MODIFIED DUNNETT'S TEST FOR A RANDOMIZED COMPLETE BLOCK DESIGN

Armando CONAGIN[1]
Décio BARBIN[2]
Clarice Garcia Borges DEMÉTRIO[2]

- *ABSTRACT: Dunnett's test is, in general, preferred to test the paired differences between treatment means and a control mean of a considered variable after analysis of variance (ANOVA). In this paper, we present a modified Dunnett's test and evaluate its performance in relation to the original Dunnett's test and to other ten tests, considering experiments in a randomized complete block design (RCBD). Monte Carlo method was used to simulate data from a normal distribution for 540 experiments in a RCBD, varying number of treatments (two groups, GI: with nine treatments and GII: with 14 treatments), number of replications (360 experiments with four blocks and 180 experiments with eight blocks) and size of the error (three coefficients of variation: 5%, 10% and 15%), giving 12 combinations. As the difference between the treatments and the control decreases there is a clear differentiation in power among the tests, and the modified tests are more efficient than their original ones. The modified Dunnett's test, in the region of small differences, is the most efficient among the tests of maximum experimentwise error rate (MEER) in comparison with all error rate type tests.*

- *KEYWORDS: Multiple comparison tests; Monte Carlo method; Simulation; Modified tests.*

1 Introduction

In general, the scientific and technological knowledge is gained slowly by observation or mainly by execution of experiments carefully planned to test a formulated hypothesis. The overall *F*-test from an ANOVA of the data allows to reject or not this hypothesis. Multiple comparison (MC) tests are then used to test the difference between subgroups of a considered variable and "making a decision about which test to use is not an easy task" (Demirhan *et al.*, 2010). In order to compare MC tests it is usually used Monte Carlo simulation procedures as discussed by Naylor *et al.* (1966), Shimizu (1984) and Podlich and Cooper (1998). A Monte Carlo simulation allows to generate pseudo-random numbers according to a specific distribution (Dachs, 1988).

The behavior of MC tests and their performance in terms of type I error rate have been evaluated by Gabriel (1964), Boardman and Moffitt (1971), O'Neill and Wetherill (1971), Bernardson (1975), Hocking (1985), Hayter (1986), Hsu (1996) and many others

[1] Instituto Agronômico de Campinas – IAC, CEP: 13012-970, Campinas, SP, Brazil.
[2] Universidade de São Paulo – USP, Escola Superior de Agricultura Luiz de Queiroz – ESALQ, Departamento de Ciências Exatas, CEP: 13418-900, Piracicaba, SP, Brasil. E-mail: *decio.barbin@usp.br* / *clarice.demetrio@usp.br*

but there are still many questions to be answered. Boardman and Moffitt (1971) compared five MC procedures, recording empirical error rates for 10,000 experiments with $t = 2(1)11$ treatments that did not differ (the global null hypothesis H_0 was true) and $r = 5$, 10, 15 observations per treatment. They showed graphically that, considering $\alpha = 0.05$, as the number of treatments increased from 2 to 11, the comparisonwise error rates: i) were around 0.05 for Student's test; ii) decreased from 0.05 to 0.025 for Duncan's test; iii) decreased from 0.05 to 0.001 for Student-Newman-Keuls (SNK) and Tukey´s tests and iv) decreased from 0.05 to almost 0 for Scheffé's test. The experimentwise error rates i) increased from 0.05 to 0.65 for Student's test; ii) increased from 0.05 to 0.35 for Duncan's test; iii) were around 0.05 for SNK and Tukey's tests and iv) decreased from 0.05 to almost 0 for Scheffé's test. Perecin and Barbosa (1988) compared five MC procedures, recording empirical error rates experiments with $t = 5$, 10, 20, 40 and 100 treatments that did not differ and with the differences of consecutive means differing by $2\sigma_{\bar{X}}$, $4\sigma_{\bar{X}}$, $6\sigma_{\bar{X}}$, and $8\sigma_{\bar{X}}$. From their study, as the number of treatments increased, the number of significant differences, i) for differences of $2\sigma_{\bar{X}}$, increased from 23.9 to 25.9 for Student's test; was almost the same for Duncan and SNK's tests; were equal to 6.6, 2.5, 0.9 and 0.0 for Tukey's test; ii) for differences of $4\sigma_{\bar{X}}$, increased from 76.9 to 82.9 for Student's test; increased from 73.6 to 79.9 for Duncan; increased from 60.4 to 69.2 for SNK's test; decreased from 42.6 to 4.70 for Tukey's test; iii) for differences of $6\sigma_{\bar{X}}$, increased from 98.8 to 99.5 for Student's test; increased from 97.6 to 99.1 for Duncan´s test; increased from 91.6 to 96.6 for SNK's tests; decreased from 87.7 to 45.5 for Tukey's test; iv) for differences of $8\sigma_{\bar{X}}$, were equal to 100.0 for Student's test; increased from 97.6 to 99.1 for Duncan; increased from 99.4 to 100.0 for SNK's test; decreased from 99.4 to 92.5 for Tukey's test. The results showed some small positive variation of power for increasing number of treatments, and for all tests, except Tukey's test, that showed a decreasing tendency, indicating that when the range of number of treatments increases it loses power. Demirhan *et al.* (2010), Ramsey *et al.* (2010) and Ramsey *et al.* (2011) studied the influence of violations of assumptions of normality and homogeneity of variances on the choice of a MC procedure.

