Center for Quality and Productivity Improvement
University of Wisconsin-Madison
610 Walnut Street
Madison, WI 53705

Report No. 20

A Critical Look at Accumulation Analysis and Related Methods

M. Hamada and C. F. J. Wu

November 1986

This research was supported by National Science Foundation Grant DMS-8420968. The simulation studies were performed on the VAX 11/780 computer of the Mathematics Research Center, University of Wisconsin--Madison. This research was aided by access to the research computer of the University of Wisconsin--Madison Statistics Department. Completed June 1986.

CENTER FOR QUALITY AND PRODUCTIVITY IMPROVEMENT
UNIVERSITY OF WISCONSIN-MADISON

Report No. 20

A CRITICAL LOOK AT ACCUMULATION ANALYSIS AND RELATED METHODS

M. Hamada and C. F. J. Wu

November 1986

PRACTICAL SIGNIFICANCE

Often in industry, quality characteristics are measured categorically rather than numerically, such as recording a response as "slight," "moderate," or "extreme." Accumulation analysis is a method proposed by Taguchi for analyzing such ordered categorical data from industrial experiments. It is used extensively in Japanese industry and is becoming popular in the United States. This paper exposes a fundamental problem with accumulation analysis in the multifactor setting, the usual industrial setting, since it is more efficient to simultaneously test many factors. A serious consequence of accumulation analysis is the spurious detection of factor effects, as demonstrated by simulation results for various situations. Furthermore, reanalysis of data from two real experiments reveals that spurious factor effects are being detected in practice. Therefore, accumulation analysis is not recommended.

Nair(1986) has proposed using the first two components of the accumulation analysis statistic, as well as simpler alternatives, to detect location and dispersion effects, respectively. This paper demonstrates that the dispersion tests are often not applicable in the industrial setting and that the location tests can be useful but are not without problems.

Key words: Accumulation analysis; ordered categorical data; multifactor experiments; location effects; dispersion effects

---

This research was supported by National Science Foundation Grant DMS-8420968. The simulation studies were performed on the VAX 11/780 computer at the Mathematics Research Center, University of Wisconsin--Madison. This research was aided by access to the research computer of the University of Wisconsin--Madison Statistics Department. Completed June 1986.

A CRITICAL LOOK AT ACCUMULATION ANALYSIS AND RELATED METHODS

M. Hamada and C. F. J. Wu

1. Introduction

Taguchi's accumulation analysis (AA) for ordered categorical data from industrial experiments has become increasingly popular (Taguchi, 1974). AA is appealing because it is easy to use; it is an ANOVA-like procedure. Unfortunately, there is a fundamental problem with AA in the multifactor setting. This problem with AA can be clearly seen in the context of fractional factorial designs, the ones most often used in industry. One consequence is that spurious effects can be detected. Furthermore, the test statistics for factor effects are dependent.

In section 2, we will describe AA in the single-factor setting. Noticing a problem with the denominator of the AA statistic, Box (1985) and Nair (1986b) proposed using the numerator sum of squares of the AA statistic. In the following, we will also consider this modified version of the AA statistic.

Box (1985) and Nair (1986b) decomposed the AA statistic into a weighted sum of score statistics. Using this decomposition, we will give a new interpretation of AA and its components in section 2. Since the scores are linear, quadratic, etc., the first two components have been interpreted as tests for location and dispersion effects. Nair (1986b) suggested using the first two components separately to increase the power for detecting location effects and for dispersion

---

The Center for Quality and Productivity Improvement cares about your reactions to our Reports. Please send comments (general or specific) to: Report Feedback, Center for Quality and Productivity Improvement, 610 Walnut Street, Madison, WI 53705. All replies will be forwarded to the authors.

effects, respectively. He also proposed simpler location and dispersion tests as alternatives; interestingly, the location test is the Kruskal-Wallis test. One consequence of our interpretation of AA is that the second component of the AA statistic and the dispersion test proposed by Nair can detect something other than a dispersion effect. Therefore qualification of their interpretation as dispersion tests will be made.

In section 3, using our interpretation of AA, we will expose a fundamental problem with AA in the multifactor setting. This problem is a result of neglecting the inherent nature of ordered categorical data which makes them harder to analyze than continuous data. The results from section 2 will be used to explain what AA is doing wrong. Three models will be examined where AA detects spurious effects. Nair (1986b) gave another reason for using the modified AA statistic: it removes one way in which the distribution of the AA statistic for a factor depends on the other factors. The distribution of the modified AA statistic will be shown to still depend on the other factors. Nair (1986a) showed that Pearson's χ^2 statistic is the unweighted sum of the components of the AA statistic. Consequently, Pearson's χ^2 test has the same problem as AA in the multifactor setting.

The results from section 3 provide a compelling reason to search for alternatives. Possibilities are the first component of the AA statistic and the Kruskal-Wallis test. However, these alternatives are not without their problems. In section 4, it will be shown that they can miss small real effects as well as detect spurious effects. In the multifactor setting, the second component of the AA statistic and the dispersion test proposed by Nair are rarely applicable as tests for dispersion. However, rather restrictive conditions will be given under which these tests can be interpreted as such.

In section 5, the consequences of using AA will be illustrated by reanalyzing data from two real experiments. Section 6 concludes with a discussion.

2. Accumulation Analysis in the Single Factor Setting

We will review AA in the single factor setting in section 2.1. In section 2.2, we offer a new interpretation of AA and its components. We will use this to show that the interpretation of the second component of the AA statistic and its simpler alternative proposed by Nair (1986b) as dispersion tests require qualification.

2.1. Review of Accumulation Analysis

Taguchi (1974) developed AA for ordered categorical data as an alternative to Pearson's χ^2 test. His argument for AA's superiority is that it accounts for the ordered nature of the categories whereas Pearson's χ^2 does not; examples were given to demonstrate this. The AA statistic in the single factor setting will be described next.

Consider a single factor A with I levels and n observations taken at each level. Let $\mathbf{y_i} = (y_{i1}, \ldots, y_{iK})^T$ denote the frequencies of K ordered categories observed at the ith level of the factor. The data can be viewed as an I×K contingency table. Let C_{ik} be the cumulative frequencies of the first k categories at the ith level and $C_{\bullet k}$ be the average of the C_{ik} across the I levels. That is, $C_{ik} = \sum_{j=1}^{k} y_{ij}$ and $C_{\bullet k} = \sum_{i=1}^{I} C_{ik} / I$.

