

ANALYZING FACTORIAL EXPERIMENT INVOLVING QUALITATIVE AND QUANTITATIVE FACTORS

¹Oyeyemi, G.M., ²Ibraheem, B.A. ³Obafemi, O. S. and ³Ige, S. O.

¹Department of Statistics, University of Ilorin, Ilorin, Nigeria.

²Department of Statistics, Kwara State polytechnic, Ilorin, Nigeria.

³Department of Mathematics and Computer Studies, Federal Polytechnic, Ado-Ekiti, Nigeria.

ABSTRACT: *In factorial experiment involving qualitative and quantitative factors, ascertaining whether there is interaction between the two factors should be the first objective in the analysis. In this paper, we explore the different possibilities that may arise when dealing with qualitative and quantitative factors in a factorial experiment. The response curves for the quantitative factor are different across the levels of the qualitative factor when there is interaction between the factors, the curves are parallel but similar with distance apart when there is no interaction but both factors are significant, while curves converged to a curve (become a single curve) when the quantitative factor is not significant.*

Key words; Qualitative factor, curve, factorial design, interaction.

INTRODUCTION

Design of experiment 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 Kempthorne, (1952). Objectives of experimental design include treatments comparisons, variables selection, response surface exploration, system optimization and system robustness, Danohue (1994). Experimentation is also 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).

Statistical design of experiment was developed by R.A. Fisher at Rothamsted agricultural experiment station, London, England in 1930 and it has come to play a vital role in many industrial and organization in term of improving process efficiency, higher quality and reducing process variability, cost of poor quality and other cost. Another important use of design of experiment is in screening, in which experimenters seek to identify significance effect (both main effects and potentially interaction effect) from a large set of candidate of effects, Montgomery (2005).

QUALITATIVE AND QUANTITATIVE FACTORS

In an experiment a factor is a variable of interest to the researcher and it is normally at different levels (at least two levels). A quantitative factor in experimental design has its level measurable on a numerical scale like temperature, density, weight and so on. The magnitude of a quantitative factor is usually used to determine the levels. Qualitative factor is not measurable but can only be categorized into classes, examples of qualitative factors are operators, cities and so on.

Assuming we have a 3^2 factorial experiment in which one factor is quantitative and the other is qualitative. Ascertaining whether there is interaction between the quantitative and qualitative factor is the first objective in the analysis. There are three possible outcomes we may consider;

1. There is interaction between the factors.
2. There is no interaction between the factors.
3. There is no interaction between the factors but the main effect of the qualitative factor is also not significant.

This paper is aimed at exploring the above mentioned possibilities when dealing with both qualitative and quantitative factors in an experiment and look at trend patterns of the qualitative factor.

EXAMPLES

A soil scientist hypothesized that the concentration of certain zinc in sewage sludge would differ among the metropolitan areas from which the sludge was obtained. He obtained sewage sludge from treatment plants located in three different metropolitan cities: A, B, and C. Corn plants were grown in sand medium to which sludge was added as fertilizer. The sludge was added to the sand at three different rates: 0.5, 1.0, and 1.5 metric tons/acre. The factorial arrangement for the treatment design consisted

of one qualitative factor, "City" with three levels and a quantitative factor, "Rate", also at three levels given 9 treatments. Each of the nine treatments was randomly assigned to four containers in a complete randomized manner to give 4 replicates and at a certain stage of growth, the zinc content in parts per million was determined for the corn grown in each of the containers.

Three sets of data were simulated from the experiment to illustrate the three possible outcomes earlier mentioned above. Table 1 gives a situation where both factors as well as their interaction are significant, the results in Table 2 are simulated such that the two factors are significant but their interaction is not significant while Table 3 gives situation where only the qualitative factor (Rate) is significant

Table 1. Both factors and their interaction are significant

City and Rate (MT/hectare)									
City	A			B			C		
Rate	0.5	1.0	1.5	0.5	1.0	1.5	0.5	1.0	1.5
	26.4	25.2	26.0	30.1	47.7	73.8	19.4	23.2	18.9
	23.5	39.2	44.6	31.0	39.1	71.1	19.3	21.3	19.8
	25.4	25.5	35.5	30.8	55.3	68.4	18.7	19.6	19.6
	22.9	31.9	38.6	32.8	50.7	77.1	19.0	19.9	21.9

Table 2. Both factors are significant but their interaction not significant

City and Rate (MT/hectare)									
City	A			B			C		
Rate	0.5	1.0	1.5	0.5	1.0	1.5	0.5	1.0	1.5
	34.4	43.4	52.4	29.3	38.3	47.3	24.3	33.3	42.3
	38.0	47.0	56.1	32.9	42.0	50.9	28.0	37.0	46.0
	37.5	46.4	55.5	32.4	41.4	50.4	27.4	36.0	45.1
	37.5	46.7	55.8	32.7	41.7	50.7	27.7	36.7	45.0

