

**UNIVERSIDADE FEDERAL DE ALFENAS**

**YACILED PAOLA MIRANDA LOPEZ**

**ABORDAGEM BAYESIANA PARA UM DELINEAMENTO *CROSSOVER*  $2 \times 2$  COM  
MEDIDAS REPETIDAS: UM ESTUDO DE SIMULAÇÃO E DADOS REAIS**

Alfenas/MG

2021

**YACILED PAOLA MIRANDA LOPEZ**

**ABORDAGEM BAYESIANA PARA UM DELINEAMENTO *CROSSOVER*  $2 \times 2$  COM  
MEDIDAS REPETIDAS: UM ESTUDO DE SIMULAÇÃO E DADOS REAIS**

Dissertação apresentada ao Programa de Pós-Graduação em Estatística Aplicada e Biometria, área de concentração em Estatística Aplicada e Biometria da Universidade Federal de Alfenas, MG, como parte dos requisitos para a obtenção do título de Mestre.

Linha de Pesquisa: Modelagem Estatística e Estatística Computacional.

Orientador: Prof. Dr. Denismar Alves Nogueira.

Coorientador: Prof. Dr. Luiz Alberto Beijo.

Alfenas/MG

2021

Dados Internacionais de Catalogação-na-Publicação (CIP)  
Sistema de Bibliotecas da Universidade Federal de Alfenas  
Biblioteca Central – Campus Sede

Miranda Lopez, Yaciled Paola  
M672a Abordagem bayesiana para um delineamento *Crossover*  $2 \times 2$  com  
medidas repetidas: um estudo de simulação e dados reais / Yaciled Paola  
Miranda Lopez– Alfenas, MG, 2021.  
75 f.: il. –

Orientador: Denismar Alves Nogueira.  
Dissertação (Mestrado em Estatística Aplicada e Biometria) – Universidade  
Federal de Alfenas, 2021.  
Bibliografia.

1. Efeitos *carryover*. 2. Dados longitudinais. 3. Modelos mistos.  
4. Distribuição a priori. I. Nogueira, Denismar Alves. II. Título.

CDD- 519.5

“ABORDAGEM BAYESIANA PARA UM DELINEAMENTO *CROSSOVER* 2 × 2 COM MEDIDAS REPETIDAS: UM ESTUDO DE SIMULAÇÃO E DADOS REAIS”

A Banca examinadora abaixo-assinada aprova a Dissertação apresentada como parte dos requisitos para a obtenção do título de Mestre em Estatística Aplicada e Biometria pela Universidade Federal de Alfenas. Área de concentração: Estatística Aplicada e Biometria.

Aprovada em: 17 de agosto de 2021.

Prof. Dr. Denismar Alves Nogueira  
Instituição: Universidade Federal de Alfenas - UNIFAL-MG

Prof. Dr. Fabyano Fonseca e Silva  
Instituição: Universidade Federal de Viçosa - UFV

Profa. Dra. Juliana Petrini  
Instituição: Universidade Federal de Alfenas - UNIFAL-MG



Documento assinado eletronicamente por **Denismar Alves Nogueira, Professor do Magistério Superior**, em 17/08/2021, às 16:07, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Fabyano Fonseca e Silva, Usuário Externo**, em 17/08/2021, às 16:11, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Juliana Petrini, Usuário Externo**, em 17/08/2021, às 16:12, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site [https://sei.unifal-mg.edu.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.unifal-mg.edu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **0575725** e o código CRC **4C3A42E5**.

*“Mas buscad primeramente el reino de Dios y su justicia, y todas estas cosas os serán añadidas”.*

*Mateo 6:33*

## AGRADECIMENTOS

*“Mira que te mando que te esfuerces y seas valiente; no temas ni desmayes, porque Jehová tu Dios estará contigo en dondequiera que vayas” ( Josue 1:9).* Agradeço a Deus por seu grande amor e fidelidade na minha vida. Por ser meu guia, fortaleza e perfeita companhia em todos meus dias.

À minha amada família, por seu amor, apoio e motivação em todo este processo.

Aos meus pais, Rafael Miranda e Mercedes Lopez, por todas as risadas brindadas à distância, suas orações e pela espera paciente. Obrigada por toda a dedicação e amor na minha formação pessoal e acadêmica.

Aos meus queridos irmãos, por todos seus cuidados e por estarem sempre para mim. A meus amados sobrinhos por todo seu carinho e alegria que me brindam.

Aos meus amigos em Colômbia, por ficarem sempre atentos a meus avanços e por cada minuto que tem compartilhado comigo. A Cristian Ricardo, por seu apoio em todo tempo, sua compreensão e todo seu carinho.

À aquelas maravilhosas pessoas que conheci no Brasil e que fizeram parte de minhas vivências em Alfenas, em especial a Matheus De Souza, por ter sido como um irmão, por sua amizade incondicional, companhia e ajuda. A Daysi Johana Fernandez e sua família, por ter-me recebido e acolhido em seu lar e por cada detalhe de carinho que tiveram comigo.

Aos meus orientadores, Prof. Dr. Denismar Alves Nogueira e Prof. Dr. Luiz Alberto Beijo, por suas orientações neste trabalho, pela paciência e todos os ensinamentos.

Aos professores da banca de qualificação e defesa, Prof. Dra. Juliana Petrini e Prof. Dr. Fabyano Fonseca e Silva, pelas sugestões e contribuições para enriquecer esta dissertação.

À instituição e ao corpo docente do Programa de Pós-graduação em Estatística Aplicada e Biometria da UNIFAL-MG por sua contribuição na minha formação.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001, a quem também agradeço.

## RESUMO

Nos delineamentos *crossover*, as unidades experimentais recebem todos os tratamentos do estudo em períodos distintos, de acordo com os grupos de sequências formados. Devido a que as unidades experimentais atuam como seu próprio controle, efeitos *carryover* podem estar presentes no modelo; dificultando as inferências sobre os efeitos dos tratamentos. Além disso, podem ser tomadas medidas repetidas da variável resposta ao longo do tempo, o qual permite examinar e comparar as tendências das respostas. No entanto, medidas tomadas na mesma unidade experimental podem estar correlacionadas. Assim, o objetivo deste trabalho foi analisar o delineamento *crossover*  $2 \times 2$  com medidas repetidas dentro de período de tratamento, via abordagem Bayesiana de modelos mistos. Foram considerados como efeitos sistemáticos: tratamentos, períodos, sequências, tempo e a interação simples entre tempo e tratamento; o efeito da unidade experimental foi considerado como aleatório através do modelo misto marginal. Um estudo de simulação foi realizado empregando conceitos frequentistas, considerando três medidas repetidas (efeitos de tempo) dentro de cada período, tamanhos de amostra de 20 e 100, dois diferentes coeficientes de variação (5% e 20%), diferença entre os tratamentos de 1 e 2 erros padrão (*EP*) entre as médias e efeitos *carryover* iguais a zero para cada tratamento. Além disso simularam-se 4 cenários considerando efeitos *carryover* iguais a  $4EP$ , sendo assim, simulou-se 28 cenários com 1000 repetições em cada um. Após a simulação, os dados foram ajustados apenas a modelos Bayesianos. Também se realizou uma aplicação com dados reais da área da fisiopatologia, considerando a frequência mediana do músculo gastrocnêmio lateral direito para avaliar se um protocolo de *exergaming* melhora a atividade muscular em pacientes com câncer. As estimativas Bayesianas *a posteriori* sobre os parâmetros desconhecidos do modelo foram obtidas sob distribuições *a priori* não informativas e função de verossimilhança normal, utilizando o amostrador de Gibbs. A taxa de erro tipo I do teste da diferença de efeitos *carryover* foi próximo a 10%, sendo menor nos cenários com amostra maior. O teste dos efeitos de tempo tende a ser liberal com amostras de 20 unidades experimentais, enquanto com tamanho amostral de 100 torna-se exato ao nível de significância de 5%; o poder do teste foi de aproximadamente 99% nos cenários em que se considerou  $6EP$  desse efeito. O modelo proposto apresentou bom desempenho com relação à acurácia, erro quadrático médio e precisão das estimativas dos efeitos *carryover*, diferenças dos efeitos do tratamento e efeitos do tempo, especialmente com amostras de 100 unidades experimentais. Por sua vez, quando os efeitos *carryover* foram iguais a  $4EP$  da diferença entre médias, as estimativas foram não viesadas e não houve perda de acurácia e precisão, embora com estimativas viesadas de efeito de período. Os resultados com os dados reais se mostraram consistentes, se aproximando dos cenários simulados com diferenças de tratamentos de  $1EP$  e efeito de tempos de 0 e 1.

**Palavras chaves:** Efeitos *carryover*; Dados longitudinais; Modelos mistos; Distribuição *a priori*.

## ABSTRACT

In the designs crossover, the subjects receive all treatments from the study at different periods, according to the groups of sequences formed. Because the subjects act as their own control, carryover effects may be present in the model, making inferences about the effects of treatment difficult. In addition, repeated measures of the response variable can be taken over time, which allows the trends of responses to be examined and compared. However, measures taken in the same subject may be correlated thus, the objective of this work was to analyze the crossover design  $2 \times 2$ , with repeated measurements within the treatment period, the Bayesian approach to mixed models. The following were considered as fixed effects: treatments, periods, sequences, time, and simple interaction between time and treatment, the effect of the subject was considered as random through the mixed marginal model. A simulation study was conducted, considering three repeated measurements (Time effects) within each period, sample sizes of 20 and 100 subjects, two different coefficients of variation (5% and 20%), a difference between treatments of 1 and 2 standard errors (SE) between means and effects carryover equal to zero for each treatment. In addition, four scenarios were simulated considering effects carryover equal to 4SE. Thus, 28 scenarios were simulated with 1000 repetitions in each one. Also, an application was also performed with real data from the area of pathophysiology, considering the median frequency of the right lateral gastrocnemius muscle to assess whether an exergaming protocol improves muscle activity in cancer patients. Bayesian estimates a posteriori of the model's unknown parameters were obtained a priori under non-informative distributions, using the Gibbs sampler. The error type I rate about effects carryover difference test carryover was close to 10%, being smaller in most subject scenarios. The test of the effects of time tends to be liberal with samples of 20 subjects, while samples of 100 subjects it becomes exact at the level of significance of 5% and the power of the test were approximately 99% in the scenarios where it was considered 6SE of this effect. The proposed model presented good performance concerning the accuracy, mean square error, and accuracy of carryover effects estimates differences carryover in treatment and time effects, especially with samples of 100 subjects. In turn, when the carryover effects were equal to 4SE of the difference between means, the estimates were unbiased and there is no loss of accuracy, although there was estimates biased period effect. The result with the real data was consistent, approaching the simulated scenarios with treatment differences of 1SE and effect times of 0 and 1.

**Keywords:** Effects carryover; Longitudinal data; Mixed models; Distribution a priori.



## LISTA DE TABELAS

Tabela1 –	Resposta esperada para as unidades experimentais em um <i>crossover</i> AB/BA. . . . .	18
Tabela2 –	Cenários (C1 até C24) simulados para o <i>crossover</i> $2 \times 2$ com 3 medidas repetidas, considerando a diferença entre as médias dos tratamentos ( $\tau$ ) iguais a $1EP$ e $2EP$ e efeitos <i>carryover</i> $\lambda_1 = \lambda_2 = 0$ . . . . .	40
Tabela3 –	Cenários (C25 até C28) simulados para o <i>crossover</i> $2 \times 2$ com 3 medidas repetidas, nos quais considerou-se a diferença entre as médias dos tratamentos ( $\tau$ ) iguais a $1EP$ ou $2EP$ entre as médias e efeitos <i>carryover</i> $\lambda_1 = \lambda_2 = 4EP_\gamma$ do coeficiente de regressão do tempo. . . . .	40
Tabela4 –	Valores médios observados da frequência mediana do músculo gastrocnêmio lateral direito com o tratamento de <i>exergaming</i> ( <i>A</i> ) ou sem <i>exergaming</i> ( <i>B</i> ) para os 38 pacientes com câncer, tomadas nos atendimentos 0, 10 e 20 de acordo com as sequências nos dois períodos de tempo. . . . .	42
Tabela5 –	Média das estatísticas descritivas das estimativas pontuais e intervalares <i>a posteriori</i> e valor mínimo e máximo dos critérios de convergência dos parâmetros do modelo <i>crossover</i> $2 \times 2$ com 3 medidas repetidas, para os cenários com amostra de 20 unidades experimentais e $cv = 5\%$ , com $\tau = 1EP$ ( C1, C2, C3 ) e $\tau = 2EP$ (C4, C5, C6). . . . .	49
Tabela6 –	Média das Estatísticas descritivas das estimativas pontuais e intervalares <i>a posteriori</i> e valor mínimo e máximo dos critérios de convergência dos parâmetros do modelo <i>crossover</i> $2 \times 2$ com 3 medidas repetidas, para os cenários com amostra de 20 unidades experimentais e $cv = 20\%$ , com $\tau = 1EP$ (C7, C8, C9) e $\tau = 2EP$ (C10, C11, C12). . . . .	50

Tabela7 –	Média das Estatísticas descritivas das estimativas pontuais e intervalares <i>a posteriori</i> e valor mínimo e máximo dos critérios de convergência dos parâmetros do modelo <i>crossover</i> $2 \times 2$ com 3 medidas repetidas, para os cenários com amostra de 100 UE e $cv = 5\%$ , com $\tau = 1EP$ (C13, C14, C15) e $\tau = 2EP$ (C16, C17, C18).	51
Tabela8 –	Média das Estatísticas descritivas das estimativas pontuais e intervalares <i>a posteriori</i> e valor mínimo e máximo dos critérios de convergência dos parâmetros do modelo <i>crossover</i> $2 \times 2$ com 3 medidas repetidas, para os cenários com amostra de 100 UE e $cv = 20\%$ , com $\tau = 1EP$ (C19, C20, C21) e $\tau = 2EP$ (C22, C23, C24).	52
Tabela9 –	Taxas de erro tipo I dos testes da diferença dos efeitos <i>carryover</i> ( $\lambda$ ) e efeito tempo ( $\gamma$ ) nos cenários em que assumiu-se que esses efeitos são nulos e, poder do teste de $\gamma$ para os cenários em que foi simulado um efeito de tempo linear de $r \cdot EP_\gamma$ , com $r = 2$ ou $6$ ; para os cenários com 20 UE, com $\tau = 1EP$ (C1, C2, C3, C7, C8, C9) e $\tau = 2EP$ (C4, C5, C6, C10, C11, C12); nos quais assumiu-se $\lambda_1 = \lambda_2 = 0$ .	53
Tabela10 –	Taxas de erro tipo I dos testes da diferença dos efeitos <i>carryover</i> ( $\lambda$ ) e efeito tempo ( $\gamma$ ) nos cenários em que assumiu-se que esses efeitos são nulos e, poder do teste de $\gamma$ para os cenários em que foi simulado um efeito de tempo linear de $w \cdot EP$ , com $w = 2$ ou $6$ ; para os cenários com 100 UE, com $\tau = 1EP$ (C13, C14, C15, C19, C20, C21) e $\tau = 2EP$ (C16, C17, C18, C22, C23, C24); nos quais assumiu-se $\lambda_1 = \lambda_2 = 0$ .	54
Tabela11 –	Média das Estatísticas descritivas das estimativas pontuais e intervalares <i>a posteriori</i> e valor mínimo e máximo dos critérios de convergência dos parâmetros do modelo <i>crossover</i> $2 \times 2$ com 3 medidas repetidas, para os cenários em que $\lambda_1 = \lambda_2 = 4EP$ , com amostra de 20 unidades experimentais C25 ( $cv = 5\%$ ) e C26 ( $cv = 20\%$ ), e para os cenários com tamanho amostal de 100 C27 ( $cv = 5\%$ ) e C28 ( $cv = 20\%$ ).	58

Tabela12 –	Taxas de erro tipo I dos testes da diferença dos efeitos <i>carryover</i> ( $\lambda$ ) e efeito tempo ( $\gamma$ ) nos cenários em que assumiu-se que esses efeitos são nulos e, poder do teste de $\gamma$ para os cenários em que foi simulado um efeito de tempo linear de $w \cdot EP$ , com $w = 2$ ou $6$ ; para os cenários com amostra de 20 C25 ( $cv = 5\%$ ) e C26 ( $cv = 20\%$ ), e para os cenários com amostras de 100 unidades experimentais C27 ( $cv = 5\%$ ) e C28 ( $cv = 20\%$ ), nos quais assumiu-se $\lambda_1 = \lambda_2 = 4EP$ . . . . .	60
Tabela13 –	Medidas de erro quadrático médio (EQM), acurácia e precisão (amplitude do intervalo HPD) nas estimativas das diferenças dos efeitos <i>carryover</i> ( $\lambda$ ), diferenças dos tratamentos ( $\tau$ ) e efeitos do tempo ( $\gamma$ ), do modelo <i>crossover</i> $2 \times 2$ com 3 medidas repetidas, para os cenários com amostra de 20 unidades experimentais C25 ( $cv = 5\%$ ) e C26 ( $cv = 20\%$ ), e para os cenários com amostras de 100 UE C27 ( $cv = 5\%$ ) e C28 ( $cv = 20\%$ ), nos quais assumiu-se $\lambda_1 = \lambda_2 = 4EP$ . . . . .	60
Tabela14 –	Resultados dos critérios de convergência da cadeia a posteriori para os parâmetros do modelo <i>crossover</i> com medidas repetidas para os dados da frequência mediana do músculo Gastrocnêmio lateral direito. . . . .	61
Tabela15 –	Estimativas pontuais e intervalares dos parâmetros, com prioris não informativas, para os dados da frequência mediana do músculo Gastrocnêmio lateral direito. . . . .	62

## SUMÁRIO

1	<b>INTRODUÇÃO</b>	12
1.1	OBJETIVOS	14
1.1.1	<b>Objetivo Geral</b>	14
1.1.2	<b>Objetivos Específicos</b>	14
2	<b>REFERENCIAL TEÓRICO</b>	15
2.1	DELINEAMENTO <i>CROSSOVER</i>	15
2.1.1	<b>Modelo do Delineamento <i>Crossover</i> <math>2 \times 2</math></b>	17
2.2	MEDIDAS REPETIDAS	19
2.2.1	<b>Estruturas de Covariâncias</b>	21
2.3	MODELOS MISTOS	25
2.3.1	<b>Modelo misto no delineamento <i>Crossover</i> <math>2 \times 2</math> com medidas repetidas</b>	27
2.4	ABORDAGEM BAYESIANA	29
2.4.1	<b>Métodos MCMC: Amostrador de Gibbs</b>	32
2.4.2	<b>Diagnóstico de Convergência</b>	33
2.4.3	<b>Abordagem Bayesiana para Modelos Mistos</b>	34
3	<b>MATERIAIS E MÉTODOS</b>	36
3.1	SIMULAÇÃO DE DADOS <i>CROSSOVER</i> $2 \times 2$ COM 3 MEDIDAS REPETIDAS	36
3.2	EXEMPLO PRÁTICO: PROTOCOLO DE <i>EXERGAMING</i> NA ATIVAÇÃO MUSCULAR EM PACIENTES COM CÂNCER	41
3.3	ANÁLISE E MODELAGEM	42
3.4	ANÁLISES DAS ESTIMATIVAS	44
3.4.1	<b>Erro tipo I e poder do teste</b>	45
3.4.2	<b>Erro Quadrático Médio (EQM)</b>	46
3.4.3	<b>Acurácia e Precisão</b>	47
4	<b>RESULTADOS E DISCUSSÃO</b>	48
4.1	ANÁLISE DOS DADOS SIMULADOS	48
4.1.1	<b>Análises de dados simulados para cenários com <math>\lambda_1 = \lambda_2 = 4EP</math></b>	58
4.2	ANÁLISE DOS DADOS REAIS DO EXEMPLO PRÁTICO	61
5	<b>CONCLUSÕES</b>	63
	<b>REFERÊNCIAS</b>	65

**APÊNDICES** ..... 70

## 1 INTRODUÇÃO

A experimentação é um procedimento mediante o qual são feitas comparações de tratamentos para determinar produtos, técnicas ou métodos mais eficientes na solução de problemas científicos. Nas investigações clínicas, psicológicas e agrícolas comumente se utilizam dos tipos de estudos experimentais: Os delineamentos de grupos paralelos e os delineamentos *crossover*. Num delineamento de grupos paralelos cada unidade experimental (UE) recebe aleatoriamente um e só um tratamento; nos delineamentos *crossover*, cada UE recebe diferentes tratamentos em distintos períodos de tempo (KROGH *et al.*, 2019; JONES; KENWARD, 2014).

Na área da saúde, os delineamentos *crossover* são adequados para estudar tratamentos envolvendo doenças crônicas como asma, enxaqueca, epilepsia, hipertensão, reumatismo, entre outras (GOMES; DANTAS, 2017; HERMANS *et al.*, 2018; MCLOUGHLIN *et al.*, 2019; NIAZI *et al.*, 2017; SAMSONSEN *et al.*, 2018). Para tratar essas doenças, são utilizados medicamentos que possuem efeito rápido, curto e irreversível, já que o objetivo principal não é curar totalmente, mas sim, minimizar seus efeitos através de tratamentos (SENN, 2002). Além disso, Jones e Kenward (2014) afirmam que a teoria sobre delineamentos *crossover* não se aplica apenas a tratamentos de doenças crônicas, citando estudos que utilizam aplicações em análise sensorial, medicina veterinária, psicologia, estresse ocupacional e indústria (não farmacêutica).

Uma das vantagens dos estudos *crossover* em relação aos estudos de grupos paralelos é que cada unidade experimental atua como seu próprio controle (JONES; KENWARD, 2014). Ademais, precisa-se de um número menor de UE do que em outros tipos de delineamentos para obter resultados eficientes e com boa precisão sobre os efeitos dos tratamentos, já que removem quaisquer variações biológicas e metodológicas, o que proporciona vantagens éticas e econômicas. No entanto, os delineamentos *crossover* geralmente são mais longos do que os experimentos de grupos paralelos (KROGH *et al.*, 2019; LI; SIVAGANESAN, 2016; SENN, 2003).

Contudo, existe uma desvantagem nos delineamentos *crossover* que são os possíveis efeitos *carryover*. Os quais são efeitos do tratamento que persistem no período seguinte e podem dificultar a obtenção de estimativas não-viesadas dos efeitos dos tratamentos e outros parâmetros do modelo *crossover*, por exemplo, a média geral, efeitos de períodos e interações

simples entre os efeitos principais. Para lidar com a possibilidade de presença de efeitos *carryover*, na literatura propõe-se realizar um período de limpeza dos efeitos do tratamento atual, o qual é chamado de *washout* (JONES; KENWARD, 2014; KROGH *et al.*, 2019). Contudo, em alguns casos, esse período não é suficiente para obter um efeito *carryover* não significativo.

Um procedimento que pode ser aplicado em delineamentos experimentais é chamado de medidas repetidas, por meio do qual são tomadas várias medidas de uma mesma unidade experimental ao longo do tempo ou espaço. Assim como nos delineamentos *crossover*, as medidas repetidas permitem avaliar as diferenças nas respostas de uma forma mais eficiente, já que possibilitam analisar a variação das resposta, ao longo do tempo, requerendo um menor número de unidades experimentais (WEST; WELCH; GALECKI, 2014). No entanto, as medidas tomadas numa mesma unidade experimental podem ser correlacionadas, devido a contribuição comum de cada UE.

Para a análise dos estudos *crossover* com medidas repetidas, propõe-se neste trabalho uma abordagem de modelos mistos, os quais são adequados para o ajuste da correlação produzida pelas medidas repetidas. A denominação de modelos mistos, vem do fato de considerar no modelo efeitos sistemáticos e aleatórios. No modelo *crossover* os efeitos de tratamentos, períodos, sequências, tempo e a interação simples entre tempo e tratamento, geralmente são considerados sistemáticos; já o efeito da unidade experimental pode ser considerado tanto sistemático como também aleatório.

A abordagem Bayesiana pode ser aplicada na análise de modelos mistos em estudos *crossover* com medidas repetidas. Através desta abordagem implementa-se informações prévias ou crenças iniciais sobre o modelo, por meio de distribuições *a priori*, e especifica-se também a função de verosimilhança para logo ser representadas pela distribuição *posteriori*. Especificamente em delineamentos *crossover*, a análise Bayesiana possibilita realizar inferências sobre os parâmetros do modelo (como efeitos de tratamentos e períodos) considerando as incertezas sobre a presença dos efeitos *carryover* (LI; SIVAGANESAN, 2016). Conforme Senn (1996), pode-se utilizar uma distribuição *a priori* para introduzir informação sobre a dependência do efeito *carryover* e o efeito direto do tratamento nos delineamentos *crossover*.

## 1.1 OBJETIVOS

Nesta seção são apresentados os objetivos deste trabalho.

### 1.1.1 Objetivo Geral

Analisar um estudo *crossover* simples com medidas repetidas dentro de período de tratamento, via abordagem Bayesiana.

### 1.1.2 Objetivos Específicos

- a) Realizar simulação de dados *crossover* com a presença de efeito de tempo, com ausência de interação tempo-tratamento e considerando efeitos *carryover* iguais para os tratamentos;
- b) Verificar se as estimativas *a posteriori* dos parâmetros do modelo correspondem aos valores simulados;
- c) Avaliar o desempenho das estimativas através das medidas erro quadrático médio, acurácia e precisão;
- d) Analisar o modelo com dados reais, através da análise de um experimento da área da fisiopatologia, o qual consistiu em avaliar o efeito de um protocolo de *exergaming* sobre a qualidade de vida e fadiga, relacionada ao câncer, realizado por [Alves \(2020\)](#).



