



**IURI DOS SANTOS MANOEL**

**SIMULAÇÃO DE MODELOS MULTINÍVEIS COM  
REGRESSÕES ADAPTATIVAS ROBUSTAS EM CONDIÇÕES  
EXPERIMENTAIS DE ARMAZENAMENTO DE CAFÉS  
ESPECIAIS EM AMBIENTES REFRIGERADOS**

**LAVRAS – MG**

**2021**

**IURI DOS SANTOS MANOEL**

**SIMULAÇÃO DE MODELOS MULTINÍVEIS COM REGRESSÕES ADAPTATIVAS  
ROBUSTAS EM CONDIÇÕES EXPERIMENTAIS DE ARMAZENAMENTO DE  
CAFÉS ESPECIAIS EM AMBIENTES REFRIGERADOS**

Dissertação de mestrado apresentado à  
Universidade Federal de Lavras, como parte  
das exigências do Programa de Pós-Graduação  
em Estatística e Experimentação Agropecuária,  
para a obtenção do título de Mestre.

Prof. DSc. Marcelo Ângelo Cirilo  
Orientador

**LAVRAS – MG**

**2021**

**Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca  
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).**

Manoel, Iuri dos Santos.

Simulação de Modelos Multiníveis com Regressões  
Adaptativas Robustas em Condições Experimentais de  
Armazenamento de Cafés Especiais em Ambientes Refrigerados /  
Iuri dos Santos Manoel. - 2021.

49 p. : il.

Orientador(a): Marcelo Ângelo Cirillo.

Dissertação (mestrado acadêmico) - Universidade Federal de  
Lavras, 2021.

Bibliografia.

1. Regressões. 2. Outliers. 3. Simulações. I. Cirillo, Marcelo  
Ângelo. II. Título.

**IURI DOS SANTOS MANOEL**

**SIMULAÇÃO DE MODELOS MULTINÍVEIS COM REGRESSÕES ADAPTATIVAS  
ROBUSTAS EM CONDIÇÕES EXPERIMENTAIS DE ARMAZENAMENTO DE  
CAFÉS ESPECIAIS EM AMBIENTES REFRIGERADOS  
USE OF UFLAMON CLASS: SAMPLE FOR USERS**

Dissertação de mestrado apresentado à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, para a obtenção do título de Mestre.

APROVADA em 26/02/2021.

Evelise Roman Corbalan Góis Freire	UFLA
Fabyano Fonseca e Silva	UFV
Jackelya Araújo da Silva	UFPI

Prof. DSc. Marcelo Ângelo Cirilo  
Orientador

**LAVRAS – MG  
2021**

## RESUMO

Inúmeros fatores contribuem para a qualidade dos cafés especiais, dentre eles, pode-se citar as condições de armazenamento e refrigeração. Desta forma, podemos conjecturar que os resultados das avaliações sensoriais poderão ser corrompidos devido a erros de mensurações, principalmente quando os provadores não são treinados, ocasionando o surgimento de observações discrepantes (*outliers*). Com essa motivação, este trabalho teve por objetivo propor cenários de simulação, considerando valores paramétricos obtidos pelo ajuste de um modelo multinível incorporando regressões adaptativas robustas a presença de *outliers* em um experimento real com cafés beneficiados e não beneficiados armazenados em diferentes períodos e temperaturas. Neste contexto, considerou-se cenários computacionalmente simulado, em que, tais erros na resposta da variável nota sensorial, poderão ser cometidos em L=5 e 10 unidades. Concluiu-se que a metodologia proposta para inferir notas sensoriais simuladas em experimento definido pelas condições experimentais de armazenamento e ambientes refrigerados foi viável por contemplar características robustas em amostras que sejam caracterizadas com até 30% de observações *outliers* nos cenários de simulação avaliados.

**Palavras-chave:** Grãos de café beneficiados. Regressões. *Outlier*. Simulação

## ABSTRACT

Numerous factors contribute to the quality of specialty coffees, including storage and refrigeration conditions. In this sense, we can conjecture that the results of sensory evaluations may be corrupted due to measurement errors, mainly when tasters are not trained, causing the appearance of observation outliers. With this motivation, this study aimed to propose simulation scenarios considering parametric values obtained by adjusting a multilevel model incorporating robust adaptive regressions to the presence of outliers in an actual experiment with processed and non-processed coffees stored at different periods and temperatures. In this context, computationally simulated scenarios, in which the errors in the response of the variable sensory score could be made at  $L = 5$  and 10 units, were considered. The proposed methodology to infer simulated sensory scores in an experiment defined by the experimental conditions of storage and refrigerated environments was feasible because it contemplates robust characteristics in samples characterized with up to 30% of outlier observations in the evaluated simulation scenarios.

**Keywords:** Processed coffee beans. Regressions. Outlier. Simulation.

## LISTA DE FIGURAS

Figura 2.1 – Estrutura hierárquica de 2 níveis. . . . .	13
Figura 2.2 – Esquema de dispersão amostral entre níveis. . . . .	31
Figura 3.1 – Esquema de hierarquia de um MLH de 3 níveis. . . . .	34
Figura 4.1 – Desempenho de regressões LTS e LMS em cafés beneficiados, considerando erro de mensuração em 5 unidades. . . . .	39
Figura 4.2 – Desempenho de regressões LTS e LMS em cafés não beneficiados, considerando erro de mensuração em 5 unidades . . . . .	39
Figura 4.3 – Desempenho de regressões LTS e LMS em cafés beneficiados, considerando erro de mensuração em 10 unidades . . . . .	41
Figura 4.4 – Desempenho de regressões LTS e LMS em cafés não beneficiados, considerando erro de mensuração em 10 unidades . . . . .	42

## LISTA DE TABELAS

Tabela 2.1 – Relações entre variáveis em uma estrutura hierárquica . . . . .	14
Tabela 2.2 – Categorização dos Modelos Lineares Hierárquicos de até 3 níveis . . . . .	19
Tabela 2.3 – Modelos Lineares Hierárquicos de 3 níveis (combinado) . . . . .	19
Tabela 4.1 – Estimativas dos coeficientes de correlação intraclasse (em %) considerando a variável nota sensorial e um erro de mensuração de 5 unidades. . . . .	37
Tabela 4.2 – Estimativas para os efeitos aleatórios considerando um erro de mensuração de 5 unidades. . . . .	38
Tabela 4.3 – Estimativas para os efeitos fixos considerando um erro de mensuração de 5 unidades. . . . .	38
Tabela 4.4 – Estimativas dos coeficientes de correlação intraclasse (em %) considerando a variável nota sensorial e um erro de mensuração em 10 unidades. . . . .	40
Tabela 4.5 – Estimativas para os efeitos aleatórios considerando um erro de mensuração em 10 unidades. . . . .	40
Tabela 4.6 – Estimativas para os efeitos fixos considerando um erro de mensuração em 10 unidades. . . . .	41

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>8</b>
<b>2</b>	<b>REFERENCIAL TEÓRICO</b>	<b>11</b>
<b>2.1</b>	<b>Armazenamento Refrigerado</b>	<b>11</b>
<b>2.2</b>	<b>Estruturas dos dados em modelos multiníveis</b>	<b>12</b>
<b>2.3</b>	<b>Formalização do Modelo Multinível</b>	<b>14</b>
<b>2.3.1</b>	<b>Equação multinível de três níveis com intercepto aleatório e inclinação fixa</b>	<b>15</b>
<b>2.3.2</b>	<b>Equação multinível de três níveis com intercepto e inclinação aleatórios</b>	<b>16</b>
<b>2.3.3</b>	<b>Suposições</b>	<b>18</b>
<b>2.3.4</b>	<b>Representação Matricial do Modelo</b>	<b>19</b>
<b>2.3.5</b>	<b>Estimando parâmetros fixos <math>\gamma</math></b>	<b>22</b>
<b>2.3.6</b>	<b>Estimando componentes de covariância (<math>G</math> e <math>R_j</math>)</b>	<b>23</b>
<b>2.3.7</b>	<b>Estimando efeitos aleatórios <math>v</math></b>	<b>23</b>
<b>2.4</b>	<b>Modelo de Regressão Adaptativo</b>	<b>23</b>
<b>2.4.1</b>	<b>Regressões Robustas</b>	<b>23</b>
<b>2.5</b>	<b>Estimadores Robustos</b>	<b>25</b>
<b>2.5.1</b>	<b>Ponto de Ruptura</b>	<b>26</b>
<b>2.5.2</b>	<b>Função Influência</b>	<b>27</b>
<b>2.5.3</b>	<b>Conjuntos Elementares</b>	<b>27</b>
<b>2.5.4</b>	<b>Regressão de Mínimos Quadrados Medianos (MQM)</b>	<b>28</b>
<b>2.5.5</b>	<b>Regressão de Mínimos Quadrados Aparados (MQA)</b>	<b>29</b>
<b>2.6</b>	<b>Coefficiente de Correlação Intraclasse (CCI)</b>	<b>30</b>
<b>2.7</b>	<b>Máxima Verossimilhança Restrita (REML)</b>	<b>32</b>
<b>3</b>	<b>MATERIAL E MÉTODOS</b>	<b>33</b>
<b>3.1</b>	<b>Estrutura dos dados</b>	<b>33</b>
<b>3.2</b>	<b>Simulação do modelo nas condições experimentais utilizados na refrigeração de cafés especiais</b>	<b>34</b>
<b>3.3</b>	<b>Incorporação das regressões adaptativas</b>	<b>35</b>
<b>4</b>	<b>RESULTADOS</b>	<b>37</b>
<b>4.1</b>	<b>Simulação do modelo considerando erro de mensuração nas avaliações das notas sensoriais especificado em 5 unidades</b>	<b>37</b>

4.2	Simulação do modelo considerando erro de mensuração nas avaliações das notas sensoriais especificado em 10 unidades . . . . .	40
5	CONCLUSÕES . . . . .	43
	<b>APENDICE A – Equação multinível de três níveis com intercepto aleatório e inclinação fixa.</b> . . . . .	44
	<b>APENDICE B – Equação multinível de três níveis com intercepto e inclinação aleatório.</b> . . . . .	45
	<b>REFERÊNCIAS</b> . . . . .	47

## 1 INTRODUÇÃO

Um dos critérios que contribuem para a produção de cafés com elevada qualidade, almejando atingir mercados internacionais, certamente refere-se aos cuidados com os grãos, não apenas, por fatores genéticos, ambientais, tecnológicos (BORÉM et al., 2019b) e ausência de grãos defeituosos (BRIGHENTI; CIRILLO, 2018) et al., mas também pela forma de armazenamento, uma vez que, cafés expostos a longos períodos de tempo sob condições ambientais desfavoráveis podem sofrer alterações em sua cor, sabor e aroma (BORÉM et al., 2019a).

Resultados de pesquisas têm demonstrado queda na qualidade inicial de cafés após armazenamento sob temperatura elevada e períodos relativamente longo de tempo (CORADI; BORÉM; OLIVEIRA, 2008). Além de alterações físicas, pode-se ter alterações químicas, como a perda de açúcares devido ao aumento da respiração do café, associada principalmente à temperaturas elevadas (JÚNIOR, 2001; ALVES et al., 2017).

Neste contexto, procedimentos alternativos que amenizam a deterioração dos grãos, frente à condições de temperatura e umidade relativa desfavoráveis têm sido empregados no desenvolvimento de novas tecnologias relacionadas a embalagens e armazenamento. Logo, novas metodologias e modelos estatísticos têm agregado informações para que novas condições de armazenamento com foco na preservação da qualidade dos cafés, beneficiados ou não beneficiados, apresentem importantes contribuições.

Abreu et al. (2017b) aplicaram a técnica de otimização de resposta simultânea, com o propósito de determinar a melhor combinação de tempo e das condições de armazenamento para a conservação de grãos de *Coffea arabica* L. colhidos no estágio de maturação cereja, processados por via úmida e natural, sendo secados até atingirem 11% de teor de água. Parte dos grãos foi beneficiada e parte não foi beneficiada antes de ser armazenada por períodos de 3, 6 e 12 meses, em duas condições de ambiente: em ar refrigerado a 10°C umidade relativa de 50% e em 25°C sem controle da umidade relativa do ar. Com estas especificações, os autores concluíram que mantendo a temperatura a 10°C, o tempo de armazenamento adequado para o café natural é aproximadamente sete meses, e três meses para o café despulpado.

Dentre as condições de armazenamento, a refrigeração é um dos métodos mais eficazes quanto à preservação da qualidade inicial de um produto agrícola (RIGUEIRA et al., 2009). Em se tratando da acurácia e precisão dos resultados sensoriais, mais especificamente a pontuação final, validada em um modelo probabilístico por Ferreira et al. (2016), sendo esta, um dos critérios principais que classifica os cafés como especiais, decorrente da influência das condições

de armazenamento e refrigeração, torna-se plausível conjecturar que determinados resultados poderão ser irreais, originando os outliers. Tal fato é corroborado por fatores subjetivos, habilidade e memória sensorial particular a cada provador e/ou consumidor, e essa heterogeneidade tem sido explorada em modelos probabilísticos e metodologias que sejam robustas a essas observações discrepantes.

Neste contexto, Santos, Cirillo e Guimarães (2021) utilizaram regressões adaptativas em modelos de equações estruturais para determinar diferentes perfis de consumidores de cafés especiais, utilizando uma abordagem metodológica na construção de um índice para validar variáveis descritoras utilizadas na composição de construtos para determinar perfis entre consumidores (SANTOS; CIRILLO, 2021).

Resende, Cirillo e Borém (2020) apresentaram um índice para validar preferências de cafés especiais por diferentes grupos de consumidores. A construção deste índice envolve técnicas de componentes principais, modelos de regressão e testes computacionalmente intensivos. Concluíram que o índice foi adequado, justamente por ter discriminado cafés especiais produzidos em diferentes altitudes, anteriormente confirmados por estudos dado por Liska et al. (2015) e Oliveira et al. (2019), ao considerar métodos de classificação automática com modelos automatizados por Machine Learning com diferentes grupos de consumidores.

