COMPARAÇÕES MÚLTIPLAS BAYESIANAS EM MODELOS NORMAIS HOMOCEDÁSTICOS E HETEROCEDÁSTICOS

Bayesian multiple comparisons in homocedastic and heterocedastic normal models

Paulo César de Resende Andrade[1], Daniel Furtado Ferreira[2]

RESUMO

Procedimentos de comparações múltiplas são utilizados para comparar médias de níveis de um fator, porém, os testes mais populares apresentam problemas de ambiguidade dos resultados e de controle do erro tipo I, além de terem seus desempenhos influenciados negativamente no caso de heterogeneidade de variâncias e não balanceamento. Objetivou-se, neste trabalho propor alternativas bayesianas para comparações múltiplas considerando os casos de homogeneidade e heterogeneidade de variâncias. A metodologia utilizada nesse trabalho foi baseada na distribuição *a posteriori t* multivariada. Foram geradas *k* cadeias de médias, utilizando o método de Monte Carlo. Foi obtida a amplitude padronizada sob H_0, obtida na distribuição *a posteriori* das médias, contemplando a possibilidade de se analisar tanto o caso de variâncias heterogêneas como o caso de variâncias homogêneas. Os procedimentos de comparações múltiplas bayesianos foram propostos com sucesso.

Termos para indexação: Simulação, erro tipo I, poder, Monte Carlo, não balanceamento.

ABSTRACT

Multiple comparison procedures are used to compare factor's levels means, since the most popular tests show problems related to ambiguous results and to the control of the type I error rates. Moreover, their performance is worst in heterocedastics and unbalanced cases. The objective of this work is to propose a Bayesian alternative for multiple comparisons considering the homocedastic and heterocedastic normal models. The methodology adopted in this paper was based on a *posteriori* multivariate t distribution. It was used k Monte Carlo chains of the mean factor to make inferences. The standardized range was obtained, under H_0, from the *posteriori* distribution of the means, for the analysis of homocedastic and heterocedastic cases. The bayesian procedures of multiple comparisons were successfully proposed.

Index terms: Simulation, type I error rate, power, Monte Carlo, unbalanced.

(Recebibo em 10 de abril de 2008 e aprovado em 23 de abril de 2010)

INTRODUÇÃO

Na experimentação, o pesquisador, geralmente, utiliza a análise de variância para avaliar a significância dos efeitos dos tratamentos, deparando-se com o problema de comparar médias de diferentes níveis do fator. Para verificar diferenças entre médias de tratamentos são sugeridos vários procedimentos de comparações múltiplas. O grande problema desses testes é a ambiguidade dos resultados (Machado et al., 2005). Um segundo problema, não menos importante, é o do controle do erro tipo I (Hochberg & Tamhane, 1987; Hsu, 1996).

O problema da ambiguidade pode ser contornado com métodos alternativos de agrupamento, mas que apresentam a desvantagem de serem válidos apenas sob normalidade. Em situações de heterogeneidade de variâncias podem ser utilizados métodos de reamostragem *bootstrap*.

Uma alternativa é o uso de procedimentos bayesianos. Uma quantidade razoável de artigos, a maioria recente, leva em consideração o problema de comparações múltiplas do ponto de vista bayesiano. Duncan (1965) dá a primeira exposição completa de uma aproximação bayesiana para o problema de comparação múltipla. Waller & Duncan (1969) ampliaram o procedimento original de Duncan (1965). Berry (1988) e Berry & Hochberg (1999) usaram o processo Dirichlet como distribuição *a priori* no contexto de multiplicidades. Gopalan & Berry (1998) desenvolveram essa aproximação, para obter probabilidades *a posteriori* para hipóteses de igualdade entre médias populacionais, com variâncias iguais.

Shaffer (1999) modificou o procedimento de Waller & Duncan (1969), postulando *prioris* não informativas. Bratcher & Hamilton (2005) apresentaram um procedimento de comparações múltiplas bayesiano para médias ranqueadas e obtidas de populações distribuídas normalmente. Cho et al. (2006) consideraram o problema específico de comparações múltiplas para populações com distribuição binomial negativa, aplicando a família Dirichlet,

[1]Universidade Federal dos Vales do Jequitinhonha e Mucuri/UFVJM – Instituto de Ciência e Tecnologia/ICT – Campus II – Rodovia MGT 367 – Km 583 – 5000 – Alto da Jacuba – Diamantina, MG – paulo.andrade@ufvjm.edu.br
[2]Universidade Federal de Lavras/UFLA – Departamento de Ciências Exatas/DEX – Lavras, MG

na forma de *prioris*. Porém, nenhum desses procedimentos enfoca o caso normal homocedástico ou heterocedástico.