## 2 REFERENCIAL TEÓRICO

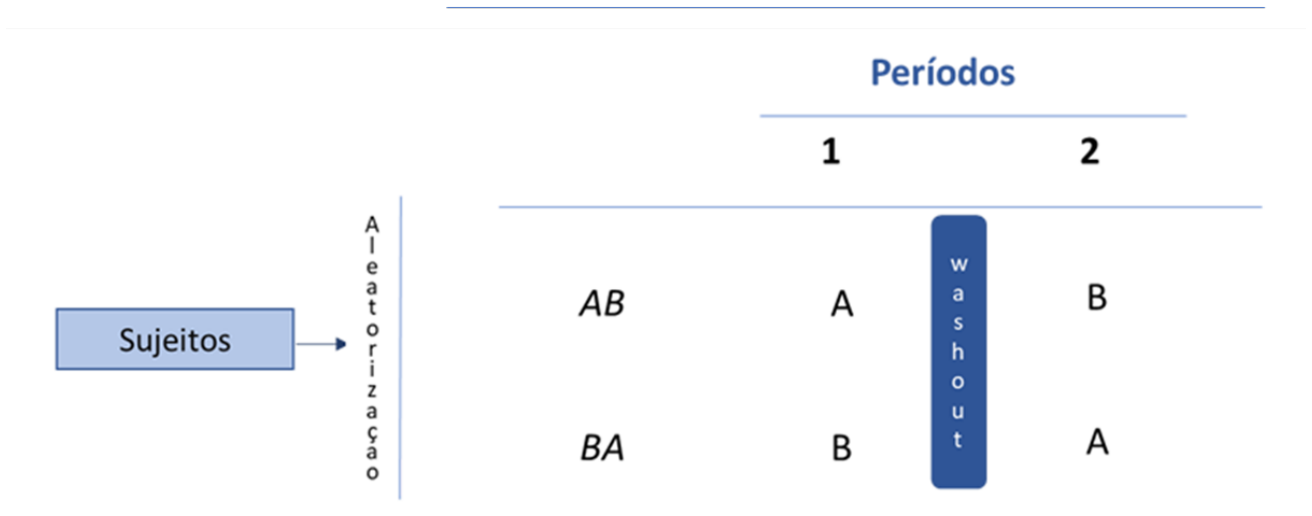
Com o propósito do desenvolvimento desta pesquisa, esta seção introduz abordagens teóricas relevantes na área da Estatística tanto clássica como Bayesiana.

A presente seção está organizada da seguinte forma: Na seção 2.1 apresenta-se o modelo de um dos delineamentos experimentais chamado *crossover*, que tem como finalidade testar a diferença de efeitos de tratamentos, considerando cada unidade experimental como seu próprio controle. Na seção 2.2 é apresentado o conceito e as vantagens de usar dados com medidas repetidas. Na seção 2.3 é apresentado o conceito de modelos mistos e sua utilidade na análise de experimentos. Por último, na seção 2.4 é apresentada uma breve descrição da abordagem Bayesiana, a qual considera e incorpora informações prévias para fazer inferências sobre parâmetros.

### 2.1 DELINEAMENTO *CROSSOVER*

Um delineamento *crossover* é um desenho experimental no qual cada unidade experimental recebe aleatoriamente todos os tratamentos de acordo com as sequências determinadas, assim, cada UE atua como seu próprio controle. Existem diferentes conjuntos de sequências dependendo do número de tratamentos. No desenho mais simples *AB/BA* se tem dois tratamentos (*A* e *B*) e é possível fazer duas sequências (*AB* e *BA*) em dois períodos distintos. Segundo os autores Senn (2002) e Jones e Kenward (2014), este desenho mais simples é chamado  $2 \times 2$  e as UE são distribuídos aleatoriamente em cada sequência; a sequência *AB* recebe o tratamento *A* no primeiro período e o tratamento *B* no segundo período, enquanto a sequência *BA* recebe o tratamento *B* no primeiro período e o tratamento *A* no segundo período, como se mostra na Figura 1.

Figura 1 – Esquema do delineamento *crossover*  $2 \times 2$ .



Fonte: Da autora.

A finalidade do delineamento *crossover* é comparar os efeitos de cada tratamento baseando-se nas informações dentro da unidade experimental e não nas sequências. “O objetivo principal do delineamento *crossover* é, portanto, eliminar do tratamento e período as comparações de quaisquer componentes relacionados com as diferenças entre as unidades experimentais” (JONES; KENWARD, 2014).

Dado que as medidas nos diferentes tratamentos são obtidas em cada UE, geram-se vantagens e desvantagens. Cada UE proporciona uma comparação direta dos tratamentos recebidos e a diferença entre estas medições minimiza os efeitos das UE e conseqüentemente a variabilidade experimental. O delineamento *crossover* precisa de um menor número de UE para fornecer as melhores estimativas não viesadas para as diferenças entre os tratamentos, quando comparado a um experimento de grupos paralelos (KROGH *et al.*, 2019; LI; SIVAGANESAN, 2016).

Uma desvantagem que se tem é que os efeitos do tratamento aplicado num período podem estar presentes no início do período seguinte, influenciando desta maneira a resposta do tratamento e dificultando as inferências sobre os efeitos dos tratamentos individualmente. Esses efeitos persistentes são chamados de efeitos *carryover*. Estes podem ser apresentados de várias maneiras, como efeitos de transição, efeitos de sequências e efeitos de interação de tratamento direto por período, ou seja, as condições presentes num período afetam o tamanho das diferenças entre os tratamentos (LUI, 2016). Os efeitos *carryover* que persistem somente em um período de tratamento são denominados efeitos *carryover* de primeira ordem. Se o efeito perdura até  $j$  períodos é denominado efeito *carryover* de ordem  $j$ . Nos desenhos  $2 \times 2$  estes

efeitos só estão presentes no segundo período.

Para lidar com os efeitos *carryover* propõe-se um período de *washout* entre a aplicação de cada tratamento (JONES; KENWARD, 2014; KROGH *et al.*, 2019). Este é um período no qual se espera que o efeito direto do tratamento anterior desapareça para não ser transferido ao tratamento seguinte. Este período de *washout* pode ser passivo, em que não se aplica nenhum tratamento às UE, período de descanso, ou ativo (sem parada).

Na literatura encontram-se trabalhos em que se utilizaram diferentes períodos de *washout*. Bahador, Lesan e Kashi (2012) realizaram um estudo *crossover* para avaliar o efeito do consumo de xilitol de curto prazo no nível de saliva de estreptococos benéficos orais; no qual utilizou-se um período de *washout* passivo de um mês. Predrazzoli-Júnior *et al.* (2017), avaliaram a bioequivalência entre duas formulações de succinato de desvenlafaxina monoidratado, em voluntários saudáveis, com um período *washout* de sete dias. Os períodos de *washout* ativo são geralmente utilizados em experimentos terapêuticos multi-doses, no qual os pacientes podem receber, quase que imediatamente, o próximo tratamento (SENN, 2003).

O período *washout* depende da natureza do tratamento. Um período de *washout* adequado deve ser suficientemente longo para minimizar qualquer interferência relevante nas respostas das UE. Sem embargo, existem casos nos quais o efeito *carryover* afeta o estado da UE, isto é, as unidades experimentais reagem de maneira diferente ante a aplicação de qualquer outro tratamento; o que influencia na resposta do tratamento apesar da aplicação do período de *washout* (KUEHL, 2000).

### 2.1.1 Modelo do Delineamento *Crossover* $2 \times 2$

Nos delineamentos *crossover*, devido ao fato dos efeitos diretos de tratamento poderem estar aliados ao efeito *carryover*, é importante eliminar tal efeito da comparação. O modelo estatístico apresentado neste trabalho, o qual considera esse efeito, é baseado nos estudos de Chow e Lui (2009) e Jones e Kenward (2014). Se  $y_{ijk}$  é o valor observado da variável resposta  $Y_{ijk}$ , então, de maneira tradicional, descreve-se o modelo linear do delineamento *crossover* como:

$$y_{ijk} = \mu + \pi_j + \tau_{t[i,j]} + \lambda_{t[i,j-1]} + S_k + e_{ijk} \quad (2.1)$$

sendo,

$\mu$ : intercepto;

$\pi_j$ : efeito associado ao período  $j$ ,  $j = 1, 2$ ;

$\tau_{t[i,j]}$ : efeito direto de tratamento aplicado no período  $j$  da sequência  $i$ ,  $i = 1, 2$ ;

$\lambda_{t[i,j-1]}$ : efeito *carryover* para o tratamento  $t$ , no qual  $\lambda_{t[i,0]} = 0$ ;

$S_k$ : efeito associado à  $k$  – ésima unidade experimental,  $k = 1, \dots, n_i$  em que  $n_i$  é a quantidade de UE na sequência  $i$ ;

$e_{ijk}$ : erro aleatório, que segue uma distribuição normal com média 0 e assumindo variâncias homogêneas e independentes  $\sigma_e^2$ .

O efeito *carryover* no modelo pode ser considerado como o efeito da sequência ou como a interação do tratamento com o período e somente se encontra presente no segundo período.

A resposta esperada para as UE num *crossover* simples, nos dois grupos de sequências, são apresentados na Tabela 1; em que  $\tau_1$  e  $\tau_2$  são os efeitos diretos dos tratamentos  $A$  e  $B$ , respectivamente;  $\pi = \pi_1 - \pi_2$  é a diferença entre os dois períodos; e  $\lambda_1$  e  $\lambda_2$  são os correspondentes efeitos *carryover*.

Tabela 1 – Resposta esperada para as unidades experimentais em um *crossover* AB/BA.

Sequências	UE	Período 1	Período 2
$AB$	$n_1$	$y_{1,1,n_1} = \mu + \pi_1 + \tau_1$	$y_{1,2,n_1} = \mu + \pi_2 + \tau_2 + \lambda_1$
$BA$	$n_2$	$y_{2,1,n_2} = \mu + \pi_1 + \tau_2$	$y_{2,2,n_2} = \mu + \pi_2 + \tau_1 + \lambda_2$

Fonte: Da autora.

O propósito do delineamento *crossover* é testar a diferença dos efeitos de tratamento, mas sabe-se que, se os efeitos *carryover* dos tratamentos estão presentes e forem diferentes entre si pode gerar estimativas viesadas. Esta é outra desvantagem dos delineamentos *crossover*: vários efeitos importantes podem ser confusos (JONES; KENWARD, 2014).

Os delineamentos de experimentos *crossover* podem ser feitos tomando medidas repetidas da resposta dos tratamentos para cada UE, em diferentes momentos, dentro de um período de tempo; a fim de observar se o efeito do tratamento é influenciado pelo tempo. Este conceito é tratado com mais detalhes na próxima seção.

## 2.2 MEDIDAS REPETIDAS

Em diferentes áreas da pesquisa têm crescido a demanda de métodos estatísticos para a análise de dados de experimento com medidas repetidas, os quais também são denominados como dados longitudinais. Este delineamento se encontra entre os mais utilizados nas investigações médicas, sociais e psicológicas (GÓMEZ *et al.*, 2012). Medidas repetidas é um procedimento que consiste em coletar dados de uma variável resposta em várias ocasiões ao longo do tempo ou em diferentes condições de estudo de uma mesma unidade experimental (WEST; WELCH; GALECKI, 2014).

O objetivo principal dos experimentos com medidas repetidas é examinar e comparar as tendências das respostas ao longo do tempo, envolver comparações do valor médio das respostas dos tratamentos ou estudar os efeitos de tratamento em tempos específicos (ISLAM; CHOWDHURY, 2017). Segundo Littell *et al.* (2006), o interesse dos estudos com medidas repetidas está em verificar se existe um efeito principal de tratamento, efeito principal do tempo e se existe uma interação de tempo e tratamento.

Nos experimentos com medidas repetidas as fontes de variabilidade entre UE são reduzidas, dado que cada UE pode ser interpretado como seu próprio controle. Assim, esta é uma vantagem tal como nos delineamentos *crossover*. Além disso, este método possui maior precisão na estimação de parâmetros do modelo de análise, melhora o poder do teste e reduz o tamanho da amostra (FERNÁNDEZ; VALLEJO, 1996).

É uma prática comum que nos experimentos *crossover* sejam coletadas uma sequência de observações de cada UE, sob o mesmo tratamento, em cada período (JONES; KENWARD, 2014). Desta forma, o modelo tradicional *crossover*  $2 \times 2$  com medidas repetidas de cada tratamento em cada período de tempo, pode ser escrito como:

$$y_{ijkl} = \mu + \pi_j + \tau_{t[i,j]} + \lambda_{t[i,j-1]} + \gamma_l + \gamma_l \tau_{t[i,j]} + S_k + e_{ijkl} \quad (2.2)$$

Nota-se que a diferença do modelo *crossover* simples (2.1), é que este incorpora o efeito sistemático do  $l$ -ésimo tempo como medida repetida ( $\gamma_l$ ) e a interação simples entre tempo e tratamento ( $\gamma_l \tau_{t[i,j]}$ ). Além disso, neste modelo não se pode assumir que os erros  $e_{ijkl}$  para a mesma UE sejam independentes. Embora os tratamentos tenham sido designados aleatoriamente para cada UE, os níveis do fator de medidas repetidas, neste caso, o tempo, não

são atribuídos aleatoriamente às UE (LITTELL *et al.*, 2006). As respostas medidas na mesma UE estão correlacionadas porque contêm uma contribuição comum destas. Sendo assim, pode-se assumir que  $e_{ijkl} \sim N(\mathbf{0}, \mathbf{R})$ , sendo  $\mathbf{R}$  uma matriz de variâncias-covariâncias. Os demais efeitos do modelo são definidos como no modelo (2.1).

Devido a correlação produzida nas mesmas UE, a suposição de independência dos erros dos modelos clássicos de análises de variância não são garantidas. Como também, as variâncias das medidas repetidas podem ser modificadas frequentemente com o tempo. Estas situações trazem perdas de precisão e capacidade de predição dos modelos ajustados às suposições clássicas (GÓMEZ *et al.*, 2012).

De acordo com Littell, Henry e Ammerman (1998), os experimentos de medidas repetidas podem ser analisados mediante vários métodos, como a análise de variância univariada (ANOVA), análise de variância multivariada (MANOVA) e modelos mistos. No método ANOVA, as medidas repetidas são tomadas como um experimento de parcelas subdivididas no tempo, associando o fator tratamento às unidades experimentais nas parcelas principais e o fator tempo às subparcelas (LITTELL *et al.*, 2006). Este método pode ser usado quando se assume que as variâncias são homogêneas nos tempos e que as covariâncias são as mesmas independentemente do intervalo de tempo entre as medidas (TORRES; NAVARRO; PÉREZ, 2003).

O método de análise de variância multivariada, naturalmente assume que existe correlação entre as medidas tomadas num mesmo indivíduo, mas não assume uma estrutura em particular para a matriz de covariância. O modelo de covariância não-estruturado requer um número de observações suficientemente grande para a estimação de seus parâmetros (TORRES; NAVARRO; PÉREZ, 2003). Além disso, esse tipo de análise não admite dados desbalanceados.

Também podem ser utilizados modelos mistos, que são considerados mais adequados por permitir analisar os dados de forma correta e eficiente. Brown e Prescott (2014), destacam vantagens da modelagem através de modelos mistos em experimentos com medidas repetidas:

- permitem que seja utilizado um modelo que possibilita a estimativa dos efeitos gerais dos tratamentos e a estimativa dos efeitos dos tratamentos em tempos específicos;
- Os erros padrão dos efeitos dos tratamentos são calculados levando em conta as informações de todas as medidas, sendo estes mais robustos do que calculados em tempos específicos;

- A ocorrência de dados faltantes não causa problema, desde que sejam unidades experimentais aleatorizadas;
- Os modelos mistos também permitem incorporar padrões de covariâncias dentro das UE, sendo esta uma grande vantagem a destacar, devido que as medições feitas na mesma UE são provavelmente mais correlacionadas do que aquelas tomadas em UE diferentes.

Ajustar uma estrutura de covariância adequada é de grande relevância para poder realizar inferências precisas quando se trabalha com dados de medidas repetidas. Segundo [Littell \*et al.\* \(2006\)](#), quando se utiliza uma estrutura de covariância muito simples, ignorando uma correlação importante, corre-se o risco de aumentar a taxa de erro tipo I e subestimar os erros padrão. Por outro lado, se a estrutura for muito complexa, sacrifica-se o poder e a eficiência.

### 2.2.1 Estruturas de Covariâncias

Para realizar uma escolha adequada da estrutura de covariância que se ajusta melhor a dados de medidas repetidas, deve analisar-se as características que cada uma possui. Nesta seção, apresenta-se diferentes modelos de covariâncias que podem ser adequados para modelar a estrutura dos erros de dados com medidas repetidas,  $\mathbf{R} = \mathbf{I} \otimes \mathbf{\Sigma}$ , de acordo com [Littell \*et al.\* \(2006\)](#), [Pinheiro e Bates \(2000\)](#) e [West, Welch e Galecki \(2014\)](#).

- Matriz Componentes de Variância (VC):

É a estrutura de covariância mais simples, da forma  $\mathbf{\Sigma} = \mathbf{I}\sigma^2$ , sendo  $\mathbf{I}$  a matriz identidade e  $\sigma^2$  um parâmetro de variância; sendo assim,

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma^2 & 0 & 0 & \cdots & 0 \\ & \sigma^2 & 0 & \cdots & 0 \\ & & \sigma^2 & \cdots & 0 \\ & & & \ddots & \vdots \\ & & & & \sigma^2 \end{pmatrix},$$

esta estrutura assume correlação nula dos erros dentro das UE e supõe variâncias homogêneas e independentes entre os componentes.

Sua aplicação se dá em experimentos em que as medidas repetidas são tomadas em intervalos de tempos suficientemente grandes e, portanto, pode-se assumir que a correlação dentro das UE não é significativa (LITTELL *et al.*, 2006).

- Matriz Não-estruturada (UN):

A estrutura de covariância mais complexa que assume uma correlação existente entre as medidas tomadas na mesma UE é a matriz não-estruturada, a qual tem a seguinte forma,

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1K} \\ & \sigma_2^2 & \sigma_{23} & \cdots & \sigma_{2K} \\ & & \sigma_3^2 & \cdots & \sigma_{3K} \\ & & & \ddots & \vdots \\ & & & & \sigma_K^2 \end{pmatrix},$$

esta estrutura se ajusta bem aos dados, pressupõe variâncias distintas para cada uma das  $n_i$  ocasiões e erros dentro da UE para cada par de medidas têm sua própria correlação. Ademais, requer a estimação de  $K + K(K - 1)/2$  parâmetros.

- Matriz de Simetria Composta (CS):

O modelo mais simples que assume correlação no conjunto de dados de medidas repetidas é a matriz de Simetria Composta,

$$\Sigma = \sigma^2 \begin{pmatrix} 1 & \rho & \rho & \cdots & \rho \\ & 1 & \rho & \cdots & \rho \\ & & 1 & \cdots & \rho \\ & & & \ddots & \vdots \\ & & & & 1 \end{pmatrix},$$

em que  $\sigma^2$  é a variância das diferentes medições e  $\rho\sigma^2$  é a covariância entre observações dentro de uma mesma UE em todos os pares de tempos. Esta estrutura pressupõe variâncias homogêneas nas  $n_i$  ocasiões e covariâncias constantes, independentemente do distanciamento de tempo entre as medidas. Requer a estimação de dois parâmetros,  $\sigma^2$  e  $\rho$ .

- Matriz Autorregressiva de Primeira Ordem (AR(1)):



De acordo com [West, Welch e Galecki \(2014\)](#), a aplicação da estrutura AR(1) é apropriada nos experimentos em que as medidas foram tomadas em distâncias iguais entre tempos nas mesmas UE. Sendo assim,

$$\Sigma = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{K-1} \\ & 1 & \rho & \cdots & \rho^{K-2} \\ & & 1 & \cdots & \vdots \\ & & & \ddots & \rho \\ & & & & 1 \end{pmatrix}$$

a qual pressupõe variâncias homogêneas nas diversas ocasiões e que a correlação entre os pontos diminui à medida em que as distâncias entre os tempos aumentam, desta forma, observações adjacentes tendem a ser mais correlacionadas do que observações afastadas.

Na estrutura AR(1), a correlação entre os erros adjacentes é  $\rho$ , diferenciando-se da matriz não-estruturada, em que cada par de erros tem sua própria correlação. A estrutura autorregressiva de primeira ordem requer estimativas de apenas dois parâmetros,  $\sigma^2$  e  $\rho$ .

- Matriz Toeplitz (TOEP):

$$\Sigma = \sigma^2 \begin{pmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{K-1} \\ & 1 & \rho_1 & \cdots & \rho_{K-2} \\ & & 1 & \cdots & \vdots \\ & & & \ddots & \rho_1 \\ & & & & 1 \end{pmatrix}$$

A Matriz Toeplitz é similar à AR(1) no sentido de que pares de erros separados por um intervalo comum compartilham a mesma correlação, no entanto, erros separados por  $d$  unidades têm correlação  $\rho_d$  em vez de  $\rho^d$ . A matriz Toeplitz permite maior flexibilidade nas correlações do que AR(1), contudo, requer a estimação de  $K$  parâmetros.

As estruturas AR(1) e Toeplitz são adequadas quando as observações são igualmente espaçadas e a estrutura de correlação não muda consideravelmente ao longo do tempo.

- Matriz Ante-dependência de Primeira Ordem (ANTE (1)):

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_1 & \sigma_1\sigma_3\rho_1\rho_2 & \cdots & \sigma_1\sigma_K\rho_1\rho_2\cdots\rho_{K-1} \\ & \sigma_2^2 & \sigma_2\sigma_3\rho_2 & \cdots & \sigma_2\sigma_K\rho_2\rho_3\cdots\rho_{K-1} \\ & & \sigma_3^2 & \cdots & \vdots \\ & & & \ddots & \sigma_{K-1}\sigma_K\rho_{K-1} \\ & & & & \sigma_K^2 \end{pmatrix},$$

a estrutura ANTE(1), permite manter as principais pressuposições das matrizes AR(1) e TOEP e, além disso, possibilita estabelecer espaçamento e mudança desiguais ao longo do tempo. Esta estrutura assume que a variação entre as observações muda ao longo do tempo e que a correlação entre pares é o produto das correlações entre os tempos adjacentes e entre as observações, para que a correlação possa mudar ao longo do tempo. A estrutura ANTE(1) requer a estimativa de  $2K - 1$  parâmetros.

Outras estruturas matriciais que podem ser consideradas na modelagem de dados com medidas repetidas são aquelas que modificam a estrutura das matrizes CS, AR(1) e TOEP, permitindo estabelecer variâncias heterogêneas ao longo do tempo (LITTELL *et al.*, 2006; WEST; WELCH; GALECKI, 2014). Estas matrizes são chamadas de Simetria Composta com heterogeneidade de variâncias (CSH), Autoregressiva de Primeira Ordem com variâncias heterogêneas (ARH(1)) e a matriz Toeplitz com heterogeneidade de variâncias (TOEPH).

Para a escolha da estrutura de covariância adequada aos dados, geralmente inicia-se testando uma estrutura de covariância Não-Estruturada e, a partir daí, testam-se outras estruturas candidatas, de estruturas mais simples de acordo com o padrão de correlação notado, visando construir o modelo mais parcimonioso (PINHEIRO; BATES, 2000).

A escolha da estrutura mais apropriada pode ser realizada através de uma análise gráfica, visualizando padrões de correlação entre as medidas nos diferentes momentos ou, através de critérios de informação, como o critério de Akaike (AIC), AIC Corrigido (AICC) e o critério de informação Bayesiano (BIC). Estes medem o ajuste relativo dos modelos concorrentes de covariância. Mais informações sobre estes procedimentos podem ser encontrados em Littell *et al.* (2006).

### 2.3 MODELOS MISTOS

Os modelos mistos são amplamente utilizados na ecologia, ciências médicas, biológicas, físicas e sociais. Este tipo de modelos é geralmente usado para a avaliação de medidas repetidas tomadas nas mesmas unidades experimentais ou naqueles nos quais são feitas medições em grupos de unidades experimentais relacionadas (JONES; KENWARD, 2014).

De acordo com Zhou (2012), os dados que são tomados dentro de um mesmo nível, geralmente são correlacionados. Os modelos mistos permitem capturar a correlação entre as medidas através dos efeitos aleatórios e suas distribuições assumidas. Além disso, o modelo pode lidar bem com dados ausentes e desbalanceamento.

Segundo Seoane (2014), a denominação de modelos mistos vem do fato destes modelos incorporarem efeitos sistemáticos e aleatórios para descreverem a relação entre a variável resposta (dependente) e as variáveis explicativas (independentes, preditoras ou covariáveis). Os efeitos sistemáticos são, na maioria dos casos, os fatores de interesse dentro do estudo. Estes são restritos aos níveis observados da variável e não se tem interesse em generalizar os resultados para outros níveis. Entretanto, os efeitos aleatórios são variáveis explicativas cujos níveis foram amostrados de uma população sobre a qual se tem interesse inferir (LITTELL *et al.*, 2006).

