

UNIVERSIDADE FEDERAL DE JUIZ DE FORA  
DEPARTAMENTO DE ESTATÍSTICA  
CURSO DE ESTATÍSTICA

Leandro Vitral Andraos

**Modelos de estrutura de covariância para  
dados longitudinais considerando o plano  
amostral complexo**

JUIZ DE FORA

2014

Leandro Vitral Andraos

# Modelos de estrutura de covariância para dados longitudinais considerando o plano amostral complexo

Monografia apresentada ao curso de Estatística da Universidade Federal de Juiz de Fora, como requisito para a obtenção do grau de Bacharel em Estatística.

Orientador: Marcel de Toledo Vieira

Doutor em Estatística – Southampton

JUIZ DE FORA

2014

Andraos, Leandro Vitral – Juiz de Fora, 2014

Modelos de estrutura de covariância para dados longitudinais considerando o plano amostral complexo/ Leandro Vitral Andraos

85.p

Monografia – Universidade Federal de Juiz de Fora e Instituto de Ciências Exatas

Orientador: Marcel de Toledo Vieira

Leandro Vitral Andraos

Modelos de estrutura de covariância para dados  
longitudinais considerando o plano amostral complexo

Monografia apresentada ao curso de  
Estatística da Universidade Federal de Juiz  
de Fora, como requisito para a obtenção do  
grau de Bacharel em Estatística.

BANCA EXAMINADORA

---

Marcel de Toledo Vieira

Ph.D. em Estatística – University of Southampton

---

Ronaldo Rocha Bastos

Ph.D. in Urban and Regional Planning - Liverpool University

---

Clécio da Silva Ferreira

Doutor em Estatística - Universidade de São Paulo

À Deus.

Aos meus pais, Mara Rúbia e Jorge,

à minha irmã querida, Luana,

Dedico com amor e carinho.

## Agradecimentos

Inicialmente agradeço a Deus por sempre iluminar meu caminho, me guiar na direção correta e me ajudar a superar todas as dificuldades encontradas.

Aos meus pais, Mara Rúbia e Jorge, e a minha querida irmã Luana, pelo amor, incentivo e apoio incondicional durante toda minha jornada. Foram vocês que sempre me motivaram a estudar e a continuar. Tenho certeza que se orgulham de mim nesse momento. Mesmo com a distância e a saudade apertando tê-los ao meu lado foi fundamental.

Aos meus avós Maria e Moisés por terem me acolhido e cuidado de mim durante todo esse período e aos meus avós em Resende Michel e Mery por toda confiança que sempre depositaram em mim desde pequeno.

A minha madrinha Margarete, por todo amor, carinho, paciência, por sempre acreditar e cuidar de mim se tornando minha segunda mãe. Até pelas broncas, pois foram importantes para meu amadurecimento.

A minha namorada Taiana pelo suporte nos momentos difíceis, pela compreensão e paciência da minha falta de tempo. Agradeço também toda ajuda que me deu durante esse tempo (inclusive na monografia), mas principalmente pelo carinho, amizade, companheirismo, e amor. Você foi importante demais.

Aos amigos Guilherme (“Didi”), Ruy (“Ruy mesmo”), Daniel (“Dani”) por todas experiências vividas, parcerias, cumplicidade e amizade. Dizem que os amigos são a família que escolhemos, tenho certeza que escolhi muito bem. Inclusive se não fosse pelo Guilherme, eu hoje poderia estar me formando em Física [ou não].

Ao meu orientador e amigo Marcel que merece todo meu reconhecimento e gratidão pela calma, disposição, compreensão e pujança durante todo o curso comigo. Além disso, agradeço muito todos os convites e oportunidades dadas a mim. Dizem por aí que o orientador acaba virando um “Pai”, realmente pude comprovar isso.

Aos amigos da Estatística Amanda (vamos fazer uma pesquisa?), Manoel

(verdadeiro Paizão, sempre com um bom conselho e preocupação comigo) e Thalita (sempre disposta a ajudar e agradeço também as caronas).

Ao grande Diego, amigo de todas as horas, sempre me ajudando com a matemática pesada. Obrigado por sempre me receber em sua casa com grande prazer.

Ao José Márcio pela oportunidade de trabalhar com qualidade de dados, de me ajudar e apresentar ao mestrado da COPPE - UFRJ.

Ao grande Márcio por toda parceria, experiência, ajuda, ensinamentos e ideias compartilhadas. Obrigado também pela amizade.

Ao Professor Ronaldo por todos os conselhos e por me aceitar com grande prazer fazer parte como bolsista voluntário de sua iniciação científica. Foi um prazer estudar e aprender outra área da Estatística. Agradeço de verdade a oportunidade.

Ao Professor Clécio por toda disposição e boa vontade em tirar quaisquer dúvidas sobre disciplinas durante o curso. Mesmo atarefado sempre parava o que estava fazendo para nos receber. Obrigado por oferecer uma disciplina de tópicos ao BI por nossa causa.

À professora Camila por ser uma excelente educadora. Ao Professor Lupércio pelos conselhos e por compartilhar um pouco de sua experiência. Ao professor Joaquim, pela vontade, disposição e empenho dado às aulas. E a todos os demais professores do Departamento de Estatística pelas aulas ministradas a mim e que me possibilitaram o âmbito do conhecimento.

A todos familiares e amigos que, direta ou indiretamente, contribuíram para a realização dessa jornada.

O meu muito obrigado.

“Aqueles que se sentem satisfeitos sentam-se e nada fazem.

Os insatisfeitos são os únicos benfeitores do mundo.”

(Walter S. Landor)

O que eu ouço, eu esqueço.

O que eu vejo, eu não lembro.

Mas o que eu faço, eu aprendo.

(Provérbio chinês)

"Há pessoas que transformam o sol numa simples mancha amarela,  
mas há também aquelas que fazem de uma simples mancha amarela o próprio sol."

(Pablo Picasso)

## Resumo

Nesta monografia são estudados métodos estatísticos com o apoio de ferramentas computacionais para o estudo de dados do tipo painel oriundos da pesquisa *British Household Panel Survey* (BHPS). Procedimentos de inferência estatística para parâmetros de modelos de efeitos mistos são avaliados considerando diferentes estruturas longitudinais de covariância. Comparações entre métodos de estimação que consideram o plano amostral complexo e aqueles que o ignoram são realizadas através de um estudo de simulação e o comportamento de estimadores que tem como base funções de ajuste é avaliado. São considerados os seguintes métodos de estimação pontual: máxima verossimilhança, máxima pseudo verossimilhança e mínimos quadrados generalizados. O desempenho dos estimadores também é avaliado considerando cenários em que os dados são gerados a partir de diferentes distribuições de probabilidade, incluindo tanto distribuições simétricas quanto assimétricas.

**Palavras-chave:** dados do tipo painel; modelos de efeitos mistos; estruturas de covariância; funções de ajuste; inferência para dados amostrais complexos; BHPS.

## Abstract

In this monograph statistical methods are studied with the support of computational tools with the aim of analysing panel data from the British Household Panel Survey (BHPS). Statistical inference procedures for parameters of mixed models are evaluated considering different longitudinal covariance structures. Comparisons between estimation methods that consider the complex sampling design and those which ignore it are carried out through a simulation study and the behavior of estimators which are based upon adjustment functions is assessed. The following estimation methods are considered: maximum likelihood, maximum pseudo-likelihood and generalized least squares. The performance of the estimators is also evaluated for scenarios which consider data being generated from different probability distributions, including both symmetrical and asymmetrical distributions.

**Keywords:** Longitudinal data, mixed models, covariance structures; fitting function; inference for complex survey data; BHPS.

# Sumário

Lista de Figuras .....	12
Lista de Tabelas .....	13
1- Introdução.....	16
1.1- Motivação e Objetivos .....	16
1.2- Dados BHPS. ....	17
1.3- Estrutura da Monografia .....	19
2- Modelo de Regressão Longitudinal.....	21
2.1- Introdução.....	21
2.2- Modelo de Efeitos Mistos.....	22
2.3.1- Modelo de Correlação Uniforme .....	25
2.3.2- Modelo Autoregressivo.....	27
3- Modelos Longitudinais de Estrutura de Covariância .....	29
3.1-Introdução.....	29
3.2-Estimação de $\underline{\beta}$ – Máxima Verossimilhança .....	30
3.3-Estimação de $\underline{\beta}$ – Máxima Pseudo-Verossimilhança.....	32
3.4- Estimação de $\Sigma$ .....	33
3.5- Estimação de $\underline{\theta}$ .....	35
3.5.1- Máxima Verossimilhança .....	36
3.5.2- Máxima Pseudo Verossimilhança.....	38
3.5.3- Mínimos Quadrados Generalizados .....	39
4- Estudo de Simulação.....	41
4.1- Introdução.....	41
4.2- Resultados.....	44
4.2.1- Resíduos com Distribuição Normal.....	46
4.2.2- Resíduos com Distribuição Normal Assimétrica.....	53
4.2.3- Resíduos com Distribuição t.....	62
4.2.4- Resíduos com Distribuição t Assimétrica.....	69
5- Considerações Finais.....	78
Referências Bibliográficas.....	81
Apêndice A .....	85

## **Lista de Figuras**

1.1 Exemplo de estrutura hierárquica.....	23
---	----

## Lista de Tabelas

1.1- Categorias das variáveis e sua descrição.....	18
5.1 - Estimação de $\hat{\theta}$ para resíduos com distribuição normal (replicações geradas pelo modelo MCU) e $n_{sim} = 100$ .....	46
5.2 - Estimação de $\hat{\theta}$ para resíduos com distribuição normal (replicações geradas pelo modelo MCU) e $n_{sim} = 500$ .....	47
5.3 - Estimação de $\hat{\theta}$ para resíduos com distribuição normal (replicações geradas pelo modelo MCU) e $n_{sim} = 10000$ .....	48
5.4 - Estimação de $\hat{\theta}$ para resíduos com distribuição normal (replicações geradas pelo modelo MCU) e $n_{sim} = 15000$ .....	49
5.5 - Estimação de $\hat{\theta}$ para resíduos com distribuição normal (replicações geradas pelo modelo MCU) e $n_{sim} = 20000$ .....	49
5.6 - Estimação de $\hat{\theta}$ para resíduos com distribuição normal (replicações geradas pelo modelo MCU-C) e $n_{sim} = 100$ .....	50
5.7 - Estimação de $\hat{\theta}$ para resíduos com distribuição normal (replicações geradas pelo modelo MCU-C) e $n_{sim} = 500$ .....	51
5.8 - Estimação de $\hat{\theta}$ para resíduos com distribuição normal (replicações geradas pelo modelo MCU-C) e $n_{sim} = 10000$ .....	51
5.9 - Estimação de $\hat{\theta}$ para resíduos com distribuição normal (replicações geradas pelo modelo MCU-C) e $n_{sim} = 15000$ .....	52
5.10 - Estimação de $\hat{\theta}$ para resíduos com distribuição normal (replicações geradas pelo modelo MCU-C) e $n_{sim} = 20000$ .....	53
5.11 - Estimação de $\hat{\theta}$ para resíduos com distribuição normal assimétrica (replicações geradas pelo modelo MCU) e $n_{sim} = 100$ .....	54
5.12 - Estimação de $\hat{\theta}$ para resíduos com distribuição normal assimétrica (replicações geradas pelo modelo MCU) e $n_{sim} = 500$ .....	55
5.13 - Estimação de $\hat{\theta}$ para resíduos com distribuição normal assimétrica (replicações geradas pelo modelo MCU) e $n_{sim} = 10000$ .....	56

5.14 - Estimaco de $\hat{\theta}$ para resduos com distribuico normal assimtrica (replicaces geradas pelo modelo MCU) e $n_{sim} = 15000$ .....	57
5.15 - Estimaco de $\hat{\theta}$ para resduos com distribuico normal assimtrica (replicaces geradas pelo modelo MCU) e $n_{sim} = 20000$ .....	57
5.16 - Estimaco de $\hat{\theta}$ para resduos com distribuico normal assimtrica (replicaces geradas pelo modelo MCU-C) e $n_{sim} = 100$ .....	58
5.17 - Estimaco de $\hat{\theta}$ para resduos com distribuico normal assimtrica (replicaces geradas pelo modelo MCU-C) e $n_{sim} = 500$ .....	59
5.18 - Estimaco de $\hat{\theta}$ para resduos com distribuico normal assimtrica (replicaces geradas pelo modelo MCU-C) e $n_{sim} = 10000$ .....	60
5.19 - Estimaco de $\hat{\theta}$ para resduos com distribuico normal assimtrica (replicaces geradas pelo modelo MCU-C) e $n_{sim} = 15000$ .....	61
5.20 - Estimaco de $\hat{\theta}$ para resduos com distribuico normal assimtrica (replicaces geradas pelo modelo MCU-C) e $n_{sim} = 20000$ .....	62
5.21 - Estimaco de $\hat{\theta}$ para resduos com distribuico $t_{\nu=4}(0,1)$ (replicaces geradas pelo modelo MCU) e $n_{sim} = 100$ .....	63
5.22 - Estimaco de $\hat{\theta}$ para resduos com distribuico $t_{\nu=4}(0,1)$ (replicaces geradas pelo modelo MCU) e $n_{sim} = 500$ .....	63
5.23 - Estimaco de $\hat{\theta}$ para resduos com distribuico $t_{\nu=4}(0,1)$ (replicaces geradas pelo modelo MCU) e $n_{sim} = 10000$ .....	64
5.24 - Estimaco de $\hat{\theta}$ para resduos com distribuico $t_{\nu=4}(0,1)$ (replicaces geradas pelo modelo MCU) e $n_{sim} = 15000$ .....	65
5.25 - Estimaco de $\hat{\theta}$ para resduos com distribuico $t_{\nu=4}(0,1)$ (replicaces geradas pelo modelo MCU) e $n_{sim} = 20000$ .....	65
5.26 - Estimaco de $\hat{\theta}$ para resduos com distribuico $t_{\nu=4}(0,1)$ (replicaces geradas pelo modelo MCU-C) e $n_{sim} = 100$ .....	66
5.27 - Estimaco de $\hat{\theta}$ para resduos com distribuico $t_{\nu=4}(0,1)$ (replicaces geradas pelo modelo MCU-C) e $n_{sim} = 500$ .....	67

5.28 - Estimaco de $\hat{\theta}$ para resduos com distribuico $t_{\nu=4}(0,1)$ (replicaes geradas pelo modelo MCU-C) e $n_{sim} = 10000$ .....	67
5.29 - Estimaco de $\hat{\theta}$ para resduos com distribuico $t_{\nu=4}(0,1)$ (replicaes geradas pelo modelo MCU-C) e $n_{sim} = 15000$ .....	68
5.30 - Estimaco de $\hat{\theta}$ para resduos com distribuico $t_{\nu=4}(0,1)$ (replicaes geradas pelo modelo MCU-C) e $n_{sim} = 20000$ .....	69
5.31 - Estimaco de $\hat{\theta}$ para resduos com distribuico $t_{\nu=4}(0,1)$ assimtrica (replicaes geradas pelo modelo MCU) e $n_{sim} = 100$ .....	70
5.32 - Estimaco de $\hat{\theta}$ para resduos com distribuico $t_{\nu=4}(0,1)$ assimtrica (replicaes geradas pelo modelo MCU) e $n_{sim} = 500$ .....	71
5.33 - Estimaco de $\hat{\theta}$ para resduos com distribuico $t_{\nu=4}(0,1)$ assimtrica (replicaes geradas pelo modelo MCU) e $n_{sim} = 10000$ .....	71
5.34 - Estimaco de $\hat{\theta}$ para resduos com distribuico $t_{\nu=4}(0,1)$ assimtrica (replicaes geradas pelo modelo MCU) e $n_{sim} = 15000$ .....	72
5.35 - Estimaco de $\hat{\theta}$ para resduos com distribuico $t_{\nu=4}(0,1)$ assimtrica (replicaes geradas pelo modelo MCU) e $n_{sim} = 20000$ .....	73
5.36 - Estimaco de $\hat{\theta}$ para resduos com distribuico $t_{\nu=4}(0,1)$ assimtrica (replicaes geradas pelo modelo MCU-C) e $n_{sim} = 100$ .....	74
5.37 - Estimaco de $\hat{\theta}$ para resduos com distribuico $t_{\nu=4}(0,1)$ assimtrica (replicaes geradas pelo modelo MCU-C) e $n_{sim} = 500$ .....	75
5.38 - Estimaco de $\hat{\theta}$ para resduos com distribuico $t_{\nu=4}(0,1)$ assimtrica (replicaes geradas pelo modelo MCU-C) e $n_{sim} = 10000$ .....	75
5.39 - Estimaco de $\hat{\theta}$ para resduos com distribuico $t_{\nu=4}(0,1)$ assimtrica (replicaes geradas pelo modelo MCU-C) e $n_{sim} = 15000$ .....	76
5.40 - Estimaco de $\hat{\theta}$ para resduos com distribuico $t_{\nu=4}(0,1)$ assimtrica (replicaes geradas pelo modelo MCU-C) e $n_{sim} = 20000$ .....	77

# Introdução

Atualmente, há uma intensa e incessante busca por fontes de dados e informações coletadas por instituições governamentais, organizações, empresas, universidades e centros de pesquisa. Muitas dessas informações precisam ser coletadas e atualizadas ao longo do tempo já que temos um mundo em constante transformação. Com isso, estão cada vez mais em voga as pesquisas longitudinais.

Pesquisas longitudinais são definidas como levantamentos em que a variável resposta é medida repetidamente para os mesmos indivíduos em diferentes momentos de tempo. Em tais estudos, as observações realizadas para o mesmo indivíduo não são, em geral, independentes, o que leva à necessidade de aplicarmos métodos estatísticos específicos que permitem a consideração de uma possível correlação entre observações repetidas. Podemos pensar em um exemplo simples sobre a renda. Em geral, espera-se que a renda de um indivíduo em um certo mês esteja correlacionada com sua renda no mês seguinte. O modelo estatístico utilizado para estudar o comportamento da renda precisa, então, captar e trabalhar bem com essa correlação temporal.

## 1.1 - Motivação e Objetivos

O que torna as pesquisas longitudinais tão populares e indispensáveis atualmente? Primeiramente, dados longitudinais são as principais fontes de informações referentes às mudanças das características demográficas, sociais, políticas e econômicas da população ao longo do tempo. Diferentemente dos dados transversais (que são coletados em apenas uma ocasião), as pesquisas longitudinais permitem o estudo de causalidade. Se analisarmos dados transversais conseguiremos estudar a associação, por exemplo, entre a ocorrência de problemas cardíacos e a ingestão de alimentos com gordura trans. No entanto, se pesquisássemos os mesmos indivíduos ao longo do tempo seríamos capazes de avaliar se o consumo de gordura trans no passado pode causar problemas cardíacos no presente ou no futuro. Assim, o método de observação longitudinal favorece o acompanhamento de mudanças que ocorrem ao longo do tempo.

Uma maneira interessante de perceber as características das pesquisas longitudinais é que as mesmas podem ser comparadas com filmes, enquanto que as pesquisas transversais se assemelham mais a fotografias. Além disso, as pesquisas longitudinais permitem a análise de mudanças nos níveis micro (nível do indivíduo) e macro (nível da população) possibilitando que se separe nas análises a variação intra-indivíduo ao longo do tempo da variação entre indivíduos ao longo do tempo.

