Afropolitan Journals

ISSN: 2730-7603X www.afropolitanjournals.com

Analysis of Variance (ANOVA) Randomized Block Design (RBD) to Test the Variability of Three Different Types of Fertilizers (NPK, UREA and SSP) on Millet Production

Nuhu Ibrahim and Aliyu Bayero Abdullahi

Department of Mathematics & Statistics, Federal Polytechnic Mubi. Adamawa State, Nigeria.

Corresponding author: aliyubayeroabdullahi@gmail.com

Abstract

This research was undertaken to test the variability of three different types of fertilizers (NPK, UREA and SSP) applied on sub-divided plots of millet farm land. Each fertilizer was applied separately on specific area at equal interval till the harvest period. The aim was to see whether the sub plots could yield the same quantity output or not as a result of the applications of the three fertilizers. The significance test conducted shows that the calculated F-Ratio was 9.18, while the critical value of F-tabulated at 5% level of significant was 3.89. This is an indication that the Alternative hypothesis (H_1) was significant and therefore the Null hypothesis (H_0) was rejected. The rejection of the null hypothesis indicates there are differences in the quantities of yields obtained from the sub-divided plots based on the quality of each fertilizer. Pair-wise comparison was conducted using three different methods (Tukey Cramer, Bonferroni and Dunn-Sidak methods) to see where the differences in production outputs lie. It was noticed that Urea was different from fertilizers one (NPK) and three (SSP). This means Urea fertilizer is less effective compared with NPK and SSP which were similar in output.

Keywords: Fertilizer, millet farm land, F-Ratio, F-tabulated, level of significant.

Introduction

Analysis of variance (ANOVA) is a statistical tool used to detect differences between experimental group means. ANOVA is warranted in experimental designs with one dependent variable that is a continuous parametric numerical outcome measures, and multiple experimental groups within one or more independent variables (Montgomery, 2003) in ANOVA terminology, independent variables are called factors, and groups within each factor are referred to as levels. The array of terms that are part and parcel of ANOVA can be intimidating to the uninitiated, such as partitioning of variance, main effects, interactions, factors, sum of squares, mean squares, F-scores, family wise alpha, multiple comparison procedures (or post hoc tests), effect size, statistical power, etc. How do these terms pertain to P values and statistical significance? Sir Ronald Fisher pioneered the development of ANOVA for analyzing results of agricultural experiments (Wackerly, 2002). Today, ANOVA is included in almost every statistical package, which makes it accessible to investigators in all experimental sciences. It is easy to impute a data set and run a simple ANOVA, but it is challenging to choose the appropriate ANOVA for different experimental designs, to examine whether data adhere to the modeling assumptions, and to interpret the results correctly (Aczel, 1974). In ANOVA terminology, independent variables are called

factors, and groups within each factor are referred to as levels. The array of terms that are part and parcel of ANOVA can be intimidating to the un-initiated, such as partitioning of variance, mean effects, interaction, factors, sum of squares, mean squares, f-scores, family wise alpha, multiple comparison procedures (or post Hoc tests), effect size, statistical power, etc. how do these terms pertain to P values and statistical significance? What precisely is meant by a "statistically significant ANOVA"? How does analyzing variance result in an inferential decision about differences in group means? Can ANOVA be performed on non-parametric data? The intent is to provide the clinician reader, whose misspent youth did not include an enthusiastic reading of statistics textbooks, an understanding of the fundamentals of this widely used form of inferential statistical analysis (Eva, 2010).

ANOVA is applicable when the aim is to infer differences in group values when there is one independent variable and more than two groups, such as one independent variable with three or more levels, or when there are two or more independent variables. Since an independent variable is called a 'factor', ANOVAs are described in terms of the number of factors; if there are two independent variables, it is a two-factor ANOVA. In the simpler case of a one factor ANOVA, the Null hypothesis asserts that the population means for each level (group) of the independent variable are equal. ANOVA evaluates differences in group means in a round-about fashion, and involves the 'partitioning of variance' from calculations of 'sum of squares' and 'Mean Squares'. Three metrics are used in calculating the ANOVA test statistic, which is called the F-score (named after R.A. Fisher, the developer of ANOVA): (i) Grand Mean, which is the mean of all scores in all groups, (ii) Sum of Squares, which are of two kinds, the sum of all squared differences between group means and the Grand Mean (between-groups sum of squares) and the sum of squared differences between individual data scores and their respective group mean (within-groups sum of squares), and (iii) Mean Squares, also of two kinds (between-groups Mean Squares, Within-groups Mean Squares), which are the average deviations of individual scores from their respective mean, calculated by dividing sum of Squares by their appropriate degrees of freedom.