Segundo West, Welch e Galecki (2014), a equação na forma matricial geral, que descreve um modelo misto, é dada da seguinte forma:

$$Y = X\beta + Zu + e, \quad (2.3)$$

$$\begin{aligned} u &\sim N(\mathbf{0}, \mathbf{G}) \\ e &\sim N(\mathbf{0}, \mathbf{R}), \end{aligned}$$

em que  $Y$  é o vetor resposta de dimensão  $n \times 1$ ,  $X$  é a matriz de incidência associada aos efeitos sistemáticos de dimensão  $n \times b$ ,  $\beta$  é o vetor de efeitos sistemáticos  $b \times 1$ ,  $Z$  é a matriz de incidência associada aos efeitos aleatórios de dimensão  $n \times q$ ,  $u$  é o vetor de efeitos aleatórios  $q \times 1$  e,  $e$  o vetor dos erros de dimensão  $q \times 1$ . Sendo  $n$  o número de observações,  $b$  o número de efeitos sistemáticos e  $q$  o número de efeitos aleatórios.

Assume-se que  $u$  e  $e$  são independentes e possuem distribuição normal multivariada com vetor de médias 0 e matriz de variância e covariâncias  $\mathbf{G}$  e  $\mathbf{R}$ , respectivamente. Matrizes que

são positivas definidas e, portanto, não singulares. Matricialmente tem-se:

$$\text{var} \begin{bmatrix} \mathbf{u} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{G} & \boldsymbol{\phi} \\ \boldsymbol{\phi} & \mathbf{R} \end{bmatrix},$$

em que  $\boldsymbol{\phi}$  é a matriz nula.

A média condicional das observações dados os efeitos aleatórios do modelo é  $E(\mathbf{Y}|\mathbf{u}) = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}$ , portanto, assume-se que a distribuição condicional  $\mathbf{Y}|\mathbf{u} \sim N(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}, \mathbf{R})$ .

A distribuição conjunta de  $\mathbf{Y}$  e  $\mathbf{u}$  pode ser obtida pelo produto da distribuição condicional de  $\mathbf{Y}$  dado  $\mathbf{u}$ , pela marginal de  $\mathbf{u}$ , isto é,

$$f(\mathbf{Y}, \mathbf{u}) = f(\mathbf{Y}|\mathbf{u}) \times f(\mathbf{u}) \quad (2.4)$$

ou seja,

$$f(\mathbf{Y}, \mathbf{u}) = \frac{1}{(2\pi)^{c/2} |\mathbf{R}|^{1/2}} \exp \left\{ -\frac{1}{2} [(\mathbf{Y} - (\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}))' \mathbf{R}^{-1} (\mathbf{Y} - (\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}))] \right\} \times \frac{1}{(2\pi)^{f/2} |\mathbf{G}|^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{u}' \mathbf{G}^{-1} \mathbf{u}) \right\} \quad (2.5)$$

em que  $c$  e  $f$  são os postos das matrizes  $\mathbf{R}$  e  $\mathbf{G}$ , respectivamente. Comumente assume-se que  $\mathbf{R} = \mathbf{I}\sigma_e^2$ , em que  $\sigma_e^2$  é a variância do erro, considerando-a como homogênea, e  $\mathbf{I}$  a matriz identidade.

O modelo misto geral, dado na equação (2.3), implica o seguinte modelo misto marginal:

$$\begin{aligned} \mathbf{Y} &= \mathbf{X}\boldsymbol{\beta} + \mathbf{e}^*, \\ \mathbf{e}^* &\sim N(\mathbf{0}, \mathbf{V}), \end{aligned} \quad (2.6)$$

sendo,

$$\begin{aligned} \mathbf{V} = \text{var}(\mathbf{Y}) &= \text{var}(\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}) \\ &= \text{var}(\mathbf{X}\boldsymbol{\beta}) + \text{var}(\mathbf{Z}\mathbf{u}) + \text{var}(\mathbf{e}) \\ &= 0 + \mathbf{Z} \text{var}(\mathbf{u}) \mathbf{Z}' + \text{var}(\mathbf{e}) \\ &= \mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R}. \end{aligned}$$

Como se tem que  $E(\mathbf{e}) = \mathbf{0}$ , a esperança marginal de  $\mathbf{Y}$  é dada por:

$$E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta},$$

de onde se obtém a distribuição marginal de  $\mathbf{Y} \sim N(\mathbf{X}\boldsymbol{\beta}, \mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R})$ . A função densidade de probabilidade de  $\mathbf{Y}$  segue uma normal multivariada, descrita da seguinte maneira:

$$f(\mathbf{Y}) = \frac{1}{(2\pi)^{n/2}[\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R}]^{1/2}} \exp \left\{ -\frac{1}{2}[(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R})^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\beta})] \right\}. \quad (2.7)$$

West, Welch e Galecki (2014), destacam que o modelo marginal de  $\mathbf{Y}$  dado em (2.6), implica o mesmo número de parâmetros de covariâncias que no modelo misto geral (2.3). Além disso, o modelo marginal é menos restrito, já que só requer que a matriz de variâncias e covariâncias  $\mathbf{V}$  seja positiva definida. Portanto, sua aplicação é importante quando no procedimento de um *software* a matriz  $\mathbf{G}$  do modelo misto geral não é positiva definida.

### 2.3.1 Modelo misto no delineamento *Crossover* $2 \times 2$ com medidas repetidas

Nos delineamentos *crossover* podem ser tomadas várias medições da resposta da UE ao tratamento dentro de um mesmo período. Isto é, são tomadas medições repetidas para analisar se o efeito do tratamento varia com o tempo. Considerando, por exemplo, três medidas repetidas ( $T1, T2$  e  $T3$ ; respectivamente); o desenho *crossover* com medidas repetidas dentro de período é ilustrado na Figura 2.

Os modelos mistos são apropriados na análise dos desenhos *crossover* com medidas repetidas. Considerando  $S_k$  como o efeito aleatório da  $k$ -ésima unidade experimental no modelo *crossover*  $2 \times 2$  com medidas repetidas de cada tratamento dentro de cada período de tempo (2.2), pode-se considerar como o modelo misto marginal:

$$y_{ijkl} = \mu + \pi_j + \tau_{t[i,j]} + \lambda_{t[i,j-1]} + \gamma_l + \gamma_l \tau_{t[i,j]} + \epsilon_{ijkl}, \quad (2.8)$$

em que,

$$\epsilon_{ijkl} = S_k + e_{ijkl},$$

assumindo que  $S_k \sim N(0, \sigma_s^2)$ , sendo  $\sigma_s^2$  a variância das unidades experimentais, e  $e_{ijkl}$  o vetor

Figura 2 – Esquema do delineamento *crossover*  $2 \times 2$  com três medidas repetidas.

		Períodos						
		1			2			
Sequências		T1	T2	T3		T1	T2	T3
AB		A	A	A	w a s h o u t	B	B	B
BA		B	B	B		A	A	A

Fonte: Da autora.

de erros com  $e_{ijkl} \sim N(\mathbf{0}, \mathbf{R})$ .

Note que, o modelo (2.8), pode ser escrito de forma matricial como um modelo misto marginal:

$$Y = X\beta + e^*,$$

em que,

$Y$  é o vetor de observações;

$X$  é a matriz de incidência associada aos efeitos sistemáticos;

$\beta$  é o vetor de efeitos sistemáticos, em que  $\beta^t = (\mu, \pi, \tau, \lambda, \gamma, \gamma\tau)$ ;

$e^*$  é o vetor aleatório de erros, dado em função dos efeitos aleatórios das unidades experimentais  $S_k$  e a parte residual do modelo,  $e^* \sim N(\mathbf{0}, \mathbf{V})$ , sendo  $\mathbf{V} = \text{var}(Y) = \mathbf{ZGZ}' + \mathbf{R}$ .

Para a análise dos modelos mistos realiza-se inferências sobre os efeitos sistemáticos, os efeitos aleatórios e os parâmetros de variância e covariância do modelo. Tais inferências podem ser feitas através da abordagem Bayesiana.

Realizar inferências sobre os modelos mistos através da abordagem Bayesiana proporciona várias vantagens como as mencionadas por [Nogueira et al. \(2003\)](#), a inferência Bayesiana permite obter estimativas intervalares para todos os parâmetros do modelo, levando em conta a incerteza destes; embora, apresente limitações computacionais. Ademais, esta abordagem permite incorporar informações prévias ao experimento, enriquecendo o processo

estatístico e permitindo realizar uma análise descritiva completa sobre os parâmetros. Desta forma, ao utilizar informações prévias informativas (privilegiam valores do espaço paramétrico), podem ser obtidas estimativas intervalares mais estreitas do que os intervalos de confiança.

## 2.4 ABORDAGEM BAYESIANA

Na evolução científica sempre existe interesse por realizar novas descobertas em distintas áreas do conhecimento, com a finalidade de gerar avanços a nível mundial. Este tipo de avanço é dado a partir de experimentos científicos. Dita atividade científica consiste em tomar uma parte da população, chamada amostra, para obter conclusão ou informações sobre a população. Generalizar os resultados obtidos a partir de uma amostra para toda a população de estudo é um processo conhecido como inferência estatística (PINTO, 2015).

Neste contexto, a abordagem Bayesiana faz o processo de inferência sobre os parâmetros populacionais desconhecidos baseando-se na distribuição *a posteriori*. Esta distribuição é obtida a partir da combinação das informações extraídas da amostra (função de verossimilhança) e as informações anteriores à amostragem, chamadas de distribuições *a priori* sobre os parâmetros de interesse (DEGROOT; SCHERVISH, 2012). Nesta abordagem, os parâmetros populacionais desconhecidos  $\theta$  são considerados como aleatórios, e a obtenção da distribuição *a posteriori* é uma aplicação do Teorema de Bayes.

Paulino, Turkman e Murteira (2003), apresentam o teorema de Bayes através da seguinte expressão:

$$P(\theta|\mathbf{y}) = \frac{P(\mathbf{y},\theta)}{P(\mathbf{y})} = \frac{P(\mathbf{y}|\theta)P(\theta)}{P(\mathbf{y})} = \frac{L(\theta|\mathbf{y})P(\theta)}{\int L(\theta|\mathbf{y})P(\theta)d\theta}, \quad (2.9)$$

em que  $\mathbf{y} = (y_1, y_2, \dots, y_n)$  corresponde a amostra. Como o denominador não depende de  $\theta$ , o teorema de Bayes toma a forma geral

$$P(\theta|\mathbf{y}) \propto L(\theta|\mathbf{y})P(\theta), \quad (2.10)$$

em que  $P(\theta|\mathbf{y})$  representa a distribuição *a posteriori* de  $\theta$  dado  $\mathbf{y}$ ,  $L(\theta|\mathbf{y})$  é a função de verossimilhança de  $\mathbf{y}$  e  $P(\theta)$  é a distribuição *a priori* do parâmetro  $\theta$ .

A utilização de informação *a priori*, na abordagem Bayesiana, requer a especificação de uma distribuição de probabilidade  $P(\theta)$  para a quantidade de interesse  $\theta$ . Esta distribuição deve representar probabilisticamente o conhecimento que se tem sobre os parâmetros populacionais ( $\theta$ ) antes da realização do experimento. A definição da distribuição *a priori* é feita pelo pesquisador, especificando a forma da distribuição e os valores dos hiperparâmetros. Denomina-se hiperparâmetros os parâmetros da distribuição *a priori*. Assim, o teorema de Bayes faz uma atualização do conhecimento prévio com a informação amostral.

Representar as informações *a prioris* não é uma tarefa fácil, uma vez que estas são tipicamente subjetivas e precisam ser elicitadas. Existem dois tipos de distribuições *a priori*, segundo a informação obtida.

O primeiro é quando não se tem informações sobre os parâmetros implicados no problema ou quando as informações que se tem são pouco relevantes em relação as informações que se extraem dos dados. Nesse caso, consideram-se distribuições *a prioris* não informativas, as quais alguns pesquisadores chamam de *prioris* vagas, com muita variabilidade (MORALES; CAUSIL, 2019). Assim, nesses casos as estimativas obtidas mediante a distribuição *a posteriori* são mais baseadas na função de verossimilhança dos dados.

O segundo caso é quando a informação prévia sobre os parâmetros é relevante, e privilegia certos valores de  $\theta$  dentro do espaço paramétrico, de forma que esse conhecimento seja incorporado na análise por meio de uma distribuição de probabilidade (PAULINO; TURKMAN; MURTEIRA, 2003). Este tipo de informação pode ser obtida por meio da opinião de especialistas ou de estudos anteriores. As distribuições *a prioris* usadas nestes casos recebem o nome de distribuições *a prioris* informativas.

Tendo definido os tipos de distribuições *a prioris* a serem usadas e a função de verossimilhança, se possibilitará a obtenção da distribuição *a posteriori* dos parâmetros  $\theta$ . Esta distribuição descreve o comportamento dos parâmetros após observados os dados amostrais, e contém toda a informação probabilística a respeito destes parâmetros (PAULINO; TURKMAN; MURTEIRA, 2003). Na prática, geralmente existe interesse sobre vários parâmetros; assim, o resultado da abordagem Bayesiana proporciona a distribuição conjunta *a posteriori* de todos os parâmetros e em seguida, se obtém a distribuição marginal de cada parâmetro. Isto é feito através da integração da distribuição *a posteriori* conjunta com relação



aos outros parâmetros  $(\theta_1, \theta_2, \dots, \theta_p)$  da mesma (NTZOUFRAS, 2011), isto é,

$$P(\theta_i|y) = \int P(\boldsymbol{\theta}|y)d\theta_{-i}$$

em que,  $i = 1, \dots, p$  e  $\theta_{-i}$  é o vetor  $\boldsymbol{\theta}$  excluindo-se  $\theta_i$ .

Mediante a distribuição *a posteriori* marginal são obtidas as estimativas pontuais do parâmetro  $\theta$ , as quais podem ser observadas na média, moda, mediana ou variância. Também pode ser analisado o comportamento do parâmetro por se conhecer sua distribuição e mediante estimativas intervalares de credibilidade, as regiões mais prováveis.

Deste modo, dada uma distribuição *a posteriori*  $P(\theta|y)$ , o intervalo  $a < \theta < b$  será chamado de intervalo de credibilidade  $100(1 - \alpha)\%$  para  $\theta$  se:

$$P(\theta \in [a, b] | y) \equiv \int_a^b P(\theta|y)d\theta \geq 1 - \alpha. \quad (2.11)$$

Os intervalos de credibilidade são definidos numericamente, ou seja, não é uma região aleatória e, podem ser interpretados como os intervalos dentro dos quais se encontra o parâmetro  $\theta$  com certa probabilidade.

Paulino, Turkman e Murteira (2003) relatam que com o uso de distribuições *a priori* informativas se obtém intervalos de credibilidade que são mais estreitos do que com distribuições *a priori* não informativas. Isto traz uma grande vantagem na obtenção das estimativas dos parâmetros, as quais serão precisas, minimizando as incertezas sobre o conhecimento do parâmetro.

Geralmente, o interesse do pesquisador se baseia na verificação de afirmações sobre o parâmetro  $\theta$ , ou seja, verificar se as hipóteses que se tem sobre o parâmetro são verdadeiras ou falsas, com o objetivo de tomar uma decisão sobre a população. Para isto, podem ser utilizados os intervalos de credibilidade. Estes devem ser empregados quando o interesse é testar hipóteses bilaterais, donde a hipótese de nulidade  $H_0 : \theta = \theta_0$  é testada contra uma hipótese alternativa  $H_1 : \theta \neq \theta_0$ . O procedimento consiste em verificar se  $\theta_0$  pertence ao intervalo, caso contrário, rejeita-se  $H_0$  (ZELLNER, 1996).

Existe um intervalo de credibilidade que contém os valores mais credíveis de  $\theta$  *a posteriori*, isto é, aquele que engloba todos os valores de  $\theta$  com menor amplitude. Este intervalo é denominado como intervalo de alta densidade *a posteriori*, abreviadamente intervalo HPD (*Highest Probability Density*). Conforme Paulino, Turkman e Murteira (2003),

existe uma infinidade de intervalos de credibilidade com o mesmo grau de credibilidade, no entanto, o intervalo HPD é único e é definido como  $[a, b] = \{\theta : P(\theta|y) \geq c_{1-\alpha}\}$  com  $c_{1-\alpha} > 0$  a maior constante tal que

$$\int_a^b P(\theta|y) d\theta \geq 1 - \alpha.$$

Existem vários métodos tradicionais para a obtenção de intervalos HPD. Costuma-se variar os limites de credibilidades até encontrar o de menor comprimento possível. Contudo, o método de Monte Carlo via Cadeias de Markov (MCMC) usa diretamente a definição HPD para a obtenção dos intervalos. Assim, a obtenção do intervalo HPD faz uso de aproximações numéricas para garantir a probabilidade em menor área ou amplitude.

#### 2.4.1 Métodos MCMC: Amostrador de Gibbs

A implementação da abordagem Bayesiana muitas vezes não é trivial, pois as integrais envolvidas para obter a distribuição *a posteriori* ou as distribuições marginais dos parâmetros *a posterioris* podem ser de alta complexidade e não ter solução analítica. Este foi um inconveniente que inviabilizou por muitos anos que a abordagem Bayesiana fosse aplicada em problemas mais complexos. O fato foi contornado a partir do uso de cálculo numérico ou de simulações baseadas em processos estocásticos (MORALES; CAUSIL, 2019).

Uma das técnicas mais utilizadas são os métodos de Monte Carlo via Cadeias de Markov (MCMC). Este é um procedimento de reamostragem, onde os métodos MCMC consistem em gerar dados aleatórios de uma cadeia de Markov,  $\theta_1, \theta_2, \dots$ , cuja distribuição estacionária de equilíbrio, após convergência do processo, é a distribuição *a posteriori*  $P(\boldsymbol{\theta}|y)$ . Cada valor gerado,  $\theta_h$ , depende unicamente de seu predecessor,  $\theta_{h-1}$ .

Um algoritmo de cadeia de Markov específico, que foi considerado útil em muitos problemas multidimensionais é o amostrador de Gibbs, o qual é definido em termos de subvetores de  $\boldsymbol{\theta}$  (GELMAN *et al.*, 2014). O amostrador de Gibbs consiste em aplicar um procedimento iterativo para gerar realizações de uma cadeia de Markov, usando Monte carlo de distribuições de probabilidades conhecidas, o qual se aproxima de sua condição de equilíbrio conforme o aumento do número de iterações (PAULINO; TURKMAN; MURTEIRA, 2003).

Seja  $\theta$  um vetor de  $h$  parâmetros com distribuições associadas  $P(\theta|y)$  e  $\theta = (\theta_1, \theta_2, \dots, \theta_h)$ . De acordo com Tobias (2014), o algoritmo de amostragem de Gibbs procede da seguinte forma:

i) Selecione um vetor de parâmetros inicial  $\theta^{(0)} = (\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_h^{(0)})$ . Este vetor inicial pode ser escolhido arbitrariamente ou a partir de um método de estimativa bruto.

(1) Amostra  $\theta_1^{(1)}$  da densidade condicional *a posteriori* completa é

$$p(\theta_1 | \theta_2 = \theta_2^{(0)}, \theta_3 = \theta_3^{(0)}, \dots, \theta_h = \theta_h^{(0)}, y)$$

(2) Amostra  $\theta_2^{(1)}$  é

$$p(\theta_2 | \theta_1 = \theta_1^{(1)}, \theta_3 = \theta_3^{(0)}, \dots, \theta_h = \theta_h^{(0)}, y)$$

⋮

(k) Amostra  $\theta_h^{(1)}$  é

$$p(\theta_h | \theta_1 = \theta_1^{(1)}, \theta_2 = \theta_2^{(1)}, \dots, \theta_{h-1} = \theta_{h-1}^{(1)}, y)$$

ii) Percorrer repetidamente (1)  $\rightarrow$  (k) para obter  $\theta^{(2)} = (\theta_1^{(2)}, \theta_2^{(2)}, \dots, \theta_h^{(2)})$ ,  $\theta^{(3)}$ , etc. Sempre condicionando os valores dos parâmetros na etapa anterior.

Quando as densidades condicionais completas *a posteriori* não possuem uma forma específica conhecida, o amostrador deve ser implementado mediante o algoritmo *Metropolis Hastings* (GAMERMAN; LOPES, 2006), podendo ser combinado com o Amostrador de Gibbs.

## 2.4.2 Diagnóstico de Convergência

A ideia de se ter convergência da cadeia para a Posteriori marginal, é ter um número suficientemente grande de iterações para que satisfaça a condição de estacionariedade com um padrão assintótico. Porém, não se trata de uma convergência formal no sentido matemático tendendo ao infinito o número de iterações, mas sim, determinar um número de iterações que possibilite possuir uma distribuição próxima à de equilíbrio (NOGUEIRA; SÁFADI; FERREIRA, 2004). Portanto, pode-se monitorar a convergência através de alguns métodos, como: critério de Geweke (GEWEKE, 1992), critério Heidelberger e Welch (HEIDELBERGER; WELCH, 1983) e critério de Raftery e Lewis (RAFTERY; LEWIS, 1992).

O critério de Geweke é utilizado para realizar um diagnóstico para a ausência de convergência das cadeias *a posteriori*. Assim, o critério propõe testar a hipótese de que não há ausência de convergência da cadeia, baseando-se no teste de igualdade de médias, através da comparação da primeira e a última parte da cadeia para verificar estabilidade. Geralmente se considera os primeiros 10% e os últimos 50% dos valores de uma cadeia. Conclui-se que não há ausência de convergência da cadeia quando o valor se encontra entre os limites de uma distribuição normal padrão, isto é, o valor seja menor do que 1,96 (estatística Z da distribuição Normal padrão) a um nível de significância maior do que 5%.

Segundo [Paulino, Turkman e Murteira \(2003\)](#), o critério de Raftery e Lewis fornece estimativas da dimensão da cadeia de Markov a ser construída, isto é, o número de iterações necessárias para que o Amostrador de Gibbs apresente convergência à distribuição estacionária; como também o espaçamento entre as iterações (*thin*) e o número de iterações para o período de aquecimento (*burn-in*) necessárias para atingir a convergência. Ademais, fornece informação sobre o fator de dependência (FD), o qual é responsável por determinar se há necessidade de aumentar o número iterações para alcançar a convergência, devido à autocorrelação dentro da cadeia. Conclui-se que não há indícios de que a cadeia não convergiu quando o Fator de dependência é próximo de 1.

O critério de Heidelberger e Welch consiste em testar a hipótese nula de estacionariedade da cadeia resultante, fazendo uso de testes estatísticos, como Cramer-Von Mises. De maneira sucessiva, o teste é aplicado primeiro em toda a cadeia, depois, após o descarte dos primeiros 10%, 20%, ... valores da cadeia, até que a hipótese nula seja aceita ou 50% da cadeia tenha sido descartada. Se mesmo nesse último caso a hipótese nula for rejeitada, tem-se uma falha do teste de estacionariedade, indicando que uma cadeia de maior comprimento é necessária ([NOGUEIRA; SÁFADI; FERREIRA, 2004](#)).

### 2.4.3 Abordagem Bayesiana para Modelos Mistos

Para realizar as inferências sobre os parâmetros de interesse do modelo misto (2.3), através da abordagem Bayesiana, deve ser considerada a distribuição *a posteriori* conjunta dos parâmetros desconhecidos  $\beta$ ,  $u$ ,  $G$  e  $R$ , segundo [Sorensen e Gianola \(2002\)](#) isto é:

$$P(\boldsymbol{\beta}, \mathbf{u}, \mathbf{G}, \mathbf{R}|\mathbf{y}) \propto P(\mathbf{y}|\boldsymbol{\beta}, \mathbf{u}, \mathbf{R})P(\boldsymbol{\beta}, \mathbf{G}, \mathbf{u}, \mathbf{R}), \quad (2.12)$$

em que,  $P(\mathbf{y}|\boldsymbol{\beta}, \mathbf{u}, \mathbf{R})$  representa a função de verossimilhança a qual segue uma distribuição normal multivariada com vetor de médias  $\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}$  e matriz de variâncias-covariâncias  $\mathbf{R}$ , e  $P(\boldsymbol{\beta}, \mathbf{G}, \mathbf{u}, \mathbf{R})$  representa a distribuição *a priori* dos parâmetros desconhecidos do modelo.

Cabe ressaltar que a abordagem Bayesiana considera todos os parâmetros como aleatórios, desta maneira a definição de efeitos sistemáticos do modelo misto não surge naturalmente nesta abordagem (SORENSEN; GIANOLA, 2002).

