

STATISTICAL PACKAGE FOR AUGMENTED DESIGNS (SPAD)

Rajender Parsad, Abhishek Rathore and V.K. Gupta
I.A.S.R.I., Library Avenue, New Delhi-110 012
rajender@iasri.res.in

1. Introduction

In agricultural experiments often the existing practices or check varieties called control treatments are compared with new varieties or germplasms collected through exotic or domestic collections, called test treatments. In some cases experimental material for test treatments is limited and it is not possible to replicate them in the design. However, adequate material is available for replicating control treatments in the design. Augmented Designs are useful for these experimental situations. An augmented design is any standard design in control treatments augmented with additional (new or test) treatments in complete or incomplete blocks in one-way heterogeneity setting. A survey of the literature reveals that generally these experiments are conducted using an augmented randomized complete block design. In an augmented randomized complete block design, the test treatments are replicated once in the design and control treatments appear exactly once in each block. However, the experimenters often like to know how many times the control treatments be replicated in each block so as to maximize the efficiency per observation for making test treatments *vs* control treatments(s) comparisons?

It has been established that the optimum replication number of each control treatment in every block of the design is given by square root (number of control treatments + number of blocks – one) multiplied by square root of number of test treatments divided by (number of control treatments multiplied by number of blocks), provided that the number of test treatments is more than the (number of control treatments + number of blocks – one). However, for single control treatment, the optimum replication number of control treatment in each block is given by square root (number of test treatments divided by number of blocks). This optimum replication number of control treatments is obtained by maximizing the efficiency per observation in the design.

There may, however, arise a situation when the optimum replication number of control treatments is not an integer. In such a situation one has to decide as to what integral value of the replication number is to be chosen. It has been established that if the optimum replication number (r) is larger than #.42, then take $r^* = \text{int}(r) + 1$ and for values of r smaller than or equal to #.42 take $r^* = \text{int}(r)$ for $u \geq 2$. For $u = 1$, the same rule applies but the cut of value of r is taken as #.45 instead of #.42.

With this end in view, a user friendly, menu driven, graphic user interface (GUI) based Statistical Package called **STATISTICAL PACKAGE FOR AUGMENTED DESIGN (SPAD)** has been developed at IASRI by a team comprising of Abhishek Rathore, Rajender Parsad and VK Gupta. The package generates randomized layout of augmented designs and performs the analysis of data generated. For given number of test treatments, number of

control treatments and number of blocks, it computes the optimum replication number of each control treatment in every block of the design such that the efficiency per observation of the test treatments *vs* control treatment(s) comparisons is maximum. The package also provides flexibility in choosing the replication number of each control treatment in every block. The user can define the replication of each control treatment in every block. Once the user defines the number of test treatments, number of control treatments, and number of blocks in the design, the randomized layout of the design is generated. The package also provides the analysis of the data generated from augmented designs. A null hypothesis on any user-defined contrast can also be tested.

The package is very useful for classroom teaching as well as for the researchers in statistics with interest in experimental designs. The package has been developed using Microsoft Visual C++ 6.0. Software is completely stand-alone and can be installed on any hardware platform with 32 Bit Microsoft Windows Operating System. Software can be executed with minimum specification of RAM for host Operating System. Installation of SPAD takes 2 MB of hard disk space and at least 1 MB free space for its working. Software is menu driven and is very user friendly. It has a rich edit control for text editor and supports cut, copy, paste, undo, find and find-replace facilities. A Context Sensitive Help with Contents, Index and Search facilities is also available. The software is designed to assist experimenters in planning and analyzing augmented designs. A window depicting the different features is given in Figure 1.1.