Nesta monografia são utilizados métodos estatísticos com o apoio de ferramentas computacionais para o estudo de dados do tipo painel oriundos da pesquisa *British Household Panel Survey* (BHPS), que será descrita na Seção 1.2. Desta forma, são estudados procedimentos de inferência para modelos de efeitos mistos considerando diferentes estruturas longitudinais de covariância.

Comparações entre métodos que consideram o plano amostral complexo e aqueles que o ignoram são realizadas através de um estudo de simulação, no que diz respeito ao comportamento de estimadores que tem como base funções de ajuste. Dados amostrais são ‘complexos’ porque as pesquisas de onde são obtidos geralmente envolvem estratificação, conglomeração, observações com pesos desiguais, ajustes para não resposta, probabilidades desiguais de seleção, dentre outras características.

São considerados os seguintes métodos de estimação pontual: máxima verossimilhança, máxima pseudo verossimilhança e mínimos quadrados generalizados. O desempenho dos estimadores também é avaliado em situações em que os dados são gerados a partir de diferentes distribuições de probabilidade, incluindo tanto distribuições simétricas quanto assimétricas.

## **1.2 - British Household Panel Survey**

Os dados utilizados nesta monografia são provenientes da BHPS que é uma pesquisa painel realizada pelo Instituto de Pesquisas Sociais e Econômicas da Universidade de Essex (ver Taylor *et al.*, 2001; e Berrington, 2002). O principal objetivo da BHPS é promover a compreensão de mudanças sociais e econômicas nos níveis do domicílio e do indivíduo na Grã-Bretanha.

A coleta dos dados desta pesquisa iniciou-se em 1991 e vem sendo realizada anualmente. O desenho amostral adotado é estratificado e conglomerado em múltiplos estágios. Na primeira ocasião (1991) foram selecionados 250 setores postais com probabilidades de seleção proporcionais ao tamanho, utilizando um procedimento sistemático com reposição para compor as unidades primárias de amostragem (UPAs). Para as unidades secundárias de amostragem (USAs) foram considerados os endereços, selecionados também por amostragem sistemática. Maiores informações sobre o desenho amostral da BHPS podem ser encontradas em Vieira (2009).

O banco de dados é composto por variáveis contínuas e categóricas sobre homens e mulheres que deram respostas completas nas ocasiões um, três, cinco, sete e nove da BHPS, coletadas respectivamente em 1991, 1993, 1995, 1997 e 1999. As variáveis coletadas consideradas nesta monografia são: pid (identificação do indivíduo), aage (idade na primeira ocasião), score91-99 (escores do papel do gênero em 1991, 1993, 1995, 1997 e 1999, descritos a seguir), asex (sexo), qualif91-99 (nível de escolaridade em 1991, 1993, 1995, 1997 e 1999), amumwk (se a mãe do respondente trabalhava quando ele tinha 14 anos de idade), eact91-99 (atividade econômica em 1991, 1993, 1995, 1997 e 1999). Maiores informações sobre as variáveis são apresentadas na Tabela 1 abaixo. Além disso, o banco de dados da BHPS inclui uma variável de peso longitudinal que permite a consideração do desenho amostral e correções para não resposta em estimadores pontuais.

Variável	Valores	Descrição
Faixa etária	1	16-21 anos
	2	22-27 anos
	3	28-33 anos
	4	Mais de 34 anos
Atividade econômica	1	Emprego em tempo integral
	2	Emprego em tempo parcial
	3	Inativa
	4	Estudante em tempo integral
	5	Dona de casa
Maternidade	0	Não
	1	Sim
Nível educacional	1	Curso superior ou mais
	2	Magistério
	3	Ensino Médio
	4	Ensino Fundamental
	5	Outras

Tabela 1.1 – Categorias das variáveis e sua descrição.

O escore de atitudes em relação ao papel do gênero é a variável resposta dos modelos ajustados nesta monografia. Essa medida foi considerada originalmente por Berrington (2002). Os respondentes expressaram suas opiniões se “concordam fortemente”, “concordam”, “não concordam e não discordam”, “discordam” ou “discordam fortemente” com cada afirmação sobre o papel do gênero no domicílio. Análise de fatores (Skinner e Vieira, 2007) e análise de correspondência (Vieira *et al.*, 2011) foram utilizados para estabelecer quais afirmações poderiam compor a medida de atitudes em relação ao papel do gênero. O escore de atitudes foi construído a partir da soma dos valores das seis primeiras afirmações consideradas pela BHPS e valores maiores representam atitudes mais igualitárias em relação ao papel do gênero. Detalhes sobre as afirmações originais são apresentados no Apêndice A.

É interessante ainda notar que como a população alvo é dinâmica, a BHPS procura pesquisar os mesmos indivíduos em todas as ocasiões e, caso haja alguma alteração nos domicílios, a entrevista é feita com todos os novos membros adultos além dos respondentes originais previamente incluídos.

### **1.3 - Estrutura da Monografia**

A presente monografia está organizada em 5 capítulos. O capítulo atual fornece informações sobre a importância e utilização de dados longitudinais e os objetivos do trabalho, além de uma breve explicação sobre as características da pesquisa da BHPS.

O restante da monografia está organizado em sequência. O Capítulo 2 aborda modelos de regressão longitudinal que levam em consideração o plano amostral. O modelo de efeitos mistos é mencionado assim como alguns de seus casos particulares que são utilizados ao longo do trabalho.

O Capítulo 3 apresenta procedimentos de estimação para parâmetros do modelo de regressão longitudinal. Os métodos discutidos são os da máxima verossimilhança, máxima pseudo verossimilhança e mínimos quadrados generalizados, sob a abordagem do plano amostral complexo. Além disso, vamos descrever os métodos de inferência para matrizes de covariância.

O Capítulo 4 apresenta as características e os resultados de um estudo de simulação, que tem o objetivo principal de avaliar as propriedades estatísticas dos procedimentos de estimação pontual discutidos no Capítulo 3. Comparamos as propriedades dos métodos estudados considerando dados gerados a partir de diferentes distribuições de probabilidade, tanto simétricas quanto assimétricas.

Por fim, o Capítulo 5 traz as conclusões e considerações finais além de perspectivas futuras de pesquisa.

## 2 - Modelo de Regressão Longitudinal

### 2.1- Introdução

Modelos de regressão longitudinal permitem relacionar características de indivíduos em um dado momento no tempo com características observadas em algum outro momento de tempo. Entre os modelos mais conhecidos na literatura da Estatística temos o modelo marginal e o modelo de efeitos mistos (também chamado de modelo de efeitos aleatórios, modelo linear misto, modelo multinível ou modelo de coeficientes aleatórios). Parâmetros de modelos de efeitos mistos são úteis para a avaliação de mudanças no nível individual enquanto parâmetros de modelos marginais são úteis para avaliação de mudanças agregadas (médias, ou diferenças médias entre grupos diferentes na população). Nessa monografia apenas o modelo de efeitos mistos é considerado.

A utilização de um bom modelo estatístico depende tanto da capacidade técnica quanto da sensibilidade do analista para captar dos dados a sua estrutura que deverá ser modelada, em relação aos efeitos aleatórios, ou seja, efeitos não observáveis. Essa modelagem pode ser feita através dos modelos de efeitos mistos que têm sido amplamente estudados por muitos autores, tais como, por exemplo, Diggle *et al.* (2002), Lian e Zeger (1986) e Vieira (2009).

Para definir o modelo começamos, então, denotando a população finita por  $\bar{v}$ , que é fixo nas ocasiões  $1, \dots, T$ . Seja  $N$  o tamanho de  $\bar{v}$  e  $N_0 = NT$ .

Seja  $\underline{Y}_i = (Y_{i1}, \dots, Y_{iT})'$  um vetor aleatório contendo  $T$  observações repetidas da variável de estudo para as unidades  $i=1, 2, \dots, N$  nas  $T$  ocasiões da pesquisa. Trabalhamos com o pressuposto de  $Y$  se comportar de acordo com o seguinte modelo (Vieira, 2009)

$$E(\underline{Y}_i) = \underline{\mu}_i(\underline{\beta}) \quad (2.1)$$

onde

$$\underline{\mu}_i(\underline{\beta}) = [\mu(\underline{x}_{i1}, \underline{\beta}), \dots, \mu(\underline{x}_{iT}, \underline{\beta})]' \quad (2.2)$$

Na equação acima,  $\underline{x}_{it}$  é um vetor  $1 \times q$  com  $q$  covariáveis fixas,  $\underline{\beta}$  é um vetor  $q \times 1$  de parâmetros desconhecidos e a pressupõem-se que  $\mu(\cdot, \cdot)$  tem forma conhecida, com  $\underline{y}_i$  sendo os valores observados de  $\underline{Y}_i$ .

Tradicionalmente, os métodos de estimação adotados no contexto longitudinal trabalham com o pressuposto de que  $\underline{Y}_i = (Y_{i1}, \dots, Y_{iT})$  tem distribuição uma normal multivariada  $T$ -dimensional, denotada por

$$\underline{Y}_i \sim N_T[\underline{\mu}_i(\underline{\beta}), \Sigma]. \quad (2.3)$$

Além disso, o modelo de estrutura de covariância populacional é dado pela seguinte matriz simétrica  $T \times T$

$$\Sigma = COV(\underline{Y}_i) = E\{[\underline{Y}_i - \underline{\mu}_i][\underline{Y}_i - \underline{\mu}_i]'\} = \Sigma(\underline{\theta}), \quad (2.4)$$

onde  $COV(\cdot)$  denota a covariância da população. Estamos pressupondo que  $\Sigma$  não depende de  $i$ .

Na Seção 2.1 o modelo de efeitos mistos é explicado enquanto nas Seções 2.2 e 2.3 dois de seus modelos particulares são discutidos.

## 2.1- Modelo de Efeitos Mistos

Em diversas aplicações é possível encontrar dados organizados de forma hierárquica, ou seja, casos agrupados em unidades de níveis mais altos que incluem as unidades de níveis mais baixos, e assim sucessivamente. Vários níveis podem ser distinguíveis. Ao analisar este tipo de dados é fundamental levar em consideração a sua estrutura hierárquica, já que ignorá-la pode implicar na má estimação dos coeficientes do modelo em questão.

Quando os dados estão dispostos de forma hierárquica, pode ser que exista uma correlação dentre os níveis de hierarquia, sendo que através de uma análise simples, essa correlação, pode não ser detectada. Isso ocorre porque essas unidades apresentam características semelhantes em cada nível, assim a suposição de independência pode ser

violada e passa a existir correlação entre as unidades. É importante, então, que o modelo seja ajustado levando em consideração a correlação resultante.

Por exemplo, ao investigarmos escores de atitudes de indivíduos que estão agrupados em domicílios é possível dizer que as atitudes não são, em geral, independentes entre indivíduos de um mesmo domicílio. Similarmente, as características de domicílios em um mesmo setor postal (ou setor censitário) podem ser dependentes. Logo, devemos considerar alguma correção nos métodos de estimação de forma a levar em consideração essa correlação. O modelo pode, por exemplo, permitir diferentes coeficientes de regressão para diferentes setores postais, ou seja, é possível que haja diferentes coeficientes de regressão para captar a variabilidade entre os indivíduos.

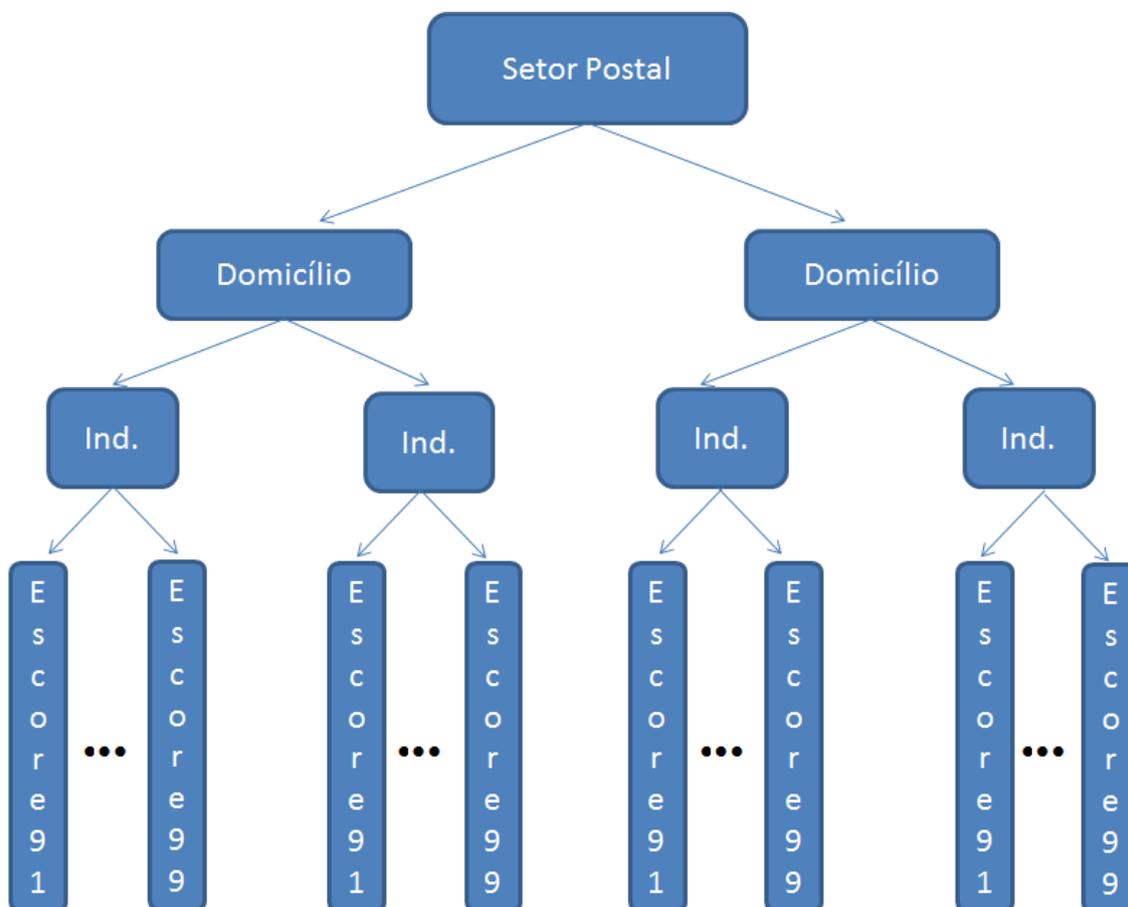


Figura 1.1: Exemplo de estrutura hierárquica.

A estrutura apresentada na Figura 1.1 (mesma considerada na BHPS, na Seção 1.2) pode ser considerada através do modelo de efeitos mistos onde as unidades de cada

nível são vistas como tendo um efeito aleatório. Estes efeitos aleatórios implicam em coeficientes aleatórios no modelo que permitem levarmos em conta a variabilidade entre unidades, seja de forma simples, através da variabilidade nos interceptos, ou de forma mais complexa, através da variabilidade nas inclinações. A ideia geral do modelo é que a variância não explicada da variável resposta  $y$  seja dividida em diferentes componentes. Uma das componentes diz respeito ao intercepto aleatório e a outra componente reflete a inclinação aleatória. Nessa monografia apenas os efeitos aleatórios do intercepto são considerados, assim como foi feito em Vieira (2009).

Outras discussões sobre modelos de efeitos mistos, incluindo análise fatorial e modelos multivariados, podem ser encontradas em Aitkin e Longford (1986).

De forma simplificada podemos descrever o modelo de efeitos mistos como:

$$Y_{it} = \underline{x}_{it} \underline{\beta} + \varepsilon_{it}, \quad (2.5)$$

onde  $\varepsilon_{it}$  é o termo de erro do modelo e  $E(\varepsilon_{it}) = 0$ .

Em (2.5),  $i$  representa o indivíduo,  $t$  a ocasião,  $x_{it}$  é um vetor  $1 \times q$  com  $q$  covariáveis,  $\beta$  é um vetor  $q \times 1$  com coeficientes desconhecidos para as covariáveis  $x$  e  $\varepsilon_{it}$  é termo de resíduo que pode ser decomposto nos diferentes efeitos aleatórios considerados no modelo.

A estrutura de covariância do modelo irá depender da forma como a correlação entre os resíduos ao longo do tempo será trabalhada. Temos diversas matrizes de correlação que podem ser consideradas para modelar os resíduos, mas nos limitamos a apresentar as estruturas uniforme (MCU) e exponencial (AR).

Pressupomos ainda que (i) as observações são igualmente espaçadas no tempo, e que (ii) o número de indivíduos na população,  $N$ , é relativamente grande em relação ao número de observações por indivíduos ( $T$ ). Supomos também que (iii) a amostra é selecionada por um método de amostragem probabilística em uma ocasião e que então as mesmas unidades amostrais retornam em cada  $T-1$  ocasiões subsequentes na pesquisa, assim como foi feito em Vieira (2009).

### 2.1.1- Modelo de Correlação Uniforme (MCU)

Um caso particular do modelo discutido acima pode ser representado por (Skinner, 2003, Seção 13.2)

$$Y_{it} = x_{it}\beta + u_i + v_{it}, \text{ com } i = 1, \dots, N \text{ e } t = 1, \dots, T. \quad (2.6)$$

Em (2.6),  $u_i$  são os efeitos aleatórios constantes (ou fatores específicos não observáveis) e  $v_{it}$  são efeitos aleatórios transitórios. Perceba que  $u_i$  não apresenta índice para o tempo, ou seja, é uma característica própria do indivíduo, já  $v_{it}$  apresenta índice  $t$ , ou seja, é uma característica do indivíduo que pode variar ao longo do tempo. Conforme mencionado anteriormente, consideramos apenas interceptos aleatórios, ou seja, nosso modelo permite que os indivíduos possuam interceptos diferentes, as quais são caracterizadas pelo efeito  $u_i$ . Desta forma, caso a variância do intercepto seja alta, estamos dizendo, que os indivíduos são muito diferentes entre si.

Para exemplificar esses dois conceitos, podemos pensar em uma situação bem simples. Vamos supor um indivíduo muito caridoso, porém pobre, mas que contribui quando pode, fazendo uma doação para Instituições de Caridade. O valor da beneficência é baixo, mas o indivíduo tem a iniciativa e propensão de ajudar o próximo (Efeito  $u_i$ ). Porém ao receber uma herança inesperada, o mesmo passa então a ajudar de forma significativa as entidades filantrópicas (efeito  $v_{it}$ ).

Assumimos que as variáveis aleatórias  $u_i$  e  $v_{it}$  são independentes com

$$E(u_i) = E(v_{it}) = 0,$$

$$VAR(u_i) = \sigma_u^2,$$

$$VAR(v_{it}) = \sigma_v^2,$$

$$COV(v_{it}, v_{it'}) = 0 \text{ (para todo } t \neq t'),$$

$$COV(u_i, v_{it}) = 0,$$

onde  $VAR(.)$  denota a variância populacional.