Como dito anteriormente, a distribuição *a priori* é escolhida pelo pesquisador de acordo com as informações que se tem sobre o experimento. Sem embargo, nos modelos mistos geralmente é considerado que existe independência *a priori* entre os parâmetros, desta forma têm-se:

$$P(\boldsymbol{\beta}, \mathbf{u}, \mathbf{G}, \mathbf{R}) = P(\boldsymbol{\beta})P(\mathbf{u}|\mathbf{G})P(\mathbf{G})P(\mathbf{R}),$$

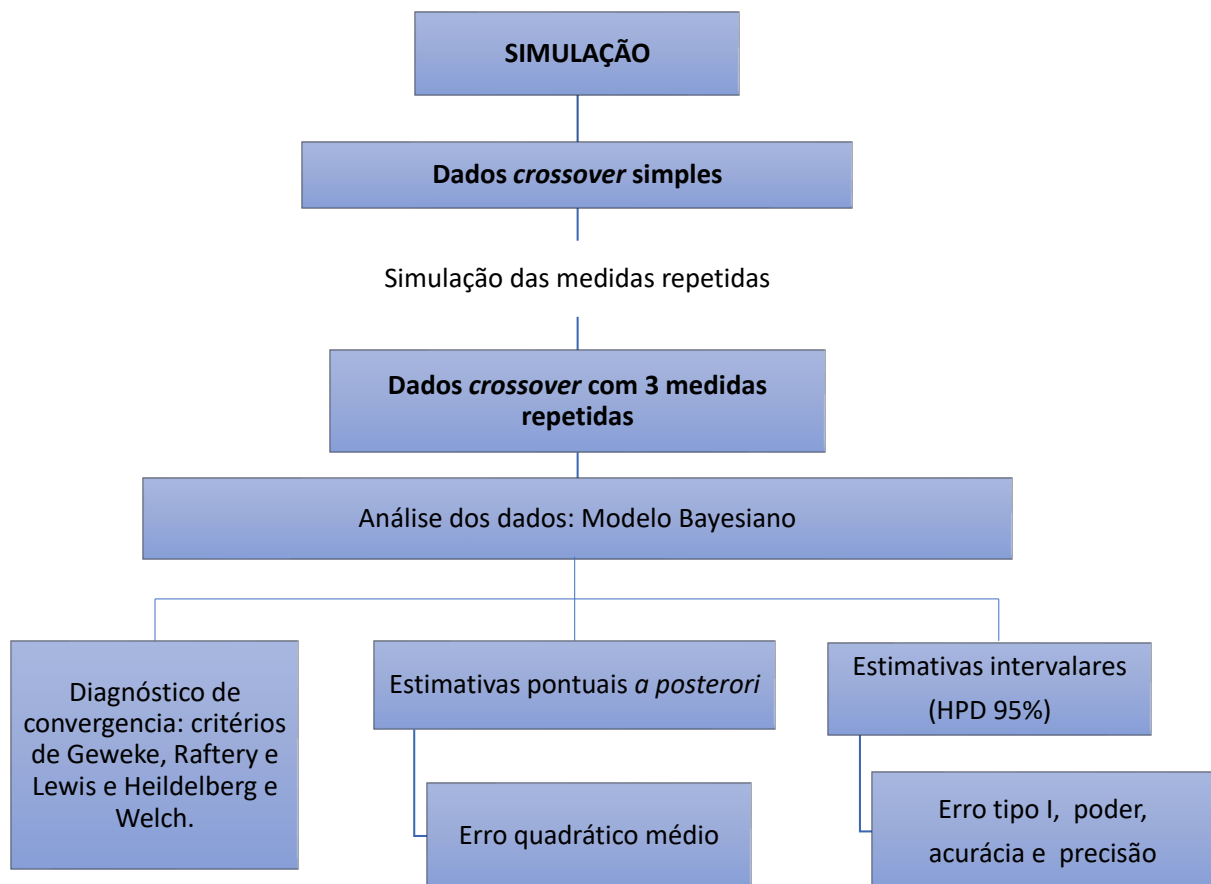
sendo  $P(\boldsymbol{\beta})$  a distribuição *a priori* referente aos efeitos “sistemáticos”, podendo ser informativa ou não;  $P(\mathbf{u}|\mathbf{G})$  é a distribuição *a priori* especificada para a parte “aleatória” do modelo (2.3), a qual pode ser uma distribuição normal multivariada com vetor de médias zero e matriz de covariâncias  $\mathbf{G}$ ; as distribuições *a priori* de  $\mathbf{G}$  e  $\mathbf{R}$  possuem diferentes estruturas dependendo da escolha dessas matrizes de variância e covariâncias. Quando a análise é feita com o uso de distribuições *a prioris* não informativas os resultados obtidos são similares aos da inferência frequentista, isto porque a maior contribuição das estimativas provem da função dos dados.

Dentro da abordagem Bayesiana, as inferências sobre os parâmetros desconhecidos é feita a partir da distribuição conjunta *a posteriori*. Logo, para obter as distribuições marginais dos parâmetros de interesse, pode ser realizada a integração da distribuição *a posteriori* (2.12) com relação a cada um dos parâmetros, ou utilizar algum procedimento numérico como MCMC, e a partir dessas marginais podem ser obtidas as medidas descritivas dos parâmetros.

### 3 MATERIAIS E MÉTODOS

Nesta seção detalha-se os métodos utilizados para realizar a simulação dos dados *crossover* simples com medidas repetidas e descreve-se uma análise deste modelo, como também os procedimentos utilizados na análise estatística. Um resumo dos passos realizados se encontra na Figura 3. Além disso apresenta-se um exemplo prático com dados reais.

Figura 3 – Fluxograma com os passos realizados para a simulação e análises dos dados.



Fonte: Da autora.

#### 3.1 SIMULAÇÃO DE DADOS *CROSSOVER* $2 \times 2$ COM 3 MEDIDAS REPETIDAS

Com a finalidade de avaliar o comportamento da simulação de dados obtidos em um experimento *crossover* foram consideradas algumas situações, as quais foram simuladas  $N = 1000$  vezes cada uma. Foram gerados 28 cenários, em que a variância foi definida em função da média geral e o coeficiente de variação das UE, isto é,  $\sigma^2 = (\mu * cv)^2$  e, portanto, a

variação não foi afetada pelo período ou tratamento. Além disso, assumiu-se covariância nula entre as UE ( $cov = 0$ ). Também, considerou-se diferentes números de amostras, coeficientes de variação e diferentes valores para os parâmetros do modelo *crossover*  $2 \times 2$  da equação (2.1), da seguinte maneira:

- Tamanhos de amostras:  $n = \{20 \text{ e } 100\}$ .
- Coeficientes de variação:  $cv = \{5\% \text{ e } 20\%\}$ .
- Média geral:  $\mu = 100$ .
- variâncias:  $\sigma^2 = \{25 \text{ e } 400\}$
- Diferença entre as médias dos tratamentos:  $\tau = \tau_1 - \tau_2 = w \cdot EP$ , com  $w = 1$  e  $2$ .
- Diferença dos períodos:  $\pi = \pi_2 - \pi_1 = 0,1$ .
- Efeitos *carryover*:  $\lambda_1 = \lambda_2 = 0$ ; em alguns casos foram considerados efeitos *carryover*  $\lambda_1 = \lambda_2 = 4EP$  (Erro Padrão para a comparação específica) entre as médias dos tratamentos; com a finalidade de avaliar o impacto nas estimativas quando os efeitos *carryover* de cada tratamento são iguais entre si, porém diferentes de zero.

Foi considerado um número igual de UE em cada sequência, assim, para amostras de 20 e 100 UE foram consideradas 10 e 50 UE em cada sequência, respectivamente. A média entre as UE foi  $\mu = 100$ . Assumiu-se que não há diferença entre os períodos quando se estabeleceu  $\pi = 0,1$ , um valor próximo de zero. Também optou-se por não haver efeito significativo da sequência ao estabelecer que os efeitos *carryover* dos tratamentos são iguais ( $\lambda_1 = \lambda_2$ ).

Adicionalmente, para estabelecer a diferença entre tratamentos, a média do tratamento  $A$  ( $\tau_A$ ) menos a média do tratamento  $B$  ( $\tau_B$ ), utilizou-se uma diferença de 1 ou 2 erros padrão entre as médias ( $\tau = 1EP$  ou  $\tau = 2EP$ ). Sendo assim, os valores paramétricos das médias populacionais que foram fixados em função da diferença dos erros padrão de médias são dados por  $\tau_B = \tau_A - w \cdot EP$ , com  $\tau_A = 167$  e  $w = 1$  e  $2$ . Portanto, as UE que recebem tratamento  $A$  alcançam, em média, 1 ou 2 erros padrão de unidades a mais do que as UE que recebem tratamento  $B$ .

Desta forma, considerou-se que os desvios padrão populacionais dos tratamentos são os mesmos para as respostas dentro de cada tratamento, ou seja,  $\sigma_1 = \sigma_2$  e calculou-se o erro

padrão entre as médias conforme [Silva e Ferreira \(2003\)](#), o qual é dado por:

$$EP = \sqrt{\sigma^2 \left( \frac{1}{n_A} + \frac{1}{n_B} \right)},$$

em que,  $\sigma^2$  é a variância em cada cenário,  $n_A$  é o número de respostas para o tratamento  $A$  e  $n_B$  é o número de respostas para o tratamento  $B$ , independente do período ou sequência em que se encontrem, ou seja,  $n_A = n_B = 20$  ou  $n_A = n_B = 100$ .

Logo após estabelecer os valores paramétricos utilizados, realizou-se a simulação dos dados de duas distribuições normais bivariadas, correspondendo aos dois grupos de sequências. De acordo com as respostas esperadas para as UE no delineamento *crossover* simples, dadas na Tabela 1, o vetor de médias da sequência  $AB$  é

$$\boldsymbol{\mu}_{AB} = (\mu + \pi_1 + \tau_1, \mu + \pi_2 + \tau_2 + \lambda_1)$$

e o vetor de médias da sequência  $BA$  é

$$\boldsymbol{\mu}_{BA} = (\mu + \pi_1 + \tau_2, \mu + \pi_2 + \tau_1 + \lambda_2).$$

Assim sendo, os dados da sequência  $AB$  seguem uma normal bivariada com média  $\boldsymbol{\mu}_{AB}$  e matriz de variância e covariância:

$$\boldsymbol{\Sigma}_{AB} = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix},$$

e os dados da sequência  $BA$  seguem uma normal bivariada com média  $\boldsymbol{\mu}_{BA}$  e com a mesma estrutura para a matriz de variância e covariância, Componentes de Variância,  $\boldsymbol{\Sigma}_{BA} = \boldsymbol{\Sigma}_{AB}$ .

Após simulado o experimento do estudo *crossover*  $2 \times 2$ , foi inserido o efeito de regressão linear nas unidades experimentais dentro de cada período, através do seguinte modelo:

$$y_{ijkl} = b_{ijk} + \gamma x_l + \epsilon_{ijkl},$$

em que:

$b_{ijk}$  é o intercepto, o qual foi considerado como a resposta do  $k$ -ésima unidade experimental para a  $i$ -ésima sequência no  $j$ -ésimo período dos dados *crossover*  $2 \times 2$  simulados;

$\gamma$  é o coeficiente de regressão associado ao tempo;

$x_l$  é o  $l$ -ésimo tempo, com  $l = 1, 2, 3$ ; sendo  $x_1 = 0, x_2 = 1$  e  $x_3 = 2$ ;



$\epsilon_{ijkl} \sim N(\mathbf{m}_k, \Sigma_k)$  é o vetor de erro associado a observação para o  $k$ -ésima UE, dado por:

$$\Sigma_k = \mathbf{I}\sigma^2 = \begin{pmatrix} 20 & 0 & 0 \\ 0 & 20 & 0 \\ 0 & 0 & 20 \end{pmatrix},$$

sendo  $\mathbf{I}$  a matriz identidade e  $\sigma^2 = 20$  a variância entre os erros. Os dados obtidos correspondem à matriz dos erros da  $k$ -ésima UE no  $l$ -ésimo tempo, com  $l = 1, 2, 3$ .

Uma vez que existe o interesse em verificar se a simulação proposta capta de maneira adequada o coeficiente de regressão do tempo estabelecido, para cada um dos cenários simulados, foram variados os coeficientes de regressão do tempo ( $\gamma$ ) de acordo com os erros padrão deste, o qual calculou-se mediante a seguinte fórmula:

$$EP_\gamma = \sqrt{\sigma^2/S_{xx}} \quad (3.1)$$

em que  $\sigma^2 = 20$  é a variância entre os erros e

$$S_{xx} = \sum_{i=1}^n x_i^2 - n(\bar{x})^2,$$

sendo  $n$  o número total de observações e  $x_i$  o tempo em cada observação. Desta forma, estabeleceu-se o coeficiente de regressão do tempo como  $\gamma = r \cdot EP_\gamma$ , com  $r = 0; 2$  e  $6$ . Para os cenários com amostras de 20 UE o  $EP_\gamma = 0,5$  assim  $\gamma$  tomou valores 0; 1 e 3 e, para os cenários com 100 UE o  $EP_\gamma = 0,22$ , portanto,  $\gamma$  foi estabelecido na simulação como 0; 0,44 e 1,32. Ademais, assumiu-se que não há interação significativa entre tempo e o tratamento,  $\gamma_l \tau_{t[i,j]} = 0$ .

A descrição dos cenários simulados para o Delineamento *crossover*  $2 \times 2$  com 3 Medidas Repetidas é apresentada nas Tabelas 2 e 3. Os dados foram simulados através de uma adaptação da função `getSimulationData()` do pacote *Reproducer* (MADEYSKI, 2017) do software R (R Core Team, 2018), na qual modificou-se a forma de se introduzir as médias das sequências ( $\mu_{AB}$  e  $\mu_{BA}$ ), de acordo com as respostas esperadas para os UE no delineamento *crossover*  $2 \times 2$  dadas na Tabela 1 e, colocar o número distinto de UE para cada sequência. Com isso, possibilitou-se gerar medidas repetidas através do efeito da regressão linear, como se encontra descrito na rotina do APÊNDICE A.

Tabela 2 – Cenários (C1 até C24) simulados para o *crossover*  $2 \times 2$  com 3 medidas repetidas, considerando a diferença entre as médias dos tratamentos ( $\tau$ ) iguais a 1EP e 2EP e efeitos *carryover*  $\lambda_1 = \lambda_2 = 0$ .

Cen.	n	cv(%)	$\lambda_1 = \lambda_2$	$\pi$	$\tau$	$\gamma$	$\gamma\tau$
C1	20	5	0	0,1	1,581 (1EP)	0,00	0
C2	20	5	0	0,1	1,581 (1EP)	1,00	0
C3	20	5	0	0,1	1,581 (1EP)	3,00	0
C4	20	5	0	0,1	3,162 (2EP)	0,00	0
C5	20	5	0	0,1	3,162 (2EP)	1,00	0
C6	20	5	0	0,1	3,162 (2EP)	3,00	0
C7	20	20	0	0,1	6,325 (1EP)	0,00	0
C8	20	20	0	0,1	6,325 (1EP)	1,00	0
C9	20	20	0	0,1	6,325 (1EP)	3,00	0
C10	20	20	0	0,1	12,649 (2EP)	0,00	0
C11	20	20	0	0,1	12,649 (2EP)	1,00	0
C12	20	20	0	0,1	12,649 (2EP)	3,00	0
C13	100	5	0	0,1	0,707 (1EP)	0,00	0
C14	100	5	0	0,1	0,707 (1EP)	0,44	0
C15	100	5	0	0,1	0,707 (1EP)	1,23	0
C16	100	5	0	0,1	1,414 (2EP)	0,00	0
C17	100	5	0	0,1	1,414 (2EP)	0,44	0
C18	100	5	0	0,1	1,414 (2EP)	1,23	0
C19	100	20	0	0,1	2,828 (1EP)	0,00	0
C20	100	20	0	0,1	2,828 (1EP)	0,44	0
C21	100	20	0	0,1	2,828 (1EP)	1,23	0
C22	100	20	0	0,1	5,657 (2EP)	0,00	0
C23	100	20	0	0,1	5,657 (2EP)	0,44	0
C24	100	20	0	0,1	5,657 (2EP)	1,23	0

Fonte: Da autora.

Nota: Cen. - cenários; n - número total de unidades experimentais; cv - coeficiente de variação; EP - Erro padrão.

Tabela 3 – Cenários (C25 até C28) simulados para o *crossover*  $2 \times 2$  com 3 medidas repetidas, nos quais considerou-se a diferença entre as médias dos tratamentos ( $\tau$ ) iguais a 1EP ou 2EP entre as médias e efeitos *carryover*  $\lambda_1 = \lambda_2 = 4EP_\gamma$  do coeficiente de regressão do tempo.

Cen.	n	cv(%)	$\lambda_1 = \lambda_2$	$\pi$	$\tau$	$\gamma$	$\gamma\tau$
C25	20	5	6,325(4EP)	0,1	1,581 (1EP)	3,00	0
C26	20	20	25,298(4EP)	0,1	6,325 (1EP)	1,00	0
C27	100	5	5,657(4EP)	0,1	1,414 (2EP)	0,00	0
C28	100	20	11,314(4EP)	0,1	2,828 (1EP)	1,32	0

Fonte: Da autora.

Nota: Cen. - cenários; n - número total de unidades experimentais; cv - coeficiente de variação; EP - Erro padrão.

### 3.2 EXEMPLO PRÁTICO: PROTOCOLO DE *EXERGAMING* NA ATIVAÇÃO MUSCULAR EM PACIENTES COM CÂNCER

É possível aplicar a teoria proposta neste trabalho a experimentos com dados reais. Nesta seção descreve-se o experimento que foi utilizado para a aplicação. Através da qual objetivou-se verificar se efeitos *carryover* estão presentes no modelo, como também determinar se há efeitos significativos do tratamento estudado.

O experimento provem de um estudo *crossover*  $2 \times 2$  com três medidas repetidas, realizado por (ALVES, 2020) o qual teve como objetivo avaliar o efeito de um protocolo de *exergaming* sobre a qualidade de vida e fadiga relacionada ao câncer, perfil térmico das pernas, equilíbrio postural estático e fadiga muscular em pacientes com câncer.

Nesse estudo foram comparados dois tratamentos, o protocolo de *exergaming* (*A*) e o tratamento controle, sem protocolo de *exergaming* (*B*). O tratamento de estudo consistiu em simular exercícios físicos através de um jogo virtual *Your Shape Fitness Evolved™* 2012 com auxílio do console *Xbox 360 Kinect™*. Os pacientes jogavam duas a três vezes na semana, alcançando 109 minutos de treino, para um total de 20 atendimentos (consultas). As avaliações foram feitas em três tempos, 0, após 10 e 20 atendimentos com protocolo de *exergaming* e sem protocolo de *exergaming*.

A amostra utilizada foi de 38 pacientes voluntários com câncer que receberam tratamento quimioterapêutico no centro de oncologia da Santa Casa de Alfenas-MG e residentes na mesma cidade. Esses pacientes foram aleatorizados para estar na sequência *AB* ou na sequência *BA*. Os pacientes que se encontravam na primeira sequência receberam o tratamento de *exergaming* no primeiro período e logo, no segundo período, passaram ao momento sem *exergaming* no qual só foram medidas as variáveis do estudo; o período de *washout* foi de um mês. Da mesma forma, procedeu-se a intervenção dos tratamentos para os pacientes da sequência *BA*, no qual, receberam tratamento *exergaming* no período 2 e momento sem *exergaming* no período 1.

Para atingir os objetivos do estudo foram tomadas medidas de 32 variáveis, neste trabalho considerou-se apenas a atividade muscular do músculo gastrocnêmio lateral direito (GLD) analisada por meio da frequência mediana (FM) do sinal electromiográfico de superfície. Para a escolha desta variável foi realizada uma pré-análise dos dados com uma abordagem frequentista, com a finalidade de verificar quais variáveis não haviam apresentado

efeitos *carryover*, por último escolheu-se a FM do músculo GLD dentro das variáveis que cumpriram essa condição. Os valores médios da FM do músculo GLD de acordo com as sequências e atendimentos se encontram na Tabela 4.

Tabela 4 – Valores médios observados da frequência mediana do músculo gastrocnêmio lateral direito com o tratamento de *exergaming* (*A*) ou sem *exergaming* (*B*) para os 38 pacientes com câncer, tomadas nos atendimentos 0, 10 e 20 de acordo com as sequências nos dois períodos de tempo.

Seq.	n	Período 1			Período 2		
		Aten. 0	Aten. 10	Aten. 20	Aten. 0	Aten. 10	Aten. 20
<i>AB</i>	17	165,91	168,06	177,94	164,42	166,91	160,93
<i>BA</i>	21	158,11	159,88	159,33	159,19	165,17	173,95

Fonte: Da autora.

Nota: Seq. - sequência; n - número de pacientes; Aten. 0 - atendimento 0; Aten. 10 - atendimento 10; Aten. 20 - atendimento 20.

### 3.3 ANÁLISE E MODELAGEM

Para realizar a análise dos dados simulados e reais, foi utilizado o *software OpenBUGS* (LUNN *et al.*, 2009) e o *software R* (R Core Team, 2018). O modelo considerado baseou-se no proposto por Zhang, Toubouti e Carlin (2017), no qual, considerando a correlação existente entre as três medições tomadas dentro de um período para o mesmo tratamento, a distribuição de probabilidade é:

$$\begin{pmatrix} Y_{ijk1} \\ Y_{ijk2} \\ Y_{ijk3} \end{pmatrix} \sim N \left( \begin{pmatrix} \mu_{ijk1} \\ \mu_{ijk2} \\ \mu_{ijk3} \end{pmatrix}, \Sigma_t \right), \quad (3.2)$$

em que  $\mu_{ijkl}$  é dado por:

$$\mu + \lambda Seq_i + \pi Per_j + \tau Trat_{t[i,j]} + \gamma Tempo_l + \gamma\tau Tempo_l \times Trat_{t[i,j]}, \quad (3.3)$$

$t$  é determinado pela  $i$ -ésima sequência e o  $j$ -ésimo período.

Neste modelo, não foi especificado o efeito da UE de forma explícita, considerando-se o modelo misto marginal ( $Y = X\beta + e^*$ ) descrito na equação (2.8). Foi definido que  $Trat_1 = 1$ , o tratamento a ser testado (*A*) e  $Trat_2 = 0$  para o tratamento de controle (*B*), sendo  $\tau$  a diferença entre o tratamento de teste e o controle. Também definiu-se que  $Seq_1 = 1$ , sendo a sequência

$AB$  e,  $Seq_2 = 0$  sendo a sequência  $BA$ , para o primeiro período  $Per_1 = 0$  e para o segundo período  $Per_2 = 1$ , además,  $Tempo_1 = 1$ ,  $Tempo_2 = 2$  e  $Tempo_3 = 3$  para os três tempos, respectivamente.

Utilizou-se distribuições *a priori* como sendo normais não informativas para os parâmetros do modelo; dado que não se tem uma informação prévia ao experimento. Desta forma, considerou-se as seguintes priors:

$$\begin{aligned} \mu &\sim N(0; 0,001); & \lambda &\sim N(0; 0,001); & \pi &\sim N(0; 0,001); \\ \tau &\sim N(0; 0,001); & \gamma &\sim N(0; 0,001); & \gamma\tau &\sim N(0; 0,001), \end{aligned}$$

as quais estão apresentadas com os hiperparâmetros da média e precisão.

Considerou-se uma priori não informativa inversa Wishart para matriz de variâncias-covariâncias  $\Sigma_t$ , de forma que se definiu a matriz de precisão  $\Sigma_t^{-1} \sim Wishart(\mathbf{R}_t, v)$ , sendo  $t = 1; 2$  e  $v = \dim(\Sigma_t) = 3$  são os graus de liberdade. Segundo [Li et al. \(2020\)](#), a dimensão de matriz  $\Sigma_t$  é o menor valor dos graus de liberdade para o qual a distribuição *a priori* não informativa de Wishart é adequada. Para a matriz  $\mathbf{R}_t$  considerou-se a seguinte estrutura:

$$\mathbf{R}_t = \mathbf{I}\sigma^2 = \begin{pmatrix} 0,001 & 0 & 0 \\ 0 & 0,001 & 0 \\ 0 & 0 & 0,001 \end{pmatrix}.$$

Para ajustar o modelo e obter as distribuições marginais *a posterioris* dos parâmetros  $\mu$ ,  $\lambda$ ,  $\pi$ ,  $\tau$ ,  $\gamma$ ,  $\gamma\tau$ , e as matrizes  $\Sigma_1$  e  $\Sigma_2$ , utilizou-se o *software* *OpenBUGS* de [Lunn et al. \(2009\)](#) através da interface do pacote *R2OpenBUGS* ([LIGGES et al., 2017](#)). Este pacote tem implementado os métodos de Monte Carlo via cadeias de Markov para obtenção das marginais. Gerou-se uma cadeia com 800000 iterações, com descarte (*burn-in*) de 775000 e salto (*thin*) de 5 iterações, obtendo-se uma cadeia final de tamanho igual a 5000 amostras.

A convergência da cadeia final, de cada cenário, foi verificada segundo os critérios de Geweke, Raftery e Lewis e Heidelberger e Welch, os quais se encontram implementados no pacote *Convergence Diagnosis and Output Analysis for MCMC* (CODA) de [Plummer et al. \(2006\)](#), do *software* R.

Foram obtidas também estimativas pontuais para cada um dos parâmetros através da média *a posteriori*, como também o intervalo de máxima densidade *a posteriori* HDP de 95% de credibilidade, com a finalidade de verificar a significância destes, considerando as seguintes

hipóteses:

para a diferença dos efeitos *carryover*:

$$\begin{cases} H_0 : \lambda = 0 \\ H_1 : \lambda \neq 0; \end{cases} \quad (3.4)$$

para a diferença dos efeitos dos períodos:

$$\begin{cases} H_0 : \pi = 0,1 \\ H_1 : \pi \neq 0,1; \end{cases} \quad (3.5)$$

para a diferença dos efeitos dos tratamentos:

$$\begin{cases} H_0 : \tau = w \cdot EP, \text{ com } w = 1 \text{ e } 2 \\ H_1 : \text{não } H_0; \end{cases} \quad (3.6)$$

para o tempo:

$$\begin{cases} H_0 : \gamma = r \cdot EP_\gamma, \text{ com } r = 0; 2 \text{ e } 6 \\ H_1 : \text{não } H_0; \end{cases} \quad (3.7)$$

para a interação tempo-tratamento:

$$\begin{cases} H_0 : \gamma\tau = 0 \\ H_1 : \gamma\tau \neq 0. \end{cases} \quad (3.8)$$

Concluiu-se sobre os testes de hipóteses da seguinte maneira: Se zero não pertence ao intervalo HDP de 95% de credibilidade, a hipóteses  $H_0$  é rejeitada, ou seja, o parâmetro é significativo e, caso contrário, se  $0 \in$  HDP de 95% de credibilidade.

