

NÚMERO DE REPETIÇÕES EM EXPERIMENTOS E SEU EFEITO NA SENSIBILIDADE
DOS TESTES TUKEY, DUNCAN E DMS.

Replication Number in Experiments and its Effect in Tukey, Duncan
and LSD Sensibility.

Ione Aydêe B. Pignataro*

RESUMO

Foi estudada a influência do número de repetições na probabilidade de ocorrência de erros tipo I e II na comparação múltipla de médias, usando técnica de simulação com computador. O número de repetições variou de 3 a 8 e os testes empregados foram Tukey, Duncan e DMS. A simulação foi feita com experimentos de 6, 10, 15 e 20 tratamentos.

O número de repetições não influenciou na probabilidade de ocorrência do erro tipo I, mas modificou a sensibilidade dos testes empregados. As diferenças relativas verdadeiras mínimas que os testes foram capazes de detectar com 95% de acerto foram $2,5\sigma$ para o Tukey, variando o número de repetições necessárias com o número de tratamentos usados; $2,0\sigma$ para o Duncan, sendo necessárias 8 repetições, e $2,0\sigma$ para o DMS, sendo necessárias 6 repetições.

UNITERMOS: Número de repetições em experimentos, testes de comparação múltipla de médias, sensibilidade.

SUMMARY

The influence of replication number in the type I and II error rates in pairwise multiple comparison of treatment means was studied using computer simulation techniques.

The replication number of treatments ranged between 3 to 8 and the procedures used were Tukey, Duncan and LSD. Simulation was carried out with 6, 10, 15 and 20 treatments.

The replication number in experiments does not have influence in type I error rate changed the test sensitivity in detecting difference among means. The minimum relative difference that were able detected with security of 95% was 2.5σ to Tukey procedure, changing the neces-

* Professora Adjunto, Departamento de Fitotecnia, C.C.Rurais, Universidade Federal de Santa Maria. 97.100 - Santa Maria, RS.

sary replication number with the treatment number, 2.0σ to Duncan procedure being necessary use 8 replications and 2.0σ to LSD procedure being necessary 6 replications.

KEY WORDS: Experimental replication number, mean multiple range tests, sensibility.

INTRODUÇÃO

É um problema difícil determinar qual o melhor número de repetições para um determinado experimento. Segundo GOMES (11), numerosas soluções já foram propostas, mas nenhuma é inteiramente satisfatória.

As repetições em um experimento têm, entre outras, as funções de providenciar a estimativa do erro experimental e aumentar a precisão do experimento, reduzindo o erro das médias dos tratamentos e aumentando a sensibilidade dos testes estatísticos (STEEL & TORRIE, 14).

O número de repetições, segundo FEDERER (10), depende do grau de precisão desejada, da variação do material experimental, dos recursos disponíveis e do tamanho e forma da parcela. Se as diferenças esperadas entre tratamentos são grandes, não é necessária uma alta precisão no experimento.

De acordo com STEEL & TORRIE (14), é importante o experimentador tomar uma decisão correta sobre o número de repetições necessárias, pois é um desperdício usar dez repetições para detectar uma diferença onde quatro repetições seriam suficientes para evidenciar esta diferença na maior parte dos casos. Também é uma perda de tempo fazer um experimento onde o número de repetições não é suficiente para detectar diferenças importantes, a não ser ocasionalmente.

Segundo CARMER & SWANSON (3), a sensibilidade ou habilidade de um teste para detectar diferenças entre tratamentos depende de diversos fatores: número de repetições, número de tratamentos, tamanho da diferença relativa entre médias e grau de homogeneidade entre médias de tratamento. O experimentador controla o número de repetições e o número de tratamentos, mas não tem conhecimentos precisos sobre a verdadeira média dos tratamentos e homogeneidade dos mesmos.

O aumento do número de repetições em um experimento aumenta o número de graus de liberdade na estimativa do erro, o que influi em muitos dos usos que são feitos com os resultados obtidos. Quando o número de graus de liberdade aumenta, diminui o intervalo de confiança pa-

ra a verdadeira diferença aumentando a probabilidade de se obter diferença significativa, aumentando assim a sensibilidade dos testes (COCHRAN & COX, 6).

O número de tratamentos em um experimento influi na precisão do mesmo e modifica o número de repetições requeridas para um determinado grau de precisão, especialmente num experimento onde o grau de liberdade do resíduo é menor do que 20 (STEEL & TORRIE, 14).

Embora o experimentador não tenha conhecimento sobre a verdadeira diferença entre tratamentos, ele pode estabelecer qual a verdadeira diferença entre médias que tem valor prático e que, portanto, deve ser detectada pelos testes usados. STEEL & TORRIE (14) lembram que o aumento do número de repetições não reduz os erros devidos a deficiências técnicas e que uma diferença estatisticamente significativa pode não ser de nenhuma importância prática. O tamanho da diferença com o valor prático deve ser julgado por alguém com conhecimento do assunto estudado. Esta medida, associada com a medida do quanto é desejável esta diferença, auxilia para determinar o número de repetições (STEEL & TORRIE, 14).

