

Exact Analysis of Means With Unequal Variances

Peter R Nelson & Edward J Dudewicz

To cite this article: Peter R Nelson & Edward J Dudewicz (2002) Exact Analysis of Means With Unequal Variances, *Technometrics*, 44:2, 152-160, DOI: [10.1198/004017002317375109](https://doi.org/10.1198/004017002317375109)

To link to this article: <https://doi.org/10.1198/004017002317375109>



Published online: 01 Jan 2012.



Submit your article to this journal [↗](#)



Article views: 148



View related articles [↗](#)



Citing articles: 6 View citing articles [↗](#)

Exact Analysis of Means With Unequal Variances

Peter R. NELSON

Department of Mathematical Sciences
Clemson University
Clemson, SC 29634-0975
(npeter@clemson.edu)

Edward J. DUDEWICZ

Department of Mathematics
Syracuse University
Syracuse, NY 13244-1150
(dudewicz@mailbox.syr.edu)

The analysis of means (ANOM) is a technique for comparing a group of treatment means to see if any one of them differs significantly from the overall mean. It can be viewed as an alternative to the analysis of variance (ANOVA) for analyzing fixed main effects in a designed experiment. The ANOM has the advantages that it identifies any treatment means that differ from the overall mean (something the ANOVA does not do), and enables a graphical display that aids in assessing practical significance. Sample size tables and power curves have previously been developed for detecting differences among I treatments when two of them differ by at least a specified multiple of the common population standard deviation. Here, we consider the heteroscedastic situation where the different processes or populations from which the samples are drawn do not necessarily have equal standard deviations. In addition, we provide power curves that enable an experimenter to design a study for detecting differences among I treatment means when any two of them differ by at least a specified amount δ , independent of these standard deviations.

KEY WORDS: Comparison of means; Critical values; Power; Unequal variances.

1. INTRODUCTION

The basic ideas of analysis of means (ANOM) date back to Laplace (1827); Ott (1967) developed its graphical form for problems in statistical quality control and statistical process control. Vaughn (1990) confirmed its value for identifying process problems (pp. 239–240), and noted its resemblance to the analysis of variance (ANOVA) (p. 263). The comments of Scheffé (1947) and Craig (1947), originally addressing the comparison of ANOVA and \bar{X} control charts, are applicable to a comparison of the ANOM with the ANOVA. Scheffé (1947) made the analogy that the ANOVA looks to see if a group taken as a whole displays any evidence of criminality, whereas the ANOM looks to see if the group harbors a criminal, and provides a method for making an identification. The ANOM is, therefore, more sensitive for detecting a single different mean. An example of this is given by Pearson and Hartley (1958, p. 51). However, as Craig (1947, p. 16) commented, “the two methods are nearly enough equivalent that both will disclose any clear lack of [equality] among averages.”

The traditional assumptions for ANOM are exactly the same as for traditional ANOVA, namely, that the observations are normally distributed and independent with means that depend on the particular treatment and common variance σ^2 . With only two treatments, ANOM and ANOVA are equivalent. While ANOM is not an optimal test in any known mathematical sense, it has the advantages that it identifies any treatment means that differ from the overall mean, and provides a graphical display that aids in assessing practical significance. Further, when using exact critical values, no price is paid for these advantages in terms of decreased power. Nelson (1983) showed that, with the same power, ANOM and ANOVA sample sizes differ by at most 2 over a broad range of parameter values for up to 11 treatments; for larger numbers of treatments, ANOVA requires substantially larger sample sizes than does ANOM.

Robustness of procedures has been a concern of statisticians for a long time. Scheffé (1959, sec. 10.6) considered robustness for ANOVA. Vaughn (1990, p. 245) noted how one might use control charts to check for outliers (a common problem, which calls for consideration), and the need “. . . to proceed no further with the ANOM until the reason for the outlier had been found.” Similar advice applies to ANOVA.

This article generalizes ANOM to the heteroscedastic situation where the population or process variances are not necessarily equal. Power curves for heteroscedastic analysis of means (HANOM) allow an experimenter to set a goal of detecting differences among the I treatment means when two of the treatment means differ by a specified amount δ that does not depend on these variances.

2. THE ONE-WAY LAYOUT

2.1 HANOM Procedure

Let I be the number of treatments being compared, and let X_{ij} be the j th observation from the i th population with $i = 1, \dots, I$; $j = 1, \dots, n_i$. We assume $X_{ij} = \mu_i + \epsilon_{ij}$, $\epsilon_{ij} \sim N(0, \sigma_i^2)$, and all of the observations are independent. Collecting the data for and performing a HANOM consists of the following steps.

1. Take a sample of initial size $n_0 (\geq 2)$ from each of the I populations, and calculate the sample means \bar{X}_{0i} and the sample variances s_i^2 .
2. Specify the level of significance α , a power γ , and the amount δ as the difference between any two treatment means that will lead to rejection of the null hypothesis (all of the

treatment means are equal) with power γ . Given α, I, γ , and $df = n_0 - 1$, find the corresponding value of w from graphs of the power curves. Figure A.1 is one example of these graphs, and others are provided in Dudewicz and Nelson (2002). Alternatively, one could choose a value of w [using Eq. (1) below] based on desired sample sizes.

3. Compute

$$n_i = \max\{n_0 + 1, \lceil (w/\delta)^2 s_i^2 \rceil + 1\} \quad (1)$$

for each i , where $\lceil y \rceil$ denotes the greatest integer in y , and take $n_i - n_0$ additional observations $X_{i, n_0+1}, \dots, X_{i, n_i}$ from population i .

4. For each i , calculate the sample mean of the second set of observations from population i :

$$\bar{X}_i = \sum_{j=n_0+1}^{n_i} X_{ij} / (n_i - n_0). \quad (2)$$

5. For each i , compute

$$b_i = \frac{n_i - n_0}{n_i} \left[1 + \sqrt{\left(\frac{n_0}{n_i - n_0} \right) \left(\left[\frac{\delta}{w} \right]^2 \frac{n_i}{s_i^2} - 1 \right)} \right] \quad (3)$$

$$\tilde{\bar{X}}_i = (1 - b_i) \bar{X}_{0i} + b_i \bar{X}_i \quad (4)$$

$$\tilde{\bar{X}}_{\bullet} = \sum_{i=1}^I \tilde{\bar{X}}_i / I. \quad (5)$$

6. Compute decision lines

$$\begin{aligned} \text{UDL} &= \tilde{\bar{X}}_{\bullet} + \mathcal{H}(\alpha; I, n_0 - 1) \frac{\delta}{w} \\ \text{LDL} &= \tilde{\bar{X}}_{\bullet} - \mathcal{H}(\alpha; I, n_0 - 1) \frac{\delta}{w} \end{aligned} \quad (6)$$

where $\mathcal{H}(\alpha; I, n_0 - 1)$ is found in Table A.1, and reject the hypothesis $H_0: \mu_1 = \dots = \mu_I$ if any of the $\tilde{\bar{X}}_i$'s falls outside these decision lines.

A numerical example with a biostatistics data set is given in Dudewicz and Nelson (2002).

