TECHNOMETRICS, FEBRUARY 1985, VOL. 27, NO. 1

# Power Curves for the Analysis of Means

**Peter R. Nelson**

Statistical, Management, and Information Sciences
School of Management
Rensselaer Polytechnic Institute
Troy, NY 12181

A brief description of the analysis of means is given with the method used to compute power curves for detecting differences among $K$ treatments at level of significance $\alpha$ when two of the treatment means differ by at least a specified amount $\Delta$ (measured in units of the process standard deviation). Power curves are given for $\alpha = .1, .05, .01$; $K = 3(1)10$; and $1 \le \Delta \le 3$. Two examples are presented.

KEY WORDS: Analysis of means; Multivariate noncentral $t$ distribution; Power curves.

## 1. INTRODUCTION

The analysis of means (ANOM) is a technique for comparing a group of treatments to see if any differences exist among the treatment means. As such it can be thought of as an alternative to the analysis of variance (ANOVA), and in fact for only two treatments the two procedures are equivalent. Ott (1967) introduced ANOM based on the multiple-significance test given by Halperin et al. (1955). It can be thought of as an extension of Shewhart-type control charts that allows for considering groups of sample means instead of one mean at a time. As such it has an advantage over ANOVA in that results can be presented graphically, making the procedure easy to explain and visualize and allowing for an assessment of practical significance as well as statistical significance. While ANOM is not an optimal test in any mathematical sense, the preceding advantages are important practical considerations, since ANOM and ANOVA behave similarly.

Both Scheffé (1947) and Craig (1947) compared control charts with ANOVA, and their comments are also applicable to a comparison of ANOM with ANOVA. Scheffé (1947) made the analogy that ANOVA looks to see if a group taken as a whole displays any evidence of criminality, whereas ANOM looks to see if the group harbors a criminal and provides a method for making an identification. ANOM is, therefore, more sensitive for detecting a single different mean, and an example of this was given by Pearson and Hartley (1958, p. 51). However, as Craig (1947) commented, "the two methods are nearly enough equivalent that both will disclose any clear lack of control among averages and when one says control is good, the other will too" (p. 16).

The purpose of this article is to present power curves for ANOM similar to those in Pearson and Hartley (1951) for ANOVA. The curves give the power for detecting differences among $K$ treatments at level of significance $\alpha$ when two of the treatment means differ by at least a specified amount $\Delta$ (measured in units of the process standard deviation) for

$$\alpha = .1, .05, .01;$$

$$K = 3(1)10;$$

and

$$1 \le \Delta \le 3.$$

These values cover the ranges of primary interest. A brief description of ANOM is given with a discussion of the method used to generate the curves, and two examples are presented.

## 2. THE BASIC PROCEDURE

Very simply ANOM is performed by computing sample means for each treatment, computing a grand mean over all treatments, and checking to see if any of the treatment means differ from the grand mean by too much. To be more specific, let us first consider the simplest case for ANOM corresponding to a one-way ANOVA with equal samples of size $n$. The means, $\bar{X}_i$, are assumed to be from $K$ normally distributed populations with unknown but equal variances, $\sigma^2$; and $S^2$, an estimate of the common population variance, is assumed to be independent of the $\bar{X}_i$ and such that $mS^2/\sigma^2$ has a chi squared distribution with $m$ degrees of freedom. When $S^2$ is computed by pooling the

Table 1. *Iron Concentration Data*

| Analyst | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* |
| 2.963 | 2.958 | 2.956 | 2.948 | 2.953 | 2.941 | 2.963 | 2.987 | 2.946 | 2.956 |
| 2.996 | 2.964 | 2.945 | 2.960 | 2.961 | 2.940 | 2.928 | 2.989 | 2.950 | 2.947 |
| 2.979 | 2.955 | 2.963 | 2.953 | 2.961 | 2.931 | 2.925 | 2.988 | 2.955 | 2.947 |
| 2.970 | 2.932 | 2.950 | 2.944 | 2.953 | 2.942 | 2.940 | 2.983 | 2.969 | 2.960 |
| 2.979 | 2.941 | 2.975 | 2.950 | 2.949 | 2.930 | 2.934 | 2.974 | 2.954 | 2.954 |

estimates of $\sigma^2$ from each of the $K$ populations, the value of $m$ would be $K(n - 1)$. If one denotes, by $\bar{A}_i$, the event

$$|\bar{X}_i - G|/[S\sqrt{(K-1)/(Kn)}] > h_{\alpha, K, m},$$

where $G = \sum_{i=1}^{K} \bar{X}_i/K$ is the grand mean and $h_{\alpha, K, m}$ is a critical point (found in L. S. Nelson 1983) depending on $\alpha$, $K$, and $m$, then ANOM rejects the hypothesis $H_0$: $\mu_1 = \cdots = \mu_K$ that all of the treatment means are the same if at least one of the $\bar{A}_i$ occur. A more detailed discussion of ANOM can be found in the January 1983 issue of the *Journal of Quality Technology*.

### Example 1

Bennett and Franklin (1954, p. 331) presented the data in Table 1 on the effects of 10 analysts on the determination of iron concentration. If these were the only analysts working in a particular lab (fixed effects), one would be interested in testing if any of the analysts are different with regard to determination of iron concentration, and if so, which ones. Bennett and Franklin (1954) performed an ANOVA and found a significant difference at the .001 level. On the other hand, one could perform an ANOM by computing

$$(\bar{X}_i - G)/[S\sqrt{(K-1)/(Kn)}]$$

for $i = 1, \ldots, 10$ and comparing these values with $\pm h_{.001, 10, 40} = \pm 4.32$, the critical value from L. S. Nelson (1983). This is the same as comparing $\bar{X}_i - G$ with $\pm h_{\alpha, K, m} S\sqrt{(K-1)/(Kn)}$, which is done graphically in Figure 1. From Figure 1 it can easily be seen that there is a statistically significant difference and that analysts one, six, seven, and eight are not as accurate as the rest. Moreover, Figure 1 allows for easy assessment of the practical significance associated with the differences.