Ao aplicarmos um teste há a possibilidade de se cometer três tipos de erro. O erro tipo I é cometido quando se afirma que duas médias verdadeiras são diferentes, quando na realidade não são. O erro tipo II é cometido quando se afirma que duas médias verdadeiras são iguais, quando na realidade são diferentes. O erro tipo III é cometido quando se afirma que a média verdadeira de A é maior que a de B, quando na realidade B é maior do que A.

A importância de um ou outro tipo de erro cabe ao experimentador determinar, levando em conta o prejuízo que pode advir.

Na discussão dos testes para comparação múltipla de médias, FEDERER (10), STEEL (13) e DUNNETT (9) deram maior ênfase ao erro tipo I. Em experimento agrícolas, segundo CHEW (5), o erro tipo II é, no mínimo, tão importante quanto o erro tipo I. CARMER (2) lembrou que, se um erro tipo II foi cometido, um tratamento foi declarado não significativamente diferente de outro, apesar de realmente ser, e o agricultor poderá ter um prejuízo econômico se escolher o tratamento inferior. Esta perda será tanto maior quando maior o tamanho da verdadeira diferença.

Os erros tipo I e II prejudicam a imagem do pesquisador, mas o ti-

po II por outro lado representa uma indecisão. Se isto envolver diferenças grandes, os usuários dos resultados podem considerar um desperdício de recursos a condução de experimentos (CARMER, 2).

Estudando a influência do número de repetições na sensibilidade dos diversos testes estatísticos CARMER & SWANSON (3) encontraram que a probabilidade do erro tipo II é modificada com número de repetições. Em geral a sensibilidade aumenta com o número de repetições, mas este aumento depende da verdadeira diferença entre médias.

CARMER & SWANSON (4) chegaram a conclusão que a diferença na sensibilidade entre os diversos testes fica menos acentuada com o aumento do número de repetições e do tamanho da diferença relativa entre médias verdadeiras.

Um número moderado de repetições pode detectar diferenças relativas grandes entre médias, mas diferenças relativas muito pequenas entre médias nem mesmo um número muito grande de repetições é capaz de evidenciar (CARMER & SWANSON, 3).

A finalidade do presente trabalho é determinar o efeito da variação do número de repetições na sensibilidade dos testes Tukey, Duncan e Diferença Mínima Significativa (DMS) em comparação múltipla de médias determinando a probabilidade de ocorrência dos erros tipos I e II para diversos valores de diferenças relativas entre médias.

MATERIAL E MÉTODOS

O estudo foi feito por simulação, sendo para isto escrito um programa, em linguagem FORTRAN IV, para o computador IBM360, que gerou os dados e analisou os experimentos.

Os valores foram gerados usando-se o método de MARSAGLIA & BRAY (12) e transformados em variáveis independentes e com distribuição normal usando o método de BOX & MULLER (1).

Os dados foram simulados a partir do modelo linear.

$$Y_{ij} = M + T_i + E_{ij}$$

onde: Y_{ij} é a observação simulada para a repetição j do tratamento i , M a média geral do experimento, que foi estabelecida igual a 100, T_i é o efeito do tratamento i , igual à diferença entre a média geral do experimento e a média verdadeira atribuída no tratamento. E_{ij} é o desvio aleatório amostrado independentemente para cada observação de uma dis-

tribuição normal com média zero e desvio padrão. σ , igual a 10, dando ao experimento um coeficiente de variação de 10%.

Os dados simulados formaram um conjunto de 500 experimentos para cada um dos 12 conjuntos de médias verdadeiras mostradas na Tabela 1, onde o número de tratamentos foi 6, 10, 15 e 20 e o número de repetições foi 3, 4, 5, 6, 7 e 8, dando um total de 33.000 experimentos. Os testes aplicados foram DMS e Tukey descritos por STEEL & TORRIE (14) e Duncan descritos por DUNCAN (8) e todos com nível $\alpha = 0,05$, com e sem exigência de um teste F significativo para aplicação do mesmo.

Foi feita a determinação da porcentagem de erros tipo I e II cometido usando os diferentes números de repetições, tomando como base o número máximo de erros possíveis para o tipo considerado, isto é, considerando todos os contrastes possíveis entre duas médias.

Para o cálculo da probabilidade do F teórico foi usado o método descrito por DAVIES (7).

RESULTADOS OBTIDOS E DISCUSSÃO

Não houve efeito do número de repetições sobre a porcentagem do erro tipo I ocorrido quando foram aplicados os testes Tukey, Duncan e DMS para 11 grupos de médias estudadas, com pelo menos duas médias iguais, (Tabela 2, 3 e 4) quando não se exigiu um F significativo para o uso do teste. Estes resultados concordam com os de CARMER & SWANSON (3, 4).

