An Analysis of Multiple Variance for The Multivariate Split-Block Design

Khawla Abdul Razzaq Swadi
Department Of Mathematics, College of Science, University of Basra, Basra Iraq

E-mail: khawalah.swadi@uobasrah.edu.iq

Abstract. This paper studies an important type of multi-factor experiments designs called (Split-Blocks experiment Design). And we’re going to use the orthogonal matrix to convert the observation matrix. Also, we study the analysis of variance for the Multivariate split-Blocks design. And we discuss the test statistics for Lawley-Hotelling test, Roy's Union-Intersection test and Walks' Lambda test.

Keywords. Split-Blocks experiment Design, MNOVA, Lawley-Hotelling Test, Roy's Union-Intersection Test, the multivariate-Wishart distribution, Walks' Lambda test.

1. Introduction

An analysis of multiple variance MANOVA is an extension of an ANOVA for more than one dependent variable. In addition to MANOVA, is a method of measuring the differences of two dependent variables, or more, depending on the classification variables that are considered independent variables [3] [5] [10] [11]. The analysis of multiple variance is used when there are several dependent variables that are related, and the researcher wishes to use one overall statistical test on this set of variables, instead of using several tests separately. The second use, which is to some extent achieves the most important purpose of using the MANOVA test, is to examine how the independent variables affect a group of dependent variables at one time and to measure several dependent variables on each experimental unit instead of one variable [9].

The assumptions of the MANOVA multiple analysis of variance:

1-The observations are statistically independent.
2- That the dependent variables follow the multivariate normal distribution.
3-The homogeneity of the variance matrices and the covariance's of all the groups.

This paper focuses on multivariate models in the analysis of multiple variance from repeated measurement designs, especially hypothesis testing. In such experiments, important variables are observed at several time points.
The idea of designing and expanding split-plot depends on the number of factors involved in the experiment, what factor or factors to focus on its accuracy, what is the least important factor, the type of design used, and the type of field in which the experiment will be carried out (laboratory, agricultural, medical)[1].

The design of split-Blocks is one of the important designs and it is also called strip plots design[2][4][6], as it deals with two or more factors and at several levels, one of which takes main -plot (main factor) and is the factor A and the other take a sub-plot (a secondary factor) and be the factor P and the goal of this design is (overlap) The interaction between the main and secondary factors and it is also called strip plots. Such experiments divide the analysis of variance into three parts, each of which has a random error effect of the interaction effect between A and P, and ((Error(p)) concerning the effect of interaction or interference between factors A and P, and often ((Error(p)) is the most accurate interference test [7][8].

2. Mathematical model
The mathematical model for the designs of split-Blocks is as follows

\[ y_{hij} = \mu + \gamma_h + \alpha_i + \tau_{hi} + \beta_j + \delta_{hj} + (\alpha\beta)_{ij} + \epsilon_{hij} \]  

Where \( h = 1,...,r \); \( i = 1,...,a \); \( j = 1,...,p \);
\( y_{hij} = [y_{h1j},...,y_{hij},...,y_{hijn}] \) represents the result of the response (h) at the sub-plot (j) within the main-plot (i), \( \mu = [\mu_1,...,\mu_h] \) represent the overall average effect of the experiment results, \( \gamma_h = [\gamma_{h1},...,\gamma_{hn}] \) the sector effect, which is an independent variable, and has a normal distribution, \( \alpha_i = [\alpha_{i1},...,\alpha_{in}] \) represent the effect of the factor A at the main-plot (i), \( \tau_{hi} = [\tau_{h1i},...,\tau_{hin}] \) the effect of random error of the factor A located at the main-plot (i) is independent and has a normal distribution, \( \beta_j = [\beta_{j1},...,\beta_{jn}] \) the effect of the factor P at the sub-plot (j), \( \delta_{hj} = [\delta_{h1j},...,\delta_{hjn}] \) the effect of random error of the factor P located at the sub-plot (j) is independent and has normal distributed, \( (\alpha\beta)_{ij} = [(\alpha\beta)_{i1j},...,\alpha\beta_{ijn}] \) it represents the effect of the interaction between factors A and P, and \( \epsilon_{hij} = [\epsilon_{h1j},...,\epsilon_{hijn}] \) the random error effect of the interaction effect between A and P is independent and normally distributed.