AA builds K-1 I×2 tables where the kth table is formed by collapsing the first k columns to form one column and by collapsing the last K-k columns to form the other. The ith row of the kth table is $(C_{ik}, n-C_{ik})$. Standard ANOVA is performed on each table as if C_{ik} ones and $n-C_{ik}$ zeros were observed at the ith level of the factor. The sums of squares for the kth table are as follows:

$$SS_A(k) = (1/n)\sum_{i=1}^{I}(C_{ik}-C_{\bullet k})^2,$$

$$SS_{tot}(k) = (I/n)C_{\bullet k}(n-C_{\bullet k}), \text{ and}$$

$$SS_e(k) = SS_{tot}(k) - SS_A(k) = (1/n)\sum_{i=1}^{I}C_{ik}(n-C_{ik}).$$

Let d_k denote $C_{\bullet k}/n$, the cumulative marginal proportions. AA then weights the sums of squares for the kth table by $(d_k(1-d_k))^{-1}$ and adds the corresponding weighted sums of squares over the K-1 tables. This yields:

$$SS_A = n\sum_{k=1}^{K-1}\sum_{i=1}^{I}(C_{ik}-C_{\bullet k})^2 \,/\, (C_{\bullet k}(n-C_{\bullet k})), \tag{2.1}$$

$$SS_{tot} = nI(K-1), \text{ and}$$

$$SS_e = SS_{tot}-SS_A = n\sum_{k=1}^{K-1}\sum_{i=1}^{I}C_{ik}(n-C_{ik}) \,/\, (C_{\bullet k}(n-C_{\bullet k})).$$

To obtain mean squares, AA uses (I-1)(K-1) and $I(n-1)(K-1) = (nI-1)(K-1)-(I-1)(K-1)$ as degrees of freedom for SS_A and SS_e, respectively. Finally, the AA statistic is an F-like statistic given by $F_A = MS_A/MS_e$. Therefore, AA is an ANOVA-like procedure.

AA's simplicity and similarity to ANOVA is appealing. Unfortunately, it does not possess ANOVA's property of independent sums of squares. Notice that $SS_e = \text{constant} - SS_A$. Box (1985) and Nair (1986b) recognized this undesirable property where SS_e depends on the effect of factor A. Consequently, they proposed using only the numerator sum of squares SS_A from (2.1). In the following we will also consider the modified statistic and will refer to it as the AA statistic T. The AA statistic T can be easily remembered as $T = \sum_{k=1}^{K-1} P_k$, where P_k is Pearson's χ^2 statistic for the kth I×2 table (Takeuchi and Hirotsu, 1982).

The AA statistic T can be expressed as follows: $T = \sum_{i=1}^{I} \mathbf{x}_i^T \mathbf{x}_i$, where $x_{ik} = \sqrt{n}(C_{ik} - C_{\bullet k}) / \sqrt{C_{\bullet k}(n - C_{\bullet k})}$ for k=1, ..., K-1. From Nair (1986c), T can be decomposed into uncorrelated components. See also Box (1985) for a somewhat different but equivalent treatment. Let $\mathbf{\Sigma}$ be the conditional variance-covariance matrix of $\mathbf{x}_i$, where $\mathbf{x}_i$ is conditioned on the marginal frequencies of the ordered categories. For $l \le k$, $\Sigma_{lk} \propto \sqrt{C_{\bullet l}(n - C_{\bullet k})} / \sqrt{(n - C_{\bullet l})C_{\bullet k}}$. Let $\mathbf{\Lambda}$ be the (K−1)×(K−1) diagonal matrix of eigenvalues of $\mathbf{\Sigma}$ and $\mathbf{R}$ be the corresponding matrix of eigenvectors. Then T can be expressed as

$$T = \sum_{i=1}^{I} \mathbf{w}_i^T \Lambda \mathbf{w}_i = \sum_{j=1}^{K-1} \lambda_j Z^2_j$$

where $\mathbf{w}_i = \mathbf{\Lambda}^{-1/2}\mathbf{R}^T\mathbf{x}_i$, $Z^2_j = \sum_{i=1}^{I} w_{ij}^2$, and $\Lambda_{jj} = \lambda_j$.

Under the null hypothesis of no factor effect the Z^2_j are asymptotically IID χ^2_{I-1} (Box, 1985 and Nair, 1986a). That is, the AA statistic T is distributed as a linear combination of IID χ^2 random variables. By matching the first two moments, the distribution of T can be approximated by $d\chi^2_v$ where $d = \sum_{j=1}^{K-1} \lambda_j^2 / \sum_{j=1}^{K-1} \lambda_j$ and $v = (I-1)(K-1)/d$.

2.2. A New Interpretation of Accumulation Analysis

In terms of the frequency data, $\mathbf{w}_i = (1/\sqrt{n})\mathbf{Q}^T(\mathbf{y}_i - \mathbf{y}_\bullet)$ for some K×(K−1) matrix $\mathbf{Q}$ where $\mathbf{y}_\bullet = \sum_{i=1}^{I} \mathbf{y}_i / I$. Then,

$$Z^2_j = \sum_{i=1}^{I} (\mathbf{q}_j^T(\mathbf{y}_i - \mathbf{y}_\bullet))^2 / n, \tag{2.2}$$

where $\mathbf{q_j}$ is the jth column of $\mathbf{Q}$. Note that the $\mathbf{q_j}$ depend only on the marginal frequencies of the ordered categories and that $Z^2{}_j$ can be viewed as score statistics using the scores $\mathbf{q_j}$.

We may interpret $(\mathbf{y_i} - \mathbf{y_\bullet})$ as a comparison of the distribution at the ith level of the factor with a reference distribution, the mixture of all the distributions at the I levels. Therefore, using the scores $\mathbf{q_j}$, $Z^2{}_j$ can be viewed as a comparison of the distributions at the I levels of the factor with a reference distribution. When there is no factor effect, the I distributions will be identical as will be the reference distribution; $Z^2{}_j$ will tend to be small. However, when there is a factor effect, some of the I distributions will be different. Consequently, the I distributions will be different from the reference distribution and $Z^2{}_j$ will tend to be large. Therefore, $Z^2{}_j$ tests whether the factor has an effect with respect to the scores $\mathbf{q_j}$. AA can then be interpreted as a weighted combination of K-1 different tests to detect a factor effect.

