{z}'_i \boldsymbol{\beta}, \lambda_i^{-1} \sigma^2)] d\boldsymbol{\lambda}_s \prod_{i \in s} N(x_i | b_0 + b_1 y_i, \sigma_\epsilon^2) \\ &\times \int \prod_{i \notin s} h(\lambda_i | \boldsymbol{\nu}) \int f(\mathbf{i} | \mathbf{y}, \mathbf{x}) \prod_{i \notin s} N(y_i, x_i | \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) d\mathbf{y}_{\bar{s}} d\boldsymbol{\lambda}_{\bar{s}}, \end{aligned}$$

onde $\boldsymbol{\mu}_i = [\mathbf{z}'_i \boldsymbol{\beta}, b_0 + b_1 \mathbf{z}'_i \boldsymbol{\beta}]'$ e

$$\boldsymbol{\Sigma}_i = \begin{bmatrix} \lambda_i^{-1} \sigma^2 & \lambda_i^{-1} \sigma^2 b_1 \\ \lambda_i^{-1} \sigma^2 b_1 & \lambda_i^{-1} \sigma^2 b_1^2 + \sigma_\epsilon^2 \end{bmatrix}.$$

Para maiores detalhes, ver Anexo A Seção A.1. Como a distribuição de \mathbf{I} é uma função de \mathbf{X} , integrando em $\mathbf{y}_{\bar{s}}$, tem-se:

$$\begin{aligned} f(\mathbf{y}_s, \mathbf{x}, \mathbf{i} | \boldsymbol{\theta}, \boldsymbol{\phi}, \mathbf{z}) &= \int \prod_{i \in s} [h(\lambda_i | \boldsymbol{\nu}) N(y_i | \mathbf{z}'_i \boldsymbol{\beta}, \lambda_i^{-1} \sigma^2)] d\boldsymbol{\lambda}_s \prod_{i \in s} N(x_i | b_0 + b_1 y_i, \sigma_\epsilon) \\ &\times \prod_{j \in s} \left(\frac{x_j^2}{\sum_{i \in U} x_i^2} \right) \int \prod_{i \notin s} [h(\lambda_i | \boldsymbol{\nu}) N(x_i | b_0 + b_1 \mathbf{z}'_i \boldsymbol{\beta}, \sigma_\epsilon^2 + \lambda_i^{-1} \sigma^2 b_1^2)] d\boldsymbol{\lambda}_{\bar{s}}. \end{aligned}$$

A equação acima pode ser reescrita considerando o seguinte modelo hierárquico:

$$\begin{aligned} X_i | \boldsymbol{\beta}, \sigma, b_0, b_1, \sigma_\epsilon, \lambda_i &\sim N(b_0 + b_1 \mathbf{z}'_i \boldsymbol{\beta}, \sigma_\epsilon^2 + \lambda_i^{-1} \sigma^2 b_1^2), \\ \lambda_i | \boldsymbol{\nu} &\sim h(\lambda_i | \boldsymbol{\nu}); \end{aligned}$$

para $i \notin s$. A função densidade de probabilidade para $\mathbf{Y}_s, \mathbf{X}, \mathbf{I}$ condicional a $\boldsymbol{\lambda}$ é dada por:

$$\begin{aligned} f(\mathbf{y}_s, \mathbf{x}, \mathbf{i} | \boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{\lambda}, \mathbf{z}) &= \prod_{i \in s} N(y_i | \mathbf{z}'_i \boldsymbol{\beta}, \lambda_i^{-1} \sigma^2) \prod_{i \in s} N(x_i | b_0 + b_1 y_i, \sigma_\epsilon) \\ &\times \prod_{j \in s} \left(\frac{x_j^2}{\sum_{i \in U} x_i^2} \right) \prod_{i \notin s} N(x_i | b_0 + b_1 \mathbf{z}'_i \boldsymbol{\beta}, \sigma_\epsilon^2 + \lambda_i^{-1} \sigma^2 b_1^2). \end{aligned} \quad (3.9)$$

O núcleo da distribuição *a posteriori* conjunta dos vetores paramétricos $\boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{\lambda}$ é dado por:

$$\begin{aligned} f(\boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{\lambda} | \mathbf{y}_s, \mathbf{x}, \mathbf{i}, \mathbf{z}) &\propto \prod_{i \in s} [h(\lambda_i | \boldsymbol{\nu}) N(y_i | \mathbf{z}'_i \boldsymbol{\beta}, \lambda_i^{-1} \sigma^2)] \prod_{i \in s} N(x_i | b_0 + b_1 y_i, \sigma_\epsilon) \\ &\times \prod_{i \notin s} [h(\lambda_i | \boldsymbol{\nu}) N(x_i | b_0 + b_1 \mathbf{z}'_i \boldsymbol{\beta}, \sigma_\epsilon^2 + \lambda_i^{-1} \sigma^2 b_1^2)] \pi(\boldsymbol{\theta}, \boldsymbol{\phi}). \end{aligned} \quad (3.10)$$

onde $\pi(\boldsymbol{\theta}, \boldsymbol{\phi})$ é a distribuição *a priori* conjunta dos vetores paramétricos $\boldsymbol{\theta}$ e $\boldsymbol{\phi}$.

No caso da amostragem aleatória simples com a variável auxiliar, o processo de obtenção da função de densidade de probabilidade $\mathbf{Y}_s, \mathbf{X}, \mathbf{I}$ é similar ao anterior. A expressão do núcleo da distribuição *a posteriori* é apresentada abaixo:

$$f(\boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{\lambda} | \mathbf{y}_s, \mathbf{x}, \mathbf{i}, \mathbf{z}) \propto \prod_{i \in s} [h(\lambda_i | \boldsymbol{\nu}) N(y_i | \mu, \lambda_i^{-1} \sigma^2)] \prod_{i \in s} N(x_i | b_0 + b_1 y_i, \sigma_\epsilon) \\ \times \prod_{i \notin s} [h(\lambda_i | \boldsymbol{\nu}) N(x_i | b_0 + b_1 \mathbf{z}'_i \boldsymbol{\beta}, \sigma_\epsilon^2 + \lambda_i^{-1} \sigma^2 b_1^2)] \pi(\boldsymbol{\theta}, \boldsymbol{\phi}). \quad (3.11)$$

Embora as expressões (3.10) e (3.11) conduzam aos mesmos estimadores para os parâmetros desconhecidos, as variáveis têm distribuições de probabilidades diferentes, uma vez que os planos amostrais são diferentes.

De maneira análoga aos cálculos apresentados anteriormente, obtém-se a função de densidade de probabilidade dos valores amostrados $\mathbf{y}_s, \mathbf{x}_s$ do plano amostragem quadrática simétrica com variável auxiliar:

$$f(\mathbf{y}_s, \mathbf{x}_s | \boldsymbol{\theta}, \boldsymbol{\phi}, \mathbf{z}) = \prod_{i \in s} \left\{ \frac{f_p(y_i | \boldsymbol{\theta}, \mathbf{z}_i) f_p(x_i | y_i, b_0, b_1, \sigma_\epsilon) (x_i - c)^2}{[b_1 E_p(Y_i | \boldsymbol{\theta}, \mathbf{z}_i) + b_0^*]^2 + b_1^2 V_p(Y_i | \boldsymbol{\theta}, \mathbf{z}_i) + \sigma_\epsilon^2} \right\}, \quad (3.12)$$

onde $b_0^* = b_0 - c$.

O núcleo da distribuição *a posteriori* dos parâmetros $\boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{\lambda}$ dado $\mathbf{y}_s, \mathbf{x}, \mathbf{i}$ é dado por:

$$f(\boldsymbol{\theta}, \boldsymbol{\phi}, \boldsymbol{\lambda} | \mathbf{y}_s, \mathbf{x}, \mathbf{i}, \mathbf{z}) \propto \prod_{i \in s} [h(\lambda_i | \boldsymbol{\nu}) N(y_i | \mathbf{z}'_i \boldsymbol{\beta}, \lambda_i^{-1} \sigma^2)] \prod_{i \in s} N(x_i | b_0 + b_1 y_i, \sigma_\epsilon) \\ \times \prod_{i \notin s} [h(\lambda_i | \boldsymbol{\nu}) N(x_i | b_0 + b_1 \mathbf{z}'_i \boldsymbol{\beta}, \sigma_\epsilon^2 + \lambda_i^{-1} \sigma^2 b_1^2)] \pi(\boldsymbol{\theta}, \boldsymbol{\phi}). \quad (3.13)$$

3.3.2 Amostragem estratificada informativa

Num cenário de dados que apresentam o comportamento de caudas pesadas, um plano amostral estratificado pode ser adotado. Este plano amostral será chamado de *amostragem estratificada informativa*. Amostragem estratificada consiste na divisão de uma população em grupos, denominados de estratos, segundo uma ou mais características conhecidas na população (Parsons, 2014). Em cada grupo são selecionadas amostras em proporções convenientes. A ideia é selecionar mais valores extremos da distribuição através da estratificação, que é feita pela variável auxiliar.

No presente trabalho, para o modelo de misturas normais com caudas pesadas, considerou-se o seguinte modelo de regressão:

$$Y_j = \mathbf{z}'_j \boldsymbol{\beta} + \xi_j, \quad j = 1, 2, \dots, N. \quad (3.14)$$

Reescrevendo ξ_j na notação de mistura, tem-se:

$$\begin{aligned} \xi_j | \lambda_j &\sim N(0, \lambda_j^{-1} \sigma_\xi^2), \quad j = 1, 2, \dots, N; \\ \lambda_j | \nu &\sim h(\lambda_j | \nu), \quad j = 1, 2, \dots, N. \end{aligned}$$

As unidades são selecionadas pela amostragem estratificada, sendo $O_j = h$, $h = 1, \dots, H$, a variável indicadora do estrato h em que a unidade amostral j pertence, que depende dos valores y_j . Em relação a construção dos estratos, considere que a variável de estratificação (X) esteja correlacionada com a variável resposta (Y) e possa ser modelada da seguinte forma:

$$X_j = b_0 + b_1 Y_j + \epsilon_j; \quad \text{onde } \epsilon_j \sim N(0, \sigma_\epsilon^2); \quad j = 1, 2, \dots, N. \quad (3.15)$$

Na prática os estratos são construídos da seguinte forma: a unidade j pertence ao estrato 1 se $x_j < c_1$; ao estrato 2 se $c_1 \leq x_j < c_2$; ...; ao estrato H se $x_j \geq c_{H-1}$, em que c_1, \dots, c_{H-1} são pontos de corte. Geralmente o analista desconhece os valores dos c_i por não ter acesso a variável X . Na maioria das ocasiões, tem-se conhecimento apenas da variável indicadora do estrato h em que a unidade amostral j pertence. Por este motivo, considerou-se que os c_i são desconhecidos.

Uma vantagem na aplicação da amostragem estratificada é que, uma vez conhecido a variável indicadora do estrato h para todas as unidades populacionais, o pesquisador não precisa conhecer os valores de X .

A equação abaixo é referente à distribuição de Y_j , $j \in S$ (Pfeffermann et al., 1998):

$$f_s(y_j | \mathbf{z}_j, \lambda_j, \boldsymbol{\theta}, \boldsymbol{\psi}) = \frac{E_p(\pi_j | y_j, \mathbf{z}_j, \lambda_j, \boldsymbol{\theta}, \boldsymbol{\psi}) f_p(y_j | \mathbf{z}_j, \lambda_j, \boldsymbol{\theta})}{E_p(\pi_j | \mathbf{z}_j, \lambda_j, \boldsymbol{\theta}, \boldsymbol{\psi})}. \quad (3.16)$$

As esperanças condicionais presentes na expressão anterior possivelmente dependem de um vetor paramétrico desconhecido adicional $\boldsymbol{\psi}$. Os cálculos das esperanças da equação

(3.16) são apresentados abaixo:

$$\begin{aligned} E_p(\pi_j|y_j, \mathbf{z}_j, \lambda_j, \boldsymbol{\theta}, \boldsymbol{\psi}) &= p(j \in s|y_j, \mathbf{z}_j, \lambda_j, \boldsymbol{\theta}, \boldsymbol{\psi}) = \sum_{k=1}^H q_k p(O_j = k|y_j, \mathbf{z}_j, \lambda_j, \boldsymbol{\theta}, \boldsymbol{\psi}) \\ &= q_1 A_1(y_j) + q_2 \{A_2(y_j) - A_1(y_j)\} + \cdots + q_H (1 - A_{H-1}(y_j)), \end{aligned}$$

onde $q_k = p(j \in s|O_j = k)$, $k = 1, 2, \dots, H$ é a fração de amostragem do estrato k e $A_k(y_j) = p(X_j < c_k|y_j, \mathbf{z}_j, \boldsymbol{\theta}, \boldsymbol{\psi}) = \Phi\{(c_k - b_0 - b_1 y_j)/\sigma_\epsilon\}$. Note que neste caso particular $\boldsymbol{\psi} = (b_0, b_1, \{c_k\}, \sigma_\epsilon)$. De forma análoga,

$$\begin{aligned} E_p(\pi_j|\mathbf{z}_j, \lambda_j, \boldsymbol{\theta}, \boldsymbol{\psi}) &= \sum_{k=1}^H q_k p(O_j = k|\mathbf{z}_j, \lambda_j, \boldsymbol{\theta}, \boldsymbol{\psi}) \\ &= q_1 B_1\{\mu(y_j)\} + q_2 [B_2\{\mu(y_j)\} - B_1\{\mu(y_j)\}] + \cdots + \\ &+ q_H [1 - B_{H-1}\{\mu(y_j)\}]; \end{aligned}$$

$$\mu(y_j) = \mathbf{z}'_j \boldsymbol{\beta};$$

$$B_k\{\mu(y_j)\} = p(X_j < c_k|\mathbf{z}_j, \boldsymbol{\theta}) = \Phi\left(\frac{c_k - b_0 - b_1 \mu\{y_j\}}{\sqrt{\sigma_\epsilon^2 + b_1^2 \lambda_j^{-1} \sigma_\xi^2}}\right).$$

Denotemos por $\mathbf{Y} = \{y_j; j = 1, 2, \dots, N\}$ e $\mathbf{O} = \{O_j; j = 1, 2, \dots, N\}$, respectivamente os valores observados e os indicadores de estrato. Os valores observados na amostra são denotados por $\mathbf{D}_{obs} = (\mathbf{Y}_{obs}, \mathbf{O}_{obs})$. Vale destacar que a função de verossimilhança do plano depende de \mathbf{D}_{obs} .

Considere a seguinte reparametrização $\eta_k = (c_k - b_0)/\sigma_\epsilon$, para $k = 1, 2, \dots, H - 1$ e $\eta_H = -(b_1/\sigma_\epsilon)$. Note que $\eta_1 < \eta_2 < \cdots < \eta_{H-1}$. Reescrevendo $A_k(y_j)$ e $B_k\{\mu(y_j)\}$ com a nova parametrização, tem-se:

$$\begin{aligned} A_k(y_j) &= \Phi(\eta_k + \eta_H y_j); \\ B_k\{\mu(y_j)\} &= \Phi\left(\frac{\eta_k + \eta_H \mu(y_j)}{\sqrt{1 + \eta_H^2 \lambda_j^{-1} \sigma_\xi^2}}\right); \end{aligned}$$

O núcleo da distribuição *a posteriori* de $\boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{\lambda}$ é dado por:

$$\begin{aligned} f(\boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{\lambda}|\mathbf{D}_{obs}, \mathbf{z}) &\propto \prod_{j=1}^N [f_s(y_j|\mathbf{z}_j, \lambda_j, \boldsymbol{\beta}, \boldsymbol{\eta}, \sigma) p_s(O_j|y_j, \lambda_j, \boldsymbol{\eta}) h(\lambda_j|\boldsymbol{\nu})] \times \\ &\pi(\boldsymbol{\theta}, \boldsymbol{\eta}); \end{aligned} \tag{3.17}$$

onde $\pi(\boldsymbol{\theta}, \boldsymbol{\eta})$ é a distribuição *a priori* conjunta do vetor paramétrico $\boldsymbol{\theta}$ e do vetor $\boldsymbol{\eta}$ e

$$p_s(O_j = k | y_j, \lambda_j, \boldsymbol{\eta}) = \frac{q_k p(O_j = k | y_j, \lambda_j, \boldsymbol{\eta})}{\sum_{l=1}^H q_l p(O_j = l | y_j, \lambda_j, \boldsymbol{\eta})}, k = 1, \dots, H;$$

$$p(O_j = 1 | y_j, \lambda_j, \boldsymbol{\eta}) = A_1(y_j); \quad p(O_j = 2 | y_j, \lambda_j, \boldsymbol{\eta}) = A_2(y_j) - A_1(y_j), \dots,$$

$$p(O_j = H | y_j, \lambda_j, \boldsymbol{\eta}) = 1 - A_{H-1}(y_j).$$

Capítulo 4

Amostragem eficiente num modelo de regressão T-Student

Neste capítulo serão apresentados detalhes sobre os planos amostrais considerados para o modelo de regressão T-Student. Entre os resultados, estão os núcleos da distribuição *a posteriori*, no contexto da inferência bayesiana. A particularização da escolha da distribuição T-Student se deve ao fato dela ser comumente utilizada para modelar dados com caudas pesadas. Entretanto, não há uma perda de generalidade da metodologia desenvolvida considerando-se os planos amostrais apresentados anteriormente. Eles poderiam ser aplicados para outras distribuições da família de distribuições com caudas pesadas.

O modelo de regressão T-Student pode ser escrito como:

$$Y_i | \boldsymbol{\beta}, \sigma, \nu, \mathbf{z}_i \sim t_\nu \left(\mathbf{z}'_i \boldsymbol{\beta}, \sigma^2 \frac{\nu}{\nu - 2} \right), i = 1, 2, \dots, N;$$
$$\mathbf{z}'_i = [1, z_{i1}, z_{i2}, \dots, z_{ik}] \quad ; \quad \boldsymbol{\beta}' = [\beta_0, \beta_1, \beta_2, \dots, \beta_k].$$

onde $\mathbf{z}'_i \boldsymbol{\beta}$ é a esperança e $\sigma^2 \frac{\nu}{\nu - 2}$ é variância da variável $Y_i | \boldsymbol{\beta}, \sigma, \nu, \mathbf{z}_i$.

Conforme visto no capítulo anterior, foram consideradas as funções densidades de probabilidades dos valores amostrados $\mathbf{y}_s, \mathbf{x}_s$ e dos valores $\mathbf{y}_s, \mathbf{x}, \mathbf{i}$. Quando for considerada para a inferência a função densidade de probabilidade dos valores amostrados $\mathbf{y}_s, \mathbf{x}_s$, denominou-se de “simplificada”, e quando for considerada a função de densidade de probabilidade dos valores $\mathbf{y}_s, \mathbf{x}_s, \mathbf{i}$ de “completa”. A Tabela 4.1 mostra as siglas que são

usadas no decorrer do trabalho.

Modelos considerados para inferência	Siglas
Amostragem aleatória simples	AAS
Amostragem aleatória simples com variável auxiliar simplificada	AASX
Amostragem quadrática com variável auxiliar simplificada	AQ
Amostragem quadrática simétrica com variável auxiliar simplificada	AQS
Amostragem aleatória simples com variável auxiliar completa	AASXC
Amostragem quadrática com variável auxiliar completa	AQC
Amostragem quadrática simétrica com variável auxiliar completa	AQSC
Amostragem estratificada informativa	AEI

Tabela 4.1: Siglas referentes aos modelos considerados para inferência.

A seguir serão apresentados os núcleos das distribuições *a posteriori* dos modelos considerados na Tabela 4.1.

Amostragem aleatória simples (AAS) A função de verossimilhança da regressão T-Student para o modelo AAS é dada por:

$$L(\boldsymbol{\beta}, \sigma, \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)^n \nu^{n\frac{\nu}{2}}}{\Gamma\left(\frac{\nu}{2}\right)^n \pi^{\frac{n}{2}} \sigma^n} \prod_{i \in s} \left[\nu + \left(\frac{y_i - \mathbf{z}'_i \boldsymbol{\beta}}{\sigma} \right)^2 \right]^{-\left(\frac{\nu+1}{2}\right)}.$$

A informação de Fisher para o vetor paramétrico $\boldsymbol{\theta} = \{\boldsymbol{\beta}, \sigma, \nu\}$ é dada por (ver [Fonseca et al. \(2008\)](#)):

$$I_1(\boldsymbol{\theta}) = \begin{bmatrix} \frac{1}{\sigma^2} \frac{(\nu+1)}{(\nu+3)} \sum_{i \in s} \mathbf{z}_i \mathbf{z}'_i & \mathbf{0}_{(k+1) \times 1} & \mathbf{0}_{(k+1) \times 1} \\ \mathbf{0}_{1 \times (k+1)} & \frac{2n}{\sigma^2} \frac{\nu}{(\nu+3)} & -\frac{2n}{\sigma} \frac{1}{(\nu+1)(\nu+3)} \\ \mathbf{0}_{1 \times (k+1)} & -\frac{2n}{\sigma} \frac{1}{(\nu+1)(\nu+3)} & \frac{n}{4} \left[\psi' \left(\frac{\nu}{2} \right) - \psi' \left(\frac{\nu+1}{2} \right) - \frac{2(\nu+5)}{\nu(\nu+1)(\nu+3)} \right] \end{bmatrix}. \quad (4.1)$$

onde $\mathbf{0}_{(k+1) \times 1}$ e $\mathbf{0}_{1 \times (k+1)}$ são vetores com todos elementos iguais a 0. Maiores detalhes referentes a matriz $\mathbf{z}_i \mathbf{z}'_i$ estão no Anexo A.

O núcleo da distribuição *a posteriori* é da forma:

$$f(\boldsymbol{\beta}, \sigma, \nu | \mathbf{y}_s, \mathbf{z}) \propto \frac{\Gamma\left(\frac{\nu+1}{2}\right)^n \nu^{n\frac{\nu}{2}}}{\Gamma\left(\frac{\nu}{2}\right)^n \pi^{\frac{n}{2}} \sigma^n} \prod_{i \in s} \left[\nu + \left(\frac{y_i - \mathbf{z}'_i \boldsymbol{\beta}}{\sigma} \right)^2 \right]^{-\left(\frac{\nu+1}{2}\right)} \pi(\boldsymbol{\beta}, \nu, \sigma); \quad (4.2)$$

onde $\pi(\boldsymbol{\beta}, \nu, \sigma)$ corresponde à distribuição *a priori* dos parâmetros da distribuição T-Student.

Amostragem aleatória simples com variável auxiliar simplificada (AASX) A partir da expressão (3.5), tem-se que a função de verossimilhança conjunta de \mathbf{x}_s e \mathbf{y}_s para o modelo AASX pode ser escrita como:

$$L(\boldsymbol{\beta}, \sigma, \nu, b_0, b_1, \sigma_\epsilon) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)^n \nu^{n\frac{\nu}{2}}}{\Gamma\left(\frac{\nu}{2}\right)^n \pi^{\frac{n}{2}} \sigma^n} \prod_{i \in s} \left[\nu + \left(\frac{y_i - \mathbf{z}'_i \boldsymbol{\beta}}{\sigma} \right)^2 \right]^{-\left(\frac{\nu+1}{2}\right)} \times$$

$$\times \frac{1}{(2\pi)^{\frac{n}{2}} \sigma_\epsilon^n} \exp \left\{ \frac{-1}{2\sigma_\epsilon^2} \sum_{i \in s} [x_i - (b_0 + b_1 y_i)]^2 \right\}. \quad (4.3)$$

A matriz de informação de Fisher para o vetor paramétrico $\boldsymbol{\gamma} = \{\boldsymbol{\theta}, b_0, b_1, \sigma_\epsilon\}$ é dada por:

$$I_2(\boldsymbol{\gamma}) = \begin{bmatrix} I_1(\boldsymbol{\theta}) & \mathbf{0}_{(k+3) \times 3} \\ \mathbf{0}_{3 \times (k+3)} & I(\boldsymbol{\phi}) \end{bmatrix}; \quad (4.4)$$

onde $\mathbf{0}_{(k+3) \times 3}$ e $\mathbf{0}_{3 \times (k+3)}$ são matrizes nulas e $I(\boldsymbol{\phi})$ é:

$$I(\boldsymbol{\phi}) = \begin{bmatrix} \frac{n}{\sigma_\epsilon^2} & \frac{\sum_{i \in s} \mathbf{z}'_i \boldsymbol{\beta}}{\sigma_\epsilon^2} & 0 \\ \frac{\sum_{i \in s} \mathbf{z}'_i \boldsymbol{\beta}}{\sigma_\epsilon^2} & \frac{1}{\sigma_\epsilon^2} \left(n\sigma^2 \frac{\nu}{\nu-2} + \sum_{i \in s} (\mathbf{z}'_i \boldsymbol{\beta})^2 \right) & 0 \\ 0 & 0 & \frac{2n}{\sigma_\epsilon^2} \end{bmatrix}.$$

Detalhes técnicos para obtenção desta matriz estão no Apêndice A Seção A.2.

