

# TESTE PARA IGUALDADE DE MÉDIAS INDEPENDENTES COM ERRO NORMAL ASSIMÉTRICO

Marcos Santos de OLIVEIRA<sup>1</sup>  
Daniel Furtado FERREIRA<sup>2</sup>

- RESUMO: Em experimentos planejados, a normalidade dos resíduos é a suposição usual quando se realizam testes para igualdade de médias independentes. No entanto, esta suposição pode ser não realista, fazendo com que as inferências realizadas a partir destes testes sejam não confiáveis. Neste trabalho, generalizou-se a suposição de normalidade, considerando que os resíduos seguem uma distribuição normal assimétrica. Dentro deste contexto, foi proposto um novo procedimento de teste de hipótese, baseado em formas quadráticas, utilizado para testar a hipótese global de igualdade de médias independentes, em uma abordagem bayesiana. Uma aplicação prática com dados reais foi feita e discutida. O desempenho do teste foi avaliado por meio de simulação de Monte Carlo em diferentes situações experimentais e os resultados comparados com o teste F usual, proveniente da análise de variância de dados normalmente distribuídos. O teste proposto apresentou bons resultados, sendo exato ou conservativo, em todas as configurações simuladas e, ao mesmo tempo, apresentou taxas de poder próximas as do teste F usual.
- PALAVRAS-CHAVE: Teste de hipóteses; distribuição normal assimétrica; inferência bayesiana; simulação Monte Carlo.

## 1 Introdução

Em muitas situações práticas, o problema de interesse é descobrir como uma variável resposta se relaciona com os vários fatores explicativos. Um exemplo típico na área agrônômica é como a produtividade de uma planta, medida pela quantidade de seus frutos, relaciona-se com a qualidade do solo, a quantidade

---

<sup>1</sup>Universidade Federal de São João del-Rei - UFSJ, Departamento de Matemática, Estatística e Ciências da Computação, CEP: 36307-352, São João del-Rei, MG, Brasil. E-mail: *mso@ufsj.edu.br*

<sup>2</sup>Universidade Federal de Lavras - UFLA, Departamento de Ciências Exatas, Caixa Postal 37, CEP: 37200-000, Lavras, MG, Brasil. E-mail: *danieljff@dex.ufla.br*. Bolsista CNPq.

de fertilizante ou mesmo com o volume de irrigação disponível. Nesses casos, os pesquisadores normalmente utilizam a teoria de planejamento de experimentos e, mais especificamente, o teste F em uma análise de variância, para comparar os vários níveis do fator, acompanhada de um teste de comparações múltiplas, como por exemplo, os testes de Tukey, Student-Newman-Keuls, Duncan, entre outros.

O teste F e os usuais testes de comparações múltiplas consideram a suposição de normalidade para os resíduos. No entanto, se esta suposição é violada, estes testes não são válidos. Do ponto de vista prático, o método mais comum adotado para alcançar a normalidade (ou simetria) é a transformação de variáveis. Embora tal método possa dar resultados empíricos razoáveis, esse deve ser evitado se um modelo mais conveniente puder ser encontrado. Azzalini e Capitanio (1999) apresentam algumas razões para evitar este procedimento: a transformação de variáveis não fornece informação útil para entender o mecanismo de geração dos dados, dificulta a interpretação dos parâmetros do modelo e, frequentemente, a transformação para um conjunto de dados pode não ser aplicável para outros conjuntos de dados. Outra alternativa seria a análise não-paramétrica.

Um enfoque alternativo à transformação de variáveis, para modelagem de dados de um ponto de vista paramétrico, consiste em considerar classes paramétricas flexíveis de distribuições que exibam assimetria e curtose diferentes da distribuição normal. A classe das distribuições elípticas (Fang et al., 1990) é provavelmente o exemplo mais conhecido deste enfoque. Essa classe inclui um vasto conjunto de distribuições, como por exemplo,  $t$  de Student, logística, exponencial potência e normal contaminada. Embora modelos elípticos representem a melhor alternativa em relação à distribuição normal, os mesmos não são apropriados em situações onde a distribuição amostral dos dados observados é assimétrica.

Nessas situações, propõe-se como alternativa a utilização de uma família mais geral de distribuições, de forma que se consiga modelar a assimetria dos dados e, além disso, incluir a distribuição normal como um caso particular, evitando assim, a necessidade de transformações. Essa família de distribuições é denominada normal assimétrica.

A distribuição normal assimétrica univariada surgiu independentemente em vários artigos estatísticos, entre os principais trabalhos pode-se destacar o de Roberts (1966) e o de O'Hagan e Leonard (1976). Entretanto, foi Azzalini (1985) que apresentou formalmente essa distribuição, estudou suas propriedades e mostrou que a distribuição tem problemas na estimação do parâmetro que controla a assimetria, pelos métodos de estimação usuais (métodos dos momentos e de máxima verossimilhança). Métodos alternativos de estimação têm sido estudados por Sartori (2006) na abordagem clássica e, na abordagem bayesiana, por Liseo e Loperfido (2006). Generalizações dessas idéias para o caso multivariado têm sido propostas por vários autores, entre eles, Azzalini e Dalla-Valle (1996), Azzalini e Capitanio (1999), Genton et al. (2001) e Sahu et al. (2003). Um excelente panorama de trabalhos na área pode ser visto em Genton (2004).

Nas últimas duas décadas, a distribuição normal assimétrica tem se mostrado benéfica no tratamento de dados assimétricos de diversos problemas aplicados

e teóricos. Por esse motivo, o objetivo deste trabalho foi sugerir o uso desta distribuição para os resíduos do modelo de planejamento e apresentar um novo procedimento de teste de hipótese, utilizado para testar a hipótese nula global de igualdade de  $k$  médias independentes.

Como ilustração, foi feita a análise de um conjunto de dados referente a planta trevo vermelho. A seguir, um estudo de simulação foi realizado para avaliar o desempenho do teste proposto com relação as taxas de erro tipo I e poder, sendo os resultados comparados com o teste F da análise de variância. A metodologia deste trabalho foi implementada no software R através de um programa particular. As rotinas *R2WinBUGS* e *rbugs* do software R são também alternativas viáveis para a implementação.

O restante do artigo está organizado como se segue. Na seção 2, apresenta-se um resumo da distribuição normal assimétrica. Na seção 3 são apresentados os métodos para a obtenção do teste proposto. A aplicação e o estudo de simulação são apresentados, respectivamente, nas seções 4 e 5, seguidas das conclusões.

## 2 Distribuição normal assimétrica

Uma variável aleatória  $Z$  tem uma distribuição normal assimétrica padrão (Azzalini, 1985) se sua função densidade de probabilidade (fdp) é dada por

$$f_Z(z) = 2\phi(z)\Phi(\lambda z), \quad z \in \mathbb{R}, \quad \lambda \in \mathbb{R}, \quad (1)$$

em que  $\phi(\cdot)$  e  $\Phi(\cdot)$  denotam a fdp e a função distribuição acumulada (fda) da normal padrão  $N(0,1)$  e  $\lambda$  é o parâmetro de assimetria. Considera-se a seguinte notação:  $Z \sim NA(\lambda)$ . Note que, se  $\lambda = 0$  então  $f_Z(z) = \phi(z)$ . Se  $\lambda < 0$  a distribuição terá uma assimetria negativa e se  $\lambda > 0$  a assimetria será positiva.

Uma variável mais flexível pode ser construída adicionando parâmetros de locação e escala, tais como  $Y = \mu + \sigma Z$ . Então,  $Y$  terá uma fdp dada por

$$f_Y(y) = \frac{2}{\sigma} \phi\left(\frac{y - \mu}{\sigma}\right) \Phi\left(\lambda \frac{y - \mu}{\sigma}\right), \quad \mu \in \mathbb{R}, \quad \sigma > 0.$$

A notação agora é  $Y \sim NA(\mu, \sigma^2, \lambda)$ , sendo chamada de modelo de três parâmetros. Neste caso, a esperança e a variância de  $Y$  são dadas por

$$E[Y] = \mu + \sqrt{\frac{2}{\pi}} \sigma \delta \quad \text{e} \quad Var[Y] = \sigma^2 \left(1 - \frac{2}{\pi} \delta^2\right).$$

em que  $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$ . Note que,  $\delta$  pode ser utilizado como uma parametrização alternativa com uma interpretação similar de  $\lambda$ , devido a característica de ser limitado ( $|\delta| < 1$ ).

Uma representação estocástica para a variável aleatória normal assimétrica  $Z \sim NA(\lambda)$  (Henze, 1986) é dada por

$$Z = \sqrt{(1 - \delta^2)}V + \delta U, \quad (2)$$

em que  $V$  tem distribuição normal padrão ( $N(0,1)$ ) e  $U$  tem distribuição normal truncada positiva ( $NT(0,1)$ ).

### 3 Modelo de planejamento com erro normal assimétrico

Considere  $k$  níveis diferentes de um único fator (tratamento) e o caso em que há um número igual de observações ( $r$ ) em cada tratamento. O modelo linear adotado (delineamento inteiramente ao acaso, sem perda de generalidade) é

$$Y_{ij} = \mu + t_i + \epsilon_{ij}, \quad \begin{cases} i = 1, \dots, k \\ j = 1, \dots, r \end{cases} \quad (3)$$

em que  $Y_{ij}$  representa a  $j$ -ésima observação tomada sob o tratamento  $i$ ,  $\mu$  é a constante geral,  $t_i$  é o efeito do  $i$ -ésimo nível do fator e  $\epsilon_{ij}$  é o efeito do erro experimental não observável. Neste trabalho, será substituída a usual suposição de normalidade para os erros, pela seguinte suposição mais geral

$$\epsilon_{ij} \sim NA(0, \sigma^2, \lambda),$$

ou seja, será admitido que os erros sejam variáveis aleatórias independentes com distribuição normal assimétrica, com parâmetros de posição, escala e assimetria iguais a 0,  $\sigma^2$  e  $\lambda$ , respectivamente. Supõe-se ainda que a variância dos erros,

$$Var[\epsilon_{ij}] = \sigma^2 \left( 1 - \frac{2}{\pi} \frac{\lambda^2}{1 + \lambda^2} \right),$$

seja constante para todos os níveis do fator. Ao fazer essa suposição não se considerará média zero para os erros, pois neste caso

$$E[\epsilon_{ij}] = \sigma \sqrt{\frac{2}{\pi}} \frac{\lambda}{\sqrt{1 + \lambda^2}} \neq 0.$$

Portanto,

$$E[Y_{ij}] = E[\mu + t_i + \epsilon_{ij}] = \mu + t_i + \sigma \sqrt{\frac{2}{\pi}} \frac{\lambda}{\sqrt{1 + \lambda^2}} = \theta_i + \left( \sigma \sqrt{\frac{2}{\pi}} \frac{\lambda}{\sqrt{1 + \lambda^2}} \right),$$

com  $\theta_i = \mu + t_i$ . No entanto, isso não afetará as comparações múltiplas entre as médias dos níveis do tratamento, devido ao fato do incremento na média da variável resposta ser um fator fixo, isto é, uma constante igual a  $\sigma \sqrt{\frac{2}{\pi}} \frac{\lambda}{\sqrt{1 + \lambda^2}}$ .