The scores $\mathbf{q_1}$ and $\mathbf{q_2}$ are linear and quadratic. Furthermore, for $I = 2$ with equal marginal frequencies, $Z^2{}_1$ and $Z^2{}_2$ are equivalent to the Wilcoxon rank-sum and Mood statistics, respectively. For $I > 2$ with equal marginal frequencies, $Z^2{}_1$ and $Z^2{}_2$ correspond to the Kruskal-Wallis statistic and the generalization of the Mood statistic. Consequently, $Z^2{}_1$ and $Z^2{}_2$ have been interpreted as tests for location and dispersion effects. Because the weights λ_j decrease rapidly in j, AA has been interpreted as a test primarily for location. Nair (1986b) proposed using $Z^2{}_1$ and $Z^2{}_2$ separately to increase the power of detecting location and dispersion effects. A simple example will show that the interpretation of $Z^2{}_2$ as a dispersion test requires qualification.

Example 1: Suppose the ordered categorical random variable is generated by an underlying continuous random variable Y, where $Y = \mu_i + \varepsilon$. Let $\mu = (-1.5, -.5, .5, 1.5)$ and ε be uniform on $[-.5, +.5]$. Four ordered categories are defined by $(-\infty, -1, 0, 1, +\infty)$. That is, if $-\infty < Y < -1$, then the observation would be in the first category, if $-1 < Y < 0$, then the observation would be in the second category, etc. Ten observations are taken at each level. Note that the factor does not have a dispersion effect. Table 2.1 presents the data for every realization of ε, yielding the AA statistic T=120.

Table 2.1: Data for Example 1

Level	Category			
	1	2	3	4
-1.5	10	0	0	0
-.5	0	10	0	0
+.5	0	0	10	0
+1.5	0	0	0	10

By decomposing the AA statistic,

$$T = \sum_{j=1}^{K-1} \lambda_j Z^2_j = 2(40) + 2/3(40) + 1/3(40) = 120.$$

Using the $d\chi^2_v$ approximation, AA would detect an effect ($d\chi^2_v(.01) = 25.4$). If the components Z^2_1 and Z^2_2 were used separately, the factor would be identified as having both a location and dispersion effect ($\chi^2_3(.01) = 11.3$). Z^2_1 correctly identifies the location effect, but Z^2_2 falsely detects a dispersion effect.

The problem with Z^2_2 as a test for dispersion is that it fails to account for the different locations at the levels of the factor. Since the scores $\mathbf{q}_2$ are functions of only the marginal frequencies, they cannot simultaneously account for the location at each level. What Z^2_2 detects is a difference between the distributions at the I levels. Therefore, Z^2_2 *can only be interpreted as a test for dispersion when there is no location effect.*

Nair (1986b) proposed simpler location and dispersion statistics, SS(**l**) and SS(**d**). Using the scores **l** and **d**,

$$SS(\mathbf{l}) = \sum_{i=1}^{I} (\mathbf{l}^T(\mathbf{y}_i - \mathbf{y}_\bullet))^2 /n \text{ and}$$

$$SS(\mathbf{d}) = \sum_{i=1}^{I} (\mathbf{d}^T(\mathbf{y}_i - \mathbf{y}_\bullet))^2 /n .$$

SS(**l**) is the Wilcoxon rank-sum test statistic when I=2 and the Kruskal-Wallis test statistic when I > 2. In the following we will refer to SS(**l**) as the Kruskal-Wallis (KW) test statistic. Asymptotically, KW and SS(**d**) are IID χ^2_{I-1} under the null hypothesis. Since these are also score statistics based on the marginal frequencies, SS(**d**) has the same problem as $Z^2{}_2$. For the example above, the dispersion scores are $(1, -1, -1, 1)$, yielding SS(**d**)=40; $Z^2{}_2$ and SS(**d**) are same when the marginal frequencies are equal. Again, a spurious dispersion effect would be detected. Therefore, SS(**d**) *can only be interpreted as a dispersion test when there is no location effect.*

Consider Example 1 where ε is now $N(0,\sigma^2)$. A study based on 10,000 simulations was performed using a $\chi^2{}_3$ approximation for $Z^2{}_2$ and SS(**d**). Table 2.2 summarizes the size of the .05 and .10 tests for a dispersion effect using $Z^2{}_2$ and SS(**d**). The same consequences can be seen when $Z^2{}_2$ or SS(**d**) are used as dispersion tests in the presence of a location effect. What is really being detected is the difference between the distributions at the I levels of the factor. As σ^2 increases, the difference between the mixtures of distributions becomes smaller and is detected less often given the same number of observations taken per run. Consequently, the size of these tests decreases as σ^2 increases.

Table 2.2: Size of .05 and .10 Tests
for Dispersion Effect Using $Z^2{}_2$ and SS(**d**)

σ^2	Nominal Level	Statistic	
		$Z^2{}_2$	SS(**d**)
1/4	.05	.98	.98
	.10	.99	.99
1/2	.05	.76	.78
	.10	.86	.87
1	.05	.33	.34
	.10	.46	.48

Nair (1986b) suggested using $Z^2{}_1$ and $Z^2{}_2$ separately to increase the power of these tests to detect location and dispersion effects. The Kruskal-Wallis statistic and SS(**d**) could also be used as simpler alternatives. The above qualification of the interpretation of $Z^2{}_2$ and SS(**d**) as dispersion tests suggests the following strategy:

First, test for a location effect using $Z^2{}_1$ or KW.

If there is no location effect, then use $Z^2{}_2$ or SS(**d**) to test for a dispersion effect.

3. Accumulation Analysis in the Multifactor Setting

In the multifactor setting, Nair (1986b) noted another reason for using the modified AA statistic. Since

$$SS_e = \text{constant} - SS_A - SS_B$$

in a two-factor main effects setting, the distribution of the original AA statistic for factor A depends on factor B. Therefore, using the modified AA statistic T from (2.1) eliminates one way in which the distribution of the AA statistic for a factor depends on the other factors.

There is a fundamental problem with AA in the multifactor setting which even this modification cannot remedy. It can be clearly seen in the context of fractional factorial designs. One consequence is that the distribution of the AA statistic for a factor still depends on the other factors. The focus of this section will be on main effects in the context of fractional factorial designs. AA in the full factorial setting will be discussed briefly.

Consider a multifactor experiment using a fractional factorial design with r runs and n observations taken at each run. Denote the frequency of the jth category for the ith run by n_{ij}. For the extension of AA to the multifactor setting, K-1 ANOVAs are performed where the data for the kth ANOVA consists of $\sum_{i=1}^{r}\sum_{j=1}^{k} n_{ij}$ ones and $rn - \sum_{i=1}^{r}\sum_{j=1}^{k} n_{ij}$ zeros. The data can also be viewed as a multiway contingency table. It can be shown that the AA statistic for a factor main effect, say factor A with I levels, is equivalently obtained by collapsing the design onto the factor, which produces an I×K contingency table, and by calculating the AA statistic for the single factor as was described in section 2.1.