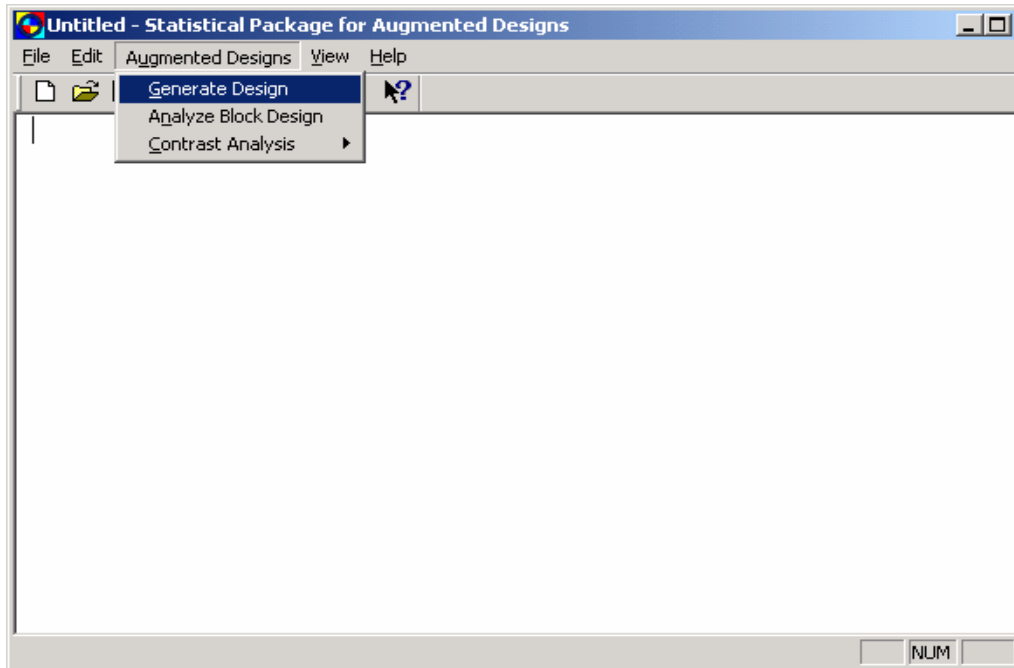


Figure 1.1: Window Depicting the Features of SPAD

These features are described in the sequel.

2. Generation of augmented design

We begin with the generation of randomized layout of augmented complete block design with each control replicated a (≥ 1) times in each block. When $a = 1$, it reduces to usual augmented randomized complete block design and when $a = r$, the number of replications of control treatments per block that maximize the efficiency per observation, then we get the randomized layout of the augmented complete block design that maximizes the efficiency per observation with respect to test treatments vs control treatment(s) contrasts. One can select the option *Augmented Designs* from the menu and then select the sub-option *Generate Design* (see figure 1.1). On selecting the sub-option *Generate Design*, a form for entering the design parameters is displayed. For generation of randomized layout of augmented design, the input in terms of number of control treatments, number of test treatments and blocks available with experimenter is required. Once the user enters the design parameters, the replication of control treatment(s) that maximizes the efficiency per observation is automatically computed and suggested to the experimenter. There is flexibility for user to change the replication number of the control treatments. To change replication of control treatments, one has to check on the “*Change Replication of Control*” check box. This will enable an edit box for replication of control treatments, where desired number of replication for control treatments can be given (see figure 2.1).

Figure 2.1: Form for entering the design parameters and change of replication of control treatment(s)

Once the desired number of replications of control treatment(s) is entered, the box for entering replication of test treatments and block sizes get activated. Software also displays the total number of plots required. The block sizes are to be entered by the user. The package accepts blocks with unequal sizes also. To make the exposition clear, we demonstrate with the help of following examples.

Example 2.1: Consider an experimental situation with $u = 2$, $w = 8$, $b = 4$ and $k = 4$. The optimum number of replications of the control treatment in each of the blocks obtained is 0.791. Suppose that the user has selected the desired replication number of control treatment(s) in each of the blocks as one. Once, the choice of r is made, the box for entering the block sizes gets activated and total number of plots required are also displayed. Once the blocks sizes are inputted, then the randomized layout of the design is generated. This randomized layout may be printed or saved in a file. Please see Figure 2.2.