A common practice in agriculture and other sciences is to compare new treatments (supposedly superior) to a control. This paper proposes a modification to Dunnett's test and compares its efficiency in relation to the original Dunnett's test, to other (unilateral and bilateral Student *t* test, Duncan, SNK, Tukey, Sidak and Bonferroni) tests commonly used and to other modified (Bonferroni, Sidak and Tukey) tests proposed by Conagin and Barbin (2006a,b) and Conagin *et al*. (2008).

2 Material and methods

Data simulation

To evaluate the proposed test and compare it with the other tests, Monte Carlo method was used to simulate data from a normal distribution for 540 experiments in a RCBD, varying number of treatments (two groups, GI: with nine treatments and GII: with 14 treatments), number of replications (360 experiments with four blocks and 180

experiments with eight blocks) and size of the error (three coefficients of variation, 5%, 10% and 15%), giving 12 combinations.

The adopted model was $y_{ij} = \mu_c + \beta_j + \tau_i + \varepsilon_{ij}$, $i = 1,...,t$, $j = 1,..., r$, where μ_c is the mean of the control, β_j is the j-th block effect, τ_i is the i-th treatment effect and ε_{ij} is the ij-th error term. Additionally, it was assumed that the ε_{ij}'s are independent and normally distributed with mean zero and common variance σ^2, that is, $\varepsilon_{ij} \sim N(0,\sigma^2)$, which can be generated using the function rannor (seed) of SAS (1999).

The mean values were defined as $\mu_i = \mu_c + \tau_i$, where μ_c is the mean of the control, τ_i is the i-th treatment effect and μ_i is the mean of the *i*-th treatment. The two groups of treatments considered here were: i) GI with nine treatments which included the control and new treatments whose means differed from the control in 40%, 35%, 30%, 25%, 20%, 15%, 10% and 5%, that is, supposing $\mu_1 = 1.4\mu_c$, $\mu_2 = 1{,}35\mu_c$, $\mu_3 = 1{,}3\mu_c$, …, $\mu_8 = 1{,}05\mu_c$ and ii) GII with 14 treatments including the ones from GI and considering additionally five treatments that did not differ from the control, that is assuming $\mu_{10} = \mu_{11} = \mu_{12} = \mu_{13} = \mu_{14} = \mu_c$. The assumed value for the control mean was $\mu_c = 4000$ and the blocks effects were $\beta_1 = -120$, $\beta_2 = -40$, $\beta_3 = 40$ and $\beta_4 = 120$.

The null hypotheses considered were $H_0 : \mu_1 = \mu_2 = ... = \mu_9$, where $\mu_9 = \mu_c$, for GI and $H_0 : \mu_1 = \mu_2 = ... = \mu_9 = ... = \mu_{14}$, for GII. The alternative hypotheses were $H_a : \mu_i > \mu_c$, $i = 1,...,8$, for GI and $i = 1,...,8,10,...,14$, for GII. It is clear that for GI, H_0 is totally false, while for GII it is partially true for $\mu_{10} = ... = \mu_{14} = \mu_c$.

From the ANOVA table if $F_0 > F_c$, where $F_0 = MST / MSE$, MST is the mean square of treatments, *MSE* is the mean square error and F_c is the percentile value of the F distribution with (*t*-1) and (*t*-1)(*r*-1) degrees of freedom for a significance level $\alpha = 0.05$ or $\alpha = 0.01$, then H_0 is rejected and at least one contrast between means is statistically significant.