A esperança de  $Y_{it}$  deve ser representada por:

$$E(Y_{it}) = \underline{x}_{it} \underline{\beta} = \underline{\mu}(x_{it}, \underline{\beta}), \text{ com } t = 1, \dots, T. \quad (2.7)$$

A partir de (2.6) e (2.7), podemos escrever a variância de  $Y_{it}$  e a covariância entre  $Y_{it}$  e  $Y_{it'}$  respectivamente como

$$\begin{aligned} \sigma^2 &= \text{VAR}(Y_{it}) = E\{\varepsilon_{it}^2\} = E[u_i + v_{it}]^2 = E[u_i^2 + 2u_i v_{it} + v_{it}^2] = \\ &\sigma_u^2 + 2E(u_i \cdot v_{it}) + \sigma_v^2 = \sigma_u^2 + \sigma_v^2, \end{aligned} \quad (2.8)$$

e

$$\begin{aligned} \text{COV}(Y_{it}, Y_{it'}) &= E\{[u_i + v_{it}] \cdot [u_i + v_{it'}]\} = \\ &= E\{[u_i]^2\} + E\{u_i \cdot v_{it}\} + E\{u_i \cdot v_{it'}\} + E\{v_{it} \cdot v_{it'}\} = E\{[u_i]^2\} = \sigma_u^2. \end{aligned} \quad (2.9)$$

Logo, a estrutura de covariância do modelo apresenta  $\sigma_u^2 + \sigma_v^2$  na diagonal principal e  $\sigma_u^2$  fora da diagonal na matriz  $\Sigma$  ( $T \times T$ ), ou seja,

$$\Sigma = \begin{bmatrix} \sigma_u^2 + \sigma_v^2 & \sigma_u^2 & \dots & \sigma_u^2 \\ \sigma_u^2 & \sigma_u^2 + \sigma_v^2 & & \sigma_u^2 \\ \vdots & & \ddots & \vdots \\ \sigma_u^2 & \sigma_u^2 & \dots & \sigma_u^2 + \sigma_v^2 \end{bmatrix}.$$

Sendo assim, a variância de  $Y$  é modelada como duas componentes separadas: (i) variância entre indivíduos,  $\sigma_u^2$ ; e a (ii) variância para o mesmo indivíduo,  $\sigma_v^2$ . Além disso, a partir de uma definição geral de correlação e das expressões (2.8) e (2.9) temos que

$$\text{CORR}(Y_{it}, Y_{it'}) = \rho = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2}, \quad (2.10)$$

onde  $\rho$  denota a correlação intra-indivíduo populacional, e  $0 \leq \rho < 1$ . Desta forma, (i)  $\rho = 0$  quando  $\sigma_u^2 = 0$ , significando que não há variância entre as unidades (o que permitiria o uso de regressão linear); e (ii)  $\rho$  é “próximo” de 1 quando  $\sigma_v^2$  está perto de 0, significando que há quase nenhuma variação no mesmo indivíduo ao longo do tempo e nenhuma medida de erro.

## 2.1.2 - Modelo AR1

Um segundo modelo considerado é o modelo autoregressivo de ordem 1 (AR1). Em tal modelo pressupõe-se que as correlações decaem ao longo do tempo, por isso é também chamado de modelo de correlação exponencial sendo mais complexo que o modelo de correlação uniforme. Os efeitos aleatórios transitórios  $v_{it}$  nesse caso são gerados por um processo auto regressivo de primeira ordem, sendo dados por

$$v_{it} = \gamma v_{it-1} + \varepsilon_{it}, \text{ com } t=1, \dots, T. \quad (2.11)$$

Podemos perceber que se  $\gamma = 0$  em (2.11), então,  $v_{it} = \varepsilon_{it}$  e o modelo AR1 se reduz ao modelo de correlação uniforme descrito em 2.1.1. O parâmetro  $\gamma$  deve pertencer ao intervalo de  $(-1,1)$ , ou seja,  $|\gamma| < 1$  e  $\varepsilon_{it}$  são os resíduos com

$$E(\varepsilon_{it}) = 0,$$

e

$$\text{VAR}(\varepsilon_{it}) = \sigma_{\varepsilon}^2.$$

De acordo com Vieira (2009) se trabalhamos com o pressuposto de que  $\text{VAR}(v_{it}) = \sigma_v^2$  e que  $v_{it}$  e  $\varepsilon_{it}$  são mutuamente independentes temos

$$\sigma_v^2 = \frac{\sigma_{\varepsilon}^2}{1 - \gamma^2}. \quad (2.12)$$

Além disso, a variância de  $Y_{it}$  e a covariância entre  $Y_{it}$  e  $Y_{it'}$  são dadas respectivamente por:

$$\sigma^2 = \text{VAR}(Y_{it}) = \text{VAR}(\underline{x}_{it} \underline{\beta} + u_i + v_{it}) = \sigma_u^2 + \sigma_v^2 = \sigma_u^2 + \frac{\sigma_{\varepsilon}^2}{1 - \gamma^2}, \quad (2.13)$$

e

$$\text{COV}(Y_{it}, Y_{it'}) = E\{\underline{u}_i \underline{u}_i'\} + E\{v_{it} \cdot v_{it'}\} = \sigma_u^2 + \gamma^{|t-t'|} \sigma_v^2. \quad (2.14)$$

Assim, o modelo também apresenta  $\sigma_u^2 + \sigma_v^2$  na diagonal principal, mas elementos diferentes daqueles do MCU fora da diagonal na matriz  $\Sigma$ , ou seja,

$$\Sigma = \begin{bmatrix} \sigma_u^2 + \sigma_v^2 & \sigma_u^2 + \gamma^{|t-t'|}\sigma_v^2 & \dots & \sigma_u^2 + \gamma^{|t-t'|}\sigma_v^2 \\ \sigma_u^2 + \gamma^{|t-t'|}\sigma_v^2 & \sigma_u^2 + \sigma_v^2 & & \sigma_u^2 + \gamma^{|t-t'|}\sigma_v^2 \\ \vdots & & \ddots & \vdots \\ \sigma_u^2 + \gamma^{|t-t'|}\sigma_v^2 & \sigma_u^2 + \gamma^{|t-t'|}\sigma_v^2 & \dots & \sigma_u^2 + \sigma_v^2 \end{bmatrix}.$$

O modelo auto-regressivo pressupõe então que a covariância intra-indivíduos diminui quando o intervalo de tempo entre as observações aumenta, já que especifica que a variável resposta depende linearmente de suas observações anteriores, podendo resultar em parâmetros com estimativas mais eficientes.

## 3 - Modelos Longitudinais de Estrutura de Covariância

### 3.1- Introdução

Jöreskog (1972), Anderson (1973) e Wiley, Schimidt e Bramble (1973) foram os pioneiros na discussão sobre a adoção de modelos para a estrutura da covariância para estimação dos componentes de variância. Mais recentemente, Pourahmadi (1999), Pourahmadi (2000) e Pan e Mackenzie (2003), dentre outros autores, discutiram a modelagem da estrutura de covariância com aplicação em dados longitudinais no caso clássico, onde se pressupõe que as observações são independentes e identicamente distribuídas (IID).

O foco deste capítulo é o estudo de procedimentos de inferência para modelos de efeitos aleatórios para dados longitudinais, primeiramente desconsiderando o plano amostral complexo e depois considerando o mesmo na estimação dos parâmetros.

Nossos parâmetros de interesse são

$\sigma_u^2$  = Variabilidade entre indivíduos (fatores individuais específicos e não observáveis),

$\sigma_v^2$  = Variabilidade intra-indivíduos (fatores aleatórios transitórios) e

$\gamma$  = Correlação temporal.

Neste momento é importante definirmos  $\underline{\theta}$  como nosso vetor paramétrico  $b \times 1$  de interesse. Para o MCU temos  $\underline{\theta} = (\sigma_u^2, \sigma_v^2)$  e para o modelo AR1 temos  $\underline{\theta} = (\sigma_u^2, \sigma_v^2, \gamma)$ , por exemplo. Nesta monografia, são comparadas e estudadas as características dos métodos de estimação de  $\underline{\theta} = (\sigma_u^2, \sigma_v^2)$  para o MCU, considerando o plano amostral complexo e também o desconsiderando. O que geralmente é feito com frequência é estimar os coeficientes ( $\underline{\beta}$ ) e suas respectivas variâncias levando em consideração alguns aspectos do plano amostral, mas ao se estimar  $\underline{\theta}$  considera-se que a amostra foi selecionada de forma aleatória simples. Diversos *software* trabalham dessa forma, mas nosso objetivo é levar em consideração o plano amostral para estimar tal parâmetro. Em geral, ao ignorar o plano amostral complexo e modelar dados como IID podemos

verificar diferenças nas estimativas pontuais, nas estimativas de variância e nas distribuições de estatísticas de teste (Vieira, 2009).

Os métodos de estimação considerados nesta monografia para  $\underline{\theta}$  são a (i) máxima verossimilhança e (ii) mínimos quadrados generalizados para o caso clássico e a (iii) máxima pseudo verossimilhança e (iv) mínimos quadrados generalizados para o caso em que se leva em consideração o plano amostral complexo.

Mesmo que a estimação de  $\underline{\beta}$  não seja o nosso principal foco de atenção nesta monografia, este assunto é considerado na próxima seção com objetivo de permitir uma melhor contextualização da questão do ajuste de modelos de regressão para dados longitudinais amostrais complexos.

### 3.2 - Estimação de $\underline{\beta}$ – Máxima Verossimilhança

Sejam  $\underline{y}_i$  vetores  $T \times 1$  com os valores observados para a variável resposta para cada indivíduo  $i$  em cada ocasião. Consideramos nesta monografia a abordagem de superpopulação (Skinner, Holt e Smith, 1989). Além disso, trabalhamos com o pressuposto de que a população seja finita e de que  $Y_1, \dots, Y_N$  são independentes. Com o produto das densidades normais define-se a função densidade conjunta do censo ou função de verossimilhança do censo, sendo dada por (Vieira, 2009)

$$\begin{aligned} f(\underline{y}_1, \dots, \underline{y}_N; \underline{\beta}) &= \ell_N[\underline{\beta}] = \prod_{i=1}^N \left\{ \frac{1}{(2\pi)^{T/2} |\Sigma|^{1/2}} e^{-[\underline{y}_i - \underline{\mu}(\underline{\beta})] \Sigma^{-1} [\underline{y}_i - \underline{\mu}(\underline{\beta})] / 2} \right\} \\ &= (2\pi)^{-N_o/2} |\Sigma|^{-N/2} e^{-\sum_{i=1}^N [\underline{y}_i - \underline{\mu}(\underline{\beta})] \Sigma^{-1} [\underline{y}_i - \underline{\mu}(\underline{\beta})] / 2}. \end{aligned} \quad (3.1)$$

O vetor de parâmetros  $\underline{\beta}$  é estimado maximizando o logaritmo da verossimilhança do censo com respeito à  $\underline{\beta}$ , ou seja,

$$\max (\log \ell_N[\underline{\beta}]). \quad (3.2)$$

Assim, o estimador de máxima verossimilhança (EMV) do censo  $\hat{\underline{\beta}}_N$  de  $\underline{\beta}$  é obtido minimizando o expoente com respeito à  $\underline{\beta}$ , ou seja,

$$\min \left( \sum_{i=1}^N [y_i - \underline{\mu}_i(\underline{\beta})]' \Sigma^{-1} [y_i - \underline{\mu}_i(\underline{\beta})] \right). \quad (3.3)$$

Note que, de acordo com a abordagem de superpopulação de Skinner, Holt e Smith (1989),  $\hat{\underline{\beta}}_N$  pode ser considerado um pseudoparâmetro enquanto  $\underline{\beta}$  é um superparâmetro ou um parâmetro da superpopulação.

A expressão acima é a soma das distâncias quadráticas multivariadas generalizadas de  $\underline{Y}_i$  a  $\underline{\mu}_i(\underline{\beta})$ . Supomos que apenas  $y_i$  foi observada para as unidades  $i$  selecionadas para a amostra  $s$ ,  $\{1, \dots, n\}$ . Logo, (3.3) pode ser estimado como

$$\frac{N}{n} \cdot \sum_{i=1}^n [y_i - \underline{\mu}_i(\underline{\beta})]' \Sigma^{-1} [y_i - \underline{\mu}_i(\underline{\beta})]. \quad (3.4)$$

Note que o estimador de máxima verossimilhança de  $\underline{\beta}$  será obtido minimizando (3.3). Alternativamente podemos obter o EMV resolvendo o sistema de equações conhecido como equações de pseudo-escore para  $\underline{\beta}$ , dadas por

$$\sum_{i=1}^n [y_i - \underline{\mu}_i(\underline{\beta})]' \Sigma^{-1} \frac{\partial \underline{\mu}_i(\underline{\beta})}{\partial \underline{\beta}} = 0. \quad (3.5)$$

Seja  $\mathbf{X}_i = (\underline{x}_{i1}', \dots, \underline{x}_{iT}')$  a matriz  $T \times q$  com as covariáveis para o indivíduo  $i$ . Quando  $\underline{\mu}_i(\underline{\beta}) = \mathbf{X}_i \underline{\beta}$ , temos

$$\frac{\partial \underline{\mu}_i(\underline{\beta})}{\partial \underline{\beta}} = \mathbf{X}_i.$$

Portanto, as equações pseudo-escore têm uma solução definida para  $\hat{\underline{\beta}}(\underline{\Sigma})$  dada por

$$\hat{\underline{\beta}}(\underline{\Sigma}) = \left( \sum_{i=1}^n \mathbf{X}_i' \underline{\Sigma}^{-1} \mathbf{X}_i \right)^{-1} \sum_{i=1}^n \mathbf{X}_i' \underline{\Sigma}^{-1} y_i. \quad (3.6)$$

Em geral,  $\underline{\Sigma}$  é desconhecida e sua estimação será discutida de forma breve na Seção 3.4.

Um caso particular de  $\hat{\underline{\beta}}(\underline{V})$  ocorre para  $\underline{\Sigma}$  definida, por exemplo, como a matriz identidade  $T \times T$ , ou seja, pressupondo que as respostas ao longo do tempo de um determinado indivíduo são independentes. Neste caso,

$$\hat{\underline{\beta}}(I) = \left( \sum_{i=1}^n \mathbf{X}'_i \mathbf{X}_i \right)^{-1} \sum_{i=1}^n \mathbf{X}'_i \underline{y}_i, \quad (3.7)$$

que é equivalente ao estimador de Mínimos Quadrados Ordinários (MQO).

### 3.3 - Estimação de $\underline{\beta}$ – Máxima Pseudo-Verosimilhança

Uma das alternativas adotadas para se levar em consideração o plano amostral complexo na estimação dos coeficientes de modelos é através da utilização de pesos amostrais que podem variar de indivíduo para indivíduo. Inicialmente vamos considerar as mesmas premissas do caso clássico. Assim, o estimador será obtido por

$$\min \sum_{i=1}^n w_i \left[ \underline{y}_i - \underline{\mu}_i(\underline{\beta}) \right]' \Sigma^{-1} \left[ \underline{y}_i - \underline{\mu}_i(\underline{\beta}) \right], \quad (3.8)$$

onde  $w_i$  são os pesos amostrais. De forma análoga ao caso da máxima verossimilhança, as equações de pseudo score têm uma solução para  $\hat{\underline{\beta}}(\Sigma)_{MPV}$  dada por (Vieira, 2009)

$$\hat{\underline{\beta}}(\Sigma)_{PMV} = \left( \sum_{i=1}^n w_i \mathbf{X}'_i \Sigma^{-1} \mathbf{X}_i \right)^{-1} \sum_{i=1}^n w_i \mathbf{X}'_i \Sigma^{-1} \underline{y}_i. \quad (3.9)$$

Se substituirmos  $\Sigma$  por uma matriz de correlação de trabalho  $T \times T$  para  $y_i$ , denotada por  $V$ , o estimador passa a ser dado por

$$\hat{\underline{\beta}}(V)_{PMV} = \left( \sum_{i=1}^n w_i \mathbf{X}'_i V^{-1} \mathbf{X}_i \right)^{-1} \sum_{i=1}^n w_i \mathbf{X}'_i V^{-1} \underline{y}_i. \quad (3.10)$$

Uma alternativa adotada com frequência para a definição de  $V$  considera que a mesma resulta do modelo MCU para o comportamento da variável resposta.

Quando o peso amostral for o mesmo para toda a amostra o estimador de MPV torna-se o estimador para o caso clássico (EMV)

$$\hat{\underline{\beta}}(V)_{PMV} = \left( \sum_{i=1}^n \mathbf{X}'_i V^{-1} \mathbf{X}_i \right)^{-1} \sum_{i=1}^n \mathbf{X}'_i V^{-1} \underline{y}_i. \quad (3.11)$$

E se  $V=I$  então o estimador de  $\beta$  no caso clássico é equivalente ao estimador de Mínimos Quadrados Ordinários (MQO).

### 3.4 - Estimação de $\Sigma$

Nesta seção consideramos a estimação da matriz de covariância para  $\underline{Y}_i$ , ou seja, estimação de  $\Sigma$ , que é um parâmetro da superpopulação ou um superparâmetro. Inicialmente, definimos  $S_N$  (pseudoparâmetro) sendo a matriz de covariância

$$S_N = \begin{pmatrix} \text{var}_N(\underline{Y}_{i1}) & & & \\ \text{cov}_N(\underline{Y}_{i2}, \underline{Y}_{i1}) & \text{var}_N(\underline{Y}_{i2}) & & \\ \vdots & & \ddots & \\ \text{cov}_N(\underline{Y}_{iT}, \underline{Y}_{i1}) & \text{cov}_N(\underline{Y}_{iT}, \underline{Y}_{i2}) & \cdots & \text{var}_N(\underline{Y}_{iT}) \end{pmatrix},$$

onde o subscrito  $N$  indica o tamanho da população finita. De acordo com Vieira (2009) para o caso clássico temos

$$\text{var}_N(\underline{Y}_{it}) = (N - q)^{-1} \sum_{i=1}^N (y_{it} - \underline{x}_{it} \hat{\underline{\beta}}_N)^2, \quad (3.12)$$

e

$$\text{cov}_N(\underline{Y}_t, \underline{Y}_{t'}) = (N - q)^{-1} \sum_{i=1}^N (y_{it} - \underline{x}_{it} \hat{\underline{\beta}}_N)(y_{it'} - \underline{x}_{it'} \hat{\underline{\beta}}_N). \quad (3.13)$$

Temos então que a matriz  $S_N$  é o estimador natural de  $\Sigma$  (ver Rencher, 1998; e Johnson e Wichern, 1998). Para simplificarmos, definimos  $N^* = (N - q)$ , assim

$$E(S_N) = E \left[ \begin{pmatrix} \text{var}_N(\underline{Y}_{i1}) & & & \\ \text{cov}_N(\underline{Y}_{i2}, \underline{Y}_{i1}) & \text{var}_N(\underline{Y}_{i2}) & & \\ \vdots & & \ddots & \\ \text{cov}_N(\underline{Y}_{iT}, \underline{Y}_{i1}) & \text{cov}_N(\underline{Y}_{iT}, \underline{Y}_{i2}) & \cdots & \text{var}_N(\underline{Y}_{iT}) \end{pmatrix} \right] =$$

$$= E \left[ \begin{pmatrix} (N^*)^{-1} \sum_{i=1}^N (y_{i1} - \underline{x}_{i1} \hat{\underline{\beta}}_N)^2 & & & \\ (N^*)^{-1} \sum_{i=1}^N (y_{i2} - \underline{x}_{i2} \hat{\underline{\beta}}_N)(y_{i1} - \underline{x}_{i1} \hat{\underline{\beta}}_N) & (N^*)^{-1} \sum_{i=1}^N (y_{i2} - \underline{x}_{i2} \hat{\underline{\beta}}_N)^2 & & \\ \vdots & & \ddots & \\ (N^*)^{-1} \sum_{i=1}^N (y_{iT} - \underline{x}_{iT} \hat{\underline{\beta}}_N)(y_{i1} - \underline{x}_{i1} \hat{\underline{\beta}}_N) & (N^*)^{-1} \sum_{i=1}^N (y_{iT} - \underline{x}_{iT} \hat{\underline{\beta}}_N)(y_{i2} - \underline{x}_{i2} \hat{\underline{\beta}}_N) & \cdots & (N^*)^{-1} \sum_{i=1}^N (y_{iT} - \underline{x}_{iT} \hat{\underline{\beta}}_N)^2 \end{pmatrix} \right] =$$

$$= \begin{pmatrix} \text{VAR}(\underline{Y}_{i1}) & & & \\ \text{COV}(\underline{Y}_{i2}, \underline{Y}_{i1}) & \text{VAR}(\underline{Y}_{i2}) & & \\ \vdots & & \ddots & \\ \text{COV}(\underline{Y}_{iT}, \underline{Y}_{i1}) & \text{COV}(\underline{Y}_{iT}, \underline{Y}_{i2}) & \cdots & \text{VAR}(\underline{Y}_{iT}) \end{pmatrix} = \Sigma. \quad (3.14)$$

Em seguida podemos supor que a amostra  $s$  tenha sido selecionada levando em consideração um desenho amostral complexo. Nesse caso, a matriz  $\Sigma$  é estimada através de uma matriz de covariância de resíduos ponderados pelos pesos amostrais, assim como feito em Vieira (2009). Temos então  $S_w$  como estimador de  $\Sigma$  sendo dada por

$$S_w = [S_w]_{it'} = \begin{pmatrix} \text{var}_w(\underline{Y}_{i1}) & & & \\ \text{cov}_w(\underline{Y}_{i2}, \underline{Y}_{i1}) & \text{var}_w(\underline{Y}_{i2}) & & \\ \vdots & & \ddots & \\ \text{cov}_w(\underline{Y}_{iT}, \underline{Y}_{i1}) & \text{cov}_w(\underline{Y}_{iT}, \underline{Y}_{i2}) & \cdots & \text{var}_w(\underline{Y}_{iT}) \end{pmatrix}$$

onde

$$\text{var}_w(\underline{Y}_{it}) = \hat{N}^{-1} \sum_{i=1}^n w_i (y_{it} - \hat{\mu}_{it})^2, \quad (3.15)$$

e

$$\text{cov}_w(\underline{Y}_{it}, \underline{Y}_{it'}) = \hat{N}^{-1} \sum_{i=1}^n w_i (y_{it} - \hat{\mu}_{it})(y_{it'} - \hat{\mu}_{it'}), \quad (3.16)$$

onde  $\hat{N} = \sum_{i=1}^n w_i$  é o estimador do tamanho da população  $N$  e para  $n$  grande temos  $\frac{\hat{N}}{N} \doteq 1$ .