Objetivou-se neste trabalho, propor alternativas bayesianas para comparações múltiplas considerando os casos de homogeneidade e heterogeneidade de variâncias em modelos com distribuição normal, balanceados ou não, ilustradas em exemplos simulados.

MATERIAL E METODOS

Considere o caso especial em que os elementos do vetor de observações $\boldsymbol{y}$ são amostras independentes $\boldsymbol{y}'_1 = (y_{11}, ..., y_{1n1}), ..., \boldsymbol{y}'_k = (y_{k1}, ..., y_{knK})$, de tamanho $n_1, ..., n_k$, $(\Sigma n_s = n)$ de k populações com médias $(\theta_1, ..., \theta_k)$, respectivamente, e variância comum σ^2. Nesse caso, a distribuição *a posteriori* de θ é a t multivariada, uma vez que foram usadas *prioris* conjugadas normal multivariada para o vetor de médias θ e Wishart invertida para a matriz de covariâncias. Assim, a distribuição a *posteriori* resultante é dada por

$$p(\theta\,|\,y) \alpha \left[1 + \frac{\sum_{i=1}^{k} n_i(\theta_i - \bar{y}_i)^2}{\nu s^2}\right]^{-\frac{\nu+k}{2}}, \; -\infty < \theta_i < \infty,\; i = 1, ..., k, \quad (1)$$

com $\bar{y}_i = \frac{1}{n_i}\sum y_{ij}$, $\nu = n - k$, e $s^2 = \frac{1}{n-k}\sum_i\sum_j\left(y_{ij} - \bar{y}_i\right)^2$ que foi usada para se realizar as comparações de médias de k populações normais. Essa distribuição *a posteriori* é a distribuição t multivariada de dimensão k com n graus de liberdade, centrada na média amostral $\bar{y}_i$, conforme detalhamento apresentado na sequência.

A idéia é, a partir da t multivariada, propor um procedimento bayesiano para se realizar as comparações múltiplas das médias.

Geração de uma amostra de tamanho *n* da t multivariada

Seja k populações com médias $\theta = (\theta_1, \theta_2, ..., \theta_k)^t$. Observações $\boldsymbol{Y} = \{Y_1, Y_2, ..., Y_k\}$ são obtidas dessas populações, em que $Y_i = (\boldsymbol{y_{i1}}, \boldsymbol{y_{i2}}, ..., \boldsymbol{y_{in_i}})^t$ é um vetor de observações $n_i \times 1$ condicionalmente independentes no tratamento i, $i = 1, 2, ..., k$ e $\sum_{i=1}^{k} n_i = n$. O problema de comparações múltiplas (PCM) é fazer inferências sobre as relações entre os θ baseados em $\boldsymbol{Y} = \{Y_1, Y_2, ..., Y_k\}$.

Pela equação (1) pode-se observar que a distribuição *a posteriori* do parâmetro θ foi utilizada para se realizar comparações múltiplas de médias normais. Considerando o modelo na forma $y_{ij} = \mu + \tau_i + \varepsilon_{ij}$ e faz-se $\theta = \mu_i = \mu + \tau_i$. A distribuição t multivariada, $t(n, \mu, \Sigma, \nu)$, possui parâmetros que podem ser especificados por:

$$\mu = \bar{y} = \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \vdots \\ \bar{y}_k \end{bmatrix}, \quad \Sigma = S_{\bar{y}} = \begin{bmatrix} \frac{s_1^2}{n_1} & 0 & \cdots & 0 \\ 0 & \frac{s_2^2}{n_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{s_k^2}{n_k} \end{bmatrix}, \quad \nu = n - k$$

em que $n = \sum_{i=1}^{k} n_i$ é o número de médias populacionais e Σ é a matriz de covariâncias das médias.

Os elementos localizados "fora" da diagonal principal indicaram covariâncias nulas entre as populações, sendo considerados independentes, ao passo que o k-ésimo elemento da diagonal de Σ representa a variância da k-ésima população sobre o número de elementos da mesma.

