

ERRO ADEQUADO PARA COMPARAÇÃO DE MÉDIAS EM CASOS DE HETEROGENEIDADE DAS VARIÂNCIAS ⁽¹⁾

TOSHIO IGUE ⁽²⁾, ARMANDO CONAGIN ⁽³⁾, VIOLETA NAGAI ⁽²⁾ e LUÍS A. AMBRÓSIO ⁽²⁾

RESUMO

Para verificar se a pequena influência da heterogeneidade das variâncias, observada em trabalho anterior, sobre os testes de Bonferroni, "t" e Tukey se devia ao fato de ter sido utilizado o erro médio que corresponde ao quadrado médio da interação tratamentos x blocos (Q.M. trat x bloco), os autores estudaram o comportamento do teste "t" com o uso de três tipos de erro: Q.M. trat x bloco, erro do contraste específico e erro ponderado dos contrastes. Foram simulados 200 experimentos, em grupos de 50, para cada uma das 56 combinações de quatro coeficientes de variação, dois efeitos de tratamentos e sete grupos de heterogeneidade, totalizando 11.200 experimentos. Foi feita a análise da variância de cada ensaio e estudados os seguintes contrastes: $m_1 - m_3$ e $m_2 - m_3$, onde m_3 representa a testemunha. "A priori" foi estabelecido que o tratamento 1 seria superior à testemunha nas quantidades de 8 e 17% (para os efeitos 1 e 2 respectivamente) e o tratamento 2, nas quantidades de 5 e 12% (para os mesmos efeitos). Assim, foi empregado o teste t, unilateral, para comparação das médias nos dois contrastes. Na presente pesquisa, são válidas as seguintes conclusões: as diferenças nas porcentagens de rejeição com os três tipos de erro praticamente inexistem; as relações de variâncias com valores maiores que 6 podem ser consideradas heterogêneas, quando utilizado o erro do contraste; a porcentagem de rejeição obtida é indiretamente proporcional ao coeficiente de variação e diretamente proporcional à magnitude dos contrastes, mesmo em condições de heterogeneidade de erros.

Termos de indexação: heterogeneidade de variâncias, erro para comparação de médias, teste t.

ABSTRACT

APPROPRIATE TEST ERROR TO COMPARE MEANS IN CASE OF HETEROGENEITY OF VARIANCES

It was observed previously a small effect of heterogeneous variances on the Bonferroni's test, Student "t" test and Tukey's test. In this paper, it was tested whether such observation was due to the fact that the error used for the tests was the treat. x block interaction mean square besides the existence of certain heterogeneous

⁽¹⁾ Trabalho apresentado no 4º Simpósio de Estatística aplicada à Experimentação Agronômica e 36ª Reunião Anual da Região Brasileira de Biometria, realizados em Goiânia (GO), de 15 a 19 de julho de 1991. Parcialmente financiado pela FAPESP e pelo CNPq. Recebido para publicação em 25 de maio e aceito em 8 de setembro de 1993.

⁽²⁾ Seção de Técnica Experimental e Cálculo, Instituto Agronômico (IAC), Caixa Postal 28, 13001-970 Campinas (SP).

variances. The authors considered the existence of the following situations: a) two groups of variance in each experiment, and b) three kinds of errors for "t" test: treat. x block interaction mean square, specific contrast error for each comparison; and, pooled contrast error. It was simulated two hundred experiments, in groups of 50, for each of 56 combinations of four coefficients of variation, two treatment effects and seven groups of heterogeneous variances, performing a total of 11,200 simulations. In the analysis of variance, for each experiment, the following contrasts between means were considered: $m_1 - m_3$ and $m_2 - m_3$, where m_3 represents the control treatments mean. The two treatment means m_1 and m_2 were considered "a priori" more productive than the control by 8% and 5% (effect 1) and by 17% and 12% (effect 2), respectively for m_1 and m_2 . Thus, the one tail "t" test was applied. From the results observed, it was concluded that: a) non significant differences occurred concerning the percentages of rejections of the null hypothesis using any of the three errors considered; b) the relation among variances greater than 6 was statistically heterogeneous, when the error used was that of specific contrast error; and c) the percentages of rejections obtained were indirectly proportional to the coefficients of variation and directly proportional to the range values of contrasts between means, even in case of heterogeneous error conditions.