Assume the conditions
\[
\begin{align*}
\sum_{i=1}^{a} \alpha_i &= 0, \\
\sum_{j=1}^{p} \beta_j &= 0, \\
\sum_{i=1}^{a} (\alpha\beta)_{ij} &= 0 \quad \text{for each } j = 1,...,p \\
\sum_{j=1}^{p} (\alpha\beta)_{ij} &= 0 \quad \text{for each } i = 1,...,a,
\end{align*}
\]

(2)

We impose that the \( \epsilon_{hij} \)'s, \( \tau_{hi} \)'s, \( \delta_{hj} \)'s and \( \gamma_h \)'s are independent with
\( \varepsilon_{hij} = [\varepsilon_{h_{i1}}, \cdots, \varepsilon_{h_{in}}] \) i.i.d. \( \sim N_n(0, \Sigma_\varepsilon) \),
\( \tau_{hi} = [\tau_{h_{i1}}, \cdots, \tau_{h_{in}}] \) i.i.d. \( \sim N_n(0, \Sigma_\tau) \),
\( \delta_{hj} = [\delta_{h_{j1}}, \cdots, \delta_{h_{jn}}] \) i.i.d. \( \sim N_n(0, \Sigma_\delta) \),
\( \gamma_{h} = [\gamma_{h_{1}}, \cdots, \gamma_{h_{n}}] \) i.i.d. \( \sim N_n(0, \Sigma_\gamma) \).

Where \( \Sigma_\varepsilon, \Sigma_\tau, \Sigma_\delta \), and \( \Sigma_\gamma \) are \( n \times n \) positive definite matrices.

Let \( Y_{hi} = [Y_{h_{i1}}, Y_{h_{i2}}, \cdots, Y_{h_{ip}}] \).

(4)

The variance-covariance matrix of \( \tilde{Y}_{hi} \) is denoted as \( \Sigma_i \),
\[ \Sigma = I_p \otimes \Sigma_\varepsilon + I_p \otimes \Sigma_\tau + I_p \otimes \Sigma_\delta + I_p \otimes \Sigma_\gamma = \begin{bmatrix} \Sigma_\varepsilon & \Sigma_\gamma & \cdots & \Sigma_\gamma \\ \Sigma_\gamma & \Sigma_\varepsilon & \cdots & \Sigma_\gamma \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_\gamma & \Sigma_\gamma & \cdots & \Sigma_\varepsilon \end{bmatrix}. \]

(5)

Where \( \Sigma_i = \Sigma_\varepsilon + \Sigma_\tau + \Sigma_\delta + \Sigma_\gamma \). \( I_p \) is \( (p \times p) \) identity matrix and \( I_p \) \( (p \times p) \) matrix of One's. \( \varepsilon_{hi} = [\varepsilon_{h_{i1}}, \cdots, \varepsilon_{h_{in}}] \) i.i.d. \( \sim N_{p \times n}(0, I_p \otimes \Sigma_\varepsilon) \),
\( \tau_{hi} = [\tau_{h_{i1}}, \cdots, \tau_{h_{in}}] \) i.i.d. \( \sim N_{p \times n}(0, I_p \otimes \Sigma_\tau) \),
\( \delta_{hi} = [\delta_{h_{j1}}, \cdots, \delta_{h_{jn}}] \) i.i.d. \( \sim N_{p \times n}(0, I_p \otimes \Sigma_\delta) \).

(6)

### 3. Multivariate Analysis of Variance (MANOVA) for the Multivariate split-Block Model:

We using an orthogonal matrix to convert the observations \( Y_{h_{ij}} \) for \( i = 1, \cdots, a \), \( j = 1, \cdots, p \).

Let \( Y'_{hi} = Y_{hi} \left( \frac{1}{\sqrt{p}} I_p V \right) \).