Neste âmbito, a discriminação desses cafés em blends formados por cafés de outras variedades, em procedimento metodológico proposto por Alves et al. (2017) para avaliar a eficiência de um painel sensorial foi promissor, uma vez que, a discriminação dos blends em relação aos cafés puros foi detectada pelo grupo de consumidores não treinados. Dada a argumentação de que observações outliers podem influenciar substancialmente os resultados das análises sensoriais e que, uma das fontes causadoras dessas observações podem ser caracterizadas por propriedades fisiológicas alteradas em função das condições de armazenamento e temperatura, tem-se como perspectiva o fato de que os modelos de efeitos aleatórios ou hierárquicos possam ser utilizados.

Diante do exposto, considerar as regressões dos fatores, contextualizados por métodos de pós-colheita influenciando na manutenção da qualidade inicial dos grãos em função do tempo de armazenamento, possam apresentar trajetórias não lineares entre os valores preditos, além de apresentar novas medidas para validação da estrutura do delineamento experimental adotado, como por exemplo, o uso do coeficiente de correlação intraclasse.

Ressalta-se que o uso destes modelos conhecidos como modelos multiníveis Vassallo, Durrant e Smith (2015), são frágeis dada quantidade excessiva de outliers, ou seja, respostas discrepantes em relação à média das avaliações. Com esta deficiência recomenda-se que as regressões utilizadas para estes modelos, sejam incorporadas procedimentos robustos a outliers (TORTI et al., 2012). Contudo, a formalização teórica com um embasamento matemático torna-se complexa e inviável de implementada na prática.

Decorrente a esta complexidade, tendo como motivação as suposições relacionadas à manutenção da qualidade sensorial de cafés em ambientes refrigerados, no tocante à ocorrência de resultados discrepantes, objetivou-se com este trabalho construir um modelo multinível adaptativo, de modo que, as estimativas das regressões robustas, obtidas pelos métodos LTS e LMS, acrônimo Least Trimmed Squares (LTS) e Least Median of Squares (LMS) sejam incorporadas (MUHLBAUER; SPICHTINGER; LOHMANN, 2009).

A validação deste modelo, mediante aos dados experimentais foi feita por Simulação Monte Carlo, com parâmetros definidos na análise de um experimento fatorial realizado com grãos de cafés especiais condicionados aos fatores: forma em que o café foi armazenado, se íntegro ou descascado; pontuações do café no início do armazenamento; temperatura do ar de armazenagem; tempo de tomada de amostras durante o armazenamento.

## 2 REFERENCIAL TEÓRICO

### 2.1 Armazenamento Refrigerado

Um dos critérios que contribuem para a produção de cafés com alta qualidade se refere aos cuidados com os grãos que, além de fatores genéticos, ambientais e tecnológicos, a forma de armazenamento é um fator relevante, visto que, a cor, sabor e aroma do café podem ter interferência quando este é submetido a condições extremas (BORÉM et al., 2019b).

A etapa de armazenamento do grão de café tem como objetivo preservar os atributos qualitativos do grão ao máximo, entre o período da colheita e a comercialização do produto (BORÉM et al., 2008). Contudo, há fatores neste processo que afetam negativamente o café, prejudicando-o qualitativamente e, até mesmo, inviabilizando seu consumo.

Pesquisas têm notificado queda na qualidade inicial de cafés após terem sido armazenados sob temperatura elevada e períodos relativamente grandes de tempo (CORADI; BORÉM; OLIVEIRA, 2008). Aspectos fisiológicos, químicos e bioquímicos são alterados durante o período de armazenamento, influenciando sua qualidade (SAATH et al., 2014).

A alteração da cor do grão de café, denominada como branqueamento, é uma alteração física apontada por Abreu et al. (2017b) que ocorre em grãos armazenados sob elevadas temperaturas. Além de alterações físicas, pode-se ter alterações químicas, como a perda de açúcares devido ao aumento da respiração do café, causada principalmente por uma temperatura elevada (JÚNIOR, 2001).

Desta forma, partindo do ponto em que o armazenamento dos grãos de café, sobretudo, a temperatura e o tempo em que estes são submetidos enquanto armazenados, interferem em sua qualidade final, métodos de armazenamento são testados e desenvolvidos com o objetivo de preservar ao máximo sua qualidade inicial.

Neste sentido, a refrigeração dos ambientes de armazenamento dos grãos se mostrou um método eficaz quanto à preservação da qualidade inicial da bebida (RIGUEIRA et al., 2009), pois os aspectos fisiológicos, os responsáveis pelos atributos qualitativos da bebida, são alterados de acordo com a temperatura e tempo de armazenamento.

O armazenamento refrigerado se constitui na retirada do calor no ambiente de armazenagem, controlando assim a temperatura e a umidade relativa, influenciando significativamente no processo de deterioração dos produtos agrícolas armazenados (SANTOS et al., 2005). Segundo SILVA et al. (2021), a redução da temperatura do ambiente de armazenamento para 10°C, du-

rante um período de até 9 meses, inferiu positivamente na preservação dos atributos qualitativos dos grãos de café.

A refrigeração é um dos métodos mais eficazes quanto à preservação da qualidade inicial de um produto agrícola em armazenamento. Por estes produtos serem sujeitos à deterioração e perda de qualidade em função da temperatura e tempo de armazenamento, o controle destes fatores favorece para a preservação dos atributos qualitativos do grão de café (BRACKETT, 1987).

## 2.2 Estruturas dos dados em modelos multiníveis

Hierarquia é uma ordem de elementos, constituindo uma sequência descendente ou ascendente, de acordo com algum critério predefinido, como relevância, valor ou subordinação. Assim, quando se tem um conjunto de dados que é composto por duas ou mais variáveis explicativas, em que estas têm influência entre si, pode-se construir uma estrutura hierárquica, e então, este conjunto é disposto em níveis ou camadas, de acordo com a subordinação de suas variáveis.

O problema ao se analisar dados com uma estrutura hierárquica, é que modelos de regressão usuais têm por suposição que os efeitos aleatórios atribuídos a cada observação são independentes entre si. Desta forma, quando estes efeitos são dependentes, inviabiliza a aplicação de modelos usuais.

Neste contexto, os modelos de regressão multiníveis são utilizados quando os dados em análise seguem uma estrutura hierárquica. Este modelo possibilita ao pesquisador estudar as relações entre variáveis explicativas em diferentes níveis, explicando melhor a variável de interesse (JR; FÁVERO, 2019). Desta forma, a organização dos dados coletados é dada segundo suas dependências, podendo assim dizer, por subordinação.

Considerando um modelo de dois níveis, sendo  $X_{ij}$  e  $W_j$  variáveis explicativas de nível um e dois, respectivamente, em que  $i$  refere-se aos indivíduos e  $j$  aos grupos, pode-se ilustrar uma estrutura hierárquica cujo  $X_{ij}$ , variável de nível inferior, é dependente de  $W_j$ , mas não o oposto. Por exemplo, se a parcela  $X_{32}$  refere-se ao terceiro indivíduo do segundo grupo e  $W_2$  ao mesmo grupo, então  $X_{32}$  é influenciado por  $W_2$ . Portanto, há um grupo de valores atribuídos à  $X_{ij}$  que são relacionados a apenas um  $W_j$ .

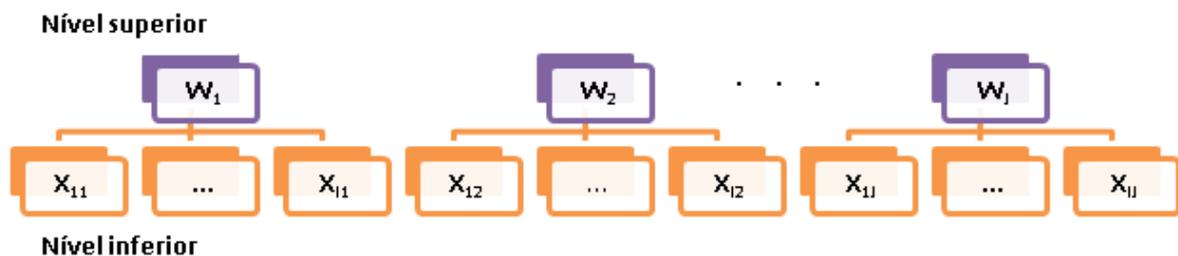
Raudenbush e Bryk (1986) realizaram um estudo cujo objetivo foi explicar o desempenho de estudantes, considerando fatores dos próprios, assim como de suas respectivas institui-

ções. Observa-se uma relação a princípio: aluno e escola. Se  $X_{ij}$  refere-se ao aluno  $i$  da escola  $j$ , pode-se denominar  $X_{ij}$  como uma variável indivíduo, assim como, se  $W_j$  refere-se a escola  $j$ , então esta é denominada como uma variável grupo, uma vez que em uma escola há  $I$  alunos.

Em outro contexto, Short et al. (2009) debateram sobre o empreendedorismo na Suécia, baseando-se em resultados estatísticos obtidos por meio da análise de alguns fatores, tais como as vendas, crescimento e sobrevivência de empresas relativamente jovens. Nesse trabalho, os setores que compõem a empresa foram estudados e classificados em níveis, possibilitando estudar a relevância de cada um destes para com as variáveis de interesse supracitadas. Assim, as observações de cada setor (variável indivíduo), estão relacionadas a uma única variável de nível superior (variável grupo).

Com o intuito de ilustrar a relação entre as variáveis explicativas de um e dois níveis, considere a Figura 2.1. Desta forma, são coletadas informações tanto do nível indivíduo, quanto

Figura 2.1 – Estrutura hierárquica de 2 níveis.



Fonte: Autoria própria.

do quanto do nível grupo, formalizando uma hierarquia. O conjunto de dados dos indivíduos é considerado juntamente com as informações de seu respectivo grupo, possibilitando uma comparação geral entre estes.

De forma análoga acontece quando se trata de uma análise no qual se podem organizar as variáveis explicativas em mais de dois níveis. Jr e Fávero (2019) discutem para o caso de três níveis definindo uma terceira variável, no qual esta é influenciada por alguma das outras duas. Assim, cada valor atribuído a esta, será relacionado a um único valor de nível maior.

Seja  $T_{ijk}$ ,  $X_{jk}$  e  $W_k$  variáveis explicativas de níveis 1, 2 e 3, respectivamente, em que  $i = 1, \dots, I$  referente aos indivíduos,  $j = 1, \dots, J$  aos grupos de segundo nível e  $k = 1, \dots, K$  aos grupos de terceiro nível. Assim, os valores de  $T$  são relacionados com seu  $X$  correspondente, que por sua vez, é relacionado ao seu  $W$ . Por exemplo, a observação  $T_{231}$  é relacionada a  $X_{31}$  e esta a  $W_1$ . De modo geral, considere a Tabela 2.1.

Tabela 2.1 – Relações entre variáveis em uma estrutura hierárquica

Variável indivíduo ( $T_{IJK}$ )	Variável do grupo de nível 2 ( $X_{JK}$ )	Variável do grupo de nível 3 ( $W_J$ )	Variável resposta
$T_{111}$	$X_{11}$	$W_1$	$Y_{111}$
$T_{211}$	$X_{11}$	$W_1$	$Y_{211}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$T_{I11}$	$X_{11}$	$W_1$	$Y_{I11}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$T_{1J1}$	$X_{J1}$	$W_1$	$Y_{1J1}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$T_{IJ1}$	$X_{J1}$	$W_1$	$Y_{IJ1}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
$T_{IJK}$	$X_{JK}$	$W_K$	$Y_{IJK}$

Fonte: Adaptado de (JR; FÁVERO, 2019).

Nas três primeiras colunas estão apresentadas o conjunto amostral de cada uma das variáveis explicativas, e na quarta coluna, as observações em relação a variável de interesse ( $Y_{ijk}$ ), que é explicada pelas variáveis precedentes. Note que, nas linhas, uma relação pode ser feita, da esquerda (nível inferior) para direita (nível superior). Esta é a mesma relação ilustrada na figura 2.1.

### 2.3 Formalização do Modelo Multinível

Os métodos estatísticos, principalmente os modelos lineares, são utilizados em diversas áreas do conhecimento, tais como indústria, educação e agropecuária. Estes têm como propósito otimizar a etapa de planejamento e analisar os dados resultantes de um experimento (RENCHEER, 2000). Para cada situação se faz uso de uma técnica específica e, no caso de dados com uma estrutura hierárquica, o modelo multinível é adequado, pois este considera os níveis das variáveis explicativas, fornecendo mais informação para explicar a variável resposta (JR; FÁVERO, 2019).

O modelo linear hierárquico (MLH) ou multinível, é considerado uma extensão do modelo linear padrão que permite especificar diferentes níveis utilizando modelos individuais, que são combinados em apenas um único modelo (RENCHEER, 2000). Desta forma, a variável resposta será explicada considerando variáveis preditoras de todos os níveis significantes. Assim,

é considerado a dependência entre as variáveis explicativas o que, conseqüentemente, melhora a estimativa da variável resposta.

Devido a grande eminência de dados com estruturas hierárquicas em empresas e instituições, o uso de MLH vêm apresentando maior destaque. Este modelo geralmente é utilizado em pesquisas sociais, como em escolas, mas aqui, no presente trabalho, ele é utilizado como suporte para análise de dados resultantes de um experimento agrário envolvendo café, no qual são considerados 3 níveis.

Como os níveis dessa estrutura hierárquica podem ser descritos por modelos lineares individuais, estes apresentam variáveis dependentes (resposta) e independentes (explicativas), coeficientes (parâmetros) e um erro experimental associado a amostra. A variável de interesse ou a variável resposta, é medida em um nível individual, enquanto as variáveis explicativas seguem a estrutura hierárquica. De modo geral, o modelo combinado pode ser visto como um sistema hierárquico de equações de regressão (PINHEIRO, 2005).

Há duas versões para o modelo multinível, uma no qual a variação é considerada no intercepto, e a outra considera também na inclinação. Estas versões são definidas nas subseções a seguir.

### 2.3.1 Equação multinível de três níveis com intercepto aleatório e inclinação fixa

Considerando  $i = 1, \dots, I$  referente às observações no nível inferior,  $j = 1, \dots, J$  às do grupo de segundo nível e  $k = 1, \dots, K$  às de terceiro, define-se os modelos destes níveis.