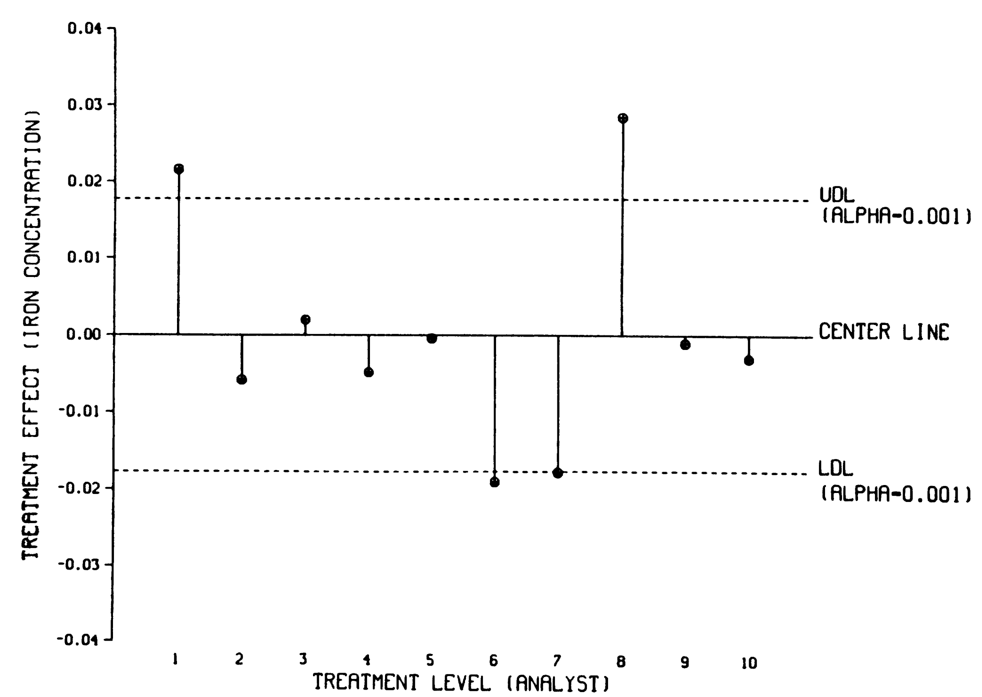


Figure 1. *ANOM Chart of Analysts (Analyst Effects).*

Table 2. *Parameter Combinations for Power Calculations*

| K | n | Δ | α | K | n | Δ | α |
|---|---|---|---|---|---|---|---|
| 3 | 3 | 1.00 | .05 | 4 | 2 | 3.00 | .05 |
| 3 | 3 | 2.00 | .01 | 5 | 4 | 1.25 | .01 |
| 3 | 3 | 2.00 | .05 | 6 | 6 | 2.00 | .10 |
| 3 | 3 | 2.00 | .10 | 7 | 8 | 1.00 | .05 |
| 3 | 3 | 3.00 | .05 | 8 | 10 | 2.50 | .01 |
| 3 | 4 | 1.00 | .10 | 9 | 20 | 1.50 | .10 |
| 3 | 5 | 2.50 | .05 | 10 | 3 | 1.00 | .05 |
| 3 | 6 | 3.00 | .01 | 10 | 3 | 2.00 | .05 |
| 3 | 7 | 3.00 | .10 | 10 | 3 | 2.00 | .10 |
| 3 | 10 | 1.75 | .05 | 10 | 3 | 3.00 | .05 |
| 3 | 20 | 1.00 | .01 | 10 | 8 | 2.00 | .05 |
| 3 | 30 | 2.00 | .10 | 10 | 24 | 3.00 | .05 |

Alcohol

| | | | | |
|---|---|---|---|---|
| Base | 91.3 | 89.9 | 89.3 | 88.1 |
| | 90.7 | 91.4 | 90.4 | 91.4 |
| | 87.3 | 89.4 | 92.3 | 91.5 |
| | 91.5 | 88.3 | 90.6 | 94.7 |

Figure 3. *Percent Yield for Alcohol/Base Combinations.*

## 3. THE POWER FUNCTION

To provide power curves sufficiently general to be useful, it is necessary to restrict the parameter space $M$ to certain subspaces and determine the power for the configuration of means where it is the lowest. This will be referred to as the least favorable configuration (LFC). The example in Scheffé (1959, p. 63) of the use of the Pearson and Hartley (1951) curves suggests that reasonable subspaces would be those where two of the population means differ by at least an amount $\Delta\sigma$. I will, therefore, use the subspaces

$$M_\Delta = \{\boldsymbol{\mu}: \max_{i,j} |\mu_i - \mu_j| \geq \Delta\sigma\}.$$

ANOM, because it subtracts the grand mean from each population mean, translates the vector of population means so that the multivariate $t$ distribution that gives the power has a vector of means that sum to zero. If one denotes these translated means by $\mu_{ti}$,

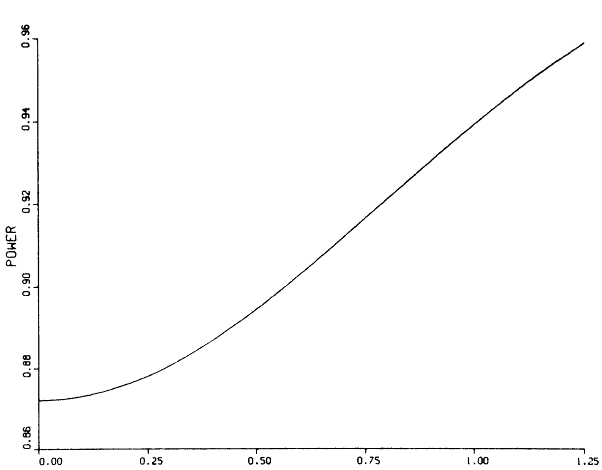


Figure 2. *ANOM Power as a Function of δ. Parameter Values: $K = 3$; $n = 5$; $\Delta = 2.50$; $\alpha = .05$.*

then

$$\mu_{ti} = \mu_i - \sum_{j=1}^{K} \mu_j / K.$$

I conjecture that on $M_\Delta$ the LFC occurs when one translated mean is at $\Delta/2$, one is at $-\Delta/2$, and the rest are at zero. This conjecture is based on two facts. First, it is possible to show (see the Appendix) that the LFC must be of the form

$$\boldsymbol{\mu} = (\Delta/2, \; -\Delta/2, \delta, \ldots, \delta). \qquad (1)$$

Second, extensive numerical results suggest that $P(\boldsymbol{\mu}) = \Pr\,[\text{rejecting } H_0 \,|\, \boldsymbol{\mu}]$ is an increasing function of $\delta$ for $\boldsymbol{\mu}$ of the form (1) and $\delta \in [0, \Delta/2]$.

