ASSET Series C (2007) 2 (1): 56-66

ASSET
An International Journal

A COMPARATIVE ANALYSIS OF WHEAT YIELD USING THE RANGE-CONTROL ANALYSIS AND THE LATIN- SQUARE DESIGN

ABRAHAM BISI ALABI-LABAIKA

Department of Business Administration
University of Lagos, Yaba, Akoka, Lagos State, Nigeria
e-mail: bisilabaika@yahoo.com
GSM 08077525261

ABSTRACT

Having looked at the use of Latin-Square Design in analyzing agricultural data, it is viewed that a relatively simple alternative can come from statistical quality control. The data set considered in this paper is for wheat yield resulting from a Latin-Square Design. Statistical quality control is basically meant for analyzing data on manufactured products. The data- presentation format for both wheat yield and manufacture data can be the same if the researcher so chooses. One of the possibilities is the data- presentation format for range- control analysis [where the sample size ,n, can be made equal to the number of samples, m] and that of the Latin -Square Design where the number of rows is equal to the number of columns is equal to the number of treatments. It is possible to make all these variables equal for the two methods, and this is what has been done in this paper. The other imposed condition is that if the sample means are significantly different, then, they are deemed collectively effective. Range-control analysis is a relatively simple method among the methods used in statistical quality control to determine whether or not the manufactured items meet a pre-manufacture set standard. For reasons of simplicity, this quality control method has been proposed as a possible alternative to the Latin-Square Design. The results of the tests conducted using the range- control analysis and the analysis of variance for the Latin –Square design lead to the same statistical conclusions: the effect of rows and columns on the wheat yield is not significant, but the effect of the treatments significantly influenced the wheat yield. Hence, it is concluded that the two methods used in the analysis are good alternatives.

Key words: range-control limits, statistical quality control, pre-manufacture set standards, imposed conditions.

INTRODUCTION

According to Brookes and Dick (1969), experiments are carried out to test the validity of a hypothesis or to estimate the magnitude of an effect. Consider, for example, an experiment in which the effects of five various fertilizers on the yield of crops is to be tested. The field on which the experiment is carried out is divided into twenty five similar plots comprising a five-by-five matrix, which are then treated with fertilizers A, B, C, D and E. Each fertilizer is used only once in each row and in each column. The data format is always a row-by-column and in the case of Latin squares, the number of rows equals the number of columns which is also equal to the number of treatments.

The Latin Square allows for testing for the

significant effect of the rows, the columns and the treatments. This test of significance is usually carried out using the analysis of Variance (ANOVA) by the F-test. When treatment effect has been found to be significant using the ANOVA, then, Duncan's Multiple Range Test is used to determi9ne treatments that are significantly different from one another (Adigun, 2002; Adigun, et al., 2003). Androulidakis et al. (1996) used Kendall's coefficient of concordance (W) to similarly analyze agriculture data. Parsimony requires that out of two or more alternative models the simplest should be preferred.

The aim of this paper, then, is to use the range-control limits analysis (RCA to test for the significance or otherwise of the effects of the rows, the columns and the treatments instead of using the ANOVA used by Moroney (1976).

The RCA is found in statistical quality control: (quality control, for short). Quality control is the application of statistical procedures to determine if a stable system of random effects is present when a process or a product is tested against a quality standard (Blank, 1980).

Further, quality control is the application of hypothesis-testing procedure each time a sample or samples are taken for investigation. To be absolutely certain that all items being produced by a manufacturer meet the specification laid down in advance and hence may be delivered to customers, it is necessary to inspect each one. If, however, only samples of these items are inspected, there is always a level of uncertainty that those ones not inspected but which are a part of a passed batch will not be suitable. The investigation and reduction of this level of uncertainty by using simple random sampling without replacement and probability is what is loosely called statistical quality control.

