

**PROPOSTA E AVALIAÇÃO DE UMA SOLUÇÃO
BAYESIANA PARA O PROBLEMA DE
BEHRENS-FISHER MULTIVARIADO**

PATRICIA DE SIQUEIRA RAMOS

2009

PATRICIA DE SIQUEIRA RAMOS

**PROPOSTA E AVALIAÇÃO DE UMA SOLUÇÃO
BAYESIANA PARA O PROBLEMA DE
BEHRENS-FISHER MULTIVARIADO**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós- Graduação em Estatística e Experimentação Agropecuária, para a obtenção do título de “Doutor”.

Orientador

Prof. Dr. Daniel Furtado Ferreira

LAVRAS
MINAS GERAIS-BRASIL
2009

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da
Biblioteca Central da UFLA**

Ramos, Patrícia de Siqueira.

Proposta e avaliação de uma solução bayesiana para o problema de Behrens-Fisher multivariado / Patrícia de Siqueira Ramos. - Lavras: UFLA, 2009.

76 p. : il.

Tese (Doutorado) - Universidade Federal de Lavras, 2009.

Orientador: Daniel Furtado Ferreira.

Bibliografia.

1. Comparação de médias. 2. Covariâncias heterogêneas. 3. Simulação Monte Carlo . 4. Erro tipo I. 5. Poder

I. Universidade Federal de Lavras. II.Título.

CDD - 519.542

PATRICIA DE SIQUEIRA RAMOS

**PROPOSTA E AVALIAÇÃO DE UMA SOLUÇÃO
BAYESIANA PARA O PROBLEMA DE
BEHRENS-FISHER MULTIVARIADO**

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós- Graduação em Estatística e Experimentação Agropecuária, para a obtenção do título de “Doutor”.

APROVADA em 18 de dezembro de 2009

Prof. Dr. Denismar Alves Nogueira	UNIFAL-MG
Prof. Dr. Eric Batista Ferreira	UNIFAL-MG
Prof. Dr. João Domingos Scalon	UFLA
Prof. Dr. Renato Ribeiro Lima	UFLA

Prof. Dr. Daniel Furtado Ferreira
UFLA
(Orientador)

LAVRAS
MINAS GERAIS-BRASIL

AGRADECIMENTOS

A minha querida mãe, meu refúgio e apoio incondicional que, com sua delicadeza e sorriso amoroso, me traz paz nas horas em que eu mais preciso.

Ao meu querido pai, sempre seguro e experiente, por procurar me mostrar com muito amor o melhor caminho a seguir.

As minhas irmãs: Paulinha, doce e sensível, por sempre torcer por mim e vibrar com cada vitória e Priscila, exemplo de perseverança e disciplina, por ajudar sempre com suas opiniões. Amo muito cada um de vocês!

Ao meu marido, Lincoln, pela paciência, compreensão, pela forma como busca me tranquilizar e perceber que as coisas são mais simples do que parecem. Com você minha vida se tornou mais leve e os meus sonhos, realidade. Eu te amo!

A minha segunda mãe, Céia, pelo incentivo, carinho e exemplo. Aos meus cunhados, Odilon, Rafael e Laís, pela torcida, amizade e momentos felizes.

Ao meu orientador, Daniel, por me passar tanta segurança, pela amizade e preocupação. Você me mostrou que devo enfrentar a sinuosa estrada do conhecimento com coragem. Eu te admiro muito!

À UFLA e ao DEX, pela oportunidade de realizar o mestrado e o doutorado. À Josi, Edila e demais funcionários, prestativos e simpáticos, por me ajudarem em tudo o que precisei. Aos professores, tão competentes e amigos, pelos ensinamentos que ficarão para sempre comigo!

À Fapemig, pela concessão da bolsa de estudos.

Às amigas e colegas Lourdinha, Nádia, Ademária, Renata, Lu e tantos outros que cruzaram o meu caminho e me ajudaram com seus conselhos, amizade e troca de conhecimento. Jamais os esquecerei!

À Professora Fátima Maria de Souza Moreira, do departamento de Ciência do Solo da UFLA por conceder os dados utilizados no exemplo adotado.

SUMÁRIO

LISTA DE TABELAS	i
LISTA DE FIGURAS	ii
RESUMO	iv
ABSTRACT	v
1 INTRODUÇÃO	1
2 REFERENCIAL TEÓRICO	4
2.1 O problema de Behrens-Fisher	4
2.2 O problema de Behrens-Fisher multivariado	5
2.3 Soluções não-bayesianas para o problema multivariado	6
2.4 Inferência bayesiana	9
2.4.1 Teorema de Bayes	10
2.4.2 Distribuições <i>a priori</i>	11
2.4.3 Distribuições <i>a posteriori</i>	14
2.5 Soluções bayesianas para o problema de Behrens-Fisher	15
2.6 Formas quadráticas	19
2.7 Distribuições multivariadas	20
2.8 Tipos de erro e poder	23
2.9 Simulação Monte Carlo	23
2.9.1 Avaliação de soluções por meio de simulação	24
2.9.2 Medidas frequentistas para avaliação de métodos bayesianos	26
2.9.3 Simulação de dados de distribuições multivariadas	27
3 METODOLOGIA	31
3.1 Solução bayesiana proposta	31
3.2 Simulação Monte Carlo	35

3.3	Teste de Nel & Merwe modificado	39
3.4	Exemplo real	40
4	RESULTADOS E DISCUSSÃO	41
4.1	Taxas de erro tipo I	41
4.2	Poder	49
4.3	Exemplo real	61
5	CONCLUSÕES	64
	REFERÊNCIAS BIBLIOGRÁFICAS	64
	ANEXOS	70

LISTA DE TABELAS

1	Dados para as amostras de solos de capoeira nova (população 1) e de pastagem da Amazônia (população 2), com tamanhos amostrais $n_1 = 30$ e $n_2 = 13$	40
2	Taxas de erro tipo I do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), em função dos tamanhos amostrais n_1 e n_2 e do grau de heterogeneidade (γ) das variâncias, para $p = 2$ variáveis e nível nominal de significância $\alpha = 0,10$, sob H_0	42
3	Taxas de erro tipo I do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), em função dos tamanhos amostrais n_1 e n_2 e do grau de heterogeneidade (γ) das variâncias, para $p = 5$ variáveis e nível nominal de significância $\alpha = 0,10$, sob H_0	46
4	Taxas de erro tipo I do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), em função dos tamanhos amostrais n_1 e n_2 e do grau de heterogeneidade (γ) das variâncias, para $p = 5$ variáveis e nível nominal de significância $\alpha = 0,05$, sob H_0	47
5	Taxas de erro tipo I do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), em função dos tamanhos amostrais n_1 e n_2 e do grau de heterogeneidade (γ) das variâncias, para $p = 5$ variáveis e nível nominal de significância $\alpha = 0,01$, sob H_0	48

LISTA DE FIGURAS

1	Taxas de erro tipo I do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), para $p = 2$, diferentes graus de heterogeneidade γ e $\alpha = 0,05$, sob H_0 . As linhas pontilhadas representam os limites em que o teste de hipótese $H_0 : \alpha = 0,05$ seria rejeitado, considerando um coeficiente de confiança de 0,99.	43
2	Taxas de erro tipo I do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), para $p = 2$, diferentes graus de heterogeneidade γ e $\alpha = 0,01$, sob H_0 . As linhas pontilhadas representam os limites em que o teste de hipótese $H_0 : \alpha = 0,01$ seria rejeitado, considerando um coeficiente de confiança de 0,99.	44
3	Poder do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), para $p = 2$, diferentes graus de heterogeneidade γ , $k = 2$ e $\alpha = 0,10$, sob H_1	50
4	Poder do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), para $p = 2$, diferentes graus de heterogeneidade γ , $k = 4$ e $\alpha = 0,10$, sob H_1	51
5	Poder do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), para $p = 2$, valores diferentes de k , $\gamma = 1$ e $\gamma = 32$ e $\alpha = 0,05$, sob H_1	53
6	Poder do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), para $p = 2$, valores diferentes de γ , $k = 2$ e $k = 16$ e $\alpha = 0,01$, sob H_1	54

7	Poder do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), para $p = 5$, diferentes graus de heterogeneidade γ , $k = 2$ e $\alpha = 0,10$, sob H_1	56
8	Poder do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), para $p = 5$, diferentes graus de heterogeneidade γ , $k = 4$ e $\alpha = 0,10$, sob H_1	57
9	Poder do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), para $p = 5$, valores diferentes de k , $\gamma = 1$ e $\gamma = 32$ e $\alpha = 0,05$, sob H_1	59
10	Poder do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), para $p = 5$, valores diferentes de γ , $k = 2$ e $k = 16$ e $\alpha = 0,01$, sob H_1	60

RESUMO

RAMOS, Patrícia de Siqueira. **Proposta e avaliação de uma solução bayesiana para o problema de Behrens-Fisher multivariado.** 2009. 76p. Tese (Doutorado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras. *

Um dos problemas mais comuns na estatística aplicada é o de comparar as médias de duas populações se a razão entre suas variâncias for desconhecida e diferente de 1 e se as populações forem normais, que é denominado problema de Behrens-Fisher. Aproximações para a estatística t foram utilizadas no caso univariado. No caso multivariado, a maior parte das soluções busca ajustar os graus de liberdade para obter uma melhor aproximação qui-quadrado ou T^2 de Hotelling. Nos dois casos, há soluções bayesianas propostas por alguns autores. Este trabalho foi realizado com os objetivos de propor uma solução bayesiana computacional para o problema de Behrens-Fisher baseada na complexa proposta analítica de Johnson & Weerahandi (1988); avaliar seu desempenho por meio de simulação Monte Carlo, em relação às taxas de erro tipo I e poder e compará-la com a melhor solução frequentista, o teste de Nel & Merwe modificado (Krishnamoorthy & Yu, 2004). As inferências foram realizadas acerca da diferença δ dos vetores de médias populacionais e foi delineado um procedimento para a obtenção da região de credibilidade $100(1 - \alpha)\%$. Utilizou-se uma distribuição *a priori* conjugada para o vetor de médias populacionais (μ_i) e para a matriz de covariâncias (Σ_i), obtendo-se uma distribuição *a posteriori* t multivariada para μ_i , para $i = 1, 2$. Um exemplo real foi utilizado para ilustrar o novo método. Em geral, o teste bayesiano foi conservativo para amostras de tamanhos diferentes e liberal em alguns casos de amostras de tamanhos iguais e pequenos. Seu poder foi igual ou superior ao de seu concorrente em amostras grandes e/ou situações de balanceamento. Como a nova solução possui vantagens, superando seu principal competidor em algumas situações, sua utilização em experimentos reais deve ser recomendada.

Palavras-chave: comparação de médias, heterogeneidade de covariâncias, simulação Monte Carlo, erro tipo I, poder.

* **Orientador:** Daniel Furtado Ferreira - UFLA

ABSTRACT

RAMOS, Patrícia de Siqueira. **Proposal and evaluation of a bayesian solution to the multivariate Behrens-Fisher problem.** 2009. 76p. Thesis (Doctor in Statistics and Agricultural Experimentation) - Federal University of Lavras, Lavras. *

One of the most common problems in applied statistics is to compare two normal population means when the ratio of variances is unknown and not equal to 1. That is the known Behrens-Fisher problem. There are many approaches for the distribution to the t -statistic in the univariate circumstance under the Behrens-Fisher problem. In the multivariate case, most solutions are based on adjusting the degrees of freedom to obtain a better chi-squared or Hotelling's T^2 approximations of the distribution of the test statistic. In both circumstances there are bayesian solutions proposed by some authors. This work aimed to propose a computational based bayesian solution to the multivariate Behrens-Fisher problem based on the complex analytical solution of Johnson & Weerahandi (1988), to evaluate its performance through Monte Carlo simulation computing the type I error rates and power and to compare it with the modified Nel & Merwe test (Krishnamoorthy & Yu, 2004), that is considered the best frequentist solution. The inferences were made to the population difference δ of the mean vectors and a procedure for obtaining a region of $100(1 - \alpha)\%$ of credibility was outlined. It was used a conjugate prior distribution for the population mean vector (μ_i) and covariance matrix (Σ_i) obtaining a posterior multivariate t distribution to μ_i , for $i = 1, 2$. In general, the bayesian test was conservative for samples of different sizes and liberal some circumstances of equal and small sample sizes and its power was equal to or greater than that of its competitor in large samples and/or in balanced circumstances. The new solution has competitive advantages and surpasses its main competitor, therefore its use in real cases should be recommended.

Keywords: means comparison, covariances heteroscedastic, Monte Carlo simulation, type I error rate, power.

* **Guidance Committee:** Daniel Furtado Ferreira - UFLA

1 INTRODUÇÃO

Um dos problemas mais comuns na estatística aplicada é o de comparar as médias de duas populações normais quando a razão entre suas variâncias for desconhecida e diferente de 1 e se as populações forem normais (Scheffé, 1970). No caso univariado, a estatística t de Student é utilizada para testar a igualdade das médias quando as variâncias das duas populações são iguais e os dados são normalmente distribuídos. Porém, a estatística t é sensível à falta de homogeneidade das variâncias e, por isso, várias aproximações foram propostas.

Behrens (1929) foi o primeiro a apresentar uma solução no contexto de teste de hipótese e, depois, Fisher (1935) mostrou que tal solução poderia ser justificada utilizando-se inferência fiducial. Por terem sido esses os primeiros autores a estudarem o problema de comparar as médias de duas populações normalmente distribuídas quando as variâncias das duas populações não são assumidas iguais, este foi batizado de problema de Behrens-Fisher. A primeira abordagem bayesiana para o problema univariado foi apresentada por Jeffreys (1940), cuja solução é equivalente à obtida pela inferência fiducial.

O problema de Behrens-Fisher multivariado ocorre quando há necessidade de comparar dois vetores de médias normais, sendo as covariâncias populacionais heterogêneas e desconhecidas. No caso multivariado, há várias alternativas propostas na literatura para contornar o problema de Behrens-Fisher. A maior parte delas busca ajustar os graus de liberdade para obter uma melhor aproximação qui-quadrado ou T^2 de Hotelling (Ferreira, 2008). Algumas dessas soluções foram propostas por James (1954), Yao (1965), Nel & Merwe (1986), Johnson & Weerahandi (1988), entre outras.

Os desempenhos de várias soluções foram avaliados, por meio de simula-

ção Monte Carlo, por Subrahmaniam & Subrahmaniam (1973), que compararam três métodos; por Christensen & Rencher (1997), que compararam sete dessas soluções aproximadas e por Lix et al. (2005), que avaliaram seis procedimentos frequentistas, sob diferentes situações, além de outros autores. De acordo com os resultados obtidos nesses trabalhos, verifica-se que não há uma solução definitiva, que apresente bom desempenho em qualquer situação.

Há poucas soluções bayesianas para o problema multivariado. Johnson & Weerahandi (1988) apresentaram uma delas, não tendo sido, porém, avaliada por simulação Monte Carlo. Alguns autores justificam o uso de métodos computacionais frequentistas para avaliar o desempenho de testes bayesianos e os utilizam em diferentes contextos. Dentre eles, destacam-se Gopalan & Berry (1998), em comparações múltiplas; Ghosh & Kim (2001), no problema de Behrens-Fisher univariado; Agresti & Min (2005), em tabelas de contingência e Biase (2009), em comparações múltiplas de proporções binomiais.

Além de o desempenho do teste bayesiano proposto por Johnson & Weerahandi (1988) não ter sido avaliado, a solução obtida é complexa, requerendo a obtenção numérica da inversa da função de distribuição *a posteriori* de uma forma quadrática que envolve as diferenças de médias paramétricas $\mu_1 - \mu_2$ e as diferenças observadas experimentalmente $\bar{Y}_1 - \bar{Y}_2$. Essa forma quadrática tem função de distribuição expressa como uma série infinita que combina uma função peso aplicada ao valor esperado de uma variável beta com a função de distribuição F . Acredita-se que o grau de dificuldade envolvido nessa solução seja muito grande, e, conseqüentemente, nenhuma referência sobre tal artigo tenha sido encontrada em trabalhos aplicados na literatura científica.

Assim, este trabalho foi realizado com o objetivo de propor uma solução bayesiana computacional para o problema de Behrens-Fisher multivariado base-

ada na complexa proposta analítica apresentada por Johnson & Weerahandi (1988). Além disso, pretende-se avaliar o desempenho do procedimento proposto por meio de simulação Monte Carlo, avaliando-se as taxas de erro tipo I e poder, sob diferentes graus de heterogeneidade das covariâncias das duas populações normais multivariadas, considerando várias combinações de tamanhos amostrais e diferenças em erros padrões entre os vetores de médias para distintos valores nominais de significância. Objetivou-se também comparar o desempenho da nova proposta com o teste de Nel & Merwe modificado, considerada a melhor solução frequentista, de acordo com Krishnamoorthy & Yu (2004) que o propuseram e o avaliaram. Um exemplo real foi utilizado para ilustrar a aplicação do novo método.

2 REFERENCIAL TEÓRICO

2.1 O problema de Behrens-Fisher

Um problema muito comum na estatística experimental é o de comparar as médias de duas populações normais quando a razão entre suas variâncias for desconhecida e diferente de 1 e se as populações forem normais (Scheffé, 1970). Behrens (1929) foi o primeiro a apresentar uma solução no contexto de teste de hipótese, $H_0 : \mu_1 = \mu_2 \times H_1 : \mu_1 \neq \mu_2$ e, depois, Fisher (1935) mostrou que tal solução poderia ser justificada utilizando inferência fiducial.

Scheffé (1943) apresentou a primeira solução frequentista para o problema univariado. Um procedimento comum na prática é testar a igualdade de variâncias e aplicar o teste t de Student usual se as variâncias forem consideradas iguais ou, caso contrário, aplicar um teste aproximado da estatística t , tais como a estatística t de Welch (1947) ou o teste t com graus de liberdade aproximados de Satterthwaite (1946). Entretanto, o procedimento de Welch não é robusto à maioria das distribuições não-normais (Reed, 2003).

Kim & Cohen (1998) apresentam algumas soluções frequentistas para o problema, tendo algumas delas sido avaliadas em relação às taxas de erro tipo I e poder. De acordo com o autor, as soluções mais importantes, do ponto de vista frequentista, são as de Welch (1947), Aspin (1948, 1949) e Tsui & Weerahandi (1989), das quais a última leva à mesma conclusão obtida pelas teorias fiducial e bayesiana. Reed (2003) desenvolveu um programa para as estatísticas sugeridas por três outros trabalhos que lida com os casos de violações às suposições.

Há também soluções não-paramétricas propostas. Geralmente, tais métodos buscam avaliar a informação nos dados, substituindo valores esperados com postos ou usando permutações dos dados, sendo alguns trabalhos citados por Leh-

mann (1975) e Munzel (1999).

Mais referências sobre soluções para o problema de Behrens-Fisher univariado podem ser encontradas em Duong & Shorrocks (1996), Ghosh & Kim (2001) e Reed (2003).

2.2 O problema de Behrens-Fisher multivariado

No caso multivariado, para se testar a hipótese nula de igualdade dos vetores de médias de duas populações normais dada por

$$H_0 : \mu_1 = \mu_2 = \mu,$$

não se pode construir um teste exato baseado na distribuição T^2 de Hotelling, extensão multivariada da distribuição t univariada. O teste T^2 baseia-se nas suposições de igualdade das matrizes de covariâncias e de normalidade das observações. É uma distribuição útil para a realização de testes de hipóteses envolvendo médias e diferenças de médias de populações normais multivariadas (Ferreira, 2008). No problema de Behrens-Fisher multivariado, como há heterogeneidade das matrizes de covariâncias, ou seja, $\Sigma_1 \neq \Sigma_2$, o teste T^2 se torna viesado, especialmente se os tamanhos das amostras forem diferentes (Lix et al., 2005). Por essa razão, várias soluções aproximadas foram propostas, como as de Bennett (1951), James (1954), Yao (1965) e Kim (1992), entre outras, em que há um ajuste que permite a aproximação da estatística T^2 de Hotelling ou qui-quadrado e seus graus de liberdade. Porém, as desvantagens desses métodos são sua complexidade e dificuldade de uso prático.

Há poucos trabalhos que propõem uma abordagem bayesiana para o problema multivariado, sendo um deles o de Johnson & Weerahandi (1988). Entre-

tanto, este envolve aproximação numérica e cálculos complexos. Outra solução bayesiana foi proposta por Buckley (2004), baseada em simulação Monte Carlo via cadeias de Markov. Nas seções seguintes, algumas dessas soluções serão explicitadas.

2.3 Soluções não-bayesianas para o problema multivariado

Um dos métodos não-paramétricos para o caso multivariado foi proposto por Brunner et al. (2002). Foram consideradas duas amostras independentes e com distribuições marginais assumidas não iguais. O modelo adotado abrangia o caso de observações empatadas e as matrizes de covariâncias foram consideradas desconhecidas. Para se testar a hipótese de efeitos iguais, os autores consideraram versões da estatística de Wald e do tipo da análise da variância, sugerida por Brunner et al. (1997) para modelos não-paramétricos univariados. Pelos resultados das simulações, os autores concluíram que a estatística do tipo da análise da variância controlou as taxas de erro tipo I, mesmo para amostras pequenas, enquanto a outra estatística tendeu a ser liberal. As duas versões foram praticamente idênticas em relação ao poder.

No caso paramétrico, geralmente, têm-se extensões do caso univariado. Bennett (1951) generalizou o procedimento univariado de Scheffé (1943), definindo um vetor envolvendo as observações multivariadas das duas populações e os tamanhos amostrais n_1 e n_2 . Para se testar a hipótese $H_0 : \mu_1 - \mu_2 = \delta_0$, uma estatística T^2 foi proposta. Sob H_0 , a estatística segue a distribuição T^2 de Hotelling, com dimensão p e $\nu = n_1 - 1$ graus de liberdade.