Na sequência foi considerada a forma geral usada para gerar variáveis aleatórias k-dimensionais t multivariada com n graus de liberdade e parâmetros μ e Σ. Seja um vetor aleatório $\boldsymbol{X}$ com distribuição $N_k(\boldsymbol{0}, \Sigma)$ e a variável escalar $\boldsymbol{U}$ com distribuição qui-quadrado com n graus de liberdade, então o vetor aleatório $\boldsymbol{Y}$ dado pela transformação $Y = \sqrt{\nu}\,\frac{X}{\sqrt{U}} + \mu$, possui distribuição t multivariada com ν graus de liberdade.

Foi aplicada a transformação anterior n vezes, a n diferentes vetores aleatórios $\boldsymbol{X}$ e a n variáveis escalares $\boldsymbol{U}$, tendo ao final do processo uma amostra de tamanho n da distribuição t multivariada com n graus de liberdade, utilizando o programa R (R Development Core Team, 2006), como ambiente de programação, em que todos os procedimentos foram implementados.

Distribuição nula da amplitude padronizada

A partir da distribuição *a posteriori* t multivariada, $t(n, \mu, \Sigma, n)$, foram geradas k cadeias de médias μ_i, utilizando o método de Monte Carlo e assumindo médias constantes, vetor $\mu = (\mu, \mu, ..., \mu)^t$, ou seja, todos os componentes iguais. Assim, sem perda de generalidade, foi assumido $\mu = 0$ (para todos os k componentes), impondo a hipótese nula H_0 no método bayesiano.

No próximo passo foi realizada a geração da amplitude padronizada da *posteriori*, sob H_0, e obtida na distribuição *a posteriori* das médias da seguinte forma:

$$q = \frac{máx(\mu_i) - mín(\mu_i)}{\sigma_h} \quad (2)$$

em que σ_h representa a média harmônica das variâncias das k médias, dada por:

$$\sigma_h = \sqrt{\frac{1}{\frac{1}{k}\left(\frac{n_1}{s_1^2} + \frac{n_2}{s_2^2} + \cdots + \frac{n_k}{s_k^2}\right)}} \quad (3)$$

para contemplar a possibilidade de se analisar tanto o caso de variâncias heterogêneas como o caso de variâncias homogêneas.

O fato de se gerar a amplitude padronizada da *posteriori*, sob H_0, obtida da distribuição *a posteriori* das médias da forma mostrada na equação (3) tem um significado muito importante, pois possibilita a análise sob heterogeneidade de variâncias e sob condições de não balanceamento, diferentemente dos testes de comparações múltiplas convencionais.

Calcula-se a seguir a diferença mínima significativa $\Delta = \sigma_h \cdot \theta_{a.}$, em que $\theta_{a.}$ é o quantil superior 100α % da distribuição *a posteriori* de q. Para qualquer amplitude maior que Δ, a diferença é significativamente considerada diferente de zero.

Para descrever os resultados finais das análises do método proposto neste trabalho foram considerados quatro exemplos simulados, sob duas situações. Na primeira situação, foram simulados dados de experimentos sob H_0. Na segunda, foram simulados dados de experimentos sob H_1, sendo que a diferença entre as médias dos níveis do fator foi de dois erros padrões. Em ambos foram consideradas situações balanceadas e homocedásticas e não balanceadas e heterocedásticas.

Foram obtidas as densidades *a posteriori* da distribuição da amplitude padronizada q e os quantis superiores 10%, 5% e 1% dessa distribuição, em cada um dos quatro exemplos utilizados, utilizando-se um estimador de densidades Kernel do programa R (R Development Core Team, 2006).

Para realizar a inferência a respeito da hipótese H_0: $\mu_i = \mu_{i'}$, $i \neq i' = 1, 2, ..., k$, considerando todos os pares e todos os exemplos, foi obtido

$$\Delta = \sigma_h \cdot q_\alpha, \quad (4)$$

para a = 10%, 5% e 1% sendo σ_h obtido a partir da expressão (3).

Da mesma forma foi obtida a média na distribuição *a posteriori* de μ e assim, para cada par $(\mu_i, \mu_{i'})$, foi obtido o intervalo de confiança bayesiano

$$\text{ICB}_{1-\alpha}\,(\mu_i - \mu_{i'}):\ \bar{\mu}_i - \bar{\mu}_{i'} \pm \Delta\,, \quad (5)$$

em que $\bar{\mu}_i$ e $\bar{\mu}_{i'}$ são as médias *posteriori* das cadeias de μ_i e $\mu_{i'}$. Todos os intervalos que contiverem zero indicarão que a hipótese H_0: $\mu_i = \mu_{i'}$, não deve ser rejeitada; se os limites do ICB_{1-a} forem positivos, indicarão que $\mu_i > \mu_{i'}$; se forem negativos, $\mu_i < \mu_{i'}$.