O modelo de nível 1 é utilizado para descrever este nível e tem a seguinte forma

$$Y_{ijk} = \pi_{0jk} + \pi_{1jk}T_{ijk} + e_{ijk}, \quad (2.1)$$

em que  $Y_{ijk}$  é a variável resposta;  $T_{ijk}$  é o preditor ou variável explicativa;  $\pi_{0jk}$  e  $\pi_{1jk}$  os coeficiente de regressão associados ao preditor  $T_{ijk}$ ; e  $e_{ijk}$  o erro aleatório associado às amostras de nível 1, cuja distribuição é normal, com média zero e variância constante ( $\sigma^2$ ). Este modelo segue a forma de um modelo de regressão usual no qual a resposta está sendo corrigida em relação a média da variável explicativa.

O modelo de nível 2 tem como finalidade atribuir ao modelo combinado a influência das amostras deste nível. Tal atribuição é feita a partir da estimação dos parâmetros do modelo de nível 1 ( $\pi_{0jk}$  e  $\pi_{1jk}$ ) ser feita baseando-se nas componentes de nível 2. Desta forma, os modelos de nível 2 podem ser escritos como

$$\begin{aligned}\pi_{0jk} &= \beta_{00k} + \beta_{01}X_{jk} + r_{0jk}; \\ \pi_{1jk} &= \beta_{10k}.\end{aligned}\tag{2.2}$$

em que  $\beta_{00k}$  e  $\beta_{10k}$  são os interceptos de nível 2;  $X_{jk}$  é o preditor ou variável explicativa de nível 2;  $\beta_{01k}$  e  $\beta_{11k}$  os coeficientes de regressão associados ao preditor  $X_{jk}$ ; e  $r_{0jk}$  e  $r_{1jk}$  são os efeitos aleatórios associados ao modelo de nível 2.

Estes modelos funcionam de forma semelhante a (2.1). Os parâmetros  $\beta_{00k}$  e  $\beta_{10k}$  são variáveis dependentes, o preditor  $X_{jk}$  é a variável independente e  $\gamma_{00}$ ,  $\gamma_{01}$ ,  $\gamma_{10}$  e  $\gamma_{11}$  são os parâmetros de nível 2.

Por conseguinte, é feito para modelos de nível maior, inclusive para o nível 3. Logo, o modelo deste nível é composto por

$$\begin{aligned}\beta_{00} &= \gamma_{000} + \gamma_{001}W_{ijk} + u_{00k}; \\ \beta_{01} &= \gamma_{100}; \\ \beta_{10} &= \gamma_{100}.\end{aligned}\tag{2.3}$$

Os interceptos são descritos por  $\gamma_{000}$ ,  $\gamma_{100}$  e  $\gamma_{100}$ , assim como a inclinação por  $\gamma_{001}$ ,  $W_{ijk}$  é a variável preditora de nível 3 e  $u_{00k}$  o efeito aleatório deste nível.

O modelo misto pode ser obtido por substituição algébrica dos parâmetros pelas equações que os explicam que, neste caso, são 2.2 e 2.3. Assim, substituindo 2.2 em 2.1, obtém-se o modelo misto de segundo nível, dado por

$$Y_{ijk} = \beta_{00k} + \beta_{10k}T_{ijk} + \beta_{01k}X_{ijk} + e_{ijk} + r_{0jk}.\tag{2.4}$$

De forma análoga, substituindo 2.3 em 2.4 obtém-se o modelo misto de nível 3, com intercepto aleatório e inclinação fixa, expresso por

$$Y_{ijk} = \gamma_{000} + \gamma_{100}T_{ijk} + \gamma_{100}X_{ijk} + \gamma_{001}W_{ijk} + e_{ijk} + r_{0jk} + u_{00k}.\tag{2.5}$$

### 2.3.2 Equação multinível de três níveis com intercepto e inclinação aleatórios

De forma análoga ao feito na subseção anterior, é feito neste. Os componentes que compõem os modelos a seguir têm a mesma interpretação, mas neste é necessário adicionar alguns outros para explicar a variabilidade da inclinação. O modelo de nível 1 é dado pela

mesma equação do modelo com apenas o intercepto aleatório, isto é, pela equação 2.1:

$$Y_{ij} = \pi_{0jk} + \pi_{1jk}T_{ijk} + e_{ijk}.$$

O modelo de nível 2 tem a finalidade de explicar as observações deste nível, considerando o confundimento entre os efeitos das variáveis explicativas  $T_{ijk}$  e  $X_{jk}$ . Reescrevendo os parâmetros  $\pi_{0jk}$  e  $\pi_{1jk}$  em função de componentes de segundo nível, tem-se

$$\begin{aligned}\pi_{0jk} &= \beta_{00k} + \beta_{01k}X_{ijk} + r_{0jk}; \\ \pi_{1jk} &= \beta_{10k} + \beta_{11k}X_{ijk} + r_{1jk}.\end{aligned}\tag{2.6}$$

Nas equações 2.6, há mais componentes que descrevem o parâmetro de nível 1,  $\pi_{1jk}$ , do que em relação a 2.2, como  $\beta_{11k}$  e  $r_{1jk}$ , inclinação e efeito aleatório de segundo nível em relação a  $\pi_{11k}$ .

Por conseguinte, obtém-se o modelo de terceiro nível por

$$\begin{aligned}\beta_{00k} &= \gamma_{000} + \gamma_{001}W_{ijk} + u_{00k} \\ \beta_{10k} &= \gamma_{100} + \gamma_{101}W_{ijk} + u_{10k} \\ \beta_{01k} &= \gamma_{010} + \gamma_{011}W_{ijk} + u_{01k} \\ \beta_{11k} &= \gamma_{110} + \gamma_{111}W_{ijk} + u_{11k}\end{aligned}\tag{2.7}$$

em que as inclinações  $\gamma_{101}$ ,  $\gamma_{011}$  e  $\gamma_{111}$  e os efeitos aleatórios  $u_{10k}$ ,  $u_{01k}$  e  $u_{11k}$  de terceiro níveis são adicionados em relação ao modelo 2.3.

O modelo misto pode ser obtido por substituição algébrica dos parâmetros pelas equações que os explicam que, neste caso, são as equações 2.6 e 2.7. Assim, substituindo a equação 2.6 na 2.1, obtém-se o modelo misto de segundo nível, dado por

$$Y_{ij} = \beta_{00} + \beta_{10}T_{ij} + \beta_{01}X_{ij} + \beta_{11}T_{ij}X_{ij} + r_{1j}T_{ij} + r_{0j} + e_{ij}.\tag{2.8}$$

De forma análoga, substituindo a equação 2.7 na 2.8, obtém-se o modelo misto de nível 3, com intercepto e inclinação aleatórios, expresso por

$$\begin{aligned}Y_{ijk} &= \gamma_{000} + \gamma_{100} \cdot T_{ijk} + \gamma_{010} \cdot X_{ijk} + \gamma_{001} \cdot W_{ijk} + \gamma_{110} \cdot T_{ijk} \cdot X_{ijk} + \gamma_{101} \cdot T_{ijk} \cdot W_{ijk} + \\ &\gamma_{011} \cdot X_{ijk} \cdot W_{ijk} + \gamma_{111} \cdot T_{ijk} \cdot X_{ijk} \cdot W_{ijk} + r_{1jk} \cdot T_{ijk} + u_{10k} \cdot X_{ijk} + u_{11k} \cdot T_{ijk} \cdot X_{ijk} + \\ &e_{ijk} + r_{0jk} + u_{00k}.\end{aligned}\tag{2.9}$$

### 2.3.3 Suposições

As suposições expostas a seguir são necessárias para que o MLH forneça uma análise confiável. Com o que diz respeito ao modelo de nível 1, assume-se que seu erro associado a amostra ( $e_{ijk}$ ) segue distribuição normal, com média zero e variância constante  $\sigma^2$ , isto é:  $E(e_{ijk}) = 0$  e  $var(\varepsilon_{ijk}) = \sigma^2$ .

No modelo de nível 2, assume-se que os parâmetros  $\beta_{0j}$  e  $\beta_{1j}$  seguem distribuição normal multivariada, com médias  $\gamma_{00}$  e  $\gamma_{10}$  e variâncias  $\tau_{00}$  e  $\tau_{11}$ . Além disso, a variância destes parâmetros e dos componentes referentes ao erro no modelo de nível 2 ( $v_{0j}$  e  $v_{1j}$ ) são iguais a observação correspondente. Por fim, assume-se que há independência entre efeitos aleatórios de níveis diferentes. Em símbolos, tem-se:

$$E(\beta_{0j}) = \gamma_{00}, var(\beta_{0j}) = var(v_{0j}) = \tau_{00};$$

$$E(\beta_{1j}) = \gamma_{10}, var(\beta_{1j}) = var(v_{1j}) = \tau_{11};$$

$$cov(\beta_{0j}, \beta_{1j}) = cov(v_{00}, v_{11}) = \tau_{01};$$

$$cov(e_{ij}, v_{0j}) = cov(e_{ij}, v_{1j}) = 0.$$

Sobre as componentes de natureza aleatória de ambos os níveis,  $e_{ij}$  e  $v_{ij}$ , seguem distribuição normal com média 0 e variância  $R$  e  $G$ , respectivamente, em que:

$$R = \sigma^2 \cdot I_{nj}$$

e

$$G = \begin{bmatrix} \gamma_{00} & \gamma_{01} \\ \gamma_{01} & \gamma_{11} \end{bmatrix}.$$

Para categorizar os diferentes tipos de modelos, considere a tabela 2.2, proposta por Sullivan(1999). Há diferença de estimação entre estes modelos, sendo necessário comparar os resultados de ambos para seleção do estimador mais adequado para o banco de dados analisado. A seguir, apresentam-se duas tabelas, nas quais são expostos os modelos definidos na seção 2.3. Na tabela 2.2, encontram-se os modelos de cada nível separadamente e em 2.3, os modelos combinados, assim como seus respectivos componentes de efeito fixo e aleatório.

Tabela 2.2 – Categorização dos Modelos Lineares Hierárquicos de até 3 níveis

Tipo de modelo	Modelo de nível 1	Modelo de nível 2	Modelo de nível 3
Modelo Multinível com intercepto aleatório e inclinação fixa	$Y_{ijk} = \pi_{0jk} + \pi_{1jk}T_{ijk} + e_{ijk}$	$\pi_{0jk} = \beta_{00k} + \beta_{01}X_{jk} + r_{0jk}; \pi_{1jk} = \beta_{10k}$	$\beta_{00} = \gamma_{000} + \gamma_{001}W_{ijk} + u_{00k}; \beta_{01} = \gamma_{100}; \beta_{10} = \gamma_{10}$
Modelo Multinível com inclinação e intercepto aleatórios	$Y_{ij} = \pi_{0jk} + \pi_{1jk}T_{ijk} + e_{ijk}$	$\pi_{0jk} = \beta_{00k} + \beta_{01k}X_{ijk} + r_{0jk}; \pi_{1jk} = \beta_{10k} + \beta_{11k}X_{ijk} + r_{1jk}$	$\beta_{00k} = \gamma_{000} + \gamma_{001}W_{ijk} + u_{00k}; \beta_{10k} = \gamma_{100} + \gamma_{101}W_{ijk} + u_{10k}; \beta_{01k} = \gamma_{010} + \gamma_{011}W_{ijk} + u_{01k}; \beta_{11k} = \gamma_{110} + \gamma_{111}W_{ijk} + u_{11k}$

Fonte: Adaptado de Sullivan(1999).

Tabela 2.3 – Modelos Lineares Hierárquicos de 3 níveis (combinado)

Tipo de modelo	Modelo combinados	Efeitos fixos	Efeitos aleatórios
Modelo Multinível com intercepto aleatório e inclinação fixa	$Y_{ijk} = \gamma_{000} + \gamma_{100}T_{ijk} + \gamma_{010}X_{ijk} + \gamma_{001}W_{ijk} + e_{ijk} + r_{0jk} + u_{00k}$	$\gamma_{000}; \gamma_{100}; \gamma_{010}; \gamma_{001}$	$e_{ijk}; r_{0jk}; u_{00k}$
Modelo Multinível com inclinação e intercepto aleatórios	$Y_{ijk} = \gamma_{000} + \gamma_{100} \cdot T_{ijk} + \gamma_{010} \cdot X_{ijk} + \gamma_{001} \cdot W_{ijk} + \gamma_{110} \cdot T_{ijk} \cdot X_{ijk} + \gamma_{101} \cdot T_{ijk} \cdot W_{ijk} + \gamma_{011} \cdot X_{ijk} \cdot W_{ijk} + \gamma_{111} \cdot T_{ijk} \cdot X_{ijk} \cdot W_{ijk} + r_{1jk} \cdot T_{ijk} + u_{10k} \cdot X_{ijk} + u_{11k} \cdot T_{ijk} \cdot X_{ijk} + e_{ijk} + r_{0jk} + u_{00k}$	$\gamma_{000}; \gamma_{100}; \gamma_{010}; \gamma_{001}; \gamma_{110}; \gamma_{101}; \gamma_{011}; \gamma_{111}$	$e_{ijk}; r_{0jk}; r_{1jk}; u_{00k}; u_{10k}; u_{11k}$

Fonte: Adaptado de Sullivan(1999).

Considerando o modelo multinível com intercepto aleatório e inclinação fixa, percebe-se que seus efeitos aleatórios ( $e_{ijk}$ ,  $r_{0jk}$ , e  $u_{00k}$ ) são gerais de cada nível da estrutura hierárquica. Entretanto, para o modelo com inclinação e intercepto aleatório, seus efeitos aleatórios são combinados também com as variáveis preditoras. Isto apresenta o efeito aleatória entre estas variáveis. Por exemplo, o produto  $u_{10k} \cdot X_{ijk}$  representa o efeito aleatório da taxa de variação da variável preditora de nível 1 ( $T_{ijk}$ ) sob o valor base da variável de nível 2 ( $X_{jk}$ ).