2.2 Determination of Critical Values and Power

Denote by A_i the event

$$\frac{|\tilde{\bar{X}}_i - \tilde{\bar{X}}_{\bullet}|}{\delta/w} > \mathcal{H}(\alpha; I, n_0 - 1).$$

Rejecting the hypothesis $H_0: \mu_1 = \dots = \mu_I$ if any of the $\tilde{\bar{X}}_i$'s falls outside the decision lines (6) is equivalent to rejecting H_0 if at least one of the events A_i occurs. Therefore, the power of the HANOM is

$$\begin{aligned} \text{Power}(\boldsymbol{\mu}) &= \Pr(A_1 \cup \dots \cup A_I \mid \boldsymbol{\mu}) \\ &= 1 - \Pr(\bar{A}_1 \cap \dots \cap \bar{A}_I \mid \boldsymbol{\mu}) \\ &= 1 - \Pr \left(\frac{|\tilde{\bar{X}}_i - \tilde{\bar{X}}_{\bullet}|}{\delta/w} \leq \mathcal{H}(\alpha; I, n_0 - 1), \right. \\ &\quad \left. \text{for all } i = 1, \dots, I \mid \boldsymbol{\mu} \right) \end{aligned}$$

$$\begin{aligned} &= 1 - \Pr \left(-\mathcal{H}(\alpha; I, n_0 - 1) \right. \\ &\quad \leq \frac{(\tilde{\bar{X}}_i - \mu_i - (\tilde{\bar{X}}_{\bullet} - \mu_{\bullet})) + (\mu_i - \mu_{\bullet})}{\delta/w} \\ &\quad \left. \leq \mathcal{H}(\alpha; I, n_0 - 1), \text{ for all } i = 1, \dots, I \right) \end{aligned}$$

where $\mu_{\bullet} = (\mu_1 + \dots + \mu_I) / I$. From Stein (1945), we know that

$$T_i = \frac{\tilde{\bar{X}}_i - \mu_i}{\delta/w} \sim t(n_0 - 1)$$

where the T_i are all independent, and therefore,

$$\begin{aligned} \text{Power}(\boldsymbol{\mu}) &= 1 - \Pr \left(-\mathcal{H}(\alpha; I, n_0 - 1) \leq T_i - \bar{T} + \frac{w}{\delta} (\mu_i - \mu_{\bullet}) \right. \\ &\quad \left. \leq \mathcal{H}(\alpha; I, n_0 - 1), \text{ for all } i = 1, \dots, I \right) \quad (7) \end{aligned}$$

where

$$\bar{T} = \frac{\tilde{\bar{X}}_{\bullet} - \mu_{\bullet}}{\delta/w} = \sum_{i=1}^I T_i / I.$$

2.2.1 Critical Values. When H_0 is true,

$$\mu_1 = \dots = \mu_I = \mu_{\bullet}.$$

and

$$\alpha = 1 - \Pr(-\mathcal{H}(\alpha; I, n_0) \leq T_i - \bar{T} \leq \mathcal{H}(\alpha; I, n_0), \text{ for all } i = 1, \dots, I). \quad (8)$$

The critical values in Table A.1 were obtained by simulating the distribution of

$$(T_1 - \bar{T}, \dots, T_I - \bar{T}) \quad (9)$$

to obtain the probability on the right-hand side in Equation (8) and searching for the value of \mathcal{H} that resulted in the desired α . The details of the simulation are discussed in a later section.

2.2.2 The Power Function. The power function (7) for HANOM, like that for ANOM and ANOVA, depends on the particular configuration of the μ_i 's. Thus, to provide power curves, we consider certain subspaces of the μ_i 's. For HANOM, reasonable subspaces are

$$M_{\delta} = \left\{ \boldsymbol{\mu} = \{\mu_1, \dots, \mu_I\} : \max_{i,j} |\mu_i - \mu_j| \geq \delta \right\}$$

and the power is then determined using the least favorable configuration (LFC) of means on these subspaces. Appendix B shows that the LFC on M_{δ} for HANOM is of the form

$$\boldsymbol{\mu} = (\delta/2, -\delta/2, 0, \dots, 0). \quad (10)$$

The power value curves in Figure A.1 were computed in a similar fashion to the critical values. The distribution of the random vector (9) was simulated to obtain the probability on the right-hand side of Equation (7) using the LFC for $\boldsymbol{\mu}$. For brevity, we present only one, typical, set of power curves for $\alpha = .1$ and $I = 12$; a larger collection for $\alpha = .1, .05, .01$ and $I = 2(1)12, 15, 20$ is given in Dudewicz and Nelson (2002).

2.2.3 The Simulations. The distribution of the random vector (9) was simulated with I independent t random variables using the method described by Kinderman and Monahan (1980). This algorithm requires uniform random variables, which were generated using the Lehmer sequence

$$x_{i+1} = 630360016x_i \bmod(2^{31} - 1) \quad (11)$$

(see Karian and Dudewicz 1999). To avoid inaccuracies in multiplying numbers with many significant digits, the sequence (11) was actually computed using

$$x_{i+1} = \{[19591x_i \bmod(2^{31} - 1)](32176)\} \bmod(2^{31} - 1).$$

Critical values $\mathcal{H}(\alpha; I, \nu)$ were found by simulating the probability on the right-hand side of Equation (8) for each fixed I and ν combination using 10^6 trials, repeating for various values of \mathcal{H} until the desired probability of $1 - \alpha$ was obtained.

In two special cases, these critical values can be computed directly. When $I = 2$, the random vector (9) reduces to the average of two independent t random variables, whose distribution is the convolution of a t distribution function and a t density function. This single numerical integration was performed using the program described in Craig (1984) for the t distribution function (accurate to 16 decimal places) and Romberg integration. When $\nu = \infty$, the HANOM critical values are $\sqrt{(I-1)/I}$ times the corresponding critical value for the ANOM. That is,

$$\mathcal{H}(\alpha; I, \infty) = \sqrt{\frac{I-1}{I}} h(\alpha; I, \infty).$$

We verified the accuracy of the simulated critical values by comparing them with the values computed using the alternative methods, and found agreement to the number of decimal places reported in the tables. Accuracy was less of a problem for the power values since the standard deviation of the power estimates with 10^6 trials was less than .0005.

3. HIGHER ORDER LAYOUTS

In designs with more than one factor, one must first test for possible interaction(s) among the factors. As in the homogeneous variance case, this is best done using ANOVA. We will consider here only a two-way layout. Details on heteroscedastic ANOVA for a two-way layout can be found in Bishop and Dudewicz (1978), and still higher order layouts are discussed in Bishop and Dudewicz (1981).