A solução proposta por James (1954) usa uma expansão assintótica de primeira ordem de formas quadráticas, permitindo utilizar uma aproximação para a distribuição da estatística

$$T^{*2} = (\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2)^\top \mathbf{S}_e^{-1} (\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2), \quad (2.1)$$

como uma soma de qui-quadrados, em que $\mathbf{S}_e = \mathbf{S}_1/n_1 + \mathbf{S}_2/n_2$. A estatística T^{*2} apresenta valor crítico $\chi_{\alpha,p}^2(A + B\chi_{\alpha,p}^2)$, em que

$$A = 1 + \frac{1}{2p} \sum_{i=1}^2 \frac{1}{n_i - 1} \left[\text{tr} \left(\frac{\mathbf{S}_e^{-1} \mathbf{S}_i}{n_i} \right) \right]^2,$$

$$B = \frac{1}{2p(p+2)} \left\{ \sum_{i=1}^2 \frac{1}{n_i - 1} 2\text{tr} \left[\left(\frac{\mathbf{S}_e^{-1} \mathbf{S}_i}{n_i} \right)^2 \right] + \left[\text{tr} \left(\frac{\mathbf{S}_e^{-1} \mathbf{S}_i}{n_i} \right) \right]^2 \right\},$$

e $\chi_{\alpha,p}^2$ é o quantil superior 100 α % da distribuição qui-quadrado com p graus de liberdade. James (1954) ainda propôs um teste de segunda ordem, porém, de maior complexidade computacional.

Yao (1965) propôs uma extensão multivariada da distribuição t de Student. Os graus de liberdade (ν) do teste são estimados por uma expressão que envolve os tamanhos amostrais, as matrizes de covariâncias e a estatística (2.1). Sob a hipótese nula $H_0 : \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 = \boldsymbol{\delta}_0$, a estatística T^{*2} possui distribuição T^2 de Hotelling com ν graus de liberdade e dimensão p . Rejeita-se H_0 se $T^{*2} > T_{\alpha,p,\nu}^2$, sendo $T_{\alpha,p,\nu}^2$ o quantil superior 100 α % da distribuição T^2 de Hotelling.

O teste de Johansen (1980) é outra solução que utiliza ajustes nos graus de liberdade. O valor crítico aproximado para o teste de $H_0 : \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 = \boldsymbol{\delta}_0$ é $F_{\alpha,p,\nu}$, quantil superior 100 α % da distribuição F com p e ν graus de liberdade.

A sugestão de Nel & Merwe (1986) também utiliza a mesma estatística utilizada por James (1954), aproximadamente distribuída como uma T^2 de Hotelling com dimensão p e ν graus de liberdade, em que

$$\nu = \frac{tr[(\mathbf{S}_e)^2] + [tr(\mathbf{S}_e)]^2}{\sum_{i=1}^2 \frac{1}{n_i-1} \left\{ tr \left[\left(\frac{\mathbf{S}_i}{n_i} \right)^2 \right] + \left[tr \left(\frac{\mathbf{S}_i}{n_i} \right) \right]^2 \right\}}. \quad (2.2)$$

Rejeita-se a hipótese nula quando $T^{*2} > \nu p F_{\alpha, p, \nu+1-p} / (\nu - 1 + p)$.

A estatística de Hwang & Paulson (1986) é parecida com a razão F e seus autores fornecem tabelas com os valores críticos de sua estatística. Esta solução e a de Nel & Merwe (1986) não são invariantes a transformações não-singulares dos dados.

Kim (1992) apresentou uma extensão dos graus de liberdade univariado de Welch (1947) em que, primeiramente, deve-se resolver um sistema de equações homogêneas envolvendo autovalores e autovetores, o que equivale à maximização de formas quadráticas. A estatística do teste é calculada por uma expressão que é função de uma transformação linear e de outras quantidades calculadas e que, sob a hipótese nula, possui distribuição F aproximada.

Krishnamoorthy & Yu (2004) propuseram uma modificação do teste de Nel & Merwe (1986) para obter um teste invariante a transformações não-singulares. O ajuste nos graus de liberdade é dado por

$$\nu = \frac{p + p^2}{\sum_{i=1}^2 \frac{1}{n_i-1} \left\{ tr \left[\left(\frac{\mathbf{S}_i \mathbf{S}_e^{-1}}{n_i} \right)^2 \right] + \left[tr \left(\frac{\mathbf{S}_i \mathbf{S}_e^{-1}}{n_i} \right) \right]^2 \right\}}. \quad (2.3)$$

A rejeição da hipótese nula se dá como no teste de Nel & Merwe (1986). Em ambos os casos, os graus de liberdade são os mesmos para $p = 1$. Assim, no caso univariado, os testes são equivalentes e reduzem-se à solução da aproximação dos graus de liberdade de Welch (1947), o que também ocorre com a aproximação

de Yao (1965).

2.4 Inferência bayesiana

Metodologias utilizando a teoria bayesiana têm avançado muito nas últimas décadas. Um dos primeiros livros a tratar de métodos bayesianos aplicados foi o de Box & Tiao (1992), cuja primeira edição é de 1973, que aborda a inferência baseando-se na distribuição normal.

Segundo Gelman et al. (2003), a inferência bayesiana busca tirar conclusões sobre quantidades não observáveis a partir de dados numéricos. Conclusões estatísticas bayesianas sobre um parâmetro θ , ou dados não observáveis \tilde{y} , são obtidas em termos de afirmações probabilísticas. Essas afirmações são condicionadas ao valor observado de y e sua notação é $p(\theta|y)$ ou $p(\tilde{y}|y)$. Esse condicionamento em dados observados diferencia a inferência bayesiana da frequentista que, por sua vez é baseada em uma avaliação retrospectiva do procedimento para estimar θ , ou \tilde{y} , dentro da distribuição dos possíveis valores de y condicionados ao valor real desconhecido de θ .

Métodos frequentistas e bayesianos diferem quanto à natureza da probabilidade, parâmetros e inferências. Segundo a teoria frequentista, apenas eventos repetíveis têm probabilidades. Sob a abordagem bayesiana, a probabilidade descreve a incerteza devido à variabilidade aleatória ou ao pouco conhecimento sobre a situação. Essa diferença na noção de probabilidade implica na noção de parâmetros. Frequentistas não fazem afirmações probabilísticas sobre parâmetros, por não os considerarem aleatórios, mas fixos. Já a perspectiva bayesiana permite formulá-las simplesmente porque parâmetros são desconhecidos. Outra diferença se encontra nas inferências. A interpretação frequentista é realizada em termos de amostragem repetida e a bayesiana é feita em termos de evidência provida pelos

dados observados (O'Hagan & Luce, 2003).

2.4.1 Teorema de Bayes

Os dados têm a função de aumentar o conhecimento e, por isso, atualizar o que pode ser dito sobre os parâmetros e hipóteses de interesse. Informação *a priori*, ou seja, o que se tem conhecimento antes da coleta dos dados, pode ser expressa na forma de uma distribuição *a priori* dos parâmetros desconhecidos do modelo estatístico. Essa informação, unida à informação proveniente dos dados, produz a distribuição *a posteriori* que expressa o que se sabe sobre os parâmetros após a observação dos dados. O teorema de Bayes é o mecanismo responsável por essa junção. Sob a abordagem bayesiana, qualquer inferência é derivada da distribuição *a posteriori* (O'Hagan & Luce, 2003).

Para produzir afirmações probabilísticas sobre θ dado y (os dados), assume-se um modelo da distribuição de probabilidade conjunta de θ e y . Essa densidade pode ser escrita como o produto de duas densidades, referidas como distribuição *a priori* $p(\theta)$ e a distribuição amostral $p(y|\theta)$ (Gelman et al., 2003):

$$p(\theta, y) = p(\theta)p(y|\theta).$$

Condicionar o valor conhecido y , usando a propriedade básica da probabilidade condicional, conhecida como regra de Bayes, leva à densidade *a posteriori* (Gelman et al., 2003):

$$p(\theta|y) = \frac{p(\theta, y)}{p(y)} = \frac{p(\theta)p(y|\theta)}{p(y)}, \quad (2.4)$$

em que $p(y) = \int p(\theta)p(y|\theta)d\theta$, se θ for contínuo. Uma forma equivalente a (2.4) omite o denominador, que não depende de θ e pode ser considerado como constante, para y fixo. Assim,

$$p(\theta|y) \propto p(\theta)p(y|\theta). \quad (2.5)$$

Utilizar a regra de Bayes com um modelo de probabilidade escolhido significa que os dados y afetam a inferência *a posteriori* (2.5) apenas pela função $p(y|\theta)$, a função de verossimilhança (Gelman et al., 2003). A função de verossimilhança $p(y|\theta)$ é aquela por meio da qual os dados modificam o conhecimento *a priori* sobre θ , ou seja, representa a informação sobre θ advinda dos dados (Box & Tiao, 1992).

2.4.2 Distribuições *a priori*

De acordo com O'Hagan & Luce (2003), a informação *a priori* representa, ao mesmo tempo, os pontos forte e fraco da abordagem bayesiana. A maior parte das críticas à análise bayesiana vem da subjetividade da informação *a priori*, o que torna a distribuição *a posteriori* e todas as inferências baseadas nela subjetivas. O ponto forte é permitir que mais informação seja incluída na análise para produzir inferências mais poderosas. Porém, a escolha de uma distribuição *a priori* inadequada pode levar a resultados ruins.

A utilização da informação *a priori* na inferência bayesiana requer a especificação de uma distribuição para a quantidade de interesse θ . Essa distribuição deve representar probabilisticamente o conhecimento que se tem sobre θ , antes da obtenção dos dados (Ehlers, 2007).

O teorema de Bayes atribui pesos à informação *a priori* e aos dados para

obter a distribuição *a posteriori*. Se a informação *a priori* é vaga, ela terá pouco peso ao se unir aos dados e a distribuição *a posteriori* será baseada inteiramente nos dados por meio de sua função de verossimilhança. À medida que mais dados são adquiridos, o teorema de Bayes atribui a eles mais peso. Essa característica reflete o processo da ciência, em que o acúmulo de evidência objetiva é o processo segundo o qual diferenças de opinião são resolvidas (O’Hagan & Luce, 2003).

Os parâmetros da distribuição *a priori* são denominados hiperparâmetros. Quando a distribuição *a posteriori* segue a mesma forma paramétrica da distribuição *a priori*, esta é chamada de “conjugada”. A família conjugada é matematicamente conveniente porque, dessa forma, a distribuição *a posteriori* segue uma forma paramétrica conhecida (Gelman et al., 2003). Se F é uma classe de distribuições amostrais $p(y|\theta)$ e P é uma classe de distribuições *a posteriori* para θ , então, a classe P é conjugada de F se

$$p(\theta|y) \in P \text{ para todo } p(\cdot|\theta) \in F \text{ e } p(\cdot) \in P.$$

O interesse está em tomar P como o conjunto de todas as densidades que apresentam a mesma forma funcional da verossimilhança. Porém, para modelos mais complexos, em que não é possível utilizar distribuições *a priori* conjugadas, devem ser utilizadas distribuições *a priori* não-conjugadas.

Quando há completa “ignorância” em relação a θ , a proposta de Laplace, princípio da razão insuficiente ou critério de Bayes-Laplace, consiste em atribuir igual probabilidade aos valores de θ , com uma distribuição *a priori* uniforme, porém, sua utilização está sujeita a críticas, por não satisfazer ao axioma de probabilidade total unitária (Paulino et al., 2003). Uma propriedade básica de uma função densidade de probabilidade $f(y)$ é que ela integre ou que a soma sobre seus valores possíveis seja 1. Porém, se $f(y) = k$ é uniforme no intervalo $(-\infty; \infty)$,

$k > 0$, $f(y)$ não será uma densidade própria, já que a integral

$$\int_{-\infty}^{\infty} f(y)dy = k \int_{-\infty}^{\infty} dy$$

é imprópria, não importando quanto k seja pequeno. Distribuições desse tipo são denominadas impróprias (Box & Tiao, 1992).

Quando não há informação *a priori* disponível ou quando há pouco conhecimento em relação à informação amostral, utilizam-se distribuições *a priori* não-informativas (vagas, difusas), no caso de informação *a priori* insuficiente que torne difícil a obtenção de uma distribuição considerada adequada. Essas distribuições também são tidas como uma forma de deduzir as crenças *a posteriori* para quem parte de um grau de conhecimento escasso e permitir a comparação com resultados da inferência frequentista que apenas utiliza a informação amostral (Paulino et al., 2003).

Jeffreys (1961) propôs uma classe de distribuições *a priori* não-informativas que é invariante a transformações um a um. O método se baseia no uso da medida de informação de Fisher sobre $\theta \in \mathbb{R}$ (Paulino et al., 2003).

Há críticas quanto ao uso de distribuições *a priori* não-informativas, mas não se pode negar seu apelo prático. Ghosh & Kim (2001) ressaltam que a crescente aceitação de técnicas bayesianas, tanto em áreas teóricas quanto práticas, pode ser atribuída ao fato de que, até mesmo com pouca ou vaga informação *a priori*, essas técnicas podem ser utilizadas com sucesso, por meio da adoção de distribuições *a priori* não-informativas.

2.4.3 Distribuições *a posteriori*

A distribuição *a posteriori* $p(\theta|y)$ é o elemento fundamental que serve de base ao desenvolvimento de toda a inferência bayesiana (Paulino et al., 2003). A distribuição $p(\theta|y)$ contém toda a informação atual sobre o parâmetro θ . Para se resumir essa informação, podem ser utilizados gráficos, medidas de posição, como média, mediana e moda da distribuição, e medidas de variabilidade, como desvio padrão e quantis (Gelman et al., 2003).

Em contraste com a inferência frequentista, cujas questões são formuladas em termos de testes de significância, intervalos de confiança e estimadores não-viesados, a inferência bayesiana pode usar a distribuição *a posteriori* de forma flexível para fornecer respostas relevantes e diretas aos mais variados tipos de questões (O'Hagan & Luce, 2003).

Métodos de Monte Carlo simples e via função de importância (Paulino et al., 2003) e métodos de reamostragem (Smith & Gelfand, 1992), como o da rejeição, são métodos de simulação em que se obtém uma amostra da distribuição *a posteriori* em um único passo. Os valores são gerados de forma independente e não há preocupação com a convergência do algoritmo, bastando que o tamanho da amostra seja suficientemente grande. Dessa forma, esses métodos são chamados não-iterativos. Porém, em muitas situações, pode ser difícil encontrar uma densidade de importância que seja uma boa aproximação da distribuição *a posteriori* e que seja fácil de ser amostrada. Os métodos de Monte Carlo via cadeias de Markov (MCMC) são alternativas em problemas complexos (Ehlers, 2007).

Simulação por cadeias de Markov é utilizada quando não é possível (ou ineficiente) amostrar θ diretamente de $p(\theta|y)$. Assim, amostra-se iterativamente de forma que, a cada passo do processo, espera-se amostrar de uma distribuição mais próxima de $p(\theta|y)$ (Gelman et al., 2003). Os métodos MCMC correspondem à

associação de algoritmos para simulação de distribuições com o método de Monte Carlo para a aproximação de integrais. O objetivo dos métodos iterativos e não-iterativos é obter uma amostra da distribuição *a posteriori* e calcular estimativas amostrais de características dessa distribuição (Ehlers, 2007).

O amostrador de Gibbs (Gelfand & Smith, 1990), caso particular dos métodos MCMC, é definido em termos de subvetores de θ . Sejam k subvetores $(\theta_1, \theta_2, \dots, \theta_k)'$ e que as distribuições condicionais de cada parâmetro θ_i , dados os outros, sejam conhecidas. Essas distribuições são denominadas condicionais completas e denotadas por $f_1(\theta_1|\theta_2, \dots, \theta_k, y), \dots, f_k(\theta_k|\theta_1, \dots, \theta_{k-1}, y)$. Após a conclusão das iterações, a amostra de valores obtidos forma uma amostra aleatória de valores de θ , utilizada para se aproximar as distribuições marginais *a posteriori* e, assim, realizar inferências (Gelman et al., 2003). Se as condicionais completas não forem conhecidas, deve-se utilizar o algoritmo de Metropolis-Hastings (Metropolis et al., 1953; Hastings, 1970).

2.5 Soluções bayesianas para o problema de Behrens-Fisher

A primeira solução bayesiana para o problema de Behrens-Fisher univariado, dada por Jeffreys (1940), mostra que um cálculo bayesiano baseado na distribuição *a priori* $\pi(\mu_1, \mu_2, \sigma_1, \sigma_2) \propto (\sigma_1\sigma_2)^{-1}$ leva a um intervalo de credibilidade para $\mu_1 - \mu_2$, que é algebricamente equivalente ao intervalo fiducial de Fisher, embora com outra interpretação. Por essa razão, Ghosh & Kim (2001) a denominam distribuição *a priori* independente de Jeffreys, diferente da regra geral de Jeffreys, que é a raiz quadrada positiva do determinante da matriz de informação de Fisher que, no exemplo, seria proporcional a $(\sigma_1\sigma_2)^{-2}$. Esses autores estudaram o problema de Behrens-Fisher do ponto de vista da teoria bayesiana, propondo uma distribuição *a priori* alternativa à de Jeffreys (1940). Pelos resultados das simu-

lações, a utilização dessa distribuição alternativa apresentou desempenho superior por levar a valores mais precisos da probabilidade de cobertura nominal, que é um critério frequentista.

Duong & Shorrok (1996) apresentaram um procedimento bayesiano para o problema de comparar duas médias normais por meio de uma generalização de uma distribuição beta de dois para três parâmetros. Uma abordagem bayesiana empírica foi utilizada para estimar o novo parâmetro da distribuição *a priori*, relacionado à razão das variâncias. O método provê uma maneira operacional de avaliar a plausibilidade de vários valores do novo parâmetro, utilizando os dados. Além de ser um método simples de se implementar, pode ser estendido para a comparação de muitas médias.

Moreno et al. (1999) propõem uma análise padrão completa bayesiana para o problema de Behrens-Fisher, convertendo distribuições *a priori* impróprias convencionais, uniformes para os parâmetros de posição e o log dos parâmetros de escala, em distribuições *a priori* para as quais o fator de Bayes seja bem definido.

Uma solução bayesiana para o problema de Behrens-Fisher multivariado foi apresentada por Johnson & Weerahandi (1988). O objetivo do procedimento foi obter regiões de credibilidade elipsoidais *a posteriori* para a diferença entre duas médias normais multivariadas. As distribuições *a priori* propostas eram não-informativas ou conjugadas. Para o primeiro caso, a distribuição foi

$$\pi(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2) = (|\boldsymbol{\Sigma}_1||\boldsymbol{\Sigma}_2|)^{-1/2}, \quad i = 1, 2,$$

em que $\boldsymbol{\mu}_i \in \mathbf{R}^p$ e $\boldsymbol{\Sigma}_i > 0$, e $\boldsymbol{\mu}_1$ e $\boldsymbol{\mu}_2$ eram independentes. A diferença $\boldsymbol{\mu}_i - \bar{\mathbf{y}}_i$ tem distribuição $T_p(\mathbf{V}_i/(n_i - p), n_i - p)$, $i = 1, 2$, em que $T_p(\mathbf{A}, \nu)$ é a distribuição t multivariada p -dimensional com parâmetros \mathbf{A} e ν .

No segundo caso, utilizaram-se distribuições *a priori* conjugadas. Considerou-se que $\boldsymbol{\mu}_i|\boldsymbol{\Sigma}_i$ é distribuída como $N_p(\mathbf{a}_i, \boldsymbol{\Sigma}_i/q_i)$ e $\boldsymbol{\Sigma}_i$ é distribuída como $W_p^{-1}(\mathbf{R}_i, r_i)$, $i = 1, 2$, em que $W_p^{-1}(\mathbf{A}, \nu)$ é a distribuição Wishart invertida com parâmetros \mathbf{A} e ν . Assim, a distribuição conjunta de $\boldsymbol{\mu}_1$ e $\boldsymbol{\mu}_2$ é definida da seguinte forma: $\boldsymbol{\mu}_1$ e $\boldsymbol{\mu}_2$ são independentes e $\boldsymbol{\mu}_i - \hat{\boldsymbol{y}}_i$ é distribuída como $T_p(\mathbf{U}_i/(n_i + r_i - 2p), n_i + r_i - 2p)$, $i = 1, 2$, em que $\hat{\boldsymbol{y}}_i = (n_i \bar{\boldsymbol{y}}_i + q_i \mathbf{a}_i)/(n_i + q_i)$ e

$$\mathbf{U}_i = \left[\frac{1}{(n_i + q_i)} \right] \left[\mathbf{W}_i + \mathbf{R}_i + \frac{n_i q_i}{(n_i + q_i)} (\mathbf{a}_i - \bar{\boldsymbol{y}}_i)(\mathbf{a}_i - \bar{\boldsymbol{y}}_i)^\top \right], i = 1, 2.$$

Para ambas as distribuições *a priori*, a inferência é realizada sobre $\boldsymbol{\delta} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$ e, para isso, é preciso calcular as probabilidades baseadas na convolução de duas distribuições *t* multivariadas. Para a inferência, os autores consideraram a forma quadrática dada por

$$Q = (\boldsymbol{\delta} - \mathbf{d})^\top \mathbf{V}^{-1} (\boldsymbol{\delta} - \mathbf{d})$$

centrada na média *a posteriori* $\mathbf{d} = \bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2$ de $\boldsymbol{\delta}$. A matriz positiva definida conhecida \mathbf{V} pode ser escolhida para facilitar os cálculos das probabilidades de Q . A função de distribuição acumulada de Q é uma combinação de distribuições *F* centrais, $F_{r,s}$, com r e s graus de liberdade, dada por

$$F_Q(q) = \sum_{j=0}^{\infty} E(\omega_j) F_{p+2j,n} \left[\frac{nq}{\theta(p+2j)} \right],$$

em que θ é uma constante arbitrária a ser escolhida de forma a obter convergência rápida do somatório, a esperança é tomada em relação à variável aleatória beta B ,

distribuída como $Be[(n_1 - p)/2, (n_2 - p)/2]$, e $n = n_1 + n_2 - 2p$. A quantidade $\omega_j (j = 1, 2, \dots)$ é definida, em termos de θ e $\lambda_j(B)$, pela relação de recursão

$$\omega_r = \frac{1}{2r} \sum_{j=0}^{r-1} H_{r-j} \omega_j,$$

em que $w_0 = \prod_{j=1}^p (\theta/\lambda_j)^{1/2}$, $H_r = \sum_{j=1}^p (1 - \theta/\lambda_j)^r (r = 1, 2, \dots)$ e $\lambda_j = \lambda_j(B)$, $j = 1, \dots, p$ são os autovalores ordenados $\lambda_1(B) \leq \dots \leq \lambda_p(B)$ da seguinte matriz:

$$\frac{1}{n_1 B} \mathbf{V}^{-1/2} \mathbf{V}_1 \mathbf{V}^{-1/2} + \frac{1}{n_2 (1 - B)} \mathbf{V}^{-1/2} \mathbf{V}_2 \mathbf{V}^{-1/2}.$$

Após obtida a distribuição de $Q (F_Q(q))$ e seus quantis, a hipótese $H_0 : \boldsymbol{\delta} = \boldsymbol{\delta}_0 = [\delta_{01}, \delta_{02}]^\top$ deve ser rejeitada se a elipsoide obtida não contiver o valor hipotético $\boldsymbol{\delta}_0$. O desempenho dessa solução não foi avaliado por simulação.