(7)

Where \( \left( \frac{1}{\sqrt{p}} I_p V \right) \) be any \( (p \times p) \) orthogonal matrix. \( I_p \) \( (p \times 1) \) vector of one's, \( V \) is \( p \times (p - 1) \) matrix, and \( V' I_p = 0 \), \( V' V = I_{p-1} \).

So \( \text{Cov}(\tilde{Y}_{hi}) = \text{Cov}\left(\left( \left( \frac{1}{\sqrt{p}} I_p V \right) \otimes I_n \right) \tilde{Y}_{hi} \right) = I_p \otimes \Sigma_\varepsilon + I_p \otimes \Sigma_\tau + I_p \otimes \Sigma_\delta + \left( \frac{1}{\sqrt{p}} I_p V \right) \otimes I_n \Sigma_\gamma I_n \)
\[ = \begin{bmatrix} \Sigma_\varepsilon + p\Sigma_\tau + a\Sigma_\delta + ap\Sigma_\gamma & 0 & \cdots & 0 \\ 0 & \Sigma_\varepsilon + p\Sigma_\tau + a\Sigma_\delta & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Sigma_\varepsilon + p\Sigma_\tau + a\Sigma_\delta \end{bmatrix}. \]

Now, we study the MANOVA for the effects of Multivariate split-Block model (1).

\( Y_{hi} = h_i \left( \frac{1}{\sqrt{p}} I_p V \right) \)

\( Y_{h_{i1}} = \left[ \frac{1}{\sqrt{p}} \sum_{j=1}^{p} Y_{h_{ij1}} \right] \)

\( Y_{h_{i1}} = \left[ \frac{1}{\sqrt{p}} \sum_{j=1}^{p} Y_{h_{ij1}} \right] \), from (1), we obtain

\( Y_{h_{i1}} = \frac{1}{\sqrt{p}} \sum_{j=1}^{p} (\mu + \gamma_{h} + \alpha_{i} + \tau_{h_{i}} + \beta_{j} + \delta_{h_{ij}} + (\alpha\beta)_{ij} + \varepsilon_{h_{ij}}) \)
\[ Y_{ij} = \sqrt{\mu} + \sqrt{\nu} + \sqrt{\rho \alpha_i} + \sqrt{\nu \pi} + \frac{1}{\sqrt{\rho}} \sum_{i=1}^{p} \delta_{hi} + \frac{1}{\sqrt{\rho}} \sum_{i=1}^{p} \varepsilon_{hi} \]  

So, the hypothesis of the same treatment effects is:

\[ H_{01} : \sqrt{\rho \alpha_1} = \ldots = \sqrt{\rho \alpha_a} = 0, \text{ and are equivalent to the hypothesis} \]

\[ H_{02} : \sqrt{\mu} + \sqrt{\nu \alpha} = \ldots = \sqrt{\nu \alpha} = 0. \]

The MANOVA based on the set of convert observations above. This leads to the following shape for the sum squares terms: \( SS_A, SS_R \) and \( SSE_A \), where

\[ SS_A = r \sum_{i=1}^{a} (Y_{i1}^2 - \bar{Y}_{i1})' (Y_{i1}^2 - \bar{Y}_{i1})' \rightarrow SS_A \sim W_n(a - 1, \Sigma + p \Sigma + \alpha \Sigma + a \pi \Sigma) \]

\[ SSE_A = \sum_{h=1}^{r} \sum_{i=1}^{p} (Y_{hi}^2 - \bar{Y}_{hi}^2 + \bar{Y}_{hi}^2 - \bar{Y}_{hi}^2)' (Y_{hi}^2 - \bar{Y}_{hi}^2)' \rightarrow SSE_A \sim W_n((a - 1)(r - 1), \Sigma + p \Sigma + \alpha \Sigma + a \pi \Sigma) \]

\[ SS_R = a \sum_{h=1}^{r} \sum_{i=1}^{p} \bar{Y}_{hi}^2 \rightarrow SS_R \sim W_n(r - 1, \Sigma + p \Sigma + \alpha \Sigma + a \pi \Sigma) \]

Where \[ \bar{Y}_{hi} = \frac{\sum_{i=1}^{p} Y_{hi}^2}{r}, \bar{Y}_{j} = \frac{\sum_{i=1}^{p} Y_{hi}^2}{a} \]

And \( W_n \) represent the multivariate-Wishart distribution.