A exigência de um F significativo para aplicação dos testes modificou a porcentagem do erro tipo I encontrada, mas o número de repetições não influenciou sobre esta porcentagem nos grupos com médias iguais (grupos 1, 4, 7 e 10) nem com médias heterogêneas (grupos 6, 9 e 12). No primeiro caso, o F teórico (Tabela 5) para todos os experimentos era 1,0 e não se modificava com o número de repetições, foram muito poucos os experimentos com F significativo. No segundo caso os F teóricos eram todos altos e praticamente todos os experimentos apresentavam F significativo.

Nos grupos 2, 5, 8 e 11 onde só duas médias diferiam das demais, o número de experimentos com F significativo aumentou o número de repetições, sendo este aumento mais acentuado nos grupos com menor número de tratamentos, o que concorda com o F teórico que, nestes grupos, é maior.

TABELA 1. Conjunto de médias verdadeiras para tratamentos usadas no estudo de simulação.

Número do grupo	Número de médias	Médias dos tratamentos											
1	6	100	100	100	100	100	100						
2	6	90	100	100	100	100	110						
3	6	85	90	95	100	110	120						
4	10	100	100	100	100	100	100	100	100	100	100	100	100
5	10	90	100	100	100	100	100	100	100	100	100	100	110
6	10	85	90	90	95	100	100	100	100	110	110	120	
7	15	100	100	100	100	100	100	100	100	100	100	100	100
8	15	90	100	100	100	100	100	100	100	100	100	100	100
9	15	85	85	90	90	95	97	98	100	100	100	105	105
10	20	100	100	100	100	100	100	100	100	100	100	100	100
11	20	90	100	100	100	100	100	100	100	100	100	100	100
12	20	85	85	90	90	90	90	90	95	95	100	100	100
		100	100	110	110	110	110	110	120	120	120	120	120

TABELA 2. Porcentagens de erro tipo I cometido usando o teste Tukey com $\alpha = 0,05$, obtidos incluindo todos os contrastes possíveis entre duas médias com 500 experimentos em cada situação.

Número de tratamentos	Grupo	Número de repetições											
		Sem F*						Com F**					
		3	4	5	6	7	8	3	4	5	6	7	8
6	1	0,016	0,150	0,110	0,090	0,160	0,120	0,120	0,110	0,080	0,070	0,130	0,070
	2	0,067	0,130	0,130	0,130	0,230	0,070	0,067	0,130	0,130	0,130	0,230	0,070
10	4	0,076	0,018	0,080	0,030	0,067	0,040	0,058	0,004	0,067	0,020	0,027	0,030
	5	0,036	0,043	0,021	0,014	0,021	0,035	0,035	0,042	0,021	0,014	0,021	0,035
15	6	0,000	0,080	0,040	0,000	0,000	0,000	0,000	0,080	0,040	0,000	0,000	0,000
	7	0,019	0,009	0,032	0,020	0,017	0,020	0,100	0,038	0,006	0,002	0,008	0,009
	8	0,023	0,000	0,028	0,010	0,008	0,010	0,010	0,010	0,028	0,013	0,008	0,013
	9	0,000	0,000	0,000	0,000	0,000	0,029	0,000	0,000	0,000	0,000	0,000	0,029
20	10	0,074	0,006	0,003	0,004	0,004	0,084	0,002	0,005	0,003	0,000	0,002	0,004
	11	0,001	0,006	0,006	0,004	0,006	0,006	0,013	0,005	0,006	0,004	0,006	0,006
	12	0,007	0,000	0,007	0,020	0,000	0,013	0,007	0,000	0,007	0,020	0,000	0,013

* Sem exigência de F com $P < 0,05$.

** Com exigência de F com $P < 0,05$.

TABELA 3. Porcentagens de erro tipo I cometido usando o teste Duncan com $\alpha = 0,05$, obtidas incluindo todos os contrastes possíveis entre duas médias com 500 experimentos em cada situação.

Número de tratamentos	Grupo	Número de repetições											
		Sem F*						Com F**					
		3	4	5	6	7	8	3	4	5	6	7	8
6	1	1,05	0,97	1,00	1,01	1,44	1,05	0,20	0,17	0,16	0,15	0,36	0,11
	2	0,53	0,70	0,53	1,10	1,30	1,10	0,40	0,63	0,43	1,07	1,30	1,10
10	4	0,73	1,04	0,80	0,20	0,76	0,71	0,12	0,13	0,16	0,07	0,10	0,11
	5	0,35	1,09	0,31	0,33	0,41	0,63	0,21	0,85	0,31	0,29	0,40	0,62
15	6	0,40	2,00	0,48	1,04	0,52	0,80	0,40	2,00	0,48	1,04	0,52	0,80
	7	0,22	0,17	0,19	0,21	0,16	0,22	0,05	0,02	0,03	0,03	0,03	0,06
	8	0,34	0,24	0,36	0,24	0,22	0,20	0,18	0,15	0,24	0,17	0,17	0,18
	9	0,37	0,46	0,54	0,63	0,97	0,89	0,37	0,46	0,54	0,63	0,97	0,89
20	10	0,31	0,71	0,30	0,30	0,31	0,29	0,00	0,06	0,01	0,01	0,00	0,01
	11	0,20	0,73	0,19	0,22	0,22	0,22	0,07	0,31	0,13	0,19	0,18	0,19
	12	0,21	0,99	0,31	0,27	0,24	0,39	0,21	0,99	0,31	0,27	0,24	0,39

* Sem exigência de F com $P < 0,05$.