Buckley (2004) propõe outra solução bayesiana para o problema de Behrens-Fisher univariado aplicada na área de ciência política, em que o autor apresenta um modelo que permite a estimação bayesiana da diferença $\delta = \mu_1 - \mu_2$ entre médias populacionais. A distribuição *a posteriori* dessa diferença, chamada de distribuição Behrens-Fisher, é obtida utilizando-se distribuições *a priori* não-informativas para as médias e para o log das variâncias das duas populações. Interpretações podem ser feitas plotando-se a distribuição *a posteriori* ou obtendo-se vários *highest posterior density* (HPDs) ou regiões de valores que contém $100(1 - \alpha)\%$ da probabilidade *a posteriori*, para diferentes porcentagens ou proporções da densidade *a posteriori*. O modelo é estendido para considerar va-

riáveis não-normais, distribuições *a priori* informativas e um teste multivariado baseado em simulação Monte Carlo via cadeias de Markov (MCMC). Para implementar o teste multivariado, o autor considerou distribuições *a priori* conjugadas para os vetores de médias e para as matrizes de covariâncias das duas populações. As médias tinham distribuições *a priori* normais independentes com média $\mathbf{0}$ e uma precisão baixa (inverso da variância) e as matrizes de precisão eram independentemente distribuídas como Wishart. A distribuição *a posteriori* da diferença entre os vetores de médias foi, então, obtida. No lugar de obter probabilidades *a posteriori* exatas usando seu modelo ou obter aproximações, o autor utilizou simulação MCMC para estimar as densidades marginais *a posteriori*. Infelizmente, as deficiências do trabalho de Buckley (2004) são que detalhes das distribuições e da implementação não são apresentados no artigo original e não foi realizada simulação Monte Carlo para avaliar sua solução, apenas exemplos de aplicação foram utilizados para ilustrar o método.

A solução proposta por Buckley (2004) foi uma tentativa de contornar a complexidade da solução de Johnson & Weerahandi (1988) e fornecer uma alternativa mais simples, embora detalhes de sua solução não sejam apresentados pelo autor. Assim, observa-se uma carência na literatura científica por uma solução bayesiana simples para o problema multivariado.

2.6 Formas quadráticas

Pode-se definir uma forma quadrática a partir de uma matriz simétrica $\mathbf{A}(n \times n)$ por (Ferreira, 2008):

$$Q(\mathbf{y}) = \mathbf{y}^\top \mathbf{A} \mathbf{y} = \sum_{i=1}^n a_{ii} y_i^2 + 2 \sum_{i=1}^{n-1} \sum_{k=i+1}^n a_{ik} y_i y_k \quad (2.6)$$

em que $\mathbf{y} \neq \mathbf{0}$ é um vetor definido em \mathbb{R}^n .

A expressão (2.6) é denominada forma quadrática por apresentar apenas termos quadráticos ou duplos produtos dos elementos do vetor \mathbf{y} . Como os elementos de \mathbf{A} são conhecidos, $Q(\mathbf{y})$ representa uma função do vetor \mathbf{y} . As propriedades das formas quadráticas são importantes para distâncias entre dois vetores, pois espera-se que estas sejam sempre positivas ou nulas (Ferreira, 2008).

A classificação das formas quadráticas pode ser feita de acordo com os resultados que produzem para uma matriz \mathbf{A} . A matriz \mathbf{A} simétrica pode ser classificada de acordo com a forma quadrática $Q(\mathbf{y})$. Tem-se a seguinte regra de classificação das formas quadráticas ($\forall \mathbf{y} \neq \mathbf{0}$):

$$Q(y) \begin{cases} > 0 & \text{positiva definida} \\ \geq 0 & \text{positiva semidefinida} \\ < 0 & \text{negativa definida} \\ \leq 0 & \text{negativa semidefinida} \end{cases} \quad (2.7)$$

A matriz \mathbf{A} correspondente a cada caso de (2.7) recebe a mesma classificação da forma quadrática (Ferreira, 2008).

2.7 Distribuições multivariadas

A família de distribuições elípticas simétricas, com propriedades similares à normal multivariada, tem sido considerada para avaliar a sensibilidade de um teste ou de um processo de estimação por região aos desvios de normalidade de dados multivariados (Ferreira, 2008).

Seja um vetor aleatório $\mathbf{Y} = [Y_1, Y_2, \dots, Y_p]^T \in \mathbb{R}^p$ com média $\boldsymbol{\mu} = [\mu_1, \mu_2, \dots, \mu_p]^T \in \mathbb{R}^p$ e matriz de covariâncias que é função da matriz simétrica e

positiva definida $\Sigma(p \times p)$, as distribuições elípticas simétricas podem ser definidas como funções da forma quadrática $(\mathbf{y} - \boldsymbol{\mu})^\top \Sigma^{-1}(\mathbf{y} - \boldsymbol{\mu})$ da seguinte maneira:

$$f_{\mathbf{Y}}(\mathbf{y}) = |\Sigma|^{-1/2} g[(\mathbf{y} - \boldsymbol{\mu})^\top \Sigma^{-1}(\mathbf{y} - \boldsymbol{\mu})],$$

em que g é uma função definida em $[0, \infty)$ que satisfaz à condição $\int_{\mathbb{R}^p} g(\mathbf{z}^\top \mathbf{z}) d\mathbf{z} = 1$ para todo $\mathbf{z} \in \mathbb{R}^p$.

Teorema 1 (Transformações lineares nas distribuições elípticas). *Se $\mathbf{Y} = [Y_1, Y_2, \dots, Y_p]^\top \in \mathbb{R}^p$ é um vetor aleatório com distribuição elíptica simétrica com parâmetros dados pelo vetor de médias $\boldsymbol{\mu} = [\mu_1, \mu_2, \dots, \mu_p]^\top \in \mathbb{R}^p$ e matriz simétrica e positiva $\Sigma(p \times p)$, então, $\mathbf{Z} = \Sigma^{-1/2}(\mathbf{Y} - \boldsymbol{\mu})$ possui distribuição elíptica simétrica com parâmetros $\boldsymbol{\mu} = 0$ e \mathbf{I} ou, mais precisamente, esférica simétrica, em que $\Sigma^{-1/2}$ é a inversa da matriz raiz quadrada de Σ , tal que $\Sigma = \Sigma^{1/2} \Sigma^{1/2}$.*

Um exemplo da família de distribuições elípticas simétricas é o caso da distribuição normal multivariada. Se a distribuição elíptica for representada por $E_p(\boldsymbol{\mu}, \Sigma, g)$, em que a função g pode ser omitida ou não, a distribuição $N_p(\boldsymbol{\mu}, \Sigma)$ é dada por

$$f_{\mathbf{Y}}(\mathbf{y}) = (2\pi)^{-p/2} |\Sigma|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{y} - \boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{y} - \boldsymbol{\mu}) \right\}$$

para $\mathbf{y} \in \mathbb{R}^p$ e $g(z) = (2\pi)^{-p/2} \exp\{-z/2\}$, $z \geq 0$.

Outra distribuição muito importante é a t de Student multivariada com ν graus de liberdade. A função de densidade do vetor aleatório $\mathbf{Y} = [Y_1, Y_2, \dots, Y_p]^\top \in \mathbb{R}^p$ com parâmetros dados pelo vetor de médias $\boldsymbol{\mu} = [\mu_1, \mu_2, \dots, \mu_p]^\top \in \mathbb{R}^p$ e

matriz simétrica e positiva definida Σ ($p \times p$) é (Giri, 2004) :

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{\Gamma(\frac{\nu+p}{2})}{(\pi\nu)^{1/(2p)}\Gamma(\nu)|\Sigma|^{1/2}} \left[1 + \frac{1}{\nu}(\mathbf{y} - \boldsymbol{\mu})^\top \Sigma^{-1}(\mathbf{y} - \boldsymbol{\mu}) \right]^{-\frac{\nu+p}{2}}.$$

em que g é dada por $g(z) = \frac{\Gamma(\frac{\nu+p}{2})}{(\pi\nu)^{1/(2p)}\Gamma(\nu)} \left(1 + \frac{z}{\nu} \right)^{-\frac{\nu+p}{2}}$. A variável aleatória \mathbf{Y} tem média $\boldsymbol{\mu}$ e matriz de covariâncias $\nu\Sigma/(\nu - 2)$, no caso de $\nu > 2$.

Efetuando-se a transformação $\mathbf{X} = \Sigma^{-1/2}(\mathbf{Y} - \boldsymbol{\mu})$, obtém-se a distribuição t multivariada esférica simétrica, cuja densidade é dada por

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{\Gamma(\frac{\nu+p}{2})}{(\pi\nu)^{1/(2p)}\Gamma(\nu)} \left[1 + \frac{1}{\nu}(\mathbf{x}^\top \mathbf{x}) \right]^{-\frac{\nu+p}{2}}$$

A variável aleatória \mathbf{X} terá vetor de médias nulo, matriz de covariâncias $\nu\mathbf{I}/(\nu - 2)$ ($\nu > 2$) e sua função densidade terá contornos de densidades constantes, esféricos (Giri, 2004).

A distribuição Wishart é a análoga multivariada da distribuição χ^2 e possui usos similares. Uma variável aleatória χ^2 é definida como a soma de variáveis aleatórias univariadas normais padrão ao quadrado (Rencher, 2002). Segundo Ferreira (2008), a distribuição Wishart é utilizada para modelar a distribuição de uma variável aleatória matricial \mathbf{W} de somas de quadrados e produtos originada de uma população normal multivariada. A distribuição Wishart invertida aparece muitas vezes na inferência multivariada. Seja \mathbf{W} uma matriz aleatória tal que $\mathbf{W} \sim W_p(\nu, \Sigma)$, então a distribuição de $\mathbf{S} = \mathbf{W}^{-1}$ é a Wishart invertida $W_p^{-1}(\nu, \Sigma)$.

2.8 Tipos de erro e poder

Ao realizar qualquer inferência, é preciso ter em mente que há um risco de tomar decisões incorretas que deverá ser considerado. Os erros podem ser classificados de três formas (Machado et al., 2005; Mood et al., 1974):

- Erro Tipo I: erro cometido ao se rejeitar a hipótese nula verdadeira, quando deveria ser aceita. É diretamente controlável pelo pesquisador. A probabilidade de se cometer esse erro é dada por: $P[\text{Erro Tipo I}] = P[\text{rejeitar } H_0 | H_0 \text{ verdadeira}] = \alpha$;
- Erro Tipo II: erro cometido ao não se rejeitar a hipótese nula falsa. Não é controlável diretamente pelo pesquisador. A probabilidade de se cometer esse erro é dada por: $P[\text{Erro Tipo II}] = P[\text{não rejeitar } H_0 | H_0 \text{ falsa}] = \beta$;
- Erro Tipo III: erro cometido ao se declarar uma média maior do que a outra quando, na verdade, ocorre o contrário (Carmer & Swanson, 1973). A probabilidade de se incorrer nesse tipo de erro é: $P[\text{Erro Tipo III}] = P[\text{rejeitar } H_0 | H_0 \text{ falsa, mas a favor de } H_1 \text{ errada}]$

O poder do teste é a probabilidade de se rejeitar H_0 , sendo H_0 falsa e é dado por $1 - \beta$. O poder corresponde à capacidade do teste de detectar diferenças.

2.9 Simulação Monte Carlo

A simulação é um processo que tenta reproduzir o comportamento de um sistema real, em geral por meio de programas de computadores, para estudar seu funcionamento sob condições alternativas. O método de Monte Carlo é uma técnica que consiste na simulação de dados por meio da geração de números pseudoaleatórios, usando algum algoritmo em alguma linguagem de programação, de acordo com determinada distribuição de probabilidade. O objetivo é estudar o

comportamento de diferentes técnicas estatísticas que poderiam ser empregadas num problema específico (Dachs, 1988).

Segundo Gentle (2003), o centro da simulação Monte Carlo é a geração desses números pseudoaleatórios. Milhões de números podem ser gerados, mas o tempo requerido para fazê-lo representa apenas uma fração do tempo total de computação. A simulação Monte Carlo é a utilização de experimentos com números aleatórios para avaliar expressões matemáticas. As unidades experimentais são os números aleatórios e as expressões podem ser integrais definidas, sistemas de equações ou modelos matemáticos complexos. O emprego de simulações é essencial para a avaliação de desempenho dos métodos devido às complexas operações matriciais envolvidas, dificultando o julgamento analítico comparativo desses métodos. O método de simulação Monte Carlo é muito utilizado para se determinar propriedades de estimadores ou para se comparar métodos e testes estatísticos.

2.9.1 Avaliação de soluções por meio de simulação

Muitos trabalhos compararam os desempenhos de diferentes soluções frequentistas para o problema de Behrens-Fisher multivariado por meio de simulação Monte Carlo e avaliaram as taxas de erro tipo I e poder dos procedimentos. Alguns desses trabalhos são apresentados a seguir.

Yao (1965) estudou uma extensão da solução de graus de liberdade aproximados de Welch e a comparou com a solução de James (1954), concluindo que a solução proposta é mais conservativa do que a de James.

Subrahmaniam & Subrahmaniam (1973) compararam os desempenhos de três métodos: os de Yao (1965), James (1954) e Bennett (1951). Os autores observaram que o procedimento de James tem o maior poder, mas é muito inferior em relação às taxas de erro tipo I e piora com o aumento da dimensão p , especial-

mente quando a amostra de menor tamanho está associada à maior covariância. A solução de Bennett é exata, mas tem baixo poder, especialmente para tamanhos de amostras muito diferentes.

Segundo Christensen & Rencher (1997), que compararam as soluções de Bennett (1951), James (1954), Yao (1965), Johansen (1980), Nel & Merwe (1986), Hwang & Paulson (1986) e Kim (1992), os procedimentos de James, Yao e Johansen apresentaram taxas de erro tipo I mais elevadas. Dentre os procedimentos que apresentaram taxas de erro tipo I menores, as soluções de Kim e Nel & Merwe apresentaram maior poder, sendo, por isso, recomendadas pelos autores.

Além de terem proposto a modificação do teste de Nel & Merwe (1986), Krishnamoorthy & Yu (2004) avaliaram seu desempenho e compararam sua solução com mais duas soluções também invariantes a transformações não-singulares, as de Yao (1965) e de Johansen (1980), por meio de simulação. A conclusão dos autores foi que a sua solução foi a melhor dentre as três para o problema de Behrens-Fisher multivariado, pois controlou o erro tipo I e apresentou alto poder. Além disso, comparado ao teste de Nel & Merwe (1986) original, a modificação teve igual poder para pequenas dimensões e foi mais poderosa para maiores dimensões. Krishnamoorthy & Yu (2004) frisam que uma solução exata com propriedades naturais não existe, havendo um contínuo interesse em desenvolver soluções aproximadas melhores. Os autores concluem que uma nova solução deverá ser comparada à sua para confirmar se há superioridade.

Cirillo & Ferreira (2004) compararam quatro métodos em suas versões original e *bootstrap* com a estatística T^2 de Hotelling. Os autores observaram que o aumento da correlação ocasionou redução no poder. No caso de homogeneidade de covariâncias, situação ideal do teste T^2 de Hotelling, o método de Kim (1992) original foi o mais apropriado para amostras de tamanhos iguais. Para amostras de

tamanhos diferentes, os métodos de Nel & Merwe (1986) e Yao (1965) apresentaram desempenhos melhores. O método de Johansen (1980) foi o único cuja versão *bootstrap* apresentou controle do erro tipo I.

Lix et al. (2005) avaliaram os procedimentos de James (1954), Yao (1965), Brown-Forsythe (1974), Johansen (1980), Nel & Merwe (1986) e Kim (1992), sob diferentes distribuições, balanceamento e não-balanceamento, e estimadores baseados em mínimos quadrados e estimadores robustos. Os autores recomendam estimadores robustos se não houver normalidade, evitar o método de Johansen (1980) se os tamanhos das amostras forem muito diferentes e utilizar a solução de Kim (1992) na presença de alta heterogeneidade de covariâncias.

2.9.2 Medidas frequentistas para avaliação de métodos bayesianos

Alguns trabalhos utilizaram medidas frequentistas como forma de avaliar o desempenho de procedimentos bayesianos. Em comparações múltiplas, Gopalan & Berry (1998) aplicaram distribuições *a priori* Dirichlet na forma de combinações *priori*/verossimilhança para obter probabilidades *a posteriori* para várias hipóteses de igualdade de médias. O amostrador de Gibbs foi utilizado devido à intratabilidade analítica das distribuições *a posteriori*. Seu procedimento foi mais poderoso do que o teste de Duncan, sob algumas hipóteses alternativas. Porém, não há muita flexibilidade na atribuição de probabilidades *a priori*.

No contexto do problema de Behrens-Fisher univariado, Ghosh & Kim (2001) utilizaram probabilidades de cobertura para avaliar seus intervalos de credibilidade, conforme apresentado na seção 2.5.

Agresti & Min (2005) utilizaram argumentos da teoria frequentista para avaliar o desempenho da inferência bayesiana utilizada para tabelas de contingência, como probabilidades de cobertura e região de confiança. Os autores analisa-

ram o desempenho de intervalos de credibilidade para a diferença de proporções, risco relativo e razão de chances em tabelas de contingência 2×2 . Os autores recomendam intervalos de credibilidade alternativos para o risco relativo e para razão de chances no lugar de intervalos HPD e recomendaram o uso de distribuição *a priori* não-informativa, como a de Jeffreys. Para amostras pequenas, os autores concluíram que utilizar intervalos de credibilidade com distribuições *a priori* mais informativas tendem a apresentar desempenho pior do que os intervalos frequentistas, demonstrando que, nem sempre, o uso da inferência bayesiana leva a resultados melhores.

2.9.3 Simulação de dados de distribuições multivariadas

A geração de amostras aleatórias de distribuições multivariadas é mais difícil de ser realizada do que a geração de amostras aleatórias de distribuições univariadas (Johnson, 1987). A seguir, serão apresentadas formas de gerar amostras aleatórias das distribuições normal e *t* multivariadas.

Para gerar uma amostra aleatória $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_n$, de tamanho n de uma normal com vetor de médias $\boldsymbol{\mu}$ e matriz de covariâncias $\boldsymbol{\Sigma}$, em que $\mathbf{Y}_j = [Y_{j1}, Y_{j2}, \dots, Y_{jp}]^\top$ é a j -ésima observação amostral, deve-se seguir os passos a seguir. Gera-se um vetor aleatório $\mathbf{Z} = [Z_1, Z_2, \dots, Z_p]^\top$ a partir de uma distribuição normal com vetor de médias $\mathbf{0}$ e matriz de covariâncias \mathbf{I} . Para gerar cada elemento desse vetor, pode-se utilizar a transformação de Box-Müller (Press et al., 1992). Sejam duas variáveis aleatórias uniformes U_1 e U_2 e as transformações

$$z_1 = \sqrt{-2 \ln(u_1)} \cos(2\pi u_2) \text{ e } z_2 = \sqrt{-2 \ln(u_1)} \sin(2\pi u_2). \quad (2.8)$$

Ao isolar U_1 e U_2 em (2.8), obtém-se

$$u_1 = e^{-\frac{1}{2}(z_1^2+z_2^2)} \text{ e } u_2 = \frac{1}{2\pi} \arctan\left(\frac{z_2}{z_1}\right).$$

O jacobiano da transformação é

$$J = \begin{vmatrix} \frac{\partial u_1}{\partial z_1} & \frac{\partial u_1}{\partial z_2} \\ \frac{\partial u_2}{\partial z_1} & \frac{\partial u_2}{\partial z_2} \end{vmatrix} = \begin{vmatrix} -z_1 e^{-0,5(z_1^2+z_2^2)} & -z_2 e^{-0,5(z_1^2+z_2^2)} \\ -\frac{z_2}{2\pi z_1^2 \left(1+\frac{z_2^2}{z_1^2}\right)} & \frac{1}{2\pi z_1 \left(1+\frac{z_2^2}{z_1^2}\right)} \end{vmatrix} = -\frac{1}{2\pi} e^{-0,5(z_1^2+z_2^2)}.$$

Dessa forma, a função de densidade conjunta de Z_1 e Z_2 é dada por $f(u_1, u_2)|J|$, sendo

$$f(z_1, z_2) = \left[\frac{1}{\sqrt{2\pi}} e^{-\frac{z_1^2}{2}} \right] \left[\frac{1}{\sqrt{2\pi}} e^{-\frac{z_2^2}{2}} \right], \quad (2.9)$$

que é a função distribuição de probabilidade conjunta de duas variáveis aleatórias independentes Z_1 e Z_2 , ambas tendo distribuição normal padrão.

Por meio de alguma função para a geração de números aleatórios, geram-se duas variáveis aleatórias uniformes em cada passo e, utilizando a equação (2.8), obtêm-se duas realizações de variáveis normais $N(0,1)$ independentes. Esse procedimento é repetido até os p componentes do vetor \mathbf{Z} serem obtidos. Outra forma de se obter o mesmo resultado é, a partir de uma variável aleatória uniforme, gerar uma variável normal padrão, pelo teorema fundamental da transformação de probabilidades. Assim, é preciso inverter a função de distribuição da variável a ser gerada no ponto u , em que u representa a realização de uma variável aleatória uniforme U . Para gerar uma variável normal padrão, deve-se gerar uma realização u de uma variável aleatória uniforme $U(0,1)$ e obter z , tal que

$$u = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{t^2}{2}\right\} dt$$

ou seja, deve-se obter $z = \Phi^{-1}(u)$, em que $\Phi^{-1}(u)$ é a inversa da função de distribuição acumulada de u . O procedimento é repetido p vezes para obter os elementos do vetor \mathbf{Z} .

