Middle-East Journal of Scientific Research 22 (7): 1056-1062, 2014
ISSN 1990-9233

DOI: 10.5829/idosi.mejsr.2014.22.07.21966

The Analysis of Two-Factor Design with Interaction

Ukponmwan. H. Nosakhare and Ajibade, F. Bright

Department of General Studies, Mathematics and Computer Science Unit,
Petroleum Training Institute, P.M.B 20 Effurun, Warri, Delta State

Abstract: In this paper an attempt is made to study two factors in the same experimental design with interaction term included in the model. By interaction, we simply mean the failure of levels of factor A to behave consistently across levels of factor B and vice versa (Ochei *et al*, 2012). The general model for a two-factor with interaction given by $y_{ilk}=\mu+\alpha_i+\beta_j+(\alpha\beta)_{ij}+\epsilon_{ijk}$, i=1,2,…,p; j=1,2,…,q; k=1,2,…,r. Where, α_I denotes the effect of the ith level of factor A, β_j is the effect of the jth level of factor B, $(\alpha\beta)_{ij}$ denotes the interaction between the ith level of factor A and the jth level of factor B and ϵ_{ijk} is the error term associated with y_{ikj}, was employed. Then the following testable hypotheses against theiralternatives were tested for significance.

H_{00}: $(\alpha\beta)_{ij} = 0$ (No interaction)
H_{01}: $\alpha_1=\alpha_2=\ldots=\alpha_p$ (No differences among the effect of levels of factor A)
H_{20}: $\beta_1=\beta_2=\ldots=\beta_q$ (No differences among the effects of levels of factor B)

In testing these hypotheses, a design with two levels of each factor and five observations per treatment combination was considered. The results obtained in this case was extended to the general settings in which there were p levels of factor A, q levels of factor B and r observations per treatment combination.

Key words: Design · Interaction · Two-factor · Hypothesis · Non-estimability · Levels · Effects · Treatment combination · Non-testability · ANOVA · Test statistic · Generalization · Reparameterization

INTRODUCTION

Up to this time one factor problems have been emphasized. Many times in practice the researcher want to study the effects of two factors in the same experiment. For instance, a chemist might want to study the effect of both pressure and temperature on the viscosity of an adhesive; an engineer might study the effect of engine speed and oil typeon the life span of a piston ring; a medical researcher might study the effect of an exercise regimen and diet on blood sugar levels in diabetes. Several designs can be used to accomplish this [1]. One of the general designs suitable for this kind of experiment is the two-factor balanced design with interaction.

In a two-factor design with interaction, two factors are studied in the same experiment and an interaction term is included in the model. The ability to detect interactions in an experiment is a major advantage of multiple factor ANOVA. Various researchers vary in their recommendations regarding the continuation of the ANOVA procedure after encountering an interaction. A significant interaction will often mask the significance of main effects [2]. In this work all these challenges are tackled accordingly. However, a lengthy discussion of interactions is available in Dallas and Franklin [3], Freund [4], Hubert [5], Cox [6], Stapleton [7], Anderson and Braak [8], Eze and Ehiwario [9].

The Hypothesis: The null hypotheses to be tested shall be of the form

H_{00}: $(\alpha\beta)_{ij}=0$ (No interaction)
H_{10}: $\alpha_1=\alpha_2=\ldots=\alpha_I$ (No differences among the effects of levels of factor A)
H_{20}: $\beta_1=\beta_2=\ldots=\beta_j$ (No difference among the effects of levels of factor B)

Corresponding Author: Ajibade, F. Bright, Department of General Studies, Mathematics and Computer Science Unit, Petroleum Training Institute, P.M.B 20 Effurun, Warri, Delta State. E-mail: equalright_bright@yahoo.com.

In this design we first test for the presence of interaction. Subsequent hypothesis-testing strategy depends on the outcome of this initial test.

The Two-factor Interaction Model: The general model for a two-factor design with interaction is

$$y_{ijk}=\mu+\alpha_i+\beta_j+(\alpha\beta)_{ij}+\epsilon_{ijk},\ i=1,2,\dots,p;\ j=1,2,\dots,q;\ k=1,2,\dots,r \qquad (1.1)$$

where, α_i denotes the effect of the i^{th} level of factor A, β_j denotes the effect of the j^{th} level of factor B, $(\alpha\beta)_{ij}$ denotes the interaction between the i^{th} level of factor A and the j^{th} level of factor B and ϵ_{ijk} is the error term associated with y_{ikj}.

