ผลงานทางวิชาการ

อำนาจการทดสอบในการทดสอบรายคู่เชิงซ้อน
Power in Multiple Testing

ดร.สุชาดา บวรกิติวงศ์

ภาควิชาวิจัยการศึกษา
คณะครุศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย
พ.ศ. 254 1

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

อำนาจการทดสอบในการทดสอบรายคู่เชิงซ้อน

Power in Multiple Testing

ดร.สุชาดา บวรกิติวงศ์*

Suchada Bowarnkitiwong, Ph.D.*

บทคัดย่อ

งานวิจัยนี้มีจุดมุ่งหมายเพื่อเปรียบเทียบอำนาจการทดสอบในการทดสอบสัมประสิทธิ์สหสัมพันธ์อย่าง
ง่ายเชิงซ้อน (เมทริกซ์สัมประสิทธิ์สหสัมพันธ์) โดยใช้ข้อมูลจำลองในการตรวจสอบอำนาจการทดสอบทั้ง 3 แบบ
คือ อำนาจการทดสอบเมื่อมีความสัมพันธ์จริงอย่างน้อย 1 คู่ เมื่อมีความสัมพันธ์จริงทุกคู่ และอำนาจการทดสอบ
โดยเฉลี่ย ผลการตรวจสอบปรากฏว่า เมื่อมีความสัมพันธ์จริงอย่างน้อย 1 คู่ วิธีการทั้ง 7 ให้ผลไม่ต่างกัน แต่ใน
กรณีที่เหลืออำนาจการทดสอบของ 6 วิธีที่ปรับปรุงจาก Bonferroni ให้ผลดีกว่าวิธีดั้งเดิม ผลที่ได้จาก 6 วิธี
ดังกล่าวไม่ต่างกันมากนัก โดยที่วิธีการของ Holm จะให้อำนาจการทดสอบต่ำสุด ในขณะที่วิธีการของ Holland-
Copenhaver (Step-up) ให้อำนาจการทดสอบสูงสุด

Abstract

Multiple hypotheses testing in the context of a correlation matrix is used to compare
the statistical power of the original Bonferroni with six modified Bonferroni procedures which
control the overall Type I error rate. Three definitions of statistical power are considered : 1)
detecting at least one true relationship, 2) detecting all true relationships, and 3) the average
power to detect true relationships. Simulation results show no difference between the seven
methods in detecting at least one true relationship; but all six modified Bonferroni procedures
are more powerful than the original Bonferroni procedure to detect all true relationship power
and average power. Among the six midified Bonferroni procedures, small differences were
observed, with the Holm procedure having the lowest power and the Holland-Copenhaver (step-
up) procedure having the highest power.

---

* Educational Research Department, Faculty of Education, Chulalongkorn University.

Introduction

Whenever an experiment involves collecting data from more than two groups or under more than two conditions we become involved in the problem of multiple comparisons—the problem of comparing each group with every other groups or arranging the results in rank order. This becomes a problem when we wish to assign a level of confidence or significance to our conclusions about the relationships among all of the populations involved. Classical methods such as the F test permit us only to reject the overall null hypothesis that all of the means are equal but they do not provide a procedure for comparing specific means with each others.

In the older psychological literature, this problem has been dealt with in a haphazard manner, without recognizing the issues involved. More recently, statistical procedures specifically designed for multiple comparisons have become available and have been discussed briefly in the psychological journals (McHugh & Ellis, 1995 : Stanley, 1957).

It has not been clear to many researchers, however, that there are several different methods with different basic assumptions or approaches. There are important questions of logic involved in the use of these methods and these issues have not been clearly faced in the behavioral and psychological literature.

Generally, whenever multiple tests are conducted simultaneously for testing individual hypothesis in a study, the criterion α level corresponding to each hypothesis is adjusted to control the experimentwise (overall) type I error rate. Many studies in this area have been done in the context of multiple comparisons of means (Hsiung & Olejnik, 1991; Seaman, et. al. 1990). However, concern for the control of experimentwise Type I error rate in other multiple testing contexts, such as correlation matrix, factorial ANOVA, multiple regression, multiple contingency tables, multiple experiments, and multiple outcome variable studies, has not received the same level of attention.

A common situation in behavioral, social, and educational research is to have data on a number of variables for a single sample of observations and to be concerned with some or all of the correlations between these variables in the population. For example, the researcher may have several predictor variables and may be interested in just the correlations of these variables with a criterion variable. Or, the researcher may have two set of variables—predictor variables

and criterion variables—and be interested in whether any relationships exist between the two sets. Or, the researcher may have several criterion variables and be interested in the extent to which these variables are related to one another. Or, in a few instances he/she may have several variables and an a priori theory about how these variables are correlated with one another. How should the researcher proceed?

Many researchers conduct a series of bivariate tests on each of the correlation coefficient in turn, testing in each case the hypothesis that the correlation between a pair of variables is equal to the hypothetical correlation. But, this procedure has two drawbacks that all too often are ignored : (a) The probability of making at least one Type I error rate increases with the number of tests performed, and (b) the respective tests in this series of tests are not statistically independent. (Larzelere & Mulaik, 1977)

R. N. Harris (1967) and Larzelere (1975) have performed Monte Carlo studies to evaluate the effects of dependency among the component tests of a series of bivariate tests of correlation. Their studies verified that the empirical estimate of the probability of making at least one Type I error rate in a family of tests increases substantially as the number of component tests increases.

When more than one correlation coefficient is tested for significance in a study, the probability of making at least one Type I error rate rises rapidly as the number of tests increases, and the probability of making a Type I error rate after a Type I error on a previous test is usually greater than the nominal significance level used in each test. To avoid excessive Type I errors with multiple tests of correlations, researchers should use procedures that answer research questions with a single statistical test and/or should use special multiple-test procedures.

Several statistical approaches for controlling the overall type I error rate are available, which could be applied generally in any multiple test situation. The Bonferroni procedure is one of the most commonly used techniques. Recently, several modified Bonferroni procedures have been developed. Halland and Copenhaver (1987) reviewed five of them. Since then, Hochberg (1988) and Hommel (1988) have proposed two additional modified Bonferroni procedures. In exchange for the higher power, the newer approaches are more complex.

Statistical Power of Modified Bonferroni Methods

There have been several discussions on the issue of controlling the overall Type I error rate in situations where multiple tests are conducted simultaneously. The simplest and perhaps the best known approach is to divide the acceptable overall risk of a Type I error rate by the number of hypotheses tested. This approach is known as the original Bonferroni method. Two advantages of this approach are that it is easy to apply and it can be used in many different multiple-testing situations (e.g. contrast analyses, univariate ANOVA tests following a significant multivariate test). A disadvantage of this approach, however, is that the statistical power to detect individual true differences can be low. A number of modifications to the original Bonferroni procedure have been developed and applied. Five of these alternatives were developed by Holm (1979), Holland and Copenhaver (1987), Hochberg (1988), Hommel (1988), and Rom (1990). The objective of these modifications is to increase the statistical power without increasing the risk of a Type I error rate.

Li, Olejnik, and Huberty (1992) compared the five modified Bonferroni procedures with the original Bonferroni using 50 correlation matrices reported in the applied research literature. The results of their study indicated very little difference in the number of hypotheses rejected by the six methods. A major limitation of their study was that since they used real data sets the true relationships among the variables could not be known. Consequently, differences between Type I errors and true relationships could not be distinguished. In addition it was not possible to study different definitions of power (Einot & Gabriel, 1975) : all true relationships, at least one true relationship, and average power.

Applications of six Modified Bonferroni Procedures

Dunnett and Tamhane (1992) categorized these procedures for adjusting significance levels into three groups : single-step (SS), step-down (SD), and step-up (SU), The SS procedure (original Bonferroni) sets a single criterion for testing all individual hypotheses. The SD and SU procedures order the hypotheses to be tested by their p-values, and then compute adjusted significance levels for each individual hypothesis. The SD procedures (Holm, Holland-Copenhaver) start the testing with the hypothesis with the smallest p-value, whereas the SU procedures (Hochberg, Hommel, Rom) start the testing with the hypothesis having the largest

p-value. The step-up procedures are based on extensions of the Simes (1986) modification of the original Bonferroni procedure which was proven to control the overall Type I error rate to be no greater than the nominal significance level if the tests are independent.

In this study, I use α to denote the overall Type I error rate per matrix, α' to denote a criterion for testing an individual hypothesis, i to indicate the order of the hypotheses, and m as the total number of hypotheses tested.

Alternatively, Wright (1992) suggested adjusting the reported p value and reporting the adjusted p value as an indicator of "how significant" the result is; which is consistent with the p-value approach to statistical testing as opposed to the fixed-α approach.

Original Bonferroni Procedure (SS)

With the original Bonferroni (Bon) procedure, the value of $\alpha'=\alpha/m$ is computed. An individual hypothesis $H_{0(i)}$ with $p_{(i)} < \alpha'$ is rejected. Thus the overall Type I error rate is divided equally among all of the possible hypotheses to be tested.

Holm Procedure (SD)

Holm (Holm) (1979) proposed sequentially setting different significance levels for rejecting each individual hypothesis : let $p_{(1)}, ..., p_{(m)}$ be the ordered p values (smallest to largest) and $H_{(1)}, ..., H_{(m)}$ be the corresponding hypotheses. The Holm procedure rejects $H_{(1)}$ to $H_{(i-1)}$ if i is the smallest integer such that $p_{(i)} > \alpha/(m-i+1)$. Statistical power is gained by sequentially increasing the criterion for statistical significance. Because any hypothesis rejected by the original Bonferroni procedure will also be rejected by the Holm procedure, the Holm procedure can never have lower power for an individual hypothesis test.

Holland and Copenhaver procedure (SD)

Holland and Copenhaver (Hc1) (1987) recommended using the Sidak (1967) inequality to set the criterion for individual hypothesis tests. Let $p_{(1)}, ..., p_{(m)}$ be the order p values and $H_{(1)}, ..., H_{(m)}$ be the corresponding hypotheses. Suppose i is the smallest integer from 1 to m such that $p(i)>1-(1-\alpha)^{1/(m-i+1)}$; the Hc1 procedure rejects H(1) to H(i-1) and retains H(i) to H(m). The

criterion used by Hc1 is slightly larger than the criterion used by the Holm procedure thus leading to slightly greater power for an individual hypothesis test.