Uma forma alternativa de escrever o modelo (3) é considerar a representação estocástica de Henze (1986) dada em (2). Nessa representação, o modelo é apresentado em uma forma hierárquica sendo mais conveniente para a implementação do amostrador de Gibbs (Geman e Geman, 1984). Alocando os

$\theta_i = \mu + t_i$ ,  $i = 1, \dots, k$  no vetor  $\boldsymbol{\theta} = [\theta_1, \dots, \theta_k]^\top$ , pode-se reescrever o modelo (3) na seguinte forma hierárquica

$$Y_{ij} = \mathbf{x}_i^\top \boldsymbol{\theta} + \sigma \left( \frac{\lambda}{\sqrt{1+\lambda^2}} U_{ij} + \frac{1}{\sqrt{1+\lambda^2}} V_{ij} \right),$$

em que  $U_{ij} \sim NT(0,1)$  e  $V_{ij} \sim N(0,1)$  são variáveis aleatórias independentes e  $\mathbf{x}_i^\top$ , de dimensão  $(1 \times k)$ , representa o vetor que possui 0 em todas as posições, exceto na  $i$ -ésima, onde o valor é igual a 1 (por exemplo,  $\mathbf{x}_1^\top = [1, 0, \dots, 0]$ ).

Fazendo  $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$ , tem-se

$$Y_{ij} = \mathbf{x}_i^\top \boldsymbol{\theta} + \sigma \delta U_{ij} + \sigma \sqrt{1 - \delta^2} V_{ij}. \quad (4)$$

Condicionando (4) em  $U_{ij}$ , tem-se

$$\begin{aligned} Y_{ij}|U_{ij} = u_{ij} &\sim N(\mathbf{x}_i^\top \boldsymbol{\theta} + \sigma \delta u_{ij}, \sigma^2(1 - \delta^2)) \\ U_{ij} &\sim NT(0,1). \end{aligned} \quad (5)$$

As variáveis aleatórias  $U_{ij}$ ,  $i = 1, \dots, k$  e  $j = 1, \dots, r$  são variáveis latentes, isto é, não observáveis. Estas são muito úteis no processo de estimação, devido à facilidade de escrever os modelos assimétricos em estruturas condicionais.

A fim de facilitar o trabalho computacional, será considerada a seguinte reparametrização

$$\eta = \sigma \delta \quad \text{e} \quad \tau = \sigma \sqrt{1 - \delta^2}.$$

Com isso, a expressão (5) fica

$$\begin{aligned} Y_{ij}|U_{ij} = u_{ij} &\sim N(\mathbf{x}_i^\top \boldsymbol{\theta} + \eta u_{ij}, \tau^2) \\ U_{ij} &\sim NT(0,1). \end{aligned}$$

A função de verossimilhança aumentada (com a inclusão de variáveis latentes) é então dada por

$$\begin{aligned} L(\boldsymbol{\theta}, \eta, \tau, \mathbf{u}_{ij} | \mathbf{Y}, \mathbf{X}) &= \prod_{i=1}^k \prod_{j=1}^r f_{\mathbf{Y}, \mathbf{U}}(y_{ij}, u_{ij}) = \prod_{i=1}^k \prod_{j=1}^r f_{\mathbf{Y} | \mathbf{U}}(y_{ij} | u_{ij}) f_{\mathbf{U}}(u_{ij}) \\ &= \prod_{i=1}^k \prod_{j=1}^r \left[ \frac{1}{\sqrt{2\pi\tau^2}} \exp \left\{ -\frac{1}{2\tau^2} (y_{ij} - \mathbf{x}_i^\top \boldsymbol{\theta} - \eta u_{ij})^2 \right\} \right. \\ &\quad \left. \times \frac{2}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} u_{ij}^2 \right\} \mathbf{I}_{[0, +\infty)}(u_{ij}) \right] \\ &\propto \frac{1}{\tau^n} \exp \left\{ -\frac{1}{2\tau^2} (\mathbf{Y} - \mathbf{X}\boldsymbol{\theta} - \eta \mathbf{U})^\top (\mathbf{Y} - \mathbf{X}\boldsymbol{\theta} - \eta \mathbf{U}) \right\} \\ &\quad \times \exp \left\{ -\frac{1}{2} \mathbf{U}^\top \mathbf{U} \right\} \prod_{i=1}^k \prod_{j=1}^r \mathbf{I}_{[0, +\infty)}(u_{ij}), \end{aligned} \quad (6)$$

em que  $\mathbf{Y} = [Y_{11}, \dots, Y_{1r}, \dots, Y_{k1}, \dots, Y_{kr}]^\top$ , de dimensão  $n \times 1$ , é o vetor aleatório da variável resposta,  $\mathbf{X}$ , de dimensão  $n \times k$ , é a matriz de planejamento,  $\mathbf{U} = [U_{11}, \dots, U_{1r}, \dots, U_{k1}, \dots, U_{kr}]^\top$ , de dimensão  $n \times 1$ , é o vetor aleatório da variável latente e  $\mathbf{I}_{(\cdot, \cdot)}$  denota a função indicadora. Note-se que, nessa terminologia,  $n$  é igual ao número de níveis do fator multiplicado pelo número de repetições (considerado igual por simplicidade), isto é,  $n = kr$ .

### 3.1 Modelagem bayesiana

Uma parte fundamental da análise bayesiana é a especificação das distribuições a priori para todos os parâmetros desconhecidos do modelo.

Uma priori subjetiva para o parâmetro de assimetria ( $\lambda$ ) deverá descrever a incerteza a respeito desse parâmetro, baseando-se no conhecimento prévio a respeito do problema. Por exemplo, uma densidade a priori com massa concentrada em valores positivos de  $\lambda$  indicaria que a assimetria dos dados deverá ser positiva. Assim, o estabelecimento de uma priori subjetiva depende do particular problema que se está tratando e deverá ser discutida caso a caso.

No entanto, para eliminar a responsabilidade da escolha de uma distribuição a priori subjetiva, pode-se, por exemplo, considerar uma priori não informativa. Neste contexto, Bayes e Branco (2007) realizaram estudos de simulação sobre aspectos inferenciais da distribuição normal assimétrica num enfoque bayesiano considerando duas especificações a priori para o parâmetro de assimetria, a saber:  $\lambda \sim t(0, 1/2; 2)$  e  $\lambda \sim t(0, \pi^2/4; 1/2)$ . Aqui,  $t(a, b; d)$  denota uma distribuição  $t$  de Student centrada em  $a$ , com parâmetro de escala  $b$  e  $d$  graus de liberdade. Esses autores concluíram que a priori

$$\lambda \sim t\left(0, \frac{1}{2}; 2\right), \quad (7)$$

é uma boa alternativa para se fazer inferência sob o parâmetro de assimetria do modelo. Bayes e Branco (2007) obtiveram a priori dada em (7) a partir da seguinte parametrização  $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$ . Como  $\delta \in (-1, 1)$ , a escolha natural seria  $\delta \sim U(-1, 1)$ , a qual induz no espaço paramétrico de  $\lambda$  uma distribuição  $t$  de Student com os hiperparâmetros especificados em (7).

Na construção da priori conjunta, seguindo a sugestão de Bayes e Branco (2007), considerou-se a priori especificada em (7) para o parâmetro de assimetria ( $\lambda$ ) e a usual priori de Jeffreys para parâmetros de posição e escala dada por  $\pi(\mu, \sigma) = \pi(\mu)\pi(\sigma) \propto \frac{1}{\sigma}$ , adaptando a idéia para o caso multivariado em questão.

Considerando  $\Theta_{(k+2 \times 1)} = [\boldsymbol{\theta}, \sigma, \lambda]^\top$  o vetor paramétrico completo e assumindo independência a priori de todos os parâmetros, a especificação a priori a ser considerada é dada por

$$\pi(\Theta) = \pi(\boldsymbol{\theta}, \sigma, \lambda) = \pi(\boldsymbol{\theta})\pi(\sigma)\pi(\lambda) \propto \frac{1}{\sigma}\pi(\lambda), \quad (8)$$

com  $\lambda \sim t(0, \sigma_t^2; \kappa)$ . No estudo de simulação foi considerado a priori induzida pela distribuição uniforme ( $\delta \sim U(-1, 1)$ ), isto é, foram considerados os seguintes valores

para os hiperpâmetros:  $\sigma_t^2 = \frac{1}{2}$  e  $\kappa = 2$ . Ao se substituir a priori atribuída a  $\lambda$  na expressão (8) obtém-se,

$$\pi(\Theta) = \pi(\theta, \sigma, \lambda) \propto \frac{1}{\sigma} \left( 1 + \frac{\lambda^2}{\kappa \sigma_t^2} \right)^{-\frac{(\kappa+1)}{2}}. \quad (9)$$

Agora, a distribuição a priori  $\lambda \sim t(0, \sigma_t^2; \kappa)$  também pode ser reescrita na seguinte forma hierárquica

$$\begin{aligned} \lambda | W = w &\sim N\left(0, \frac{\sigma_t^2}{w}\right) \\ W &\sim \text{Gama}\left(\frac{\kappa}{2}, \frac{\kappa}{2}\right), \end{aligned} \quad (10)$$

com  $W$  sendo considerada outra variável aleatória latente.

