Review

Specific residue: application of orthogonal contrasts when heteroscedasticity is present

Maria Cristina Stolf Nogueira

USP/ESALQ - Depto. de Ciências Exatas, C.P.9 - 13418-900 - Piracicaba, SP - Brasil - e-mail <mcsnogue@esalq.usp.br>.

ABSTRACT: When experimental data are submitted to analysis of variance, the assumption of data homoscedasticity (variance homogeneity among treatments), associated to the adopted mathematical model must be satisfied. This verification is necessary to ensure the correct test for the analysis. In some cases, when data homoscedascity is not observed, errors may invalidate the analysis. An alternative to overcome this difficulty is the application of the specific residue analysis, which consists of the decomposition of the residual sum of squares in its components, in order to adequately test the correspondent orthogonal contrasts of interest between treatment means. Although the decomposition of the residual sum of squares is a seldom used procedure, it is useful for a better understanding of the residual mean square nature and to validate the tests to be applied. The objective of this review is to illustrate the specific residue application as a valid and adequate alternative to analyze data from experiments following completely randomized and randomized complete block designs in the presence of heteroscedasticity.
Key words: analysis of variance, completely randomized design, randomized complete block design

Resíduo específico: aplicação de contrastes ortogonais na presença da heterocedasticidade

RESUMO: Ao realizar-se a análise da variância de um conjunto de dados, pressupõe-se que o critério de homocedasticidade (homogeneidade de variâncias entre tratamentos), associada ao modelo matemático adotado, seja satisfeito. Esta verificação se faz necessária para a correta aplicação dos testes de significância. Quando não é satisfeita, em certos casos, compromete a normalidade dos erros. Uma alternativa para contornar essa deficiência é a aplicação do resíduo específico, que consiste em decompor a soma de quadrados do resíduo em componentes, correspondentes aos contrastes ortogonais de interesse, apropriados para testar cada contraste ortogonal entre médias de tratamentos. A decomposição da soma de quadrados do resíduo é um procedimento pouco utilizado, mas é útil para melhor compreensão da natureza do quadrado médio residual e garantir a validade dos testes aplicados. Nessa revisão avaliou-se a aplicação dos resíduos específicos como alternativa válida e adequada, na análise de dados obtidos de experimentos que seguem a estrutura dos delineamentos inteiramente casualizados e em blocos casualizados, na presença da heterocedasticidade.
Palavras-chave: análise da variância, delineamento inteiramente casualizado, delineamento em blocos casualizados

Introduction

The analysis of variance of experimental data requires that the assumption of homoscedasticity (similar variances among treatments), associated to the adopted mathematical model is satisfied. This verification is necessary for a correct significance of the test application. When this condition is not met the heteroscedasticity is prevailing (variance heterogeneity).

The heteroscedasticity can be classified as regular and irregular according to Steel and Torrie (1981) based on Cochran (1947). The regular type is generally originated from data non-normality and some type of relationship between means and variance treatments. In this case, the data may be transformed to have variance stability among treatments and, as a consequence, the errors will fit into an approximately normal distribution. The irregular type is characterized by certain treatments showing significantly higher variability compared to others, not necessarily presenting a relation between means and variances. In this case, Cochran and Cox (1957, 1971) recommended that such high variability treatments are omitted or that treatments are subdivided into homocedasticity groups in such way that they may present similar variances; or yet, to subdivide the residual sum of squares (SSResidual) in applicable components for the several comparisons of interest, thus obtaining specific residues.

When an analysis of variance is performed, the sum of squares of the treatments (SSTreatment) can be decomposed into components corresponding to orthogonal contrasts; in the same way, the residual sum of squares (SSResidual) can also be decomposed into their orthogonal contrast components, giving origin to the specific residues that are appropriate to test each contrast between treatment means.

The residual sum of squares (SSResidual) decomposition is not a usual procedure as the treatment sum of squares (SSTreatment) decomposition, but according to Cochran and Cox (1957, 1971), it can be applied when there are reasons suggesting the presence of irregular types of heteroscedasticity. In this case, the SSResidual decomposition is useful to better understand the residual mean square (MSResidual) nature and validate the tests to be applied.

A residual sum of squares (SSResidual) decomposition for experimental data of a randomized complete block design was presented by Steel and& Torrie (1981); initially, they established an orthogonal contrast grouping for treatments and thereafter they obtained the value of each contrast for each block. The authors concluded that if the randomized complete block design is valid, any comparison within each block is not influenced by the general level of the block. As a consequence, the variance for any comparison within blocks is appropriate to test contrasts between treatment means. The procedure was numerically shown.