Hochberg Procedure (SU)

Hochberg (Hb) (1988) developed a step-up procedure based on an extension of the Simes (1986) enhancement of the original Bonferroni procedure. Rather than beginning with the smallest p value, Hb begins with the largest p value and proceeds sequentially through the hypotheses with decreasing p values. The Hochberg procedure rejects $H_{(1)}$ to $H_{(i)}$ for any i=m, m-1, ...1 if $p_{(i)} \leq \alpha/(m-i+1)$. The step-up procedure uses the same criterion for individual hypotheses as does the Holm procedure but tests hypotheses with the larger p values first. The consequence of this procedure is that Hb will test, and possibly reject, hypotheses not examined by the Holm procedure while rejecting the same hypotheses that are rejected by the Holm procedure.

Hommel Procedure (SU)

Hommel (Homm) (1988), extended the Simes (1986) test to individual hypothesis by suggesting a procedure that requires two stages. Let $J=\{i' \in \{1, ..., m\} : p_{(m-i+k)} > k\alpha/i'; k=1, ..., i'\}$. The first stage uses the obtained p values to compute the number of members in J. The second stage obtains the significance level of rejection using $\alpha'=\alpha/j'$, where j' is the largest number in J. Then, if J is nonempty, $H_{(i)}$ is rejected whenever $p_{(i)} \leq \alpha/j'$. If J is empty, all $H_{(i)}$ (i=1, ...m) are rejected.

An example application can help clarify the Hommel procedure. Suppose four hypotheses (m=4) are to be tested with the overall $\alpha=.05$ having the following ordered p-values : $p_{(1)}=.0011$, $p_{(2)}=.0145$, $p_{(3)}=.0291$, and $p_{(4)}=.0562$. To determine j' in the first stage, several steps are needed involving one or more tests (denoted by k). At the first step i'=1; i' is then increased 1 at each step (k=1 ...i' within each step).

Step 1 only contains one test. With m=4, i'=1, and k=1, $p_{(m-i'+k)}=p_{(4)}=.0562$. The criterion value is $k\alpha/i'=1(.05)/1=.05$. Because $p_{(4)}>.05$, the procedure continues with j'=1.

Step 2 contains two tests. For the first test in step 2, $i'=2$, $k=1$, $p_{(m-i'+k)}=p_{(4-2+1)}=p_{(3)}=.0291$, and $k\alpha/i'=1(.05)/2=.025$. Because $p_{(3)}>.025$, the procedure continues. For the second test in step 2, $i'=2$, $k=2$, $p_{(m-i'+k)}=p_{(4-2+2)}=p_{(4)}=.0562$, and $k\alpha/i'=2(.05)/2=.05$, with $p_{(4)}>.05$. Because at step 2 both $p_{(3)}$ and $p_{(4)}$ are larger than the corresponding criterion values, the procedure continues with $j'=2$.

Step 3 contains three tests. For the first test in step 3, $i'=3$, $k=1$, $p_{(m-i'+k)}=p_{(4-3+1)}=P_{(2)}=.0145$, and $k\alpha/i'(.05)/3=.0167$. Because $p_{(2)}<.0167$, the procedure at the first stage stops. The largest number in J is 2, so $j'=2$.

In general, for each test, the obtained p value, $p_{(m-i'+k)}$, is compared with a computed criterion values $k\alpha/i'$. If $p_{(m-i'+k)}>k\alpha/i'$, the calculations move on to the next $p_{(m-i'+k)}$. If $p_{(m-i'+k)}>k\alpha$ i' for all the tests within a step, j' increases by one; otherwise the calculations stop.

For the second stage of the Hommel procedure, reject $H_{(i)}$ whenever $p_{(i)} \leq \alpha/j'$. In the current example, $j'=2$, $\alpha/j'=.05/2=.025$. Therefore hypotheses $H_{(1)}$ and $H_{(2)}$ would be rejected. If $p_{(4)}< .05$ at step 1 all hypotheses would have been rejected.

The Hommel procedure not only considers the order of the tests but also takes the obtained p values into the calculation while computing the α'.

Rom Procedure (SU)

Rom (Rom) (1990) suggested another alternative to the original Bnferroni procedure that enhances the statistical power, but his procedure is computationally intensive as the number of hypotheses increase. With the Rom procedure, I denote $H_{(1)}$ as the hypothesis with the largest p value and $H_{(m)}$ as the hypothesis with the smallest p value.

The testing starts by comparing $p_{(1)}$ with $\alpha_{(1)}'$ and stops when $p_{(i)} < \alpha_{(i)}'$. Then $H_{(1)}$ to $H_{(i+1)}$ are retained and $H_{(i)}$ to $H_{(m)}$ are rejected. The computation needed to determine the $\alpha_i's$ can be divided into three parts. The first part is : $\alpha^1 + \alpha^2 + ... + \alpha^{i-1}$ and the second part is : $\binom{i}{1}(\alpha'_{(2)}{}^{i-1}) + \binom{i}{2}(\alpha'_3{}^{i-2}) + ... + \binom{i}{i-2}(\alpha_{(i-1)}{}'^2)$. The third part is to solve for α_i', which is obtained by subtracting the second part from the first part, and dividing the difference by i. Using the example p-values presented in the previous section but reordered ($p_{(4)}=.0011$, $p_{(3)}=.0145$, $p_{(2)}=.0291$, and $p_{(1)}=.0562$) the Rom procedure can be demonstrated. For

the Rom procedure $\alpha_{(1)}'$ is always equal to the overall α. Because $p_{(1)}>.05$, $H_{(1)}$ is not rejected and the procedure continues.

For the second hypothesis $\alpha_{(2)}$ is always equal to $\alpha_{(1)}'/2$. Because $p_{(2)}>.025$, $H_{(2)}$ is not rejected and the procedure continues. For the third hypothesis $\alpha_{(3)}'=[\alpha+\alpha^2-3\ (\alpha_{(2)}'^2)]/3=.0169$. (The 3 that precedes $(\alpha_{(2)}'^2)$ is the result of $\binom{i}{3}$.) Because $p_{(3)}<\alpha_{(3)}'$, $H_{(3)}$ and all hypotheses with p smaller than $p_{(3)}$ (i.e. $p_{(4)}$) are rejected, and the remaining hypotheses are retained. Rom (1990) provides adjusted critical values for up to 10 tests when the overall α equals .05 and .01. The increased statistical power for an individual hypothesis test is obtained by controlling the overall Type I error rate exactly at the nominal significance level. The alternatives to the original Bonferroni procedure proposed by Holm (1979), Holland and Copenhaver (1987), Hochberg (1988) and Hommel (1988) are conservative in limiting the overall Type I error rate to be no greater than the nominal level.

Holland and Copenhaver procedure (SU)

An approach not previously considered is an application of the Holland and Copenhaver as a step-up (Hc2) procedure. The Holland-Copenhaver step-up procedure may be described as follow : let $p_{(1)}$, ..., $p_{(m)}$ be the order p-values and $H_{(1)}$, ..., $H_{(m)}$ be the corresponding hypotheses. Suppose i is the largest integer from 1 to m such that $p_{(i)} < 1-(1-\alpha)^{1/(m-i+1)}$. The Holland-Copenhaver step-up procedure rejects $H_{(1)}$ to $H_{(i-1)}$ and retain $H_{(i)}$ to $H_{(m)}$.

Li, Olejnik, and Huberty (1992) demonstrated the numerical examples for Bonferroni and five modified Bonferroni procedures.

Purpose

The purpose of the present study is to address the limitations of the previous investigation by studying the Type I error rate and the three conceptualizations of power using computer simulation methods. In addition a sixth modified Bonferroni method is introduced. The Holland-Copenhaver approach uses a step down method. That is after ordering the p-values, hypotheses are tested from the smallest to largest p-values. I test the hypotheses from the largest to the smallest p-values, thus a step-up approach. This step-up approach is similar

to the Hochberg method. The present study also uses the correlation matrix as the context for multiple tests.

Related Literature

Li, Olejnik, and Huberty (1992) demonstrated the application of these six methods excluding Holland and Copenhaver (step-up) method for controlling the overall Type I error rate using multiple hypothesis tests associated with a correlation matrix. A SAS macro computer program was developed to compute the adjusted criteria for the original Bonferroni, Hochberg (1988), Holland & Copenhaver (step-down) (1987), Holm (1979), and Hommel (1988) procedures. Using a single data example, Wright (1992) demonstrated the increased statistical power for individual hypothesis tests in the context of contrasts among means in a oneway ANOVA by presenting the adjusted p values associated with the original Bonferroni and the Holm (1979), Hochberg (1988), and Hommel (1988) procedures. A computer program for computing the adjusted p values recommended by Wright is available in PROC MULTTEST (SAS, 1992) and is discussed by Westfall and Young (1993).

Analytically, it can be seen that the Holm procedure is more powerful than the original Bonferroni and that the Holland-Copenhaver procedure is more powerful than the Holm procedure. Hochberg and Benjamini (1990) showed that the Hochberg procedure is also more powerful than the Holm procedure. Hommel (1989) showed that his procedure is as least as powerful as the Hochberg procedure and in general can be more powerful. Holland (1991) compared the Holland-Copenhaver, the Hommel, and the Rom procedures in the context of all pairwise comparisons for a single factor repeated measures design. The results indicated that the Hommel procedure rejected more hypotheses than either the Rom or Holland copenhaver procedures; the difference in the number of tests rejected was, however, very small. Holland suggested the need for further research in situations where there is no logical interrelationship among the hypotheses tested. Finaly, Dunnett and Tamhane (1992) reported that the increase in power for individual hypothesis tests provided by the Hommel and Rom procedures over the Hochberg approach is at best marginal, with the Rom procedure having only a slight advantage.