O núcleo da distribuição *a posteriori* é da forma:

$$f(\boldsymbol{\beta}, \sigma, \nu, b_0, b_1, \sigma_\epsilon | \mathbf{y}_s, \mathbf{x}_s, \mathbf{z}) \propto \frac{\Gamma\left(\frac{\nu+1}{2}\right)^n \nu^{n\frac{\nu}{2}}}{\Gamma\left(\frac{\nu}{2}\right)^n \pi^{\frac{n}{2}} \sigma^n} \prod_{i \in s} \left[\nu + \left(\frac{y_i - \mathbf{z}'_i \boldsymbol{\beta}}{\sigma} \right)^2 \right]^{-\left(\frac{\nu+1}{2}\right)}$$

$$\times \frac{1}{(2\pi)^{\frac{n}{2}} \sigma_\epsilon^n} \exp \left\{ \frac{-1}{2\sigma_\epsilon^2} \sum_{i \in s} [x_i - (b_0 + b_1 y_i)]^2 \right\} \pi(\boldsymbol{\beta}, \sigma, \nu, b_0, b_1, \sigma_\epsilon); \quad (4.5)$$

onde $\pi(\boldsymbol{\beta}, \sigma, \nu, b_0, b_1, \sigma_\epsilon)$ corresponde à distribuição *a priori* conjunta dos parâmetros da distribuição T-Student e dos parâmetros $b_0, b_1, \sigma_\epsilon$.

Amostragem quadrática com variável auxiliar simplificada (AQ) Considerando que $\nu > 2$, tem-se que $V_p(Y_i|\boldsymbol{\beta}, \sigma, \nu, \mathbf{z}_i) = \sigma^2 \frac{\nu}{\nu-2}$. De (3.7), tem-se que a função de verossimilhança para o modelo AQ é dada por:

$$L(\boldsymbol{\beta}, \sigma, \nu, b_0, b_1, \sigma_\epsilon) = \prod_{i \in s} \left[\frac{f_p(y_i|\boldsymbol{\beta}, \sigma, \nu, \mathbf{z}_i) f_p(x_i|y_i, b_0, b_1, \sigma_\epsilon) x_i^2}{(b_1 \mathbf{z}'_i \boldsymbol{\beta} + b_0)^2 + b_1^2 \sigma^2 \frac{\nu}{\nu-2} + \sigma_\epsilon^2} \right]. \quad (4.6)$$

Aplicando-se o logaritmo em (4.6) tem-se:

$$\begin{aligned} \log [L(\boldsymbol{\beta}, \sigma, \nu, b_0, b_1, \sigma_\epsilon)] &= \sum_{i \in s} \log [f_p(y_i|\boldsymbol{\beta}, \sigma, \nu, \mathbf{z}_i)] + \sum_{i \in s} \log [f_p(x_i|y_i, b_0, b_1, \sigma_\epsilon)] \\ &+ \sum_{i \in s} \log(x_i^2) - \sum_{i \in s} \log \left[(b_1 \mathbf{z}'_i \boldsymbol{\beta} + b_0)^2 + b_1^2 \sigma^2 \frac{\nu}{\nu-2} + \sigma_\epsilon^2 \right]. \end{aligned} \quad (4.7)$$

A informação de Fisher deste modelo é composta pela soma de duas matrizes. A primeira é a matriz de Fisher do modelo AASX e a segunda é uma matriz simétrica referente à correção da informatividade.

$$I_3(\boldsymbol{\gamma}) = I_2(\boldsymbol{\gamma}) + A(\boldsymbol{\gamma}). \quad (4.8)$$

A matriz $A(\boldsymbol{\gamma})$ é apresentada no Anexo A Seção A.2.

Uma observação pertinente com respeito ao presente modelo é que $E \left[-\frac{\partial a(\boldsymbol{\gamma})}{\partial \nu^2} \right] = a_{33} > 0$. Note que:

$$b_1^2 \sigma^2 (\nu - 2)^{-1} \sum_{i \in s} \frac{1}{k_i^2} < \sum_{i \in s} \frac{1}{k_i};$$

onde $k_i = (b_1 \mathbf{z}'_i \boldsymbol{\beta} + b_0)^2 + b_1^2 \sigma^2 \frac{\nu}{\nu-2} + \sigma_\epsilon^2$. Para se demonstrar a relação anterior, basta observar que $b_1^2 \sigma^2 (\nu - 2)^{-1} < b_1^2 \sigma^2 \frac{\nu}{\nu-2}$, uma vez que $\nu > 2$. Logo:

$$\begin{aligned} b_1^2 \sigma^2 (\nu - 2)^{-1} &< (b_1 \mathbf{z}'_i \boldsymbol{\beta} + b_0)^2 + b_1^2 \sigma^2 \frac{\nu}{\nu-2} + \sigma_\epsilon^2; \\ b_1^2 \sigma^2 (\nu - 2)^{-1} &< k_i; \\ b_1^2 \sigma^2 (\nu - 2)^{-1} \frac{1}{k_i^2} &< \frac{1}{k_i}; \forall i. \end{aligned}$$

É importante destacar que considerando a função de verossimilhança marginal de ν , a informação de Fisher para o modelo AQ é sempre maior do que a informação de Fisher

para o modelo AASX. A amostragem proposta traz mais informação, de fato, a respeito de ν .

O núcleo da distribuição *a posteriori* é da forma:

$$f(\boldsymbol{\beta}, \sigma, \nu, b_0, b_1, \sigma_\epsilon | \mathbf{y}_s, \mathbf{x}_s, \mathbf{z}) \propto \prod_{i \in s} \left[\frac{f_p(y_i | \boldsymbol{\beta}, \sigma, \nu, \mathbf{z}_i) f_p(x_i | y_i, b_0, b_1, \sigma_\epsilon) x_i^2}{(b_1 \mathbf{z}'_i \boldsymbol{\beta} + b_0)^2 + b_1^2 \sigma^2 \frac{\nu}{\nu-2} + \sigma_\epsilon^2} \right] \times \pi(\boldsymbol{\beta}, \sigma, \nu, b_0, b_1, \sigma_\epsilon). \quad (4.9)$$

Amostragem quadrática simétrica com variável auxiliar simplificada (AQS)

De (3.12), tem-se a função de verossimilhança do modelo AQS para a distribuição T-Student é apresentada abaixo:

$$L(\boldsymbol{\beta}, \sigma, \nu, b_0, b_1, \sigma_\epsilon) = \prod_{i \in s} \left[\frac{f_p(y_i | \boldsymbol{\beta}, \sigma, \nu, \mathbf{z}_i) f_p(x_i | y_i, b_0, b_1, \sigma_\epsilon) (x_i - c)^2}{(b_1 \mathbf{z}'_i \boldsymbol{\beta} + b_0 - c)^2 + b_1^2 \sigma^2 \frac{\nu}{\nu-2} + \sigma_\epsilon^2} \right].$$

Considerando $b_0^* = b_0 - c$, tem-se que $X_i | Y_i = y_i, b_0^*, b_1, \sigma_\epsilon \sim N(b_0^* + c + b_1 y_i, \sigma_\epsilon^2)$. Logo:

$$L(\boldsymbol{\beta}, \sigma, \nu, b_0^*, b_1, \sigma_\epsilon) = \prod_{i \in s} \left[\frac{f_p(y_i | \boldsymbol{\beta}, \sigma, \nu, \mathbf{z}_i) f_p(x_i | y_i, b_0^*, b_1, \sigma_\epsilon) (x_i - c)^2}{(b_1 \mathbf{z}'_i \boldsymbol{\beta} + b_0^*)^2 + b_1^2 \sigma^2 \frac{\nu}{\nu-2} + \sigma_\epsilon^2} \right]. \quad (4.10)$$

Aplicando-se o logaritmo na expressão (4.10), tem-se:

$$\begin{aligned} \log[L(\boldsymbol{\beta}, \sigma, \nu, b_0^*, b_1, \sigma_\epsilon)] &= \sum_{i \in s} \log[f_p(y_i | \boldsymbol{\beta}, \sigma, \nu, \mathbf{z}_i)] + \sum_{i \in s} \log[f_p(x_i | y_i, b_0^*, b_1, \sigma_\epsilon)] \\ &+ \sum_{i \in s} \log(x_i - c)^2 - \sum_{i \in s} \log \left\{ (b_1 \mathbf{z}'_i \boldsymbol{\beta} + b_0^*)^2 + b_1^2 \sigma^2 \frac{\nu}{\nu-2} + \sigma_\epsilon^2 \right\}. \end{aligned} \quad (4.11)$$

Note que a expressão do logaritmo da função de verossimilhança em (4.11) é similar à expressão (4.7). Logo a informação de Fisher deste modelo é similar à informação do modelo anterior, com a diferença que b_0 é substituído por b_0^* e $X_i | Y_i = y_i, b_0^*, b_1, \sigma_\epsilon \sim N(b_0^* + c + b_1 y_i, \sigma_\epsilon^2)$.

O núcleo da distribuição *a posteriori* é da forma:

$$f(\boldsymbol{\beta}, \sigma, \nu, b_0^*, b_1, \sigma_\epsilon | \mathbf{y}_s, \mathbf{x}, \mathbf{i}, \mathbf{z}) \propto \prod_{i \in s} \left[\frac{f_p(y_i | \boldsymbol{\beta}, \sigma, \nu, \mathbf{z}_i) f_p(x_i | y_i, b_0^*, b_1, \sigma_\epsilon) (x_i - c)^2}{(b_1 \mathbf{z}'_i \boldsymbol{\beta} + b_0^*)^2 + b_1^2 \sigma^2 \frac{\nu}{\nu-2} + \sigma_\epsilon^2} \right] \times \pi(\boldsymbol{\beta}, \sigma, \nu, b_0^*, b_1, \sigma_\epsilon). \quad (4.12)$$

Considerou-se \bar{X} como a constante de centralização, apesar de ser uma variável aleatória. Esta consideração é apropriada, uma vez que geralmente o tamanho da

população N é muito grande, tornando \bar{X} aproximadamente uma constante. Assim $c = \bar{X} = \frac{\sum_{i \in U} x_i}{N}$.

Levando em consideração a aleatoriedade de \bar{X} , uma aproximação adotada, utilizando-se a lei dos grandes números é $E(\bar{X}) \approx \bar{X}$. Considerando-se o modelo linear $X_i = Y_i + \epsilon_i$ com $E(Y_i) = \mu$, tem-se que a distribuição conjunta de Y_i e X_i pode ser aproximada por:

$$f_s(y_i, x_i | \mu, \sigma, \nu, \sigma_\epsilon) \approx \frac{f_p(y_i | \mu, \sigma, \nu) f_p(x_i | y_i, \sigma_\epsilon) [x_i - E(\bar{X})]^2}{\int f_p(y_i | \mu, \sigma, \nu) E_{X_i|Y_i} \left\{ [X_i - E(\bar{X})]^2 \right\} dy_i}. \quad (4.13)$$

Note que:

$$E(\bar{X}) = E[E(\bar{X} | \mathbf{Y})] = E\left[\frac{1}{N} \sum_{i \in U} E(X_i | Y_i)\right] = \frac{1}{N} \sum_{i \in U} E(Y_i) = \mu.$$

Logo:

$$\begin{aligned} E_{X_i|Y_i} [(X_i - \mu)^2] &= E_{X_i|Y_i} (X_i^2) - 2\mu E_{X_i|Y_i} (X_i) + \mu^2 =; \\ &= y_i^2 + \sigma_\epsilon^2 - 2\mu y_i + \mu^2. \end{aligned} \quad (4.14)$$

Substituindo (4.14) em (4.13), tem-se:

$$\begin{aligned} f_s(y_i, x_i | \mu, \sigma, \nu, \sigma_\epsilon) &\approx \frac{f_p(y_i | \mu, \sigma, \nu) f_p(x_i | y_i, \sigma_\epsilon) (x_i - \mu)^2}{\int f_p(y_i | \mu, \sigma, \nu) (y_i^2 + \sigma_\epsilon^2 - 2\mu y_i + \mu^2) dy_i} = \\ &= \frac{f_p(y_i | \mu, \sigma, \nu) f_p(x_i | y_i, \sigma_\epsilon) (x_i - \mu)^2}{[E_p(Y_i^2) + \sigma_\epsilon^2] - 2\mu E_p(Y_i) + \mu^2}. \end{aligned}$$

Tem-se então:

$$\begin{aligned} f_s(y_i, x_i | \mu, \sigma, \nu, \sigma_\epsilon) &\approx \frac{f_p(y_i | \mu, \sigma, \nu) f_p(x_i | y_i, \sigma_\epsilon) (x_i - \mu)^2}{[\mu^2 + \sigma^2 \frac{\nu}{\nu-2} + \sigma_\epsilon^2] - 2\mu\mu + \mu^2} = \\ &= \frac{f_p(y_i | \mu, \sigma, \nu) f_p(x_i | y_i, \sigma_\epsilon) (x_i - \mu)^2}{\sigma^2 \frac{\nu}{\nu-2} + \sigma_\epsilon^2}, \end{aligned} \quad (4.15)$$

desde que $\nu > 2$. A função de verossimilhança é dada por:

$$L(\mu, \sigma, \nu, \sigma_\epsilon) \approx \prod_{i \in S} \left[\frac{f_p(y_i | \mu, \sigma, \nu) f_p(x_i | y_i, \sigma_\epsilon) (x_i - \mu)^2}{\sigma^2 \frac{\nu}{\nu-2} + \sigma_\epsilon^2} \right]. \quad (4.16)$$

Aplicando-se o logaritmo na expressão (4.16), tem-se:

$$\log [L(\mu, \sigma, \nu, \sigma_\epsilon)] \approx \sum_{i \in s} \log [f_p(y_i | \mu, \sigma, \nu)] + \sum_{i \in s} \log [f_p(x_i | y_i, \sigma_\epsilon)] + \sum_{i \in s} \log (x_i - \mu)^2 -$$

$$-n \log \left(\sigma^2 \frac{\nu}{\nu-2} + \sigma_\epsilon^2 \right). \quad (4.17)$$

A informação de Fisher deste modelo também é composta pela soma de duas matrizes simétricas:

$$I_4(\boldsymbol{\gamma}) = I_2(\boldsymbol{\gamma}) + B(\boldsymbol{\gamma}), \quad (4.18)$$

com $B(\boldsymbol{\gamma})$ de dimensão 4×4 e elementos

$$\begin{aligned} b_{11} &= 2 \sum_{i \in s} E_p \left[\frac{1}{(X_i - \mu)^2} \right]; b_{12} = b_{21} = b_{13} = b_{31} = b_{14} = b_{41} = 0; \\ b_{22} &= \frac{2n\nu [\sigma_\epsilon^2(\nu-2) - \sigma^2\nu]}{[\sigma^2\nu + \sigma_\epsilon^2(\nu-2)]^2}; b_{23} = \frac{-4n\sigma\sigma_\epsilon^2}{[\sigma^2\nu + \sigma_\epsilon^2(\nu-2)]^2}; b_{24} = \frac{-4n\nu(\nu-2)\sigma\sigma_\epsilon}{[\sigma^2\nu + \sigma_\epsilon^2(\nu-2)]^2}; b_{33} = \frac{4n\sigma^2[\sigma^2(\nu-1) + \sigma_\epsilon^2(\nu-2)]}{[\sigma^2\nu(\nu-2) + \sigma_\epsilon^2(\nu-2)]^2}; \\ b_{34} &= \frac{4n\sigma^2\sigma_\epsilon}{[\sigma^2\nu + \sigma_\epsilon^2(\nu-2)]^2}; b_{44} = \frac{2n(\nu-2)[\sigma^2\nu - \sigma_\epsilon^2(\nu-2)]}{[\sigma^2\nu + \sigma_\epsilon^2(\nu-2)]^2}. \end{aligned}$$

Destaca-se que $E_p \left[\frac{1}{(X_i - \mu)^2} \right]$ não é integrável (maiores detalhes ver Anexo A Seção A.3).

Pode-se ainda obter a expressão da função de verossimilhança considerando-se que \bar{X} não seja uma constante (ver Anexo A Seção A.4):

$$f_s(y_i, x_i | \mu, \sigma, \nu, \sigma_\epsilon) = \frac{(N-1)}{N} \frac{f_p(y_i | \mu, \sigma, \nu) f_p(x_i | y_i, \sigma_\epsilon) (x_i - \mu)^2}{\sigma^2 \frac{\nu}{\nu-2} + \sigma_\epsilon^2} + \frac{f_p(y_i | \mu, \sigma, \nu) f_p(x_i | y_i, \sigma_\epsilon)}{N}.$$

Vale destacar que a expressão aproximada, obtida pela lei dos grandes números, é similar à expressão da verossimilhança encontrada quando N for suficientemente grande. Uma vantagem de considerar $\bar{X} = c$ é justamente a existência da informação de Fisher e a possibilidade da sua utilização para obtenção da *priori* de Jeffreys, ver [Migon \(2015\)](#). Então para o estudo de simulação que será apresentado no Capítulo 5 considerou-se $\bar{X} = c$.

Amostragem aleatória simples com variável auxiliar completa (AASXC) A expressão do núcleo da distribuição *a posteriori* do modelo AASXC é:

$$\begin{aligned} f(\boldsymbol{\beta}, \sigma, \nu, b_0, b_1, \sigma_\epsilon, \boldsymbol{\lambda} | \mathbf{y}_s, \mathbf{x}, \mathbf{i}, \mathbf{z}) &\propto \prod_{i \in s} \left[g \left(\lambda_i \left| \frac{\nu}{2}, \frac{\nu}{2} \right. \right) N(y_i | \mathbf{z}'_i \boldsymbol{\beta}, \lambda_i^{-1} \sigma^2) \right] \prod_{i \in s} N(x_i | b_0 + b_1 y_i, \sigma_\epsilon) \\ &\times \prod_{i \notin s} \left[g \left(\lambda_i \left| \frac{\nu}{2}, \frac{\nu}{2} \right. \right) N(x_i | b_0 + b_1 \mathbf{z}'_i \boldsymbol{\beta}, \sigma_\epsilon^2 + \lambda_i^{-1} \sigma^2 b_1^2) \right] \pi(\boldsymbol{\beta}, \sigma, \nu, b_0, b_1, \sigma_\epsilon). \end{aligned} \quad (4.19)$$

Amostragem quadrática com variável auxiliar completa (AQC) A expressão do núcleo da distribuição *a posteriori* do modelo AQC é:

$$f(\boldsymbol{\beta}, \sigma, \nu, b_0, b_1, \sigma_\epsilon, \boldsymbol{\lambda} | \mathbf{y}_s, \mathbf{x}, \mathbf{i}, \mathbf{z}) \propto \prod_{i \in s} \left[g\left(\lambda_i \left| \frac{\nu}{2}, \frac{\nu}{2} \right.\right) N(y_i | \mathbf{z}'_i \boldsymbol{\beta}, \lambda_i^{-1} \sigma^2) \right] \prod_{i \in s} N(x_i | b_0 + b_1 y_i, \sigma_\epsilon) \\ \times \prod_{i \notin s} \left[g\left(\lambda_i \left| \frac{\nu}{2}, \frac{\nu}{2} \right.\right) N(x_i | b_0 + b_1 \mathbf{z}'_i \boldsymbol{\beta}, \sigma_\epsilon^2 + \lambda_i^{-1} \sigma^2 b_1^2) \right] \pi(\boldsymbol{\beta}, \sigma, \nu, b_0, b_1, \sigma_\epsilon). \quad (4.20)$$

Amostragem quadrática simétrica com variável auxiliar completa (AQSC) A expressão do núcleo da distribuição *a posteriori* do modelo AQSC é:

$$f(\boldsymbol{\beta}, \sigma, \nu, b_0, b_1, \sigma_\epsilon, \boldsymbol{\lambda} | \mathbf{y}_s, \mathbf{x}, \mathbf{i}, \mathbf{z}) \propto \prod_{i \in s} \left[g\left(\lambda_i \left| \frac{\nu}{2}, \frac{\nu}{2} \right.\right) N(y_i | \mathbf{z}'_i \boldsymbol{\beta}, \lambda_i^{-1} \sigma^2) \right] \prod_{i \in s} N(x_i | b_0 + b_1 y_i, \sigma_\epsilon) \\ \times \prod_{i \notin s} \left[g\left(\lambda_i \left| \frac{\nu}{2}, \frac{\nu}{2} \right.\right) N(x_i | b_0 + b_1 \mathbf{z}'_i \boldsymbol{\beta}, \sigma_\epsilon^2 + \lambda_i^{-1} \sigma^2 b_1^2) \right] \pi(\boldsymbol{\beta}, \sigma, \nu, b_0, b_1, \sigma_\epsilon). \quad (4.21)$$

Amostragem estratificada informativa (AEI) Referente ao modelo AEI, considerando a expressão (3.17), tem-se que o núcleo da distribuição *a posteriori* é dado por:

$$f(\boldsymbol{\beta}, \sigma, \nu, \boldsymbol{\eta}, \boldsymbol{\lambda} | \mathbf{D}_{obs}, \mathbf{z}) \propto \prod_{j=1}^N \left[\frac{E_p(\pi_j | y_j, \boldsymbol{\beta}, \sigma, \nu, \boldsymbol{\eta}, \mathbf{z}_j, \lambda_j) f_p(y_j | \boldsymbol{\beta}, \sigma, \nu, \mathbf{z}_j, \lambda_j)}{E_p(\pi_j | \boldsymbol{\beta}, \sigma, \nu, \boldsymbol{\eta}, \mathbf{z}_j, \lambda_j)} \times p_s(O_j | y_j, \lambda_j, \boldsymbol{\eta}) f_s(\lambda_j | \nu) \right] \pi(\boldsymbol{\beta}, \sigma, \nu, \boldsymbol{\eta}); \quad (4.22)$$

com:

$$p_s(O_j = k | y_j, \lambda_j, \boldsymbol{\eta}) = \frac{q_k p(O_j = k | y_j, \lambda_j, \boldsymbol{\eta})}{\sum_{l=1}^H q_l p(O_j = l | y_j, \lambda_j, \boldsymbol{\eta})} \\ p(O_j = 1 | y_j, \lambda_j, \boldsymbol{\eta}) = A_1(y_j); \quad p(O_j = 2 | y_j, \lambda_j, \boldsymbol{\eta}) = A_2(y_j) - A_1(y_j), \dots, \\ p(O_j = H | y_j, \lambda_j, \boldsymbol{\eta}) = 1 - A_{H-1}(y_j); \\ E_p(\pi_j | y_j, \boldsymbol{\beta}, \sigma, \nu, \boldsymbol{\eta}, \mathbf{z}_j, \lambda_j) = q_1 A_1(y_j) + q_2 \{A_2(y_j) - A_1(y_j)\} + \dots + q_H (1 - A_{H-1}(y_j)); \\ E_p(\pi_j | \boldsymbol{\beta}, \sigma, \nu, \boldsymbol{\eta}, \mathbf{z}_j, \lambda_j) = q_1 B_1\{\mu(y_j)\} + q_2 [B_2\{\mu(y_j)\} - B_1\{\mu(y_j)\}] + \dots \\ + q_H [1 - B_{H-1}\{\mu(y_j)\}];$$

onde:

$$A_k(y_j) = \Phi(\eta_k + \eta_H y_j); \quad B_k\{\mu(y_j)\} = \Phi\left(\frac{\eta_k + \eta_H \mu(y_j)}{\sqrt{1 + \eta_H^2 \lambda_j^{-1} \sigma_\epsilon^2}}\right); \quad \mu(y_j) = \mathbf{z}'_j \boldsymbol{\beta}.$$

Nota-se que as expressões dos núcleos das distribuições *a posteriori* dos modelos apresentados não serão analiticamente tratáveis. Uma vantagem dos modelos simplificados em comparação aos modelos completos é o baixo custo computacional exigido para obtenção das amostras das distribuições *a posteriori*, uma vez que eles não necessitam simular as cadeias do vetor λ . Entretanto, os modelos completos possuem uma quantidade muito maior de dados para a estimação dos parâmetros desconhecidos.

Para o estudo de simulação apresentado no Capítulo 5, para cada modelo apresentado, foi considerado que as distribuições *a priori* conjuntas serão obtidas pelo produto das suas respectivas distribuições *a priori* marginais.

Capítulo 5

Estudo de simulação

Nesta seção são apresentados estudos de simulação para avaliar a qualidade na estimação dos parâmetros do modelo de mistura T-Student quando se introduzem os desenhos amostrais propostos descritos nas seções anteriores. Vale ressaltar que cada plano amostral tem sua parametrização particular. Serão apresentados somente os resultados dos parâmetros referentes a distribuição T-Student, parâmetros de interesse do presente estudo e que são comuns a todos os planos amostrais apresentados.

Duas abordagens foram consideradas para a inferência dos modelos: a clássica e a bayesiana, com destaque maior para a última. O estudo sob enfoque clássico foi feito para se ter uma análise preliminar do desempenho dos modelos AQ e AQS comparados aos modelos AASX e AAS.