The MANOVA based on the set of convert observations above. This leads to the following forms for the sum squares terms:

\[ SSE_{AP} = \sum_{i=1}^{a} \sum_{j=1}^{p} (Y_{ij}^2 - \bar{Y}_{ij}^2)' (Y_{ij}^2 - \bar{Y}_{ij}^2)' \rightarrow SSE_{AP} \sim W_n((a - 1)(p - 1), \Sigma + \alpha \Sigma) \]

\[ SS_{AP} = r \sum_{i=1}^{a} \sum_{j=1}^{p} (Y_{ij}^2 - \bar{Y}_{ij}^2)' (Y_{ij}^2 - \bar{Y}_{ij}^2)' \rightarrow SS_{AP} \sim W_n((a - 1)(r - 1), \Sigma + \alpha \Sigma) \]

Now we test the following hypotheses

\[ H_01 : \beta_{p1}^2 = \ldots = \beta_{p1}^2 = 0, \quad H_02 : (\alpha \beta)_{i2} = \ldots = (\alpha \beta)_{ip} = 0. \]

The MANOVA based on the set of convert observations above. This leads to the following forms for the sum squares terms:

\[ SSE_{AP} = \sum_{i=1}^{a} \sum_{j=1}^{p} (Y_{ij}^2 - \bar{Y}_{ij}^2)' (Y_{ij}^2 - \bar{Y}_{ij}^2)' \rightarrow SSE_{AP} \sim W_n((a - 1)(p - 1), \Sigma + \alpha \Sigma) \]

\[ SS_{AP} = r \sum_{i=1}^{a} \sum_{j=1}^{p} (Y_{ij}^2 - \bar{Y}_{ij}^2)' (Y_{ij}^2 - \bar{Y}_{ij}^2)' \rightarrow SS_{AP} \sim W_n((a - 1)(r - 1), \Sigma + \alpha \Sigma) \]

Where \[ \bar{Y}_{ij} = \frac{Y_{hi}^2}{r}, \quad i = 1, \ldots, a, \quad j = 2, \ldots, p, \quad \bar{Y}_{ij} = \frac{Y_{hi}^2}{a}, \quad j = 2, \ldots, p, \quad \bar{Y}_{hi} = \frac{Y_{hi}^2}{p}, \quad h = 1, \ldots, r, \quad i = 1, \ldots, a. \]

Let \((E)\) the error sum of squares matrix and \((H)\) the hypothesis sum of squares matrix. Then the Multivariate split-Block ANOVA with Lawley-Hotelling test are given in the following table.
Table (1): The Multivariate split-Block MANOVA with Lawley-Hotelling test.

| Source of Variation | d.F | S.S | Lawley-Hotelling Criterion |
|---------------------|-----|-----|---------------------------|
| Replicates          | r-1 | $SS_R = a \sum_{h=1}^{r} (Y_{h1}^* - \bar{Y}_1^*)(Y_{h1}^* - \bar{Y}_1^*)'$ | trace $(SS_A \quad SSE_A^{1/2})$ |
| Factor (A)          | a-1 | $SS_A = r \sum_{i=1}^{a} (Y_{i1}^* - \bar{Y}_1^*)(Y_{i1}^* - \bar{Y}_1^*)'$ | trace $(SS_P \quad SSE_P^{1/2})$ |
| Error (A)           | (r-1)(a-1) | $SSE_A = \sum_{h=1}^{r} \sum_{i=1}^{a} (Y_{h1}^* - \bar{Y}_1^*)(Y_{h1}^* - \bar{Y}_1^*)'$ | trace $(SS_{AP} \quad SSE_{AP}^{1/2})$ |
| Factor (P)          | (p-1) | $SS_P = r a \sum_{j=1}^{p} (Y_{j}^*)(Y_{j}^*)'$ | trace $(SS_P \quad SSE_P^{1/2})$ |
| Error (P)           | (r-1)(p-1) | $SSE_P = a \sum_{h=1}^{r} \sum_{j=1}^{p} (Y_{h1}^* - \bar{Y}_1^*)(Y_{h1}^* - \bar{Y}_1^*)'$ | trace $(SS_{AP} \quad SSE_{AP}^{1/2})$ |
| $A \times P$        | (a-1)(p-1) | $SS_{AP} = r \sum_{i=1}^{a} \sum_{j=1}^{p} (Y_{i1}^* - \bar{Y}_1^*)(Y_{i1}^* - \bar{Y}_1^*)'$ | trace $(SS_{AP} \quad SSE_{AP}^{1/2})$ |
| Error (AP)          | (r-1)(a-1)(p-1) | $SSE_{AP} = \sum_{h=1}^{r} \sum_{i=1}^{a} \sum_{j=1}^{p} (Y_{h1}^* - \bar{Y}_1^*)(Y_{h1}^* - \bar{Y}_1^*)'$ | trace $(SS_{AP} \quad SSE_{AP}^{1/2})$ |
| Total               | rap-1 | $SS_T$ | trace $(SS_{AP} \quad SSE_{AP}^{1/2})$ |

