E

TWC/25/11
ORIGINAL: English
DATE: August 13, 2007

UPOV

INTERNATIONAL UNION FOR THE PROTECTION OF NEW VARIETIES OF PLANTS

GENEVA

TECHNICAL WORKING PARTY ON AUTOMATION AND COMPUTER PROGRAMS

Twenty-Fifth Session
Sibiu, Romania, September 3 to 6, 2007

MULTIPLE RANGE TESTS

Document prepared by an expert from Australia

MULTIPLE RANGE TESTS

Statisticians have developed many procedures to compare between multiple means. In general, these tests are labeled as **multiple range tests**. All are based on two or more critical values when comparing between multiple means and they are more powerful and reliable than an LSD, which is based on a single critical value.

There are a number of multiple range tests available for mean comparison. These include: Duncan's Multiple Range (DMR) test, Student-Newman-Keuls (SNK) test, Tukey's Range Procedure, Dunnett's Procedure etc. Two multiple range tests relevant for plant breeders' rights testing are discussed with examples:

These are:

(1) Duncan's Multiple Range (DMR) test and
(2) Student-Newman-Keuls (SNK) test

Before performing these tests we have to keep in mind that their legitimate use only applies when an Analysis of Variance (ANOVA) F test reveals significant differences between the treatment means.

(1) Duncan's Multiple Range test

In 1955, Duncan devised a method to compare each treatment mean with every other treatment mean. The procedure is simple and powerful and has become very popular among researchers, especially in the plant science area.

The following example will be helpful to understand the procedure. There are 2 candidates (variety A and variety B) and 4 comparators (varieties C to F) in a randomized complete block design with four replications. We want to compare the mean leaf length of these varieties. The data and the Analysis of Variance (ANOVA) are presented below.

Table 1: Leaf length (mm)

Variety	**A**	**B**	**C**	**D**	**E**	**F**
Rep 1	49.6	71.2	67.6	53.2	73.3	55.5
Rep 2	47.5	68.6	70.3	59.6	68.5	54.3
Rep 3	45.2	69.3	65.2	56.2	63.3	50.2
Rep 4	48.7	70.5	64.2	62.7	70.4	49.3
Variety mean	**47.75**	**69.90**	**66.83**	**57.93**	**68.88**	**52.33**

Table 2: ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*
Replication	46.62	3	15.54	1.90	
Varieties	1737.88	5	347.57	42.64	P ≤0.01
Error	122.25	15	**8.15**		
Total	1906.76	23			

From the ANOVA F test, it is evident that there are significant differences between the variety means: however, it does not in any way indicate which means are different or the magnitude of differences. A pairwise t test could be used to compare each candidate with each comparator or a DMR test could be performed to find the differences between the various mean combinations.

Step 1: Rank the means in ascending order.

Rank1	Rank2	Rank3	Rank4	Rank5	Rank6
A	F	D	C	E	B
47.75	52.33	57.93	66.83	68.88	69.90

Step 2: Calculate the standard error of the mean (s_x) derived from the Mean Square of Error (MSE) highlighted in Table 2 as:

$s_x = \sqrt{MSE/n}$ where, n= number of replications

$s_x = \sqrt{8.15/4} = 1.43$

Step 3: Use the Duncan's Multiple Range Table (Appendix 1) for ranked order of means at p = 0.01. This table is found in the back of almost every statistics text book. First, look vertically for the appropriate df value which is the df value for error in the ANOVA table (in this example 15) then move horizontally to find the r_p values for the ranked order of mean (in this example look for Rank 2 to Rank 6)

The tabulated r_p values for Rank 2 to Rank 6 for df 15 at p = 0.01 are given below:

	Rank2	Rank3	Rank4	Rank5	Rank6
r_p	4.168	4.347	4.463	4.547	4.610

Step 4: Calculate the critical value Dp, what Duncan has termed as the "shortest significant range" by using the formula $Dp = r_p \times s_x$

	Rank2	Rank3	Rank4	Rank5	Rank6
r_p	4.168	4.347	4.463	4.547	4.610
s_x	1.43	1.43	1.43	1.43	1.43
Dp	**5.96**	**6.22**	**6.38**	**6.50**	**6.59**

It may be noted that Duncan's critical value Dp gradually increases as the varieties are ranked further apart. It means that the "protection level" of the test decreases with increasing number of means making DMR test a more powerful tool than the LSD. Consequently, there is a high probability of declaring a difference when there is actually a difference between the variety means. Its reliability is one of the reasons that Duncan's procedure has been extremely popular among researchers.