### 2.3.4 Representação Matricial do Modelo

Na estatística é comum representar modelos com matrizes por sua praticidade. Em MLH não é diferente. Dito isto, segue a representação matricial dos componentes do modelo de nível

1 dado pela equação (2.1):  $Y_j = \begin{bmatrix} Y_{1j} \\ Y_{2j} \\ \vdots \\ Y_{nj} \end{bmatrix}$ ,  $X_j = \begin{bmatrix} 1 & X_{1j} \\ 1 & X_{2j} \\ \vdots & \vdots \\ 1 & X_{nj} \end{bmatrix}$ ,  $\beta_j = \begin{bmatrix} \beta_{0j} \\ \beta_{1j} \end{bmatrix}$  e  $\varepsilon_j = \begin{bmatrix} \varepsilon_{1j} \\ \varepsilon_{2j} \\ \vdots \\ \varepsilon_{nj} \end{bmatrix}$ , em que

$Y_j$  é o vetor resposta,  $X_j$  é a matriz de entrada (cujo componentes são conhecidos) relacionados aos dados de nível 1,  $\beta_j$  é o vetor dos parâmetros (intercepto e inclinação) e  $\varepsilon_j$  o vetor de erros aleatórios associado ao modelo.

De forma análoga, o modelo de nível 2 que escreve os parâmetros  $\beta_{0j}$  e  $\beta_{1j}$  em função das variáveis  $\gamma$ ,  $v$  e  $W$ , que é dado pela equação (2.2), é descrito matricialmente por:  $\beta_j =$

$$\begin{bmatrix} \beta_{0j} \\ \beta_{1j} \end{bmatrix}, W_j = \begin{bmatrix} 1 & W_j & 0 & 0 \\ 0 & 0 & 1 & W_j \end{bmatrix}, \gamma_j = \begin{bmatrix} \gamma_{00} \\ \gamma_{01} \\ \gamma_{10} \\ \gamma_{11} \end{bmatrix} \text{ e } v_j = \begin{bmatrix} v_{0j} \\ v_{1j} \end{bmatrix}, \text{ em que } W_j \text{ é a matriz de entrada}$$

(semelhante a matriz  $X_j$  descrita anteriormente) referente aos dados de nível 2,  $\gamma_j$  o vetor de efeitos fixos e  $v_j$  o vetor de efeitos aleatórios.

Daí vem que o modelo combinado dos dois níveis pode ser escrito por:

$$Y_j = X_j \cdot (W_j \cdot \gamma + v_j) + \varepsilon_j$$

$$\Rightarrow \begin{bmatrix} Y_{1j} \\ Y_{2j} \\ \vdots \\ Y_{nj} \end{bmatrix} = \begin{bmatrix} 1 & X_{1j} \\ 1 & X_{2j} \\ \vdots & \vdots \\ 1 & X_{nj} \end{bmatrix} \cdot \left( \begin{bmatrix} 1 & W_j & 0 & 0 \\ 0 & 0 & 1 & W_j \end{bmatrix} \cdot \begin{bmatrix} \gamma_{00} \\ \gamma_{01} \\ \gamma_{10} \\ \gamma_{11} \end{bmatrix} + \begin{bmatrix} v_{0j} \\ v_{1j} \end{bmatrix} \right) + \begin{bmatrix} \varepsilon_{1j} \\ \varepsilon_{2j} \\ \vdots \\ \varepsilon_{nj} \end{bmatrix}.$$

Resolvendo o lado esquerdo da equação matricial anterior, tem-se:

$$\begin{bmatrix} Y_{1j} \\ Y_{2j} \\ \vdots \\ Y_{nj} \end{bmatrix} = \begin{bmatrix} 1 & X_{1j} \\ 1 & X_{2j} \\ \vdots & \vdots \\ 1 & X_{nj} \end{bmatrix} \cdot \left( \begin{bmatrix} \gamma_{00} + W_j \cdot \gamma_{01} \\ \gamma_{10} + W_j \cdot \gamma_{11} \end{bmatrix} + \begin{bmatrix} v_{0j} \\ v_{1j} \end{bmatrix} \right) + \begin{bmatrix} \varepsilon_{1j} \\ \varepsilon_{2j} \\ \vdots \\ \varepsilon_{nj} \end{bmatrix}$$

$$\begin{aligned}
\Rightarrow \begin{bmatrix} Y_{1j} \\ Y_{2j} \\ \vdots \\ Y_{nj} \end{bmatrix} &= \begin{bmatrix} 1 & X_{1j} \\ 1 & X_{2j} \\ \vdots & \vdots \\ 1 & X_{nj} \end{bmatrix} \cdot \begin{bmatrix} \gamma_{00} + W_j \cdot \gamma_{01} + \nu_{0j} \\ \gamma_{10} + W_j \cdot \gamma_{11} + \nu_{1j} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1j} \\ \varepsilon_{2j} \\ \vdots \\ \varepsilon_{nj} \end{bmatrix} \\
\Rightarrow \begin{bmatrix} Y_{1j} \\ Y_{2j} \\ \vdots \\ Y_{nj} \end{bmatrix} &= \begin{bmatrix} \gamma_{00} + W_j \cdot \gamma_{01} + \nu_{0j} + X_{1j} \cdot (\gamma_{10} + W_j \cdot \gamma_{11} + \nu_{1j}) \\ \gamma_{00} + W_j \cdot \gamma_{01} + \nu_{0j} + X_{2j} \cdot (\gamma_{10} + W_j \cdot \gamma_{11} + \nu_{1j}) \\ \vdots \\ \gamma_{00} + W_j \cdot \gamma_{01} + \nu_{0j} + X_{nj} \cdot (\gamma_{10} + W_j \cdot \gamma_{11} + \nu_{1j}) \end{bmatrix} + \begin{bmatrix} \varepsilon_{1j} \\ \varepsilon_{2j} \\ \vdots \\ \varepsilon_{nj} \end{bmatrix} \\
\text{Portanto} \begin{bmatrix} Y_{1j} \\ Y_{2j} \\ \vdots \\ Y_{nj} \end{bmatrix} &= \begin{bmatrix} \gamma_{00} + W_j \cdot \gamma_{01} + \nu_{0j} + X_{1j} \cdot (\gamma_{10} + W_j \cdot \gamma_{11} + \nu_{1j}) + \varepsilon_{1j} \\ \gamma_{00} + W_j \cdot \gamma_{01} + \nu_{0j} + X_{2j} \cdot (\gamma_{10} + W_j \cdot \gamma_{11} + \nu_{1j}) + \varepsilon_{2j} \\ \vdots \\ \gamma_{00} + W_j \cdot \gamma_{01} + \nu_{0j} + X_{nj} \cdot (\gamma_{10} + W_j \cdot \gamma_{11} + \nu_{1j}) + \varepsilon_{nj} \end{bmatrix}.
\end{aligned}$$

Agora o vetor referente as observações  $Y_{nj}$  está escrito em função de um único vetor de ordem  $nx1$ , cuja cada uma de suas entradas está em função de componentes de nível 1 e nível 2. Nesta igualdade de matrizes, obtêm-se cada uma das observações  $Y_{nj}$  desde que se tenha a linha correspondente. Desta forma, igualando cada entrada da matriz resposta com a matriz que contém os componentes que explicam o modelo, obtêm-se um sistema de equações. Este conjunto é dado por:

$$Y_{1j} = \gamma_{00} + W_j \cdot \gamma_{01} + \nu_{0j} + X_{1j} \cdot (\gamma_{10} + W_j \cdot \gamma_{11} + \nu_{1j}) + \varepsilon_{1j}$$

$$Y_{2j} = \gamma_{00} + W_j \cdot \gamma_{01} + \nu_{0j} + X_{2j} \cdot (\gamma_{10} + W_j \cdot \gamma_{11} + \nu_{1j}) + \varepsilon_{2j}$$

$$\vdots$$

$$Y_{nj} = \gamma_{00} + W_j \cdot \gamma_{01} + \nu_{0j} + X_{nj} \cdot (\gamma_{10} + W_j \cdot \gamma_{11} + \nu_{1j}) + \varepsilon_{1j}.$$

Como os modelos têm diferentes tipos de parâmetros (aleatórios ou não), é necessário fazer uso de diversas técnicas de estimação. Já o parâmetro  $\beta_{0j}$  deste mesmo modelo é escrito por:  $\beta_{0j} = \gamma_{00} + \gamma_{01} \cdot W_j + \nu_{0j}$ . Sabe-se que  $\gamma_{00}$  e  $\gamma_{01}$  são efeitos fixos e  $W_j$  é a variável independente, entretanto,  $\nu_{0j}$  é um efeito aleatório, logo  $\beta_{0j}$  é de natureza aleatória. Portanto, para estimar os parâmetros  $\beta_{0j}$  e  $\beta_{1j}$  é necessário fazer uso de técnicas diferentes, pois têm naturezas distintas.

Antes de apresentar as técnicas de estimação dos parâmetros fixos e aleatórios, é necessário definir uma matriz. Seja  $A_j$  uma matriz dada pelo produto das matrizes de entrada de ambos os níveis, isto é :

$$A_j = X_j \cdot W_j$$

$$\Rightarrow A_j = \begin{bmatrix} 1 & X_{1j} \\ 1 & X_{2j} \\ \vdots & \vdots \\ 1 & X_{nj} \end{bmatrix} \cdot \begin{bmatrix} 1 & W_j & 0 & 0 \\ 0 & 0 & 1 & W_j \end{bmatrix}$$

$$\Rightarrow A_j = \begin{bmatrix} 1 & W_j & X_{1j} & W_j \cdot X_{1j} \\ 1 & W_j & X_{2j} & W_j \cdot X_{2j} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & W_j & X_{nj} & W_j \cdot X_{nj} \end{bmatrix}$$

Perceba que, assim como  $X_{nj}$  e  $W_j$ , a matriz  $A_j$  também é conhecida e sua ordem é  $n \times 4$  (pois  $X_{nj}$  tem ordem  $n \times 2$  e  $W_j$   $2 \times 4$ ).

### 2.3.5 Estimando parâmetros fixos $\gamma$

Os componentes de efeito fixo ( $\gamma$ ) podem ser estimados fazendo uso do Método de Mínimos Quadrados Ponderado ou Método de Mínimos Quadrados Generalizado. Aplicando-os, obtêm-se:

$$\hat{\gamma} = (A^T \cdot \hat{V} \cdot A)^{-1} \cdot A^T \cdot \hat{V}^{-1} \cdot Y \quad (2.10)$$

em que

$$V = \text{var}(Y_{ij}) = X_{nj} \cdot G \cdot X_{nj}^T + R_j. \quad (2.11)$$

As variâncias  $G$  e  $R_j$  descritas em (2.11), podem ser estimadas por máxima verossimilhança.

A variância do estimador  $\hat{\gamma}$  é dada por:

$$\text{var}(\hat{\gamma}) = (A^T \cdot \hat{V}^{-1} \cdot A)^{-1}. \quad (2.12)$$

Entretanto, quando a variância entre as componentes de nível 2 não são homogêneas, é recomendado usar a seguinte variância de  $\hat{y}$ :

$$\text{var}(\hat{y}) = (A^T \cdot \hat{V}^{-1} \cdot A)^{-1} \cdot A^T \cdot \hat{V}^{-1} \cdot (Y - A \cdot \hat{y})^T \cdot \hat{V}^{-1} \cdot A \cdot (A^T \cdot V^{-1} \cdot A)^{-1}. \quad (2.13)$$

### 2.3.6 Estimando componentes de covariância ( $G$ e $R_j$ )

É imprescindível estimar os parâmetros de covariância  $G$  e  $R_j$ , pois  $V$  depende destes (veja (2.11)). Estes parâmetros são estimados por máxima verossimilhança. Este processo, é feito a partir da maximização da seguinte função logaritmica:

$$l_{ML}(G, R) = -\frac{1}{2} \log|V| - \frac{N}{2} \log r^T V^{-1} r - \frac{N}{2} \left[ 1 + \log \frac{2\pi}{N} \right] \quad (2.14)$$

em que  $p$  é o  $\text{rank}(A)$  (ou  $\text{posto}(A)$ ) e  $r$  explicado por

$$r = Y - A \cdot (A^T \cdot V^{-1} \cdot A)^T A^T V^{-1} \cdot Y.$$

### 2.3.7 Estimando efeitos aleatórios $v$

Para estimar os efeitos aleatórios ( $v$ ) de nível 2, faz-se uso dos estimadores de encolhimento. O grau de encolhimento depende da variância das componentes do nível 2. Desta forma, quanto menor o número de unidades de nível 1 na unidade de nível 2, mais próximo de  $\hat{\gamma}_0$  a estimativa de  $v$  estará. Mas, quanto maior o número destas mesmas unidades, mais próximo da média  $\bar{Y}_j$  estará  $v$ .

A estimativa dos efeitos aleatórios  $v$  é dada por

$$\hat{v} = \hat{G} \cdot X^T \cdot \hat{V}^{-1} \cdot (Y - A \cdot \hat{y}). \quad (2.15)$$

## 2.4 Modelo de Regressão Adaptativo

### 2.4.1 Regressões Robustas

Assumindo o modelo de regressão linear dado por  $Y = X\beta + E$  e tendo como objetivo obter a estimativa do vetor de parâmetros, representado por  $\hat{\beta}$ , em geral, utiliza-se o método de Mínimos Quadrados Ordinários, com o propósito de minimizar a soma de quadrados residual.

Assim, dada uma amostra  $y_1, \dots, y_n$ , tem-se o resíduo  $e_i = y_i - \hat{y}_i$ , portanto, enuncia-se a estimativa de mínimos quadrados por

$$\hat{\beta}_{ols} = \operatorname{argmin}_{\beta} \sum_{i=1}^n (e_i)^2. \quad (2.16)$$

Convém ressaltar que, há situações em que observações outliers poderão estar presentes na amostra, as estimativas  $\hat{\beta}_{ols}$  poderão ser usadas, logo, como alternativa a esta questão, torna-se apropriado o uso de métodos robustos, em aspectos mais teóricos, tratando o problema de regressão em um espaço  $L_p$ . Desta forma, enuncia-se

$$\hat{\beta}_{DA} = \operatorname{argmin}_{\beta} \sum_{i=1}^n |y_i - \hat{y}_{DA}|^p, \quad (2.17)$$

em que  $\hat{y}_{DA}$  correspondem a  $i$ -ésima observação predita por algum estimador robusto de regressão que utiliza como princípio minimizar os resíduos absolutos, justificando a notação em subscrito.