Modified Dunnett's test

Dunnett's test controls the Type I experimentwise error for comparisons of all $(t-1)$ treatment means against a control mean. The statistics of the test is given by

$$d_i = \frac{m_i - m_c}{\sqrt{\frac{2}{r} MSE}},$$

where, m_c is the estimate of the mean of the control and m_i is the estimate of the mean of the *i*-th treatment, $i = 1,...,t-1$, *MSE* is the mean square error as computed from the analysis of variance and *r* is the sample size of the experimental treatments and the control. If $d_i > t_D$, where t_D is the α critical value from a table of Dunnett's test with

$p = t - 1$ and error number of degrees of freedom (Dunnett, 1955), the null hypothesis $H_0 : \mu_i = \mu_c$ is rejected.

The procedure to develop the modified tests follows the steps:

i) perform the ANOVA for the data and if $F_0 > F_c$, the global null hypothesis $H_0 : \mu_1 = \mu_2 = ... = \mu_t$ is rejected, and at least one contrast between means is statistically significant;

ii) calculate the least significant difference (*lsd*),

$$lsd = t_\alpha \sqrt{\frac{2}{r} MSE},$$

where t_α is the α (0.05 or 0.01) quantile of the Student distribution with error number of degrees of freedom, $\upsilon = (t-1)(r-1)$;

iii) calculate $\hat{a}$ as the number of significant differences between the treatment means, m_i, and the control mean, m_c, that is, number of times that $m_i > m_c + lsd$. This number $\hat{a}$ is an estimate of the number a of real parametric differences (the result is conservative because small differences will not be possibly detected, influenced by the random variability).

iv) calculate the number *k* of comparisons to be made as $k = t - 1$, for comparisons with the control, and $k = C_t^2 = t(t-1)/2$, for all paired comparisons. Considering that α is the global probability of a type I error for the total number of differences to be studied, the probability for each comparison will be given by $\alpha' = \alpha/(k - \hat{a})$. Here, $k' = (k - \hat{a})$ represent the number of differences in consonance with the next null hypothesis H_0 (due to $\hat{a}$).

The argument to accept that $\hat{a}$ is generally smaller than a is: if the treatments are ranked then the treatments that are situated far apart have differences that are probably statistically significant. Nevertheless, two treatments that are consecutive in the ordered set, due to the size of experimental error or smaller number of replications or other causes, have differences generally not significant. It is sufficient to have at least one or more situations like this to cause $\hat{a}$ to be smaller than a.

Using this procedure, Conagin and Barbin (2006a) proposed the modified Bonferroni's and Sidak's tests and Conagin *et al.* (2008) proposed the modified Tukey´s test. The probability of type I error will be $\alpha_B^{'} = \alpha/(k - \hat{a})$, where $k = t - 1$, for modified Bonferroni's test and $\alpha_S^{'} = 1 - (1 - \alpha)^{1/(k-\hat{a})}$ for modified Sidak´s test. For the modified Tukey´s test it is necessary to obtain the α point of the distribution of the Studentized range of t normal variables, and use $q(t - \hat{a}, \upsilon, \alpha)$ instead of $q(t, \upsilon, \alpha)$, where υ is the number of degrees of freedom of the error of the ANOVA table.

To have a modified unilateral Dunnett´s test, we need to obtain, from a Dunnett's test table, the value t_D with $p' = p - \hat{a}$, where $p = t - 1$, and error number of degrees of freedom, $\upsilon = (t-1)(r-1)$.

3 Results and discussion

For the simulated data, $a = 8$ for GI group because it was supposed that 8 means are greater than the control mean; also, $a = 8$ for GII group, by the same reason (although it was supposed that five others means are equal to the control mean).

For all experiments of a type, the values obtained for $\hat{a} = 8$ was not greater than $n = 14$ for GI group and not greater than $n = 12$ for GII group. So $\hat{a}$ can be safely used as an appropriate estimate of a in the modified tests. All the others $\hat{a}$ from the experiments are smaller than 8 and so $\hat{a}$ is a conservative estimate of a.