Method

Computer programs were written using SAS/IML (1988) to generate correlation matrices for the purpose of comparing the statistical power and Type I error rates among the seven methods. Five factors were considered : number of variables, sample size, overall Type I error rate, strength of the relationship between variables, and the number of true non-zero relationships in a given matrix. Data were generated for four and six variables each having a standard normal distribution. The null hypotheses were tested ($H_0 : \rho_{ij}=0$ for all $i \neq j$) with the overall Type I error rate set at .05 and .20. The true zero and non-zero relationships among the variables were simulated for two situations. First, to study the Type I error rates all variables generated were independent of each other ($\rho=0$ for all pairs of variables). Second, to compare the seven procedures under each of the three definitions of power, matrices were generated for which pairs of variables were correlated .60, .40, or .20 when the number of variables equaled four. When the number of variables in the matrix equaled six, pairs of variables were correlated .40. The partial Type I error rates were also considered. That is, for the subset of hypotheses within a matrix where the variables were independent, the proportion of times when the null hypothesis was rejected was recorded. Table 1 presents three example target population correlation matrices involving four and six variables where three pairs of variables in the former are correlated: a) .40 or .20, b) .60, and c) the five pairs of variables are correlated .40 with the remaining pairs of variables being independent ($\rho=0$).

Table 1 Example Target Population Correlation Matrices for Four and Six Variables

	Variable				Variable				Variable				
	V2	V3	V4		V2	V3	V4		V2	V3	V4	V5	V6
V1	.40	0	0	V1	.60	0	0	V1	.40	0	0	0	0
V2		.20	0	V2		.60	0	V2		.40	0	0	0
V3			.40	V3			.60	V3			.40	0	0
								V4				.40	0
								V5					.40
	(a)				(b)				(c)				

The SAS-RANNOR function was used to generate the normal random numbers. The matrices containing true non-zero relationships were generated using the procedure suggested by Kaiser and Dickman (1962). To check the data generating procedure, 100,000 observations were generated for four variables with all pairs of variables being correlated .40. The absolute value of the difference between the sample correlations and the population correlations was no greater than .006. The procedure was repeated for six variables with the same results.

For each condition studied, 10,000 replications were generated. The percent of hypotheses rejected by each procedure was recorded. Under the complete null, the proportion of matrices rejecting at least one hypothesis was recorded. For non-null matrices the proportion of matrices in which all true relationships were identified were recorded as well as the proportion of matrices in which at least one true relationship was detected. Finally, when two or more true non-zero relationships were present within a matrix, the average power for detecting each of the true non-zero relationships was also recorded.

Results

Type I Error Rates. Table 2 presents the proportion of matrices in which at least one correlation was declared significant when there were no true relationship among any of the variables. All seven methods provided empirical Type I error rates less than the nominal significance levels of .05 and .20 when the number of variables equaled four and six. These results provide a partial check of the computer programs.

Table 2 Type I error Rates

k	α	n	Bon	Holm	Hc1	Hb	Homm	Rom	Hc2
4	.05	10	.051	.051	.053	.051	.051	.052	.053
		30	.046	.046	.047	.046	.047	.047	.047
		50	.051	.051	.052	.051	.051	.052	.052
		100	.047	.047	.048	.047	.048	.048	.048
	.20	10	.178	.178	.196	.179	.184	.197	.197
		30	.182	.182	.201	.183	.189	.201	.202
		50	.178	.178	.194	.178	.183	.194	.195
		100	.189	.189	.203	.189	.194	.204	.204
6	.05	10	.050	.050	.051	.050	.050	.051	.051
		30	.047	.047	.048	.047	.047	.048	.048
		50	.049	.049	.050	.049	.049	.049	.050
		100	.051	.051	.052	.051	.051	.052	.052
	.20	10	.182	.182	.199	.182	.184	.199	.199
		30	.179	.179	.199	.180	.182	.200	.200
		50	.187	.187	.204	.187	.189	.205	.204
		100	.179	.179	.196	.179	.181	.196	.196

Bon=Original Bonferroni procedure

Holm=Holm procedure

Hc1=Holland-Copenhaver (step-down) procedure

Hb=Hochberg procedure

Homm=Hommel procedure

Rom=Rom procedure

Hc2=Holland-Copenhaver (step-up) procedure

All True Relationship Power. The proportions of matrices in which all of the true non-zero relationships were detected by the seven procedures are reported in Tables 3 through 6.

As expected, as the overall risk of a Type I error rate increases from .05 to .20 and as the sample sizes increase, power increases, but as the number of variables in a matrix increases, the probability of rejecting all of the non-null hypotheses decreases. All six of the enhancements were more sensitive than the original Bonferroni procedure in detecting all true non-zero relationships. The difference between the original Bonferroni procedure and the

enhancements increased as the number of true non-zero relationships increased. Very small differneces in statistical power, however, were found among the six enhancements to the original Bonferroni procedure. The Holm procedure consistently had the lowest sensitivity in detecting all true non-zero relationships while the Holland-Copenhaver (step-up) procedure had the greatest power. When all of the correlations were non-zero the Hochberg, Hommel, and Rom procedures had the same estimated power. The pattern of differences in power were similar when the variables were correlated .40 or .60.

Table 3 All true relationship power for alpha=.05 k =4

#of sig. correlations	n	Bon	Holm	Hc1	Hb	Homm	Rom	Hc2
1	10	.058	.059	.059	.059	.059	.059	.059
	30	.332	.334	.337	.334	.334	.337	.337
	50	.613	.614	.617	.615	.615	.617	.617
	100	.943	.943	.944	.943	.943	.943	.944
2	10	.004	.005	.005	.005	.005	.005	.005
	30	.116	.133	.135	.133	.134	.135	.135
	50	.370	.402	.405	.404	.405	.406	.407
	60	.508	.541	.545	.541	.542	.545	.545
	70	.643	.675	.680	.676	.677	.680	.681
	80	.747	.771	.773	.772	.773	.774	.774
	90	.821	.840	.842	.840	.841	.842	.842
	100	.882	.896	.898	.896	.896	.898	.898
3	30	.004	.008	.008	.008	.009	.009	.009
	50	.026	.043	.044	.044	.044	.044	.045
	100	.211	.263	.264	.264	.265	.265	.265
	200	.583	.643	.645	.644	.645	.645	.646
	250	.708	.755	.757	.755	.756	.757	.757
	300	.805	.843	.845	.844	.845	.846	.846
	350	.872	.904	.906	.905	.905	.905	.906
	400	.915	.937	.937	.938	.938	.938	.938
	450	.947	.963	.963	.963	.963	.963	.964

Table 4 All true relationship power for alpha=.20 k =4

#of sig. correlations	n	Bon	Holm	Hc1	Hb	Homm	Rom	Hc2
1	10	.160	.164	.174	.165	.167	.174	.175
	30	.548	.552	.568	.553	.559	.569	.569
	50	.785	.789	.801	.791	.794	.802	.803
	100	.981	.981	.983	.981	.982	.983	.983
2	30	.292	.339	.359	.345	.351	.362	.363
	40	.465	.510	.528	.515	.522	.531	.532
	50	.615	.657	.672	.660	.665	.674	.676
	60	.741	.771	.782	.774	.777	.783	.785
	70	.840	.862	.872	.864	.867	.873	.874
	80	.896	.914	.919	.915	.917	.919	.920
3	40	.069	.111	.121	.119	.124	.127	.130
	50	.126	.185	.197	.193	.199	.203	.206
	60	.177	.244	.262	.253	.258	.266	.270
	70	.240	.319	.335	.326	.332	.338	.342
	80	.309	.389	.407	.397	.404	.410	.414
	90	.364	.449	.467	.456	.463	.470	.473
	100	.415	.498	.514	.505	.510	.518	.521
	150	.626	.695	.708	.700	.704	.710	.713
	200	.761	.819	.826	.823	.826	.829	.830
	250	.859	.895	.901	.898	.900	.902	.903
	300	.916	.942	.946	.944	.945	.947	.947

Table 5 All true relationship power for alpha=.05 k =6

#of sig. correlations	n	Bon	Holm	Hc1	Hb	Homm	Rom	Hc2
1	10	.030	.030	.031	.030	.030	.031	.031
	30	.220	.220	.224	.220	.220	.223	.224
	50	.496	.496	.500	.497	.497	.499	.500
	100	.892	.892	.893	.892	.892	.893	.893
3	50	.113	.127	.130	.128	.128	.129	.130
	60	.220	.238	.242	.239	.239	.240	.242
	70	.355	.379	.384	.380	.380	.383	.384
	80	.481	.507	.511	.508	.508	.511	.512
	90	.607	.631	.635	.632	.632	.634	.635
	100	.716	.736	.739	.736	.736	.738	.739
	120	.857	.867	.868	.867	.867	.867	.868
	130	.905	.915	.917	.915	.915	.916	.917
5	200	.211	.253	.256	.253	.253	.257	.256
	250	.374	.418	.421	.418	.418	.423	.422
	300	.509	.557	.559	.557	.557	.560	.559
	350	.634	.678	.682	.678	.678	.683	.682
	400	.750	.785	.786	.785	.785	.787	.786
	450	.828	.853	.856	.854	.854	.857	.856
	500	.890	.910	.912	.910	.910	.912	.912
	550	.929	.943	.944	.943	.943	.944	.944
	600	.954	.964	.965	.964	.964	.965	.965

Table 6 All true relationship power for alpha=.20 k =6

#of sig. correlations	n	Bon	Holm	Hc1	Hb	Homm	Rom	Hc2
1	10	.085	.086	.093	.086	.087	.093	.093
	30	.397	.399	.415	.399	.401	.416	.415
	50	.672	.673	.687	.673	.675	.687	.687
	100	.955	.956	.960	.956	.956	.960	.960
3	50	.315	.344	.362	.344	.348	.363	.364
	60	.465	.499	.521	.500	.503	.521	.521
	70	.606	.636	.654	.637	.640	.655	.655
	80	.719	.744	.760	.745	.747	.761	.761
	90	.807	.825	.837	.825	.827	.837	.837
	100	.875	.890	.897	.890	.892	.898	.898
5	60	.014	.023	.026	.023	.023	.026	.026
	100	.090	.122	.132	.122	.123	.132	.132
	150	.250	.296	.311	.296	.301	.311	.311
	200	.421	.478	.498	.479	.482	.498	.498
	250	.590	.642	.659	.643	.646	.659	.659
	300	.731	.775	.788	.776	.778	.788	.788
	350	.819	.851	.861	.851	.853	.861	.861

Although a very limited, specific sets of conditions were studied in the present study in order to compare power estimates among adjustment procedures, Tables 3 through 6 do provide some interesting results on sample size requirements for the conditions I simulated. For example, when using the modified Bonferroni procedures, approximately the same sample size is needed to detect all true non-zero relationships when all of the pairs of variables are correlated .60 as when only one pair of variables is correlated .60. Using Rom with k=4 and $\alpha=.20$, a sample size of 20 will result in power of .806 when only one pair of variables is correlated .60 and power of .814 when all 6 pairs of variables are correlated .60. This result is a function of the dependence among the hypothesis tests when all pairs of variables are correlated. If fewer pairs of variables are correlated, a greater difference in sample sizes will result. For example, with k=4 and $\alpha=.20$, power will equal .806 when N=20 when using the

Rom procedure with one pair of variables is correlated .60. But with k=4 and α=.20 and four pairs of variables are correlated .60 a sample size of 25 would be needed for the estimated power to equal .80.