A recomendação do procedimento de minimização baseados em normas- $L_p$  para um problema de regressão, é justificada pela distribuição dos resíduos estar caracterizada por elevada assimetria ou excesso de curtose, denotando uma distribuição com caudas mais pesadas.

Convém ressaltar que os métodos robustos são limitados a proporção de outliers contida na amostra, com propriedades particulares na especificação do ponto de ruptura. Diante do exposto, uma alternativa a este problema foi proposta por (MAITY; SHERMAN, 2008) sugerindo as regressões adaptativas.

Para uma melhor compreensão desta nova abordagem, assume-se a notação  $\hat{\beta}_{Lp}$  como estimativas obtidas por algum procedimento robusto, definido por

$$\hat{\beta}_{Lp} = \operatorname{MIN}_{\beta} \sum_{i=1}^n |y_i - \hat{y}_{AD}|^p. \quad (2.18)$$

Naturalmente, para  $p = 1$  temos que  $\hat{\beta}_{Lp} = \hat{\beta}_{ols}$ , porém, para outras métricas a escola de  $p$  em problema de regressão, conforme Sposito (1990), deve-se considerar a distribuição dos resíduos, e relação ao efeito de cauda. Logo, (MAITY; SHERMAN, 2008) propuseram um procedimento denominado como regressão adaptativa, no qual as estimativas de  $\beta$  são obtidas em 2.19

$$\hat{\beta}_{ADPT} = w\hat{\beta}_{ols} + (1 - w)\hat{\beta}_{Lp}, \quad (2.19)$$

em  $w$  é definido por  $w > (1 - w)$  para que  $\hat{\beta}_{ADPT} \cong \hat{\beta}_{ols}$ , ou seja, a distribuição dos resíduos apresenta um comportamento com cauda menos prolongada. Para  $w < (1 - w)$ , então  $\hat{\beta}_{ADPT} \cong \hat{\beta}_{ols}$ , isto é, para tal situação a distribuição dos resíduos apresenta um elevado excesso de curtose.

A conexão de  $w$ , a princípio especificado arbitrariamente com o coeficiente de curtose, dado por  $k$  remete ao cálculo da curtose média  $\bar{k}$  obtido por

$$h_{2(ols)} = \frac{1}{n} \sum_{i=1}^n \left[ \frac{e_{i(ols)} - \bar{e}_{i(ols)}}{s} \right]^4 - 3 \quad (2.20)$$

e

$$h_{2(Lp)} = \frac{1}{n} \sum_{i=1}^n \left[ \frac{e_{i(ols)} - \bar{e}_{i(Lp)}}{s} \right]^4 - 3. \quad (2.21)$$

Assim,  $\bar{k} = \frac{h_{2(ols)} - h_{2(Lp)}}{2}$ . Desta forma

$$w(k) = \begin{cases} 1, se & k \leq 3, \\ \frac{3}{k}, se & k > 3. \end{cases}$$

A exemplificar, assume-se  $w(k) \leq 3$ , então  $\hat{\beta}_{ALR} = w\hat{\beta}_{ols} + (1 - w)\hat{\beta}_{Lp} = \hat{\beta}_{ols}$ , logo  $\hat{\beta}_{ALR} = \hat{\beta}_{ols}$ . Note que, quanto maior o efeito de cauda, maior será o coeficiente de curtose,  $h_{2(Lp)}$ , implicando em menor  $\bar{k}$ , e por consequência, um menor  $w$ .

## 2.5 Estimadores Robustos

Em uma análise estatística, principalmente quando há várias variáveis regressoras, é comum que haja valores extremos (outliers). Estes podem comprometer o estimador, tornando os resultados pouco representativos. A decisão simplista de excluir estes valores pode comprometer a capacidade da generalização dos resultados (FARCOMENI; VENTURA, 2010).

Uma outra alternativa para solucionar eventuais problemas com os dados é a transformação dos mesmos. Todavia, isto implicaria na mudança de escala dos valores, podendo dificultar ainda mais a análise (CHAGAS, 2011).

A regressão múltipla é uma das ferramentas mais utilizadas na estatística aplicada. Entretanto, esta é excepcionalmente suscetível a outliers. Por esta razão, é preferível métodos muito robustos, por serem resistentes a estes valores (TORTI et al., 2012).

Os métodos robustos são uma alternativa válida para analisar conjuntos propícios a outliers, já que os mesmos são limitados a uma fração de valores extremos contidos no conjunto

em questão (CHAGAS, 2011). Estes métodos proporcionam uma análise mais precisa do conjunto, como a identificação dos valores extremos. Os estimadores "resistentes", são essenciais para a inferência robusta, assim o desenvolvimento deste tipo de estimador se tornou o objetivo de muitos pesquisadores.

Neste trabalho, foram considerados dois estimadores: Least Median Squares (LMS) e Least Trimmed Squares (LTS). Antes de prosseguir com estes estimadores, é pertinente apresentar algumas definições que validam e qualificam um método robusto.

### 2.5.1 Ponto de Ruptura

É definido pela quantidade máxima de pontos discrepantes que um estimador robusto pode suportar, de tal forma que, ainda possa fornecer informações confiáveis sobre o parâmetro em questão. Este é um dos meios para se medir a robustez de uma estatística (FARCOMENI; VENTURA, 2010).

Seja  $Z = (X_{11}, \dots, X_{1k}, Y_1), \dots, (X_{n1}, \dots, X_{nk}, Y_n)$  um conjunto de dados e  $T$  um estimador de regressão. Quando  $T$  é aplicado à amostra  $Z$ , tem-se como resultado um vetor de parâmetros  $\hat{\beta}$ , isto é,  $T(Z) = \hat{\beta}$ , em que  $\hat{\beta} = (\beta_1, \dots, \beta_k)$ .

Defindo  $Z^*$  como um conjunto cuja composição seja feita a partir de  $m$  ( $1 \leq m \leq n$ ) substituições de valores originais por valores possivelmente discrepantes, pode-se definir uma fração de contaminação. Logo, como  $Z^*$  tem  $n$  elementos, a fração de contaminação pode ser escrita por  $\delta = \frac{m}{n}$ . É interessante salientar que há inúmeros conjuntos  $Z^*$  que possam ser definidos desta forma.

O vício, isto é, a diferença entre o valor esperado e o valor verdadeiro, pode ser expresso como  $b(m; T, Z)$ . Desta forma, o vício máximo que uma contaminação  $\delta$  pode causar a amostra, é dada por

$$b(m; T, Z) = \sup \|T(Z^*) - T(Z)\|. \quad (2.22)$$

Se o vício for finito, os valores extremos têm um efeito arbitrário no estimador  $T$ , que por sua vez, perde a robustez a partir de uma certa quantidade destes. Isto é conhecido como a ruptura de  $T$ , ou melhor, o ponto de ruptura. Em símbolos este é definido por

$$\delta_n^*(T, Z) = \min(b(m; T, Z)) = \infty. \quad (2.23)$$

### 2.5.2 Função Influência

A função de influência (do inglês, Influence Function, sigla IF) mede o efeito que uma amostra infinitesimalmente contaminada infere no estimador  $T$ . Isso permite que o pesquisador tenha ideia de como o estimador reagirá à contaminações pontuais. A seguir, tem-se a definição matemática da IF, segundo (CHAGAS, 2011).

Sejam  $F$ ,  $G$  e  $H$  distribuições e  $T(G)$  um estimador de regressão. Considere a relação

$$G = (1 - \delta) \cdot F + \delta \cdot H \quad (2.24)$$

em que  $0 \leq \delta \leq 1$ . Neste caso, a distribuição  $G$  está sendo definida por uma distribuição principal  $F$  e por uma distribuição de contaminação  $H$ , com probabilidade  $\delta$ .

Pela equação 2.24, pode-se afirmar que o estimador  $G$  será robusto caso se mantenha estável independente dos valores que a componente  $\delta \cdot H$  assumir. Um caso específico de  $G$  ocorre quando sempre há contaminação.

Seja  $\Delta_h$  uma distribuição na qual o valor  $h$  tem ocorrência certa. Desta forma,  $X$  tem distribuição  $\Delta_h$  e ainda, tem-se que  $P(X \leq x) = 0$ , quando  $x < h$ , com  $E(X) = h$ . Isso resulta na distribuição  $G_{h,\delta}$  que, por sua vez, pode simplificar uma amostra contaminada por uma proporção  $\delta$  em  $h$ . Em símbolos, tem-se

$$G_{h,\delta} = (1 - \delta)F + \delta\Delta_h. \quad (2.25)$$

Quando  $h$  ocorre com probabilidade  $\delta$ , a função  $G_{h,\delta}$  expõe a reação de  $F$  numa função ou estimador. A influência relativa de um estimador  $T(F)$  quando  $h$  ocorre pode ser escrito por  $\frac{T(G_{h,\delta}) - T(F)}{\delta}$ .

Se considerar que a contaminação por  $h$  tenda a zero, isto é, considerando um  $\delta$  tão pequeno quanto queira, pode-se definir a função influência. Portanto, uma IF mensura a influência relativa que uma contaminação infinitesimal  $h$  causa em um estimador  $T(F)$  e é descrita por

$$IF = \lim_{\delta \rightarrow 0} \frac{T(G_{h,\delta}) - T(F)}{\delta}. \quad (2.26)$$

### 2.5.3 Conjuntos Elementares

Um algoritmo foi proposto na década de 50, por Theil (1950) e consiste em dividir o conjunto amostral em subamostras de tamanho  $k$ . Estas subamostras devem ser grande o bastante

para que todos os parâmetros considerados no modelo possam ser estimados (MACHADO, 1997).

Este é um método computacional que tem como objetivo estimar a média e a matriz de covariância com alto ponto de ruptura. Desta forma, o algoritmo supracitado pode ser aplicado em métodos resistentes à observações discrepantes, isto é, métodos robustos (CHAGAS, 2011).

Seja o conjunto  $Z_{n \times (k+p)} = (X_{n \times k} \ Y_{n \times p})$ , em que  $X_{n \times k}$  e  $Y_{n \times p}$  são as matrizes das variáveis predictoras e resposta, respectivamente. Particionando  $Z_{n \times (k+p)}$  tem-se

$$Z_{n \times (k+p)} = \begin{pmatrix} X_J & Y_J \\ X_{n-J} & Y_{n-J} \end{pmatrix}, \quad (2.27)$$

em que  $X_J$  e  $Y_J$  são submatrizes que representam as subamostras que foram citadas anteriormente.

Considere  $S$  um conjunto cuja composição é feita a partir de todos os possíveis subconjuntos da amostra. Desta forma,  $S$  pode ser definido por  $\{\{1, \dots, k\}, \dots, \{j_1, \dots, j_k\}, \dots, \{(n - k + p), \dots, n\}\}$ , em que  $J = \{j_1, \dots, j_k\}$ . Para obter-se os subconjuntos, é necessário fixar  $k$  e  $n$ .

Cada subconjunto  $Z_J$  é definido da seguinte forma

$$Z_J = \begin{pmatrix} X_{J_1 1} & \cdots & X_{J_1 k} & Y_{J_1 1} & \cdots & Y_{J_1 p} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ X_{J_k 1} & \cdots & X_{J_k k} & Y_{J_k 1} & \cdots & Y_{J_k p} \end{pmatrix}. \quad (2.28)$$

Cada um dos subconjuntos ( $Z_J$ ) é denominado conjunto elementar (CHAGAS, 2011).

#### 2.5.4 Regressão de Mínimos Quadrados Medianos (MQM)

Quando se trata da estimação de parâmetros por meio de um modelo de regressão, muitos estimadores visam minimizar o quadrado dos resíduos para obter um melhor resultado. O método dos mínimos quadrados é muito utilizado por ser simples e prático, entretanto, sua vulnerabilidade em relação a valores extremos é um problema. Tendo isto em vista, SIEGEL (1982) propôs um estimador que se constitui em aplicar a mediana repetidas vezes junto a um ponto de ruptura (ROUSSEEUW, 1984).

Considerando o modelo de regressão  $y_i = x_{i1}\theta_1 + \dots + x_{ip}\theta_p + e_i$ , tem-se em mente que o objetivo é obter  $\hat{\theta}$ , em que  $\theta = (\theta_1, \dots, \theta_p)^T$  é o vetor de estimativas a partir dos dados, que podem ser expressos por  $(x_{i1}, \dots, x_{ip}, y_i)$ .

Para qualquer número de parâmetros  $p^*$ , determina-se um vetor de parâmetros exclusivos cuja a  $j$ -ésima coordenada pode ser expressa por  $\theta(i_1, \dots, i_p)$ . Assim, a mediana repetida desta coordenada é definida como  $\hat{\theta}_j = med_{i_1}(\dots (med_{p-1}(med_p \theta_j(i_1, \dots, i_p))) \dots)$ .

Neste ponto, tendo em vista o menor resíduo possível, propõe-se minimizar a mediana dos quadrados dos resíduos, isto é *minimize med  $r_i^2$* , em que  $r_i$  pode ser expresso por  $y_i - x_{i1}\hat{\theta}_1 \dots x_{ip}\hat{\theta}_p$ . Essa proposta é essencialmente baseada numa ideia de (HAMPEL, 1975).

O ponto de ruptura  $h$  de um estimador LMS pode ser expresso por

$$h = \frac{n + p^* + 1}{2} \quad (2.29)$$

(ROUSSEEUW; DRIESSEN, 2006). Isto quer dizer que pode-se substituir até uma quantidade  $h$  de valores das observações por valores habitários sem que prejudique o estimador. Sem dúvidas, este é um ponto forte para que um estimador LMS seja utilizado.

### 2.5.5 Regressão de Mínimos Quadrados Aparados (MQA)

Este estimador tem uma taxa de convergência maior do que o LMS, isto é, quanto maior for a amostra, maior será a diferença entre estes. A taxa de convergência do estimador LMS é  $n^{-\frac{1}{3}}$ , enquanto para o LTS é  $n^{-\frac{1}{2}}$  (ROUSSEEUW, 1984).