Considering $t = 9$, the critical values, in relation to the control mean, for $r = 4$, were 20%, 40% and above 40%, respectively, for coefficients of variation of 5%, 10% and 15%, while for $r = 8$, they were 15%, 35% and 40%. Considering $t = 14$, the critical values, in relation to the control mean, for $r = 4$, were 20%, 40% and above 40%, respectively, for coefficients of variation of 5%, 10% and 15%, while for $r = 8$, they were 15%, 30% and 40%. For all true differences between control and treatments above the critical values, all tests show equal power in detecting the differences; below the critical values, as the true differences between control and treatments decrease, there is an increasing difference in power between the tests. For the case of true differences below the critical values, the ranking of the tests (the first three with comparisonwise error rate type and the others with experimentwise error rate type) in power were in order: Student unilateral, Student bilateral, Duncan, modified unilateral Dunnett, modified Sidak, modified Bonferroni, modified Tukey, unilateral Dunnett, SNK, Tukey, Sidak and Bonferroni. The modified Dunnett's test revealed its superiority to the other three modified tests, to the SNK and to the remaining tests of experimentwise error rate type.

The results for the power of the tests studied, expressed as percentage of the number of significant differences in relation to the total number of simulations, are presented in Tables 1 to 6. They show that as differences between the mean of the treatments and the mean of the control decrease there are more differences in efficiency between tests. In terms of power the tests are decreasingly ordered as unilateral Student, bilateral Student, Duncan, modified unilateral Dunnett, modified Sidak, modified Bonferroni, modified Tukey, unilateral Dunnett, SNK, Tukey, Sidak and Bonferroni.

The efficiency of the modified tests in relation to the correspondent original ones and of the modified Dunnett's test in relation to other modified tests are presented in Tables 7 and 8. In all cases the modified tests are more efficiency than the original ones. The efficiency increased always from the critical level point, growing as the percentage of differences in relation to the control mean decreases; this is a good advantage because in making a good choice of the test to be used the researcher will be confident that will detect a greater number of small differences.

Table 1 - Power (%) of several tests for treatments differing from the control by 5 to 40%, CV = 5% and for four and eight replications (nine treatments)

Tests	*r* = 4								*r* = 8							
	40	**35**	**30**	**25**	**20**	**15**	**10**	**5**	**40**	**35**	**30**	**25**	**20**	**15**	**10**	**5**
T uni.	100	100	100	100	100	98	68	23	100	100	100	100	100	100	90	43
T bil.	100	100	100	100	100	92	60	20	100	100	100	100	100	100	90	37
Duncan	100	100	100	100	100	92	55	17	100	100	100	100	100	100	90	37
SNK	100	100	100	100	100	80	47	13	100	100	100	100	100	100	90	40
Tukey	100	100	100	100	90	65	12	2	100	100	100	100	100	97	50	7
Sidak	100	100	100	100	88	57	10	2	100	100	100	100	100	97	43	3
Bonfer	100	100	100	100	85	57	10	2	100	100	100	100	100	97	43	3
Dun uni	100	100	100	100	100	85	43	7	100	100	100	100	100	100	80	23
BonferM	100	100	100	100	100	92	58	25	100	100	100	100	100	100	87	50
DunnM	100	100	100	100	100	92	60	22	100	100	100	100	100	100	80	47
TukeyM	100	100	100	100	100	83	53	20	100	100	100	100	100	100	90	37
SidakM	100	100	100	100	100	88	60	23	100	100	100	100	100	100	90	50

Table 2 - Power (%) of several tests for treatments differing from the control by 5 to 40%, CV = 10% and for four and eight replications (nine treatments)

Tests	*r* = 4								*r* = 8							
	40	**35**	**30**	**25**	**20**	**15**	**10**	**5**	**40**	**35**	**30**	**25**	**20**	**15**	**10**	**5**
T uni.	100	100	98	88	73	62	30	20	100	100	100	100	100	100	53	13
T bil.	100	100	97	83	65	45	15	10	100	100	100	97	93	77	43	3
Duncan	100	98	88	73	62	37	13	10	100	100	100	97	87	70	37	3
SNK	95	78	67	47	42	20	5	7	100	100	100	97	73	47	17	3
Tukey	93	78	62	30	32	12	2	5	100	100	97	83	60	27	3	3
Sidak	93	73	50	22	23	7	2	5	100	97	93	83	63	27	3	3
Bonfer	93	74	48	22	23	7	2	5	100	97	93	83	63	27	3	3
Dun uni	100	92	78	67	50	27	5	5	100	100	100	97	83	63	7	3
BonferM	100	95	78	70	57	38	15	13	100	100	100	97	90	77	33	10
DunnM	100	97	85	72	57	38	15	12	100	100	100	100	97	83	47	13
TukeyM	100	85	70	60	48	25	10	10	100	100	100	97	83	60	33	7
SidakM	100	92	80	68	55	37	15	13	100	100	100	97	90	73	40	7