In presence of the heteroscedasticity among experiments, when a group of experiments is considered, the interaction effects involving experiments (assumed as randomized effects) are influenced. An appropriate alternative to analyze the experimental data is the application of the specific residue method. With the objective to illustrate this case, Oliveira and Nogueira (2007) applied the specific residue method on sugarcane yield (t ha^{-1}) experimental data obtained from a group of eleven experiments characterized by the presence of heteroscedasticity among experiments. Each experiment had a randomized incomplete block design, arranged in a 3^3 NPK factorial (27 treatments = three blocks × nine experimental units). The confounding of two degrees of freedom corresponding to the block effects plus NPK interaction effects was considered. No replication was applied to blocks.

The objective of this review is to illustrate the application of specific residues as an alternative procedure to analyze data showing heteroscedasticity among treatments.

Material and Methods

The methods, definitions and concepts on orthogonal contrasts applied to obtain specific residues can be found in Nogueira (2004). To bypassthe irregular heteroscedasticity present in the experimental data of a randomized complete block design, Ferreira (1978) presenteda mathematical procedure to obtain the specific residue sum of squares, correspondent to the appropriate components for comparisons (orthogonal contrasts) of interest, using the orthogonal transformation method. Thus, the specific residue sum of squares of the Y_h component ($SSR(Y_h)$) is given by

$$SSR(Y_h) = \frac{1}{\sum_{i=1}^{I} c_{hi}^2} \left[\sum_{j=1}^{J} \hat{Y}_{hj}^2 - \frac{\hat{Y}_h^2}{J} \right],$$

with (J-1) degrees of freedom and $\hat{Y}_{hj}$ is the Y_{hj} contrast estimate, correspondent to the Y_h contrast application within block j, for j = 1, ..., J,

$$\hat{Y}_{hj} = \sum_{i=1}^{I} c_{hi}\, y_{ij} \quad , \text{ such that } \sum_{i=1}^{I} c_{hi} = 0 \text{ and}$$

$$\hat{Y}_h = \sum_{j=1}^{J} \hat{Y}_{hj} = \sum_{i=1}^{I} c_{hi} \bar{y}_i = \frac{1}{J} \sum_{i=1}^{I} c_{hi}\, y_{i.}$$

where I is the total number of treatments, for i = 1, ..., I; c_{hi} is the associated coefficient of the i-esimal treatment mean in the h-esimal contrast; $\hat{Y}_h$ is the h-esimal contrast estimate, for h = 1, ..., (I-1) ; y_{ij} is the observed value to i-esimal treatment in j-esimal block; $y_{i.} = \sum_{j=1}^{J} y_{ij}$ the total sum of the i-esimal treatment and $\bar{y}_i = \frac{y_{i.}}{J}$ the mean of the i-esimal treatment. Two contrasts are orthogonal when $\sum_{i=1}^{I} c_{hi} c_{h'i} = 0$, for h ≠ h'= 1, ..., (I-1).

$\sum_{h=1}^{I-1} SSR(Y_h) = SSResidual$ has (I-1) (J-1) degrees of freedom and the residual mean square for Y_h, $MSR(Y_h) = \frac{1}{J-1} SSR(Y_h)$ has (J-1) degrees of freedom.

Thus, the hypothesis $H_0: Y_h = 0$ vs. $H_a : Y_h \neq 0$, for h = 1, ..., (I - 1), is tested by the application of the F test, and $F_h = \frac{MS(Y_h)}{MSR(Y_h)} \sim F_{(1,(J-1))}$. where $MS(Y_h)$ is the mean square referred to the Y_h component, with one degree of freedom, obtained as follows:

$$MS(Y_h) = \frac{SS(Y_h)}{1} \text{ and}$$

$$SS(Y_h) = \frac{J\,\hat{Y}_h^2}{\sum_{i=1}^{I} c_{hi}^2} = \frac{J(\sum_{i=1}^{I} c_{hi} \bar{y}_{i.})^2}{\sum_{i=1}^{I} c_{hi}^2} = \frac{J(\sum_{i=1}^{I} c_{hi} \frac{y_{i.}}{J})^2}{\sum_{i=1}^{I} c_{hi}^2} = \frac{(\sum_{i=1}^{I} c_{hi} y_{i.})^2}{J\sum_{i=1}^{I} c_{hi}^2}$$

In the case of a completely randomized design experiment in presence of irregular heteroscedascity SSResidual is decomposed in specific residues as shown by Nogueira (1984) and Nogueira and Campos (1985). These authors developed the decomposition of SSResidual and presented appropriate specific residues to test each contrast, and also identified how the specific residue sum of squares refers to the Y_h component ($SSR(Y_h)$). The development of the specific residue sum of squares in relationto the Y_h component was obtained by applying the mathematical expectance (E) on $SSR(Y_h)$ of the randomized complete block design experiment, as follows:

$$E[SSR(Y_h)] = E\left[\frac{1}{\sum_{i=1}^{I} c_{hi}^2}\left[\sum_{j=1}^{J} \hat{Y}_{hj}^2 - \frac{\hat{Y}_h^2}{J}\right]\right],$$