O parâmetro  $\beta_{LTS}$  é estimado diretamente por quadrados mínimos. A soma de quadrados, por sua vez é calculada utilizando um grupo restrito, as amostras centrais. Estas são delimitadas por um subconjunto  $H$  de tamanho  $h = \frac{n+p+1}{2}$ . Desta forma,  $\hat{\beta}_{LTS}$  é estimado por

$$\text{minimize } SS_T\{\hat{\beta}(h)\} = \sum_{i=1}^h e_i^2\{\hat{\beta}(h)\}, \quad (2.30)$$

em que  $\hat{\beta}(h)$  é estimado por mínimos quadrados.

A variância é baseada na soma de quadrados dos resíduos supracitada, porém é necessário dimensioná-la já que 2.30 é aplicada em uma porção da amostra. Torti et al. (2012) ajusta

a variação da distribuição normal truncada, referente a parte  $\frac{h}{n}$  da seguinte forma

$$\sigma_T^2(h) = 1 - \frac{2n}{2h} \Phi^{-1} \left( \frac{n+h}{2n} \right) \phi \left\{ \Phi^{-1} \left( \frac{n+h}{2n} \right) \right\}, \quad (2.31)$$

em que  $\phi$  é a densidade normal padrão e  $\Phi^{-1}$  a função de distribuição acumulada.

A variância estimada em função os parâmetros do LTS é dada por  $\sigma_{LTS}^2 = \frac{SS_T}{\beta_{LTS}(h) \times \sigma_T^2(h)}$ .

## 2.6 Coeficiente de Correlação Intraclasse (CCI)

O Coeficiente de Correlação Intraclasse (CCI) é uma ferramenta estatística utilizada para verificação da homogeneidade do conjunto de dados, sendo interpretado como uma razão de variâncias (SHROUT; FLEISS, 1979). Tratando-se de níveis ou camadas entre variáveis explicativas, o CCI pode ser aplicado para inferir a respeito da relevância de um nível para explicar a variável de interesse.

Este coeficiente é definido pelo quociente entre a variância do nível superior e a variância total, podendo ser expresso, no caso de 2 níveis, por

$$\rho = \frac{\sigma_2}{\sigma_1 + \sigma_2}, \quad (2.32)$$

em que  $|\rho| \leq 1$  é o CCI,  $\sigma_1$  a variância entre as observações de primeiro nível e  $\sigma_2$  a variância entre os grupos de segundo nível.

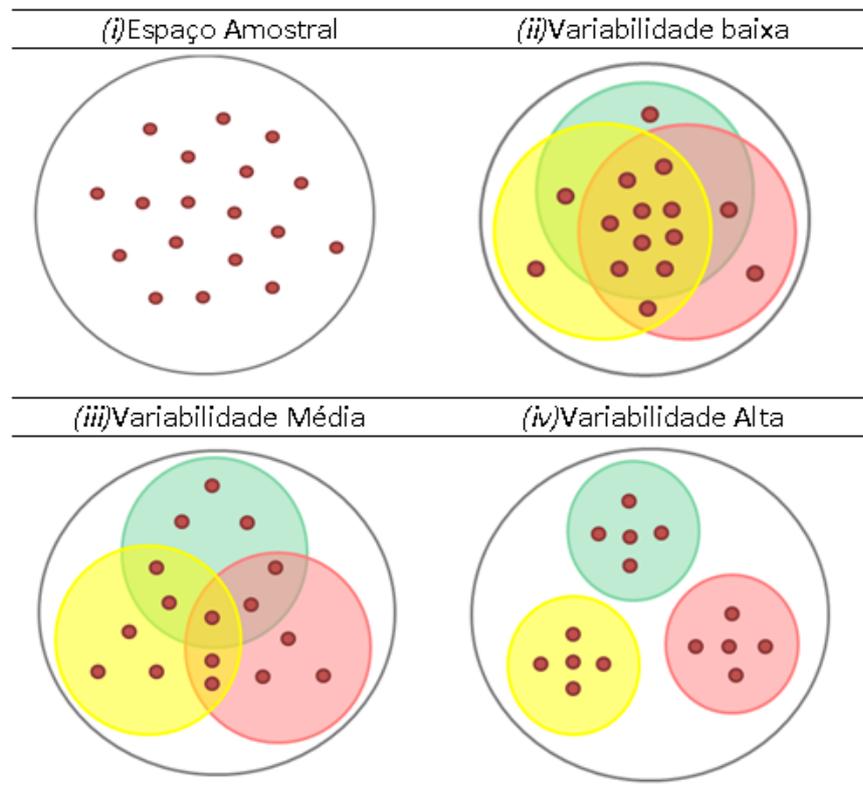
Observa-se que, quanto menor for a variância do nível 1 ( $\sigma_1$ ), implica em  $\rho \rightarrow 1$ . Por exemplo, supondo que  $\sigma_1$  seja infinitesimal, podendo ser escrito por  $\sigma_1 \approx 0$ , então  $\rho \approx \frac{\sigma_2}{\sigma_2+0} \approx \frac{\sigma_2}{\sigma_2}$ , isto é,  $\rho \approx 1$ .

Neste contexto, afirma-se que, quanto menor a variabilidade dos dados do nível inferior, maior é sua subordinação em relação ao nível superior. Por outro lado, se a variabilidade do nível inferior for relativamente alta,  $\rho$  será menor, indicando que a variável de nível superior não influencia tanto a inferior.

Twisk (2006) ilustra e discute sobre situações nas quais a variância entre os grupos de segundo nível são diferentes. A figura 2.2 apresenta algumas situações hipotéticas em relação ao conjunto amostral, salientando dois níveis.

Na Figura 2.2, tem-se quatro possíveis situações. Primeiramente, em (i) apresenta um espaço amostral comum, sem níveis. Entretanto, em (ii), (iii) e (iv), encontram-se três subcon-

Figura 2.2 – Esquema de dispersão amostral entre níveis.



Fonte: Autoria própria.

juntos (ou grupos) que contêm as informações observadas. Esses grupos podem ser interpretados como as observações de segundo nível.

Perceba que os grupos representados na Figura (ii) têm uma grande quantidade de observações de nível 1 em comum (denotadas pelas interseções) fazendo com que a variabilidade entre estas observações contidas em cada grupo não seja alta. Assim, não faz sentido considerar um nível superior, já que a diferença entre seus grupos não é expressiva.

Analogamente, na Figura (iii), o mesmo acontece, mas com menos intensidade, visto que a interseção entre os grupos é menor do que o visto em *ii*. Neste ponto, pode ser que faça sentido considerar um segundo nível, já que seus grupos contêm informações diferentes.

Por fim, na Figura (iv), percebe-se que não há observações de nível 1 em comum entre os grupos, isto é, cada grupo fornece informações diferentes do outro, viabilizando a definição de um segundo nível. Este pensamento se estende para casos que apresentem mais de 2 níveis.

Portanto, o coeficiente de correlação intraclasses pode ser utilizado como critério para seleção de níveis em um delineamento multinível. O quão próximo de 1 o coeficiente  $\rho$  deve

estar para que se admita um nível é uma questão subjetiva, isto é, o pesquisador decide em considera-lo ou não.

## 2.7 Máxima Verossimilhança Restrita (REML)

Desenvolvido apenas para casos isolados por Anderson e Bancroft (1952) e Russell e Bradley (1958), a máxima verossimilhança restrita (sigla REML, em inglês) foi aperfeiçoada por outros estudiosos, como Thompson (1962) para dados balanceados e por Patterson e R. Thompson (1971) para modelos mistos generalizados.

De modo geral, a estimativa REML se baseia em estimar os componentes de variância considerando os resíduos após a parte fixa do modelo ser ajustada por mínimos quadrados ordinários. Assim, o REML não considera o vetor de observações, mas sim uma combinação linear dos elementos deste, de modo que não haja nenhum efeito fixo (SEARLE; CASELLA; MCCULLOCH, 1992).

As combinações lineares consideradas entre as observações são equivalentes aos resíduos obtidos após o ajuste dos efeitos fixos. Desta forma, não há perda de combinação e a análise é feita sobre observações equivalentes às originais.

Neste sentido, considere uma amostra aleatória  $(x_1, x_2, \dots, x_n)$ , distribuída de forma idêntica e independente, em que segue distribuição normal cuja esperança seja  $\mu$  e variância  $\sigma^2$ . E ainda, seja  $y$  o vetor de observações originais e  $k'y$  o vetor referente as combinações lineares dos elementos de  $y$ , em que  $k'$  é escolhido de tal forma que satisfaça a equação  $K'y = k'X\beta + k'Z$  no qual  $\beta$  e  $Z$  são as matrizes dos componentes fixos e aleatórios, respectivamente. Todas as entradas de  $\beta$  devem ser nulas, isto é,  $K'X\beta = 0$ .

A aplicação da máxima verossimilhança restrita é feita de forma computacional, utilizando *softwares* apropriados para a análise estatística de dados.

### 3 MATERIAL E MÉTODOS

Em concordância com os objetivos propostos, a estrutura metodológica utilizada neste trabalho é descrita nas seções a seguir: descrição dos experimentos e fatores (3.1), procedimento de análise (3.2), confundimento dos efeitos e adaptação das regressões multiníveis (3.3), construção do modelo multinível adaptativo (3.4).

#### 3.1 Estrutura dos dados

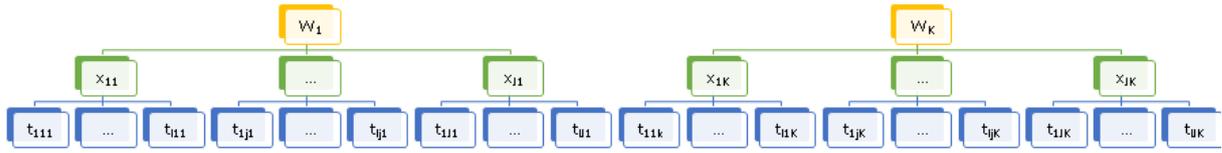
Os dados são provenientes de um experimento que foi realizado na Universidade Federal de Lavras, mais especificamente no laboratório de sementes, cujo objetivo foi de comparar cafés especiais de diferentes pontuações de qualidade, armazenado em ambientes com diferentes temperaturas e períodos de tempo. Utilizou-se cafés cereja despulpado armazenados da forma íntegra e descascados, com 3 pontuações de qualidades diferentes.

Foram considerados quatro fatores: Benefício; Qualidade inicial; Temperatura; Época. Estes são detalhados a seguir.

- Benefício: se refere à forma como os cafés foram armazenados, se íntegros (com a presença do pergaminho), ou beneficiados (sem o pergaminho; descascado);
- Qualidade inicial: pontuação dos cafés no início do experimento (perfil), na avaliação sensorial, realizada de acordo com o protocolo de degustação do café, da Specialty Coffee Association of America - SCAA (LINGLE, 2011);
- Temperatura: temperatura do ar no ambiente de armazenagem dos cafés, controlada em câmara fria a  $10^{\circ}\text{C}$  ou em ambiente sem controle de temperatura, com temperatura média de  $25^{\circ}\text{C}$ ;
- Época: se refere aos tempos de tomada de amostras para a avaliação da qualidade sensorial, durante os doze meses de armazenamento nos dois ambientes; 1 - zero meses; 2 - três meses, 3 - seis meses, 4 - nove meses, e 5 - doze meses.

Para cada tipo de café, o modelo multinível, seguindo três hierarquias, foi ajustado. Para fim de ilustração, considere a figura 3.1, em que  $t_{ijk}$  refere-se ao fator época (nível 1),  $x_{jk}$  à temperatura (nível 2) e  $w_k$  à qualidade (nível 3).

Figura 3.1 – Esquema de hierarquia de um MLH de 3 níveis.



Fonte: Autoria própria.

### 3.2 Simulação do modelo nas condições experimentais utilizados na refrigeração de cafés especiais

Seguindo a estrutura do delineamento hierárquico (3.1), ajustou-se o modelo multinível para os dados reais, assumindo intercepto e inclinação aleatória. Desta forma, a resposta referente a nota sensorial ( $y_{ijk}$ ) foi obtida em função de duas componentes  $\eta_{ijk}$  e  $\gamma_{ijk}$ , referindo-se respectivamente aos termos com efeitos fixos e aleatórios.

$$\hat{y}_{ijk} = \hat{\eta}_{ijk} + \hat{\gamma}_{ijk} \quad (3.1)$$

$$\hat{\eta}_{ijk} = \hat{\beta}_0 + \hat{\beta}_1 Q \quad (3.2)$$

$$\hat{\gamma}_{ijk} = \hat{\beta}_{0ij}^* + \hat{Z}_{ij} Q + \hat{Z}_{ij}^* T \quad (3.3)$$

Na equação (3.2), tem-se o intercepto e o coeficiente de inclinação referente às notas iniciais dadas aos níveis do fator qualidade ( $Q$ ). De forma análoga, em (3.3) o intercepto aleatório  $\beta_{0ij}$  e os coeficientes de inclinação aleatórios  $Z_{ij}$  e  $Z_{ij}^*$ , estes estão associados aos níveis dos fatores qualidade ( $Q$ ) e tempo ( $T$ ), respectivamente.

Com estas especificações, após o ajuste do modelo, considerou-se as estimativas obtidas como valores paramétricos do experimento a ser simulado via Monte Carlo, de modo que, a respostas simuladas fossem coerentes com a condições experimentais especificadas na refrigeração de cafés especiais.

Contudo, mantendo a hipótese de que as avaliações pudessem ser susceptíveis a erros de mensuração em 5 unidades a mais ou a menos em relação a nota real observada, incorporou-se uma quantidade de valores discrepantes, controlados no processo pela equação (3.4), através da especificação arbitrária pelo parâmetro  $\alpha$ , interpretado como probabilidade de mistura, fixados em 0, 10, 0, 20 e 0, 30. Desta forma, obtém-se

$$y_{ijk}^* = u(\alpha) + (1 - \alpha)u, \quad (3.4)$$

em que  $u \sim U[0, 1]$  e  $0 \leq \alpha \leq 1$ .