Com o intuito de simplificar o estudo de simulação, considerou-se $E(Y_i) = \mu, \forall i \in U$, fixando o valor $\mu = 1$. Em relação a σ , foi fixado o valor $\sigma = 1$. Por fim, fixou-se $\nu = 5$, uma vez que os planos propostos foram desenvolvidos para dados com caudas pesadas.

Para todos os cenários de simulações analisados, exceto o caso AEI sob abordagem bayesiana, considerou-se o modelo simplificado:

$$X_i = Y_i + \epsilon_i; \epsilon_i \sim N(0, \sigma_\epsilon^2), i = 1, 2, 3, \dots, N.$$

Um exemplo em que a hipótese de uma relação linear entre Y e X é plausível no contexto de inferência em populações finitas seria quando X e Y são as mesmas variáveis, mas X é obtida por meio de um censo ou de um registro administrativo e Y por meio de

uma pesquisa amostral obtida num momento posterior. Devido ao alto custo financeiro, censos são realizados em grandes intervalos de tempo e as vezes deseja-se pesquisar uma determinada variável num período posterior ao censo. Neste caso, tem-se acesso aos valores de X_i para toda população e o modelo $X_i = Y_i + \epsilon_i$.

Para avaliar o efeito nas estimativas dos parâmetros obtidas para cada plano quando a correlação linear entre X e Y , ρ_{XY} , for considerada baixa, média ou alta, consideraram-se, respectivamente, os seguintes valores de ρ_{XY} : 0,2; 0,5 e 0,8. No caso do estudo de simulação da amostragem estratificada informativa, considerou-se apenas $\rho_{XY} = 0,7$, uma vez que espera-se que a variável de estratificação tenha alta correlação com a variável de interesse.

Vale destacar que, para melhorar a estimação do parâmetro ν (principalmente sob abordagem clássica) e garantir pressupostos dos modelos AQ, AQS e AQSC, assumiu-se que o parâmetro ν está contido no intervalo $(2, 50]$.

5.1 Primeiro estudo de simulação

No presente estudo foram considerados os modelos AQ, AQS, AASX e AAS sob abordagem clássica. Foram simuladas 1000 populações de uma distribuição T-Student, $N = 5000$ de parâmetros $\mu = 1$, $\sigma = 1$ e $\nu = 5$. Para cada uma das 1000 populações simuladas, foram simulados três conjuntos de valores de X , fixando-se $\rho_{XY} = 0,2; 0,5; 0,8$ respectivamente. Para avaliar a estimação dos parâmetros em função do tamanho de amostra, considerou-se $n = 20, 60, 100$. Uma vez definidos os valores de μ , σ e ρ_{XY} , pode-se obter σ_ϵ pela expressão:

$$\sigma_\epsilon = \sigma \sqrt{\frac{\nu}{\nu - 2} \left(\frac{1}{\rho_{XY}^2} - 1 \right)}.$$

Para cada amostra obtida foram calculadas as estimativas dos parâmetros via método de máxima verossimilhança. Além disso, para cada modelo considerado, foram calculados o viés, variância e as raízes dos erros quadráticos médios para o estimador de máxima verossimilhança de cada parâmetro do modelo. Dividiu-se cada medida calculada pelo valor verdadeiro do respectivo parâmetro. Os cálculos destes sumários sobre os estimadores

podem ser obtidos por meio das seguintes fórmulas:

$$\begin{aligned} \text{Viés}_*(\theta_i) &= \frac{\text{Viés}(\theta_i)}{|\theta_i|} = \frac{\frac{\sum_{r=1}^R \hat{\theta}_i^{(r)}}{R} - \theta_i}{|\theta_i|}; \\ \text{Var}_*(\theta_i) &= \frac{\text{Var}(\theta_i)}{\theta_i^2} = \frac{1}{R} \sum_{r=1}^R \left(\frac{\hat{\theta}_i^{(r)} - \frac{\sum_{r=1}^R \hat{\theta}_i^{(r)}}{R}}{\theta_i} \right)^2; \\ \sqrt{\text{EQM}_*(\theta_i)} &= \sqrt{\frac{\text{EQM}(\theta_i)}{\theta_i^2}} = \sqrt{\frac{\sum_{r=1}^R \left(\frac{\hat{\theta}_i^{(r)} - \theta_i}{\theta_i} \right)^2}{R}} = \sqrt{\text{Viés}_*^2(\theta_i) + \text{Var}_*(\theta_i)}; \end{aligned}$$

onde θ_i representa genericamente os parâmetros verdadeiros da distribuição T-Student e $\hat{\theta}_i^{(r)}$ é a estimativa de máxima verossimilhança do parâmetro θ_i obtida na amostra $r = 1, 2, \dots, R$; com $R = 1000$. Todas as simulações foram realizadas no software [R Core Team \(2020\)](#).

Na Tabela 5.1, são apresentados os resultados obtidos das simulações para os parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,2$ e $n = 20, 60, 100$.

Parâmetros	Métodos	$n = 20$			$n = 60$			$n = 100$		
		Viés*	Var.*	$\sqrt{\text{EQM}_*}$	Viés*	Var.*	$\sqrt{\text{EQM}_*}$	Viés*	Var.*	$\sqrt{\text{EQM}_*}$
ν	AQ	3,9213	19,0721	5,8693	1,6033	9,6494	3,4957	0,7972	4,3735	2,2381
	AQS	4,0884	19,1496	5,9887	1,5899	9,5324	3,4728	0,7348	3,8234	2,0889
	AASX	4,3724	20,0041	6,2547	2,2619	13,585	4,3245	1,276	7,9000	3,0868
	AAS	4,5444	20,4563	6,4115	2,1586	13,1262	4,2173	1,2619	7,9659	3,0917
μ	AQ	-0,0057	0,0688	0,2624	0,0048	0,0206	0,1435	0,0041	0,0129	0,1136
	AQS	-0,0028	0,0690	0,2627	0,0019	0,0227	0,1506	0,0006	0,0138	0,1176
	AASX	0,0002	0,0673	0,2594	-0,0065	0,0235	0,1533	-0,0020	0,0131	0,1145
	AAS	-0,0013	0,0678	0,2604	0,0017	0,0232	0,1524	0,0008	0,0143	0,1196
σ	AQ	0,0175	0,0558	0,2370	0,0249	0,0243	0,1578	0,0172	0,0150	0,1237
	AQS	0,0222	0,0545	0,2345	0,0251	0,0235	0,1553	0,0202	0,0141	0,1203
	AASX	-0,0054	0,0628	0,2506	0,0150	0,0261	0,1623	0,0130	0,0159	0,1268
	AAS	-0,0068	0,0606	0,2463	0,0060	0,0249	0,1579	0,0066	0,0158	0,1258

Tabela 5.1: Resultados das simulações para os modelos AQ, AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,2$ e $n = 20, 60, 100$ (abordagem clássica).

Na Tabela 5.2, são apresentados os resultados obtidos das simulações para os parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,5$ e $n = 20, 60, 100$.

Parâmetros	Métodos	$n = 20$			$n = 60$			$n = 100$		
		Viés*	Var.*	$\sqrt{EQM_*}$	Viés*	Var.*	$\sqrt{EQM_*}$	Viés*	Var.*	$\sqrt{EQM_*}$
ν	AQ	2,3936	14,1966	4,4639	0,5324	2,5686	1,6888	0,2395	0,6543	0,8436
	AQS	2,3093	13,3429	4,3215	0,4763	2,1283	1,5347	0,2017	0,4363	0,6906
	AASX	4,3724	20,0042	6,2547	2,2619	13,5851	4,3245	1,2760	7,8999	3,0868
	AAS	4,5444	20,4563	6,4115	2,1586	13,1262	4,2173	1,2619	7,9659	3,0917
μ	AQ	0,0034	0,0729	0,2700	-0,0012	0,0240	0,1551	0,0035	0,0143	0,1196
	AQS	0,0119	0,0566	0,2382	0,0034	0,0190	0,1380	-0,0002	0,0119	0,1089
	AASX	0,0002	0,0673	0,2594	-0,0065	0,0235	0,1533	-0,0020	0,0131	0,1145
	AAS	-0,0013	0,0678	0,2604	0,0017	0,0232	0,1524	0,0008	0,0143	0,1196
σ	AQ	0,0251	0,0500	0,2251	0,0196	0,0186	0,1377	0,0133	0,0114	0,1075
	AQS	0,0321	0,0516	0,2294	0,0193	0,0174	0,1332	0,0134	0,0103	0,1022
	AASX	-0,0054	0,0628	0,2506	0,0150	0,0261	0,1623	0,0130	0,0159	0,1268
	AAS	-0,0068	0,0606	0,2463	0,0060	0,0249	0,1579	0,0066	0,0158	0,1258

Tabela 5.2: Resultados das simulações para os modelos AQ, AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,5$ e $n = 20, 60, 100$ (abordagem clássica).

Na Tabela 5.3, são apresentados os resultados obtidos das simulações para os parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,8$ e $n = 20, 60, 100$.

Parâmetros	Métodos	$n = 20$			$n = 60$			$n = 100$		
		Viés*	Var.*	$\sqrt{EQM_*}$	Viés*	Var.*	$\sqrt{EQM_*}$	Viés*	Var.*	$\sqrt{EQM_*}$
ν	AQ	1,4552	8,7765	3,3006	0,2435	0,8573	0,9574	0,1389	0,2035	0,4720
	AQS	1,1627	7,0792	2,9036	0,2176	0,5631	0,7813	0,1287	0,1326	0,3862
	AASX	4,3724	20,0043	6,2547	2,2619	13,585	4,3245	1,2760	7,8999	3,0868
	AAS	4,5444	20,4563	6,4115	2,1586	13,1262	4,2173	1,2619	7,9659	3,0917
μ	AQ	0,0104	0,0831	0,2884	0,0016	0,0298	0,1726	0,0010	0,0176	0,1325
	AQS	-0,0193	0,0496	0,2236	-0,0085	0,0148	0,1218	-0,0052	0,0092	0,0962
	AASX	0,0002	0,0673	0,2594	-0,0065	0,0235	0,1533	-0,0020	0,0131	0,1145
	AAS	-0,0013	0,0678	0,2604	0,0017	0,0232	0,1524	0,0008	0,0143	0,1196
σ	AQ	0,0316	0,0512	0,2284	0,0137	0,0166	0,1295	0,0158	0,0100	0,1015
	AQS	0,0539	0,0529	0,2363	0,0266	0,0179	0,1364	0,0195	0,0111	0,1071
	AASX	-0,0054	0,0628	0,2506	0,0150	0,0261	0,1623	0,0130	0,0159	0,1268
	AAS	-0,0068	0,0606	0,2463	0,0060	0,0249	0,1579	0,0066	0,0158	0,1258

Tabela 5.3: Resultados das simulações para os modelos AQ, AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,8$ e $n = 20, 60, 100$ (abordagem clássica).

Nas Figuras 5.1, 5.2 e 5.3 são apresentados os boxplots das estimativas obtidas dos parâmetros ν , μ e σ , respectivamente. Destaca-se a Figura 5.1, na qual observa-se uma

melhora apreciável nas estimativas de ν ao adotar-se os modelos AQ e AQS em comparação aos demais modelos.

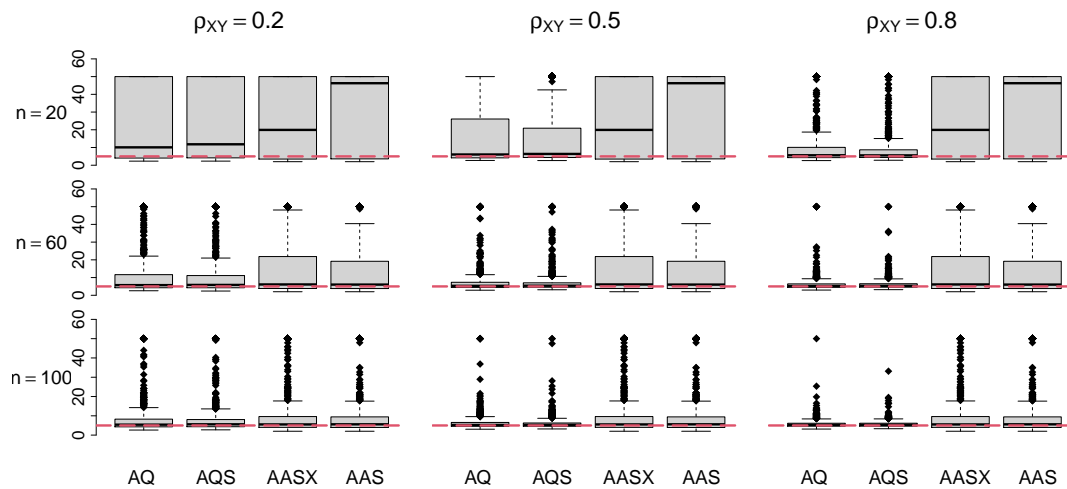


Figura 5.1: Boxplots das estimativas de máxima verossimilhança de ν obtidas nas simulações, considerando-se os modelos AQ, AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,2; 0,5; 0,8$ e $n = 20, 60, 100$.

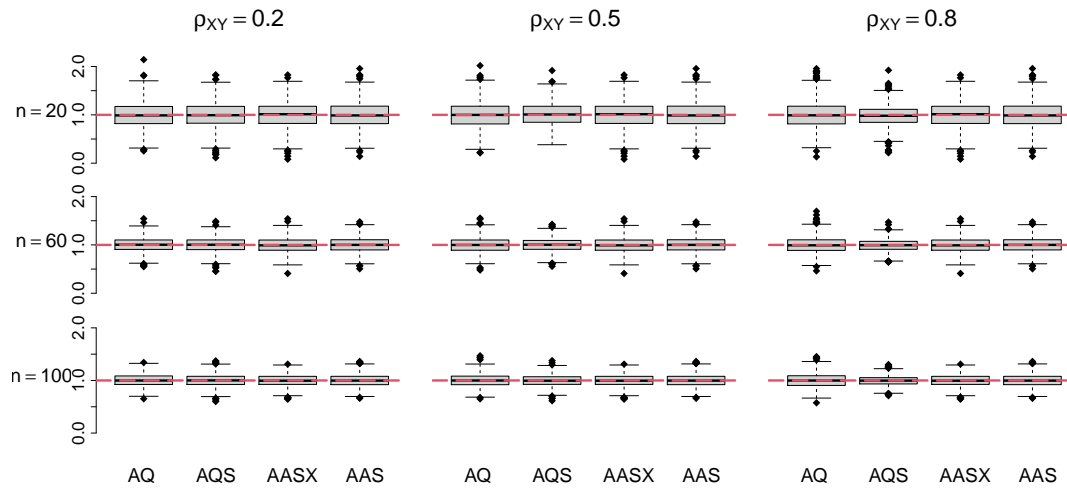


Figura 5.2: Boxplots das estimativas de máxima verossimilhança de μ obtidas nas simulações, considerando-se os modelos AQ, AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,2; 0,5; 0,8$ e $n = 20, 60, 100$.

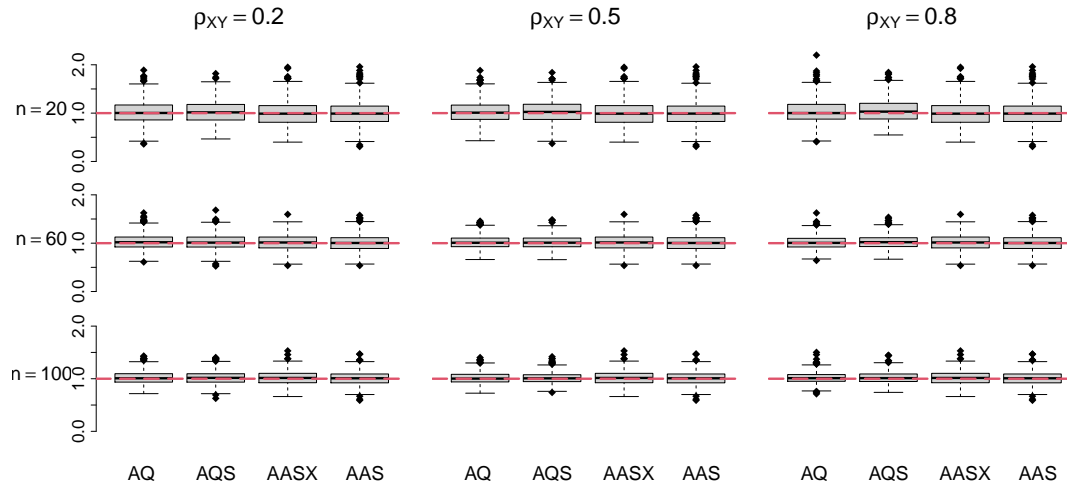


Figura 5.3: Boxplots das estimativas de máxima verossimilhança de σ obtidas nas simulações, considerando-se os modelos AQ, AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,2; 0,5; 0,8$ e $n = 20, 60, 100$.

O modelo AQS teve o melhor desempenho em comparação aos demais. Observando-se as Tabelas 5.1, 5.2 e 5.3, nota-se que à medida que ρ_{XY} e o tamanho da amostra aumentam, há uma melhora apreciável nas estimativas de ν , com uma redução significativa tanto no viés quanto na variância. Como esperado, destaque para o caso particular $\rho_{XY} = 0,8$ e $n = 100$, com um desempenho muito melhor na estimação de ν em comparação a AAS. Analisando as Figuras 5.2 e 5.3, nota-se que não houve perda de eficiência na estimação dos demais parâmetros. Em relação ao modelo AQ, embora tenha melhorado as estimativas de ν , houve uma perda de eficiência na estimação do parâmetro μ . À medida que ρ_{XY} aumenta, há uma piora nas estimativas de μ , com o aumento da variância. Entretanto, não há um aumento significativo no viés.

5.2 Segundo estudo de simulação

Esta seção será inteiramente dedicada à apresentação dos resultados obtidos nas simulações sob os modelos AEI, AQS, AQSC, AASX, AASXC e AAS considerando-se a estimação sob o enfoque bayesiano. Para o modelo AQ não se obteve um bom desempenho nas simulações apresentadas sob a abordagem clássica, principalmente com relação

a estimação de μ . Por este motivo ele não foi considerado nas simulações que serão apresentadas nesta seção.

Uma vantagem da abordagem bayesiana em relação à abordagem clássica é a possibilidade de serem introduzidas informações previamente conhecidas pelo pesquisador através da distribuição *a priori*. Por exemplo, em relação a ν , o pesquisador pode considerar uma distribuição *a priori* que tem massa de probabilidade mais concentrada em valores pequenos de ν , permitindo então lidar com expectativas de caudas mais pesadas. Introduzir esta informação na distribuição *a priori* pode ajudar na estimação dos parâmetros. Contudo, neste estudo de simulação foi considerada uma distribuição *a priori* para ν pouco informativa, proposta por Geweke (1993), $\pi(\nu) \propto \kappa \exp\{-\kappa\nu\} I_{(a,b]}(\nu)$, $\kappa = 0,1$. Para os parâmetros μ e σ foram consideradas as distribuições *a priori*: $\mu \sim N(m_0, s_0^2)$; $m_0 = 0$, $s_0^2 = 100^2$ e $\sigma \sim \text{Half-Cauchy}(m_1, s_1^2)$; $m_1 = 0$ e $s_1^2 = 25^2$ (Gelman, 2006).

Todas as simulações que serão apresentadas foram realizadas no software R Core Team (2020), utilizando o pacote RStan (Stan Development Team, 2022).

Assim como Fonseca et al. (2008), no presente trabalho foi considerada como estimativa pontual para os parâmetros da distribuição T-Student a mediana *a posteriori*.

Por fim, destaca-se que outra distribuição *a priori* para ν também foi considerada. Para maiores detalhes, ver Anexo B.

5.2.1 Comparação entre os modelos AQS, AASX e AAS, e comparação entre os modelos AQSC e AASXC

Nesta seção são apresentados os resultados obtidos para os modelos AQS, AASX, AAS, AQSC e AASXC sob abordagem bayesiana. Com relação aos modelos AQS, AASX, AQSC e AASXC, foi empregada a seguinte distribuição *a priori*, proposta por Gelman (2006), $\sigma_\epsilon \sim \text{Half-Cauchy}(m_1, s_1^2)$, com $m_1 = 1$ e $s_1^2 = 25^2$.

No presente cenário de simulação, foram utilizadas as mesmas populações simuladas anteriormente da distribuição T-Student no estudo sob abordagem clássica, com os mesmos parâmetros fixados anteriormente e considerando-se os mesmos cenários. Em uma primeira etapa, compararam-se os resultados dos modelos AQS, AASX e AAS, depois

foram comparados os resultados dos modelos AQSC e AASXC.

Para cada população r simulada, foram obtidas duas cadeias, com valores iniciais diferentes para cada parâmetro. Foram realizadas 40000 iterações, utilizando-se um período de aquecimento de 10000 iterações, ou seja, foram descartadas as primeiras 10000 iterações de ambas as cadeias. Para eliminar a correlação entre os valores restantes foi aplicado o espaçamento de 30 iterações em cada cadeia. Ao final, amostras das distribuições *a posteriori* de tamanho 2000 foram obtidas para μ , σ e ν para cada uma das 1000 amostras para cada modelo considerado.

Foram calculadas as medianas, desvios-padrão, comprimentos dos intervalos HPDs de 95% (Migon et al., 2014) de cada uma das 1000 amostras das distribuições *a posteriori* obtidas dos parâmetros da distribuição T-Student. Assim, para cada parâmetro, tem-se um conjunto de 1000 valores de cada uma das medidas mencionadas acima para os modelos AQS, AASX e AAS. Por fim foram calculadas a média das medianas, média dos desvios-padrão, média dos comprimentos do intervalo, a proporção de quantos intervalos HPD contêm os verdadeiros valores dos parâmetros e o erro absoluto relativo para as estimativas obtidas. Os cálculos destes sumários sobre os estimadores podem ser obtidos por meio das seguintes fórmulas:

$$\begin{aligned} \text{Med}_*(\theta_i) &= \frac{1}{R} \sum_{r=1}^R \hat{\theta}_i^{(r)}; \quad \text{DP}_*(\theta_i) = \frac{1}{R} \sum_{r=1}^R \text{DP} \left(\theta_i^{(r)} | \mathbf{y}_s, \mathbf{x}_s \right); \\ \text{Comp}_*(\theta_i) &= \frac{1}{R} \sum_{r=1}^R \text{Comp} \left(\theta_i^{(r)} | \mathbf{y}_s, \mathbf{x}_s \right); \quad \text{Cob}_*(\theta_i) = \frac{1}{R} \sum_{r=1}^R \text{Cob} \left(\theta_i^{(r)} | \mathbf{y}_s, \mathbf{x}_s \right); \\ \text{EAR}(\theta_i) &= \frac{1}{R} \sum_{r=1}^R \left| \frac{\hat{\theta}_i^{(r)} - \theta_i}{\theta_i} \right|; \end{aligned}$$

onde θ_i representa genericamente os parâmetros verdadeiros da distribuição T-Student, $\hat{\theta}_i^{(r)}$ é a mediana *a posteriori* do parâmetro θ_i obtida na amostra r , $\text{DP} \left(\theta_i^{(r)} | \mathbf{y}_s, \mathbf{x}_s \right)$ e $\text{Comp} \left(\theta_i^{(r)} | \mathbf{y}_s, \mathbf{x}_s \right)$ são respectivamente o desvio-padrão e o comprimento do intervalo HPD da distribuição *a posteriori* do parâmetro θ_i referente à amostra r e $\text{Cob} \left(\theta_i^{(r)} | \mathbf{y}_s, \mathbf{x}_s \right)$ é uma função indicadora que assume o valor 1 caso o intervalo HPD contenha o verdadeiro valor θ_i .

Nas Tabelas 5.4, 5.5 e 5.6 são apresentados os resultados obtidos.

Parâmetros	Métodos	$n = 20$					$n = 60$					$n = 100$				
		Med*	DP*	Comp*	Cob*	EAR	Med*	DP*	Comp*	Cob*	EAR	Med*	DP*	Comp*	Cob*	EAR
ν	AQS	9,8735	9,3000	28,7221	0,9980	1,0048	8,6449	6,9111	21,4203	0,9870	0,7906	7,7805	5,3525	16,6899	0,9830	0,6211
	AASX	9,2794	8,5554	27,6396	0,9990	0,8860	8,7197	7,1452	23,2498	0,9930	0,8227	8,0844	6,0323	19,6614	0,9820	0,7158
	AAS	9,2992	8,5270	27,5867	0,9990	0,8925	8,6878	7,0945	23,0979	0,9930	0,8187	8,0557	6,0011	19,5419	0,9840	0,7112
μ	AQS	0,9972	0,2728	1,068	0,9530	0,2074	1,0019	0,1521	0,5922	0,9450	0,1176	1,0004	0,1168	0,4544	0,9470	0,0929
	AASX	0,9989	0,2843	1,1145	0,9590	0,2056	0,9932	0,1550	0,6044	0,9500	0,1202	0,9978	0,1186	0,4621	0,9520	0,0917
	AAS	0,9981	0,2839	1,1134	0,9600	0,2024	1,0007	0,1542	0,6008	0,9430	0,1197	1,0004	0,1180	0,4597	0,9390	0,0947
σ	AQS	1,1135	0,2317	0,8867	0,9580	0,1913	1,0751	0,1386	0,5363	0,9210	0,1247	1,0604	0,1109	0,4296	0,9350	0,1013
	AASX	1,1132	0,2540	0,9630	0,9570	0,1989	1,0660	0,1441	0,5583	0,9490	0,1217	1,0517	0,1153	0,4469	0,9410	0,0998
	AAS	1,1094	0,2526	0,9578	0,9580	0,1974	1,0586	0,1431	0,5537	0,9470	0,1165	1,0460	0,1148	0,4451	0,9470	0,0960

Tabela 5.4: Resultados das simulações para os modelos AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,2$ e $n = 20, 60, 100$ (abordagem bayesiana).