assuming that $E(t_i) = t_i$, $E(t_i^2) = t_i^2$, $E(e_{ij}) = 0$ and $E(e_{ij}^2) = \sigma_i^2$; where t_i is the i-esimal treatment effect, e_{ij} is the experimental error associated to y_{ij}. The specific residue sum of squares for Y_h ($SSR(Y_h)$) obtained is presented as follows:

$$SSR(Y_h) = \frac{J-1}{\sum_{i=1}^{I} c_{hi}^2} \sum_{i=1}^{I} c_{hi}^2 S_i^2 \text{ , for h = 1,..., (I-1)}$$

with (J-1) degrees of freedom, and

$$S_i^2 = \frac{1}{J-1}\underbrace{[\sum_{j=1}^{J} y_{ij}^2 - \frac{(\sum_{j=1}^{J} y_{ij})^2}{J}]}_{SST_i} = \frac{SST_i}{J-1},$$

where SST_i is the i-esimal treatment sum of squares. Thus, the residual mean square for Y_h ($MSR(Y_h)$) is given by:

$$MSR(Y_h) = \frac{1}{J-1}[\frac{1}{\sum_{i=1}^{I} c_{hi}^2} \sum_{i=1}^{I} c_{hi}^2 S_i^2],$$

with n_h degrees of freedom, obtained by the application of the Satterthwaite (1941,1946) formula, and thus,

$$n_h = \frac{(\sum_{i=1}^{I} c_{hi}^2 S_i^2)^2}{\sum_{i=1}^{I} \frac{(c_{hi}^2 S_i^2)^2}{J-1}}.$$

and

$$SSResidual = \sum_{h=1}^{I-1} SSR(Y_h) + SSR(\text{among replications}),$$

with I(J-1) degrees of freedom, and the SSR(among replications) is the residual sum of squares among replications, so that

$$SSR(\text{among replications}) = \frac{J-1}{I}\sum_{i=1}^{I} S_i^2,$$

with (J-1) degrees of freedom and that residual mean square among replications (MSR(among replications) is

$$MSR(\text{among replications}) = \frac{1}{J-1}[\frac{J-1}{I}\sum_{i=1}^{I} S_i^2] = \frac{\sum_{i=1}^{I} S_i^2}{I}$$

It was also observed that

$$MSResidual = MSR(\text{among replications}) = \frac{1}{I-1}\sum_{h=1}^{(I-1)} MSR(Y_h).$$

Therefore, the hypotheses $H_o: Y_h = 0$ vs. $H_a: Y_h \neq 0$, for h = 1, ..., (I - 1) were tested by the application of the F test, and the calculated F value was obtained through the expression:

$$F_h = \frac{MS(Y_h)}{MSR(Y_h)} \sim F_{(1,n_h)},$$

where $MS(Y_h)$ is the mean square of the Y_h component, with one degree of freedom, obtained as follows:

$$MS(Y_h) = \frac{SS(Y_h)}{1};\ SS(Y_h) = \frac{\hat{Y}_h^2}{J\sum_{i=1}^{I} c_{hi}^2} = \frac{(\sum_{i=1}^{I} c_{hi} y_{i.})^2}{J\sum_{i=1}^{I} c_{hi}^2}, \text{ and}$$

the followed the approximated F distributions with one degree of freedom was referred to $MS(Y_h)$ with n_h degrees of freedom obtained by the Satterthwaite (1941, 1946) formula and to $MSR(Y_h)$ as verified by Nogueira (1984). The verification was accomplished through the application of the simulation method developed by Godoi (1978), based on Box and Miller (1958), to variables with normal and one-dimensional distributions. The Chi-square test was applied to verify the adherence of F_h with the $F_{(1,n_h)}$ distributions.

Results and Discussion

Completely randomized design

The experimental data shown in Table 1, cited by Nogueira (1984), refer to sorghum total dry matter yield, first cropping (g per pot) obtained from a completely randomized design experiment, with eight treatments and four replications, so that: Total for each treatment

$$\rightarrow y_{i.} = \sum_{j=1}^{J=4} y_{ij}$$

Sum of squares error for each treatment

$$\rightarrow SST_i = \sum_{j=1}^{J=4} y_{ij}^2 - \frac{(\sum_{j=1}^{J=4} y_{ij})^2}{4},$$

Table 1 - Sorghum plant total dry matter yields (g per pot), mean deviation sum of squares and variance estimate for each treatment (eight treatments, average of four replications).