In this connection, after considering cost implications, samples of equal and small sizes, each size usually being less than or equal to ten, in most cases, each time, are taken for cases of variable-testing, small or large samples also of equal sizes may be taken for attribute-testing. Acceptance or rejection sampling may also require a small or large sample. There is no fixed relationship between the population size and the sample size in range- control analysis; the principle is that there is massive production of units having an infinitely large population out of which small samples are taken without replacement.

After samples have been taken, an appropriate control chart is then used in accordance with the quality-control technique being used. Some of the control charts include those of acceptance requiring single or double or multiple sampling plans, such as the Dodge-and-Rominq sampling scheme (Moroney, 1976), the proportion or fraction defective or the P-chart (Blank, 1980); the mean chart (or the X-chart), the range chart or the R-chart (Moroney, 1976; Summers et al., 1977; Blank, 1980; Banjoko, 1989). In this paper the range-control scale consisting of range-control limits and the sample ranges has been used because doing so sufficiently explains the thought of the process.

In the end if all the appropriate sample statistic (e.g. sample means, sample ranges) lie between the appropriate control limits

of the control chart, the process or the batch of its finished products is said to be at a desired level of good quality and hence it is acceptable as far as the study variable or attribute goes. However, if at one of the sample statisticslies outside the appropriate control limits of the chart, then, something in the process has changed and it requires investigation (Caswell, 1982). He also stated that it is usual to look upon the following occurrences as signals that some interfering factor is influencing the process and hence investigation is required:

A point outside the action-control limits.

i. Two successive points between the warning- and the action- control limits.

ii Five successive points in control but on the same side of the mean line.

MATERIALS AND METHODS

Item	Sample number				Row or Sample
	1	2. ..j … m			Range R_i
1	X_{11}	X_{12} … X_{1j} … X_{1m}			R_1
2	X_{21}	X_{22} … X_2j … X_{2m}			R_2
.		.	.	.	.
.				.	.
.				.	.
I	X_{i1}	X_{i2} … X_{ij} … X_{im}			R_i
.	.	.		.	.
.	.	.		.	.
.	.	.		.	.
N	X_{n1}	X_{n2} … X_nj		X_{nm}	R_n
Column or Sample range R_j	R_1	R_2…	R_j…	R_m	

It is known that this data-presentation format is adaptable to that of the Latin Square by letting the number of rows be equal to the number of columns equal to the number of treatments in experimental design in agriculture. The objective here is to compare the range-control limits analysis with Latin Square in analyses of agricultural data. In testing of a hypothesis in statistics, various methods may be used, but the conclusion to accept or reject the tested hypothesis must be the same for all such methods. This is what is in focus in this paper. For example, if Latin Square is used, and the range-control limits analysis is also used and the two methods lead to the same conclusion-accepting or rejecting the tested hypothesis, then, the two methods can be said to be good alternatives.

The sample range is the difference between the largest and the smallest values in the sample. Control of item-to-item variability

in quality work is achieved by setting up a control chart for the range. The sampling distribution of the ranges has been worked out and tables have been produced for the purpose [Alabi-Labaika, 2005].

In the data format samples each of size m or n are taken and the sample ranges are also taken and denoted as $R_{1.}$, $R_{2.}$, …,Ri., …, $R_{n.}$ for the row arrangement of data or $R._{1}$, $R._{2}$, …, $R._{j}$, …, $R._{m}$ for the column arrangement. The mean of such sample ranges, R, is calculated from:

$R_r = \frac{\Sigma \overline{R}_{i.}}{r}$ for rows , for columns $\overline{R}_c = \frac{\Sigma R_i}{c}$

for treatments $\overline{Rt} = \frac{\Sigma R_{i.}}{t}$

There are several alternative tables of range factors and we shall use the one having four factors or coefficients for a given sample size n. The factors are denoted by DUA DUW, DLW, DLA corresponding to the four range-control limits usually used. This alternative of four factors is preferred because it has warning factors. Warning prevents damage that may later occur otherwise. Each coefficient is multiplied by a given mean of sample ranges, R, (Moroney, 1976) as