Para se gerar uma variável aleatória com distribuição normal $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, deve-se aplicar a transformação linear do vetor \mathbf{Z} dada por $\mathbf{Y} = \boldsymbol{\Sigma}^{1/2} \mathbf{Z} + \boldsymbol{\mu}$, em que $\boldsymbol{\Sigma}^{1/2}$ é a matriz raiz quadrada de $\boldsymbol{\Sigma}$. De acordo com o teorema das combinações lineares, \mathbf{Y} possui distribuição normal multivariada com média $\boldsymbol{\mu}$ e matriz de covariâncias $\boldsymbol{\Sigma}$. Se o processo for repetido n vezes, a amostra desejada terá sido obtida. Há a possibilidade de se utilizar a decomposição de Cholesky de $\boldsymbol{\Sigma} = \boldsymbol{\Gamma} \boldsymbol{\Gamma}^\top$ e a combinação linear $\mathbf{X} = \boldsymbol{\Gamma} \mathbf{Z} + \boldsymbol{\mu}$ para se gerar n vetores aleatórios amostrais (Ferreira, 2008).

Para se gerar variáveis aleatórias p -dimensionais da distribuição t multivariada com ν graus de liberdade e parâmetros $\boldsymbol{\mu}$ e $\boldsymbol{\Sigma}$, Ferreira (2008) sugere proceder-se da forma a seguir. Seja um vetor aleatório \mathbf{Z} com distribuição $N_p(\mathbf{0}, \mathbf{I})$ e a variável aleatória U com distribuição qui-quadrado com ν graus de liberdade, então o vetor aleatório \mathbf{X} , dado pela transformação

$$\mathbf{X} = \sqrt{\nu} \frac{\mathbf{Z}}{\sqrt{U}}, \quad (2.10)$$

possui distribuição t multivariada esférica com ν graus de liberdade. O vetor \mathbf{Y} obtido pela transformação linear

$$\mathbf{Y} = \Sigma^{1/2} \mathbf{X} + \boldsymbol{\mu} \quad (2.11)$$

possui distribuição t multivariada elíptica com ν graus de liberdade e parâmetros $\boldsymbol{\mu}$ e Σ (Teorema 1). Dessa forma, devem-se aplicar (2.10) e (2.11) n vezes para se obter uma amostra de tamanho n da distribuição t multivariada.

Outra maneira de se obterem amostras da distribuição t multivariada elíptica é utilizando-se a decomposição de Cholesky. Simula-se o vetor \mathbf{X} da distribuição t multivariada esférica com vetor de médias $\mathbf{0}$, parâmetro de covariação \mathbf{I} e ν graus de liberdade, ou seja, $\mathbf{X} \sim t_p(\mathbf{0}, \mathbf{I}, \nu)$. Tendo-se $\Sigma = \Gamma\Gamma^\top$, faz-se a transformação $\mathbf{Y} = \Gamma\mathbf{X} + \boldsymbol{\mu}$. Dessa forma, obtém-se $\mathbf{Y} \sim t_p(\boldsymbol{\mu}, \Sigma, \nu)$.

Uma das dificuldades do método é obter n variáveis qui-quadrado. Estas podem ser geradas utilizando o teorema fundamental da transformação de probabilidades, o que pode ser feito por meio de algoritmos para se inverter a função de distribuição qui-quadrado ou utilizarem-se os métodos de aceitação/rejeição (Press et al., 1992).

3 METODOLOGIA

Uma solução bayesiana foi proposta para o problema de Behrens-Fisher multivariado, inspirada na solução apresentada por Johnson & Weerahandi (1988). O desempenho do método proposto neste trabalho foi avaliado em relação às taxas de erro tipo I e poder. Para isso, foram realizadas simulações Monte Carlo sob a hipótese nula (H_0) e sob a hipótese alternativa (H_1), em situações de diferentes graus de heterogeneidade de covariâncias das duas populações.

3.1 Solução bayesiana proposta

Sejam Y_{11}, \dots, Y_{1n_1} e Y_{21}, \dots, Y_{2n_2} amostras aleatórias independentes de tamanhos n_1 e n_2 obtidas de duas populações normais p -variadas $N_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ e $N_p(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$, respectivamente, em que $\boldsymbol{\mu}_i$ é o vetor de médias $p \times 1$ e $\boldsymbol{\Sigma}_i$ a matriz de covariâncias $p \times p$ da i -ésima população, $i = 1, 2$. Para caracterizar o problema de Behrens-Fisher, as matrizes $\boldsymbol{\Sigma}_1$ e $\boldsymbol{\Sigma}_2$ são diferentes e desconhecidas. Para a realização de inferências sobre o vetor resultante da diferença dos vetores de médias populacionais $\boldsymbol{\delta} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$, foi proposto um teste bayesiano para se testar a hipótese nula

$$H_0 : \boldsymbol{\delta} = \boldsymbol{\delta}_0 = \begin{bmatrix} \delta_{01} \\ \delta_{02} \\ \vdots \\ \delta_{0p} \end{bmatrix} \quad (3.1)$$

e um procedimento para a obtenção de uma região de credibilidade $100(1 - \alpha)\%$. Particularmente, o interesse no teste da hipótese nula (3.1) foi direcionado para

o caso em que $\boldsymbol{\delta}_0 = [0, 0, \dots, 0]^\top$, ou seja, de igualdade dos vetores de médias das duas populações. Porém, o valor de $\boldsymbol{\delta}_0$ pode ser escolhido de acordo com o interesse do pesquisador.

Para a construção do teste, utilizou-se o procedimento descrito em Johnson & Weerahandi (1988) e apresentado na sequência. Inicialmente, foram obtidos os vetores de médias amostrais e as matrizes de somas de quadrados e produtos por

$$\bar{\mathbf{Y}}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \mathbf{Y}_{ij}, \quad (3.2)$$

$$\mathbf{V}_i = \sum_{j=1}^{n_i} (\mathbf{Y}_{ij} - \bar{\mathbf{Y}}_i)(\mathbf{Y}_{ij} - \bar{\mathbf{Y}}_i)^\top, \quad (3.3)$$

em que $i = 1, 2$.

Foi utilizada uma distribuição *a priori* conjugada para $\boldsymbol{\mu}_i$ e $\boldsymbol{\Sigma}_i$, conforme apresentado por Johnson & Weerahandi (1988), em que $\boldsymbol{\mu}_i | \boldsymbol{\Sigma}_i \sim N_p(\mathbf{a}_i, \boldsymbol{\Sigma}_i/q_i)$ e $\boldsymbol{\Sigma}_i \sim W_p^{-1}(\mathbf{R}_i, r_i)$, sendo W_p^{-1} a distribuição Wishart invertida com dimensão p . Nesse caso, \mathbf{a}_i , q_i , \mathbf{R}_i e r_i são hiperparâmetros. Fazendo $\mathbf{a}_i = \mathbf{0}$, $q_i \rightarrow 0$ e $\mathbf{R}_i \rightarrow \mathbf{0}$ ($p \times p$), então, a distribuição *a posteriori* de $\boldsymbol{\mu}_i$ é $T_p(\bar{\mathbf{Y}}_i, \mathbf{V}_i/[n_i(n_i + r_i - 2p)], n_i + r_i - 2p)$, em que $T_p(\mathbf{A}, \nu)$ é uma distribuição p -dimensional t multivariada com parâmetro de covariância \mathbf{A} e graus de liberdade ν . O hiperparâmetro r_i foi escolhido de tal forma que $n_i + r_i - 2p$ variasse entre $n_i - p$ e $n_i - 1$. Assim, r_i deve variar entre p e $2p - 1$. Verificou-se que, quanto mais próximo de n_i o valor de p se encontrava, pior era o desempenho do teste proposto. O teste se tornava liberal se $n_i + r_i - 2p$ fosse substituído por $n_i - 1$. De forma semelhante, se o valor de $n_i + r_i - 2p$ se aproximasse de $n_i - p$, o teste se tornava conservativo.

Foi verificado que o desempenho do teste caminhava de liberal para conservativo de uma forma não-linear dependente de r_i . Assim, a informação necessária para se estabelecer o valor do hiperparâmetro r_i estava contida em p e n_i , expressa como uma função de p/n_i . Portanto, o valor de r_i foi definido empiricamente por

$$r_i = (2p - 1) \left(\frac{p}{2p - 1} \right)^{\left(\frac{p}{n_i} \right)^3}, \quad (3.4)$$

que varia entre p e $2p - 1$.

Para se realizar inferência sobre $\boldsymbol{\delta} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$ é necessário obter-se a convolução de duas t multivariadas, o que é uma estratégia muito difícil. Assim, optou-se por utilizar uma função de $\boldsymbol{\delta}$ e de $\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2$ e obter sua distribuição *a posteriori* computacionalmente por meio de simulação Monte Carlo. Isso é o que diferencia o presente trabalho da proposta de Johnson & Weerahandi (1988), que obtiveram uma distribuição analítica dessa função e calcularam seus quantis *a posteriori* por meio de métodos numéricos, utilizando a distribuição exata e algumas aproximações. Essa função é dada pela forma quadrática

$$q = (\boldsymbol{\delta} - \mathbf{d})^\top \mathbf{V}_p^{-1} (\boldsymbol{\delta} - \mathbf{d}),$$

em que $\mathbf{d} = \bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2$, $\boldsymbol{\delta} = \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2$ e \mathbf{V}_p (*V pooled*) é uma combinação linear das matrizes \mathbf{V}_1 e \mathbf{V}_2 dada por

$$\mathbf{V}_p = \frac{(n_1 - 1)\mathbf{V}_1 + (n_2 - 1)\mathbf{V}_2}{n_1 + n_2 - 2}. \quad (3.5)$$

Como a função de distribuição de Q *a posteriori* obtida por Johnson & Weerahandi (1988) é muito complexa e envolve uma série infinita, a proposta deste trabalho é obtê-la computacionalmente. Assim, dadas as amostras multivariadas das duas populações normais consideradas $\mathbf{Y}_{11}, \dots, \mathbf{Y}_{1n_1}$ e $\mathbf{Y}_{21}, \dots, \mathbf{Y}_{2n_2}$, foram estimadas as médias ($\bar{\mathbf{Y}}_i$) e as matrizes de somas de quadrados e produtos (\mathbf{V}_i) utilizando (3.2) e (3.3), $i = 1, 2$. Utilizando-se $\bar{\mathbf{Y}}_i$ como média da distribuição *a posteriori* de $\boldsymbol{\mu}_i$, que foi considerada uma $T_p(\bar{\mathbf{Y}}_i, \mathbf{V}_i/[n_i(n_i + r_i - 2p)], n_i + r_i - 2p)$, simularam-se amostras aleatórias de tamanho N . Sendo $\boldsymbol{\mu}_{ij}$ a j -ésima observação p -variada dessa distribuição *a posteriori* correspondente à i -ésima população, com $j = 1, 2, \dots, N$ e $i = 1, 2$, foi obtida a quantidade

$$q_j = (\boldsymbol{\mu}_{1j} - \boldsymbol{\mu}_{2j} - \mathbf{d})^\top \mathbf{V}_p^{-1} (\boldsymbol{\mu}_{1j} - \boldsymbol{\mu}_{2j} - \mathbf{d}). \quad (3.6)$$

O valor esperado de $\boldsymbol{\mu}_{1j} - \boldsymbol{\mu}_{2j} - \mathbf{d}$ é o vetor nulo $p \times 1$. Assim, a distribuição de q_j corresponde à distribuição sob a hipótese nula de igualdade das médias populacionais. Para se aplicar o teste da hipótese nula $H_0 : \boldsymbol{\delta} = \boldsymbol{\delta}_0$, foi obtido o valor da quantidade

$$q_c = (\mathbf{d} - \boldsymbol{\delta}_0)^\top \mathbf{V}_p^{-1} (\mathbf{d} - \boldsymbol{\delta}_0). \quad (3.7)$$

O valor da probabilidade empírica, denotado por credibilidade empírica C , usado como evidência para decidir se a hipótese nula deve ou não ser rejeitada foi obtido por

$$C = \frac{\sum_{j=1}^N I(q_c \leq q_j)}{N},$$

em que $I(q_c \leq q_j)$ é a função indicadora tal que $I(q_c \leq q_j) = 1$ se q_j supera o valor original q_c ou 0, caso contrário. O valor de N considerado foi de 2000.

Para se obter uma região de $100(1 - \alpha)\%$ de credibilidade, os valores de q_j foram ordenados e o quantil superior q_α da distribuição *a posteriori* foi obtido.

A região de $100(1 - \alpha)\%$ de credibilidade para δ foi estabelecida por

$$RC_{1-\alpha}(\delta) : \{\delta \in \mathbb{R}^p | (\delta - \mathbf{d})^\top \mathbf{V}_p^{-1} (\delta - \mathbf{d}) \leq q_\alpha\}.$$

3.2 Simulação Monte Carlo

Para se avaliar o desempenho da proposta realizada empregou-se simulação Monte Carlo utilizando o programa R (R Development Core Team, 2009). Duas etapas foram consideradas. Na primeira, avaliou-se o desempenho em relação ao erro tipo I e, na segunda, em relação ao poder, simulando-se amostras sob H_0 e sob H_1 , respectivamente.

Sob a hipótese nula de igualdade dos vetores de médias, foram realizadas $M = 2000$ simulações Monte Carlo para a obtenção das taxas de erro tipo I. Foram simuladas amostras de duas distribuições normais $N_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ e $N_p(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$, com tamanhos n_1 e n_2 , em que p é a dimensão (número de variáveis) e $\boldsymbol{\Sigma}_i$ ($p \times p$) é a matriz de covariâncias da i -ésima população, assumida positiva definida. Sob a hipótese nula

$$H_0 : \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2, \quad (3.8)$$

e, sem perda de generalidade, os vetores de médias foram considerados iguais ao vetor nulo p -dimensional $\mathbf{0}$.

Sob a hipótese alternativa $H_1 : \boldsymbol{\mu}_1 \neq \boldsymbol{\mu}_2$, foram realizadas M simulações Monte Carlo em uma segunda etapa para o cálculo do poder. Novamente, foram simuladas amostras de duas distribuições normais $N_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ e $N_p(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$, com tamanhos n_1 e n_2 , em que as médias diferiam entre si por k erros padrões ($k = 2, 4, 8, 16$). A média da população 2, $\boldsymbol{\mu}_2$, foi obtida a partir da média da população 1, $\boldsymbol{\mu}_1$:

$$\boldsymbol{\mu}_2 = \boldsymbol{\mu}_1 + k\sqrt{\text{diag}(\boldsymbol{\Sigma}_{\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2})}$$

sendo $\boldsymbol{\Sigma}_{\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2} = \boldsymbol{\Sigma}_1/n_1 + \boldsymbol{\Sigma}_2/n_2$ a matriz de covariâncias da diferença entre as médias e, sem perda de generalidade, $\boldsymbol{\mu}_1$ foi fixado como o vetor nulo, ou seja, $\boldsymbol{\mu}_1 = [0 \cdots 0]^\top$ e

$$\sqrt{\text{diag}(\boldsymbol{\Sigma}_{\bar{\mathbf{Y}}_1 - \bar{\mathbf{Y}}_2})} = \begin{bmatrix} \sqrt{\frac{\sigma_{11}^1}{n_1} + \frac{\sigma_{11}^2}{n_2}} \\ \sqrt{\frac{\sigma_{22}^1}{n_1} + \frac{\sigma_{22}^2}{n_2}} \\ \vdots \\ \sqrt{\frac{\sigma_{pp}^1}{n_1} + \frac{\sigma_{pp}^2}{n_2}} \end{bmatrix}$$

em que $\sigma_{\ell\ell}^i$ corresponde à variância da ℓ -ésima variável ($\ell = 1, \dots, p$) na i -ésima

população ($i = 1, 2$), de tamanho n_i .

Em ambos os casos (sob H_0 e H_1), as covariâncias foram

$$\Sigma_1 = \sigma^2 \begin{bmatrix} 1 & \rho & \rho & \cdots & \rho \\ \rho & 1 & \rho & \cdots & \rho \\ \rho & \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \rho & \cdots & 1 \end{bmatrix}, \Sigma_2 = \gamma_p \Sigma_1, \quad (3.9)$$

em que $\gamma_p = \sqrt{\gamma}$ e $\gamma = |\Sigma_2|/|\Sigma_1|$ refere-se ao grau de heterogeneidade das covariâncias entre as duas populações, sendo considerados os seguintes valores de γ : 1, 2, 8, 16 e 32. O valor $\gamma = 1$ determina o caso de homogeneidade das matrizes de covariâncias das duas populações e, portanto, não contempla o problema de Behrens-Fisher multivariado, sendo utilizado apenas para comparação do desempenho com os outros valores de γ .

Após obtidas as amostras de ambas as populações em cada uma das M simulações Monte Carlo, os seguintes testes foram aplicados: o teste bayesiano proposto na seção 3.1 e o teste Nel & Merwe modificado apresentado por Krishnamoorthy & Yu (2004).

Sob H_0 , as taxas de erro tipo I foram obtidas pela proporção de casos, em relação às M simulações realizadas, em que a hipótese nula foi rejeitada, ou seja,

$$\text{taxa de erro tipo I} = \frac{\sum_{j=1}^M I(q_{c_j} \geq q_{\alpha_j})}{M} \quad (3.10)$$

em que $I(\cdot)$ é a função indicadora que retorna 1 se a desigualdade se verifica e

0, caso contrário, q_{cj} corresponde à estatística do teste obtida sob H_0 e $q_{\alpha j}$ é o quantil $100\alpha\%$, obtidos na j -ésima simulação.

Testes binomiais exatos considerando um nível nominal de significância de 0,01 para as hipóteses $H_0 : \alpha = \alpha_0$ e $H_1 : \alpha \neq \alpha_0$ foram aplicados para cada um dos valores obtidos das taxas de erro tipo I, sendo os valores testados de α_0 iguais a 0,10, 0,05 e 0,01. Se a hipótese nula for rejeitada e as taxas de erro tipo I observadas forem consideradas significativamente (valor-p $<$ 0,01) inferiores ao nível nominal de significância considerado, o teste é conservativo. Se a hipótese nula for rejeitada e as taxas de erro tipo I observadas forem consideradas significativamente (valor-p $<$ 0,01) superiores ao nível nominal de significância, o teste é liberal. Porém, se as taxas de erro tipo I observadas não forem consideradas significativamente (valor-p $>$ 0,01) diferentes do nível nominal de significância considerado, o teste é exato.

Sob a hipótese alternativa H_1 , as simulações foram realizadas para avaliar o poder do teste, ou seja, a proporção de casos, em relação às M simulações realizadas, em que a hipótese nula foi rejeitada corretamente, ou seja,

$$\text{poder} = \frac{\sum_{j=1}^M I(q_{cj} \geq q_{\alpha j})}{M}$$

em que q_{cj} corresponde à estatística do teste obtida sob H_1 e $q_{\alpha j}$ é o quantil $100\alpha\%$, obtidos na j -ésima simulação.

Foram considerados, nas simulações Monte Carlo, valores de p iguais a 2 e 5 e fixados $\rho = 0,5$ e $\sigma^2 = 1$, uma vez que os testes são invariantes a transformações não-singulares. Os valores dos tamanhos amostrais foram fixados em $n_1, n_2 = 8, 15, 20, 30$ e 100, restritos a $n_1 > p$ e $n_2 > p$ para garantir a singularidade da

matriz Σ . Para o erro tipo I, foram consideradas as combinações de n_1 e n_2 dadas por: (8,8), (15,15), (20,20), (30,30), (100,100), (8,15), (15,8), (8,30) e (30,8). Para o poder, as combinações consideradas foram (15,15), (30,30), (100,100), (8,15), (15,8), (8,30) e (30,8). Um total de 90 configurações foi simulado para o erro tipo I e de 280 para o poder. O programa desenvolvido em R (R Development Core Team, 2009) para a simulação e para a aplicação do teste de Nel & Merwe modificado (Krishnamoorthy & Yu, 2004) se encontra no Anexo A e para a aplicação do teste bayesiano proposto se encontra no Anexo B.

3.3 Teste de Nel & Merwe modificado

Para aplicar o teste de Nel & Merwe modificado (Krishnamoorthy & Yu, 2004), alteração do trabalho original de Nel & Merwe (1986), é necessário calcular inicialmente o valor da estatística

$$T_c^2 = (\bar{Y}_1 - \bar{Y}_2 - \delta_0)^\top \left(\frac{V_1}{n_1(n_1 - 1)} + \frac{V_2}{n_2(n_2 - 1)} \right)^{-1} (\bar{Y}_1 - \bar{Y}_2 - \delta_0)$$

e a hipótese nula de igualdade dos vetores de médias das duas populações (3.8) deve ser rejeitada, ou seja, com $\delta_0 = \mathbf{0}$, se $T_c^2 > \nu p F_{\alpha, p, \nu+1-p} / (\nu + 1 - p)$, em que ν é calculado pela expressão (2.3) e $F_{\alpha, p, \nu+1-p}$ é o quantil superior 100 α % da distribuição F com p e $\nu + 1 - p$ graus de liberdade.

Esse teste foi aplicado a cada situação simulada e os resultados da simulação Monte Carlo foram utilizados para fins de comparação com a proposta de teste apresentada neste trabalho.

3.4 Exemplo real

Para ilustrar a aplicação do teste bayesiano proposto e do teste de Nel & Merwe modificado (Krishnamoorthy & Yu, 2004), foi utilizado um exemplo real. O interesse do estudo original era testar a hipótese de igualdade dos vetores de médias das variáveis medidas nos solos de capoeira nova e de pastagem da Amazônia, sendo elas: Ca (cálcio), Mg (magnésio) e SB (saturação de bases). Assim, $p = 3$ e os tamanhos amostrais diferentes eram $n_1 = 30$ de capoeira nova e $n_2 = 13$ de pastagem. Os dados se encontram na Tabela 1, a seguir:

TABELA 1 Dados para as amostras de solos de capoeira nova (população 1) e de pastagem da Amazônia (população 2), com tamanhos amostrais $n_1 = 30$ e $n_2 = 13$.