Developing the Appropriate Test Statistic: In seeing how to test our hypotheses in section 1.1, we consider in detail a design with two levels of each factor and two observations per treatment combination to formulate the appropriate test statistic. The result obtained in this specific case are extended easily to the general setting in which there are p levels of factor A, q levels of factor B and r observations per treatment combination.

The data layout for the 2x2 experimental design is given in Table 1.1

Table 1.1: Data layout for a two-factor design with p=q=r=2

		1	2
Factor A	1	y_{111} y_{112}	y_{211} y_{212}
	2	y_{121} y_{122}	y_{221} y_{222}

In matrix form, the model is

$$Y=X\beta+\epsilon \qquad (1.2)$$

where,

$$X=\begin{array}{c} \begin{matrix} \mu & \alpha_1 & \alpha_2 & \beta_1 & \beta_2 & (\alpha\beta)_{11} & (\alpha\beta)_{12} & (\alpha\beta)_{21} & (\alpha\beta)_{22} \end{matrix} \\ \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array} \qquad (1.3)$$

$$\beta=\begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ (\alpha\beta)_{11} \\ (\alpha\beta)_{12} \\ (\alpha\beta)_{21} \\ (\alpha\beta)_{22} \end{bmatrix} (1.4),\ y=\begin{bmatrix} y_{111} \\ y_{112} \\ y_{121} \\ y_{122} \\ y_{211} \\ y_{212} \\ y_{221} \\ y_{222} \end{bmatrix} (1.5) \text{ and } \epsilon=\begin{bmatrix} \varepsilon_{111} \\ \varepsilon_{112} \\ \varepsilon_{121} \\ \varepsilon_{122} \\ \varepsilon_{211} \\ \varepsilon_{212} \\ \varepsilon_{221} \\ \varepsilon_{222} \end{bmatrix} (1.6)$$

The matrix X, is of order 8x9. To determine its rank, we let C_i denote the i^{th} column of X. Bear in mind that each column of X can be expressed as a linear combination of column 6 (C_6), through column 9 (C_9), the columns that code the interaction terms and that these columns are linearly independent as demonstrated by Myers & Milton, (1991). Hence the rank of X is four, which proves that the matrix is less than full rank.

Test for Interaction: Our next assignment is to develop a test statistic for detecting the presence of interaction. It is risky to assume that if no interaction exists, then each of the interaction effects $(\alpha\beta)_{ij}$, is zero and that at least one of these is non-zero otherwise. The trouble arises from the fact that the interaction effects are not estimable. Hence the hypothesis

$$H_0\colon C\beta = 0 \qquad (1.7)$$

where

$$c=\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad (1.8)$$

is not testable. It can be proved by showing that $CQ \neq C$ given that $Q = (X'X)^c X'X$ where Q is

$$Q=\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad (1.9)$$

and $(X'X)^c$ is the conditional inverse of $X'X$. One of the ways of overcoming this challenge of non-estimability of the interaction effect and the non-testability of the hypothesis of equation (1.7)is by reparameterization. It is easy to verify by substitution that the definition of no interaction can be rephrased in terms of the original model parameters. In particular, no interaction exists if and only if

$$[(\alpha\beta)_{ij}-(\alpha\beta)_{ij'}]-[(\alpha\beta)_{i'j}-(\alpha\beta)_{i'j'}] = 0 \qquad (1.10)$$

for all i, i′, j and j′. In general, this criterion generates pq(p-1)(q-1) equations of which all but (p-1)(q-1) are redundant. From this it can be seen that to test for no interaction based on the original model parameters, we test a null hypothesis of the form

H_0: $C\beta^* = 0$ (1.11)

where C is an appropriately chosen matrix of ones and zeros of dimension (p-1)(q-1) x(p+q+pq+1).

In a two-factor design with p=q=r=2, the null hypothesis of no interaction is expressed as

H_0: $[(\alpha\beta)_{11}-(\alpha\beta)_{12}]-[(\alpha\beta)_{21}-(\alpha\beta)_{22}] = 0$ (1.12)

In matrix form, this null hypothesis of (1.11) becomes

H_0: $C\beta^* = 0$ (1.13)

where

$$C = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & -1 & -1 & 1 \end{bmatrix} \quad (1.14)$$

and

$$\beta = \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ (\alpha\beta)_{11} \\ (\alpha\beta)_{12} \\ (\alpha\beta)_{21} \\ (\alpha\beta)_{22} \end{bmatrix} \quad (1.15)$$

It can be proved that the null hypothesis of equation (1.13) is testable by showing that CQ = C where $Q = (X'X)^c X'X$.

