

IDENTIFICATION OF SIGNIFICANT TREATMENT EFFECTS IN UNREPLICATED FACTORIAL EXPERIMENT

IBRAHIM, B. A, ADELEKE, B. L, OYEYEMI, G. M
Department of Statistics
University of Ilorin
Ilorin, Nigeria

ABSTRACT

Unreplicated factorial designs are used in industrial and other real life experiments due to costs consideration or some other reasons. Normal plot has been an important technique of identifying significant treatment effects. However, two other methods of identifying significant effects namely, Lenth and Step-Down Lenth methods are examined in this paper. The three methods are compared in terms of their performance in identifying significant effects by simulating unreplicated 2^4 factorial designs at two different settings. In evaluating the three methods, we observed that Normal plot has the overall best performance, but the three methods are similar when the magnitude of each of the effects is high.

Key words: Pseudo-Standard Error; Significant Effect; Treatment Contrast, Unreplicated Factorial Experiment

1 INTRODUCTION

Experimentation is one of the most common activities that people engage in. It covers a wide range of applications from household work like food preparation to

technological innovation in material, science, agriculture, and so on. Experimentation is used to understand and/or improve a system. It allows an investigator to find out what happens to output or response when the settings of the input variables in the system are purposely changed (Hohn 1984, Danohue 1994).

Statistical design of experiment was developed by R.A. Fisher at Rothamsted agricultural experiment station, London, England in 1930 (Fisher 1935) and it has come to play a vital role in many industries and organizations in terms of improving process efficiency, higher quality and reducing process variability, cost of poor quality and other costs. Another important use of design of experiment is in screening, in which experimenters seek to identify significant effects (both main effects and potentially interaction effects) from a large set of possible effects.

2 FACTORIAL DESIGNS

The common design of experiments includes; complete randomized design (CRD), complete randomized block design (CRBD), Lattice Designs, Youden Square Designs; and any one of them may be considered for either simple or factorial treatment structure. Factorial Design, that is a design with factorial treatment structure, is most widely used, in varied industrial applications, in some other areas like agricultural and other fields of human endeavors. A factorial experiment written, as F^K is an experiment involving K factors each at F levels.

A full factorial experiment is obtained when all the possible combinations of the factors effects are run in a single experiment and we have a fractional factorial experiment when a part or fraction of all the combinations are considered or run.

Un-replicated factorial designs are frequently used in industrial experiment in order to cut cost or due to some other important reasons (Hunter 1977). In many industrial settings, replication is sacrificed for run size and this can present serious

difficulties in the analysis of such design. Replication is the allocation of treatment to a number of units that are representative of the population.

Replication enables us to estimate the experimental error and as well increase the power to detect important effects, by decreasing the variance of the treatment effect estimates. When there is no estimate for the experimental error, the higher order interactions are often sacrificed for the estimate of error term, which are then used for computing the required statistics in testing for the significance of the design factors.

3 FACTOR EFFECT ESTIMATES

A very simple way of calculating effect estimates in 2^k factorial design is the technique devised by Yates (Reynar 1967) and it is called Yates' algorithm or Yates' method. It is also used in determining the sum of squares for the treatment effects. The procedure can sometimes be done manually if the value of k is not large. (Montgomery, 2005)

It has been pointed out that in the regression model representation of the 2^k factorial experiment, the regression coefficients are exactly one-half the effect estimates. Therefore it is straight forward to show that the model coefficients (and hence the effect estimates) are the least square estimates.

In particular for 2^2 factorial designs, the regression model is given as;

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_{12} x_{i1} x_{i2} + e_i \quad i = 1, 2, 3, 4 \quad \dots \dots \dots (1)$$

The least square estimates of the model parameters in Equation (1) are chosen to minimize the sum of squares of the error terms (L), which is expressed as follows

$$L = \sum_{i=1}^4 (y_i - \beta_0 - \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_{12} x_{i1} x_{i2})^2 \quad \dots \dots \dots (2)$$

