Journal of Basic Physical Research ISSN: 2141-8403 Prints, 2141-8411 Online
Vol.12, Issue 2, October, 2024, pp. 38 - 51. Available online at https://journals.unizik.edu.ng/jbpr/

A COMPARATIVE STATISTICAL ANALYSIS OF 2^2 FACTORIAL EXPERIMENT AND TWO-WAY ANALYSIS OF VARIANCE INTERACTIVE MODEL

F. C. Eze and G. U. Ortutu
Department of Statistics, Nnamdi-Azikiwe University, Awka

Abstract
A statistical comparative analysis of 2^2 factorial experiments were compared with Two-Way analysis of variance (ANOVA) interactive model. The three models for Two-Way ANOVA were considered with respect to when to test for the main effects. The three models are; fixed effect, random and mixed effect models. When the two factors namely; factor A and factor B are fixed, the common denominator for testing for the main effects is mean square error (MSe). Conversely, when the two factors are random, the common denominator for testing for the main effects is mean square interaction (MS_λ). However, when factor A is fixed and factor B is random, the denominator for testing for factor A is MS_λ while that of factor B is MSe. Conversely, when factor A is random and factor B is fixed, the denominator for testing for factor A is MSe while that of factor B is MS_λ. The 2^2 factorial experiments have no model specifications. After the statistical analysis, the results from the fixed effect model for the Two-Way ANOVA gave the same result with that of 2^2 factorial experiments using the Yates' technique or any other technique. The researcher therefore recommends that when both factors are fixed, either 2^2 factorial design or Two-Way analysis of variance (ANOVA) interactive model could be used. But when both factors are random or mixed, the Two-Way analysis of variance (ANOVA) interactive model is highly recommended.
Keywords: *Interactions, Mixed effect model, Random effect model, Mixed effect model.*

Introduction

A 2^k factorial experiment is a factorial experiment with k factors observed at 2 levels. The levels of the factors are 0 and 1. The low factor is observed at 0 level while the higher factor is 1. Roman capital letters are used to denote the factors while small letters are used to denote the levels of the factors. (1) is used to indicate that all factors involved in the experiment occur at their lowest level.

A **2×2 factorial design (2^2)** is a type of experimental design that allows researchers to understand the effects of two independent variables (each with two levels) on a single dependent variable. The k factors in this paper are factor A and factor B each observed at two levels 0 and 1. These factors A and B are compared with two-way analysis of experiments balanced design with more than one observation per cell.

When the levels of any of the factors is more than 2, the implementation of the 2^2 factorial experiments become complicated. The model of the experiment cannot be determined if they are fixed, random or mixed. A factor is said to be fixed if the entire treatment levels are observed. It is random when a random sample of both factors are observed and it is mixed when either factor is random or fixed and vice versa. The estimations of the treatment effects of 2^2 factorial design are determined by any of the following methods: Expansion of products, Even and odd rule, Sign Table and Yates Techniques.

Two-way ANOVA (analysis of variance) is a statistical test used to determine the differences between means in two variables. It can also be defined as an ANOVA test used to analyze the difference between the means of more than two groups. A two-way ANOVA is used to estimate how the mean of a quantitative variable changes according to the levels of two categorical variables.

Two-way ANOVA helps researchers answer questions about more complex relationships than those explored with just one independent variable.

The model for a 2-way ANOVA is

$$X_{ijk} = \mu + \alpha_i + \beta_j + \lambda_{ij} + \varepsilon_{ijk} \begin{cases} i = 1, 2, \ldots, p \\ j = 1, 2, \ldots, q \\ k = 1, 2, \ldots r \end{cases} \qquad (1)$$

Where X_{ijk} is the kth observation in ijth cell,
μ is a constant,
α_i is the average effect of factor A,
β_j is the average effect of factor B,
λ_{ij} is the interaction that exists between factor A and factor B,
ε_{ijk} is the error associated with X_{ijk}.