As the number of levels of factor A and B increases, the system of equations needed to express the notion of no interaction becomes more complex (Myers &Milton, 1991). As a result of this, it is convenient to reparameterize the model in such a way that interaction can be expressed more simply by means of the new parameters. Specifically, we want reparameterize in such a way that the reparameterized "interaction" terms will be estimable. By so doing, interaction can be tested by considering the numerical values of these terms directly. The guide to the general technique can be provided [10].

Considering the two-factor model of equation (1.1)

$y_{ijk}=\mu+\alpha_i+\beta_j+(\alpha\beta)_{ij}+\epsilon_{ijk}$, i=1,2,…,p; j=1,2,…,q; k=1,2,…,r and $u_{ij}=\mu+\alpha_i+\beta_j+(\alpha\beta)_{ij}$ (1.16)

We define $\bar{\alpha}$, $\bar{\beta}$, $\bar{\mu}$, $\overline{(\alpha\beta)}_{..}$, $\overline{(\alpha\beta)}_{i.}$ and $\overline{(\alpha\beta)}_{.j}$ by

$$\left.\begin{aligned} &\bar{\alpha} = \textstyle\sum_{i=1}^{p}\alpha_i/p && \overline{(\alpha\beta)}_{..} = \textstyle\sum_{i=1}^{p}\sum_{j=1}^{q}(\alpha\beta)_{ij}/pq \\ &\bar{\beta} = \textstyle\sum_{j=1}^{q}\beta_i/q && \overline{(\alpha\beta)}_{i.} = \textstyle\sum_{j=1}^{q}(\alpha\beta)_{ij}/q \\ &\bar{\mu} = \textstyle\sum_{i=1}^{p}\sum_{j=1}^{q}\mu_{ij}/pq && \overline{(\alpha\beta)}_{.j} = \textstyle\sum_{i=1}^{p}(\alpha\beta)_{ij}/p \end{aligned}\right\} \quad (1.17)$$

The model can be expressed via these parameters as

$$y_{ijk} = [\mu+\bar{\alpha}+\bar{\beta}+\overline{(\alpha\beta)}_{..}]+[\alpha_i-\bar{\alpha}+\overline{(\alpha\beta)}_{i.}-\overline{(\alpha\beta)}_{..}]+[\beta_j-\bar{\beta}+\overline{(\alpha\beta)}_{.j}-\overline{(\alpha\beta)}_{..}]+[\overline{(\alpha\beta)}_{ij}-\overline{(\alpha\beta)}_{i.}-\overline{(\alpha\beta)}_{.j}+\overline{(\alpha\beta)}_{..}]+\varepsilon_{ijk} \quad (1.18)$$

which can be written as

$$y_{ijk}^* = \mu^*+\alpha_i^*+\beta_j^*+(\alpha\beta)_{ij}^*+\epsilon_{ijk} \quad (1.19)$$

where, $\mu^*=\mu+\bar{\alpha}+\bar{\beta}+\overline{(\alpha\beta)}_{..}$, $\alpha_i^*=\alpha_i-\bar{\alpha}+\overline{(\alpha\beta)}_{i.}-\overline{(\alpha\beta)}_{..}$, $\beta_j^*=\beta_j-\bar{\beta}+\overline{(\alpha\beta)}_{.j}-\overline{(\alpha\beta)}_{..}$, $(\alpha\beta)_{ij}^*=\overline{(\alpha\beta)}_{ij}-\overline{(\alpha\beta)}_{i.}-\overline{(\alpha\beta)}_{.j}+\overline{(\alpha\beta)}_{..}$

It can be shown by mathematical induction that each of these new parameters is estimable. Hence it is reasonable to expect that the null hypothesis of no interaction can be expressed simply in terms of these redefined parameters. Specifically, it can be shown that no interaction exists if and only if $(\alpha\beta)_{ij}^*=0$ for each iand j [11].