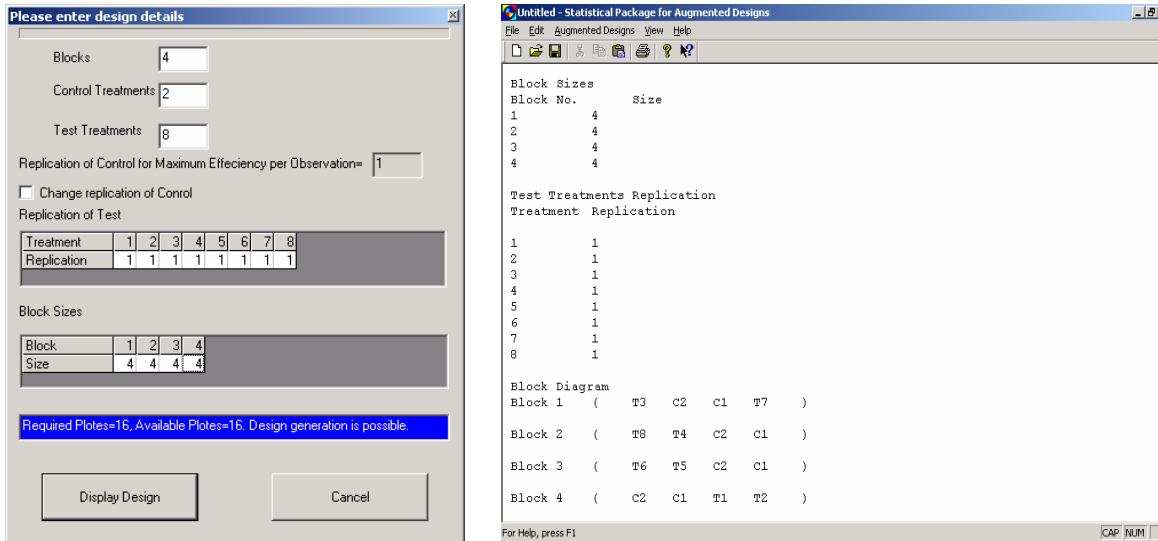


Figure 2.2: Layout for the situation where control treatment (s) are replicated once in each block

Example 2.2: Through this example we illustrate the case of the experimental situation where the replication of the control treatment(s) in each of the blocks that maximizes the efficiency per observation is greater than one. Let $u = 2$, $w = 19$, $b = 3$, $k_1 = 10$, $k_2 = 10$, $k_3 = 11$. The results obtained are displayed in figure 2.3.

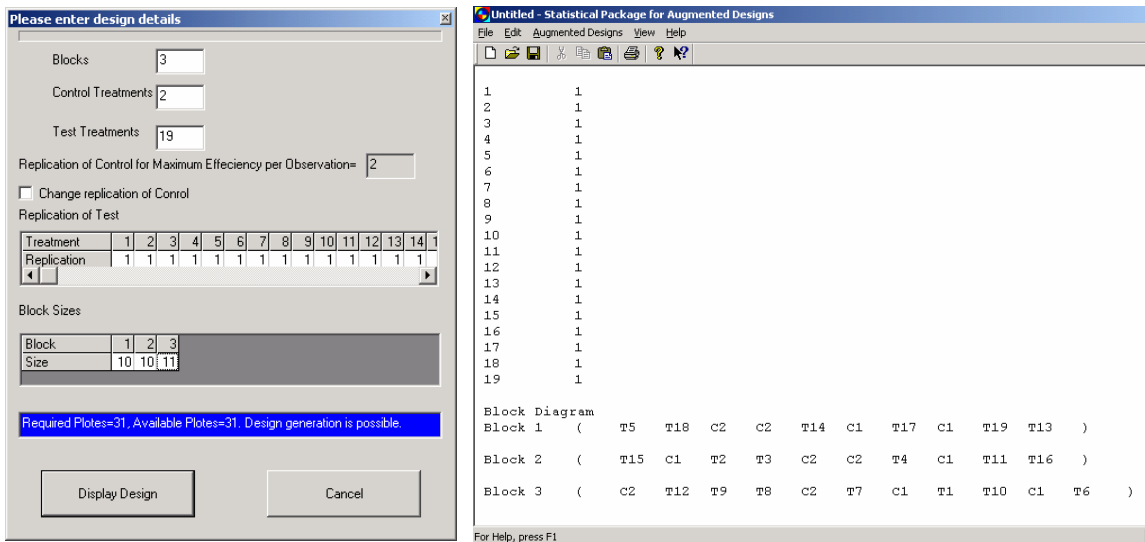


Figure 2.3: Design layout for the situation where control treatment(s) are replicated twice in each block