Using our interpretation of AA, the fundamental problem with AA in the multifactor setting is that it is no longer comparing I distributions but I mixtures of distributions. This suggests the possibility of the I mixtures of distributions being different even when there is no factor A effect. Note that this scenario depends on the other factors. A simple example demonstrating this problem with AA in the multifactor setting will be examined next.

Example 2: For an experiment with a 2^{3-1} fractional factorial design (I=123), suppose that the ordered categorical random variable is generated by an underlying continuous random variable

Y, where $Y = A + B + C + \varepsilon$. Let A=0, B=±1, C=±.5, and ε be uniform on $[-.5, +.5]$. Four ordered categories are defined by $(-\infty, -1, 0, 1, +\infty)$ and ten observations are taken at each design setting. Note that factor A has neither a location nor a dispersion effect. Table 3.1 presents the data for the collapsed design onto factor A for every realization of ε, yielding the AA statistic $T_A = 26.7$. That is, the AA statistic has a degenerate distribution. AA would spuriously detect an effect for factor A ($d\chi^2_v(.01) = 13.8$).

Table 3.1: Data for Example 2 for Factor A

Level	Category			
	1	2	3	4
+	10	0	0	10
-	0	10	10	0

The decomposition of the AA statistic aids in explaining what AA is doing wrong in the multifactor setting. For this example, the second component of the AA statistic can be interpreted as a dispersion test since the locations of the underlying mixtures of distributions are the same. Although AA heavily weights the first component, the second component can also be important, which is true for this example. By decomposing the AA statistic T_A,

$$T_A = \sum_{j=1}^{K-1} \lambda_j Z^2_j = 2(0) + 2/3(40) + 1/3(0) = 26.7\,.$$

Here AA is detecting the difference in dispersion between the I mixtures of distributions at the levels of factor A. It is not detecting a factor A effect. Notice that there is a difference between the I mixtures of distributions because there are both factor B and C effects.

Example 2 illustrates the fundamental problem with AA in a main effects location model. Further, consider a main effects dispersion model and a main effects location-dispersion model. For the latter model, some factors can have both a location effect and a dispersion effect, whereas other factors may have only a location effect or only a dispersion effect. Similarly, AA can falsely detect an effect for factor A when there is no factor A effect because the I mixtures of distributions can be different. A simulation study using these three models for the same 2^{3-1} fractional factorial design will be described next.

For the location model, $Y = A + B + C + \varepsilon$. Let A=0, B=$\pm$1, C=$\pm$.5, and ε be $N(0, \sigma^2)$. Four ordered categories are defined by $(-\infty, -1, 0, 1, +\infty)$. For the dispersion model, $Y = \exp(A + B + C)\varepsilon$, where A=0, B=$\pm$.75, C=$\pm$.5, ε is $N(0, \sigma^2)$, and $(-\infty, -1.25, 0, 1.25, +\infty)$ define the categories. For the location-dispersion model, $Y = \sigma\varepsilon$ where ε is exponential with mean exp(A+B+C). Let A=0, B=$\pm$.7, C=$\pm$.4, and $(0, 1, 2, 3, +\infty)$ define the categories. Ten observations are taken at each design setting. Table 3.2 summarizes the size of the .05 and .10 tests for factor A based on 10,000 simulations using a $d\chi^2_v$ approximation for AA. Note that factor A has neither a location nor a dispersion effect in all three models.

Table 3.2: Size of .05 and .10 Tests for
Factor A Effect Using AA

σ^2	Nominal Level	Model		
		Location	Dispersion	Location-Dispersion
1/4	.05	.85	.13	.12
	.10	.94	.33	.26
1/2	.05	.39	.12	.13
	.10	.60	.26	.25
1	.05	.08	.06	.09
	.10	.19	.13	.19

In these three scenarios, AA is detecting the difference between the two mixtures of distributions. For the location model and dispersion model, AA is detecting the difference between the dispersion of the two mixtures of distributions. For the location-dispersion model, AA is detecting the difference between both the location and dispersion of the two mixtures.

The substantial difference in the dispersion of the mixtures of distributions explains the dramatic results for the location model. For the dispersion model, AA appears to be performing quite well for large σ^2. However, there is still a (smaller) difference in dispersion between the mixtures of distributions. AA is not detecting this difference because the ten observations per run and five ordered categories do not provide enough information to do so.

In the three scenarios above, the mixtures of distributions at the levels of factor A are different because there are both factor B and C effects. Therefore, the distribution of the AA statistic for a factor depends on the other factors. This dependence on the other factors can also be seen by studying the asymptotic distribution of the AA statistic.

We next derive the asymptotic distribution of the AA statistic and its components for the 2^{3-1} fractional factorial design. The four design settings for (A, B, C) are given in Table 3.3.

Table 3.3: 2^{3-1} Fractional Factorial Design

Run	Factor		
	A	B	C
1	+	+	+
2	+	-	-
3	-	+	-
4	-	-	+

Let $\mathbf{n}_i = (n_{i1}, \ldots, n_{iK})^T$, where $\sum_{j=1}^{K} n_{ij} = n$. Denote the category probabilities at the ith run by $\mathbf{p}_i = (p_{i1}, \ldots, p_{iK})^T$, where $\mathbf{p}_\bullet$ is the average of the $\mathbf{p}_i$. For $I = 2$, the AA statistic T can be expressed as

$$T = \mathbf{Z}^T \Lambda \mathbf{Z}\,, \text{ where } \mathbf{Z} = \mathbf{Q}^T (1/\sqrt{n})(\mathbf{y}_1 - \mathbf{y}_\bullet)\,.$$