Index terms: heterogeneity of variances, error for means comparisons, t test.

1. INTRODUÇÃO

A homogeneidade das variâncias é uma das hipóteses para validade da análise e para assegurar o nível de significância dos testes estatísticos. Cochran (1947), em razão de heterogeneidade das variâncias em dados experimentais de um ensaio de calagem, dividiu o erro em componentes homogêneos, concluindo que a heterogeneidade pode afetar certos tratamentos ou parte dos dados a uma extensão imprevisível. Em análise conjunta, na aplicação dos testes estatísticos, normalmente é utilizado o erro ponderado dos experimentos.

Neter & Wasserman (1974) comentam que, quando as variâncias que compõem o erro são diferentes, mas as amostras são de mesmo tamanho, o teste F, para igualdade de médias, num modelo de efeito fixo, é pouco afetado. Variâncias diferentes alteram o nível de significância levemente acima do especificado. De modo semelhante, o teste de comparação múltipla de Scheffé, baseado na distribuição F, também não é muito prejudicado por variâncias diferentes se os tamanhos das amostras forem iguais. Comparações simples entre dois fatores podem, no entanto, ser afetadas de tal forma pela heterogeneidade das variâncias que o intervalo de confiança real e o estimado podem ser bem diferentes. Em modelos aleatórios,

mesmo em amostras de igual tamanho, a heterogeneidade das variâncias pode ter efeito pronunciado nas inferências sobre os componentes de variância. Conforme Cochran & Cox (1957), à semelhança dos tratamentos, a soma dos quadrados do erro pode ser partida em componentes. Embora tais subdivisões não sejam frequentes, como no caso dos tratamentos, apresentam inúmeras utilidades. Às vezes, há razões para acreditar que o quadrado médio do erro não seja homogêneo, isto é, o quadrado médio residual não tem a mesma variância, σ^2 , como foi postulado no modelo matemático. Nesse caso, a subdivisão do erro se faz necessária para validade do teste "t". Segundo Little & Hills (1978), a heterogeneidade das variâncias pode ser contornada, em certas condições, utilizando transformações de dados. Uma questão importante é decidir sobre que tipo de transformação deve ser utilizada. Conagin et al. (1990) mostraram que a heterogeneidade das variâncias tem pouca influência nos testes de comparação de médias (Bonferroni, "t" e Tukey) quando se emprega o erro ponderado (interação trat x bloco) da análise da variância. Para verificar se esse resultado se devia ao fato de ter sido usado o erro médio, ou seja, o quadrado médio da interação tratamentos x blocos (Q.M. trat x bloco), os autores estudaram o comportamento

do teste "t" com o emprego de três tipos de erros: Q.M. trat x bloco, erro do contraste específico estudado e erro ponderado dos contrastes.

O presente trabalho pretende verificar, dos três tipos de erro, o mais adequado para uso do teste "t", quando nos experimentos existem dois grupos de variância.

2. MATERIAL E MÉTODO

Foram simulados 200 experimentos, em grupos de 50, para cada uma das 56 combinações de quatro coeficientes de variação (CV); dois efeitos de tratamentos (Ef) e sete relações de variância (G), com os seguintes valores:

CV: 10, 15, 20 e 25%;

Ef-(Ef1): 120, 75, 0, -60, -135;

(Ef2): 180, 100, -80, -90, -110;

G: 1,00; 2,04; 4,00; 6,25; 9,00; 16,00 e 22,56.

No total, foram simulados 11.200 experimentos, sempre em condições de normalidade dos erros experimentais e considerando uma média de 1.500 kg/ha.