Utilizando-se de (10), a priori proposta em (9) pode ser reescrita como

$$\pi(\theta, \sigma, \lambda, w) \propto \frac{1}{\sigma} \exp\left(-\frac{1}{2} \frac{w \lambda^2}{\sigma_t^2}\right) w^{\frac{(\kappa+1)}{2}-1} \exp\left(-\frac{\kappa w}{2}\right). \quad (11)$$

Considerando novamente a reparametrização  $\eta = \sigma \delta$  e  $\tau = \sigma \sqrt{1 - \delta^2}$  e utilizando de métodos padrões de transformações de variáveis, a distribuição a priori em (11) pode ser reescrita como

$$\pi(\theta, \eta, \tau, w) \propto \frac{1}{\tau^2} \exp\left(-\frac{1}{2} \frac{\eta^2 w}{\tau^2 \sigma_t^2}\right) w^{\frac{(\kappa+1)}{2}-1} \exp\left(-\frac{\kappa}{2} w\right). \quad (12)$$

Finalmente, a distribuição a posteriori aumentada é obtida multiplicando a priori aumentada dada em (12) pela função de verossimilhança aumentada dada em (6) sendo dada por

$$\begin{aligned} \pi(\theta, \eta, \tau, w, u_{ij} | \mathbf{Y}, \mathbf{X}) &\propto \pi(\theta, \eta, \tau, w) L(\theta, \eta, \tau, u_{ij} | \mathbf{Y}, \mathbf{X}) \\ &\propto \frac{1}{\tau^2} \exp\left(-\frac{1}{2} \frac{\eta^2 w}{\tau^2 \sigma_t^2}\right) w^{\frac{(\kappa+1)}{2}-1} \exp\left(-\frac{\kappa}{2} w\right) \\ &\quad \times \frac{1}{\tau^n} \exp\left\{-\frac{1}{2\tau^2} (\mathbf{Y} - \mathbf{X}\theta - \eta \mathbf{U})^\top (\mathbf{Y} - \mathbf{X}\theta - \eta \mathbf{U})\right\} \\ &\quad \times \exp\left\{-\frac{1}{2} \mathbf{U}^\top \mathbf{U}\right\} \prod_{i=1}^k \prod_{j=1}^r \mathbf{I}_{[0, +\infty)}(u_{ij}), \end{aligned} \quad (13)$$

resultando em formas conhecidas para todas as distribuições condicionais completas

a posteriori, as quais são representadas pelas seguintes expressões

$$\begin{aligned}
w|\boldsymbol{\theta},\eta,\tau,\mathbf{U},\mathbf{Y},\mathbf{X} &\sim \text{Gama}\left(\frac{\kappa+1}{2},\frac{1}{2}\left(\frac{\eta^2}{\tau^2\sigma_t^2}+\kappa\right)\right), \\
\mathbf{U}|\boldsymbol{\theta},\eta,\tau,w,\mathbf{Y},\mathbf{X} &\sim NT_n\left(\frac{(\mathbf{Y}-\mathbf{X}\boldsymbol{\theta})\eta}{\eta^2+\tau^2},\left(\frac{\tau^2}{\eta^2+\tau^2}\right)\mathbf{I}_n\right), \\
\boldsymbol{\theta}|\eta,\tau,w,\mathbf{U},\mathbf{Y},\mathbf{X} &\sim N_k\left((\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top(\mathbf{Y}-\eta\mathbf{U}),\tau^2(\mathbf{X}^\top\mathbf{X})^{-1}\right), \\
\eta|\boldsymbol{\theta},\tau,w,\mathbf{U},\mathbf{Y},\mathbf{X} &\sim N\left(\frac{(\mathbf{Y}-\mathbf{X}\boldsymbol{\theta})^\top\mathbf{U}}{\frac{w}{\sigma_t^2}+\mathbf{U}^\top\mathbf{U}},\frac{\tau^2}{\frac{w}{\sigma_t^2}+\mathbf{U}^\top\mathbf{U}}\right), \\
\frac{1}{\tau^2}\boldsymbol{\theta},\eta,w,\mathbf{U},\mathbf{Y},\mathbf{X} &\sim \text{Gama}(a,b), \text{ com} \\
a = \frac{n+1}{2} \quad \text{e} \quad b &= \frac{1}{2}\left((\mathbf{Y}-\mathbf{X}\boldsymbol{\theta}-\eta\mathbf{U})^\top(\mathbf{Y}-\mathbf{X}\boldsymbol{\theta}-\eta\mathbf{U})+\frac{w\eta^2}{\sigma_t^2}\right).
\end{aligned} \tag{14}$$

A distribuição condicional completa de  $\mathbf{U}$  dada em (14) é uma normal truncada multivariada à esquerda de zero, de dimensão  $n$ , isto é, cada componente do vetor  $(U_{ij})$  é uma normal truncada à esquerda de zero, para todo  $i = 1, \dots, k$  e  $j = 1, \dots, r$ , com  $\mathbf{I}_n$  representando a matriz identidade de ordem  $n \times n$ . A demonstração da obtenção das distribuições condicionais completas se encontram no apêndice.

Sob a hipótese nula,  $H_0 : \theta_1 = \theta_2 = \dots = \theta_k$ , o vetor paramétrico  $\boldsymbol{\theta} = [\theta_1, \dots, \theta_k]^\top$  se reduz somente a um parâmetro  $\theta$ . Com isso, o modelo reduzido pode ser escrito na seguinte forma matricial

$$\mathbf{Y} = \mathbf{1}\theta + \boldsymbol{\epsilon},$$

com  $\mathbf{Y}$ , de dimensão  $n \times 1$ , sendo o vetor aleatório da variável resposta,  $\boldsymbol{\epsilon}$ , de dimensão  $n \times 1$ , sendo o vetor de erros e  $\mathbf{1}$ , de dimensão  $n \times 1$ , representando um vetor composto de elementos iguais a 1. Com isso, a função de verossimilhança aumentada dada em (6), a distribuição a priori aumentada dada em (12) e a distribuição a posteriori aumentada dada em (13) serão as mesmas para o modelo reduzido, porém, com as devidas adaptações, isto é,  $\theta$  no lugar de  $\boldsymbol{\theta}$  e  $\mathbf{1}$  no lugar de  $\mathbf{X}$ , resultando nas seguintes condicionais completas



$$\begin{aligned}
w|\theta,\eta,\tau,\mathbf{U},\mathbf{Y},\mathbf{1} &\sim \text{Gama}\left(\frac{\kappa+1}{2}, \frac{1}{2}\left(\frac{\eta^2}{\tau^2\sigma_t^2} + \kappa\right)\right), \\
\mathbf{U}|\theta,\eta,\tau,w,\mathbf{Y},\mathbf{1} &\sim NT_n\left(\frac{(\mathbf{Y}-\mathbf{1}\theta)\eta}{\eta^2+\tau^2}, \left(\frac{\tau^2}{\eta^2+\tau^2}\right)\mathbf{I}_n\right), \\
\theta|\eta,\tau,w,\mathbf{U},\mathbf{Y},\mathbf{1} &\sim N\left((\mathbf{1}^\top\mathbf{1})^{-1}\mathbf{1}^\top(\mathbf{Y}-\eta\mathbf{U}), \tau^2(\mathbf{1}^\top\mathbf{1})^{-1}\right), \\
\eta|\theta,\tau,w,\mathbf{U},\mathbf{Y},\mathbf{1} &\sim N\left(\frac{(\mathbf{Y}-\mathbf{1}\theta)^\top\mathbf{U}}{\frac{w}{\sigma_t^2} + \mathbf{U}^\top\mathbf{U}}, \frac{\tau^2}{\frac{w}{\sigma_t^2} + \mathbf{U}^\top\mathbf{U}}\right), \\
\frac{1}{\tau^2}\Big|\theta,\eta,w,\mathbf{U},\mathbf{Y},\mathbf{1} &\sim \text{Gama}(a,b), \text{ com} \\
a = \frac{n+1}{2} \text{ e } b &= \frac{1}{2}\left((\mathbf{Y}-\mathbf{1}\theta-\eta\mathbf{U})^\top(\mathbf{Y}-\mathbf{1}\theta-\eta\mathbf{U}) + \frac{w\eta^2}{\sigma_t^2}\right).
\end{aligned} \tag{15}$$

Note-se que, sob  $H_0$ , a distribuição condicional completa de  $\theta$  dada em (15) se torna agora uma distribuição normal univariada.

### 3.2 Teste para hipótese nula global - TG

A formalização deste teste foi inspirada na estatística multivariada de razão de verossimilhanças para vetores de médias normais, que é uma forma quadrática (Johnson e Weerahandi, 1988; Ferreira, 2008).

Considere a seguinte hipótese nula global, de igualdade de  $k$  médias independentes,

$$H_0 : \theta_1 = \theta_2 = \dots = \theta_k.$$

A função do teste baseado na estatística multivariada para testar  $H_0$  é dada por

$$q_c = (\boldsymbol{\theta}_p - \boldsymbol{\theta}_0)^\top \boldsymbol{\Sigma}_p^{-1} (\boldsymbol{\theta}_p - \boldsymbol{\theta}_0), \tag{16}$$

em que  $\boldsymbol{\theta}_p$  é o vetor de médias da distribuição a posteriori de  $\boldsymbol{\theta}$  sob  $H_1$  dada por

$$\boldsymbol{\theta}_p = \begin{bmatrix} E[\theta_1|\mathbf{Y}] \\ E[\theta_2|\mathbf{Y}] \\ \vdots \\ E[\theta_k|\mathbf{Y}] \end{bmatrix}_{k \times 1} = E[\boldsymbol{\theta}|\mathbf{Y}], \tag{17}$$

$\theta_0$  é o vetor constituído de  $k$  repetições da média a posteriori de  $\theta$  sob  $H_0$ , dado por

$$\theta_0 = \begin{bmatrix} E[\theta|\mathbf{Y}] \\ E[\theta|\mathbf{Y}] \\ \vdots \\ E[\theta|\mathbf{Y}] \end{bmatrix}_{k \times 1} = \mathbf{1}E[\theta|\mathbf{Y}], \quad (18)$$

e, por fim,  $\Sigma_p^{-1}$  é a inversa da matriz de covariância da distribuição a posteriori de  $\theta$  sob  $H_1$ , sendo dada por

$$\begin{aligned} \Sigma_p^{-1} &= \begin{bmatrix} \text{Var}[\theta_1|\mathbf{Y}] & 0 & \cdots & 0 \\ 0 & \text{Var}[\theta_2|\mathbf{Y}] & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \text{Var}[\theta_k|\mathbf{Y}] \end{bmatrix}^{-1} \\ &= [\text{diag}(\text{Var}[\theta_i|\mathbf{Y}])]^{-1}, \end{aligned}$$

com  $\text{Var}[\theta_i|\mathbf{Y}]$  representando a variância a posteriori de  $\theta_i$ ,  $i = 1, \dots, k$ , sob  $H_1$  e  $\text{diag}(\cdot)$  representando uma matriz diagonal.