$(\mathbf{y}_1 - \mathbf{y}_\bullet) = \sum_{i=1}^{4} \pm \mathbf{n}_i / 2$ and the sign of $\mathbf{n}_i$ for each factor is given in Table 3.3. Since $\mathbf{n}_i$ is distributed as multinomial(n, $\mathbf{p}_i$),

$$(1/\sqrt{n})\mathbf{n}_i \xrightarrow{L} N(\mathbf{p}_i, \Sigma_i)\,,$$

where $\Sigma_i = \mathrm{diag}(\mathbf{p}_i) - \mathbf{p}_i \mathbf{p}_i^T$. Furthermore, $\Lambda \xrightarrow{p} \overline{\Lambda}$ and $\mathbf{Q} \xrightarrow{p} \overline{\mathbf{Q}}$ where $\overline{\mathbf{Q}}$ satisfies $\overline{\mathbf{Q}}^T \Sigma_{\mathbf{p}_\bullet} \overline{\mathbf{Q}} = I$ and $\Sigma_{\mathbf{p}_\bullet} = \mathrm{diag}(\mathbf{p}_\bullet) - \mathbf{p}_\bullet \mathbf{p}_\bullet^T$. Note that $\overline{\Lambda}$ and $\overline{\mathbf{Q}}$ are functions of $\mathbf{p}_\bullet$. Denote the average of the Σ_i

by $\Sigma_\bullet$. Then,

$$\mathbf{Z} \xrightarrow{L} N(\mu, \Sigma)\,, \tag{3.1}$$

where $\mu = \overline{\mathbf{Q}}^T\eta$, $\eta = \sum_{i=1}^{4} \pm\mathbf{p_i}/2$, and $\Sigma = \overline{\mathbf{Q}}^T\Sigma_\bullet\overline{\mathbf{Q}}$. Therefore,

$$Z^2_j \xrightarrow{L} \Sigma_{jj}\chi^2(\delta^2_j) \text{ where } \delta^2_j = \mu^2_j/\Sigma_{jj}\,. \tag{3.2}$$

Under the null hypothesis that *none of the factors has an effect,* the $\mathbf{p_i}$ are equal so that $\overline{\mathbf{Q}}^T\Sigma_\bullet\overline{\mathbf{Q}} = \overline{\mathbf{Q}}^T\Sigma_{\mathbf{p}_\bullet}\overline{\mathbf{Q}} = I$ and $\delta^2_j = 0$. Therefore, Z^2_j are IID χ^2_1 and the AA statistic T is distributed as a linear combination of independent χ^2_1 random variables. Note that the critical values are based on this null hypothesis. Since the proper null hypothesis should be that the mixtures of distributions at the two levels are the same, the test could be either liberal or conservative. Preliminary work suggests that the tests using the AA statistic and its components are conservative.

Under the alternative hypothesis, (3.1) states that the Z^2_j for all the factors are dependent. From (3.2), Z^2_j is distributed as a constant multiple of a noncentral χ^2_1 random variable. Therefore, the AA statistic T is distributed as a linear combination of dependent noncentral χ^2_1 random variables.

Since $\eta = \sum_{i=1}^{4} \pm\mathbf{p_i}/2 = (1/n)E(\mathbf{y_1} - \mathbf{y_\bullet})$, $\pm\eta$ is the asymptotic quantity which compares the mixtures of distributions at each level of the factor with a reference distribution, the marginal distribution of the ordered categories. From (3.2), for a given j, η is the only quantity required to compare the asymptotic performance of the Z^2_j for all the factors since $\overline{\mathbf{q_j}}$ and Σ_{jj} remain the same for each factor.

The asymptotic distribution of the AA statistic T depends on all factors. This is so because it depends on $\overline{\Lambda}$, $\overline{\mathbf{Q}}$, $\Sigma_{\bullet}$ and η which are functions of $\mathbf{p}_l$ and $\mathbf{p}_l$ depends on all the factors.

Furthermore, the AA statistics for the factors are dependent. This is even true asymptotically. For any two factors, say A and B, the $(\mathbf{y}_{A,1} - \mathbf{y}_{\bullet})$ and $(\mathbf{y}_{B,1} - \mathbf{y}_{\bullet})$ are asymptotically dependent. Therefore, $Z^2{}_{A,j}$ and $Z^2{}_{B,j}$ are dependent, and consequently T_A and T_B are dependent. Remember that T is the sum of squares (2.1) so that in contrast with ANOVA, the sums of squares for all the factors in AA are dependent.

Nair (1986a) showed that Pearson's χ^2 test statistic is the sum of the components of the AA statistic. Therefore, if Pearson's χ^2 test is applied in the multifactor setting by collapsing the design onto the factor, it will have the same problem as AA. This may not seem so surprising because it is well known that there are problems with collapsing multiway contingency tables in the context of testing for independence. See Agresti (1984). The size of the .05 and .10 tests for factor A effect using Pearson's χ^2 test from the same simulation study described above for the three models appears in Table 3.4. Since Pearson's χ^2 test is more powerful than AA for detecting dispersion effects, the size of the test for the location and dispersion models is larger than that of AA.

Table 3.4: Size of .05 and .10 Tests
for Factor A Effect Using Pearson's χ^2 Test

σ^2	Nominal Level	Model		
		Location	Dispersion	Location-Dispersion
1/4	.05	.98	.19	.03
	.10	.99	.37	.10
1/2	.05	.78	.17	.07
	.10	.87	.31	.17
1	.05	.34	.07	.11
	.10	.48	.15	.22

The three examples above suggest what could happen in a full factorial setting. Since factor A has no effect, these examples can be viewed as a 2^2 factorial experiment in factors B and C. The test for factor A would correspond to a test for BC interaction. The above examples show that *AA can falsely detect an interaction effect.*

From the description of AA in section 2.1, AA is an ANOVA-like procedure. However, AA has none of its properties. That is, the distribution of the AA statistic for a factor depends on the other factors and the sums of squares are dependent. This is dramatically seen in the location model where ANOVA is applicable if continuous data were observed. It is interesting that ANOVA also compares I mixtures of distributions. That is, it too collapses the design onto the factor. However, ANOVA uses the additivity of the model and the orthogonality of the design to cancel out the effects of the other factors. Thus, the inherent nature of ordered categorical data which makes them harder to analyze is revealed.

The problem with AA is a result of neglecting this inherent nature of ordered categorical data. The examples above demonstrate how spurious effects can be detected. This result also suggests that the power of the AA test is artificially high. In the examples above, if factor A had a very small effect, most of the power of the test for a factor A effect would have come from the difference in the mixtures of distributions because of the large factor B and C effects. Therefore, AA can detect spurious effects as well as detect real effects for the wrong reason.

4. Methods Related to Accumulation Analysis

The results in section 3 provide a compelling reason to search for alternatives to AA in the multifactor setting. Nair (1986b) suggested using the first two components of the AA statistic separately to increase the power of detecting location and dispersion effects. Also, Nair proposed using the Kruskal-Wallis (KW) statistic and SS(**d**) as simpler alternatives to $Z^2{}_1$ and $Z^2{}_2$. Note that the Kruskal-Wallis test has been used to analyze ordered categorical data in the single factor setting for a long time. It appears that this test has not been studied in the multifactor setting when it is applied to data from the collapsed design.