Duas alternativas foram avaliadas. Na primeira, considerando todos os pares $(\mu_i, \mu_{i'})$, foi calculada a seguinte probabilidade *a posteriori*

$$P\,(\mu_i > \mu_{i'}). \quad (6)$$

Esta probabilidade serviu de forma descritiva como uma evidência a favor ou contra H_0.

A segunda alternativa seria obter uma nova cadeia com os limites inferior ($LI^{ii'}$) e superior ($LS^{ii'}$) de um intervalo *a posteriori* para cada par de médias, da seguinte forma

$$\begin{cases} LI^{ii'} = \mu_{ij} - \mu_{i'j} - q_j\ \sigma_h \\ LS^{ii'} = \mu_{ij} - \mu_{i'j} + q_j\ \sigma_h \end{cases} \quad (7)$$

Como medida de evidência a favor ou contra H_0: $\mu_i = \mu_{i'}$, calculou-se a probabilidade *a posteriori* dos intervalos conterem o valor zero - $P(ICB \supset 0)$. Seja $I\,(0 \in [LI_j^{ii'}, LS_j^{ii'}])$, a função indicadora para verificar se o valor 0 pertence ao intervalo na j-ésima unidade amostral de Monte Carlo da cadeia *a posteriori*, logo,

$$P(ICB \supset 0) = \frac{\sum_{j=1}^{n} I\left(0 \in \left[LI_j^{ii'}, LS_j^{ii'}\right]\right)}{n} \quad (8)$$

Uma última análise foi realizada utilizando o HPD nas cadeias *a posteriori* da distribuição das diferenças de duas médias. Para cada cadeia $\mu_i - \mu_{i'}$, foi obtido o HPD utilizando um coeficiente de credibilidade de 95%, utilizando se para isso o pacote BOA, um aplicativo do programa R (R Development Core Team, 2006). Os resultados de todos os procedimentos foram comparados em experimentos simulados.

Foram simulados dois exemplos sob a hipótese nula, ou seja, todas as médias dos níveis do fator foram consideradas iguais. No primeiro caso, considerou-se uma situação balanceada e homocedástica. Foi considerado o valor $k = 6$ e $r = 5$, com $\sigma^2 = 1$. Assim, as observações y_{ij} foram geradas de uma normal N(0, 1), para i = 1, 2,..., k e j = 1, 2,..., r.

No segundo caso, considerou-se que as amostras de cada população seriam geradas de distribuições normais com média 0 e variâncias $\sigma_1^2 = 1$, $\sigma_2^2 = 4$, $\sigma_3^2 = 7$, $\sigma_4^2 = 10$, $\sigma_5^2 = 13$ e $\sigma_6^2 = 16$. Isso foi realizado para que houvesse uma heterogeneidade de variâncias máxima de $\delta = \sigma_6^2 / \sigma_1^2 = 16$. Assim, os valores y_{ij} foram gerados de $N(0, \sigma_i^2)$, considerando os valores das médias dos n_i's dados por 3, 5, 4, 3, 2 e 4, para i = 1, 2, 3, 4, 5 e 6, respectivamente.

Foram simulados dois exemplos sob a hipótese alternativa, sendo o primeiro balanceado e homocedástico e o segundo não balanceado e heterocedástico. Os valores de k, r ou n_i e σ^2 ou σ_i^2 foram os mesmos. A diferença é que as amostras das populações consideradas possuíam diferentes médias. A média da população 1 foi considerada igual a zero ($\mu_i = 0$) sem perda de generalidade. As demais médias μ_i foram definidas por:

$$\mu_{i+1} = \mu_i + 2\,\sigma_h \,, \qquad (9)$$

sendo σ_h a raiz quadrada da média harmônica das variâncias das médias μ_i, dada por:

$$\sigma_h = \sqrt{\frac{1}{\frac{1}{k}\left(\frac{n_1}{\sigma_1^2} + \frac{n_2}{\sigma_2^2} + \cdots + \frac{n_k}{\sigma_k^2}\right)}} \qquad (10)$$

RESULTADOS E DISCUSSÃO

Foram considerados H_0 e H_1 em dois casos, experimentos balanceados e homocedásticos (comportados) e não balanceados e heterocedásticos (não comportados).