The Multivariate split-Block ANOVA with Roy's Union-Intersection test are given in the following table.

Table (2): The Multivariate split-Block MANOVA with Roy's Union-Intersection Test.

| Source of Variation | d.F | S.S | Roy's Union-Intersection Criterion |
|---------------------|-----|-----|-----------------------------------|
| Replicates          | r-1 | $SS_R = a \sum_{h=1}^{r} (Y_{h1}^* - \bar{Y}_1^*)(Y_{h1}^* - \bar{Y}_1^*)'$ | Largest characteristics root of $(SS_A \quad SSE_A^{1/2})$ |
| Factor (A)          | a-1 | $SS_A = r \sum_{i=1}^{a} (Y_{i1}^* - \bar{Y}_1^*)(Y_{i1}^* - \bar{Y}_1^*)'$ | Largest characteristics root of $(SS_P \quad SSE_P^{1/2})$ |
| Error (A)           | (r-1)(a-1) | $SSE_A = \sum_{h=1}^{r} \sum_{i=1}^{a} (Y_{h1}^* - \bar{Y}_1^*)(Y_{h1}^* - \bar{Y}_1^*)'$ | Largest characteristics root of $(SS_{AP} \quad SSE_{AP}^{1/2})$ |
| Factor (P)          | (p-1) | $SS_P = r a \sum_{j=1}^{p} (Y_{j}^*)(Y_{j}^*)'$ | Largest characteristics root of $(SS_P \quad SSE_P^{1/2})$ |
| Error (P)           | (r-1)(p-1) | $SSE_P = a \sum_{h=1}^{r} \sum_{j=1}^{p} (Y_{h1}^* - \bar{Y}_1^*)(Y_{h1}^* - \bar{Y}_1^*)'$ | Largest characteristics root of $(SS_{AP} \quad SSE_{AP}^{1/2})$ |
| $A \times P$        | (a-1)(p-1) | $SS_{AP} = r \sum_{i=1}^{a} \sum_{j=1}^{p} (Y_{i1}^* - \bar{Y}_1^*)(Y_{i1}^* - \bar{Y}_1^*)'$ | Largest characteristics root of $(SS_{AP} \quad SSE_{AP}^{1/2})$ |
| Error (AP)          | (r-1)(a-1)(p-1) | $SSE_{AP} = \sum_{h=1}^{r} \sum_{i=1}^{a} \sum_{j=1}^{p} (Y_{h1}^* - \bar{Y}_1^*)(Y_{h1}^* - \bar{Y}_1^*)'$ | Largest characteristics root of $(SS_{AP} \quad SSE_{AP}^{1/2})$ |
| Total               | rap-1 | $SS_T$ | Largest characteristics root of $(SS_{AP} \quad SSE_{AP}^{1/2})$ |

And the Multivariate split-Block ANOVA with Walks Lambda test are given in the following table.
Table (3): The Multivariate split-Block MANOVA with Walks Lambda test.