Assumptions of ANOVA

Assumptions for ANOVA pertain to the underlying mathematics of general linear models, specifically; a data set should meet the following criteria before being subjected to ANOVA.

Parametric data: A parametric ANOVA, requires parametric data (ratio or interval measures). There are non-parametric, one-factor versions of ANOVA for non-parametric ordinal (ranked) data, specifically the Kruskal-Wallis test for independent groups and the Friedman test for repeated measures analysis (Scot, 2013).

Normally distributed data within each group: ANOVA can be thought of as a way to infer whether the normal distribution curves of different data sets are best thought of as being from the same population or different populations. It follows that a fundamental assumption of parametric ANOVA is that each group of data (each level) be normally distributed. The Shapiro-Wilk test[2] is commonly used to test for normality for group sample

size (N) less than 50; D'Agnostino's modification[3]is useful for larger samplings (N>50) (Wilk, 1965).

A normally distributed curve can be described by whether it has symmetry about the mean and the appropriate width and height. These attributes are defined statistically by 'Skewness' and 'kurtoses respectively. A normally distributed curve will have Skewness =0 and Kurtosis =3. (Note that an alternative definition of Kurtosis subtract 3 from the final value so that a normal distribution will have Kurtosis = 0. This 'minus 3' Kurtosis value is sometimes referred to as 'excess Kurtosis' to distinguish it from the value obtained with the standard Kurtosis function. The Kurtosis value calculated by many statistical programs is the 'minus3' variant but is referred to somewhat misleadingly, as 'Kurtosis. Normality of a data set can be assessed with a z-test in reference to the standard error of Skewness (estimated as $\sqrt{\frac{6}{N}}$) and the standard error of Kurtosis (estimated as $\sqrt{\frac{24}{N}}$)[4]. A conservative alpha of 0.01 ($z \geq 2.56$) is appropriate, due to the overly sensitive nature of these tests, especially for large sample sizes (> 100)[4]. As a computational example, for N=20, the estimation of standard error of Skewness value greater than $\pm 2.56 \times 0.55 = \pm 1.41$ would indicate non-normality. Perhaps the best 'test' is what always should be done: examine a histogram of the distribution of the data. In practice, any distribution that resembles a bell-shaped curve will be 'normal enough' to pass normality tests, especially if the samples size is adequate.

Homogeneity of variance within each group: Referring again to the notion that ANOVA compares normal distribution curves of data sets, these curves need to be similar to each other in shape and width for the comparison to be valid. In other words, the amount of data dispersion (variance) needs to be similar between groups. Two commonly invoked tests of homogeneity of variance are by Levene, Brown & Forsthye (forsthye, 2013).

Independent Observations: a general assumption of parametric analysis is that the value of each observation for each subject is independent of the value of any other observation. For independent groups designs, this issue is addressed with random sampling, random assignment to groups, and experimental control of extraneous variables. This assumption is an inherent concern for repeated measures designs, in which an assumption of sphericity comes into play (Shapiro & Wilk, 1965). When subjects are exposed to all levels of an independent variable (e.g., all treatments), it is conceivable that the effects of a treatment can persist and affect the response to subsequent treatments. For example, if a treatment effect for one level has a long half-time (analogous to drug effect) and there is inadequate 'wash out' time between exposures to different levels (treatments), there will be a carryover effect. A well designed and executed cross-over experimental design can mitigate carry-over effects. Mauchly's test of sphericity is commonly employed to test the assumption of independence in repeated measures designs. If the Mauchly test is statistically significant, corrections to the F score calculation are warranted.