3. Analysis of data generated from augmented design

The data pertaining to an augmented block design is analyzed as per procedure of analysis of general block designs. The treatment sum of squares is partitioned into different components of interest *viz.* (i) among test treatments, (ii) among control treatments and (iii) among test treatments and control treatments. The pairwise comparisons of treatment means can be simplified for an augmented complete block design in which each of the control treatments appear in each block '*a*' times.

For an augmented incomplete block design, the significance of all possible pairwise treatment comparisons can be tested by automatically generating all the possible elementary treatment contrasts.

For performing the analysis of data generated through an augmented block design, an ASCII data file in a specified format is required. The existing ASCII data file can be opened in the SPAD window using File-Open options. A new data file can also be created in the SPAD window using File-New option. One can also copy and paste data into SPAD editor from any windows based software like Excel or which supports clipboard operations. For creation of data file in a specified format, the treatments are renumbered as 1, 2, ..., *u*, *u* + 1, ..., *u* + *w*. Here first *u* treatments are the control treatments and *u*+1, ..., *u* + *w* are the test treatments. Data file contains at least three columns; first column represents **block number**, second column represents **treatment number** and third column consists of **observed value** of character. If there is more than one character to be analyzed, then the characters can be entered from fourth column onwards. There is no limitation on the number of characters present in the file. All these data values must be separated by a SPACE or a TAB.

For performing the analysis of the data generated through an augmented block design, one can select the sub-option *Analyze Block Design* from Option *Augmented Design* in the menu. A click on sub-option *Analyze Block Design* displays a dialog box. In this dialog box user must specify the character to analyze this time. This box will only appear if data file has more than one character. Once a character is selected for the analysis, complete analysis with two ANOVA tables; one for testing the equality of treatment effects and another for testing the equality of block effects, R^2 , Coefficient of Variation, Root Mean Square Error (RMSE), General Mean and adjusted treatment means is generated. For partitioning the treatment sum of squares into components of interest *viz.* (i) among test treatments, (ii) among control treatments and (iii) among test treatments *vs* control treatments, one can select the sub-option *Contrast Analysis*. There are three options within the contrast analysis *viz.* (i) Augmented CB design, (ii) GBD for Tests *vs* Control(s) and (iii) User Defined Contrasts. Here *Tests* is used for test treatments and *Controls* for control treatments. If the data is generated from an augmented design in which each control treatment appears equally often in all the blocks, then the option *Augmented CB* design can be used for obtaining partitioned sum of squares and critical differences for performing all possible pairwise treatment comparisons. If the data is generated from an augmented incomplete block design, then the option *GBD for Tests vs Controls* may be used. In this option, the exact probability levels of significance of all possible pairwise treatment comparisons are given in a $(u + w) \times (u + w)$ matrix. A null hypothesis on any other contrast of interest can be tested using *User Defined Contrasts*. To make the exposition clear, we illustrate the analysis with the help of following examples.

Example 3.1: This example has been taken from Federer (1956). Four hypothetical check varieties (control treatments) are denoted by C1, C2, C3, C4 and eight hypothetical seedlings (new accessions) or test treatments are denoted by T1, T2, ..., T8 in 3 blocks of sizes 7, 6 and 7 respectively. The layout is given below.

	Experimental Units						
	1	2	3	4	5	6	7
Block 1	T8 (74)	C3 (78)	C4 (78)	T3 (70)	C1 (83)	C2 (77)	T7 (75)
Block 2	C4 (91)	C2 (81)	C1 (79)	C3 (81)	T1 (79)	T5 (78)	
Block 3	T4 (96)	C3 (87)	C1 (92)	T2 (89)	C4 (81)	C2 (79)	T6 (82)

The values in the parenthesis are the observed values on the character under investigation.