Table 3 - Power (%) of several tests for treatments differing from the control by 5 to 40%, CV = 15% and for four and eight replications (nine treatments)

Tests	*r* = 4								*r* = 8							
	40	**35**	**30**	**25**	**20**	**15**	**10**	**5**	**40**	**35**	**30**	**25**	**20**	**15**	**10**	**5**
T uni.	100	85	82	78	58	32	25	15	100	97	93	90	80	57	37	20
T bil.	87	78	65	65	33	18	10	8	100	93	93	87	70	43	23	10
Duncan	87	80	58	57	32	17	12	7	100	90	90	83	70	40	17	7
SNK	52	40	23	18	5	3	8	2	93	77	70	70	40	20	13	3
Tukey	47	30	17	13	3	2	2	0	93	77	67	53	17	7	7	0
Sidak	35	18	18	8	5	0	0	0	90	70	53	43	17	3	3	0
Bonfer	35	18	17	7	5	0	0	0	90	70	53	43	17	3	3	0
Dun uni	82	67	50	42	22	8	7	2	100	93	87	83	63	30	17	3
BonferM	78	73	53	48	28	13	8	7	100	90	90	83	70	43	27	13
DunnM	85	82	67	65	40	20	10	8	100	90	87	83	73	43	33	17
TukeyM	63	48	35	37	17	7	5	8	97	83	70	73	50	33	17	10
SidakM	82	80	57	52	30	17	8	8	100	90	87	83	70	40	27	13

Table 4 - Power (%) of several tests for treatments differing from the control by 5 to 40%, CV = 5% and for four and eight replications (14 treatments)

Tests	*r* = 4								*r* = 8							
	40	**35**	**30**	**25**	**20**	**15**	**10**	**5**	**40**	**35**	**30**	**25**	**20**	**15**	**10**	**5**
T uni.	100	100	100	100	100	100	65	18	100	100	100	100	100	100	97	40
T bil.	100	100	100	100	100	100	62	12	100	100	100	100	100	100	93	33
Duncan	100	100	100	100	100	100	43	8	100	100	100	100	100	100	87	33
SNK	100	100	100	100	97	83	30	2	100	100	100	100	100	100	80	20
Tukey	100	100	100	100	85	45	13	2	100	100	100	100	100	100	63	3
Sidak	100	100	100	100	78	43	12	2	100	100	100	100	100	97	63	3
Bonfer	100	100	100	100	78	43	12	2	100	100	100	100	100	97	63	3
Dun uni	100	100	100	100	100	87	38	3	100	100	100	100	100	100	90	10
BonferM	100	100	100	100	100	83	33	3	100	100	100	100	100	100	90	30
DunnM	100	100	100	100	100	90	38	5	100	100	100	100	100	100	90	30
TukeyM	100	100	100	100	93	72	30	2	100	100	100	100	100	93	73	17
SidakM	100	100	100	100	100	85	43	5	100	100	100	100	100	100	90	37

Table 5 - Power (%) of several tests for treatments differing from the control by 5 to 40%, CV = 10% and for four and eight replications (14 treatments)

Tests	*r* = 4								*r* = 8							
	40	**35**	**30**	**25**	**20**	**15**	**10**	**5**	**40**	**35**	**30**	**25**	**20**	**15**	**10**	**5**
T uni.	100	100	100	100	82	60	35	13	100	100	100	100	97	67	43	17
T bil.	100	98	97	97	75	47	23	10	100	100	100	100	90	53	30	10
Duncan	100	98	97	92	70	35	15	7	100	100	100	97	77	30	7	0
SNK	100	92	73	48	37	13	7	2	100	100	97	90	63	13	3	0
Tukey	97	85	60	37	23	8	2	0	100	100	97	90	50	13	0	0
Sidak	95	82	50	35	20	8	2	0	100	100	97	90	50	13	0	0
Bonfer	97	80	48	35	20	8	2	0	100	100	100	100	87	37	13	7
Dun uni	100	95	93	77	60	20	12	5	100	100	100	100	87	43	20	10
BonferM	98	93	88	75	60	23	10	5	100	100	100	100	90	50	37	10
DunnM	100	95	95	83	70	37	13	5	100	100	100	100	97	56	40	13
TukeyM	100	88	72	45	35	10	8	3	100	100	100	93	73	30	7	7
SidakM	100	97	93	87	65	25	12	5	100	100	100	100	87	43	33	7