### 3.4 ANÁLISES DAS ESTIMATIVAS

Nesta seção são descritas as medidas calculadas para avaliar a inferência realizada pelas estimativas obtidas com as simulações.

### 3.4.1 Erro tipo I e poder do teste

Para verificar o comportamento de nossa abordagem para uma confiança nominal de 5%, calculou-se as taxas de erro tipo I para os parâmetros da diferença dos efeitos *carryover* ( $\lambda$ ), efeito de tempo ( $\gamma$ ), como também o poder dos testes deste.

Na tomada de decisões através de um teste de hipótese para um parâmetro, existe a possibilidade de cometer certos erros. Desta forma, há a probabilidade de rejeitar a hipóteses nula ( $H_0$ ) sendo verdadeira. Este tipo de erro é conhecido como o erro tipo I ( $\alpha$ ) o qual, normalmente, é estabelecido pelo pesquisador como uma probabilidade muito pequena de acontecer. Outra situação que não é desejável de acontecer na realização de um teste de hipóteses é não rejeitar a hipótese nula quando esta é falsa, este tipo de erro é conhecido como erro tipo II (GIRARDI; CARGNELUTTI; STORCK, 2009).

Por outro lado, uma decisão que pode ser tomada é rejeitar a hipótese nula quando esta é falsa. A probabilidade de tomar essa decisão é o poder do teste, o qual é desejável que seja um valor alto, assim o poder é o complementar do erro tipo II e este vai ser máximo quando a probabilidade de cometer o erro tipo II seja mínima (BOLFARINE; SANDOVAL, 2001).

Então, considerando as hipóteses para a diferença dos efeitos *carryover* dadas na equação (3.4), isto é:

$$\begin{cases} H_0 : \lambda = 0 \\ H_1 : \lambda \neq 0; \end{cases}$$

para o  $n$ -ésimo conjunto de dados, definiu-se  $I_i = 1$  se o intervalo HDP de 95% de credibilidade para  $\lambda$  não continha o valor 0, caso contrário,  $I_i = 0$ . Então, a taxa de erro tipo I foi estimada empiricamente como:

$$\left(\frac{1}{N} \sum_{i=1}^N I_i\right), \quad (3.9)$$

isto é, calculou-se o número de vezes que se rejeitou  $H_0$  em  $N=1000$  simulações, quando o verdadeiro valor de  $\lambda$  foi zero.

Para o coeficiente de regressão do tempo, tem-se as seguintes hipóteses:

$$\begin{cases} H_0 : \gamma = 0 \\ H_1 : \gamma \neq 0; \end{cases} \quad (3.10)$$

então, definiu-se também  $I_i = 1$  se o intervalo HDP de 95% de credibilidade para  $\gamma$  não continha o zero, caso contrário,  $I_i = 0$ . Assim, nos cenários em que se assumiu  $\gamma = 0$ , a taxa de erro tipo I foi calculada com a equação (3.9). No entanto, quando se assumiu  $\gamma = 2EP_\gamma$  ou  $6EP_\gamma$  a equação (3.9) proporciona o poder do teste, uma vez que deve ser rejeitada  $H_0$  quando, de fato, for falsa.

Com a finalidade de verificar se as estimativas das taxas de erros tipo I, dos testes para as diferenças dos efeitos *carryover* e do tempo, são estatisticamente iguais ao nível de significância estabelecido ( $\alpha = 5\%$ ), utilizou-se os limites do intervalo de confiança exato de 99% para a proporção  $p_\alpha = 0,05$ , o qual foi calculado através do *software* Sisvar desenvolvido por Ferreira (2011).

Desta forma, assumiu-se que as taxas de erro tipo I que se encontraram dentro do intervalo não foram estatisticamente diferentes ao nível de significância e, nesses casos, o teste foi considerado exato (GIRARDI; CARGNELUTTI; STORCK, 2009). Nos casos em que a taxa de erro tipo I foi estatisticamente superior ao nível nominal, isto é, o valor da taxa de erro tipo I foi maior que o limite superior do intervalo  $IC_{99\%}(p_\alpha)$ , o teste foi considerado liberal (RIBOLDI *et al.*, 2014).

### 3.4.2 Erro Quadrático Médio (EQM)

Segundo Bolfarine e Sandoval (2001), o erro quadrático médio de um estimador  $\hat{\theta}$  do parâmetro populacional  $\theta$  é dado por:

$$EQM(\hat{\theta}) = E[(\hat{\theta} - \theta)^2]; \quad (3.11)$$

isso significa que a equação anterior proporciona o risco quadrático que um estimador de um parâmetro têm a respeito do valor paramétrico.

Neste trabalho calculou-se o *EQM* das estimativas dos parâmetros  $\lambda$ ,  $\tau$  e  $\gamma$ , em cada um dos  $n$  conjuntos de dados, de maneira empírica, através da fórmula:

$$EQM(\hat{\theta}) = \frac{1}{N} \sum_{i=1}^N (\hat{\theta}_i - \theta)^2,$$

sendo  $N$  o número de simulações para cada um dos cenários,  $\hat{\theta}_i$  a  $i$ -ésima estimativa pontual



da média *a posteriori* e,  $\theta$  o valor simulado para o parâmetro.

### 3.4.3 Acurácia e Precisão

Calculou-se a acurácia nas estimativas dos parâmetros  $\lambda$ ,  $\tau$  e  $\gamma$ , com a finalidade de verificar se estimador está estimando o parâmetro, valor este estabelecido na simulação. Desta forma, utilizou-se o intervalo HDP de 95% de credibilidade e, calculou-se a acurácia através da contagem de vezes que o intervalo continha o valor paramétrico estabelecido para cada cenário.

Outra medida de grande importância na análise estatística é a precisão. Esta permite verificar o quão próximo está a estimativa que foi obtida em relação ao valor do parâmetro simulado. Por conseguinte, calculou-se a precisão das estimativas  $\lambda$ ,  $\tau$  e  $\gamma$  através da amplitude dos intervalos HDP de 95% de credibilidade.

## 4 RESULTADOS E DISCUSSÃO

Nesta seção são apresentados e discutidos os resultados obtidos das estimativas dos parâmetros para cada um dos cenários simulados e para o exemplo prático com dados reais. Também apresenta-se os resultados das medidas que avaliaram a simulação realizada, sendo a taxa de erro tipo I e poder dos testes, erro quadrático médio, acurácia e precisão.

### 4.1 ANÁLISE DOS DADOS SIMULADOS

Tendo em vista que em cada cenário simulado foram realizadas 1000 repetições, são apresentadas as médias das estatísticas descritivas para as estimativas *a posteriori* dos seguintes parâmetros: média geral mais o efeito do tratamento  $B$  ( $\mu_B$ ), diferença dos efeitos *carryover* ( $\lambda$ ), diferença dos períodos ( $\pi$ ), diferença dos tratamentos ( $\tau$ ), efeito do tempo ( $\gamma$ ) e a interação tempo por tratamento ( $\gamma\tau$ ). Para as estimativas intervalares dos parâmetros tem-se o valor médio, em 1000 simulações, dos limites do intervalo HDP de 95% de credibilidade *a posteriori*.

Também é dado o valor mínimo e o valor máximo dos resultados obtidos para os critérios de Geweke, Raftery e Lewis e Heidelberger e Welch nas simulações, em cada cenário. Com base nas Tabelas 5, 6, 7 e 8 verificou-se para todos parâmetros do modelo *crossover*  $2 \times 2$  com 3 medidas repetidas, em cada cenário, que não há evidências que indicam a não convergência das cadeias *a posteriori*. Pelo critério de Geweke, tem-se que todos os valores da estatística  $|Z|$  de Geweke são menores que 1,96. Analisando-se o fator de dependência de Raftery e Lewis, observou-se que os valores estão próximos de 1. Pelo critério Heidelberger e Welch, foi constatado que as cadeias são estacionárias (valor-p  $\geq 0,05$ ).

Nas Tabelas 5 e 6 encontram-se os resultados das estimativas e critérios de convergência dos cenários que utilizaram amostras de 20 UE e coeficientes de variação de 5% e 20%, respectivamente. Nas Tabelas 7 e 8 apresentam-se os resultados dos cenários com amostras de 100 unidades experimentais com *cv* de 5% e 20%, respectivamente.

Tabela 5 – Média das estatísticas descritivas das estimativas pontuais e intervalares *a posteriori* e valor mínimo e máximo dos critérios de convergência dos parâmetros do modelo *crossover* 2 × 2 com 3 medidas repetidas, para os cenários com amostra de 20 unidades experimentais e  $cv = 5\%$ , com  $\tau = 1EP$  (C1, C2, C3) e  $\tau = 2EP$  (C4, C5, C6).

Cen.	Par.	VR	Média	Md	LI	LS	Mín. G	Máx. G	Mín. RL	Máx. RL	Mín. HW	Máx. HW
C1	$\mu_B$	265,42	263,96	263,95	259,35	268,30	0,001	1,938	0,951	1,190	0,051	1,000
	$\lambda$	0,00	0,42	0,34	-3,01	3,82	0,003	1,958	0,944	1,110	0,050	1,000
	$\pi$	0,10	0,49	0,45	-2,95	4,00	0,000	1,946	0,951	1,100	0,051	0,997
	$\tau$	1,58	2,64	2,59	-2,56	7,93	0,001	1,950	0,944	1,160	0,051	1,000
	$\gamma$	0,00	0,30	0,30	-1,09	1,72	0,003	1,950	0,944	1,150	0,051	1,000
	$\gamma\tau$	0,00	-0,30	-0,29	-2,23	1,68	0,000	1,947	0,951	1,100	0,051	0,999
C2	$\mu_B$	265,42	262,93	263,00	258,18	267,37	0,000	1,953	0,951	1,320	0,050	0,999
	$\lambda$	0,00	0,39	0,32	-3,08	3,82	0,002	1,957	0,944	1,100	0,051	1,000
	$\pi$	0,10	0,56	0,55	-2,92	4,11	0,002	1,957	0,951	1,100	0,050	1,000
	$\tau$	1,58	2,77	2,80	-2,57	8,18	0,001	1,919	0,951	1,100	0,051	1,000
	$\gamma$	1,00	1,26	1,27	-0,14	2,69	0,007	1,955	0,944	1,120	0,051	0,998
	$\gamma\tau$	0,00	-0,32	-0,30	-2,25	1,69	0,002	1,943	0,944	1,120	0,052	0,998
C3	$\mu_B$	265,42	260,93	260,97	256,22	265,36	0,000	1,939	0,944	1,230	0,050	0,999
	$\lambda$	0,00	0,38	0,35	-3,09	3,83	0,000	1,949	0,944	1,120	0,050	1,000
	$\pi$	0,10	0,47	0,42	-3,01	4,04	0,000	1,956	0,944	1,120	0,052	0,999
	$\tau$	1,58	2,73	2,86	-2,51	8,09	0,004	1,937	0,951	1,130	0,050	0,999
	$\gamma$	3,00	3,32	3,32	1,93	4,73	0,001	1,959	0,951	1,130	0,050	1,000
	$\gamma\tau$	0,00	-0,30	-0,28	-2,22	1,68	0,000	1,957	0,944	1,120	0,050	1,000
C4	$\mu_B$	263,84	262,41	262,47	257,71	266,82	0,000	1,937	0,951	1,210	0,050	0,998
	$\lambda$	0,00	0,40	0,44	-3,04	3,82	0,004	1,953	0,944	1,110	0,050	1,000
	$\pi$	0,10	0,52	0,49	-2,93	4,05	0,001	1,949	0,944	1,120	0,050	0,998
	$\tau$	3,16	4,24	4,26	-0,98	9,57	0,001	1,953	0,944	1,120	0,051	0,999
	$\gamma$	0,00	0,28	0,27	-1,11	1,70	0,000	1,960	0,944	1,140	0,051	0,999
	$\gamma\tau$	0,00	-0,30	-0,29	-2,23	1,68	0,003	1,948	0,951	1,150	0,051	0,999
C5	$\mu_B$	263,84	261,19	261,17	256,49	265,61	0,000	1,946	0,951	1,190	0,050	0,997
	$\lambda$	0,00	0,49	0,54	-2,94	3,90	0,000	1,960	0,951	1,100	0,051	1,000
	$\pi$	0,10	0,43	0,40	-3,02	3,95	0,001	1,950	0,951	1,130	0,050	1,000
	$\tau$	3,16	4,41	4,39	-0,85	9,75	0,000	1,953	0,944	1,100	0,051	0,999
	$\gamma$	1,00	1,35	1,34	-0,05	2,77	0,000	1,952	0,944	1,120	0,051	0,999
	$\gamma\tau$	0,00	-0,30	-0,28	-2,23	1,68	0,003	1,932	0,944	1,100	0,054	0,998
C6	$\mu_B$	263,84	259,37	259,33	254,68	263,82	0,001	1,945	0,951	1,210	0,050	1,000
	$\lambda$	0,00	0,32	0,32	-3,12	3,75	0,001	1,960	0,944	1,120	0,051	0,999
	$\pi$	0,10	0,52	0,52	-2,93	4,06	0,001	1,925	0,944	1,080	0,051	0,998
	$\tau$	3,16	4,26	4,21	-1,00	9,61	0,000	1,937	0,951	1,120	0,050	1,000
	$\gamma$	3,00	3,29	3,29	1,90	4,72	0,001	1,956	0,944	1,130	0,050	1,000
	$\gamma\tau$	0,00	-0,30	-0,29	-2,24	1,69	0,000	1,953	0,944	1,130	0,050	0,999

Fonte: Da autora.

Nota: Par - Parâmetro, Cen - cenário, VR - valor real, Md - mediana, LI - limite inferior do HPD de 95% de credibilidade, LS - limite superior do HPD de 95%, Mín. G - valor mínimo do critério Geweke, Máx. G - valor máximo do Geweke, Mín. RL - valor mínimo do critério Raftery e Lewis, Máx. RL - valor máximo do critério Raftery e Lewis, Mín. HW - valor mínimo do critério Heildelberg e Welch e, Máx. HW - valor máximo do critério Heildelberg e Welch.

Nas Tabelas 5, 6, 7 e 8 pode ser verificado que o valor médio dos valores estimados para os parâmetros se encontram próximos aos valores reais simulados. Através da média dos limites dos intervalos HPD de 95% de credibilidade dos parâmetros, nas 1000 simulações em cada um dos cenários, constatou-se que o efeito da média geral adicionado o efeito do tratamento *B* foi significativo, sendo portanto diferente de zero.

Além disso, não houve significância para a diferença dos efeitos *carryover* ( $\lambda$ ), nem para a diferença dos efeitos dos períodos ( $\pi$ ), como também para a interação de tempo por tratamento ( $\gamma\tau$ ), dado que o zero está contido na média dos limites dos intervalos; conforme esperado, já

que foram condições estabelecidas nas simulações.

Tabela 6 – Média das Estatísticas descritivas das estimativas pontuais e intervalares *a posteriori* e valor mínimo e máximo dos critérios de convergência dos parâmetros do modelo *crossover*  $2 \times 2$  com 3 medidas repetidas, para os cenários com amostra de 20 unidades experimentais e  $cv = 20\%$ , com  $\tau = 1EP$  (C7, C8, C9) e  $\tau = 2EP$  (C10, C11, C12).

Cen.	Par.	VR	Média	Md	LI	LS	Mín. G	Máx. G	Mín. RL	Máx. RL	Mín. HW	Máx. HW
C7	$\mu_B$	260,68	248,77	248,86	232,48	263,74	0,000	1,948	0,951	2,770	0,053	1,000
	$\lambda$	0,00	5,23	5,10	-7,71	18,45	0,000	1,952	0,951	1,120	0,050	1,000
	$\pi$	0,10	5,47	5,60	-7,64	18,83	0,000	1,960	0,944	1,140	0,051	1,000
	$\tau$	6,32	13,14	13,13	-0,96	27,82	0,002	1,960	0,944	1,120	0,051	0,998
	$\gamma$	0,00	0,28	0,28	-1,21	1,76	0,001	1,957	0,944	1,120	0,050	1,000
	$\gamma\tau$	0,00	-0,29	-0,26	-2,32	1,74	0,000	1,942	0,944	1,120	0,050	0,999
C8	$\mu_B$	260,68	247,93	247,81	231,63	262,93	0,000	1,912	0,951	1,120	0,052	0,998
	$\lambda$	0,00	5,44	5,46	-7,52	18,66	0,002	1,946	0,951	1,150	0,052	0,994
	$\pi$	0,10	4,89	5,05	-8,21	18,27	0,002	1,951	0,944	1,130	0,051	0,998
	$\tau$	6,32	12,91	13,02	-1,15	27,52	0,000	1,945	0,944	1,140	0,050	1,000
	$\gamma$	1,00	1,26	1,27	-0,20	2,72	0,001	1,958	0,944	1,120	0,051	0,999
	$\gamma\tau$	0,00	-0,27	-0,25	-2,26	1,73	0,000	1,956	0,951	1,180	0,051	0,998
C9	$\mu_B$	260,68	245,62	245,53	229,18	260,74	0,000	1,936	0,951	1,140	0,051	1,000
	$\lambda$	0,00	5,59	5,49	-7,45	18,93	0,001	1,938	0,944	1,100	0,052	1,000
	$\pi$	0,10	5,72	5,86	-7,49	19,21	0,000	1,958	0,944	1,100	0,051	0,999
	$\tau$	6,32	12,53	12,50	-1,57	27,21	0,000	1,958	0,944	1,100	0,051	1,000
	$\gamma$	3,00	3,23	3,23	1,77	4,69	0,000	1,955	0,951	1,100	0,051	0,999
	$\gamma\tau$	0,00	-0,27	-0,25	-2,27	1,73	0,001	1,951	0,944	1,120	0,050	0,998
C10	$\mu_B$	254,35	242,78	243,27	226,81	257,48	0,002	1,958	0,951	1,130	0,050	1,000
	$\lambda$	0,00	5,18	4,97	-7,62	18,28	0,001	1,957	0,951	1,110	0,050	0,999
	$\pi$	0,10	5,38	5,38	-7,59	18,61	0,003	1,955	0,944	1,100	0,050	0,999
	$\tau$	12,65	18,61	18,74	4,69	33,07	0,000	1,955	0,951	1,140	0,050	1,000
	$\gamma$	0,00	0,24	0,22	-1,24	1,71	0,001	1,955	0,944	1,140	0,050	0,999
	$\gamma\tau$	0,00	-0,26	-0,25	-2,29	1,76	0,002	1,957	0,951	1,110	0,050	1,000
C11	$\mu_B$	254,35	241,87	242,01	226,00	256,51	0,001	1,954	0,944	1,120	0,054	1,000
	$\lambda$	0,00	5,03	4,96	-7,80	18,16	0,001	1,922	0,944	1,120	0,051	1,000
	$\pi$	0,10	5,19	5,39	-7,85	18,42	0,000	1,940	0,951	1,120	0,050	0,998
	$\tau$	12,65	18,63	18,76	4,74	33,08	0,000	1,944	0,951	1,100	0,050	0,998
	$\gamma$	1,00	1,27	1,26	-0,20	2,72	0,000	1,958	0,944	1,120	0,050	1,000
	$\gamma\tau$	0,00	-0,24	-0,23	-2,24	1,76	0,001	1,960	0,944	1,130	0,051	0,999
C12	$\mu_B$	254,35	240,33	240,81	224,52	254,95	0,000	1,959	0,951	1,130	0,050	1,000
	$\lambda$	0,00	5,16	5,31	-7,56	18,21	0,002	1,928	0,944	1,120	0,050	0,999
	$\pi$	0,10	4,82	4,77	-8,09	17,99	0,008	1,955	0,951	1,140	0,051	0,999
	$\tau$	12,65	18,44	18,24	4,65	32,80	0,005	1,958	0,951	1,140	0,050	0,999
	$\gamma$	3,00	3,27	3,31	1,80	4,74	0,000	1,955	0,951	2,620	0,051	0,999
	$\gamma\tau$	0,00	-0,26	-0,25	-2,28	1,76	0,002	1,954	0,944	1,100	0,051	0,999

Fonte: Da autora.

Nota: Par - parâmetro, Cen - cenário, VR - valor real, Md - mediana, LI - limite inferior do HPD de 95% de credibilidade, LS - limite superior do HPD de 95%, Mín. G - valor mínimo do critério Geweke, Máx. G - valor máximo do Geweke, Mín. RL - valor mínimo do critério Raftery e Lewis, Máx. RL - valor máximo do critério Raftery e Lewis, Mín. HW - valor mínimo do critério Heildelberg e Welch e, Máx. HW - valor máximo do critério Heildelberg e Welch.

Nos cenários C3, C6, C9, C12 ( $n = 20$ ) e C15, C18, C21, C24 ( $n = 100$ ), em que foi simulado um efeito de tempo linear correspondente a  $6EP_\gamma$ , a variável tempo apresentou efeito significativo. Entretanto, nos cenários C2, C5, C8, C11 ( $n = 20$ ) e C14, C17, C20, C23 ( $n = 100$ ), em que se simulou efeito de tempo de  $2EP_\gamma$  este foi, em média, não significativo.

Também, pode-se constatar que, embora tenha sido simulado uma diferença para os efeitos dos tratamentos de  $1EP$ , essa diferença não foi estatisticamente significativa em nenhum desses cenários (C1, C2, C3, C7, C8, C9, C13, C14, C15, C19, C20 e C21). No entanto,

quando se simulou  $\tau = 2EP$  não foi verificado efeito significativo nas médias dos resultados dos cenários com  $cv = 5\%$  (C4, C5, C6, C16, C17 e C18), porém houve significância para a diferença dos tratamentos nos cenários com  $cv = 20\%$  (C10, C11, C12, C22, C23 e C24). Este resultado deve-se a que o erro padrão depende da variabilidade e esta sendo maior implica em maiores valores de erros padrão e consequentemente maior diferença dos tratamentos simulada.

Tabela 7 – Média das Estatísticas descritivas das estimativas pontuais e intervalares *a posteriori* e valor mínimo e máximo dos critérios de convergência dos parâmetros do modelo *crossover*  $2 \times 2$  com 3 medidas repetidas, para os cenários com amostra de 100 UE e  $cv = 5\%$ , com  $\tau = 1EP$  (C13, C14, C15) e  $\tau = 2EP$  (C16, C17, C18).

Cen.	Par.	VR	Média	Md	LI	LS	Mín. G	Máx. G	Mín. RL	Máx. RL	Mín. HW	Máx. HW
C13	$\mu_B$	266,29	265,95	265,98	263,98	267,87	0,001	1,938	0,951	1,370	0,052	0,998
	$\lambda$	0,00	0,10	0,07	-1,48	1,64	0,001	1,958	0,944	1,160	0,050	1,000
	$\pi$	0,10	0,21	0,21	-1,35	1,77	0,001	1,958	0,944	1,100	0,050	1,000
	$\tau$	0,71	0,91	0,94	-1,42	3,23	0,000	1,955	0,951	1,120	0,052	1,000
	$\gamma$	0,00	0,07	0,07	-0,58	0,67	0,000	1,958	0,951	1,140	0,050	0,999
	$\gamma\tau$	0,00	-0,06	-0,06	-0,92	0,82	0,002	1,950	0,944	1,160	0,050	0,999
	C14	$\mu_B$	266,29	265,62	265,64	263,65	267,54	0,000	1,935	0,958	1,330	0,050
$\lambda$		0,00	0,06	0,04	-1,52	1,60	0,004	1,950	0,951	1,130	0,051	0,998
$\pi$		0,10	0,16	0,14	-1,40	1,73	0,000	1,959	0,951	1,100	0,050	1,000
$\tau$		0,71	0,93	0,93	-1,40	3,24	0,000	1,958	0,951	1,120	0,052	1,000
$\gamma$		0,44	0,48	0,49	-0,16	1,09	0,004	1,952	0,951	1,140	0,050	0,999
$\gamma\tau$		0,00	-0,06	-0,06	-0,92	0,82	0,002	1,957	0,951	1,100	0,051	0,999
C15		$\mu_B$	266,29	264,62	264,60	262,65	266,54	0,002	1,954	0,958	1,350	0,050
	$\lambda$	0,00	0,07	0,09	-1,50	1,61	0,000	1,959	0,944	1,140	0,051	1,000
	$\pi$	0,10	0,20	0,21	-1,35	1,77	0,001	1,945	0,951	1,110	0,052	0,999
	$\tau$	0,71	0,92	0,94	-1,41	3,23	0,002	1,947	0,951	1,120	0,051	1,000
	$\gamma$	1,32	1,40	1,39	0,76	2,01	0,005	1,958	0,944	1,120	0,050	0,996
	$\gamma\tau$	0,00	-0,06	-0,06	-0,92	0,82	0,000	1,942	0,944	1,110	0,052	0,998
	C16	$\mu_B$	265,59	265,29	265,23	263,32	267,21	0,000	1,958	0,951	1,370	0,051
$\lambda$		0,00	0,08	0,09	-1,50	1,62	0,001	1,951	0,951	1,130	0,052	0,999
$\pi$		0,10	0,17	0,18	-1,38	1,74	0,002	1,955	0,951	1,100	0,051	1,000
$\tau$		1,41	1,62	1,61	-0,71	3,94	0,001	1,957	0,944	1,100	0,050	1,000
$\gamma$		0,00	0,06	0,07	-0,58	0,67	0,000	1,948	0,951	1,120	0,050	0,999
$\gamma\tau$		0,00	-0,06	-0,06	-0,92	0,82	0,000	1,958	0,944	1,140	0,051	0,996
C17		$\mu_B$	265,59	264,85	264,87	262,88	266,76	0,004	1,959	0,951	1,360	0,050
	$\lambda$	0,00	0,08	0,10	-1,49	1,62	0,001	1,945	0,944	1,160	0,050	0,999
	$\pi$	0,10	0,21	0,24	-1,34	1,77	0,001	1,953	0,944	1,100	0,052	1,000
	$\tau$	1,41	1,60	1,59	-0,73	3,92	0,007	1,952	0,944	1,090	0,050	1,000
	$\gamma$	0,44	0,51	0,52	-0,13	1,12	0,003	1,944	0,944	1,120	0,050	0,999
	$\gamma\tau$	0,00	-0,06	-0,06	-0,92	0,82	0,003	1,926	0,951	2,700	0,051	0,998
	C18	$\mu_B$	265,59	263,97	263,98	262,00	265,88	0,000	1,924	0,958	1,330	0,051
$\lambda$		0,00	0,09	0,10	-1,49	1,63	0,001	1,959	0,944	1,130	0,050	0,999
$\pi$		0,10	0,22	0,22	-1,33	1,79	0,001	1,949	0,944	1,110	0,052	0,998
$\tau$		1,41	1,63	1,63	-0,69	3,94	0,002	1,952	0,944	1,100	0,050	1,000
$\gamma$		1,32	1,36	1,36	0,72	1,97	0,003	1,954	0,944	1,120	0,050	1,000
$\gamma\tau$		0,00	-0,06	-0,06	-0,92	0,82	0,002	1,959	0,944	1,120	0,050	0,997

Fonte: Da autora.