Pop.	Ca	Mg	SB	Pop.	Ca	Mg	SB	Pop.	Ca	Mg	SB
1	3,1	2,1	5,3	1	5,1	2,2	7,4	1	6,4	2,6	9,1
1	4,1	2,4	6,8	1	7,3	1,7	9,2	2	2,3	1,7	4,1
1	8,3	2,6	11,1	1	5,7	1,2	7,0	2	2,5	2,5	5,1
1	5,2	2,3	7,7	1	8,5	3,6	12,3	2	1,8	2,1	4,1
1	8,5	4,1	12,8	1	7,3	2,2	9,6	2	3,4	2,5	6,1
1	4,2	2,5	6,9	1	5,2	1,9	7,3	2	1,8	1,1	3,0
1	10,5	4,8	15,5	1	5,0	2,5	7,7	2	3,7	1,4	5,2
1	6,9	3,1	10,4	1	7,8	2,2	10,3	2	1,4	0,7	2,2
1	3,4	1,6	5,3	1	6,2	1,8	8,2	2	1,5	0,6	2,2
1	2,6	1,8	4,5	1	4,4	2,1	6,6	2	2,8	2,2	5,1
1	7,6	2,1	10,0	1	7,9	3,1	11,2	2	1,4	0,8	2,3
1	5,4	3,6	9,2	1	7,0	2,3	9,6	2	1,8	0,6	2,5
1	4,9	2,1	7,2	1	5,0	2,5	7,6	2	1,9	1,7	3,7
1	10,4	4,8	15,5	1	2,9	1,2	4,4	2	2,8	0,8	3,7
1	7,1	2,2	9,5								

FONTE: Ferreira (2008, p. 300).

Os dois testes foram aplicados a esses dados, adotando-se os níveis nominais de significância de 0,10, 0,05 e 0,01.

4 RESULTADOS E DISCUSSÃO

A seguir serão apresentados os resultados das taxas de erro tipo I e poder obtidos com as simulações sob H_0 e H_1 e com diferentes números de variáveis p . Foram considerados níveis nominais de significância de 0,10, 0,05 e 0,01, diferentes graus de heterogeneidade de variâncias para as duas populações e diferentes tamanhos amostrais. Em todos os casos simulados sob a hipótese nula, testes para a hipótese $H_0 : \alpha = \alpha_0$ foram realizados considerando-se $\alpha_0 = 0,10, 0,05$ e $0,01$ e um coeficiente de confiança de 0,99.

4.1 Taxas de erro tipo I

Na Tabela 2 são apresentadas as taxas de erro tipo I dos testes bayesiano proposto (TB) e Nel & Merwe modificado (TNVM) (Krishnamoorthy & Yu, 2004), em função de n_1, n_2 e γ , sob H_0 , para $p = 2$ e $\alpha = 0,10$. Para tamanhos amostrais iguais, ou seja, $n_1 = n_2$, os dois testes controlaram as taxas de erro para todos os graus de heterogeneidade de covariâncias, ou seja, os testes apresentaram taxas de erro próximas ao nível nominal estabelecido para todos os valores de γ . Para tamanhos amostrais diferentes, $n_1 \neq n_2$, o TB apresentou comportamento conservativo em algumas situações, sendo esse fato mais evidente no caso em que a amostra de maior tamanho estava associada à população com menor covariância ($n_1 = 30, n_2 = 8$), em que as taxas de erro foram menores do que o nível nominal de significância adotado de 0,10 para todos os valores de γ . Deve-se ressaltar que menor covariância implica em menor variância generalizada, ou seja, menor $|\Sigma_i|$, $i = 1, 2$. O TNVM controlou as taxas de erro em todos os casos, com tamanhos iguais ao valor nominal.

Na Figura 1 são apresentadas as taxas de erro dos testes bayesiano e Nel & Merwe modificado (Krishnamoorthy & Yu, 2004) para $p = 2, \alpha = 0,05$ e

TABELA 2 Taxas de erro tipo I do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), em função dos tamanhos amostrais n_1 e n_2 e do grau de heterogeneidade (γ) das variâncias, para $p = 2$ variáveis e nível nominal de significância $\alpha = 0,10$, sob H_0 .

γ	$n_1 = 8, n_2 = 8$		$n_1 = 15, n_2 = 15$		$n_1 = 20, n_2 = 20$	
	TB	TNVM	TB	TNVM	TB	TNVM
1	0,0925	0,1035	0,0885	0,0950	0,0945	0,1045
2	0,0870	0,0980	0,1025	0,1075	0,0995	0,1055
8	0,0945	0,0945	0,1085	0,1130	0,1035	0,1085
16	0,0940	0,0965	0,1025	0,1030	0,0960	0,0925
32	0,1105	0,1075	0,1015	0,0990	0,1040	0,1030
γ	$n_1 = 30, n_2 = 30$		$n_1 = 100, n_2 = 100$		$n_1 = 8, n_2 = 15$	
	TB	TNVM	TB	TNVM	TB	TNVM
1	0,0980	0,1050	0,0925	0,0945	0,0970	0,1170
2	0,0910	0,0930	0,1130	0,1145	0,0860	0,1030
8	0,0955	0,0970	0,1105	0,1105	0,0720 ⁺⁺	0,0910
16	0,0970	0,0960	0,0950	0,0930	0,0940	0,1035
32	0,0890	0,0870	0,1075	0,1075	0,0895	0,0975
γ	$n_1 = 15, n_2 = 8$		$n_1 = 8, n_2 = 30$		$n_1 = 30, n_2 = 8$	
	TB	TNVM	TB	TNVM	TB	TNVM
1	0,0845	0,0970	0,0730 ⁺⁺	0,1035	0,0740 ⁺⁺	0,0965
2	0,0860	0,1045	0,0725 ⁺⁺	0,1065	0,0680 ⁺⁺	0,1065
8	0,0790 ⁺⁺	0,0975	0,0780 ⁺⁺	0,1005	0,0675 ⁺⁺	0,1085
16	0,0920	0,1095	0,0850	0,1100	0,0700 ⁺⁺	0,0975
32	0,0940	0,1045	0,0840	0,1015	0,0790 ⁺⁺	0,0970

⁺⁺ Significativamente diferente (valor- $p < 0,01$) e considerado menor do que o nível de significância $\alpha = 0,10$.

** Significativamente diferente (valor- $p < 0,01$) e considerado maior do que o nível de significância $\alpha = 0,10$.

diferentes γ , sob H_0 . Nos casos de tamanhos amostrais iguais (Figuras 1a, 1b e 1c), os testes controlaram as taxas de erro tipo I, com apenas uma exceção, em que o TB se mostrou conservativo para $n_1 = n_2 = 8$ e $\gamma = 2$ (Figura 1a). A situação mais extrema, de $n_1 = 30$ e $n_2 = 8$, foi ilustrada na Figura 1d, em que o TB foi conservativo e o TNVM apresentou tamanho de teste não significativamente diferente (valor- $p < 0,01$) do nível nominal de significância para todos os graus de heterogeneidade γ , para qualquer grau de heterogeneidade.

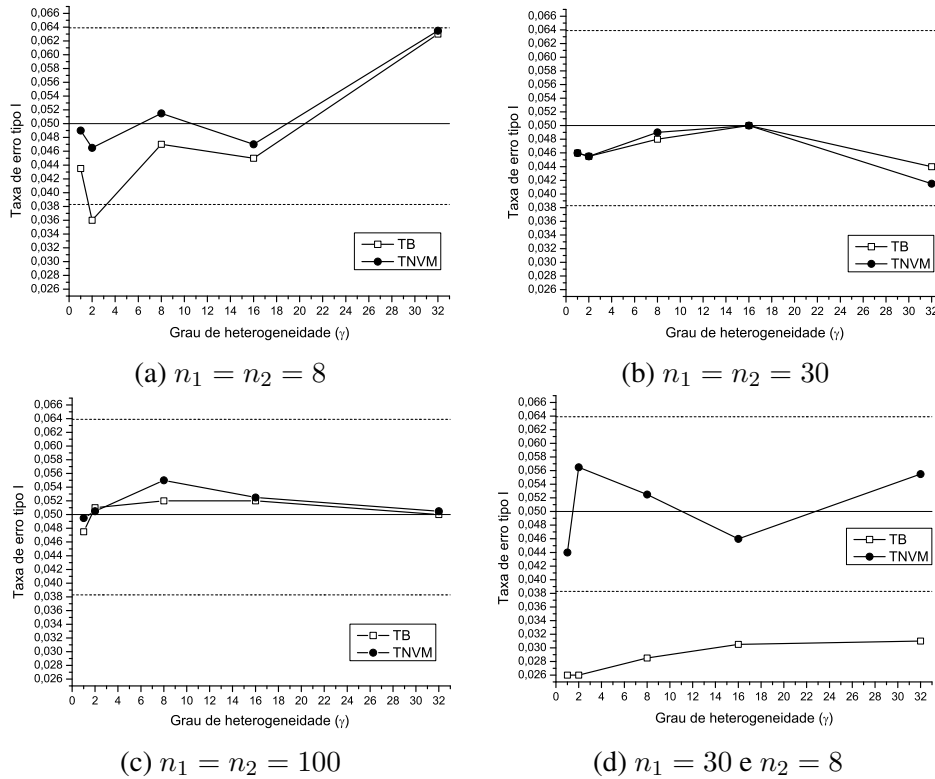


FIGURA 1 Taxas de erro tipo I do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), para $p = 2$, diferentes graus de heterogeneidade γ e $\alpha = 0,05$, sob H_0 . As linhas pontilhadas representam os limites em que o teste de hipótese $H_0 : \alpha = 0,05$ seria rejeitado, considerando um coeficiente de confiança de 0,99.

Na Figura 2 são apresentadas as taxas de erro dos testes bayesiano e Nel & Merwe modificado (Krishnamoorthy & Yu, 2004) para $p = 2$ e $\alpha = 0,01$, considerando diferentes γ , sob H_0 . O mesmo comportamento observado na Figura 1 foi apresentado nessa situação, com os dois testes controlando as taxas de erro nos casos de tamanhos amostrais iguais (Figuras 2a, 2b e 2c) e com comportamentos muito semelhantes, tendo o TB quase sempre sido conservativo para $\gamma = 1$ e 2, com $n_1 = n_2 = 8$ (Figura 2a). A característica conservativa do teste bayesiano

se manteve na situação ilustrada na Figura 2d, em que os tamanhos amostrais são diferentes.

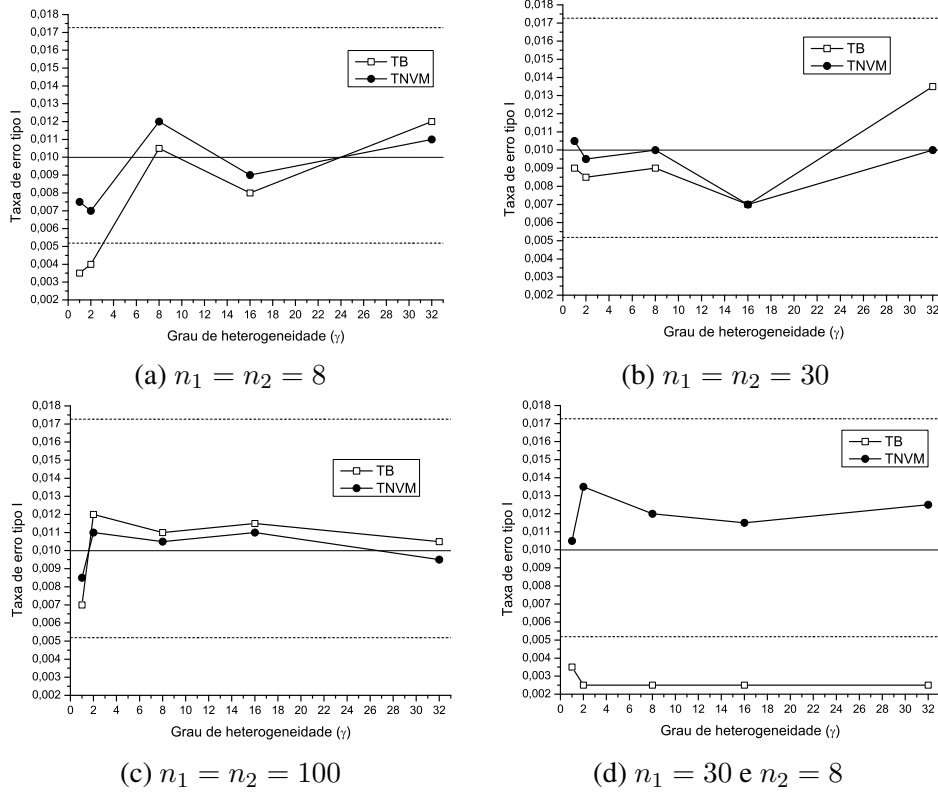


FIGURA 2 Taxas de erro tipo I do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), para $p = 2$, diferentes graus de heterogeneidade γ e $\alpha = 0,01$, sob H_0 . As linhas pontilhadas representam os limites em que o teste de hipótese $H_0 : \alpha = 0,01$ seria rejeitado, considerando um coeficiente de confiança de 0,99.

Na Tabela 3 são apresentadas as taxas de erro tipo I dos testes bayesiano proposto e Nel & Merwe modificado (Krishnamoorthy & Yu, 2004), em função de n_1, n_2 e γ , sob H_0 , para $p = 5$ e $\alpha = 0,10$. Para tamanhos amostrais iguais, os dois testes apresentaram taxas de erro próximas ao nível nominal estabelecido, tendo o teste bayesiano sido considerado liberal em algumas situações e conservativo ape-

nas no caso de tamanhos amostrais pequenos ($n_1 = n_2 = 8$) e sob homogeneidade de covariâncias ($\gamma = 1$). Porém, para tamanhos amostrais diferentes, o comportamento conservativo do TB observado para $p = 2$ e $\alpha = 0,10$ (Tabela 2) foi mais intenso e essa característica se acentuou com o aumento da diferença entre n_1 e n_2 e se tornou mais evidente no caso de $n_1 = 30$ e $n_2 = 8$. Nessa situação, o TNVM, que se apresentou liberal apenas em algumas situações de tamanhos amostrais diferentes, tornou-se liberal para todos os graus de heterogeneidade de covariâncias γ .

Nas situações de grandes diferenças entre os tamanhos amostrais, como $n_1 = 8$, $n_2 = 30$ e $n_1 = 30$, $n_2 = 8$, o TNVM foi, em geral, liberal, mesmo sendo as covariâncias populacionais homogêneas. Quando a população de maior covariância, população 2, estava associada ao menor tamanho amostral ($n_2 = 8$), as taxas de erro tipo I cresceram com o aumento da heterogeneidade entre as covariâncias, atingindo uma diferença de 3,95 pontos percentuais em relação ao nível nominal de 0,10. Em todas essas situações, o TB proposto apresentou resultados opostos no sentido de ser mais conservativo. Em nenhuma situação, para $n_1 \neq n_2$, o novo teste foi considerado liberal.

Na Tabela 4 são apresentadas as taxas de erro obtidas sob H_0 para diferentes graus de heterogeneidade γ e tamanhos amostrais n_1 e n_2 , considerando $p = 5$ e $\alpha = 0,05$. O comportamento dos dois testes foi semelhante ao apresentado para $\alpha = 0,10$, com a diferença que o TNVM se apresentou mais conservativo para tamanhos amostrais iguais menores (8 e 15), associados a menores graus de heterogeneidade γ de covariâncias. A característica conservativa do teste bayesiano para tamanhos amostrais diferentes e a liberal do teste de Nel & Merwe modificado (Krishnamoorthy & Yu, 2004) para $n_1 = 30$ e $n_2 = 8$ se mantiveram.

Na Tabela 5 são apresentadas as taxas de erro tipo I dos testes bayesiano

TABELA 3 Taxas de erro tipo I do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), em função dos tamanhos amostrais n_1 e n_2 e do grau de heterogeneidade (γ) das variâncias, para $p = 5$ variáveis e nível nominal de significância $\alpha = 0,10$, sob H_0 .

γ	$n_1 = 8, n_2 = 8$		$n_1 = 15, n_2 = 15$		$n_1 = 20, n_2 = 20$	
	TB	TNVM	TB	TNVM	TB	TNVM
1	0,0830 ⁺⁺	0,0780 ⁺⁺	0,1090	0,0840	0,1135	0,0890
2	0,0865	0,0845	0,1195 ^{**}	0,1015	0,1175	0,0990
8	0,0990	0,0895	0,1300 ^{**}	0,1045	0,1110	0,0850
16	0,1100	0,1000	0,1165	0,0885	0,1245 ^{**}	0,0955
32	0,0975	0,0890	0,1255 ^{**}	0,0915	0,1180	0,0955
γ	$n_1 = 30, n_2 = 30$		$n_1 = 100, n_2 = 100$		$n_1 = 8, n_2 = 15$	
	TB	TNVM	TB	TNVM	TB	TNVM
1	0,1140	0,1030	0,1070	0,1040	0,0560 ⁺⁺	0,1020
2	0,1055	0,0930	0,0915	0,0905	0,0540 ⁺⁺	0,0925
8	0,1060	0,0965	0,1065	0,0980	0,0575 ⁺⁺	0,0915
16	0,1095	0,0920	0,0935	0,0905	0,0695 ⁺⁺	0,0975
32	0,1095	0,0930	0,1040	0,1010	0,0670 ⁺⁺	0,1000
γ	$n_1 = 15, n_2 = 8$		$n_1 = 8, n_2 = 30$		$n_1 = 30, n_2 = 8$	
	TB	TNVM	TB	TNVM	TB	TNVM
1	0,0515 ⁺⁺	0,1060	0,0190 ⁺⁺	0,1290 ^{**}	0,0185 ⁺⁺	0,1200 ^{**}
2	0,0620 ⁺⁺	0,1120	0,0280 ⁺⁺	0,1105	0,0170 ⁺⁺	0,1270 ^{**}
8	0,0540 ⁺⁺	0,1045	0,0245 ⁺⁺	0,1230 ^{**}	0,0215 ⁺⁺	0,1295 ^{**}
16	0,0605 ⁺⁺	0,1250 ^{**}	0,0220 ⁺⁺	0,1095	0,0165 ⁺⁺	0,1395 ^{**}
32	0,0610 ⁺⁺	0,1175	0,0255 ⁺⁺	0,1100	0,0205 ⁺⁺	0,1365 ^{**}

⁺⁺ Significativamente diferente (valor- $p < 0,01$) e considerado menor do que o nível de significância $\alpha = 0,10$.

^{**} Significativamente diferente (valor- $p < 0,01$) e considerado maior do que o nível de significância $\alpha = 0,10$.

proposto e Nel & Merwe modificado (Krishnamoorthy & Yu, 2004), em função de n_1 , n_2 e γ , sob H_0 e para $p = 5$ e $\alpha = 0,01$. Nessa situação de $\alpha = 0,01$ e tamanhos amostrais iguais, $n_1 = n_2$, os dois testes controlaram as taxas de erro tipo I para todos os graus de heterogeneidade considerados, com apenas uma exceção ($n_1 = n_2 = 15$ e $\gamma = 2$), em que o teste bayesiano se mostrou liberal, porém, com taxa de erro pouco maior do que o nível nominal de 0,01 adotado. Para tamanhos amostrais diferentes, o TB manteve sua característica conservativa

TABELA 4 Taxas de erro tipo I do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), em função dos tamanhos amostrais n_1 e n_2 e do grau de heterogeneidade (γ) das variâncias, para $p = 5$ variáveis e nível nominal de significância $\alpha = 0,05$, sob H_0 .

γ	$n_1 = 8, n_2 = 8$		$n_1 = 15, n_2 = 15$		$n_1 = 20, n_2 = 20$	
	TB	TNVM	TB	TNVM	TB	TNVM
1	0,0345 ⁺⁺	0,0335 ⁺⁺	0,0535	0,0340 ⁺⁺	0,0585	0,0460
2	0,0385	0,0375 ⁺⁺	0,0680 ^{**}	0,0490	0,0595	0,0485
8	0,0460	0,0460	0,0705 ^{**}	0,0460	0,0535	0,0395
16	0,0510	0,0475	0,0590	0,0440	0,0595	0,0495
32	0,0460	0,0410	0,0640	0,0485	0,0620	0,0455
γ	$n_1 = 30, n_2 = 30$		$n_1 = 100, n_2 = 100$		$n_1 = 8, n_2 = 15$	
	TB	TNVM	TB	TNVM	TB	TNVM
1	0,0610	0,0520	0,0575	0,0520	0,0165 ⁺⁺	0,0540
2	0,0570	0,0465	0,0435	0,0415	0,0175 ⁺⁺	0,0440
8	0,0590	0,0475	0,0595	0,0535	0,0210 ⁺⁺	0,0470
16	0,0590	0,0490	0,0505	0,0485	0,0190 ⁺⁺	0,0425
32	0,0535	0,0420	0,0460	0,0430	0,0260 ⁺⁺	0,0470
γ	$n_1 = 15, n_2 = 8$		$n_1 = 8, n_2 = 30$		$n_1 = 30, n_2 = 8$	
	TB	TNVM	TB	TNVM	TB	TNVM
1	0,0170 ⁺⁺	0,0550	0,0035 ⁺⁺	0,0665 ^{**}	0,0050 ⁺⁺	0,0680 ^{**}
2	0,0230 ⁺⁺	0,0520	0,0045 ⁺⁺	0,0555	0,0010 ⁺⁺	0,0660 ^{**}
8	0,0170 ⁺⁺	0,0555	0,0055 ⁺⁺	0,0610	0,0040 ⁺⁺	0,0725 ^{**}
16	0,0205 ⁺⁺	0,0605	0,0035 ⁺⁺	0,0540	0,0025 ⁺⁺	0,0740 ^{**}
32	0,0205 ⁺⁺	0,0635	0,0080 ⁺⁺	0,0470	0,0020 ⁺⁺	0,0755 ^{**}

⁺⁺ Significativamente diferente (valor- $p < 0,01$) e considerado menor do que o nível de significância $\alpha = 0,05$.

^{**} Significativamente diferente (valor- $p < 0,01$) e considerado maior do que o nível de significância $\alpha = 0,05$.

para todos os valores de γ , enquanto o TNVM continuou liberal apenas para dois casos de $n_1 = 30$ e $n_2 = 8$.

O TB apresentou o mesmo comportamento, para diferentes níveis nominais de significância α , embora tenha sido um pouco mais conservativo quando se considerou $p = 5$ variáveis em relação a $p = 2$. Da mesma forma, o TNVM apresentou diferenças em relação ao número de variáveis envolvidas. Para um menor número de variáveis p , há controle das taxas de erro tipo I, mas, com o aumento de

TABELA 5 Taxas de erro tipo I do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), em função dos tamanhos amostrais n_1 e n_2 e do grau de heterogeneidade (γ) das variâncias, para $p = 5$ variáveis e nível nominal de significância $\alpha = 0,01$, sob H_0 .