For the 24 different combinations of $K$, $n$, $\Delta$, and $\alpha$ listed in Table 2, $P(\boldsymbol{\mu})$ was plotted as a function of $\delta$ using the program described in P. R. Nelson (1982). All of the resulting graphs were increasing in $\delta$, and a typical example is shown in Figure 2.

The power curves in Figures 5–28 were computed using the LFC of one mean at $\Delta/2$, one at $-\Delta/2$, and the rest at zero. The degrees of freedom used were based on the assumption that $\sigma$ would be estimated using the pooled root-mean-square estimator. The program in P. R. Nelson (1982) was used to compute the power at values of $\Delta = 1, 1.25, 1.5, 1.75, 2, 2.5$, and 3, and the curves are the result of interpolation using splines.

### Example 2

One step in a particular chemical production process consists of first a reaction with an alcohol and then a reaction with a base. There are two alcohols and two bases that can be used in the reaction. The experimenter wished to know if there were any differences with regard to percent yield, and if so, what the optimal alcohol/base combination was. Since time and resources were limited, the experimenter wished to determine the minimum number of experiments

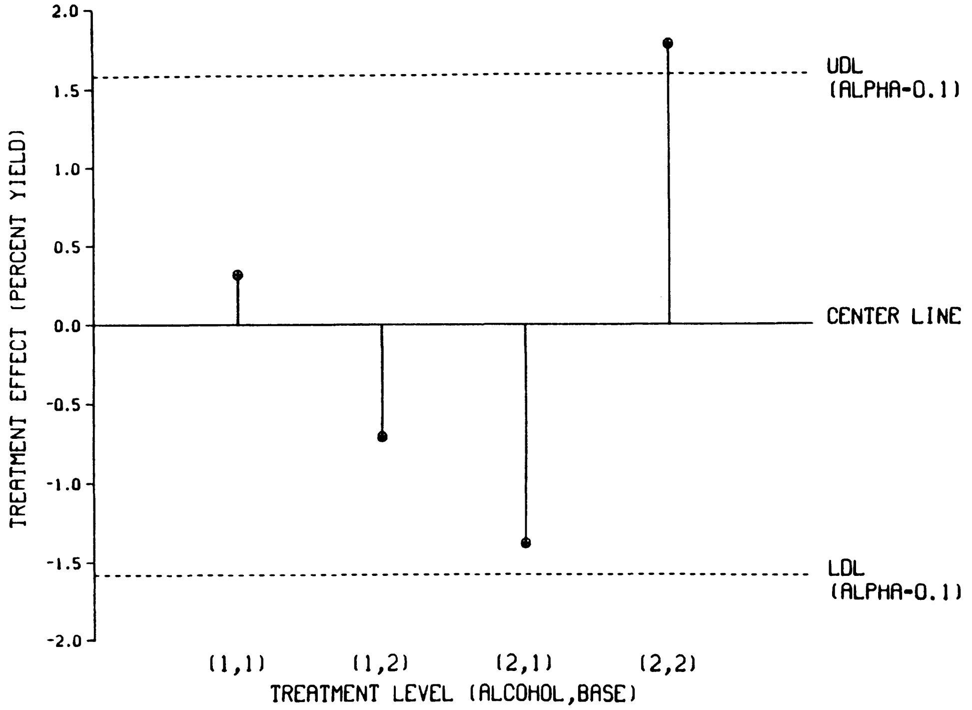


Figure 4. *ANOM Chart for Chemical Yield (Alcohol/Base Effects).*

that could be run and still be assured of detecting differences of reasonable size. Using a significance level of .1, the experimenter was interested in detecting differences in the yields of the four alcohol/base combinations that were at least twice the standard deviation of the experimental error. From Figure 6 one finds that three observations with each combination would result in a power of .49, and four observations with each combination would result in a power of .65. Therefore, a $2 \times 2$ factorial experiment with four replicates per cell was designed and run. The results are recorded in Figure 3. The mean squared error (MSE) for the experiment is 2.209 with 12 degrees of freedom, and there is a significant interaction effect. The interaction necessitates considering the alcohol/base combinations as four populations, and the differences of the four sample means from the grand mean would be compared with $(2.45)\sqrt{2.209}\sqrt{3/16} = 1.58$. Figure 4 shows this comparison graphically, from which one sees that the (2, 2) combination is statistically different and provides an approximate 1.8% increase in yield.

## ACKNOWLEDGMENT

The author thanks Donald L. Crady and Lloyd S. Nelson for assistance in generating the curves and Jason C. Hsu for helpful discussions on majorization.

