

SEQUENTIAL ANALYSIS OF MEANS FOR TESTING EQUALITY OF SEVERAL MEANS

G. M. OYEYEMI

Department of Statistics

University of Ilorin

Nigeria.

gmoveyemi@gmail.com, matanmi@unilorin.edu.ng

+2348052278655

B. L. ADELEKE

Department of Statistics

University of Ilorin

Nigeria.

bladeleke@unilorin.edu.ng

Keywords: ANOM, Decision Lines, Homogenous, SANOM, Treatments.

Abstract

The of Analysis of Means (ANOM) and Sequential Analysis of Means (SANOM) as graphical statistical techniques for comparing group of treatment means were examined in this paper. Results arising from the ANOM and SANOM were compared with the conventional (parametric) approach, Analysis of Variance (ANOVA) and its Post Hoc analysis. Apart from its simplicity for non-statistician, ANOM and SANOM were found to be more efficient, precise and unambiguous especially in its post hoc analysis (SANOM) for producing non-overlapping homogeneous subgroups in the separation of means.

1.0 Introduction

The analysis of means (ANOM) is a graphical and statistical method used in illustrating important variations among group of means (populations) and it is commonly employed in statistical quality control. ANOM is a technique originally developed by Ott (1967) for comparing a group of treatment means in order to ascertain if any of them differs significantly from the overall mean at a specified significance level. ANOM methodology compares the mean of each group to overall mean to detect statistically significant differences among means. It is not limited to comparison of group means only but also used to compare rates, proportions and variances. While analysis of variance (ANOVA) is the most commonly used method to compare the means of several groups, analysis of means (ANOM) serves as alternative procedure for comparing means.

Though ANOM cannot be used in the same settings as the ANOVA, ANOM has the advantages of being much more intuitive and providing easily understood graphical results, which clearly indicates any means that are different from the overall mean and allows for easy assessment of practical as well as statistical significance. The graphical results obtained from ANOM are easy for non-statisticians to understand and offers a clear advantage over ANOVA in that it sheds light on the nature of the differences among the groups. Ott (1967) observed that ANOVA was difficult for non-statisticians to understand. His objective was to give quality control practitioners an alternative procedure that they could use. ANOM as a procedure is more effective than ANOVA because graphical display of data is a part of the procedure. Nevertheless, the concept of the two procedures ANOM and ANOVA are basically the same and the conclusions are usually the same. While ANOVA basically examines if the group taken as a whole displays any criminality, the ANOM examines if the group harbors a criminal, and thereafter identifies the individual criminal.

Often the experimenter performs one of the multiple comparison procedures when the null hypothesis is rejected in analysis of variance test, the experimenter further wants to know what the different homogeneous groups under study are. In such a case, the experimenter performs one of the multiple comparison procedures, among which are the Turkey's test, Scheffe's test, S-N-K test or Duncan's test. These multiple comparison procedures often yield overlapping homogeneous groups of means. In ANOM graphical method of comparing means, rates and proportions, a new procedure called Sequential analysis of means (SANOM) is formulated to separate a set of treatment means into non overlapping homogeneous groups. This in turn helps the researcher in the interpretation of the results of the experiment and ultimately facilitates managerial decisions.

2.0 Analysis of Means (ANOM)

The analysis of variance (ANOVA) is a good tool to determine if there are differences between several sample means, but it does not determine where the differences come from and so, it does not show which samples are so disparate that the null hypothesis has to be rejected. To know where the difference originates from, it is necessary to conduct further analysis after rejecting the null hypothesis. Turkey, Fisher and Duncan procedures are examples of comparisons that can help situate the sources of variations among means (see e.g. Mason *et al.* (2003)). Analysis of means is an alternative to ANOVA for factors involving fixed effects but is inappropriate for factors involving random effects.

When conducting ANOM, we determine first the upper and lower decision limits. If all the sample means fall within those boundaries, we can say with 100 $(1 - \alpha)$ percent confidence that there is no ground to reject the null hypothesis, that is, there is no significant difference among the sample means. If at least one mean falls outside the limits, we reject the null hypothesis. The upper and lower decision limits depend on several factors, namely; the sample means, the overall mean of the sample means, the standard deviation, the level of significance, the number of samples, and the sample size for determining the degree of freedom.