** Com exigência de F com $P < 0,05$.

TABELA 4. Porcentagem de erro tipo I cometido usando o teste DMS com $\alpha = 0,05$, obtidos incluindo todos os contrastes possíveis entre duas médias com 500 experimentos em cada situação.

Número de tratamentos	Grupo	Número de repetições											
		Sem F*						Com F**					
		3	4	5	6	7	8	3	4	5	6	7	8
6	1	2,52	2,45	2,37	2,44	3,23	2,81	0,49	0,39	0,40	0,24	0,65	0,41
	2	2,30	2,40	2,36	2,60	3,20	3,13	1,37	2,24	2,00	2,57	3,20	3,13
10	4	2,99	2,56	3,16	2,72	3,18	2,79	0,31	0,16	0,38	0,24	0,37	0,24
	5	2,37	2,57	2,83	3,01	2,79	3,04	1,08	1,74	2,13	0,30	2,45	2,85
15	6	1,92	3,20	2,72	3,48	2,32	2,52	1,92	3,20	2,72	3,48	2,32	2,52
	7	2,90	2,50	2,80	2,77	2,80	2,94	0,13	0,34	0,20	0,10	0,13	0,20
20	8	3,07	2,58	3,30	2,59	2,71	2,63	1,13	1,39	1,90	1,84	2,02	2,33
	9	2,40	2,54	2,66	2,77	3,43	2,66	2,40	2,54	2,66	2,77	3,43	2,66
20	10	2,53	2,52	2,65	2,59	2,62	2,52	0,07	0,12	0,07	0,06	0,03	0,03
	11	2,70	2,49	2,65	2,80	2,77	2,72	0,57	0,74	1,15	1,77	1,71	2,02
	12	2,90	2,42	2,50	1,51*	2,65	2,92	2,90	2,42	2,50	1,51	2,65	2,92

* Sem exigência de F com $P < 0,05$.

** Com exigência de F com $P < 0,05$.

TABELA 5. Porcentagem de experimento com F significativo (%), F teórico (F) e probabilidade de ocorrer ao acaso o F teórico (P).

Grupo	Número de repetições																	
	3			4			5			6			7			8		
	%	F	P	%	F	P	%	F	P	%	F	P	%	F	P	%	F	P
1	1,6*	1,00	0,4588	1,4	1,00	0,4464	1,6	1,00	0,4397	0,8	1,00	0,4354	2,4	1,00	0,4325	1,6	1,00	0,4303
2	23,8	2,50	0,0895	41,0	3,00	0,0383	57,0	3,50	0,0162	74,2	4,00	0,0068	82,8	4,50	0,0028	90,0	5,00	0,0012
3	90,8	7,36	0,0024	99,4	9,48	0,0002	99,8	11,60	0,0000	100,0	13,72	0,0000	100,0	15,84	0,0000	100,0	17,96	0,0000
4	1,2	1,00	0,4214	0,6	1,00	0,4616	1,8	1,00	0,4562	1,2	1,00	0,4528	1,8	1,00	0,4504	1,2	1,00	0,4487
5	14,0	1,66	0,1649	33,3	1,89	0,0922	39,8	2,11	0,0513	59,2	2,33	0,0280	68,0	2,55	0,0150	79,8	2,77	0,0079
6	91,6	4,48	0,0025	100,0	4,64	0,0007	99,8	5,80	0,0000	100,0	6,96	0,0000	100,0	8,12	0,0000	100,0	9,28	0,0000
7	0,8	1,00	0,4779	0,4	1,00	0,4698	1,0	1,00	0,4653	0,6	1,00	0,4625	0,8	1,00	0,4606	1,2	1,00	0,4591
8	11,6	1,43	0,1994	23,2	1,57	0,1255	31,2	1,71	0,0771	43,2	1,85	0,0463	50,2	1,99	0,0272	65,4	2,14	0,0151
9	98,8	4,51	0,0003	100,0	5,68	0,0000	100,0	6,85	0,0000	100,0	8,02	0,0000	100,0	9,19	0,0000	100,0	10,36	0,0000
10	0,4	1,00	0,4814	0,8	1,00	0,4743	0,4	1,00	0,4704	0,4	1,00	0,4680	0,2	1,00	0,4663	0,2	1,00	0,4650
11	6,6	1,31	0,2306	12,4	1,42	0,1520	20,6	1,52	0,1008	37,6	1,63	0,0630	41,2	1,73	0,0399	51,4	1,84	0,0238
12	99,6	4,31	0,0000**	100,0	5,42	0,0000	100,0	6,52	0,0000	100,0	7,63	0,0000	100,0	8,73	0,0000	100,0	9,84	0,0000

* Cada percentagem foi calculada com 500 experimentos.