The data file is created by renumbering the treatments from 1 to 12 where treatments 1, 2, 3, 4 are the control treatments and 5, 6, 7, 8, 9, 10, 11 and 12 are the test treatments. The format of the data file for the above example is given as

Block No	Treatment No	Observed Value
1	1	83
1	2	77
1	3	78
1	4	78
1	7	70
1	11	75
1	12	74
2	1	79
2	2	81
2	3	81
2	4	91
2	5	79
2	9	78
3	1	92
3	2	79
3	3	87
3	4	81
3	8	96
3	6	89
3	10	82

Once data file is prepared and opened in the SPAD window, execute analysis module from menu by selecting Option Analyze Block Design. As there is only one character to be analyzed, therefore, a click on the Analyze Block Design option displays the analysis consisting of ANOVA tables (Block Adjusted and Treatments Adjusted), R^2 , Coefficient of Variation, Root MSE, General Mean and adjusted treatment means (see Fig 3.1.)

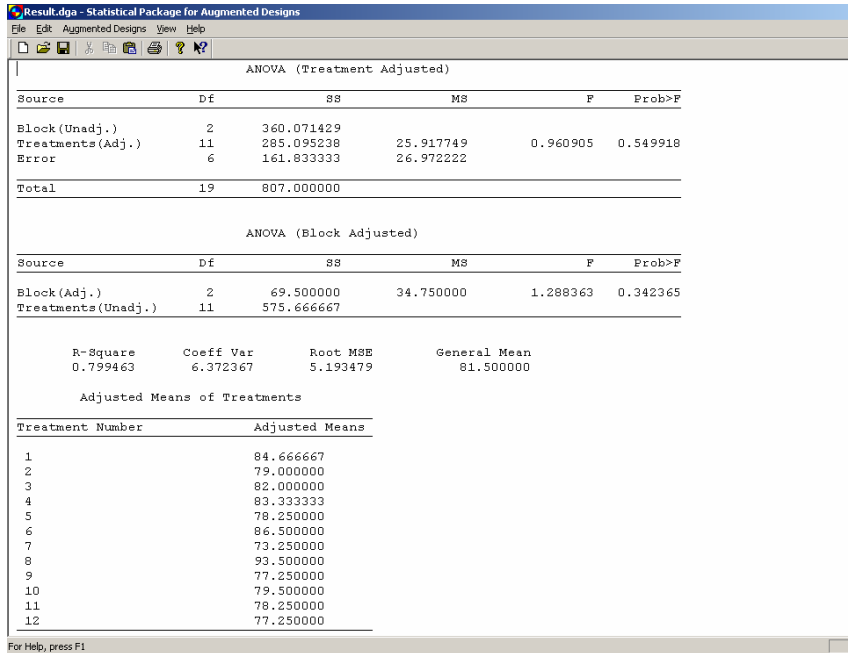


Figure 3.1 Analysis Results

Besides usual analysis of variance, experimenter is also interested in three types of contrasts namely, Among- Control (s), Among-Test(s) and Tests-Vs-Control(s) Contrasts. On selecting the Contrast Analysis option “*Augmented CB Design*” user can see Sum of Squares, Mean square, F-Cal and Prob>F of these contrasts. Four Critical Differences (CD) are also be listed at 1% and 5% level of significance see figure 3.2.

