# DISCRIMINATIVE POWER OF A MODIFIED BONFERRONI'S TEST

**Armando Conagin**[1]

## INTRODUCTION

The problem of choosing the more adequate statistical test for comparison of means of two treatments in research work is dependent on the type of error adopted, whether comparisonwise or experimentwise, the latter under general and/or partial null hypothesis.

It is well known that LSD's and Duncan's multiple range test are of the **comparisonwise type**, SNK's is of the **experimentwise type** under general null hypothesis, and that Tukey's, Bonferroni and Dunnett's tests are of the **experimentwise type** under general and/or partial null hypothesis (SAS, 1990).

Comparisonwise Type I error rate ($H_0$ true) is the ratio of the number of comparisons incorretly declared significant, divided by the total number of comparisons tested (STEEL & TORRIE, 1981).

Experimentwise Type I error rate ($H_0$ true) is the ratio of the number of experiments with one or more comparisons incorrectly declared significant, divided by the total number of experiments conducted (STEEL & TORRIE, 1981).

The objective of the present research is to evaluate the discriminative power of a new test, called **Modified Bonferroni's Test** and compare it with the tests cited above.

From now on we are denoting Fisher's Test by (LSD), Duncan's by (D), Student Newman-Keuls by (SNK), Tukey's by (Tu), Bonferroni's by (B), Dunnett's by (Du) and the Modified

[1] Agronomist and applied statistician, retired from the Instituto Agronômico, Caixa Postal 28, 13001-970, Campinas, SP, Brazil.

Bonferroni's by (BM). Scheffé's test used more effectively for complex contrasts and Waller-Duncan's test, not very much used so far, were not included is this research. We also performed the LSD (protected t test) and Duncan's protected test (DP).

PERECIN & BARBOSA (1988) carried out a simulation on various tests and appointed conclusions that are corroborated for some of the tests here utilized.

## MATERIAL AND METHODS

The design used was of the randomized block type.

The theoretical model used was:

$Y_{ij} = M + T_i + B_j + E_{ij}$ ; i = 1,2,... t; j = 1,2,...r.

The values of the parameters were: M = 3000, the $T_1$'s values were 40%, 30%, 20%, 10%, etc., of the M value.

Four groups, $G_1$, $G_2$, $G_3$ and $G_4$ were studied. For $G_1$. $T_1$ = 40%, $T_2$ = 10%, $T_3$ = 0% and $T_4$ = 0% (control), where k = 2 ("a priori" selected comparisons with the control), t=4, r=3 and r=6 and CV = 10%, CV = 15% and CV = 20% (coefficients of variation) of M. For $G_2$, $T_1$ = 30%, $T_2$ = 20%, $T_3$ = 0% and $T_4$ = 0% (control), in all combinations, as above (k=2).

For $G_3$, $T_1$ = 40%, $T_2$ = 30%, $T_3$ = 20%, $T_4$ = 10%, $T_5$ = 0% and $T_6$ = 0% (control). Now k = 4, t = 6, and all combinations were performed.

For $G_4$, $T_1$ = 40%, $T_2$ = 30%, $T_3$ = 20%, $T_4$ = 10% , $T_5$ = 0%, $T_6$ = 0% (control), $T_7$ = 5%, $T_8$ = 3%, $T_9$ = 1%, $T_{10}$ = -5%, $T_{11}$ = -3% and $T_{12}$ = -1%. Here, again, k=4, t=12, and all combinations of CV's and r's were performed.

The $B_j$'s values were the same in all sets.

The $E'_{ij}s$ were obtained by the RANNOR FUNCTIONS of the SAS, 1990.

We performed 200 experiments for each combination when r=3 and 100 experiments for each combination when r=6, taking together 100 pairs of the single experiments.

On the whole 2,400 experiments for r=3 were carried out and 1,200 experiments for r=6 type. We used SAS Program (SAS, 1990) to perform the simulations, the analyses of variance and the tests already specified.