Remember that these score tests are comparing I mixtures of distributions. Therefore, there are only certain situations where these are applicable as tests for location or dispersion effects. The examples in section 3 showed that AA can detect spurious effects when there is a difference in dispersion between the mixtures of distributions because of the large $Z^2{}_2$ component. Note that in all the examples factor A had no dispersion effect. In the single factor setting, $Z^2{}_2$ and SS(**d**) are applicable as dispersion tests when there is no location effect. In the multifactor setting, $Z^2{}_2$ and SS(**d**) are rarely applicable as dispersion tests. The following must hold in order to interpret them as such: (1) the factor has no location effect, (2) at most one of the other factors has a location effect, and (3) at most one of the other factors has a dispersion effect. Using Z^2_2 or SS(**d**) when these conditions are violated can result in the detection of spurious dispersion effects. Table 4.1 presents the size of the .05 and .10 tests for a factor A dispersion effect using Z^2_2 and SS(**d**) from the simulation study in section 3 under the location and dispersion models. Although factor A has neither a location nor a dispersion effect, the size of these tests is substantially larger than the nominal level.

Table 4.1: Size of .05 and .10 Tests for Dispersion Effect for Factor A Using Z^2_2 and SS(**d**) Under the Location and Dispersion Models

σ^2	Nominal Level	Model			
		Location		Dispersion	
		Z^2_2	SS(**d**)	Z^2_2	SS(**d**)
1/4	.05	1.00	1.00	.45	.49
	.10	1.00	1.00	.64	.66
1/2	.05	.92	.92	.34	.34
	.10	.96	.96	.49	.50
1	.05	.55	.56	.12	.12
	.10	.67	.68	.22	.22

For the dispersion model, Z^2_1 and KW can detect spurious location effects when $\mathbf{p}_\bullet$, the marginal expected proportions, are heavily skewed. For the same 2^{3-1} fractional factorial design, let $Y = \exp(A + B + C)\varepsilon$, where ε has a standard logistic distribution. The ordered categories are defined by $(-.08, .28, .85, 2.09)$. For A=0, and two (B, C) combinations, $(\pm.5, \pm.25)$ and $(\pm1, \pm.5)$, the corresponding $\mathbf{p}_\bullet$ are $(.48, .10, .13, .16, .13)$ and $(.47, .15, .13, .10, .15)$, respectively. Note that factor B has no location effect. Table 4.2 presents the size of .05 and .10 tests for a factor B location effect using Z^2_1 and KW based on 10,000 simulations and a χ^2_1

approximation for Z^2_1 and KW. These results can be explained by the heavily skewed $\mathbf{p}_{\bullet}$. Although the underlying mixtures of distributions have the same location, the transformation into ordered categories can make them appear to have different locations. It is this difference that Z^2_1 and KW are detecting. Because $\mathbf{p}_{\bullet}$ is more skewed for (B, C) = (±1, ±.5), the transformation has a greater effect which can be seen in Table 4.2. Observe that KW is less sensitive to this transformation effect. This reveals another aspect of the nature of ordered categorical data. Therefore, Z^2_1 and KW have the potential of spuriously detecting location effects when the marginal expected proportions are heavily skewed under the dispersion model.

Table 4.2: Size of .05 and .10 Tests for Factor B Location Effect
Using Z^2_1 and KW Under the Dispersion Model

(B, C)	Nominal Level	Statistic	
		Z^2_1	KW
(±1, ±.5)	.05	.35	.16
	.10	.50	.25
(±.5, ±.25)	.05	.14	.10
	.10	.23	.17

The performance of Z^2_1 and KW under the location model will be the focus of the remainder of this section. As was seen above, the transformation into ordered categories can make mixtures of distributions with the same location appear to have different locations. Similarly, when there are a small number of categories, mixtures of distributions with different locations can be transformed into ones which appear to have similar locations. For example, if there were only two categories in Example 2 of section 3, the factor C effect would not be detected.

Table 4.3 presents the data for factor C from Example 2 for the original four categories and for the two new categories formed from them.

Table 4.3: Data for Factor C from Example 2
for the Original and New Categories

Level	Category					
	Original				New	
	1	2	3	4	1&2	3&4
+	0	10	0	10	10	10
-	10	0	10	0	10	10

This suggests that $Z^2{}_1$ and KW have the potential for missing real effects and for detecting spurious effects. From section 3, $Z^2{}_1$ for all the factors are dependent. Similarly, KW for all the factors are dependent. The effect that the transformation into ordered categories and the dependence between the statistics of all the factors has on the performance of $Z^2{}_1$ and KW will be discussed next.

The dependence of the statistics for the factors result in small location effects being missed in the presence of factors with larger effects. This can be seen asymptotically by examining the noncentrality parameters of their asymptotic distributions. For I=2, the noncentrality parameter depends on $\mathbf{s}^T\boldsymbol{\eta}$, where $\mathbf{s}$ represents the scores $\mathbf{q}_1$ or $\mathbf{l}$ and $\boldsymbol{\eta} = (1/n)E(\mathbf{y}_1 - \mathbf{y}_\bullet)$. Remember that $\pm\boldsymbol{\eta}$ compares how different the mixtures of distributions at the levels of the factor are with a reference distribution. Note that $(1/n)E(\mathbf{y}_\bullet)$ is the same regardless of which factor is being considered, so that for a factor with a large effect, $(1/n)E(\mathbf{y}_1)$ will be more different from $(1/n)E(\mathbf{y}_\bullet)$ than it would be for a factor with a small effect. Therefore, the noncentrality parameter for a factor with a large effect will be larger than that for a factor with a small effect. In fact, if one factor has a substantially larger effect than another, the mixtures of distributions of the factor with the smaller effect will be very similar. That is, $\eta \approx \mathbf{0}$. Therefore, factors with small effects in the presence of factors with large effects can be missed. Table 4.4 presents the simulation results for the .05 and .10 tests using $Z^2{}_1$ and KW under the location model from section 3, where A=0, B=±1, and C=±.5 and ε is $N(0,\sigma^2)$. These results are the size of the tests for factor A effect and the power of the tests for factor B and C effects. The large factor B effect explains why the power of the tests for factor C and the size for the tests for factor A are smaller than the power of the tests for factor B, especially when the underlying error variance is large. Observe that if factor A had a small effect, it would be missed most of the time. The similar performance of $Z^2{}_1$ and KW seems to be true for symmetric boundaries with a symmetric error distribution. It is not true for asymmetric boundaries as can be seen in Tables 4.5 and 4.6.