Seja  $E_p(\cdot)$  o valor esperado com respeito à distribuição de uma estatística considerando todas as amostras  $s$  possíveis de serem selecionadas. Além disso, seja  $z_i$  uma variável qualquer, que não depende da escolha da amostra  $s$ . Assim, podemos considerar o seguinte resultado geral (Isaki e Fuller, 1982; Vieira, 2009)

$$E_p \left( \sum_{i=1}^n w_i z_i \right) = E_p \left( \sum_{i=1}^n \pi_i^{-1} z_i \right) = \sum_{i=1}^N z_i.$$

A partir deste resultado temos,

$$E_p(\mathbf{S}_w) = E_p \left( \begin{array}{cccc} \hat{N}^{-1} \sum_{i=1}^n w_i (y_{i1} - \hat{\mu}_{i1})^2 & & & \\ \hat{N}^{-1} \sum_{i=1}^n w_i (y_{i2} - \hat{\mu}_{i2})(y_{i1} - \hat{\mu}_{i1}) & \hat{N}^{-1} \sum_{i=1}^n w_i (y_{i2} - \hat{\mu}_{i2})^2 & & \\ \vdots & \vdots & \ddots & \\ \hat{N}^{-1} \sum_{i=1}^n w_i (y_{iT} - \hat{\mu}_{iT})(y_{i1} - \hat{\mu}_{i1}) & \hat{N}^{-1} \sum_{i=1}^n w_i (y_{iT} - \hat{\mu}_{iT})(y_{i2} - \hat{\mu}_{i2}) & \cdots & \hat{N}^{-1} \sum_{i=1}^n w_i (y_{iT} - \hat{\mu}_{iT})^2 \end{array} \right) \doteq$$

$$\doteq \left( \begin{array}{cccc} N^{-1} E_p \left[ \sum_{i=1}^n w_i (y_{i1} - \hat{\mu}_{i1})^2 \right] & & & \\ N^{-1} E_p \left[ \sum_{i=1}^n w_i (y_{i2} - \hat{\mu}_{i2})(y_{i1} - \hat{\mu}_{i1}) \right] & N^{-1} E_p \left[ \sum_{i=1}^n w_i (y_{i2} - \hat{\mu}_{i2})^2 \right] & & \\ \vdots & \vdots & \ddots & \\ N^{-1} E_p \left[ \sum_{i=1}^n w_i (y_{iT} - \hat{\mu}_{iT})(y_{i1} - \hat{\mu}_{i1}) \right] & N^{-1} E_p \left[ \sum_{i=1}^n w_i (y_{iT} - \hat{\mu}_{iT})(y_{i2} - \hat{\mu}_{i2}) \right] & \cdots & N^{-1} E_p \left[ \sum_{i=1}^n w_i (y_{iT} - \hat{\mu}_{iT})^2 \right] \end{array} \right) \doteq$$

Então,

$$E_p(\mathbf{S}_w) \doteq \frac{(N-q)}{N} \mathbf{S}_N \doteq \mathbf{S}_N$$

e pressupondo  $\hat{\mu}_{it}$  como um estimador consistente de  $\hat{\mu}_{Nit}$  temos

$$E(E_p(\mathbf{S}_w)) \doteq E(\mathbf{S}_N) = \Sigma(\underline{\theta}). \quad (3.17)$$

Para maiores detalhes, ver Vieira (2009).

Logo,  $S_w$  é então um estimador consistente de  $\Sigma$  em relação ao modelo e ao plano amostral. Quando estivermos trabalhando com o caso clássico os pesos amostrais são constantes. Neste caso,  $S_w$  se reduz para  $S$ , que é o estimador da matriz de covariância que pressupõe amostragem aleatória simples.

### 3.5 - Estimação de $\underline{\theta}$

Na seção anterior trabalhamos com a estrutura de covariância dos resíduos ao longo do tempo. A matriz de covariância pode ser definida de acordo com uma estrutura paramétrica definida a partir do modelo considerado para  $y_{it}$ , ou seja,  $\Sigma = \Sigma(\underline{\theta})$ .

Sejam  $F(S_w, \Sigma)$  e  $F(S, \Sigma)$  funções de ajuste (ou funções de discrepância) que a partir de  $\hat{\underline{\theta}}$  minimizam a distancia entre  $S_w$  e  $\Sigma$ , e entre  $S$  e  $\Sigma$ , respectivamente. Como  $S_w$

e  $S$  são estimadores de  $\Sigma$ , procuramos estimadores ótimos que minimizem tais distâncias. Para o caso de  $S_w$ , por exemplo, a função de ajuste é dada por

$$F(S_w, \Sigma(\hat{\underline{\theta}})) = \min_{\underline{\theta} \in \Theta} F(S_w, \Sigma(\underline{\theta})). \quad (3.18)$$

Se o ajuste for perfeito, então  $F(\underline{\theta}) = 0$ , assim quanto menor for o valor de  $F(\underline{\theta})$  melhor será o ajustamento dos dados ao modelo teórico.

De acordo com Browne (1982, 1984) a função de ajuste apresenta certas propriedades como: (i)  $F(\underline{\theta})$  é escalar; (ii)  $F(\underline{\theta}) \geq 0$ ; (iii)  $F(S_w, \Sigma(\underline{\theta})) = 0$  se e somente se  $S_w = \Sigma$ ; e (iv)  $F(\underline{\theta})$  é uma função contínua diferenciável de  $S_w$  e  $\Sigma$ .

### 3.5.1 - Máxima Verossimilhança

Mais uma vez, consideramos o método de Máxima Verossimilhança (MV) agora para o caso do estimador de  $\underline{\theta}$  que tem como base uma função de ajuste. Estimamos  $\underline{\theta}$  a partir do valor que maximiza a verossimilhança da amostra. Sendo assim, o estimador de MV de  $\underline{\theta}$  é função dos dados amostrais  $Y_1, \dots, Y_n$  que torna máxima a função  $l(y; \underline{\theta})$  em  $\underline{\theta}$ . No caso clássico, a inferência é baseada no modelo  $f(y; \underline{\theta})$ , os dados recebem pesos iguais e consequentemente o plano amostral é ignorado. Com o pressuposto de  $Y_i$  na forma multivariada normal, a função de log verossimilhança com respeito à  $\underline{\beta}$  e  $\underline{\theta}$  será dada por (Vieira, 2009)

$$L[\underline{\beta}, \underline{\theta}] = \log \ell[\underline{\beta}, \underline{\theta}] = -\frac{1}{2} \left[ n_o \log 2\pi + n \log |\Sigma(\underline{\theta})| + \sum_{i=1}^n [y_i - \mu_i(\underline{\beta})] \Sigma(\underline{\theta})^{-1} [y_i - \mu_i(\underline{\beta})] \right]. \quad (3.19)$$

Para maximizar a função acima com respeito à  $\underline{\theta}$  e  $\underline{\beta}$  precisamos minimizar

$$\begin{aligned} & n \cdot \log |\Sigma(\underline{\theta})| + \sum_{i=1}^n [y_i - \mu_i(\underline{\beta})] \Sigma(\underline{\theta})^{-1} [y_i - \mu_i(\underline{\beta})] = \\ & = \text{tr} \left[ \Sigma(\underline{\theta})^{-1} \sum_{i=1}^n (y_i - \mu_i)(y_i - \mu_i)' \right] + n \cdot \log |\Sigma(\underline{\theta})| = \\ & = n \cdot \text{tr} \left[ \Sigma(\underline{\theta})^{-1} n^{-1} \sum_{i=1}^n (y_i - \mu_i)(y_i - \mu_i)' \right] + n \cdot \log |\Sigma(\underline{\theta})|. \end{aligned} \quad (3.20)$$

Na igualdade acima o traço aparece devido a seguinte propriedade

$$a' \times U \times b = tr(U \times b \times a'),$$

onde  $a$  e  $b$  são dois vetores quaisquer e  $U$  é uma matriz (respeitando a dimensão de multiplicação).

A partir da definição da matriz  $S$  na seção 3.4, podemos reescrever (3.20) como

$$n \cdot \{tr[S\Sigma(\underline{\theta})^{-1}] + \log|\Sigma(\underline{\theta})|\}. \quad (3.21)$$

As matrizes  $S$  e  $\Sigma(\underline{\theta})$  devem ser positivas definidas, e  $|\Sigma(\underline{\theta})|$  denota o determinante de  $\Sigma(\underline{\theta})$  (esse determinante advém da função de densidade normal multivariada).

De acordo com Jöreskog (1970), Wiley, Schimidt e Bramble (1973) a partir da equação (3.21) a função de ajuste de máxima verossimilhança é dada por

$$F(\underline{\theta})_{MV} = \log|\Sigma(\underline{\theta})| + tr[S\Sigma(\underline{\theta})^{-1}]. \quad (3.22)$$

Nesse momento vamos incluir o termo  $-\log|S| - T$  na equação acima, pois assim permitimos que a função de ajuste seja igual a zero quando  $\Sigma(\hat{\underline{\theta}}) = S$ . Portanto, a função de ajuste é dada por

$$F(\underline{\theta})_{MV} = \log|\Sigma(\underline{\theta})| + tr[S\Sigma(\underline{\theta})^{-1}] - \log|S| - T,$$

ou de forma análoga (3.23)

$$F(\underline{\theta})_{MV} = tr[S\Sigma(\underline{\theta})^{-1}] - \log|S\Sigma(\underline{\theta})^{-1}| - T.$$

Esse termo que foi adicionado não depende de  $\beta$  e  $\underline{\theta}$  por isso não altera o valor dos mesmos  $\beta$  e  $\underline{\theta}$  na função de minimização. Ao incluirmos esse termo garantimos que  $F(\underline{\theta})_{MV} = 0$  quando  $\Sigma(\hat{\underline{\theta}}) = S$ , porque  $S$  e  $\Sigma(\hat{\underline{\theta}})$  teriam o mesmo determinante e  $S\Sigma(\underline{\theta})^{-1} = I$ . Assim o traço de  $I$  tem valor igual a  $T$ .

A partir da definição da função de ajuste podemos finalmente minimizá-la. A função de ajuste é minimizada quando  $\hat{\underline{\theta}}_{MV}$  é obtido pela diferenciação de  $F(\underline{\theta})_{MV}$  com

respeito a cada um dos parâmetros (Wiley, Schmidt e Bramble, 1973). Mas, obtemos  $\hat{\underline{\theta}}_{MV}$  de uma forma alternativa, a partir da solução de um sistema de equações

$$\frac{\partial F(\underline{\theta})_{ML}}{\partial \theta_b} = tr \left\{ \Sigma(\underline{\theta})^{-1} \cdot [\Sigma(\underline{\theta}) - S] \cdot \Sigma(\underline{\theta})^{-1} \cdot \frac{\partial \Sigma(\underline{\theta})}{\partial \theta_b} \right\} = 0,$$

onde  $\underline{\theta}_b = (\sigma_u^2, \sigma_v^2)$ .

Duas estratégias poderiam ser utilizadas no processo de minimização de  $F(\underline{\theta})_{MV}$ . A primeira opção seria estimar  $\underline{\beta}$  separadamente adotando a equação 3.6 e então minimizar a função de ajuste com respeito apenas a  $\underline{\theta}$ . Uma segunda alternativa a ser adotada seria minimizar simultaneamente a função com respeito à  $\underline{\beta}$  e a  $\underline{\theta}$ . Adotamos o primeiro método como forma de minimização de  $F(\underline{\theta})_{MV}$  assim como foi feito em Vieira (2009). Além disso, os estimadores de  $\sigma_u^2$  e  $\sigma_v^2$  são obtidos nesta monografia por soluções numéricas a partir do método de Newton-Raphson uma vez que as soluções analíticas não podem ser encontradas.

Para conhecimentos adicionais nesta área ver Pourahmadi (1999), Pourahmadi (2000), e Pan e Mackenzie (2003) que utilizam a estimação por máxima verossimilhança no contexto de modelos de estruturas de covariância com dados longitudinais.

### 3.5.2 - Máxima Pseudo Verossimilhança

Nesta seção apresentamos o estimador de Máxima Pseudo Verossimilhança (MPV) para obter o vetor  $\hat{\underline{\theta}}$  que minimize a função de ajuste. Este método é empregado para que possamos considerar o plano amostral complexo na obtenção das estimativas dos parâmetros do modelo.

O vetor de parâmetros  $\underline{\theta}$  será estimado maximizando a função de pseudo log verossimilhança, isto é, minimizando

$$\begin{aligned}
& \sum_{i=1}^n w_i \cdot \log|\Sigma(\underline{\theta})| + tr \left[ \sum_{i=1}^n w_i [y_i - \mu_i(\underline{\beta})][y_i - \mu_i(\underline{\beta})] \Sigma(\underline{\theta})^{-1} \right] = \\
& \sum_{i=1}^n w_i \cdot \log|\Sigma(\underline{\theta})| + \sum_{i=1}^n w_i \cdot tr \left[ \left( \sum_{i=1}^n w_i \right)^{-1} \cdot \sum_{i=1}^n w_i [y_i - \mu_i(\underline{\beta})][y_i - \mu_i(\underline{\beta})] \Sigma(\underline{\theta})^{-1} \right] = \\
& \sum_{i=1}^n w_i \cdot \left\{ \log|\Sigma(\underline{\theta})| + tr \left[ \left( \sum_{i=1}^n w_i \right)^{-1} \cdot \sum_{i=1}^n w_i [y_i - \mu_i(\underline{\beta})][y_i - \mu_i(\underline{\beta})] \Sigma(\underline{\theta})^{-1} \right] \right\}.
\end{aligned} \tag{3.24}$$

Da mesma forma que fizemos na seção anterior iremos adicionar o termo  $-\log|S_w| - T$ . Portanto, a função de ajuste  $F(\underline{\theta})_{PML}$  é dada por

$$F(\underline{\theta})_{PML} = tr[S_w \Sigma(\underline{\theta})^{-1}] + \log|\Sigma(\underline{\theta})| - \log|S_w| - T,$$

Ou de forma análoga (3.25)

$$F(\underline{\theta})_{PML} = tr[S_w \Sigma(\underline{\theta})^{-1}] - \log|S_w \Sigma(\underline{\theta})^{-1}| - T.$$

Repare que  $F(\underline{\theta})_{PML}$  reduz-se para  $F(\underline{\theta})_{ML}$  quando os pesos  $w_i$  são constantes.

Novamente, os estimadores de  $\sigma_u^2$ ,  $\sigma_v^2$  e  $\gamma$  são obtidos nesta monografia por soluções numéricas através do método de Newton-Raphson.

### 3.5.3 - Mínimos Quadrados Generalizados

O terceiro método utilizado é o de Mínimos Quadrados Generalizados (MQG). Esse método, proposto por Jöreskog e Goldberger (1972) no contexto dos modelos para estrutura de covariâncias, estima os parâmetros ponderando pelos resíduos. Desta forma, são adotados pesos que correspondem ao inverso da matriz de covariância dos resíduos. Os elementos da matriz  $\Sigma$  que tem maior variância amostral apresentam menor peso no modelo. Desta forma, obtêm-se estimadores mais precisos do que os de Mínimos Quadrados Não Ponderados de Bentler e Weeks (1980).

Consideramos a seguir duas formas deste método de estimação. A primeira não leva em consideração o desenho amostral (mínimos quadrados generalizados não ponderados) e possui função de ajuste dada por

$$F(\underline{\theta})_{MQG} = \{vech[S] - vech[\Sigma(\underline{\theta})]\}' \hat{C}^{-1} \{vech[S] - vech[\Sigma(\underline{\theta})]\}. \quad (3.26)$$

onde  $vech[S]$  e  $vech[\Sigma(\underline{\theta})]$  são vetores  $k \times 1$  que incluem os elementos não duplicados das respectivas matrizes. Assim temos

$$vech[S] = [S_{11}, S_{21}, \dots, S_{T1}, S_{22}, \dots, S_{T2}, \dots, S_{TT}],$$

e

$$(3.27)$$

$$vech[\Sigma(\underline{\theta})] = [\Sigma_{11}, \Sigma_{21}, \dots, \Sigma_{T1}, \Sigma_{22}, \dots, \Sigma_{T2}, \dots, \Sigma_{TT}].$$

Além disso,  $\hat{C}$  é uma matriz  $k \times k$  positiva definida que estima a matriz de covariância da matriz  $E$ ,  $T \times T$ , sem levar em consideração o desenho amostral. Note que  $E = vech[S] - vech[\Sigma(\underline{\theta})]$ . Para maiores detalhes, ver Vieira (2009).