An ANOM graph is conceptually similar to a control chart; it portrays decision lines so that both statistical and practical significance of samples may be assessed simultaneously. The

main difference between Statistical Quality Control Charts and ANOM charts is that the value of k (number of samples being compared) is usually as large as 20 or more in control charts to compute the control limits, whereas $k \geq 2$ in ANOM chart is enough to compute the decision lines. In essence, ANOM is only useful for the improvement of quality rather than the control of quality.

For testing the equality of several population means, ott (1967) introduced the analysis of means (ANOM) procedure which is based on the multiple significance test proposed by Halperin et al (1955). Halperin (1966) computed upper and lower bounds to the upper 5% and 1% points of the distribution of d , where $d = \max\{|x_i - \bar{x}|/S\} i = 1, 2, 3, \dots, k$, the conjectured true values, were close to the lower bound. Ott (1967) considered the averages of the above upper and lower bounds as H_α . For comparing the individual sample means with the overall mean, the following decision lines are computed:

$$\bar{\bar{x}} \pm H_\alpha \check{\check{d}}_{\bar{x}}, \text{ where } \bar{\bar{x}} = \text{average of } k \text{ sample means.}$$

$$H_\alpha = \max \left[\frac{\bar{x}_{(n)} - \bar{\bar{x}}}{\check{\check{d}}_{\bar{x}}}, \frac{\bar{\bar{x}} - \bar{x}_{(1)}}{\check{\check{d}}_{\bar{x}}} \right], \quad \check{\check{d}}_{\bar{x}} = \frac{\check{\check{d}}}{\sqrt{n}}, \quad \check{\check{d}} = \frac{\bar{R}}{d_2^*} \quad (2.1)$$

$\bar{x}_{(i)}$ is the ordered mean, $\bar{R}$ the average of k sample ranges, d_2^* is a factor for estimating σ from $\bar{R}$ and it depends on k (where as the usual d_2 factor in the control chart is independent of k) and n is the sample size.

Tables of H_α for $\alpha = 0.05$ and 0.01 have been developed by Ott (1967) for selected values of k and selected degrees of freedom for error $\nu = 0.90k(n-1)$. Schilling (1973a) used the critical values h_α in computing the decision lines

$$\bar{\bar{x}} \pm h_\alpha \sqrt{\frac{k-1}{k}} \check{\check{d}}_{\bar{x}}, \text{ where } h_\alpha = H_\alpha \sqrt{\frac{k}{(k-1)}}. \quad (2.2)$$

The upper bound from Halperin et al. (1955) was considered as H_α by Schilling (1973a), who developed the tables of h_α for $\alpha = 0.05$ and 0.01 . Nelson (1974) extended the tables of h_α for $\alpha = 0.10$ and 0.001 . Sheesley (1981) presented simplified factors A_α and computed the decision lines as;

$$\bar{\bar{x}} \pm A_\alpha \bar{R}, \text{ where } A_\alpha = \frac{H_\alpha}{d_2^* \sqrt{n}} \quad (2.3)$$

Nelson (1982) obtained the exact critical point for h_α and used the decision lines as;

$$\bar{\bar{x}} \pm h_\alpha S \sqrt{\frac{k-1}{kn}}, \quad (2.4)$$

Where S^2 is the pooled sample variance and h_α is the critical point (which depends on k , v and α) which can be obtained as the upper $(\alpha/2k)$ percentage point of the student's t -distribution with $v = k(n-1)$ degrees of freedom. Ott (1967) and Schilling (1973a) used the range statistic d_2^* to estimate σ in their procedure. It is important to note that this procedure gives an approximate result. By using S in the estimation of σ , Nelson (1982) obtained exact critical values of h_α for $\alpha = 0.01, 0.05$ and 0.10 in the case of equal sample sizes.

3.0 SANOM Procedure for multiple comparison of treatment means

This procedure is an extension of ANOM procedure given in Nelson (1983a) for variable data. The new procedure is a divisible procedure in which a set of means can be separated into three groups. In this procedure, a pair of decision lines, the Lower Decision Line (LDL) and the Upper Decision Line (UDL) are drawn in different ANOM charts constructed in different stages. In the first stage, we assume that the means falling between the pair of decision lines form the first group and are homogeneous, whereas the means falling above the UDL form a second group and the means falling below the LDL form a third group. The ANOM procedure is then repeated on the second and third groups only until no further division is possible. The sequential analysis of means (SANOM) procedure for the multiple comparison of k means computed from a sample of size n is given in the following steps.