Equation (1) is said to be a fixed effect model or model 1

$$\text{if } \sum_i \alpha_i = \sum_j \beta_j = \sum_i \lambda_i = \sum_j \lambda_j = 0 \text{ and } \varepsilon_{ijk} \sim N\left(0, \sigma_\varepsilon^2\right). \qquad (2)$$

similarly, equation (1) is said to be a random effect model or model 2 if

$$\alpha_i \sim N\left(0, \sigma_\alpha^2\right); \beta_j \sim N\left(0, \sigma_\beta^2\right); \lambda_{ij} \sim N\left(0, \sigma_\lambda^2\right) \text{ and } \varepsilon_{ijk} \sim N\left(0, \sigma_\varepsilon^2\right) \qquad (3)$$

For mixed effect model or model 3 we have two cases:

For case 1: When factor A is fixed while factor B is random, we have

$$\sum_i \alpha_i = \sum_i \lambda_{ij} = 0; \beta_j \sim N\left(0, \sigma_\beta^2\right) \text{ and } \varepsilon_{ijk} \sim N\left(0, \sigma_\varepsilon^2\right). \qquad (4)$$

For case 2: When factor A is random while factor B is fixed, we have

$$\alpha_i \sim N\left(0, \sigma_\alpha^2\right) \; \sum_j \beta_j = \sum_j \lambda_{ij} = 0; \text{and } \varepsilon_{ijk} \sim N\left(0, \sigma_\varepsilon^2\right). \qquad (5)$$

There are varying F-ratios for testing for the main effects in equation (1) depending on if the model is fixed, random or mixed effect. Details of these will be discussed in the methodology.

When the levels of any of the factors are more than 2 levels, 2^2 factorial experiments become a problem. The model of the experiment is assumed to be fixed since it involves 2 levels.

Conversely, two-way ANOVA has more than 2 levels and the model can be determined if it is a fixed effect, random effect or mixed effect model. The problem arises when to use a 2^2 factorial experiments or 2-way ANOVA.

Stampfer, Burning, Willet, Rosner, Eberlein and Hennekens (1985) worked on the 2 x 2 factorial design which calls for randomizing each participant to treatment A or B to address one question and further assignment at random within each group to treatment C or D to examine a second issue, permitting the simultaneous test of two different hypotheses. The Physicians' Health Study, a randomized trial of aspirin and beta-carotene among U.S. physicians, illustrates some features and potential problems in the design and analysis of a factorial trial. The most common concern, interaction between treatments, is generally an advantage rather than a limitation of this design. They concluded that if the interaction is sufficiently severe, however, then loss of power is possible.

Further, Hennekens and Eberlin (1985) worked on the 2 x 2 factorial design: its application to a randomized trial of aspirin and carotene in U.S. physicians. In their study, the Physicians' Health Study, a randomized trial of aspirin and beta-carotene among U.S. physicians, illustrates some features and potential problems in the design and analysis of a factorial trial. The most common concern, interaction between treatments, is generally an advantage rather than a limitation of this design. Although such interactions are relatively uncommon, this design provides a means to measure an effect which otherwise might not be apparent. If the interaction is sufficiently severe, however, then loss of power is possible.

Satterfield, Greco, Goldhaber, Stampfer, Swartz, Stein, Kaplan and Hennekens (1990) worked on Biochemical markers of compliance in the Physicians' Health Study. The Physicians' Health Study is a randomized, double-blind, placebo-controlled trial using a 2 x 2 factorial design to test the effects of low-dose aspirin on risk of cardiovascular disease and beta-carotene supplementation on the incidence of cancer. They concluded that there was a highly significant positive correlation between levels of these biochemical markers and the self-reports of compliance ($r = 0.65$ for thromboxane B2 and $r = 0.69$ for beta-carotene, P-value less than .0001). These findings support the validity of the self-reported compliance in the Physicians' Health Study.

Whether 2x2 factorial experiments or 2-way (ANOVA) with more than one observation per cell, the study of the presence of interactions is very necessary. If interaction is present, it could obscure the validity of the experiment.

Aylin and KurtView (2021) studied Testing non-additivity (interaction) in two-way ANOVA tables with no replication. In their work, they argued that testing for any significant interaction between two variables depends on the number of replicates in each cell of the two-way table and structure of the interaction. When there are several observations taken at each level combination of two variables, testing non-additivity can easily be done by usual two-way ANOVA method which cannot be used when there is only one observation per cell they concluded.