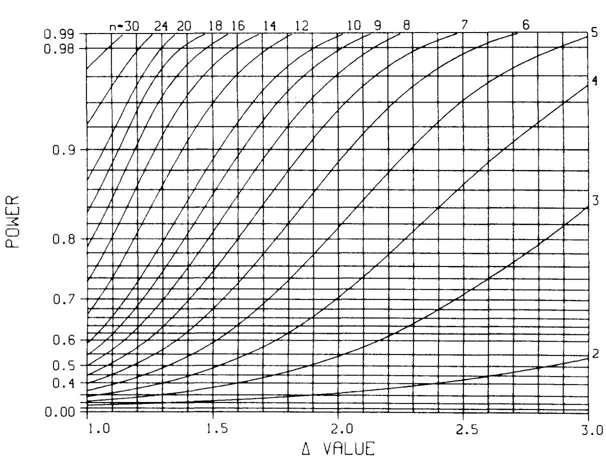


Figure 5. *ANOM Power Curves (K = 3 and α = .10).*

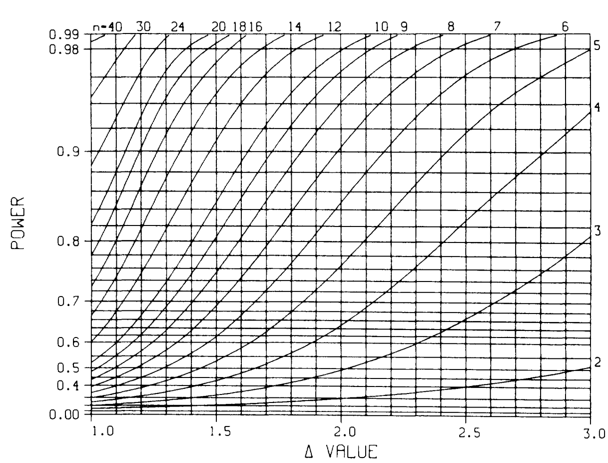


Figure 6. *ANOM Power Curves (K = 4 and α = .10).*

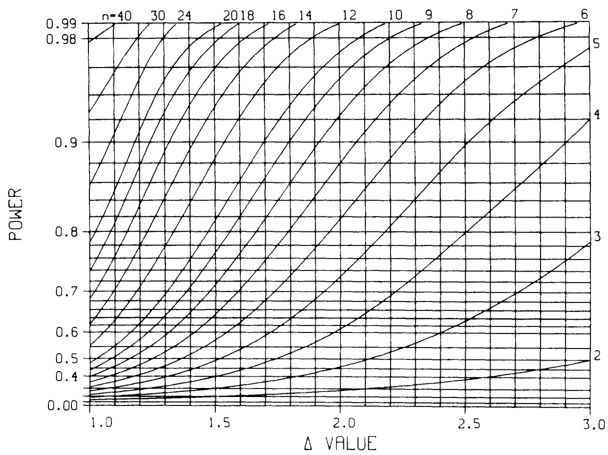


Figure 7. *ANOM Power Curves* ($K = 5$ *and* $\alpha = .10$).

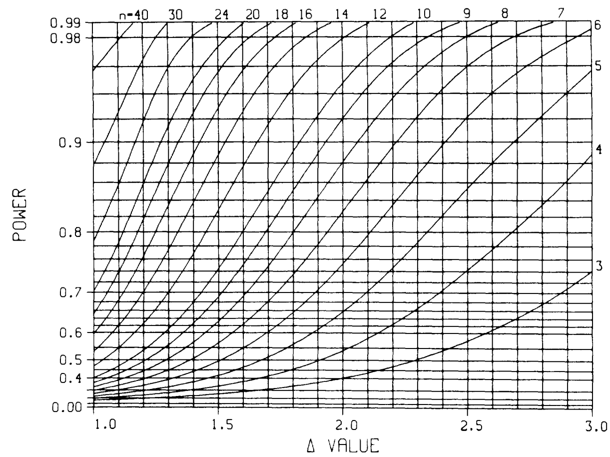


Figure 10. *ANOM Power Curves* ($K = 8$ *and* $\alpha = .10$).

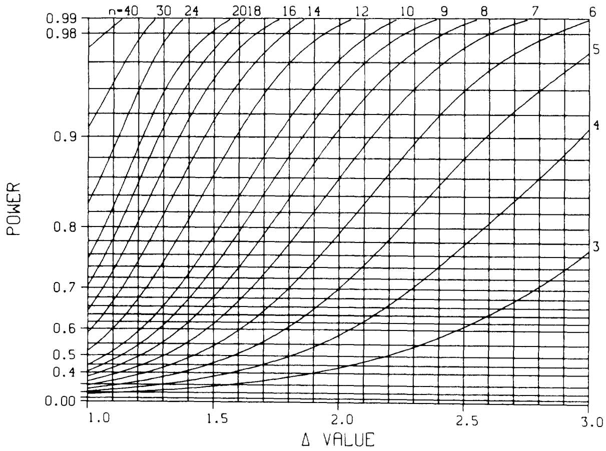


Figure 8. *ANOM Power Curves* ($K = 6$ *and* $\alpha = .10$).

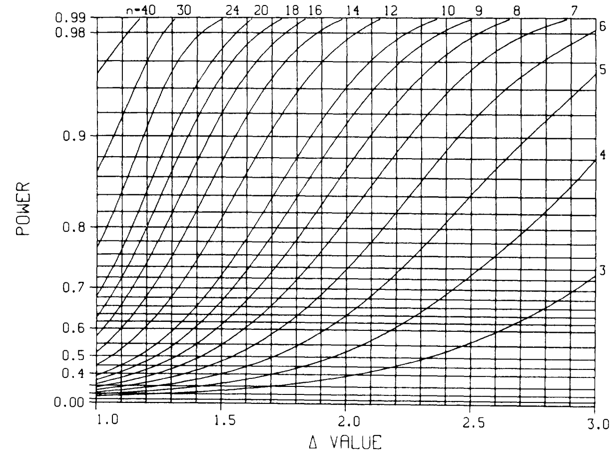


Figure 11. *ANOM Power Curves* ($K = 9$ *and* $\alpha = .10$).