Another interesting result is that when k=4 and α=.20, using Rom, the sample size needed to detect all true relationships would be slightly less than 20 for power to equal .80 when all pairs of variables are correlated .60. If a researcher had used the power tables provided by Cohen (1988) to detect a correlation of .60 with α=.20 and power equal to .80, only 11 participants would be needed. This result demonstrates a limitation of using the Cohen power charts to determine the sample size needed if multiple hypotheses are tested and the overall Type I error rate is controlled.

The partial Type I error rate for all matrices containing pairs of variables which were independent of each other was less than the nominal significance levels of .05 and .20.

At Least One True Relationship Power. Tables 7 and 8 present the proportions of matrices in which at least one true non-zero relationship was detected when the significance level equaled .05 and .20. Again predictable power results as a function of α, k, and N were obtained. Power estimates to reject at least one true relationship were always greater than the power estimates provided in Tables 3 through 6 to reject all true non-zero relationships for the same conditions. The results indicate almost no difference between the original Bongferroni and each of the enhancements. Finally, the sample size needed to detect at least one true non-zero relationship in the matrix was approximately the same as indicated by the sample size chart provided by Cohen (1988); that is, for α=.20 and ρ=.40, a sample size of 28 is needed for power to equal .80.

Table 7 At least one true relationship power for alpha=.05

k	#of sig correlations	n	Bon	Holm	Hc1	Hb	Homm	Rom	Hc2
4	2	10	.117	.117	.118	.117	.118	.119	.119
		30	.566	.568	.571	.568	.572	.572	.572
		50	.849	.850	.853	.851	.855	.854	.854
		100	.996	.996	.997	.997	.997	.997	.997
	3	10	.123	.124	.126	.124	.126	.126	.127
		30	.527	.528	.531	.528	.534	.532	.532
		50	.824	.824	.827	.825	.830	.827	.828
		100	.994	.994	.994	.994	.994	.994	.994
6	3	10	.085	.085	.087	.085	.085	.086	.087
		30	.538	.539	.544	.539	.542	.543	.544
		50	.868	.868	.872	.869	.871	.871	.872
		100	.999	.999	.999	.999	.999	.999	.999
	5	10	.095	.095	.097	.095	.096	.097	.097
		30	.564	.565	.569	.565	.568	.569	.569
		50	.881	.882	.884	.882	.885	.883	.884
		100	.999	.999	.999	.999	1.00	.999	.999

Table 8 At least one true relationship power for alpha=.20

k	#of sig correlations	n	Bon	Holm	Hc1	Hb	Homm	Rom	Hc2
4	2	10	.293	.299	.319	.303	.310	.321	.321
		30	.790	.793	.808	.796	.806	.811	.811
		50	.956	.957	.962	.957	.962	.962	.962
		100	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	3	10	.312	.316	.337	.318	.327	.339	.340
		30	.778	.781	.797	.784	.797	.799	.799
		50	.952	.953	.956	.954	.959	.957	.957
		100	1.00	1.00	1.00	1.00	1.00	1.00	1.00
6	3	10	.227	.230	.247	.230	.234	.248	.247
		30	.782	.784	.798	.785	.790	.799	.799
		50	.965	.965	.970	.965	.967	.970	.970
		100	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	5	10	.269	.271	.289	.272	.277	.290	.289
		30	.808	.810	.826	.810	.817	.826	.826
		50	.975	.975	.979	.975	.978	.979	.979
		100	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Average Power. The average proportion of true non-zero relationships rejected per matrix are presented in Tables 9 and 10. The original Bonferroni procedure had the lowest average power but the enhancement procedures offered only a small, generally between two and three percent, increase in average power. Using the Cohen (1988) sample size charts would result in an underestimate in the needed sample size to achieve the desired level of power.

Table 9 Average power for alpha=.05

k	#of sig correlations	n	Bon	Holm	Hc1	Hb	Homm	Rom	Hc2
4	2	30	.341	.350	.353	.351	.353	.353	.354
		50	.609	.626	.629	.627	.630	.630	.630
		60	.713	.730	.733	.731	.732	.733	.733
		70	.801	.818	.821	.818	.820	.821	.821
		80	.865	.877	.878	.877	.878	.878	.878
		90	.906	.915	.916	.915	.916	.916	.916
		100	.939	.946	.947	.946	.947	.947	.947
	3	30	.212	.219	.221	.220	.223	.222	.222
		50	.397	.414	.417	.415	.418	.417	.418
		100	.686	.709	.711	.710	.712	.711	.711
		200	.860	.881	.882	.881	.882	.882	.882
		250	.903	.918	.919	.918	.919	.919	.919
		300	.935	.948	.948	.948	.948	.949	.949
6	3	30	.226	.229	.232	.230	.231	.232	.232
		50	.486	.495	.499	.495	.498	.499	.499
		60	.604	.614	.617	.614	.616	.617	.617
		70	.708	.719	.722	.720	.721	.722	.722
		80	.783	.794	.796	.794	.795	.797	.797
		90	.846	.856	.858	.856	.857	.857	.858
		100	.895	.902	.903	.902	.902	.903	.903
		120	.950	.954	.954	.954	.954	.954	.954
	5	30	.147	.150	.152	.151	.152	.152	.152
		50	.315	.323	.325	.323	.325	.325	.325
		100	.605	.618	.620	.618	.620	.620	.620
		200	.784	.798	.799	.798	.798	.800	.799
		250	.843	.856	.857	.856	.856	.857	.857
		300	.885	.897	.898	.897	.897	.898	.898
		350	.919	.929	.930	.929	.929	.930	.930
		400	.947	.954	.945	.954	.954	.955	.955

Table 10 Average power for alpha=.20

k	#of sig correlations	n	Bon	Holm	Hc1	Hb	Homm	Rom	Hc2
4	2	10	.160	.168	.180	.171	.176	.183	.183
		30	.541	.566	.584	.570	.579	.586	.587
		40	.680	.704	.716	.707	.714	.719	.719
		50	.786	.807	.817	.809	.814	.818	.819
		60	.860	.875	.882	.876	.879	.882	.883
		70	.916	.927	.933	.928	.930	.933	.934
	3	10	.116	.123	.133	.126	.131	.135	.136
		30	.347	.399	.413	.403	.413	.416	.418
		50	.571	.606	.617	.611	.619	.620	.622
		70	.690	.726	.735	.729	.734	.736	.738
		80	.735	.768	.777	.772	.776	.778	.780
		90	.765	.799	.806	.801	.805	.808	.809
		100	.791	.822	.828	.824	.827	.829	.830
		150	.875	.898	.902	.900	.901	.903	.904
		200	.920	.940	.942	.941	.942	.943	.943
6	3	30	.398	.409	.423	.409	.414	.424	.424
		50	.681	.695	.712	.696	.699	.709	.709
		60	.775	.790	.801	.790	.793	.802	.802
		70	.846	.858	.866	.858	.860	.866	.966
		80	.896	.905	.912	906	.907	.912	.912
		90	.932	.938	.942	.938	.939	.942	.942
	5	30	.245	.274	.287	.275	.280	.287	.287
		40	.375	.390	.402	.390	.395	.403	.403
		50	.457	.474	.485	.474	.479	.486	.486
		60	.531	.549	.560	.550	.554	.561	.561
		100	.702	.720	.728	.721	.723	.728	.728
		150	.797	.813	.819	.813	.815	.819	.819
		200	.859	.875	.880	.875	.876	.881	.881
		250	.906	.920	.924	.920	.921	.924	.924

Conclusions

The Bonferroni method for controlling the Type I error rate over a series of hypothesis tests has been popular among researchers because of its computational simplicity and wide applicability. Its major limitation has been a reduction in statistical power for the hypothesis tests as the number of tests increase. In recent years several efforts have been made to increase the statistical power of the Bonferroni method. Analytic studies of these enhancements have shown that they do provide greater sensitivity to true relationships than the original Bonferroni but the magnitude of that difference has not been clear. Similarly, comparisons between the enhancements have been shown analytically that some alternatives are more powerful than others but again the magnitude of the difference has not been clear. The greater statistical power has generally come as a result of increase computational difficulty.

Li, Olejnik, and Huberty (1992) raised some question as to the utility of the enhancements when they showed very small differences between the alternatives and only modest increases in power over the original Bonferroni. They used real data sets where true relationships could not be distinguished from Type I errors.