** P = 0,0000 significa $P < 0,0001$.

Para seis tratamentos a probabilidade de obter por acaso um F igual ao F teórico é menor do que 0,05 para quatro ou mais repetições. Já com 20 tratamentos essa situação ocorre só com sete ou mais repetições.

O aumento do número de repetições modificou a redução do erro tipo I obtida com a exigência de um F significativo anterior ao teste. Tomando-se como exemplo o teste Duncan para o grupo 5 (Tabela 3), verifica-se que com três repetições a porcentagem de erro tipo I foi de 0,35% quando se exigiu um F significativo anterior ao teste, baixando para 0,21% quando o mesmo foi exigido. Com oito repetições, a porcentagem de erro tipo I sem exigências de teste foi de 0,63% quando não se exigiu o teste F anterior a 0,62% quando o mesmo foi exigido.

A porcentagem de ocorrência do erro tipo II diminuiu com o aumento de repetições, o que concorda com STEEL & TORRIE (14), mas a intensidade deste efeito dependeu da maior ou menor homogeneidade dos tratamentos. Para os grupos mais homogêneos (grupos 2, 5, 8 e 11), resultados apresentados nas Tabelas 6, 7 e 8, esta redução dependeu não só do número de repetições, mas também da exigência ou não de um F significativo anterior ao teste, bem como do tamanho da diferença relativa entre médias verdadeiras, o que concorda com CARMER & SWANSON (3) quando dizem que a sensibilidade de um teste depende do número de repetições, número de tratamentos e tamanho da diferença relativa entre médias.

Para médias verdadeiras com diferenças iguais a 2 σ , o aumento do número de repetições diminuiu a porcentagem do erro tipo II para os três testes; como exemplo temos o teste Tukey para 20 tratamentos: sem exigência de F significativo anterior, apresentou 83,4% de erro tipo II com 3 repetições e baixou para 12,8% com 8 repetições. A exigência de um F significativo anterior aumentou para 88,3% o erro tipo II com 3 repetições e 22,0% com 8 repetições.

Para a diferença de 1 σ , a redução da porcentagem de erro tipo II pelo aumento do número de repetições foi bem menos acentuada, mostrando que, à medida que diminui a diferença relativa entre médias, o aumento do número de repetições é menos eficiente na redução do erro tipo II, o que está de acordo com CARMER & SWANSON (3), quando dizem que diferenças relativas muito pequenas nem mesmo um número muito grande de repetições é capaz de evidenciar. Para diferenças de 1 σ não houve influên-

TABELA 6. Porcentagens de erro tipo II cometido nos grupos 2, 5, 8 e 11 usando o teste Tukey com $\alpha = 0,05$, obtidas incluindo todos os constrates possíveis entre duas médias com 500 experimentos em cada situação.

Diferença real entre tratamen- tos ()	Grupo	Nº de tra- tamentos	Número de repetições															
			Sem F*								Com F**							
			3	4	5	6	7	8			3	4	5	6	7	8		
2	2	6	73,8	57,4	40,4	21,2	15,0	10,2			78,4	63,8	47,2	27,4	19,2	11,0		
	5	10	83,8	64,4	52,8	30,0	25,0	12,8			88,8	73,0	61,2	42,6	35,6	22,0		
	8	15	88,4	70,0	61,4	39,0	33,0	21,2			93,8	81,6	71,8	60,0	52,2	36,6		
	11	20	90,0	77,8	65,0	43,0	36,0	24,2			96,6	90,0	81,6	65,6	61,4	51,0		
1	2	6	97,2	94,9	91,7	87,1	85,1	79,8			97,3	95,0	91,8	87,2	85,1	79,9		
	5	10	98,7	97,0	96,6	93,3	91,7	87,8			98,9	97,1	96,8	93,4	91,8	87,8		
	8	15	99,3	98,3	97,8	96,2	94,8	92,6			99,5	98,5	99,4	96,4	95,0	92,8		
	11	20	99,7	99,1	98,3	97,1	96,6	95,5			99,8	99,7	98,6	97,3	96,8	95,9		

* Sem exigência de F com $P < 0,05$.

** Com exigência de F com $P < 0,05$.

TABELA 7. Porcentagens de erro tipo II cometido nos grupos 2, 5, 8 e 11, usando o teste Duncan com $\alpha = 0,05$, obtidas incluindo todos os contrastes possíveis entre duas médias com 500 experimentos em cada situação.

Diferença real entre tratamen- tos ()			Nº de tra- tamentos	Número de repetições											
				Sem F*								Com F**			
				3	4	5	6	7	8	3	4	5	6	7	8
2	2	6	53,4	30,2	18,8	6,4	4,6	2,0	77,8	62,0	44,4	26,4	17,2	10,2	
	5	10	53,8	21,2	24,0	18,0	6,6	3,2	86,4	66,8	61,2	40,8	32,6	20,6	
	8	15	72,8	46,0	34,2	19,0	14,4	7,6	90,8	77,6	68,0	57,2	50,2	34,6	
	11	20	65,5	27,6	29,2	16,6	12,4	6,4	94,4	87,8	80,2	63,0	59,2	48,8	
1	2	6	93,5	90,2	84,9	77,2	73,4	67,4	95,4	91,5	85,9	77,7	73,7	67,5	
	5	10	95,0	78,9	88,5	80,3	82,6	71,0	97,0	85,8	90,8	82,2	82,2	71,6	
	8	15	95,5	91,6	89,5	84,3	81,6	75,8	97,8	94,9	92,5	87,8	84,9	78,3	
	11	20	96,5	83,7	80,0	85,8	84,6	80,7	98,8	94,4	94,5	89,4	88,0	84,5	

* Sem exigência de F com $P < 0,05$.