Treatment (i)	Fertilization (N+K+S+Mg +Micro)	P-rates (Triple-superphosphate (ppm)	P-fertilizer application	Total y_i	SST_i	S^2_i with 3 DF
1	no	0		2.68	0.083	0.027
2	yes	0		1.91	0.049	0.016
3	yes	50	Located	181.13	52.141	17.380
4	yes	100	Located	226.93	144.207	48.069
5	yes	200	Located	236.98	37.045	12.348
6	yes	50	Incorporated	169.91	4.265	1.422
7	yes	100	Incorporated	195.85	45.030	15.010
8	yes	200	Incorporated	216.51	537.337	179.112

with (4 -1) degrees of freedom, where y_{ij} is the observed value (g per pot) of the i-esimal treatment in the j-esimal replication.

The variance for each treatment is given by $S_i^2 = \frac{1}{4\text{-}1} SST_i$, with (4-1) degrees of freedom and i = 1, ..., 8.

Preliminary analyses of variance results are presented in Table 2. Seven degrees of freedom for treatments and the sum of squares for treatments were decomposed according to the following group of orthogonal contrasts of interest: Y_1: control treatments versus located and incorporated P-rates; Y_2: among controls; Y_3: Located versus incorporated P-rates; Y_4: Linear effect of located P-rates; Y_5: Quadratic effect of located P-rates; Y_6: Linear effect of incorporated P-rates; Y_7: Quadratic effect of incorporated P-rates.

Contrasts Y_4 and Y_5 provided the located-P treatment effect and contrasts Y_6 and Y_7, the incorporated-P treatment effect. The coefficients of applied contrasts and some results are shown in Table 3. As P-rates are not equidistant, the coefficients attributed to Y_4, Y_5, Y_6 and Y_7 contrasts were obtained using the orthogonal polynomial coefficient procedure for non-equidistant levels developed by Nogueira (1978) and cited by Nogueira (2007). The new analysis of variance with F test results without specific residue application is presented in Table 4.

Table 2 - Preliminary analysis of variance for the sorghum experiment

Sources of Variation	DF	SS	MS
Treatment	7	16,220.4155	2,317.2022
Residual	24	820.1616	34.1734
Total	31	17,040.5771	

Note: DF is degrees of freedom; SS is Sum of Squares; MS is Mean Square.

If the model homoscedasticity assumption is satisfied, that is, if it is possible to consider that statistically $S_1^2 = S_2^2 = \cdots = S_8^2 = S^2 =$ MSResidual, the analysis presented in Table 4 is perfectly valid.

In order to verify the experimental data homoscedasticity, the Bartlett test was applied (among other tests), which is appropriate to test the following hypotheses:

$$H_0 : \sigma_1^2 = \cdots = \sigma_8^2 \quad \text{vs.} \quad H_a : \sigma_i^2 \neq \sigma_{i'}^2 \text{ , for at least one } i \neq i'.$$

The hypothesis $H_0 : \sigma_1^2 = \cdots = \sigma_8^2$ was rejected at *p*-value < 0.005 significance level, evidencing significant differences among variances due to the replications within treatments, characterizing the presence of heteroscedasticity. Once heteroscedasticity was evidenced, a procedure should be applied to overcome this situation. One alternative was the use of the specific residue as the F test denominator, to test each contrast defined in Table 3. This procedure consisted of the decomposition of all residual degrees of freedom (24), and consequently, the residual sum of squares obtaining the specific residue for each contrast:

$$SSR(Y_h) = \frac{4\text{-}1}{\sum_{i=1}^{I=8} c_{hi}^2} \sum_{i=1}^{I=8} c_{hi}^2 S_i^2 \text{ , for h = 1,..., (8 -1)}$$

with (4 -1) degrees of freedom and

$$MSR(Y_h) = \frac{SSR(Y_h)}{4-1} = \frac{1}{4\text{-}1}\left[\frac{(4\text{-}1)}{\sum_{i=1}^{I=8} c_{hi}^2} \sum_{i=1}^{I=8} c_{hi}^2 S_i^2\right], \text{ with } n_h$$

degrees of freedom obtained through the application of the Satterthwaite (1941, 1946) formula

$$n_h = \frac{[\sum_{i=1}^{I=8} c_{hi}^2 S_i^2]^2}{\sum_{i=1}^{I=8} \frac{[c_{hi}^2 S_i^2]^2}{4\text{-}1}}$$

Table 3 - Application of orthogonal contrasts to the sorghum experiment.

Treatment (i)	$\bar{y}_i$	Orthogonal contrast coefficients, c_{hi}, h = 1,..., 7.						
		c_{1i}	c_{2i}	c_{3i}	c_{4i}	c_{5i}	c_{6i}	c_{7i}
1	0.67	3	1	0	0	0	0	0
2	0.45	3	-1	0	0	0	0	0
3	45.28	-1	0	1	-4	18	0	0
4	56.73	-1	0	1	-1	-27	0	0
5	59.25	-1	0	1	5	9	0	0
6	42.48	-1	0	-1	0	0	-4	18
7	48.96	-1	0	-1	0	0	-1	-27
8	54.13	-1	0	-1	0	0	5	9
$\hat{Y}_h$		-303.3800	0.1900	15.6900	19.4500	-8.7400	17.2600	-3.3500
$SS(Y_h)$		15,340.4097	0.0741	164.1697	324.3982	118.7572	255.2015	17.4052
$DF(Y_h)$		1	1	1	1	1	1	1

Note: $\bar{y}_i = \frac{y_{i.}}{4}$ is the i-esimal treatment mean; $y_{i.} = \sum_{j=1}^{J=4} y_{ij}$, total of the i-esimal treatment.