The comparisons of treatments 1 (40%), 2 (30%), 3 (20%) and 4 (10%) with the control 6 (0%) evaluates the discriminative power of the respective contrast; the comparison of treatment 5 (0%) with the control 6 (0%) evaluates the size of type I error of the test between treatments 5 and 6. Since treatments 5 and 6 are equal, their difference is a simultaneous measure of comparisonwise and experimentwise type I error.

## THE MODIFIED BONFERRONI'S TEST

The usual Bonferroni's Test is used to establish confidence intervals or carry out a statistical test of significance for **k** comparisons chosen "a priori".

The means of two treatments, $\bar{y}_1$ and $\bar{y}_2$, are considered different according to Bonferroni's test, if

$$|\bar{y}_1 - \bar{y}_2| \geq t_{(\gamma_B, df)}.s\sqrt{2/r} \qquad \text{(SAS, 1990).}$$

The **t** is the Student **t** value for $\gamma_B$ level of significance, **df** degrees of freedom, **s** the standard deviation and **r** the number of replications for each mean.

To calculate $\gamma_B$, if $\alpha$ is the joint level of significance for the **k** comparisons (with values 0.05 or 0.01), then, for each of the **k** comparisons, $\gamma_B = \alpha/k$.

If the number of treatments in the experiment is **t**, the $\gamma_B$ value adequate for comparisons of all possible differences between two means is $\gamma_B = \alpha/t(t-1) \div 2$, where $\mathbf{k} = C_t^2$ is the number of all possible combinations of **t** means taken two at a time.

In the modified Bonferroni's Test, for the comparisons of **k** differences between two means, we should start testing the general null hypothesis in the Analysis of Variance. If the null hypothesis is rejected at $\alpha$ level, that is, if $F_o$ obtained is greater than the $F_c$ critical level, we perform the Modified Bonferroni's Test.

The test assumes that $\gamma_{BM} = \alpha\ (\ 1 + P(F))/k$ (A) .

The calculation of P(F) is as follows:

Let us assume that $F_o$ = MS treatments/MS residual in the analysis of variance of the experiment.

We know that:

EMS treatments $= \sigma^2 + (r \Sigma t_i^2)/(t-1)$,

EMS residual $= \sigma^2$.

If $H_o$ (general null hypothesis) is true, then

$T_1 = T_2 = ... = T_t = 0$.

If $H_a$ (general alternative hypothesis) is true, at least one of the Ti's is different from zero.

The non-centrality parameter $\lambda = r \Sigma T_i^2/\sigma^2$ is a parameter of the non-central F distribution (WINNER **et al.**, 1991).

In the analysis of the experiment,

$F_o = (s^2 + r \Sigma T_i^2 / (t-1))/s^2 = 1 + r \Sigma T_i^2/s^2 (t-1)$ .

An estimate $\hat{\lambda}$ of $\lambda$ is $\hat{\lambda} = r \Sigma T_i^2/s^2 = (F_o - 1)(t - 1)$ .

The critical value of F for the rejection of Ho is the value $F_c$ . If $F_o > F_c$ then $H_o$ is rejected.

The probability of values smaller than $F_o$ if $H_a$ is true is:

$$P(F_o): \int_0^{Fo} g(F_o, t\text{-}1), (t\text{-}1)\,(r\text{-}1), \lambda\,)\,dx .$$

The probability of values smaller than $F_c$ if $H_a$ is true is:

$$P(F_c) = \int_0^{Fc} g(F_c, t\text{-}1, (t\text{-}1)\,(r\text{-}1), \lambda)dx.$$

The values may be calculated from PROBF FUNCTION of the SAS Program (SAS, 1990).

We define

$$P(F) = P(F_o) - P(F_c)$$

$$= \int_{Fc}^{Fo} g(F_c, t\text{-}1, (t\text{-}1)\,(r\text{-}1), \lambda)\,dx.$$

which shows that $0 < P(F) < 1$.