Foram considerados sete grupos (G) de heterogeneidade, obtidos multiplicando-se o valor aleatório do erro, no modelo para delineamento em blocos ao acaso, por coeficientes escolhidos, de forma a obter as sete relações de variância mencionadas.

Dentro de cada grupo, os tratamentos 1 e 2 tinham variâncias iguais e maiores que os tratamentos 3, 4 e 5, também com variâncias iguais. Assim, cada grupo estudado tinha dois conjuntos de tratamentos com variâncias diferentes.

Foi efetuada a análise da variância de cada ensaio e estudados os seguintes contrastes: m1 - m3 e m2 - m3, onde m3 representa o valor médio da testemunha em qualquer dos grupos considerados. "A priori", foi estabelecido que o tratamento 1 seria superior à testemunha nas quantidades de 8 e 17% (para os efeitos 1 e 2 respectivamente) e o tratamento 2, nas quantidades de 5 e 12% (para os mesmos efeitos) empregado o teste t, unilateral, para comparação das médias nos dois contrastes.

Para emprego do teste, consideraram-se os seguintes tipos de erro: o erro médio (Ep), erro

Quadro 1. Grupos de heterogeneidade de variância (G_i) definidos pelos coeficientes aplicados aos diferentes tratamentos (T_i) e contrastes (C_i) entre os grupos de heterogeneidade

Tratamentos e Contrastes	Grupos (G _i)						
	G1 1,00	G2 2,04	G3 4,00	G4 6,25	G5 9,00	G6 16,00	G7 22,56
T _i							
T1	1	1,15	1,35	1,42	1,48	1,50	1,50
T2	1	1,15	1,35	1,42	1,48	1,50	1,50
T3	1	0,84	0,67	0,57	0,49	0,37	0,32
T4	1	0,84	0,67	0,57	0,49	0,37	0,32
T5	1	0,84	0,67	0,57	0,49	0,37	0,32
C _i							
C1	-4	-4	-4	3	3	3	3
C2	-1	1	0	0	0	0	0
C3	-1	-1	2	0	0	0	0
C4	0	0	0	-1	1	0	0
C5	0	0	0	0	0	-1	1

da diferença entre duas médias (Ed) e o erro ponderado pelas variâncias dos grupos de médias em estudo (Ew), ou seja, a variância considerando os dois primeiros tratamentos, com 7 graus de liberdade, e a correspondente às três últimas médias, com 14 graus de liberdade. O erro médio (Ep) corresponde à interação tratamentos x blocos na análise da variância; o erro da diferença (Ed) foi obtido considerando a diferença entre os valores x_{ij} para $i = 1$ e 2 e $j = 1, 2, \dots, 8$. O valor de t , quando usado o erro ponderado (Ew) foi calculado, de acordo com Cochran & Cox (1957), por $t = (w_1 t_1 + w_2 t_2) / (w_1 + w_2)$, sendo t_1 e t_2 os valores críticos do teste t , com n_1 e n_2 graus de liberdade correspondentes a s_1^2 e s_2^2 , ou seja, os valores críticos de t são ponderados com as respectivas variâncias das médias. Os valores críticos de t foram, respectivamente, 1,9; 1,81 e 1,76 para 7, 10 e 14 graus de liberdade; $w_1 = s_1^2/n_1$ e $w_2 = s_2^2/n_2$.

Para cada conjunto de 50 experimentos, foi calculada a porcentagem de rejeição da hipótese de nulidade nos dois contrastes considerados, utilizando os três tipos de erro; foram calculadas também as porcentagens de rejeição de variâncias iguais entre grupos, ou seja, a variância referente ao grupo com os tratamentos 1 e 2 em relação ao grupo com os tratamentos 3, 4 e 5. Sobre essas porcentagens de rejeição, foi feita análise da variância, após transformação em arco seno $(p)^{1/2}$, considerando como fontes de variação o coeficiente de variação, os efeitos de tratamentos e de grupos de heterogeneidade e as interações respectivas. Os grupos foram estudados considerando os contrastes apresentados no quadro 1.