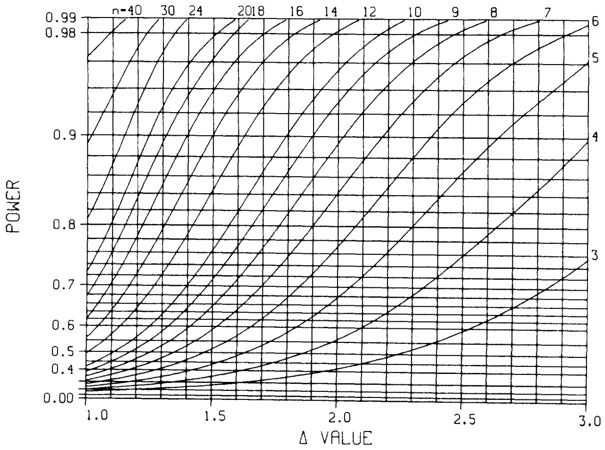


Figure 9. *ANOM Power Curves* ($K = 7$ *and* $\alpha = .10$).

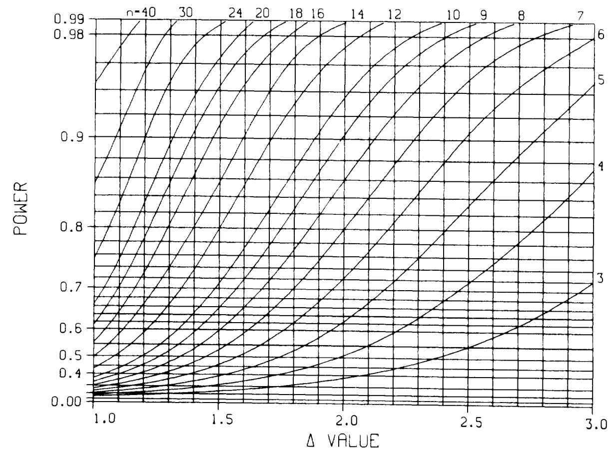


Figure 12. *ANOM Power Curves* ($K = 10$ *and* $\alpha = .10$).

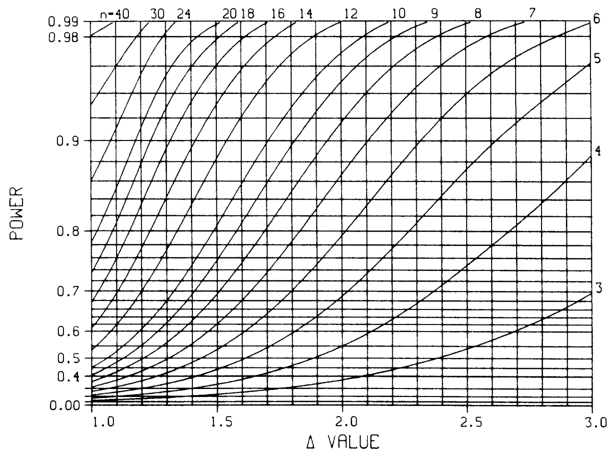


Figure 13. *ANOM Power Curves* ($K = 3$ *and* $\alpha = .05$).

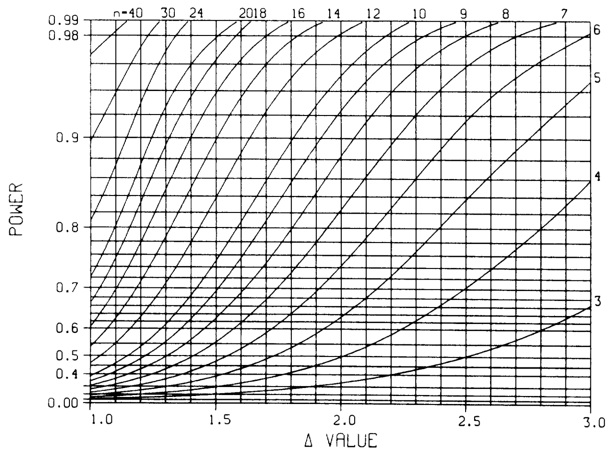


Figure 14. *ANOM Power Curves* ($K = 4$ *and* $\alpha = .05$).

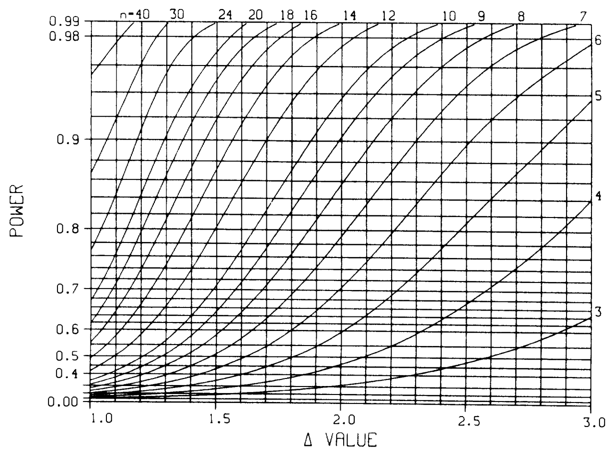


Figure 15. *ANOM Power Curves* ($K = 5$ *and* $\alpha = .05$).

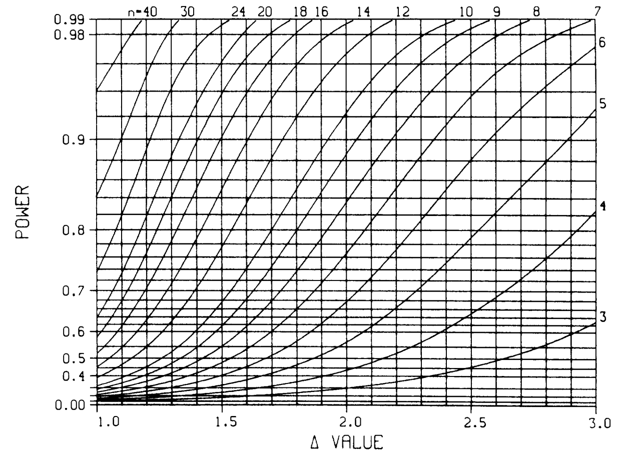


Figure 16. *ANOM Power Curves* ($K = 6$ *and* $\alpha = .05$).