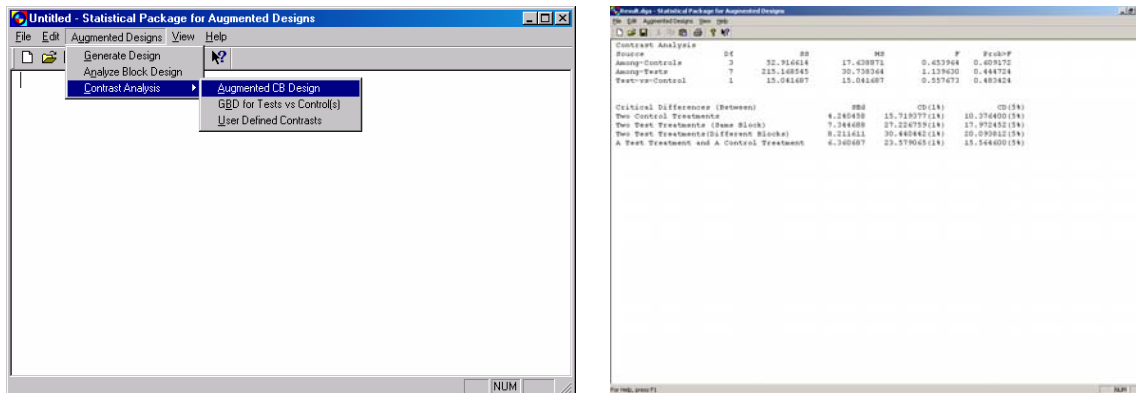


Figure 3.2: Results on Selecting Augmented CB design from Contrast Analysis option

In addition to *Augmented CB Design*, SPAD can perform two more types of contrast analysis. These are *GBD for Tests vs Control(s)* and *User Defined Contrast*. *GBD for Tests vs Control(s)* option gives user an opportunity to carry out analysis of design for test vs control comparisons for the data generated from an augmented incomplete block design. Option *User Defined Contrasts* offers a facility to user to test any desired treatment contrast. On selecting *User Defined Contrast* option a form for taking contrasts appears see figure 3.3.

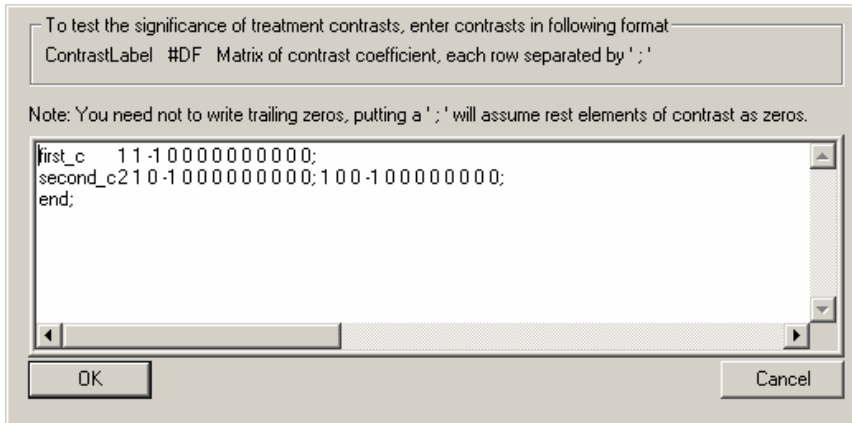


Figure 3.3: User Defined Contrasts Analysis

In this form user need to enter contrasts in a specific format, any number of contrasts can be tested simultaneously. To test a contrast, first enter its label, degree of freedom and then contrast coefficient matrix, each row separated by a semicolon ';

e.g.

```
first_c 1 1 -1 0 0 0 0 0 0 0 0 0 0 0;
second_c 2 1 0 -1 0 0 0 0 0 0 0 0 0 0; 1 0 0 -1 0 0 0 0 0 0 0 0 0 0;
_c
end;
```

Here *first_c* is the name of contrast, with 1 degrees of freedom, where as *second_c* is another contrast with 2 degree of freedom. User can enter as many contrasts as desired. Each new contrast must start with a new line. After entering contrasts, declaration must be terminated by writing *end;* in the last line. Now a click on OK button will display result in SPAD window, see Figure 3.4.

Contrast Analysis						
Source	Df	SS	MS	F	Prob>F	
first_c	1	48.166618	48.166618	1.785786	0.229885	
second_c	2	10.666656	5.333328	0.197734	0.825727	

Figure 3.4: Results of User Defined Contrasts Analysis

Example 3.2: This illustration has been taken to demonstrate the analysis of augmented complete block design where each of the control treatment(s) is replicated in each of the block more than once. For this purpose, we modify the data given in illustration 1 by taking C1 and C2 as C1 and C3 and C4 as C2. Corresponding experimental situation will have parameters $u = 2$, $w = 8$, $b = 3$, $k_1 = 7$, $k_2 = 6$ and $k_3 = 7$ and will look like as given below.