O segundo estimador leva em consideração o desenho amostral (mínimos quadrados generalizados ponderados) sendo ponderado pelos pesos  $w_i$ . Nesse caso, a função de ajuste a minimizar é dada por

$$F(\underline{\theta})_{MQGC} = \{vech[S_w] - vech[\Sigma(\underline{\theta})]\}' \hat{C}_c^{-1} \{vech[S_w] - vech[\Sigma(\underline{\theta})]\}. \quad (3.28)$$

onde  $\hat{C}_c$  é a matriz de covariância da matriz  $E$ , proposto por Vieira (2009) e que leva em consideração o desenho amostral e que não depende de pressupostos sobre a distribuição da variável de interesse (*asymptotically distribution free* - ADF).

Entretanto, de acordo com Bollen (1989, p. 432), Satorra (1992), Muthén e Satorra (1995), Satorra (2000), e Satorra e Bentler (2001) métodos do tipo ADF devem ser utilizados com cautela no contexto de modelos para estrutura de covariâncias já que exigem tamanhos da amostra grandes para se tornarem estáveis e estatisticamente robustos à violações aos pressupostos sobre a distribuição.

## 4 - Estudo de Simulação

### 4.1 - Introdução

Neste capítulo são apresentados as características e os resultados de um estudo de simulação que tem como objetivo principal avaliar as propriedades estatísticas dos procedimentos de estimação pontual discutidos no capítulo anterior. São considerados o viés, o coeficiente de variação e o erro quadrático médio (EQM).

O estudo envolve simular  $d = 1, \dots, D$  replicações amostrais. O número de replicações consideradas é o mesmo para todos os cenários ( $D = 1000$ ). Os valores das covariáveis  $x$  são fixos e os valores dos escores  $Y_{it}$  são simulados a partir do conjunto de dados reais da BHPS com 1340 indivíduos considerando diferentes modelos, de forma independente para cada replicação. No decorrer deste capítulo, o subscrito *sim* é utilizado para denotar simulação.

Para avaliar o efeito do tamanho da amostra repetimos a simulação para diferentes tamanhos de amostra,  $n_{sim} = 100$ ,  $n_{sim} = 500$ ,  $n_{sim} = 10000$ ,  $n_{sim} = 15000$ , e  $n_{sim} = 20000$ . Os dois primeiros resultados são obtidos por amostragem aleatória simples sem reposição enquanto os três últimos com reposição.

Inicialmente adotamos o modelo de correlação uniforme (MCU),

$$Y_{it} = \underline{x}_{it} \underline{\beta} + u_i + v_{it}, \quad (4.1)$$

conforme definido na seção 2.1 para gerar os valores simulados de  $Y_{it}$ . Além disso, para os outros parâmetros do estudo temos inicialmente

$$u_i \sim N(0, \sigma_u^2),$$

e

$$v_{it} \sim N(0, \sigma_v^2).$$

O segundo modelo considerado leva em consideração os efeitos do plano amostral complexo, mais especificamente os impactos da conglomeração. Esse modelo é chamado de MCU-C sendo representado por

$$Y_{ijt} = \underline{x}_{ijt} \underline{\beta} + \eta_j + u_{ij} + v_{ijt}. \quad (4.2)$$

Além disso, temos inicialmente

$$\eta_j \sim N(0, \sigma_\eta^2),$$

$$u_{ij} \sim N(0, \sigma_u^2),$$

e

$$v_{ijt} \sim N(0, \sigma_v^2),$$

onde  $Y_{ijt}$  é o escore para o indivíduo  $i=1, \dots, n_{sim}$ , no conglomerado  $j=1, \dots, m_{sim}$ , na ocasião  $t$  da pesquisa. Temos também que  $\eta_j$  são os efeitos aleatórios de conglomeração, com  $m_{sim}$  denotando o número de conglomerados considerados em cada simulação e  $\underline{\beta}$ ,  $u_i$  e  $v_{it}$  definidos anteriormente.

Antes de obter as estimativas dos parâmetros de interesse, faremos o ajuste do modelo de acordo com as expressões (4.1) e (4.2). O parâmetro simulado  $\underline{\beta}$  é um vetor  $q \times 1$ , onde  $q$  é o número de covariáveis fixas e é estimado por (ver expressão 3.6)

$$\hat{\underline{\beta}}_{sim}(\mathbf{V}) = \left( \sum_{i=1}^{1340} \mathbf{X}'_i \mathbf{V}^{-1} \mathbf{X}_i \right)^{-1} \sum_{i=1}^{1340} \mathbf{X}'_i \mathbf{V}^{-1} \underline{y}_i, \quad (4.3)$$

onde  $\underline{y}_i$  é um vetor  $T \times 1$  com os valores observados das variáveis respostas para cada indivíduo  $i$  incluído no banco de dados da BHPS,  $\mathbf{X}_i$  são matrizes  $T \times q$  com valores das  $q$  covariáveis para cada indivíduo  $i$  em cada ocasião, e  $\mathbf{V}$  é uma matriz de covariância permutável de ordem  $T \times T$ .

As variáveis adotadas são as mesmas definidas na Seção 1.2. Para a variável resposta temos o escore de atitude do papel do gênero e as covariáveis são: (i) número de ocasiões, (ii) faixa etária (com quatro categorias), (iii) atividade econômica (com cinco categorias) e (iv) nível educacional (com cinco categorias).

Para o modelo MCU os coeficientes estimados foram:

$$\underline{\beta} = \underline{\hat{\beta}}_{sim} = \begin{bmatrix} 22.18 \\ -0.04 \\ -0.70 \\ -0.87 \\ -1.00 \\ -0.91 \\ -0.73 \\ 0.17 \\ -2.05 \\ -0.51 \\ -0.60 \\ -0.43 \\ -1.17 \end{bmatrix}. \quad (4.4)$$

Além disso, consideramos que

$$\sigma_u^2 = \sigma_{u,sim}^2 \cong 6.744,$$

$$\sigma_v^2 = \sigma_{v,sim}^2 \cong 4.965,$$

$$\rho = \rho_{sim} \cong 0.576.$$

Os valores acima são calculados pelo método Swamy-Arora, como em Vieira (2009), e obtidos pelo *software* Stata.

Analogamente para o MCU-C temos:

$$\underline{\beta} = \underline{\hat{\beta}}_{sim,C} = \begin{bmatrix} 22.16 \\ -0.04 \\ -0.70 \\ -0.87 \\ -1.00 \\ -0.91 \\ -0.73 \\ 0.17 \\ -2.05 \\ -0.51 \\ -0.60 \\ -0.42 \\ -1.17 \end{bmatrix}, \quad (4.5)$$

Além disso, consideramos que

$$\begin{aligned}
\sigma_{\eta}^2 &= \sigma_{\eta \text{ sim,C}}^2 \cong 0.088, \\
\sigma_u^2 &= \sigma_{u \text{ sim,C}}^2 \cong 7.135, \\
\sigma_v^2 &= \sigma_{v \text{ sim,C}}^2 \cong 4.981, \\
\rho &= \rho_{\text{sim,C}} \cong 0.589.
\end{aligned}
\tag{4.6}$$

Esses valores foram calculados por Vieira (2009), aplicando o método do MQG descritos por Goldstein (1995) e Rasbash *et al.* (2002).

Neste exercício de simulação, para os mesmos modelos apresentados acima consideramos adicionalmente resíduos com as seguintes distribuições diferentes da Normal: (i) a densidade t Student com  $\nu = 4$  graus de liberdade; (ii) Normal assimétrica com parâmetro de assimetria  $\lambda = 3$ ; e (iii) t assimétrica também com  $\nu = 4$  graus de liberdade e  $\lambda = 3$ . Esse valor foi adotado, pois caracteriza uma considerável assimetria aos resíduos gerados.

Para a geração dos valores da variável de interesse considerando os modelos e distribuições descritos acima, consideramos metodologia adotada por Vieira (2009). Para todos os cenários os modelos ajustados foram sempre do tipo MCU.

## 4.2 - Resultados

Utilizamos os mesmos critérios adotados por Vieira (2009) para avaliar o comportamento dos estimadores. O EQM é calculado como

$$\text{EQM}(\hat{\theta}_j) = \text{VAR}(\hat{\theta}_j) + [\text{VIÉS}(\hat{\theta}_j)]^2,
\tag{4.7}$$

onde  $\hat{\theta}_j$  denota os  $b$  elementos incluídos no vetor de parâmetros  $\underline{\theta}$ ,  $\text{VAR}(\hat{\theta}_j)$  é a variância verdadeira de  $\hat{\theta}_j$ ,  $\text{VIÉS}(\hat{\theta}_j)$  denota o viés verdadeiro do estimador sendo estimado por

$$\text{Viés}(\hat{\theta}_j) = \hat{E}(\hat{\theta}_j) - \theta_j,
\tag{4.8}$$

onde  $\hat{E}(\hat{\theta}_j)$  é o valor esperado do estimador de  $\hat{\theta}_j$ .

Além disso, calculamos o viés relativo de  $\hat{\theta}_j$ , tal que

$$\text{Viés relativo } (\hat{\theta}_j) = \frac{\text{Viés}(\hat{\theta}_j)}{\theta_j} \cdot 100. \quad (4.9)$$

Em cada replicação calculamos  $\hat{\theta}_j^{(d)}$ , para cada um dos métodos de estimação estudados. Ou seja,

$$\hat{E}(\hat{\theta}_j) = \frac{1}{D} \sum_{d=1}^D \hat{\theta}_j^{(d)} \quad (4.10)$$

Temos também a variância do estimador sendo dada por

$$\text{var}(\hat{\theta}_j) = \frac{1}{D-1} \sum_{d=1}^D [\hat{\theta}_j^{(d)} - \hat{E}(\hat{\theta}_j)]^2 \quad (4.11)$$

onde  $\text{var}(\hat{\theta}_j)$  é o estimador de  $\text{VAR}(\hat{\theta}_j)$ .

O erro quadrático médio é estimado por

$$\text{Eqm}(\hat{\theta}_j) = \text{var}(\hat{\theta}_j) + [\text{viés}(\hat{\theta}_j)]^2 \quad (4.12)$$

e o coeficiente de variação de  $\hat{\theta}_j$  é estimado por

$$\text{cv}[\hat{\theta}_j] = \frac{\sqrt{\text{var}(\hat{\theta}_j)}}{\hat{E}(\hat{\theta}_j)} \cdot 100 \quad (4.13)$$

Para a implementação e obtenção dos resultados deste estudo de simulação utilizamos o *software* R (versão 2.15.3 para o sistema operacional Windows), que é um ambiente de programação distribuído gratuitamente sob a *General Public License* e está disponível em <http://www.r-project.org/>. Na geração dos dados assimétricos utilizamos especificamente o pacote *sn* do R (Azzalini e Capitanio, 2014). Além disso, para a aplicação do método de minimização de Newton-Raphson utilizamos o pacote do R *nlm* (R Core Team, 2013). Para maiores informações sobre distribuições assimétricas ver Ferreira, Bolfarine e Lachos (2013).

O computador utilizado nesta monografia possui as seguintes especificações: (i) *Intel Core 3 i5*, (ii) 8 *Gigabytes (GB)* de memória *ram*, (iii) 1 *Terabyte (TB)* de armazenamento no *Hard Disk (HD)*. Grande parte das simulações foram inicializadas

por meio de conexões remotas. Dessa forma, mesmo não estando presente no laboratório foi possível acessar o computador em qualquer local por meio de outro, apenas sendo necessário conexão a internet.

As próximas seções apresentam os resultados do estudo de simulação em formato de tabelas. Cada tabela apresenta resultados produzidos para cada um dos quatro estimadores estudados nesta monografia, sendo que os dois primeiros são não ponderados enquanto os dois últimos são estimadores ponderados pelos pesos amostrais longitudinais da BHPS. Desta forma os seguintes estimadores são considerados,

- (i) mínimos quadrados generalizados não ponderados ( $\hat{\theta}_{MQG}$ , ver expressão 3.26),
- (ii) máxima verossimilhança ( $\hat{\theta}_{MV}$ , ver expressão 3.23),
- (iii) mínimos quadrados generalizados ponderados ( $\hat{\theta}_{MQGC}$ , ver expressão 3.28) e
- (iv) máxima pseudo verossimilhança ( $\hat{\theta}_{MPV}$ , ver expressão 3.25).

Note que (i) e (iii) são métodos ADF.

#### 4.2.1 - Resíduos com Distribuição Normal

Os resultados apresentados nessa subseção consideram que os resíduos seguem a distribuição normal. Os modelos considerados para gerar os dados das replicações são o MCU (Tabelas 5.1 a 5.5) e o MCU-C (Tabelas 5.6 a 5.10).

Parâmetro		Viés Relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-11.52 %	20.84%	2414.551
	$\hat{\sigma}_v^2$	-10.03%	9.54%	613.226
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-10.06%	15.32%	1612.963
	$\hat{\sigma}_v^2$	1.10%	6.96%	125.939
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-13.43%	18.48%	2768.062
	$\hat{\sigma}_v^2$	-8.47%	9.10%	361.047
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-4.67%	17.52%	1527.071
	$\hat{\sigma}_v^2$	-2.67%	7.63%	154.887

Tabela 5.1 – Estimação de  $\hat{\theta}$  para resíduos com distribuição normal (replicações geradas pelo modelo MCU) e  $n_{sim} = 100$ .

Ao observar a Tabela 5.1 pode-se notar que, em geral, há diferenças ao compararmos o EQM dos métodos ADF em relação aos demais métodos de estimação. Essas diferenças se justificam uma vez que para o presente cenário os resíduos são normais. Por outro lado, de acordo com Satorra (1992) e Yuan e Bentler (1997) deve-se utilizar métodos ADF com bastante cuidado em situações em que o tamanho da amostra é menor que 5000, para modelos complexos como no caso da presente monografia. Além disso, podemos concluir que a variância é a principal componente que contribui para o EQM, quando comparada ao viés.

Além disso, de acordo com os resultados da monografia, métodos ponderados parecem ter viés ligeiramente inferior quando comparados aos não ponderados. Isto deve tornar-se mais claro nas subseções posteriores, quando o tamanho da amostra é maior. No entanto, não se deve supor que isso seja sempre verdade, pois não há qualquer razão teórica clara para este resultado dado que o método de seleção foi amostragem aleatória simples para todas as replicações.

Parâmetro		Viés Relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-2.75%	8.04%	332.051
	$\hat{\sigma}_v^2$	-3.09%	3.50%	53.430
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-2.02%	7.24%	261.871
	$\hat{\sigma}_v^2$	-0.21%	3.09%	24.693
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-2.27%	7.74%	347.921
	$\hat{\sigma}_v^2$	-2.21%	3.31%	35.274
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-1.36%	7.56%	250.732
	$\hat{\sigma}_v^2$	-0.72%	3.27%	27.501

Tabela 5.2 – Estimação de  $\hat{\theta}$  para resíduos com distribuição normal (replicações geradas pelo modelo MCU) e  $n_{sim} = 500$ .

A partir dos resultados incluídos na Tabela 5.2, podemos fazer observações muito semelhantes às que são feitas para os resultados apresentados na Tabela 5.1, em termos de comparação entre os métodos. Mais uma vez, o estimador MQGC teve o maior EQM entre os métodos para  $\sigma_u^2$  e o estimador MQG o maior EQM para  $\sigma_v^2$ ,

embora em termos de c.v. sejam muito semelhantes aos outros métodos. Os resultados da Tabela 5.2 indicam diminuição de viés, c.v. e EQM quando comparados com os da Tabela 5.1, provocados por um aumento do tamanho da amostra.

Parâmetro		Viés Relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-1.00%	1.74%	17.257
	$\hat{\sigma}_v^2$	-0.34%	0.75%	2.032
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-0.98%	1.63%	16.899
	$\hat{\sigma}_v^2$	-0.21%	0.70%	1.473
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-1.02%	1.64%	17.310
	$\hat{\sigma}_v^2$	-0.28%	0.71%	1.677
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-0.93%	1.74%	16.743
	$\hat{\sigma}_v^2$	-0.24%	0.75%	1.734

Tabela 5.3 – Estimação de  $\hat{\theta}$  para resíduos com distribuição normal (replicações geradas pelo modelo MCU) e  $n_{sim} = 10000$ .

Pela Tabela 5.3 percebemos grandes diferenças em relação às tabelas 5.1 e 5.2 decorrente do incremento do tamanho de amostra. Os valores de viés, c.v. e EQM diminuíram substancialmente. Praticamente não há diferenças entre os estimadores indicando que para  $n$  grande os estimadores se comportam de forma muito semelhante. Além disso, o c.v. é praticamente igual para os estimadores, havendo então uma redução da dispersão em relação à média.

Parâmetro		Viés Relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-0.93%	1.39%	16.256
	$\hat{\sigma}_v^2$	-0.32%	0.60%	1.422
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-0.93%	1.29%	16.081
	$\hat{\sigma}_v^2$	-0.22%	0.56%	1.052
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-0.96%	1.28%	16.345
	$\hat{\sigma}_v^2$	-0.28%	0.55%	1.185
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-0.89%	1.38%	15.931
	$\hat{\sigma}_v^2$	-0.25%	0.60%	1.237

Tabela 5.4 – Estimação de  $\hat{\theta}$  para resíduos com distribuição normal (replicações geradas pelo modelo MCU) e  $n_{sim} = 15000$ .

A Tabela 5.4 apresenta resultados muito similares à tabela 5.3. Porém, em termos do coeficiente de variação, o método MQGC passa a apresentar menores valores do que os demais métodos. Os valores de viés são muito semelhantes entre si e há pequenas diferenças entre a eficiência dos estimadores, com diminuição dos valores de c.v. e EQM. Praticamente não há diferenças entre os estimadores indicando que há uma tendência para n grande dos estimadores se comportarem de forma muito semelhante e parece haver também uma tendência das estimativas convergirem para um dado valor.

Parâmetro		Viés Relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-0.87%	1.28%	16.198
	$\hat{\sigma}_v^2$	-0.23%	0.52%	1.071
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-0.84%	1.18%	16.051
	$\hat{\sigma}_v^2$	-0.15%	0.48%	0.833
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-0.93%	1.18%	16.021
	$\hat{\sigma}_v^2$	-0.22%	0.48%	0.902
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-0.82%	1.27%	15.899
	$\hat{\sigma}_v^2$	-0.22%	0.52%	0.944

Tabela 5.5 – Estimação de  $\hat{\theta}$  para resíduos com distribuição normal (replicações geradas pelo modelo MCU) e  $n_{sim} = 20000$ .

A Tabela 5.5 apresenta resultados semelhantes à tabela 5.3 e 5.4. Novamente, o c.v. do método MQGC é um dos menores obtidos juntamente com o MV. Os valores de viés diminuem, porém continuam muito semelhantes entre si e há pequenas diferenças entre a eficiência dos estimadores. Com esse último aumento no tamanho da amostra parece haver uma tendência dos métodos ADF em melhorarem sua eficiência de forma mais evidente quando comparados aos demais métodos. Como há uma necessidade de tamanhos grandes de amostras para o método livre de distribuição parece que conforme esse número aumenta esse método melhora cada vez mais sua eficiência.

As Tabelas 5.6 a 5.10 incluem resultados para o ajuste do modelo MCU a partir de dados gerados pelo modelo MCU-C.