| Source of Variation | d.F | S.S | Walks Criterion |
|---------------------|-----|-----|-----------------|
| Replicates          | r-1 | SS_R = a \sum_{h=1}^{r}(\bar{Y}_{h1}^{*} - \bar{Y}_{i1}^{*})(\bar{Y}_{h1}^{*} - \bar{Y}_{i1}^{*})^\prime | | |
| Factor (A)          | a-1 | SS_A = r \sum_{i=1}^{a}(\bar{Y}_{i1}^{*} - \bar{Y}_{i1}^{*})^\prime | | |
| Error (A) (r-1)(a-1)|     | SSE_A = \sum_{h=1}^{r}\sum_{i=1}^{a}(\bar{Y}_{h1}^{*} - \bar{Y}_{i1}^{*} + Y_{h1}^{*})(\bar{Y}_{h1}^{*} - \bar{Y}_{i1}^{*} + Y_{i1}^{*})^\prime | | |\frac{|SSE_A|}{SSE_A + SS_A}|
| Factor (P) (p-1)    |     | SS_P = r a \sum_{j=1}^{p}(\bar{Y}_{j1}^{*} - \bar{Y}_{j1}^{*})^\prime | | |\frac{|SSE_P|}{SSE_P + SS_P}|
| Error (P) (r-1)(p-1)|     | SSE_P = a \sum_{h=1}^{r}\sum_{j=1}^{p}(\bar{Y}_{hj}^{*} - \bar{Y}_{j1}^{*})(\bar{Y}_{hj}^{*} - \bar{Y}_{j1}^{*})^\prime | | |\frac{|SSE_AP|}{SSE_AP + SS_AP}|
| A x P (a-1)(p-1)    |     | SSE_AP = r a \sum_{i=1}^{a}\sum_{j=1}^{p}(\bar{Y}_{ij}^{*} - \bar{Y}_{ij}^{*})^\prime | | |\frac{|SSE_AP|}{SSE_AP + SS_AP}|
| Error (AP) (r-1)(a-1)(p-1) | | SSE_AP = \sum_{h=1}^{r}\sum_{i=1}^{a}\sum_{j=1}^{p}(\bar{Y}_{hij}^{*} - \bar{Y}_{ihj}^{*} - \bar{Y}_{hij}^{*} + \bar{Y}_{ihj}^{*} - \bar{Y}_{hij}^{*} + \bar{Y}_{ihj}^{*} + \bar{Y}_{ihj}^{*})^\prime | | |\frac{|SSE_AP|}{SSE_AP + SS_AP}|
| Total               | rap-1 | SSE_T |

References:

[1] Bingham, D.R., Sitter, R.R. Design issues in fractional factorial split-plot experiments. Journal of Quality Technology 2001; 33(1):2–15.
[2] Federer, Walter T. & King, Freedom (2007) "Variation on split plot and split Block Experiment Designs" John Wiley& Sons, Inc. New York.
[3] Gabriel, K. R. (1967). Simultaneous test procedures in multivariate analysis of variance. Biometrika, 56,489-504.
[4] Government of Kenya (2010), "Agriculture Sector Development Strategy (ASDS)", 2010-2020.
[6] Johnson, R.A. and Wichern, D. W. (1988). "Applied Multivariate Statistical Analysis", 2nd Edition. Prentice-Hall International, Inc., London.
[7] Kowalski SM, Cornell JA, Vining GG. Split-plot designs and estimation methods for mixture experiments with process variables. Technometrics 2002; 44(1):72–79.
[8] Montgomery DC. Design and Analysis of Experiments (5th edn). Wiley: New York, 2001.
[9] N.H. Timm (1980). Multivariate Analysis of repeated measurements. In P.R Keishmaiah (Ed.), handbook of statistics, vol.1 (Analysis of Variance), 41 - 87. North-Holland.
[10] Sahni NS, Aastveit A, Næs T. In-line process and product control using spectroscopy and multivariate calibration. Journal of Quality Technology 2005; 37(1):1–20.
[11] Seber, G.A.F. (1983). "Multivariate Observations". Wiley series in Probability and Statistics.