As an example, assume that in an experiment, r=3, df=10, t=6, $F_o$=9.72, $F_c$ = 3.33 and k=4. Since $F_o > F_c$ the null hypothesis shall be rejected. The calculation needed is:

$$\hat{\lambda} = (t\text{-}1)\,(F_o\text{-}1) = 5(9.72 - 1) = 43.60 ,$$

$$P(F_o) = \int_0^{Fo} g(9.72, 5, 10, 43.60)\,dx = 0.4682 ,$$

$$P(F_c) = \int_0^{Fo} g(3.33, 5, 10, 43.60)dx = 0.0155.$$ Then:

$$P(F) = P(F_o) - P(F_c) = 0.4527 .$$

Using formula (A) ,

$$\gamma_{BM} = 0.05\,(1 + 0.4527)/4 = 0.01816 .$$

For the Bonferroni's test we have:

$$\gamma_B = 0.05/4 = 0.0125 .$$

The interval of variation of $\gamma_{BM}$ is: $\gamma_B < \gamma_{BM} < 2\gamma_B$ .

The argument to use $\gamma_{BM}$ is based on the following reasoning: "If the F ratio MS treat/MS residual is large ($F_o > F_c$), the null hypothesis is rejected and there is "evidence" of the existence of heterogenous treatments; therefore we do not need to be so rigorous in using $\alpha = 0.05$ for the joint comparisons because we already accepted that $H_o$ is false. We consider P(F) as the weight of "evidence"that $H_a$ is true. So we use $\alpha(1+P(F))$ as the joint level for **k** "a priori" comparisons.

Then $\gamma_{BM} = \alpha\ (1 + P(F))/k$ .

This type of reasoning was already used in the Waller-Duncan's test (CHEW, 1977).

To be consistent we should use the Modified Bonferroni's Test (BM) only if $H_o$ is rejected ($H_o$ if false). The structure of the test proposed incorporates the "weight of evidence" of the veracity of $H_a$ through the introduction of the $F_o$ value obtained in the analysis of variance for the calculation of the **t** value, used to judge **k** chosen "a priori" comparisons of means. The results presented in **Tables 1** to 6 regarding BM were obtained in this way.

In performing their test, Waller-Duncan calculate the test even if $H_o$ is not rejected. It would be possible in the Modified Test to use P(F) in the same way. As $P(F) = P(F_o) - P(F_c)$, if $F_o < F_c$ ($H_o$ not rejected), P(F) will be negative, and in this case $\gamma_{BM} < \gamma_B$; the test would be then more restrictive than Bonferroni's. When P(F) = 0, the two tests are identical.

## RESULTS

We define as **Discriminative Power** of a test as the percentage of significant contrasts among **k** taken "a priori" contrasts of two means. In the Tables the cells represent the significant percentage of the comparison between the specified mean and the control mean obtained in a large group of experiments. The **Tables 1** to **6** present the **discriminative power** of all tests performed, based in the conditions established in the research, the following conclusions are possible.

**Table 1.** Results obtained when r=3, CV = 10%, for k=2 and k=4, for groups $G_1 + G_2$, $G_3$ and $G_4$, and for df=6 (t=4), df=10 (t=6) and df=22 (t=12), for differences of 40%, 30%, 20%, 10%, and 0% between the respective treatment mean and the control mean. The values in the cells represent the **discriminative power** in percentage based on 200 simulated experiments.

| | $T_1 - T_6$ (40%) | | | $T_2 - T_6$ (30%) | | | $T_3 - T_6$ (20%) | | |
|---|---|---|---|---|---|---|---|---|---|
| df | 6 | 10 | 22 | 6 | 10 | 22 | 6 | 10 | 22 |
| LSD | 98 | 99 | 99 | 85 | 94 | 92 | 51 | 56 | 60.5 |
| LSDP | 92.5 | 97.5 | 98.5 | 75 | 93 | 92 | 49 | 56 | 60 |
| D | 98 | 99 | 97.5 | 83.5 | 91.5 | 88 | 50.5 | 51 | 49.5 |
| DP | 92.5 | 97.5 | 97.5 | 74.5 | 90.5 | 88 | 49 | 51 | 49.5 |
| SNK | 89.5 | 93.5 | 92 | 66.5 | 73 | 63 | 35 | 30.5 | 26 |
| Tu | 85.5 | 91 | 87.5 | 64 | 62.5 | 52 | 22.5 | 17 | 16.5 |
| B | 76.5 | 80.5 | 87 | 51 | 50 | 44.5 | 16 | 10.5 | 14 |
| Du | 91.5 | 95.5 | 94 | 69 | 71.5 | 73.5 | 30.5 | 26 | 33 |
| BM | 93 | 97.5 | 97 | 71 | 78 | 84.5 | 39 | 33 | 45 |