Step 1. In the first stage, compute the upper and lower decision lines based on k treatment means

$$UDL_1 = \bar{\bar{x}} + h_{\alpha,k,v} S \sqrt{(k-1)/kn} \quad (3.1)$$

$$LDL_1 = \bar{\bar{x}} - h_{\alpha,k,v} S \sqrt{(k-1)/kn} \quad (3.2)$$

$$\text{Where } \bar{\bar{x}} = \frac{1}{k} \sum_{i=1}^k \bar{x}_i, \quad \bar{x}_i = \frac{1}{n} \sum_{j=1}^n x_{ij}, \quad i = 1, 2, \dots, k$$

$$S = \sqrt{\sum_{i=1}^k S_i^2 / k}, \quad S_i^2 = \frac{1}{(n-1)} \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2,$$

And $h_{\alpha,k,v}$ is the upper $(\alpha/2k)$ percent point of the t -distribution with v degree of freedom and is observed from the tables of Nelson (1993) for the number of treatment means (k), the degrees of freedom $v = k(n-1)$ in S and the level of significance α at each stage.

Step 2. Plot the means $\bar{x}_i, i = 1, 2, \dots, k$ against the decision lines.

- If all the means fall between UDL_1 and LDL_1 , conclude that a statistically significance does not exist among the means. In other words, all the treatment means belong to a homogeneous group. In this case, we stop the analysis.
- If one or more means fall outside the decision lines (i.e. above UDL_1 and/or below LDL_1), conclude that a statistically significant difference exists among the means. In other words, the means falling between decision lines form homogeneous

group I, and the means falling outside the decision lines do not belong to this homogeneous group. In this case, we proceed further.

Step 3 Let k_1 treatment means ($2 \leq k_1 < k$) fall above UDL_1 and we want to know whether these k_1 means form a homogeneous group or not. In the second stage, compute the upper and lower decision lines as

$$UDL_2 = \bar{\bar{x}} + h_{\alpha, k_1, v} S \sqrt{(k_1 - 1)/k_1 n} \quad (\text{iii})$$

$$LDL_2 = \bar{\bar{x}} - h_{\alpha, k_1, v} S \sqrt{(k_1 - 1)/k_1 n} \quad (\text{iv})$$

In (iii) and (iv), $\bar{\bar{x}}$ and S are computed from k_1 means and $v_1 = k_1 (n-1)$: plot the means $\bar{x}_i$, $i = 1, 2, \dots, k_1$ against the decision lines. If these k_1 means fall between UDL_2 and LDL_2 , conclude that these k_1 means form homogeneous group II. If $k_1 = 1$, step 3 does not occur and we simply assume that this mean belongs to homogeneous group II.

Step 4 Let k_2 treatment means ($2 \leq k_2 < k$) fall below LDL_1 and we wish to know whether these k_2 means form a homogeneous group or not. In the third stage, compute the upper and lower decision lines as

$$UDL_3 = \bar{\bar{x}} + h_{\alpha, k_2, v} S \sqrt{(k_2 - 1)/k_2 n} \quad (\text{v})$$

$$LDL_3 = \bar{\bar{x}} - h_{\alpha, k_2, v} S \sqrt{(k_2 - 1)/k_2 n} \quad (\text{vi})$$

In (v) and (vi), $\bar{\bar{x}}$ and S are computed from k_2 means and $v_2 = k_2 (n-1)$: plot the means $\bar{x}_i$, $i = 1, 2, \dots, k_2$ against the decision lines. If these k_2 means fall between UDL_3 and LDL_3 , conclude that these k_2 means form homogeneous group III. If $k_2 = 1$, step 3 does not occur and we simply assume that this mean belongs to homogeneous group III.

Step 5 If some of the k_1 (at least 2) means in step 3 and/or some of the k_2 (at least 2) means in step 4 still fall outside the respective decision lines, we continue the analysis by further computing decision lines, till these means form into distinct homogeneous groups.

4.0 Illustrative example

4.1 Using Duncan Test and Sequential Analysis of Means

A product development engineer is interested to know whether the cotton weight percentage in a synthetic fiber used to make cloth for men's shirt affects the tensile strength of the fiber.