The primary assumptions in applying ANOVA are that the response must follow a normal distribution. The variance of the response is Constant and they are independent and randomly distributed.

The ANOVA test is computed based on decomposition of the total variances, α^2 into two distinct components. The first component is known as the between group variance, and it involves finding the variance of the means. The second component, the within group or (levels) variance, by computing the variance using all data and is not affected by differences in the means, (Brown, 1974). If there is no difference in the COD means, the between group variance estimate will be approximately equal to the within group variance estimate, and the F statistics test will be equal to 1. The null hypothesis will not be rejected.

ANOVA General Linear Models

ANOVA is based mathematically on linear regression and general linear models that quantify the relationship between the dependent variable and the independent variable(s) (Wackerly, 2002). There are three different general linear models for ANOVA: (i) Fixed effects model. This makes inferences that are specific and valid only to the populations and treatments of the study. (ii) Random effects model. This makes inferences about levels of the factor that are not used in the study, such as a continuum of drug doses when the study only used three doses. This model pertains to random effects within levels, and makes inferences about a population's random variation. (iii) Mixed effects model. This can contain both Fixed and Random effects.

In most types of orthopedic rehabilitation clinical research, the fixed effect model is relevant since the statistical inferences being sought are fixed to the levels of the experimental design (Brown, 2012).

One-way AVOVA test procedure

A one-way analysis of Variance is used when the data are divided into groups according to only one factor.

Assume that the data $x_{11}, x_{12}, x_{13}. . . , x_{1n1}$ are sample from population 1, $x_{21}, x_{22}, x_{23}, . . . x_{2n2}$ are sample from population 2, . . . , $x_{k1}, x_{k2}, x_{k3}, . . . , x_{knk}$ are sample from population k. let x_{ij} denote the data from the sample i^{th} group (level) and j^{th} observation.

We have values of independent normal random variables x_{ij}, I = 1, 2. . . K and j = 1, 2. . . n_i with mean μ and constant standard deviation δ, $x_{ij} - N(\mu, \delta)$. Alternatively, each $x_{ij} = \mu_1 + \sum_{ij}$ Where $\sum_{ij}$ are normally distributed independent random errors, $\sum_{ij} \rightarrow$ N (o, δ). Let $N = n_1 + n_2 + . . . + n_k$ is the total number of observations (the total sample size across all groups), where n_i is sample size for the i^{th} group.

The parameters of this model are the population means $\mu_1, \mu_2, ...\mu_k$ and the common standard deviation δ (Fisher, 1925).

Statement of the Problem

It is widely believed that the applications of fertilizer increase the yield of crops. Researches have shown that this widely believed assertion is true since it has been proved beyond reasonable doubt scientifically. There are different types of fertilizer in use worldwide and the level of their applications depends on the discretion of the user and the type of soil. However, it has not be ascertained which of the fertilizer is the best among the fertilizer in terms of production output. This research intends to find out if the three fertilizers used here are the same in terms of production output or if there are differences among the fertilizers.

Objectives

(i) To estimate the data obtained.

(ii) To investigate if there are variability among the three fertilizers.

Research Question

Which of the fertilizers is more effective than the others in term of crops output or can we conclude that they are the same in crop yield output?

Hypothesis

H_o: $\mu_1 = \mu_2 = \mu_3 = \ldots \mu_k$

Against the alternative hypothesis

H_1: $\mu_1 \neq \mu_2 \neq \mu_3 \ldots \mu_k$ (there is at least one pair of unequal mean)

Using many separate two-sample t-test to compare many pairs of means is a bad idea because we don't get a p-value or a confidence level for the complete set of comparisons together (Dimas, 2012).

Materials and Methods

Let $\overline{X}_i$ represent the mean sample i(i = 1, 2, . . ., k):

$$\overline{X} = \frac{1}{n_i}\sum_{j=1}^{n_i} X_{ij}, \tag{1}$$

$\overleftarrow{X}$ Represent the grand mean, the mean of all the data points.

$$\overline{X} = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}, \tag{2}$$

S_i^2 Represent the sample variance:

$$S_i^{\,2} = \frac{1}{n_i - 1}\sum_{j=1}^{n_i}(X_{ij} - \overline{X})^2, \tag{3}$$

And s^2 = MSE is an estimate of the variance δ^2 common to all k populations.

$$S^2 = \frac{1}{N-K}\sum_{i=1}^{k}(n_i - 1).S_i{}^2. \quad (4)$$

ANOVA is concerned around the idea to compare the variation between groups (levels) and the variation within samples by analyzing their variances (Rafter, 2002).