A discussão é prejudicada pela escassez de artigos sobre o problema.

As médias originais e *a posteriori* de experimentos simulados sob H_0 e H_1, comportado e não comportado apresentaram resultados similares, indicando que a simulação foi consistente. É conveniente salientar que cada média simulada difere da outra por uma diferença de dois erros padrões, sob H_1. Assim, as diferenças podem ser atribuídas apenas ao erro de Monte Carlo, obtido nas 10.000 simulações realizadas.

As densidades *a posteriori* permitem que se afirme que houve uma grande flexibilidade do método, indicado pelo ajuste ao conjunto de dados e as suas características. Assim, diferentemente do caso clássico, os quantis superiores da amplitude padronizada modificam-se em função do não balanceamento e da heterogeneidade de variâncias que porventura possam existir nos conjuntos de dados submetidos à análise.

Na Tabela 1 são apresentados os quantis superiores 10%, 5% e 1% da distribuição da amplitude padronizada q. Os valores dos quantis da amplitude padronizada são muito similares aos da amplitude *estudentizada q* do teste clássico de comparações múltiplas de Tukey para médias de 6 fatores com 24 graus de liberdade, cujos valores são 3,90, 4,37 e 5,37 para 10%, 5% e 1%, respectivamente, principalmente com os do caso comportado. Os valores observados sob H_1 comportado e não comportado indicam uma pequena diferença, sendo um pouco maiores nesse último caso.

Considerando as diferenças entre todos os pares de médias e comparando com o Δ, obtido a partir da equação (4), verifica-se que realmente não há diferença significativa entre as médias dos fatores nas situações sob H_0 comportado e não comportado. Esse resultado é muito interessante, pois ambas as situações foram simuladas sob a hipótese nula em que as médias dos fatores são iguais e com isso não se esperava detectar diferenças entre as médias. Verifica-se também que há diferenças significativas entre as médias dos fatores nas situações sob H_1 comportado e não comportado, mas que muitas diferenças foram não significativas, indicando que se cometeu o erro tipo II. A heterogeneidade de variâncias afetou consideravelmente o poder do teste. As diferenças mínimas significativas foram muito discrepantes, sendo uma delas mais de

Tabela 1 – Quantis superiores da amplitude padronizada q da distribuição a posteriori.

$1 - \alpha$	Sob H_0		Sob H_1	
	Comportado	Não Comportado	Comportado	Não Comportado
90%	3,877603	4,278497	3,933255	4,286779
95%	4,373322	4,849312	4,421675	4,873247
99%	5,389085	6,114813	5,374668	6,203168

quatro vezes superior à outra. Isso é um indicativo que o teste quando aplicado em situações heterocedásticas, deve ser menos poderoso.

Observando os intervalos de confiança bayesianos dos testes de comparações múltiplas, utilizando os quantis 95% da distribuição da amplitude padronizada *a posteriori* q sob H_0 comportado e não comportado, verifica-se que todos os intervalos englobaram o zero, que é um indicativo de que a hipótese H_0: $\mu_i = \mu_{i'}$, não deva ser rejeitada. Esse procedimento é alternativo ao procedimento anterior, considerando a amplitude padronizada *a posteriori*. Outro fato interessante a ser destacado é que os intervalos de confiança bayesianos para a situação não comportada são mais amplos. Isso é decorrente da maior diferença mínima bayesiana nesse caso, em consequência da maior variabilidade no caso heterocedástico.

Assim, o procedimento é adequado tanto para situações comportadas como para situações em que as condições não seriam favoráveis à aplicação de um teste clássico de comparações múltiplas. Em ambos os casos, no entanto, a suposição de normalidade não foi violada. Sob H_1 comportado e não comportado, intervalos contiveram o valor zero, que é um indicativo de que está se cometendo o erro tipo II. Cada média populacional, nesse caso, difere da sua vizinha por dois erros padrões. Médias vizinhas de um salto igual a 2, possuem valores distantes em quatro erros padrões e assim sucessivamente. Assim, médias mais distantes tendem a ser diferenciadas. As oscilações nos valores das médias experimentais são devidas aos maiores erros associados à situação denominada não comportada.