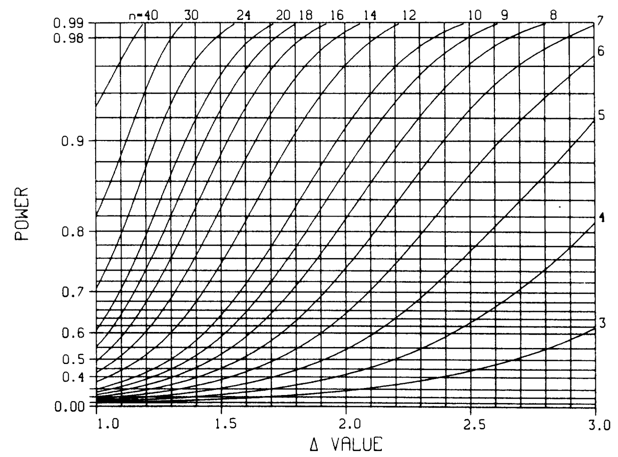


Figure 17. *ANOM Power Curves* ($K = 7$ *and* $\alpha = .05$).

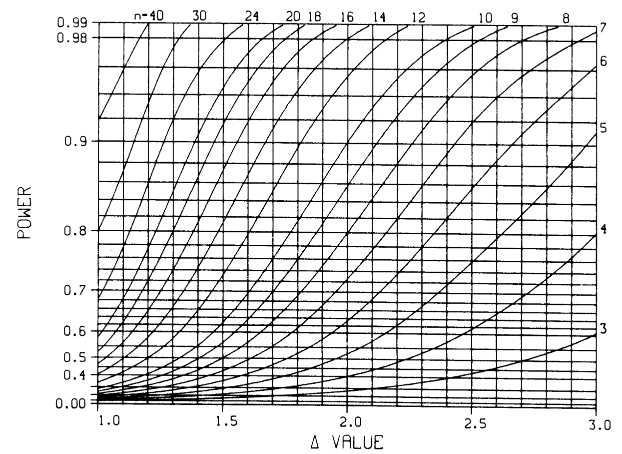


Figure 18. *ANOM Power Curves* ($K = 8$ *and* $\alpha = .05$).

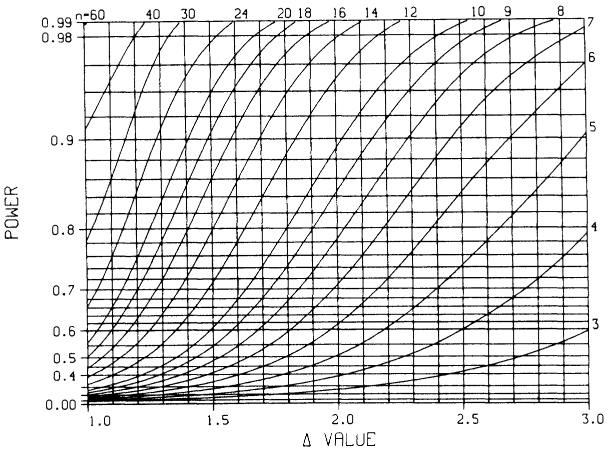


Figure 19. *ANOM Power Curves* ($K = 9$ *and* $\alpha = .05$).

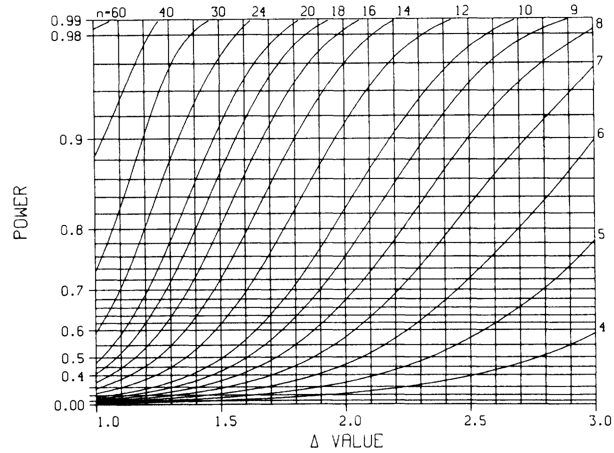


Figure 22. *ANOM Power Curves* ($K = 4$ *and* $\alpha = .01$).

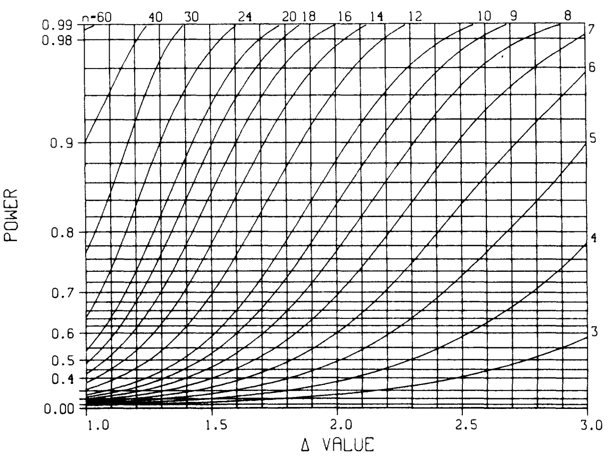


Figure 20. *ANOM Power Curves* ($K = 10$ *and* $\alpha = .05$).

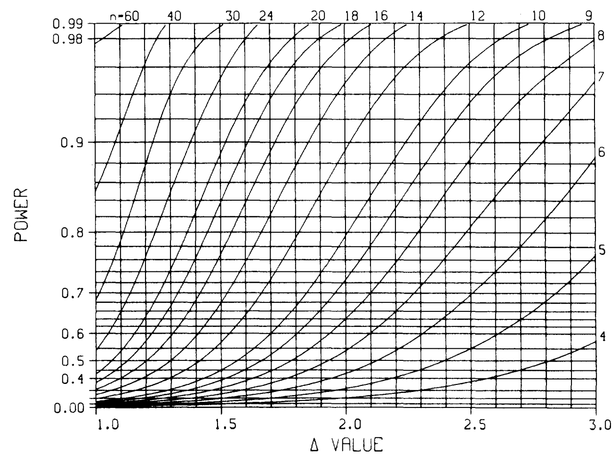


Figure 23. *ANOM Power Curves* ($K = 5$ *and* $\alpha = .01$).