The normal equations arising from Equation (2) are given below

$$4\hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^4 x_{i1} + \hat{\beta}_2 \sum_{i=1}^4 x_{i2} + \hat{\beta}_{12} \sum_{i=1}^4 x_{i1} x_{i2} = (1) + a + b + ab$$

$$\hat{\beta}_0 \sum_{i=1}^4 x_{i1} + \hat{\beta}_1 \sum_{i=1}^4 x_{i1}^2 + \hat{\beta}_2 \sum_{i=1}^4 x_{i1}x_{i2} + \hat{\beta}_{12} \sum_{i=1}^4 x_{i1}^2x_{i2} = -(1) + a - b + ab \quad \dots\dots\dots (3)$$

$$\hat{\beta}_0 \sum_{i=1}^4 x_{i2} + \hat{\beta}_1 \sum_{i=1}^4 x_{i1}x_{i2} + \hat{\beta}_2 \sum_{i=1}^4 x_{i2}^2 + \hat{\beta}_{12} \sum_{i=1}^4 x_{i1}x_{i2}^2 = -(1) - a + b + ab$$

$$\hat{\beta}_0 \sum_{i=1}^4 x_{i1}x_{i2} + \hat{\beta}_1 \sum_{i=1}^4 x_{i1}^2x_{i2} + \hat{\beta}_2 \sum_{i=1}^4 x_{i1}x_{i2}^2 + \hat{\beta}_{12} \sum_{i=1}^4 x_{i1}^2x_{i2}^2 = (1) - a - b + ab$$

The expressions on the right hand side (RHS) of the Equations (3) are the contrasts of the factor effects, and since $\sum_{i=1}^4 x_{i1} = \sum_{i=1}^4 x_{i2} = \sum_{i=1}^4 x_{i1}x_{i2} = \sum_{i=1}^4 x_{i1}^2x_{i2} = \sum_{i=1}^4 x_{i1}x_{i2}^2 = 0$ because the design is orthogonal, the normal equations reduce to a very simple form as;

$$\begin{aligned} 4\hat{\beta}_0 &= (1) + a + b + ab \\ 4\hat{\beta}_1 &= -(1) + a - b + ab \quad \dots\dots\dots (4) \\ 4\hat{\beta}_2 &= -(1) - a + b + ab \\ 4\hat{\beta}_{12} &= (1) - a - b + ab \end{aligned}$$

Therefore;

$$\begin{aligned} \hat{\beta}_0 &= \frac{(1) + a + b + ab}{4} \\ \hat{\beta}_1 &= \frac{-(1) + a - b + ab}{4} \quad \dots\dots\dots (5) \\ \hat{\beta}_2 &= \frac{-(1) - a + b + ab}{4} \\ \hat{\beta}_{12} &= \frac{(1) - a - b + ab}{4} \end{aligned}$$

These regression model coefficients are exactly one-half the factor effects estimates.

4 METHODS OF IDENTIFICATION OF SIGNIFICANT TREATMENT EFFECTS

In what follows we shall consider three methods namely, Half-Normal Plot, Lenth and Step-Down Lenth methods. These methods are indeed suitable for identification of significant treatment effects

4.1 HALF-NORMAL PLOT.

When a factorial experiment is not replicated, the common method of analyzing the design in order to identify active or significant effects from such experiment is by making use of half-normal plot. (Daniel 1959). Under the assumption that the values originate from normal population with mean zero, the plotted absolute value of estimated effects should lie approximately on a straight line. Effect that are lying far off the line are said to be significant. This method work nicely when

- i.) There are not many significant effects
- ii.) When all effect estimates have the same variance.

It turns out that all contrasts computed from a 2^K design (and hence all effects estimate) have the same variance even if the individual observations have different variances (Montgomery, 2005.)

The graphical method entails plotting on normal probability paper;

$$\Phi^{-1}(0.5 + 0.5(i + 0.5)/(n - 1)) \text{ Against } |\hat{\beta}_i| \text{ for } i = 1, 2, \dots, n$$

Where $|\hat{\beta}_i|$ are the ordered absolute effects, and $\Phi(\bullet)$ is the cumulative standard normal distribution. Assuming that all β_i are zero, the resulting plot should be linear apart from some random variations. Any point that falls off the line is declared significant.

A normal probability plot can also be constructed on ordinary graph paper in the absence of normal probability plotting paper. This is done by plotting standardized

normal z-scores $z_{(j)}$ against the ranked contrast effects, where the standardized normals z-scores are obtained by utilizing the expression given as Equation (6) below.

$$p(Z \leq z_j) = \Phi(z_j) = \frac{j-0.5}{n} \dots\dots\dots 6$$

For example, if $(j-0.5)/n = 0.05$, then $\Phi(z_j) = 0.05$ implies that $z_j = -1.64$.