Eze, Adimonye, Nnanwa and Ezeani (2013) worked on An Appropriate F-Test for Two-Way Balanced Interactive Model. The presence of interaction in a set of data/ model in a two-way interactive model may lead to a biased result when testing for the main effects. The nuisance parameter which is the interaction was removed from the data without distorting the assumption of homogeneity condition of analysis of variance. This is done by a linear combination such that the differences between the corresponding yield row-wise as well as column-wise difference is a constant and yet the total sum of the yield remains unchanged.

A key issue in various applications of analysis of variance (ANOVA) is testing for the interaction and the interpretation of resulting ANOVA tables. Jesper and Sven (2008) demonstrated that for a two-way ANOVA, whether interactions are incorporated or not may have a dramatic influence when considering the usual statistical tests for normality of residuals.

Iskandar, Noprianto, Bahtiar, Benfano, and Raymondus (2016) in their work Two-Way ANOVA with Interaction Approach to Compare Content Creation Speed Performance in Knowledge Management System examined the speed performance of content creation in four modules of BINUS University KMS: Documents, Video, Material, and Binuspedia. The examined mean value of speed performance was conducted in 3 location campuses of BINUS University: Anggrek Campus, Syahdan Campus and Alam Sutera Campus, each was done with five repetitions. The speed performance was analyzed using Two-way ANOVA with interaction approach. According to the experiment results, F calculated for module is greater than F table, F calculated for campus is less than F table and F calculated for interaction is less than F table. Thus, the conclusions are: (1) there are differences in the average value of speed performance for each module, (2) there is no difference in the average value of speed performance for each campus, and (3) there is no interaction between campuses and modules on the speed performance.

Gulab (2019) studied Methodology and Application of Two-way ANOVA and concluded that there are commonly two types of ANOVA tests for univariate analysis-One way ANOVA and Two-Way ANOVA. One way-ANOVA is used when one can interested in studying the effect of one independent variable factor on population, whereas Two-way ANOVA is used for studying the effects of two factors on population at the same time. He concluded that in many statistical applications in business administration, psychology, social science, and the natural sciences we need to compare more than two groups. For hypothesis testing more than two population means scientists have developed ANOVA method. The analysis of variance assumes that the observations are normally and independently distributed with the same variance for each treatment or factor level.

Abdulhafedh (2023) in analyzing the Impact of Age and Gender on COVID-19 Deaths Using Two-Way ANOVA can effectively determine whether the age and gender are significant factors in COVID-19 death cases in the US. The dependent variable in the analysis is the number of COVID deaths in the entire US, and the two independent variables are the age groups and gender (sex). The age groups consist of 11 subgroups or levels ranging from babies to elderly people. The sex is either male or female. Results showed that age group is a significant factor in COVID deaths, while gender was found to be insignificant factor in the mortality of COVID.

2^2 factorial experiment and 2-way ANOVA must obey the assumptions of ANOVA. ANOVA has certain assumptions that need to be met in order for the test to be valid. One of the most important assumptions of ANOVA is the normality assumption. This assumption states that the data should be normally distributed within each group. Normality is important because ANOVA is based on the assumption that the errors are normally distributed. If the data is not normally distributed, the results of ANOVA may not be accurate. Another important assumption of ANOVA is the homogeneity of variance assumption. This assumption states that the variance of the data should be equal across all groups. Homogeneity of variance is important because ANOVA assumes that the variances are equal across all groups. If the variances are not equal, the results of ANOVA may not be accurate. There are several ways to check for homogeneity of variance, including the Levene's test and the Bartlett's test. The independence assumption states that the observations within each group should be independent of each other. This means that the observations should not be influenced by each other. Independence is important because ANOVA assumes that the observations are independent of each other. If the observations are not independent, the results of ANOVA may not be accurate.

A valid interpretation of most statistical techniques requires that one or more assumptions be met. In published articles, however, little information tends to be reported on whether the data satisfy the assumptions underlying the statistical techniques used. This could be due to self-selection: Only manuscripts with data fulfilling the assumptions are submitted. Another explanation could be that violations of assumptions are rarely checked for in the first place Rink, Henk and Addie (2012).

Any statistical method not tested may led to a serious problem when analyzing the data (Olsen 2003 and Choi 2005) which can influence Type 1 and Type 11 errors and this can cause overestimation or underestimation of the inferential measures and effect sizes (Osborne and Waters, 2002). Keselman *et al.* (1998) argues that "The applied researcher who routinely adopts a traditional procedure without giving thought to its associated assumptions may unwittingly be filling the literature with non-replicable results."