In the present study I used computer generated data to investigate Type I error rates and three definitions of statistical power to compare five suggested enhancements to the Bonferroni and I proposed still another enhancement that has not been previously considered based on the Holland-Copenhaver approach. Results show that for all three definitions of power, the new step-up Holland-Copenhaver and the Rom procedures tend to have the highest power among the seven procedures. Because the Rom procedure is more complicated, I recommend the Holland-Copenhaver (step-up) procedure to be used. Moreover, results from this study are consistent with 1) Hommel (1989) that the Hommel procedure is at least as powerful as the Hochberg procedure, and in general more powerful. The power difference however is in the third decimal place. 2) Dunnett and Tamhane (1992) that power increases yielded by the Hommel and the Rom procedures over the Hochberg procedure are marginal at the best, with the Rom procedure being slightly superior. c) Hochberg and Benjamin (1990) that Holm procedure is sharper than the original Bonferroni and the Hochberg procedure is sharper than the Holm procedure. All six modified Bonferroni procedures are supetior to the original Bonferroni procedure under all true relationship power and average power. However, I

gain less than 2% when alpha is small (i.e., .05) and less than 5% when alpha is .20 under the average power definition. Under all true relationship power I gain greater power only when alpha is large (.20) and there are large number of true relationships in the matrix. Therefore, while I agree with Holland and Copenhaver (1987) that a modified Bonferroni procedure should be used in situations where the original Bonferroni would otherwise be the method of choice, I have been disappointed with the magnitude of the power increase.

Controlling the overall Type I error rate over a series of hypothesis tests is an important topic of interest to applied researchers and data analysis. A considerable effort has gone into modifying the Bonferroni method in order to increase statistical power. Results of the present study appears that this effort has not been too successful in improving the statistical power. Additional research in this area is needed to develop still other alternatives that may be more sensitive to true relationships than the current enhancements and the original Bonferroni.

References

Barton, R.M., & Huberty, C.J. (1989, April). **Multiple testing and correlation matrices : An application of three procedures**. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Cohen, J. (1988). **Statistical power analysis for the behavioral sciences**. (2nd ed.) Hillsdale, NJ : Erlbaum.

Dunnett, C.W., & Tamhane, A.C. (1992). A step-up multiple test procedure. **Journal of the American Statistical Association, 87**, 162-170.

Einot, I., & Gabriel, K.R. (1975). A study of the powers of several methods of multiple comparisons. **Journal of the American Statistical Association, 70**, 574-583.

Hochberg, Y. (1988). A sharper Bonferrroni procedure for multiple tests of significance. **Biometrika, 75**, 800-802.

Hochberg, Y., & Benjamini, Y. (1990). More powerful procedures for multiple significance testing. **Statistics in Medicine, 9**, 811-818.

Holland, B.S., & copenhaver, M.D. (1987). An improved sequentially rejective Bonferroni test procedure. **Biometrics, 43**, 417-423.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. **Scandinavian Journal of Statistics, 6**, 65-70.

Hommel, G. (1988). A stagewise rejective multiple test procedure based on a modified Bonferroni test. **Biometrika, 75**, 383-386.

Hommel, G. (1989). A comparison of two modified Boferroni procedures. **Biometrika, 76**, 624-625.

Hsiung, T.H., & Olejnik, S. (1991, April). **Power of pairwise multiple comparisons in the unequal variance case**. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Kaiser, H.F., & Dickman, K. (1962). Sample and population score matrices and sample correlation matrices from and arbitrary population correlation matrix. **Psychometrika, 27**, 179-182.

Keppel, G. (1991). **Design and analysis : A researcher's handbook** (3rd ed.). Englewood Cliffs, NJ : Prentice Hall.

Larzelere, R.E., & Mulaik, S.A. (1977). Single-Sample Tests for Many Correlations. **Psychological Bulletin, 84**, 557-569.

Li, J., Olejnik, S., & Huberty, C.J. (1992, April). **Multiple testing with modified Bonferroni methods**. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Maxwell S.E., & Delaney, H.D. (1990). **Designing experiments and analyzing data : A Model comparison perspective**. Belmont, CA : Wadsworth.

McHugh, R.B, & Ellis, D.S. (1955). The post mortem testing of experimental comparison. **Psychological Bulletin, 52**, 425-428.

Rom, D.M. (1990). A sequentially rejective test procedure based on a modified Bonferroni inequality. **Biometrika, 77**, 663-665.

Ryan, T.A. (1959). Multiple comparisons in psychological research. **Psychological Bulletin, 56**, 26-47.

SAS Institute Inc. (1992). **SAS/STAT Software : changes and Enhancements. Release 6.07**. SAS Technical Report P-229. Cary North Carolina : SAS Institute Inc.

SAS Institute Inc. (1988). **SAS/IML User's Guide**. Cary, North Carolina : SAS Institute Inc.

Seaman, M.A., Levin, J.R., Serlin, R.C., & Franke, T.M. (1990, April). **A modified algorithm and an examination of power for Peritz's pairwise multiple-comparison procedure**. Paper presented at the annual meeting of American Educational Research Association, Boston, MA.

Sidak, Z. (1967). Rectangular confidence regions for the means of multivariate normal distributions. **Journal of the American Statistical Association, 62**, 623-633.

Simes, R.J. (1986). An improved Bonferroni procedure for multiple tests of significance. **Biometrika, 73**, 751-754.

Stanley, J.C. (1957). Additional post mortem tests of experimental comparisons. **Psychological Bulletin, 54**, 128-130.

Toothaker, L.E. (1991). **Multiple comparisons for researchers**. Newbury, CA : Sage.

Westfall, P.H., & Young, S.S. (1993). On adjusting p-values for multiplicity. **Bicmetrics, 49**, 941-945.

Wright, S.P. (1992). Adjusted p-values for simultaneous inference. **Biometrics, 48**, 1005-1013.
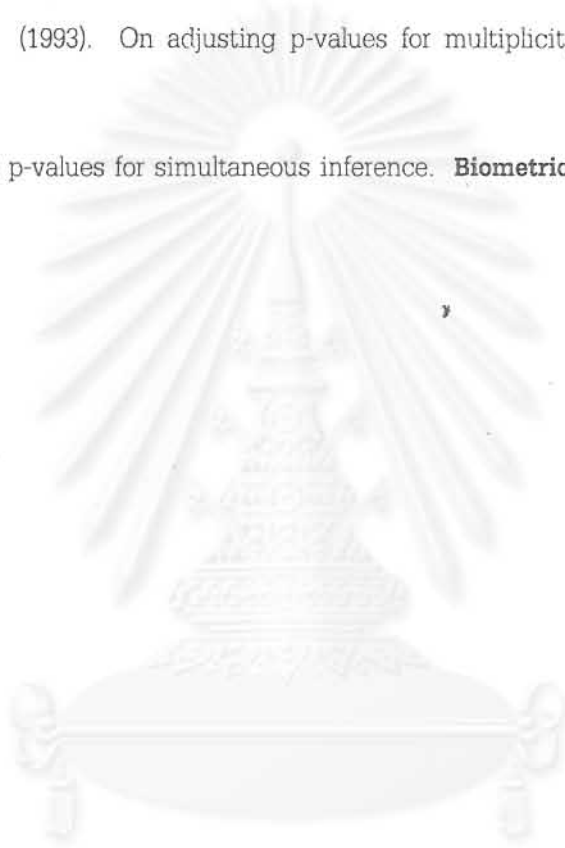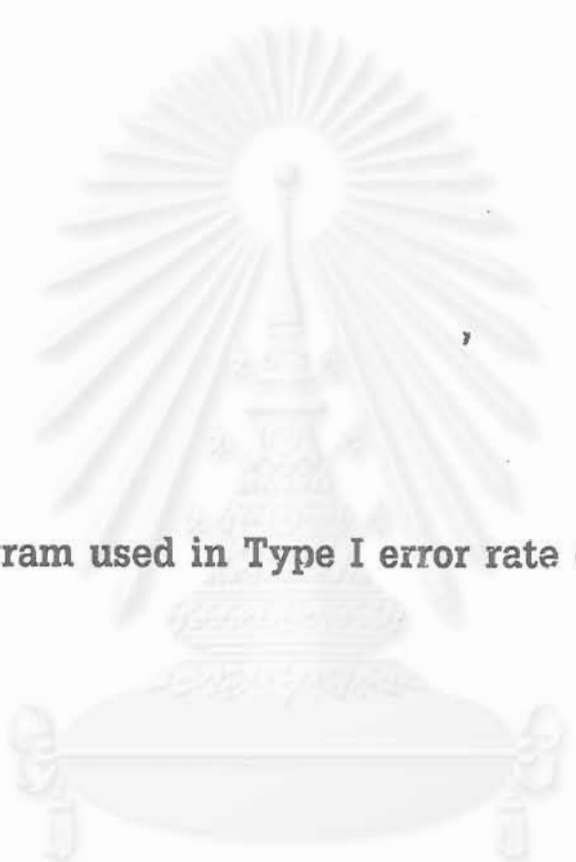
Appendix