The usual model for the two-way layout is

$$X_{ijk} = \mu + \alpha_i^A + \alpha_j^B + \alpha_{ij}^{AB} + \epsilon_{ijk}$$

where factor A has I levels and factor B has J levels. In the heteroscedastic case, one assumes that $\epsilon_{ijk} \sim N(0, \sigma_{ij}^2)$. Initially, one would take a sample of size $n_0 \geq 2$ from each of the IJ treatment combinations, and compute the usual unbiased estimates s_{ij}^2 of the σ_{ij}^2 and the averages of the observations for each treatment combination \bar{X}_{0ij} . If it is known a priori that the σ_{ij}^2 are not all equal, or if a test on the variances such as the analysis of means for variances (ANOMV) (see Wludyka and Nelson 1997) indicates that the variances are not all equal, then heteroscedastic ANOVA and HANOM

are appropriate. Sample sizes at the second stage are computed using Equation (1) where s_i is replaced by s_{ij} and n_i is replaced by n_{ij} , namely,

$$n_{ij} = \max\{n_0 + 1, \lceil[(w/\delta)^2 s_{ij}^2] + 1\rceil\}. \quad (12)$$

An appropriate value for w depends on whether there is significant interaction, and what one is interested in comparing. For fixed α , γ , and δ , the value of w increases as the number of means being compared increases. Thus, a conservative choice for w would be based on comparing IJ means, which is the situation that would occur if one found significant interaction and were then interested in comparing the IJ treatment combinations. Alternatively, if one were interested in the main effects of a factor or if it were known a priori that there was no interaction, then it would make sense to choose w based on $\max(I, J)$.

The heteroscedastic ANOVA test of $H_{AB} : \alpha_{ij}^{AB} = 0$ for all (i, j) is conducted using essentially the usual interaction sum of squares with \bar{X}_{ij} replaced by $\tilde{\bar{X}}_{ij}$. Specifically, for each (i, j) , one would compute

$$b_{ij} = \frac{n_{ij} - n_0}{n_{ij}} \left[1 + \sqrt{\left(\frac{n_0}{n_{ij} - n_0}\right) \left(\left[\frac{\delta}{w}\right]^2 \frac{n_{ij}}{s_{ij}^2} - 1\right)} \right] \quad (13)$$

$$\tilde{\bar{X}}_{ij} = (1 - b_{ij}) \bar{X}_{0ij} + b_{ij} \bar{X}_{ij} \quad (14)$$

$$\tilde{\bar{X}}_{i\bullet} = \frac{1}{J} \sum_{j=1}^J \tilde{\bar{X}}_{ij} \quad (15)$$

$$\tilde{\bar{X}}_{\bullet j} = \frac{1}{I} \sum_{i=1}^I \tilde{\bar{X}}_{ij} \quad (16)$$

$$\tilde{\bar{X}}_{\bullet\bullet} = \frac{1}{IJ} \sum_{i=1}^I \sum_{j=1}^J \tilde{\bar{X}}_{ij}. \quad (17)$$

The hypothesis H_{AB} is tested using

$$\tilde{F}_{AB} = \left(\frac{w}{\delta}\right)^2 \sum_{i=1}^I \sum_{j=1}^J (\tilde{\bar{X}}_{ij} - \tilde{\bar{X}}_{i\bullet} - \tilde{\bar{X}}_{\bullet j} + \tilde{\bar{X}}_{\bullet\bullet})^2 \quad (18)$$

which is compared with the appropriate quantile of an $((n_0 - 1)/(n_0 - 3))\chi_{(I-1)(J-1)}^2$ distribution.

If no significant interaction is found, then the main effects of factor A can be tested by comparing the $\tilde{\bar{X}}_{i\bullet}$ with the decision lines

$$\begin{aligned} \text{UDL} &= \tilde{\bar{X}}_{\bullet\bullet} + \mathcal{H}(\alpha; I, n_0 - 1) \frac{\delta}{w} \\ \text{LDL} &= \tilde{\bar{X}}_{\bullet\bullet} - \mathcal{H}(\alpha; I, n_0 - 1) \frac{\delta}{w}. \end{aligned} \quad (19)$$

Similarly, the main effects of factor B can be tested by comparing the $\tilde{\bar{X}}_{\bullet j}$ with the decision lines

$$\begin{aligned} \text{UDL} &= \tilde{\bar{X}}_{\bullet\bullet} + \mathcal{H}(\alpha; J, n_0 - 1) \frac{\delta}{w} \\ \text{LDL} &= \tilde{\bar{X}}_{\bullet\bullet} - \mathcal{H}(\alpha; J, n_0 - 1) \frac{\delta}{w}. \end{aligned} \quad (20)$$

If significant interaction is found, then one can either compare the IJ treatment combinations using the $\tilde{\bar{X}}_{ij}$ and

Table 1. Summary Statistics From the First-Stage and Second-Stage Samples

<i>i</i>	<i>j</i>	First stage			Second stage			Overall
		n_0	\bar{X}_{0ij}	s_{ij}^2	n_{ij}	\bar{X}_{ij}	b_{ij}	\bar{X}_{ij}
1	1	6	5.083	2.2496	7	2.3	.5543	3.541
1	2	6	3.467	1.8827	7	1.1	.6184	2.003
1	3	6	5.250	5.8270	8	7.1	.3479	5.894
1	4	6	8.683	2.3496	7	9.0	.5389	8.854
2	1	6	4.250	2.1870	7	4.1	.5643	4.165
2	2	6	2.250	1.2590	7	1.7	.7744	1.824
2	3	6	6.850	5.6190	8	2.35	.3799	5.140
2	4	6	9.367	3.2546	7	10.9	.4243	10.017
3	1	6	3.217	3.1576	7	7.5	.4351	5.080
3	2	6	1.333	1.2547	7	3.2	.7758	2.782
3	3	6	7.433	14.4386	19	5.338	.7245	5.915
3	4	6	9.067	1.8067	7	9.4	.6336	9.278

decision lines

$$\begin{aligned} \text{UDL} &= \bar{X}_{..} + \mathcal{H}(\alpha; IJ, n_0 - 1) \frac{\delta}{w} \\ \text{LDL} &= \bar{X}_{..} - \mathcal{H}(\alpha; IJ, n_0 - 1) \frac{\delta}{w} \end{aligned} \quad (21)$$

or (assuming $I > J$) compare the I levels of factor A separately for each level j' of factor B using the $\bar{X}_{ij'}$ and decision lines

$$\begin{aligned} \text{UDL} &= \bar{X}_{.j'} + \mathcal{H}(\alpha; I, n_0 - 1) \frac{\delta}{w} \\ \text{LDL} &= \bar{X}_{.j'} - \mathcal{H}(\alpha; I, n_0 - 1) \frac{\delta}{w}. \end{aligned} \quad (22)$$

Example: An experiment was conducted to study the effects at three different temperatures (factor A) of four different types of insulation (factor B) on the insulation's ability to maintain a fixed temperature. Initial samples of size 6 were taken for each of the 12 treatment combinations. The values in Table A.2 are the rise in temperature ($^{\circ}\text{F}$) after a fixed amount of time. The sample means and variances are given in Table 1. The ANOMV chart given in Figure 1 shows

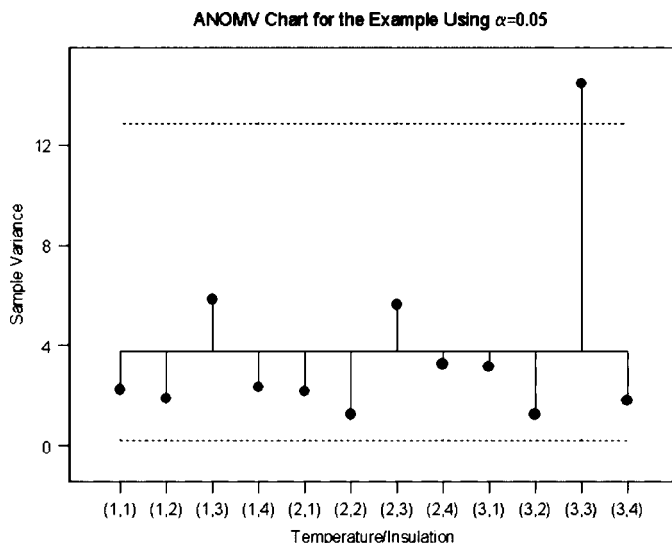


Figure 1. ANOMV Chart for the Example.