Meanwhile, many computer/Statistical packages present the normal probability plot this way.

One of the drawbacks or limitations of using half normal plot in identifying significant effect is it's subjectively. Also if the effects contrasts are so close, this method may prove inefficient in identifying significant factors. Also, the efficiency of normal plot depends on the number of significant effects among the total effects under consideration.

4.2 LENTH METHOD

To overcome the subjectivity and other problems of using half normal plot, many formal testing methods have been proposed among which is Lenth method (Hamada and Balakrishnan, 1998)

Let $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k$ denote k mutually orthogonal estimated factorial effects (that is, contrasts of factorial design). Assuming that there are only a few active effects, Lenth (1989) uses Pseudo Standard Error (PSE) to estimate the standard deviation of $\hat{\theta}_i$, $i = 1, 2, \dots, k$

$$PSE = 1.5 \bullet median_{|\hat{\theta}_i| < 2.5 \bullet S_0} |\hat{\theta}_i|$$

$$\text{where } S_0 = 1.5 \bullet median |\hat{\theta}_i|$$

Lenth (1988) then calculated t-like statistic by dividing the $\hat{\theta}_i$ by PSE, which he referred to as Lenth statistics, and expressed as Equation (7).

$$t_{\text{Lenth}, i} = \frac{\hat{\theta}_i}{PSE} \dots\dots\dots (7)$$

He suggested a t-distribution with $k/3$ degrees of freedom as an approximate to the t_{Lenth} reference distribution for controlling the individual error rate. Any effect that has its t_{Lenth} greater than the critical value is declared to be significant; otherwise such effect is not significant.

4.3 STEP-DOWN LENTH METHOD

Voss (1989) proposes a step-down lenth method with little modification on original lenth method. Let $|\hat{\theta}|_{(1)} \leq |\hat{\theta}|_{(2)} \leq \dots \leq |\hat{\theta}|_{(k)}$ be the order statistics of the k absolute contrasts.

The corresponding test statistic is expressed as Equation (8) below:

$$t_i = \frac{|\hat{\theta}|_{(i)}}{PSE_i} \dots\dots\dots (8)$$

Where PSE_i is the Pseudo Standard Error of $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_k$ the signed contrast corresponding to the absolute contrasts $|\hat{\theta}|_{(1)} \leq |\hat{\theta}|_{(2)} \leq \dots \leq |\hat{\theta}|_{(k)}$. If we denote c'_α to be the critical value at the significant level α and using the original Lenth method with i th contrasts. If the first largest contrast is declared significant, the PSE_i is recalculated for the reaming $k-1$ contrasts and the second largest contrast is tested for significance. This procedure continues until no contrast is found to be significant.

5 SIMULATION EXPERIMENT

In order to evaluate the performance of the three methods of analyzing unreplicated factorial design which are described in section 4 above, a 2^4 factorial designs will be used. As a consequence of this 15 distinct contrast effects are to be estimated. The 15 contrast effects are simulated using the following plans;

For $C = 1$: $\theta_1 = \theta_2 = \theta_3 = \dots = \theta_{14} = 0$, $\theta_{15} = \Delta$

For $C = 2$: $\theta_1 = \theta_2 = \theta_3 = \dots = \theta_{12} = 0$, $\theta_{13} = \theta_{14} = \theta_{15} = \Delta$

For $C = 3$: $\theta_1 = \theta_2 = \theta_3 = \dots = \theta_{10} = 0$, $\theta_{11} = \theta_{12} = \theta_{13} = \theta_{14} = \theta_{15} = \Delta$

For $C = 4$: $\theta_1 = \theta_2 = \theta_3 = \dots = \theta_8 = 0$, $\theta_9 = \theta_{10} = \theta_{11} = \theta_{12} = \theta_{13} = \theta_{14} = \theta_{15} = \Delta$

where Δ ranges from 1 to 8. For each of the possible values of C ($C = 1, 2, 3, 4$), ten sets of 15 contrast effects were simulated. Result of simulation experiment for $\Delta = 1, 2, 5, 6$ for $C = 1$ only are shown in the Appendix. The number of significant contrasts vary with the values of C ; in that when $C = 1$, we have only one significant contrast, when $C = 2$, we have three significant contrasts, while for $C = 3$ and $C = 4$ we have five and seven significant contrasts respectively.