Como a distribuição da função do teste dada em (16) é desconhecida, foi usado o seguinte procedimento para obtenção da distribuição nula. Inicialmente foi gerado uma única cadeia de tamanho  $\ell = 51000$  da distribuição a posteriori dos parâmetros  $\theta$ ,  $\sigma$  e  $\lambda$  sob  $H_0$ , constituindo uma matriz de dimensão  $3 \times 51000$ . Como são conhecidas somente as condicionais completas a posteriori, foi utilizado simulação Monte Carlo via cadeias de Markov para obtenção dessa cadeia. Foram considerados os valores de *burn-in* e *jump* iguais a 1000 e 50, respectivamente, definidos após simulações preliminares, através do critério de Raftery e Lewis (1992) por meio do pacote CODA, que está implementado no software R. Com isso, foi obtido um valor de  $N$  igual a 1000, isto é, foi obtida uma amostra de tamanho 1000 da distribuição de cada um dos 3 parâmetros, constituindo uma nova matriz de dimensão  $3 \times N$ . Os valores gerados  $\theta_j, \sigma_j, \lambda_j$ ,  $j = 1, \dots, N$ , são realizações das distribuições a posterioris referentes a  $j$ -ésima unidade da distribuição de equilíbrio a posteriori.

A partir da matriz de dimensão  $3 \times N$  (constituída dos  $N$  valores gerados para  $\theta$ ,  $\sigma$  e  $\lambda$  sob  $H_0$ ), foi construída uma nova matriz de dimensão  $((k+2) \times N)$ , de tal forma que as  $k$  primeiras linhas desta matriz contivessem valores de  $\theta$  (emulando uma situação de  $k$  diferentes populações) e as últimas duas linhas contivessem os valores gerados para os parâmetros  $\sigma$  e  $\lambda$ . O procedimento para a obtenção dessa matriz é descrito como segue: a 1ª linha foi constituída dos mesmos valores gerados para  $\theta$  sob  $H_0$ ; a 2ª até a  $k$ -ésima linha foram preenchidas a partir de permutações dos valores gerados para  $\theta$ ; por fim, as últimas duas linhas ( $(k+1)$ ª e  $(k+2)$ ª) foram preenchidas com os mesmos valores gerados para os parâmetros  $\sigma$  e  $\lambda$  sob  $H_0$ .

Após a obtenção da matriz de dimensão  $((k + 2) \times N)$ , para cada coluna, isto é, para cada unidade vetorial  $j$  da distribuição de equilíbrio a posteriori sob  $H_0$  simulada, foi obtida a forma quadrática

$$q_j = (\boldsymbol{\theta}_j - \boldsymbol{\theta}_0)^\top \boldsymbol{\Sigma}_0^{-1} (\boldsymbol{\theta}_j - \boldsymbol{\theta}_0), \quad (19)$$

em que  $\boldsymbol{\theta}_j^\top = [\theta_{1j}, \theta_{2j}, \dots, \theta_{kj}]$  para  $j = 1, \dots, N$ ,  $\boldsymbol{\theta}_0$  é como dado em (18) e  $\boldsymbol{\Sigma}_0^{-1}$ , a inversa da matriz de covariâncias da distribuição a posteriori de  $\theta$  sob  $H_0$ , é dada por

$$\boldsymbol{\Sigma}_0^{-1} = \begin{bmatrix} \text{Var}[\theta|\mathbf{Y}] & 0 & \dots & 0 \\ 0 & \text{Var}[\theta|\mathbf{Y}] & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \text{Var}[\theta|\mathbf{Y}] \end{bmatrix}^{-1} = [\text{diag}(\text{Var}[\theta|\mathbf{Y}])]^{-1}.$$

em que  $\text{Var}[\theta|\mathbf{Y}]$  é a variância a posteriori de  $\theta$  sob  $H_0$ .

Os  $N$  valores de  $q_j$  obtidos formam a distribuição nula da função dos parâmetros dada por (16). Ordenando esses valores, pode-se obter o quantil superior  $100\alpha\%$  ( $q_\alpha$ ) para diferentes valores de  $\alpha$ , como, por exemplo, para  $\alpha = 10\%$ ,  $\alpha = 5\%$  ou  $\alpha = 1\%$ . A decisão contrária ou não à hipótese  $H_0$  é tomada de acordo com o seguinte critério: se o valor da função  $q_c$  obtida em (16) for maior que  $q_\alpha$  obtida em (19), rejeita-se  $H_0$ .

## 4 Aplicação

Steel e Torrie (1980) apresentaram um experimento no qual foi medida a quantidade de nitrogênio, em miligramas, contida em plantas do tipo trevo vermelho (*Trifolium pratense*) inoculadas com culturas de bactérias (fixadoras de nitrogênio) dos tipos *Rhizobium trifolii* e um composto com cinco culturas de *Rhizobium meliloti*. Cada uma das cinco culturas de *Rhizobium trifolii*, que é própria da planta trevo vermelho, foi testada individualmente com cinco culturas de *Rhizobium meliloti*, típica da alfafa. Adicionalmente, um composto de culturas de trevo vermelho foi também testado com um composto de culturas da alfafa, totalizando seis níveis de tratamentos. O experimento foi conduzido em uma casa de vegetação utilizando um delineamento inteiramente casualizado com cinco repetições por tratamento, sendo cada uma constituída de um pote. O objetivo foi comparar os níveis de nitrogênio para os diferentes tratamentos. Os dados dos teores de nitrogênio, em miligramas, são apresentados na Tabela 1.

Na Tabela 2 é apresentada a análise de variância para os dados do teor de nitrogênio.

Steel e Torrie (1980) analisaram esses dados por meio de metodologias frequentistas, isto é, assumiram distribuições normais independentes para os resíduos do modelo. A proposta deste trabalho é analisar novamente este conjunto de dados, porém, assumindo a distribuição normal assimétrica para os resíduos

Tabela 1 - Teor de nitrogênio, em mg, de plantas trevo vermelho inoculadas com combinações de culturas de bactérias do tipo *Rhizobium trifolii* e *Rhizobium meliloti*

Repetição	Tratamentos					
	1 3DOK13	2 3DOK4	3 Composto	4 3DOK7	5 3DOK5	6 3DOK1
1	14,3	17,0	17,3	20,7	17,7	19,4
2	14,4	19,4	19,4	21,0	24,8	32,6
3	11,8	9,1	19,1	20,5	27,9	27,0
4	11,6	11,9	16,9	18,8	25,2	32,1
5	14,2	15,8	20,8	18,6	24,3	33,0
Média	13,26	14,64	18,70	19,92	23,98	28,82
D. Padrão	1,43	4,12	1,60	1,13	3,78	5,80

Tabela 2 - Análise de variância para os dados de teor de nitrogênio

Fonte de variação	GL	SQ	QMR	F
Tratamentos	5	847,05	169,41	14,37
Resíduo	24	282,93	11,79	
Total	29	1.129,98		

do modelo de planejamento adotado. Todos os cálculos computacionais foram realizados no software livre R.

Os valores da distribuição F tabelado para 5 e 24 graus de liberdade são 2,62 e 3,90 para 5% e 1%, respectivamente. Como o valor de F calculado a partir dos dados excede o valor de F tabelado a 1%, pode-se concluir que o experimento fornece evidências significativas de diferenças entre as médias dos tratamentos.

Na sequência, apresentam-se os resultados obtidos aplicando-se a metodologia proposta neste trabalho.

Inicialmente, realizaram-se alguns testes de normalidade aos resíduos do modelo. Os resultados são apresentados na Tabela 3.

Tabela 3 - Testes de normalidade para os resíduos do modelo

Teste	Estatísticas		Valor- <i>p</i>	
Shapiro-Wilk	W	0,910027	Pr < W	0,0149
Kolmogorov-Smirnov	D	0,174133	Pr > D	0,0205
Cramer-von Mises	W-Sq	0,155870	Pr > W-Sq	0,0198
Anderson-Darling	A-Sq	0,908188	Pr > A-Sq	0,0194

A partir dos valores- $p$  apresentados na Tabela 3, conclui-se pela rejeição da hipótese nula, isto é, há indícios significativos ( $p < 0,01$ ) de que a distribuição dos resíduos não é uma normal. Na Figura 1 ilustra-se o histograma dos resíduos a partir do ajuste da distribuição normal.

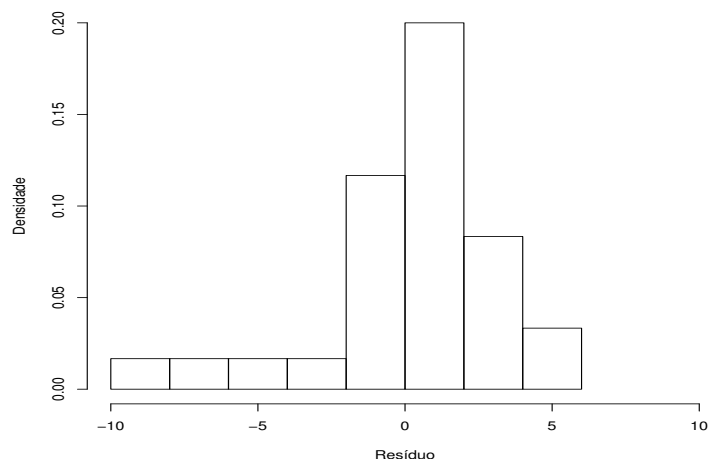


Figura 1 - Histograma dos resíduos do ajuste do modelo de planejamento inteiramente casualizado com distribuição normal para o experimento da planta trevo vermelho.

Constata-se, a partir da Figura 1 que a distribuição dos resíduos não é simétrica, mas sim, assimétrica à esquerda. Portanto, justifica-se considerar uma distribuição mais flexível para os resíduos. Em particular, será assumido que os resíduos possuem uma distribuição normal assimétrica.

Após realizar o procedimento descrito na Seção 3.2, a fim de obter a distribuição nula da função dos parâmetros baseado na estatística multivariada dada por (19), obteve-se

$$q_{0,05} = 11,09 \quad \text{e} \quad q_{0,01} = 14,98. \quad (20)$$

O valor da função do teste, dada em (16), encontrado na posteriori sob  $H_1$  foi

$$q_c = 34,34. \quad (21)$$

Como  $q_c = 34,34 > q_{0,05} = 11,09$  (e também maior que  $q_{0,01} = 14,98$ ), conclui-se pela rejeição da hipótese nula, confirmando o mesmo resultado do teste F.

## 5 Estudo de simulação

Um estudo de simulação Monte Carlo foi realizado para estudar o desempenho das taxas de erro tipo I e poder do teste TG. Os resultados foram comparados com o do teste F da análise de variância.