Table 1: Data on tensile strength of synthetic fiber of different percentage cotton

cotton wh1	cotton wh2	cotton wh3	cotton wh4	cotton wh5	cotton wh6
19.40	17.70	17.00	20.70	14.30	17.30
32.60	24.80	19.40	21.00	14.40	19.40
27.00	27.90	9.10	20.50	11.80	19.10
32.10	25.20	11.90	18.80	11.60	16.90
33.00	24.30	15.80	18.60	14.20	20.80

3.42 Applying Duncan's Test for Multiple Comparison

The following Tables present the summary of both ANOVA and Duncan Multiple Range Test (DMRT) results.

Table 2(a) ANOVA Table

Tensile strength

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	847.047	5	169.409	14.371	.000
Within Groups	282.928	24	11.789		
Total	1129.975	29			

Table 2(b)

Tensile strength

Duncan^a

Cotton	N	Subset for alpha = .05				
		1	2	3	4	5
wh5	5	13.2600				
wh3	5	14.6400	14.6400			
wh6	5		18.7000	18.7000		
wh4	5			19.9200	19.9200	
wh2	5				23.9800	
wh1	5					28.8200
Sig.		.531	.074	.579	.074	1.000

a. Uses Harmonic Mean Sample Size = 5.000.

4.2 Sequential Analysis of Means

There are six treatments denoted by wh_1 , wh_2 , wh_3 , wh_4 , wh_5 , wh_6 with each treatment replicated five times; $k = 6$, $n = 5$. The means and variances of the six treatments are presented in Table 3.

Table 3: Means and variances of the six treatments

<i>Treatments</i>	<i>Means ($\bar{X}_i$)</i>	<i>Variances (S_i^2)</i>
wh_1	28.82	33.64
wh_2	23.98	14.27
wh_3	14.64	16.94
wh_4	19.92	1.28
wh_5	13.26	2.04
wh_6	18.70	2.57

Stage 1:

From Table 3, we compute $\bar{\bar{X}} = 19.89$ and $s = 3.434$ using all the six means and variances. We observed $h_{\alpha,k,v} = h_{0.05,6,24} = 2.83$, $\alpha = 0.05$, $k=6$, $v=24$

$$V = k(n-1) = 6(5-1) = 24$$

Compute decision lines with $k=6$ and $n=5$

$$\begin{aligned} UDL_1 &= \bar{\bar{X}} + h_{\alpha,k,v} S \sqrt{(k-1)/kn} \\ &= 19.89 + h_{0.05,6,24} (3.434) \sqrt{(6-1)/6(5)} \\ &= 23.86 \end{aligned}$$

$$\begin{aligned} LDL_1 &= 19.89 - 3.97 \\ &= 15.92 \end{aligned}$$

The decision lines UDL_1 and LDL_1 are drawn and the means $\bar{X}_i = 1, 2, \dots, 6$ are plotted against the decision lines. Treatment wh_4 and wh_6 fall between the decision lines and the rest of the means fall outside the decision lines. Hence, the treatment wh_4 and wh_6 form homogeneous group I (see Figure 1).

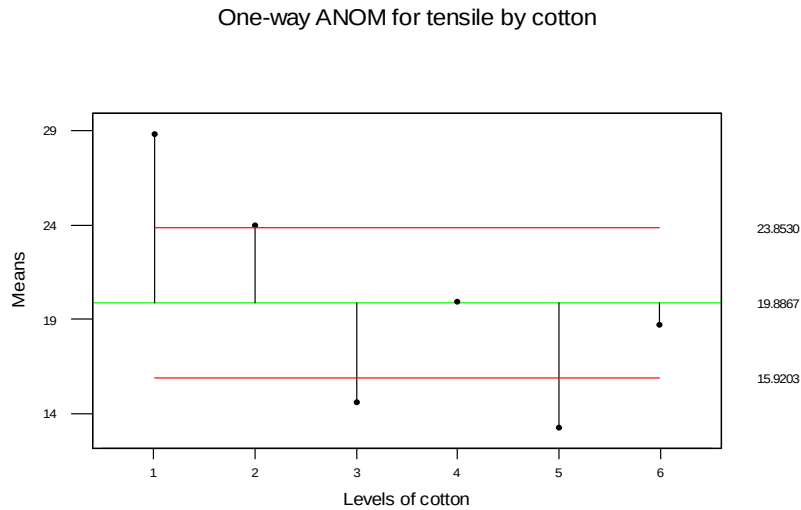


Figure 1: One-way ANOM plot of tensile strength of different levels of cotton.