	Experimental Units						
	1	2	3	4	5	6	7
Block 1	T8 (74)	C2 (78)	C2 (78)	T3 (70)	C1 (83)	C1 (77)	T7 (75)
Block 2	C2 (91)	C1 (81)	C1 (79)	C2 (81)	T1 (79)	T5 (78)	
Block 3	T4 (96)	C2 (87)	C1 (92)	T2 (89)	C2 (81)	C1 (79)	T6 (82)

On renumbering of treatments the data file will look like this.

Block No	Treatment No	Observed Value
1	1	83
1	1	77
1	2	78
1	2	78
1	5	70
1	9	75
1	10	74
2	1	79
2	1	81
2	2	81
2	2	91
2	3	79
2	7	78
3	1	92
3	1	79
3	2	87
3	2	81
3	6	96
3	4	89
3	8	82

To carry out this type of analysis first use the same steps as explained in example 3.2.1. This displays ANOVA tables (Block Adjusted and Treatments Adjusted), R^2 , Coefficient of Variation, Root MSE, General Mean and Adjusted treatment means. Four Critical Differences (CD) are listed at 1% and 5% level of significance. The exact probability levels of significance of all possible paired treatment comparisons can be obtained in a $(u + w) \times (u + w)$ matrix by selecting GBD for Tests vs Control(s) submenu from Contrast Analysis. The results obtained are displayed in Figures 3.5(a) and 3.5(b).

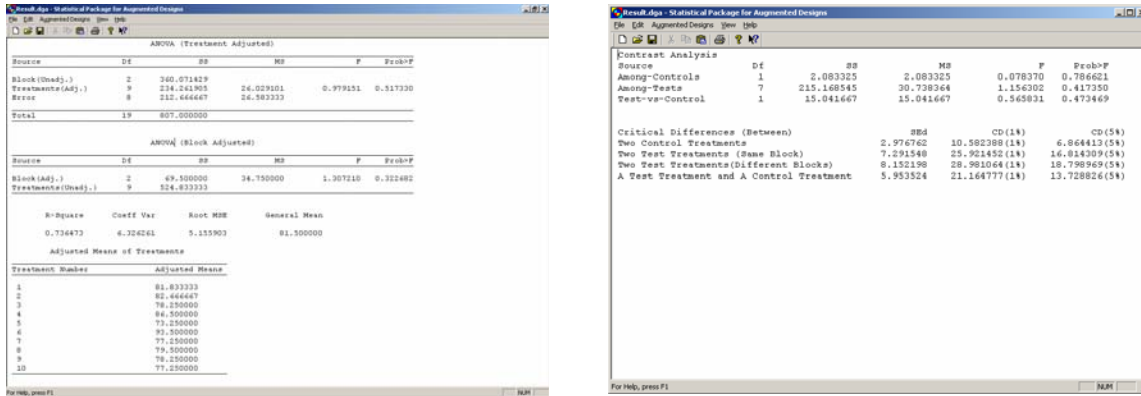


Figure 3.5(a). ANOVA Tables, Adjusted Treatment Means and Critical differences for example 3.2