Table 6 - Power (%) of several tests for treatments differing from the control by 5 to 40%, CV = 15% and for four and eight replications (14 treatments)

Tests	*r* = 4								*r* = 8							
	40	**35**	**30**	**25**	**20**	**15**	**10**	**5**	**40**	**35**	**30**	**25**	**20**	**15**	**10**	**5**
T uni.	98	85	80	70	55	30	30	12	100	93	93	87	73	40	33	13
T bil.	93	83	70	53	35	27	18	7	100	90	83	77	67	27	23	3
Duncan	87	78	57	35	18	17	12	2	100	93	83	77	57	20	17	0
SNK	63	28	28	13	7	2	0	0	80	80	60	40	23	7	0	0
Tukey	57	23	18	10	5	0	0	0	93	73	57	40	20	3	0	0
Sidak	48	25	18	5	2	0	0	0	77	73	50	17	17	3	0	0
Bonfer	48	18	15	3	0	0	0	0	77	73	50	17	17	3	0	0
Dun uni	82	65	50	28	8	12	7	0	100	93	73	67	43	10	3	0
BonferM	73	55	52	27	10	12	8	0	100	93	73	63	37	10	3	0
DunnM	82	62	52	33	15	13	10	0	100	90	73	70	50	17	7	0
TukeyM	63	48	40	12	10	3	2	0	87	83	67	43	23	7	0	0
SidakM	77	60	47	28	10	12	8	0	100	93	73	63	40	10	3	0

Table 7 - Efficiency of the various modified tests in relation to the original ones and of the modified Dunnett's test in relation to the Tukey, Sidak and Bonferroni tests, for four and eight replications, coefficients of variation 5, 10 and 15%, for treatments differing from the control by 10 to 40% (nine treatments)

CV	Tests	*r* = 4				*r* = 8			
		40%	**30%**	**20%**	**10%**	**40%**	**30%**	**20%**	**10%**
	TuM/Tu	1.36	2.10	3.30	3.00	1.04	1.33	1.85	1.50
	SiM / Si	2.33	3.09	6.00	10.00	1.11	1.63	4.20	-
	BoM/Bo	2.23	3.10	8.00	10.00	1.11	1.67	4.20	-
15%	DuM/Du	1.04	1.33	1.85	1.50	1.00	1.00	1.16	2.00
	DuM/Tu	1.82	4.00	12.00	6.00	1.07	1.80	4.40	5.00
	DuM/Si	2.42	3.64	8.00	12.00	1.11	1.63	4.40	-
	DuM/Bo	2.42	4.00	8.00	12.00	1.11	1.63	4.40	-
	TuM/Tu	1.07	1.14	1.53	6.00	1.00	1.03	1.25	-
	SiM/Si	1.07	1.60	2.36	9.00	1.00	1.07	1.42	-
	BoM/Bo	1.07	1.62	2.43	9.00	1.00	1.07	1.42	-
10%	DuM/Du	1.00	1.09	1.13	3.00	1.00	1.00	1.16	7.00
	DuM/Tu	1.07	1.38	1.79	9.00	1.00	1.03	1.45	-
	DuM/Si	1.07	1.70	2.43	9.00	1.00	1.07	1.53	-
	DuM/Bo	1.07	1.76	2.43	9.00	1.00	1.07	1.53	-
	TuM/Tu	1.00	1.00	1.11	4.57	1.00	1.00	1.00	1.80
	SiM/Si	1.00	1.00	1.13	6.00	1.00	1.00	1.00	2.08
	BoM/Bo	1.00	1.00	1.18	5.83	1.00	1.00	1.00	2.00
5%	DuM/Du	1.00	1.00	1.00	1.44	1.00	1.00	1.00	1.00
	DuM/Tu	1.00	1.00	1.11	5.14	1.00	1.00	1.00	1.60
	DuM/Si	1.00	1.00	1.13	6.00	1.00	1.00	1.00	1.85
	DuM/Bo	1.00	1.00	1.18	6.00	1.00	1.00	1.00	1.85