SSResidual $= \sum_{h=1}^{8-1} SSR(Y_h)$ + SSR(among replications), with 8(4 -1) degrees of freedom. And SSR(among replications) $= \frac{4-1}{8}\sum_{i=1}^{I=8} S_i^2$, with (4 – 1) degrees of freedom and MSR (among replications) $= \frac{1}{4-1}[\frac{4-1}{8}\sum_{i=1}^{I=8} S_i^2] = \frac{\sum_{i=1}^{I=8} S_i^2}{8}$.

Thus, the hypothesis $H_0: Y_h = 0$ vs. $H_a: Y_h \neq 0$, for h=1, ..., (8 - 1), will be tested by the application of the F test and that $F_h = \frac{MS(Y_h)}{MSR(Y_h)} \sim F_{(1,n_h)}$, as observed by Nogueira (1984). Results are shown in Table 5, where the values in [ ], found in DF (degrees of freedom) col-

Table 4 - Analysis of variance with treatment decomposition of seven degrees of freedom decomposition in orthogonal contrasts without specific residue application.

Sources of Variation	DF	SS	MS	F test
Controls vs P-treatments (Y1)	1	15,340.4097	15,340.4097	448.90**
Between controls (Y2)	1	0.0741	0.0741	0.00
Located-P vs Incorporated-P (Y3)	1	164.1697	164.1697	4.80*
Located-P (Y4 + Y5)	2	443.1554	221.5778	6.48**
Located- P linear effect (Y4)	1	324.3982	324.3982	9.49**
Located-P quadratic effect (Y5)	1	118.7572	118.7572	3.48
Incorporated-P (Y6 + Y7)	2	272.6066	136.3033	3.99*
Incorporated-P linear effect (Y6)	1	255.2015	255.2015	7.47*
Incorporated-P quadratic effect (Y7)	1	17.4052	17.4052	0.51
(Treatments)	(7)	(16,220.4155)	2,317.2022	67.81
Residual	24	820.1616	34.1734	
Total	31	17,040.5771		

Note: *significance by (0.01< *p*-Value ≤ 0.05); **significance by (*p*-Value ≤ 0.01).

Table 5 - Analysis of variance with specific residue application.

Sources of Variation	DF	SS	MS	F test
Controls vs P-treatments (Y1)	1	15,340.4097	15,340.4097	1,344.95**
Between controls (Y2)	1	0.0741	0.0741	3.34
Located-P vs Incorporated-P (Y3)	1	164.1697	164.1697	3.60
Located-P (Y4 + Y5)	2	443.1554	221.5778	8.54*
Located- P linear effect (Y4)	1	324.3982	324.3982	21.46*
Located-P quadratic effect (Y5)	1	118.7572	118.7572	3.23
Incorporated-P (Y6 + Y7)	2	272.6066	136.3033	2.09
Incorpor.-P linear effect (Y6)	1	255.2015	255.2015	2.37
Incorpor.-P quadratic effect (Y7)	1	17.4052	17.4052	0.76
(Treatments)	(7)	(16,220.4155)	2,317.2022	67.81
$R(Y_1)$	3 [7]	34.2177	11.4059	
$R(Y_2)$	3 [6]	0.0664	0.0222	
$R(Y_3)$	3 [7]	136.6710	45.5571	
R(Located-P)	6[5]	155.5971	25.9321	
$R(Y_4)$	3[7]	45.3481	15.1160	
$R(Y_5)$	3[4]	110.2490	36.7496	
R(Incorporated-P)	6[8]	391.0889	65.1814	
$R(Y_6)$	3[3]	322.5410	107.5140	
$R(Y_7)$	3[6]	68.5479	22.8493	
R(among replications)	3	102.520	34.1734	
(Residual)	(24)	(820.1616)	34.1734	
Total	31	17,040.5771		

Note: *significance by (0.01 < *p*-Value ≤ 0.05); **significance by (*p*-Value ≤ 0.01).

umn refer to the effective degrees of freedom - n_h , obtained by the Satterthwaite formula and applied in the F test.

It was observed that

MSResidual = MSR (among replications) = $\frac{1}{7}\sum_{h=1}^{(8-1)}$ MSR (Y_h) = 34.1734.