Table 2. Average Outputs for Temperature/Insulation Combinations Together With Row Means, Column Means, and the Grand Mean

Insulation	Temperature			
	1	2	3	
1	3.540	4.165	5.080	4.262
2	2.003	1.824	2.782	2.203
3	5.894	5.140	5.915	5.650
4	8.854	10.017	9.278	9.383
	5.073	5.286	5.764	5.374

that, at the .05 level, the σ_{ij}^2 are not all equal, indicating that heteroscedastic ANOVA and HANOM are appropriate.

The experimenters were interested in conducting tests using $\alpha = .1$, and being able to detect any effect of $\delta = 7$ or more with a power of .8. From Figure A.1 ($\alpha = .1$ and $I = 12$), one finds that, for a power of .8 and $df = 5$, the value of w is approximately 8. Using Equation (12), the sample sizes for the second stage were computed, and are given in Table 1. The second-stage data are also given in Table A.2, and the statistics computed using Equations (13)–(17) are given in Table 1. For convenience, the row means, column means, and grand mean (of the \bar{X}_{ij}) are given in Table 2. Using Equation (18), one then would compute $\tilde{F}_{AB} = (8/7)^2(1.7994) = 2.350$ which, when compared with $(5/3)\chi^2(.05; 6) = (5/3)(12.592) = 21.0$, is found to be not significant. Thus, the main effect of the two factors can be studied using HANOM.

For factor A , one computes the decision lines as in (19):

$$\begin{aligned} \text{UDL} &= 5.374 + \mathcal{H}(.1; 3, 5)(7/8) \\ &= 5.374 + (2.16)(7/8) = 7.264 \\ \text{LDL} &= 5.374 - \mathcal{H}(.1; 3, 5)(7/8) \\ &= 5.374 - (2.16)(7/8) = 3.484. \end{aligned}$$

The HANOM chart in Figure 2 suggests no effect due to the temperature at the .1 level. Similarly, for factor B , one obtains

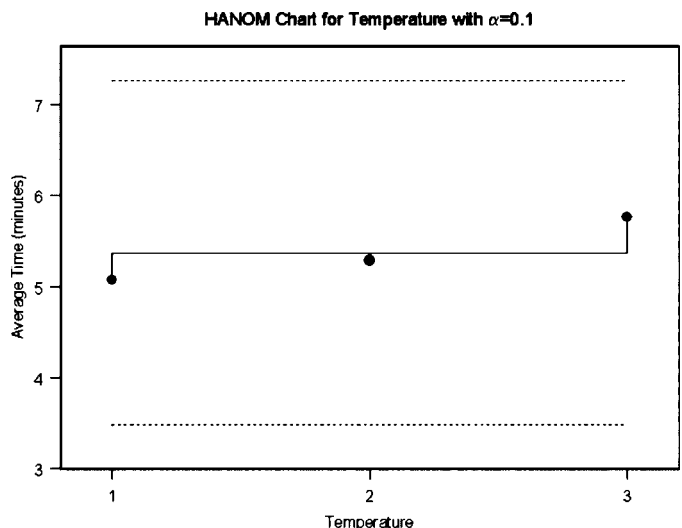


Figure 2. HANOM Chart for Temperature.

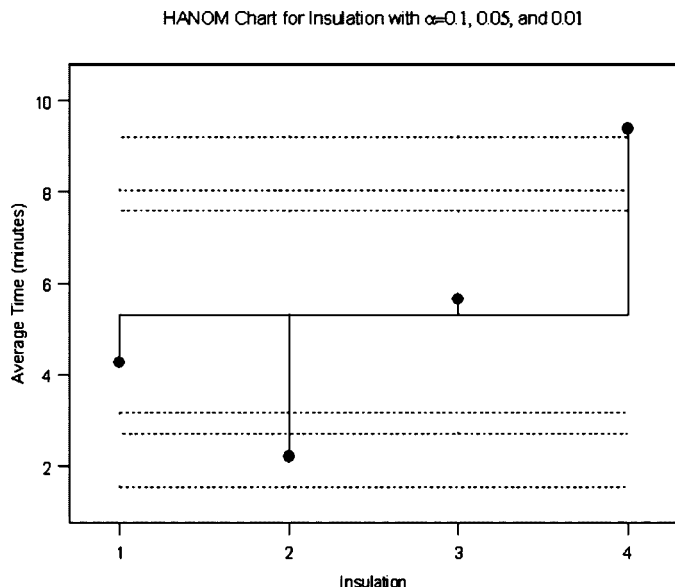
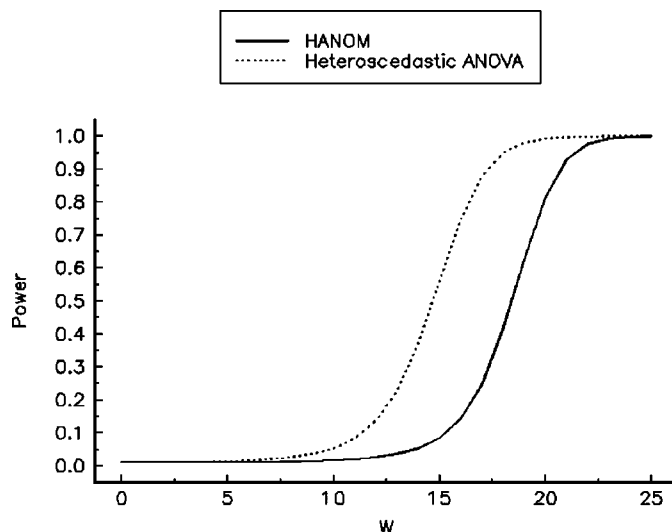


Figure 3. HANOM Chart for Insulation.

Figure 4. Power Curves for HANOM and Heteroscedastic ANOVA When $I = 20$, $n_0 = 5$, and $\alpha = .01$.

the HANOM chart in Figure 3, from which one sees that the insulations have an effect at the .01 level.