Stage 2:

From Figure 1, the means of treatments wh_1 and wh_2 fall above UDL_1 . We compute $\bar{\bar{X}} = 26.4$ and $s = 4.894$. The second stage decision line is computed from these two treatment means and variance. $k_1=2$ and $n=5$.

$$UDL_2 = 26.4 + (2.306) (4.894) (0.316) = 29.97 \quad LDL_2 = 26.4 \pm 3.57 = 22.83$$

Where $h_{\alpha,k,v} = h_{0.05,2,8} = t_{0.05,8} = 2.306$, $\alpha=0.05$, $k=6$, $v=24$ The h_{α} values may be seen from student's t -table where there are only two means to be compared.

The two means are plotted against decision lines UDL_2 and LDL_2 . The two means fall between the decision lines and hence form a homogeneous group II.

Stage 3:

It is observed that treatment means S_3 and S_5 fall below LDL_1 . We compute $\bar{\bar{X}} = 13.95$ and $s = 3.081$ for treatment wh_3 and wh_5 . The third stage decision line UDL_3 and LDL_3 are computed as follows $k_2=2$ and $n=5$

$$UDL_3 = 13.95 + (2.306) (3.081) (0.316) = 16.20$$

$$LDL_3 = 13.95 \pm 2.25 = 11.70$$

The two means are plotted against decision lines UDL_3 and LDL_3 . It is observed that these two means fall between the decision lines and form a homogeneous group III.

Stage 4:

The three homogeneous groups (wh_6, wh_4) , (wh_2, wh_1) and (wh_5, wh_3) are non-overlapping.

4.1 Interpretation of Results

In the example, the ANOVA result shows that the average tensile strength are not equal, while the ANOM graphical method of Nelson and Dudewicz (2002) shows that some of the points fall outside the decision lines. Hence, the ANOM $\bar{X}$ chart gives additional information that the average tensile strength of cottons wh_1 and wh_2 are higher than the overall mean while those of wh_3 , and wh_5 are lower.

Duncan's test for multiple comparisons shows overlapping of treatment means in wh_3 , wh_4 and wh_6 . Applying sequential analysis on means procedure to the same data, we do observed that each treatment belongs to only one homogeneous groups, and the three homogeneous groups (wh_6, wh_4) , (wh_2, wh_1) and (wh_5, wh_3) are non-overlapping.

4.2 Conclusion

The ANOM procedure for comparing averages, proportions and rates as demonstrated in this study combines both the features of ANOVA and other post-hoc analysis and could be used without an in-depth knowledge of statistics. Furthermore, ANOM is easy to interpret and more effective due to the fact that it does more than revealing differences among the population means as it also presents Graphical display of data, which consequently identify individual treatments means that have significant contributions.

REFERENCES

- Halperin, M., Greenhouse, S. W., Cornfield, J. and Zalokar, J. (1955).** Tables of percentage points for studentized maximum absolute deviate in normal samples. *Journal of American Statistical Association* 50, 185-195.
- Halperin, M. (1966).** Confidence interval from censored samples II. *Technometrics* 8, 291-307.
- Mason, R. L., Gunst, R. F., & Hess, J. L. (2003):** *Statistical Design and Analysis of Experiments; With Application to Engineering and Science.* 2nd ed. John Wiley & Sons, Inc. Hoboken, New Jersey.
- MINITAB** release 14 (Statistical Software).
- Nelson, L. S. (1983a):** Exact critical values for use with the analysis of means. *Technometrics* 30: 53-61.
- Nelson, P. R. (1982):** Exact critical points for the analysis of means. *Communication in statistics-theory and methods*, 11: 699-709.
- Nelson, P. R. (1993):** Additional uses for analysis of means and extended tables of critical values. *Technometrics*, 35: 61-71.
- Nelson, P. R. (1974):** Exact analysis of means with unequal variances. *Technometrics* 44, 152-160.
- Nelson, P. R & Dudewicz, E. J. (2002).** Exact Analysis of Means with unequal variances. *Technometrics* 44, 152 ï 160.
- Ott, E. R. (1967):** Analysis of means-A graphical procedure. *Industrial Quality Control* 24: 101-109.
- Schilling, E. G. (1973a).** A systematic approach to the analysis of means. Part I. *industrial quality control* 40:101-205
- Sheesley, J. H. (1981).** Simplified factors for analysis of means when the standard deviation is estimated from the range. *Journal of Quality Technology.*, 13, 184-185.
- SPSS 15.0** for windows evaluation versions.