** Com exigência de F com $P < 0,05$.

TABELA 8. Porcentagem de erro tipo II cometido nos grupos 2, 5, 8 e 11 usando o teste DMS com $\alpha = 0,05$.

Diferença real entre tratamen- tos ()	Grupo	Nº de tra- tamentos	Número de repetições															
			Sem F*								Com F**							
			3	4	5	6	7	8			3	4	5	6	7	8		
2	2	6	39,4	17,0	9,6	3,0	1,6	1,0			76,2	54,2	43,0	25,8	17,2	10,0		
	5	10	32,6	11,2	8,6	5,0	1,6	0,0			86,0	66,8	60,4	40,8	32,0	20,2		
	8	15	30,8	14,8	7,8	1,8	1,4	1,4			88,4	87,3	66,8	56,8	49,8	34,6		
	11	20	30,4	13,2	7,2	3,2	1,2	1,0			93,6	87,6	79,4	62,4	58,8	48,6		
1	2	6	81,9	73,3	65,4	55,6	52,0	43,2			89,0	80,4	71,1	59,3	54,7	44,6		
	5	10	81,4	70,1	66,0	54,7	52,3	42,0			93,2	82,7	79,0	64,3	59,3	47,7		
	8	15	80,8	69,8	64,9	56,1	51,4	44,1			94,8	93,9	82,2	74,3	67,9	56,1		
	11	20	80,1	70,7	64,8	54,0	72,9	45,1			96,8	92,7	87,6	75,3	83,2	65,2		

* Sem exigência de F com $P < 0,05$.** Com exigência de F com $P < 0,05$.

cia da exigência ou não de um F significativo anterior ao teste na porcentagem de erro tipo II.

Nos grupos heterogêneos (3, 6, 9 e 12) com médias verdadeiras com diferenças de até $3,5\sigma$, o aumento do número de repetições reduziu a porcentagem do erro tipo II (Tabelas 9, 10, 11), mas esta redução dependeu do tamanho da diferença relativa entre as médias verdadeiras e nestes grupos não houve efeito da exigência ou não de F significativo antes do teste, pois todos os experimentos apresentaram F significativo. O tamanho da diferença real que foi possível detectar variou com o número de repetições usado, o teste escolhido e a probabilidade de acertos desejada.

Usando o teste Tukey e desejando-se 95% ou mais de acertos a menor diferença relativa entre médias que foi possível detectar foi $2,5\sigma$ mesmo usando oito repetições. O número de repetições necessárias variou com o número de tratamentos, para seis tratamentos foram necessários sete repetições, para 10 tratamentos necessitou-se de sete repetições e para 15 e 20 tratamentos foram necessárias oito repetições.

Com a mesma exigência, 5% ou menos de erro, o teste Duncan permitiu detectar uma diferença de 2σ quando se usou oito repetições para todos os números de tratamento usado, embora a porcentagem de erro tipo II aumente com o número de tratamentos. A diferença mínima relativa entre médias detectada pelo DMS foi 2σ , usando-se seis repetições e não sofreu influência do número de tratamentos usados.

Um experimentador que deseje detectar, por exemplo uma diferença entre médias reais de tratamentos de 300 kg/ha para uma cultura com produção média de 2.000 kg/ha, em um experimento com 10 tratamentos e com coeficiente de variação de 10% poderá, baseado nas Tabelas 9, 10 e 11, escolher o número de repetições baseado na proporção do erro. Esta diferença corresponde a $1,5\sigma$. Usando o teste Duncan verifica-se de acordo com a Tabela 10, que, com 4 repetições, a porcentagem de erro tipo II é de 50,1%, isto significa que em 50,1% das vezes que o tratamento tiveram médias reais com esta diferença ele concluirá que estes tratamentos tem médias iguais. Usando oito repetições a porcentagem de erro será 29,5%. Diferenças de $1,5\sigma$ ou menos são difíceis de serem detectadas pelos testes com probabilidade de erro menor do que 20%, mesmo usando um número grande de repetições.

Cabe ao experimentador determinar o tamanho da diferença que tem

TABELA 9. Porcentagens do erro tipo II cometido nos grupos 3, 5, 6 e 12, usando o teste Tukey com $\alpha = 0,05$, obtidos incluindo todos os contrastes possíveis entre duas médias com 500 experimentos em cada situação.