4. COMPARISON WITH HETEROSCEDASTIC ANOVA

Since the critical points and powers reported in Bishop and Dudewicz (1978) for heteroscedastic ANOVA were based on simulations using only 10^4 trials, a more fair comparison with HANOM required resimulated critical points and power values using 10^6 trials. As in the homogeneous case, heteroscedastic ANOVA and HANOM are equivalent when $I = 2$, and critical values for heteroscedastic ANOVA are $2[\mathcal{H}(\alpha; 2, \nu)]^2$. For more than two means, a comparison of the powers for the two procedures was made. It showed (for full details, see Dudewicz and Nelson 2001) that neither procedure is uniformly better, and with one exception, the differences are not of practical significance. When $I = 20$, the differences are the most extreme. For $I = 20$ and $n_0 = 25$, HANOM becomes more powerful when powers above about .7 are desired. However, for $I = 20$ and $n_0 = 5$, heteroscedastic ANOVA is much more powerful. The biggest difference occurs when $I = 20$, $n_0 = 5$, and $\alpha = .01$ (see Table 3); when the variability is large enough to require second samples

greater than 1, one would need about 1.5 times as many second-stage observations with HANOM for equivalent power. The key quantity for calculating this ratio of sample sizes is w^2 [see eq. (1) for HANOM; the equation for HANOVA is similar; and note that when a second sample greater than 1 is required, $(ws/\delta)^2 > n_0$, so the 1 in eq. (1) can be ignored]. From Figure 4, for a specific power (such as .7), one finds that, for HANOM, $w = 19$, for HANOVA, $w = 15.5$, and $(19/15.5)^2 = 1.50$.

As noted in the Introduction, the comparison of ANOM and ANOVA is well established in the homoscedastic case (Nelson 1983). These results can be regarded as similar. An Associate Editor noted that, "ANOM is relatively well established among some audiences . . . I'd stop short of presenting it as an omnibus test that competes with ANOVA, and instead emphasize its descriptive advantages." We concur. While the two procedures *can* test the same hypotheses, and with relatively comparable sample sizes and power, our results say to statisticians and practitioners the following.

- If you have been using ANOM (or, now, use HANOM), you have nothing to gain from ANOVA (or HANOVA) in terms of sample size or power, so stay with ANOM (or HANOM) for the graphical advantages.

Table 3. Maximum Power Difference Between HANOM and Heteroscedastic ANOVA.
Power for HANOM (HANOM Power Minus Heteroscedastic ANOVA Power $\times 10^3$)

I	$n_0 = 5$			$n_0 = 15$			$n_0 = 25$		
	α			α			α		
	.10	.05	.01	.10	.05	.01	.10	.05	.01
3	.661(−15)	.600(−27)	.526(−95)	.769(−6)	.713(−9)	.638(−27)	.733(−7)	.727(−9)	.667(−21)
4	.561(−39)	.561(−69)	.617(−172)	.626(−20)	.546(−26)	.565(−57)	.661(−16)	.598(−20)	.563(−43)
5	.504(−67)	.491(−109)	.407(−276)	.517(−28)	.547(−38)	.524(−74)	.581(−24)	.584(−28)	.536(−55)
10	.442(−152)	.415(−226)	.288(−486)	.449(−48)	.413(−57)	.469(−104)	.365(−34)	.388(−37)	.456(−62)
20	.386(−218)	.362(−321)	.209(−635)	.310(−55)	.280(−55)	.391(−93)	.252(−32)	.191(−30)	.225(−31)

• If you have been using ANOVA (or HANOVA), you are not wasting sample size or power compared with ANOM (or HANOM). However, you could gain in communication of results by the use of ANOM (or HANOM).

5. CONCLUSIONS

The usual procedures for testing to see if a group of treatment means contains any means that are different require the populations being compared to have equal variances. When the variances are not all equal, and it is either inconvenient or impossible to transform the data to make the variances equal, heteroscedastic procedures are needed. We have presented a heteroscedastic form of ANOM that allows an experimenter to set a goal of detecting differences among I treatment means when two of them differ by at least a specified amount δ , which does not depend on the variances of the populations.

A comparison of HANOM with heteroscedastic ANOVA indicates that neither procedure is uniformly better, and the

differences in power are not great enough to make a practical difference, except when I is large and both n_0 and α are small ($I = 20$, $n_0 = 5$, and $\alpha = .01$), and there is enough variability that second samples of more than one observation are needed. However, if one anticipated large variability and wanted to use $\alpha = .01$, initial samples larger than $n_0 = 5$ would generally be chosen. HANOM also has the same advantages as the ANOM; namely, it points out which means are different from the grand mean if any of them is, and it makes it easy to judge the practical significance of any differences.

ACKNOWLEDGMENTS

This research was begun while E. J. Dudewicz was Visiting Distinguished Professor at Clemson University, whose support is gratefully acknowledged. We thank the Editor (Karen Kafadar), Associate Editor, and two referees for their helpful comments on presentation, which we believe have helped the article's clarity and crispness.

APPENDIX A: CRITICAL VALUES

Table A.1. HANOM Critical Values $\mathcal{H}(\alpha; k, \nu)$ Where $\alpha = .1$ and $\nu = n_0 - 1$

	Number of means being compared, k																			
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
ν																				
1	6.32	12.5	18.5	24.5	30.6	36.7	42.9	48.9	54.9	61.0	66.9	73.0	79.0	85.1	91.0	97.2	103.	109.	115.	
2	2.28	3.69	4.69	5.56	6.32	7.01	7.65	8.23	8.78	9.30	9.79	10.3	10.7	11.1	11.6	11.9	12.3	12.7	13.1	
3	1.75	2.68	3.24	3.70	4.08	4.42	4.71	4.98	5.23	5.46	5.67	5.87	6.05	6.22	6.39	6.55	6.70	6.85	6.99	
4	1.55	2.33	2.76	3.10	3.37	3.61	3.81	3.99	4.16	4.30	4.44	4.57	4.69	4.80	4.91	5.00	5.10	5.19	5.27	
5	1.45	2.16	2.53	2.81	3.04	3.23	3.39	3.53	3.66	3.78	3.88	3.98	4.07	4.16	4.24	4.32	4.39	4.45	4.52	
6	1.39	2.06	2.40	2.65	2.85	3.01	3.15	3.27	3.38	3.48	3.57	3.65	3.73	3.80	3.87	3.93	3.99	4.05	4.10	
7	1.36	1.99	2.31	2.54	2.72	2.87	3.00	3.11	3.20	3.29	3.37	3.45	3.52	3.58	3.64	3.69	3.74	3.79	3.84	
8	1.33	1.94	2.25	2.46	2.63	2.77	2.89	2.99	3.08	3.16	3.23	3.30	3.36	3.42	3.47	3.52	3.57	3.61	3.66	
9	1.31	1.91	2.20	2.41	2.57	2.70	2.81	2.91	2.99	3.06	3.13	3.20	3.25	3.31	3.35	3.40	3.44	3.48	3.52	
10	1.29	1.88	2.17	2.37	2.52	2.65	2.75	2.84	2.92	2.99	3.06	3.12	3.17	3.22	3.26	3.31	3.35	3.39	3.42	
11	1.28	1.86	2.14	2.33	2.48	2.60	2.70	2.79	2.87	2.94	3.00	3.05	3.10	3.15	3.19	3.24	3.27	3.31	3.34	
12	1.27	1.84	2.12	2.31	2.45	2.57	2.67	2.75	2.82	2.89	2.95	3.00	3.05	3.10	3.14	3.18	3.21	3.25	3.28	
13	1.26	1.83	2.10	2.28	2.42	2.54	2.63	2.72	2.79	2.85	2.91	2.96	3.01	3.05	3.09	3.13	3.16	3.20	3.23	
14	1.25	1.82	2.08	2.26	2.40	2.51	2.61	2.69	2.76	2.82	2.88	2.93	2.97	3.01	3.05	3.09	3.12	3.16	3.19	
15	1.24	1.81	2.07	2.25	2.38	2.49	2.59	2.66	2.73	2.79	2.85	2.90	2.94	2.98	3.02	3.06	3.09	3.12	3.15	
16	1.24	1.80	2.06	2.23	2.37	2.48	2.57	2.65	2.71	2.77	2.82	2.87	2.91	2.96	2.99	3.03	3.06	3.09	3.12	
17	1.23	1.79	2.05	2.22	2.35	2.46	2.55	2.63	2.69	2.75	2.80	2.85	2.89	2.93	2.97	3.00	3.03	3.07	3.09	
18	1.23	1.78	2.04	2.21	2.34	2.45	2.54	2.61	2.68	2.73	2.79	2.83	2.87	2.91	2.95	2.98	3.01	3.04	3.07	
19	1.23	1.78	2.03	2.20	2.33	2.44	2.52	2.60	2.66	2.72	2.77	2.81	2.85	2.89	2.93	2.96	2.99	3.02	3.05	
20	1.22	1.77	2.02	2.19	2.32	2.43	2.51	2.59	2.65	2.70	2.75	2.80	2.84	2.88	2.91	2.95	2.98	3.00	3.03	
24	1.21	1.76	2.00	2.17	2.29	2.39	2.48	2.55	2.61	2.66	2.71	2.75	2.79	2.83	2.86	2.89	2.92	2.95	2.97	
30	1.20	1.74	1.98	2.14	2.26	2.36	2.44	2.51	2.57	2.62	2.67	2.71	2.75	2.78	2.81	2.84	2.87	2.89	2.92	
40	1.19	1.72	1.96	2.12	2.24	2.33	2.41	2.47	2.53	2.58	2.62	2.66	2.70	2.73	2.76	2.79	2.82	2.84	2.87	
60	1.18	1.71	1.94	2.09	2.21	2.30	2.38	2.44	2.49	2.54	2.58	2.62	2.66	2.69	2.72	2.74	2.77	2.79	2.81	
120	1.17	1.69	1.92	2.07	2.18	2.27	2.34	2.40	2.46	2.50	2.54	2.58	2.61	2.64	2.67	2.70	2.72	2.74	2.77	
∞	1.16	1.68	1.90	2.05	2.15	2.24	2.31	2.37	2.42	2.47	2.50	2.54	2.57	2.60	2.63	2.65	2.68	2.70	2.72	
																			(continued)	