Parâmetros	Métodos	$n = 20$					$n = 60$					$n = 100$				
		Med*	DP*	Comp*	Cob*	EAR	Med*	DP*	Comp*	Cob*	EAR	Med*	DP*	Comp*	Cob*	EAR
ν	AQS	9,3247	7,7262	24,0033	0,9970	0,9118	7,2586	4,0953	12,9337	0,9750	0,5082	6,4265	2,5228	8,2425	0,9620	0,3462
	AASX	9,2796	8,5497	27,6221	0,9990	0,8861	8,7238	7,1572	23,3010	0,9940	0,8240	8,0703	6,0287	19,6588	0,9810	0,7137
	AAS	9,2992	8,5270	27,5867	0,9990	0,8925	8,6878	7,0945	23,0979	0,9930	0,8187	8,0557	6,0011	19,5419	0,9840	0,7112
μ	AQS	1,0111	0,2462	0,9605	0,9500	0,1914	1,0031	0,1389	0,5406	0,9550	0,1095	0,9997	0,1067	0,4152	0,9450	0,0863
	AASX	0,9988	0,2844	1,1152	0,9640	0,2051	0,9930	0,1549	0,6036	0,9500	0,1203	0,9976	0,1187	0,4624	0,9510	0,0916
	AAS	0,9981	0,2839	1,1134	0,9600	0,2024	1,0007	0,1542	0,6008	0,9430	0,1197	1,0004	0,1180	0,4597	0,9390	0,0947
σ	AQS	1,1180	0,2083	0,8026	0,9430	0,1869	1,0686	0,1275	0,4930	0,9300	0,1159	1,0469	0,0998	0,3862	0,9190	0,0880
	AASX	1,1128	0,2539	0,9632	0,9530	0,1988	1,0659	0,1442	0,5583	0,9510	0,1217	1,0518	0,1153	0,4470	0,9410	0,0998
	AAS	1,1094	0,2526	0,9578	0,9580	0,1974	1,0586	0,1431	0,5537	0,9470	0,1165	1,0460	0,1148	0,4451	0,9470	0,0960

Tabela 5.5: Resultados das simulações para os modelos AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,5$ e $n = 20, 60, 100$ (abordagem bayesiana).

Parâmetros	Métodos	$n = 20$					$n = 60$					$n = 100$				
		Med*	DP*	Comp*	Cob*	EAR	Med*	DP*	Comp*	Cob*	EAR	Med*	DP*	Comp*	Cob*	EAR
ν	AQS	8,2966	5,9180	18,3711	0,9780	0,7188	6,4509	2,4656	8,0364	0,9500	0,3581	5,9385	1,4595	5,0280	0,9190	0,2540
	AASX	9,2756	8,5418	27,5931	0,9990	0,8854	8,7263	7,1552	23,2990	0,9930	0,8248	8,0746	6,0276	19,6397	0,9820	0,7140
	AAS	9,2992	8,5270	27,5867	0,9990	0,8925	8,6878	7,0945	23,0979	0,9930	0,8187	8,0557	6,0011	19,5419	0,9840	0,7112
μ	AQS	0,9807	0,2292	0,8947	0,9420	0,1769	0,9915	0,1249	0,4860	0,9530	0,0987	0,9947	0,0954	0,3714	0,9350	0,0755
	AASX	0,9986	0,2843	1,1146	0,9590	0,2052	0,9932	0,1550	0,6042	0,9460	0,1204	0,9976	0,1186	0,4623	0,9540	0,0915
	AAS	0,9981	0,2839	1,1134	0,9600	0,2024	1,0007	0,1542	0,6008	0,9430	0,1197	1,0004	0,1180	0,4597	0,9390	0,0947
σ	AQS	1,1574	0,2131	0,8187	0,8950	0,2085	1,0741	0,1247	0,4816	0,8940	0,1226	1,0488	0,0960	0,3719	0,8960	0,0931
	AASX	1,1126	0,2538	0,9621	0,9550	0,1988	1,0661	0,1442	0,5584	0,9480	0,1217	1,0518	0,1153	0,4472	0,9410	0,0998
	AAS	1,1094	0,2526	0,9578	0,9580	0,1974	1,0586	0,1431	0,5537	0,9470	0,1165	1,0460	0,1148	0,4451	0,9470	0,0960

Tabela 5.6: Resultados das simulações para os modelos AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,8$ e $n = 20, 60, 100$ (abordagem bayesiana).

Nas Figura 5.4, Figura 5.5 e Figura 5.6 são apresentados os boxplots das estimativas obtidas dos parâmetros ν , μ e σ respectivamente.

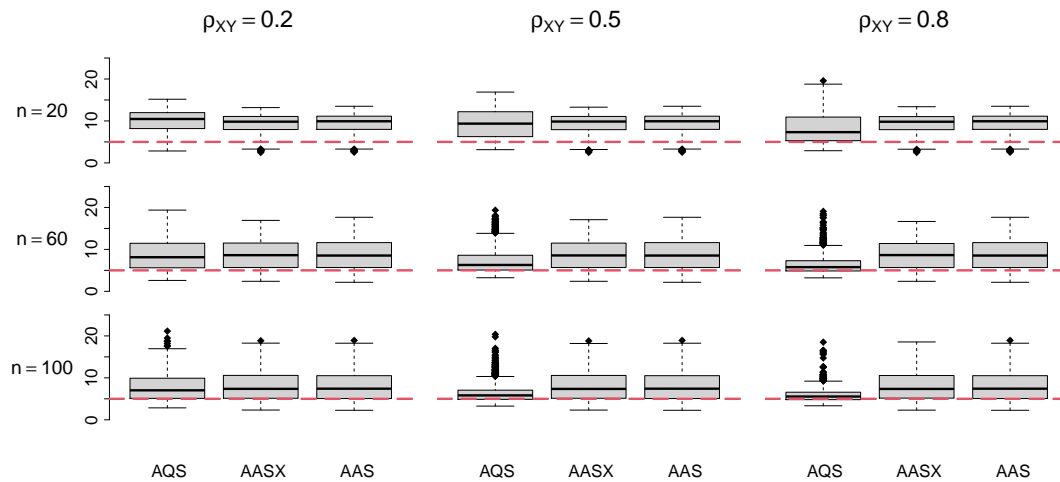


Figura 5.4: Boxplots das estimativas bayesianas pontuais de ν obtidas nas simulações, considerando-se os modelos AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,2; 0,5; 0,8$ e $n = 20, 60, 100$.

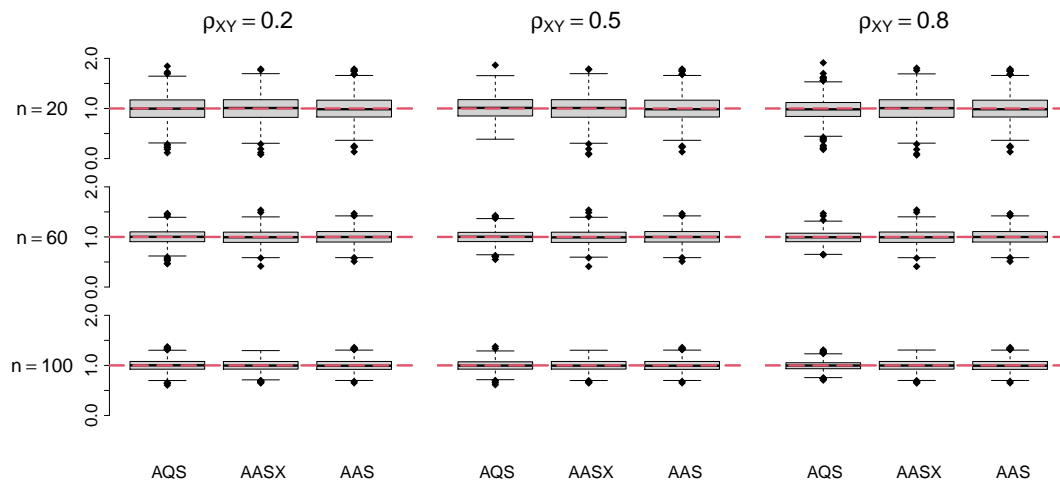


Figura 5.5: Boxplots das estimativas bayesianas pontuais de μ obtidas nas simulações, considerando-se os modelos AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,2; 0,5; 0,8$ e $n = 20, 60, 100$.

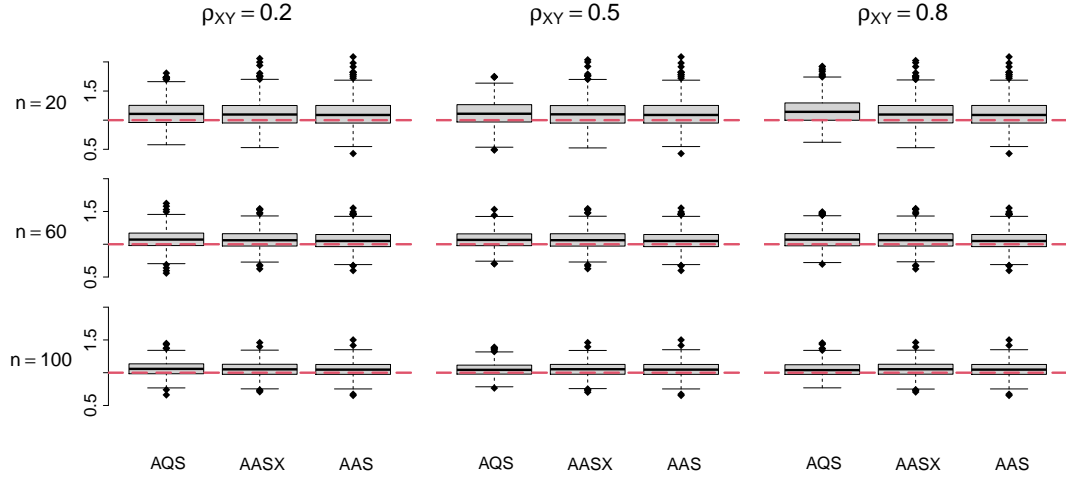


Figura 5.6: Boxplots das estimativas bayesianas pontuais de σ obtidas nas simulações, considerando-se os modelos AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,2$; $0,5$; $0,8$ e $n = 20, 60, 100$.

Analisando-se os resultados apresentados nas Tabelas 5.4, 5.5 e 5.6, nota-se que, no geral, o modelo AQS obteve o melhor desempenho. Vale ressaltar que, com a introdução da distribuição *a priori* pouco informativa para ν , houve uma melhora significativa na estimação deste parâmetro em comparação à abordagem clássica. Esta melhora pode ser verificada comparando-se as Figuras 5.1 e 5.4. Também não houve uma perda significativa na estimação de μ . Referente ao parâmetro σ , foi constatada uma redução na cobertura, à medida que ρ_{XY} aumenta. Entretanto, o comprimento médio dos intervalos são menores em comparação aos modelos AAS e AASX.

Em relação à comparação dos modelos AQSC e AASXC, considerou-se o caso $\nu = 5$, $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0.8$ e $n = 20$. Para este cenário, o modelo AQS não obteve um bom desempenho sob abordagem bayesiana. Em virtude da alta quantidade de cadeias simuladas referentes aos parâmetros λ_i , o custo computacional dos modelos completos são muito maiores que os simplificados. Das 1000 populações simuladas, referente ao parâmetro ν , foram considerados seis cenários para a simulação: os dois cenários em que se obtiveram os menores EAR; os dois cenários em que se obtiveram os maiores EAR; e os dois cenários medianos, referente ao modelo AQS.

Casos	Parâmetros	Métodos	Maiores EAR					Medianos EAR					Menores EAR				
			Med	DP	Comp	Cob	EAR	Med	DP	Comp	Cob	EAR	Med	DP	Comp	Cob	EAR
1	ν	AQS	19,5796	11,3289	38,1376	1	2,9159	7,3228	5,8214	16,6906	1	0,4646	5,0009	2,0759	6,3460	1	0,0002
		AQSC	6,5048	0,9323	3,5146	1	0,3010	4,5459	0,7612	2,9140	1	0,0908	4,6376	0,6289	2,4252	1	0,0725
		AASXC	6,0180	0,9126	3,4484	1	0,2036	4,6076	0,8930	3,3668	1	0,0785	5,7723	0,8996	3,3764	1	0,1545
	μ	AQS	1,0438	0,2464	0,9820	1	0,0438	1,1574	0,2246	0,8844	1	0,1574	0,7116	0,2364	0,9288	1	0,2884
		AQSC	1,0127	0,02217	0,0862	1	0,01274	0,9909	0,0220	0,0860	1	0,0091	0,9935	0,0223	0,0862	1	0,0065
		AASXC	1,0129	0,0226	0,0883	1	0,0129	0,9931	0,0226	0,0867	1	0,0069	0,9955	0,0224	0,0853	1	0,0045
	σ	AQS	1,5122	0,17067	0,6522	0	0,5122	1,0731	0,2061	0,8051	1	0,0731	1,0916	0,2132	0,7983	1	0,0916
		AQSC	1,1522	0,0565	0,2148	0	0,1522	0,8721	0,0923	0,3562	1	0,1279	0,9724	0,0746	0,2854	1	0,0276
		AASXC	1,1097	0,0770	0,3035	1	0,1097	0,8770	0,1216	0,4634	1	0,1230	1,1338	0,0886	0,3422	1	0,1338
2	ν	AQS	18,7739	11,4346	38,7856	1	2,7548	7,3280	4,9842	15,3527	1	0,4656	5,0016	2,7860	7,6920	1	0,0003
		AQSC	4,9181	0,6501	2,5118	1	0,0164	4,0719	0,6031	2,2987	1	0,1856	5,1998	0,8129	3,0527	1	0,0400
		AASXC	4,0718	0,6674	2,5850	1	0,1856	5,1628	0,91980	3,5674	1	0,03256	5,8654	1,0377	3,9712	1	0,1731
	μ	AQS	1,1259	0,2290	0,9021	1	0,12592	1,2614	0,2334	0,9175	1	0,2614	0,9951	0,2232	0,8746	1	0,0049
		AQSC	0,9754	0,0221	0,0866	1	0,0246	1,0104	0,0225	0,0872	1	0,0104	0,9754	0,0221	0,0866	1	0,02456
		AASXC	0,9755	0,0216	0,0826	1	0,0245	1,0084	0,02165	0,0827	1	0,0084	0,9886	0,0218	0,0846	1	0,0114
	σ	AQS	1,3969	0,1694	0,6666	0	0,3969	1,1445	0,2031	0,7694	1	0,1445	1,0734	0,2424	0,9190	1	0,0734
		AQSC	0,9860	0,0711	0,2766	1	0,0140	0,8523	0,0922	0,3497	1	0,1477	0,9933	0,0818	0,3159	1	0,0067
		AASXC	0,8454	0,1064	0,4059	1	0,1546	1,0302	0,1165	0,4395	1	0,0302	1,0830	0,1006	0,3757	1	0,0830

Tabela 5.7: Comparação dos modelos AQSC e AASXC com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,8$ e $n = 20$ (abordagem bayesiana).

Como esperado, os resultados das simulações para os modelos completos AQSC e AASXC foram mais satisfatórios do que os obtidos pelos modelos simplificados. Comparando-se os modelos AQSC e AASXC, verificou-se que, de cinco das seis populações simuladas, o modelo AQSC obteve comprimento dos seus intervalos de credibilidade menores para os parâmetros ν e σ . Novamente, comparando-se os modelos AQSC e AASXC não foi detectada nenhuma diferença significativa na estimação de μ .

5.2.2 Comparação entre os modelos AEI, AQS e AAS

Nesta seção são apresentados os resultados obtidos referentes aos modelos AEI, AQS e AAS sob abordagem bayesiana. Em relação ao modelo AEI, considerou-se as seguintes distribuições *a priori* relativamente vagas para os parâmetros: $\eta_k \sim N(m_0, s_0^2)$, $k = 1, \dots, H - 1$; e $\eta_H \sim N(m_0, s_0^2)$; $m_0 = 0$, $s_0^2 = 100^2$.

Definiu-se a quantidade de estratos (H) igual a 4 e 6 para o presente estudo de simulação. No caso $H = 4$ (AEI1) os pontos de cortes c_k foram definidos pelos quantis da distribuição de X . No caso $H = 6$ (AEI2) os pontos de corte c_k foram definidos pelos quantis e os percentis 10 e 90 da distribuição de X . Os tamanhos das amostras n_h dos estratos considerados para os casos em que $H = 4, 6$ são apresentados na Tabela 5.8.

Modelos	n_h	n		
		20	60	100
AEI1	n_1	8	20	40
	n_2	2	10	10
	n_3	2	10	10
	n_4	8	20	40
AEI2	n_1	5	15	25
	n_2	3	10	15
	n_3	2	5	10
	n_4	2	5	10
	n_5	3	10	15
	n_6	5	15	25

Tabela 5.8: Tamanhos das amostras n_h dos estratos referente ao plano amostral amostragem estratificada informativa.

A título de comparação com o modelo AEI, referente ao modelo AQS, foi considerado o seguinte modelo linear:

$$X_i = b_0 + b_1 Y_i + \epsilon_i; \epsilon_i \sim N(0, \sigma_\epsilon^2), i = 1, 2, 3, \dots, N.$$

onde o valor de σ_ϵ pode ser obtido como:

$$\sigma_\epsilon = b_1 \sigma \sqrt{\frac{\nu}{\nu - 2} \left(\frac{1}{\rho_{XY}^2} - 1 \right)}.$$

Foram consideradas as seguintes distribuições *a priori* para b_0 e b_1 : $b_0 \sim N(m_0, s_0^2)$; $b_1 \sim N(m_1, s_1^2)$; onde $m_0 = 0$, $s_0^2 = 100^2$.

Foram simuladas $R = 1000$ populações da distribuição T-Student com tamanho populacional $N = 5000$, fixando-se os seguintes valores para os parâmetros do modelo: $b_0 = 1$; $b_1 = 2$; $\sigma = 1$; $\nu = 5$ e $\rho_{XY} = 0,7$. Consideraram-se os mesmos tamanhos de amostra empregados para os outros modelos.

Referentes aos casos AEI1 e AEI2, para cada população r , obtiveram-se duas cadeias com valores iniciais diferentes para cada parâmetro. Foram realizadas 15000 iterações, utilizando-se um período de aquecimento de 5000 iterações. Considerou-se um espaçamento de 10 iterações para cada cadeia. Referente aos casos AQS e AAS foram obtidas duas cadeias de 40000 iterações cada, com aquecimento de 10000 e o espaçamento de 30 em cada cadeia. Ao final, em todos os casos, amostras das distribuições *a posteriori* de tamanho 2000 foram obtidas para ν , μ e σ em 1000 populações simuladas.

Foram calculadas as medianas, desvios-padrão, comprimentos dos intervalos HPDs de cada uma das 1000 amostras das distribuições *a posteriori* obtidas dos parâmetros da distribuição T-Student. Assim, para cada parâmetro, tem-se um conjunto de 1000 valores de cada uma das medidas mencionadas acima para os modelos AEI1, AEI2, AQS e AAS. Por fim foram calculadas a média das medianas, média dos desvios-padrão, média dos comprimentos do intervalo, a proporção de quantos intervalos HPDs que contêm os verdadeiros valores dos parâmetros e o erro absoluto relativo para as estimativas obtidas.

Na Tabela 5.9 são apresentados os resultados obtidos nas simulações.

Parâmetros	Métodos	$n = 20$					$n = 60$					$n = 100$				
		Med*	DP*	Comp*	Cob*	EAR	Med*	DP*	Comp*	Cob*	EAR	Med*	DP*	Comp*	Cob*	EAR
ν	AEI1	8,7263	8,1913	26,1957	0,9980	0,7950	8,2188	6,7194	21,7045	0,9920	0,7396	7,0895	5,0899	16,3639	0,9820	0,5582
	AEI2	8,2096	7,7969	24,7343	0,9970	0,7126	6,9675	5,5679	17,6169	0,9780	0,5620	6,4620	4,3301	13,8755	0,9620	0,4672
	AQS	8,8438	6,8309	21,2116	0,9870	0,8171	6,6400	2,9321	9,3603	0,9590	0,3909	6,0759	1,7661	5,9446	0,9300	0,2769
	AAS	9,2992	8,5270	27,5867	0,9990	0,8925	8,6878	7,0945	23,0979	0,9930	0,8187	8,0557	6,0011	19,5419	0,9840	0,7112
μ	AEI1	1,0013	0,2728	1,0688	0,9830	0,1710	1,0019	0,1497	0,5834	0,9780	0,0990	1,0040	0,1137	0,4424	0,9710	0,0791
	AEI2	0,9964	0,2593	1,0147	0,9680	0,1805	0,9987	0,1469	0,5724	0,9720	0,1002	1,0016	0,1130	0,4398	0,9720	0,0808
	AQS	0,9881	0,2803	1,0959	0,9460	0,2110	0,9986	0,1538	0,5987	0,9450	0,1216	0,9970	0,1178	0,4580	0,9440	0,0968
	AAS	0,9981	0,2839	1,1134	0,9600	0,2024	1,0007	0,1542	0,6008	0,9430	0,1197	1,0004	0,1180	0,4597	0,9390	0,0947
σ	AEI1	1,1424	0,2471	0,9417	0,9680	0,1976	1,0720	0,1400	0,5418	0,9430	0,1161	1,0604	0,1123	0,4349	0,9240	0,0976
	AEI2	1,1365	0,2287	0,8761	0,9470	0,1884	1,0949	0,1388	0,5365	0,9380	0,1248	1,0786	0,1099	0,4256	0,9090	0,1041
	AQS	1,1717	0,2352	0,9013	0,9260	0,2288	1,0796	0,1351	0,5215	0,9060	0,1305	1,0556	0,1037	0,4012	0,9150	0,0980
	AAS	1,1094	0,2526	0,9578	0,9580	0,1974	1,0586	0,1431	0,5537	0,9470	0,1165	1,0460	0,1148	0,4451	0,9470	0,0960

Tabela 5.9: Resultados das simulações para os modelos AEI1, AEI2, AQS e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,7$ e $n = 20, 60, 100$ (abordagem bayesiana).

A Figura 5.7 apresenta boxplots das estimativas obtidas dos parâmetros ν , μ e σ respectivamente.

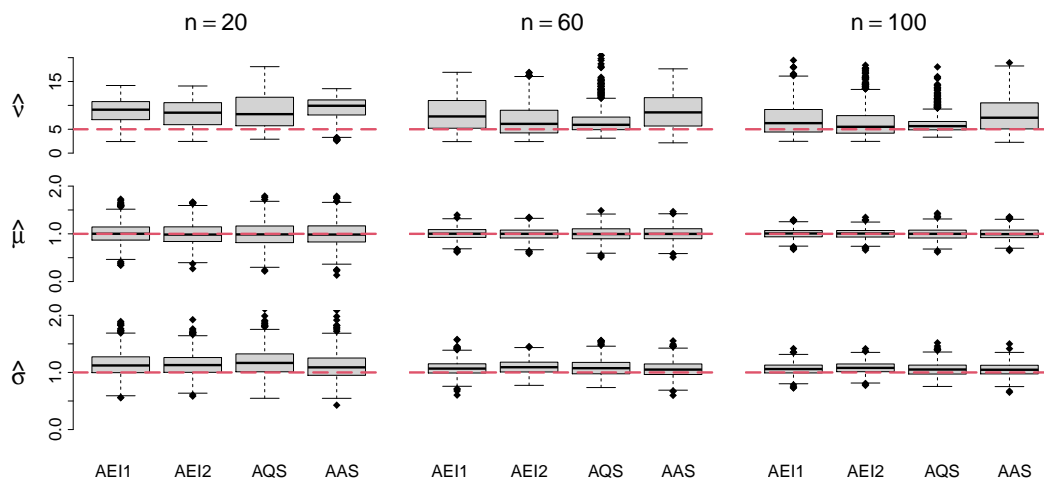


Figura 5.7: Boxplots das estimativas bayesianas pontuais de ν , μ e σ obtidas nas simulações, considerando-se os modelos AEI1, AEI2, AQS e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,7$ e $n = 20, 60, 100$.