Define the total sum of squares SST, sum of squares for error (or within groups) SSE, and the sum of squares for treatments (or between groups) SSC:

$$\text{SST} = \sum_{i=1}^{k}\sum_{j=1}^{nk}(n_{ij} - \overline{X})^2 \quad (5)$$

$$\text{SSE} = \sum_{i=1}^{k}\sum_{j=1}^{n}(X_{ij} - \overline{X}) = \sum_{i=1}^{k}(n_i - 1)S_i{}^2 \quad (6)$$

$$\text{SSC} = \sum_{i=1}^{n}\sum_{j=1}^{k}(\overline{X_i} - \overline{X})^2 = \sum_{i=1}^{k} n_i(\overline{X}_I - \overline{X})^2 \quad (7)$$

Consider the deviation from an observation to the grand mean written in the following way:

$$X_{ij} - \overline{X} = (X_{ij} - \overline{X_i}) + (\overline{X_i} - \overline{X}). \quad (8)$$

Notice that the left side is at the heart of SST, and the right side has the analogous pieces of SSE and SSC. It actually works out that:

$$\text{SST} = \text{SSE} + \text{SSC}. \quad (9)$$

The total mean sum of squares MST, the mean sums of squares for error MSE, and the mean sum of squares for treatment MSC are:

$$\text{MST} = \frac{SST}{df(SST)} = \frac{SST}{N-1} \quad (10)$$

$$\text{MSE} = \frac{SSE}{df(SSE)} = \frac{SSE}{N-K} \quad (11)$$

$$\text{MSC} = \frac{SSC}{df(SSC} = \frac{SSC}{K-1} \quad (12)$$

The one-way ANOVA, assuming the test conditions are satisfied uses the following test statistic:

$$\text{F} = \frac{MSC}{NSE} \quad (13)$$

Under H_o, this statistic has Fisher's distribution F (K-1, N-K). In case it holds for the test criteria:

$$\text{F} > F_{1\text{-a}}, \text{K-1, N-K}, \quad (14)$$

Where $F_{1\text{-a}}$, k-1, N-k is (1- α) quartile of F-distribution with K-1 and N-K degrees of freedom, then hypothesis H_o is rejected on significance level α (Aczel,,1989).

Results and Discussions

The results of the computations that lead to the F-statistic are presented in an ANOVA

table; the form of which is shown in table 1.1

Table 1.1 Basic one way ANOVA table

Source of variation	Sum of squares (SS)	degrees of Freedom	Mean Square MS	F-statistic	Tail area above F
Between	SSC	K-1	MSC	MSC/MSE	P-value
Within	SSE	N-K	MSE	;	-
Total	SST	N-1	-	;	-

Where K is the number of variables involved in the experiment and N is the total population (Shapiro & Wilk, 1965)

The P-value says the probability of rejecting the null hypothesis in case the null hypothesis holds, in case $p < \alpha$; where α is chosen significant level, is the null hypothesis rejected with probability greater than (1- α) 100% probability (Kruskal, 1952).The quantity of wheat harvested from each sub-plot are indicated in the table 2.

The results of the calculation obtained from table 1.2 are summarized in table 1.3.

Table 1.2: The quantity of wheat yield from each sub-plot

Plots	1	2	3	4	5
NPK	31	28	26	30	25
Urea	29	27	25	31	25
SSP	35	37	32	33	31

(Adewale et al., 2019)

Table 1.3 summary table of the one way ANOVA for the three types of fertilizer.