The F test values presented in Table 4 were obtained having MSResidual as denominator, with 24 degrees of freedom. The results presented in Tables 4 and 5 are different as well as some of the conclusions. This fact is important due to the presence of heteroscedasticity, because in Table 4, the MSResidual corresponds to the MSR(Y_h) arithmetic mean; and in Table 5, the values obtained for MSR(Y_h) were different. In the presence of homoscedasticity the values obtained for MSR(Y_h) are very close to the ones obtained for MSResidual. The use of the specific residue procedure showed to be an interesting alternative to be applied when irregular heteroscedasticity is present, providing trustworthy results.

Randomized complete block design

In order to illustrate the specific residue procedure application on data analyses of a randomized complete block design experiment, the following experimental data were considered: yields of eight potato varieties (t ha^{-1}) distributed in five blocks (Table 6).

The Bartlett test was applied to verify the variance homogeneity hypothesis, which was rejected, thus evidencing the presence of variance heterogeneity among treatments. Due to this fact and considering that experimental errors followed a normal distribution, the specific residue procedure was applied as an alternative for this data analysis. The initial analysis of variance is shown in Table 7.

Seven degrees of freedom and the variety sum of squares were decomposed in a group of orthogonal contrasts according to the high and low productivity criterion. Then, the potato varieties were divided into two groups and the high productivity potato group consisted of the varieties: (3) B1-52, (4) Huinkul, (5) B116-51; (6) B72-53 A and (7) S. Rafaela; and the low productivity potato group consisted of the varieties: (1) Kennebec, (2) B25-50E and (8) Buena Vista. Thus, the group of orthogonal contrasts built up according to the productivity criterion was: Y_1: High productivity varieties (varieties 3, 4, 5, 6 and 7) versus Low productivity varieties (varieties 1, 2 and 8); Y_2: Variety 7 versus varieties 3, 4, 5 and 6; Y_3: Varieties 4 and 6 versus varieties 3 and 5; Y_4: Between varieties 4 and 6; Y_5: Between varieties 3 and 5; Y_6: Variety 1 versus varieties 2 and 8; Y_7: Between varieties 2 and 8.

The orthogonal contrasts Y_2, Y_3, Y_4 and Y_5 provided the high productivity variety effect with four degrees of freedom, and the contrasts Y_6 and Y_7 provided the low productivity variety effect with two degrees of freedom. The coefficients of the applied contrasts, the contrast estimates and the sum of squares obtained are shown in Table 8.

Twenty eight degrees of freedom and the residual sum of squares were decomposed according to the Y(h) components, resulting the Y(h) specific residues given by:

$$SSR(Y_h) = \frac{1}{\sum_{i=1}^{I=8} c_{hi}^2}\left[\sum_{j=1}^{J=5}\hat{Y}_{hj}^2 - \frac{\hat{Y}_h^2}{5}\right],$$

with (5-1) = 4 degrees of freedom and $\hat{Y}_{hj}$ is the Y_{hj} contrast estimate, corresponding to the Y_h contrast application in the block j, for j = 1, ..., J = 5 ,

$$\hat{Y}_{hj} = \sum_{i=1}^{I=8} c_{hi}\, y_{ij}, \text{ such that } \sum_{i=1}^{I} c_{hi} = 0 \text{ and}$$

$$\hat{Y}_h = \sum_{j=1}^{J=5}\hat{Y}_{hj} = \sum_{i=1}^{I=8} c_{hi}\, y_{i.},$$

where y_{ij} is the observed value related to variety i in block j; $\hat{Y}_h$ is the h-esimal contrast estimate, for h = 1, ..., (8-1)=7 and $y_{i.} = \sum_{j=1}^{J=5} y_{ij}$. The values referred to y_{ij} and the Y_h coefficients for the $\hat{Y}_{hj}$ calculus are presented in Table 9.

The results referred to $\hat{Y}_{hj}$ and $\hat{Y}_h$ estimates and SSR(Y_h) values are presented in Table 10, as follows:

It was observed that $\sum_{h=1}^{8-1}$ SSR(Y_h) = SQResidual = 348.324, with (8-1)(5-1)=28 degrees of freedom.

Also that MSR(Y_h) = $\frac{1}{5-1}$ SSR(Y_h), with (5-1) = 4 degrees of freedom.

Thus, the hypotheses H_0:Y_h = 0 vs. H_a : $Y_h \neq 0$, for

Table 6 - Potato variety yields (t ha^{-1}).

Variety (i)	Blocks					Variance
	1	2	3	4	5	S_i^2
(1) Kennebec	11.750	7.950	10.700	12.050	12.300	31.0798
(2) B25-50 E	12.075	16.250	16.500	8.950	14.575	2.1561
(3) B1-52	16.150	30.025	22.275	10.975	23.500	10.5313
(4) Huinkul	20.550	30.125	25.050	15.600	23.225	2.2617
(5) B116-51	15.275	30.575	22.500	13.075	23.200	9.3295
(6) B72-53 A	17.350	27.800	22.800	12.975	19.925	3.1558
(7) S. Rafaela	17.125	29.400	25.450	12.200	23.225	7.5243
(8) Buena Vista	11.925	10.650	12.425	12.400	13.000	21.0424

Table 7 - Analysis of variance of potato yield.