3. RESULTADOS E DISCUSSÃO

No quadro 2, observa-se que as magnitudes dos contrastes entre médias (D), correspondentes às diferenças percentuais de 5, 8, 12 e 17%, bem como a precisão do experimento (CV), influíram significativamente nas porcentagens de rejeição obtidas com o uso dos três tipos de erro.

As porcentagens de rejeição cresceram com a magnitude das diferenças e diminuíram significativamente com o aumento do coeficiente de variação.

O fator grupo de heterogeneidade (G) só foi significativo quando utilizado o "erro da diferença" (Ed). Esse fator foi, também, significativo na rejeição de hipóteses da variância (V) iguais.

A interação D x CV foi significativa com qualquer dos erros utilizados. Isso mostra que as porcentagens de rejeição devidas à magnitude dos contrastes, D, em cada um dos níveis de CV, não foram de mesma proporção para os três grupos. Os componentes linear e quadrático dos fatores D e CV (Dl, Dq, CVI e CVq), bem como a interação D x CV, foram significativos com o uso dos três tipos de erro. Somente no caso em que se considerou o "erro da diferença" (Ed), o contraste C1 (grupos 1, 2 e 3 vs. grupos 4, 5, 6 e 7) foi significativo, enquanto o contraste C2 (G1 vs. G2) foi significativo quando utilizado o erro ponderado (Ew). Todos os seis contrastes entre grupo de heterogeneidade foram significativos na rejeição de hipótese de variâncias iguais, e pode ser observado pelo resultado da análise da variância na coluna V - Quadro 2.

Em condições de homogeneidade, isto é, $G = 1,0$, as porcentagens de rejeição foram de 38,8; 34,4 e 36,4%, correspondentes, respectivamente, ao erro médio com 28 graus de liberdade (GL), erro da diferença com 7 GL e erro ponderado com 14 GL. Essas porcentagens permaneceram em níveis semelhantes, conservando a mesma ordem para os níveis de heterogeneidade de 2,04; 4,00; 6,25; 9,00; 16,00 e 22,56. Essas pequenas diferenças numéricas podem ser atribuídas às precisões das estimativas dos erros, que estão relacionados aos diferentes graus de liberdade. No caso do erro da diferença (Ed), relações de variância acima de 6 podem ser consideradas como heterogêneas, pois o contraste C1 foi significativo e os demais não o foram. Pimentel Gomes (1985) considera como homogêneas, para fins de análise de grupos de experimentos, variâncias cujo quociente entre o maior e o menor valor do quadrado médio residual seja menor que 7.

Quadro 2. Porcentagens médias de rejeição referentes aos contrastes (Y_p , Y_d , Y_w) quando foram empregados três tipos de erro (E_p , E_d e E_w) e de homogeneidade de variância (V), considerando três fontes de variação: magnitude de diferenças entre médias (D), coeficiente de variação (CV) e relação de heterogeneidade de variância (G)

D	CV	G	Y_p	Y_d	Y_w	V
75	-	-	18,0	13,9	14,5	68,8
120	-	-	27,7	21,8	23,8	68,8
180	-	-	44,7	36,8	38,1	67,6
260	-	-	64,7	55,6	59,6	67,6
-	10	-	59,4	51,0	53,6	67,8
-	15	-	40,2	33,6	35,6	67,8
-	20	-	30,7	24,7	26,2	68,2
-	25	-	24,8	18,8	20,6	69,1
-	-	1,00	38,8	34,4	36,4	5,6
-	-	2,05	37,8	33,5	33,6	28,4
-	-	4,00	37,5	32,2	33,4	66,2
-	-	6,25	38,4	30,4	33,0	85,8
-	-	9,00	39,4	31,7	33,6	93,8
-	-	16,00	39,4	30,9	34,2	98,2
-	-	22,56	40,2	31,1	33,9	99,3
D			*	*	*	
CV			*	*	*	
G				*		*
D x CV			*	*	*	
D x G						
CV x G						
D1			*	*	*	*
Dq			*	*	*	
CV1			*	*	*	
CVq			*	*	*	
D x CV			*	*	*	
C1				*		*
C2					*	*
C3						*
C4						*
C5						*
C6						*
Média			38,2	33,7	35,1	59,0
CV%			7,4	6,4	6,9	4,1

* F significativo ($P \leq 0,05$).