Nota: Par - parâmetros, Cen - cenário, VR - valor real, Md - mediana, LI - limite inferior do HPD de 95% de credibilidade, LS - limite superior do HPD de 95%, Mín. G - valor mínimo do critério Geweke, Máx. G - valor máximo do Geweke, Mín. RL - valor mínimo do critério Raftery e Lewis, Máx. RL - valor máximo do critério Raftery e Lewis, Mín. HW - valor mínimo do critério Heildelberg e Welch e, Máx. HW - valor máximo do critério Heildelberg e Welch.

Tabela 8 – Média das Estatísticas descritivas das estimativas pontuais e intervalares *a posteriori* e valor mínimo e máximo dos critérios de convergência dos parâmetros do modelo *crossover*  $2 \times 2$  com 3 medidas repetidas, para os cenários com amostra de 100 UE e  $cv = 20\%$ , com  $\tau = 1EP$  (C19, C20, C21) e  $\tau = 2EP$  (C22, C23, C24).

Cen.	Par.	VR	Média	Md	LI	LS	Mín. G	Máx. G	Mín. RL	Máx. RL	Mín. HW	Máx. HW
C19	$\mu_B$	264,17	262,14	262,24	256,21	267,89	0,002	1,951	0,951	1,220	0,050	0,999
	$\lambda$	0,00	0,80	0,86	-4,75	6,24	0,000	1,959	0,951	1,100	0,050	0,998
	$\pi$	0,10	1,11	1,05	-4,41	6,52	0,002	1,951	0,944	1,130	0,050	0,999
	$\tau$	2,83	4,01	4,04	-1,87	10,23	0,000	1,957	0,951	1,120	0,050	0,999
	$\gamma$	0,00	0,06	0,07	-0,57	0,68	0,001	1,953	0,951	1,100	0,050	0,996
	$\gamma\tau$	0,00	-0,04	-0,04	-0,96	0,83	0,003	1,956	0,951	1,120	0,050	1,000
C20	$\mu_B$	264,17	261,73	261,70	255,77	267,52	0,000	1,958	0,951	1,170	0,051	0,999
	$\lambda$	0,00	0,81	0,84	-4,75	6,28	0,000	1,959	0,951	2,700	0,050	0,999
	$\pi$	0,10	1,07	1,12	-4,49	6,48	0,000	1,951	0,944	1,100	0,050	0,999
	$\tau$	2,83	3,99	4,15	-1,90	10,24	0,000	1,954	0,951	1,140	0,052	0,998
	$\gamma$	0,44	0,50	0,50	-0,13	1,11	0,000	1,941	0,944	1,100	0,051	1,000
	$\gamma\tau$	0,00	-0,04	-0,04	-0,96	0,83	0,001	1,953	0,944	1,100	0,050	0,999
C21	$\mu_B$	264,17	260,77	260,83	254,86	266,56	0,002	1,960	0,951	1,210	0,052	0,999
	$\lambda$	0,00	0,78	0,77	-4,78	6,23	0,000	1,958	0,944	1,120	0,051	0,998
	$\pi$	0,10	1,20	1,36	-4,33	6,61	0,005	1,953	0,944	1,110	0,050	0,999
	$\tau$	2,83	3,92	3,89	-1,95	10,15	0,000	1,958	0,944	1,100	0,050	0,999
	$\gamma$	1,32	1,36	1,36	0,73	1,98	0,001	1,935	0,944	1,140	0,050	0,998
	$\gamma\tau$	0,00	-0,04	-0,04	-0,96	0,83	0,000	1,958	0,951	1,120	0,051	0,995
C22	$\mu_B$	261,34	259,47	259,49	253,54	265,24	0,002	1,950	0,951	1,180	0,051	0,998
	$\lambda$	0,00	0,75	0,73	-4,82	6,21	0,000	1,958	0,944	1,140	0,050	1,000
	$\pi$	0,10	1,02	1,02	-4,52	6,44	0,000	1,955	0,944	1,140	0,052	0,999
	$\tau$	5,66	6,67	6,70	0,80	12,90	0,001	1,953	0,944	1,100	0,054	0,994
	$\gamma$	0,00	0,04	0,03	-0,59	0,65	0,000	1,960	0,944	2,430	0,050	1,000
	$\gamma\tau$	0,00	-0,04	-0,04	-0,96	0,83	0,000	1,928	0,951	1,100	0,054	0,997
C23	$\mu_B$	261,34	258,79	258,74	252,87	264,55	0,000	1,949	0,944	1,180	0,051	1,000
	$\lambda$	0,00	0,93	0,94	-4,61	6,38	0,000	1,959	0,944	1,100	0,051	0,998
	$\pi$	0,10	1,22	1,24	-4,31	6,62	0,001	1,959	0,951	1,100	0,050	0,999
	$\tau$	5,66	6,71	6,61	0,86	12,94	0,000	1,950	0,944	1,120	0,050	1,000
	$\gamma$	0,44	0,49	0,50	-0,14	1,11	0,000	1,946	0,944	1,120	0,052	0,999
	$\gamma\tau$	0,00	-0,04	-0,04	-0,96	0,83	0,000	1,955	0,951	1,120	0,052	0,998
C24	$\mu_B$	261,34	258,04	257,98	252,12	263,83	0,000	1,950	0,951	1,170	0,051	1,000
	$\lambda$	0,00	0,76	0,82	-4,80	6,22	0,003	1,959	0,951	2,700	0,050	1,000
	$\pi$	0,10	1,04	1,12	-4,50	6,46	0,000	1,946	0,944	1,100	0,050	1,000
	$\tau$	5,66	6,68	6,68	0,80	12,92	0,000	1,954	0,951	1,120	0,051	1,000
	$\gamma$	1,32	1,37	1,37	0,74	1,99	0,000	1,943	0,951	1,120	0,051	1,000
	$\gamma\tau$	0,00	-0,04	-0,04	-0,96	0,83	0,000	1,960	0,944	1,110	0,050	0,999

Fonte: Da autora.

Nota: Par - parâmetro, Cen - cenário, VR - valor real, Md - mediana, LI - limite inferior do HPD de 95% de credibilidade, LS - limite superior do HPD de 95%, Mín. G - valor mínimo do critério Geweke, Máx. G - valor máximo do Geweke, Mín. RL - valor mínimo do critério Raftery e Lewis, Máx. RL - valor máximo do critério Raftery e Lewis, Mín. HW - valor mínimo do critério Heildelberg e Welch e, Máx. HW - valor máximo do critério Heildelberg e Welch.

Por outro lado, considerando o teste das hipóteses (3.4), para a diferença dos efeitos *carryover* ( $\lambda$ ), em que se assumiu uma hipótese de nulidade do efeito ser nulo, nas Tabelas 9 e 10, estão apresentados os resultados das taxas de erros tipo I que foram calculadas para todos os cenários com 20 e 100 unidades experimentais, respectivamente. Além disso, encontram-se os resultados da taxa de erro tipo I do teste das hipóteses (3.10) de efeito do tempo ( $\gamma$ ), somente nos cenários C1, C4, C7, C10 ( $n = 20$ ) e C13, C16, C19, C22 ( $n = 100$ ), nos quais foi simulado a não existência do efeito de tempo. O poder do teste para o tempo está dado nos cenários C2, C3, C5, C6, C8, C9, C11, C12 ( $n = 20$ ) e C14, C15, C17, C18, C20, C21, C23, C24 ( $n = 100$ ),

já que nestas situações foi simulado um efeito de tempo linear de  $r \cdot EP_\gamma$ , com  $r = 2$  ou  $6$ .

Tabela 9 – Taxas de erro tipo I dos testes da diferença dos efeitos *carryover* ( $\lambda$ ) e efeito tempo ( $\gamma$ ) nos cenários em que assumiu-se que esses efeitos são nulos e, poder do teste de  $\gamma$  para os cenários em que foi simulado um efeito de tempo linear de  $r \cdot EP_\gamma$ , com  $r = 2$  ou  $6$ ; para os cenários com 20 UE, com  $\tau = 1EP$  (C1, C2, C3, C7, C8, C9) e  $\tau = 2EP$  (C4, C5, C6, C10, C11, C12); nos quais assumiu-se  $\lambda_1 = \lambda_2 = 0$ .

Cen. ( $cv = 5\%$ )	Par.	Taxa erro tipo I	Poder	Cen. ( $cv = 20\%$ )	Par.	Taxa erro tipo I	Poder
C1	$\lambda$	0,122	-	C7	$\lambda$	0,132	-
	$\gamma$	0,085	-		$\gamma$	0,080	-
C2	$\lambda$	0,114	-	C8	$\lambda$	0,128	-
	$\gamma$	-	0,423		$\gamma$	-	0,390
C3	$\lambda$	0,104	-	C9	$\lambda$	0,142	-
	$\gamma$	-	0,995		$\gamma$	-	0,991
C4	$\lambda$	0,116	-	C10	$\lambda$	0,130	-
	$\gamma$	0,090	-		$\gamma$	0,068	-
C5	$\lambda$	0,133	-	C11	$\lambda$	0,136	-
	$\gamma$	-	0,477		$\gamma$	-	0,399
C6	$\lambda$	0,116	-	C12	$\lambda$	0,134	-
	$\gamma$	-	0,994		$\gamma$	-	0,989

Fonte: Da autora.

Nota: '-' a medida não foi calculada.

Na Tabela 9, pode-se observar que as taxas de erro tipo I da diferença dos efeitos *carryover* para os cenários com  $cv = 5\%$  e  $20\%$  são similares. Considerando o intervalo de confiança exato (99%) para a proporção  $p_\alpha = 0,05$ ,  $IC_{99\%}(p_\alpha) = [0,034; 0,068]$ , verificou-se que todas as taxas de erros tipo I para  $\lambda$  são estatisticamente superiores ao nível de significância de 5%, já que os valores foram maiores que o limite superior do intervalo. Desse modo, o teste para a diferença dos efeitos *carryover* foi liberal para os cenários com 20 UE.

No cenário C10 a taxa de erro tipo I para o teste de hipóteses do efeito do tempo foi de 6,8%, ou seja, não foi estatisticamente diferente ao nível de significância de 5% por ser um valor que não difere estatisticamente do possível valor paramétrico  $IC_{99\%}(p_\alpha)$ , indicando que o teste foi exato nesse cenário. Já para os cenários C1, C4, e C7 o teste comportou-se de maneira liberal.

Desses resultados ressalta-se que as taxas de erro tipo I do teste da diferença dos efeitos *carryover* com amostras de 20 UE foram todos maiores a 10%, que foi o nível de significância sugerido por Grizzle (1965) para realizar o teste do efeito *carryover*, em duas etapas, em que, se a hipótese nula de que a diferença dos efeitos *carryover* é igual a zero for rejeitada, deve-se

apenas utilizar os dados do primeiro período para a análise, como nos delineamentos de grupos paralelos (JONES; KENWARD, 2014).

Embora os resultados aqui apresentados sejam de cunho teórico, fundamentados em simulação de dados com medidas repetidas, há estudos aplicados como o realizado por Yin *et al.* (2017), no qual estudou-se a tendência de tempo da toxicidade de cinco doses de um medicamento utilizando um tamanho amostral de 36 unidades experimentais. Neste estudo, os resultados da taxa de erro tipo I do teste dos efeitos de tempos como medidas repetidas foi menor que 13%.

Ademais, para os cenários C2, C5, C8 e C11, que foram considerados um efeito de tempo de  $2EP_\gamma = 1$ , os resultados de poder foram menores que 50%, indicando que nas 1000 simulações feitas nesses cenários o teste rejeitou a hipótese nula ( $H_0 : \gamma = 0$ ) em um pouco mais de 500 simulações, sendo este um resultado que não era esperado. Por outro lado, para os cenários C3, C6, C9 e C12 em que  $\gamma = 6EP_\gamma = 3$  o poder foi de aproximadamente de 99% como era esperado, por ter considerado uma diferença maior entre os valores de  $\gamma$ .

Tabela 10 – Taxas de erro tipo I dos testes da diferença dos efeitos *carryover* ( $\lambda$ ) e efeito tempo ( $\gamma$ ) nos cenários em que assumiu-se que esses efeitos são nulos e, poder do teste de  $\gamma$  para os cenários em que foi simulado um efeito de tempo linear de  $w \cdot EP$ , com  $w = 2$  ou 6; para os cenários com 100 UE, com  $\tau = 1EP$  (C13, C14, C15, C19, C20, C21) e  $\tau = 2EP$  (C16, C17, C18, C22, C23, C24); nos quais assumiu-se  $\lambda_1 = \lambda_2 = 0$ .

Cen. ( $cv = 5\%$ )	Par.	Taxa erro tipo I	Poder	Cen. ( $cv = 20\%$ )	Par.	Taxa erro tipo I	Poder
C13	$\lambda$	0,089	-	C19	$\lambda$	0,075	-
	$\gamma$	0,058	-		$\gamma$	0,052	-
C14	$\lambda$	0,066	-	C20	$\lambda$	0,054	-
	$\gamma$	-	0,307		$\gamma$	-	0,346
C15	$\lambda$	0,078	-	C21	$\lambda$	0,057	-
	$\gamma$	-	0,995		$\gamma$	-	0,994
C16	$\lambda$	0,085	-	C22	$\lambda$	0,065	-
	$\gamma$	0,049	-		$\gamma$	0,054	-
C17	$\lambda$	0,071	-	C23	$\lambda$	0,065	-
	$\gamma$	-	0,340		$\gamma$	-	0,330
C18	$\lambda$	0,077	-	C24	$\lambda$	0,061	-
	$\gamma$	-	0,987		$\gamma$	-	0,988

Fonte: Da autora.

Nota: '-' a medida não foi calculada.

Na Tabela 10, De acordo com o intervalo  $IC_{99\%}(p_\alpha) = [0,034; 0,068]$ , tem-se que com uma amostra de  $n = 100$  o teste para  $\lambda$  foi exato nos cenários C14, C20, C21, C22, C23 e



C24, dado que os valores da taxa de erro tipo I não são estatisticamente diferentes do possível valor paramétrico. Além disso, contrastando os resultados para os cenários de amostras de  $n = 20$ , pode-se observar que com amostras de 100 UE todas as taxas de erro tipo I para o teste das hipóteses do *carryover* foram valores inferiores a 10%, o qual foi o nível de significância sugerido por Grizzle (1965).

De acordo com o  $IC_{99\%}(p_\alpha)$ , também verifica-se que o teste de hipótese do efeito de tempo, foi exato em todos os cenários em que simulou-se  $\gamma = 0$  (C13, C16, C19 e C22). Em um estudo aplicado realizado por Li *et al.* (2020), no qual foi feita uma simulação com um delineamento *crossover*  $4 \times 4$  com medidas repetidas, utilizando tamanhos amostrais maiores e iguais a 100 unidades experimentais, encontrou que as taxas de erro tipo I para o teste do efeito do tempo se encontram próximas ao nível de significância de 5%.

De maneira similar aos cenários com  $n = 20$ , o poder do teste para o efeito do tempo foi maior a 98% nos casos em que simulou-se  $\gamma = 6EP_\gamma$ , a dizer, nos cenários C15, C18, C21 e C24. Estes resultados evidenciam que o tamanho da amostra não influenciou o poder do teste para  $\gamma$  quando existe efeito de tempo, contudo, observou-se aumento do poder quando o efeito do tempo considerado foi maior.

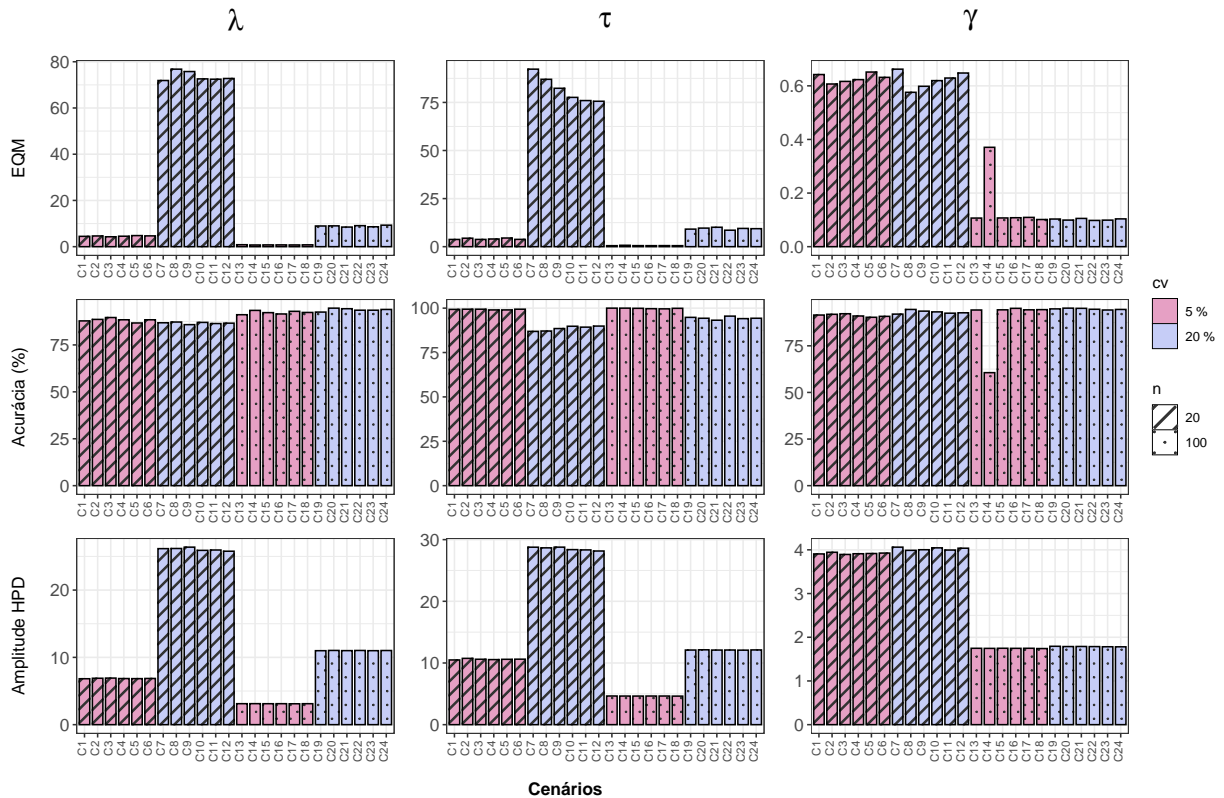
Pode-se destacar, que em relação ao erro tipo I e poder dos testes da diferença dos efeitos *carryover* e efeitos do tempo, obteve-se melhores resultados nos cenários com amostras de 100 UE e  $\gamma = 6EP_\gamma$ , ressaltando os cenários com  $cv = 20\%$  (C21 e C24).

Adicionalmente, na Figura 4 encontram-se os resultados das medidas de erro quadrático médio, acurácia e amplitude do intervalo HPD de 95%, através da qual verificou-se a precisão, nas estimativas dos parâmetros da diferença dos efeitos *carryover*, diferença de tratamentos e efeitos de tempo, em cada um dos cenários com amostras de 20 e 100 unidades experimentais.

Nota-se que nos cenários em que o coeficiente de variação foi de 5% obteve-se menor erro quadrático médio, como também, menor amplitude dos intervalos HPD 95% de credibilidade para as estimativas dos parâmetros  $\lambda$  e  $\tau$ , do que nos cenários em que foram utilizados coeficientes de variação de 20%. Já para o efeito do tempo, o erro quadrático médio e a amplitude dos intervalos HPD foram similares em todos os cenários com o mesmo tamanho amostral; sendo de aproximadamente 0,60 e 4, respectivamente, para os cenários com  $n = 20$  e, o EQM foi 0,11 e a amplitude do intervalo HPD de 1,75 para os cenários com  $n = 100$ .

Quanto à acurácia, obteve-se um maior porcentagem para os parâmetros  $\lambda$  e  $\gamma$  nos cenários com  $cv = 20\%$ . Para a diferença dos efeitos dos tratamentos ( $\tau$ ) a maior porcentagem

Figura 4 – Medidas de erro quadrático médio (EQM), acurácia (%) e precisão (amplitude do intervalo HPD) nas estimativas das diferenças dos efeitos *carryover* ( $\lambda$ ), diferenças dos tratamentos ( $\tau$ ) e efeitos do tempo ( $\gamma$ ), do modelo *crossover*  $2 \times 2$  com 3 medidas repetidas, para os cenários  $n = 20$ , com  $\tau = 1EP$  (C1, C2, C3, C7, C8, C9) e  $\tau = 2EP$  (C4, C5, C6, C10, C11, C12) e cenários com  $n = 100$ , com  $\tau = 1EP$  (C13, C14, C15, C19, C20, C21) e  $\tau = 2EP$  (C16, C17, C18, C22, C23, C24).



Fonte: Da autora.

de acurácia foi para cenários com  $cv = 5\%$ . No entanto, em todos os cenários obteve-se alta acurácia, sendo superiores a 85%.

Além disso, os resultados para os cenários com mesmo coeficiente de variação foram similares quando se consideram a diferença dos tratamentos  $\tau = 1EP$  ou  $\tau = 2EP$ , este resultado se deve ao fato de que o erro padrão entre as médias não foram valores elevados, sendo de 1,58 para os cenários com  $cv = 5\%$  e 6,32 para os cenários com  $cv = 20\%$ . Assim, espera-se que há maior diferença nas estimativas dos parâmetros para cenários em que a diferença entre os  $\tau$  simulados sejam maiores.

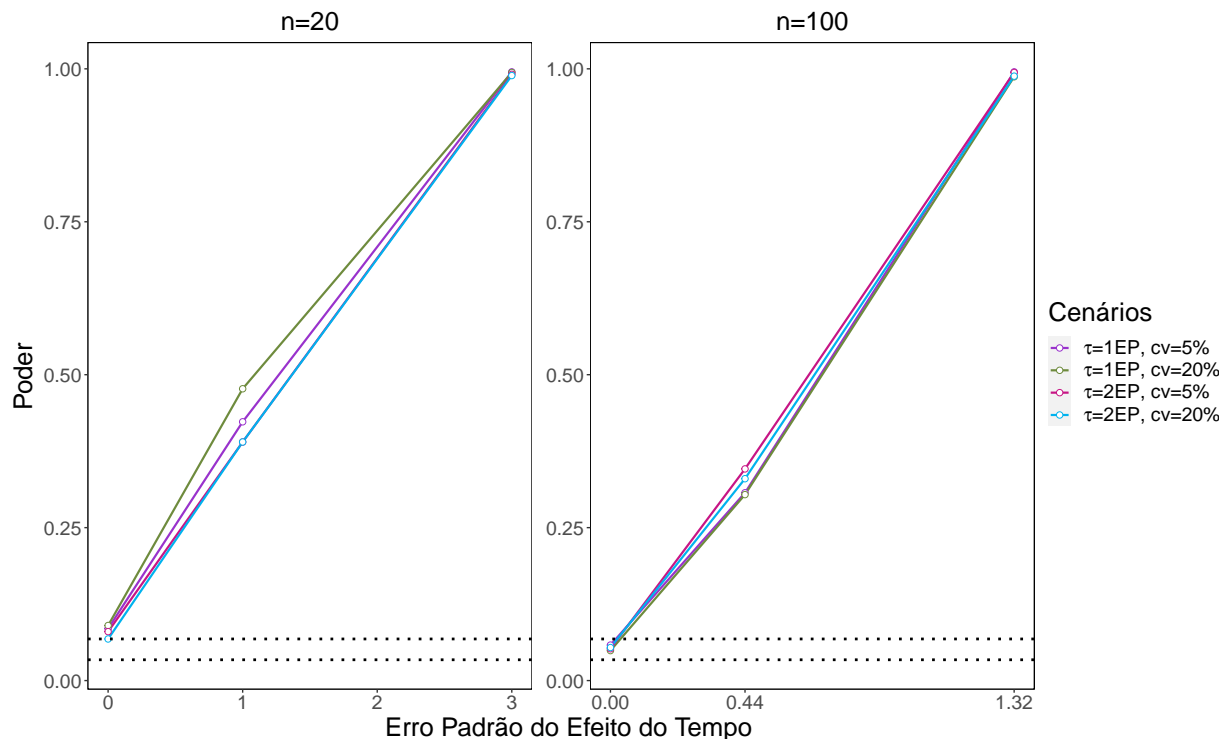
Sendo assim, ao se comparar os resultados com distinto tamanho amostral, na Figura 4, pode-se afirmar que os valores obtidos para as medidas de análise das estimativas dos parâmetros do modelo *crossover* com 3 medidas repetidas, como EQM, acurácia e precisão, foram melhores para os cenários com amostras de 100 UE do que com amostras de 20 UE, dado que obteve-se menor erro quadrático médio e maior porcentagem de acurácia, como

também menor amplitude do intervalo HPD de 95% de credibilidade nas estimativas dos parâmetros da diferença dos efeitos dos tratamentos, diferença nos efeitos *carryover* e efeitos de tempo.