Example of program used in Type I error rate calculating

```
options pagesize=54;
proc iml wrkspace=300;
title 'type I error rate with n=10,30,50,100 k=4   alpha=.05';
n=10;
c=4;
alpha=.05;
count=j(1,8,0);
pmcount=j(1,8,0);
do ii=1 to 10000;
x =rannor(j(n ,c,0));
sum=x[+,];
xpx=x *x-sum *sum/n;
s=diag(1/sqrt(vecdiag(xpx)));
r =  s*xpx*s;

nc=(c*(c-1))/2;

cp=j(nc,1,0);

cp[1,1]=r[2,1];
cp[2,1]=r[3,1];
cp[3,1]=r[3,2];
cp[4,1]=r[4,1];
cp[5,1]=r[4,2];
cp[6,1]=r[4,3];

*rint cp;
top=cp#(sqrt(n-2));
bo=sqrt(1-cp##2);
t=top/bo;
p=(1-probt(abs(t),(n-2)))*2;
*rint top bo t p r x;

b=p;
p[ rank(p),]=b;
*rint b p;

*={.0336,.0338,.0520,.0720,.0900,.2500};
*={.0400,.0300,.0520,.0800,.1100,.2200};

*** no control of type I error ***;
nocntl=j(nc,1,alpha);
cl=j(nc,1,0);
cl=p<nocntl;
*rint nocntl cl p;

*** original Bonferroni procedure ***;
bonf=j(nc,1,alpha);
bonf=bonf/nc;
c2=j(nc,1,0);
c2=p<bonf;
*rint c2 bonf p;

*** Holm procedure ***;
holm=j(nc,1,0);
*rint holm;
do i=1 to nc;
holm[i,1]=alpha/(nc-i+1);
 end;
_c3=j(nc,1,0);
```

```
c3=p<holm;
*rint c3 p holm;

*** Holland-Copenhaver procedure ***;
hoco=j(nc,1,0);
do i=1 to nc;
xx=1/(nc-i+1);
hoco[i,1]=1-(1-alpha)##xx;
end;
c4=j(nc,1,0);
c4=p<hoco;
*rint c4 hoco p;

*** Hochberg procedure ***;
hoch=j(nc,1,0);
do i=1 to nc;
hoch[i,1]=alpha/(nc-i+1);
end;
c5=j(nc,1,0);
c5=p<hoch;
*rint c5 p hoch;

*** Hommel procedure ***;
p2=j(nc,1,0);
p2[1,1]=p[6,1];
p2[2,1]=p[5,1];
p2[3,1]=p[4,1];
p2[4,1]=p[3,1];
p2[5,1]=p[2,1];
p2[6,1]=p[1,1];

*={.2670,.0633,.0633,.0350,.0255,.0129,.0089,.0011,.0001,.0000};
*c=10;
*lpha=.10;

hn=nc;
m=-1;
pp=1;
z=nc;
do i=1 to nc while(pp > m);
do k=1 to i   while(pp > m);
m=k*alpha/i;
 hn=nc-i+k;
pp= p[hn,1];
if pp<=m then z=i-1;
end;
if pp<=m then i=i-1;
end;

homm=j(nc,1,alpha);
if z>0 then
homm=homm/z;
c6=j(nc,1,0);
c6=p<homm;
*rint c6 p homm i;

*** Rom procedure ***;
rom=j(nc,1,0);
rom[1,1]=.0085;
rom[2,1]=.0102;
```

```
rom[3,1]=.0127;
rom[4,1]=.0169;
rom[5,1]=.0250;
rom[6,1]=.0500;
c7=p<rom;
*rint c7 p rom;


***holland copenhaver step-up***;
hoc2=j(nc,1,0);
do i=1 to nc;
xx=1/(nc-i+1);
hoc2[i,1]=1-(1-alpha)##xx;
end;
c8=j(nc,1,0);
c8=p<hoc2;
*rint c8 hoc2 p;

ct1=c1[+, ];
ct2=c2[+, ];

ct3=0;
do i=1 to nc while (c3[i,1]=1);
                ct3=ct3+1;
end;

ct4=0;
do i=1 to nc while (c4[i,1]=1);
                ct4=ct4+1;
end;

ct5=nc;
do i=nc to 1 by -1 while(c5[i,1]=0);
                ct5=ct5-1;
end;

ct6=nc;
do i=nc to 1 by -1 while(c6[i,1]=0);
                ct6=ct6-1;
end;

ct7=nc;
do i=nc to 1 by -1 wnile(c7[i,1]=0);
                ct7=ct7-1;
end;
ct8=nc;
do i=nc to 1 by -1 while(c8[i,1]=0);
                ct8=ct8-1;
end;

count[1,1]=count[1,1]+ct1;
count[1,2]=count[1,2]+ct2;
count[1,3]=count[1,3]+ct3;
count[1,4]=count[1,4]+ct4;
count[1,5]=count[1,5]+ct5;
count[1,6]=count[1,6]+ct6;
count[1,7]=count[1,7]+ct7;
count[1,8]=count[1,8]+ct8;

if ct1=0 then pmcount[1,1]=pmcount[1,1]+0;
```

```
            else pmcount[1,1]=pmcount[1,1]+1;
 if ct2=0 then pmcount[1,2]=pmcount[1,2]+0;
            else pmcount[1,2]=pmcount[1,2]+1;
 if ct3=0 then pmcount[1,3]=pmcount[1,3]+0;
            else pmcount[1,3]=pmcount[1,3]+1;
 if ct4=0 then pmcount[1,4]=pmcount[1,4]+0;
            else pmcount[1,4]=pmcount[1,4]+1;
 if ct5=0 then pmcount[1,5]=pmcount[1,5]+0;
            else pmcount[1,5]=pmcount[1,5]+1;
 if ct6=0 then pmcount[1,6]=pmcount[1,6]+0;
            else pmcount[1,6]=pmcount[1,6]+1;
 if ct7=0 then pmcount[1,7]=pmcount[1,7]+0;
            else pmcount[1,7]=pmcount[1,7]+1;
 if ct8=0 then pmcount[1,8]=pmcount[1,8]+0;
            else pmcount[1,8]=pmcount[1,8]+1;
 end;
 print  count pmcount n ii alpha nc                    ;
 n=30;
 c=4;
 alpha=.05;
 count=j(1,8,0);
 pmcount=j(1,8,0);
 do ii=1 to 10000;
 x =rannor(j(n ,c,0));
 sum=x[+,];
 xpx=x`*x-sum`*sum/n;
 s=diag(1/sqrt(vecdiag(xpx)));
 r = s*xpx*s;

 nc=(c*(c-1))/2;

 cp=j(nc,1,0);

 cp[1,1]=r[2,1];
 cp[2,1]=r[3,1];
 cp[3,1]=r[3,2];
 cp[4,1]=r[4,1];
 cp[5,1]=r[4,2];
 cp[6,1]=r[4,3];

 *rint cp;
 top=cp#(sqrt(n-2));
 bo=sqrt(1-cp##2);
 t=top/bo;
 p=(1-probt(abs(t),(n-2)))*2;
 *rint top bo t p r x;

 b=p;
 p[ rank(p),]=b;
 *rint b p;

 *={.0330,.0338,.0520,.0720,.0900,.2500};
 *={.0400,.0300,.0520,.0800,.1100,.2200};

 *** no control of type I error ***;
 nocntl=j(nc,1,alpha);
 cl=j(nc,1,0);
 cl=p<nocntl;
 *rint nocntl cl p;
```

```
*** original Bonferroni procedure ***;
bonf=j(nc,1,alpha);
bonf=bonf/nc;
c2=j(nc,1,0);
c2=p<bonf;
*rint c2 bonf p;

*** Holm procedure ***;
holm=j(nc,1,0);
*rint holm;
do i=1 to nc;
holm[i,1]=alpha/(nc-i+1);
end;
c3=j(nc,1,0);
c3=p<holm;
*rint c3 p holm;

*** Holland-Copenhaver procedure ***;
hoco=j(nc,1,0);
do i=1 to nc;
xx=1/(nc-i+1);
hoco[i,1]=1-(1-alpha)##xx;
end;
c4=j(nc,1,0);
c4=p<hoco;
*rint c4 hoco p;

*** Hochberg procedure ***;
hoch=j(nc,1,0);
do i=1 to nc;
hoch[i,1]=alpha/(nc-i+1);
end;
c5=j(nc,1,0);
c5=p<hoch;
*rint c5 p hoch;

*** Hommel procedure ***;
p2=j(nc,1,0);
p2[1,1]=p[6,1];
p2[2,1]=p[5,1];
p2[3,1]=p[4,1];
p2[4,1]=p[3,1];
p2[5,1]=p[2,1];
p2[6,1]=p[1,1];

*={.2670,.0633,.0633,.0350,.0255,.0129,.0089,.0011,.0301,.0000};
*c=10;
*lpha=.10;

hn=nc;
m=-1;
pp=1;
z=nc;
do i=1 to nc while(pp > m);
do k=1 to i  while(pp > m);
m=k*alpha/i;
 hn=nc-i+k;
pp= p[hn,1];
if pp<=m then z=i-1;
end;
```

```
if pp<=m then i=i-1;
end;

homm=j(nc,1,alpha);
if z>0 then
homm=homm/z;
c6=j(nc,1,0);
c6=p<homm;
*rint c6 p homm i;

*** Rom procedure ***;
rom=j(nc,1,0);
rom[1,1]=.0085;
rom[2,1]=.0102;
rom[3,1]=.0127;
rom[4,1]=.0169;
rom[5,1]=.0250;
rom[6,1]=.0500;
c7=p<rom;
*rint c7 p rom;


***holland copenhaver step-up***;
hoc2=j(nc,1,0);
do i=1 to nc;
xx=1/(nc-i+1);
hoc2[i,1]=1-(1-alpha)##xx;
end;
c8=j(nc,1,0);
c8=p<hoc2;
*rint c8 hoc2 p;

ct1=c1[+, ];
ct2=c2[+, ];

ct3=0;
do i=1 to nc while (c3[i,1]=1);
                ct3=ct3+1;
end;

ct4=0;
do i=1 to nc while (c4[i,1]=1);
                ct4=ct4+1;
end;

ct5=nc;
do i=nc to 1 by -1 while(c5[i,1]=0);
                ct5=ct5-1;
end;

ct6=nc;
do i=nc to 1 by -1 while(c6[i,1]=0);
                ct6=ct6-1;
end;

ct7=nc;
do i=nc to 1 by -1 while(c7[i,1]=0);
                ct7=ct7-1;
end;
ct8=nc;
```

```
do i=nc to 1 by -1 while(c8[i,1]=0);
                      ct8=ct8-1;
end;

count[1,1]=count[1,1]+ct1;
count[1,2]=count[1,2]+ct2;
count[1,3]=count[1,3]+ct3;
count[1,4]=count[1,4]+ct4;
count[1,5]=count[1,5]+ct5;
count[1,6]=count[1,6]+ct6;
count[1,7]=count[1,7]+ct7;
count[1,8]=count[1,8]+ct8;

if ct1=0 then pmcount[1,1]=pmcount[1,1]+0;
         else pmcount[1,1]=pmcount[1,1]+1;
if ct2=0 then pmcount[1,2]=pmcount[1,2]+0;
         else pmcount[1,2]=pmcount[1,2]+1;
if ct3=0 then pmcount[1,3]=pmcount[1,3]+0;
         else pmcount[1,3]=pmcount[1,3]+1;
if ct4=0 then pmcount[1,4]=pmcount[1,4]+0;
         else pmcount[1,4]=pmcount[1,4]+1;
if ct5=0 then pmcount[1,5]=pmcount[1,5]+0;
         else pmcount[1,5]=pmcount[1,5]+1;
if ct6=0 then pmcount[1,6]=pmcount[1,6]+0;
         else pmcount[1,6]=pmcount[1,6]+1;
if ct7=0 then pmcount[1,7]=pmcount[1,7]+0;
         else pmcount[1,7]=pmcount[1,7]+1;
if ct8=0 then pmcount[1,8]=pmcount[1,8]+0;