Range Upper Action Control Limit (RUACL) = DUA$\overline{X}$R
Range Upper Warning Control Limit (RUWCL) = DUW $\overline{X}$ R
Range Lower Warning Control Limit (RLWCL) = DLW $\overline{X}$ R
Range Lower Action Control Limit (RLACL) = DLA X $\overline{R}$

Decisions are then taken as explained by Caswell (1982). That is, if all the sample ranges lie between the warning limits, then, the sample range is under control and we assert that there is an insignificant difference among sample ranges, and consequently variability in output is owing to chance or that the treatment under test is not significantly effective. However, if a sample range lies between the upper warning control limit and the upper action control limit or between their respective lower counterparts, then the sample range is not under control and we conclude that there is a warning that there may be an upset that requires investigation. If a sample range lies beyond any of the two actions limits, then, undoubtedly something is wrong with the process that needs to be investigated and corrected. Test samples are taken after correction to make sure that steady-state conditions have resumed before full-scale production is recommenced. However, if it is a treatment that is under test, then we conclude that such a treatment is highly significant or effective.

RESULTS AND DISCUSSION

Consider the agricultural data on yields of wheat in kilograms resulting form five manurial treatments A, B, C, D and E applied to 25 plots of land shown in Table 1.

Table 1: Wheat yields from treatments A, B, C,D and E with row and column ranges

	A	B	C	D	E	Row range
	13	9	21	7	6	15
	D	E	A	B	C	
	9	8	15	7	16	9
	B	C	D	E	A	
	11	17	8	10	17	9
	E	A	B	C	D	
	8	15	7	10	7	8
	C	D	E	A	B	
	11	9	8	15	11	7
Column range	5	9	14	8	11	

Source of data: Moroney (1976)

We now conduct tests of significance or otherwise of the effects of the five rows, the five columns and the five treatments A, B, C, D and E as given in the table. The three corresponding null hypothesis are:

i. H_o: the effect of rows is not significant an hence, it may be neglected
ii. H_o: the effect of columns is not significant and hence it may be neglected.
iii. H_o: the effect of treatments is not significant and hence it may be neglected.

We test these null hypotheses by using range-control limits.

i. The five rows:
Row sample ranges R_r are 21-6, 16-7, 17-8, 15-7, 15-8 or 15,9,9,8,7.
The average of the five ranges, R^{-r}, is

$$R_r = \frac{15 + 9 + 9 + 8 + 7}{5} = 9.6$$

Here n = 5 rows, $R^{-}_r = 9.6$, m =5

From quality-control tables of coefficients of R_r, D_i 's
We have, as previously noted, these usual four range control limits as:

RUACL = DUA $\bar{R}_r$ =2.34 (9.6) =22.464
RUWCL = DUW $\bar{R}$-$_r$ =1.81 (9.6) =17.376
RLWCL = DLW $\bar{R}_r$ =0.37 (9.6) =3.552
RLACL = DLA $\bar{R}_r$ = 0.16 (9.6) =1.536

where: DUA=2.34,DUW=1.81,DLW=0.37, DLA=0.16 for sample size 5 are obtained from statistical- quality- control tables for sample ranges (Alabi-Labaika, 2005).

ii. The five columns:
Column sample ranges R_c, are:
13-8, 17-8, 21-7, 15-7, 17-6 OR 5, 9, 14, 8, 11.
The average of the sample ranges, R_c is:

$$\bar{R}_c = \frac{5 + 9 + 14 + 8 + 11}{5} = 9.4$$

Here n = 5, R_c = 9.4

From the quality control table of coefficients of R_c,

$D_{i's}$ we have the usual four range control limits as:

RUACL = DUA $\bar{R}_c$ = 2.34 (9.4) = 21.996
RUWCL = DUW $\bar{R}_c$ = 1.81 (9.4) = 17.014
RLWCL = DLW $\bar{R}_c$ = 0.37 (9.4) = 3.478
RLACL = DLA $\bar{R}_c$ = 0.16 (9.4) = 1.504

iii. The five treatments A, B, C, D and E. Here we collect all the yields for each one of the five treatments and get:

Table 2: Yield sample ranges for treatment A, B, C, D and E

Yield number	A	B	C	D	E
1	13	9	21	7	6
2	15	7	16	9	8
3	17	11	17	8	10
4	15	7	10	7	8
5	15	11	11	9	8
Yield sample range	4	4	11	2	4

The five treatments yield sample ranges $\bar{R}_t$ are
17-13, 11-7, 21-10, 9-7, 10-6
OR 4, 4, 11, 2, 4
The mean of the treatment yield sample range R_t is (t = 5)

$$Rt = \frac{4 + 4 + 11 + 2 + 4}{5} = 5$$

Here again n = 5, $\bar{R}_t$ = 5
From the quality control table of coefficients of sample range mean, R_t, Di's, we have the following usual four control limits.

RUACL	=	DUA $\bar{R}_t$	=	2.34 (5)	=	11.7
RUWCL	=	DUW $\bar{R}_t$	=	1.81 (5)	=	9.05
RLWCL	=	DLW $\bar{R}_t$	=	0.37 (5)	=	1.85
RLACL	=	DLA $\bar{R}_t$	=	0.16 (5)	=	0.80

Collecting all the results for rows, columns and treatments separately and arranging them in a decreasing order of magnitude in each case we have for:

i. Rows

Table 3: Fixing row sample ranges in the range-control limits

Serial number	Row number	Range-control limits and sample ranges
1	-	22.46 = RUACL
2	-	17.38 = RUWCL
3	1	15
4	-	9.6 = R_r row men range
5	2	9
6	3	9
7	4	8
8	5	7
9	-	3.55 = RLWCL
10	-	1.54 = RLACL

In the orderly arrangement or the range control scale all the row sample rages are between the two warning limits (17.38=RUWCL and 3.55=RLWCL). We conclude ,consequently, that sample ranges are statistically controlled. Hence, we accept the null hypothesis that the effect is not significant and hence, it may be neglected.

(ii) Columns:

Table 4: Fixing column sample ranges in the range-control limits

Serial number	Column number	Range-control limits and sample ranges
1	-	22.00 = RUACL
2	-	17.01 = RUWCL
3	3	14
4	5	11
5	-	9.4 = R_c column mean range
6	2	9
7	4	8
8	1	5
9	-	3.48 = RLWCL
10	-	1.50 = RLACL

In the range- control limit scale, all column sample ranges lie between the two warning limits (17.38=RUWCL and 3.55=RLWCL). Hence, we accept that the column sample ranges are under control and consequently accept the hypothesis of the insignificant effect of the columns.

iii. Treatments:

Table 5: Fixing treatment sample ranges in the range-control limits

Serial number	Treatment letter	Range-control limits and sample ranges	
1	-	11.70	= RUACL
2	C	11	
3	-	9.05	= RUWCL
4	-	5	= $\bar{R}_t$ treatment mean range
5	A	4	
6	B	4	
7	E	4	
8	D	2	
9	-	1.85	= RLWCL
10	-	0.80	= RLACL

In the treatment range-control scale, the sample range for the treatment C is outside the warning limits, 9.05= RUWCL and1.85 = RLWCL, and approaches the upper action limit, 11.70= RUACL. This is a warning that there may be a problem or significance of the treatment C using Caswell's decision criterion mentioned earlier that one point on or outside the action control line is usually looked upon as a signal that some interfering factor is influencing that process. We may then conclude that the five treatments significantly influenced the yield of wheat as we have it. To be perfectly right in this case we should say that if a sample range lies between the upper warning limit and the upper action limit or between their respective lower counterparts, then, there is significance.