Step 5: Compare the ranked means in all possible combinations. Beginning with the largest mean, each variety mean is compared to the smallest mean using the appropriate critical value (Dp). For example, the largest mean (Var B) and the smallest mean (Var A) are 6 steps apart in the ranked order therefore the correct critical value for comparing them would be 6.59. The mean difference between these two varieties is 22.15, which is higher than the Duncan's critical value therefore we can conclude that these two means are significantly different at P=0.01. Similarly, when comparing the second largest mean (Var E) with the smallest mean (Var A) we would use the critical value of 6.50 because these two means are 5 steps apart in the ranked order. The results of mean comparisons in all possible combinations using the appropriate Dp values are summarized below:

Comparison	**Mean Difference**	**Steps apart**	**Dp Value**	**Conclusion**
B - A	69.90 - 47.75 = 22.15	6	6.59	significantly different
E - A	68.88 - 47.75 = 21.13	5	6.50	significantly different
C - A	66.83 - 47.75 = 19.08	4	6.38	significantly different
D - A	57.93 - 47.75 = 10.18	3	6.22	significantly different
F - A	52.33 - 47.75 = 4.58	2	5.96	not significantly different
B - F	69.90 - 52.33 = 17.57	5	6.50	significantly different
E - F	68.88 - 52.33 = 16.55	4	6.38	significantly different
C - F	66.83 - 52.33 = 14.50	3	6.22	significantly different
D - F	57.93 - 52.33 = 5.60	2	5.96	not significantly different
B - D	69.90 - 57.93 = 11.97	4	6.38	significantly different
E - D	68.88 - 57.93 = 10.95	3	6.22	significantly different
C - D	66.83 - 57.93 = 8.90	2	5.96	significantly different
B - C	69.90 - 66.83 = 3.07	3	6.22	not significantly different
E - C	68.88 - 66.83 = 2.05	2	5.96	not significantly different
B - E	69.90 - 68.88 = 1.02	2	5.96	not significantly different

Step 6: We can now summarize the results from the above comparison as:

There are no significant differences between varietal means of B, E and C: however, they are significantly different from D, F and A.

There is no significant difference between the varietal means of D and F; there is also no significant difference between the varietal means of F and A: therefore, variety F falls within the grouping range of both variety D and A. However, variety D is significantly different from variety A.

This is usually presented in the following format:

Variety B 69.90[a]
Variety E 68.88[a]
Variety C 66.83[a]
Variety D 57.93[b]
Variety F 52.33[bc]
Variety A 47.75[c]

Note: Superscripts a, b and c are the grouping ranges within which the varietal mean values are not significantly different.

(2) Student-Newman-Keuls (SNK) Test

This test was independently developed by Newman (1939) and Keuls (1952). The name 'Student' likewise is associated as the procedure makes use of the studentised range values. The same general method and type of calculations are performed in both SNK and DMR tests. The same example and data from Table 1 is used to demonstrate the procedure.

Step 1: Rank the means in ascending order.

Rank1	Rank2	Rank3	Rank4	Rank5	Rank6
A	F	D	C	E	B
47.75	52.33	57.93	66.83	68.88	69.90

Step 2: Calculate the standard error of the mean (s_x) derived from the Mean Square of Error (MSE) as:

$s_x = \sqrt{MSE/n}$ where, n= number of replications

$s_x = \sqrt{8.15/4} = 1.43$

Step 3: Use the table (Appendix 2) for Studentised Range Values (s_r) for ranked order of mean at p = 0.01. This table is found in the back of almost every statistics text book. For convenience that table is also reproduced as Appendix 2. First, look vertically for the appropriate df value which is the df value for error in the ANOVA table (in this example 15), then move horizontally to find the s_r values for the ranked order of mean (in this example look for Rank 2 to Rank 6)

The tabulated s_r values for Rank 2 to Rank 6 for df 15 at p = 0.01 are given below :

	Rank2	Rank3	Rank4	Rank5	Rank6
s_r	4.17	4.84	5.25	5.56	5.80

Step 4: Calculate the critical value S-N-K, by using the formula S-N-K = s_r x s_x

	Rank2	Rank3	Rank4	Rank5	Rank6
s_r	4.17	4.84	5.25	5.56	5.80
s_x	1.43	1.43	1.43	1.43	1.43
S-N-K	**5.96**	**6.92**	**7.50**	**7.95**	**8.29**

Like DMR test, the critical values for S-N-K also increases with the increasing number of ranks, however the S-N-K procedure is more conservative (i.e. the means have to be further apart to be deemed significant) than DMR. This is reflected in the higher critical values for S-N-K test. A comparison between the critical values of the two tests are given below:

	Rank2	Rank3	Rank4	Rank5	Rank6
S-N-K	**5.96**	**6.92**	**7.50**	**7.95**	**8.29**
DMRT	**5.96**	**6.22**	**6.38**	**6.50**	**6.59**

Some researchers prefer the S-N-K procedure over the DMR test because they are more confident that the differences described actually exist.