Na comparação dos modelos AEI, AQS e AAS, o modelo AQS também teve o melhor desempenho na estimação de ν em comparação aos demais. Analisando-se a Figura 5.7, nota-se que à medida que o tamanho da amostra aumenta, as estimativas pontuais obtidas de ν são melhores em comparação aos demais modelos. Além disso, em todos os cenários, o modelo AQS teve o menor comprimento médio dos intervalos, conforme verificado na Tabela 5.9. Vale ressaltar que o modelo AEI com seis estratos teve um melhor desempenho na estimação de ν do que o modelo AEI com quatro estratos. Isso vem reforçar a hipótese da necessidade de se amostrar mais valores das caudas em dados com caudas pesadas.

Capítulo 6

Aplicação

Considerou-se para aplicação dos planos amostrais propostos o conjunto de dados do censo de agricultura realizado nos Estados Unidos (Lohr Sharon, 1999). Este conjunto de dados é composto por $N = 3078$ unidades amostrais (condados¹) e classificadas hierarquicamente em estados. As seguintes variáveis podem ser extraídas deste arquivo para os anos de 1982, 1987 e 1992: números de fazendas; áreas dedicadas às fazendas; números de fazendas com menos de 9 acres e números de fazendas com mais de 1000 acres. Destaca-se que uma unidade acre é definida como um pedaço de terra com aproximadamente $4047 m^2$ (Melo et al., 2018).

Assim como em Lohr Sharon (1999), a fim de facilitar a estimação dos parâmetros de interesse e melhorar a apresentação dos resultados, no presente estudo dividiram-se as variáveis quantitativas por 1500000.

A abordagem considerada para a inferência foi a bayesiana. Novamente todas as amostras das distribuições *a posteriori* foram obtidas no software R Core Team (2020), utilizando o pacote RStan (Stan Development Team, 2022).

A variável de interesse Y considerada foi a área dedicada a fazenda no ano de 1992 e a covariável X utilizada foi esta mesma variável medida no ano de 1987. Ressalta-se que existem 34 condados com pelos menos um valor destas variáveis ausentes, sendo excluídos do estudo. Abaixo na Figura 6.1 é apresentado o gráfico de dispersão das respectivas variáveis.

¹Regiões administrativas dos EUA que correspondem aos municípios no Brasil.

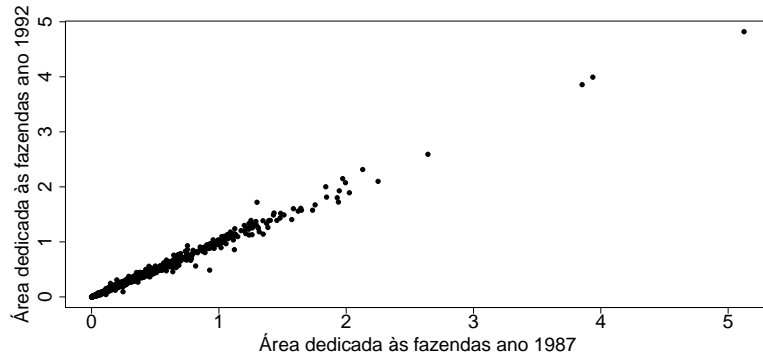


Figura 6.1: Gráfico de dispersão das variáveis área dedicada às fazendas no ano de 1992 e área dedicada às fazendas no ano de 1987.

Em uma análise preliminar, sob a abordagem clássica, ajustou-se um modelo de regressão linear simples considerando que os erros ϵ_i , $i = 1, 2, \dots, N$, são independentes e identicamente distribuídos, com distribuição Normal (Jorgensen, 2019). Abaixo na Figura 6.2 é apresentado o histograma dos resíduos da respectiva regressão:

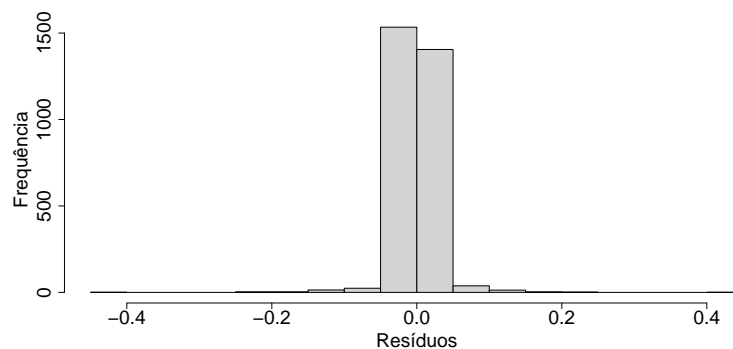


Figura 6.2: Histograma dos resíduos do modelo de regressão linear simples considerando como variável dependente a área dedicada às fazendas no ano de 1992 e covariável a área dedicada às fazendas no ano de 1987.

Nota-se que as caudas do histograma dos resíduos são pesadas. Considerou-se no presente estudo que os erros ϵ_i , $i = 1, 2, \dots, N$, são independentes e identicamente distribuídos, com distribuição T-Student.

A variável auxiliar X que servirá também para construção dos estratos considerada

foi a área dedicada à fazenda no ano de 1982. Existem no conjunto de dados 3 condados sem informação desta variável. Tais unidades foram excluídos do estudo. Na Figura 6.3 é apresentado o gráfico de dispersão das variáveis:

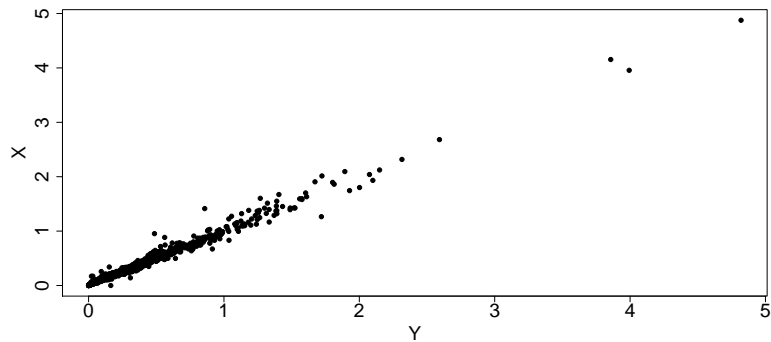


Figura 6.3: Gráfico de dispersão das variáveis área dedicada às fazendas no ano de 1992 e área dedicada às fazendas no ano de 1982.

Analisando-se a Figura 6.3 percebe-se que a relação linear entre as variáveis é plausível.

Abaixo na Figura 6.4 é apresentado o histograma da variável X :



Figura 6.4: Histograma das áreas dedicadas às fazendas no ano de 1982.

Verifica-se graficamente que a variável auxiliar não é simétrica. Um dos pressupostos do modelo AQS é a simetria da variável. Por este motivo o modelo AQS não foi considerado no presente estudo. Destaca-se que o modelo AEI não pressupõe a simetria da variável de estratificação.

Conforme mencionado no Capítulo 2 geralmente não se dispõe de todos os valores do vetor \mathbf{y} . No presente estudo, em particular, tem-se acesso a todos os dados populacionais das variáveis de interesse. Numa primeira etapa ajustou-se uma regressão T-Student considerando todos os dados populacionais das variáveis Y_i . Em seguida, foram selecionadas amostras da população de tamanho $n = 20, 60$ e 100 considerando os modelos AEI1, AEI2, AASX e AAS. Por fim, verificou-se qual modelo obteve as estimativas dos parâmetros ν, β_0, β_1 e σ mais próximas daquelas encontradas considerando todos os dados populacionais.

O conjunto de distribuições *a priori* consideradas foram: $\pi(\nu) \propto \kappa \exp\{-\kappa\nu\} I_{(a,b)}(\nu)$, com $a = 2, b = 50$ e $\kappa = 0,1$; $\beta_i \sim N(m_0, s_0^2)$, para $i = 0,1$, com $m_0 = 0, s_0^2 = 100^2$; $\sigma \sim \text{Half-Cauchy}(m_1, s_1^2)$, com $m_1 = 0$ e $s_1^2 = 25^2$. As distribuições *a priori* consideradas para os demais parâmetros específicos de cada modelo foram as mesmas utilizadas no estudo de simulação apresentado no capítulo anterior.

Foram obtidas duas cadeias, com valores iniciais diferentes para cada parâmetro. Foram realizadas 40000 iterações, utilizando-se um período de aquecimento de 10000 iterações, ou seja, foram descartadas as primeiras 10000 iterações de ambas as cadeias. Para eliminar a correlação entre os valores restantes foi aplicado o espaçamento de 30 iterações em cada cadeia. Ao final, amostras das distribuições *a posteriori* de tamanho 2000 foram obtidas para ν, β_0, β_1 e σ .

Abaixo, na Tabela 6.1, são apresentadas as medianas, médias, desvios-padrão e intervalos HPDs das distribuições *a posteriori* obtidas dos parâmetros da distribuição T-Student considerando todos os dados populacionais da variável de interesse.

Parâmetros	Med	Méd	DP	HPD
ν	2,0046	2,0067	0,0067	[2,0000; 2,0199]
β_0	-0,0026	-0,0026	0,0002	[-0,0029; -0,0022]
β_1	0,9915	0,9915	0,0013	[0,9889; 0,9942]
σ	0,0064	0,0064	0,0001	[0,0061; 0,0066]

Tabela 6.1: Resultados do ajuste da regressão de T-Student considerando os dados populacionais.

Na Figura 6.5 são apresentados os boxplots referentes as amostras das distribuições *a posteriori* dos parâmetros ν , β_0 , β_1 e σ dos modelos AEI1, AEI2, AASX e AAS. As linhas vermelhas são as médias *a posteriori* dos parâmetros referentes ao ajuste da regressão T-Student considerando os dados populacionais.

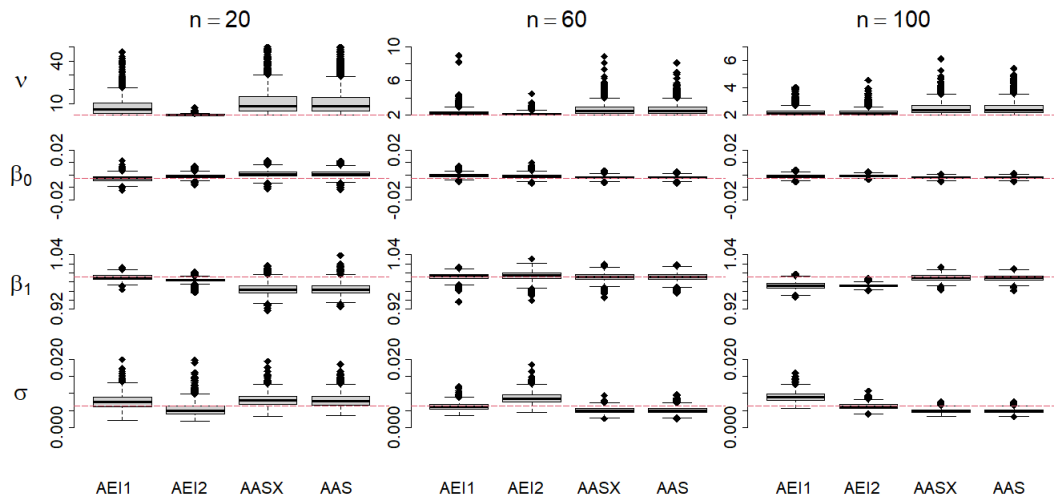


Figura 6.5: Boxplots referentes as amostras das distribuições *a posteriori* dos parâmetros ν , β_0 , β_1 e σ , considerando-se os modelos AEI1, AEI2, AASX e AAS e $n = 20, 60, 100$.

Analisando-se a Figura 6.5, referente ao parâmetro ν , nota-se que os valores das amostras das distribuições *a posteriori* do modelo AEI2 estão mais próximos da linha vermelha em comparação aos demais.

Em relação ao intercepto, nota-se que no caso $n = 20$ o modelo AE2 obteve valores amostrados da distribuição *a posteriori* mais próximos da linha vermelha em comparação aos demais, fato que não se repetiu nos casos $n = 60$ e 100 , onde constatou-se um pequeno desvio em comparação aos demais modelos.

Referente ao parâmetro β_1 , no caso $n = 20$, destaca-se que os modelos AE1 e AE2 tiveram valores amostrados da distribuição *a posteriori* mais próximos da linha vermelha em comparação aos modelos AASX e AAS. No caso $n = 60$ não foi constatado diferença significativa na comparação dos modelos. O caso $n = 100$ verifica-se que os boxplots dos modelos AEI1 e AEI2 ficaram deslocados da linha vermelha em relação aos demais.

Por fim, em relação ao parâmetro σ , nos casos $n = 20$ e $n = 60$ constata-se uma

flutuação dos boxplots dos modelos em relação a linha vermelha. No caso $n = 100$ verifica-se que o modelo AEI2 foi o único que obteve valores amostrados da distribuição *a posteriori* próximos da linha vermelha.

A seguir na Figura 6.6 são apresentados os HPDs de 95% das estimativas obtidas dos parâmetros ν , β_0 , β_1 e σ respectivamente.

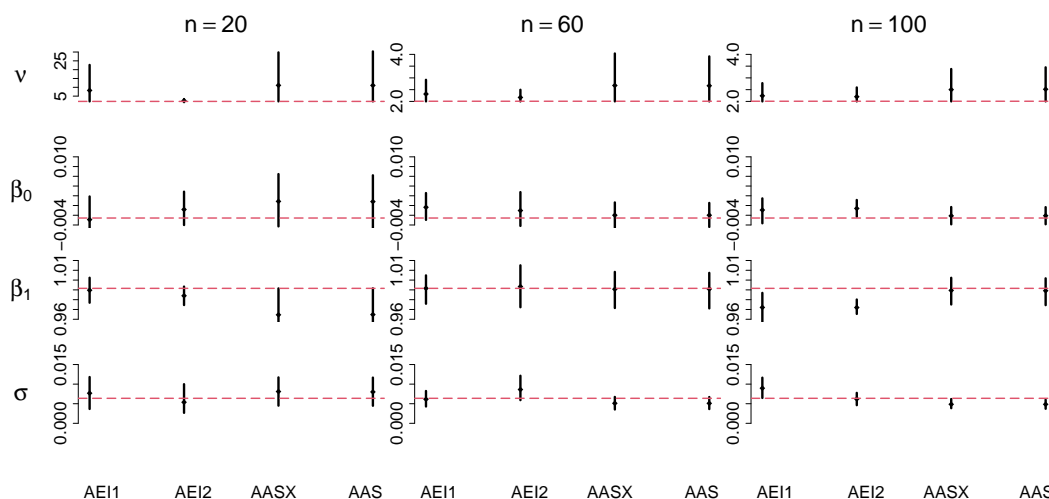


Figura 6.6: Intervalos HPDs de 95% dos parâmetros ν , β_0 , β_1 e σ , considerando-se os modelos AEI1, AEI2, AASX e AAS e $n = 20, 60, 100$.

Analisando-se a Figura 6.6 nota-se que o modelo AEI2 teve um desempenho superior na estimação pontual e intervalar do parâmetro ν em relação aos demais. Destaca-se que nos casos $n = 20, 60$ e 100 , os intervalos HPD compreenderam a estimação pontual do ajuste da regressão T-Student considerando os dados populacionais, além de terem comprimentos significativamente menores em comparação aos modelos AEI1, AASX e AAS. Nota-se que a estimação dos demais parâmetros não foi afetada na maioria das ocasiões. Como um fator negativo, destaca-se o caso $n = 100$ em que o modelo obteve uma estimação pontual do parâmetro β_1 pior em relação aos planos AASX e AAS, além do intervalo HPD não compreender a estimativa pontual referente ao ajuste com os dados populacionais.

Destaca-se que para o caso em que $n = 20$, o modelo AEI2 obteve os intervalos de credibilidade com os menores comprimentos para todos os parâmetros em relação aos

demais modelos, com exceção do parâmetro σ , no qual obteve uma diferença insignificante no comprimento do intervalo de 0.0001 em relação ao modelo AAS. Ressalta-se que em todos os casos os intervalos compreenderam as médias *a posteriores* dos parâmetros obtidas no ajuste do modelo T-Student para os dados populacionais.

Assim como nas simulações apresentadas no capítulo anterior, o plano AEI2 obteve estimativas melhores em comparação ao plano AEI1, com destaque para a estimação pontual e intervalar do parâmetro ν no caso $n = 20$. Novamente a hipótese da necessidade de se selecionar mais valores extremos é corroborada, principalmente quando o tamanho da amostra é relativamente pequeno.

Capítulo 7

Conclusões e trabalhos futuros

Ao longo deste trabalho, foram propostos planos amostrais que proporcionam uma maior probabilidade de se selecionarem valores da variável resposta mais extremos presentes nas caudas, visando estimar de forma eficiente os graus de liberdade em modelos com caudas pesadas. Foram realizados estudos de simulação no contexto clássico e bayesiano a fim de verificar-se o desempenho dos planos propostos. No geral, constatou-se que os planos propostos tiveram um desempenho melhor nas simulações em comparação à amostragem aleatória simples. Destaca-se o plano de amostragem quadrática simétrica, que obteve resultados significativamente melhores em relação à amostragem aleatória simples, principalmente na estimação do parâmetro ν , além de ser facilmente implementado. Na aplicação aos dados reais notou-se que no geral o plano que obteve o melhor desempenho foi a amostragem estratificada informativa com 6 estratos. Neste sentido, uma contribuição do presente trabalho é a melhora da estimação do parâmetro ν , sem grande impacto nas estimativas dos demais parâmetros, utilizando-se os planos propostos.

7.1 Trabalhos futuros

Visando-se melhorar ainda mais a estimação do parâmetro ν e também dos demais parâmetros, pretende-se estudar como obter a alocação ótima da amostra nos estratos dos modelos AEI1 e AEI2.

Pretende-se considerar a seguinte relação linear entre a variável de interesse Y e a

variável auxiliar X :

$$X_i = b_0 + b_1 Y_i + \epsilon_i; \epsilon_i \sim t_\phi \left(0, \sigma_\epsilon^2 \frac{\phi}{\phi - 2} \right), i = 1, 2, \dots, N. \quad (7.1)$$

No caso do plano amostragem quadrática simétrica com variável auxiliar, considerando a estrutura (7.1), tem-se que a função de densidade de probabilidade dos valores amostrados $\mathbf{y}_s, \mathbf{x}_s$ é:

$$f(\mathbf{y}_s, \mathbf{x}_s | \boldsymbol{\theta}, \phi, \mathbf{z}) = \prod_{i \in s} \left\{ \frac{f_p(y_i | \boldsymbol{\theta}, \mathbf{z}_i) f_p(x_i | y_i, b_0, b_1, \phi, \sigma_\epsilon) (x_i - c)^2}{[b_1 E_p(Y_i | \boldsymbol{\theta}, \mathbf{z}_i) + b_0 - c]^2 + b_1^2 V_p(Y_i | \boldsymbol{\theta}, \mathbf{z}_i) + \sigma_\epsilon^2 \frac{\phi}{\phi - 2}} \right\}, \quad (7.2)$$

Uma extensão natural da metodologia proposta é considerar que os erros sejam heterocedásticos nos modelos propostos.

Pretende-se fazer um estudo de simulação considerando outras distribuições para as variáveis de interesses, em particular os casos *Slash* e *Variance Gama*.

Embora a função quadrática tenha tido um bom desempenho nas simulações e também tenha facilitado a obtenção das funções densidades, outras funções poderiam ser adotadas, de acordo com a natureza dos modelos utilizados.

Avaliar a viabilidade da implementação da *priori* de Jeffreys para os planos amostrais informativos pode ser investigado.

Apêndice A

Resultados referentes aos planos amostrais apresentados

A.1 Resultados referentes aos modelos AASX, AQ, AQS e AQC

Em relação ao plano amostragem quadrática com variável auxiliar, foi mencionado na Subseção 3.3.1 que:

$$N(x_i | b_0 + b_1 y_i, \sigma_\epsilon^2) N(y_i | \mathbf{z}'_i \boldsymbol{\beta}, \lambda_i^{-1} \sigma^2) \equiv N(y_i, x_i | \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i),$$

onde $\boldsymbol{\mu}_i = [\mathbf{z}'_i \boldsymbol{\beta}, b_0 + b_1 \mathbf{z}'_i \boldsymbol{\beta}]'$ e

$$\boldsymbol{\Sigma}_i = \begin{bmatrix} \lambda_i^{-1} \sigma^2 & \lambda_i^{-1} \sigma^2 b_1 \\ \lambda_i^{-1} \sigma^2 b_1 & \lambda_i^{-1} \sigma^2 b_1^2 + \sigma_\epsilon^2 \end{bmatrix}.$$

Para verificar, basta:

$$\begin{aligned} & \frac{1}{2\pi} \frac{1}{\lambda_i^{-0.5} \sigma \sigma_\epsilon} \exp \left\{ -\frac{1}{\sigma_\epsilon^2} [x_i - (b_0 + b_1 y_i)]^2 - \frac{\lambda_i}{2\sigma^2} (y_i - \mathbf{z}'_i \boldsymbol{\beta})^2 \right\} \\ &= \frac{1}{2\pi} \frac{1}{\lambda_i^{-0.5} \sigma \sigma_\epsilon} \exp \left\{ -\frac{1}{2} \left[y_i^2 \left(\frac{b_1^2}{\sigma_\epsilon^2} + \frac{1}{\lambda^{-1} \sigma^2} \right) - 2 \mathbf{z}'_i \boldsymbol{\beta} y_i \left(\frac{1}{\lambda^{-1} \sigma^2} \right) + (\mathbf{z}'_i \boldsymbol{\beta})^2 \left(\frac{1}{\lambda^{-1} \sigma^2} \right) \right] \right\} \\ &\times \exp \left\{ -\frac{1}{2} \left[x_i^2 \left(\frac{1}{\sigma_\epsilon^2} \right) - 2x_i \left(\frac{b_0}{\sigma_\epsilon^2} \right) + b_0^2 \left(\frac{1}{\sigma_\epsilon^2} \right) \right] \right\} \times \exp \left\{ \frac{b_1 x_i y_i}{\sigma_\epsilon^2} - \frac{b_0 b_1 y_i}{\sigma_\epsilon^2} \right\} \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2\pi} \frac{1}{\lambda_i^{-0.5} \sigma \sigma_\epsilon} \exp \left\{ -\frac{1}{2} \left[y_i^2 \left(\frac{b_1^2}{\sigma_\epsilon^2} + \frac{1}{\lambda^{-1} \sigma^2} \right) - 2 \mathbf{z}'_i \boldsymbol{\beta} y_i \left(\frac{b_1^2}{\sigma_\epsilon^2} + \frac{1}{\lambda^{-1} \sigma^2} \right) + (\mathbf{z}'_i \boldsymbol{\beta})^2 \left(\frac{b_1^2}{\sigma_\epsilon^2} + \frac{1}{\lambda^{-1} \sigma^2} \right) \right] \right\} \\
&\times \exp \left\{ -\frac{1}{2} \left[x_i^2 \left(\frac{1}{\sigma_\epsilon^2} \right) - 2 x_i (b_0 + b_1 \mathbf{z}'_i \boldsymbol{\beta}) \left(\frac{1}{\sigma_\epsilon^2} \right) + (b_0 + b_1 \mathbf{z}'_i \boldsymbol{\beta})^2 \left(\frac{1}{\sigma_\epsilon^2} \right) \right] \right\} \\
&\times \exp \left\{ \frac{b_1 x_i y_i}{\sigma_\epsilon^2} - \frac{b_0 b_1 y_i}{\sigma_\epsilon^2} - \frac{\mathbf{z}'_i \boldsymbol{\beta} y_i b_1^2}{\sigma_\epsilon^2} + \frac{1}{2} \frac{(\mathbf{z}'_i \boldsymbol{\beta})^2 b_1^2}{\sigma_\epsilon^2} - \frac{\mathbf{z}'_i \boldsymbol{\beta} x_i b_1}{\sigma_\epsilon^2} + \frac{\mathbf{z}'_i \boldsymbol{\beta} b_0 b_1}{\sigma_\epsilon^2} + \frac{1}{2} \frac{(\mathbf{z}'_i \boldsymbol{\beta})^2 b_1^2}{\sigma_\epsilon^2} \right\} \\
&= \frac{1}{2\pi} \frac{1}{\lambda_i^{-0.5} \sigma \sigma_\epsilon} \exp \left\{ -\frac{1}{2} \left(\frac{b_1^2}{\sigma_\epsilon^2} + \frac{1}{\lambda^{-1} \sigma^2} \right) (y_i - \mathbf{z}'_i \boldsymbol{\beta})^2 - \frac{1}{2\sigma_\epsilon^2} [x_i - (b_0 + b_1 \mathbf{z}'_i \boldsymbol{\beta})]^2 \right\} \\
&\times \exp \left\{ b_1 \left[\frac{(\mathbf{z}'_i \boldsymbol{\beta})^2 b_1}{\sigma_\epsilon^2} - \frac{\mathbf{z}'_i \boldsymbol{\beta} y_i b_1}{\sigma_\epsilon^2} - \frac{\mathbf{z}'_i \boldsymbol{\beta} x_i}{\sigma_\epsilon^2} + \frac{\mathbf{z}'_i \boldsymbol{\beta} b_0}{\sigma_\epsilon^2} - \frac{b_0 y_i}{\sigma_\epsilon^2} + \frac{x_i y_i}{\sigma_\epsilon^2} \right] \right\} \\
&= \frac{1}{2\pi} \frac{1}{\lambda_i^{-0.5} \sigma \sigma_\epsilon} \exp \left\{ -\frac{1}{2} \left(\frac{b_1^2}{\sigma_\epsilon^2} + \frac{1}{\lambda^{-1} \sigma^2} \right) (y_i - \mathbf{z}'_i \boldsymbol{\beta})^2 - \frac{1}{2\sigma_\epsilon^2} [x_i - (b_0 + b_1 \mathbf{z}'_i \boldsymbol{\beta})]^2 \right\} \\
&\times \exp \left\{ \frac{b_1}{\sigma_\epsilon^2} (y_i - \mathbf{z}'_i \boldsymbol{\beta}) [x_i - (b_0 + b_1 \mathbf{z}'_i \boldsymbol{\beta})] \right\} \\
&= \frac{1}{2\pi} \frac{1}{\sigma^* \sigma_\epsilon^* \sqrt{1 - \rho_{XY}^2}} \exp \left\{ -\frac{1}{2(1 - \rho_{XY}^2)} \left\{ \left(\frac{y_i - \mathbf{z}'_i \boldsymbol{\beta}}{\sigma^*} \right)^2 - 2 \frac{\rho_{XY} (y_i - \mathbf{z}'_i \boldsymbol{\beta}) [x_i - (b_0 + b_1 \mathbf{z}'_i \boldsymbol{\beta})]}{\sigma^* \sigma_\epsilon^*} + \left[\frac{x_i - (b_0 + b_1 \mathbf{z}'_i \boldsymbol{\beta})}{\sigma_\epsilon^*} \right]^2 \right\} \right\};
\end{aligned}$$

com

$$\sigma^* = \frac{\sigma_\epsilon \rho_{XY}}{b_1 \sqrt{1 - \rho_{XY}^2}}; \quad \sigma_\epsilon^* = \frac{\sigma_\epsilon}{\sqrt{1 - \rho_{XY}^2}}; \quad \rho_{XY} = \frac{1}{\sqrt{\frac{1}{b_1^2} \frac{\sigma_\epsilon^2}{\lambda^{-1} \sigma^2} + 1}}.$$

