อำนาจการทดสอบในการทดสอบรายคู่เชิงซ้อน

Power in Multiple Testing

*ดร.สุชาดา บวรกิติวงศ์**

*Suchada Bowarnkitiwong, Ph. D.**

บทคัดย่อ

งานวิจัยนี้มีจุดมุ่งหมายเพื่อเปรียบเทียบอำนาจการทดสอบในการทดสอบสัมประสิทธิ์สหสัมพันธ์อย่างง่ายเชิงซ้อน (เมทริกซ์สัมประสิทธิ์สหสัมพันธ์) โดยใช้ข้อมูลจำลองในการตรวจสอบอำนาจการทดสอบทั้ง 3 แบบ คือ อำนาจการทดสอบเมื่อมีความสัมพันธ์จริงอย่างน้อย 1 คู่ เมื่อมีความสัมพันธ์จริงทุกคู่ และอำนาจการทดสอบโดยเฉลี่ย ผลการตรวจสอบปรากฏว่า เมื่อมีความสัมพันธ์จริงอย่างน้อย 1 คู่ วิธีการทั้ง 7 ให้ผลไม่ต่างกัน แต่ในกรณีที่เหลือ อำนาจการทดสอบของ 6 วิธีที่ปรับปรุงจาก Bonferroni ให้ผลดีกว่าวิธีดั้งเดิม ผลที่ได้จาก 6 วิธีดังกล่าวไม่ต่างกันมากนัก 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 Rom and the Holland-Copenhaver (step-up) methods having the highest power.

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 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

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

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.

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 in 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 significant levels for each individual hypothesis. The SD (Holm, Holland-Copenhaver) procedures start the testing with the hypothesis with the smallest p-value, whereas the SU (Hochberg, Hommel, Rom) start the testing with the hypothesis with the largest p-value. In this study, we 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.

Original Bonferroni Procedure (SS)

The original Bonferroni procedure computes $\alpha'=\alpha/m$. The hypotheses with $p < \alpha'$ are rejected.

Holm Procedure (SD)

Holm (1979) proposed sequentially setting different significance levels for rejecting each individual hypothesis : let $p_{(1)},\ldots,p_{(m)}$ be the ordered p-values and $H_{(1)},\ldots,H_{(m)}$ be the corresponding hypotheses. Holm procedure rejects $H_{(1)}$ to $H_{(i-1)}$ if i is the smallest integer such that $p_{(i)} > \alpha/(m-i+1)$.

Holland and Copenhaver procedure (SD)

Let $P_{(1)},\ldots,p_{(m)}$ be the order p values and $H_{(1)},\ldots,H_{(m)}$ be the corresponding hypotheses. Suppose i is the smallest integer. The Holland-Copenhaver procedure rejects $H_{(1)}$ to $H_{(i-1)}$ such that $p_{(i)} > 1-(1-\alpha)^{1/(m-i+1)}$

Hochberg Procedure (SU)

Hochberg (1988) developed the first step-up approach. Hochberg procedure rejects $H_{(1)}$ to $H_{(i)}$ for any i=m, m-1,...,1 if $p_{(i)} < \alpha/(m-i+1)$.

Hommel Procedure (SU)

This procedure includes two stages. 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. The uniqueness of the Hommel procedure is that it not only considers the order of the tests but also takes the obtained p-values into the calculation while computing the α'.

Let $J = \{i' \in \{1,...,m\} : p_{(m-i'+k)} > k\alpha/i' ; k=1,...,i'\}$. Then if J is nonempty, reject $H_{(i)}$ whenever $p_{(i)} \leq \alpha/j'$ where j' is the largest number in J. If J is empty, reject all $H_{(i)}$ (i=1,...,m).

Rom Procedure (SU)

Rom (1990) developed a very complicated procedure. With this procedure, we 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 computing equation for solving α_i's can be divided into three parts. The first part is $\alpha^1+\alpha^2+...+\alpha^{i-1}$ and the second part is $\binom{i}{1}(\alpha'^{\,i-1}_{(2)})+\binom{i}{2}(\alpha'^{\,i-2}_{3})+...+\binom{i}{i-2}(\alpha'^{\,2}_{(i-1)})$. The third part is to solve for α', which subtracts the second part from the first part, and divide the difference by i.

Holland and Copenhaver procedure (SU)

An approach not previously considered is an application of the Holland and Copenhaver as a step-up 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 the $p_{(i)} < 1-(1-\alpha)^{1/(m-i+1)}$. 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. We 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.