         else pmcount[1,8]=pmcount[1,8]+1;
end;
print  count pmcount n ii alpha nc                    ;
n=50;
c=4;
alpha=.05;
count=j(1,8,0);
pmcount=j(1,8,0);
do ii=1 to 10000;
x =rannor(j(n ,c,0));
sum=x[+,];
xpx=x *x-sum *sum/n;
s=diag(1/sqrt(vecdiag(xpx)));
r =  s*xpx*s;

nc=(c*(c-1))/2;

cp=j(nc,1,0);

cp[1,1]=r[2,1];
cp[2,1]=r[3,1];
cp[3,1]=r[3,2];
cp[4,1]=r[4,1];
cp[5,1]=r[4,2];
cp[6,1]=r[4,3];

*rint cp;
top=cp#(sqrt(n-2));
bo=sqrt(1-cp##2);
t=top/bo;
p=(1 probt(abs(t),(n-2)))*2;
*rint top bo_t p r x;
```

```
b=p;
p[ rank(p),]=b;
*rint b p;

*={.0330,.0338,.0520,.0720,.0900,.2500};
*={.0400,.0300,.0520,.0800,.1100,.2200};

*** no control of type I error ***;
nocntl=j(nc,1,alpha);
c1=j(nc,1,0);
c1=p<nocntl;
*rint nocntl c1 p;

*** original Bonferroni procedure ***;
bonf=j(nc,1,alpha);
bonf=bonf/nc;
c2=j(nc,1,0);
c2=p<bonf;
*rint c2 bonf p;

*** Holm procedure ***;
holm=j(nc,1,0);
*rint holm;
do i=1 to nc;
holm[i,1]=alpha/(nc-i+1);
end;
c3=j(nc,1,0);
c3=p<holm;
*rint c3 p holm;

*** Holland-Copenhaver procedure ***;
hoco=j(nc,1,0);
do i=1 to nc;
xx=1/(nc-i+1);
hoco[i,1]=1-(1-alpha)##xx;
end;
c4=j(nc,1,0);
c4=p<hoco;
*rint c4 hoco p;

*** Hochberg procedure ***;
hoch=j(nc,1,0);
do i=1 to nc;
hoch[i,1]=alpha/(nc-i+1);
end;
c5=j(nc,1,0);
c5=p<hoch;
*rint c5 p hoch;

*** Hommel procedure ***;
p2=j(nc,1,0);
p2[1,1]=p[6,1];
p2[2,1]=p[5,1];
p2[3,1]=p[4,1];
p2[4,1]=p[3,1];
p2[5,1]=p[2,1];
p2[6,1]=p[1,1];

*={.2670,.0633,.0633,.0350,.0255,.0129,.0089,.0011,.0001,.0000};
```

```
*c=10;
*lpha=.10;

hn=nc;
m=-1;
pp=1;
z=nc;
do i=1 to nc while(pp > m);
do k=1 to i  while(pp > m);
m=k*alpha/i;
 hn=nc-i+k;
pp= p[hn,1];
if pp<=m then z=i-1;
end;
if pp<=m then i=i-1;
end;

homm=j(nc,1,alpha);
if z>0 then
homm=homm/z;
c6=j(nc,1,0);
c6=p<homm;
*rint c6 p homm i;

*** Rom procedure ***;
rom=j(nc,1,0);
rom[1,1]=.0085;
rom[2,1]=.0102;
rom[3,1]=.0127;
rom[4,1]=.0169;
rom[5,1]=.0250;
rom[6,1]=.0500;
c7=p<rom;
*rint c7 p rom;


***holland copenhaver step-up***;
hoc2=j(nc,1,0);
do i=1 to nc;
xx=1/(nc-i+1);
hoc2[i,1]=1-(1-alpha)##xx;
end;
c8=j(nc,1,0);
c8=p<hoc2;
*rint c8 hoc2 p;

ct1=c1[+, ];
ct2=c2[+, ];

ct3=0;
do i=1 to nc while (c3[i,1]=1);
                ct3=ct3+1;
end;

ct4=0;
do i=1 to nc while (c4[i,1]=1);
                ct4=ct4+1;
end;

ct5=nc;
```

```
do i=nc to 1 by -1 while(c5[i,1]=0);
                 ct5=ct5-1;
end;

ct6=nc;
do i=nc to 1 by -1 while(c6[i,1]=0);
                 ct6=ct6-1;
end;

ct7=nc;
do i=nc to 1 by -1 while(c7[i,1]=0);
                 ct7=ct7-1;
end;
ct8=nc;
do i=nc to 1 by -1 while(c8[i,1]=0);
                 ct8=ct8-1;
end;

count[1,1]=count[1,1]+ct1;
count[1,2]=count[1,2]+ct2;
count[1,3]=count[1,3]+ct3;
count[1,4]=count[1,4]+ct4;
count[1,5]=count[1,5]+ct5;
count[1,6]=count[1,6]+ct6;
count[1,7]=count[1,7]+ct7;
count[1,8]=count[1,8]+ct8;

if ct1=0 then pmcount[1,1]=pmcount[1,1]+0;
         else pmcount[1,1]=pmcount[1,1]+1;
if ct2=0 then pmcount[1,2]=pmcount[1,2]+0;
         else pmcount[1,2]=pmcount[1,2]+1;
if ct3=0 then pmcount[1,3]=pmcount[1,3]+0;
         else pmcount[1,3]=pmcount[1,3]+1;
if ct4=0 then pmcount[1,4]=pmcount[1,4]+0;
         else pmcount[1,4]=pmcount[1,4]+1;
if ct5=0 then pmcount[1,5]=pmcount[1,5]+0;
         else pmcount[1,5]=pmcount[1,5]+1;
if ct6=0 then pmcount[1,6]=pmcount[1,6]+0;
         else pmcount[1,6]=pmcount[1,6]+1;
if ct7=0 then pmcount[1,7]=pmcount[1,7]+0;
         else pmcount[1,7]=pmcount[1,7]+1;
if ct8=0 then pmcount[1,8]=pmcount[1,8]+0;
         else pmcount[1,8]=pmcount[1,8]+1;
end;
print  count pmcount n ii alpha nc                    ;
n=100;
c=4;
alpha=.05;
count=j(1,8,0);
pmcount=j(1,8,0);
do ii=1 to 10000;
x =rannor(j(n ,c,0));
sum=x[+,];
xpx=x *x-sum *sum/n;
s=diag(1/sqrt(vecdiag(xpx)));
r = s*xpx*s;

nc=(c*(c-1))/2;

cp=j(nc,1,0);
```

```
cp[1,1]=r[2,1];
cp[2,1]=r[3,1];
cp[3,1]=r[3,2];
cp[4,1]=r[4,1];
cp[5,1]=r[4,2];
cp[6,1]=r[4,3];

*rint cp;
top=cp#(sqrt(n-2));
bo=sqrt(1-cp##2);
t=top/bo;
p=(1-probt(abs(t),(n-2)))*2;
*rint top bo t p r x;

b=p;
p[ rank(p),]=b;
*rint b p;

*={.0330,.0338,.0520,.0720,.0900,.2500};
*={.0400,.0300,.0520,.0800,.1100,.2200};

*** no control of type I error ***;
nocntl=j(nc,1,alpha);
c1=j(nc,1,0);
c1=p<nocntl;
*rint nocntl c1 p;

*** original Bonferroni procedure ***;
bonf=j(nc,1,alpha);
bonf=bonf/nc;
c2=j(nc,1,0);
c2=p<bonf;
*rint c2 bonf p;

*** Holm procedure ***;
holm=j(nc,1,0);
*rint holm;
do i=1 to nc;
holm[i,1]=alpha/(nc-i+1);
end;
c3=j(nc,1,0);
c3=p<holm;
*rint c3 p holm:

*** Holland-Copenhaver procedure ***;
hoco=j(nc,1,0);
do i=1 to nc;
xx=1/(nc-i+1);
hoco[i,1]=1-(1-alpha)##xx;
end;
c4=j(nc,1,0);
c4=p<hoco;
*rint c4 hoco p:

*** Hochberg procedure ***;
hoch=j(nc,1,0);
do i=1 to nc;
hoch[i,1]=alpha/(nc-i+1);
end;
```

```
c5=j(nc,1,0);
c5=p<hoch;
*rint c5 p hoch;

*** Hommel procedure ***;
p2=j(nc,1,0);
p2[1,1]=p[6,1];
p2[2,1]=p[5,1];
p2[3,1]=p[4,1];
p2[4,1]=p[3,1];
p2[5,1]=p[2,1];
p2[6,1]=p[1,1];

*={.2670,.0633,.0633,.0350,.0255,.0129,.0089,.0011,.0001,.0000};
*c=10;
*lpha=.10;

hn=nc;
m=-1;
pp=1;
z=nc;
do i=1 to nc while(pp > m);
do k=1 to i  while(pp > m);
m=k*alpha/i;
 hn=nc-i+k;
pp= p[hn,1];
if pp<=m then z=i-1;
end;
if pp<=m then i=i-1;
end;

homm=j(nc,1,alpha);
if z>0 then
homm=homm/z;
c6=j(nc,1,0);
c6=p<homm;
*rint c6 p homm i;

*** Rom procedure ***;
rom=j(nc,1,0);
rom[1,1]=.0085;
rom[2,1]=.0102;
rom[3,1]=.0127;
rom[4,1]=.0169;
rom[5,1]=.0250;
rom[6,1]=.0500;
c7=p<rom;
*rint c7 p rom;

***holland copenhaver step-up***;
hoc2=j(nc,1,0);
do i=1 to nc;
xx=1/(nc-i+1);
hoc2[i,1]=1-(1-alpha)##xx;
end;
c8=j(nc,1,0);
c8=p<hoc2;
*rint c8 hoc2 p;
```

```
ct1=c1[+, ];
ct2=c2[+, ];

ct3=0;
do i=1 to nc while (c3[i,1]=1);
                 ct3=ct3+1;
end;

ct4=0;
do i=1 to nc while (c4[i,1]=1);
                 ct4=ct4+1;
end;

ct5=nc;
do i=nc to 1 by -1 while(c5[i,1]=0);
                 ct5=ct5-1;
end;

ct6=nc;
do i=nc to 1 by -1 while(c6[i,1]=0);
                 ct6=ct6-1;
end;

ct7=nc;
do i=nc to 1 by -1 while(c7[i,1]=0);
                 ct7=ct7-1;
end;
ct8=nc;
do i=nc to 1 by -1 while(c8[i,1]=0);
                 ct8=ct8-1;
end;

count[1,1]=count[1,1]+ct1;
count[1,2]=count[1,2]+ct2;
count[1,3]=count[1,3]+ct3;
count[1,4]=count[1,4]+ct4;
count[1,5]=count[1,5]+ct5;
count[1,6]=count[1,6]+ct6;
count[1,7]=count[1,7]+ct7;
count[1,8]=count[1,8]+ct8;

if ct1=0 then pmcount[1,1]=pmcount[1,1]+0;
         else pmcount[1,1]=pmcount[1,1]+1;
if ct2=0 then pmcount[1,2]=pmcount[1,2]+0;
         else pmcount[1,2]=pmcount[1,2]+1;
if ct3=0 then pmcount[1,3]=pmcount[1,3]+0;
         else pmcount[1,3]=pmcount[1,3]+1;
if ct4=0 then pmcount[1,4]=pmcount[1,4]+0;
         else pmcount[1,4]=pmcount[1,4]+1;
if ct5=0 then pmcount[1,5]=pmcount[1,5]+0;
         else pmcount[1,5]=pmcount[1,5]+1;
if ct6=0 then pmcount[1,6]=pmcount[1,6]+0;
         else pmcount[1,6]=pmcount[1,6]+1;
if ct7=0 then pmcount[1,7]=pmcount[1,7]+0;
         else pmcount[1,7]=pmcount[1,7]+1;
if ct8=0 then pmcount[1,8]=pmcount[1,8]+0;
         else pmcount[1,8]=pmcount[1,8]+1;
end;
print  count pmcount n ii alpha nc          ;
quit;
```