(continued)

Table A.1. (continued). HANOM Critical Values $\mathcal{H}(\alpha; k, \nu)$ Where $\alpha = .05$ and $\nu = n_0 - 1$

	Number of means being compared, k																			
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
ν																				
1	12.7	25.3	37.8	50.1	62.5	74.7	87.1	99.4	112.	125.	137.	150.	162.	175.	187.	200.	212.	225.	237.	
2	3.28	5.24	6.72	7.97	9.07	10.0	11.0	11.8	12.6	13.3	14.0	14.7	15.35	16.0	16.6	17.2	17.7	18.3	18.8	
3	2.29	3.43	4.17	4.75	5.25	5.67	6.05	6.40	6.71	7.00	7.28	7.53	7.76	7.99	8.20	8.39	8.58	8.77	8.94	
4	1.97	2.86	3.39	3.80	4.13	4.41	4.65	4.87	5.07	5.25	5.41	5.57	5.70	5.83	5.96	6.08	6.19	6.30	6.40	
5	1.81	2.60	3.04	3.37	3.62	3.84	4.03	4.20	4.35	4.49	4.61	4.72	4.82	4.91	5.01	5.09	5.18	5.25	5.33	
6	1.72	2.45	2.83	3.12	3.34	3.52	3.68	3.82	3.95	4.06	4.16	4.25	4.34	4.42	4.49	4.56	4.62	4.69	4.75	
7	1.66	2.35	2.71	2.96	3.16	3.33	3.46	3.59	3.70	3.80	3.89	3.97	4.04	4.11	4.17	4.23	4.29	4.34	4.39	
8	1.62	2.29	2.62	2.86	3.04	3.19	3.32	3.43	3.53	3.62	3.69	3.77	3.83	3.89	3.95	4.00	4.06	4.10	4.15	
9	1.59	2.24	2.55	2.78	2.95	3.09	3.21	3.31	3.40	3.48	3.56	3.62	3.69	3.74	3.79	3.84	3.89	3.93	3.97	
10	1.56	2.20	2.50	2.71	2.88	3.01	3.13	3.22	3.31	3.39	3.45	3.52	3.57	3.63	3.67	3.72	3.76	3.80	3.84	
11	1.55	2.17	2.46	2.67	2.83	2.96	3.06	3.15	3.24	3.31	3.37	3.43	3.49	3.54	3.58	3.62	3.66	3.70	3.74	
12	1.53	2.14	2.43	2.63	2.78	2.91	3.01	3.10	3.18	3.24	3.31	3.36	3.41	3.46	3.51	3.55	3.58	3.62	3.65	
13	1.52	2.12	2.40	2.60	2.75	2.87	2.97	3.06	3.13	3.19	3.25	3.31	3.36	3.40	3.44	3.48	3.52	3.56	3.59	
14	1.51	2.10	2.38	2.57	2.72	2.83	2.93	3.02	3.09	3.15	3.21	3.26	3.31	3.35	3.39	3.43	3.47	3.50	3.53	
15	1.50	2.09	2.36	2.55	2.69	2.81	2.90	2.98	3.05	3.12	3.17	3.22	3.27	3.31	3.35	3.39	3.42	3.45	3.48	
16	1.49	2.08	2.35	2.53	2.67	2.78	2.88	2.96	3.02	3.09	3.14	3.19	3.24	3.28	3.32	3.35	3.38	3.42	3.44	
17	1.48	2.07	2.33	2.52	2.65	2.76	2.85	2.93	3.00	3.06	3.11	3.16	3.21	3.25	3.28	3.32	3.35	3.38	3.41	
18	1.48	2.06	2.32	2.50	2.64	2.75	2.84	2.91	2.98	3.04	3.09	3.14	3.18	3.22	3.26	3.29	3.32	3.35	3.38	
19	1.47	2.05	2.31	2.49	2.62	2.73	2.82	2.89	2.96	3.02	3.07	3.12	3.16	3.20	3.23	3.27	3.30	3.33	3.36	
20	1.47	2.04	2.30	2.48	2.61	2.72	2.80	2.88	2.94	3.00	3.05	3.10	3.14	3.17	3.21	3.25	3.28	3.31	3.33	
24	1.45	2.02	2.27	2.44	2.57	2.67	2.76	2.83	2.89	2.94	2.99	3.04	3.08	3.11	3.15	3.18	3.21	3.23	3.26	
30	1.44	2.00	2.24	2.41	2.53	2.63	2.71	2.78	2.84	2.89	2.94	2.98	3.01	3.05	3.08	3.11	3.14	3.16	3.19	
40	1.43	1.98	2.22	2.38	2.50	2.59	2.67	2.73	2.79	2.84	2.88	2.92	2.96	2.99	3.02	3.05	3.07	3.10	3.12	
60	1.41	1.95	2.19	2.35	2.46	2.55	2.63	2.69	2.74	2.79	2.83	2.87	2.90	2.93	2.96	2.99	3.01	3.04	3.06	
120	1.40	1.93	2.16	2.31	2.43	2.51	2.58	2.64	2.70	2.74	2.78	2.82	2.85	2.88	2.90	2.93	2.95	2.97	3.00	
∞	1.39	1.91	2.14	2.29	2.39	2.48	2.54	2.60	2.65	2.70	2.73	2.77	2.80	2.83	2.85	2.88	2.90	2.92	2.94	