| | $T_4 - T_6$ (10%) | | | $T_5 - T_6$ (0%) | | |
|---|---|---|---|---|---|---|
| df | 6 | 10 | 22 | 6 | 10 | 22 |
| LSD | 20.5 | 19 | 22.5 | 5.5 | 3.5 | 3.5 |
| LSDP | 20.5 | 19 | 22.5 | 5.5 | 3.5 | 3.5 |
| D | 20 | 18.5 | 16.5 | 5.25 | 3 | 2.5 |
| DP | 20 | 18.5 | 16.5 | 5.25 | 3 | 2.5 |
| SNK | 13 | 11.5 | 3.5 | 5 | 2 | 0 |
| Tu | 8 | 2 | 2.5 | 1 | 0 | 0 |
| B | 6.5 | 0.5 | 2 | 0.75 | 0 | 0 |
| Du | 12 | 8.5 | 6 | 1.75 | 1 | 0 |
| BM | 16 | 9 | 10 | 4 | 4 | 3 |

**Table 2.** Results obtained when r=6, CV = 10%, for k=2 and k=4, for groups $G_1 + G_2$, $G_3$ and $G_4$, and for df=12 (t=4), df=20 (t=6) and df=44 (t=12), for differences of 40%, 30%, 20%, 10%, and 0% between the respective treatment mean and the control mean. The values in the cells represent the **discriminative power** in percentage based on 100 experiments each one resultant of grouping two single experiments.

| | $T_1 - T_6$ (40%) | | | $T_2 - T_6$ (30%) | | | $T_3 - T_6$ (20%) | | |
|---|---|---|---|---|---|---|---|---|---|
| df | 12 | 20 | 44 | 12 | 20 | 44 | 12 | 20 | 44 |
| LSD | 100 | 100 | 100 | 100 | 100 | 100 | 86 | 90 | 92 |
| LSDP | 100 | 100 | 100 | 100 | 100 | 100 | 86 | 90 | 92 |
| D | 100 | 100 | 100 | 100 | 99 | 99 | 85 | 87 | 87 |
| DP | 100 | 100 | 100 | 100 | 99 | 99 | 85 | 87 | 87 |
| SNK | 100 | 100 | 100 | 100 | 98 | 99 | 81 | 72 | 67 |
| Tu | 100 | 100 | 99 | 100 | 98 | 90 | 61 | 56 | 48 |
| B | 100 | 99 | 99 | 98 | 97 | 89 | 59 | 47 | 39 |
| Du | 100 | 100 | 100 | 100 | 98 | 99 | 71 | 72 | 66 |
| BM | 100 | 100 | 100 | 100 | 98 | 99 | 81 | 77 | 81 |

| | $T_4 - T_6$ (10%) | | | $T_5 - T_6$ (0%) | | |
|---|---|---|---|---|---|---|
| df | 12 | 20 | 44 | 12 | 20 | 44 |
| LSD | 44 | 35 | 41 | 5.5 | 7 | 9 |
| LSDP | 44 | 35 | 41 | 5.5 | 7 | 9 |
| D | 41 | 35 | 36 | 5.5 | 7 | 1 |
| DP | 41 | 35 | 36 | 5.5 | 7 | 1 |
| SNK | 34 | 26 | 10 | 5.5 | 5 | 0 |
| Tu | 16 | 7 | 7 | 0 | 2 | 0 |
| B | 14 | 4 | 5 | 0 | 2 | 0 |
| Du | 28 | 12 | 20 | 1.5 | 2 | 0 |
| BM | 38 | 17 | 35 | 3.5 | 2 | 0 |