As probabilidades *a posteriori* $P(\mu_i > \mu_{i'})$, considerando todos os pares $(\mu_i, \mu_{i'})$ da distribuição conjunta *a posteriori*, são apresentadas na Tabela 2. Quando a probabilidade $P(\mu_i > \mu_{i'}) \leq 0{,}025$ ou $P(\mu_i > \mu_{i'}) \geq 0{,}975$ tem-se uma evidência contra H_0, caso $0{,}025 < P(\mu_i > \mu_{i'}) < 0{,}975$, tem-se uma evidência a favor de H_0. Essa evidência foi considerada neste estudo por meio da escolha de 95% de confiança, para termos uma situação comparável ao normalmente aplicado aos testes clássicos de comparações múltiplas. Verifica-se que, sob H_0 comportado, todas as probabilidades evidenciam que a hipótese H_0 não deve ser rejeitada. Contudo, sob H_0 não comportado, há uma situação de evidência contra H_0, parâmetros μ_4 e μ_6, que indicam que o procedimento é mais liberal, com mais chance de cometer erros. Verifica-se também que, sob H_1 comportado, há 4 (quatro) casos evidenciando que a hipótese H_0 não deve ser rejeitada, e que, sob H_1 não comportado, há 5 (cinco) casos. Esses resultados, quando comparados com aqueles que utilizaram

Tabela 2 – Probabilidades a posteriori $P(\mu_i > \mu_{i'})$, em porcentagem, considerando todos os pares $(\mu_i, \mu_{i'})$ distribuição conjunta a posteriori.

Parâmetros	Sob H_0		Sob H_1	
	Comportado	Não Comportado	Comportado	Não Comportado
μ_1 e μ_2	68,77	58,56	22,29	32,50
μ_1 e μ_3	80,41	75,15	6,41	1,89
μ_1 e μ_4	40,65	92,84	0,00	0,03
μ_1 e μ_5	55,21	37,54	0,00	15,73
μ_1 e μ_6	74,08	21,90	0,00	0,01
μ_2 e μ_3	65,21	70,24	21,88	2,35
μ_2 e μ_4	23,79	92,23	0,06	0,00
μ_2 e μ_5	36,12	29,55	0,00	24,12
μ_2 e μ_6	56,75	13,52	0,00	0,00
μ_3 e μ_4	13,74	82,42	0,53	0,86
μ_3 e μ_5	23,38	17,95	0,00	76,70
μ_3 e μ_6	41,31	6,35	0,00	0,12
μ_4 e μ_5	64,12	5,92	1,89	99,29
μ_4 e μ_6	81,04	1,45	0,13	26,51
μ_5 e μ_6	69,74	35,79	8,76	0,08

a distribuição *a posteriori* da amplitude padronizada permitem que se infira que este método é mais poderoso do que o anterior. No entanto, sob a hipótese nula, principalmente sob situações não comportadas, esse procedimento foi considerado mais liberal. Assim, a vantagem em relação ao maior poder deve ser olhada com ressalva. Esse procedimento é o que os bayesianos normalmente fazem, ou seja, tomam suas decisões baseadas na distribuição *a posteriori* conjunta dos parâmetros e numa medida probabilística a favor de uma hipótese ou de outra.

Assim, este procedimento, aplicado para fins de comparação, não é equivalente ao da proposta de trabalho que considera a distribuição *a posteriori* da amplitude padronizada para construir o critério de decisão. O teste de Tukey seria a alternativa equivalente ao procedimento bayesiano que considera a amplitude *a posteriori* padronizada e que controlaria as taxas de erro por experimento. Portanto, a maior liberalidade dessa alternativa seria esperada, como acontece com os seus pares clássicos.

As probabilidades *a posteriori* dos intervalos de confiança conter o valor zero - $P(ICB \supset 0)$, são apresentadas na Tabela 3. A diferença entre esse método e o anterior é que as cadeias das diferenças de médias consideram os valores de Δ determinados em cada ponto amostral dessa cadeia. Assim, espera-se uma maior proteção do erro tipo I quando se utiliza tal procedimento. Quando a probabilidade for menor ou igual a 5% tem-se uma evidência contra H_0, caso contrário, a favor de H_0.