Variation above source	Sum of square (SS)	degrees of Freedom	Mean Square MS	F-statistic	Tail area
Between	116.93	2	58.47	9.18	0.005
Within	76.4	12.0	6.37	-	-
Total	193.33	14	-	-	-

Hypothesis testing: F_{α}, (K-1, N-K) df = $F_{.05}$ (2, 12) = 3.89 (Adewale et al., 2019)

Comparing the F-calculated with the critical or table value, we could see that F-calculated is greater than F-tabulated for this analysis. Therefore, we would reject H_0, the null hypothesis. What this means is that there are differences among the means for at least one of the three types of fertilizers.

Pair-Wise Multiple Comparison Tests

Ordinarily, the F-test can only show whether a difference exists among three or more means. It cannot reveal where the difference lies. When the null hypothesis is rejected using the F-test in ANOVA, we want to know where the difference among the means is. To

determine which pairs of means are significantly different, and which are not, we can use the pairwise multiple comparison tests and the best-established methods to determine where the differences lie are Turkey Kramer, Bonferroni and Dunn Sidak methods (Rykov, 2010) as shown in tables 4, 5 and 6 below:

Table 4: Results using Tukey Kramer method to detect the differences

Pairs I, j	difference	Lower limit	Upper limit
1.3	0.7776	-4.4565	4'8007
1.2	-6.0000	-8.3575	-1.7646
3.2	-6.7778	-8.6008	-2.6335

(Abubakar et al., 2018, McIntosh, 2015)

Table 5: Results using Bonferroni's method to detect the differences

Pairs I, j	Difference	Lower limit	Upper limit
1.3	0.7776	-4.5520	4.9210
1.2	-6.0000	-8.4552	-1.8624
3.2	-6.6667	-8.5425	-2.6532

(Abubakar et al., 2018, McIntosh, 2015)

Table 6: Results using Dunn-Sidak method to detect the differences

Pairs I, j	Difference	Lower limit	Upper limit
1.3	0.7776	-4.5543	4.9007
1.2	-6.0000	-8.4553	-1.8564
3.2	-6.6667	-8.5425	-2.6453

(Abubakar et al., 2018, McIntosh, 2015)

The three different methods for multiple comparisons include Tukey-Kramer, Bonferroni and Dunn-Sidak used above show similar result. The little discrepancies of the results is negligible. This indicates that any one of them is reliable to compute multiple comparisons. Using all the three multiple comparison methods (as shown above) we discovered that Urea, which is, fertilizer 2 takes significantly longer time to yield than fertilizers 1 and 3 which are similar. The result shows that the application of Urea fertilizer may results into low production output because the nutrients contain therein takes longer time to decay for crops to absolve the composition to produce well. For the Urea fertilizer to yield good results, it could take two to three years to have proper decay of the nutrients. However, when the first and third types of fertilizers are not affordable or available the Urea fertilizer can be applied.

From the analysis of the multiple comparisons, it is noticed that the first type of fertilizer, that is, NPK is more effective when compared with the third fertilizer, that is, SSP. To ensure bumper harvest the first type of fertilizer should be applied to wheat crops. The second type of fertilizer could be applied when the first and the third types are unavailable.

It is also advisable to use the first and second types at different intervals as this could improve the general production of crops. In some instances, the two could be mixed to gather and apply on crops at the same time.

Conclusion

The two-way analysis of variance is an extension of the one-way analysis of variance. Both are used to compare means of several groups or samples. Most of the statistical applications in social science, natural science, business administration, psychology etc., use comparison methods to compare two or more related groups. For hypothesis testing of more than two population means_statisticians have invented and developed ANOVA method. The technique is used to test for significant difference between class means, and this is done by analyzing the variances. The application of one-way or two-way analysis of variance (ANOVA) depends on the data available to the researcher and the objective(s) of the research. The researcher should be able to decide or choose which of the will suit his research to achieve the desired accurate result which will further lead to the appropriate final decision based on the outcome of the analysis. The ANOVA test procedure compares the variation in observations between samples (sum of squares for groups, SSC) to the variation within samples (sum of squares for error, SSE). The ANOVA F-test rejects the null hypothesis that the mean responses are equal in all groups if SSC is large relative to SSE.The analysis of variance (ANOVA) assumes that the observations are normally and independently distributed with the same variance for each treatment, factor or replication level (Montgomery, et al, 2003), if the normality assumption of the one-way ANOVA T-test is not met, we can use the Kruskal-Wallis (a non- parametric and assumption free test) rank test. The above analysis practically reveals all the requirements of normality, equal variance and other assumptions associated with analysis of variance. The researcher can confidently say that the objectives of the research have been achieved since it was observed that the three different fertilizers applied does not yield equal production output and therefore the null hypotheses (H_0) was rejected. This means that one of the fertilizers is more effective than the others as seen when the pairwise comparison was used after the rejection of the null hypothesis.