Table 3. Only the qualitative factor is significant

City and Rate (MT/hectare)									
City	A			B			C		
Rate	0.5	1.0	1.5	0.5	1.0	1.5	0.5	1.0	1.5
	39.4	48.4	57.4	39.3	48.3	57.3	39.4	48.3	57.3
	43.0	52.0	61.0	43.0	52.0	61.0	42.9	52.0	60.9
	42.5	51.5	60.5	42.4	51.0	60.4	42.4	51.9	60.5
	42.8	61.8	60.7	42.7	51.1	60.7	42.0	51.7	60.7

ESTIMATION OF LINEAR AND QUADRATIC EFFECTS

Since there is a quantitative factor in the design we need to test for the significance of such factor both at linear and quadratic levels. We therefore define two contrasts C_1 and C_2 for linear and quadratic effects respectively,

where $C_1 = [-1 \ 0 \ 1]$ and $C_2 = [1 \ -2 \ 1]$ and their sum of squares for the possible contrasts are given as follows;

$$SS[\text{Rate Linear}] = ra \left[\sum_j C_{1j} \bar{y}_{.j} \right]^2 / \sum_j C_{1j}^2$$

$$SS[\text{Rate Quadratic}] = ra \left[\sum_j C_{2j} \bar{y}_{.j} \right]^2 / \sum_j C_{2j}^2$$

$$SS[\text{Rate Linear by City}] = r \sum_i \left[\sum_j C_{1j} \bar{y}_{.j} \right]^2 / \sum_j C_{1j}^2 - SS[\text{Rate linear}]$$

$$SS[\text{Rate Quadratic by City}] = r \sum_j \left[\sum_j C_{2j} \bar{y}_{.j} \right]^2 / \sum_j C_{2j}^2 - SS[\text{Rate Quadratic}]$$

$i = 1, 2, 3; j = 1, 2, 3; \text{ and } r = 1, 2, 3, 4.$

ANALYSIS

The data were analyzed using SAS Code "Proc GLM". The first GLM runs the usual ANOVA for the two factors and their interaction while the second GLM runs the factorial design to obtain the sum of squares for linear, quadratic and interaction effects of the qualitative and quantitative factors.

The SAS Code for the data analysis;
data new;
array Y Y1-Y4;
INPUT F1 \$ F2 \$ TRT \$ Y1-Y4;
do over Y;
X=Y;
output;

```
end;
drop Y1-Y4;
label F1= 'CITY' F2= 'RATE' X='ZINC-CONC'
TRT='CITY-RATE';
cards;
A 0.5 A5 26.4 23.5 25.4 22.9
....
....;
proc glm;
class F1 F2;
model X = F1 F2 F1*F2/SS3;
means F1 F2 F1*F2;
output out=ASSUMP r=RESID p=MEANS;
proc GLM;
CLASS TRT;
MODEL X=TRT/SS3;
CONTRAST 'LIN-A'
CONTRAST 'LIN-B'
CONTRAST 'LIN-C'
CONTRAST 'QUAD-A'
CONTRAST 'QUAD-B'
CONTRAST 'QUAD-C'
CONTRAST 'City-A'

CONTRAST 'City-B'

CONTRAST 'City-C'

End.
```

Table 4. Results when both factors and their interaction are significant

Source	DF	Squares	Mean Square	F Value	Pr > F
Model	8	9475.515	1184.4394	61.75	<.0001
Error	27	517.865	19.1802		
Corrected Total	35	9993.38			

Type III					
Source	DF	SS	Mean Square	F Value	Pr > F
F1	2	5720.672	2860.3358	149.13	<.0001
F2	2	1945.445	972.7225	50.71	<.0001
F1*F2	4	1809.398	452.3496	23.58	<.0001

Dependent Variable: X ZINC-CONC

Contrast					
Contrast	DF	SS	Mean Square	F Value	Pr > F
LIN-A	1	270.2813	270.2813	14.09	0.0008
LIN-B	1	3432.061	3432.0613	178.94	<.0001
LIN-C	1	1.805	1.805	0.09	0.7614
QUAD-A	1	0.0204	0.0204	0	0.9742
QUAD-B	1	36.2604	36.2604	1.89	0.1804
QUAD-C	1	14.415	14.415	0.75	0.3936
City-A	2	270.3017	135.1508	7	0.0034
City-B	2	3468.322	1734.1608	90.41	<.0001
City-C	2	16.22	8.11	0.42	0.6595