γ	$n_1 = 8, n_2 = 8$		$n_1 = 15, n_2 = 15$		$n_1 = 20, n_2 = 20$	
	TB	TNVM	TB	TNVM	TB	TNVM
1	0,0065	0,0065	0,0090	0,0055	0,0135	0,0085
2	0,0070	0,0080	0,0180**	0,0135	0,0165	0,0125
8	0,0075	0,0090	0,0140	0,0060	0,0110	0,0075
16	0,0100	0,0100	0,0130	0,0080	0,0105	0,0070
32	0,0100	0,0095	0,0120	0,0070	0,0170	0,0115
γ	$n_1 = 30, n_2 = 30$		$n_1 = 100, n_2 = 100$		$n_1 = 8, n_2 = 15$	
	TB	TNVM	TB	TNVM	TB	TNVM
1	0,0160	0,0135	0,0100	0,0075	0,0000 ⁺⁺	0,0085
2	0,0130	0,0095	0,0085	0,0075	0,0015 ⁺⁺	0,0075
8	0,0135	0,0100	0,0140	0,0115	0,0000 ⁺⁺	0,0070
16	0,0145	0,0090	0,0085	0,0080	0,0015 ⁺⁺	0,0095
32	0,0100	0,0065	0,0115	0,0115	0,0020 ⁺⁺	0,0130
γ	$n_1 = 15, n_2 = 8$		$n_1 = 8, n_2 = 30$		$n_1 = 30, n_2 = 8$	
	TB	TNVM	TB	TNVM	TB	TNVM
1	0,0005 ⁺⁺	0,0070	0,0000 ⁺⁺	0,0135	0,0000 ⁺⁺	0,0135
2	0,0000 ⁺⁺	0,0105	0,0000 ⁺⁺	0,0115	0,0000 ⁺⁺	0,0125
8	0,0015 ⁺⁺	0,0135	0,0005 ⁺⁺	0,0120	0,0000 ⁺⁺	0,0175**
16	0,0020 ⁺⁺	0,0100	0,0000 ⁺⁺	0,0120	0,0000 ⁺⁺	0,0160
32	0,0010 ⁺⁺	0,0150	0,0000 ⁺⁺	0,0090	0,0000 ⁺⁺	0,0180**

⁺⁺ Significativamente diferente (valor- $p < 0,01$) e considerado menor do que o nível de significância $\alpha = 0,01$.

** Significativamente diferente (valor- $p < 0,01$) e considerado maior do que o nível de significância $\alpha = 0,01$.

p (de 2 para 5) e nas situações de tamanhos amostrais bem diferentes e com a maior covariância associada à população de menor tamanho ($n_1 = 30$ e $n_2 = 8$), esse teste se mostrou liberal. Tal característica pode ser uma vantagem do teste bayesiano, por apresentar taxas de erro tipo I menores do que o nível nominal adotado, mesmo quando são considerados tamanhos amostrais bem diferentes, enquanto o TNVM se torna liberal em situações desse tipo. Krishnamoorthy & Yu (2004), ao avaliarem os testes de Nel & Merwe original (1986) e o modificado, observaram

que a versão modificada sempre controlou o erro tipo I, porém, os autores, apesar de terem simulado situações de tamanhos amostrais diferentes, não atribuíram a maior covariância ao menor tamanho amostral. Esta situação foi a que apresentou comportamento liberal, especialmente para maior valor de p , no presente trabalho.

4.2 Poder

Na Figura 3 estão apresentados os valores de poder do TB e do TNVM para $p = 2$, em função do grau de heterogeneidade γ e dos tamanhos amostrais n_1 e n_2 , fixada a diferença entre os vetores de médias em $k = 2$ erros padrões para $\alpha = 0,10$. Para amostras médias (30) ou grandes (100) de tamanhos iguais, o TB e o TNVM apresentaram valores de poder iguais e indistinguíveis (Figuras 3c e 3d). Nesses casos, o grau de heterogeneidade praticamente não teve efeito no poder de ambos os testes já que as curvas de poder foram praticamente paralelas à abscissa. Os valores de poder oscilaram em torno de 0,60 a 0,65, sendo que, com $n_1 = n_2 = 100$ houve valores de poder próximos a 0,65 e, para $n_1 = n_2 = 30$, os valores ficaram próximos de 0,60.

Para o primeiro caso de amostras de tamanhos diferentes, $n_1 = 8$ e $n_2 = 30$ (Figura 3a), em que a menor amostra está associada à população de menor variância generalizada, ou menor covariância, o TB teve poder inferior ao do TNVM para todos os graus de heterogeneidade. A diferença foi praticamente constante em todos os casos, girando em torno de 5 pontos percentuais. Convém salientar que o TB foi conservativo no controle do erro tipo I nessas situações relativas ao tamanho amostral e grande heterogeneidade de covariâncias para $p = 2$. No segundo caso, $n_1 = 30$ e $n_2 = 8$ (Figura 3b), houve uma pequena redução dos valores de poder em relação à situação anterior, o que condiz com o comentário feito por Kim (1992) de que o poder de alguns testes é afetado quando os dados associados às maiores covariâncias estão associados ao menor tamanho amostral.

O desempenho relativo do TB em relação ao TNVM melhorou. Somente para os casos homogêneos ($\gamma = 1$) e de baixa heterogeneidade ($\gamma = 2$), o poder do TB foi menor do que o do TNVM. Nos demais casos, os valores de poder foram iguais. Esse resultado é surpreendente, pois foi a situação em que o TB foi mais conservativo e o TNVM apresentou tamanhos exatos de teste, exceto pelo erro de Monte Carlo.

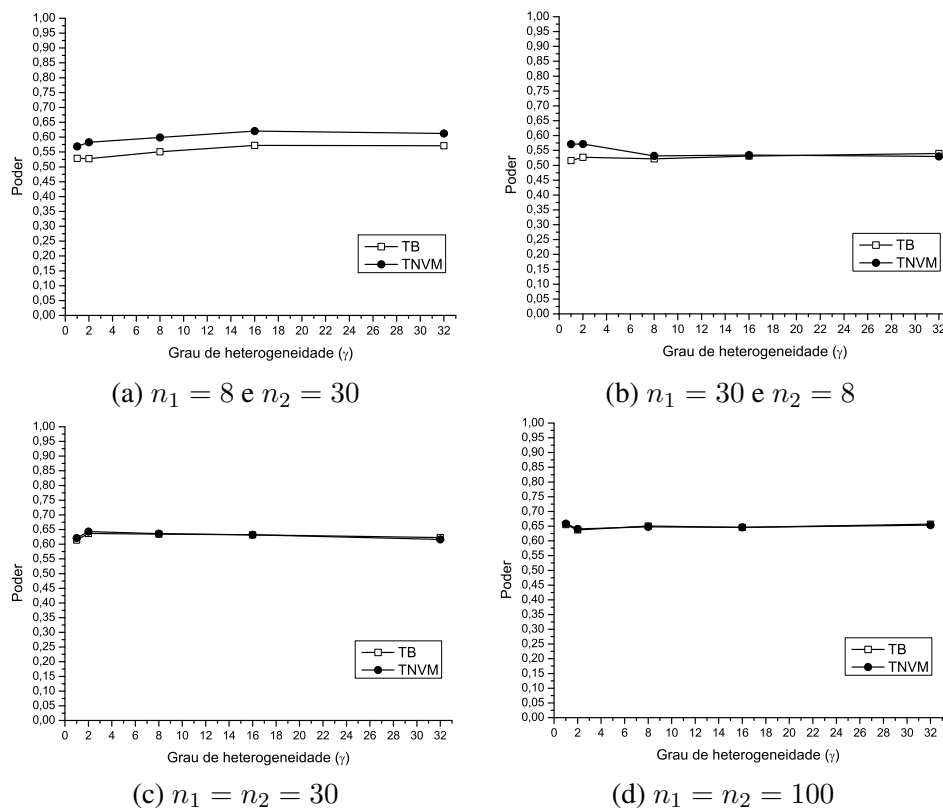


FIGURA 3 Poder do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), para $p = 2$, diferentes graus de heterogeneidade γ , $k = 2$ e $\alpha = 0,10$, sob H_1 .

Na Figura 4 os valores de poder dos testes foram apresentados para as mesmas configurações da Figura 3, exceto pelo fato de o valor de k ser igual a

4. Independentemente dos tamanhos amostrais e do grau de heterogeneidade, os valores de poder de ambos os testes foram praticamente iguais e muito próximos de 1. Para o caso particular de $n_1 = 30$ e $n_2 = 8$, houve uma pequena superioridade do TB em relação ao TNVM para valores maiores de γ ($\gamma \geq 16$), embora essa diferença possa ser atribuída aos erros de Monte Carlo.

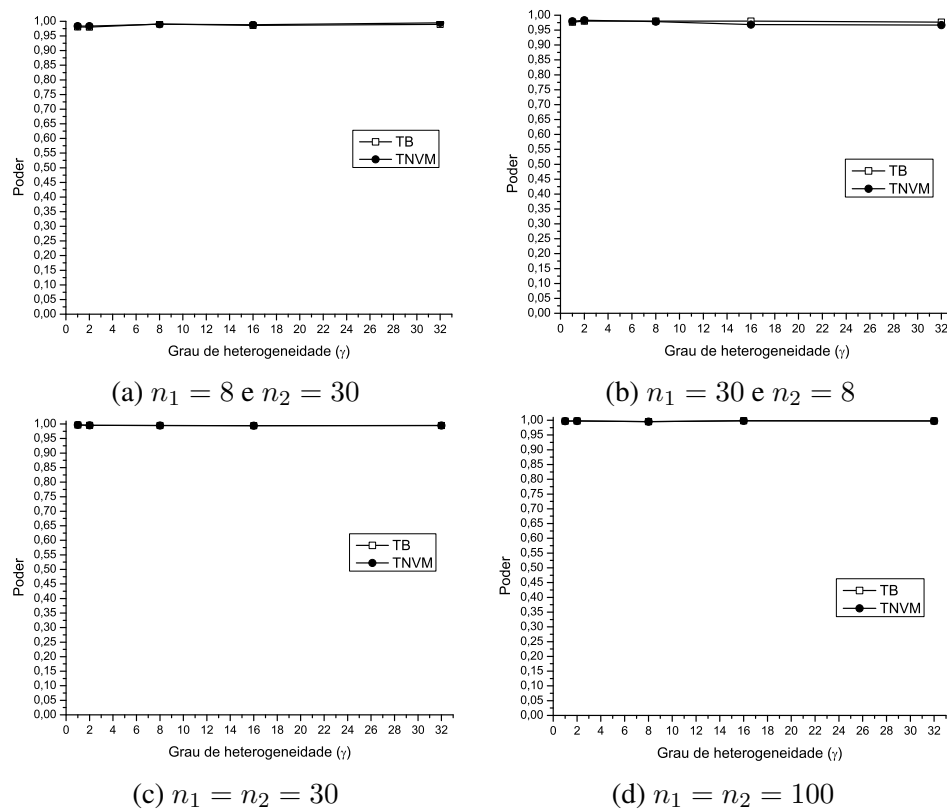


FIGURA 4 Poder do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), para $p = 2$, diferentes graus de heterogeneidade γ , $k = 4$ e $\alpha = 0,10$, sob H_1 .

Para $k = 8$ e $\alpha = 0,10$, considerando diferentes tamanhos de amostras e graus de heterogeneidade das covariâncias, os dois testes tiveram desempenhos iguais e ótimos, com poder de 1, para $p = 2$ e $p = 5$, confirmando o que Subrah-

maniam & Subrahmaniam (1973) obtiveram que, para grandes diferenças entre os vetores de médias, o poder tende a ser 1. Além desse patamar de diferenças entre as médias populacionais ($k > 8$), os valores de poder dos testes também foram sempre iguais a 1, fixados $p = 2$ e $\alpha = 0,10$ e, por essa razão, igualmente não foram apresentados.

Devido ao fato de os resultados de poder para $\alpha = 0,05$ serem semelhantes aos de $\alpha = 0,10$, apenas algumas situações foram retratadas na Figura 5. Foram escolhidas situações homogêneas ($\gamma = 1$) e de grande heterogeneidade ($\gamma = 32$), sendo os valores de poder expressos em função da diferença k entre os vetores de médias das populações e fixado o valor de $p = 2$. Para amostras de mesmo tamanho, médias ($n_1 = n_2 = 30$) ou grandes ($n_1 = n_2 = 100$), os testes tiveram o mesmo desempenho, independentemente da heterogeneidade e do valor de k (Figuras 5c e 5d).

Para amostras de tamanhos diferentes, $n_1 = 8$ e $n_2 = 30$ (Figura 5a), o TB apresentou menor poder do que o TNVM para $k \leq 4$, com diferenças em torno de 5 pontos percentuais para $k = 2$ e de 1 ponto percentual para $k = 4$. Para $k \geq 8$, os valores de poder foram equivalentes nos dois testes e iguais a 1. A heterogeneidade provocou pequenos aumentos nos valores de poder dos testes, o que é, de certa forma, surpreendente, pois, o que se espera é que ocorra o contrário (Figura 5a). Uma possível explicação pode ser dada pelo fato de a maior amostra estar associada à maior covariância populacional. Essa característica demonstra grande robustez de ambos os testes, embora as diferenças no poder tenham sido muito pequenas. Para o caso de $n_1 = 30$ e $n_2 = 8$ (Figura 5b), diferenças nos valores de poder entre os dois testes diminuiram, sendo praticamente inexpressivas em todas as situações considerando-se k e γ .

Na Figura 6 estão apresentados os valores de poder do TB e do TNVM,

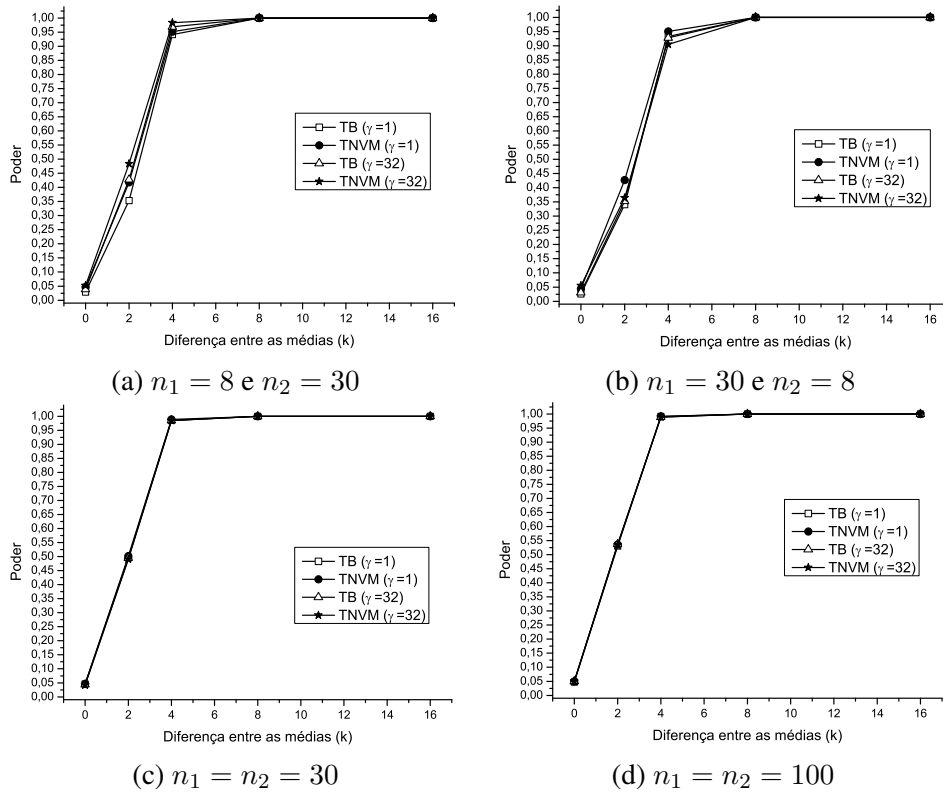
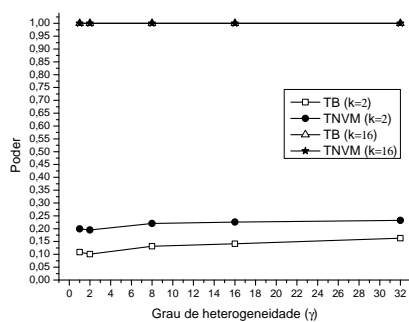


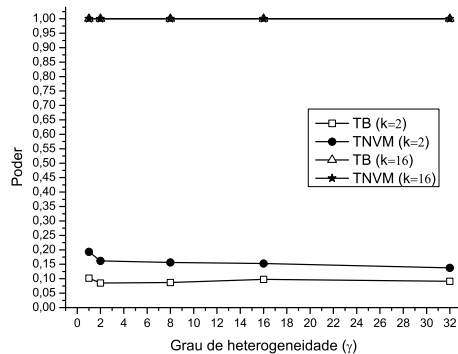
FIGURA 5 Poder do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), para $p = 2$, valores diferentes de k , $\gamma = 1$ e $\gamma = 32$ e $\alpha = 0,05$, sob H_1 .

em função de γ , k e dos tamanhos de amostras para $\alpha = 0,01$ e $p = 2$. O mesmo padrão de desempenhos relativos das situações de $\alpha = 0,10$ e $\alpha = 0,05$ foram observados para $\alpha = 0,01$. O único destaque ocorreu para $n_1 = 30$ e $n_2 = 8$ (Figura 6b), em que as diferenças entre o TB e o TNVM se acentuaram em relação às diferenças observadas para $\alpha = 0,10$ (Figuras 3b e 4b) e $\alpha = 0,05$ (Figura 5b). Para $k \geq 8$, cujos resultados não foram apresentados, os valores de poder foram sempre superiores a 0,97 e praticamente iguais nos dois testes.

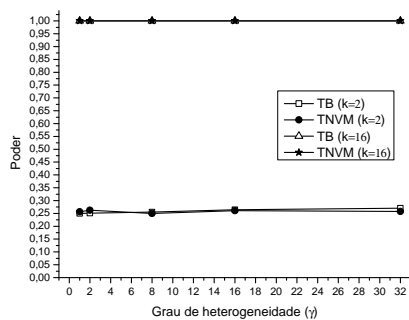
Na Figura 7 estão apresentados os valores de poder do TB e do TNVM,



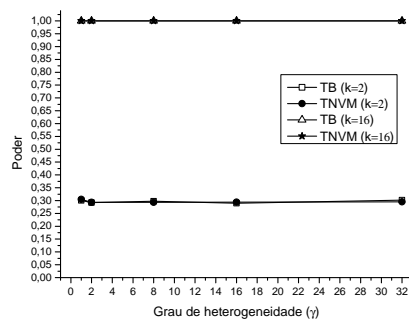
(a) $n_1 = 8$ e $n_2 = 30$



(b) $n_1 = 30$ e $n_2 = 8$



(c) $n_1 = n_2 = 30$



(d) $n_1 = n_2 = 100$

FIGURA 6 Poder do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), para $p = 2$, valores diferentes de γ , $k = 2$ e $k = 16$ e $\alpha = 0,01$, sob H_1 .

em função de γ e tamanhos amostrais para $k = 2$, $p = 5$ e $\alpha = 0,10$. Para amostras iguais e medianas, $n_1 = n_2 = 30$, o TB foi superior ao TNVM e, para $n_1 = n_2 = 100$, as diferenças entre os dois testes foram inexpressivas (Figuras 7c e 7d). Para amostras diferentes, o TB apresentou poder inferior ao do TNVM (Figuras 7a e 7b), sendo que as diferenças foram maiores do que as observadas para $p = 2$ (Figuras 3a e 3b). É importante destacar que, nessas mesmas situações, sob H_0 , o TNVM foi, quase sempre, liberal, enquanto o TB foi conservativo (Tabela 3). O mesmo comportamento do TNVM, nesse caso, foi obtido por Christensen

& Rencher (1997) e Subrahmaniam & Subrahmaniam (1973) para a solução proposta por James (1954), que apresentou grande poder, porém, com taxas de erro tipo I altas, especialmente quando a maior covariância estava associada à população de menor tamanho e para maior p . Cirillo & Ferreira (2004) observaram comportamento semelhante para a solução de Johansen (1980) que, apesar de ter sido poderoso com $\alpha = 0,05$ e $p = 5$, não controlou o erro tipo I, para amostras de tamanhos diferentes ($n_1 = 10$ e $n_2 = 15$). Assim, a vantagem do desempenho do TNVM em relação ao poder não deve ser considerada, pois, em situações reais, o pesquisador não sabe se está sob H_0 ou sob H_1 . Assim, uma rejeição da hipótese nula de igualdade dos vetores de médias poderia ser atribuída ao erro tipo I, se estiver sob H_0 , pois o teste é liberal. A probabilidade desse erro é, de fato, maior do que o nível nominal α utilizado, o que não ocorre com o TB.

As mesmas situações para o poder da Figura 7, exceto pelo fato de que $k = 4$, são apresentadas na Figura 8. Com uma maior diferença entre os vetores de médias ($k = 4$), o TB e o TNVM tiveram valores de poder iguais e próximos de 1 quando as amostras tinham tamanhos amostrais iguais (Figuras 8c e 8d). O grau de heterogeneidade das covariâncias praticamente não teve efeito no desempenho dos testes, com retas paralelas às abscissas. Para amostras de tamanhos diferentes (Figuras 8a e 8b), as diferenças nos desempenhos dos testes diminuíram em relação à situação de $k = 2$ (Figuras 7a e 7b). Como é preconizado pela teoria, o aumento das diferenças entre os vetores de médias provocou um aumento considerável no poder, como pode ser observado na comparação entre as Figuras 7 e 8.