As simulações foram divididas em duas etapas. Na primeira foram simuladas situações sob  $H_0$  completa (todas as médias dos tratamentos idênticas), a fim de se avaliar a taxa de erro tipo I e, na segunda, foram simuladas situações sob  $H_0$  parcial (pelo menos uma média diferente das demais), para avaliar o poder do teste.

Os passos para a simulação de Monte Carlo são descritos a seguir. Como têm-se disponíveis somente as distribuições condicionais completas dos parâmetros do modelo, foi utilizada a técnica do amostrador de Gibbs para obtenção das cadeias de cada parâmetro. Com isso, o primeiro passo antes de dar início ao estudo de simulação, foi definir valores coerentes para as quantidades *burn-in* e *jump*, as quais foram fixadas para todo o estudo de simulação. A partir de simulações preliminares, definiu-se um *burn-in* de 1000 e um *jump* de 50, por meio do critério de Raftery e Lewis (1992). Com isso, para cada configuração adotada no estudo de simulação foram geradas duas longas realizações da cadeia, uma sob o modelo restrito e a outra sob o modelo irrestrito, sendo o comprimento ( $\ell$ ) de cada uma fixada em 51000, a fim de sempre ter uma amostra efetiva ( $N$ ) de tamanho 1000.

Foram considerados então,  $M = 1000$  simulações Monte Carlo de  $k$  populações não estruturadas e qualitativas. Em cada simulação geraram-se amostras de tamanho  $r$  de cada uma das  $k$  populações (níveis do fator), sendo considerados três modelos probabilísticos para os resíduos do modelo: a normal padrão ( $\lambda = 0$ ), a normal assimétrica padrão com parâmetro de assimetria  $\lambda = 1$  e a normal assimétrica padrão com parâmetro de assimetria  $\lambda = 5$ .

Para as simulações sobre  $H_0$  completa foram consideradas  $k$  populações com médias idênticas e fixadas com valor 100, sem perda de generalidade. Foram consideradas algumas configurações envolvendo o número de níveis do fator ( $k$ ) e o número de repetições ( $r$ ). Mais especificamente, foram utilizados valores de  $k$  iguais a 5, 10 e 20 e valores de  $r$  iguais a 4 e 10. Em todas as configurações e nos  $M$  experimentos gerados de cada uma delas, os testes TG e F foram aplicados considerando os níveis nominais de credibilidades  $\alpha$  iguais a 1%, 5% e 10%, sendo computadas as taxas de erro tipo I.

Para as simulações sobre  $H_0$  parcial, foram considerados dois grupos iguais internamente, mas diferentes entre si. Nesse caso, as diferenças entre as médias dos tratamentos dos dois grupos considerados ( $\delta$ ) foram iguais a 2, 4 e 8. O valor de  $\delta$  representa o número de erros padrões da diferença entre duas médias consecutivas (levando-se em conta que se está assumindo a distribuição normal assimétrica para os resíduos). Adotaram-se agora, somente duas distribuições para os resíduos do modelo: a normal padrão ( $\lambda = 0$ ) e a normal assimétrica padrão com parâmetro de assimetria  $\lambda = 5$ . Novamente, foram utilizados valores de  $r$  iguais a 4 e 10; no entanto, para o número de tratamento  $k$  foram adotados somente 5 e 20. Levando-se em conta que sob  $H_0$  parcial o interesse se concentra no poder dos testes, os mesmos

foram computados para cada configuração considerando os mesmos níveis nominais de credibilidades, isto é,  $\alpha = 1\%$ ,  $5\%$  e  $10\%$ .

Em todas as simulações, os hiperparâmetros da distribuição a priori foram fixados em  $\sigma_t^2 = \frac{1}{2}$  e  $\kappa = 2$ , seguindo a sugestão de Bayes e Branco (2007).

Considerando que as taxas de erro tipo I serão estimadas via simulação Monte Carlo, as mesmas não estão livres de erros. Com isso, seguindo a mesma idéia realizada por Biase e Ferreira (2009), aplicou-se um teste binomial exato considerando um nível nominal de significância de  $1\%$  para as hipóteses

$$\begin{array}{lll} H_0 : \alpha = 10\% & H_0 : \alpha = 5\% & H_0 : \alpha = 1\% \\ H_1 : \alpha \neq 10\% & H_1 : \alpha \neq 5\% & H_1 : \alpha \neq 1\% \end{array}$$

sendo estes aplicados para as taxas de erro tipo I. Se a hipótese nula for rejeitada e a taxa de erro tipo I for considerada significativamente ( $p < 0,01$ ) inferior ao nível nominal, o teste será considerado conservativo; se a hipótese nula for rejeitada e a taxa de erro tipo I for considerada significativamente ( $p < 0,01$ ) superior ao nível nominal, o teste será considerado liberal; e se a taxa de erro tipo I observada for não significativamente ( $p > 0,01$ ) diferente ao nível nominal, o teste será considerado exato. Se  $y$  representa o número de hipóteses nulas rejeitadas nas  $M = 1000$  simulações de Monte Carlo ao nível nominal de significância  $\alpha$ , então, pode-se obter a estatística do teste usando a relação entre as distribuições  $F$  e binomial, com probabilidade de sucesso  $p = \alpha$  por meio de

$$F = \left( \frac{y + 1}{M - y} \right) \left( \frac{1 - \alpha}{\alpha} \right),$$

que, sob a hipótese nula, segue uma distribuição  $F$  com  $\nu_1 = 2(M - y)$  e  $\nu_2 = 2(y + 1)$  graus de liberdade. Se  $F \leq F_{0,005}$  ou  $F \geq F_{0,995}$ , a hipótese nula deve ser rejeitada ao nível de  $1\%$ , em que  $F_{0,005}$  e  $F_{0,995}$  são os quantis da distribuição  $F$  com  $\nu_1$  e  $\nu_2$  graus de liberdade.

### 5.1 Erro tipo I sob hipótese nula completa

O tamanho é uma fundamental característica da performance de um teste. Um teste é considerado exato se tem tamanho real igual ao tamanho nominal  $\alpha$ . Testes que tenham um tamanho real menor que o nominal são considerados conservativos e aqueles com taxas de erros tipo I maiores que o nível nominal  $\alpha$  são liberais.

As taxas de erro tipo I para os testes TG e F da análise de variância são apresentadas na Tabela 4 em função do número de repetições ( $r$ ), do número de tratamentos ( $k$ ) e dos níveis nominais de credibilidades ( $\alpha$ ), sob  $H_0$  completa, considerando-se a distribuição normal padrão ( $\lambda = 0$ ) para os resíduos, obtidas em 1000 simulações Monte Carlo.

De modo geral, pode-se observar que o teste TG é considerado conservativo, ou seja, as taxas empíricas de erros tipo I ficaram bem abaixo dos valores nominais estabelecidos.

Tabela 4 - Taxas de erro tipo I (%) para os testes TG e F, em função do número de repetições ( $r$ ), do número de tratamentos ( $k$ ) e dos níveis nominais de credibilidades ( $\alpha$ ), sob  $H_0$  completa, considerando-se a distribuição normal padrão ( $\lambda = 0$ ) para os resíduos

$r$	$k$	$\alpha = 1\%$		$\alpha = 5\%$		$\alpha = 10\%$	
		TG	F	TG	F	TG	F
4	5	0,0*	1,2	0,2*	5,0	0,9*	10,2
	10	0,1*	1,3	0,1*	5,6	0,2*	11,3
	20	0,1*	1,0	0,5*	4,7	0,7*	8,8
10	5	0,0*	0,9	0,0*	5,3	0,1*	10,8
	10	0,2*	0,8	0,3*	5,7	0,4*	10,2
	20	0,1*	0,9	0,1*	4,4	0,1*	10,0

\* significativamente diferente ( $p < 0,01$ ) do nível de credibilidade nominal.

Em nenhuma das configurações avaliadas o TG apresentou desempenho liberal. Isto mostra que este teste, considerando-se a distribuição normal padrão ( $\lambda = 0$ ) para os resíduos, controlou o erro tipo I sob a hipótese  $H_0$  completa, embora em forma conservativa.

Nota-se ainda que o teste F apresentou em todas as configurações consideradas taxas de erro tipo I muito próximas dos níveis nominais de credibilidades adotados, o que era esperado, pois está se considerando neste caso a distribuição normal padrão para os resíduos.

Os resultados obtidos para as taxas de erro tipo I dos testes TG e F, considerando a distribuição normal assimétrica com  $\lambda = 1$  e  $\lambda = 5$  para os resíduos do modelo são apresentados nas Tabelas 5 e 6, respectivamente.

Tabela 5 - Taxas de erro tipo I (%) para os testes TG e F, em função do número de repetições ( $r$ ), do número de tratamentos ( $k$ ) e dos níveis nominais de credibilidades ( $\alpha$ ), sob  $H_0$  completa, considerando-se a distribuição normal assimétrica padrão ( $\lambda = 1$ ) para os resíduos

$r$	$k$	$\alpha = 1\%$		$\alpha = 5\%$		$\alpha = 10\%$	
		TG	F	TG	F	TG	F
4	5	0,1*	1,1	0,1*	5,6	0,4*	10,5
	10	0,1*	1,0	0,4*	5,7	0,6*	11,1
	20	0,1*	1,5	0,2*	5,1	0,9*	9,2
10	5	0,0*	1,0	0,0*	4,3	0,0*	9,1
	10	0,1*	1,2	0,2*	5,5	0,3*	10,3
	20	0,2*	1,1	0,4*	5,6	0,5*	11,8

\* significativamente diferente ( $p < 0,01$ ) do nível de credibilidade nominal.



Tabela 6 - Taxas de erro tipo I (%) para os testes TG e F, em função do número de repetições ( $r$ ), do número de tratamentos ( $k$ ) e dos níveis nominais de credibilidades ( $\alpha$ ), sob  $H_0$  completa, considerando-se a distribuição normal assimétrica padrão ( $\lambda = 5$ ) para os resíduos

$r$	$k$	$\alpha = 1\%$		$\alpha = 5\%$		$\alpha = 10\%$	
		TG	F	TG	F	TG	F
4	5	0,2*	1,6	0,9*	5,9	1,2*	9,9
	10	0,3*	0,8	0,8*	4,8	1,9*	9,4
	20	1,4	0,7	5,4	4,4	10,7	9,3
10	5	0,0*	0,7	0,1*	5,1	0,2*	10,2
	10	0,0*	1,1	0,3*	6,0	1,5*	9,9
	20	0,4*	0,8	1,9*	5,1	3,6*	9,4

\* significativamente diferente ( $p < 0,01$ ) do nível de credibilidade nominal.