No entanto, obteve-se bons resultados das medidas de análises das estimativas para amostras de 20 UE. Destacando uma das vantagens dos delineamentos *crossover*, precisa-se de um menor número de UE que em outros delineamentos para obter um poder de teste alto e maior precisão das estimativas (JONES; KENWARD, 2014; LARSEN *et al.*, 2018).

Ademais, em todos os cenários com amostras de 100 UE houve menor taxa de erro tipo I, para todos os parâmetros. Os resultados do poder para o teste de efeito do tempo foram similares, tanto para os cenários com amostras de 20 UE quanto para cenários com 100 UE, como se apresenta na Figura 5. Além disso, Carriere e Huang (2000) destacam que os delineamentos *crossover* proporcionam um maior poder do teste, devido a que a variabilidade entre as UE é reduzida.

Figura 5 – Gráfico do poder do efeito do tempo ( $\gamma$ ) para os cenários com amostras de 20 e 100 unidades experimentais.



Fonte: Da autora.

Nota: As linhas pontilhadas delimitam o intervalo de confiança exato de 99% para o nível de significância  $\alpha$ .

Note-se que, na Figura 5, as curvas de poder de  $\gamma$  e taxas de Erro tipo I, para os cenários com amostras de 20 UE partem de valores superiores ao limite superior do  $IC_{99\%}(p_\alpha)$ , não

conseguindo garantir o nível de significância de 5% pre estabelecido. Há exceção nos cenários com  $\tau = 2EP$  e  $cv = 20\%$ . Enquanto as curvas de poder de  $\gamma$  para os cenários com 100 UE aumentam com o incremento do efeito de tempo e é considerado que a taxa de erro tipo I foi controlada e portanto o teste se mostrou exato nessas situações de aumento da amostra.

#### 4.1.1 Análises de dados simulados para cenários com $\lambda_1 = \lambda_2 = 4EP$

Na Tabela 11 apresenta-se os valores médios das estimativas dos parâmetros do modelo *crossover*  $2 \times 2$  com 3 medidas repetidas, em que os efeitos *carryover* de cada tratamento foram simulados com  $4EP$  da média dos tratamentos, como descrito na Tabela 3,  $\lambda_1 = \lambda_2 = 6,325$ ;  $\lambda_1 = \lambda_2 = 25,29$ ;  $\lambda_1 = \lambda_2 = 5,657$ ;  $\lambda_1 = \lambda_2 = 11,314$ , para os cenários C25, C26, C27 e C28, respectivamente. Note que, a diferença entre os efeitos *carryover*, na simulação, é igual a zero ( $\lambda = \lambda_2 - \lambda_1 = 0$ ).

Tabela 11 – Média das Estatísticas descritivas das estimativas pontuais e intervalares *a posteriori* e valor mínimo e máximo dos critérios de convergência dos parâmetros do modelo *crossover*  $2 \times 2$  com 3 medidas repetidas, para os cenários em que  $\lambda_1 = \lambda_2 = 4EP$ , com amostra de 20 unidades experimentais C25 ( $cv = 5\%$ ) e C26 ( $cv = 20\%$ ), e para os cenários com tamanho amostral de 100 C27 ( $cv = 5\%$ ) e C28 ( $cv = 20\%$ ).

Cen	Par	VR	media	med	LI	LS	min G	max G	min RL	max RL	min HW	max HW
C25	$\mu_B$	265,42	260,94	260,99	256,24	265,38	0,001	1,957	0,951	1,230	0,051	0,999
	$\lambda$	0,00	0,38	0,36	-3,10	3,83	0,000	1,949	0,944	1,130	0,050	1,000
	$\pi$	0,10	6,78	6,73	3,29	10,34	0,000	1,956	0,944	1,120	0,052	0,999
	$\tau$	1,58	2,73	2,86	-2,51	8,09	0,003	1,937	0,951	1,130	0,050	0,999
	$\gamma$	3,00	3,32	3,32	1,93	4,73	0,001	1,959	0,951	1,130	0,050	1,000
	$\gamma\tau$	0,00	-0,30	-0,28	-2,22	1,68	0,000	1,957	0,944	1,120	0,050	1,000
	C26	$\mu_B$	260,68	247,95	247,97	232,09	262,61	0,001	1,952	0,944	1,130	0,051
$\lambda$		0,00	5,54	5,31	-7,15	18,52	0,001	1,956	0,944	1,160	0,051	1,000
$\pi$		0,10	30,04	29,81	17,15	43,14	0,001	1,946	0,944	1,110	0,051	0,999
$\tau$		6,32	13,01	12,97	-0,91	27,54	0,002	1,959	0,944	1,120	0,050	0,999
$\gamma$		1,00	1,28	1,29	-0,20	2,76	0,001	1,956	0,944	1,100	0,051	1,000
$\gamma\tau$		0,00	-0,27	-0,26	-2,30	1,75	0,000	1,945	0,944	1,100	0,051	0,998
C27		$\mu_B$	265,59	265,28	265,26	263,31	267,21	0,003	1,947	0,951	1,320	0,051
	$\lambda$	0,00	0,10	0,07	-1,47	1,65	0,000	1,957	0,951	1,120	0,053	0,999
	$\pi$	0,10	5,82	5,84	4,27	7,39	0,001	1,958	0,944	1,100	0,050	0,999
	$\tau$	1,41	1,61	1,62	-0,72	3,93	0,003	1,937	0,944	1,120	0,051	0,999
	$\gamma$	0,00	0,07	0,06	-0,57	0,68	0,000	1,947	0,944	1,120	0,052	1,000
	$\gamma\tau$	0,00	-0,06	-0,06	-0,93	0,83	0,000	1,945	0,944	1,110	0,051	1,000
	C28	$\mu_B$	264,17	260,78	260,76	254,88	266,54	0,004	1,959	0,951	1,190	0,052
$\lambda$		0,00	0,85	0,86	-4,70	6,30	0,002	1,946	0,951	1,120	0,051	1,000
$\pi$		0,10	12,36	12,39	6,82	17,75	0,001	1,950	0,944	1,140	0,051	1,000
$\tau$		2,83	3,87	3,88	-1,99	10,09	0,000	1,956	0,944	1,140	0,050	0,997
$\gamma$		1,32	1,36	1,36	0,73	1,98	0,000	1,955	0,944	1,100	0,052	1,000
$\gamma\tau$		0,00	-0,04	-0,04	-0,96	0,83	0,001	1,958	0,944	1,120	0,051	0,998

Fonte: Da autora.

Nota: Cen - cenário, VR - valor real, Md- mediana, LI - limite inferior do HPD de 95% de credibilidade, LS - limite superior do HPD de 95%, Mfn. G - valor mínimo do critério Geweke, Máx. G - valor máximo do Geweke, Mfn. RL - valor mínimo do critério Raftery e Lewis, Máx. RL - valor máximo do critério Raftery e Lewis, Mfn. HW - valor mínimo do critério Heildelberg e Welch e, Máx. HW - valor máximo do critério Heildelberg e Welch.

Mediante a Tabela 11, pode-se observar que os valores médios das estimativas dos parâmetros estão próximos aos valores reais simulados, a exceção do parâmetro da diferença de efeitos de períodos ( $\pi$ ), a qual foi significativa em todos os cenários. De acordo com Li *et al.* (2020) quando encontram-se um efeito do período significativo a interpretação do modelo pode ser um tanto problemática, no entanto isso não inviabiliza a realização do teste de efeitos de tratamento.

Pode-se notar que as estimativas de  $\pi$  se encontram próximas aos valores simulados para o efeito *carryover* do tratamento recebido no primeiro período,  $\lambda_1$  para a sequência *AB* e  $\lambda_2$  para a sequência *BA* (Tabela 11). Estes resultados provavelmente se devem a que os valores simulados para os efeitos *carryover* ( $\lambda_1 = \lambda_2$ ), diferentes de zero, porém iguais entre si, foram captados pela diferença dos efeitos dos períodos ( $\pi$ ); pois os efeitos *carryover* se encontram no segundo período somente. Segundo Jones e Kenward (2014), não se podem testar efeitos *carryover*  $\lambda_1 = \lambda_2 = 0$  quando existe efeitos de período no modelo. Desta forma pode-se inferir que o efeitos *carryover* se misturam com efeitos de períodos.

Além disso, note que o efeito da diferença dos *carryover* ( $\lambda$ ) foram não significativos em todos os cenários. O efeito da média geral mais o tratamento B foi significativo, em todos os cenários, mantendo o padrão encontrado nas simulações. Note que, como  $\lambda_1 = \lambda_2$ , a estimativa da diferença dos efeitos dos tratamentos não é viesada, sendo próxima ao valor simulado e não significativa em todos os cenários. Também, obteve-se efeitos de tempos significativos nos cenários C25 e C28, nos quais foram considerados efeitos de tempo linear de  $\gamma = 6EP_\gamma$ . Para a interação tempo por tratamento os efeitos foram não significativos.

Nas Tabelas 12 e 13 apresentam-se os resultados das medidas de análise dos resultados, ou seja, as taxas de erros tipo I e poder dos testes da diferença dos efeitos *carryover* ( $\lambda$ ) e efeito tempo ( $\gamma$ ), erro quadrático médio, acurácia e precisão nas estimativas das diferenças dos efeitos *carryover*, diferenças dos tratamentos e efeitos do tempo do modelo.

De acordo com o intervalo  $IC_{99\%}(p_\alpha) = [0,034; 0,068]$ , tem-se que as taxas de erro tipo I são estatisticamente maiores que nível de significância estabelecido (5%) para os testes da diferença dos efeitos *carryover* nos C25, C26 e C27, concluindo-se que nesses casos o teste se mostrou de comportamento liberal. Ademais, os testes da diferença dos efeitos *carryover* e efeitos do tempo foram exatos nos cenários C27 e C28. Note-se também que o poder do teste dos efeitos de tempo foi maior que 98% nos cenários em que simulou-se  $\gamma = P_\gamma$ , isto é, nos cenários C25 e C28.

Tabela 12 – Taxas de erro tipo I dos testes da diferença dos efeitos *carryover* ( $\lambda$ ) e efeito tempo ( $\gamma$ ) nos cenários em que assumiu-se que esses efeitos são nulos e, poder do teste de  $\gamma$  para os cenários em que foi simulado um efeito de tempo linear de  $w \cdot EP$ , com  $w = 2$  ou  $6$ ; para os cenários com amostra de 20 C25 ( $cv = 5\%$ ) e C26 ( $cv = 20\%$ ), e para os cenários com amostras de 100 unidades experimentais C27 ( $cv = 5\%$ ) e C28 ( $cv = 20\%$ ), nos quais assumiu-se  $\lambda_1 = \lambda_2 = 4EP$ .

Cen. (n=20)	Par.	Taxa erro tipo I	Poder	Cen. (n=100)	Par.	Taxa erro tipo I	Poder
C25	$\lambda$	0,105	-	C27	$\lambda$	0,086	-
	$\gamma$	-	0,995		$\gamma$	0,058	-
C26	$\lambda$	0,136	-	C28	$\lambda$	0,061	-
	$\gamma$	-	0,405		$\gamma$	-	0,986

Nota: '-' a medida não foi calculada, n- tamanho da amostra.

Fonte: Da autora.

Tabela 13 – Medidas de erro quadrático médio (EQM), acurácia e precisão (amplitude do intervalo HPD) nas estimativas das diferenças dos efeitos *carryover* ( $\lambda$ ), diferenças dos tratamentos ( $\tau$ ) e efeitos do tempo ( $\gamma$ ), do modelo *crossover*  $2 \times 2$  com 3 medidas repetidas, para os cenários com amostra de 20 unidades experimentais C25 ( $cv = 5\%$ ) e C26 ( $cv = 20\%$ ), e para os cenários com amostras de 100 UE C27 ( $cv = 5\%$ ) e C28 ( $cv = 20\%$ ), nos quais assumiu-se  $\lambda_1 = \lambda_2 = 4EP$ .

Cen. (n=20)	Par.	EQM	Acurácia (%)	Amplitude HPD	Cen. (n=100)	Par.	EQM	Acurácia (%)	Amplitude HPD
C25	$\lambda$	4,29	89,50	6,92	C27	$\lambda$	0,79	91,40	3,11
	$\tau$	3,78	99,40	10,60		$\tau$	0,53	99,70	4,66
	$\gamma$	0,62	92,30	3,89		$\gamma$	0,10	94,20	1,75
C26	$\lambda$	76,85	86,40	25,67	C28	$\lambda$	9,03	93,9	11,00
	$\tau$	88,85	86,90	28,45		$\tau$	8,93	93,9	12,09
	$\gamma$	0,70	90,50	4,05		$\gamma$	0,10	94,2	1,78

Fonte: Da autora.

Na Tabela 13 nota-se que em todos os cenários com amostras de 100 UE obteve-se maior acurácia, menor erro quadrático médio e menor amplitude dos intervalos HPD de 99% de credibilidade para os parâmetros  $\lambda$ ,  $\tau$  e  $\gamma$  do que nos cenários com 20 UE. No entanto, nos cenários com amostras menores também obteve-se resultados satisfatórios em termos de EQM, acurácia e precisão nas estimativas dos parâmetros.

Os resultados obtidos estão em consonância com os resultados esperados; dado que ao compará-los com resultados dos cenários em que simulou-se  $\lambda_1 = \lambda_2 = 0$ , isto é, comparando C3 com C25, C8 com C26, C16 com C27 e C21 com C28, pode-se notar que as estimativas para os parâmetros  $\mu_B$ ,  $\lambda$ ,  $\tau$ ,  $\gamma$  e  $\gamma\tau$  foram semelhantes. Além disso, obteve-se valores semelhantes das taxas de erros tipo I e poder dos testes da diferença dos efeitos *carryover* ( $\lambda$ ) e efeito tempo ( $\gamma$ ), como também das medidas de erro quadrático médio, acurácia e precisão nas estimativas

das diferenças dos efeitos *carryover*, diferenças dos tratamentos e efeitos do tempo do modelo.

Sendo assim, conforme observado na literatura, os efeitos *carryover* afetam na realização das inferências sobre os tratamentos de estudo quando não são iguais ( $\lambda_1 \neq \lambda_2$ ), pois podem haver estimativas viesadas dos demais parâmetros do modelo.

## 4.2 ANÁLISE DOS DADOS REAIS DO EXEMPLO PRÁTICO

Nesta seção são apresentados as estimativas obtidas através da abordagem Bayesiana do modelo *crossover* com 3 medidas repetidas para a análise dos dados da frequência mediana do músculo gastrocnêmio lateral direito, medida nos tempos 0, após 10 e 20 atendimentos dos 38 pacientes voluntários com câncer para o estudo do efeito de um protocolo de *exergaming* na ativação muscular.

Na Tabela 14, verifica-se para todos os parâmetros, que não há evidência que indicam a não convergência da cadeia a posteriori. Pelo critério de Geweke tem-se que o valor absoluto da estatística ( $|z|$ ) é menor que 1,96 para todas as cadeias, indicando que não há indícios de ausência de convergência. Analisando o fator de dependência (FD) de Raftery e Lewis, observa-se que os valores estão próximos de 1, o que indicaria independência entre as iterações ocorridas no processo de amostragem MCMC e portanto uma possível convergência. Pelo critério de Heidelberg-Welch, foi constatado que a série é estacionária, dado que os valores-p foram maiores que nível de significância de 5%.

Tabela 14 – Resultados dos critérios de convergência da cadeia a posteriori para os parâmetros do modelo *crossover* com medidas repetidas para os dados da frequência mediana do músculo Gastrocnêmio lateral direito.

Parâmetros	Geweke ( Z )	Raftery e Lewis (FD)	Heidelberg-Welch (p-valor)
$\mu_B$	0,62	1,02	0,92
$\lambda$	0,07	0,99	0,70
$\pi$	0,55	0,99	0,87
$\tau$	1,17	1,03	0,46
$\gamma$	0,81	0,98	0,45
$\gamma\tau$	0,14	0,97	0,39

Fonte: Da autora.

De acordo com a Tabela 15 pode-se observar que embora a média *a posteriori* da

diferença entre os efeitos *carryover* dos tratamentos foi de 6,49, esta não é estatisticamente significativa dado que o intervalo  $HPD_{95\%}(\lambda) = [-3,93; 16,7]$  contém o zero, portanto, a sequência em que os pacientes receberam os tratamentos não teve influência nos seus efeitos. Da mesma forma, nota-se que não há efeito significativo da diferença entre os períodos ( $\pi$ ), isto implica que a resposta do paciente ao tratamento não foi influenciada pelo período em que o recebeu. Além disso, não há efeito significativo da interação tempo por tratamento ( $\gamma\tau$ ), assim, o efeito de cada tratamento não variou com o tempo. O efeito do tempo ( $\gamma$ ) também foi não significativo, ou seja, a frequência mediana do músculo GLD não apresentou variação ao longo do tempo.

Tabela 15 – Estimativas pontuais e intervalares dos parâmetros, com priors não informativas, para os dados da frequência mediana do músculo Gastrocnêmio lateral direito.

Parâmetros	Média Posteriori	$LI_{HPD95\%}$	$LS_{HPD95\%}$
$\mu_B$	154,15	141,4	166,1
$\lambda$	6,49	-3,93	16,7
$\pi$	3,77	-6,03	13,74
$\tau$	-1,94	-16,67	13,14
$\gamma$	0,51	-2,99	4,33
$\gamma\tau$	5,06	-0,33	10,11

Fonte: Da autora.

Nota:  $LI_{HPD95\%}$  - Limite Inferior do intervalo HPD de 95% de credibilidade;  $LS_{HPD95\%}$  - Limite Superior do intervalo HPD de 95% de credibilidade.

A média *a posteriori* da diferença do tratamento com protocolo de *exergaming* e sem protocolo de *exergaming* foi de -1,94, sendo não significativa de acordo ao intervalo  $HPD_{95\%}(\tau) = [-16,67; 13,14]$ . Sendo assim, pode-se concluir que o efeito do protocolo de *exergaming* não é estatisticamente diferente do tratamento controle. O que não foi possível afirmar sobre a melhora na ativação muscular do músculo gastrocnêmio lateral direito proporcionada pela intervenção.

Por meio das estimativas obtidas para efeitos da diferença dos tratamentos e efeitos do tempo e considerando que o erro padrão dos dados utilizados foi 1,83; este estudo real pode-se comparar aos cenários C1 e C2, nos quais simulou-se  $\tau = 1EP$  e  $\gamma = 0$  e 1, respectivamente. Sendo assim, neste estudo com dados reais pode-se esperar que a taxa de erro tipo I para a diferença dos efeitos *carryover* seja de aproximadamente 10% e poder do teste de efeitos de tempo seja maior do que 50%. Além disso, é esperado que as estimativas *a posteriori* obtidas para os parâmetros  $\lambda$ ,  $\tau$  e  $\gamma$  sejam acuradas, com erros quadráticos médios baixos e boa precisão.



## 5 CONCLUSÕES

Mediante os resultados encontrados, é possível concluir que:

- Sob abordagem Bayesiana, o modelo apresentou bom desempenho em relação à acurácia, erro quadrático médio e precisão das estimativas com tamanho amostral de 20 unidades experimentais. Com tamanho amostral 100, os resultados dessas medidas foram ainda melhores. Além disso, com amostras maiores obtém-se testes exatos da diferença dos efeitos *carryover* e efeitos do tempo. No entanto, o teste do efeito de tempo mostrou-se poderoso mesmo com amostras pequenas.
- Para minimizar as taxas de erros tipo I e obter estimativas mais acuradas e precisas, recomenda-se utilizar um tamanho amostral mínimo de 100 unidades experimentais em experimentos *crossover*  $2 \times 2$  com medidas repetidas.
- Considerar efeitos *carryover* diferentes de zero para cada tratamento,  $\lambda_1 = \lambda_2 = 4EP$ , não interfere na análise da diferença dos efeitos dos tratamentos. Ademais, não há perda da acurácia e precisão das estimativas dos parâmetros  $\lambda$ ,  $\tau$  e  $\gamma$  do modelo. No entanto, Precisa-se ter cuidado especial respeito à interpretação da estimativa da diferença dos períodos, a qual capta os valores dos efeitos *carryover* presentes em cada tratamento.
- Se existirem efeitos do período presentes no modelo e os efeitos *carryover* forem distintos de zero e iguais entre si, torna-se inviável inferir sobre as estimativas dos efeitos de períodos e as estimativas dos efeitos *carryover* para cada tratamento, devido ao confundimento destes dois efeitos importantes.
- A partir da análise dos dados reais usando a metodologia proposta neste trabalho, não foi possível afirmar sobre a melhora na ativação muscular do músculo gastrocnêmio lateral direito proporcionada pela intervenção com o protocolo de *exergaming* nos pacientes com câncer.
- Como trabalhos futuros pretende-se considerar tamanhos de amostras de 30 ou 50 unidades experimentais, distintos valores para a variância entre os erros  $\epsilon_{ijkl}$ , como também diferentes estruturas de covariância. Considerar efeito das unidades experimentais como sendo aleatório no modelo misto também será avaliado nos

próximos trabalhos. Ademais, pode-se implementar o fator de Bayes para realizar as inferências sobre a significância dos parâmetros do modelo e não apenas o intervalo HPD.

## REFERÊNCIAS

- ALVES, R. S. **Análise da fadiga relatada, qualidade de vida, ativação muscular, perfil térmico e equilíbrio estático de voluntários com câncer em quimioterapia submetidos a prática de exergaming: estudo randomizado e crossover.** 2020. Tese (Doutorado) — 2020. 156 f. Tese (Doutorado em Biociências Aplicadas à Saúde) - Universidade Federal de Alfenas.
- BAHADOR, A.; LESAN, S.; KASHI, N. Effect of xylitol on cariogenic and beneficial oral streptococci: a randomized, double-blind crossover trial. **Iranian journal of microbiology**, v. 4, n. 2, p. 75, 2012. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3434645/>>. Acesso em: 27 jun. 2019.
- BOLFARINE, H.; SANDOVAL, M. C. **Introdução à inferência estatística.** [S.l.]: SBM, 2001. v. 2. 159 p.
- BROWN, H.; PRESCOTT, R. **Applied mixed models in medicine.** Chichester: John Wiley & Sons, 2014. 455 p.
- CARRIERE, K.; HUANG, R. Crossover designs for two-treatment clinical trials. **Journal of Statistical Planning and Inference**, S.l., v. 87, n. 1, p. 125–134, 2000. Disponível em: <[https://doi.org/10.1016/S0378-3758\(99\)00185-8](https://doi.org/10.1016/S0378-3758(99)00185-8)>. Acesso em: 12 maio 2019.
- CHOW, S. C.; LUI, J. P. **Design and Analysis of Bioavailability and Bioequivalence Studies.** 3. ed. USA: Chapman & Hall/CRC, 2009. 733 p.
- DEGROOT, M. H.; SCHERVISH, M. J. **Probability and statistics.** Boston: Addison-Wesley, 2012. v. 4. 893 p.
- FERNÁNDEZ, P.; VALLEJO, G. . Diseño de medidas repetidas con dependencia serial en el error bajo la violación de la asunción de homogeneidad. **Anales de Psicología**, v. 12, n. 1, p. 87–106, 1996. Disponível em: <<https://revistas.um.es/analesps/article/view/30281>>. Acesso em: 4 abr. 2020.
- FERREIRA, D. F. Sisvar: a computer statistical analysis system. **Ciência e Agrotecnologia (UFPA)**, v. 35, n. 6, p. 1039–1042, 2011. Disponível em: <<https://doi.org/10.1590/S1413-70542011000600001>>. Acesso em: 17 abr. 2021.
- GAMERMAN, D.; LOPES, H. F. **Markov chain Monte Carlo: stochastic simulation for bayesian inference.** [S.l.]: Boca Raton: CRC Press, 2006. 323 p. Acesso em: 9 fev. 2021.
- GELMAN, A. *et al.* **Bayesian data analysis.** 3. ed. [S.l.]: Chapman and Hall/CRC, 2014. 675 p.
- GEWEKE, J. Evaluating the accuracy of sampling-based approaches to the calculations of posterior moments. **Federal Reserve Bank of Minneapolis**, v. 4, p. 641–649, 1992. Disponível em: <<https://ideas.repec.org/p/fip/fedmsr/148.html>>. Acesso em: 13 fev. 2021.
- GIRARDI, L. H.; CARGNELUTTI, F. A.; STORCK, L. Erro tipo 1 e poder de cinco testes de

comparação múltipla de médias. **Revista Brasileira de Biometria**, São Paulo, v. 27, n. 1, p. 23–36, 2009. Disponível em: <[http://jaguar.fcav.unesp.br/RME/fasciculos/v27/v27\\_n1/A2\\_Alberto.pdf](http://jaguar.fcav.unesp.br/RME/fasciculos/v27/v27_n1/A2_Alberto.pdf)>. Acesso em: 19 abr. 2021.