Mantidos os valores de $k = 4$ e $\alpha = 0,10$, porém, considerando-se $p = 2$ (Figura 4), para tamanhos amostrais iguais, quase não há diferenças entre os valores de poder dos dois testes, ficando sempre acima de 0,99, quando são considera-

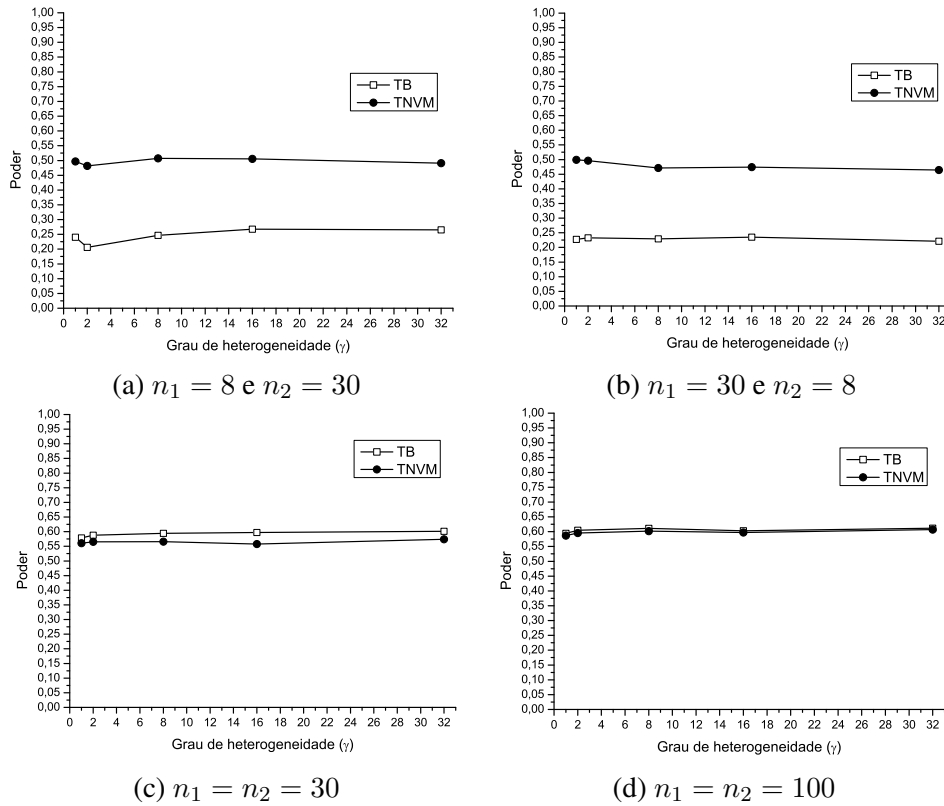


FIGURA 7 Poder do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), para $p = 5$, diferentes graus de heterogeneidade γ , $k = 2$ e $\alpha = 0,10$, sob H_1 .

das 5 variáveis no lugar de 2, o que pode ser constatado comparando-se as Figuras 4c e 4d com as Figuras 8c e 8d. Porém, para tamanhos amostrais diferentes, há uma queda no poder com o aumento de variáveis para os dois testes, sendo o TB mais afetado do que o TNVM (Figuras 4a e 4b e Figuras 8a e 8b). Esse fato confirma o que Subrahmaniam & Subrahmaniam (1973) observaram, ou seja, o poder dos testes diminui com o aumento de p , para uma dada diferença entre os vetores de médias.

Na Figura 9 estão apresentados os valores de poder para os testes baye-

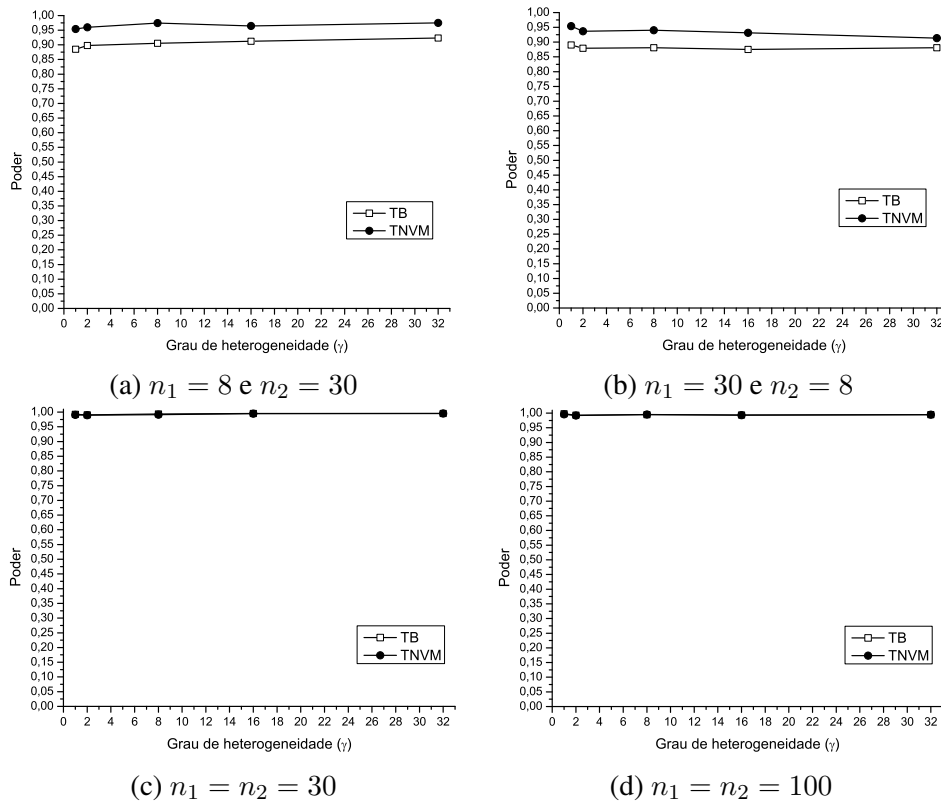


FIGURA 8 Poder do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), para $p = 5$, diferentes graus de heterogeneidade γ , $k = 4$ e $\alpha = 0,10$, sob H_1 .

siano e Nel & Merwe modificado (Krishnamoorthy & Yu, 2004), para $p = 5$ e $\alpha = 0,05$, considerando $\gamma = 1$ e $\gamma = 32$ e diferentes k , incluindo também o caso de $k = 0$, fornecendo informação a respeito do erro tipo I. Comparando-se os valores de poder com os da Figura 5, que corresponde às mesmas configurações, mas com $p = 2$, observou-se que, para tamanhos amostrais diferentes (Figuras 9a e 9b), as diferenças entre os dois testes se acentuaram, com o TB apresentando valores menores do que os do TNVM para $k = 2$ e $k = 4$. Diferentemente do que ocorreu com as situações ilustradas nas Figuras 5a e 5b, em que a heterogeneidade

ocasionou aumento de poder para os testes, no caso da Figura 9, a heterogeneidade não afetou consideravelmente o desempenho do TB e, para o TNVM, ao contrário, no caso de $k = 2$ e $n_1 = 30$ e $n_2 = 8$ (Figura 9b), a homogeneidade de covariâncias ($\gamma = 1$) apresentou o maior poder, o que se espera que ocorra. Porém, esse resultado deve ser interpretado com cuidado pois, para $k = 0$, as taxas de erro tipo I foram maiores do que o nível nominal de significância. Para diferenças entre os vetores de médias maiores do que 8 erros padrões, os dois testes apresentaram valores de poder iguais a 1 e iguais entre si. Para tamanhos amostrais iguais (Figuras 9c e 9d), o TB e o TNVM apresentaram o mesmo comportamento em relação ao poder, com a falta de influência da heterogeneidade, sendo este afetado apenas pelo valor de k . Esse fato demonstra a robustez desses testes para lidar com severas situações do problema de Behrens-Fisher, já que as linhas para $\gamma = 1$ e $\gamma = 32$ são praticamente indistinguíveis.

Na Figura 10 estão apresentados os valores de poder do TB e do TNVM, em função de γ , k e dos tamanhos amostrais para $\alpha = 0,01$ e $p = 5$. Comparando-se os valores de poder com aqueles das mesmas configurações, porém, para $p = 2$ (Figura 6), observa-se queda no desempenho do TB em relação ao TNVM, essencialmente para amostras de tamanhos diferentes (Figuras 10a e 10b), em que o TB apresentou valores de poder baixos. Houve também redução dos valores de poder do TNVM, entretanto, em menor escala. Cabe ressaltar que, em relação às taxas de erro tipo I, o TB foi conservativo para essas situações de tamanhos amostrais diferentes e o TNVM foi liberal em alguns casos (Tabela 5). Para tamanhos amostrais iguais (Figuras 10c e 10d), o desempenho do TB foi levemente superior ao do TNVM. Para $k \geq 8$, cujos resultados não foram apresentados, os valores de poder foram sempre superiores a 0,97 e praticamente iguais nos dois testes.

No geral, o poder dos dois testes foi diminuído com o aumento do número

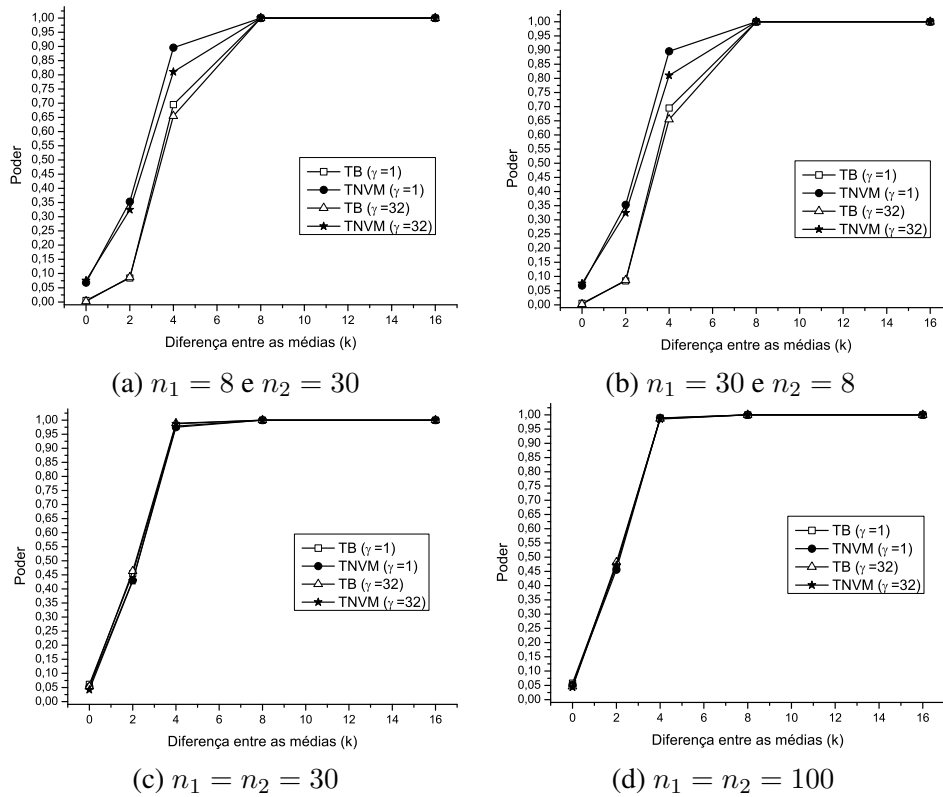


FIGURA 9 Poder do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), para $p = 5$, valores diferentes de k , $\gamma = 1$ e $\gamma = 32$ e $\alpha = 0,05$, sob H_1 .

de variáveis de 2 para 5, tendo sido esse efeito mais pronunciado no TB e para tamanhos amostrais diferentes. A heterogeneidade praticamente não afetou os valores de poder de nenhum dos testes. Considerando-se diferenças entre os vetores de médias maiores ou iguais a 4 erros padrões, os valores de poder foram muito altos para amostras de tamanhos iguais para os dois valores de p (valores acima de 0,97) e para amostras de tamanhos diferentes com $p = 2$ (valores acima de 0,90) e apenas moderados para amostras de tamanhos diferentes com $p = 5$ (valores acima de 0,65). Em todos os casos, o poder foi mais baixo quando a maior

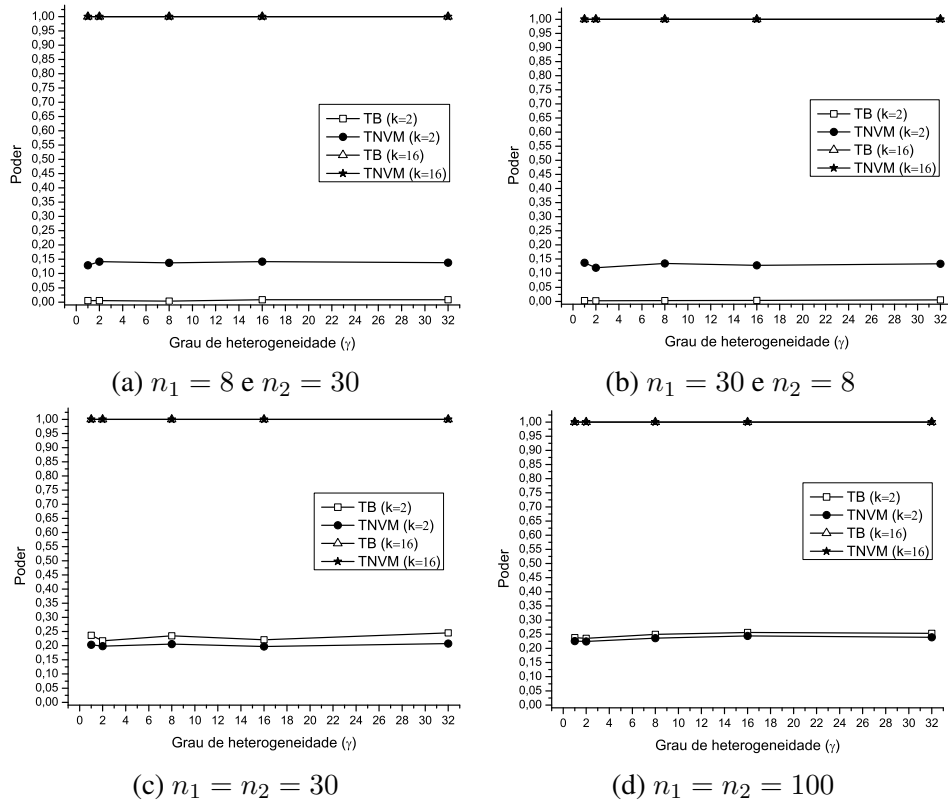


FIGURA 10 Poder do teste bayesiano (TB) e de Nel & Merwe modificado (TNVM), para $p = 5$, valores diferentes de γ , $k = 2$ e $k = 16$ e $\alpha = 0,01$, sob H_1 .

amostra estava associada à menor covariância ($n_1 = 30$ e $n_2 = 8$).

Algumas combinações de tamanhos amostrais foram simuladas, com $(n_1, n_2) = (15,15)$, $(8,15)$ e $(15,8)$, porém, esses resultados não foram inteiramente apresentados por terem comportamento semelhante às situações de n_1 e n_2 dados, respectivamente, por $(30,30)$, $(8,30)$ e $(30,8)$. Porém, como as diferenças entre os tamanhos amostrais das populações 1 e 2 nas situações não apresentadas eram menores, os valores de poder obtidos foram mais altos, indicando que quanto maior for o balanceamento, melhor será o desempenho dos testes e mais parecidos eles

se tornam.

O teste bayesiano apresentou comportamento parecido com o observado por Christensen & Rencher (1997) para a solução de Kim e de Nel & Merwe original, por terem alto poder e serem conservativas. Como o TB apresentou valores de poder altos, foi conservativo para as situações adversas (tamanhos amostrais diferentes e com a maior amostra associada à menor covariância) e controlou o erro tipo I em boa parte das situações, mesmo com alta heterogeneidade, é um teste que deve ser recomendado. Além do mais, em várias situações sob H_0 , o TNVM se mostrou liberal, embora com taxas de erro tipo I não tão expressivamente superiores aos valores nominais de significância. Nessas situações, mas sob H_1 , seu poder foi superior ao do TB. Essa vantagem não é real pois, sob H_0 , não houve controle do erro tipo I. Nessas condições de maiores diferenças entre os tamanhos amostrais, o pesquisador deveria considerar a possibilidade de utilizar o TB por ser conservativo no controle das taxas de erro tipo I e apresentar poder competitivo.

Devido ao maior controle do erro tipo I, na maior parte das situações simuladas, o teste bayesiano proposto é recomendado, principalmente para ser utilizado em situações de balanceamento pois, nessas situações, seu poder foi, no mínimo, igual ao do TNVM. Novas avaliações considerando escolhas diferentes para os hiperparâmetros podem ser realizadas em trabalhos futuros, tendo como perspectiva a melhoria do desempenho do TB nas situações de não balanceamento ($n_1 \neq n_2$).

4.3 Exemplo real

Os dados dos solos de capoeira nova e de pastagem da Amazônia do exemplo descrito na seção 3.4 se encaixam na situação de tamanhos amostrais diferentes ($n_1 = 30$ e $n_2 = 8$), com a maior covariância associada à população de maior tamanho amostral (população 1). Para verificar que os dados do exemplo seguiam as pressuposições necessárias para caracterizar o problema de Behrens-Fisher, ou

seja, que a distribuição dos dados era normal multivariada e as matrizes de covariâncias eram heterogêneas, Ferreira (2008) realizou os testes necessários e confirmou as duas condições.

Para aplicação do teste bayesiano, foram obtidos os vetores de médias amostrais (\bar{Y}_1 e \bar{Y}_2) e as matrizes de somas de quadrados e produtos (V_1 e V_2):

$$\bar{Y}_1 = \begin{bmatrix} 6,1 \\ 2,5 \\ 8,8 \end{bmatrix}, \quad \bar{Y}_2 = \begin{bmatrix} 2,2 \\ 1,4 \\ 3,8 \end{bmatrix},$$

$$V_1 = \begin{bmatrix} 124,40 & 37,68 & 163,21 \\ 37,68 & 23,56 & 61,60 \\ 163,21 & 61,60 & 226,49 \end{bmatrix} \quad \text{e} \quad V_2 = \begin{bmatrix} 6,83 & 3,45 & 10,35 \\ 3,45 & 6,29 & 9,91 \\ 10,35 & 9,91 & 20,53 \end{bmatrix}.$$

O vetor das diferenças entre as médias amostrais (d), e a matriz V_p , obtida por meio da expressão (3.5) são

$$d = \begin{bmatrix} 3,9 \\ 1,1 \\ 5,0 \end{bmatrix} \quad \text{e} \quad V_p = \begin{bmatrix} 89,99 & 27,66 & 118,47 \\ 27,66 & 18,50 & 46,47 \\ 118,47 & 46,47 & 166,21 \end{bmatrix}$$

Os valores dos hiperparâmetros obtidos por meio de (3.4) para se gerarem as amostras *a posteriori* foram $r_1 = 4,99$ e $r_2 = 4,97$. Em seguida, foram geradas as $N = 2000$ amostras, os valores das diferenças δ foram obtidos e os valores da quantidade q_j foram calculados pela expressão (3.6). Ordenaram-se os valores de q_j e obtiveram-se os quantis 0,90, 0,95 e 0,99, que foram iguais a 0,022, 0,027 e 0,048, respectivamente. Foi calculado o valor de q_c , de acordo com a expressão (3.7) e que foi igual a 0,191. Como o valor de q_c supera os valores dos quantis obtidos, conclui-se que os vetores de médias dos solos de capoeira nova e de pastagem da Amazônia diferem, pelo TB, com credibilidade de 0,90, 0,95 e 0,99.

O TNVM também foi aplicado aos dados do exemplo. Foram obtidos os valores de $\nu = 36,954$, $T_c^2 = 94,289$ e valor- $p = 9,907305 \times 10^{-10}$. Assim, os vetores de médias dos solos de capoeira nova e de pastagem da Amazônia também diferem pelo TNVM.

5 CONCLUSÕES

A solução bayesiana foi obtida com sucesso e foi competitiva com a melhor solução frequentista existente. Em geral, o teste proposto foi conservativo para amostras de tamanhos diferentes e liberal em alguns casos de amostras de tamanhos iguais e pequenos. Nas situações em que o teste bayesiano foi liberal, a diferença entre a taxa de erro tipo I empírica e o valor nominal de significância foi praticamente inexpressiva.

O teste bayesiano foi conservativo em casos desbalanceados quando o número de variáveis foi mais alto ($p = 5$), enquanto o teste frequentista de Nel & Merwe modificado (Krishnamoorthy & Yu, 2004) foi liberal.

Como o teste bayesiano possui vantagens competitivas, superando seu principal concorrente em algumas situações, como em casos de tamanhos amostrais diferentes e maior número de variáveis, deve ser recomendada a sua utilização em experimentos reais.

REFERÊNCIAS BIBLIOGRÁFICAS

AGRESTI, A.; MIN, Y. Frequentist performance of bayesian confidence intervals for comparing proportions in 2×2 contingency tables. **Biometrics**, Oxford, v. 61, n. 2, p. 515–523, June 2005.

ASPIN, A. A. An examination and further development of a formula arising in the problem of comparing two means values. **Biometrika**, London, v. 35, n. 1/2, p. 88–96, May 1948.

ASPIN, A. A. Tables for use in comparisons whose accuracy involves two variances, separately estimated. **Biometrika**, London, v. 36, n. 3/4, p. 290–296, Dec. 1949.

BEHRENS, W. U. A contribution to error estimation with few observations. **Landwirtschaftliche Jahrbücher**, Bonn, v. 68, p. 807–837, 1929.

BENNETT, B. M. Note on a solution of generalized Behrens-Fisher problem. **Annals of the Institute of Statistical Mathematics**, Tokyo, v. 2, n. 1, p. 87–90, Dec. 1951.

BOX, G. E. P.; TIAO, G. C. **Bayesian inference in statistical analysis**. New York: J. Wiley Classics, 1992.

BROWN, M. B.; FORSYTHE, A. B. The small sample behavior of some statistics which test the equality of several means. **Technometrics**, Washington, v. 16, n. 1, p. 385–389, 1974.

BRUNNER, E.; DETTE, H.; MUNK, A. Box-type approximations in nonparametric factorial designs. **Journal of American Statistical Association**, Washington, v. 92, n. 1, p. 1494–1502, Dec. 1997.

BRUNNER, E.; MUNZEL, U.; PURI, M. L. The multivariate nonparametric Behrens-Fisher problem. **Journal of Statistical Planning and Inference**, Amsterdam, v. 108, n. 2, p. 37–53, Nov. 2002.

BUCKLEY, J. Simple bayesian inference for qualitative political research. **Political Analysis**, Oxford, v. 12, n. 4, p. 386–399, 2004.

CARMER, S. G.; SWANSON, M. R. An evaluation of ten pairwise multiple comparison procedures by Monte Carlo methods. **Journal American Statistical Association**, Washington, v. 68, n. 341, p. 66–74, Mar. 1973.