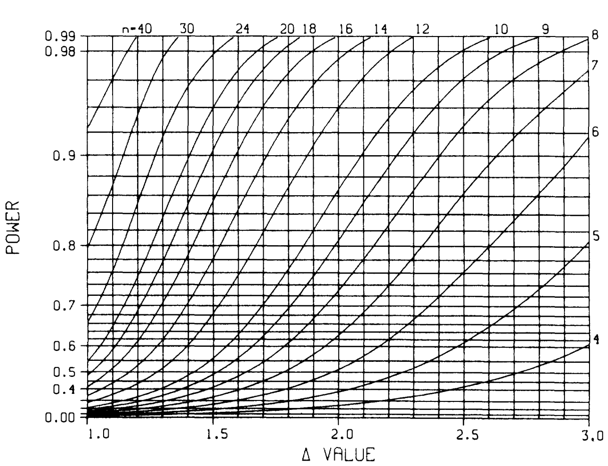


Figure 21. *ANOM Power Curves* ($K = 3$ *and* $\alpha = .01$).

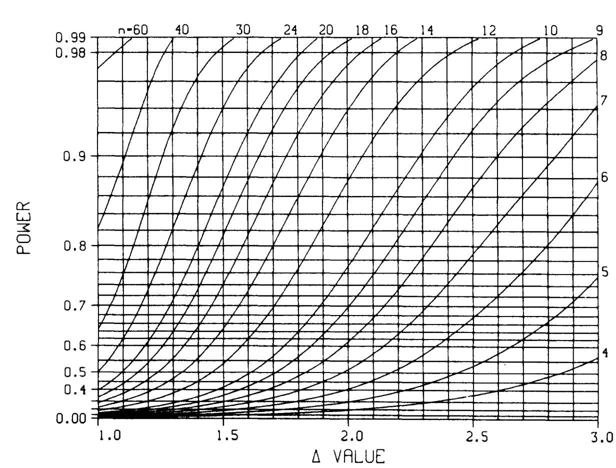


Figure 24. *ANOM Power Curves* ($K = 6$ *and* $\alpha = .01$).

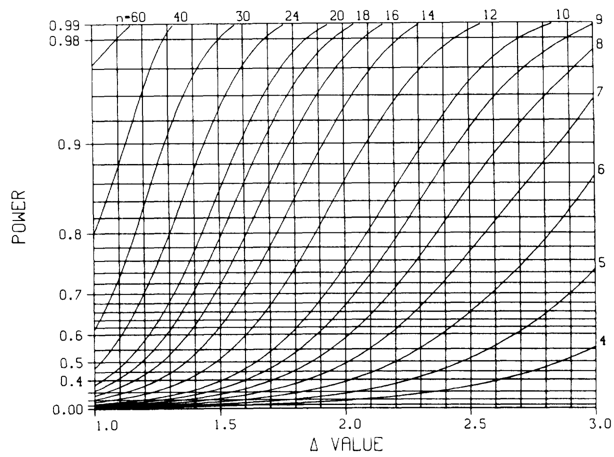


Figure 25. *ANOM Power Curves* ($K = 7$ *and* $\alpha = .01$).

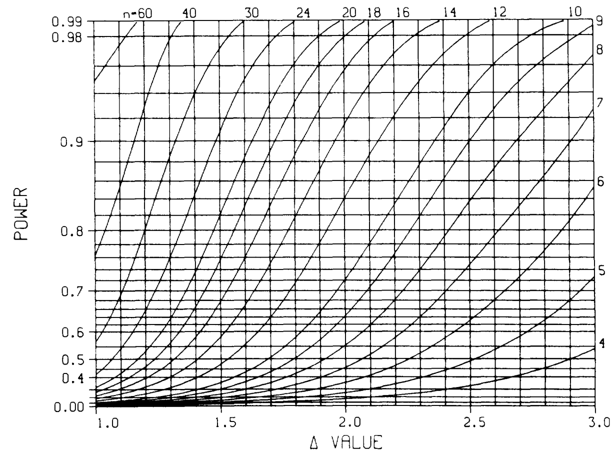


Figure 27. *ANOM Power Curves* ($K = 9$ *and* $\alpha = .01$).

## APPENDIX

The method for showing that the LFC must be of the form (1) is based on the idea of majorization, so the following definitions and results are needed.

*Definition 1.* In $K$ dimensions, a vector $\boldsymbol{\theta}$ is said to majorize a vector $\boldsymbol{\gamma}$ (written $\boldsymbol{\gamma} \prec \boldsymbol{\theta}$) if, when the vector components are ordered such that

$$\theta_{(1)} \geq \theta_{(2)} \geq \cdots \geq \theta_{(K)}$$

and

$$\gamma_{(1)} \geq \gamma_{(2)} \geq \cdots \geq \gamma_{(K)},$$

one has

$$\sum_{i=1}^{k} \gamma_{(i)} \leq \sum_{i=1}^{k} \theta_{(i)}, \quad \text{for } k = 1, \ldots, K,$$

and

$$\sum_{i=1}^{K} \gamma_{(i)} = \sum_{i=1}^{K} \theta_{(i)}.$$

*Definition 2.* A function $f$ is said to be Schur-concave if $\boldsymbol{\gamma} \prec \boldsymbol{\theta}$ implies $f(\boldsymbol{\gamma}) \geq f(\boldsymbol{\theta})$.

*Definition 3.* A set $A \subset \mathscr{R}^K$ is said to be Schur-concave if $\mathbf{y} \in A$ and $\mathbf{x} \prec \mathbf{y}$ implies $\mathbf{x} \in A$. Alternatively, $A$ is Schur-concave if it is a permutationally invariant convex set.