Das Tabelas 5 e 6, nota-se que os valores obtidos para as taxas de erro tipo I para os testes TG e F não foram relativamente afetados quando se considerou a distribuição normal assimétrica para os resíduos, isto é, o teste TG se manteve muito conservativo e o teste F apresentou novamente taxas de erro tipo I muito próximas dos níveis nominais adotados. Isso provavelmente ocorreu devido ao fato de a distribuição normal assimétrica não apresentar comportamento muito diferente da distribuição normal.

Observando-se com maior detalhe a Tabela 6, verifica-se que para uma específica configuração ( $r = 4$ ,  $k = 20$  e normal assimétrica com  $\lambda = 5$  para os resíduos), o teste TG apresentou-se exato, com taxas de erro tipo I muito próximas aos valores nominais, tendo, portanto, desempenho semelhante ao teste F nessa configuração. Há de se ressaltar que o teste F é considerado robusto a violações da suposição de normalidade dos resíduos, como observado por vários autores, entre eles, Ringland (1983) e Driscoll (1996).

Nota-se ainda na Tabela 6 que houve um pequeno aumento nos valores empíricos das taxas de erro tipo I para o teste TG, independente da configuração, comparados aos resultados da Tabela 4.

Com base nestes resultados, provavelmente os valores de poder do teste TG serão inferiores aos valores de poder do teste F. No entanto, se isso não ocorrer ou, se pelo menos o poder destes testes forem próximos, será possível afirmar que o teste TG caminha para um teste ideal, isto é, baixas taxas de erro tipo I e alto poder.

## 5.2 Poder sobre sob hipótese nula parcial

As estimativas empíricas de poder para os testes TG e F são apresentadas na Figura 2, em função do número de repetições ( $r$ ) e do número de tratamentos

( $k$ ), considerando-se a distribuição normal padrão ( $\lambda = 0$ ) para os resíduos, para  $\alpha = 5\%$ .

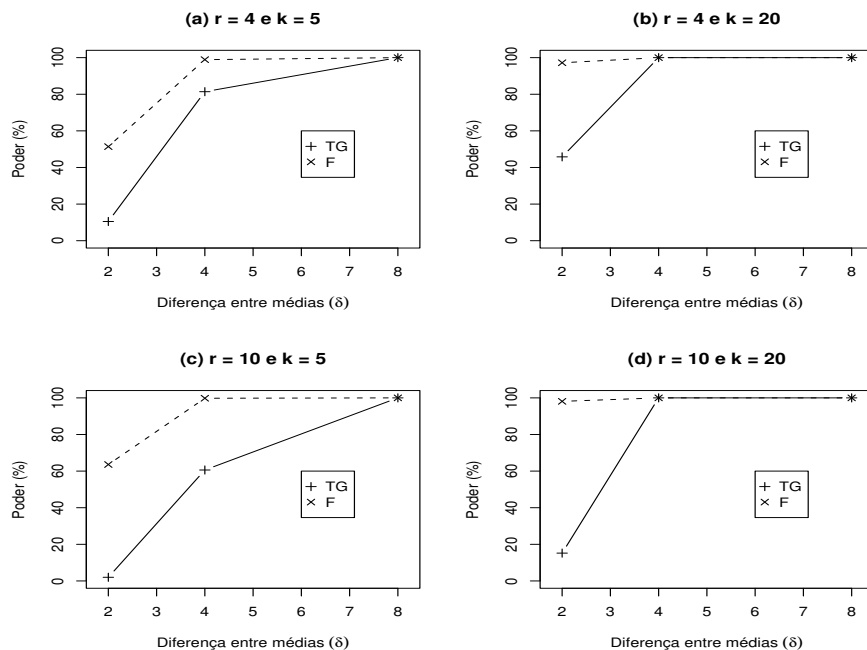


Figura 2 - Poder dos testes TG e F, em função do número de repetições ( $r$ ) e do número de tratamentos ( $k$ ), considerando-se a distribuição normal padrão ( $\lambda = 0$ ) para os resíduos, para  $\alpha = 5\%$ .

Nota-se na Figura 2 que o teste TG apresentou em todas as configurações consideradas, taxas empíricas de poder inferiores ao do teste F, isto para pequenas diferenças entre médias ( $\delta = 2$ ). Para diferenças maiores ( $\delta = 4$  e  $\delta = 8$ ), nota-se que o teste TG apresentou desempenho similar ao teste F, o que de certo ponto induz que este teste é uma alternativa viável quando a distribuição dos resíduos é a distribuição normal padrão. A inferioridade do poder do TG para  $\delta$  pequeno já era esperado, devido ao fato do TG apresentar desempenho mais conservativo do que o teste F, com relação as taxas empíricas de erro tipo I, como descrito anteriormente.

Outro fato esperado quando se analisam taxas empíricas de poder é que, com o aumento da diferença ( $\delta$ ) entre as médias dos tratamentos dos dois grupos, o poder de todos os testes analisados aumentem. Este fato se verifica tanto para o teste TG quanto para o teste F, como pode ser constatado na Figura 2, indicando uma possível comprovação da exatidão do programa de simulação.

Por fim, ainda na Figura 2, nota-se de maneira geral, que ambos os testes tiveram um pequeno aumento no poder quando os parâmetros  $r$  ou  $k$  aumentaram,

e que, com o aumento de  $k$ , o teste TG melhorou seu desempenho relativo mais rapidamente com o aumento de  $\delta$  do que com o aumento de  $r$ .

Na Figura 3 são apresentados os resultados quando assumiu-se a distribuição normal assimétrica padrão com parâmetro de assimetria  $\lambda = 5$  para os resíduos.

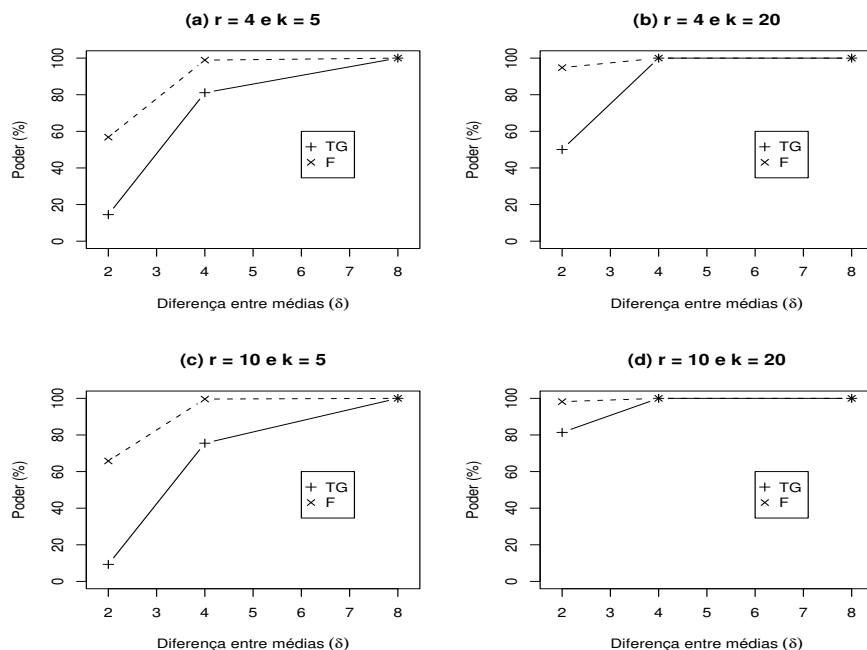


Figura 3 - Poder dos testes TG e F, em função do número de repetições ( $r$ ) e do número de tratamentos ( $k$ ), considerando-se a distribuição normal assimétrica padrão ( $\lambda = 5$ ) para os resíduos, para  $\alpha = 5\%$ .

De maneira geral pode-se afirmar que não houve mudança no comportamento das taxas empíricas de poder de ambos os testes quanto da troca da distribuição dos resíduos de  $N(0,1)$  para  $NA(5)$ . No entanto, observa-se na Figura 3 que as taxas empíricas de poder do teste TG ficaram relativamente maiores das obtidas na Figura 2. Por exemplo, para a configuração  $r = 10$ ,  $k = 20$  e  $\delta = 2$ , em que se considera a distribuição  $N(0,1)$  para os erros (Figura 2 (d)), o poder do teste TG está em torno de 20%, enquanto que, ao trocar a distribuição dos erros para  $NA(5)$ , nesta mesma configuração (Figura 3 (d)), a taxa empírica de poder deste teste salta para aproximadamente 80%. Em particular, para esta específica configuração, pode-se afirmar que o teste TG teve uma bom desempenho.

Novamente, como observado na Figura 2, constata-se na Figura 3 que as taxas

empíricas dos testes TG e F aumentaram com o aumento da diferença ( $\delta$ ) entre as médias dos tratamentos dos dois grupos, com o aumento do número de repetições ( $r$ ) e, também, com o aumento do número de tratamentos ( $k$ ), fixados os demais parâmetros, com era esperado.

## Conclusões

O teste para hipótese nula global de igualdade de médias supondo erros normais assimétricos foi proposto com sucesso, devido ao fato do mesmo apresentar as taxas de erro tipo I iguais ou inferiores ao níveis de credibilidades adotados sob todas as configurações simuladas e apresentar poder similar ao teste F. Comprovou-se novamente a grande robustez do teste F de Fisher. Acredita-se ser mais realista ao conjunto de dados da aplicação prática, a suposição de que os erros do modelo tenham distribuição normal assimétrica, o que de certa forma, faz com que à abordagem do teste de hipótese proposto neste trabalho seja mais confiável.

OLIVEIRA, M. S.; FURTADO, D. F. Test for equality of independent means with skew-normal error. *Rev. Bras. Biom.*, São Paulo, v.28, n.2, p.1-26, 2010.