Table A.1. (continued). HANOM Critical Values $\mathcal{H}(\alpha; k, \nu)$ Where $\alpha = .01$ and $\nu = n_0 - 1$

	Number of means being compared, k																			
	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
ν																				
1	63.6	128.	190.	253.	318.	381.	446.	508.	570.	634.	698.	758.	824.	885.	950.	1015.	1077.	1145.	1210.	
2	7.24	11.7	15.1	18.0	20.5	22.7	24.9	26.8	28.5	30.2	31.9	33.3	34.8	36.3	37.7	38.9	40.2	41.4	42.6	
3	3.95	5.89	7.21	8.24	9.15	9.88	10.5	11.1	11.7	12.2	12.6	13.1	13.4	13.8	14.2	14.5	14.8	15.2	15.5	
4	3.07	4.37	5.20	5.83	6.34	6.77	7.15	7.49	7.78	8.05	8.30	8.53	8.75	8.95	9.13	9.31	9.47	9.64	9.80	
5	2.69	3.74	4.38	4.85	5.22	5.53	5.79	6.04	6.24	6.43	6.59	6.75	6.89	7.03	7.15	7.27	7.39	7.50	7.59	
6	2.48	3.41	3.93	4.32	4.63	4.87	5.08	5.26	5.42	5.58	5.71	5.83	5.94	6.04	6.13	6.23	6.31	6.39	6.46	
7	2.35	3.20	3.66	4.01	4.27	4.47	4.65	4.81	4.95	5.08	5.19	5.28	5.38	5.46	5.54	5.61	5.69	5.76	5.82	
8	2.27	3.06	3.48	3.79	4.02	4.21	4.36	4.50	4.63	4.73	4.83	4.92	4.99	5.07	5.14	5.20	5.26	5.32	5.38	
9	2.20	2.96	3.35	3.63	3.84	4.01	4.16	4.29	4.40	4.50	4.59	4.66	4.74	4.80	4.86	4.92	4.97	5.02	5.08	
10	2.15	2.88	3.25	3.52	3.72	3.87	4.01	4.13	4.23	4.32	4.40	4.47	4.53	4.59	4.65	4.70	4.75	4.79	4.85	
11	2.12	2.83	3.18	3.42	3.61	3.77	3.89	4.00	4.10	4.18	4.25	4.33	4.38	4.44	4.49	4.54	4.59	4.63	4.66	
12	2.08	2.78	3.12	3.36	3.53	3.67	3.80	3.90	3.99	4.07	4.14	4.20	4.26	4.31	4.36	4.41	4.45	4.49	4.53	
13	2.06	2.74	3.07	3.30	3.47	3.61	3.72	3.82	3.91	3.98	4.05	4.11	4.17	4.22	4.26	4.30	4.35	4.38	4.42	
14	2.04	2.71	3.03	3.25	3.42	3.55	3.66	3.75	3.83	3.91	3.97	4.03	4.08	4.13	4.17	4.21	4.25	4.29	4.33	
15	2.02	2.69	3.00	3.21	3.37	3.50	3.61	3.70	3.78	3.85	3.91	3.96	4.01	4.06	4.11	4.14	4.18	4.22	4.25	
16	2.01	2.67	2.97	3.17	3.34	3.46	3.56	3.65	3.72	3.79	3.85	3.91	3.95	4.00	4.04	4.08	4.11	4.15	4.18	
17	1.99	2.64	2.94	3.14	3.30	3.42	3.52	3.61	3.68	3.75	3.81	3.86	3.91	3.95	3.99	4.02	4.06	4.09	4.13	
18	1.98	2.62	2.93	3.12	3.27	3.39	3.49	3.58	3.65	3.71	3.77	3.82	3.87	3.91	3.94	3.98	4.01	4.05	4.08	
19	1.98	2.61	2.90	3.10	3.25	3.37	3.46	3.54	3.62	3.68	3.73	3.78	3.83	3.87	3.91	3.94	3.97	4.01	4.04	
20	1.96	2.59	2.89	3.08	3.23	3.34	3.44	3.52	3.59	3.65	3.71	3.75	3.80	3.83	3.87	3.91	3.94	3.97	4.00	
24	1.94	2.55	2.83	3.02	3.16	3.26	3.36	3.43	3.50	3.56	3.61	3.65	3.70	3.73	3.77	3.80	3.83	3.86	3.89	
30	1.91	2.51	2.79	2.96	3.09	3.20	3.28	3.35	3.41	3.47	3.52	3.56	3.60	3.63	3.67	3.70	3.73	3.75	3.77	
40	1.89	2.48	2.74	2.91	3.03	3.13	3.21	3.27	3.33	3.38	3.43	3.47	3.51	3.54	3.57	3.60	3.62	3.65	3.67	
60	1.86	2.44	2.69	2.85	2.97	3.07	3.14	3.20	3.26	3.31	3.35	3.38	3.42	3.45	3.47	3.50	3.53	3.55	3.57	
120	1.84	2.41	2.64	2.80	2.92	3.00	3.07	3.13	3.18	3.23	3.27	3.30	3.33	3.36	3.39	3.41	3.44	3.45	3.48	
∞	1.82	2.38	2.61	2.76	2.87	2.95	3.01	3.07	3.12	3.16	3.20	3.23	3.26	3.28	3.31	3.33	3.35	3.37	3.39	

Table A.2. Data^a for the Example

	A1			A2			A3		
B1	6.6	4.0	2.3	4.5	2.2	4.1	2.3	.9	7.5
	2.7	6.2		5.5	2.7		5.6	4.9	
	6.0	5.0		4.8	5.8		2.2	3.4	
B2	3.0	3.2	1.1	3.0	1.5	1.7	1.3	3.3	3.2
	2.1	4.1		2.5	2.6		.5	1.1	
	5.9	2.5		.4	3.5		1.7	.1	
B3	5.7	4.4	5.3	8.9	7.7	.5	2.6	9.9	^b
	3.2	3.2	8.9	7.0	7.3	4.2	11.5	10.5	
	5.3	9.7		8.0	2.2		3.4	6.7	
B4	7.0	8.9	9.0	12.0	9.7	10.9	8.3	8.0	9.4
	7.3	9.0		8.5	10.8		10.4	9.7	
	8.6	11.3		7.9	7.3		10.6	7.4	

^a First-stage data are in regular font [e.g., 6.6 in cell (A1, B1)]. Second-stage data are in bold [e.g., **2.3** in cell (A1, B1)].

^b 5.9, 3.3, 1.9, 4.3, 13.5, 12.8, .2, 2.8, 7.9, 1.3, 3.1, 3.2, 9.2.