Many authors have written on the robustness of some methods as regards the violations of assumptions of ANOVA (see Kohr and Games, 1974; Bradley, 1980; Sawilowsky and Blair,1992; Wilcox and Keselman,2003; Bathke,2004),

It is a common practice for researchers to test for assumptions of ANOVA to see if it satisfies or violate the assumptions of ANOVA. (Schucany and Ng, 2006) argued that that it is not appropriate to check assumptions by means of tests (such as Levene's test) carried out before deciding on which statistical analysis technique to use because such tests compound the probability of making a Type I error. Even if one desires to check whether or not an assumption is met, two problems stand in the way. First, assumptions are usually about the population, and in a sample the population is by definition not known. For example, it is usually not possible to determine the exact variance of the population in a sample-based study, and therefore it is also impossible to determine that two population variances are equal, as is required for the assumption of equal variances (also referred to as the assumption of homogeneity of variances) to be satisfied. Second, because assumptions are usually defined in a very strict way (e.g., all groups have equal variances in the population, or the variable is normally distributed in the population), the assumptions cannot reasonably be expected to be satisfied. Given these complications, researchers can usually only examine whether assumptions are not violated "too much" in their sample; for deciding on what is too much, information about the robustness of the technique with regard to violations of the assumptions is necessary.

Montgomery (2001) demonstrated that interaction occurs between two factors when the difference in response between the levels of one factor is not the same at all levels of the other factors. According to Montgomery (2001), a significant interaction can mask the significance of the main effects. This means that, when interaction is present, the main effects of the factors involved in the interaction may not have much meaning.

Methodology

1. The 2^2 Design. A 2^k factorial design is a k-factor design such that

(i) Each factor has two levels (coded -1 and +1) or (0 and 1).

(ii) The 2^k experimental runs are based on the 2^k combinations of the $\pm$ factor levels. The simplest 2^k design is the 2^2 designs. This is a special case of a two-factor factorial design with factors A and B having two levels. A 2^2 design has only 4 runs, therefore several (n) replications are taken. Notationally, lowercase letters a, b, ab, and (1) to indicate the sum of the responses for all replications at each of the corresponding levels of A and B. If the lower-case letter appears, then that factor is at its high (+1) level. If the lower-case letter does not appear, then that factor is at its low (1) level. This can be demonstrated below:

Table 1
2^2 factorial Experiment

$$\begin{array}{c} A \\ \begin{pmatrix} a_0 & & a_1 \\ b_0 & (1) & a \\ b_1 & b & ab \end{pmatrix} \end{array}$$

This can be further demonstrated as follows:

Factor level combination	**Coded levels**	**Replicates 1, 2, … n**	**Sum of n replicates**
A low, B low	-1 -1	xxx xxx …xxx	(1) = y_{11}
A high, B low	+1 -1	xxx xxx …xxx	a = y_{21}
A low, B high	-1 +1	xxx xxx …xxx	b = y_{12}
A high, B high	+1 +1	xxx xxx …xxx	ab = y_{22}

The notation A^+ and A^{-1} to represent the set of observations with factor A at its high (+1) and its low (-1) levels, respectively. The same notation applies to B^+ and B for factor B^{-1}.

a and ab correspond to A^+ and (1) and b correspond to A^{-1}.

b and ab correspond to B^+ and (1) and a correspond to B^{-1}. yA+ and yA are the means of all observations when A = +1 and A = -1, respectively.

$\bar{y}_{A+}$ and $\bar{y}_{A-}$ are the means of all observations when $A=+1$ and $A=-1$ respectively.

$\bar{y}_{B+}$ and $\bar{y}_{B-}$ are the means of all observations when $B=+1$ and $B=-1$ respectively.

The average effect of a factor is the average change in the response produced by a change in the level of that factor averaged over the levels of the other factor. For a 2^2 design with r replicates, the

Average effect of factor A is $A=\bar{y}_{A+}-\bar{y}_{A-}=\frac{1}{2^2*r}\left(ab+a-b-(1)\right)$.

Average effect of factor B is $B=\bar{y}_{B+}-\bar{y}_{B-}=\frac{1}{2^2*r}\left(ab-a+b-(1)\right)$.