- **ABSTRACT:** *In design studies, the normality of the error terms is the usual assumption when tests for equality of independent means are performed. However, this assumption could be unrealistic and the inferences which come from these tests can not be reliable. In this paper, the normality assumption was encapsulated by the skew normally distributed error assumption. By the way, a new test for mean equality was proposed, based on quadratic forms. It will be used to test the overall equality hypothesis of independent means in the Bayesian approach. A practical application with a real data set was developed and discussed. On the evaluation of the performance of the proposed test, it was presented a Monte Carlo simulation in several experimental conditions. Moreover, the simulation results were compared with the usual F test. The proposed test has presented good performance by being exact or conservative in all simulated designs and, at the same time, has presented close power rates to the usual F test.*
- **KEYWORDS:** *Testing of hypothesis; skew-normal distribution; Bayesian inference; Monte Carlo simulation.*

## Referências

- AZZALINI, A. A class of distributions which includes the normal ones. *Scand. J. Stat.*, Oxford, v.12, n.2, p.171-178, 1985.
- AZZALINI, A.; CAPITANIO, A. Statistical applications of the multivariate skew normal distribution. *J. R. Stat. Soc. Ser. B: Stat. Methodol.*, London, v.61, n.3, p.579-602, 1999.

- AZZALINI, A.; DALLA-VALLE, A. The multivariate skew-normal distribution. *Biometrika*, London, v.83, n.4, p.715-726, 1996.
- BAYES, C. L.; BRANCO, M. D. Bayesian inference for the skewness parameter of the scalar skew-normal distribution. *Braz. J. Prob. Stat.*, São Paulo, v.21, n.2, p.141-163, 2007.
- BIASE, N. G.; FERREIRA, D. F. Comparações múltiplas e testes simultâneos para parâmetros binomiais de  $k$  populações independentes. *Rev. Bras. Biom.*, São Paulo, v.27, n.3, p.301-323, 2009.
- DRISCOLL, W. D. Robustness of the ANOVA and Tukey-Kramer statistical tests. *Comp. Ind. Eng.*, New York, v.31, n.1/2, p.265-268, 1996.
- FANG, K. T.; KOTZ, S.; NG, K. W. *Symmetric multivariate and related distributions*. New York: Chapman and Hall, 1990. 220p.
- FERREIRA, D. F. *Estatística multivariada*. Lavras: UFLA, 2008. 662p.
- GEMAN, S.; GEMAN, D. Stochastic relaxation, Gibbs distribution and the Bayesian restoration of imagens. *IEEE Trans. Patt. Anal. Machine Intell.*, New York, v.6, n.6, p.721-741, 1984.
- GENTON, M. G. *Skew-elliptical distributions and their applications: a journey beyond normality*. Boca Raton: CRC, 2004. 416p.
- GENTON, M. G.; HE, L.; LIU, X. Moments of skew-normal random vectors and their quadratic forms. *Stat. Prob. Lett.*, Amsterdam, v.51, n.4, p.319-325, 2001.
- HENZE, N. A probabilistic representation of the skew-normal distribution. *Scand. J. Stat.*, Oxford, v.13, n.4, p.271-275, 1986.
- JOHNSON, R. A.; WEERAHANDI, S. A Bayesian solution to the multivariate Behrens-Fisher problem. *J. Am. Stat. Assoc.*, New York, v.83, n.401, p.145-149, 1988.
- LISEO, B.; LOPERFIDO, N. A note on reference priors for the scalar skew-normal distribution. *J. Stat. Plan. Infer.*, Amsterdam, v.136, n.2, p.373-389, 2006.
- O'HAGAN, A.; LEONARD, T. Bayes estimation subject to uncertainty about parameter constraints. *Biometrika*, London, v.63, p.201-203, 1976.
- RAFTERY, A. E.; LEWIS, S. How many iterations in the Gibbs sampler? In: BERNARDO, J. M.; BERGER, J. O.; DAWID, A. P.; SMITH, A. F. M. (Ed.). *Bayesian statistics*. 4.ed. Oxford: University of Oxford, 1992. p.763-773.
- RINGLAND, J. T. Robust multiple comparisons. *J. Am. Stat. Assoc.*, New York, v.78, n.381, p.145-151, 1983.
- ROBERTS, C. A correlation model useful in study of twins. *J. Am. Stat. Assoc.*, New York, v.61, n.316, p.1184-1190, 1966.
- SAHU, S. K.; DEY, D. K.; BRANCO, M. A new class of multivariate skew distributions with applications to bayesian regression models. *Can. J. Stat.*, Toronto, v.31, n.2, p.129-150, 2003.

SARTORI, N. Bias prevention of maximum likelihood estimates for scalar skew normal and skew t distributions. *J. Stat. Plan. Infer.*, Amsterdam, v.136, n.12, p.4259-4275, 2006.

STEEL, R. G. D.; TORRIE, J. H. *Principles and procedures of statistics*. 2.ed. New York: McGraw-Hill, 1980. 633p.

Recebido em 07.12.2009.

Aprovado após revisão em 11.05.2010.

## Apêndice

Apresentam-se neste apêndice a demonstração das obtenções das distribuições condicionais completas a posteriori apresentadas na seção 3.1.

A distribuição condicional completa para cada parâmetro  $(\boldsymbol{\theta}, \eta, \tau, w$  e  $u_{ij})$  foi obtida a partir da distribuição conjunta a posteriori dada em (13), considerando somente os termos em que o respectivo parâmetro aparece. Como na expressão (13) apresentou-se apenas o núcleo da distribuição conjunta a posteriori, será obtido portanto, somente o núcleo da distribuição condicional completa, ao qual, se poderá identificar a distribuição que a mesma pertence.

Apresenta-se na sequência cada distribuição condicional completa por meio de proposições com as respectivas demonstrações.

**Proposição 1.** *A distribuição condicional completa do parâmetro  $w$  é dada por*

$$w|\boldsymbol{\theta}, \eta, \tau, \mathbf{U}, \mathbf{Y}, \mathbf{X} \sim \text{Gama} \left( \frac{\kappa + 1}{2}, \frac{1}{2} \left( \frac{\eta^2}{\tau^2 \sigma_t^2} + \kappa \right) \right).$$

**Demonstração:** Considerando somente os termos em que o parâmetro  $w$  aparece na expressão (13) tem-se

$$\begin{aligned} \pi(w|\boldsymbol{\theta}, \eta, \tau, \mathbf{U}, \mathbf{Y}, \mathbf{X}) &\propto \exp \left( -\frac{1}{2} \frac{\eta^2 w}{\tau^2 \sigma_t^2} \right) w^{\frac{(\kappa+1)}{2}-1} \exp \left( -\frac{\kappa}{2} w \right) \\ &\propto w^{\frac{(\kappa+1)}{2}-1} \exp \left\{ -\frac{1}{2} \left( \frac{\eta^2}{\tau^2 \sigma_t^2} + \kappa \right) w \right\}, \end{aligned}$$

que é o núcleo da distribuição  $\text{Gama} \left( \frac{\kappa+1}{2}, \frac{1}{2} \left( \frac{\eta^2}{\tau^2 \sigma_t^2} + \kappa \right) \right)$ .

**Proposição 2.** *A distribuição condicional completa do parâmetro  $\mathbf{U}$  é dada por*

$$\mathbf{U}|\boldsymbol{\theta}, \eta, \tau, w, \mathbf{Y}, \mathbf{X} \sim \text{NT}_n \left( \frac{(\mathbf{Y} - \mathbf{X}\boldsymbol{\theta})\eta}{\eta^2 + \tau^2}, \left( \frac{\tau^2}{\eta^2 + \tau^2} \right) \mathbf{I}_n \right).$$

**Demonstração:** Inicialmente, será provado que cada  $u_{ij}$ ,  $i = 1, \dots, k$  e  $j = 1, \dots, r$  tem distribuição normal truncada positiva, com parâmetro de posição  $(y_{ij} - \mathbf{x}_i^\top \boldsymbol{\theta})\eta / (\eta^2 + \tau^2)$  e de escala  $\tau^2 / (\eta^2 + \tau^2)$ , com  $\mathbf{x}_i^\top$ , de dimensão  $(1 \times k)$ , representando o vetor que possui 0 em todas as posições, exceto na  $i$ -ésima, onde o valor é igual a 1 (por exemplo,  $\mathbf{x}_1^\top = [1, 0, \dots, 0]$ ).

Considerando somente os termos em que o parâmetro  $u_{ij}$  aparece na expressão (13) tem-se

$$\pi(u_{ij}|\boldsymbol{\theta}, \eta, \tau, w, \mathbf{Y}, \mathbf{X}) \propto e^{-\frac{1}{2\tau^2}(y_{ij} - \mathbf{x}_i^\top \boldsymbol{\theta} - \eta u_{ij})^2} \times e^{-\frac{1}{2} u_{ij}^2} \mathbf{I}_{[0, +\infty)}(u_{ij})$$

$$\begin{aligned}
&\propto e^{-\frac{1}{2\tau^2}((y_{ij}-\mathbf{x}_i^\top\boldsymbol{\theta})^2-2(y_{ij}-\mathbf{x}_i^\top\boldsymbol{\theta})\eta u_{ij}+\eta^2 u_{ij}^2)-\frac{1}{2}u_{ij}^2} \\
&\quad \times \mathbf{I}_{[0,+\infty)}(u_{ij}) \\
&\propto e^{-\frac{1}{2\tau^2}((\eta^2+\tau^2)u_{ij}^2-2(y_{ij}-\mathbf{x}_i^\top\boldsymbol{\theta})\eta u_{ij})} \mathbf{I}_{[0,+\infty)}(u_{ij}) \\
&\propto e^{-\frac{1}{2}\frac{\eta^2+\tau^2}{\tau^2}\left(u_{ij}-\frac{(y_{ij}-\mathbf{x}_i^\top\boldsymbol{\theta})\eta}{\eta^2+\tau^2}\right)^2} \mathbf{I}_{[0,+\infty)}(u_{ij}), \tag{22}
\end{aligned}$$

para todo par  $(i,j)$ . Note-se que (22) é o núcleo da distribuição normal truncada positiva, isto é,

$$u_{ij}|\boldsymbol{\theta},\eta,\tau,w,\mathbf{Y},\mathbf{X} \sim N\left(\frac{(y_{ij}-\mathbf{x}_i^\top\boldsymbol{\theta})\eta}{\eta^2+\tau^2},\frac{\tau^2}{\eta^2+\tau^2}\right)\mathbf{I}_{[0,+\infty)}(u_{ij}),$$

para todo par  $(i,j)$ . Alocando as variáveis latentes  $u_{ij}$  em um vetor  $(\mathbf{U})$  e levando-se em conta que as mesmas são mutuamente independentes, pode-se escrever a distribuição condicional completa do vetor latente  $\mathbf{U}$  como uma distribuição normal truncada positiva multivariada dada por

$$\mathbf{U}|\boldsymbol{\theta},\eta,\tau,w,\mathbf{Y},\mathbf{X} \sim NT_n\left(\frac{(\mathbf{Y}-\mathbf{X}\boldsymbol{\theta})\eta}{\eta^2+\tau^2},\left(\frac{\tau^2}{\eta^2+\tau^2}\right)\mathbf{I}_n\right),$$

finalizando a demonstração.