Parâmetro		Viés Relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-12.53%	20.68%	2465.472
	$\hat{\sigma}_v^2$	-10.09%	9.78%	617.351
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-10.08%	16.60%	1652.289
	$\hat{\sigma}_v^2$	2.21%	7.45%	156.308
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-16.91%	19.34%	2770.621
	$\hat{\sigma}_v^2$	- 8.52%	9.48%	367.1422
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-4.09%	17.56%	1530.605
	$\hat{\sigma}_v^2$	-1.74%	7.62%	146.763

Tabela 5.6 – Estimação de  $\hat{\theta}$  para resíduos com distribuição normal (replicações geradas pelo modelo MCU-C) e  $n_{sim} = 100$ .

Conclusões semelhantes podem ser utilizadas quando se comparam os resultados da Tabela 5.6 com aqueles da Tabela 5.1. Podemos destacar aqui o desempenho do método MPV na estimação dos parâmetros, com viés muito menor do que os métodos restantes. Além disso, podemos notar que a maioria dos métodos teve uma diminuição em seu EQM em comparação com o caso que desconsidera o desenho amostral, principalmente devido a uma diminuição modesta do viés. Novamente os métodos ADF são piores do que os demais já que voltamos a considerar um pequeno tamanho de amostra. Em relação ao c.v. todos os métodos apresentam valores elevados embora os métodos ADF apresentem resultados mais elevados que os demais.

Parâmetro		Viés Relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-2.80%	8.14%	335.962
	$\hat{\sigma}_v^2$	-3.10%	3.56%	53.434
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-2.47%	7.26%	267.197
	$\hat{\sigma}_v^2$	0.37%	3.16%	25.349
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-2.90%	7.75%	298.134
	$\hat{\sigma}_v^2$	-1.79%	3.33%	35.114
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	0.83%	7.76%	313.532
	$\hat{\sigma}_v^2$	-0.47%	3.36%	27.841

Tabela 5.7 – Estimação de  $\hat{\theta}$  para resíduos com distribuição normal (replicações geradas pelo modelo MCU-C) e  $n_{sim} = 500$ .

Podemos fazer observações muito semelhantes às que são feitas para os resultados apresentados na Tabela 5.6. O estimador MQG teve o maior EQM entre os métodos para  $\sigma_u^2$  e o estimador MQGC o maior EQM para  $\sigma_v^2$ , embora em termos de viés sejam ligeiramente superiores aos outros métodos. Os resultados da Tabela 5.7 indicam diminuição tanto de viés quanto de variância, quando comparados com os da Tabela 5.6, provocados por um aumento do tamanho da amostra. Houve redução do c.v. embora os métodos ADF possuam os maiores valores.

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	0.78%	1.70%	21.474
	$\hat{\sigma}_v^2$	-0.15%	0.73%	1.372
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	0.81%	1.61%	20.386
	$\hat{\sigma}_v^2$	0.03%	0.67%	1.138
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	0.78%	1.61%	19.560
	$\hat{\sigma}_v^2$	-0.07%	0.67%	1.143
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	0.85%	1.69%	22.566
	$\hat{\sigma}_v^2$	-0.01%	0.72%	1.315

Tabela 5.8 – Estimação de  $\hat{\theta}$  para resíduos com distribuição normal (replicações geradas pelo modelo MCU-C) e  $n_{sim} = 10000$ .

Pela Tabela 5.8 percebemos grandes diferenças em relação às Tabelas 5.6 e 5.7 decorrentes do incremento do tamanho de amostra. Os valores de viés, c.v. e EQM diminuíram, embora isso tenha ocorrido de forma expressiva no c.v. e na eficiência. Praticamente não há diferenças entre os estimadores indicando que para  $n$  grande os estimadores se comportam de forma muito semelhante. Além disso, o c.v. é praticamente igual para os estimadores, havendo então uma redução da dispersão em relação ao valor esperado. Porém, diferentemente da Tabela 5.3, o método MQGC aparece com o menor EQM para  $\sigma_u^2$  e com valor muito semelhante a MV para a estimativa de  $\sigma_v^2$ .

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	0.75%	1.44%	16.889
	$\hat{\sigma}_v^2$	-0.07%	0.61%	0.9517
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	0.74%	1.39%	15.830
	$\hat{\sigma}_v^2$	0.03%	0.58%	0.858
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	0.70%	1.39%	15.300
	$\hat{\sigma}_v^2$	-0.03%	0.58%	0.855
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	0.79%	1.44%	17.640
	$\hat{\sigma}_v^2$	0.01%	0.61%	0.927

Tabela 5.9 – Estimação de  $\hat{\theta}$  para resíduos com distribuição normal (replicações geradas pelo modelo MCU-C) e  $n_{sim} = 15000$ .

A Tabela 5.9 traz resultados muito semelhantes à Tabela 5.8. Os valores de viés, c.v. e EQM diminuíram, embora isso não tenha ocorrido de forma tão perceptível. Praticamente não há diferenças entre os estimadores indicando que para  $n$  grande os estimadores se comportam de forma muito semelhante. Além disso, o c.v. e o EQM são praticamente iguais para os estimadores, havendo então uma ligeira vantagem do MQGC sobre os demais na estimação dos parâmetros.

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	0.69%	1.29%	15.668
	$\hat{\sigma}_v^2$	-0.10%	0.54%	0.7552
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	0.70%	1.21%	14.698
	$\hat{\sigma}_v^2$	-0.001%	0.51%	0.6587
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	0.67%	1.21%	14.250
	$\hat{\sigma}_v^2$	-0.05%	0.51%	0.6584
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	0.72%	1.29%	16.235
	$\hat{\sigma}_v^2$	-0.03%	0.54%	0.7314

Tabela 5.10 – Estimação de  $\hat{\theta}$  para resíduos com distribuição normal (replicações geradas pelo modelo MCU-C) e  $n_{sim} = 20000$ .

Conclusões semelhantes podem ser utilizadas quando se compara a Tabela 5.10 com a Tabela 5.9. Novamente os valores de viés, c.v. e EQM diminuiram decorrente do aumento do tamanho da amostra e o método MQGC apresenta melhor eficiência na estimação dos parâmetros.

#### 4.2.2- Resíduos com Distribuição Normal Assimétrica

Os resultados apresentados nessa subseção dizem respeito às simulações com os resíduos seguindo a distribuição de probabilidade normal assimétrica com parâmetro de assimetria  $\lambda = 3$  e os modelos utilizados são o MCU (Tabelas 5.11 a 5.15) e o MCU-C (Tabelas 5.16 a 5.20).

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-31.01%	57.98%	11435.352
	$\hat{\sigma}_v^2$	-31.81%	10.39%	5101.655
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-31.20%	18.81%	10543.932
	$\hat{\sigma}_v^2$	-28.99%	7.56%	4014.731
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-32.14%	23.31%	11486.872
	$\hat{\sigma}_v^2$	-30.21%	9.62%	4776.655
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-30.91%	19.67%	9871.854
	$\hat{\sigma}_v^2$	-29.07%	8.15%	4245.025

Tabela 5.11 – Estimação de  $\hat{\theta}$  para resíduos com distribuição normal assimétrica (replicações geradas pelo modelo MCU) e  $n_{sim} = 100$ .

Ao observar cada linha da Tabela 5.11 notam-se grandes diferenças em relação à Tabela 5.1. O fato dos dados serem gerados de forma assimétrica inflacionou de forma expressiva o viés e a variância dos estimadores. Há certa diferença ao comparar o EQM dos métodos ADF em relação aos demais métodos de estimação. Essas diferenças são perceptíveis e esperadas já que é esperado tamanho de amostra maior para que os métodos ADF apresentem melhores resultados. Além disso, podemos observar que o impacto do viés é maior quando comparado com o c.v., ou seja, para os dados assimétricos o viés é o principal componente que contribui para o EQM.

Parâmetro		Viés Relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-30.44%	8.90%	9465.365
	$\hat{\sigma}_v^2$	-29.50%	3.99%	4321.681
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-30.24%	8.08%	9338.452
	$\hat{\sigma}_v^2$	-28.66%	3.44%	4292.247
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-30.62%	8.42%	9574.505
	$\hat{\sigma}_v^2$	-30.22%	3.72%	4240.097
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-29.99%	8.50%	9190.636
	$\hat{\sigma}_v^2$	-28.83%	3.67%	4128.928

Tabela 5.12 – Estimação de  $\hat{\theta}$  para resíduos com distribuição normal assimétrica (replicações geradas pelo modelo MCU) e  $n_{sim} = 500$ .

A partir dos resultados incluídos na Tabela 5.12, podemos fazer observações muito semelhantes às que são feitas para os resultados apresentados na Tabela 5.11, em termos de comparação entre os métodos. O estimador MQGC teve o maior EQM entre os métodos para  $\sigma_u^2$  e o estimador MQG o maior EQM para  $\sigma_v^2$ , embora em termos de c.v. sejam muito semelhantes aos outros métodos. Os resultados da Tabela 5.12 indicam uma pequena diminuição de viés e EQM enquanto o c.v. diminui de forma mais expressiva quando comparados com os da Tabela 5.11, provocados por um aumento no número das observações. E novamente o viés é o principal componente que contribui para o EQM.

Parâmetro		Viés Relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-29.86%	1.85%	9080.332
	$\hat{\sigma}_v^2$	-28.76%	0.82%	4105.075
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-29.85%	1.72%	9074.775
	$\hat{\sigma}_v^2$	-28.71%	0.77%	4091.637
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-29.87%	1.72%	9077.453
	$\hat{\sigma}_v^2$	-28.74%	0.77%	4100.001
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-29.83%	1.84%	9063.757
	$\hat{\sigma}_v^2$	-28.72%	0.82%	4094.455

Tabela 5.13 – Estimação de  $\hat{\theta}$  para resíduos com distribuição normal assimétrica (replicações geradas pelo modelo MCU) e  $n_{sim} = 10000$ .

A partir dos resultados incluídos na Tabela 5.13, podemos fazer observações muito semelhantes às que são feitas para os resultados apresentados na Tabela 5.11 e na Tabela 5.12, em termos de comparação entre os métodos. Não há uma grande diferença entre o viés dos métodos na Tabela 5.13 quando comparado às Tabelas 5.12 e 5.11. Porém há uma expressiva diminuição do c.v. . Essa diminuição mostra a semelhança entre o MQGC e o MV no que diz respeito ao comportamento. Em termos do EQM os estimadores ADF apresentam os maiores valores embora possuam ganhos de eficiência superiores aos outros com o incremento do tamanho da amostra.

Parâmetro		Viés Relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-29.82%	1.49%	9055.825
	$\hat{\sigma}_v^2$	-28.74%	0.68%	4099.578
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-29.81%	1.40%	9051.753
	$\hat{\sigma}_v^2$	-28.71%	0.64%	4090.885
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-29.83%	1.40%	9053.785
	$\hat{\sigma}_v^2$	-28.73%	0.65%	4096.492
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-29.80%	1.50%	9048.431
	$\hat{\sigma}_v^2$	-28.71%	0.68%	4092.345

Tabela 5.14 – Estimação de  $\hat{\theta}$  para resíduos com distribuição normal assimétrica (replicações geradas pelo modelo MCU) e  $n_{sim} = 15000$ .

Os resultados incluídos na Tabela 5.14 são muito semelhantes aos apresentados na tabela 5.13 em termos de comparação entre os métodos. O viés praticamente se manteve constante com o c.v., diminuindo de forma muito pequena. Os EQMs dos métodos foram muito similares, mas podemos perceber que os métodos ADF se aproximaram dos outros, praticamente deixando de haver diferenças entre eles. Parece haver uma tendência dos métodos à convergência. Porém os métodos ADF aparecem estar diminuindo a diferença em relação aos demais.

Parâmetro		Viés Relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-29.81%	1.33%	9067.81
	$\hat{\sigma}_v^2$	-28.74%	0.59%	4100.91
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-29.80%	1.25%	9062.78
	$\hat{\sigma}_v^2$	-28.72%	0.54%	4095.00
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-29.81%	1.25%	9064.38
	$\hat{\sigma}_v^2$	-28.73%	0.54%	4095.21
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-29.83%	1.33%	9063.55
	$\hat{\sigma}_v^2$	-28.72%	0.59%	4095.47

Tabela 5.15 – Estimação de  $\hat{\theta}$  para resíduos com distribuição normal assimétrica (replicações geradas pelo modelo MCU) e  $n_{sim} = 20000$ .

Conclusões semelhantes podem ser utilizadas quando se compara a Tabela 5.15 com as Tabelas 5.14 e 5.13. Porém, diferente das outras o viés parece ter estabilizado com inclusive pequenos aumentos. Além disso, o estimador MQGC apresentou um valor muito próximo de EQM do estimador de MV para  $\sigma_u^2$  e  $\sigma_v^2$ . Isso confirma a necessidade de um grande tamanho de amostra para que os métodos ADF possam superar em termos de eficiência os outros, já que seus coeficientes de variação já são relativamente menores.

As Tabelas 5.16 a 5.20 incluem resultados para o ajuste do modelo MCU a partir de dados gerados pelo modelo MCU-C com os resíduos seguindo a distribuição normal assimétrica.

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-31.17%	23.09%	10089.13
	$\hat{\sigma}_v^2$	-31.89%	10.44%	10131.24
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-30.71%	18.47%	9733.64
	$\hat{\sigma}_v^2$	-28.22%	7.85%	3968.17
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-32.16%	23.43%	10719.02
	$\hat{\sigma}_v^2$	-30.77%	9.75%	9431.69
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-29.39%	19.15%	8958.12
	$\hat{\sigma}_v^2$	-29.13%	8.56%	4227.465

Tabela 5.16 – Estimação de  $\hat{\theta}$  para resíduos com distribuição normal assimétrica (replicações geradas pelo modelo MCU-C) e  $n_{sim} = 100$ .

A Tabela 5.16 inclui os resultados para o modelo MCU que é ajustado aos dados simulados de  $Y_{it}$  gerados pelo modelo MCU-C. Conclusões razoavelmente semelhantes podem ser utilizadas quando se comparam os resultados da Tabela 5.16 para aqueles na Tabela 5.11. Além disso, podemos notar que a maioria dos métodos tiveram uma diminuição em seu EQM em comparação com o caso que desconsidera o desenho amostral. Novamente os métodos ADF apresentam um desempenho pior do que os demais já que temos um pequeno tamanho de amostra. Em relação ao c.v. todos os métodos apresentam valores elevados embora os métodos ADF apresentem resultados mais elevados que os demais.

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-29.02%	8.61%	8613.595
	$\hat{\sigma}_v^2$	-29.39%	3.85%	4289.91
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-28.82%	7.68%	8485.339
	$\hat{\sigma}_v^2$	-28.56%	3.44%	4051.163
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-29.20%	7.99%	8711.42
	$\hat{\sigma}_v^2$	-29.11%	3.65%	4210.37
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-28.58%	8.24%	8350.56
	$\hat{\sigma}_v^2$	-28.72%	3.67%	4097.95

Tabela 5.17 – Estimação de  $\hat{\theta}$  para resíduos com distribuição normal assimétrica (replicações geradas pelo modelo MCU-C) e  $n_{sim} = 500$ .

A partir dos resultados incluídos na Tabela 5.17, podemos fazer observações muito semelhantes às que são feitas para os resultados apresentados na Tabela 5.16, em termos de comparação entre os métodos. Mais uma vez, o estimador MQGC apresentou o maior EQM entre os métodos para  $\sigma_u^2$  e o estimador MQG o maior EQM para  $\sigma_v^2$ , embora em termos de c.v. sejam muito semelhantes aos outros métodos. Os resultados da Tabela 5.17 indicam uma pequena diminuição de viés, c.v. e EQM comparados com os da Tabela 5.16.

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{GLS}$	$\hat{\sigma}_u^2$	-28.42%	1.88%	8226.845
	$\hat{\sigma}_v^2$	-28.69%	0.84%	4086.221
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-28.41%	1.77%	8219.867
	$\hat{\sigma}_v^2$	-28.65%	0.77%	4073.164
$\hat{\theta}_{GLSC}$	$\hat{\sigma}_u^2$	-28.43%	1.78%	8227.231
	$\hat{\sigma}_v^2$	-28.67%	0.78%	4080.598
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-28.39%	1.87%	8210.551
	$\hat{\sigma}_v^2$	-28.65%	0.84%	4075.743

Tabela 5.18 – Estimação de  $\hat{\theta}$  para resíduos com distribuição normal assimétrica (replicações geradas pelo modelo MCU-C) e  $n_{sim} = 10000$ .

A partir dos resultados incluídos na Tabela 5.18, percebemos uma grande diminuição do c.v. em comparação com a Tabela 5.17, porém o viés e o EQM não diminuíram de forma expressiva. Podemos fazer observações muito semelhantes às que são feitas para os resultados apresentados na Tabela 5.17, em termos de comparação entre os métodos. Apesar do estimador MQGC ainda apresentar uma maior EQM, o mesmo teve uma melhoria de eficiência maior do que os demais. Podemos perceber um indício novamente de que para grandes tamanhos de amostra os estimadores ADF possuem melhor eficiência do que os outros.

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{GLS}$	$\hat{\sigma}_u^2$	-28.41%	1.68%	8216.915
	$\hat{\sigma}_v^2$	-28.66%	0.70%	4078.353
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-28.45%	1.55%	8214.342
	$\hat{\sigma}_v^2$	-28.65%	0.65%	4069.866
$\hat{\theta}_{GLSC}$	$\hat{\sigma}_u^2$	-28.41%	1.55%	8214.788
	$\hat{\sigma}_v^2$	-28.65%	0.65%	4073.439
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-28.39%	1.68%	8210.310
	$\hat{\sigma}_v^2$	-28.64%	0.70%	4071.419

Tabela 5.19 – Estimação de  $\hat{\theta}$  para resíduos com distribuição normal assimétrica (replicações geradas pelo modelo MCU-C) e  $n_{sim} = 15000$ .

A partir dos resultados incluídos na Tabela 5.19, percebemos uma leve diminuição do c.v. e do EQM em comparação com a Tabela 5.18, porém o viés praticamente não se alterou. Parece que os estimadores chegaram próximos de uma convergência. Porém o EQM dos métodos ADF diminuiram mais do que os demais. Podemos fazer observações muito semelhantes às que são feitas para os resultados apresentados na Tabela 5.18, em termos de comparação entre os métodos. A diferença dos métodos ADF em relação aos outros vem diminuindo conforme o tamanho da amostra aumenta apesar do c.v. do método MQGC ser semelhante ao MV.

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-28.39%	1.36%	8212.255
	$\hat{\sigma}_v^2$	-28.67%	0.58%	4080.435
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-28.39%	1.29%	8209.350
	$\hat{\sigma}_v^2$	-28.64%	0.55%	4072.974
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-28.40%	1.29%	8207.815
	$\hat{\sigma}_v^2$	-28.66%	0.55%	4071.142
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-28.39%	1.36%	8207.298
	$\hat{\sigma}_v^2$	-28.65%	0.59%	4075.120

Tabela 5.20 – Estimação de  $\hat{\theta}$  para resíduos com distribuição normal assimétrica (replicações geradas pelo modelo MCU-C) e  $n_{sim} = 20000$ .