Table 4.4: Size of Tests for Factor A and Power of Tests
for Factors B and C Using $Z^2{}_1$ and KW Under the Location Model
Category Boundaries = $(-\infty, -1, 0, 1, +\infty)$

σ^2	Level	Factor				
		A		B	C	
		$Z^2{}_1$	KW	$Z^2{}_1$ & KW	$Z^2{}_1$	KW
1/4	.05	.00	.00	1.00	.83	.84
	.10	.00	.00	1.00	.95	.96
1/2	.05	.00	.00	1.00	.63	.63
	.10	.00	.00	1.00	.81	.82
1	.05	.00	.00	1.00	.46	.46
	.10	.02	.01	1.00	.64	.64

Tables 4.5 and 4.6 demonstrate how the transformation into ordered categories affects the performance of $Z^2{}_1$ and KW. The transformation has changed the locations of the mixtures of

distributions of factor A. It is this difference in locations that $Z^2{}_1$ detects so that a factor A effect would often be spuriously detected. This also suggests that the power of the tests for factor B and C effects is artificially high. However, KW performs differently. The size of the factor A test is lower than its nominal value. Furthermore, the test for factor C has substantially smaller power than that for factor B especially when the error variance is small. What appears to be happening is that the transformation accentuates the factor B effect which decreases the power of the test for factor C. Despite these problems with $Z^2{}_1$ and KW, this does suggest that they can be useful in situations where they perform similarly. This was seen in Table 4.4 where the transformation appears to have little effect on the performance of $Z^2{}_1$ and KW. In practice, this would mean $Z^2{}_1$ and KW can be useful when they have nearly the same value.

Table 4.5: Size of Tests for Factor A and Power of Tests
for Factors B and C Using $Z^2{}_1$ and KW Under the Location Model
Category Boundaries = $(-\infty, 0, 1, 2, +\infty)$

σ^2	Level	Factor				
		A		B	C	
		$Z^2{}_1$	KW	$Z^2{}_1$ & KW	$Z^2{}_1$	KW
1/4	.05	.24	.01	1.00	.65	.31
	.10	.54	.06	1.00	.85	.58
1/2	.05	.15	.01	1.00	.58	.42
	.10	.34	.05	1.00	.77	.64
1	.05	.08	.01	1.00	.45	.41
	.10	.18	.05	1.00	.62	.58

When the ordered categories are coarse, less information is available to detect a difference between the locations of the mixtures of distributions. In situations where the transformation effect is substantial, this means that spurious effects will be declared less often when $Z^2{}_1$ is used, whereas real effects will be missed more often when KW is used. Table 4.6 demonstrates this when three categories are used instead of four and reveals that Z^2_1 is more sensitive to the loss of information. The amount of information provided by the number of ordered categories used has implications during the design phase of the experiment: refine the categories as much as possible to make the most information available to detect real effects.

Table 4.6: Size of Tests for Factor A and Power of Tests
for Factors B and C Using Z^2_1 and KW Under the Location Model
Category Boundaries = $(-\infty, 0, 1, +\infty)$

σ^2	Level	Factor				
		A		B	C	
		Z^2_1	KW	Z^2_1 & KW	Z^2_1	KW
1/4	.05	.12	.01	1.00	.52	.30
	.10	.31	.06	1.00	.75	.55
1/2	.05	.06	.01	1.00	.46	.39
	.10	.15	.04	1.00	.66	.60
1	.05	.03	.01	1.00	.40	.38
	.10	.08	.04	1.00	.56	.55

Finally, we compare Z^2_1 and KW with ANOVA. ANOVA uses the additivity of the model and the orthogonality of the design to cancel out the effects of the other factors. Even with orthogonality, the distributions of Z^2_1 and KW for a factor depend on all the other factors. Furthermore, an unbalanced design has a more serious effect on Z^2_1 and KW than ANOVA. Without a balanced design, the realizations of the mixtures of distributions appear to have different locations even when there is no effect. Suppose in Example 2 of section 3 that the sample sizes corresponding to $A + B + C = (-1.5, -.5, .5, 1.5)$ were $\mathbf{n} = (5, 15, 5, 15)$. Table 4.7 presents the data for factor A for every realization of ε. In contrast with Table 3.1, the realizations of the mixtures of distributions of factor A appear to have different locations. It is this difference that Z^2_1 and KW would detect, resulting in spurious effects declared. Similarly, real effects could be missed because the realizations of the mixtures of distributions could appear to have the same location. Therefore, when Z^2_1 and KW are used, an equal number of observations should be taken at each design setting.

Table 4.7: Data for Example 2 for Factor A With Unequal Sample Sizes

Level	Category			
	1	2	3	4
+	5	0	0	15
-	0	15	5	0

5. Examples from Real Experiments

In section 3, it was shown that AA can detect spurious effects. This problem with AA will be demonstrated by reanalyzing data from two real experiments.

5.1. An Arc Welding Experiment

We reanalyze data from an arc welding experiment performed by the National Railway Corporation of Japan. The data and original analysis appear in Taguchi and Wu (1980). One facet of the experiment was to find the important factors which affect the workability of an arc welded section between two steel plates. We interpret this to mean the degree of difficulty in welding the two steel plates together. Workability was classified into three categories: easy, normal, and difficult. The experimenters were initially interested in nine factors and four two-factor interactions. Some of the factors were welding method, angle of welding device, current, and type of welding rod. An experiment using a 2^{9-5} fractional factorial design was performed with one observation per run. Table 5.1 presents the results of AA, its first two components, $Z^2{}_1$ and $Z^2{}_2$, and KW. Using AA, the original analysis concluded that the main effects for factors D, F, and G were significant. However, using $Z^2{}_1$ or KW, only the main effects for factors D and F are important. Therefore, AA spuriously detected a factor G main effect. Because of the small $Z^2{}_1$ component and the large $Z^2{}_2$ component of the AA statistic for factor G, AA detected the difference in dispersion between the mixtures of distributions at the two levels of factor G, not a factor G main effect. Because of the confounding patterns of the design, the main effect for factor G is confounded with the DF interaction effect. This example demonstrates the results of section 3 which showed that AA can detect spurious interactions when the corresponding main effects are significant.