Step 5: `As in DMR test, compare the ranked means in all possible combinations. Beginning with the largest mean, each variety mean is compared to the smallest mean using the appropriate critical value (S-N-K). For example, the largest mean (Var B) and the smallest mean (Var A) are 6 steps apart in the ranked order: therefore, the right critical value for comparing them would be 8.29. The mean difference between these two varieties is 22.15 which is higher than the S-N-K critical value: therefore, we can conclude that these two means are significantly different at P=0.01. Similarly, when comparing the second largest mean
(Var E) with the smallest mean (Var A) we would use the critical value of 7.95 because these two means are 5 steps apart in the ranked order. The results of mean comparisons in all possible combinations using the appropriate S-N-K values are summarized below:

Comparison	**Mean Difference**	**Steps apart**	**S-N-K Value**	**Conclusion**
B - A	69.90 - 47.75 = 22.15	6	8.29	significantly different
E - A	68.88 - 47.75 = 21.13	5	7.95	significantly different
C - A	66.83 - 47.75 = 19.08	4	7.50	significantly different
D - A	57.93 - 47.75 = 10.18	3	6.92	significantly different
F - A	52.33 - 47.75 = 4.58	2	5.96	not significantly different
B - F	69.90 - 52.33 = 17.57	5	7.95	significantly different
E - F	68.88 - 52.33 = 16.55	4	7.50	significantly different
C - F	66.83 - 52.33 = 14.50	3	6.92	significantly different
D - F	57.93 - 52.33 = 5.60	2	5.96	not significantly different
B - D	69.90 - 57.93 = 11.97	4	7.50	significantly different
E - D	68.88 - 57.93 = 10.95	3	6.92	significantly different

C - D	66.83 - 57.93 = 8.90	2	5.96	significantly different
B - C	69.90 - 66.83 = 3.07	3	6.92	not significantly different
E - C	68.88 - 66.83 = 2.05	2	5.96	not significantly different
B - E	69.90 - 68.88 = 1.02	2	5.96	not significantly different

Step 6: We can now summarize our results from the above comparisons as:

There are no significant differences between varietal means of B, E and C: however, they are significantly different from D, F and A.

There is no significant difference between the varietal means of D and F: there is also no significant difference between the varietal means of F and A: therefore, variety F falls within the grouping range of both variety D and A. However, variety D is significantly different from variety A.

This is usually presented in the following format:

Variety B 69.90^{a}
Variety E 68.88^{a}
Variety C 66.83^{a}
Variety D 57.93^{b}
Variety F 52.33^{bc}
Variety A 47.75^{c}

Note: Superscripts a, b and c are the grouping ranges within which the varietal mean values are not significantly different.

Although we got the same result, S-N-K is more conservative and powerful than DMRT because each of the mean differences were tested against a comparatively higher critical value.

Glossary:

df =	Degrees of freedom
Dp =	Shortest significant range
MS =	Mean square
MSE =	Mean square of Error
r_p =	Duncan’s multiple range values
Rep =	Replication
s_r =	Studentised range values
SS =	Sum of squares
s_x =	Standard error of mean
Var =	Variety