Número de tratamentos	Número de repetições	Diferença verdadeira entre tratamentos - expressa em desvio padrão*						
		0,5	1,0	1,5	2,0	2,5	3,0	3,5
6	3	99,3	97,5	89,8	73,4	51,5	33,0	15,6
	4	99,0	93,6	80,2	53,8	26,6	9,8	2,6
	5	98,9	92,4	70,8	41,7	14,6	4,0	0,0
	6	98,4	88,5	59,9	23,3	3,8	1,0	0,0
	7	98,0	85,6	50,9	16,8	2,9	0,0	0,0
	8	97,3	79,9	35,4	7,3	1,0	0,2	0,0
10	3	99,7	99,0	94,6	84,5	65,9	43,9	25,0
	4	99,2	97,4	88,4	66,4	36,0	13,7	3,8
	5	99,8	96,4	83,8	45,2	21,6	5,1	1,0
	6	99,5	93,1	68,6	33,6	8,1	1,1	0,2
	7	99,4	91,4	62,4	22,9	3,5	0,1	0,0
	8	99,3	89,3	52,8	14,0	1,9	0,1	0,0
15	3	99,9	99,4	96,5	87,8	72,0	47,7	27,4
	4	99,9	98,7	91,8	71,1	40,3	17,3	4,0
	5	99,9	98,0	86,5	60,4	27,1	8,3	1,5
	6	99,7	95,7	77,0	4,7	11,4	1,3	0,2
	7	99,5	94,9	42,0	31,2	7,2	0,3	0,0
	8	99,6	92,2	66,8	18,8	1,4	0,1	0,0
20	3	99,9	99,5	97,7	91,1	77,0	54,5	31,0
	4	99,9	99,4	94,0	76,6	46,3	20,1	4,8
	5	99,9	98,5	90,1	65,3	32,8	9,5	1,5
	6	99,9	97,4	83,6	47,0	16,2	2,3	0,1
	7	99,8	96,9	76,7	38,5	8,7	0,6	0,0
	8	99,7	94,1	66,2	22,4	2,5	0,1	0,0

* (Diferença entre duas médias verdadeiras)/(desvio padrão verdadeiro = 10)

TABELA 10. Porcentagens do erro tipo II cometido nos grupos 3, 5, 6 e 17 usando o teste Duncan com $\alpha = 0,05$, obtidos incluindo todos os contrastes possíveis entre duas médias com 500 experimentos em cada situação.

Número de tratamentos	Número de repetições	Diferença verdadeira entre tratamentos - expressa em desvio padrão*						
		0,5	1,0	1,5	2,0	2,5	3,0	3,5
6	3	97,9	90,2	76,4	60,4	38,6	24,6	8,8
	4	96,2	85,0	60,5	35,6	14,5	5,4	1,0
	5	96,1	78,1	51,5	26,7	7,9	1,0	0,0
	6	94,0	69,8	36,8	11,7	1,3	0,4	0,0
	7	93,0	66,5	29,7	7,0	0,9	0,0	0,0
	8	90,4	57,4	20,8	3,5	0,4	0,2	0,0
10	3	98,2	94,7	84,8	65,8	43,3	22,6	10,8
	4	97,6	88,1	70,1	28,5	8,7	0,9	0,2
	5	97,5	85,6	62,1	28,0	8,7	1,7	0,0
	6	95,6	76,5	42,8	15,4	3,1	0,0	0,0
	7	95,8	75,1	39,0	8,7	1,1	0,0	0,0
	8	94,1	68,0	24,5	3,6	0,7	0,1	0,0
15	3	98,8	95,0	85,5	52,4	42,9	21,5	8,4
	4	98,3	91,4	63,5	41,2	16,2	4,6	1,0
	5	98,2	87,7	61,1	29,3	9,3	1,6	0,2
	6	96,6	80,9	46,0	14,6	2,5	0,0	0,0
	7	96,0	76,9	38,6	9,1	1,1	0,05	0,0
	8	95,1	70,9	27,5	4,1	0,1	0,0	0,0
20	3	99,2	96,1	88,2	69,5	46,9	23,2	9,9
	4	98,9	91,2	73,1	34,4	10,2	1,7	0,0
	5	98,2	89,3	65,7	30,9	9,7	1,5	0,2
	6	97,9	83,5	53,0	16,5	2,9	0,3	0,0
	7	97,3	80,1	40,1	10,8	1,3	0,1	0,0
	8	96,4	76,2	30,6	4,5	0,2	0,0	0,0

* (Diferença entre duas médias verdadeiras)/(desvio padrão verdadeiro = 10)

TABELA 11. Porcentagens do erro tipo II cometido nos grupos 3, 6, 9 e 12 usando o teste DMS com $\alpha = 0,05$, obtidos incluindo todos os contrastes possíveis entre duas médias com 500 experimento em cada situação.