Note ainda que:

$$\sigma^{*2} = \lambda^{-1} \sigma^2; \quad \sigma_\epsilon^{*2} = \lambda^{-1} \sigma^2 b_1^2 + \sigma_\epsilon^2; \quad \text{COV}(X, Y) = \lambda^{-1} \sigma^2 b_1.$$

Calculemos o cálculo da correlação $\rho_{X_i Y_i}$ referentes aos modelos AASX, AQ e AQS, onde $Y_i | \boldsymbol{\beta}, \sigma, \nu, \mathbf{z}_i \sim t_\nu(\mathbf{z}'_i \boldsymbol{\beta}, \sigma^2 \frac{\nu}{\nu-2})$ e $X_i | Y_i, b_0, b_1, \sigma_\epsilon \sim N(b_0 + b_1 y_i, \sigma_\epsilon^2)$:

$$\begin{aligned}
E(X_i Y_i) &= \int_{-\infty}^{+\infty} y_i f(y_i | \boldsymbol{\beta}, \sigma, \nu, \mathbf{z}_i) \int_{-\infty}^{+\infty} x_i f(x_i | y_i, b_0, b_1, \sigma_\epsilon) dx_i dy_i \\
&= \int_{-\infty}^{+\infty} (b_0 y_i + b_1 y_i^2) f(y_i | \mu, \sigma, \nu, \mathbf{z}_i) dy_i = b_1 E(Y_i^2) + b_0 E(Y_i) \\
&= b_1 \left[(\mathbf{z}'_i \boldsymbol{\beta})^2 + \sigma^2 \frac{\nu}{\nu-2} \right] + b_0 \mathbf{z}'_i \boldsymbol{\beta};
\end{aligned}$$

$$\begin{aligned} COV(X,Y) &= E(XY) - E\{E(X|Y)\}E(Y) = b_1E(Y^2) + b_0E(Y) - [b_0 + b_1E(Y)]E(Y) \\ &= b_1V(Y) = b_1\sigma^2\frac{\nu}{\nu-2}; \end{aligned}$$

$$\rho_{xy} = \frac{b_1V(Y)}{\sqrt{V(X)V(Y)}} = \frac{b_1\sqrt{V(Y)}}{\sqrt{E\{V(X|Y)\} + V\{E(X|Y)\}}} = \frac{b_1}{\sqrt{\frac{\sigma_\epsilon^2}{\sigma^2}\frac{\nu-2}{\nu} + b_1^2}}.$$

Note que da expressão de ρ_{XY} , pode-se concluir que:

$$\sigma_\epsilon = \sigma\sqrt{b_1^2\frac{\nu}{\nu-2}\left(\frac{1}{\rho_{XY}^2} - 1\right)}.$$

A.2 Informação de Fisher dos modelos AASX e AQ

Para o cálculo da informação de Fisher do modelo AASX apresentado no Capítulo 4, considere o seguinte lema:

Lema A.1: Considere o modelo de regressão $Y \sim t_\nu(\mathbf{z}'\boldsymbol{\beta}, \sigma^2)$. Defina:

$$r = \frac{y - \mathbf{z}'\boldsymbol{\beta}}{\sigma}; a_k = E\left[\frac{1}{(\nu + R^2)^k}\right]; b_q = E\left[\frac{(Y - \mathbf{z}'\boldsymbol{\beta})^q}{\nu + R^2}\right]; c_q = E\left[\frac{(Y - \mathbf{z}'\boldsymbol{\beta})^q}{(\nu + R^2)^2}\right];$$

$$c(\nu, \sigma^2) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)\nu^{\frac{\nu}{2}}}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi\sigma^2}}.$$

De [Fonseca et al. \(2008\)](#), tem-se os seguintes resultados:

1. $E(R) = 0$, se $\nu > 1$; $E(R^2) = \frac{\nu}{\nu-2}$, se $\nu > 2$; $E(R^3) = 0$, se $\nu > 3$;
 $E(R^4) = \frac{3\nu^2}{(\nu-2)(\nu-4)}$, se $\nu > 4$.
2. $a_k = E\left[\frac{1}{(\nu+R^2)^k}\right] = \int_{-\infty}^{+\infty} c(\nu, \sigma^2) (\nu + r^2)^{-\frac{(\nu+2k+1)}{2}} dy = \left[\frac{(\nu+2k)}{\nu}\right]^{\frac{\nu+2k+1}{2}} \frac{c(\nu, \sigma^2)}{c\left(\nu+2k, \frac{\nu\sigma^2}{\nu+2k}\right)}$.
 Em particular, $a_1 = \frac{1}{\nu+1}$ e $a_2 = \frac{\nu+2}{\nu(\nu+1)(\nu+3)}$.
3. $b_q = E\left[\frac{(Y - \mathbf{z}'\boldsymbol{\beta})^q}{\nu + R^2}\right] = \int_{-\infty}^{+\infty} (y - \mathbf{z}'\boldsymbol{\beta})^q c(\nu, \sigma^2) (\nu + r^2)^{-\frac{\nu+3}{2}} dy = a_1 E[(Y - \mathbf{z}'\boldsymbol{\beta})^q]$,
 onde $Y \sim t_{\nu+2}\left(\mathbf{z}'\boldsymbol{\beta}, \frac{\nu\sigma^2}{\nu+2}\right)$.
4. $c_q = E\left[\frac{(Y - \mathbf{z}'\boldsymbol{\beta})^q}{(\nu+R^2)^2}\right] = \int_{-\infty}^{+\infty} (y - \mathbf{z}'\boldsymbol{\beta})^q c(\nu, \sigma^2) (\nu + r^2)^{-\frac{\nu+5}{2}} dy = a_2 E[(Y - \mathbf{z}'\boldsymbol{\beta})^q]$,
 onde $Y \sim t_{\nu+4}\left(\mathbf{z}'\boldsymbol{\beta}, \frac{\nu\sigma^2}{\nu+4}\right)$.

As derivadas de primeira ordem do plano amostral AASX são apresentados abaixo:

$$\frac{\partial l(\gamma)}{\partial \beta} = \frac{\nu+1}{\sigma^2} \sum_{i \in s} \left[z_i \frac{y_i - \mathbf{z}'_i \beta}{\nu + \frac{(y_i - \mathbf{z}'_i \beta)^2}{\sigma^2}} \right];$$

$$\frac{\partial l(\gamma)}{\partial \sigma} = -\frac{n}{\sigma} + \frac{\nu+1}{\sigma^3} \sum_{i \in s} \left[\frac{(y_i - \mathbf{z}'_i \beta)^2}{\nu + \frac{(y_i - \mathbf{z}'_i \beta)^2}{\sigma^2}} \right];$$

$$\frac{\partial l(\gamma)}{\partial \nu} = \frac{n}{2} \left[\psi \left(\frac{\nu+1}{2} \right) - \psi \left(\frac{\nu}{2} \right) \right] + \frac{1}{2} \sum_{i \in s} \left\{ -\log \left[\frac{\nu + \frac{(y_i - \mathbf{z}'_i \beta)^2}{\sigma^2}}{\nu} \right] \right\};$$

$$\frac{\partial l(\gamma)}{\partial b_0} = \frac{1}{\sigma_\epsilon^2} \left(\sum_{i \in s} x_i - b_1 \sum_{i \in s} y_i - n b_0 \right);$$

$$\frac{\partial l(\gamma)}{\partial b_1} = \frac{1}{\sigma_\epsilon^2} \left(\sum_{i \in s} x_i y_i - b_1 \sum_{i \in s} y_i^2 - b_0 \sum_{i \in s} y_i \right);$$

$$\frac{\partial l(\gamma)}{\partial \sigma_\epsilon} = -\frac{n}{\sigma_\epsilon} + \frac{1}{\sigma_\epsilon^3} \sum_{i \in s} [x_i - (b_0 + b_1 y_i)]^2.$$

As derivadas de segunda ordem são:

$$\frac{\partial^2 l(\gamma)}{\partial \beta_0^2} = \frac{(\nu+1)}{\sigma^4} \sum_{i \in s} \left\{ \frac{(y_i - \mathbf{z}'_i \beta)^2}{\left[\nu + \frac{(y_i - \mathbf{z}'_i \beta)^2}{\sigma^2} \right]^2} \right\} - \frac{v(\nu+1)}{\sigma^2} \sum_{i \in s} \left\{ \frac{1}{\left[\nu + \frac{(y_i - \mathbf{z}'_i \beta)^2}{\sigma^2} \right]^2} \right\};$$

$$\frac{\partial^2 l(\gamma)}{\partial \beta_j \partial \beta_0} = \frac{(\nu+1)}{\sigma^4} \sum_{i \in s} \left\{ \frac{z_{ij} (y_i - \mathbf{z}'_i \beta)^2}{\left[\nu + \frac{(y_i - \mathbf{z}'_i \beta)^2}{\sigma^2} \right]^2} \right\} - \frac{v(\nu+1)}{\sigma^2} \sum_{i \in s} \left\{ \frac{z_{ij}}{\left[\nu + \frac{(y_i - \mathbf{z}'_i \beta)^2}{\sigma^2} \right]^2} \right\};$$

$$\frac{\partial^2 l(\gamma)}{\partial \beta_j^2} = \frac{(\nu+1)}{\sigma^4} \sum_{i \in s} \left\{ \frac{z_{ij}^2 (y_i - \mathbf{z}'_i \beta)^2}{\left[\nu + \frac{(y_i - \mathbf{z}'_i \beta)^2}{\sigma^2} \right]^2} \right\} - \frac{v(\nu+1)}{\sigma^2} \sum_{i \in s} \left\{ \frac{z_{ij}^2}{\left[\nu + \frac{(y_i - \mathbf{z}'_i \beta)^2}{\sigma^2} \right]^2} \right\};$$

$$\frac{\partial^2 l(\boldsymbol{\gamma})}{\partial \beta_i \partial \beta_j} = \frac{(\nu+1)}{\sigma^4} \sum_{i \in S} \left\{ \frac{z_{ij} z_{il} (y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\left[\nu + \frac{(y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\sigma^2} \right]^2} \right\} - \frac{\nu(\nu+1)}{\sigma^2} \sum_{i \in S} \left\{ \frac{z_{ij} z_{il}}{\left[\nu + \frac{(y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\sigma^2} \right]^2} \right\};$$

Considere a seguinte matriz simétrica:

$$\mathbf{z}_i \mathbf{z}'_i = \begin{bmatrix} 1 & z_{i1} & z_{i2} & z_{i3} & \cdot & \cdot & \cdot & z_{ik-1} & z_{ik} \\ z_{i1} & z_{i1}^2 & z_{i1} z_{i2} & z_{i1} z_{i3} & \cdot & \cdot & \cdot & z_{i1} z_{ik-1} & z_{i1} z_{ik} \\ z_{i2} & z_{i1} z_{i2} & z_{i2}^2 & z_{i2} z_{i3} & \cdot & \cdot & \cdot & z_{i2} z_{ik-1} & z_{i2} z_{ik} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ z_{ik} & z_{i1} z_{ik} & z_{i2} z_{ik} & z_{i3} z_{ik} & \cdot & \cdot & \cdot & z_{ik-1} z_{ik} & z_{ik}^2 \end{bmatrix}.$$

Tem-se então:

$$\frac{\partial^2 l(\boldsymbol{\gamma})}{\partial \boldsymbol{\beta}^2} = \frac{(\nu+1)}{\sigma^4} \sum_{i \in S} \left\{ \frac{\mathbf{z}_i \mathbf{z}'_i (y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\left[\nu + \frac{(y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\sigma^2} \right]^2} \right\} - \frac{\nu(\nu+1)}{\sigma^2} \sum_{i \in S} \left\{ \frac{\mathbf{z}_i \mathbf{z}'_i}{\left[\nu + \frac{(y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\sigma^2} \right]^2} \right\};$$

$$\frac{\partial^2 l(\boldsymbol{\gamma})}{\partial \sigma \partial \boldsymbol{\beta}} = \frac{-2\nu(\nu+1)}{\sigma^3} \sum_{i \in S} \left\{ \mathbf{z}_i \frac{(y_i - \mathbf{z}'_i \boldsymbol{\beta})}{\left[\nu + \frac{(y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\sigma^2} \right]^2} \right\};$$

$$\frac{\partial^2 l(\boldsymbol{\gamma})}{\partial \nu \partial \boldsymbol{\beta}} = \frac{1}{\sigma^2} \sum_{i \in S} \left\{ \mathbf{z}_i \frac{(y_i - \mathbf{z}'_i \boldsymbol{\beta})}{\left[\nu + \frac{(y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\sigma^2} \right]} \right\} - \frac{(\nu+1)}{\sigma^2} \sum_{i \in S} \left\{ \mathbf{z}_i \frac{(y_i - \mathbf{z}'_i \boldsymbol{\beta})}{\left[\nu + \frac{(y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\sigma^2} \right]^2} \right\};$$

$$\frac{\partial^2 l(\boldsymbol{\gamma})}{\partial b_0 \partial \boldsymbol{\beta}} = \frac{\partial^2 l(\boldsymbol{\gamma})}{\partial b_1 \partial \boldsymbol{\beta}} = \frac{\partial^2 l(\boldsymbol{\gamma})}{\partial \sigma_\epsilon \partial \boldsymbol{\beta}} = \mathbf{0}_{(k+1) \times 1};$$

$$\frac{\partial^2 l(\boldsymbol{\gamma})}{\partial \sigma^2} = \frac{n}{\sigma^2} - \frac{3\nu(\nu+1)}{\sigma^4} \sum_{i \in S} \left\{ \frac{(y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\left[\nu + \frac{(y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\sigma^2} \right]^2} \right\} - \frac{(\nu+1)}{\sigma^6} \sum_{i \in S} \left\{ \frac{(y_i - \mathbf{z}'_i \boldsymbol{\beta})^4}{\left[\nu + \frac{(y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\sigma^2} \right]^2} \right\};$$

$$\frac{\partial^2 l(\boldsymbol{\gamma})}{\partial \nu \partial \sigma} = \frac{1}{\sigma^3} \sum_{i \in s} \left\{ \frac{(y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\left[\nu + \frac{(y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\sigma^2} \right]} \right\} - \frac{(\nu+1)}{\sigma^3} \sum_{i \in s} \left\{ \frac{(y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\left[\nu + \frac{(y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\sigma^2} \right]^2} \right\};$$

$$\frac{\partial^2 l(\boldsymbol{\gamma})}{\partial b_0 \partial \sigma} = 0; \quad \frac{\partial^2 l(\boldsymbol{\gamma})}{\partial b_1 \partial \sigma} = 0; \quad \frac{\partial^2 l(\boldsymbol{\gamma})}{\partial \sigma_\epsilon \partial \sigma} = 0;$$

$$\frac{\partial^2 l(\boldsymbol{\gamma})}{\partial \nu^2} = \frac{n}{4} \left[\psi' \left(\frac{\nu+1}{2} \right) - \psi' \left(\frac{\nu}{2} \right) \right] + \frac{1}{2\nu\sigma^4} \sum_{i \in s} \left\{ \frac{(y_i - \mathbf{z}'_i \boldsymbol{\beta})^4}{\left[\nu + \frac{(y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\sigma^2} \right]^2} \right\} + \frac{1}{2} \sum_{i \in s} \left\{ \frac{1}{\left[\nu + \frac{(y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\sigma^2} \right]^2} \right\};$$

$$\frac{\partial^2 l(\boldsymbol{\gamma})}{\partial b_0 \partial \nu} = 0; \quad \frac{\partial^2 l(\boldsymbol{\gamma})}{\partial b_1 \partial \nu} = 0; \quad \frac{\partial^2 l(\boldsymbol{\gamma})}{\partial \sigma_\epsilon \partial \nu} = 0;$$

$$\frac{\partial^2 l(\boldsymbol{\gamma})}{\partial b_0^2} = -\frac{n}{\sigma_\epsilon^2}$$

$$\frac{\partial^2 l(\boldsymbol{\gamma})}{\partial b_1 \partial b_0} = -\frac{\sum_{i \in s} y_i}{\sigma_\epsilon^2};$$

$$\frac{\partial^2 l(\boldsymbol{\gamma})}{\partial \sigma_\epsilon \partial b_0} = -\frac{2}{\sigma_\epsilon^3} \left(\sum_{i \in s} x_i - b_1 \sum_{i \in s} y_i - nb_0 \right);$$

$$\frac{\partial^2 l(\boldsymbol{\gamma})}{\partial b_1^2} = -\frac{\sum_{i \in s} y_i^2}{\sigma_\epsilon^2};$$

$$\frac{\partial^2 l(\boldsymbol{\gamma})}{\partial \sigma_\epsilon \partial b_1} = -\frac{2}{\sigma_\epsilon^3} \left(\sum_{i \in s} x_i y_i - b_1 \sum_{i \in s} y_i^2 - b_0 \sum_{i \in s} y_i \right);$$

$$\frac{\partial^2 l(\boldsymbol{\gamma})}{\partial \sigma_\epsilon^2} = \frac{n}{\sigma_\epsilon^2} - \frac{3}{\sigma_\epsilon^4} \left(\sum_{i \in s} x_i^2 - 2b_0 \sum_{i \in s} x_i - 2b_1 \sum_{i \in s} x_i y_i + nb_0^2 + 2b_0 b_1 \sum_{i \in s} y_i + \sum_{i \in s} y_i^2 \right);$$

Considere os seguintes resultados derivados do Lema 1:

$$E \left\{ \frac{1}{\left[\nu + \frac{(Y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\sigma^2} \right]^2} \right\} = \frac{(\nu+2)}{\nu(\nu+1)(\nu+3)}; \quad E \left\{ \frac{(Y_i - \mathbf{z}'_i \boldsymbol{\beta})}{\left[\nu + \frac{(Y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\sigma^2} \right]} \right\} = \frac{1}{\nu+1} E (Y_{*i} - \mathbf{z}'_i \boldsymbol{\beta}) = 0;$$

$$E \left\{ \frac{(Y_i - \mathbf{z}'_i \boldsymbol{\beta})}{\left[\nu + \frac{(Y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\sigma^2} \right]^2} \right\} = \frac{(\nu+2)}{\nu(\nu+1)(\nu+3)} E (Y_{**i} - \mathbf{z}'_i \boldsymbol{\beta}) = 0;$$

$$E \left\{ \frac{(Y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\left[\nu + \frac{(Y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\sigma^2} \right]} \right\} = \frac{1}{(\nu+1)} E [(Y_{*i} - \mathbf{z}'_i \boldsymbol{\beta})^2] = \frac{1}{(\nu+1)} \frac{\psi \sigma^2 (\nu+2)}{(\nu+2) \psi} = \frac{\sigma^2}{(\nu+1)};$$

$$E \left\{ \frac{(Y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\left[\nu + \frac{(Y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\sigma^2} \right]^2} \right\} = \frac{(\nu+2)}{\nu(\nu+1)(\nu+3)} E [(Y_{**i} - \mathbf{z}'_i \boldsymbol{\beta})^2] = \frac{(\nu+2)}{\psi(\nu+1)(\nu+3)} \frac{\psi \sigma^2 (\nu+4)}{(\nu+4) (\nu+2)} = \frac{\sigma^2}{(\nu+1)(\nu+3)};$$

$$E \left\{ \frac{(Y_i - \mathbf{z}'_i \boldsymbol{\beta})^4}{\left[\nu + \frac{(Y_i - \mathbf{z}'_i \boldsymbol{\beta})^2}{\sigma^2} \right]^2} \right\} = \frac{(\nu+2)}{\nu(\nu+1)(\nu+3)} E [(Y_{**i} - \mathbf{z}'_i \boldsymbol{\beta})^4] = \frac{(\nu+2)}{\psi(\nu+1)(\nu+3)} \frac{\psi \sigma^4 3(\nu+4)^2}{(\nu+4)^2 (\nu+2) \psi} = \frac{3\sigma^4}{(\nu+1)(\nu+3)}.$$

Para o cálculo da Informação de Fisher, basta calcular o valor esperado $E \left(-\frac{\partial^2 l(\boldsymbol{\gamma})}{\partial^2 \boldsymbol{\gamma}} \right)$.