Appendix 1

Values for Duncan's Multiple Range Test (r_p) at p = 0.01

the rank order of means

df	2	3	4	5	6	7	8	9	10
2	14.040	14.040	14.040	14.040	14.040	14.040	14.040	14.040	14.040
3	8.261	8.321	8.321	8.321	8.321	8.321	8.321	8.321	8.321
4	6.512	6.677	6.740	6.756	6.756	6.756	6.756	6.756	6.756
5	5.702	5.893	5.989	6.040	6.065	6.074	6.074	6.074	6.074
6	5.243	5.439	5.549	5.614	5.655	5.680	5.694	5.701	5.703
7	4.949	5.145	5.260	5.334	5.383	5.416	5.439	5.454	5.464
8	4.746	4.939	5.057	5.135	5.189	5.227	5.256	5.276	5.291
9	4.596	4.787	4.906	4.986	5.043	5.060	5.118	5.142	5.160
10	4.482	4.671	4.790	4.871	4.931	4.975	5.010	5.037	5.058
11	4.392	4.579	4.697	4.780	4.841	4.887	4.924	4.952	4.975
12	4.320	4.504	4.622	4.706	4.767	4.815	4.852	4.883	4.907
13	4.260	4.442	4.560	4.644	4.706	4.755	4.793	4.824	4.850
14	4.210	4.391	4.508	4.591	4.654	4.704	4.743	4.775	4.802
15	4.168	4.347	4.463	4.547	4.610	4.660	4.700	4.733	4.760
16	4.131	4.309	4.425	4.509	4.572	4.622	4.663	4.696	4.724
17	4.099	4.275	4.391	4.475	4.539	4.589	4.630	4.664	4.693
18	4.071	4.246	4.362	4.445	4.509	4.560	4.601	4.635	4.664
19	4.046	4.220	4.335	4.419	4.483	4.534	4.575	4.610	4.639
20	4.024	4.197	4.312	4.395	4.459	4.510	4.552	4.587	4.617
24	3.956	4.126	4.239	4.322	4.386	4.437	4.480	4.516	4.546
30	3.889	4.056	4.168	4.250	4.314	4.366	4.409	4.445	4.477
40	3.825	3.988	4.098	4.180	4.244	4.296	4.339	4.376	4.408
60	3.762	3.922	4.031	4.111	4.174	4.226	4.270	4.307	4.340
120	3.702	3.858	3.965	4.044	4.107	4.158	4.202	4.239	4.272
∞	3.643	3.796	3.900	3.978	4.040	4.091	4.135	4.172	4.205

Appendix 2

Studentised Range Values (S_r) at p = 0.01

the rank order of means

df	2	3	4	5	6	7	8	9	10
1	90.03	135.00	164.30	185.60	202.20	215.80	227.20	237.00	245.60
2	14.04	19.02	22.29	24.72	26.63	28.20	29.53	30.68	31.69
3	8.26	10.62	12.17	13.33	14.24	15.00	15.64	16.20	16.69
4	6.51	8.12	9.17	9.96	10.58	11.10	11.55	11.93	12.27
5	5.70	6.98	7.80	8.42	8.91	9.32	9.67	9.97	10.24
6	5.24	6.33	7.03	7.56	7.97	8.32	8.61	8.87	9.10
7	4.95	5.92	6.54	7.01	7.37	7.68	7.94	8.17	8.37
8	4.75	5.64	6.20	6.62	6.96	7.24	7.47	7.68	7.86
9	4.60	5.43	5.96	6.35	6.66	6.91	7.13	7.33	7.49
10	4.48	5.27	5.77	6.14	6.43	6.67	6.87	7.05	7.21
11	4.39	5.15	5.62	5.97	6.25	6.48	6.67	6.84	6.99
12	4.32	5.05	5.50	5.84	6.10	6.32	6.51	6.67	6.81
13	4.26	4.96	5.40	5.73	5.98	6.19	6.37	6.53	6.67
14	4.21	4.89	5.32	5.63	5.88	6.08	6.26	6.41	6.54
15	4.17	4.84	5.25	5.56	5.80	5.99	6.16	6.31	6.44
16	4.13	4.79	5.19	5.49	5.72	5.92	6.08	6.22	6.35
17	4.10	4.74	5.14	5.43	5.66	5.85	6.01	6.15	6.27
18	4.07	4.70	5.09	5.38	5.60	5.79	5.94	6.08	6.20
19	4.05	4.67	5.05	5.33	5.55	5.73	5.89	6.02	6.14
20	4.02	4.64	5.02	5.29	5.51	5.69	5.84	5.97	6.09
24	3.96	4.55	4.91	5.17	5.37	5.54	5.69	5.81	5.92
30	3.89	4.45	4.80	5.05	5.24	5.40	5.54	5.65	5.76
40	3.82	4.37	4.70	4.93	5.11	5.26	5.39	5.50	5.60
60	3.76	4.28	4.59	4.82	4.99	5.13	5.25	5.36	5.45
120	3.70	4.20	4.50	4.71	4.87	5.01	5.12	5.21	5.30
inf	3.64	4.12	4.40	4.60	4.76	4.88	4.99	5.08	5.16

[End of Appendices and of document]