Method

Computer programs were written using SAS/IML (1990) to generate correlation matrices for the purpose of comparing the differences in statistical power and Type I error rates among the seven methods. Four factors are considered : number of variables, sample size, overall Type I error rate, and the number of true relationships in a given matrix. Data were generated for 4 and 6 variables with the overall Type I error rate set at .05 and .20. The true relationships among the variables were simulated for the following situations : To study the Type I error rates all variables generated are independent of each other ; To study the three conceptualizations of power, matrices are generated in which one, two, three, or five pairs of variables are correlated .4 or .2 while the others are independent. The partial Type I error rates are also considered.

The SAS-RANNOR function is used to generate the normal random numbers. The matrices containing true relationships are generated using the procedure suggested by Kaiser & Dickman (1962). For each condition, 10,000 replications are generated. The program includes the following modules : (1) Compute a correlation coefficient from the generated random numbers, (2) Compute a p-value corresponding to each correlation coefficient, (3) Sort all computed p-values in a correlation matrix by ascending order, (4) Apply original Bonferroni procedure and the six modified Bonferroni procedures to each of the correlation matrices, using overall Type I error rate per matrix of .05 and .20. The number of hypotheses rejected by each procedure is recorded. Under the complete null, the proportion of matrices rejecting at least one hypothesis is recorded. For non-null matrices the proportion of matrices in which all true relationships are identified are recorded as well as the proportion of matrices in which at least one true relationship is detected and the average power for detecting the true relationships.

Results

Type I Error Rates. Table 1 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 our computer programs.

All True Relationship Power. The proportion of matrices in which all of the true relationships were detected by the seven procedures are reported in Tables 2 through 5. All six of the enhancements to the Bonferroni procedure were more sensitive than the original Bonferroni approach in detecting all true relationships. The difference between the original Bonferroni and the enhancements increases as the number of true relationships increase. Very small differences in statistical power however were found between the six enhancements to the original Bonferroni method. The Holm procedure consistently had the lowest sensitivity in detecting all true relationships while the Rom and the Holland-Copenhaver step-up procedures had the greatest power.

At Least One True Relationship Power. Tables 6 and 7 presents the proportion of matrices in which at least one true relationship was detected when the significance level equaled .05 and .20 respectively. The results indicate almost no difference between the original Bonferroni and the enhancements.

Average Power. The average proportion of true relationships detected per matrix is presented in Tables 8 and 9. 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.

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 enchancements 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 we used computer generated data to investigate Type I error rates and three definitions of statistical power to compare six suggested enhancements to the Bonferroni and we proposed still another enhancement that has not been previously considered based on the Holland-Copenhaver approach. Our 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, we 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 superior to the original Bonferroni procedure under all true relationship power and average power. However, we 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 we gain greater power only when alpha is large (.20) and there are large number of true relationships in the matrix. Therefore, while we 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, we 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 analysts. 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.

Table 1 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

Hcl=Holland-Copenhaver (step-down) procedure

Hb=Hochberg procedure

Homm=Hommel procedure

Rom=Rom procedure

Hc2=Holland-Copenhaver (step-up) procedure

Table 2 All true relationship power for alpha =.05 k = 4

#ofsig. correlation	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	.906	.906
	400	.915	.937	.937	.938	.938	.938	.938
	450	.947	.963	.963	.963	.963	.963	.964

Table 3 All true relationship power for alpha = .20 k = 4

#ofsig. correla tion	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 4 All true relationship power for alpha = .05 k = 6

#ofsig. correla tion	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	.496	.497	.499	.500
	100	.892	.892	.893	.892	.892	.893	.893
3	50	.113	.127	.130	.127	.128	.129	.130
	60	.220	.238	.242	.239	.239	.240	.242
	70	.355	.379	.384	.380	.380	.383	.384
	80	.481	.507	.511	.507	.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 5 All true relationship power for alpha = .20 k = 6

#ofsig. correla tion	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

Table 6 At least one true relationship power for alpha = .05

k	#ofsig. correlation	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 7 At least one true relationship power for alpha = .20

k	#ofsig. correlation	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

Table 8 Average power for alpha = .05

k	#ofsig. correla tion	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 9 Average power for alpha = .20

k	#ofsig. correla tion	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	.866
		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

References

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 Bonferroni 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 Joumal 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 Bonferroni procedures. **Biometrika, 76,** 624-625.

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

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

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