Logo, adotou-se a seguinte regra

$$\begin{cases} u > \alpha \rightarrow y_{ijk}^* = y_{ijk} - 5 \\ u < \alpha \rightarrow y_{ijk}^* = y_{ijk}. \end{cases}$$

Decorrente ao fato de que, na amostra simulada foi incorporada a presença de outliers, ajustou-se as regressões LMS e LTS assumindo efeito fixo entre as variáveis. Desta forma, as regressões ajustadas foram representadas nas equações (3.5) e (3.6).

$$y_{ijk}^* = \hat{\beta}_{0(LMS)} + \hat{\beta}_{1(LMS)}Q, \quad (3.5)$$

$$y_{ijk}^* = \hat{\beta}_{0(LTS)} + \hat{\beta}_{1(LTS)}Q. \quad (3.6)$$

### 3.3 Incorporação das regressões adaptativas

As regressões adaptativas no contexto deste trabalho, foram implementadas pela combinação de coeficientes aleatórios: intercepto e inclinação, estimados entre as regressões multi-níveis ajustadas pelo método de máxima verossimilhança restrita, com o acrônimo de (REML) e robustas (LMS e LTS) para os efeitos fixos, conforme encontram-se descritos pelas equações (3.7) e (3.8).

Entretanto,  $w$  originalmente interpretado como um parâmetro de perturbação dos resíduos, foi estimado através de uma inspeção gráfica de valores atribuídos arbitrariamente em um intervalo de (0,1) em função do erro de predição obtido com o ajuste do modelo completo (.).

Nestas condições, duas relações foram realizadas e descritas a seguir

$$\hat{\beta}_{0(ADPT)} = w\hat{\beta}_{0(REML)} + (1 - w)\hat{\beta}_{0(.)}, \quad (3.7)$$

$$\hat{\beta}_{1(ADPT)} = w\hat{\beta}_{1(REML)} + (1 - w)\hat{\beta}_{1(.)}. \quad (3.8)$$

Em ambos os casos, a regra para seleção dos componentes foi definida pela geração de um número aleatório  $u$ , de modo que, se  $u \geq w$ , então  $\beta_{0(ADPT)} = \beta_{0(.)}$ , caso contrário,

$\hat{\beta}_{0(ADPT)} = \hat{\beta}_{0(REML)}$ , sendo,  $\hat{\beta}_{0(\cdot)}$  e  $\hat{\beta}_{1(\cdot)}$ , respectivamente, as estimativas dos interceptos e coeficiente de inclinação obtida pelos métodos robustos LMS e LTS.

Desta forma, ao considerar o confundimento das estimativas robustas com o intercepto aleatório, o modelo multinível com regressões adaptativas foi apresentado na equação (3.9), no qual os termos  $\hat{\beta}_{0(ADPT)}$  e  $\hat{\beta}_{0}E_i$  referem-se aos efeitos aleatórios.

$$y_{ijk}^* = \hat{\beta}_{0(ADPT)} + \hat{\beta}_{0(T_i/Aleat.)} + \hat{\beta}_{0(Fixo)} + \hat{\beta}_{1(Fixo)}Q_i + (\hat{\beta}_{1(ADPT)} + \hat{\beta}_{1(Q_i/Aleat.)}T_i). \quad (3.9)$$

Por fim, para a realização deste procedimento foram elaboradas funções e um script para entrada de dados e realização da simulação no software R.

## 4 RESULTADOS

Considerando os dados reais sobre o beneficiamento e não beneficiamento do café, denominados por cafés íntegros, estimou-se a correlação intraclasse e foi inspecionado o valor de  $w$  na regressão adaptativa em função do erro experimental ao considerar erros de mensuração em 5 e 10 unidades. A seguir, tem-se as análises referentes a dois modelos (modelo com intercepto aleatório e inclinação fixa e modelo intercepto aleatório e inclinação aleatória).

### 4.1 Simulação do modelo considerando erro de mensuração nas avaliações das notas sensoriais especificado em 5 unidades

Tabela 4.1 – Estimativas dos coeficientes de correlação intraclasse (em %) considerando a variável nota sensorial e um erro de mensuração de 5 unidades.

	Época	Temperatura	Qualidade
<b>Modelos com intercepto aleatório e inclinação fixa</b>			
Nível 2 (dados:"beneficiado")	56,90	43,09	-
Nível 2 (dados:"íntegro")	63,20	36,79	-
Nível 3 (dados:"beneficiado")	32,94	40,99	26,05
Nível 3 (dados:"íntegro")	43,21	33,54	23,24
<b>Modelos com intercepto fixo e inclinação aleatória</b>			
Nível 2 (dados:"beneficiado")	55,11	44,88	-
Nível 2 (dados:"íntegro")	61,94	38,05	-
Nível 3 (dados: "beneficiado")	19,94	3,34	72,03
Nível 3 (dados: "íntegro")	31,12	29,40	38,47

Fonte: Autoria própria.

Os coeficientes expostos na tabela 4.1, evidenciam que ao considerar a estrutura com três níveis, as avaliações das notas sensoriais que mais corroboram para justificar o modelo multinível é considerado com três níveis, com maior ênfase aos dados do cafés beneficiados (72,03%), tendo como referência a estimativa do coeficiente de correlação intraclasse, portanto, há evidências de que as notas foram homogêneas em relação a uma mesma hierarquia, porém heterogêneas entre elas.

Neste sentido, notou-se que quanto maior a correlação intra-Qualidade, maior dependência entre as respostas das notas sensoriais, logo, a construção de um modelo que respeite a estrutura de agregação dos dados é plausível de ser considerada. Analogamente, segue-se as mesmas observações em relação ao café íntegro, entretanto, dada a baixa estimativa do coeficiente de correlação intraclasse (38,47%), este modelo não garante ganhos expressivos de informação.

A partir dos resultados para os coeficientes de correlação intraclasse, foram efetuadas as estimativas dos coeficientes aleatórios e fixos, descritos nas tabelas 4.2 e 4.3, respectivamente, considerando a estrutura hierárquica e as respostas das avaliações sensoriais, utilizadas como valores paramétricos para estudos de simulação que envolvem a construção de regressões adaptativas.

Tabela 4.2 – Estimativas para os efeitos aleatórios considerando um erro de mensuração de 5 unidades.

Fator	Coeficiente	Variância	Desvio-padrão	Correção
Época	Intercepto	2,341	1,530066	
	Inclinação	0,000009082	0,003014	0,84
Temperatura	Intercepto	3,613	1,900889	
	Inclinação	0,0001631	0,012773	-1
Residual		2,02	1,424601	

Fonte: Autoria própria.

Tabela 4.3 – Estimativas para os efeitos fixos considerando um erro de mensuração de 5 unidades.

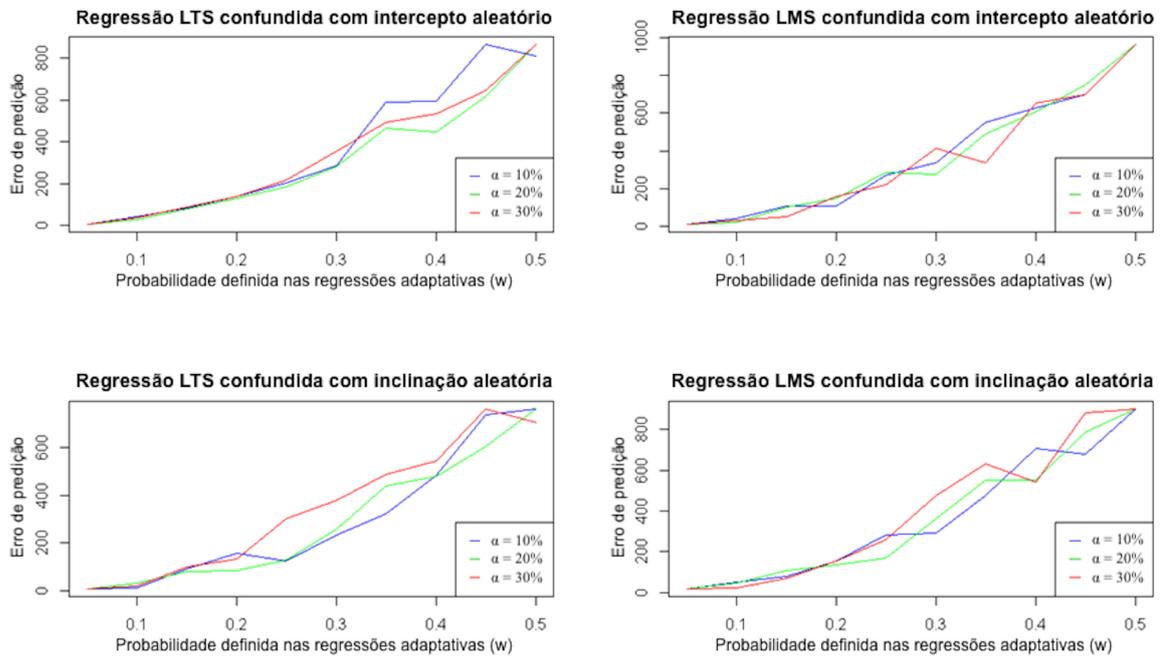
Coeficiente	Estimativa	Desvio-padrão residual	Valor $t$	Correlação
Intercepto	21,36362	6,23305	3,427	
Inclinação	0,71851	0,07089	10,135	-0,988

Fonte: Autoria própria.

Os resultados ilustrados nas figuras 4.1 e 4.2, permitiram concluir que para ambas formas de confundimento das estimativas robustas LMS e LTS, tratadas neste trabalho identificaram sequência de valores de  $w$  que conduz a implementação das regressões adaptativas no modelo multinível com intercepto e inclinação aleatória, de tal forma que, para  $w \leq 0,10$ , o erro de predição do modelo é minimizado. Convém ressaltar que resultado tem forte impacto, se considerando uma taxa de contaminação de outliers em valores elevados, como  $\alpha = 0,30$ .

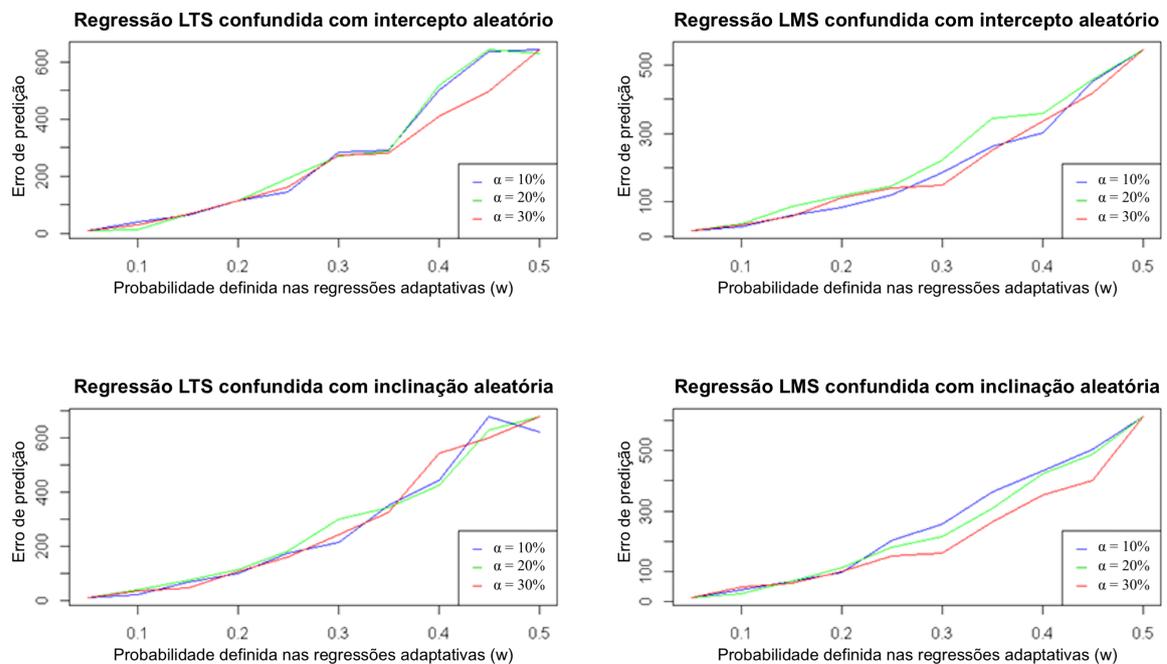
No tocante da discriminação do método ser LMS ou LTS, os resultados simulados considerando um erro de mensuração de 5 unidades, indicaram que em ambas abordagens, o comportamento das curvas em relação a inspeção de  $w$  e o erro experimental similares, com pequenas oscilações atribuídas ao erro de Monte Carlo.

Figura 4.1 – Desempenho de regressões LTS e LMS em cafés beneficiados, considerando erro de mensuração em 5 unidades.



Fonte: Autoria própria.

Figura 4.2 – Desempenho de regressões LTS e LMS em cafés não beneficiados, considerando erro de mensuração em 5 unidades



Fonte: Autoria própria.

## 4.2 Simulação do modelo considerando erro de mensuração nas avaliações das notas sensoriais especificado em 10 unidades

Considerando um erro de mensuração de 10 unidades, foi estimado os coeficientes de correlação intraclasse para dois e três níveis, conforme exposto na tabela 4.4.

Tabela 4.4 – Estimativas dos coeficientes de correlação intraclasse (em %) considerando a variável nota sensorial e um erro de mensuração em 10 unidades.

	Época	Temperatura	Qualidade
<b>Modelos com intercepto aleatório e inclinação fixa</b>			
Nível 2 (dados:"beneficiado")	61,13	38,87	-
Nível 2 (dados:"íntegro")	59,05	40,95	-
Nível 3 (dados:"beneficiado")	37,25	23,67	39,08
Nível 3 (dados:"íntegro")	35,80	24,83	39,37
<b>Modelos com intercepto e inclinação aleatória</b>			
Nível 2 (dados:"beneficiado")	4,49	95,51	-
Nível 2 (dados:"íntegro")	42,69	57,31	-
Nível 3 (dados: "beneficiado")	20,74	42,98	36,28
Nível 3 (dados: "íntegro")	30,88	37,95	30,99

Não foi obtido um valor expressivo para justificar uma abordagem com três níveis, entretanto, para dois níveis, tem-se 95,51% para os dados que se referem aos cafés beneficiados, utilizando um modelo com intercepto e inclinação aleatórios. Ainda para os dados beneficiados, mas utilizando o modelo com intercepto aleatório e inclinação fixa, obteve-se 61,13%.