Verifica-se que, sob H_0, todas as probabilidades constituíram-se em uma forte evidência a favor de H_0. Esse procedimento é equivalente ao intervalo de confiança bayesiano usando o Δ e a diferença entre eles é que o intervalo bayesiano é obtido como um sumário da cadeia da diferença entre duas médias e a obtenção das probabilidades, envolve o cálculo do intervalo em cada ponto simulado da cadeia *a posteriori*. Verifica-se também que, sob H_1 existem probabilidades que evidenciam uma decisão a favor de H_0, ou seja, indicativo de se estar cometendo erro do tipo II. Esse procedimento apresentou piores resultados do que os obtidos utilizando o intervalo baseado na média *a posteriori* das cadeias da diferença das médias e com a amplitude padronizada, pois houve mais erros do tipo II no presente caso.

Os HPD's de 95% para cada cadeia $\mu_i = \mu_{i'}$ são apresentados na Tabela 4. Novamente é interessante que se saliente que esse é um procedimento que seria genuinamente adotado pelos estatísticos bayesianos.

Esse procedimento deve ser melhor do que os procedimentos que utilizam a distribuição *a posteriori* das diferenças de médias apresentados anteriormente, mas não contorna o problema do controle do erro tipo I global nas diferentes análises realizadas simultaneamente.

Tabela 3 – Probabilidades, em porcentagem, do intervalo de confiança bayesiano (56) a posteriori conter o valor zero.

Parâmetros	Sob H_0		Sob H_1	
	Comportado	Não Comportado	Comportado	Não Comportado
μ_1 e μ_2	84,28	87,92	78,90	86,29
μ_1 e μ_3	77,63	81,19	58,13	44,61
μ_1 e μ_4	86,40	57,03	2,89	2,82
μ_1 e μ_5	86,74	80,27	0,02	67,23
μ_1 e μ_6	81,35	78,52	0,00	0,88
μ_2 e μ_3	85,61	86,83	79,24	59,60
μ_2 e μ_4	79,95	63,98	8,74	4,33
μ_2 e μ_5	85,43	80,44	0,16	78,00
μ_2 e μ_6	86,53	76,48	0,01	1,34
μ_3 e μ_4	71,73	76,06	22,49	28,68
μ_3 e μ_5	79,28	70,60	0,92	76,40
μ_3 e μ_6	86,39	62,41	0,10	13,74
μ_4 e μ_5	85,26	45,61	38,68	14,62
μ_4 e μ_6	76,87	33,07	8,77	82,51
μ_5 e μ_6	83,76	80,92	64,70	6,85

Tabela 4 – HPD's de 95% para cada cadeia μ_i - $\mu_{i'}$.

Parâmetros	Sob H_0				Sob H_1			
	Comportado		Não Comportado		Comportado		Não Comportado	
	LIC	LSC	LIC	LSC	LIC	LSC	LIC	LSC
μ_1 e μ_2	-0,81	1,341	-2,78	3,213	-1,36	0,620	-4,56	2,917
μ_1 e μ_3	-0,57	1,552	-1,95	4,092	-1,72	0,222	-7,67	0,117
μ_1 e μ_4	-1,16	0,978	-0,77	5,801	-3,01	-1,05	-12,7	-4,46
μ_1 e μ_5	-0,97	1,160	-4,39	3,002	-4,08	-2,12	-7,10	2,158
μ_1 e μ_6	-0,77	1,387	-4,17	1,954	-4,75	-2,80	-13,8	-6,05
μ_2 e μ_3	-0,87	1,249	-1,92	3,497	-1,36	0,581	-6,54	0,330
μ_2 e μ_4	-1,43	0,673	-0,86	5,139	-2,63	-0,65	-11,5	-4,01
μ_2 e μ_5	-1,26	0,889	-4,19	2,559	-3,66	-1,73	-5,72	2,929
μ_2 e μ_6	-0,92	1,222	-4,24	1,176	-4,39	-2,42	-12,5	-5,64
μ_3 e μ_4	-1,63	0,506	-1,75	4,415	-2,33	-0,37	-8,67	-0,81
μ_3 e μ_5	-1,41	0,735	-5,02	1,980	-3,35	-1,38	-2,69	6,120
μ_3 e μ_6	-1,20	0,934	-5,17	0,627	-3,96	-1,99	-9,70	-2,51
μ_4 e μ_5	-0,89	1,248	-6,86	0,627	-2,02	-0,09	1,723	11,13
μ_4 e μ_6	-0,54	1,588	-6,65	-0,39	-2,70	-0,74	-5,11	2,546
μ_5 e μ_6	-0,80	1,334	-4,16	2,916	-1,62	0,298	-11,8	-3,09