6 EVALUATION OF THE METHODS

The effectiveness of correctly detecting significant contrasts from the simulated contrasts for the three methods under consideration was evaluated. This is done by obtaining the relative frequencies (or equivalently the probabilities) of correctly detecting significant contrasts. The results of the evaluation are presented in Tables 1 through 4 for the three methods for $C = 1$, $C = 2$, $C = 3$ and $C = 4$. Graphical presentations of the results of evaluation of the three methods are given in Figures 1 through 4, by plotting the relative frequencies against Δ .

Table 1. The probabilities of the three methods when $C = 1$

Delta	Original Lenth	Step-Down Lenth	Normal plot
1	0.10	0.10	0.38
2	0.30	0.30	0.60
3	0.40	0.40	0.62
4	0.50	0.50	0.70
5	0.90	0.90	1.00
6	0.90	0.90	1.00
7	1.00	1.00	1.00
8	1.00	1.00	1.00

Table 2. The probabilities of the three methods when $C = 2$

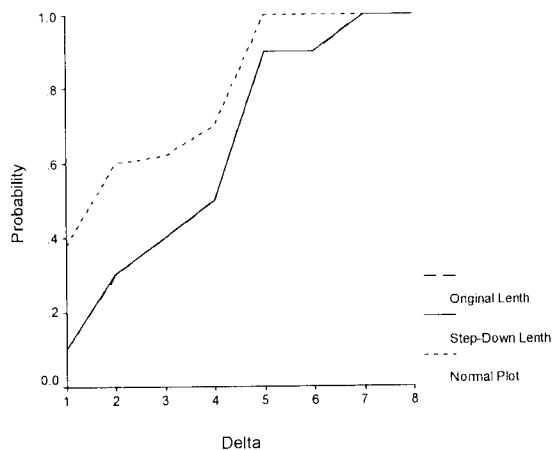
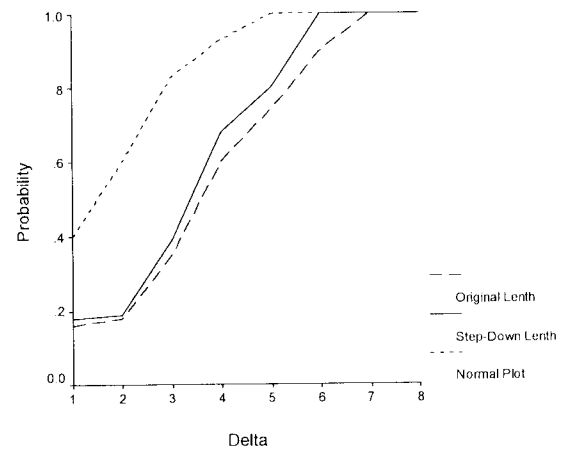
Delta	Original Lenth	Step-Down Lenth	Normal plot
1	0.16	0.18	0.40
2	0.18	0.19	0.60
3	0.35	0.39	0.83
4	0.60	0.68	0.93
5	0.74	0.80	1.00
6	0.90	1.00	1.00
7	1.00	1.00	1.00
8	1.00	1.00	1.00

Table 3. The probabilities of the three methods when $C = 3$

Delta	Original Lenth	Step-Down Lenth	Normal plot
1	0.18	0.20	0.40
2	0.22	0.30	0.72
3	0.45	0.54	0.88
4	0.58	0.68	0.98
5	0.66	0.86	1.00
6	0.90	0.94	1.00
7	1.00	1.98	1.00
8	1.00	1.00	1.00

Table 4. The probabilities of the three methods when $C = 4$

Delta	Original Lenth	Step-Down Lenth	Normal plot
1	0.25	0.31	0.42
2	0.38	0.42	0.72
3	0.45	0.58	0.84
4	0.60	0.69	0.96
5	0.70	0.78	1.00
6	0.95	0.98	1.00
7	1.00	1.00	1.00
8	1.00	1.00	1.00

Fig 1. Probability graph when $C = 1$ Fig 2. Probability graph when $C = 2$ Fig 3. Probability graph when $C = 3$ 