GOMES, R. P.; DANTAS, F. F. O. Influência do tamanho da massa muscular ativa na hipotensão pós-exercício em idosas hipertensas. **Revista UNI-RN**, Rio Grande do Norte, v. 16-17, p. 228–238, 2017. Disponível em: <<http://revistas.unirn.edu.br/index.php/revistaunirn/article/view/393>>. Acesso em: 16 nov. 2020.

GÓMEZ, S. *et al.* Procedimientos estadísticos más utilizados en el análisis de medidas repetidas en el tiempo en el sector agropecuario. **Revista Cubana de Ciencia Agrícola**, La Habana, v. 46, n. 1, p. 1–7, 2012. Disponível em: <<http://www.redalyc.org/articulo.oa?id=193024313001>>. Acesso em: 20 dez. 2020.

GRIZZLE, J. E. The two-period change-over design and its use in clinical trials. **Biometrics**, Washington, p. 467–480, 1965. Disponível em: <[www.jstor.org/stable/2528104](http://www.jstor.org/stable/2528104)>. Acesso em: 21 jun. 2021.

HEIDELBERGER, P.; WELCH, P. D. Simulation run length control in the presence of an initial transient. **Operations Research**, Catonsville, v. 31, n. 6, p. 1109–1144, 1983. Disponível em: <<https://doi.org/10.1287/opre.31.6.1109>>. Acesso em: 13 fev. 2021.

HERMANS, L. *et al.* Influence of morphine and naloxone on pain modulation in rheumatoid arthritis, chronic fatigue syndrome/fibromyalgia, and controls: a double-blind, randomized, placebo-controlled, cross-over study. **Pain Practice**, USA, v. 18, n. 4, p. 418–430, 2018. Disponível em: <<https://doi.org/10.1111/papr.12613>>. Acesso em: 25 jan. 2020.

ISLAM, M. A.; CHOWDHURY, R. I. **Analysis of repeated measures data**. Singapore: Springer, 2017. 250 p.

JONES, B.; KENWARD, M. G. **Design and analysis of cross-over trials**. 2. ed. EUA: Chapman & Hall/CRC, 2014. 438 p.

KROGH, H. B. *et al.* Methodological advantages and disadvantages of parallel and crossover randomised clinical trials on methylphenidate for attention deficit hyperactivity disorder: a systematic review and meta-analyses. **BMJ open**, GB, v. 9, n. 3, p. e026478, 2019. Disponível em: <<https://doi.org/10.1136/bmjopen-2018-026478>>. Acesso em: 30 jun. 2020.

KUEHL, R. O. O. **Designs of experiments: statistical principles of research design and analysis**. EUA: Duxbury Press, 2000. 666 p.

LARSEN, M. S. *et al.* Impact of trial design on the estimation of drug potency and power in clinical trials of haemophilia with inhibitors. **European Journal of Pharmaceutical Sciences**, Elsevier, Amsterdam, v. 123, p. 531–538, 2018. Disponível em: <<https://doi.org/10.1016/j.ejps.2018.07.056>>. Acesso em: 22 jul. 2021.

LI, B. *et al.* Testing treatment-by-period interaction in four-period crossover trials. **Pharmaceutical statistics**, Wiley Online Library, [S.l.], v. 19, n. 2, p. 145–163, 2020.

Disponível em: <<https://doi.org/10.1002/pst.1975>>. Acesso em: 1 jul. 2021.

LI, D.; SIVAGANESAN, S. An objective bayesian analysis of a crossover design via model selection and model averaging. **Statistics in medicine**, Chichester, v. 35, n. 25, p. 4509–4527, 2016. Disponível em: <<https://doi.org/10.1002/sim.7015>>. Acesso em: 23 mar. 2020.

LIGGES, U. *et al.* Package ‘r2openbugs’. January 2017. Disponível em: <<https://cran.biodisk.org/web/packages/R2OpenBUGS/R2OpenBUGS.pdf>>. Acesso em: 6 mar. 2020.

LITTELL, R. C.; HENRY, P. R.; AMMERMAN, C. B. Statistical analysis of repeated measures data using sas procedures. **Journal of Animal Science**, USA, v. 76, n. 4, p. 1216–1231, 1998. Disponível em: <<https://doi.org/10.2527/1998.7641216x>>. Acesso em: 25 ago. 2019.

LITTELL, R. C. *et al.* **SAS for Mixed Models**. 2. ed. Cary: SAS Institute, 2006. 814 p.

LUI, K. **Crossover designs : testing, estimation, and sample size**. Reino Unido: John Wiley & Sons, 2016. 249 p.

LUNN, D. *et al.* The bugs project: Evolution, critique and future directions. **Statistics in medicine**, v. 28, n. 25, p. 3049–3067, 2009. Disponível em: <<https://doi.org/10.1002/sim.3680>>. Acesso em: 18 dez. 2020.

MADEYSKI, L. Reproducible: Reproduce statistical analyses and meta-analyses. 2017. Disponível em: <<http://CRAN.R-project.org/package=reproducer>>. Acesso em: 14 ago. 2020.

MCCLOUGHLIN, R. *et al.* Soluble fibre supplementation with and without a probiotic in adults with asthma: A 7-day randomised, double blind, three way cross-over trial. **EBioMedicine**, Amsterdam, v. 46, p. 473–485, 2019. Disponível em: <<https://doi.org/10.1016/j.ebiom.2019.07.048>>. Acesso em: 14 jul. 2020.

MORALES, J. C. C.; CAUSIL, C. J. B. **Introducción a la Estadística Bayesiana**. Colombia: Instituto Tecnológico Metropolitano, 2019. 222 p.

NIAZI, M. *et al.* Efficacy of topical rose (rosa damascena mill.) oil for migraine headache: A randomized double-blinded placebo-controlled cross-over trial. **Complementary therapies in medicine**, Edinburgh, v. 34, p. 35–41, 2017. Disponível em: <<https://doi.org/10.1016/j.ctim.2017.07.009>>. Acesso em: 17 abr. 2021.

NOGUEIRA, D.; SÁFADI, T.; FERREIRA, D. Avaliação de critérios de convergência para o método de monte carlo via cadeias de markov. **Revista Brasileira de Estatística**, Rio de Janeiro, v. 65, n. 224, p. 59–88, 2004. Disponível em: <<https://www.researchgate.net/publication/339149423>>. Acesso em: 13 fev. 2020.

NOGUEIRA, D. A. *et al.* Análises clássica e bayesiana de um modelo misto aplicado ao melhoramento animal: uma ilustração. **Ciência e Agrotecnologia**, Lavras, v. 27, p. 1614–1624, 2003. Disponível em: <[http://www2.esalq.usp.br/departamentos/lce/arquivos/aulas/2012/LCE5872/ANALISES\\_CLASSICA\\_E\\_BAYESIANA\\_DE\\_UM\\_MODELO\\_](http://www2.esalq.usp.br/departamentos/lce/arquivos/aulas/2012/LCE5872/ANALISES_CLASSICA_E_BAYESIANA_DE_UM_MODELO_)>

[MISTO\\_melhoramento\\_animal.PDF](#)>. Acesso em: 28 ago. 2019.

NTZOUFRAS, I. **Bayesian modeling using WinBUGS**. New Jersey: John Wiley & Sons, 2011. 520 p.

PAULINO, D. C.; TURKMAN, M. A. A.; MURTEIRA, B. **Estatística Bayesiana**. Lisboa: FCG, 2003. 446 p.

PINHEIRO, J. C.; BATES, D. M. **Mixed-effects models in S and S-PLUS**. New York: Springer-Verlang, 2000. 528 p.

PINTO, P. **Introdução à Análise Estatística**. Faro, Portugal: Sílabas & Desafios, 2015. v. 2. 237 p.

PLUMMER, M. *et al.* "coda: Convergence diagnosis and output analysis for mcmc.". **R News**, v. 6, n. 1, p. 7–11, 2006. Disponível em: <[https://cran.r-project.org/doc/Rnews/Rnews\\_2006-1.pdf#page=7](https://cran.r-project.org/doc/Rnews/Rnews_2006-1.pdf#page=7)>. Acesso em: 1 ago. 2020.

PREDRAZZOLI-JÚNIOR, J. *et al.* Bioequivalence evaluation of two desvenlafaxine succinate monohydrate formulations 50-mg extended-release coated tablet administered in fasted and fed conditions in healthy male and female volunteers. **Jornal brasileiro de economia da saúde**, São Paulo, v. 9, n. 2, p. 198–206, 2017. Disponível em: <<https://doi.org/10.21115/JBES.v9.n2.p198-206>>. Acesso em: 29 jan. 2020.

R Core Team. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2018. Disponível em: <<https://www.R-project.org/>>. Acesso em: 18 maio 2019.

RAFTERY, A. E.; LEWIS, S. Comment: One long run with diagnostics: implementation strategies for markov chain monte carlo. **Statistical Science**, v. 7, n. 4, p. 493–497, 1992. Disponível em: <<https://doi.org/10.1214/ss/1177011143>>. Acesso em: 13 fev. 2021.

RIBOLDI, J. *et al.* Precisão e poder de testes de homocedasticidade paramétricos e não-paramétricos avaliados por simulação. **Revista Brasileira de Biomedicina**, v. 32, n. 3, p. 334–344, 2014. Disponível em: <[http://jaguar.fcav.unesp.br/RME/fasciculos/v32/v32\\_n3/A2\\_JoaoRiboldi.pdf](http://jaguar.fcav.unesp.br/RME/fasciculos/v32/v32_n3/A2_JoaoRiboldi.pdf)>. Acesso em: 13 jun. 2021.

SAMSONSEN, C. *et al.* The seizure precipitating effect of alcohol: A prospective observational cross-over study. **Epilepsy research**, Amsterdam, v. 143, p. 82–89, 2018. Disponível em: <<https://doi.org/10.1016/j.eplepsyres.2018.04.007>>. Acesso em: 25 jul. 2020.

SENN, S. J. The AB/BA crossover: how to perform the two-stage analysis if you can't be persuaded that you shouldn't. **Liber Amicorum Roel van Strik**, Rotterdam, p. 93–100, 1996. Disponível em: <<ftp://biostat.wisc.edu/pub/chappell/641/notes.week5and6/senn1.pdf>>. Acesso em: 12 jun. 2019.

\_\_\_\_\_. **Cross-over Trials in Clinical Research**. 2. ed. Chichester: John Wiley & Sons, Ltd, 2002.

\_\_\_\_\_. Crossover design. In: CHOW, S. C. (Ed.). **Encyclopedia of Biopharmaceutical Statistics**. London: Taylor & Francis, 2003. v. 1, p. 255–262.

SEOANE, J. ¿modelos mixtos (lineales)? una introducción para el usuario temeroso.

**Etologuía: boletín de la Sociedad Española de Etología**, n. 24, p. 15–37, 2014. Disponível em: <[http://ecoevo.uvigo.es/web-see/pdfs/Etologuia\\_vol.24\\_\(2014\).pdf](http://ecoevo.uvigo.es/web-see/pdfs/Etologuia_vol.24_(2014).pdf)>. Acesso em: 24 jan. 2020.

SILVA, R. B. V.; FERREIRA, D. F. Alternativas para o teste t com variâncias heterogêneas avaliadas por meio de simulação. **Ciência e Agrotecnologia**, Lavras, v. 27, n. 1, p. 185–191, 2003. Disponível em: <<https://doi.org/10.1590/S1413-70542003000100023>>. Acesso em: 21 dez. 2020.

SORENSEN, D.; GIANOLA, D. Uncertainty about functions of random variables. In: **Likelihood, Bayesian, and MCMC Methods in Quantitative Genetics**. New York: Springer, 2002. p. 77–116.

TORRES, V.; NAVARRO, J.; PÉREZ, T. Modelos estadísticos para el procesamiento de experimentos con mediciones repetidas en la misma unidad experimental. **Revista Cubana de Ciencia Agrícola**, La Habana, v. 37, n. 3, p. 227–232, 2003. Disponível em: <<https://www.redalyc.org/articulo.oa?id=193018048002>>. Acesso em: 20 dez. 2020.

WEST, B. T.; WELCH, K. B.; GALECKI, A. T. **Linear mixed models: a practical guide using statistical software**. 2. ed. New York: CRC Press, 2014. 414 p.

YIN, J. *et al.* Repeated measures dose-finding design with time-trend detection in the presence of correlated toxicity data. **Clinical Trials**, London, v. 14, n. 6, p. 611–620, 2017. Disponível em: <<https://doi.org/10.1177/1740774517723829>>. Acesso em: 28 jun. 2021.

ZELLNER, A. **Introduction to Bayesian inference in econometrics**. John Wiley: Nova Iorque, 1996. 413 p.

ZHANG, Q.; TOUBOUTI, Y.; CARLIN, B. P. Design and analysis of bayesian adaptive crossover trials for evaluating contact lens safety and efficacy. **Statistical methods in medical research**, London, v. 26, n. 3, p. 1216–1236, 2017. Disponível em: <<https://doi.org/10.1177/0962280215572272>>. Acesso em: 10 fev. 2021.

ZHOU, L. **Application of linear mixed-effects models to crossover designs**. 2012. Tese (Doutorado) — 2012. 36 f. Thesis ( Department of Biostatistics and Bioinformatics) - University of Louisville.

## APÊNDICES

**APÊNDICE A - Link no Github da rotina utilizada para simular os dados *crossover*, com 3 medidas repetidas dentro do período, para o cenário C2; através do *software R* e modelo no *software Openbugs* utilizado para a análises dos dados através da inferência Bayesiana.**

[https://github.com/yacimiranda/Crossover\\_codex.git](https://github.com/yacimiranda/Crossover_codex.git)

**APÊNDICE B - Estimativas das matrizes de variâncias - covariâncias para os tratamentos A e B,  $\hat{\Sigma}_1$  e  $\hat{\Sigma}_2$ , respectivamente. As entradas das matrizes correspondem ao valor médio (em 1000 simulações) das estimativas de média *a posteriori*. Nas diagonais das matrizes temos as estimativas das variâncias para os tempos 1, 2 e 3 e, por fora das diagonais as covariâncias entre os tempos.**

Cenário C1:

$$\hat{\Sigma}_1 = \begin{pmatrix} 47,21 & 26,28 & 26,36 \\ 26,28 & 47,22 & 26,41 \\ 26,36 & 26,41 & 47,30 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 47,49 & 26,95 & 27,22 \\ 26,95 & 48,07 & 27,68 \\ 27,22 & 27,68 & 48,35 \end{pmatrix}.$$

Cenário C2:

$$\hat{\Sigma}_1 = \begin{pmatrix} 49,98 & 28,25 & 27,90 \\ 28,25 & 48,48 & 27,10 \\ 27,90 & 27,10 & 48,27 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 49,01 & 28,07 & 27,72 \\ 28,07 & 48,83 & 27,66 \\ 27,72 & 27,66 & 48,41 \end{pmatrix}.$$

Cenário C3:

$$\hat{\Sigma}_1 = \begin{pmatrix} 49,51 & 28,39 & 28,26 \\ 28,39 & 48,93 & 28,25 \\ 28,26 & 28,25 & 48,82 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 47,55 & 27,39 & 26,98 \\ 27,39 & 48,67 & 27,93 \\ 26,98 & 27,93 & 47,81 \end{pmatrix}.$$



Cenário C4:

$$\hat{\Sigma}_1 = \begin{pmatrix} 48,80 & 27,77 & 27,87 \\ 27,77 & 48,95 & 27,87 \\ 27,87 & 27,87 & 48,72 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 47,13 & 25,98 & 26,71 \\ 25,98 & 46,86 & 26,65 \\ 26,71 & 26,65 & 47,67 \end{pmatrix}.$$

Cenário C5:

$$\hat{\Sigma}_1 = \begin{pmatrix} 49,42 & 27,93 & 27,78 \\ 27,93 & 48,80 & 27,53 \\ 27,78 & 27,53 & 48,18 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 47,02 & 26,60 & 26,28 \\ 26,60 & 48,31 & 27,11 \\ 26,28 & 27,11 & 47,16 \end{pmatrix}.$$

Cenário C6:

$$\hat{\Sigma}_1 = \begin{pmatrix} 49,43 & 27,84 & 27,58 \\ 27,84 & 48,58 & 27,36 \\ 27,58 & 27,36 & 47,99 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 47,43 & 26,24 & 26,83 \\ 26,24 & 47,12 & 26,99 \\ 26,83 & 26,99 & 48,08 \end{pmatrix}.$$

Cenário C7:

$$\hat{\Sigma}_1 = \begin{pmatrix} 548,37 & 522,98 & 521,53 \\ 522,98 & 540,43 & 518,17 \\ 521,53 & 518,17 & 538,17 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 446,28 & 427,09 & 425,53 \\ 427,09 & 450,51 & 428,19 \\ 425,53 & 428,19 & 447,52 \end{pmatrix}.$$

Cenário C8:

$$\hat{\Sigma}_1 = \begin{pmatrix} 542,62 & 521,58 & 518,57 \\ 521,58 & 543,20 & 518,46 \\ 518,57 & 518,46 & 536,43 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 438,21 & 418,29 & 417,34 \\ 418,29 & 440,80 & 418,20 \\ 417,34 & 418,20 & 437,69 \end{pmatrix}.$$

Cenário C9:

$$\hat{\Sigma}_1 = \begin{pmatrix} 554,50 & 531,01 & 530,05 \\ 531,01 & 549,76 & 527,49 \\ 530,05 & 527,49 & 547,96 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 444,60 & 424,83 & 422,99 \\ 424,83 & 447,04 & 423,96 \\ 422,99 & 423,96 & 443,08 \end{pmatrix}.$$

Cenário C10:

$$\hat{\Sigma}_1 = \begin{pmatrix} 538,57 & 516,04 & 513,00 \\ 516,04 & 535,90 & 512,24 \\ 513,00 & 512,24 & 530,60 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 442,83 & 420,14 & 421,58 \\ 420,14 & 439,57 & 420,48 \\ 421,58 & 420,48 & 442,86 \end{pmatrix}.$$

Cenário C11:

$$\hat{\Sigma}_1 = \begin{pmatrix} 537,26 & 514,23 & 513,88 \\ 514,23 & 533,21 & 512,20 \\ 513,88 & 512,20 & 533,12 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 446,35 & 425,64 & 424,83 \\ 425,64 & 446,67 & 425,28 \\ 424,83 & 425,28 & 445,34 \end{pmatrix}.$$

Cenário C12:

$$\hat{\Sigma}_1 = \begin{pmatrix} 536,18 & 512,14 & 509,94 \\ 512,14 & 530,84 & 507,49 \\ 509,94 & 507,49 & 526,69 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 433,07 & 415,03 & 412,86 \\ 415,03 & 439,45 & 416,22 \\ 412,86 & 416,22 & 435,02 \end{pmatrix}.$$

Cenário C13:

$$\hat{\Sigma}_1 = \begin{pmatrix} 45,65 & 25,28 & 25,34 \\ 25,28 & 45,55 & 25,36 \\ 25,34 & 25,36 & 45,47 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 45,74 & 25,44 & 25,39 \\ 25,44 & 45,59 & 25,36 \\ 25,39 & 25,36 & 45,38 \end{pmatrix}.$$

Cenário C14:

$$\hat{\Sigma}_1 = \begin{pmatrix} 45,55 & 25,40 & 25,33 \\ 25,40 & 45,72 & 25,36 \\ 25,33 & 25,36 & 45,48 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 45,49 & 25,48 & 25,50 \\ 25,48 & 45,75 & 25,56 \\ 25,50 & 25,56 & 45,79 \end{pmatrix}.$$

Cenário C15:

$$\hat{\Sigma}_1 = \begin{pmatrix} 45,34 & 25,22 & 25,17 \\ 25,22 & 45,44 & 25,41 \\ 25,17 & 25,41 & 45,44 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 45,43 & 25,32 & 25,17 \\ 25,32 & 45,37 & 25,32 \\ 25,17 & 25,32 & 45,28 \end{pmatrix}.$$

Cenário C16:

$$\hat{\Sigma}_1 = \begin{pmatrix} 45,70 & 25,41 & 25,41 \\ 25,41 & 45,31 & 25,36 \\ 25,41 & 25,36 & 45,50 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 45,41 & 25,54 & 25,34 \\ 25,54 & 45,67 & 25,61 \\ 25,34 & 25,61 & 45,59 \end{pmatrix}.$$

Cenário C17:

$$\hat{\Sigma}_1 = \begin{pmatrix} 45,40 & 25,05 & 25,04 \\ 25,05 & 45,20 & 24,91 \\ 25,04 & 24,91 & 45,14 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 45,44 & 25,21 & 25,12 \\ 25,21 & 45,29 & 25,00 \\ 25,12 & 25,00 & 45,18 \end{pmatrix}.$$

Cenário C18:

$$\hat{\Sigma}_1 = \begin{pmatrix} 45,39 & 25,16 & 25,21 \\ 25,16 & 44,97 & 25,07 \\ 25,21 & 25,07 & 45,22 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 45,40 & 25,57 & 25,39 \\ 25,57 & 45,60 & 25,55 \\ 25,39 & 25,55 & 45,49 \end{pmatrix}.$$

Cenário C19:

$$\hat{\Sigma}_1 = \begin{pmatrix} 425,48 & 405,33 & 404,51 \\ 405,33 & 425,61 & 404,66 \\ 404,51 & 404,66 & 424,13 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 424,50 & 403,86 & 404,22 \\ 403,86 & 423,73 & 403,90 \\ 404,22 & 403,90 & 424,66 \end{pmatrix}.$$

Cenário C20:

$$\hat{\Sigma}_1 = \begin{pmatrix} 430,11 & 409,37 & 408,48 \\ 409,37 & 428,78 & 407,85 \\ 408,48 & 407,85 & 427,19 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 425,49 & 405,17 & 405,14 \\ 405,17 & 425,11 & 404,97 \\ 405,14 & 404,97 & 425,27 \end{pmatrix}.$$

Cenário C21:

$$\hat{\Sigma}_1 = \begin{pmatrix} 427,63 & 406,47 & 407,23 \\ 406,47 & 425,86 & 406,44 \\ 407,23 & 406,44 & 427,29 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 423,59 & 402,95 & 403,34 \\ 402,95 & 422,95 & 403,11 \\ 403,34 & 403,11 & 423,68 \end{pmatrix}.$$

Cenário C22:

$$\hat{\Sigma}_1 = \begin{pmatrix} 427,66 & 407,36 & 407,26 \\ 407,36 & 427,03 & 407,10 \\ 407,26 & 407,10 & 427,18 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 425,94 & 405,38 & 405,91 \\ 405,38 & 424,88 & 405,53 \\ 405,91 & 405,53 & 426,35 \end{pmatrix}.$$

Cenário C23:

$$\hat{\Sigma}_1 = \begin{pmatrix} 427,11 & 406,83 & 406,76 \\ 406,83 & 426,83 & 406,62 \\ 406,76 & 406,62 & 426,56 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 423,02 & 402,92 & 403,12 \\ 402,92 & 423,19 & 403,20 \\ 403,12 & 403,20 & 423,52 \end{pmatrix}.$$

Cenário C24:

$$\hat{\Sigma}_1 = \begin{pmatrix} 427,78 & 407,11 & 406,64 \\ 407,11 & 426,44 & 405,85 \\ 406,64 & 405,85 & 425,58 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 425,40 & 405,42 & 405,30 \\ 405,42 & 425,53 & 405,24 \\ 405,30 & 405,24 & 425,42 \end{pmatrix}.$$

Cenário C25:

$$\hat{\Sigma}_1 = \begin{pmatrix} 49,52 & 28,41 & 28,29 \\ 28,41 & 48,95 & 28,27 \\ 28,29 & 28,27 & 48,85 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 47,55 & 27,40 & 27,00 \\ 27,40 & 48,66 & 27,94 \\ 27,00 & 27,94 & 47,83 \end{pmatrix}.$$

Cenário C26:

$$\hat{\Sigma}_1 = \begin{pmatrix} 540,42 & 515,69 & 514,99 \\ 515,69 & 533,68 & 512,27 \\ 514,99 & 512,27 & 532,83 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 436,64 & 417,89 & 416,02 \\ 417,89 & 441,57 & 419,05 \\ 416,02 & 419,05 & 438,00 \end{pmatrix}.$$

Cenário C27:

$$\hat{\Sigma}_1 = \begin{pmatrix} 46,10 & 25,27 & 25,59 \\ 25,27 & 45,32 & 25,22 \\ 25,59 & 25,22 & 45,88 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 45,61 & 25,09 & 25,21 \\ 25,09 & 45,25 & 25,04 \\ 25,21 & 25,04 & 45,52 \end{pmatrix}.$$

Cenário C28:

$$\hat{\Sigma}_1 = \begin{pmatrix} 425,25 & 405,61 & 405,53 \\ 405,61 & 425,75 & 405,84 \\ 405,53 & 405,84 & 425,99 \end{pmatrix}, \quad \hat{\Sigma}_2 = \begin{pmatrix} 425,23 & 405,00 & 404,95 \\ 405,00 & 424,63 & 404,70 \\ 404,95 & 404,70 & 424,98 \end{pmatrix}.$$