**Table 3.** Results obtained when r=3, CV = 10%, for k=2 and k=4, for groups $G_1 + G_2$, $G_3$ and $G_4$, and for df=6 (t=4), df=10 (t=6) and df=22 (t=12), for differences of 40%, 30%, 20%, 10%, and 0% between the respective treatment mean and the control mean. The values in the cells represent the **discriminative power** in percentage based in 200 simulated experiments.

| | $T_1 - T_6$ (40%) | | | $T_2 - T_6$ (30%) | | | $T_3 - T_6$ (20%) | | |
|---|---|---|---|---|---|---|---|---|---|
| df | 6 | 10 | 22 | 6 | 10 | 22 | 6 | 10 | 22 |
| LSD | 77.5 | 86 | 90.5 | 58 | 58.5 | 66 | 29 | 31.5 | 39 |
| LSDP | 58.5 | 74 | 83.5 | 42.5 | 54 | 63 | 25 | 28 | 37.5 |
| D | 75 | 62 | 87 | 57 | 54 | 57.5 | 28 | 26.5 | 30.5 |
| DP | 57.5 | 53 | 81.5 | 42 | 50.5 | 55 | 24.5 | 24 | 29.5 |
| SNK | 53 | 54 | 55 | 34 | 28 | 21 | 16 | 11.5 | 6 |
| Tu | 47.5 | 48.5 | 43 | 29.5 | 20 | 14 | 14 | 6.5 | 4.5 |
| B | 36.5 | 27.5 | 32 | 21.5 | 16 | 10 | 8.5 | 4 | 2 |
| Du | 56 | 61 | 67.5 | 37.5 | 33.5 | 34 | 19 | 12.5 | 11 |
| BM | 56 | 62 | 78.5 | 41 | 34.5 | 45 | 17 | 13.5 | 18.5 |

| | $T_4 - T_6$ (10%) | | | $T_5 - T_6$ (0%) | | |
|---|---|---|---|---|---|---|
| df | 6 | 10 | 22 | 6 | 10 | 22 |
| LSD | 12 | 6.5 | 14.5 | 5.75 | 4.5 | 6 |
| LSDP | 12 | 6 | 14.5 | 5.5 | 4 | 6 |
| D | 12 | 6 | 10 | 5.5 | 3 | 2 |
| DP | 12 | 5.5 | 10 | 5.5 | 2.5 | 2 |
| SNK | 8.5 | 2.5 | 1 | 4 | 2 | 0.5 |
| Tu | 5 | 0.5 | 0.5 | 1.25 | 0 | 0 |
| B | 3 | 0.5 | 0.5 | 1.25 | 0 | 0 |
| Du | 7.5 | 2 | 5 | 2 | 0 | 0 |
| BM | 9.5 | 3 | 7.5 | 2.5 | 1 | 2 |

**Table 4.** Results obtained when r=6, CV = 15%, for k=2 and k=4, for groups $G_1 + G_2$, $G_3$ and $G_4$, and for df=12 (t=4), df=20 (t=6) and df=44 (t=12), for differences of 40%, 30%, 20%, 10%, and 0% between the respective treatment mean and the control mean. The values in the cells represent the **discriminative power** in percentage based on 100 experiments from the grouping of two single experiments.

| | $T_1 - T_6$ (40%) | | | $T_2 - T_6$ (30%) | | | $T_3 - T_6$ (20%) | | |
|---|---|---|---|---|---|---|---|---|---|
| df | 12 | 20 | 44 | 12 | 20 | 44 | 12 | 20 | 44 |
| LSD | 98 | 100 | 100 | 92 | 95 | 95 | 67 | 55 | 62 |
| LSDP | 93 | 100 | 100 | 89 | 95 | 95 | 67 | 55 | 62 |
| D | 98 | 100 | 99 | 91 | 93 | 91 | 64 | 49 | 53 |
| DP | 93 | 100 | 99 | 89 | 93 | 91 | 64 | 49 | 53 |
| SNK | 91 | 95 | 97 | 81 | 75 | 75 | 51 | 31 | 24 |
| Tu | 87 | 93 | 95 | 78 | 65 | 51 | 33 | 20 | 16 |
| B | 85 | 90 | 95 | 70 | 57 | 46 | 27 | 15 | 12 |
| Du | 93 | 96 | 98 | 80 | 73 | 76 | 45 | 30 | 29 |
| BM | 91 | 98 | 98 | 84 | 80 | 87 | 55 | 32 | 45 |