Sources of Variation	DF	SS	MS
Blocks	4	542.2406	
Varieties	7	793.9257	113.4180
Residual	28	348.3238	12.4401
Total	39	1,684.6119	

h = 1, ..., (8 - 1), were then tested by the application of the F test, $F_h = \frac{MS(Y_h)}{MSR(Y_h)} \sim F_{(1,(5-1))}$.

The analysis of variance obtained with the specific residue procedure application is presented in Table 11. Significant F test values for Y_1 and Y_4 contrasts were observed, evidencing they differ from zero.

Table 8 - Coefficients of contrasts, estimates and contrast sum of squares for the potato yield experiment.

Treatment (i)	Total for 5 blocks	Orthogonal contrast coefficients, c_{hi}, h = 1,..., 7.						
		c_{1i}	c_{2i}	c_{3i}	c_{4i}	c_{5i}	c_{6i}	c_{7i}
(1) Kennebec	54.750	-5	0	0	0	0	2	0
(2) B25-50 E	68.350	-5	0	0	0	0	-1	1
(3) B1-52	102.350	3	-1	-1	0	1	0	0
(4) Huinkul	114.550	3	-1	1	1	0	0	0
(5) B116-51	104.625	3	-1	-1	0	-1	0	0
(6) B72-53 A	100.850	3	-1	1	-1	0	0	0
(7) S. Rafaela	107.400	3	4	0	0	0	0	0
(8) Buena Vista	60.700	-5	0	0	0	0	-1	-1
$\hat{Y}_h$		672.050	6.650	7.850	13.700	-1.700	-19.550	7.650
$SS(Y_h)$		752.752	0.442	3.081	18.769	0.289	12.740	5.852
$DF(Y_h)$		1	1	1	1	1	1	1

Table 9 - Observed values (y_{ij}) and Y_h-coefficients for $\hat{Y}_{hj}$ estimation.

Variety (i)	Observed values, y_{ij}					Y_h coefficients						
	Blocks (j)											
	1	2	3	4	5	c_{1i}	c_{2i}	c_{3i}	c_{4i}	c_{5i}	c_{6i}	c_{7i}
(1) Kennebec	11.750	7.950	10.700	12.050	12.300	-5	0	0	0	0	2	0
(2) B25-50 E	12.075	16.250	16.500	8.950	14.575	-5	0	0	0	0	-1	1
(3) B1-52	16.150	30.025	22.275	10.975	23.500	3	-1	-1	0	1	0	0
(4) Huinkul	20.550	30.125	25.050	15.600	23.225	3	-1	1	1	0	0	0
(5) B116-51	15.275	30.575	22.500	13.075	23.200	3	-1	-1	0	-1	0	0
(6) B72-53 A	17.350	27.800	22.800	12.975	19.925	3	-1	1	-1	0	0	0
(7) S. Rafaela	17.125	29.400	25.450	12.200	23.225	3	4	0	0	0	0	0
(8) Buena Vista	11.925	10.650	12.425	12.400	13.000	-5	0	0	0	0	-1	-1

Table 10 - Estimation of $\hat{Y}_{hj}$ and $\hat{Y}_h$ and $SSR(Y_h)$ values.

(h) $\hat{Y}_h$	$\hat{Y}_{hj} = \sum_{i=1}^{I=8} c_{hi}\, y_{ij}$ Blocks (j)					$SSR(Y_h)$
	1	2	3	4	5	
(1) 672.05	80.600	269.525	156.1	27.475	138.35	275.605
(2) 6.65	-0.825	-0.925	9.175	-3.825	3.050	5.04027
(3) 7.85	6.475	-2.675	3.075	4.525	-3.550	19.8226
(4) 13.70	3.200	2.325	2.250	2.625	3.300	0.47538
(5) -1.70	0.875	-0.550	-0.225	-2.100	0.300	2.52037
(6) -19.55	-0.500	-1.100	-7.525	2.750	-3.275	19.9539
(7) 7.65	0.150	5.600	4.075	-3.450	1.275	24.9059
SSResidual						348.324

Table 11 - Analyses of variance with specific residue procedure application.