Interaction effect between Factors A and B, denoted AB, is the difference between (i) the average change in response when the levels of Factor A are changed given Factor B is at its high level and (ii) the average change in response when the levels of Factor A are changed given Factor B is at its low level:

$$AB = \left(\bar{y}_{A^+B^+} - \bar{y}_{A^+B^+}\right) - \left(\bar{y}_{A^+B^-} - \bar{y}_{A^-B^-}\right) = \frac{ab - a - b + (1)}{2^2 * r}$$

This average effect is the same if we use the rules of Expansion of products, Even and odd rule, Sign Table or Yates Techniques and divide the effects by 2^k x r where k is the number of factors and r is the number of replicates.

When estimating the effects for A, B and AB, the following contrasts are used.

$$\Gamma A = ab + a - b - (1),\ \Gamma B = ab - a + b - (1),\ \Gamma AB = ab - a - b + (1)$$

$\Gamma A, \Gamma B,$ and ΓAB are used to estimate A, B and AB and they are orthogonal contrasts.

The coefficient vectors for the contrasts are [1 1 -1 -1] for A, [1- 1 1 -1] for B, and [1 -1 -1 1] for AB. Note the dot product of any two vectors = 0. This is why they are called orthogonal contrasts.

The sum of squares for contrast Γ are:
For a replicated 2^2 design, this is equivalent to:

$$SS_A = \frac{\left[ab + a - b - (1)\right]}{4r}, SS_B = \frac{\left[ab - a + b - (1)\right]}{4r}, SS_{AB} = \frac{\left[ab - a - b + (1)\right]}{4r}$$

There are two levels for both factors, the degree of freedom associated with each sum of squares is 1. Thus $SS_A = MS_A, SS_B = MS_B$ and $SS_{AB} = MS_{AB}$

There are r replicates for each of four A*B treatment combinations, there are 4(r-1) degree of freedom for error for the four parameters interaction model.

2. The Two-way Analysis of Variance (ANOVA)

Equations (1) to (5) defines the model for Two-way analysis of variance with more than one observation per cell.

Equation (2) defines the conditions for a fixed effect model of equation (1). The expected mean squares are shown in Table 2 below (Montogomery (1991)).

Table 2
Expected mean squares for fixed effect 2-way ANOVA

Factor	Expected mean square
α_i	$\sigma_\varepsilon^2 + \frac{qr\sum_i \alpha_i^2}{p-1}$
β_j	$\sigma_\varepsilon^2 + \frac{pr\sum_j \beta_j^2}{q-1}$
λ_{ij}	$\sigma_\varepsilon^2 + \frac{r\sum_{ij} \lambda_{ij}^2}{(p-1)(q-1)}$
ε_{ijk}	σ_ε^2

From Table 3.2, if we are interested to test for the main effect for model 1 (fixed effect model) we shall have:

$$H_{01}:\alpha_1=\alpha_2=...=\alpha_p;\ H_{02}:\beta_1=\beta_2=...=\beta_q;\ H_{03}:\lambda_{11}=\lambda_{12}=...=\lambda_{pq}$$

The $F-$ratio for H_{01} is $\frac{MS_\alpha}{MS\varepsilon}$; The $F-$ratio for H_{02} is $\frac{MS_\beta}{MS\varepsilon}$; The $F-$ratio for H_{03} is $\frac{MS_\lambda}{MS\varepsilon}$

Similarly, the expected mean squares for random effect model (model 2) are shown in Table 3.

Table 3
Expected mean squares for random effect 2-way ANOVA

Factor	Expected mean square
α_i	$\sigma_\varepsilon^2+r\sigma_\lambda^2+qr\sigma_\alpha^2$
β_j	$\sigma_\varepsilon^2+r\sigma_\lambda^2+pr\sigma_\beta^2$
λ_{ij}	$\sigma_\varepsilon^2+r\sigma_\lambda^2$
ε_{ijk}	σ_ε^2

Under model 2, if we are interested to test for the random effect we shall have

$$H_{01}:\sigma_\alpha^2=0;\ H_{02}:\sigma_\beta^2=0;\ H_{03}:\sigma_\lambda^2=0$$

The corresponding F-ratios are

The $F-$ratio for H_{01} is $\frac{MS_\alpha}{MS_\lambda}$; The $F-$ratio for H_{02} is $\frac{MS_\beta}{MS_\lambda}$; The $F-$ratio for H_{03} is $\frac{MS_\lambda}{MS\varepsilon}$

Under mixed effect model (model 3), there are two cases viz:

Case 1: Here factor A is fixed and factor B is random. The expected mean squares are shown in Table 4.