Verifica-se que esse procedimento foi equivalente ao das probabilidades *a posteriori* P($\mu_i > \mu_{i'}$), nesse caso. Espera-se que resultados similares entre esses dois procedimentos sejam observados com muita frequência em situações práticas reais, mas não há garantias de equivalência teórica entre esses dois métodos. O que, no entanto, fica claro é que ambos os métodos não possuem garantias de controle do erro tipo I simultaneamente nas múltiplas inferências realizadas nas cadeias *a posteriori* das diferenças entre médias. Assim, observando-se o erro tipo I no caso não comportado, tem-se uma evidência de que esse procedimento seja mais liberal que os que utilizam a distribuição *a posteriori* da amplitude padronizada.

CONCLUSÕES

O procedimento que encontra os intervalos de confiança ICB$_{1-\alpha}$ (μ_i - $\mu_{i'}$) usando a amplitude padronizada Δ é mais conservativo e flexível a situações de heterocedasticidade. Apresenta resultados praticamente equivalentes ao método que calcula as probabilidades do intervalo *a posteriori* conter o zero - $P(ICB \supset 0)$.

O procedimento que calcula as probabilidades *a posteriori* P($\mu_i > \mu_{i'}$) e os HPD's, foram equivalentes entre si. São também mais liberais, identificando diferenças em situações em que elas não existem.

Os procedimentos propostos apresentam a desvantagem de ainda não estarem implementados. Em compensação apresentam vantagens em relação aos testes convencionais, no sentido de não haver necessidade de balanceamento dos dados, que é muito significativo do ponto de vista prático, e de poderem ser utilizados em situações homo e heterocedásticas.

Os procedimentos baseados na amplitude padronizada foram superiores aos demais procedimentos estudados por terem controlado, nos exemplos simulados, o erro tipo I e detectadas a maior parte das diferenças sob H_1.

Os procedimentos de comparações múltiplas bayesianos foram propostos com sucesso para situações de normalidade, com ou sem homogeneidade de variâncias e com ou sem balanceamento.

REFERÊNCIAS BIBLIOGRÁFICAS

BERRY, D.A. Multiple comparisons, multiple tests, and data dredging: a Bayesian perspective: with discussion. In: BERNARDO, J.M.; DEGROOT, M.H.; LINDLEY, D.V.; SMITH, A.F.M. **Bayesian statistics.** Oxford: Oxford University, 1988. v.3, p.79-94.

BERRY, D.A.; HOCHBERG, Y. Bayesian perspectives on multiple comparisons. **Journal of Statistical Planning and Inference,** Amsterdam, v.82, p.215-227, 1999.

BRATCHER, T.; HAMILTON, C. A Bayesian multiple comparison procedure for ranking the means of normally distributed data. **Journal of Statistical Planning and Inference,** Amsterdam, v.133, p.23-32, 2005.

CHO, J.S.; MASOOM, M.A.; BEGUM, M. Bayesian multiple comparisons with nonparametric Dirichlet process prioris for negative binomial populations. **Pakinson Journal Statistical,** v.22, n.2, p.89-98, 2006.

DUNCAN, D.B. A Bayesian approach to multiple comparisons. **Technometrics**, Washington, v.7, p.171-222, 1965.

GOPALAN, R.; BERRY, D.A. Bayesian multiple comparisons using Dirichlet process priors. **Journal American Statistical Association,** Washington, v.93, n.443, p.1130-1139, 1998.

HOCHBERG, Y.; TAMHANE, A.C.T. **Multiple comparison procedures**. New York: J.Wiley, 1987.

HSU, J.C. **Multiple comparisons**: theory and methods. London: CRC, 1996.

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. **Anais...** Londrina: ISBN, 2005.

R DEVELOPMENT CORE TEAM. **R**: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2005. Disponível em: <http://www.R-project.org>. Acesso em: 10 jun. 2006.

SHAFFER, P.J. A semi-Bayesian study of Duncan's Bayesian multiple comparison procedure. **Journal of Statistical Planning and Inference**, Amsterdam, v.82, p.197-213, 1999.

WALLER, R.A.; DUNCAN, D.B. A Bayes rule for the symmetric multiple comparisons problem. **Journal of the American Statistical Association**, Washington, v.64, n.328, p.1484-1503, 1969.