Com os resultados obtidos na Tabela 5.20, percebemos que parece haver uma convergência do viés dos estimadores e uma leve diminuição do c.v. e do EQM em comparação com a Tabela 5.19. A diferença dos métodos ADF em relação aos outros diminuiu ainda mais com o presente tamanho da amostra. Além disso, o c.v. do método MQGC é novamente semelhante ao MV. Inclusive o método da MPV diminuiu sua eficiência por conta de uma pequena inflação do viés e o método MQGC apresentou melhor eficiência para  $\sigma_v^2$  e um valor próximo a MPV para  $\sigma_u^2$ .

#### 4.2.3- Resíduos com Distribuição t

Os resultados apresentados nessa subseção dizem respeito às simulações com os resíduos seguindo a distribuição de probabilidade  $t_{\nu=4}(0,1)$  e os modelos utilizados são o MCU e o MCU-C.

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-25.95%	28.45%	5035.740
	$\hat{\sigma}_v^2$	-18.05%	14.66%	2229.370
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-17.15%	28.15%	3410.110
	$\hat{\sigma}_v^2$	-10.18%	15.54%	597.516
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-26.15%	28.47%	5650.340
	$\hat{\sigma}_v^2$	-19.91%	13.97%	1699.636
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-22.89%	29.07%	4796.464
	$\hat{\sigma}_v^2$	-13.62%	17.04%	702.241

Tabela 5.21 – Estimação de  $\hat{\theta}$  para resíduos com distribuição  $t_{\nu=4}(0,1)$  (replicações geradas pelo modelo MCU) e  $n_{sim} = 100$ .

A Tabela 5.21 inclui os resultados para o modelo MCU que é ajustado aos dados simulados de  $Y_{it}$  gerados pelo modelo MCU com resíduos distribuídos como uma  $t_{\nu=4}(0,1)$ . Podemos notar grandes diferenças entre os métodos. O MV e o MPV apresentaram os melhores resultados em termos de viés, c.v. e EQM, com grandes diferenças em relação a este último. Já os métodos ADF sofreram maiores impactos devido ao pequeno tamanho de amostra.

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-15.94%	12.76%	1725.07
	$\hat{\sigma}_v^2$	-10.67%	6.56%	368.28
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-13.78%	12.06%	1393.45
	$\hat{\sigma}_v^2$	7.24%	8.44%	178.12
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-16.48%	11.91%	1732.09
	$\hat{\sigma}_v^2$	-9.20%	6.23%	289.62
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-13.00%	12.66%	1357.37
	$\hat{\sigma}_v^2$	-8.54%	9.36%	215.91

Tabela 5.22 – Estimação de  $\hat{\theta}$  para resíduos com distribuição  $t_{\nu=4}(0,1)$  (replicações geradas pelo modelo MCU) e  $n_{sim} = 500$ .

Com os resultados obtidos na Tabela 5.22, percebemos a melhoria expressiva dos métodos ADF em termos viés, c.v. e EQM. Ambos os métodos tiveram grandes melhorias devido ao aumento da amostra, inclusive o método MQGC apresenta o menos c.v. em relação aos outros.

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-7.15%	2.89%	745.636
	$\hat{\sigma}_v^2$	-1.99%	1.68%	16.654
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-6.88%	2.87%	720.832
	$\hat{\sigma}_v^2$	-1.30%	5.46%	15.941
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-7.10%	2.77%	725.164
	$\hat{\sigma}_v^2$	-1.85%	1.56%	14.409
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-6.90%	2.94%	724.283
	$\hat{\sigma}_v^2$	-1.34%	4.63%	16.200

Tabela 5.23 – Estimação de  $\hat{\theta}$  para resíduos com distribuição  $t_{\nu=4}(0,1)$  (replicações geradas pelo modelo MCU) e  $n_{sim} = 10000$ .

Com os resultados obtidos na Tabela 5.23, e o aumento do tamanho da amostra percebemos que o método MQGC apresenta a melhor eficiência em termos de  $\sigma_v^2$ . Para  $\sigma_u^2$  o método MV apresenta o menor EQM. Porém os métodos ADF aparentam uma tendência de melhoria decorrente do tamanho da amostra. É importante notar também que os menores coeficientes de variação são encontrados no método MQGC.

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-5.84%	2.40%	703.830
	$\hat{\sigma}_v^2$	-1.71%	1.45%	12.303
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-5.72%	2.32%	685.150
	$\hat{\sigma}_v^2$	-0.43%	1.83%	8.755
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-5.86%	2.26%	686.030
	$\hat{\sigma}_v^2$	-1.55%	1.35%	8.443
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-5.67%	2.43%	679.910
	$\hat{\sigma}_v^2$	-0.47%	1.84%	8.937

Tabela 5.24 – Estimação de  $\hat{\theta}$  para resíduos com distribuição  $t_{\nu=4}(0,1)$  (replicações geradas pelo modelo MCU) e  $n_{sim} = 15000$ .

A partir dos resultados incluídos na Tabela 5.24, podemos fazer observações muito semelhantes às que são feitas para os resultados apresentados na Tabela 5.23, em termos de comparação entre os métodos. Embora os estimadores ADF tenham apresentando os maiores EQM's para  $\sigma_v^2$ , o método MQGC apresentou uma maior eficiência para  $\sigma_u^2$  em relação aos demais. Além disso, o estimador MQGC apresenta os menores valores de coeficiente de variação.

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-4.87%	2.07%	701.250
	$\hat{\sigma}_v^2$	-1.39%	1.23%	8.504
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-4.75%	1.97%	687.289
	$\hat{\sigma}_v^2$	-0.32%	1.47%	5.626
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-4.75%	1.97%	685.934
	$\hat{\sigma}_v^2$	-0.28%	1.18%	5.454
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-4.76%	2.06%	684.071
	$\hat{\sigma}_v^2$	-0.32%	1.58%	6.443

Tabela 5.25 – Estimação de  $\hat{\theta}$  para resíduos com distribuição  $t_{\nu=4}(0,1)$  (replicações geradas pelo modelo MCU) e  $n_{sim} = 20000$ .

Com os resultados obtidos na Tabela 5.25, percebemos novamente que parece haver uma convergência do viés dos estimadores para os métodos MV e PMV e com pequenas melhorias para os métodos ADF e do EQM em comparação com a Tabela 5.24. A diferença dos métodos ADF em relação aos outros diminuiu de modo que o método MQGC passou a apresentar resultados semelhantes à MPV.

As Tabelas 5.26 a 5.30 incluem resultados para o ajuste do modelo MCU a partir de dados gerados pelo modelo MCU-C com os resíduos seguindo a distribuição  $t_{\nu=4}(0,1)$ .

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-27.69%	28.78%	5605.83
	$\hat{\sigma}_v^2$	-27.23%	14.21%	2105.61
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-21.63%	27.24%	4315.52
	$\hat{\sigma}_v^2$	12.39%	19.83%	1037.65
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-30.93%	28.10%	6231.07
	$\hat{\sigma}_v^2$	-22.88%	13.58%	1571.93
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-16.44%	28.98%	4003.28
	$\hat{\sigma}_v^2$	-11.61%	18.45%	824.46

Tabela 5.26 – Estimação de  $\hat{\theta}$  para resíduos com distribuição  $t_{\nu=4}(0,1)$  (replicações geradas pelo modelo MCU-C) e  $n_{sim} = 100$ .

A Tabela 5.26 inclui os resultados para o modelo MCU que é ajustado aos dados simulados de  $Y_{it}$  gerados pelo modelo MCU-C. Conclusões razoavelmente semelhantes podem ser utilizadas quando se comparam os resultados da Tabela 5.26 para aqueles na Tabela 5.21. Além disso, podemos notar que a maioria dos métodos teve um aumento em seu EQM comparado com o caso que desconsidera o desenho amostral, principalmente devido a um aumento modesto do viés. Novamente os métodos ADF são piores do que os demais já que temos um pequeno tamanho de amostra.

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-21.69%	12.30%	2433.124
	$\hat{\sigma}_v^2$	-10.89%	8.77%	384.942
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-16.65%	11.61%	2012.101
	$\hat{\sigma}_v^2$	-6.51%	8.00%	157.492
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-22.31%	11.58%	2503.717
	$\hat{\sigma}_v^2$	-9.44%	8.52%	308.016
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-18.60%	12.51%	1902.044
	$\hat{\sigma}_v^2$	-5.08%	8.49%	178.086

Tabela 5.27 – Estimação de  $\hat{\theta}$  para resíduos com distribuição  $t_{\nu=4}(0,1)$  (replicações geradas pelo modelo MCU-C) e  $n_{sim} = 500$ .

A partir dos resultados incluídos na Tabela 5.27, podemos fazer observações muito semelhantes às que são feitas para os resultados apresentados na Tabela 5.26, em termos de comparação entre os métodos. Mais uma vez, o estimador MQGC apresentou o maior EQM entre os métodos para  $\sigma_u^2$  e o estimador MQG o maior EQM para  $\sigma_v^2$ , embora em termos de c.v. sejam muito semelhantes aos outros métodos. Os resultados da Tabela 5.27 indicam uma diminuição de viés, c.v. e EQM em relação à Tabela 5.26.

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-12.33%	3.03%	754.020
	$\hat{\sigma}_v^2$	-1.71%	1.69%	14.160
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-12.04%	3.13%	733.250
	$\hat{\sigma}_v^2$	-1.09%	2.16%	11.610
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-12.25%	2.96%	743.430
	$\hat{\sigma}_v^2$	-1.12%	1.61%	12.030
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-12.08%	3.18%	729.230
	$\hat{\sigma}_v^2$	-1.04%	2.22%	12.300

Tabela 5.28 – Estimação de  $\hat{\theta}$  para resíduos com distribuição  $t_{\nu=4}(0,1)$  (replicações geradas pelo modelo MCU-C) e  $n_{sim} = 10000$ .

A partir dos resultados incluídos na Tabela 5.28, percebemos grandes diferenças em relação à tabela 5.27 e 5.26 no que diz respeito aos valores de viés, c.v. e EQM entre os métodos. Não há uma grande diferença entre o viés dos métodos na Tabela 5.13 quando comparado às Tabelas 5.12 e 5.11. Porém há uma expressiva diminuição do c.v.. Essa diminuição mostra a semelhança entre o MQGC e o MV no que diz respeito ao comportamento. Em termos do EQM os estimadores ADF apresentam os maiores valores embora possuam ganhos de eficiência superiores aos outros com o incremento do tamanho da amostra.

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-12.16%	2.53%	706.500
	$\hat{\sigma}_v^2$	-0.97%	1.43%	9.720
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-11.88%	2.38%	695.890
	$\hat{\sigma}_v^2$	-0.83%	1.72%	7.350
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-12.01%	2.38%	705.750
	$\hat{\sigma}_v^2$	-0.82%	1.36%	8.260
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-11.84%	2.46%	705.370
	$\hat{\sigma}_v^2$	-0.73%	1.77%	7.870

Tabela 5.29 – Estimação de  $\hat{\theta}$  para resíduos com distribuição  $t_{\nu=4}(0,1)$  (replicações geradas pelo modelo MCU-C) e  $n_{sim} = 15000$ .

A Tabela 5.29 traz resultados muito semelhantes à Tabela 5.28. Os valores de viés, c.v. e EQM diminuíram, embora isso não tenha ocorrido de forma tão perceptível. Há pequenas diferenças entre os estimadores indicando que para  $n$  grande os estimadores se comportam de forma muito semelhante. Além disso, o c.v. e o EQM são praticamente iguais para os estimadores, havendo uma ligeira vantagem da MV sobre os demais na estimação dos parâmetros.

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-12.00%	2.13%	705.250
	$\hat{\sigma}_v^2$	-0.85%	1.26%	6.910
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-12.01%	1.97%	689.650
	$\hat{\sigma}_v^2$	0.84%	1.91%	9.100
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-12.00%	1.97%	691.990
	$\hat{\sigma}_v^2$	-0.56%	1.18%	5.740
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-12.01%	2.05%	691.460
	$\hat{\sigma}_v^2$	-0.71%	1.76%	7.760

Tabela 5.30– Estimação de  $\hat{\theta}$  para resíduos com distribuição  $t_{\nu=4}(0,1)$  (replicações geradas pelo modelo MCU-C) e  $n_{sim} = 20000$ .

A partir dos resultados da Tabela 5.30 percebemos uma ligeira diminuição do c.v. em comparação com a Tabela 5.29. Além disso, o estimador MQGC apresentou um valor muito próximo de EQM do estimador de MV para  $\sigma_u^2$  e um menor valor para  $\sigma_v^2$ . Isso confirma a necessidade de um grande tamanho de amostra para que os métodos ADF possam superar em termos de eficiência os outros, já que seus coeficientes de variação já são relativamente menores.

#### 4.2.4- Resíduos com Distribuição t Assimétrica

Os resultados apresentados nessa subseção dizem respeito às simulações com os resíduos seguindo a distribuição de probabilidade  $t_{\nu=4}(0,1)$  assimétrica com parâmetro de assimetria  $\lambda = 3$  e os modelos utilizados são o MCU (Tabelas 5.31 a 5.35) e o MCU-C (Tabelas 5.36 a 5.40).

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-27.88%	31.91%	9609.960
	$\hat{\sigma}_v^2$	-34.31%	18.83%	330.960
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-10.17%	29.82%	8244.550
	$\hat{\sigma}_v^2$	14.35%	3.87%	42.860
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-31.95%	32.18%	9196.430
	$\hat{\sigma}_v^2$	-29.54%	18.68%	259.770
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-18.82%	30.71%	8436.940
	$\hat{\sigma}_v^2$	15.70%	7.31%	143.883

Tabela 5.31– Estimação de  $\hat{\theta}$  para resíduos com distribuição  $t_{v=4}(0,1)$  assimétrica (replicações geradas pelo modelo MCU) e  $n_{sim} = 100$ .

Ao observar a Tabela 5.31 notam-se grandes diferenças em relação à Tabela 5.21. O fato dos dados serem gerados de forma assimétrica inflacionou de forma expressiva o viés e a variância dos estimadores. Há certa diferença ao comparar o EQM dos métodos ADF em relação aos demais métodos de estimação. Essas diferenças são perceptíveis e esperadas já que é esperado tamanho de amostra maior para que os métodos ADF apresentem melhores resultados. Porém, ao comparar a geração dos dados da distribuição  $t_{v=4}(0,1)$  assimétrica com a normal percebemos grandes diferenças no que diz respeito ao comportamento dos estimadores. A Tabela 5.11 apresenta maiores valores de viés, c.v. e EQM. Percebemos que essa diferença é ainda maior com o aumento do tamanho da amostra, nos resultados das tabelas seguintes.

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-16.18%	24.82%	3537.190
	$\hat{\sigma}_v^2$	-15.85%	9.75%	70.873
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-8.16%	25.37%	3103.160
	$\hat{\sigma}_v^2$	1.16%	1.49%	5.570
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-16.23%	23.99%	3399.560
	$\hat{\sigma}_v^2$	-14.26%	9.48%	66.534
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-5.82%	18.71%	3136.940
	$\hat{\sigma}_v^2$	-2.70%	5.11%	13.343

Tabela 5.32– Estimação de  $\hat{\theta}$  para resíduos com distribuição  $t_{v=4}(0,1)$  assimétrica (replicações geradas pelo modelo MCU) e  $n_{sim} = 500$ .

A partir dos resultados incluídos na Tabela 5.32, podemos fazer observações muito semelhantes às que são feitas para os resultados apresentados na Tabela 5.31, em termos de comparação entre os métodos. O estimador MQGC teve o maior EQM entre os métodos para  $\sigma_u^2$  e o estimador MQG o maior EQM para  $\sigma_v^2$ . Os resultados da Tabela 5.12 indicam uma diminuição de viés e EQM sobressaindo os métodos de MV e PMV sobre os estimadores ADF.

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-6.74%	2.96%	512.449
	$\hat{\sigma}_v^2$	-3.59%	2.29%	7.061
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-5.82%	2.77%	515.790
	$\hat{\sigma}_v^2$	-0.30%	0.96%	3.327
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-6.69%	2.68%	516.528
	$\hat{\sigma}_v^2$	-3.28%	1.29%	5.848
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-5.56%	2.83%	511.769
	$\hat{\sigma}_v^2$	-0.39%	1.31%	4.610

Tabela 5.33– Estimação de  $\hat{\theta}$  para resíduos com distribuição  $t_{v=4}(0,1)$  assimétrica (replicações geradas pelo modelo MCU) e  $n_{sim} = 10000$ .

A partir dos resultados incluídos na Tabela 5.33, podemos fazer observações muito semelhantes às que são feitas para os resultados apresentados na Tabela 5.32 e na Tabela 5.31, em termos de comparação entre os métodos. Com o aumento do tamanho da amostra temos uma semelhança entre o MQGC e o MV no que diz respeito ao comportamento. Em termos do EQM os estimadores ADF apresentam os maiores valores embora possuam ganhos de eficiência superiores aos outros com o incremento do tamanho da amostra.

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-5.34%	1.06%	388.242
	$\hat{\sigma}_v^2$	-2.93%	1.36%	3.440
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-5.59%	1.57%	360.430
	$\hat{\sigma}_v^2$	-0.31%	0.25%	3.100
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-5.60%	1.17%	363.300
	$\hat{\sigma}_v^2$	-2.69%	1.01%	3.600
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-5.57%	1.76%	361.428
	$\hat{\sigma}_v^2$	-0.33%	1.12%	3.400

Tabela 5.34– Estimação de  $\hat{\theta}$  para resíduos com distribuição  $t_{\nu=4}(0,1)$  assimétrica (replicações geradas pelo modelo MCU) e  $n_{sim} = 15000$ .

Os resultados incluídos na Tabela 5.34 são muito semelhantes aos apresentados na Tabela 5.13 (normal assimétrica) em termos de comparação entre os métodos. O viés praticamente se manteve constante com o c.v. diminuindo de forma muito pequena. Os EQMs dos métodos foram muito similares, mas podemos perceber que os métodos ADF se aproximaram dos outros, praticamente deixando de haver diferenças entre eles.

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-4.41%	1.03%	331.805
	$\hat{\sigma}_v^2$	-2.43%	1.34%	2.630
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-5.05%	1.25%	299.117
	$\hat{\sigma}_v^2$	-0.31%	0.21%	1.365
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-5.28%	1.12%	303.470
	$\hat{\sigma}_v^2$	-2.24%	0.95%	1.560
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-5.04%	1.45%	302.896
	$\hat{\sigma}_v^2$	-0.38%	0.78%	1.870

Tabela 5.35– Estimação de  $\hat{\theta}$  para resíduos com distribuição  $t_{\nu=4}(0,1)$  assimétrica (replicações geradas pelo modelo MCU) e  $n_{sim} = 20000$ .

Conclusões semelhantes podem ser utilizadas quando se compara a Tabela 5.35 com as Tabelas 5.34 e 5.33. Além disso, o estimador MQGC apresentou um valor muito próximo de EQM do estimador de MPV para  $\sigma_u^2$  e um valor menor para  $\sigma_v^2$ . Isso confirma a necessidade de um grande tamanho de amostra para que os métodos ADF possam superar em termos de eficiência os outros, já que seus coeficientes de variação já são menores.

As Tabelas 5.36 a 5.40 incluem resultados para o ajuste do modelo MCU a partir de dados gerados pelo modelo MCU-C com os resíduos seguindo a distribuição  $t_{\nu=4}(0,1)$  assimétrica.