| | $T_4 - T_6$ (10%) | | | $T_5 - T_6$ (0%) | | |
|---|---|---|---|---|---|---|
| df | 12 | 20 | 44 | 12 | 20 | 44 |
| LSD | 23 | 13 | 21 | 3.5 | 3 | 3 |
| LSDP | 23 | 13 | 21 | 3.5 | 3 | 3 |
| D | 20 | 11 | 20 | 3.5 | 3 | 3 |
| DP | 20 | 11 | 20 | 3.5 | 3 | 3 |
| SNK | 15 | 4 | 6 | 3.5 | 2 | 1 |
| Tu | 8 | 2 | 2 | 1.5 | 0 | 0 |
| B | 7 | 1 | 2 | 1 | 0 | 0 |
| Du | 12 | 4 | 9 | 3 | 1 | 1 |
| BM | 18 | 5 | 12 | 3.5 | 2 | 1 |

**Table 5.** Results obtained when r=3, CV = 20%, for k=2 and k=4, for groups $G_1 + G_2$, $G_3$ and $G_4$, and for df=6 (t=4), df=10 (t=6) and df=22 (t=12), for differences of 40%, 30%, 20%, 10%, and 0% between the respective treatment mean and the control mean. The values in the cells represent the **discriminative power** in percentage based on 200 simulated experiments.

| | $T_1 - T_6$ (40%) | | | $T_2 - T_6$ (30%) | | | $T_3 - T_6$ (20%) | | |
|---|---|---|---|---|---|---|---|---|---|
| df | 6 | 10 | 22 | 6 | 10 | 22 | 6 | 10 | 22 |
| LSD | 52.5 | 57.5 | 67 | 32 | 41 | 42.5 | 15.5 | 22.5 | 16.5 |
| LSDP | 35 | 39.5 | 44 | 18 | 30 | 30.5 | 9 | 18.5 | 12.5 |
| D | 49.5 | 53 | 57 | 28.5 | 36 | 31.5 | 14.5 | 20.5 | 11.5 |
| DP | 35 | 38 | 41 | 17 | 28.5 | 23 | 9 | 16.5 | 9.5 |
| SNK | 30.5 | 27.5 | 19.5 | 12 | 14.5 | 3.5 | 8 | 8 | 1.5 |
| Tu | 24 | 25 | 16.5 | 8.5 | 12 | 3 | 6.5 | 4 | 1.5 |
| B | 18.5 | 20.5 | 12.5 | 5.5 | 8.5 | 1.5 | 3.5 | 2.5 | 1 |
| Du | 33 | 36.5 | 32.5 | 14 | 17.5 | 9.5 | 8 | 8.5 | 6 |
| BM | 32 | 35 | 39.5 | 14 | 19.5 | 17.5 | 7.5 | 9 | 9 |

| | $T_4 - T_6$ (10%) | | | $T_5 - T_6$ (0%) | | |
|---|---|---|---|---|---|---|
| df | 6 | 10 | 22 | 6 | 10 | 22 |
| LSD | 9 | 10.5 | 11.5 | 3.75 | 6 | 5.5 |
| LSDP | 9 | 9 | 8.5 | 3.75 | 4 | 5 |
| D | 9 | 9 | 6 | 3.25 | 5.5 | 3 |
| DP | 9 | 7.5 | 5.5 | 3.25 | 4 | 2.5 |
| SNK | 4 | 3 | 0.5 | 2 | 2.5 | 0 |
| Tu | 0.5 | 1.5 | 0.5 | 1 | 0.5 | 0 |
| B | 0 | 0.5 | 0 | 0.75 | 0 | 0 |
| Du | 2.5 | 4 | 1 | 1.5 | 1.5 | 1 |
| BM | 6 | 3.5 | 4 | 2 | 3 | 3 |