**Proposição 3.** *A dist. condicional completa do parâmetro  $\boldsymbol{\theta}$  é dada por*

$$\boldsymbol{\theta}|\eta,\tau,w,\mathbf{U},\mathbf{Y},\mathbf{X} \sim N_k\left((\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}(\mathbf{Y}-\eta\mathbf{U}),\tau^2(\mathbf{X}^\top\mathbf{X})^{-1}\right).$$

**Demonstração:** Considerando somente os termos em que o parâmetro  $\boldsymbol{\theta}$  aparece na expressão (13) tem-se

$$\pi(\boldsymbol{\theta}|\eta,\tau,w,\mathbf{U},\mathbf{Y},\mathbf{X}) \propto e^{-\frac{1}{2\tau^2}\underbrace{(\mathbf{Y}-\mathbf{X}\boldsymbol{\theta}-\eta\mathbf{U})^\top(\mathbf{Y}-\mathbf{X}\boldsymbol{\theta}-\eta\mathbf{U})}_{\star}}. \tag{23}$$

Agora, trabalhando somente a forma quadrática dada em  $(\star)$ , tem-se

$$\begin{aligned}
\star &= (\mathbf{Y}-\mathbf{X}\boldsymbol{\theta}-\eta\mathbf{U})^\top(\mathbf{Y}-\mathbf{X}\boldsymbol{\theta}-\eta\mathbf{U}) \\
&= ((\mathbf{Y}-\eta\mathbf{U})-\mathbf{X}\boldsymbol{\theta})^\top((\mathbf{Y}-\eta\mathbf{U})-\mathbf{X}\boldsymbol{\theta}) \\
&= (\mathbf{Y}-\eta\mathbf{U})^\top(\mathbf{Y}-\eta\mathbf{U})-(\mathbf{Y}-\eta\mathbf{U})^\top\mathbf{X}\boldsymbol{\theta} \\
&\quad -\boldsymbol{\theta}^\top\mathbf{X}^\top(\mathbf{Y}-\eta\mathbf{U})+\boldsymbol{\theta}^\top\mathbf{X}^\top\mathbf{X}\boldsymbol{\theta} \\
&\propto \boldsymbol{\theta}^\top\mathbf{X}^\top\mathbf{X}\boldsymbol{\theta}-2\boldsymbol{\theta}^\top\mathbf{X}^\top(\mathbf{Y}-\eta\mathbf{U}) \\
&\propto [\boldsymbol{\theta}-(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top(\mathbf{Y}-\eta\mathbf{U})]^\top(\mathbf{X}^\top\mathbf{X})[\boldsymbol{\theta}-(\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top(\mathbf{Y}-\eta\mathbf{U})],
\end{aligned}$$



de onde se encontra que (23) com a substituição de ( $\mathbf{X}$ ) é o núcleo da distribuição normal multivariada, a saber

$$N_k \left( (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top (\mathbf{Y} - \eta \mathbf{U}), \tau^2 (\mathbf{X}^\top \mathbf{X})^{-1} \right).$$

**Proposição 4.** *A distribuição condicional completa do parâmetro  $\eta$  é dada por*

$$\eta | \boldsymbol{\theta}, \tau, w, \mathbf{U}, \mathbf{Y}, \mathbf{X} \sim N \left( \frac{(\mathbf{Y} - \mathbf{X}\boldsymbol{\theta})^\top \mathbf{U}}{\frac{w}{\sigma_t^2} + \mathbf{U}^\top \mathbf{U}}, \frac{\tau^2}{\frac{w}{\sigma_t^2} + \mathbf{U}^\top \mathbf{U}} \right).$$

**Demonstração:** Considerando somente os termos em que o parâmetro  $\eta$  aparece na expressão (13) tem-se

$$\begin{aligned} \pi(\eta | \boldsymbol{\theta}, \tau, w, \mathbf{U}, \mathbf{Y}, \mathbf{X}) &\propto e^{-\frac{1}{2} \frac{\eta^2 w}{\tau^2 \sigma_t^2}} \times e^{-\frac{1}{2\tau^2} (\mathbf{Y} - \mathbf{X}\boldsymbol{\theta} - \eta \mathbf{U})^\top (\mathbf{Y} - \mathbf{X}\boldsymbol{\theta} - \eta \mathbf{U})} \\ &\propto e^{-\frac{1}{2} \frac{w}{\tau^2 \sigma_t^2} \eta^2 - \frac{1}{2\tau^2} ((\mathbf{Y} - \mathbf{X}\boldsymbol{\theta}) - \eta \mathbf{U})^\top ((\mathbf{Y} - \mathbf{X}\boldsymbol{\theta}) - \eta \mathbf{U})} \\ &\propto e^{-\frac{1}{2} \frac{w}{\tau^2 \sigma_t^2} \eta^2 - \frac{1}{2\tau^2} [-2(\mathbf{Y} - \mathbf{X}\boldsymbol{\theta})^\top \eta \mathbf{U} + \mathbf{U}^\top \mathbf{U} \eta^2]} \\ &\propto e^{-\frac{1}{2} \left[ \left( \frac{w}{\tau^2 \sigma_t^2} + \frac{\mathbf{U}^\top \mathbf{U}}{\tau^2} \right) \eta^2 - \frac{2(\mathbf{Y} - \mathbf{X}\boldsymbol{\theta})^\top \mathbf{U}}{\tau^2} \eta \right]} \\ &\propto e^{-\frac{1}{2} \frac{w + \sigma_t^2 \mathbf{U}^\top \mathbf{U}}{\tau^2 \sigma_t^2} \left( \eta - \frac{(\mathbf{Y} - \mathbf{X}\boldsymbol{\theta})^\top \mathbf{U}}{\frac{w + \sigma_t^2 \mathbf{U}^\top \mathbf{U}}{\sigma_t^2}} \right)^2} \\ &\propto e^{-\frac{1}{2} \frac{\frac{w}{\sigma_t^2} + \mathbf{U}^\top \mathbf{U}}{\tau^2} \left( \eta - \frac{(\mathbf{Y} - \mathbf{X}\boldsymbol{\theta})^\top \mathbf{U}}{\frac{w}{\sigma_t^2} + \mathbf{U}^\top \mathbf{U}} \right)^2}, \end{aligned}$$

que é o núcleo da distribuição  $N \left( \frac{(\mathbf{Y} - \mathbf{X}\boldsymbol{\theta})^\top \mathbf{U}}{\frac{w}{\sigma_t^2} + \mathbf{U}^\top \mathbf{U}}, \frac{\tau^2}{\frac{w}{\sigma_t^2} + \mathbf{U}^\top \mathbf{U}} \right)$ .

**Proposição 5.** *A distribuição condicional completa do parâmetro  $1/\tau^2$  é dada por*

$$\begin{aligned} \frac{1}{\tau^2} \Big| \boldsymbol{\theta}, \eta, w, \mathbf{U}, \mathbf{Y}, \mathbf{X} &\sim \text{Gama}(a, b), \quad \text{com} \\ a = \frac{n+1}{2} \quad e \quad b &= \frac{1}{2} \left( (\mathbf{Y} - \mathbf{X}\boldsymbol{\theta} - \eta \mathbf{U})^\top (\mathbf{Y} - \mathbf{X}\boldsymbol{\theta} - \eta \mathbf{U}) + \frac{w\eta^2}{\sigma_t^2} \right). \end{aligned}$$

**Demonstração:** Considerando somente os termos em que o parâmetro  $\tau$  aparece

na expressão (13) tem-se

$$\begin{aligned}\pi(\tau|\boldsymbol{\theta}, \eta, w, \mathbf{U}, \mathbf{Y}, \mathbf{X}) &\propto \frac{1}{\tau^2} e^{-\frac{1}{2} \frac{\eta^2 w}{\tau^2 \sigma_t^2}} \times \frac{1}{\tau^n} e^{-\frac{1}{2\tau^2} (\mathbf{Y} - \mathbf{X}\boldsymbol{\theta} - \eta\mathbf{U})^\top (\mathbf{Y} - \mathbf{X}\boldsymbol{\theta} - \eta\mathbf{U})} \\ &\propto \left(\frac{1}{\tau^2}\right)^{\frac{n+2}{2}} e^{-\frac{1}{2} \left( (\mathbf{Y} - \mathbf{X}\boldsymbol{\theta} - \eta\mathbf{U})^\top (\mathbf{Y} - \mathbf{X}\boldsymbol{\theta} - \eta\mathbf{U}) + \frac{w\eta^2}{\sigma_t^2} \right) \frac{1}{\tau^2}} \\ &\propto \left(\frac{1}{\tau^2}\right)^{a+\frac{1}{2}} \exp\left\{-b\frac{1}{\tau^2}\right\},\end{aligned}$$

com

$$a = \frac{n+1}{2} \quad e \quad b = \frac{1}{2} \left( (\mathbf{Y} - \mathbf{X}\boldsymbol{\theta} - \eta\mathbf{U})^\top (\mathbf{Y} - \mathbf{X}\boldsymbol{\theta} - \eta\mathbf{U}) + \frac{w\eta^2}{\sigma_t^2} \right). \quad (24)$$

No entanto, o interesse está na distribuição condicional completa a posteriori de  $1/\tau^2$ . Para tanto, considere-se a seguinte transformação de variável  $T = g(\tau) = 1/\tau^2$ . Logo,  $\tau = h(T) = T^{-1/2}$ . O jacobiano da transformação (em relação a  $t$ ) fica

$$J(\tau, t) = \frac{dh(t)}{dt} = -\frac{1}{2}t^{-3/2}.$$

Com isso, a distribuição condicional completa a posteriori de  $T$ , isto é, de  $1/\tau^2$  será

$$\begin{aligned}\pi_T(T|\boldsymbol{\theta}, \eta, w, \mathbf{U}, \mathbf{Y}, \mathbf{X}) &= \pi_\tau(h(t)|\boldsymbol{\theta}, \eta, w, \mathbf{U}, \mathbf{Y}, \mathbf{X}) \times |J(\tau, t)| \\ &\propto t^{a+1/2} \exp\{-bt\} \times \left| -\frac{1}{2}t^{-3/2} \right| \\ &\propto t^{a-1} \exp\{-bt\}\end{aligned}$$

que é o núcleo da distribuição *Gama*( $a, b$ ) com  $a$  e  $b$  dados na expressão (24), finalizando a demonstração.