It is pertinent to mention that by defining α_i^*, β_j^* and $(\alpha\beta)_{ij}^*$ as has been done, certain restrictions have been induced which states that:

$$\textstyle\sum_{i=1}^{p}\alpha_i^* = \sum_{j=1}^{q}\beta_j^* = \sum_{i=1}^{p}(\alpha\beta)_{ij}^* = \sum_{j=1}^{q}(\alpha\beta)_{ij}^* = \sum_{i=1}^{p}\sum_{j=1}^{q}(\alpha\beta)_{ij}^* = 0 \quad (1.20)$$

These constraints make it easy to solve the normal equations for the reparameterized model.The design matrix for the general two-factor model is a matrix of ones and zeros which are generalization of equations (1.3), (1.4), (1.5) and (1.6) given respectively by

$$X = \begin{array}{c} \mu\alpha_1\alpha_2 \;\ldots\; \alpha_p|\beta_1\beta_2 \;\ldots\; \beta \\ \left[\begin{array}{c|cccc|cccc}
1 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\
1 & 1 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & & \vdots \\
1 & 1 & 0 & \cdots & 0 & 0 & 0 & \vdots & 1 \\
\hline
1 & 0 & 1 & \cdots & 0 & 1 & 0 & \cdots & 0 \\
1 & 0 & 1 & \cdots & 0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1 & 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 1 \\
\hline
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\hline
1 & 0 & 0 & \cdots & 1 & 1 & 0 & \cdots & 0 \\
1 & 0 & 0 & \cdots & 1 & 0 & 0 & \vdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\
1 & 0 & 0 & \cdots & 1 & 0 & 0 & & 1
\end{array}\right]\end{array}, \; \beta = \underline{\mu}, \; y = \begin{bmatrix} y_{11} \\ \vdots \\ y_{1q} \\ - \\ y_{21} \\ \vdots \\ y_{2q} \\ - \\ \vdots \\ - \\ y_{p1} \\ \vdots \\ y_{pq} \end{bmatrix} \text{ and } \varepsilon = \begin{bmatrix} \varepsilon_{11} \\ \vdots \\ \varepsilon_{1q} \\ - \\ \varepsilon_{21} \\ \vdots \\ \varepsilon_{2q} \\ - \\ \vdots \\ - \\ \varepsilon_{p1} \\ \vdots \\ \varepsilon_{pq} \end{bmatrix} \tag{1.21}$$

As usual, it is assumed that ϵ is a normally distributed random vector with mean 0 and variance $\sigma^2{}_{I.}$ So, the vectors,$X'X$ which is an (p+q+1)x(p+q+1) matrix of rank (p+q-1) used to establish the estimability of contrasts in the α's; $X'Y$; and C which is the (p-1)x(p+q+1) matrix forming a testable hypothesis of (1.13), are respectively given by [12]:

$$X'X = \begin{array}{c} \mu\alpha_1\alpha_2 \;\ldots\; \alpha_p|\beta_1\beta_2 \;\ldots\; \beta_q \\ \mu \\ \alpha \end{array}, \quad X'Y = \begin{bmatrix} y_{\cdots} \\ y_{1\cdot\cdot} \\ \vdots \\ - \\ y_{\cdot 1\cdot} \\ y_{\cdot 2\cdot} \\ \vdots \\ - \\ y_{\cdot 11} \\ y_{\cdot 12} \\ \vdots \\ - \\ y_{\cdot 21} \\ y_{\cdot 22} \\ \vdots \\ - \\ y_{\cdot 31} \\ y_{\cdot 32} \\ \vdots \\ - \end{bmatrix},$$

$$C = \begin{array}{c} \mu\alpha_1\alpha_2\alpha_3 \;\ldots\; \alpha_{p-1}\alpha_p\beta_1\beta_2 \;\ldots\; \beta_q \\ \begin{bmatrix}
0 & 1 & -1 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\
0 & 1 & 1 & -2 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 1 & 1 & 1 & \cdots & 1 & p-1 & 0 & 0 & \cdots & 0
\end{bmatrix}\end{array} \tag{1.22}$$

where

$$\left.\begin{aligned}
y_{\ldots} &= \sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{r} y_{ijk} \\
y_{i\ldots} &= \sum_{j=1}^{q}\sum_{k=1}^{r} y_{ijk} \\
y_{\cdot j\cdot} &= \sum_{i=1}^{p}\sum_{k=1}^{r} y_{ijk} \\
y_{ij\cdot} &= \sum_{k=1}^{r} y_{ijk}
\end{aligned}\right\} \quad (1.23) \tag{1.23}$$