APPENDIX B: LFC FOR HANOM

First note that the power function (7) depends on $\boldsymbol{\mu}$ only through $\boldsymbol{\mu}_i = (\mu_1 - \mu_*, \dots, \mu_i - \mu_*)$. In order to show that the LFC for HANOM is of the form (10), we will show that $1 - \text{Power}(\boldsymbol{\mu}_i)$ can be expressed as the integral of a log-concave density over a permutationally symmetric translation-invariant convex set. We will need the following two results.

Theorem 1 (Marshall and Olkin 1979, p. 73). If g is a continuous nonnegative function, then

$$f(\mathbf{x}) = \prod_{i=1}^n g(x_i)$$

is Schur concave if and only if g is log concave.

Theorem 2 (Giani and Finner 1991). Let f be a symmetric, permutation invariant, and log-concave probability density with respect to the Lebesgue measure on \mathbb{R}^k . If the region

$R \subset \mathbb{R}^k$ is convex, symmetric about the origin, translation invariant, and permutation invariant, then

$$\sup_{\theta \in M_\delta} \int_R f(\mathbf{x} + \theta) d\mathbf{x} = \int_R f(\mathbf{x} + \boldsymbol{\mu}_{\text{LFC}}) d\mathbf{x}$$

where $M_\delta = \{\boldsymbol{\mu} = \{\mu_1, \dots, \mu_l\} : \max_{i,j} |\mu_i - \mu_j| \geq \delta\}$ and $\boldsymbol{\mu}_{\text{LFC}} = (\delta/2, -\delta/2, 0, \dots, 0)$.

It follows from (7) that

$$1 - \text{Power}(\boldsymbol{\mu}_i) = \int_E f\left(\mathbf{x} + \frac{w}{\delta} \boldsymbol{\mu}_i\right) d\mathbf{x}$$

where

$$E = \{\mathbf{x} : |x_i - \bar{x}| \leq \mathcal{H} \text{ for all } i\}$$

and $f(\mathbf{x})$ is the joint density of l independent $t(n_0 - 1)$ random variables. Since the random variables are independent,

$$f(\mathbf{x}) = \prod_{i=1}^l f_{T_i}(x_i)$$

where $f_{T_i}(x_i)$ is a univariate t density, which is proportional to

$$g(x) = (1 + x^2)^{-\nu}.$$

Letting

$$\begin{aligned} \phi(x) &= \ln(g(x)) \\ &= -\nu \ln(1 + x^2) \end{aligned}$$

it follows that

$$\phi'(x) = -2\nu x / (1 + x^2)$$

and

$$\phi''(x) = -2\nu(1 + x^2 - x) / (1 + x^2)^2 \leq 0.$$

Thus, ϕ is concave, g (and, therefore, f_{T_i}) is log concave, and Theorem 1 implies that f is Schur concave. Since f is Schur concave, $\log(f)$ is Schur concave, and f is log concave. Finally, since E is a permutationally symmetric translation-invariant convex set, Theorem 2 implies that the LFC is given by configuration (10).

[Received August 2000. Revised October 2001.]

REFERENCES

- Bishop, T. A., and Dudewicz, E. J. (1978), "Exact Analysis of Variance With Unequal Variances: Test Procedures and Tables," *Technometrics*, 20, 419-430.
- (1981), "Heteroscedastic ANOVA," *Sankhyā, Series B*, 43, 40-57.
- Craig, C. C. (1947), "Control Charts Versus the Analysis of Variance in Process Control by Variables," *Industrial Quality Control*, 3(4), 14-16.
- Craig, R. J. (1984), "Normal Family Distribution Functions: FORTRAN and BASIC Programs," *Journal of Quality Technology*, 16, 232-236.
- Dudewicz, E. J., and Nelson, P. R. (2002), "Heteroscedastic Analysis of Means," *American Journal of Mathematical and Management Sciences*, to appear.
- Giani, G., and Finner, H. (1991), "Some General Results on Least Favorable Parameter Configurations With Special Reference to Equivalence Testing and the Range Statistic," *Journal of Statistical Planning and Inference*, 28, 33-47.

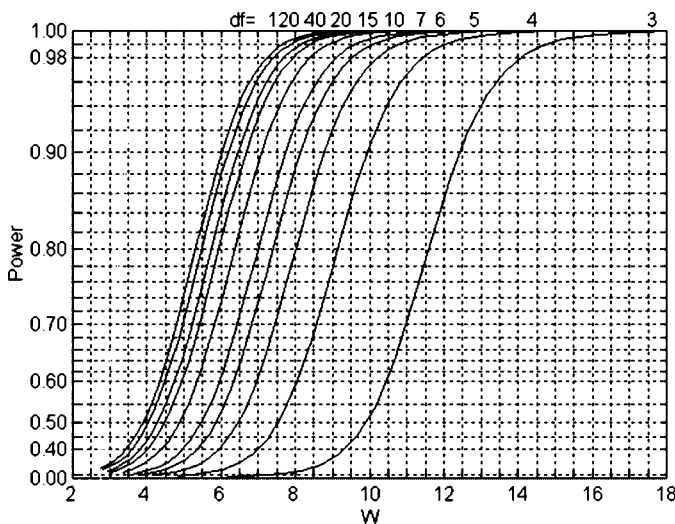


Figure A.1. Heteroscedastic ANOM Power Curves for $\alpha = .1$ and $l = 12$.

- Karian, Z. A., and Dudewicz, E. J. (1999), *Modern Statistical, Systems, and GPSS Simulation* (Second Edition), Boca Raton, London, New York, and Washington, DC: CRC Press LLC.
- Kinderman, A. J., and Monahan, J. F. (1980), "New Methods for Generating Student's t and Gamma Variables," *Computing*, 25, 369–377.
- Laplace, P. S. (1827), "Mémoire sur le Flux et Reflux Lunaire Atmospheric," in *Connaissance des Temps pour l'an 1830*, 3–18.
- Marshall, A. W., and Olkin, I. (1979), *Inequalities: Theory of Majorization and Its Applications*, New York: Academic Press.
- Nelson, P. R. (1983), "A Comparison of Sample Sizes for the Analysis of Means and the Analysis of Variance," *Journal of Quality Technology*, 15, 33–39.
- Ott, E. R. (1967), "Analysis of Means—A Graphical Procedure," *Industrial Quality Control*, 24, 101–109. Reprinted in *Journal of Quality Technology*, (1983), 15, 10–18.
- Pearson, E. S., and Hartley, H. O. (eds.) (1958), *Biometrika Tables for Statisticians, Vol. 1*. New York: Cambridge University Press.
- Scheffé, H. (1947), "The Relationship of Control Charts to the Analysis of Variance and Chi-Square Tests," *Journal of the American Statistical Association*, 42, 425–431.
- (1959), *The Analysis of Variance*. New York: Wiley.
- Stein, C. (1945), "A Two-Sample Test for a Linear Hypothesis Whose Power is Independent of the Variance," *Annals of Mathematical Statistics*, 16, 243–258.
- Vaughn, R. C. (1990), *Quality Assurance*, Ames, IA: Iowa State University Press.
- Wludyka, P. S., and Nelson, P. R. (1997), "An Analysis of Means Type Test for Variances From Normal Populations," *Technometrics*, 39, 274–285.