Table 5. Results for both linear and quadratic contrasts for the two factors

Dependent Variable: X ZINC-CONC

Contrast	DF	ContrastSS	Mean Square	F Value	Pr > F
LIN-A	1	270.2813	270.2813	14.0900	0.0008
LIN-B	1	3432.0610	3432.0613	178.9400	<.0001
LIN-C	1	1.8050	1.8050	0.09	0.7614
QUAD-A	1	0.0204	0.0204	0.00	0.9742
QUAD-B	1	36.2604	36.2604	1.8900	0.1804
QUAD-C	1	14.4150	14.4150	0.7500	0.3936
City-A	2	70.3017	135.1508	7.05	0.0034
City-B	2	468.3217	1734.1608	90.41	<.0001
City-C	2	6.2200	8.1100	0.42	0.6595

RESULTS AND CONCLUSION

The SAS outputs of the first case were shown in Tables 4 and 5, while the response curves (Trend) for the three cases were shown in figures 1 – 3. In all the cases the quadratic effects were not significant while the linear effects were significant. In the first case, the responses to the quantitative factor were found to be different across the levels of the qualitative factor, therefore, the response curves for the quantitative factor must be estimated separately for each level of qualitative factor (fig.1). Parallel lines that are shifted from one another will describe the response curves of the quantitative factor in case there is no interaction between the factors but when both factors are significant (fig.2). A single response curve is good enough to describe the outcome when there is no interaction and the main effect of the qualitative factor is not significant (fig.3).

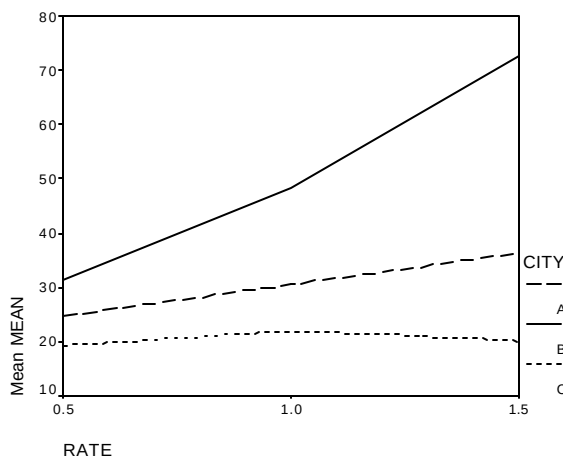


Figure 1. Response curves when there is interaction between the two factors.

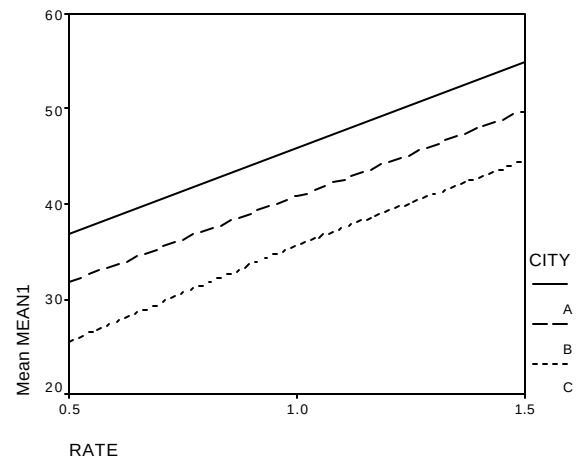


Figure 2. Response curves when there is no interaction between the factors but the two factors are significant

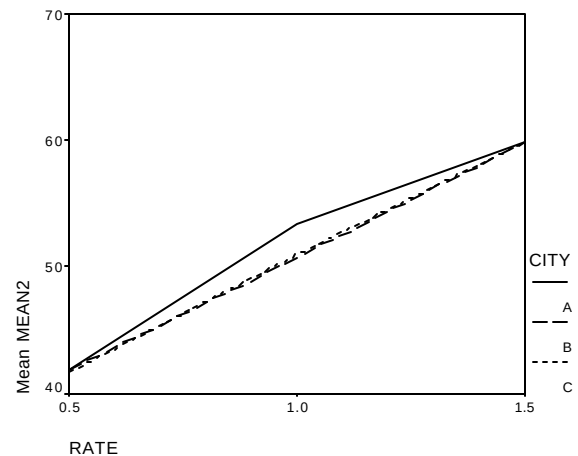


Figure 3. Response curves when only quantitative factor is significant

REFERENCE

- Danohue, J. M. (1994). Experimental Designs for Simulation. Proceedings of the Winter Simulation Conference, pp 200 – 206.
- Kempthorne, O. (1952). The Design and Analysis of Experiments. John Wiley. New York. NY, pp 234-251
- Hohn, G. J. (1984). Experimental Design in a Complex World. Technometrics, vol. 26, pp 19 – 31
- Montgomery, D. C. (2005). Design and Analysis of Experiments. Wiley, New York, NY, pp, 1-8