Example of program used in statistical power calculating

```
options pagesize=54;
proc iml wrkspace=300;
N=50 ;
c=4;
alpha=.20;
title 'n= 60 c=4 alpha=.20 second method 3 corr';
****  Define matrices for counting  ***;
p12__cr  =j(1,8,0);
p13__cr  =j(1,8,0);
p14__cr  =j(1,8,0);
p23__cr  =j(1,8,0);
p24__cr  =j(1,8,0);
p34__cr  =j(1,8,0);
p1234cr  =j(1,8,0);
patlecr  =j(1,8,0);
F={.47712   .67175  .52188 -.22080,
   .77199   .41516 -.32254  .35727,
   .77199 -.41516 -.32254 -.35727,
   .47712 -.67175  .52188  .22080};
***  number of iteration  ***;

DO II=1 TO 10000;

***  compute correlation matrix  ***;

x =rannor(j(n ,c,0));
xhat=t(x);
yhat=f*xhat;
y=t(yhat);
sum=y[+,];
ypy=y *y-sum *sum/n;
s=diag(1/sqrt(vecdiag(ypy)));
r =  s*ypy*s;
nc=(c*(c-1))/2;        ***  nc is the number of correlations  ***;
cp=j(nc,1,0);          ***  define matrix cp with nc rows, 1 col. 0 value;

****  put individual correlations into cp  in a defined order  ***;

cp[1,1]=r[2,1];
cp[2,1]=r[3,1];
cp[3,1]=r[3,2];
cp[4,1]=r[4,1];
cp[5,1]=r[4,2];
cp[6,1]=r[4,3];

***  computing p values for the correlation  ***;

top=cp#(sqrt(n-2));
bo=sqrt(1-cp##2);
t=top/bo;
p=(1-probt(abs(t),(n-2)))*2;

***  ranking the p values  ***;
```

```
rp=rank(p);
corl2=rp[1,1];
cor23=rp[3,1];
cor34=rp[6,1];
b=p;
p[ rank(p),]=b;
*rint r p corl2 cor23 cor34;

****                                            ***
****    computing criteria value for each method   ***
****                                            ***;

*** no control of type I error ***;
nocntl=j(nc,1,alpha);
cl=j(nc,1,0);
cl=p<nocntl;
*rint cl p nocntl;

*** original Bonferroni procedure ***;
bonf=j(nc,1,alpha);
bonf=bonf/nc;
c2=j(nc,1,0);                         *** c2 is comparison of obtained;
c2=p<bonf;                            *** p value with criteria;
*rint c2 bonf p;

*** Holm procedure ***;
holm=j(nc,1,0);
*rint holm;
do i=1 to nc;
holm[i,1]=alpha/(nc-i+1);
end;
c3=j(nc,1,0);
c3=p<holm;
*rint c3 p holm;

*** Holland-Copenhaver procedure ***;
hoco=j(nc,1,0);
do i=1 to nc;
xx=1/(nc-i+1);
hoco[i,1]=1-(1-alpha)##xx;
end;
c4=j(nc,1,0);
c4=p<hoco;
*rint c4 hoco p;

*** Hochberg procedure ***;
hoch=j(nc,1,0);
do i=1 to nc;
hoch[i,1]=alpha/(nc-i+1);
end;
c5=j(nc,1,0);
c5=p<hoch;
*rint c5 p hoch;
```

```
*** Hommel procedure ***;
p2=j(nc,1,0);
p2[1,1]=p[6,1];
p2[2,1]=p[5,1];
p2[3,1]=p[4,1];
p2[4,1]=p[3,1];
p2[5,1]=p[2,1];
p2[6,1]=p[1,1];
hn=nc;
m=-1;
pp=1;
z=nc;
do i=1 to nc while(pp > m);
do k=1 to i  while(pp > m);
m=k*alpha/i;
 hn=nc-i+k;
pp= p[hn,1];
if pp<=m then z=i-1;
end;
if pp<=m then i=i-1;
end;
homm=j(nc,1,alpha);
if z>0 then
homm=homm/z;
c6=j(nc,1,0);
c6=p<homm;
*rint c6 p homm i;
*** Rom procedure ***;
rom=j(nc,1,0);
rom[1,1]=.0364;
rom[2,1]=.0434;
rom[3,1]=.0537;
rom[4,1]=.0700;
rom[5,1]=.1000;
rom[6,1]=.2000;
c7=p<rom;
*rint c7 p rom;

*** Holland-Copenhaver procedure (step up) ****;
hoc2=j(nc,1,0);
do i=1 to nc;
xx=1/(nc-i+1);
hoc2[i,1]=1-(1-alpha)##xx;
end;
c8=j(nc,1,0);
c8=p<hoc2;
*rint c8 hoc2 p;

ct1=j(nc,1,0);
ct2=j(nc,1,0);
ct3=j(nc,1,0);
ct4=j(nc,1,0);
ct5=j(nc,1,1);
ct6=j(nc,1,1);
```

```
ct7=j(nc,1,1);
ct8=j(nc,1,1);

ct1=c1;
ct2=c2;

do i=1 to nc while (c3[i,1]=1);
                   ct3[i,1]=1;
end;

do i=1 to nc while (c4[i,1]=1);
                   ct4[i,1]=1;
end;

do i=nc to 1 by -1 while(c5[i,1]=0);
                   ct5[i,1]=0;
end;

do i=nc to 1 by -1 while(c6[i,1]=0);
                   ct6[i,1]=0;
end;

do i=nc to 1 by -1 while(c7[i,1]=0);
                   ct7[i,1]=0;
end;

do i=nc to 1 by -1 while(c8[i,1]=0);
                   ct8[i,1]=0;
end;

*** counting within the matrices correct reject and type ii error;

c12__cr=j(1, 8,0);     **c12__   is the count for r12 sig;
c23__cr=j(1, 8,0);     **c23__   is the count for r23 sig;
c34__cr=j(1, 8,0);     **c34__   is the count for r34 sig;
c1234cr=j(1, 8,0);     **c1234 for r12 and r23 and r34 sig;

***Start counting r1234 sig for 8 methods;

IF(CT1[COR12,1]=1 & CT1[COR23,1]=1
                              & CT1[COR34,1]=1) THEN
                   c1234cr[1,1]=c1234cr[1,1]+1;
IF(CT2[COR12,1]=1 & CT2[COR23,1]=1
                              & CT2[COR34,1]=1) THEN
                   c1234cr[1,2]=c1234cr[1,2]+1;
IF(CT3[COR12,1]=1 & CT3[COR23,1]=1
                              & CT3[COR34,1]=1) THEN
                   c1234cr[1,3]=c1234cr[1,3]+1;
IF(CT4[COR12,1]=1 & CT4[COR23,1]=1
                              & CT4[COR34,1]=1) THEN
                   c1234cr[1,4]=c1234cr[1,4]+1;
IF(CT5[COR12,1]=1 & CT5[COR23,1]=1
                              & CT5[COR34,1]=1) THEN
                   c1234cr[1,5]=c1234cr[1,5]+1;
```

47

FILE: RD9   SAS   A1 UGA VM/ESA2.1 CMS   PAGE 00005

```
IF(CT6[COR12,1]=1 & CT6[COR23,1]=1
                              & CT6[COR34,1]=1) THEN
                    c1234cr[1,6]=c1234cr[1,6]+1;
IF(CT7[COR12,1]=1 & CT7[COR23,1]=1
                              & CT7[COR34,1]=1) THEN
                    c1234cr[1,7]=c1234cr[1,7]+1;
IF(CT8[COR12,1]=1 & CT8[COR23,1]=1
                              & CT8[COR34,1]=1) THEN
                    c1234cr[1,8]=c1234cr[1,8]+1;
p1234cr  =p1234cr   +c1234cr;
end;
print p1234cr ;
print n ii alpha c;
quit;
run;
```