Recommendations

- The essence of undertaking statistical analysis is to discover new ways of improving the present from the past and to project to the future on the long-run.
- From the ongoing research, the null hypothesis was rejected indicating that the means of the three variables are not the same for at least one mean.
- The multiple comparisons made using Tukey method shows that the third type of fertilizer, that is, the local manure is not as effective as the other types of fertilizers. With this in mind the following recommendations are made.
- NPK fertilizer seems to be more effective than the rest of the fertilizers and therefore should be applied to millet crops.

- Urea fertilizer seems to be less effective than SSP but could also be applied with the two used together.
- The Urea can also be used when the other two types are not affordable or available.
- To achieve better results, the SSP could be used intermittently with the other two.
- To achieve better result two fertilizers can be combined or mixed and applied to the millet crop.
- Since type one fertilizer appears to be more productive from Tukey, Bonferroni and Dunn Sidak tests, any of the other two types can be used with it to improve the quantity of millet production.

References

Abubakar, H., Abubakar, H. D., & Salisu, A. (2018). One way anova: Concepts and application in agricultural system. *CEUR Workshop Proceedings*, *2152*(3), 49–53.

Adewale, A. A., Ibidunni, A. S., Atayero, A. A., John, S. N., Okesola, O., & Ominiabohs, R. R. (2019). Nigeria's preparedness for internet of everything: A survey dataset from the work-force population. *Data in Brief*, *23*, 103807.

McIntosh, M. S. (2015). Can analysis of variance be more significant? *Agronomy Journal*, *107*(2), 706–717.

Shapiro, S. S., & Wilk, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, *52*(3/4), 591.

Aczel, A.D., *Complete Business Statistics, Irwin,* 1989.

Barnas, J.A., Manaret, A.B., Jenabib, A.K., (2010). *"Analysis of variance. A critical Approach"* Vol. 18, pp. 235-360.

Brown, M.B., & Forsyth, A.B., (2012). *Robust tests for the equality of Variances*. Journal of the American Statistical Association. 69: 364-367.

Dimas, A.T., Patrina D.C., & Johan, W.C., (2012). "*Statistical Methods and Scientific Approach to Design of experiment.* Pages (375-425).

Eva, O., & Oskar, O. (2010). Department of mathematics and Theoretical Informatics*, Faculty of Electrical Engineering and Informatics,* Technical University of Kosice, Nen covej 32, 04200Kosice, Slovac republic.

Fisher, R.A., (1925). *Statistical Methods for Research Workers;* Edinburgh, United Kingdom: Oliver& Boyd.

Kruskal, W.H., & Wallis, W.A., (1952). Use of ranks in one-criterion analysis of variance. *J Am Stat Assoc.* 1952' 47: 583-621.

Montgomery, D.C., & Runger, G.C., (2003). *Applied Statistics and Probability for Engineers,* John Wiley & Sons.

Rafter, J.A., ABELL, M. L., & Braselton, J. P., (2002). "*Multiple Comparison Methods for Means,* "SIAM Review, 44 (2), 259-278.

Rykov, V.V., Balakrishman, N., & Nikulin, M., S., (2010). *Mathematical and Statistical Models and Methods in Reliability,* Springer.

Scot, M. Flaherty, D., &UCurally, J. (2013). Statistics: *Dealing with categorical Data.* J. Small Anim Piae. 54: 3-8.

Wackerly, D.D., Mendenhall, W. 111, & Scheaffer R.L. (2002). *Mathematical Statistics with Applications. 6th ed.* Pacific Grove, CA: Druxbury Press.

Wilk, M.B, & Farakas, H. M., (1965). *An Analysis of Variance (ANOVA) test for Normality*