Table 4
Expected mean squares for mixed effect 2-way ANOVA (case 1)

Factor	Expected mean square
α_i	$\sigma_\varepsilon^2+r\sigma_\lambda^2+qr\sigma_\alpha^2$
β_j	$\sigma_\varepsilon^2+pr\sigma_\beta^2$
λ_{ij}	$\sigma_\varepsilon^2+r\sigma_\lambda^2$
ε_{ijk}	σ_ε^2

Here, the hypotheses under mixed effect case 1, we have

$$H_{01}:\alpha_1=\alpha_2=...=\alpha_p;\ H_{02}:\sigma_\beta^2=0;\ H_{03}:\sigma_\lambda^2=0$$

The corresponding F-ratios are

The $F-$ratio for H_{01} is $\frac{MS_\alpha}{MS_\lambda}$; The $F-$ratio for H_{02} is $\frac{MS_\beta}{MS_\varepsilon}$; The $F-$ratio for H_{03} is $\frac{MS_\lambda}{MS_\varepsilon}$ Case 2: Here factor A is random and factor B is fixed. The expected mean squares are shown in Table 3.5.

Table 5
Expected mean squares for mixed effect 2-way ANOVA (case 2)

Factor	Expected mean square
α_i	$\sigma_\varepsilon^2 + qr\sigma_\alpha^2$
β_j	$\sigma_\varepsilon^2 + r\sigma_\lambda^2 + pr\frac{\sum_j \beta_j^2}{q-1}$
λ_{ij}	$\sigma_\varepsilon^2 + r\sigma_\lambda^2$
ε_{ijk}	σ_ε^2

The hypotheses under mixed effect case 2, we have

$$H_{01}:\sigma_\alpha^2;\ H_{02}:\beta_1=\beta_2=...=\beta_q;\ H_{03}:\sigma_\lambda^2=0$$

The corresponding F-ratios are

The $F-$ratio for H_{01} is $\frac{MS_\alpha}{MS_\varepsilon}$; The $F-$ratio for H_{02} is $\frac{MS_\beta}{MS_\lambda}$; The $F-$ratio for H_{03} is $\frac{MS_\lambda}{MS_\varepsilon}$

There are varying denominators for testing for the main effects for all the models.

3. Discussions and Results

Response variable(s) in any experiment can be found to be affected by a number of factors in the overall system some of which are controlled or maintained at desired levels in the experiment. An experiment in which the treatments consist of all possible combinations of the selected levels in two or more factors is referred as a factorial experiment. For example, an experiment on rooting of cuttings involving two factors, each at two levels, such as two hormones at two doses, is referred to as a 2 x 2 or a 2^2 factorial experiment. Its treatments consist of the following four possible combinations of the two levels in each of the two factors.

Table 6: 2^2 factorial experiments

	Treatment combination	
Treatment number	**Hormone**	**Dose (ppm)**
1	NNA	10
2	NNA	20
3	IBA	10
4	IBA	20

Source: https://www.fao.org

When the treatments include all combinations of the selected levels of the factor we term it complete factorial experiment. On the other hand, when only a fraction of all the combinations is tested, we term it fractional factorial experiment.

If the factors appear at more than two levels, the above procedure becomes complicated since there are many ways to express the differences between the average responses.

For example:

Table 7. Data from a 2x2 factorial experiment

Factor B

	Level	b_1	b_2
	a_1	20	30
Factor A			
	a_2	40	52

Source: https://www.fao.org

The main effect of factor A is the difference between the mean response at the first level of A and the mean response at the second level of A. This is

$$A = \frac{40+52}{2} - \frac{20+30}{2} = 21$$

This means increasing factor A from level 1 to level to 2 causes an average increase in the response by 21 units. Similarly, the main effect of B is

$$B = \frac{30+52}{2} - \frac{20+40}{2} = 11$$

The interaction between the factors can be checked. In some cases, the difference in response between the levels of one factor is not the same at all levels of the other factors. When this occurs, there is an interaction between the factors.