*Proposition 1.* (Marshall and Olkin 1974). If $X_1, \ldots, X_K$ have a joint density $f$ that is Schur-concave and $A$ is a Schur-concave set, then

$$\int_{A+\boldsymbol{\theta}} f(\mathbf{x})\, d\mathbf{x}$$

is a Schur-concave function of $\boldsymbol{\theta}$, where $\{\mathbf{x} \in A + \boldsymbol{\theta}\} \equiv \{\mathbf{x} - \boldsymbol{\theta} \in A\}$.

*Proposition 2.* (Marshall and Olkin 1974). The equicorrelated multivariate central $t$ distribution has a density that is Schur-concave.

We are now in a position to prove the following theorem.

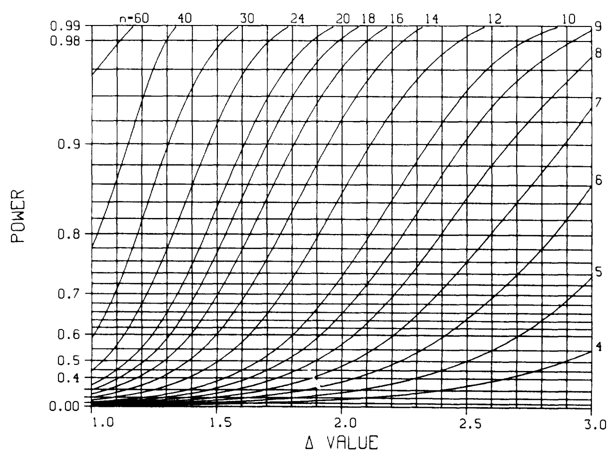


Figure 26. *ANOM Power Curves* ($K = 8$ *and* $\alpha = .01$).

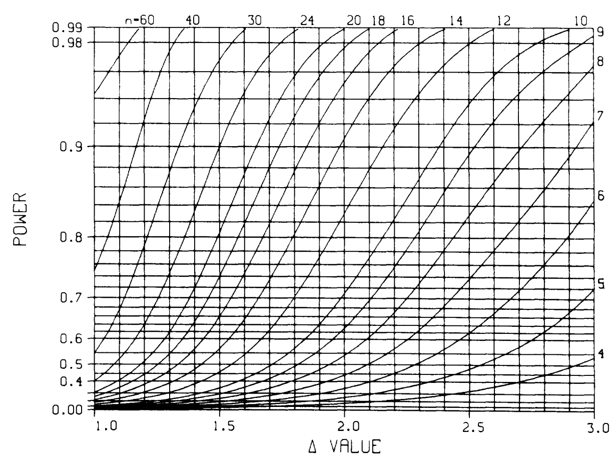


Figure 28. *ANOM Power Curves* ($K = 10$ *and* $\alpha = .01$).

*Theorem 1.* $P(\boldsymbol{\mu}) \geq P(\bar{\boldsymbol{\mu}})$, where $\bar{\boldsymbol{\mu}} = (\mu_{(1)}, \bar{\mu}, \ldots, \bar{\mu}, \mu_{(K)})$ and $\bar{\mu} = 1/(K-2)\sum_{i=2}^{K-1}\mu_{(i)}$.

*Proof.* Let $E = \{(x_1, \ldots, x_K): |x_i| \leq h_\alpha \text{ for all } i\}$, and note that $E$ is Schur-concave. If $f$ is an equicorrelated multivariate central $t$ density, then

$$P(\boldsymbol{\mu}) = 1 - \int_E f(\mathbf{x} + \boldsymbol{\mu}_t)\, d\mathbf{x}$$
$$= 1 - \int_{E+\boldsymbol{\mu}_t} f(\mathbf{x})\, d\mathbf{x},$$

and, from Propositions 1 and 2, it follows that $1 - P(\boldsymbol{\mu})$ is a Schur-concave function of $\boldsymbol{\mu}_t$. Finally, since $\bar{\boldsymbol{\mu}}_t \prec \boldsymbol{\mu}_t$, it follows that $P(\boldsymbol{\mu}) \geq P(\bar{\boldsymbol{\mu}})$. Therefore, we can restrict our attention to $\boldsymbol{\mu} = (\Delta/2, -\Delta/2, \delta, \ldots, \delta)$.

[*Received March 1984. Revised July 1984.*]

## REFERENCES

BENNETT, C. A., and FRANKLIN, N. L. (1954), *Statistical Analysis in Chemistry and the Chemical Industry*, New York: John Wiley.

CRAIG, C. C. (1947), "Control Charts Versus the Analysis of Variance in Process Control by Variables," *Industrial Quality Control*, 3, 4, 14–16.

HALPERIN, M., GREENHOUSE, S. W., CORNFIELD, J., and ZALOKAR, J. (1955), "Tables of Percentage Points for the Studentized Maximum Absolute Deviate in Normal Samples," *Journal of the American Statistical Association*, 50, 185–195.

MARSHALL, A. W., and OLKIN, I. (1974), "Majorization in Multivariate Distributions," *Annals of Statistics*, 2, 1189–1200.

NELSON, L. S. (1983), "Exact Critical Values for Use with the Analysis of Means," *Journal of Quality Technology*, 15, 40–44.

NELSON, P. R. (1982), "Multivariate Normal and $t$ Distributions With $\rho_{jk} = \alpha_j \alpha_k$," *Communications in Statistics, Part B—Simulation and Computation*, *11*, 239–248.

OTT, E. R. (1967), "Analysis of Means—A Graphical Procedure," *Industrial Quality Control*, 24, 101–109.

PEARSON, E. S., and HARTLEY, H. O. (1951), "Charts of the Power Function of the Analysis of Variance Tests, Derived From the Non-Central $F$-Distribution," *Biometrika*, 38, 112–130.

——— (eds.) (1958), *Biometrika Tables for Statisticians* (Vol. 1), New York: Cambridge University Press.

SCHEFFÉ, H. (1947), "The Relationship of Control Charts to Analysis of Variance and Chi-Square Tests," *Journal of the American Statistical Association*, 42, 425–431.

——— (1959), *The Analysis of Variance*, New York: John Wiley.