**Table 6.** Results obtained when r=6, CV = 20%, for k=2 and k=4, for groups $G_1 + G_2$, $G_3$ and $G_4$, and for df=12 (t=4), df=20 (t=6) and df=44 (t=12), for differences of 40%, 30%, 20%, 10%, and 0% between the respective treatment mean and the control mean. The values in the cells represent the **discriminative power** in percentage based on 100 experiments from the grouping of two single experiments.

| | $T_1 - T_6$ (40%) | | | $T_2 - T_6$ (30%) | | | $T_3 - T_6$ (20%) | | |
|---|---|---|---|---|---|---|---|---|---|
| df | 12 | 20 | 44 | 12 | 20 | 44 | 12 | 20 | 44 |
| LSD | 90 | 91 | 93 | 71 | 68 | 72 | 31 | 40 | 32 |
| LSDP | 83 | 80 | 92 | 56 | 62 | 72 | 31 | 38 | 32 |
| D | 90 | 87 | 88 | 68 | 63 | 60 | 28 | 35 | 24 |
| DP | 83 | 78 | 87 | 55 | 60 | 60 | 28 | 34 | 24 |
| SNK | 77 | 65 | 60 | 42 | 43 | 25 | 22 | 20 | 7 |
| Tu | 71 | 62 | 50 | 34 | 34 | 17 | 17 | 15 | 5 |
| B | 65 | 57 | 43 | 30 | 30 | 14 | 15 | 14 | 3 |
| Du | 78 | 74 | 75 | 52 | 46 | 36 | 20 | 19 | 12 |
| BM | 78 | 72 | 79 | 48 | 48 | 52 | 26 | 27 | 20 |

| | $T_4 - T_6$ (10%) | | | $T_5 - T_6$ (0%) | | |
|---|---|---|---|---|---|---|
| df | 12 | 20 | 44 | 12 | 20 | 44 |
| LSD | 13 | 15 | 17 | 2 | 8 | 4 |
| LSDP | 13 | 15 | 17 | 2 | 8 | 4 |
| D | 12 | 14 | 8 | 2 | 8 | 4 |
| DP | 12 | 14 | 8 | 2 | 8 | 4 |
| SNK | 6 | 11 | 1 | 2 | 6 | 0 |
| Tu | 2 | 9 | 0 | 1 | 2 | 0 |
| B | 2 | 6 | 0 | 1 | 2 | 0 |
| Du | 3 | 11 | 3 | 1.5 | 2 | 0 |
| BM | 8 | 12 | 6 | 2.5 | 3 | 3 |

The Tables show that the **discriminative power** of Fisher's LSD and Duncan's test under the form unprotected and protected are always higher than that of other tests (because they are of "comparisonwise type").

With the exception of a few cases (that we attribute to sampling variation), the Modified Bonferroni's test exhibits **discriminative power** greater than Dunnett's and SNK's tests, and much higher **discriminative power** than Tukey's and Bonferroni's tests.

According with the Statistical Theory we observe that for larger differences (40% and 30%) all tests exhibit high **discriminative power** for all conditions.

Higher coefficients of variation (CV) are associated with lower **discriminative power** for all statistical tests.

Greater number of replications (r=6) exhibits higher **discriminative power** in relation to r=3, for all tests.

Smaller coefficient of variation (10%) and higher number of replications (r=6) guarantee **discriminative power** superior to 80% for almost all types of test for differences equal or higher than 20%, with significant results in 4 of 5 experiments.

When differences are smaller (20% and 10%) the **discriminative power** is lower for all tests in almost all cases.

Many research workers prefer to use the **comparisonwise type** of error when testing differences between means (CHEW, 1977).

In cases in which the error of rejecting $H_0$ when it is true is judged important and must be avoided the **experimentwise error** should be preferred.

For the present situation, based in results when k=2 and k=4 and for different number of treatments, replications and degrees of freedon of the residual, we may, assume that the Modified Bonferroni's test behaved better than the group of **experimentwise tests** because the new test improved the **discriminative power,**

showing higher efficiency in the detection of the significance of the differences studied.