Sources of Variation	DF	SS	MS	F Test	$\hat{Y}_h$
Blocks	4	542.2406			
Varieties	7	793.9257	113.4180		
Y_1	1	752.7520		10.92	672.05**
Y_2	1	0.4420		0.35	6.65
Y_3	1	3.0810		0.62	7.85
Y_4	1	18.7690		157.93	13.7**
Y_5	1	0.2890		0.46	-1.7
Y_6	1	12.7400		2.55	-19.55
Y_7	1	5.8520		0.94	7.65
Residual	28	348.3240	12.4401		
$R(Y_1)$	4	275.6050	68.9014		
$R(Y_2)$	4	5.0403	1.26007		
$R(Y_3)$	4	19.8226	4.95566		
$R(Y_4)$	4	0.4754	0.11884		
$R(Y_5)$	4	2.5204	0.63009		
$R(Y_6)$	4	19.9539	4.98847		
$R(Y_7)$	4	24.9059	6.22647		
Total	39	1,684.6119			

Note: *significance by ($0.01 < p$-Value ≤ 0.05); **significance by (p-Value ≤ 0.01).

Table 12 - Analyses of variance without specific residue procedure application.

Sources of Variation	DF	SS	MS	F Test	$\hat{Y}_h$
Blocks	4	542.2406			
Varieties	7	793.9257	113.4180		
Y_1	1	752.7520		60.51	672.05 **
Y_2	1	0.4422		0.04	6.65
Y_3	1	3.0811		0.25	7.85
Y_4	1	18.7690		1.51	13.7
Y_5	1	0.2890		0.02	-1.7
Y_6	1	12.7401		1.02	-19.55
Y_7	1	5.8522		0.47	7.65
Residual	28	348.3238	12.4401		
Total	39	1,684.6119			

Note: *significance by ($0.01 < p$-Value ≤ 0.05); **significance by (p-Value ≤ 0.01).

The analysis of variance without the specific residue procedure was also obtained (Table 12) in order to be compared to the previous analysis (Table 11). Significant F value was obtained for the Y_1 contrast when calculated with MSResidual as denominator, with 28 degrees of freedom, evidencing that it significantly differed from zero. When the specific residue procedure was applied (Table 11), significant F values were obtained for the Y_1 and Y_4 contrasts.

Conclusion

The use of the specific residue procedure is a valid and efficient alternative when heteroscedasticity is present, because it validates the applied tests and also allows a better understanding of the residual mean square nature. The MSResidual corresponds to the $MSR(Y_h)$ arithmetic mean, although the values obtained for $MSR(Y_h)$ can be different. In the presence of homoscedasticity the values obtained for $MSR(Y_h)$ are very close to those obtained for MSResidual.

References

Box, G.E.P; Miller, M.E. 1958. A note on the generation of random normal deviates. Annals of Mathematics Statistics 29: 610-611.

Cochran, W.G. 1947. Some consequences when the assumptions for the analysis of variance are not satisfied. Biometrics 3: 22–38.

Cochran, W.G.; Cox, G.M. 1957. Experimental Designs. 2ed. John Wiley, New York, NY, USA.

Cochran, W.G.; Cox, G.M. 1971. Diseños Experimentales. Editorial Trillas, Ciudad de México, México.

Godoi, C.R.M. 1978. Um algoritmo eficiente para simulação de vetores com distribuição multinormal. Ciência e Cultura 20: 701-705.

Ferreira, L.E.P. 1978. A decomposição do resíduo em casos de heterocedasticidade nas análises de variância de ensaios em blocos casualizados. MSc Dissertation. Universidade de São Paulo, Piracicaba, SP, Brazil. (in Portuguese with summary in English).

Nogueira, I.R. 1978. Método geral para obtenção de tabelas de polinômios ortogonais. Revista da Agricultura 53: 269-279.

Nogueira, M.C.S. 1984. Resíduo específico para contraste de tratamentos no delineamento inteiramente casualizado. Dr. Thesis. Universidade de São Paulo, Piracicaba, SP, Brazil (in Portuguese with summary in English).

Nogueira, M.C.S.; Campos, H. 1985. Resíduo específico para contraste de tratamentos no delineamento inteiramente casualizado. Anais do Simpósio de Estatística Aplicada à Experimentação Agronômica 1. Fundação Cargill, Piracicaba, SP, Brazil.

Nogueira, M.C.S. 2004. Orthogonal contrasts: Definitions and concepts. Scientia Agricola 61: 118-124.

Nogueira, M.C.S. 2007. Experimentação agronômica I. Conceitos, planejamento e análise de dados. Editora MCSNogueira, Piracicaba, SP, Brazil.

Oliveira, W.; Nogueira, M.C.S. 2007. Aplicação do resíduo específico na análise de grupos de experimentos. Bragantia 66: 737-744.

Satterthwaite, F.E. 1941. Synthesis of variance. Psychometrika 6: 309-316.

Satterthwaite, F.E. 1946. An approximate distribution of estimates of variance components. Biometrics Bulletin 2: 110-114.

Steel, R.G.D.; Torrie, J.H. 1981. Principles and Procedures of Statistics. 2ed. McGraw-Hill, New York, NY, USA.

Received June 15, 2007
Accepted August 21, 2009