Uma abordagem com três níveis não forneceria muita informação além de uma estrutura com dois níveis, visto que os coeficientes de correlação intraclasse correspondentes aos modelos de três níveis não foram maiores do que 40%.

Nas tabelas 4.5 e 4.6 estão expostos as estimativas dos efeitos aleatórios e fixos, respectivamente, considerando um modelo com intercepto e inclinação aleatórios com três níveis.

Tabela 4.5 – Estimativas para os efeitos aleatórios considerando um erro de mensuração em 10 unidades.

Fator	Coeficiente	Variância	Desvio-padrão	Correção
Época	Intercepto	4.9508	2.2250	
	Inclinação	0,0000959	0,009793	0,84
Temperatura	Intercepto	3.7058	1.9250	
	Inclinação	0,0001656	0,01286	-1
Residual		3.2405	1.8001	

Fonte: Autoria própria.

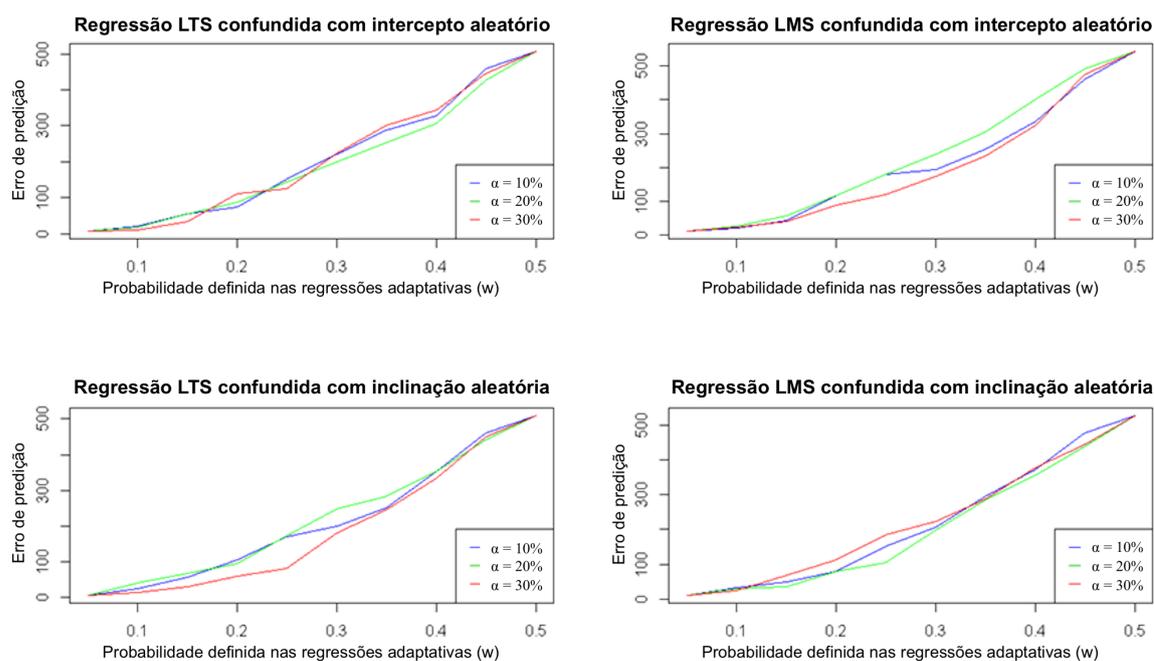
Tabela 4.6 – Estimativas para os efeitos fixos considerando um erro de mensuração em 10 unidades.

Coefficiente	Estimativa	Desvio-padrão residual	Valor $t$	Correlação
Intercepto	4.5857	10.7408	0.427	
Inclinação	0.8140	0.1244	6.545	-0.997

Fonte: Autoria própria.

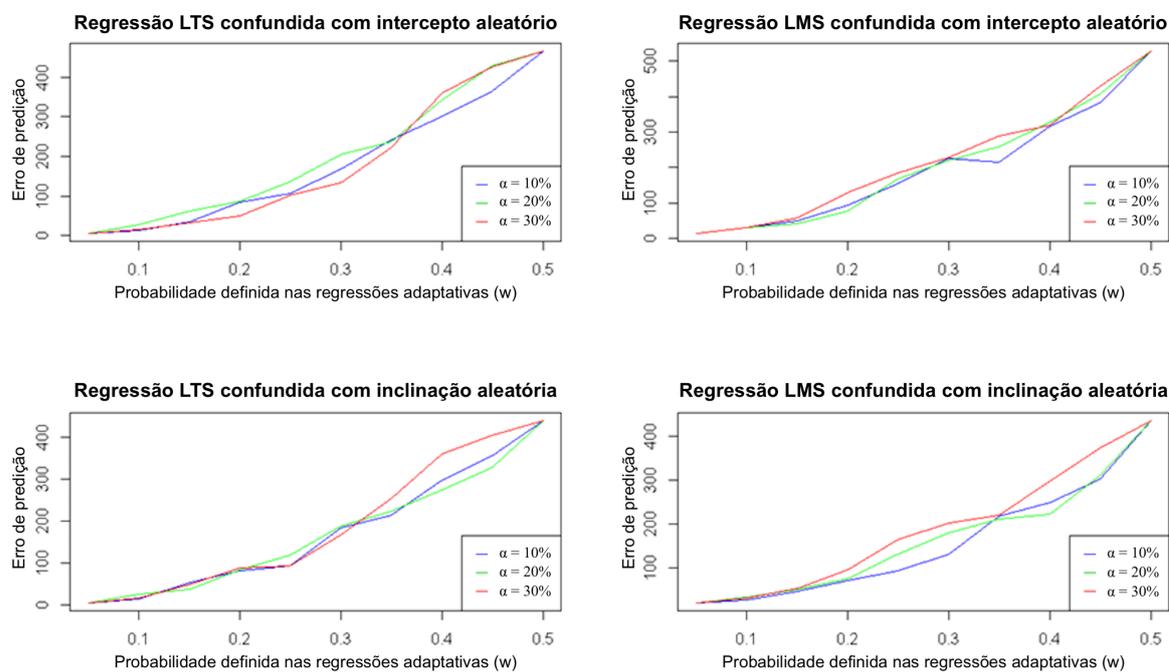
Considerando os dados simulados com erro de mensuração em 10 unidades, percebe-se a partir das figuras 4.3 e 4.4 que, independente da proporção de contaminação de outliers nos dados, o erro experimental não apresenta variação expressiva. E ainda, valores baixos para o erro experimental estão relacionados a menores valores de  $w$  ( $\leq 0,10$ ).

Figura 4.3 – Desempenho de regressões LTS e LMS em cafés beneficiados, considerando erro de mensuração em 10 unidades



Fonte: Autoria própria.

Figura 4.4 – Desempenho de regressões LTS e LMS em cafés não beneficiados, considerando erro de mensuração em 10 unidades



Fonte: Autoria própria.

Comparando os erros experimentais gerados pelos dados considerando um erro de mensuração em 5 e 10 unidades, constata-se que, para um erro de 10 unidades o erro experimental é menor em relação ao  $w$ . Isto indica que o método pode ser mais eficiente quando o erro de mensuração é maior.

## 5 CONCLUSÕES

Conclui-se que a regressão adaptativa aplicada aos métodos de estimação robustos (LMS e LTS) e ao estimador por máxima verossimilhança é viável quando se trata de amostras com proporção elevada de outliers (até 30%), possibilitando uma análise com erro experimental relativamente pequeno diante a parcela de contaminação.

Se tratando dos modelos, salienta-se que, para ambas as condições em que o café foi sujeito e erros de mensuração considerados, o intercepto de seus respectivos modelos não teve contribuição significativa, ao contrário de suas inclinações, exceto na estimação dos efeitos fixos referentes aos cafés não beneficiados, considerando um erro de mensuração de  $L = 10$  unidades, em que o intercepto e a inclinação apresentaram contribuição significativa no poder de predição do modelo.

Evidenciou-se ainda que, ao considerar a estrutura com três níveis, a avaliação das notas sensoriais que mais corroboram para justificar o modelo multinível tem maior ênfase aos dados do café beneficiado, considerando um erro de mensuração sensorial de até  $L = 5$  unidades (65, 15%). Além disso, quanto maior a correlação intra-Qualidade, maior a dependência entre as respostas das notas sensoriais, logo, a construção de um modelo que respeite a estrutura de agregação dos dados é plausível de ser considerada.

Portanto, considerando análises sensoriais de café cujo os avaliadores não têm treinamento, a regressão adaptativa aplicada aos métodos de estimação robustos (LMS e LTS) é apropriada, visto que esta proporciona uma análise com erro experimental relativamente pequeno, mesmo quando valores discrepantes constituem grande parte da amostra.

**APÊNDICE A – Equação multinível de três níveis com intercepto aleatório e inclinação fixa.**

- Modelo de nível 1: O modelo de nível 1 é dado por uma equação de regressão usual, como

$$Y_{ij} = \pi_{0j} + \pi_{1j}X_{ij} + e_{ij}. \quad (1)$$

- Modelo de nível 2: Escreve-se os parâmetros da equação 1 utilizando outros parâmetros para descrevê-lo, como mostrado a seguir

$$\begin{aligned} \pi_{0j} &= \beta_{00} + \beta_{01}W_{ij} + r_{0j}; \\ \pi_{1j} &= \beta_{10}. \end{aligned} \quad (2)$$

Substituindo 2 em 1, obtem-se

$$Y_{ij} = \beta_{00} + \beta_{01}W_{ij} + r_{0j} + \beta_{10}X_{ij} + e_{ij},$$

desta forma o modelo de nível 2 pode ser escrito como

$$Y_{ij} = \beta_{00} + \beta_{10}X_{ij} + \beta_{01}W_{ij} + e_{ij} + r_{0j}. \quad (3)$$

- Modelo de nível 3: De forma análoga ao feito anteriormente para obter o modelo de nível 2, será feito para obter o de nível 3. Reescrevendo os parâmetros  $\beta$  em 3, têm-se

$$\begin{aligned} \beta_{00} &= \gamma_{000} + \gamma_{001}T_{ijt} + u_{00t}; \\ \beta_{01} &= \gamma_{100}; \\ \beta_{10} &= \gamma_{010}. \end{aligned} \quad (4)$$

Substituindo as igualdades em 4 na equação 3, tem-se

$$Y_{ijt} = \gamma_{000} + \gamma_{001}T_{ijt} + u_{00t} + \gamma_{100}X_{ij} + \gamma_{010}W_{ij} + e_{ij} + r_{0j}.$$

Portanto, o modelo de nível 3 com intercepto aleatório de e inclinação fixa é dado por

$$Y_{ijt} = \gamma_{000} + \gamma_{100}X_{ij} + \gamma_{010}W_{ij} + \gamma_{001}T_{ijt} + e_{ij} + r_{0j} + u_{00t}. \quad (5)$$

**APÊNDICE B – Equação multinível de três níveis com intercepto e inclinação aleatório.**

- Modelo de nível 1: O modelo de nível 1 é dado por uma equação de regressão usual, como

$$Y_{ij} = \pi_{0j} + \pi_{1j}X_{ij} + e_{ij}. \quad (6)$$

- Modelo de nível 2: Reescrevendo os parâmetros de 6 como

$$\begin{aligned} \pi_{0j} &= \beta_{00} + \beta_{01}W_{ij} + r_{0j}; \\ \pi_{1j} &= \beta_{10} + \beta_{11}W_{ij} + r_{1j}. \end{aligned} \quad (7)$$

Substituindo 7 em 6, tem-se

$$Y_{ij} = \beta_{00} + \beta_{01}W_{ij} + r_{0j} + (\beta_{10} + \beta_{11}W_{ij} + r_{1j})X_{ij} + e_{ij}.$$

Logo o modelo de nível 2 é dado por

$$Y_{ij} = \beta_{00} + \beta_{10}X_{ij} + \beta_{01}W_{ij} + \beta_{11}X_{ij}W_{ij} + r_{1j}X_{ij} + r_{0j} + e_{ij}. \quad (8)$$

- Modelo de nível 3: Novamente, os parâmetros do modelo que explica o nível que precede este ( $\beta$ ) são escritos em função de outras componentes. Logo, tem-se

$$\begin{aligned} \beta_{00} &= \gamma_{000} + \gamma_{001}T_{ijt} + U_{00t} \\ \beta_{10} &= \gamma_{100} + \gamma_{101}T_{ijt} + U_{10t} \\ \beta_{01} &= \gamma_{010} + \gamma_{011}T_{ijt} + U_{01t} \\ \beta_{11} &= \gamma_{110} + \gamma_{111}T_{ijt} + U_{11t} \end{aligned} \quad (9)$$

Substituindo 9 em 8, obtem-se

$$\begin{aligned} Y_{ij} &= \gamma_{000} + \gamma_{001}T_{ijt} + U_{00t} + (\gamma_{100} + \gamma_{101}T_{ijt} + U_{10t})X_{ij} + (\gamma_{010} + \gamma_{011}T_{ijt} + U_{01t})W_{ij} + \\ & \quad (\gamma_{110} + \gamma_{111}T_{ijt} + U_{11t})X_{ij}W_{ij} + r_{1j}X_{ij} + r_{0j} + e_{ij}. \end{aligned} \quad (10)$$

Portanto, o modelo de nível 3 com intercepto e inclinação aleatórios é expresso por

$$\begin{aligned}
 Y_{ijt} = & \gamma_{000} + \gamma_{100} \cdot X_{ijt} + \gamma_{010} \cdot W_{ijt} + \gamma_{001} \cdot T_{ijt} + \gamma_{110} \cdot X_{ijt} \cdot W_{ijt} + \gamma_{101} \cdot X_{ijt} \cdot T_{ijt} + \\
 & \gamma_{011} \cdot W_{ijt} \cdot T_{ijt} + \gamma_{111} \cdot X_{ijt} \cdot W_{ijt} \cdot T_{ijt} + r_{1j} \cdot X_{ijt} + u_{10t} \cdot W_{ijt} + u_{11t} \cdot X_{ijt} \cdot W_{ijt} + \\
 & e_{ijt} + r_{0jt} + u_{00t}. \quad (11)
 \end{aligned}$$