CONCLUSION

Range-control analysis: the effects of the rows and the columns are insignificant, but the effect of the treatments is significant. This conclusion is to be compared with that of the analysis of variance for the Latin-Square design in the next section.

ANOVA with Latin Square

The following points are to be noted at this transition from the range-control limits analysis of agricultural data to the use of the Latin Square Design based on the analysis of variance:

The above-given data was analyzed as if they were manufacturing data: the objective is to see if a statistical quality control technique can be applied to agricultural data of that form. The range-control limits analysis has been chosen out of many quality control techniques because it is relatively simple compared to some others in the group. Further, it is also a part of the objective to give a rather simple method as an alternative to the classical agricultural methods of analyses of such data..

The range- control limits scale has been used to determine the degree of effectiveness of each treatment, that is, which

treatments perform better than when there is no treatment by looking at the positions of the appropriate ranges on the scale; the degree of the effectiveness of every row and of every column is similarly determined. However, the analysis of variance in this paper determines the collective effectiveness or otherwise of the treatments by hypothesizing that they have the same effect and by extension that they have equal yield means. If the yield means of the treatments are significantly different, the effects of the treatments are collectively significant. Further, it is required that the sample size ,n, be equal to the number of samples ,m, equal to the number of rows ,r, equal to the number of columns ,c, equal to the number of treatments ,t, for the comparative analysis being proposed here to be workable.

It is the range- control limits only that are used and not the range- chart [or graph] because rang- control limits alone are sufficient for the required analysis in our case here.

The range- control limits parameters are tabulated for the range - control limits analysis [Alabi-Labaika, 2005], they are not used in the course of the analysis of variance of the Latin- Square Design.

Again, the link between the range- control analysis [RCL] with the analysis of variance [ANOVA] is:
Number of rows, r ,= the number of columns, c,= the number of treatments, t,= the sample size, n,=the number of samples, m
The joint effects of the treatments are in focus in the tests here and not individual treatment effects.

Null Hypothesis (Ho) for Latin Square Design -Analysis of Variance [ANOVA]

1. Ho : For the treatments the sample yield means are equal $x_1 = x_2 = x_3 = \ldots =$ or not significantly different.
2. Ho: The effect of r rows is not significant
3. Ho : The effect of c columns is not significant

Calculation for the tests:

Total sum of squares $= \Sigma x_i^2 - [\Sigma x_i]^2 /rc = 13^2+9^2+21^2+\ldots+11^2+7^2 - [13+9+21+\ldots+11+7]^2 = 3413-3025=388$

Rows sum of squares $= \Sigma T^2 i./ni. - [\Sigma x_i]^2/rc = 56^2/5 + 55^2/5 + 63^2/5 + 47^2/5 + 54^2/5 - 275^2/25 = 15255/5 - 3025 = 3051 - 3025 = 26$

Columns sum of squares $= \Sigma T^2.j/n.j - \Sigma[x_i]^2/rc = 52^2/5 + 58^2/5 + 59^2/5 + 49^2/5 + 57^2/5 - 275^2/25 = 3039.8 - 3025 = 14.8$

Treatments sum of squares $= 75^2/5 + 45^2/5 + 75^2/5 + 40^2/5 + 40^2/5 - 275^2/25 = 16475/5 - 3025 = 3295 - 3025 = 270$

Mean sum of squares, mss=sum of squares/degree of freedom=ss/df
The degree of freedom df for r rows is r-1, for c columns, is c-1, t-1 for t treatments and rc-1 for the total sum of squares.