The graph in Table 7 is shown below

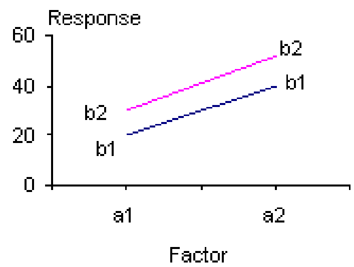


Figure 1*: The graph of Table 7*

Table 8: Data for 2x2 factorial experiment

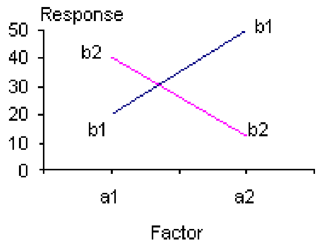
Factor B

Levels	b_1	b_2

	a_1	20	40
Factor A			
	a_2	50	12

Source: https://www.fao.org

The graph in Table 8 is shown below



F*igure 2: The graph in Table 8 showing the presence of interaction*.

Suppose an exploration on how a detergent and a particular water temperature affect the dirt removal of laundry. Interest is on how to check if the combined effect of detergent and water temperature can affect the dirt removal. Here two types of detergents, namely x and y is assumed. Three types of water temperatures, the levels of which are cold, warm and hot were used. Here the detergent and the water temperature are independent variables, while the amount of dirt removed is the dependent variable.

With each combination, one would wash five loads, which shall be replicates. Suppose the information is as follows:

Table 9
Two-way ANOVA with three levels

	Water Temperature		
Detergent	**Cold**	**Warm**	**Hot**
Detergent x	4	7	10
	5	8	11
	5	9	12
	6	12	19
	5	3	15
Detergent y	4	12	10
	4	12	12
	6	13	13
	6	15	13
	5	13	12

Source: Wallstreetmojo Team, Statistics Guides (2024)

Since the interest is on 2 levels, the level for warm temperatures will be deleted to have cold and hot temperatures. The cold temperature represents low level while the hot temperatures represent the high temperatures.

Table 10
Two-way ANOVA with three levels

Detergent	**Cold**	**Hot**
Detergent x	4	10
	5	11
	5	12

	6 5	19 15
Detergent y	4 4 6 6 5	10 12 13 13 12

The normality test for the data in Table 10 was tested and presented in Table 11.

Table 11

Tests of Normality

	Temperature	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Observations	cold	.200	10	.200*	.832	10	.035
	warm	.255	10	.064	.845	10	.051

Table 11 shows that the data is normally distributed.

The test for constant variance is shown Table 12.

Table 12
Test of Homogeneity of Variances
Observations

Levene Statistic	df1	df2	Sig.
4.239	1	18	.064

Table 12 shows that the data has constant variance,
Finally, the test for independent is shown in Table 13.

Table 13
Chi-square Tests

	Value	df	Asymptotic Significance(2-sides)
Pearson Chi-Square	60.000[a]	64	.619
Likelihood Ratio	49.781	64	.904
Linear-by-Linear	.001	1	.976
Association	20		
N of Valid Cases			

The result shows that the data is independent since the p-value is greater than 0.05.
The analysis of the data in Table 10 using Yates' technique is shown in Table 14.

Table 14
Yates' technique

Treatment combinations	Yield	Col 1	Col 2	SS col
(1)	25	92	177	1566.45
a	67	85	77	296.45
b	25	42	-7	2.45
ab	60	35	-7	2.45

In Table 14, 'a' is the water temperature while 'b' is detergent. 'ab' is the interactions between water temperatures and detergent.

$SS_{(1)}$ is $\frac{177^2}{2^2 * 5} = 1566.45$

$$SS_A = \frac{77^2}{2^2 * 5} = 296.45; SS_B = \frac{-7^2}{2^2 * 5} = 2.45; SS_{AB} = \frac{-7^2}{2^2 * 5} = 2.45$$

The ANOVA Table is presented in Table 15.

Table 15
ANOVA Table for 2^2 factorial Experiments

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	301.350[a]	3	100.450	24.650	.000
Intercept	1566.450	1	1566.450	384.405	.000
Temperature	296.450	1	296.450	72.748	.000
Detergent	2.450	1	2.450	.601	.449
Temperature * Detergent	2.450	1	2.450	.601	.449
Error	65.200	16	4.075		
Total	1933.000	20			
Corrected Total	366.550	19			

Looking at Tables 14 and 15, the sum of squares for both 2^2 factorial experiments and Two-way ANOVA are the same when both factors are fixed as well as the conclusion for the test. When factor A and factor B are random, the common denominator for testing for the main effect is the mean square interaction $\left(MS_{\lambda}\right)$. The ANOVA Table is shown in Table 16.