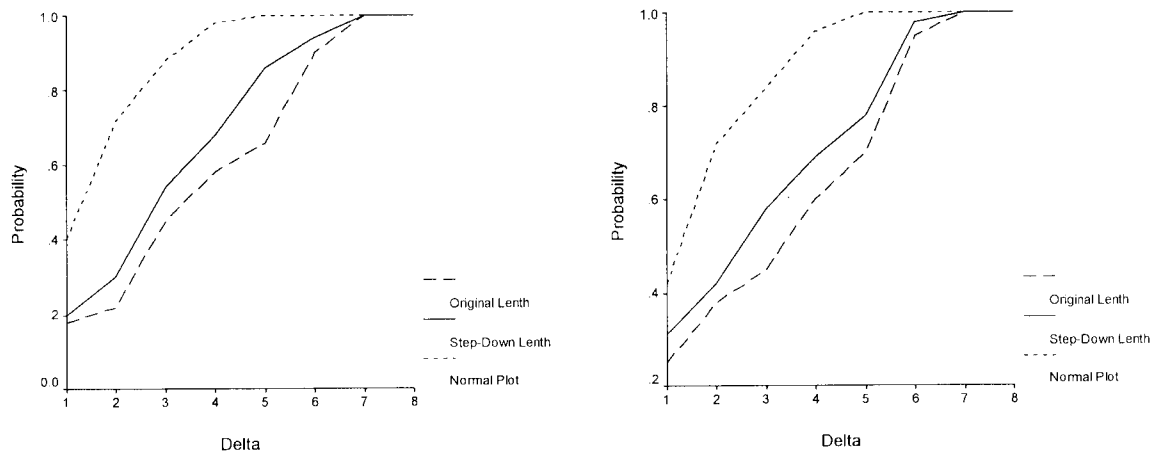


Fig 4. Probability graph when $C = 4$

7 DISCUSSION OF RESULTS AND CONCLUSIONS

Half-Normal Plot, Lenth and Step-Down Lenth methods were evaluated by obtaining the relative frequency of correctly identifying true significant effect and the results are presented in Tables 1 through 4. Interestingly, all the methods improve in performance when the magnitude of the treatment effects is increased, though in all the cases, Normal plot has the overall best performance, followed by Step-Down Lenth with the Original Lenth coming last.

Graphical presentation of the probability of correctly detecting significant treatment contrasts in Figures 1 through 4 present visually, the relative performance of the three methods, with the Normal plot method being the overall best.

In conclusion Normal plot is superior to the other two methods presented in this paper, namely, Original Lenth and Step-Down Lenth. However, the Normal plot method unlike the Step-Down Lenth and Original Lenth, has some inherent shortcomings in terms of subjectivity and non-existence of rigorous statistical tests.

REFERENCE

- Daniels, C. (1959).** "Use of Half-Normal Plots in interpreting Factorial Two-level Experiments". *Technometrics* 1. pp 311 - 340.
- Danohue, J. M. (1994).** "Experimental Designs for Simulation". *Proceedings of the Winter Simulation Conference*, pp 200 – 206.
- Fisher (1935)

-
- Hamada, M. and Balakrishnan, N. (1998).** "Analyzing Unreplicated Factorial Experiments: A review with some New Proposals". *Statistica Sinica* 8. pp1 – 41
- Hohn, G. J. (1984).** "Experimental Design in a Complex World". *Technometrics*, vol. 26, pp 19 – 31
- Hunter, W. G. (1977).** "Some ideas About Teaching Design of Experiments" *The American Statistician*, vol. 31, pp 12 – 17
- Lenth, R. V. (1989).** "Quick and Easy Analysis of Unreplicated Factorials". *Technometrics* 31, pp 469 – 47
- Montgomery, D. C. (2005).** *Design and Analysis of Experiments*. Wiley, New York, NY.
- Rayner, A. A. (1967).** "The Square summing check on the Main Effects and Interaction in a 2^k Experiments as calculated by Yates' algorithm". *Biometrics*, vol. 23, pp 571 – 573
- STATA8 Manual.** www.stata.com.
- Voss, D. T. (1988).** "Generalized Modulus-Ratio Tests for Analysis of Factorial Designs with Zero Degrees of Freedom for Error". *Communication in Statistics-Theory and Methods* 17, pp. 3345–3359.