No quadro 3, observa-se grande efeito do coeficiente de variação (CV) sobre a rejeição da hipótese de nulidade dos contrastes $m_1 - m_3$ e $m_2 - m_3$ para os dois efeitos de tratamentos.

Considerando a porcentagem de rejeição Y_w , as médias foram de 53,9; 35,5; 27,7 e 20,6%, respectivamente, para os coeficientes de variação de 10, 15, 20 e 25%. Nota-se que as diferenças entre as porcentagens de rejeição para os grupos de heterogeneidade G1, G4 e G7, dentro de cada CV e para os tipos de erro E_p , E_d e E_w , foram pequenas.

Quadro 3. Médias de porcentagens de contrastes (Y_p , Y_d , Y_w) significativos quando foram utilizados o erro médio (E_p), o erro da diferença (E_d) e o erro ponderado (E_w) e de rejeição de homogeneidade das variâncias (V), nos três grupos de relações de variância, G1, G4 e G7, e quatro valores de CV%

Grupos	Rejeição			
	Y_p	Y_d	Y_w	V
CV = 10%				
G1	60,8	53,8	56,8	5,5
G4	57,9	48,4	51,2	87,5
G7	59,7	49,1	53,8	99,2
Média	59,5	50,0	53,9	64,1
CV = 15%				
G1	39,2	35,2	37,0	4,5
G4	41,4	33,0	35,4	84,8
G7	42,2	32,8	34,1	99,5
Média	41,0	33,7	35,5	62,9
CV = 20%				
G1	32,0	29,5	31,1	6,0
G4	28,8	22,4	23,9	84,0
G7	34,2	25,2	28,1	99,5
Média	31,7	25,8	27,7	64,2
CV = 25%				
G1	23,2	19,0	20,9	6,2
G4	25,5	19,0	21,5	84,0
G7	24,5	17,1	19,5	99,0
Média	24,4	18,4	20,6	63,1

4. CONCLUSÕES

1. As diferenças de porcentagens de rejeição de hipótese de nulidade dos contrastes entre médias, quando o teste foi feito utilizando o erro médio, erro da diferença ou erro ponderado, foram pequenas e os resultados, proporcionais aos números de graus de liberdade dos erros respectivos.

2. Utilizando o erro da diferença, uma relação de variâncias maior que 6 pode ser considerada heterogênea.

3. A precisão do experimento, dada pelo coeficiente de variação, teve grande influência na porcentagem de rejeição de contrastes, mesmo em condições de heterogeneidade das variâncias.

4. A porcentagem de rejeição de hipótese de nulidade de contraste entre médias foi diretamente proporcional à magnitude desse contraste, mesmo em condições de heterogeneidade dos erros.

REFERÊNCIAS BIBLIOGRÁFICAS

- COCHRAN, W.G. Some consequences when the assumptions for the analysis of variance are not satisfied. *Biometrics*, Washington, D.C., 3(1):22-38, 1947.
- COCHRAN, W.G. & COX, G.M. *Experimental designs*. 2.ed. London, John Wiley, 1957. 611p.
- CONAGIN, A.; NAGAI, V. & IGUE, T. Poder discriminativo de diferentes testes de comparação de médias. *Revista de Agricultura*, Piracicaba, 65(2): 203-214, 1990.
- LITTLE, T.M. & HILLS, F.J. *Agricultural experimentation: design and analysis*. Nova York, John Wiley, 1978. 350p.
- NETER, J. & WASSERMAN, W. *Applied linear statistical models: regression, analysis of variance, and experimental design*. Georgetown, Richard D. Irwin, 1974. 842p.
- PIMENTEL GOMES, F. *Curso de estatística experimental*. 11.ed. São Paulo, Nobel, 1985. 466p.