Table 16
ANOVA Table for Two-way random model

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	Hypothesis	1566.450	1	1566.450	5.284	.261
	Error	296.410	1.000	296.450[a]		
Temperature	Hypothesis	296.450	1	296.450	121.000	.058
	Error	2.450	1	2.450[b]		
Detergent	Hypothesis	2.450	1	2.450	1.000	.500
	Error	2.450	1	2.450[b]		
Temperature * Detergent	Hypothesis	2.450	1	2.450	.601	.449
	Error	65.200	16	4.075[c]		

When factor A is fixed and factor B is random and vice versa the ANOVA result will be the same as in Table 16, since they have the same mean square.

Conclusion

The analytical results obtained from 2^k factorial experiment and that of Two-Way analysis of variance when both factors are fixed gave the same results. The results when both factors are random or mixed gave varying results.

Having carefully studied the analytical comparisons between the 2^2 factorial experiment and that of Two-Way analysis of variance, it is recommended to either use 2^2 factorial experimental method or Two-Way analysis of variance when both factors are fixed. However, when both factors are random or mixed, the Two-Way ANOVA should be used.

References

Aylin Alim and S. Kurt View (2021). Testing non-additivity (interaction) in two-way ANOVA tables with no interaction. Sage Journals vol. 15 issue 1 14-20.

Azad, Abdulhafedh (2023). Analyzing the Impact of Age and Gender on Covid-19 Deaths Using Two-Way ANOVA. Open Access Library Journal vol. 10 Issue 1 25-35.

Bathke A. (2004). The ANOVA F test can still be used in some unbalanced designs with unequal variances and nonnormal data. J. Stat. Plan. Inference 126, 413–422 10.1016/j.jspi.2003.09.010.

Bradley J. V. (1980). Nonrobustness in Z, t, and F tests at large sample sizes. Bull. Psychon. Soc. 16, 333–336.

FC Eze, FO Adimonye, CP Nnawa and M I Ezeani (2013). An Appropriate Test for Two-Way Balanced Interactive Model. International Journal of Computational Engineering Research. Vol. 3 issue 2 121-127.

Gulab M Nibrad (2019). Method and Application of Two-Way ANOVA. International Journal of Marketing and Technology vol. 9 issue 6 56-65.

Hennekens CH and K Eberlein (1985). A randomized trial of aspirin and beta carotene among U.S. physicians. Pre. Med journal 14(2) 165-8.

Iskandar, Noprianto, Bahtiar, Benfano, and Raymondus(2016). Two-Way ANOVA with interaction approach to compare content creation speed performance in knowledge management system. Conference Paper· November 2016 DOI: 10.1109/ KICSS.2016.7951453

Kohr R. L., Games P. A. (1974). Robustness of the analysis of variance, the Welch procedure and a Box procedure to heterogeneous variances. J. Exp. Educ. 43, 61–69

Montgomery, D.C. (2001). Design and Analysis of Experiments. John Wiley & Sons New York p.176.

Osborne J. W., Waters E. (2002). Four assumptions of multiple regression that researchers should always test. Pract. Assess. Res. Evalu. 8

Rink Hoekstra, Henk A.L, and Addie Johnson (2012). Are Assumptions of Well-Known Statistical Techniques Checked, and Why (Not)? Front Psychol 3. 137

Satterfield S , P J Greco, S Z Goldhaber, M J Stampfer, S L Swartz, E A Stein, L Kaplan, C H Hennekens (1990). Biochemical markers of compliance in the Physicians' Health Study. Am J Prev Med 6(5):290-4.

Sawilowsky S. S., Blair R. C. (1992). A more realistic look at the robustness and type II error properties of the t test to departures from population normality. Psychol. Bull. 111, 352–360 10.1037/0033-2909.111.2.352

Schucany W. R., Ng H. K. T. (2006). Preliminary goodness-of-fit tests for normality do not validate the one-sample Student *t*. Commun. Stat. Theory Methods 35, 2275–2286 10.