CHRISTENSEN, W. F.; RENCHER, A. C. A comparison of type I error rates and power levels for seven solutions to the multivariate Behrens-Fisher problem. **Communications in Statistics Simula**, New York, v. 26, n. 4, p. 1251–1273, Apr 1997.

CIRILLO, M. A.; FERREIRA, D. F. Extensão do teste para normalidade univariado baseado no coeficiente de correlação quantil-quantil para o caso multivariado. **Revista de Matemática e Estatística**, Jaboticabal, v. 21, n. 3, p. 57–75, 2003.

DACHS, J. N. W. **Estatística computacional**: uma introdução em turbo Pascal. Rio de Janeiro: [s.n.], 1988. 236p.

DUONG, Q. P.; SHORROCK, R. W. On Behrens-Fisher solutions. **The Statistician**, v. 45, n. 1, p. 57–63, Jan. 1996.

EHLERS, R. S. **Introdução à inferência bayesiana**. Curitiba: Departamento de Estatística - UFPR, 2007. Disponível em: <<http://leg.ufpr.br/ehlers/bayes>>. Acesso em: agosto de 2008.

FERREIRA, D. F. **Estatística Multivariada**. Lavras: UFLA, 2008. 662p.

FISHER, R. A. The fiducial argument in statistical inference. **Annals of Eugenics**, London, v. 6, n. 3, p. 391–398, 1935.

GELFAND, A. E.; SMITH, A. F. M. Sampling-based approaches to calculating marginal densities. **Journal of the American Statistical Association**, Washington, v. 85, n. 410, p. 398–409, Apr 1990.

GELMAN, A.; CARLIN, J. B.; STERN, H. S.; RUBIN, D. B. **Bayesian Data Analysis**. 2. ed. London: Chapman and Hall/CRC, 2003. 668p.

GENTLE, J. E. **Random number generation and Monte Carlo methods**. 2. ed. New York: Springer, 2003. 381p.

GHOSH, M.; KIM, Y. The Behrens-Fisher problem revisited: a Bayes-frequentist synthesis. **The Canadian Journal of Statistics**, Montreal, v. 29, n. 1, Jan. 2001.

GIRI, N. C. **Multivariate statistical analysis**. 2. ed. New York: Marcel Dekker, 2004. 558p.

GOPALAN, R.; BERRY, D. A. Bayesian multiple comparisons using Dirichlet process priors. **Journal of the American Statistical Association**, New York, v. 93, n. 443, p. 1130–1139, Sept. 1998.

HASTINGS, W. K. Monte Carlo sampling methods using Markov chains and their applications. **Biometrika**, London, v. 57, n. 1, p. 97–109, Apr 1970.

HWANG, H. L.; PAULSON, A. S. **Some methods for the multivariate two-sample location problem**. Oak Ridge: Unpublished manuscript, 1986.

JAMES, G. S. Tests of linear hypotheses in univariate and multivariate analysis when the ratios of the population variances are unknown. **Biometrika**, London, v. 41, n. 1/2, p. 19–43, June 1954.

JEFFREYS, H. Note on the Behrens-Fisher formula. **Annals of Eugenics**, London, v. 10, p. 48–51, 1940.

JEFFREYS, H. **Theory of probability**. 3. ed. London: Oxford University Press, 1961.

JOHANSEN, S. The Welch-James approximation to the distribution of the residual sum of squares in a weighted linear regression. **Biometrika**, London, v. 67, n. 1, p. 85–92, 1980.

JOHNSON, M. E. **Multivariate statistical simulation: a guide to selecting and generating continuous multivariate distributions**. New York: J. Wiley, 1987. 240p.

JOHNSON, R. A.; WEERAHANDI, S. A bayesian solution to the multivariate Behrens-Fisher problem. **Journal of the American Statistical Association**, v. 83, n. 401, p. 145–149, Mar. 1988.

KIM, S. A practical solution to the multivariate Behrens-Fisher problem. **Biometrika**, London, v. 79, n. 1, p. 171–176, Mar. 1992.

KRISHNAMOORTHY, K.; YU, J. Modified Nel and Van der Merwe test for the multivariate Behrens-Fisher problem. **Statistics & Probability Letters**, Amsterdam, v. 66, n. 2, p. 161–169, Aug. 2004.

LEHMANN, E. L. **Nonparametrics: statistical methods based on ranks**. São Francisco: Holden Day, 1975.

LIX, L. M.; KESELMAN, H. J.; HINDS, A. M. Robust tests for the multivariate Behrens-Fisher problem. **Computer Methods and programs in Biomedicine**, Amsterdam, v. 77, n. 2, p. 129–139, Feb. 2005.

MACHADO, A. A.; DEMÉTRIO, C. G. B.; FERREIRA, D. F.; SILVA, J. G. C. Estatística experimental: uma abordagem fundamentada no planejamento e no uso de recursos computacionais. In: REUNIÃO ANUAL DA REGIÃO BRASILEIRA DA SOCIEDADE INTERNACIONAL DE BIOMETRIA, 50.; SIMPÓSIO DE ESTATÍSTICA APLICADA À EXPERIMENTAÇÃO AGRONÔMICA, 11., 2005, Londrina. In: **Anais...** Londrina: SEAGRO, 2005. p. 290.

METROPOLIS, N.; ROSENBLUTH, A. W.; ROSENBLUTH, M. N.; TELLER, A. H.; TELLER, E. Equation of state calculations by fast computing machines. **The Journal Chemical Physics**, New York, v. 21, n. 6, p. 1087–1092, 1953.

MOOD, A. M.; GRAYBILL, F. A.; BOES, D. C. **Introduction to the theory of statistics**. 3. ed. Singapore: McGraw-Hill, 1974. 564p.

MORENO, E.; BERTOLINO, F.; RACUGNO, W. Default bayesian analysis of the Behrens-Fisher problem. **Journal of the Statistical Planning and Inference**, Amsterdam, v. 81, n. 2, p. 323–333, Mar. 1999.

MUNZEL, U. Nonparametric methods for paired samples. **Statistica Neerlandica**, v. 53, n. 3, p. 277–286, Nov. 1999.

NEL, D. G.; MERWE, C. A. van der. A solution to the multivariate Behrens-Fisher problem. **Communication in Statistics: Theory and Methods**, v. 15, n. 1, p. 3719–3735, 1986.

O'HAGAN, A.; LUCE, B. R. **A primer on bayesian statistics in health economics and outcomes research**. London: MEDTAP International, 2003. Disponível em <<http://www.bayesian-initiative.com>>. Acesso em: 20 mar. 2009.

PAULINO, C. D.; TURKMAN, M. A. A.; MURTEIRA, B. **Estatística bayesiana**. Lisboa: Fundação Calouste Gulbekian, 2003. 444p.

PRESS, W. H.; FLANNERY, B. P.; TEUKOLSKY, S. A.; VETTERLING, W. T. **Numerical recipes in Fortran: the art of scientific computing**. Cambridge: Cambridge University, 1992. 994p.

R Development Core Team. **R: A language and environment for statistical computing**. Vienna: R Foundation for Statistical Computing, 2009. Disponível em: <<http://www.r-project.org>>. Acesso em: 15 fev. 2009.

REED, J. F. Solutions to the Behrens-Fisher problem. **Computer Methods and Programs in Biomedicine**, Amsterdam, v. 70, n. 3, p. 259–263, Mar. 2003.

RENCHE, A. C. **Methods of multivariate analysis**. 2. ed. New York: J. Wiley, 2002. 708p.

SATTHELTHWAITE, F. E. An approximate distribution of estimates of variance components. **Biometrics Bulletin**, Washington, v. 2, n. 6, p. 110–114, Dec. 1946.

SCHEFFÉ, H. On solutions of the Behrens-Fisher problem based on t-distribution. **Annals of Mathematical Statistics**, Ann Arbor, v. 10, p. 129–137, Dec. 1943.

SCHEFFÉ, H. Practical solutions of the Behrens-Fisher problem. **Journal of the American Statistical Association**, v. 65, n. 332, p. 1501–1508, Dec. 1970.

SMITH, A. F. M.; GELFAND, A. E. Bayesian statistics without tears. **American Statistician**, Washington, v. 46, n. 2, p. 84–88, May 1992.

SUBRAHMANIAM, K.; SUBRAHMANIAM, K. On the multivariate Behrens-Fisher problem. **Biometrika**, London, v. 60, n. 1, p. 107–111, Apr 1973.

TSUI K.-H.; WHEERAHANDI, S. Generalized p-values in significance testing of hypotheses in the presence of nuisance parameters. **Journal of the American Statistical Association**, Washington, v. 84, n. 406, p. 602–607, June 1989.

WELCH, B. L. The generalisation of Student's problem when several different population variances are involved. **Biometrika**, London, v. 6, n. 1/2, p. 28–35, Jan. 1947.

YAO, Y. An approximate degrees of freedom solution to the Behrens-Fisher problem. **Biometrika**, London, v. 52, n. 1/2, p. 139–147, June 1965.

ANEXOS

	Página
PROGRAMA A: Programa R de simulação utilizado para computar o erro tipo I e o poder dos testes TB e TNVM, sob H_0 e sob H_1 , conforme a situação avaliada	71
PROGRAMA B: Programa R para realizar os testes TB e TNVM para o problema de Behrens-Fisher multivariado	75

PROGRAMA A: Programa R de simulação utilizado para computar o erro tipo I e o poder dos testes TB e TNVM, sob H_0 e sob H_1 , conforme a situação avaliada

```
#Função para gerar amostras da distribuição t multivariada
tpostmult = function(n, Ybi, Vi, nu)
{
  y = rmvnorm(n, sigma=diag(length(Ybi))/sqrt(rchisq(n, nu)/nu));
  y = y %*% chol(Vi) +
      matrix(rep(Ybi, times=n), n, length(Ybi), byrow=T)
  return(y)
}

#Função para simular de amostras normais
simulanorms = function(n, mu, sigma)
{
  return(rmvnorm(n, mu, sigma))
}

#Função para aplicação do teste bayesiano:
#recebe o N para gerar as amostras a posteriori, as amostras
#normais obtidas para as duas populações, os tamanhos
#amostrais e o vetor delta0 a ser testado
TB = function(N, Y1, Y2, n1, n2, delta0)
{
  p = ncol(Y1) #n° variáveis
  xb1 = apply(Y1, 2, mean) #obter média da pop.1
  xb2 = apply(Y2, 2, mean) #obter média da pop.2
  V1 = (n1-1)*var(Y1) #obter matriz SQP1
  V2 = (n2-1)*var(Y2) #obter matriz SQP2
  d = xb1 - xb2 #vetor das diferenças das médias amostrais
  Vp = ((n1-1)*V1 + (n2-1)*V2)/(n1 + n2 - 2) #V pooled
  Vpinv = solve(Vp) #inversa da V pooled
  r1 = (2*p-1)*(p/(2*p-1))^((p/n1)^3) #hiperparâmetro r1
  r2 = (2*p-1)*(p/(2*p-1))^((p/n2)^3) #hiperparâmetro r2
  #obter amostras a posteriori
  amostras.post.1 = tpostmult(N, xb1, V1/((n1 + r1 - 2*p)*n1),
                              n1 + r1 - 2*p),
  amostras.post.2 = tpostmult(N, xb2, V2/((n2 + r2 - 2*p)*n2),
                              n2 + r2 - 2*p)

  delta = amostras.post.1 - amostras.post.2
  #obter os quantis q
  q = (t(delta[1,] - d))%*%Vpinv%*%(delta[1,] - d)
  for(ii in 2:N)
  {
    q.t = (t(delta[ii,] - d))%*%Vpinv%*%(delta[ii,] - d)
    q = rbind(q, q.t)
  }
  q.ord = sort(q)
  #quantidade qc necessária ao teste da hipótese nula
}
```

```

qc = (t(delta0 - d))%%Vpinv%%(delta0 - d)
return(list(q.ord=q.ord, qc=qc))
}

#Função para aplicação do teste Nel & Van der Merwe modificado:
#recebe a amostra da pop. 1 em Y1 e da pop. 2 em Y2
#H0:  $\mu_1 - \mu_2 = \text{delta0}$  (delta0, deve ser passado para a função
#default = 0 (p x 1)
TNVMM = function(Y1, Y2, delta0=matrix(0, ncol(Y1), 1))
{
  n1 = nrow(Y1); n2 = nrow(Y2)
  p = ncol(Y1)
  if (n1 > n2)
  {
    aux1 = Y1
    Y1 = Y2
    Y2 = aux1
    delta0 = -1*delta0
    n1 = nrow(Y1); n2 = nrow(Y2)
  } # evitar trocas de dimensões n1 > n2
  S1=var(Y1); S2=var(Y2)
  Y1b = apply(Y1, 2, mean); Y2b = apply(Y2, 2, mean)
  SeI=solve(S1/n1+S2/n2)
  aux=Y1b - Y2b - delta0
  T2 = t(aux)%%SeI%%aux
  V1 = ((S1/n1)%%SeI); V2 = ((S2/n2)%%SeI)
  v2 = p + p^2
  aux = (sum(diag(V1%%V1)) + (sum(diag(V1)))^2)/(n1-1)
  aux = aux + (sum(diag(V2%%V2)) + (sum(diag(V2)))^2)/(n2-1)
  v2=v2/aux
  v1 = p
  F=(v2+1-p)*T2/(v2*p)
  pr.F=1 - pf(F, v1, v2+1-p)
  return(list(f1=v1, f2=v2+1-p, F = F, pr.F=pr.F, T2=T2, nu=v2))
}

#Função para avaliar teste - erros tipo I e II
#Retorna as taxas de erro obtidas para alpha = 10, 5 e 1%
SMC = function(N, M, mu1, mu2, Sigma1, Sigma2, n1, n2, delta0)
{
  res = matrix(0, 3, 2) #matriz para armazenar os resultados
                          #de 10, 5 e 1%
  dimnames(res) = list(c("10%", "5%", "1%"), c("TB", "TNVMM"))
  for (i in 1:M)
  {
    Y1 = simulanorms(n1, mu1, sigma=Sigma1) #obter pop. 1 normal
    Y2 = simulanorms(n2, mu2, sigma=Sigma2) #obter pop. 2 normal

```

```

#aplicar TB e TNVM
B = TB(N, Y1, Y2, n1, n2, delta0)
T2NWM = TNVMM(Y1, Y2, delta0);
#obtenção da proporção de casos em que H0 foi rejeitada
#para o TB
if (B$qc >= B$q.ord[0.90*M]) res[1,1] = res[1,1] + 1.0/M
if (B$qc >= B$q.ord[0.95*M]) res[2,1] = res[2,1] + 1.0/M
if (B$qc >= B$q.ord[0.99*M]) res[3,1] = res[3,1] + 1.0/M
#para o TNVM
if (T2NWM$pr.F <= 0.10) res[1,2] = res[1,2] + 1.0/M
if (T2NWM$pr.F <= 0.05) res[2,2] = res[2,2] + 1.0/M
if (T2NWM$pr.F <= 0.01) res[3,2] = res[3,2] + 1.0/M
}
return(res)
}

#Simulação dos dados
library(mvtnorm)
N = 2000; M = 2000 #número de amostras e simulações
sig2 = 1 #valor utilizado para obter Sigma1
k = c(0,2,4,8,16) #diferentes erros padrões
gama = c(1,2,8,16,32) #diferentes graus de
#heterogeneidade das covariâncias
rho = 0.5 #correlação entre as variáveis

#Exemplo de obtenção das taxas de erro tipo I para:
#k = 0 (médias iguais), gama = 1 (homogeneidade), p = 2, n1=8,n2=8
p = 2
n1 = 8;n2 = 8
delta0 = vector(mode = "integer", length = p) #hipótese testada
#(delta=vetor de zeros)

#matrizes de variâncias e covariâncias
Sigma1 = sig2*((1 - rho)*diag(p) + rho*matrix(1,p,p))
Sigma2 = gama[1]^(1/p)*Sigma1
#situação de médias iguais
mu1 = vector(mode = "integer", length = p)
mu2 = mu1
#chamada à função que computa a taxa de erro tipo I para TB e TNVM
res1 = SMC(N,M,mu1,mu2,Sigma1,Sigma2,n1,n2,delta0)
res1

#Exemplo de obtenção do poder para:
#k = 2 erros padrões, gama = 1 (homogeneidade), p = 2,
#n1=8, n2=15
p = 2
n1 = 8;n2 = 15
delta0 = vector(mode = "integer", length = p) #hipótese testada
#Sdif é a matriz de covariâncias das diferenças entre as médias

```

```

Sdif = vector(mode = "integer", length = p)
#matrizes de variâncias e covariâncias Sigma1 e Sigma2
Sigma1 = sig2*((1 - rho)*diag(p) + rho*matrix(1,p,p))
Sigma2 = gama[1]^(1/p)*Sigma1
for(i in 1:p)
{
  Sdif[i] = sqrt(Sigma1[i,i]/n1 + Sigma2[i,i]/n2)
}
#situação de vetores de médias diferentes por k=1 erro padrão
mu1 = vector(mode = "integer", length = p)
mu2 = mu1 + k[2]*Sdif
res2 = SMC(N,M,mu1,mu2,Sigma1,Sigma2,n1,n2,delta0)
res2

```

PROGRAMA B: Programa R para realizar os testes TB e TNVM para o problema de Behrens-Fisher multivariado

```

#Função para aplicação do teste bayesiano (TB):
#recebe o N para gerar as amostras a posteriori, as amostras
#normais obtidas para as duas populações, os tamanhos
#amostrais, o vetor delta0 a ser testado e o valor
#da credibilidade nominal a ser usada
AplicaTB = function(N,Y1,Y2,n1,n2,delta0,c.nominal = 0.05)
{
  p = ncol(Y1) #nº variáveis
  yb1 = apply(Y1,2,mean) #obter média da pop.1
  yb2 = apply(Y2,2,mean) #obter média da pop.2
  V1 = (n1-1)*var(Y1) #obter matriz SQP1
  V2 = (n2-1)*var(Y2) #obter matriz SQP2
  d = yb1 - yb2 #vetor das diferenças das médias amostrais
  Vp = ((n1-1)*V1 + (n2-1)*V2)/(n1 + n2 - 2) #V pooled
  Vpinv = solve(Vp) #inversa da V pooled
  r1 = (2*p-1)*(p/(2*p-1))^((p/n1)^3) #hiperparâmetro r1
  r2 = (2*p-1)*(p/(2*p-1))^((p/n2)^3) #hiperparâmetro r2
  #obter amostras a posteriori
  y1 = rmvnorm(N,sigma=diag(length(yb1)))/
      sqrt(rchisq(N,n1 + r1 - 2*p)/(n1 + r1 - 2*p))
  amostras.post.1 = y1 %*% chol(V1/((n1 + r1 - 2*p)*n1)) +
      matrix(rep(yb1,times=N),N,length(yb1),byrow=T)
  y2 = rmvnorm(N,sigma=diag(length(yb2)))/
      sqrt(rchisq(N,n2 + r2 - 2*p)/(n2 + r2 - 2*p))
  amostras.post.2 = y2 %*% chol(V2/((n2 + r2 - 2*p)*n2)) +
      matrix(rep(yb2,times=N),N,length(yb2),byrow=T)
  delta = amostras.post.1 - amostras.post.2
  #obter os quantis q
  q = (t(delta[1,] - d))%*%Vpinv%*(delta[1,] - d)
  for(ii in 2:N)
  {
    q.t = (t(delta[ii,] - d))%*%Vpinv%*(delta[ii,] - d)
    q = rbind(q,q.t)
  }
  q.ord = sort(q)
  quantil = q.ord[(1 - c.nominal)*N]
  #quantidade qc necessária ao teste da hipótese nula
  qc = (t(delta0 - d))%*%Vpinv%*(delta0 - d)
  C = length(q[q >= as.numeric(qc)])/N #valor da
      #credibilidade empírica
  return(list(quantil=quantil,qc=qc,c.nominal=c.nominal,
      c.value=C))
}

#Exemplo de aplicação
#dados de Ferreira, 2008, p. 241 e 254

```

```

#as populações eram solos de pastagem e capoeira nova
#para as variáveis Ca, Mg e SB (cálcio , magnésio e
#saturação de bases)
#havia 13 amostras de pastagem e 30 de capoeira nova
library(mvtnorm)
N = 2000
n1 = 13; n2 = 30
p = 3
delta0 = vector(mode = "integer", length = p) #hipótese testada
                                                #(delta=vetor de zeros)
c.nominal = 0.05 #valor da credibilidade nominal a
                 #ser utilizada para realizar o teste
#dados inseridos por coluna (em que as linhas representam a
                             #amostra e cada coluna representa uma variável)
Y1 = matrix(c(2.3,2.5,1.8,3.4,1.8,3.7,1.4,1.5,2.8,1.4,1.8,1.9,2.8,
              1.7,2.5,2.1,2.5,1.1,1.4,0.7,0.6,2.2,0.8,0.6,1.7,0.8,
              4.1,5.1,4.1,6.1,3.0,5.2,2.2,2.2,5.1,2.3,2.5,3.7,3.7)
            ,n1,p)
Y1
Y2 = matrix(c(3.1,4.1,8.3,5.2,8.5,4.2,10.5,6.9,3.4,2.6,7.6,5.4,
              4.9,10.4,7.1,5.1,7.3,5.7,8.5,7.3,5.2,5.0,7.8,6.2,
              4.4,7.9,7.0,5.0,2.9,6.4,2.1,2.4,2.6,2.3,4.1,2.5,
              4.8,3.1,1.6,1.8,2.1,3.6,2.1,4.8,2.2,2.2,1.7,1.2,
              3.6,2.2,1.9,2.5,2.2,1.8,2.1,3.1,2.3,2.5,1.2,2.6,
              5.3,6.8,11.1,7.7,12.8,6.9,15.5,10.4,5.3,4.5,10.0,
              9.2,7.2,15.5,9.5,7.4,9.2,7.0,12.3,9.6,7.3,7.7,
              10.3,8.2,6.6,11.2,9.6,7.6,4.4,9.1),n2,p)
Y2
AplicaTB(N,Y1,Y2,n1,n2,delta0,c.nominal)
#Se o valor de c.value for alto, há evidência para não rejeitar H0
#Se o valor de c.value for baixo, há evidência para rejeitar H0
#Como o valor de c.value obtido para esse exemplo foi muito baixo,
#a hipótese H0 de igualdade dos vetores de médias dos solos
#deve ser rejeitada

```