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-31.50%	30.66%	10405.070
	$\hat{\sigma}_v^2$	-33.12%	25.05%	3500.453
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	-28.83%	28.71%	10438.600
	$\hat{\sigma}_v^2$	4.72%	3.99%	48.300
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-31.72%	33.14%	13966.020
	$\hat{\sigma}_v^2$	-36.80%	27.16%	4126.398
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	-10.02%	28.30%	11216.610
	$\hat{\sigma}_v^2$	-6.54%	7.84%	161.048

Tabela 5.36– Estimação de  $\hat{\theta}$  para resíduos com distribuição  $t_{v=4}(0,1)$  assimétrica (replicações geradas pelo modelo MCU-C) e  $n_{sim} = 100$ .

A Tabela 5.36 inclui os resultados para o modelo MCU dado que os valores simulados de  $Y_{it}$  são gerados pelo modelo MCU-C. Conclusões razoavelmente semelhantes podem ser utilizadas quando se comparam os resultados da Tabela 5.36 para aqueles na Tabela 5.31. Além disso, podemos notar que a maioria dos métodos teve um aumento em seu EQM em comparação com o caso que desconsidera o desenho amostral. Novamente os métodos ADF são piores do que os demais já que temos um pequeno tamanho de amostra. Em relação ao c.v. todos os métodos apresentam valores elevados embora os métodos ADF apresentem resultados mais elevados que os demais.

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-7.87%	22.07%	2421.680
	$\hat{\sigma}_v^2$	-5.69%	9.44%	468.032
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	8.92%	25.59%	3401.590
	$\hat{\sigma}_v^2$	2.36%	1.41%	5.358
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-8.80%	23.41%	3484.731
	$\hat{\sigma}_v^2$	-3.83%	9.26%	433.347
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	5.91%	25.91%	3470.228
	$\hat{\sigma}_v^2$	-2.59%	3.84%	37.002

Tabela 5.37– Estimação de  $\hat{\theta}$  para resíduos com distribuição  $t_{v=4}(0,1)$  assimétrica (replicações geradas pelo modelo MCU-C) e  $n_{sim} = 500$ .

A partir dos resultados da Tabela 5.37, podemos fazer observações muito semelhantes às que são feitas para os resultados apresentados na Tabela 5.36, em termos de comparação entre os métodos. O estimador MQGC teve o maior EQM entre os métodos para  $\sigma_u^2$  e o estimador MQG o maior EQM para  $\sigma_v^2$ . Os resultados da Tabela 5.37 indicam uma diminuição expressiva de viés e EQM comparados com os da Tabela 5.36, provocados por um aumento no número das observações.

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	-6.39%	6.77%	232.760
	$\hat{\sigma}_v^2$	-2.96%	2.30%	14.210
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	5.72%	7.45%	289.440
	$\hat{\sigma}_v^2$	2.02%	0.30%	0.236
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	-0.31%	6.60%	221.095
	$\hat{\sigma}_v^2$	-2.66%	2.18%	16.756
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	0.97%	8.63%	391.790
	$\hat{\sigma}_v^2$	-0.08%	1.08%	2.908

Tabela 5.38– Estimação de  $\hat{\theta}$  para resíduos com distribuição  $t_{v=4}(0,1)$  assimétrica (replicações geradas pelo modelo MCU-C) e  $n_{sim} = 10000$ .

A partir dos resultados incluídos na Tabela 5.38, percebemos uma grande diminuição do c.v. e do EQM em comparação com a Tabela 5.37, porém o viés não diminui de forma expressiva. Com esse aumento de amostra o estimador MQGC apresenta o menor EQM para  $\sigma_u^2$ , mas essa diferença vem diminuindo. Percebemos novamente o indício de que para grandes tamanhos de amostra os estimadores ADF possuem melhor eficiência do que os outros.

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	0.12%	5.23%	226.908
	$\hat{\sigma}_v^2$	-2.63%	2.26%	12.639
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	1.30%	9.35%	382.237
	$\hat{\sigma}_v^2$	0.01%	0.25%	0.167
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	0.22%	6.18%	214.588
	$\hat{\sigma}_v^2$	-2.32%	2.19%	10.1560
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	0.25%	8.25%	370.650
	$\hat{\sigma}_v^2$	-0.07%	1.16%	3.354

Tabela 5.39 – Estimação de  $\hat{\theta}$  para resíduos com distribuição  $t_{\nu=4}(0,1)$  assimétrica (replicações geradas pelo modelo MCU-C) e  $n_{sim} = 15000$ .

A partir dos resultados incluídos na Tabela 5.39, percebemos uma leve diminuição do c.v. e do EQM em comparação com a Tabela 5.38. Podemos fazer observações muito semelhantes às que são feitas para os resultados apresentados na Tabela 5.38, em termos de comparação entre os métodos. O método MQGC apresenta aumento de eficiência para  $\sigma_u^2$ .

Parâmetro		Viés relativo	$cv(\hat{\theta})$	EQM ( $\times 1000$ )
$\hat{\theta}_{MQG}$	$\hat{\sigma}_u^2$	0.10%	5.60%	160.965
	$\hat{\sigma}_v^2$	-2.02%	1.78%	7.758
$\hat{\theta}_{MV}$	$\hat{\sigma}_u^2$	1.04%	6.54%	227.987
	$\hat{\sigma}_v^2$	<0.00%	0.22%	0.130
$\hat{\theta}_{MQGC}$	$\hat{\sigma}_u^2$	0.20%	5.60%	161.039
	$\hat{\sigma}_v^2$	-1.59%	1.67%	5.605
$\hat{\theta}_{MPV}$	$\hat{\sigma}_u^2$	0.80%	6.41%	220.560
	$\hat{\sigma}_v^2$	0.02%	0.79%	1.552

Tabela 5.40– Estimação de  $\hat{\theta}$  para resíduos com distribuição  $t_{\nu=4}(0,1)$  assimétrica (replicações geradas pelo modelo MCU-C) e  $n_{sim} = 20000$ .

Com os resultados obtidos na Tabela 5.40, uma diminuição do c.v. e do EQM em comparação com a Tabela 5.39. A diferença dos métodos ADF em relação aos outros diminuiu ainda mais com esse último tamanho da amostra. Além disso, o c.v. do método MQGC é o menor de todos para  $\sigma_u^2$  e a diferença para  $\sigma_v^2$  também diminuiu.

## 5 - Considerações Finais

Nesta monografia, investigamos o comportamento de métodos propostos por Vieira (2009) para a análise de dados longitudinais levando em consideração o desenho amostral complexo. Técnicas estatísticas especiais são necessárias ao analisarmos dados longitudinais, uma vez que a correlação entre as observações para o mesmo indivíduo deve ser levada em consideração. Além disso, desenhos amostrais complexos resultam, geralmente, em uma correlação entre diferentes indivíduos o que viola pressupostos clássicos da inferência estatística. Por isso, é importante que os analistas de dados longitudinais levem em consideração as características de planos amostrais complexos.

Inicialmente, fizemos uma revisão sobre procedimentos de estimação para parâmetros e modelos longitudinais de regressão considerando a situação clássica, que pressupõe amostragem aleatória simples e observações IID, e o contexto dos dados amostrais complexos. Os métodos de máxima verossimilhança, pseudo máxima verossimilhança e mínimos quadrados generalizados foram discutidos no contexto de estimação dos parâmetros dos modelos longitudinais considerando o plano amostral complexo.

Resultados de uma investigação empírica conduzida no âmbito desta monografia utilizando dados longitudinais amostrais da *British Household Panel Survey* (BHPS) também foram apresentados. As características da BHPS, do banco de dados analisado, e de seu desenho amostral foram apresentadas.

O estudo de simulação teve como objetivo avaliar os métodos de estimação discutidos em cenários que não haviam sido considerados por Vieira (2009), que havia considerados apenas distribuições simétricas para os resíduos dos modelos adotados para geração dos dados. Comparamos os métodos em termos de viés, coeficiente de variação (c.v.) e erro quadrático médio (EQM), para dados com diferentes estruturas (com efeito da conglomeração e sem o efeito da conglomeração), gerados a partir de diferentes distribuições de probabilidade (normal e t) tanto simétricas quanto assimétricas e considerando vários tamanhos de amostra (100, 500, 10000, 15000, 20000).

Para os resíduos seguindo a distribuição normal percebemos que os métodos estudados se comportaram de maneira muito semelhante, confirmando resultados de

Vieira (2009). Porém, para cenários em que o modelo MCU-C foi adotado para geração dos dados, os métodos ADF apresentaram melhores medidas de eficiências quando comparados aos demais sobretudo para um tamanho de amostra menor. Conclusões semelhantes podem ser encontradas nos cenários em que os resíduos seguiam a distribuição  $t_{\nu=4}(0,1)$ , porém, com resultados apresentando viés e EQM mais elevados.

Para cenários que consideram resíduos gerados de forma assimétrica, percebemos grandes diferenças nos valores do viés, c.v. e EQM. Para a normal assimétrica, os resultados apresentam altos valores de viés e uma baixa eficiência. Com a distribuição  $t_{\nu=4}(0,1)$ , os resultados apresentam menores valores de viés e de EQM. O fato da distribuição  $t$  apresentar caudas mais pesadas (atribuindo assim maior probabilidade a eventos que ocorrem em suas caudas do que as distribuições de cauda mais curta) parece causar menos problemas aos métodos de estimação do que a distribuição normal.

Além disso, nossos resultados indicam que os métodos ponderados parecem ter viés ligeiramente inferior quando comparados com os não ponderados, embora não há nenhuma razão teórica clara para esse resultado, como já mencionado anteriormente.

Os métodos da MV (e MPV) apresentam estimadores, em geral, com bons desempenhos em termos de viés e variância independentes do tamanho da amostra. Já os métodos MQG e MQGC, apresentam bons desempenhos em termos de viés e variância, em geral, para amostras de tamanho acima de 10000. Portanto, recomendamos que os métodos ADF devam ser utilizados com cautela em situações onde apenas amostras de pequeno tamanho são disponíveis.

Outro fato observado é que na maioria dos casos o EQM das estimativas de  $\sigma_u^2$  é muito superior ao viés de  $\sigma_v^2$ . O parâmetro  $\sigma_u^2$  mede a variabilidade dos efeitos aleatórios permanentes e  $\sigma_v^2$  mede a variabilidade dos efeitos aleatórios transitórios. Desta forma, o segundo parâmetro, por variar no tempo, considera  $NT$  observações enquanto que o outro considera apenas  $N$  observações.

Os métodos ADF tiveram um bom desempenho geral em várias situações, apresentando, em geral, os menores valores de coeficiente de variação. Para entender

plenamente as razões do por que disso acontecer precisamos de novas investigações no futuro.

Adicionalmente, podemos pensar em trilhar diversos caminhos em estudos futuros. Por exemplo, alterando o parâmetro de assimetria  $\lambda = 3$  é possível estudar os impactos de diferentes graus de assimetria na estimação dos parâmetros. Além disso, outras distribuições assimétricas com caudas mais pesadas poderiam ser utilizadas como a slash, normal contaminada, dentre outras.

Outros tamanhos de amostra poderiam ser utilizados assim como outros métodos de estimação considerados também por Vieira (2009) e não contemplados nesta monografia. Em relação à distribuição t poderíamos alterar seus graus de liberdade e estudar o impacto na estimação dos parâmetros. Poderíamos também alterar o modelo ajustado aos dados, pois além do MCU poderíamos ajustar o MCU-C e também o AR1. O modelo em si, também poderia ter sua complexidade ampliada permitindo, além de interceptos aleatórios, que os demais coeficientes fossem aleatórios entre os indivíduos.

## Referências Bibliográficas

- Aitkin, M. and Longford, N. (1986) Statistical modelling in school effectiveness studies (with discussion). *Journal of the Royal Statistical Society, Series A*, 149, 1–43
- Anderson, T. W. (1973) Asymptotically Efficient Estimation of Covariance Matrices with Linear Structure. *Annals of Statistics*, Vol. 1, N. 1, 135-141.
- Azzalini, A. e Capitanio, A. (2014). *The Skew-Normal and Related Families*. Cambridge University Press, IMS Monographs series.
- Bentler, P. M. and Weeks, D. G. (1980) Linear Structural Equations with Latent Variables. *Psychometrika*, Vol. 45, N. 3, 289-308.
- Berrington, A. (2002) Exploring Relationships Between Entry Into parenthood and Gender Role Attitudes: Evidence from the British Household Panel Study. In Lesthaeghe, R. ed *Meaning and Choice: Value Orientations and Life Course Decisions*. Brussels, NIDI.
- Bollen, K. A. (1989) *Structural Equations with Latent Variables*. New York, John Wiley & Sons.
- Browne, M. W. (1982) Covariance Structures. In Hawkins, D. M. eds. *Topics in Applied Multivariate Analysis*. Cambridge, Cambridge University Press.
- Browne, M. W. (1984) Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, 37, 62-83.
- Chambers, R. L. and Skinner, C. J. eds. (2003) *Analysis of Survey Data*. Chichester, John Wiley & Sons.
- Diggle, P. J., Heagerty, P., Liang, K. & Zeger, S. L. (2002) *Analysis of Longitudinal Data*. 2<sup>nd</sup> ed. Oxford, Oxford University Press.
- Ferreira, C.; Lachos, V.H.; Bolfarine, H. Inference and Diagnostics in Skew Scale mixtures of Normal Regression Models (2013). *Journal of Statistical Computation and Simulation*. DOI: <http://dx.doi.org/10.1080/00949655.2013.828057>
- Goldstein, H. (1995) *Multilevel Statistical Models*. 2<sup>nd</sup> ed., London, Arnold.

- Højsgaard, S., Halekoh, U. & Yan J. (2005) The R Package geepack for Generalized Estimating Equations *Journal of Statistical Software*, 15, 2, pp1—11.
- Isaki, C. T. and Fuller, W. A. (1982) Survey Design Under the Regression Superpopulation Model. *Journal of the American Statistical Association*. Vol. 77, n. 377, 89-96.
- Johnson, R. A. and Wichern, D. W. (1998) *Applied Multivariate Statistical Analysis*. 4<sup>th</sup> ed., Upper Saddle River, Prentice-Hall.
- Jöreskog, K. G. (1970) A General Method for Analysis of Covariance Structures. *Biometrika*, Vol. 57, N. 2, 239-251.
- Jöreskog, K. G. and Goldberger, A. S. (1972) Factor Analysis by Generalized Least Squares. *Psychometrika*, 37, 243-260.
- Liang, and Zeger, S.L. (1986). Longitudinal Data Analysis Using Generalised Linear Models. *Biometrika*, 73, 13–22.
- Muthén, B. O. and Satorra, A. (1995) Complex Sample Data in Structural Equation Modelling. *Sociological Methodology*, Vol. 25, 267-316.
- Pan, J. and Mackenzie, G. (2003) On Modelling Mean-covariance Structures in Longitudinal Studies. *Biometrika*, Vol. 90, n. 1, 239-2449
- Pfeffermann, D. and LaVange, L. (1989) Regression Models for Stratified Multi-Stage Cluster Samples. In Skinner, C. J., Holt, D. and Smith, T. M. F. eds. *Analysis of Complex Surveys*. Chichester, John Wiley & Sons.
- Pourahmadi, M. (1999) Joint Mean-covariance with Applications to Longitudinal Data: Unconstrained Parameterisation. *Biometrika*, Vol. 86, n. 3, 677-690.
- Pourahmadi, M. (2000) Maximum Likelihood Estimation of Generalised Linear Models for Multivariate Normal Covariance Matrix. *Biometrika*, Vol. 87, n. 2, 425-435.
- R Development Team (2013) *The R Environment for Statistical Computing and Graphics – Reference Index*, Version 2.15.3, The R Foundation for Statistical Computing.
- Rasbash, J., Browne, W., Goldstein, H., Yang, M., Plewis, I., Healy, M., Woodhouse, G., Draper, D., Langford, I., and Lewis, T. (2002) *A user's guide to MLwiN – Version 2.1d for use with MLwiN 1.10*. London, Centre for Multilevel Modelling, Institute of Education, University of London.

- Rencher, A. C. (1998) *Multivariate Statistical Inference and Applications*. New York, John Wiley & Sons.
- Satorra, A. (1992) Asymptotic Robust Inferences in the Analysis of Mean and Covariance Structures. *Sociological Methodology*, Vol. 22, p. 249-278.
- Satorra, A. (2000) Scaled and Adjusted Restricted Tests in Multi-sample Analysis of Moment Structures. In Heijmans, D. D. H., Pollock, D. S. G. and Satorra, A. eds. *Innovations in Multivariate Statistical Analysis: A Festschrift for Henz Neudecker*. Dordrecht, Kluwer Academic Publishers.
- Satorra, A. and Bentler, P. M. (2001) A Scaled Difference Chi-square Test Statistic for Moment Structure Analysis. *Psychometrika*, Vol. 66, N. 4, 507-514.
- Skinner, C. J (2003) Introduction to Part D – Longitudinal Data. In Chambers, R. L. and Skinner, C. J. eds. *Analysis of Survey Data*. Chichester, John Wiley & Sons.
- Skinner, C. J., Holt, D. and Smith, T. M. F. eds. (1989) *Analysis of Complex Surveys*. Chichester, John Wiley & Sons.
- Skinner, C. J. e Vieira, M. D. T. (2007) Variance estimation in the analysis of clustered longitudinal survey data. *Survey Methodology*, Vol. 33, No. 1, pp. 3-12
- Taylor, M. F. ed, with Brice J., Buck, N. and Prentice-Lane E. (2001) *British Household Panel Survey - User Manual - Volume A: Introduction, Technical Report and Appendices*. Colchester, University of Essex.
- Twisk, J. W. R. (2003) *Applied Longitudinal Data Analysis for Epidemiology: A Practical Guide*. Cambridge: Cambridge University Press.
- Vieira, M. D. T. (2009) *Analysis of Longitudinal Survey Data*. 1. ed. Saarbrücken: VDM Verlag Dr. Müller.
- Vieira, M. D. T., Bastos, R. R., Souza, A. C., Hippert, H. S. (2011) On the use of exploratory and confirmatory longitudinal data analysis. *Advances and Applications in Statistics*, v.22, p. 129-156.
- Wiley, D. E., Schmidt, W. H. and Bramble, W. J. (1973) Studies of a Class of Covariance Structure Models. *Journal of the American Statistical Association*, Vol. 68, N. 342, 317-323.

Yuan, K-H. and Bentler, P. M. (1997a) Mean and Covariance Structure Analysis: Theoretical and Practical Improvements. *Journal of the American Statistical Association*, Vol. 92, N. 438, 767-774.

## Apêndice A

### Afirmações sobre atitudes em relação ao papel do gênero

A Tabela A.1 apresenta um conjunto de afirmações sobre a família, o papel da mulher e trabalhos domésticos, que foram incluídos ao questionário das ocasiões um, três, cinco, sete e nove da BHPS.

---

1	A pre-school child is likely to suffer if his or her mother works
2	All in all, family life suffers when the woman has a full-time job
3	A woman and her family would all be happier if she goes out to work
4	Both the husband and wife should contribute to the household income
5	Having a full-time job is the best way for a woman to be an independent person
6	A husband's job is to earn money; a wife's job is to look after the home and family
7	Children need a father to be as closely involved in their upbringing as the mother
8	Employers should make special arrangements to help mothers combine jobs and childcare
9	A single parent can bring up children as well as a couple

---

Fonte: Berrington (2002)

Tabela A.1 – Afirmações sobre atitudes em relação ao papel do gênero.