Table 8 - Efficiency of the various modified tests in relation to the original ones and of the modified Dunnett's test in relation to the Tukey, Sidak and Bonferroni tests, for four and eight replications, coefficients of variation 5, 10 and 15%, for treatments differing from the control by 10 to 40% (14 treatments)

CV	Tests	*r* = 4				*r* = 8			
		40%	**30%**	**20%**	**10%**	**40%**	**30%**	**20%**	**10%**
15%	TuM/Tu	1.11	2.19	2.00	5.00	1.00	1.18	1.17	1.00
	SiM/Si	1.58	2.54	6.00	-	1.30	1.47	2.40	2.00
	BoM/Bo	1.52	3.44	-	-	1.30	1.47	2.20	2.00
	DuM/Du	1.00	1.03	1.80	1.50	1.00	1.00	1.15	2.00
	DuM/Tu	1.44	2.82	3.00	6.00	1.15	1.29	2.50	4.00
	DuM/Si	1.69	2.82	-	-	1.30	1.47	3.00	4.00
	DuM/Bo	1.69	3.44	-	-	1.30	1.47	3.00	4.00
10%	TuM/Tu	1.03	1.19	1.50	5.00	1.00	1.03	1.47	4.00
	SiM/Si	1.05	1.87	3.25	7.00	1.00	1.03	1.73	-
	BoM/Bo	1.02	1.83	3.00	6.00	1.00	1.00	1.04	2.75
	DuM/Du	1.00	1.02	1.17	1.14	1.00	1.00	1.12	2.67
	DuM/Tu	1.03	1.58	3.00	8.00	1.00	1.03	1.93	-
	DuM/Si	1.05	1.73	3.50	8.00	1.00	1.03	1.93	-
	DuM/Bo	1.03	1.97	3.50	8.00	1.00	1.00	1.11	4.00
5%	TuM/Tu	1.00	1.00	1.10	2.25	1.00	1.00	1.00	1.16
	SiM/Si	1.00	1.00	1.28	3.71	1.00	1.00	1.00	1.42
	BoM/Bo	1.00	1.00	1.28	2.86	1.00	1.00	1.00	1.42
	DuM/Du	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	DuM/Tu	1.00	1.00	1.18	2.87	1.00	1.00	1.00	1.42
	DuM/Si	1.00	1.00	1.28	3.28	1.00	1.00	1.00	1.42
	DuM/Bo	1.00	1.00	1.28	3.28	1.00	1.00	1.00	1.42

Conclusions

Student's unilateral test was more efficient than Student bilateral, Duncan's test and to the other tests and should be preferred if the researcher wants to accept the comparisonwise type I error criterion.

The modified tests were, always, more efficient than the correspondent original tests for differences between means and the control mean in the region below the critical value.

Modified Dunnett's test, in the region of small differences, is the most efficient test in comparison with all remaining experimentwise error rate type tests included in this study.

CONAGIN, A.; BARBIN, D.; DEMÉTRIO, C. G. B. Teste de Dunnett modificado para delineamentos em blocos completos casualizados. *Rev. Bras. Biom.*, São Paulo, v.29, n.4, p.599-610, 2011.

- RESUMO: O teste de Dunnett, em geral, é preferido para a comparação de diferenças pareadas entre médias de tratamentos e um controle após uma análise de variância (ANOVA). Neste artigo, apresentamos o teste de Dunnett modificado e avaliamos sua performance em relação ao teste original de Dunnett e a outros dez testes, considerando experimentos em blocos casualizados. O método de simulação Monte Carlo foi usado para gerar dados de uma distribuição normal para 540 experimentos, variando o número de tratamentos (dois grupos, GI: com nove tratamentos e GII: com 14 tratamentos), número de repetições (360 experimentos com quatro blocos e 180 experimentos com oito blocos) e tamanho do erro (três coeficientes de variação: 5%, 10% e 15%), resultando 12 combinações. Verificou-se que à medida que diminui a diferença entre as medias de tratamentos e o controle há uma clara diferenciação em poder entre os testes, e que os testes modificados são mais eficientes que os testes originais. O teste de Dunnett modificado, na região de pequenas diferenças, é o mais eficiente entre os testes de taxa de erro por experimento (MEER) em comparação com todos os outros testes desse tipo.
- PALAVRAS-CHAVE: Testes de comparações múltiplas; Método Monte Carlo; Simulação; Testes modificados.