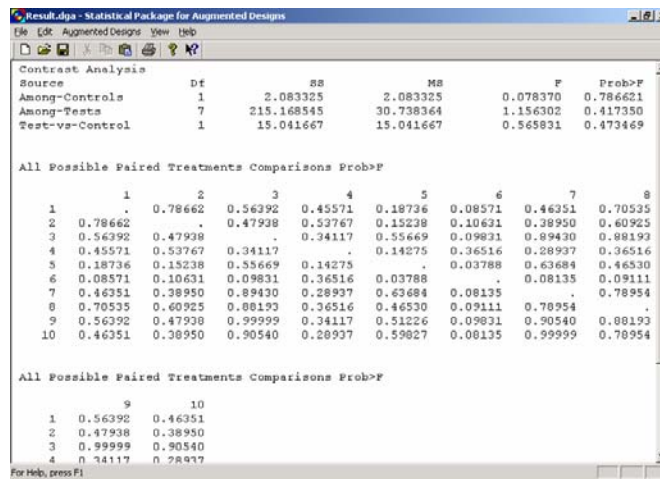


Figure 3.5(b). Exact probability levels of significance matrix

Example 3.3: This illustration demonstrates the analysis of augmented incomplete block design. In the situation each block does not have all control treatment(s). For example consider an experiment with three blocks, 8 test treatments and 4 control treatments. Let block -1 contains control treatments 2, 3, 4, block-2 have 1, 3, 4 and block-3 have 1, 2 and 4 control treatments This situation can be obtained by deleting control treatments 1, 2 and 3 from block 1, 2 and 3 respectively from data set used in example 3.2.1. The parameters of the design are $u = 4, w = 8, b = 3, k_1 = 6, k_2 = 5, k_3 = 6$. The layout is given below:

	Experimental Units					
	1	2	3	4	5	6
Block 1	T8 (74)	C3 (78)	C4 (78)	T3 (70)	C2 (77)	T7 (75)
Block 2	C4 (91)	C1 (79)	C3 (81)	T1 (79)	T5 (78)	
Block 3	T4 (96)	C1 (92)	T2 (89)	C4 (81)	C2 (79)	T6 (82)

On renumbering of treatments the data file will look like this.

Block No	Treatment No	Observed Value
1	2	77
1	3	78
1	4	78
1	7	70
1	11	75
1	12	74
2	1	79
2	3	81
2	4	91
2	5	79
2	9	78
3	1	92
3	2	79
3	4	81
3	8	96
3	6	89
3	10	82

The analysis is carried out on the similar steps as in example 3.2. For this experimental situation, the use of critical differences for paired treatment comparisons is not as simple as in case of augmented complete block designs. Therefore, the exact probability levels of significance of all possible paired treatment comparisons can be obtained in a $(u + w) \times (u + w)$ matrix by selecting GDB for Tests vs Control (s) submenu from Contrast Analysis Menu. Analysis Results are presented in Figure 3.6.

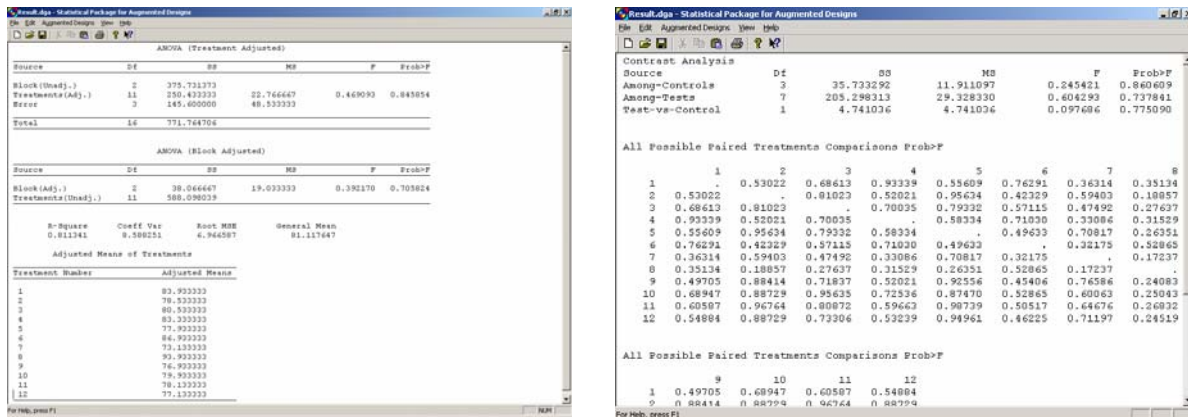


Figure 3.6. Analysis results of an Augmented Incomplete Block Design

SPAD has facility to navigate between Data File, Analysis Results and Contrast Analysis Results from view menu. Alternatively one can use shortcut keys F4, F5 and F6 for these operations respectively.