The normal equations are given by

$$(X'X)b = X'Y \tag{1.24}$$

Then the estimators for μ^*, α_i^*, β_j^* and $(\alpha\beta)_{ij}^*$ are

$$\left.\begin{aligned} \hat{\mu}^* &= \bar{y}_{\dots} \\ \hat{\alpha}_i^* &= \bar{y}_{i\cdot\cdot} - \bar{y}_{\dots} \\ \hat{\beta}_j^* &= \bar{y}_{\cdot j\cdot} - \bar{y}_{\dots} \\ \widehat{(\alpha\beta)}_{ij}^* &= \bar{y}_{ij\cdot} - \bar{y}_{i\cdot\cdot} - \bar{y}_{\cdot j\cdot} + \bar{y}_{\dots} \end{aligned}\right\} \tag{1.25}$$

The regression sum of squares for the reparameterized full rank model is thus given by

$$SS_{Reg\ (Full)} = b^* X'Y = \sum_{i=1}^{p}\sum_{j=1}^{q} \frac{y_{ij\cdot}^2}{r} \tag{1.26}$$

where, $b^* = [\hat{\mu}^* \hat{\alpha}_i^* \hat{\beta}_j^* \widehat{(\alpha\beta)}_{ij}^*]$. In the reduced model, it is assumed that there is no interaction. The regression sum of squares for this model is given by

$$SS_{Reg\ (Reduced)} = \sum_{i=1}^{p} y_{i\cdot\cdot}^2/qr + \sum_{j=1}^{q} y_{\cdot j\cdot}^2/pr - y_{\dots}^2/pqr \tag{1.27}$$

By subtraction, $SS_{Reg\ (Hypothesis)} = SS_{Reg\ (Full)} - SS_{Reg\ (Reduced)} = \sum_{i=1}^{p}\sum_{j=1}^{q} \frac{y_{ij\cdot}^2}{r} - \sum_{i=1}^{p} \frac{y_{i\cdot\cdot}^2}{qr} - \sum_{j=1}^{q} \frac{y_{\cdot j\cdot}^2}{pr} + \frac{y_{\dots}^2}{pqr}$ (1.28)

The degrees of freedom associated with these sums of squares are pq, p+q-1 and (p-1)(q-1) respectively. The F-ratio used to test H_0 is

$$F_{(p-1)(q-1),\ pqr-pq} = \frac{SS_{Reg\ (Hypothesis)}/(p-1)(q-1)}{SS_{Res}/(pqr-pq)} \tag{1.29}$$

The ANOVAfor interaction test is summarized in Table 1.2

Tests for Main Effects and Interaction: To test for interaction only we test the hypothesis

$$H_0: (\alpha\beta)_{ij}^* = 0 \tag{1.30}$$

If interaction is detected in a two-factor design, it is suggested that factors be compared row by row or column by column using one-way classification procedure[10].The ANOVA Table for this procedure is summarized in Table 1.3

Application of the Procedure

Illustrative Example:A study of the solubility of two solutes in two different solvents is conducted. The study is aimed at testing the effect of the two solvents on the time required for the solutes to dissolve. The experiment is repeated for five consecutive times resulting in five observations for each treatment combinations. This was done to detect possible presence of interaction between the two factors. The data obtained is displayedin Table 1.4.

From the data using the procedural formulae, we have that;

$\bar{y}_{11\cdot} = 49.2$, $\bar{y}_{1\cdot\cdot} = 45.2$, $\bar{y}_{\dots} = 42.55$, $\bar{y}_{12\cdot} = 41.2$, $\bar{y}_{2\cdot\cdot} = 39.9$, $\bar{y}_{21\cdot} = 36.0$, $\bar{y}_{\cdot 1\cdot} = 42.6$, $\bar{y}_{22\cdot} = 43.8$, $\bar{y}_{\cdot 2\cdot} = 42.5$

The estimates of the parameters $(\alpha\beta)_{ij}^*$ are as follows:

$(\alpha\beta)_{11}^* = \bar{y}_{11\cdot} - \bar{y}_{1\cdot\cdot} - \bar{y}_{\cdot 1\cdot} + \bar{y}_{\dots} = 3.95$; $(\alpha\beta)_{12}^* = \bar{y}_{12\cdot} - \bar{y}_{1\cdot\cdot} - \bar{y}_{\cdot 2\cdot} + \bar{y}_{\dots} = -3.95$; $(\alpha\beta)_{21}^* = \bar{y}_{21\cdot} - \bar{y}_{2\cdot\cdot} - \bar{y}_{\cdot 1\cdot} + \bar{y}_{\dots} = -3.95$; $(\alpha\beta)_{22}^* = \bar{y}_{22\cdot} - \bar{y}_{2\cdot\cdot} - \bar{y}_{\cdot 2\cdot} + \bar{y}_{\dots} = 3.95$.

Table 1.2: ANOVA Table used to Test for Interaction in a Two Factor design

Sources of variation	Degrees of freedom	Sum of squares	Mean square	F-ratio
Regression	pq	$SS_{Reg\ (Full)}$	$MS_{Reg\ (Full)}$	
Reduced model	p+q-1	$SS_{Reg\ (Reduced)}$	$MS_{Reg\ (Reduced)}$	$F = \frac{SS_{Reg(Hypothesis)}/(p-1)(q-1)}{SS_{Res}/pqr-pq}$
Hypothesis	(p-1)(q-1)	$SS_{Reg\ (Hypothesis)}$	$MS_{Reg\ (Hypothesis)}$	
Residual	pqr-pq	SS_{Res}	MS_{Res}	
Total	Pqr	SS_{Total}		

Table 1.3: ANOVA Table for the General Two-factor Design with Interaction

Sources of variation	Degrees of freedom	Sum of squares	Mean square	F-ratio
Regression (Full)	pq	$\sum_{i=1}^{p}\sum_{j=1}^{q}\frac{y_{ij\cdot}^2}{r}$		
Mean	1	$y_{\ldots}^2/pqr$	$SS_{Factor\ A}/(p-1)$	$MS_{Factor\ A}/MS_{Res}$
Factor A	p-1	$\sum_{i=1}^{p}\frac{y_{i\cdot\cdot}^2}{qr}-\frac{y_{\ldots}^2}{pqr}$	$SS_{Factor\ B}/(q-1)$	$MS_{Factor\ B}/MS_{Res}$
Factor B	q-1	$\sum_{j=1}^{q}\frac{y_{\cdot j\cdot}^2}{pr}-\frac{y_{\ldots}^2}{pqr}$	$SS_{Interaction}/(p-1)\ (q-1)$	$MS_{Interaction}/MS_{Res}$
Interaction	(p-1)(q-1)	$SS_{Reg(Full)}-SS_{Mean}-SS_{FactorA}-SS_{Factor\ B}$	$SS_{Res}/(pqr-pq)$	
Residual	pqr-pq	$SS_{Total}-SS_{Reg(Full)}$		
Total (Uncorrected)	pqr	$\sum_{i=1}^{p}\sum_{j=1}^{q}\sum_{k=1}^{r}y_{ijk}^2$		

Table 1.4: Data for the Repeated Solubility Experiment of Two Solutes in Two Solvents

		Solvent (Factor I)	
		Water	Kerosene
Solute (Factor II)	Chalk	39 49 63 45 50	31 36 38 33 42
	Laterite	47 39 41 43 36	44 47 42 41 45

Table 1.5: ANOVA Test for Interaction in the Two-Factor Design

Sources of variation	Degrees of freedom	Sum of squares	Mean square	F-ratio
Regression	4	36,662.6		
Reduced model	3	36,350.55	312.05	$F=\frac{312.05}{29.9}=10.44$
Hypothesis	1	312.05	29.9	
Residual	16	478.4		
Total (Uncorrected)	20	37,141		

Table 1.6: One-Way ANOVA for Comparing Solutes in Water

Sources of variation	Degrees of freedom	Sum of squares	Mean square	F-ratio
Regression	2	20,590.4		
Reduced model	1	20,430.4	160.0	$F=\frac{160.0}{47.7}=3.35$
Hypothesis	1	160.0	47.7	
Residual	8	381.6		
Total (Uncorrected)	10	20,972.0		

Since the estimates are all far from zero, the data suggest the presence of interaction between factor I and factor II. To further analytically test for the significant presence or otherwise of the interaction effect, we test our hypothesis of (1.30). The row, column and cell totals of our illustrative data are thus:

$y_{11\cdot}=246$, $y_{12\cdot}=206$, $y_{21\cdot}=180$, $y_{22\cdot}=219$, $y_{1\cdot\cdot}=452$, $y_{2\cdot\cdot}=399$, $y_{21\cdot}=426$, $y_{\cdot 2\cdot}=425$.