Referente ao plano AQ, a matriz $A(\boldsymbol{\gamma})$ é apresentada abaixo:

$$A(\boldsymbol{\gamma}) = \begin{bmatrix} \mathbf{a}_{11} & \mathbf{a}_{12} & \mathbf{a}_{13} & \mathbf{a}_{14} & \mathbf{a}_{15} & \mathbf{a}_{16} \\ \mathbf{a}_{21} & \mathbf{a}_{22} & \mathbf{a}_{23} & \mathbf{a}_{24} & \mathbf{a}_{25} & \mathbf{a}_{26} \\ \mathbf{a}_{31} & \mathbf{a}_{32} & \mathbf{a}_{33} & \mathbf{a}_{34} & \mathbf{a}_{35} & \mathbf{a}_{36} \\ \mathbf{a}_{41} & \mathbf{a}_{42} & \mathbf{a}_{43} & \mathbf{a}_{44} & \mathbf{a}_{45} & \mathbf{a}_{46} \\ \mathbf{a}_{51} & \mathbf{a}_{52} & \mathbf{a}_{53} & \mathbf{a}_{54} & \mathbf{a}_{55} & \mathbf{a}_{56} \\ \mathbf{a}_{61} & \mathbf{a}_{62} & \mathbf{a}_{63} & \mathbf{a}_{64} & \mathbf{a}_{65} & \mathbf{a}_{66} \end{bmatrix},$$

$$\begin{aligned}
\mathbf{A}_{11} &= 2b_1^2 \sum_{i \in s} \left(\frac{1}{k_i} z_i z'_i \right) - 4b_1^4 \sum_{i \in s} \left[\frac{1}{k_i^2} (z'_i \beta)^2 z_i z'_i \right] - 8b_0 b_1^3 \sum_{i \in s} \left(\frac{1}{k_i^2} z'_i \beta z_i z'_i \right) - 4b_0^2 b_1^2 \sum_{i \in s} \left(\frac{1}{k_i^2} z_i z'_i \right); \\
\mathbf{a}_{12} &= -4b_1^4 \sigma \frac{\nu}{\nu-2} \sum_{i \in s} \left(z_i \frac{1}{k_i^2} z'_i \beta \right) - 4b_0 b_1^3 \sigma \frac{\nu}{\nu-2} \sum_{i \in s} \left(z_i \frac{1}{k_i^2} \right); \\
\mathbf{a}_{13} &= 4b_1^4 \sigma^2 (\nu-2)^{-2} \sum_{i \in s} \left(z_i \frac{1}{k_i^2} z'_i \beta \right) + 4b_0 b_1^3 \sigma^2 (\nu-2)^{-2} \sum_{i \in s} \left(z_i \frac{1}{k_i^2} \right); \\
\mathbf{a}_{14} &= 2b_1 \sum_{i \in s} \left(z_i \frac{1}{k_i} \right) - 4b_0^2 b_1 \sum_{i \in s} \left(z_i \frac{1}{k_i^2} \right) - 8b_0 b_1^2 \sum_{i \in s} \left(z_i \frac{1}{k_i^2} z'_i \beta \right) - 4b_1^3 \sum_{i \in s} \left[z_i \frac{1}{k_i^2} (z'_i \beta)^2 \right]; \\
\mathbf{a}_{15} &= 2b_0 \sum_{i \in s} \left(z_i \frac{1}{k_i} \right) + 4b_1 \sum_{i \in s} \left(z_i \frac{1}{k_i} z'_i \beta \right) - 4b_0 b_1^2 \sigma^2 \frac{\nu}{\nu-2} \sum_{i \in s} \left(z_i \frac{1}{k_i^2} \right) - 4 \left(b_1^3 \sigma^2 \frac{\nu}{\nu-2} + b_0^2 b_1 \right) \sum_{i \in s} \left(z_i \frac{1}{k_i^2} z'_i \beta \right) - \\
&\quad - 8b_0 b_1^2 \sum_{i \in s} \left[z_i \frac{1}{k_i^2} (z'_i \beta)^2 \right] - 4b_1^3 \sum_{i \in s} \left[z_i \frac{1}{k_i^2} (z'_i \beta)^3 \right]; \\
\mathbf{a}_{16} &= -4b_0 b_1 \sigma_\epsilon \sum_{i \in s} \left(z_i \frac{1}{k_i^2} \right) - 4b_1^2 \sigma_\epsilon \sum_{i \in s} \left(z_i \frac{1}{k_i^2} z'_i \beta \right); \\
a_{22} &= 2b_1^2 \frac{\nu}{\nu-2} \sum_{i \in s} \frac{1}{k_i} - 4b_1^4 \sigma^2 \frac{\nu^2}{(\nu-2)^2} \sum_{i \in s} \frac{1}{k_i^2}; \\
a_{23} &= -4b_1^2 \sigma (\nu-2)^{-2} \sum_{i \in s} \frac{1}{k_i} + 4b_1^4 \sigma^3 \nu (\nu-2)^{-3} \sum_{i \in s} \frac{1}{k_i^2}; \\
a_{24} &= -4b_0 b_1^2 \sigma \frac{\nu}{\nu-2} \sum_{i \in s} \frac{1}{k_i^2} - 4b_1^3 \sigma \frac{\nu}{\nu-2} \sum_{i \in s} \left(\frac{1}{k_i^2} z'_i \beta \right); \\
a_{25} &= 4b_1 \sigma \frac{\nu}{\nu-2} \sum_{i \in s} \frac{1}{k_i} - 4b_1^3 \sigma^3 \frac{\nu^2}{(\nu-2)^2} \sum_{i \in s} \frac{1}{k_i^2} - 4b_0 b_1^2 \sigma \frac{\nu}{\nu-2} \sum_{i \in s} \left(\frac{1}{k_i^2} z_i \beta \right) - 4b_1^3 \sigma \frac{\nu}{\nu-2} \sum_{i \in s} \left[\frac{1}{k_i^2} (z_i \beta)^2 \right]; \\
a_{26} &= -4b_1^2 \sigma \frac{\nu}{\nu-2} \sigma_\epsilon \sum_{i \in s} \frac{1}{k_i^2}; \\
a_{33} &= 4b_1^2 \sigma^2 (\nu-2)^{-3} \sum_{i \in s} \frac{1}{k_i} - 4b_1^4 \sigma^4 (\nu-2)^{-4} \sum_{i \in s} \frac{1}{k_i^2}; \\
a_{34} &= 4b_0 b_1^2 \sigma^2 (\nu-2)^{-2} \sum_{i \in s} \frac{1}{k_i^2} + 4b_1^3 \sigma^2 (\nu-2)^{-2} \sum_{i \in s} \left(\frac{1}{k_i^2} z'_i \beta \right); \\
a_{35} &= -4b_1 \sigma^2 (\nu-2)^{-2} \sum_{i \in s} \frac{1}{k_i} + 4b_1^3 \sigma^4 \nu (\nu-2)^{-3} \sum_{i \in s} \frac{1}{k_i^2} + 8b_0 b_1^2 \sigma^2 (\nu-2)^{-2} \sum_{i \in s} \left(\frac{1}{k_i^2} z_i \beta \right) + 4b_1^3 \sigma^2 (\nu-2)^{-2} \sum_{i \in s} \left[\frac{1}{k_i^2} (z_i \beta)^2 \right]; \\
a_{36} &= 4b_1^2 \sigma^2 \sigma_\epsilon (\nu-2)^{-2} \sum_{i \in s} \frac{1}{k_i^2}; \\
a_{44} &= 2 \sum_{i \in s} \frac{1}{k_i} - 4 \sum_{i \in s} \left[\frac{1}{k_i^2} (b_1 z_i \beta + b_0)^2 \right]; \\
a_{45} &= 2 \sum_{i \in s} \left(\frac{1}{k_i} z'_i \beta \right) - 4b_0 b_1 \sigma^2 \frac{\nu}{\nu-2} \sum_{i \in s} \frac{1}{k_i^2} - 4 \left(b_0^2 + b_1^2 \sigma^2 \frac{\nu}{\nu-2} \right) \sum_{i \in s} \left(\frac{1}{k_i^2} z'_i \beta \right) - 8b_0 b_1 \sum_{i \in s} \left[\frac{1}{k_i^2} (z'_i \beta)^2 \right] - 4b_1^2 \sum_{i \in s} \left[\frac{1}{k_i^2} (z'_i \beta)^3 \right]; \\
a_{46} &= -4b_0 \sigma_\epsilon \sum_{i \in s} \frac{1}{k_i^2} - 4b_1 \sigma_\epsilon \sum_{i \in s} \left(\frac{1}{k_i^2} z'_i \beta \right); \\
a_{55} &= 2\sigma^2 \frac{\nu}{\nu-2} \sum_{i \in s} \frac{1}{k_i} + 2 \sum_{i \in s} \left[(z'_i \beta)^2 \frac{1}{k_i} \right] - 4 \sum_{i \in s} \left\{ \frac{1}{k_i^2} \left[b_1 \left((z'_i \beta)^2 + \sigma^2 \frac{\nu}{\nu-2} \right) + b_0 z'_i \beta \right]^2 \right\}; \\
a_{56} &= -4b_1 \sigma^2 \frac{\nu}{\nu-2} \sigma_\epsilon \sum_{i \in s} \frac{1}{k_i^2} - 4(b_0 + b_1) \sigma_\epsilon \sum_{i \in s} \left(\frac{1}{k_i^2} z'_i \beta \right); \\
a_{66} &= 2 \sum_{i \in s} \frac{1}{k_i} - 4\sigma_\epsilon^2 \sum_{i \in s} \frac{1}{k_i^2}.
\end{aligned}$$

onde:

$$k_i = (b_1 z'_i \beta + b_0)^2 + b_1^2 \sigma^2 \frac{\nu}{\nu-2} + \sigma_\epsilon^2.$$

A.3 Informação de Fisher do plano AQS - aproximação pela Lei dos grandes números

Considere o plano AQS. Referente à aproximação pela Lei dos grandes números, afirmou-se que $E \left[\frac{1}{(X_i - \mu)^2} \right]$ não é integrável. A verificação desta afirmação é apresentada abaixo.

$$\begin{aligned} E \left[\frac{1}{(X_i - \mu)^2} \right] &= \int \int \frac{1}{(X_i - \mu)^2} f_p(x_i, y_i | \mu, \sigma, \nu, \sigma_\epsilon) dx_i dy_i \\ &= \int f_p(y_i | \mu, \sigma, \nu) \int \frac{1}{(X_i - \mu)^2} f_p(x_i | y_i, \sigma_\epsilon) dx_i dy_i. \end{aligned}$$

Calcular a integral à direita da expressão anterior é equivalente a calcular:

$$E_{X_i|Y_i} \left[\frac{1}{(X_i - \mu)^2} \right]; X_i|Y_i \sim N(y_i, \sigma_\epsilon).$$

Considere a seguinte transformação linear da variável $X_i|Y_i$ (para facilitar a notação, foi omitido $|Y_i$):

$$R_i = \frac{X_i - \mu}{\sigma_\epsilon} \sim N\left(\frac{y_i - \mu}{\sigma_\epsilon}, 1\right).$$

Então:

$$E_{X_i|Y_i} \left[\frac{1}{(X_i - \mu)^2} \right] = \frac{1}{\sigma_\epsilon^2} E_{R_i|Y_i} \left(\frac{1}{R_i^2} \right).$$

A variável $W_i = R_i^2$ tem distribuição qui-quadrada não centrada, com parâmetros graus de liberdade $k = 1$ e $\lambda_i = \left(\frac{y_i - \mu}{\sigma_\epsilon}\right)^2$, parâmetro de não centralidade relacionado à média da distribuição normal R_i . Logo:

$$\begin{aligned} E_{X_i|Y_i} \left[\frac{1}{(X_i - \mu)^2} \right] &= \frac{1}{\sigma_\epsilon^2} E_{R_i|Y_i} \left(\frac{1}{R_i^2} \right) = \frac{1}{\sigma_\epsilon^2} E_{W_i|Y_i} \left(\frac{1}{W_i} \right) \\ &= \int \frac{1}{w_i} e^{-\lambda_i/2} {}_0F_1\left(; k/2; \lambda_i w_i/4\right) \frac{1}{2^{k/2} \gamma(k/2)} e^{w_i/2} w_i^{k/2-1} dw_i \\ &= \int e^{-\lambda_i/2} {}_0F_1\left(; k'/2 + 1; \lambda_i w_i/4\right) \frac{1}{2^{k/2} \gamma(k/2)} e^{w_i/2} w_i^{k'/2-1} dw_i. \end{aligned}$$

onde $k' = k - 2$. O cálculo da integral acima envolve uma distribuição qui-quadrada não centrada com parâmetros k' e λ_i . Uma observação importante referente aos parâmetros da distribuição qui-quadrada não centrada é que $k > 0$ e $\lambda > 0$. Logo, como $k' = -1$, segue que a esperança não é integrável.

A.4 Função de verossimilhança referente ao plano AQS - considerando que \bar{X} não seja uma constante

A expressão da função verossimilhança referente ao plano AQS considerando-se que \bar{X} não seja uma constante é dada por:

$$\begin{aligned} f_s(y_i, x_i | \mu, \sigma, \nu, \sigma_\epsilon) &= \frac{P(I_i = 1 | x_i, y_i) f_p(x_i, y_i | \mu, \sigma, \nu, \sigma_\epsilon)}{P(I_i = 1)} \\ &= \frac{E_p(\pi_i | x_i, y_i) f_p(x_i, y_i | \mu, \sigma, \nu, \sigma_\epsilon)}{\int \int E_p(\pi_i | x_i, y_i) f_p(x_i, y_i | \mu, \sigma, \nu, \sigma_\epsilon) dx_i dy_i}. \end{aligned}$$

O valor da esperança pode ser calculado como:

$$\begin{aligned} E_p(\pi_i | x_i, y_i) &= E_p[(x_i - \bar{X})^2 | x_i, y_i] = E_p(x_i^2 - 2x_i\bar{X} + \bar{X}^2 | x_i, y_i) \\ &= x_i^2 - 2x_i E_p(\bar{X} | x_i, y_i) + E_p(\bar{X}^2 | x_i, y_i) \\ &= x_i^2 - 2x_i E_p(\bar{X} | x_i, y_i) + V_p(\bar{X} | x_i, y_i) + E_p(\bar{X} | x_i, y_i)^2 \\ &= [x_i - E_p(\bar{X} | x_i, y_i)]^2 + V_p(\bar{X} | x_i, y_i). \end{aligned}$$

O cálculo de $E_p(\bar{X} | x_i, y_i)$ e $V_p(\bar{X} | x_i, y_i)$ pode ser obtido como:

$$E_p(\bar{X} | x_i, y_i) = \frac{1}{N} \sum_{j \in U} E_p(X_j | x_i, y_i) = \frac{1}{N} \left[x_i + \sum_{j \neq i} E_p(X_j) \right] = \frac{1}{N} [x_i + (N-1)\mu],$$

uma vez que $E_p(X_j | x_i, y_i) = E_p(X_j) = E_p[E_p(X_j | Y_j)] = E_p(Y_j) = \mu, \forall j \neq i$.

$$V_p(\bar{X} | x_i, y_i) = \frac{1}{N^2} \sum_{j \in U} V_p(X_j | x_j, y_j) = \frac{1}{N^2} \sum_{j \neq i} V_p(X_j) = \frac{N-1}{N^2} \left(\sigma^2 \frac{\nu}{\nu-2} + \sigma_\epsilon^2 \right),$$

uma vez que $V_p(X_j) = E_p[V_p(X_j | Y_j)] + V_p[E_p(X_j | Y_j)] = E_p(\sigma_\epsilon) + V_p(Y_j) = \sigma^2 \frac{\nu}{\nu-2} + \sigma_\epsilon$.

Então:

$$\begin{aligned} E_p(\pi_i | x_i, y_i) &= \left\{ x_i - \frac{1}{N} [x_i + (N-1)\mu] \right\}^2 + \frac{N-1}{N^2} \left(\sigma^2 \frac{\nu}{\nu-2} + \sigma_\epsilon^2 \right) \\ &= \left(\frac{N-1}{N} \right)^2 (x_i - \mu)^2 + \frac{N-1}{N^2} \left(\sigma^2 \frac{\nu}{\nu-2} + \sigma_\epsilon^2 \right) \\ &= a(x_i - \mu)^2 + b, \end{aligned}$$

onde $a = \left(\frac{N-1}{N}\right)^2$ e $b = \frac{N-1}{N^2} \left(\sigma^2 \frac{\nu}{\nu-2} + \sigma_\epsilon^2\right)$. Tem-se:

$$\begin{aligned} f_s(y_i, x_i | \mu, \sigma, \nu, \sigma_\epsilon) &= \frac{f_p(x_i, y_i | \mu, \sigma, \nu, \sigma_\epsilon) [a(x_i - \mu)^2 + b]}{\int \int [a(x_i - \mu)^2 + b] f_p(x_i, y_i | \mu, \sigma, \nu, \sigma_\epsilon) dx_i dy_i} \\ &= \frac{f_p(x_i, y_i | \mu, \sigma, \nu, \sigma_\epsilon) [a(x_i - \mu)^2 + b]}{\int f_p(y_i | \mu, \sigma, \nu) \int [a(x_i - \mu)^2 + b] f_p(x_i | y_i, \sigma_\epsilon) dx_i dy_i}; \end{aligned}$$

Resolvendo-se a integral:

$$\begin{aligned} &\int [a(x_i - y_i + y_i - \mu)^2 + b] f_p(x_i | y_i, \sigma_\epsilon) dx_i \\ &= \int \{a[(x_i - y_i)^2 + 2(x_i - y_i)(y_i - \mu) + (y_i - \mu)^2] + b\} f_p(x_i | y_i, \sigma_\epsilon) dx_i \\ &= a \int (x_i - y_i)^2 f_p(x_i | y_i, \sigma_\epsilon) dx_i + 2a(y_i - \mu) \int (x_i - y_i) f_p(x_i | y_i, \sigma_\epsilon) dx_i \\ &\quad + a(y_i - \mu)^2 \int f_p(x_i | y_i, \sigma_\epsilon) dx_i + b \int f_p(x_i | y_i, \sigma_\epsilon) dx_i = \\ &= a\sigma_\epsilon^2 + a(y_i - \mu)^2 + b. \end{aligned}$$

Assim,

$$\begin{aligned} f_s(y_i, x_i | \mu, \sigma, \nu, \sigma_\epsilon) &= \frac{f_p(y_i | \mu, \sigma, \nu) f_p(x_i | y_i, \sigma_\epsilon) [a(x_i - \mu)^2 + b]}{\int f_p(y_i | \mu, \sigma, \nu) [a\sigma_\epsilon^2 + a(y_i - \mu)^2 + b] dy_i} \\ &= \frac{f_p(y_i | \mu, \sigma, \nu) f_p(x_i | y_i, \sigma_\epsilon) [a(x_i - \mu)^2 + b]}{a\sigma_\epsilon^2 + aV(Y_i) + b} \\ &= \frac{f_p(y_i | \mu, \sigma, \nu) f_p(x_i | y_i, \sigma_\epsilon) [a(x_i - \mu)^2 + b]}{a\sigma_\epsilon^2 + a\sigma^2 \frac{\nu}{\nu-2} + b}. \end{aligned}$$

Substituindo-se os valores de a e b na expressão anterior:

$$\begin{aligned} f_s(y_i, x_i | \mu, \sigma, \nu, \sigma_\epsilon) &= \frac{f_p(y_i | \mu, \sigma, \nu) f_p(x_i | y_i, \sigma_\epsilon) \left[\left(\frac{N-1}{N}\right)^2 (x_i - \mu)^2 + \frac{N-1}{N^2} \left(\sigma^2 \frac{\nu}{\nu-2} + \sigma_\epsilon^2\right) \right]}{\left(\frac{N-1}{N}\right)^2 \sigma_\epsilon^2 + \left(\frac{N-1}{N}\right)^2 \sigma^2 \frac{\nu}{\nu-2} + \frac{N-1}{N^2} \left(\sigma^2 \frac{\nu}{\nu-2} + \sigma_\epsilon^2\right)} \\ &= \frac{f_p(y_i | \mu, \sigma, \nu) f_p(x_i | y_i, \sigma_\epsilon) [(N-1)(x_i - \mu)^2 + \sigma^2 \frac{\nu}{\nu-2} + \sigma_\epsilon^2]}{N \left(\sigma^2 \frac{\nu}{\nu-2} + \sigma_\epsilon^2\right)} \\ &= \frac{(N-1) f_p(y_i | \mu, \sigma, \nu) f_p(x_i | y_i, \sigma_\epsilon) (x_i - \mu)^2}{N \left(\sigma^2 \frac{\nu}{\nu-2} + \sigma_\epsilon^2\right)} + \frac{f_p(y_i | \mu, \sigma, \nu) f_p(x_i | y_i, \sigma_\epsilon)}{N}. \end{aligned}$$

Apêndice B

Análise de sensibilidade da distribuição *a priori* do parâmetro graus de liberdade (ν)

Conforme mencionado na Seção 5.2, foi realizado um estudo de simulação considerando outra distribuição *a priori* para ν . A distribuição em questão foi proposta por Juárez e Steel (2010), uma distribuição Gama(2,1/10). Como em alguns planos amostrais propostos necessita-se a existência do momento de segunda ordem, truncou-se a distribuição *a priori*, sendo $\pi(\nu) \propto \frac{\nu}{100} \exp\left\{-\frac{\nu}{10}\right\} I_{(a,b]}(\nu)$, $a = 2$ e $b = 50$. As demais distribuições *a priori* referentes aos outros parâmetros foram as mesmas dos estudos de simulações anteriores, sendo $\mu \sim N(m_0, s_0^2)$; $m_0 = 0$, $s_0^2 = 100^2$; $\sigma \sim \text{Half-Cauchy}(m_1, s_1^2)$; $m_1 = 0$ e $s_1^2 = 25^2$; $\sigma_\epsilon \sim \text{Half-Cauchy}(m_1, s_1^2)$, com $m_1 = 1$ e $s_1^2 = 25^2$.

No presente cenário de simulação, foram utilizadas as mesmas populações anteriormente simuladas da distribuição T-Student no estudo sob abordagem clássica, com os mesmos parâmetros fixados anteriormente e considerando-se os mesmos cenários. Compararam-se os modelos AQS, AASX e AAS.

Novamente, para cada população r simulada, foram obtidas duas cadeias, com valores iniciais diferentes para cada parâmetro. Foram realizadas 40000 iterações, utilizando-se um período de aquecimento de 10000 iterações. Para eliminar a correlação entre os valores restantes foi aplicado o espaçamento de 30 iterações em cada cadeia. Ao final, amostras

das distribuições *a posteriori* de tamanho 2000 foram obtidas para μ , σ e ν para cada uma das 1000 amostras para cada plano amostral considerado.

Foram calculadas as medianas, desvios-padrão, comprimentos dos intervalos HPDs de cada uma das 1000 amostras das distribuições *a posteriori* obtidas dos parâmetros da distribuição T-Student. Assim, para cada parâmetro, tem-se um conjunto de 1000 valores de cada uma das medidas mencionadas acima para os modelos AQS, AASX e AAS. Por fim foram calculadas a média das medianas, média dos desvios-padrão, média dos comprimentos do intervalo, a proporção de quantos intervalos HPD contêm os verdadeiros valores dos parâmetros e o erro absoluto relativo para as estimativas obtidas. Na Figura B.1 é apresentado o boxplot das estimativas obtidas referente ao parâmetro ν .

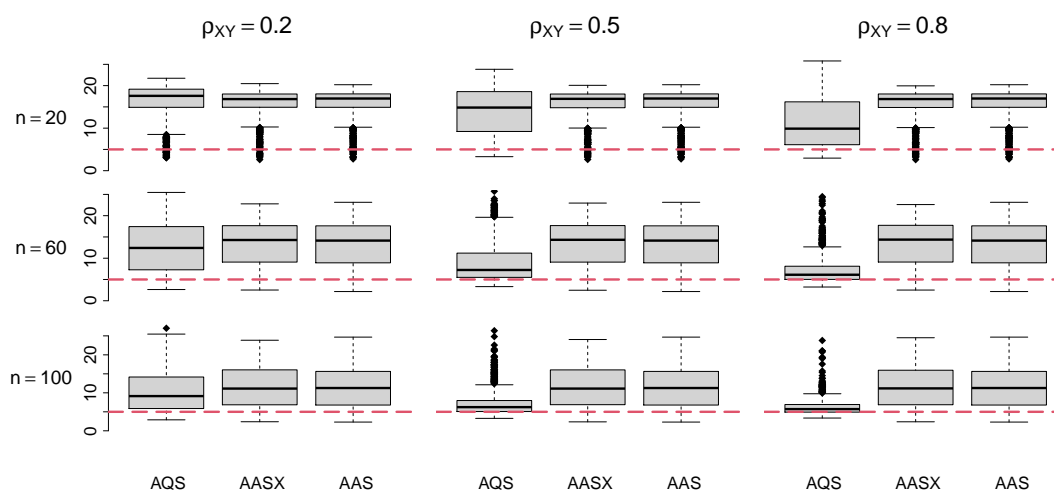


Figura B.1: Boxplots das estimativas bayesianas pontuais de ν obtidas nas simulações, considerando-se os modelos AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,2; 0,5; 0,8$ e $n = 20, 60, 100$.

Destaca-se que na maioria dos cenários o plano amostral AQS obteve um desempenho melhor em relação aos demais modelos. Nota-se que à medida que o tamanho da amostra e correlação aumentam, as estimativas obtidas ao adotar-se o plano amostral AQS são melhores em comparação aos demais modelos.

Nas Tabelas B.1, B.2 e B.3 são apresentados os resultados obtidos referentes aos modelos AQS, AASX e AAS.

Parâmetros	Métodos	$n = 20$					$n = 60$					$n = 100$				
		Med*	DP*	Comp*	Cob*	EAR	Med*	DP*	Comp*	Cob*	EAR	Med*	DP*	Comp*	Cob*	EAR
ν	AQS	16,2373	13,0241	42,4111	1,0000	2,2588	12,5369	9,9246	31,7443	0,9730	1,5438	10,3535	7,6836	24,3158	0,9640	1,1125
	AASX	15,8145	10,6857	37,1352	0,9990	2,1704	13,3796	9,2710	31,7690	0,9520	1,7096	11,5283	7,9763	27,0964	0,9270	1,3587
	AAS	15,7951	10,6688	37,0590	0,9990	2,1666	13,2834	9,2234	31,6093	0,9440	1,6930	11,4651	7,9699	27,0034	0,9240	1,3425
μ	AQS	0,9963	0,2777	1,0873	0,9590	0,2104	1,0017	0,1544	0,6016	0,9510	0,1186	1,0007	0,1181	0,4593	0,9430	0,0937
	AASX	0,9981	0,2900	1,1377	0,9630	0,2090	0,9929	0,1570	0,6121	0,9560	0,1210	0,9976	0,1201	0,4680	0,9570	0,0919
	AAS	0,9975	0,2892	1,1353	0,9640	0,2030	1,0005	0,1563	0,6095	0,9490	0,1199	1,0000	0,1196	0,4654	0,9440	0,0953
σ	AQS	1,1724	0,2304	0,8825	0,8990	0,2241	1,1183	0,1387	0,5378	0,8620	0,1498	1,0930	0,1119	0,4335	0,8630	0,1209
	AASX	1,1698	0,2511	0,9526	0,9180	0,2309	1,1110	0,1422	0,5519	0,9000	0,1452	1,0891	0,1146	0,4449	0,8840	0,1201
	AAS	1,1647	0,2507	0,9490	0,9220	0,2262	1,1032	0,1415	0,5493	0,9050	0,1373	1,0830	0,1144	0,4437	0,8920	0,1140

Tabela B.1: Resultados das simulações para os modelos AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,2$ e $n = 20,60,100$ (abordagem bayesiana).