## RESUMO

O poder discriminativo de um novo teste estatístico, designado por **Teste de Bonferroni Modificado**, foi comparado com o dos testes mais utilizados na comparação de médias. Mostra-se que para **k** contrastes, escolhidos "a priori", entre a média de um tratamento e a do controle (testemunha) o novo teste é mais discriminativo que os testes de Student – Newman-Keuls (SNK), de Tukey (Tu), de Bonferroni (B) e de Dunnett (Du), todos eles do tipo de erro **por experimento**, sendo, porém, menos discriminativo que os testes de Fisher (LSD), o de comparação múltipla de Duncan (D) e os testes LSD e de Duncan protegidos (LSDP e DP), que são do tipo de erro **por comparação**.

Os resultados do poder discriminativo dos testes estatísticos comparados para diferentes valores de **k,** de coeficiente de variação (CV) e de números de repetições (r) constam das Tabelas 1 a 6.



## SUMMARY

The discriminative power of a new statistical test, called **Modified Bonferroni's Test**, was compared with that of several of the most utilized statistical tests adequate for pairwise comparisons.

Is is shown that for **k** comparisons chosen "a priori", the test is more discriminative than Student-Newman-Keuls, Tukey's, Bonferroni's and Dunnett's, all of **experimentwise** type of error, but less discriminative than Fisher's LDS test and Duncan's multiple range test, without or with protection, these of the **comparisonwise** type.

## REFERENCES

CHEW, V., 1977. **Comparison Among Treatment Means in an Analysis of Variance.** USDA, Agricultural Research Service. 60p.

PERECIN, D. & J.C. BARBOSA, 1988. Uma Avaliação de Seis Procedimentos para Comparações Múltiplas. **Rev. Mat. Estat.,** São Paulo, **6:** 95-103.

SAS, 1990. **The GLM Procedure.** Cap. 24, p.892-996.

SAS, 1990. **SAS Functions.** Caps. 11, 12, p.521-634.

STEEL, R.G.D. & J.H. TORRIE, 1981. **Principles and Procedures of Statistics.** McGraw-Hill. 663p.

WINNER, B.J.; D.R. BROWN & K.M. MICHELS, 1991. **Statistical Principles in Design.** McGraw-Hill. 1057p. (Series in Psychology).

## ADDITIONAL READING

BERNHARDSON, C.A., 1975. Type I Error Rates When Multiple Comparison Procedures Follow a Significant Test of Anova. **Biometrics, 31:** 229-232.

CARNER, S.G. & M.R. SWANSON, 1973. Evaluation of Tem Pairwase Multiple Comparison Procedures by Monte Carlo Methods. **Jour. Amer. Stat. Assoc., 68:** 66-74.

DIXON, D.O. & D.B. DUNCAN, 1975. Minimum Bases Risk t-Intervals for Multiple Comparisons. **Jour. Amer. Stat. Assoc., 70:** 822-831.

DUNCAN, D.B., 1955. Multiple Range and Multiple F Tests. **Biometrics, 11:** 1-42.

DUNCAN, D.B., 1965. A Bayesian Approach to Multiple Comparisons. **Technometrics, 7:** 171-222.

DUNCAN, D.B., 1975. t Test and Intervals for Comparisons Suggested by the Data. **Biometrics, 31:** 339-359.

DUNN, O.J., 1961. Multiple Comparisons Among Means. **Jour. Amer. Stat. Assoc., 56:** 52-64.

DUNNETT, C.W., 1955. A Multiple Comparison Procedure for Comparing Several Treatment With a Control. **Jour. Amer. Stat. Assoc., 50:** 1096-1121.

FEDERER WALTER, T., 1955. **Experimental Design.** New York, MacMillan Co. 544p.

GABRIEL, K.R., 1964. A Procedure for Testing the Homogeneity of All Sets of Means in Analysis of Variance. **Biometrics, 20:** 459-477.

PIMENTEL-GOMES, F., 1985. **Curso de Estatística Experimental.** 13ª ed. São Paulo, Nobel. 466p.

SHAFER, W.D. & G.B. McREADY, 1975. A Modification of the Bonferroni Procedure on Contrasts Which Are Grouped Into Internally Independent Sets. **Biometrics, 31:** 227-228.