Hence,

$SS_{Total(Uncorrected)} = \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{5} y_{ijk}^2 = 37{,}141$; $SS_{Reg(Full)} = \sum_{i=1}^{2}\sum_{j=1}^{2}\frac{y_{ij\cdot}^2}{5} = 36{,}662.6$; $SS_{Reg(Reduced)} = \sum_{i=1}^{2}\frac{y_{i\cdot\cdot}^2}{10} + \sum_{j=1}^{2}\frac{y_{\cdot j\cdot}^2}{10} - \frac{y_{\cdot\cdot\cdot}^2}{20}$ = 36,350.55.

The ANOVA test is summarized in Table 1.5.

Remark: Since $F_{Cal}(10.44) > F_{1,16}(4.49)$ with $p<0.01$ we therefore reject the null hypothesis of no interaction. This agrees with our initial results from our estimates that there is interaction between the two factors.

Based on the suggestion by scholars that if interaction is detected in a two-factor design, thenthe factors will be compared row by row or column by column using the one-way classification procedure discussed in Section (1.5). This test would be to see if there is a difference in the average time required for the two solutes to dissolve in water. We have that;

$SS_{Total} = \sum_{i=1}^{2}\sum_{k=1}^{5} y_{ijk}^2 = 20{,}972$, $y_{11\cdot}=246$, $y_{12\cdot}=206$, $y_{1\cdot\cdot}=452$, $\sum_{j=1}^{2}\frac{y_{ij}^2}{5} = 20{,}590.4$

The One-way ANOVA test for the main effects when interaction is detected is summarized in Table 1.6

Remark: Since $F_{Cal}(3.35) < F_{1,8}(3.46)$, $p>0.1$, we would accept that there is no difference in the average time required for these two solutes to dissolve in water.

CONCLUSION

In the study the researcher has successfully carried out the analysis of a two-way balanced design in the face of significant interaction between the factors. Specific methodology of transforming a non-estimable parameter and a non-testable hypothesis with a less than full rank model into estimable and testable full rank model by reparameterization was employed.

REFERENCES

1. Myers, R.H. and J.S. Milton, 1991. A First Course in the Theory of Linear Statistical Models, PWS-KENT Publishing Company, 20 Park Plaza, Boston, Massachusetts.
2. Montgomery, D.C., 2001. Design and analysis of experiments, 5th ed., John Wiley & Sons, Inc., New York.
3. Dallas, E.J. and A.G. Franklin, 1972. Estimation of variance in a two-way classification with interaction. Journal of the American Statistical Association, 67(338): 557-560.
4. Fruend, J.E., 1971. Mathematical Statistics. Second Edition. PRENTICE-HALL, INC, Englewood Cliffs, New Jersey.
5. Hubert, M.B., 1979. Social statistics.2nd edition. McGraw-Hill, Inc., Washington.
6. Cox, D.R., 1992. Planning experiments, John Wiley and Sons Inc., New York.
7. Stapleton, J.H., 1995. Linear statistical models, John Wiley and Sons, Inc., 605 Third Avenue, New York.
8. Anderson, M.J. and C.J.F. Ter Braak, 2003. Permutation tests for multi-factorial analysis of variance, Journal of Statistical Computation and Simulation, 73(2): 85-113.
9. Eze, F.C. and J.C. Ehiwario, 2009.Common f-test denominator for two-way interactive balanced design, Natural and Applied Sciences Journal, 10: 128-137.
10. Milton, J.S and J. Arnold, 1990. Introduction to Probability and Statistics: Principles and Applications for Engineering and the Computing Sciences. New York: McGraw-Hill.
11. Ochei, L.S, F.C. Eze and B.F. Ajibade, 2012. Analysis of Unbalanced Fixed-Effect Non-Interactive Model: An intra-Factor Approach.Journal of the Nigerian Association of Mathematical Physics, 22: 199-206.
12. Freund, J.E., 1974. Modern elementary statistics. 4th Edition. Prentice Hall Inc., London.