References

BERNARDSON, C. S. Type I error rates when multiple comparison procedures follow a significant F test of ANOVA. *Biometrics*, Oxford, v.31, n.1, p.229-232, 1975.

BOARDMAN, T. J.; MOFFITT, D. R. Graphical Monte Carlo type I error rates for multiple comparison procedures. *Biometrics*, Oxford, v.27, n.3, p.738-744, 1971.

CONAGIN, A.; BARBIN, D. Bonferroni's and Sidak's modified tests. *Sci. Agric.*, Piracicaba, v.63, n.1, p.70-76, 2006a.

CONAGIN, A.; BARBIN, D. Poder e eficiência dos diferentes testes estatísticos para comparações múltiplas. *Rev. Agric.*, Piracicaba, v.81, p.118-137, 2006b.

CONAGIN, A.; BARBIN, D.; DEMÉTRIO, C. G. B. Modifications for the Tukey test procedure and evaluation of the power and efficiency of multiple comparison procedures. *Sci. Agric.*, Piracicaba, v.65, n.3, p.428-432, 2008.

DACHS, J. N. *Estatística computacional*: uma introdução em turbo pascal. Rio de Janeiro: Livros Técnicos e Científicos, 1988. 126p.

DEMIRHAN, H.; DOLGUN, N. A.; DEMIRHAN, P.; DOLGUN, M. O. Performance of some multiple comparison tests under heteroscedasticity and dependency. *J. Stat. Comput. Simul.*, Essex, v.80, n.10, p.1083-1100, 2010.

DUNNETT, A. A multiple comparison procedure for comparing several treatments with a control. *J. Am. Stat. Assoc.*, Essex, v.50, n.2, p.1096-1121, 1955.

GABRIEL, K. R. A. A procedure for treating the homogeneity of all set of means in analysis of variance. *Biometrics*, Oxford, v.20, n.3, p.459-477, 1964.

HAYTER, A. J. The Maximum Familywise error rate of Fisher's least significant difference test. *J. Am. Stat. Assoc.*, Essex, v.81, n.396, p.1000-1004, 1986.

HOCKING, R. R. *The analysis of linear models.* Belmont: Brooks/Cole, 1985. p.385.

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

NAYLOR, T. H.; BALINTFY, J. L.; BURDICK, D. S., CHU, K. 1966. *Computer simulation techniques*. New York : John Wiley Sons, 1996.

O'NEILL, R.; WETHERILL, G. B. The present state of multiple comparison methods. *J. R. Stat. Soc., Ser. B Methodol.*, Oxford, v.33, n.2, p.218-250, 1971.

PERECIN, D.; BARBOSA, J. C. Uma avaliação de seis procedimentos para comparações múltiplas. *Rev. Mat. Estat.*, São Paulo, v.6, p.95-103, 1988.

PODLICH, D. W.; COOPER, M. *QU-GENE*: a platform for quantitative analysis of genetic models. *Bioinformatics*, Oxford, v.14, n.7, p.632-653, 1998.

RAMSEY, P. H.; RAMSEY, P. P.; BARRERA, K. Choosing the best pairwise comparisons of means from nom-normal populations, with unequal variances, but equal sample sizes. *J. Stat. Comput. Simul.*, Essex, v.80, n.5, p.595-608, 2010.

RAMSEY, P. H.; BARRERA, K.; HACHIMINE-SEMPREBOM, P.; LIU, C.-C. Pairwise comparisons of means under realistic non-normality, unequal variances, outliers and equal sample sizes. *J. Stat. Comput. Simul.*, Essex, v.81, n.2, p.125-135, 2011.

SAS INSTITUTE. System for Microsoft Windows, release 8. Cary, 1999.

SANTOS, C.; FERREIRA, D. F.; BUENO FILHO, J. S. S. Novas alternativas de testes de agrupamentos avaliadas por meio de simulação Monte Carlo. *Ciênc. Agrotec.*, Lavras, v.25, n.6, p.1382-1392, 2001.

SHIMIZU, T. *Pesquisa operacional em engenharia, economia e administração*: modelos básicos e métodos computacionais. Rio de Janeiro: Guanabara Dois, 1984. 360p.

Received in 30.12.2011

Approved after revised in 05.03.2012