Número de tratamentos	Número de repetições	Diferença verdadeira entre tratamentos - expressa em desvio padrão*						
		0,5	1,0	1,5	2,0	2,5	3,0	3,5
6	3	93,8	79,8	56,5	34,0	16,0	4,0	1,4
	4	91,2	71,7	39,3	15,4	4,1	0,8	0,0
	5	90,3	66,1	44,1	11,4	1,4	0,2	0,0
	6	88,1	56,9	19,2	3,2	0,4	0,0	0,0
	7	87,3	50,8	14,7	2,1	0,0	0,0	0,0
	8	84,0	41,6	8,5	0,9	0,7	0,0	0,0
10	3	93,3	81,3	59,8	34,4	15,3	4,8	1,0
	4	92,0	71,9	40,3	14,4	3,4	0,7	0,2
	5	90,5	64,6	30,5	8,3	1,1	0,1	0,0
	6	87,5	55,4	20,0	3,17	0,3	0,0	0,0
	7	86,6	50,9	14,8	1,7	0,1	0,0	0,0
	8	84,4	43,7	9,7	0,8	0,1	0,0	0,0
15	3	93,2	80,1	57,8	32,3	13,0	3,4	0,7
	4	91,6	71,4	38,6	13,3	2,4	0,5	0,0
	5	90,0	64,6	30,5	8,3	1,1	0,1	0,0
	6	87,4	54,1	18,7	2,8	0,4	0,0	0,0
	7	86,5	50,5	15,1	1,6	0,1	0,0	0,0
	8	84,5	43,4	8,1	0,7	0,0	0,0	0,0
20	3	93,5	80,0	58,3	31,5	13,7	3,8	0,6
	4	91,0	70,9	37,0	12,6	2,6	0,2	0,0
	5	90,2	64,9	30,4	7,9	1,0	0,1	0,0
	6	88,8	55,7	20,6	3,2	0,2	0,0	0,0
	7	86,2	51,2	13,5	1,5	0,1	0,0	0,0
	8	84,2	41,9	8,5	0,6	0,0	0,0	0,0

* (Diferença entre duas médias verdadeiras)/(desvio padrão verdadeiro = 10)

importância prática para seu estudo e a margem de erro com que deseja trabalhar. Baseado nestes dados, deve escolher então o teste e o número de repetições que permita detectar esta diferença com a precisão desejada que pode variar com a importância econômica do erro (CARMER, 2), pois é uma perda de tempo e dinheiro realizar um experimento, provavelmente, não vai ser encontrada como significativa uma diferença entre médias considerada importante.

CONCLUSÕES

Para experimentos com número de repetições variando de 3 a 8 e possuindo de 6 a 20 tratamentos:

1. A modificação do número de repetições só influencia sobre a porcentagem do erro tipo I cometido quando se exige um F significativo anterior ao teste em experimentos onde poucos tratamentos diferem dos demais.

2. Nessa situação a probabilidade de erro tipo I e II diminui com o aumento do número de repetições.

3. A diferença mínima entre médias verdadeiras mostrada, usando $\alpha = 0,05$ pelo teste Tukey é de $2,5 \sigma$ sendo necessária oito repetições pelos testes Duncan e DMS é 2σ sendo necessárias oito e seis repetições respectivamente. O número de repetições necessárias aumenta com o número de tratamentos para o teste Tukey e Duncan.

BIBLIOGRAFIA CITADA

1. BOX, G.E. & MULLER, M.E. A note on the generation of random normal deviates. *Ann. Math. Stat.*, 29:610-611, 1958.
2. CARMER, S.G. Optimal significance levels for application of the Least significance difference in crop performance trials. *Crop. Sci.*, 16:95-99, 1976.
3. CARMER, S.G. & SWANSON, M.R. Detection of differences between mean: a Monte Carlo study of five pairwise multiple comparison procedures. *Agron. J.*, 63:940-945, 1971.
4. CARMER, S.G. & SWANSON, M.R. An evaluation of ten pairwise multiple comparison procedures by Monte Carlo methods. *J. Amer. Stat. Assoc.*, 68:67-74, 1973.
5. CHEW, V. Comparing treatments means. A compedium. *Hort Science*, 11:348-356, 1976.
6. COCHRAN, W.G. & COX, G.M. *Experimental Designs*. New York, John Wiley Sons, 1953. 453 p.
7. DAVIES, R.G. *Computer Programing in Quantitative Biology*. London,

- Academy Press, 1971. 492 p.
8. DUNCAN, D.B. Multiple range and multiple F teste. *Biometrics*, 11:1-42, 1955.
 9. DUNNETT, C.V. Query 272. Multiple comparison tests. *Biometrics*, 26:139-141, 1970.
 10. FEDERER, W.T. *Experimental Design, Theory and Application*. New York, Mcmillan Co, 1955, 544 p.
 11. GOMES, F.P. *Curso de Estatística Experimental*. São Paulo, Livraria Nobel, 1976. 420 p.
 12. MARSAGLIA, G. & BRAY, T.A. One line random number generation and their use in combinations. *Comm. Assoc. Comp. Mach.*, 11:757-759, 1968.
 13. STEEL, R.G.D. Query 163. Error rates in multiple comparison. *Biometrics*, 17:326-328, 1961.
 14. STEEL, R.G.D. & TORRIE, J.H. *Principles and Procedures of Statistics*. New York, McGraw Hill, 1960. 481 p.