ANOVA Table

Source of variation:	df	ss	mss	F-ratio
1. Rows	4	26	6.50	1.01
2.Columns	4	14.8	3.70	0.58
3.Treatments	4	270	67.50	10.50
4. Error	12	77.20	6.43	-
Total	24	388	-	-

where: ss = sum of squares, df =degrees of freedom, mss= mean sum of squares

From the table of the F-distribution, F[4, 12, .01] = 5.41 and F[4,12, .05]=3.26 Comparing the F-ratios with the tabulated F's, the F-ratios of rows[1.01] and of columns [0.58] are each less than those from tables at .01 probability level, [5.4From the table of the F-distribution, F [4, 12, .01] = 5.41 and F[4,12, .05]=3.26 Comparing the F-ratios with the tabulated F's, the F-ratios of rows[1.01] and of columns [0.58] are each less than those from tables at .01 probability level, [5.41] and at .05 probability level, [3.26]. It is concluded, then, that the effects of rows and columns do not have significant effects on the wheat yield. However, the F-ratio of treatments [10.50] is higher than the tabulated ones:[5.41 and 3.26]. Hence, the treatments significantly influenced the wheat yield.

The over- all conclusions are that with the range-control analysis [RCA] the effects of rows and columns are not significant on the wheat yield. Similarly the effects of rows and columns are not significant with analysis of variance [ANOVA]. Further, the effect of the treatments is significant with both RCA and ANOVA methods of analysis. In that case the two methods, RCA and ANOVA, are good alternatives for analyzing agricultural data presented according to the format given in this paper, i.e., the number of rows = the number of columns= the number of treatments = sample size= number of samples [for the Latin- Square Design and Range-Control Analysis]. At this point the set objective of the paper has been achieved.

The RCA may also be used to analyze market data in a row-by-column form, say, in agricultural economics.

REFERENCES

Adigun, J.A. 2002. "Chemical weed control in transplanted rain-fed tomato … in the forest savanna transition zone of South Western Nigeria" ASSET Journal, University of Agriculture, Abeokuta, Nigeria, Series A, Vol.2 No. 2 pp. 141-150

Adigun. J.A. Lagoke, S.T.O., Adekpe, I.D. 2003. "Efficacy of selected herbicides for weed control in rain-fed upland rice in the Nigerian Northern Guinea Savanna." ASSET, Journal, University of Agriculture, Abeokuta, Nigeria, Series A, Vol.3 No.1 pp. 43-51

Alabi-Labaika, A. B. 2005. Principles and Administration of statistical quality control, University of Lagos Press, Lagos, pp. 24-30

Androulidakis, S.I., Siardos, G.C. 1996. "Agriculture Extension Agents Perceptions Regarding Their Relevance and Compe-

tence in Certain Professional Task Areas" European Journal of Agricultural Educational Extension Vol 1 No 3 pp. 1-14

Banjoko, S. A. 1989. Techniques of production management, Lagos, Development Press Ltd. Pp.255-272

Blank, L.1980. Statistical Procedures for Engineering, Management and Science, New York McGraw-Hill Book Co pp. 256-549

Brookes, B. C., Dick, W. F. L. 1969. Introduction to Statistical Methods, 2nd edition, London: Heinemann pp.264-267

Caswell, F. 1982. Success in Statistics, John Murray Publishers Ltd. Pp258-263

Moroney, M.J. 1976. Facts From Figures, London, Cox and Wyman Ltd:. Pp. 141-215 & 390-394

Neter. J., Wasserman, W., Kutner, M. H. 1990. Applied Linear Statistical Methods 3rd Edition, Boston: IRWIN, pp. 522 -560

Summers, G.W., Peters, I.S., Armstrong, C.P. 1977. Basic Statistics in Business and Economics, Belmont, Wadsworth Publishing. PP. 378-385.

Moroney, M.J. 1976. Facts From Figures, London, Cox and Wyman Ltd:. Pp. 141-215 & 390-394

Neter. J., Wasserman, W., Kutner, M. H. 1990. Applied Linear Statistical Methods 3rd Edition, Boston: IRWIN, pp. 522 -560

Summers, G.W. Peters, I.S., Armstrong, C.P. 1977. Basic Statistics in Business and Economics, Belmont, Wadsworth Publishing. PP. 378-385.