Parâmetros	Métodos	$n = 20$					$n = 60$					$n = 100$				
		Med*	DP*	Comp*	Cob*	EAR	Med*	DP*	Comp*	Cob*	EAR	Med*	DP*	Comp*	Cob*	EAR
ν	AQS	13,8792	11,0076	35,3717	0,9910	1,8010	8,9430	5,7843	18,2133	0,9470	0,8301	7,2180	3,3605	10,7864	0,9380	0,4908
	AASX	15,8046	10,6841	37,1342	0,9990	2,1682	13,3936	9,2733	31,7622	0,9550	1,7125	11,5196	7,9736	27,0586	0,9210	1,3574
	AAS	15,7951	10,6688	37,0590	0,9990	2,1666	13,2834	9,2234	31,6093	0,9440	1,6930	11,4651	7,9699	27,0034	0,9240	1,3425
μ	AQS	1,0097	0,2502	0,9768	0,9560	0,1928	1,0032	0,1398	0,5439	0,9510	0,1095	0,9995	0,1073	0,4174	0,9470	0,0867
	AASX	0,9982	0,2896	1,1359	0,9630	0,2092	0,9929	0,1572	0,6131	0,9530	0,1214	0,9979	0,1201	0,4672	0,9550	0,0919
	AAS	0,9975	0,2892	1,1353	0,9640	0,2030	1,0005	0,1563	0,6095	0,9490	0,1199	1,0000	0,1196	0,4654	0,9440	0,0953
σ	AQS	1,1852	0,2088	0,8067	0,874	0,2273	1,1033	0,1305	0,5041	0,8690	0,1381	1,0679	0,1022	0,3951	0,8780	0,1005
	AASX	1,1697	0,2514	0,9528	0,9180	0,2313	1,1110	0,1421	0,5513	0,8990	0,1452	1,0890	0,1147	0,4456	0,8860	0,1201
	AAS	1,1647	0,2507	0,9490	0,9220	0,2262	1,1032	0,1415	0,5493	0,9050	0,1373	1,0830	0,1144	0,4437	0,8920	0,1140

Tabela B.2: Resultados das simulações para os modelos AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,5$ e $n = 20,60,100$ (abordagem bayesiana).

Parâmetros	Métodos	$n = 20$					$n = 60$					$n = 100$				
		Med*	DP*	Comp*	Cob*	EAR	Med*	DP*	Comp*	Cob*	EAR	Med*	DP*	Comp*	Cob*	EAR
ν	AQS	11,2796	8,5004	26,9356	0,9510	1,2956	7,2081	3,2805	10,4774	0,9280	0,4981	6,2659	1,7623	5,9166	0,9030	0,311
	AASX	15,8008	10,6870	37,1462	0,9990	2,1677	13,3793	9,2698	31,7883	0,9490	1,7095	11,5316	7,9687	27,0533	0,9250	1,3596
	AAS	15,7951	10,6688	37,0590	0,9990	2,1666	13,2834	9,2234	31,6093	0,9440	1,6930	11,4651	7,9699	27,0034	0,9240	1,3425
μ	AQS	0,9812	0,2316	0,9037	0,9500	0,1766	0,9913	0,1253	0,4879	0,9610	0,0986	0,9949	0,0956	0,3720	0,9410	0,0755
	AASX	0,9982	0,2896	1,1362	0,967	0,2081	0,9930	0,1572	0,6129	0,9560	0,1211	0,9978	0,1201	0,4670	0,9610	0,0920
	AAS	0,9975	0,2892	1,1353	0,9640	0,2030	1,0005	0,1563	0,6095	0,9490	0,1199	1,0000	0,1196	0,4654	0,9440	0,0953
σ	AQS	1,2276	0,2179	0,8390	0,8080	0,2594	1,1006	0,1280	0,4946	0,8550	0,1403	1,0636	0,0978	0,3792	0,8790	0,1021
	AASX	1,1696	0,2514	0,9536	0,9170	0,2311	1,1111	0,1421	0,5512	0,8930	0,1454	1,0892	0,1146	0,4449	0,8950	0,1201
	AAS	1,1647	0,2507	0,9490	0,9220	0,2262	1,1032	0,1415	0,5493	0,9050	0,1373	1,0830	0,1144	0,4437	0,8920	0,1140

Tabela B.3: Resultados das simulações para os modelos AQS, AASX e AAS com parâmetros $\nu = 5$; $\mu = 1$; $\sigma = 1$; $\rho_{XY} = 0,8$ e $n = 20,60,100$ (abordagem bayesiana).

Nota-se que à medida que o tamanho da amostra e correlação aumentam houve uma diminuição média dos desvios-padrão, média dos comprimentos do intervalo e erro absoluto relativo para as estimativas obtidas ao adotar-se o plano amostral AQS.

Constatou-se que ao adotar esta distribuição *a priori* houve uma piora nos resultados obtidos em comparação aos resultados apresentados na subseção 5.2.1. Segundo [Juárez e Steel \(2010\)](#), esta distribuição cobre uma ampla faixa de valores relevantes dos graus de liberdade. Talvez por isso a distribuição em questão não lidou com expectativas de caudas mais pesadas.

Apêndice C

Resultados da aplicação aos dados reais

São apresentados na Figura C1 as cadeias referentes a aplicação aos dados reais do ajuste considerando os dados populacionais da variável de interesse Y . Vale destacar que todas as cadeias atingiram a convergência desejada.

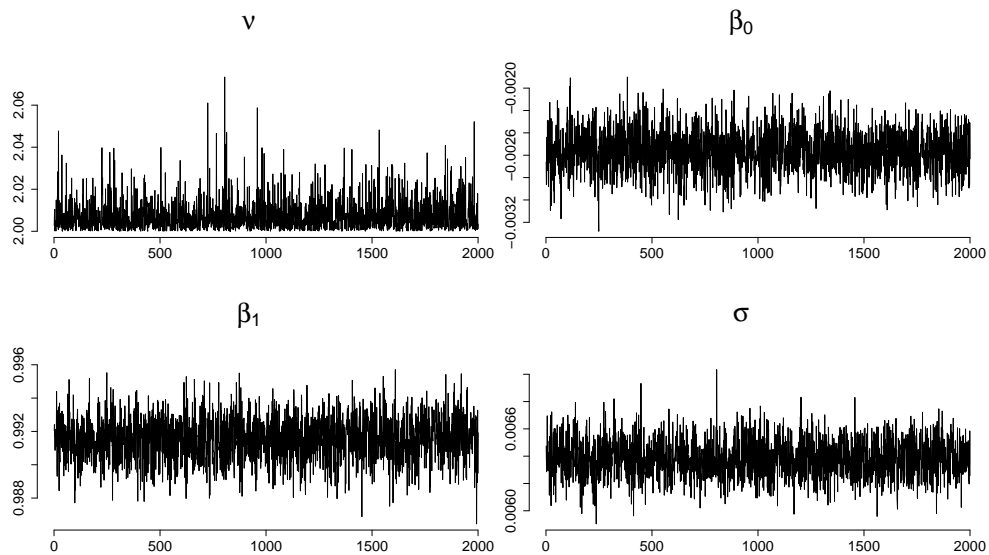


Figura C.1: Amostras obtidas das distribuições *a posteriori* dos parâmetros ν , β_0 , β_1 , β_2 e σ em relação ao ajuste considerando todos os dados populacionais.

São apresentadas nas Figuras C2, C3, C4 e C5 as cadeias referentes a aplicação aos dados reais considerando os modelos AE1, AE2, AASX e AAS. Novamente todas as cadeias atingiram a convergência desejada.

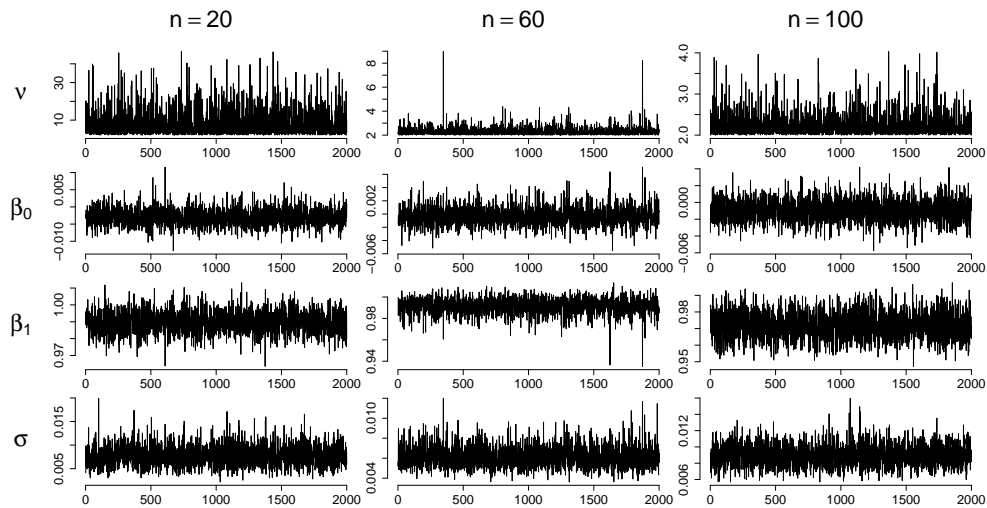


Figura C.2: Amostras obtidas das distribuições *a posteriori* dos parâmetros ν , β_0 , β_1 , β_2 e σ em relação ao modelo AEI1.

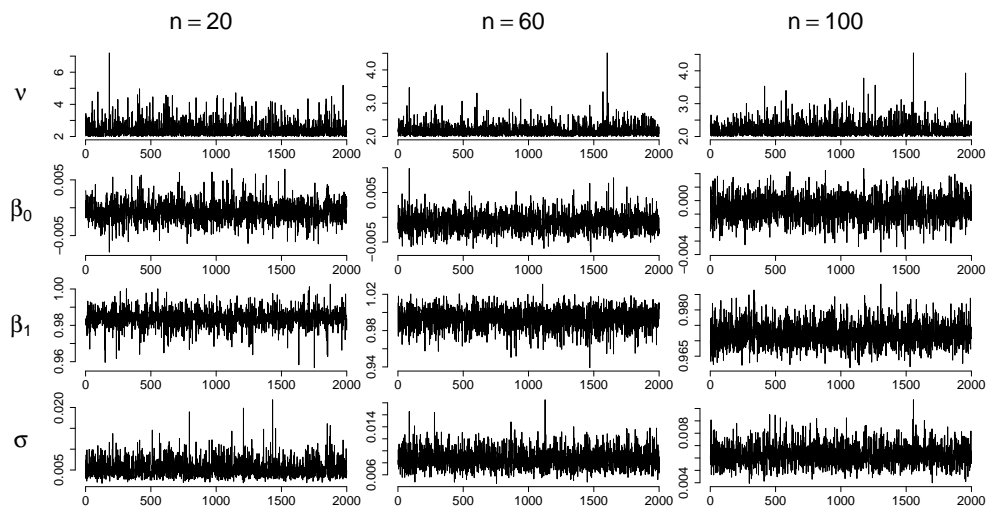


Figura C.3: Amostras obtidas das distribuições *a posteriori* dos parâmetros ν , β_0 , β_1 , β_2 e σ em relação ao modelo AEI2.

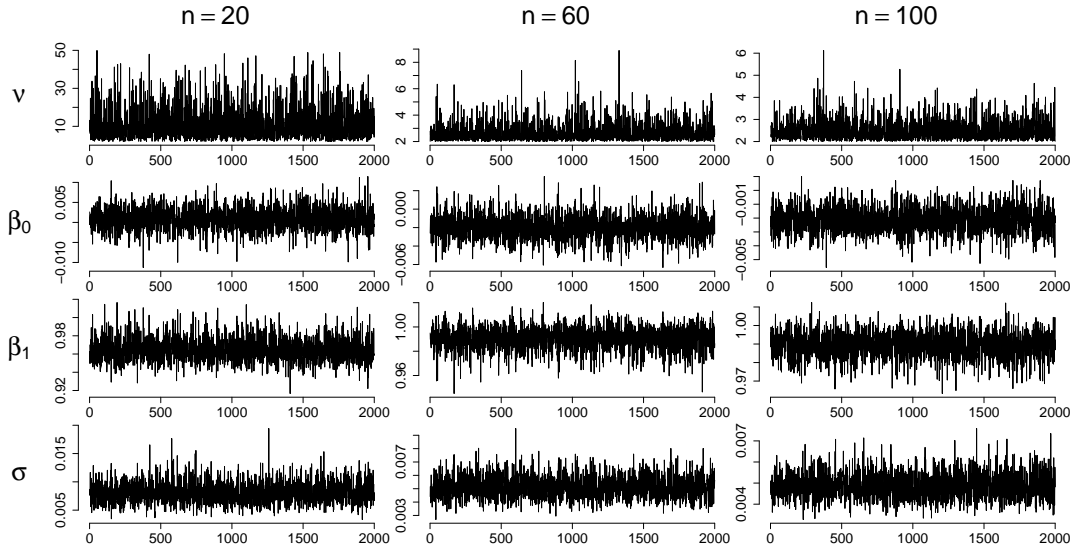


Figura C.4: Amostras obtidas das distribuições *a posteriori* dos parâmetros ν , β_0 , β_1 , β_2 e σ em relação ao modelo AASX.

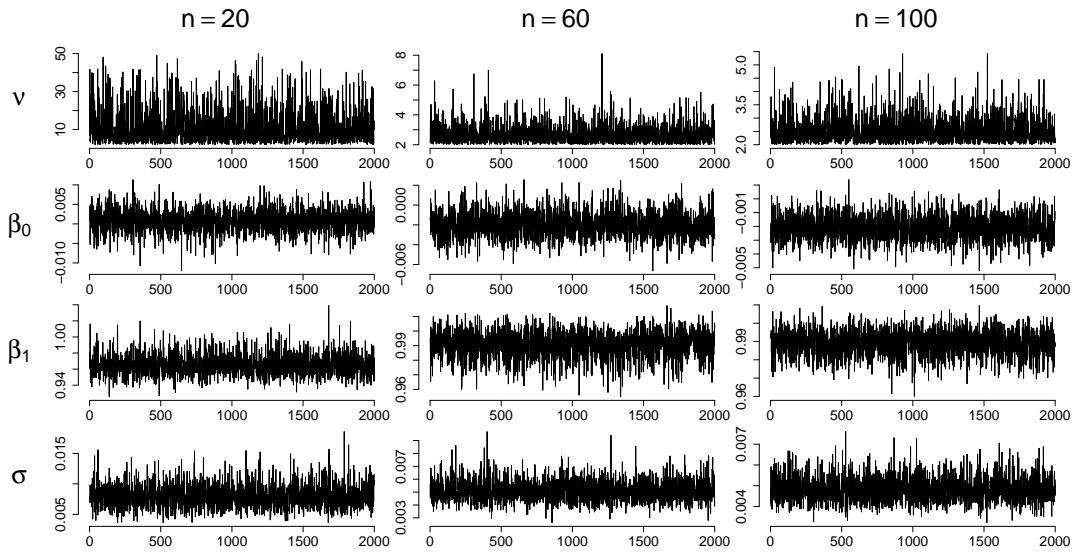


Figura C.5: Amostras obtidas das distribuições *a posteriori* dos parâmetros ν , β_0 , β_1 , β_2 e σ em relação ao modelo AAS.

Abaixo na Tabela C.1 são apresentados as médias, medianas, desvios-padrão e HPDs de 95% dos amostras das distribuições *a posteriori* dos modelos AEI1, AEI2, AASX e AAS:

Parâmetros	Métodos	$n = 20$				$n = 60$				$n = 100$			
		Med	Méd	DP	HPD	Med	Méd	DP	HPD	Med	Méd	DP	HPD
ν	AEI1	5,7805	8,2304	6,9454	[2,0018; 22,6897]	2,2145	2,3142	0,3742	[2,0004; 2,9192]	2,1496	2,2379	0,2734	[2,0001; 2,7767]
	AEI2	2,2527	2,4063	0,4596	[2,0003; 3,3395]	2,1135	2,1689	0,1753	[2,0000; 2,4958]	2,1317	2,1994	0,2132	[2,0003; 2,5943]
	AASX	8,1521	11,1177	8,7091	[2,0063; 29,7874]	2,4808	2,6776	0,6982	[2,0003; 4,0390]	2,3723	2,5010	0,4509	[2,0000; 3,3789]
	AAS	8,5191	11,0731	8,8245	[2,0059; 30,3478]	2,4760	2,6662	0,6369	[2,0001; 3,9143]	2,3819	2,5168	0,4801	[2,0002; 3,4496]
β_0	AEI1	-0,0030	-0,0029	0,0025	[-0,0077; 0,0019]	-0,0004	-0,0004	0,0014	[-0,0029; 0,0026]	-0,0009	-0,0009	0,0013	[-0,0037; 0,0015]
	AEI2	-0,0008	-0,0008	0,0017	[-0,0040; 0,0028]	-0,0011	-0,0010	0,0018	[-0,0042; 0,0028]	-0,0006	-0,0006	0,0009	[-0,0023; 0,0011]
	AASX	0,0008	0,0008	0,0028	[-0,0043; 0,0065]	-0,0020	-0,0020	0,0013	[-0,0046; 0,0006]	-0,0021	-0,0021	0,0009	[-0,0039; -0,0003]
	AAS	0,0009	0,0008	0,0028	[-0,0047; 0,0062]	-0,0020	-0,0020	0,0013	[-0,0044; 0,0005]	-0,0021	-0,0021	0,0009	[-0,0039; -0,0003]
β_1	AEI1	0,9895	0,9896	0,0065	[0,9769; 1,0024]	0,9921	0,9916	0,0074	[0,9758; 1,0048]	0,9721	0,9721	0,0081	[0,9555; 0,9869]
	AEI2	0,9841	0,9840	0,0047	[0,9745; 0,9934]	0,9942	0,9934	0,0110	[0,9723; 1,0151]	0,9717	0,9720	0,0038	[0,9654; 0,9802]
	AASX	0,9635	0,9645	0,0128	[0,9420; 0,9915]	0,9916	0,9907	0,0093	[0,9715; 1,0084]	0,9897	0,9894	0,0069	[0,9751; 1,0024]
	AAS	0,9632	0,9648	0,0131	[0,9399; 0,9916]	0,9917	0,9909	0,0092	[0,9711; 1,0074]	0,9896	0,9891	0,0070	[0,9744; 1,0018]
σ	AEI1	0,0075	0,0077	0,0022	[0,0037; 0,0118]	0,0061	0,0061	0,0011	[0,0043; 0,0083]	0,0089	0,0090	0,0014	[0,0065; 0,0117]
	AEI2	0,0050	0,0054	0,0021	[0,0027; 0,0100]	0,0085	0,0086	0,0016	[0,0059; 0,0121]	0,0061	0,0062	0,0008	[0,0046; 0,0078]
	AASX	0,0079	0,0081	0,0019	[0,0045; 0,0117]	0,0050	0,0051	0,0008	[0,0035; 0,0067]	0,0048	0,0049	0,0006	[0,0039; 0,0062]
	AAS	0,0079	0,0080	0,0019	[0,0045; 0,0117]	0,0050	0,0051	0,0008	[0,0036; 0,0067]	0,0048	0,0048	0,0006	[0,0037; 0,0060]

Tabela C.1: Resultados obtidos para os planos amostrais AEI1, AEI2, AASX e AAS com $n = 20, 60$ e 100 (abordagem bayesiana).

Referências Bibliográficas

- Abanto-Valle, C. A., Bandyopadhyay, D., Lachos, V. H. e Enriquez, I. (2010) Robust bayesian analysis of heavy-tailed stochastic volatility models using scale mixtures of normal distributions. *Computational Statistics & Data Analysis*, **54**, 2883–2898.
- Anderson, P. L. e Meerschaert, M. M. (1998) Modeling river flows with heavy tails. *Water Resources Research*, **34**, 2271–2280.
- Bolfarine, H. (2005) *Elementos de amostragem*. Blucher.
- Bolfarine, H. e Zacks, S. (2012) *Prediction theory for finite populations*. Springer Science & Business Media.
- Boris Choy, S. e Chan, J. S. (2008) Scale mixtures distributions in statistical modelling. *Australian & New Zealand Journal of Statistics*, **50**, 135–146.
- Cassel, C.-M., Särndal, C. E. e Wretman, J. H. (1977) Foundations of inference in survey sampling.
- Ericson, W. A. (1969) Subjective bayesian models in sampling finite populations. *Journal of the Royal Statistical Society: Series B (Methodological)*, **31**, 195–224.
- Fonseca, T. C., Ferreira, M. A. e Migon, H. S. (2008) Objective Bayesian analysis for the student-t regression model. *Biometrika*, **95**, 325–333.
- Gardiner, J. C., Luo, Z., Tang, X. e Ramamoorthi, R. (2014) Fitting heavy-tailed distributions to health care data by parametric and Bayesian methods. *Journal of Statistical Theory and Practice*, **8**, 619–652.

- Gelman, A. (2006) Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian analysis*, **1**, 515–534.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A. e Rubin, D. B. (2013) *Bayesian data analysis*. CRC press.
- Geweke, J. (1993) Bayesian treatment of the independent student-t linear model. *Journal of applied econometrics*, **8**, S19–S40.
- Gneiting, T., Balabdaoui, F. e Raftery, A. E. (2007) Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **69**, 243–268.
- Gómez, Y. M., Bolfarine, H. e Gómez, H. W. (2019) Gumbel distribution with heavy tails and applications to environmental data. *Mathematics and Computers in Simulation*, **157**, 115–129.
- Horvitz, D. G. e Thompson, D. J. (1952) A generalization of sampling without replacement from a finite universe. *Journal of the American statistical Association*, **47**, 663–685.
- Jorgensen, B. (2019) *Theory of linear models*. Routledge.
- Juárez, M. A. e Steel, M. F. (2010) Model-based clustering of non-gaussian panel data based on skew-t distributions. *Journal of Business & Economic Statistics*, **28**, 52–66.
- Kim, D.-H., Lee, I.-S., Sohn, J.-K. e Cho, J.-S. (1996) Bayesian analysis under heavy-tailed priors in finite population sampling. *Communications for Statistical Applications and Methods*, **3**, 225–233.
- Lange, K. L., Little, R. J. e Taylor, J. M. (1989) Robust statistical modeling using the t distribution. *Journal of the American Statistical Association*, **84**, 881–896.
- Lee, S.-A., Suh, S.-H. e Kim, D.-H. (2006) Robust bayesian analysis in finite population sampling with auxiliary information. *Journal of the Korean Data and Information Science Society*, **17**, 1309–1317.

- Lohr Sharon, L. (1999) Sampling: Design and analysis, brooks.
- Melo, W. J. d., Melo, G. M. P. d., Bertipaglia, L. M. A., Brennecke, K., Quintans, N. e Melo, V. P. d. (2018) Boletim técnico da produção animal, 38/universidade brasil: unidades de medida de a a z: conceitos, símbolos, valores e relações.
- Migon, H.S.; Gamerman, D. L. T. (2015) *Statistical Inference. An Integrated Approach*. CRC Press. Taylor& Francis Group, LLC.
- Migon, H. S., Gamerman, D. e Louzada, F. (2014) *Statistical inference: an integrated approach*. CRC press.
- Parsons, V. L. (2014) Stratified sampling. *Wiley StatsRef: Statistics Reference Online*, 1–11.
- Pfeffermann, D., Da Silva Moura, F. A. e Do Nascimento Silva, P. L. (2006) Multi-level modelling under informative sampling. *Biometrika*, **93**, 943–959.
- Pfeffermann, D., Krieger, A. M. e Rinott, Y. (1998) Parametric distributions of complex survey data under informative probability sampling. *Statistica Sinica*, 1087–1114.
- R Core Team (2020) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL<http://www.R-project.org>. ISBN 3-900051-07-0.
- Singh, S. (2003) Simple random sampling. Em *advanced sampling theory with applications*, 71–136. Springer.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P. e Linde, A. (2002) Bayesian measures of model complexity and fit (with discussion). *Journal of Royal Statistical Society B*, **64**, 583–639.
- Stan Development Team (2022) RStan: the R interface to Stan. URL<https://mc-stan.org/>. R package version 2.21.5.
- Valliant, R., Dorfman, A. H., Royall, R. M. et al. (2000) *Finite population sampling and inference: a prediction approach*. No. 04; QA276. 6, V3. John Wiley New York.