Table 5.1: Results of AA, $Z^2{}_1$, $Z^2{}_2$, and KW for the Welding Experiment

Factor	AA	$Z^2{}_1$	$Z^2{}_2$	KW
A	1.74	.10	2.23	.18
B	.41	.28	.16	.13
C	1.74	1.26	.18	1.32
D	5.03	3.70	.41	3.52
E	.41	.16	.28	.13
F	9.03	7.01	.10	7.09
G	5.03	.23	6.55	.09
H	.41	.16	.28	.13
I	1.74	1.26	.18	1.32
AG	.41	.16	.28	.13
AH	1.74	.10	2.23	.18
AC	.41	.16	.28	.13
GH	1.74	.10	2.23	.18

5.2. A Contact Stain Experiment

A company produces a rubber product which must meet a contact stain specification; it should not stain or mar the painted panel to which it is attached. The data and original analysis are found in Lear and Stanton (1985). A 2^{4-1} fractional factorial experiment (1=234) was performed to determine the important factors affecting the contact stain characteristic of the product. The four factors were chemical compounds used in making the rubber product. For each run, one product was attached to a painted metal panel and subjected to high temperature for three days. The panel was then inspected and its contact stain characteristic was classified as one of the following: none to very slight, slight to moderate, and moderately severe to severe. Table 5.2 presents the results of AA, its first two components, $Z^2{}_1$ and $Z^2{}_2$, and KW.

Table 5.2: Results of AA, $Z^2{}_1$, $Z^2{}_2$, and KW for the Stain Experiment

Factor	AA	$Z^2{}_1$	$Z^2{}_2$	KW
A	3.20	.28	5.05	.11
B	.53	.18	.48	.24
C	7.47	5.05	.28	5.23
BC	.53	.18	.48	.24
D	.53	.18	.48	.24
BD(AC)	3.20	1.93	.74	1.71
CD	.53	.18	.48	.24

Based on the unmodified AA statistics, the original analysis declared the main effects for factors A and C and their interaction significant. Based on the modified AA statistics from Table 5.2 and the $d\chi^2{}_v$ approximation, the factor A main effect and the AC interaction effect are not significant. Nevertheless, these two effects could be declared relatively important. However, using $Z^2{}_1$ or KW, only the factor C main effect is significant, although the AC interaction effect may be relatively important. The difference in the conclusion about the factor A main effect can be explained by the small $Z^2{}_1$ component and the large $Z^2{}_2$ component of its AA statistic. Notice that the factor A main effect is confounded with the AC×C effect. What AA detected then was the difference in dispersion between the two mixtures of distributions caused by the significant factor C main effect and the relatively important AC interaction effect.

The original conclusions reveal another concern when fractional factorial designs are used: as was true for the arc welding experiment, confounding must be considered. In the original analysis by Lear and Stanton (1985), the BD(AC) interaction effect, which was declared significant, was labeled as the BD interaction neglecting the confounding with the AC interaction. Since the main effects for factors A and C were declared significant, the AC interaction effect is more plausible. Generally, labeling it as a BD interaction can affect the recommended combination of the factor levels. Moreover, additional runs must be performed to unconfound the effects in question. However, using the correct analysis based on $Z^2{}_1$ or KW, confounding need not be considered since only the factor C main effect is significant.

Using the $\mathbf{q}_1$ or $\mathbf{l}$ scores, the high level of factor C is the optimum level. In the original analysis, because AA also detected a factor A main effect and a BD interaction effect, the original recommendations (Lear and Stanton, 1985) chose the high level of factor A and the low levels of factor B and D. Based on the correct analysis, these additional recommendations are unnecessary. Although these additional recommendations will not affect the contact stain characteristic, they can have significant financial implications.

Other aspects of both experiments are worth mentioning. Only one observation was taken per run and the contact stain characteristic and workability were classified into three ordered categories. These may not provide enough information for detecting smaller yet important factors. More observations per run should be taken and if possible more refined categories should be used to provide more information to detect real effects.

6. Discussion

Although popular for analyzing ordered categorical data from industrial experiments, there is a fundamental problem with AA in the multifactor setting. For a factor main effect, AA collapses the design onto the factor, and computes the AA statistic as in the single factor setting. The problem is that AA is no longer comparing I distributions but I mixtures of distributions, where the factor has I levels. AA can detect the difference between the I mixtures of distribution. However, this difference does not imply that the factor has an effect, because whether the I mixtures of distributions are different depends on the other factors. This problem with AA is a result of neglecting the inherent nature of ordered categorical data which makes them harder to analyze. Spurious effects can be detected. Furthermore, small real effects can be detected for the wrong reason.

The first component of the AA statistic, ${Z^2}_1$, and the Kruskal-Wallis test provide somewhat unsatisfactory alternatives for detecting location effects. Under the dispersion model, they can detect spurious effects. Under the location model, ${Z^2}_1$ can still detect spurious effects whereas the Kruskal-Wallis test misses real effects. However, they can be useful under the location model, where they perform similarly. ${Z^2}_2$ and SS(**d**), as dispersion tests, are rarely applicable in the multifactor setting.

The results in the previous sections suggest important details which must be considered during the design and analysis phases of the experiment. In the design phase, a balanced design should be used and the ordered categories should be refined as much as possible. In the analysis phase, AA has been applied to mixed categorical-continuous data by grouping the continuous data. This practice should be avoided because valuable information for detecting effects is lost. Furthermore, the transformation into ordered categories can change the location of the mixtures of distributions so that either real effects can be missed or spurious effects detected. When a fractionated design is used, the confounding patterns must be considered and additional runs must be made so that the effects in question can be unconfounded.

Acknowledgements

This research was supported by National Science Foundation Grant DMS-8420968. The simulation studies were performed on the VAX 11/780 computer of the Mathematics Research Center, University of Wisconsin--Madison. This research was aided by access to the research computer of the University of Wisconsin--Madison Statistics Department.

References

Agresti, A. (1984), *Analysis of Ordinal Categorical Data,* New York, NY: Wiley.

Box, G. (1985), "A Note on Accumulation Analysis." (unpublished manuscript)

Lear, C. and Stanton, J. (1985), "Contact Stain Requirement," in *Third Symposium on Taguchi Methods,* Romulus, MI: American Supplier Institute, Inc.

Nair, V. N. (1986a), "Components of Cumulative Chi-Square Type Tests for Ordered Alternatives in Contingency Tables," Statistical Research Reports #18, Murray Hill, NJ: AT&T Bell Laboratories.

Nair, V. N. (1986b), "Testing in Industrial Experiments with Ordered Categorical Data," *Technometrics,* 28, 283-291.

Nair, V. N. (1986c), personal communication.

Taguchi, G. (1974), "A New Statistical Analysis for Clinical Data, the Accumulating Analysis, in Contrast with the Chi-Square Test," *Saishin Igaku,* 29, 806-813.

Taguchi, G. and Wu Y. (1980), *Introduction to Off-Line Quality Control,* Japan Quality Control Association.

Takeuchi, K. and Hirotsu, C. (1982), "The Cumulative Chi-Square Method Against Ordered Alternatives in Two-Way Contingency Tables," *Rep. Statist. Appl. Res. Union Jap. Sci. Eng.,* 29, 1-13.